

Matematika jegyzet

Óbudai Egyetem, AMK

Halmazok

Halmaznak elemek egy bizonyos sokaságát nevezzük. Elsőként nagyrészt csak bizonyos számhalmazokkal (és némi geometriával) foglalkozunk.

A valós számok bevezetése

A matematika tárgyalását a számok bevezetésével, ezek közül is a valós számok bevezetésével kezdjük meg. A valós számokat egy egyenes, az ún. számegyenes pontjainak feleltetjük meg. A valós számok körében két műveletet értelmezünk, ezek lesznek az összeadás és a szorzás. Le kell szögeznünk, hogy ezen műveletek nem vezetnek ki a valós számok köréből.

Először bevezetjük a nulla számot, aminek az a lényeges tulajdonsága, hogy bármely x valós számhoz hozzáadva a nullát, eredményül x -et kapunk.

Hasonlóan bevezetjük az 1-et mint számot (ő lesz a valós számok körében az ún. egységelem), aminek az a lényeges tulajdonsága, hogy bármely y valós számot 1-gyel szorozva eredményül y -t kapunk.



A valós számok műveleti tulajdonságait az alábbiakban foglaljuk össze (valamiről már eddig is esett szó):

Az összeadásra vonatkozó tulajdonságok:

1. *Az összeadás kommutatív művelet, azaz mindegy, hogy milyen sorrendben adunk össze két valós számot. Ez formálisan azt jelenti, hogy bármely x, y valós számok esetén teljesül, hogy $x+y = y+x$.*

2. Az összeadás asszociatív művelet, azaz valós számok összeadásakor tetszőlegesen zárójeljelezhetünk. Ez formálisan azt jelenti, hogy bármely x, y, z valós számok esetén teljesül, hogy $(x+y)+z = x + (y+z)$.
3. Létezik úgynevezett nullelem (ami a nulla mint szám). Bármely x valós számhoz hozzáadva a nullát, eredményül x -et kapunk.
4. Minden valós számnak létezik ellentettje, azaz bármely valós számhoz létezik olyan valós szám, amivel őt összeadva nullát kapunk eredményül.

A fentiekhez hasonló módon írhatók le a szorzás fő tulajdonságai.

Az szorzásra vonatkozó tulajdonságok:

1. Az szorzás kommutatív művelet, azaz mindegy, hogy milyen sorrendben szorzunk össze két valós számot. Ez formálisan azt jelenti, hogy bármely x, y valós számok esetén teljesül, hogy $xy = yx$.
2. Az szorzás asszociatív művelet, azaz valós számok szorzásakor tetszőlegesen zárójeljelezhetünk. Ez formálisan azt jelenti, hogy bármely x, y, z valós számok esetén teljesül, hogy $(xy)z = x(yz)$.
3. Létezik úgynevezett egységelem (ami az 1 mint szám). Bármely y valós számot 1-gyel megszorozva, eredményül y -t kapunk.
4. Minden nullától különböző valós számnak létezik reciproka, azaz bármely nullával nem egyenlő valós számhoz létezik olyan valós szám, amivel őt összeszorozza 1-et kapunk eredményül.

Vegyük észre, hogy a fenti négy-négy tulajdonság mindegyike vagy csak az összeadásra, vagy csak a szorzásra vonatkozott. Ahhoz, hogy az összeadást és a szorzást valamilyen módon összekapcsolhassuk, szükség lesz egy további tulajdonságra is. Ezt középiskolában egyszerűen kiemelési szabálynak szokták nevezni, mi itt disztributív tulajdonságként hivatkozunk rá:

Disztributív tulajdonság:

Tetszőleges x, y, z valós számok esetén $x(y+z)=xy+xz$.

Néhány általános megjegyzés axiómákról és következményeiről

A valós számokra vonatkozó fenti kilenc tulajdonságot csak önkényesen bevezettük, de egyáltalán nem is próbáltuk meg ellenőrizni, hogy tényleg igazak-e. Ezzel nincs is semmi baj, az olyan állításokat, melyeket mindenféle kritika nélkül igaznak fogadunk el, **axiómáknak** nevezzük. Ezekre kiindulópontként mindig szükségünk van.

Természetesen az axiómák megváltoztatásával az egész vizsgált struktúra (akár a számok) tulajdonságai is jelentősen megváltozhatnak. Erre egy jó példa a Bolyai János-féle geometria, amely nem ugyanúgy használja az axiómákat, mint a hagyományosnak mondható euklideszi geometria, amiről általában középiskolában tanulnak. Mivel ebbe a témakörbe itt nem tudunk belemenni részletesen, ezért csak irodalmat ajánlok. Érdekes dolgokat találhatnak a néhai Dr. Kálmán Attila „Nemeuklideszi geometriák elemei” című könyvében.

Ahhoz, hogy a számok bevezetésével kapcsolatban is lássunk egy egyszerű kis érvelést, tegyük fel magunknak azt a kérdést, hogy mit kapunk eredményül, ha egy valós számot nullával szorzunk. Valószínűleg mindenki rögtön rávágja, hogy az eredmény nulla lesz. Ez persze igaz, csak hogy az axiómáink között ez nem szerepel, ezért ha teljesen korrektül szeretnénk eljárni, akkor ezt mint állítást be kell bizonyítanunk az axiómák segítségével.

***Állítás:** Bármely valós számot nullával szorozva nullát kapunk eredményül.*

Az állítás bizonyítása:

Tegyük fel, hogy az állítással ellentétben mégis létezik olyan x valós szám, amit nullával megszorozva az eredmény nullától különböző valós szám lesz. Legyen $x \cdot 0 = y \neq 0$.

A szorzásra vonatkozó negyedik axióma miatt az y -nak létezik reciproka, legyen ez z .

Ekkor tehát $(x \cdot 0) \cdot z = 1$, innen felhasználva a szorzás kommutativitását kapjuk, hogy

$(0 \cdot x) \cdot z = 1$. Az asszociativitás miatt ez átírható úgy, hogy $0 \cdot (x \cdot z) = 1$.

Ebből az következne, hogy a nullának van reciproka, amit viszont a szorzás negyedik axiómájánál kizártunk. Tehát a megfogalmazott állításnak igaznak kell lennie, különben sérülne a szorzás negyedik axiómája. ■

Látható, hogy ezt az állítást ugyan nem volt nehéz bebizonyítani, mégis már egy ilyen egyszerű állítás belátásához is dolgoznunk kellett.

Figyelemreméltó az alkalmazott bizonyítási módszer. Feltettük, hogy az állítás hamis, és ebből ellentmondásra jutottunk. Az ilyen típusú bizonyítások a matematikában nagyon gyakoriak, összefoglaló néven ezeket **indirekt bizonyításoknak** hívjuk. Az ún. indirekt feltevés ilyenkor mindig a bizonyítandó állítás tagadása, és ha ebből ellentmondásra tudunk jutni, akkor ezzel beláttuk, hogy az eredeti (bizonyítandó) állítás igaz. Az ilyen típusú bizonyítási eljárást teljesen

korrektnek fogadjuk el, logikailag nem kell foglalkozni a háttérének tisztázásával (persze a józan ész is azt diktálja, hogy ennek az érvelésnek korrektnek kell lennie).

Testek

Az olyan számhalmazokat, melyekben megvan az összeadás és a szorzás, és ezek nem vezetnek ki a kiinduló számhalmazból, valamint eleget tesznek a már megfogalmazott összesen kilenc darab axiómának (lsd. korábban), összefoglaló néven testeknek nevezzük. Tehát a valós számok testet alkotnak.

A különbség és a hányados

Fontos megjegyeznünk, hogy a valós számok körében alapvetően csak két műveletet, az összeadást és a szorzást értelmeztük. Az axiómáknak köszönhetően azonban tudjuk értelmezni két valós szám különbségét (lsd. az ellentett létezésére vonatkozó axiómát) és valós számok hányadosát is, amennyiben nem nullával szeretnénk osztani (lsd. a reciprok létezésére vonatkozó axiómát).

Az egyenlőségi reláció, ekvivalenciarelációk

Külön szót kell ejtenünk az egyenlőségről, mint relációról.

Az egyenlőségnek három alapvető tulajdonsága az alábbi:

1. *Az egyenlőség reflexív, azaz minden x esetén $x=x$*
2. *Az egyenlőség szimmetrikus, azaz ha $x=y$, akkor $y=x$ is teljesül.*
3. *Az egyenlőség tranzitív, azaz amennyiben $x=y$ és $y=z$ egyszerre teljesül, akkor szükségképpen $x=z$.*

Az olyan típusú relációkat, amiknek megvan a három itt felsorolt tulajdonsága, gyűjtőnéven **ekvivalenciarelációknak** nevezzük.

A valós számok rendezettsége, rendezési axiómák

A valós számok úgymond rendezett testet alkotnak, ami azt jelenti, hogy nagyság szerint növekvő sorrendbe tudjuk rendezni őket. Ez egész egyszerűen annak köszönhető, hogy egy számegyenes pontjainak feleltettük meg őket, és itt megállapodás szerint mondhatjuk azt, hogy a számegyenesen

jobbra haladva a megfelelő pontok egyre nagyobb valós számoknak felelnek meg (tehát a rendezés a számegyenes pontjai között a haladási iránynak megfelelően történik).

A nullánál nagyobb (tehát a számegyenesen tőle jobbra elhelyezkedő) számokat **pozitív** valós számoknak, míg a nullánál kisebb számokat **negatív** valós számoknak nevezzük.

Azt, hogy az x valós szám nem nagyobb az y valós számnál, úgy jelöljük, hogy $x \leq y$.

Azt, hogy az x valós szám kisebb az y valós számnál, úgy jelöljük, hogy $x < y$.

A valós számok rendezésére vonatkozóan az alábbi axiómákból indulunk ki:

1. *A rendezési reláció reflexív, azaz minden x valós számra $x \leq x$*
2. *Bármely két x és y , egymástól különböző valós számra az $x \leq y$ és az $y \leq x$ relációk közül pontosan egy teljesül*
3. *A rendezési reláció tranzitív, azaz ha $x \leq y$, és $y \leq z$, akkor $x \leq z$.*
4. *Ha $x \leq y$ valós számok, akkor tetszőleges z valós számra $x + z \leq y + z$.*
5. *Ha $x \leq y$ valós számok, akkor tetszőleges $z \geq 0$ valós számra $x \cdot z \leq y \cdot z$.*

Az axiómákból következik, hogy pozitív számok szorzata mindig pozitív szám. Legyenek ugyanis y és z pozitív valós számok. Az ötödik axióma miatt $0 \cdot z \leq y \cdot z$, tehát $y \cdot z$ biztosan nemnegatív. Ugyanakkor $y \cdot z \neq 0$, hiszen biztosan létezik reciproka.

Hasonlóan bizonyítható, hogy két negatív szám szorzata pozitív, illetve az, hogy egy negatív és egy pozitív szám szorzata negatív.

Számhalmazok alsó és felső korlátai, korlátosság

Szükségképpen szót kell ejtenünk az olyan számhalmazokról, amelyeknek elemei nem lehetnek akármilyen nagyok, vagy éppen nem lehetnek akármilyen kicsik.

Először definiáljuk a **felülről korlátos** számhalmazokat.

Definíció: Egy valós számhalmazt felülről korlátosnak nevezünk, ha létezik olyan c valós szám, melynél a halmaz egyik eleme sem nagyobb. Egy ilyen c szám az adott számhalmaznak ún. felső korlátja.

Hasonlóan definiálандók az **alulról korlátos** számhalmazok:

Definíció: Egy valós számhalmazt alulról korlátosnak nevezünk, ha létezik olyan c valós szám, melynél a halmaz egyik eleme sem kisebb. Egy ilyen c szám az adott számhalmaznak ún. alsó korlátja.

Amennyiben egy számhalmaz alulról és felülről szintén korlátos, akkor egész egyszerűen **korlátosnak** fogjuk nevezni.

Nagyon fontos az ún. felsőhatár- és alsóhatár-axióma, ami a fentiekkel kapcsolatos.

Felsőhatár-axióma: Bármely felülről korlátos számhalmaznak van legkisebb felső korlátja. Ezt **felsőhatárnak**, másnéven **szuprémumnak** nevezzük.

Alsóhatár-axióma: Bármely alulról korlátos számhalmaznak van legnagyobb alsó korlátja. Ezt **alsóhatárnak**, másnéven **infimumnak** nevezzük.

A Cantor-axióma

Szintén a rendezési kérdésekhez tartozik a **Cantor-axióma**. Ez annyit mond ki, hogy egymásba skatulyázott szakaszoknak (azaz egymásba skatulyázott zárt intervallumoknak) van legalább egy közös pontja.

A természetes számok bevezetése, a teljes indukció elve

Külön meg kell említenünk a valós számok egy speciális osztályát, a természetes számokat.

A természetes számokat az alábbi két szabály szerint vezetjük be:

1. A nulla legyen természetes szám.
2. Ha egy szám természetes szám, akkor a nála eggyel nagyobb szám is legyen természetes szám.

Könnyen meggondolható, hogy ez a két utasítás együtt éppen azt jelenti, hogy a természetes számok növekvő sorrendben felsorolva $0, 1, 2, 3, \dots$

Tehát azt lehet mondani, hogy a természetes számokat úgy vezetjük be, hogy megnevezzük a legkisebbet (ami a nulla), és innen egyesével felfelé lépkedve mindegyik számot bevesszük a halmazba.

Ezen eljáráson felbuzdulva akár egy új bizonyítási stratégiát is kidolgozhatunk. Ez lesz az ún. **teljes indukció**. Legyenek adva olyan állítások, amik meg vannak számozva a természetes számokkal. Szeretnénk bebizonyítani ezen állítások mindegyikének helyességét. Ehhez elég lenne az alábbi két dolgot megtennünk:

1. Leellenőrizni, hogy a nulladik állítás igaz.

2. *Bebizonyítani, hogy mindegyik állításból logikailag következik a nála eggyel nagyobb sorszámú állítás.*

Persze egyáltalán nem biztos, hogy ez sikerülni fog (mivel nem biztos, hogy a megfogalmazott állítások mind igazak), de ha sikerül, akkor a természetes számok bevezetésénél látott elv miatt igaznak fogadhatjuk el az összes állításunkat (hiszen az elv ugyanaz, mint amit a természetes számok bevezetésénél láttunk). Az ilyen típusú bizonyításokat teljes indukciós bizonyításoknak nevezzük.

Ahhoz, hogy demonstráljuk ezt a típusú bizonyítást, az alábbi közismert állítást fogjuk vele bebizonyítani:

Állítás: Amennyiben n természetes szám, akkor $0+1+2+\dots+n=\frac{n(n+1)}{2}$.

Az állítás bizonyítása:

Látható, hogy itt minden egyes n természetes számra van egy-egy állításunk, amiket megpróbálhatunk bizonyítani teljes indukcióval. A nulladik állítás az, hogy $0=0$, ez pedig igaz.

Be kellene látni, hogy minden állításból logikailag következik a nála eggyel nagyobb sorszámú állítás.

Induljunk ki a k -adik állításból, ami azt mondja ki, hogy $0+1+2+\dots+k=\frac{k(k+1)}{2}$.

Ebből szeretnénk logikailag következtetni a $(k+1)$ -edik állításra, ami azt mondja ki, hogy

$$0+1+2+\dots+k+(k+1)=\frac{(k+1)(k+2)}{2}.$$

A k -adik állításnál szereplő jobb oldalt beírhatjuk a $(k+1)$ -edik állítás bal oldalán szereplő képletbe:

$$0+1+2+\dots+k+(k+1)=\frac{k(k+1)}{2}+(k+1).$$

Innen már csak számolnunk kell. A $(k+1)$ -et kiemelhetjük:

$$\frac{k(k+1)}{2}+(k+1)=(k+1)\left(\frac{k}{2}+1\right)=(k+1)\left(\frac{k}{2}+\frac{2}{2}\right)=(k+1)\left(\frac{k+2}{2}\right)=\frac{(k+1)(k+2)}{2}.$$

Ezzel tehát azt is beláttuk, hogy minden állítás helyessége maga után vonja a következő állítás helyességét. Mivel a nulladik állítás igaz volt, ezért ezzel az összes állítás helyességét bebizonyítottuk. ■

Fontos megjegyeznünk, hogy a teljes indukciós bizonyításoknak az imént ismertetett csupán a **gyenge formája**.

A teljes indukció erős formája azt jelenti, hogy (a természetes számokkal megszámozott) állításainkat az alábbi módon szeretnénk bebizonyítani:

1. *Leellenőrizzük, hogy a nulladik állítás igaz.*
2. *Bebizonyítjuk, hogy mindegyik állítás helyessége logikailag következik a nála kisebb indexű állítások helyességéből.*

Amennyiben az állítások nem mind igazak, akkor persze ez az egész valahol elakad. De amennyiben mindegyik igaz, akkor ennek a módszernek megegyezés szerint van létjogosultsága.

Végtelen sok elemből álló halmazok

A természetes számok halmaza, illetve a valós számok halmaza is **végtelen sok elemű halmaz**. Ez azt jelenti, hogy ezen halmazok elemszáma bármely c valós számnál nagyobb.

Arkhimédész axiómája

A természetes számok bevezetése után meg kell említenünk az arkhimédészi axiómát. Ez azt mondja ki, hogy **minden valós számnál van nála nagyobb természetes szám**.

Számtani és mértani sorozatok

Ahhoz hasonlóan, ahogyan bevezettük a természetes számokat, tudjuk értelmezni a számtani sorozatokat.

Definíció (számtani sorozat) Legyenek w és d valós számok. A „ w ” kezdőelemű, „ d ” differenciájú számtani sorozatot az alábbi módon adjuk meg:

1. *A sorozat kezdőeleme w .*
2. *Ha egy szám eleme a sorozatnak, akkor a nála d -vel nagyobb szám is eleme a sorozatnak.*

A szomszédos természetes számok összegéről belátott állítás segítségével rögtön megkaphatjuk a középiskolában megismert számtani sorozatos összegképletet. Legyen a számtani sorozatunk kezdőeleme w , a differenciája d . Legyen n tetszőleges nullánál nagyobb természetes szám. Ekkor ezen számtani sorozat első n darab elemének összege

$$\begin{aligned} w + (w+d) + (w+2d) + \dots + [w + (n-1)d] &= nw + d[0 + 1 + 2 + \dots + (n-1)] = \\ &= nw + d \cdot \frac{(n-1)n}{2} \end{aligned}$$

Az előzőekhez hasonlóan járhatunk el a mértani sorozatok értelmezése során.

Definíció (mértani sorozat) Legyenek w és q valós számok. A „ w ” kezdőelemű, „ q ” kvóciensű mértani sorozatot az alábbi módon adjuk meg:

1. A sorozat kezdőeleme w .

2. Ha egy szám eleme a sorozatnak, akkor az q -szoros is eleme a sorozatnak.

A mértani sorozatokra vonatkozó, középiskolában megismert összegképletet később fogjuk bebizonyítani.

Az egész számok halmaza, oszthatóság, maradékos osztás

Az egész számok halmazát a természetes számok és azok ellentettjei együttevén alkotják.

Az egész számok körében az alábbi módon értelmezzük az oszthatóságot:

Definíció: Az n egész szám osztható az m egész számmal, ha létezik olyan k egész szám, melyre $km=n$ teljesül. Ezt másképp úgy is mondjuk, hogy ebben az esetben az n szám többszöröse az m számnak, vagy úgy, hogy az n szám osztható az m számmal (jelölés $m|n$).

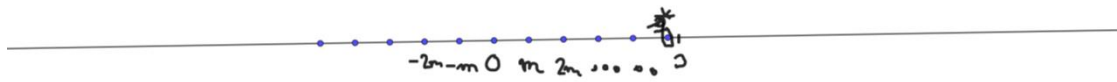
Az egész számok közül a kettő többszörőseit **párosnak**, a kettővel nem oszthatókat **páratlannak** szoktuk nevezni.

Érdekes azon elgondolkodnunk, hogy hogyan ábrázolhatjuk egy m nullától különböző egész szám összes többszörösét a számegyenesen. A nulla minden egész számnak többszöröse, ezért a nulla biztosan az m többszörősei között van. A többi többszörőst pedig úgy találhatjuk meg, hogy a nullából akár jobbra, akár balra elindulva elkezdünk lépegetni, mindegyik lépésben pontosan m egységnyit. Minden ilyen lépésben az m egész szám egy-egy többszörösét találjuk meg. Ezen szemlélet segítségével könnyen bizonyítható az alábbi tétel:

A maradékos osztás tétele: Minden s egész számot fel tudunk írni $s=jm+r$ alakban, ahol $m > 0$ egész, j és r egész számok, továbbá $0 \leq r < m$.

Bizonyítás:

Legyen s adott. Ha s többszöröse m -nek, akkor természetesen felírható a kívánt alakban (ekkor $r=0$ lesz). Ha s nem többszöröse m -nek, akkor s -től keressük meg balra a számegyenesen a legelső olyan pontot, ami az m szám egy többszörösének felel meg, legyen ez s^* . Az s^* távolsága s -től biztosan kisebb m -nél, és mivel s^* többszöröse m -nek, ezért biztosan létezik a kívánt felírás.



Érdemes megfigyelnünk, hogy a maradékos osztás tételében megfogalmazott felírás mindig egyértelmű. Ez abból következik, hogy ha a bizonyításban szereplő s^* helyett az m számnak egy másik többszörösét választanánk, akkor az r értéke vagy negatív lenne, vagy pedig nagyobb lenne $(m-1)$ -nél.

Ez tehát azt jelenti, hogy tetszőleges m pozitív egész szám esetén fennáll, hogy bármely s egész szám egyértelműen osztható maradékosan m -mel. A tételben szereplő r -et az m szerinti **osztási maradék**nak nevezzük.

Prímszámok és összetett számok, a számelmélet alaptétele

A pozitív egész számok halmazát oszthatóság szempontjából az alábbi három különböző (azaz diszjunkt) részhalmazra osztjuk:

1. Az egységelem (az 1 mint szám önmagában).
2. Prímszámok: azok az egynél nagyobb egész számok, melyeknek az egyen és saját magukon kívül nincsen más pozitív osztójuk.
3. Összetett számok: azok a pozitív egész számok, amik nem egyenlők eggyel, és nem is prímszámok.

Könnyen látható, hogy mindegyik pozitív egész szám a fenti három közül pontosan az egyik kategóriába sorolható.

Az összetett számok (mint ahogy az a fentiekből leolvasható) olyan egynél nagyobb egész számok, amiknek van olyan pozitív osztójuk, ami egytől és az adott számtól különböző.

Prímszámok a 2, 3, 5, 7, stb.

Összetett számok például a 24, 60, 85. Az általános iskolában és a középiskolában szó esett a törzstényezős felbontásról, ami azt jelentette, hogy a pozitív egész számokat felírtuk prímszámok szorzataként. Például $2=2$, $5=5$, $24=2 \cdot 2 \cdot 2 \cdot 3$, $60=2 \cdot 2 \cdot 3 \cdot 5$. Megjegyezzük, hogy az egyet üres szorzatként tudjuk felírni (nem szerepel benne szorzótényező). A prímszámok szorzataként történő felírással kapcsolatosan következő tétel alapvető fontosságú (középiskolában valószínűleg kimondták, de nem bizonyították).

A számelmélet alaptétele: Minden egynél nagyobb egész szám a szorzótényezők sorrendjétől eltekintve egyértelműen felírható pozitív prímszámok szorzataként.

Bizonyítás:

Az előállíthatóság bizonyítása:

Azt, hogy minden egynél nagyobb egész szám felírható így, (erős) teljes indukcióval igazoljuk. A 2 számot fel tudjuk írni pozitív prímszámok szorzataként. Tegyük fel, hogy a 2, 3, ..., $k-1$ számok mindegyikére beláttuk már ezt. Ebből szeretnénk következtetni arra, hogy a k szám is felírható így. Ha a k szám prímszám, akkor persze felírható a kívánt alakban. Ha k összetett, akkor vannak olyan i és j egynél nagyobb egészek, melyek szorzata $ij=k$. Mivel ezek egynél nagyobbak, és a szorzatuk k , ezért mindketten kisebbek k -nál, és így őket biztosan fel tudjuk írni valamilyen pozitív prímszámok szorzataként. Ezáltal a szorzatuk, $ij=k$ is valamilyen pozitív prímszámok szorzataként áll elő, amit bizonyítani akartunk.

Az egyértelműség bizonyítása:

Hátravan még annak igazolása, hogy a pozitív prímszámok szorzataként való előállítás a sorrendtől eltekintve egyértelmű. Ugyanúgy (erős) teljes indukcióval bizonyítunk. A 2 számra igaz az állítás. Tegyük fel, hogy a 2, 3, ..., $k-1$ számok mindegyikére beláttuk már az egyértelmű felírást. Ez alapján kellene bizonyítanunk, hogy a k számra is igaz ez az egyértelmű felírás. Tegyük fel indirekt, hogy a k számot legalább két különböző módon fel lehet írni pozitív prímszámok szorzataként, legyen két ilyen felírás $k=p_1 p_2 \dots p_s = q_1 q_2 \dots q_t$.

Ha $p_1=q_1$, akkor egyből ellentmondásra jutunk, hiszen ekkor teljesül, hogy

$1 < p_2 \dots p_s = q_2 \dots q_t < k$, és az indukciós feltevés szerint ebben az esetben a $p_2, p_3, \dots, p_s$ prímszámok valamilyen sorrendben egybeesnek a $q_2, q_3, \dots, q_t$ prímszámokkal.

Tehát feltehető, hogy $p_1 \neq q_1$, legyen mondjuk $p_1 < q_1$.

Tekintsük most a $k - p_1(q_2 \dots q_t)$ számot. Ez egy olyan egész szám, ami k -nál kisebb. Valamint az is biztos, hogy ez a szám egynél nagyobb, hiszen $k - p_1(q_2 \dots q_t) = p_1 p_2 \dots p_s - p_1(q_2 \dots q_t) =$

$= p_1(p_2 \dots p_s - q_2 \dots q_t)$, ahol $(p_2 \dots p_s - q_2 \dots q_t)$ biztosan egy pozitív egész szám, valamint p_1 pozitív prímszám. Ennek a számnak az indukciós feltevés szerint egyértelmű a pozitív prímszámok szorzataként történő előállítása, és abban is biztosak lehetünk, hogy a p_1 szerepel az előállításban. Ugyanakkor ez a szám felírható úgy is, hogy

$$k - p_1(q_2 \dots q_t) = q_1 q_2 \dots q_t - p_1(q_2 \dots q_t) = (q_1 - p_1)(q_2 \dots q_t).$$

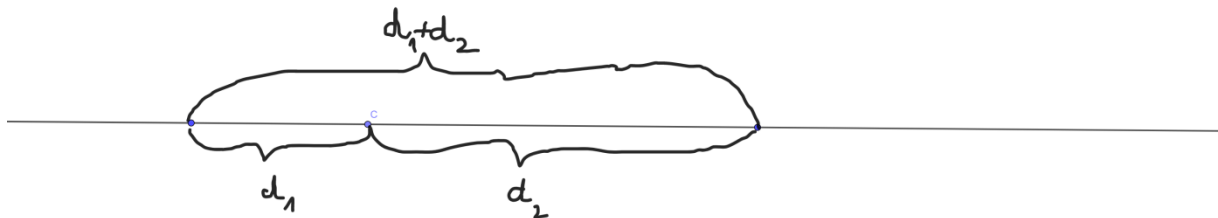
Ha a $q_2, q_3, \dots, q_t$ prímszámok közül valamelyik is megegyezik p_1 -gyel, akkor ugyanúgy ellentmondásra jutunk, mint amikor megvizsgáltuk a $p_1=q_1$ esetet. Tehát feltehetjük, hogy ezen prímek egyike sem egyenlő p_1 -gyel. Ugyanakkor a $q_1 \cdot p_1$ szám nem osztható p_1 -gyel, hiszen ha osztható lenne, akkor $(q_1 \cdot p_1) : p_1 = q_1$ is osztható lenne p_1 -gyel, ami nyilvánvalóan lehetetlen, amennyiben p_1 és q_1 különböző prímek.

Tehát mindenképpen ellentmondásra jutunk abból a feltevésünkből, hogy a k számot legalább két különböző módon sikerült felírunk pozitív prímszámok szorzataként. ■

A mérték fogalma (avagy miért éppen ezekkel az axiómákkal dolgozunk?)

Ebben a részben megpróbálunk magyarázatot adni arra, hogy miért éppen a már ismertett axiómákkal dolgozunk, és talán sikerül is adnunk néhány szemléletes példát.

A bevezetést követően rátérünk arra, hogy mi alapján tekinthetjük természetesnek az eddig kimondott axiómákat. Érdemes először arra gondolnunk, hogy a valós számokat egy egyenesen (a számegyenesen) bizonyos szakaszok hosszának feleltettük meg. Ez egy ún. mérték. A szemlélet alapján elvárjuk, hogy két csatlakozó, közös belső pont nélküli szakasz összhossza megegyezzen a két eredeti szakasz hosszának összegével. Innen viszont a szemlélet alapján „el kell fogadnunk” az összeadás kommutatív tulajdonságát.



Hasonló megfontolásokból célszerű elfogadnunk az összeadás asszociativitását.

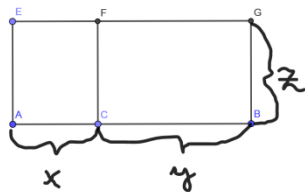
A nullelem bevezetése talán a „nullányi távolság” miatt tekinthető kézenfekvőnek, míg az ellentett elemek bevezetése a kivonás mint művelet bevezetése miatt szükséges.

A szorzás axiómáinak létjogosultságát talán jól megvilágíthatjuk a terület mint mérték segítségével. Ha kiindulunk abból, hogy egy téglalap területét (ami egy valós számmal mérhető) egyértelműen meghatározza két nem párhuzamos oldalának hossza (hiszen ezek szorzata méri a területet), akkor már el kell fogadnunk a szorzás kommutatív tulajdonságát. Az asszociatív tulajdonság létjogosultsága hasonlóan magyarázható a térben, téglatestek térfogatának segítségével.

Az egységnyi terület azon négyzet területe, aminek oldala egységnyi hosszú.

A reciprok létezése (ha nem nulla a kiinduló elem) a fentiek segítségével „magyarázható”.

A disztributív tulajdonság megvilágítására tekintsük az alábbi ábrát:



Az ábrán látható téglalapok területe xz , illetve yz . A nagy, összeragasztott téglalap területe $(x+y)z$, ami ugyanaz, mint a két kis téglalap területének összege, azaz $xz+yz$. Így szemléltettük, hogy valóban célszerű elfogadnunk a disztributivitást.

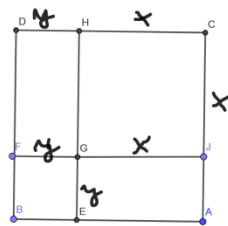
Néhány nevezetes azonosság és alkalmazásaik

Kéttagú összeg négyzete:

Az axiómák segítségével könnyedén levezethető a kéttagú összeg négyzetére vonatkozó képlet.

$$(x + y)^2 = (x + y)(x + y) = x(x + y) + y(x + y) = x^2 + xy + yx + y^2 = x^2 + 2xy + y^2.$$

Ehhez szemléletesen tekinthetjük az alábbi ábrát:



A nagy négyzet területe $(x + y)^2$, ami összeadódik a két kisebb négyzet, illetve a két kisebb téglalap területéből, amiknek értéke x^2 , y^2 , xy , xy .

Előjelcserével kapjuk, hogy $(x - y)^2 = x^2 - 2xy + y^2$.

Kéttagú összeg köbe:

Az axiómák (és a négyzetes képlet) segítségével két összeadandó tag köbére nézve is könnyen kapunk azonosságot:

$$\begin{aligned} (x + y)^3 &= (x + y)^2 \cdot (x + y) = (x^2 + 2xy + y^2) \cdot (x + y) = \\ &= (x^2 + 2xy + y^2)x + (x^2 + 2xy + y^2)y = \\ &= x^3 + 2xy \cdot x + y^2x + x^2y + 2xy \cdot y + y^3 = x^3 + y^3 + 3x^2y + 3y^2x. \end{aligned}$$

Ha ugyanezt a képletet geometriai szemlélet segítségével szeretnénk megkapni, akkor (a négyzetes képlethez hasonlóan) egy $(x+y)$ oldalú kocka térfogatából kell kiindulnunk. Ezt az érdeklődő olvasóra bízunk.

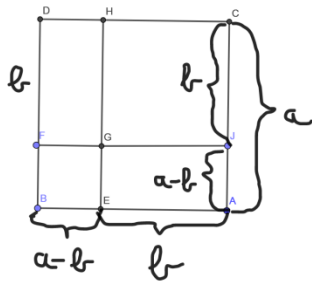
Előjelcserével kapjuk, hogy

$$(x - y)^3 = x^3 - y^3 - 3x^2y + 3y^2x.$$

Négyzetek különbsége:

Itt először a geometriai szemléletet vesszük elő.

Tekintsük az alábbi ábrát:



Nézzük meg, mit kapunk eredményül, ha az „a” oldalú négyzet területéből levonjuk a „b” oldalú négyzet területét. Ami megmarad, az három téglalap. Formálisan

$$a^2 - b^2 = (a - b)^2 + b(a-b) + b(a-b).$$

Itt kiemelhetünk (a-b)-t:

$$a^2 - b^2 = (a - b)^2 + b(a-b) + b(a-b) = (a-b)[(a - b) + b + b] = (a-b)(a+b).$$

Ezzel tehát azt kaptuk, hogy $a^2 - b^2 = (a-b)(a+b)$.

Ez persze az axiómák segítségével is könnyűszerrel ellenőrizhető:

$$(a-b)(a+b) = a(a+b) - b(a+b) = a^2 + ab - ba - b^2 = a^2 - b^2.$$

Többszörös összeg négyzete

Ha nem kéttagú az összeg, hanem n-tagú, ahol n valamilyen pozitív egész szám, akkor egy ilyen összeg négyzete úgy számolható ki, hogy vennünk kell az eredeti összeadandó tagok négyzeteinek összegét, és ehhez hozzá kell adnunk az eredeti összeadandó tagok kettős szorzatainak dupláját.

$$\text{Tehát pl. } (x + y + z)^2 = x^2 + y^2 + z^2 + 2xy + 2yz + 2zx.$$

Általánosan a képlet teljes indukcióval bizonyítható. Kéttagú összegre az állítás igaz. Bebizonyítjuk, hogy ha valamilyen k pozitív egészre igaz, akkor igaz (k+1)-re is.

Tekintsük az $(x_1 + x_2 + \dots + x_{k+1})^2$ kifejezést. Ez átírható úgy, hogy

$$(x_1 + x_2 + \dots + x_k)^2 + 2x_{k+1}(x_1 + x_2 + \dots + x_k) + x_{k+1}^2.$$

Az indukciós feltevés szerint $(x_1 + x_2 + \dots + x_k)^2$ úgy írható fel, hogy összeadjuk az $x_1, x_2, \dots, x_k$ változók négyzeteit, és ehhez hozzáadjuk azon különböző kettős szorzatok kétszeresét, amelyekben az x_{k+1} nem szerepel..

Ugyanakkor $2x_{k+1}(x_1 + x_2 + \dots + x_k)$ -ban éppen azok a kéttényezős szorzatok fordulnak elő, melyekben x_{k+1} is előfordul, ráadásul mindegyik pontosan kétszer. Végül ehhez hozzáadjuk x_{k+1}^2 -t.

Látható, hogy az állítás helyessége öröklődik k -ról $(k+1)$ -re, ezzel az állítást beláttuk.

Ilyen típusú kifejezéseknél már célszerű szummás alakot használni, a szummajelhez mindig odaírva, hogy milyen index(ek) szerint összegzünk.

Például a többtagú összeg négyzetére vonatkozó képletünket egyszerűen úgy írhatjuk le, hogy

$$(\sum_{i=1}^n x_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{j \neq k} 2 x_j x_k.$$

Azonos kitevőjű hatványok különbsége, egyenlőtlenségek hatványozása és a mértani sorozat összegképlete

A négyzetek különbségét már felírtuk szorzatként, ezt a formulát próbáljuk meg most általánosítani. Legyen n pozitív egész szám. Azt állítjuk, hogy általánosan igaz, hogy amennyiben x és y valós számok, akkor

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1}).$$

Ezt az axiómák segítségével igazoljuk:

$$\begin{aligned} & (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1}) = \\ & = x(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1}) - \\ & - y(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1}) = \\ & = x^n + x^{n-1}y + x^{n-2}y^2 + \dots + x^2y^{n-2} + xy^{n-1} - \\ & - (x^{n-1}y + x^{n-2}y^2 + x^{n-3}y^3 + \dots + xy^{n-1} + y^n) = \\ & = x^n - y^n. \end{aligned}$$

Például ha $n=3$, akkor azt az azonosságot kapjuk, hogy $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$.

Egyenlőtlenségek hatványozása

Az új általános azonosságunkkal könnyen belátható, hogy amennyiben $0 \leq y < x$ nemnegatív valós számok és n pozitív egész, akkor $y^n < x^n$. Ugyanis ekkor az

$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1})$ szorzat mindkét szorzótényezője pozitív.

Vegyük észre, hogy ez az állítás megfordítható, azaz ha n pozitív egész szám, valamint az y , illetve x nemnegatív valós számokra teljesül, hogy $y^n < x^n$, akkor $y < x$.

A mértani sorozat összegképlete

Tekintsük most azt a speciális esetet, amikor $y=1$. Ebben az esetben a kapott azonosság szerint igaz, hogy

$$x^n - 1 = x^n - 1^n = (x - 1)(x^{n-1} + x^{n-2} + x^{n-3} + \dots + x^2 + x + 1).$$

$$\text{Tehát ha } x \neq 1, \text{ akkor } \frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + x^{n-3} + \dots + x^2 + x + 1.$$

Ezt a speciális képletet fel tudjuk használni a **mértani sorozatokra vonatkozó összegképlet** levezetéséhez.

Legyen k a mértani sorozat kezdőtagja, és q a kvóciense, ahol $q \neq 1$. Ekkor az első n darab tag összege

$$\begin{aligned} k + kq + kq^2 + \dots + kq^{n-1} &= k(q^{n-1} + q^{n-2} + q^{n-3} + \dots + q^{n-2} + 1) = \\ &= k \cdot \frac{q^n - 1}{q - 1}. \end{aligned}$$

Azonos kitevőjű hatványok összege, ha a kitevő páratlan

Ha n páratlan pozitív egész, akkor könnyen meggyőződhetünk róla, hogy $(-1)^n = -1$. Ez azzal magyarázható, hogy hatványozáskor páratlan sokszor történik előjelváltás, ezért a hatványozás elvégzése után az előjel változatlan marad.

Legyen tehát n páratlan pozitív egész szám, és $n=2k+1$, ahol k egész szám.

Az eddigiek alapján kapjuk azt, hogy

$$\begin{aligned} x^{2k+1} + y^{2k+1} &= x^{2k+1} - (-y^{2k+1}) = \\ &= x^{2k+1} - (-y)^{2k+1} = [x - (-y)](x^{2k} - x^{2k-1}y + x^{2k-2}y^2 - \dots - xy^{2k-1} + y^{2k}) \\ &= [x + y](x^{2k} - x^{2k-1}y + x^{2k-2}y^2 - \dots - xy^{2k-1} + y^{2k}) \end{aligned}$$

Valós számok négyzete, a Cauchy-Schwarz-Bunyakovszkij-egyenlőtlenség

Érdekes külön szót ejtenünk valós számok négyzetéről (tehát amikor egy valós számot önmagával szorzunk). A rendezési relációra vonatkozó ötödik axióma miatt minden nemnegatív szám négyzete legalább nullával egyenlő. Ha pedig negatív számot emelünk négyzetre, akkor az ugyanaz, mintha ennek a negatív számnak az ellentettjét emelnénk négyzetre (lsd. ellentett létezése), ezért ekkor is legalább nullát kapunk eredményül.

Tehát konklúzióként kijelenthetjük, hogy bármely valós szám négyzete legalább nullával egyenlő (és csak akkor lesz nulla, ha a kiinduló szám értéke nulla).

Ezen egyszerű észrevétel felhasználásával bebizonyítjuk az alábbi alapvető egyenlőtlenséget:

Tétel (Cauchy-Schwarz-Bunyakovszkij-egyenlőtlenség): Legyen n tetszőleges pozitív egész szám. Ha $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$ tetszőleges valós számok, akkor

$$(x_1^2 + x_2^2 + \dots + x_n^2)(y_1^2 + y_2^2 + \dots + y_n^2) \geq (x_1y_1 + x_2y_2 + \dots + x_ny_n)^2.$$

Bizonyítás:

Ha a jobb oldalon elvégezzük a négyzetre emelést, akkor megjelennek a zárójelben szereplő összeadandó tagok négyzetei, illetve az ezekből képezhető vegyes szorzatok, mindegyik kettővel szorozva. Tehát

$$(x_1y_1 + x_2y_2 + \dots + x_ny_n)^2 = \sum_{i=1}^n (x_iy_i)^2 + \sum_{j \neq k} 2x_jy_jx_ky_k$$

A bal oldal (mivel mindent mindennel szorzunk)

$$\begin{aligned} (x_1^2 + x_2^2 + \dots + x_n^2)(y_1^2 + y_2^2 + \dots + y_n^2) &= \\ &= \sum_{i=1}^n (x_iy_i)^2 + \sum_{j \neq k} x_j^2 y_k^2 = \\ &= \sum_{i=1}^n (x_iy_i)^2 + \sum_{j < k} x_j^2 y_k^2 + \sum_{j > k} x_j^2 y_k^2. \end{aligned}$$

Azt kellene tehát belátnunk, hogy

$$\sum_{i=1}^n (x_iy_i)^2 + \sum_{j < k} x_j^2 y_k^2 + \sum_{j > k} x_j^2 y_k^2 \geq \sum_{i=1}^n (x_iy_i)^2 + \sum_{j \neq k} 2x_jy_jx_ky_k.$$

Ez azzal ekvivalens, hogy

$$\sum_{j < k} x_j^2 y_k^2 + \sum_{j > k} x_j^2 y_k^2 \geq \sum_{j \neq k} 2x_jy_jx_ky_k, \text{ vagyis}$$

$$\sum_{j < k} x_j^2 y_k^2 + \sum_{j > k} x_j^2 y_k^2 - \sum_{j \neq k} 2x_jy_jx_ky_k \geq 0.$$

Vegyük észre, hogy a bal oldal itt éppen $\sum_{j \neq k} (x_jy_k - x_ky_j)^2$, ami természetesen legalább nullával egyenlő, hiszen ez valós számok négyzeteinek összege. ■

A racionális számok teste

Racionális számoknak nevezzük azon számokat, amiket elő tudunk állítani két egész szám hányadosaként.

Mivel a valós számok halmaza testet alkot (ami nekünk itt fontos, hogy bármely nullától különböző elemmel lehet benne osztani), ezért a racionális számok mind valósak is egyben.

Racionális szám például az összes egész szám, vagy például az $1/2$, $-5/6$, de akár a $0,2$ ($=1/5$) is.

A racionális számok halmaza szintén testet alkot. Ehhez elég ellenőrizni a kilenc darab testaxióma teljesülését. Mivel ez nem nehéz, hanem csupán egy technikai feladat, ezért ezt itt nem részletezzük.

Racionális kitevőjű hatványok, gyökvonás

Hatványozni könnyű, ha a kitevő egész szám. Ha ugyanis a kitevő nulla, akkor az eredmény egy. Ha a kitevő valamilyen pozitív egész, akkor az alapot annyiszor kell önmagával összeszoroznunk, amennyi a kitevő értéke. Ha pedig a kitevő negatív egész szám, akkor az alapot először ennek a kitevőnek az ellentettjére kell emelni, majd vennünk kell a kapott szám reciprokát.

Kérdés, hogy mi a helyzet akkor, ha a kitevő nem egész szám, de racionális. Legyen a kitevő r/s , ahol r és s valamilyen egész számok. Itt persze az is feltehető, hogy s pozitív (hiszen azt megtehetjük, hogy a számlálóban és a nevezőben egyszerre végzünk előjelváltást). Ilyen esetben az (r/s) -edik hatványt úgy értelmezzük, hogy először r -edekre emelünk, majd vesszük a kapott szám s -edik gyökét. Csakhogy az s -edik gyök nem feltétlenül létezik (legalábbis ilyen axiómánk nem volt). Tehát ezzel kapcsolatban be kell bizonyítanunk a megfelelő tételt, amit már középiskolában tanultunk.

Tétel: *Ha s pozitív egész, illetve x valamilyen nemnegatív valós szám, akkor létezik pontosan egy olyan nemnegatív valós szám, amelynek s -edik hatványa egyenlő x -szel.*

Bizonyítás:

Ha $x=0$, akkor az egyetlen megfelelő szám a nulla lesz (ez könnyen belátható az axiómákból). Legyen x pozitív.

Az arkhimédeszi axióma miatt van olyan természetes N szám, amelyre teljesül, hogy $x \leq N^s$.

Ugyanígy az arkhimédeszi axióma miatt van olyan n természetes szám, amelyre $\frac{1}{x} \leq n^s$, ami azt jelenti, hogy $\frac{1}{n^s} \leq x$. A két egyenlőtlenség együtt azt jelenti, hogy $\frac{1}{n^s} \leq x \leq N^s$, azaz $\left(\frac{1}{n}\right)^s \leq x \leq N^s$.

Tehát ha az x számnak van nemnegatív s -edik gyöke, akkor annak biztosan $\frac{1}{n}$ és N közé kell esnie.

Ezzel persze még nem találtuk meg a számot, amit keresünk, de az első lépést már megtettük.

Próbáljuk meg ezt az ötletünket finomítani. Felezzük el a számegyenesen az $\frac{1}{n}$ és N számokat összekötő szakaszt, a felezőpontnak megfelelő szám legyen f .

Ha $f^s \leq x$, akkor $f^s \leq x \leq N^s$, ha pedig $f^s \geq x$, akkor $\left(\frac{1}{n}\right)^s \leq x \leq f^s$. Bármelyik eset is teljesül, a kérdéses s -edik gyököt máris egy feleakkora intervallumon kell csak keresnünk.

Az ötlet az, hogy ezt a lépést akármeddig ismételhetjük. Ha véletlenül valamelyik lépésben az aktuális felezőpont egyenlő a keresett s -edik gyökkel, akkor megállunk. Ellenkező esetben végtelen sok lépést hajtunk végre, és keresendő s -edik gyököt mindig egy feleakkora zárt intervallumon kell keresnünk, mint az előző lépésben, és ráadásul az így kapott zárt intervallumok egymásba vannak skatulyázva. Tehát a Cantor-axióma miatt ezeknek az intervallumoknak biztosan van legalább egy közös pontja. Az is világos, hogy ez most egyetlen pont lesz, mert az intervallumhossz minden lépésben feleakkora lett, ezért legfeljebb egyetlen pont lehet, ami közös. Jelölje ezt az egyetlen közös pontot c . Belátjuk, hogy $c^s = x$. Ahhoz, hogy ezt bebizonyítsuk, jelöljük a k -adik felezés után nyert intervallum végpontjait v -vel és w -vel, ahol $v < w$.

A feltételeink szerint $v^s < x < w^s$, valamint $w - v = \frac{N - \frac{1}{n}}{2^k}$. Vizsgáljuk meg w^s -nek az x -től való eltérését.

$$\begin{aligned} w^s - x &< w^s - v^s = (w - v) \cdot (w^{s-1} + w^{s-2}v + w^{s-3}v^2 + \dots + wv^{s-2} + v^{s-1}) = \\ &= \frac{N - \frac{1}{n}}{2^k} \cdot (w^{s-1} + w^{s-2}v + w^{s-3}v^2 + \dots + wv^{s-2} + v^{s-1}) < \\ &< \frac{N - \frac{1}{n}}{2^k} \cdot (N^{s-1} + N^{s-2} \cdot N + N^{s-3} \cdot N^2 + \dots + N \cdot N^{s-2} + N^{s-1}) = \frac{N - \frac{1}{n}}{2^k} \cdot s \cdot N^{s-1}. \end{aligned}$$

Itt $N - \frac{1}{n}$, s , illetve N^{s-1} is konstans, viszont 2^k értéke tetszőlegesen nagy lehet (az arkhimédieszi axióma miatt), így w^s és x eltérése tetszőlegesen kicsi pozitív érték lehet. Természetesen ugyanez igaz v^s és x eltérésére. Mivel $v^s < c^s < w^s$, ezért biztosak lehetünk abban, hogy $c^s = x$.

Tehát létezik (nemnegatív) s -edik gyöke az x -nek. Ennek egyértelmősége azonnal következik az egyenlőtlenségekről korábban belátott hatványozási tulajdonságból. Ha ugyanis $0 \leq y < z$ közül mindkettő s -edik gyöke lenne az adott x számnak, akkor

$x = y^s < z^s = x$ teljesülne, ami nyilvánvalóan ellentmondás. ■

Fontos ellenőriznünk, hogy hatványozáskor mindegy, hogy a racionális kitevőt milyen törtalakban írjuk fel, mindig ugyanazt az eredményt kapjuk. Legyenek ugyanis r , s , t , v olyan egész számok, melyekre teljesül, hogy r/s egyenlő t/v -vel. Tegyük fel indirekt, hogy valamilyen x pozitív valós szám esetén $x^{r/s} \neq x^{t/v}$. Legyen mondjuk a jobb oldali hatvány a nagyobb, azaz $0 < x^{r/s} < x^{t/v}$. Az egyenlőtlenségek érvényben maradnak, ha mindent felemelünk az (sv) -edik hatványra:

$0 < x^{rv} < x^{ts}$. Mivel $rv = ts$, ezért ez ellentmondás. Tehát azt kaptuk, hogy valóban mindegy az, hogy a racionális kitevőnknek melyik alakját használjuk hatványozáskor.

Az irracionális számok halmaza

Vannak olyan valós számok, melyeket nem lehet felírni két egész szám hányadosaként. Ezeket **irracionális számoknak** nevezzük.

Először belátjuk, hogy léteznek irracionális számok. Ehhez elég egy konkrét irracionális számot mutatnunk, ami most a kettő négyzetgyöke lesz.

Állítás: A kettő négyzetgyöke irracionális.

Bizonyítás:

Az állítás bizonyítása indirekt fog történni.

Tegyük fel indirekt, hogy a kettő négyzetgyöke felírható két egész szám hányadosaként, legyen egy ilyen felírás k/n , ahol k és n egész számok. Azt is feltehetjük, hogy k és n egyaránt pozitív, mivel a kettő négyzetgyöke pozitív.

Ekkor $\frac{k}{n} \cdot \frac{k}{n} = \frac{k^2}{n^2} = 2$. Ezt átrendezhetjük úgy, hogy $2n^2 = k^2$.

A számelmélet alaptétele szerint a k és az n is egyértelműen felírható pozitív prímszámok szorzataként, legyen $k = p_1 p_2 \dots p_s$, $n = q_1 q_2 \dots q_t$.

Ha $2n^2 = k^2$, akkor ezt az imént bevezetett prímszámokkal felírva kapjuk, hogy

$$2q_1^2 q_2^2 \dots q_t^2 = p_1^2 p_2^2 \dots p_s^2.$$

Tekintsük az egyenlőség jobb oldalát. Mivel itt mindegyik prímszám négyzeten szerepel, ezért ebben a szorzatban a kettesek száma biztosan páros. Viszont a bal oldalon a kettesek száma biztosan páratlan, hiszen az első szorzótényező egy kettes, és utána minden prímszám második hatványon szerepel. Ez viszont azt jelenti, hogy a bal és a jobb oldal nem egyezhet meg egymással, hiszen ez ellentmondana a számelmélet alaptételében szereplő egyértelmű előállításnak.

Ezzel ellentmondásra jutottunk és beláttuk, hogy a kettő négyzetgyöke irracionális. ■

A fentihez hasonlóan belátható, hogy ha r és s olyan pozitív egész számok, melyekre teljesül, hogy az r szám nem s -edik hatványa egy egész számnak, akkor az r szám s -edik gyöke szükségképpen irracionális.

Irracionális számok közelítése racionális számok segítségével

Legyen adva egy irracionális szám. Válasszunk egy tetszőleges n pozitív egész számot. Ugyanúgy, mint ahogy a maradékos osztásnál láttuk, tekintsük a számegyenesen azokat a pontokat, amiket a nullából úgy érhetünk el, hogy jobbra vagy balra $\frac{1}{n}$ -nyi hosszúakat lépegetünk. Ezzel tehát megkapjuk az $\frac{i}{n}$ alakú számok halmazát, ahol i valamilyen egész számot jelöl. Tekintsük ezek közül a legnagyobbat, ami a kijelölt irracionális számunknál nem nagyobb. Ez nyilvánvalóan kevesebb, mint $\frac{1}{n}$ -nyi távolságra fekszik a kijelölt irracionális számunktól (ugyanúgy látható, mint a maradékos osztás tételénél ismerttetett gondolatmenetnél). Az arkhimédeszi axióma miatt n növekedtével ez az $\frac{1}{n}$ távolság tetszőlegesen kicsi pozitív érték lehet. Ez azt jelenti, hogy a kiszemelt irracionális számunk úgy is jellemezhető, mint a nála kisebb racionális számok halmazának a szupréma.

Természetesen ezt csinálhattuk volna fordítva is. Például a Cantor-axióma segítségével könnyen belátható, hogy az irracionális számunkat úgy is jellemezhetnénk, hogy ő a nála nagyobb racionális számok halmazának az infimuma.

A fenti gondolatmenetből következik, hogy minden irracionális szám tetszőlegesen közelíthető racionális számokkal.

Sőt ugyanígy az is könnyen bizonyítható, hogy a racionális számok a számegyenesen mindenütt sűrűn helyezkednek el. Ezalatt azt kell értenünk, hogy tetszőleges x valós szám és bármely $d > 0$ távolság esetén létezik végtelen sok olyan racionális szám, aminek x -től vett távolsága kisebb d -nél.

Irracionális kitevőjű hatványok bevezetése, azonos alapú hatványok rendezése

Legyen A egy rögzített, egynél nagyobb valós szám, továbbá legyenek r, s, t, v olyan egész számok, melyekre teljesül, hogy $\frac{r}{s} < \frac{t}{v}$. Természetesen feltehető, hogy s és v is pozitív (ui. ha valamelyik nem az, akkor az adott törtben elég megváltoztatnunk a számláló és a nevező előjelét.)

Hasonlítsuk össze $A^{\frac{r}{s}}$ -et és $A^{\frac{t}{v}}$ -t. Azt fogjuk bebizonyítani, hogy $A^{\frac{r}{s}} < A^{\frac{t}{v}}$.

Tegyük fel indirekt, hogy $A^{\frac{r}{s}} \geq A^{\frac{t}{v}}$. Emeljük az egyenlőtlenség mindkét oldalát sv -edik hatványra. Azt kapjuk, hogy $A^{rv} \geq A^{st}$. Ha ezt leosztunk A^{st} -vel, akkor kapjuk, hogy $A^{rv-st} \geq 1$.

Azonban (mivel s és v is pozitív) az $\frac{r}{s} < \frac{t}{v}$ egyenlőtlenség azt jelenti, hogy $rv < st$, azaz $(rv-st)$ értéke negatív. Mivel $A > 1$, ezért A -t akárhányszor önmagával szorozva egynél nagyobb értéket kapunk, így A -nak minden negatív egész kitevőjű hatványa kisebb 1-nél. Ez ellentmondás, ami azt bizonyítja, hogy valóban az $A^{\frac{r}{s}} < A^{\frac{t}{v}}$ egyenlőtlenségnek kell teljesülnie.

Tehát minél nagyobb racionális kitevőre emeljük az egynél nagyobb A valós számot, annál nagyobb lesz a kapott szám értéke.

Tisztázandó még, hogy miként értelmezzük az A szám irracionális kitevőjű hatványait. Legyen w valamilyen irracionális szám. Tudjuk, hogy w megegyezik a nála kisebb racionális számok halmazának szuprémumával. Tekintsük A azon hatványait, amelyekben a kitevő valamilyen w -nél kisebb racionális szám (és az alap A). Ezen számok halmaza felülről korlátos (hiszen a fenti észrevétel miatt egyik eleme sem nagyobb, mint A -nak egy olyan racionális kitevőjű hatványa, amelynek racionális kitevője nagyobb w -nél), így a felsőhatár-axióma miatt van szuprémuma.

Tehát ha $A > 1$, akkor általánosságban A^w értelmezhető az A szám olyan hatványainak szuprémumaként, amelyeknek kitevője w -nél nem nagyobb racionális szám.

Így tehát tudjuk értelmezni bármely $A > 1$ tetszőleges valós kitevőjű hatványait. Sőt az értelmezés alapján az is világos, hogy amennyiben $x < y$ valós számok, akkor $A^x < A^y$.

Hasonlóan tudjuk értelmezni A valós kitevőjű hatványait az egyéb esetekben. **Amennyiben $0 < A < 1$, akkor A^w értelmezhető az A szám olyan hatványainak infimumaként, amelyeknek kitevője w -nél nem nagyobb racionális szám.** Itt természetesen az fog teljesülni, hogy amennyiben $x < y$ valós számok, akkor $A^x > A^y$.

A számtani és mértani közepekre vonatkozó egyenlőtlenség

Gyakran felmerülő kérdés, hogy ha például ismerjük bizonyos nemnegatív számoknak az összegét, akkor legfeljebb mekkora lehet a szorzatuk. Az eddigi ismereteink alapján készen állunk ennek a kérdésnek a megválaszolására.

Tétel (a számtani és mértani középére vonatkozó egyenlőtlenség): Véges sok nemnegatív valós szám számtani közepe legalább akkora, mint a mértani közepük. Formálisan ez azt jelenti, hogy tetszőleges $x_1; x_2; x_3 \dots x_n \geq 0$ valós számok esetén teljesül az, hogy

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}$$

Bizonyítás:

Az természetesen feltehető, hogy a számaink az indexelés szerint vannak nagyság szerint rendezve, azaz $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$. Jelöljük ezen számok számtani közepét A -val.

Először is rögzített A és n mellett belátjuk, hogy a számok szorzatának értéke akkor lesz a legnagyobb, ha mind az n darab szám értékét A -nak, tehát éppen az átlaguknak választjuk meg. Ez persze azt is mutatja, hogy rögzített számtani közép mellett a mértani közép akkor lesz maximális, ha mindegyik elem ugyanakkora (itt használjuk fel az egyenlőtlenségek hatványozásáról tanultakat).

Amennyiben $x_1 = x_n$, akkor a számaink nagyság szerinti rendezése miatt mindegyikük egyenlő, vagyis mindegyikük egyenlő az átlaggal, A -val.

Tehát elegendő azt az esetet vizsgálnunk, amikor $x_1 < x_n$.

Ekkor cseréljük ki az x_1 és x_n számokat A -ra és $(x_1 + x_n - A)$ -ra, a többi számot pedig hagyjuk változatlanul. A cserét követően az n darab számunk összege (és számtani közepe) változatlan marad, továbbá az új számaink is mind nemnegatívak, hiszen A eleve nem lehetett negatív, valamint

$(x_1 + x_n - A)$ sem lehet negatív, ugyanis az x_n értéke legalább A -val egyenlő.

Azt fogjuk megvizsgálni, hogy az n darab szám szorzata miként változik a cserét követően. Mivel csak az x_1 és x_n számok változtak meg, ezért elég lenne az $x_1 x_n$ és az $A (x_1 + x_n - A)$ számokat összehasonlítani.

Azt fogjuk belátni, hogy $x_1 x_n < A (x_1 + x_n - A)$.

Ez ugyanis azzal ekvivalens, hogy $x_1 x_n - A x_1 - A x_n + A^2 < 0$,

azaz $x_1 (x_n - A) - A (x_n - A) < 0$,

ami szorzattá alakítás után úgy írható, hogy $(x_1 - A) (x_n - A) < 0$.

Mivel x_1 a legkisebb elem, és nem mindegyik szám egyenlő, ezért x_1 értéke kisebb az átlagnál, azaz A -nál. Hasonlóan (ahogy erre már utaltunk), mivel x_n a legnagyobb elem, ezért az ő értéke nagyobb A -nál. Ez alapján biztosan igaz, hogy $(x_1 - A)(x_n - A) < 0$, és ezáltal bizonyítottuk, hogy $x_1 x_n < A(x_1 + x_n - A)$.

Tehát amennyiben nem az összes számunk egyenlő, akkor a fenti csere segítségével a számok szorzata növelhető. Az ilyen típusú cseréket véges sokszor végrehajtva eljutunk ahhoz az állapothoz, hogy mindegyik számunk értéke az átlaggal egyenlő. Ebből tehát az következik, hogy rögzített összeg mellett a szorzat akkor lesz maximális, ha a nemnegatív számaink mindegyikét egyformának választjuk, és ebben az esetben a számtani és a mértani közép meg fog egyezni.

Ezzel a tétel állítását beláttuk, és azt is bizonyítottuk, hogy egyenlőség csak akkor teljesül, ha

$$x_1 = x_2 = x_3 = \dots = x_n.$$

Torlódási pontok, pozitív sugarú környezet

Azt, hogy a racionális számok az egész számegyenesen sűrűn helyezkednek el, kényelmesebben meg tudjuk fogalmazni az ún. **torlódási pont** fogalmát használva.

Definíció: Egy számhalmaznak **torlódási pontja** a t valós szám, ha bármely $d > 0$ távolság esetén a halmaznak végtelen sok olyan eleme létezik, melyeknek a t -től vett távolsága kisebb d -nél.

A torlódási pont fogalmát használva tehát kijelenthető, hogy a racionális számok halmazának minden valós szám torlódási pontja. Fontos megjegyeznünk, hogy a torlódási pont nem feltétlenül eleme a kiinduló halmaznak (mint ezt a mostani példánk is mutatja).

A torlódási pont fogalma hasonlóan értelmezhető olyan esetekben, amikor tudunk távolságot mérni (tehát nemcsak a számegyenesen, hanem pl. a síkban és a térben is).

A torlódási ponttal kapcsolatos a környezet fogalma, amit az alábbi módon vezetünk be:

Definíció (környezet): Egy u valós számnak a pozitív d sugarú környezetén azoknak a számoknak a halmazát értjük, amelyeknek az u -tól való eltérése kisebb, mint d .

Természetesen a környezet fogalma is érdekesebbé válik, ha nem csupán a számegyenesen, hanem olyan pontthalmazokban vizsgálódunk, melyekben tudjuk értelmezni a távolságot.

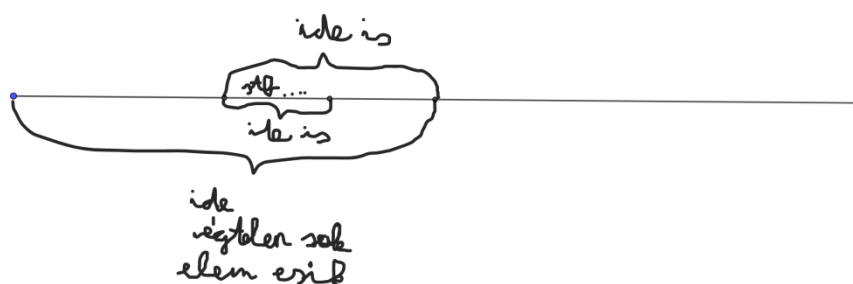
Torlódási pontok létezése

Belátjuk, hogy egy számhalmaznak nagyon sok esetben létezik torlódási pontja. Az alábbi tételt fogjuk bebizonyítani:

A Bolzano-Weierstrass-tétel számhalmazokra: Bármely végtelen sok elemből álló, korlátos számhalmaznak létezik torlódási pontja.

Bizonyítás:

A bizonyítás (mint ahogy a gyökök létezésénél) felezési eljárással fog történni. Először is tekintsünk egy olyan szakaszt (zárt intervallumot) a számegyenesen, amely a halmazunk minden elemét tartalmazza. Felezzük el ezt a zárt intervallumot. Az egyik felébe biztosan végtelen sok halmazbeli elem fog esni. Ezt (vagy az egyik ilyen) a felét tartsuk meg, a másik felét dobjuk el. A maradék zárt intervallumot ugyanígy felezzük el, és alkalmazzuk az előző lépést.



Az ilyen típusú lépéseink akárhányszor alkalmazhatóak. A Cantor-axióma miatt van olyan pont, ami minden ilyen zárt intervallumban benne van. Ráadásul egyetlen ilyen pont van, mert az intervallumok hossza mindig feleakkora lett. Az egyetlen pont, ami mindegyik intervallumunk eleme, biztosan torlódási pontja lesz a halmaznak, hiszen az intervallumokat minden egyes lépés során úgy értelmeztük, hogy bennük végtelen sok halmazbeli elem található. ■

Egy speciális halmaz torlódási pontjáról

Tekintsük azoknak a valós számoknak a halmazát, amelyek valamilyen egynél nagyobb n pozitív egész segítségével úgy írhatók fel, hogy $\sqrt[n]{n}$. Az világos, hogy ennek a halmaznak minden eleme nagyobb egynél. Azt fogjuk bebizonyítani, hogy az 1 ennek a számhalmaznak az egyetlen torlódási pontja. Az imént megfogalmazott állítás bizonyítása végett alkalmazzuk a számtani és mértani közepekre vonatkozó egyenlőtlenséget a következő n darab számra: az első $(n-2)$ db szám mindegyike legyen egyenlő 1-gyel, az utolsó két szám legyen egyenlő $\sqrt{n}$ -nel.

