Bevezetés az Algebrába - 1
Jegyzőkönyv, BME, 2025 ősz

Sági Gábor*

2025. december 27.

HUN-REN Rényi Alfréd Matematikai Kutatóintézet és
BME Matematika Intézet, Algebra és Geometria Tanszék
e-mail: sagi@renyi.hu

Minden visszajelzést szívesen veszek.
1.1. Verzió.

*Ez az anyag néhány hallgatóm (névsor szerinti sorrendben): Balog Benjámin, Domokos Norbert, Gegő Levente, Kárpáti Noémi, Sándor Ákos és Sándor Zétény hathatós segítségével jött létre. Precíz munkájukért, lelkesedésükért, nagyon hálás vagyok, hatalmas köszönet érte.

Előszó

Mi ez az anyag? Ez az iromány egy elektronikus segédlet, mely a BME-n, 2025 őszén, elsőéves matematikus hallgatók számára tartott „Bevezetés az algebrába - 1" kurzushoz készült. Első, nyers változatát néhány lelkes hallgatóm készítette; ezzel kapcsolatban az alábbi, „Hogy jött létre ez az anyag?" kezdetű bekezdésre utalok. Az iromány egyetlen célja, hogy - a lehetőségekhez képest - minél pontosabban lehessen rekonstruálni, hogy mi hangzott el az egyes előadásokon. Azt gondolom, ezt a célt nagymértékben (nyilván nem tökéletesen) sikerült is megvalósítani. A későbbi, javított változatokon folyamatosan dolgozunk, amint elkészül egy újabb változat, elérhetővé teszem azt is.

Ez az iromány nem könyv. Rengeteg kiváló, magyar nyelven elérhető bevezető algebra és bevezető számelmélet könyv kapható jelenleg is; ezekkel értelmetlen lenne konkurálni, már csak azért is, mert számos esetben egykor én is ezekből tanultam, ezekből inspirálódtam...

Ez az iromány nem egyetemi jegyzet abban az értelemben sem, hogy az áttekinthetőség, tömörség érdekében a szövegbe rövidítések, sőt akár képletek ékelődnek úgy, ahogy egy klasszikus, nyomtatott anyagban megengedhetetlen lenne.

Ahogy azt alább, a „Hogy jött létre ez az anyag?" kezdetű bekezdésben is ki fogom fejteni, ennek a szövegnek a nyers változatát több hallgató önkéntes munkával állította elő. Emiatt az egyes fejezetek kidolgozottsága eltérő, néhányan rövidítéseket használtak, képleteket írtak, míg mások kiírták a megfelelő szófordulatokat. Nem egységes a szöveg nyelvi szempontból sem (kisbetű-nagybetű használata egyes elnevezésekben, egybeírás-különírás egyes kifejezésekben sajnos nem következetes és emiatt nyilvánvaló, hogy a szöveg egyes részletei nyelvtanilag hibásak (hiszen, pl. egyes elnevezésekben a kisbetűk-nagybetűk eltérő használata esetén legalább az egyik változat hibás)). Mivel a vizsgaidőszak előtt el kellett készülnie ennek az iránynak, az ilyen jellegű szöveggondozást (nyelvtani hibák, következetlenségek javítását) későbbre halasztjuk, és mégegyszer hangsúlyozzuk, hogy egy valamirevaló egyetemi jegyzetben ez a „hanyagság" elfogadhatatlan lenne! Tisztában vagyok vele, hogy az erre vonatkozó kritikák teljesen jogosak.

Ez az iromány a **Jegyzőkönyv** címet kapta, ami arra utal, hogy egyetlen célunk annak rögzítése volt, hogy - a táblaképen túl - mi hangzott el előadáson; természetesen az apróbb hibákat (index-elírásokat, előjel-tévesztéseket, stb.) igyekeztem korrigálni. Kritikusan olvassátok, mert minden igyekezetem ellenére bizonyára maradtak ilyen jellegű pontatlanságok ebben a szövegben is. Az ezzel kapcsolatos felelősség engem terhel. Ha valaki hibát talál, megköszönöm, ha visszajelzi nekem. Az egyetlen

lényeges változtatás az, hogy az egyes előadások legelején szereplő emlékeztetőket elhagytuk, hiszen ebben a Jegyzőkönyvben az elhagyott emlékeztetők néhány oldallal korábban teljes terjedelmükben megtalálhatók.

Ebben az irományban tehát (az előadások legelején szereplő rövid emlékeztetők elhagyásán kívül) mindent pontosan abban a sorrendben, abban a formában igyekeztünk rekonstruálni, ahogy azt az előadáson feldolgoztuk. Ez sok esetben - kisebb-nagyobb mértékben - eltér attól a felépítéstől, ami a témakörök belső logikája szempontjából ideális lenne, hiszen - mint minden egyetemi kurzus esetében - tekintettel kellett lennem az óraelmaradásokra, a gyakorlatokra mindig biztosítani kellett annyi új ismeretet, ami nem sok, de nem is kevés a heti haladás szempontjából, koordinálnom kellett, hogy - a néhány, nem is egyforma mértékben elmaradó gyakorlat ellenére - a két gyakorlati csoport nagyjából ugyanott tartson, stb. Az előadásokon elhangzottakhoz képest ebbe az anyagba további ismereteket szándékosan nem illesztettem, de a szóban elhangzott, intuitív magyarázatokat, történeti megjegyzéseket, „meséket" jobbára beleírtam, mert azt gondolom, hogy egyrészt ezek a kitérők - az önmagukban is megőrzendő értéket képviselő - matematikai folklór részét képezik, másrészt ezek nemcsak hasznos, nemcsak szórakoztató kitérők, hanem az anyag lényegéhez tartoznak abból a szempontból is, hogy segítik a mélyebb megértést, orientálnak, hogy mit miért vizsglunk, az egyes részeket miért úgy vizsgáljuk, továbbá az informális megjegyzéseknek önmagukban is komoly magyarázóereje lehet.

Hogyan használd ezt az anyagot?

> „Most elmagyarázzuk, hogyan kell olvasnod ezt a könyvet. A helyes mód az, hogy nappal az íróasztalodra, este a párnád alá teszed, és teljesen az olvasásnak szenteled magad... egészen addig, amíg szívből, kívülről nem tudod az egészet."
>
> Saharon Shelah: Classification Theory.

A fenti idézettel ellentétben nem vagyok biztos benne, hogy a vizsgára készülés idejében ezt az anyagot a legelső oldalától a legutolsóig teljes terjedelmében végig kell olvasni. Ezt talán a vizsga után (pl. a következő regisztrációs héten) tedd meg, hogy az őszi kurzus hangulatát felidézd, az előadásokon közösen eltöltött idő szép emlékké nemesedjen és a sikeres vizsga, majd az azt követő megérdemelt pihenés után ráhangolódj a tavaszi kurzusra. Egyes részletek szépsége akkor lesz szembetűnő, ha az anyagot nagyjából ismered már.

A vizsgaidőszakban ezt az Előszót :-) és a tartalomjegyzéket, jelölésekről szóló nulladik fejezetet érdemes alaposabban átböngészni, hogy később tudd, mit hol találsz. A vizsgára készüléskor újabb források használata kifejezetten zavaró lehet az eltérő jelölések, eltérő felépítés, sőt az óhatatlanul eltérő matematikai tartalom miatt. Ez az oka annak is, hogy ebbe a Jegyzőkönyvbe nem írtam bele olyan új ismereteket,

melyek az előadásokon nem hangzottak el. A vizsgára készüléshez a legjobb forrás az előadásokon készített jegyzeted. Ha valamelyik előadásról hiányoztál, vagy jelen voltál ugyan, de a jegyzeteidből nem tudsz mindent rekonstruálni, akkor érdemes megnézni ennek az iromány nak a megfelelő részletét. Hangsúlyozom továbbá, hogy a vizsgán számonkért anyag valamivel kisebb annál, mint ami előadáson elhangzott (tehát a vizsgaanyag valamivel kisebb annál, mint ami ebben a Jegyzőkönyvben is szerepel). Ugyanakkor azonban, a szokásoknak megfelelően, a vizsgán számonkért anyag szerkezete kismértékben eltér attól, ahogy az anyag előadáson elhangzott. Ez az eltérés szándékos: azt a célt szolgálja, hogy a vizsgára készülés során Te magad azonosítsd, hogy az egyes vizsgatételek mikor és hogyan hangzottak el előadáson; ez a munka nagymértékben segíti a mélyebb megértést, ezért nem is célszerű megpróbálni elkerülni.

Ha ennek a Jegyzőkönyvnek a segítségével sem sikerül rekonstruálni egyes részleteket, kérlek keress, és konzultálunk!

Hogy jött létre ez az anyag? Ez az anyag úgy jött létre, hogy spontán, önszerveződő módon néhány hallgatóm szerkeszthető LaTex file-á alakította az előadásokon készített kézzel írt jegyzeteit. Ezt véletlenül tudtam meg. Elkötelezettségük lenyűgözött, tiszteletet ébresztett bennem, és lelkesedésük annyira meghatott, hogy felajánlottam: az általuk készített LaTex file-okat átnézem, szükség esetén korrigálom, kiegészítem, egybeszerkesztem, és minden hallgatóm számára elérhetővé teszem. Ennek a munkának az eredménye ez a Jegyzőkönyv.

Ezt az előszót azzal zárom, hogy köszönetet mondok azoknak a hallgatóimnak, akik ezt a rendkívül nagy munkát elvégezték. Tehát **HATALMAS KÖSZÖNET** (névsor szerinti sorrendben)

Balog Benjáminnak, Domokos Norbertnek, Gegő Leventének,
Kárpáti Noéminek, Sándor Ákosnak és Sándor Zéténynek.

Hathatós segítségükért, precíz munkájukért, lelkesedésükért nagyon hálás vagyok! Az általuk átadott nyers szöveget nagy örömmel, szeretettel csinosítottam tovább; azt kívánom, Ti is ilyen örömmel olvassátok.

Budapest, 2025 December.

Sági Gábor

Tartalomjegyzék

0. Alapfogalmak és jelölések

Gyorsírásos jelölések, rövidítések

- TFH: Tegyük fel, hogy;
- ACSA: Akkor és csak akkor;
- $\exists$: Létezik;
- $\forall$: Minden;
- $\exists!$: Pontosan egy van;
- ... $\Rightarrow$ ...: Ha... akkor... .

Halmazelméleti alapfogalmak

- $\in$: eleme;
- $\subset$ illetve $\subseteq$: valódi részhalmaz, illetve részhalmaz;
- $\cap$: metszet;
- $\cup$: unió;
- $A \setminus B$ vagy $A - B$: az A és B halmazok különbsége;
- $\mathbb{N}$: természetes számok halmaza ($\mathbb{N} = \{0, 1, 2, 3, \dots\}$);
- $\mathbb{N}^+$: pozitív egész számok halmaza ($\mathbb{N}^+ = \{1, 2, 3, \dots\}$);
- $\mathbb{Z}$: egész számok halmaza;
- $\mathbb{Q}$: racionális számok halmaza;
- $\mathbb{R}$: valós számok halmaza.

1. Relációk, ekvivalenciarelációk és rendezési relációk

Ebben a részben bevezetünk néhány olyan fogalmat, melyekre később léptennyomon szükségünk lesz (ráadásul nem csak algebrából).

1.1. Definíció (Descartes-szorzat)**.** *Az A és B halmazok* ***Descartes-szorzata****:*

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

1.2. Példa. $\{0, 1\} \times \{1, 2, 3\} = \{(0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3)\}.$

1.3. Definíció (Reláció)**.** *Legyen A egy halmaz. Az $A \times A$ Descartes-szorzat egy tetszőleges R részhalmazáról azt mondjuk, hogy* ***kétváltozós reláció*** *A-n.*

1.4. Definíció (Ekvivalenciareláció)**.** *Egy $R \subseteq A \times A$ reláció* ***ekvivalenciareláció****, ha teljesül, hogy:*

1. $\forall x \in A : (x, x) \in R$ *(reflexivitás);*
2. $\forall x, y \in A : (x, y) \in R \Rightarrow (y, x) \in R$ *(szimmetria);*
3. $\forall x, y, z \in A : (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$ *(tranzitivitás).*

1.5. Példa. *Az alábbi első 3 példa ekvivalenciareláció, az utolsó nem az.*

1. $R = \{(x, y) \mid x, y \in \mathbb{N}\}$ *ekvivalenciareláció $\mathbb{N}$-en;*
2. $R = \{(x, y) \mid x, y \in \mathbb{N}, x - y$ *páros*$\}$ *ekvivalenciareláció $\mathbb{N}$-en;*
3. $A = \{$*egyenesek*$\}$, $R = \{(e, f) \mid e, f \in A, e \parallel f\}$ *ekvivalenciareláció a sík egyeneseinek halmazán;*
4. $R = \{(a, b) \mid a, b \in \mathbb{N}, b = a + 1\}$ ***NEM*** *ekvivalenciareláció $\mathbb{N}$-en.*

1.6. Definíció (Ekvivalencia-osztály)**.** *Legyen R ekvivalenciareláció A-n és $a \in A$. Ekkor az a* ***ekvivalencia-osztálya****:*

$$a/R = \{b \in A \mid (a, b) \in R\}.$$

1.7. Tétel (Osztályfelbontás)**.** *Ha R ekvivalenciareláció A-n, akkor $\{a/R \mid a \in A\}$ partíciót alkot A-n:*

$$\bigcup_{a \in A} a/R = A \quad \text{és} \quad a/R \neq b/R \Rightarrow a/R \cap b/R = \emptyset.$$

Bizonyítás. Az előző sorban szereplő két állítást az alábbi pontokban ellenőrizzük.

- Ha $a \in A$, akkor $a \in a/R$ (reflexivitás miatt), tehát

$$A \subseteq \bigcup_{a \in A} a/R \subseteq A,$$

ezért $\bigcup_{a \in A} a/R = A$.

- Tegyük fel, hogy $a/R \cap b/R \neq \emptyset$, azaz $\exists c \in a/R \cap b/R$. Ekkor $(a,c) \in R$ és $(b,c) \in R$, tehát $(a,b) \in R$ (szimmetria és tranzitivitás miatt), ezért $a/R = b/R$. Azt kaptuk, hogy ha R két ekvivalencia-osztályának nem üres a metszete, akkor a két ekvivalencia-osztály azonos egymással. Emiatt R különböző ekvivalencia-osztályai páronként diszjunktak egymástól.

□

Az előző állítás megfordítható: az A halmaz egy tetszőleges osztályfelbontásához (partíciójához) tekintsük azt az R relációt A-n, melyre $(a,b) \in R$ pontosan akkor teljesül, ha a és b ugyanabban a partíció-blokkban van. Ekkor R ekvivalenciareláció lesz, és a kiindulásul vett partíciót határozza meg.

1.8. Definíció (Rendezési reláció)**.**

- *Egy $R \subseteq A \times A$ reláció **rendezési reláció** A-n, ha teljesül, hogy:*
  1. $\forall a \in A : (a,a) \in R$ *(reflexivitás);*
  2. $\forall a,b \in A : (a,b) \in R \wedge (b,a) \in R \Rightarrow a = b$ *(antiszimmetria);*
  3. $\forall a,b,c \in A : (a,b) \in R \wedge (b,c) \in R \Rightarrow (a,c) \in R$ *(tranzitivitás).*

Az A halmazt a rendezés alaphalmazának nevezzük.

- *Egy rendezési reláció **teljes rendezés**, ha:*

$$\forall a,b \in A: \quad (a,b) \in R \quad \vee \quad a = b \quad \vee \quad (b,a) \in R \quad \textit{(trichotómia).}$$

- *Egy rendezési reláció **jólrendezés**, ha az alaphalmaz minden nemüres részhalmazának van legkisebb eleme.*

1.9. Példa. *$\mathbb{N}$ a szokásos rendezésével jólrendezett; $\mathbb{Z}$ a szokásos rendezésével teljesen rendezett, de nem jólrendezett.*

1.1. Archimédészi axióma, egészrészek, sűrűségi és approximációs tételek

1.10. Definíció (Archimédészi axióma)**.**

$$\forall a \in \mathbb{R}\ \exists n \in \mathbb{N} : a < n.$$

1.11. Definíció (Felső egészrész)**.** *Legyen $x \in \mathbb{R}$. Ekkor $\lceil x \rceil$ a legkisebb $n \in \mathbb{Z}$, melyre $x \leq n$.*

1.12. Definíció (Alsó egészrész)**.** *Legyen $x \in \mathbb{R}$. Ekkor $\lfloor x \rfloor$ a legnagyobb $n \in \mathbb{Z}$, melyre $n \leq x$.*

1.13. Tétel. *Minden valós számnak létezik felső egészrésze.*

Bizonyítás. Legyen $x \in \mathbb{R}$ és $A = \{n \in \mathbb{N} \mid x \leq n\}$. Az Archimédészi axióma miatt $A \neq \emptyset$. Mivel $\mathbb{N}$ jólrendezett, A-nak van legkisebb eleme. □

1.14. Megjegyzés. *Ebből következik, hogy minden valós szám felírható*

$$x = \lfloor x \rfloor + \{x\}$$

alakban, ahol $\{x\} = x - \lfloor x \rfloor$ *az* x *törtrésze (*$0 \leq \{x\} < 1$*). Emiatt minden valós számnak van alsó egészrésze is.*

1.15. Tétel (Racionális számok sűrűsége)**.**

$$\big(\forall a, b \in \mathbb{R}, a < b\big)\ \exists c \in \mathbb{Q} : a < c < b.$$

Bizonyítás. Az Archimédészi axióma miatt $\exists n \in \mathbb{N} : \frac{1}{n} < b - a$. Osszuk fel az $[\lfloor a \rfloor, \lfloor a \rfloor + 1]$ intervallumot n darab, egyenként $\frac{1}{n}$ hosszú, (balról zárt, jobbról nyílt) részintervallumra:

$$\Big[\lfloor a \rfloor, \lfloor a \rfloor + \frac{1}{n}\Big), \quad \Big[\lfloor a \rfloor + \frac{1}{n}, \lfloor a \rfloor + \frac{2}{n}\Big) \ ...$$

Ezek közül pontosan az egyik (mondjuk a k-adik) kis intervallum fogja a-t tartalmazni. Az n választása miatt a és b távolsága $\frac{1}{n}$-nél nagyobb, ezért az a-t tartalmazó (k-adik) kis intervallum nem tartalmazza b-t. Tehát $\lfloor a \rfloor + \frac{k}{n}$ racionális szám a és b között.

□

1.16. Tétel (Dirichlet approximációs tétele)**.** *Legyen* $\alpha \in \mathbb{R}$ *és* $n \in \mathbb{N}$*,* $n > 0$*. Ekkor* $\exists p, q \in \mathbb{Z}$*:*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{nq} \quad \textit{és} \quad 1 \leq q \leq n.$$

Bizonyítás. Osszuk fel a $[0, 1)$ intervallumot n darab, egyenként $\frac{1}{n}$ hosszú, balról zárt, jobbról nyílt kis részintervallumra:

$$\Big[0, \frac{1}{n}\Big), \quad \Big[\frac{1}{n}, \frac{2}{n}\Big), ...$$

Tekintsük a $0, \{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}$ számokat, ahol $\{x\} = x - \lfloor x \rfloor$. Ez $n + 1$ darab szám a $[0, 1)$ intervallumban. Ezért a skatulyaelv miatt lesz köztük két szám (mondjuk $\{i\alpha\}$ és $\{j\alpha\}$, ahol $0 \leq i < j \leq n$), melyek ugyanabba a kis, $\frac{1}{n}$ hosszú intervallumba esnek. Ekkor

$$|\{j\alpha\} - \{i\alpha\}| < \frac{1}{n}.$$

De

$$\{j\alpha\} - \{i\alpha\} = (j\alpha - \lfloor j\alpha \rfloor) - (i\alpha - \lfloor i\alpha \rfloor) = (j-i)\alpha - (\lfloor j\alpha \rfloor - \lfloor i\alpha \rfloor).$$

Legyen $q = j - i$ és $p = \lfloor j\alpha \rfloor - \lfloor i\alpha \rfloor$. Ezekkel az előző sor szerint

$$|q\alpha - p| < \frac{1}{n} \quad \text{ezért} \qquad \left|\alpha - \frac{p}{q}\right| < \frac{1}{nq},$$

ahogy állítottuk. □

1.17. Megjegyzés. Dirichlet előző (1.16) tételében, ha n értékét rögzítjük, akkor csak véges sok olyan pozitív q van, ami nevezőként felléphet, ezért csak véges sok (p, q) párra teljesülhet a Dirichlet-tétel következménye. Az alábbi, 1.18 Tételben azt mutatjuk meg, hogy minden irracionális α számhoz végtelen sok olyan racionális $\frac{p}{q}$ szám van, melyek $\frac{1}{q^2}$-nél közelebb vannak α-hoz.

1.18. Tétel (Irracionális számok approximációja)**.** *Ha $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, akkor végtelen sok különböző (p, q) számpár van, amelyre:*

$$(*) \qquad \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Bizonyítás. Indirekt tegyük fel, hogy csak véges sok ilyen (p, q) pár van. Legyen

$$\varepsilon = \min\left\{ \left|\alpha - \frac{p}{q}\right| : (p,q) - \text{ra} \;\; (*) \;\; \text{teljesül} \right\}.$$

Mivel $\alpha \notin \mathbb{Q}$, ezért $\varepsilon > 0$. Az Archimédészi axióma miatt $\exists n \in \mathbb{N} : \frac{1}{n} < \varepsilon$. A Dirichlet-tétel szerint ehhez az n-hez is van $p', q' \in \mathbb{Z}$:

$$\left|\alpha - \frac{p'}{q'}\right| < \frac{1}{nq'} \leq \frac{1}{n} < \varepsilon.$$

Ez ellentmondás, hiszen ε definíciója szerint a (p', q') számpár különbözik minden olyan (p, q) számpártól, melyre $(*)$ teljesül (mert $\frac{p'}{q'}$ ε-nál szigorúan közelebb van α-hoz).

Végül, a teljesség kedvéért megjegyezzük, hogy ha $\alpha \in \mathbb{Q}$, mondjuk $\alpha = \frac{p_0}{q_0}$, akkor nagyon könnyű végtelen sok olyan számpárt találni, melyre a tétel következménye fennáll. Ugyanis ekkor tetszőleges pozitív $n \in \mathbb{N}$-re

$$\alpha - \frac{np_0}{nq_0} = \alpha - \frac{q_0}{q_0} = 0,$$

ezért a végtelen sok (np_0, nq_0) számpár mindegyike kielégíti $(*)$-ot. □

1.2. Jólrendezések, rekurziók és a teljes indukció

1.19. Tétel. *Legyen $(A, \leq)$ rendezett halmaz. A következő két állítás ekvivalens:*

1. *$(A, \leq)$ jólrendezett;*
2. *A-ban nincs végtelen leszálló lánc, azaz A elemeiből nem képezhető egy szigorúan monoton csökkenő $a_1 > a_2 > a_3 > \cdots$ végtelen sorozat.*

Bizonyítás. (1) $\Rightarrow$ (2): Ellentmondást keresve tegyük fel, hogy - (2)-vel ellentétben - van végtelen leszálló lánc $a_1 > a_2 > a_3 > \cdots$. Mivel nincs legnagyobb természetes szám, ezért az $\{a_1, a_2, a_3, \dots\}$ halmaznak nincs legkisebb eleme, ellentmondás.

(2) $\Rightarrow$ (1): Legyen $\emptyset \neq B \subseteq A$ tetszőleges. (1)-hez azt kell megmutatnunk, hogy B-nek van legkisebb eleme. Mivel B nem üres, ezért van egy $b_0 \in B$ elem.

- Ha b_0 nem a legkisebb B-beli, akkor $\exists b_1 \in B : b_1 < b_0$.
- Ha b_1 nem a legkisebb B-beli, akkor $\exists b_2 \in B : b_2 < b_1$, ... stb.

A (2) feltétel miatt ezt az eljárást csak véges sok lépésen át lehet folytatni, ezért B-ben van legkisebb elem. □

Teljes indukció

1.20. Tétel (Teljes indukció elve)**.** *Legyen P a természetes számok egy tulajdonsága és legyen $n_0 \in \mathbb{N}$. Ha*

1. *$P(n_0)$ igaz, és*
2. *$\forall n \geq n_0 : P(n) \Rightarrow P(n+1)$,*

akkor $\forall n \geq n_0 : P(n)$ igaz.

Az állítás szerint, ha meg akarjuk mutatni, hogy minden n_0-nál nagyobb, vagy egyenlő természetes szám rendelkezik a P tulajdonsággal, akkor ehhez elég megmutatni, hogy

- $P(n_0)$ fennáll (ezt nevezzük **Alaplépés**nek) és
- ha P fennáll valamely $n \geq n_0$ természetes számra, akkor P fennáll $n+1$-re is (ezt nevezzük **Indukciós lépés**nek).

Bizonyítás. Legyen

$$A = \{n \in \mathbb{N} \mid n \geq n_0 \text{ és } P(n) \text{ nem igaz}\}.$$

Ellentmondást keresve tegyük fel, hogy $A \neq \emptyset$. Mivel $\mathbb{N}$ jólrendezett, A-nak van legkisebb eleme, legyen ez n. Ekkor $P(n)$ hamis (mert $n \in A$) és $n > n_0$, mert $P(n_0)$ igaz. Ezért $n - 1 \geq n_0$ és $P(n-1)$ igaz (mert n volt a legkisebb természetes szám, melyre P nem igaz). De az indukciós lépés miatt $P(n-1)$ igazságából $P(n)$ igazsága is következik; ez ellentmondás.

Tehát $A = \emptyset$, azaz $\forall n \geq n_0 : P(n)$ igaz. □

1.21. Definíció (Rekurzió)**.** *Egy sorozatot **rekurzívan** definiálunk, ha a sorozat tagjait saját, kisebb indexű tagjai felhasználásával állítjuk elő.*

Rekurzív sorozatokra klasszikus példa a Fibonacci-számok sorozata. Részletesebben, a nulladik és első Fibonacci-számok az $f_0 = 0$ és $f_1 = 1$; a sorozat további tagjait az előző (kisebb indexű) tagjai segítségével az $f_{n+1} = f_n + f_{n-1}$ rekurzióval definiáljuk. A XIII. században ezt a sorozatot Fibonacci a nyulak szaporodásának modellezése érdekében tanulmányozta, de nem Ő vezette be (korábban is ismert volt már).

Egy második példa bemutatása érdekében emlékeztetünk rá, hogy adott n természetes számra $n!$ jelöli az első n darab pozitív egész szám szorzatát: $n! = 1 \cdot 2 \cdot ... \cdot n$ továbbá az $n!$ kifejezést „n-faktoriálisának" nevezzük. A faktoriális függvény is kényelmesen definiálható rekurzióval: legyen $1! = 1$, ekkor nagyobb számok faktoriálisai megadhatók az

$$(n + 1)! = (n + 1) \cdot n!$$

alakban; tehát a fenti módon ebben a sorozatban is, a korábbi n. tagja segítségével adtuk meg a sorozat későbbi $(n + 1)$-edik tagját.

2. Algebrai struktúrák általában

Nyugtalanító, és nehezen megválaszolható az a kérdés, hogy „a számok micsodák?" Pl. a természetes számok kavicskupacok? Vagy pénzkötegek? Vagy...? Az ilyen irányú vizsgálódásokból az derült ki, hogy az esetek többségében MINDEGY IS, hogy konkrétan a természetes számok micsodák, a fontos az, hogy - bármik is legyenek pl. a természetes számok - milyen összefüggések teljesülnek a rajtuk végzett műveletekre (vagy még sarkosabban: milyen számolási szabályok vonatkoznak rájuk). A kavicskupacokkal és pénzkötegekkel kapcsolatos érvelésekben az pl. közös, hogy

„x kupac + y kupac kavics összesen ugyanannyi, mint y kupac + x kupac kavics",

és ugyanígy,

„x köteg + y köteg bankjegy összesen ugyanannyi, mint y köteg + x köteg bankjegy".

Tehát nem az a fontos, hogy kavicskupacokat, vagy pénzkötegeket adunk össze, hanem az, hogy a rájuk vonatkozó számolási szabályok azonosak. Ennek megfelelően, az algebra a „műveletek tulajdonságait" vizsgálja. Ehhez tisztázni kell, mit tekintünk műveleteknek. Erre egy meglehetősen általános válasz a következő.

2.1. Definíció (Általános algebra)**.**
Legyen A nemüres halmaz, legyen I indexhalmaz, és minden $i \in I$-re legyen f_i egy (esetleg sokváltozós) függvény A-n. Ekkor az

$$\mathcal{A} = \langle A, f_i \rangle_{i \in I}$$

rendszert algebrai struktúrának (vagy általános algebrának) nevezzük. Itt A az alaphalmaz, az f_i-k pedig az $\mathcal{A}$ algebra műveletei.

Az előző definíció jelöléseit megtartva tehát egy $\mathcal{A}$ általános algebrát úgy adunk meg, hogy tisztázzuk, mi az alaphalmaza, hány művelet van rajta (azaz mi az I indexhalmaz), és hogy az egyes műveletek külön-kölön hány változósak. **Semmilyen megkötés nincs** I méretére: I lehet üres (ekkor nincs művelet az algebránkban), I lehet véges halmaz, de I lehet akár végtelen nagy is (lehet megszámlálhatóan végtelennél is nagyobb). Mindegyik f_i művelet lehet 1-változós, vagy 2-változós, vagy akár mégtöbb változós; általános algebrákban a műveletek változói számára mindössze azt a korlátot szabjuk, hogy minden műveletnek csak véges sok változója van. Azt azonban fontos hangsúlyozni, hogy minden művelet az alaphalmaz minden olyan sorozatán értelmezve van, melynek hossza megegyezik a művelet változóinak számával (azaz minden 2-változós művelet értelmezve van A^2 minden elemén, minden 3-változós értelmezve van A^3 minden elemén, stb.) Megengedünk 0-változós műveleteket is, ezeket az A alaphalmaz kitüntetett elemeivel azonosíthatjuk. (Ez

az azonosítás csak egy nem túl fontos technikai részlet, kezelhetnénk a kitüntetett elemeket külön, de akkor sok érvelés hosszabb lenne). Egy általános algebrában lehet pl. 10 darab 5-változós + megszámlálhatóan végtelen 17-változós, + 45 darab 70-változós művelet. A klasszikus esetekben I-nek 1, 2, 3, vagy 4 eleme van; a fenti értelemben pl. minden félcsoport, csoport, gyűrű, stb. egyúttal általános algebra is.

Az általános algebrák absztrakciós szintjén is sok érdekes és távolról sem triviális tétel igazolható. Mi azonban ebben a félévben a klasszikus algebrai vizsgálatokban felmerülő algebrai struktúrákkal foglalkozunk, melyeket elsősorban a számelmélet, az egyenletek megoldása, vagy a közönséges sík, illetve tér geometriai vizsgálódásai inspiráltak. Algebrai kalandozásainkat a gyűrűkkel kezdjük, és számelméleti indíttatású kérdésekkel motiváljuk.

2.2. Definíció (Gyűrű)**.** *Egy* $(A, +, \cdot, 0, 1)$ *struktúra* ***gyűrű****, ha* $+$ *és* $\cdot$ *2-változós,* 0 *és* 1 *pedig 0-válzozós műveletek, és teljesül, hogy*

1. $\forall x, y, z \in A : (x + y) + z = x + (y + z)$ *(azaz* a *+ művelet asszociatív);*
2. $\forall x, y \in A : x + y = y + x$ *(azaz* a *+ művelet kommutatív);*
3. $\forall x \in A : 0 + x = x$ *(azaz létezik additív neutrális elem);*
4. $\forall x \in A\ \exists y \in A : x + y = 0$ *(azaz létezik additív inverz);*
5. $\forall x, y, z \in A : (x \cdot y) \cdot z = x \cdot (y \cdot z)$ *(azaz* a *·* *művelet asszociatív);*
6. A *·* *művelet disztributív* a *+* *műveletre:*

$$\forall x, y, z \in A : x(y + z) = xy + xz \quad \text{és} \quad (x + y)z = xz + yz.$$

A gyűrű ***kommutatív****, ha még* $\forall x, y \in A : xy = yx$ *is teljesül.*

2.3. Példa.

- $(\mathbb{N}, +, \cdot, 0, 1)$ *nem gyűrű (nincs additív inverz);*
- $(\mathbb{Z}, +, \cdot, 0, 1)$ *kommutatív gyűrű;*
- $(\mathbb{R}, +, \cdot, 0, 1)$ *kommutatív gyűrű;*
- $\{n \in \mathbb{Z} \mid n$ *páratlan*$\}$ *nem gyűrű (nem zárt az összeadásra).*

2.4. Tétel ($\sqrt{2}$ irracionalitása)**.** $\sqrt{2} \notin \mathbb{Q}$.

Erre a jólismert állításra olyan bizonyítást adunk, mely nem használja a Számelmélet Alaptételét (melyet csak később, a 3.34 tételben igazolunk majd). Ehelyett a természetes számok jólrendezettségére fogunk alapozni.

Bizonyítás. Tegyük fel indirekt, hogy $\sqrt{2} \in \mathbb{Q}$. Ekkor $\exists p, q \in \mathbb{Z}^+$ relatív prímek: $\sqrt{2} = \frac{p}{q}$. Ezekre $p = \sqrt{2}q$. Tekintsük a következő halmazt:

$$A = \{n \in \mathbb{N}^+ \mid \sqrt{2}n \in \mathbb{N}\}.$$

$A \neq \emptyset$, mert $q \in A$. Mivel $\mathbb{N}$ jólrendezett, A-nak van legkisebb eleme, legyen ez a. Mivel $a \in A$, ezért $\sqrt{2}a \in \mathbb{N}$. Továbbá $(\sqrt{2} - 1)a < a$, mert $\sqrt{2} - 1 < 1$. Ezért ellentmondást kapunk, ha megmutatjuk, hogy $(\sqrt{2} - 1)a \in A$ (mert A-ban a volt a

legkisebb elem).

Az A halmaz definíciója szerint $(\sqrt{2}-1)a \in A$-hoz a következő két állítást kell igazolnunk:

(1) $(\sqrt{2}-1)a \in \mathbb{N}^+$ és
(2) $\sqrt{2}((\sqrt{2}-1)a) \in \mathbb{N}$.

(1) teljesül, mert

$$(\sqrt{2}-1)a = \sqrt{2}a - a \overset{a\in A}{\in} \mathbb{N}^+.$$

(2) is teljesül, mert

$$\sqrt{2}(\sqrt{2}-1)a) = 2a - \sqrt{2}a \overset{a\in A}{\in} \mathbb{N},$$

és ezzel készen vagyunk. □

2.1. Gyűrűk és oszthatóság

2.5. Definíció (Oszthatóság)**.** *Legyen R gyűrű, $a, b \in R$. Azt mondjuk, hogy a **osztója** b-**nek** (jelölés: $a \mid b$), ha $\exists c \in R : ac = b$.*

Legyen R gyűrű. Figyeljük meg, hogy 0-nak minden $a \in R$ osztója, hiszen ha $a \in R$ tetszőleges, akkor $c = 0$ mutatja, hogy van olyan c, melyre $ac = 0$. A másik végletet az olyan elemek alkotják, melyek minden gyűrűelemnek osztói, ezeket nevezzük egységeknek:

2.6. Definíció (Egység)**.** *Az $e \in R$ gyűrűelem **egység**, ha $\forall x \in R : e \mid x$ (minden elemet oszt).*

2.7. Tétel. *$\mathbb{Z}$-ben a szokásos műveletekkel $e \in \mathbb{Z}$ egység $\Leftrightarrow e = 1$ vagy $e = -1$.*

Bizonyítás. Ha $e \in \mathbb{Z}$ egység, akkor $e \mid 1$, tehát $|e| \leq 1$. Mivel 0 nem osztója 1-nek, ezért 0 nem lehet egység. Tehát $e = 1$ vagy $e = -1$, ezeken kívül más egység nem lehet $\mathbb{Z}$-ben. Fordítva: világos, hogy $e = 1$ és $e = -1$ valóban egységek. □

2.8. Példa. *Legyen $A = \{x \in \mathbb{Z} \mid x \text{ páros}\}$ és $R = (A, +, \cdot)$ a szokásos műveletekkel.*

- *R gyűrű (tényleg).*
- *R-ben $2 \not\mid 10$ mert - mivel 4 nem osztója 10-nek - 2-t páros számokkal (azaz A elemeivel) szorozva sosem kaphatunk 10-et. Hasonlóan, $-2 \not\mid 10$.*
- *R-ben nincs egység, mert ha $e \in R$ egység lenne, akkor $e \mid 2$ azaz $e = 2$ vagy $e = -2$ következne, de az előbb láttuk, hogy ezek egyike sem osztója 10-nek, tehát nem lehetnek egységek.*

2.9. Definíció (Egységelemes gyűrű)**.** *R gyűrű **egységelemes**, ha*

$$\exists 1 \in R : \forall x \in R : 1 \cdot x = x \cdot 1 = x.$$

2.10. Definíció (Nullosztómentes gyűrű)**.** *R gyűrű* ***nullosztómentes****, ha*

$$\forall x, y \in R : xy = 0 \Rightarrow (x = 0 \vee y = 0).$$

2.11. Definíció (Valódi osztó)**.** *a valódi osztója b-nek, ha $a \mid b$, $a \neq b$ és a nem egység.*

2.12. Tétel. Az Oszthatóság alaptulajdonságai. *Legyen R kommutatív, egységelemes, nullosztómentes gyűrű. Ekkor az oszthatóság relációja:*

1. *Reflexív: $\forall x \in R : x \mid x$;*
2. *„Majdnem antiszimmetrikus":*

$$\forall x, y \in R : (x \mid y \wedge y \mid x) \Leftrightarrow \big(\exists e, f \in R \ \text{ egység: } \ x = ey,\ y = fx\big);$$

3. *Tranzitív: $\forall x, y, z \in R : (x \mid y \wedge y \mid z) \Rightarrow x \mid z$;*
4. *$\forall x, y, a, b \in R : (x \mid y \wedge x \mid z) \Rightarrow x \mid (ay + bz)$.*

Bizonyítás. $x = 1 \cdot x$ mutatja: $x \mid x$, ezért 1. teljesül.

3. igazolásához TFH $x \mid y$, $y \mid z$. Ekkor $\exists u, v : xu = y$, $yv = z$, tehát

$$x(uv) = (xu)v = yv = z,$$

azaz uv mutatja, hogy $x \mid z$.

2. igazolását két részre bolntjuk.
$\Rightarrow$: Ha $x = y = 0$, akkor $e = f = 1$-el $x = ey,\ y = fx$ nyilván teljesül. Emiatt feltehetjük, hogy $x \neq 0$. TFH $x \mid y$ és $y \mid x$. Ez azt jelenti, hogy vannak olyan $u, v \in R$ elemek, hogy $y = ux$ és $x = vy$. De ekkor $x = vy = vux$ azaz $0 = (vu-1)x$. Mivel $x \neq 0$ és R nullosztómentes, ezért $vu - 1 = 0$, azaz $vu = 1$, speciálisan $u \mid 1$ és $v \mid 1$, és mivel 1 egység, ezért a már igazolt 3. miatt u is, v is egység. Ezért az $e = v, f = u$ egységekkel $x = ey,\ y = fx$ teljesül.

$\Leftarrow$: TFH e, f olyan egységek, hogy $ex = y$, $fy = x$. Ezekből $x \mid y$ és $y \mid x$ azonnal adódik (az első oszthatóságot e, a másodikat f mutatja).

Végül 4.-hez TFH $x \mid y$, $x \mid z$. Ekkor $\exists u, v : xu = y$, $xv = z$, amiből

$$ay + bz = axu + bxv = x(au + bv),$$

tehát $x \mid (ay + bz)$. □

2.13. Megjegyzés. *$\mathbb{N}$-ben (csak 1 az egység) a $P = \{(x, y) \in \mathbb{N} \mid x \mid y\}$ reláció reflexív, antiszimmetrikus, tranzitív, tehát részbenrendezés.*

3. Az egész számok számelméletének alapjai

E fejezet szerepe kettős: egyrészt szeretnénk összefoglalni néhány alapvető számelméleti fogalmat és eredményt, másrészt szeretnénk előkészíteni és motiválni a későbbi algebrai vizsgálódásainkat.

3.1. Maradékos osztás, számrendszerek

3.1. Tétel (Maradékos osztás $\mathbb{N}$-ben)**.** *Legyen $a, b \in \mathbb{N}$, $b > 0$. Ekkor $\exists! q, r \in \mathbb{N}$:*

$$a = qb + r \quad \textit{és} \quad 0 \leq r < b.$$

Bizonyítás. **Létezés**: Legyen $q = \max\{n \in \mathbb{N} \mid nb \leq a\}$, $r = a - qb$. Ekkor $0 \leq r < b$, mert $(q+1)b > a$.

Egyértelműség: TFH $a = q_1 b + r_1 = q_2 b + r_2$, ahol $0 \leq r_1, r_2 < b$. Ekkor $0 = (q_1 - q_2)b + (r_1 - r_2)$, tehát $b \mid (r_1 - r_2)$. De $|r_1 - r_2| < b$, ezért $r_1 = r_2$ és emiatt $q_1 = q_2$. □

3.2. Tétel (Számrendszerbeli alak)**.** *Legyen $t \in \mathbb{N}$, $t \geq 2$, $b \in \mathbb{N}$. Ekkor $\exists n \in \mathbb{N}$ és $a_0, \ldots, a_n \in \mathbb{N}$, $0 \leq a_i < t$:*

$$b = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0.$$

Bizonyítás. Teljes indukció b-re.

Alaplépés: $b = 0$, ekkor $n = 0$, $a_0 = 0$ jó.

Indukciós lépés: Legyen a_0 a b t-vel való osztási maradéka, $b' = \frac{b - a_0}{t}$. Az indukciós feltevés szerint $b' = a_n t^{n-1} + \cdots + a_1$, tehát $b = a_n t^n + \cdots + a_1 t + a_0$. □

3.3. Példa (10-esből 3-as számrendszerbe)**.** *Átalakítás 525-öt 3-as számrendszerbe:*

b_m	*525*	*175*	*58*	*19*	*6*	*2*
a_m	*0*	*1*	*1*	*1*	*0*	*2*

Tehát $525 = 2 \cdot 3^5 + 0 \cdot 3^4 + 1 \cdot 3^3 + 1 \cdot 3^2 + 1 \cdot 3^1 + 0 \cdot 3^0 = 201110_3$.

3.4. Példa (3-asból 10-es számrendszerbe)**.** *Átalakítás* 201110_3*-ból 10-esbe Horner-módszerrel:*

$$((((2 \cdot 3 + 0) \cdot 3 + 1) \cdot 3 + 1) \cdot 3 + 1) \cdot 3 + 0 = 525.$$

3.5. Tétel (Horner-módszer)**.** *Legyen $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$. Ekkor:*

$$f(x) = (((a_n x + a_{n-1})x + a_{n-2})x + \cdots + a_1)x + a_0$$

3.6. Példa. *Számítsuk ki $f(5) = 3x^3 + 2x^2 - 7x + 1$ értékét Horner-módszerrel:*

	3	*2*	*-7*	*1*
5	↓	*15*	*85*	*390*
	3	*17*	*78*	*391*

Tehát $f(5) = 391$.

3.2. Közös osztók és többszörösök és az euklideszi algoritmus

3.7. Definíció (Közös osztók)**.** *Legyen R gyűrű, $a, b \in R$. Ekkor az a és b* ***közös osztóinak halmaza****:*

$$K(a,b) = \{c \in R \mid c \mid a \wedge c \mid b\}.$$

3.8. Definíció (Kitüntetett közös osztó)**.** *$d \in R$* ***kitüntetett közös osztója*** *a-nak és b-nek (jelölés: $d = (a,b)$), ha:*

1. *$d \in K(a,b)$;*
2. *$\forall c \in K(a,b) : c \mid d$.*

3.9. Definíció (Legnagyobb közös osztó $\mathbb{Z}$-ben)**.** *$\mathbb{Z}$-ben:*

$$\mathrm{LNKO}(a,b) = \max(K(a,b) \cap \mathbb{N}).$$

3.10. Megjegyzés.

- *Kitüntetett közös osztó értelmezhető kizárólag a gyűrű-műveletekkel, de a legnagyobb közös osztó fogalmához a gyűrűn külön meg kell adni egy további rendezési relációt is.*
- *$\mathrm{LNKO}(0,0)$-nak nincs legnagyobb eleme, mert $K(0,0) = \mathbb{Z}$.*
- *Konvenció:* $\mathrm{LNKO}(0,0) = 0$.
- *Ha $a \neq 0$ vagy $b \neq 0$, akkor $\mathbb{Z}$-ben $K(a,b)$ véges.*
- *Alább, 3.13-ben lesz: $\mathbb{Z}$-ben minden számpárnak van nemnegatív kitüntetett közös osztója.*

3.11. Definíció (Közös többszörösök)**.** *Legyen R gyűrű, $a, b \in R$. Ekkor a* ***közös többszörösök halmaza****:*

$$KT(a,b) = \{c \in R \mid a \mid c \wedge b \mid c\}.$$

3.12. Definíció (Kitüntetett közös többszörös)**.** *$m \in R$ az a és b* ***kitüntetett közös többszöröse*** *(jelölés: $m = [a,b]$), ha:*

1. *$m \in KT(a,b)$ és*
2. *$\forall c \in KT(a,b) : m \mid c$.*

Legyen $a, b \in \mathbb{Z}$, $b \neq 0$. Definiáljuk az $r_0, r_1, \ldots$ sorozatot az alábbi rekurzív módon:

$$\begin{aligned} r_0 &= a; \\ r_1 &= b; \\ r_{k+1} &= r_{k-1} \text{ osztási maradéka } r_k\text{-val.} \end{aligned}$$

Euklideszi algoritmusnak azt az algoritmust nevezzük, melynek bemenete az (a,b) számpár, és az előző rekurziót követve felépíti az $r_0, r_1, \ldots$ sorozatot a legkisebb olyan n-ig, melyre $r_n = 0$; ekkor az algoritmus eredménye (kimenete) az r_{n-1} érték. Alább látjuk majd, hogy mindíg van olyan n, melyre $r_n = 0$, ezért az euklideszi algoritmus mindig befejeződik.

3.13. Tétel (Euklideszi algoritmusra vonatkozó tétel)**.** *Akármilyen $a, b \in \mathbb{Z}$, $b \neq 0$ számokon is indítjuk el az euklideszi algoritmust, az véges sok lépésben mindig befejeződik, és r_{n-1} kimenetére teljesül, hogy $r_{n-1} = \text{LNKO}(a, b)$.*

Bizonyítás. Tetszőleges módon válasszuk és rögzítsük az $a, b \in \mathbb{Z}$, $b \neq 0$ bemenetet. Az euklideszi algoritmus során előálló értékekre $|r_0| > |r_1| > |r_2| > \cdots$. Mivel $\mathbb{N}$ jólrendezett, ezért az 1.19 Tétel miatt nincs benne végtelen leszálló lánc. Emiatt az euklideszi algoritmus véges sok lépésben befejeződik (mondjuk n a legkisebb szám, melyre $r_n = 0$). Mivel minden lépésben maradékos osztásokat végeztünk, ezért vannak olyan $q_1, q_2...$ számok, hogy minden $1 \leq k < n$-re

$$r_{k-1} = q_k r_k + r_{k+1} \text{ azaz}$$
$$r_{k+1} = r_{k-1} - q_k r_k.$$

Ezért

$$K(r_0, r_1) = K(r_1, r_2) = \cdots = K(r_{n-1}, r_n) = K(r_{n-1}, 0)$$

és emiatt

$$\text{LNKO}(r_0, r_1) = \text{LNKO}(r_{n-1}, 0) = r_{n-1}.$$

Tehát a végeredmény valóban az a és b legnagyobb közös osztója. □

3.14. Tétel (Kiterjesztett euklideszi algoritmus)**.** *Tetszőleges $a, b \in \mathbb{Z}$-hez van $\delta, \beta \in \mathbb{Z}$, hogy*

$$(a, b) = \delta a + \beta b.$$

Bizonyítás. m-re vonatkozó indukcióval belátjuk:

$$\forall m \ \exists \delta_m, \beta_m \in \mathbb{Z} : r_m = \delta_m \cdot a + \beta_m \cdot b.$$

$m = 0$-ra: $r_0 = a = 1 \cdot a + 0 \cdot b \quad \Rightarrow \delta_0 = 1, \ \beta_0 = 0.$
$m = 1$-re: $r_1 = b = 0 \cdot a + 1 \cdot b \quad \Rightarrow \delta_1 = 0, \ \beta_1 = 1.$

Indukciós lépés: TFH $\forall k \leq m$-re tudjuk, hogy $\exists \delta_k, \beta_k$:

$$r_k = \delta_k \cdot a + \beta_k \cdot b.$$

Ezt a 3.13 Tétel bizonyításában bevezetett jelölésekkel kombinálva:

$$r_{m+1} = r_{m-1} - q_m r_m = (\delta_{m-1} a + \beta_{m-1} b) - q_m(\delta_m a + \beta_m b)$$
$$= (\delta_{m-1} - q_m \delta_m) a + (\beta_{m-1} - q_m \beta_m) b.$$

Tehát

$$\delta_{m+1} = \delta_{m-1} - q_m \delta_m, \quad \beta_{m+1} = \beta_{m-1} - q_m \beta_m$$

mutatják, hogy állításunk $m + 1$-re is igaz, és ezzel készen vagyunk az indukcióval. □

3.15. Definíció (Relatív prímek)**.** *Legyen R gyűrű, $a, b \in R$. a és b* ***relatív prímek****, ha $(a, b) = 1$.*

3.16. Tétel (Relatív prímek szabálya)**.** *Ha $a, b, c \in \mathbb{Z}$, $a \mid bc$ és $(a, b) = 1$, akkor $a \mid c$.*

Bizonyítás. Mivel $(a, b) = 1$, $\exists x, y \in \mathbb{Z} : ax+by = 1$. Szorozzuk c-vel: $acx+bcy = c$. Nyilván $a \mid acx$ és feltettük, hogy $a \mid bc$, tehát $a \mid bcy$; ezekből $a \mid c$. □

3.3. Lineáris diofantoszi egyenletek

Diofantoszi egyenleteknek olyan (esetleg többváltozós) egyenleteket nevezünk, melyeknek a megoldásait az egész számok körében keressük. Speciálisan, a **lineáris diofantoszi egyenletek** a következők: legyen $a, b, c \in \mathbb{Z}$, $a \neq 0$, $b \neq 0$. Keressük az összes olyan $x, y \in \mathbb{Z}$ párt, hogy

$$ax + by = c.$$

3.17. Tétel. *Mint előbb, legyen $a, b, c \in \mathbb{Z}$, $a \neq 0$, $b \neq 0$.*

1. *$ax + by = c$-nek van megoldása $\mathbb{Z}$-ben $\Leftrightarrow (a, b) \mid c$;*
2. *Ha (x_0, y_0) megoldás, akkor az összes megoldás:*

$$x = x_0 + \frac{b}{d}t, \quad y = y_0 - \frac{a}{d}t, \quad t \in \mathbb{Z},$$

ahol $d = (a, b)$.

