ALGEBRA I

Sügis 2021

Lektor: Valdis Laan

Konspekt: Valdis Laan ja Lauri Tart

Sisukord

Eessõna

Algebra kui matemaatikaharu võib jagada kaheks suureks osaks: lineaaralgebraks ja abstraktseks algebraks. Käesolev kursus koosneb põhiliselt lineaaralgebrast: vaatleme maatrikseid, determinante, lineaarvõrrandisüsteeme, vektorruume ja lineaarkujutusi. Abstraktne algebra uurib algebralisi struktuure. Struktuuridest tutvume vaid väga põgusalt kõige tähtsamatega: rühma, ringi ja korpusega.

Kursuse jooksul eeldame, et üliõpilane on tuttav selliste hulgateooria põhimõistetega nagu alamhulk, hulkade otsekorrutis, ühisosa, ühend, kujutus, binaarne seos, ekvivalentsiseos, ekvivalentsiklass. Samuti eeldame, et kuulajale on tuttavad naturaalarvude, täisarvude, ratsionaalarvude ja reaalarvude omadused.

Seda teksti lugedes panete tähele, et mõned kohad tekstis on väiksemas kirjas kui ülejäänud tekst. Nende kohtade lugemine ei ole muust materjalist arusaamiseks vajalik.

Kursuse jooksul kasutame mitmeid matemaatilisi ja matemaatilise loogika sümboleid, mille tähendused on järgmised:

$\forall a \in A$ — iga elemendi a korral hulgast A ehk hulga A iga elemendi a korral;

$\exists a \in A$ — leidub element a hulgas A ehk leidub hulga A element a;

$A \Rightarrow B$ — A-st järeldub B;

$A \Leftrightarrow B$ — A kehtib parajasti siis, kui kehtib B, ehk A kehtib siis ja ainult siis, kui kehtib B;

$\mathbb{N}$ — naturaalarvude hulk;

$\mathbb{Z}$ — täisarvude hulk;

$\mathbb{Q}$ — ratsionaalarvude hulk;

$\mathbb{R}$ — reaalarvude hulk.

1. Algebralised struktuurid

Algebraline struktuur on hulk, millel on defineeritud mingid tehted, mis rahuldavad teatud tingimusi. Käesolevas kursuses tutvume vaid kõige tähtsamate ja klassikalisemate algebraliste struktuuride definitsioonidega. Nendeks on rühm, ring, korpus ja vektorruum. Algebralisi struktuure uurib algebra haru, mida kutsutakse abstraktseks algebraks.

1.1. Rühmoid

Definitsioon 1.1. Kahekohaline (ehk **binaarne) algebraline tehe** hulgal A on kujutus hulgast $A \times A$ hulka A.

Märkus 1.2. Sõna "algebraline" eelmises definitsioonis rõhutab seda, et tehte tulemus kuulub ka hulka A. Põhimõtteliselt võib vaadelda ka kahekohalisi tehteid $A \times A \to B$, kus $B \neq A$. Näiteks kolmemõõtmelise ruumi vabavektorite liitmine on algebraline tehe $\mathbb{E}_3 \times \mathbb{E}_3 \to \mathbb{E}_3$, aga skalaarkorrutamine on kahekohaline tehe $\mathbb{E}_3 \times \mathbb{E}_3 \to \mathbb{R}$, mis ei ole algebraline.

Märkus 1.3. Algebras võib vaadelda mitte ainult kahekohalisi, vaid ka suvalisi n-kohalisi algebralisi tehteid, kus $n \in \{0, 1, 2, \ldots\}$. n-kohaline algebraline tehe hulgal A on kujutus $A^n \to A$. Muuhulgas ühekohalised algebralised tehted on kujutused $A \to A$ ja nullkohalised algebralised tehted on kujutused $A^0 \to A$. Kuna A^0 on hulk, milles on üks element, siis kujutuse $A^0 \to A$ defineerimine tähendab sisuliselt ühe konkreetse elemendi väljavalimist hulgast A.

Niisiis kahekohaline algebraline tehe hulgal A on eeskiri, mis igale hulga A elementide järjestatud paarile (a, b) seab vastavusse hulga A mingi elemendi. Kahekohalisest tehtest rääkides kirjutatakse tehtemärk harilikult hulga elementide vahele. Seega näiteks kahekohalise tehte $* : A \times A \to A$ korral kirjutatakse $*((a, b))$ asemel harilikult $a * b$ ja võib öelda, et tehe $*$ seab paarile (a, b) vastavusse hulga A elemendi $a * b$.

Näide 1.4. Naturaalarvude hulgal[1] $\mathbb{N}$ võib vaadelda kahekohalisi algebralisi tehteid, mis on defineeritud näiteks järgmiste eeskirjadega:

- $(a, b) \mapsto a + b$;
- $(a, b) \mapsto a^b$;
- $(a, b) \mapsto b + 5$;
- $(a, b) \mapsto ab$;
- $(a, b) \mapsto b^a$;
- $(a, b) \mapsto 3$.

Lahutamistehe ei ole algebraline tehe hulgal $\mathbb{N}$, sest näiteks $1 - 2 \notin \mathbb{N}$.

Definitsioon 1.5. Rühmoid on hulk koos sellel defineeritud kahekohalise algebralise tehtega.

Kui hulk on A ja tehe sellel $*$, siis sellist rühmoidi tähistatakse ka järjestatud paarina $(A, *)$. Näiteks võib rääkida rühmoidist $(\mathbb{N}, +)$ või $(\mathbb{Z}, -)$. Samas $(\mathbb{N}, -)$ ei ole rühmoid.

Kui räägitakse rühmoididest üldiselt (s.t. kui ei peeta silmas ühtegi konkreetset hulka ega tehet), siis kutsutakse tehet $*$ kokkuleppeliselt korrutamiseks ja elementi $a * b$ elementide a ja b korrutiseks. Veelgi enam, tihti jäetakse tehtemärk üldse ära ja kirjutatakse $a * b$ asemel ab.

Klassikaliselt on algebras tavaks eeldada, et algebraline struktuur on mittetühi. Siiski on teatud olukordades otstarbekas lubada ka tühje struktuure. Käesolevas kursuse me loeme ka tühja hulga koos tühja kujutusega $\emptyset \times \emptyset \to \emptyset$ rühmoidiks.

[1] Käesolevas kursuses loeme, et $\mathbb{N} = \{1, 2, 3, \ldots\}$.

1.2. Poolrühm ja monoid

Definitsioon 1.6. Kahekohalist algebralist tehet $*$ hulgal A nimetatakse

1. **assotsiatiivseks**, kui $(a * b) * c = a * (b * c)$ iga $a, b, c \in A$ korral;
2. **kommutatiivseks**, kui $a * b = b * a$ iga $a, b \in A$ korral.

Definitsioon 1.7. Poolrühm on rühmoid, mille tehe on assotsiatiivne.

Näide 1.8. Poolrühmadeks on näiteks $(\mathbb{N}, +)$, $(\mathbb{R}, \cdot)$ ja mingi hulga X kõigi teisenduste hulk $\mathcal{T}(X)$ nende teisenduste järjestrakendamise tehte $\circ$ suhtes.

Kui meil on kahekohaline algebraline tehe $*$ hulgal A ja rohkem kui kaks hulga A elementi, siis sulgude abil saab ära näidata, millises järjekorras me tehet $*$ neile elementidele peame rakendama. Näiteks kirjapanek $a * (b * c)$ näitab, et kõigepäält tuleb leida element $b * c$ ning seejärel rakendada tehet $*$ elementidele a ja $b * c$. Tulemuseks on hulga A mingi element. Kolme elemendi korral on kaks võimalust sulgude paigutamiseks: $(a * b) * c$ ja $a * (b * c)$. Kui elemente on rohkem, siis on ka sulgude paigutamise võimalusi rohkem, näiteks $(a * b) * (c * d)$, $((a * b) * c) * d$, $a * ((b * c) * d)$ jne. Siin me vaatleme ainult selliseid avaldisi, kus sulgude paigutus on korrektne, s.t. et sulud määravad ära, millises järjekorras tuleb tehteid sooritada. Näiteks avaldis $a * b * (c * d)$ ei ole korrektne, sest ei ole selge, kas enne tuleb leida $a * b$ või $b * (c * d)$.

Osutub, et assotsiatiivse tehte korral ei sõltu kõigi tehete sooritamise järel tulemuseks saadav element sulgude paigutusest.

Lause 1.9. *Tehte rakendamise tulemus poolrühmas ei sõltu sulgude paigutusest.*

Tõestus. Olgu meil tegemist poolrühmaga $(A, *)$. Tõestame matemaatilise induktsiooniga, et iga $n \in \mathbb{N}$ ja mistahes elementide $a_1, \ldots, a_n \in A$ korral ei sõltu tehte $*$ neile elementidele rakendamise tulemus sulgude paigutusest.

Kui $n = 1$ või $n = 2$, siis on väide ilmne. Kui $n = 3$, siis järeldub väide tehte $*$ assotsiatiivsusest. Olgu nüüd $n > 3$ ja eeldame, et väide kehtib, kui elemente on vähem kui n. Näitame, et see kehtib siis ka n korral. Tähistame iga $k \in \mathbb{N}$ ja $b_1, \ldots, b_k \in A$ korral

$$b_1 * b_2 * \ldots * b_k := (\ldots((b_1 * b_2) * b_3) * \ldots * b_{k-1}) * b_k. \tag{1}$$

(See tähendab, et $b_1 * b_2 * \ldots * b_k$ on element, mille saame, kui rakendame tehet $*$ niiöelda vasakult paremale.) Paneme tähele, et

$$b_1 * b_2 * \ldots * b_k := (b_1 * b_2 * \ldots * b_{k-1}) * b_k, \tag{2}$$

kui $k \geq 2$. Induktsiooni eelduse põhjal on mistahes sulgude paigutusega avaldis vähem kui n elemendist võrdne avaldisega, mis on kujul (1). Tõestuse lõpetamiseks piisab näidata, et sama omadus on ka igal n elemendist moodustatud avaldisel. Vaatleme sellist avaldist ja leiame selles üles tehte $*$ viimase rakendamise koha:

$$(a_1 * \ldots * a_k) * (a_{k+1} * \ldots * a_n). \tag{3}$$

Siin sulgudes olevad avaldised võib kirjutada kujul (1), sest neis esineb vähem kui n elementi. On kaks võimalust.

1) $k = n - 1$. Siis on avaldis (3) kujul $(a_1 * \ldots * a_{n-1}) * a_n = a_1 * \ldots * a_{n-1} * a_n$, s.t. kujul (1).

2) $k < n-1$. Siis

$$\begin{aligned}(a_1 * \ldots * a_k) * (a_{k+1} * \ldots * a_n) &= (a_1 * \ldots * a_k) * ((a_{k+1} * \ldots * a_{n-1}) * a_n) && \text{(omadus (2))}\\ &= ((a_1 * \ldots * a_k) * (a_{k+1} * \ldots * a_{n-1})) * a_n && \text{(assotsiatiivsus)}\\ &= (a_1 * \ldots * a_k * a_{k+1} * \ldots * a_{n-1}) * a_n && \text{(induktsiooni eeldus)}\\ &= a_1 * \ldots * a_k * a_{k+1} * \ldots * a_{n-1} * a_n, && \text{(omadus (2))}\end{aligned}$$

s.t. jällegi on vaadeldav avaldis võrdne avaldisega kujul (1). □

Arvestades lauset 1.9 jäetakse assotsiatiivse tehte korral sulud tihti üldse kirjutamata. Rühmoidis võib leiduda teatud eriliste omadustega elemente.

Definitsioon 1.10. Rühmoidi $(A, *)$ elementi e nimetatakse **ühikelemendiks**, kui

$$a * e = a \quad \text{ja} \quad e * a = a$$

iga $a \in A$ korral.

Lause 1.11. *Rühmoidis võib olla ainult üks ühikelement.*

Tõestus. Olgu e ja f rühmoidi $(A, *)$ ühikelemendid. Kuna e on ühikelement, siis $f = e * f$. Et f on ühikelement, siis $e * f = e$. Järelikult $f = e * f = e$. □

Definitsioon 1.12. Monoidiks nimetatakse poolrühma, milles leidub ühikelement.

Näide 1.13. Monoidideks on näiteks

- $(\mathbb{N}, \cdot)$ (ühikelement on arv 1),
- $(\mathbb{Z}, +)$ (ühikelement on arv 0) ja
- $(\mathcal{T}(X), \circ)$ (ühikelement on hulga X samasusteisendus).

Samas $(\mathbb{N}, +)$ on poolrühm, mis ei ole monoid.

Definitsioon 1.14. Olgu $(A, *)$ monoid ühikelemendiga e. Selle monoidi elementi a nimetatakse **pööratavaks**, kui leidub selline element $b \in A$, et

$$a * b = e \quad \text{ja} \quad b * a = e.$$

Sellisel juhul öeldakse, et elemendid a ja b on teineteise **pöördelemendid**.

Näide 1.15. Monoidi $(\mathbb{N}, \cdot)$ ainus pööratav element on 1, selle pöördelement on 1 ise.
Monoidi $(\mathbb{Z}, \cdot)$ pööratavad elemendid on 1 ja -1, sest $1 \cdot 1 = 1$ ja $(-1) \cdot (-1) = 1$.
Monoidi $(\mathcal{T}(X), \circ)$ pööratavad elemendid on hulga X bijektiivsed teisendused.

Lause 1.16. *Kui monoidi element on pööratav, siis selle elemendi pöördelement on üheselt määratud.*

TÕESTUS. Olgu $(A, *)$ monoid ühikelemendiga e ja a selle monoidi mingi element. Oletame, et elemendid b ja c on elemendi a pöördelemendid. Siis

$$c = e * c = (b * a) * c = b * (a * c) = b * e = b.$$

Sellega oleme näidanud, et elemendi a pöördelement on üheselt määratud. □

Monoidi pööratava elemendi a (üheselt määratud) pöördelementi tähistatakse sümboliga a^{-1}. Pöördelemendi definitsioonist on selge, et pööratava elemendi a korral

$$\left(a^{-1}\right)^{-1} = a.$$

Järgnev lause annab eeskirja korrutise pöördelemendi leidmiseks, kui meil on teada mõlema teguri pöördelemendid.

Lause 1.17. *Monoidi $(A, *)$ mistahes pööratavate elementide a ja b korral on ka element $a * b$ pööratav ja*

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

TÕESTUS. Paneme tähele, et

$$\begin{aligned}
(a * b) * (b^{-1} * a^{-1}) &= ((a * b) * b^{-1}) * a^{-1} && \text{(assotsiatiivsus)} \\
&= (a * (b * b^{-1})) * a^{-1} && \text{(assotsiatiivsus)} \\
&= (a * e) * a^{-1} && \text{(pöördelemendi def.)} \\
&= a * a^{-1} && \text{(ühikelemendi def.)} \\
&= e. && \text{(pöördelemendi def.)}
\end{aligned}$$

Analoogiliselt saab tõestada võrduse $(b^{-1} * a^{-1}) * (a * b) = e$. Seega $a * b$ ja $b^{-1} * a^{-1}$ on teineteise pöördelemendid. □

1.3. Rühm ja Abeli rühm

Üheks klassikalisemaks algebraliseks struktuuriks on rühm.

Definitsioon 1.18. Rühm on hulk G koos kahekohalise algebralise tehtega $*$, mis rahuldab järgmisi tingimusi:

G1. $(a * b) * c = a * (b * c)$ iga $a, b, c \in G$ korral;

G2. leidub element $e \in G$ nii, et $a * e = a = e * a$ iga $a \in G$ korral;

G3. iga $a \in G$ korral leidub element $b \in G$ nii, et $a * b = e = b * a$.

Teiste sõnadega võib öelda, et rühm on monoid, mille kõik elemendid on pööratavad.

Näide 1.19. 1. Nullist erinevate ratsionaalarvude ja nullist erinevate reaalarvude hulgad on rühmad korrutamise suhtes. Seega võime vaadelda rühmi $(\mathbb{R} \setminus \{0\}, \cdot)$ ja $(\mathbb{Q} \setminus \{0\}, \cdot)$. Samas $(\mathbb{R}, \cdot)$ ja $(\mathbb{Q}, \cdot)$ ei ole rühmad (miks?).

2. Positiivsete ratsionaalarvude ja positiivsete reaalarvude hulgad on rühmad korrutamise suhtes.

3. Hulgad $\mathbb{Z}, \mathbb{Q}$ ja $\mathbb{R}$ on rühmad liitmise suhtes.

4. Hulk $\{1, -1\}$ on rühm korrutamise suhtes.

5. Mittetühja hulga X bijektiivsete teisenduste hulk $\mathcal{S}(X)$ on rühm kujutuste järjestrakendamise suhtes. Selle rühma ühikelement on hulga X samasusteisendus ja teisenduse f pöördelement on selle teisenduse pöördteisendus.

6. Kui hulgas X on vähemalt kaks elementi, siis $(\mathcal{T}(X), \circ)$ on monoid, mis ei ole rühm. Näiteks konstantsed kujutused on sellisel juhul mittepööratavad.

Rühma ühikelementi tähistatakse tihti ka sümboliga 1.

Tänu lausele 1.16 võib öelda, et rühma igal elemendil on täpselt üks pöördelement. Nii nagu monoididegi korral tähistatakse rühma elemendi a (üheselt määratud) pöördelementi sümboliga a^{-1}.

Märkus 1.20. Arvestades märkust 1.3 võib rühma vaadelda ka hulgana G, millel on antud kolm algebralist tehet:

- kahekohaline tehe $(a, b) \mapsto a * b$ (korrutamine),
- ühekohaline tehe $a \mapsto a^{-1}$ (pöördelemendi võtmine),
- nullkohaline tehe (ühikelemendi fikseerimine).

Definitsioon 1.21. Rühma nimetatakse **kommutatiivseks** ehk **Abeli rühmaks**, kui selle tehe on kommutatiivne.

Näide 1.22. Näite 1.19 neljas esimeses punktis loetletud rühmad on kommutatiivsed. Bijektiivsete teisenduste rühm $\mathcal{S}(\{1, 2, 3\})$ ei ole kommutatiivne.

Abeli rühmadest üldiselt kõneldes on tavaks kasutada nn. aditiivset sümboolikat. Tehtemärgina kasutatakse märki + ja tehet kutsutakse liitmiseks, ühikelementi kutsutakse nullelemendiks ja kasutatakse sümbolit 0, elemendi a pöördelementi kutsutakse vastandelemendiks ja tähistatakse sümboliga $-a$. Seega Abeli rühma definitsioon sõnastatakse harilikult järgmisel kujul.

Definitsioon 1.23. Abeli rühm on hulk A koos kahekohalise algebralise tehtega $+$, mis rahuldab järgmisi tingimusi:

AG1. $(a + b) + c = a + (b + c)$ iga $a, b, c \in A$ korral;

AG2. leidub element $0 \in A$ nii, et $a + 0 = a = 0 + a$ iga $a \in A$ korral;

AG3. iga $a \in A$ korral leidub element $-a \in A$ nii, et $a + (-a) = 0 = (-a) + a$;

AG4. $a + b = b + a$ iga $a, b \in A$ korral.

Abeli rühma korral räägitakse ka elementide a ja b **vahest**, mis defineeritakse võrdusega

$$a - b := a + (-b).$$

Tingimuse AG3 põhjal on selge, et iga a korral

$$a - a = 0.$$

Järgmist omadust läheb Abeli rühmades arvutades väga tihti vaja.

Lause 1.24. *Kui* $(A,+)$ *on Abeli rühm ja* $a,b,c \in A$*, siis*

$$a+b=c \implies a=c-b.$$

TÕESTUS. Kehtigu võrdus $a+b=c$, kus $a,b,c \in A$. Liites selle võrduse mõlemale poolele elemendi $-b$ saame

$$(a+b)+(-b)=c+(-b)=c-b.$$

Kuna

$$(a+b)+(-b) \underset{AG1}{=} a+(b+(-b)) \underset{AG3}{=} a+0 \underset{AG2}{=} a,$$

siis saamegi võrduse $a=c-b$. □

Tõestatud omadust võib sõnastada ka nii: *Abeli rühmas võib elemendi viia võrduse ühelt poolelt teisele poolele vastandmärgiga.*

1.4. Ring ja korpus

Vaatleme nüüd selliseid struktuure, milles on kaks kahekohalist algebralist tehet. Näiteks arvuhulkade puhul on loomulik vaadelda nii liitmist kui korrutamist, samuti ruutmaatriksite korral. Nendel tehetel on mitmeid häid omadusi.

Definitsioon 1.25. Ring on hulk R koos kahe kahekohalise algebralise tehtega $+$ ja $\cdot$, mis rahuldavad järgmisi tingimusi:

R1. $(a+b)+c=a+(b+c)$ iga $a,b,c \in R$ korral;

R2. leidub element $0 \in R$ nii, et $a+0=a=0+a$ iga $a \in R$ korral;

R3. iga $a \in R$ korral leidub element $-a \in R$ nii, et $a+(-a)=0=(-a)+a$;

R4. $a+b=b+a$ iga $a,b \in R$ korral;

R5. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ iga $a,b,c \in R$ korral;

R6. leidub element $1 \in R$ nii, et $a \cdot 1 = a = 1 \cdot a$ iga $a \in R$ korral;

R7. $a \cdot (b+c) = a \cdot b + a \cdot c$ iga $a,b,c \in R$ korral;

R8. $(a+b) \cdot c = a \cdot c + b \cdot c$ iga $a,b,c \in R$ korral.

Harilikult kirjutatakse ringi puhul $a \cdot b$ asemel lühemalt ab. Definitsiooni tingimustest R1–R4 näeme, et ring peab liitmise suhtes olema Abeli rühm. Tingimusi R7 ja R8 kutsutakse **distributiivsuse seadusteks**.

Märkus 1.26. Algebras esineb erinevaid ringi definitsioone. Vahel ei nõuta tingimust R6, mõnikord isegi tingimust R5. Seega algebralisi tekste lugedes peab hoolega jälgima, milliseid ringe silmas peetakse. Käesolevas kursuses vaatleme vaid selliseid ringe, mis rahuldavad kõiki tingimusi R1–R8, s.t. ühikelemendiga assotsiatiivseid ringe.

Ringides on terve rida omadusi, mida küll definitsioonis ei nõuta, kuid mis järelduvad definitsioonist lihtsasti ja mis aitavad arvutusi ringi elementidega läbi viia. Loetleme neist mõned.

Lause 1.27. *Kui R on ring ja $a, b, c \in R$, siis*

1. $0a = 0 = a0$;
2. $(-a)b = a(-b) = -ab$;
3. $a(b-c) = ab - ac$;
4. $(a-b)c = ac - bc$.

Tõestus. 1. Olgu $a \in R$. Siis

$$0a \underset{R2}{=} (0+0)a \underset{R8}{=} 0a + 0a.$$

Kasutades lauset 1.24 saame võrduse

$$0a = 0a - 0a \underset{R3}{=} 0.$$

Võrduse $a0 = 0$ saab tõestada analoogiliselt.

2. Olgu $a, b \in R$. Siis

$$0 = 0b \underset{R3}{=} ((-a) + a)b \underset{R8}{=} (-a)b + ab.$$

Lause 1.24 põhjal

$$(-a)b = 0 - ab \underset{R2}{=} -ab.$$

Võrduse $a(-b) = -ab$ saab tõestada analoogiliselt.

3. Olgu $a, b, c \in R$. Siis

$$a(b-c) = a(b+(-c)) \underset{R7}{=} ab + a(-c) \underset{om.\ 2}{=} ab + (-ac) = ab - ac.$$

4. Võrduse $(a-b)c = ac - bc$ saab tõestada analoogiliselt. □

Definitsioon 1.28. Ringi nimetatakse **kommutatiivseks**, kui tema korrutamistehe on kommutatiivne.

Näide 1.29. $(\mathbb{Z}, +, \cdot)$ on kommutatiivne ring.

Definitsioon 1.30. Ringi $(R, +, \cdot)$ elementi nimetatakse **pööratavaks**, kui tal leidub pöördelement selle ringi korrutamistehte suhtes, s.t. kui see element on pööratav monoidis $(R, \cdot)$. Ringi R kõigi pööratavate elementide hulka tähistatakse sümboliga $U(R)$.

Näide 1.31. 1. Ringi $\mathbb{Z}$ korral $U(\mathbb{Z}) = \{1, -1\}$.

2. Ringi $\mathbb{R}$ korral $U(\mathbb{R}) = \mathbb{R} \setminus \{0\}$.

3. Reaalarvuliste elementidega n-ndat järku ruutmaatriksite hulk $\mathrm{Mat}_n(\mathbb{R})$ on ring maatriksite liitmise ja korrutamise suhtes. Ruutmaatriksite ringi nullelement on nullmaatriks ja ühikelement on ühikmaatriks. Kui $n \geqslant 2$, siis see ring ei ole kommutatiivne. Selles ringis on näiteks nullmaatriksist erinev maatriks

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

mittepööratav.

Definitsioon 1.32. Korpus on kommutatiivne ring, milles on vähemalt kaks elementi ja mille kõik nullelemendist erinevad elemendid on pööratavad.

Märkus 1.33. Ringi puhul on võimalik, et $1 = 0$ (ühikelement võrdub nullelemendiga). Sellisel juhul see ring koosnebki ainult ühest elemendist, sest mistahes elemendi a korral $a = a1 = a0 = 0$. Järelikult, kui ringis on vähemalt kaks elementi, siis selles ringis $1 \neq 0$. Muuhulgas korpuses on alati ühikelement nullelemendist erinev.

Arvestades tingimusi R5 ja R6 näeme, et korpuse nullelemendist erinevad elemendid moodustavad rühma korrutamise suhtes.

Märkus 1.34. Eestikeelses kirjanduses (nt. [1], [2]) ei ole tihti nõutud, et korpuse korrutamine oleks kommutatiivne. Käesolevas kursuses me seda siiski nõuame ja loodame, et sellest ei teki segadust.

Näide 1.35. 1. $(\mathbb{Z}, +, \cdot)$ on kommutatiivne ring, mis ei ole korpus (nt. elemendil $2 \in \mathbb{Z}$ ei leidu ringis $\mathbb{Z}$ pöördelementi korrutamise suhtes).

2. $(\mathbb{Q}, +, \cdot)$ ja $(\mathbb{R}, +, \cdot)$ on korpused.

Märkus 1.36. Korpuste puhul (nt. $\mathbb{Q}$ ja $\mathbb{R}$ korral) räägitakse sageli jagamisest. Nimelt öeldakse, et korpuse K elemendi a ja nullist erineva elemendi b jagatis on element ab^{-1} (elemendi a ja elemendi b pöördelemendi korrutis) ning tähistatakse seda elementi sümboliga $\frac{a}{b}$. Sellise tähistuse korral saab kasutada harilikke murdudega arvutamise reegleid. Näiteks $\frac{a}{b} \cdot c = \frac{ac}{b}$. Tõepoolest,

$$\frac{a}{b} \cdot c = (ab^{-1})c = a(b^{-1}c) = a(cb^{-1}) = (ac)b^{-1} = \frac{ac}{b}.$$

Lugeja võiks proovida iseseisvalt veenduda, et korpuses

$$\frac{a}{b} = \frac{c}{d} \iff ad = bc.$$

Definitsioon 1.37. Ringi R nullist erinevaid elemente a ja b nimetatakse **nullitegureiks**, kui $ab = 0$. Ring on **nullitegureita**, kui temas ei ole nullitegureid.

Lause 1.38. *Korpuses ei ole nullitegureid.*

Tõestus. Olgu K korpus ja oletame vastuväiteliselt, et $a, b \in K$ on nullist erinevad elemendid, mille korral $ab = 0$. Korrutades selle võrduse mõlemaid pooli paremalt elemendiga b^{-1}, saame

$$a = a1 = a(bb^{-1}) = (ab)b^{-1} = 0b^{-1} = 0,$$

mis on vastuolus eeldusega. □

Ka ring $\mathbb{Z}$, mis ei ole korpus, on nullitegureita. Nullitegureid leidub näiteks ringis $\mathrm{Mat}_2(\mathbb{R})$.

Definitsioon 1.39. Öeldakse, et ring R on **taandamisega**, kui mistahes $a, b, c \in R$, $c \neq 0$ korral võrdusest $ac = bc$ järeldub võrdus $a = b$ ja võrdusest $ca = cb$ järeldub võrdus $a = b$.

Lause 1.40. *Ring on taandamisega parajasti siis, kui ta on nullitegureita.*

Tõestus. Tarvilikkus. Olgu R taandamisega ring. Oletame, et $ab = 0$, kus $a, b \in R$. Kui $a \neq 0$, siis võrdusest $ab = a0$ elementi a taandades saame võrduse $b = 0$. Seega võrdusest $ab = 0$ järeldub, et kas $a = 0$ või $b = 0$, mis tähendab, et R on nullitegureita.
Piisavus. Olgu R nullitegureita ring, $a, b, c \in R$, $c \neq 0$ ja $ac = bc$. Siis $ac - bc = 0$ ehk $(a-b)c = 0$. Kuna $c \neq 0$ ja nullitegurid puuduvad, siis $a - b = 0$ ehk $a = b$. Teise implikatsiooni saab tõestada analoogiliselt. □

1.5. Jäägiklassiringid

Selles paragrahvis tutvume teatud lõplike ringidega, mida kasutatakse palju nii arvuteoorias kui arvutiteaduses.

Fikseerime mingi naturaalarvu $n \geqslant 2$. On teada, et iga täisarvu a jaoks leiduvad üheselt määratud täisarvud q, r nii, et

$$a = qn + r \quad \text{ja} \quad 0 \leqslant r < n.$$

Arvu r nimetatakse **jäägiks**, mis tekib arvu a jagamisel arvuga n ning arvude q ja r leidmist kutsutakse **jäägiga jagamiseks.** Näiteks

$$17 = 3 \cdot 5 + 2, \quad \text{kus} \quad 0 \leqslant 2 < 5,$$

mis tähendab, et 17 jagamisel 5-ga saame jäägi 2.

Definitsioon 1.41. Öeldakse, et täisarv b **jagab** täisarvu a (ja kirjutatakse $b \mid a$), kui leidub selline täisarv c, et $bc = a$.

Teiste sõnadega võib öelda, et arv b jagab arvu a, kui a jagamisel arvuga b tekib jääk 0.

Lihtne on veenduda, et kui $a, b, c, d \in \mathbb{Z}$, $a \mid b$ ja $a \mid c$, siis $a \mid b \pm c$ ja $a \mid bd$.

Definitsioon 1.42. Öeldakse, et täisarvud a ja b on **kongruentsed** mooduli n järgi, kui $n \mid a - b$. Tähistus: $a \equiv b \pmod{n}$.

Lause 1.43. *Täisarvud a ja b on kongruentsed mooduli n järgi parajasti siis, kui nad annavad arvuga n jagades sama jäägi.*

Tõestus. Tarvilikkus. Olgu $a = qn + r$, kus $0 \leqslant r < n$. Eeldame, et $a \equiv b \pmod{n}$. Siis $n \mid a - b$, mis tähendab, et $nk = a - b$ mingi $k \in \mathbb{Z}$ korral. Järelikult

$$b = a - nk = qn + r - nk = (q - k)n + r.$$

Siit näeme, et ka b annab arvuga n jagamisel jäägi r.

Piisavus. Eeldame, et $a = q_1 n + r$ ja $b = q_2 n + r$, kus $q_1, q_2, r \in \mathbb{Z}$ ja $0 \leqslant r < n$. Siis $a - b = (q_1 - q_2)n$, mis tähendab, et $n \mid a - b$. □

Kuna seos "a ja b annavad arvuga n jagades sama jäägi" on ilmselt ekvivalentsiseos täisarvude hulgal, siis on ka kongruentsusseos (mooduli n järgi) ekvivalentsiseos täisarvude hulgal. Iga ekvivalentsiseose puhul võib vaadelda ekvivalentsiklasse selle seose järgi. Kongruentsusseose puhul tähistatakse arvu $a \in \mathbb{Z}$ ekvivalentsiklassi sümboliga $\overline{a}$ ja nimetatakse arvu a **jäägiklassiks** mooduli n järgi. Niisiis

$$\overline{a} \;=\; \{b \in \mathbb{Z} \mid a \equiv b \pmod{n}\} = \{b \in \mathbb{Z} \mid a \text{ ja } b \text{ annavad } n\text{-ga jagamisel sama jäägi}\}.$$

Nagu hulgateooriast teada, võib ekvivalentsiklassi esindajaks valida mistahes elemendi sellest ekvivalentsiklassist. Näiteks kui $n = 5$, siis $\overline{2} = \overline{17} = \overline{-8}$.

Et erinevaid jääke, mis n-ga jagamisel saab tekkida, on täpselt n tükki (need jäägid on $0, 1, \ldots, n-1$), siis mooduli n järgi on täpselt n jäägiklassi. Nende jäägiklasside hulka tähistatakse

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

Defineerime nüüd tehted jäägiklasside hulgal.

Definitsioon 1.44. Jäägiklasside $\overline{a}, \overline{b} \in \mathbb{Z}_n$ **summa** ja **korrutis** defineeritakse võrdustega

$$\begin{aligned} \overline{a} + \overline{b} &= \overline{a+b}, \\ \overline{a} \cdot \overline{b} &= \overline{a \cdot b}. \end{aligned}$$

Lemma 1.45. *Jäägiklasside liitmine ja korrutamine on korrektselt defineeritud.*

Tõestus. Korrektsuse tõestamiseks peame näitama, et tehte tulemus ei sõltu sellest, millised täisarvud me jäägiklasside esindajateks valime. Oletame, et $\overline{a_1} = \overline{a_2}$ ja $\overline{b_1} = \overline{b_2}$. Siis $n \mid a_1 - a_2$ ja $n \mid b_1 - b_2$. Järelikult

$$n \mid (a_1 - a_2) + (b_1 - b_2) = (a_1 + b_1) - (a_2 + b_2),$$

s.t. $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$ ja $\overline{a_1 + b_1} = \overline{a_2 + b_2}$. Et

$$a_1 b_1 - a_2 b_2 = a_1(b_1 - b_2) + b_2(a_1 - a_2)$$

ja n jagab selle võrduse paremat poolt, siis $n \mid a_1 b_1 - a_2 b_2$ ning järelikult $a_1 b_1 \equiv a_2 b_2 \pmod{n}$ ehk $\overline{a_1 b_1} = \overline{a_2 b_2}$. □

Teoreem 1.46. *Hulk $\mathbb{Z}_n$ on eespool defineeritud tehete suhtes kommutatiivne ring. See ring on korpus parajasti siis, kui n on algarv.*

Tõestus. Tehete definitsioonidest ja täisarvude omadustest järeldub kergesti, et $\mathbb{Z}_n$ on kommutatiivne ring. Üksikasjade kontrollimise jätame lugejale.

Tõestame teise väite. Oletame, et $\mathbb{Z}_n$ on korpus, kuid n ei ole algarv, s.t. leiduvad sellised $a, b \in \mathbb{N}$, $1 < a, b < n$, et $n = ab$. Siis $\overline{a}\overline{b} = \overline{n} = \overline{0}$, kuid $\overline{a} \neq \overline{0}$ ja $\overline{b} \neq \overline{0}$. Kuna korpuses ei ole nullitegureid (lause 1.38), siis oleme saanud vastuolu. Seega n peab olema algarv.

Oletame nüüd, et n on algarv. Võtame nullist erineva elemendi $\overline{a}$, kus $1 \leqslant a \leqslant n-1$, ringis $\mathbb{Z}_n$. Siis $\text{SÜT}(a, n) = 1$. On teada (vt. [1], lause 6.3.5), et sellisel juhul leiduvad $x, y \in \mathbb{Z}$ nii, et $ax + ny = 1$. Järelikult

$$\overline{1} = \overline{ax + ny} = \overline{ax} + \overline{ny} = \overline{a}\,\overline{x} + \overline{0}\,\overline{y} = \overline{a}\,\overline{x} + 0 = \overline{a}\,\overline{x},$$

mis tähendab, et element $\overline{a}$ on pööratav. Kuna iga nullist erinev element on pööratav, siis oleme näidanud, et $\mathbb{Z}_n$ on korpus. □

Definitsioon 1.47. Ringe $\mathbb{Z}_n$, kus $n \in \mathbb{N}$, kutsutakse **jäägiklassiringideks**. Korpusi $\mathbb{Z}_p$, kus p on algarv, kutsutakse **jäägiklassikorpusteks**.

Näide 1.48. Kuna 5 on algarv, siis $\mathbb{Z}_5$ on korpus. Selles korpuses

$$\begin{aligned} \overline{1}^{-1} &= \overline{1}, \quad \text{sest } \overline{1} \cdot \overline{1} = \overline{1}, \\ \overline{2}^{-1} &= \overline{3}, \quad \text{sest } \overline{2} \cdot \overline{3} = \overline{1}, \\ \overline{3}^{-1} &= \overline{2}, \quad \text{sest } \overline{2} \cdot \overline{3} = \overline{1}, \\ \overline{4}^{-1} &= \overline{4}, \quad \text{sest } \overline{4} \cdot \overline{4} = \overline{1}. \end{aligned}$$

Ring $\mathbb{Z}_6$ ei ole korpus, sest näiteks nullist erineval elemendil $\overline{2}$ ei leidu pöördelementi. Selles ringis leidub nullitegureid: $\overline{2} \cdot \overline{3} = \overline{0}$, kuigi $\overline{2} \neq \overline{0}$ ja $\overline{3} \neq \overline{0}$.

2. Kompleksarvud

2.1. Kompleksarvude korpus

Meile hästituntud ratsionaalarvude hulk $\mathbb{Q}$ ja reaalarvude hulk $\mathbb{R}$ on korpused. Samuti nägime, et iga algarvu p jaoks on olemas p-elemendiline jäägiklassikorpus $\mathbb{Z}_p$. Kuigi reaalarvude korpusel on palju häid omadusi, on tal ka üks oluline puudus: temas ei ole võimalik leida võrrandi

$$x^2 = -1$$

lahendit (ja, nagu me teame, ka mitmete teiste ruutvõrrandite lahendeid). Sellest puudusest üle saamiseks konstrueeritakse üks suurem arvuhulk, mis sisaldab reaalarvude hulka ja kus antud võrrand on lahenduv. Seda suuremat hulka kutsutakse kompleksarvude hulgaks. Kompleksarvude defineerimiseks on mitmeid võimalusi. Meie teeme seda defineerides järjestatud reaalarvupaaride hulgal $\mathbb{R}^2$ tehted sobival viisil.

Lause 2.1. *Hulk* $\mathbb{R}^2 = \{(a,b) \mid a,b \in \mathbb{R}\}$ *on korpus liitmise ja korrutamise suhtes, mis on defineeritud võrdustega*

$$\begin{aligned}(a,b)+(c,d) &= (a+c,b+d),\\ (a,b)(c,d) &= (ac-bd,bc+ad).\end{aligned}$$

Tõestus. Kui $a,b,c,d \in \mathbb{R}$, siis $a+c, b+d, ac-bd, bc+ad \in \mathbb{R}$. Seega on defineeritud liitmine ja korrutamine algebralised tehted hulgal $\mathbb{R}^2$. Lihtne on veenduda, et $(\mathbb{R}^2,+)$ on Abeli rühm, kus nullelemendiks on paar $(0,0)$ ja paari (a,b) vastandelement on paar $(-a,-b)$. Kui $a,b,c,d,e,f \in \mathbb{R}$, siis

$$\begin{aligned}((a,b)(c,d))(e,f) &= (ac-bd,bc+ad)(e,f)\\ &= ((ac-bd)e-(bc+ad)f,(bc+ad)e+(ac-bd)f)\\ &= (ace-bde-bcf-adf,bce+ade+acf-bdf)\\ &= (a(ce-df)-b(de+cf),b(ce-df)+a(de+cf))\\ &= (a,b)(ce-df,de+cf)\\ &= (a,b)((c,d)(e,f)),\\ (a,b)(c,d) &= (ac-bd,bc+ad) = (ca-db,da+cb) = (c,d)(a,b),\end{aligned}$$

mis tähendab, et korrutamine on assotsiatiivne ja kommutatiivne. Et mistahes $a,b \in \mathbb{R}$ korral

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

ja kommutatiivsuse tõttu ka $(1,0)(a,b) = (a,b)$, siis $(1,0)$ on ühikelement korrutamise suhtes. Kui $a,b,c,d,e,f \in \mathbb{R}$, siis

$$\begin{aligned}(a,b)((c,d)+(e,f)) &= (a,b)(c+e,d+f)\\ &= (a(c+e)-b(d+f),b(c+e)+a(d+f))\\ &= (ac+ae-bd-bf,bc+be+ad+af)\\ &= (ac-bd,bc+ad)+(ae-bf,be+af)\\ &= (a,b)(c,d)+(a,b)(e,f)\end{aligned}$$

ja seega kehtib esimene distributiivsuse seadus. Teise kehtimine järeldub jällegi korrutamise kommutatiivsusest. Sellega oleme näidanud, et hulk $\mathbb{R}^2$ on defineeritud tehete suhtes ring.

Olgu nüüd $(a,b) \in \mathbb{R}^2$ nullelemendist $(0,0)$ erinev element, s.t. $a \neq 0$ või $b \neq 0$, millest järeldub, et $a^2+b^2 \neq 0$. Kuna

$$(a,b) \cdot \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right) = \left(\frac{a^2+b^2}{a^2+b^2}, \frac{ba-ab}{a^2+b^2} \right) = (1,0),$$

siis

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right) \in \mathbb{R}^2.$$

Sellega oleme näidanud, et $\mathbb{R}^2$ on defineeritud tehete suhtes korpus. □

Lauses 2.1 konstrueeritud korpust tähistame sümboliga $\mathbb{C}$ ja nimetame **kompleksarvude korpuseks.** Selle korpuse elemente kutsutakse **kompleksarvudeks**.

Meie järgmiseks eesmärgiks on leida korpuses $\mathbb{C}$ midagi, mis oleks teatud mõttes sarnane reaalarvude korpusega.

Lause 2.2. *Hulga $\mathbb{R}^2$ alamhulk $\mathbb{R}' = \{(a,0) \mid a \in \mathbb{R}\}$ on korpus lauses 2.1 defineeritud tehete suhtes.*

Tõestus. Kuna

$$(a,0) + (b,0) = (a+b,0)$$

ja

$$(a,0)(b,0) = (ab-0, 0+0) = (ab,0)$$

mistahes $a,b \in \mathbb{R}$ korral, siis defineeritud liitmine ja korrutamine on algebralised tehted hulgal $\mathbb{R}'$. Need tehted on ilmselt ka assotsiatiivsed, kommutatiivsed ja seotud distributiivsuse seadustega. Elemendid $(0,0)$ ja $(1,0)$ kuuluvad hulka $\mathbb{R}'$, seega on hulgas $\mathbb{R}'$ olemas ka nullelement ja ühikelement. Kui $(a,0) \in \mathbb{R}'$, siis on sellel elemendil olemas vastandelement $(-a,0) \in \mathbb{R}'$. Kui $(a,0) \in \mathbb{R}'$ ei ole nullelement (s.t. $a \neq 0$), siis on sellel elemendil hulgas $\mathbb{R}'$ olemas pöördelement $\left(\frac{1}{a},0\right)$. Kokkuvõttes oleme veendunud, et $\mathbb{R}'$ on korpus. □

Kui kahe sama tüüpi algebralise struktuuri vahel leidub üksühene vastavus (bijektsioon), mis säilitab tehteid, siis neid struktuure nimetatakse isomorfseteks. Ringide korral võtab see definitsioon järgmise kuju.

Definitsioon 2.3. Ringe R ja R' nimetatakse **isomorfseteks**, kui leidub bijektiivne kujutus $f : R \to R'$ nii, et

RH1. $f(a+b) = f(a) + f(b)$ mistahes $a,b \in R$ korral (s.t. f säilitab liitmist);

RH2. $f(ab) = f(a)f(b)$ mistahes $a,b \in R$ korral (s.t. f säilitab korrutamist);

RH3. $f(1) = 1$ (s.t. f säilitab ühikelementi).

Sellist kujutust f nimetatakse **isomorfismiks** ringist R ringi R'. Kahte korpust nimetatakse **isomorfseteks**, kui nad on isomorfsed ringidena.

Märkus 2.4. Kujutust f, mis rahuldab tingimusi RH1–RH3 nimetatakse **ringide homomorfismiks**. Seega ringid on isomorfsed, kui nende vahel leidub bijektiivne homomorfism. Saab näidata, et kui f on bijektiivne ja rahuldab tingimust RH2, siis $f(1) = 1$, seega tingimuse RH3 võiks definitsioonist 2.3 ka välja jätta.

Lause 2.5. *Reaalarvude korpus $\mathbb{R}$ on isomorfne korpusega $\mathbb{R}'$ lausest 2.2.*

TÕESTUS. Kujutus

$$f : \mathbb{R} \to \mathbb{R}', \quad a \mapsto (a, 0)$$

on ilmselt bijektiivne. Kuna mistahes $a, b \in \mathbb{R}$ korral

$$\begin{aligned} f(a+b) &= (a+b, 0) = (a, 0) + (b, 0) = f(a) + f(b), \\ f(ab) &= (ab, 0) = (a \cdot b - 0 \cdot 0, 0 \cdot b + a \cdot 0) = (a, 0)(b, 0) = f(a)f(b), \\ f(1) &= (1, 0), \end{aligned}$$

siis f on isomorfism. □

Arvestades isomorfismi korpuste $\mathbb{R}$ ja $\mathbb{R}'$ vahel samastatakse enamasti reaalarv a korpuse $\mathbb{C}$ elemendiga $(a, 0)$. Seda samastamist arvestades võime korpuse $\mathbb{C}$ mistahes elemendi (a, b) kirjutada üles kujul

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + bi,$$

kus oleme tähistanud $i := (0, 1)$. Paneme tähele, et elemendi $i \in \mathbb{C}$ ruut on -1:

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

Seega võrrandil $x^2 = -1$ on korpuses $\mathbb{C}$ olemas lahend. Kompleksarvu i nimetatakse **imaginaarühikuks**.

Harilikult esitataksegi kompleksarve kujul

$$a + bi,$$

kus $a, b \in \mathbb{R}$. Sellist kuju nimetatakse kompleksarvu **algebraliseks kujuks**.

Kui b on negatiivne, siis harilikult kirjutatakse $a+bi$ asemel $a-|b|i$. Näiteks $2+(-3)i$ asemel kirjutatakse harilikult $2 - 3i$.

Definitsioon 2.6. Kui $z = a + bi$ on algebralisel kujul esitatud kompleksarv, siis

- reaalarvu a nimetatakse z **reaalosaks** (tähistatakse $a = \operatorname{Re} z$);
- kompleksarvu bi nimetatakse z **imaginaarosaks**;
- reaalarvu b nimetatakse z **imaginaarosa kordajaks** (tähistatakse $b = \operatorname{Im} z$).

Definitsioon 2.7. Kompleksarvu, mis ei ole reaalarv, nimetatakse **imaginaararvuks**. Imaginaararvu, mille reaalosa on 0, nimetatakse **puhtimaginaararvuks**.

Seega näiteks $5 + 2i$ on imaginaararv ja $-3i$ on puhtimaginaararv.

Kaks algebralisel kujul esitatud kompleksarvu $a + bi$ ja $c + di$ on võrdsed, kui järjestatud paarid (a, b) ja (c, d) on võrdsed. Seega

$$a + bi = c + di \iff a = c \text{ ja } b = d.$$

Teiste sõnadega: kaks kompleksarvu on võrdsed parajasti siis, kui nende kompleksarvude reaalosad on võrdsed ja imaginaarosa kordajad on võrdsed.

Algebralisel kujul näevad tehete definitsioonid välja järgmised:

$$\begin{aligned}(a+bi)+(c+di) &= (a+c)+(b+d)i,\\ (a+bi)(c+di) &= (ac-bd)+(bc+ad)i.\end{aligned}$$

Samuti kehtivad eelpoolöeldut arvestades võrdused

$$\begin{aligned}-(a+bi) &= -a-bi,\\ (a+bi)^{-1} &= \frac{a}{a^2+b^2}+\frac{-b}{a^2+b^2}i=\frac{1}{a^2+b^2}\cdot(a-bi).\end{aligned}$$

Ka kompleksarvude korpuses võib rääkida nullist erineva arvuga jagamisest pidades selle all silmas pöördelemendiga korrutamist (vt. märkust 1.36). Seega

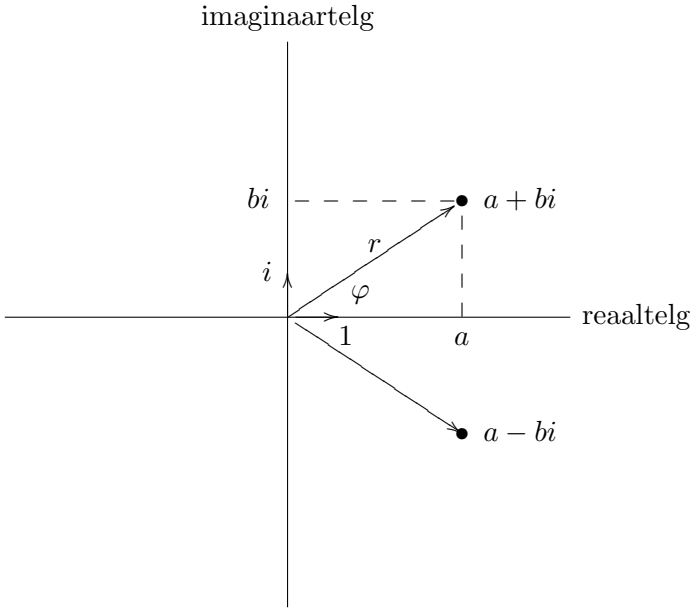
$$\frac{c+di}{a+bi}=(c+di)\cdot(a+bi)^{-1}=\frac{(c+di)(a-bi)}{a^2+b^2}.$$

Selle eeskirja võib sõnastada järgmiselt: *selleks, et jagada kompleksarv $c+di$ nullist erineva kompleksarvuga $a+bi$ tuleb arv $c+di$ korrutada kompleksarvuga $a-bi$ ja tulemus reaalarvuga* $\frac{1}{a^2+b^2}$.