Ezekre a számokra alkalmazva a számtani és mértani közepekre vonatkozó egyenlőtlenséget, kapjuk, hogy

$$\sqrt[n]{\sqrt{n} \cdot \sqrt{n} \cdot 1^{n-2}} \leq \frac{2 \cdot \sqrt{n} + (n-2) \cdot 1}{n}, \text{ ez egy kicsit egyszerűbben írva}$$

$$\sqrt[n]{n} \leq \frac{2 \cdot \sqrt{n} + (n-2) \cdot 1}{n}.$$

Azt már megállapítottuk, hogy minden pozitív egész n szám esetén $1 \leq \sqrt[n]{n}$, ezért

$$1 \leq \sqrt[n]{n} \leq \frac{2 \cdot \sqrt[n]{n} + (n-2)}{n}.$$

Ezzel az $\sqrt[n]{n}$ értékét két másik érték közé szorítottuk be. Elég lenne annyit belátnunk, hogy n növekedtével $\frac{2 \cdot \sqrt[n]{n} + (n-2)}{n}$ értéke egyre közelebb kerül 1-hez, és ennek az 1-től való távolsága tetszőlegesen kicsi lehet.

Ehhez vizsgáljuk a $\frac{2 \cdot \sqrt[n]{n} + (n-2)}{n} - 1$ különbséget. Ez másképp

$$\frac{2 \cdot \sqrt[n]{n} + (n-2) \cdot 1}{n} - 1 = \frac{2 \cdot \sqrt[n]{n} + (n-2)}{n} - \frac{n}{n} = \frac{2 \cdot \sqrt[n]{n} - 2}{n}.$$

Erre pedig teljesül, hogy $0 \leq \frac{2 \cdot \sqrt[n]{n} - 2}{n} < \frac{2 \cdot \sqrt[n]{n}}{n} = \frac{2}{\sqrt[n]{n}}$. Világos, hogy ez a mennyiség n növekedtével egyre

kisebb lesz, és tetszőlegesen közel kerül a nullához (tehát a nulla lesz a $\frac{2}{\sqrt[n]{n}}$ alakú számokból álló halmaz egyetlen torlódási pontja).

Ezzel beláttuk, hogy az $\sqrt[n]{n}$ alakú számokból álló halmaz egyetlen torlódási pontja az 1.

A mérték fogalma

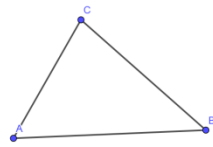
Ahogy arra már utaltunk, a valós számokat a gyakorlatban sokszor arra használjuk, hogy valamit mérjünk velük. Ez lehet például távolság, kerület, terület, felszín, de akár térfogat is. Célszerű összefoglalnunk, hogy mit várunk el ezektől a mértékektől (ezek lesznek az axiómáink).

A távolságra vonatkozó axiómáink:

1. Két pont távolsága a térben akkor és csak akkor nulla, ha a két pont egybeesik.

2. A térben bármely A és B pontra teljesül, hogy A-nak B-től vett távolsága megegyezik B-nek A-tól vett távolságával.

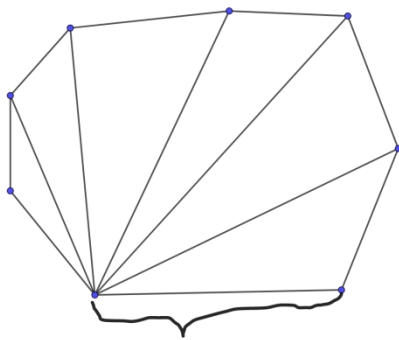
3. A térben bármely A, B, C ponthármasra fennáll, hogy az A-nak és B-nek a távolsága nem nagyobb, mint B-nek a C-től vett távolságának, illetve C-nek az A-tól vett távolságának összege.



$$AB \leq BC + CA$$

A távolságra vonatkozó axiómák közül a harmadik az, ami igazán figyelemreméltó. Ez ugyanis a középiskolában már megismert háromszög-egyenlőtlenség.

A háromszög-egyenlőtlenség többszöri alkalmazásával könnyen következik, hogy egy sokszög bármely oldala rövidebb, mint a többi oldal hosszának az összege. Ennek igazolásához elég háromszögekre vágnunk az adott sokszöget, és többször, egymás után alkalmazni a háromszög-egyenlőtlenséget.



A távolság tulajdonságait persze speciálisan vizsgálhatjuk a valós számegyenesen is. Egy x valós számnak a nullától vett távolságát az x szám abszolútértékének nevezzük, és $|x|$ -szel jelöljük.

Általánosan felírva, a valós számegyenesen az x és y pontok távolsága $|x-y|$ -nal egyenlő (ugyanis x -nek és y -nak az egymástól számított távolsága ugyanannyi, mint $(x-y)$ -nak a nullától vett távolsága).

Legyenek x és y tetszőleges valós számok. Ha a háromszög-egyenlőtlenséget alkalmazzuk a valós számegyenesen az x , $(-y)$, illetve nulla pontokra, akkor azt kapjuk, hogy

$|x-(-y)| \leq |x-0| + |-y-0|$, ami egyszerűbben felírva $|x+y| \leq |x| + |y|$. Ez a valós számok körében a háromszög-egyenlőtlenség.

Szedjük össze a területre vonatkozó axiómáinkat is:

A területre vonatkozó axiómák:

- 1. A síkban egy adott alakzat területe nem lehet negatív.*
- 2. Egybevágó alakzatok területe megegyezik.*
- 3. Ha egy alakzatot két, közös belső pont nélküli alakzatra vágunk, akkor a két alakzat területének összege megegyezik az eredeti alakzat területével.*
- 4. Az egységnyi oldalú négyzet területe egy.*

Az első három területaxióma elég kézenfekvő. A negyedikkel kapcsolatban érdemes megjegyezni, hogy ez az axióma nem általánosan a téglalapok területére vonatkozik (úgy, ahogy a középiskolában), hanem csupán az egységnyi oldalú négyzetre. Ebből kell most levezetnünk a középiskolában is használt, téglalapokra vonatkozó területképletet.

Tétel: Egy téglalap területe megegyezik oldalai hosszának szorzatával.

Bizonyítás:

Ha a téglalap oldalainak hossza egy-egy egész szám, akkor vágjuk fel a téglalapot egységnyi oldalú négyzetekre. A harmadik és negyedik területaxióma miatt ekkor a nagy téglalap területe pontosan akkora, ahány egységnégyzetre sikerült őt felválnunk. Ebben az esetben tehát a tétel állítása igaz.

Ha a téglalap oldalai közül mindkettő racionális, akkor jelölje az egyik oldal hosszát r/s , a másikat t/v , ahol r, s, t, v pozitív egész számok. Ennek a téglalapnak sv darab példányából össze tudunk rakni egy olyan téglalapot, aminek oldalai r , illetve t hosszúak. Mivel r és t egész hosszúak, ezért tudjuk, hogy a nagy téglalap területe rt . Használva az axiómáinkat, kapjuk, hogy az eredeti kis téglalap területe $\frac{rt}{sv}$, hiszen a nagy téglalapot a kis téglalapnak sv darab példányából raktuk össze.

Általános esetben, ha a téglalap oldalainak hossza A , illetve B , akkor közelítsük az A , illetve a B értéket is racionális számokkal (azért, mert racionális oldalú téglalapokra már beláttuk a tétel állítását).

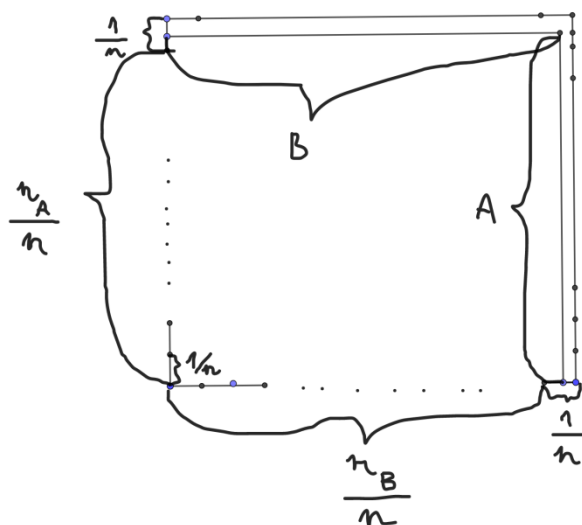
Jelöljük a téglalap meghatározandó területét T -vel.

Legyen n egy olyan pozitív egész szám, amelyre teljesül, hogy $\frac{1}{n}$ értéke az A és B értékénél is kisebb.

Ekkor (ahogy a számok racionális számokkal való közelítésénél láttuk) valamilyen n_A és n_B pozitív egészekre teljesül, hogy

$$\frac{n_A}{n} \leq A < \frac{n_A+1}{n} \text{ és } \frac{n_B}{n} \leq B < \frac{n_B+1}{n}.$$

Ekkor tehát egy $\frac{n_A}{n}$ és $\frac{n_B}{n}$ oldalú téglalap befér az eredeti, T területű téglalapba, míg egy $\frac{n_A+1}{n}$ és $\frac{n_B+1}{n}$ oldalú téglalapba befoglalható az eredeti téglalap.



Így a keresett T területre teljesül, hogy

$$\frac{n_A \cdot n_B}{n \cdot n} \leq T < \frac{n_A+1}{n} \cdot \frac{n_B+1}{n}.$$

Ugyanakkor tudjuk azt is, hogy minden elegendően nagy pozitív egész n számra

$$\frac{n_A \cdot n_B}{n \cdot n} \leq AB < \frac{n_A+1}{n} \cdot \frac{n_B+1}{n}.$$

Innen kellene bizonyítanunk, hogy $AB=T$. Ehhez elég lesz megnézni, hogy mekkora az egyenlőtlenséglánc két szélén elhelyezkedő téglalapterület különbsége:

$$\begin{aligned} \frac{n_A+1}{n} \cdot \frac{n_B+1}{n} - \frac{n_A \cdot n_B}{n \cdot n} &= \frac{(n_A+1)(n_B+1) - n_A n_B}{n^2} = \frac{n_A n_B + n_A + n_B + 1 - n_A n_B}{n^2} = \frac{n_A + n_B + 1}{n^2} = \frac{n_A}{n^2} + \frac{n_B}{n^2} + \frac{1}{n^2} = \\ &= \frac{1}{n} \cdot \frac{n_A}{n} + \frac{1}{n} \cdot \frac{n_B}{n} + \frac{1}{n^2} \leq \frac{1}{n} \cdot A + \frac{1}{n} \cdot B + \frac{1}{n^2}. \end{aligned}$$

Mivel n-et akármilyen nagynak választhatjuk, valamint A és B rögzítve van, ezért világos, hogy az

$\frac{1}{n} \cdot A + \frac{1}{n} \cdot B + \frac{1}{n^2}$ kifejezés infimuma nullával egyenlő.

Ezért szükségképpen $AB=T$, azaz a téglalapunk területe biztosan AB-vel egyenlő. ■

A térfogatra vonatkozó axiómák a területre vonatkozó axiómák analóg megfogalmazásai:

1. A térben egy adott test térfogata nem lehet negatív.
2. Egybevágó testek térfogata megegyezik.
3. Ha egy testet két, közös belső pont nélküli testre vágunk, akkor a két test térfogatának összege megegyezik az eredeti test térfogatával.
4. Az egységnyi oldalú kocka térfogata egy.

Ugyanazzal a módszerrel, amit a téglalapoknál használtunk, a téglatestekről is be lehet bizonyítani, hogy térfogatuk egyenlő a három oldalélük szorzatával.

Pithagorasz tétele és a tétel megfordítása

Most elérkeztünk oda, hogy be tudjuk bizonyítani a középiskolában már megismert Pithagorasz-tételt, a megfordításával együtt.

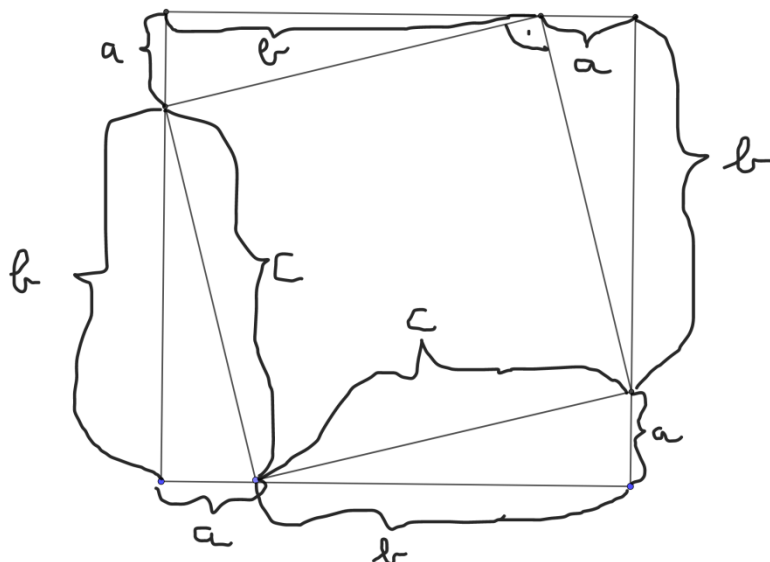
Pithagorasz tétele: Egy háromszög pontosan akkor derékszögű, ha két oldalának négyzetösszege megegyezik a harmadik oldal négyzetével.

Bizonyítás:

Két irányt kell bizonyítanunk. Először bebizonyítjuk, hogy ha a háromszög derékszögű, akkor a befogók (tehát az egymással derékszöget bezáró oldalak) négyzeteinek összege egyenlő a harmadik oldal négyzetével.

Legyenek a derékszögű háromszögünk oldalhosszai „a”, „b” és „c”, ahol a „c” hosszúságú oldallal szemben fekszik a derékszög (tehát ez az átfogó).

Szerkesszünk egy olyan négyzetet, aminek oldalhossza (a+b). A négyzet oldalain valamilyen irányban végighaladva jelöljük ki a csúcsoktól „a” távolságra fekvő osztópontokat az ábrán látható módon:



A belső osztópontok egy c oldalhosszúságú négyzetet határoznak meg, és ezen négyzet mellett létrejön még négy olyan derékszögű háromszög, melyeknek oldalhosszai „a”, „b” és „c”.

A nagy négyzet területe egyenlő a belső kisebb négyzet területének, illetve a négy kis egybevágó derékszögű háromszög területének összegével. Két ilyen derékszögű háromszögből összerakható egy téglalap, aminek oldalai „a” és „b” hosszúak, ezért a négy derékszögű háromszög összterülete $2ab$.

Tehát a nagy négyzet területe kétféleképpen kifejezve

$$(a + b)^2 = c^2 + 2ab, \text{ innen}$$

$$a^2 + 2ab + b^2 = c^2 + 2ab, \text{ azaz } a^2 + b^2 = c^2. \text{ Ezzel az egyik irányt beláttuk.}$$

Annyit kell még bizonyítani, hogy ha két oldal négyzetösszege egyenlő a harmadik oldal négyzetével, akkor a háromszög biztosan derékszögű. Legyen tehát adva egy olyan háromszög, amelynek oldalai „a”, „b” illetve „c” hosszúak, és amelyben $a^2 + b^2 = c^2$.

Tekintsük azt a háromszöget, amelynek két oldala „a” és „b” hosszú, ahol ezek az oldalak derékszöget zárnak be. Legyen a harmadik oldal hossza x. Az előzőek miatt tudjuk, hogy $a^2 + b^2 = x^2$.

Tehát ebből kapjuk, hogy $c^2 = x^2$. Mivel c és x egyaránt távolságokat jelöltek, ezért $c=x$ (hiszen a nemnegatív számok körében a gyökvonás egyértelmű). Ezért a háromszögnek valóban derékszögűnek kell lennie, ahogy azt állítottuk. ■

Külön szót kell ejtenünk az ilyen típusú tételekről, amelyekben több irányt kell bebizonyítanunk.

Két állítást akkor nevezünk **ekvivalensnek (egyenértékűnek)**, ha azok egymásból logikailag következnek. Ilyenkor élünk a „pontosan akkor”, másképpen az „akkor és csak akkor” szóhasználattal.

Fent tehát azt bizonyítottuk, hogy az alábbi két állítás ekvivalens (egymásból következnek):

1. *A háromszög derékszögű*

2. *A háromszögben két oldal négyzetének összege megegyezik a harmadik oldal négyzetével*

Azért volt tehát szükség két irányra a bizonyításunk során, mert be kellett látnunk, hogy az 1. állítás maga után vonja a 2. állítást, és a 2. állítás maga után vonja az 1. állítást.

A téglatest testátlójának hosszúsága

A Pithagorasz-tétel segítségével könnyen kifejezhetjük egy téglatest testátlójának (azaz két szemközti csúcst összekötő szakasznak) a hosszát a téglatest élének felhasználásával. Amennyiben a téglatest élei „a”, „b”, illetve „c” hosszúak, akkor először fejezzük ki csak egy olyan lapnak a lapátlóját, melynek oldalai „a” és „b” hosszúak. Ez a lapátló a Pithagorasz-tétel miatt $\sqrt{a^2 + b^2}$. Ahhoz, hogy a testátlót kifejezzük, észre kell vennünk, hogy a testátló átfogó egy olyan derékszögű háromszögben, melynek befogói $\sqrt{a^2 + b^2}$, illetve c hosszúak.

Így a testátló hossza (a Pithagorasz-tétel újabb felhasználásával) éppen

$$\sqrt{(\sqrt{a^2 + b^2})^2 + c^2} = \sqrt{a^2 + b^2 + c^2}.$$

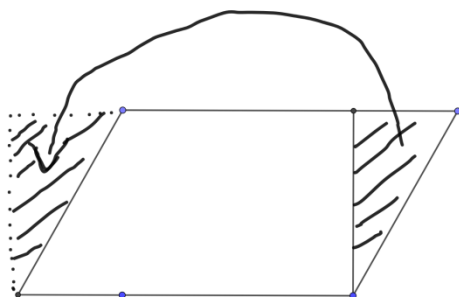
A paralellogramma és a háromszög területe

Parallelogrammának az olyan négyszögeket nevezzük, melyeknek szemközi oldalai párhuzamosak. Eddigi eredményeink alapján könnyen be tudjuk bizonyítani az alábbi tételt:

Tétel: *A parallelogramma területe megegyezik egy oldalának és az ezen oldallal szemközi oldaltól vett távolságnak a szorzatával.*

Bizonyítás:

Egy parallelogrammát át tudunk darabolni téglalalappá az ábrán látható módon:



Innen a tétel állítása (a területre vonatkozó axiómákat használva) következik. ■

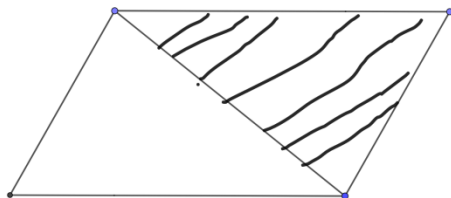
A szemközi oldalak távolságát a parallelogrammában **magasságnak** szoktuk nevezni (ezekből kettő van, mert két szemközi oldalpár van).

Az előző tételből már nagyon egyszerűen kapható a háromszögre vonatkozó területképlet.

Tétel: *Egy háromszög területe fele egy adott oldalának és az ezen oldallal szemközt elhelyezkedő csúcsnak a kiszemelt oldaltól vett távolságának szorzatával.*

Bizonyítás:

Bármely háromszöget ki tudunk egészíteni egy nála kétszer akkora területű parallelogrammává:



A paralelogramma területképletéből adódik az állítás. ■

Egy háromszögben egy csúsból a vele szemkötti oldalra merőleget állítva, a háromszög egy **magasságát** kapjuk (ezekből egy háromszögben három darab van).

Néhány megjegyzés vektorokról

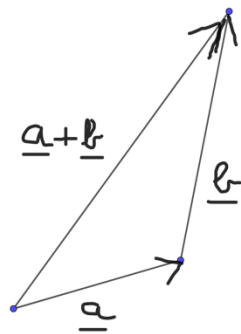
A vektoroknak a teljesen axiomatikus bevezetésére később lesz lehetőségünk, most csak szemlélet alapján bevezetjük a fontosabb kapcsolódó alapfogalmakat.

Térben vektoroknak nevezzük az ún. **irányított szakaszokat**. Tehát ezek olyan szakaszok, melyeknek van egy **kezdőpontja**, illetve egy **végpontja**. Két vektort pontosan akkor tekintünk egyenlőnek, ha párhuzamosak, egyforma hosszúak, és az irányításuk is azonos.

A vektor hosszát másnéven a vektor **abszolútértékének** nevezzük.

A vektorok körében is létezik ún. **nullvektor**, illetve a vektorokat szorozhatjuk ún. **skalárokkal**. Ezek a skalárok általában valamilyen valós számok. Negatív számmal történő szorzáskor a vektor irányítása megváltozik, a vektor hossza pedig a skalár abszolútértékének megfelelően változik.

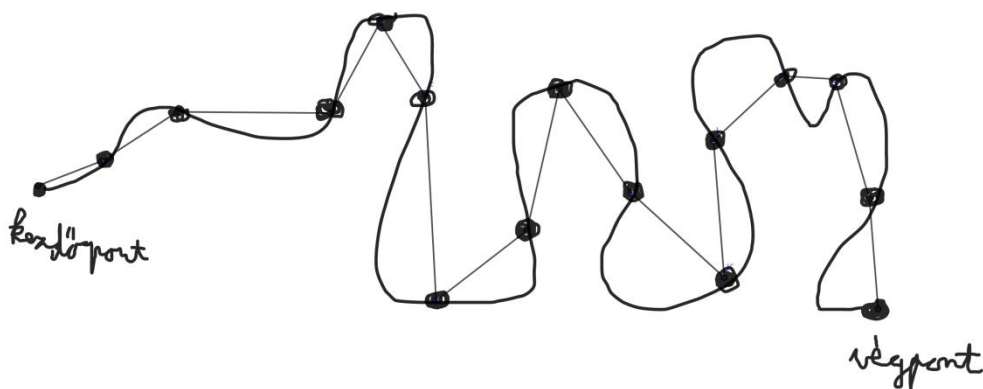
Mivel a vektorok szemléletesen valamilyen eltolást, illetve térbeli elmozdulást írnak le, ezért vektorok összegét rendszerint a **láncszabály** segítségével értelmezzük. Ez azt jelenti, hogy geometriailag megnézzük, hogy az összeadandó vektorok együttesen milyen eltolást eredményeznek.



Görbék ívhosszáról

Mivel két pont távolságát tudjuk mérni, ezért viszonylag kézenfekvő, hogyan kell értelmeznünk pl. sokszögek területét. Azonban sok esetben szükség lehet arra, hogy bizonyos görbék ívhosszát meghatározzuk.

Legyen adva egy térbeli görbe, aminek van egy irányítása (ezt az egyszerűbb tárgyalás kedvéért vezetjük be). A térbeli irányított görbénknek legyen egy kezdő- és egy végpontja, melyek akár egybe is eshetnek. A kezdőponttól a végpontig haladva vegyünk fel egymás után véges sok pontot a görbén, melyeknek számozása a kezdőponttól a végpontig haladva egyesével növekszik.



Kössünk össze minden egyes kiválasztott pontot a rákövetkező ponttal, és mindig mérjük meg az így kapott szakasz hosszát. Adjuk össze ezeket a hosszúságokat, így a görbe hosszának egy közelítő értékét kapjuk.

A kulcskérdés az, hogy miként kellene értelmeznünk a görbe ívhosszát. A válasz a szemlélet alapján rendkívül egyszerű. Mivel a szakaszok összhossza a szemlélet miatt legfeljebb olyan hosszú, mint a közelített ívhossz, ezért definíció szerint az imént bevezetett (alulról) **közelítő értékek halmazának szuprémuma** lesz a görbe ívhossza (természetesen ez konszenzus alapján történik így).

Töröttvonalak és ívhosszak

Egymáshoz kapcsolódó szakaszok úgynevezett **töröttvonalat** alkotnak. Az irányított görbe ívhosszát a töröttvonalak segítségével úgy lehetne definiálni, mint az **irányított görbén haladás szerint egymás utáni pontokat összekötő töröttvonalak hosszának szuprémumát**.

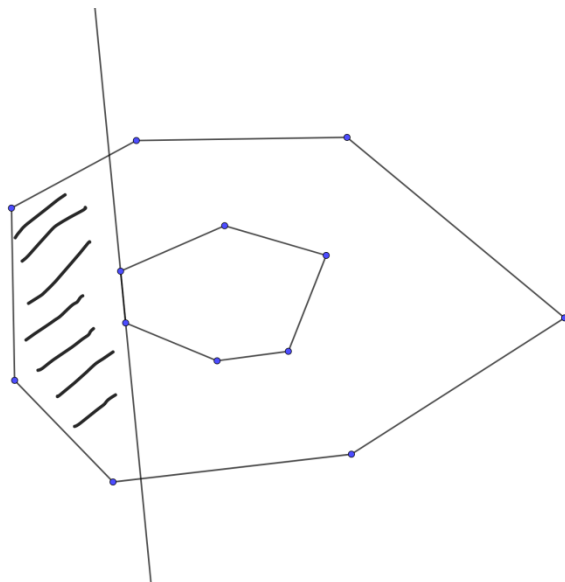
Egy hasznos tétel egymást tartalmazó konvex sokszögekkel kapcsolatosan

A későbbiekben többször is szükségünk lesz egymást tartalmazó síkbeli alakzatok területének összehasonlítására. Ezzel kapcsolatosan fontos és alapvető tétel a következő:

Tétel: *Ha a síkban egy konvex sokszög tartalmaz egy másik konvex sokszöget, akkor a tartalmazó sokszög területe legalább akkora, mint a tartalmazott sokszögé.*

Bizonyítás:

Először is a tartalmazott sokszög egyik oldalának egyenesével levágunk egy darabot a tartalmazó sokszögből. Miután egy ilyen lépést elvégzünk, a háromszög-egyenlőtlenség következtében a tartalmazó sokszög területe nem növekedhet, a tartalmazotté pedig változatlan marad.



Világos, hogy a fenti lépés véges sokszori ismételtetésével eljutunk a tartalmazott sokszöghöz, és ezzel a tételt beláttuk. ■

A kör kerülete, a π szám bevezetése

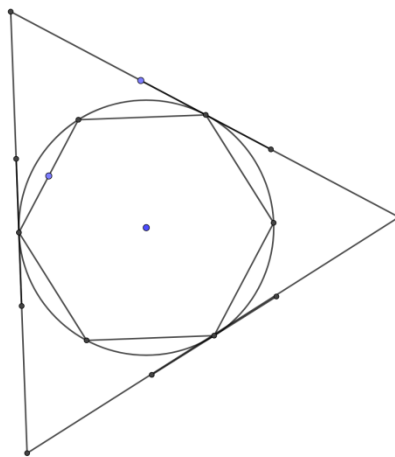
A hasonlósági transzformációk részletes tárgyalását mellőzve abból indulunk ki, hogy bármely két kör hasonló egymáshoz, ezáltal a kör kerülete egyenesen arányosan változik a sugarának hosszával.

Tulajdonképpen annyi dolgunk van, hogy bebizonyítsuk, hogy **egy egységnyi sugarú körnek valóban véges hosszú a kerülete**.

Elég annyit belátnunk, hogy az egységkörnek mint görbének véges az ívhossza.

A görbék ívhosszáról elmondottak alapján az egységkör kerülete a belé írt sokszögek kerületeinek szuprénuma. Tekintsük most a beírt sokszögeket. Ezen beírt sokszögek mindegyikének kerülete kisebb, mint a kör kerülete.

Tekintsük ugyanakkor az egységkör köré írt szabályos háromszöget is. Az egymást tartalmazó konvex sokszögek kerületéről szóló tétel miatt a körülírt szabályos háromszög kerülete nagyobb, mint akármelyik beírt sokszög kerülete. Ebből viszont következik, hogy a beírt sokszögek kerületeinek szuprénuma nem lehet nagyobb, mint a körülírt szabályos háromszög kerülete.



Ha ugyanis a beírt sokszögek kerületeinek szuprénuma nagyobb lenne, mint a körülírt szabályos háromszög kerülete, akkor lenne olyan beírt sokszög, aminek kerülete nagyobb a körülírt szabályos háromszög kerületénél, ez pedig az egymást tartalmazó konvex sokszögek kerületéről szóló tételnek ellentmondana.

Tehát a beírt sokszögek kerületeinek halmaza felülről korlátos, ezért ennek a halmaznak van szuprénuma, vagyis az egységkörnek létezik véges kerülete.

Ezáltal lehetőségünk nyílik a π szám bevezetésére.

Definíció: Az egységkör kerületének hossza 2π .

Körívek és középponti szögek (fejezet a szögek mértékéről)

A teljes körhöz mint körívhez tartozó középponti szöget **teljesszögnek** nevezzük. A félkörhöz mint körívhez tartozó középponti szöget **egyenesszögnek** nevezzük.

Alapvetően fontos eredmény, hogy a körben egy körív hossza és a hozzá tartozó középponti szög egymással egyenes arányosságban van. Ezt mondja ki a következő tétel:

Tétel: Adott körben két körív hossza úgy aránylik egymáshoz, mint hozzájuk tartozó középponti szögek.

Bizonyítás:

Ezt a tételt lényegében ugyanúgy bizonyíthatjuk, mint a téglalapok területéről szóló tételt.

Induljunk ki az egységnyi sugarú körből, melynek kerülete 2π .

Legyen először n valamilyen pozitív egész szám. Tekintsünk egy olyan körívet, amihez tartozó középponti szögnek a mértéke a teljesszög $1/n$ -edrésze. Világos, hogy egy ilyen körív hossza $2\pi/n$, ugyanis n darab ilyen körívből rakható ki az egész körkerület.

Most legyenek $0 < k < n$ egész számok, és tekintsünk egy olyan körívet, amihez tartozó középponti szögnek a mértéke a teljesszög k/n -edrésze. Az előzőekből világos, hogy az ehhez tartozó körív hossza $2k\pi/n$.

Végül legyen $0 < \alpha < 1$ irracionális szám, és vegyünk egy olyan körívet, amihez tartozó középponti szögnek a mértéke a teljesszög α -szorosa. Legyen n olyan pozitív egész szám, aminek reciproka kisebb α -nál. A számok racionális számokkal való közelítéséről szóló tételhez hasonlóan kapjuk, hogy létezik olyan k pozitív egész szám, amelyre $\frac{k}{n} \leq \alpha < \frac{k+1}{n}$. Tehát az adott középponti szöghöz tartozó körív hossza biztosan $\frac{k}{n} \cdot 2\pi$ és $\frac{k+1}{n} \cdot 2\pi$ közé esik. Mivel n -et akármilyen nagynak választhatjuk, ezért a

szóban forgó körív hossza biztosan $\alpha \cdot 2\pi$ -vel egyenlő, ugyanis $\frac{k}{n}$ és $\frac{k+1}{n}$ mindig közrefogja az α számot, és ezen két szám különbsége $\frac{k+1}{n} - \frac{k}{n} = \frac{1}{n}$, ami tetszőlegesen kicsi lehet. ■

A most bizonyított tétel alapján van létjogosultsága annak, hogy a teljesszöget is 2π -vel jelöljük (az egyenesszöget pedig természetesen π -vel).

Egyenletek és egyenletrendszerek megoldási módszereiről

Ebben a fejezetben megpróbáljuk áttekinteni az egyenletek, illetve egyenletrendszerek megoldásának alapvető módszereit.

Egyenletek megoldása során tudnunk kell, hogy mindkét oldalt ugyanannyival növelve (vagy csökkentve) a kiinduló egyenlettel egyenértékű (azaz ekvivalens) egyenlethez jutunk. Ugyanez a

helyzet, ha az egyenlet mindkét oldalát ugyanazzal a nullától különböző valós számmal szorozzuk. Ezek könnyen következnek a valós számokról megfogalmazott axiómákból.

Vigyáznunk kell azonban egyéb lépések alkalmazása során. Például ha mindkét oldalt négyzetre emeljük, akkor a kapott új egyenlet már nem feltétlenül lesz ekvivalens az eredetivel.

Egyváltozós polinomok

Elsőfokú egyenleteket könnyűszerrel meg tudunk oldani (tehát olyanokat, mint pl. $3x-6=0$).

A **másodfokú egyenletek megoldásához** már megoldóképletet használtunk a középiskolában. Ezt a megoldóképletet itt most levezetjük. Legyen a megoldandó egyenletünk

$Ax^2 + Bx + C = 0$, ahol A, B, C adott valós számok, és x az ismeretlenünk.

Feltehetjük, hogy $A \neq 0$, különben ez egy szimpla elsőfokú egyenlet. Osszuk le mindkét oldalt A -val.

Az így kapott ekvivalens egyenlet

$$x^2 + \frac{B}{A}x + \frac{C}{A} = 0.$$

Használva a kéttagú összeg négyzetére vonatkozó azonosságot, ez átírható úgy, hogy

$$\left(x + \frac{B}{2A}\right)^2 - \left(\frac{B}{2A}\right)^2 + \frac{C}{A} = 0, \text{ ami átrendezve } \left(x + \frac{B}{2A}\right)^2 = \left(\frac{B}{2A}\right)^2 - \frac{C}{A}.$$

Innen x -et könnyen kifejezhetjük, amennyiben ismerjük azokat a valós számokat, amiknek négyzete egyenlő a $\left(\frac{B}{2A}\right)^2 - \frac{C}{A}$ különbséggel.

Azt már tudjuk, hogy amennyiben létezik olyan valós x , ami megoldása az egyenletnek, akkor

$$\left(\frac{B}{2A}\right)^2 - \frac{C}{A} \geq 0.$$

Ha elfogadjuk a gyökvonásról középiskolában tanultakat, akkor a megoldásokat úgy fejezhetjük ki, hogy

$$x_{1,2} = -\frac{B}{2A} \pm \sqrt{\left(\frac{B}{2A}\right)^2 - \frac{C}{A}}$$

Közös nevezőre hozás után kapjuk az $x_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$ megoldóképletet. Itt $(B^2 - 4AC)$ az egyenlet ún. **diszkriminánsa**.

Most megpróbáljuk az eddigieket valamilyen formában általánosítani.

Ha n természetes szám, és $c_0, c_1, \dots, c_n$ valamilyen számok, akkor a

$p(x) = c_n \cdot x^n + c_{n-1} \cdot x^{n-1} + \dots c_1 \cdot x + c_0$ alakú kifejezéseket **egyváltozós polinomoknak** nevezzük. Ennek változója itt az x , a $c_0, c_1, \dots, c_n$ számok az úgynevezett együtthatók. Ha $c_n \neq 0$, akkor a **polinom ún. foka** egyenlő n -nel (amennyiben $n=0$ és $c_0=0$, akkor a polinomnak definíció szerint nincsen foka).

Eddig az első- és másodfokú polinomok ún. **zérushelyeit/gyökeiket** kerestük (tehát azokat a helyeket, ahol a polinom a nulla értéket veszi fel).

Természetesen a zérushelyek keresése jóval nehezebb, ha a polinom foka nagyobb kettőnél.

Ha a polinom foka három vagy négy, akkor létezik megoldóképlet, amivel megtalálhatjuk a zérushelyeket (ezek az ún. **Cardano-féle képletek**).

Mély algebrai eszközökkel bizonyítható, hogy négynél magasabb fokú polinomokra nézve nem létezik „rendes” megoldóképlet (lényegében ezt mondja ki az **Abel-Ruffini-tétel**).

Persze ez nem jelenti azt, hogy bizonyos magasabb fokú polinomoknak sosem tudjuk megtalálni a zérushelyeit. Egyszerűen csak annyit jelent, hogy nincsen kész képlet, amibe be tudnánk helyettesíteni.

Adunk egy példát olyan ötödfokú polinomra, melynek zérushelyeit meg tudjuk találni.

Legyen $p(x) = x^5 - 5x^4 + 6x^3$. Ha ennek a polinomnak a zérushelyeit keressük, akkor meg kell oldanunk az $x^5 - 5x^4 + 6x^3 = 0$ egyenletet. Vegyük észre, hogy x^3 -t ki tudunk emelni.

$$x^3(x^2 - 5x + 6) = 0.$$

Egy szorzat csak úgy lehet nulla, ha valamelyik szorzótényezője nulla. Tehát x^3 -nak, vagy $(x^2 - 5x + 6)$ -nak egyenlőnek kell lennie nullával. Amennyiben $x^3 = 0$, akkor $x=0$.

Ha $x^2 - 5x + 6 = 0$, akkor a másodfokú megoldóképlet felhasználásával kapjuk az $x=2$, illetve $x=3$ megoldásokat. Tehát az adott polinomnak három valós zérushelye van, ezek $x=0,2,3$.

A Bézout-tétel

A $p(x) = x^5 - 5x^4 + 6x^3$ polinom zérushelyeinek keresésekor a nullát mint zérushelyet könnyű volt megtalálni. Ez egész egyszerűen annak köszönhető, hogy az adott polinomunkban a c_0 konstans értéke nulla, hiszen $x^5 - 5x^4 + 6x^3 = x^5 - 5x^4 + 6x^3 + 0 \cdot x^2 + 0 \cdot x + 0$, azaz x -et biztosan ki lehet emelni, így $x=0$ zérushely.

Ezen észrevételünk általánosítása az alábbi tétel:

Bézout-tétel: Legyen n pozitív egész szám. Amennyiben a $p(x) = c_n \cdot x^n + c_{n-1} \cdot x^{n-1} + \dots c_1 \cdot x + c_0$ polinomnak zérushelye az r szám, akkor ebből a polinomból kiemelhető $(x-r)$.

Bizonyítás:

A Bézout-tételt legegyszerűbben a polinom fokára vonatkozó teljes indukcióval bizonyíthatjuk be. Ha a polinom elsőfokú, akkor ez egyszerűen belátható. Legyen ugyanis adott a $p(x) = Ax+B$ elsőfokú polinom (tehát itt $A \neq 0$), aminek zérushelye r , azaz $Ar+B=0$. Innen átrendezéssel kapjuk, hogy $r = -\frac{B}{A}$.

Tehát $Ax+B = A\left(x + \frac{B}{A}\right) = A(x-r)$, ezzel az állítást elsőfokú polinomokra beláttuk.

Tegyük fel, hogy az állítást beláttuk, ha a polinom foka legfeljebb k , és most szeretnénk $(k+1)$ -edfokú polinomokra is bizonyítani.

Legyen tehát $q(x) = c_{k+1} \cdot x^{k+1} + c_k \cdot x^k + \dots c_1 \cdot x + c_0$ olyan polinom, aminek zérushelye az r szám.

Vonjuk ki $q(x)$ -ből $c_{k+1} \cdot (x-r)^{k+1}$ -et, ekkor kapjuk a

$c_{k+1} \cdot x^{k+1} - c_{k+1} \cdot (x-r)^{k+1} + c_k \cdot x^k + \dots c_1 \cdot x + c_0$ polinomot.

Ennek az új polinomnak szintén zérushelye az r , mivel $q(x)$ -nek és $(x-r)^{k+1}$ -nek is zérushelye az r . Ugyanakkor az új polinom foka kisebb $(k+1)$ -nél, mivel a kivonáskor x^{k+1} teljesen kiesik.

Az indukciós feltevés szerint az új polinomból kiemelhető $(x-r)$, a kivont $c_{k+1} \cdot (x-r)^{k+1}$ -ből szintén kiemelhető $(x-r)$, ezért $q(x)$ -ből is kiemelhető $(x-r)$, és ezt kellett bizonyítani. ■

Tehát azt kaptuk, hogy egy polinom zérushelyeinek keresése lényegében megfelel annak, hogy az adott polinomból szeretnénk kiemelni elsőfokú szorzótényezőket.

A már vizsgált $p(x) = x^5 - 5x^4 + 6x^3$ polinomot pl. teljesen szorzattá tudjuk bontani, ugyanis

$x^5 - 5x^4 + 6x^3 = (x-0)^3 \cdot (x-2) \cdot (x-3)$. Megfigyelhetjük, hogy a zérushelyeknek van egy ún. **multiplicitása** (ez tehát azt fejezi ki, hogy hányszoros gyök/zérushely az adott zérushely). Ennél az ötödfokú polinomnál a nullának mint zérushelynek a multiplicitása 3, a kettő és a három multiplicitása pedig 1.

Függvényekről általánosan

A függvényekről azt tanuljuk középiskolában, hogy ezek valamilyen egyértelmű hozzárendelések. Az alábbi definícióban részletesen elmagyarázzuk, hogy ez pontosan mit takar.

Definíció (Függvény): Függvény alatt egyértelmű hozzárendelést értünk, ami az alábbi módon jellemezhető:

1. adott egy A halmaz (ami az ún. **értelmezési tartomány**), illetve ettől függetlenül egy B halmaz.
2. az A halmaz minden eleméhez hozzárendeljük B -nek pontosan egy elemét.

*A függvény ún. **értékkészlete** a B halmaz azon elemeinek halmaza, amik hozzá lettek rendelve legalább egy A-beli elemhez.*

Az ilyen típus függvényekre nézve általánosan az $f: A \rightarrow B$ jelölést szoktuk használni (tehát A az értelmezési tartomány, az értékkészlet elemei pedig B-ből valók).

Fontos megjegyeznünk, hogy itt nem biztos, hogy minden B-beli elem hozzá lett rendelve valamelyik A-belihez, és létezhet olyan B-beli elem, ami több A-belihez is hozzá lett rendelve.

*Az olyan függvényeket, amiknek értelmezési tartománya a valós számoknak egy részhalmaza, és a felvett értékeik is valós számok, a továbbiakban a rövidség kedvéért egyszerűen **valós függvényeknek** fogjuk nevezni.*

Néhány valós függvény, amiket már ismerünk

Egyszerű és alapvető függvények az alábbiak (a teljesség igénye nélkül):

1. Az **identitásfüggvény**, ami az értelmezési tartomány minden egyes elemére ezt az adott elemet veszi fel értéként (formálisan tehát $f(x)=x$ teljesül minden értelmezési tartománybeli x -re)

2. A **konstansfüggvények** (ezek mindenütt ugyanazt az értéket veszik fel)

3. A **polinomok**

4. A **hatványfüggvények**, azaz az $f(x)=x^c$ előírással kapható függvények, ahol c valamilyen konstans, és x a változó (ilyenkor általánosan az értelmezési tartomány a nemnegatív számok halmaza)

5. Az **exponenciális függvények**, azaz az $f(x)=c^x$ előírással kapható függvények, ahol c valamilyen pozitív konstans, és x a változó

Az itt leírt függvényeket már gond nélkül tudjuk értelmezni az eddig tanultak segítségével.

Növekedés és csökkenés

Definíció: Egy valós függvényt **szigorúan monoton növekvőnek** nevezünk, ha a behelyettesített értékek növekedtével a függvény által felvett értékek is egyre növekednek.

Jó példa szigorúan monoton növekvő függvényre a valós számok halmazán értelmezett identitásfüggvény. De ismerünk ennél általánosabb példákat is. A hatványok kiterjesztésével foglalkozó részben láthattuk, hogy tetszőleges, egynél nagyobb alapú exponenciális függvény szigorúan monoton növekvő, azaz tetszőleges $A > 1$ rögzített valós szám esetén a valós $f(x)=A^x$ függvény szigorúan monoton növekvő.

Definíció: Egy valós függvényt **monoton növekvőnek** nevezünk, ha a behelyettesített értékek növekedtével a a függvény által felvett értékek nem csökkennek.

Jó példa monoton növekvő függvényre bármely konstansfüggvény.

Definíció: Egy valós függvényt **szigorúan monoton csökkenőnek** nevezünk, ha a behelyettesített értékek növekedtével a függvény által felvett értékek egyre csökkennek.

Jó példa szigorúan monoton csökkenő függvényre bármely exponenciális függvény, melynek alapja 0 és 1 közé esik, azaz tetszőleges $0 < A < 1$ rögzített valós szám esetén a valós $f(x)=A^x$ függvény szigorúan monoton csökkenő.

Definíció: Egy valós függvényt **monoton csökkenőnek** nevezünk, ha a behelyettesített értékek növekedtével a függvény által felvett értékek nem növekednek.

Jó példa monoton csökkenő függvényre bármely konstansfüggvény.

Szélsőértékek, minimumok és maximumok

Az ún. **abszolút (másnéven globális) szélsőértékek** definíciója kézenfekvő, először ezeket tekintjük át.

Definíció: Ha egy valós függvénynek van legnagyobb felvett értéke, akkor ezt az értéket a függvény **abszolút (másnéven globális) maximumának** nevezzük. Azokat a helyeket, amelyeken az ilyen függvény a maximális értékét felveszi, **maximumhelyeknek** hívjuk.

Például a valós $f(x) = 1-x^2$ függvénynek abszolút maximuma az 1. Az egyetlen maximumhely az $x=0$ pont, mivel a függvény csak itt veszi fel az 1 értéket. Ennek a függvénynek nincsen minimuma, mivel értékkészlete alulról nem korlátos.

Definíció: Hasonlóan, ha egy valós függvénynek van legkisebb felvett értéke, akkor ezt az értéket a függvény **abszolút (másnéven globális) minimumának** nevezzük. Azokat a helyeket, amelyeken az ilyen függvény a minimális értékét felveszi, **minimumhelyeknek** hívjuk.

Például a valós $g(x) = (x - 1)^2$ függvénynek abszolút minimuma a 0 érték, és az egyetlen minimumhelye van, ez pedig az $x=1$ (hiszen csak itt lesz a függvény értéke 0). Ennek a függvénynek viszont nincsen maximuma, mivel értékkészlete felülről nem korlátos.

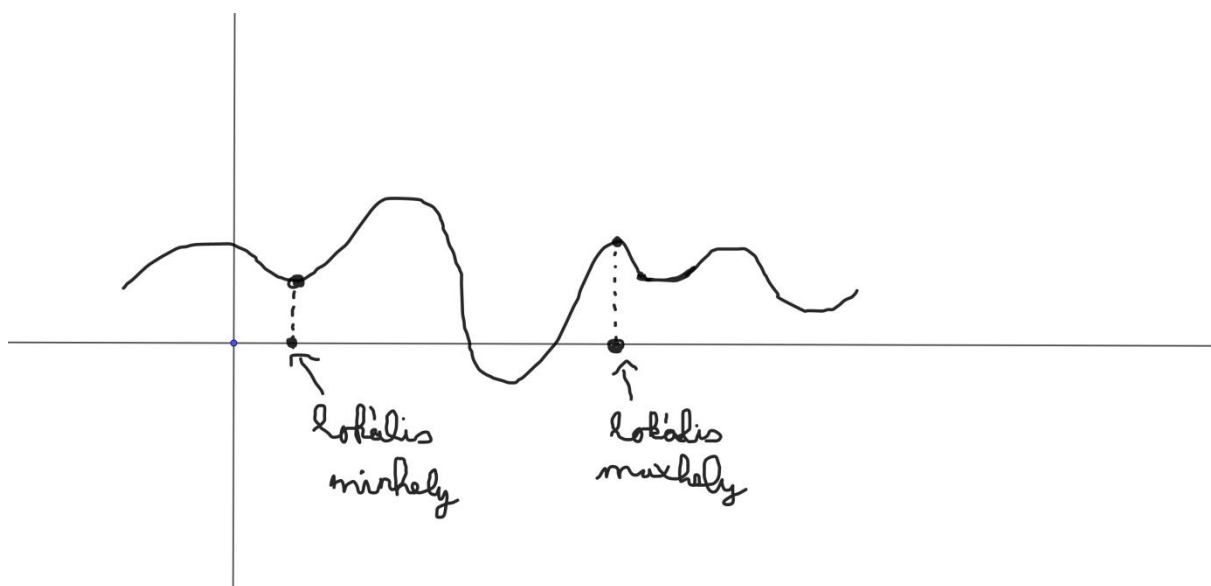
Némi magyarázatra szorul a **lokális (másnéven helyi) szélsőértékek** értelmezése.

Definíció: Az $f(x)$ valós függvény értelmezési tartományának u pontja lokális maximumhelye a szóban forgó függvénynek, ha van az u -nak olyan pozitív sugarú környezete, hogy bármely olyan értelmezési tartománybeli x -re, ami ebből a környezetből való, teljesül, hogy $f(x) \leq f(u)$.

Az előzőhöz hasonló módon vezethető be a lokális minimumhely fogalma is:

Definíció: Az $f(x)$ valós függvény értelmezési tartományának u pontja lokális minimumhelye a szóban forgó függvénynek, ha van az u -nak olyan pozitív sugarú környezete, hogy bármely olyan értelmezési tartománybeli x -re, ami ebből a környezetből való, teljesül, hogy $f(u) \leq f(x)$.

Ahhoz, hogy az olvasó kapjon egy szemléletes képet a lokális szélsőértékek jelentéséről, célszerű elemeznie az alábbi ábrát:



Műveletek függvényekkel

Amennyiben az $f(x)$ és a $g(x)$ valós függvények értelmezési tartománya megegyezik, akkor rájuk nézve természetes módon tudjuk értelmezni a hagyományos műveleteket, amiket a valós számoknál megismertünk.

Definíció (valós függvények közti műveletek):

Legyenek $f(x)$ és $g(x)$ olyan valós függvények, amelyeknek értelmezési tartománya azonos.

Ezen függvények összege az a valós függvény, aminek az x helyen felvett értéke $f(x) + g(x)$.

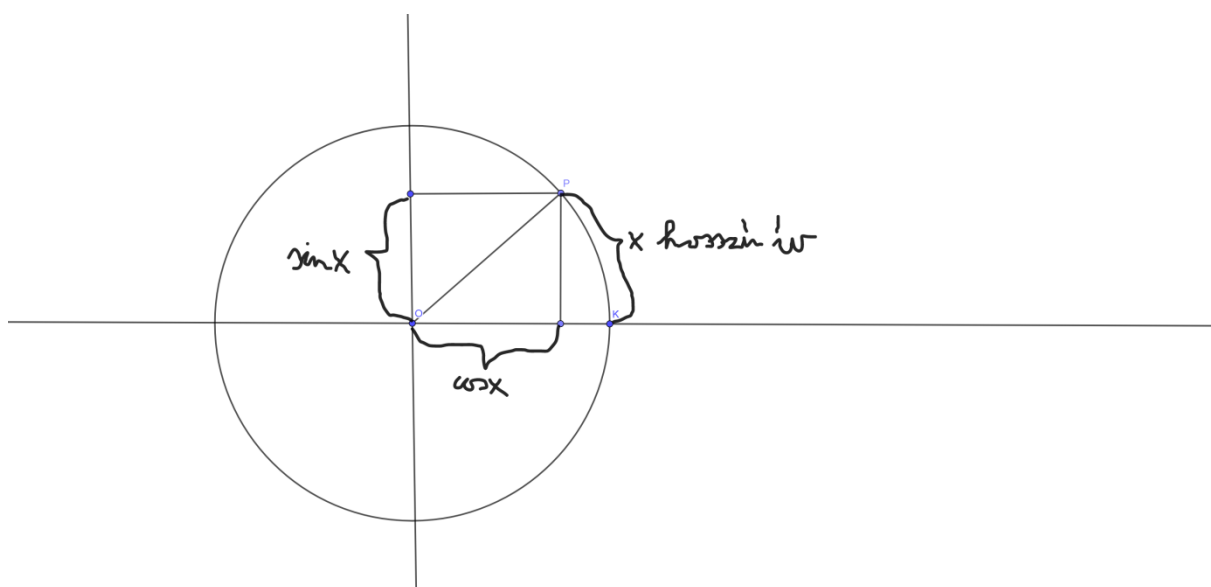
Ezen függvények szorzata az a valós függvény, aminek az x helyen felvett értéke $f(x) \cdot g(x)$.

Ezen függvények f/g hányadosfüggvénye olyan x -ekre van értelmezve, melyekre $g(x) \neq 0$, és ekkor a hányadosfüggvény értéke egyenlő $f(x)/g(x)$ -szel.

A szögfüggvények

A függvények egy fontos osztályát alkotják a szögfüggvények. Ezek a legtermészetesebb úton geometriailag értelmezhetők. A szögfüggvények értelmezéséhez a Descartes-féle derékszögű koordináta-rendszert fogjuk használni.

Az origó (O pont) körül rajzolunk egy egységnyi sugarú kört a koordináta-rendszerben. Messe ez a kör a vízszintes tengelyt a K pontban. Tekintsük az egységnyi sugarú kör egy tetszőleges pontját, legyen ez P. Legyen az egységgörön a K-tól P-ig az óramutató járásával ellentétes irányban vezető ív hossza x .



A $\sin x$ értéke definíció szerint a P pontnak a függőleges tengelyen vett koordinátája, míg a $\cos x$ értéke definíció szerint a P pontnak a vízszintes tengelyen vett koordinátája lesz. Így módon definiáltuk a szinuszfuggvényt és a koszinuszfuggvényt is.

Az ábrán csak az első síknegyedben szemléltettük a szinusz- és a koszinuszfuggvény értelmezését, de természetesen a összes síknegyedben ugyanígy tudunk eljárni.

Mivel a kör, ahonnan kiválasztottuk a P pontot, egységnyi sugarú, ezért a Pithagorasz-tétel felhasználásával kapjuk, hogy minden egyes x -szögre

$\sin^2 x + \cos^2 x = 1$. Ezt az összefüggést trigonometrikus Pithagorasz-tételként is szokták emlegetni.

A tangens- és kotangensfüggvényeket már tudjuk értelmezni a szinusz- és a koszinuszfuggvény segítségével.

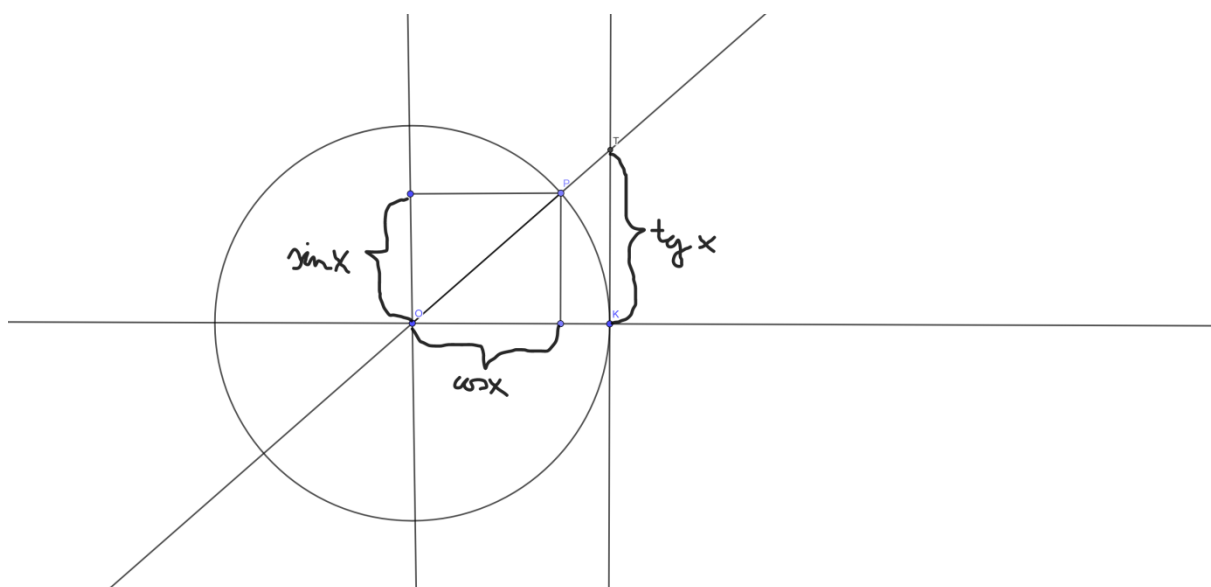
A tangens értelmezése formálisan a $\operatorname{tg} x = \frac{\sin x}{\cos x}$, a kotangensé $\operatorname{ctg} x = \frac{\cos x}{\sin x}$ előírással történik. Fontos

azonban arra figyelniük, hogy ezek csak akkor vannak értelmezve, ha a nevezőbeli mennyiségek nullától különbözőek. Tehát a kotangensfüggvény nincsen értelmezve a π egész számú többszörösein, a kotangensfüggvény pedig nincs értelmezve azokon a helyeken, amik $\pi/2$ -től a π -nek valamilyen egész számú többszörösével térnek el.

A geometriai értelmezés alapján könnyen beláthatók az alábbi összefüggések:

$$\sin(\pi - x) = \sin x, \cos(\pi - x) = -\cos x, \sin\left(\frac{\pi}{2} - x\right) = \cos x$$

Hasznos lehet, ha a tangens- és kotangensértékeket is tudjuk geometriailag értelmezni. Húzzuk meg az egységgörnek az y-tengellyel párhuzamos érintőit. Ekkor az OP szakasznak a P ponton túli meghosszabbítása (ha a tangens értelmezve van) az érintők közül pontosan az egyiket metszi, legyen a metszéspont T.



Az ábrán látható hasonló háromszögek előjeles oldalhosszából kapjuk, hogy $\frac{\sin x}{\cos x} = \frac{KT}{OK} = KT$, azaz a KT-nek az előjeles hossza éppen $\operatorname{tg} x$ -szel egyenlő.

Ugyanígy értelmezhetjük geometriailag a kotangenst, csak ott az x-tengellyel párhuzamos érintőket kell megszerkesztenünk.

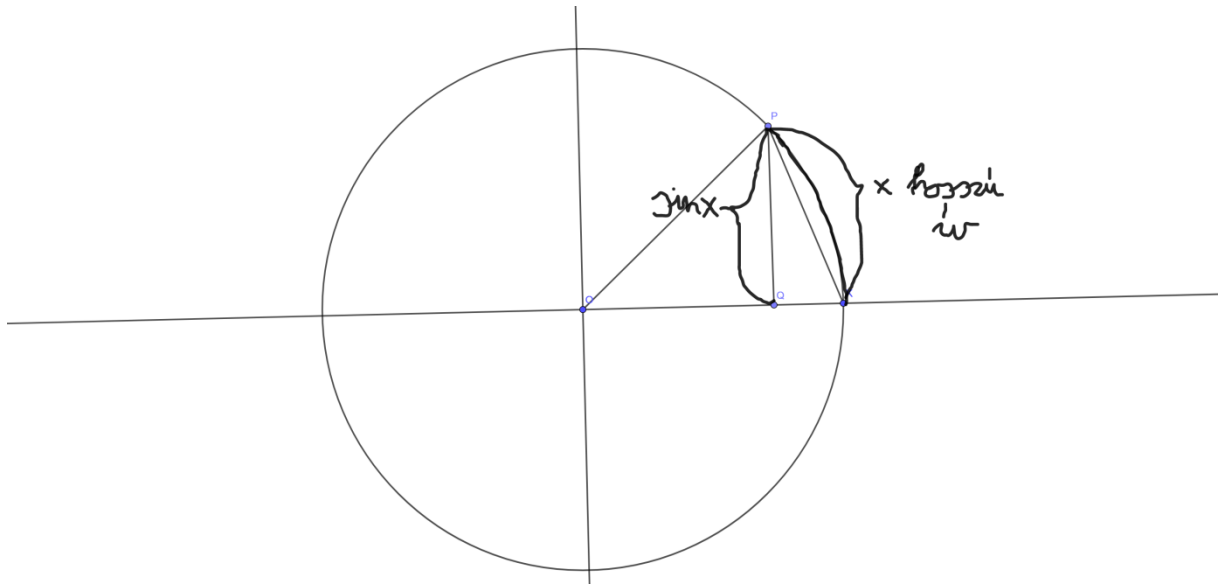
Néhány fontos egyenlőtlenség szögfüggvényekkel kapcsolatban

Belátjuk az alábbi, a szinuszra és a tangensre vonatkozó egyenlőtlenségláncot.

Tétel: Amennyiben $|x| < \frac{\pi}{2}$, akkor $|\sin x| \leq |x| \leq |\operatorname{tg} x|$.

Bizonyítás:

Tekintsük az origó középpontú egységgörte, az origó legyen O, az egységgörte a vízszintes tengely pozitív felét messe a K pontban. Adva egy P pont a körvonalon úgy, hogy ennek a P-nek a vízszintes tengelyre vonatkozó Q vetülete a tengely pozitív felére esik. Ez azt jelenti, hogy a K pontból a P pontba legfeljebb $\frac{\pi}{2}$ hosszú, előjelesen vett köríven tudunk eljutni, legyen ennek a körívnek a hossza $|x|$.



A PQK háromszög derékszögű, ebben az egyik befogó PQ, aminek hossza egyenlő $|\sin x|$ -szel.

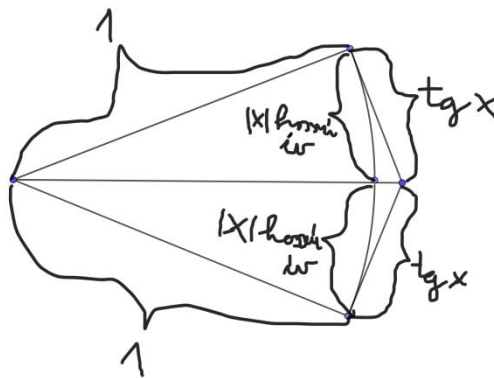
A derékszögű háromszög egy befogója nem hosszabb az átfogónál, ezért

$|\sin x| = PQ \leq PK$. Ugyanakkor az ívhossz értelmezése miatt a PK biztosan nem hosszabb, mint a P és K pontokat összekötő, $|x|$ hosszúságú körív. Ebből kapjuk, hogy $|\sin x| \leq |x|$.

A tangensre vonatkozó egyenlőtlenség igazolásához is a geometriai értelmezést használjuk. Egy olyan derékszögű háromszögből kell kiindulnunk, aminek egyik befogója egységnyi, a másik pedig $|\tan x|$ hosszúságú. Az $|x|$ hosszú körív itt az ábrán látható módon kapható:



Tükrözzük ezt az egész alakzatot az átfogó egyenesére:



Azt fogjuk bebizonyítani, hogy az így kapott deltoid kerülete nagyobb, mint annak az alakzatnak a kerülete, amit a két egység hosszúságú szakasz, illetve a $2|x|$ hosszúságú körív meghatároz.

Közelítsük ugyanis a $2|x|$ hosszúságú körívet töröttvonalakkal oly módon, ahogy azt a görbék ívhosszáinak értelmezésénél láttuk. Az ilyen töröttvonalak hosszának szuprémuma definíció szerint éppen $2|x|$. Az ilyen töröttvonalak a két egységnyi hosszúságú oldallal együtt mindig olyan konvex sokszöget határoznak meg, amik a deltoidba lettek beleírva. Tehát minden ilyen beírt sokszög kerülete kisebb a deltoid kerületénél, azaz $(2+2|tg x|)$ -nél. Ebből viszont az is következik, hogy ezen beírt sokszögek kerületeinek szuprémuma sem lehet nagyobb, mint $(2+2|tg x|)$.

Tehát azt kaptuk, hogy $2+2|x| \leq 2+2|tg x|$, azaz $|x| \leq |tg x|$. ■

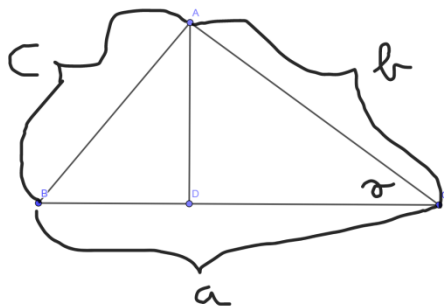
A koszinusztétel

A szögfüggvények felhasználásával könnyebben tudunk általános összefüggéseket keresni egy háromszög oldalai és szögei között. Az egyik alapvető eredmény az ún. koszinusztétel.

Koszinusztétel: Ha egy háromszögben az oldalak hosszai „a”, „b” és „c”, valamint a „c” hosszúságú oldallal szemközi szög mértéke γ , akkor $c^2 = a^2 + b^2 - 2ab \cdot \cos \gamma$.

Bizonyítás:

Húzzuk meg az „a” hosszú oldalhoz tartozó magasságot. Először tegyük fel, hogy ez a szemközi oldalt valamely belső pontjában metszi, legyen ez a belső pont D.



A magasság hossza ekkor $b \cdot \sin \gamma$, a D talppontot a másik (tehát a „b” hosszú oldallal szemközi) csúccsal összekötő szakasz hossza pedig $(a - b \cdot \cos \gamma)$.

A Pithagorasz-tétel szerint $(b \cdot \sin \gamma)^2 + (a - b \cdot \cos \gamma)^2 = c^2$.

A bal oldalon elvégezve a négyzetre emeléseket, kapjuk, hogy

$$\begin{aligned} b^2 \cdot \sin^2 \gamma + a^2 - 2ab \cdot \cos \gamma + b^2 \cdot \cos^2 \gamma &= b^2 \cdot (\sin^2 \gamma + \cos^2 \gamma) + a^2 - 2ab \cdot \cos \gamma = \\ &= b^2 + a^2 - 2ab \cdot \cos \gamma. \end{aligned}$$

Ezzel beláttuk az állítást, ha az „a” hosszú oldalhoz tartozó magasság a szemközi oldalt valamely belső pontjában metszi.

Hasonlóan bizonyítható a tétel állítása, ha ez a talppont valamilyen külső pont.

Ha a magasságtalppont az „a” hosszú oldalnak a B ponton túli meghosszabbításán van, akkor az $(a - b \cdot \cos \gamma)$ mennyiség helyett a $(b \cdot \cos \gamma - a)$ mennyiséget kell négyzetre emelnünk, ami formálisan persze ugyanaz.