Bizonyítás. 1. **Megoldhatóság**:

- TFH x_0, y_0 megoldás. Azt szeretnénk belátni, hogy $d \mid c$. Mivel

$$d \mid a \quad \text{és} \quad d \mid b \;\; \text{ezért} \;\; d \mid (ax_0 + by_0) = c.$$

- Fordítva, TFH $d \mid c$. Azt szeretnénk belátni, hogy van megoldás. A kiterjesztett euklideszi algoritmus miatt van $\delta, \beta \in \mathbb{Z}$:

$$d = \delta a + \beta b.$$

Továbbá, $d \mid c$, ezért $\exists t : dt = c$. Ezekkel

$$c = dt = (\delta a + \beta b)t = (\delta t)a + (\beta t)b$$

Tehát

$$x_0 = \delta t, \quad y_0 = \beta t \quad \text{megoldás.}$$

2. **TFH** (x_0, y_0) **megoldás.** Legyen t tetszőleges. Belátjuk, hogy $x = x_0 + \frac{b}{d}t$, $y = y_0 - \frac{a}{d}t$ is megoldás. Helyettesítsünk be:

$$a\left(x_0 + \frac{b}{d}t\right) + b\left(y_0 - \frac{a}{d}t\right) = ax_0 + a\frac{b}{d}t + by_0 - b\frac{a}{d}t = ax_0 + by_0 = c,$$

az utolsó lépésben azt használtuk, hogy (x_0, y_0) megoldás.

3. **TFH** (x_0, y_0) **és** (x_1, y_1) **is megoldások.** Belátjuk:

$$\exists t : x_1 = x_0 + \frac{b}{d}t, \quad y_1 = y_0 - \frac{a}{d}t.$$

Mivel (x_0, y_0) és (x_1, y_1) megoldások, ezért

$$\text{I. } ax_0 + by_0 = c, \quad \text{II. } ax_1 + by_1 = c.$$

Vonjuk ki II-ből I-t:

$$a(x_1 - x_0) + b(y_1 - y_0) = 0, \quad \text{azaz}$$

$$a(x_1 - x_0) = b(y_0 - y_1).$$

Osszunk le d-vel:

$$(*) \qquad \frac{a}{d}(x_1 - x_0) = \frac{b}{d}(y_0 - y_1).$$

Mivel $d \mid a$, $d \mid b$, ezért $\frac{a}{d}$ és $\frac{b}{d}$ egész számok. Továbbá, $\frac{a}{d}$ osztója a baloldalnak, ezért osztója a jobboldalnak is. De $\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, ezért a relatív prímek szabálya (3.16 Tétel) szerint

$$\frac{a}{d} \mid (y_0 - y_1) \quad \Rightarrow \quad \exists t : y_0 - y_1 = \frac{a}{d}t \ \text{ azaz } \ y_1 = y_0 - \frac{a}{d}t.$$

Az előző sorban az is kijött, hogy $y_0 - y_1 = \frac{a}{d}t$. Ezt $(*)$-be helyettesítve:

$$x_1 - x_0 = \frac{b}{d}t \quad \Rightarrow \quad x_1 = x_0 + \frac{b}{d}t.$$

□

3.4. Kongruenciák és maradékosztály-gyűrűk

3.18. Definíció. *Legyen $m \in \mathbb{N}^+$, $a, b \in \mathbb{Z}$. Azt mondjuk, hogy a kongruens b-vel modulo m, ha a és b m-el osztva ugyanannyit ad maradékul, azaz $m \mid (a-b)$. Jelölés:*

$$a \equiv b \pmod{m}.$$

Nyilván, az $m \geq 2$ eset az érdekes.

3.19. Tétel. *A* $\pmod m$ *kongruencia reláció ekvivalenciareláció* $\mathbb{Z}$*-n.*

Bizonyítás. Rögzítsük m-et és tekintsük azt a $\varrho : \mathbb{Z} \to \mathbb{N}$ függvényt, mely minden $a \in \mathbb{Z}$-hez a-nak az m-el való osztási maradékát rendeli. Világos, hogy

$$a \equiv b \pmod m \;\Leftrightarrow\; \varrho(a) = \varrho(b),$$

az pedig ismert (vagy könnyen ellenőrízhető), hogy tetszőleges f függvényre az $\{(a, b) : f(a) = f(b)\}$ reláció ekvivalenciareláció f értelmezési tartományán. □

3.20. Tétel. *Legyen* $m \in \mathbb{Z}$, $m \geq 2$, *legyenek* $a, a', b, b' \in \mathbb{Z}$ *és TFH* $a \equiv a' \pmod m$ *és* $b \equiv b' \pmod m$. *Ekkor*

1. $a + b \equiv a' + b' \pmod m$;
2. $ab \equiv a'b' \pmod m$.

Bizonyítás. Mivel $a \equiv a' \pmod m$ és $b \equiv b' \pmod m$, ezért $m \mid (a - a')$ és $m \mid (b - b')$. Ezért van $u, v \in \mathbb{Z}$: $um = a - a'$, $vm = b - b'$.

1. bizonyításával kezdjük. Az előző két sor szerint:

$$(a + b) - (a' + b') = (a - a') + (b - b') = um + vm = (u + v)m,$$

tehát $m \mid [(a + b) - (a' + b')] \Rightarrow a + b \equiv a' + b' \pmod m$, ezért 1. teljesül.

2. igazolására térve (és a bizonyítás első két sorában rögzített jelöléseket továbbra is használva) $a = a' + um$ és $b = b' + vm$. Ezekből

$$ab = (a' + mu)(b' + mv) = a'b' + a'mv + b'mu + m^2uv,$$

és ezért $ab - a'b' = m(a'v + b'u + muv)$. Ez mutatja: $m \mid (ab - a'b') \Rightarrow ab \equiv a'b' \pmod m$. □

Az $a \in \mathbb{Z}$ szám $\pmod m$ szerinti maradékosztálya:

$$\overline{a} = \{x \in \mathbb{Z} \mid x \equiv a \pmod m\}.$$

Világos, hogy m darab maradékosztály van.

3.21. Definíció. *Legyen* $m \in \mathbb{N}$, $m \geq 2$. $\mathbb{Z}_m$ *a következő struktúra (a 2.1 Definíció értelmében):*

- *Alaphalmaz: mod* m *maradékosztályok (*m *db);*
- *Műveletek: tetszőleges* $\overline{a}, \overline{b}$ *maradékosztályokra*

$$\overline{a} + \overline{b} = \overline{a + b};$$
$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

3.22. Tétel. $\mathbb{Z}_m$ *kommutatív gyűrű.*

Bizonyítás. Asszociativitás: $\forall a, b, c \in \mathbb{Z}$:

$$(\overline{a} + \overline{b}) + \overline{c} = \overline{a+b} + \overline{c} = \overline{(a+b)+c} \quad \text{és}$$
$$\overline{a} + (\overline{b} + \overline{c}) = \overline{a} + \overline{b+c} = \overline{a+(b+c)}.$$

De $(a+b)+c = a+(b+c)$, tehát az előző két sor baloldalai egyenlők. Hasonlóan ellenőrízhető a többi gyűrű-axióma. □

3.23. Példa ($\mathbb{Z}_4$ művelettáblája)**.**

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Itt $3 \cdot 3 = 1 \pmod 4$*, mert* $9 \equiv 1 \pmod 4$.

3.24. Példa.
- $\mathbb{Z}_6$*-ban:* $2 \cdot 3 = 0 \pmod 6$*, tehát 2 és 3 nullosztók.*
- $\mathbb{Z}_{34}$*-ben:* $5 \cdot 7 = 1 \pmod{34}$*, ezért* $\mathbb{Z}_{34}$*-ben 5 és 7 egymás inverzei.*

3.5. Lineáris kongruenciák és a kínai maradéktétel

3.25. Definíció (Lineáris kongruenciák)**.** *Legyen* $a, b \in \mathbb{Z}$*,* $m \in \mathbb{N}$*,* $m \geq 2$*.*

$$ax \equiv b \pmod m$$

összes megoldása? (Az ismeretlen x*).*

3.26. Tétel (Kongruenciák egyszerűsítése)**.**

$$ac \equiv bc \pmod m \Leftrightarrow a \equiv b \pmod{\frac{m}{(c,m)}}.$$

Bizonyítás. TFH: $ac \equiv bc \pmod m$. Ekkor $0 \equiv ac - bc \pmod m$, azaz

$$0 \equiv (a-b)c \pmod m, \quad \text{azaz}$$
$$m \mid (a-b)c, \quad \text{azaz} \quad \exists t \in \mathbb{Z} : m \cdot t = (a-b)c.$$

Osszuk le (c, m)-mel:

$$\frac{m}{(c,m)} \cdot t = (a-b) \cdot \frac{c}{(c,m)}.$$

De $\left(\frac{m}{(c,m)}, \frac{c}{(c,m)}\right) = 1$, ezért a relatív prímek szabálya (3.16 Tétel) miatt

$$\frac{m}{(c,m)} \mid (a-b) \quad \text{azaz} \quad a \equiv b \pmod{\frac{m}{(c,m)}}.$$

Megfordítva: TFH $a \equiv b \pmod{\frac{m}{(c,m)}}$, azaz $\frac{m}{(c,m)} \mid (a-b) \Rightarrow \exists t \in \mathbb{Z} : \frac{m}{(c,m)} \cdot t = a - b$. Ekkor:

$$(a-b)c = \frac{m}{(c,m)} \cdot t \cdot c = m \cdot t \cdot \frac{c}{(c,m)}.$$

mutatja: $m \mid (a-b)c$. Tehát $(a-b)c \equiv 0 \pmod m \Rightarrow ac \equiv bc \pmod m$. □

3.27. Tétel (Lineáris kongruencia megoldása)**.**

1. *$ax \equiv b \pmod{m}$ megoldható $\Leftrightarrow (a,m) \mid b$;*
2. *Ha x_0 megoldás, akkor az összes megoldás:*

$$x \equiv x_0 + t \cdot \frac{m}{(a,m)} \pmod{m}, \quad t = 0, 1, \ldots, (a,m) - 1.$$

Bizonyítás. (1) $ax \equiv b \pmod{m}$ megoldható

$$\begin{aligned} &\Leftrightarrow \exists x : ax \equiv b \pmod{m} \\ &\Leftrightarrow \exists x : m \mid (ax - b) \\ &\Leftrightarrow \exists x, y : m \cdot y = ax - b \\ &\Leftrightarrow \exists x, y : b = ax - m \cdot y \\ &\Leftrightarrow ax - m \cdot y = b \quad (x, y \text{ megoldása az előbbi lineáris diofantoszi egyenletnek}) \\ &\Leftrightarrow (a,m) \mid b. \end{aligned}$$

(2) Egyrészt, ha x_0 megoldás, akkor $\forall t$:

$$a\left(x_0 + t \cdot \frac{m}{(a,m)}\right) = ax_0 + a \cdot t \cdot \frac{m}{(a,m)} \equiv b \pmod{m}$$

mert x_0 megoldás $\Rightarrow x_0 + t \cdot \frac{m}{(a,m)}$ is megoldás.

Másrészt TFH x megoldás. Belátjuk: $\exists t \in \mathbb{Z} : x = x_0 + t \cdot \frac{m}{(a,m)}$. Mivel x és x_0 megoldás, ezért $ax \equiv b \pmod{m}$ és $ax_0 \equiv b \pmod{m}$, tehát

$$ax \equiv ax_0 \pmod{m}.$$

Ebből a kongruenciák egyszerűsítési szabálya (3.26 Tétel) miatt: $x \equiv x_0 \pmod{\frac{m}{(a,m)}}$, azaz $\exists t : x - x_0 = t \cdot \frac{m}{(a,m)}$ vagyis $x = x_0 + t \cdot \frac{m}{(a,m)}$. □

Kínai maradéktétel

3.28. Tétel (Kínai maradéktétel)**.** *Ha $m_1, m_2, \ldots, m_r \in \mathbb{N}$ páronként relatív prímek és $a_1, \ldots, a_r \in \mathbb{Z}$, akkor $\exists! x \in \mathbb{Z}$ modulo $M = m_1 m_2 \cdots m_r$:*

$$x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2}, \quad \ldots, \quad x \equiv a_r \pmod{m_r}.$$

Bizonyítás. Legyen minden k-ra $M_k = \frac{M}{m_k}$. Tetszőleges k-ra, $M_k \equiv 1 \pmod{m_k}$ megoldható, mert $(M_k, m_k) = 1$. Legyen x_k olyan, hogy $M_k x_k \equiv 1 \pmod{m_k}$.

Válasszuk X-t így:

$$X = a_1 M_1 x_1 + a_2 M_2 x_2 + \cdots + a_r M_r x_r.$$

Ekkor tetszőleges k-ra

$$X \equiv a_k M_k x_k \equiv a_k \cdot 1 \equiv a_k \pmod{m_k}$$

mert minden $j \neq k$-ra $M_j \equiv 0 \pmod{m_k}$ és x_k választása miatt $M_k x_k \equiv 1 \pmod{m_k}$. □

3.29. Megjegyzés. *Ha $m_1, \ldots, m_r$ nem relatív prímek, akkor az*

$$x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2}, \quad \ldots, \quad x \equiv a_r \pmod{m_r}$$

kongruencia-rendszer akkor és csak akkor oldható meg, ha bármely két tagjának van közös megoldása.

3.6. Felbonthatatlan elemek, prím elemek és a számelmélet alaptétele

3.30. Definíció. *Legyen R gyűrű, $t \in R$.*

- *t **felbonthatatlan**, ha $t \neq 0$, t nem egység, és*

$$\forall a, b \in R : t = ab \Rightarrow (a \text{ egység vagy } b \text{ egység} \quad);$$

- *t **prím**, ha $t \neq 0$, t nem egység, és $\forall a, b \in R : t \mid ab \Rightarrow t \mid (a \text{ vagy } t \mid b)$.*

3.31. Példa. *Legyen R páros egészek szokásos gyűrűje. Ebben 6 felbonthatatlan, de nem prím a következők miatt. R-ben 6 felbonthatatlan, mert nem bontható fel 2 páros szám szorzatára. Ugyanakkor R-ben 6 nem prím, mert $6 \mid 2 \cdot 18$, de R-ben $6 \nmid 2$ és $6 \nmid 18$, mert $\frac{2}{6}$ és $\frac{18}{6}$ nem R-beli elemek.*

3.32. Tétel. *$\mathbb{Z}$ tetszőleges eleme akkor és csak akkor felbonthtatlan, ha prím.*

Bizonyítás. $\Rightarrow$ Először TFH $u \in \mathbb{Z}$ felbonthatatlan (kell: u prím). TFH: $u \mid ab$. Mivel $(u, a) \mid u$, ezért $\exists q \in \mathbb{Z} : q(u, a) = u$. De u felbonthatatlan, ezért $(u, a) = 1$ vagy $q = \pm 1$ (a 2.7 Tétel miatt $\mathbb{Z}$-ben ezek az egységek).
1. eset: $(u, a) = 1$. Ekkor u és a relatív prímek, ezért a relatív prímek szabálya (3.16 Tétel) miatt $u \mid b$.
2. eset: $q = \pm 1$. Ekkor $(u, a) = \pm u$, speciálisan u közös osztója u-nak és a-nak, tehát $u \mid a$.

$\Leftarrow$ TFH u prím, azt szeretnénk megmutatni, hogy u felbonthatatlan. TFH: $u = ab$ (kell: a egység, vagy b egység). Mivel u prím, ezért $u \mid a$ vagy $u \mid b$; szimmetria miatt feltehető, hogy $u \mid a$. Mivel $u = ab$, ezért azt kaptuk, hogy $ab \mid a$, azaz $b = \pm 1$ egység $\mathbb{Z}$-ben. Tehát u valóban felbonthatatlan. □

3.33. Megjegyzés. *Ha u prím és $u \mid a_1 a_2 \cdots a_n$, akkor $\exists i : u \mid a_i$.*

3.34. Tétel (Számelmélet alaptétele $\mathbb{Z}$-re)**.** *Ha $n \in \mathbb{Z} \setminus \{0, \pm 1\}$, akkor:*

1. *Vannak $p_1, \ldots, p_m$ (nem feltétlenül különböző) prímek, hogy $n = p_1 \cdots p_m$;*
2. *ez a felbontás egységszorzóktól, sorrendtől eltekintve egyértelmű;*
3. *az előző felbontásban az azonos prímtényezők szorzatát hatványalakba írva*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

ezt nevezzük a felbontás kanonikus alakjának.

Bizonyítás. Nyilvánvaló, hogy a biznyítást elég pozitív n-ekre elvégezni (-1-el szorozva ebből az eredmény negatív n-ekre is adódik). Az $n \geq 2$ feltétel mellett teljes indukciót alkalmazunk n-re és (1)-et és (2)-t együtt igazoljuk.
Alaplépés: $n = 2$. Ekkor $m = 1$, $p_1 = n$ megfelelő felbontás.
Indukciós lépés: TFH: (1),(2) igaz $\forall u < n$-re. Ha n felbonthatatlan, akkor mint előbb, $m = 1$, $p_1 = n$ megfelelő felbontás. Ha n nem felbonthatatlan, akkor $\exists a, b \in \mathbb{Z} - \{0, \pm 1\} : n = a \cdot b$. Feltehető, hogy $a, b > 0$, ekkor nyilván $a, b < n$, ezért az indukció miatt $a = p_1 \cdots p_m$, $b = q_1 \cdots q_k$, így tehát $n = p_1 \cdots p_m \cdot q_1 \cdots q_k$ és (1) teljesül n-re.

(2) is teljesül, mert TFH: $n = p_1 \cdots p_m = q_1 \cdots q_k$. $|n| = 2$ mintájára, ha az u felbonthatatlan, akkor $m = k = 1$, $p_1 = q_1$. Ha n nem felbonthatatlan (azaz $m \geq 2$), akkor $p_1 \mid n = q_1 \cdots q_k$. Mivel p_1 prím, ezért $\exists j : p_1 \mid q_j$. De q_j is prím, így a 3.32 Tétel miatt q_j felbonthatatlan, ezért $p_1 = \pm q_j$. Ezért $p_2 \cdots p_m = q_1 \cdots q_{j-1}(\pm 1) q_{j+1} \cdots q_k$. Az indukciós feltevésünk szerint viszont ez a felbontás sorrendtől és egységszorzóktól eltekintve egyértelmű. Ezért ugyanez n felbontásaira is fennáll.

Végül 3. triviális. □

3.35. Megjegyzés. *Pozitív egészek azonosíthatók a véges multihalmazokkal: pl. $n = p_1^{e_1} \cdots p_r^{e_r}$, n az a multihalmaz, melynek p_1 e_1-szeres eleme, p_2 e_2-szeres eleme,... stb.*

3.36. Következmény.

1. *Ha $(n, a) = 1$, $(m, a) = 1$, akkor $(nm, a) = 1$;*
2. *ha $a \equiv b \pmod{m}$, akkor $K(a, m) = K(b, m)$ (emiatt: $(a, m) = (b, m)$).*

Bizonyítás. 1. A számelmélet alaptétele (3.34 Tétel) és a véges multihalmazokra vonatkozó megjegyzés miatt világos. 2.-höz TFH $a \equiv b \pmod{m}$. Ekkor $m \mid a - b$ azaz $\exists q : m \cdot q = a - b$. Ebből

(i) $a = mq + b$ és
(ii) $b = a - mq$.

Ezért, ha $c \in K(m, b)$, akkor $c \mid m$ és $c \mid b$ így (i) miatt $c \mid a$, tehát $c \in K(m, a)$, ezért $K(m, b) \subseteq K(m, a)$. A fordított irányú, $K(m, a) \subseteq K(m, b)$ tartalmazás (ii) segítségével hasonlóan ellenőrízhető. □

3.7. Maradékrendszerek, Euler-féle φ függvény és az Euler-Fermat tételkör

3.37. Definíció (Euler-féle φ függvény)**.** *Ha $n \in \mathbb{N} \setminus \{0\}$, akkor*

$$\varphi(n) = |\{k \in \mathbb{N} \mid 1 \leq k \leq n, (k, n) = 1\}|,$$

vagyis $\varphi(n)$ a nemnegatív, n-nél nem nagyobb, n-hez relatív prímek száma.

3.38. Definíció (Teljes maradékrendszer)**.** $T \subseteq \mathbb{N}$ ***teljes maradékrendszer*** *(röviden TMR) modulo m, ha*

- $|T| = m$ *és*
- $\forall t, t' \in T, t \neq t' : t \not\equiv t' \pmod{m}$.

3.39. Definíció (Redukált maradékrendszer)**.** $R \subseteq \mathbb{N}$ ***redukált maradékrendszer*** *(röviden RMR) modulo m, ha*

- $\forall x \in R : (x, m) = 1$*;*
- $|R| = \varphi(m)$ *és*
- $\forall t, t' \in R, t \neq t' : t \not\equiv t' \pmod{m}$.

Könnyű meggondolni, hogy ha $R \subseteq \mathbb{N}$ egy RMR modulo m, akkor minden, m-hez relatív prím $y \in \mathbb{N}$-hoz $\exists! x \in R : x \equiv y \pmod{m}$: x unicitása abból adódik, hogy R elemei páronként nem kongruensek modulo m, x egzisztenciája pedig a 3.36 Következményből látható be.

3.40. Tétel. *TFH $(a, m) = 1$, b tetszőleges.*

1. *Ha $\{t_1, \ldots, t_m\}$ TMR, akkor $\{at_1 + b, \ldots, at_m + b\}$ is TMR;*
2. *ha $\{r_1, \ldots, r_{\varphi(m)}\}$ RMR, akkor $\{ar_1, \ldots, ar_{\varphi(m)}\}$ is RMR.*

Bizonyítás. 1. igazolásához a 3.38 Definíció két pontját ellenőrizzük. Az világos, hogy az $\{at_1 + b, \ldots, at_m + b\}$ halmazban m darab elem van (azaz a felsorolt elemek páronként különbözők). Továbbá páronként nem-kongruensek egymással, mert THF $at_i + b \equiv at_j + b \pmod{m}$. Ekkor $at_i \equiv at_j \pmod{m}$. Mivel $(a, m) = 1$, ezért a kongruenciák egyszerűsítési szabálya (3.26 Tétel) miatt $t_i \equiv t_j \pmod{m}$. De $\{t_1, \ldots, t_m\}$ TMR, ezért $i = j$.

Hasonlóan, 2. igazolásához a 3.39 Definíció pontjait ellenőrizzük. A

$$\{ar_1, \ldots, ar_{\varphi(m)}\}$$

halmazban valóban $\varphi(m)$ darab elem van. TFH $ar_i \equiv ar_j \pmod{m}$. Ekkor, mint az 1. pont bizonyításában, $r_i \equiv r_j \pmod{m}$. De $\{r_1, \ldots, r_{\varphi(m)}\}$ RMR, ezért $i = j$. Kell még: $\forall r \in R : (ar, m) = 1$. De ez teljesül: mivel $(a, m) = 1$ és $(r, m) = 1$, ezért a 3.36 Következmény 1. pontja miatt valóban $(ar, m) = 1$. Ezzel azt kaptuk, hogy a 3.39 Definíció összes pontja valóban teljesül a $\{ar_1, \ldots, ar_{\varphi(m)}\}$ halmazra is. □

3.41. Tétel. *φ multiplikatív, azaz ha $(n, m) = 1$, akkor*

$$\varphi(n \cdot m) = \varphi(n) \cdot \varphi(m).$$

Bizonyítás. Tekintsük a következő $m \times n$ táblázatot:

$$\begin{array}{ccccc}
1 & 2 & 3 & \cdots & n \\
n+1 & n+2 & n+3 & \cdots & 2n \\
2n+1 & 2n+2 & 2n+3 & \cdots & 3n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(m-1)n+1 & (m-1)n+2 & (m-1)n+3 & \cdots & m \cdot n
\end{array}$$

Első sor: $\{1, 2, \ldots, n\}$, ez TMR modulo n ezért az első sorban $\varphi(n)$ db n-hez relatív prím szám van.

Oszlopok: az i-edik oszlop: $i, n+i, 2n+i, \ldots, (m-1)n+i$, ezeket az elemeket úgy kapjuk, hogy a $\{0, 1, ..., m-1\}$ halmaz elemeit n-el szorozzuk, majd hozzáadunk még i-t. Ezért a 3.40 Tétel miatt minden oszlop TMR modulo m. Ezért minden oszlopban $\varphi(m)$ db m-hez relatív prím szám van.

Relatív prímek: most két módon is leszámoljuk a táblázatban szereplő, $m \cdot n$-hez relatív prím számokat. Egyrészt, egy szám akkor és csak akkor relatív prím $n \cdot m$-hez, ha relatív prím n-hez és m-hez is. A fentiek szerint az első sor $\varphi(n)$ db n-hez relatív prím számot tartalmaz. Mivel egyazon oszlopon belül az elemek kongruensek modulo n, ezért $\varphi(n)$ db oszlop tartalmaz n-hez relatív prím számokat és a 3.36 Következmény 2. pontja miatt minden ilyen oszlop minden eleme relatív prím n-hez. Továbbá, fentebb átgondoltuk, hogy minden oszlopban $\varphi(m)$ db m-hez relatív prím szám van. Ezért összesen $\varphi(n) \cdot \varphi(m)$ db szám van a táblázatunkban, melyek relatív prímek $n \cdot m$-hez.

Másrészt definíció szerint: a táblázatban $\varphi(n \cdot m)$ db ilyen szám van. Ezek alapján valóban $\varphi(n) \cdot \varphi(m) = \varphi(n \cdot m)$. □

3.42. Tétel (Euler-féle φ függvény kiszámolása)**.** *Ha* $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$*, akkor*

$$\varphi(n) = (p_1 - 1)p^{e_1 - 1} \cdot ... \cdot (p_r - 1)p^{e_r - 1} = n \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right).$$

Bizonyítás. Először az $r = 1$ esetet vizsgáljuk (vagyis azt az esetet, amikor $n = p_1^{e_1}$ prímhatvány). Ekkor pontosan a p_1-el osztható számok nem lesznek relatív prímek n-hez, és $\frac{n}{p_1} = p_1^{e_1 - 1}$ darab n-nél nem nagyobb, p-vel osztható szám van. Ezért

$$\varphi(n) = p_1^{e_1} - p_1^{e_1 - 1} = p_1^{e_1 - 1}(p_1 - 1).$$

Az általános esetet (amikor n nem feltétlenül 1 darab, hanem r darab - páronként különböző alapú - prímhatvány szorzata) φ multipikativitásával (azaz a 3.41 Tétellel) intézzük el:

$$\varphi(n) = \varphi(p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}) = \varphi(p_1^{e_1})\varphi(p_2^{e_2}) \cdots \varphi(p_r^{e_r}) = (p_1 - 1)p_1^{e_1 - 1} \cdot ... \cdot (p_r - 1)p_r^{e_r - 1}.$$

Végül

$$\varphi(n) = (p_1 - 1)p_1^{e_1 - 1} \cdot ... \cdot (p_r - 1)p_r^{e_r - 1} = (p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}) \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right)$$

$$= n \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right).$$

□

3.43. Tétel (Euler-Fermat)**.** *Ha* $(a, m) = 1$*, akkor* $a^{\varphi(m)} \equiv 1 \pmod{m}$*.*

Bizonyítás. Tekintsük a $\{t_1, t_2, \ldots, t_{\varphi(m)}\}$ redukált maradékrendszert modulo m. A 3.40 Tétel miatt $\{at_1, at_2, \ldots, at_{\varphi(m)}\}$ is redukált maradékrendszer modulo m. Ezért minden 1 és $\varphi(m)$ közé eső i-hez pontosan 1 darab (1 és $\varphi(m)$ közés eső) j van, melyre $at_i \equiv t_j \pmod{m}$. Emiatt

$$at_1 \cdot at_2 \cdots at_{\varphi(m)} \equiv t_1 \cdot t_2 \cdots t_{\varphi(m)} \pmod{m}$$

azaz

$$a^{\varphi(m)}(t_1 t_2 \cdots t_{\varphi(m)}) \equiv t_1 t_2 \cdots t_{\varphi(m)} \pmod{m}.$$

Mivel $\forall i : (t_i, m) = 1$, azaz $t_1 t_2 \cdots t_{\varphi(m)}$ relatív prím m-hez, ezért (a 3.26 Tétel szerint) egyszerűsíthetünk vele:

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

□

3.44. Tétel (Kis Fermat-tétel)**.** *Legyen* p *prímszám.*

(1) Ha $p \nmid a$*, akkor* $a^{p-1} \equiv 1 \pmod{p}$*.*

(2) Tetszőleges a *egész számra* $a^p \equiv a \pmod{p}$*.*

Bizonyítás. (1) Ha p prím és $p \nmid a$, akkor $(a, p) = 1$ és $\varphi(p) = p - 1$, így a 3.43 Euler-Fermat tétel miatt:

$$a^{p-1} \equiv 1 \pmod{p}.$$

(2)-t esetszétválasztással bizonyítjuk.

1. eset: $p \nmid a$.

Ekkor (1) alapján $a^{p-1} \equiv 1 \pmod{p}$. Mindkét oldalt a-val szorozva:

$$a^p \equiv a \pmod{p}.$$

2. eset: $p \mid a$

Ekkor $a \equiv 0 \pmod{p}$, tehát:

$$a^p \equiv 0 \pmod{p} \quad \text{és} \quad a \equiv 0 \pmod{p}$$

így $a^p \equiv a \pmod{p}$. □

4. Speciális algebrai struktúrák (félcsoportok, csoportok, ... stb.)

4.1. Definíció (Félcsoport)**.** *Az* $(A, \cdot)$ *struktúra* ***félcsoport****, ha a szorzás asszociatív:*

$$\forall x, y, z \in A : x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

4.2. Definíció (Monoid)**.** *Az* $(A, \cdot, 1)$ *struktúra* ***monoid****, ha:*

- $(A, \cdot)$ *félcsoport, és*
- $\forall x \in A : 1 \cdot x = x \cdot 1 = x$ *(azaz létezik egységelem).*

4.3. Definíció (Csoport)**.** *Az* $(A, \cdot, 1)$ *struktúra* ***csoport****, ha:*

- $(A, \cdot, 1)$ *monoid, és*
- $\forall x \in A \; \exists x^{-1} \in A : x \cdot x^{-1} = x^{-1} \cdot x = 1$ *(azaz minden elemnek létezik inverze).*

4.4. Definíció (Abel-csoport)**.** *Egy csoport* ***Abel-csoport****, ha:*

$$\forall x, y \in A : x \cdot y = y \cdot x$$

(azaz a szorzás művelete kommutatív).

4.5. Definíció (Test)**.** *Az* $(A, +, \cdot, 0, 1)$ *struktúra* ***test****, ha:*

- $(A, +, 0)$ *Abel-csoport,*
- $(A \setminus \{0\}, \cdot, 1)$ *csoport és*
- *A szorzás disztributív az összeadásra nézve.*

4.6. Példa. *Példák:*

- $(\mathbb{Z}, +, 0)$*: Abel-csoport,*
- $(\mathbb{N}, \cdot, 1)$*: kommutatív monoid,*
- $(\mathbb{Z}, +, \cdot, 0, 1)$*: gyűrű, de nem test,*
- $(\mathbb{Q}, +, \cdot, 0, 1)$*: test.*

Függvények tulajdonságai

4.7. Definíció. *Legyen* $f : A \to B$ *függvény.*

- *Az* f *függvény* ***injektív****, ha:*

$$\forall x, y \in A : x \neq y \Rightarrow f(x) \neq f(y).$$

- *Az* f *függvény* ***szürjektív****, ha:*

$$\forall y \in B \; \exists x \in A : f(x) = y.$$

- *Az* f *függvény* ***bijektív****, ha injektív és szürjektív (kölcsönösen egyértelmű leképezés).*

Struktúramegőrző leképezések

A következő definíciókban azt szeretnénk precízzé tenni, hogy két struktúra ugyanolyan (lényegében azonos), vagy legalább hasonlít egymásra abból a szempontból, hogy elemeik olyan módon is megfeleltethetők egymásnak, hogy a megfeleltetés a műveletekkel kompatibilis legyen. Ilyen esetekben teljesül ugyanis, hogy a két srtruktúrában „ugyanúgy", vagy legalább „nagyon hasonlóan" kell számolni.

4.8. Definíció (Izomorfizmus)**.** *Legyenek $\mathcal{A} = (A, f_i)_{i\in I}$ és $\mathcal{B} = (B, g_i)_{i\in I}$ azonos típusú struktúrák.*
Egy $\alpha : A \to B$ leképezés ***izomorfizmus****, ha:*

1. *α bijekció, és*
2. *α művelettartó:*

$$\forall i \in I, \forall x_1, \dots, x_n \in A : \alpha(f_i(x_1, \dots, x_n)) = g_i(\alpha(x_1), \dots, \alpha(x_n)).$$

Egy $\alpha : A \to B$ leképezés ***homomorfizmus****, ha a művelettartás teljesül (de az 1. feltételt nem kötjük ki).*

4.9. Megjegyzés. Középiskolás tanulmányokból ismerős lehet a gráfok izomorfiájának fogalma: két gráf akkor és csak akkor izomorf, ha csúcsaik között meg lehet adni egy „éleket megőrző" bijekciót (amely tehát élt élbe, nem-élt pedig nem élbe visz). Az izomorf gráfokat pedig azonosnak (vagy legalább lényegében azonosnak) tekinthetjük. A 4.8 Definíció mögött nagyon hasonló gondolat húzódik meg: két algebrai struktúra izomorf (lényegében azonos), ha elemeik között van művelettartó bijekció. A 4.8 Definíció tehát a gráfok izomorfizmus-fogalmának algebrai struktúrákra vonatkozó természetes adaptációja.

4.1. Struktúrák és azonosságok

4.10. Definíció (Változók és termek)**.** *Legyen $V \neq \emptyset$ egy változóhalmaz (általában $V = \{x, y, z, \dots\}$, de lehet véges is).*
A ***termek*** *halmazát a következőképpen definiáljuk:*
(1) Ha x változó vagy kitüntetett konstans, akkor x term;
(2) Ha $t_1, \dots, t_n$ termek és f n-változós függvény, akkor $f(t_1, \dots, t_n)$ is term;
(3) Minden term az előző két szabály véges sokszori alkalmazásával áll elő.

4.11. Példa. *Példák:*

- *$x(y + z)$ egy gyűrű-term;*
- *$x((y+)$ nem term;*
- *$x(y + z)$ nem csoport-term (csoportban 1 db kétváltozós művelet van).*

4.12. Definíció (Azonosság)**.** *A $t_1 = t_2$ szimbólumsorozatot* ***azonosságnak*** *nevezzük, ha t_1, t_2 termek.*

A 2.2 Definícióban szereplő gyűrűaxiómák, csakúgy mint a 4.1-4.4 Definíciókban szereplő félcsoport-, csopartaxiómák,... stb. mind mind azonosságok. De pl. testekben a multiplikatív inverzre vonatkozó feltételek nem azonosságok (mert a nullelemnek nincs reciproka, ezért a multiplikatív inverz képzése nem olyan művelet, amely minden test-elemen értelmezve van).

4.13. Tétel (Azonosságok megőrzése)**.** *Legyenek $\mathcal{A} = (A, f_i)_{i\in I}$ és $\mathcal{B} = (B, g_i)_{i\in I}$ azonos típusú struktúrák. Tegyük fel, hogy $\alpha : A \to B$ szürjektív homomorfizmus.*

(1) Tetszőleges $t(x_1, \dots, x_n)$ termre és $a_1, \dots, a_n \in A$ esetén:

$$\alpha(t^{\mathcal{A}}(a_1, \dots, a_n)) = t^{\mathcal{B}}(\alpha(a_1), \dots, \alpha(a_n))$$

ahol pl. $t^{\mathcal{A}}(a_1, \dots, a_n)$ az $\mathcal{A}$ struktúrának azt az elemét jelöli, amelyet a t term ad eredményül, ha a benne szereplő műveleteket $\mathcal{A}$-ban elvégezzük az $a_1, \dots, a_n$ elemekre.

(2) Ha $t_1 = t_2$ azonosság igaz $\mathcal{A}$-ban, akkor $\mathcal{B}$-ben is igaz.

Bizonyítás. (1) A t termben előforduló függvényjelek m száma szerinti indukciót alkalmazunk.

Alaplépés ($m = 0$): t valamelyik változó vagy konstans. Ekkor (1) mindkét oldalán ugyanannak az A-beli elemnek az α szerinti képe áll, ezek nyilván egyenlők. (Kicsit részletesebben, pl. ha a t term az x változó, melybe $a \in A$-t helyettesítünk, akkor (1) baloldalát így értelmezhetjük: „$\mathcal{A}$-ban a-val további műveleteket nem végzünk, majd az eredménynek (vagyis a-nak) vesszük az α szerinti képét". Hasonlóan, ebben az esetben (1) jobboldala így értelmezhető: „vesszük az a elem α szerinti képét, és ezzel a képpel $\mathcal{B}$-ben további műveleteket már nem végzünk". Világos, hogy a baloldal és a jobboldal előző értelmezései szerint mindkét esetben $\alpha(a)$-t kapunk eredményül, tehát az Alaplépés esetében (1) baloldala és jobboldala valóban egyenlő.)

Indukciós lépés: Tegyük fel, hogy $t = f(t_1, \dots, t_k)$ és (1)-et tudjuk már a $t_1, \dots, t_k$ termekre. Ekkor

$$\begin{aligned}
\alpha(t^{\mathcal{A}}(a_1, \dots, a_n)) &= \alpha(f^{\mathcal{A}}(t_1^{\mathcal{A}}(\bar{a}), \dots, t_k^{\mathcal{A}}(\bar{a}))) \quad \text{(de } \alpha \text{ homomorfizmus)} \\
&= f^{\mathcal{B}}(\alpha(t_1^{\mathcal{A}}(\bar{a})), \dots, \alpha(t_k^{\mathcal{A}}(\bar{a}))) \quad \text{(ind. feltevés)} \\
&= f^{\mathcal{B}}(t_1^{\mathcal{B}}(\alpha(\bar{a})), \dots, t_k^{\mathcal{B}}(\alpha(\bar{a}))) \\
&= t^{\mathcal{B}}(\alpha(a_1), \dots, \alpha(a_n)).
\end{aligned}$$

(2) Tegyük fel, hogy $t_1 = t_2$ igaz $\mathcal{A}$-ban. Kell: igaz $\mathcal{B}$-ben is. Legyen $b_1, \dots, b_n \in B$ tetszőleges. Mivel α szürjektív, ezért vannak olyan $a_1, \dots, a_n \in A$ elemek, melyekre $\alpha(a_1) = b_1, \dots, \alpha(a_n) = b_n$.

Ekkor:

$$\begin{aligned} t_1^{\mathcal{B}}(b_1,\ldots,b_n) &= t_1^{\mathcal{B}}(\alpha(a_1),\ldots,\alpha(a_n)) \qquad ((1) \text{ alapján}) \\ &= \alpha(t_1^{\mathcal{A}}(a_1,\ldots,a_n)) \qquad (t_1 = t_2 \text{ igaz } \mathcal{A}\text{-ban}) \\ &= \alpha(t_2^{\mathcal{A}}(a_1,\ldots,a_n)) \qquad (\text{megint } (1) \text{ alapján}) \\ &= t_2^{\mathcal{B}}(\alpha(a_1),\ldots,\alpha(a_n)) = t_2^{\mathcal{B}}(b_1,\ldots,b_n). \quad \checkmark \end{aligned}$$

□

4.14. Következmény. *Nincs olyan gyűrűazonosság, amely pontosan a nullosztómentes gyűrűkben igaz.*

Bizonyítás. Ellentmondást keresve tegyük fel, hogy van olyan $t_1 = t_2$ gyűrűazonosság, amely pontosan a nullosztómentes gyűrűkben igaz. Tekintsük a következő struktúrákat:

- $\mathbb{Z}$: nullosztómentes gyűrű;
- $\mathbb{Z}_4$: nem nullosztómentes gyűrű.

Legyen $\alpha : \mathbb{Z} \to \mathbb{Z}_4$, $f(n) = n \pmod 4$.
Ez szürjektív homomorfizmus. Mivel $\mathbb{Z}$ nullosztómentes, ezért $t_1 = t_2$ igaz $\mathbb{Z}$-ben. Ezért az előző 4.13 Tétel miatt $t_1 = t_2$ igaz $\mathbb{Z}_4$-ben is. De $\mathbb{Z}_4$-ben van nullosztó, ellentmondás. □

5. Komplex számok

5.1. A komplex számok bevezetésének motivációja

Leopold Kronecker (1823-1891) korának neves német matematikusa szerint

„A természetes számokat Isten teremtette, minden más az ember műve."

Valóban, egészen ősi kultúrák ismerték, és használták a természetes számokat. Ugyanakkor, ahogy azt a 2. fejezet elején láttuk, meglepően nehéz kérdés annak tisztázása, hogy pontosan mik is a természetes számok. Ezért, ezen a kurzuson a természetes számok fogalmát ismertnek tekintettük és tekintjük (illetve további vizsgálatukat más kurzusokra halasztjuk), de kiemeljük a következőket.

• Egyrészt a természetes számok azonosíthatók egy (egyirányban végtelen) félegyenes pontjaival a szokásos módon (úgy, hogy szomszédos természetes számoknak egységnyi távolságú pontok felelnek meg).
• Másrészt, a természetes számok körében pl. az

$$(1) \qquad x + 3 = 1$$

egyenlet nem oldható meg: akármelyik természetes számot adjuk is értékül x-nek, a baloldal eredménye sosem lesz 3-nál kevesebb, ezért x értékét nem tudjuk úgy választani, hogy a jobboldalon szereplő 1-et kapjuk eredményül.

Az (1)-hez hasonló típusú egyenletek megoldhatósága érdekében érdemes bevezetni az egész számokat: ezek között lehetnek negatívak is. Negatív számokat is megengedve könnyen megtalálhatjuk az (1) egyenlet megoldását. Ugyanakkor, a negatív egész számok fogalma már nem annyira szemléletes: pl. egy buszon akkor utazik -2 darab utas, ha 2 utasnak fel kell szállni ahhoz, hogy a busz üres legyen. Ezt ebben a formában nehéz elképzelni (de pl. az egészen szemléletes, hogy negatív számokat is megengedve egyszerűen tudjuk modellezni a „profit és költség" vagy „tőke és adósság" ellentétes fogalompárjait). Az egész számok halmaza egészen konkrétan megkonstruálható a természetes számok halmazából, e konstrukcióval kapcsolatban a 2025 őszi kurzus 4. adag házi feladatai közül a 7. feladatra utalunk. E konstrukció során meg kellett adni, hogy

1. Mik az egész számok;
2. Mikor tekintünk két egész számot egyenlőnek;
3. Hogyan kell a szokásos műveleteket elvégezni az egész számokon.

Az 1. és 2. pontokkal kapcsolatban (a részletek felidézése nélkül) azt jegyezzük meg, hogy az egész számok azonosíthatók a természetes számok rendezett párjainak bizonyos ekvivalenciaosztályaival. A konstrukció végén azt mondhatjuk, hogy ha

tudnánk, mik a természetes számok, akkor tudnánk azt is, hogy mik az egész számok. Ismét kiemeljük a következő aspektusokat.

• Mindenek előtt a természetes számok összeadásra és szorzásra vett félcsoportjai természetes módon beágyazhatók az egész számok összeadásra és szorzásra vett félcsoportjaiba (azaz van olyan injektív függvény, mely a természetes számokat képezi az egész számok halmazába, és megtartja az összeadás és szorzás műveletét, tehát durván fogalmazva van injektív homomorfizmus $\mathbb{N}$-ből $\mathbb{Z}$-be).
• Az egész számok azonosíthatók egy (mindkét irányban végtelen) egyenes pontjaival a szokásos módon (úgy, hogy szomszédos egész számoknak egységnyi távolságú pontok felelnek meg).
• Végül, az egész számok körében pl. a

$$2x + 4 = 1 \tag{2}$$

egyenlet nem oldható meg: akármelyik egész számot adjuk is értékül x-nek, a baloldal eredménye páros lesz, ezért x értékét nem tudjuk úgy választani, hogy eredményül a jobboldalon szereplő páratlan 1-et kapjuk.

Az előző bekezdésben szereplő (2)-höz hasonló egyenletek megoldhatósága érdekében ismét érdemes újfajta számokat bevezetni, ezek lesznek a racionális számok, melyek akár tört értékűek is lehetnek. A racionális számok körében könnyen meg tudjuk oldani a (2)-höz hasonló egyenleteket, emellett a (pozitív) racionális számok könnyen el is képzelhetők: pl. egy almát 3 egyenlő részre felszeletelve viszonylag pontos (bár a szeletelés óhatatlan pontatlansága miatt nem teljesen pontos) módon (fizikailag, érzékszervi tapasztalással is) érzékelhetünk $\frac{1}{3}$ darab almát. Amint azt a 2025 őszi kurzus 4. heti gyakorlatra szánt 8. feladatában vázoltuk, az egész számok esetéhez hasonlóan, a racionális számok teste is egészen konkrétan megkonstruálható az egész számok gyűrűjéből. Ismét azt kell tisztázni, hogy

1. Mik a racionális számok;
2. Mikor tekintünk két racionális számot egyenlőnek;
3. Hogyan kell a szokásos műveleteket elvégezni racionális számokon.

Ahogy azt általános iskolában megtanultuk, a racionális számok egész számokból álló (számlálóra és nevezőre tagolt) rendezett párokból képzett ekvivalenciaosztályok (két egész számokból álló pár ekvivalens, ha közös nevezőre bővítve őket, a számlálóik, mint egész számok egyenlőkké válnak). A racionális számok konstrukciója végén ismét azt mondhatjuk, hogy ha tudnánk, mik az egész számok, akkor tudnánk azt is, hogy mik a racionális számok. Ismét kiemeljük a következőket.

• Van olyan injektív függvény, mely az egész számokat képezi a racionális számok halmazába, és megtartja az összeadás és szorzás műveletét (azaz, a részleteket elnagyolva: van injektív homomorfizmus $\mathbb{Z}$-ből $\mathbb{Q}$-ba).

• A racionális számok azonosíthatók egy (mindkét irányban végtelen) egyenes bizonyos pontjaival a szokásos módon, és egy ilyen azonosítás után a racionális pontok halmaza sűrű lesz a számegyenesen: bármely két különböző racionális szám között van további (valójában van végtelen sok további) racionális szám.