2.2. Kompleksarvude geomeetriline tõlgendus

Olgu antud tasand koos ristkoordinaadistikuga. Siis tekib üksühene vastavus selle tasandi punktide ja reaalarvupaaride (a,b) (nende punktide koordinaatide) vahel. Seega on ka kompleksarvud üksüheses vastavuses selle tasandi punktidega: kompleksarvule $a+bi$ vastab punkt koordinaatidega (a,b). Tasandit koos kirjeldatud vastavusega nimetatakse **komplekstasandiks**. Koordinaadistiku x-telge nimetatakse **reaalteljeks** ja y-telge **imaginaarteljeks**. Reaalteljel asuvad parajasti reaalarvudele vastavad punktid ning imaginaarteljel puhtimaginaararvudele ja 0-le vastavad punktid. Koordinaatide alguspunkt vastab kompleksarvule 0.

Võib vaadelda ka selle tasandi vabavektorite hulka. Iga punkti koordinaadid on selle punkti kohavektori koordinaadid. Teades, et vektorite liitmisel tuleb liita nende vastavad koordinaadid, võime öelda ka, et kompleksarvude liitmisele vastab nendele arvudele vastavate punktide kohavektorite liitmine.



Definitsioon 2.8. Kompleksarvu $z = a+bi$ **kaaskompleksarvuks** nimetatakse kompleksarvu $\overline{z} = a - bi$.

Seega antud kompleksarvule ja tema kaaskompleksarvule vastavad punktid komplekstasandil asetsevad sümmeetriliselt reaaltelje suhtes.

Vaatleme mõningaid kaaskompleksarvude omadusi.

Lause 2.9. *Mistahes kompleksarvude z, w korral*

1. $\overline{\overline{z}} = z$,
2. $\overline{z} = z$ *parajasti siis, kui* z *on reaalarv,*
3. $\overline{z+w} = \overline{z} + \overline{w}$,
4. $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.

TÕESTUS. Olgu $z = a+bi$ ja $w = c+di$. Väited 1 ja 2 järelduvad vahetult definitsioonist. Lisaks sellele

$$\begin{aligned} \overline{z+w} &= \overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = (a-bi)+(c-di) = \overline{z}+\overline{w}, \\ \overline{z \cdot w} &= \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i = (a-bi)\cdot(c-di) = \overline{z}\cdot\overline{w}. \end{aligned}$$

□

Olgu nüüd komplekstasandil antud veel polaarkoordinaadistik, mis koosneb ühest fikseeritud punktist (poolusest) ja sellest punktist algavast kiirest (polaarteljest). Olgu poolus ja polaartelg valitud nii, et pooluseks on reaal- ja imaginaartelje lõikepunkt ja et polaartelg langeb kokku reaaltelje positiivse suunaga. Igale kompleksarvule $z = a + bi$ vastab komplekstasandil punkt, mille ristkoordinaadid on (a, b). Sellise punkti polaarkoordinaatideks on

1) polaarraadius ehk punkti kaugus r poolusest,

2) polaarnurk ehk nurk φ polaartelje ja vaadeldava punkti kohavektori vahel mõõdetuna kellaosuti liikumise vastassuunas (ehk vastupäeva).

Reaalarvu r nimetatakse kompleksarvu z **mooduliks** ja nurka φ tema **argumendiks**. Kasutatakse tähistusi

$$r = |z| \quad \text{ja} \quad \varphi = \arg(z).$$

Kompleksarvu 0 moodul on 0, aga argument ei ole määratud.

Pythagorase teoreemi põhjal (vt. joonist selle paragrahvi alguses) on selge, et

$$\boxed{r = \sqrt{a^2 + b^2}.}$$

Arvule z vastava komplekstasandi punkti ristkoordinaatide ja polaarkoordinaatide vahel kehtivad järgmised seosed:

$$\begin{aligned} a &= r\cos\varphi, \\ b &= r\sin\varphi. \end{aligned}$$

Siit näeme, et kui $a \neq 0$ (s.t. kui kompleksarvule vastav punkt ei asu imaginaarteljel), siis

$$\boxed{\tan\varphi = \frac{b}{a}.}$$

Kui $a = 0$ ja $b \neq 0$, siis $\varphi = \frac{\pi}{2}$ või $\varphi = \frac{3\pi}{2}$ sõltuvalt sellest, kas $b > 0$ või $b < 0$.

Samuti võime öelda, et $z = a + bi = r\cos\varphi + i \cdot r\sin\varphi = r(\cos\varphi + i\sin\varphi)$. Kompleksarvu z esitust kujul

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

nimetatakse selle kompleksarvu **trigonomeetriliseks kujuks**.

Arvestades siinus- ja koosinusfunktsiooni perioodilisust võib öelda, et

$$r(\cos\varphi + i\sin\varphi) = r(\cos(\varphi + 2k\pi) + i\sin(\varphi + 2k\pi))$$

iga $k \in \mathbb{Z}$ korral. Seda silmas pidades on mugav lugeda arvu z trigonomeetriliseks kujuks ka need avaldised, kus φ asemel on $\varphi + 2k\pi$, kus $k \in \mathbb{Z}$. Siis võime öelda, et

$$r_1(\cos\varphi_1 + i\sin\varphi_1) = r_2(\cos\varphi_2 + i\sin\varphi_2) \iff r_1 = r_2 \wedge (\exists k \in \mathbb{Z})(\varphi_1 = \varphi_2 + 2k\pi). \quad (4)$$

Sõnades: *trigonomeetrilisel kujul antud kompleksarvud on võrdsed parajasti siis, kui nende moodulid on võrdsed ja argumendid erinevad täispöörde täisarvkordse võrra.*

Näide 2.10. Kompleksarvu $z = \frac{3\sqrt{3}}{2} + \frac{3}{2}i$ võib trigonomeetrilisel kujul esitada näiteks järgnevalt:

$$z = 3\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right) = 3\left(\cos\frac{13\pi}{6} + i\sin\frac{13\pi}{6}\right) = 3\left(\cos\left(-\frac{11\pi}{6}\right) + i\sin\left(-\frac{11\pi}{6}\right)\right).$$

Uurime, kuidas korrutada trigonomeetrilisel kujul olevaid kompleksarve. Olgu meil arvud $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ ja $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$. Siis kasutades summa siinuse ja koosinuse valemeid saame, et

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

Seega *kompleksarvude korrutise moodul on tegurite moodulite korrutis ja korrutise argument on tegurite argumentide summa*:

$$\boxed{z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)).} \quad (5)$$

Kui $z = r(\cos\varphi + i\sin\varphi) \neq 0$, siis

$$z \cdot \left(\frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi))\right) = \frac{r}{r}(\cos(\varphi - \varphi) + i\sin(\varphi - \varphi)) = \cos 0 + i\sin 0 = 1,$$

mis tähendab, et

$$\boxed{z^{-1} = \frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)).}$$

Järelikult, kasutades eelnevaid tähistusi võime öelda, et $z_2 \neq 0$ korral

$$\boxed{\frac{z_1}{z_2} = z_1 z_2^{-1} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)).}$$

Niisiis: *kompleksarvude z_1 ja z_2 jagatise moodul on nende arvude moodulite jagatis ja argument on nende arvude argumentide vahe.*

Kui $z = r(\cos\varphi + i\sin\varphi)$ ja $n \in \mathbb{N}$, siis valemit (5) korduvalt rakendades saame järgmise valemi kompleksarvu astendamiseks:

$$\boxed{z^n = r^n(\cos n\varphi + i\sin n\varphi).}$$

Seda valemit kutsutakse **Moivre'i**[2] **valemiks**.

[2] Abraham de Moivre (1667–1754) — prantsuse matemaatik.

2.3. Kompleksarvude juurimine

Kui kompleksarvu astendamisega on tänu Moivre'i valemile asi lihtne, siis juurimise puhul on olukord tunduvalt keerulisem.

Definitsioon 2.11. Olgu n naturaalarv. Kompleksarvu w nimetatakse **n-nda astme juureks** kompleksarvust z, kui $w^n = z$.

Kuna nullist erinevate kompleksarvude korrutis ei saa olla null, siis ainsaks n-nda astme juureks kompleksarvust 0 on 0 ise. Edasises tegeleme nullist erinevate kompleksarvude juurte uurimisega. Osutub, et nullist erineval kompleksarvul z võib olla mitu n-nda astme juurt. Tähistame kõigi nende juurte hulka sümboliga $\sqrt[n]{z}$. Niisiis

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

Kui r on kompleksarvu $z \neq 0$ moodul, siis sümboliga $\sqrt[n]{r}$ tähistame sellist positiivset reaalarvu, mille n-s aste on r. Hulga $\sqrt[n]{z}$ elemendid saab leida järgmise teoreemi abil.

Teoreem 2.12. *Kui $n \in \mathbb{N}$ ja $z = r(\cos\varphi + i\sin\varphi) \in \mathbb{C} \setminus \{0\}$, siis n-nda astme juuri arvust z on täpselt n tükki ja nende hulk on*

$$\boxed{\sqrt[n]{z} = \left\{ \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right) \,\middle|\, k \in \{0,1,\ldots,n-1\} \right\}.} \tag{6}$$

Tõestus. Kui võtame suvalise arvu tõestatava võrduse paremal poolel olevast hulgast ja kasutades Moivre'i valemit tõstame selle astmesse n, siis saame arvu z. Seega paremal poolel olev hulk sisaldub hulgas $\sqrt[n]{z}$.

Näitame, et kehtib vastupidine sisalduvus. Olgu $w = s(\cos\psi + i\sin\psi) \in \sqrt[n]{z}$. Siis $w^n = z$. Moivre'i valemi põhjal

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

Kriteeriumi (4) tõttu $s^n = r$ ja leidub selline $l \in \mathbb{Z}$, et $n\psi = \varphi + 2l\pi$. Järelikult $s = \sqrt[n]{r}$ ja $\psi = \frac{\varphi+2l\pi}{n}$. Jagades arvu l jäägiga arvuga n saame leida sellised $q, k \in \mathbb{Z}$, et $l = qn + k$ ja $0 \leqslant k < n$. Seega

$$\psi = \frac{\varphi+2l\pi}{n} = \frac{\varphi+2(qn+k)\pi}{n} = \frac{\varphi+2k\pi}{n} + 2q\pi$$

ja

$$\begin{aligned} w &= s(\cos\psi + i\sin\psi) = \sqrt[n]{r}\left(\cos\frac{\varphi+2l\pi}{n} + i\sin\frac{\varphi+2l\pi}{n}\right) \\ &= \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right), \end{aligned}$$

kus $k \in \{0,1,\ldots,n-1\}$ ja viimase võrduse juures kasutasime siinuse ja koosinuse perioodilisust. Sellega oleme tõestanud nõutud hulkade võrduse.

Näitame lõpuks, et n-nda astme juuri on täpselt n tükki. Neid ei saa olla rohkem kui n, sest võrduse (6) paremal poolel olevas hulgas on ülimalt n elementi. Veendume, et neid on vähemalt n. Selleks näitame, et argumentide

$$\frac{\varphi}{n}, \frac{\varphi+2\pi}{n}, \frac{\varphi+4\pi}{n}, \ldots, \frac{\varphi+2(n-1)\pi}{n}$$

hulgas ei ole selliseid, mis erineks täispöörde täisarvkordse võrra (siis kriteeriumi (4) tõttu peavad vastavad kompleksarvud olema paarikaupa erinevad). Oletame vastuväiteliselt, et $0 \leqslant k < l \leqslant n-1$ ja

$$2\pi u = \frac{\varphi + 2l\pi}{n} - \frac{\varphi + 2k\pi}{n} = \frac{(l-k)\cdot 2\pi}{n}$$

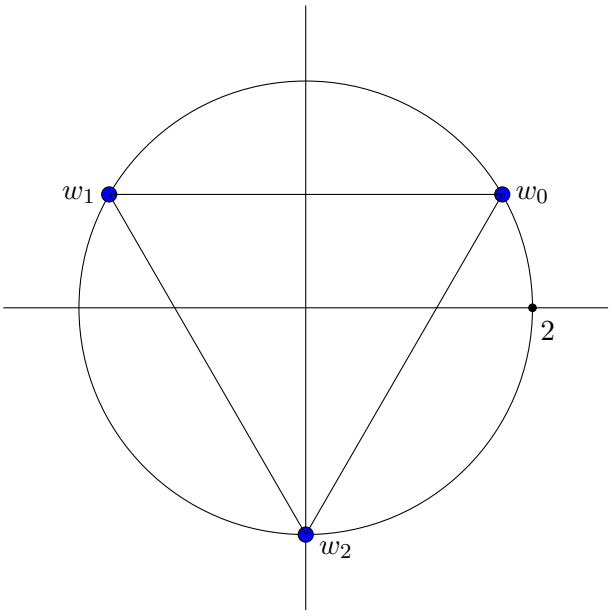
mingi $u \in \mathbb{Z}$ korral. Järelikult $nu = l - k > 0$, millest saame, et $u \neq 0$. Järelikult $u \geqslant 1$ ja $l - k \geqslant n$, mis on vastuolus eeldusega. □

Valemi (6) põhjal võib geomeetriliselt öelda, et *n-nda astme juured kompleksarvust* $z = r(\cos\varphi + i\sin\varphi) \in \mathbb{C}\setminus\{0\}$ *asuvad komplekstasandil sellise korrapärase n-nurga tippudes, mille ümberringjoone raadius on* $\sqrt[n]{r}$.

Näide 2.13. Olgu $z = 8i = 8(\cos 90^\circ + i\sin 90^\circ)$. Siis $\sqrt[3]{z} = \{w_0, w_1, w_2\}$, kus

$$\begin{aligned} w_0 &= 2(\cos 30^\circ + i\sin 30^\circ), \\ w_1 &= 2(\cos 150^\circ + i\sin 150^\circ), \\ w_2 &= 2(\cos 270^\circ + i\sin 270^\circ). \end{aligned}$$

Seega kuupjuured w_0, w_1, w_2 asuvad korrapärase kolmnurga tippudes, mille ümberringjoone raadius on 2.



Definitsioon 2.14. **n-nda astme ühejuur** on n-nda astme juur kompleksarvust 1.

Kuna $1 = \cos 0 + i\sin 0$, siis valemi (6) põhjal

$$\boxed{\sqrt[n]{1} = \left\{ \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \;\middle|\; k \in \{0, 1, \ldots, n-1\} \right\}.}$$

Tähistame iga $l \in \mathbb{Z}$ korral

$$\varepsilon_l := \cos\frac{2l\pi}{n} + i\sin\frac{2l\pi}{n}$$

ja toome sisse ka tähistuse

$$H_n := \{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}\} = \sqrt[n]{1}.$$

Kasutades samasugust mõttekäiku nagu oli teoreemi 2.12 tõestuses võime öelda, et

$$H_n = \{\varepsilon_l \mid l \in \mathbb{Z}\}. \tag{7}$$

Paneme veel tähele, et Moivre'i valemist järeldub, et iga $k \in \{0, 1, \ldots, n-1\}$ korral

$$\varepsilon_k = \varepsilon_1^k,$$

s.t. kõik n-nda astme ühejuured avalduvad ε_1 astmetena. Samuti võib öelda, et mistahes $k, l \in \mathbb{Z}$ korral

$$\begin{aligned} \varepsilon_k = \varepsilon_l &\iff (\exists u \in \mathbb{Z})\left(2\pi \cdot u = \frac{2k\pi}{n} - \frac{2l\pi}{n} = \frac{(k-l)\cdot 2\pi}{n}\right) \\ &\iff (\exists u \in \mathbb{Z})(nu = k-l) \iff n \mid k-l \iff k \equiv l \pmod{n}. \end{aligned}$$

Teoreem 2.15. *n-nda astme ühejuurte hulk H_n on rühm kompleksarvude korrutamise suhtes.*

Tõestus. Kui $\varepsilon_k, \varepsilon_l \in H_n$, siis ka

$$\varepsilon_k \cdot \varepsilon_l = \cos\frac{2(k+l)\pi}{n} + i\sin\frac{2(k+l)\pi}{n} = \varepsilon_{k+l} \in H_n$$

tänu võrdusele (7), seega korrutamine on algebraline tehe hulgal H_n. Kuna kõigi kompleksarvude korrutamine on assotsiatiivne, siis on seda ka ühejuurte korrutamine. Samuti on ühejuurte korrutamine kommutatiivne. Ühikelemendiks on kompleksarv $1 = \varepsilon_0 \in H_n$. Et mistahes $k \in \mathbb{Z}$ korral

$$\varepsilon_k \cdot \varepsilon_{n-k} = \varepsilon_{k+n-k} = \varepsilon_n = \cos 2\pi + i\sin 2\pi = 1,$$

siis $\varepsilon_k^{-1} = \varepsilon_{n-k}$. Seega $(H_n, \cdot)$ on rühm. □

Tuleb välja, et ühejuurte rühmad on väga sarnased jäägiklassirühmadega.

Definitsioon 2.16. Rühmad $(G, *)$ ja $(H, \circ)$ on **isomorfsed**, kui leidub bijektiivne kujutus $f : G \to H$ nii, et iga $a, b \in G$ korral

$$f(a * b) = f(a) \circ f(b).$$

Lause 2.17. *Rühmad $(H_n, \cdot)$ ja $(\mathbb{Z}_n, +)$ on isomorfsed.*

Tõestus. Defineerime kujutuse $f : \{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}\} \longrightarrow \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ võrdusega

$$f(\varepsilon_k) := \overline{k}.$$

On selge, et see kujutus on sürjektiivne. Kui $\overline{k} = \overline{l}$, siis $k \equiv l \pmod{n}$ ja $\varepsilon_k = \varepsilon_l$, mis näitab, et f on injektiivne. Olgu $k, l \in \mathbb{Z}$ ja olgu r jääk, mis tekib arvu $k + l$ jagamisel arvuga n. Siis $k + l \equiv r \pmod{n}$, $\varepsilon_{k+l} = \varepsilon_r$ ja

$$f(\varepsilon_k \cdot \varepsilon_l) = f(\varepsilon_{k+l}) = f(\varepsilon_r) = \overline{r} = \overline{k+l} = \overline{k} + \overline{l} = f(\varepsilon_k) + f(\varepsilon_l).$$

Seega kujutusel f on definitsioonis 2.16 nõutud omadused. □

3. Maatriksid

3.1. Maatriksi mõiste

Definitsioon 3.1. Olgu m ja n naturaalarvud. $(m \times n)$**-maatriks üle korpuse** K on m reast ja n veerust koosnev tabel, mille iga rea ja iga veeru lõikekohal on mingi korpuse K element ja mis on ümbritsetud ümarsulgudega. Neid K elemente nimetatakse **maatriksi elementideks**. Kõigi $(m \times n)$-maatriksite hulka üle korpuse K tähistatakse sümboliga $\mathrm{Mat}_{m,n}(K)$.

Näide 3.2. Näiteks

$$\begin{pmatrix} 2 & 3 & 7 & 0 \\ 4 & -1 & -7 & 5 \end{pmatrix}, \quad \begin{pmatrix} 6 & -4 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

on vastavalt (2×4)-, (2×2)- ja (3×1)-maatriksid üle reaalarvude korpuse $\mathbb{R}$ (või ka üle ratsionaalarvude korpuse $\mathbb{Q}$). Neist esimese maatriksi 1. rea ja 3. veeru element (ehk element kohal $(1, 3)$) on 7.

Maatrikseid tähistatakse harilikult suurte ladina tähtede $A, B, C, \ldots$ abil. Rääkides maatriksist üldiselt tähistatakse tema elemente harilikult väikese ladina tähe abil, millel on kaks indeksit. Neist esimene näitab, millises reas vaadeldav element asub ja teine näitab, millises veerus see element on. Näiteks $(m \times n)$-maatriks A, mille elemendid on a_{ij}, $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, esitatakse harilikult kujul

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

või

$$A = (a_{ij})_{i=1,\ldots,m, j=1,\ldots,n}.$$

Kui kontekstist on selge, millised on A mõõtmed (s.o. ridade arv ja veergude arv), siis kirjutatakse lühidalt ka

$$A = (a_{ij}).$$

Märkus 3.3. Põhimõtteliselt on maatriksit $A \in \mathrm{Mat}_{m,n}(K)$ võimalik vaadelda kui kujutust

$$A : \{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow K.$$

Sellise lähenemise korral on a_{ij} järjestatud paari (i, j) kujutis, s.t. $a_{ij} = A(i, j)$. Siiski praktika on näidanud, et maatriksite käsitlemine tabelitena on palju mugavam ja otstarbekam.

Märgime veel, et tihti kasutatakse maatriksite puhul ümarsulgude asemel nurksulge, kirjutades näiteks

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

Definitsioon 3.4. Kaks maatriksit on **võrdsed**, kui nende ridade arvud on võrdsed, veergude arvud on võrdsed ja vastavatel kohtadel olevad elemendid on võrdsed.

Seega maatriksid $A = (a_{ij})$ ja $B = (b_{ij})$ hulgast $\mathrm{Mat}_{m,n}(K)$ on võrdsed parajasti siis, kui $a_{ij} = b_{ij}$ iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$ korral.

Definitsioon 3.5. Ruutmaatriks on maatriks, mille ridade arv on võrdne veergude arvuga. Kui maatriksis on n rida ja n veergu, siis öeldakse, et see on **n-ndat järku ruutmaatriks.**

Kõigi n-ndat järku ruutmaatriksite hulka üle korpuse K tähistatakse $\mathrm{Mat}_n(K)$. Näiteks $\begin{pmatrix} 5 & 1 \\ -1 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$.

Definitsioon 3.6. Kui $A = (a_{ij}) \in \mathrm{Mat}_n(K)$, siis öeldakse, et elemendid $a_{11}, a_{22}, \ldots, a_{nn}$ moodustavad maatriksi A **peadiagonaali**.

Eelmise näitemaatriksi peadigonaal koosneb seega arvudest 5 ja 3.

Iga maatriksiga saab loomulikul viisil siduda veel kaks maatriksit: transponeeritud maatriksi ja vastandmaatriksi.

Transponeerimine tähendab maatriksi ridade ja veergude ümbervahetamist.

Definitsioon 3.7. Maatriksi A **transponeeritud maatriks** on maatriks, mille esimeseks reaks on maatriksi A esimene veerg, teiseks reaks maatriksi A teine veerg jne. Tähistus: A^T või A^t.

Definitsioonist on selge, et kui $A \in \mathrm{Mat}_{m,n}(K)$, siis $A^T \in \mathrm{Mat}_{n,m}(K)$. Samuti on ilmne, et

$$(A^T)^T = A.$$

Näide 3.8. Näiteks

$$\begin{pmatrix} 2 & 3 & 7 & 0 \\ 4 & -1 & -7 & 5 \end{pmatrix}^T = \begin{pmatrix} 2 & 4 \\ 3 & -1 \\ 7 & -7 \\ 0 & 5 \end{pmatrix}.$$

Definitsioon 3.9. Maatriksi $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ **vastandmaatriksiks** $-A$ nimetatakse maatriksit, mille elementideks on maatriksi A vastavate elementide vastandelemendid, s.t. maatriksit

$$-A = \begin{pmatrix} -a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & -a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{m1} & -a_{m2} & \ldots & -a_{mn} \end{pmatrix}.$$

Definitsioonist on selge, et kui $A \in \mathrm{Mat}_{m,n}(K)$, siis $-A \in \mathrm{Mat}_{m,n}(K)$ ja $-(-A) = A$.

Näide 3.10. Näiteks kui

$$A = \begin{pmatrix} \overline{2} & \overline{3} & \overline{7} & \overline{0} \\ \overline{4} & \overline{1} & \overline{8} & \overline{5} \end{pmatrix} \in \mathrm{Mat}_{2,4}(\mathbb{Z}_{11}),$$

siis

$$-A = \begin{pmatrix} \overline{9} & \overline{8} & \overline{4} & \overline{0} \\ \overline{7} & \overline{10} & \overline{3} & \overline{6} \end{pmatrix} \in \mathrm{Mat}_{2,4}(\mathbb{Z}_{11}).$$

Nende mõistete abil defineeritakse sümmeetrilised ja kaldsümmeetrilised maatriksid.

Definitsioon 3.11. Ruutmaatriks A on **sümmeetriline**, kui $A^T = A$. Ruutmaatriks A on **kaldsümmeetriline**, kui $A^T = -A$.

Näide 3.12. Maatriks

$$A = \begin{pmatrix} 2 & 3 & 7 \\ 3 & -1 & -7 \\ 7 & -7 & 4 \end{pmatrix}$$

on sümmeetriline ja maatriks

$$B = \begin{pmatrix} 0 & -3 & 1 \\ 3 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix}$$

on kaldsümmeetriline.

Niisiis ruutmaatriks on sümmeetriline, kui tema peadiagonaali (kui mõttelise joone) suhtes sümmeetriliselt asuvad elemendid on võrdsed.

3.2. Ringi elementide summeerimisest

Käesolevas paragrahvis tähistagu R alati ringi.

Toome sisse ühe tähistuse, mida matemaatikas kasutatakse palju ja mis aitab meil mitmetes kohtades materjali lihtsamalt esitada. Nimelt summat $s_1 + s_2 + s_3 + \ldots + s_n$, kus $n \in \mathbb{N}$ ja $s_1, s_2, \ldots, s_n \in R$, tähistatakse lühidalt

$$\sum_{i=1}^{n} s_i. \tag{8}$$

Siin $\sum$ (kreeka suurtäht sigma) on summeerimismärk ja i on summeerimisindeks. Arvud 1 ja n näitavad ära, millistes piirides summeerimisindeks muutub ja summeerimisindeks omandab kõik naturaalarvulised väärtused 1-st n-ni. Igale i väärtusele vastab üks liidetav vaadeldavas summas. Liitmine toimub ringis R. Avaldist (8) võiks lugeda järgmiselt: "summa, kus i muutub ühest n-ni, s_i-dest" või "s_i-de summa, kus i muutub ühest n-ni".

Näiteks kui $R = \mathbb{Z}$, $s_1 = 3$, $s_2 = 6$, $s_3 = -4$ ja $s_4 = 5$, siis

$$\sum_{i=1}^{4} s_i = 3 + 6 - 4 + 5 = 10.$$

Kui aga $R = \mathbb{Z}_7$, $s_1 = \overline{3}$, $s_2 = \overline{6}$, $s_3 = \overline{-4}$ ja $s_4 = \overline{5}$, siis

$$\sum_{i=1}^{4} s_i = \overline{3} + \overline{6} - \overline{4} + \overline{5} = \overline{10} = \overline{3}.$$

Liidetav s_i võib olla mingi avaldis, mis sõltub arvust i. Näiteks

$$\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + 3^2 + \ldots + (n-1)^2 + n^2.$$

Võib vaadelda ka summasid, kus liidetavad sõltuvad kahest indeksist. Selliseid liidetavaid võib summeerida kas ühe, teise või mõlema indeksi järgi. Näiteks võib vaadelda summat

$$\sum_{i=1}^{m} \sum_{j=1}^{n} s_{ij},$$

kus liidetavaid s_{ij} summeritakse enne j ja siis i järgi. Sellise summa erijuhuks on näiteks

$$\sum_{i=1}^{2}\sum_{j=1}^{3} j^i = \sum_{i=1}^{2}(1^i + 2^i + 3^i) = (1^1 + 2^1 + 3^1) + (1^2 + 2^2 + 3^2) = 6 + 14 = 20.$$

Summeerimismärki võib kasutada ka sellistel juhtudel, kui on vaja summeerida mingisse hulka kuuluvaid elemente. Kui näiteks $A \subseteq \mathbb{Z}$ on kõigi 10-st väiksemate algarvude hulk, siis

$$\sum_{a\in A}(a-5) = (2-5) + (3-5) + (5-5) + (7-5) = -3 - 2 + 0 + 2 = -3.$$

Lause 3.13. *Summeerimisel ringis R on järgmised omadused:*

SO1.

$$\boxed{\sum_{i=1}^{n} ts_i = t\sum_{i=1}^{n} s_i, \quad \sum_{i=1}^{n} s_i t = \left(\sum_{i=1}^{n} s_i\right) t}$$

(s.t. konstandi, mis ei sõltu summeerimisindeksist, võib tuua summa märgi alt välja);

SO2.

$$\boxed{\sum_{i=1}^{n}(s_i + t_i) = \sum_{i=1}^{n} s_i + \sum_{i=1}^{n} t_i;}$$

SO3.

$$\boxed{\sum_{i=1}^{m}\sum_{j=1}^{n} s_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} s_{ij}}$$

(s.t. kõrvutiolevad summa märgid võib ära vahetada).

Tõestus. SO1 järeldub distributiivsuse seadustest R7 ja R8.

SO2. Tänu liitmise assotsiatiivsusele ja kommutatiivsusele

$$\begin{aligned}
\sum_{i=1}^{n}(s_i + t_i) &= (s_1 + t_1) + (s_2 + t_2) + \ldots + (s_n + t_n) \\
&\underset{R1,R4}{=} (s_1 + s_2 + \ldots + s_n) + (t_1 + t_2 + \ldots + t_n) \\
&= \sum_{i=1}^{n} s_i + \sum_{i=1}^{n} t_i.
\end{aligned}$$

SO3. Paneme tähele, et

$$\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{n} s_{ij} &= \sum_{i=1}^{m}(s_{i1} + \ldots + s_{in}) \\
&= (s_{11} + \ldots + s_{1n}) + (s_{21} + \ldots + s_{2n}) + \ldots + (s_{m1} + \ldots + s_{mn}) \\
&\underset{R1,R4}{=} (s_{11} + s_{21} + \ldots + s_{m1}) + \ldots + (s_{1n} + s_{2n} + \ldots + s_{mn}) \\
&= \sum_{j=1}^{n}(s_{1j} + \ldots + s_{mj}) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{m} s_{ij}.
\end{aligned}$$

□

Märkus 3.14. Kuna liidetavaid s_{ij} võime vaadelda kui $(m\times n)$-maatriksi elemente, siis omadust SO3 võib tõlgendada nii, et maatriksi kõigi elementide summa ei sõltu sellest, kas me liidame neid järjest ridade kaupa või veergude kaupa.

Arvestades omadust SO3 kirjutatakse juhul, kui $m = n$, mõnikord lühemalt

$$\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij} = \sum_{i,j=1}^{n} s_{ij}.$$

Tänu liitmise kommutatiivsusele ringis kehtib järgmine tulemus.

Lause 3.15. *Kui $s_1, \ldots, s_n$ on ringi R elemendid ja $f : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ on bijektiivne kujutus, siis*

$$s_1 + s_2 + \ldots + s_n = s_{f(1)} + s_{f(2)} + \ldots + s_{f(n)}.$$

3.3. Maatriksite liitmine ja maatriksi korrutamine skalaariga

Enne tehete juurde asumist peatume veelkord sellel, kuidas on võimalik maatrikseid esitada. Me võime näiteks vaadelda maatriksit $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$, mille element a_{ij}, kus $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$, on antud mingi valemiga, mis võib sõltuda indeksitest i ja j. Näiteks maatriks $A = (a_{ij}) \in \mathrm{Mat}_{3,4}(\mathbb{R})$, kus

$$a_{ij} = \min(4, i + j),$$

näeb välja nii:

$$A = \begin{pmatrix} 2 & 3 & 4 & 4 \\ 3 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{pmatrix}.$$

Tihti kirjutatakse sama asja veel lühemalt: $A = (\min(4, i + j)) \in \mathrm{Mat}_{3,4}(\mathbb{R})$. See kirjapilt väljendab järgmist asjaolu: A on (3×4)-maatriks, mille i-ndas reas ja j-ndas veerus on reaalarv $\min(4, i + j)$.

Kui $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(\mathbb{R})$, siis $B = (a_{ij} + 3) \in \mathrm{Mat}_{m,n}(\mathbb{R})$ on $(m \times n)$-maatriks, mille i-ndas reas ja j-ndas veerus on arv $a_{ij} + 3$. Samasugust kokkulepet kasutades võime öelda, et

$$-A = (-a_{ij}) \in \mathrm{Mat}_{m,n}(K)$$

ja

$$A^T = (a_{ji}) \in \mathrm{Mat}_{n,m}(K).$$

Väga tihti läheb lineaaralgebras vaja järgmisi mingis mõttes hästi lihtsaid maatrikseid.

Definitsioon 3.16. Nullmaatriks on maatriks, mille kõik elemendid on võrdsed korpuse nullelemendiga.

Definitsioon 3.17. Ühikmaatriks on ruutmaatriks, mille peadiagonaali elemendid võrduvad korpuse ühikelemendiga ja kõik muud elemendid on võrdsed korpuse nullelemendiga.

$(m\times n)$-nullmaatriksit tähistame sümboliga $\Theta_{m,n}$ või lihtsalt Θ (kreeka suurtäht teeta). Tihti kasutatakse nullmaatriksi tähisena ka lihtsalt sümbolit 0.

n-ndat järku ühikmaatriksit tähistame sümboliga E_n või lihtsalt E. Kasutatakse ka tähiseid I_n ja I.

Näide 3.18. Näiteks

$$\Theta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{ja} \quad E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

on vastavalt (3×2)-nullmaatriks ja kolmandat järku ühikmaatriks üle korpuse $\mathbb{R}$.

Üldjuhul võib kirjutada ka

$$\Theta_{m,n} = (\theta_{ij}), \text{ kus } \theta_{ij} = 0 \text{ iga } i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\} \text{ korral}$$

ja

$$E_n = (\delta_{ij}), \text{ kus } \delta_{ij} = \begin{cases} 1, & \text{kui } i = j, \\ 0, & \text{kui } i \neq j, \end{cases} \text{ iga } i, j \in \{1, \ldots, n\} \text{ korral.}$$

Sümbolit δ_{ij} tuntakse matemaatikas kui **Kroneckeri**[3] **deltat**.

Defineerime nüüd maatriksite summa.

Definitsioon 3.19. Maatriksite $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja $B = (b_{ij}) \in \mathrm{Mat}_{m,n}(K)$ **summa** on maatriks $A + B = (c_{ij}) \in \mathrm{Mat}_{m,n}(K)$, kus $c_{ij} = a_{ij} + b_{ij}$ iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$ korral.

Kui kontekstist on selge, milliste mõõtmetega maatriksitega on tegu, võib maatriksite liitmise definitsiooni anda lühemal kujul:

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}).$$

Tabelite kujul näeb liitmisreegel välja nii:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{m1} & b_{m2} & \ldots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \ldots & a_{mn} + b_{mn} \end{pmatrix},$$

s.t. *maatriksite liitmisel liidetakse nende vastavatel kohtadel olevad elemendid.* Rõhutame veelkord, et liita saab vaid samade mõõtmetega maatrikseid.

Liitmise ja vastandmaatriksi abil saab defineerida maatriksite lahutamise. **Maatriksite** $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja $B = (b_{ij}) \in \mathrm{Mat}_{m,n}(K)$ **vahe** on maatriks

$$A - B := A + (-B).$$

Seega $A - B = (c_{ij}) \in \mathrm{Mat}_{m,n}(K)$, kus $c_{ij} = a_{ij} + (-b_{ij}) = a_{ij} - b_{ij}$ iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$ korral.

Definitsioon 3.20. Maatriksi $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ **ja skalaari** $k \in K$ **korrutis** on maatriks $kA = (c_{ij}) \in \mathrm{Mat}_{m,n}(K)$, kus $c_{ij} = ka_{ij}$ iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$ korral.

Seega

$$k(a_{ij}) = (ka_{ij})$$

[3]Leopold Kronecker (1823–1891) — saksa matemaatik

ehk

$$k\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix},$$

s.t. *maatriksi korrutamisel skalaariga k korrutatakse selle maatriksi kõik elemendid skalaariga k.* Muuhulgas

$$(-1)A = -A,$$

sest $(-1)a_{ij} = -(1a_{ij}) = -a_{ij}$ tänu lausele 1.27(2).

Näide 3.21. Näiteks

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -2 & -2 & 1 \\ -2 & -3 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 & 5 \\ 2 & 0 & 1 & 5 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -2 & -2 & 1 \\ -2 & -3 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 4 & 5 & 3 \\ 6 & 6 & 3 & -3 \end{pmatrix},$$

$$3 \cdot \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 0 & -3 \end{pmatrix}.$$

Defineeritud tehetel on terve rida häid omadusi.

Lause 3.22. *Mistahes $A, B, C \in \mathrm{Mat}_{m,n}(K)$ ja $k, l \in K$ korral*

1. $(A + B) + C = A + (B + C)$;
2. $A + \Theta_{m,n} = A = \Theta_{m,n} + A$;
3. $A + (-A) = \Theta_{m,n} = (-A) + A$;
4. $A + B = B + A$;
5. $k(A + B) = kA + kB$;
6. $(k + l)A = kA + lA$;
7. $(kl)A = k(lA)$;
8. $1A = A$.

Tõestus. 1. Olgu $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}) \in \mathrm{Mat}_{m,n}(K)$. Siis

$$\begin{aligned} (A + B) + C &= ((a_{ij}) + (b_{ij})) + (c_{ij}) \underset{\text{Def. 3.19}}{=} (a_{ij} + b_{ij}) + (c_{ij}) \underset{\text{Def. 3.19}}{=} ((a_{ij} + b_{ij}) + c_{ij}) \\ &\underset{\text{R1}}{=} (a_{ij} + (b_{ij} + c_{ij})) \underset{\text{Def. 3.19}}{=} (a_{ij}) + (b_{ij} + c_{ij}) \underset{\text{Def. 3.19}}{=} (a_{ij}) + ((b_{ij}) + (c_{ij})) \\ &= A + (B + C). \end{aligned}$$

2. Olgu $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja vaatleme $(m \times n)$-nullmaatriksit $\Theta_{m,n} = (\theta_{ij})$, kus $\theta_{ij} = 0$ iga i ja j korral. Siis

$$A + \Theta_{m,n} = (a_{ij}) + (\theta_{ij}) \underset{\text{Def. 3.19}}{=} (a_{ij} + \theta_{ij}) = (a_{ij} + 0) \underset{R2}{=} (a_{ij}) = A.$$

Teise võrduse saab tõestada anaoogiliselt.

5. Olgu $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja $k \in K$. Siis

$$\begin{aligned} k(A+B) &= k((a_{ij}) + (b_{ij})) \underset{\text{Def. 3.19}}{=} k(a_{ij} + b_{ij}) \underset{\text{Def. 3.20}}{=} (k(a_{ij} + b_{ij})) \\ &\underset{R7}{=} (ka_{ij} + kb_{ij}) \underset{\text{Def. 3.19}}{=} (ka_{ij}) + (kb_{ij}) \underset{\text{Def. 3.20}}{=} k(a_{ij}) + k(b_{ij}) = kA + kB. \end{aligned}$$

Nagu näeme, on omaduste 1, 2 ja 5 tõestamiseks vaja kasutada vaid maatriksite liitmise ja skalaariga korrutamise definitsiooni ning korpuse omadusi. Ka ülejäänud väidete tõestamine taandub tehete definitsioonide ja korpuse omaduste kasutamisele. Need tõestused jätame läbi mõtlemiseks lugejale. □

3.4. Maatriksite korrutamine

Maatriksite korrutise definitsioon on mõnevõrra keerulisem kui maatriksite summa definitsioon. Kahte maatriksit saab korrutada ainult siis, kui esimese maatriksi veergude arv on võrdne teise maatriksi ridade arvuga.

Definitsioon 3.23. Maatriksite $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja $B = (b_{ij}) \in \mathrm{Mat}_{n,p}(K)$ **korrutiseks** nimetatakse maatriksit $C = (c_{ij}) \in \mathrm{Mat}_{m,p}(K)$, kus iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, p\}$ korral

$$\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.}$$

Niisiis *selleks, et leida korrutise C element, mis asub i-ndas reas ja j-ndas veerus, tuleb maatriksi A i-nda rea elemendid korrutada maatriksi B j-nda veeru vastavate elementidega ja tulemused liita.*

Harilikult kirjutatakse korrutise C asemel AB.

Näide 3.24. Näiteks

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & -2 & 4 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ 1 & 2 \\ 5 & -1 \end{pmatrix} = \begin{pmatrix} 2+3+5 & -3+6-1 \\ 0-2+20 & 0-4-4 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 18 & -8 \end{pmatrix},$$

$$\begin{pmatrix} 2 & -3 \\ 1 & 2 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 1 \\ 0 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 12 & -10 \\ 1 & -1 & 9 \\ 5 & 17 & 1 \end{pmatrix}.$$

Definitsioonist on kohe selge, et leidub maatrikseid, mille korral korrutis AB on olemas, aga korrutist BA ei leidu. Isegi siis, kui AB ja BA leiduvad, ei pruugi nad võrdsed olla, nagu näha eelnenud näitest. Seega maatriksite korrutamine ei ole kommutatiivne. Siiski on maatriksite korrutamisel rida muid omadusi.

Lause 3.25. *Maatriksite korrutamisel on järgmised omadused.*

1. *Mistahes $A \in \mathrm{Mat}_{m,n}(K)$, $B \in \mathrm{Mat}_{n,p}(K)$ ja $C \in \mathrm{Mat}_{p,q}(K)$ korral*

$$(AB)C = A(BC).$$

2. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$ *korral*

$$E_m A = A \quad ja \quad AE_n = A,$$

kus $E_m \in \mathrm{Mat}_m(K)$ *ja* $E_n \in \mathrm{Mat}_n(K)$ *on vastavat järku ühikmaatriksid.*

3. *Mistahes* $A, B \in \mathrm{Mat}_{m,n}(K)$, $C \in \mathrm{Mat}_{n,p}(K)$ *korral*

$$(A+B)C = AC + BC.$$

4. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$, $B, C \in \mathrm{Mat}_{n,p}(K)$ *korral*

$$A(B+C) = AB + AC.$$

5. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$, $B \in \mathrm{Mat}_{n,p}(K)$ *ja* $k \in K$ *korral*

$$k(AB) = (kA)B = A(kB).$$

6. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$ *ja* $p, q \in \mathbb{N}$ *korral*

$$\Theta_{q,m} A = \Theta_{q,n} \quad ja \quad A\Theta_{n,p} = \Theta_{m,p}.$$

Tõestus. 1. Olgu $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$, $B = (b_{ij}) \in \mathrm{Mat}_{n,p}(K)$ ja $C = (c_{ij}) \in \mathrm{Mat}_{p,q}(K)$. Toome sisse tähised maatriksite $AB, (AB)C, BC$ ja $A(BC)$ elementide jaoks:

$$\begin{aligned} AB &= (u_{ij}) \in \mathrm{Mat}_{m,p}, \\ (AB)C &= (v_{ij}) \in \mathrm{Mat}_{m,q}, \\ BC &= (w_{ij}) \in \mathrm{Mat}_{n,q}, \\ A(BC) &= (t_{ij}) \in \mathrm{Mat}_{m,q}. \end{aligned}$$

Kasutades maatriksite korrutamise definitsiooni, summeerimise omadusi ja korpuse korrutamise assotsiatiivsust võime kirjutada:

$$u_{ik} \underset{\text{Def. 3.23}}{=} \sum_{l=1}^{n} a_{il} b_{lk},$$

$$v_{ij} \underset{\text{Def. 3.23}}{=} \sum_{k=1}^{p} u_{ik} c_{kj} = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} a_{il} b_{lk} \right) c_{kj} \underset{SO1}{=} \sum_{k=1}^{p} \sum_{l=1}^{n} (a_{il} b_{lk})\, c_{kj} \underset{R5}{=} \sum_{k=1}^{p} \sum_{l=1}^{n} a_{il} (b_{lk} c_{kj}),$$

$$w_{lj} \underset{\text{Def. 3.23}}{=} \sum_{k=1}^{p} b_{lk} c_{kj},$$

$$t_{ij} \underset{\text{Def. 3.23}}{=} \sum_{l=1}^{n} a_{il} w_{lj} = \sum_{l=1}^{n} a_{il} \left(\sum_{k=1}^{p} b_{lk} c_{kj} \right) \underset{SO1}{=} \sum_{l=1}^{n} \sum_{k=1}^{p} a_{il} \left(b_{lk} c_{kj} \right) \underset{SO3}{=} \sum_{k=1}^{p} \sum_{l=1}^{n} a_{il} \left(b_{lk} c_{kj} \right).$$

Kuna $v_{ij} = t_{ij}$ iga $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, q\}$ korral, siis on maatriksid $(AB)C$ ja $A(BC)$ võrdsed, $(AB)C = A(BC)$.

2. Olgu $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja olgu $E_m = (\delta_{ij})$ m-ndat järku ühikmaatriks. Siis korrutise $E_m A \in \mathrm{Mat}_{m,n}(K)$ i-ndas reas ja j-ndas veerus on element

$$\sum_{k=1}^{m} \delta_{ik} a_{kj} = \delta_{i1} a_{1j} + \delta_{i2} a_{2j} + \ldots + \delta_{im} a_{mj} = \delta_{ii} a_{ij} = 1 \cdot a_{ij} = a_{ij}.$$

Järelikult $E_m A = A$, sest nende maatriksite vastavatel kohtadel olevad elemendid on võrdsed. Võrduse $AE_n = A$ saab tõestada analoogiliselt.

3. Olgu $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}_{m,n}(K)$, $C = (c_{ij}) \in \mathrm{Mat}_{n,p}(K)$. Toome sisse tähised maatriksite $(A+B)C, AC$ ja BC elementide jaoks:

$$\begin{aligned} (A+B)C &= (u_{ij}) \in \mathrm{Mat}_{m,p}(K), \\ AC &= (v_{ij}) \in \mathrm{Mat}_{m,p}(K), \\ BC &= (w_{ij}) \in \mathrm{Mat}_{m,p}(K). \end{aligned}$$

Et maatriksis $A+B$ kohal (i,k) on element $a_{ik}+b_{ik}$, siis

$$u_{ij} \underset{\text{Def.3.23}}{=} \sum_{k=1}^{n}(a_{ik}+b_{ik})c_{kj} \underset{R7}{=} \sum_{k=1}^{n}(a_{ik}c_{kj}+b_{ik}c_{kj}) \underset{SO2}{=} \sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj} \underset{\text{Def.3.23}}{=} v_{ij} + w_{ij}$$

iga $i \in \{1,\ldots,m\}$ ja $j \in \{1,\ldots,p\}$ korral. Kuna maatriksites $(A+B)C$ ja $AC+BC$ on vastavatel kohtadel samad elemendid, siis on need maatriksid võrdsed.

Ülejäänud omadused saab tõestada analoogiliselt. □

Lause 3.22 ja lause 3.25 põhjal võime öelda, et kehtib järgmine tulemus.

Lause 3.26. *Hulk* $\mathrm{Mat}_n(K)$ *on ring maatriksite liitmise ja korrutamise suhtes.*

3.5. Transponeerimise omadused

Uurime nüüd, kuidas on transponeerimine seotud maatriksite liitmisega, maatriksi skalaariga korrutamisega ja maatriksite korrutamisega.

Lause 3.27. *Maatriksite transponeerimisel on järgmised omadused.*

1. *Mistahes* $A, B \in \mathrm{Mat}_{m,n}(K)$ *korral*

$$(A+B)^T = A^T + B^T.$$

2. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$ *ja* $k \in K$ *korral*

$$(kA)^T = kA^T.$$

3. *Mistahes* $A \in \mathrm{Mat}_{m,n}(K)$ *ja* $B \in \mathrm{Mat}_{n,p}(K)$ *korral*

$$(AB)^T = B^T A^T.$$

Tõestus. Tõestame neist omadustest viimase (ülejäänud jäävad jälle lugejale läbi mõtlemiseks). Olgu $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ ja $B = (b_{ij}) \in \mathrm{Mat}_{n,p}(K)$. Paneme tähele, et $B^T \in \mathrm{Mat}_{p,n}(K)$ ja $A^T \in \mathrm{Mat}_{n,m}(K)$, seega on korrutis $B^T A^T$ olemas ja $B^T A^T$ on $(p \times m)$-maatriks, nagu ka $(AB)^T$. Maatriksi $(AB)^T$ i-ndas reas ja j-ndas veerus on maatriksi AB j-nda rea ja i-nda veeru element, s.t. summa

$$\sum_{k=1}^{n} a_{jk}b_{ki}.$$

Maatriksi B^TA^T i-ndas reas ja j-ndas veerus on element

$$\sum_{k=1}^{n} u_{ik}v_{kj},$$

kus u_{ik} on B^T i-nda rea ja k-nda veeru element ja v_{kj} on A^T k-nda rea ja j-nda veeru element. Maatriksi transponeerimise definitsiooni kohaselt $u_{ik} = b_{ki}$ ja $v_{kj} = a_{jk}$. Seega

$$\sum_{k=1}^{n} u_{ik}v_{kj} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki},$$

kuna korrutamine korpuses K on kommutatiivne. Järelikult kehtib võrdus $(AB)^T = B^TA^T$. $\square$

4. Determinandid

4.1. Permutatsioonid ja substitutsioonid

Definitsioon 4.1. Olgu n naturaalarv ja olgu A n-elemendiline hulk. **Permutatsioon** hulga A elementidest on selline n-elemendiline järjend, milles hulga A iga element esineb täpselt ühe korra.