Ha a magasságtalppont az „a” hosszú oldalnak a C ponton túli meghosszabbításán van, akkor a γ szög nagyobb, mint derékszög. Ebben az esetben alkalmazhatjuk a $\sin(\pi - \gamma) = \sin \gamma$, illetve $\cos(\pi - \gamma) = -\cos \gamma$ összefüggéseket, és ugyanúgy járhatunk el, mint eredetileg tettük. ■

Vektorok koordinátái

Mivel a vektorok valamiféle elmozdulást fejeznek ki, ezért természetes ötlet, ha megpróbáljuk őket koordinátázni. A térben tekintsük a hagyományos koordináta-rendszert, ami azt jelenti, hogy három tengely van (mondjuk x-, y- és z-tengely), melyek mindegyike átmegy az origón, és ezek a tengelyek egymásra páronként merőlegesek. Az origó mindhárom koordinátája nulla. Bármely térbeli ponthoz egyértelműen tudunk három koordinátát rendelni. Az első koordinátát úgy kapjuk meg, hogy megnézzük, hogy az adott pontnak az x-tengelyre eső vetülete hová esik. Hasonlóan kapható meg a másik két koordináta is.

Tekintsünk most egy adott vektort. Ha öt koordinátákkal szeretnénk leírni, akkor feltehetjük, hogy a kezdőpontja az origóban van, és ezzel a megkötéssel élve már azzal tudjuk jellemezni, hogy a tér melyik pontja az ő végpontja (egyébként az origóból induló vektorokat **helyvektoroknak** szoktuk nevezni).

Az előzőek alapján a tér pontjait, és velük együtt a térbeli vektorok mindegyikét le tudjuk írni három koordináta segítségével.

Vektorok összeadásakor **koordinátánként történik az összeadás** (hiszen így tudjuk leírni, hogy tengelyenként milyen elmozdulás történik, miután a vektorainkat összeadjuk a láncszabály szerint).

A hasonlósági tulajdonságok alkalmazásával az is világos, hogy ha egy **vektort megszorunk egy valós számmal** (azaz vesszük az adott vektornak valamilyen **skalárszorosát**), akkor ez azt jelenti, hogy **meg kell szoroznunk minden koordinátát ugyanazzal a számmal**.

Vektorok skaláris szorzata

A vektorok körében fontos szerepet tölt be az ún. skaláris szorzás. Ezt mint műveletet az alábbi módon definiáljuk:

Definíció (skaláris szorzat): Az $\mathbf{a} = (a_1, a_2, a_3)$ és $\mathbf{b} = (b_1, b_2, b_3)$ térbeli vektorok skaláris szorzata $\mathbf{a}\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$.

Ebből a definícióból könnyen ellenőrizhető, hogy a skaláris szorzás mint művelet kommutatív (azaz mindegy a sorrend), illetve disztributív is a vektorösszeadásra nézve, azaz $\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{a}\mathbf{b} + \mathbf{a}\mathbf{c}$.

A skaláris szorzat a térben leírható ennél szemléletesebben is. Könnyű lesz bebizonyítani az alábbi tételt:

Tétel: Két térbeli vektor skaláris szorzata kiszámolható úgy is, hogy az adott vektorok hosszát (azaz abszolútértéküket) először összeszorozzuk, majd a kapott értéket megszorozzuk a két vektor által bezárt szög koszinuszával.

Bizonyítás:

Legyen a két vektor $\mathbf{a} = (a_1, a_2, a_3)$ és $\mathbf{b} = (b_1, b_2, b_3)$, ezek bezárt szöge legyen ε . A téglatest testátlójának hosszáról szóló eredményünk miatt $|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$, $|\mathbf{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2}$,

$$|\mathbf{a} - \mathbf{b}| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.$$

A koszinusztételt felírva kapjuk, hogy

$$\begin{aligned} & \left(\sqrt{a_1^2 + a_2^2 + a_3^2} \right)^2 + \left(\sqrt{b_1^2 + b_2^2 + b_3^2} \right)^2 - 2 \cdot \sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \sqrt{b_1^2 + b_2^2 + b_3^2} \cdot \cos \varepsilon = \\ & = \left(\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2} \right)^2. \end{aligned}$$

Innen átrendezés után kapjuk, hogy

$$\begin{aligned} & \sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \sqrt{b_1^2 + b_2^2 + b_3^2} \cdot \cos \varepsilon = \\ & = \frac{\left(\sqrt{a_1^2 + a_2^2 + a_3^2} \right)^2 + \left(\sqrt{b_1^2 + b_2^2 + b_3^2} \right)^2 - \left(\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2} \right)^2}{2} = \\ & = \frac{a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - [a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - 2(a_1 b_1 + a_2 b_2 + a_3 b_3)]}{2} = \end{aligned}$$

$= a_1 b_1 + a_2 b_2 + a_3 b_3$, ez utóbbi pedig éppen a skaláris szorzat definíciója. Ezzel az állítást beláttuk. ■

Az addíciós tételek

Ebben a fejezetben megvizsgáljuk, hogyan lehet kifejezni szögek összegének, illetve különbségének szögfüggvényeit. Be fogjuk bizonyítani az alábbi tételt:

Tétel (a szinuszra és a koszinuszra vonatkozó addíciós tételek):

Tetszőleges α és β szögekre teljesülnek az alábbiak:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$$

$$\cos(\beta - \alpha) = \sin \alpha \cdot \sin \beta + \cos \alpha \cdot \cos \beta$$

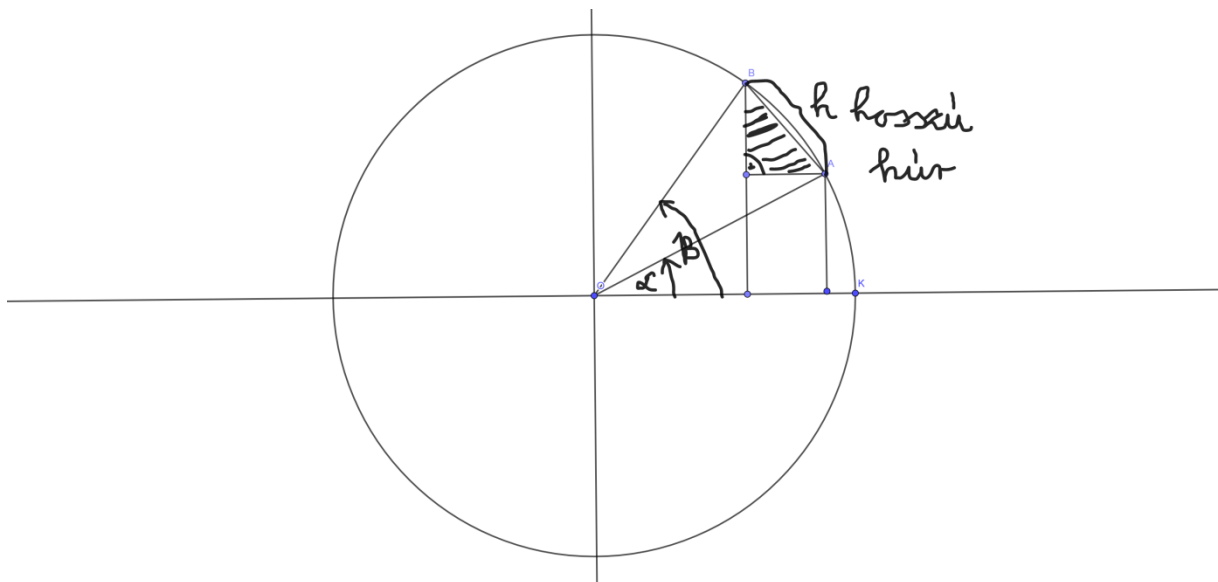
$$\sin(\beta - \alpha) = \sin \beta \cos \alpha - \cos \beta \sin \alpha$$

Bizonyítás:

A szokásos módon tekintsük a koordinátarendszerben az origó középpontú, egységnyi sugarú kört. Az origó legyen az O pont, és legyen K az (1,0) pont. Legyenek α és β valamilyen szögek. Legyen az $\overrightarrow{OK}$ vektornak az O körüli α szöggel való elforgatottja az $\overrightarrow{OA}$, míg az $\overrightarrow{OK}$ vektornak az O körüli β szöggel való elforgatottja az $\overrightarrow{OB}$ vektor.

Az AB szakasz hosszát jelölje h.

Ez a h hosszú szakasz átfogó egy olyan derékszögű háromszögben, melynek befogói $|\sin \alpha - \sin \beta|$, illetve $|\cos \alpha - \cos \beta|$ hosszúságúak.



A Pithagorasz-tétel miatt $(|\sin \alpha - \sin \beta|)^2 + (|\cos \alpha - \cos \beta|)^2 = h^2$.

Ugyanakkor h^2 kifejezhető úgy is, hogy az ABO egyenlő szárú háromszögben felírjuk a koszinusztételt:

$$h^2 = 1^2 + 1^2 - 2 \cdot 1 \cdot 1 \cdot \cos(\beta - \alpha)$$

Az előző két egyenletet összehasonlítva kapjuk, hogy

$$1^2 + 1^2 - 2 \cdot 1 \cdot 1 \cdot \cos(\beta - \alpha) = (|\sin \alpha - \sin \beta|)^2 + (|\cos \alpha - \cos \beta|)^2.$$

Tehát $1^2 + 1^2 - 2 \cdot 1 \cdot 1 \cdot \cos(\beta - \alpha) = (\sin \alpha - \sin \beta)^2 + (\cos \alpha - \cos \beta)^2$, azaz

$$2 - 2 \cos(\beta - \alpha) = \sin^2 \alpha + \sin^2 \beta - 2 \sin \alpha \cdot \sin \beta + \cos^2 \alpha + \cos^2 \beta - 2 \cos \alpha \cdot \cos \beta = \sin^2 \alpha + \sin^2 \beta - 2 \sin \alpha \cdot \sin \beta + \cos^2 \alpha + \cos^2 \beta - 2 \cos \alpha \cdot \cos \beta = 2 - 2 \sin \alpha \cdot \sin \beta - 2 \cos \alpha \cdot \cos \beta.$$

Innen $2 - 2 \cos(\beta - \alpha) = 2 - 2 \sin \alpha \cdot \sin \beta - 2 \cos \alpha \cdot \cos \beta$, ami azzal ekvivalens, hogy

$$\cos(\beta - \alpha) = \cos \alpha \cdot \cos \beta + \sin \alpha \cdot \sin \beta.$$

A szögek összegére vonatkozó koszinuszos addíciós tétel innen már könnyedén levezethető. Ugyanis a szögek különbségére vonatkozó formulát használva kapjuk, hogy

$$\cos(\beta + \alpha) = \cos[\beta - (-\alpha)] = \cos(-\alpha) \cdot \cos \beta + \sin(-\alpha) \cdot \sin \beta = \cos \alpha \cos \beta - \sin \alpha \sin \beta.$$

A szögek összegére vonatkozó szinuszos addíciós tétel megkapható a koszinuszos addíciós tételekből, az alábbi módon:

$$\begin{aligned} \sin(\alpha + \beta) &= \cos\left[\frac{\pi}{2} - (\alpha + \beta)\right] = \cos\left[\left(\frac{\pi}{2} - \alpha\right) - \beta\right] = \cos\left(\frac{\pi}{2} - \alpha\right)\cos(\beta) + \sin\left(\frac{\pi}{2} - \alpha\right)\sin(\beta) = \\ &= \sin \alpha \cos \beta + \cos \alpha \sin \beta. \end{aligned}$$

A különbségre vonatkozó szinuszos addíciós tétel innen előjelváltással könnyen nyerhető:

$$\sin(\beta - \alpha) = \sin[\beta + (-\alpha)] = \sin \beta \cos(-\alpha) + \cos \beta \sin(-\alpha) = \sin \beta \cos \alpha - \cos \beta \sin \alpha. \blacksquare$$

Sorozatok és torlódási pontjaik

Eddig csak kétféle sorozattal találkoztunk: számtani és mértani sorozatokkal. Valójában a sorozat fogalma ennél jóval általánosabb. A **sorozat** olyan függvény, aminek értelmezési tartománya a természetes számok halmaza.

Ez azt jelenti, hogy a sorozat egyes elemeit egyesével végig tudjuk indexelni a természetes számokkal, de arra vonatkozóan, hogy mik legyenek a sorozat elemei, igazából nincsen semmiféle szigorú megkötés.

A mi praxisunkban többnyire olyan sorozatok fognak előfordulni, amik számokból (esetleg vektorokból) állnak.

A sorozatok esetén szintén fontos lesz számunkra a **torlódási pont** fogalma, amit formálisan ugyanúgy vezetünk be, mint azt a számhalmazoknál tettük:

Definíció: Egy számsorozatnak **torlódási pontja** a t valós szám, ha bármely $d > 0$ távolság esetén a sorozatnak végtelen sok olyan eleme létezik, melyeknek a t -től vett távolsága kisebb d -nél.

A torlódási pont fogalma –bár sorozatok esetében formálisan ugyanaz, mint a halmazoknál –egy kicsit eltérő a sorozatok esetében. Ugyanis halmazoknál alapvetően úgy tekintjük, hogy minden elem

egyszer van felsorolva, viszont sorozatoknál egy elem akárhányszor szerepelhet. Ezáltal pl. egy konstans sorozatnak is van torlódási pontja (és ennek csak egyféle eleme van), ugyanakkor egy egyelemű halmaznak nincsen torlódási pontja.

Természetesen a korlátosság, infimum és szuprémum ugyanúgy értelmezendő sorozatoknál, mint ahogy azt halmazok esetében láttuk. A monotonitás pedig ugyanúgy értelmezhető, mint a függvényeknél (hiszen a sorozatok a függvények speciális fajtái).

Ahogy a számhalmazoknál tettük, számsorozatok esetén ugyanúgy (felezéssel) bizonyítható a Bolzano-Weierstrass-tétel:

Tétel (Bolzano-Weierstrass-tétel számsorozatokra): Bármely korlátos, valós számsorozatnak létezik torlódási pontja.

Konvergens és divergens sorozatok, sorozatok határértéke

A torlódási pont segítségével kényelmesen definiálhatjuk sorozatoknak egy nagyon fontos osztályát, az ún. konvergens sorozatokat.

Definíció (konvergencia, határérték): Egy valós számsorozat **konvergens**, ha korlátos, és pontosan egy torlódási pontja van. Adott konvergens sorozat egyetlen torlódási pontját a szóban forgó **sorozat határértékének** nevezzük.

Azokat a sorozatokat, melyek nem konvergensek, **divergensnek** szoktuk nevezni.

A határérték fogalmát az eddigiek során már több ízben használtuk, csak még sosem neveztük meg, hogy igazából erről van szó. Gondoljunk vissza például az valós számoknak a racionális számokkal való közelítéséről szóló eredményünkre. Ezt másképpen úgy is megfogalmazhattuk volna, hogy **bármely valós számot elő lehet állítani valamilyen konvergens, racionális elemű sorozat határértékeként.**

Ha jól meggondoljuk, akkor a hatványozásnál és gyökvonásnál, a téglalap területképletének bizonyításakor, illetve a körív hossza és a középponti szög egymáshoz való viszonyának vizsgálata során is a háttérben a határérték fogalmát használtuk.

Sőt az $\sqrt[n]{n}$ alakú számok vizsgálatakor azt láttuk be, hogy az $x_n = \sqrt[n]{n}$ (ahol $n=2, 3, 4, \dots$) előírással definiált sorozat konvergens, és határértéke 1-gyel egyenlő.

A konvergencia egy ekvivalens értelmezése

A számsorozatok konvergenciájának egy másik definíciója az alábbi:

Definíció (konvergenca, határérték): Az $x_1, x_2, x_3, \dots$ valós számsorozat konvergens, és határértéke az A valós szám, ha tetszőleges pozitív d távolsághoz található valamilyen k pozitív egész (küszöb) index oly módon, hogy bármely k -nál nagyobb n index esetén az x_n elemnek az A -tól vett távolsága kisebb d -nél.

Szó sincs arról, hogy ez a definíció egy másfajta határértéket írna le. Mindjárt tisztázni fogjuk, hogy a konvergenciának ez a definíciója ekvivalens a régebbi, torlódási pont segítségével bevezetett definícióval. Csupán kényelmi okokból vezetjük be ezt a másik értelmezést, ugyanis valamikor egyszerűbb lesz ezt használnunk.

Az alábbi tételből világosan ki fog derülni, hogy a konvergenciájának a két eddig ismertett definíciója egymással ekvivalens:

Tétel: Egy $x_1, x_2, x_3, \dots$ valós számsorozatra nézve az alábbi kijelentések ekvivalensek:

- (i) létezik olyan A valós szám, melynek bármely pozitív sugarú környezetén kívülre a sorozatnak csak véges sok eleme esik
- (ii) a sorozat korlátos és pontosan egy torlódási pontja van
- (iii) létezik olyan A valós szám, amelyre igaz, hogy tetszőleges pozitív d távolsághoz található valamilyen k pozitív egész index oly módon, hogy bármely k -nál nagyobb n index esetén az x_n elemnek az A -tól vett távolsága kisebb d -nél.

Bizonyítás:

A fenti három kijelentés ekvivalenciáját kell bizonyítanunk. Ezt úgy fogjuk megtenni, hogy először belátjuk (i) és (ii) ekvivalenciáját, majd utána (i) és (iii) ekvivalenciáját bizonyítjuk. Ebből az ekvivalenciarelációk tranzitivitási tulajdonsága miatt adódik, hogy a három kijelentés egymással ekvivalens.

Az (i) és (ii) ekvivalenciájának bizonyítása:

Bebizonyítjuk, hogy az (i) és az (ii) kijelentések ekvivalensek egymással.

Ehhez első lépésként azt látjuk be, hogy minden olyan sorozat, ami rendelkezik az (i)-ben leírt tulajdonsággal, az rendelkezik az (ii)-beli tulajdonsággal is. Tegyük fel ugyanis, hogy a sorozatunkhoz létezik egy olyan A valós szám, melynek bármely pozitív sugarú környezetén kívülre a sorozatnak csak véges sok eleme esik. Legyen d tetszőleges pozitív távolság. A feltétel szerint csak véges sok olyan sorozatbeli elem van, ami $(A-d)$ -nél kisebb, vagy $(A+d)$ -nél nagyobb, ezért a sorozatunk egészen biztosan korlátos. Az A pont biztosan torlódási pontja a sorozatnak, hiszen ennek bármely pozitív sugarú környezetén belülre a sorozatnak végtelen sok eleme esik. Elég tehát annyit megmutatnunk, hogy A az egyetlen torlódási pont. Tegyük fel indirekt, hogy létezik a sorozatnak egy A -tól különböző t torlódási pontja is. Legyen az A -nak és a t -nek a távolsága D . Ekkor – lévén, hogy t torlódási pont – a t -nek a $D/3$ sugarú környezetébe a sorozatnak végtelen sok eleme esik. Ugyanakkor az itteni végtelen sok elem mind kívül esik az A -nak a $D/3$ sugarú környezetén, ez pedig ellentmondás. Tehát az (i) valóban maga után vonja az (ii)-t.

Most azt bizonyítjuk, hogy minden olyan sorozat, ami rendelkezik az (ii)-beli tulajdonsággal, egyben rendelkezik az (i)-beli tulajdonsággal is. Tegyük fel tehát most, hogy a sorozatunk korlátos, és pontosan egy torlódási pontja van. Jelölje ezt az egyetlen torlódási pontot A . Bebizonyítjuk, hogy az A -nak bármely pozitív sugarú környezetén kívül a sorozatnak csak véges sok eleme esik. Ezt indirekt bizonyítjuk. Tegyük fel indirekt, hogy mégis létezik olyan d távolság, amelyre teljesül, hogy a sorozatnak az A -nak a d sugarú környezetén kívülre végtelen sok eleme esik a sorozatnak. Vegyük most figyelembe, hogy a sorozatunk korlátos. A Bolzano-Weierstrass-tételből következik, hogy a sorozat azon elemeinek, amik az A -nak a d sugarú környezetén kívülre esnek, szintén létezik egy torlódási pontja, és világos, hogy ez a torlódási pont A -tól különböző. Ezzel ellentmondásra jutottunk az indirekt feltevésből, és így igazoltuk, hogy az (ii) maga után vonja (i)-t.

Az (i) és (iii) ekvivalenciájának bizonyítása:

Most azt látjuk be, hogy (i) és (iii) kijelentések is ekvivalensek egymással.

Először belátjuk, hogy minden olyan sorozat, ami rendelkezik az (i)-ben leírt tulajdonsággal, az rendelkezik az (iii)-beli tulajdonsággal is. Tegyük fel, hogy valamilyen A számra teljesül, hogy az A -nak bármely pozitív sugarú környezetén kívülre a sorozatnak csak véges sok eleme esik.

Legyen d tetszőleges pozitív távolság. A feltétel szerint a sorozatnak csak véges sok olyan eleme van, ami nem $(A-d)$ és $(A+d)$ közé esik. Mivel csak véges sok ilyen elem van a sorozatban, ezért létezik olyan k pozitív egész index, ami minden ilyen típusú sorozatbeli elem indexénél nagyobb. Ezért ha n olyan index, ami nagyobb k -nál, akkor az x_n elem biztosan az A -nak a d sugarú környezetébe esik.

Még azt bizonyítjuk, hogy minden olyan sorozat, ami rendelkezik az (iii)-ben leírt tulajdonsággal, az rendelkezik az (i)-beli tulajdonsággal is. Tegyük fel ugyanis, hogy valamilyen A valós szám mellett a sorozatunk teljesíti az (iii)-beli feltételt. Legyen d tetszőleges pozitív távolság. Ehhez a feltétel szerint létezik olyan k index, hogy minden k -nál nagyobb n index esetén az x_n elem az A -nak a d sugarú környezetébe esik. Ez tehát azt jelenti, hogy az adott d mellett legfeljebb k darab (tehát véges sok) olyan elem létezhet, ami nem esik bele az A -nak a d sugarú környezetébe. Mivel d tetszőlegesen választható pozitív érték, ezért ezzel beláttuk, hogy (iii) maga után vonja (i)-t. ■

Monoton, korlátos sorozatok

Az alsóhatár- és a felsőhatár-axiómából könnyen következik az alábbi tétel:

Tétel: *Bármely monoton, korlátos sorozat konvergens.*

Bizonyítás:

A bizonyítást monoton növekedő, korlátos sorozatokra írjuk le, monoton csökkenő sorozatok esetén a bizonyítás ugyanígy működik.

Legyenek a sorozatunk elemei $x_0 \leq x_1 \leq x_2 \leq x_3 \leq \dots$

A korlátossága miatt ennek a sorozatnak van felső korlátja, így a felsőhatár-axióma miatt van szuprémuma, legyen ez S .

Először belátjuk, hogy S a sorozatnak torlódási pontja. Legyen d rögzített pozitív távolság. Mivel S a sorozat elemeinek szuprémuma, ezért van olyan eleme a sorozatnak, ami nagyobb $(S-d)$ -nél. A sorozat monoton növekedése miatt az összes ennél nagyobb indexű elem is nagyobb $(S-d)$ -nél, valamint az is világos, hogy egyikük sem lehet nagyobb S -nél, mivel S a szuprémum. Mivel ez tetszőleges pozitív d -ről elmondható, ezért S valóban torlódási pontja a sorozatnak.

Még annyit kell belátnunk, hogy S a sorozatnak az egyetlen torlódási pontja. Tegyük fel indirekt, hogy van a sorozatnak egy másik torlódási pontja is, jelölje ezt t . Mivel a sorozat elemei közül egyik sem lehet S -nél nagyobb, ezért $t < S$. Ugyanakkor a sorozatnak biztosan létezik olyan eleme, ami t -nél nagyobb, legyen egy ilyen mondjuk az x_k . A monoton növekedés miatt a sorozatnak legfeljebb k darab, azaz véges sok olyan eleme létezik, ami nem nagyobb t -nél, a többi elem pedig mind nagyobb t -nél.

Legyen $\frac{x_k - t}{2} = D$, és tekintsük a t pontnak a D sugarú környezetét. Az előbb elmondottak alapján ebbe a környezetbe a sorozatnak csak véges sok eleme eshet, ami ellentmond annak, hogy a t pontot mint torlódási pontot értelmeztük. Tehát S az egyetlen torlódási pont, és ezzel a tétel állítását beláttuk. ■

Az Euler-szám bevezetése

Adjunk hozzá az 1-hez egy nagyon kicsi pozitív értéket, és a kapott számot emeljük fel akkora hatványra, amekkora az 1-hez hozzáadott kicsi pozitív szám reciproka. Ezáltal egy 1-hez közeli, 1-nél nagyobb értéket emelünk fel egy nagyon nagy hatványra. Ennek a hatványnak a viselkedését fogjuk vizsgálni.

Az előzőeket formálisan felírva az lesz a feladatunk, hogy vizsgáljuk az $e_n = \left(1 + \frac{1}{n}\right)^n$ sorozatot.

Be fogjuk bizonyítani, hogy ez a sorozat konvergens. Ahhoz, hogy ezt megtehessük, bevezetjük az

$E_n = \left(1 + \frac{1}{n}\right)^{n+1}$ sorozatot is.

Bebizonyítjuk az alábbi tételt:

Tétel:

(i) Az $e_n = \left(1 + \frac{1}{n}\right)^n$ sorozat szigorúan monoton növekvő.

(ii) Az $E_n = \left(1 + \frac{1}{n}\right)^{n+1}$ sorozat szigorúan monoton csökkenő.

(iii) Pontosan egy olyan valós szám létezik, ami az (e_n) sorozat minden eleménél nagyobb, és az (E_n) sorozat minden eleménél kisebb.

Bizonyítás:

Az (i) bizonyítása:

Legyen n rögzített pozitív egész szám. Vegyünk $(n+1)$ darab számot úgy, hogy ezek közül n darab egyenlő $\left(1 + \frac{1}{n}\right)$ -nel, az $(n+1)$ -edik pedig 1-gyel egyenlő. Ezekre a számokra alkalmazzuk a számtani és mértani közepek közt fennálló egyenlőtlenséget. Figyeljük meg, hogy itt most nem állhat fent egyenlőség, hiszen nem az összes szám egyenlő egymással. Az egyenlőtlenséget alkalmazva tehát azt kapjuk, hogy

$$\sqrt[n+1]{1 \cdot \left(1 + \frac{1}{n}\right)^n} < \frac{1+n \cdot \left(1 + \frac{1}{n}\right)}{n+1} = \frac{1+n+1}{n+1} = \frac{n+1+1}{n+1} = 1 + \frac{1}{n+1}.$$

$$\text{Tehát } \sqrt[n+1]{\left(1 + \frac{1}{n}\right)^n} < 1 + \frac{1}{n+1}.$$

Az egyenlőtlenség mindkét oldala pozitív, ezért mindkét oldalt $(n+1)$ -edik hatványra emelhetjük. Így kapjuk, hogy

$$\left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n+1}\right)^{n+1}.$$

Ez pontosan azt jelenti, hogy $e_n < e_{n+1}$. Mivel n tetszőleges pozitív egész számot jelölt, ezért ezzel beláttuk az (i) állítás helyességét.

Az (ii) bizonyítása:

Legyen $n > 1$ rögzített pozitív egész szám. Az (i) bizonyításához hasonlóan ismét a számtani és mértani közepekre vonatkozó egyenlőtlenséget fogjuk alkalmazni $(n+1)$ darab számra. Ezen $(n+1)$ darab szám közül n darabnak az értéke legyen egyenlő $\left(1 - \frac{1}{n}\right)$ -nel, az $(n+1)$ -edik pedig legyen egyenlő 1-gyel.

Az egyenlőtlenségben (ahogy az (i)-nél) nem teljesülhet egyenlőség. Felírva a számtani és mértani közepekre vonatkozó egyenlőtlenséget, az alábbiakat kapjuk:

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

$$\text{A bal oldalon álló mennyiség } \sqrt[n+1]{1 \cdot \left(1 - \frac{1}{n}\right)^n} = \sqrt[n+1]{\left(1 - \frac{1}{n}\right)^n} = \sqrt[n+1]{\left(\frac{n-1}{n}\right)^n} =$$

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

Tehát tudjuk, hogy $\sqrt[n+1]{\left(\frac{1}{1+\frac{1}{n-1}}\right)^n} < \frac{1}{1+\frac{1}{n}}.$

Az egyenlőtlenség mindkét oldala pozitív, ezért mindkét oldalt $(n+1)$ -edik hatványra emelhetjük. Így kapjuk, hogy

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

A bal oldal úgy is írható, hogy $\frac{1}{\left(1+\frac{1}{n-1}\right)^n} = \frac{1}{E_{n-1}}$, a jobb oldal pedig $\frac{1}{\left(1+\frac{1}{n}\right)^{n+1}} = \frac{1}{E_n}.$

Tehát $\frac{1}{E_{n-1}} < \frac{1}{E_n}.$

Mivel a nevezőkben álló mennyiségek pozitívak, ezért a velük való szorzás után az egyenlőtlenséggel ugyanolyan irányban fog állni. Mindkét oldalt szorozva $E_{n-1} \cdot E_n$ -nel kapjuk, hogy

$$E_n < E_{n-1}.$$

Mivel ez minden $n > 1$ indexre fennáll, ezért az (ii) állítást ezzel beláttuk.

Az (iii) bizonyítása:

Vizsgáljuk meg az $(E_n - e_n)$ különbségeket.

$$E_n - e_n = \left(1 + \frac{1}{n}\right)^{n+1} - \left(1 + \frac{1}{n}\right)^n = \left(1 + \frac{1}{n}\right)^n \cdot \left[\left(1 + \frac{1}{n}\right) - 1\right] = \left(1 + \frac{1}{n}\right)^n \cdot \frac{1}{n}.$$

Ebből azonnal leolvasható, hogy minden n indexre $e_n < E_n$. Az (i)-ben és (ii)-ben belátott állítások miatt ebből következik, hogy minden n indexre $e_n < e_{n+1} < E_{n+1} < E_n$

Ez egyszerűbben megfogalmazva azt jelenti, hogy az $[e_n; E_n]$ zárt intervallumok az indexelés sorrendjében egymást tartalmazzák (itt $n=1, 2, 3, 4, \dots$). A Cantor-axióma miatt ezeknek a zárt intervallumoknak biztosan van közös pontja. Még annyit kell belátnunk, hogy ezeknek az intervallumoknak egyetlen közös pontjuk van. Ehhez elég ismét az $(E_n - e_n)$ különbségeket vizsgálnunk, méghozzá abból a szempontból, hogy ezek milyen nagyok lehetnek.

$$0 < E_n - e_n = \left(1 + \frac{1}{n}\right)^n \cdot \frac{1}{n} < E_1 \cdot \frac{1}{n} = \left(1 + \frac{1}{1}\right)^2 \cdot \frac{1}{n} = \frac{4}{n}.$$

Mivel n növekedtével a $\frac{4}{n}$ pozitív érték tetszőlegesen kicsi lesz (másképp fogalmazva a neki megfelelő sorozat határértéke a nulla), ezért biztosak lehetünk benne, hogy az $[e_n; E_n]$ zárt intervallumoknak csupán egy közös pontja van. ■

A fent bizonyított tételből következik, hogy az $e_n = \left(1 + \frac{1}{n}\right)^n$ sorozat konvergens, és a határértéke az az egyetlen pozitív valós szám, amelyet az $[e_n; E_n]$ zárt intervallumok mindegyike tartalmaz.

Definíció: Az $e_n = \left(1 + \frac{1}{n}\right)^n$ sorozat határértékét **Euler-számnak** nevezzük, és inntől kezdve e -vel fogjuk jelölni.

Természetesen az $E_n = \left(1 + \frac{1}{n}\right)^{n+1}$ sorozat szintén konvergens, és a határértéke ugyancsak az Euler-szám.

Az Euler-szám numerikus értéke $e = 2,7182.....$

Végtelenhez tartás

Mielőtt továbbmegyünk, beszélnünk kell a végtelenhez tartás fogalmáról.

Definíció (végtelenhez tartás): Egy valós számsorozat **tart a plusz végtelenbe**, ha bármely c valós szám esetén teljesül, hogy a sorozatnak csak véges sok eleme nem nagyobb c -nél.

Egy valós számsorozat **tart a mínusz végtelenbe**, ha bármely c valós szám esetén teljesül, hogy a sorozatnak csak véges sok eleme nem kisebb c -nél.

Láthatjuk, hogy a végtelenhez tartó sorozatok egyike sem konvergens, viszont a végtelenhez tartás is egy, a konvergenciájához hasonló viselkedést takar.

Az nyilvánvaló, hogy egy végtelenbe tartó sorozat nem lehet korlátos. Azonban ha egy sorozat felülről vagy alulról nem korlátos, még egyáltalán nem biztos, hogy a (plusz vagy mínusz) végtelenbe tartana. Tekintsük ugyanis azt a sorozatot, aminek elemeit az alábbi módon képezzük:

1, 2, 1, 3, 1, 4, 1, 5, 1, 6, 1, 7, 1, 8, 1, 9, 1, 10, 1, 11,

Ez a sorozat felülről nem korlátos, de nem tart a plusz végtelenbe.

Megállapodás szerint úgy tekintjük, hogy a plusz végtelenbe tartó sorozatoknak a plusz végtelen torlódási pontja, a mínusz végtelenbe tartó sorozatoknak pedig mínusz végtelen torlódási pontja.

Hasonlóan a felülről nem korlátos számhalmazoknak torlódási pontja a plusz végtelen, és az alulról nem korlátos számhalmazoknak torlódási pontja a mínusz végtelen.

A függvényhatárérték fogalma

A konvergens sorozatok határértékének fogalma általánosítható függvényekre nézve. Ehhez ki kell választanunk a függvény értelmezési tartományának valamely torlódási pontját, és meg kell néznünk, hogy ha valamilyen értelmezési tartományból származó sorozattal konvergálunk ehhez a torlódási ponthoz, akkor hogyan viselkednek a sorozat elemei a függvénybe helyettesítve.

A definíció egészen pontosan az alábbi lesz:

Definíció (függvényhatárérték): Legyen az u pont torlódási pontja az f függvény értelmezési tartományának. Azt mondjuk, hogy az f függvény határértéke az u pontban az A valós szám, ha tetszőleges olyan (x_n) sorozatra, melynek elemei f értelmezési tartományából valók, és (x_n) sorozat konvergál u -hoz, akkor $(f(x_n))$ sorozat is konvergens, és határértéke A -val egyenlő.

Vegyük észre, hogy az u pont nem feltétlenül van benne a függvény értelmezési tartományában.

Nagyon fontos megértenünk, hogy a definíció tetszőleges u -hoz tartó (x_n) sorozatra vonatkozik, aminek elemeit be tudjuk helyettesíteni az f függvénybe, és ha a behelyettesítés után kapott sorozat mindig ugyanahhoz az A értékhez konvergál, akkor beszélünk arról, hogy az f függvénynek az u pontbeli határértéke az A . Ezt úgy szokás jelölni, hogy $\lim_{x \rightarrow u} f(x) = A$.

A definíció természetes módon kiterjeszthető azokra az esetekre is, amikor u vagy A értéke nem véges (tehát ezek közül bármelyik lehet akár plusz, vagy mínusz végtelen is). Ez teljesen összhangban van azzal a konszenzuson alapuló ténnyel, hogy a plusz (illetve mínusz) végtelen torlódási pontja a felülről (illetve alulról) nem korlátos számhalmazoknak.

A fentiekben megfogalmazottak alapján az is világos, hogy a **sorozathatárérték fogalma egy speciális függvényhatárértéket takar**. A sorozatnak, amely egy függvény, ami a természetes számok halmazán van értelmezve, ilyenkor keressük az esetleges határértékét a plusz végtelenben.

A határérték és a műveletek

Felmerül a kérdés, mi történik a határértékekkel akkor, ha konvergens sorozatok között végzünk műveleteket. Hasonlóan az is kérdés, hogy mi történik akkor, ha olyan függvények között végzünk alpműveleteket, melyeknek ugyanabban a pontban van határértékük.

Először is sorozatokra bebizonyítjuk az alábbi tételt:

Tétel:

- (i) Konvergens sorozatok összege konvergens, és az összegként kapott sorozat határértéke ugyanannyi, mint a két összeadott sorozat határértékeinek összege.
- (ii) Konvergens sorozatok szorzata konvergens, és a szoratként kapott sorozat határértéke ugyanannyi, mint a két összeszorozott sorozat határértékeinek szorzata.

(iii) Konvergens sorozatok hányadosa konvergens, amennyiben a nevezőbe kerülő sorozat nem tart nullához. Ekkor a hányadosként kapott sorozat határértéke ugyanannyi, mint a kiinduló sorozataink határértékeinek hányadosa.

Bizonyítás:

Először teszünk néhány alapvető észrevételt.

Legyen (a_n) olyan konvergens valós számsorozat, aminek határértéke A , és (b_n) olyan konvergens valós számsorozat, aminek határértéke B . Legyen $d > 0$ tetszőleges távolság. A határértékre vonatkozó ekvivalens állításokat megfogalmazó tételből tudjuk, hogy az (a_n) sorozatnak csak véges sok eleme nincsen benne az A szám d sugarú környezetében, illetve a (b_n) sorozatnak szintén csak véges sok eleme nincsen benne a B szám d sugarú környezetében. Ez azt jelenti, hogy valamilyen k pozitív egészre fennáll, hogy minden k -nál nagyobb m index esetén az a_m elem az A -nak, a b_m elem pedig a B -nek a d sugarú környezetébe esik, azaz ezekre az m indexekre $A-d < a_m < A+d$; illetve $B-d < b_m < B+d$.

Ezek ekvivalens módon úgy is felírhatók, hogy $|(a_m - A)| < d$, illetve $|(b_m - B)| < d$.

Ezen észrevételeink után készen állunk a tételbeli állítások bizonyítására.

Az (i) bizonyítása:

A fenti k -nál nagyobb m indexekre (felhasználva azt, hogy egyenlőtlenségeket összeadhatunk) kapjuk, hogy

$$(A - d) + (B - d) < a_m + b_m < (A + d) + (B + d).$$

Ez azzal ekvivalens, hogy $-2d < (a_m + b_m) - (A + B) < 2d$, azaz

$|(a_m + b_m) - (A + B)| < 2d$, azaz $(a_m + b_m)$ -nek és $(A + B)$ -nek az egymástól való távolsága kisebb, mint $2d$.

Összefoglalva: bármely pozitív d -hez létezik olyan k index, melynél nagyobb m -ek mindegyikére teljesül, hogy $(a_m + b_m)$ -nek és $(A + B)$ -nek az egymástól való távolsága kisebb, mint $2d$. Ez viszont azt jelenti, hogy (i) igaz.

Az (ii) bizonyítása:

Tekintsük az előzetes észrevételben bevezetett k -nál nagyobb m indexeket.

Ahogy az (i) bizonyításában az $(a_m + b_m)$ -nek és $(A + B)$ -nek az egymástól való távolságát vizsgáltuk, úgy most $a_m \cdot b_m$ -nek és $A \cdot B$ -nek az egymástól való távolságát kell vizsgálni.

Ez a távolság másképpen $|a_m \cdot b_m - A \cdot B|$.

Ahhoz, hogy ezt a távolságot hatékonyan tudjuk becsülni, egy átalakítást végzünk rajta. Vegyük észre, hogy $|a_m \cdot b_m - A \cdot B| = |a_m \cdot b_m - a_m \cdot B + a_m \cdot B - A \cdot B|$.

Ezt az átalakítást azért végeztük el, mert így már ki tudunk emelni:

$$|a_m \cdot b_m - a_m \cdot B + a_m \cdot B - A \cdot B| = |a_m \cdot (b_m - B) + B \cdot (a_m - A)|.$$

Az $|x + y| \leq |x| + |y|$ háromszög-egyenlőtlenség alapján kapjuk, hogy

$$|a_m \cdot (b_m - B) + B \cdot (a_m - A)| \leq |a_m \cdot (b_m - B)| + |B \cdot (a_m - A)|.$$

Itt a jobb oldalt átalakíthatjuk úgy, hogy

$$|a_m \cdot (b_m - B)| + |B \cdot (a_m - A)| = |a_m| \cdot |(b_m - B)| + |B| \cdot |(a_m - A)|.$$

Ezekre érvényesek az alábbi egyenlőtlenségek:

$$|a_m| \cdot |(b_m - B)| + |B| \cdot |(a_m - A)| < |a_m| \cdot d + |b_m| \cdot d = (|a_m| + |b_m|) \cdot d.$$

Az (a_n) és (b_n) sorozatok konvergenciája miatt mindkét sorozat korlátos, ezért abszolútértékeik összege is korlátos. Így az $(|a_m| + |b_m|) \cdot d$ mennyiség, ami nagyobb az $a_m \cdot b_m$ -nek és $A \cdot B$ -nek az egymástól való távolságánál, tetszőlegesen kicsi lehet, ha d értékét elég kicsinek választjuk.

Ezzel az (ii)-beli állítást beláttuk.

Az (iii) bizonyítása:

Legyen a nevezőbe kerülő sorozat (b_n) , amiről feltesszük, hogy nem tart nullához, azaz eddigi jelöléseink szerint $B \neq 0$. Mivel nem tart nullához, ezért legfeljebb véges sok eleme lehet nullával egyenlő, ezért formálisan az $\left(\frac{a_n}{b_n}\right)$ sorozatnak csak véges sok eleme nem értelmezhető, így van létjogosultsága, hogy ezen sorozatnak a konvergenciáját vizsgáljuk.

Legyen d olyan pozitív távolság, ami $\left|\frac{B}{2}\right|$ -nél kisebb. Tudjuk, hogy a (b_n) sorozatnak csak véges sok eleme nem esik a B -nek a d sugarú környezetén kívül, így létezik olyan k index, hogy minden k -nál nagyobb m index esetén b_m -nak a B -től vett távolsága kisebb d -nél (és így kisebb $\left|\frac{B}{2}\right|$ -nél is). Rögzítsünk egy ilyen m indexet. Az ilyen m indexekre tehát teljesül, hogy $|b_m - B| < d$, illetve az is, hogy $|b_m| > \left|\frac{B}{2}\right|$. Az utóbbi azért igaz, mert a d távolságot úgy választottuk, hogy kisebb legyen $\left|\frac{B}{2}\right|$ -nél.

Először vizsgáljuk speciálisan az $\left(\frac{1}{b_n}\right)$ sorozat konvergenciáját. Azt kellene bizonyítanunk, hogy ez a sorozat konvergens, és határértéke $\frac{1}{B}$.

Ehhez elég a már fent bevezett m mellett az $\left|\frac{1}{b_m} - \frac{1}{B}\right|$ -et, azaz $\frac{1}{b_m}$ és $\frac{1}{B}$ távolságát vizsgálnunk.

$$\left|\frac{1}{b_m} - \frac{1}{B}\right| = \left|\frac{b_m - B}{b_m \cdot B}\right| = |b_m - B| \cdot \left|\frac{1}{b_m \cdot B}\right| = |b_m - B| \cdot \left|\frac{1}{b_m}\right| \cdot \left|\frac{1}{B}\right|.$$

Erre teljesülnek az alábbi egyenlőtlenségek:

$$|b_m - B| \cdot \left|\frac{1}{b_m}\right| \cdot \left|\frac{1}{B}\right| < d \cdot \left|\frac{2}{B}\right| \cdot \left|\frac{1}{B}\right|.$$

Világos, hogy ha a d pozitív értéket elég kicsinek választjuk, akkor $d \cdot \left| \frac{2}{B} \right| \cdot \left| \frac{1}{B} \right|$ pozitív mennyiség is tetszőlegesen kicsi lehet.

Ezzel beláttuk, hogy az $\left(\frac{1}{b_n} \right)$ sorozat konvergens, és határértéke $\frac{1}{B}$.

Az (ii)-beli tulajdonság miatt az $\left(\frac{a_n}{b_n} \right)$ sorozat is konvergens, és határértéke $\frac{A}{B}$, ugyanis

$\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$. Ezzel az (iii)-beli állítást is bebizonyítottuk. ■

Általánosan a függvényhatárértékekre is bebizonyítható a fenti tétel megfelelője, a bizonyítás során szó szerint ugyanúgy lehet eljárni, ahogy azt az előzőekben tettük, ezért ennek a tételnek a bizonyítását itt nem részletezzük.

Tétel (A függvényhatárérték és a műveletek) :

Legyenek az f és g olyan valós függvények, amik ugyanott vannak értelmezve, továbbá legyen u torlódási pontja a közös értelmezési tartományuknak. Amennyiben valamilyen A , illetve B valós számokra teljesül, hogy $\lim_{x \rightarrow u} f(x) = A$, $\lim_{x \rightarrow u} g(x) = B$, akkor:

$$(i) \lim_{x \rightarrow u} (f(x) + g(x)) = A + B$$

$$(ii) \lim_{x \rightarrow u} f(x) \cdot g(x) = AB$$

$$(iii) \text{feltéve, hogy } B \neq 0, \text{ teljesül, hogy } \lim_{x \rightarrow u} \frac{f(x)}{g(x)} = \frac{A}{B}.$$

A hatványozás és gyökvonás műveleti tulajdonságai

A határérték műveleti tulajdonságainak segítségével már képesek leszünk arra, hogy igazoljuk a hatványozásnak és gyökvonásnak a középiskolában megismert (de ott nem bizonyított) műveleti tulajdonságait. Természetesen beszélhetünk csak szigorúan csak a hatványozás tulajdonságairól is, hiszen minden gyökvonás egyfajta hatványozást jelent, ugyanis tetszőleges 1-nél nagyobb n egész esetén egy nemnegatív szám n -edik gyöke ugyanaz, mint az adott nemnegatív számnak az $\frac{1}{n}$ -edik hatványa. Az utóbbiról már felezéses eljárással bizonyítottuk, hogy minden esetben létezik, ráadásul értéke egyértelműen meghatározott.

Mielőtt rátérnénk a hatványozás műveleti tulajdonságainak általános vizsgálatára, a könnyebb tárgyalás kedvéért először azokat csak racionális kitevőkre bizonyítjuk, egy segédttételben összefoglalva.

Segédttétel: *Ha A és B pozitív valós számok, illetve x és y tetszőleges racionális számok, akkor a következők teljesülnek:*

$$(i) (AB)^x = A^x \cdot B^x$$

$$(ii) A^x \cdot A^{-x} = 1$$

$$(iii) (A^x)^y = A^{xy}$$

$$(iv) A^{x+y} = A^x \cdot A^y$$

Bizonyítás:

Mivel x és y is racionális, ezért külön-külön felírhatók két egész szám hányadosaként. Legyen $x = \frac{r}{s}$, illetve $y = \frac{t}{v}$. Nyilván az is feltehető, hogy s és v egyaránt pozitív egészek (hiszen ha valamelyik nem az, akkor az adott törtben egyszerre megcserélhetjük a számláló és a nevező előjelét).

Az (i) bizonyítása:

$(AB)^x = (AB)^{\frac{r}{s}}$, ennek s -edik hatványa $(AB)^r$.

$$A^x \cdot B^x = A^{\frac{r}{s}} \cdot B^{\frac{r}{s}}, \text{ ennek } s\text{-edik hatványa } \left(A^{\frac{r}{s}}\right)^s \cdot \left(B^{\frac{r}{s}}\right)^s = A^r \cdot B^r.$$

Az, hogy egész r -ekre $(AB)^r = A^r \cdot B^r$ teljesül, közvetlenül következik az egész kitevőkre vonatkozó definícióból. Tehát $(AB)^{\frac{r}{s}}$ -nek az s -edik hatványa megegyezik $A^{\frac{r}{s}} \cdot B^{\frac{r}{s}}$ -nek az s -edik hatványával. Az s -edik gyökre vonatkozó egyértelműségi tétel miatt ebből következik, hogy $(AB)^{\frac{r}{s}} = A^{\frac{r}{s}} \cdot B^{\frac{r}{s}}$.

Az (ii) bizonyítása:

$$\text{Az } A^x \cdot A^{-x} = A^{\frac{r}{s}} \cdot A^{-\frac{r}{s}} \text{ szám } s\text{-edik hatványa } \left(A^{\frac{r}{s}}\right)^s \cdot \left(A^{-\frac{r}{s}}\right)^s = A^r \cdot A^{-r}.$$

Az egész kitevőjű hatványok definíciójából adódik, hogy $A^r \cdot A^{-r} = 1$, ezért (használva az s -edik gyökre vonatkozó egyértelműségi tételt) kapjuk, hogy $A^{\frac{r}{s}} \cdot A^{-\frac{r}{s}} = 1$.

Az (iii) bizonyítása:

$$A^{xy} = A^{\frac{r}{s} \cdot \frac{t}{v}} = A^{\frac{rt}{sv}}, \text{ ennek az } sv\text{-edik hatványa } A^{rt}.$$

$$(A^x)^y = \left(A^{\frac{r}{s}}\right)^{\frac{t}{v}}, \text{ ennek az } sv\text{-edik hatványa az } \left[\left(A^{\frac{r}{s}}\right)^{\frac{t}{v}}\right]^v \text{ számnak az } s\text{-edik hatványa.}$$

$$\left[\left(A^{\frac{r}{s}}\right)^{\frac{t}{v}}\right]^v = \left(A^{\frac{r}{s}}\right)^t, \text{ ennek az } s\text{-edik hatványa pedig } \left[\left(A^{\frac{r}{s}}\right)^s\right]^t = (A^r)^t = A^{rt}.$$

Ugyanúgy, mint az előzőekben, az állítás helyessége a gyökvonásra vonatkozó egyértelműségi tételből következik.

Az (iv) bizonyítása:

$A^x \cdot A^y = A^{\frac{r}{s} + \frac{t}{v}}$, ennek az sv -edik hatványa $\left[\left(A^{\frac{r}{s}}\right)^s\right]^v \cdot \left[\left(A^{\frac{t}{v}}\right)^v\right]^s = A^{rv} \cdot A^{ts} = A^{rv+ts}$.

$A^{x+y} = A^{\frac{r}{s} + \frac{t}{v}} = A^{\frac{rv+ts}{sv}}$, ennek az sv -edik hatványa A^{rv+ts} .

Ugyanúgy, mint az előző pontokban, az állítás helyessége a gyökvonásra vonatkozó egyértelműségi tételből következik. ■

Tétel: Legyen A pozitív valós szám, valamint legyen (r_n) olyan racionális tagú sorozat, ami konvergens, és határértéke a c valós szám. Ekkor az (A^{r_n}) sorozat is konvergens, és határértéke A^c .

Bizonyítás:

Azt az esetet tárgyaljuk, amikor $A > 1$. A $0 < A < 1$ eset ugyanúgy tárgyalható.

Legyen tehát $A > 1$. Emlékezzünk vissza, hogy tetszőleges w valós szám esetén ekkor A^w úgy volt értelmezve, mint az A szám olyan racionális kitevőjű hatványainak szupréma, melyekben a kitevő w -nél nem nagyobb racionális szám.

Legyen s tetszőleges pozitív egész szám, amelyre teljesül, hogy $\frac{1}{s} < |c|$, valamint $s > A$.

A racionális számokkal való közelítésnél látottak szerint létezik olyan i egész szám, amelyre teljesül, hogy $\frac{i}{s} \leq c < \frac{i+1}{s}$.

Mivel az (r_n) racionális tagú sorozat konvergál a c számhoz, ezért biztosan létezik olyan k küszöbindex, hogy a k -nál nagyobb m indexek bármelyikére teljesül, hogy $\frac{i-1}{s} < r_m < \frac{i+1}{s}$.

Mivel az egynél nagyobb alapú exponenciális függvények szigorúan monoton növekvőek, ezért

a fenti m indexekre teljesül, hogy A^c és A^{r_n} értéke egyaránt $A^{\frac{i-1}{s}}$ és $A^{\frac{i+1}{s}}$ közé esik.

Ezért A^c és A^{r_n} egymástól való távolsága kisebb, mint $A^{\frac{i}{s}}$ és $A^{\frac{i+1}{s}}$ egymástól való távolsága, azaz

$$|A^c - A^{r_n}| < \left| A^{\frac{i-1}{s}} - A^{\frac{i+1}{s}} \right|.$$

Elég tehát ennek az egyenlőtlenségnek a jobb oldalát vizsgálnunk. Használva a racionális kitevőkre bizonyított $A^{x+y} = A^x \cdot A^y$ tulajdonságot, kapjuk, hogy

$$\left| A^{\frac{i-1}{s}} - A^{\frac{i+1}{s}} \right| = \left| A^{\frac{i-1}{s}} - A^{\frac{i-1}{s}} \cdot A^{\frac{2}{s}} \right| = \left| A^{\frac{i-1}{s}} \cdot \left(1 - A^{\frac{2}{s}} \right) \right| = \left| A^{\frac{i-1}{s}} \right| \cdot \left| 1 - A^{\frac{2}{s}} \right| = A^{\frac{i-1}{s}} \cdot \left(A^{\frac{2}{s}} - 1 \right).$$

Ismét felhasználva a szigorú monotonitást, illetve az $s > A$ feltételt, kapjuk, hogy

$$A^{\frac{i-1}{s}} \cdot \left(A^{\frac{2}{s}} - 1 \right) < A^c \cdot \left(s^{\frac{2}{s}} - 1 \right).$$

Mivel $s^{\frac{2}{s}} = \left(s^{\frac{1}{s}}\right)^2$ sorozat határértéke az s szám végtelenbe tartásával az 1, ezért s növekedtével az $|A^c - A^{r_n}|$ tetszőlegesen kicsi lehet, ami bizonyítja a tétel helyességét az $A > 1$ esetben.

Ha $0 < A < 1$, akkor szó szerint ugyanígy járhatunk el, csak szuprérum helyett infimumról kell beszélnünk, illetve a megfelelő exponenciális függvény ebben az esetben szigorúan monoton csökkenő. ■

Az imént belátott tételt használva a hatványoknak egy újabb értelmezését nyertük, amit a határérték fogalmára támaszkodva lehet megfogalmazni.

Most már be tudjuk bizonyítani a hatványozás általános műveleti tulajdonságait. Ezeket egy tételben szedjük össze.

Tétel: Ha A és B pozitív valós számok, illetve x és y tetszőleges valós számok, akkor a következők teljesülnek:

$$(i) \quad (AB)^x = A^x \cdot B^x$$

$$(ii) \quad A^x \cdot A^{-x} = 1$$

$$(iii) \quad (A^x)^y = A^{xy}$$

$$(iv) \quad A^{x+y} = A^x \cdot A^y$$

Bizonyítás:

Az állítások mindegyikét beláttuk már akkor, amikor x és y is racionális.

Legyenek tehát x és y tetszőleges valós számok. Legyen (x_n) egy olyan racionális tagú sorozat, ami x -hez tart, illetve (y_n) egy olyan racionális tagú sorozat, ami y -hoz tart.

Az (i) bizonyítása:

Tudjuk, hogy $(AB)^{x_n} = A^{x_n} \cdot B^{x_n}$ teljesül minden n indexre. Az előző tétel miatt $(AB)^{x_n}$ konvergál $(AB)^x$ -hez, A^{x_n} konvergál A^x -hez, míg B^{x_n} konvergál B^x -hez. A konvergens sorozatok szorzatáról szóló tétel miatt az $(A^{x_n} \cdot B^{x_n})$ sorozat konvergál $A^x \cdot B^x$ -hez.

Az (ii) bizonyítása:

Tudjuk, hogy minden n -re $A^{x_n} \cdot A^{-x_n} = 1$, azaz $A^{x_n} = \frac{1}{A^{-x_n}}$. Ebből következik, hogy a bal oldalon álló

sorozat határértéke ugyanaz, mint a jobb oldalon álló sorozat határértéke. A bal oldali sorozat határértéke A^x , a jobb oldalié pedig (használva az előző tételt, illetve a konvergens sorozatok hányadosáról szóló tételt) egyenlő $\frac{1}{A^{-x}}$ -nel. Innen az állítás adódik.

Az (iii) bizonyítása:

Elegendő az állítást abban az esetben bizonyítanunk, amikor x és y is pozitív. Ha ugyanis valamelyikük negatív, akkor áttérhetünk pozitív kitevőre a (ii)-ben bizonyított azonosság segítségével.

Tehát legyen x és y is pozitív, valamint tegyük fel, hogy $A > 1$.

Legyen s pozitív egész szám, amelynek reciproka kisebb x -nél és y -nél is. A valós számok racionális számokkal való közelítésénél látottaknak megfelelően léteznek olyan i és j pozitív egész számok, amelyekre teljesül, hogy $\frac{i}{s} \leq x < \frac{i+1}{s}$, illetve $\frac{j}{s} \leq y < \frac{j+1}{s}$.

Használva a szigorú monotonitást, megállapítható, hogy

$$\left(A^{\frac{i}{s}}\right)^{\frac{j}{s}} \leq (A^x)^y < \left(A^{\frac{i+1}{s}}\right)^{\frac{j+1}{s}}, \text{ valamint } A^{\frac{i}{s} \cdot \frac{j}{s}} \leq A^{xy} < A^{\frac{i+1}{s} \cdot \frac{j+1}{s}}.$$

Azt már tudjuk (hiszen racionális kitevőkre beláttuk az állítást), hogy

$$\left(A^{\frac{i}{s}}\right)^{\frac{j}{s}} = A^{\frac{i}{s} \cdot \frac{j}{s}}, \text{ illetve } \left(A^{\frac{i+1}{s}}\right)^{\frac{j+1}{s}} = A^{\frac{i+1}{s} \cdot \frac{j+1}{s}}$$

Tehát összességében megállapítható, hogy $(A^x)^y$ és A^{xy} értéke is $A^{\frac{i}{s} \cdot \frac{j}{s}}$ és $A^{\frac{i+1}{s} \cdot \frac{j+1}{s}}$ közé esik.

$$\text{Ebből következik, hogy } |(A^x)^y - A^{xy}| \leq \left| A^{\frac{i+1}{s} \cdot \frac{j+1}{s}} - A^{\frac{i}{s} \cdot \frac{j}{s}} \right|.$$

A jobb oldalon álló mennyiség (felhasználva azt is, hogy $A > 1$), úgy alakítható át, hogy

$$\left| A^{\frac{i+1}{s} \cdot \frac{j+1}{s}} - A^{\frac{i}{s} \cdot \frac{j}{s}} \right| = \left| A^{\frac{i+j+1}{s^2}} - A^{\frac{ij}{s^2}} \right| = \left| A^{\frac{ij}{s^2}} \cdot \left| A^{\frac{i+j+1}{s^2}} - 1 \right| \right| = A^{\frac{ij}{s^2}} \cdot \left(A^{\frac{i+j+1}{s^2}} - 1 \right) = A^{\frac{i}{s} \cdot \frac{j}{s}} \cdot \left[\left(A^{\frac{i+j+1}{s}} \right)^{\frac{1}{s}} - 1 \right],$$

erre pedig teljesül az az egyenlőtlenség, hogy

$$A^{\frac{i}{s} \cdot \frac{j}{s}} \cdot \left[\left(A^{\frac{i+j+1}{s}} \right)^{\frac{1}{s}} - 1 \right] < A^{x \cdot y} \cdot \left[\left(A^{\frac{i+1}{s} + \frac{j+1}{s}} \right)^{\frac{1}{s}} - 1 \right] < A^{x \cdot y} \cdot \left[(A^{2x+2y})^{\frac{1}{s}} - 1 \right].$$

Ha s tart a végtelenbe, akkor $(A^{2x+2y})^{\frac{1}{s}}$ határértéke 1-gyel egyenlő, ebből következően ebben az esetben az $A^{x \cdot y} \cdot \left[(A^{2x+2y})^{\frac{1}{s}} - 1 \right]$ a nullához tart. Ezek alapján csak az lehetséges, hogy

$$|(A^x)^y - A^{xy}| = 0, \text{ ami ekvivalens azzal, amit bizonyítanunk kellett.}$$

Ha $0 < A < 1$, akkor a bizonyítás szó szerint ugyanígy vihető végig.

Az (iv) bizonyítása:

Tudjuk, hogy minden n -re $A^{x_n+y_n} = A^{x_n} \cdot A^{y_n}$. A bal oldalon álló sorozat határértéke (az előző tételt használva) A^{x+y} , a jobb oldali szorzatban szereplőké pedig A^x , illetve A^y . Innen az állítás adódik. ■

Rendőrelv

Egyszerű, de gyakran jól használható kritériumot ad meg az ún. rendőrelv bizonyos sorozatok konvergenciájára nézve. A kapcsolódó tétel gyakorlatilag azt mondja ki, hogy ha két sorozat ugyanoda tart, és ezek közrefognak egy harmadik sorozatot (mint két rendőr a gyanúsítottat), akkor a közrefogott sorozatnak is ugyanoda kell konvergálnia. Természetesen ugyanez függvényekre is igaz lesz.

Tétel (rendőrelv)

(i) Ha (a_n) és (c_n) konvergens sorozatok határértéke ugyanaz a h valós szám, illetve ha (b_n) olyan sorozat, hogy $a_n \leq b_n \leq c_n$ minden n -re, akkor a (b_n) sorozat szintén konvergens, és határértéke h .

(ii) Ha $f(x)$, $g(x)$ és $h(x)$ ugyanott értelmezett valós függvények,

$$\text{melyekre fennáll hogy } \lim_{x \rightarrow u} f(x) = \lim_{x \rightarrow u} h(x) = A,$$

valamint tudjuk, hogy $f(x) \leq g(x) \leq h(x)$ teljesül minden megfelelő x -re, akkor $\lim_{x \rightarrow u} g(x) = A$.

Bizonyítás: Elég az (i)-t bizonyítani, ugyanis az (ii) automatikusan következik belőle a függvényhatárérték sorozatokkal megfogalmazott definíciója miatt. Legyen $d > 0$. Mivel (a_n) és (c_n) sorozat is tart a h számhoz, ezért az (a_n) és a (c_n) sorozatnak egyaránt csak véges sok eleme esik a h szám d sugarú környezetén kívülre. Ez azt jelenti, hogy létezik olyan k küszöbindex, hogy minden $m \geq k$ esetén az a_m és c_m elem is a h számnak a d sugarú környezetébe esik. Így az ilyen m indexekre teljesül, hogy b_m szintén a h -nak a d sugarú környezetébe esik, hiszen a feltételek szerint $a_n \leq b_n \leq c_n$ teljesül minden n indexre. Ezzel beláttuk, hogy a (b_n) sorozat konvergens, és határértéke h . ■

A rendőrelv egy alkalmazása

Szükségünk lesz a későbbiekben az alábbi alapvető tételre, ami lényegében arról szól, hogy amennyiben egy hatványalap egy pozitív számhoz, míg a kitevő valamilyen valós számhoz tart, akkor az egész hatvány ahhoz a számhoz tart, amit az előzőek értelemszerű behelyettesítésével kapunk.

Tétel: Legyen (a_n) olyan pozitív tagú, valós számsorozat, melynek határértéke az A szám. Legyen továbbá (b_n) olyan valós számsorozat, melynek határértéke a B szám. Ekkor az $(a_n^{b_n})$ sorozat konvergens, és határértéke A^B .

Bizonyítás:

Az A szám biztosan nem negatív, hiszen egy pozitív elemekből álló sorozatnak a határértéke.

Vizsgáljuk először azt az esetet, amikor A pozitív.

Alakítsuk át $a_n^{b_n}$ -et az alábbi módon:

$$a_n^{b_n} = [A + (a_n - A)]^{B + (b_n - B)}$$

Ezt elosztva a konstans A^B -nel, majd alkalmazva a hatványozásnak már bizonyított műveleti tulajdonságait, kapjuk, hogy

$$\begin{aligned} \frac{a_n^{b_n}}{A^B} &= \frac{[A + (a_n - A)]^{B + (b_n - B)}}{A^B} = \frac{[A + (a_n - A)]^B}{A^B} \cdot [A + (a_n - A)]^{(b_n - B)} = \\ &= \left[\frac{A + (a_n - A)}{A} \right]^B \cdot [A + (a_n - A)]^{(b_n - B)} = \left[1 + \frac{(a_n - A)}{A} \right]^B \cdot [a_n]^{(b_n - B)} \end{aligned}$$

A most kapott értéket szeretnénk valahogy közrefogni, hogy alkalmazhassuk a rendőrelvet.

Az biztos, hogy $-|a_n - A| \leq a_n - A \leq |a_n - A|$, így (figyelembe véve a hatványok rendezéséről tanultakat is) kapjuk, hogy

$$\begin{aligned} \left[1 - \frac{|a_n - A|}{A} \right]^B \cdot [a_n]^{(b_n - B)} &\leq \left[1 + \frac{(a_n - A)}{A} \right]^B \cdot [a_n]^{(b_n - B)} \leq \\ &\leq \left[1 + \frac{|a_n - A|}{A} \right]^B \cdot [a_n]^{(b_n - B)} \end{aligned}$$

Legyen m az az egész szám, amelyre teljesül, hogy $m \leq B < m+1$.

Vegyük észre, hogy minden n indexre teljesül, hogy $\left[1 - \frac{|a_n - A|}{A} \right]^{m+1} \leq \left[1 - \frac{|a_n - A|}{A} \right]^B \leq \left[1 - \frac{|a_n - A|}{A} \right]^m$.

A bal szélén álló $\left(\left[1 - \frac{|a_n - A|}{A} \right]^{m+1} \right)$ sorozat konvergens, és határértéke 1, hiszen ez $|m+1|$ darab 1-hez konvergáló sorozatnak a szorzata.

Ugyanígy látható, hogy az $\left(\left[1 - \frac{|a_n - A|}{A} \right]^{m+1} \right)$ sorozat konvergens, és határértéke 1.

A rendőrelv miatt tehát a középen álló $\left(\left[1 - \frac{|a_n - A|}{A} \right]^B \right)$ sorozat határértéke is 1.

Ugyanígy belátható, hogy az $\left(\left[1 + \frac{|a_n - A|}{A} \right]^B \right)$ sorozat szintén 1-hez tart.