• A valós számok bevezetésének nem algebrai okai vannak. Noha a racionális számok a számegyenes sűrű részhalmazát adják, még mindig nem kaptuk meg a számegyenes összes pontját. A valós számok bevezetésének oka és célja az, hogy kitöltsük a számegyenesen azokat a hézagokat, amelyek a racionális számok bevezetése után még mindig megmaradtak. A racionális számok egy másfajta bővítését indokolná, ha pl. a

$$x^2 = -1 \qquad (3)$$

egyenletet szeretnénk megoldani. A racionális számok halmazának ilyen bővítéseivel kapcsolatban az alábbi 2 lehetőséget említjük meg:

Algebrai számok: racionális együtthatós polinomok gyökei,

$$\text{pl.: } 5 + 2\sqrt{2}, \qquad \sqrt[5]{\sqrt{3} + \sqrt{2}} + 1.$$

Gyökkifejezések: termek a $(\mathbb{Q}, +, \cdot, 0, 1, \sqrt[n]{\ })_{n \in \mathbb{N}^+}$ struktúrában.

A 2.4 Tétel szerint $\mathbb{Q}$ valódi részhalmaza a gyökkifejezéseknek és meg lehet mutatni, hogy a gyökkifejezések valódi részhalmazát alkotják az algebrai számoknak (és mellesleg: az algebrai számok nem is feltétlenül valósak). Az algebrai számok fogalmára és a gyökkifejezésekre alább még vissza fogunk térni.

A (3)-hoz hasonló egyenletek még a valós számok körében sem oldhatók meg: minden valós szám négyzete nemnegatív, ezért nem tudunk olyan valós értéket adni x-nek, hogy (3) baloldala egyenlő legyen a negatív jobboldalával. A (3)-hoz hasonló egyenletek megoldása (akár a számfogalom bővítése árán) egészen a reneszánsz koráig komoly fejtörést okozott a matematikusoknak, és további évszázadokra volt szükség a kérdéskör pontos tisztázásához. Ebben a fejezetben alább a komplex számokat és alaptulajdonságaikat ismertetjük. Az előző bekezdésekben ismétlődő mintázat szerint kibővítjük a számfogalmat a valós számokon túlra, mégpedig azért, hogy a (3)-hoz hasonló egyenletek megoldhatók legyenek. Az újfajta számokat komplex számoknak fogjuk nevezni. A komplex számokat is valós számpárokkal azonosíthatjuk, és egészen pontosan meg fogjuk majd adni, hogy az újfajta (komplex) számokon hogyan kell a szokásos műveleteket elvégeznünk. Konstrukciónk ismét olyan lesz, hogy egy alkalmas injektív homomorfizmus a valós számok $\mathbb{R}$ testét beágyazza a komplex számok $\mathbb{C}$ testébe. Ismét elmondhatjuk majd, hogy ha tudnánk, hogy mik azok a valós számok, akkor azt is tudnánk, hogy a komplex számok micsodák. Emellett - mivel a valós számok már teljes mértékben kitöltik a számegyenest - a komplex számokat nem egy egyenes, hanem egy sík pontjaival tudjuk azonosítani.

Felmerül a kérdés, hogy esetleg még bonyolultabb egyenletek megoldhatóságának

érdekében további számfogalom-bővítésre lenne-e szükség, és ha ez így van, akkor a számfogalom-bővítéseinknek végére érünk-e valaha. Ezzel kapcsolatban jó hírünk van, ugyanis igaz a következő:

5.1. Tétel (Az Algebra Alaptétele)**.** [1] *Minden komplex-együtthatós, legalább elsőfokú polinomnak van gyöke $\mathbb{C}$-ben.*

Bizonyítás. Bizonyítás év végén $\rightarrow$ sőt másik félévben. □

Ez a tétel a természet egy kisebbfajta csodája: azt állítja, hogy a polinomegyenletek megoldhatósága szempontjából megállhatunk a komplex számoknál még akkor is, ha polinomegyenleteinkben szerepelhetnek az újfajta komplex számaink is.

Végül megjegyezzük, hogy a számfogalom bővíthető a végtelen nagy (vagy akár a végtelen pici) mennyiségek irányába is; a végtelenség más-más aspektusait megragadva a végtelen nagy mennyiségekkel (számosságokkal, rendszámokkal) a halmazelmélet, a végtelen pici mennyiségekkel a (nemsztenderd) analízis foglalkozik. Ezek (akár több) önálló tantárgyat kitöltő területek; ebben a kurzusban a végtelen mennyiségeken végzett műveletekkel és azok tulajdonságaival nem foglalkozunk.

5.2. Komplex számok definíciója

5.2. Definíció (Komplex számok halmaza és egyenlősége)**.** $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ *és*

$$(a, b) = (c, d) \Leftrightarrow (a = c) \wedge (b = d).$$

Jelölés: az $(a, b) \in \mathbb{C}$ komplex számot $a + bi$ alakban is írjuk, és a kérdéses komplex szám algebrai alakjának nevezzük.

Az előző Definíció jelöléseit megtartva hangsúlyozzuk a következőket. A fenti, szigorú algebrai alak mellett $0 + bi$ helyett írhatunk egyszerűen bi-t, és hasonlóan, $a + 0i$ helyett írhatunk egyszerűen a-t is.

Egy komplex szám tehát egy valós számpár: a $z = a + ib$ komplex számnak a a valós, és b a képzetes része, ezekre alább az 5.5 Definícióban vezetünk be jelöléseket. A racionális számok esetében a törtvonal különíti el a számlálót a nevezőtől. Ehhez némileg hasonlóan, az algebrai alakban az i szorzótényező jelöli meg a komplex számunk képzetes részét. A következő definícióban megadjuk, hogyan kell összeadni és összeszorozni két komplex számot. Látjuk majd, hogy az algebrai alak jelölései kényelmesek lesznek az egyszerűbb műveletek elvégzésére. Lesz egy új számolási szabályunk is: $i^2 = -1$. Ezek szerint az új i szám megoldása lesz a valósak körében megoldhatatlan $x^2 = -1$ egyenletnek; emiatt néha az $i = \sqrt{-1}$ jelölést is alkalmazzák (ami egy picit pontatlan: egyrészt egyelőre nem definiáltuk negatív szám négyzetgyökét; másrészt, ha már definiáltuk is volna, egy ilyen jelöléshez akkor is be kéne látni, hogy -1-nek pontosan 1 darab négyzetgyöke van. Ezzel szemben az derül

[1] Ezt a tételt újabban a Klasszikus Algebra Alaptételének nevezik.

majd ki, hogy két olyan kompex szám is van, melyek négyzete -1: az egyik ilyen szám i a másik pedig $-i$ azaz szigorúan vett algebrai alakban $0+1i$ és $0+(-1)i$). Emiatt a 5.12 Definícióban további magyarázattal fogunk majd szolgálni, hogy pl. az $\sqrt{-1}$ jelölést hogyan alkalmazzuk.

5.3. Definíció (Műveletek)**.**

- ***Összeadás****:* $(a,b)+(c,d)=(a+c,b+d)$ *azaz*

$$a+bi+c+di=(a+c)+(b+d)i.$$

- ***Szorzás****:* $(a,b)\cdot(c,d)=(ac-bd,ad+bc)i$ *azaz*

$$(a+bi)(c+di)=ac+adi+bci+bdi^2=(ac-bd)+(ad+bc)i,$$

ahol tehát $i^2=-1$*.*

A műveletek előbbi definíciója alapján valóban praktikus az algebrai alak használata: komplex számokkal „úgy számolunk, mint a szokásos kéttagú összegekkel, azzal a szabállyal kiegészítve, hogy $i^2=-1$".

5.4. Tétel. $(\mathbb{C},+,\cdot,0,1)$ *test.*

Bizonyítás. A testaxiómák ellenőrzésével. A bizonyításra egy másik (rövidebb, de absztraktabb, és további előkészületi lépéseket igénylő) lehetőség az lehetne, hogy igazoljuk: $\mathbb{C}$ izomorf egy gyűrű (polinomgyűrű) speciális homomorf képével. Mivel a 4.13 Tétel szerint a gyűrűhomomorfizmusok megőrzik a gyűrűazonosságok igazságát is, ezért ebből azonnal adódna, hogy $\mathbb{C}$ egy gyűrű. Azt pedig külön (viszonylag röviden) meg lehetne gondolni, hogy a homomorfizmust olyan speciális módon is választhatjuk, hogy értékkészletében (azaz $\mathbb{C}$-ben) a nemnulla elemek mind invertálhatók legyenek. □

5.3. Komplex számok algebrai alakja és a műveletek alaptulajdonságai

5.5. Definíció (Algebrai alak)**.** *Ha* $z=(a,b)\in\mathbb{C}$*, akkor* $a+bi$ *a* z ***algebrai alakja****.*

5.6. Definíció (Valós és képzetes rész)**.** $z=a+bi$ *esetén:*

- $\Re(z)=a$*:* ***valós rész****;*
- $\Im(z)=b$*:* ***képzetes rész****;*
- $|z|=\sqrt{a^2+b^2}$*: a* z *komplex szám abszolút értéke (=* 0*-tól való távolsága);*
- $\overline{z}=a-bi$*:* z ***konjugáltja****.*

Ahogy korábban megjegyeztük, a számegyenes minden pontja megfelel egy valós számnak. Mivel egyetlen valós szám négyzete sem negatív, ezért az i számot nem tudjuk elhelyezni a számegyenesen, hanem azon kívül helyezzük el: felveszünk egy, a

számegyenesre merőleges, origón átmenő újabb egyenest – amit **képzetes tengelynek** nevezünk – és i-t a képzetes tengely egyik olyan pontjával azonosítjuk, melynek 1 a távolsága az origótól. A valós számok eredeti számegyenese és a képzetes tengely együtt megad egy derékszögű koordinátarendszert. Ennek segítségével az $a+ib$ komplex számot a sík (a, b) koordinátájú pontjával azonosítjuk, és az azonosítás után a komplex számok síkjáról (vagy számsíkról) beszélünk – pontosan ugyanúgy, ahogy a valós számok esetében használjuk a „számegyenes" kifejezést. A komplex számok összeadása egybeesik a nekik megfelelő helyvektorok, mint síkvektorok összeadásával. Alább, a 5.14 Megjegyzésben látjuk majd, hogy a komplex számok szorzásának is van egy nagyon világos geometriai jelentése, amely azonban nem teljesen nyilvánvaló; ismertetése további előkészületeket igényel.

5.7. Tétel (Konjugálás tulajdonságai)**.** *Legyen* $z \in \mathbb{C}$.

1. $\overline{\overline{z}} = z$;
2. $z \cdot \overline{z} = |z|^2 \in \mathbb{R}$.

Bizonyítás. (1) $\overline{\overline{a+bi}} = \overline{a-bi} = a+bi$ ✓.

(2) $z \cdot \overline{z} = (a+bi)(a-bi) = a^2 - abi + abi + b^2 = a^2 + b^2 = |z|^2$. ✓ □

5.8. Példa (Osztás algebrai alakban)**.**

$$\frac{5+2i}{1+i} = \frac{(5+2i)(1-i)}{(1+i)(1-i)} = \frac{5-5i+2i-2i^2}{2} = \frac{7-3i}{2} = \frac{7}{2} - \frac{3}{2}i.$$

5.9. Tétel. *A konjugálás automorfizmus. Részletesebben, az*

$$f : \mathbb{C} \to \mathbb{C},\ f(z) = \overline{z}$$

függvény egy gyűrűizomorfizmus $\mathbb{C}$*-ből* $\mathbb{C}$*-be, azaz:*

1. $f(z_1 + z_2) = f(z_1) + f(z_2)$;
2. $f(z_1 z_2) = f(z_1) f(z_2)$.

Bizonyítás. Legyen $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$. Ki fogjuk számítani 1. bal- és jobboldalát, és látjuk majd, hogy valóban egyenlőek.

$$f(z_1 + z_2) = f((a_1 + a_2) + (b_1 + b_2)i) = (a_1 + a_2) - (b_1 + b_2)i$$
$$f(z_1) + f(z_2) = (a_1 - b_1 i) + (a_2 - b_2 i) = (a_1 + a_2) - (b_1 + b_2)i. \quad ✓$$

2. Hasonlóan igazolható. □

5.4. Komplex számok trigonometrikus alakja

HIÁNYZIK A SíKBELI POLÁRKOORDINÁTA RENDSZER !!!!!!!!!!!!

5.10. Definíció (Trigonometrikus alak)**.** *Legyen $z = a + bi \in \mathbb{C}$, $z \neq 0$. Ekkor z felírható* ***trigonometrikus alakban****:*

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

ahol:

- $r = |z| = \sqrt{a^2 + b^2}$ *a komplex szám* ***abszolút értéke****;*
- φ *a komplex szám* ***argumentuma*** *(szöge), amelyre:*

$$\cos\varphi = \frac{a}{r}, \quad \sin\varphi = \frac{b}{r}$$

Az argumentum csak 2π egész számú többszörösei erejéig van meghatározva.

5.11. Tétel (Szorzás trigonometrikus alakban)**.** *Legyenek $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ és $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ komplex számok. Ekkor:*

$$z_1 z_2 = r_1 r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)].$$

Bizonyítás.

$$\begin{aligned} z_1 z_2 &= r_1 r_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) \\ &= r_1 r_2[(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)] \\ &= r_1 r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)]. \end{aligned}$$

□

5.5. Komplex gyökök és egységgyökök

5.12. Definíció (Komplex gyökök)**.** *Legyen $z \in \mathbb{C}$, $n \in \mathbb{N}$, $n \geq 1$. Ekkor:*

$$\sqrt[n]{z} = \{w \in \mathbb{C} \mid w^n = z\}.$$

Ezek szerint a z komplex szám n-edik gyöke egy halmaz: az $\sqrt[n]{z}$ halmaz azokat a komplex számokat tartalmazza, melyek n-edik hatványa éppen z.

5.13. Tétel (Osztás és gyökvonás trigonometrikus alakban)**.** *Legyenek*

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1), \qquad z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2) \neq 0$$

komplex számok és legyen $n \in \mathbb{N}$.

1. $\dfrac{z_1}{z_2} = \dfrac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$*;*

2. $z^n = r^n(\cos(n\varphi) + i\sin(n\varphi))$ *(Moivre-formula);*

3.
$$\sqrt[n]{z} = \left\{ \sqrt[n]{r}\left(\cos\left(\frac{\varphi+2k\pi}{n}\right) + i\sin\left(\frac{\varphi+2k\pi}{n}\right)\right) \mid k = 0,1,...,n-1 \right\}.$$

Bizonyítás. 1. Ellenőrizhető:
$$\frac{r_1}{r_2}(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2)) \cdot z_2 = z_1.$$

2. Triviális (indukció n-re).

3. Tetszőleges $k \in \mathbb{N}$ esetén:
$$\begin{aligned}&\left[\sqrt[n]{r}\left(\cos\left(\frac{\varphi+2k\pi}{n}\right) + i\sin\left(\frac{\varphi+2k\pi}{n}\right)\right)\right]^n \\ &= r(\cos(\varphi+2k\pi) + i\sin(\varphi+2k\pi)) \\ &= r(\cos\varphi + i\sin\varphi) = z.\end{aligned}$$

Ezért a jobboldali halmaz elemei mind benne vannak az $\sqrt[n]{z}$ halmazban. Ezek páronként különbözőek, mert a sin és cos függvények legkisebb közös periódusa 2π. Mivel 2π közös periódusa sin-nek és cos-nak, ezért nagyobb k-k nem adnak új gyököket: $k = 0,\ldots,n-1$. Végül, $\sqrt[n]{z}$-nek más elemei a már igazolt 2. miatt nem lehetnek.

□

5.14. Megjegyzés. A 5.13 Tételből látszik, hogy tetszőleges $z \in \mathbb{C}$-vel való szorzás: nyújtva forgatás (nyújtás $|z|$-szeresre + origó körüli forgatás z argumentumával). Speciálisan, ha a $z \in \mathbb{C}$ komplex szám abszolútértéke 1, akkor a z-vel való szorzás geometriai hatása éppen a z argumentumával való elforgatás (0 körül).

5.15. Példa. *Legyen $n \in \mathbb{N}^+$; ekkor az $\sqrt[n]{1}$ halmaz elemei a következők:*
$$\begin{aligned}\varepsilon_0 &= 1(\cos 0 + i\sin 0) = 1; \\ \varepsilon_1 &= 1\left(\cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)\right); \\ \varepsilon_2 &= 1\left(\cos\left(\frac{4\pi}{n}\right) + i\sin\left(\frac{4\pi}{n}\right)\right); \\ &\vdots \\ \varepsilon_{n-1} &= 1\left(\cos\left(\frac{2(n-1)\pi}{n}\right) + i\sin\left(\frac{2(n-1)\pi}{n}\right)\right).\end{aligned}$$

Megjegyezzük még, hogy az ε_1-el való szorzás $\frac{2\pi}{n}$-nel való elforgatás (0 körül).

A következő definícióban az előző példát általánosítjuk.

5.16. Definíció (Rend)**.** *Legyen $z \in \mathbb{C}$. A z **rendje** az a legkisebb $n \in \mathbb{N}$, $n \geq 1$ (ha van ilyen), melyre $z^n = 1$.*

- *Ha $\forall n \in \mathbb{N}$-re $z^n \neq 1$, akkor $rend(z) = \infty$;*
- *Ha z-nek van rendje, akkor $|z| = 1$.*

$\varepsilon \in \mathbb{C}$: **$n$-edik egységgyök**, ha $\varepsilon^n = 1$ (speciálisan $\varepsilon = 1$ is n-edik egységgyök). További példák: i 4. egységgyök, hiszen $i^4 = 1$.

5.17. Definíció. *$\varepsilon \in \mathbb{C}$ **primitív n-edik egységgyök**, ha ε n-edik egységgyök és rendje n.*

5.18. Megjegyzés. Rögzített n-re legyen

$$A = \{\varepsilon \in \mathbb{C} \mid \varepsilon^n = 1\} \text{ (n-edik egységgyökök halmaza) és}$$

tekintsük az $\mathcal{A} = (A, \cdot)$ struktúrát, ahol $\cdot$ a $\mathbb{C}$-ből öröklődő szorzás művelete. Ekkor $\mathcal{A}$ izomorf a $(\mathbb{Z}_n, +)$ csoporttal.

Továbbá, ha ε a legkisebb pozitív szögű eleme A-nak, és $z \in \mathbb{C}$, akkor

$$\sqrt[n]{z} = \{\varepsilon^k z_0 \mid k = 0, 1, \ldots, n-1\}.$$

hiszen az ε-al való szorzás hatása: $\frac{2\pi}{n}$-nel való 0 körüli forgatás.

5.19. Példa. *Legyen $n \in \mathbb{N}$ rögzített, $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}$ az összes n-edik egységgyök. Ekkor*

$$\sum_{k=0}^{n-1} \varepsilon_k = 0.$$

1. bizonyítás. Legyen ε a legkisebb pozitív szögű n-edik egységgyök:

$$\varepsilon = \varepsilon_1 = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right).$$

Ekkor $\varepsilon_k = \varepsilon^k$. Legyen

$$A = \sum_{k=0}^{n-1} \varepsilon_k = \sum_{k=0}^{n-1} \varepsilon^k.$$

Ekkor

$$\varepsilon A = \sum_{k=0}^{n-1} \varepsilon^{k+1} = \sum_{k=1}^{n} \varepsilon^k = A.$$

Tehát $\varepsilon A = A$, azaz $(\varepsilon - 1)A = 0$. Mivel $\varepsilon \neq 1$, ezért $A = 0$. □

2. bizonyítás. A mértani sorozatok összegére vonatkozó képlet szerint

$$A = \sum_{k=0}^{n-1} \varepsilon^k = \frac{\varepsilon^n - 1}{\varepsilon - 1} = \frac{1 - 1}{\varepsilon - 1} = 0.$$

□

5.6. Komplex számok és síkgeometria

Ebben az alfejezetben a komplex számokat – algebrai alakjukat derékszögű koordinátákként használva – azonosítjuk a sík pontjaival. Ennek megfelelően pl. közvetlenül fogunk egyes komplex számok mértani helyéről beszélni (ahelyett, hogy a komplex számunknak megfelelő pont mértani helyéről beszélnénk).

5.20. Példa. *Hol vannak azok a $z \in \mathbb{C}$ komplex számok, melyekre*

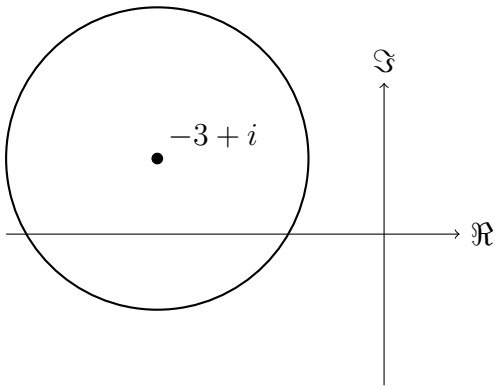
$$|z + 3 - i| = 2?$$

Megoldás. Átalakítjuk:

$$|z - (-3 + i)| = 2.$$

Tehát azoknak a z kompelx számoknak a mértani helyét keressük, melyek távolsága $-3 + i$-től 2. Ez egy kör:

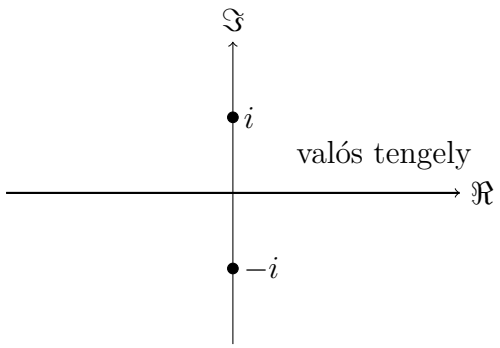
- Középpont: $-3 + i$;
- Sugár: 2.



5.21. Példa. *Hol vannak azok a $z \in \mathbb{C}$ komplex számok, melyekre*

$$|z + i| = |z - i|?$$

1. Megoldás. z távolsága i-től és $-i$-től azonos. Ezért a kérdéses z komplex számok mértani helye az $(i, -i)$ szakasz felezőmerőlegese (mint egyenes).



2. Megoldás. Vegyük fel z-t algebrai alakban: $z = x + iy$ és fogalmazzuk át a feltételünket:

$$\begin{aligned} |x + iy + i| &= |x + iy - i| \\ |x + i(y+1)| &= |x + i(y-1)| \\ \sqrt{x^2 + (y+1)^2} &= \sqrt{x^2 + (y-1)^2} \quad \text{(abszolútérték nem negatív valós).} \end{aligned}$$

Négyzetre emelünk:

$$\begin{aligned} x^2 + (y+1)^2 &= x^2 + (y-1)^2 \\ x^2 + y^2 + 2y + 1 &= x^2 + y^2 - 2y + 1 \\ 2y &= -2y \\ 4y &= 0 \\ y &= 0. \end{aligned}$$

Tehát z a valós tengelyen van (és a valós tengely mindegyik pontja megfelelő).

5.7. Binomiális együtthatók és a Binomiális Tétel

5.22. Definíció (Binomiális együttható)**.** *Legyen $n \in \mathbb{N}$, $k \in \mathbb{N}$, $0 \leq k \leq n$. Az $\binom{n}{k}$ szimbólum (melyet binomiális együtthatónak nevezünk) azt a számot jelöli, ahány k elemű részhalmaza van egy n elemű halmaznak. Ismertnek tekintjük, hogy*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

5.23. Példa (Pascal-háromszög)**.**

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 1 & & 1 & & \\ & & 1 & 2 & 1 & & \\ & 1 & 3 & 3 & 1 & & \\ & 1 & 4 & 6 & 4 & 1 & \\ 1 & 5 & 10 & 10 & 5 & 1 & \end{array}$$

A Pascal-háromszög szélein 1-esek vannak, beljebb minden elem egyenlő a felette levő két elem összegével. A Pascal-háromszög sorait, és a sorokon belül az elemeket 0-tól sorszámozzuk. Ismertnek tekintjük, hogy ezzel a konvencióval a Pascal-háromszög n. sorának k. eleme éppen $\binom{n}{k}$. Például, a Pascal-háromszög 2. sorának 0. eleme $\binom{2}{0} = 1$, 2. sorának 1. eleme $\binom{2}{1} = 2$, stb.

5.24. Tétel (Binomiális tétel)**.** *Tetszőleges kommutatív gyűrű tetszőleges a, b elemeire és $n \in \mathbb{N}^+$-ra*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

(ahol egy c gyűrűelem az $m \in \mathbb{N}$ számmal való mc szorzatán az m-tagú $c+c+\ldots+c$ ismételt összeadás eredményét értjük).

Bizonyítás. $(a+b)^n = (a+b)(a+b)\cdots(a+b)$ (n db) zárójelek felbontásával: hányszor kapunk $a^{n-k}b^k$-t? Pontosan annyiszor, ahányféleképpen ki tudunk választani k darab b-t az n darab tényező közül, azaz $\binom{n}{k}$-szor. □

5.25. Megjegyzés (Pascal-tetraéder)**.** *A $(a+b+c)^n = ((a+b)+c)^n$ kifejtéséhez hasonlóan definiálhatjuk a trinomiális együtthatókat.*

A Binomiális tétel alkalmazásai

5.26. Példa.

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 1^k = \sum_{k=0}^{n} \binom{n}{k}.$$

5.27. Példa.

$$0 = (1-1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} (-1)^k = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \quad \textit{ha } n > 0.$$

5.28. Példa. *Egyrészt a Binomiális Tételt (5.24 Tételt) alkalmazva:*

$$(1+i)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} i^k = \sum_{k=0}^{n} \binom{n}{k} i^k.$$

Másrészt a hatványozást trigonometrikus alakban elvégezve:

$$(1+i)^n = (\sqrt{2})^n \left(\cos\frac{\pi}{4} + i \sin\frac{\pi}{4}\right)^n = 2^{n/2} \left(\cos\frac{n\pi}{4} + i \sin\frac{n\pi}{4}\right).$$

A valós és képzetes részek összehasonlításával:

$$\sum_{\substack{k=0 \\ k \text{ páros}}}^{n} (-1)^{k/2} \binom{n}{k} = 2^{n/2} \cos\frac{n\pi}{4}, \qquad \sum_{\substack{k=0 \\ k \text{ páratlan}}}^{n} (-1)^{(k-1)/2} \binom{n}{k} = 2^{n/2} \sin\frac{n\pi}{4}.$$

5.29. Példa. *Legyen $\varphi \in \mathbb{R}$, $n \in \mathbb{N}$ tetszőleges. Fejezzük ki $\sin(n\varphi)$-t és $\cos(n\varphi)$-t $\sin(\varphi), \cos(\varphi)$-vel.*

Az ötlet az, hogy a $(\cos\varphi + i\sin\varphi)$ komplex szám n-edik hatványát két módon (egyszer a Binomiális Tétel segítségével, egyszer pedig a trigonometrikus alakokra vonatkozó hatványozással) számoljuk ki; mivel ugyanazt a dolgot számoljuk ki kétféle módon, a két eredmény egyenlő lesz...
Egyrészt Binomiális tétellel:

$$\begin{aligned}(\cos\varphi + i\sin\varphi)^n &= \sum_{k=0}^{n} \binom{n}{k} (\cos\varphi)^{n-k} (i\sin\varphi)^k \\ &= \sum_{k=0}^{n} \binom{n}{k} i^k \cos^{n-k}\varphi \sin^k\varphi.\end{aligned}$$

Másrészt trigometrikus alak hatványaként:

$$(\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi).$$

A valós részek összehasonlításából:

$$\cos(n\varphi) = \sum_{\substack{k=0 \\ k \text{ páros}}}^{n} (-1)^{k/2} \binom{n}{k} \cos^{n-k}\varphi \sin^k\varphi;$$

hasonlóan, a képzetes részek összehasonlításából:

$$\sin(n\varphi) = \sum_{\substack{k=0 \\ k \text{ páratlan}}}^{n} (-1)^{(k-1)/2} \binom{n}{k} \cos^{n-k}\varphi \sin^k\varphi.$$

5.30. Példa ($n = 3$ eset)**.**

$$\begin{aligned}(\cos\varphi + i\sin\varphi)^3 &= \binom{3}{0}\cos^3\varphi(i\sin\varphi)^0 + \binom{3}{1}\cos^2\varphi(i\sin\varphi)^1 \\ &\quad + \binom{3}{2}\cos\varphi(i\sin\varphi)^2 + \binom{3}{3}\cos^0\varphi(i\sin\varphi)^3 \\ &= \cos^3\varphi + 3i\cos^2\varphi\sin\varphi - 3\cos\varphi\sin^2\varphi - i\sin^3\varphi \\ &= (\cos^3\varphi - 3\cos\varphi\sin^2\varphi) + i(3\cos^2\varphi\sin\varphi - \sin^3\varphi).\end{aligned}$$

Másrészt: $(\cos\varphi + i\sin\varphi)^3 = \cos(3\varphi) + i\sin(3\varphi)$.
A valós részek egyenlőségéből:

$$\cos(3\varphi) = \cos^3\varphi - 3\cos\varphi\sin^2\varphi.$$

A képzetes részek egyenlőségéből:

$$\sin(3\varphi) = 3\cos^2\varphi\sin\varphi - \sin^3\varphi.$$

6. Polinomok

Alapdefiníciók

Polinomokkal korábban ebben az anyagban is találkoztunk már, sőt középiskolai tanulmányainkból is ismerős lehet a polinomok fogalma. Eddigi ismereteink szerint az x-et, mint határozatlant tartalmazó polinomok

$$a_0 + a_1x + a_2x^2 + ...$$

alakú véges összegek; egy polinomot tehát úgy szoktunk megadni, hogy megadjuk az egyes x-hatványok együtthatóit. Mivel a polinomok véges összegek, az első néhány együttható felsorolása után elég azt rögzíteni, hogy a további (nagyobb kitevős x-hatványokhoz tartozó) együtthatók mind nullák. Más szavakkal, egy polinomot azonosíthatunk a nemnulla együtthatói véges sorozatával (úgy, hogy a sorozatunk nulladik tagja x^0 együtthatóját, első tagja x^1 együtthatóját, stb. adja meg). Ezt teszi precízzé a következő definíció, melyben nem az a lényeg, hogy konkrétan hogyan definiáltuk a polinomokat, hanem az az igazán fontos, hogy a polinomokat így-vagy-úgy, de teljesen precízen definiáltuk. Később látjuk majd, hogy az alábbi definíció praktikus jelölést biztosít rengeteg kérdés tisztázásához.

6.1. Definíció (Polinomok halmaza)**.** *Legyen R kommutatív, egységelemes gyűrű. Az R feletti* ***polinomok halmazát*** *a következőképpen definiáljuk:*

$$A = \{a : \mathbb{N} \to R \mid \exists N \in \mathbb{N} : \forall n > N : a(n) = 0\}.$$

A ***műveletek****:*

- ***Összeadás****:* $(a + b)(k) = a(k) + b(k) \quad (\forall k \in \mathbb{N})$*;*
- ***Szorzás****:* $(ab)(k) = \sum_{m=0}^{k} a(m)b(k-m) \quad (\forall k \in \mathbb{N})$*.*

Az így kapott $(A, +, \cdot)$ struktúrát $R[x]$-el jelöljük, és megelőlegezve a 6.5 Tételt, R feletti ***egyváltozós polinomgyűrűnek*** *nevezzük.*

6.2. Definíció (Jelölés és alakok)**.** *Ha $a \in A$ egy polinom, akkor a következő alakban írhatjuk:*

$$a = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

ahol $a_k = a(k)$ a k. ***együttható*** *és x a* ***határozatlan****.*

Általában x^0-t nem írjuk ki és ha $a_r = 0$, akkor nem mindig írjuk ki a megfelelő tagot.

Polinomok foka és speciális tulajdonságai

6.3. Definíció (Fok)**.** *Legyen $a \in R[x]$. Ekkor az a polinom* ***fokát*** *a következőképpen definiáljuk:*

$$\deg(a) = \begin{cases} -\infty, & \text{ha } \forall k : a_k = 0; \\ \max\{k \mid a_k \neq 0\}, & \text{különben.} \end{cases}$$

Ha $b = b_0 + b_1 x + \cdots + b_n x^n$ *és* $b_n \neq 0$*, akkor*

- b_n *a* b *polinom* ***főegyütthatója****;*
- *ha* $b_n = 1$*, akkor a* b *polinom* ***monikus****;*
- *a* $b_k x^k$ *alakú összeadandókat a* b *polinom* ***tagjainak*** *nevezzük.*

6.4. Tétel (Alaptulajdonságok)**.**

1. *A* $\varphi : R \to R[x]$*,* $\varphi(a) = a$ *(jobboldalon a konstans polinom van) függvény injektív homomorfizmus;*
2. *Ha* $f, g \in R[x]$*, akkor*
$$\deg(f + g) \leq \max\{\deg(f), \deg(g)\};$$
3. *Ha* R *nullosztómentes, akkor*
$$\deg(f \cdot g) = \deg(f) + \deg(g).$$

Bizonyítás. Az állítások $\mathbb{R}[x]$-ben való megfelelőjét középiskolában már megismertük; a bizonyítás $\mathbb{R}$ helyett tetszőleges R gyűrűre hasonlóan végezhető el. □

6.5. Tétel (Gyűrűtulajdonságok)**.**

1. $R[x]$ *kommutatív egységelemes gyűrű;*
2. *Ha* R *nullosztómentes, akkor* $R[x]$ *is nullosztómentes.*

Bizonyítás.

1. A gyűrűaxiómák könnyen (de hosszadalmasan) ellenőrízhetők, ezért nem részletezzük, hogy:
   - Az összeadás kommutatív csoportot alkot;
   - A szorzás asszociatív és kommutatív;
   - A disztributivitás teljesül;
   - Az 1 konstans polinom egységelem.
2. Tegyük fel, hogy R nullosztómentes. Legyenek $f, g \in R[x]$, $f \neq 0$, $g \neq 0$.
   Ekkor $\deg(f) \geq 0$ és $\deg(g) \geq 0$, így az előző tétel (3) pontja miatt:
   $$\deg(f \cdot g) = \deg(f) + \deg(g) \geq 0.$$
   Ezért $f \cdot g \neq 0$.

□

6.1. Polinomok oszthatósága és maradékos osztása

6.6. Definíció (Oszthatóság)**.** *Legyenek $f, g \in R[x]$. A 2.5 Definícióra emlékeztetve, az $R[x]$ gyűrű esetében is azt mondjuk, hogy f* ***osztója*** *g-**nek** (jelölés: $f \mid g$), ha*

$$\exists h \in R[x] : g = f \cdot h.$$

6.7. Definíció (Maradékos osztás)**.** *Azt mondjuk, hogy $f \in R[x]$* ***maradékosan osztható*** *$g \in R[x]$-szel, ha*

$$\exists q, r \in R[x] : f = g \cdot q + r \quad \textit{és} \quad \deg(r) < \deg(g).$$

Ekkor q-t ***hányadosnak****, r-et* ***maradéknak*** *nevezzük.*

6.8. Példa (Ellenpélda)**.** *$\mathbb{Z}[x]$-ben x^2 nem osztható maradékosan $2x$-szel:*

Tegyük fel, hogy $x^2 = (2x)h + r$, ahol $\deg(r) < 1$.

Ekkor $\deg(r) \leq 0$ és ezért $r = 0$, de h főegyütthatója is egész szám kellene, hogy legyen, emiatt nem található megfelelő h.

6.9. Tétel (Maradékos osztás feltétele)**.** *Legyen R kommutatív, egységelemes gyűrű és $f \in R[x]$. A következő állítások ekvivalensek:*

1. *f-fel lehet maradékosan osztani, azaz*

$$\forall g \in R[x]\ \exists h, r \in R[x] : g = f \cdot h + r, \quad \deg(r) < \deg(f);$$

2. *f főegyütthatója invertálható R-ben.*

Bizonyítás. **(1) ⇒ (2)**: (1) miatt $f \neq 0$, legyen $n = \deg(f)$. (1) alapján a $g = x^n$ polinomot is el lehet maradékosan osztani f-el, ezért

$$\exists h, r \in R[x] : x^n = f \cdot h + r, \text{ ahol } \deg(r) < \deg(f).$$

A főegyütthatók összehasonlításából:

$$1 = a_n \cdot (h \text{ főegyütthatója}),$$

mert $\deg(r) < \deg(f)$. Ezért a_n invertálható R-ben.

(2) ⇒ (1): Legyen $f = a_n x^n + \cdots + a_1 x + a_0$, ahol a_n invertálható. Ellentmondást keresve tegyük fel, hogy van olyan $g \in R[x]$, amit nem lehet f-fel maradékosan osztani; legyen m a legkisebb fokszám, amelyre van ilyen g és legyen $g \in R[x]$ olyan, hogy $\deg(g) = m$ és g-t nem lehet f-fel maradékosan osztani. Vegyük fel g együtthatóit:

$$g = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0.$$

Figyeljük meg, hogy $m \geq n$ (különben $g = 0 \cdot f + g$ jó maradékos osztás lenne). Tekintsük a következő polinomot:

$$\tilde{g} = g - (a_n^{-1} b_m x^{m-n}) f.$$

Ekkor $\deg(\tilde{g}) < m$, ezért m választása (illetve g fokszámának minimalitása) miatt $\tilde{g}$ maradékosan osztható f-fel:

$$\exists h, r \in R[x] : \tilde{g} = h \cdot f + r, \quad \deg(r) < \deg(f).$$

Ezért

$$g = \tilde{g} + (a_n^{-1} b_m x^{m-n}) f = (h + a_n^{-1} b_m x^{m-n}) f + r.$$

Ez ellentmondás, hiszen az előző sor szerint g is maradékosan osztható f-fel. □

6.10. Tétel (Maradékos osztás egyértelműsége)**.** *Ha $f \in R[x]$ főegyütthatója invertálható, akkor a maradékos osztás egyértelmű.*

Bizonyítás. Tegyük fel, hogy

$$g = h_1 f + r_1 = h_2 f + r_2$$

ahol $\deg(r_1) < \deg(f)$ és $\deg(r_2) < \deg(f)$. Ekkor

$$(*) \qquad (h_1 - h_2) f = r_2 - r_1.$$

Ellentmondást keresve tegyük fel, hogy $h_1 \neq h_2$; ebből:

- $\deg((h_1 - h_2) f) \geq \deg(f)$ (mert f főegyütthatója invertálható);
- $\deg(r_2 - r_1) < \deg(f)$.

Ezek szerint $(*)$ baloldalának foka legalább $deg(f)$, jobboldalának foka pedig szigorúan kisebb ennél. Ez ellentmondás, ezért mégiscsak $h_1 = h_2$ és ebből $r_1 = r_2$. □

6.11. Következmény. *Ha R test, akkor $R[x]$-ben minden nem nulla polinommal lehet maradékosan osztani.*

Bizonyítás. Testben minden nem nulla elem invertálható, így minden nem nulla polinom főegyütthatója is invertálható. □

6.2. Polinomfüggvények

6.12. Definíció. *Legyen $f \in R[x]$, $f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ és $c \in R$. Ekkor f **c-nél vett helyettesítési értéke**:*

$$f(c) = a_n c^n + a_{n-1} c^{n-1} + \cdots + a_0.$$

*Az f-hez tartozó **polinomfüggvény**:*

$$p_f : R \to R, \quad p_f(c) = f(c).$$

6.13. Példa. *Legyen $f_1 = x$, $f_2 = x^2$ a $\mathbb{Z}_2$ felett.*

c	$f_1(c)$	$f_2(c)$
0	*0*	*0*
1	*1*	*1*

Itt $f_1 \neq f_2$ mint polinomok, de $p_{f_1} = p_{f_2}$ mint függvények.

6.14. Definíció. *$c \in R$ **gyöke** f-nek, ha $f(c) = 0$.*

6.15. Megjegyzés. Klasszikus kérdés: Adott $f \in R[x]$-nek keressük meg a gyökeit! Hogyan keressük a gyököket? $c \in$?? (Általában, – nem mindig – ha $f \in R[x]$, akkor c-t R-ben keressük).

6.16. Példa. *Legyenek $f_1(x) = x^n - 2$ és $f_2(x) = x^2 + 1$. Hány gyökük van? A különböző $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ testekben és a $\mathbb{Z}$ gyűrűben:*

	$\mathbb{Q}$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{Z}$
f_1	*0*	*1-2*	*n*	*0*
f_2	*0*	*0*	*2*	*0*

6.17. Megjegyzés. A 6.21 Következményben látjuk majd, hogy kummutatív, nullosztómentes gyűrűk feletti polinomgyűrűkben a gyökök száma mutat valamiféle rendezettséget (legfeljebb annyi gyök van, mint a polinom fokszáma). A 6.22 pontban látunk majd patologikus példákat olyan (nullosztót tartalmazó) kommutatív gyűrűre, melyben az előző állítás nem igaz; ott a gyökök számának furcsa viselkedése a nullosztók létezésén múlik. A gyökök számának szempontjából nemcsak a nullosztók, hanem az alapgyűrű kommutativitása is lényeges: a 2025 őszi kurzus 10. heti gyakorlatára szánt 5. feladatában láttuk, hogy a kvaterniók NEM-KOMMUTATíV teste felett van olyan másodfokú polinom, melynek végtelen sok gyöke van; itt a problémát a kommutativitás hiánya okozza (hiszen minden test nullosztómentes).

A reneszánsz idejében már intenzíven vizsgálták a harmadfokú, sőt még magasabbfokú polinom-egyenleteket is. Ugyanakkor a negatív számok fogalma nem volt még teljesen rendbe téve. Emiatt, a mai jelöléseinket használva, a másodfokú

$$f(x) = x^2 + ax + b \in \mathbb{R}[x]$$

polinom vizsgálata is nehézkes volt. Mivel együtthatóként neagtív számokat nem használtak még, az előző f-re az $f(x) = 0$ egyenletet 4 különböző esetre bontották, és külön kezelték az

$$\begin{aligned} x^2 &= ax + b; \\ x^2 + ax &= b; \\ x^2 + b &= ax; \\ x^2 + ax + b &= 0 \end{aligned}$$

alakú egyenleteket (úgy értve, hogy mind a 4 esetben a, b, c nemnegatív számok). Látjuk majd, hogy e technikai nehézségek ellenére meglepően messzire jutottak a harmad- és negyedfokú egyenletek megoldása terén.

6.3. Polinomok gyökei és faktorizáció

6.18. Tétel (A behelyettesítés homomorfizmus). [2] *Legyen R kommutatív gyűrű, $c \in R$ és legyen*

$$\Phi_c : R[x] \to R, \quad \Phi_c(f) = f(c).$$

Ekkor Φ_c (azaz a c-t behelyettesítő függvény) gyűrűhomomorfizmus.

Bizonyítás. Azt kell ellenőríznünk, hogy Φ_c megtartja a gyűrű-műveleteket. Ehhez legyenek $f, g \in R[x]$, mondjuk

$$f = \sum_k a_k x^k, \quad g = \sum_k b_k x^k.$$

Ekkor egyrészt

$$\Phi_c(f+g) = \Phi_c\big(\sum_k a_k x^k \ + \ \sum_k b_k x^k\big) = \Phi_c\big(\sum_k (a_k + b_k) x^k\big) =$$

$$\sum_k (a_k + b_k) c^k = \sum_k a_k c^k + \sum_k b_k c^k = \Phi_c(f) + \Phi_c(g).$$

Másrészt, hasonlóan

$$\Phi_c(fg) = \Phi_c\Big(\big(\sum_k a_k x^k\big)\big(\sum_k b_k x^k\big)\Big) = \Phi_c\Big(\sum_k \big(\sum_{j=0}^{k} a_j b_{k-j}\big) x^k\Big) =$$

$$\sum_k \sum_{j=0}^{k} \big(a_j b_{k-j}\big) c^k = \big(\sum_k a_k c^k\big)\big(\sum_k b_k c^k\big) = \Phi_c(f)\Phi_c(g).$$

□

6.1. Konvenció.
A továbbiakban R mindig kommutatív, egységelemes gyűrű.

6.19. Tétel. *Legyen $f \in R[x]$, $c \in R$. Ekkor ekvivalensek:*

(a) $f(c) = 0$ (azaz c gyöke f-nek);

(b) $(x - c) \mid f$ (azaz $x - c$ osztója f-nek).

[2] Ez a tétel teljesen kimaradt abból az anyagból, amit a hallgatóktól kaptam. Leírom a bizonyítást újra, de ebben a bizonyításban a jelölések nem feltétlenül lesznek azonosak azzal, ahogy előadáson (emlékeim szerint 2025 október 17.-én) elmondtam.

Bizonyítás. **(a) ⇒ (b)**: $x - c$ főegyütthatója 1, ezért a 6.9 Tétel miatt lehet vele maradékosan osztani: vannak olyan $q, r \in R[x]$ polinomok, melyekre

$$f = (x - c)q + r, \quad \deg(r) < 1.$$

Mindkét oldalba c-t helyettesítve, a 6.18 Tétel miatt

$$0 = f(c) = (c - c)q(c) + r(c) = r(c),$$

tehát $r = 0$.
(b) ⇒ (a): Tudjuk, hogy $\exists g \in R[x] : f = (x - c)g$. Ismét, mindkét oldalba c-t helyettesítve és a 6.18 Tételre hivatkozva azt kapjuk, hogy $f(c) = (c - c)g(c) = 0$, azaz c gyöke f-nek. □

6.20. Tétel. *Ha R nullosztómentes és $c_1, \ldots, c_k$ különböző gyökei f-nek, akkor*

$$(x - c_1)(x - c_2) \cdots (x - c_k) \mid f.$$

Bizonyítás. Indukció k-ra.
Alaplépés: $k = 1$, az előző, 6.19 Tétel miatt.
Indukciós lépés: Tegyük fel, hogy $c_1, \ldots, c_k$ különböző gyökei f-nek. Az előző, 6.19 Tétel szerint: $(x - c_1) \mid f$, azaz $\exists g \in R[x] : f = (x - c_1)g$. Ismét a 6.18 Tétel miatt, tetszőleges $2 \leq i \leq k$-ra:

$$f(c_i) = (c_i - c_1)g(c_i) = 0.$$

Mivel a $c_1, \ldots c_k$ gyökök páronként különböznek, ezért $c_i - c_1 \neq 0$, és így a nullosztómentesség miatt $g(c_i) = 0$. Ezért $c_2, \ldots, c_k$ gyökei g-nek, és alkalmazható az indukció. □

6.21. Következmény. *Ha R kommutatív, nullosztómentes gyűrű, $f \in R[x]$, $\deg(f) = n$, akkor f-nek legfeljebb darab n különböző gyöke lehet.*

Bizonyítás. Legyenek f gyökei $c_1, c_2, \cdots \in R$. Az 6.20 Tétel ismételt alkalmazásával f-ből az $(x - c_1), (x - c_2), \ldots$ lineáris tényezőket egyesével ki lehet emelni. Minden ilyen kiemelés 1-el csökkenti a fokszámot, tehát a kiemeléseket legfeljebb $deg(f) = n$-szer lehet megismételni. Ezért a gyökök száma nem lehet nagyobb n-nél. □

6.22. Példa. A 6.21 Következmény nullosztót tartalmazó gyűrűkben nem marad érvényben. Pl. a másodfokú $x^2 - 1$-nek $\mathbb{Z}_8$-ban gyöke $1, 3, 5, 7$, ez összesen 4 darab (ezzel kapcsolatban még a 6.55 Példára is utalunk).

6.4. Gyökök és együtthatók kapcsolata

Mint említettük már, a polinomegyenletek megoldásának klasszikus problémája az, hogy adott polinom együtthatóiból állítsuk elő a gyökeit. Erre a kérdésre vissza fogunk térni, ebben az alfejezetben azonban a fordított, és jóval egyszerűbb kérdést vizsgáljuk: hogyan állnak elő a gyökökből az együtthatók, azaz adott $c_1, \dots c_n$ gyűrűelemekhez hogyan találhatjuk meg egy olyan polinom együtthatóit, melynek pont a $c_1, \dots c_n$ gyűrűelemek a gyökei. Ez a fordított, és könnyebb irány hasznos lesz az eredeti, gyökök előállítására vonatkozó kérdések vizsgálatánál is.

6.23. Tétel (Viète-formulák)**.** *Legyen* $f = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ *monikus polinom. Ha* $f = (x - c_1)(x - c_2)\cdots(x - c_n)$*, akkor:*

$$\begin{aligned}
a_{n-1} &= -(c_1 + c_2 + \dots + c_n);\\
a_{n-2} &= c_1c_2 + c_1c_3 + \dots + c_{n-1}c_n;\\
&\vdots\\
a_1 &= (-1)^{n-1}(c_2c_3\cdots c_n + c_1c_3\cdots c_n + \dots + c_1c_2\cdots c_{n-1});\\
a_0 &= (-1)^n c_1c_2\cdots c_n.
\end{aligned}$$

Bizonyítás. A szorzatot kibontva:

$$(x - c_1)(x - c_2)\cdots(x - c_n) = x^n - (c_1 + \dots + c_n)x^{n-1} + \dots + (-1)^n c_1\cdots c_n.$$

Az a_{n-k} együttható egyenlő $(-1)^k$-szor a $\{c_1, \dots c_n\}$ halmaz összes k-elemű részhalmaza szorzatainak összegével. □

Gyakorlati szempontból az előző tétel azt mutatja, hogy szorzat alakban adott polinomok esetében a szorzást (zárójelek felbontását) csak indokolt esetben célszerű elvégezni, mert ezt bármikor, viszonylag könnyen megtehetjük. Ugyanakkor egy együtthatóival adott polinom szorzattá alakítása sok fejtörést okozhat.