Näide 4.2. Tüüpiline mängukaartide pakk on 36-elemendiline hulk. Selle paki iga segamise tulemusel tekib permutatsioon nendest kaartidest.

Enamasti vaadeldakse matemaatikas permutatsioone hulga $A = \{1, 2, \ldots, n\}$ elementidest. Sellist permutatsiooni tähistame $(i_1, i_2, \ldots, i_n)$. Tihti kirjutatakse ka lihtsalt $i_1 i_2 \ldots i_n$ (näiteks raamatus [1]). Lihtne on aru saada, et permutatsioone n elemendist on $n!$ tükki.

Näiteks $(4, 1, 3, 5, 2)$ on permutatsioon 5-st elemendist, aga $(4, 1, 3, 4, 2)$ ja $(2, 5, 4, 3)$ ei ole.

Definitsioon 4.3. Permutatsiooni $(1, 2, \ldots, n)$ nimetatakse **loomulikuks permutatsiooniks** n elemendist.

Definitsioon 4.4. Öeldakse, et üks permutatsioon on saadud teisest **transpositsiooni** abil, kui see esimene permutatsioon on saadud teisest kahe elemendi äravahetamise teel.

Näide 4.5. Permutatsioon $(4, 1, 3, 5, 2)$ on saadud permutatsioonist $(4, 1, 2, 5, 3)$ kolmanda ja viienda elemendi äravahetamise teel.

Lause 4.6. *Kõik permutatsioonid n elemendist on võimalik järjestada niiviisi, et iga järgnev permutatsioon on eelnevast saadav transpositsiooni abil, kusjuures esimeseks võib valida suvalise permutatsiooni.*

Tõestus. Tõestame lause matemaatilise induktsiooniga elementide arvu n järgi. Kui $n = 1$, siis on väide ilmne. Kui $n = 2$ ja esimene permutatsioon on (i_1, i_2), siis teine permutatsioon (i_2, i_1) on esimesest saadav transpositsiooni abil. Rohkem permutatsioone 2-st elemendist pole. Oletame nüüd, et $n \geqslant 3$ ja lause väide kehtib permutatsioonide jaoks $n-1$ elemendist. Võtame suvalise permutatsiooni

$$(i_1, i_2, i_3, \ldots, i_n)$$

n elemendist. Vaatleme kõiki permutatsioone n elemendist, mis algavad elemendiga i_1. Kui neist esimene komponent i_1 ära jätta, siis saame kõik permutatsioonid $n-1$ elemendist $i_2, \ldots, i_n$. Induktsiooni eelduse põhjal võime need järjestada sellisel viisil, et iga järgnev on saadud eelnevast transpositsiooni abil. Olgu sellise järjestuse viimane permutatsioon

$$(i_1, j_2, j_3, \ldots, j_n).$$

Transpositsiooni abil, mis vahetab ära i_1 ja j_2 saame permutatsiooni

$$(j_2, i_1, j_3, \ldots, j_n).$$

Järjestame nüüd nõutaval viisil kõik permutatsioonid elementidest $i_1, j_3, \ldots, j_n$. Lisades neile ette j_2 saame vajaliku järjestuse kõigi permutatsioonide jaoks, mille esimene komponent on j_2. Olgu selles järjestuses viimane permutatsioon

$$(j_2, k_2, k_3, \ldots, k_n).$$

Leiame elementide $k_2, k_3, \ldots, k_n$ hulgast sellise elemendi k_s, mis ei kuulu hulka $\{i_1, j_2\}$. Vahetame ära j_2 ja k_s (s.t. teeme transpositsiooni) ja järjestame nõutaval viisil kõik permutatsioonid, mille esimene komponent on k_s. Nii jätkates saame nõutaval viisil ära järjestada kõik permutatsioonid, mille esimene komponent on $1, 2, \ldots, n$, s.t. kõikvõimalikud permutatsioonid elementidest $1, 2, \ldots, n$. □

Näide 4.7. Võttes esimeseks permutatsiooniks $(2, 1, 3)$ võime kõik 6 permutatsiooni kolmest elemendist lause tõestuses kasutatud meetodi abil järjestada nii:

$$(2,1,3),\ \ (2,3,1),\ \ (3,2,1),\ \ (3,1,2),\ \ (1,3,2),\ \ (1,2,3).$$

Definitsioon 4.8. Öeldakse, et elemendid i_k ja i_l moodustavad **inversiooni** permutatsioonis $(i_1, i_2, \ldots, i_k, \ldots, i_l, \ldots, i_n)$, kui $k < l$ ja $i_k > i_l$. Inversioonide koguarvu permutatsioonis $(i_1, i_2, \ldots, i_n)$ tähistame sümboliga $I(i_1, i_2, \ldots, i_n)$.

Permutatsiooni nimetatakse **paarispermutatsiooniks**, kui inversioonide koguarv selles permutatsioonis on paarisarv. Vastasel juhul nimetatakse seda permutatsiooni **paarituks permutatsiooniks**.

Näide 4.9. Permutatsioonis $(4, 1, 3, 5, 2)$ moodustavad inversiooni elementide paarid $(4, 1)$, $(4, 3)$, $(4, 2)$, $(3, 2)$ ja $(5, 2)$. Seega inversioone on 5 tükki, $I(4, 1, 3, 5, 2) = 5$, ja tegemist on paaritu permutatsiooniga.

Permutatsioonis $(1, 2, 3, 4, 5)$ on aga 0 inversiooni ja seega on tegu paarispermutatsiooniga.

Lause 4.10. *Transpositsioon muudab permutatsiooni paarsust.*

Tõestus. Vaatleme esialgu juhtumit, kus permutatsioonis vahetatakse ära kõrvutiasetsevad elemendid i ja j. Sellise transpositsiooni käigus ei muutu nende inversioonide arv, mida i ja j moodustavad ülejäänud elementidega. Kui i ja j enne transpositsiooni ei moodustanud inversiooni, siis pärast transpositsiooni nad moodustavad, ja vastupidi. Seega inversioonide koguarv kas suureneb või väheneb ühe võrra ning sellega paarsus muutub.

Nüüd vaatleme olukorda, kus äravahetatavate elementide i ja j vahel on m elementi $i_1, \ldots, i_m$, s.t. et permutatsioon on kujul

$$(\ldots, i, i_1, \ldots, i_m, j, \ldots). \tag{9}$$

Sel juhul kujutame i ja j äravahetamist ette järgmiselt. Vahetame ära i ja i_1, siis i ja i_2 jne., kuni vahetame ära i ja i_m ning seejärel i ja j. Sellega oleme jõudnud permutatsioonini

$$(\ldots, i_1, \ldots, i_m, j, i, \ldots).$$

Liikudes paremalt vasakule vahetame ära j ja i_m, j ja i_{m-1}, jne. kuni on vahetatud j ja i_1. Tulemuseks on permutatsioon

$$(\ldots, j, i_1, \ldots, i_m, i, \ldots). \tag{10}$$

Seega näeme, et i-d ja j-i vahetava transpositsiooni saab esitada $2m + 1$ kõrvutiasetsevate elementide transpositsiooni järjestrakendamisena. Tõestuse esimese osa põhjal teame, et niimoodi muutub permutatsiooni paarsus $2m+1$ korda, mis aga tähendabki, et permutatsiooni (9) paarsus erineb permutatsiooni (10) paarsusest. □

Definitsioon 4.11. Substitutsiooniks lõplikul hulgal M nimetatakse hulga M bijektiivset teisendust, s.t. üksühest pealekujutust $M \to M$.

Märkus 4.12. Tihti (eriti ingliskeelses kirjanduses) kasutatakse ka lõpliku hulga bijektiivsetest teisendustest rääkides sõna "permutatsioon". Selles kursuses me üritame sellist lähenemist vältida.

Harilikult vaadeldakse substitutsioone hulgal $M = \{1, 2, \ldots, n\}$ ja kutsutakse neid **substitutsioonideks n elemendist**. Kõigi substitutsioonide hulka n-elemendist tähistatakse sümboliga S_n. Substitutsioone tähistatakse tavaliselt väikeste kreeka tähtedega $\sigma, \tau, \ldots$.

Substitutsiooni esitamiseks kasutatakse tihti 2-realist ja n-veerulist tabelit (see tähendab $(2 \times n)$-maatriksit), mille esimeses reas on hulga $\{1, 2, \ldots, n\}$ elemendid mingis järjekorras ja teises reas on esimeses reas olevate elementide kujutised, seega

$$\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ \sigma(i_1) & \sigma(i_2) & \ldots & \sigma(i_n) \end{pmatrix}.$$

Nii sellise tabeli esimese rea kui ka teise rea elemendid moodustavad permutatsiooni.

Definitsioon 4.13. Substitutsiooni nimetatakse **paarissubstitutsiooniks**, kui inversioonide koguarv permutatsioonides tema esituses tabelina on paarisarv. Vastasel korral nimetatakse seda substitutsiooni **paarituks substitutsiooniks**.

Osutub, et see definitsioon on korrektne selles mõttes, et substitutsiooni paarsus ei sõltu tema esitusest tabelina. Selle näitamiseks oletame, et substitutsioon σ on esitatud kahel viisil:

$$\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ \sigma(i_1) & \sigma(i_2) & \ldots & \sigma(i_n) \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \ldots & j_n \\ \sigma(j_1) & \sigma(j_2) & \ldots & \sigma(j_n) \end{pmatrix}.$$

Kui vahetame tabelis ära kaks veergu, siis saame sama substitutsiooni uue esituse. Selle vahetuse käigus toimub nii ülemises kui alumises permutatsioonis transpositsioon, mis muudab kummagi permutatsiooni paarsust. Inversioonide koguarvu paarsus tabelis aga ei muutu. Järjestame nüüd permutatsioonid n elemendist lauses 4.6 kirjeldatud viisil nii, et alustame permutatsioonist $(i_1, i_2, \ldots, i_n)$. Selles järjestuses peab esinema ka permutatsioon $(j_1, j_2, \ldots, j_n)$. Tehes vastavad transpositsioonid tabeli veergudega saame esimesest tabelist teise, kusjuures inversioonide koguarvu paarsus ühegi sammu käigus ei muutu. See näitabki, et inversioonide koguarvu paarsus ei sõltu substitutsiooni esitusest tabelina.

Näide 4.14. Leiame substitutsiooni

$$\sigma = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{pmatrix} \in S_4$$

paarsuse. Kuna $I(2,3,4,1) + I(3,4,1,2) = 3 + 4 = 7$ on paaritu arv, siis σ on paaritu substitutsioon.

Iga substitutsiooniga saab siduda tema "märgi" (+ või −). Täpsemalt öeldes võib vaadelda kujutust

$$\operatorname{sign} : \bigcup_{n \in \mathbb{N}} S_n \to \{1, -1\}$$

(ladinakeelsest sõnast *signum*, mis tähendab märki), mis on defineeritud võrdusega

$$\operatorname{sign}(\sigma) := \begin{cases} 1, & \text{kui } \sigma \text{ on paarissubstitutsioon,} \\ -1, & \text{kui } \sigma \text{ on paaritu substitutsioon.} \end{cases}$$

Kõige sagedamini esitatakse substitutsioon tabelina nii, et selle esimeses reas on loomulik permutatsioon $(1, 2, 3, \ldots, n)$:

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Sellist tabelit nimetatakse substitutsiooni σ **normaalkujuks**.

Tuletame meelde, et hulga M teisendusi on võimalik korrutada (s.t. järjest rakendada). Kui τ ja σ on hulga M teisendused, siis nende korrutis $\tau\sigma$ (vahel tähistatakse ka $\tau \circ \sigma$) on hulga M teisendus, mis on defineeritud eeskirjaga

$$\boxed{(\tau\sigma)(m) := \tau(\sigma(m))}$$

iga $m \in M$ korral. Seega saame korrutada ka substitutsioone ning on teada, et selline korrutamine on assotsiatiivne.

Näide 4.15. Leiame substitutsioonide

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \quad \text{ja} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

korrutise $\tau\sigma$. Kuna $(\tau\sigma)(1) = \tau(\sigma(1)) = \tau(2) = 4$, $(\tau\sigma)(2) = \tau(\sigma(2)) = \tau(3) = 3$, $(\tau\sigma)(3) = 1$ ja $(\tau\sigma)(4) = 2$, siis

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}.$$

Samasusteisendust 1_M nimetame hulga M **ühiksubstitutsiooniks** ja tähistame sümboliga ε (kreeka täht epsilon). Seega kui $M = \{1, 2, \ldots, n\}$, siis

$$\varepsilon = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

On selge, et mistahes $\sigma \in S_n$ korral $\varepsilon\sigma = \sigma = \sigma\varepsilon$.

Hulgateooriast teame, et hulga M igal bijektiivsel teisendusel σ on olemas pöördteisendus, s.t. teisendus σ^{-1}, mis rahuldab seoseid $\sigma\sigma^{-1} = 1_M$ ja $\sigma^{-1}\sigma = 1_M$. Substitutsiooni $\sigma \in S_n$ pöördteisendust nimetatakse σ **pöördsubstitutsiooniks** ja tähistatakse sümboliga σ^{-1}. Eelnevat arvestades võime sõnastada järgmise tulemuse.

Lause 4.16. *Iga naturaalarvu n korral on S_n rühm substitutsioonide korrutamise suhtes.*

Rühmi S_n, kus $n \in \mathbb{N}$, nimetatakse **substitutsioonirühmadeks** ehk **sümmeetrilisteks rühmadeks**.

On selge, et kui

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}, \tag{11}$$

siis

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \ldots & \sigma(n) \\ 1 & 2 & \ldots & n \end{pmatrix}, \tag{12}$$

s.t. pöördsubstitutsiooni saamiseks võib σ esituses tabelina read ära vahetada.

Näide 4.17. Kui

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},$$

siis

$$\sigma^{-1} = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

Lause 4.18. *Substitutsioon ja tema pöördsubstitutsioon on sama paarsusega.*

Tõestus. Inversioonide koguarv tabelites (11) ja (12) on sama. □

4.2. Determinandi definitsioon

Seome nüüd iga ruutmaatriksiga ühe korpuse elemendi.

Definitsioon 4.19. Ruutmaatriksi $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ **determinandiks** nimetatakse korpuse K elementi

$$\boxed{|A| := \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}.} \tag{13}$$

n-ndat järku ruutmaatriksi determinanti nimetatakse **n-ndat järku determinandiks**. Korrutisi $a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$ nimetatakse **determinandi $|A|$ liikmeteks**. Maatriksi $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ determinandist rääkides kasutatakse tihti tähistust

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

samuti kirjutatakse $|A|$ asemel $\det(A)$.

Kommenteerime pisut seda definitsiooni. Selles summas liidetakse märgiga ($\mathrm{sign}(\sigma)$) varustatud korrutisi n-st maatriksi A elemendist, kusjuures korrutises on igast reast ja igast veerust üks tegur. Seda, et igast veerust on võetud täpselt üks tegur, näitab see, et elementide veeruindeksid moodustavad permutatsiooni $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. Kuna permutatsioonis $(1, 2, \ldots, n)$ on 0 inversiooni, siis $\mathrm{sign}(\sigma)$ sõltub inversioonide arvust permutatsioonis $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. Kui see on paarisarv, siis esineb summas liidetav $a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$, vastasel korral aga korrutise $a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$ vastandelement korpuses K. Summeerimine toimub üle kõigi substitutsioonide hulgal $\{1, 2, \ldots, n\}$, s.t. liita tuleb kõikvõimalikud sellised korrutised.

Kui tähistame sümboliga $P(n)$ kõigi permutatsioonide hulga n elemendist, siis võib determinandi definitsiooni anda ka järgmisel kujul:

$$|A| = \sum_{(i_1, \ldots, i_n) \in P(n)} (-1)^{I(i_1, \ldots, i_n)} \cdot a_{1i_1} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n},$$

sest

$$(-1)^{I(i_1, \ldots, i_n)} = \mathrm{sign} \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}.$$

Definitsiooni põhjal on lihtne veenduda, et kui $A = (a) \in \mathrm{Mat}_1(K)$, siis $|A| = a$, ja et

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Kuna $|S_n| = n!$, siis definitsiooni järgi arvutades tuleb näiteks 4-ndat järku determinandi puhul liita $4! = 24$ korrutist, 5-ndat järku determinandi puhul $5! = 120$ korrutist jne. On selge, et vähegi suurema järgu korral on determinandi arvutamine definitsiooni järgi väga töömahukas. Õnneks on determinandil mitmeid omadusi, mis tema arvutamist lihtsustavad. Vaatlemegi neid omadusi lähemalt.

4.3. Determinandi omadused

Teoreem 4.20. *Transponeerimisel maatriksi determinant ei muutu.*

TÕESTUS. Olgu $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ ja $A^T = (u_{ij})$. Me peame tõestama, et

$$\boxed{|A^T| = |A|.}$$

Kuna iga $\sigma \in S_n$ korral $(\sigma^{-1})^{-1} = \sigma$, siis kujutus

$$f : S_n \longrightarrow S_n, \quad \sigma \mapsto \sigma^{-1}$$

on pööratav ($ff = 1_{S_n}$ ehk $f^{-1} = f$) ja seega bijektiivne. Tähistame

$$s_\sigma := \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$$

ja meenutame, et

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \ldots & \sigma(n) \\ 1 & 2 & \ldots & n \end{pmatrix}. \tag{14}$$

Nüüd

$$\begin{aligned}
|A| &= \sum_{\sigma \in S_n} s_\sigma && (|A| \text{ def.}) \\
&= \sum_{\sigma \in S_n} s_{f(\sigma)} && (\text{lause } 3.15) \\
&= \sum_{\sigma \in S_n} s_{\sigma^{-1}} && (f \text{ def.}) \\
&= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma^{-1}) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(n)n} && (\text{võrdus } (14)) \\
&= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(n)n} && (\text{lause } 4.18) \\
&= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot u_{1\sigma(1)} \cdot u_{2\sigma(2)} \cdot \ldots \cdot u_{n\sigma(n)} && (u_{ij} = a_{ji}) \\
&= |A^T|. && (|A^T| \text{ def.})
\end{aligned}$$

□

Lause 4.21. *Kui ruutmaatriks sisaldab nullidest koosnevat rida, siis tema determinant on* 0*.*

TÕESTUS. Kui maatriksi $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ k-s rida koosneb ainult nullidest, siis igas korrutises $a_{1\sigma(1)}a_{2\sigma(2)} \ldots a_{n\sigma(n)}$ on k-s tegur võrdne nulliga ja seega on summa (13) kõik liidetavad nullid. □

Järeldus 4.22. *Kui ruutmaatriks sisaldab nullidest koosnevat veergu, siis tema determinant on* 0*.*

TÕESTUS. Kui ruutmaatriks A sisaldab nullidest koosnevat veergu, siis maatriks A^T sisaldab nullidest koosnevat rida. Lause 4.21 põhjal $|A^T| = 0$. Teoreemi 4.20 kasutades saame, et $|A| = |A^T| = 0$. □

Märkus 4.23. Edaspidi sõnastame ja tõestame veel terve rea determinantide omadusi ridade abil. Samasugused omadused saaks sõnastada ka veergude abil ja tõestada need Teoreemi 4.20 kasutades. Me ei hakka neid omadusi siin kirja panema ega tõestama, kuid vajaduse korral kasutame neid determinandi arvutamisel.

Lause 4.24. *Kui ruutmaatriksi mingi rea kõik elemendid korrutada elemendiga c, siis tema determinant korrutub ka elemendiga c.*

TÕESTUS. Olgu $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ ja olgu B maatriks, mis on saadud maatriksist A k-nda rea elementide korrutamisel elemendiga $c \in K$. Siis kasutades korpuse korrutamise kommutatiivsust ja summeerimise omadust SO1 saame, et

$$\begin{aligned} |B| &= \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{k-1,\sigma(k-1)} \cdot ca_{k\sigma(k)} \cdot a_{k+1,\sigma(k+1)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &= c \left(\sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \right) = c \cdot |A|. \end{aligned}$$

□

Näide 4.25. Kui maatriksi

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \end{pmatrix}$$

teine rida korrutada arvuga 3, siis saadava maatriksi determinant on $3 \cdot |A|$:

$$\begin{vmatrix} 2 & 0 & 1 \\ 3 & 6 & -3 \\ 1 & 3 & 1 \end{vmatrix} = 3 \cdot \begin{vmatrix} 2 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \end{vmatrix}.$$

Lause 4.26. *Kui ruutmaatriksis vahetada ära kaks rida, siis determinant muudab märki.*

TÕESTUS. Olgu $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ ja olgu $B = (b_{ij})$ maatriks, mis on saadud maatriksist A k-nda ja l-nda rea äravahetamisel, kus $k < l$, $k, l \in \{1, \ldots, n\}$. Siis

$$b_{ij} = \begin{cases} a_{lj}, & \text{kui } i = k, \\ a_{kj}, & \text{kui } i = l, \\ a_{ij}, & \text{kui } i \notin \{k, l\}. \end{cases} \tag{15}$$

Iga $\sigma \in S_n$ korral vaatleme substitutsiooni

$$\sigma' = \begin{pmatrix} 1 & \ldots & k & \ldots & l & \ldots & n \\ \sigma(1) & \ldots & \sigma(l) & \ldots & \sigma(k) & \ldots & \sigma(n) \end{pmatrix}, \tag{16}$$

mille ülemine permutatsioon on loomulik permutatsioon ja alumine permuatsioon on saadud permutatsioonist $(\sigma(1),\dots,\sigma(k),\dots,\sigma(l),\dots,\sigma(n))$ transpositsiooni abil, mis vahetab ära k-nda ja l-nda elemendi. Et transpositsioon muudab permutatsiooni paarsust (lause 4.10), siis $\operatorname{sign}(\sigma') = -\operatorname{sign}(\sigma)$ ehk $\operatorname{sign}(\sigma) = -\operatorname{sign}(\sigma')$.

Kuna iga $\sigma \in S_n$ korral $(\sigma')' = \sigma$, siis kujutus

$$f : S_n \longrightarrow S_n, \quad \sigma \mapsto \sigma'$$

on pööratav ($f^{-1} = f$) ja seega bijektiivne. Tähistame

$$s_\sigma := \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)} .$$

Nüüd

$$\begin{aligned}
|A| &= \sum_{\sigma\in S_n} s_\sigma && (|A| \text{ def.})\\
&= \sum_{\sigma\in S_n} s_{f(\sigma)} && (\text{lause } 3.15)\\
&= \sum_{\sigma\in S_n} s_{\sigma'} && (f \text{ def.})\\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma') \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{k\sigma(l)} \cdot \ldots \cdot a_{l\sigma(k)} \cdot \ldots \cdot a_{n\sigma(n)} && (\text{võrdus } (16))\\
&= \sum_{\sigma\in S_n} (-1) \cdot \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{k\sigma(l)} \cdot \ldots \cdot a_{l\sigma(k)} \cdot \ldots \cdot a_{n\sigma(n)} && (\operatorname{sign}(\sigma') = -\operatorname{sign}(\sigma))\\
&= -\sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{k\sigma(l)} \cdot \ldots \cdot a_{l\sigma(k)} \cdot \ldots \cdot a_{n\sigma(n)} && (\text{SO1})\\
&= -\sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \cdot b_{1\sigma(1)} \cdot \ldots \cdot b_{l\sigma(l)} \cdot \ldots \cdot b_{k\sigma(k)} \cdot \ldots \cdot b_{n\sigma(n)} && (\text{def. } 15)\\
&= -\sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \cdot b_{1\sigma(1)} \cdot \ldots \cdot b_{k\sigma(k)} \cdot \ldots \cdot b_{l\sigma(l)} \cdot \ldots \cdot b_{n\sigma(n)} && (\text{korrutamise kommut.})\\
&= -|B| . && (|B| \text{ def.})
\end{aligned}$$

ehk $|B| = -|A|$. □

Lause 4.27. *Kui ruutmaatriksis on kaks võrdset rida, siis tema determinant on* 0.

Tõestus. Olgu maatriksis $A = (a_{ij}) \in \operatorname{Mat}_n(K)$ k-s ja l-s rida võrdsed, $k < l$, $k, l \in \{1,\dots,n\}$. Iga $\sigma \in S_n$ korral olgu

$$s_\sigma = \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k\sigma(k)} \dots a_{l\sigma(l)} \dots a_{n\sigma(n)}.$$

Siis $|A| = \sum_{\sigma\in S_n} s_\sigma$. Iga $\sigma \in S_n$ korral olgu

$$\sigma' = \begin{pmatrix} 1 & \dots & k & \dots & l & \dots & n \\ \sigma(1) & \dots & \sigma(l) & \dots & \sigma(k) & \dots & \sigma(n) \end{pmatrix},$$

s.t. σ' on substitutsioon, mille tabelina esitus saadakse σ esitusest alumises reas k-nda ja l-nda elemendi vahetamisel. Siis $\operatorname{sign}(\sigma) = -\operatorname{sign}(\sigma')$. Kuna k-s ja l-s rida on võrdsed, siis $a_{k\sigma(k)} =$

$a_{l\sigma(k)}$ ja $a_{k\sigma(l)} = a_{l\sigma(l)}$. Seega

$$\begin{aligned} s_\sigma &= \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{l\sigma(k)} \dots a_{k\sigma(l)} \dots a_{n\sigma(n)} && \text{(asendused)} \\ &= \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k\sigma(l)} \dots a_{l\sigma(k)} \dots a_{n\sigma(n)} && \text{(korrutamise kommut.)} \\ &= -\operatorname{sign}(\sigma') a_{1\sigma'(1)} \dots a_{k\sigma'(k)} \dots a_{l\sigma'(l)} \dots a_{n\sigma'(n)} && (\sigma' \text{ def.}) \\ &= -s_{\sigma'} && (s_{\sigma'} \text{ def.}) \end{aligned}$$

ehk

$$s_\sigma + s_{\sigma'} = 0.$$

Kui $\tau \notin \{\sigma, \sigma'\}$, siis ka $\tau' \notin \{\sigma, \sigma'\}$ ja $\{\sigma, \sigma'\} \cap \{\tau, \tau'\} = \emptyset$. Järelikult jaguneb hulk S_n kaheelemendiliste hulkade $\{\sigma, \sigma'\}$ lõikumatuks ühendiks. Iga sellise paari korral kehtib võrdus $s_\sigma + s_{\sigma'} = 0$, seega ka $|A| = \sum_{\sigma \in S_n} s_\sigma = 0$. □

Järgmises lauses mõtleme maatriksi rea elemendiga c korrutamise all seda, et selle rea kõik elemendid korrutatakse elemendiga c ning ridade liitmise all mõeldakse seda, et omavahel liidetakse nende ridade vastavad elemendid. Näiteks maatriksi

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ -2 & -4 & -5 \end{pmatrix}$$

kolmandale reale arvuga 2 korrutatud esimese rea liitmisel saame maatriksi

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Lause 4.28. *Kui ruutmaatriksi mingile reale liita suvalise korpuse elemendiga korrutatud teine rida, siis selle maatriksi determinant ei muutu.*

Tõestus. Olgu $A = (a_{ij}) \in \operatorname{Mat}_n(K)$ ja olgu B maatriks, mis on saadud maatriksist A selle k-ndale reale elemendiga c korrutatud l-nda rea liitmisel, kus $k \neq l$. Peame näitama, et $|A| = |B|$.

Eeldame, et $k < l$. (Kui $k > l$, siis on tõestus analoogiline.) Maatriksi B k-s rida koosneb elementidest

$$a_{k1} + ca_{l1},\ a_{k2} + ca_{l2},\ \dots\ , a_{kn} + ca_{ln};$$

kõik ülejäänud B elemendid on samad, mis A vastavatel kohadel olevad elemendid. Seega

$$|B| = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1,\sigma(k-1)} (a_{k\sigma(k)} + ca_{l\sigma(k)}) a_{k+1,\sigma(k+1)} \dots a_{l\sigma(l)} \dots a_{n\sigma(n)}.$$

Kasutades distributiivsuse seadusi ja summeerimise omadusi võime kirjutada

$$\begin{aligned} |B| &= \sum_{\sigma \in S_n} \big(\operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1,\sigma(k-1)} a_{k\sigma(k)} a_{k+1,\sigma(k+1)} \dots a_{l\sigma(l)} \dots a_{n\sigma(n)} \\ &\quad + \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1,\sigma(k-1)} (ca_{l\sigma(k)}) a_{k+1,\sigma(k+1)} \dots a_{l\sigma(l)} \dots a_{n\sigma(n)} \big) \\ &= |A| + c \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{k-1,\sigma(k-1)} a_{l\sigma(k)} a_{k+1,\sigma(k+1)} \dots a_{l\sigma(l)} \dots a_{n\sigma(n)} \\ &= |A| + c \cdot 0 \\ &= |A|, \end{aligned}$$

kus altpoolt kolmandas reas olev summa on 0 sellepärast, et ta on sellise maatriksi

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{k-1,1} & a_{k-1,2} & \dots & a_{k-1,n} \\ a_{l1} & a_{l2} & \dots & a_{ln} \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,n} \\ \dots & \dots & \dots & \dots \\ a_{l1} & a_{l2} & \dots & a_{ln} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

determinant, milles k-s ja l-s rida on võrdsed. □

Lause 4.28 on väga kasulik. Selle ja tema analoogi abil saame ridu või veerge omavahel liita nii, et tekiks juurde nulliga võrduvaid elemente, aga determinant samal ajal ei muutu. Mida rohkem on maatriksis nulle, seda lihtsam on leida tema determinandi väärtust.

4.4. Laplace'i teoreem

Selle paragrahvi eesmärk on tutvuda Laplace'i[4] teoreemiga, mis lubab n-ndat järku determinandi arvutamise taandada madalamat järku determinantide arvutamisele. Selleks on meil vaja miinori mõistet ja sellega seotud mõisteid.

Definitsioon 4.29. Maatriksi A **alammaatriksiks** nimetatakse maatriksit, mis saadakse, kui maatriksis A valitakse välja mingid read ja veerud ning moodustatakse uus maatriks elementidest, mis asuvad väljavalitud ridade ja veergude lõikekohtades, kusjuures nende elementide omavahelist asendit ei muudeta. **Alamruutmaatriks** on alammaatriks, milles on sama arv ridu ja veerge.

Definitsioon 4.30. Maatriksi A **k-ndat järku miinor** on tema k-ndat järku alamruutmaatriksi determinant.

Kui miinor on sellise alamruutmaatriksi determinant, mis on saadud ridade $i_1, \dots, i_k$ ja veergude $j_1, \dots, j_k$ väljavalimisel, siis ütleme, et see *miinor asub ridades* $i_1, \dots, i_k$ *ja veergudes* $j_1, \dots, j_k$.

Definitsioon 4.31. Kui ruutmaatriksi A miinor M asub ridades $i_1, \dots, i_k$ ja veergudes $j_1, \dots, j_k$, siis selle miinori **täiendusmiinoriks** nimetatakse miinorit $\tilde{M}$, mis asub ridades ja veergudes, mis jäävad järele ridade $i_1, \dots, i_k$ ja veergude $j_1, \dots, j_k$ väljajätmisel maatriksist A. Korpuse elementi

$$(-1)^{i_1+\dots+i_k+j_1+\dots+j_k}\tilde{M}$$

nimetatakse miinori M **algebraliseks täiendiks**.

Kui me tahame rõhutada, et vaadeldav miinor asub ridades $i_1, \dots, i_k$ ja veergudes $j_1, \dots, j_k$, siis kasutame selle miinori, tema täiendusmiinori ja algebralise täiendi jaoks järgmisi tähistusi:

$$M^{j_1,\dots,j_k}_{i_1,\dots,i_k}, \qquad \tilde{M}^{j_1,\dots,j_k}_{i_1,\dots,i_k}, \qquad A^{j_1,\dots,j_k}_{i_1,\dots,i_k}.$$

[4] Pierre-Simon Laplace (1749–1827) — prantsuse matemaatik.

Näide 4.32. Kui maatriksis

$$A = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix}$$

valime välja read numbritega 2 ja 4 ning veerud numbritega 1 ja 4, siis saame, et

$$\begin{aligned} M_{2,4}^{1,4} &= \begin{vmatrix} e & h \\ m & p \end{vmatrix} = ep - hm, \\ \tilde{M}_{2,4}^{1,4} &= \begin{vmatrix} b & c \\ j & k \end{vmatrix} = bk - cj, \\ A_{2,4}^{1,4} &= (-1)^{2+4+1+4} \cdot \tilde{M}_{2,4}^{1,4} = -(bk - cj) = cj - bk. \end{aligned}$$

Järgneva teoreemi tõestas prantsuse matemaatik Pierre-Simon Laplace 1772. aastal. Selles kursuses ei jõua me selle teoreemi tõestust anda. Tõestuse võib leida näiteks raamatust [1].

Teoreem 4.33 (Laplace'i teoreem). *Olgu maatriksis $A \in \mathrm{Mat}_n(K)$ fikseeritud read $i_1,\ldots,i_k$. Siis*

$$\boxed{|A| = \sum_{1 \leqslant j_1 < \ldots < j_k \leqslant n} M_{i_1,\ldots,i_k}^{j_1,\ldots,j_k} \cdot A_{i_1,\ldots,i_k}^{j_1,\ldots,j_k},}$$

s.t. A determinant on võrdne ridades $i_1, \ldots, i_k$ asuvate kõikvõimalike k-ndat järku miinorite ja nende miinorite algebraliste täiendite korrutiste summaga.

Kui Laplace'i teoreemi rakendatakse ridade $i_1, \ldots, i_k$ korral, siis räägitakse maatriksi A *determinandi arendamisest ridade $i_1, \ldots, i_k$ järgi.* Kuna transponeerimisel determinant ei muutu (vt. teoreemi 4.20), siis võib determinanti arendada ka veergude järgi.

Eriti kasulik on determinanti arendada mingite k rea järgi siis, kui neis ridades on ainult üks nullist erinev miinor. Sellisel juhul tuleb summasse ainult üks liidetav.

Näide 4.34. Järgneva determinandi kahes esimeses reas on ainult üks nullist erinev teist järku miinor (1. ja 3. veerus). Seega kahe esimese rea järgi arendades saame:

$$\begin{vmatrix} 1 & 0 & 2 & 0 \\ 3 & 0 & 4 & 0 \\ 9 & 5 & 10 & 6 \\ 11 & 7 & 12 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \cdot (-1)^{1+2+1+3} \cdot \begin{vmatrix} 5 & 6 \\ 7 & 8 \end{vmatrix} = (-2) \cdot (-1) \cdot (-2) = -4.$$

Loomulikult võib determinanti arendada ka ühe rea või veeru järgi. Esimest järku miinor maatriksi A i-ndas reas ja j-ndas veerus on võrdne sellel kohal oleva elemendiga a_{ij}. Sellise miinori algebralist täiendit tähistatakse A_i^j asemel harilikult sümboliga A_{ij} ja nimetatakse **elemendi a_{ij} algebraliseks täiendiks**. Niisiis võime Laplace'i teoreemi põhjal öelda, et arendades i-nda rea järgi

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = \sum_{j=1}^{n} a_{ij}A_{ij} \qquad (17)$$

ja arendades j-nda veeru järgi

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}.$$

Eelöeldut kokku võttes võib sõnastada järgmise meetodi n-ndat järku ruutmaatriksi A determinandi arvutamiseks.

1. Valime determinandis välja ühe rea või veeru (soovitavalt sellise, kus juba on nulle). Kui see rida või veerg koosneb ainult nullidest, siis $|A| = 0$. Vastasel korral võtame ühe nullist erineva elemendi ja muudame selle abil kõik ülejäänud elemendid antud reas või veerus nulliks kasutades liitmisteisendust. Determinant selle käigus ei muutu.

2. Arendame determinanti valitud rea või veeru järgi. Laplace'i teoreemi põhjal taandub $|A|$ arvutamine ühe $(n-1)$-st järku determinandi arvutamisele.

3. Kordame seda protseduuri kuni jõuame esimest järku determinandini.

4.5. Maatriksite korrutise determinant

Selles paragrahvis näitame, et determinandi leidmine on kooskõlas maatriksite korrutamisega.

Teoreem 4.35. *Sama järku ruutmaatriksite korrutise determinant on võrdne tegurite determinantide korrutisega.*

Tõestus. Olgu $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}_n(K)$. Meie eesmärk on tõestada, et

$$\boxed{|AB| = |A| \cdot |B|.}$$

Vaatleme $2n$-järku determinanti

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix}.$$

Nagu näha, on D sellise maatriksi determinant, mis koosneb neljast nn. blokist: maatriksitest A, Θ, $-E$ ja B. Arendades seda determinanti D esimese n rea järgi näeme, et

$$D = |A| \cdot |B|,$$

sest $(1 + 2 + \ldots + n) + (1 + 2 + \ldots + n)$ on paarisarv. Teisendame nüüd seda determinanti nii, et kõik elemendid b_{ij} muutuksid nullideks. Selleks liidame $(n+1)$-sele veerule

- b_{11}-kordse esimese veeru,
- b_{21}-kordse teise veeru,
- $\ldots$,
- b_{n1}-kordse n-nda veeru.

Selle tulemusena ei jää $(n+1)$-se veeru esimesed n elementi enam nullideks, vaid omandavad uued väärtused

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1}, \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} + \ldots + a_{2n}b_{n1}, \\ &\ldots \\ c_{n1} &= a_{n1}b_{11} + a_{n2}b_{21} + \ldots + a_{nn}b_{n1}. \end{aligned}$$

Teisendades analoogiliselt determinandis D ka ülejäänud elemendid b_{ij} nullideks saame, et

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{vmatrix}, \tag{18}$$

kus

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Tähistades $C = (c_{ij}) \in \mathrm{Mat}_n(K)$ näeme, et maatriks C on maatriksite A ja B korrutis, $AB = C$. Arendame nüüd determinanti D viimase n rea järgi kasutades selleks avaldist (18). Saame

$$\begin{aligned} D &= (-1)^{1+2+\ldots+n+(n+1)+\ldots+2n} \cdot \begin{vmatrix} -1 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & -1 \end{vmatrix} \cdot |C| = (-1)^{1+2+\ldots+2n} \cdot (-1)^n \cdot |C| \\ &= (-1)^{1+2+\ldots+2n+n}|C| = |C|, \end{aligned}$$

sest järgnev summa on paarisarv:

$$1 + 2 + \ldots + 2n + n = \frac{(1+2n)2n}{2} + n = (1+2n)n + n = (2+2n)n = 2n(n+1).$$

□

5. Pöördmaatriks

Definitsioon 5.1. Maatriksi $A \in \mathrm{Mat}_n(K)$ **pöördmaatriksiks** nimetatakse sellist maatriksit $B \in \mathrm{Mat}_n(K)$, mille korral

$$AB = E \quad \text{ja} \quad BA = E.$$

Maatriksit $A \in \mathrm{Mat}_n(K)$ nimetatakse **pööratavaks**, kui tal leidub pöördmaatriks.

Teiste sõnadega võib öelda, et pööratavad maatriksid on monoidi $(\mathrm{Mat}_n(K), \cdot)$ pööratavad elemendid.

On selge, et mitte kõik ruutmaatriksid ei ole pööratavad. Näiteks nullmaatriksil puudub pöördmaatriks. Samas ühikmaatriks on alati pööratav, tema pöördmaatriks on ta ise, sest kehtib võrdus $EE = E$.

Lause 1.16 põhjal võime öelda, et kehtib järgmine tulemus.

Lause 5.2. *Kui ruutmaatriksil leidub pöördmaatriks, siis on see üheselt määratud.*

Maatriksi A pöördmaatriksit tähistatakse sümboliga A^{-1}. On selge, et

$$\boxed{(A^{-1})^{-1} = A.}$$

Lausest 1.17 saame järeldada järgmise fakti.

Lause 5.3. *Kui ruutmaatriksid $A, B \in \mathrm{Mat}_n(K)$ on pööratavad, siis ka maatriks AB on pööratav, kusjuures*

$$\boxed{(AB)^{-1} = B^{-1}A^{-1}.}$$

Kõigi n-ndat järku pööratavate maatriksite (üle korpuse K) hulka tähistatakse harilikult $\mathrm{GL}_n(K)$ (inglise keeles *general linear group*).

Lause 5.4. *Hulk $\mathrm{GL}_n(K)$ on rühm maatriksite korrutamise suhtes.*

Tõestus. Tänu lausele 5.3 on maatriksite korrutamine algebraline tehe hulgal $\mathrm{GL}_n(K)$. Lause 3.25 põhjal on korrutamine assotsiatiivne ja ühikmaatriks $E \in \mathrm{GL}_n(K)$ on selle suhtes ühikelement. Kuna pööratava maatriksi pöördmaatriks on ka pööratav, siis hulga $\mathrm{GL}_n(K)$ igal elemendil on samas hulgas olemas pöördelement korrutamise suhtes. Seega on rühma definitsiooni kõik tingimused täidetud. □

Definitsioon 5.5. Ruutmaatriksit A nimetatakse **regulaarseks**, kui $|A| \neq 0$.

Lause 5.6. *Iga pööratav ruutmaatriks on regulaarne.*

Tõestus. Olgu maatriks $A \in \mathrm{Mat}_n(K)$ pööratav. Siis tal leidub pöördmaatriks A^{-1} nii, et $AA^{-1} = E$. Kuna teoreemi 4.35 põhjal on korrutise determinant võrdne tegurite determinantide korrutisega, siis $1 = |E| = |AA^{-1}| = |A| \cdot |A^{-1}|$. Kui oleks $|A| = 0$, siis oleks $1 = |A| \cdot |A^{-1}| = 0$, mis aga korpuses pole võimalik. Järelikult $|A| \neq 0$. □

Defineerime nüüd elementaarteisendused maatriksi ridadega. Selliseid teisendusi kasutatakse paljude praktiliste ülesannete lahendamisel (pöördmaatriksi leidmine, maatriksi astaku leidmine, lineaarvõrrandisüsteemi lahendamine jne.).

Definitsioon 5.7. Elementaarteisendused maatriksi ridadega on järgmised teisendused:

1. maatriksi kahe rea äravahetamine;
2. maatriksi rea korrutamine nullist erineva korpuse elemendiga;
3. maatriksi mingile reale mingi korpuse elemendiga korrutatud teise rea liitmine.

Analoogiliselt defineeritakse elementaarteisendused maatriksi veergudega.

Lause 5.8. *Maatriksi ridade äravahetamise saab taandada teistsugust tüüpi elementaarteisendustele ridadega*

Tõestus. Olgu meil vaja vahetada ära maatriksi A i-s ja j-s rida. Toimime selleks järgmiselt. Liidame i-ndale reale j-nda rea. Siis liidame saadud maatriksis j-ndale reale (-1)-ga korrutatud i-nda rea. Seejärel liidame i-ndale reale j-nda rea. Sellega oleme teinud kolm 3. tüüpi elementaarteisendust. Lõpuks korrutame j-nda rea (-1)-ga ja saamegi tulemuseks nõutava maatriksi. □

Loomulikult kehtib analoogiline tulemus ka veergude kohta.

Näitame nüüd, et elementaarteisenduste sooritamiseks maatriksi ridade või veergudega võib seda maatriksit korrutada teatud erikujuliste maatriksitega, mis on küllaltki sarnased ühikmaatriksiga.

Olgu n fikseeritud naturaalarv. Iga $c \in K \setminus \{0\}$ ja $i \in \{1, 2, \ldots, n\}$ korral olgu $E_i(c)$ n-ndat järku ruutmaatriks, mille peadiagonaali i-s element on c, teised peadiagonaali elemendid on 1-d ja kõik ülejäänud elemendid on 0-d, s.t. $E_i(c)$ on maatriks kujul

$$E_i(c) = \begin{pmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 1 \end{pmatrix}.$$

Iga $c \in K$ ja $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$, korral olgu $E_{ij}(c)$ n-ndat järku ruutmaatriks, mille peadiagonaalil on 1-d, i-nda rea ja j-nda veeru element on c ja kõik ülejäänud elemendid on 0-d. Seega juhul, kui $i < j$, on maatriks $E_{ij}(c)$ kujul

$$E_{ij}(c) = \begin{pmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & \ldots & c & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{pmatrix}.$$

Definitsioon 5.9. Maatrikseid $E_i(c)$ ja $E_{ij}(c)$ nimetatakse **n-indat järku elementaarmaatriksiteks**.

Lause 5.10. *Elementaarteisenduste tegemine maatriksi ridadega (veergudega) on samaväärne maatriksi korrutamisega vasakult (paremalt) teatud arvu elementaarmaatriksitega.*

Tõestus. Vaatleme maatriksit $A \in \mathrm{Mat}_n(K)$. Leides korrutise $E_i(c)A$ näeme, et see on maatriks, mis on saadud maatriksist A i-nda rea korrutamisel c-ga. Samuti on lihtne veenduda, et $E_{ij}(c)A$ on maatriks, mis on saadud maatriksist A i-ndale reale c-kordse j-nda rea liitmisel. Lause 5.8 põhjal taandub ridade vahetus teatud arvule 2. ja 3. tüüpi elementaarteisendustele, seega samuti elementaarmaatriksitega vasakult korrutamisele.

Tõestus veergude jaoks on analoogiline. □

Lause 5.11. *Kõik elementaarmaatriksid on pööratavad, kusjuures ka nende pöördmaatriksid on elementaarmaatriksid.*

Tõestus. Vahetu kontroll näitab, et

$$(E_i(c))^{-1} = E_i\left(c^{-1}\right)$$

ja

$$(E_{ij}(c))^{-1} = E_{ij}(-c).$$

□

Lause 5.12. *Kui A on regulaarne maatriks ja maatriks B on saadud maatriksist A ridade või veergude elementaarteisenduste abil, siis ka B on regulaarne.*

Tõestus. Kui B on saadud maatriksist A ridade elementaarteisenduste abil, siis

$$B = E_k \dots E_2 E_1 A,$$

kus $E_1, E_2, \dots, E_k$ on elementaarmaatriksid ja seega peavad nad olema regulaarsed. Kuna $|B| = |E_k| \dots |E_2||E_1||A|$, tegurid viimases korrutises on nullist erinevad ning korpuses ei ole nullitegureid, siis ka $|B| \neq 0$. □

Lause 5.13. *Iga regulaarse maatriksi saab ridade elementaarteisenduste abil teisendada ühikmaatriksiks.*

Tõestus. Järelduse 4.22 tõttu peab regulaarse maatriksi esimeses veerus leiduma nullist erinev element. Ridade vahetamise abil võime saavutada olukorra, kus see element asub kohal $(1,1)$. Vaatleme nüüd regulaarset maatriksit $A = (a_{ij}) \in \mathrm{Mat}_n(K)$, kus $a_{11} \neq 0$. Liidame maatriksi A i-ndale reale, kus $i \in \{2, \dots, n\}$, elemendiga $-\frac{a_{i1}}{a_{11}}$ korrutatud esimese rea (vt. märkust 1.36). Korrutades veel esimese rea nullist erineva elemendiga $\frac{1}{a_{11}}$ jõuame maatriksini B, mis on kujul

$$B = \begin{pmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & b_{22} & b_{23} & \dots & b_{2n} \\ 0 & b_{32} & b_{33} & \dots & b_{3n} \\ \dots & \dots & \dots & \dots & \\ 0 & b_{n2} & b_{n3} & \dots & b_{nn} \end{pmatrix}.$$

Lause 5.12 põhjal on B regulaarne. Arendades B determinanti esimese veeru järgi näeme, et $|B| = |B'|$, kus

$$B' = \begin{pmatrix} b_{22} & b_{23} & \dots & b_{2n} \\ b_{32} & b_{33} & \dots & b_{3n} \\ \dots & \dots & \dots & \\ b_{n2} & b_{n3} & \dots & b_{nn} \end{pmatrix}.$$

Seega ka B' on regulaarne maatriks ja tema esimeses veerus (s.t. elementide $b_{22}, b_{32}, \dots, b_{n2}$ hulgas) peab leiduma nullist erinev element. Tõstes sarnasel moel nullist erinevaid elemente peadiagonaalile ja muutes neist allapoole jäävaid elemente nullideks jõuame ridade elementaarteisenduste abil maatriksini

$$C = \begin{pmatrix} 1 & c_{12} & c_{13} & \dots & c_{1,n-1} & c_{1n} \\ 0 & 1 & c_{23} & \dots & c_{2,n-1} & c_{2n} \\ 0 & 0 & 1 & \dots & c_{3,n-1} & c_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & c_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

s.t. ülemise kolmnurkmaatriksini, mille peadiagonaalil on ühikelemendid. Liigume nüüd veerge vaadeldes paremalt vasakule. Liidame maatriksi C i-ndale reale, kus $i \in \{1, 2, \dots, n-1\}$, elemendiga $-c_{in}$ korrutatud n-nda rea. Sellega muutuvad nulliks kõik elemendid viimases veerus, välja arvatud viimane element. Tulemuseks on maatriks kujul

$$D = \begin{pmatrix} 1 & d_{12} & d_{13} & \dots & d_{1,n-1} & 0 \\ 0 & 1 & d_{23} & \dots & d_{2,n-1} & 0 \\ 0 & 0 & 1 & \dots & d_{3,n-1} & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Samamoodi toimime eelviimase, üle-eelviimase jne. veeruga, kuni jõuame ühikmaatriksini. □

Järgmine tulemus ütleb, et maatriksi pööratavuse üle saab otsustada tema determinandi põhjal.