Menjünk most tovább, és vizsgáljuk az $([a_n]^{(b_n-B)})$ sorozatot. Itt az alap A-hoz, a kitevő nullához tart. Legyen N pozitív egész. Elég nagy n indexekre ez a hatvány $\left(\frac{1}{N}\right)^{\frac{1}{N}}$ és $N^{\frac{1}{N}}$ közé esik. Mivel az N pozitív egész növekedtével az $\left(\frac{1}{N}\right)^{\frac{1}{N}}$, illetve az $N^{\frac{1}{N}}$ is 1-hez tart, ezért a rendőrelv miatt az

$([a_n]^{(b_n-B)})$ sorozat is 1-hez tart.

Az előzőekből a rendőrelv segítségével következik, hogy az

$\left(\left[1 + \frac{(a_n-A)}{A}\right]^B \cdot [a_n]^{(b_n-B)}\right) = \left(\frac{a_n^{b_n}}{A^B}\right)$ sorozat szintén 1-hez tart, hiszen ezt két 1-hez tartó

sorozat közé sikerült beszorítanunk.

Ezért az $(a_n^{b_n})$ sorozat határértéke valóban A^B -nel egyenlő, amennyiben az A pozitív.

Ha $A = 0$, akkor szintén alkalmazhatjuk a rendőrelvet. Legyen N valamilyen pozitív egész szám, illetve legyen m az az egész szám, amelyre teljesül, hogy $m \leq B < m+1$.

Ekkor elég nagy n indexekre teljesül, hogy $0 \leq a_n^{b_n} < \left(\frac{1}{N}\right)^m$. Innen a rendőrelv miatt az $(a_n^{b_n})$ sorozat határértéke nulla (ami az $A=0$ esetben persze éppen A^B). ■

Nevezetes határértékek

Az eddigieket felhasználva már nagyon sokféle típusú határértéket meg tudunk határozni. A nevezetesebbeket egy tételben gyűjtjük össze.

Tétel:

$$(i) \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

$$(ii) \lim_{x \rightarrow +\infty} \left(x^{\frac{1}{x}}\right) = 1$$

$$(iii) \lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$$

$$(iv) \lim_{x \rightarrow +\infty} \left(1 + \frac{c}{x}\right)^x = e^c, \text{ ahol } c \text{ valamilyen valós konstans jelöl}$$

Bizonyítás:

Az (i) bizonyítása:

Legyen $|x| < \frac{\pi}{2}$. Már korábban beláttuk, hogy ha $|x| < \frac{\pi}{2}$, akkor $|\sin x| \leq |x|$.

Ha osztunk $|\sin x|$ -szel, akkor nyerjük azt az egyenlőtlenséget, hogy

$$1 \leq \left| \frac{x}{\sin x} \right|.$$

Szintén a $|\sin x| \leq |x|$ egyenlőtlenségből kiindulva, négyzetre emeléssel adódik, hogy a megfelelő x -ekre $\sin^2 x \leq x^2$ teljesül, amiből kapjuk, hogy

$$1 - x^2 \leq 1 - \sin^2 x = \cos^2 x.$$

Ez utóbbi egyenlőtlenség azzal ekvivalens, hogy $\frac{1}{\cos^2 x} \leq \frac{1}{1 - x^2}$. Innen gyökvonással adódik, hogy

$$\frac{1}{|\cos x|} \leq \frac{1}{\sqrt{1 - x^2}}.$$

Korábban már azt is beláttuk, hogy amennyiben $|x| < \frac{\pi}{2}$, akkor $|x| \leq |tg x|$. Ebből $|\sin x|$ -szel történő osztással kapjuk, hogy $\left| \frac{x}{\sin x} \right| \leq \frac{1}{|\cos x|}$.

Az eddigieket összevetve az alábbi egyenlőtlenség-lánc írható fel, amennyiben $|x| < \frac{\pi}{2}$:

$$1 \leq \left| \frac{x}{\sin x} \right| \leq \frac{1}{|\cos x|} \leq \frac{1}{\sqrt{1 - x^2}}.$$

Nyilvánvalóan $\lim_{x \rightarrow 0} 1 = 1$, illetve $\lim_{x \rightarrow 0} \frac{1}{\sqrt{1 - x^2}} = 1$, ezért a rendőrelv miatt $\lim_{x \rightarrow 0} \left| \frac{x}{\sin x} \right| = 1$.

Mivel $|x| < \frac{\pi}{2}$ esetén az x -nek és a $\sin x$ -nek ugyanaz az előjele, ezért ilyenkor $\left| \frac{x}{\sin x} \right| = \frac{x}{\sin x}$.

Azaz azt is beláttuk, hogy $\lim_{x \rightarrow 0} \frac{x}{\sin x} = 1$, ezért ha vesszük az adott függvény reciprokát, a határérték

szintén 1 lesz, és nekünk ezt kellett bizonyítanunk.

Az (ii) bizonyítása:

Ebben az esetben az x változó a plusz végtelenbe tart. Először tegyük fel, hogy valamilyen n pozitív egészre teljesül, hogy $n \leq x < n+1$.

Ebben az esetben az alábbi egyenlőtlenség-lánc lesz érvényben:

$$n^{\frac{1}{n+1}} \leq x^{\frac{1}{x}} \leq (n+1)^{\frac{1}{n}}.$$

$$n^{\frac{1}{n+1}} = \left(n^{\frac{1}{n}} \right)^{\frac{n}{n+1}}, \text{ ahol az alap } 1\text{-hez, és a kitevő is } 1\text{-hez tart, ezért } \lim_{n \rightarrow +\infty} \left(n^{\frac{1}{n+1}} \right) = 1.$$

$$\text{Hasonlóan látható be, hogy } \lim_{n \rightarrow +\infty} \left((n+1)^{\frac{1}{n}} \right) = 1$$

A rendőrelv alkalmazásával ebből éppen a (ii)-beli állítást nyerjük.

Az (iii) bizonyítása:

Itt az x változó a nullához tart. Tegyük fel először, hogy x értéke pozitív, és kisebb 1-nél. Legyen n az a természetes szám, amelyre fennáll, hogy $\frac{1}{n+1} \leq x < \frac{1}{n}$. Ekkor természetesen $n < \frac{1}{x} \leq n+1$, ezért az alábbi egyenlőtlenség-lánc lesz érvényben:

$$\left(1 + \frac{1}{n+1}\right)^n \leq (1+x)^{\frac{1}{x}} \leq \left(1 + \frac{1}{n}\right)^{n+1}.$$

Felhasználva, hogy $\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^{n+1} = \lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n+1}\right)^n = e$, a rendőrelv alapján kapjuk,

hogy ha a pozitív x változó nullához tart, akkor $(1+x)^{\frac{1}{x}}$ tart az e -hez.

Ez a tulajdonság hasonlóan bizonyítható, amennyiben x negatív, és úgy tart nullához.

Mivel negatív és pozitív x -ekre is teljesül a kívánt tulajdonság, ezért a (iii)-at is beláttuk.

Az (iv) bizonyítása:

Ezt visszavezetjük a (iii)-ra. Ha x tart a plusz végtelenbe, akkor $\frac{c}{x}$ tart a nullához. A (iii)-beli tulajdonság miatt innen adódik, hogy

$$\lim_{x \rightarrow +\infty} \left(1 + \frac{c}{x}\right)^{\frac{x}{c}} = e.$$

Ebből c -edik hatványra emeléssel nyerhető a (iv)-es állítás. ■

Függvények folytonossága

A függvényhatárérték segítségével könnyedén megfogalmazhatjuk a pontbeli folytonosság értelmezését.

Definíció (Függvény pontbeli folytonossága): Legyen u az $f(x)$ valós függvény értelmezési tartományának egy olyan pontja, ami egyben torlódási pontja is a függvény értelmezési tartományának. Amennyiben minden olyan u -hoz tartó (x_n) sorozatra, melynek elemei szintén az $f(x)$ függvény értelmezési tartományából valók, az is teljesül, hogy az $(f(x_n))$ sorozat konvergens, és határértéke az $f(u)$ szám, akkor az $f(x)$ függvényt az u pontban folytonosnak nevezzük.

Tekintsünk először egy konkrét példát, mondjuk a koszinuszfüggvényt az $u=0$ pontban. Ha a koszinuszfüggvényről szeretnénk ellenőrizni, hogy folytonos-e az $u=0$ pontban, akkor le kell ellenőriznünk, hogy tetszőleges, nullához tartó (x_n) sorozat esetén igaz-e, hogy a $(\cos(x_n))$ sorozat tart $\cos 0$ -hoz, vagyis 1-hez. Burkoltan ezt a tulajdonságot már ellenőriztük, amikor vizsgáltuk azt,

hogy van-e határértéke a $\frac{\sin x}{x}$ -nek a 0-ban. Az ott bizonyított $1 \leq \frac{1}{|\cos x|} \leq \frac{1}{\sqrt{1-x^2}}$ egyenlőtlenség-láncból leolvasható, hogy a koszinuszfüggvény folytonos az $x=0$ helyen.

Most bevezetjük általánosan a folytonosság fogalmát függvényekre nézve.

Definíció (folytonosság): Egy valós függvényt folytonosnak nevezünk, ha értelmezési tartományának bármely pontjában folytonos.

A pontbeli folytonosságot a függvényhatárérték segítségével vezettük be. Emiatt (felhasználva a függvényhatárértéknek a műveletekre vonatkozó tulajdonságait) érvényes az alábbi tétel:

Tétel (folytonos függvények és a műveletek): Ugyanott értelmezett folytonos függvények összege, szorzata, különbsége, illetve hányadosa is folytonos (értelmszerűen a hányadost csak azon pontokban értelmezzük, ahol a nevezőbeli behelyettesített érték nem nulla).

Hangsúlyozzuk, hogy ezt a tételt már nem kell külön bizonyítanunk, hiszen következik a függvényhatárértékekről korábban belátott tulajdonságokból.

Ugyanígy a hatványozás kapcsán belátott, határértékekre vonatkozó eredményeinkből adódik az alábbi tétel:

Tétel: Ha az $f(x)$ pozitív értékű valós függvény, illetve a $g(x)$ valós függvény mindketten folytonosak, akkor az $(f(x))^{g(x)}$ függvény is folytonos.

Immár készen állunk arra, hogy függvényeknek egy nagy csoportjáról belássuk, hogy ezek folytonosak.

Tétel: A polinomok, a hatványfüggvények, az exponenciális függvények, valamint a trigonometrikus függvények mind folytonosak.

Bizonyítás:

Világos, hogy a hatványfüggvények és az exponenciális függvények folytonosak, hiszen ezek olyan hatványként írhatók fel, melyeknek alapja és kitevője is egy-egy folytonos függvény.

Az egyváltozós polinomok általában úgy írhatók fel, hogy $p(x) = c_n \cdot x^n + c_{n-1} \cdot x^{n-1} + \dots + c_1 \cdot x + c_0$. Itt a folytonos függvények szorzatára vonatkozó tétel miatt mindegyik $c_k \cdot x^k$ alakú összeadandó tag folytonos, ezért (felhasználva a folytonos függvények összegéről szóló eredményt) maga a polinom is biztosan folytonos.

Hátravan még annak igazolása, hogy a trigonometrikus függvények is mind folytonosak.

Először a szinuszfüggvény folytonosságát bizonyítjuk. Legyen u tetszőleges valós szám, és (x_n) egy olyan valós számsorozat, ami u -hoz tart. Azt kellene belátnunk, hogy a $(\sin x_n)$ sorozat tart a $(\sin u)$ -hoz. Tekintsük $(\sin x_n)$ -nek a $(\sin u)$ -tól vett távolságát, majd az addíciós tételek alapján végezzük el az alábbi átalakításokat:

$$\begin{aligned} |\sin x_n - \sin u| &= \left| \sin\left(\frac{x_n + u}{2} + \frac{x_n - u}{2}\right) - \sin\left(\frac{x_n + u}{2} - \frac{x_n - u}{2}\right) \right| = \\ &= \left| \sin\left(\frac{x_n + u}{2}\right) \cdot \cos\left(\frac{x_n - u}{2}\right) + \cos\left(\frac{x_n + u}{2}\right) \cdot \sin\left(\frac{x_n - u}{2}\right) \right. \\ &\quad \left. - \left[\sin\left(\frac{x_n + u}{2}\right) \cos\left(\frac{x_n - u}{2}\right) - \cos\left(\frac{x_n + u}{2}\right) \cdot \sin\left(\frac{x_n - u}{2}\right) \right] \right| = \\ &= \left| 2 \cos\left(\frac{x_n + u}{2}\right) \cdot \sin\left(\frac{x_n - u}{2}\right) \right| \end{aligned}$$

Triviálisan teljesül minden n indexre, hogy $\left| 2 \cos\left(\frac{x_n + u}{2}\right) \right| \leq 2$, továbbá elég nagy n indexekre fennáll, hogy $\left| \frac{x_n - u}{2} \right| < \frac{\pi}{2}$, ezért az ismert szinuszos egyenlőtlenségnek megfelelően $\left| \sin\left(\frac{x_n - u}{2}\right) \right| \leq \left| \frac{x_n - u}{2} \right|$.

Az előzőek miatt elég nagy n indexek esetén $\left| 2 \cos\left(\frac{x_n + u}{2}\right) \cdot \sin\left(\frac{x_n - u}{2}\right) \right| \leq 2 \cdot \left| \frac{x_n - u}{2} \right|$. Innen a rendőrelv alapján következik, hogy a $(\sin x_n)$ sorozat tart a $(\sin u)$ -hoz.

A koszinuszfüggvény folytonossága ugyanígy látható be.

A tangens- és kotangensfüggvény folytonossága következik abból, hogy ezek felírhatók két folytonos függvény hányadosaként ($\operatorname{tg} x = \frac{\sin x}{\cos x}$, $\operatorname{ctg} x = \frac{\cos x}{\sin x}$). ■

Mielőtt továbbmennénk a folytonos függvények vizsgálatával, szükségünk lesz az összetett függvény, illetve az inverzfüggvény fogalmára.

Összetett függvények, függvények inverze

Ha egy függvénybe behelyettesítünk egy másikat, akkor ún. **összetett függvényt** kapunk. Ilyenkor számít a sorrend, azaz általában nem mindegy, hogy melyik a külső, és melyik a belső függvény. Természetesen az $f(x)$ külső, illetve a $g(x)$ belső függvényekből létrehozott $f(g(x))$ összetett függvény akkor lesz értelmezhető, ha $g(x)$ értékkészlete részhalmaza az $f(x)$ függvény értelmezési tartományának. Ha az $f(g(x))$ függvény létezik, akkor ezt az $f(x)$ és $g(x)$ függvények **kompozíciójának** is szokás nevezni.

Nézzünk egy példát. Legyen $f(x) = x^2 + x$, $g(x) = x^3$. Ekkor $f(g(x)) = (x^3)^2 + x^3$, ugyanakkor

$g(f(x)) = (x^2 + x)^3$. Már ebből a példából is világos, hogy az összetett függvények képzése során a sorrend számít.

Az f és g függvények kompozícióját (ebben a sorrendben véve) kényelmi okokból valamikor úgy is szoktuk jelölni, hogy $f \circ g$. Tehát ezen jelöléssel $(f \circ g)(x) = f(g(x))$ minden megfelelő x -re.

A függvénykompozíció képzését egyfajta műveletnek is tekinthetjük, amit függvények között tudunk elvégezni. Fontos kiemelnünk, hogy ez egy asszociatív művelet, amit a következő segéd-tételben bizonyítunk be:

Állítás: *A függvénykompozíció képzése mint művelet asszociatív.*

Bizonyítás:

Legyenek f , g és h olyan függvények, melyekre értelmezve vannak az $(f \circ g)$ és $(g \circ h)$ függvénykompozíciók. Azt kell belátnunk, hogy az $((f \circ g) \circ h)$, illetve az $(f \circ (g \circ h))$ függvények egybeesnek. Legyenek x_1, x_2, x_3, x_4 olyan elemek, hogy $f(x_3) = x_4$, $g(x_2) = x_3$, valamint $h(x_1) = x_2$ teljesüljön.

Ekkor $g(h(x_1)) = g(x_2) = x_3$, ezért $(f \circ (g \circ h))(x_1) = f(x_3) = x_4$.

Hasonlóan $f(g(x_2)) = f(x_3) = x_4$, ezért $((f \circ g) \circ h)(x_1) = f(g(x_2)) = f(x_3) = x_4$.

Tehát az $((f \circ g) \circ h)$, illetve az $(f \circ (g \circ h))$ függvények egybeesnek, és ezt kellett belátnunk. ■

Speciális szerepe van az identitásfüggvénynek, ha összetett függvényekről van szó. Ugyanis ilyenkor az identitásfüggvény teljesen úgy viselkedik, mint a valós számok szorzásakor az 1 (akár külső, akár belső függvénynek választjuk az identitásfüggvényt, az a másik függvényt nem fogja „eltorzítani”).

Ez alapján természetesen merül fel a kérdés, hogy adott függvényhez mint belső függvényhez mikor találhatunk olyan másik (külső) függvényt, hogy ezek kompozíciója az identitásfüggvény legyen. Így jutunk el az inverzfüggvény fogalmához.

Definíció (inverzfüggvény): *A valós $g(x)$ függvény inverzfüggvénye az az $f(x)$ valós függvény, aminek értelmezési tartománya megegyezik $g(x)$ értékkészletével, továbbá teljesül rá, hogy $f(g(x)) = x$ minden olyan x -re, ahol g értelmezett.*

Az invertálható $g(x)$ függvény inverzét általában $g^{-1}(x)$ -szel jelöljük (itt nem hatványozásról van tehát szó!).

Tegyük itt néhány fontos észrevételt. Azt állítjuk, hogy csak olyan függvényeknek létezik inverze, amik minden felvett értéküket csak egyetlenegyszer veszik fel (ezek az ún. **injektív** függvények).

Sőt ennél sokkal többet is bizonyítunk az alábbi tételben:

Tétel: *Egy invertálható $g(x)$ függvényre az alábbiak teljesülnek:*

- (i) *a $g(x)$ függvény injektív*
- (ii) *a $g^{-1}(x)$ függvény injektív*

(iii) $g^{-1}(x)$ függvény invertálható, és inverze az eredeti $g(x)$ függvény

(iv) ha $g(x)$ szigorúan monoton, akkor $g^{-1}(x)$ függvény is az.

Bizonyítás:

Az (i) bizonyítása:

Tegyük fel indirekt, hogy valamely g függvény invertálható, viszont nem injektív, azaz valamilyen $A \neq B$ számokra $g(A) = g(B)$ teljesül. Ekkor az inverz definíciója miatt

$g^{-1}(g(A)) = A$, $g^{-1}(g(B)) = B$, viszont $g(A) = g(B)$ miatt $g^{-1}(g(A)) = g^{-1}(g(B))$. Ez ellentmond annak, hogy $A \neq B$.

Az (ii) bizonyítása:

Tegyük fel indirekt, hogy valamilyen $g(x)$ függvényre és valamilyen $C \neq D$ számokra $g^{-1}(C) = g^{-1}(D)$ teljesül. Ekkor biztosan vannak olyan c és d számok, melyekre $g(c) = C$, illetve $g(d) = D$ teljesül, hiszen

az inverzfüggvény definíciójában szerepelt az a feltétel, hogy az értelmezési tartománya egyezzen meg az eredeti függvény értékkészletével.

Ekkor tehát $g^{-1}(C) = g^{-1}(g(c)) = c$, valamint $g^{-1}(D) = g^{-1}(g(d)) = d$. Innen a $g^{-1}(C) = g^{-1}(D)$ feltétel miatt kapjuk, hogy $c = d$, ezért $C = g(c) = g(d) = D$, ami ellentmond annak, hogy $C \neq D$.

Az (iii) bizonyítása:

Legyen u egy eleme a $g(x)$ függvény értelmezési tartományának. Ekkor definíció szerint

$$g^{-1}(g(u)) = u.$$

Azt kellene megvizsgálnunk, hogy mivel egyenlő $g(g^{-1}(u))$ értéke. Először is lássuk be, hogy ez mindig értelmezett. A $g^{-1}(x)$ függvény értékkészlete csakis olyan számokból áll, amik benne vannak $g(x)$ értelmezési tartományában (hiszen ezen x -ek esetén lesz igaz, hogy $g^{-1}(g(x)) = x$), így

$g(g^{-1}(u))$ mindig értelmezett.

Most megnézzük, hogy mi lehet $g(g^{-1}(u))$ értéke. Ehhez helyettesítsük be ezt az értéket a $g^{-1}(x)$ függvénybe. Ekkor tehát megkapjuk a $(g^{-1} \circ (g \circ g^{-1}}))(u)$ értéket, ami az asszociativitás miatt egyenlő $((g^{-1} \circ g) \circ g^{-1})(u)$ -val. Az inverzfüggvény definíciója alapján

$$((g^{-1} \circ g) \circ g^{-1})(u) = g^{-1}(u). \text{ Az egyenlőség tranzitivitása miatt } g^{-1}(u) = (g^{-1} \circ (g \circ g^{-1}}))(u).$$

A (ii)-ben beláttuk, hogy $g^{-1}(x)$ függvény injektív, ezért biztosan $g(g^{-1}(u)) = u$. Mivel ez minden u -ra igaz, ami a $g(x)$ függvény értelmezési tartományából való, ezért ezzel a (iii)-ast is beláttuk.

A (iv) bizonyítása:

Azt bizonyítjuk, hogy a $g^{-1}(x)$ függvény is szigorúan monoton, ráadásul monotonitása azonos $g(x)$ -ével. Legyenek ugyanis $u < v$ olyan helyek, ahol g értelmezve van. Az egyszerűség kedvéért tegyük

fel, hogy g szigorúan monoton növekvő (a csökkenő eset ugyanígy lenne vizsgálható). Mivel $u < v$, ezért $g(u) < g(v)$. Azt állítjuk, hogy ekkor szükségképpen $g^{-1}(u) < g^{-1}(v)$. Tegyük fel indirekt, hogy $g^{-1}(u) \geq g^{-1}(v)$. Ekkor g szigorú monoton növekedése folytán

$g(g^{-1}(u)) \geq g(g^{-1}(v))$ lenne, ami azt jelentené, hogy $u \geq v$, és ez nyilvánvalóan ellentmondás. ■

Az imént bizonyított tételben figyelemreméltó tulajdonságot kaptunk a (iii)-ban. Ez szavakkal elmondva lényegében ugyanis azt jelenti, hogy **függvény inverzének inverze mindig az eredeti függvény**. Ugyancsak a (iii)-as pontból kristálytiszta látszik, hogy **amennyiben egy $g(x)$ függvény injektív, akkor biztosan létezik inverze**. Ugyanis tetszőleges értelmezési tartománybeli u esetén a $g^{-1}(u)$ -nak meg kell egyeznie azzal az értékkel, amelyet a g függvénybe helyettesítve u -t kapunk.

Nézzünk most néhány példát. Ha c valós szám, akkor a pozitív számok halmazán értelmezett $g(x) = x^c$ hatványfüggvény szigorúan monoton, tehát van inverze. A keresett $g^{-1}(x)$ inverzfüggvényre tehát teljesül, hogy minden pozitív x esetén $g(g^{-1}(x)) = x$, azaz $(g^{-1}(x))^c = x$. Innen $g^{-1}(x) = x^{\frac{1}{c}}$.

Tudjuk, hogy $A > 0, A \neq 1$ konstans esetén a $g(x) = A^x$ függvény szigorúan monoton. Ennek a függvénynek definíció alapján az inverze az **A alapú logaritmusfüggvény**, melynek műveleti tulajdonságai egyenesen következnek a hatványozás azonosságából.

Tétel: Legyen A pozitív, 1-től különböző valós szám. Ekkor bármely x és y pozitív valós, valamint t valós számokra fennállnak az alábbi azonosságok:

$$(i) \log_A(xy) = \log_A(x) + \log_A(y)$$

$$(ii) \log_A\left(\frac{1}{x}\right) = -\log_A(x)$$

$$(iii) \log_A(x^t) = t \cdot \log_A(x)$$

Bizonyítás:

Az (i) bizonyítása:

$\log_A(xy)$ úgy értelmezhető, mint az a hatványkitevő, amelyre $A^{\log_A(xy)} = xy$ teljesül.

Ugyanígy $A^{\log_A(x)} = x$, $A^{\log_A(y)} = y$.

Tehát xy felírható egyrészt úgy, hogy $A^{\log_A(xy)} = xy$, másrészt úgy is, hogy $A^{\log_A(x)} \cdot A^{\log_A(y)} = xy$.

A második felírás a hatványazonosságok miatt $A^{\log_A(x) + \log_A(y)} = xy$.

A szigorú monotonitás miatt ebből kapjuk, hogy $\log_A(xy) = \log_A(x) + \log_A(y)$.

Az (ii) bizonyítása:

Tudjuk, hogy $A^{\log_A(\frac{1}{x})} = \frac{1}{x}$.

Ha megcseréljük a kitevő előjelét, akkor kapjuk, hogy $A^{-\log_A(\frac{1}{x})} = x$.

Ugyanakkor tudjuk, hogy $A^{\log_A(x)} = x$. Ismét a szigorú monotonitás miatt adódik, hogy

$$-\log_A\left(\frac{1}{x}\right) = \log_A(x).$$

Az (iii) bizonyítása:

Tudjuk, hogy $A^{\log_A(x^t)} = x^t$, illetve $A^{\log_A(x)} = x$.

A hatványazonosságokat használva $x^t = (A^{\log_A(x)})^t = A^{t \cdot \log_A(x)}$.

A szigorú monotonitás miatt $t \cdot \log_A(x) = \log_A(x^t)$. ■

A szögfüggvények nem szigorúan monoton függvények, viszont ha az értelmezési tartományukat leszűkítjük mondjuk a $(0, \frac{\pi}{2})$ intervallumra, akkor ott már azok lesznek, és ezért ott invertálhatók.

Így nyerjük az arkusz szinuszt, az arkusz koszinuszt, az arkusz tangenst és az arkusz kotangenst mint függvényeket.

A függvény és inverzének geometriai kapcsolata

Érdekes dolgot fedezhetünk fel, ha ábrázolunk egy injektív függvényt az inverzával együtt a koordinátarendszerben. Ezt az alábbi állításban tárgyaljuk részletesen.

Állítás: Egy injektív függvénynek és az inverzének grafikonja egybevágó görbék, méghozzá egymás tükörképei az $x=y$ egyenletű egyenesre nézve (ez az az egyenes, ami felezi az első és a harmadik síknegyedbeli derékszöget).

Bizonyítás:

Legyen az invertálható függvényünk $g(x)$, valamint jelölje a bizonyítás során tetszőleges T pontnak az $x=y$ egyenesre vonatkozó tükörképét T' .

Először is azt fogjuk belátni, hogy az alábbi két állítás ekvivalens egymással:

(i) P rajta van g grafikonján

(ii) P' rajta van g^{-1} grafikonján

Először is lássuk be, hogy az (i)-esből következik a (ii)-es.

Legyen a kiinduló invertálható függvényünk $g(x)$, aminek grafikonján rajta van a P pont, aminek koordinátái $(u, g(u))$. Ennek a pontnak az $x=y$ egyenesre vett P' tükörképének koordinátái $(g(u), u)$, ez pedig biztosan rajta van g inverzének grafikonján, hiszen $g^{-1}(g(u)) = u$.

Most lássuk be azt, hogy a (ii)-esből következik az (i)-es. Használjuk most fel, hogy inverzfüggvény inverze az eredeti függvény. Ugyanazzal a gondolatmenettel, amit az előbb alkalmaztunk, kapjuk, hogy ha P' rajta van g^{-1} grafikonján, akkor $(P')'$ rajta van g grafikonján. Viszont a tengelyes tükrözés tulajdonságai miatt $(P')' = P$, és ezzel készen is vagyunk.

Beláttuk tehát, hogy (i) és (ii) ekvivalensek. Mivel tetszőleges P pontra $(P')' = P$, ezért a két állítás egyenértékűsége pontosan azt jelenti, hogy a két grafikon egymásból az $x=y$ egyenesre vonatkozó tengelyes tükrözéssel nyerhető. ■

Kevésbé formálisan megfogalmazva a fenti bizonyításban az történt, hogy a síkbeli pontokat párokba rendeztük. Minden pontnak az a pont a párja, ami az ő $x=y$ egyenesre vonatkozó tükörképe. Ezzel valóban egyértelműen párokat hoztunk létre (az $x=y$ egyenes pontjai önmagukkal alkotnak párt). Ezek után azt láttuk be, hogy egy páros egyik tagja akkor és csak akkor van rajta g grafikonján, ha a páros másik tagja rajta van g^{-1} grafikonján. Ez pedig éppen azt jelenti, hogy a két grafikon egymásból az $x=y$ egyenesre vonatkozó tengelyes tükrözéssel nyerhető.

Az összetett függvények és a folytonosság

Először is belátjuk, hogy a folytonos függvényekből képezett összetett függvények mind folytonosak.

Tétel: Ha f és g valós folytonos függvényeknek létezik $f(g(x))$ kompozíciója, akkor $f(g(x))$ is folytonos.

Bizonyítás:

Legyen u valós szám, és (x_n) olyan sorozat, ami u -hoz tart, és aminek elemei benne vannak g értelmezési tartományában. Ekkor a g folytonossága miatt a $(g(x_n))$ sorozat tart $g(u)$ -hoz, az f folytonossága miatt pedig az $(f(g(x_n)))$ sorozat tart $f(g(u))$ -hoz. ■

Most azt fogjuk bebizonyítani, hogy invertálható folytonos függvényeknek az inverze is folytonos.

Tétel: Szigorúan monoton, folytonos függvény inverze szintén folytonos.

Bizonyítás:

Legyen u az $f(x)$ szigorúan monoton, folytonos függvény értelmezési tartományának egy pontja, ami egyben torlódási pont is. Tekintsünk egy u -hoz konvergáló (y_n) sorozatot, aminek elemei benne vannak f^{-1} értelmezési tartományában. Azt kellene belátnunk, hogy $f^{-1}(y_n)$ sorozat konvergens, és

határértéke $f^{-1}(u)$. Azt tudjuk, hogy az $f(f^{-1}(y_n))$ sorozat határértéke $f(f^{-1}(u))$ -val egyenlő.

Használjuk fel, hogy f függvény szigorúan monoton, feltehető például, hogy szigorúan monoton növekvő.

Ez alapján belátjuk, hogy $f^{-1}(y_n)$ tart $f^{-1}(u)$ -hoz. Tegyük fel indirekt, hogy létezik olyan (y_n) sorozat, ami u -hoz tart, de $f^{-1}(y_n)$ nem tart $f^{-1}(u)$ -hoz. Ez azt jelenti, hogy létezik olyan $\alpha > 0$, hogy végtelen sok n indexre $|f^{-1}(y_n) - f^{-1}(u)| \geq \alpha$. Tehát végtelen sok n -re teljesül, hogy

$$y_n = f(f^{-1}(y_n)) \leq f(f^{-1}(u) - \alpha) < f(f^{-1}(u)) = u, \text{ vagy végtelen sok } n\text{-re teljesül, hogy}$$

$$y_n = f(f^{-1}(y_n)) \geq f(f^{-1}(u) + \alpha) > f(f^{-1}(u)) = u.$$

Ekkor $f(f^{-1}(y_n))$ sorozat nem tarthat $f(f^{-1}(u))$ -hoz, ami ellentmondás. ■

Az imént bizonyított tétel következménye, hogy a **logaritmusfüggvények folytonosak**.

Ennek tudatában könnyedén igazolhatunk egy logaritmusokra vonatkozó nevezetes határértéket.

Induljunk ki abból a már bizonyított tényből, hogy $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$.

A logaritmusfüggvény folytonossága miatt $\lim_{x \rightarrow 0} \log_e(1+x)^{\frac{1}{x}} = \log_e e = 1$, ami másképp írva

$$\lim_{x \rightarrow 0} \frac{\log_e(1+x)}{x} = 1.$$

Az Euler-szám alapú logaritmust $\ln$ -nel jelölve ez éppen azt jelenti, hogy

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1.$$

Számtani és mértani közepsorozatok, néhány nevezetes határérték

Tudva, hogy a logaritmusfüggvény folytonos, egyszerűen tárgyalható bizonyos számsorozatok konvergenciája. Először bevezetünk két elnevezést.

Az (x_n) számsorozat **számtani közepsorozata** az

$$s_n = \frac{x_1 + x_2 + \dots + x_n}{n} \text{ sorozat.}$$

Ha az (x_n) számsorozat elemei nemnegatív valós számok, akkor formálisan ugyanilyen elven tudjuk értelmezni a **mértani közepsorozatát**, ami a

$$g_n = \sqrt[n]{x_1 x_2 \cdot \dots \cdot x_n} \text{ sorozat.}$$

Ezekkel kapcsolatban belátjuk az alábbi tételt:

Tétel:

a, Ha (x_n) konvergens számsorozat, aminek határértéke A , akkor számtani közepsorozatának határértéke is A .

b, Ha a pozitív tagú (y_n) sorozatnak a határértéke h , akkor mértani közepsorozatának határértéke is h .

Bizonyítás:

a, Legyen $\varepsilon > 0$ tetszőleges. A határérték definíciója miatt ehhez az ε számhoz létezik olyan k küszöbindex, hogy minden k -nál nagyobb n index esetén $A - \varepsilon < x_n < A + \varepsilon$ teljesül. Ekkor minden ilyen n indexre

$$\frac{x_1 + x_2 + \dots + x_k + (n-k) \cdot (A - \varepsilon)}{n} < \frac{x_1 + x_2 + \dots + x_n}{n} <$$

$$< \frac{x_1 + x_2 + \dots + x_k + (n-k) \cdot (A + \varepsilon)}{n}, \text{ azaz}$$

$$\frac{x_1 + x_2 + \dots + x_k}{n} + \frac{(n-k) \cdot (A - \varepsilon)}{n} < \frac{x_1 + x_2 + \dots + x_n}{n} <$$

$$< \frac{x_1 + x_2 + \dots + x_k}{n} + \frac{(n-k) \cdot (A + \varepsilon)}{n}.$$

Ha n tart végtelenbe, akkor az egyenlőtlenség lánc bal oldala $(A - \varepsilon)$ -hoz, a jobb oldala pedig $(A + \varepsilon)$ -hoz konvergál. Mivel ε -t tetszőlegesen kicsi pozitív számnak választhatjuk, ezért innen következik, hogy az

$s_n = \frac{x_1 + x_2 + \dots + x_n}{n}$ sorozat konvergens, és határértéke A .

Megjegyzendő, hogy gondolatmenetünk akkor is érvényes, ha $A = +\infty$ vagy $-\infty$.

b, A tételnek ezt az állítását egyszerűen visszavezethetjük az a, pontra. Tekintsük az

$x_n = \ln y_n$ sorozatot. Az (x_n) sorozat számtani közepsorozata

$$s_n = \frac{x_1 + x_2 + \dots + x_n}{n} = \ln \left(\sqrt[n]{y_1 y_2 \cdot \dots \cdot y_n} \right).$$

Felhasználva a logaritmusfüggvény folytonosságát, és az a, pontban bizonyítottakat, azt kapjuk, hogy pozitív h esetén a $\ln(\sqrt[n]{y_1 y_2 \cdot \dots \cdot y_n})$ sorozat konvergens és határértéke $\ln h$, míg $h=0$ esetén ez a sorozat $(-\infty)$ -be tart (h nem lehet negatív, hiszen mindegyik y szám nemnegatív). Szintén a logaritmusfüggvény folytonossága miatt ez azt jelenti, hogy (y_n) sorozat mértani közepsorozatának határértéke h . ■

Alkalmazások: az előző tétel lehetőséget nyújt arra, hogy néhány nevezetes határértéket kényelmesen meghatározzunk.

(i) Az $y_n = n$ sorozatra alkalmazva a tételt, kapjuk, hogy $\sqrt[n]{n!}$ sorozat a $(+\infty)$ -be tart.

(ii) Alkalmazzuk a tételt az $e_n = \left(1 + \frac{1}{n}\right)^n$ sorozatra.

Ennek a mértani közepsorozata

$$\sqrt[n]{\left(1 + \frac{1}{1}\right)^1 \cdot \left(1 + \frac{1}{2}\right)^2 \cdots \left(1 + \frac{1}{n}\right)^n} = \sqrt[n]{\left(\frac{2}{1}\right)^1 \cdot \left(\frac{3}{2}\right)^2 \cdots \left(\frac{n+1}{n}\right)^n}.$$

Vegyük észre, hogy a szorzásnál „keresztbe” tudunk egyszerűsíteni, hiszen a számlálóban k^{k-1} , míg a nevezőben k^k alakú szorzótényezők szerepelnek.

Az egyszerűsítések után kapjuk, hogy

$$\sqrt[n]{\left(\frac{2}{1}\right)^1 \cdot \left(\frac{3}{2}\right)^2 \cdots \left(\frac{n+1}{n}\right)^n} = \sqrt[n]{\frac{(n+1)^n}{n!}} = \frac{n+1}{\sqrt[n]{n!}}.$$

A bizonyított tétel szerint ez a sorozat e -hez tart.

Néhány alapvető tétel folytonos függvényekkel kapcsolatban

Kimondunk és bizonyítunk két olyan tételt folytonos függvényekkel kapcsolatban, melyeket a későbbiekben többször is felhasználunk majd.

Weierstrass-tétel: Ha f egy $\boxed{\text{zárt}}$ $[a,b]$ intervallumon értelmezett folytonos függvény, akkor az f függvénynek az adott intervallumon van maximuma és minimuma is.

Bizonyítás:

Csak a maximumra vonatkozó állítást bizonyítjuk, mivel a minimumra vonatkozó ugyanúgy megy.

Legyen S az f függvény értékkészletének szuprémuma az $[a,b]$ zárt intervallumon. Azt kellene belátni, hogy S is benne van a függvény értékkészletében.

Ha a függvény nem venné fel az S -et, akkor a szuprémum definíciója miatt létezne olyan (x_n) sorozat, melynek elemei az $[a,b]$ -ből valók, és az $(f(x_n))$ sorozat konvergál S -hez. Az (x_n) korlátos sorozatra alkalmazható a Bolzano-Weierstrass-tétel, ami szerint az (x_n) sorozatnak létezik torlódási pontja, legyen ez a z pont. Vizsgáljuk a függvény értékét a z helyen.

Mivel a z torlódási pontja az (x_n) sorozatnak, ezért van olyan $n_1 < n_2 < n_3 < \dots$ indexekből álló sorozat, amelyre fennáll, hogy az $x_{n_1}, x_{n_2}, x_{n_3}, \dots$ sorozat konvergál z -hez.

Ekkor persze az is teljesül, hogy az $(f(x_{n_1})), (f(x_{n_2})), (f(x_{n_3})), \dots$ sorozat az S -hez konvergál. Ugyanakkor a folytonosság miatt az $(f(x_{n_1})), (f(x_{n_2})), (f(x_{n_3})), \dots$ sorozat az $f(z)$ -hez konvergál. Innen következik, hogy $f(z)=S$, tehát a függvény mindenképp felveszi az S értéket. ■

Vegyük észre, hogy ki lett használva az a tény, hogy a folytonos függvényünk zárt intervallumon volt értelmezve. Ugyanis a zártság miatt a z torlódási pontnak bele kellett esnie az $[a, b]$ zárt intervallumba. Olyan intervallumon, amely nem zárt, általánosan nem igaz a tétel állítása, pl. a $(0, 1]$ -en értelmezett $f(x) = \frac{1}{x}$ függvénynek nincsen maximuma.

Zárt intervallumon értelmezett folytonos függvények értékkészletével kapcsolatos az ún. Bolzano-tétel. Ennek az a lényege, hogy ha egy folytonos függvény felvesz egy-egy értéket a zárt intervallum két végpontjában, akkor az intervallumon belül minden köztes értéket is fel kell vennie.

Bolzano-tétel: Legyen f a zárt $[a, b]$ -on értelmezett folytonos függvény.

- (i) Ha $f(a)$ és $f(b)$ különböző előjelűek, akkor az f függvény az $[a, b]$ intervallumon valahol felveszi a 0 értéket is.
- (ii) Az (i) feltételeitől függetlenül az f függvény az $[a, b]$ -on felvesz minden $f(a)$ és $f(b)$ közötti értéket.

Bizonyítás:

Az (i) bizonyítása:

Tegyük fel indirekt, hogy van olyan $[a, b]$ zárt intervallum és olyan f függvény, amire teljesül az, hogy az $[a, b]$ zárt intervallumon folytonos, és amire $f(a)$ és $f(b)$ különböző előjelű, de az f függvény nem veszi fel a 0 értéket az $[a, b]$ -on. Elég azt az esetet vizsgálnunk, amikor $f(a) < 0 < f(b)$. Mielőtt továbbmennénk, bebizonyítjuk a következő észrevételt.

Észrevétel: Ha valamilyen $[a, b]$ -beli r számra $f(r)$ pozitív, akkor van olyan $\delta > 0$, hogy az $(r - \delta, r + \delta)$ nyílt intervallumban a függvény minden értéke pozitív. (Megjegyzés: ugyanez érvényes, ha $f(r)$ negatív, csak itt a környezetben felvett értékek negatívak lesznek.)

Üi: Ha az észrevétel nem lenne igaz, akkor minden $\delta > 0$ esetén létezne olyan x , amelyre $|r - x| < \delta$ és $f(x) \leq 0$. Ekkor gyárthatunk egy sorozatot az ilyen rossz x -ekből, legyen ez a sorozat (x_n) sorozat. Ha $\delta \rightarrow 0$, akkor az is teljesül, hogy $x_n \rightarrow r$. Ekkor a folytonosság miatt $f(x_n) \rightarrow f(r)$. Ez itt egy ellentmondás, hiszen az $f(x_n)$ sorozat elemei nem lehetnek nagyobbak nullánál, de $f(r)$ pozitív. Ezért az észrevétel állítása igaz.

Az észrevételünk miatt van olyan $\delta > 0$, hogy $(a + \delta < b)$ és a függvény az $[a, a + \delta)$ -on negatív értéket vesz fel.

Tudjuk, hogy $f(b)$ pozitív, ezért az előző δ számoknak van egy szupréma, mondjuk D . Vizsgáljuk meg a függvényt az $(a+D)$ helyen.

Először nézzük meg, hogy lehet-e $f(a+D)$ értéke negatív. Ez nem lehetséges, hiszen akkor az észrevételünk alapján lenne olyan $\delta' > 0$, hogy az f függvény csupa negatív értéket venne fel az $[a, a+D+\delta')$ -n.

Most nézzük meg, hogy lehet-e $f(a+D)$ értéke negatív. Ez sem lehetséges, hiszen ekkor lenne olyan $\delta' > 0$, amire az teljesül, hogy f pozitív értékeket vesz fel az $(a+D-\delta', a+D+\delta')$ -n. Tehát csakis $f(a+D)=0$, ezzel (i)-t beláttuk.

Az (ii) bizonyítása:

Ha $f(a)=f(b)$, akkor nincs mit bizonyítanunk.

Ha $f(a) \neq f(b)$, akkor legyen w egy olyan érték, ami $f(a)$ és $f(b)$ közé esik, és mindkettőtől különböző. Megmutatjuk, hogy a függvény az adott intervallumon felveszi a w helyettesítési értéket.

Tekintsük ugyanis az $f(x)$ függvény helyett a $g(x)=f(x)-w$ függvényt.

Nyilván $g(x)$ folytonos az $[a, b]$ -on, illetve $g(a)$ és $g(b)$ ellentétes előjelűek.

(i) alapján van olyan $v \in [a, b]$, amire $g(v)=f(v)-w=0$.

De ekkor $f(v)=w$. Mivel w tetszőleges $f(a)$ és $f(b)$ közötti érték, ezért készen vagyunk. ■

A Bolzano-tételből könnyen következik, hogy pl. a valós számok halmazán értelmezett exponenciális függvényeknek az összes pozitív valós szám benne van az értékkészletében.

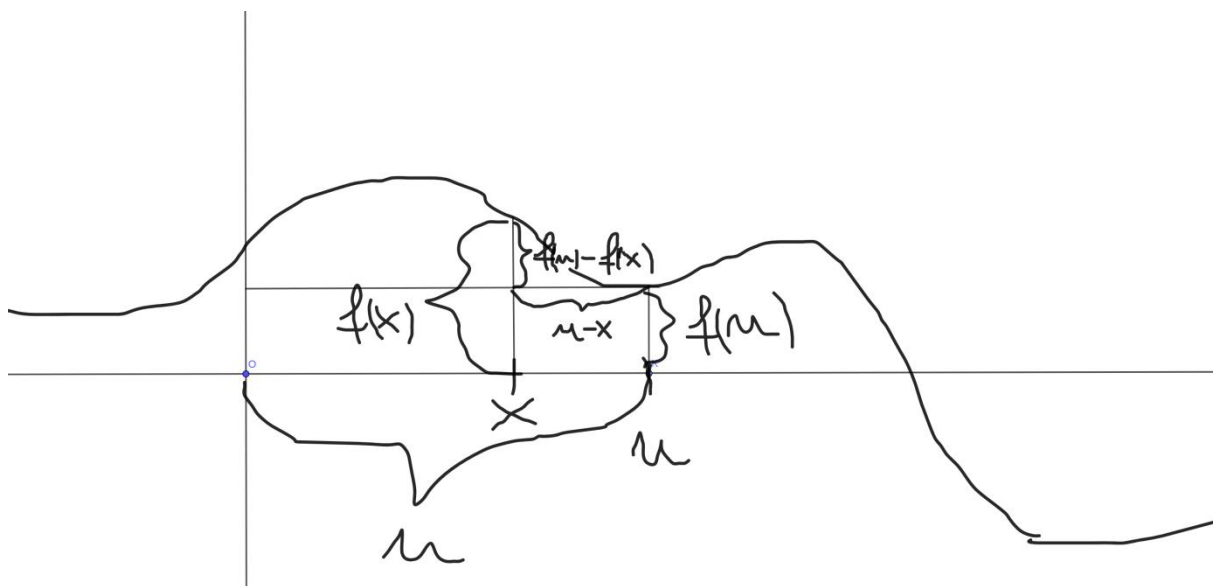
Differenciálszámítás

Már minden eszközünk megvan ahhoz, hogy a függvények differenciálhatóságával foglalkozzunk. Először a pontbeli differenciálhatóság (másképpen pontbeli deriválhatóság) fogalmát tisztázzuk.

Definíció (pontbeli differenciálhatóság/deriválhatóság): Azt mondjuk, hogy az f függvény differenciálható az értelmezési tartományának egy u pontjában, ha létezik és véges a $\lim_{x \rightarrow u} \frac{f(x)-f(u)}{x-u}$ határérték. Ha létezik ez a határérték, akkor ezt $f'(u)$ -val jelöljük.

Azokat a függvényeket, amik az értelmezési tartományuk minden pontjában differenciálhatóak, röviden **differenciálhatóak**nak fogjuk nevezni.

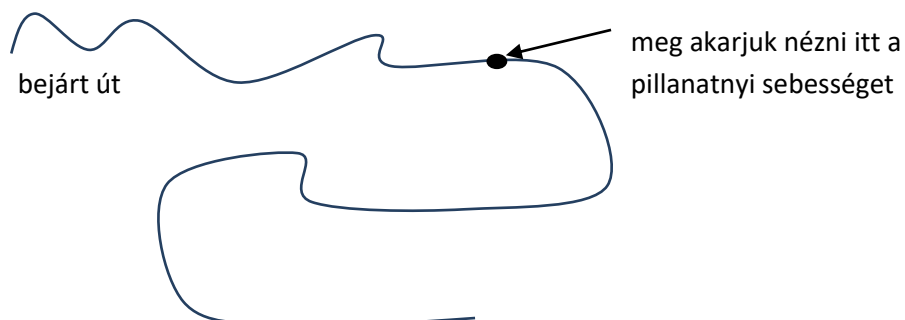
Célszerű átgondolnunk, hogy mi a deriválnak a szemléletes jelentése. Ehhez tekintsük az alábbi ábrát:



Az $\frac{f(x) - f(u)}{x - u}$ tört az ábra szerint egy iránytangens. Ha x tart az u -hoz, akkor úgy sejtethető, hogy ezeknek az iránytangenseknek a határértéke az $(u, f(u))$ pontban húzott érintő iránytangensével egyezik meg.

Valójában ez lesz a **pontbeli érintő iránytangensének** definíciója: az adott pontbeli derivált értéke lesz az ott húzott érintő iránytangense.

Felmerülhet a kérdés, hogy mi adta a motivációt a derivált definíciójához. Ennek demonstrálásához a fizikában értelmezett pillanatnyi sebességet vesszük alapul.



Ha $\vec{s}(t)$ vektor jelöli, hogy t idő elteltével hol tartózkodunk (tehát mik a koordinátáink), és a t időpontban vizsgáljuk a pillanatnyi sebességet, akkor éppen azt vizsgáljuk, hogy létezik-e

$$\lim_{t_0 \rightarrow t} \frac{\vec{s}(t) - \vec{s}(t_0)}{t - t_0} \text{ határérték.}$$

Hasonlóan kapható, hogy a gyorsulás a sebesség idő szerinti deriváltja, ugyanis ekkor a

$\lim_{t_0 \rightarrow t} \frac{\bar{v}(t) - \bar{v}(t_0)}{t - t_0}$ határértéket kell vizsgálni.

A differenciálhatóság és a folytonosság kapcsolata

Tétel: ha az f függvény differenciálható az értelmezési tartományának u pontjában, akkor ott folytonos is.

Bizonyítás: Indirekt tegyük fel, hogy létezik olyan f függvény és olyan u pont, amire f differenciálható u -ban, de ott nem folytonos. Az, hogy f nem folytonos u -ban, azt jelenti, hogy van olyan (x_n) sorozat, aminek elemei f értelmezési tartományába esnek, (x_n) tart u -hoz, viszont $(f(x_n))$ sorozat nem tart $f(u)$ -hoz. Ekkor létezik „rossz” $\varepsilon > 0$, amihez nem találunk jó küszöbindexet, tehát végtelen sok n indexre $|f(x_n) - f(u)| \geq \varepsilon$.

Ezekre az n -ekre $\left| \frac{f(x_n) - f(u)}{x_n - u} \right| \geq \frac{\varepsilon}{|x_n - u|} \rightarrow +\infty$, ez pedig ellentmondás. ■

Megjegyezzük, hogy a tétel megfordítása általánosan nem igaz. Meg lehet adni akár olyan valós függvényt is, ami minden pontjában folytonos, de sehol sem differenciálható.

Elemi függvények deriváltfüggvényei

Értelemszerűen a **deriváltfüggvény** azt a függvényt jelenti, amit egy minden pontjában differenciálható függvény minden pontjában elvégzett differenciálás után kapunk.

Tétel:

(i) Ha $c \in \mathbb{R}$ konstans, akkor a pozitív számok halmazán értelmezett $h(x) = x^c$ függvény deriváltfüggvénye $(x^c)' = c x^{c-1}$.

(ii) $(e^x)' = e^x$.

(iii) A pozitív számok halmazán értelmezett logaritmusfüggvény deriváltfüggvénye $(\ln x)' = \frac{1}{x}$.

(iv) $(\sin x)' = \cos x$.

(v) $(\cos x)' = -\sin x$.

Bizonyítás:

Az (i) bizonyítása:

Legyen $u > 0$. Átalakítások után, majd felhasználva a $\lim_{t \rightarrow 0} \frac{\ln(1+t)}{t} = 1$ nevezetes határértéket, kapjuk, hogy

$$\lim_{x \rightarrow u} \frac{x^c - u^c}{x - u} = \lim_{x \rightarrow u} \frac{u^c \left[\left(\frac{x}{u} \right)^c - 1 \right]}{u \left(\frac{x}{u} - 1 \right)} = \lim_{x \rightarrow u} u^{c-1} \frac{\left(\frac{x}{u} \right)^c - 1}{\frac{x}{u} - 1} = \lim_{x \rightarrow u} u^{c-1} \cdot \frac{\left(\frac{x}{u} \right)^c - 1}{\ln \left(\frac{x}{u} \right)^c} \cdot \frac{\ln \left(\frac{x}{u} \right)^c}{\frac{x}{u} - 1} = \lim_{x \rightarrow u} u^{c-1} \cdot \frac{\left(\frac{x}{u} \right)^c - 1}{\ln \left(\frac{x}{u} \right)^c} \cdot c \cdot \frac{\ln \frac{x}{u}}{\frac{x}{u} - 1} = c \cdot u^{c-1}.$$

Az (ii) bizonyítása:

Legyen u valós szám. Átalakítások után, majd felhasználva a $\lim_{t \rightarrow 0} \frac{\ln(1+t)}{t} = 1$ nevezetes határértéket, kapjuk, hogy

$$\lim_{x \rightarrow u} \frac{e^x - e^u}{x - u} = \lim_{x \rightarrow u} e^u \cdot \frac{e^{x-u} - 1}{x - u} = \lim_{x-u=t \rightarrow 0} e^u \cdot \frac{e^t - 1}{t} = \lim_{e^t=T \rightarrow 1} e^u \cdot \frac{T - 1}{\ln T} = e^u$$

Az (iii) bizonyítása:

Legyen $u > 0$. Átalakítások után, majd felhasználva a $\lim_{t \rightarrow 0} \frac{\ln(1+t)}{t} = 1$ nevezetes határértéket, kapjuk, hogy

$$\lim_{x \rightarrow u} \frac{\ln x - \ln u}{x - u} = \lim_{x \rightarrow u} \frac{\ln \frac{x}{u}}{u \left(\frac{x}{u} - 1 \right)} = \lim_{x \rightarrow u} \frac{1}{u} \cdot \frac{\ln \frac{x}{u}}{\frac{x}{u} - 1} = \lim_{x \rightarrow u} \frac{1}{u} \cdot \frac{\ln \frac{x}{u}}{\frac{x}{u} - 1} = \lim_{\frac{x}{u}=t \rightarrow 1} \frac{1}{u} \cdot \frac{\ln t}{t - 1} = \frac{1}{u}$$

Az (iv) bizonyítása:

Legyen u valós szám. A $(\sin x - \sin u)$ kifejezést úgy fejezzük ki az addíciós tételek segítségével, hogy az $x = \frac{x+u}{2} + \frac{x-u}{2}$, illetve $u = \frac{x+u}{2} - \frac{x-u}{2}$ azonosságokat alkalmazzuk. Az átalakítások után, majd felhasználva a $\lim_{t \rightarrow 0} \frac{\sin t}{t} = 1$ nevezetes határértéket, illetve a koszinusz folytonosságát, kapjuk, hogy

$$\lim_{x \rightarrow u} \frac{\sin x - \sin u}{x - u} =$$

$$= \lim_{x \rightarrow u} \frac{2 \cdot \cos \left(\frac{x+u}{2} \right) \cdot \sin \left(\frac{x-u}{2} \right)}{x - u} = \lim_{x \rightarrow u} \frac{\cos \left(\frac{x+u}{2} \right) \cdot \sin \left(\frac{x-u}{2} \right)}{\frac{x-u}{2}} = \lim_{x \rightarrow u} \cos \left(\frac{x+u}{2} \right) \cdot \frac{\sin \left(\frac{x-u}{2} \right)}{\frac{x-u}{2}} =$$

$$= \cos u \cdot 1 = \cos u.$$

A (v) bizonyítása:

Hasonlóan megy, mint (iv) bizonyítása, nem részletezzük. ■

Differenciálási szabályok

Ha differenciálható függvényekkel alapszabályokat végzünk, akkor általában differenciálható függvényeket kapunk, és a kapott differenciálható függvények deriváltjai is viszonylag kényelmesen jellemezhetők. Erről szól az alábbi tétel.

Tétel (differenciálási szabályok): Ha f és g differenciálható függvények értelmezési tartománya azonos, akkor az értelmezési tartomány tetszőleges u pontjában az teljesül, hogy

1. $(f + g)'(u) = f'(u) + g'(u)$
2. $(f \cdot g)'(u) = f'(u) \cdot g(u) + f(u) \cdot g'(u)$
3. $\left(\frac{f}{g}\right)'(u) = \frac{f'(u) \cdot g(u) - f(u) \cdot g'(u)}{g^2(u)}$; feltéve, hogy $g(u) \neq 0$.

Bizonyítás:

Az összegre vonatkozóan:

$$(f + g)'(u) = \lim_{x \rightarrow u} \frac{(f(x) + g(x)) - (f(u) + g(u))}{x - u} = \lim_{x \rightarrow u} \left(\frac{f(x) - f(u)}{x - u} + \frac{g(x) - g(u)}{x - u} \right)$$

Mivel a zárójelben az adott törteknek az u -ban is van határértéke, ezért a kifejezés kettéválasztható az alábbi módon:

$$\underbrace{\lim_{x \rightarrow u} \frac{f(x) - f(u)}{x - u}}_{f'(u)} + \underbrace{\lim_{x \rightarrow u} \frac{g(x) - g(u)}{x - u}}_{g'(u)}$$

A szorzatra vonatkozóan:

$$(f \cdot g)'(u) = \lim_{x \rightarrow u} \frac{f(x)g(x) - f(u)g(u)}{x - u} = \lim_{x \rightarrow u} \frac{f(x)g(x) - f(u)g(x) + f(u)g(x) - f(u)g(u)}{x - u} =$$

$$\lim_{x \rightarrow u} \left[\frac{g(x) \cdot (f(x) - f(u))}{x - u} + \frac{f(u) \cdot (g(x) - g(u))}{x - u} \right].$$

Tudjuk, hogy $\lim_{x \rightarrow u} \frac{f(x) - f(u)}{x - u} = f'(u)$; $\lim_{x \rightarrow u} \frac{g(x) - g(u)}{x - u} = g'(u)$, illetve $g(x)$ folytonossága miatt

$\lim_{x \rightarrow u} g(x) = g(u)$. Innen a fenti határérték létezik és értéke $g(u) \cdot f'(u) + f(u) \cdot g'(u)$.

A hányadosra vonatkozóan:

$$\begin{aligned} \left(\frac{f}{g}\right)'(u) &= \lim_{x \rightarrow u} \frac{\frac{f(x)}{g(x)} - \frac{f(u)}{g(u)}}{x - u} = \lim_{x \rightarrow u} \frac{\frac{f(x)g(u) - g(x)f(u)}{g(x) \cdot g(u)}}{x - u} = \lim_{x \rightarrow u} \frac{f(x)g(u) - g(x)f(u)}{g(x)g(u) \cdot (x - u)} \\ &= \lim_{x \rightarrow u} \frac{\frac{f(x)g(u) - g(x)f(u)}{x - u}}{g(x)g(u)} = \lim_{x \rightarrow u} \frac{\frac{f(x)g(u) - f(u)g(u) + g(u)f(u) - g(x)f(u)}{x - u}}{g(x)g(u)} \\ &= \lim_{x \rightarrow u} \frac{g(u) \cdot \frac{f(x) - f(u)}{x - u} - f(u) \cdot \frac{g(x) - g(u)}{x - u}}{g(x)g(u)} = \frac{g(u) \cdot f'(u) - f(u) \cdot g'(u)}{g^2(u)}. \blacksquare \end{aligned}$$

Megjegyzendő, hogy a szorzatra vonatkozó szabályból könnyen látható, hogy tetszőleges c konstans esetén $(c \cdot f(x))' = c \cdot f'(x)$.

Differenciálható függvényekből képzett összetett függvény is általában differenciálható. Erről szól a következő tétel.

Tétel: Ha f és g olyan differenciálható függvények, amelyeknek létezik az $f(g(x))$ kompozíciója, akkor $f(g(x))$ is differenciálható és a deriváltfüggvénye: $f(g(x))' = f'(g(x)) \cdot g'(x)$.

Bizonyítás:

Ha u olyan pont, ahol g differenciálható, és az f függvény differenciálható a $g(u)$ -ban, akkor

$$\lim_{x \rightarrow u} \frac{f(g(x)) - f(g(u))}{x - u} = \lim_{x \rightarrow u} \frac{f(g(x)) - f(g(u))}{g(x) - g(u)} \cdot \frac{g(x) - g(u)}{x - u} = f'(g(u)) \cdot g'(u).$$

Itt felhasználtuk, hogy g folytonos az u -ban, és ezért $\lim_{x \rightarrow u} g(x) = g(u)$, illetve felhasználtuk az f -re és g -re vonatkozó deriválhatósági feltételeket. $\blacksquare$

Tekintsünk most néhány példát:

1. $e^{x \cdot \sin x}$

$$f(x) = e^x; g(x) = x \cdot \sin x$$

$$f(g(x)) = e^{x \cdot \sin x}$$

$$(e^{x \cdot \sin x})' = f'(g(x)) \cdot g'(x) = e^{x \cdot \sin x} \cdot (x \cdot \sin x)' = e^{x \cdot \sin x} \cdot (1 \cdot \sin x + x \cdot \cos x)$$

$$2. \quad 2017^x = (e^{\ln 2017})^x = e^{x \cdot \ln 2017}$$

$$f(x) = e^x; \quad g(x) = x \cdot \ln 2017$$

$$f(g(x)) = e^{x \cdot \ln 2017}$$

$$(e^{x \cdot \ln 2017})' = f'(g(x)) \cdot g'(x) = e^{x \cdot \ln 2017} \cdot (x \cdot \ln 2017)' = 2017^x \cdot \ln 2017$$

Általánosításként kapjuk, hogy amennyiben $a > 0$, akkor az $(a^x)' = a^x \cdot \ln a$.

$$3. \quad \ln(\cos^2 x)$$

$$f(x) = \ln x; \quad g(x) = \cos^2 x$$

$$f(g(x)) = \ln(\cos^2 x)$$

$$(\ln(\cos^2 x))' = f'(g(x)) \cdot g'(x) = \frac{1}{\cos^2 x} \cdot (\cos^2 x)' =$$

A $\cos^2 x$ is összetett függvény, $r(x) = x^2$, $s(x) = \cos x$ függvényekből.

$$r(s(x)) = \cos^2 x$$

$$(\cos^2 x)' = r'(s(x)) \cdot s'(x) = 2 \cos x (-\sin x)$$

$$= \frac{1}{\cos^2 x} = 1 \cos x \cdot (-\sin x)$$

$$4. \quad r(x)^{s(x)} = (e^{\ln r(x)})^{s(x)} = e^{s(x) \cdot \ln r(x)}$$

$$\begin{aligned} (e^{s(x) \cdot \ln r(x)})' &= e^{s(x) \cdot \ln r(x)} \cdot (s(x) \cdot \ln r(x))' = \\ &= r(x)^{s(x)} \cdot [s'(x) \cdot \ln r(x) + s(x) \cdot (\ln r(x))'] = \\ &= r(x)^{s(x)} \cdot \left[s'(x) \cdot \ln r(x) + s(x) \cdot \frac{1}{r(x)} \cdot r'(x) \right] \end{aligned}$$

Inverzfüggvény deriváltja

Az alábbi tétel differenciálható függvény inverzéről szól (amennyiben az létezik).

Tétel: Ha f intervallumon értelmezett, invertálható és differenciálható függvény, akkor inverze is differenciálható.

Bizonyítás:

Legyen u az intervallum tetszőleges pontja.

$$\lim_{x \rightarrow u} \frac{f^{-1}(x) - f^{-1}(u)}{x - u} = \lim_{x \rightarrow u} \frac{f^{-1}(x) - f^{-1}(u)}{f(f^{-1}(x)) - f(f^{-1}(u))} = \lim_{x \rightarrow u} \frac{1}{\frac{f(f^{-1}(u)) - f(f^{-1}(x))}{f^{-1}(u) - f^{-1}(x)}}$$

Az f függvény folytonos, hiszen differenciálható, valamint f inverze is folytonos. Ezek alapján

$$\lim_{x \rightarrow u} \frac{1}{\frac{f(f^{-1}(u)) - f(f^{-1}(x))}{f^{-1}(u) - f^{-1}(x)}} = \frac{1}{f'(f^{-1}(u))}. \blacksquare$$

Látható, hogy a bizonyítás során az inverzfüggvény deriváltjára nézve képletet is kaptunk.

Példaként határozzuk meg a szinuszfüggvény inverzének deriváltját (itt x végig az első síknegyedből való). A kapott képlet szerint, illetve felhasználva a trigonometrikus Pithagorasz-tételt, kapjuk, hogy

$$(\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1 - x^2}}.$$

A többi nevezetes differenciálható függvény inverzének deriváltfüggvénye (a megfelelő értelmezési tartományon belül) hasonlóan kapható meg, ezeket itt nem részletezzük.

Függvények lokális monotonitása

Definíció (lokális monotonitás): Azt mondjuk, hogy az f függvény lokálisan szigorúan monoton növekedő az értelmezési tartományának egy u pontjában, ha létezik olyan $\delta > 0$, amire az teljesül, hogy

- (1) minden $u - \delta < x < u$ értelmezési tartományba eső x esetén $f(x) < f(u)$
- (2) minden $u < x < u + \delta$ értelmezési tartományba eső x esetén $f(u) < f(x)$

Megjegyzés: a pontbeli (lokális) szigorúan monoton csökkenés analóg módon definiálható.

Észrevétel: A pontbeli lokális szigorú monotonitás nem vonja maga után, hogy a függvény valamilyen intervallumon is monoton lenne. Legyen ugyanis

$$f(x) = \begin{cases} x, & \text{ha } x \text{ racionális} \\ 2x, & \text{ha } x \text{ irracionális} \end{cases}$$

Ez a függvény 0-ban lokálisan szig. mon. növ., de semmilyen (nemelfajuló) intervallumon nem monoton.