6.24. Példa. Oldjuk meg $\mathbb{C}$-n:

$$\begin{aligned}
u + v &= 4;\\
uv &= 3.
\end{aligned}$$

A Viète-formulák miatt u és v gyökei az

$$x^2 - 4x + 3 = 0$$

másodfokú egyenletnek. Ennek gyökei: $x = 1$ és $x = 3$, tehát $(u, v) = (1, 3)$ vagy $(3, 1)$.

6.5. Harmadfokú egyenletek $\mathbb{C}$ felett

Általános alak és redukció

- Általános alak: $ax^3 + bx^2 + cx + d = 0$.
- Redukált alak: $y^3 + py + q = 0$.

6.25. Tétel (Harmadfokú egyenlet redukálása)**.** *Minden harmadfokú*

$$ax^3 + bx^2 + cx + d = 0, \quad a \neq 0$$

alakú egyenlet átalakítható $y^3 + py + q = 0$ *alakúra.*

Bizonyítás. A főegyütthatóval osztva feltehetjük, hogy polinomunk monikus:

$$x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = 0.$$

A baloldali polinomot rendezzük át $x + \frac{b}{3a}$ hatványai szerint:

$$x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = \Big(x + \frac{b}{3a}\Big)^3 - x\Big(\frac{b^2}{3a^2}\Big) - \Big(\frac{b}{3a}\Big)^3 + \frac{c}{a}\Big(x + \frac{b}{3a}\Big) - \frac{bc}{3a^2} + \frac{d}{a} =$$

$$\Big(x + \frac{b}{3a}\Big)^3 + \Big(\frac{c}{a} - \frac{b^2}{3a^2}\Big)\Big(x + \frac{b}{3a}\Big) + \frac{b^3}{9a^3} - \frac{b^3}{27a^3} - \frac{bc}{3a^2} + \frac{d}{a},$$

ez, az

$$x = y - \frac{b}{3a}, \quad p = \frac{c}{a} - \frac{b^2}{3a^2}, \quad q = \frac{2b^3}{27a^3} - \frac{bc}{3a^2} + \frac{d}{a}$$

jelölésekkel $y^3 + py + q = 0$ alakú. □

Az $y^3 + py + q = 0$ egyenlet megoldása $\mathbb{C}$-ben.

Feltehetjük, hogy $p \neq 0$, mert $p = 0$ esetén y értékeit azonnal megkapnánk a $-q$-ból való (komplex) köbgyökvonással.

Keressük y-t $y = u + v$ alakban. A Binomiális tételből (5.24 Tételből)

$$y^3 = (u + v)^3 = u^3 + 3u^2v + 3uv^2 + v^3 =$$

$$u^3 + v^3 + 3uv(u + v) = u^3 + v^3 + 3uv \cdot y,$$

azaz

$$y^3 - 3uv \cdot y - (u^3 + v^3) = 0,$$

és ez mindig igaz, ha $y = u + v$. Annak érdekében, hogy az előző sorban szereplő egyenletet kapcsolatba hozzuk az $y^3 + py + q = 0$ egyenlettel, következő lépésként olyan u, v-t keresünk, melyekre $p = -3uv$ és $q = -(u^3 + v^3)$, azaz

(1) $uv = -\frac{p}{3}$ és
(2) $u^3 + v^3 = -q$.

Az előző (1) feltételből következik, hogy

(3) $u^3v^3 = -\left(\frac{p}{3}\right)^3$,

de ez a következtetés nem fordítható meg: $w^3 = -\left(\frac{p}{3}\right)^3$-ből $w = -\left(\frac{p}{3}\right)$ nem következik, mert három (különböző) komplex szám is van, melyek köbe $-\left(\frac{p}{3}\right)^3$ (hiszen $-\left(\frac{p}{3}\right)$-et a harmadik egységgyökökkel szorozva mindhárom szám köbe $-\left(\frac{p}{3}\right)^3$ lesz). Ezt fejben tartva, keressünk olyan u, v-t, melyre (2) és (3) teljesül. Láthatjuk, hogy u^3 és v^3 összegére és szorzatára vonatkozó előírásokat kell kielégítenünk, ezért a másodfokú egyenletekre vonatkozó Viète-formulák (azaz a 6.23 Tétel) alapján u^3 és v^3 gyökei a következő másodfokú egyenletnek:

$$z^2 + qz - \left(\frac{p}{3}\right)^3 = 0.$$

Ennek megoldásai:

$$z = -\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}.$$

Tehát:

$$(4) \qquad u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}, \qquad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

Ebből $u+v$-re 9 db lehetőség is adódhat, mert a komplex köbgyökvonás eredményei általában 3-elemű halmazok. Ugyanakkor a 6.21 Következmény alapján harmadfokú egyenletünknek (legfeljebb) 3 darab gyöke lehet. A hamis-gyökök megjelenését az okozza, hogy az (1) $\rightarrow$ (3) következtetés nem fordítható meg (azaz, amint azt átgondoltuk, (3)-ból nem következik (1)). Tudjuk azonban, hogy u-nak és v-nek nemcsak a fenti (3) feltételt, hanem az ennél erősebb (1) feltételt is teljesítenie kell, azaz teljesülnie kell $uv = -\frac{p}{3}$-nak is.

Válasszuk u értékét a (4)-ből adódó 3 lehetőség közül tetszőlegesen; $u \neq 0$, mert (1) alapján $u = 0$-ból $p = 0$ következne. Legyen $v' = -\frac{p}{3u}$, ekkor $u \cdot v' = -\frac{p}{3}$ és

$$(v')^3 = (-\frac{p}{3u})^3 = -\left(\frac{p}{3}\right)^3 \cdot \frac{1}{u^3} = v^3,$$

tehát u, v'-re igaz, hogy $u + v'$ gyöke $y^3 + py + q$-nak.

Legyen ε a 120°-hoz tartozó harmadik egységgyök, ekkor a lehetséges értékek

u-ra: $u_0,\ u_0 \cdot \varepsilon$ és $u_0 \cdot \varepsilon^2$, ezekből meghatározhatók v megfelelő értékei u_0-hoz, $u_0 \cdot \varepsilon$-hoz és $u_0 \cdot \varepsilon^2$-hez. Ott tartunk tehát, hogy harmadfokú egyenletünkhöz találtunk 3 gyököt, de azt még nem tudjuk, hogy megtaláltuk-e az összes gyököt (hiszen elvileg elképzelhető lenne, hogy a fenti módszerrel megtalált 3 gyökünk közül néhány akkor is egyenlő lenne egymással, ha a harmadfokú egyenletünknek 3 különböző gyöke van). Szerencsére nem ez a helyzet, a fenti módszerrel elő tudjuk állítani harmadfokú egyenletünk összes gyökét. Ezt úgy lehetne ellenőrizni, hogy az

$$(x - (u_0 + v_0))(x - (u_1 + v_1))(x - (u_2 + v_2))$$

polinomban felbontva a zárójeleket (elvégezve a szorzásokat), éppen $y^3 + py + q$ adódna, ami mutatná, hogy a fenti módon minden gyököt megkaptunk (ezt a hosszadalmas számolást elhagyjuk, vizsgán sem kell tudni).

Összefoglalva, azt kaptuk, hogy $y^3 + py + q = 0$ megoldása a következő. Ha $p = 0$, akkor $y = \sqrt[3]{-q}$. Ha $p \neq 0$, akkor $y = u + v$, ahol

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

és $v = -\frac{p}{3u}$; ezeket nevezik Cardano-képleteknek. □

A harmadfokú egyenletek most ismeretett megoldási módszerének nagyon érdekes története van, előadáson meséltem róla. Javaslom, hogy nézz utána a „harmadfokú egyenlet", „ Gerolamo Cardano" és „Niccolo Tartaglia" Wikipewdia-oldalainak (Cardano személyét ne keverd össze a Cardano elnevezésű kriptovalutával). Ugyancsak elmeséltem, hogy $\mathbb{R}$, illetve $\mathbb{C}$ felett a negyedfokú egyenletek megoldhatók úgy, hogy minden ($\mathbb{R}$, illetve $\mathbb{C}$ feletti) negyedfokú polinom két másodfokú szorzatára bomlik. Közelebbről, az egyik másodfokú tényező együtthatóit ügyesen választva, a másik másodfokú tényező határozatlan együtthatóira harmadfokú feltételek adódnak, melyeket a Cardano-képletekkel meg lehet határozni. Meséltem arról is, hogy a legalább ötödfokú egyenleteknek nincs szokásos megoldóképlete, egész egyszerűen azért, mert (az 5.1 alfejezet elején, a racionális számok után bevezetett terminológiával) nem minden algebrai szám gyökkifejezés: pl. az

$$x^5 - 4x - 2$$

polinom gyökei algebrai számok (hiszen minden racionális együtthatós polinom minden gyöke algebrai), de meg lehet mutatni, hogy az előző polinom egyetlen gyöke sem gyökkifejezés, ezért e gyököket egész egyszerűen nem tudjuk megnevezni racionális számok, az alapműveletek, és gyökvonások segítségével - emiatt megoldóképlet sem létezhet. Röviden meséltem még Galois-ról, Abel-ről, és arról, hogy bizonyos szerkesztési feladatok (kockakettőzés, szögharmadolás) miért nem oldhatók meg.

A harmadfokú egyenletek megoldásának első lépése az volt, hogy polinomunkból

kiküszöböltük a másodfokú tagot, azaz y^3+py+q alakúra hoztuk úgy, hogy alkalmasan választott u paraméterrel átrendeztük $x-u$ hatványai szerint. Alfejezetünket azzal zárjuk, hogy ez mindig megtehető.

6.26. Tétel (Polinomok $(x-c)$ szerinti felírása)**.** *Legyen R kommutatív, egységelemes gyűrű, $f \in R[x]$, $c \in R$. Ekkor f felírható $(x-c)$ hatványai szerint:*

$$f = b_n(x-c)^n + b_{n-1}(x-c)^{n-1} + \cdots + b_1(x-c) + b_0.$$

Bizonyítás. Teljes indukció $\deg(f)$-re.
Alaplépés: $\deg(f) = 0$.

$$f = a_0 = 0 \cdot (x-c) + a_0 \text{ (azaz } b_0 = a_0)$$

megfelelő átrendezés.
Indukciós lépés: Osszuk el f-et maradékosan $(x-c)$-vel:

$$f = (x-c)g + d, \quad \deg(g) < \deg(f).$$

Az indukciós feltevés szerint g felírható így:

$$g = b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1.$$

Ebből

$$f = (x-c)[b_n(x-c)^{n-1} + \cdots + b_1] + d = b_n(x-c)^n + \cdots + b_1(x-c) + b_0.$$

ahol $b_0 = d$. □

6.6. Szimmetrikus Polinomok

Ebben az alfejezetben olyan polinomokkal fogunk foglalkozni, melyekben több változó (határozatlan) is szerepelhet. A többhatározatlanú polinomok gyűrűinek rekurzív definíciójával kezdünk.

6.27. Definíció. *Legyenek $x_1, ...x_n$ határozatlanok. Az $R[x_1]$ egyhatározatlanú polinomgyűrűt definiáltuk már a 6.1 Definícióban. Ha az $(n-1)$-határozatlanú*

$$R[x_1, \ldots, x_{n-1}]$$

polinomgyűrűt definiáltuk már, akkor legyen

$$R[x_1, \ldots, x_n] \;=\; R[x_1, \ldots, x_{n-1}][x_n].$$

Az előző definícióban számít a változók sorrendje: pl. $R[x][y]$ és $R[y][x]$ formálisan nézve különböző gyűrűk (az elsőben y-nak olyan polinomjai vannak, melyekben az együtthatók x polinomjai, a másodikban pedig x olyan polinomjai vannak, melyekben az együtthatók y polinomjai). Noha ezek a gyűrűk szigorú értelemben különböznek, annyira természetes módon izomorfak, hogy azonosítani fogjuk őket, és úgy számolunk bennük, hogy az x és y határozatlanok „egyenrangúak", függetlenül attól, hogy az R alapgyűrűt a határozatlanok milyen sorrendjében bővítettük sokhatározatlanú polinomgyűrűvé. Például, a kéthatározatlanú

$$x^2y^2 + xy^2 + 7 \in R[x, y]$$

polinom tekinthető az $(y^2)x^2 + (y^2)x + 7 \in R[y][x]$ polinomnak, de tekinthető az $(x^2+x)y^2+7 \in R[x][y]$ polinomnak is. Ez a nagyvonalúság a gyakorlatban semmilyen problémát nem fog okozni.

6.28. Definíció. *Egy $f(x_1, \ldots, x_n)$ polinom **szimmetrikus**, ha bármely két változóját felcserélve a polinom nem változik.*

6.29. Példa. $f(x, y) = x^2 + y^2 - 3(x^2 + y^2)^7$ *szimmetrikus.*

A 2025 őszi kurzus 11. adag házi feladataiban a 7. feladat szerint minden permutáció előállítható transzpozíciók kompozíciójaként. Ezért az $f \in R[x_1, \ldots, x_n]$ polinom akkor és csak akkor szimmetrikus, ha az $\{1, 2, \ldots, n\}$ halmaz tetszőleges π permutációjára

$$f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)}).$$

6.30. Definíció (Elemi szimmetrikus polinomok)**.**

$$\sigma_1 = x_1 + x_2 + \cdots + x_n; \qquad \binom{n}{1} = n \text{ db összeadandó}$$

$$\sigma_2 = x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n; \qquad \binom{n}{2} \text{ db összeadandó}$$

$$\vdots$$

$$\sigma_n = x_1x_2 \cdots x_n \qquad \binom{n}{n} = 1 \text{ db összeadandó.}$$

Általában: σ_k-t úgy kapjuk, hogy vesszük az $\{x_1, \ldots, x_n\}$ halmaz összes k-elemű részhalmazát, e k-elemű részhalmazok elemeit összeszorozzuk, és az összes ilyen (k-tényezős) szorzatot összeadjuk.

6.31. Példa. *$x_1x_2x_3 + x_1 + x_2 + x_3$ szimmetrikus, de nem elemi (hiszen elemi szimmetrikus polinomokban minden összeadandó ugyanannyi-tényezős szorzat).*

Következő célunk a Szimmetrikus Polinomok Alaptételének (6.37 Tétel) igazolása, mely durván szólva úgy foglalható össze, hogy minden szimmetrikus polinom előállítható (ráadásul pontosan 1 módon állítható elő) elemi szimmetrikus polinomok polinomjaként. E tétel igazolásához további előkészületekre lesz szükségünk.

Lexikografikus rendezés

6.32. Definíció. *Legyen $(A, \leq)$ rendezett halmaz, $I = \{1, \ldots, n\}$ véges halmaz. Ekkor I **lexikografikus hatványa** az a rendezés, melynek*

- *alaphalmaza*

$$A^I = \{A \text{ elemeiből alkotott, } I\text{-vel indexelt sorozatok}\};$$

- *rendezése: ha $s \neq t$, $s_i = t_i$ minden $i = 1, \ldots, d-1$-re és $s_d < t_d$, akkor $s < t$.*

A lexikografikus rendezés tehát olyan, mint a szigorú névsor szerinti rendezés: s-ben és t-ben megkeressük az első eltérést (az előző definícióban ez a d. pozíciónál van), és s pontosan akkor előzi meg t-t, ha az első eltérés pozíciójánál s-ben kisebb érték van, mint t-ben.

6.33. Tétel. *Ha $(A, \leq)$ jólrendezett és I véges, akkor a lexikografikus hatvány is jólrendezett.*

Bizonyítás. Ezt a tételt nem bizonyítottuk, ebben a kurzusban semmit nem kell tudni ennek a bizonyításáról. Annyit hadartam el, hogy ha a lexikografikus hatvány nem lenne jólrendezett, akkor lenne benne végtelen leszálló lánc. E lánc párjait színezzük meg I színnel aszerint, hogy melyik pozícióban térnek el. Ramsey tétele szerint kapnánk a végtelen leszálló láncunkban egy olyan végtelen részsorozatot, melyben az egymást követő elemek mindig ugyanabban a pozícióban (mondjuk az i. pozícióban) térnének el, ezért az előbbi részsorozatban szereplő sorozatok i. koordinátái A-ban is adnának egy végtelen leszálló láncot, ami ellentmondana A jólrendezettségének. □

6.34. Definíció. *Többhatározatlanú polinomjainkat a kitevőik szerint lexikografikusan rendezzük. $f \in R[x_1, \ldots, x_n]$ **főtagja** $L(f)$ az a tag, amelynek kitevő-sorozata lexikografikusan a legnagyobb.*

6.35. Tétel (Szorzat főtagja)**.** *Ha $f, g \in R[x_1, \ldots, x_n]$, akkor $L(f \cdot g) = L(f) \cdot L(g)$.*

Bizonyítás. Legyen

$$L(f) = ax_1^{d_1} \cdots x_n^{d_n} \quad \text{és} \quad L(g) = bx_1^{e_1} \cdots x_n^{e_n}.$$

Ekkor egyrészt $L(f) \cdot L(g)$ kitevő-sorozata $(d_1 + e_1, \ldots, d_n + e_n)$. Másrészt, vegyük f és g két tetszőleges tagját, az f-ből választott tag kitevő-sorozata legyen $(d_1', \ldots, d_n')$ és a g-ből választott tag kitevő-sorozata legyen $(e_1', \ldots, e_n')$. Ha kiszemelt tagjainkat összeszorozzuk, szorzatuk kitevő-sorozata $(d_1' + e_1', \ldots, d_n' + e_n')$ lesz, és ez lexikografikusan kisebb $(d_1 + e_1, \ldots, d_n + e_n)$-nél, hiszen a főtag definíciója értelmében $(d_1', \ldots, d_n')$ lexikografikusan kisebb, vagy egyenlő $(d_1, \ldots, d_n)$-nél és $(e_1', \ldots, e_n')$ lexikografikusan kisebb, vagy egyenlő $(e_1, \ldots, e_n)$-nél. Emiatt fg főtagjának kitevő-sorozata $L(fg) = (d_1 + e_1, \ldots, d_n + e_n)$, ahogy állítottuk. □

6.36. Példa (Elemi szimmetrikus polinomok főtagjai)**.**

$$\begin{aligned}
L(\sigma_1) &= L(x_1 + \cdots + x_n) = x_1; \\
L(\sigma_2) &= L(x_1x_2 + \cdots + x_{n-1}x_n) = (x_1x_2) = x_1x_2; \\
&\vdots \\
L(\sigma_n) &= L(x_1 \cdots x_n) = x_1 \cdots x_n.
\end{aligned}$$

Szimmetrikus polinomok előállítása

6.37. Tétel (Szimmetrikus Polinomok Alaptétele)**.** *Ha $f \in R[x_1, \ldots, x_n]$ szimmetrikus polinom, akkor $\exists! g \in R[y_1, \ldots, y_n]$:*

$$f(x_1, \ldots, x_n) = g(\sigma_1, \ldots, \sigma_n).$$

Bizonyítás. Bizonyításunk egyúttal módszert is ad g előállítására. Tegyük fel, hogy $L(f) = r \cdot x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}$. Mivel f szimmetrikus polinom, ezért változóit tetszőlegesen permutálva nem változik meg. Emiatt az $\{1, \ldots n\}$ halmaz minden π permutációjára igaz, hogy f-ben van olyan tag, melynek kitevő-rendszere $(\pi(m_1), \pi(m_2), \ldots, \pi(m_n))$. Ezek a permutált kitevőrendszerek lexikografikusan mind kisebbek, vagy egyenlőek f főtagjának kitevő-rendszerénél, ami azt jelenti, hogy az $m_1, m_2 \ldots m_n$ számokra teljesül, hogy

$$(*) \qquad m_1 \geq m_2 \geq \cdots \geq m_{n-1} \geq m_n.$$

Válasszuk a $k_1, \ldots, k_n \in \mathbb{N}$ számokat úgy, hogy:

$$\begin{aligned}
k_1 + k_2 + \cdots + k_n &= m_1; \\
k_2 + \cdots + k_n &= m_2; \\
&\vdots \\
k_{n-1} + k_n &= m_{n-1}; \\
k_n &= m_n.
\end{aligned}$$

Ilyen $k_1, \ldots, k_n$-et az előző egyenleteken alulról felfele haladva $(*)$ miatt valóban tudunk választani. Tekintsük most a $g_1 = r \cdot x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ polinomot. Ekkor:

$$L(g_1(\sigma_1, \ldots, \sigma_n)) = r \cdot x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n} = L(f).$$

Így tehát $f_1 = f - g_1(\sigma_1, \ldots, \sigma_n)$ egy olyan szimmetrikus polinom, melynek főtagja lexikografikusan kisebb, mint f főtagja. Ezt az eljárást ismételgetve a főtag kitevősorozata lexikografikusan csökken, ezért a lexikografikus hatvány jólrendezettsége (6.33 Tétel) miatt véges sok lépésben befejeződik és megkapjuk g-t.

Egyértelműség. Ellentmondást keresve tegyük fel, hogy $g_1, g_2 \in R[y_1, \ldots, y_n]$, $g_1 \neq g_2$ de

$$g_1(\sigma_1, \ldots, \sigma_n) = g_2(\sigma_1, \ldots, \sigma_n).$$

Ekkor

(a) $(g_1 - g_2)(\sigma_1, \ldots, \sigma_n) = 0$, de
(b) $g_1 - g_2 \neq 0$.

Ezért (b) miatt $L(g_1 - g_2) \neq 0$. A 6.18 Tétel szerint a behelyettesítés homomorfizmus, ezért $(L(g_1 - g_2))(\sigma_1, ..., \sigma_n) \neq 0$, amiből viszont $(g_1 - g_2)(\sigma_1, ..., \sigma_n) \neq 0$ következik, ellentmondva (a)-nak, és ezzel készen vagyunk. □

A szimmetrikus polinomok vizsgálatával meglepő összefüggéseket találhatunk a gyökkifejezések között. Egy ilyet (a kíváncsiság felkeltése érdekében) bizonyítás nélkül közlünk:

6.38. Példa. *$A = B$, ahol*

$$A = \sqrt{5} + \sqrt{22 + 2\sqrt{5}}$$

és

$$B = \sqrt{11 + 2\sqrt{29}} + \sqrt{16 - 2\sqrt{29} + 2\sqrt{55 - 10\sqrt{29}}}.$$

6.7. Oszthatóság polinomgyűrűkben

Ebben az alfejezetben szisztematikusan megvizsgáljuk a polinomgyűrűk oszthatósági relációit. Egy rövid emlékeztetővel kezdünk.

Emlékeztető

Legyen R tetszőleges gyűrű (nem feltétlenül polinomgyűrű), és legyen $c \in R$ olyan gyűrűelem, amely 0-tól és R egységeitől is különbözik. A 3.30 Definíció szerint

- c prím, ha $c \mid ab \Rightarrow c \mid a$ vagy $c \mid b$ és
- c felbonthatatlan (irreducibilis), ha $c = ab \Rightarrow a$ egység vagy b egység.

Továbbá, ha R kommutatív, egységelemes, nullosztómentes gyűrű, akkor a 6.5 Tétel szerint $R[x]$ is kommutatív, egységelemes, nullosztómentes gyűrű. Ezért a 2.12 Tétel alkalmazható $R[x]$-re, és azt kapjuk, hogy az $R[x]$-beli oszthatóság is reflexív, „majdnem antiszimmetrikus", tranzitív reláció, mely megörződik lineáris kombinációra, azaz igazak az alábbiak:

1. $\forall f \in R[x] : f \mid f$;
2. $\forall f, g \in R[x] : (f \mid g,\ g \mid f) \Leftrightarrow \exists e_1, e_2$ egységek: $f = e_1 g$, $g = e_2 f$;
3. $\forall f, g, h \in R[x] : (f \mid g,\ g \mid h) \Rightarrow f \mid h$ és
4. $\forall f, g, h, u, v \in R[x] : (f \mid g,\ f \mid h) \Rightarrow f \mid (ug + vh)$.

Egységek polinomgyűrűben

6.39. Tétel. *Legyen R kommutatív, egységelemes, nullosztómentes gyűrű. Ekkor:*

$$f \in R[x] \text{ egység } R[x]\text{-ben} \Leftrightarrow f \in R \text{ és } f \text{ egység } R\text{-ben.}$$

Bizonyítás. ($\Rightarrow$:) Tegyük fel, hogy $g \in R[x]$ egység, azaz $\forall h \in R[x] : g \mid h$. Emiatt (mivel $1 \in R[x]$), $g \mid 1$. Speciálisan $\deg(g) = 0$, azaz $g \in R$. Továbbá, 1 egység R-ben ezért 1 osztója R összes elemének. Tehát az R-beli oszthatóság tranzitivitása miatt g is osztója R minden elemének, vagyis g egység R-ben.

($\Leftarrow$) Tegyük fel, hogy $g \in R$ egység R-ben. Ekkor $g \mid 1$ R-ben és – mivel 1 egység $R[x]$-ben is, ezért – minden $h \in R[x] :\ 1 \mid h$. Ezért (most az $R[x]$-beli) oszthatóság tranzitivitása miatt g osztója $R[x]$ összes elemének, vagyis g egység $R[x]$-ben. □

Kitüntetett közös osztó

6.40. Definíció. *Legyen $f, g \in R[x]$, $h \in R[x]$ kitüntetett közös osztója f-nek és g-nek, ha:*

1. *$h \mid f$ és $h \mid g$;*
2. *$\forall r \in R[x] : r \mid f$ és $r \mid g \Rightarrow r \mid h$.*

6.41. Tétel. *Legyen R kommutatív, egységelemes, nullosztómentes gyűrű, $f, g \in R[x]$. Ekkor f és g kitüntetett közös osztója egységszeres erejéig egyértelmű.*

Bizonyítás. Tegyük fel, hogy h_1, h_2 kitüntetett közös osztói f-nek és g-nek. Ekkor $h_1 \mid h_2$ és $h_2 \mid h_1$, vagyis h_1 és h_2 kölcsönösen osztják egymást, emiatt (az előző Emlékeztető 2. pontja szerint) egymás egységszeresei. □

Test feletti polinomok

6.42. Tétel. *Ha K test, $f, g \in K[x]$ ($f \neq 0$ vagy $g \neq 0$), akkor egységszeres erejéig egyértelműen létezik f és g legnagyobb közös osztója (jelölés: (f, g)).*

Bizonyítás. **Egyértelműség**: A 6.41 Tétel szerint ez nemcsak testekben, hanem kommutatív, egységelemes, nullosztómentes gyűrűkben is igaz.

Létezés: A 6.9 Tétel miatt testekben minden (nemnulla) polinommal lehet maradékosan osztani. $K[x]$-ben a maradékos osztás elvégzésekor a maradék-polinom foka szigorúan kisebb, mint az osztó polinom foka. Ezért ($\mathbb{N}$ jólrendezettsége miatt) $K[x]$-ben az euklideszi algoritmus véges sok lépésben befejeződik. Ugyanúgy, ahogy azt a 3.13 Tétel bizonyításában láttuk, az euklideszi algoritmus során képződő utolsó nem nulla maradék kitüntetett közös osztó. □

6.43. Megjegyzés. A fokszámok összevetéséből azonnal adódik, hogy (mindegyik) kitüntetett közös osztó egyben maximális fokú közös osztó. Ezért néha polinomgyűrűk esetében is használjuk a legnagyobb közös osztó kifejezést (amin legnagyobb fokú közös osztót értünk).

Továbbá, a $\mathbb{Z}$-re megismert kiterjesztett euklideszi algoritmus és a rá vonatkozó tétel (3.14 Tétel) – szó szerint megismételhető bizonyítással – érvényben marad $K[x]$-re is:

$$\forall f, g \in K[x]\ ((f, g) \neq 0)\ \exists u, v \in K[x] : (f, g) = uf + vg.$$

Felbonthatatlan és prím polinomok

6.44. Tétel. *Legyen K test, $f \in K[x]$ felbonthatatlan (irreducibilis) $\iff$ f prím.*

Bizonyítás. $\Leftarrow$: Tegyük fel, hogy f prím, és $f = g \cdot h$. Azt kell belátnunk, hogy g vagy h valamelyik egység $K[x]$-ben. Speciálisan, $f \neq 0$, ezért $g \neq 0$ és $h \neq 0$. Egyrészt, mivel f prím, ezért $f \mid g$ vagy $f \mid h$; mondjuk $f \mid g$ (az $f \mid h$ eset teljesen szimmetrikusan kezelhető).

Másrészt, a bizonyítás első mondata szerint $g \mid f$. Tehát f és g kölcsönösen osztják egymást, emiatt egymás egységszeresei; speciálisan, a 6.39 Tétel miatt $deg(f) = deg(g)$. Mivel $f = g \cdot h$, ez csak úgy lehet, ha $deg(h) = 0$, azaz $h \in K$. Mivel azonban K test, ezért h egység K-ban, így a 6.39 Tétel miatt $K[x]$-ben is.

$\Rightarrow$:Tegyük fel: $f \in K[x]$ felbonthatatlan és $f \mid g \cdot h$. Most azt kell belátnunk, hogy $f \mid g$ vagy $f \mid h$. De $(f,g) \mid f$ ezért $\exists u \in K[x] : f = u \cdot (f,g)$. Mivel f felbonthatatlan, ezért u egység vagy (f,g) egység.
1. eset: u egység. Ekkor f és (f,g) egymás egységszeresei, ezért $f \mid (f,g)$. Mivel $(f,g) \mid g$, ezért az oszthatóság tranzitivitása miatt $f \mid g$.
2. eset: (f,g) egység. Ekkor $\exists v, w \in K[x] : (f,g) = 1 = v \cdot f + w \cdot g$. Ezért

$$(*) \qquad h = 1 \cdot h = (v \cdot f + w \cdot g) \cdot h = v \cdot f \cdot h + w \cdot g \cdot h.$$

Az előző sor legvégén lévő összegben nyilván $f \mid (v \cdot f \cdot h)$. Továbbá a bekezdés elején feltettük, hogy $f \mid g \cdot h$, ezért $f \mid w \cdot g \cdot h$. Tehát f osztója $(*)$ jobboldalának, emiatt h-nak is. $\square$

Számelmélet alaptétele test feletti polinomokra

6.45. Tétel. *Legyen K test. A számelmélet alaptétele érvényes $K[x]$-re, azaz ha $f \in K[x]$, $f \neq 0$ és f nem egység, akkor:*
1. *Van felbonthatatlan $g_1, \ldots, g_m \in K[x] : f = g_1 \cdots g_m$;*
2. *Ez a felbontás sorrendtől és egységszorzóktól eltekintve egyértelmű.*

Bizonyítás. Indukció $n = \deg(f)$-re.
1. Alaplépés: $n = 1$, azaz $deg(f) = 1$. Ebben az esetben f-et egy elsőfokú és egy nulladfokú polinom szorzatára lehet csak felbontani. De minden ilyen felbontásban a nulladfokú tényező egység, mert K test. Ezért f irreducibilis, és önmagának egy 1-tényezős felbontása. Továbbá az előző gondolatmenet azt is adja, hogy f minden felbontásában valamelyik tényező egység, ezért a felbontás egység-szorzó erejéig egyértelmű.

2. Indukciós lépés: TFH: állításunk igaz minden n-nél kisebb, vagy egyenlő fokú polinomra, és THF $deg(f) = n + 1$.
1. eset: f felbonthatatlan. Ekkor, az $n = 1$ esethez teljesen hasonlóan, f sajátmagának egy 1-tényezős felbontása, és ez egységszorzó erejéig egyértelmű is.
2. eset: f nem irreducibilis, azaz van $g, h \in K[x]$:

$$f = g \cdot h, \qquad \deg(g), \deg(h) < \deg(f).$$

Az indukciós feltevés miatt g és h felbontható: vannak $g_1, \ldots g_u, g_{u+1}, \ldots g_{u+v}$ irreducibilis polinomok, melyekre:

$g = g_1 \cdots g_u$ és
$h = g_{u+1} \cdots g_{u+v}$.

De ekkor

$$f = g \cdot h = g_1 \cdots g_u \cdot g_{u+1} \cdots g_{u+v}$$

f egy felbontása irreducibilis tényezőkre. Ezzel a felbontás létezését beláttuk. Kell még (2), azaz az egyértelműség. TFH $f = g_1 \cdots g_p = h_1 \cdots h_m$ az f két felbontása irreducibilis tényezők szorzatára. Kell: $p = m$ és $\forall i\ \exists j : g_i$ és h_j egymás egységszeresei. Világos, hogy $g_1 \mid h_1 \cdots h_m$. Mivel g_1 irreducibilis, ezért a 6.44 Tétel miatt prím is. Emiatt valamelyik k-ra $g_1 \mid h_k$. De h_k irreducibilis, ezért g_1 és h_k egymás egységszeresei. Osszunk g_1-el:

$$g_2 \cdots g_n = h_1 \cdots h_{j-1} \cdot e \cdot h_{j+1} \cdots h_m,$$

ahol e valamelyik egység, és a két oldalon álló polinom(ok) közös foka kisebb, mint $deg(f)$. Ezért az indukciós feltevés miatt ez a felbontás sorrendtől és egységszorzóktól eltekintve egyértelmű. Ezért ugyanez f felbontásaira is igaz. □

6.46. Tétel. *Legyen K test.*

1. *Ha $f \in K[x]$, $\deg(f) = 1$, akkor f irreducibilis;*
2. *Ha $f \in K[x]$, $\deg(f) = 2$ vagy $\deg(f) = 3$, akkor (f irreducibilis $\iff$ nincs gyöke $K[x]$ − ben).*

Bizonyítás. (1) Bizonyítását lényegében elvégeztük már a 6.45 Tétel bizonyításának alaplépésében: THF $f \in K[x]$, $\deg(f) = 1$ és TFH: $f = g \cdot h$. Azt kell belátnunk, hogy g egység vagy h egység. Mivel $\deg(f) = \deg(g) + \deg(h) = 1$, ezért a fokszámok összevetéséből: $\deg(g) = 0$ vagy $\deg(h) = 0 \Rightarrow$ valamelyik 0-tól különböző testelem $K[x]$-ben, de K test $\Rightarrow g$ vagy h egység.

(2) $\Rightarrow$: TFH $f \in K[x]$, $2 \leq \deg(f) \leq 3$, és $f(c) = 0$, $c \in K$. Ekkor a 6.20 Tétel szerint f osztható $x-c$-vel, vagyis $\exists h : f = (x-c) \cdot h$, ekkor $\deg(h) = \deg(f) - 1 \geq 1$. Ezért $(x - c)$ nem egység és h nem egység, vagyis f nem irreducibilis. Mellesleg itt $deg(f)$-re a felső korlátot nem is használtuk: azt kaptuk, hogy ha egy legalább másodfokú polinomnak van gyöke K-ban, akkor a polinom nem irreducibilis K felett.

$\Leftarrow$: TFH: $2 \leq \deg(f) \leq 3$, f-nek nincs gyöke. TFH: f felbontható, azaz vannak olyan (nullától és) egységektől különböző $g, h \in K[x]$, melyekre $f = g \cdot h$. Ekkor g és h fokaira kizárólag a következő lehetőségek adódnak: $\deg(g) = 1$, $\deg(h) = 2$, vagy fordítva. Ha pl. g az elsőfokú, mondjuk $g = ax + b$, akkor könnyen adódik, hogy g-nek van gyöke K-ban: $ax + b = 0 \Leftrightarrow x = -\frac{b}{a}$; emiatt f-nek is van gyöke K-ban. □

Az előző tétel 3-nál nagyobbfokú polinomokra nem marad érvényben: pl. a negyedfokú $((x^2 + 1)^2$ ($\mathbb{R}$ feletti) polinomnak nincs gyöke $\mathbb{R}$-ben, de mégsem irreducubilis.

6.8. Formális deriváltak és többszörös gyökök

6.47. Definíció (Többszörös gyök)**.** *$f \in K[x]$-nek $c \in K$ n-szeres gyöke, ha*

$$(x - c)^n \mid f,$$

vagyis f kanonikus alakjában $(x - c)$ kitevője legalább n.

6.48. Definíció. *Legyen R kommutatív gyűrű, $a \in R$, $n \in \mathbb{N}$. Ekkor $na = a+\cdots+a$ (n-szer).*
Hangsúlyozzuk, hogy n nem feltétlenül gyűrűbeli szám, ezért $n \cdot a$ nem feltétlenül a gyűrűbeli szorzás.

Formális derivált

6.49. Definíció. *Legyen R kommutatív, egységelemes gyűrű. Ha*

$$f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in R[x],$$

*akkor f **formális deriváltja**:*

$$f' = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1.$$

6.50. Példa. *Legyen $f(x) = 2x^3 + 2x^2$.*

- *Ha $f \in \mathbb{C}[x]$ akkor f' a szokásos: $f' = 6x^2 + 4x$.*
- *Ha $f \in \mathbb{Z}_3[x]$ akkor $f' = 0x^2 + x = x$ (mert $6 \equiv 0 \pmod 3$).*

A formális deriváltról megjegyezzük még a következőket. A fenti értelemben csak polinomoknak definiáltuk a formális deriváltját, másfajta függvényeknek nem. Itt a „formális" jelző arra utal, hogy a definícióhoz csak olyan eszközöket használtunk, melyek minden gyűrűben rendelkezésre állnak; nem használtunk semmilyen távolságfogalmat, nem említettünk különbségi hányadosokat, mégkevésbé ezek határértékeit (mert egy absztrakt gyűrűben legalábbis nem világos, hogy hogyan kéne a határétrék fogalmát értreni, ha nincs beépítve, nincs - a gyűrűműveletek mellett – külön megadva valamilyen távolság-fogalom). Definíciónkat az motiválja, hogy a valós test feletti polinomok esetében az imént definiált formális derivált egybeesik a polinomfüggvények, mint valós-valós függvények szokásos deriváltfüggvényeivel. Ugyanakkor vannak olyan deriválható valós-valós függvények, melyek nem polinom-függvények, ebből a szempontból a mi definíciónk kevésbé általános. Viszont a formális deriváltakat nem csak $\mathbb{R}$ feletti polinomok esetében vizsgálhatjuk, hanem tetszőleges (kommutatív, egységelemes) gyűrűk felett, ilyen értelemben a formális derivált általánosabb, mint a bevezető analízis derivált-fogalma. Valós polinomok esetében régóta ismertek különböző összefüggések polinomok többszörös gyökei és deriváltjaik gyökei között. Alább ezekből az összefüggésekből mutatunk be néhányat, melyek tetszőleges (kommutatív) test felett érvényben maradnak. Végig tisztán algebrai keretek között maradunk; következő célunk vizsgálataink technikai előkészítése (azzal, hogy a szokásos deriváltakra vonatkozó néhány számolási szabályra algebrai - emiatt minden kommutatív test felett érvényes - bizonyítást adunk).

6.51. Tétel (Deriválási szabályok)**.** *Legyen K test, $f, g \in K[x]$, $c \in K$. Ekkor:*

1. $(c)' = 0$;
2. $(cf)' = c(f')$;

3. $(f+g)' = f' + g'$;
4. $(fg)' = f'g + fg'$;
5. $((x-c)^m)' = m(x-c)^{m-1}$.

Bizonyítás. (1): $c = c \cdot x^0 \Rightarrow (c)' = 0 \cdot c \cdot x^{-1} = 0$.

(2) egyszerű számolással adódik:

$$(cf)' = \left(\sum_{k=0}^{n} ca_k x^k\right)' = \sum_{k=1}^{n} kca_k x^{k-1} = c\sum_{k=1}^{n} ka_k x^{k-1} = c(f').$$

(3) igazolásához legyen $f = a_0 + a_1 x + \cdots + a_n x^n$, $g = b_0 + b_1 x + \cdots + b_m x^m$ és legyen $N = \max(n, m)$. Ezekre

$$(f+g)' = \left(\sum_{k=0}^{N} (a_k + b_k)x^k\right)' = \sum_{k=1}^{N} k(a_k+b_k)x^{k-1} = \sum_{k=1}^{N} ka_k x^{k-1} + \sum_{k=1}^{N} kb_k x^{k-1} = f'+g'.$$

(4) igazolásához legyenek

$$f = \sum_{i=0}^{n} a_i x^i, \quad g = \sum_{j=0}^{m} b_j x^j.$$

A két polinom szorzata:

$$f \cdot g = \sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j x^{i+j}.$$

A szorzat deriváltja (azaz (4) baloldala):

$$(f \cdot g)' = \left(\sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j x^{i+j}\right)' = \sum_{i=0}^{n}\sum_{j=0}^{m} (i+j) a_i b_j x^{i+j-1}.$$

Számítsuk ki a jobb oldal tagjait külön-külön.

A deriváltak:

$$f' = \sum_{i=1}^{n} ia_i x^{i-1}, \quad g' = \sum_{j=1}^{m} jb_j x^{j-1}.$$

Ezért (4) jobboldalának első összeadandója

$$f' \cdot g = \left(\sum_{i=1}^{n} ia_i x^{i-1}\right)\left(\sum_{j=0}^{m} b_j x^j\right) = \sum_{i=1}^{n}\sum_{j=0}^{m} ia_i b_j x^{i+j-1}$$

és hasonlóan, második összeadandója:

$$f \cdot g' = \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=1}^{m} j b_j x^{j-1}\right) = \sum_{i=0}^{n}\sum_{j=1}^{m} j a_i b_j x^{i+j-1}.$$

E két tag összege (azaz (4) jobboldala):

$$\begin{aligned} f' \cdot g + f \cdot g' &= \sum_{i=1}^{n}\sum_{j=0}^{m} i a_i b_j x^{i+j-1} + \sum_{i=0}^{n}\sum_{j=1}^{m} j a_i b_j x^{i+j-1} \\ &= \sum_{i=0}^{n}\sum_{j=0}^{m} (i+j) a_i b_j x^{i+j-1}. \end{aligned}$$

Ezek alapján (4) baloldala és jobboldala megegyezik, és ezzel (4) bizonyításával készen vagyunk. (5)-höz először is vegyük észre, hogy

$$(*) \quad \binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n}{k} \cdot \frac{(n-1)!}{(k-1)!\,(n-k)!} = \frac{n}{k}\binom{n-1}{k-1}.$$

Ekkor:

$$\begin{aligned} ((x+c)^n)' &= \left(\sum_{k=0}^{n} \binom{n}{k} x^k c^{n-k}\right)' = \sum_{k=0}^{n} \binom{n}{k} k\, x^{k-1} c^{n-k} \\ &\overset{(*)}{=} \sum_{k=0}^{n} \frac{n}{k}\binom{n-1}{k-1} k\, x^{k-1} c^{n-k} \\ &= n \sum_{k=0}^{n} \binom{n-1}{k-1} x^{k-1} c^{n-k} \\ &= n \sum_{k=0}^{n} \binom{n-1}{k-1} x^{k-1} c^{(n-1)-(k-1)} \overset{\text{Binomiális Tétel}}{=} n(x+c)^{n-1}. \end{aligned}$$

□

Többszörös gyökök és gyökteszt

6.52. Tétel. *Legyen $f \in K[x]$, $c \in K$, $k \in \mathbb{N}$, $k \geq 2$. Ekkor ekvivalensek:*

1. *c f-nek k-szoros gyöke;*
2. *c f-nek és f'-nek $(k-1)$-szeres gyöke.*

Bizonyítás. (1) ⇒ (2): TFH c f-nek k-szoros gyöke. Ekkor $\exists g \in K[x] : f = (x-c)^k g$. Ezért

$$f' = k(x-c)^{k-1} g + (x-c)^k g' = (x-c)^{k-1}(kg + (x-c)g'),$$

ez mutatja: c f'-nek $(k-1)$-szeres gyöke (nyilván f-nek is).

(2) $\Rightarrow$ (1): TFH c f-nek és f'-nek $(k-1)$-szeres gyöke. Speciálisan, $\exists g \in K[x]$: $f = (x-c)^{k-1}g$. Mindkét oldalt deriválva:

$$(*) \qquad f' = (k-1)(x-c)^{k-2}g + (x-c)^{k-1}g'.$$

Továbbá, (2) szerint c f'-nek $(k-1)$-szeres gyöke, ezt $(*)$ jobboldalával összevetve $(x-c) \mid g$. De ekkor $(x-c)^k \mid f$, vagyis c f-nek k-szoros gyöke. □

6.53. Megjegyzés.
1. Gyökkeresésnél érdemes redukálni a fokszámot (mert legfeljebb 4-edfokú polinomokra van megoldóképlet).
2. Az $\frac{f}{(f,f')}$ polinom gyökei egyszeresek, azonosak f gyökeivel, ez hasznos fokszám-redukció lehet.
3. A gyökök pontos előállítása helyett próbálkozhatunk közelítő módszerekkel.

6.54. Tétel (Rolle-féle gyökteszt)**.** *Ha $f \in \mathbb{Z}[x]$, $f = a_0 + a_1x + \cdots + a_nx^n$, $p, q \in \mathbb{Z}$ relatív prímek, és $f\left(\frac{p}{q}\right) = 0$, akkor $q \mid a_n$ és $p \mid a_0$.*

Szavakban: ha egy egész együtthatós polinomnak a tovább már nem egyszerűsíthető $\frac{p}{q}$ racionális szám gyöke, akkor a p számláló osztója a polinom konstans tagjának, és a q nevező osztója a főeggyüthatónak.

A tétel segítségével meg tudjuk keresni egy egész együtthós polinom összes racionális gyökeit: soroljuk fel a főegyüttható és a konstans tag osztóinak hányadosait, így véges sok racionális számot kapunk. Ezeket egyenként kipróbálva (visszahelyettesítve) megtalálhatjuk polinomunk összes racionális gyökét. A főegyüttható és a konstans tag osztóinak hányadosaiból álló lista tartalmazhat olyan racionális számokat is, melyek nem gyökei a polinomunknak (ennek így kell lennie, ha a főegyütthatónak, vagy a konstans tagnak sok osztója van a polinom fokához képest). De a megvizsgálandó racionális számok halmaza véges, ez elméleti szempontból azért érdekes, mert a racionális számok végtelen halmazát le tudtuk szűkíteni egy véges halmazra, mely minden racionális gyököt tartalmaz. Ennek az a gyakorlati jelentősége, hogy a racionális gyökök megkeresésére javasolt módszerünk véges sok lépésben befejeződik.

Bizonyítás. $f\left(\frac{p}{q}\right) = 0$ miatt

$$a_0 + a_1\frac{p}{q} + \cdots + a_n\frac{p^n}{q^n} = 0.$$

Szorozzuk be mindkét oldalt q^n-el:

$$q^na_0 + q^{n-1}a_1p + \cdots + a_np^n = 0.$$

Mivel a jobboldali 0 osztható p-vel és q-val is, ezért a baloldal is osztható p-vel és q-val. Mivel $q \mid q^na_0 + q^{n-1}a_1p + \cdots + a_{n-1}p^{n-1}q$, ezért $q \mid a_np^n$. De $(q, p^n) = 1$, emiatt $q \mid a_n$.

Hasonlóan, $p \mid q^na_0$, amiből $p \mid a_0$ adódik. □

6.55. Példa (Polinomok gyűrűkben)**.** *Gyűrűkben nem feltétlenül igaz az egyértelmű szorzatra bontás. Ennek illusztrálására mutatunk néhány perverz példát.*

- $\mathbb{Z}_8[x]$*-ben:* $x^2 - 1 = (x+1)(x-1) = (x+3)(x-3)$.

A 6.46 Tételben láttuk, hogy testek felett minden elsőfokú polinom irreducibilis. Gyűrűk felett ez sem feltétlenül marad érvényben:

- $\mathbb{Z}_6[x]$*-ben:* $x = (2x+3)(3x+2)$.