Teoreem 5.14. *Ruutmaatriks on pööratav parajasti siis, kui ta on regulaarne.*

Tõestus. Tarvilikkus. See on tõestatud lauses 5.6.
Piisavus. Olgu $A \in \mathrm{Mat}_n(K)$ regulaarne maatriks. Nagu näitasime lauses 5.13, saab maatriksi A ridade elementaarteisenduste abil teisendada ühikmaatriksiks. Lause 5.10 põhjal on elementaarteisenduste tegemine maatriksi ridadega samaväärne maatriksi korrutamisega vasakult teatud arvu elementaarmaatriksitega. Seega leiduvad sellised elementaarmaatriksid $E_1, E_2, \dots, E_k$, et

$$E_k \dots E_2 E_1 A = E.$$

Tähistame $B := E_k \dots E_2 E_1$, siis $BA = E$. Et elementaarmaatriksid on pööratavad (lause 5.11), siis on ka B pööratav ja lauset 5.3 $k-1$ korda rakendades saame, et

$$B^{-1} = E_1^{-1} E_2^{-1} \dots E_k^{-1}.$$

Korrutame võrduse $BA = E$ mõlemad pooled vasakult maatriksiga B^{-1}. See annab meile võrduse $B^{-1}(BA) = B^{-1}E$, millest järeldub, et

$$A = EA = (B^{-1}B)A = B^{-1}(BA) = B^{-1}E = B^{-1},$$

s.t. A on B pöördmaatriks. Definitsiooni põhjal on siis ka B maatriksi A pöördmaatriks ja seega maatriks A on pööratav. □

Järeldus 5.15. *Iga regulaarse maatriksi saab esitada elementaarmaatriksite korrutisena.*

Tõestus. Teoreemi 5.14 tõestuses nägime, et kui A on regulaarne, siis $A = E_1^{-1}E_2^{-1}\dots E_k^{-1}$, kus maatriksid $E_1, E_2, \dots, E_k$ on elementaarmaatriksid. Tänu lausele 5.11 on ka $E_1^{-1}, E_2^{-1}, \dots, E_k^{-1}$ elementaarmaatriksid. □

Märkus 5.16. Viimast fakti kasutatakse muuhulgas matemaatilises analüüsis kordsete integraalide muutujavahetuse juures.

Teoreemi 5.14 tõestuses nägime, et

$$A^{-1} = E_k \dots E_2E_1 = E_k \dots E_2E_1E.$$

See tähendab, et maatriksi A pöördmaatriksi saamiseks tuleb ühikmaatriksi E ridadega teha täpselt samad teisendused (ja täpselt samas järjekorras), mis maatriksi A ridadegagi, et saada A-st ühikmaatriks. See annab meile järgmise algoritmi A pöördmaatriksi leidmiseks. Koostame $(n \times 2n)$-indat järku maatriksi nii, et kirjutame A kõrvale paremale n-indat järku ühikmaatriksi:

$$(A|E).$$

Teeme selle maatriksi ridadega elementaarteisendusi eesmärgiga saada vasakule poole ühikmaatriks. Eelpoolöeldut arvestades jõuame lõpuks maatriksini

$$(E|A^{-1}),$$

mille põhjal saame välja kirjutada A pöördmaatriksi.

Pöördmaatriksi leidmiseks on ka veel teine võimalus, nimelt determinantide abil.

Teoreem 5.17. *Kui maatriks $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ on regulaarne, siis*

$$\boxed{A^{-1} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},}$$

kus A_{ij} on maatriksi A elemendi a_{ij} algebraline täiend.

Tõestus. Tähistame

$$A^* := \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}.$$

Vastavalt maatriksite korrutamise eeskirjale on korrutise AA^* i-nda rea ja i-nda veeru ($i \in \{1, \ldots, n\}$) element

$$a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} = |A|,$$

kus viimane võrdus kehtib Laplace'i teoreemi põhjal, sest tegemist on $|A|$ arendisega i-nda rea järgi (vt. võrdust (17)). Kui $i \neq j$, $i, j \in \{1, \ldots, n\}$, siis korrutise AA^* i-nda rea ja j-inda veeru element on

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \sum_{k=1}^{n} a_{ik}A_{jk}.$$

Moodustame maatriksi B, mille kõik read, välja arvatud j-is rida on samad, mis maatriksil A, j-indaks reaks on aga maatriksi A i-s rida, s.t.

$$B = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \ldots & \ldots & \ldots & \ldots \\ a_{j-1,1} & a_{j-1,2} & \ldots & a_{j-1,n} \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ a_{j+1,1} & a_{j+1,2} & \ldots & a_{j+1,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Et maatriksis B on kaks võrdset rida, siis $|B| = 0$. Teisest küljest, arendades maatriksi B determinanti j-inda rea järgi saame võrduse $|B| = \sum_{k=1}^{n} a_{ik}A_{jk}$ (sest elemendi a_{ik} algebraline täiend maatriksis B on sama, mis elemendi a_{jk} algebraline täiend maatriksis A, s.t. A_{jk}). Seega $\sum_{k=1}^{n} a_{ik}A_{jk} = 0$, mis tähendab, et maatriksi AA^* kõik väljaspool peadiagonaali asuvad elemendid on 0-d. Sellega oleme näidanud, et

$$AA^* = \begin{pmatrix} |A| & 0 & \ldots & 0 \\ 0 & |A| & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & |A| \end{pmatrix} = |A| \cdot E.$$

Arutledes analoogiliselt veergudega saame võrduse $A^*A = |A| \cdot E$. Et A on regulaarne, siis $|A| \neq 0$ ja on olemas element $|A|^{-1} = \frac{1}{|A|}$. Korrutades selle elemendiga võrduste $AA^* = |A| \cdot E$ ja $A^*A = |A| \cdot E$ mõlemaid pooli ja kasutades lauset 3.25(5) saame

$$A \cdot \left(\frac{1}{|A|} \cdot A^* \right) = E = \left(\frac{1}{|A|} \cdot A^* \right) \cdot A,$$

mis tänu pöördmaatriksi definitsioonile tähendabki, et

$$A^{-1} = \frac{1}{|A|} \cdot A^*.$$

□

Tõestatud teoreemist saab lihtsa vaevaga tuletada valemi teist järku ruutmaatriksi pöördmaatriksi leidmiseks.

Järeldus 5.18. *Kui $ad - bc \neq 0$, siis*

$$\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.}$$

Näide 5.19. Leiame maatriksi

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 1 & 2 \\ 3 & 5 & 7 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

pöördmaatriksi. Arvutades maatriksi A determinandi ja kõigi elementide algebralised täiendid, saame

$$\begin{aligned} |A| &= \begin{vmatrix} 2 & 3 & 4 \\ 1 & 1 & 2 \\ 3 & 5 & 7 \end{vmatrix} \overset{-2\mathrm{I}}{=} \begin{vmatrix} 2 & 3 & 0 \\ 1 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1, \\ A^{-1} &= \frac{1}{|A|}(A_{ij})^T = \frac{1}{-1} \begin{pmatrix} \begin{vmatrix} 1 & 2 \\ 5 & 7 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ 5 & 7 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 1 & 2 \\ 3 & 7 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 3 & 7 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 1 & 2 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ 3 & 5 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} \end{pmatrix} \\ &= -\begin{pmatrix} -3 & -1 & 2 \\ -1 & 2 & 0 \\ 2 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & -2 \\ 1 & -2 & 0 \\ -2 & 1 & 1 \end{pmatrix}. \end{aligned}$$

Lõpuks näitame, kuidas on omavahel seotud pöördmaatriksi leidmine ja transponeerimine.

Lause 5.20. *Kui A on pööratav ruutmaatriks, siis*

$$\boxed{\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.}$$

Tõestus. Kuna

$$\begin{aligned} \left(A^{-1}\right)^T A^T &= \left(AA^{-1}\right)^T = E^T = E, \\ A^T \left(A^{-1}\right)^T &= \left(A^{-1}A\right)^T = E^T = E, \end{aligned}$$

siis definitsiooni põhjal on $\left(A^{-1}\right)^T$ maatriksi A^T pöördmaatriks. □

6. Vektorruum. Lineaarne sõltuvus

Koolist on teada, et nii tasandil kui kolmemõõtmelises ruumis võib vaadelda vabavektoreid. Neid vabavektoreid saab omavahel liita ja arvuga korrutada ning nendel tehetel on terve rida häid omadusi. Nende struktuuride üldistamiseks on lineaaralgebras kasutusele võetud vektorruumi mõiste. Kui vabavektoreid saab korrutada reaalarvudega, siis üldisema vektorruumi elemente saab korrutada korpuse elementidega.

6.1. Vektorruumi mõiste

Definitsioon 6.1. Hulka V nimetatakse **vektorruumiks** ehk **lineaarseks ruumiks** üle korpuse K, kui on defineeritud kujutused

$$\begin{aligned} V \times V \to V, &\quad (a,b) \mapsto a+b, \\ K \times V \to V, &\quad (k,a) \mapsto ka, \end{aligned}$$

nii et

VR1. $(a+b)+c = a+(b+c)$ iga $a,b,c \in V$ korral;

VR2. leidub element $0 \in V$ nii, et iga $a \in V$ korral $a+0 = a = 0+a$;

VR3. iga elemendi $a \in V$ korral leidub element $-a \in V$ nii, et $a+(-a) = 0 = (-a)+a$;

VR4. $a+b = b+a$ iga $a,b \in V$ korral;

VR5. $k(a+b) = ka+kb$ iga $a,b \in V$ ja $k \in K$ korral;

VR6. $(k+l)a = ka+la$ iga $a \in V$ ja $k,l \in K$ korral;

VR7. $(kl)a = k(la)$ iga $a \in V$ ja $k,l \in K$ korral;

VR8. $1a = a$ iga $a \in V$ korral.

Vektorruumi V elemente on tavaks nimetada **vektoriteks** ja korpuse K elemente **skalaarideks**. Elementi $a+b \in V$ nimetatakse vektorite a ja b **summaks** ning elementi $ka \in V$ skalaari k ja vektori a **korrutiseks**. Elementi $0 \in V$ tingimuses VR2 nimetatakse **nullvektoriks** ja elementi $-a \in V$ tingimuses VR3 nimetatakse vektori a **vastandvektoriks**.

Märkus 6.2. Tingimustest VR1-VR4 näeme, et vektorruum on liitmistehte suhtes Abeli rühm.

Märkus 6.3. Matemaatilise induktsiooni abil on lihtne näidata, et kui kehtivad tingimused VR5 ja VR6, siis iga naturaalarvu n ja mistahes $a_1, \ldots, a_n, a \in V$, $k, k_1, \ldots, k_n \in K$ korral

$$\begin{aligned} k(a_1 + \ldots + a_n) &= ka_1 + \ldots + ka_n, \\ (k_1 + \ldots + k_n)a &= k_1a + \ldots + k_na. \end{aligned}$$

Näide 6.4. 1. Tasandi vabavektorite hulk $\mathbb{E}_2$ ja kolmemõõtmelise ruumi vabavektorite hulk $\mathbb{E}_3$ on vektorruumid.

2. Iga korpus K on vektorruum üle iseenda, kui kujutused $K \times K \to K$ on selle korpuse liitmis- ja korrutamistehe.

3. Iga korpuse K ja naturaalarvu n korral võib hulka K^n vaadelda vektorruumina üle K, kui tehted defineerida n.ö. komponenthaaval:

$$\begin{aligned}(k_1,\ldots,k_n)+(l_1,\ldots,l_n) &:= (k_1+l_1,\ldots,k_n+l_n),\\ k(k_1,\ldots,k_n) &:= (kk_1,\ldots,kk_n)\end{aligned}$$

mistahes $k,k_1,\ldots,k_n,l_1,\ldots,l_n\in K$ korral.

4. $(m\times n)$-maatriksite hulk $\mathrm{Mat}_{m,n}(K)$ üle korpuse K on vektorruum üle K, kui tehetena vaadelda maatriksite liitmist ja maatriksi korrutamist korpuse K elementidega. See on tõestatud lauses 3.22. Vektorruum K^n on sisuliselt sama, mis vektorruum $\mathrm{Mat}_{1,n}(K)$.

5. Hulk $\mathbb{C}$ on vektorruum üle korpuse $\mathbb{R}$, kui liitmisena vaadelda kompleksarvude liitmist ning reaalarvu c ja kompleksarvu $a+bi$ korrutis defineeritakse kui $c(a+bi):=ca+cbi$.

Lause 6.5. *Mistahes vektorruumis V üle suvalise korpuse K kehtivad järgmised arvutusreeglid.*

1. *Iga $a,b,c\in V$ korral, kui $a+b=c$, siis $a=c-b$.*
2. *$0a=0$ iga $a\in V$ korral. (Selle võrduse vasakul poolel olev 0 tähistab korpuse K nullelementi ja paremal poolel olev 0 vektorruumi V nullelementi.)*
3. *$k0=0$ iga $k\in K$ korral. (Selles võrduses on mõlemad 0-d V elemendid.)*
4. *$(-1)a=-a$ iga $a\in V$ korral. (Siin -1 on korpuse K ühikelemendi vastandelement.)*
5. *$(-k)a=k(-a)=-(ka)$ iga $k\in K$ ja $a\in V$ korral.*
6. *$k(a-b)=ka-kb$ iga $k\in K$ ja $a,b\in V$ korral.*
7. *$(k-l)a=ka-la$ iga $k,l\in K$ ja $a\in V$ korral.*

Tõestus. Kõik need omadused saab tõestada väga sarnaselt sellele, kuidas on tõestatud ringide kohta käiv lause 1.27. □

6.2. Vektorruumi alamruum

Algebras kutsutakse algebralise struktuuri mittetühje tehete suhtes kinniseid alamhulki alamstruktuurideks. Nii võib rääkida näiteks alamrühmadest, alamringidest ja alamkorpustest. Meie uurime põhjalikumalt vektorruumide alamstruktuure.

Definitsioon 6.6. Vektorruumi V mittetühja alamhulka U nimetatakse V **alamruumiks**, kui

AR1 iga $a,b\in U$ korral $a+b\in U$ (s.t. U on kinnine liitmise suhtes);

AR2 iga $a\in U$ ja $k\in K$ korral $ka\in U$ (s.t. U on kinnine skalaariga korrutamise suhtes).

Näide 6.7. 1. Iga vektorruumi V korral on V ise ja nullvektorist koosnev alamhulk $\{0\}$ selle vektorruumi alamruumid.

2. Mistahes korpuse K ja naturaalarvu n korral on n-ndat järku sümmeetriliste maatriksite (üle K) hulk alamruum vektorruumis $\mathrm{Mat}_n(K)$.

3. Hulk $U=\{(k,0,l)\mid k,l\in\mathbb{R}\}$ on alamruum vektorruumis $\mathbb{R}^3$ (üle korpuse $\mathbb{R}$).

4. Fikseeritud sirgega paralleelsete vabavektorite hulk on alamruum tasandi vabavektorite vektorruumis $\mathbb{E}_2$.

Lause 6.8. *Vektorruumi iga alamruum sisaldab nullvektorit.*

TÕESTUS. Olgu U vektorruumi V alamruum. Kuna U on mittetühi, siis leidub mingi $a \in U$. Tingimuse AR2 ja lause 6.5(2) tõttu $0 = 0a \in U$. □

Lause 6.9. *Vektorruumi V iga alamruum on ise ka vektorruum tehete suhtes, mis on defineeritud samamoodi nagu vektorruumi V tehted.*

TÕESTUS. Olgu U vektorruumi V alamruum. Tingimused AR1 ja AR2 tagavad selle, et võime vaadelda kujutusi $U \times U \to U$, $(a, b) \mapsto a+b$, ja $K \times U \to U$, $(k, a) \mapsto ka$. On selge, et tingimused VR1 ja VR4-VR8 on U korral täidetud. Lausest 6.8 järeldub, et ka tingimus VR2 kehtib U jaoks. Kui $a \in U$, siis lause 6.5 põhjal võib öelda, et $(-1)a = -a$. Kuna AR2 tõttu $(-1)a \in U$, siis ka $-a \in U$. Seega U rahuldab tingimust VR3. □

Definitsioon 6.10. Olgu V vektorruum üle korpuse K ja $a_1, \ldots, a_s \in V$. Mistahes avaldist

$$k_1a_1 + k_2a_2 + \ldots + k_sa_s, \tag{19}$$

kus $k_1, \ldots, k_s \in K$, aga ka selle avaldise poolt määratud V elementi, nimetatakse vektorite $a_1, \ldots, a_s$ **lineaarkombinatsiooniks.** Skalaare $k_1, \ldots, k_s$ nimetatakse selle lineaarkombinatsiooni **kordajateks**.

Näide 6.11. Olgu $a = (3, -1, 2)$ ja $b = (1, 4, 0)$ vektorruumi $\mathbb{R}^3$ vektorid ning $k = 2$ ja $l = -3$. Arvutame lineaarkombinatsiooni $ka + lb$:

$$ka + lb = 2 \cdot (3, -1, 2) + (-3) \cdot (1, 4, 0) = (6, -2, 4) + (-3, -12, 0) = (3, -14, 4).$$

Definitsioon 6.12. Vektorruumi V alamhulka, mis koosneb vektorite $a_1, \ldots, a_s$ kõigist lineaarkombinatsioonidest, nimetatakse vektorite $a_1, \ldots, a_s$ **lineaarseks katteks** ehk **lineaarkatteks** ja tähistatakse kas $L(a_1, \ldots, a_s)$, $\langle a_1, \ldots, a_s \rangle$ või $\operatorname{span}(a_1, \ldots, a_s)$. Käesolevas kursuses eelistame esimest tähistust.

Niisiis

$$L(a_1, \ldots, a_s) = \{k_1a_1 + \ldots + k_sa_s \mid k_1, \ldots, k_s \in K\}.$$

Lineaarkatete moodustamine annab ühe kasuliku viisi alamruumide konstrueerimiseks.

Lause 6.13. *Vektorite $a_1, \ldots, a_s$ lineaarne kate on vähim alamruum, mis neid vektoreid sisaldab.*

TÕESTUS. Olgu $a_1, \ldots, a_s$ vektorruumi V (üle korpuse K) vektorid. Kuna $0 = 0a_1 + \ldots + 0a_s \in L(a_1, \ldots, a_s)$, siis $L(a_1, \ldots, a_s)$ ei ole tühi. Kui $k_1a_1 + \ldots + k_sa_s, l_1a_1 + \ldots + l_sa_s \in L(a_1, \ldots, a_s)$ ja $k \in K$, siis

$$\begin{aligned}(k_1a_1 + \ldots + k_sa_s) + (l_1a_1 + \ldots + l_sa_s) &= (k_1 + l_1)a_1 + \ldots + (k_s + l_s)a_s \in L(a_1, \ldots, a_s),\\ k(k_1a_1 + \ldots + k_sa_s) &= (kk_1)a_1 + \ldots + (kk_s)a_s \in L(a_1, \ldots, a_s).\end{aligned}$$

Sellega oleme näidanud, et $L(a_1, \ldots, a_s)$ on V alamruum. Et

$$a_i = 0a_1 + \ldots + 0a_{i-1} + 1a_i + 0a_{i+1} + \ldots + 0a_s \in L(a_1, \ldots, a_s)$$

iga $i \in \{1, \ldots, s\}$ korral, siis $L(a_1, \ldots, a_s)$ sisaldab vektoreid $a_1, \ldots, a_s$. Kui U on mistahes alamruum, mis sisaldab vektoreid $a_1, \ldots, a_s$, siis ta peab sisaldama ka kõiki vektoreid $k_1a_1, \ldots, k_sa_s$, kus $k_1, \ldots, k_s \in K$, ning ka selliste vektorite summasid. Seega $L(a_1, \ldots, a_s) \subseteq U$, s.t. $L(a_1, \ldots, a_s)$ on vähim V alamruum, mis sisaldab vektoreid $a_1, \ldots, a_s$. □

Näide 6.14. Vektorruumi $\mathbb{R}^3$ vektorite $a = (1, 0, 0)$ ja $b = (0, 0, 1)$ lineaarne kate on alamruum $L(a, b) = \{(k, 0, l) \mid k, l \in \mathbb{R}\}$.

Lause 6.15. *Vektorite $a_1, \ldots, a_s$ ja $b_1, \ldots, b_t$ lineaarsed katted on võrdsed, $L(a_1, \ldots, a_s) = L(b_1, \ldots, b_t)$, parajasti siis, kui*

$$a_1, \ldots, a_s \in L(b_1, \ldots, b_t) \quad ja \quad b_1, \ldots, b_t \in L(a_1, \ldots, a_s).$$

Tõestus. Selle lihtsa tõestuse jätame läbimõtlemiseks lugejale. □

6.3. Lineaarne sõltumatus

Vektoritest rääkides kasutatakse tihti terminit **vektorite süsteem**. Selle all mõeldakse sellist vektorite kogumit, mis erineb vektorite hulgast selle poolest, et üks vektor võib selles esineda mitu korda. Samuti on vektorite süsteemi puhul oluline vektorite järjekord. Näiteks kui V on vektorruum ja $a, b, c \in V$, siis võib vaadelda vektorite süsteemi c, a, b, a, c. Käesolevas kursuses vaatleme vaid lõplikke vektorite süsteeme. Samuti eeldame, et vektorite süsteem ei ole tühi, s.t. ta sisaldab vähemalt ühte vektorit.

Formaalselt võib vektorite süsteemi $a_1, a_2, \ldots, a_s$ vaadelda hulga V^s elemendina, kus $s \in \mathbb{N}$. Harilikult seda vaatepunkti siiski ei rõhutata.

Anname nüüd lineaarse sõltumatuse definitsiooni.

Definitsioon 6.16. Vektorruumi V (üle korpuse K) vektorite süsteemi $a_1, a_2, \ldots, a_s$ nimetatakse **lineaarselt sõltumatuks**, kui mistahes $k_1, k_2, \ldots, k_s \in K$ korral võrdusest

$$k_1a_1 + k_2a_2 + \ldots + k_sa_s = 0$$

järeldub, et

$$k_1 = k_2 = \ldots = k_s = 0.$$

Vektorite süsteemi nimetatakse **lineaarselt sõltuvaks**, kui ta ei ole lineaarselt sõltumatu.

Niisiis vektorite süsteem $a_1, a_2, \ldots, a_s$ on lineaarselt sõltuv, kui leiduvad sellised skalaarid $k_1, k_2, \ldots, k_s \in K$, et

$$k_1a_1 + k_2a_2 + \ldots + k_sa_s = 0$$

ja vähemalt üks elementidest $k_1, k_2, \ldots, k_s$ on nullist erinev.

Definitsioon 6.17. Lineaarkombinatsiooni nimetatakse **triviaalseks**, kui kõik tema kordajad on nullid. Kui vähemalt üks kordaja on nullist erinev, siis öeldakse, et see lineaarkombinatsioon on **mittetriviaalne**.

Seega lineaarse sõltumatuse definitsiooni võib anda ka järgmisel kujul: vektorite süsteem on lineaarselt sõltumatu, kui nullvektoriga on võrdne vaid selle süsteemi triviaalne lineaarkombinatsioon. Vektorite süsteem on lineaarselt sõltuv, kui mingi mittetriviaalne lineaarkombinatsioon selle süsteemi vektoritest on võrdne nullvektoriga.

Näide 6.18. Vektorruumi $\mathbb{R}^3$ vektorite süsteem

$$\begin{aligned} a_1 &= (1, 1, -2), \\ a_2 &= (-2, -2, 4) \end{aligned}$$

on lineaarselt sõltuv, sest $2a_1 + a_2 = (0, 0, 0)$. Süsteem

$$\begin{aligned} b_1 &= (1, 0, 0), \\ b_2 &= (0, 2, 0), \\ b_3 &= (0, 0, 4) \end{aligned}$$

on aga lineaarselt sõltumatu, sest kui

$$(0, 0, 0) = k_1b_1 + k_2b_2 + k_3b_3 = (k_1, 0, 0) + (0, 2k_2, 0) + (0, 0, 4k_3) = (k_1, 2k_2, 4k_3),$$

siis $0 = k_1 = 2k_2 = 4k_3$, millest $0 = k_1 = k_2 = k_3$. Analoogiliselt saab näidata, et mistahes järku ühikmaatriksi reavektorite süsteem on lineaarselt sõltumatu.

Märkus 6.19. Kuna vektorruumi liitmistehe on kommutatiivne, siis vektorite süsteemi lineaarne sõltuvus või sõltumatus ei sõltu vaadeldavate vektorite järjestusest.

Tuleb välja, et lisaks definitsioonile on veel terve rida tingimusi, mille abil saab otsustada vektorite süsteemi lineaarse sõltuvuse või sõltumatuse üle. Järgnevas vaatleme neist mõningaid.

Järeldus 6.20. *Iga vektorite süsteem, mis sisaldab nullvektorit, on lineaarselt sõltuv.*

Tõestus. Oletame, et süsteem sisaldab nullvektorit. Moodustame lineaarkombinatsiooni, kus nullvektori kordaja on korpuse ühikelement ja kõik ülejäänud vektorid on kordajaga 0. See on nulliga võrduv mittetriviaalne lineaarkombinatsioon. Järelikult süsteem on lineaarselt sõltuv. □

Lause 6.21. *Ühestainsast vektorist a koosnev süsteem on lineaarselt sõltumatu parajasti siis, kui $a \neq 0$.*

Tõestus. Tarvilikkus. Olgu vektorist a koosnev süsteem lineaarselt sõltumatu. Kui oletame, et $a = 0$ ja võtame näiteks korpuse ühikelemendi 1, siis lause 6.5(3) põhjal $1a = 0 \in V$, mis on vastuolus süsteemi lineaarse sõltumatusega. Järelikult $a \neq 0$.
Piisavus. Olgu $a \neq 0$. Oletame, et $ka = 0$, kus k on korpuse element. Kui $k \neq 0$, siis leidub pöördelement k^{-1}. Korrutades sellega võrduse $ka = 0$ pooli saame

$$0 \underset{lause\ 6.5(3)}{=} k^{-1}0 = k^{-1}(ka) \underset{VR7}{=} (k^{-1}k)a = 1a \underset{VR8}{=} a,$$

mis on vastuolus eeldusega. Järelikult $k = 0$. Definitsiooni põhjal on vektorist a koosnev süsteem lineaarselt sõltumatu. □

Lause 6.22. *Nullist erinevate vektorite süsteemi $a_1, \ldots, a_s$, kus $s \geqslant 2$, jaoks on järgmised väited samaväärsed.*

1. *See süsteem on lineaarselt sõltuv.*
2. *Selles süsteemis leidub vektor, mis avaldub eelnevate vektorite lineaarkombinatsioonina.*
3. *Selles süsteemis leidub vektor, mis avaldub ülejäänud vektorite lineaarkombinatsioonina.*

TÕESTUS. 1. $\Rightarrow$ 2. Olgu süsteem $a_1, \ldots, a_s$ lineaarselt sõltuv. Siis leiduvad $k_1, \ldots, k_s \in K$, mis ei ole kõik nullid, nii et

$$k_1 a_1 + \ldots + k_s a_s = 0.$$

Olgu k_l viimane nullist erinev kordaja eelmise võrduse vasakul poolel. Siis

$$k_1 a_1 + \ldots + k_l a_l = 0$$

Paneme tähele, et $l \neq 1$. Tõepoolest, kui $l = 1$, siis $k_1 a_1 = 0$, millest elemendiga k_1^{-1} korrutades saame võrduse $a_1 = 0$. Viimane on vastuolus eeldusega. Niisiis $l > 1$. Kasutades arvutusreegleid vektorruumis saame avaldada

$$a_l = -k_l^{-1} k_1 a_1 - \ldots - k_l^{-1} k_{l-1} a_{l-1},$$

mis tähendab, et a_l on talle eelnevate vektorite $a_1, \ldots, a_{l-1}$ lineaarkombinatsioon.

2. $\Rightarrow$ 3. Ilmne.

3. $\Rightarrow$ 1. Oletame, et vektor a_i avaldub ülejäänud vektorite lineaarkombinatsioonina:

$$a_i = l_1 a_1 + \ldots + l_{i-1} a_{i-1} + l_{i+1} a_{i+1} + \ldots + l_s a_s,$$

kus $l_1, \ldots, l_{i-1}, l_{i+1}, \ldots, l_s \in K$. Siis

$$l_1 a_1 + \ldots + l_{i-1} a_{i-1} + (-1) a_i + l_{i+1} a_{i+1} + \ldots + l_s a_s = 0,$$

kusjuures lineaarkombinatsioon viimase võrduse vasakul poolel on mittetriviaalne, sest a_i kordaja ei ole 0. Järelikult on süsteem $a_1, \ldots, a_s$ lineaarselt sõltuv. □

Järeldus 6.23. *Nullist erinevate vektorite süsteemi $a_1, \ldots, a_s$, kus $s \geqslant 2$, jaoks on järgmised väited samaväärsed.*

1. *See süsteem on lineaarselt sõltumatu.*
2. *Selle süsteemi ükski vektor ei avaldu eelnevate vektorite lineaarkombinatsioonina.*
3. *Selle süsteemi ükski vektor ei avaldu ülejäänud vektorite lineaarkombinatsioonina.*

Näide 6.24. Tänu järeldusele 6.23 on näiteks selge, et vektorruumi $\mathbb{R}^4$ vektorite süsteem

$$\begin{aligned} a_1 &= (3,0,0,0), \\ a_2 &= (5,2,1,0), \\ a_3 &= (0,4,0,2) \end{aligned}$$

on lineaarselt sõltumatu. Tõepoolest, vektorit a_2 ei saa esitada vektori a_1 kordsena ja vektorit a_3 ei saa esitada a_1 ja a_2 lineaarkombinatsioonina.

Järeldus 6.25. *Kahest vektorist koosnev süsteem on lineaarselt sõltuv parajasti siis, kui üks vektor on teisest saadav skalaariga korrutades.*

Järeldus 6.26. *Iga vektorite süsteem, mis sisaldab kahte võrdset vektorit, on lineaarselt sõltuv.*

Lause 6.27. *Kui vektorite süsteemi mingi alamsüsteem on lineaarselt sõltuv, siis ka terve süsteem on lineaarselt sõltuv.*

TÕESTUS. Kuna vektorite järjekorra muutmine ei muuda süsteemi lineaarset sõltuvust/sõltumatust, siis võime eeldada, et süsteemi $a_1, \ldots, a_s$ lineaarselt sõltuv alamsüsteem koosneb selle süsteemi t ($t \leqslant s$) esimesest vektorist $a_1, \ldots, a_t$. Siis leiduvad $k_1, \ldots, k_t \in K$, millest vähemalt üks on nullist erinev, nii et

$$k_1 a_1 + \ldots + k_t a_t = 0.$$

Siis kehtib ka võrdus

$$k_1 a_1 + \ldots + k_t a_t + 0 a_{t+1} + \ldots + 0 a_s = 0,$$

kusjuures viimase võrduse vasakul poolel on mittetriviaalne lineaarkombinatsioon vektoritest $a_1, \ldots, a_s$. Järelikult on süsteem $a_1, \ldots, a_s$ lineaarselt sõltuv. □

Lause 6.28. *Lineaarselt sõltumatu vektorite süsteemi iga alamsüsteem on ka lineaarselt sõltumatu.*

TÕESTUS. Olgu süsteem $a_1, \ldots, a_s$ lineaarselt sõltumatu. Kui selle süsteemi mingi alamsüsteem $a_{i_1}, \ldots, a_{i_t}$ oleks lineaarselt sõltuv, siis lause 6.27 põhjal oleks ka $a_1, \ldots, a_s$ lineaarselt sõltuv, mis on vastuolus eeldusega. Seega on kõik alamsüsteemid lineaarselt sõltumatud. □

6.4. Moodustajate süsteem

Definitsioon 6.29. Vektorruumi V vektorite süsteemi M nimetatakse **moodustajate süsteemiks** ehk **tekitajate süsteemiks**, kui vektorruumi V iga vektor avaldub süsteemi M kuuluvate vektorite lineaarkombinatsioonina.

Enamasti on vektorruumil palju moodustajate süsteeme, mõned neist suuremad, mõned väiksemad. Eriti kasulikud on moodustajate süsteemid, mille kaudu iga vektori saab avaldada täpselt ühel viisil. Neid me uurida soovimegi.

Lepime kokku, et **edasises vaatleme selle kursuse jooksul ainult selliseid vektorruume, millel on olemas lõplik moodustajate süsteem.** Vastavalt definitsioonidele võime öelda, et süsteem $a_1, \ldots, a_s$ on vektorruumi V moodustajate süsteem parajasti siis, kui V on selle süsteemi lineaarne kate:

$$V = L(a_1, \ldots, a_s).$$

Alustuseks tõestame ühe abitulemuse.

Lemma 6.30. *Olgu $a_1, \ldots, a_s$ vektorruumi V moodustajate süsteem. Kui selle süsteemi mingi vektor avaldub ülejäänud vektorite lineaarkombinatsioonina, siis selle vektori väljajätmisel süstemist $a_1, \ldots, a_s$ saame jällegi V moodustajate süsteemi.*

TÕESTUS. Oletame, et vektor a_1 avaldub lineaarkombinatsioonina

$$a_1 = k_2 a_2 + \ldots + k_s a_s,$$

kus $k_2, \ldots, k_s \in K$. (Kui ülejäänute kaudu avaldub mõni teine vektor, on tõestus analoogiline.) Olgu nüüd $a \in V$ suvaline vektor. Et $a_1, \ldots, a_s$ on moodustajate süsteem, siis leiduvad sellised $l_1, \ldots, l_s \in K$, et

$$a = l_1 a_1 + l_2 a_2 + \ldots + l_s a_s.$$

Asendades viimasesse võrdusse a_1 avaldise ja kasutades vektorruumi omadusi saame, et

$$a = l_1(k_2 a_2 + \ldots + k_s a_s) + l_2 a_2 + \ldots + l_s a_s = (l_1 k_2 + l_2) a_2 + \ldots + (l_1 k_s + l_s) a_s,$$

s.t. a avaldub vektorite $a_2, \ldots, a_s$ lineaarkombinatsioonina. Seega $a_2, \ldots, a_s$ on ka moodustajate süsteem. □

Kuna süsteem $a_1, \ldots, a_n$ on lineaarse katte $L(a_1, \ldots, a_n)$ moodustajate süsteem, siis saame eelmisest lemmast teha järgmise järelduse.

Järeldus 6.31. *Kui vektor a_i, kus $i \in \{1, \ldots, n\}$, avaldub süsteemi $a_1, \ldots, a_n$ ülejäänud vektorite lineaarkombinatsioonina, siis*

$$L(a_1, \ldots, a_n) = L(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n).$$

Näiteks $L(a, b, a, c, b, b) = L(a, b, c)$ mistahes vektorite $a, b, c \in V$ korral.

6.5. Vektorruumi baas

Definitsioon 6.32. Vektorruumi **baas** on selle vektorruumi lineaarselt sõltumatu moodustajate süsteem.

Kuna baasis ei saa ükski vektor esineda kaks korda (muidu oleks süsteem järelduse 6.26 tõttu lineaarselt sõltuv), siis baasivektorid moodustavad hulga. Baasidest rääkides kasutamegi vastavat tähistust, nt. kirjutame

$$e = \{e_1, \ldots, e_n\},$$

kui räägime baasist e, mis koosneb vektoritest $e_1, \ldots, e_n$. Nii nagu kõigi vektorite süsteemide korral on ka baasi puhul vektorite järjekord oluline ja me loeme, et e_1 on baasi e esimene vektor, e_2 on teine vektor jne.

Näide 6.33. 1. Vektorruumi K^n vektorite süsteem

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0, 0), \\ e_2 &= (0, 1, 0, \ldots, 0, 0), \\ &\ldots \\ e_n &= (0, 0, 0, \ldots, 0, 1) \end{aligned}$$

on baas. Tõepoolest, kui oletame, et

$$\begin{aligned} (0, 0, \ldots, 0) &= k_1 e_1 + k_2 e_2 + \ldots + k_n e_n \\ &= k_1(1, 0, \ldots, 0) + k_2(0, 1, \ldots, 0) + \ldots + k_n(0, 0, \ldots, 1) \\ &= (k_1, 0, \ldots, 0) + (0, k_2, \ldots, 0) + \ldots + (0, 0, \ldots, k_n) \\ &= (k_1, k_2, \ldots, k_n), \end{aligned}$$

siis $k_1 = k_2 = \ldots = k_n = 0$, mis tähendab, et süsteem $e_1, e_2, \ldots, e_n$ on lineaarselt sõltumatu. Samuti, iga vektor $(l_1, l_2, \ldots, l_n) \in K^n$ on avaldatav lineaarkombinatsioonina

$$(l_1, l_2, \ldots, l_n) = l_1 e_1 + l_2 e_2 + \ldots + l_n e_n$$

ja seega $e_1, e_2, \ldots, e_n$ on moodustajate süsteem.

2. Eelneva põhjal teame, et $e_1 = (1,0), e_2 = (0,1)$ on baas vektorruumis $\mathbb{R}^2$. Toome näite ühest teistsugusest $\mathbb{R}^2$ baasist. Vaatleme näiteks vektorite süsteemi

$$\begin{aligned} f_1 &= (1,1), \\ f_2 &= (0,1). \end{aligned}$$

Kui $(0,0) = k_1 f_1 + k_2 f_2 = (k_1, k_1) + (0, k_2) = (k_1, k_1 + k_2)$, siis $k_1 = 0$ ja $k_1 + k_2 = 0$, millest saame, et ka $k_2 = 0$. Seega f_1, f_2 on lineaarselt sõltumatu süsteem. Mistahes vektori $(l_1, l_2) \in \mathbb{R}^2$ saame avaldada kujul

$$(l_1, l_2) = l_1 f_1 + (l_2 - l_1) f_2,$$

s.t. f_1, f_2 on moodustajate süsteem. Järelikult f_1, f_2 on baas.

3. Vektorruumis $\mathbb{E}_2$ moodustavad baasi mistahes 2 mittekollineaarset vektorit.

4. Vektorruumis $\mathbb{E}_3$ moodustavad baasi mistahes 3 mittekomplanaarset vektorit.

Teoreem 6.34. *Kui vektorruumis on vähemalt kaks vektorit, siis on selles vektorruumis olemas baas.*

Tõestus. Olgu $a_1, \ldots, a_s$ vektorruumi V moodustajate süsteem. (Me oleme eeldanud, et vaadeldavates vektorruumides leidub lõplik moodustajate süsteem.) Et vektorruumis on vähemalt kaks vektorit, siis leidub selles vähemalt üks nullist erinev vektor x. Kui kõik vektorid $a_1, \ldots, a_s$ oleks nullvektorid, siis me ei saaks vektorit x selle süsteemi kaudu avaldada ja see ei oleks moodustajate süsteem. Seega on selles süsteemis vähemalt üks nullist erinev vektor a_i. Kui süsteemis $a_1, \ldots, a_s$ sisaldub nullvektor, siis avaldub see ülejäänud vektorite lineaarkombinatsioonina ja nullvektorit välja jättes saame lemma 6.30 põhjal ikkagi V moodustajate süsteemi. Seega võime üldisust kitsendamata eeldada, et süsteemis $a_1, \ldots, a_s$ ei ole nullvektoreid.

Konstrueerime nüüd süsteemi $a_1, \ldots, a_s$ põhjal uue süsteemi (selle süsteemi alamsüsteemi), mis oleks V baasiks. Selleks jätame vaadeldavast süsteemist $a_1, \ldots, a_s$ välja n.ö. mittevajalikud vektorid. Täpsemalt toimime järgmiselt. Vaatleme vektorit a_2. Kui see vektor avaldub talle eelneva vektori a_1 lineaarkombinatsioonina, siis jätame selle süsteemist välja. Lemma 6.30 tõttu on saadav süsteem ka moodustajate süsteem. Kui a_2 ei avaldu vektori a_1 lineaarkombinatsioonina, siis me teda välja ei jäta. Edasi vaatleme saadud süsteemis vektorit a_3. Kui a_3 avaldub talle eelnevate vektorite lineaarkombinatsioonina, siis jätame selle vektori süsteemist välja. Pärast a_3 vaatlemist on meil ikkagi tegemist moodustajate süsteemiga. Niimoodi jätkates jõuame lõpuks viimase vektorini. Vastavalt konstruktsioonile on lõpuks saadud süsteem $S = \{a_1, a_{i_2}, \ldots, a_{i_n}\}$ (selles süsteemis ei esine ükski vektor kaks korda!) vektorruumi V moodustajate süsteem. Konstruktsiooni põhjal ei avaldu süsteemi S ükski vektor eelnevate vektorite lineaarkombinatsioonina, mis järelduse 6.23 põhjal tähendab seda, et S on lineaarselt sõltumatu. Kokkuvõttes võime öelda, et S on baas. □

Märkus 6.35. Kui vektorruum koosneb ainult ühest vektorist, siis see vektor on nullvektor. Sellises vektorruumis ei ole baasi, sest ei ole lineaarselt sõltumatuid vektorite süsteeme.

Lause 6.36. *Iga lõpliku mittetühja lineaarselt sõltumatu vektorite süsteemi saab täiendada vektorruumi baasiks.*

Tõestus. Olgu $a_1, \ldots, a_k$ ($k \in \mathbb{N}$) lineaarselt sõltumatu vektorite süsteem vektorruumis V ja olgu $e_1, \ldots, e_n$ selle vektorruumi baas. Siis süsteem

$$a_1, \ldots, a_k, e_1, \ldots, e_n$$

on moodustajate süsteem vektorruumis V. Eraldame sellest välja baasi kasutades teoreemi 6.34 tõestuses kirjeldatud meetodit. Siis saadav baas sisaldab kindlasti vektoreid $a_1, \ldots, a_k$, sest ükski neist ei avaldu eelnevate vektorite lineaarkombinatsioonina. □

Meie järgmiseks eesmärgiks on anda veel kaks kirjeldust baasidele.

Definitsioon 6.37. Vektorruumi V vektorite süsteemi S nimetatakse **maksimaalseks lineaarselt sõltumatuks süsteemiks**, kui see süsteem on lineaarselt sõltumatu, aga iga süsteem, mis saadakse süsteemist S ühe vektori lisamisel, on lineaarselt sõltuv.

Definitsioon 6.38. Vektorruumi V vektorite süsteemi S nimetatakse **minimaalseks moodustajate süsteemiks**, kui see süsteem on moodustajate süsteem, aga iga süsteem, mis saadakse süsteemist S ühe vektori väljajätmisel, ei ole moodustajate süsteem.

Teoreem 6.39. *Vektorruumi V (üle korpuse K) lõpliku vektorite süsteemi S korral on järgmised väited samaväärsed.*

1. *S on baas.*
2. *S on maksimaalne lineaarselt sõltumatu süsteem.*
3. *S on minimaalne moodustajate süsteem.*

Tõestus. 1. $\Rightarrow$ 2. Olgu süsteem $e_1, \ldots, e_n$ baas, s.t. lineaarselt sõltumatu moodustajate süsteem. Kui võtame suvalise vektori $a \in V$, siis a avaldub süsteemi $e_1, \ldots, e_n$ lineaarkombinatsioonina. Seega süsteem $e_1, \ldots, e_n, a$ on lause 6.22 põhjal lineaarselt sõltuv. Järelikult $e_1, \ldots, e_n$ on maksimaalne lineaarselt sõltumatu süsteem.

2. $\Rightarrow$ 1. Olgu $e_1, \ldots, e_n$ maksimaalne lineaarselt sõltumatu süsteem. Peame näitama, et see on moodustajate süsteem. Selleks võtame suvalise $a \in V$ ja vaatleme süsteemi

$$e_1, \ldots, e_n, a.$$

Eelduse põhjal on see süsteem lineaarselt sõltuv. Järelikult leidub selles süsteemis vektor, mis avaldub eelnevate lineaarkombinatsioonina. Kuna see ei saa olla ükski vektoritest $e_1, \ldots, e_n$ (muidu oleks süsteem $e_1, \ldots, e_n$ lineaarselt sõltuv), siis see vektor peab olema a. Sellega oleme näidanud, et $e_1, \ldots, e_n$ on moodustajate süsteem.

1. $\Rightarrow$ 3. Olgu süsteem $e_1, \ldots, e_n$ baas. Siis $e_1, \ldots, e_n$ on moodustajate süsteem. Näitame, et süsteem $e_2, \ldots, e_n$ ei ole moodustajate süsteem. (Kui süsteemist jätta välja mõni teine vektor, on tõestus analoogiline.) Kui $e_2, \ldots, e_n$ oleks moodustajate süsteem, siis peaks e_1 avalduma $e_2, \ldots, e_n$ kaudu, mis on vastuolus $e_1, \ldots, e_n$ lineaarse sõltumatusega (vt. järeldust 6.23).

3. $\Rightarrow$ 1. Olgu $e_1, \ldots, e_n$ minimaalne moodustajate süsteem. Peame näitama, et see süsteem on lineaarselt sõltumatu. Kui näiteks $e_1 = k_2e_2 + \ldots + k_ne_n$, siis avaldub iga $a \in V$ kujul

$$a = l_1e_1 + l_2e_2 + \ldots + l_ne_n = l_1(k_2e_2 + \ldots + k_ne_n) + l_2e_2 + \ldots + l_ne_n \in L(e_2, \ldots, e_n),$$

mis tähendab, et ka $e_2, \ldots, e_n$ on moodustajate süsteem. Samamoodi saame vastuolu eeldusega, kui mõni teine vektor süsteemis $e_1, \ldots, e_n$ avaldub ülejäänute kaudu. Järelikult $e_1, \ldots, e_n$ on lineaarselt sõltumatu järelduse 6.23 põhjal. □

Tuleb välja, et vektorite arv baasis ei sõltu baasi valikust.

Teoreem 6.40. *Vektorruumi mistahes kahes baasis on sama palju vektoreid.*

TÕESTUS. Olgu $A = \{a_1, \ldots, a_m\}$ ja $B = \{b_1, \ldots, b_n\}$ vektorruumi V kaks baasi. Oletame vastuväiteliselt, et $m < n$. Vaatleme süsteemi

$$b_1, a_1, a_2, \ldots, a_m. \tag{T_1}$$

Et b_1 avaldub vektorite $a_1, \ldots, a_m$ lineaarkombinatsioonina (A on moodustajate süsteem,) siis on süsteem (T_1) lineaarselt sõltuv. Seega lause 6.22 põhjal peab süsteemi (T_1) mingi vektor avalduma eelnevate lineaarkombinatsioonina. See peab olema üks vektoritest $a_1, \ldots, a_m$, sest vektorile b_1 ei eelne ühtegi vektorit. Eeldame, et see vektor on a_r Lemma 6.30 põhjal on

$$b_1, a_1, \ldots, a_{r-1}, a_{r+1}, \ldots, a_m \tag{U_1}$$

ka vektorruumi V moodustajate süsteem. Süsteemi U_1 võib vaadelda kui süsteemi, mis on saadud süsteemist A ühe vektori asendamisel süsteemi B esimese vektoriga. Oletame, et me oleme selliseid asendusi tehes jõudnud moodustajate süsteemini

$$b_k, b_{k-1}, \ldots, b_1, a_{i_1}, a_{i_2}, \ldots, a_{i_{m-k}} \tag{U_k}$$

($k \in \{1, \ldots, m-1\}$), s.t. oleme süsteemi A k vektorit asendanud vektoritega $b_1, \ldots, b_k$. Siis muuhulgas ka vektor b_{k+1} avaldub süsteemi U_k kaudu, mis tähendab, et süsteem

$$b_{k+1}, b_k, b_{k-1}, \ldots, b_1, a_{i_1}, a_{i_2}, \ldots, a_{i_{m-k}} \tag{T_{k+1}}$$

on lineaarselt sõltuv. Järelikult peab üks selle süsteemi vektoritest avalduma eelnevate lineaarkombinatsioonina. See vektor ei saa kuuluda hulka $\{b_{k+1}, b_k, b_{k-1}, \ldots, b_1\}$, sest muidu peaks süsteem B olema lineaarselt sõltuv. Seega peab eelnevate lineaarkombinatsioonina avalduma üks vektoritest $a_{i_1}, a_{i_2}, \ldots, a_{i_{m-k}}$. Selle vektori väljajätmisel jääb järele V moodustajate süsteem, mille kohta võib öelda, et ta on saadud süsteemist A $k+1$ vektori asendamisel süsteemi B $k+1$ vektoriga. Niiviisi jätkates ja vektoreid a_j välja vahetades jõuame m-ndal sammul moodustajate süsteemini

$$b_m, b_{m-1}, \ldots, b_1. \tag{U_m}$$

Siis aga süsteem $B = \{b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\}$ ei ole minimaalne moodustajate süsteem, mis on vastuolus teoreemiga 6.39. Järelikult m ei saa olla väiksem kui n. Analoogiliselt saab näidata, et n ei ole väiksem kui m ja seega $m = n$. □

Fakt, et baasivektorite arv ei sõltu baasi valikust lubab meil anda järgmise väga olulise definitsiooni.

Definitsioon 6.41. Vektorruumi **mõõtmeks** ehk **dimensiooniks** nimetatakse vektorite arvu selle vektorruumi mingis baasis. Ainult nullvektorist koosneva vektorruumi mõõtmeks loetakse arv 0. Vektorruumi V mõõdet tähistatakse järgmiselt: $\dim(V)$.

Näide 6.42. 1. Kui K on korpus ja $n \in \mathbb{N}$, siis vektorruumi K^n mõõde on n, sest üheks baasiks selles vektorruumis on näites 6.33 toodud baas $e_1, \ldots, e_n$.

2. Vektorruumi $\mathrm{Mat}_{m,n}(K)$ mõõde on mn, sest baasivektoreiks võib võtta kõik maatriksid, milles on ühel kohal 1 ja ülejäänud kohtadel 0-d.