Fontos észrevétel (ami majd kapcsolatot teremt a deriváltakkal): Ha f az egész $\mathbb{R}$ -en értelmezett és az u pontban lok. szig. mon. növ., akkor létezik olyan $\delta > 0$, amire igaz, hogy

$$(i) \quad u - \delta < x < u \text{ esetén } f(x) - f(u) < 0 \text{ pontosan akkor teljesül, ha } \frac{f(x) - f(u)}{x - u} > 0$$

(ii) $u < x < u + \delta$ esetén $f(x)-f(u) > 0$ pontosan akkor teljesül, ha $\frac{f(x)-f(u)}{x-u} > 0$.

TÉTEL: Legyen $f: R \rightarrow R$ függvény, ami u -ban differenciálható.

(i) Ha az $f'(u) > 0$, akkor f az u -ban lokálisan szigorúan monoton növekedő.

(ii) Ha az $f'(u) < 0$, akkor f az u -ban lokálisan szigorúan monoton csökkenő.

(iii) Ha f lokálisan monoton növekvő u -ban, akkor $f'(u) \geq 0$ (itt a feltételben elengedtük a szigorúságot, tehát egyenlőség is megengedett).

(iv) Ha f lokálisan monoton csökkenő u -ban, akkor $f'(u) \leq 0$ (a fenti zárójeles megjegyzés szintén érvényes).

(v) Ha f -nek lokális szélsőértéke van u -ban, akkor $f'(u) = 0$.

Bizonyítás:

Az (i) bizonyítása:

Tudjuk, hogy $f'(u) = \lim_{x \rightarrow u} \frac{f(x)-f(u)}{x-u} > 0$.

Léteznie kell olyan $\delta > 0$ számnak, amire igaz, hogy minden olyan x -re, ami $|x - u| < \delta$, az teljesül, hogy $\frac{f(x)-f(u)}{x-u} > 0$. Ui: Ha nem lenne ilyen δ , akkor létezne olyan $x_n \rightarrow u$ sorozat,

amire az teljesülne, hogy $\frac{f(x_n)-f(u)}{x_n-u} \leq 0$. Viszont ekkor ha létezik is a $\lim_{n \rightarrow \infty} \frac{f(x_n)-f(u)}{x_n-u}$

határérték, akkor az nem lehet pozitív, és ez ellentmond az u -beli differenciálhatóságnak.

Tehát van ilyen $\delta > 0$.

Az előző megjegyzés miatt az f -nek u -ban lokálisan szigorú monoton növekvőnek kell lennie.

Az (ii) bizonyítása:

Ugyanaz, mint az (i)-nél, csak előjelcserével.

Az (iii) bizonyítása:

Egy alkalmas intervallumon $\frac{f(x)-f(u)}{x-u} \geq 0$, ezért $\lim_{x \rightarrow u} \frac{f(x)-f(u)}{x-u} \geq 0$.

A (iv) bizonyítása:

Ugyanaz, mint a (iii)-nél, csak előjelcserével.

A (v) bizonyítása:

Az u -ban lokális szélsőértékhely van.

Ha balról tartunk x -szel az u -hoz, akkor $\frac{f(x)-f(u)}{x-u}$ előjele ellentétes, mintha jobbról tartanánk x -szel az u -hoz. Ezért $\lim_{x \rightarrow u} \frac{f(x)-f(u)}{x-u}$ nem lehet sem pozitív, sem negatív. Így tehát ez a határérték szükségképpen 0. ■

Megjegyzés: A (v) megfordítása nem igaz. Abból, hogy a differenciálható függvénynek valamely pontjában 0 a deriváltja, még nem következik, hogy ott lokális szélsőértéke van. Pl. tekintsük az $f(x) = x^3$ függvényt az $u=0$ pontban.

Megjegyzés: Olyan többváltozós függvényeknél, amik minden változójukban parciálisan differenciálhatók (az adott változó szerint deriválunk), azokra az 5. feltétel szolgáltat (szükséges feltételt) pontbeli lokális szélsőérték létezésére. Ha az $f(x_1, x_2, \dots, x_n)$ eleget tesz a fenti feltételeknek és az $(u_1, u_2, \dots, u_n)$ helyen lokális szélsőértéke van, akkor

$$f'_{x_1}(u_1, u_2, \dots, u_n) = f'_{x_2}(u_1, \dots, u_n) = \dots = f'_{x_n}(u_1, \dots, u_n) = 0.$$

Középértéktételek

ROLLE-TÉTEL: Legyen f az $[a, b]$ -on értelmezett folytonos függvény, ami a belső pontokban differenciálható is, illetve legyen $f(a)=f(b)$. Ekkor létezik olyan $c \in (a, b)$, amire $f'(c)=0$.

Bizonyítás: Ha a függvény konstans függvény, akkor bármelyik belső pont jó, mert a konstans függvény deriváltfüggvénye az azonosan 0 függvény. Amennyiben nem ez a helyzet, akkor van olyan $x_0 \in (a, b)$, amire $f(x_0) \neq f(a)=f(b)$.

I. eset: ha $f(x_0) > f(a)=f(b)$. Ebben az esetben a Weierstrass-tételt alkalmazva találunk olyan $c \in [a, b]$ számot, ahol az f függvény felveszi a maximumát. Nyilván $f(c) \geq f(x_0)$, ezért $c \neq a, b$. Mivel c lokális szélsőértékhely, ezért $f'(c)=0$.

II. eset: ha $f(x_0) < f(a)=f(b)$, akkor ugyanúgy járunk el, csak minimumhellyel. ■

LAGRANGE-KÖZÉPÉRTÉKTÉTEL: Ha f folytonos az $[a, b]$ -on és a belső pontokban differenciálható, akkor létezik olyan $c \in (a, b)$, amire $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Bizonyítás:

Legyen $F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a} \cdot (x - a)$.

Tudjuk, hogy ez a $F(x)$ folytonos az $[a, b]$ -on és differenciálható (a, b) -on. Számítsuk ki a $F(a)$ és $F(b)$ értékeket.

$$F(a) = f(a) - f(a) - \frac{f(b)-f(a)}{b-a} \cdot (a-a) = 0$$

$$F(b) = f(b) - f(a) - \frac{f(b)-f(a)}{b-a} \cdot (b-a) = 0$$

Tehát az F függvényre alkalmazható a Rolle-tétel. A Rolle-tétel szerint létezik olyan $c \in (a, b)$, amire $0 = F'(c)$.

Az $F(x)$ deriváltfüggvénye $F'(x) = \left[f(x) - f(a) - \frac{f(b)-f(a)}{b-a} \cdot (x-a) \right]' = f'(x) - \frac{f(b)-f(a)}{b-a}$, innen

$$f'(c) - \frac{f(b)-f(a)}{b-a} = 0. \blacksquare$$

Megjegyzés: Szemléletesen ez azt jelenti, hogy a $(c, f(c))$ pontban húzott érintőnek kell párhuzamosnak lennie az $(a, f(a))$ és $(b, f(b))$ pontokat összekötő egyenessel.

CAUCHY-KÖZÉPÉRTÉKTÉTEL: f és g az $[a, b]$ -on folytonos, a belső pontokban differenciálható függvények, továbbá a $g'(x)$ függvény nem vesz fel 0 értéket a belső pontokban.

Ekkor létezik olyan $c \in (a, b)$, amire $\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}$.

Előzetes megjegyzés: automatikusan teljesül, hogy $g(b) \neq g(a)$, ugyanis ha egyenlők lennének, akkor a Rolle-tétel miatt lenne olyan $y \in (a, b)$, amire $g'(y) = 0$.

Bizonyítás: Legyen $F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot (g(x) - g(a))$

A $F(x)$ folytonos az $[a, b]$ -on és differenciálható a belső pontokban.

$$F(a) = f(a) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot (g(a) - g(a)) = 0$$

$$F(b) = f(b) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot (g(b) - g(a)) = 0$$

Az F függvényre alkalmazva a Rolle-tételt, kapjuk, hogy létezik olyan $c \in (a, b)$, amire $F'(c) = 0$.

$$F'(x) = f'(x) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot g'(x)$$

Innen kapjuk, hogy $0 = F'(c) = f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot g'(c)$, amiből átrendezéssel adódik, hogy

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

A Lagrange-közéértéktétel fontos következménye: Ha f folytonos az $[a, b]$ -on és differenciálható (a, b) -on, valamint minden $x \in (a, b)$ -re $f'(x) = 0$, akkor f szükségképpen konstans $[a, b]$ -n.

Bizonyítás:

Lagrange-közéértéktételt használunk.

Legyen $d \in (a,b)$. A $[d,b]$ -on értelmezett f függvényre alkalmazzuk a Lagrange-közéértéktételt. Léteznie kell egy $c \in (d,b)$ számnak, amire $f'(c) = \frac{f(d)-f(b)}{d-b}$. Innen következik, hogy $f(d)=f(b)$. Mivel d tetszőleges belső pont volt, ezért a belső pontokban konstans a függvény. A folytonosság miatt ez persze a végpontokra is kiterjeszthető, és így beláttuk, hogy a függvény konstans az adott intervallumon. ■

Észrevétel: Ha f és g folytonos az $[a,b]$ -on, és differenciálható (a,b) -on, valamint minden $x \in (a,b)$ -re $f'(x)=g'(x)$, akkor $(f(x)-g(x))$ egy konstans függvény.

Bizonyítás: Az előző következményre visszavezethető az alábbi módon:

$$f'(x) = g'(x) \leftrightarrow f'(x) - g'(x) = 0 \leftrightarrow (f(x) - g(x))' = 0. \blacksquare$$

TÉTEL: Legyen az f folytonos az $[a,b]$ -on és differenciálható (a,b) -on. Ekkor

(i) ***f pontosan akkor*** monoton növekedő az $[a,b]$ -on, ha

$$f'(x) \geq 0 \text{ minden } x \in (a,b)\text{-re.}$$

(ii) ***f pontosan akkor*** szigorúan monoton növekedő az $[a,b]$ -on, ha $f'(x) > 0$ minden $x \in (a,b)$ -re és az $[a,b]$ -nek semmilyen (nemelfajuló) részintervallumán az $f'(x)$ nem azonosan 0.

Bizonyítás:

(i) **Egyik irány:** $f'(x) \geq 0$ minden $x \in (a,b)$ -re.

Legyen $x_1, x_2 \in [a,b]$, $x_1 < x_2$. A Lagrange-közéértéktételt alkalmazzuk az $[x_1, x_2]$ -on értelmezett f -re. Van olyan $c \in (x_1, x_2)$, amire $f'(c) = \frac{f(x_1)-f(x_2)}{x_1-x_2} \geq 0$. Mivel (x_1-x_2) negatív, ezért $f(x_1)-f(x_2) \leq 0$, innen $f(x_1) \leq f(x_2)$. Mivel x_1 és x_2 tetszőleges belső pontok, ezért bizonyítottuk.

Másik irány: ha f monoton növekvő az $[a,b]$ -on.

Ennek az iránynak bizonyításakor hivatkozhatunk a közéértéktételek előtti ismereteinkre. Amennyiben f mon. növ. $[a,b]$ -on, akkor minden (a,b) -beli pontban lokálisan monoton növekvő. Ebből következett, hogy minden $x \in (a,b)$ -re $f'(x) \geq 0$.

(ii) Ugyanúgy, mint az 1-nél, azzal a különbséggel, azzal a különbséggel, hogy fel kell használnunk, hogy ha egy nem elfajuló intervallumon a derivált azonosan 0, akkor a függvény azon az intervallumon konstans. ■

Megjegyzés: Monoton csökkenésre és szigorú monoton csökkenésre ugyanígy megfogalmazható a tétel, csak az egyenlőtlenségek fordítva állnak.

Példa: a monotonitási feltételek (és az előtte szereplő tételek) alkalmazása:

$f(x) = x + \frac{1}{x}$. Határozzuk meg a lokális szélsőértékeit!

Megoldás: Az $f(x)$ differenciálható függvény (a 0-ban nincs értelmezve). Mivel lokális szélsőérték(hely)eket keresünk, ezért először meg kell néznünk, hogy $f'(x)$ mikor vesz fel 0 értéket, illetve mikor nincsen értelmezve. Értelmezni csak az $x = 0$ helyen nem tudjuk.

Keressük meg azokat a helyeket, ahol a derivált eltűnik.

$$f'(x) = 1 + (x^{-1})' = 1 + x^{-2} = 1 - \frac{1}{x^2}$$

Meg kell oldanunk az $1 - \frac{1}{x^2} = 0$ egyenletet. $1 = \frac{1}{x^2}; x^2 = 1; x_1 = 1; x_2 = -1$

Az 1 és a (-1) lehetséges szélsőérték helyek, de ellenőrizni kell, hogy tényleg azok-e.

Célszerű egy számegyenesen feltüntetni a deriváltfüggvény zérushelyeit.

A deriváltfüggvény, tehát $f'(x) = 1 - \frac{1}{x^2}$ folytonos függvény, ezért a Bolzano-tételt használva nem vált előjelet a $(-\infty; -1) - en, a (-1; 0) - an, a (0; 1) - en és az (1; +\infty) - en sem$.

Az f függvény a $(-\infty; -1)$ -en szigorúan monoton növekvő, a $(-1; 0)$ -on szigorúan monoton csökkenő, ezért a (-1) lokális maximumhely.

Az f függvény a $(0; 1)$ -en szigorúan monoton csökkenő, az $(1; +\infty)$ -on szigorúan monoton növekvő, ezért az 1 lokális minimumhely.

Deriváltak alkalmazása a határértékszámításban

Sok esetben megkönnyíti a határértékek kiszámítását a deriváltak ismerete. Az ún. L'Hospital-szabályt azokban az esetekben tudjuk alkalmazni, amikor a meghatározandó határérték "nulla per nulla", vagy pedig abszolútértékben véve "végtelen per végtelen"-típusú.

TÉTEL (L'HOSPITAL-SZABÁLY): Legyenek f és g olyan függvények, amik differenciálhatók $I \setminus \{\alpha\}$ -on, ahol I valamilyen intervallum, $\alpha \in I$, és lehet akár $\alpha = \pm\infty$ is.

Tegyük fel, hogy $g(x) \neq 0$ semmilyen x -re, illetve $g'(x) \neq 0$ semmilyen x -re.

Ha

1. $\lim_{x \rightarrow \alpha} f(x) = \lim_{x \rightarrow \alpha} g(x) = 0$, vagy
2. $\lim_{x \rightarrow \alpha} |g(x)| = +\infty$, akkor a következő igaz:
amennyiben $\lim_{x \rightarrow \alpha} \frac{f'(x)}{g'(x)} = \beta$ létezik, akkor $\lim_{x \rightarrow \alpha} \frac{f(x)}{g(x)} = \beta$

Bizonyítás:

A bizonyítást először csak egy speciális esetben végezzük el, majd ez alapján áttérünk az általános eset bizonyítására.

I. (egyszerűbb) eset: $\alpha \in \mathbb{R}, f(\alpha) = g(\alpha) = 0$

Ekkor természetesen bármely x -re $\frac{f(x)}{g(x)} = \frac{f(x)-f(\alpha)}{g(x)-g(\alpha)}$.

Vegyük fel x -et és α -t úgy, hogy elég közel legyenek egymáshoz.

A Cauchy-féle középértéktétel szerint találunk α és x között egy olyan c -t, amire

$$\frac{f(x)-f(\alpha)}{g(x)-g(\alpha)} = \frac{f'(c)}{g'(c)}.$$

Legyen (x_k) egy α -hoz konvergáló sorozat.

A fenti észrevétel alapján x_k és α között mindig találunk olyan y_k számot, amire

$$\frac{f(x_k)-f(\alpha)}{g(x_k)-g(\alpha)} = \frac{f'(y_k)}{g'(y_k)}. \text{ A rendőrelv miatt persze az } (y_k) \text{ sorozat } \alpha\text{-hoz konvergál.}$$

A feltétel szerint, felhasználva az előbbi átalakítást is, kapjuk, hogy

$$\beta = \lim_{k \rightarrow \infty} \frac{f'(y_k)}{g'(y_k)} = \lim_{k \rightarrow \infty} \frac{f(x_k)-f(\alpha)}{g(x_k)-g(\alpha)} = \lim_{k \rightarrow \infty} \frac{f(x_k)}{g(x_k)} = \lim_{x \rightarrow \alpha} \frac{f(x)}{g(x)},$$

amit bizonyítani kellett.

II. (általános) eset: előfordulhat, hogy

- (1) f vagy g nincs értelmezve α -ban
- (2) értelmezve van ugyan, de valamelyik nem 0
- (3) $\alpha = \pm\infty$.

Az általános eset bizonyításában a fenti három lehetőség mindegyikét le fogjuk fedni.

Legyen most $x \neq y$. Végezzük el az alábbi átalakításokat:

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(x)-g(y)}{g(x)} + \frac{f(y)}{g(x)} = \\ &= \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} \end{aligned}$$

A Cauchy-középértéktétel szerint van x és y között olyan c , amire

$$\frac{f'(c)}{g'(c)} = \frac{f(x)-f(y)}{g(x)-g(y)}, \text{ azaz ezzel a } c\text{-vel kifejezve azt kapjuk, hogy}$$

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)} \cdot \left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}.$$

Legyen (x_k) egy α -hoz konvergáló sorozat, amelyben az α nem fordul elő elemként.

Megmutatjuk, hogy mindhárom esetben létezik olyan (y_k) , ami α -hoz konvergál, valamint egyszerre teljesíti a

$$\lim_{k \rightarrow \infty} \frac{g(y_k)}{g(x_k)} = 0, \text{ illetve } \lim_{k \rightarrow \infty} \frac{f(y_k)}{g(x_k)} = 0 \text{ feltételeket. Ebből már következni fog a tétel állítása,}$$

hiszen ez pontosan azt vonja maga után, hogy $\lim_{x \rightarrow \alpha} \frac{f(x)}{g(x)}$ és $\lim_{x \rightarrow \alpha} \frac{f'(x)}{g'(x)}$ határértékek megegyeznek.

Ha $\lim_{x \rightarrow \alpha} f(x) = 0$, $\lim_{x \rightarrow \alpha} g(x) = 0$, akkor legyen (x_k) egy α -hoz konvergáló sorozat, amelyben az α nem fordul elő elemként.

$$\text{Legyen most tehát } \lim_{k \rightarrow \infty} f(x_k) = \lim_{k \rightarrow \infty} g(x_k) = 0.$$

A szóban forgó sorozatok nullához tartása miatt tudunk választani egy olyan (y_k) részsorozatot az (x_k) sorozatból, amire tetszőleges k index esetén fennáll, hogy

$$\left| \frac{f(y_k)}{g(x_k)} \right| < \frac{1}{k}, \text{ illetve } \left| \frac{g(y_k)}{f(x_k)} \right| < \frac{1}{k}, \text{ és ezzel az állítást az első esetben igazoltuk.}$$

Amennyiben $\lim_{x \rightarrow \alpha} |g(x)| = +\infty$, akkor ugyanilyen módon járhatunk el, azaz a két hányados „kicsire” választható, és a szóban forgó határértékek megegyeznek, ahogy azt bizonyítani kellett. ■

Példák:

1. $\lim_{x \rightarrow +\infty} \frac{x^{100}}{e^x} = \lim_{x \rightarrow +\infty} \frac{100x^{99}}{e^x} = \dots = \lim_{x \rightarrow +\infty} \frac{100!}{e^x} = 0$
2. $\lim_{x \rightarrow 1} \frac{x^2+x-2}{5x^2-6x+1} = \lim_{x \rightarrow 1} \frac{2x+1}{10x-6} = \frac{3}{4}$

Primitív függvények (határozatlan integrálok)

DEFINÍCIÓ (PRIMITÍV FÜGGVÉNY): Legyen I valamilyen intervallum a valós számegyenesen.

Az $f: I \rightarrow \mathbf{R}$ függvénynek primitív függvénye az $F: I \rightarrow \mathbf{R}$, ha minden $x \in I$ esetén $F'(x) = f(x)$.

A definícióból egyből kitűnik, hogy primitív függvények keresésekor a deriváláshoz képest visszafelé kell gondolkodnunk. Nézzünk egy egyszerű példát. Az $f(x) = x$ -nek primitív függvénye

$$F(x) = \frac{x^2}{2}, \text{ vagy } F(x) = \frac{x^2}{2} + e, \text{ de akár } F(x) = \frac{x^2}{2} + \pi \text{ is.}$$

Természetes módon kerül fel az a kérdés, hogy miként írható le egy adott függvénynek az összes primitív függvénye (persze feltételezve azt, hogy létezik legalább egy primitív függvénye). Erre vonatkozik az alábbi tétel.

TÉTEL: Ha az $f: I \rightarrow \mathbf{R}$ függvénynek F_1 és F_2 is primitív függvénye, akkor $(F_1 - F_2)$ konstans függvény.

Bizonyítás:

Minden $x \in I$ -re $F_1'(x) = f(x) = F_2'(x)$, azaz ezekre az x -ekre $F_1'(x) = F_2'(x)$.

$F_1'(x) - F_2'(x) = 0$, ezért $(F_1(x) - F_2(x))' = 0$ minden x -re az adott intervallumon.

Viszont ekkor $(F_1(x) - F_2(x))$ konstans függvény, ahogyan az következett a Lagrange-középértéktételből (ld. az ottani következményt). ■

Tehát amennyiben az f függvény egy primitív függvénye az F , akkor az f függvény összes primitív függvényeinek halmaza az $(F+c)$ alakú függvényekből áll, ahol c konstans. Ezt a halmazt $\int f$, vagy a változó kiírása esetén $\int f(x) dx$ fogja jelölni, és ezeket nevezzük a f függvény **határozatlan integráljának**.

Primitív függvények keresésének néhány módszere:

1. **Parciális „integrálás”**: induljunk ki az $(f \cdot g)' = f' \cdot g + f \cdot g'$ differenciálási szorzásszabályból, és tekintsük az $\int (f \cdot g)'$ halmazt. Ez éppen az $(f \cdot g + konstans)$ alakú függvényekből áll. Ez alapján

$$f \cdot g + c = \int (f' \cdot g + f \cdot g')$$

A deriváltak összegéről szóló tétel miatt $\int (f' \cdot g + f \cdot g') = \int f' \cdot g + \int f \cdot g'$, amennyiben az itteni primitív függvények mind léteznek. Így formálisan azt kapjuk, hogy

$$f \cdot g + c = \int f' \cdot g + \int f \cdot g', \text{ azaz } f \cdot g - \int f \cdot g' = \int f' \cdot g$$

A most belátott képletet nevezzük a parciális integrálás képletének.

Nézzünk egy példát a parciális integrálás alkalmazására:

$$\int x \cdot e^x dx = \int (e^x)' \cdot x dx = e^x \cdot x - \int e^x \cdot (x)' dx = e^x \cdot x - \int e^x dx = e^x \cdot x - e^x + c$$

Természetesen nem mindegy, hogyan osztjuk ki a szerepeket. Ugyanis ha ebben a példában fordítva próbálkoztunk volna, akkor az alábbi kapjuk (ami az eredeti feladatnál „rosszabb”, hiszen az új integrál után másodfokú polinom van megszorozva egy exponenciálissal):

$$\int \left(\frac{x^2}{2}\right)' \cdot e^x = \frac{x^2}{2} \cdot e^x - \int \frac{x^2}{2} \cdot e^x dx.$$

2. **Helyettesítéses integrálás:**

Legyen $F(x)$ egy differenciálható függvény. Ekkor definíció szerint $\int F'(x) dx = F(x) + c$.

Írjuk most fel az x változót egy másik t változó függvényeként, azaz legyen $x=G(t)$, ahol a G függvényről feltesszük, hogy differenciálható, valamint létezik inverze is (természetesen $t = G^{-1}(x)$).

Ekkor az $F(G(t))$ és az $F(x)$ függvények formálisan egybeesnek. Deriváljuk most az $F(G(t))$ függvényt a „t” változó szerint. Ekkor a deriválás után megkapjuk az $F'(G(t)) \cdot G'(t)$ függvényt. Mivel az $F(G(t))$ és az $F(x)$ függvények formálisan egybeesnek, ezért

$$\int F'(G(t)) \cdot G'(t) dt = \int F'(x) dx. \text{ Ebből kapjuk, hogy}$$

$$F(x) = \int F'(G(t)) \cdot G'(t) dt, \text{ ahol } t = G^{-1}(x).$$

Az imént kapott formula a helyettesítéses integrálás képlete.

Nézzünk egy konkrét példát, amikor jól jön a helyettesítés. A példában a $\sqrt{1-x^2}$ függvény primitív függvényeit keressük. Közvetlenül persze nem tudnánk mit kezdeni a feladattal, de ha eszünkbe jut, hogy $0 \leq t \leq \pi/2$ esetén $\sqrt{1-\sin^2 t} = \cos t$, akkor $0 \leq x \leq 1$ esetén alkalmazhatjuk a $\sin t = x$ helyettesítést, és az alábbi módon találhatjuk meg a primitív függvényeket:

$$\int \sqrt{1-x^2} dx = \int \sqrt{1-\sin^2 t} \cdot \cos t dt = \int \cos t \cdot \cos t dt = \int \cos^2 t dt =$$

$$= \int \frac{1+\cos(2t)}{2} dt = \frac{1}{2}t + \frac{\sin(2t)}{4} + C = \frac{1}{2} \arcsin x + \frac{\sin(2\arcsin x)}{4} + c$$

Értelemszerűen ugyanígy járhatunk el, ha $-1 \leq x \leq 0$, ezt itt nem részletezzük.

A határozott integrál

Először Riemann adott erre pontos definíciót, de mi a **Darboux-féle integráldefiníciót** fogjuk használni.

A célunk a területfogalom pontos kialakítása.

DEFINÍCIÓ (DARBOUX-FÉLE INTEGRÁLFOGALOM): Legyen $f: [a, b] \rightarrow \mathbb{R}$. Osszuk fel az $[a, b]$ -ot véges sok részre, az osztópontok legyenek $a = x_0 < x_1 < x_2 < \dots < x_n = b$. Ezzel az $[a, b]$ -ot n db részre vágtuk fel. Minden kis részintervallumon határozzuk meg a függvény értékkészletének szuprémumát és ezt szorozzuk meg az aktuális intervallum hosszával (itt feltesszük, hogy f felülről korlátos).

Ezeket a szorzatokat adjuk össze, azaz tekintsük az adott beosztáshoz tartozó

$$S = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \sup f \Big|_{[x_{i-1}, x_i]}$$

összeget, amit az adott beosztáshoz tartozó **FELSŐ KÖZELÍTŐ ÖSSZEG**nek hívunk.

Hasonló módon definiáljuk az adott beosztáshoz tartozó **ALSÓ KÖZELÍTŐ ÖSSZEG**et, azaz:

$$s = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \inf f \Big|_{[x_{i-1}, x_i]}$$

A felső közelítő tehát olyan téglalapok területösszege, amiknek nincsen közös belső pontja, valamint uniójuk tartalmazza az egész görbe alatti tartományt. Az alsó közelítő összegek esetén szintén olyan téglalapok területeit adjuk össze, melyeknek nincsen közös belső pontjuk, csak éppen itt a téglalapok unióját tartalmazza a görbe alatti tartomány.

Az S -ek infimumát az f függvény **FELSŐ INTEGRÁL**jának, az s -ek szupremumának az f függvény **ALSÓ INTEGRÁL**jának fogjuk nevezni. Amennyiben ezek egybeesnek, azt fogjuk mondani, hogy f integrálható az $[a, b]$ -on és az integrál értéke megegyezik az alsó és a felső integrál közös értékével.

Tisztázandó az alsó és a felső integrál kapcsolata.

Észrevétel: Új osztópont beiktatásakor az alsó közelítő összeg nem csökkenhet, a felső közelítő összeg nem növekedhet.

Bizonyítás: A felső közelítő összeg esetében csak azon a kis részintervallumon kell vizsgálni a változást, ahová az új osztópont bekerült. Legyen $[r, s]$ az a részintervallum, ahová beraktunk egy új u osztópontot (r és s eredetileg szomszédos osztópontok voltak).

$$\text{A változás: } (u - r) \cdot \sup f \Big|_{[r, u]} + (s - u) \cdot \sup f \Big|_{[u, s]} - (s - r) \cdot \sup f \Big|_{[r, s]}$$

Mivel $[r, u]$ és $[u, s]$ részintervallumai $[r, s]$ -nek, ezért $\sup f \Big|_{[r, u]} \leq \sup f \Big|_{[r, s]}$, illetve

$$\sup f \Big|_{[u, s]} \leq \sup f \Big|_{[r, s]}.$$

A változás az előzőek miatt legfeljebb

$$(u - r) \cdot \sup f \Big|_{[r, s]} + (s - u) \cdot \sup f \Big|_{[r, s]} - (s - r) \cdot \sup f \Big|_{[r, s]} = \sup f \Big|_{[r, s]} ((u - r) + (s - u) - (s - r)) = 0$$

azaz a felső közelítő összeg nem növekedhetett.

Ugyanúgy jön ki, hogy az alsó közelítő összeg nem csökkenhet új osztópont beiktatásakor. ■

Állítás: Bármely felső közelítő összeg (azaz tetszőleges beosztáshoz tartozó) legalább akkora, mint egy tetszőleges (akár másik) beosztáshoz tartozó alsó közelítő összeg.

Bizonyítás: Vegyünk egy tetszőleges beosztást, az ehhez tartozó felső közelítő összeg legyen S^* . Vegyünk egy tetszőleges (akár másik) beosztást, az ehhez tartozó alsó közelítő összeg legyen s' . Most az előző két beosztásból csinálunk egy újat: összefésüljük őket, azaz vesszük azt a beosztást, ami az előző osztópontok uniójából áll. Az új összefésült beosztásban legyen a felső közelítő összeg S , akkor az alsó közelítő összeg s . Az előző észrevétel miatt $s \leq S^*$, illetve $s' \leq s$. Felhasználva, hogy $s \leq S$ kapjuk, hogy $s' \leq s \leq S \leq S^*$, innen $s' \leq S^*$. ■

Következmény: Az alsó integrál nem lehet nagyobb a felső integrálnál.

DEFINÍCIÓ: Azt mondjuk, hogy egy beosztás δ -nál finomabb, ha a szomszédos osztópontok távolsága kisebb δ -nál.

TÉTEL (DARBOUX): Legyen $f: [a, b] \rightarrow \mathbb{R}$ korlátos függvény. Ekkor bármely $\varepsilon > 0$ esetén létezik olyan $\delta > 0$, hogy az $[a, b]$ bármely δ -nál finomabb beosztása esetén:

1. az alsó közelítő összegek az alsó integráltól nem térnek el ε -nál jobban,
2. a felső közelítő összegek a felső integráltól nem térnek el ε -nál jobban.

Bizonyítás:

Elég a másodikat bizonyítani, mert az első ugyanúgy jön ki. Mivel a felső integrál (jelölje most F) a felső közelítő összegek infimuma, ezért nyilván van olyan beosztás, amire az teljesül, hogy a hozzá tartozó felső közelítő összeg „elég közel” van F -hez. Azaz tetszőleges $\varepsilon' > 0$ esetén van olyan beosztás, hogy az adott beosztáshoz tartozó felső közelítő összegre (amit most S jelöl) teljesül, hogy $F \leq S \leq F + \varepsilon'$. Legyen az ε' -höz egy ilyen beosztás $a = x_0 < x_1 < x_2 < \dots < x_n = b$. Jelölje ezt a beosztást B_1 .

Mivel a B_1 felosztásban véges sok kis intervallum van, ezért van köztük legrövidebb. Legyen $\delta' > 0$ olyan távolság, ami a legrövidebb kis intervallumnál is rövidebb. Tekintsünk egy δ' -nél finomabb beosztást, legyen ez a B_2 felosztás.

Azt próbáljuk megbecsülni, hogy ha kicseréljük a B_1 felosztást a B_2 felosztásra, akkor mennyit változik a felső közelítőösszegre. A cserét két lépésben hajtjuk végre. Először vesszük a B_1 és B_2 felosztások egyesítését, majd arról térünk át a B_2 felosztásra.

Az egyesített beosztás minden x_i pontot tartalmaz, ezért (ahogy azt korábban láttuk) a B_1 és B_2 egyesítéséhez tartozó felső közelítő összeg közelebb van F -hez, mint ε' . Most térjünk át az egyesített felosztásról a B_2 felosztásra. Ha a B_2 -ben szerepel az összes x_i pont, akkor már nincs dolgunk. Tehát feltételezhetjük, hogy van olyan i index, amelyre a B_2 beosztásban nem szerepel az x_i osztópontként. Jelölje r és s a B_2 beosztásban azon osztópontokat, amelyekre teljesül, hogy $r < x_i < s$.

Mivel a δ' távolság definíciónk szerint rövidebb a szomszédos x_j -k távolságánál, ezért az világos, hogy $x_{i-1} < r < x_i$, $x_i < s < x_{i+1}$, így tehát az $[r, s]$ intervallumon csak egyetlen olyan pont helyezkedik el, ami a B_1 beosztásban osztópont volt (és ez persze az x_i).

Most csak az $[r, s]$ -en nézzük a felső közelítő összeg változását, ha x_i -t kivesszük, r -et és s -t hozzáadjuk (tehát amikor az egyesített beosztásról áttérünk a B_2 -re).

A B_2 beosztásra nézve a felső közelítő összegnél az $[r, s]$ intervallumon a megfelelő szorzat egyenlő $(s - r) \cdot \sup f|_{[r, s]}$ -sel.

A B_1 és B_2 egyesítésében a megfelelő rész

$$(x_i - r) \cdot \sup f|_{[r, x_i]} + (s - x_i) \cdot \sup f|_{[x_i, s]}.$$

Becsüljük felülről most ezek eltérését (azt tudjuk, hogy az előző legalább akkora, mint az utóbbi, hiszen ezt már egy korábbi észrevételben beláttuk):

$$\begin{aligned} & (s - r) \cdot \sup f|_{[r, s]} - \left[(x_i - r) \cdot \sup f|_{[r, x_i]} + (s - x_i) \cdot \sup f|_{[x_i, s]} \right] \leq \\ & \leq (s - r) \cdot \sup f|_{[a, b]} - \left[(x_i - r) \cdot \inf f|_{[a, b]} + (s - x_i) \cdot \inf f|_{[a, b]} \right] = \\ & = (s - r) \cdot \left(\sup f|_{[a, b]} - \inf f|_{[a, b]} \right) < \delta' \cdot \left(\sup f|_{[a, b]} - \inf f|_{[a, b]} \right) \end{aligned}$$

Tehát ha δ' -t kicsire választjuk, akkor ez az eltérés akármilyen kicsi lehet. Ezt a típusú egyenlőtlenséget legfeljebb n -szer (azaz véges sokszor) kell biztosítanunk, ezért biztosan elérhető, hogy elég finom beosztások mindegyikére az F -től való eltérés kisebb legyen, mint

$\varepsilon' + n \cdot \delta' \cdot \left(\sup f \Big|_{[a, b]} - \inf f \Big|_{[a, b]} \right)$. Itt ε' tetszőlegesen kicsi távolság lehet, az n és a

$\left(\sup f \Big|_{[a, b]} - \inf f \Big|_{[a, b]} \right)$ rögzítve vannak, a δ' pedig szintén akármilyen kicsi pozitív távolság lehet (az n értéke persze függött ε' -től, de ez nem baj, mert a δ' értékét utána szabadon választottuk meg). ■

DEFINÍCIÓ (OSZCILLÁCIÓS ÖSSZEG): Adott beosztáshoz tartozó felső és alsó közelítő összeg különbségét az adott beosztáshoz tartozó oszcillációs összegnek nevezzük.

A Darboux-tétel felhasználásával közvetlenül adódik az alábbi tétel:

TÉTEL: Egy korlátos $f: [a, b] \rightarrow \mathbb{R}$ függvény pontosan akkor (Darboux-) integrálható, ha bármely $\varepsilon > 0$ esetén van olyan $\delta > 0$, hogy az $[a, b]$ bármilyen δ -nál finomabb beosztása esetén az oszcillációs összeg kisebb ε -nál.

TÉTEL (NEWTON-LEIBNIZ): Ha $f: [a, b] \rightarrow \mathbb{R}$ korlátos függvény integrálható $[a, b]$ -on és ugyanitt primitív függvénye is létezik (mondjuk legyen egy primitív függvénye $F(x)$), akkor f -nek az $[a, b]$ -on vett integrálja éppen $F(b) - F(a)$.

Bizonyítás:

Legyen az $[a, b]$ egy felosztása $a = x_0 < x_1 < x_2 < \dots < x_n = b$.

A Lagrange-középtértéktételt alkalmazzuk minden kis intervallumon: minden nyílt (x_{i-1}, x_i) intervallumon van olyan y_i , amire $F'(y_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}$.

Felhasználva, hogy F primitívfüggvénye f -nek, és szorozva mindkét oldalt a részintervallum hosszával, kapjuk, hogy $f(y_i) \cdot (x_i - x_{i-1}) = F(x_i) - F(x_{i-1})$. Adjuk össze ezeket az egyenlőségeket minden lehetséges i -re:

$$\begin{aligned} \sum_{i=1}^n f(y_i) (x_i - x_{i-1}) &= \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = \\ &= F(x_1) - F(x_0) + F(x_2) - F(x_1) + F(x_3) - F(x_2) + \dots + F(x_n) - F(x_{n-1}) = \\ &= F(x_n) - F(x_0) = F(b) - F(a) \end{aligned}$$

A bal oldalon álló kifejezést alulról és felülről becsljük:

$$\sum_{i=1}^n \inf f \Big|_{[x_{i-1}, x_i]} \cdot (x_i - x_{i-1}) \leq \sum_{i=1}^n f(y_i) (x_i - x_{i-1}) \leq \sum_{i=1}^n \sup f \Big|_{[x_{i-1}, x_i]} \cdot (x_i - x_{i-1})$$

Mivel feltettük, hogy f integrálható, ezért az alsó közelítő összegek szuprémuma megegyezik a felső közelítő összegek infimumával, ez a közös érték a rendőrelv miatt $F(b) - F(a)$, amit bizonyítani kellett. ■

Néhány megjegyzés a Newton-Leibniz-tételhez:

1. Ez a tétel valamilyen kapcsolatot teremtett a határozott és a határozatlan integrál között.
2. A tételben külön nem emeltük ki, de igaz, hogy teljesen mindegy, melyik primitív függvényt választjuk az integrál kiszámításához, hiszen $(F(b) + c) - (F(a) + c) = F(b) - F(a)$.

A kör területe

Már megvannak az eszközeink ahhoz, hogy a kör területképletét felírjuk. Ha a $\sqrt{1-x^2}$ függvényt integráljuk a $[0,1]$ zárt intervallumon, akkor megkapjuk az egységsugarú kör egynegyedének területét.

Felhasználva az ezen függvény primitív függvényére vonatkozóan korábban levezetett formulát, valamint a Newton-Leibniz-tételt, az alábbiakat kapjuk:

$$\int_0^1 \sqrt{1-x^2} = \frac{1}{2} \arcsin 1 + \frac{\sin(2\arcsin 1)}{4} - \frac{1}{2} \arcsin 0 - \frac{\sin(2\arcsin 0)}{4} = \frac{\pi}{4} + 0 - 0 - 0 = \frac{\pi}{4}.$$

Ez alapján az egységsugarú kör területe π -vel egyenlő. Ebből kapjuk, hogy az r sugarú kör területe $r^2\pi$ -vel egyenlő. Ez következik a hasonlóságból is, illetve úgy is kiszámolható, hogy a $\sqrt{r^2-x^2}$ függvényt integráljuk a $[0,r]$ -en (ugyanúgy járhatunk el, mint a fenti esetben, csak most pl. az $x=r \cdot \sin t$ helyettesítést célszerű elvégeznünk).

Görbék ívhossza

Legyen adva egy görbe, ami leírható a koordinátarendszerben az $f(x)$ függvénnyel egy $[a,b]$ zárt intervallumon. Próbáljuk meghatározni ezen görbe ívhosszát. Ehhez osszuk fel az $[a,b]$ zárt intervallumot véges sok osztóponttal, legyen egy ilyen felosztás

$$a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

Az ezen felosztáshoz tartozó pontok a görbén $(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_n, f(x_n))$.

Írjuk fel annak a töröttvonalnak a hosszát, amit ezek a pontok a görbén meghatároznak. Ez a Pithagorasz-tétel miatt éppen

$$\sum_{k=1}^n \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2}.$$

Ez átalakítható úgy, hogy

$$\sum_{k=1}^n \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2} =$$

$$= \sum_{k=1}^n (x_k - x_{k-1}) \cdot \sqrt{1 + \left(\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \right)^2}.$$

A Lagrange-középértéktétel miatt minden egyes $[x_{k-1}, x_k]$ intervallumon létezik olyan y_k pont, amelyre az teljesül, hogy $\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(y_k)$.

Az ilyen y_k pontokkal tehát a töröttvonal hossza, ami közelíti a görbe ívhosszát, éppen

$$\sum_{k=1}^n (x_k - x_{k-1}) \cdot \sqrt{1 + (f'(y_k))^2}.$$

Tegyük fel azt, hogy $f(x)$ differenciálható, továbbá azt is, hogy $\sqrt{1 + (f'(x))^2}$ integrálható $[a, b]$ -on.

A $\sum_{k=1}^n (x_k - x_{k-1}) \cdot \sqrt{1 + (f'(y_k))^2}$ összeg biztosan $\sqrt{1 + (f'(x))^2}$ -nek az adott beosztáshoz

tartozó alsó és felső (integrálra vonatkozó) közelítőösszege közé esik, ezért a beosztás

finomításával ennek létezik határértéke, ami szükségképpen egyenlő $\int_a^b \sqrt{1 + (f'(x))^2} dx$ -szel.

Ezzel tehát kaptunk egy kész formulát görbék ívhosszára vonatkozóan (amiknek néhány tulajdonságát differenciálhatóság és integrálhatóság tekintetében feltételeztük).

A határozott integrál további néhány tulajdonsága, integrálfüggvények

TÉTEL: Az $f: [a, b] \rightarrow \mathbb{R}$ monoton függvény integrálható.

Bizonyítás:

Elég ezt abban az esetben bizonyítani, amikor a függvény monoton növekedő (csökkenő esetben ugyanúgy bizonyítható).

A szokásos módon legyenek $a = x_0 < x_1 < x_2 < \dots < x_n = b$ osztópontok.

A felső közelítő összeg $\sum_{i=1}^n f(x_i) \cdot (x_i - x_{i-1})$, az alsó közelítő összeg $\sum_{i=1}^n f(x_{i-1}) \cdot (x_i - x_{i-1})$, hiszen monoton növekedő a függvény.

Az oszcillációs összeg $\sum_{i=1}^n (f(x_i) - f(x_{i-1})) \cdot (x_i - x_{i-1})$.

Ha a beosztás δ -nál finomabb, akkor

$$\sum_{i=1}^n (f(x_i) - f(x_{i-1})) \cdot (x_i - x_{i-1}) < \sum_{i=1}^n (f(x_i) - f(x_{i-1})) \cdot \delta =$$

$$= \delta \cdot (f(x_1) - f(x_0) + f(x_2) - f(x_1) + f(x_3) - f(x_2) + \dots + f(x_n) - f(x_{n-1})) =$$

$= \delta \cdot (f(x_n) - f(x_0)) = \delta \cdot (f(b) - f(a))$. Mivel δ akármilyen kicsi pozitív távolság lehet, ezért az állítás adódik. ■

Szeretnénk megmutatni, hogy a zárt intervallumon folytonos függvények integrálhatók is ezen az intervallumon. Először belátjuk az alábbi segédteét:

Segédteét (Heine): Ha $f: [a, b] \rightarrow \mathbb{R}$ folytonos, akkor bármely $\varepsilon > 0$ esetén létezik olyan $\delta > 0$, hogy minden olyan $x, y \in [a, b]$ -beli elemekre, melyek távolsága $|x - y| < \delta$, az is teljesül, hogy

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

Bizonyítás (segédteét): Indirekt bizonyítunk. Tegyük fel, hogy van olyan $\varepsilon > 0$, amihez nem találunk „jó” $\delta > 0$ számot, azaz van olyan $\varepsilon > 0$, hogy bármely $\delta > 0$ esetén találunk olyan $x, y \in [a, b]$ elemeket, melyekre $|x - y| < \delta$ és $|f(x) - f(y)| \geq \varepsilon$.

A δ fusson végig az $\frac{1}{n}$ sorozat elemein. Tehát bármely n -re van olyan $x_n, y_n \in [a, b]$, hogy $|x_n - y_n| < \frac{1}{n}$ és $|f(x_n) - f(y_n)| \geq \varepsilon$.

A Bolzano-Weierstrass-tétel miatt (zárt intervallumon vagyunk) az (x_n) sorozatnak van konvergens részsorozata, legyen ez $z_n \rightarrow u \in [a, b]$. Az (y_n) sorozatnak azok az elemei, amik a (z_n) sorozat elemeihez tartoznak, ugyanúgy ehhez az u -hoz tartanak, legyen y_n ez a részsorozata $t_n \rightarrow u$. Ekkor a feltétel miatt $|f(z_n) - f(t_n)| \geq \varepsilon$ minden n -re, ugyanakkor a folytonosság miatt $f(z_n) \rightarrow f(u)$, $f(t_n) \rightarrow f(u)$, azaz $|f(z_n) - f(t_n)| \rightarrow 0$, ez ellentmondás. ■

1. Megjegyzés: A Heine-tételt másképp úgy szokták megfogalmazni, hogy zárt intervallumon folytonos függvény szükségképpen **egyenletesen folytonos**. Azaz az egyenletes folytonosság a következőt jelenti: Az f függvény egyenletesen folytonos az I intervallumon, ha minden $\varepsilon > 0$ esetén létezik olyan $\delta > 0$, hogy bármely $x, y \in I$ -re, amelyek távolsága $|x - y| < \delta$, az is teljesül, hogy $|f(x) - f(y)| < \varepsilon$.

2. Megjegyzés: Az egyenletes folytonosság erősebb feltételt jelent, mint a folytonosság. Az egyenletes folytonosságból mindig következik a folytonosság, de fordítva ez nem feltétlenül igaz.

TÉTEL: Ha $f: [a, b] \rightarrow \mathbb{R}$ folytonos, akkor integrálható is $[a, b]$ -on.

Bizonyítás: A Heine-tétel miatt egy ilyen f függvény egyenletesen folytonos is. Legyen $\varepsilon > 0$. Ehhez van olyan $\delta > 0$, hogy minden $x, y \in [a, b]$ -re, amik eleget tesznek $|x - y| < \delta$ feltételnek, az is igaz, hogy $|f(x) - f(y)| < \varepsilon$.

Vegyük az $[a, b]$ intervallum egy δ -nál finomabb beosztását (véges sok osztóponttal). Legyen ez most $a = x_0 < x_1 < x_2 < \dots < x_n = b$. Ekkor minden kis $[x_{i-1}, x_i]$ részintervallumon teljesül, hogy $s, t \in [x_{i-1}, x_i]$ -re $|f(s) - f(t)| < \varepsilon$. Ez alapján adunk felső becslést az oszcillációs összegre. Az előző miatt minden kis $[x_{i-1}, x_i]$ intervallumon teljesül, hogy

$$\sup f \Big|_{[x_{i-1}, x_i]} - \inf f \Big|_{[x_{i-1}, x_i]} < 2\varepsilon.$$

Az oszcillációs összeg értéke így legfeljebb $2\varepsilon \cdot \sum_{i=1}^n (x_i - x_{i-1}) = 2\varepsilon(b - a)$. Mivel az ε -t akármilyen kicsire választhatjuk, az állítás adódik. ■

Fontos észrevétel: Ha f függvény integrálható $[a, b]$ -on, akkor ennek minden részintervallumán is integrálható.

Bizonyítás: (Lényegében rendőrelvet használunk.) Ha $[a, b]$ -n veszünk egy osztópontsorozatot, akkor tetszőleges $c, d \in [a, b]$ -re (ahol $c < d$) az $[a, b]$ -re vonatkozó teljes oszcillációs összeg éppen az $[a, c]$ -re, $[c, d]$ -re, illetve $[c, b]$ -re vonatkozó oszcillációs összegek összege (egy adott felosztás mellett). Elég finomra választva a beosztást, az $[a, b]$ -re vonatkozó oszcillációs összeg kicsi lesz, hiszen a feltétel szerint f integrálható az $[a, b]$ intervallumon. Ekkor az $[a, c]$, $[c, d]$, $[d, b]$ -re vett oszcillációs összegek még kisebbek, hiszen az oszcillációs összeg értéke soha nem lehet negatív. ■

Az észrevétünket felhasználva bevezethető az **integrálfüggvény** fogalma:

DEFINÍCIÓ (INTEGRÁLFÜGGVÉNYEK HALMAZA): Legyen f az $[a, b]$ -n integrálható függvény. Ekkor minden $x \in [a, b]$ -re definiálható az $\int_a^x f(t)dt =: F(x)$, azaz $F(x)$ az f integrálja az $[a, x]$ -en. Az $f(x)$ függvény integrálfüggvényeinek halmaza álljon az $F(x)+c$ alakú függvényekből, ahol c valamilyen konstans.

1. Megjegyzés: Az integrálfüggvények az észrevételünk miatt voltak bevezethetők.

2. Megjegyzés: $F(a) = \int_a^a f(t)dt = 0$.

Jelölés: $\int_r^s f(y)dy$ azt jelenti, hogy az y változós f függvényt integráljuk az $[r, s]$ -on.

Mielőtt továbbmegyünk, belátjuk az alábbi tételt:

Tétel:

(i) Ha $c \in \mathbb{R}$, akkor $c \cdot \int_a^b f(x)dx = \int_a^b c \cdot f(x)dx$.

(ii) Ha f integrálható $[a, b]$ -on és $a \leq c \leq b$, akkor $\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx$.

Az (i) bizonyítása:

A konvergencia szorzásra vonatkozó tulajdonságaiból egyszerűen következik.

Az (ii) bizonyítása:

Ha az f függvény integrálható $[a, b]$ -on, akkor tudjuk, hogy ennek bármely részintervallumán is integrálható, azaz $[a, c]$ -on, illetve $[c, b]$ -on is. Vegyük egy olyan beosztását az $[a, b]$

intervallumnak, amiben c nem osztópont. Ehhez a beosztáshoz a c pontot hozzáadva megkapjuk az $[a,c]$, illetve $[c,b]$ intervallumok egy-egy felosztásának egyesítését.

Elég lenne tehát annyit meggondolni, hogy mennyit változik az oszcillációs összeg, ha a c pontot elhagyjuk.

A Darboux-tétel bizonyítása során már láttuk, hogy egy osztópont elhagyásakor ez a változás

legfeljebb $\delta' \cdot \left(\sup f \Big|_{[a,b]} - \inf f \Big|_{[a,b]} \right)$ lesz, ahol δ' -nél finomabb a c -től különböző

osztópontok sűrűsége. Ebből az állítás egyből következik, hiszen δ' értéke tetszőlegesen választható. ■

Az imént bizonyított tétel additivitási része miatt az integrálfüggvény az alábbi módon is jellemezhető:

Az integrálfüggvény definíciójának átfogalmazása: Legyen f az $[a,b]$ -on integrálható függvény. Az ugyanitt értelmezett G függvény pontosan akkor integrálfüggvénye f -nek, ha tetszőleges $a \leq c \leq d \leq b$ esetén $G(d)-G(c) = \int_c^d f(x)dx$.

Bizonyítás:

Két irányt kell vizsgálnunk. Ha G integrálfüggvénye f -nek, akkor felhasználva az integrál intervallum szerinti additivitását, azt kapjuk, hogy

$$G(d)-G(c) = \int_a^d f(x)dx - \int_a^c f(x)dx = \int_c^d f(x)dx.$$

Most vizsgáljuk a másik irányt. Legyen $G(a) = c$. Tekintsük az $F(x) = G(x)-c$ függvényt. Ekkor persze $F(a) = 0 = \int_a^a f(x)dx$.

A feltétel szerint tetszőleges $[a,b]$ -beli x -re

$F(x)-F(a) = (G(x)-c) - (G(a)-c) = G(x) - G(a) = \int_a^x f(t)dt$. Ezért $G(x)$ valóban integrálfüggvénye f -nek. ■

Többször is hasznos lehet az integrál értékére vonatkozó alábbi becslés:

Tétel: Legyen $f: [a,b] \rightarrow \mathbb{R}$ integrálható függvény. Ekkor $\inf f \Big|_{[a,b]} \leq \frac{\int_a^b f(x)dx}{b-a} \leq \sup f \Big|_{[a,b]}$.

Bizonyítás:

Tetszőleges $a = x_0 < x_1 < x_2 < \dots < x_n = b$ beosztás esetén az alsó közelítőösszeget alulról

becsüljük, ha mindegyik téglalap magasságát kicseréljük $\inf f \Big|_{[a,b]}$ -re.

Hasonlóan felső becslést kapunk a felső közelítőösszegre, ha mindegyik téglalap magasságát egyszerűen $\sup f \Big|_{[a,b]}$ -ra cseréljük.

Ebből formálisan az alábbi kapjuk:

$$(b-a) \cdot \inf_{[a,b]} f = \sum_{i=1}^n \inf_{[a,b]} f \cdot (x_i - x_{i-1}) \leq \text{alsó közelítőösszeg} \leq \int_a^b f(x) dx \leq \text{felső közelítőösszeg} \leq \sum_{i=1}^n \sup_{[a,b]} f \cdot (x_i - x_{i-1}) = (b-a) \cdot \sup_{[a,b]} f$$

Leosztva az $[a,b]$ intervallum hosszával éppen a bizonyítandó állítást kapjuk. ■

Készen állunk arra, hogy ismertessük az integrálfüggvény néhány egyszerű tulajdonságát.

TÉTEL (AZ INTEGRÁLFÜGGVÉNY TULAJDONSÁGAI): Ha $f(x)$ korlátos függvény integrálható $[a,b]$ -on, és $F(x) = \int_a^x f(t)dt$, akkor az alábbiak érvényesek:

- (i) A $F(x)$ folytonos minden $x \in [a,b]$ -re.
- (ii) Minden olyan $u \in [a,b]$ pontban, ahol f folytonos, ott F differenciálható is és $F'(u) = f(u)$.

Az (i) bizonyítása:

Legyen $u \in [a,b]$, és $x_n \rightarrow u$.

$F(x_n) - F(u) = \int_a^{x_n} f(t)dt \rightarrow 0$, mert f függvény korlátossága miatt az itteni oszcillációs összeg is 0-hoz tart, hiszen az oszcillációs összeg $< |x_n - u| \cdot 2K$, ahol $K > |f(x)|$ minden $x \in [a,b]$ -re.

A (ii) bizonyítása:

Tudjuk, hogy az f folytonos az $u \in [a,b]$ pontban. Az alábbi kifejezést kell vizsgálnunk, ha $x \rightarrow u$:

$$\frac{F(x) - F(u)}{x - u} = \frac{\int_u^x f(t)dt}{x - u}$$

Legyen tehát $x_n \rightarrow u$. Ekkor már tudunk becslést $\frac{\int_u^{x_n} f(t)dt}{x_n - u}$ - re vonatkozóan (ezt korábban beláttuk):

$$\inf_{[u, x_n]} f \leq \frac{\int_u^{x_n} f(t)dt}{x_n - u} \leq \sup_{[u, x_n]} f$$

Az f függvény u -beli folytonossága miatt teljesül, hogy $x_n \rightarrow u$ esetén

$$\inf_{[u, x_n]} f \rightarrow f(u), \text{ illetve } \sup_{[u, x_n]} f \rightarrow f(u).$$

A rendőrelv miatt innen $\lim_{x_n \rightarrow u} \frac{F(x_n) - F(u)}{x_n - u} = f(u)$, azaz $F'(u) = f(u)$. ■

TÉTEL: Ha $f(x)$ folytonos az I zárt intervallumon, akkor itt van primitívfüggvénye, és a primitívfüggvényei egybeesnek az ugyanezen az intervallumon definiált integrálfüggvényeivel.

Bizonyítás: Azt már beláttuk, hogy egy zárt intervallumon értelmezett folytonos függvény ugyanott integrálható is, sőt a zárt intervallum bármely részintervallumán integrálható. Ez tehát azt jelenti, hogy ilyenkor lesz integrálfüggvénye az $f(x)$ -nek az I intervallumon, azaz olyan $G(x)$ függvény, hogy minden $a, b \in I$ -re $\int_a^b f(x)dx = G(b) - G(a)$. Szintén láttuk már, hogy folytonos függvény esetén $G'(x) = f(x)$, azaz $G(x)$ primitív függvénye lesz a folytonos f függvénynek. Mivel az integrálfüggvények egymáshoz képest konstanssal vannak eltolva (mint ahogy a primitívfüggvények), ezért a két halmaz biztosan egybeesik. ■

Kettős integrál, többszörös integrálok

Először csak olyan függvényekkel foglalkozunk, amik a sík pontjainak egy bizonyos részhalmazán értelmezettek és a felvett értékek halmaza valós számokból áll. Sőt, először csak téglalap alakú tartományon értelmezett ilyen típusú függvényeket vizsgálunk.

Vegyünk egy f korlátos függvényt, ami értelmezve van az $a \leq x \leq b$, $a' \leq y \leq b'$ pontok által kifeszített téglalap alakú tartományon.

Felosztjuk a téglalapot véges sok kis olyan téglalapra, melyeknek minden oldala párhuzamos valamelyik tengellyel.

Legyenek a kis téglalapok területei $t_1, t_2, \dots, t_n$ (tehát összesen n db téglalap van). A Darboux-integrálhoz hasonló módon készítjük el az alsó és felső közelítő összegeket. Az adott felosztáshoz tartozó alsó közelítő összeg

$$\sum_{i=1}^n t_i \cdot \inf f \Big|_{\text{az } i - \text{edik téglalapra nézve}}$$

a felső pedig

$$\sum_{i=1}^n t_i \cdot \sup f \Big|_{\text{az } i - \text{edik téglalapra nézve}}$$

Felosztásoknak egy olyan sorozatát vesszük, ahol a kis téglalapok közül a maximális átmérőjű (átlójú) átlójának hossza a nullához tart.

Hasonlóan az egyváltozós esethez az alsó közelítő összegeknek lesz egy szuprénuma, ami az alsó integrál, és a felső közelítő összegeknek lesz egy infimuma, ami a felső integrál.

Ugyanúgy, mint az egyváltozós esetben is láttuk, ha az alsó és a felső integrál értéke megegyezik, akkor a függvény téglalapon vett integrálja ez a közös érték.

Megjegyzendő, hogy az egyváltozós esethez hasonlóan itt is igaz lesz, hogy amennyiben sűrítjük a beosztást (tehát a felosztásban szereplő téglalapokat darabolunk tovább), akkor az alsó közelítőösszeg nem csökken, a felső pedig nem növekszik. Ezt hasonlóan bizonyíthatjuk, mint ahogy ott láttuk.

Most teszünk ezzel kapcsolatban egy egyszerű észrevételt. Legyen adott egy téglalap, ami szerint integrálunk. Tekintsük most ezen téglalap egy tetszőleges, kisebb téglalapokra való felosztását. Ezt a felosztást bővítsük úgy, hogy a felbontásbeli téglalapok vízszintes és függőleges oldalegyeneseit is meghosszabbítjuk, és térjünk át a kapott konfiguráció szerinti felosztásra. Ezáltal egy speciálisabb típusú felosztást nyerünk, amiben vízszintes és függőleges egyenesek metszéspontjai határozzák meg a téglalapokat, de az előző észrevétel miatt az ilyen speciálisabb felosztásokhoz tartozó alsó, illetve felső közelítőösszegeket is elég vizsgálnunk az integrálhatóság szempontjából.

A kettős integrál kiterjesztése egyszeresen összefüggő tartományokra („pacnikra”):

Egy ilyen tartományt körbeveszünk egy téglalappal. Ha az f kétváltozós függvény értelmezve volt a pacnin, akkor elkészítjük a téglalapon értelmezett F függvényt az alábbi módon:

$F(x, y) = f(x, y)$, ha (x, y) pont rajta van a pacnin, és 0 egyébként.

Ekkor az F függvény téglalapon vett integrálja éppen az f függvény pacnin vett integrálja.

Jelölés: ha ω síkbeli egyszeresen összefüggő tartomány („pacni”), és f egy kétváltozós, valós értékű, ω -on integrálható függvény, akkor az adott tartományon vett integrált $\iint_{\omega} f(x, y)$ jelöli.

Bizonyítás nélkül közöljük az alábbi tételt:

TÉTEL: Ha D téglalap alakú tartomány, $f: D \rightarrow \mathbf{R}$ folytonos, kivéve esetleg olyan helyeket, amik lefedhetők véges sok folytonos görbével, akkor f integrálható D -n.

Szukcesszív integrálás

Csak bizonyos alakú tartományokon működő módszer, az a lényege, hogy egyváltozós integrálásokra vezetjük vissza a kettős (vagy akár a többszörös) integrálokat.

TÉTEL: Tegyük fel, hogy D olyan téglalap alakú tartomány, ahol $f: D \rightarrow \mathbf{R}$ integrálható.

A D téglalap legyen az $a \leq x \leq b$, $a' \leq y \leq b'$ pontok által meghatározva. Tegyük fel, hogy bármely rögzített x -re az $f(x, y)$ az y -tól függő egyváltozós függvény is integrálható.

Ekkor:

1. $F(x) = \int_{a'}^{b'} f(x, y) dy$ integrálható $[a, b]$ -on.
2. $\iint_D f(x, y) = \int_a^b \left(\int_{a'}^{b'} f(x, y) dy \right) dx$

Példa:

$$f(x, y) = x^2 y$$

$$\iint_D f(x, y) = \int_1^2 \left(\int_2^4 x^2 y \, dy \right) dx$$

$$\int_1^2 \left(x^2 \left[\frac{y^2}{2} \right]_2^4 \right) dx == \int_1^2 \left(x^2 \cdot \left(\frac{16}{2} - \frac{4}{2} \right) \right) dx == \int_1^2 6x^2 \, dx = 6 \cdot \left[\frac{x^3}{3} \right]_1^2 == 6 \cdot \left(\frac{2^3}{3} - \frac{1^3}{3} \right) = 6 \cdot \frac{7}{3} = 14$$

Bizonyítás:

Legyen az $[a, b]$ egy felosztása $a = x_0 < x_1 < \dots < x_n = b$, és legyen $[a', b']$ egy felosztása $a' = y_0 < y_1 < \dots < y_m = b'$.

Ezekkel a osztópontokkal a nagy téglalapot kis téglalapokra bontottuk, legyenek ezek a téglalapok a D_{ij} téglalapok (tehát D_{ij} az i -edik sor j -edik téglalapja). Tudjuk, hogy az integrálhatóság szempontjából elég csak az ilyen speciális típusú felbontásokat vizsgálnunk (lsd. az ehhez kapcsolódó korábbi megjegyzésünket).

Legyen $x_{i-1} \leq \varphi_i \leq x_i$, azaz minden egyes $[x_{i-1}, x_i]$ intervallumon vegyünk fel egy belső φ_i pontot.

A D_{ij} kis téglalapon az f infimuma legyen m_{ij} , szuprémuma pedig M_{ij} .

Az integrálokra vonatkozó korábbi egyszerű becslésünk alapján tudjuk azt, hogy

$$m_{ij}(y_j - y_{j-1}) \leq \int_{y_{j-1}}^{y_j} f(\varphi_i, y) \, dy \leq M_{ij}(y_j - y_{j-1}).$$

Az integrál intervallumonkénti additivitása miatt az is világos, hogy

$$F(\varphi_i) = \int_{a'}^{b'} f(\varphi_i, y) \, dy = \sum_{j=1}^m \left(\int_{y_{j-1}}^{y_j} f(\varphi_i, y) \, dy \right)$$

Ezek alapján kapjuk, hogy $\sum_{j=1}^m m_{ij}(y_j - y_{j-1}) \leq F(\varphi_i) \leq \sum_{j=1}^m M_{ij}(y_j - y_{j-1})$

A kapott egyenlőtlenséget felszorozzuk $(x_i - x_{i-1})$ -gyel, majd i szerint is összegzünk:

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^m m_{ij}(y_j - y_{j-1})(x_i - x_{i-1}) &\leq \sum_{i=1}^n F(\varphi_i) (x_i - x_{i-1}) \leq \\ &\leq \sum_{i=1}^n \sum_{j=1}^m M_{ij}(y_j - y_{j-1})(x_i - x_{i-1}). \end{aligned}$$

Figyeljük meg, hogy itt a $\sum_{i=1}^n F(\varphi_i) (x_i - x_{i-1})$ egy integrálközelítő összeg (éppen az $F(x)$ integrálját közelíti az $[a, b]$ intervallumon), a két szélső kifejezés pedig az $\iint_D f(x, y)$ integrál alsó, illetve felső közelítő összege. Mivel az utóbbi integrál a feltételek szerint létezik, ezért az állítás adódik. ■

A szukcesszív integrálás egy általánosítása: ún. **normáltartományokon** integrálunk. Egy normáltartományt két y -tengellyel párhuzamos egyenes és olyan $\psi_1(x)$ és $\psi_2(x)$ folytonos görbék határoznak meg, melyekre $\psi_1(x) < \psi_2(x)$ a megfelelő x -ek mellett (nyilván tengelycserénél csak a változók szerepe cserélődik).

Legyen ω a kiszemelt normálttartomány, $f: \omega \rightarrow \mathbf{R}$ függvény, ami integrálható ω -on. Foglalkozunk vele ω -t egy D téglalapba. Legyen F az f függvény olyan kiterjesztése, hogy

$F(x, y) = f(x, y)$, ha (x, y) pont rajta van a pacnin, és 0 egyébként.