Sőt $\mathbb{Z}$*-ben 6 nem irreducibilis:* $6 = 2\cdot 3$*, bár 6* $\mathbb{Q}$*-ban sem irreducibilis (*$\mathbb{Q}$*-ban egység).*

6.9. Irreducibilis polinomok $\mathbb{C}[x]$-ben és $\mathbb{R}[x]$-ben

Irreducibilis polinomok $\mathbb{C}[x]$-ben

6.56. Tétel. $f \in \mathbb{C}[x]$ *irreducibilis* $\Leftrightarrow \deg(f) = 1$.

Bizonyítás. $\Leftarrow$: $\mathbb{C}$ test ezért a 6.46 Tétel szerint $\mathbb{C}$ felett minden elsőfokú polinom irreducibilis.
$\Rightarrow$: TFH $f \in \mathbb{C}[x]$ irreducibilis: ha $\deg(f) \leq 0$ akkor $f = 0$ vagy f egység $\mathbb{C}[x]$-ben, ezért nem irreducibilis. Ha $\deg(f) \geq 2$, akkor az algebra alaptétele miatt $\deg(f)$ db gyöke van: $c_1, \ldots, c_{\deg(f)} \in \mathbb{C}$, és a megfelelő gyöktényezők f-ből kiemelhetők:

$$f = a_n(x - c_1)(x - c_2)\cdots(x - c_{\deg(f)}).$$

Ezért f nem irreducibilis. Azt kaptuk, hogy ha f irreducibilis, akkor foka kizárólag 1 lehet. □

Irreducibilis polinomok $\mathbb{R}[x]$-ben

6.57. Tétel. *Ha* $f = \sum a_k x^k \in \mathbb{R}[x]$*,* $c \in \mathbb{C}$*,* $f(c) = 0$*, akkor* $f(\overline{c}) = 0$ *(azaz, ha egy valós együtthatós* f *polinomnak a* c *komplex szám gyöke, akkor* $\overline{c}$ *is gyöke* f*-nek).*

Bizonyítás. Mivel f mindegyik a_k együtthatója valós, ezért minden k-ra $\overline{a_k} = a_k$. Ezt azzal kombinálva, hogy (az 5.9 Tétel szerint) a konjugálás automorfizmus:

$$0 = \overline{0} = \overline{f(c)} = \overline{\sum a_k c^k} = \sum \overline{a_k}\overline{c}^k = \sum a_k \overline{c}^k = f(\overline{c}),$$

vagyis $\overline{c}$ valóban gyöke f-nek. □

6.58. Tétel. *Ha* $c \in \mathbb{C}$*, akkor* $(x - c)(x - \overline{c})$ *valós együtthatós.*

Bizonyítás.

$$(x - c)(x - \overline{c}) = x^2 - (c + \overline{c})x + c\overline{c} = x^2 - 2\Re(c)x + |c|^2.$$

Mivel $\Re(c) \in \mathbb{R}$, $|c|^2 \in \mathbb{R}$, ezért készen vagyunk. □

6.59. Tétel. *Ha* $f \in \mathbb{R}[x]$, $deg(f) \geq 1$, *akkor van* $g_1, \ldots, g_k \in \mathbb{R}[x]$, $\forall i$:

- $\deg(g_i) = 1$ *vagy*
- $\deg(g_i) = 2$ *és* g_i*-nek nincs gyöke* $\mathbb{R}[x]$*-ben.*

Szavakban: $\mathbb{R}$ *felett minden legalább elsőfokú polinom olyan szorzattá bontható, melyben minden tényező vagy elsőfokú, vagy olyan másodfokú, melynek nincs valós gyöke.*

Bizonyítás. Tudjuk, hogy f-nek annyi komplex gyöke van, mint a foka. Legyenek ezek: $d_1, \ldots, d_n \in \mathbb{R}$, $\beta_1, \overline{\beta_1}, \ldots, \beta_m, \overline{\beta_m} \in \mathbb{C} \setminus \mathbb{R}$. Legyen c az f főegyütthatója. Ekkor:

$$f = c(x - d_1) \cdots (x - d_n)(x - \beta_1)(x - \overline{\beta_1}) \cdots (x - \beta_m)(x - \overline{\beta_m}).$$

Minden j-re legyen $g_j = (x - \beta_j)(x - \overline{\beta_j})$. Ekkor $\deg(g_j) = 2$, a 6.58 Tétel miatt g_j valós együtthatós, és (nyilván) nincs gyöke $\mathbb{R}$-ben. □

6.60. Tétel. $f \in \mathbb{R}[x]$ *ekvivalensek:*

1. f *irreducibilis* $\mathbb{R}[x]$*-ben;*
2. $\deg(f) = 1$ *vagy (*$\deg(f) = 2$ *és* f*-nek nincs valós gyöke* $\mathbb{R}$*-ben).*

Bizonyítás. (1) ⇒ (2): A 6.59 Tételben szereplő felbontásból:

$$f = p_1 \cdots p_n q_1 \cdots q_m.$$

Ezért, ha f irreducibilis, akkor $1 \leq \deg(f) \leq 2$. Ha $\deg(f) = 2$, akkor a 6.46 Tétel szerint nincs gyöke $\mathbb{R}$-ben.

(2) ⇒ (1): $\mathbb{R}$ test, ezért (ismét a 6.46 Tétel szerint) $\mathbb{R}$ felett minden elsőfokú polinom irreducibilis és ha $\deg(f) = 2$ és f-nek nincs gyöke $\mathbb{R}$-ben, akkor f irreducibilis. □

6.10. Irreducibilis polinomok $\mathbb{Z}[x]$-ben és $\mathbb{Q}[x]$-ben

6.61. Definíció (Primitív polinom)**.** $f = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ ***primitív****, ha* $LNKO(a_0, a_1, \ldots, a_n) = 1$.

6.62. Megjegyzés. *Ha* f *irreducibilis, akkor primitív, ugyanis ha* c $(\neq \pm 1)$ *valódi közös osztója lenne* $a_0, \ldots, a_n$*-nek, akkor* $f = c \sum \frac{a_k}{c} x^k$ *az* f *egy olyan felbontása lenne, melyben a jobboldal egyik tényezője sem egység (hiszen* c *nem egység* $\mathbb{Z}$*-ben).*

6.63. Példa. $\mathbb{Z}[x]$*-ben:* $2x^2 + 4 = 2(x^2 + 2)$.
$\mathbb{Z}[x]$*-ben ez olyan felbontás, melyben egyik tényező sem egység (az ilyen felbontásokat röviden valódi felbontásoknak, vagy nemtriviális felbontásoknakis nevezik).*

6.64. Definíció (Redukció modulo n)**.** *Legyen* $n \in \mathbb{N}$, $\mu_n : \mathbb{Z} \to \mathbb{Z}_n$,

$$\mu_n(x) = x \pmod{n},$$

és legyen $\varphi_n : \mathbb{Z}[x] \to \mathbb{Z}_n[x]$,

$$\varphi_n\Big(\sum a_k x^k\Big) = \sum \mu_n(a_k) x^k.$$

Szavakban: tetszőleges $f \in \mathbb{Z}[x]$*-re* $\varphi_n(f)$*-et úgy kapjuk, hogy* f *együtthatóit külön-külön* $(\mathrm{mod}\ n)$ *vesszük.*

6.65. Példa. *Ha* $f = 3x^2 - 7x + 4$ *és* $n = 3$*, akkor*

$$\varphi_3(f) = 0x^2 - 1x + 1.$$

6.66. Tétel. *Előző jelölésekkel, tetszőleges* $n \geq 1$*-re:*

$$\varphi_n : \mathbb{Z}[x] \to \mathbb{Z}_n[x] \quad \textit{homomorfizmus.}$$

Bizonyítás. Legyen $f, g \in \mathbb{Z}[x]$, $f = \sum a_k x^k$, $g = \sum b_k x^k$

$$\begin{aligned}\varphi_n(f+g) &= \varphi_n\left(\sum (a_k + b_k)x^k\right) = \sum \mu_n(a_k + b_k)x^k \\ &= \sum (\mu_n(a_k) + \mu_n(b_k))x^k = \sum \mu_n(a_k)x^k + \sum \mu_n(b_k)x^k \\ &= \varphi_n(f) + \varphi_n(g).\end{aligned}$$

Hasonlóan,

$$\mu_n(f \cdot g) = \varphi_n\left(\sum_k \left(\sum_{j=0}^{k} a_i b_j\right) x^k\right) =$$

$$\sum_k \mu_n\left(\sum_{j=0}^{k} a_j b_{k-j}\right) x^k = \sum_k \sum_{j=0}^{k} \mu(a_j) \cdot \mu(b_{k-j}) x^k = \varphi(f) \cdot \varphi(g).$$

□

6.67. Tétel. *Ha* $f \in \mathbb{Q}[x]$, $f \neq 0$*, akkor* $\exists! r \in \mathbb{Q}^+$, $\exists! g \in \mathbb{Z}[x]$*:* $f = r \cdot g$*,* g *primitív. Ha* $f \in \mathbb{Z}[x]$*, akkor* $r \in \mathbb{Z}$*.*

Szavakban: minden nem-nulla racionális együtthatós f *polinom előállítható egy* r *pozitív racionális szám és egy* $g \in \mathbb{Z}[x]$ *primitív polinom szorzataként, az előállítás egyértelmű, és ha* f *egész-együtthatós, akkor* r *egész szám.*

Bizonyítás. Az ötlet a következő: f-et felszorozzuk együtthatói nevezőinek legkisebb közös többszörösével, majd ebből kiemeljük az együtthatók legnagyobb közös osztóját – amit kapunk, primitív polinom lesz. Részletesebben, legyen B az f együtthatói nevezőinek legkisebb közös többszöröse, és legyen A az (egész-együtthatós) $B \cdot f$ polinom együtthatóinak legnagyobb közös osztója. Végül legyen $r = \frac{A}{B} > 0$. Ezekkel:

$$f = \frac{1}{B} \cdot B \cdot f = \frac{A}{B} \cdot \frac{B}{A} f = r \cdot \frac{B}{A} f.$$

Világos, hogy A és B választása miatt $\frac{B}{A}f$ egész együtthatós polinom, sőt primitív polinom (mert $B \cdot f$ együtthatóinak legnagyobb közös osztója épp A volt). Továbbá, ha f már eredetileg is egész együtthatós volt, akkor $B = 1$, és emiatt $r \in \mathbb{Z}$ is teljesül.

Be kell még látnunk a felbontás egyértelműségét. Ehhez TFH: $f = r_1 g_1$ és $f = r_2 g_2$ az f két felbontása (ahol tehát $r_1, r_2 \in \mathbb{Q}^+$ és $g_1, g_2 \in \mathbb{Z}[x]$ primitív polinomok). Ezekből

$$g_1 = \frac{1}{r_1} f = \frac{r_2}{r_1} g_2.$$

Itt $\frac{r_2}{r_1}$ egy pozitív racionális szám, ezért vannak olyan $a, b \in \mathbb{N}$ számok, hogy $\frac{r_2}{r_1} = \frac{a}{b}$, és feltehető, hogy a jobboldali tört tovább már nem egyszerűsíthető, azaz a és b relatív prímek. Tehát

$$(*) \qquad g_1 = \frac{a}{b} g_2.$$

Mivel g_1 együtthatói egész számok, ezért $\frac{a}{b} g_2$ együtthatói is egész számok, ami csak úgy lehet, ha b (közös) osztója g_2 mindegyik együtthatójának. De g_2 primitív polinom, ezért együtthatói legnagyobb közös osztója 1 és emiatt $b = 1$ (hiszen $b > 0$). Ekkor viszont $(*)$ miatt a közös osztója g_1 együtthatóinak, amiből g_1 primitívsége miatt $a = 1$ következik. Ezek szerint $a = b = 1$, tehát $(*)$ alapján $g_1 = g_2$ (és emiatt $r_1 = r_2$). □

Gauss-lemma és irreducibilitás

6.68. Tétel (Gauss-lemma)**.** *Ha $f, g \in \mathbb{Z}[x]$ primitív polinomok, akkor $f \cdot g$ is primitív.*

Bizonyítás. Indirekt tegyük fel, hogy $f \cdot g$ nem primitív. Ekkor $\exists p$ prímszám: p osztja $f \cdot g$ összes együtthatóját. Tekintsük a 6.64 Definícióban szereplő $\varphi_p : \mathbb{Z}[x] \to \mathbb{Z}_p[x]$ modulo p redukáló függvényt; ez a 6.66 Tétel miatt homomorfizmus.

Ekkor egyrészt p választása miatt $\varphi_p(f \cdot g) = 0$. Másrészt viszont

$$(*) \qquad \varphi_p(f \cdot g) = \varphi_p(f) \cdot \varphi_p(g).$$

Mivel f és g primitívek, ezért $\varphi_p(f) \neq 0$ és $\varphi_p(g) \neq 0$. Továbbá $\mathbb{Z}_p$ test, tehát nullosztómentes, ezért $(*)$ jobboldala nem lehet nulla. Ez ellentmond e bekezdés első mondatának. □

A $\mathbb{Z}[x]$-beli és $\mathbb{Q}[x]$-beli irreducibilitás közti kapcsolat

6.69. Tétel. *Legyen $f \in \mathbb{Z}[x]$ primitív polinom. Ekkor:*

$$f \text{ irreducibilis } \mathbb{Z}[x]\text{-ben} \Leftrightarrow f \text{ irreducibilis } \mathbb{Q}[x]\text{-ben.}$$

Bizonyítás. Nyilván elég azt belátni, hogy

$$f \text{ felbontható } \mathbb{Z}[x]\text{-ben} \Leftrightarrow f \text{ felbontható } \mathbb{Q}[x]\text{-ben.}$$

$\Rightarrow$: Ha f felbontható $\mathbb{Z}[x]$-ben, akkor ugyanez a felbontás mutatja, hogy felbontható $\mathbb{Q}[x]$-ben is.

$\Leftarrow$: Tegyük fel, hogy f felbontható $\mathbb{Q}[x]$-ben: $\exists g_1, g_2 \in \mathbb{Q}[x] : f = g_1 \cdot g_2$, ahol $0 < \deg(g_1), \deg(g_2) < \deg(f)$. A 6.67 Tétel szerint g_1 és g_2 előállítható primitív polinomok segítségével:

$$\begin{aligned} g_1 &= r_1 \cdot g_1^*, \quad r_1 \in \mathbb{Q}^+, \quad g_1^* \in \mathbb{Z}[x] \text{ primitív,} \\ g_2 &= r_2 \cdot g_2^*, \quad r_2 \in \mathbb{Q}^+, \quad g_2^* \in \mathbb{Z}[x] \text{ primitív.} \end{aligned}$$

Ezekkel $f = (r_1 r_2) \cdot (g_1^* g_2^*)$. De a Gauss-lemma (6.68 Tétel) miatt $g_1^* g_2^*$ primitív. Mivel $f = 1 \cdot f$ is primitív, ezért a 6.67 Tétel (egyértelműségre vonatkozó része) szerint $r_1 r_2 = 1$ és $f = g_1^* \cdot g_2^*$; ez f egy $\mathbb{Z}[x]$-beli felbontását adja. □

Karakterizációk

6.70. Tétel (Irreducibilis polinomok $\mathbb{Z}[x]$-ben)**.** *Egy $f \in \mathbb{Z}[x]$ polinom pontosan akkor irreducibilis, ha:*

1. $\deg(f) = 0$ *és f prímszám, vagy*
2. $\deg(f) \geq 1$*, f primitív és irreducibilis $\mathbb{Q}[x]$-ben.*

Bizonyítás. $\Rightarrow$:

- Ha $\deg(f) = 0$, akkor $f \in \mathbb{Z}$. Itt az irreducibilitás megegyezik a prímséggel.
- Ha $\deg(f) \geq 1$ és f nem primitív, akkor legyen c az f egyxüttható inak $LNKO$-ja. Ezzel $f = c \cdot \frac{1}{c} f$ ahol $c > 1$ és $\frac{1}{c} f$ primitív, tehát f felbontható.
- Ha $\deg(f) \geq 1$ és f primitív, akkor az előző tétel (6.69 Tétel alapján f $\mathbb{Q}[x]$-ben is irreducibilis.

$\Leftarrow$: ha $deg(f) = 0$ (azaz $f \in \mathbb{Z}$) prím, akkor f $\mathbb{Z}[x]$-ben is nullad fokúak szorzataként áll elő, de ilyen nemtriviális felbontás nem létezik. Végül, ellentmondást keresve TFH f primitív, $deg(f) \geq 1$ és irreducibilis $\mathbb{Q}[x]$-ben, de f-nek $\mathbb{Z}[x]$-ben van nemtrivi felbontása. Egy ilyen felbontás mindkét tényezője legalább elsőfokú, mert f primitív, de ekkor ez a $\mathbb{Z}[x]$-beli felbontás egyúttal $\mathbb{Q}[x]$-beli nemtrivi felbontás is lenne, ellentmondás. □

6.71. Tétel (Irreducibilis polinomok $\mathbb{Q}[x]$-ben)**.** *Egy $f \in \mathbb{Q}[x]$ polinom pontosan akkor irreducibilis, ha:*

$$\exists g \in \mathbb{Z}[x] \textit{ primitív, irreducibilis polinom és } \quad \exists r \in \mathbb{Q} : f = r \cdot g.$$

Bizonyítás. ($\Rightarrow$:) TFH $f \in \mathbb{Q}[x]$ irreducibilis $\mathbb{Q}[x]$-ben. A 6.67 Tétel szerint f felírható $f = r \cdot g$ alakban, ahol $r \in \mathbb{Q}$ és $g \in \mathbb{Z}[x]$ primitív. Mivel f irreducibilis

$\mathbb{Q}[x]$-ben, ezért g-nek is irreducibilisnek kell lennie $\mathbb{Q}[x]$-ben (mert r egység $\mathbb{Q}$-ban). Ezért a 6.69 Tétel miatt g irreducibilis $\mathbb{Z}[x]$-ben is.

($\Leftarrow$:) TFH g primitív, irreducibilis $\mathbb{Z}[x]$-ben és $r \in \mathbb{Q}$. Azt kell belátnunk, hogy $f = r \cdot g$ irreducibilis $\mathbb{Q}[x]$-ben. A 6.69 Tétel miatt g irreduciblis $\mathbb{Q}[x]$-ben is. Továbbá r egység $\mathbb{Q}$-ban, ezért f valóban irreducibilis $\mathbb{Q}[x]$-ben. □

Számelmélet alaptétele $\mathbb{Z}[x]$-ben

6.72. Tétel. *$\mathbb{Z}[x]$-ben is igaz a számelmélet alaptétele:*
Ha $f \in \mathbb{Z}[x]$, $\deg(f) \geq 1$, akkor $\exists! f = g_1 \cdots g_k \in \mathbb{Z}[x]$ irreducibilis faktorokra bontás.

Bizonyítás. **Létezés.** Az alapötlet a következő: először bontsuk fel f-et egy egész szám és egy primitív polinom szorzatára, a primitív tényezőt bontsuk irreducbilis tényezőkre $\mathbb{Q}$ felett, végül gondoljuk meg, hogy ebből hogyan kapunk $\mathbb{Z}$ feletti felbontást.

Részletesebben, a 6.67 Tétel szerint van olyan $r \in \mathbb{Q}$ és $f^* \in \mathbb{Z}[x]$ primitív polinom, hogy $f = r \cdot f^*$. De f egész-együtthatós, ezért a 6.67 Tétel szerint r is egész szám.

Mivel $\mathbb{Q}$ test, ezért (a 6.45 Tétel szerint) $\mathbb{Q}[x]$-ben igaz a Számelmélet Alaptétele. Tehát vannak $\mathbb{Q}[x]$-ben irreducibilis $h_1, \ldots, h_k \in \mathbb{Q}[x]$ polinomok, melyekre:

$$f = h_1 \cdots h_k.$$

A (6.67 Tétel szerint) minden h_j-hez ($j = 1, \ldots, k$) $\exists r_j \in \mathbb{Q}$, $\exists g_j \in \mathbb{Z}[x]$ primitív polinom:

$$h_j = r_j \cdot g_j.$$

Mivel mindegyik h_j irreducibils $\mathbb{Q}[x]$-ben és r_j egység $\mathbb{Q}$-ban, ezért mindegyik g_j is irreducibils $\mathbb{Q}[x]$-ben. De mindegyik g_j primitív is, ezért a 6.70 Tétel miatt mindegyik g_j irreducibilis $\mathbb{Z}[x]$-ben is. De ekkor:

$$f = h_1 \cdots h_k = (r_1 \cdot g_1) \cdots (r_k \cdot g_k) = (r_1 \cdots r_k) \cdot (g_1 \cdots g_k).$$

Továbbá, a Gauss-lemma (6.68 Tétel) szerint a $g_1 \cdots g_k$ szorzat egy primitív polinom. De $f = r \cdot f^*$ is f egy primitívvé bontása, ezért a 6.67 Tétel egyértelműségre vonatkozó része miatt $r = r_1 \cdots r_k$ egy egész szám. Ezt $\mathbb{Z}$-ben prímek szorzatára bonthatjuk:

$$r_1 \cdots r_k = p_1 \cdots p_\ell.$$

Ezekkel f felbontása $\mathbb{Z}[x]$-ben:

$$f = p_1 \cdots p_\ell \cdot g_1 \cdots g_k$$

ahol tehát

- $p_1, \ldots, p_\ell \in \mathbb{Z}$ prímek (0-adfokú irreducibilis polinomok);

- $g_1, \ldots, g_k \in \mathbb{Z}[x]$ (primitív) irreducibilis polinomok.

Egyértelműség.[3] Először tegyük fel, hogy $f \in \mathbb{Z}[x]$ primitív és f két irreducibilis tényezőkre bontása:

$$f = g_1 \cdots g_k = h_1 \cdots h_m.$$

Mivel f primitív, ezért mindegyik g_i és h_j legalább elsőfokú. Mivel mindegyik g_i és h_j irreducibilis $\mathbb{Z}[x]$-ben, ezért primitívek is. Ezért a 6.69 Tétel miatt mindegyik g_i és h_j irreducibilis $\mathbb{Q}[x]$-ben is.

De a 6.45 Tétel miatt a Számelmélet Alaptétele test feletti polinomgyűrűkben igaz, ezért f megadott két felbontása sorrendtől és $\mathbb{Q}$-beli egységszorzóktól eltekintve egyértelmű. Mivel azonban mindegyik g_i és h_j primitív, ezért (a 6.67 Tétel egyéretlműségre vonatkozó része miatt) a szóbajövő $\mathbb{Q}$-beli egységszorzók csak ± 1-ek lehetnek. Ezzel primitív f-ekre a felbontás $\mathbb{Z}[x]$-beli egyértelműségét beláttuk.

Ha $f \in \mathbb{Z}[x]$ tetszőleges (nem feltétlenül primitív) polinom, akkor a 6.67 Tétel szerint pontosan egy olyan r szám és primitív f^* van, hogy $f = r \cdot f^*$. Mivel f egész együtthatós, ezért r egész szám; ez $\mathbb{Z}$-ben egyértelműen bontható fel prímszámok szorzatára. Végül, az előző bekezdés szerint, a primitív f^* egyértelműen bontható fel irreducibilisek szorzatára $\mathbb{Z}[x]$-ben. □

Láttuk, hogy $\mathbb{Z}$-ben és testek feletti polinomgyűrűkben be lehet vezetni, és el lehet végezni a maradékos osztást. Erre alapozva a kiterjesztett euklideszi algoritmus segítségével igazolható a kitüntetett közös osztók létezése, és az, hogy a prím és irreducibilis elemek egybeesnek; a Számelmélet Alaptételének analogonjait mindig ezek segítségével igazoltuk.

Ugyanakkor a 6.8 példában láttuk, hogy $\mathbb{Z}[x]$-ben a maradékos osztás elvégzése nem lehetséges mindig, tehát az előző bekezdésben összefoglalt sémát nem tudjuk adaptálni $\mathbb{Z}[x]$-re. A 6.72 Tétel szerint a Számelmélet Alaptétele mégis érvényben marad $\mathbb{Z}[x]$-ben: bizonyításunk azon múlt, hogy a 6.45 Tétel értelmében test feletti polinomgyűrűkben (tehát $\mathbb{Q}[x]$-ben is) a Számelmélet Aleptétele igaz, és $\mathbb{Z}[x]$ és $\mathbb{Q}[x]$ valamilyen értelemben nagyon közel vannak egymáshoz. Végül pedig azt jegyezzük meg, hogy a Számelmélet Alaptételéből egyszerűen következik, hogy bármely két elemnek van kitüntetett közös osztója. Ezért kitüntetett közös osztók $\mathbb{Z}[x]$-ben is léteznek, noha a maradékos osztást nem lehet mindig elvégezni.

Kronecker kitalált egy viszonylag egyszerű (de nem túl hatékony) algoritmust, melynek bemenete $\mathbb{Z}[x]$ egy eleme, az algoritmus mindig véges sok lépésben megáll, és helyes választ ad arra a kérdésre, hogy bemenete irreducibilis-e $\mathbb{Z}[x]$-ben. Ezzel együtt, $\mathbb{Q}[x]$, $\mathbb{Z}[x]$ irreducibilis elemeire szép jellemzés (szükséges-és-elégséges feltétel) nem ismert. Ezért azzal folytatjuk, hogy megadjuk a $\mathbb{Q}[x]$-beli irreducibilitás néhány - a gyakorlatban hasznosnak bizonyult - elégséges feltételét.

[3]Emlékeim szerint előadáson ezt időhiány miatt nem részleteztem. A hallgatóktól kapott anyagban mégis volt egy elég zavaros vázlat erről, ezért a rend kedvéért kiegészítettem, de az egyértelműségről vizsgán elég annyit tudni, hogy a „$\mathbb{Q}[x]$-beli felbontás egyértelműségéből következik a $\mathbb{Z}[x]$-beli felbontás egyértelműsége". A létezés bizonyítását - természetesen - részletesen, pontosan kell tudni.

Schönemann-Eisenstein kritérium

6.73. Tétel. *Legyen* $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$.

1. **Schönemann-Eisenstein kritérium.** *Ha* $\exists p \in \mathbb{N}$ *olyan prím, hogy:*
   (a) $p \mid a_0, p \mid a_1, \ldots, p \mid a_{n-1}$*;*
   (b) $p \nmid a_n$ *és*
   (c) $p^2 \nmid a_0$*,*
   akkor f *irreducibilis* $\mathbb{Q}[x]$*-ben.*
2. **Fordított Schönemann-Eisenstein kritérium.** *Ha* $\exists p \in \mathbb{N}$ *olyan prím, hogy:*
   (a) $p \mid a_1, \ldots, a_n$*,*
   (b) $p \nmid a_0$ *és*
   (c) $p^2 \nmid a_n$*,*
   akkor f *irreducibilis* $\mathbb{Q}[x]$*-ben.*

Schönemann-Eisenstein kritérium bizonyítása. A 6.67 Tétel szerint van $r \in \mathbb{Q}$ és primitív $f^* \in \mathbb{Z}[x]$, hogy $f = r \cdot f^*$. Mivel f egész-együtthatós, ezért r egész szám.

Ellentmondást keresve TFH: f nem irreducibilis $\mathbb{Q}[x]$-ben $\Rightarrow$ f^* sem irreducibilis $\mathbb{Q}[x]$-ben. De f^* primitív, ezért a 6.69 Tétel szerint f^* $\mathbb{Z}[x]$-ben sem irreducibilis: $\exists g, h \in \mathbb{Z}[x] : f^* = g \cdot h$. Tehát

$$f = r \cdot f^* = r \cdot g \cdot h,$$

$\deg(f) = \deg(g) + \deg(h)$. Továbbá f^* primitív, tehát egyetlen szorzatra bontásában sem lehetnek nulladfokú tényezők (mert a nulladfokú tényező osztaná f^* összes együtthatóját). Emiatt $1 \leq \deg(g), \deg(h)$ és így $\deg(g), \deg(h) < n$. Vegyük fel g és h együtthatóit, legyen mondjuk

$$g = b_m x^m + \cdots + b_1 x + b_0,$$
$$h = c_k x^k + \cdots + c_1 x + c_0.$$

Mint a 6.64 Definícióban, legyen $\mu_p : \mathbb{Z} \to \mathbb{Z}_p$, $\mu_p(x) = x \pmod p$ és legyen

$$\varphi_p : \mathbb{Z}[x] \to \mathbb{Z}_p[x],\ \varphi_p(f) = f \text{ együtthatói } \pmod p.$$

Az állításban megadott feltételek miatt (és mivel a 6.66 Tétel szerint φ_p homomorfizmus):

- $\varphi_p(f) = \mu_p(a_n)x^n = \mu_p(a_n)x^n + 0x^{n-1} + \cdots + 0$ (mert p osztja f összes többi együtthatóját) és
- $\varphi_p(f) = \varphi_p(r \cdot g \cdot h) = \varphi_p(r \cdot g) \cdot \varphi_p(h)$.

Ezért $\varphi_p(r \cdot g) \mid \mu_p(a_n)x^n$ és $\varphi_p(h) \mid \mu_p(a_n)x^n$. De $\mathbb{Z}_p[x]$-ben $\mu_p(a_n)x^n$ egy irreducibilis tényezőkre bontása

$$\mu_p(a_n)x^n = (\mu_p(a_n)x) \cdot x \cdots x,$$

ezért az Alaptétel miatt osztói is konstans-szor x-hatvány alakúak, azaz $\exists u, t \in \mathbb{Z}_p$:

$$\varphi_p(r \cdot g) = u \cdot x^m,$$
$$\varphi_p(h) = t \cdot x^k,$$

és egyrészt $m + k = n$, másrészt $m, k < n$ miatt $m, k \geq 1$. Tehát $\varphi_p(r \cdot g)$ konstans tagja 0 és ezért $r \cdot g$ konstans tagja (azaz b_0) osztható p-vel. Hasonlóan: h konstans tagja (azaz c_0) szintén osztható p-vel. Végül: f-nek a_0 a konstans tagja és $a_0 = b_0 \cdot c_0$. Emiatt a_0 osztható p^2-el is; ez ellentmond a tétel feltételeinek. □

Fordított Schönemann-Eisenstein bizonyítása. Az előző bizonyítás jelöléseit megtartva, ellentmondást keresve TFH: f nem irreducibilis $\mathbb{Q}[x]$-ben. Ebből (pontosan ugyanúgy, mint az előző bizonyításban) azt kapjuk, hogy $\exists r \in \mathbb{Z}$, $g, h \in \mathbb{Z}[x]$:

$$f = (r \cdot g) \cdot h, \quad 1 \leq \deg(g), \deg(h) < n.$$

A feltételek szerint:

$$\varphi_p(f) = \mu_p(a_0) \neq 0 \text{ és}$$
$$\varphi_p(f) = \varphi_p(r \cdot g) \cdot \varphi_p(h).$$

Ezek szerint $\varphi_p(r \cdot g) \mid \mu_p(a_0)$ és $\varphi_p(h) \mid \mu_p(a_0)$. Tehát $\varphi_p(r \cdot g)$ és $\varphi_p(h)$ nulladfokú polinomok. Mivel azonban $r \cdot g$ és h legalább elsőfokúak, ez csak úgy lehet, ha $r \cdot g$ főegyütthatója (azaz b_m) és h főegyütthatója (azaz c_k) oszthatók p-vel. Ekkor azonban f főegyütthatója (azaz a_n) az $a_n = b_m c_k$ egyenlőség miatt osztható p^2-el, ami ellentmond az állításban felsorolt feltételeknek. □

6.74. Megjegyzés. Ha $f = a_n x^n + \cdots + a_1 x + a_0$, akkor

$$S(f) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

az f reciprok polinomja. Pl.: ha $f = 3x^2 - 7x + 2$, akkor $S(f) = 2x^2 - 7x + 3$. Könnyen ellenőrízhető, hogy rögzített $n \in \mathbb{N}$-re (tetszőleges kommutatív, nullosztómentes gyűrű felett) a legfeljebb n-edfokú polinomok között S művelettartó bijekció. Erre alapozva a Schönemann-Eisenstein kritérium és a fordított Schönemann-Eisenstein kritérium gyorsan bebizonyítható egymásból.

A Schönemann-Eisenstein kritériumok $\mathbb{Q}[x]$-beli irreducibilitást garantálnak. Ha $f \in \mathbb{Z}[x]$ primitív és teljesül valamelyik Schönemann-Eisenstein feltétel, akkor a 6.69 Tétel értelmében f $\mathbb{Z}[x]$-ben is irreducibilis. A Schönemann-Eisenstein feltétel elégséges, de nem szükséges az irreducibilitáshoz. Például

$$27x^3 + 8$$

irreducibilis $\mathbb{Q}[x]$-ben (mert egyetlen gyöke sem racionális, azaz nincs gyöke $\mathbb{Q}$-ban, és harmadfokú). De egyik kritérium sem teljesül: konstans tagja 2-hatvány, ezért a Schönemann-Eisenstein kritérium esetében csak $p = 2$ jön szóba; főegyütthatója pedig 3-hatvány, ezért a fordított Schönemann-Eisenstein kritérium esetében csak $p = 3$ jön szóba. De

- $p = 2$: $2^2 = 4 \mid 8 \Rightarrow$ Schönemann-Eisenstein nem alkalmazható;
- $p = 3$: $3^2 = 9 \mid 27 \Rightarrow$ fordított Schönemann-Eisenstein nem alkalmazható.

6.75. Megjegyzés (Algoritmikus aspektus)**.** Léteznek hatékony irreducibilis tényezőkre bontó algoritmusok (LLL-algoritmus: Lenstra-Lenstra-Lovász).

Irreducibilis polinomok fokszámairól

Láttuk, hogy $\mathbb{C}[x]$-ben az irreducibilis polinomok elsőfokúak, $\mathbb{R}[x]$-ben pedig legfeljebb másodfokúak. A következő tételben megmutatjuk, hogy $\mathbb{Q}[x]$-ben, illetve véges testek feletti polinomgyűrűkben az irreducibilis polinomok fokai nem maradnak korlátosak.

6.76. Tétel.

1. *$\mathbb{Q}[x]$-ben minden $n \in \mathbb{N}^+$-ra van n-edfokú irreducibilis polinom.*
2. *Legyen $I = \{f \in \mathbb{Z}_p[x] \mid f$ irreducibilis$\}$. Ekkor I végtelen és $\mathbb{Z}_p[x]$-ben az irreducibilis polinomok fokszáma nem marad egy közös korlát alatt.*

Bizonyítás. (1)-hez elég az $x^n + 2$ polinomot vizsgálni: ennek irreducibilisa következik a Schönemann-Eisenstein kritériumból ($p = 2$-vel).

(2) igazolása hasonló lesz Eukleidész, a prímszámok végtelenségére vonatkozó jól ismert bizonyításához. Ellentmondást keresve TFH I véges. Legyen $g = \left(\prod_{f \in I} f\right) + 1$. Mivel $\mathbb{Z}_p$ test, g is irreducibilis tényezőkre bontható, de g nem osztható I egyetlen elemével sem, ez ellentmondás. Tehát I végtelen. De rögzített n-re $\mathbb{Z}_p[x]$-ben csak véges sok n-edfokú polinom van, ezért:

$$\{\deg(f) : f \in \mathbb{Z}_p[x] \text{ irreducibilis}\} \text{ nem korlátos.}$$

Végül megjegyezzük, hogy az előző érvelés szó szerint érvényben marad $\mathbb{Z}_p[x]$ helyett tetszőleges véges testre. □

6.77. Példa (Schönemann-Eisenstein alkalmazása)**.**

- $x^4 + 4$ *irreducibilis?* $p = 2$*:* $2 \mid 4$*,* $2 \nmid 1$*,* $2^2 = 4 \mid 4 \Rightarrow$ *nem alkalmazható.*
- $x^3 - 2$ *irreducibilis?* $p = 2$*:* $2 \mid 2$*,* $2 \nmid 1$*,* $2^2 = 4 \nmid 2 \Rightarrow$ *irreducibilis.*
- $2x^3 + 3x^2 + 6x + 12$ *irreducibilis?* $p = 3$*:* $3 \mid 3, 6, 12$*,* $3 \nmid 2$*,* $3^2 = 9 \nmid 12 \Rightarrow$ *irreducibilis.*

7. A lineáris algebra alapjai

Előkészületek

7.1. Definíció (Permutációcsoport)**.** *Ha I halmaz, akkor*

$$Sym(I) = \{f : I \to I \mid f \text{ bijekció}\}.$$

A csoportműveletek a függvény-kompozíció, és a függvény-invertálás (ezekkel tényleg csoportot kapunk).

Ha $n \in \mathbb{N}$, akkor $[n] = \{1, 2, 3, \ldots, n\}$.

$Sym([n])$ elemei megadhatók úgy, hogy $[n]$ elemeit sorba tesszük (mindegyiket pontosan 1-szer használva).

7.2. Példa. *Ha $n = 5$, akkor*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix} \text{ egy permutáció.}$$

Mivel a felső sor mindig ugyanaz, általában elhagyjuk.

7.3. Definíció (Inverzió)**.** *Legyen $\sigma = a_1 a_2 \ldots a_n$ az $[n]$ egy permutációja. Ebben (i, j) inverziót alkot, ha $i < j$ de $a_i > a_j$.*

7.4. Példa. 5 1 2 3 4*-ben inverziók:* $(1, 5), (2, 5), (3, 5), (4, 5)$.

7.1. Vektorterek fogalma

7.5. Definíció (Vektortér)**.** *Legyen K test, V halmaz (vektorok).*
V vektortér K felett, ha:

- $+$*:* $V \times V \to V$ *(vektorok összeadása),*
- $\cdot$*:* $K \times V \to V$ *(testelemmel szorzás)*

olyan függvények, hogy:

1. $(V, +)$ *Abel-csoport;*
2. $\forall \alpha \in K, \forall u, v \in V : \alpha(u + v) = \alpha u + \alpha v$*;*
3. $\forall \alpha, \beta \in K, \forall u \in V : (\alpha + \beta)u = \alpha u + \beta u$*;*
4. $\forall \alpha, \beta \in K, \forall u \in V : \alpha(\beta u) = (\alpha\beta)u$*;*
5. $\forall u \in V : 1u = u$ *(1 az alaptest multiplikatív egysége).*

7.6. Megjegyzés. A vektor-összeadás kummutativitása következik a többi vektortér-axiómából, ugyanis tetszőleges $u, v \in V$-re egyrészt

$$(1 + 1)(u + v) = (u + v) + (u + v) = u + v + u + v,$$

másrészt

$$(1 + 1)(u + v) = (1 + 1)u + (1 + 1)v = u + u + v + v.$$

Ezekből $u + v + u + v = u + u + v + v$, ahonnan (balról u-t, jobbról v-t kivonva) $v + u = u + v$ adódik.

7.7. Példa.

- $K = \mathbb{R}$ *(esetleg* $K = \mathbb{C}$ *test);*
- $n = 2$ *(vagy* $n = 3$*) mellett* $V = K^n$*; műveletek koordinátákként:*

  ha $(a_1, a_2) \in \mathbb{R}^2$*,* $(b_1, b_2) \in \mathbb{R}^2$*,* $c \in \mathbb{R}$*, akkor*

  $$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$$

  és

  $$c(a_1, a_2) = (ca_1, ca_2),$$

  $\mathbb{R}^2$ *ezekkel vektorteret alkot.*

Lineáris kombináció

7.8. Definíció (Lineáris kombináció)**.** *Ha adottak* $v_1, \ldots, v_n \in V$ *vektorok, akkor a számolási szabályok (a 7.5 Definícióban megadott vektortér-axiómák) miatt ezekből pontosan* $c_1v_1+c_2v_2+\cdots+c_nv_n$ *alakú vektorok állíthatók elő (*$c_1, \ldots, c_n \in K$*). Ezeket a kifejezéseket nevezzük* $v_1, \ldots, v_n$ *lineáris kombinációinak.*

7.2. Lineáris egyenletrendszerek

7.9. Definíció (Lineáris egyenletrendszer)**.**

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Itt $a_{11} \ldots a_{mn}, b_1, \ldots b_m$ *adottak, és feladatunk az* $x_1, \ldots x_n$ *ismeretlenek értékeinek meghatározása.*

Mivel a lineáris egyenletrendszerek nagyon hasonló alakúak, az egészet összefoglalhatjuk egy táblázatba (mátrixba). Táblázatunk i-edik sorának j-edik pozíciójába a_{ij}-t írva, majd a táblázat jobb szélét egy további oszloppal kiegészítve, melybe felülről lefele $b_1, b_2, \ldots$ értékeit írjuk, minden információt rögzítettünk (abban az értelemben, hogy egy ilyen táblázatból fel tudjuk írni egyenletrendszerünk előző definícióban megadott alakját).

7.10. Megjegyzés.

- ***Sormodell**: Látjuk majd, hogy minden egyes lineáris egyenlet (azaz egyenletrendszerünk minden egyes sora) egy hipersíkot ír le, ezért az egyenletrendszer megoldása hipersíkok metszete lesz.*
- ***Oszlopmodell**: Legyen*

$$u_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix},\ u_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} \cdots, u_n = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- *Az egyenletrendszerünk megoldásása során valójában az a feladatunk, hogy állítsuk elő b-t $u_1, \ldots, u_n$ lineáris kombinációjaként: olyan $x_1, \ldots, x_n$ kell, hogy $x_1u_1 + \cdots + x_nu_n = b$ teljesüljön.*

Lineáris egyenletrendszerek megoldása

7.11. Példa. *Oldjuk meg a következő egyenletrendszert:*

$$\begin{cases} x + y + 2z = 0 \\ 2x + 2y + 3z = 2 \\ x + 3y + 3z = 4 \\ x + 2y + z = 5. \end{cases}$$

Megoldás lépésenként. Egy 1-ismeretlenes lineáris egyenletrendszert nagyon könnyen meg tudunk oldani. Ha több ismeretlenünk van, akkor próbálkozhatunk azzal, hogy valamelyik ismeretlent kiküszöböljük, így egy olyan egyenletrendszert kapunk, amiben kevesebb ismeretlen van. Ezt ismételgethetjük, amíg 1-ismeretlenes egyenletrendszert nem kapunk. Hogyan küszöböljünk ki ismeretleneket? Ha az egyik egyenletünk valahányszorosát egy másik egyenletünkhoz adjuk, akkor az új egyenlet következni fog a korábbiakból (abban az értelemben, hogy ami a régi egyenletrendszernek megoldása volt, az megoldása marad az új egyenletrendszernek is). Ezzel a módszerrel úgy tudunk kiejteni (kiküszöbölni) ismeretleneket, hogy az egyik (mondjuk az i-edik) egyenletnek olyan számszorosát (mondjuk c-szeresét) adjuk egy másik egyenlethez (mondjuk a j-edikhez), hogy valamelyik ismeretlen együtthatója a j-edik egyenletben épp $-c$-szerese legyen ugyanennek az ismeretlennek az i-edik egyenletben szereplő együtthatójának. Példánkon ezt az ötletet fogjuk illusztrálni; a módszer szisztematikusabb változatait Gauus-eliminációnak nevezik.

Írjuk fel a kibővített mátrixot és végezzük el az elemi sorátalakításokat:

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 2 & 2 & 3 & 2 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 1 & 5 \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 2 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 1 & 5 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 2 & 1 & 4 \\ 1 & 2 & 1 & 5 \end{array}\right]$$

$$\xrightarrow{R_4 \to R_4 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 2 & 1 & 4 \\ 0 & 1 & -1 & 5 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 5 \end{array}\right] \xrightarrow{R_4 \to R_4 - \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -\frac{3}{2} & 3 \end{array}\right]$$

$$\xrightarrow{R_4 \to R_4 - \frac{3}{2}R_3} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Az utolsó sor: $0x + 0y + 0z = 0$ mindig teljesül.
Most oldjuk meg visszafele (lemtről felfele) haladva:

$$-1z = 2 \Rightarrow z = -2;$$
$$2y + z = 4 \Rightarrow 2y - 2 = 4 \Rightarrow 2y = 6 \Rightarrow y = 3;$$
$$x + y + 2z = 0 \Rightarrow x + 3 - 4 = 0 \Rightarrow x - 1 = 0 \Rightarrow x = 1.$$

Tehát a megoldás: $x = 1$, $y = 3$, $z = -2$. □

Elemi sorátalakítások

7.12. Definíció (Elemi sorműveletek)**.**

1. ***Sorcsere****: két sor felcserélése;*
2. ***Sor szorzása****: egy sor szorzása $\neq 0$ skalárral;*
3. ***Sorok összeadása****: egyik sorhoz hozzáadjuk egy másik sor skalárszorosát.*

7.13. Megjegyzés. Az elemi sorátalakítások nem változtatják meg az egyenletrendszer megoldáshalmazát. Ez a következők miatt van így. Az világos, hogy az elemi sorátalakításokkal mindig olyan egyenletrendszerhez jutunk, hogy eredeti egyenletrendszerünk minden megoldása megoldása marad az új egyenletrendszernek is.

De minden elemi sorátalakítás „visszacsinálható": egy másik, alkalmasan választott elemi sorátalakítással az új egyenletrendszerünkből visszakaphatjuk az eredetit. Ezért az előző bekezdés gondolata azt is adja, hogy új egyenletrendszerünk minden megoldása egyúttal megoldása az eredeti egyenletrendszernek is.

Az előző két bekezdés szerint az eredeti és az új egyenletrendszerünk megoldáshalmazai kölcsönösen tartalmazzák egymást, ezért az eredeti és az új egyenletrendszer megoldásainak halmaza egyenlő egymással.

7.14. Megjegyzés. *A fenti példában:*

- *A negyedik egyenlet redundáns volt (a többi következménye);*
- *Pontosan 1 megoldás volt (azaz a megoldás egyértelmű):* $(1, 3, -2)$.

7.3. Vektorterekről bővebben

7.15. Definíció. *Legyenek V és W vektorterek a K test felett. Ekkor W* **altere** *V-nek, ha $W \subseteq V$ (azaz W elemei egyúttal V-nek is elemei) és a W-beli vektorokon minden műveletnek ugyanaz az eredménye, mintha V-ben végeznénk el őket.*

Tipikus esetben adott egy V vektortér és vektoroknak egy $W \subseteq V$ részhalmaza, és a V-ben adott vektortér-műveleteket akarjuk W elemein elvégezni. Mikor lesz ezekkel a műveletekkel W a V egy altere? A vektortér-axiómák nyilvánvalóan kivétel nélkül érvényben maradnak, hiszen, ha néhány W-beli vektor megsértené pl. a vektorösszeadásra vonatkozó asszociativitási szabályt, akkor ugyanezek a vektork V-ben is ellenpéldát szolgáltatnának az asszociativitásra (ami nyilván lehetetlen, mert V vektortér). Az egyetlen problémát az jelentheti, hogy minden vektortérben bármely két vektort össze lehet adni és bármely vektort meg lehet szorozni bármelyik testelemmel. Ha tehát W is vektorteret alkot a V-beli műveletekkel, akkor szükséges, hogy

- W bármely két vektorával együtt azok V-ben számolt összegét is tartamazza (azaz $\forall u, v \in W : \ (u +^V v \in W)$;
- W bármely vektorával együtt annak minden (V-ben számolt) skalárszorosát is tartamazza (azaz $\forall u \in W, c \in K : \ (c \cdot^V u \in W)$.

Ezekre a feltételekre együtt úgy hivatkozunk, hogy W zárt a V-beli műveletekre.

Az előző két bekezdés szerint $W \subseteq V$ a V-ből örökölt műveletekkel akkor és csak akkor lesz altere V-nek, ha W zárt a V-beli műveletekre. Mostantól - picit pontatlanul - egy vektortér altereinek mindig csak az alaphalmazát adjuk meg, és hallgatólagosan úgy értjük, hogy a vektortér-műveletek mindig V-ből öröklődnek.

7.16. Definíció (Generált altér)**.** *Legyen K test, V vektortér K felett, $n \in \mathbb{N}$, $\{v_1, \ldots, v_n\} \subseteq V$. Ekkor a $\{v_1, \ldots, v_n\}$ által generált altér:*

$$Span\{v_1, \ldots, v_n\} = \{\lambda_1 v_1 + \cdots + \lambda_n v_n : \lambda_i \in K\}.$$

Könnyű meggondolni, hogy az előző definícióban szereplő $Span\{v_1, \ldots, v_n\}$ zárt a V-beli műveletekre, ezért altere V-nek. Továbbá, ha $W \subseteq V$ olyan altér, melyre $\{v_1, \ldots, v_n\} \subseteq W$, akkor $Span\{v_1, \ldots, v_n\} \subseteq W$ is teljesül. Ezért $Span\{v_1, \ldots, v_n\}$ a legkisebb altere V-nek, ami tartalmazza a $v_1, \ldots, v_n$ vektorokat.