Lause 6.43. *n-mõõtmelises vektorruumis on iga n lineaarselt sõltumatust vektorist koosnev süsteem baas.*

TÕESTUS. Olgu $a_1, \ldots, a_n$ lineaarselt sõltumatu vektorite süsteem n-mõõtmelises vektorruumis V. Lausest 6.36 teame, et iga lineaarselt sõltumatu vektorite süsteemi saab täiendada baasiks. Kuna aga kõigis baasides on n vektorit, siis peabki süsteem $a_1, \ldots, a_n$ olema juba baas. □

6.6. Vektori koordinaadid

Ilmselt on lugeja kooliprogrammist tuttav analüütilise geomeetria algetega, kus kesksel kohal on vektori koordinaatide mõiste. Tuleb välja, et koordinaatidest saab rääkida mitte ainult geomeetrilistes vektorruumides $\mathbb{E}_2$ ja $\mathbb{E}_3$ vaid suvalistes mittetriviaalsetes vektorruumides.

Lause 6.44. *Vektorruumi $V \neq \{0\}$ iga vektor on üheselt avaldatav baasivektorite lineaarkombinatsioonina.*

Tõestus. Olgu V vektorruum üle korpuse K ja olgu $e = \{e_1, e_2, \ldots, e_n\}$ selle vektorruumi baas. Kuna e on moodustajate süsteem, siis saab iga vektori avaldada vektorite $e_1, \ldots, e_n$ lineaarkombinatsioonina. Oletame nüüd, et vektor $a \in V$ on avaldatav kahel viisil:

$$a = a_1e_1 + a_2e_2 + \ldots + a_ne_n = b_1e_1 + b_2e_2 + \ldots + b_ne_n,$$

kus $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$. Siis

$$(a_1 - b_1)e_1 + (a_2 - b_2)e_2 + \ldots + (a_n - b_n)e_n = 0,$$

millest lineaarse sõltumatuse tõttu saame, et $a_1 - b_1 = a_2 - b_2 = \ldots = a_n - b_n = 0$ ehk

$$a_1 = b_1,\ a_2 = b_2,\ \ldots,\ a_n = b_n.$$

□

Tõestatud lause lubab defineerida vektori koordinaadid baasi suhtes.

Definitsioon 6.45. Olgu V vektorruum üle korpuse K, olgu $e = \{e_1, e_2, \ldots, e_n\}$ selle vektorruumi baas ja $a \in V$. Kui $a = a_1e_1 + a_2e_2 + \ldots + a_ne_n$, siis skalaare $a_1, a_2, \ldots, a_n \in K$ nimetatakse vektori a **koordinaatideks** baasi e suhtes.

Lause 6.46. *Olgu V vektorruum üle korpuse K ja olgu $e = \{e_1, e_2, \ldots, e_n\}$ selle vektorruumi baas. Kui vektori a koordinaadid baasi e suhtes on $a_1, \ldots, a_n$ ja vektori b koordinaadid baasi e suhtes on $b_1, \ldots, b_n$, siis vektori $a + b$ koordinaadid baasi e suhtes on $a_1 + b_1, \ldots, a_n + b_n$ ja vektori ka (kus $k \in K$) koordinaadid baasi e suhtes on $ka_1, \ldots, ka_n$.*

Tõestus. Kui $a = a_1e_1 + a_2e_2 + \ldots + a_ne_n$ ja $b = b_1e_1 + b_2e_2 + \ldots + b_ne_n$, siis

$$a + b = (a_1e_1 + \ldots + a_ne_n) + (b_1e_1 + \ldots + b_ne_n) = (a_1 + b_1)e_1 + \ldots + (a_n + b_n)e_n$$

ja

$$ka = k(a_1e_1 + \ldots + a_ne_n) = (ka_1)e_1 + \ldots + (ka_n)e_n.$$

□

7. Astak

Selles peatükis tutvume astaku mõistega ja selle arvutamise meetoditega. Astakust saab rääkida nii maatriksite kui ka suvaliste vektorite süsteemide korral.

7.1. Vektorite süsteemi astak

Definitsioon 7.1. Vektorite süsteemi **astakuks** nimetatakse selle vektorite süsteemi lineaarse katte mõõdet.

Süsteemi $a_1, \ldots, a_s$ astakut tähistatakse sümboliga $\operatorname{rank}(a_1, \ldots, a_s)$. Definitsiooni põhjal

$$\boxed{\operatorname{rank}(a_1, \ldots, a_s) = \dim(L(a_1, \ldots, a_s)).}$$

Lause 7.2. *Vektorite süsteemi astak on võrdne vektorite arvuga selle süsteemi mistahes maksimaalses lineaarselt sõltumatus alamsüsteemis.*

Tõestus. Vaatleme vektorruumi V vektorite süsteemi $a_1, \ldots, a_s$. Olgu selle süsteemi astak r, s.t. $r = \dim(L(a_1, \ldots, a_s))$. Vaatleme süsteemi $a_1, \ldots, a_s$ mingit maksimaalset lineaarselt sõltumatut alamsüsteemi. Vajaduse korral vektoreid ümber nummerdades võime eeldada, et see süsteem koosneb vektoritest $a_1, \ldots, a_t$, kus $t \leqslant s$. Peame näitama, et $t = r$.

Kuna $a_1, \ldots, a_t$ on maksimaalne lineaarselt sõltumatu alamsüsteem, siis lisades süsteemile $a_1, \ldots, a_t$ vektori a_i, kus $i \in \{t+1, \ldots, s\}$, saame lineaarselt sõltuva süsteemi, milles mingi vektor peab avalduma eelnevate lineaarkombinatsioonina. See saab olla vaid a_i. Seega vektorid $a_{t+1}, \ldots, a_s$ avalduvad vektorite $a_1, \ldots, a_t$ kaudu, mis järelduse 6.31 põhjal tähendab, et $L(a_1, \ldots, a_t, a_{t+1}, \ldots, a_s) = L(a_1, \ldots, a_t)$. Seega

$$r = \dim(L(a_1, \ldots, a_s)) = \dim(L(a_1, \ldots, a_t)) = t,$$

sest $a_1, \ldots, a_t$ on $L(a_1, \ldots, a_t)$ lineaarselt sõltumatu moodustajate süsteem ehk baas. □

Näide 7.3. Kui vektorid a, b, c on lineaarselt sõltumatud, siis a, b, c on süsteemi $a, b, b, c, 2a+c, a$ maksimaalne lineaarselt sõltumatu alamsüsteem ja $\operatorname{rank}(a, b, b, c, 2a + c, a) = 3$.

Definitsioon 7.4. Vektorruumi V kahte vektorite süsteemi nimetatakse **ekvivalentseteks**, kui nende süsteemide lineaarsed katted on võrdsed.

Seega süsteemid $a_1, \ldots, a_s$ ja $b_1, \ldots, b_t$ on ekvivalentsed, kui

$$L(a_1, \ldots, a_s) = L(b_1, \ldots, b_t).$$

On lihtne aru saada, et ekvivalentsuse seos vektorruumi V vektorite süsteemide vahel on refleksiivne, sümmeetriline ja transitiivne.

Nii nagu maatriksi ridade puhulgi (vt. definitsiooni 5.7), võib suvalise vektorite süsteemi korral vaadelda järgmisi teisendusi.

Definitsioon 7.5. Vektorite süsteemi **elementaarteisendused** on järgmised teisendused:

1. süsteemi kahe vektori äravahetamine;
2. süsteemi vektori korrutamine nullist erineva skalaariga;

3. süsteemi mingile vektorile mingi skalaariga korrutatud teise vektori liitmine.

Lause 7.6. *Kui vektorruumi V vektorite süsteem T on saadud süsteemist S elementaarteisenduste abil, siis süsteemid S ja T on ekvivalentsed.*

TÕESTUS. Tänu ekvivalentsusseose transitiivsusele piisab, kui vaatleme süsteemi T, mis on saadud süsteemist S ühe elementaarteisenduse abil.

1. On selge, et vektorite järjekorra muutmine ei muuda süsteemi lineaarset katet.

2. Teist tüüpi teisenduste korral jätame tõestuse läbimõtlemiseks lugejale.

3. Näitame, et mistahes skalaari k, vektorite $a_1, \ldots, a_s$ ja indeksite $i, j \in \{1, \ldots, s\}$, $i < j$ korral

$$L(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_s) = L(a_1, \ldots, a_i + ka_j, \ldots, a_j, \ldots, a_s). \tag{20}$$

Kuna mistahes skalaaride $k_1, \ldots, k_s$ ja $l_1, \ldots, l_s$ korral

$$\begin{aligned}
&k_1a_1 + \ldots + k_ia_i + \ldots + k_ja_j + \ldots + k_sa_s \\
&= k_1a_1 + \ldots + k_i(a_i + ka_j) + \ldots + (k_j - kk_i)a_j + \ldots + k_sa_s, \\
&l_1a_1 + \ldots + l_i(a_i + ka_j) + \ldots + l_ja_j + \ldots + l_sa_s \\
&= l_1a_1 + \ldots + l_ia_i + \ldots + (l_ik + l_j)a_j + \ldots + l_sa_s,
\end{aligned}$$

siis võrduse (20) vasakul poolel olev hulk sisaldub paremal poolel olevas hulgas ja vastupidi. □

7.2. Maatriksi astak

Maatriksi korral võib rääkida tema rea- ja veeruvektoritest.

Definitsioon 7.7. Maatriksi $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$ i-ndaks **reavektoriks** nimetatakse vektorit

$$A_i := (a_{i1}, a_{i2}, \ldots, a_{in}) \in K^n$$

ja i-ndaks **veeruvektoriks** nimetatakse vektorit

$$(a_{1i}, a_{2i}, \ldots, a_{mi}) \in K^m.$$

Seega põhimõtteliselt võiks maatriksit $A \in \mathrm{Mat}_{m,n}(K)$ vaadelda kui hulga $(K^n)^m$ elementi kirjutades selle ridade kaupa kujul

$$A = ((a_{11}, a_{12}, \ldots, a_{1n}), (a_{21}, a_{22}, \ldots, a_{2n}), \ldots, (a_{m1}, a_{m2}, \ldots, a_{mn})) = (A_1, A_2, \ldots, A_n),$$

või kui hulga $(K^m)^n$ elementi, kui kirjutame A veergude kaupa. Esimest viisi maatriksi esitamiseks kasutatakse mitmetes arvutialgebra süsteemides (nt. Maple, Mathematica).

Definitsioon 7.8. Maatriksi astakuks nimetatakse selle maatriksi reavektorite süsteemi astakut. Maatriksi A astakut tähistatakse sümboliga $\mathrm{rank}(A)$.

Formaalselt võib kirjutada, et

$$\mathrm{rank}(A) = \dim(L(A_1, \ldots, A_m)).$$

Tänu lausele 7.2 on maatriksi A astak võrdne lineaarselt sõltumatute reavektorite maksimaalse arvuga. Teiste sõnadega: *maatriksi A astak on r siis ja ainult siis, kui*

1. *sellel maatriksil leidub r lineaarselt sõltumatut reavektorit,*

2. *neile r vektorile mistahes reavektori lisamisel saame lineaarselt sõltuva süsteemi.*

Näide 7.9. Maatriksi

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & 1 \\ 0 & 2 & 1 & 3 & -1 \\ 2 & 2 & 0 & -2 & 2 \end{pmatrix}$$

astak on 2, sest reavektorid A_2 ja A_3 on lineaarselt sõltumatud, aga süsteemid A_1, A_2, A_3 ja A_2, A_3, A_4 on lineaarselt sõltuvad (Miks?).

Kehtib järgmine oluline teoreem (mille tõestust me käesolevas kursuses ei anna).

Teoreem 7.10 (Teoreem maatriksi astakust). *Maatriksi astak on võrdne selle maatriksi nullist erinevate miinorite kõrgeima järguga.*

Niisiis: *maatriksi astak on r siis ja ainult siis, kui*

1. *selles maatriksis leidub r-ndat järku nullist erinev miinor,*

2. *kõik suuremat järku miinorid on võrdsed nulliga.*

Kuna alamruutmaatriksi B determinant on võrdne maatriksi B^T determinandiga, siis maatriksi A transponeerimisel tema nullist erinevate miinorite kõrgeim järk ei muutu. Seega saame teoreemist 7.10 järgmise järelduse.

Järeldus 7.11. *Maatriksi astak ei muutu transponeerimisel, s.t.*

$$\boxed{\operatorname{rank}(A) = \operatorname{rank}(A^T).}$$

Järeldus 7.12. *Maatriksi astak on võrdne selle maatriksi veeruvektorite süsteemi astakuga.*

Tõestus. Kuna $\operatorname{rank}(A) = \operatorname{rank}(A^T)$, siis $\operatorname{rank}(A)$ on võrdne maatriksi A^T reavektorite süsteemi astakuga. Viimane aga on võrdne maatriksi A veeruvektorite süsteemi astakuga. □

Edasises läheb meil vaja ka järgmisi tulemusi.

Lause 7.13. *Ruutmaatriksi reavektorid on lineaarselt sõltuvad parajasti siis, kui selle maatriksi determinant on* 0.

Tõestus. Tarvilikkus. Kui maatriksi A reavektorid on lineaarselt sõltuvad, siis leidub nende hulgas mingi reavektor, mis avaldub eelnevate lineaarkombinatsioonina. Olgu näiteks $A_i = k_1A_1 + \ldots + k_{i-1}A_{i-1}$. Liites sellele reale sobiva skalaariga korrutatud eelnevad read saame i-nda rea muuta nullreaks ja seega on $|A| = 0$.
Piisavus. Kui $A \in \operatorname{Mat}_n(K)$ ja $|A| = 0$, siis selles maatriksis ei leidu n-ndat järku nullist erinevat miinorit. Seega teoreemi 7.10 põhjal $\operatorname{rank}(A) < n$, mis tähendab, et maatriksi A reavektorite süsteemi astak on väiksem kui n. Seega reavektorite hulgas ei saa olla n lineaarselt sõltumatut vektorit. Järelikult A reavektorite süsteem on lineaarselt sõltuv. □

Lause 7.14. *Maatriksite korrutise astak ei ületa tegurite astakuid.*

Tõestus. Olgu $A = (a_{ij}) \in \mathrm{Mat}_{mn}(K)$ ja $B = (b_{ij}) \in \mathrm{Mat}_{np}(K)$. Me tõestame, et

$$\mathrm{rank}(AB) \leqslant \mathrm{rank}(A) \quad \text{ja} \quad \mathrm{rank}(AB) \leqslant \mathrm{rank}(B).$$

Olgu $AB =: C = (c_{ij}) \in \mathrm{Mat}_{mp}(K)$. Siis maatriksi C i-s reavektor avaldub kujul

$$\begin{aligned} C_i &= (c_{i1}, \ldots, c_{ip}) = (a_{i1}b_{11} + \ldots + a_{in}b_{n1}, \ldots, a_{i1}b_{1p} + \ldots + a_{in}b_{np}) \\ &= (a_{i1}b_{11}, \ldots, a_{i1}b_{1p}) + \ldots + (a_{in}b_{n1}, \ldots, a_{in}b_{np}) = a_{i1}B_1 + \ldots + a_{in}B_n \\ &\in L(B_1, \ldots, B_n). \end{aligned}$$

Järelikult $L(C_1, \ldots, C_m) \subseteq L(B_1, \ldots, B_n)$ ja

$$\mathrm{rank}(AB) = \mathrm{rank}(C) = \dim(L(C_1, \ldots, C_m)) \leqslant \dim(L(B_1, \ldots, B_n)) = \mathrm{rank}(B).$$

Lisaks sellele

$$\mathrm{rank}(AB) = \mathrm{rank}((AB)^T) = \mathrm{rank}(B^T A^T) \leqslant \mathrm{rank}(A^T) = \mathrm{rank}(A).$$

□

7.3. Astaku arvutamisest

Põhiliseks meetodiks maatriksi astaku leidmisel on **elementaarteisenduste meetod**. Selle meetodi puhul tehakse elementaarteisendusi maatriksi ridade ja veergudega, et muuta maatriks "lihtsamaks" (selle lihtsuse all mõeldakse enamasti seda, et maatriksis oleks võimalikult palju nulle). Harilikult on eesmärgiks viia maatriks nn. astmelisele kujule.

Definitsioon 7.15. Öeldakse, et maatriks on **astmelisel kujul**, kui

1. nullidest koosnevad read on nullist erinevaid elemente sisaldavatest ridadest allpool;
2. iga $i \geqslant 2$ korral i-nda rea esimene nullist erinev element (kui see leidub) on kaugemal (s.t. tema veeruindeks on suurem), kui $(i-1)$-se rea esimene nullist erinev element.

Niisiis maatriks on astmelisel kujul, kui tal on kuju

$$\begin{pmatrix} 0 & \ldots & 0 & a_{1j_1} & \ldots & a_{1,j_2-1} & a_{1j_2} & \ldots & a_{1,j_3-1} & a_{1j_3} & \ldots & \ldots & \ldots & \ldots & a_{1n} \\ 0 & \ldots & 0 & 0 & \ldots & 0 & a_{2j_2} & \ldots & a_{2,j_3-1} & a_{2j_3} & \ldots & \ldots & \ldots & \ldots & a_{2n} \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & a_{3j_3} & \ldots & \ldots & \ldots & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & a_{rj_r} & \ldots & a_{rn} \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}, \tag{21}$$

kus

$$1 \leqslant j_1 < j_2 < j_3 < \ldots < j_r \leqslant n$$

ja elemendid $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ on nullist erinevad. Sellise kuju puhul ütleme, et astmelises kujus on r astet ja et astmekohad on $(1, j_1), (2, j_2), \ldots, (r, j_r)$.

Näide 7.16. Maatriks

$$\begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

on astmelisel kujul, kusjuures astmeid on 3 ja astmekohad on $(1,1)$, $(2,2)$ ja $(3,4)$.

Lause 7.17. *Iga maatriksi saab ridade elementaarteisenduste abil viia astmelisele kujule.*

TÕESTUS. Otsime maatriksis A üles esimese veeru, milles leidub nullist erinevaid elemente. Olgu selle veeru number j_1. Ridade järjekorra vahetamisega võime saavutada olukorra, kus j_1-se veeru esimene element on nullist erinev. Selle elemendi abil saab kolmandat tüüpi teisenduste abil muuta nulliks kõik ülejäänud elemendid j_1-ses veerus. Leiame nüüd järgmise veeru, mis sisaldab nullist erinevaid elemente esimesest reast allpool. Olgu selle veeru number j_2. Ridade vahetamisega võime saavutada, et kohal $(2, j_2)$ on nullist erinev element b. Muudame liitmisteisendusega nulliks kõik elemendid b all. Jätkame samas vaimus kuni jõuame viimase veeruni. □

Lause 7.18. *Maatriksi astak on võrdne astmete arvuga selle maatriksi astmelises kujus.*

TÕESTUS. Kuna lause 7.6 tõttu elementaarteisendused ei muuda maatriksi astakut, siis on maatriksi ja tema astmelise kuju astakud võrdsed. Kui maatriksi astmeline kuju on (21), siis selles maatriksis ridades $1, 2, \ldots, r$ ja veergudes $j_1, j_2, \ldots, j_r$ on nullist erinev r-ndat järku miinor

$$\begin{vmatrix} a_{1j_1} & a_{1j_2} & \ldots & a_{1j_r} \\ 0 & a_{2j_2} & \ldots & a_{2j_r} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{rj_r} \end{vmatrix} = a_{1j_1}a_{2j_2}\ldots a_{rj_r}.$$

Samas kõik suuremat järku miinorid on võrdsed nulliga, sest nad sisaldavad vähemalt ühe nullidest koosneva rea. Teoreemi 7.10 tõttu on antud maatriksi astak r, mis ühtlasi on võrdne astmete arvuga astmelises kujus. □

Teine võimalus maatriksi astaku arvutamiseks on kasutada niinimetatud **miinorite ääristamise meetodit**. Öeldakse, et maatriksi A miinor M' **ääristab** maatriksi A miinorit M, kui M' on saadud miinorist M ühe rea ja ühe veeru lisamisel. Meetod põhineb järgmisel faktil (mida me siinkohal ei tõesta).

Lause 7.19. *Kui M on maatriksi A r-ndat järku nullist erinev miinor ja kõik miinorit M ääristavad miinorid on võrdsed nulliga, siis* $\operatorname{rank}(A) = r$.

8. Lineaarvõrrandisüsteemid

Selles peatükis uurime lineaarvõrrandisüsteemide lahendamist. Otsime vastust järgmistele küsimustele.

- Millal on lineaarvõrrandisüsteem lahenduv?
- Kui palju on lineaarvõrrandisüsteemil lahendeid?
- Kuidas neid lahendeid leida?

Üldiste lineaarvõrrandisüsteemide kõrval uurime ka teatud erikujulisi süsteeme, näiteks homogeenseid süsteeme.

8.1. Ülesande püstitus

Definitsioon 8.1. Lineaarvõrrandisüsteem üle korpuse K on võrrandisüsteem kujul

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\quad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{22}$$

kus $x_1, \ldots, x_n$ on tundmatud ja $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in K$. Elemente $a_{11}, \ldots, a_{mn} \in K$ nimetatakse selle **võrrandisüsteemi kordajateks** ja elemente $b_1, \ldots, b_m \in K$ nimetatakse **vabaliikmeteks**. Märgime, et nii võrrandite arv m, kui tundmatute arv n võivad olla suvalised naturaalarvud.

Definitsioon 8.2. Võrrandisüsteemi (22) **lahendiks** (ehk **erilahendiks**) nimetatakse hulga K^n iga elementi $(k_1, \ldots, k_n)$, mille korral

$$a_{i1}k_1 + a_{i2}k_2 + \ldots + a_{in}k_n = b_i,$$

kus $i = 1, \ldots, m$.

Seega $(k_1, \ldots, k_n)$ on süsteemi (22) lahend, kui iga i korral asendades $k_1, \ldots, k_n$ süsteemi i-ndas võrrandis tundmatute $x_1, \ldots, x_n$ asemele ja arvutades välja vastava K elemendi saame tulemuseks b_i.

Lineaarvõrrandisüsteemi **lahendamise** all peetakse silmas selle süsteemi kõigi lahendite leidmist.

Märkus 8.3. Lineaarvõrrandisüsteemide lahendamisel ei paku meile huvi sellised süsteemid, kus mingi tundmatu x_i kordaja kõigis võrrandites on null. Tõepoolest, kui me oskaksime lahendada süsteeme, kus selliseid tundmatuid ei ole, siis oskaksime vajaduse korral lahendada ka süsteeme, kus selliseid tundmatuid on. Näiteks vaatleme süsteemi, mis koosneb ainult ühest võrrandist $x + y + 0 \cdot z = 0$. Leiame kahe tundmatuga süsteemi $x + y = 0$ kõik lahendid. Nendeks on paarid $(k, -k)$, kus $k \in K$. Siis süsteemi $x + y + 0 \cdot z = 0$ lahenditeks on kõik kolmikud $(k, -k, l)$, kus $k, l \in K$.

Seega **eeldame edaspidises, et kõigis vaadeldavates süsteemides on iga tundmatu vähemalt ühes võrrandis nullist erineva kordajaga**.

Definitsioon 8.4. Lineaarvõrrandisüsteemi nimetatakse

- **mittelahenduvaks** ehk **vasturääkivaks**, kui tal ei ole ühtegi lahendit,
- **lahenduvaks** ehk **kooskõlaliseks**, kui tal on vähemalt üks lahend,
- **üheselt lahenduvaks** ehk **määratuks**, kui tal on täpselt üks lahend.

Definitsioon 8.5. Maatriksit

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

nimetatakse **lineaarvõrrandisüsteemi** (22) **maatriksiks**. Maatriksit

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

nimetatakse **lineaarvõrrandisüsteemi** (22) **laiendatud maatriksiks**. Üheveerulist maatriksit

$$\overline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

nimetatakse **lineaarvõrrandisüsteemi** (22) **tundmatute veeruks** ja üheveerulist maatriksit

$$\overline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}$$

nimetatakse **lineaarvõrrandisüsteemi** (22) **vabaliikmete veeruks**.

Kuna süsteemi (22) kõigi lahendite $k = (k_1, \dots, k_n) \in K^n$ leidmine on samaväärne kõigi selliste maatriksite

$$\overline{k} = \begin{pmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{pmatrix} \in \mathrm{Mat}_{n,1}(K)$$

leidmisega, mille korral

$$A\overline{k} = \overline{b},$$

siis räägitakse ka süsteemi (22) **maatrikskujust**

$$A\overline{x} = \overline{b}.$$

8.2. Gaussi meetod

Selles paragrahvis anname meetodi suvalise lineaarvõrrandisüsteemi lahendamiseks.

Definitsioon 8.6. Kahte n tundmatuga lineaarvõrrandisüsteemi üle korpuse K nimetatakse **ekvivalentseteks**, kui neil on ühed ja samad lahendid.

On selge, et lineaarvõrrandisüsteemide ekvivalentsuse seos on refleksiivne, sümmeetriline ja transitiivne.

Lause 8.7. *Kui lineaarvõrrandisüsteemi laiendatud maatriksi reavektorite süsteemiga teha elementaarteisendusi või jätta sellest välja nullvektorid, siis tulemuseks saadavale maatriksile vastav lineaarvõrrandisüsteem on ekvivalentne esialgsega.*

Tõestus. Meenutame (vt. definitsiooni 5.7), et elementaarteisendusi maatriksi ridadega on kolme tüüpi ja nende tegemine on samaväärne maatriksi korrutamisega vasakult elementaarmaatriksitega (lause 5.10). Seega peame näitama, et kui lineaarvõrrandisüsteemi laiendatud maatriks on $A' \in \mathrm{Mat}_{m,n+1}(K)$ ja me korrutame selle vasakult elementaarmaatriksiga C, siis maatriksitele A' ja CA' vastavad lineaarvõrrandisüsteemid on ekvivalentsed. Kui esialgse lineaarvõrrandisüsteemi maatriks on A ja vabaliikmete veerg $\overline{b}$, siis selle süsteemi maatrikskuju on

$$A\overline{x} = \overline{b}. \tag{23}$$

Pärast elementaarteisendust saadava süsteemi maatrikskuju on

$$(CA)\overline{x} = C\overline{b}. \tag{24}$$

Kui nüüd $\overline{k} \in \mathrm{Mat}_{n,1}(K)$ on süsteemi (23) lahend, siis $A\overline{k} = \overline{b}$. Korrutades selle võrduse mõlemad pooled maatriksiga C saame $CA\overline{k} = C\overline{b}$, mis tähendab, et $\overline{k}$ on ka süsteemi (24) lahend. Vastupidi, kui $\overline{k} \in \mathrm{Mat}_{n,1}(K)$ on süsteemi (24) lahend, siis $CA\overline{k} = C\overline{b}$. Kuna kõik elementaarmaatriksid on pööratavad, siis võime viimase võrduse pooli korrutada vasakult maatriksiga C^{-1}. See annab meile võrduse $A\overline{k} = \overline{b}$, kust näeme, et $\overline{k}$ on süsteemi (23) lahend. Seega on neil kahel süsteemil täpselt samad lahendid.

On ka selge, et kui süsteemist jätta välja võrrand $0 \cdot x_1 + \ldots + 0 \cdot x_n = 0$ (ehk laiendatud maatriksist jätta välja nullidest koosnev rida), siis saadaval süsteemil on samad lahendid, mis esialgsel. □

Tõestatud lause lubab lineaarvõrrandisüsteemi lihtsustada nii, et lahendite hulk selle käigus ei muutu.

Järgnevalt leiame tarviliku ja piisava tingimuse selleks, et lineaarvõrrandisüsteem oleks lahenduv. Osutub, et selle üle saab otsustada astakute põhjal.

Teoreem 8.8 (Kroneckeri-Capelli teoreem).[5] *Lineaarvõrrandisüsteem on lahenduv parajasti siis, kui selle süsteemi maatriksi astak on võrdne selle süsteemi laiendatud maatriksi astakuga.*

Tõestus. Vaatleme lineaarvõrrandisüsteemi, mille maatriks on A ja laiendatud maatriks on A'. Teeme maatriksi A' ridadega selliseid elementaarteisendusi, mis viivad maatriksi A astmelisele kujule. See on võimalik tänu lausele 7.17. Tulemuseks on maatriks kujul

[5] Leopold Kronecker (1823–1891) — saksa matemaatik,
Alfredo Capelli (1855–1910) — itaalia matemaatik.

$$\left(\begin{array}{cccccccccccc}
a_{11} & \dots & a_{1,j_2-1} & a_{1j_2} & \dots & a_{1,j_{3-1}} & a_{1j_3} & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\
0 & \dots & 0 & a_{2j_2} & \dots & a_{2,j_{3-1}} & a_{2j_3} & \dots & \dots & \dots & \dots & a_{2n} & b_2 \\
0 & \dots & 0 & 0 & \dots & 0 & a_{3j_3} & \dots & \dots & \dots & \dots & a_{3n} & b_3 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & a_{rj_r} & \dots & a_{rn} & b_r \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & b_{r+1} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & b_m
\end{array}\right), \qquad (25)$$

kus

$$1 < j_2 < j_3 < \ldots < j_r \leqslant n$$

ja elemendid $a_{11}, a_{2j_2}, \ldots, a_{rj_r}$ on nullist erinevad. Juhime tähelepanu, et siin $a_{11} \neq 0$ tänu meie kokkuleppele, et x_1 kordaja mingis võrrandis peab olema nullist erinev. Lause 8.7 põhjal on esialgne lineaarvõrrandisüsteem ekvivalentne maatriksile (25) vastava lineaarvõrrandisüsteemiga. Muuhulgas on nad samaaegselt kas lahenduvad või mittelahenduvad. Seega piisab kui vaatleme maatriksile (25) vastava süsteemi lahenduvust.

Astmete arvu järgi näeme, et süsteemi maatriksi astak on r (vt. lauset 7.18). Oletame, et laiendatud maatriksi astak ei võrdu süsteemi maatriksi astakuga r. Siis laiendatud maatriksi astak peab olema suurem kui r, s.t. mõni elementidest $b_{r+1}, \ldots, b_m$ peab olema nullist erinev. Siis on meil süsteemis võrrand $0 \cdot x_1 + \ldots + 0 \cdot x_n = b_i$, kus $b_i \neq 0$. Sellisel võrrandil ei ole ühtegi lahendit ja seega süsteem ei ole lahenduv. Järelikult kui süsteem on lahenduv, siis $b_{r+1} = \ldots = b_m = 0$ ja süsteemi maatriksi astak on võrdne selle süsteemi laiendatud maatriksi astakuga.

Tõestame nüüd vastupidise implikatsiooni. Olgu süsteemi maatriksi astak võrdne selle süsteemi laiendatud maatriksi astakuga, s.t. $b_{r+1} = \ldots = b_m = 0$. Näitame, et sellel süsteemil leidub lahend $(k_1, \ldots, k_n) \in K^n$, kus $k_i = 0$ iga $i \in \{1, 2, \ldots, n\} \setminus \{1, j_2, \ldots, j_r\}$ korral. Sellise lahendi saame leida, kui oskame ära lahendada süsteemi

$$\begin{cases}
a_{11}x_1 + a_{1j_2}x_{j_2} + \ldots + a_{1j_{r-1}}x_{j_{r-1}} + a_{1j_r}x_{j_r} = b_1 \\
a_{2j_2}x_{j_2} + \ldots + a_{2j_{r-1}}x_{j_{r-1}} + a_{2j_r}x_{j_r} = b_2 \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \ldots \\
a_{r-1,j_{r-1}}x_{j_{r-1}} + a_{r-1,j_r}x_{j_r} = b_{r-1}. \\
a_{rj_r}x_{j_r} = b_r.
\end{cases}$$

Viimane süsteem on aga tõesti lahenduv. Viimasest võrrandist arvutame $x_{j_r} = a_{rj_r}^{-1} b_r$. Siis asendame saadud väärtuse eelviimasesse võrrandisse ja arvutame välja $x_{j_{r-1}}$ väärtuse jne. □

Näitame nüüd kuidas saab leida lahenduva süsteemi lahendeid. Oletame, et meil on tegemist lahenduva süsteemiga, mille laiendatud maatriks on viidud astmelisele kujule (25), kus $b_{r+1} = \ldots = b_m = 0$. Jätame sellest maatriksist välja nullidest koosnevad read. Lause 8.7 tõttu ei muuda see võrrandisüsteemi lahendite hulka. Niisiis vaatleme lineaarvõrrandisüsteemi laiendatud maatriksiga

$$\left(\begin{array}{cccccccccccc}
a_{11} & \dots & a_{1,j_2-1} & a_{1j_2} & \dots & a_{1,j_{3-1}} & a_{1j_3} & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\
0 & \dots & 0 & a_{2j_2} & \dots & a_{2,j_{3-1}} & a_{2j_3} & \dots & \dots & \dots & \dots & a_{2n} & b_2 \\
0 & \dots & 0 & 0 & \dots & 0 & a_{3j_3} & \dots & \dots & \dots & \dots & a_{3n} & b_3 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & a_{rj_r} & \dots & a_{rn} & b_r
\end{array}\right), \qquad (26)$$

On kaks võimalust.

1) $r = n$. See tähendab, et $j_2 = 2$, $j_3 = 3$, ..., $j_n = n$ ning maatriksile (26) vastav süsteem on kujul

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{22}x_2 + \ldots + a_{2,n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \ldots \\ + a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1}. \\ a_{nn}x_n = b_n. \end{cases}$$

Viimasest võrrandist näeme, et $x_n = a_{nn}^{-1}b_n$. Asendades selle eelviimasesse võrrandisse saame välja arvutada x_{n-1} väärtuse jne. Niisiis sellisel juhul on süsteemil täpselt üks lahend ehk *süsteem on üheselt lahenduv.*

2) $r < n$. Osutub, et sellisel juhul on süsteemil rohkem kui üks lahend (lõpmatu K korral on neid lõpmata palju). Niisuguses olukorras kõigi lahendite esitamiseks jagatakse süsteemi tundmatud kahte rühma: vabadeks ja sõltuvateks tundmatuteks. Nimelt tundmatuid $x_1, x_{j_2}, \ldots, x_{j_r}$ (need on tundmatud, mis vastavad n.ö. astmekohtadele maatriksis (26)) nimetatakse **sõltuvateks tundmatuteks** ja kõiki ülejäänud tundmatuid **vabadeks tundmatuteks**. Paneme tähele, et *sõltuvate tundmatute arv r on võrdne lineaarvõrrandisüsteemi maatriksi astakuga ja vabade tundmatute arv on $n-r$, s.t. kõigi tundmatute arvu ja lineaarvõrrandisüsteemi maatriksi astaku vahe.*

Kuna x_{j_i} ($i \in \{2, \ldots, r\}$) kordaja a_{ij_i} on nullist erinev, siis liites sobiva skalaariga korrutatud i-ndat rida eelnevatele ridadele on j_i-ndas veerus võimalik kõik ülejäänud elemendid nulliks muuta. Viies seejärel vabade tundmatutega liikmed paremale poole võrdusmärki ja korrutades iga võrrandi mõlemaid pooli sobiva skalaariga saame sõltuvad tundmatud avaldada vabade tundmatute ja vabaliikmete kaudu:

$$\begin{cases} x_1 = c_1 + d_{1,r+1}x_{j_{r+1}} + \ldots + d_{1n}x_{j_n} \\ x_{j_2} = c_2 + d_{2,r+1}x_{j_{r+1}} + \ldots + d_{2n}x_{j_n} \\ \quad \ldots \\ x_{j_r} = c_r + d_{r,r+1}x_{j_{r+1}} + \ldots + d_{rn}x_{j_n}, \end{cases} \tag{27}$$

kus $c_1, \ldots, c_r$ on vabaliikmed ja $\{j_{r+1}, \ldots, j_n\} = \{1, 2, \ldots, n\} \setminus \{1, j_2, \ldots, j_r\}$, s.t. $x_{j_{r+1}}, \ldots, x_{j_n}$ on vabad tundmatud. Kui anda nüüd vabadele tundmatutele (n.ö. vabalt) mistahes väärtused $k_{r+1}, \ldots, k_n \in K$ siis saame seoste (27) abil välja arvutada sõltuvate tundmatute $x_1, x_{j_2}, \ldots, x_{j_r}$ väärtused ja seega saame kätte ühe vaadeldava süsteemi lahendi. Teisest küljest on selge, et sellisel viisil saame me kätte kõik vaadeldava süsteemi lahendid, sest kõik lahendid peavad rahuldama seoseid (27). Kui lineaarvõrrandisüsteemi lahendid on antud seoste (27) abil, siis öeldakse, et tegemist on selle süsteemi **üldlahendiga vabade tundmatute kaudu**. Kirjeldatud meetodit lineaarvõrrandisüsteemi lahendamiseks kutsutakse **Gaussi**[6] **meetodiks**.

Märkus 8.9. Maatriksit võib astmelisele kujule viia mitmel erineval viisil ja sellest tulenevalt võib ka saada mitmeid erinevaid sõltuvate (ja vabade) tundmatute komplekte. Kui lahenduva lineaarvõrrandisüsteemi maatriksi astak on r ja meil on selles maatriksi mingid r lineaarselt sõltumatut rida (sellised tingimata leiduvad), siis valides nendes ridades mingi r-ndat järku nullist erineva miinori saame selle miinori teisendada diagonaalkujule ja seega tundmatud, mille kordajate veergudest see miinor on moodustatud, võib võtta sõltuvateks tundmatuteks, sest neid on võimalik ülejäänud tundmatute kaudu avaldada.

[6] Carl Friedrich Gauss (1777–1855) — saksa matemaatik.

Näide 8.10. Lahendame lineaarvõrrandisüsteemi

$$\begin{cases} 2x_1 - 3x_2 + 5x_3 + 7x_4 = 1 \\ 4x_1 - 6x_2 + 2x_3 + 3x_4 = 2 \\ 2x_1 - 3x_2 - 11x_3 - 15x_4 = 1 \end{cases} \tag{28}$$

(üle $\mathbb{R}$) Gaussi meetodil.

Moodustame võrrandisüsteemi laiendatud maatriksi ja teisendame selle astmelisele kujule jättes ära nullidest koosnevad read:

$$\left(\begin{array}{cccc|c} 2 & -3 & 5 & 7 & 1 \\ 4 & -6 & 2 & 3 & 2 \\ 2 & -3 & -11 & -15 & 1 \end{array}\right)\begin{matrix} \\ -2\mathrm{I} \\ -\mathrm{I} \end{matrix} \to \left(\begin{array}{cccc|c} 2 & -3 & 5 & 7 & 1 \\ 0 & 0 & -8 & -11 & 0 \\ 0 & 0 & -16 & -22 & 0 \end{array}\right)\begin{matrix} \\ \\ -2\mathrm{II} \end{matrix} \to$$

$$\left(\begin{array}{cccc|c} 2 & -3 & 5 & 7 & 1 \\ 0 & 0 & -8 & -11 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)\cdot\left(-\tfrac{1}{8}\right) \to \left(\begin{array}{cccc|c} 2 & -3 & 5 & 7 & 1 \\ 0 & 0 & 1 & \frac{11}{8} & 0 \end{array}\right)-5\mathrm{II} \to$$

$$\left(\begin{array}{cccc|c} 2 & -3 & 0 & \frac{1}{8} & 1 \\ 0 & 0 & 1 & \frac{11}{8} & 0 \end{array}\right).$$

Näeme, et nii süsteemi maatriksi kui ka laiendatud maatriksi astak on 2 ja seega süsteem on lahenduv. Sõltuvaid tundmatuid on 2 (sest astak on 2) ja vabu tundmatuid on $4 - 2 = 2$. Astmekohtade järgi võime sõltuvateks tundmatuteks valida x_1 ja x_3, seega vabadeks tundmatuteks jäävad x_2 ja x_4. Avaldades viimast maatriksit kasutades sõltuvad tundmatud vabade kaudu, saame üldlahendi vabade tundmatute kaudu:

$$\begin{cases} x_1 = \frac{1}{2} + \frac{3}{2}x_2 - \frac{1}{16}x_4 \\ x_3 = -\frac{11}{8}x_4 \end{cases}.$$

Seda tuleb tõlgendada nii, et andes vabadele tundmatutele x_2 ja x_4 kõikvõimalikud reaalarvulised väärtused saame välja arvutada vastavad x_1 ja x_3 väärtused ja niimoodi kätte kõik esialgse süsteemi lahendid. See tähendab, et süsteemi kõigi lahendite hulk on

$$L = \left\{ \left(\frac{1}{2} + \frac{3}{2}k - \frac{1}{16}l\ ,\ k\ ,\ -\frac{11}{8}l\ ,\ l \right) \middle|\ k, l \in \mathbb{R} \right\} \subseteq \mathbb{R}^4.$$

Näiteks võttes $k = l = 0$ saame süsteemi (28) üheks erilahendiks $d_* = (\frac{1}{2}, 0, 0, 0)$.

8.3. Crameri peajuht

Vaatleme nüüd lineaarvõrrandisüsteemide lahendamist ühel tähtsal erijuhul.

Definitsioon 8.11. Öeldakse, et lineaarvõrrandisüsteemi puhul on tegemist **Crameri**[7] **peajuhuga**, kui

1. võrrandite arv on võrdne tundmatute arvuga ja
2. süsteemi maatriks on regulaarne.

[7] Gabriel Cramer (1704–1752) — šveitsi matemaatik.

Seega Crameri peajuhuga on tegemist siis, kui lineaarvõrrandisüsteem on kujul

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{cases} \tag{29}$$

kus $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ on regulaarne (s.t. $|A| \neq 0$).

Lause 8.12. *Kui lineaarvõrrandisüsteemi puhul on tegemist Crameri peajuhuga, siis on sellel süsteemil täpselt üks lahend.*

Tõestus. Meenutame, et vektor $k = (k_1, \ldots, k_n) \in K^n$ on süsteemi (29) lahend parajasti siis, kui üheveeruline maatriks

$$\overline{k} = \begin{pmatrix} k_1 \\ \cdots \\ k_n \end{pmatrix} \in \mathrm{Mat}_{n,1}(K)$$

rahuldab võrdust

$$A\overline{k} = \overline{b}.$$

Et A on regulaarne, siis leidub pöördmaatriks A^{-1}. Järelikult

$$A(A^{-1}\overline{b}) = (AA^{-1})\overline{b} = E\overline{b} = \overline{b},$$

mis tähendab, et süsteemil leidub vähemalt üks lahend (selle komponentideks on üheveerulise maatriksi $A^{-1}\overline{b}$ elemendid). Teisest küljest, kui $k, l \in K^n$ on süsteemi lahendid, siis $A\overline{k} = \overline{b} = A\overline{l}$. Korrutades võrduse $A\overline{k} = A\overline{l}$ mõlemaid pooli vasakult maatriksiga A^{-1} saame võrduse $\overline{k} = \overline{l}$, millest järeldub, et $k = l$. Seega süsteemil on ülimalt üks lahend ja kokkuvõttes peab tal olema täpselt üks lahend. □

Tuleb välja, et Crameri peajuhu korral saab süsteemi lahendi leida determinantide abil.

Nagu nägime lause 8.12 tõestuses peab süsteemi (29) lahendi $k = (k_1, \ldots, k_n) \in K^n$ korral kehtima võrdus

$$\begin{pmatrix} k_1 \\ \cdots \\ k_n \end{pmatrix} = A^{-1}\overline{b}.$$

Teoreemi 5.17 põhjal $A^{-1} = |A|^{-1}(A_{ji})$ (see on maatriks, mille i-ndas reas ja j-ndas veerus on element $|A|^{-1}A_{ji}$, kus A_{ji} on maatriksi A elemendi a_{ji} algebraline täiend). Vastavalt maatriksite korrutamise ja maatriksite võrduse definitsioonile võime kirjutada, et

$$k_i = \sum_{j=1}^{n} \left(|A|^{-1}A_{ji}\right) b_j = |A|^{-1} \cdot \sum_{j=1}^{n} A_{ji}b_j$$

iga $i \in \{1, \ldots, n\}$ korral. Olgu $D = |A|$ ja olgu D_i ($i \in \{1, \ldots, n\}$) sellise maatriksi determinant, mis on saadud maatriksist A selle i-nda veeru asendamisel süsteemi (29) vabaliikmete veeruga, s.t.

$$D_i = \begin{vmatrix} a_{11} & \ldots & a_{1,i-1} & b_1 & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & b_2 & a_{2,i+1} & \ldots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,i-1} & b_n & a_{n,i+1} & \ldots & a_{nn} \end{vmatrix}$$

Arendades determinanti D_i i-nda veeru järgi saame summa $\sum_{j=1}^{n} A_{ji}b_j$. Järelikult

$$k_i = D^{-1}D_i$$

iga $i \in \{1, \ldots, n\}$ korral. Sellega oleme tõestanud järgmise tulemuse.

Teoreem 8.13. *Kui lineaarvõrrandisüsteemi (29) puhul on tegemist Crameri peajuhuga, siis selle süsteemi lahendi $k = (k_1, \ldots, k_n) \in K^n$ i-s komponent $k_i = D^{-1}D_i$, kus D on selle süsteemi maatriksi determinant ja D_i on sellise maatriksi determinant, mis on saadud süsteemi maatriksist selle i-nda veeru asendamisel süsteemi vabaliikmete veeruga.*

Märkus 8.14. Kui on tegemist Crameri peajuhuga, siis $D \neq 0$ ja arvestades märkust 1.36 võib lahendi leidmise valemid esitada kujul

$$\boxed{k_i = \frac{D_i}{D}, \quad i = 1, \ldots, n.}$$

Neid valemeid kutsutakse **Crameri valemiteks**.

8.4. Homogeenne lineaarvõrrandisüsteem

Definitsioon 8.15. Lineaarvõrrandisüsteemi nimetatakse **homogeenseks**, kui selle süsteemi kõik vabaliikmed on nullid.

Niisiis homogeensel lineaarvõrrandisüsteemil on kuju

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\quad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \tag{30}$$

Maatrikskujul võib homogeense lineaarvõrrandisüsteemi esitada järgmiselt:

$$A\overline{x} = \overline{0},$$

kus

$$\overline{0} = \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix} \in \mathrm{Mat}_{m,1}(K).$$

On selge, et nullvektor $(0, \ldots, 0) \in K^n$ on alati homogeense lineaarvõrrandisüsteemi lahend, seega kõik homogeensed lineaarvõrrandisüsteemid on lahenduvad.

Kui mittehomogeense lineaarvõrrandisüsteemi kõigi lahendite hulk on lihtsalt hulga K^n alamhulk, siis homogeensete süsteemide korral võib öelda enamat.

Lause 8.16. *n tundmatuga homogeense lineaarvõrrandisüsteemi (üle korpuse K) kõigi lahendite hulk on alamruum vektorruumis K^n.*

Tõestus. Olgu

$$L = \{k \in K^n \mid A\overline{k} = \overline{0}\}$$

süsteemi (30) kõigi lahendite hulk. Nagu mainitud, see hulk ei ole tühi. Olgu $k, l \in L$ ja $c \in K$. Siis

$$A(\overline{k+l}) = A(\overline{k} + \overline{l}) = A\overline{k} + A\overline{l} = \overline{0} + \overline{0} = \overline{0}$$

ja

$$A(c\overline{k}) = c(A\overline{k}) = c\overline{0} = \overline{0},$$

mis tähendab, et ka $k + l, ck \in L$. □

Vektorruumi alamruum on ise ka vektorruum (vt. lauset 6.9). See lubab anda järgmise definitsiooni.

Definitsioon 8.17. Homogeense lineaarvõrrandisüsteemi lahendite fundamentaalsüsteemiks nimetatakse selle süsteemi lahendite alamruumi baasi.

Teoreem 8.18. *Kui homogeenses lineaarvõrrandisüsteemis on n tundmatut ja selle süsteemi maatriksi astak on r, siis selle süsteemi lahendite fundamentaalsüsteemis on $n - r$ lahendit.*

Ka selle teoreemi tõestust me käesolevas kursuses anda ei jõua.

Näide 8.19. Lahendades homogeense lineaarvõrrandisüsteemi

$$\begin{cases} 2x_1 - 3x_2 + 5x_3 + 7x_4 = 0 \\ 4x_1 - 6x_2 + 2x_3 + 3x_4 = 0 \\ 2x_1 - 3x_2 - 11x_3 - 15x_4 = 0 \end{cases} \tag{31}$$

analoogiliselt sellega, kuidas lahendasime süsteemi (28) saame üldlahendiks vabade tundmatute kaudu

$$\begin{cases} x_1 = \frac{3}{2}x_2 - \frac{1}{16}x_4 \\ x_3 = -\frac{11}{8}x_4 \end{cases}. \tag{32}$$

Selle süsteemi lahendite fundamentaalsüsteemi saame kui anname vabadele tundmatutele x_2 ja x_4 kaks komplekti väärtusi mingi regulaarse teist järku ruutmaatriksi, nt. $\begin{pmatrix} 2 & 0 \\ 0 & 16 \end{pmatrix}$, ridadest ning arvutame kummalgi juhul võrduste (32) abil sõltuvate tundmatute x_1 ja x_3 väärtused:

$$\begin{aligned} d_1 &= (3, \mathbf{2}, 0, \mathbf{0}), \\ d_2 &= (-1, \mathbf{0}, -22, \mathbf{16}). \end{aligned}$$

Süsteemi (31) kõik lahendid avalduvad kujul $t_1 d_1 + t_2 d_2$, kus $t_1, t_2 \in \mathbb{R}$, s.t. et kõigi lahendite hulk on

$$L_h = \{t_1 d_1 + t_2 d_2 \mid t_1, t_2 \in \mathbb{R}\}.$$

8.5. Mittehomogeenne lineaarvõrrandisüsteem

Vaatleme üldist lineaarvõrrandisüsteemi

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{33}$$

Süsteemile (33) vastavaks homogeenseks lineaarvõrrandisüsteemiks nimetatakse süsteemi

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases} \tag{34}$$

mis on saadud süsteemist (33) kõigi vabaliikmete asendamisel nullidega. Meenutame, et maatrikskujul võime need kaks süsteemi kirja panna järgmiselt: $A\overline{x} = \overline{b}$ ja $A\overline{x} = \overline{0}$. Osutub, et nende kahe süsteemi lahendid on omavahel tihedalt seotud.