Ha F integrálható D -n, akkor megfelelő feltételek mellett szukcesszívan is integrálható.

$$\iint_{\omega} f = \iint_{\omega} F = \iint_D F = \int_r^s \left(\int_{a'}^{b'} F(x, y) dy \right) dx = \int_a^b \left(\int_{\psi_1(x)}^{\psi_2(x)} f(x, y) dy \right) dx$$

Megjegyzés: Ugyanezek hasonlóan megfogalmazhatók hármass integrálokra, sőt többszörös integrálokra (a felvett értékek valósak).

Alkalmazás:

1. Ha az f függvény az azonosan 1 függvény (identitás), akkor a D lemezen $\iint_D f$ a terület, a V testre nézve $\iiint_V f$ a V térfogata.
2. Vékony H lemez tömege: $m = \iint_H \rho(x, y)$, ahol ρ a sűrűséget jelöli.
3. Test tömege (V a test): $m = \iiint_V \rho(x, y, z)$
4. Vékony H lemez súlypontjának koordinátái: $s_x = \frac{\iint_H x \cdot \rho(x, y)}{m}$, $s_y = \frac{\iint_H y \cdot \rho(x, y)}{m}$.
5. Ugyanígy a test súlypontja.

Vonalintegrál

Olyan görbe mentén fogunk integrálni, amelynek van véges ívhossza (ún. **rektifikálható görbe**). A görbék, amik mentén az integrálás történik, valamilyen folytonos térgörbék lesznek.

A függvények, amiket ilyen térgörbék mentén integrálni szeretnénk, ún. **vektormezők** lesznek, azaz olyan függvények, amelyek R^3 egy bizonyos részhalmazán értelmezettek és értékkészletük is R^3 -beli pontokból áll.

A vektormezők egy speciális részhalmazát alkotják az ún. **skalármezők**, ezek olyan függvények, melyek értelmezési tartománya szintén R^3 egy részhalmaza, de az értékkészlet valós számokból áll.

A térgörbékét általában konkrét feladatoknál paraméteres alakban adjuk meg, azaz egy változó függvényeként (valamikor nem elég egy). Emellett a térgörbéknek kell adnunk egy irányítást is (nem mindegy, hogy melyik a kezdőpont, és melyik a végpont).

Pl. egy egységnyi sugarú körben egy negyed körívnek egy paraméterezése lehet $[\cos t, \sin t, 0]$, ahol $t \in \left[0, \frac{\pi}{2}\right]$.

Definíció (vonalintegrál): Legyen $G \subseteq R^3$ és $\underline{v}: G \rightarrow R^3$ vektormező. Továbbá legyen C egy olyan görbe, ami G -ben rektifikálható, és legyen ennek a görbének egy paraméterezése $\underline{r}: [a, b] \rightarrow G$, ami szigorúan monoton módon ugyanolyan bejárási irányban paraméterez (itt a és b valós számok).

Vegyük az $[a, b]$ egy felosztását a szokásos módon, $a = t_0 < t_1 < t_2 < \dots < t_n = b$.

Az $r(t_{i-1})$ és $r(t_i)$ között vegyünk fel minden i -re egy c_i pontot. A közelítő összeg legyen

$$k = \sum_{i=1}^n \underline{v}(c_i) \cdot (\underline{r}(t_i) - \underline{r}(t_{i-1}))$$

Ha a beosztás finomításával kapott k közelítőösszegeknek van egy határértéke, akkor az lesz a vonalintegrál (görbe menti integrál) értéke.

Speciálisan, ha a térgörbe, ami mentén integrálunk, olyan, hogy a kezdőpont egybeesik a végponttal, akkor **körintegrálról** beszélünk.

Jelölés:

$$\oint_C \underline{v} \, d\underline{r}$$

A vonalintegrál néhány tulajdonsága:

1. A görbe irányítását megcserélve az integrál értéke az ellentettjére változik.
2. Ha két azonos irányítású görbét összeragasztunk, akkor a megfelelő integrálokat összeadva kapjuk az ezen görbék uniója mentén vett integrált.

A körintegrál értéke nem mindig nulla, de sokszor az.

A vonalintegrál konkrét kiszámolása az alábbi módon történhet:

$$\int_C \underline{v} \, d\underline{r} = \int_a^b \underline{v}(\underline{r}(t)) \cdot \dot{\underline{r}}(t) \, dt$$

ahol $\dot{\underline{r}}(t)$ az $\underline{r}(t)$ gradiense, ami azt jelenti, hogy koordinátánként deriválunk.

Komplex számok

A valós számtestet fogjuk bővíteni. Vannak ugyanis olyan valós együtthatós polinomok, amiknek nem létezik valós gyöke. Ilyen pl. a $p(x) = x^2 + 1$.

Általánosabban tehát a probléma polinomok gyökeinek (zérushelyeinek) keresése.

Ismétlésképpen álljon itt, hogy mit értünk polinom alatt:

DEFINÍCIÓ (POLINOM): A $p(z) = a_0 + a_1z + a_2z^2 + \dots + a_kz^k$ kifejezés a z változó egy polinomja (ahol az a_i -k rögzített számok, a z pedig a változó).

Az általános megoldáshoz vezető út az, hogy bővítjük a valós számtestet. Először bevezetjük az ún. i számot, amire az teljesül, hogy $i^2 = -1$. Ezáltal készen állunk arra, hogy értelmezzük a komplex számok halmazát.

DEFINÍCIÓ (KOMPLEX SZÁMOK HALMAZA): Az $(a + bi)$ alakú számok, ahol a és b valós számokat jelölnek, alkotják a komplex számok halmazát.

Azt kijelentettük, hogy a komplex számok halmaza testet alkot, de le kell ellenőrizni, hogy tényleg teljesülnek-e a testaxiómák.

Az összeadás persze komponensenként (pontosabban koordinátánként) történik:

$$(a + bi) + (c + di) = a + c + (b + d)i$$

Az összeadás alábbi négy tulajdonságát könnyű ellenőrizni:

Összeadás tulajdonságai:

1. *kommutativitás (öröklődik a valós számok testéből)*
2. *asszociatív (öröklődik a valós számok testéből)*
3. *létezik nullelem (ami a hagyományos 0)*
4. *létezik mindenkinek ellentettje: $-(ai + bi) = -ai - bi$*

A szorzást is értelemszerűen terjesztjük ki (a kiemelési szabályt megtartjuk):

$$(a + bi)(c + di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i$$

A szorzásnak meglesznek az alábbi tulajdonságai:

Szorzás tulajdonságai:

1. *kommutatív*
2. *asszociatív*
3. *egységelem: a hagyományos 1*
4. *minden nullától különböző elemnek van reciproka.*

A negyediket fogjuk csak leellenőrizni, a többi triviális. Legyenek a és b olyan valós számok, hogy nem mindkettőjük nulla. Ez azzal ekvivalens, hogy a négyzetösszegük pozitív. Végezzük el az alábbi gyöktelenítést:

$$\frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 - (bi)^2} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} \cdot i$$

Mivel itt $a^2 + b^2 > 0$, ezért biztosan lehet vele osztani.

Következésként kijelenthetjük, hogy a komplex számok halmaza valóban egy test, ami tartalmazza a valós számtestet.

Az $(a+bi)$ alakot **algebrai alak**nak nevezzük. Az algebrai alakban az " a " az ún. **valós rész**, a " b " pedig a **képzetes rész**.

A $z=a+bi$ komplex szám ún. **konjugáltja** a $\bar{z} = a-bi$ szám. A szorzás 4. tulajdonságánál már láttuk, hogy $z \cdot \bar{z} = a^2 + b^2$. Könnyen ellenőrizhető továbbá, hogy tetszőleges v és w komplex számok esetén

$$\overline{\bar{v}} = v, \overline{v + w} = \bar{v} + \bar{w}, \text{ valamint } \overline{v \cdot w} = \bar{v} \cdot \bar{w}.$$

Komplex számok trigonometrikus és exponenciális alakja

A komplex számokat persze kétdimenziós vektorként is felfoghatjuk (azért, mert a és b egymástól függetlenül megválasztott valós számok). Az $a+bi$ komplex számnak feleljen meg az origóból az (a,b)-be mutató helyvektor.

A vektorainkat most jellemezzük a hosszúságuk (abszolútértékük) és a valós (azaz az x-tengellyel) bezárt szögük szerint.

A hosszúság a Pithagorasz-tétel miatt $r = \sqrt{a^2 + b^2}$. Legyen φ irányított szög, amit a szokásos módon mérünk (az x-tengely jobb félegyeneséből kiindulva). Ekkor az alábbiak teljesülnek:

$$\frac{a}{r} = \cos\varphi$$

$$\frac{b}{r} = \sin\varphi$$

$$a = r \cdot \cos\varphi$$

$$b = r \cdot \sin\varphi$$

A fentiekből tehát

$$a + bi = r \cdot \cos\varphi + i \cdot r \cdot \sin\varphi, \text{ azaz}$$

$$a + bi = r \cdot (\cos\varphi + i \cdot \sin\varphi), \text{ és ez lesz a komplex számok } \mathbf{\text{trigonometrikus alakja.}}$$

A trigonometrikus alak előnye, hogy szorzáskor kényelmes vele számolni. Felhasználva az addíciós tételeket, az alábbi nyerjük:

$$\begin{aligned} r_1(\cos\varphi_1 + i \cdot \sin\varphi_1) \cdot r_2(\cos\varphi_2 + i \cdot \sin\varphi_2) &= \\ = r_1 r_2 [(\cos\varphi_1 \cdot \cos\varphi_2 - \sin\varphi_1 \cdot \sin\varphi_2) + i \cdot (\sin\varphi_1 \cos\varphi_2 + \sin\varphi_2 \cdot \cos\varphi_1)] &= \\ = r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \cdot (\sin(\varphi_1 + \varphi_2))] &= \end{aligned}$$

Ebből az következik, hogy a szorzásnál a hosszúságokat összeszorozzuk, a szögeket pedig összeadjuk.
Pl. $2017(\cos 60^\circ + i \cdot \sin 60^\circ) \cdot 2(\cos 34^\circ + i \cdot \sin 34^\circ) = 4034(\cos 94^\circ + i \cdot \sin 94^\circ)$

$$\text{Ezért persze az osztás is könnyű, pl. } \frac{100 \cdot (\cos 147^\circ + i \cdot \sin 147^\circ)}{4 \cdot (\cos 92^\circ + i \cdot \sin 92^\circ)} = 25 \cdot (\cos 55^\circ + i \cdot \sin 55^\circ)$$

Ezt az egészet lehet még egyszerűbben írni oly módon, hogy hatványokat vezetünk be (ez vezet el az exponenciális alakban történő felíráshoz).

DEFINÍCIÓ: $e^{i\varphi} = \cos\varphi + i \cdot \sin\varphi$, ahol φ tetszőleges valós szám.

Az **exponenciális alak** ez alapján tehát $r \cdot e^{i\varphi}$, ahol r a hosszúság (abszolútérték), φ pedig a megfelelő szög, radiánban kifejezve. Fontos megjegyeznünk, hogy ezt egyelőre elfogadjuk definícióként, később azonban ki fog derülni, hogy csakis így lehet értelmesen definiálni.

Most nézzük meg, hogyan vonhatunk n -edik gyököt nullától különböző komplex számokból. Ehhez meg kell oldanunk a

$z^n = r \cdot (\cos \varphi + i \cdot \sin \varphi)$ egyenletet, ahol r , φ , n adott, és z az ismeretlen.

Ha $z = h(\cos \omega + i \cdot \sin \omega)$, akkor $z^n = h^n \cdot [\cos(n \cdot \omega) + i \cdot \sin(n \cdot \omega)] = r \cdot (\cos \varphi + i \cdot \sin \varphi)$.

Ez alapján $h^n = r$, ezért $h = \sqrt[n]{r}$ (a valós n -edik gyök), valamint $n \cdot \omega = \varphi + k \cdot 2\pi$, ahol k egész.

Innen $\omega = \frac{\varphi}{n} + \frac{k}{n} \cdot 2\pi$, és itt a k értékét n -féleképp választhatjuk meg, ilyenkor mindig különböző megoldásokat kapva.

Tehát minden nullától különböző komplex számnak pontosan n db különböző n -edik gyöke van.

Határérték és folytonosság a komplex számsíkon

A komplex számokat (mivel mindegyikük egyértelműen megfeleltethető egy rendezett valós számpárnak) megfeleltethetjük a sík pontjainak, így módon kapva a **komplex számsíkot**.

A komplex számsíkra természetes módon kiterjeszthetjük a **távolság** és a **környezet** fogalmát (a környezet egy olyan körlemez lesz, aminek a kerülete már nem eleme a halmaznak, ezt másképp **nyílt körlemeznek** szokás nevezni).

Mivel a távolság és a környezet fogalma átvihető a komplex számokra, ezért a határérték, illetve a folytonosság, valamint a deriválhatóság is ugyanúgy értelmezhető, mint a valós esetben, ezért ezeket itt külön nem is részletezzük. Sőt a **határértékre és folytonosságra vonatkozó alapvető tételek is érvényben maradnak** (a differenciálhatóságnál lesznek új dolgok, de erről majd később lesz szó).

Az **abszolútérték-függvénnyel** nem foglalkoztunk folytonosság szempontjából a valós függvényektárgyalása során. Itt példaként megmutatjuk, hogy ezek **mint komplex függvények is folytonosak** (azért, mert erre később úgyis szükségünk lesz). Tekintsük tehát a komplex számok halmazán értelmezett $p(z)=|z|$ függvényt. Legyen $c=a+bi$ tetszőleges komplex szám algebrai alakban felírva, valamint legyen (c_n) olyan komplex számsorozat, ami c -hez konvergál. Írjuk fel a sorozat elemeit algebrai alakban, legyen minden n -re $c_n = a_n + b_n \cdot i$. Ekkor tehát az $(a_n + b_n \cdot i)$ komplex számsorozat konvergál $(a + b \cdot i)$ -hez, vagyis a távolságuk, azaz $\sqrt{(a - a_n)^2 + (b - b_n)^2}$ konvergál nullához. Ekkor a konvergencia sorozatokról tanultak alapján $(a - a_n)^2 + (b - b_n)^2$ is konvergál nullához. Mivel itt valós értékekről van szó, ezért ez csak úgy lehetséges, hogy az $(a - a_n)$ és a $(b - b_n)$ sorozatok egyaránt nullához konvergálnak, azaz $a_n \rightarrow a$, $b_n \rightarrow b$. Ebből a konvergencia sorozatokról tanultak alapján következik, hogy a $\left(\sqrt{a_n^2 + b_n^2}\right)$ sorozat konvergál $\sqrt{a^2 + b^2}$ -hez, ami ugyanazt jelenti, hogy $|c_n| \rightarrow |c|$, és ezt akartuk belátni.

Kényelmi okokból kimondjuk a határértékre vonatkozó Cauchy-féle kritériumot, és belátjuk, hogy ez ekvivalens a konvergenciáról belátott eddigi három definícióval.

Tétel (Cauchy-féle konvergenciakritérium): A komplex (c_n) számsorozat pontosan akkor konvergens, ha bármely d pozitív távolság esetén létezik olyan k küszöbindex, hogy bármely k -nál nagyobb m és n indexek esetén $|c_m - c_n| < d$ teljesül.

Bizonyítás:

Két irányt kell bizonyítanunk. Először is azt látjuk be, hogy amennyiben a (c_n) komplex számsorozat konvergens, akkor teljesíti a Cauchy-féle konvergenciakritériumot. Amennyiben ugyanis a számsorozat konvergens, akkor egyetlen komplex torlódási pontja van, legyen ez t . Tudjuk, hogy bármely pozitív d távolság esetén létezik olyan k küszöbindex, hogy bármely k -nál nagyobb index esetén $|t - c_s| < \frac{d}{2}$. Ekkor persze tetszőleges $k < m, n$ indexekre fennáll, hogy

$$|t - c_m| < \frac{d}{2}, \text{ illetve } |t - c_n| < \frac{d}{2}.$$

A háromszög egyenlőtlenség alapján $|c_n - c_m| = |c_n - t + t - c_m| \leq |c_n - t| + |t - c_m| < \frac{d}{2} + \frac{d}{2} = d$, azaz a konvergencia maga után vonja a Cauchy-féle kritériumot.

Most a másik irányt látjuk be, azaz azt, hogy a Cauchy-féle kritérium teljesülése esetén a sorozat biztosan konvergens. Tegyük fel tehát, hogy a (c_n) komplex számsorozat eleget tesz a Cauchy-kritérium feltételeinek. A konvergenciát indirekt módon bizonyítjuk. Tegyük fel először indirekt, hogy ennek a sorozatnak két különböző komplex torlódási pontja is van, legyenek ezek t_1 és t_2 . Legyen d olyan pozitív távolság, ami kisebb, mint t_1 és t_2 távolságának az egyharmada, azaz $d < \left| \frac{t_1 - t_2}{3} \right|$. Ekkor persze ehhez a d -hez is találunk olyan k küszöbindexet, hogy minden k -nál nagyobb m és n indexre

$|c_n - c_m| < d$. Mivel t_1 és t_2 is torlódási pontok, ezért m és n értéke megválasztható úgy, hogy

$|c_n - t_1| < d$ és $|c_m - t_2| < d$ egyszerre teljesüljenek. Ekkor átalakításokat végezve, valamint használva a háromszög-egyenlőtlenséget, kapjuk, hogy

$$|t_1 - t_2| = |t_1 - c_n + c_m - t_2 + c_n - c_m| \leq |t_1 - c_n| + |c_m - t_2| + |c_n - c_m| < d + d + d <$$

$$< 3 \cdot \left| \frac{t_1 - t_2}{3} \right| = |t_1 - t_2|, \text{ ami nyilvánvalóan ellentmondás.}$$

Most már csak annyit kell igazolnunk, hogy a sorozat korlátos is. Ehhez tekintsük a feltételt a $d=1$ választás mellett. Az elemek közül véges sok egy egységnyi sugarú körön kívül, míg az összes többi mind ezen egységsugarú körön belül helyezkedik el, ezért a sorozat biztosan korlátos.

Tehát a másik irányban beláttuk, hogy a sorozat korlátos, valamint egyetlen torlódási pontja van, így tehát biztosan konvergens. ■

Az algebra alaptétele

Ami nekünk itt fontos lesz, az az, hogy **Weierstrass szélsőértékekre vonatkozó tétele is érvényben marad** abban az értelemben, hogy ha egy zárt körlemezen (tehát ami a kerületi pontokat is

tartalmazza) értelmezünk egy valós értékű komplex függvényt, ami folytonos is, akkor ennek a függvénynek a zárt körlemezen biztosan van legkisebb és legnagyobb értéke. Ezt ugyanúgy lehet bizonyítani, mint ahogy a valós esetben megtettük. A lényeg az, hogy a zárt körlemezen bármely olyan pontsorozat, ami eleget tesz a Cauchy-féle konvergenciakritériumnak, a körlemez valamelyik pontjához konvergál (ez igazából a **zárttság** általános megfogalmazása).

Ezen észrevétel alapján tudjuk bizonyítani a matematika egyik alapvető eredményét, az algebra alaptételét, ami komplex együtthatós polinomokról mond ki szép és általános eredményt.

AZ ALGEBRA ALAPTÉTELE: *Tetszőleges, legalább elsőfokú komplex együtthatós polinomnak létezik komplex gyöke.*

Bizonyítás:

Legyen $p(z) = c_n \cdot z^n + c_{n-1} \cdot z^{n-1} + \dots + c_1 z + c_0$ komplex együtthatós polinom, aminek foka $n \geq 1$ (azaz itt $c_n \neq 0$).

Vegyük először észre, hogy $\lim_{|z| \rightarrow +\infty} |p(z)| = +\infty$ (tehát nagy abszolútértékű komplex számokat behelyettesítve az eredmény abszolútértéke is nagy). Ez abból következik, hogy

$c_n \cdot z^n + c_{n-1} \cdot z^{n-1} + \dots + c_1 z + c_0 = z^n \cdot \left(c_n + \frac{c_{n-1}}{z} + \dots + \frac{c_0}{z^n} \right)$, ahol $|z| \rightarrow +\infty$ esetén a zárójelbeli kifejezés határértéke c_n (amiről feltettük, hogy nem nulla, hiszen a polinom n -edfokú).

Legyen m a $|p(z)|$ értékkészletének egy eleme (nyilván m valós, és legalább nullával egyenlő). A fenti határértéktulajdonság miatt van olyan R sugarú, nulla középpontú kör, amin kívül $|p(z)|$ mindig nagyobb m -nél.

A szóban forgó R sugarú körlemez zárt és korlátos, $|p(z)|$ pedig egy folytonos, valós értékű, komplex változós függvény ezen zárt körlemezen értelmezve, ezért $|p(z)|$ -nek a körlemezen van minimuma. Legyen ez egy z_0 komplex pontban.

Ebben az z_0 -ban a $|p(z)|$ -nek globális minimuma van, hiszen az R sugarú körlemezen kívül csak m -nél nagyobb értékeket vesz fel, az R sugarú zárt körlemezen pedig van legkisebb értéke, ami persze nem lehet m -nél nagyobb (ugyanis az m -et itt fel is veszi értékként).

Be fogjuk látni azt, hogy szükségképpen $|p(z_0)| = 0$, ami ugyanazt jelenti, hogy $p(z_0) = 0$, azaz z_0 egy zérushely (ne feledjük, hogy legalább egy zérushely létezésének bizonyítása a célunk).

Tegyük fel indirekt, hogy ez mégsem feltétlenül igaz, azaz tegyük fel, hogy $|p(z_0)| > 0$.

A könnyebb számolás kedvéért az eredeti polinom helyett vesszük a $g(z) = \frac{p(z+z_0)}{p(z_0)}$ polinomot.

Az új polinomra $g(0) = 1$ teljesül (azaz a konstans tagja 1), valamint $g(z)$ szintén n -edfokú, és ráadásul 0 a $|g(z)|$ -nek globális minimumhelye.

Legyen k az a pozitív egész, melyre $g(z) = 1 + a_k \cdot z^k + a_{k+1} \cdot z^{k+1} + \dots + a_n \cdot z^n$ teljesül.

Új változót vezetünk be a z helyett, legyen ez mondjuk t . Ezt formálisan úgy csináljuk, hogy a z változót kicseréljük $\sqrt[k]{\frac{-1}{a_k}} \cdot t$ -re, ahol valamelyik komplex k -adik gyökkel szoroztunk.

Ezáltal az újabb komplex polinom úgy néz ki, hogy $h(t) = 1 - t^k + b_{k+1} \cdot t^{k+1} + \dots + b_n \cdot t^n$.

A $|h(t)|$ függvény globális minimumértéke 1, hiszen $h(t)$ értékkészlete ugyanaz, mint a $g(z)$ polinomé.

Legyen most t valós, pozitív, és „kicsi” (tehát tartsunk vele jobbról a nullához).

Ekkor persze $|1 - t^k| = 1 - t^k < 1$, ugyanakkor

$|b_{k+1} \cdot t^{k+1} + \dots + b_n \cdot t^n| = t^k \cdot |b_{k+1} \cdot t^1 + \dots + b_n \cdot t^{n-k}| < t^k$, amennyiben a t értékét elég kicsinek választjuk.

Ilyen t számokra a háromszög-egyenlőtlenség értelmében

$$|h(t)| = |1 - t^k + b_{k+1} \cdot t^{k+1} + \dots + b_n \cdot t^n| \leq |1 - t^k| + |b_{k+1} \cdot t^{k+1} + \dots + b_n \cdot t^n| <$$

$< |1 - t^k| + t^k = 1$, ami ellentmond annak, hogy a $|h(t)|$ függvénynek a globális minimumértéke 1.

Tehát z_0 zérushelye a $p(z)$ -nek, és ezt akartuk belátni. ■

Bézout tétele ugyanúgy érvényben marad komplex együtthatós polinomokra is. Ezt nem kell külön bizonyítanunk, hiszen a komplex számok a valós számokhoz hasonlóan szintén testet alkotnak, tehát a bizonyítás szó szerint ugyanúgy működik komplex együtthatós polinomokra is. Így tehát az algebra alaptételének közvetlen következménye, hogy **minden legalább elsőfokú komplex együtthatós polinom előállítható annyi elsőfokú komplex együtthatós polinom szorzataként, amennyi a foka.**

Sőt az is igaz lesz, hogy **tetszőleges, legalább elsőfokú komplex együtthatós polinomnak annyi komplex zérushelye van, amennyi a foka** (itt a zérushelyeket multiplicitással kell számolni).

Végtelen sorok

Szeretnénk tisztázni, hogy miként lehetne értelmezni egy **végtelen sok tagú összeget**. Egészen pontosan az a kérdés, hogy hogyan kellene értenünk azt, amikor **egy sorozat elemeit szeretnénk összeadni**. Legyen adva egy (c_n) komplex számokból álló sorozat. Készítsük el ennek a sorozatnak az ún. részletösszeg sorozatát, azaz azt az (s_n) komplex számsorozatot, aminek k -adik elemét úgy kapjuk, hogy a (c_n) sorozat első k db elemét összeadjuk. Tehát $s_1 = c_1$, $s_2 = c_1 + c_2$, $s_3 = c_1 + c_2 + c_3$, és általánosan $s_k = c_1 + c_2 + \dots + c_k$.

A részletösszeg-sorozat konvergenciája fogja meghatározni, hogy a $c_1 + c_2 + c_3 + \dots$ végtelen sok tagú összeget értelmezzük-e, és amennyiben igen, akkor miképpen értelmezzük.

DEFINÍCIÓ: Azt mondjuk, hogy a (c_n) komplex számsorozatból készített $c_1 + c_2 + c_3 + \dots$ végtelen sok tagú összeg (végtelen sor) konvergens, ha a (c_n) sorozatból elkészített (s_k) részletösszegek sorozata is

konvergens. Ha az (s_n) részletösszeg-sorozat konvergens, és határértéke az s komplex szám, akkor a végtelen sor összegét s -nek tekintjük. Ha az (s_n) sorozat divergens, akkor a végtelen sort is divergensnek tekintjük és általában nem rendelünk hozzá értéket (kivéve esetleg $(+\infty)$ -t vagy $(-\infty)$ -t).

Példák:

1.

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \frac{1}{4 \cdot 5} + \dots$$

Itt az összeadandó tagok általánosan az alábbi módon írhatók át:

$$\frac{1}{x(x+1)} = \frac{(x+1) - x}{x(x+1)} = \frac{x+1}{x(x+1)} - \frac{x}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$$

Tehát általában a k -adik részletösszeg

$$s_k = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{k} - \frac{1}{k+1} = 1 - \frac{1}{k+1}. \text{ Ez tart } 1\text{-hez, tehát a sor összege } 1.$$

Megjegyzés: az olyan típusú összegeket, ahol a tagok a fent látott módon néhány kivételtől eltekintve kiesnek, **teleszkopikusnak** nevezzük (és persze vannak ilyen szorzatok is).

2. *Harmonikus sor:*

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$$

Végezzük el az alábbi csoportosításokat az összegben (az összeadandó tagok száma zárójelről zárójelre a duplájára változik):

$$1 + \left(\frac{1}{2} + \frac{1}{3}\right) + \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right) + \left(\frac{1}{8} + \frac{1}{9} + \dots + \frac{1}{15}\right) + \left(\frac{1}{16} + \frac{1}{17} + \dots + \frac{1}{31}\right) + \dots$$

Mivel a k -adik zárójelben az összeg $\frac{1}{2^k} + \frac{1}{1+2^k} + \frac{1}{2+2^k} + \dots + \frac{1}{2 \cdot 2^{k-1}} > \frac{1}{2 \cdot 2^k} \cdot 2^k = \frac{1}{2}$, ezért világos, hogy a

harmonikus sor összege plusz végtelennel egyenlő.

3. *A prímek reciprokösszege:*

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \dots$$

Később belátjuk, hogy a prímek reciprokösszege is egyenlő plusz végtelennel (ebből leolvasható, hogy a prímek viszonylag sűrűn helyezkednek el a természetes számok között).

4. *A négyzetszámok reciprokösszege:*

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \frac{1}{36} + \dots$$

Később belátjuk, hogy a négyzetszámok reciprokösszege $\frac{\pi^2}{6}$.

5. *A faktoriálisok reciprokösszege:*

$$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots$$

Később kiderül, hogy a faktoriálisok reciprokösszege az Euler-számmal egyenlő.

Mértani sorok és alkalmazásaik

A végtelen sorok speciális osztályát alkotják azok, melyeket (végtelen hosszú) mértani sorozatok elemeiből készítünk. Ezeket **mértani soroknak** nevezzük. Itt az elemek persze tetszőleges, komplex elemű mértani sorozat elemei lehetnek.

Egy mértani sor tehát $a + aq + aq^2 + aq^3 + \dots$

alakú, ahol „a” a kiinduló mértani sorozat (komplex) kezdőeleme, és q a (komplex) kvóciens.

Vizsgáljuk meg, hogy mikor konvergens egy ilyen sor, és ha konvergens, akkor határozzuk is meg az összeget.

A részletösszeg-sorozat k-adik eleme, azaz a sor első k db elemének összege

$s_k = a + aq + aq^2 + \dots + aq^{k-1}$, aminek értéke a korábban megismert összegformula szerint

$$s_k = a \cdot \frac{q^k - 1}{q - 1}.$$

Ha $q=1$, akkor a fenti tört nincs értelmezve, de ebben az esetben könnyen meggyőződhetünk róla, hogy a sor ilyenkor csak úgy lesz konvergens, ha $a=0$. Ha $a \neq 0$ és $q=1$, akkor a sor divergens, mivel minden eleme egyenlő, és nullától különböző.

Amennyiben $q \neq 1$, $a \neq 0$, akkor az $s_k = a \cdot \frac{q^k - 1}{q - 1} = a \cdot \frac{1 - q^k}{1 - q}$ tört értelmezve van. Amennyiben k tart a plusz végtelenbe és $|q| > 1$, akkor $|1 - q^k|$ is tart a plusz végtelenbe, ezért ilyenkor a sor divergens.

Ha k tart a plusz végtelenbe és $|q| < 1$, akkor $(1 - q^k)$ a nullához konvergál, azaz (s_k) sorozat konvergál az $\frac{a}{1-q}$ számhoz. Tehát $|q| < 1$ esetben azt kapjuk, hogy a mértani sor konvergens, és az összege $\frac{a}{1-q}$.

Nézzünk meg néhány problémát és alkalmazást, amik a mértani sorokkal kapcsolatosak:

1. Zénón „paradoxonja” Akhilleuszról és a teknősbékáról

Egy nagyon hosszú, egyenes pályán versenyzik egymással Akhilleusz és a teknős. Mindketten egyenletes sebességgel haladnak előre, ugyanabba az irányba. Mivel Akhilleusz tízszer olyan gyors, mint a teknős, ezért úgy dönt, hogy nagyvonalú lesz: ad a teknősnek száz lábnyi előnyt.

Zénón azt javasolja, hogy bontsuk időegységekre a futás folyamatát: tekintsük először azt a pillanatot, amikor Akhilleusz abba a pontba ér, ahol a teknős eredetileg tartózkodott. Ekkor a teknősnek tíz lábnyi előnye van Akhilleusszal szemben.

Másodszor ismétljük az előző gondolatmenetet, és tekintsük azt a pillanatot, amikor Akhilleusz elér abba a pontba, ahol a teknős tartózkodott az első vizsgált pillanatban. A teknős ekkor is előnyben van, még hozzá előnye egy lábnyi.

Ha ezt a gondolatmenetet végtelen sokszor megismétljük, akkor azt kapjuk, hogy minden vizsgált időpontban a teknős előnyben van, még hozzá az előnye mindig tizedére csökken az előző vizsgált időponthoz képest. Ez viszont (látszólag) azt jelenti, hogy Akhilleusz sosem lesz képes megelőzni a teknőst.

Hogyan lehet feloldani ezt a paradoxont?

Úgy, hogy észrevevessük, hogy hiába van végtelen sok vizsgált pillanat, azok egy véges időintervallumba esnek, ugyanis a vizsgált időpillanatokat egy $1/10$ kvóciensű mértani sorozat részletösszegei adják, ami konvergens.

2. A prímek reciprokösszege

Euler 1748-ban egy zseniális gondolatmenettel bizonyította, hogy a prímek reciprokösszege plusz végtelen (amiből persze az is következik, hogy végtelen sok prímszám van).

Induljunk ki abból, hogy az $\left(1 + \frac{1}{n}\right)^n$ sorozat szigorúan monoton növekvő, és határértéke e .

Legyen p tetszőleges prímszám. Ekkor $\left(1 + \frac{1}{p-1}\right)^{p-1}$ értéke e -nél kisebb.

A természetes alapú logaritmusfüggvény szigorú monotonitása miatt tudjuk, hogy

$$\ln\left(1 + \frac{1}{p-1}\right)^{p-1} < \ln e = 1, \text{ azaz}$$

$$(p-1) \cdot \ln\left(1 + \frac{1}{p-1}\right) < 1, \text{ ami azzal egyenértékű, hogy}$$

$$\ln\left(1 + \frac{1}{p-1}\right) < \frac{1}{p-1}.$$

Vegyük észre, hogy

$$1 + \frac{1}{p-1} = \frac{p}{p-1} = \frac{1}{1-\frac{1}{p}}.$$

Most jön a kulcsészrevétel: mivel $\frac{1}{p}$ egy 1-nél kisebb abszolútértékű szám, ezért $\frac{1}{1-\frac{1}{p}}$ felírható egy mértani sor összegeként az alábbi módon:

$$\frac{1}{1-\frac{1}{p}} = 1 + \frac{1}{p} + \frac{1}{p^2} + \dots$$

Legyen $n \geq 2$ egész szám, és legyenek az n -nél nem nagyobb (pozitív) prímek $p_1, p_2, \dots, p_k$.

A fentiek felhasználásával

$$\frac{1}{1-\frac{1}{p_1}} \cdot \frac{1}{1-\frac{1}{p_2}} \cdot \dots \cdot \frac{1}{1-\frac{1}{p_k}} =$$

$$\left(1 + \frac{1}{p_1} + \dots\right) \cdot \left(1 + \frac{1}{p_2} + \dots\right) \cdot \dots \cdot \left(1 + \frac{1}{p_k} + \dots\right)$$

A számelmélet alaptétele miatt ennek a szorzatnak az értéke legalább

$1 + \frac{1}{2} + \dots + \frac{1}{n}$ lesz. Mivel a harmonikus sor divergens, ezért ha n tart plusz végtelenbe, akkor a fenti szorzat értéke (amint egyre több prímet használunk fel) is tart plusz végtelenbe. Sőt ekkor nyilvánvalóan

$\ln\left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right)$ is a plusz végtelenbe tart. A logaritmusfüggvény folytonosságát, és a szigorú monotonitást felhasználva kapjuk, hogy

$$\ln\left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right) \leq$$

$$\leq \ln\left(1 + \frac{1}{p_1} + \dots\right) \cdot \left(1 + \frac{1}{p_2} + \dots\right) \cdot \dots \cdot \left(1 + \frac{1}{p_k} + \dots\right) =$$

$$= \ln\left(\frac{1}{1-\frac{1}{p_1}} \cdot \frac{1}{1-\frac{1}{p_2}} \cdot \dots \cdot \frac{1}{1-\frac{1}{p_k}}\right) =$$

$$= \ln \frac{1}{1-\frac{1}{p_1}} + \ln \frac{1}{1-\frac{1}{p_2}} + \dots + \ln \frac{1}{1-\frac{1}{p_k}} =$$

$$= \ln\left(1 + \frac{1}{p_1-1}\right) + \ln\left(1 + \frac{1}{p_2-1}\right) + \dots + \ln\left(1 + \frac{1}{p_k-1}\right) <$$

$$< \frac{1}{p_1-1} + \frac{1}{p_2-1} + \dots + \frac{1}{p_k-1} < 2 \cdot \left(\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_k} \right).$$

Ebből következik, hogy a prímek reciprokkösszege plusz végtelennel egyenlő. ■

Végtelen sorokra vonatkozó néhány konvergenciakritérium

Tétel (majoráns kritérium): Ha (a_n) és (b_n) olyan pozitív elemű valós számsorozatok, melyekre teljesül, hogy $a_n \leq b_n$ minden n indexre, továbbá tudjuk, hogy a (b_n) sorozat elemeiből készített $\sum b_n$ sor konvergens, akkor $\sum a_n$ sor is konvergens.

Bizonyítás:

Legyen az (a_n) sorozatból készített végtelen sorhoz tartozó részletösszeg-sorozat (A_k) , azaz legyen $\sum_{n=1}^k a_n = A_k$ minden k index esetén. Hasonlóan legyen $\sum_{n=1}^k b_n = B_k$ minden k index esetén.

A feltételek szerint a (B_k) részletösszeg-sorozat szigorúan monoton növekvő, és konvergens. A minden n indexre fennálló $a_n \leq b_n$ feltétel miatt az a (A_k) részletösszeg-sorozat felülről korlátos, hiszen a (B_k) sorozat szuprémuma mindenképpen felső korlátja. Ugyanakkor (lévén, hogy a sorozataink elemei pozitívak) az (A_k) sorozat szigorúan monoton növekvő. Az előzőekből következik, hogy az (A_k) sorozat konvergens, azaz a $\sum a_n$ sor is konvergens. ■

Tétel (minoráns kritérium): Ha (a_n) és (b_n) olyan pozitív elemű valós számsorozatok, melyekre teljesül, hogy $a_n \geq b_n$ minden n indexre, továbbá tudjuk, hogy a (b_n) sorozat elemeiből készített $\sum b_n$ sor divergens, akkor $\sum a_n$ sor is divergens.

Bizonyítás:

Világos, hogy a pozitív tagú $\sum b_n$ sor csak úgy lehet divergens, hogy az elemek összege plusz végtelen. Ez azt jelenti, hogy a $\sum b_n$ -hez elkészített (B_k) részletösszeg-sorozat a plusz végtelenbe tart. Viszont a minden n indexre fennálló $a_n \geq b_n$ feltétel miatt az (A_k) sorozat szintén a plusz végtelenbe tart, tehát $\sum a_n$ sor valóban divergens. ■

Most szükségünk lesz egy tételre, ami **abszolút konvergens sorokra** vonatkozik (tehát **amelyek elemeinek abszolútértékeinek összege konvergens**):

Abszolút konvergens sorok konvergenciája: Ha az (c_n) komplex számsorozatból készített végtelen sor abszolút konvergens, azaz $\sum_{n=1}^{\infty} |c_n|$ konvergens, akkor a sor elemeiből készített $\sum_{n=1}^{\infty} c_n$ sor is konvergens.

Bizonyítás:

Ha $\sum_{n=1}^{\infty} |c_n|$ konvergens, akkor részletösszeg-sorozata eleget tesz a Cauchy-féle konvergenciakritériumnak. Ez azt jelenti, hogy tetszőleges, pozitív d távolság esetén létezik olyan k index, hogy minden k -nál nagyobb $i < j$ indexek esetén

$$|c_i| + |c_{i+1}| + \dots + |c_j| < d.$$

A háromszög egyenlőtlenség miatt $|c_i + c_{i+1} + \dots + c_j| \leq |c_{i+1}| + \dots + |c_j|$, azaz ekkor a $\sum_{n=1}^{\infty} c_n$ -hez tartozó részletösszeg-sorozat szintén eleget tesz a Cauchy-konvergenciakritériumnak, ezért $\sum_{n=1}^{\infty} c_n$ is biztosan konvergens. ■

Az alábbiakban a mértani sorokról tanultak alapján fogalmazunk meg két konvergenciakritériumot, majd a Cantor-axióma segítségével is bizonyítunk egy kritériumot váltakozó előjelű, valós sorokra.

Tétel (hányadoskritérium): Amennyiben a (c_n) komplex számsorozathoz található olyan pozitív $q < 1$ szám, hogy $|\frac{c_{n+1}}{c_n}| \leq q$ minden n indexre, akkor $\sum c_n$ konvergens.

Bizonyítás:

A feltételekből adódik, hogy $|c_2| \leq |c_1| \cdot q$, $|c_3| \leq |c_1| \cdot q^2$, stb.

Általánosan könnyen belátható (pl. teljes indukcióval), hogy $|c_{n+1}| \leq |c_1| \cdot q^n$ teljesül minden n nemnegatív egészre.

A háromszög-egyenlőtlenséget is felhasználva kapjuk, hogy

$$|c_1 + c_2 + \dots + c_k| \leq |c_1| + |c_2| + \dots + |c_k| \leq |c_1| \cdot (1 + q + q^2 + \dots + q^{k-1}).$$
 Innen a majoráns

kritérium miatt adódik, hogy $\sum_{k=1}^{\infty} |c_k|$ konvergens, és az abszolút konvergens sorokra vonatkozó tétel miatt $\sum c_n$ is konvergens. ■

Tétel (gyökkritérium): Amennyiben az (c_n) komplex számsorozathoz található olyan pozitív $q < 1$ szám, hogy $\sqrt[n]{|c_n|} \leq q$ minden n indexre, akkor $\sum c_n$ konvergens.

Bizonyítás:

A feltételek közt szereplő egyenlőtlenség azzal ekvivalens, hogy $|c_n| \leq q^n$ minden n indexre, innen a háromszög-egyenlőtlenséget is felhasználva kapjuk, hogy

$$|c_1 + c_2 + \dots + c_k| \leq |c_1| + |c_2| + \dots + |c_k| \leq q + q^2 + \dots + q^k.$$

Mivel q abszolútértéke 1-nél kisebb, ezért $\sum_{k=1}^{\infty} |c_k|$ konvergens, amiből az abszolút konvergens sorokra vonatkozó tétel miatt következik, hogy $\sum c_n$ szintén konvergens. ■

Tétel (Leibniz-kritérium): Bármely nullához tartó, váltakozó előjelű, abszolútértékben szigorúan monoton csökkenő valós számsorozat elemeiből készített végtelen sor konvergens.

Bizonyítás:

Legyen (a_n) a szóban forgó sorozat. Erről tudjuk, hogy nullához tart, váltakozó előjelű, és azt is, hogy $|a_{n+1}| < |a_n|$ minden n pozitív egész esetén. Az természetesen feltehető, hogy a_1 pozitív (ha negatív lenne, akkor csak meg kellene cserélni az előjeleket a bizonyításban).

Legyen tetszőleges k pozitív egész esetén $s_k = a_1 + a_2 + \dots + a_k$.

Vegyük észre, hogy az előjelváltások miatt $s_1 > s_2, s_3 > s_4, s_5 > s_6, \dots$

Világos, hogy bármely k pozitív egész esetén $s_{2k+1} \geq s_{2k}$.

Mivel az elemek abszolútértékeiből készített sorozat szigorúan monoton csökkenő, ezért tetszőleges j pozitív egész esetén

$$s_{2j+1} = s_{2j-1} + a_{2j} + a_{2j+1} < s_{2j-1}, \text{ és hasonlóan}$$

$$s_{2j+2} = s_{2j} + a_{2j+1} + a_{2j+2} > s_{2j}.$$

Ez azt jelenti, hogy az $[s_{2k}, s_{2k+1}]$ zárt intervallumok eleget tesznek a Cantor-axióma feltételeinek. Ezen intervallumok pontosan egy számot határoznak meg, hiszen az (a_n) sorozat nullához tart. Az egyetlen szám, amit az imént definiált egymásba skatulyázott zárt intervallumok tartalmaznak, a $\sum(a_n)$ végtelen sor értéke. ■

A Leibniz-kritérium következménye, hogy a harmonikus sorból készített váltakozó előjelű

$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots$ végtelen sor konvergens. Sőt később ennek a sornak a pontos értékét is meg tudjuk majd határozni.

Végtelen sorok átrendezése, műveletek végtelen sorokkal

Kézenfekvő kérdés, hogy ugyanúgy viselkedik-e egy végtelen sor, ha benne a tagok sorrendjét megváltoztatjuk. Lévén, hogy végtelen sok tagú összegekről van szó, a kérdés megválaszolása

egyáltalán nem triviális. Azonban abszolút konvergencia sorok esetén szép, általános eredmény fogalmazható meg.

Tétel: Legyen (c_n) olyan komplex számsorozat, melyre teljesül, hogy $\sum_{n=1}^{\infty} |c_n|$ konvergens. Ha a (d_n) komplex számsorozat a (c_n) sorozatnak egy átrendezése (azaz a két sorozatban pontosan ugyanazok az elemek fordulnak elő, és persze ugyanannyiszor, csak más sorrendben), akkor $\sum_{n=1}^{\infty} d_n$ konvergens, és értéke megegyezik akkor $\sum_{n=1}^{\infty} c_n$ értékével.

Bizonyítás:

Azt már tárgyaltuk, hogy $\sum_{n=1}^{\infty} |c_n|$ konvergenciája maga után vonja $\sum_{n=1}^{\infty} c_n$ konvergenciáját. Tehát azt kellene belátnunk, hogy $\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} d_n$.

Mivel $\sum_{n=1}^{\infty} |c_n|$ és $\sum_{n=1}^{\infty} c_n$ is konvergens, ezért tetszőleges d rögzített pozitív távolsághoz létezik olyan k index, hogy bármely k -nál nagyobb m index esetén

$$||c_1| + |c_2| + \dots + |c_m| - \sum_{n=1}^{\infty} |c_n|| = \sum_{n=m+1}^{\infty} |c_n| < d, \text{ valamint}$$

$$|(c_1 + c_2 + \dots + c_m) - \sum_{n=1}^{\infty} c_n| < d \text{ egyaránt teljesül.}$$

Rögzítsük most a d távolságot, illetve ehhez egy alkalmas k küszöbindexet, ami a fentieknek eleget tesz. Legyen továbbá $m \geq k$, és tekintsük a $c_1, c_2, \dots, c_m$ elemeket. Az átrendezési tulajdonság miatt biztosan létezik olyan N pozitív egész szám, hogy teljesül az, hogy a $d_1, d_2, d_3, \dots, d_N$ elemek között előfordul a $c_1, c_2, c_3, \dots, c_m$ elemek mindegyike (ez nyilván multiplicitással értendő, tehát mindegyik elem legalább annyiszor szerepel, ahányszor kell).

Legyen K egy N -nél nagyobb index. Adjunk most felső becslést a $(d_1 + d_2 + \dots + d_K)$ és a $\sum_{n=1}^{\infty} c_n$ eltérésére. A háromszög-egyenlőtlenség felhasználásával kapjuk, hogy

$$\begin{aligned} |(d_1 + d_2 + \dots + d_K) - \sum_{n=1}^{\infty} c_n| &= \\ &= |(d_1 + d_2 + \dots + d_K) - (c_1 + c_2 + \dots + c_m) + (c_1 + c_2 + \dots + c_m) - \sum_{n=1}^{\infty} c_n| \leq \\ &\leq |(d_1 + d_2 + \dots + d_K) - (c_1 + c_2 + \dots + c_m)| + |(c_1 + c_2 + \dots + c_m) - \sum_{n=1}^{\infty} c_n|. \end{aligned}$$

Itt a bizonyítás elején elmondottak szerint $|(c_1 + c_2 + \dots + c_m) - \sum_{n=1}^{\infty} c_n| < d$.

Belátjuk, hogy $|(d_1 + d_2 + \dots + d_K) - (c_1 + c_2 + \dots + c_m)|$ is kisebb d -nél. Vegyük észre ugyanis, hogy az átrendezési tulajdonság miatt $(d_1 + d_2 + \dots + d_K) - (c_1 + c_2 + \dots + c_m)$ olyan c_j számok összege, melyek mind m -nél nagyobb indexűek, ezért a háromszög-egyenlőtlenség alkalmazásával azt kapjuk, hogy

$$|(d_1 + d_2 + \dots + d_K) - (c_1 + c_2 + \dots + c_m)| \leq \sum_{n=m+1}^{\infty} |c_n| < d.$$

Tehát tetszőleges d -hez létezik olyan N küszöbindex, hogy minden N -nél nagyobb K indexre

$$|(d_1 + d_2 + \dots + d_K) - \sum_{n=1}^{\infty} c_n| < 2d, \text{ ez pedig éppen azt jelenti, hogy } \sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} d_n. \blacksquare$$

A fenti tételből azt kaptuk tehát, hogy ha egy sor abszolút konvergens, akkor már az elemei önmagukban egyértelműen meghatározzák a sor összegét, az elemek sorrendjére való tekintet nélkül. Ebből következményként adódik, hogy **abszolút konvergens sorok összege is abszolút konvergens, és a szóban forgó sorösszegek egyszerűen összeadódnak.**

Egy konvergens sort **feltételesen konvergensnek** nevezünk, ha az elemek abszolútértékeiből alkotott sor már nem konvergens. Ilyenre már láttunk példát. Ugyanis a

$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots$ sor konvergens, de az $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$ nem konvergens.

Megmutatjuk, hogy a feltételesen konvergens sorok igen szeszélyesen viselkednek, ha átrendezzük őket. Ezt csak valós sorokra nézzük meg (komplex sorokra az itteni ötlet átvihető).

Tétel (Riemann): Legyen (x_n) valós számsorozat, melyre $\sum_{n=1}^{\infty} x_n$ konvergens, valamint

$\sum_{n=1}^{\infty} |x_n| = +\infty$, legyen továbbá s tetszőleges valós szám. Ekkor az (x_n) sorozatnak létezik olyan

(y_n) átrendezése, melyre $\sum_{n=1}^{\infty} y_n = s$, továbbá létezik olyan (z_n) átrendezése is, melyre $\sum_{n=1}^{\infty} z_n$ divergens.

Bizonyítás:

Először is rendezzük külön sorozatba az (x_n) pozitív, illetve negatív elemeit: legyen a pozitív elemek sorozata (p_n) , a negatívaké pedig (q_n) .

Először is vegyük észre, hogy $\sum_{n=1}^{\infty} p_n = +\infty$. Ugyanis ha a pozitív elemek összege konvergálna, akkor a negatív elemek összege is konvergálna, és $\sum_{n=1}^{\infty} x_n$ abszolút konvergens lenne, ami ellentmond a feltételeknek. Természetesen ugyanígy igazolható, hogy $\sum_{n=1}^{\infty} q_n = -\infty$.

Legyen most s rögzített nemnegatív valós szám. Kezdjük el összeadogatni sorrendben a (p_n) sorozat elemeit. Amikor az részletösszeg túllépi az s értéket, akkor kezdjük el hozzáadogatni sorrendben a negatív elemeket, egészen addig, amíg az részletösszeg ismét s alá nem süllyed. Ha ezt az eljárást vég nélkül folytatjuk, akkor ezáltal a sor egy olyan átrendezését definiáljuk, aminek összege s . Ez következik abból, hogy az (x_n) sorozat határértéke nulla (hiszen $\sum_{n=1}^{\infty} x_n$ konvergens).

Ha az előző ötletet úgy módosítjuk, hogy a pozitív elemekkel mindig addig lépkedünk, amíg a részletösszeg 1 fölé nem nő, majd a negatív elemekkel mindig addig lépkedünk, amíg a részletösszeg 0 alá nem csökken, akkor értelemszerűen a sor egy olyan átrendezéséhez jutunk, ami divergál. ■

Az előző eredményekből jól látszik, hogy egyáltalán nem mindegy, hogyan értelmezzük végtelen sorok szorzatát. Az ún. Cauchy-féle szorzatot fogjuk használni (ez a leginkább elterjedt megközelítés).

A komplex $\sum_{k=1}^{\infty} v_k$ és $\sum_{m=1}^{\infty} w_m$ **sorok Cauchy-szorzata** úgy lesz értelmezve mint a

$\sum_{n=1}^{\infty} (v_1 w_n + v_2 w_{n-1} + v_3 w_{n-2} + \dots + v_{n-1} w_2 + v_n w_1)$ végtelen sor.

Konvergens sorok szorzatára vonatkozik Mertens alábbi tétele.

Tétel (Mertens): Egy abszolút konvergens sornak egy konvergens sorral vett Cauchy-szorzata konvergens, és a szorzatként kapott sor értéke egyenlő a két sor értékének szorzatával.

Bizonyítás:

Legyenek (w_n) és (v_n) olyan komplex számsorozatok, melyekre $\sum_{r=1}^{\infty} v_r$ és $\sum_{s=1}^{\infty} |w_s|$ konvergens.

Azt kellene bizonyítanunk, hogy

$$\sum_{k=1}^{\infty} (v_1 w_k + v_2 w_{k-1} + v_3 w_{k-2} + \dots + v_{k-1} w_2 + v_k w_1) = (\sum_{r=1}^{\infty} v_r) \cdot (\sum_{s=1}^{\infty} w_s).$$

Vezessük be a megfelelő részletösszeg-sorozatokot, azaz tetszőleges n index esetén legyen

$$\sum_{r=1}^n v_r = V_n, \text{ valamint } \sum_{s=1}^n w_s = W_n.$$

A feltételek szerint a (V_n) részletösszeg-sorozat konvergens, jelölje a határértékét V .

Hasonlóan a (W_n) részletösszeg-sorozat is konvergens, jelölje a határértékét W .

Ezekkel a jelölésekkel az alábbi kapjuk:

$$\begin{aligned} \sum_{k=1}^n (v_1 w_k + v_2 w_{k-1} + v_3 w_{k-2} + \dots + v_{k-1} w_2 + v_k w_1) &= \\ &= w_1 \cdot \sum_{r=1}^n v_r + w_2 \cdot \sum_{r=1}^{n-1} v_r + w_3 \cdot \sum_{r=1}^{n-2} v_r + \dots + w_n \cdot \sum_{r=1}^1 v_r = \\ &= w_1 \cdot V_n + w_2 \cdot V_{n-1} + w_3 \cdot V_{n-2} + \dots + w_n \cdot V_1 = \sum_{j=1}^n w_{n-j} \cdot V_j \\ &= \sum_{j=1}^n w_{n-j} \cdot (V_j - V) + W_n \cdot V. \end{aligned}$$

Legyen d tetszőleges pozitív távolság. Mivel a (V_n) részletösszeg-sorozat határértéke V , ezért létezik olyan N index, hogy minden N -nél nem kisebb n index esetén $|V_n - V| < d$.

Mivel a (w_n) sorozat határértéke nulla, ezért létezik olyan $N^* > N$ index, hogy minden N^* -nál nagyobb n -re fennáll az, hogy

$$|w_n| < \frac{d}{N \cdot \left[1 + \max_{1 \leq j \leq N-1} |V_j - V| \right]}$$

Mivel a (W_n) részletösszeg-sorozat tart W -hez, ezért létezik olyan $N^{**} > N^*$ index, hogy bármely N^{**} -nál nagyobb n esetén $|W_n - W| < d$.

Legyen most n egy olyan index, ami nagyobb $(N + N^{**})$ -nál. Ekkor végezzük el az alábbi átalakítást:

$$\begin{aligned} \left| \sum_{k=1}^n (v_1 w_k + v_2 w_{k-1} + v_3 w_{k-2} + \dots + v_{k-1} w_2 + v_k w_1) - V \cdot W \right| &= \\ &= \left| \sum_{j=1}^n w_{n-j} \cdot (V_j - V) + W_n \cdot V - V \cdot W \right| = \left| \sum_{j=1}^n w_{n-j} \cdot (V_j - V) + (W_n - W) \cdot V \right| \end{aligned}$$

Ezt a kifejezést kell felülről becsülnünk. A háromszög-egyenlőtlenség miatt erre az alábbi felső becslés adható:

$$\begin{aligned} & \left| \sum_{j=1}^n w_{n-j} \cdot (V_j - V) + (W_n - W) \cdot V \right| \leq \\ & \leq \sum_{j=0}^{N-1} |w_{n-j}| \cdot |V_j - V| + \sum_{j=N}^n |w_{n-j}| \cdot |V_j - V| + |(W_n - W) \cdot V|. \end{aligned}$$

Ezt a háromtagú összeget tagonként vizsgáljuk.

Mivel $n > N + N^{**}$, ezért az első összeadandó tagban, a $\sum_{j=0}^{N-1} |w_{n-j}| \cdot |V_j - V|$ szummában az $(n-j)$ indexek mind nagyobbak N^{**} -nál, így

$$\sum_{j=0}^{N-1} |w_{n-j}| \cdot |V_j - V| \leq \sum_{j=0}^{N-1} \frac{d}{N \cdot \left[1 + \max_{1 \leq j \leq N-1} |V_j - V|\right]} \cdot |V_j - V| < d.$$

A második összeadandó tagban, a $\sum_{j=N}^n |w_{n-j}| \cdot |V_j - V|$ szummában a j indexek mindegyike legalább N -nel egyenlő, ezért

$$\sum_{j=N}^n |w_{n-j}| \cdot |V_j - V| < \sum_{j=N}^n |w_{n-j}| \cdot d \leq \sum_{j=1}^{\infty} |w_{n-j}| \cdot d$$

A harmadik összeadandó tag, vagyis $|(W_n - W) \cdot V|$ egyszerűen felülről becsülhető úgy, hogy

$$|(W_n - W) \cdot V| < |d \cdot V|, \text{ hiszen az } n \text{ index nagyobb } N^{**}\text{-nál.}$$

Összességében azt a becslést kaptuk, hogy az elég nagy n -ek mindegyikére

$$\left| \sum_{j=1}^n w_{n-j} \cdot (V_j - V) + (W_n - W) \cdot V \right| < d + \sum_{j=1}^{\infty} |w_{n-j}| \cdot d + |d \cdot V|.$$

Mivel $\sum_{j=1}^{\infty} |w_{n-j}|$ konvergens, ezért a $(d + \sum_{j=1}^{\infty} |w_{n-j}| \cdot d + |d \cdot V|)$ tetszőlegesen kicsi lehet.

Ezzel a Mertens-tételt bebizonyítottuk. ■

Hatványsorok

A mértani sorokra vonatkozó képlet következménye, hogy tetszőleges, egynél kisebb abszolútértékű z komplex szám esetén

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots$$

Ez úgy is megfogalmazható, hogy az $\frac{1}{1-z}$ komplex függvény előállítható a z változó hatványainak összegeként, amennyiben $|z| < 1$.

Általánosan is ezt szeretnénk tenni. Bizonyos komplex $p(z)$ függvényeket fel szeretnénk írni

$$p(z) = c_0 + c_1 \cdot (z - c) + c_2 \cdot (z - c)^2 + c_3 \cdot (z - c)^3 + \dots = \sum_{k=0}^{\infty} c_k \cdot (z - c)^k$$

ún. **hatványsor**-alakban, ahol a (c_n) komplex számsorozat elemei az együtthatók, és a c rögzített komplex szám a hatványsor ún. **középpontja**.

Figyeljük meg, hogy a hatványsor fogalma általánosabb, mint a polinomé, hiszen a hatványsorban akár végtelen sok összeadandó tag is szerepelhet.

Hatványsorok konvergenciája

Először is megvizsgáljuk, hogy mikor konvergens egy hatványsor.

Ennek leírásához bevezetjük a **limesz superior** fogalmát: egy **valós számsorozat** (vagy akár számhalmaz) limesz superiorja alatt a **torlódási pontjainak szuprémumát** értjük. Hangsúlyozzuk, hogy ez az elnevezés most kizárólag valós számsorozatokra, illetve valós számhalmazokra vonatkozik, ugyanis ilyenkor definiáltunk egyértelmű rendezést az elemek között.

Természetesen hasonló módon értelmezhető a **limesz inferior** fogalma, de erre most nem lesz szükségünk.

A hatványsorok konvergenciáját a Cauchy-Hadamard-tétellel írjuk le:

Tétel (Cauchy-Hadamard):

A $\sum_{k=0}^{\infty} c_k \cdot (z - c)^k$ komplex hatványsor abszolút konvergens a c középpontú,

$R = \frac{1}{\limsup \sqrt[n]{|c_n|}}$ sugarú kör belső pontjaiban, a körön kívüli pontokban viszont divergens.

Bizonyítás:

Az nyilvánvalóan mindegy, hogy a középpontot hogyan választjuk meg, ezért legyen az egyszerűség kedvéért a hatványsor középpontja $c = 0$. Tehát azt szeretnénk vizsgálni, hogy milyen z komplex számok esetén konvergens a $\sum_{k=0}^{\infty} c_k \cdot z^k$ összeg.

Legyen $R = \frac{1}{\limsup \sqrt[n]{|c_n|}}$, és tekintsük a komplex síkon a $c=0$ középpontú, R sugarú kört.

Legyen először z olyan pont, amit ezen kör belsejéből vettünk, azaz legyen $|z| < R$.

Ekkor egy küszöbtől kezdve bármely k -ra

$$\left| \frac{1}{c_k \cdot z^k} \right| = \left| \left(\frac{1}{z} \cdot \frac{1}{\sqrt[k]{|c_k|}} \right)^k \right| \geq \left| \left(\frac{1}{z} \cdot \frac{1}{\limsup \sqrt[n]{|c_n|}} \right)^k \right| = \left| \left(\frac{1}{z} \cdot R \right)^k \right| = \left| \left(\frac{R}{z} \right)^k \right|, \text{ azaz}$$

$$|c_k \cdot z^k| \leq \left| \left(\frac{z}{R} \right)^k \right|. \text{ A gyökkritérium miatt } (q = \left| \frac{z}{R} \right| < 1 \text{ választás mellett}) \text{ kapjuk, hogy } \sum_{k=0}^{\infty} \left| \left(\frac{z}{R} \right)^k \right|$$

konvergens. Innen a majoráns kritérium alapján $\sum_{k=0}^{\infty} c_k \cdot z^k$ abszolút konvergens.

Ha a z pontot a körön kívül választjuk (tehát $\left| \frac{z}{R} \right| > 1$), akkor végtelen sok k index esetén fordított egyenlőtlenség teljesül, azaz végtelen sok k -ra $|c_k \cdot z^k| \geq \left| \left(\frac{z}{R} \right)^k \right| \geq 1$. Ekkor persze a $(c_k \cdot z^k)$ sorozat nem tarthat nullához, holott ez a Cauchy-féle konvergenciakritérium miatt szükséges feltétele annak, hogy a $\sum_{k=0}^{\infty} c_k \cdot z^k$ összeg konvergens legyen. Tehát a körön kívül a hatványsor biztosan divergens. ■

A Cauchy-Hadamard-tételben szereplő c középpontú, $R = \frac{1}{\limsup \sqrt[n]{|c_n|}}$ sugarú kör az adott hatványsorhoz tartozó **konvergenciakör**, melynek R hosszú sugarát **konvergenciasugárnak** szokás nevezni. A konvergenciakör határpontjaival kapcsolatban a tétel nem állít semmit. Ez nem véletlen, ugyanis a határpontok bármelyikén elvileg a hatványsor lehet konvergens, de akár divergens is.

Amennyiben $\limsup \sqrt[n]{|c_n|} = 0$, akkor a hatványsor konvergenciasugara plusz végtelennel egyenlő, azaz bármely a hatványsor bármely komplex z esetén konvergens.

Ha $\limsup \sqrt[n]{|c_n|}$ egyenlő plusz végtelennel, akkor a hatványsor konvergenciasugara nulla hosszúságú, és ekkor a hatványsor csakis a középpontjában konvergens.

Hatványsorok differenciálása

Szép eredményt kapunk, ha a hatványsorok deriváltjait vizsgáljuk. Ezt az alábbi tételben foglaljuk össze.

Tétel: *Bármely komplex hatványsor a konvergenciakörének belsejében tagonként differenciálható, és a deriválást követően kapott hatványsor konvergenciasugara ugyanakkora, mint az eredeti hatványsor konvergenciasugara.*

Bizonyítás:

Legyen $\sum_{k=0}^{\infty} c_k \cdot (z - c)^k$ komplex hatványsor. Ismét feltehetjük, hogy a középpont, azaz c , egyenlő nullával.

Ha formálisan, tagonként differenciáljuk a $\sum_{k=0}^{\infty} c_k \cdot z^k$ hatványsort, akkor a $\sum_{k=0}^{\infty} k \cdot c_k \cdot z^{k-1}$ hatványsort kapjuk. Ha ez utóbbit még z -vel szorozzuk, akkor persze a konvergenciasugara ugyanakkora marad.

A Cauchy-Hadamard-tétel szerint a kapott $\sum_{k=0}^{\infty} k \cdot c_k \cdot z^k$ hatványsor konvergenciasugara $\frac{1}{\limsup \sqrt[k]{|k \cdot c_k|}}$, ami megegyezik $\frac{1}{\limsup \sqrt[k]{|c_k|}}$ -val, hiszen $\sqrt[k]{k}$ határértéke 1-gyel egyenlő, amennyiben k tart végtelenbe.

Ugyanezzel a gondolatmenettel kapható, hogy a konvergenciasugár akkor sem változik, ha a második formális deriváltat képezzük, azaz a $\sum_{k=0}^{\infty} k \cdot (k-1) \cdot c_k \cdot z^{k-2}$ hatványsor konvergenciasugara is ugyanakkora, mint az eredeti hatványsoré (erre az észrevételre a későbbiekben még szükségünk lesz).

Még annyit be kell látnunk, hogy a hatványsor differenciálható a megfelelő pontokban, és a deriválás után kapott hatványsor valóban tagonkénti differenciálással kapható meg belőle.

A Cauchy-Hadamard-tételből tudjuk, hogy a konvergenciakörön belül található bármely z_0 pontban $\sum_{k=0}^{\infty} c_k \cdot z_0^k$ abszolút konvergens. Legyen tehát z_0 tetszőleges pont a konvergenciakör belsejében.