7.17. Definíció (Lineáris függetlenség)**.** *$F \subseteq V$ lineárisan független, ha egyik $f \in F$ sem fejezhető ki a többivel (azaz $\forall f \in F : f \notin Span(F \setminus \{f\})$).*

7.18. Definíció (Generátorrendszer)**.** *$G \subseteq V$ generátorrendszer, ha $\forall v \in V$ kifejezhető G-beliek lineáris kombinációjaként (azaz $\forall v \in V : v \in Span(G)$).*

7.19. Definíció (Bázis)**.** *$B \subseteq V$ bázis, ha B független generátorrendszer.*

7.20. Tétel (Lineáris függetlenség ekvivalens feltételei)**.** *Legyen V vektortér K felett és legyen $\{v_1, \ldots, v_n\} \subseteq V$. Ekvivalensek:*

1. *$\{v_1, \ldots, v_n\}$ lineárisan független;*
2. *Ha $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ (nullvektor), akkor $\lambda_1 = \cdots = \lambda_n = 0$ (vagyis a nullvektort csak a triviális lineáris kombnináció val lehet előállítani).*

Bizonyítás. (1) $\Rightarrow$ (2): Ellentmondást keresve tegyük fel, hogy van nem csupa-nulla $\lambda_1, \ldots, \lambda_n$ (azaz $\exists j : \ \lambda_j \neq 0$), hogy

$$\sum_{i=1}^{n} \lambda_i v_i = 0.$$

Ekkor

$$v_j = -\frac{1}{\lambda_j} \sum_{\substack{i=1 \\ i \neq j}}^{n} \lambda_i v_i,$$

ami – ellentmondva (1)-nek – mutatja, hogy $v_j \in Span(\{v_1, \ldots, v_n\} \setminus \{v_j\})$.
(2) $\Rightarrow$ (1): Tegyük fel, hogy $\{v_1, \ldots, v_n\}$ nem független. Ekkor $\exists j$, hogy v_j kifejezhető a többi lineáris kombinációjaként, azaz van olyan $\lambda_1, \ldots \lambda_{j-1}, \lambda_{j+1}, \ldots \lambda_n$, hogy

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n.$$

Ekkor (mindkét oldalból v_j-t kivonva)

$$\lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} - 1 \cdot v_j + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n = 0$$

ahol v_j együtthatója -1 (tehát, (2)-nek ellentmondva, v_j „λ-ja" nem nulla, mégis a nullvektort kaptuk). □

Bázis ekvivalens tulajdonságai

A generált alterek bevezetésénél már használtuk egy V vektortér részhalmazaira a „legkisebb" kifejezést. Általában is, egy alaphalmaz részhalmazait a tartalmazás szerint rendezzünk: az egyik részhalmaz nagyobb a másiknál, ha részhalmazként tartalmazza azt. Vigyázzunk: ez csak részbenrendezés: lehetnek olyan halmazok, melyek egyike sem részhalmaza a másiknak, ezért egyik sem kisebb/nagyobb a másiknál (vagyis a szóbanforgó halmazok összehasonlíthatatlanok). Ugyanebben az értelemben használjuk majd a „minimális" és „maximális" kifejezéseket. Halmazok egy $\mathcal{B}$ összességében az X halmaz maximális, ha nincs olyan $\mathcal{B}$-beli halmaz, melynek X részhalmaza lenne (azaz $\mathcal{B}$-ben nincs X-nél nagyobb halmaz). Teljesen hasonlóan, X minimális, ha nincs $\mathcal{B}$-ben olyan halmaz, mely részhalmaza lenne X-nek. Mégegyszer hangsúlyozzuk azonban, hogy $\mathcal{B}$-ben lehetnek olyan halmazok,

melyek összehasonlíthatatlanok egy minimális, illetve maximális halmazzal. Tehát pl. egy maximális halmazra csak az igaz, hogy nincs nála nagyobb halmaz $\mathcal{B}$-ben, de az nem feltétlenül igaz, hogy egy maximális halmaz minden $\mathcal{B}$-beli halmaznál nagyobb.

A generált alterek esetében a legkisebb kifejezést jogosan használtuk: legyen V vektortér, és legyen

$$\mathcal{B} = \{W \subseteq V : v_1, \dots, v_n \in W,\ W \text{altere } V\text{-nek}\}.$$

Ekkor $Span\{v_1, \dots, v_n\}$ nemcsak minimális részhalmaz $\mathcal{B}$-ben, hanem valóban legkisebb: ugyanis $Span\{v_1, \dots, v_n\}$ tényleg részhalmaza $\mathcal{B}$ összes elemének.

7.21. Tétel (Bázis ekvivalens jellemzései)**.** *Legyen $B = \{v_1, \dots, v_m\} \subseteq V$. Ekvivalensek:*

1. *B bázis;*
2. *B minimális generátorrendszer;*
3. *B maximális független rendszer;*
4. *$\forall v \in V : \exists! \lambda_1, \dots, \lambda_m \in K$ hogy $v = \lambda_1 v_1 + \dots + \lambda_m v_m$.*

7.22. Megjegyzés. *A (4)-es tulajdonságnál $[\lambda_1, \dots, \lambda_m]$ a v vektor koordináta-vektora a B bázisra vonatkozóan.*

7.23. Példa. *$\mathbb{R}^2$-ben:*

- $B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ *bázis, ez a szokásos koordinátákat adja:*
- *a* $v = \begin{pmatrix} x \\ y \end{pmatrix}$ *vektor B-re vonatkozó koordinátá-vektora:* $[x, y]$.

Bizonyítás. (ciklikus bizonyítás)

(1) $\Rightarrow$ (2): B generátorrendszer, mert definíció szerint minden bázis generátorrendszer. B minimális generátorrendszer, mert $\forall i$-re $v_i \notin Span(B \setminus \{v_i\})$, mert B független. Ez mutatja, hogy B valódi részhalmazai nem generátorrendszerek, tehát B a generátorrendszerek között valóban minimális halmaz.

(2) $\Rightarrow$ (3): B független, mert TFH valamilyen nem csupa-nulla $\lambda_1, \dots, \lambda_m$-re:

$$\lambda_1 v_1 + \dots + \lambda_m v_m = 0.$$

Legyen j olyan, hogy $\lambda_j \neq 0$. Ekkor:

$$v_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i v_i \Rightarrow v_j \in Span(B \setminus \{v_j\}),$$

ezért B nem lenne minimális generátorrendszer (hisz $B \setminus \{v_j\}$ generálná B-t, ami generálná az egész vektorteret).

Kell még, hogy B a független vektorhalmazok között maximális.
Ha lenne olyan $v \in V$, hogy $B \cup \{v\}$ független, akkor $v \notin Span(B)$, ezért, (2)-vel ellentétben, B nem lenne generátorrendszer.

(3) $\Rightarrow$ (4):

(3) miatt B maximális független, ezért $\forall v \in V$: $v \in Span(B)$. Ez (4)-ből a megfelelő λ-k létezését mutatja. Kell még a λ-k egyértelműsége. Legyen $\lambda_1, \ldots, \lambda_m$ és $\mu_1, \ldots, \mu_m$ úgy, hogy $v = \sum \lambda_i v_i = \sum \mu_i v_i$. Ekkor $\sum(\lambda_i - \mu_i)v_i = 0$, és mivel B független, $\lambda_i = \mu_i$ minden i-re, vagyis (4)-ben a λ-k valóban egyértelműen léteznek.

(4) $\Rightarrow$ (1): B generátorrendszer, mert $\forall v \in V$-re (4) miatt van $\lambda_1, \ldots, \lambda_m$: $v = \sum \lambda_i v_i$.

B független is, mert TFH: v_j előáll a többi lineáris kombinációjaként, azaz van $\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_m$:

$$v_j = \mu_1 v_1 + \cdots + \mu_{j-1} v_{j-1} + 0 \cdot v_j + \mu_{j+1} v_{j+1} + \cdots + \mu_m v_m.$$

De $v_j = 0 \cdot v_1 + \cdots + 0 \cdot v_{j-1} + 1 \cdot v_j + 0 \cdot v_{j+1} + \cdots + 0 \cdot v_m$ a v_j egy másik előállítása (hiszen az előző sorban v_j együtthatója nem nulla); ez ellentmond annak, hogy (4) szerint a koordináták egyértelműek. Ezért (1) valóban teljesül B-re. □

Bázisok és dimenzió

7.24. Tétel (Bázisok elemszáma)**.** *Legyen K test, $n \in \mathbb{N}$, V altere K^n-nek.*

1. *Ha $B_1, B_2 \subseteq V$ bázisok V-ben, akkor elemszámuk ugyanannyi: $|B_1| = |B_2|$;*
2. *V-ben van (véges elemszámú) bázis.*

Bizonyítás. Először egy önmagában is érdekes lemmát igazolunk.

Kicserélési lemma: Ha $F = \{f_1, \ldots, f_k\} \subseteq V$ független, $G = \{g_1, \ldots, g_m\}$ generátorrendszer, akkor minden i-hez van j, hogy F-ben f_i-t g_j-re cserélve továbbra is független rendszert kapunk (azaz $(F \setminus \{f_i\}) \cup \{g_j\}$ független).

A Kicserélési Lemma bizonyítása. Tegyük fel indirekt, hogy adott i-re f_i-hez nincs jó g_j, azaz $\forall j$: $\{f_1, \ldots, f_{i-1}, g_j, f_{i+1}, \ldots, f_k\}$ nem független. Ekkor, a 7.20 Tétel miatt az előző vektorrendszernek van nem csupa-nulla súlyokkal képzett lineáris kombinációja, mely a nullvektort adja. Egy ilyen lineáris kombinációban g_j együtthatója nem lehet nulla (mert F független volt). Ezért g_j kifejezhető a többi lineáris kombinációjával: $g_j \in Span\{f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_k\}$. Mivel ez minden j-re igaz, ezért $Span\{g_1, \ldots, g_m\} \subseteq Span\{f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_k\}$ generátorrendszer, ezért $f_i \in Span\{f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_k\}$ adódna, vagyis F nem lenne független. Ezzel az ellentmondással a kicserélési lemma be van bizonyítva. □

A 7.24 Tétel bizonyítását a Kicserélési Lemma egy következményével folytatjuk. Az alábbiak szerint, a Kicserélési Lemmából azonnal adódik, hogy ha F tetszőleges független vektorrendszer, G pedig tetszőleges véges generátorrendszer, akkor $|F| \leq |G|$. Ez azért van így, mert a Kicserélési Lemma szerint F elemeit sorban, egymás után (egészen addig, amíg el nem fogynak), ki tudjuk cserélni G elemeire. A Kicserélési Lemma garantálja, hogy menet közben F (illetve az a vektorrendszer, melyet az átalakítások során F-ből kaptunk) mindig független, ezért a Kicserélési Lemma végig alkalmazható marad, és szintén az „aktuális F" függetlensége garantálja, hogy az egyes lépésekben G-nek mindig más és más elemét tesszük át F-be (ha nem így lenne, akkor az aktuális F nem maradna független). Mivel F elemeit ki tudjuk cserélni G páronként különböző elemeire, ezért valóban $|F| \leq |G|$.

Következő lépésben (2)-t látjuk be. Minden $i \leq n$-re legyen

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad i\text{-edik helyen.}$$

$\{e_1, \ldots, e_n\}$ generátor K^n-ben. Emiatt, ha $\{b_1, \ldots, b_k\}$ független V-ben, akkor persze független K^n-ben is, és ezért az előző bekezdés szerint $k \leq n$. Azt kaptuk, hogy V-ben minden független vektorrendszer legfeljebb n elemű. Emiatt V-ben van maximális független rendszer (ami tehát bázis), és ez véges (legfeljebb n-elemű).

Ezek után (1) bizonyítása a következő. Legyenek B_1, B_2 bázisok V-ben. A korábbiak szerint $|B_1|, |B_2| \leq n$. Mivel B_1 független, és B_2 generátorrendszer, ezért $|B_1| \leq |B_2|$. De B_2 is független, és B_1 is generátorrendszer, ezért $|B_2| \leq |B_1|$. Tehát $|B_1| = |B_2|$. □

7.25. Definíció. *Az előző jelölések megtartásával, ha V altere K^n-ek, akkor V bázisainak közös méretét V* **dimenzió***jának nevezzük, és $\dim(V)$-vel jelöljük.*

7.26. Megjegyzés. A korábbi jelöléseket megtartva megjegyezzük, hogy a Kicserélési Lemmából, illetve a 7.24 Tételből könnyen adódik, hogy minden véges dimenziós vektortérben:

1. Minden generátorrendszer tartalmaz bázist;
2. Minden lineárisan független rendszer kiegészíthető bázissá;
3. $\dim(\mathbb{R}^n) = n$;
4. $\dim(\mathbb{C}^n) = n$ (komplex vektortérként);
5. $\{e_1, \ldots, e_n\}$ bázis K^n-ben (ezt nevezzük sztenderd bázisnak);
6. $\dim P_n = n + 1$ (itt P_n a K feletti, legfeljebb n-edfokú polinomok vektortere).

Megjegyezzük még, hogy minden (nem feltétlenül véges dimenziós) vektortérnek van bázisa, ezt idén nem bizonyítjuk (mert a bizonyítás további halmazelméleti előkészületeket igényelne).

7.27. Definíció (Affin altér)**.** *$W \subseteq K^n$ affin altér, ha van $u \in K^n$ és $V \subseteq K^n$ altér, úgy, hogy*

$$W = u + V = \{u + a : a \in V\}.$$

Itt u az eltolásvektor.

7.28. Definíció (Hipersík)**.** *$W = u + V$ hipersík, ha olyan affin altér, amiben V bázisai $(n-1)$ eleműek.*

7.29. Példa.

1. *$\mathbb{R}^2$-ben: az egyenesek affin alterek (sőt hipersíkok);*
2. *$\mathbb{R}^3$-ban: a síkok hipersíkok ($\dim = 2$);*
3. *$\mathbb{R}^3$-ban: az egyenesek (1-dimenziós) affin alterek.*

Affin alterek és lineáris egyenletrendszerek

7.30. Definíció (Homogén és inhomogén egyenletrendszer)**.** *$Az\ Ax = b$ lineáris egyenletrendszer:*

- ***homogén**, ha $b = 0$;*
- ***inhomogén**, különben.*

Jelölés:

$$M_H = \{u \in K^n : Au = 0\} \quad \text{homogén rész megoldáshalmaza;}$$
$$M_I = \{u \in K^n : Au = b\} \quad \text{inhomogén rész megoldáshalmaza.}$$

7.31. Megjegyzés. *Tetszőleges $A \in K^{m\times n}$, $u, v \in K^n$, $\lambda \in K$ esetén:*

$$A(u + v) = Au + Av;$$
$$A(\lambda u) = \lambda(Au).$$

7.32. Tétel (Megoldáshalmazok szerkezete)**.** *Ha $A \in K^{n\times m}$, akkor:*

1. *M_H altere $K^{n\times m}$-nek;*
2. *M_I affin altere $K^{n\times m}$-nek, sőt tetszőleges $u \in M_I$-re:*

$$M_I = u + M_H.$$

Bizonyítás. (1) M_H altér: a 7.15 Definíció utáni megjegyzés szerint azt fogjuk ellenőrizni, hogy M_H zárt a vektorok összeadására és skalárral szorzására. Ha $u, v \in M_H$, $\lambda \in K$, akkor:

$$A(u + v) = Au + Av = 0 + 0 = 0 \Rightarrow u + v \in M_H;$$
$$A(\lambda u) = \lambda(Au) = \lambda \cdot 0 = 0 \Rightarrow \lambda u \in M_H.$$

Ez mutatja M_H zártságát a megfelelő műveletekre.

(2) Legyen $u \in M_I$ (azaz $Au = b$). Ekkor egyrészt

$$\forall w \in M_H\colon\ A(u+w) = Au + Aw = b + 0 = b \Rightarrow u + w \in M_I \text{ ezért } u + M_H \subseteq M_I.$$

Másrészt fordítva:

$$\forall v \in M_I\colon\ v = u + (v-u), \text{ ahol } v - u \in M_H,$$

mert:

$$A(v-u) = Av - Au = b - b = 0.$$

Tehát $M_I = u + M_H$. □

7.33. Példa.

- $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$:

$$M_H = \{\lambda(-2,1) : \lambda \in \mathbb{R}\}, \quad M_I = (1,1) + M_H.$$

- $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$: $M_I = \emptyset$ *(ellentmondásos).*

7.4. Mátrixok és mátrixműveletek

Legyen K test, $n, m \in \mathbb{N}^+$. Az olyan téglalap alakú táblázatokat, melyeknek n sora, m oszlopa van, és a táblázatban K elemei szerepelnek, K feletti $n \times m$-es mátrixoknak nevezzük; az összes ilyen mátrix hamlazát $K^{n\times n}$-el jelöljük. Picit precízebben (de a mi céljaink szempontjából talán túlzott, a lényeget enyhén elfedő prcizitással) $K^{n\times m}$ elemei azonosíthatók azokkal a kétváltozós, K-ba képező függvényekkel, melyek értelmezési tartománya $[n] \times [m]$: hiszen egy mátrixban pontosan az lesz fontos, hogy az $i.$ sora $j.$ pozíciójában K melyik eleme van - ezt megadjatjuk egy függvénnyel (és végsősoron ezeket a függvényeket akár azonosíthatjuk is a megfelelő méretű mátrixokkal).

Az előző bekezdéssel összhangban, ha $A \in K^{n\times m}$ egy mátrix, akkor A i-edik sorának j-edik elemét $(A)_{ij}$-vel jelöljük (ami utalhat a megfelelő függvény helyettesítési értékére is). Emellett használjuk még a következő jelöléseket is:

- A i-edik sora: $A_{i*} = [A_{i1} \dots A_{im}] \in K^m$;
- A j-edik oszlopa: $A_{*j} = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{pmatrix} \in K^n$.

Mátrixokkal találkoztunk már a lineáris egyenletrendszerek vizsgálata során. Most szisztematikusan bezevetjük, és megvizsgáljuk az alapvető mátrix-műveleteket.

7.34. Definíció (Transzponált)**.** *Ha $A \in K^{n\times m}$, akkor $A^T \in K^{m\times n}$ az a mátrix, melyet A sorainak és oszlopainak felcserélésével kapunk, azaz*

$$(A^T)_{ij} = (A)_{ji}.$$

Az A^T mátrixot A **transzponáltjának** *nevezzük.*

7.35. Definíció (Mátrixszorzás)**.** *Ha $A \in K^{n\times m}$, $B \in K^{m\times k}$, akkor $AB \in K^{n\times k}$ a következő:*

$$(AB)_{ij} = \sum_{l=1}^{m} A_{il}B_{lj}$$

7.36. Megjegyzés. A mátrixszorzás nem kommutatív, mert a mátrixok mérete akár olyan is lehet, hogy egyik sorrendben összeszorozhatóak, de a másikban nem, ilyen esetekben fel sem merül, hogy a különböző sorrendekben vett szorzatok egyenlőek-e.

7.37. Definíció. *A négyzetes mátrix, ha sorai száma = oszlopai száma. Ekkor a szorzás elvégezhető (mindig).*

$K^{n\times n}$ a mátrixok összeadásával, szorzásával egy egységelemes, nem kommutatív gyűrűt alkot. A gyűrűaxiómák minden nehézség nélkül, de hosszadalmasan ellenőrizhetők. A szorzás asszociativitását később, a 7.51 Tételben bebizonyítjuk. A multiplikatív egységelemről és az inverzről most csak röviden teszünk néhány megjegyzést, később vissza fogunk térni ezekre.

7.38. Definíció (Egységmátrix)**.** *$I_n \in K^{n\times n}$ az egységmátrix:*

$$(I_n)_{ij} = \delta_{ij} = \begin{cases} 1 & ha\ i = j; \\ 0 & ha\ i \neq j. \end{cases}$$

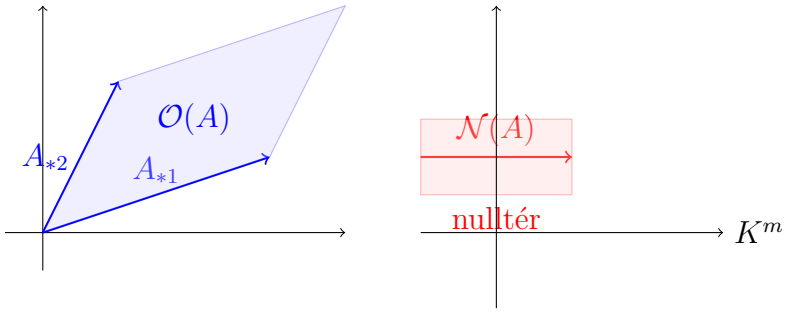
7.39. Definíció (Inverz mátrix)**.** *$A \in K^{n\times n}$-nek B inverze, ha $AB = BA = I_n$. Jelölés: $B = A^{-1}$.*

Látjuk majd, hogy nem minden mátrixnak van inverze. Továbbá - mivel a mátrix-szorzás még négyzetes mátrixok esetében sem kommutatív - kétféle inverzzel (balinverzzel és jobbinverzzel) is foglalkoznunk kéne. A balinverzek és jobbinverzek között szerencsére van kapcsolat. A részleteket később a 7.52 Definícióban, és az azt követő tételekben dolgozzuk ki. Az inverzzel való ismerkedésünket most azzal zárjuk, hogy a B inverzet megpróbálhatjuk meghatározni az $AX = I_n$ mátrixegyenlet megoldásával, bár ez nem is olyan rossz módszer, lesznek más módszerek is.

7.40. Definíció (Mátrix oszlop- és sortere)**.** *Legyen $A \in K^{n\times m}$.*

7.41. Definíció (Oszloptér, sortér, nulltér)**.**

$$\mathcal{O}(A) = Span\{A_{*1}, A_{*2}, \ldots, A_{*m}\} \quad A \textit{ oszloptere;}$$
$$\mathcal{S}(A) = Span\{A_{1*}, A_{2*}, \ldots, A_{n*}\} \quad A \textit{ sortere;}$$
$$\mathcal{N}(A) = \{u \in K^m : Au = 0\} = M_H \quad A \textit{ nulltere.}$$



Mátrixokhoz társított alterek bázisai

Hogyan találunk bázist ezekben a tereken?

I. Bázis $\mathcal{N}(A)$-ban (nulltér.)

Megoldjuk az $Ax = 0$ homogén egyenletrendszert.
Bázist úgy kapunk, hogy a szabad változók helyébe paramétereket írunk, a megoldást a paraméterek szerint rendezzük; ekkor a paraméterek vektor-együtthatói adnak egy bázist.

II. Bázis $\mathcal{S}(A)$-ban (sortér.)

Hozzuk redukált lépcsős alakra:

$$A = A_0 \to A_1 \to A_2 \to \cdots \to A_r = L.$$

A bázist L-nek a nem csupa-nulla sorai adják, ezt a következő tételben igazoljuk.

7.42. Tétel. $\mathcal{S}(A)$ *egy bázisa* L *nem csupa 0 sorai.*

Bizonyítás. Az elemi sorátalakítások olyanok, hogy A_{i+1} sorai $= A_i$ sorainak lineáris kombinációja. Ezért $\mathcal{S}(A_{i+1}) \subseteq \mathcal{S}(A_i)$. Mivel azonban az $A_i \to A_{i+1}$ lépés megfordítható, ezért $\mathcal{S}(A_i) \subseteq \mathcal{S}(A_{i+1})$, így végülis $\mathcal{S}(A_i) = \mathcal{S}(A_{i+1})$. Tehát

$$\mathcal{S}(A) = \mathcal{S}(L) = Span(L \text{ nem csupa 0 sorai}).$$

Ezért L nem csupa-nulla sorai generálják $\mathcal{S}(A)$-t.

Továbbá, L nem csupa 0 sorai lineárisan függetlenek a következők miatt. Legyen $\lambda_1, \ldots, \lambda_r \in K$ olyan, hogy

$$\lambda_1 L_{1*} + \lambda_2 L_{2*} + \cdots + \lambda_r L_{r*} = 0.$$

Ekkor $\lambda_1 = 0$, különben L_{1*} vezéreleme nem nullázódik le.
Hasonlóan, $\lambda_2 = 0$, $\lambda_3 = 0$, ... stb. Tehát L nem csupa 0 sorai lineárisan függetlenek. □

III. Bázis $\mathcal{O}(A)$-ban (oszloptér.)

Egy rövid válasz az lehet, hogy $\mathcal{O}(A)$ nyilvánvalóan izomorf $\mathcal{S}(A^T)$-vel, ezért $\mathcal{O}(A)$-ban úgy találhatunk bázist, ha a fenti módon $\mathcal{S}(A^T)$-ben keresünk egyet. Az alábbi tétel ennél több információt ad.

7.43. Tétel. *Legyen A (egy) redukált lépcsős alakra hozása L. Ha L-ben a $j_1, j_2, \ldots, j_r$ oszlopokban van vezérelem, akkor*

$$A_{*j_1}, A_{*j_2}, \ldots, A_{*j_r}$$

bázis $\mathcal{O}(A)$-ban.

Bizonyítás. Az $\{A_{*j_1}, \ldots, A_{*j_r}\}$ vektorrendszer generálja $\mathcal{O}(A)$-t, mert ha az A_{*k} oszlop nincs ezek között, akkor az

$$[A_{*j_1} \ldots A_{*j_r} | A_{*k}]$$

egyenletrendszer megoldható (lépcsős alakra hozás mutatja).

Továbbá, $\{A_{*j_1}, \ldots, A_{*j_r}\}$ lineárisan független a következők miatt. Tegyük fel, hogy

$$(*) \qquad \lambda_1 A_{*j_1} + \lambda_2 A_{*j_2} + \cdots + \lambda_r A_{*j_r} = 0.$$

Ekkor azonban $\lambda_1 = 0$, különben A_{*j_1} vezéreleme $(*)$-ban nem nullázódna ki. Hasonlóan $\lambda_2 = 0, \ldots, \lambda_r = 0$. □

7.44. Példa. *Legyen* $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 2 & 3 \end{pmatrix}$

Redukált lépcsős alak:

$$L = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- $\mathcal{S}(A)$ *egy bázisa:* $\{(1,0,1,2), (0,1,1,1)\}$*;*
- $\mathcal{O}(A)$ *egy bázisa:* $A_{*1} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $A_{*2} = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$*;*
- $\mathcal{N}(A)$ *egy bázisa:* $\{(-1,-1,1,0), (-2,-1,0,1)\}$.

Mátrix rangja és oszlopmodell

7.45. Definíció (Oszloprang és sorrang)**.** *Legyen* $A \in K^{n\times m}$*.*

- A ***oszloprangja****:* $r_o(A) = \dim(\mathcal{O}(A))$ *(= hány lineárisan független oszlopa van* A*-nak);*
- A ***sorrangja****:* $r_s(A) = \dim(\mathcal{S}(A))$*.*

7.46. Tétel (Oszloprang = sorrang)**.** $r_o(A) = r_s(A)$ *(ezért* r_s *felesleges, nem használjuk).*

Bizonyítás. $r_s(A) = \dim(\mathcal{S}(A)) = L$ nem csupa 0 sorainak száma.

$r_o(A) = \dim(\mathcal{O}(A)) = L$-ben ennyi oszlopban van vezérelem.

Ł egy nem csupa 0 sorához rendeljük hozzá azt az oszlopot, melyben a sor vezéreleme van. Ez kölcsönösen egyértelmű megfeleltetés L nem csupa 0 sorai és azon oszlopai között, melyekben van vezérelem. Ez a megfeleltetés mutatja, hogy valóban:

$$\dim(\mathcal{S}(A)) = \dim(\mathcal{O}(A)).$$

□

7.47. Definíció (Rang)**.**

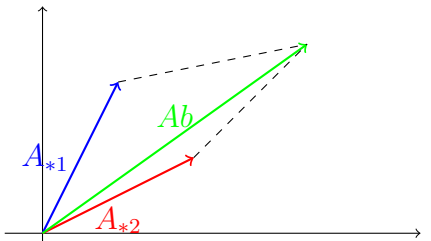
$$rang(A) = r_o(A) = r_s(A).$$

7.48. Tétel (Oszlopmodell értelmezése)**.** $Ax = c$ *megoldható* $\Leftrightarrow c \in \mathcal{O}(A)$*.*

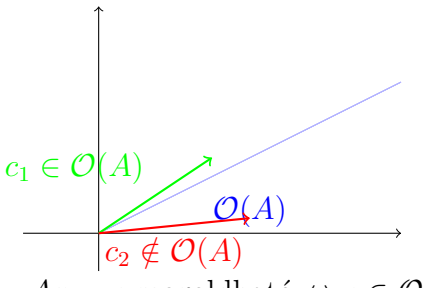
Bizonyítás. $Ax = c$ megoldható $\Leftrightarrow \exists x \in K^m : Ax = c$
$\Leftrightarrow c$ előáll A oszlopainak lineáris kombinációjaként
$\Leftrightarrow c \in \mathcal{O}(A)$. □

7.49. Példa. *Legyen* $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$

$$Ab = 5\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 6\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 15 \end{pmatrix} + \begin{pmatrix} 12 \\ 24 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}.$$



Oszlopmodell: $Ab = b_1A_{*1} + b_2A_{*2}$



$Ax = c$ megoldható $\Leftrightarrow c \in \mathcal{O}(A)$

7.50. Tétel (Lineéris egyenletrendszerek egoldhatóságának mátrixrangos jellemzése)**.**

1. *$Ax = c$ megoldható $\Leftrightarrow$ rang(A) = rang$(A|c)$;*
2. *Ha rang(A) = rang$(A|c)$ < változók száma, akkor sok megoldás van (vannak szabad változók);*
3. *Ha rang(A) = rang$(A|c)$ = változók száma, akkor pontosan 1 megoldás van (nincsenek szabad változók);*
4. *Nem fordulhat elő, hogy rang(A) = rang$(A|c)$ > változók száma, mert rang(A) = $\dim(\mathcal{O}(A)) \leq$ oszlopok száma = változók száma.*

7.5. Bővebben a mátrixok invertálhatóságáról

Ebben a részben a mátrixok inverzének létezésével, egyértelműségével, kiszámításának módjaival foglalkozunk. Egy négyzetes $A \in K^{n\times n}$ mátrix **főátlóján** az

$$a_{11}, a_{22}, \ldots, a_{nn}$$

elemeit értjük. Speciálisan: az I_n egységmátrix főátlójában 1-esek vannak. A mátrix-szorzás asszociativitásának ellenőrzésével kezdünk.

7.51. Tétel (Asszociativitás)**.** *$A, B, C \in K^{n\times n}$ esetén:*

$$(AB)C = A(BC).$$

Bizonyítás. Tetszőleges $i, j \leq n$-re

$$[(AB)C]_{ij} = \sum_{l=1}^{n} (AB)_{il} C_{lj} = \sum_{l=1}^{n} \left(\sum_{k=1}^{n} A_{ik} B_{kl} \right) C_{lj} = \sum_{k=1}^{n} A_{ik} \left(\sum_{l=1}^{n} B_{kl} C_{lj} \right) = [A(BC)]_{ij}.$$

□

Mint említettük, a mátrix-szorzás még a négyzetes mátrixok körében sem kommutatív, ezért e nem-kommutatív szorzásnak kétféle inverze lehet. Részletesebben:

7.52. Definíció (mátrix balinverze, jobbinverze)**.** *$A \in K^{n\times n}$-nek:*
- *Y jobbinverze, ha $AY = I_n$;*
- *Y balinverze, ha $YA = I_n$.*

A 7.39 Definícióban megadtuk már az A mátrix inverzének definícióját: Y inverze A-nak, ha $AY = YA = I_n$, vagyis ha Y egyszerre balinverze és jobbinverze is A-nak. E különböző inverzek között teremt kapcsolatot a következő tétel.

7.53. Tétel (Jobbinverz, inverz létezésének feltétele)**.** *Legyen $A \in K^{n\times n}$.*
1. *A-nak van jobbinverze $\Leftrightarrow$ rang$(A) = n$;*
2. *A-nak legfeljebb 1 db jobbinverze van;*
3. *Ha Y jobbinverz, akkor inverz is.*

Bizonyítás. (1) $\Rightarrow$: Tegyük fel: Y jobbinverze A-nak: $AY = I_n$. Ekkor tetszőleges j-re $AY_{*j} = j$-edik oszlop I_n-ben. Ezért $e_1, e_2, \ldots, e_n \in \mathcal{O}(A)$. De $\{e_1, e_2, \ldots, e_n\}$ bázis K^n-ben. ezért $\dim(\mathcal{O}(A)) = n$, vagyis $\text{rang}(A) = n$ (nagyobb nem lehet).

$\Leftarrow$: Fordítva: ha $\text{rang}(A) = n$, akkor tetszőleges $e_j \in K^n$-re $Ax = e_j$ megoldható (hiszen ekkor $\mathcal{O}(A) =$ az egész K^n tér). Tehát $\forall j$: $Ax = e_j$ megoldható, e megoldás, mint oszlopvektor lesz a jobbinverz j-edik oszlopa.

(2): Ha $\text{rang}(A) < n$, akkor (1) szerint nulla darab jobbinverz van, ha $\text{rang}(A) = n$, akkor az inverzre (1)-ben adott konstrukció egyértelmű (a mátrixrangos jellemzés (7.50 Tétel) miatt $\forall j$: $Ax = e_j$-nek pontosan 1 megoldása van, és ennek a megoldásnak kell lennie a jobbinverz j-edik oszlopának).

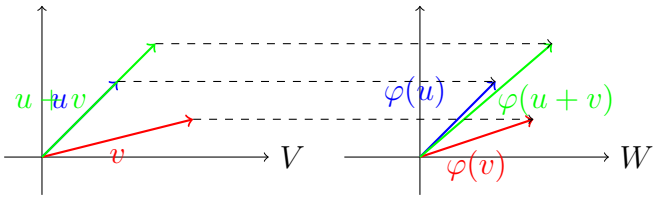
(3) Tegyük fel: Y jobbinverz: $AY = I_n \Rightarrow AYA = A$ azaz YA megoldása az

$$(*) \qquad AX = A$$

mátrixegyenletnek. Egyrészt $\text{rang}(A) = n$ (és a 7.50 Tétel) miatt e mátrixegyenletnek 1 megoldása van: tetszőleges j-re X_{*j} csak az $Ay = A_{*j}$ lineáris egyenletrendszer egyetlen megoldása lehet. Másrészt, $AI_n = A$ miatt $(*)$-nak $X = I_n$ megoldása, ezért $YA = I_n$. □

7.54. Definíció (Lineáris leképezés)**.** *Legyenek V, W vektorterek K felett.*
$\varphi : V \to W$ vektortér homomorfizmus (= lineáris leképezés), ha $\forall u, v \in V$, $c \in K$:

$$\varphi(u + v) = \varphi(u) + \varphi(v);$$
$$\varphi(cu) = c\varphi(u).$$



Lineáris leképezés: $\varphi(u + v) = \varphi(u) + \varphi(v)$.

Lineáris leképezésekre könnyű példát adni: legyen $A \in K^{m \times n}$ rögzített mátrix. Könnyen ellenőrízhető, hogy ekkor a $\varphi : K^n \to K^m$, $\varphi(x) = Ax$ függvény lineáris leképezés lesz (valójában ezt az ellenőrzést elvégeztük már a 7.31 Megjegyzésben is). A következő tétel szerint véges dimenziós esetben nincs is másfajta lineáris leképezés: mindegyik mátrixszorzásból származtatható!

7.55. Tétel (Lineáris leképezések mátrixreprezentációja)**.**

1. *Ha $f : K^n \to K^m$ lineáris, akkor $\exists A \in K^{m \times n}$: $\forall x \in K^n$: $f(x) = Ax$.*

2. *(Előírhatósági-tétel) Tetszőleges $w_1, \ldots, w_n \in K^m$-re $\exists$ lineáris f: $f(e_1) = w_1$, $\ldots$, $f(e_n) = w_n$.*
3. *Ha V véges dimenziós vektortér K felett, akkor van olyan $n \in \mathbb{N}$, hogy V izomorf K^n-el.*

(1)-ben valójában pontosan 1 ilyen A mátrix van, (2)-ben valójában pontosan 1 ilyen f függvény van. Az egyértelműség - egyébként rövid - bizonyításait a vizsgára nem kell tudni.

Bizonyítás. (1) Az A mátrix oszlopvektorai legyenek az $f(e_1), \ldots, f(e_n)$ oszlopvektorok:

$$A = [f(e_1) \mid f(e_2) \mid \cdots \mid f(e_n)] \in K^{m \times n}.$$

Ekkor $\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$:

$$f(x) = f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i) = Ax.$$

(2) Legyen $A = [w_1 \mid w_2 \mid \cdots \mid w_n] \in K^{m \times n}$ (oszlopok). Definiáljuk az f függvényt így: $f(x) = Ax$, ez jó.

(3) Mivel V véges dimenziós, van benne egy $\{b_1, \ldots, b_n\}$ véges bázis. $\forall v \in V$-hez $\exists! \lambda_1, \ldots, \lambda_n \in K$, hogy

$$v = \sum_{i=1}^{n} \lambda_i b_i.$$

$[\lambda_1, \ldots, \lambda_n]$ koordinátavektora v-nek $\{b_1, \ldots, b_n\}$-re. Definiáljuk φ-t így:

$$\varphi : V \to K^n,\ \varphi(v) = [\lambda_1, \ldots, \lambda_n]\ (v \text{ koordinátái } \{b_1, \ldots, b_n\}\text{-ben}).$$

Könnyen ellenőrízhető, hogy φ lineáris leképezés. Végül φ bijekció mert V minden elemének pontosan 1 koordináta-vektora van és K^n tetszőleges $[\lambda_1, \ldots, \lambda_n]$ eleme koordináta-vektora valamelyik V-beli vektornak (konkrétan a $\sum_{k=1}^{n} \lambda_k b_k$ vektornak). □

7.56. Megjegyzés.

1. (1), (3) végtelen dimenzióban nem marad érvényben, de (2) igen.
2. (1)-ben a konstrukció bázisfüggő (ha a sztenderd bázis helyett más bázist választunk, akkor más mátrixot kapunk, erre még vissza fogunk térni a 9.6 Tételben).

7.57. Példa (Forgatás $\mathbb{R}^2$-ben)**.**
Legyen $V = W = \mathbb{R}^2$, $f : V \to W$ α szögű (pozitív irányú) forgatás.

f mátrixa a sztenderd bázisban:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

(Az oszlopok a sztenderd bázisvektorok elfogatottjainak koordinátái a sztenderd bázisban.)

7.58. Példa (Polinomok terében)**.**
Legyen $V = \mathbb{R}[X]$ feletti legfeljebb 2-edfokú polinomok vektortere.

$f : V \to V$ ami a polinomokat $(x-1)$ hatványai szerint rendezi át: ha $ax^2 + bx + c = a'(x-1)^2 + b'(x-1) + c'$, akkor f az $[a, b, c]$ vektort $[a', b', c']$-re képezi. Ez az f lineáris.

f mátrixa a sztenderd $\{x^2, x, 1\}$ bázisban:

$$\begin{aligned} f(x^2) = x^2 &= 1 \cdot (x-1)^2 + 2 \cdot (x-1) + 1; \\ f(x) = x &= 0 \cdot (x-1)^2 + 1 \cdot (x-1) + 1; \\ f(1) = 1 &= 0 \cdot (x-1)^2 + 0 \cdot (x-1) + 1. \end{aligned}$$

Tehát:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

8. Determinánsok

Permutációk inverzióit bevezettük már a 7.3 Definícióban. Ebben a részben intenzíven használni fogjuk az inverziók további tulajdonságait, az egyik ilyen fontos tulajdonsággal kezdünk.

8.1. Tétel. *Ha $\pi \in Sym([n])$ az $[n]$ egy permutációja, $i < j \leq n$, akkor π-ben az i. és j. tagot felcserélve az inverziószám paritása megváltozik.*

Bizonyítás. Legyen:

$$\pi : \pi(1), \pi(2), \ldots, \pi(i), \ldots, \pi(j), \ldots, \pi(n)$$
$$\pi' : \pi(1), \pi(2), \ldots, \pi(j), \ldots, \pi(i), \ldots, \pi(n).$$

Először tegyük fel, hogy $j = i + 1$. Ekkor

$$\mathrm{inv}(\pi') - \mathrm{inv}(\pi) = \pm 1,$$

mert $\pi(i)$ és $\pi(i+1)$ felcserélésekor $\pi(i)$ és $\pi(i+1)$ egymással való inverziós viszonya megváltozik, de sem $\pi(i)$, sem $\pi(i+1)$ inverziós viszonya a permutáció semelyik más tagjával nem változik meg.

Ha i, j tetszőleges (nem feltétlenül egymás utáni) számok, akkor (a szomszédos elemek felcserélésével)

- Mozgassuk $\pi(i)$-t $\pi(j)$-ig: ez $j - i - 1$ lépés;
- Cseréljük fel $\pi(i)$-t $\pi(j)$-vel; ez 1 lépés;
- Mozgassuk vissza $\pi(j)$-t $\pi(i)$ helyére: ez $j - i - 1$ lépés.

Összesen $2(j - i - 1) + 1$-szer változik a paritás, ez páratlan szám. □

8.2. Megjegyzés. Ha $i \neq j$, akkor $[i, j] \in Sym([n])$ az a permutáció, amely i-t és j-t felcseréli, a többi elemet változatlanul hagyja:

$$\begin{pmatrix} 1 & 2 & \ldots & i-1 & i & i+1 & \ldots j-1 & j & j+1 & \ldots n \\ 1 & 2 & \ldots & i-1 & j & i+1 & \ldots j-1 & i & j+1 & \ldots n \end{pmatrix}.$$

A 2 elemet felcserélő, többi elemet fixen hagyó permutációkat transzpozícióknak is nevezik. A 8.1 Tételben azt láttuk be, hogy ha $\pi \in Sym([n])$, akkor $inv(\pi)$ és $inv(\pi \circ [i, j])$ eltérő paritású, itt $\pi \circ [i, j]$ az a permutáció, amit a függvénykompozíció ad: π-ben $\pi(i)$-t és $\pi(j)$-t felcseréljük.

8.1. A determináns geometriai jelentése, létezése, egyértelműsége

Vektortereinkben szeretnénk bizonyos halmazok térfogatát értelmezni, vagyis szeretnénk olyan függvényeket megadni, melyek vektortereink (bizonyos) részhalmazaihoz az alaptest elemeit rendeli – amennyire csak lehet – hasonlóan ahhoz, ahogy pl. $\mathbb{R}^2$ illetve $\mathbb{R}^3$ (egyes) részhalmazainak területét, térfogatát értelmeztük.

A térfogat-fogalom ilyen általánosításának szándéka a következő kihívásokat veti fel: tetszőleges test felett szeretnénk dolgozni, eredményeink olyan testekben is érvényesek lesznek, amiket jelenleg el sem bírunk képzelni. Emiatt mégkevésbé lehet intuíciónk arról, hogy (jelenlegi tudásunk alapján) ismeretlen testek feletti vektorterekben hogyan lenne érdemes definiálni a térfogatot. De hasonló nehézségekbe ütközünk akkor, ha egy jól ismert test (pl. $\mathbb{R}$) felett egy – mondjuk – 17-dimenziós vektortérben próbáljuk a térfogatot értelmezni. Ha a vektortér dimenziója elég nagy, még jól ismert test feletti, véges dimenziós vektorterek esetében sincs jól motiválható elképzelésünk a sokdimenziós halmazok térfogatairól, mely kizárólag a vektorterek algebrai szerkezetére alapoz. További probléma, hogy azok a halmazok, melyeknek térfogatot szeretnénk tulajdonítani, meglehetősen bonyolult szerkezetűek lehetnek (sokdimenziós görbe felületek határolhatják őket, de akár még ennél is szabálytalanabbak lehetnek). Ezért a következő szempontok szerint egy picit kevésbé ambíciózus célt tűzünk ki:

- ebben a kurzusban végig véges dimenziós (de tetszőleges, (0 karakterisztikájú) test feletti) vektorterekben fogunk dolgozni;
- csak szögletes testek térfogatát akarjuk értelmezni, ezek közül is csak a paralelepipedonokét (ezek definíciója alább lesz);
- ugyanakkor térfogat-fogalmunk előjeles lesz, ez később hasznosnak fog bizonyulni (pl. egyes alkalmazásokban, ha az a kérdés, hogy adott felületen időegység alatt mennyi folyadék áramlik át, akkor a térfogat előjelébe kódoltan meg fogjuk tudni különböztetni, hogy a kakaó felénk áramlik-e, vagy a másik irányban);
- intuíció hiányában a térfogatot axiomatikusan próbáljuk definiálni, valahogy így: azt ugyan (jelenleg még) nem tudjuk, mi a 17-dimenziós paralelepipedonok előjeles térfogata, de bármi is legyen ez, azt azért térfogat-fogalmunknak tudnia kell, hogy ... és itt felsorolunk néhány plauzibilis tulajdonságot, amit minden „valamirevaló” térfogat-fogalomnak tudnia kellene. Ezek mind könnyen motiválható, szemléletünk számára természetes tulajdonságok lesznek.
- Végül a 8.5 Tételben igazoljuk majd, hogy pontosan 1 darab olyan függvény van, mely sokdimenziós paralelepipedonokhoz az alaptest elemeit rendeli úgy, hogy az előző pontban említett előírásaink teljesülnek. Ezek után azt mondhatjuk majd, hogy a sokdimenziós (előjeles) térfogat az az egyetlen függvény, amely eleget tesz a térfogat-fogalommal kapcsolatos természetes előírásainknak. Ezt követően módszereket adunk majd térfogat-függvényünk kiszámításához.

A sokdimenziós (de véges-dimenziós) paralelepipedonok definíciójával folytatjuk.

8.3. Definíció (Paralelepipedonok)**.** *Legyen K tetszőleges test, $n \in \mathbb{N}^+$. Az $a_1, ..., a_n \in K^n$ vektorok által kifeszített n-dimenziós paralelepipedon csúcsai:*

$$\varepsilon_1 a_1 + \varepsilon_2 a_2 + \cdots + \varepsilon_n a_n,$$

ahol $\varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}$.

8.4. Definíció (Determináns függvény)**.** *$D : (K^n)^n \to K$ determináns függvény, ha $\forall a_1, \dots, a_n \in K^n$ (sorvektorok), $\forall c \in K$, $\forall b \in K^n$:*

1. *$D(a_1, \dots, ca_i, \dots, a_n) = cD(a_1, \dots, a_i, \dots, a_n)$;*
2. *$D(a_1, \dots, a_i + b, \dots, a_n) = D(a_1, \dots, a_i, \dots, a_n) + D(a_1, \dots, b, \dots, a_n)$;*
3. *Ha $i \neq j$, akkor $D(a_1, \dots, a_i \dots a_j, \dots a_n) = -D(a_1, \dots, a_j \dots a_i, \dots a_n)$;*
4. *$D(I_n) = 1$ ahol I_n az $n \times n$-es egységmátrix.*

A determináns függvény adja majd a sokdimenziós előjeles térfogat fogalmát. Mielőtt folytatnánk, végigmegyünk az előző definícióban szereplő pontokon, és összevetjük a térfogat-fogalommal kapcsolatos intuíciónkkal.

- Az (1) pont szerint, ha egy paralelepipedon egyetlen élét c-szeresére nyújtjuk, de a többi élét változatlanul hagyjuk, akkor a térfogat c-szeresre változik. Ez elég plauzibilis: ha pl. két egybevágó kockát egymásra teszünk úgy, hogy egy-egy lapjuk pontosan fedje egymást, akkor egyrészt olyan téglatestet kapunk, melynek egyik éle a kockák élhosszának duplája lesz, többi éle változatlan marad; másrészt a 2 kockából álló téglatest térfogata valóban duplája az eredeti kockák térfogatának.
- A (2) pont is hasonlóan intuitív, javaslom, hogy rajzold le 2-dimenzióban.
- (3) a térfogat előjelével van kapcsolatban. A paralelepipedonok irányítását azzal lehet kódolni, hogy milyen sorrendben adjuk meg az oldaléleit.
- Végül (4)-nek csak skálázási, „mértékegység-választási" szerepe van: azt rögzíti, hogy az egységnyi élhosszúságú kocka térfogata egységnyi.

8.5. Tétel (Determináns létezése és egyértelműsége)**.** *Véges n-re pontosan 1 db determináns függvény van.*

Bizonyítás. (1) Unicitás: Tegyük fel, hogy D determináns függvény.

Legyenek $a_1 \dots a_n \in K^n$, vegyük fel ezeket a vektorokat koordinátásan is: $a_i = \sum_{j=1}^{n} a_{ij} e_j$.

Ekkor:

$$D(a_1, \dots, a_n) = D\left(\sum_{i_1=1}^{n} a_{1i_1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{ni_n} e_{i_n}\right) \overset{8.4\ (1),(2)}{=}$$

$$\sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{1i_1} \cdots a_{ni_n} D(e_{i_1}, \dots, e_{i_n}) = (*).$$

Vizsgáljuk meg az utolsó összegben előforduló $D(e_{i_1}, \dots, e_{i_n})$ mennyiségeket.

1. eset: van $j \neq k$: $i_j = i_k$. Ekkor $D(e_{i_1}, \dots, e_{i_n})$-ben a j-edik és k-adik sort felcserélve egyrészt nem történik változás (mert $e_{i_j} = e_{i_k}$), másrészt cserekor 8.4 (3) miatt kifejezésünk előjelet vált. Ez csak úgy lehet, hogy $D(e_{i_1}, \dots, e_{i_n}) = 0$. Ezért $(*)$-ban minden olyan összeadandó nulla (tehát elhagyható), melyek az 1. esethez tartoznak.

2. eset: az $i_1, i_2, \dots, i_n$ értékek páronként különbözők, azaz $(i_1, \dots, i_n)$ permutációja

$[n]$-nek; jelöljük ezt a permutációt π-vel. Ez azt jelenti, hogy az az F mátrix, melynek sorai $e_{i_1}, e_{i_2}, \ldots, e_{i_n}$ az egységmátrix sorainak permutálásával áll elő. Az egységmátrixból F megkapható úgy, hogy az egységmátrixban pontosan annyi sort cserélünk fel, mint amennyi π inverziószáma[4]. Ezért ebben a 2. esetben (8.4 (3)-at és a 8.1 Tételt is figyelembe véve)

$$D(e_{i_1}, \ldots, e_{i_n}) = (-1)^{inv(\pi)} D(e_1, \ldots, e_n) \overset{8.4\ (4)}{=} (-1)^{inv(\pi)}.$$

Az előző 2 eset szerint

$$(*) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}.$$

Azt kaptuk, hogy ha létezik egyáltalán D determináns függvény, akkor az csak

$$D(a_1, \ldots, a_n) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$$

lehet.