Kui U on vektorruumi V alamruum ja $v \in V$, siis tähistatakse

$$v + U := \{v + u \mid u \in U\}.$$

Tähistame süsteemi (33) kõigi lahendite hulka tähega L ja süsteemi (34) kõigi lahendite hulka sümboliga L_h.

Teoreem 8.20. *Süsteemide (33) ja (34) lahendihulkade L ja L_h vahel kehtib iga $d_* \in L$ korral seos*

$$\boxed{L = d_* + L_h.}$$

Tõestus. Olgu $d_* \in L$ süsteemi (33) suvaline lahend.

Näitame, et $d_* + L_h \subseteq L$. Kui $l \in L_h$, siis

$$A(\overline{d_* + l}) = A(\overline{d_*} + \overline{l}) = A\overline{d_*} + A\overline{l} = \overline{b} + \overline{0} = \overline{b},$$

mis tähendab, et $d_* + l \in L$. Seega $d_* + L_h \subseteq L$.

Võtame nüüd vektori $k \in L$. Siis

$$A(\overline{k - d_*}) = A(\overline{k} - \overline{d_*}) = A\overline{k} - A\overline{d_*} = \overline{b} - \overline{b} = \overline{0}$$

ehk $k - d_* \in L_h$. Järelikult $k = d_* + (k - d_*) \in d_* + L_h$, millega oleme näidanud, et $L \subseteq d_* + L_h$. Kokkuvõttes $L = d_* + L_h$. □

Märkus 8.21. Teoreem 8.20 sõnastatakse tihti kujul: *mittehomogeense lineaarvõrrandisüsteemi üldlahend on võrdne selle süsteemi mingi erilahendi ja vastava homogeense lineaarvõrrandisüsteemi üldlahendi summaga.*

Näide 8.22. Vaatleme jälle lineaarvõrrandisüsteemi

$$\begin{cases} 2x_1 - 3x_2 + 5x_3 + 7x_4 = 1 \\ 4x_1 - 6x_2 + 2x_3 + 3x_4 = 2 \\ 2x_1 - 3x_2 - 11x_3 - 15x_4 = 1. \end{cases} \tag{35}$$

Sellele süsteemile vastava homogeense lineaarvõrrandisüsteemi (31) lahendite fundamentaalsüsteem on $d_1 = (3, 2, 0, 0), d_2 = (-1, 0, -22, 16)$. Nagu nägime näites 8.10 on süsteemi (35) üheks erilahendiks $d_* = (\frac{1}{2}, 0, 0, 0)$. Seega võime öelda, et süsteemi (35) kõigi lahendite hulk on

$$L = d_* + L_h = \{d_* + t_1 d_1 + t_2 d_2 \mid t_1, t_2 \in \mathbb{R}\}.$$

Teiste sõnadega: süsteemi (35) lahendid on kujul $d_* + t_1 d_1 + t_2 d_2$, kus $t_1, t_2 \in \mathbb{R}$.

9. Polünoomid

Lugeja on kindlasti tuttav polünoomidega, mille kordajateks on reaalarvud. Käesolevas peatükis vaatleme polünoome üle ringide. Alustuseks konstrueerime polünoomide ringi.

9.1. Polünoomide ring

Olgu R ring. Me hakkame uurima polünoome üle ringi R. Polünoomidest rääkides eeldame kõikjal, et **ringis R on vähemalt 2 elementi.** Seda eeldust ei hakka me iga kord eraldi välja tooma. Meenutame (vt. märkust 1.33), et sellises ringis on ühikelement ja nullelement erinevad, $1 \neq 0$. Kui $R = \{0\}$ on ühe-elemendiline ring, siis üle sellise ringi on vaid üks polünoom ja selline olukord ei paku meile huvi.

Vaatleme kõigi selliste jadade hulka, mille komponendid kuuluvad ringi R ja mille komponendid on mingist kohast alates kõik võrdsed ringi R nullelemendiga. Tähistame selle hulga sümboliga $R[X]$. Seega

$$R[X] = \bigcup_{n \in \mathbb{N} \cup \{0\}} \{(a_0, a_1, \ldots, a_n, 0, 0, \ldots) \mid a_0, a_1, \ldots, a_n \in R\}.$$

Kaks hulka $R[X]$ kuuluvat jada on võrdsed parajasti siis, kui nende vastavad komponendid on võrdsed.

Definitsioon 9.1. Hulka $R[X]$ kuuluvate jadade $a = (a_0, a_1, a_2, \ldots)$ ja $b = (b_0, b_1, b_2, \ldots)$ **summa** defineeritakse võrdusega

$$a + b := c = (c_0, c_1, c_2, \ldots),$$

kus

$$c_k = a_k + b_k$$

iga $k \in \mathbb{N} \cup \{0\}$ korral, ja **korrutis** defineeritakse võrdusega

$$ab := d = (d_0, d_1, d_2, \ldots),$$

kus

$$d_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_{k-1} b_1 + a_k b_0 = \sum_{i+j=k} a_i b_j$$

iga $k \in \mathbb{N} \cup \{0\}$ korral.

Teoreem 9.2. *Hulk $R[X]$ on eelpool defineeritud tehete suhtes ring. Kui R on kommutatiivne, siis ka ring $R[X]$ on kommutatiivne.*

Tõestus. Lihtne on aru saada, et kui jadade a ja b komponendid on mingist kohast alates nullid, siis on seda ka jadade $a + b$ ja ab komponendid. Seega on liitmine ja korrutamine algebralised tehted hulgal $R[X]$.

Jadade liitmine on assotsiatiivne ja kommutatiivne tänu sellele, et liitmine on defineeritud komponenthaaval ja ringi R elementide liitmine on assotsiatiivne ja kommutatiivne. Nullelemendiks on jada, mille kõik komponendid on nullid. Jada $a = (a_0, a_1, a_2, \ldots)$ vastandelemendiks on jada $-a = (-a_0, -a_1, -a_2, \ldots)$.

Kui $a = (a_0, a_1, a_2, \ldots) \in R[X]$ on suvaline, siis $(1, 0, 0, \ldots) \cdot a = (d_0, d_1, d_2, \ldots)$, kus

$$d_k = 1 \cdot a_k + 0 \cdot a_{k-1} + \ldots + 0 \cdot a_1 + 0 \cdot a_0 = a_k$$

iga $k \in \mathbb{N} \cup \{0\}$ korral. Seega $(1,0,0,\ldots)\cdot a = a$ ja analoogiliselt $a\cdot(1,0,0,\ldots) = a$, mis tähendab, et jada $(1,0,0,\ldots)$ on ringi $R[X]$ ühikelement.

Veendume, et korrutamine on assotsiatiivne. Olgu

$$\begin{aligned} a &= (a_0, a_1, a_2, \ldots), \\ b &= (b_0, b_1, b_2, \ldots), \\ c &= (c_0, c_1, c_2, \ldots), \\ ab &= (d_0, d_1, d_2, \ldots), \\ (ab)c &= (e_0, e_1, e_2, \ldots). \end{aligned}$$

Paneme tähele, et

$$\begin{aligned} \sum_{i+j+l=m} a_i b_j c_l &= \sum_{i+j=m} a_i b_j c_0 + \sum_{i+j=m-1} a_i b_j c_1 + \ldots + \sum_{i+j=0} a_i b_j c_m \\ &= \left(\sum_{i+j=m} a_i b_j\right) c_0 + \left(\sum_{i+j=m-1} a_i b_j\right) c_1 + \ldots + \left(\sum_{i+j=0} a_i b_j\right) c_m \\ &= \sum_{k+l=m} \left(\sum_{i+j=k} a_i b_j\right) c_l = \sum_{k+l=m} d_k c_l = e_m. \end{aligned}$$

Analoogiliselt saab näidata, et korrutise $a(bc)$ komponent indeksiga m on võrdne summaga $\sum_{i+j+l=m} a_i b_j c_l$. Järelikult $(ab)c = a(bc)$.

Näitame veel, et $(a+b)c = ac + bc$. Selleks olgu

$$\begin{aligned} a &= (a_0, a_1, a_2, \ldots), \\ b &= (b_0, b_1, b_2, \ldots), \\ c &= (c_0, c_1, c_2, \ldots), \\ a+b &= (d_0, d_1, d_2, \ldots), \\ (a+b)c &= (e_0, e_1, e_2, \ldots), \\ ac &= (u_0, u_1, u_2, \ldots), \\ bc &= (v_0, v_1, v_2, \ldots), \\ ac+bc &= (w_0, w_1, w_2, \ldots). \end{aligned}$$

Siis

$$e_k = \sum_{i+j=k} d_i c_j = \sum_{i+j=k} (a_i + b_i) c_j = \sum_{i+j=k} a_i c_j + \sum_{i+j=k} b_i c_j = u_k + v_k = w_k.$$

Kuna $e_k = w_k$ iga $k \in \mathbb{N} \cup \{0\}$ korral, siis kehtibki distributiivsuse seadus $(a+b)c = ac + bc$. Võrduse $a(b+c) = ac + bc$ saab tõestada analoogiliselt. Seega on $R[X]$ ring.

Olgu nüüd R kommutatiivne ring ja $a = (a_0, a_1, a_2, \ldots), b = (b_0, b_1, b_2, \ldots) \in R[X]$. Siis $\sum_{i+j=k} a_i b_j = \sum_{j+i=k} b_j a_i$, millest järeldub, et $ab = ba$. □

Definitsioon 9.3. Ringi $R[X]$ nimetatakse **polünoomide ringiks üle ringi** R ning tema elemente nimetatakse **polünoomideks**.

Lause 9.4. *Hulk* $R' = \{(r, 0, 0, \ldots) \mid r \in R\} \subseteq R[X]$ *on ring definitsioonis 9.1 defineeritud tehete suhtes. See ring on isomorfne ringiga* R*.*

TÕESTUS. Lihtne on veenduda, et R' on ring. Samuti ei ole keeruline näha, et kujutus $\varphi : R \to R'$, mis on defineeritud võrdusega

$$\varphi(r) := (r, 0, 0, \ldots),$$

$a \in R$, on ringide isomorfism. □

Arvestades lauset 9.4 samastatakse jada $(r, 0, 0, \ldots)$ ringi R elemendiga r.

Tähistame nüüd ringi $R[X]$ elemendi $(0, 1, 0, 0, \ldots)$ sümboliga X:

$$X = (0, 1, 0, 0, \ldots).$$

Kasutades korrutamise definitsiooni saame arvutada, et

$$\begin{aligned} X^1 &= (0, 1, 0, 0, 0, \ldots), \\ X^2 &= (0, 0, 1, 0, 0, \ldots), \\ X^3 &= (0, 0, 0, 1, 0, \ldots), \end{aligned}$$

jne. Matemaatilise induktsiooni abil saab näidata, et iga $n \in \mathbb{N}$ korral on X^n selline jada, mille komponent kohal $n+1$ on 1 (s.t. ringi R ühikelement) ja kõik ülejäänud komponendid on nullid (s.t. ringi R nullelemendid). Kokkuleppeliselt loetakse, et $X^0 = (1, 0, 0, \ldots)$.

Olgu nüüd $f = (a_0, a_1, a_2, \ldots, a_n, 0, \ldots)$ suvaline nullist erinev polünoom ringist $R[X]$ ja olgu a_n polünoomi f kui jada viimane nullist erinev komponent. Kuna

$$\begin{aligned} a_0 &= (a_0, 0, 0, 0, \ldots), \\ a_1X &= (0, a_1, 0, 0, \ldots), \\ a_2X^2 &= (0, 0, a_2, 0, \ldots) \end{aligned}$$

ja nii edasi, siis on polünoom f esitatav kujul

$$f = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n.$$

Harilikult esitataksegi polünoome just sellisel kujul. Liidetavaid $a_0, a_1X, a_2X^2, \ldots, a_nX^n$ nimetatakse polünoomi f **liikmeteks**, liiget a_0 tema **vabaliikmeks** ja liiget a_nX^n tema **pealiikmeks**. Ringi R elemente $a_0, a_1, \ldots, a_n$ nimetatakse **polünoomi** f **kordajateks**. Polünoomi $X \in R[X]$ nimetatakse **muutujaks** ning ringi $R[X]$ kohta öeldakse ka, et see on **polünoomide ring üle ringi** R **muutuja** X **suhtes**. Kui $f \in R[X]$, siis öeldakse, et f on **polünoom üle ringi** R. Nullpolünoomi pealiige ei ole defineeritud. Polünoomi, mille pealiikmel on kuju X^n, s.t., mille pealiikme kordaja on 1, nimetatakse **unitaarseks polünoomiks**.

Märkus 9.5. 1. See, millist tähte kasutatakse polünoomides muutuja tähisena, on kokkuleppe küsimus ja teooria seisukohalt ei ole vahet, kas kasutatakse tähte X või mõnda muud tähte. Teatud olukordades võib olla mugav kasutada mõnda teist tähte, näiteks Y, x või hoopis λ.

2. Vahetu kontrolliga võib veenduda, et mistahes $a, b \in R$ ja $m, n \in \mathbb{N} \cup \{0\}$ korral

$$aX^m \cdot bX^n = abX^{m+n},$$

kusjuures polünoomi aX^0 tõlgendame kui konstantset polünoomi.

9.2. Polünoomi aste

Üheks tähtsamaks polünoomi iseloomustavaks suuruseks on tema aste.

Kui f ei ole nullpolünoom ja tal on kuju

$$f = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n,$$

kusjuures $a_n \neq 0$, siis öeldakse, et polünoomi f **aste** on n. Nullpolünoomi astmeks loetakse $-\infty$. Polünoomi f astet tähistatakse sümboliga $\deg(f)$. Niisiis $\deg(f) \in \{0, 1, 2, \ldots\} \cup \{-\infty\}$, kusjuures me loeme, et $-\infty < n$ iga $n \in \{0, 1, 2, \ldots\}$ korral.

Polünoomi $f \in R[X]$ aste on 0 parajasti siis, kui $f = a_0 \neq 0$, s.t. kui f on nullist erinev ringi R element. Polünoome kujul $f = a_0 \in R$ kutsutakse **konstantseteks polünoomideks**. Kui $\deg(f) = 1$, siis öeldakse, et f on **lineaarpolünoom**. Kui $\deg(f) = 2$, siis öeldakse, et f on **ruutpolünoom**.

Mistahes polünoomide $f, g \in R[X]$ korral järeldub liitmise definitsioonist, et

$$\boxed{\deg(f + g) \leqslant \max(\deg(f), \deg(g)).} \tag{36}$$

Kuna $f - g = f + (-g)$ ja $\deg(-g) = \deg(g)$, siis kehtib ka võrratus

$$\boxed{\deg(f - g) \leqslant \max(\deg(f), \deg(g)).} \tag{37}$$

Võib juhtuda, et polünoomide summa aste on rangelt väiksem liidetavate astmetest. Näiteks kui $f = 1 + 2X^2$ ja $g = 3 + X - 2X^2$, siis $\deg(f + g) = \deg(4 + X) = 1$.

Uurime, kuidas on polünoomide korrutise aste seotud tegurite astmetega.

Lause 9.6. *Mistahes ringi R ja nullist erinevate polünoomide $f, g \in R[X]$ korral*

$$\boxed{\deg(fg) \leqslant \deg(f) + \deg(g).}$$

Kui R on nullitegureita ring, siis

$$\boxed{\deg(fg) = \deg(f) + \deg(g).}$$

TÕESTUS. Olgu $f = (a_0, a_1, \ldots, a_n, 0, \ldots)$ ja $g = (b_0, b_1, \ldots, b_m, 0, \ldots)$, kus a_n ja b_m on viimased nullist erinevad komponendid nendes jadades. Siis $\deg(f) = n$ ja $\deg(g) = m$. Olgu $fg = (d_0, d_1, d_2, \ldots)$. Definitsiooni 9.1 põhjal $d_k = \sum_{i+j=k} a_ib_j$ iga $k \in \mathbb{N} \cup \{0\}$ korral. Kui $k > n + m$ ja $i + j = k$, siis kas $i > n$ või $j > m$ ning seega $a_i = 0$ või $b_j = 0$. Järelikult kõik liidetavad summas $\sum_{i+j=k} a_ib_j$ on nullid, mis tähendab, et $d_k = 0$. Seega viimane nullist erinev komponent jadas fg peab olema mingi d_l, mille korral $l \leqslant n + m$. Järelikult

$$\deg(fg) = l \leqslant n + m = \deg(f) + \deg(g).$$

Kuna $d_{n+m} = \sum_{i+j=n+m} a_ib_j = a_nb_m$, siis nullitegureita ringi R korral $d_{n+m} \neq 0$, sest $a_n, b_m \neq 0$. See tähendab, et

$$\deg(fg) = n + m = \deg(f) + \deg(g).$$

□

Lause 9.6 abil saame lihtsasti tõestada järgmise tulemuse.

Teoreem 9.7. *Kui ring R on nullitegureita, siis ka polünoomide ring $R[X]$ on nullitegureita.*

Tõestus. Kui $f, g \in R[X]$ on nullist erinevad polünoomid, siis $\deg(f) \geqslant 0$ ja $\deg(g) \geqslant 0$. Järelikult $\deg(fg) = \deg(f) + \deg(g) \geqslant 0$, mis tähendab, et ka fg ei ole nullpolünoom. □

Selle teoreemi põhjal võib öelda, et nii ring $\mathbb{Z}[X]$ kui ka ringid $K[X]$, kus K on korpus (nt. $\mathbb{Q}$, $\mathbb{R}$ või $\mathbb{C}$), on nullitegureita.

Teeme nüüd kindlaks, millised on polünoomide ringi pööratavad elemendid. Ka selle juures saame kasutada polünoomi astet.

Lause 9.8. *Olgu R nullitegureita ring. Siis polünoomide ringi $R[X]$ pööratavad elemendid on parajasti ringi R pööratavad elemendid (vaadelduna konstantsete polünoomidena).*

Tõestus. On selge, et kui $a \in R$ on pööratav, siis on ta pööratav ka kui konstantne polünoom.

Oletame nüüd, et $f \in R[X]$ on pööratav polünoom. Siis leidub polünoom $g \in R[X]$ nii, et $fg = 1$. On selge, et f ja g ei ole nullpolünoomid, sest muidu oleks nende korrutis nullpolünoom, mis erineb polünoomist 1 (vt. märkust 1.33). Oletame, et $\deg(f) \geqslant 1$. Siis

$$0 = \deg(1) = \deg(fg) = \deg(f) + \deg(g) \geqslant 1,$$

sest $\deg(g) \geqslant 0$. See vastuolu näitab, et $\deg(f) = 0$, s.t. et f on konstantne polünoom. Analoogiliselt peab g olema konstantne polünoom ja seega f on pööratav ringis R. □

Järeldus 9.9. *Kui K on korpus, siis ringi $K[X]$ pööratavad elemendid on parajasti nullist erinevad konstantsed polünoomid.*

Tõestus. Meenutame, et korpuses on kõik nullist erinevad elemendid pööratavad. □

Näide 9.10. Ringis $\mathbb{R}[X]$ on pööratavad elemendid polünoomid kujul $f = c$, kus $c \in \mathbb{R} \setminus \{0\}$. Ringis $\mathbb{Z}[X]$ on pööratavaid polünoome vaid kaks, nendeks on konstantsed polünoomid 1 ja -1.

9.3. Polünoomide jäägiga jagamine

Nagu hästi teada, saab iga täisarvu jagada jäägiga naturaalarvuga. Tuleb välja, et midagi sarnast saab teha ka polünoomidega.

Teoreem 9.11. *Olgu R nullitegureita ring ning olgu $f, g \in R[X]$, kusjuures polünoomi g pealiikme kordaja on pööratav ringis R. Siis leiduvad üheselt määratud polünoomid $q, r \in R[X]$ nii, et*

$$f = gq + r \quad \textit{ja} \quad \deg(r) < \deg(g).$$

Tõestus. Olgu R nullitegureita ring ja olgu

$$g = b_m X^m + \ldots + b_1 X + b_0 \in R[X]$$

polünoom, mille pealiikme kordaja b_m on pööratav ringis R. Muuhulgas tähendab see, et g ei ole nullpolünoom ja $m = \deg(g) \geqslant 0$. Tõestame, et iga $f \in R[X]$ jaoks leiduvad sellised $q, r \in R[X]$, et $f = gq + r$ ja $\deg(r) < \deg(g)$. Teeme seda matemaatilise induktsiooniga polünoomi f astme järgi.

Kui $f = 0$, siis sobivateks polünoomideks on $q = r = 0$. Kui $\deg(f) = 0$, siis f on nullist erinev konstantne polünoom, $f = a_0 \in R \setminus \{0\}$. Sel juhul, kui g on mittekonstantne polünoom,

siis võib võtta $q = 0$ ja $r = f$. Kui aga $g = b_0 \in R \setminus \{0\}$ ($b_0 \neq 0$, sest g pealiikme kordaja on pööratav), siis

$$a_0 = b_0(b_0^{-1}a_0) + 0$$

ja me võime võtta $r = 0$, $q = b_0^{-1}a_0$. Sellega oleme tõestanud induktsiooni aluse.

Teeme nüüd induktsiooni sammu. Eeldame, et polünoomi

$$f = a_nX^n + \ldots + a_1X + a_0 \in R[X]$$

aste $n > 0$ ja iga polünoomi $k \in R[X]$ jaoks, mille aste on väiksem kui n, leiduvad sellised $q, r \in R[X]$, et $k = gq + r$ ja $\deg(r) < \deg(g)$. On kaks võimalust.

1) $\deg(f) < \deg(g)$. Siis võib võtta $q = 0$ ja $r = f$.

2) $\deg(f) \geqslant \deg(g)$. Siis $n \geqslant m$. Olgu

$$g_1 := g\left(b_m^{-1}a_nX^{n-m}\right) = (b_mX^m + \ldots + b_1X + b_0)\left(b_m^{-1}a_nX^{n-m}\right) \in R[X].$$

Siis g_1 on polünoom, mille pealiige on

$$b_mX^m \cdot b_m^{-1}a_nX^{n-m} = b_mb_m^{-1}a_nX^{m+n-m} = a_nX^n,$$

s.t. sama, mis polünoomi f pealiige. (Üksliige $b_m^{-1}a_nX^{n-m}$ valitaksegi nii, et sellega polünoomi g korrutades tekiks polünoom, mille pealiige langeks kokku f pealiikmega.) Järelikult polünoomi

$$f_1 := f - g_1$$

aste on madalam kui polünoomi f aste n. Rakendades induktsiooni eeldust saab leida sellised $q_1, r \in R[X]$, et $f_1 = gq_1 + r$ ja $\deg(r) < \deg(g)$. Siis aga

$$f = g_1 + f_1 = g\left(b_m^{-1}a_nX^{n-m}\right) + gq_1 + r = g\left(b_m^{-1}a_nX^{n-m} + q_1\right) + r.$$

Tähistades $q := b_m^{-1}a_nX^{n-m} + q_1 \in R[X]$ olemegi saanud sellised q ja r nagu vaja.

Tõestuse lõpetamiseks tuleb veel näidata, et q ja r on üheselt määratud f ja g poolt. Selleks oletame, et

$$f = gq_1 + r_1,\ \deg(r_1) < \deg(g) \quad \text{ja} \quad f = gq_2 + r_2,\ \deg(r_2) < \deg(g).$$

Siis $gq_1 + r_1 = gq_2 + r_2$, millest saame võrduse

$$g(q_1 - q_2) = r_2 - r_1.$$

Kuna $\deg(r_1) < \deg(g)$ ja $\deg(r_2) < \deg(g)$, siis $\deg(r_2 - r_1) < \deg(g)$ tänu võrratusele (37). Oletame vastuväiteliselt, et $q_1 \neq q_2$. Siis $q_1 - q_2 \neq 0$ ja $\deg(q_1 - q_2) \geqslant 0$. Lause 9.6 põhjal

$$\deg(g(q_1 - q_2)) = \deg(g) + \deg(q_1 - q_2).$$

Järelikult

$$\deg(r_2 - r_1) = \deg(g) + \deg(q_1 - q_2) \geqslant \deg(g).$$

See on aga vastuolus võrratusega $\deg(r_2 - r_1) < \deg(g)$. Järelikult $q_1 = q_2$, millest tänu võrdusele $0 = g(q_1 - q_2) = r_2 - r_1$ saame, et ka $r_1 = r_2$. □

Teoreemis 9.11 kirjeldatud polünoome q ja r nimetatakse vastavalt **jagatiseks** ja **jäägiks**, mis tekivad polünoomi f jagamisel polünoomiga g.

9.4. Jaguvus nullitegureita kommutatiivsetes ringides

Kogu käesoleva paragrahvi jooksul olgu R kommutatiivne nullitegureita ring. Märgime, et eesti keeles kutsutakse selliseid ringe ka *integriteetkondadeks* ja inglise keeles kasutatakse terminit *integral domain*. Need on ringid, mille omadused on sarnased täisarvude ringi omadustega.

Integriteetkondades saab rääkida elementide jaguvusest. Täisarvude jaguvusega seotud asju (algarvud, SÜT ja VÜK, modulaarne aritmeetika jne.) uuritakse peamiselt matemaatika valdkonnas, mis kannab nime arvuteooria.

Definitsioon 9.12. Olgu $a, b \in R$. Öeldakse, et element a **jagab** elementi b (ja tähistatakse $a \mid b$), kui leidub selline $c \in R$, et $ac = b$.

Kui element a ei jaga elementi b, siis kirjutatakse $a \nmid b$.

Paneme tähele, et element $a \in R$ on pööratav, parajasti siis, kui a jagab ringi R ühikelementi.

Märkus 9.13. Kui R on korpus, siis mistahes nullist erinevate elementide a ja b korral $a \mid b$, sest $a(a^{-1}b) = b$. Seega korpuste puhul on jaguvuse teooria triviaalne. Küll aga võib jaguvuse teooria ringides (nt. täisarvude ringis) olla vägagi huvitav. Seda näitab kasvõi arvuteooria jätkuv populaarsus kaasaegses matemaatikas.

Lihtne on veenduda, et kehtib järgmine lause.

Lause 9.14. *Jaguvusseosel ringis R on järgmised omadused.*

1. *Kui $a \mid b$ ja $b \mid c$, siis $a \mid c$;*
2. *kui $a \mid b$ ja $a \mid c$, siis $a \mid b \pm c$;*
3. *kui $a \mid b$, siis $a \mid bc$*

iga $a, b, c \in R$ korral.

Definitsioon 9.15. Öeldakse, et ringi R elemendid a ja b on **assotsieeritud** (ja tähistatakse $a \sim b$), kui $a \mid b$ ja $b \mid a$.

Lihtne on aru saada, et assotsieerituse seos on ekvivalentsiseos ringil R. Samuti on lihtne veenduda, et nullelemendiga on assotsieeritud ainult nullelement ise.

Lause 9.16. *Olgu $a, b \in R \setminus \{0\}$. Elemendid a ja b on assotsieeritud parajasti siis, kui leidub mingi pööratav element $u \in R$ nii, et $a = bu$.*

Tõestus. Tarvilikkus. Eeldame, et $a \mid b$ ja $b \mid a$. Siis leiduvad $c, d \in R$ nii, et $ac = b$ ja $bd = a$. Järelikult $a = acd$. Meenutame, et nullitegureita ring on taandamisega (lause 1.40). Kuna $a \neq 0$, siis võime võrdusest $a1 = acd$ elemendi a taandada ja saame, et $1 = cd$. Seega d on pööratav.
Piisavus. Olgu $a = bu$, kus u on pööratav. Siis $b = au^{-1}$. Järelikult $a \mid b$ ja $b \mid a$. □

Näide 9.17. 1. Ringis $\mathbb{Z}$ on elemendid a ja b assotsieeritud parajasti siis, kui $a = b$ või $a = -b$, sest ringi $\mathbb{Z}$ pööratavad elemendid on 1 ja -1.

2. Korpuses on mistahes kaks nullist erinevat elementi a ja b assotsieeritud, sest $a(a^{-1}b) = b$, kus $a^{-1}b$ on pööratav.

3. Kui $f \in R[X]$ ja c on pööratav element ringis R, siis ka konstantne polünoom c on pööratav ringis $R[X]$ (vt. lauset 9.8). Seega polünoomid f ja fc on assotsieeritud. Näiteks ringis $\mathbb{R}[X]$ on polünoomid $4X^3 + 6X - 10$ ja $2X^3 + 3X - 5$ assotsieeritud.

Definitsioon 9.18. Elementi $d \in R$ nimetatakse elementide a ja b **suurimaks ühisteguriks** (ja tähistatakse $d = \text{SÜT}(a, b)$), kui

1. $d \mid a$ ja $d \mid b$;
2. kui $d' \in R$, $d' \mid a$ ja $d' \mid b$, siis $d' \mid d$.

Definitsioon 9.19. Öeldakse, et ringi R elemendid a ja b on **ühistegurita**, kui $\text{SÜT}(a, b) = 1$.

Definitsioon 9.20. Elementi $m \in R$ nimetatakse elementide a ja b **vähimaks ühiskordseks** (ja tähistatakse $m = \text{VÜK}(a, b)$), kui

1. $a \mid m$ ja $b \mid m$;
2. kui $m' \in R$, $a \mid m'$ ja $b \mid m'$, siis $m \mid m'$.

Lause 9.21. *Kui $d \in R$ on elementide a ja b suurim ühistegur ja $d \sim c$, siis ka c on elementide a ja b suurim ühistegur.*

Tõestus. Olgu $d = \text{SÜT}(a, b)$ ja $d \sim c$. Siis $d \mid c$ ja $c \mid d$. Veendume, et c rahuldab elementide a ja b suurima ühisteguri definitsiooni tingimusi.
1. Kuna $c \mid d$, $d \mid a$ ja $d \mid b$, siis ka $c \mid a$ ja $c \mid b$ lause 9.14(1) põhjal.
2. Oletame, et $d' \in R$ on selline, et $d' \mid a$ ja $d' \mid b$. Siis $d' \mid d$. Kuna $d' \mid d$ ja $d \mid c$, siis $d' \mid c$. □

Analoogiliselt saab tõestada järgmise tulemuse.

Lause 9.22. *Kui $m \in R$ on elementide a ja b vähim ühiskordne ja $m \sim n$, siis ka n on elementide a ja b vähim ühiskordne.*

Näide 9.23. Täisarvudel 6 ja -9 on kaks suurimat ühistegurit ringis $\mathbb{Z}$, need on arvud 3 ja -3. Samuti on neil kaks vähimat ühiskordset: 18 ja -18. Kuna kahel ringi elemendil võib olla mitu suurimat ühistegurit, siis iga suurimat ühistegurit sisaldavat võrdust tuleb tõlgendada assotsieerituse täpsuseni. Näiteks võrdust $\text{SÜT}(6, -9) = 3$ tuleks rangelt võttes lugeda nii: täisarvude 6 ja -9 suurimaks ühisteguriks on arv 3 ja kõik arvuga 3 assotsieeritud arvud. Või näiteks korpuse K ja polünoomide $f, g, d \in K[X]$ korral tähendab võrdus $\text{SÜT}(f, g) = d$ seda, et polünoomide f ja g suurimaks ühisteguriks on polünoom d ja kõik sellised polünoomid, mis on saadud d korrutamisel nullist erineva K elemendiga.

Märkus 9.24. On hästi teada, et mistahes kahel täisarvul leidub suurim ühistegur ja vähim ühiskordne. Suvalises ringis ei pruugi kõigil elemendipaaridel suurimat ühistegurit ja vähimat ühiskordset leiduda.

Kerge on veenduda, et kehtivad järgmised omadused.

Lause 9.25. *Suurimal ühisteguril ja vähimal ühiskordsel ringis R on järgmised omadused: iga $a, b, c \in R$ korral*

1. $\text{SÜT}(a, b) = a$ *siis ja ainult siis, kui* $a \mid b$,
2. $\text{SÜT}(a, 0) = a$,
3. $\text{SÜT}(\text{SÜT}(a, b), c) = \text{SÜT}(a, \text{SÜT}(b, c))$,
4. $\text{VÜK}(a, b) = b$ *siis ja ainult siis, kui* $a \mid b$,
5. $\text{VÜK}(a, 0) = 0$,
6. $\text{VÜK}(\text{VÜK}(a, b), c) = \text{VÜK}(a, \text{VÜK}(b, c))$,

kui kõik siin esinevad suurimad ühistegurid või vähimad ühiskordsed on olemas,

7. $\text{SÜT}(a,b) = 0$ *siis ja ainult siis, kui* $a = 0$ *ja* $b = 0$*,*

8. $\text{VÜK}(a,b) = 0$ *siis ja ainult siis, kui kas* $a = 0$ *või* $b = 0$*.*

Suurim ühistegur ja vähim ühiskordne on omavahel seotud.

Teoreem 9.26. *Olgu* a *ja* b *nullist erinevad elemendid ringist* R*,* $m = \text{VÜK}(a,b)$ *ning kehtigu võrdus* $ab = md$*, kus* $d \in R$*. Siis* $d = \text{SÜT}(a,b)$*.*

Tõestus. Kontrollime suurima ühisteguri definitsiooni tingimusi. Kuna $m = \text{VÜK}(a,b)$, siis $a \mid m$ ja seega leidub selline $a' \in R$, et $aa' = m$. Võrdusest $ab = md$ saame, et $ab = aa'd$. Taandades võrdusest $ab = aa'd$ elemendi a saame võrduse $b = a'd$. Seega $d \mid b$. Analoogiliselt saab näidata, et $d \mid a$.

Olgu nüüd $f \in R$ selline, et $f \mid a$ ja $f \mid b$. Siis leiduvad $a', b' \in R$ nii, et $fa' = a$ ja $fb' = b$. Vaatleme elementi $c := fa'b'$. Siis $c = ab'$ ja $c = ba'$. Seega $a \mid c$ ja $b \mid c$, millest vähima ühiskordse definitsiooni teise tingimuse põhjal järeldub, et $m \mid c$. Viimane aga tähendab, et $mc' = c$ mingi $c' \in R$ korral. Siis

$$mfc' = fmc' = fc = ffa'b' = fa'fb' = ab = md.$$

Kuna $a \neq 0$ ja $b \neq 0$, siis lause 9.25(8) põhjal ka $m \neq 0$. Taandades elemendi m võrdusest $mfc' = md$ saame võrduse $fc' = d$, mis tähendab seda, et $f \mid d$. □

Järeldus 9.27. *Kui ringi* R *nullist erinevatel elementidel* a *ja* b *leidub vähim ühiskordne, siis neil elementidel leidub ka suurim ühistegur.*

Tõestus. Olgu $a, b \neq 0$ ja $m = \text{VÜK}(a,b)$. Kuna ab on elementide a ja b ühine kordne, siis vähima ühiskordse definitsiooni põhjal $m \mid ab$. Seega leidub $d \in R$ nii, et $md = ab$. Teoreemi 9.26 põhjal $d = \text{SÜT}(a,b)$. □

Suurima ühisteguri leidumisest ei pruugi järelduda vähima ühiskordse olemasolu.

Teeme veel ühe järelduse teoreemist 9.26.

Järeldus 9.28. *Kui ringi* R *nullist erinevate elementide* a *ja* b *korral* $m = \text{VÜK}(a,b)$ *ja* $d = \text{SÜT}(a,b)$*, siis* $ab \sim md$*.*

Tõestus. Kuna $d = \text{SÜT}(a,b)$, siis leiduvad sellised $a', b' \in R$, et $da' = a$ ja $db' = b$. Vaatleme elementi $da'b'$. Et $a \mid da'b'$ ja $b \mid da'b'$, siis $da'b'$ on a ja b ühine kordne ja seega $m \mid da'b'$. Järelikult $mm' = da'b'$ mingi $m' \in R$ korral. Nüüd

$$ab = da'db' = dda'b' = dmm' = m(dm').$$

Teoreemi 9.26 põhjal dm' on a ja b suurim ühistegur. Kuna ka d on a ja b suurim ühistegur, siis $dm' \mid d$ ehk $dm'u = d$, kus $u \in R$. Lause 9.25(7) põhjal $d \neq 0$. Taandades elemendi d võrdusest $dm'u = d$ saame, et $m'u = 1$, mis tähendab, et m' on pööratav. Kuna $ab = (md)m'$, siis lause 9.16 tõttu $ab \sim md$. □

Näide 9.29. Vaatleme ringi $\mathbb{Z}$. On teada, et selles ringis on suvalisel kahel elemendil a ja b olemas nii suurim ühistegur d kui ka vähim ühiskordne m. Olgu a ja b mõlemad nullist erinevad. Siis järelduse 9.28 põhjal $ab \sim md$, mis täisarvude korral tähendab seda, et $ab = md$ või $ab = -md$. Oletame, et me oleme mingil viisil arvutanud d. Lihtne on leida korrutis ab ja jagada see d-ga. Tulemus $\frac{ab}{d}$ on võrdne kas arvuga m või arvuga $-m$, mõlemal juhul oleme saanud a ja b ühe vähima ühiskordse. Niisiis, kui oskame leida täisarvude suurimat ühistegurit, siis oskame leida ka vähimat ühiskordset, ja tegelikult ka vastupidi.

9.5. Eukleidese ringid

Nagu oleme näinud selles kursuses, saab polünoome teatud eeldustel jäägiga jagada. Samuti saab jäägiga jagada täisarve. Selles paragrahvis vaatleme ühte ringide klassi, mis sisaldab nii täisarvude ringi kui polünoomide ringe üle korpuste.

Definitsioon 9.30. Nullitegureita kommutatiivset ringi R nimetatakse **Eukleidese ringiks**, kui leidub kujutus

$$\delta : R \setminus \{0\} \to \mathbb{N} \cup \{0\}$$

(δ seab igale nullist erinevale ringi elemendile vastavusse mingi mittenegatiivse täisarvu), mis rahuldab tingimusi:

ER1. iga $a, b \in R \setminus \{0\}$ korral $\delta(ab) \geq \delta(a)$,

ER2. iga $a \in R$ ja iga $b \in R \setminus \{0\}$ korral leiduvad sellised $q, r \in R$, et

$$a = bq + r, \quad \text{kusjuures} \quad r = 0 \quad \text{või} \quad \delta(r) < \delta(b).$$

Elementi q nimetatakse harilikult elementide a ja b **jagatiseks** ning elementi r jagamisel tekkivaks **jäägiks**.

Lause 9.31. *Järgmised ringid on Eukleidese ringid:*

1. *täisarvude ring $\mathbb{Z}$,*
2. *polünoomide ring $K[X]$, kus K on korpus.*

Tõestus. 1. Ringi $\mathbb{Z}$ korral võib kujutuse $\delta : \mathbb{Z} \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defineerida võrdusega

$$\delta(a) := |a|$$

($\delta(a)$ on täisarvu a absoluutväärtus). On lihtne näha, et tingimused ER1 ja ER2 on täidetud.

2. Ringi $K[X]$ korral defineerime kujutuse $\delta : K[X] \setminus \{0\} \to \mathbb{N} \cup \{0\}$ võrdusega

$$\delta(f) := \deg(f)$$

(nullist erineva polünoomi aste on mittenegatiivne täisarv). Tingimus ER1 on täidetud tänu lausele 9.6 ja tingimus ER2 tänu teoreemile 9.11. □

Osutub, et sarnaselt täisarvudega saab Eukleidese ringides suurima ühisteguri leida kasutades niinimetatud *Eukleidese*[8] *algoritmi*. Kirjeldame seda.

Olgu R Eukleidese ring ja $a, b \in R$. Kui $a = 0$, siis definitsiooni põhjal $\text{SÜT}(a, b) = b$ ja kui $b = 0$, siis $\text{SÜT}(a, b) = a$. Eeldame edasises, et $a \neq 0$ ja $b \neq 0$.

Jagame elemendi a jäägiga elemendiga b:

$$a = bq_1 + r_1, \qquad r_1 = 0 \text{ või } \delta(r_1) < \delta(b).$$

Kui $r_1 = 0$, siis $b \mid a$ ja seega $\text{SÜT}(a, b) = b$. Kui $r_1 \neq 0$, siis jagame elemendi b elemendiga r_1:

$$b = r_1 q_2 + r_2, \qquad r_2 = 0 \text{ või } \delta(r_2) < \delta(r_1).$$

[8] Kreeka matemaatiku Eukleidese (u. 365 – u. 300 e.m.a.) järgi

Kui $r_2 = 0$, siis lõpetame; vastasel juhul jagame elemendi r_1 elemendiga r_2:

$$r_1 = r_2 q_3 + r_3, \qquad r_3 = 0 \text{ või } \delta(r_3) < \delta(r_2).$$

Niimoodi jätkame senikaua kui saame mingil sammul jäägiks $r_{n+1} = 0$. Varem või hiljem peab see juhtuma, sest $\delta(b) > \delta(r_1) > \delta(r_2) > \ldots$ ja ei leidu lõpmatuid kahanevaid naturaalarvujadasid. Osutub, et suurimaks ühisteguriks on viimane nullist erinev jääk r_n. Tehtud sammud võib kokku võtta järgmise tabelina:

$$\begin{array}{rcll} a & = & bq_1 + r_1, & \delta(r_1) < \delta(b), \\ b & = & r_1 q_2 + r_2, & \delta(r_2) < \delta(r_1), \\ r_1 & = & r_2 q_3 + r_3, & \delta(r_3) < \delta(r_2), \\ & \ldots & & \\ r_{n-3} & = & r_{n-2} q_{n-1} + r_{n-1} & \delta(r_{n-1}) < \delta(r_{n-2}), \\ r_{n-2} & = & r_{n-1} q_n + r_n & \delta(r_n) < \delta(r_{n-1}), \\ r_{n-1} & = & r_n q_{n+1} + 0. & \end{array} \tag{38}$$

Näitame, et $r_n = \text{SÜT}(a, b)$. Selleks kontrollime definitsiooni tingimusi.

1. Tabeli (38) viimasest reast näeme, et $r_n \mid r_{n-1}$. Kuna $r_n \mid r_n$ ja $r_n \mid r_{n-1}$, siis tabeli eelviimasest reast saame, et $r_n \mid r_{n-2}$. Kuna $r_n \mid r_{n-1}$ ja $r_n \mid r_{n-2}$, siis tabeli üle-eelviimasest reast saame, et $r_n \mid r_{n-3}$. Niimoodi ülespoole liikudes näeme lõpuks, et $r_n \mid b$ ja $r_n \mid a$.

2. Oletame nüüd, et $c \in R$ on selline element, et $c \mid a$ ja $c \mid b$. Tabeli esimese rea põhjal $c \mid a - bq_1 = r_1$. Teisest reast saame, et $c \mid r_2$. Järjest allapoole liikudes näeme lõpuks, et $c \mid r_n$.

Suurima ühisteguri leidmise algoritmi, mida eespool kirjeldasime, nimetatakse **Eukleidese algoritmiks**.

Lause 9.32. *Kui R on Eukleidese ring, $a, b \in R$ ja $d = \text{SÜT}(a, b)$, siis leiduvad sellised $u, v \in R$, et*

$$ua + vb = d.$$

Tõestus. Olgu elementide a ja b suurim ühistegur leitud Eukleidese algoritmi abil, s.t. kehtigu võrdused (38) ja olgu $d = r_n$. Liikudes tabelis altpoolt üles saame

$$\begin{aligned} r_n &= r_{n-2} - r_{n-1} q_n = r_{n-2} - (r_{n-3} - r_{n-2} q_{n-1}) q_n \\ &= r_{n-3}(-q_n) + r_{n-2}(1 + q_{n-1} q_n) = \ldots = au + bv = ua + vb. \end{aligned}$$

□

Järeldus 9.33. *Kui R on Eukleidese ring ja $a, b \in R$, siis $\text{SÜT}(a, b) = 1$ parajasti siis, kui leiduvad sellised $u, v \in R$, et*

$$ua + vb = 1.$$

Tõestus. Tarvilikkus on tõestatud lauses 9.32. Piisavuse näitamiseks oletame, et leiduvad $u, v \in R$ nii, et $ua + vb = 1$. On selge, et $1 \mid a$ ja $1 \mid b$. Oletame, et $c \mid a$ ja $c \mid b$. Siis lause 9.14 põhjal

$$c \mid ua + vb = 1.$$

Sellega oleme tõestanud, et $1 = \text{SÜT}(a, b)$. □

9.6. Faktoriaalsed ringid ja taandumatud polünoomid

Definitsioon 9.34. Kommutatiivse nulliteguriteta ringi R mittepööratavat elementi $p \neq 0$ nimetatakse **taandumatuks**, kui võrdusest $p = ab$, $a, b \in R$, järeldub, et kas a või b on pööratav.

Kui $R = K[X]$, kus K on korpus, siis on eelnev definitsioon samaväärne sellega, et polünoom p on mittekonstantne ja seda ei saa esitada kahe mittekonstantse polünoomi korrutisena.

Näide 9.35. 1. Iga lineaarpolünoom $aX+b \in K[X]$ (kus $a \neq 0$ ja K on korpus) on taandumatu.
2. Polünoom $X^2 + 1 \in \mathbb{R}[X]$ on taandumatu.
3. Polünoom $X^2 - 1 = (X - 1)(X + 1) \in \mathbb{R}[X]$ ei ole taandumatu.
4. Polünoom $X^2 - 2$ on taandumatu üle korpuse $\mathbb{Q}$, kuid taanduv üle korpuse $\mathbb{R}$: $X^2 - 2 = (X + \sqrt{2})(X - \sqrt{2})$.

Taandumatud elemendid mängivad nulliteguriteta kommutatiivsetes ringides samasugust rolli nagu täisarvude ringis seda teevad algarvud ja nende vastandarvud. Tuletame meelde, et aritmeetika põhiteoreemi kohaselt saab iga täisarvu (tegurite järjekorda ja märgi täpsust arvestamata üheselt) esitada korrutisena, kus tegurid on algarvud või nende vastandarvud.

Definitsioon 9.36. Nulliteguriteta kommutatiivset ringi R nimetatakse **faktoriaalseks**, kui iga nullist erinev mittepööratav element $a \in R$ on esitatav korrutisena

$$a = p_1 p_2 \dots p_s, \quad s \in \mathbb{N},\ p_1, \dots, p_s \text{ on taandumatud},$$

ja see esitus on ühene selles mõttes, et mistahes teise esituse

$$a = q_1 q_2 \dots q_t, \quad t \in \mathbb{N},\ q_1, \dots, q_t \text{ on taandumatud},$$

korral $t = s$ ja leidub hulga $\{1, 2, \dots s\}$ substitutsioon σ nii, et p_i ja $q_{\sigma(i)}$ on assotsieeritud iga $i \in \{1, \dots, s\}$ korral.

Näide 9.37. 1. Täisarvude ring $\mathbb{Z}$ on aritmeetika põhiteoreemi kohaselt faktoriaalne.
2. Kompleksarvude korpuse alamring

$$\mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Z}\}$$

ei ole faktoriaalne, aga selles saab iga nullist erinevat mittepööratavat elementi esitada (mitte tingimata üheselt!) taandumatute elementide korrutisena. Näiteks

$$4 = (1 + \sqrt{-3})(1 - \sqrt{-3}) = 2 \cdot 2$$

ja saab näidata, et elemendid $2, 1 + \sqrt{-3}, 1 - \sqrt{-3}$ on taandumatud.
3. Algebraliste täisarvude ringis

$$\{a \in \mathbb{C} \mid \exists \text{ unitaarne } f \in \mathbb{Z}[X] : f(a) = 0\}$$

ei saa iga nullist erinevat mittepööratavat elementi esitada taandumatute elementide korrutisena.

Osutub, et faktoriaalsuse definitsioonis esinev ühesuse mõiste on samaväärne sellega, et vaadeldavas ringis kehtib Eukleidese lemma. Seda tulemust me käesolevas kursuses tõestada ei jõua.

Teoreem 9.38. *Olgu R nulliteguriteta kommutatiivne ring, mille iga nullist erinev mittepööratav element on esitatav taandumatute elementide korrutisena. Siis ring R on faktoriaalne parajasti siis, kui*

$$(\forall a, b, p \in R)(p \text{ on taandumatu ja } p \mid ab \Longrightarrow (p \mid a \text{ või } p \mid b)\,.$$

Teoreem 9.39. *Eukleidese ring on faktoriaalne ring.*

Tõestus. Olgu R Eukleidese ring ja $\delta : R \setminus \{0\} \to \mathbb{N} \cup \{0\}$ sellele vastav kujutus. Näitame esiteks, et δ väheneb mittepööratavatel teguritel, s.t. kui $R \ni a = bc$, kus $b, c \in R$ on nullist erinevad mittepööratavad elemendid, siis $\delta(a) > \delta(b)$. Definitsiooni kohaselt $\delta(a) \geq \delta(b)$, seega tuleb meil välistada juht $\delta(a) = \delta(b)$. Oletame vastuväiteliselt, et viimane võrdus siiski kehtib ja leiame Eukleidese ringi jäägiga jagamise abil $q, r \in R$ nii, et

$$b = aq + r, \quad \text{kus } \delta(r) < \delta(a) \text{ või } r = 0.$$

Vaatleme mõlemat juhtu.