Írjuk fel a megfelelő hányadost, amiből megállapítható a derivált, és vonjuk ki belőle

$\sum_{k=0}^{\infty} k \cdot c_k \cdot z_0^{k-1}$ -et, majd alakítsuk át a szóban forgó különbséget. Azt kell belátnunk, hogy $z \rightarrow z_0$ esetén ez az eltérés nullához tart. A számolás során a lényeg az lesz, hogy megmutassuk, ebből a különbségből kiemelhető $(z - z_0)$, ami persze tart nullához (hiszen $z \rightarrow z_0$). A másik szorzótényezőről, ami egy végtelen sok tagú összeg lesz, elég lesz annyit belátnunk, hogy abszolútértékben korlátos. Végezzük el először csak az alábbi átalakítást:

$$\frac{(\sum_{k=0}^{\infty} c_k \cdot z^k) - (\sum_{k=0}^{\infty} c_k \cdot z_0^k)}{z - z_0} - \sum_{k=0}^{\infty} k \cdot c_k \cdot z_0^{k-1} = \sum_{k=0}^{\infty} c_k \cdot \left(\frac{z^k - z_0^k}{z - z_0} - k \cdot z_0^{k-1} \right)$$

Bontsuk tovább a feladatot, és vizsgáljuk egyelőre csak a $\left(\frac{z^k - z_0^k}{z - z_0} - k \cdot z_0^{k-1} \right)$ alakú kifejezéseket, amik a fenti zárójelekben szerepelnek:

$$\begin{aligned} \frac{z^k - z_0^k}{z - z_0} - k \cdot z_0^{k-1} &= (z^{k-1} + z^{k-2} \cdot z_0 + z^{k-3} \cdot z_0^2 + \dots + z_0^{k-1}) - k \cdot z_0^{k-1} = \\ &= (z^{k-1} - z_0^{k-1}) + (z^{k-2} \cdot z_0 - z_0^{k-1}) + (z^{k-3} \cdot z_0^2 - z_0^{k-1}) + \dots + (z \cdot z_0^{k-2} - z_0^{k-1}) = \\ &= (z - z_0) \cdot (z^{k-2} + z^{k-3} \cdot z_0 + z^{k-4} \cdot z_0^2 + \dots + z_0^{k-2}) + \\ &+ (z - z_0) \cdot (z^{k-3} \cdot z_0 + z^{k-4} \cdot z_0^2 + \dots + z_0^{k-2}) + \\ &+ (z - z_0) \cdot (z^{k-4} \cdot z_0^2 + z^{k-5} \cdot z_0^3 + \dots + z_0^{k-2}) + \\ &+ \dots \\ &\dots \\ &\dots \\ &+ (z - z_0) \cdot z_0^{k-2}. \end{aligned}$$

Látható, hogy a $(z - z_0)$ szorzótényező valóban kiemelhető lesz. A másik szorzótényezőben összevonások után kapjuk, hogy a fenti kifejezés másképpen felírva úgy néz ki, hogy

$$(z - z_0) \cdot [(k-1) \cdot z_0^{k-2} + (k-2) \cdot z_0^{k-3} \cdot z + \dots + 2 \cdot z_0 \cdot z^{k-3} + z^{k-2}].$$

Mivel $z \rightarrow z_0$, ezért elég a másik (hosszú) szorzótényező abszolútértékére felső becslést adnunk. Legyen u olyan komplex szám, aminek abszolútértéke $|z_0|$ -nál és $|z|$ -nél egyaránt nagyobb.

A háromszög-egyenlőtlenséget felhasználva kapjuk, hogy

$$\begin{aligned} &|(k-1) \cdot z_0^{k-2} + (k-2) \cdot z_0^{k-3} \cdot z + \dots + 2 \cdot z_0 \cdot z^{k-3} + z^{k-2}| \leq \\ &\leq |(k-1) \cdot z_0^{k-2}| + |(k-2) \cdot z_0^{k-3} \cdot z| + |2 \cdot z_0 \cdot z^{k-3}| + |z^{k-2}| \leq \\ &\leq |(k-1) \cdot u^{k-2}| + |(k-2) \cdot u^{k-2}| + \dots + |2 \cdot u^{k-2}| + |u^{k-2}| = \end{aligned}$$

$$= |u^{k-2}| \cdot \frac{(k-1) \cdot k}{2}.$$

Az u megválasztható úgy, hogy $\sum_{k=0}^{\infty} |c_k \cdot u^{k-2}| \cdot \frac{(k-1) \cdot k}{2}$ konvergens legyen, hiszen a bizonyítás elején megjegyeztük, hogy a $\sum_{k=0}^{\infty} k \cdot (k-1) \cdot c_k \cdot z^{k-2}$ hatványsor konvergenciasugara ugyanakkora, mint az eredeti hatványsoré, az u számmal szemben pedig csak annyi elvárásunk van, hogy az abszolútértéke nagyobb legyen $|z_0|$ -nál és $|z|$ -nél (ahol $z \rightarrow z_0$). Tehát u megválasztható a konvergenciakör belsejében, és ebből adódik, hogy a kiemelt $(z - z_0)$ szorzótényező egy abszolútértékben korlátos kifejezéssel van megszorozva, és ezt akartuk belátni. ■

Az imént bizonyított tétel következménye, hogy **bármely hatványsor akárhányszor differenciálható, és a deriválások során a konvergenciasugar változatlan marad**. Sőt amennyiben egy $p(z)$ komplex függvényről tudjuk, hogy hatványsorba fejthető egy körön belül, mondjuk formálisan $p(z) = \sum_{k=0}^{\infty} c_k \cdot (z - c)^k$, akkor a hatványsor együtthatóit egymás utáni deriválásokkal egyértelműen meg tudjuk határozni. Ekkor ugyanis

$$c_0 = p(c), c_1 = p'(c), c_2 = \frac{p''(c)}{2} \text{ stb.}$$

Általánosan azt kapjuk, hogy $c_k = \frac{p^{(k)}(c)}{k!}$, ahol $p^{(k)}(z)$ a $p(z)$ komplex függvény k -adik deriváltját jelöli. Ellenőrizzük ezt le. Deriváljuk ugyanis k -szor egymás után a $\sum_{n=0}^{\infty} c_n \cdot (z - c)^n$ hatványsort. A k -adik deriváltban a konstans tag $c_k \cdot k!$, tehát ha a k -adik deriváltba c -t helyettesítünk, akkor a felvett érték éppen $c_k \cdot k!$ (hiszen a többi összeadandó tag mind kinullázódik, ha c -t helyettesítünk be). Ugyanakkor a $p(z) = \sum_{k=0}^{\infty} c_k \cdot (z - c)^k$ feltétel miatt $p^{(k)}(c) = c_k \cdot k!$, ami valóban azt jelenti, hogy

$$c_k = \frac{p^{(k)}(c)}{k!}.$$

Ennek az észrevételnek közvetlen következménye, hogy ha egy függvényt valamilyen körön belül hatványsorba tudunk fejteni, akkor ez a **hatványsorba történő fejtes egyértelmű**, hiszen az együtthatók egyértelműen kiszámolhatók az imént kapott képletekből.

Már készen állunk arra, hogy a nevezetesebb valós függvényeket sorba fejtsük. Kezdjük az exponenciális függvénnyel.

Tétel: $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ tetszőleges x valós szám esetén.

Bizonyítás:

Előrebocsátjuk, hogy az együtthatókat könnyen megtalálhatjuk egymás utáni deriválásokkal. Azt kell igazából tisztázni, hogy az exponenciális függvény tényleg hatványsorba fejthető.

Legyen $f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Először is határozzuk meg a konvergenciasugarat. Ehhez meg kell határoznunk $\limsup \frac{1}{\sqrt[k]{k!}}$ értékét. A mértani középsorozatoknál beláttuk, hogy $\sqrt[k]{k!}$ tart végtelenbe, ha k tart végtelenbe. Itt erre kiegészítésképpen adunk egy közvetlen bizonyítást is (a mértani középsorozat fogalmát nem használva):

$$\text{Tudjuk, hogy } \sqrt[k]{k!} = \sqrt[k]{1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot (k-1) \cdot k}.$$

Legyen m az a legnagyobb pozitív egész, ami $k/2$ -nél még nem nagyobb (azaz legyen m a $k/2$ -nek az egészrésze). Ekkor

$$\begin{aligned} \sqrt[k]{1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot (k-1) \cdot k} &\geq \sqrt[k]{1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot (2m-1) \cdot 2m} = \\ \sqrt[k]{\prod_{j=1}^m (m+1+j) \cdot (m+1-j)} &= \\ = \sqrt[k]{\prod_{j=1}^m [(m+1)^2 - j^2]} &\geq \sqrt[k]{\prod_{j=1}^m [(m+1)^2 - m^2]} = \sqrt[k]{\prod_{j=1}^m [2m+1]} = \\ = (2m+1)^{\frac{m}{k}} &\geq (k)^{\frac{\frac{k}{2}-1}{k}}. \end{aligned}$$

Nilván $(k)^{\frac{\frac{k}{2}-1}{k}}$ tart végtelenbe, ha k tart végtelenbe, ezért $\sqrt[k]{k!}$ is végtelenbe tart. Ebből kapjuk, hogy a konvergenciasugár végtelen hosszú, azaz az $f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ sor minden valós x -re konvergens.

Vegyük észre, hogy az $f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ sor deriváltja önmaga (hiszen $\frac{x^k}{k!}$ -nek az x szerinti deriváltja $\frac{x^{k-1}}{(k-1)!}$), valamint $f(0)=1$.

Bebizonyítjuk, hogy ebből a két feltételből következik, hogy bármely valós x -re $f(x) = e^x$.

Tekintsük ugyanis az $\frac{f(x)}{e^x}$ függvényt a valós számok halmazán. Ennek deriváltfüggvénye

$\frac{f'(x) \cdot e^x - f(x) \cdot e^x}{e^{2x}} = \frac{(f'(x) - f(x)) \cdot e^x}{e^{2x}}$, ami azonosan nulla, hiszen az $f(x)$ függvény deriváltfüggvénye önmaga.

Ezért $\frac{f(x)}{e^x}$ függvény biztosan konstans, ráadásul az $f(0)=1$ feltétel miatt csakis azonosan 1 lehet.

Ezzel tehát bizonyítottuk, hogy minden valós x -re $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. ■

Vegyük észre azt, hogy ezzel nem csupán a valós exponenciális függvényt jellemeztük. A konvergenciasugár ugyanis akkor is végtelen hosszú marad, ha a hatványsorba komplex számokat is helyettesíthetünk (hiszen az együtthatók nem változnak). Így tehát az **exponenciális függvényt mint komplex függvényt** úgy tudjuk értelmezni, hogy tetszőleges z komplex szám esetén előírjuk, hogy

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!} \text{ teljesüljön.}$$

Sőt az is kijön, hogy az exponenciális függvényre vonatkozó $e^{x+y} = e^x \cdot e^y$ azonosság is érvényben marad. Tetszőleges x és y valós számokra ugyanis teljesül, hogy $e^{x+y} = e^x \cdot e^y$. Ezért tehát bármely x és y valós számokra igaz, hogy $\left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right) \cdot \left(\sum_{m=0}^{\infty} \frac{y^m}{m!}\right) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}$.

Azt kellene megmutatnunk, hogy tetszőleges v és w komplex számokra is ugyanez érvényes.

Ezt egy tételben fogalmazzuk meg.

Tétel: Tetszőleges v és w komplex számokra $e^{v+w} = e^v \cdot e^w$.

Bizonyítás:

Tekintsük a $\left(\sum_{k=0}^{\infty} \frac{v^k}{k!}\right) \cdot \left(\sum_{m=0}^{\infty} \frac{w^m}{m!}\right)$ szorzatot, ahol v és w komplex számok.

Ez a szorzat, ha mindent mindennel szorzunk, a Cauchy-szorzat definíciója szerint éppen

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{v^k \cdot w^{n-k}}{k! \cdot (n-k)!} \right).$$

Elég lenne tehát annyit bizonyítani, hogy bármely n természetes számra

$$\sum_{k=0}^n \frac{v^k \cdot w^{n-k}}{k! \cdot (n-k)!} = \frac{(v+w)^n}{n!}.$$

A bizonyítandó állítás azzal ekvivalens, hogy $\sum_{k=0}^n \frac{n!}{k! \cdot (n-k)!} \cdot v^k \cdot w^{n-k} = (v+w)^n$.

Ha a jobb oldalon elvégezzük az n -edik hatványra emelést, akkor a $(v+w)$ tényezőt n -szer szorozzuk össze önmagával, tehát valóban $v^k \cdot w^{n-k}$ típusú összedandó tagokat kapunk, a kérdés csak az, hogy adott k mellett hányszor kapjuk meg a $v^k \cdot w^{n-k}$ szorzatot (ez lesz ugyanis $v^k \cdot w^{n-k}$ együtthatója). Természetesen annyszor, ahányféleképpen ki tudunk jelölni k darab zárójelet az n darab zárójel közül úgy, hogy azokból a v betűt választjuk, a maradék $(n-k)$ darab zárójelből pedig a w betűt választjuk. Az első v betűt n darab zárójelből választhatjuk, a másodikat $(n-1)$ darab zárójelből, ..., a k -adikat $(n-k+1)$ darab zárójelből. Ez összesen $n \cdot (n-1) \cdot \dots \cdot (n-k+1)$ -féle lehetőség lenne, de ezt osztanunk kell $k!$ -sal, mert a zárójelek egymás közötti elhelyezkedésére nem voltunk tekintettel. Tehát

$$\sum_{k=0}^n \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!} \cdot v^k \cdot w^{n-k} = (v+w)^n.$$

Ugyanakkor ha a $\frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}$ törtet $(n-k)!$ -sal bővítjük, akkor éppen $\frac{n!}{k! \cdot (n-k)!}$ -t kapunk eredményül. Ezzel az állítást beláttuk. ■

Az iménti második bizonyításban beláttott és felhasznált $\sum_{k=0}^n \frac{n!}{k! \cdot (n-k)!} \cdot v^k \cdot w^{n-k} = (v+w)^n$ azonosságot **binomiális tételnek** szokás nevezni.

Menjünk most tovább, és fejtsük hatványsorba a trigonometrikus függvényeket (először valósban, majd ezeket is kiterjesztjük komplexre).

Tétel:

(i) $\cos x = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k}}{(2k)!}$ tetszőleges x valós szám esetén.

(ii) $\sin x = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!}$ tetszőleges x valós szám esetén.

Bizonyítás:

Legyen $g(x) = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k}}{(2k)!}$.

A konvergenciasugár végtelen hosszúságú, ami triviálisan következik abból az érvelésből, amit az exponenciális függvény hatványsoránál láttunk.

Könnyen ellenőrizhető, hogy bármely valós x -re $g''(x) = -g(x)$, valamint $g(0)=1$, $g'(0)=0$.
Megmutatjuk, hogy legfeljebb egy olyan $g(x)$ valós függvény létezik, amire teljesül, hogy

$$g''(x) = -g(x), \text{ valamint } g(0)=1, g'(0)=0.$$

Tegyük fel indirekt, hogy van két különböző $g_1(x)$ és $g_2(x)$ valós függvény, amikre ezek teljesülnek.

Legyen $d(x) = g_1(x) - g_2(x)$, ami az indirekt feltevés szerint nem azonosan nulla.

Ekkor $d''(x) = [g_1(x) - g_2(x)]'' = g_1''(x) - g_2''(x) = -g_1(x) + g_2(x) = -d(x)$, valamint

$$d(0) = d'(0) = 0.$$

A $d''(x) = -d(x)$ egyenlet mindkét oldalát szorozzuk meg $2 \cdot d'(x)$ -szel.

Ekkor $2 \cdot d'(x) \cdot d''(x) = -d(x) \cdot 2 \cdot d'(x)$, ami azzal ekvivalens, hogy a

$2 \cdot d'(x) \cdot d''(x) + d(x) \cdot 2 \cdot d'(x)$ függvény azonosan nulla. Ez azonban éppen a $\left[(d'(x))^2 + (d(x))^2\right]$ függvény deriváltfüggvénye.

Ezért $\left[(d'(x))^2 + (d(x))^2\right]$ konstansfüggvény, és a $d(0) = d'(0) = 0$ feltételek miatt csakis azonosan nulla lehet. Mivel valós számok négyzetösszege csakis akkor lehet nulla, ha mindegyikük nulla, ezért kapjuk, hogy $d(x)$ azonosan nulla, ami ellentmond annak, hogy $g_1(x)$ és $g_2(x)$ nem azonos függvények.

Ezért legfeljebb egy valós $g(x)$ függvény felel meg a feltételeknek. De ilyen valós függvény biztosan van, ez pedig a koszinuszfüggvény. Így tehát egészen biztosak lehetünk abban, hogy $g(x) = \cos x$ (hiszen másik jelölt nem jön szóba).

Az (ii)-beli állítás szó szerint ugyanígy igazolható (csak mások a kezdőértékek). ■

A most belátott hatványsor-előállítások néhány fontos következményét vizsgáljuk.

Legyen z komplex szám. Írjuk fel először e^{iz} -t:

$$\begin{aligned} e^{iz} &= \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} = \sum_{k=0}^{\infty} \frac{(iz)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(iz)^{2k+1}}{(2k+1)!} = \sum_{k=0}^{\infty} i^{2k} \cdot \frac{z^{2k}}{(2k)!} + \sum_{k=0}^{\infty} i^{2k+1} \cdot \frac{z^{2k+1}}{(2k+1)!} = \\ &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k}}{(2k)!} + i \cdot \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k+1}}{(2k+1)!}. \end{aligned}$$

Hasonlóan felírjuk e^{-iz} -et:

$$e^{-iz} = e^{i \cdot (-z)} = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{(-z)^{2k}}{(2k)!} + i \cdot \sum_{k=0}^{\infty} (-1)^k \cdot \frac{(-z)^{2k+1}}{(2k+1)!} =$$

$$= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k}}{(2k)!} - i \cdot \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k+1}}{(2k+1)!}.$$

$$\text{Tehát } \frac{e^{iz} + e^{-iz}}{2} = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k}}{(2k)!}, \text{ valamint } \frac{e^{iz} - e^{-iz}}{2i} = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k+1}}{(2k+1)!}.$$

A $\sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k}}{(2k)!}$ és a $\sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k+1}}{(2k+1)!}$ hatványsorok konvergenciasugara végtelen hosszú,

ezért tetszőleges z komplex szám szinuszt tudjuk értelmezni úgy, hogy

$$\sin z = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k+1}}{(2k+1)!}.$$

a koszinuszt pedig úgy, hogy

$$\cos z = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{z^{2k}}{(2k)!}.$$

Ezáltal a **szinuszt és a koszinuszt komplex értékekre is kiterjesztettük**. Az exponenciális összefüggések segítségével könnyen igazolható, hogy az **addíciós tételek komplexben is érvényesek maradnak**.

Sőt ezzel **igazoltuk a komplex számok exponenciális alakjára vonatkozó definíció létjogosultságát**, hiszen a hatványsorokból leolvasható, hogy minden komplex z számra

$$e^{iz} = \cos z + i \cdot \sin z.$$

A logaritmusfüggvény komplex kiterjesztése, illetve hatványsora

A logaritmusfüggvény ugyanúgy terjeszthető ki komplex értelemben, mint ahogy valós függvényként értelmeztük, vagyis a komplex exponenciális függvény inverzeként. Itt azonban arra vigyázni kell, hogy a komplex exponenciális függvény nem injektív, így csak alkalmasan megválasztott komplex tartomány(ko)n tudjuk a komplexlogaritmust egyértelműen értelmezni.

A logaritmusfüggvényt is hatványsorba tudjuk fejteni (először valósban). Az alábbi tételt bizonyítjuk be:

Tétel: Tetszőleges egynél kisebb abszolútértékű valós x szám esetén teljesül, hogy

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

Bizonyítás:

$$\text{Legyen } p(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \dots$$

A $p(x)$ konvergenciasugara a Cauchy-Hadamard-tétel szerint $\frac{1}{\limsup \sqrt[n]{|a_n|}} = 1$, tehát $|x| < 1$ esetén a

$p(x)$ hatványsor konvergens. Vizsgáljuk $p(x)$ deriváltját:

$$p'(x) = 1 - x + x^2 - x^3 + \dots$$

Azt már korábban tudjuk, hogy $|x| < 1$ esetén $1 - x + x^2 - x^3 + \dots = \frac{1}{1+x}$.

Tehát a $p(x)$ függvényről tudjuk, hogy $p(0)=1$, valamint $|x| < 1$ esetén $p'(x) = \frac{1}{1+x}$.

A primitív függvényekről tanultak alapján ebből kapjuk, hogy $p(x) = \ln(1+x) + c$ minden $|x| < 1$ esetén, ahol c alkalmas konstans. Viszont mindenképpen $c=0$, hiszen $p(0)=1$. Ezzel a tétel állítását beláttuk. ■

A most látott hatványsorba fejtés evidens módon kiterjeszthető komplex értelemben is a nulla középpontú egységkör belsejében, de nem terjeszthető ki az egész komplex síkon (hiszen a konvergenciasugár hossza egységnyi). Így tehát példát adtunk olyan függvényre, melynek hatványsora nem állítja elő a függvényt az egész komplex számsíkon.

Megemlítjük, hogy az $\arctg x$ függvény deriváltfüggvénye $\frac{1}{1+x^2}$, így a fentihez teljesen hasonló módon igazolható, hogy $|x| < 1$ esetén teljesül, hogy

$$\arctg x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \dots$$

Abel folytonossági tétele

Amikor a hatványsorok konvergenciaköréről beszéltünk, egyáltalán nem foglalkoztunk a kör kerületi pontjaival. Ezt a hiányosságot némileg kárpótolja Abel folytonossági tétele, aminek lesz néhány fontos következménye.

Abel-féle folytonossági tétel: Legyen az $f(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$ függvény értelmezési tartománya mindazon z valós számok halmaza, melyekre a fenti hozzárendelésben megadott hatványsor konvergens. Amennyiben a hatványsor konvergenciasugara nullától különböző, akkor az $f(z)$ valós függvény az értelmezési tartományának minden pontjában szükségképpen folytonos is.

Bizonyítás:

Legyen a konvergenciasugár $R > 0$. A konvergenciakör belsejében az állítás komplex értelemben is teljesül, hiszen ott már korábban beláttuk, hogy a hatványsor differenciálható is (tagonként). Ezért elég az állítást a konvergenciakör kerületének olyan valós pontjaiban bizonyítani, amelyekben a hatványsor konvergens. Csak két lehetséges vizsgálandó pont van, ezek az R és a $(-R)$. Tegyük fel mondjuk, hogy a hatványsor R -ben konvergens, és legyen $f(R)=s$. Legyen a q olyan valós változó, aminek abszolútértéke egynél kisebb. Minden ilyen q értékre $f(Rq)$ értelmezett, és

$$f(Rq) = c_0 + (c_1 \cdot R) \cdot q + (c_2 \cdot R^2) \cdot q^2 + (c_3 \cdot R^3) \cdot q^3 + \dots$$

Mivel $|q| < 1$, ezért

$$\frac{1}{1-q} = 1 + q + q^2 + q^3 + \dots$$

Tekintsük most $f(Rq)$ -t és $\frac{1}{1-q}$ -t is egy-egy q változó szerinti hatványsornak a $(-1,1)$ intervallumon. A korábbi ismereteink alapján tudjuk, hogy itt ezek abszolút konvergensek, így Mertens tétele miatt Cauchy-szorzatuk is konvergens.

A szóban forgó hatványsorok Cauchy-szorzata úgy néz ki, hogy

$$\frac{f(Rq)}{1-q} = \sum_{n=0}^{\infty} (c_0 + c_1 \cdot R + c_2 \cdot R^2 + \dots + c_n \cdot R^n) \cdot q^n$$

Vezessük most be az $s_n = c_0 + c_1 \cdot R + c_2 \cdot R^2 + \dots + c_n \cdot R^n$ sorozatot.

$$\text{Ezzel a jelöléssel } f(Rq) = (1 - q) \cdot \sum_{n=0}^{\infty} s_n \cdot q^n$$

$$\text{Vegyük észre, hogy } s \text{ átírható úgy, hogy } s = (1 - q) \cdot \sum_{n=0}^{\infty} s \cdot q^n.$$

Az $f(Rq)$ és s eltérését fogjuk vizsgálni.

$$s - f(Rq) = (1 - q) \cdot \sum_{n=0}^{\infty} s \cdot q^n - (1 - q) \cdot \sum_{n=0}^{\infty} s_n \cdot q^n = (1 - q) \cdot \sum_{n=0}^{\infty} (s - s_n) \cdot q^n$$

Legyen d rögzített pozitív távolság. Mivel $s_n \rightarrow s$, ezért van v olyan pozitív egész, hogy minden $N \geq v$ esetén $|s - s_N| < d$.

Ekkor írjuk át az $|(1 - q) \cdot \sum_{n=0}^{\infty} (s - s_n) \cdot q^n|$ kifejezést az alábbi módon:

$$|s - f(Rq)| = |(1 - q) \cdot \sum_{n=0}^{\infty} (s - s_n) \cdot q^n| =$$

$$|(1 - q) \cdot \sum_{n=0}^{v-1} (s - s_n) \cdot q^n + (1 - q) \cdot \sum_{n=v}^{\infty} (s - s_n) \cdot q^n|.$$

Ez a háromszög-egyenlőtlenség segítségével felülről becsülhető úgy, hogy

$$|(1 - q) \cdot \sum_{n=0}^{v-1} (s - s_n) \cdot q^n + (1 - q) \cdot \sum_{n=v}^{\infty} (s - s_n) \cdot q^n| \leq$$

$$\leq |(1 - q) \cdot \sum_{n=0}^{v-1} (s - s_n) \cdot q^n| + |(1 - q) \cdot \sum_{n=v}^{\infty} (s - s_n) \cdot q^n|.$$

Mivel az R -beli folytonosságot vizsgáljuk, ezért az feltehető, hogy $q \rightarrow 1$. Így tehát

$|(1 - q) \cdot \sum_{n=0}^{v-1} (s - s_n) \cdot q^n|$ alkalmas q -ra kisebb lesz d -nél, valamint (használva a háromszög-egyenlőtlenséget is) kapjuk, hogy

$$|(1 - q) \cdot \sum_{n=v}^{\infty} (s - s_n) \cdot q^n| \leq \sum_{n=v}^{\infty} |(1 - q) \cdot \sum_{n=v}^{\infty} (s - s_n) \cdot q^n| =$$

$$= |1 - q| \cdot \sum_{n=v}^{\infty} |s - s_n| \cdot |q^n| < |1 - q| \cdot d \cdot \frac{1}{1 - |q|} = d \cdot \frac{|1 - q|}{1 - |q|} = d.$$

Tehát azt kaptuk, hogy az $|s - f(Rq)|$ eltérés kicsi lesz, ha $q \rightarrow 1$, és ez éppen az R -beli folytonosságot jelenti.

Természetesen ugyanígy járhatunk el, ha a hatványsor $(-R)$ -ben konvergens. ■

A bizonyításból könnyen leolvasható, hogy Abel tételét általánosíthattuk volna a komplex síkon azzal a gyengébb megkötéssel, hogy a bizonyítás során olyan komplex q értékeket használunk, amikre $\frac{|1-q|}{1-|q|}$ felülről korlátos (és persze $q \rightarrow 1$, valamint $|q| < 1$). Az ilyen q értékek ún. Stolz-tartományokat írnak le a komplex síkon.

Abel tételének egy fontos következménye, hogy meg tudjuk határozni az $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots$ sor összegét. Erről ugyanis a Leibniz-kritérium alapján biztosan tudjuk, hogy konvergens, de értékét még nem sikerült meghatároznunk. Használjuk fel azt a már bizonyított tételt, hogy $-1 < x < 1$ esetén

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

A fenti hatványsor az $x=1$ helyen konvergens, így Abel tétele miatt ugyanott folytonos is. De ez azt jelenti, hogy az ottani értéke egyenlő $\lim_{x \rightarrow 1} \ln(1+x)$ -gyel, ami éppen $\ln 2$.

Egy ugyanilyen alkalmazáshoz használjuk fel azt, hogy minden $|x| < 1$ esetén teljesül az, hogy

$$\arctg x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \dots$$

Ebből a fentihez hasonló gondolatmenettel kapható az ún. Leibniz-féle sor, azaz

$$\frac{\pi}{4} = \arctg 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots$$

Ezzel többek között lehetőségünk nyílt a π szám „geometriamentes” bevezetésére is (egy anekdota szerint Leibniz annyira megörült ennek a képletnek, hogy emiatt hagyta ott a jogi pályát, és onnantól kezdve matematikával foglalkozott). A Leibniz-féle formulának az a gyengéje, hogy viszonylag lassan közelíthető vele a π , ezért gyakorlati alkalmazásokban rendszerint nem ezt a formulát használják.

Lineáris algebra

Először lineáris egyenletrendszerekkel foglalkozunk, azaz olyan egyenletrendszerekkel, ahol az egyenletekben minden ismeretlen az első hatványon szerepel, és ezek nincsenek egymással összeszorozva.

Egy lineáris egyenletrendszer általános alakja:

$$a_{11} \cdot x_1 + a_{12} \cdot x_2 + \dots + a_{1n} \cdot x_n = b_1$$

$$a_{21} \cdot x_1 + a_{22} \cdot x_2 + \dots + a_{2n} \cdot x_n = b_2$$

.....

.....

.....

$$a_{k1} \cdot x_1 + a_{k2} \cdot x_2 + \dots + a_{kn} \cdot x_n = b_k$$

Természetesen adódik a kérdés, miszerint hogyan oldunk meg lineáris egyenletrendszereket.

Arra fogunk törekedni, hogy a felírt lineáris egyenletrendszerből gyártsunk egy egyszerűbbet. Erre fog szolgálni az ún. **Gauss-elimináció (kiküszöbölés)** módszere. Az eljárás során minden lépésben arra törekszünk, hogy az egyenletek száma eggyel, az ismeretlenek száma pedig legalább eggyel csökkenjen.

Válaszunk egy ismeretlent, amit eliminálni szeretnénk, legyen ez mondjuk x_1 . Az egyenleteinket két csoportra osztjuk: az első csoportbeliekben szerepel az x_1 (nem nulla az együtthatója), a második csoportbelieknél az x_1 együtthatója nulla. Egyelőre a második csoportbelieket jegeljük. Az első csoportbeli egyenletek mindegyikét megszorozzuk egy-egy alkalmas számmal úgy, hogy szorzás után az x_1 -ek együtthatója mindenütt megegyezzen. Ezekre az egyenletekre nézve végezzük el a következő lépéseket: vonjuk ki az elsőből a másodikat, a másodikból a harmadikat, stb... Így (az első csoportbelieknél) eggyel kisebb új egyenlethez jutottunk. A második csoportbeli (elraktározott) egyenletekkel együtt ezek egy olyan lineáris egyenletrendszert alkotnak, ahol az eredetiekhez képest 1-gyel kevesebb egyenlet van, és x_1 nem szerepel.

Mátrixműveletek

A mátrixok téglalap alakú táblázatok, ahol a mezőket (általában számokkal) kitöltjük.

Pl. $\begin{pmatrix} 1 & 3 & 4 \\ -1 & 2 & 5 \end{pmatrix}$ egy (2×3) -as méretű mátrix, tehát megegyezés szerint először a sorok, majd az oszlopok számát adjuk meg.

A vektoroknál látotthoz hasonló módon történik **mátrixok összeadása**: két azonos méretű mátrix összege is egy ugyanilyen méretű mátrix lesz, amit úgy kapunk, hogy a megfelelő elemeket összeadjuk. Például:

$$\begin{pmatrix} 1 & 3 & 4 \\ -1 & 2 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 5 & 0 \\ 1 & 5 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 8 & 4 \\ 0 & 7 & 10 \end{pmatrix}.$$

A **skalárral való szorzás** is ugyanúgy megy, mint a vektoroknál láttuk, minden elemet végigszorozunk az adott skalárral. Például:

$$5 \cdot \begin{pmatrix} 1 & 3 & 4 \\ -1 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 15 & 20 \\ -5 & 10 & 25 \end{pmatrix}.$$

A **skaláris szorzást** általánosan két egyforma hosszú vektor között fogjuk értelmezni, a térbeli vektorokkal analóg módon. Az $(x_1, x_2, \dots, x_n)$ és $(y_1, y_2, \dots, y_n)$ vektorok skaláris szorzata legyen $\sum_{j=1}^n x_j \cdot y_j$.

A **mátrixok szorzása** már jóval bonyolultabb, konszenzuson alapuló művelet. Eleve nem kommutatív, tehát számít, hogy melyik mátrix kerül a bal, és melyik a jobb oldalra. A méretre is vannak megkötések. Ebben a sorrendben egy $(n \times k)$ -es mátrixot tudunk összeszorozni egy $(k \times l)$ -essel, azaz a bal oldalon állónak annyi oszlopa van, ahány sora van a jobb oldalon állónak. Ebben az esetben az eredmény egy $(n \times l)$ -es mátrix, amiben az i -edik sor j -edik eleme úgy áll elő, mint a bal oldali mátrix i -edik sorvektorának, illetve a jobb oldali mátrix j -edik oszlopának skaláris szorzata.

Példa:

$$\begin{pmatrix} 1 & -2 & 0 & 3 \\ 5 & 1 & -1 & 2 \\ 0 & 3 & 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 7 & 2 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + (-2) \cdot 1 + 0 \cdot 7 + 3 \cdot 0 & 1 \cdot 1 + (-2) \cdot (-1) + 0 \cdot 2 + 3 \cdot 8 \\ 5 \cdot 2 + 1 \cdot 1 + (-1) \cdot 7 + 2 \cdot 0 & 5 \cdot 1 + 1 \cdot (-1) + (-1) \cdot 2 + 2 \cdot 8 \\ 0 \cdot 2 + 3 \cdot 1 + 1 \cdot 7 + 5 \cdot 0 & 0 \cdot 1 + 3 \cdot (-1) + 1 \cdot 2 + 5 \cdot 8 \end{pmatrix}$$

A mátrixszorzás nem kommutatív, mint ahogy az alábbi egyszerű példa is mutatja:

$$\begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 6 & 6 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 10 & 11 \end{pmatrix}, \text{ viszont } \begin{pmatrix} 6 & 6 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 12 \\ -2 & -3 \end{pmatrix}.$$

Viszont mátrixokra nézve teljesül az asszociatív tulajdonság.

Tétel: A mátrixszorzás asszociatív.

Bizonyítás:

Legyen A egy $(n \times k)$ -as, B egy $(k \times l)$ -es, C pedig egy $(l \times t)$ -es mátrix.

Ekkor $(AB)C$ és $A(BC)$ egyaránt $(n \times t)$ -es méretű mátrixok. Azt fogjuk belátni, hogy az $(AB)C$ és $A(BC)$ mátrixok első sorának első eleme megegyezik. A többi elemre a bizonyítás átindexeléssel ugyanígy működik.

Legyen A mátrix i -edik sorának j -edik eleme a_{ij} , B mátrix u -adik sorának v -edik eleme b_{uv} , illetve C mátrix r -edik sorának s -edik eleme az c_{rs} .

Az AB mátrix első sorának w -edik eleme $\sum_{i=1}^k a_{1i}b_{iw}$, míg a C mátrix első oszlopának w -edik eleme c_{w1} . Ezért az $(AB)C$ mátrix első sorának első eleme

$$\sum_{w=1}^l \left(c_{w1} \left(\sum_{i=1}^k a_{1i} b_{iw} \right) \right).$$

BC mátrix első oszlopának i -edik eleme $\sum_{w=1}^k b_{iw}c_{w1}$, ezért az $A(BC)$ mátrix első sorának első eleme $\sum_{i=1}^k \left(a_{1i} \left(\sum_{w=1}^k b_{iw}c_{w1} \right) \right).$

A $\sum_{w=1}^l \left(c_{w1} \left(\sum_{i=1}^k a_{1i} b_{iw} \right) \right)$ illetve a $\sum_{i=1}^k \left(a_{1i} \left(\sum_{w=1}^k b_{iw}c_{w1} \right) \right)$ kifejezések egyenlők, mivel mindkettőben az $a_{1i}b_{iw}c_{w1}$ alakú szorzatok vannak összeadva (mindegyik pontosan egyszer szerepel). Ezzel a tétel állítását beláttuk. ■

Mátrixműveletek kapcsolata lineáris egyenletrendszerekkel

A lineáris egyenletrendszereket mátrixos alakban is meg tudjuk adni, a régi jelöléseket használva.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

Később majd azzal is foglalkozunk, hogy ha mátrixos alakban van megadva az egyenletrendszer, hogyan lehet automatikusan megoldást találni.

Inverz

A mátrixszorzásra nézve próbáljuk meg megkeresni az egységelemet.

Erre jó jelölt az az $(n \times n)$ -es mátrix, aminek főátlójában 1-esek állnak, a többi elem 0.

Jelöljük ezt a mátrixot E_n -nel, és nevezzük el $(n \times n)$ -es **egységmátrix**nak.

Azért lett ez elkeresztelve egységmátrixnak, mert minden olyan A, B mátrixra (ahol a szorzás elvégezhető), az alábbi igaz:

$$A \cdot E_n = A$$

$$E_n \cdot B = B$$

Azt az $(n \times k)$ -as mátrixot, aminek minden eleme nulla, az $(n \times k)$ -as **nullmátrix**nak fogjuk hívni.

Megjegyzés: Ez egy természetes elnevezés, hiszen **egy nullmátrix szorzata egy másik (megfelelő méretű) mátrixsal nullmátrix lesz. De ha két mátrix szorzata nullmátrix, abból nem következik, hogy közülük valamelyik is nullmátrix lenne.**

Mivel már vannak egységmátrixaink, ezért van értelme a mátrixok körében is „reciprokot” keresni. Ezt inverznek fogjuk nevezni.

DEFINÍCIÓ: Az $A \in T^{n \times k}$ mátrix balinverze az $A_b \in T^{k \times n}$, ha $A_b \cdot A = E_k$. Az $A \in T^{n \times k}$ mátrix jobbinverze az $A_j \in T^{k \times n}$, ha $A \cdot A_j = E_n$.

Nézzünk egy példát inverzkeresésre:

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix}$$

Balinverz keresése:

$$\begin{pmatrix} x & a \\ y & b \\ z & c \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{cases} x + 3a = 1 \\ -x + 4a = 0 \\ 2x + 5a = 0 \end{cases}$$

$$\begin{cases} y + 3b = 0 \\ -y + 4b = 0 \\ 2y + 5b = 0 \end{cases}$$

$$\begin{cases} z + 3c = 0 \\ -z + 4c = 0 \\ 2z + 5c = 1 \end{cases}$$

$$\begin{cases} 2x + 6a = 2 \\ 2x - 8a = 0 \\ 2x + 5a = 0 \end{cases}$$

$$(I) - (II): 14a = 2$$

$$a = \frac{2}{14} = \frac{1}{7}$$

$$(II) - (III): -13a = 0$$

$$a = 0$$

Az egyenletrendszer nem megoldható, nincs balinverz.

Jobbinverz keresése:

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} x & a \\ y & b \\ z & c \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{cases} x - y + 2z = 1 \\ 3x + 4y + 5z = 0 \end{cases}$$

$$\begin{cases} a - b + 2c = 0 \\ 3a + 4b + 5c = 1 \end{cases}$$

$$\begin{cases} 3x - 3y + 6z = 3 \\ 3x + 4y + 5z = 0 \end{cases}$$

$$-7y + z = 3$$

Két ismeretlen marad, ezért végtelen sok megoldás van. Ugyanez a helyzet a,b,c-re is.

Most tekintsünk egy négyzetes mátrixot példaként:

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 5 & 0 & 1 \end{pmatrix}$$

Balinverz keresése:

$$\begin{pmatrix} a & b & c \\ x & y & z \\ r & s & t \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 5 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{cases} a - b + 5c = 1 \\ 2a + b = 0 \\ 3a + 4b + c = 0 \end{cases}$$

$$\begin{cases} a - b + 5c = 1 \\ 15a + 20b + 5c = 0 \end{cases}$$

$$\begin{cases} -14a - 21b = 1 \\ 2a + b = 0 \end{cases}$$

$$\begin{cases} 2a + 3b = -\frac{1}{7} \\ 2a + b = 0 \end{cases}$$

$$2b = -\frac{1}{7}$$

$$b = -\frac{1}{14}$$

$$2a = \frac{1}{14}$$

$$a = \frac{1}{28}$$

$$\frac{1}{28} + \frac{1}{14} + 5c = 1$$

$$\frac{3}{28} + 5c = 1$$

$$5c = -\frac{25}{28}$$

$$c = -\frac{5}{28}$$

A többi ugyanúgy számolhatjuk.

Jobbinverz keresése:

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 5 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ x & y & z \\ r & s & t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{cases} a + 2x + 3r = 1 \\ -a + x + 4r = 0 \\ 5a + r = 0 \end{cases}$$

$$\begin{cases} a + 2x + 3r = 1 \\ -2a + 2x + 8r = 0 \end{cases}$$

$$\begin{cases} 3a - 5r = 1 \\ 5a + r = 0 \end{cases}$$

$$\begin{cases} 15a - 25r = 5 \\ 15a + 3r = 0 \end{cases}$$

$$-28r = 5$$

$$r = -\frac{5}{28}$$

$$3a + \frac{25}{28} = 1$$

$$3a = \frac{3}{28}$$

$$a = \frac{1}{28}$$

Ha tovább számolunk, kijön, hogy a bal- és jobbinverz megegyezik.

Kérdés, hogy véletlen-e, amit most tapasztaltunk. Ezt mindjárt megvizsgáljuk. Tegyük fel, hogy az A négyzetes mátrixnak van bal és jobbinverze is. Az eddigi jelölések mellett számítsuk ki az

$A_b \cdot A \cdot A_j$ szorzatot.

A mátrixszorzás asszociatív, azaz akárhogy zárójelezhetünk:

$$A_j = (A_b \cdot A) \cdot A_j = A_b \cdot (A \cdot A_j) = A_b$$

Tehát a négyzetes esetben a bal- és a jobbinverz (amennyiben mindkettő létezik) azonos egymással.

Később az alábbi, általánosabb tétel is bizonyítva lesz:

TÉTEL: Ha egy négyzetes mátrixnak van jobbinverze, akkor van balinverze is, és ezek egybeesnek.

Természetesen analóg módon igaz lesz, hogy ha van balinverz, akkor van jobbinverz is, és ezek egybeesnek. Ezek alapján a négyzetes mátrixoknál bal-, illetve jobbinverz helyett elég lesz simán inverzről beszélni.

Vektorterek

Veszünk egy T testet (ebből választjuk majd a skalárokat). V lesz majd az általunk vizsgált vektorok halmaza (jóval bővebb, mint a középiskolai értelemben vett vektor).

A két- és háromdimenziós vektorokra vonatkozó műveleti tulajdonságokat fogjuk összeszedni ahhoz, hogy egy új algebrai struktúrát (a vektorteret) definiáljuk.

DEFINÍCIÓ (VEKTORTÉR): A V vektorokból álló halmaz vektorteret alkot a T test fölött, ha az alábbiak teljesülnek:

Összeadás axiómái: van egy összeadás nevű művelet (két vektor között értelmezzük), két V -beli vektor összege szintén V -beli vektor.

1. Kommutatív: $\underline{a}, \underline{b} \in V$ esetén $\underline{a} + \underline{b} = \underline{b} + \underline{a}$;
2. Asszociatív: $\underline{a}, \underline{b}, \underline{c} \in V$ esetén $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c})$;
3. Létezik nullvektor (jelölés $\underline{0}$), amire teljesül, hogy minden $\underline{v} \in V$ esetén $\underline{v} + \underline{0} = \underline{v}$;
4. Minden V -beli vektornak van V -ben ellentettje, az előbbi nullvektorra nézve.

Skalárral való szorzás axiómái: minden $\underline{v} \in V$ és $\lambda \in T$ esetén $\lambda \cdot \underline{v} \in V$

1. $\lambda, \mu \in T, \underline{v} \in V$ esetén $(\lambda + \mu) \cdot \underline{v} = \lambda \cdot \underline{v} + \mu \cdot \underline{v}$
2. $\lambda \in T, \underline{a}, \underline{b} \in V$ esetén $\lambda \cdot (\underline{a} + \underline{b}) = \lambda \cdot \underline{a} + \lambda \cdot \underline{b}$
3. $\lambda, \mu \in T, \underline{v} \in V$ esetén $\lambda \cdot \mu \cdot \underline{v} = \lambda \cdot (\mu \cdot \underline{v}) = (\lambda \cdot \mu) \cdot \underline{v}$

4. 1-gyel jelölve a T egységelemét, teljesül, hogy minden $\underline{v} \in V$ -re $1 \cdot \underline{v} = \underline{v}$

Megjegyzés: az előbb leírt 8 tulajdonság alkotja az ún. **vektortér-axiómákat**.

Példák vektortérre:

1. két- és háromdimenziós vektorok $\mathbf{R}$ felett
2. n -dimenziós valós koordinátájú vektorok szintén $\mathbf{R}$ felett (jelölés: $\mathbf{R}^n$)
3. n -dimenziós komplex koordinátájú vektorok $\mathbf{C}$ felett (jelölés: $\mathbf{C}^n$)
4. Az $(n \times k)$ -as valós számokkal kitöltött mátrixok $\mathbf{R}$ felett (jelölés: $\mathbf{R}^{n \times k}$)
5. Az $(n \times k)$ -as komplex számokkal kitöltött mátrixok $\mathbf{C}$ felett (jelölés: $\mathbf{C}^{n \times k}$)

A szóhasználattal kapcsolatos problémák:

1. a skaláris szorzás nem keverendő össze a skalárral való szorzással
2. Nem szabad elfeledkeznünk arról, hogy egy halmaz csak egy bizonyos test felett alkothat vektorteret, ezért nem elég annyit mondani, hogy pl. a „kétdimenziós vektorok tere”, hanem azt is hozzá kell tennünk, hogy melyik test felett tekintjük ezt

DEFINÍCIÓ (LINEÁRIS KOMBINÁCIÓ): A T test feletti V vektortérben adott $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k \in V$ vektorok egy lineáris kombinációja $\lambda_1 \cdot \underline{v}_1 + \lambda_2 \cdot \underline{v}_2 + \dots + \lambda_k \cdot \underline{v}_k \in V$ (ahol $\lambda_1, \lambda_2, \dots, \lambda_k \in T$).

Könnyen látható, hogy ha $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$, akkor a lineáris kombináció $\underline{0}$ lesz.

Fordítva nehezebb a dolog: próbáljuk meg azt vizsgálni, hogy adott V -beli vektorok milyen lineáris kombinációi állítják elő a $\underline{0}$ -t.

Sokszor vannak olyan előállításai a $\underline{0}$ -nak, ahol nem csak a csupa 0 választása működik, pl.

$$1 \cdot (1,2) + 2 \cdot (0,4) + (-1) \cdot (1,1,0) = (0,0)$$

Ha a nullvektort úgy állítjuk elő adott vektorok lineáris kombinációjaként, hogy minden skalárt 0-nak választunk, akkor **triviális előállításról** beszélünk.

DEFINÍCIÓ: Egy adott vektortérben vektorok egy kijelölt halmaza **lineárisan független** vektorrendszert alkot, ha az ő lineáris kombinációjaként a $\underline{0}$ csak triviálisan állítható elő.

Ha van a triviálistól eltérő előállítás, akkor a vektorrendszer **összefüggő**.

DEFINÍCIÓ: Egy adott vektortérben vektorok egy kijelölt halmaza **generátorrendszert** alkot, ha az ő lineáris kombinációjaként a vektortér összes eleme előállítható.

DEFINÍCIÓ: Ha egy vektortérben egy vektorrendszer **lineárisan független és generátorrendszer**, akkor **bázisnak** nevezzük.

Nézzünk néhány gyakorlati példát. Először is vizsgáljuk azt, hogy milyen módon dönthető el egy vektorrendszerrel (egy adott vektortéren belül), hogy a benne szereplő vektorok lineárisan függetlenek-e. Mondjuk nézzük a térvektorok vektorterét a valós test felett, ahol el akarjuk dönteni, hogy lineárisan független-e a következő vektorrendszer:

$(1,2,3), (-4,7,3), (2, -5,8).$

Az alábbi egyenlet megoldásait kell vizsgálnunk:

$$x \cdot (1,2,3) + y \cdot (-4,7,3) + z \cdot (2, -5,8) = (0,0,0)$$

Ez az egyenlet ekvivalens az alábbi lineáris egyenletrendszerrel:

$$\begin{cases} x - 4y + 2z = 0 \\ 2x + 2y - 5z = 0 \\ 3x + 3y + 8z = 0 \end{cases}$$

Gauss-eliminációt alkalmazva megoldjuk az egyenletrendszert:

$$\begin{cases} 6x - 24y + 12z = 0 \\ 6x + 21y - 15z = 0 \\ 6x + 6y + 16z = 0 \end{cases}$$

$$I - II.: -45y + 27z = 0$$

$$II - III.: 15y - 31z = 0$$

$$\begin{cases} I'. 15y - 9z = 0 \\ II'. 15y - 31z = 0 \end{cases}$$

$$I' - II'. 22z = 0$$

$$x = 0$$

$$y = 0$$

Tehát a példabeli vektoraink lineárisan függetlenek.

Most nézzünk példát generátorrendszerre is. Generátorrendszert alkotnak a síkvektorok körében (a valós test felett) pl. az $(1,0), (0,1)$ vektorok. Egyébként bármely két nem párhuzamos síkvektor is jó lenne példának. Vegyük észre, hogy ez a példa bázist is szolgáltat.

Általánosítsuk ezt a példánkat, amikor több koordináta van (mondjuk $\mathbf{R}^N$ -ben). Az $\mathbf{R}^N$ -beli **standard bázis** a következőképpen néz ki:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

A standard bázisról könnyű ellenőrizni, hogy tényleg rendelkezik a bázistulajdonságokkal.

Bázisok elemszáma, a dimenzió fogalma

Egy vektortérnek általában sok bázisa van, de ezeknek lesz egy fontos közös tulajdonsága.

TÉTEL: Adott vektortérben bármely két bázis azonos elemszámú.

Megjegyzés: ezt a közös elemszámot nevezzük a vektortér **dimenziójának**.

SEGÉDTÉTEL: Ha egy lineáris egyenletrendszernek (a valós vagy a komplex testben) pontosan egy megoldása van, akkor az egyenleteinek száma nem lehet kisebb, mint az ismeretlenek száma.

Bizonyítás:

Azt fogjuk belátni, hogy amennyiben egy lineáris egyenletrendszerek pontosan egy megoldása van, akkor a Gauss-elimináció során minden egyes lépésben az ismeretlenek száma pontosan eggyel csökken. Tegyük fel ugyanis indirekt, hogy a Gauss-elimináció során valamelyik lépésben legalább kettővel csökkent az ismeretlenek száma. Legyen " x_1 " és " x_2 " két kieső ismeretlen. Ha ez a kettő egyszerre kiesik, akkor minden egyenletben az " x_1 " és " x_2 " által meghatározott rész $\lambda \cdot (\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2)$, ahol λ valamilyen konstans (lehet 0 is), α_1 és α_2 pedig valamilyen nullától különböző konstansok (hiszen csak így eshet ki egyszerre " x_1 " és " x_2 " is). Ekkor azonban $(\alpha_1 \cdot x_1 + \alpha_2 \cdot x_2)$, egy új ismeretlennek tekinthető (mondjuk legyen ez y), és így az eredeti egyenletrendszernek egynél több megoldása lenne, ami ellentmondás.

Ha pontosan egy megoldás van, akkor tehát a Gauss-elimináció végére egyetlen ismeretlenünk marad (aminek az értékét itt le tudjuk olvasni), és legalább egy egyenletünk, amiben csak ez az utolsó ismeretlen szerepel. Mivel az elimináció minden egyes lépése során legalább eggyel csökkent az egyenletek száma, és pontosan eggyel az ismeretleneké, ezért világos, hogy eredetileg legalább annyi egyenlet volt, mint amennyi ismeretlen. ■

TÉTEL: Adott vektortérben bármely két bázis azonos elemszámú (csak arra az esetre bizonyítjuk, amikor a vektortérben van véges sok elemű bázis).

Bizonyítás:

Legyen $\underline{f}_1, \underline{f}_2, \dots, \underline{f}_k$ egy véges sok elemből álló bázis az adott vektortérben, és legyen $\underline{g}_1, \underline{g}_2, \dots, \underline{g}_n$ egy másik bázis. Megmutatjuk, hogy $k \geq n$. Tekintsük a

$\underline{0} = \lambda_1 \cdot \underline{f}_1 + \lambda_2 \cdot \underline{f}_2 + \dots + \lambda_k \cdot \underline{f}_k$ egyenletet, ahol a λ skalárok az ismeretlenek. A bázis definíciója miatt az egyetlen megoldás $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$. Írjuk le az előző egyenleteinket a $\underline{g}$ vektorok segítségével. Állítsuk elő az $\underline{f}$ -típusú vektorok mindegyikét a $\underline{g}$ -típusúak lineáris kombinációjaként, legyen: $\underline{f}_i = a_{i1} \cdot \underline{g}_1 + a_{i2} \cdot \underline{g}_2 + \dots + a_{in} \cdot \underline{g}_n$.

Ezekkel a jelölésekkel

$$\underline{0} = (\lambda_1 \cdot a_{11} + \lambda_2 \cdot a_{21} + \lambda_3 \cdot a_{31} + \dots + \lambda_k \cdot a_{k1}) \cdot \underline{g}_1 +$$

$$\begin{aligned}
& +(\lambda_1 \cdot a_{12} + \lambda_2 \cdot a_{22} + \lambda_3 \cdot a_{32} + \dots + \lambda_k \cdot a_{k2}) \cdot \underline{g_2} + \dots \\
& \dots \\
& \dots \\
& \dots \\
& \dots \\
& +(\lambda_1 \cdot a_{1n} + \lambda_2 \cdot a_{2n} + \lambda_3 \cdot a_{3n} + \dots + \lambda_k \cdot a_{kn}) \cdot \underline{g_n}
\end{aligned}$$

Mivel a $\underline{g}$ -típusú vektorok bázist alkotnak, ezért itt a zárójelekben az együtthatók mind nullák. Tehát a zárójelekből kiolvashatunk egy n egyenletből álló, k -ismeretlenes lineáris egyenletrendszer, aminek pontosan egy megoldása van. Így a segédétel miatt $k \geq n$.

Viszont az $\underline{f}$ -típusú vektorok és a $\underline{g}$ -típusú vektorok szerepcseréjével ugyanígy kijön az is, hogy $k \leq n$. Ebből következik, hogy $k = n$. ■

Determinánsok

Négyzetes mátrixokhoz fogunk rendelni bizonyos értékeket. Ehhez szükségünk lesz a permutáció fogalmára, és a permutációkat is majd két csoportba fogjuk sorolni.

DEFINÍCIÓ (PERMUTÁCIÓ): *Megkülönböztethető elemek egy sorbarendezése.*

Könnyen ellenőrizhető, hogy n db elemnek $n!$ db permutációja van ($n! = 1 \cdot 2 \cdot \dots \cdot n$).

Szükségünk lesz az inverzió fogalmára:

DEFINÍCIÓ (INVERZIÓ): *n db valós szám egy permutációjában két elem inverzióban áll egymással (inverziós párt alkot), ha a nagyobb megelőzi a kisebbet.*

Példa: $n = 5$, $H = \{1, 2, 3, 4, 5\}$, legyen egy permutáció $4, 1, 5, 3, 2$.

Inverziós párok: $\{4, 1\}, \{5, 3\}, \{3, 2\}, \{5, 2\}, \{4, 2\}, \{4, 3\}$.

Az inverziók esetében arra leszünk kíváncsiak, hogy egy adott permutációban páros számú vagy páratlan számú inverzió (azaz inverziós pár) szerepel. Ez alapján azokat a permutációkat, amikben **páros sok inverzió van, párosnak, amikben páratlan sok inverzió van, páratlannak** nevezzük.

Az előbb bevezetett fogalmakat négyzetes mátrixok esetén fogjuk használni.

Legyen adott egy $(n \times n)$ -es mátrix. Ebből válasszunk ki n db mezőt úgy, hogy semelyik kettő se essen azonos sorba, vagy oszlopba. Ha ezt úgy csináljuk, hogy soronként haladunk és mindig egy-egy elemet választunk ki, akkor minden egyes kiválasztott mező után egy oszlopot is elhasználunk. Kicsit átfogalmazva ez azt jelenti, hogy kiválasztjuk az $a_{1,p(1)}, a_{2,p(2)}, \dots, a_{n,p(n)}$ elemeket a mátrixból, ahol $p(1), p(2), \dots, p(n)$ egy permutációja az $\{1, 2, 3, \dots, n\}$ halmaz elemeinek. Haladjunk végig az

oszlopok összes ($n!$ db) permutációján, mindig szorozzuk össze a kiválasztott $a_{1,p(1)}, a_{2,p(2)}, \dots, a_{n,p(n)}$ elemeket, a kapott szorzatot szorozzuk 1-gyel, ha az aktuális (oszlop)permutáció páros volt, egyébként (-1) -gyel. Az így kapott előjellel ellátott értékeket összeadva kapjuk meg a determináns értékét.

Példák:

1. $n = 1, A = (5) \rightarrow \det(A) = \det(5) = 5$

2. $n = 2, A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix} \rightarrow 1 \cdot 1 \cdot 5 + (-1) \cdot 3 \cdot 4 = -7$

3. $n = 3, A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & -1 \\ 2 & 5 & 1 \end{pmatrix}$

$$1 \cdot (1 \cdot 4 \cdot 1) + (-1) \cdot (1 \cdot (-1) \cdot 5) + (-1) \cdot ((-1) \cdot 3 \cdot 1) + 1 \cdot ((-1) \cdot (-1) \cdot 2) + 1 \cdot (2 \cdot 3 \cdot 5) + (-1) \cdot (2 \cdot 4 \cdot 2) = 4 + 5 + 3 + 2 + 30 - 16 = 28$$

Látszólag az oszlopoknak kiemelt szerepe van a determináns definíciójában. Nézzük meg, mi lenne, ha sorokra csinálnánk ugyanezt. Megmutatjuk, hogy akkor is ugyanaz az érték jön ki.

Ki kell választanunk $a_{1,p(1)}, a_{2,p(2)}, \dots, a_{n,p(n)}$ elemeket a mátrixból, ahol $p(1), p(2), \dots, p(n)$ az $\{1, 2, \dots, n\}$ elemeinek egy permutációja. Ha mondjuk $p(i)$ és $p(j)$ inverzióban áll egymással, akkor $i < j$ és $p(i) > p(j)$. Ha most oszloponként felírnánk a kiválasztott elemeket, akkor az eredeti értelemben is egy inverziós párt kapnánk. Ugyanígy, ha $p(i)$ és $p(j)$ nem áll inverzióban, akkor $p(i) < p(j)$ és $i < j$. Ilyenkor ismét oszloponként felírva a kiválasztott elemeket, az eredeti értelemben sem alkotnának inverziós párt.

Tehát az “új” értelmezésünkénél (ahol a sorokat permutáltuk) minden permutációhoz ugyanannyi inverzió tartozik, mint a régi (oszlopos) értelemben tartozott volna, ezért a determináns értéke ugyanannyi lenne.

Determinánsok tulajdonságai:

1. Ha a négyzetes mátrixot λ -vel szorzunk, akkor a determinánsa λ^n -szerese lesz (n a sorok és oszlopok száma).

Bizonyítás: Ez nyilvánvalóan igaz, hiszen minden összeadandó tag λ^n -nel szorzódik.

2. Egy determináns kéttagú összegre beosztható sor- vagy oszlopvektorai szerint az alábbi módon:

$$\det \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \text{vmi} \quad x_3 + y_3 \quad \text{még vmi} \\ \vdots \\ x_n + y_n \end{pmatrix} = \det \begin{pmatrix} x_1 \\ x_2 \\ \text{vmi} \quad x_3 \quad \text{még vmi} \\ \vdots \\ x_n \end{pmatrix} + \det \begin{pmatrix} y_1 \\ y_2 \\ \text{vmi} \quad y_3 \quad \text{még vmi} \\ \vdots \\ y_n \end{pmatrix}$$

Ez a disztributivitás miatt igaz (ki lehet emelni), azaz: $(x + y) \cdot z = x \cdot z + y \cdot z$

Megjegyzés: a 2. tulajdonság persze hasonlóan sorokra nézve is igaz.

3. Ha a négyzetes mátrixban van két azonos sor (vagy két azonos oszlop), akkor a determinánsa nulla.

Bizonyítás:

Legyen két azonos oszlopunk. Azt fogjuk megnézni, hogy ha a bal oldali kitüntetett oszlopunkból veszünk ki x_i -t, a jobb oldali kitüntetett oszlopból x_j -t, valamint még $(n - 2)$ db megfelelő elemet az előírások szerint, akkor az ebből kapott előjelezett szorzat hogyan viszonyul ahhoz, amit abból kapunk, ha a bal oldaliból x_j -t, a jobb oldaliból x_i -t, és a maradék $(n - 2)$ db elemet pedig ugyanúgy választjuk. Ki fog derülni, hogy két ilyen előjelezett szorzat egymás ellentettje.

Ahhoz, hogy az előzőt belássuk, elég megmutatni, hogy két elem cseréjekor egy permutáció paritása megváltozik. Ha két szomszédos elemet cserélünk ki, akkor az inverziók száma 1-gyel változik, tehát ekkor igaz az állítás. Ha két nem szomszédos elemet cserélünk ki, mondjuk a-t és b-t, és közéjük k db elem esik: összesen $k + (k + 1)$ db szomszédos párcserével érhető el, hogy a és b helyet cseréljen. Mivel $k + (k + 1) = 2k + 1$ páratlan, ezért a paritás megváltozik.

4. Ha egy négyzetes mátrixban felcserélünk két sort (vagy felcserélünk két oszlopot), akkor a determináns értéke az ellentettjére változik.

Bizonyítás:

Cseréljünk ki mondjuk két sort: Az előbbieik alapján egy adott (x_i, y_j) párra nézve (amik különböző sorba esnek) a sorcsere annyit jelent, hogy minden olyan szorzatban, ahol x_i és y_j is szerepel, az előjelezés megváltozik. Ezért a determináns is az ellentettjére változik a csere után.

5. Ha egy négyzetes mátrix egy sorához hozzáadjuk egy másik sor konstansszorosát (vagy egy oszlopához egy másik oszlop konstansszorosát), a determináns nem változik.

Bizonyítás:

Adjuk hozzá (az egyszerűbb jelölés kedvéért) az első oszlophoz a második λ -szorosát:

$$\begin{aligned} \det(\underline{\sigma}_1 + \lambda \underline{\sigma}_2, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n) &= \det(\underline{\sigma}_1, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n) + \det(\lambda \underline{\sigma}_2, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n) = \\ &= \det(\underline{\sigma}_1, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n) + \lambda \cdot \det(\underline{\sigma}_2, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n) = \\ &= \det(\underline{\sigma}_1, \underline{\sigma}_2, \underline{\sigma}_3, \dots, \underline{\sigma}_n). \blacksquare \end{aligned}$$

Kifejtés és ferde kifejtés

ELŐJELES ALDETERMINÁNS FOGALMA: A négyzetes mátrix i -edik sorát és j -edik oszlopát kivágjuk. Az így kapott (eggyel kisebb méretű) négyzetes mátrix determinánsát $(-1)^{i+j}$ -gyel szorozzuk. Az így kapott előjeles érték az eredeti négyzetes mátrix i -edik sorának j -edik oszlopához tartozó előjeles aldeterminánsa.

Jelölés: Ha A az eredeti mátrix, akkor az i -edik sor j -edik eleméhez tartozó előjeles aldeterminánst A_{ij} -vel jelöljük.

Példa:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 9 & 0 \\ 7 & 8 & 6 \end{pmatrix}$$

$$A_{11} = (-1)^{1+1} \cdot \det \begin{pmatrix} 9 & 0 \\ 8 & 6 \end{pmatrix}$$

$$A_{12} = (-1)^{1+2} \cdot \det \begin{pmatrix} 2 & 0 \\ 7 & 6 \end{pmatrix}$$

KIFEJTÉSI TÉTEL: Tetszőleges négyzetes mátrix determinánsa kiszámolható úgy, hogy egy adott sor vagy oszlop szerint kifejtünk, ami azt jelenti, hogy az adott sorban vagy oszlopban minden egyes elemet megszorozunk a neki megfelelő előjeles aldeterminánssal, és a kapott értékeket összeadjuk.

Bizonyítás (kifejtési tétel):

Az első sorra fogjuk csak belátni, a többi sorra, illetve oszlopra ugyanúgy menne a bizonyítás. A szokásos jelöléseket használjuk. Be kell látni, hogy $\det A = a_{11} \cdot A_{11} + a_{12} \cdot A_{12} + \dots + a_{1n} \cdot A_{1n}$. Megmutatjuk, hogy mindkét oldalon ugyanazok a szorzatok szerepelek, méghozzá ugyanolyan előjellel. Legyen az $a_{1k} \cdot A_{1k}$ az egyik összeadandó tag a jobb oldalról. Itt olyan szorzatok fordulnak

elő, amikben $a_{1k}, a_{2,q(1)}, a_{3,q(2)}, \dots, a_{n,q(n-1)}$ szorzótényezők szerepelnek, ahol $q(1), q(2), \dots, q(n-1)$ az $\{1, 2, \dots, n\} \setminus \{k\}$ halmaz elemeinek egy permutációja áll, és az ilyen szorzatokat kell előjelezni. Tekintsük a $q(1), q(2), \dots, q(n-1)$ és a $k, q(1), q(2), \dots, q(n-1)$ permutációkat. A második esetben $(k-1)$ -gyel több inverziós pár van, mint az első esetben. Minden szorzathoz tartozik még egy előjel is, ami $(-1)^{k+1}$. A $(k-1)$ és $(k+1)$ paritása azonos, ezért az előjelezés is mindkét oldalon megegyezik, és ezt kellett belátni. ■

Példa a kifejtésre:

$$\det \begin{pmatrix} 1 & 2 & -3 & 5 \\ 8 & 7 & 6 & 1 \\ 2 & 3 & 5 & 9 \\ 1 & -2 & 3 & 7 \end{pmatrix} = 1 \cdot (-1)^{1+1} \cdot \det \begin{pmatrix} 7 & 6 & 1 \\ 3 & 5 & 9 \\ -2 & 3 & 7 \end{pmatrix} + 2 \cdot (-1)^{1+2} \cdot \det \begin{pmatrix} 8 & 6 & 1 \\ 2 & 5 & 9 \\ 1 & 3 & 7 \end{pmatrix} +$$

$$+ (-3) \cdot (-1)^{1+3} \cdot \det \begin{pmatrix} 8 & 7 & 1 \\ 2 & 3 & 9 \\ 1 & -2 & 7 \end{pmatrix} + 5 \cdot (-1)^{1+4} \cdot \det \begin{pmatrix} 8 & 7 & 6 \\ 2 & 3 & 5 \\ 1 & -2 & 3 \end{pmatrix}.$$

FERDE KIFEJTÉSI TÉTEL: Ha két különböző sort (vagy oszlopot) veszünk és az egyik sor (vagy oszlop) elemeit rendre végigszorozzuk a másik megfelelő elemeihez tartozó előjeles aldeterminánsokkal, és a kapott szorzatokat összeadjuk, akkor eredményként 0-át kapunk.