(2) Egzisztencia: Legyen

$$J(a_1, \ldots a_n) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}.$$

Megmutatjuk, hogy ez determináns függvény (azaz teljesülnek rá a 8.4 Definícióban megadott tulajdonságok).

- 8.4 (1) teljesül, mert:

$$J(a_1, \ldots, ca_i, \ldots, a_n) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} \cdots ca_{i\pi(i)} \cdots a_{n\pi(n)} =$$

$$c \cdot \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} = cJ(a_1, \ldots, a_i, \ldots, a_n).$$

- 8.4 (2) teljesül, mert:

$$J(a_1, \ldots, a_i + b, \ldots, a_n) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} \cdots (a_{i\pi(i)} + b_{\pi(i)}) \cdots a_{n\pi(n)} =$$

$$\sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{n\pi(n)} + \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{\pi(1)} \cdots b_{\pi(i)} \cdots a_{n\pi(n)} =$$

$$J(a_1, \ldots, a_i, \ldots, a_n) + J(a_1, \ldots, b, \ldots, a_n).$$

[4]Nyilván megkaphatjuk F-t másképp is: pl. először feleslegesen oda-vissza cserélgetjük a sorokat, és miután visszakaptuk az egységmátrixot, előállítjuk F-t... De az, hogy páros sok, vagy páratlan sok csere kellett-e F előállításához, az minden esetben ugyanúgy lesz.

- 8.4 (3) teljesül a következők miatt. Ha $\pi \in Sym([n])$ tetszőleges permutáció, akkor legyen π' az a permutáció, melyet π-ből $\pi(i)$ és $\pi(j)$ felcserélésével kapunk (azaz $\pi' = \pi \circ [i,j]$). Figyeljük meg, hogy ekkor $a_{j\pi(i)}a_{i\pi(j)} = a_{i\pi'(i)}a_{j\pi'(j)}$. A 8.1 Tétel szerint π és π' inverziószáma eltérő paritású. Ezeket fejben tartva (a legutolsó előtti lépést a számolás után magyarázva):

$$J(a_1,\dots,a_j,\dots,a_i,\dots,a_n) = \sum_{\pi\in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)}\cdots a_{j\pi(i)}\cdots a_{i\pi(j)}\cdots a_{n\pi(n)} =$$

$$- \sum_{\pi\in Sym([n])} (-1)^{inv(\pi')} a_{1\pi(1)}\cdots a_{j\pi(i)}\cdots a_{i\pi(j)}\cdots a_{n\pi(n)} =$$

$$- \sum_{\pi\in Sym([n])} (-1)^{inv(\pi')} a_{1\pi'(1)}\cdots a_{i\pi'(i)}\cdots a_{j\pi'(j)}\cdots a_{n\pi'(n)} =$$

$$- \sum_{\pi'\in Sym([n])} (-1)^{inv(\pi')} a_{1\pi'(1)}\cdots a_{i\pi'(i)}\cdots a_{j\pi'(j)}\cdots a_{n\pi'(n)} = -J(a_1,\dots,a_n).$$

  Az utolsó előtti lépésben annyi történt, hogy π helyett π'-re összegeztünk. Mivel a $\pi \mapsto \pi'$ függvény bijekció $Sym([n])$-ről sajátmagába, ezért a két sorban a szummák belsejében pontosan ugyanazok az összeadandók jelennek meg, csak más sorrendben. Ezért az összegük tényleg egyenlő; az utolsó előtti lépés is helyes.
- 8.4 (4) teljesül, mert:

$$J(e_1,\dots,e_n) = \sum_{\pi\in Sym([n])} (-1)^{inv(\pi)} e_{1\pi(1)}\cdots e_{n\pi(n)}.$$

  De tetszőleges i-re az e_i egységvektornak csak egyetlen koordinátája (éppen az i-edik koordinátája) különbözik nullától. Ha tehát $\pi(1) \neq 1$ vagy $\pi(2) \neq 2$ vagy ... vagy $\pi(n) \neq n$, akkor $e_{1\pi(1)}\cdots e_{n\pi(n)}$ mindig nulla lesz. Egyetlen permutáció esetén nem lesz nulla ez a szorzat: akkor, ha $\pi(1) = 1, \pi(2) = 2, \dots, \pi(n) = n$, és ekkor a kérdéses szorzat minden tényezője 1 lesz. Ezért $J(e_1,\dots,e_n) = 1$.

□

8.6. Definíció (Determináns)**.** *Adott $n \in \mathbb{N}$-re a 8.5 Tétel szerint egyértelműen létező determináns függvényt* det*-tel jelöljük:*

$$\det(A) = \sum_{\pi\in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}.$$

8.2. A determináns kiszámítása

8.7. Tétel. $\forall A \in K^{n\times n}$: $\det(A) = \det(A^T)$.

Bizonyítás. Vegyük észre: ha $\pi \in Sym([n])$ egy permutáció, akkor

$$\mathrm{inv}(\pi) = \mathrm{inv}(\pi^{-1})$$

mert tetszőleges $i, j \leq n$-re

i, j inverzióban van π-ben, $\Leftrightarrow$
$i < j$, de $\pi(i) > \pi(j) \Leftrightarrow$
$\pi(j) < \pi(i)$, de $\pi^{-1}(\pi(i)) < \pi^{-1}(\pi(j)), \Leftrightarrow$
$\pi(j), \pi(i)$ inverzióban van π^{-1}-ben.

Legyen $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$. Figyeljük meg, hogy tetszőleges $\pi \in Sym([n])$-re

$$a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)} = a_{\pi^{-1}(1)1}a_{\pi^{-1}(2)2}\cdots a_{\pi^{-1}(n)n}$$

(hiszen ugyanazokat a tényezőket szorozzuk, csak más sorrendben). Ekkor

$$\det(A) = \sum_{\pi\in Sym([n]} (-1)^{inv(\pi)} a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)} =$$

$$\sum_{\pi\in Sym([n]} (-1)^{inv(\pi^{-1})} a_{\pi^{-1}(1)1}a_{\pi^{-1}(2)2}\cdots a_{\pi^{-1}(n)n} =$$

$$\sum_{\pi^{-1}\in Sym([n]} (-1)^{inv(\pi^{-1})} a_{\pi(1)1}a_{\pi(2)2}\cdots a_{\pi(n)n} = \det(A^T).$$

□

Ha n értéke kicsi, akkor $[n]$ összes permutációit könnyen át tudjuk tekinteni. Ennek segítségével n kicsi értékeire képleteket fogunk megadni a determináns kiszámítására.

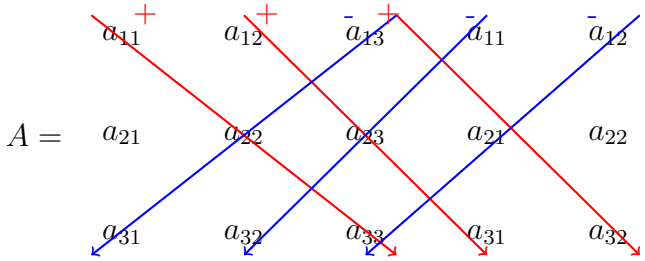
- $n = 1$: $\det(a) = a$, hiszen [1]-nek egyetlen permutációja van, melyben nincs inverzió. Ennek megfelelően, a 8.6 Definícióban szereplő szumma 1-tagú. A $\det(a) = a$ képlet összhangban van a geometriai szemléletünkkel is: az 1-dimenziós paralelepipedonok szakaszok; oldalélük hossza (vagyis a) egyúttal az 1-dimenziós térfogatuk is.
- $n = 2$: $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$, hiszen [2]-nek két permutációja van: $(0, 1)$ és $(1, 0)$. Az elsőben nincs inverzió, a második inverziószáma pedig 1. Ezért, az $n = 2$ esetben a 8.6 Definícióban szereplő szumma két összeadandóból áll, és

értéke $ad - bc$. A 2025 nov. 21.-i előadáson mutattam geometriai bizonyítást arra, hogy a síkon az (origóból induló) (a, b) és (c, d) pontokba mutató vektorok által kifeszített parallelogramma területe (azaz 2-dimenziós térfogata) valóban $ad - bc$ (erről a hallgatóktól elektronikus anyagot nem kaptam; nem nagy baj, a vizsgára e területképlet geometriai biznyítását úgysem kell tudni).

A Sarrus-szabály egy módszer 3×3-as mátrixok determinánsának kiszámítására. Ez is azon múlik, hogy a [3] halmaz 6 darab permutációját még mindig viszonylag könnyű áttekinteni. **A Sarrus-szabály 3×3-asnál nagyobb mátrixokra nem marad érvényben**, illetve hasonló, könnyen megjegyezhető, egyszerű módszerek nincsenek. Tehát a Sarrus-szabály:

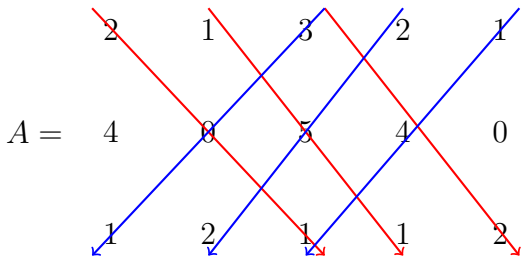
1. Írjuk le a mátrixot, majd az első két oszlopot ismételjük meg jobb oldalon;
2. A **pozitív tagok** a főátlóval párhuzamosan balról jobbra haladó nyilak mentén;
3. A **negatív tagok** a mellékátlóval párhuzamosan jobbról balra haladó nyilak mentén.

$$\det(A) = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}).$$



8.8. Példa. *Számítsuk ki a következő mátrix determinánsát:*

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 1 & 2 & 1 \end{pmatrix}.$$



$$\begin{aligned} \det(A) &= (2 \cdot 0 \cdot 1 + 1 \cdot 5 \cdot 1 + 3 \cdot 4 \cdot 2) - (3 \cdot 0 \cdot 1 + 2 \cdot 5 \cdot 2 + 1 \cdot 4 \cdot 1) \\ &= (0 + 5 + 24) - (0 + 20 + 4) = 29 - 24 = 5. \end{aligned}$$

Sorátalakítások és determináns

3×3-asnál nagyobb mátrixok determinánsait úgy is meghatározhatjuk, hogy speciális alakúra hozzuk őket. Az átalakítások (pl. elemi sor- és oszlopátalakítások) során kontrolláljuk, hogyan változik a determináns, és bizonyos speciális alakú (pl. az alább definiált alsó- vagy felső-háromszög) mátrixok determinánsait könnyen meg fogjuk tudni határozni.

8.9. Tétel (Elemi sorátalakítások hatása)**.**

1. *Ha az i-edik sort c-vel szorozzuk: a determináns c-szeresére változik;*
2. *Ha az i-edik és j-edik sort felcseréljük: a determináns -1-szeresére változik;*
3. *Ha van 2 egyenlő sor, akkor a determináns értéke* 0*;*
4. *Ha egyik sor c-szeresét a másikhoz adjuk: a determináns értéke nem változik.*

Bizonyítás. (1) és (2) azonnal következik a 8.4 Definícióból, azért illesztettük ezeket is állításaink közé, hogy minden elemi sorátalakítás hatását megadjuk.

(3)-at (picit kevésbé általános formában) megfigyeltük már a 8.5 Tétel unicitásra vonatkozó részének 1. esetében. A teljesség (és a teljes általánosság) kedvéért megismételjük ugyanazt a gondolatot: tegyük fel, hogy valamilyen $i \neq j$-re mátrixunk i-edik és j-edik sora egyenlő. Ekkor egyrészt, az i-edik és j-edik sort felcserélve nem történik változás, másrészt 8.4 (3) miatt a determináns értéke -1-szeresére változik. Mátrixunk determinánsa tehát egyenlő sajátmaga (-1)-szeresével; ez csak úgy lehet, ha a determináns érétke 0.

$$(4): \qquad \det \begin{pmatrix} \vdots \\ a_i + ca_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{8.4\ (1),(2)}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + c \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{(3)}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + 0. \qquad \square$$

8.10. Megjegyzés. Az elemi oszlopműveletekre ugyanúgy változik a determináns, mint a megfelelő sorműveletekre, mert a 8.7 Tétel szerint tetszőleges (négyzetes) mátrixnak és transzponáltjának ugyanannyi a determinánsa.

8.11. Definíció (Háromszögmátrix)**.**
*$A \in K^{n\times n}$ **felső háromszögmátrix**, ha a főátló alatti elemek 0-k.*
*Hasonlóan, A **alsó háromszögmátrix**, ha a főátló feletti elemek 0-k.*

8.12. Tétel (Háromszögmátrix determinánsa)**.** *Ha $A \in K^{n\times n}$ felső (vagy alsó) háromszögmátrix, akkor a determinánsa a főátló elemeinek szorzata:*

$$\det(A) = a_{11}a_{22}\cdots a_{nn}.$$

Bizonyítás. Jelöljük *id*-el az $(1, 2, 3, \ldots, n)$ permutációt (*id* az „identitás”-t rövidíti arra utalva, hogy ez a permutáció minden elemet fixen hagy (nem csinál semmit)).

Felső háromszögre: ha $\pi \in Sym([n])$, $\pi \neq id$, akkor van i, melyre $\pi(i) < i$. Egy ilyen i-re $a_{i\pi(i)} = 0$, mert A felső háromszög. Egyrészt

$$\det(A) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)},$$

másrészt a bizonyítás eleje szerint, ha $\pi \neq id$, akkor az $a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$ szorzatban valamelyik tényező nulla, ezért a szorzat is nulla. Tehát $\det(A)$ szummás alakjában mindegyik, $\pi \neq id$-hez tartozó összeadandó nulla, így csak $\pi = \mathrm{id}$ marad: $\det(A) = a_{11} a_{22} \cdots a_{nn}$. □

8.13. Példa.

$$\det \begin{pmatrix} 2 & 1 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} = 2 \cdot 4 \cdot 6 = 48.$$

8.3. A determináns kiszámítása kifejtéssel

8.14. Definíció (Aldeterminány)**.** *Ha $A \in K^{n \times n}$ és $1 \leq i, j \leq n$, akkor $A_{ij} \in K^{(n-1)\times(n-1)}$ az a mátrix, melyet A-ból úgy kapunk, hogy A-ból kitöröljük az i-edik sort és j-edik oszlopot.*

8.15. Definíció (Kifejtések)**.**

- *A i-edik sor szerinti kifejtése:*

$$SK_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij});$$

- *A j-edik oszlop szerinti kifejtése:*

$$OK_j(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Az előbbi összegekben megjelenő $(-1)^{i+j}$ előjeleket a sakktábla-szabály segít memorizálni: a mátrix első sorának első elemét pozitív előjellel látjuk el, a többi elemet pedig úgy, hogy előjele különbözzön a tőle balra, illetve a felette elhelyezkedő elemhez társított előjeltől.

8.16. Tétel (Kifejtési tétel)**.** $\forall i, j$*:* $\det(A) = SK_i(A) = OK_j(A)$.

Bizonyítás. $SK_i(A) = \det(A)$ a következők miatt: A i-edik sorát felírhatjuk $\sum_{j=1}^{n} a_{ij} e_j$ alakban, ahol $a_{ij} e_j = [0, \ldots, 0, a_{ij}, 0, \ldots, 0]$, tehát a jobboldali összeadandó vektorok mind olyanok, hogy legfeljebb 1 koordinátájuk nem nulla. Ezt figyelembe véve

$$\det(A) \overset{8.4\ (2)}{=} \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & 0 \cdots 0 & 0 \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + \cdots + \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ 0 & 0 \cdots 0 & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

ezért elég belátni, hogy tetszőleges j-re

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & \ddots & \vdots & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} = (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Az i-edik sort mozgassuk az 1. sorba: ez $(i-1)$ db sorcsere, ezzel a determináns értéke $(-1)^{i-1}$-szeresére változik. Ezután a j-edik oszlopot mozgassuk az 1. oszlopba: ez újabb $(j-1)$ db oszlopcsere, ezzel a determináns előjele tovább változik az előző $(-1)^{j-1}$-szeresére, az előjel összesen $(-1)^{i+j-2} = (-1)^{i+j}$-szeresére változott. A sor- és oszlopcserékkel azt kaptuk, hogy

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & \ddots & \vdots & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} = (-1)^{i+j} \det\begin{pmatrix} a_{ij} & 0 & \cdots & 0 \\ a_{1j} & a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \\ \vdots & \ddots & \vdots & \\ a_{nj} & \cdots & \cdots & a_{nn} \end{pmatrix},$$

az előző sor jobboldalán álló mátrixot jelöljük B-vel. B jobb alsó $(n-1)\times(n-1)$-es része épp A_{ij}. Ekkor

$$\det(B) = \sum_{\pi \in Sym([n])} (-1)^{inv(\pi)} b_{1\pi(1)} \cdots b_{n\pi(n)}.$$

De B első sorában az első elem kivételével minden elem 0, ezért az előbbi összegben minden olyan összeadandó 0 lesz, melyben $\pi(1) \neq 1$. Ezért elég az $(n-1)$-elemű $\{2, \ldots, n\}$ halmaz permutációira összegezni:

$$\det(B) = a_{ij} \sum_{\pi \in Sym([n]-\{1\})} (-1)^{inv(\pi)} b_{2\pi(2)} \cdots b_{n\pi(n)} = a_{ij} \det(A_{ij}),$$

és emiatt valóban

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Az $OK_j(A) = \det(A)$ állítást hasonlóan, vagy akár a transzponált determinánsára vonatkozó 8.7 Tétellel lehet elintézni. □

9. További tudnivalók, alkalmazások

9.1. Rang, inverz, determináns néhány tulajdonsága

9.1. Tétel (Rang és mátrixműveletek)**.** *Legyenek* $A, B \in K^{m\times n}$.

1. $rang(A+B) \le rang(A) + rang(B)$;
2. $rang(AB) \le \min\{rang(A), rang(B)\}$.

Bizonyítás. (1): Legyen $\{u_1, \dots, u_{\mathrm{rang(A)}}\}$ bázis $\mathcal{O}(A)$-ban és $\{v_1, \dots, v_{\mathrm{rang(B)}}\}$ bázis $\mathcal{O}(B)$-ben. Ekkor $\{u_1, \dots, a_{\mathrm{rang(A)}}, v_1, \dots, v_{\mathrm{rang(B)}}\}$ generátorrendszere $\mathcal{O}(A+B)$-nek, így:

$$\mathrm{rang}(A+B) = \dim \mathcal{O}(A+B) \le \mathrm{rang}(A) + \mathrm{rang}(B).$$

(2): AB j-edik oszlopa:

$$(AB)_{*j} = \sum_{k=1}^{n} b_{kj} A_{*k}.$$

Tehát AB oszlopai A oszlopainak lineáris kombinációi, ezért $\mathcal{O}(AB) \subseteq \mathcal{O}(A)$ emiatt

$$\mathrm{rang}(AB) = \dim \mathcal{O}(AB) \le \dim \mathcal{O}(A) = \mathrm{rang}(A).$$

Hasonlóan: AB sorai B sorainak lineáris kombinációi, ezért $\mathcal{S}(AB) \subseteq \mathcal{S}(B)$ és

$$\mathrm{rang}(AB) = \dim \mathcal{S}(AB) \le \dim \mathcal{S}(B) = \mathrm{rang}(B).$$

□

Inverz mátrix tulajdonságai

9.2. Tétel. *Legyenek* $A, B, C, D \in K^{n\times n}$*, és tegyük fel, hogy* A *és* B *invertálható mátrixok. Ekkor:*

1. $(A^{-1})^{-1} = A$*, továbbá* $(C^T)^T = C$;
2. $(AB)^{-1} = B^{-1}A^{-1}$*, továbbá* $(CD)^T = D^T C^T$;
3. *Minden* $k \in \mathbb{N}$ *esetén* $(A^k)^{-1} = (A^{-1})^k$*, továbbá* $(C^k)^T = (C^T)^k$.

Bizonyítás.

1. Mivel
$$A \cdot A^{-1} = A^{-1} \cdot A = I_n,$$
ezért A^{-1} inverze éppen A, azaz $(A^{-1})^{-1} = A$. A transzponálás definíciójából közvetlenül következik, hogy $(C^T)^T = C$.
2. Vizsgáljuk:
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n.$$
Hasonlóan:
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_n.$$

Tehát $B^{-1}A^{-1}$ valóban az AB inverze, azaz $(AB)^{-1} = B^{-1}A^{-1}$.
Most legyen $i, j \leq n$ tetszőleges. Ekkor

$$((CD)^T)_{ij} = (CD)_{ji} = \sum_{k=1}^{n} C_{jk}D_{ki}.$$

Másrészt:

$$(D^TC^T)_{ij} = \sum_{k=1}^{n} (D^T)_{ik}(C^T)_{kj} = \sum_{k=1}^{n} D_{ki}C_{jk}.$$

A két összeg megegyezik, tehát $(CD)^T = D^TC^T$.

3. Indukcióval bizonyítjuk $(A^k)^{-1} = (A^{-1})^k$-t.
Alaplépés: $k = 1$ esetén triviális.
Indukciós lépés: Tegyük fel, hogy $(A^k)^{-1} = (A^{-1})^k$. Ekkor:

$$(A^{k+1})^{-1} = (A^kA)^{-1} = A^{-1}(A^k)^{-1} = A^{-1}(A^{-1})^k = (A^{-1})^{k+1}.$$

A transzponálásra vonatkozó állítás hasonló módon igazolható.

□

Négyzetes mátrixok invertálhatóságának ekvivalens feltételei

9.3. Tétel. *$A \in K^{n\times n}$ esetén ekvivalensek:*

1. *rang$(A) = n$ (maximális rang);*
2. *$\dim(\mathcal{O}(A)) = n$ (oszloptér teljes dimenziós);*
3. *$\dim(\mathcal{S}(A)) = n$ (sortér teljes dimenziós);*
4. *A invertálható;*
5. *$\det(A) \neq 0$.*

Bizonyítás. (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3): A 7.47 Definíció szerint $\text{rang}(A) = \dim(\mathcal{O}(A))$ és a 7.46 Tétel szerint $\dim(\mathcal{O}(A)) = \dim(\mathcal{S}(A))$.

(1) $\Leftrightarrow$ (4): Ez a 7.53 Tétel (egy részének) átfogalmazása.

(4) $\Rightarrow$ (5): Ha A invertálható, akkor $A^{-1}A = I_n$, így a Determinánsok szorzástétele (alábbi 9.4 Tétel) szerint:

$$1 = \det(I_n) = \det(A^{-1}A) = \det(A^{-1})\det(A) \Rightarrow \det(A) \neq 0.$$

(5) $\Rightarrow$ (1): Hozzuk A-t redukált lépcsős alakra elemi sorátalakításokkal. A 8.9 Tétel szerint mindegyik elemi sorátalakítás nemnulla determinánsú mátrixból nemnulla determinánsú mátrixot csinál. Ezért, a redukált lépcsős alakot L-el jelölve, ha $\det(A) \neq 0$, akkor $\det(L) \neq 0$, tehát L főátlójában nincs 0 elem. Így L minden oszlopában van vezérelem, ezért (a 7.43 Tétel alapján) $\text{rang}(A) = n$. □

Determinánsok szorzástétele

9.4. Tétel (Determinánsok szorzástétele)**.** *$A, B \in K^{n\times n}$ esetén* $\det(AB) = \det(A)\det(B)$.

Bizonyítás. 1. eset: $\det(A) = 0$. Ekkor $\operatorname{rang}(A) < n$, ezért a 9.1 Tétel miatt $\operatorname{rang}(AB) \leq \operatorname{rang}(A) < n$, tehát pl. a 9.3 Tétel miatt $\det(AB) = 0$.
2. eset: $\det(A) \neq 0$. Tekintsük a következő függvényt:

$$f(B) = \frac{\det(AB)}{\det(A)}.$$

Ellenőrizhető, hogy f determináns függvény (azaz teljesülnek rá a 8.4 Definíció kitételei). Pl. $f(I_n) = \frac{\det(A)}{\det(A)} = 1$, stb. Mivel a 8.5 Tétel szerint csak egy determináns függvény van $K^{n\times n}$-en, ezért $f(B) = \det(B)$. Tehát:

$$\frac{\det(AB)}{\det(A)} = \det(B) \Rightarrow \det(AB) = \det(A)\det(B).$$

□

9.2. Bázistranszformáció

Ebben az alfejezetben azt vizsgáljuk, hogy ha egy (véges dimenziós) vektortérben adott két bázis, és ismerjük egy vektor koordinátáit az egyik bázisban, akkor hogyan határozhatjuk meg **ugyanannak** a vektornak a koordinátáit a másik bázisban, vagyis hogyan térhetünk át egyik bázisból a másikba. Könnyű meggondolni (de az alábbi, 9.5 Tételben mindjárt be is látjuk), hogy az a függvény, amely a vektorok egyik bázisban adott koordinátáihoz a másik bázisbeli koordinátáit rendeli, lineáris. Ezért a koordináták bázisok közti átszámolása mátrixszorzással megoldható; a megfelelő mátrixot meg is adjuk a 9.5 tételben.

9.5. Tétel. *Mint eddig is, legyen V vektortér a K test felett. Tegyük fel, hogy $B = \{b_1, \ldots, b_n\} \subseteq V$ és $D = \{d_1, \ldots, d_n\} \subseteq V$ bázisok V-ben.*

1. *A B-ről D-re való áttérés mátrixa*

$$T_{D\leftarrow B} = \big([b_1]_D \ \ldots \ [b_n]_D\big),$$

azaz az oszlopvektorok a b_j vektorok D-beli koordinátái. Ez azt jelenti, hogy Minden $v \in V$ esetén:

$$[v]_D = T_{D\leftarrow B}\,[v]_B.$$

2. $T^{-1}_{D\leftarrow B} = T_{B\leftarrow D}$.

Bizonyítás.

1. Legyen

$$v = [v_1 \ \dots \ v_n]_B, \qquad T_{D\leftarrow B} = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ \vdots & & \vdots \\ t_{n1} & \dots & t_{nn} \end{pmatrix}.$$

Ekkor a mátrix j-edik oszlopa:

$$(T_{D\leftarrow B})_{*j} = \begin{pmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{pmatrix} = [b_j]_D.$$

Mivel

$$v = \sum_{j=1}^{n} v_j b_j \quad \text{és} \quad b_j = \sum_{k=1}^{n} t_{kj} d_k,$$

ezért

$$v = \sum_{j=1}^{n} v_j \sum_{k=1}^{n} t_{kj} d_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} v_j t_{kj} \right) d_k.$$

Így

$$[v]_D = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ \vdots & & \vdots \\ t_{n1} & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = T_{D\leftarrow B}\,[v]_B.$$

2. Elég belátni, hogy

$$T_{B\leftarrow D}\,T_{D\leftarrow B} = I_n.$$

Legyen $E = \{e_1, \dots, e_n\}$ a sztenderd bázis, ahol

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}.$$

Ekkor:

$$(T_{B\leftarrow D}\,T_{D\leftarrow B})e_j = T_{B\leftarrow D}(T_{D\leftarrow B}e_j) = T_{B\leftarrow D}[b_j]_D = [b_j]_B = e_j.$$

Mivel ez minden j-re teljesül, ezért

$$T_{B\leftarrow D}\,T_{D\leftarrow B} = I_n,$$

tehát $T_{D\leftarrow B}^{-1} = T_{B\leftarrow D}$.

□

9.6. Tétel (Lineáris leképezés mátrixának megváltozása bázisváltáskor)**.**
Legyen $\varphi\colon V \to V$ lineáris leképezés, és $A \in K^{n\times n}$ a φ mátrixa a sztenderd bázisban. Legyen $B = \{b_1, \dots, b_n\} \subseteq V$ egy bázis, $E = \{e_1, \dots, e_n\}$ a sztenderd bázis, és $T = T_{E\leftarrow B}$. Ekkor minden $v \in V$ esetén:

$$[\varphi(v)]_B = (T^{-1}AT)\,[v]_B.$$

Bizonyítás.

$$[T^{-1}AT]\,[v]_B = T^{-1}A[T[v]_B] = T^{-1}A[v]_E = T^{-1}[\varphi(v)]_E = [\varphi(v)]_B.$$

□

9.7. Példa. *Legyen $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ egy bázis $\mathbb{R}^2$-ben, ahol $\mathbf{b}_1 = \begin{bmatrix}1\\0\end{bmatrix}$ és $\mathbf{b}_2 = \begin{bmatrix}1\\1\end{bmatrix}$. Adjuk meg a $T_{\mathcal{B}\leftarrow\mathcal{E}}$ mátrixot, ahol $\mathcal{E}$ a sztenderd bázis.*

1. Megoldás. *A bázisváltás mátrix oszlopai a régi bázis ($\mathcal{E}$) vektorainak az új bázis ($\mathcal{B}$) szerinti koordinátái.*

$$\mathbf{e}_1 = 1\cdot\mathbf{b}_1 + 0\cdot\mathbf{b}_2 \quad\Rightarrow\quad [\mathbf{e}_1]_{\mathcal{B}} = \begin{bmatrix}1\\0\end{bmatrix},$$
$$\mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix} = (-1)\cdot\begin{bmatrix}1\\0\end{bmatrix} + 1\cdot\begin{bmatrix}1\\1\end{bmatrix} = (-1)\mathbf{b}_1 + 1\mathbf{b}_2 \quad\Rightarrow\quad [\mathbf{e}_2]_{\mathcal{B}} = \begin{bmatrix}-1\\1\end{bmatrix}.$$

Tehát:

$$T_{\mathcal{B}\leftarrow\mathcal{E}} = \begin{bmatrix}[\mathbf{e}_1]_{\mathcal{B}} & [\mathbf{e}_2]_{\mathcal{B}}\end{bmatrix} = \begin{bmatrix}1 & -1\\0 & 1\end{bmatrix}.$$

9.3. Polinom-interpoláció

9.8. Definíció (Interpolációs feladat)**.** Adottak a páronként különböző alappontok: $a_0, a_1, \dots, a_n \in K$ és a hozzájuk tartozó értékek: $b_0, b_1, \dots, b_n \in K$. Az interpolációs feladat: Keressünk olyan $f \in K[x]$ polinomot, amelyre:

$$f(a_0) = b_0, \quad f(a_1) = b_1, \quad \dots, \quad f(a_n) = b_n.$$

9.9. Definíció (Vandermonde-mátrix)**.** *Legyenek $a_0, a_1, \dots, a_n \in K$ skalárok. A hozzájuk tartozó* ***Vandermonde-mátrix****:*

$$V(a_0, a_1, \dots, a_n) = \begin{bmatrix}1 & a_0 & a_0^2 & \dots & a_0^n\\1 & a_1 & a_1^2 & \dots & a_1^n\\\vdots & \vdots & \vdots & \ddots & \vdots\\1 & a_n & a_n^2 & \dots & a_n^n\end{bmatrix} \in K^{(n+1)\times(n+1)}.$$

9.10. Tétel (Vandermonde-mátrix determinánsa)**.**

1. *A $V(a_0, a_1, \dots, a_n)$ Vandermonde-mátrix determinánsa akkor és csak akkor* 0*, ha generáló-elemei között vannak egyenlők (azaz van $i \neq j : a_i = a_j$).*

2. *Sőt, a Vandermonde-mátrix determinánsára érvényes:*

$$\det\big(V(a_0, a_1, \dots, a_n)\big) = \prod_{0 \leq i < j \leq n} (a_j - a_i).$$

Ebből is adódik: $\det(V(a_0, a_1, \dots, a_n)) = 0$ *pontosan akkor, ha létezik $i \neq j$ indexpár úgy, hogy $a_i = a_j$.*

Előadáson én 1.-et bizonyítottam be (aminek van egy viszonylag rövid és csinos bizonyítása), 2.-t a 12. heti gyakorlat 1. feladataként tűztem ki. A hallgatók által küldött anyagban ki volt mondva 2. is, ezért benne hagytam. Vizsgán elég 1.-et tudni...

Bizonyítás. (1) rövid bizonyítása a következő. $\Leftarrow$: ha valamely $i \neq j$-re $a_i = a_j$, akkor $V(a_0, a_1, \dots, a_n)$-nak van két egyenlő sora, ezért $\det\big(V(a_0, a_1, \dots, a_n)\big) = 0$ (pl. a 8.9 Tétel 3. pontja miatt).

$\Rightarrow$: Tegyük fel: $\det\big(V(a_0, a_1, \dots, a_n)\big) = 0$. Ekkor a 9.3 Tétel miatt $V(a_0, a_1, \dots, a_n)$ oszlopai lineárisan összefüggők: van nem csupa-nulla $\lambda_0, \dots \lambda_n \in K$:

$$(*) \qquad 0 = \lambda_0 \begin{pmatrix} a_0^0 \\ a_1^0 \\ \vdots \\ a_n^0 \end{pmatrix} + \lambda_1 \begin{pmatrix} a_0^1 \\ a_1^1 \\ \vdots \\ a_n^1 \end{pmatrix} + \cdots \lambda_n \begin{pmatrix} a_0^n \\ a_1^n \\ \vdots \\ a_n^n \end{pmatrix}.$$

Legyen $f(x) = \lambda_0 + \lambda_0 x + \dots \lambda_n x^n$, ez nem a konstans-0 polinom, mert valamelyik $\lambda_k \neq 0$. Továbbá, $(*)$ első sora szerint a_0 gyöke az f polinomnak, $(*)$ második sora szerint a_1 is gyöke f-nek, ... $(*)$ utolsó sora szerint a_n is gyöke f-nek. Ez összesen összesen $n+1$ db gyök, de f legfeljebb n-edfokú. Mivel a 6.21 Következmény szerint egy legfeljebb n-edfokú polinomnak legfeljebb n különböző gyöke lehet, ezért az $a_0, a_1, \dots a_n$ gyökök között vannak egyenlők. □

Lagrange-interpoláció

9.11. Tétel (Lagrange Interpolációs Tétele)**.** *Ha az $(a_0, \dots, a_n)$ alappontok páronként különbözők, akkor pontosan 1 darab n-edfokú $f \in K[x]$ van, amely megoldja az interpolációs feladatot.*

Bizonyítás. Legyen adott még $b_0, b_1, \dots, b_n \in K$, olyan f polinomot kell találnunk, melyre $f(a_0) = b_0, \dots f(a_n) = b_n$. Vegyük fel f-et határozatlan (egyelőre ismeretlen) együtthatókkal:

$$f(x) = c_0 + c_1 x + \cdots + c_n x^n.$$

Mivel

$$\begin{aligned}
f(a_0) = b_0 \quad &\Leftrightarrow \quad c_0 + c_1 a_0 + \cdots + c_n a_0^n = b_0, \\
f(a_1) = b_1 \quad &\Leftrightarrow \quad c_0 + c_1 a_1 + \cdots + c_n a_1^n = b_1, \\
&\cdots \\
f(a_n) = b_n \quad &\Leftrightarrow \quad c_0 + c_1 a_n + \cdots + c_n a_n^n = b_n,
\end{aligned}$$

ezért f akkor és csak akkor megoldása az interpolációs feladatnak, ha együtthatói kielégítik az előző lineáris egyenletrendszert. Ezt az egyenletrendszert mátrix-alakban írva:

$$(*) \qquad \begin{bmatrix} 1 & a_0 & a_0^2 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & \dots & a_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^n \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Az egyenletrendszer mátrixa a $V(a_0, a_1, \dots, a_n)$ Vandermonde-mátrix. Továbbá, mivel feltettük, hogy az $(a_0, \dots, a_n)$ alappontok páronként különbözők, ezért a 9.10 Tétel miatt $\det\big(V(a_0, a_1, \dots, a_n)\big) \neq 0$. Ezért pl. a 9.3 Tétel miatt $V(a_0, a_1, \dots, a_n)$ invertálható és emiatt $(*)$ egyetlen megoldása

$$\big(V(a_0, a_1, \dots, a_n)\big)^{-1} \cdot \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Azt kaptuk, hogy egy legfeljebb n-edfokú polinom együtthatóit pontosan egyféleképpen választhatjuk meg, ha a polinom megoldása az interpolációs feladatunknak. □

9.12. Példa. *Keressünk másodfokú $g \in \mathbb{R}[x]$ polinomot, amelyre $g(0) = 2$, $g(1) = 4$ és $g(2) = 8$.*

2. Megoldás. *Legyen $g(x) = \mu_0 + \mu_1 x + \mu_2 x^2$. Felírjuk az egyenletrendszert:*

$$\begin{cases} g(0) = \mu_0 = 2, \\ g(1) = \mu_0 + \mu_1 + \mu_2 = 4, \\ g(2) = \mu_0 + 2\mu_1 + 4\mu_2 = 8. \end{cases}$$

Mátrix alakban:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}.$$

Ennek megoldása: $\mu_0 = 2$, $\mu_1 = 1$, $\mu_2 = 1$. Tehát a keresett polinom:

$$g(x) = 2 + x + x^2.$$

Alkalmazás: Titokmegosztás

Tekintsük a következő feladatot: adott egy s titok (pl. egy szám, vagy egy rögzített K test egy eleme). Van n titokgazda. A rendszer úgy működjön, hogy legalább k darab titokgazdának ($k \leq n$) együtt kell működnie a titok visszanyeréséhez, de kevesebb, mint k titokgazda ne tudjon semmit a titokról, akkor sem, ha összejátszanak.

3. Megoldás (Shamir k-küszöbös titokmegosztási módszere)**.** Legyen $s \in K$ a titok.

1. Válasszunk véletlenszerűen egy $f \in K[x]$ polinomot, melyre $\deg(f) \leq k-1$ és $f(0) = s$. (Ezt megtehetjük, ha véletlenül választjuk a $b_1, \ldots, b_{k-1}$ értékeket, majd megoldjuk a $(0, s),\ (1, b_1), (2, b_2), \ldots, (k-1, b_{k-1})$ interpolációs feladatot).
2. $(i = 1, \ldots, n)$-re az i-edik titokgazdával közöljük az $(i, f(i))$ párt (vagyis egy $x = i$ pontbeli függvényértéket).
3. **Titok visszanyerése:** Mivel Lagrange interpolációs tétele (9.11 Tétel) miatt k pont egyértelműen meghatároz egy legfeljebb $(k-1)$-edfokú polinomot, ezért ha legalább k titokgazda összejön, az ő $(i, f(i))$ pontjaik alapján interpolációval rekonstruálhatják f-et, majd kiolvashatják a titkot: $s = f(0)$.
4. **Biztonság:** Kevesebb, mint k titokgazda birtokában az f polinomra sok lehetőség marad mert bármelyik s'-re megoldható a $(0, s'),\ (i_1, b_{i_1}), (i_2, b_{i_2}), \ldots, (i_{k-1}, b_{i_{k-1}})$ interpolációs feladat is, ezek a megoldások a $k-1$ darab konspiráló titokgazda számára megkülönböztethetetlenek. Ezért számukra a titok (vagyis $f(0)$) teljesen bizonytalan.

Másik konstrukció az interpolációs polinomra

Az interpolációs feladat megoldására adunk egy másik módszert it.

9.13. Definíció (Lagrange-féle alappolinomok)**.** *Legyenek $a_1, \ldots, a_n \in K$ páronként különböző alappontok. Ha $i \leq n$, akkor az i-edik Lagrange-féle alappolinom az a legfeljebb $n-1$-edfokú $\ell_i \in K[x]$, melyre*

$$\ell_i(a_0) = 0, \ldots \ell_i(a_i) = 1, \ldots \ell_i(a_n) = 0$$

($\ell_i(a_i) = 1$, és ha $i \neq j$, akkor $\ell_i(a_j) = 0$). Ilyen ℓ_i van:

$$\ell_i(x) = \frac{(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)} \in K[x].$$

9.14. Tétel (Interpolációs polinom Lagrange-alakja)**.** *A $g(a_i) = b_i$ $(i = 1, \ldots, n)$ interpolációs feltételeket kielégítő, legfeljebb $(n-1)$-edfokú polinom egyértelműen megadható a Lagrange-alappolinomok lineáris kombinációjaként:*

$$g(x) = \sum_{i=1}^{n} b_i\, \ell_i(x).$$

Bizonyítás. Ez valóban megoldás, mert $g(a_j) = \sum_{i=1}^{n} b_i \ell_i(a_j) = b_j$. □

Mivel n adott alappont esetén az interpolációs feladatot pontosan 1 legfeljebb $n-1$-edfokú f polinom oldja meg, ezért a 9.11 Tételben megadott polinom ugyanaz, mint amit a 9.14 Tételben konstruáltunk. Valójában– megelőlegezve a 9.15 Tételt – arról van szó, hogy f együtthatóit a legfeljebb $n-1$-edfokú polinomok vektorterében más és más bázisban írtuk fel. A 9.11 Tétel bizonyításában a hatványfüggvények alkotta szokásos

$$\{1, x, x^2, \ldots, x^{n-1}\}$$

bázist használtuk és f együtthatóit egy egyenletrendszer megoldásának árán tudtuk meghatározni. A 9.14 Tételben az $\{\ell_1, \ldots \ell_n\}$ bázist használtuk. Ebben a bázisban f koordinátái villámgyorsan leolvashatók; ha f együtthatóit (vagyis koordinátáit a hatványfüggvények alkotta bázisban) is meg akarjuk határozni, akkor az f-re kapott kifejezést x hatványai szerint rendeznünk kell (illetve át kell térnünk az egyik bázisból a másikba). Azzal folytatjuk, hogy az imént említett polinomhalmazok valóban bázisok a megfelelő vektortérben.

A Lagrange-alappolinomok bázist alkotnak

9.15. Tétel. *Legyen $V = \{f \in K[x] \mid \deg(f) < n\}$. Ekkor V n-dimenziós vektortér K felett, és a következő halmazok mind bázisai V-nek:*

1. $\mathcal{B}_1 = \{\ell_1, \ell_2, \ldots, \ell_n\}$ *(Lagrange-alappolinomok),*
2. $\mathcal{B}_2 = \{1, x, x^2, \ldots, x^{n-1}\}$ *(sztenderd bázis).*

Bizonyítás.

1. **Linearitás függetlenség:** Tegyük fel, hogy $\sum_{i=1}^{n} c_i \ell_i(x) = 0$. Helyettesítsünk $x = a_j$-t:
$$0 = \sum_{i=1}^{n} c_i \ell_i(a_j) = c_j \quad \text{minden } j = 1, \ldots, n\text{-re.}$$
Tehát $c_1 = \cdots = c_n = 0$, a Lagrange-alappolinomok lineárisan függetlenek. Mivel $\dim V = n$, ezért bázist alkotnak.
2. **Sztenderd bázis:** A $\{1, x, \ldots, x^{n-1}\}$ halmaz nyilván generálja V-t. Lineáris függetlenségüket így láthatjuk be: legyen $c_0, \ldots, c_{n-1} \in K$ tetszőleges, és vegyük $\{1, x, \ldots, x^{n-1}\}$-nek a $c_0, \ldots, c_{n-1}$ súlyokkal képzett lineáris kombinációját. Ha erre
$$\sum_{k=0}^{n-1} c_k x^k = 0,$$
de a c_k együtthatók nem lennének mind nullák, akkor a baloldalon egy nem azonosan nulla polinom lenne, a jobboldalon viszont a 0 polinom; ez ellentmondást adna. Azt kaptuk, hogy $\{1, x, \ldots, x^{n-1}\}$-nek csak a csupa-nulla lineáris

kombinációja lehet 0 (= konstans-0 polinom), ezért $\{1, x, \ldots, x^{n-1}\}$ lineárisan független.

□

Newton-interpoláció (rekurzív megközelítés)

9.16. Tétel (Newton-interpoláció)**.** *Legyenek $a_1, \ldots, a_n \in K$ páronként különböző alappontok, $b_1, \ldots, b_n \in K$ tetszőleges értékek. Ekkor van olyan $f \in K[x]$, $\deg(f) < n$, amelyre $f(a_i) = b_i$ minden $i = 1, \ldots, n$ esetén.*

Bizonyítás. Teljes indukciót alkalmazunk n-re.
Alaplépés ($n = 1$): Az $f(x) = b_1$ konstans polinom nyilván megfelel (és nyilván ez az egyetlen jó nulladfokú polinom).
Indukciós lépés: Tegyük fel, hogy $n-1$ ponthoz már tudunk konstruálni interpolációs polinomot. Legyen $g \in K[x]$ olyan polinom, hogy $\deg(g) < n-1$ és $g(a_i) = b_i$ minden $i = 1, \ldots, n-1$ esetén (ez az indukciós feltevés szerint létezik). Legyen $f(x) = g(x) + (b_n - g(a_n))\ell_n$, ahol ℓ_n a 9.13 Definíció szerinti n-edik alappolinom. Ekkor $\deg(\ell_n) = n-1$, és $\ell_n(a_i) = 0$ minden $i = 1, \ldots, n-1$-re. Ezért, ha $i < n$, akkor

$$f(a_i) = g(a_i) + (b_n - g(a_n))\ell_n(a_i) = g_(a_i) = b_i.$$

Végül

$$f(a_n) = g(a_n) + (b_n - g(a_n))\ell_n(a_n) = g(a_n) + (b_n - g(a_n)) = b_n.$$

Így $f(x)$ kielégíti az összes feltételünket (beleértve a fokszámra vonatkozót is). □

9.17. Megjegyzés (Lagrange vs. Newton interpoláció)**.**

- *A **Lagrange-forma** akkor előnyös, ha sok különböző $(b_1, \ldots, b_n)$ értékkészlethez szeretnénk interpolálni ugyanazon $(a_1, \ldots, a_n)$ alappontok mellett, mert az $\ell_i(x)$ alappolinomokat (illetve az alappontok által generált Vandermonde-mátrix inverzét) csak egyszer kell kiszámítani.*
- *A **Newton-forma** (és a fenti rekurzív konstrukció) akkor hatékony, ha az alappontok száma (n) folyamatosan növekszik (pl. sorban kapjuk az (a_i, b_i) párokat), mivel az új pont hozzáadása csak egy új tag hozzáadását igényli a már meglévő polinomhoz.*

Taylor-polinom

Sok esetben szembesülünk a következő feladattal: adott egy bonyolult f függvény, szeretnénk egy egyszerűbb (ezért könnyebben kezelhető) másik, g függvényre, mondjuk egy g polinomfüggvényre áttérni úgy, hogy g azért „hasonlítson” az eredeti f-re.

Azt, hogy mit éretünk függvények hasonlóságán, további céljaink szerint sok módon precízzé lehet tenni. Egy lehetőség az, hogy f és g helyettesítési értékei egyezzenek meg előre adott véges sok pontban. Ezt a változatot polinom-interpolációval könnyen meg tudjuk már oldani; és valóban, sok esetben hasznos, ha egy bonyolult (fizikai, közgazdaságtani, stb. jelenséghez társítható) függvényt néhány helyen megmérünk, a mérési eredményeken polinom-interpolációt végzünk, majd az eredményül kapott polinommal közelítjük („jósoljuk meg") a bonyolult függvény helyettesítési értékeit olyan pontokban, ahol nem mértük még meg (pl. fizikai, vagy gazdasági kérdések esetén akár a jövőben).

Azt, hogy az f függvény „hasonlít" g-re, más módon is értelmezhetjük: megkövetelhetjük pl. azt, hogy egyetlen előre rögzített a pontban f és g első néhány deriváltjának egyezzen meg a helyettesítési értéke (a nulladik dferivált maga a függvény). A feladat, amit meg akarunk oldani, tehát az, hogy adott f-hez, a-hoz és $n \in \mathbb{N}$-hez találjunk g polinomot, melyre

$$f(a) = g(a),\ f'(a) = g'(a), \ldots, f^{(n)}(a) = g^{(n)}(a).$$

A Taylor-polinomok erre a feladattípusra adják meg a választ.

9.18. Definíció (Taylor-polinom)**.** *Legyen $I \subseteq \mathbb{R}$ nyílt intervallum, $a \in I$, és $f : I \to \mathbb{R}$ n-szer differenciálható az a pontban. Az f függvény a pont körüli n-**edfokú Taylor-polinomja** az a legfeljebb n-edfokú $T_{n,a,f} \in \mathbb{R}[x]$ polinom, amelyre:*

$$T^{(k)}_{n,a,f}(a) = f^{(k)}(a) \quad \textit{minden } k = 0, 1, \ldots, n \textit{ esetén},$$

ahol $f^{(0)} \equiv f$, és $T^{(k)}_{n,a,f}$ a polinom k-adik deriváltját jelöli. Az f-re vonatkozó deriválhatósági feltételek csak arra szolgálnak, hogy értelmes legyen a követelményrendszerünk (tudjunk a nem feltétlenül polinomfüggvény f sokadik deriváltjairól beszélni).