1) $r = 0$. Siis $b = aq$, kust $a = bc = aqc$. Kuna $a \neq 0$ tänu nullitegurite puudumisele, siis võime temaga taandada ja leida, et $1 = qc$, s.t. c on pööratav. See on vastuolus meie eeldusega.

2) $\delta(r) < \delta(a)$. Kuna $1 - cq \neq 0$ (sest c ei ole pööratav) ja $b \neq 0$, siis saame kasutada Eukleidese ringi definitsiooni esimest tingimust ja leida, et

$$\delta(a) > \delta(r) = \delta(b - aq) = \delta(b - bcq) = \delta(b(1 - cq)) \geq \delta(b) = \delta(a).$$

Oleme jälle jõudnud vastuoluni, seega meie oletus $\delta(a) = \delta(b)$ ei kehti ja $\delta(a) > \delta(b)$.

Näitame nüüd, et iga nullist erinev mittepööratav element $a \in R$ esitub taandumatute elementide korrutisena. Kui a on taandumatu, on ta vaadeldav ühest taandumatust tegurist koosneva korrutisena. Kui a ei ole taandumatu, siis $a = bc$, kus $b, c \in R \setminus \{0\}$ on mittepööratavad. Kui b ja c on taandumatud, on meil esitus olemas. Kui üks neist, üldisust kitsendamata b, ei ole taandumatu, siis $b = ef$, kus $e, f \in R \setminus \{0\}$ on mittepööratavad. Jälle, kas e ja f on mõlemad taandumatud või me saame ühe või mõlemad neist esitada mittepööratavate elementide korrutisena. Kuna tõestuse esimese osa põhjal on δ väärtused nullist erinevatel mittepööratavatel teguritel rangelt kahanevad ja δ minimaalne väärtus on 0, siis seda teguriteks lahutamise protsessi ei saa lõpmatult jätkata ja lõpliku arvu sammude järel on a esitatav taandumatute elementide korrutisena.

Viimaks on meil vaja näidata, et nullist erineva mittepööratava elemendi esitus taandumatute elementide korrutisena on ühene definitsiooni 9.36 mõttes. Teoreemi 9.38 tõttu piisab, kui me näitame, et ringis R kehtib Eukleidese lemma. Olgu $p \in R$ taandumatu ja $p \mid ab$, $a, b \in R$.

Tähistame $d := \text{SÜT}(p, a)$. Siis $d \mid p$ ehk leidub $d' \in R$ nii, et $dd' = p$. Kuna p on taandumatu, siis on kaks võimalust.

1) d on pööratav. Järelduse 9.33 põhjal teame, et leiduvad $u, v \in R$ nii, et $pu + av = d$. Korrutades selle võrduse mõlemaid pooli elemendiga b saame võrduse

$$pub + abv = bd.$$

Kuna eelduse $p \mid ab$ ja lause 9.14 tõttu p jagab selle võrduse vasakut poolt, siis ta jagab ka elementi bd. Seega leidub $c \in R$ nii, et $pc = bd$. Siit saame $p(cd^{-1}) = b$ ehk $p \mid b$.

2) d' on pööratav ehk $d \sim p$. Kuna $p \mid d$ ja $d \mid a$, siis transitiivsuse tõttu ka $p \mid a$. □

Kahe erineva algarvu suurim ühistegur on 1. Analoogiline omadus on ka taandumatutel polünoomidel.

Lemma 9.40. *Olgu K korpus. Kui $p, q \in K[X]$ on mitteassotsieeritud taandumatud polünoomid, siis* $\mathrm{SÜT}(p, q) = 1$.

Tõestus. Oletame vastuväiteliselt, et polünoomide p ja q suurim ühistegur d ei ole 1. See tähendab, et d ei ole konstantne polünoom, $\deg(d) \geqslant 1$. Kuna $d \mid p$ ja $d \mid q$, siis leiduvad polünoomid $p_1, q_1 \in K[X]$ nii, et $dp_1 = p$ ja $dq_1 = q$. Et $p, q \neq 0$, siis ka $p_1, q_1 \neq 0$. On järgmised võimalused.
1. $\deg(p_1) = \deg(q_1) = 0$. Siis p_1, q_1 on konstantsed polünoomid ja seega $p \sim d \sim q$, kust $p \sim q$, mis on vastuolus eeldusega.
2. $\deg(p_1) \geqslant 1$. Siis p on kahe mittekonstantse polünoomi (d ja p_1) korrutis, mis on vastuolus p taandumatusega.
3. $\deg(q_1) \geqslant 1$. Sel juhul tekib vastuolu q taandumatusega. □

Lause 9.31 osast 2., teoreemist 9.39 ja lemmast 9.40 saab nüüd tuletada aritmeetika põhiteoreemi analoogi ringi $K[X]$ jaoks.

Järeldus 9.41. *Olgu K korpus. Iga mittekonstantse polünoomi $f \in K[X]$ saab (tegurite järjekorra täpsuseni üheselt) esitada korrutisena*

$$f = ap_1^{k_1} p_2^{k_2} \dots p_m^{k_m},$$

kus $a \in K$, $p_1, p_2, \dots, p_m$ on paarikaupa ühistegurita unitaarsed taandumatud polünoomid ja $k_1, \dots, k_m \in \mathbb{N}$.

9.7. Jagatiste korpus

Selles paragrahvis uurime konstruktsiooni, mille abil muuhulgas konstrueeritakse ratsionaalarvud lähtudes täisarvudest.

Olgu $R \neq \{0\}$ nullitegureita kommutatiivne ring. Defineerime hulgal

$$R \times (R \setminus \{0\}) = \{(a, b) \mid a, b \in R, b \neq 0\}$$

binaarse seose ρ järgmiselt:

$$(a, b)\rho(c, d) \iff ad = bc.$$

Veendume, et ρ on ekvivalentsiseos.

Kuna $ab = ba$, siis $(a, b)\rho(a, b)$ ja seega ρ on refleksiivne. Et

$$(a, b)\rho(c, d) \iff ad = bc \iff cb = da \iff (c, d)\rho(a, b),$$

siis on see seos ka sümmeetriline. Transitiivsuse tõestamiseks oletame, et kehtib $(a, b)\rho(c, d)$ ja $(c, d)\rho(e, f)$. Siis $ad = bc$ ja $cf = de$. Järelikult $adf = bcf = bde$. Taandades elemendi d võrdusest $adf = bde$ (paneme tähele, et $d \neq 0$) saame võrduse $af = be$, mis tähendab, et $(a, b)\rho(e, f)$. Sellega on ka transitiivsus näidatud ja ρ on ekvivalentsiseos.

Paari $(a, b) \in R \times (R \setminus \{0\})$ ekvivalentsiklassi seose ρ järgi tähistame sümboliga $\frac{a}{b}$, s.t.

$$\frac{a}{b} = \{(c, d) \in R \times (R \setminus \{0\}) \mid (a, b)\rho(c, d)\}.$$

Kahe paari ekvivalentsiklassid on võrdsed parajasti siis, kui need paarid on seoses ρ, seega

$$\boxed{\frac{a}{b} = \frac{c}{d} \iff ad = bc.}$$

Paneme tähele, et

$$\boxed{\frac{ac}{bc} = \frac{a}{b}}$$

mistahes $a \in R$ ja $b, c \in R \setminus \{0\}$ korral, sest $acb = bca$.

Vaatleme faktorhulka seose ρ järgi:

$$Q(R) := (R \times (R \setminus \{0\}))/\rho = \left\{ \frac{a}{b} \;\middle|\; a, b \in R, b \neq 0 \right\}.$$

Hulgal $Q(R)$ defineerime liitmise ja korrutamise valemitega

$$\begin{array}{rcl} \dfrac{a}{b} + \dfrac{c}{d} & := & \dfrac{ad + bc}{bd}, \\ \dfrac{a}{b} \cdot \dfrac{c}{d} & := & \dfrac{ac}{bd}. \end{array}$$

Tuleb veenduda, et need definitsioonid on korrektsed. Olgu

$$\frac{a}{b} = \frac{a_1}{b_1} \quad \text{ja} \quad \frac{c}{d} = \frac{c_1}{d_1}.$$

Siis $ab_1 = ba_1$ ja $cd_1 = dc_1$. Korrutades esimest neist võrdustest elemendiga dd_1 ja teist elemendiga bb_1 saame

$$\begin{array}{rcl} adb_1d_1 & = & a_1d_1bd, \\ bcb_1d_1 & = & b_1c_1bd, \end{array}$$

millest vastavate poolte liitmisel järeldub, et $adb_1d_1 + bcb_1d_1 = a_1d_1bd + b_1c_1bd$. Kasutades distributiivsust võime kirjutada $(ad + bc)b_1d_1 = bd(a_1d_1 + b_1c_1)$, mis tähendab, et

$$\frac{ad + bc}{bd} = \frac{a_1d_1 + b_1c_1}{b_1d_1}.$$

Seega on liitmine korrektselt defineeritud. Korrutamise korrektsuse kontrolli jätame lugejale läbimõtlemiseks.

Teoreem 9.42. *Mistahes nullitegureita kommutatiivse ringi $R \neq \{0\}$ korral on $Q(R)$ korpus.*

Tõestus. Et $0 \cdot 1 \neq 1 \cdot 1$, siis $\frac{0}{1} \neq \frac{1}{1}$ ja seega hulgas $Q(R)$ on vähemalt 2 elementi. Kuna

$$\frac{a}{b} + \frac{0}{1} = \frac{a \cdot 1 + b \cdot 0}{b \cdot 1} = \frac{a}{b} = \frac{0}{1} + \frac{a}{b}$$

mistahes $\frac{a}{b} \in Q(R)$ korral, siis $\frac{0}{1}$ on nullelement liitmise suhtes. Saab näidata, et $\frac{1}{1}$ on ühikelement korrutamise suhtes ja et elemendi $\frac{a}{b}$, kus $a, b \neq 0$, pöördelement on $\frac{b}{a}$. Üksikasjade läbitegemise jätame lugeja hooleks. □

Definitsioon 9.43. Korpust $Q(R)$ nimetatakse nullitegureita kommutatiivse ringi R **jagatiste korpuseks**.

Konstrueerides täisarvude ringi $\mathbb{Z}$ jagatiste korpuse $Q(\mathbb{Z})$ saame ratsionaalarvude korpuse $\mathbb{Q}$.

Nii nagu täisarv a samastatakse ratsionaalarvuga $\frac{a}{1}$ saab ka ringi R elemendi a samastada korpuse $Q(R)$ elemendiga $\frac{a}{1}$.

Lause 9.44. *Olgu R nullitegureita kommutatiivne ring. Siis korpuse $Q(R)$ alamhulk*

$$R' = \left\{ \frac{a}{1} \mid a \in R \right\}$$

on ise ka ring hulgal $Q(R)$ defineeritud tehete suhtes. See ring on isomorfne ringiga R.

TÕESTUS. Esimese väite tõestuse jätame läbimõtlemiseks lugejale.

Kujutuse $f : R \to R'$ defineerime võrdusega

$$f(a) := \frac{a}{1}$$

mistahes $a \in R$ korral. On selge, et see kujutus on sürjektiivne. Kui $f(a) = f(b)$, siis $\frac{a}{1} = \frac{b}{1}$, millest $a1 = 1b$ ehk $a = b$. Seega on f ka injektiivne. Ka definitsiooni 2.3 ülejäänud tingimuste kontroll pole keeruline. □

9.8. Ratsionaalmurdude korpus

Kui K on korpus, siis polünoomide ring $K[X]$ on nullitegureita kommutatiivne ring. Seega võime konstrueerida tema jagatiste korpuse

$$\boxed{Q(K[X]) = \left\{ \frac{f}{g} \mid f, g \in K[X], g \neq 0 \right\}}$$

nii nagu seda tegime eelmises paragrahvis. Korpust $Q(K[X])$ nimetatakse **ratsionaalmurdude korpuseks üle korpuse** K ning tähistatakse $K(X)$. Korpuse $K(X)$ elemente nimetatakse **ratsionaalmurdudeks**.

Ratsionaalmurdude jaoks saab defineerida astme, mis on teatud täisarv (võib olla ka negatiivne).

Definitsioon 9.45. Ratsionaalmurru $\frac{f}{g} \in K(X)$, kus $f, g \neq 0$, **aste** on f astme ja g astme vahe:

$$\boxed{\deg\left(\frac{f}{g}\right) := \deg(f) - \deg(g).}$$

Veendume, et see definitsioon on korrektne. Oletame, et $\frac{f_1}{g_1} = \frac{f_2}{g_2}$. Siis $f_1 g_2 = f_2 g_1$. Kuna korpuses ei ole nullitegureid, siis lause 9.6 põhjal

$$\deg(f_1) + \deg(g_2) = \deg(f_1 g_2) = \deg(f_2 g_1) = \deg(f_2) + \deg(g_1).$$

Siit aga järeldub võrdus $\deg(f_1) - \deg(g_1) = \deg(f_2) - \deg(g_2)$ ehk $\deg\left(\frac{f_1}{g_1}\right) = \deg\left(\frac{f_2}{g_2}\right)$. Seega on ratsionaalmurru astme definitsioon tõepoolest korrektne.

Definitsioon 9.46. Nullist erinevat ratsionaalmurdu $\frac{f}{g} \in K(X)$ nimetatakse **lihtmurruks**, kui tema aste on negatiivne (s.t. f aste on väiksem kui g aste).

Näide 9.47. Ratsionaalmurd

$$\frac{2X^3 - 4X + 5}{X^5 - 3X^2} \in \mathbb{R}(X)$$

on lihtmurd, sest

$$\deg\left(\frac{2X^3 - 4X + 5}{X^5 - 3X^2}\right) = \deg(2X^3 - 4X + 5) - \deg(X^5 - 3X^2) = 3 - 5 = -2 < 0.$$

Lause 9.48. *Iga nullist erinev ratsionaalmurd on üheselt esitatav polünoomi ja lihtmurru summana.*

TÕESTUS. Vaatleme ratsionaalmurdu $\frac{f}{g} \in K(X)$, kus $f, g \neq 0$. Jagades polünoomi f jäägiga polünoomiga g saame leida polünoomid q ja r nii, et $f = gq + r$ ja $\deg(r) < \deg(g)$. Siis aga

$$\frac{f}{g} = \frac{gq+r}{g} = \frac{gq}{g} + \frac{r}{g} = \frac{q}{1} + \frac{r}{g} = q + \frac{r}{g},$$

kus $\frac{r}{g}$ on lihtmurd. (Siin kasutasime seda, et ratsionaalmurd $\frac{q}{1}$ on samastatud polünoomiga q.)

Näitame nüüd, et sellised esitused on ühesed. Oletame, et

$$\frac{f}{g} = p_1 + \frac{q_1}{r_1} = p_2 + \frac{q_2}{r_2},$$

kus p_1, p_2 on polünoomid ja $\frac{q_1}{r_1}, \frac{q_2}{r_2}$ on lihtmurrud. Siis

$$p_1 - p_2 = \frac{q_2}{r_2} - \frac{q_1}{r_1} = \frac{r_1 q_2 - r_2 q_1}{r_1 r_2}.$$

Kasutades võrratust (37) ja lauset 9.6 saame, et

$$\begin{aligned}
\deg(p_1 - p_2) &= \deg\left(\frac{r_1 q_2 - r_2 q_1}{r_1 r_2}\right) \\
&= \deg(r_1 q_2 - r_2 q_1) - \deg(r_1 r_2) \\
&\leqslant \max(\deg(r_1 q_2), \deg(r_2 q_1)) - \deg(r_1 r_2) \\
&= \max(\deg(r_1 q_2) - \deg(r_1 r_2), \deg(r_2 q_1) - \deg(r_1 r_2)) \\
&= \max(\deg(q_2) - \deg(r_2), \deg(q_1) - \deg(r_1)) \\
&< 0.
\end{aligned}$$

Ainuke polünoom, mille aste on negatiivne, on nullpolünoom, seega $p_1 - p_2 = 0$ ehk $p_1 = p_2$. Võrdusest $p_1 + \frac{q_1}{r_1} = p_2 + \frac{q_2}{r_2}$ järeldub siis, et ka $\frac{q_1}{r_1} = \frac{q_2}{r_2}$. Seega ratsionaalmurd esitub ainult ühel viisil polünoomi ja lihtmurru summana. □

Definitsioon 9.49. Nullist erinevat ratsionaalmurdu $\frac{f}{g} \in K(X)$ nimetatakse **algmurruks**, kui $g = p^n$, kus n on naturaalarv, p on taandumatu polünoom ja $\deg(f) < \deg(p)$.

Näide 9.50. Ratsionaalmurrud

$$\frac{4}{3X+5} \in \mathbb{R}(X) \quad \text{ja} \quad \frac{2X+3}{(X^2+1)^7} \in \mathbb{R}(X)$$

on algmurrud, aga ratsionaalmurd

$$\frac{2X^2+3}{(X^2+1)^2} \in \mathbb{R}(X)$$

ei ole algmurd.

Järgmise teoreemi tõestamiseks läheb meil vaja kahte tulemust polünoomide kohta.

Lause 9.51. *Olgu $f, g_1, \ldots, g_m \in K[X]$, kus K on korpus. Kui* $\SÜT(f, g_i) = 1$ *iga* $i \in \{1, \ldots, m\}$ *korral, siis ka* $\SÜT(f, g_1 \ldots g_m) = 1$.

Tõestus. Oletame vastuväiteliselt, et $d := \SÜT(f, g_1 \ldots g_m) \neq 1$. Siis ka $d \neq 0$, sest vastasel korral $0 \mid f$ ja $0 \mid g_1 \ldots g_m$, kust $f = 0$, $g_1 \ldots g_m = 0$ ja nullitegurite puudumise tõttu $g_i = 0$ mingi $i \in \{1, \ldots, m\}$ korral. Aga nüüd $\SÜT(f, g_i) = 0 \neq 1$, vastuolu. Kuna korpuses on kõik konstandid peale 0 omavahel assotsieeritud, siis lause 9.21 tõttu on d mittekonstantne ja järelduse 9.41 põhjal peab tal olema vähemalt üks taandumatu tegur $p \in K[X]$. Kuna $p \mid d$ ja $d \mid g_1 \ldots g_m$, siis ka $p \mid g_1 \ldots g_m$. Eukleidese ringis $K[X]$ kehtib teoreemide 9.38 ning 9.39 tõttu Eukleidese lemma, järelikult $p \mid g_j$ mingi $j \in \{1, \ldots, m\}$ korral. Aga nüüd $p \mid d$, $d \mid f$, seega ka $p \mid f$ ja $p \mid g_j$. Suurima ühisteguri definitsiooni põhjal $p \mid \SÜT(f, g_j) = 1$, ehk p on pööratav. Samas, p on taandumatu ja taandumatu element ei saa olla pööratav. Oleme jõudnud vastuoluni, mistõttu peab ikkagi kehtima $d = 1$. □

Järeldus 9.52. *Kui polünoomid $f, g \in K[X]$ on ühistegurita ja $k, l \in \mathbb{N}$, siis ka polünoomid $f^k, g^l \in K[X]$ on ühistegurita.*

Tõestus. Rakendame lauset 9.51 juhule $f = f$, $g_i = g$, $i = 1, \ldots, l$, mida me saame teha eelduse $\SÜT(f, g) = 1$ tõttu. Siis $\SÜT(f, g^l) = 1$. Võttes analoogiliselt $f = g^l$ ja $g_i = f$, $i = 1, \ldots, k$, saame $\SÜT(g^l, f) = 1$ abil, et $\SÜT(g^l, f^k) = 1$, mida oligi tarvis tõestada. □

Üheks olulisemaks tulemuseks ratsionaalmurdude kohta on järgmine teoreem.

Teoreem 9.53. *Iga lihtmurd üle korpuse K on esitatav algmurdude summana.*

Tõestus. 1) Näitame kõigepäält, et kui $\frac{f}{g} \in K(X)$ on lihtmurd ja $g = g_1 g_2$, kus $\SÜT(g_1, g_2) = 1$, siis $\frac{f}{g}$ on esitatav kujul $\frac{f}{g} = \frac{v_1}{g_1} + \frac{v_2}{g_2}$, kus $\frac{v_1}{g_1}$ ja $\frac{v_2}{g_2}$ on lihtmurrud. Kuna $K[X]$ on Eukleidese ring ja $\SÜT(g_1, g_2) = 1$, siis lause 9.32 põhjal saab leida sellised polünoomid $u_1, u_2 \in K[X]$, et $1 = u_1 g_1 + u_2 g_2$. Korrutades selle võrduse pooli polünoomiga f saame võrduse $f = f u_1 g_1 + f u_2 g_2$. Jagame polünoomi $f u_1$ jäägiga polünoomiga g_2. Tekivad jagatis q ja jääk v_2 nii, et

$$f u_1 = g_2 q + v_2 \quad \text{ja} \quad \deg(v_2) < \deg(g_2).$$

Siis

$$f = f u_1 g_1 + f u_2 g_2 = (g_2 q + v_2) g_1 + (f u_2) g_2 = (g_1 q + f u_2) g_2 + v_2 g_1 = v_1 g_2 + v_2 g_1,$$

kus oleme tähistanud $g_1 q + f u_2 =: v_1$. Järelikult

$$\frac{f}{g} = \frac{v_1 g_2 + v_2 g_1}{g_1 g_2} = \frac{v_1 g_2}{g_1 g_2} + \frac{v_2 g_1}{g_1 g_2} = \frac{v_1}{g_1} + \frac{v_2}{g_2},$$

kus $\frac{v_2}{g_2}$ on lihtmurd.

Näitame, et ka $\frac{v_1}{g_1}$ on lihtmurd, see tähendab, et $\deg(v_1) < \deg(g_1)$. Selleks oletame vastuväiteliselt, et $\deg(v_1) \geqslant \deg(g_1)$. Siis

$$\deg(v_1 g_2) = \deg(v_1) + \deg(g_2) \geqslant \deg(g_1) + \deg(g_2) = \deg(g_1 g_2) = \deg(g).$$

Võrratusest $\deg(v_2) < \deg(g_2)$ aga järeldub, et

$$\deg(v_2 g_1) < \deg(g_2 g_1) = \deg(g) \leqslant \deg(v_1 g_2).$$

Seega

$$\deg(f) = \deg(v_1 g_2 + v_2 g_1) = \deg(v_1 g_2) \geqslant \deg(g).$$

Võrratus $\deg(f) \geqslant \deg(g)$ on aga vastuolus eeldusega, et $\frac{f}{g}$ on lihtmurd. Järelikult peab kehtima võrratus $\deg(v_1) < \deg(g_1)$ ning $\frac{v_1}{g_1}$ on lihtmurd.

2) Et $\frac{f}{g}$ on lihtmurd, siis g on mittekonstantne polünoom. Tänu järeldusele 9.41 saame g esitada kujul

$$g = c p_1^{k_1} p_2^{k_2} \dots p_m^{k_m},$$

kus $c \in K$, $p_1, p_2, \dots, p_m$ on paarikaupa ühistegurita taandumatud polünoomid ja $k_1, k_2, \dots, k_m$ on naturaalarvud. Näitame matemaatilise induktsiooni abil, et iga lihtmurd $\frac{h}{p_1^{k_1} \dots p_r^{k_r}}$, kus $r \in \{2, \dots m\}$, on esitatav summana

$$\frac{h}{p_1^{k_1} \dots p_r^{k_r}} = \frac{a_1}{p_1^{k_1}} + \dots + \frac{a_r}{p_r^{k_r}},$$

kus liidetavad on lihtmurrud. Tänu järeldusele 9.52 on polünoomid $p_1^{k_1}$ ja $p_2^{k_2}$ ühistegurita, seega kasutades tõestuse esimest osa näeme, et väide kehtib $r = 2$ korral. Oletame nüüd, et $r > 2$ ja et väide kehtib $r-1$ korral. Kuna SÜT $\left(p_i^{k_i}, p_r^{k_r}\right) = 1$ iga $i \in \{1, \dots, r-1\}$ korral, siis lausest 9.51 järeldub, et SÜT $\left(p_1^{k_1} \dots p_{r-1}^{k_{r-1}}, p_r^{k_r}\right) = 1$. Seega tõestuse esimese osa põhjal

$$\frac{h}{p_1^{k_1} \dots p_{r-1}^{k_{r-1}} p_r^{k_r}} = \frac{a}{p_1^{k_1} \dots p_{r-1}^{k_{r-1}}} + \frac{b}{p_r^{k_r}},$$

kus liidetavad on lihtmurrud. Kasutades induktsiooni eeldust saame murru $\frac{a}{p_1^{k_1} \dots p_{r-1}^{k_{r-1}}}$ esitada lihtmurdude summana ning seega on meil vajalik summa olemas ka murru $\frac{h}{p_1^{k_1} \dots p_{r-1}^{k_{r-1}} p_r^{k_r}}$ jaoks, s.t. väide kehtib r korral.

Tõestatud väitest järeldub $r = m$ ja $h = c^{-1} f$ korral, et

$$\frac{f}{g} = \frac{c c^{-1} f}{c p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}} = \frac{c^{-1} f}{p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}} = \frac{a_1}{p_1^{k_1}} + \frac{a_2}{p_2^{k_2}} + \dots + \frac{a_m}{p_m^{k_m}},$$

kus kõik liidetavad on lihtmurrud.

3) Tõestuse lõpetamiseks peame näitama, et mistahes lihtmurd $\frac{a}{p^k}$, kus p on taandumatu polünoom, on esitatav algmurdude summana. Olgugi $\frac{a}{p^k}$ selline lihtmurd. Jagame polünoomi a jäägiga polünoomiga p^{k-1}:

$$a = q_1 p^{k-1} + v_1, \quad \deg(v_1) < \deg(p^{k-1}).$$

Näitame, et

$$\deg(q_1) < \deg(p).$$

Kui oletaksime vastuväiteliselt, et $\deg(q_1) \geqslant \deg(p)$, siis

$$\deg(a) = \deg(q_1 p^{k-1} + v_1) = \deg(q_1 p^{k-1}) \geqslant \deg(p p^{k-1}) = \deg(p^k),$$

mis on vastuolus eeldusega, et $\frac{a}{p^k}$ on lihtmurd.

Jagame nüüd jäägi v_1 jäägiga polünoomiga p^{k-2}. Tekivad jagatis q_2 ja jääk v_2 nii, et

$$v_1 = q_2 p^{k-2} + v_2, \quad \deg(v_2) < \deg(p^{k-2}), \quad \deg(q_2) < \deg(p).$$

Analoogiliselt jätkates jõuame võrduseni ja võrratusteni

$$v_{k-2} = q_{k-1} p + v_{k-1}, \quad \deg(v_{k-1}) < \deg(p), \quad \deg(q_{k-1}) < \deg(p).$$

Siis aga

$$a = q_1 p^{k-1} + q_2 p^{k-2} + \ldots + q_{k-1} p + v_{k-1},$$

millest

$$\frac{a}{p^k} = \frac{q_1}{p} + \frac{q_2}{p^2} + \ldots + \frac{q_{k-1}}{p^{k-1}} + \frac{v_{k-1}}{p^k}.$$

Kuna liidetavad viimases summas on algmurrud, siis on teoreem tõestatud. □

Märkus 9.54. Eelmise teoreemi tõestuses peitub ka algoritm lihtmurru esitamiseks algmurdude summana. Märgime, et selle algoritmi keerukaim osa on polünoomi g tegureiks lahutamine.

Seda teoreemi kasutatakse matemaatilises analüüsis ratsionaalfunktsioonide integreerimisel juhul kui $K = \mathbb{R}$. Täpsemalt võib selle kohta lugeda näiteks järgmistest allikatest:

- Gunnar Kangro, Matemaatiline analüüs I, 1982, lk. 313–329,
- Leiki Loone ja Virge Soomer, Matemaatilise analüüsi algkursus, 2007, lk. 173–180,
- Elmar Reimers, Matemaatilise analüüsi praktikum I, 1988, lk. 139–144.

9.9. Polünoomi juured

Käesolevas paragrahvis eeldame kõikjal, et R on nullitegureita kommutatiivne ring, milles on vähemalt kaks elementi. Teatud põhjustel on juurtest rääkides otstarbekas nummerdada polünoomi kordajaid vastupidises järjekorras, s.t. nii, et pealiikme kordaja on a_0, järgmise liikme kordaja a_1 jne.

Definitsioon 9.55. Olgu R nullitegureita kommutatiivne ring,

$$f(X) = a_0 X^n + a_1 X^{n-1} + \ldots + a_{n-1} X + a_n$$

polünoom ringist $R[X]$ ja $c \in R$. Ringi R elementi

$$a_0 c^n + a_1 c^{n-1} + \ldots + a_{n-1} c + a_n$$

nimetatakse polünoomi $f(X)$ **väärtuseks** kohal c ja tähistatakse $f(c)$.

Näide 9.56. Polünoomi $f(X) = X^3 + \overline{3}X + \overline{4} \in \mathbb{Z}_5[X]$ väärtus kohal $c = \overline{2}$ on

$$f(\overline{2}) = \overline{2}^3 + \overline{3} \cdot \overline{2} + \overline{4} = \overline{8} + \overline{6} + \overline{4} = \overline{3} + \overline{1} + \overline{4} = \overline{3}.$$

Definitsioon 9.57. Elementi $c \in R$ nimetatakse polünoomi $f(X)$ **juureks**, kui $f(c) = 0$.

Teoreem 9.58 (Bezout' teoreem).[9] *Polünoomi $f(X) \in R[X]$ väärtus kohal $c \in R$ on võrdne jäägiga, mis tekib polünoomi $f(X)$ jagamisel polünoomiga $X - c$.*

[9] Étienne Bézout (1730–1783) — prantsuse matemaatik

Tõestus. Jagame polünoomi $f(X) \in R[X]$ jäägiga lineaarpolünoomiga $X - c$, kus $c \in R$:

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

Kuna $\deg(r(X)) < 1$, siis $r(X) = r \in R$ on konstantne polünoom. Kui kaks polünoomi (antud juhul $f(X)$ ja $(X - c)q(X) + r(X)$) on võrdsed, siis on võrdsed ka nende väärtused kohal c. Järelikult

$$f(c) = (c - c) \cdot q(c) + r = 0 + r = r.$$

□

Bezout' teoreemist ja jaguvuse definitsioonist järeldub järgmine tulemus.

Järeldus 9.59. *Element $c \in R$ on polünoomi $f(X) \in R[X]$ juur parajasti siis, kui $X - c \mid f(X)$.*

Tõestus. Tarvilikkus. Kui $f(c) = 0$, siis jääk $f(X)$ jagamisel polünoomiga $X - c$ on 0. Seega $f(X) = (X - c)q(X)$, kus $q(X) \in R[X]$. Järelikult $X - c \mid f(X)$.
Piisavus. Selle jätame läbimõtlemiseks lugejale. □

Polünoomi väärtuse leidmiseks kohal c võib muidugi kasutada definitsiooni 9.55 ja arvutada nii nagu näites 9.56, kuid tuleb välja, et on ka veidi lihtsam võimalus. Olgu polünoomi $f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$ ja polünoomi $X - c$ jagatis

$$q(X) = b_0X^{n-1} + b_1X^{n-2} + \ldots + b_{n-2}X + b_{n-1}$$

ning jääk $r \in R$. Siis kehtib võrdus

$$f(X) = (X - c)q(X) + r.$$

Selle võrduse paremal poolel on polünoom

$$b_0X^n + (b_1 - cb_0)X^{n-1} + (b_2 - cb_1)X^{n-2} + \ldots + (b_{n-1} - cb_{n-2})X + (r - cb_{n-1}).$$

See polünoom peab võrduma polünoomiga $f(X)$, mis tähendab, et X vastavate astmete kordajad peavad olema samad. Seega peavad kehtima võrdused

$$\begin{array}{rcl} a_0 &=& b_0, \\ a_1 &=& b_1 - cb_0, \\ a_2 &=& b_2 - cb_1, \\ &\ldots& \\ a_{n-1} &=& b_{n-1} - cb_{n-2}, \\ a_n &=& r - cb_{n-1}, \end{array} \quad \text{ehk} \quad \begin{array}{rcl} b_0 &=& a_0, \\ b_1 &=& a_1 + cb_0, \\ b_2 &=& a_2 + cb_1, \\ &\ldots& \\ b_{n-1} &=& a_{n-1} + cb_{n-2}, \\ r &=& a_n + cb_{n-1}. \end{array}$$

Jagatise $q(X)$ ja jäägi r leidmist nende võrduste abil kutsutakse **Horneri**[10] **skeemiks**. Harilikult esitatakse Horneri skeem järgneva tabeli kujul:

$$\begin{array}{c|cccccc} & a_0 & a_1 & a_2 & \ldots & a_{n-1} & a_n \\ c & & cb_0 & cb_1 & \ldots & cb_{n-2} & cb_{n-1} \\ \hline & b_0 & b_1 & b_2 & \ldots & b_{n-1} & r \end{array}.$$

Selle tabeli teise rea elemendid saadakse kolmanda rea elementide korrutamisel c-ga ja joone all olevad elemendid saadakse joone kohal olevate elementide liitmisel. Tihti jäetakse selle tabeli teine rida üldse ära.

[10] William George Horner (1786–1837) — inglise matemaatik

Näide 9.60. Rakendame Horneri skeemi näites 9.56 olnud polünoomi $f(X) = X^3 + \overline{3}X + \overline{4} \in \mathbb{Z}_5[X]$ jaoks:

$$\begin{array}{c|cccc} & \overline{1} & \overline{0} & \overline{3} & \overline{4} \\ \overline{2} & & \overline{2} & \overline{4} & \overline{4} \\ \hline & \overline{1} & \overline{2} & \overline{2} & \overline{3} \end{array},$$

kust näeme, et $f(\overline{2}) = \overline{3}$.

Polünoomi juure korral võib rääkida selle kordsusest.

Definitsioon 9.61. Olgu k naturaalarv. Elementi $c \in R$ nimetatakse polünoomi $0 \neq f(X) \in R[X]$ **k-kordseks juureks**, kui $(X-c)^k \mid f(X)$, aga $(X-c)^{k+1} \nmid f(X)$ ringis $R[X]$.

Näide 9.62. Arv 1 on polünoomi $f(X) = (X-1)^2(X+2) = X^3 - 3X + 2 \in \mathbb{Z}[X]$ kahekordne juur ja arv -2 on selle polünoomi ühekordne juur.

Näide 9.63. Polünoomi $f(X) = X^4 - 4X^3 + 5X^2 - 4X + 4 \in \mathbb{R}[X]$ saab esitada kujul

$$f(X) = (X-2)^2(X^2+1).$$

Siit näeme, et sellel polünoomil on kahekordne juur 2. Kui vaatleksime seda polünoomi üle korpuse $\mathbb{C}$, siis oleks tal veel ka ühekordsed juured i ja $-i$.

Lause 9.64. *Olgu $0 \neq f(X) \in R[X]$, kus R on nullitegureita kommutatiivne ring. Olgu $c_1, \ldots, c_m$ polünoomi $f(X)$ vastavalt $k_1, \ldots, k_m$-kordsed juured, mis on paarikaupa erinevad. Siis leidub selline polünoom $g(X) \in R[X]$, et*

$$f(X) = (X-c_1)^{k_1} \ldots (X-c_m)^{k_m} g(X),$$

kusjuures ükski elementidest $c_1, \ldots, c_m$ ei ole polünoomi $g(X)$ juur. Lisaks sellele

$$k_1 + \ldots + k_m \leqslant \deg(f(X)).$$

Selle lause teine pool sõnastatakse tihti ka kujul: *polünoomi $f(X)$ juurte koguarv (kordsust arvestades) ei ületa selle polünoomi astet.*

Tõestus. Tõestame väite induktsiooniga m järgi. Juhul $m = 1$ kehtib see juure kordsuse definitsiooni põhjal.

Eeldame nüüd, et väide kehtib $m-1$ juure korral ja näitame, et see kehtib ka m juure $c_1, \ldots, c_m$ korral, mille kordsused on vastavalt $k_1, \ldots, k_m$. Eelduse põhjal leidub polünoom $h(X)$ nii, et

$$f(X) = (X-c_1)^{k_1} \ldots (X-c_{m-1})^{k_{m-1}} h(X),$$

kusjuures ükski ringi R elementidest $c_1, \ldots, c_{m-1}$ ei ole polünoomi $h(X)$ juur. Kuna $c_m - c_1 \neq 0, \ldots, c_m - c_{m-1} \neq 0$ ja R on nullitegureita, siis $(c_m - c_1)^{k_1} \ldots (c_m - c_{m-1})^{k_{m-1}} \neq 0$. Võrdusest $f(c_m) = 0$ järeldub, et $h(c_m) = 0$. Järelduse 9.59 põhjal $X - c_m \mid h(X)$. Olgu s suurim naturaalarv, mille korral $(X-c_m)^s \mid h(X)$. Siis leidub $g(X) \in R[X]$ nii, et

$$h(X) = (X-c_m)^s g(X)$$

ja $g(c_m) \neq 0$. Kuna c_m on polünoomi $f(X)$ k_m-kordne juur, siis $s \leqslant k_m$. Samal põhjusel leidub polünoom $u(X) \in R[X]$ nii, et

$$f(X) = (X - c_m)^{k_m} u(X)$$

ja $u(c_m) \neq 0$. Oletame vastuväiteliselt, et $s < k_m$. Et ring $R[X]$ on nullitegureita, siis on ta taandamisega ja seega võrdusest

$$(X - c_m)^{k_m} u(X) = (X - c_1)^{k_1} \dots (X - c_{m-1})^{k_{m-1}} (X - c_m)^s g(X)$$

järeldub võrdus

$$(X - c_m)^{k_m - s} u(X) = (X - c_1)^{k_1} \dots (X - c_{m-1})^{k_{m-1}} g(X),$$

kus $k_m - s \geqslant 1$. Asendades c_m selle võrduse mõlemal poolel olevasse polünoomi saame, et

$$0 = (c_m - c_1)^{k_1} \dots (c_m - c_{m-1})^{k_{m-1}} g(c_m),$$

kust $g(c_m) = 0$, mis on vastuolus eelnevaga. Järelikult $s = k_m$, mis annabki meile võrduse

$$f(X) = (X - c_1)^{k_1} \dots (X - c_m)^{k_m} g(X).$$

Kuna $f(X) \neq 0$, siis ka $g(X) \neq 0$ ja seega $\deg(g(X)) \geqslant 0$. Lauset 9.6 kasutades saame

$$\deg(f(X)) = k_1 + \ldots + k_m + \deg(g(X)),$$

millest võrratuse $\deg(g(X)) \geqslant 0$ tõttu järeldub, et

$$\deg(f(X)) \geqslant k_1 + \ldots + k_m.$$

□

Meenutame, et lineaarpolünoomid on esimese astme polünoomid, see tähendab polünoomid kujul $aX + b$, kus $a \neq 0$.

Lause 9.65. *Olgu R nullitegureita kommutatiivne ring ja olgu $f(X) \in R[X]$ n-nda astme polünoom, kus $n \in \mathbb{N}$. Kui polünoomil $f(X)$ on ringis R kordsusi arvestades n juurt, siis $f(X)$ on esitatav lineaarpolünoomide korrutisena.*

Kui R on korpus ja polünoom $f(X)$ on esitatav lineaarpolünoomide korrutisena, siis on polünoomil $f(X)$ korpuses R n juurt.

Tõestus. Olgu $\deg(f(X)) = n \in \mathbb{N}$ ja olgu $c_1, \dots, c_n$ polünoomi $f(X)$ juured (siin võib $c_1, \dots, c_n$ hulgas olla võrdseid elemente). Lause 9.64 põhjal leidub polünoom $g(X) \in R[X]$ nii, et

$$f(X) = (X - c_1) \dots (X - c_n) g(X).$$

Tänu lausele 9.6 peab $g(X)$ olema nullist erinev konstantne polünoom. Seega $f(X)$ on lineaarpolünoomide korrutis.

Oletame nüüd, et R on korpus ja et $f(X)$ on esitatav lineaarpolünoomide korrutisena. Kuna $f(X)$ aste on n, siis peab ka neid lineaarpolünoome olema n tükki, s.t.

$$f(X) = (a_1 X + b_1)(a_2 X + b_2) \dots (a_n X + b_n).$$

Siit näeme, et polünoomil $f(X)$ on korpuses R n juurt $c_i := -a_i^{-1} b_i$, $i \in \{1, \dots, n\}$. □

Nagu eespool nägime, ei pruugi n-nda astme reaalarvuliste kordajatega polünoomil olla n reaalarvulist juurt. Osutub, et kui reaalarvude asemel vaadelda kompleksarve, siis see probleem kaob.

Teoreem 9.66 (Algebra põhiteoreem). *Kui $n \in \mathbb{N}$ ja $f(X)$ on n-nda astme polünoom üle korpuse $\mathbb{C}$, siis on tal (kordsusi arvestades) n kompleksarvulist juurt.*

Algebra põhiteoreemi tõestus ei mahu käesoleva kursuse raamesse. Selle võib leida näiteks raamatust [1], lk. 222.

9.10. Lagrange'i interpolatsioonivalem

Järgmine lause annab ühe piisava tingimuse kahe polünoomi võrdumiseks.

Lause 9.67. *Olgu R nullitegureita kommutatiivne ring, olgu $c_0, c_1, \ldots, c_n \in R$ paarikaupa erinevad elemendid ja olgu $f, g \in R[X]$ sellised polünoomid, et $\deg(f), \deg(g) \leqslant n$. Kui $f(c_i) = g(c_i)$ iga $i \in \{0, 1, \ldots, n\}$ korral, siis $f = g$.*

Tõestus. Polünoomi $h := f - g$ aste ei ole suurem kui n. Kuna $h(c_i) = f(c_i) - g(c_i) = 0$ iga $i \in \{0, 1, \ldots, n\}$ korral, siis polünoomil h on $n + 1$ juurt. Tänu lausele 9.64 ei ole võimalik, et h on nullist erinev polünoom. Seega $h = 0$ ehk $f = g$. □

Lause 9.68. *Olgu K korpus ja $b_0, \ldots, b_n, c_0, \ldots, c_n \in K$, kusjuures $c_0, \ldots, c_n$ on paarikaupa erinevad. Siis leidub täpselt üks ülimalt n-nda astme polünoom $f \in K[X]$ nii, et*

$$f(c_0) = b_0,\ f(c_1) = b_1,\ \ldots,\ f(c_n) = b_n.$$

Tõestus. Eelmine lause ütleb, et selliseid polünoome ei saa olla rohkem kui üks. Tuleb välja, et neid on ka vähemalt üks. Selliseks polünoomiks on

$$f = \sum_{i=0}^{n} b_i \cdot \frac{(X - c_0) \ldots (X - c_{i-1})(X - c_{i+1}) \ldots (X - c_n)}{(c_i - c_0) \ldots (c_i - c_{i-1})(c_i - c_{i+1}) \ldots (c_i - c_n)}. \tag{39}$$

Kuna $c_0, \ldots, c_n$ on paarikaupa erinevad, siis

$$d_i := (c_i - c_0) \ldots (c_i - c_{i-1})(c_i - c_{i+1}) \ldots (c_i - c_n) \in K \setminus \{0\}.$$

Seega d_i on pööratav element ja

$$f = \sum_{i=0}^{n} b_i d_i^{-1} (X - c_0) \ldots (X - c_{i-1})(X - c_{i+1}) \ldots (X - c_n)$$

on summa n-nda astme polünoomidest, seega $\deg(f) \leqslant n$. Kui $j \in \{0, 1, \ldots, n\}$, siis $f(c_j)$ arvutamisel muutub summas b_j kordaja 1-ks ja kõik ülejäänud b_i-de kordajad nullideks, sest murru lugejas esineb $c_j - c_j$. Seega $f(c_j) = b_j$, mida oligi tarvis tõestada. □

Valemit (39) kutsutakse **Lagrange'i interpolatsioonivalemiks**.

Interpolatsiooniülesande graafiline tõlgendus on järgmine: meil on ristkoordinaatidega tasandil antud $n + 1$ punkti

$$(c_0, b_0), (c_1, b_1), \ldots, (c_n, b_n)$$

ja eesmärgiks on leida ülimalt n-nda astme polünoom, mille graafik kõiki neid punkte läbiks. Nagu nägime, selline ülesanne on üheselt lahenduv.

9.11. Kordsete tegurite eraldamine

Selles paragrahvis vaatleme lihtsuse mõttes polünoome üle korpuse $\mathbb{R}$. Samasugused tulemused kehtivad siiski ka üldisemal juhul, muuhulgas korpuste $\mathbb{Q}$ ja $\mathbb{C}$ korral.

Definitsioon 9.69. Olgu $f \in \mathbb{R}[X]$, olgu $p \in \mathbb{R}[X]$ taandumatu polünoom ja $k \in \mathbb{N}$. Polünoomi p nimetatakse polünoomi f **k-kordseks teguriks**, kui $p^k \mid f$, aga $p^{k+1} \nmid f$ ringis $\mathbb{R}[X]$.

Definitsioon 9.70. Olgu $f = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n \in \mathbb{R}[X]$. Polünoomi f **tuletiseks** nimetatakse polünoomi

$$f' = na_0X^{n-1} + (n-1)a_1X^{n-2} + \ldots + a_{n-1} \in \mathbb{R}[X].$$

Lihtne on veenduda, et tuletise leidmisel kehtivad järgmised reeglid:

$$\begin{aligned} (f+g)' &= f' + g', \\ (fg)' &= f'g + fg', \\ (f^k)' &= kf^{k-1}f' \end{aligned}$$

mistahes $f, g \in \mathbb{R}[X]$ ja $k \in \mathbb{N}$ korral.

Definitsioonist on näha, et kui $\deg(f) = n$, siis $\deg(f') = n - 1$.

Teoreem 9.71. *Kui taandumatu polünoom $p \in \mathbb{R}[X]$ on polünoomi $f \in \mathbb{R}[X]$ k-kordne tegur, siis on ta tuletise f' $(k-1)$-kordne tegur.*

Tõestus. Olgu $f, p \in \mathbb{R}[X]$, p taandumatu, $p^k \mid f$ ja $p^{k+1} \nmid f$. Siis leidub selline $g \in \mathbb{R}[X]$, et $f = p^kg$ ja $p \nmid g$. Siis aga

$$f' = kp^{k-1}p'g + p^kg' = p^{k-1}(kp'g + pg'),$$

kust näeme, et $p^{k-1} \mid f'$.

Jääb veel näidata, et $p^k \nmid f'$. Oletame vastuväiteliselt, et $p^k \mid f'$, s.t. et leidub $h \in \mathbb{R}[X]$ nii, et

$$p^{k-1}(kp'g + pg') = p^kh.$$

Taandades polünoomi p^{k-1} saame võrduse $kp'g + pg' = ph$. Nüüd $p \mid ph - pg' = (kp')g$. Kuna $\mathbb{R}[X]$ on Eukleidese ring, siis on ta faktoriaalne. Teoreem 9.38 ütleb, et $p \mid kp'$ või $p \mid g$. Esimene võimalus on vastuolus sellega, et $\deg(p) > \deg(kp')$ ja teine võimalus sellega, et $p \nmid g$. Saadud vastuolu näitab, et $p^k \nmid f'$. □

Lineaarsel polünoomil üle $\mathbb{R}$ on üks ühekordne juur, seega kordseid juuri ei ole.

Vaatleme edasises polünoomi $f \in \mathbb{R}[X]$, mille aste on vähemalt 2. Tuletame meelde (vt. järeldust 9.41), et sellise polünoomi saab tegurite järjekorra täpsuseni üheselt esitada kujul

$$f = ap_1^{k_1}p_2^{k_2}\ldots p_m^{k_m} \in \mathbb{R}[X],$$

kus $a \in \mathbb{R}$, $k_1, \ldots, k_m \in \mathbb{N}$, $p_1, \ldots, p_m \in \mathbb{R}[X]$ on taandumatud polünoomid ja $\text{SÜT}(p_i, p_j) = 1$, kui $i \neq j$.

Lemma 9.72. *Kehtib*

$$\text{SÜT}(f, f') = ap_1^{k_1-1}p_2^{k_2-1}\ldots p_m^{k_m-1}.$$

TÕESTUS. Ka polünoom f' esitub taandumatute polünoomide astmete korrutisena, kusjuures tänu teoreemile 9.71 peavad nende hulgas olema $p_1^{k_1-1}, \ldots, p_m^{k_m-1}$. Seega leidub polünoom $h \in \mathbb{R}[X]$ nii, et

$$f' = p_1^{k_1-1} \ldots p_m^{k_m-1} h.$$

Tähistame

$$d := p_1^{k_1-1} \ldots p_m^{k_m-1}$$

Piisab tõestada, et d on f ja f' suurim ühistegur, sest polünoomid d ja $ap_1^{k_1-1} \ldots p_m^{k_m-1}$ on assotsieeritud. Me kasutame suurima ühisteguri definitsiooni. On selge, et $d \mid f$ ja $d \mid f'$.

Oletame, et ka polünoom d_1 on selline, et $d_1 \mid f$ ja $d_1 \mid f'$. Siis $d_1 \neq 0$, sest $f \neq 0$. Kui d_1 on konstantne, siis $d_1 \mid d$.