Ferde kifejtési tétel bizonyítása:

Vegyük az A négyzetes mátrixot, és módosítsuk úgy, hogy a j -edik sorába átmásoljuk az i -edik sorát. Legyen a módosítás után kapott mátrix A' .

$\det A' = 0$, mivel van két azonos sor az A' mátrixban.

Ugyanakkor a j -edik sor szerint kifejtve $\det A' = a'_{j1} \cdot A'_{j1} + a'_{j2} \cdot A'_{j2} + \dots + a'_{jn} \cdot A'_{jn} =$

$$= a'_{j1} \cdot A_{j1} + a'_{j2} \cdot A_{j2} + \dots + a'_{jn} \cdot A_{jn} = a_{j1} \cdot A_{j1} + a_{j2} \cdot A_{j2} + \dots + a_{jn} \cdot A_{jn}. \blacksquare$$

Példa a ferde kifejtésre:

$$\det \begin{pmatrix} 1 & -2 & 1 \\ 5 & -4 & 3 \\ -2 & 1 & 2 \end{pmatrix} = 1 \cdot (-1)^{1+1} \cdot \det \begin{pmatrix} 5 & 3 \\ -2 & 2 \end{pmatrix} + 3 \cdot (-1)^{2+2} \cdot \det \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix} + 2 \cdot (-1)^{3+2} \cdot \det \begin{pmatrix} 1 & 1 \\ 5 & 3 \end{pmatrix} =$$

$$= -(10 - (-6)) + 3 \cdot (2 - (-2)) - 2 \cdot (3 - 5) = 0$$

DETERMINÁNSOK SZORZÁSTÉTELE, INVERZ LÉTEZÉSE

DETERMINÁNSOK SZORZÁSTÉTELE: Azonos méretű A és B négyzetes mátrixokra teljesül, hogy $\det(AB) = (\det A) \cdot (\det B)$.

Első bizonyítás (trükkös):

Vegyük először észre azt, hogy ha $C \in R^{r \times r}$, $D \in R^{s \times s}$ négyzetes mátrixok, akkor

$$\det \begin{pmatrix} C & \text{bármilyen} \\ 0 & D \end{pmatrix} = \det C \cdot \det D, \text{ és hasonlóan } \det \begin{pmatrix} C & 0 \\ \text{bármilyen} & D \end{pmatrix} = \det C \cdot \det D.$$

Legyenek most A és B azonos méretű, $(n \times n)$ -es négyzetes mátrixok.

$\det A \cdot \det B = \det \begin{pmatrix} A & 0 \\ -E_n & B \end{pmatrix}$ az észrevételünk miatt. Determinánstartó átalakításokkal nullázzuk ki a jobb alsó blokkot (tehát a B helyét).

Mivel $B \cdot (-E_n) + B = 0$, ezért a jobb felső blokk az átalakítások után $0 + A \cdot B = AB$ lesz.

Az új mátrix $\begin{pmatrix} A & AB \\ -E_n & 0 \end{pmatrix}$, ebből n db oszlopcserével kapunk egy olyan mátrixot, hogy $\begin{pmatrix} AB & A \\ 0 & -E_n \end{pmatrix}$,

aminek determinánsa $\det(-E_n) \cdot \det(AB) = (-1)^n \cdot \det(AB)$.

Minden oszlopcserénél előjelváltás történt, ezért $\det(AB) = \det(A) \cdot \det(B)$.

Második bizonyítás (közvetlen számolással):

Az $\{1, 2, \dots, n\}$ halmaz elemeinek tetszőleges $p(1), p(2), \dots, p(n)$ permutációja esetén legyen

$e_{p(1), p(2), \dots, p(n)}$ értéke $(+1)$ vagy (-1) attól függően, hogy a szóban forgó permutáció páros, illetve páratlan.

Ez alapján a $\det A \cdot \det B$ szorzat úgy írható fel, hogy

$$\left(\sum e_{u(1), u(2), \dots, u(n)} \cdot a_{1, u(1)} \cdot a_{2, u(2)} \cdot \dots \cdot a_{n, u(n)} \right) \cdot \left(\sum e_{t(1), t(2), \dots, t(n)} \cdot b_{1, t(1)} \cdot b_{2, t(2)} \cdot \dots \cdot b_{n, t(n)} \right),$$

ahol a szummák az összes lehetséges permutáción futnak végig.

Legyen $C = AB$. Ugyanúgy, mint fent is láttuk, a szokásos jelölésekkel $\det C$ értéke éppen

$$\sum e_{r(1), r(2), \dots, r(n)} \cdot c_{1, r(1)} \cdot c_{2, r(2)} \cdot \dots \cdot c_{n, r(n)},$$

ahol a szumma szintén a lehetséges permutációk szerint fut végig.

Mivel $C=AB$, ezért az előző éppen

$$\sum e_{r(1), r(2), \dots, r(n)} \cdot \left(\sum_{i=1}^n a_{1, i} \cdot b_{i, r(1)} \right) \cdot \left(\sum_{i=1}^n a_{2, i} \cdot b_{i, r(2)} \right) \cdot \dots \cdot \left(\sum_{i=1}^n a_{n, i} \cdot b_{i, r(n)} \right)$$

alakban írható. Itt, ha mindent mindennel összeszorozunk, akkor azt kapjuk, hogy az

$$e_{r(1),r(2),\dots,r(n)} \cdot a_{1,u(1)} \cdot b_{u(1),r(1)} \cdot a_{2,u(2)} \cdot b_{u(2),r(2)} \cdot \dots \cdot a_{n,u(n)} \cdot b_{u(n),r(n)}$$

típusú kifejezéseket kell összeadnunk, ahol az $(r(1), r(2), \dots, r(n))$ -nel végigfutunk az $\{1, 2, \dots, n\}$ halmaz elemeinek összes permutációján, és az $u(1), u(2), \dots, u(n)$ mindegyike végigfut 1-től n -ig. Vizsgáljuk meg, mi történik akkor, ha az $u(i)$ számok között van két egyenlő, legyen mondjuk a könnyebb indexelés kedvéért $u(1)=u(2)=u$.

Ebben az esetben, rögzített $r(1), r(2), \dots, r(n)$ permutáció mellett

$$e_{r(1),r(2),\dots,r(n)} \cdot a_{1,u} \cdot b_{u,r(1)} \cdot a_{2,u} \cdot b_{u,r(2)} \cdot a_{3,u(3)} \cdot b_{u(3),r(3)} \cdot \dots \cdot a_{n,u(n)} \cdot b_{u(n),r(n)}$$

értéke éppen ellentettje lesz

$$e_{r(2),r(1),\dots,r(n)} \cdot a_{1,u} \cdot b_{u,r(1)} \cdot a_{2,u} \cdot b_{u,r(2)} \cdot a_{3,u(3)} \cdot b_{u(3),r(3)} \cdot \dots \cdot a_{n,u(n)} \cdot b_{u(n),r(n)}$$

értékének. Ugyanígy járhatnánk el bármely két $u(i)$ egyenlőségének esetében.

Ez tehát azt jelenti, hogy $\det(AB)$ kiszámításakor azok az összeadandó tagok mind nullaszer szerepelnek, melyeknél valamelyik két $u(i)$ megegyezik. Ezért $\det(AB)$ úgy is írható, mint az

$$e_{r(1),r(2),\dots,r(n)} \cdot a_{1,u(1)} \cdot b_{u(1),r(1)} \cdot a_{2,u(2)} \cdot b_{u(2),r(2)} \cdot \dots \cdot a_{n,u(n)} \cdot b_{u(n),r(n)}$$

típusú kifejezések összege, ahol $(r(1), r(2), \dots, r(n))$ -nel, és az $(u(1), u(2), \dots, u(n))$ -nel is végigfutunk az $\{1, 2, \dots, n\}$ halmaz elemeinek összes permutációján.

Visszatérve $\det A \cdot \det B$ -hez, ott elvégezve a szorzást (mindent mindennel szorzunk),

$$e_{u(1),u(2),\dots,u(n)} \cdot a_{1,u(1)} \cdot a_{2,u(2)} \cdot \dots \cdot a_{n,u(n)} \cdot e_{t(1),t(2),\dots,t(n)} \cdot b_{1,t(1)} \cdot b_{2,t(2)} \cdot \dots \cdot b_{n,t(n)}$$

típusú összeadandó tagokat kapunk, ahol $(t(1), t(2), \dots, t(n))$ -nel, és az $(u(1), u(2), \dots, u(n))$ -nel is végigfutunk az $\{1, 2, \dots, n\}$ halmaz elemeinek összes permutációján. Ezt átrendezve úgy, ahogy a $\det(AB)$ kiszámolásánál láttuk, kapjuk, hogy ezek az összeadandó tagok éppen

$$e_{u(1),u(2),\dots,u(n)} \cdot e_{t(1),t(2),\dots,t(n)} \cdot a_{1,u(1)} \cdot b_{u(1),t(u(1))} \cdot a_{2,u(2)} \cdot b_{u(2),t(u(2))} \cdot \dots \cdot a_{n,u(n)} \cdot b_{u(n),t(u(n))}$$

alakúak lesznek.

Annyit kell már csak belátnunk, hogy amennyiben

$$t(u(1))=r(1), t(u(2))=r(2), \dots, t(u(n))=r(n) \text{ teljesül, akkor}$$

$e_{u(1),u(2),\dots,u(n)} \cdot e_{t(1),t(2),\dots,t(n)} = e_{r(1),r(2),\dots,r(n)}$, hiszen ez éppen azt jelenti, hogy mindkét oldalon minden összeadandó tag (ami nem esett ki) ugyanolyan típusú, és ugyanolyan előjellel szerepel.

Azt kellene tehát belátni, hogy az $\{1, 2, \dots, n\}$ halmaz elemeinek tetszőleges $(t(1), t(2), \dots, t(n))$ és $(u(1), u(2), \dots, u(n))$ permutációja esetén teljesül, hogy

$$e_{u(1),u(2),\dots,u(n)} \cdot e_{t(1),t(2),\dots,t(n)} = e_{t(u(1)),t(u(2)),\dots,t(u(n))}$$

Amennyiben $t(1)=1, t(2)=2, \dots, t(n)=n$, akkor ez könnyen látható módon teljesül. Azt fogjuk megmutatni, hogy ha a $(t(1), t(2), \dots, t(n))$ permutációt úgy módosítjuk, hogy két elemet cserélünk, ez

a tulajdonság megmarad. Ez már bizonyítja az állításunk helyességét, hiszen alkalmas cserékkel bármely permutációhoz eljuthatunk.

Amennyiben a $t(i)$ és $t(j)$ helyét megcseréljük ($i \neq j$), akkor a bizonyítandó bal oldalán álló szorzat az ellentettjére változik. Egyértelműen létezik két olyan különböző a és b szám az $\{1, 2, \dots, n\}$ halmazban, melyekre $i = u(a)$ és $j = u(b)$. Ekkor a jobb oldalhoz tartozó permutációban $t(u(a))$ és $t(u(b))$ helye felcserélődik, ezért a jobb oldal is ellentettjére változik. Ezzel a bizonyításunk teljes. ■

Invertálhatóság

Visszakanyarodunk kicsit a négyzetes mátrixok inverzéhez. Belátjuk az alábbi tételt:

TÉTEL: Egy négyzetes mátrix pontosan akkor invertálható, ha determinánsa nem 0.

Bizonyítás:

Két irányt kell igazolnunk.

I. Ha $\det A \neq 0$, akkor a szokásos jelölésekkel készítsük el az

$$A^* = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots & A_{n1} \\ A_{12} & A_{22} & A_{32} & & \\ A_{13} & A_{23} & & \ddots & \\ \vdots & & & & \\ A_{1n} & & & & A_{nn} \end{pmatrix}$$

mátrixot, melynek elemei A előjeles aldeterminánsai, szokásoshoz képest fordított indexeléssel. Ekkor

$$\begin{aligned} A \cdot A^* &= \begin{pmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{n1} \\ a_{12} & a_{22} & a_{32} & & \\ a_{13} & a_{23} & & \ddots & \\ \vdots & & & & \\ a_{1n} & & & & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots & A_{n1} \\ A_{12} & A_{22} & A_{32} & & \\ A_{13} & A_{23} & & \ddots & \\ \vdots & & & & \\ A_{1n} & & & & A_{nn} \end{pmatrix} = \\ &= \begin{pmatrix} \det A & 0 & 0 & 0 & 0 \\ 0 & \det A & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \det A & 0 \end{pmatrix} \end{aligned}$$

Az $(A \cdot A^*)$ mátrix főátlójában a kifejtési tétel miatt végig $\det A$ szerepel, míg a ferde kifejtési tétel miatt ugyanebben a mátrixban a többi elem mind nulla. Ez alapján

$\left(\frac{1}{\det A} \cdot A^*\right)$ mátrix éppen az A inverze.

II. Tegyük fel, hogy az A négyzetes mátrix invertálható, legyen az inverze A^{-1} , ekkor tehát

$E_n = A \cdot A^{-1}$. Vegyük mindkét oldal determinánsát, és alkalmazzuk a determinánsok

szorzástételét: $\det(E_n) = \det(A) \cdot \det(A^{-1})$, amiből világos, hogy $\det A \neq 0$. ■

NÉHÁNY EKVIVALENS TULAJDONSÁG INVERTÁLHATÓ MÁTRIXOKRA

TÉTEL: Egy A négyzetes mátrixra az alábbi állítások egyenértékűek (ekvivalensek):

- (i) $\det A \neq 0$
- (ii) A sorai függetlenek
- (iii) az olyan $\underline{x}$ oszlopvektorokra, amelyeknek annyi koordinátája van, ahány sora van A -nak, az $A\underline{x} = \underline{0}$ egyenletnek $\underline{x}$ -ben pontosan egy megoldása van (az itteni $\underline{0}$ is ugyanolyan hosszú, mint $\underline{x}$).
- (iv) A invertálható

Bizonyítás:

Azt már korábbról tudjuk, hogy (i) és (iv) ekvivalens. Azt fogjuk belátni, hogy (i), (ii) és (iii) ekvivalens. Ez úgy fog történni, hogy bebizonyítjuk, hogy a (iii)-asból következik az (i)-es, az (i)-esből a (ii)-es, illetve a (ii)-esből a (iii)-as.

Első lépés ((iii)-ból következik (i)):

Először induljunk ki tehát a (iii)-asból. Próbáljuk meg ezt lefordítani az egyenletrendszerek nyelvére (a jelölések a szokásosak):

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Ha (iii) teljesül, akkor az alábbi egyenletrendszernek van (pontosan egy) megoldása:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = 0 \end{cases}$$

Akármilyen konstansokat is írunk a jobb oldalra, a módosított egyenletrendszernek ugyanúgy létezne megoldása (hiszen a Gauss-elimináció az együtthatókra nézve ugyanúgy fut le):

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$$

Ebből következik, hogy A invertálható, hiszen a jobb oszlopvektort rendre a standard bázis elemeinek választhatjuk, és mindig kapunk megoldást. Viszont azt tudjuk, hogy az invertálható négyzetes mátrixoknak nem lehet nulla a determinánsa, ezért a (iii) maga után vonja (i)-et.

Második lépés ((i)-ből következik (ii)):

Indirekt tegyük fel, hogy van olyan A négyzetes mátrix, amire $\det A \neq 0$, és a mátrix sorai összefüggők. Determinánstartó átalakításokkal (itt sorhoz adva másik sor konstansszorosát) az összefüggőség miatt elérhető, hogy legyen csupa nulla sor a mátrixban. Ekkor viszont $\det A = 0$ lenne, ami ellentmondás.

Harmadik lépés ((ii)-ből következik (iii)):

Az $A\underline{x} = \underline{0}$ egyenletet visszaírjuk lineáris egyenletrendszerre (a szokásos jelölésekkel):

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = 0 \end{cases}$$

A feltételek szerint az $(a_{11}, a_{12}, \dots, a_{1n}), (a_{21}, a_{22}, \dots, a_{2n}), \dots, (a_{n1}, a_{n2}, \dots, a_{nn})$ vektorok függetlenek. Ezek egy n -dimenziós vektortér vektorai, hiszen n db koordináta van.

Ez tehát egy bázisa az adott n -dimenziós vektortérnek, jelölje az egyes vektorokat $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$.

Az $x_1, x_2, \dots, x_n$ ismertleneket tartalmazó oszlopvektor legyen $\underline{x}$.

Tehát a feltételek szerint $\underline{v}_1 \cdot \underline{x} = \underline{v}_2 \cdot \underline{x} = \underline{v}_3 \cdot \underline{x} = \dots = \underline{v}_n \cdot \underline{x} = 0$.

Állítsuk elő az $(1, 0, 0, \dots, 0)$ vektort a $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ vektorok lineáris kombinációjaként. Mivel $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ generátorrendszer, ezért ez biztosan megtehető, tehát vannak olyan $\lambda_1, \lambda_2, \dots, \lambda_n$ együtthatók, melyekre $\lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 + \dots + \lambda_n \underline{v}_n = (1, 0, 0, \dots, 0)$.

Sőt, ezek a $\lambda_1, \lambda_2, \dots, \lambda_n$ számok egyértelműek. Ugyanis ha lenne még egy ilyen szám n -esünk, mondjuk a $\lambda_1', \lambda_2', \dots, \lambda_n'$ számokból, akkor

$$\lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 + \dots + \lambda_n \underline{v}_n = \lambda_1' \underline{v}_1 + \lambda_2' \underline{v}_2 + \dots + \lambda_n' \underline{v}_n, \text{ ami átrendezve}$$

$(\lambda_1 - \lambda_1') \underline{v}_1 + (\lambda_2 - \lambda_2') \underline{v}_2 + \dots + (\lambda_n - \lambda_n') \underline{v}_n = \underline{0}$ lenne, viszont a $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ függetlensége miatt itt minden együtthatónak 0-nak kell lennie. Ezért $\lambda_1 = \lambda_1', \lambda_2 = \lambda_2', \dots, \lambda_n = \lambda_n'$, ahogy azt állítottuk.

Tehát pontosan egyféleképpen találunk együtthatókat, melyekre

$$\lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 + \dots + \lambda_n \underline{v}_n = (1, 0, 0, \dots, 0) \text{ teljesül.}$$

Ha mindkét oldalt skalárisan szorozzuk az ismeretleneket tartalmazó oszlopvektorral, akkor az alábbiakat kapjuk:

$$(\lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 + \dots + \lambda_n \underline{v}_n) \cdot \underline{x} = (1, 0, 0, \dots, 0) \cdot \underline{x}$$

A jobb oldal egyenlő x_1 -gyel, amire a fenti egyértelmű előállítás miatt egyértelmű (azaz pontosan egy) megoldást kaptunk. Természetesen ugyanígy belátható, hogy a többi x_j -re is pontosan egy megoldást kapunk. ■

Fontos megjegyeznünk, hogy ebből a tételből már következik, hogy **négyzetes mátrix inverze egyértelmű (vagy nincs, vagy pontosan egy darab van).**

Sajátvektor, sajátérték

DEFINÍCIÓ (SAJÁTVEKTOR, SAJÁTÉRTÉK): Az $(n \times n)$ -es A mátrix sajátvektora az $(n \times 1)$ -es $\underline{v} (\neq \underline{0})$ vektor, ha valamilyen λ (komplex vagy valós) számra $A \cdot \underline{v} = \lambda \cdot \underline{v}$. Az adott $\underline{v} (\neq \underline{0})$ sajátvektorhoz tartozó λ a sajátértéke $\underline{v}$ -nek.

Könnyen látható, hogy **adott sajátértékhez tartozó sajátvektorok egy vektorteret (ún. sajátalteret) alkotnak**. Ehhez elég leellenőrizni, hogy a skalárral való szorzás és az összeadás nem visz ki a halmazból.

Legyen $\underline{v}$ és $\underline{v}'$ az A sajátvektora a λ sajátértékkel, azaz $A \cdot \underline{v} = \lambda \cdot \underline{v}$, valamint $A \cdot \underline{v}' = \lambda \cdot \underline{v}'$.

Ekkor $A(\alpha \underline{v}) = \alpha \cdot A \cdot \underline{v} = \alpha \cdot \lambda \cdot \underline{v} = \lambda \cdot (\alpha \underline{v})$, valamint

$A(\underline{v} + \underline{v}') = A\underline{v} + A\underline{v}' = \lambda \underline{v} + \lambda \underline{v}' = \lambda(\underline{v} + \underline{v}')$, ahogyan állítottuk.

Tekintsünk most példaként egy (2×2) -es mátrixot, és próbáljuk meghatározni a hozzá tartozó sajátvektorokat, illetve sajátértékeket.

Példa:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

A definíció alapján a sajátvektorra és a sajátértékre nézve az alábbi összefüggést kapjuk:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

Ez ekvivalens az alábbi egyenletrendszerrel:

$$\begin{cases} x + 2y = \lambda x \\ 2x + y = \lambda y \end{cases}$$

Ez így még kellemetlen, mert a λ értéke még nincs meg. Az a kérdés, hogy miként lehetne λ lehetséges értékeit meghatározni.

Általánosan annyit tudunk, hogy

$$A \cdot \underline{v} = \lambda \cdot \underline{v}, \text{ ami átrendezve}$$

$$A \cdot \underline{v} - \lambda \cdot \underline{v} = \underline{0}$$

Szorozattá alakítva a bal oldalon kapjuk, hogy

$$(A - \lambda \cdot E_n) \cdot \underline{v} = \underline{0}.$$

$$\text{Legyen } (A - \lambda \cdot E_n) = B.$$

A $B\underline{v} = \underline{0}$ egyenletnek egy megoldása mindig van, méghozzá a triviális $\underline{v} = \underline{0}$. Ha találunk sajátvektort, akkor – lévén, hogy a $\underline{0}$ hivatalból nem számít sajátvektornak –, a $B\underline{v} = \underline{0}$ egyenletnek több megoldása is van.

Ezért tehát ha létezik sajátvektor (ami definíció szerint nem lehet nullvektor), akkor szükségképpen $\det B = 0$, így $\det(A - \lambda \cdot E_n) = 0$.

Ez az összefüggés már segít hatékonyan kiszűrni a lehetséges sajátértékeket.

Ezen észrevétel után térjünk vissza most a példánkhoz:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

$$(A - \lambda \cdot E_2) = \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{pmatrix}$$

$$\det \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{pmatrix} = 0$$

$$(1 - \lambda)^2 - 2^2 = 0$$

$$(1 - \lambda)^2 = 4$$

$$1 - \lambda = 2, \text{ ekkor } \lambda = -1$$

$$1 - \lambda = -2, \text{ ekkor } \lambda = 3$$

$$\text{I. } \lambda = -1$$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}$$

$$\begin{cases} x + 2y = -x \\ 2x + y = -y \end{cases}$$

$$\begin{cases} 2x + 2y = 0 \\ 2x + 2y = 0 \end{cases}$$

Az itteni sajátvektorok $\begin{pmatrix} x \\ -x \end{pmatrix}$ alakúak.

$$\text{II. } \lambda = 3$$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$$

$$\begin{cases} x + 2y = 3x \\ 2x + y = 3y \end{cases}$$

$$\begin{cases} -2x + 2y = 0 \\ 2x - 2y = 0 \end{cases}$$

$$\begin{cases} 2x - 2y = 0 \\ 2x - 2y = 0 \end{cases}$$

Az itteni sajátvektorok $\begin{pmatrix} x \\ x \end{pmatrix}$ alakúak.

Karakterisztikus polinom

A fentiekben megismert $\det(A - \lambda E_n)$ kifejezés a λ -nak egy n -edfokú polinomja, aminek az algebra alaptétele miatt n darab komplex gyöke van. Ezt a polinomot az **A négyzetes mátrixhoz tartozó karakterisztikus polinom**nak nevezzük.

Összefoglalva, általánosan kiderült az alábbi összefüggés (ami megkönnyíti sajátvektorok keresését):

TÉTEL: *A négyzetes mátrix sajátértékei gyökei a mátrixhoz tartozó karakterisztikus polinomnak.*

A lineáris algebra néhány geometriai alkalmazása

Tétel: Két, síkban megadott $Q_1(x_1, y_1)$ és $Q_2(x_2, y_2)$ ponton átmenő egyenes egyenlete:

$$\det \begin{pmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix} = 0$$

Bizonyítás: Bármely síkbeli egyenes egyenlete $ax+by+c=0$ alakra hozható, ezért kézenfekvőnek tűnik az egyenes egyenletét

$$\det \begin{pmatrix} x & y & 1 \\ r & s & t \\ u & v & w \end{pmatrix} = 0 \text{ alakban keresnünk, ahol } r,s,t,u,v,w \text{ valamilyen konstansok, és } x, \text{ illetve } y$$

a változók. Ez azért ideális választás, mert ha az itt szereplő determinánst kiszámoljuk, éppen egy $ax+by+c$ alakú kifejezést kapunk.

A kérdés az, hogy miként érdemes megválasztani az r,s,t,u,v,w konstansok értékét.

Tudjuk, hogy az egyenes egyenletébe $x=x_1-t$, $y=y_1-t$ helyettesítve, nullát kapunk. Ugyanígy az egyenletbe $x=x_2-t$, $y=y_2-t$ helyettesítve, szintén nullát kapunk.

Tudjuk azt, hogy ha egy négyzetes mátrixnak van két azonos sora, akkor determinánsának értéke nulla. Ez alapján

$$\det \begin{pmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix} = 0 \text{ tényleg a szóban forgó egyenes egyenlete lesz.}$$

Megjegyzés: Hasonlóan látható be az előző eredményünk térbeli általánosítása:

Tétel: Térben három adott ponton, a $T_1(x_1, y_1, z_1)$, a $T_2(x_2, y_2, z_2)$, és a $T_3(x_3, y_3, z_3)$ pontokon átmenő sík egyenlete:

$$\det \begin{pmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{pmatrix} = 0$$

A fentiekhez hasonlóan a körre, illetve a gömbre vonatkozóan is készítünk determinánsos egyenleteket. A középiskolában tanult köregyenlet úgy néz ki, hogy

$(x - u)^2 + (y - v)^2 = r^2$, ahol (u, v) a kör középpontja, és r a sugara. Persze ha ezt megszorozzuk egy nullától különböző konstanssal, és elvégezzük a négyzetre emeléseket, akkor a rendezés után kapott egyenlet ugyanazt a kört írja le. Tehát egy kicsit általánosabban a köregyenletek olyan alakúak, hogy

$A(x^2 + y^2) + Bx + Cy + D = 0$ (fontos, hogy x^2 és y^2 együtthatója azonos marad a konstanssal való szorzást követően is).

Tétel: Három adott ponton, $P_1(x_1, y_1)$ -en, $P_2(x_2, y_2)$ -n, $P_3(x_3, y_3)$ -on áthaladó kör egyenlete

$$\det \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix} = 0$$

Bizonyítás:

Ugyanúgy járunk el, mint az egyenes egyenleténél. Észrevesszük, hogy az egyenletet célszerű

$$\det \begin{pmatrix} x^2 + y^2 & x & y & d \\ r & s & t & f \\ u & v & w & g \\ i & j & k & h \end{pmatrix} = 0 \text{ alakban keresni, ahol } x \text{ és } y \text{ a változók, a többi betű pedig valamilyen}$$

konstanst jelöl. Ez azért célszerű választás, mert az itteni determináns kifejtve éppen

$A(x^2 + y^2) + Bx + Cy + D$ alakú lesz. Tudjuk, hogy a kör mely három ponton megy át, tehát ezeket az egyenletbe helyettesítve nullát kell kapnunk. Ismét felhasználva a determinánsok azon tulajdonságát, hogy amennyiben van két azonos soruk, akkor értékük nullával egyenlő, kapjuk, hogy

$$\det \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix} = 0 \text{ valóban a szóban forgó kör egyenlete.}$$

Megjegyzés: Ha a pontok egy egyenesen vannak, akkor a determináns értéke automatikusan nulla (hiszen ekkor összefüggőek a sorok), tehát a semmitmondó, $0=0$ egyenletet kapjuk.

Jobbrendszerek, vektoriális szorzás

Azt fogjuk mondani, hogy az $\mathbf{x}$, $\mathbf{y}$ és $\mathbf{z}$ térbeli vektorok ebben a sorrendben **jobbrendszert** alkotnak, ha a (3×3) -as

$\det \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{pmatrix}$ determináns értéke nemnegatív.

Legyenek a térbeli standard bázis elemei $\mathbf{i}=(1,0,0)$, $\mathbf{j}=(0,1,0)$, $\mathbf{k}=(0,0,1)$.

Bevezetünk egy új műveletet, amit vektoriális szorzásnak fogunk nevezni.

Az $\mathbf{a}=(a_1, a_2, a_3)$ és $\mathbf{b}=(b_1, b_2, b_3)$ térbeli vektorok vektoriális szorzata legyen úgy definiálva, hogy

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}.$$

A determinánsokról belátott tulajdonságok alapján egyből leolvashatjuk, hogy a vektoriális szorzás ún. antikommutatív művelet, azaz az $\mathbf{a} \times \mathbf{b}$ vektoriális szorzat mindig az ellentettje a $\mathbf{b} \times \mathbf{a}$ vektoriális szorzatnak.

Szintén a determinánsok tulajdonságai alapján világos, hogy a vektoriális szorzás bilineáris tulajdonságú, azaz tetszőleges r, s valós számokra (skalárookra), illetve térbeli $\mathbf{a}, \mathbf{b}, \mathbf{c}$ vektorokra teljesül, hogy

$$\mathbf{a} \times (r\mathbf{b} + s\mathbf{c}) = r \cdot (\mathbf{a} \times \mathbf{b}) + s \cdot (\mathbf{a} \times \mathbf{c}).$$

A vektoriális szorzás legfontosabb tulajdonságait az alábbi tételben foglaljuk össze:

Tétel: Legyenek $\mathbf{a}=(a_1, a_2, a_3)$ és $\mathbf{b}=(b_1, b_2, b_3)$ tetszőleges térbeli vektorok, melyeknek egymással bezárt szöge $\Psi < \pi$. Ekkor az $\mathbf{a} \times \mathbf{b}$ vektoriális szorzat szintén egy térbeli vektor, ami merőleges az $\mathbf{a}$ és $\mathbf{b}$ vektorokra, hossza egyenlő $|\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \Psi$ -val, illetve az $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ vektorok ebben a sorrendben jobbrendszert alkotnak.

Bizonyítás:

A kifejtési tétel felhasználásával kapjuk, hogy

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = (a_2 b_3 - a_3 b_2) \mathbf{i} - (a_1 b_3 - a_3 b_1) \mathbf{j} + (a_1 b_2 - a_2 b_1) \mathbf{k} =$$

$= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$, tehát az eredmény valóban egy térbeli vektor.

Ellenőrizzük le, hogy ez a vektor merőleges $\mathbf{a}$ -ra és $\mathbf{b}$ -re is. Ehhez elég annyit bizonyítanunk, hogy ennek a vektornak az $\mathbf{a}$ -val, illetve $\mathbf{b}$ -vel vett skaláris szorzata nulla (itt felhasználjuk burkoltan a skaláris szorzatnak a második, azaz a koszinuszos értelmezését is).

$$\begin{aligned}
& (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) \cdot (a_1, a_2, a_3) = \\
& = a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_2a_1b_3 + a_3a_1b_2 - a_3a_2b_1 = \\
& = (a_1a_2b_3 - a_1a_2b_3) + (a_3a_1b_2 - a_1a_3b_2) + (a_2a_3b_1 - a_3a_2b_1) = 0.
\end{aligned}$$

Ebből tehát adódik, hogy $\mathbf{a} \times \mathbf{b}$ merőleges $\mathbf{a}$ -ra.

Természetesen ugyanígy (betűcserével) kijön, hogy $\mathbf{a} \times \mathbf{b}$ merőleges $\mathbf{b}$ -re is.

Most belátjuk, hogy $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \Psi$. Ez azzal ekvivalens, hogy

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 \cdot \sin^2 \Psi.$$

A bal oldalon álló kifejezés koordinátákkal kifejezve úgy néz ki, hogy

$$|\mathbf{a} \times \mathbf{b}|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2.$$

A jobb oldalon álló kifejezés pedig

$$\begin{aligned}
|\mathbf{a}|^2 \cdot |\mathbf{b}|^2 \cdot \sin^2 \Psi &= |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 \cdot (1 - \cos^2 \Psi) = |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 - |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 \cos^2 \Psi = |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 - |\mathbf{a}\mathbf{b}|^2 = \\
&= (a_1^2 + a_2^2 + a_3^2) \cdot (b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 = \\
&= (a_1^2b_1^2 + a_1^2b_2^2 + a_1^2b_3^2 + a_2^2b_1^2 + a_2^2b_2^2 + a_2^2b_3^2 + a_3^2b_1^2 + a_3^2b_2^2 + a_3^2b_3^2) - \\
&- (a_1^2b_1^2 + a_2^2b_2^2 + a_3^2b_3^2 + 2a_1b_1a_2b_2 + 2a_1b_1a_3b_3 + 2a_2b_2a_3b_3) = \\
&= (a_1^2b_2^2 - 2a_1b_1a_2b_2 + a_2^2b_1^2) + (a_1^2b_3^2 - 2a_1b_1a_3b_3 + a_3^2b_1^2) + \\
&+ (a_2^2b_3^2 - 2a_2b_2a_3b_3 + a_3^2b_2^2) = \\
&= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2.
\end{aligned}$$

Így tehát beláttuk, hogy valóban $|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 \cdot \sin^2 \Psi$.

Még annyit kell bizonyítanunk, hogy az $\mathbf{a}$, $\mathbf{b}$, $\mathbf{a} \times \mathbf{b}$ vektorok ebben a sorrendben jobbbrendszert alkotnak.

Ehhez meg kell határoznunk a

$$\det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ a_2b_3 - a_3b_2 & a_3b_1 - a_1b_3 & a_1b_2 - a_2b_1 \end{pmatrix} \text{ előjelét.}$$

Mivel ez kifejtve éppen $(a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2$, így valós koordináták mellett biztos, hogy a determináns értéke nemnegatív, és ezt kellett belátnunk. ■

A bizonyítás során, amikor a hosszúságot kiszámoltuk, ugyanazt az azonosságot használtuk három dimenzióban, mint amit a Cauchy-Schwarz-Bunyakovszkij-egyenlőtlenség bizonyításánál használtunk általánosabban (az ún. Lagrange-azonosságot).

Két térbeli vektor, $\mathbf{a}$ és $\mathbf{b}$ egyértelműen meghatároz egy paralellogrammát, ha őket egy közös kezdőpontba toljuk el. Vegyük észre, hogy az így keletkezett paralellogramma területe éppen $|\mathbf{a} \times \mathbf{b}|$. Ezen észrevétel alapján vezetjük majd be a felületi integrál fogalmát.

A jobbkéz-szabály, avagy a jobbrendszerek geometriai átfogalmazása

Azt, hogy három térvektor adott sorrendben jobbrendszert alkot, megpróbáljuk szemléletesen is leírni. A fizikában is gyakran használt jobbkézszabály a következőt takarja: az $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ térvektorok ebben a sorrendben eleget tesznek a **jobbkéz-szabálynak**, ha felvehetők oly módon, hogy az $\mathbf{x}$ vektor legyen egyirányú a jobb kezünk hüvelykujjával, az $\mathbf{y}$ vektor a jobb kezünk mutatóujjával, míg a $\mathbf{z}$ vektor a jobb kezünk középső ujjával.

Ezt matematikailag úgy fogjuk leírni, hogy az $\mathbf{x}$ vektort az $\mathbf{y}$ vektorba pozitív irányú, pínél kisebb szögű nyújtva forgatás viszi (ez biztosítja, hogy ezek megfelelnek a jobb kezünk hüvelyk-, illetve mutatóujjának), valamint az $\mathbf{x}$ -re és $\mathbf{y}$ -ra egyaránt merőleges $\mathbf{n}(\mathbf{x}, \mathbf{y})$ egységvektor, ami ugyanabba a féltérbe mutat, mint ilyenkor a középső ujjunk, a $\mathbf{z}$ -vel legfeljebb derékszöget zár be.

Lényegében azt fogjuk belátni, hogy **az adott sorrendben jobbrendszert alkotó térbeli vektorhármasok egyúttal eleget tesznek a jobbkéz-szabálynak, és ez visszafelé is érvényes** (tehát a szemléletes definíció, valamint a determinánsos definíció egymással ekvivalens).

Az alábbi tételt bizonyítjuk be:

Tétel: *A független $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ térbeli vektorok ebben a sorrendben pontosan akkor alkotnak jobbrendszert, ha ebben a sorrendben eleget tesznek a jobbkéz-szabálynak.*

Bizonyítás:

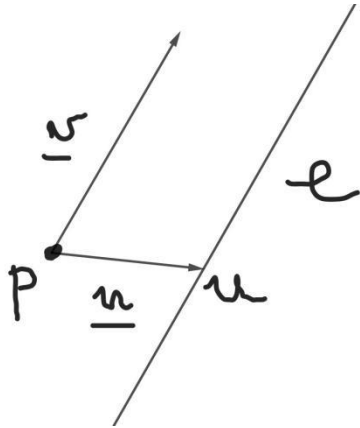
Tekintsünk három olyan $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ vektort, amik ebben a sorrendben eleget tesznek a jobbkéz-szabálynak.

A lényeges észrevételünk az lesz, hogy ha hozzáadjuk a három vektorunk közül valamelyikhez is egy másiknak valamilyen skalárszorosát (azaz valamilyen valós számszorosát), a jobbkéz-szabály ugyanezen sorrend megtartása mellett érvényben marad.

Bizonyítsuk be a fenti észrevételünket. Először tekintsük azt az esetet, amikor az $\mathbf{u}$ vektorhoz adjuk a $\mathbf{v}$ vektor egy skalárszorosát. Legyen az $\mathbf{u}$ és $\mathbf{v}$ vektorok közös kezdőpontja P , az $\mathbf{u}$ vektor végpontja U .

Ekkor (ha P -t vesszük origónak) az $\mathbf{u} + \lambda \mathbf{v}$ vektorok mindegyike a P -ből olyan pontba mutat, ami rajta van azon az egyenesen, ami $\mathbf{v}$ vektorral párhuzamos, és átmegy az U ponton. Ezért ebben az esetben

$\mathbf{n}(\mathbf{u}, \mathbf{v}) = \mathbf{n}(\mathbf{u} + \lambda \mathbf{v}, \mathbf{v})$, így természetesen $\mathbf{n}(\mathbf{u} + \lambda \mathbf{v}, \mathbf{v})$ is legfeljebb derékszöget zár be $\mathbf{w}$ -vel.



Persze ugyanúgy kezelhető az az eset, amikor $\mathbf{v}$ -hez adjuk $\mathbf{u}$ -nak egy skalárszorosát.

Ha $\mathbf{w}$ -hez adjuk hozzá $\mathbf{v}$ egy skalárszorosát, mondjuk a $\lambda \mathbf{v}$ vektort, akkor azt kell ellenőriznünk, hogy a $\mathbf{w} + \lambda \mathbf{v}$ és az $\mathbf{n}(\mathbf{u}, \mathbf{v})$ vektorok szöge legfeljebb derékszög. Ehhez elég leellenőriznünk a skaláris szorzatuk előjelét:

$(\mathbf{w} + \lambda \mathbf{v}) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) = \mathbf{w} \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) + \lambda \mathbf{v} \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) = \mathbf{w} \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) + 0$, azaz ha a $\mathbf{w} \cdot \mathbf{n}(\mathbf{u}, \mathbf{v})$ skaláris szorzat eredetileg nem volt negatív, a $(\mathbf{w} + \lambda \mathbf{v}) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v})$ sem lehet az.

Ugyanígy ellenőrizhető az az eset, amikor $\mathbf{w}$ -hez az $\mathbf{u}$ egy skalárszorosát adjuk hozzá.

Ha $\mathbf{u}$ -hoz adjuk a $\mathbf{w}$ egy skalárszorosát, mondjuk a $\lambda \mathbf{w}$ vektort, akkor azt kell belátnunk, hogy

$\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v})$ és $\mathbf{w}$ egymással legfeljebb derékszöget zár be.

Bontsuk fel az $\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v})$ vektort két komponensre: legyen

$$\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v}) = c(\lambda) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) + d(\lambda) \cdot \mathbf{n}(\mathbf{w}, \mathbf{v}).$$

Ez mindig egyértelműen megtehető, hiszen $\mathbf{n}(\mathbf{u}, \mathbf{v})$ és $\mathbf{n}(\mathbf{w}, \mathbf{v})$ függetlenek.

Ha koordinátákkal írjuk fel a vektorainkat, könnyen látható, hogy $c(\lambda)$ és $d(\lambda)$ olyan függvények, amik a λ változóban folytonosak. Ráadásul tudjuk, hogy $c(0)=1$.

Megmutatjuk, hogy $c(\lambda)$ nem vehet fel semmilyen negatív értéket. Ha ugyanis valamilyen λ -ra $c(\lambda)$ negatív lenne, akkor a Bolzano-tétel miatt létezne olyan $r \neq 0$ szám, melyre $c(r)=0$ lenne. Ezen r mellett

$\mathbf{n}(\mathbf{u} + r\mathbf{w}, \mathbf{v}) = d(r) \cdot \mathbf{n}(\mathbf{w}, \mathbf{v})$ teljesülne, azaz $\mathbf{n}(\mathbf{u} + r\mathbf{w}, \mathbf{v})$ merőleges lenne a $\mathbf{v}$, $\mathbf{w}$, $\mathbf{u} + r\mathbf{w}$ vektorok mindegyikére, ami lehetetlen, hiszen $\mathbf{v}$, $\mathbf{w}$, $\mathbf{u} + r\mathbf{w}$ vektorok függetlenek (nem egy síkban fekszenek, ezért nem létezik olyan egyenes, ami mindhármukra merőleges lenne).

Tehát $c(\lambda)$ értéke biztosan mindig pozitív. Vizsgáljuk ennek tudatában az $\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v}) \cdot \mathbf{w}$ és $\mathbf{n}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w}$ skaláris szorzatok előjelét:

$$\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v}) \cdot \mathbf{w} = c(\lambda) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w} + d(\lambda) \cdot \mathbf{n}(\mathbf{w}, \mathbf{v}) \cdot \mathbf{w} = c(\lambda) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w} + d(\lambda) \cdot 0 = c(\lambda) \cdot \mathbf{n}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w}.$$

Ebből az olvasható le, hogy az $\mathbf{n}(\mathbf{u} + \lambda \mathbf{w}, \mathbf{v}) \cdot \mathbf{w}$ és $\mathbf{n}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w}$ skaláris szorzatok mindig azonos előjelűek lesznek, és ezt kellett belátni.

Természetesen betűcserével ugyanígy kezelhető az az eset, amikor $\mathbf{v}$ -hez adjuk $\mathbf{w}$ egy skalárszorosát.

Most bizonyítsuk be a tételbeli ekvivalenciát. Legyen $\mathbf{u}$, $\mathbf{v}$ és $\mathbf{w}$ három olyan független térbeli vektor, amik eleget tesznek a jobbkéz-szabálynak. Készítsük el a (3×3) -as

$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix}$ mátrixot. Ezen mátrix determinánsának előjelét szeretnénk megtudni. A Gauss-elimináció segítségével, determinánstartó átalakításokkal elérhető, hogy a mátrix diagonális legyen, azaz

$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ alakú. Ekkor a fenti észrevételünk miatt az $(a,0,0)$, $(0,b,0)$, $(0,0,c)$ vektorok is biztosan

eleget tesznek a jobbkéz-szabálynak. Ugyanakkor könnyen meggyőződhetünk róla, hogy az $(a,0,0)$, $(0,b,0)$, $(0,0,c)$ vektorok pontosan akkor alkotnak jobbbrendszert, ha eleget tesznek a jobbkéz-szabálynak (itt csak az a , b , c lehetséges előjeleit kell vizsgálni).

Ezzel a tételt beláttuk. ■

Felületi integrál

A felületi integrál értelmezéséhez térben elképzelünk egy összefüggő, zárt térrészt, és ennek vesszük a határoló felületét. Az integrálást a felület szerint (a „krumpli héján” akarjuk elvégezni. A függvény, amit a felület szerint integrálunk, egy vektormező, ami értelmezve van a felület pontjain.

Először szemléletesen elmondjuk, hogyan készül a közelítő összeg.

A felületet feldaraboljuk kb. paralelogrammákra (pontosabban egy paralelogrammaráccsal közelítünk).

Minden lapból áll kifelé egy merőleges tűske, aminek a hossza az adott paralelogramma területe.

A függvénybe helyettesítjük a felület egy pontját, majd a helyettesítési értéket (ami egy vektor) skalárisan megszorozzuk a ponthoz tartozó felületi darabból kiálló tűskével mint vektorral. Ezen skaláris szorzatok összege lesz a közelítő összeg.

Pontosabban, ha $\underline{u}$ az integrálandó vektormező, és $\underline{r}$ a felület egy paraméterezése az u és v változókkal, akkor a felületi integrál úgy számolandó ki, hogy:

$$\iint_T \underline{u}(\underline{r}) \cdot (\underline{r}'_u \times \underline{r}'_v) du dv$$

Alkalmazás (a Maxwell-egyenletek integrálos alakja):

1. $\oint \underline{B} d\underline{F} = 0$, ahol $\underline{B}$ a mágneses indukció, $\underline{F}$ pedig egy felület,
2. $\epsilon_0 \oint \underline{E} d\underline{F} = \iiint \rho dv$, ahol $\underline{E}$ elektromos térerősség, ρ a töltéssűrűség, ϵ_0 konstans, v egy térrész, $\underline{F}$ ennek a felülete
3. $\frac{1}{\mu} \oint \underline{B} d\underline{x} = \iint \left(\underline{j} + \epsilon_0 \cdot \frac{\partial \underline{E}}{\partial t} \right) d\underline{F}$

$$4. \oint E \, d\mathbf{x} = \iint \frac{\partial k}{\partial t} dF$$

Differenciálegyenletek

A differenciálegyenletek olyan egyenletek, ahol az ismeretlen valamilyen függvény, és ennek valamilyen deriváltjai is szerepelnek az egyenletben (szerepelhet akár többszörös derivált, vagy akár parciális derivált is).

Példa: $y'(x) = 2x^2 + 5$

$$\operatorname{tg} x \cdot y(x) - \arcsin x \cdot y'(x) + x^2 = e^{e^x}$$

$$\sin x \cdot y(x) + \operatorname{cth} x \cdot y'(x) - \operatorname{th} x \cdot y''(x) + x^2 \cdot y'''(x) = \operatorname{ctg}(x^2)$$

Megjegyzés: A differenciálegyenleteket általában nehéz megoldani, de bizonyos típusúakat viszonylag egyszerűen meg lehet.

I. Szétválasztható változójú: Ezek $y'(x) = \frac{f(x)}{g(y(x))}$ alakúak (vagy erre az alakra hozhatók).

Megoldásuk: $g(y(x)) \cdot y'(x) = f(x)$.

Ha $(G(x))' = g(x)$, akkor a bal oldal úgy írható, mint $(G(y(x)))'$. Ha megkeressük $f(x)$ egy primitív függvényét (amennyiben van), az $F(x)$ -et, akkor az egyenlet $(G(y(x)))' = (F(x))'$. Ebből az következik, hogy a $G(y(x)) - F(x) = C$ konstans.

Példa:

$$1. y'(x) \cdot \operatorname{ctg} x + y(x) = 2$$

$$y'(x) \cdot \operatorname{ctg} x = 2 - y(x)$$

$$y'(x) = (2 - y(x)) \cdot \operatorname{tg} x$$

$$\frac{y'(x)}{2 - y(x)} = \operatorname{tg} x$$

$$(-\ln|2 - y(x)|)' = (-\ln|\cos x|)'$$

$$-\ln|2 - y(x)| + \ln|\cos x| = C$$

$$e^{\ln|\frac{\cos x}{2 - y(x)}|} = e^{C^*} \rightarrow C_0 > 0 \text{ konstans}$$

$$|\cos x| = C_0 \cdot |2 - y(x)|$$

Az $y(x)$ függvény folytonossága miatt csak két eset van:

$$a) \cos x = C_0 \cdot (2 - y(x))$$

$$y(x) = \frac{\cos x - 2C_0}{-C_0}$$

$$b) -\cos x = C_0 \cdot (2 - y(x))$$

$$y(x) = \frac{-\cos x - 2C_0}{-C_0}$$

$$f(x) = \operatorname{tg} x$$

$$g(x) = \frac{1}{2 - x}$$

$$g(y(x)) = \frac{1}{2 - y(x)} \rightarrow \frac{1}{g(y(x))} = 2 - y(x)$$

Tehát ez tényleg I. típusú.

$$\int \frac{1}{2 - x} dx = -\ln|2 - x| + C$$

$$\text{A bal oldal: } (-\ln|2 - y(x)|)'$$

$$\int \operatorname{tg} x dx = -\ln|\cos x| + C$$

$$2. y'(x) = 3 \cdot \sqrt[3]{y^2(x)} \quad f(x) = 3$$

$$\frac{y'(x)}{\sqrt[3]{y^2(x)}} = 3$$

$$(3 \cdot \sqrt[3]{x})' = (3x)'$$

$$3 \cdot \sqrt[3]{x} - 3x = C_0$$

$$y(x) = \left(\frac{3x + C_0}{3} \right)^3 = \left(x + \frac{1}{3}C_0 \right)^3$$

$$g(x) = \frac{1}{\sqrt[3]{x^2}} = x^{-2/3}$$

$$\int x^{-2/3} dx = \frac{x^{1/3}}{1/3} + c = 3 \cdot \sqrt[3]{x} + c$$

$$\int 3 dx = 3x + c$$

II. $y'(x) = f\left(\frac{Ax+By(x)+C}{ax+by(x)+c}\right)$ alakú egyenletek: Az ilyen típusú egyenleteket megpróbáljuk

viSSZavezetni az I. típusra. Tekintsük az $\begin{cases} Ax + By(x) + C = 0 \\ ax + by(x) + c = 0 \end{cases}$ egyenletrendszert. Ha ennek pontosan egy megoldása van, mondjuk $x = x_0, y = y_0$, akkor vezessük be az új $y(x) - y_0 = v(x)$ függvényt, illetve az új $x - x_0 = u$ ismeretlent.

$$y'(x) = f\left(\frac{Ax + By(x) + C}{ax + by(x) + c}\right) = f\left(\frac{A(x - x_0) + B(y(x) - y_0)}{a(x - x_0) + b(y(x) - y_0)}\right)$$

Példa:

$$1. y'(x) = \frac{2x-5y(x)+3}{2x+4y(x)-6}$$

$$f(x) = x \quad A = 2, B = -5, C = 3$$

$$a = 2, b = 4, c = -6$$

$$y'(x) = \frac{2(x-1)-5(y(x)-1)}{2(x-1)+4(y(x)-1)} = \frac{2-5\frac{y(x)-1}{x-1}}{2+4\frac{y(x)-1}{x-1}}$$

$$\begin{cases} 2x - 5y + 3 = 0 \\ 2x + 4y - 6 = 0 \end{cases}$$

$$\frac{y(x)-1}{x-1} = w(x)$$

$$\text{I-II.} \quad -9y + 9 = 0$$

$$[(x-1) \cdot w(x)]' = \frac{2-5w(x)}{2+4w(x)}$$

$$y_0 = 1, x_0 = 1$$

$$w(x) + (x-1) \cdot w'(x) = \frac{2-5w(x)}{2+4w(x)}$$

Ez már I. típusú.

2. Ha nem egyértelmű a megoldása az egyenletrendszernek:

$$y'(x) = \left(\frac{x + 2y(x) + 1}{x + 2y(x) + 2} \right)^3 = \left(1 - \frac{1}{x + 2y(x) + 2} \right)^3 = \left(1 - \frac{1}{z(x)} \right)^3$$

$$\frac{z'(x)}{2} - 1 = \left(1 - \frac{1}{z(x)} \right)^3$$

III. Elsőrendű differenciálegyenlet: $y'(x) + y(x) \cdot p(x) = q(x)$

Gyakran működő ötlet az, hogy megszorozzuk mindkét oldalt egy alkalmas $s(x)$ függvénnyel, amit majd ügyesen választunk meg. Ekkor tehát

$$s(x) \cdot y'(x) + s(x) \cdot p(x) \cdot y(x) = s(x) \cdot q(x)$$

Ha úgy választjuk $s(x)$ -et, hogy $s'(x) = s(x) \cdot p(x)$ legyen, akkor a bal oldal átírható a szorzásszabály segítségével (hiszen a bal oldalon $s(x) \cdot y(x)$ deriváltja fog szerepelni).

Az $s'(x) = s(x) \cdot p(x)$ feltétel éppen azt jelenti, hogy $\frac{s'(x)}{s(x)} = p(x)$, vagyis $(\ln |s(x)|)' = p(x)$.

Példa:

$$y'(x) + y(x) = 1$$

$$s(x) \cdot y'(x) + s(x) \cdot y(x) = s(x) \rightarrow (s(x))' = s(x) \rightarrow s(x) = e^x$$

$$e^x \cdot y'(x) + e^x \cdot y(x) = e^x$$

$$(e^x \cdot y(x))' = (e^x)'$$

$$e^x \cdot y(x) - e^x = C$$

$$e^x(y(x) - 1) = C$$

$$y(x) = \frac{C}{e^x} + 1$$

Homogén lineáris differenciálegyenlet-rendszerek

Az alábbi alakú egyenletrendszerekről van szó:

$$\begin{cases} a_{11}(x) \cdot f_1(x) + a_{12}(x) \cdot f_2(x) + \dots + a_{1k}(x) \cdot f_k(x) + g_1(x) = f_1'(x) \\ a_{21}(x) \cdot f_1(x) + a_{22}(x) \cdot f_2(x) + \dots + a_{2k}(x) \cdot f_k(x) + g_2(x) = f_2'(x) \\ \vdots \\ a_{n1}(x) \cdot f_1(x) + a_{n2}(x) \cdot f_2(x) + \dots + a_{nk}(x) \cdot f_k(x) + g_n(x) = f_n'(x) \end{cases}$$

Az $f_1(x), f_2(x), \dots, f_k(x)$ az ismeretlenek.

Példa:

$$\begin{cases} \sin x \cdot f_1(x) - 2^x \cdot f_2(x) + \operatorname{cth} x \cdot f_3(x) - \ln x = f_1'(x) \\ -\cos^2 x \cdot f_1(x) + \sin(3x) \cdot f_2(x) - 5 \cdot f_3(x) + \pi^2 = f_2'(x) \\ (x^4 - x^2) \cdot f_1(x) + (3^x - x^2) \cdot f_2(x) - \sin^2 x \cdot f_3(x) = f_3'(x) \end{cases}$$

TÉTEL: Egy homogén lineáris differenciálegyenlet-rendszer megoldásai vektorteret alkotnak. Amennyiben $n=k$ (tehát ugyanannyi egyenlet van, mint ismeretlen függvény), akkor ennek a vektortérnek a dimenziója n .

Csak olyan esetben fogjuk felírni a megoldásokat, amikor az együtthatók konstansok, pl.

$$\begin{cases} 2f(x) - g(x) + h(x) = f'(x) \\ f(x) + 2g(x) - h(x) = g'(x) \\ f(x) - g(x) + 2h(x) = h'(x) \end{cases}$$

TÉTEL: Legyen A az a mátrix, ami egy konstans együtthatós lineáris differenciálegyenlet-rendszerhez tartozik. Tegyük fel, hogy ez egy négyzetes mátrix. A minden sajátértéke valós és különböző. Legyenek a sajátértékek $\lambda_1, \lambda_2, \dots, \lambda_N$; illetve a megfelelő sajátvektorok $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_N$. Ebben az esetben $(e^{\lambda_i \cdot x} \cdot \underline{v}_i)$ függvények egyúgynevezett alrendszeret alkotnak, amiből generálhatóak a differenciálegyenlet-rendszer megoldásai.

Példa megoldása:

$$\begin{cases} 2f(x) - g(x) + h(x) = f'(x) \\ f(x) + 2g(x) - h(x) = g'(x) \\ f(x) - g(x) + 2h(x) = h'(x) \end{cases}$$

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

Karakterisztikus polinom:

$$\begin{aligned} c(x) &= \det \begin{pmatrix} 2-x & -1 & 1 \\ 1 & 2-x & -1 \\ 1 & -1 & 2-x \end{pmatrix} = \det \begin{pmatrix} 2-x & 0 & 1 \\ 1 & 1-x & -1 \\ 1 & 1-x & 2-x \end{pmatrix} = \\ &= -0 + (1-x) \cdot \det \begin{pmatrix} 2-x & 1 \\ 1 & 2-x \end{pmatrix} - (1-x) \cdot \det \begin{pmatrix} 2-x & 1 \\ 1 & -1 \end{pmatrix} = \\ &= (1-x) \cdot \left[\det \begin{pmatrix} 2-x & 1 \\ 1 & 2-x \end{pmatrix} - \det \begin{pmatrix} 2-x & 1 \\ 1 & -1 \end{pmatrix} \right] = \\ &= (1-x) \cdot \det \begin{pmatrix} 2-x & 0 \\ 1 & 3-x \end{pmatrix} = \\ &= (1-x) \cdot (2-x) \cdot (3-x) \end{aligned}$$

Sajátértékek: 1, 2, 3.

Meg kell vizsgálni, hogy ezekhez milyen sajátvektorok tartoznak.

I. $\lambda = 1$

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{cases} 2x - y + z = x \\ x + 2y - z = y \\ x - y + 2z = z \end{cases}$$

$$\begin{cases} x - y + z = 0 \\ x + y - z = 0 \\ x - y + z = 0 \end{cases}$$

$$\begin{cases} x - y + z = 0 \\ x + y - z = 0 \end{cases}$$

$$-2y + 2z = 0$$

$$z - y = 0$$

$$z = y$$

$$x = 0$$

A sajátvektorok $\begin{pmatrix} 0 \\ x \\ x \end{pmatrix}$ alakúak.

II. $\lambda = 2$

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}$$

$$\begin{cases} 2x - y + z = 2x \\ x + 2y - z = 2y \\ x - y + 2z = 2z \end{cases}$$

$$\begin{cases} -y + z = 0 \\ x - z = 0 \\ x - y = 0 \end{cases}$$

$$x = y = z$$

A sajátvektor $\begin{pmatrix} x \\ x \\ x \end{pmatrix}$ alakú.

III. $\lambda = 3$

$$\begin{cases} 2x - y + z = 3x \\ x + 2y - z = 3y \\ x - y + 2z = 3z \end{cases}$$

$$\begin{cases} -x - y + z = 0 \\ x - y - z = 0 \\ x - y - z = 0 \end{cases}$$

$$\begin{cases} -x - y + z = 0 \\ x - y - z = 0 \end{cases}$$

$$-x + 2z = 0$$

$$z = x$$

$$y = 0$$

A sajátvektor $\begin{pmatrix} x \\ 0 \\ x \end{pmatrix}$ alakú.

A tétel felhasználásával a differenciálegyenletrendszer megoldásai:

$$e^{1x} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ e^x \\ e^x \end{pmatrix}; e^{2x} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{2x} \\ e^{2x} \\ e^{2x} \end{pmatrix}; e^{3x} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{3x} \\ 0 \\ e^{3x} \end{pmatrix}$$

Az összes megoldás:

$$\begin{pmatrix} f(x) \\ g(x) \\ h(x) \end{pmatrix} = C_1 \cdot \begin{pmatrix} 0 \\ e^x \\ e^x \end{pmatrix} + C_2 \cdot \begin{pmatrix} e^{2x} \\ e^{2x} \\ e^{2x} \end{pmatrix} + C_3 \cdot \begin{pmatrix} e^{3x} \\ 0 \\ e^{3x} \end{pmatrix}$$

Magasabb rendű homogén lineáris differenciálegyenletek

Az alábbi alakban írhatók: $y^N(x) + a_1 x \cdot y^{N-1}(x) + \dots + a_N x \cdot y(x) = 0$, ahol $y(x)$ az ismeretlen függvény, aminek a magasabb rendű deriváltjai is szerepelnek az egyenletben.

Példa: $y^5(x) - \ln x \cdot y^4(x) + (x^5 + x^6) \cdot y'''(x) - \operatorname{tg} x \cdot y''(x) + (\operatorname{cth} x - e^x) \cdot y'(x) + 2^x \cdot y(x) = 0$

TÉTEL: Egy N -edrendű homogén lineáris differenciálegyenlet megoldásai egy N -dimenziós vektorteret alkotnak.

Azzal az esettel foglalkozunk, amikor az együtthatók konstans függvények.

Tehát az $y^N(x) + a_1 \cdot y^{N-1}(x) + \dots + a_N \cdot y(x) = 0$ alakú differenciálegyenleteket vizsgáljuk. Ennek megoldásai a $p(z) = z^N + a_1 \cdot z^{N-1} + \dots + a_{N-1} \cdot z + a_N$ polinom gyökeinek segítségével állíthatók elő.

Az algebra alaptétele miatt ennek a polinomnak multiplicitással számolva N db gyöke van.

Először tekintsük ennek a polinomnak azokat a gyökeit, amik nem valósak. Ha $p(z)$ valós együtthatós és $p(z_0) = 0$, akkor $p(\overline{z_0}) = \overline{p(z_0)} = 0$, azaz teljesül, hogy amennyiben z_0 gyök, akkor $\overline{z_0}$ is gyök.

A nem valós gyökök legyenek:

$$\begin{array}{ll} z_1 = C_1 + i \cdot d_1 & z_2 = C_1 - i \cdot d_1 + \\ z_3 = C_2 + i \cdot d_2 & z_4 = C_2 - i \cdot d_2 \\ \vdots & \vdots \\ z_{2K-1} = C_K + i \cdot d_K & z_{2K} = C_K - i \cdot d_K \end{array}$$

A valós gyökök legyenek $z_{2K+1}, z_{2K+2}, \dots, z_M$. Ezeknek a multiplicitása legyen rendre $m_{2K+1}, m_{2K+2}, \dots, m_M$.

TÉTEL: Az $y^N(x) + a_1 \cdot y^{N-1}(x) + \dots + a_N \cdot y(x) = 0$ differenciálegyenlet alapmegoldásai úgy írhatók, hogy $f_1(x) = e^{C_1 x} \cdot \cos(d_1 x)$, $f_2(x) = e^{C_1 x} \cdot \sin(d_1 x)$, $f_3(x) = x \cdot e^{C_1 x} \cdot \cos(d_1 x)$, $f_4(x) = x \cdot e^{C_1 x} \cdot \sin(d_1 x)$

$\vdots$

$$f_{2m_1-1}(x) = x^{m_1-1} \cdot e^{C_1 x} \cdot \cos(d_1 x), f_{2m_1}(x) = x^{m_1-1} \cdot e^{C_1 x} \cdot \sin(d_1 x)$$

$\vdots$

$$f_{2(m_1+m_2+\dots+m_k)-1}(x) = x^{m-1} \cdot e^{C_K x} \cdot \cos(d_K x), f_{2(m_1+m_2+\dots+m_k)}(x) = x^{m-1} \cdot e^{C_K x} \cdot \sin(d_K x)$$

$\vdots$

$$f_{2(m_1+m_2+\dots+m_k)+1}(x) = e^{m_{2K+1} \cdot x}$$

$\vdots$

$$f_M(x) = e^{m_M \cdot x}$$

Példa:

$$y^5(x) + 8 \cdot y^3(x) + 16 \cdot y'(x) = 0$$

$$p(z) = z^5 + 8 \cdot z^3 + 16 \cdot z = z(z^4 + 8z^2 + 16) = z \cdot (z^2 + 4)^2 = z(z + 2i)^2(z - 2i)^2$$

nem valós gyökök: $(-2i)$ és $(2i)$, kettő multiplicitással

valós gyökök: 0, 1 multiplicitással

$$f_1(x) = e^{0 \cdot x} \cdot \cos(2x), f_2(x) = e^{0 \cdot x} \cdot \sin(2x), f_3(x) = x \cdot e^{0 \cdot x} \cdot \cos(2x),$$

$$f_4(x) = x \cdot e^{0 \cdot x} \cdot \sin(2x), f_5(x) = e^{0 \cdot x} = 1$$

$$\text{Megoldások: } C_1 \cdot \cos(2x) + C_2 \cdot \sin(2x) + C_3 \cdot x \cdot \cos(2x) + C_4 \cdot x \cdot \sin(2x) + C_5$$