9.19. Tétel (Taylor-polinom együtthatói)**.** *Minden $n \in \mathbb{N}$-re: az n-edfokú Taylor-polinom létezik, egyértelmű, és a következő alakban írható fel:*

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

Bizonyítás. Rögzítsük n-t és keressük $T_{n,a,f}$-et az $\{1, (x-a), (x-a)^2, \ldots, (x-a)^n\}$ bázisban határozatlan (egyelőre ismeretlen) $\mu_0, ..., \mu_n$ együtthatókkal:

$$T_{n,a,f}(x) = \mu_0 + \mu_1(x-a) + \mu_2(x-a)^2 + \cdots + \mu_n(x-a)^n.$$

Ugyanúgy, mint az interpolációs feladat megoldása során, a $T_{n,a,f}$-re előírt tulajdonságokból elő fogunk állítani egy lineáris egyenletrendszert az ismeretlen $\mu_0, ..., \mu_n$

együtthatókra. Ehhez először számítsuk ki $T_{n,a,f}$ (formális) deriváltjait:

$$\begin{aligned}
T_{n,a,f}(x) &= 0! \cdot \mu_0 + \mu_1(x-a) + \mu_2(x-a)^2 + \cdots + \mu_n(x-a)^n,\\
T'_{n,a,f}(x) &= 1! \cdot \mu_1 + 2\mu_2(x-a)^1 + \cdots + n\mu_n(x-a)^{n-1},\\
T''_{n,a,f}(x) &= 2! \cdot \mu_2 + \cdots + n(n-1)\mu_n(x-a)^{n-2},\\
T'''_{n,a,f}(x) &= 3! \cdot \mu_3 + \cdots + n(n-1)(n-2)\mu_n(x-a)^{n-3}\\
&\vdots\\
T^{(k)}_{n,a,f}(x) &= k! \cdot \mu_k + \cdots + n(n-1)(n-2)\cdots(n-k+1)\mu_n(x-a)^{n-k}\\
&\vdots\\
T^{(n)}_{n,a,f}(a) &= n! \cdot \mu_n.
\end{aligned}$$

Ezek segítségével számítsuk ki $T_{n,a,f}$ deriváltjait az $x = a$ helyen:

$$\begin{aligned}
T_{n,a,f}(a) &= 0! \cdot \mu_0,\\
T'_{n,a,f}(a) &= 1! \cdot \mu_1,\\
T''_{n,a,f}(a) &= 2! \cdot \mu_2,\\
T'''_{n,a,f}(a) &= 3! \cdot \mu_3,\\
&\vdots\\
T^{(k)}_{n,a,f}(a) &= k! \cdot \mu_k,\\
&\vdots\\
T^{(n)}_{n,a,f}(a) &= n! \cdot \mu_n.
\end{aligned}$$

A definíció szerint $T^{(k)}_{n,a,f}(a) = f^{(k)}(a)$, tehát:

$$\mu_k = \frac{f^{(k)}(a)}{k!} \quad \text{minden } k = 0, 1, \ldots, n \text{ esetén.}$$

Ezzel az együtthatók egyértelműen meghatározottak, és a polinom a fenti összeg alakot ölti. □

Felmerül a kérdés, hogy egy (elegendően sokszor deriválható) f és a hozzá tartozó n-edik $T_{n,a,f}$ Taylor-polinom helyettesítési értékei mennyire maradnak közel egymáshoz. Az a pontban persze megegyeznek a helyettesítési értékeik, de mi van, ha a-ból arrébb megyünk? Vajon közel marad-e, milyen h-ra marad közel $T_{n,a,f}(a+h)$ és $f(a+h)$? Bevezetve az $R_n = f(a+h) - T_{n,a,f}(a+h)$ jelölést, a kérdés az, hogy mit mondhatunk az R_n eltérésről, „maradéktagról". Ezt a kérdést tisztázzák a **Taylor-formulák**, melyeknek sok változata van.

9.20. Tétel (Taylor-formula Lagrange-féle maradéktaggal)**.** *Legyen* $f : [a, a+h] \to \mathbb{R}$ *(*$h > 0$*)* $(n+1)$*-szer folytonosan differenciálható. Ekkor létezik olyan* $\xi \in (a, a+h)$*, hogy*

$$f(a+h) = T_{n,a,f}(a+h) + R_n,$$

ahol a ***Lagrange-féle maradéktag***

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}.$$

Azaz pontosan:

$$f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}.$$

Bizonyítás. Kalkulusból volt (pl. a Cauchy-féle középérték Tétel ismételt alkalmazásával). □

Ez a tétel tehát azt mondja ki, hogy ha a-tól h távolságra elmozdulunk, akkor van olyan $\xi \in (a, a+h)$, hogy az $f(a+h) - T_{n,a,f}(a+h)$ eltérés éppen

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}.$$

Ez utóbbi kifejezés hasonlít a Taylor-polinomok együtthatóira: olyan, mint az $(n+1)$-edik Taylor-polinom főtagja, de az együtthatóban az f függvény $(n+1)$-edik deriváltját nem az a helyen, hanem a ξ helyen vesszük (amiről csak annyit tudunk, hogy az $(a, a+h)$ intervallumban van), és nem $(x-a)^{n+1}$-el szorzunk, hanem h^{n+1}-el. Összefoglalva:

9.21. Megjegyzés.

1. Az interpolációs feladatban sok alappont van, és csak a helyettesítési értékekre (nulladik deriváltakra) van előírás. Ezzel szemben a Taylor-polinomok esetében egyetlen alappont van, viszont nemcsak a helyettesítési értékre, hanem az első néhány deriváltra van előírás. Későbbi félévekben megvizsgáljuk e két feladat közös általánosítását (több alappont van, mindegyik alappontban a helyettesítési értékek mellett elő van még írva az első néhány derivált értéke is).
2. Tudatosítsuk, hogy a Taylor-polinom, Taylor-formula és Taylor-sor három különböző dolog!
   (a) A Taylor-polinom a 9.18 Definíció szerint olyan polinom, melynek első néhány deriváltja egyetlen előre adott a pontban megegyezik egy adott f függvény első néhány deriváltjával az a pontban;
   (b) A Taylor-formula azt tisztázza, hogy f értékei és az a pontban vett Taylor-polinomjai értékei mennyire maradnak egymáshoz közel, ha a-ból arrébb megyünk;

(c) A Taylor-sor pedig azt a kérdést vizsgálja, hogy ha adott f-re adott a-ban minden n-re elkészítjük a $T_{n,a,f}$ Taylor-polinomokat, akkor e végtelen sok polinom együtt mutat-e valamiféle rendezettséget, pl. ha n végtelenbe tart, akkor ez a polinomsorozat (valamilyen értelemben) visszakonvergál-e az eredeti f-hez. Ezt a kérdést más kurzusokon még sokat fogjátok vizsgálni; a válasz az lesz, hogy általában persze nem konvergálnak vissza a Taylor-polinomok az eredeti f-hez, de ha f „szép" függvény, akkor mégis visszakonvergálnak.

3. Tehát a 9.20 Tétel azt mondja ki, hogy a függvény értéke mindegyik $(a + h)$ pontban felírható a Taylor-polinom (közelítés) és egy maradéktag (hiba) összegeként.
4. A maradéktag pontos formája (itt Lagrange-féle) lehetővé teszi a közelítés hibájának becslését, ha tudjuk korlátozni az $(n+1)$-edik deriváltat az intervallumon. Ennek egy alkalmazása az, ahogy a zsebszámológépek kiszámítják pl. a sin függvényt: valójában nem a sin függvényt számolják ki, hanem valamilyen előre rögzített jó nagy n-re a $T_{n,0,\sin}$ Taylor-polinomot (ez az alapműveletek ismételgetésével elvi nehézségek nélkül bármelyik pontban kiszámolható). A zsebszámológépek tervezésekor n-et előre rögzítik úgy, hogy a kapott közelítés előre adott (mondjuk 8 tizedesjegy) pontossággal megegyezzen a sin függvény értékeivel. Ez megtehető, mert a sin függvény összes deriváltja az összes pontban közös korlát alatt marad, ezért a maradéktag abszolútértéke előre felülbecsülhető.
5. Ha f maga is polinom, és $n \geq \deg(f)$, akkor $R_n = 0$, hiszen $f^{(n+1)} \equiv 0$. Ez persze abból is látszik, hogy ha $n \geq \deg(f)$, akkor a 9.18 Definíció értelmében f sajátmagának n-edik Taylor-polinomja (minden a-ra), ezért a 9.19 Tétel egyértelműségre vonatkozó része miatt $f = T_{n,a,f}$.

9.22. Példa (Az e^x függvény Taylor-polinomja $a = 0$ körül)**.** *Legyen* $f(x) = e^x$. *Ekkor* $f^{(k)}(x) = e^x$, *így* $f^{(k)}(0) = 1$ *minden* $k \geq 0$ *esetén. Az* n*-edfokú Taylor-polinom* $a = 0$ *körül:*

$$T_{n,0,e^x}(x) = \sum_{k=0}^{n} \frac{1}{k!}x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}.$$

A Taylor-formula szerint tetszőleges x*-re:*

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!}x^{n+1}, \quad \textit{valamely } \xi \in (0, x)\textit{-re.}$$

9.4. LU-felbontás

Ebben az alfejezetben elégséges feltételt adunk arra, hogy egy (négyzetes) mátrix előálljon egy alsó háromszögmátrix és egy felső háromszögmátrix szorzataként. Ehhez megvizsgáljuk, hogy az elemi sorátalakítások hogyan végezhetők el mátrix-szorzásokkal. Emlékeztetőül, az elemi sorátalakítások a következők voltak.

(α) Az i-edik sort szorozzuk egy nem nulla $c \in K$ skalárral.
(β) Az i-edik és j-edik sort felcseréljük.
(γ) Az i-edik sor c-szeresét hozzáadjuk a j-edik sorhoz ($i \neq j$).

9.23. Tétel. *Minden elemi sorátalakítás előáll úgy, hogy az $A \in \mathbb{K}^{n\times n}$ mátrixot balról megszorozzuk azzal a mátrixszal, amelyet az I_n egységmátrixból az adott sorátalakítással kapunk. Az α típusú és $i < j$ esetén a γ típusú sorátalakításokhoz tartozó mátrixok alsó háromszögmátrixok.*

Bizonyítás. Szisztematikusan megvizsgáljuk az elemi sorátalakítások típusait.

- (α) **típus:** Az i-edik sor megszorzása nemnulla skalárral. Ez megfelel az alábbi diagonális, így alsó háromszögmátrixszal való szorzásnak.

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

- (β) **típus:** Az i-edik és j-edik sor felcserélése. Könnyű ellenőrizni, hogy ez is előáll egy invertálható elemi mátrixszal való balról szorzással:

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 1 & \cdots & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots 1 \end{bmatrix}.$$

- (γ_{ij}) **típus:** Az i-edik sor c-szeresét hozzáadjuk a j-edik sorhoz. Könnyen ellenőrizhető, hogy ez az átalakítás végrehajtható az alábbi mátrixal való balról szorzással (az i-edik és a j-edik sort írjuk ki, és az $i < j$ esetet szemléltettük):

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 1 & \cdots & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & c & \cdots & 1 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots 1 \end{bmatrix}.$$

  Ez szintén alsó háromszögmátrix, ha $i < j$.

□

9.24. Tétel (LU-felbontás létezése)**.** *Ha $B \in \mathbb{K}^{n\times n}$ olyan, hogy Gauss-eliminációval sorcsere nélkül lépcsős alakra hozható, akkor léteznek*

$$L \text{ alsó háromszögmátrix}, \qquad U \text{ felső háromszögmátrix},$$

melyekre

$$B = LU.$$

9.25. Megjegyzés. *A 9.24 Tétel elégséges feltételt ad arra, hogy egy $B \in K^{n\times n}$ mátrix előállítható legyen $B = LU$ alakban, ahol:*

- *L* ***alsó háromszögmátrix*** *(lower triangular)*
- *U* ***felső háromszögmátrix*** *(upper triangular).*

Bizonyítás. A tétel feltételei miatt B lépcsős alakra hozásánál a Gauss-elimináció során elég α típusú és olyan γ típusú sorátalakításokat végezni, melyekre $i < j$. A 9.23 Tétel miatt ezek az átalakítások alsó háromszögmátrixokkal való balról szorzásnak felelnek meg. Ezért léteznek $L_1, \dots, L_k$ alsó háromszögmátrixok úgy, hogy

$$L_k \cdots L_1 B = U,$$

ahol U felső háromszögmátrix (hiszen lépcsős alakú).

Mivel minden elemi sorátalakítás visszacsinálható elemi sorátalakításokkal, ezért mindegyik L_i invertálható és L_i^{-1} is alsó háromszögmátrix (mert L_i^{-1} is α, vagy megfelelő alakú γ típusú sorátalakításnak felel meg). Ezért

$$B = (L_1 \cdots L_k)^{-1} U = (L_k^{-1} \cdots L_1^{-1})U.$$

De alsó háromszögmátrixok szorzata ismét alsó háromszögmátrix ⇒ készen vagyunk.
□

9.5. Ferde kifejtés, determináns és inverz, Cramer-szabály

Emlékeztetünk rá, hogy a 8.14 Definíció szerint ha $A \in K^{n\times n}$ és $1 \le i, j \le n$, akkor $A_{ij} \in K^{(n-1)\times(n-1)}$ az a mátrix, melyet A-ból úgy kapunk, hogy A-ból elhagyjuk az i-edik sort és j-edik oszlopot. Az A_{ij} jelölés összekeverhető azzal, ahogy az A mátrix i-edik sorának j-edik elemét is jelöltük, ezért ebben az alfejezetben $A(i, j)$-vel fogjuk jelölni az A-ból i-edik sora és j-edik oszlopa kitörlésével előálló $(n-1)\times(n-1)$-es mátrixot.

9.26. Tétel (Ferde kifejtés és Cramer-szabály)**.**

1. *Ferde kifejtés*[5]*: ha $j \neq i$, akkor*

$$\sum_{k=1}^{n} (-1)^{j+k} a_{ik} \det(A(j,k)) = 0.$$

2. *Legyen A^* az a mátrix, melyet A-ból úgy kapunk, hogy A-ban minden elemet kicserélünk a hozzá tartozó előjelezett aldeterminánsra, majd az így kapott mátrixot transzponáljuk:*

$$(A^*)_{ij} = (-1)^{i+j} \det(A(j,i)).$$

Ekkor

$$AA^* = \begin{pmatrix} \det(A) & & 0 \\ & \ddots & \\ 0 & & \det(A) \end{pmatrix} = \det(A)\, I_n.$$

3. *Cramer-szabály: Ha A invertálható, akkor az*

$$Ax = b$$

egyenlet egyetlen megoldása

$$x = A^{-1}b.$$

Sőt, minden i-re

$$x_i = \frac{\det(A_{*1}, \ldots, A_{*i-1}, b, A_{*i+1}, \ldots, A_{*n})}{\det(A)},$$

ahol a számlálóban A i-edik oszlopát b-re cseréljük.

Bizonyítás. (1) igazolásával kezdünk. Legyen A' az a mátrix, amelyet A-ból úgy kapunk, hogy a j-edik sort az i-edik sorra cseréljük (de minden másik sort, az i-ediket is, változatlan formában megtartunk; ezáltal A'-ben az i-edik és j-edik sor egyenlő lesz). Egyrészt,

$$\det(A') = 0,$$

mivel van neki két egyenlő sora. Másrészt fejtsük ki $\det(A')$-t a j-edik sor szerint:

$$0 = \det(A') = \sum_{k=1}^{n} (-1)^{j+k} a'_{jk} \det(A'(j,k)) = \sum_{k=1}^{n} (-1)^{j+k} a_{ik} \det(A(j,k)),$$

és itt a sor elején 0, a sor végén pedig a ferde kifejtés szerepel.

[5] Az alábbi összeg hasonlít arra, mintha A-t az i-edik sora szerint kifejtenénk, azonban az i-edik sor a_{ik} elemeit nem a saját előjelezett aldeterminánsával szorozzuk, hanem a j-edik sor megfelelő a_{jk} eleméhez tartozó előjelezett aldeterminánssal. Erre utal a „ferde" kifejezés.

(2) igazolása a következő. Tetszőleges i, j-re:

$$(AA^*)_{ij} = \sum_{k=1}^{n} a_{ik}(A^*)_{kj} = \sum_{k=1}^{n} a_{ik}(-1)^{k+j} \det(A(j,k)).$$

Ha $i \neq j$, akkor ez 0 a ferde kifejtés (előbb igazolt (1)) miatt. Ha $i = j$, akkor

$$(AA^*)_{ii} = \sum_{k=1}^{n}(-1)^{i+k} a_{ik} \det(A(i,k)) = \det(A),$$

az i-edik sor szerinti kifejtés miatt.

Tehát

$$AA^* = \det(A) I_n,$$

vagyis

$$(*) \qquad A^{-1} = \frac{1}{\det(A)} A^*.$$

Végül (3) bizonyítása a következő. Az nyilvánvaló, hogy ha A invertálható, akkor $Ax = b$ egyetlen megoldása

$$x = A^{-1}b,$$

ezt kombinálva a (2)-ben (néhány sorral feljebb) igazolt $(*)$-al, tetszőleges i-re:

$$x_i = (A^{-1}b)_i = \sum_{k=1}^{n}(A^{-1})_{ik} b_k = \frac{1}{\det(A)} \sum_{k=1}^{n}(-1)^{i+k} \det(A(k,i)) b_k,$$

és a sor végén álló összeg $\det(A_{*1}, \ldots, A_{*i-1}, b, A_{*i+1}, \ldots, A_{*n})$ kifejtése az az i-edik oszlopa szerint. □

A Cramer-szabály egy elméleti jelentőségű megoldóképlet. Megoldóképlet abban az értelemben, hogy (egy meglehetősen speciális esetben) a lineáris egyenletrendszer megoldását egy áttekinthető képlet formájában adja meg; ez hasznos lesz későbbi félévekben, amikor erre alapozva fogunk tételeket igazolni. Ugyanakkor a jelentősége inkább elméleti, semmint gyakorlati a következők miatt.

1. Ha az egyenletrendszer mátrixa nem négyzetes, akkor fel sem merül, hogy a Cramer-szabályt alkalmazzuk, mert csak négyzetes mátrixnak van determinánsa;
2. Ha az egyenletrendszer mátrixa négyzetes ugyan, de a determináns nulla (azaz nem pontosan 1 megoldás van), a Cramer-szabály akkor sem alkalmazható (mert 0-val kéne osztani);
3. Ha az egyenletrendszer mátrixa négyzetes, és determinánsa nem 0, akkor a Cramer-szabály alkalmazható ugyan, de nagyon művelet-igényes: n ismeretlen esetében $n + 1$ darab $n \times n$-es determinánst kellene kiszámolni, míg

a Gauss-elimináció műveletigénye nagyjából 1 darab $n \times n$-es determináns kiszámításával arányos (hiszen a gyakorlatban a determinánsokat sem kifejtéssel, hanem pl. elemi sorátalakításokat használva háromszög-alakra hozással érdemes kiszámolni).

9.6. Skaláris szorzás K^n-ben

A 7.55 Tételben láttuk, hogy a K test feletti minden véges dimenziós vektortérhez van $n \in \mathbb{N}$, hogy V izomorf K^n-el. Továbbá, K^n szerkezete elég konkrét ahhoz, hogy egyszerűen be lehessen vezetni rajta a **skaláris szorzás** műveletét, melyet az $\mathbb{R}^2$-ben és $\mathbb{R}^3$-ban középiskolában megismert skaláris szorzás motivál. Nem véges dimenziós terekben a skaláris szorzás bevezetése további előkészületeket igényelne, ezt későbbi félévekre halasztjuk. Ebben a kurzusban a véges dimenziós esettel ismerkedünk meg, és egy alkalmazást is be fogunk mutatni: meg fogunk oldani megoldhatatlan egyenletrendszereket (pontos megoldás helyett nyilván csak – valamilyen egészen világosan definiált értelemben – optimális közelítő megoldásokat fogunk tudni találni).

9.27. Definíció. *Ha $a, b \in K^n$, akkor*

$$\langle a, b \rangle = \sum_{k=1}^{n} a_k b_k$$

az a és b skaláris szorzata.

Hangsúlyozzuk, hogy K^n-ben az előbbi módon **definiáltuk** a skaláris szorzást. $\mathbb{R}^2$-ben és $\mathbb{R}^3$-ban másképp, geometriai megfontolásokkal szokás definiálni a skaláris szorzást, abban a felépítésben az **tétel**, hogy két vektor skaláris szorzata a koordinátáikból úgy számolható ki, ahogy azt az előző definíció körülírja.

9.28. Definíció. *Az $a, b \in K^n$ vektorok* **merőlegesek** *(=***ortogonálisak***) egymásra (jelölés $a \perp b$), ha skaláris szorzatuk 0:*

$$a \perp b \iff \langle a, b \rangle = 0.$$

9.29. Tétel (A skaláris szorzás tulajdonságai)**.** *Minden $a, b, c \in K^n$ és $\lambda \in K$ esetén:*

1. $\langle a, b \rangle = \langle b, a \rangle$,
2. $\langle a, \lambda b \rangle = \lambda \langle a, b \rangle$,
3. $\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$.

Bizonyítás.

1.
$$\langle a, b \rangle = \sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} b_k a_k = \langle b, a \rangle.$$

2.
$$\langle a, \lambda b\rangle = \sum_{k=1}^{n} a_k(\lambda b_k) = \lambda \sum_{k=1}^{n} a_k b_k = \lambda\langle a, b\rangle.$$

3.
$$\langle a, b + c\rangle = \sum_{k=1}^{n} a_k(b_k + c_k) = \sum_{k=1}^{n} a_k b_k + \sum_{k=1}^{n} a_k c_k.$$

□

Direkt és ortogonális kiegészítés

9.30. Definíció. *Legyen V vektortér, $W_1, W_2 \subseteq V$ alterek. Azt mondjuk, hogy W_1 és W_2 egymás direkt kiegészítői, ha teljesülnek az alábbi 1. és 2. pontok:*

1. $W_1 \cap W_2 = \{0\}$,
2. $\forall x \in V \ \exists\, w_1 \in W_1,\ w_2 \in W_2 :\ x = w_1 + w_2$.

3. *Azt mondjuk, hogy W_1 és W_2* **merőlegesek** *(=***ortogonálisak***) egymásra (jelölés: $W_1 \perp W_2$), ha*
$$\forall w_1 \in W_1,\ \forall w_2 \in W_2 :\quad w_1 \perp w_2.$$
4. *Azt mondjuk, hogy W_1 és W_2 egymás ortogonális kiegészítői V-ben, ha az előző 1., 2. és 3. pontok mindegyike teljesül rájuk (tehát 1. és 2. miatt direkt kiegészítők is).*

9.31. Megjegyzés. $\mathbb{R}$-ben (1) következik (3)-ból, mert ha $x \in W_1 \cap W_2$, akkor (3) miatt $x \perp x$, vagyis
$$\langle x, x\rangle = \sum_{k=1}^{n} x_k^2 = 0,$$
amiből $x = 0$ következik. A komplex test (és még sok más test) felett ez a következtetés nem marad érvényben: tekintsük pl. az $[1, i] \in \mathbb{C}^2$ vektort. Erre
$$\langle [1, i], [1, i]\rangle = 1 + i^2 = 0,$$
tehát $[1, i] \perp [1, i]$, noha nyilvánvalóan $0 \neq [1, i]$.

9.32. Példa. *$\mathbb{R}^2$-ben:*
$$W_1 = \{ke_1 \mid k \in \mathbb{R}\}, \qquad W_2 = \{k(1, 1) \mid k \in \mathbb{R}\}.$$
Ezek nem ortogonális kiegészítők, viszont direkt kiegészítők.

W_1-nek végtelen sok direkt kiegészítője van (W_2-ben az $(1, 1)$ vektor használata esetleges, bármilyen nemnulla α-val az origón átmenő, $(1, \alpha)$ irányvektorú egyenes direkt kiegészítője lenne W_1-nek). De W_1-nek csak egy ortogonális kiegészítője van: az e_2 által kifeszített egyenes.

9.7. Lineáris leképezések magtere és képtere

9.33. Definíció (Magtér és képter)**.** *Legyen V és W vektorterek, $\varphi : V \to W$ lineáris leképezés.*

- $\ker(\varphi) = \{v \in V \mid \varphi(v) = 0\} \subseteq V$ *a* φ ***magtere****.*
- $\mathrm{im}(\varphi) = \{\varphi(v) \mid v \in V\} \subseteq W$ *a* φ ***képtere****.*

Ha $\varphi : V \to V$ lineáris leképezés, akkor $[\varphi]$ jelöli φ mátrixát a sztenderd bázisban.

9.34. Tétel. *Legyen V véges dimenziós (mondjuk n-dimenziós) vektortér a K test felett, $\varphi, \psi : V \to V$ lineáris leképezések. Ekkor*

1. *$[\varphi \circ \psi] = [\varphi][\psi]$ (itt a jobboldalon mátrixszorzás van);*
2. *ha φ bijekció, akkor $[\varphi^{-1}] = [\varphi]^{-1}$;*
3. *$ker(\varphi)$ és $im(\varphi)$ alterei V-nek;*
4. *Dimenziótétel: $dim(ker(\varphi)) + dim(im(\varphi)) = n$;*
5. *Tetszőleges $\mathbb{R} \in K^{n\times n}$-re $\mathcal{S}(A)$ és $\mathcal{N}(A)$ ortogonális kiegészítők.*

Ebből vizsgára elég tudni az első 4 pont állítását (bizonyítás nélkül), és az 5. pontot bizonyítással.

Bizonyítás. (1): Legyen $b \in V$ tetszőleges vektor, $[b]$ a koordináta-vektora (a sztenderd bázisban). Ekkor

$$[\varphi \circ \psi][b] = [\varphi(\psi(b))] = [\varphi([\psi][b])] = [\varphi]([\psi][b]) = ([\varphi][\psi])[b].$$

Ez minden b vektorra igaz, speciálisan igaz a $b = e_1, \ldots, e_n$ sztenderd bázis vektoraira is. A $b = e_j$ vektorra alkalmazva azt kapjuk, hogy a baloldali $[\varphi \circ \psi]$ mátrix j-edik oszlopa egyenlő a jobboldali $[\varphi][\psi]$ szorzatmátrix j-edik oszlopával, és ez minden j-re igaz. Ezért valóban, $[\varphi \circ \psi] = [\varphi][\psi]$.

(2):

$$[\varphi^{-1}][\varphi] \overset{(1)}{=} [\varphi^{-1} \circ \varphi] = I_n,$$

ezért $[\varphi^{-1}] = [\varphi]^{-1}$.

(3): Ha $v_1, v_2 \in \ker(\varphi)$ és $\lambda \in K$, akkor

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = 0$$

és

$$\varphi(\lambda v_1) = \lambda\varphi(v_1) = 0,$$

tehát $v_1 + v_2 \in \ker(\varphi)$ és $\lambda v_1 \in \ker(\varphi)$. Ez mutatja, hogy $ker(\varphi)$ zárt a vektortérműveletekre, ezért altér.

Hasonlóan, ha $v_1, v_2 \in im(\varphi)$ és $\lambda \in K$, akkor van $w_1, w_2 \in V :\ \varphi(w_1) = v_1$ és $\varphi(w_2) = v_2$. De ekkor

$$v_1 + v_2 = \varphi(w_1) + \varphi(w_2) = \varphi(w_1 + w_2)$$

és

$$\lambda v_1 = \lambda\varphi(w_1) = \varphi(\lambda w_1),$$

ezért $v_1 + v_2 \in im(\varphi)$ és $\lambda v_1 \in im(\varphi)$, azaz $im(\varphi)$ is zárt a vektortér műveletekre, tehát altér.

(4): Emlékeztetünk rá, hogy tételünk kimondásakor rögzítettük vektorterünk dimenzióját: $\dim(V) = n$. Válasszunk egy $\{b_1, \ldots, b_k\} \subseteq \ker(\varphi)$ bázist a magtérben. Egészítsük ki ezt V egy B bázisává: $B = \{b_1, \ldots, b_k, b_{k+1}, \ldots, b_n\}$. Megmutatjuk, hogy $\{\varphi(b_{k+1}), \ldots, \varphi(b_n)\}$ bázisa $im(\varphi)$-nek.

- **Generálás:** Legyen $w \in im(\varphi)$. Ekkor van olyan $v \in V$, hogy $\varphi(v) = w$. Állítsuk elő v-t a B bázisban: $v = \sum_{i=1}^{n} \lambda_i b_i$. Ekkor

$$w = \varphi(v) = \varphi\Big(\sum_{i=1}^{n} \lambda_i b_i\Big) = \sum_{i=1}^{n} \lambda_i \varphi(b_i) = \sum_{i=k+1}^{n} \lambda_i \varphi(b_i),$$

  mert $\varphi(b_1) = \cdots = \varphi(b_k) = 0$, hiszen $b_1, \ldots b_k$ mindegyike φ magterében van. Tehát $\{\varphi(b_{k+1}), \ldots, \varphi(b_n)\}$ generátorrendszer $im(\varphi)$-ben.

- **Lineáris függetlenség:** Tegyük fel, hogy $\sum_{i=k+1}^{n} \mu_i \varphi(b_i) = 0$. Ekkor

$$0 = \sum_{i=k+1}^{n} \mu_i \varphi(b_i) = \varphi\left(\sum_{i=k+1}^{n} \mu_i b_i\right) \text{ azaz } \sum_{i=k+1}^{n} \mu_i b_i \in \ker(\varphi).$$

  Mivel $\{b_1, \ldots, b_k\}$ bázis a magtérben, léteznek $\alpha_1, \ldots, \alpha_k$ skalárok, hogy

$$\sum_{i=k+1}^{n} \mu_i b_i = \sum_{j=1}^{k} \alpha_j b_j.$$

  Átrendezve:

$$\sum_{j=1}^{k} (-\alpha_j) b_j + \sum_{i=k+1}^{n} \mu_i b_i = 0.$$

  Mivel $B = \{b_1, \ldots, b_n\}$ lineárisan független, ezért $\alpha_1 = \cdots = \alpha_k = 0$ és $\mu_{k+1} = \cdots = \mu_n = 0$. Tehát $\{\varphi(b_{k+1}), \ldots, \varphi(b_n)\}$ lineárisan független.

Ezek szerint $\dim(im(\varphi)) = n - k = n - \dim(ker(\varphi))$, ami épp a dimenziók (4) szerinti (bizonyítandó) összefüggését adja.

(5) A 9.30 Definíció pontjait kell ellenőriznünk. Először azt mutatjuk meg, hogy $\mathcal{S}(A) \perp \mathcal{N}(A)$. Ehhez legyen $b \in \mathcal{N}(A)$ és $d \in \mathcal{S}(A)$ tetszőleges, azt kell belátnunk, hogy $b \perp d$. De

$$(*) \qquad b \in \mathcal{N}(A) \;\Leftrightarrow\; Ab = 0 \;\Leftrightarrow\; \forall i: \; (Ab)_i = 0 \;\Leftrightarrow\; \forall i: \; \langle A_{i*}, b\rangle = 0.$$

Mivel $d \in \mathcal{S}(A)$, ezért előáll A sorainak lineáris kombinációjaként: vannak olyan $\alpha_1, \dots \alpha_n \in \mathbb{R}$ számok, hogy $d = \sum_{i=1}^{n} \alpha_i A_{i*}$. Viszont ekkor

$$\langle d, b\rangle = \langle \sum_{i=1}^{n} \alpha_i A_{i*}, b\rangle = \sum_{i=1}^{n} \alpha_i \langle A_{i*}, b\rangle \overset{(*) \text{ jobboldala}}{=} 0,$$

vagyis valóban $b \perp d$.

Az világos, hogy $\mathcal{S}(A) \cap \mathcal{N}(A) = \{0\}$: tegyük fel, hogy a $b = [b_1, \dots, b_n]$ vektor $\mathcal{S}(A) \cap \mathcal{N}(A)$-ban van. Ekkor az előbb igazolt rész szerint $b \perp b$, azaz

$$0 = \langle b, b\rangle = \sum_{i=1}^{n} b_i^2,$$

amiből $b_1 = b_2 = \dots = b_n = 0$.

Legyenek $B = \{b_1, \dots, b_k\} \subseteq \mathcal{S}(A)$ és $D = \{d_1, \dots, d_m\} \subseteq \mathcal{N}(A)$ bázisok A sorterében és nullterében. Állítjuk, hogy $B \cup D$ lineárisan független. Legyenek ugyanis $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m \in \mathbb{R}$ tetszőlegesek. Ha

$$0 = \sum_{i=1}^{k} \lambda_i b_i + \sum_{j=1}^{m} \mu_j d_j,$$

akkor

$$-\sum_{i=1}^{k} \lambda_i b_i = \sum_{j=1}^{m} \mu_j d_j$$

adódik, de a baloldal $\mathcal{S}(A)$-ban, a jobboldal $\mathcal{N}(A)$-ban van, vagyis – egyenlőségük miatt – midkét oldal benne van $\mathcal{S}(A) \cap \mathcal{N}(A)$-ban. Ezért az előző bekezdés szerint mindkét oldal 0. Viszont B és D bázisok, emiatt

$$\lambda_1 = \dots = \lambda_k = \mu_1 = \dots = \mu_m = 0;$$

ez mutatja, hogy $B \cup D$ tényleg lineárisan független.

Az előző bekezdés miatt a bizonyítás befejezéséhez elég azt meggondolni, hogy

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{S}(A)) = n.$$

Ehhez legyen $\varphi : \mathbb{R}^n \to \mathbb{R}^n, \varphi(x) = Ax$. Ekkor, mivel $\mathcal{N}(A) = ker(\varphi)$ és $\mathcal{O}(A) = im(\varphi)$, ezért

$$\dim(\mathcal{N}(A)) + dim(\mathcal{S}(A)) \overset{7.46 \text{ Tétel}}{=} \dim(\mathcal{N}(A)) + dim(\mathcal{O}(A)) =$$

$$\dim(ker(\varphi)) + \dim(im(\varphi)) \overset{(4) \text{ (Dimenziótétel)}}{=} n,$$

és ezzel készen vagyunk. □

9.8. Ellentmondásos (túlhatározott) lineáris egyenletrendszerek optimális közelítő megoldása

Tegyük fel, hogy meg kell oldanunk egy olyan $Ax = b$ lineáris egyenletrendszert, melyek egyenletei ellentmondanak egymásnak (pl. az egyenletrendszer jobboldala mérési eredményekből származik, de a mérések pontatlanok). A lineáris egyenletrendszerek megoldhatóságának mátrixrangos jellemzéséből (7.50 Tételből) tudjuk, hogy az $Ax = b$ egyenletrendszer pontosan akkor ellentmondásos (azaz pontosan akkor nincs neki megoldása), ha b nincs benne A oszlopterében. Mit tehetünk ilyenkor? Megpróbálhatjuk megkeresni $\mathcal{O}(A)$-ban a b-hez legközelebbi b' vektort, és az eredeti egyenletrendszer helyett $Ax = b'$-t oldjuk meg. Mivel b'-t $\mathcal{O}(A)$-ból választottuk, ennek az egyenletrendszernek biztosan lesz megoldása. Továbbá b' választása miatt igaz lesz, hogy az Ay alakú vektorok közül (vagyis $\mathcal{O}(A)$ elemei közül) épp a mi x megoldásunk lesz az, melyre Ax (vagyis b') a lehető legközelebb lesz b-hez (speciálisan, ha b mégis benne van $\mathcal{O}(A)$-ban, akkor b és b' távolsága 0 lesz, ezért a fenti közelítő megoldás igazi, pontos megoldás lesz).

A fenti ötlet keresztülviteléhez már csak a megfelelő távolságfogalom bevezetése van hátra, illetve az, hogy adott b-hez megtaláljuk az $\mathcal{O}(A)$ altér legközelebbi elemét. Geometriai ismereteink alapján tudjuk, hogy $\mathbb{R}^2$ illetve $\mathbb{R}^3$ adott altereiben b-nek - a szóbanforgó altérre vett - merőleges vetülete van a legközelebb b-hez. Ez adja az ötletet, hogy sokdimenziós esetben is a merőleges vetülettel próbálkozzunk. Ebben az alfejezetben tehát

- bevezetjük sokdimenziós vektorok hosszát (és ezzel pontpárok távolságát),
- kidolgozzuk a merőleges vetület megfelelő általánosítását,
- ellenőrizzük, hogy intuíciónk helyes: adott altérben adott b-hez tényleg b merőleges vetülete van legközelebb,
- példát adunk egy ellentmondásos lineáris egyenletrendszer közelítő optimális megoldására.

A következő félévben folytatjuk majd ilyen irányú vizsgálatainkat, lesznek további (hatékonyabb, szebb módszerek).

9.35. Definíció. *Ha $b \in K^n$, akkor b* **hossza***: $|b| = \sqrt{\langle b, b\rangle}$.*

Az előző definíció jelöléseinek megtartásával vegyük fel b-t koordinátás alakban: $b = [b_1, \ldots, b_n]$. Ekkor

$$\langle b, b\rangle = \sum_{i=1}^{n} b_i^2,$$

ez – a Pitagorasz-tétel miatt – $\mathbb{R}^2$-ben illetve $\mathbb{R}^3$-ben a síkvektorok, illetve térvektorok szokásos hossz-négyzetét adja. Ennek a természetes általánosítása szerepel az előző definícióban.

9.36. Definíció. *Legyen W a K^n vektortér altere, és legyen $v \in K^n$. Azt mondjuk, hogy a $v' \in W$ vektor* **merőleges vetülete** *v-nek W-re, ha*

$$v - v' \perp W.$$

9.37. Tétel. *Legyen $n \in \mathbb{N}$ és legyen $V \subseteq \mathbb{R}^n$ egy altér. Ekkor*

1. *Pitagorasz-tétel: Ha $u, v \in \mathbb{R}^n$ és $u \perp v$, akkor $|u+v|^2 = |u|^2 + |v|^2$;*
2. *$\forall u \in \mathbb{R}^n \exists u' \in V$: u' merőleges vetülete u-nak V-re;*
3. *V-ben u merőleges vetülete van legközelebb u-hoz.*

Bizonyítás. (1) igazolása a következő:

$$|u+v|^2 = \langle u+v, u+v\rangle = \langle u, u+v\rangle + \langle v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = (*).$$

Mivel azonban $u \perp v$, ezért $\langle u, v\rangle = 0$ és $\langle v, u\rangle = 0$. Ezért

$$(*) = \langle u, u\rangle + \langle v, v\rangle = |u|^2 + |v|^2,$$

ahogy állítottuk.

(2) Igazolásához vegyünk fel egy $B = \{b_1, \ldots, b_k\} \subseteq V$ bázist V-ben. Legyen A az a mátrix, melynek a $\{b_1, \ldots, b_k\}$ vektorok a sorai. Vegyünk fel egy $D = \{b_{k+1}, \ldots, b_n\} \subseteq \mathcal{N}(A)$ bázist $\mathcal{N}(A)$-ban. A 9.34 Tétel 5. pontja miatt $B \cup D$ bázis $\mathbb{R}^n$-ben. Ezért vannak $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ számok, melyekkel

$$u = \sum_{i=1}^{n} \alpha_i b_i.$$

Legyen $u' = \sum_{i=1}^{k} \alpha_i b_i$ (tehát u' előállításához u koordinátáiból csak a V-be eső vektorok koordináit használjuk). Azt állítjuk, hogy ez az u' merőleges vetülete u-nak. Az világos, hogy $u' \in V$. Továbbá, mivel

$$u - u' = \sum_{i=k+1}^{n} \alpha_i b_i,$$

ezért $u - u' \in \mathcal{N}(A)$. Mivel $\mathcal{N}(A)$ merőleges $\mathcal{S}(A) = V$-re ezért $u - u' \perp V$ is teljesül, tehát u' a 9.36 Definíció szerint u' valóban merőleges vetülete u-nak.

A merőleges vetület egyértelműségét és 3.-at egyszerre látjuk be. Legyen $v \in V$ tetszőleges és legyen u' tetszőleges merőleges vetülete u-nak V-re. Ekkor $u-u' \perp V$ és nyilván $u'-v \in V$. Ezért $u-u' \perp u'-v$. Ezt az (1)-ben igazolt Pitagorasz-tétellel kombinálva azt kapjuk, hogy

$$(*) \qquad |u-v|^2 = |(u-u') + (u'-v)|^2 \overset{(1)\ \text{Pit.Tétel}}{=} |u-u'|^2 + |u'-v|^2.$$

Ezek szerint $|u-v|$ akkor és csak akkor minimális, ha $|u'-v|^2 = 0$, vagyis ha $v = u'$. Ez egyrészt mutatja, hogy V-ben u merőleges vetületei vannak legközelebb u-hoz. Másrészt, u'-n kívül, u-nak nem lehet másik u'' merőleges vetülete V-re, mert akkor ezt a másik u''-t adva értékül $(*)$-ban v-nek, nem a minimális távolság adódna. □

9.38. Példa. *Legyen W az x-tengely $\mathbb{R}^2$-ben: $W = \{(x, 0) \mid x \in \mathbb{R}\}$, és $\mathbf{b} = (3, 4)$. Ekkor $\mathbf{b}$ ortogonális vetülete W-re: $(3, 0)$. Távolság: $|\mathbf{b} - (3, 0)| = 4$. Valóban, bármely más $(x, 0)$ pontra:*

$$|(3, 4) - (x, 0)|^2 = (3 - x)^2 + 16 \geq 16,$$

és egyenlőség csak $x = 3$ esetén teljesül.

Példa: egy ellentmondásos egyenletrendszer megoldása

Ebben az alfejezetben a a következő egyenletrendszert vizsgáljuk.

$$(*)\begin{cases}\frac{1}{2}x-\frac{1}{2}y=5\\ y=7\\ x=-1\end{cases}$$

Ez nyilvánvalóan ellentmondásos: a két utolsó egyenletből $x-y=-8$ következik, ami ellentmond az első egyenletnek. Mielőtt megoldjuk, leírom, milyen geometriai megfontolások alapján találtam ki ezt az egyenletrendszert. Ez a megoldáshoz nem lenne szükséges, de talán segíthet megérteni a sok számolás mögött rejlő gondolatokat.

Találomra választottam egy síkot:

$$S: \qquad 2x+y-z=0.$$

Egyetlen szempontom az volt, hogy x, y és z együtthatói páronként különbözzenek, és kicsi egész számok legyenek. Adjuk meg ezt a síkot paraméteresen: S egyenletében y és z szabad változók, ezért S egy paraméteres előállítása:

$$\begin{bmatrix}-\frac{1}{2}s+\frac{1}{2}t\\ s\\ t\end{bmatrix}=s\cdot\begin{bmatrix}-\frac{1}{2}\\ 1\\ 0\end{bmatrix}+t\cdot\begin{bmatrix}\frac{1}{2}\\ 0\\ 1\end{bmatrix}.$$

Válasszunk egy b' pontot az S síkon, mondjuk az $s=4, t=2$ paraméter-értékekhez tartozó pont legyen b'. Itt megint csak arra figyeltem, hogy s, t értékei különbözzenek, párosak legyenek, hogy a menet közben adódó törtek egészekké váljanak, és ne legyenek túl nagyok. Tehát $b'=[-1,4,2]$.

Legyen e a b' ponton átmenő, S-re merőleges egyenes. Ennek egy irányvektora megegyezik S egy normálvektorával: $[2,1,-1]$. Ebből az e egyenes paraméteres egyenletrendszere:

$$\begin{bmatrix}-1\\ 4\\ 2\end{bmatrix}+u\cdot\begin{bmatrix}2\\ 1\\ -1\end{bmatrix}.$$

Legyen b az e egyenes (mondjuk) az $u=3$ paraméter-értékhez tartozó pontja: $b=[5,7,-1]$. Ez nincs benne S-ben, de S-re való merőleges vetülete b'. Az egyenletrendszerünk fejezze ki azt, hogy ezt a b-t akarjuk az S paraméterezésében szereplő két vektor lineáris kombinációjaként előállítani. Vagyis keressünk olyan x, y-nt, melyekre

$$y\cdot\begin{bmatrix}-\frac{1}{2}\\ 1\\ 0\end{bmatrix}+x\cdot\begin{bmatrix}\frac{1}{2}\\ 0\\ 1\end{bmatrix}=\begin{bmatrix}5\\ 7\\ -1\end{bmatrix}.$$

Ez nyilvánvalóan lehetetlen, mert b nincs az S síkban. Feltételeinket egyenletrendszerré alakítva kapjuk az alfejezet elején megadott $(*)$ egyenletrendszert.

Ezek után oldjuk meg $(*)$-ot a korábbi fejezetek ismeretei alapján (még egyszer hangsúlyozom, hogy későbbi félévekben lesz hatékonyabb módszer).

9.39. Példa (Megoldás vetítési módszerrel)**.** Az világos, hogy $(*)$ ellentmondásos.
1. lépés: Mátrix alakba írás

$$A = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 7 \\ -1 \end{bmatrix}, \quad A\mathbf{x} = \mathbf{b},\ \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Meg kéne keresnünk az egyenletrendszer jobboldalának (b-nek) merőleges vetületét A oszlopterére.
2. lépés: Oszloptér ($\mathcal{O}(A)$) bázisának megválasztása Az A mátrix oszlopvektorai:

$$\mathbf{a}_1 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} \frac{-1}{2} \\ 1 \\ 0 \end{bmatrix}.$$

Ellenőrizve: $\mathbf{a}_2$ nem skalárszorosa $\mathbf{a}_1$-nek, tehát lineárisan függetlenek. Így

$$\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2\} \quad \text{bázis } \mathcal{O}(A)\text{-ban.}$$

3. lépés: Ortogonális kiegészítő bázisának megválasztása vagyis bázist kell keresnünk $\mathcal{N}(A^T)$-ban: Keressük azokat a $\mathbf{d} \in \mathbb{R}^3$ vektorokat, amelyekre $A^T\mathbf{d} = \mathbf{0}$:

$$A^T = \begin{bmatrix} \frac{1}{2} & 0 & 1 \\ \frac{-1}{2} & 1 & 0 \end{bmatrix}.$$

Gauss-eliminációval

$$\left[\begin{array}{ccc|c} \frac{1}{2} & 0 & 1 & 0 \\ \frac{-1}{2} & 1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} \frac{1}{2} & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right],$$

ez lépcsős alakú. Megoldása:

$$\begin{bmatrix} -2s \\ -s \\ s \end{bmatrix} = s \cdot \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}.$$

Ezek szerint $\mathcal{N}(A^T)$ egy bázisa: $[-2, -1, 1]$ (vegyük észre, hogy ez párhuzamos a feladat elkészítésénél használt e egyenes irányvektorával!) Tehát a megfelelő teljes bázis $\mathbb{R}^3$-ban: $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{d}\}$.
4. lépés: b felírása az előbbi bázisban Keressük $\lambda_1, \lambda_2, \mu$ skalárokat úgy, hogy:

$$\mathbf{b} = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \mu\mathbf{d}$$

azaz

$$\begin{bmatrix} 5 \\ 7 \\ -1 \end{bmatrix} = \lambda_1 \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} \frac{-1}{2} \\ 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}.$$

Ez lineáris egyenletrendszer, előadáson Gauss-elminációval megoldottuk, a hallgatóktól kapott anyag erről az előadásról nagyon eltért attól, ami elhangzott, ezért újragépeltem. A Gauss-elimináció nem túl érdekes lépéseit nem gépelem be, csak a megoldást írom ide: $\lambda_1 = 2,\ \lambda_2 = 4,\ \mu = -3$. (visszahelyettesítéssel ellenőrizhető, hogy ez tényleg megoldás, és mivel $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{d}\}$ bázis, ezért több megoldás nincs is). Tehát:

$$\mathbf{b} = 2\mathbf{a}_1 + 4\mathbf{a}_2 + -3\mathbf{d}.$$

5. lépés: Vetület meghatározása. A vetület az $\mathcal{O}(A)$-ba eső rész, azaz a $\mathbf{d}$ irányú komponens elhagyásával:

$$\mathbf{b}' = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2$$

azaz

$$\mathbf{b}' = 2 \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} \frac{-1}{2} \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}.$$

Vegyük észre, hogy ez a b' pontosan az, amit a feladat kitalálása során az S síkon választottunk!

6. lépés: Az $A\mathbf{x} = \mathbf{b}'$ egyenlet megoldása. Az

$$\begin{cases} \frac{1}{2}x - \frac{1}{2}y = -1 \\ y = 4 \\ x = 2 \end{cases}$$

egyenletrendszert kell megoldani. Ez menne pl. Gauss-eliminációval, de most szerencsére rögtön látszik a megoldás: a két utolsó egyenletből $x = 2,\ y = 4$, és ez kielégíti az első egyenletet is. Azt kaptuk, hogy az optimális közelítő megoldás $x = 2,\ y = 4$, ami azt is jelenti, hogy az egyenletrendszerünk A mátrixának oszlopterében az $A \cdot [2, 4]^T$ vektor van az eredeti jobboldalhoz ($[5, 7, -1]$-hez) a lehető legközelebb.

Mellesleg, vegyük észre, hogy $A \cdot [2, 4]^T = [-1, 4, 2]$ éppen az a b' vektor, amit a feladat kitalálásakor választottunk az S síkon, és az optimális közelítő megoldást adó $[2, 4]^T$ sem véletlenül esik egybe a feladat kitalálása során a b' megválasztásához használt $s = 4,\ t = 2$ paraméter-értékekkel.