Kui d_1 on mittekonstantne, siis sellest, et $d_1 \mid f$, järeldub, et d_1 taandumatud tegurid peavad olema hulgast $\{p_1, \ldots, p_m\}$. Seega $d_1 = bp_1^{l_1} \ldots p_m^{l_m}$, kus $b \in \mathbb{R}$ ja $l_1, \ldots, l_m \in \mathbb{N} \cup \{0\}$, kusjuures $l_i \leqslant k_i$ iga $i \in \{1, \ldots, m\}$ korral. Kui oletada, et leidub mingi $j \in \{1, \ldots, m\}$ nii, et $l_j = k_j$, siis $p_j^{k_j} \mid d_1 \mid f'$, millest $p_j^{k_j} \mid f'$. Viimane on vastuolus sellega, et p_j on polünoomi f' $(k_j - 1)$-kordne tegur. Seega

$$l_1 \leqslant k_1 - 1, \ldots, l_m \leqslant k_m - 1,$$

kust

$$d_1 \left(b^{-1} p_1^{k_1-1-l_1} \ldots p_m^{k_m-1-l_m} \right) = p_1^{k_1-1} \ldots p_m^{k_m-1} = d$$

ehk $d_1 \mid d$. □

Niisiis $\text{SÜT}(f, f') \mid f$ ning polünoomide f ja $\text{SÜT}(f, f')$ jagatiseks on polünoom

$$\frac{f}{\text{SÜT}(f, f')} = p_1 p_2 \ldots p_m =: g,$$

millel on samad taandumatud tegurid nagu polünoomil f, aga kordsusega 1. Sellise polünoomi g leidmist nimetatakse polünoomi f **kordsete tegurite eraldamiseks**.

Niisiis: *polünoomi kordsete tegurite eraldamiseks tuleb see polünoom jagada tema ja tema tuletise suurima ühisteguriga.* Suurim ühistegur leitakse reeglina Eukleidese algoritmi abil.

Kui c on polünoomi f juur, siis taandumatu polünoom $X - c$ jagab nii polünoomi f kui ka polünoomi g. Seega c on ka polünoomi g juur. Kui meil õnnestub leida polünoomi g juured, siis oleme leidnud ka polünoomi f juured. Kuna üldiselt on polünoomi g aste väiksem kui polünoomi f aste, siis on juurte leidmine polünoomi g põhjal lihtsam kui esialgse polünoomi f põhjal.

9.12. Polünoomi juured ja tuletised

Selles paragrahvis tõestame tulemuse, mis lubab polünoomi juurte kordsusi kindlaks teha tuletise abil. Alustuseks tõestame abitulemuse.

Lemma 9.73. *Mistahes polünoomi $f \in \mathbb{R}[X]$, elemendi $c \in \mathbb{R}$ ja naturaalarvude k, l korral*

$$(X - c)^l \mid f' \ \wedge \ (X - c)^k \mid f \ \wedge \ k \leqslant l \implies (X - c)^{k+1} \mid f.$$

TÕESTUS. Eeldame, et $(X - c)^l \mid f'$, $(X - c)^k \mid f$ ja $k \leqslant l$. Siis leiduvad polünoomid g ja h nii, et

$$f = (X - c)^k g \ \text{ ja } \ f' = (X - c)^l h.$$

Kasutades korrutise tuletise leidmise eeskirja saame

$$k(X-c)^{k-1}g + (X-c)^k g' = f' = (X-c)^l h.$$

Taandades polünoomi $(X-c)^{k-1}$ saame võrduse

$$kg + (X-c)g' = (X-c)^{l-k+1}h.$$

(Paneme tähele, et $l-k+1 \geqslant 1$.) Avaldades polünoomi g näeme, et

$$X - c \mid \frac{1}{k}(X-c)^{l-k+1}h - \frac{1}{k}(X-c)g' = g.$$

Kuna $X-c \mid g$, siis $(X-c)^{k+1} \mid f$. □

Järeldus 9.74. *Mistahes polünoomi $f \in \mathbb{R}[X]$, elemendi $c \in \mathbb{R}$ ja naturaalarvu l korral*

$$(X-c)^l \mid f' \wedge (X-c) \mid f \implies (X-c)^{l+1} \mid f.$$

Tõestus. Eeldame, et $(X-c)^l \mid f'$. Kasutades korduvalt eelmist lauset saame, et

$$(X-c) \mid f \implies (X-c)^2 \mid f \implies (X-c)^3 \mid f \implies \ldots \implies (X-c)^l \mid f \implies (X-c)^{l+1} \mid f.$$

□

Teoreem 9.75. *Element $c \in \mathbb{R}$ on polünoomi $f \in \mathbb{R}[X]$ k-kordne juur parajasti siis, kui*

$$f(c) = f^{(1)}(c) = \ldots = f^{(k-1)}(c) = 0 \quad \textit{ja} \quad f^{(k)}(c) \neq 0.$$

(Siin $f^{(j)}$ tähistab polünoomi f j-ndat tuletist.)

Tõestus. Tarvilikkus. See järeldub teoreemist 9.71.
Piisavus. Eeldame, et $f(c) = f^{(1)}(c) = \ldots = f^{(k-1)}(c) = 0$ ja $f^{(k)}(c) \neq 0$. Siis $X-c$ jagab polünoome $f, f^{(1)}, f^{(2)}, \ldots, f^{(k-1)}$. Kasutades korduvalt järeldust 9.74 saame, et

$$(X-c) \mid f^{(k-1)} \wedge (X-c) \mid f^{(k-2)} \implies (X-c)^2 \mid f^{(k-2)} \implies (X-c)^3 \mid f^{(k-3)} \implies \ldots$$
$$\implies (X-c)^{k-1} \mid f^{(1)} \implies (X-c)^k \mid f.$$

Oletame vastuväiteliselt, et $(X-c)^{k+1} \mid f$. Siis teoreemi 9.71 põhjal $(X-c) \mid f^{(k)}$ ehk $f^{(k)}(c) = 0$, vastuolu. Seega $(X-c)^{k+1} \nmid f$. Oleme tõestanud, et c on f k-kordne juur. □

10. Lineaarkujutused

10.1. Lineaarkujutuse definitsioon

Algebraliste struktuuride uurimisel on suur tähtsus tehteid säilitavatel kujutustel sama tüüpi struktuuride vahel. Meie vaatleme selliseid kujutusi vektorruumide korral.

Definitsioon 10.1. Olgu V_1 ja V_2 vektorruumid üle korpuse K. Kujutust $\varphi : V_1 \to V_2$ nimetatakse **lineaarkujutuseks**, kui

LK1. $\varphi(a+b) = \varphi(a) + \varphi(b)$ iga $a, b \in V_1$ korral (s.t. φ säilitab liitmist);

LK2. $\varphi(ka) = k\varphi(a)$ iga $a \in V_1$ ja $k \in K$ korral (s.t. φ säilitab skalaaridega korrutamist).

Definitsioon 10.2. Lineaarkujutust vektorruumist V iseendasse nimetatakse vektorruumi V **lineaarteisenduseks**.

Kõigi lineaarkujutuste hulka vektorruumist V_1 vektorruumi V_2 tähistatakse sümboliga $\mathrm{Hom}(V_1, V_2)$. Vektorruumi V kõigi lineaarteisenduste hulka tähistatakse sümboliga $\mathrm{End}(V)$. (Hom tuleb sõnast "homomorfism", End sõnast "endomorfism".)

Näide 10.3. 1. Iga vektorruumi V samasusteisendus 1_V on lineaarteisendus.

2. Mistahes vektorruumide V_1 ja V_2 korral on kujutus $V_1 \to V_2$, mis viib kõik V_1 vektorid V_2 nullvektoriks, lineaarkujutus. Sellist kujutust nimetatakse **nullkujutuseks** ja tähistatakse tihti sümboliga 0.

3. Olgu $A \in \mathrm{Mat}_{m,n}(K)$ fikseeritud maatriks. Kujutus $\varphi : K^n \to K^m$, mis on defineeritud võrdusega

$$\varphi(x) := Ax,$$

$x \in K^n$, on lineaarkujutus. Siin me samastame vektori $x \in K^n$ talle vastava veeruvektoriga.

4. Tasandi vabavektorite vektorruumi $\mathbb{E}_2$ üheks lineaarteisenduseks on teisendus, mis peegeldab iga vektorit fikseeritud koordinaatsüsteemi y-telje suhtes.

5. Kui defineerida polünoomi $a_nX^n + \ldots + a_1X + a_0 \in \mathbb{R}[X]$ ja reaalarvu k korrutis võrdusega $k(a_nX^n + \ldots + a_1X + a_0) := ka_nX^n + \ldots + ka_1X + ka_0$, siis on hulk $\mathbb{R}[X]$ vektorruum üle korpuse $\mathbb{R}$. Diferentseerimiskujutus

$$\varphi : \mathbb{R}[X] \to \mathbb{R}[X], \quad f(X) \mapsto f'(X),$$

on lineaarne.

Lineaarkujutuse definitsioonist järeldub, et lineaarkujutusel on veel teisigi omadusi.

Lause 10.4. *Olgu $\varphi : V_1 \to V_2$ lineaarkujutus. Siis*

1. $\varphi(0) = 0$;
2. $\varphi(-a) = -\varphi(a)$ *iga* $a \in V_1$ *korral.*
3. $\varphi(a-b) = \varphi(a) - \varphi(b)$ *iga* $a, b \in V_1$ *korral.*

Teiste sõnadega: lineaarkujutus säilitab nullelementi, vastandelemendi võtmist ja lahutamist.

TÕESTUS. 1. Definitsiooni põhjal

$$\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0).$$

Liites selle võrduse mõlemale poolele $-\varphi(0)$ saamegi võrduse $0 = \varphi(0)$.

2. Kuna

$$\varphi(a) + \varphi(-a) = \varphi(a-a) = \varphi(0) = 0,$$

siis vastandelemendi definitsiooni põhjal $\varphi(-a) = -\varphi(a)$ iga $a \in V_1$ korral.

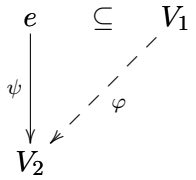
3. Kui $a, b \in V_1$, siis osas 2 tõestatu põhjal

$$\varphi(a-b) = \varphi(a+(-b)) = \varphi(a) + \varphi(-b) = \varphi(a) + (-\varphi(b)) = \varphi(a) - \varphi(b).$$

□

Tuleb välja, et lineaarkujutuse defineerimiseks piisab, kui näitame ära, kuidas see kujutus tegutseb baasivektoritel.

Lause 10.5. *Olgu V_1 ja V_2 vektorruumid üle korpuse K, olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V_1 baas ja olgu $\psi : e \to V_2$ suvaline kujutus. Siis leidub üheselt määrtatud lineaarkujutus $\varphi : V_1 \to V_2$ nii, et $\varphi(e_i) = \psi(e_i)$ iga $i \in \{1, \ldots, n\}$ korral.*



TÕESTUS. Defineerime kujutuse $\varphi : V_1 \to V_2$ järgmiselt:

$$\varphi(a) = \varphi\left(\sum_{i=1}^{n} a_i e_i\right) := \sum_{i=1}^{n} a_i \psi(e_i)$$

mistahes vektori $a = \sum_{i=1}^{n} a_i e_i$ korral. Kuna vektorruumi V_1 iga element a on tänu lausele 6.44 üheselt esitatav kujul $a = \sum_{i=1}^{n} a_i e_i$, siis on see definitsioon korrektne. On selge, et $\varphi(e_i) = \psi(e_i)$ iga $i \in \{1, \ldots, n\}$ korral.

Kujutus φ on lineaarkujutus, sest mistahes vektorite $a = \sum_{i=1}^{n} a_i e_i$ ja $b = \sum_{i=1}^{n} b_i e_i$ ning skalaari $k \in K$ korral

$$\begin{aligned}
\varphi\left(\sum_{i=1}^{n} a_i e_i + \sum_{i=1}^{n} b_i e_i\right) &= \varphi\left(\sum_{i=1}^{n} (a_i + b_i) e_i\right) = \sum_{i=1}^{n} (a_i + b_i)\psi(e_i) \\
&= \sum_{i=1}^{n} a_i \psi(e_i) + \sum_{i=1}^{n} b_i \psi(e_i) = \varphi\left(\sum_{i=1}^{n} a_i e_i\right) + \varphi\left(\sum_{i=1}^{n} b_i e_i\right), \\
\varphi\left(k \sum_{i=1}^{n} a_i e_i\right) &= \varphi\left(\sum_{i=1}^{n} (k a_i) e_i\right) = \sum_{i=1}^{n} (k a_i)\psi(e_i) = k \sum_{i=1}^{n} a_i \psi(e_i) = k\varphi\left(\sum_{i=1}^{n} a_i e_i\right).
\end{aligned}$$

Nende võrduste juures kasutasime φ definitsiooni ja erinevaid vektorruumi omadusi.

Kui ka $\chi : V_1 \longrightarrow V_2$ on selline lineaarkujutus, et $\chi(e_i) = \psi(e_i)$ iga $i \in \{1, \ldots, n\}$ korral, siis

$$\varphi(a) = \sum_{i=1}^{n} a_i \psi(e_i) = \sum_{i=1}^{n} a_i \chi(e_i) = \chi\left(\sum_{i=1}^{n} a_i e_i\right) = \chi(a)$$

ja seega $\varphi = \chi$. □

10.2. Lineaarkujutuse tuum ja kujutis

Iga lineaarkujutusega on loomulikul viisil seotud kaks alamruumi.

Definitsioon 10.6. Lineaarkujutuse $\varphi : V_1 \to V_2$

1. **tuumaks** nimetatakse hulka

$$\operatorname{Ker}\varphi = \{a \in V_1 \mid \varphi(a) = 0\} \subseteq V_1,$$

2. **kujutiseks** nimetatakse hulka

$$\operatorname{Im}\varphi = \{\varphi(a) \mid a \in V_1\} \subseteq V_2.$$

Lause 10.7. *Olgu V_1 ja V_2 vektorruumid üle korpuse K. Lineaarkujutuse $\varphi : V_1 \to V_2$*

1. *tuum on vektorruumi V_1 alamruum,*

2. *kujutis on vektorruumi V_2 alamruum.*

Tõestus. Kuna $\varphi(0) = 0$, siis V_1 nullvektor kuulub hulka $\operatorname{Ker}\varphi$ ja V_2 nullvektor hulka $\operatorname{Im}\varphi$. Seega need hulgad on mittetühjad.
1. Olgu $a, b \in \operatorname{Ker}\varphi$ ja $k \in K$. Siis

$$\varphi(a+b) = \varphi(a) + \varphi(b) = 0 + 0 = 0$$

ja

$$\varphi(ka) = k\varphi(a) = k0 = 0.$$

Järelikult $a + b, ka \in \operatorname{Ker}\varphi$ ja $\operatorname{Ker}\varphi$ on vektorruumi V_1 alamruum.
2. Selle osa tõestusest jätame läbimõtlemiseks lugejale. □

Osutub, et tuuma põhjal saab kindlaks teha, millal on lineaarkujutus üksühene.

Lause 10.8. *Lineaarkujutus $\varphi : V_1 \to V_2$ on üksühene parajasti siis, kui* $\operatorname{Ker}\varphi = \{0\}$.

Tõestus. Tarvilikkus. Olgu φ üksühene ja $a \in \operatorname{Ker}\varphi$. Kuna $\varphi(a) = 0 = \varphi(0)$, siis üksühesuse tõttu $a = 0$. Seega $\operatorname{Ker}\varphi \subseteq \{0\}$. Kuna vastupidine sisalduvus on ilmne, siis peabki kehtima võrdus $\operatorname{Ker}\varphi = \{0\}$.
Piisavus. Eeldame, et $\operatorname{Ker}\varphi = \{0\}$. Kehtigu võrdus $\varphi(a) = \varphi(b)$, $a, b \in V_1$. Siis

$$\varphi(a-b) = \varphi(a) - \varphi(b) = 0.$$

Järelikult $a - b \in \operatorname{Ker}\varphi = \{0\}$, s.t. $a - b = 0$ ehk $a = b$. Sellega oleme näidanud, et φ on üksühene. □

Lause 10.9. *Lineaarkujutus $\varphi : V_1 \to V_2$ on pealekujutus parajasti siis, kui* $\operatorname{Im}\varphi = V_2$.

Tõestus. See järeldub vahetult definitsioonidest. □

Definitsioon 10.10. Olgu V_1 ja V_2 vektoruumid üle korpuse K. Neid vektorruume nimetatakse **isomorfseteks**, kui leidub bijektiivne lineaarkujutus $f : V_1 \to V_2$. Tähistatakse $V_1 \cong V_2$.

Näide 10.11. Vektorruumid K^n ja $\mathrm{Mat}_{1n}(K)$ on isomorfsed. Selle isomorfismi realiseerib kujutus

$$\varphi : K^n \to \mathrm{Mat}_{1n}(K), \quad (k_1, k_2, \ldots, k_n) \mapsto (k_1\ k_2\ \ldots\ k_n).$$

Teoreem 10.12. *Iga n-mõõtmeline vektorruum üle korpuse K on isomorfne vektorruumiga K^n.*

Tõestus. Olgu $\dim V = n$ ja olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V baas. Defineerime kujutuse $\varphi : K^n \to V$ võrdusega

$$\varphi((k_1, \ldots, k_n)) := k_1e_1 + \ldots + k_ne_n.$$

Kui $(k_1, \ldots, k_n), (l_1, \ldots, l_n) \in K^n$, siis

$$\begin{aligned}
\varphi((k_1, \ldots, k_n) + (l_1, \ldots, l_n)) &= \varphi((k_1 + l_1, \ldots, k_n + l_n)) = (k_1 + l_1)e_1 + \ldots + (k_n + l_n)e_n \\
&= k_1e_1 + l_1e_1 + \ldots + k_ne_n + l_ne_n \\
&= k_1e_1 + \ldots + k_ne_n + l_1e_1 + \ldots + l_ne_n \\
&= \varphi((k_1, \ldots, k_n)) + \varphi((l_1, \ldots, l_n)).
\end{aligned}$$

Analoogiliselt saab näidata, et $\varphi(c(k_1, \ldots, k_n)) = c\varphi((k_1, \ldots, k_n))$ iga $c \in K$ korral. Seega φ on lineaarkujutus.

Kuna e on moodustajate süsteem, siis φ on pealekujutus. Kuna e on lineaarselt sõltumatu, siis

$$k_1e_1 + \ldots + k_ne_n = 0 \implies k_1 = \ldots k_n = 0,$$

mis näitab, et $\mathrm{Ker}\,\varphi$ koosneb ainult vektorruumi K^n nullvektorist $(0, \ldots, 0)$. Lause 10.8 põhjal on φ üksühene. □

Tõestatud teoreem on väga kasulik. Sisuliselt ütleb ta seda, et selle asemel, et teha arvutusi vektorruumi V vektoritega võib arvutusi teha nende koordinaatide vektoritega vektorruumis K^n. Sellele asjaolule põhineb näiteks suures osas analüütiline geomeetria: selle asemel, et arvutada suunatud sirglõikude ekvivalentsiklassidega tehakse arvutusi nende koordinaatide, s.t. järjestatud reaalarvupaaride või -kolmikutega.

Järeldus 10.13. *Kaks vektorit n-mõõtmelises vektorruumis V üle korpuse K on võrdsed parajasti siis, kui nende koordinaatide vektorid on võrdsed vektorruumis K^n.*

Tõestus. Olgu $\dim V = n$ ja olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V baas. Vaatleme vektoreid $a = a_1e_1 + \ldots + a_ne_n$ ja $b = b_1e_1 + \ldots + b_ne_n$ vektorruumis V. Kasutades eelmise teoreemi kujutust φ võime öelda, et

$$a = b \iff \varphi((a_1, \ldots, a_n)) = \varphi((b_1, \ldots, b_n)) \iff (a_1, \ldots, a_n) = (b_1, \ldots, b_n).$$

□

10.3. Lineaarkujutuse maatriks

Kui vektorruumides V_1 ja V_2 on fikseeritud mingid baasid, siis saab iga lineaarkujutusega $\varphi : V_1 \to V_2$ siduda teatud maatriksi.

Definitsioon 10.14. Olgu V_1 ja V_2 vektorruumid üle korpuse K, olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V_1 baas ja $e' = \{e'_1, \ldots, e'_m\}$ vektorruumi V_2 baas ning olgu $\varphi : V_1 \to V_2$ lineaarkujutus. **Lineaarkujutuse φ maatriksiks baaside e ja e' suhtes** nimetatakse maatriksit

$A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$, mille i-nda veeruvektori ($i \in \{1, \ldots, n\}$) komponendid on vektori $\varphi(e_i)$ koordinaadid baasi e' suhtes. Seda maatriksit tähistatakse sümboliga $A_\varphi^{e,e'}$.

Kui φ on vektorruumi V lineaarteisendus ja e on vektorruumi V baas, siis maatriksit $A_\varphi^{e,e}$ nimetatakse **lineaarteisenduse φ maatriksiks baasi e suhtes** ja tähistatakse A_φ^e.

Vastavalt sellele definitsioonile peavad lineaarkujutuse φ korral kehtima võrdused

$$\varphi(e_i) = a_{1i}e'_1 + \ldots + a_{mi}e'_m = \sum_{j=1}^{m} a_{ji}e'_j, \tag{40}$$

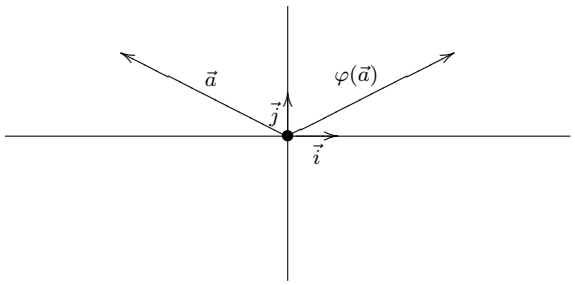
$i = 1, \ldots, n$.

Lineaarkujutuse maatriksi leidmiseks tuleb
1. leida vektorid $\varphi(e_1), \ldots, \varphi(e_n)$,
2. avaldada need baasi e' kaudu,
3. saadud koordinaate veergudesse paigutades moodustada maatriks A.

Näide 10.15. 1. Samasusteisenduse maatriks mistahes baasi suhtes on ühikmaatriks.

2. Nullkujutuse maatriks misthes baaside suhtes on nullmaatriks.

3. Olgu tasandi vabavektorite vektorruumis $\mathbb{E}_2$ fikseeritud mingi ristbaas $e = \{\vec{i}, \vec{j}\}$. Vaatleme lineaarteisendust $\varphi : \mathbb{E}_2 \to \mathbb{E}_2$, mis seisneb vektorite peegeldamises y-telje suhtes.



Kuna

$$\begin{aligned} \varphi(\vec{i}) &= -\vec{i} = (-1)\vec{i} + 0\vec{j}, \\ \varphi(\vec{j}) &= \vec{j} = 0\vec{i} + 1\vec{j}, \end{aligned}$$

siis

$$A_\varphi^e = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Tuleb välja, et lineaarkujutuse rakendamise vektorile võib taandada selle vektori koordinaatide veeru korrutamisele lineaarkujutuse maatriksiga. Vektori x koordinaatide veergu baasi $e = \{e_1, \ldots, e_n\}$ suhtes tähistame sümboliga $\overline{x}_e$. Seega

$$\overline{x}_e = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix} \iff x = x_1e_1 + \ldots + x_ne_n.$$

Lause 10.16. *Olgu V_1 ja V_2 vektorruumid üle korpuse K, olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V_1 baas ja $e' = \{e'_1, \ldots, e'_m\}$ vektorruumi V_2 baas ning olgu $\varphi : V_1 \to V_2$ lineaarkujutus. Siis iga vektori $x \in V_1$ korral*

$$\boxed{\overline{\varphi(x)}_{e'} = A_\varphi^{e,e'}\overline{x}_e.}$$

TÕESTUS. Olgu $x = x_1e_1 + \ldots + x_ne_n = \sum_{i=1}^n x_ie_i$, kus $x_1, \ldots, x_n \in K$, olgu $A_\varphi^{e,e'} = (a_{ij})$ ja $A_\varphi^{e,e'}\overline{x}_e = \begin{pmatrix} b_1 \\ \cdots \\ b_m \end{pmatrix}$. Siis

$$\begin{aligned}
\varphi(x) &= \varphi\left(\sum_{i=1}^n x_ie_i\right) \\
&= \sum_{i=1}^n x_i\varphi(e_i) && (\varphi \text{ on lineaarkujutus}) \\
&= \sum_{i=1}^n x_i\left(\sum_{j=1}^m a_{ji}e'_j\right) && (\text{võrdused (40)}) \\
&= \sum_{i=1}^n\sum_{j=1}^m x_ia_{ji}e'_j && (\text{SO1}) \\
&= \sum_{j=1}^m\sum_{i=1}^n x_ia_{ji}e'_j && (\text{SO3}) \\
&= \sum_{j=1}^m\left(\sum_{i=1}^n a_{ji}x_i\right)e'_j && (\text{märkus 6.3}) \\
&= \sum_{j=1}^m b_je'_j, && (\text{maatriksite korrutise definitsioon})
\end{aligned}$$

mida oligi tarvis tõestada. □

10.4. Lineaarkujutuste vektorruum

Lineaarkujutuste hulga $\mathrm{Hom}(V_1, V_2)$ saab loomulikult viisil muuta vektorruumiks.

Definitsioon 10.17. Olgu V_1 ja V_2 vektorruumid üle korpuse K, olgu $\varphi, \psi \in \mathrm{Hom}(V_1, V_2)$ ja $k \in K$. Defineerime kujutused $\varphi + \psi, k\varphi : V_1 \to V_2$ võrdustega

$$\begin{aligned}
(\varphi + \psi)(a) &:= \varphi(a) + \psi(a), \\
(k\varphi)(a) &:= k\varphi(a),
\end{aligned}$$

$a \in V_1$.

Selliste definitsioonide puhul öeldakse, et lineaarkujutuste liitmine ning lineaarkujutuse ja skalaari korrutis on defineeritud *punktiviisiliselt*. Selleks et leida $\varphi + \psi$ n.ö. punktis a, leiame φ ja ψ punktis a ning liidame tulemused. Analoogiliselt $k\varphi$ korral.

Lause 10.18. *Kui V_1 ja V_2 on vektorruumid üle korpuse K, siis hulk* $\mathrm{Hom}(V_1, V_2)$ *on vektorruum eespool defineeritud liitmise ja skalaaridega korrutamise suhtes.*

TÕESTUS. Kontrollime kõigepealt, et $\varphi+\psi$ ja $k\varphi$ on lineaarkujutused. Tõepoolest,

$$\begin{aligned}
(\varphi+\psi)(a+b) &= \varphi(a+b)+\psi(a+b)=\varphi(a)+\varphi(b)+\psi(a)+\psi(b)\\
&= \varphi(a)+\psi(a)+\varphi(b)+\psi(b)=(\varphi+\psi)(a)+(\varphi+\psi)(b),\\
(\varphi+\psi)(la) &= \varphi(la)+\psi(la)=l\varphi(a)+l\psi(a)=l(\varphi(a)+\psi(a))=l((\varphi+\psi)(a))\\
(k\varphi)(a+b) &= k(\varphi(a+b))=k(\varphi(a)+\varphi(b))=k\varphi(a)+k\varphi(b)=(k\varphi)(a)+(k\varphi)(b),\\
(k\varphi)(la) &= k(\varphi(la))=k(l\varphi(a))=(kl)\varphi(a)=(lk)\varphi(a)=l(k\varphi(a))=l((k\varphi)(a))
\end{aligned}$$

iga $a,b\in V_1$ ja $l\in K$ korral. Seega on meil tegemist algebraliste tehetega hulgal $\mathrm{Hom}(V_1,V_2)$.

Veendume, et on täidetud vektorruumi definitsiooni tingimused. Märgime esiteks, et hulk $\mathrm{Hom}(V_1,V_2)$ ei ole tühi, sest ta sisaldab nullkujutust $0:V_1\to V_2$.

VR1. Olgu $\varphi,\psi,\chi\in\mathrm{Hom}(V_1,V_2)$. Siis iga $a\in V_1$ korral

$$\begin{aligned}
((\varphi+\psi)+\chi)(a) &= (\varphi+\psi)(a)+\chi(a)=(\varphi(a)+\psi(a))+\chi(a)\\
&= \varphi(a)+(\psi(a)+\chi(a))=\varphi(a)+(\psi+\chi)(a)=(\varphi+(\psi+\chi))(a),
\end{aligned}$$

mis tähendab, et $(\varphi+\psi)+\chi=\varphi+(\psi+\chi)$. Analoogiliselt saab näidata, et $\varphi+\psi=\psi+\varphi$, s.t. et kehtib VR4.

VR2. Olgu $\varphi\in\mathrm{Hom}(V_1,V_2)$. Siis iga $a\in V_1$ korral

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

ehk $\varphi+0=\varphi$. See tähendab, et nullkujutus on nullelement lineaarkujutuste liitmise suhtes.

VR3. Olgu $\varphi\in\mathrm{Hom}(V_1,V_2)$. Defineerime kujutuse $-\varphi:V_1\to V_2$ võrdusega

$$(-\varphi)(a):=-\varphi(a),$$

$a\in V_1$. Siis iga $a\in V_1$ korral

$$(\varphi+(-\varphi))(a)=\varphi(a)+(-\varphi)(a)=\varphi(a)+(-\varphi(a))=0,$$

mis tähendab, et $\varphi+(-\varphi)=0$ ja seega $-\varphi$ on φ vastandelement liitmise suhtes.

VR5. Olgu $\varphi,\psi\in\mathrm{Hom}(V_1,V_2)$ ja $k\in K$. Siis iga $a\in V_1$ korral

$$\begin{aligned}
(k(\varphi+\psi))(a) &= k((\varphi+\psi)(a))=k(\varphi(a)+\psi(a))=k\varphi(a)+k\psi(a)\\
&= (k\varphi)(a)+(k\psi)(a)=(k\varphi+k\psi)(a)
\end{aligned}$$

ja seega $k(\varphi+\psi)=k\varphi+k\psi$.

Ülejäänud tingimuste kontroll on analoogiline. □

Nii saadud lineaarkujutuste vektorruumid ja maatriksite vektorruumid on omavahel väga tihedalt seotud.

Teoreem 10.19. *Olgu V_1 n-mõõtmeline ja V_2 m-mõõtmeline vektorruum üle korpuse K. Siis vektorruumid* $\mathrm{Hom}(V_1,V_2)$ *ja* $\mathrm{Mat}_{m,n}(K)$ *on isomorfsed.*

TÕESTUS. Olgu $e=\{e_1,\dots,e_n\}$ vektorruumi V_1 baas ja $e'=\{e'_1,\dots,e'_m\}$ vektorruumi V_2 baas. Defineerime kujutuse

$$f:\mathrm{Hom}(V_1,V_2)\longrightarrow\mathrm{Mat}_{m,n}(K)$$

võrdusega

$$f(\varphi) := A_{\varphi}^{e,e'},$$

$\varphi \in \mathrm{Hom}(V_1, V_2)$. Veendume, et f on lineaarkujutus.

LK1. Olgu $\varphi, \psi \in \mathrm{Hom}(V_1, V_2)$. Siis lause 6.46 põhjal

$$\overline{(\varphi+\psi)(e_i)}_{e'} = \overline{\varphi(e_i)+\psi(e_i)}_{e'} = \overline{\varphi(e_i)}_{e'} + \overline{\psi(e_i)}_{e'}$$

iga $i \in \{1, \ldots, n\}$ korral, s.t. maatriksi $A_{\varphi+\psi}^{e,e'}$ i-s veerg on maatriksite $A_{\varphi}^{e,e'}$ ja $A_{\psi}^{e,e'}$ i-ndate veergude summa. Järelikult $A_{\varphi+\psi}^{e,e'} = A_{\varphi}^{e,e'} + A_{\psi}^{e,e'}$ ja

$$f(\varphi+\psi) = A_{\varphi+\psi}^{e,e'} = A_{\varphi}^{e,e'} + A_{\psi}^{e,e'} = f(\varphi) + f(\psi).$$

LK2. Selle tingimuse kontroll on analoogiline.

Oletame nüüd, et $\varphi \in \mathrm{Ker}\, f$. Siis $f(\varphi) = A_{\varphi}^{e,e'}$ on nullmaatriks, mis tähendab, et $\varphi(e_i) = 0$ iga $i \in \{1, \ldots, n\}$ korral. Kui nüüd $a = a_1 e_1 + \ldots + a_n e_n \in V_1$ on suvaline vektor, siis

$$\varphi(a) = a_1\varphi(e_1) + \ldots + a_n\varphi(e_n) = 0 + \ldots + 0 = 0.$$

See tähendab, et φ on nullkujutus ja me oleme näidanud, et $\mathrm{Ker}\, f = \{0\}$. Lause 10.8 põhjal on f üksühene.

Olgu lõpuks $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(K)$. Tänu lausele 10.5 leidub selline lineaarkujutus $\varphi : V_1 \to V_2$, mille korral

$$\varphi(e_j) = \sum_{i=1}^{m} a_{ij} e_i',$$

kus $j = 1, \ldots, n$. Selle lineaarkujutuse korral $f(\varphi) = A$. Seega oleme näidanud, et f on pealekujutus ja kokkuvõttes on f vektorruumide isomorfism. □

10.5. Lineaarteisenduste ring

Olgu V vektorruum üle korpuse K. Vektorruumi V lineaarteisenduste korrutamine defineeritakse järjestrakendamise abil:

$$(\psi\varphi)(a) := \psi(\varphi(a))$$

$\psi, \varphi \in \mathrm{End}(V)$, $a \in V$.

Lause 10.20. *Hulk* $\mathrm{End}(V)$ *on ring lineaarteisenduste liitmise ja korrutamise suhtes.*

Tõestus. Veendume, et lineaarteisenduste korrutis on ka lineaarteisendus. Kui V on vektorruum üle korpuse K, $\varphi, \psi \in \mathrm{End}(V)$, $a, b \in V$ ja $k \in K$, siis

$$\begin{aligned}(\psi\varphi)(a+b) &= \psi(\varphi(a+b)) = \psi(\varphi(a)+\varphi(b)) = \psi(\varphi(a)) + \psi(\varphi(b)) = (\psi\varphi)(a) + (\psi\varphi)(b),\\ (\psi\varphi)(ka) &= \psi(\varphi(ka)) = \psi(k\varphi(a)) = k\psi(\varphi(a)) = k((\psi\varphi)(a)),\end{aligned}$$

mis tähendab, et $\psi\varphi \in \mathrm{End}(V)$.

Lause 10.18 tõestuses nägime, et $(\mathrm{End}(V), +)$ on Abeli rühm. Samuti teame, et hulga teisenduste korrutamine on assotsiatiivne ja samasusteisendus on selle korrutamise suhtes ühikelement. Seega jääb veel kontrollida, et kehtivad distributiivsuse seadused.

Olgugi $\varphi, \psi, \chi \in \mathrm{End}(V)$. Siis iga $a \in V$ korral

$$\begin{aligned}(\varphi(\psi+\chi))(a) &= \varphi((\psi+\chi)(a)) = \varphi(\psi(a)+\chi(a)) = \varphi(\psi(a)) + \varphi(\chi(a)) \\ &= (\varphi\psi)(a) + (\varphi\chi)(a) = (\varphi\psi + \varphi\chi)(a), \\ ((\psi+\chi)\varphi)(a) &= (\psi+\chi)(\varphi(a)) = \psi(\varphi(a)) + \chi(\varphi(a)) = (\psi\varphi)(a) + (\chi\varphi)(a) \\ &= (\psi\varphi + \chi\varphi)(a).\end{aligned}$$

Järelikult $\varphi(\psi+\chi) = \varphi\psi + \varphi\chi$ ja $(\psi+\chi)\varphi = \psi\varphi + \chi\varphi$. □

Lineaarteisenduste ringid on seotud ruutmaatriksite ringidega.

Teoreem 10.21. *Olgu V n-mõõtmeline vektorruum üle korpuse K. Siis ringid* $\mathrm{End}(V)$ *ja* $\mathrm{Mat}_n(K)$ *on isomorfsed.*

Tõestus. Olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V baas. Defineerime kujutuse

$$f : \mathrm{End}(V) \longrightarrow \mathrm{Mat}_n(K)$$

võrdusega

$$f(\varphi) := A^e_\varphi,$$

$\varphi \in \mathrm{End}(V)$. Teoreemi 10.19 tõestuses nägime, et selline kujutus on bijektiivne ja säilitab liitmist. On selge, et $f(1_V) = E$, kus ühikmaatriks E on ringi $\mathrm{Mat}_n(K)$ ühikelement.

Jääb veel näidata, et mistahes $\varphi, \psi \in \mathrm{End}(V)$ korral $f(\psi\varphi) = f(\psi)f(\varphi)$ ehk $A^e_\psi A^e_\varphi = A^e_{\psi\varphi}$. Maatriksi $A^e_{\psi\varphi}$ i-s veerg on definitsiooni järgi $\overline{(\psi\varphi)(e_i)}_e$. Lause 10.16 põhjal

$$\overline{(\psi\varphi)(e_i)}_e = \overline{\psi(\varphi(e_i))}_e = A^e_\psi \overline{\varphi(e_i)}_e = A^e_\psi(A^e_\varphi \overline{e_i}_e) = (A^e_\psi A^e_\varphi)\overline{e_i}_e,$$

kus $\overline{e_i}_e$ on baasivektori e_i koordinaatide veerg baasi e suhtes. Kuna

$$e_i = 0 \cdot e_1 + \ldots + 0 \cdot e_{i-1} + 1 \cdot e_i + 0 \cdot e_{i+1} + \ldots + 0 \cdot e_n,$$

siis maatriks $\overline{e_i}_e$ on veerg, mille i-ndal kohal on 1 ja kõik ülejäänud komponendid on nullid. Korrutades maatriksi $A^e_\psi A^e_\varphi$ paremalt sellise veeruga saame maatriksi $A^e_\psi A^e_\varphi$ i-nda veeru. Seega maatriksite $A^e_{\psi\varphi}$ ja $A^e_\psi A^e_\varphi$ vastavad veerud on võrdsed, mis tähendab, et ka need maatriksid on võrdsed. □

Võtame lõpuks veelkord lühidalt kokku tähtsamad seosed lineaarkujutuste ja maatriksite vahel. Niisiis

$$\boxed{A^{e,e'}_{\varphi+\psi} = A^{e,e'}_\varphi + A^{e,e'}_\psi,}$$

$$\boxed{A^{e,e'}_{k\varphi} = kA^{e,e'}_\varphi,}$$

$$\boxed{A^e_{\psi\varphi} = A^e_\psi A^e_\varphi.}$$

10.6. Sarnased maatriksid

Definitsioon 10.22. Maatrikseid $A, B \in \mathrm{Mat}_n(K)$ nimetatakse **sarnasteks** (ja kirjutatakse $A \sim B$), kui leidub selline regulaarne maatriks $T \in \mathrm{Mat}_n(K)$, et $B = T^{-1}AT$.

Lemma 10.23. *Maatriksite sarnasuse seos on ekvivalentsiseos hulgal* $\mathrm{Mat}_n(K)$.

Tõestus. Iga maatriksi $A \in \mathrm{Mat}_n(K)$ korral $A = E^{-1}AE$, kusjuures ühikmaatriks E on regulaarne. Seega $A \sim A$ ja seos $\sim$ on refleksiivne.

Olgu $A \sim B$. Siis $B = T^{-1}AT$, kus $T \in \mathrm{Mat}_n(K)$ on regulaarne maatriks. Järelikult

$$A = TBT^{-1} = (T^{-1})^{-1}BT^{-1},$$

kus ka T^{-1} on regulaarne. Seega $B \sim A$ ja seos $\sim$ on sümmeetriline.

Kui $A \sim B$ ja $B \sim C$, siis leiduvad sellised regulaarsed maatriksid $T, U \in \mathrm{Mat}_n(K)$, et $B = T^{-1}AT$ ja $C = U^{-1}BU$. Järelikult

$$C = U^{-1}BU = U^{-1}(T^{-1}AT)U = (U^{-1}T^{-1})A(TU) = (TU)^{-1}A(TU),$$

kus ka maatriks TU on regulaarne. See tähendab, et $A \sim C$ ja seos $\sim$ on transitiivne. □

Näide 10.24. Olgu

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{ja} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Siis

$$T^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

ja

$$T^{-1}AT = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -4 \\ 3 & 7 \end{pmatrix}.$$

Seega maatriksid

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{ja} \quad \begin{pmatrix} -2 & -4 \\ 3 & 7 \end{pmatrix}$$

on sarnased.

Ühes vektorruumis võib vaadelda mitut erinevat baasi. Nendevahelisi seoseid kirjeldab üleminekumaatriks, mille kohe defineerime.

Definitsioon 10.25. Olgu V n-mõõtmeline vektorruum üle korpuse K ja olgu $e = \{e_1, \ldots, e_n\}$, $e' = \{e'_1, \ldots, e'_n\}$ vektorruumi V kaks baasi. **Üleminekumaatriksiks** baasilt e baasile e' nimetatakse maatriksit, mille i-ndas veerus ($i \in \{1, \ldots, n\}$) on vektori e'_i koordinaadid baasi e suhtes. Seda maatriksit tähistatakse $T^{e,e'}$.

Seega kui $T^{e,e'} = (t_{ij}) \in \mathrm{Mat}_n(K)$, siis kehtivad seosed

$$e'_i = t_{1i}e_1 + \ldots + t_{ni}e_n = \sum_{j=1}^{n} t_{ji}e_j, \tag{41}$$

$i \in \{1, \ldots, n\}$.

Lemma 10.26. *Üleminekumaatriks ühelt baasilt teisele on regulaarne.*

TÕESTUS. Vaatleme üleminekumaatriksit $T := T^{e,e'} = (t_{ij})$. Olgu $T^1, T^2, \ldots, T^n$ üheveerulised maatriksid, mille elementideks on T vastava veeru elemendid. Tänu lause 7.13 analoogile veergude jaoks võime öelda, et maatriks $T^{e,e'}$ on regulaarne, kui tema veeruvektorite süsteem on lineaarselt sõltumatu. Viimane kehtib parajasti siis, kui ükski veeruvektor ei avaldu eelmiste lineaarkombinatsioonina. Oletame vastuväiteliselt, et i-s veeruvektor avaldub eelmiste lineaarkombinatsioonina. Siis leiduvad sellised skalaarid $k_1, \ldots, k_{i-1} \in K$, et

$$\begin{pmatrix} t_{1i} \\ \cdots \\ t_{ni} \end{pmatrix} = T^i = k_1 T^1 + \ldots + k_{i-1} T^{i-1} = \begin{pmatrix} k_1 t_{11} + \ldots + k_{i-1} t_{1,i-1} \\ \cdots \\ k_1 t_{n1} + \ldots + k_{i-1} t_{n,i-1} \end{pmatrix}.$$

Järelikult

$$\begin{aligned} e'_i &= t_{1i} e_1 + \ldots + t_{ni} e_n \\ &= (k_1 t_{11} + \ldots + k_{i-1} t_{1,i-1}) e_1 + \ldots + (k_1 t_{n1} + \ldots + k_{i-1} t_{n,i-1}) e_n \\ &= k_1 (t_{11} e_1 + \ldots + t_{n1} e_n) + \ldots + k_{i-1} (t_{1,i-1} e_1 + \ldots + t_{n,i-1} e_n) \\ &= k_1 e'_1 + \ldots + k_{i-1} e'_{i-1}, \end{aligned}$$

mis tähendab, et vektor e'_i avaldub vektorite $e'_1, \ldots, e'_{i-1}$ lineaarkombinatsioonina. Seda aga ei saa olla, sest baasivektorid on lineaarselt sõltumatud. □

Tõestame nüüd tulemuse, mis näitab, kuidas on omavahel seotud lineaarteisenduse φ maatriksid erinevate baaside e ja e' suhtes.

Teoreem 10.27. *Olgu V vektorrum üle korpuse K, olgu $e = \{e_1, \ldots, e_n\}$ ja $e' = \{e'_1, \ldots, e'_n\}$ selle vektorruumi baasid ja olgu φ selle vektorruumi lineaarteisendus. Siis*

$$\boxed{A^{e'}_{\varphi} = T^{-1} A^{e}_{\varphi} T,}$$

kus $T = T^{e,e'}$.

TÕESTUS. Vastavalt teoreemile 10.21 on kujutus

$$f : \mathrm{End}(V) \to \mathrm{Mat}_n(K), \quad \varphi \mapsto A^e_\varphi$$

ringide isomorfism. Lemma 10.26 põhjal on maatriks T regulaarne ja seega leidub tal pöördmaatriks T^{-1}. Kuna f on pealekujutus, siis leiduvad sellised $\psi, \chi \in \mathrm{End}(V)$, et

$$T = f(\psi) = A^e_\psi, \quad T^{-1} = f(\chi) = A^e_\chi.$$

Siis

$$f(1_V) = E = TT^{-1} = f(\psi) f(\chi) = f(\psi\chi)$$

ja analoogiliselt $f(1_V) = f(\chi\psi)$. Kuna f on üksühene, siis $\psi\chi = 1_V = \chi\psi$ ehk teiste sõnadega $\chi = \psi^{-1}$. Niisiis $f(\psi^{-1}) = T^{-1}$.

Paneme tähele, et

$$T^{-1} A^e_\varphi T = f(\psi^{-1}) f(\varphi) f(\psi) = f(\psi^{-1}\varphi\psi) = A^e_{\psi^{-1}\varphi\psi}.$$

Vaja oleks näidata, et $A^e_{\psi^{-1}\varphi\psi} = A^{e'}_{\varphi}$. Olgu $A^{e'}_{\varphi} = (b_{ij})$, s.t.

$$\varphi(e'_i) = \sum_{j=1}^{n} b_{ji} e'_j$$

iga $i \in \{1, \ldots, n\}$ korral. Kuna $A^e_{\psi} = T$, siis võrreldes nende maatriksite i-ndaid veerge ja kasutades võrdust (41) saame, et

$$\psi(e_i) = t_{1i}e_1 + \ldots + t_{ni}e_n = e'_i \tag{42}$$

iga $i \in \{1, \ldots, n\}$ korral ning järelikult

$$\psi^{-1}(e'_i) = \psi^{-1}(\psi(e_i)) = 1_V(e_i) = e_i$$

iga $i \in \{1, \ldots, n\}$ korral. Seda arvestades võime kirjutada

$$\begin{aligned}(\psi^{-1}\varphi\psi)(e_i) &= \psi^{-1}(\varphi(\psi(e_i))) = \psi^{-1}(\varphi(e'_i)) = \psi^{-1}\left(\sum_{j=1}^{n} b_{ji}e'_j\right) \\ &= \sum_{j=1}^{n} b_{ji}\psi^{-1}\left(e'_j\right) = \sum_{j=1}^{n} b_{ji}e_j.\end{aligned}$$

See tähendab, et maatriksi (b_{ij}) i-ndas veerus on vektori $(\psi^{-1}\varphi\psi)(e_i)$ koordinaadid baasi e suhtes. Järelikult $A^e_{\psi^{-1}\varphi\psi} = (b_{ij})$ ehk $T^{-1}A^e_{\varphi}T = A^{e'}_{\varphi}$. □

Järeldus 10.28. *Maatriksid $A, B \in \mathrm{Mat}_n(K)$ on sarnased parajasti siis, kui nad on mingi vektorruumi V (üle K) mingi lineaarteisenduse maatriksid mingite baaside suhtes.*

Tõestus. Piisavus. See on tõestatud teoreemis 10.27.
Tarvilikkus. Olgu maatriksid $A, B \in \mathrm{Mat}_n(K)$ sarnased, s.t. $B = T^{-1}AT$, kus $T = (t_{ij}) \in \mathrm{Mat}_n(K)$ on regulaarne maatriks. Vaatleme vektorruumi K^n ja selle baasi e, mis koosneb vektoritest

$$\begin{aligned}e_1 &= (1, 0, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ &\ldots \\ e_n &= (0, 0, 0, \ldots, 1).\end{aligned}$$

Siis teoreemi 10.21 põhjal (täpsemalt kujutuse f sürjektiivsuse põhjal) leidub lineaarteisendus $\varphi \in \mathrm{End}(K^n)$ nii, et $A = f(\varphi) = A^e_{\varphi}$. Iga $i \in \{1, \ldots, n\}$ korral olgu

$$e'_i := t_{1i}e_1 + \ldots + t_{ni}e_n = (t_{1i}, \ldots, t_{ni}),$$

s.t. $e'_i \in K^n$ on maatriksi T i-s veeruvektor. Kuna T on regulaarne, siis on tema veeruvektorid $e'_1, \ldots, e'_n$ lineaarselt sõltumatud ning järelikult on $e' = \{e'_1, \ldots, e'_n\}$ vektorruumi K^n baas. Üleminekumaatriksi definitsiooni põhjal $T = T^{e,e'}$. Seega

$$B = T^{-1}AT = T^{-1}A^e_{\varphi}T = A^{e'}_{\varphi}$$

ning A ja B on lineaarteisenduse φ maatriksid baaside e ja e' suhtes. □

10.7. Karakteristlik polünoom

Seome nüüd iga maatriksiga teatud polünoomi.

Definitsioon 10.29. Maatriksi $A \in \mathrm{Mat}_n(K)$ **karakteristlikuks polünoomiks** nimetatakse polünoomi $|A - \lambda E| \in K[\lambda]$.

Näeme, et A karakteristliku polünoomi kordajateks on K elemendid ja muutujaks on λ.

Näide 10.30. Maatriksi $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$ karakteristlik polünoom on

$$\begin{aligned} |A - \lambda E| &= \left| \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = \left| \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = \begin{vmatrix} 2-\lambda & 1 \\ -1 & -\lambda \end{vmatrix} \\ &= \lambda^2 - 2\lambda + 1. \end{aligned}$$

Üldjuhul, kui $A = (a_{ij}) \in \mathrm{Mat}_n(K)$, siis

$$|A - \lambda E| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}.$$

Determinandi definitsioonist järeldub kergesti, et maatriksi $A \in \mathrm{Mat}_n(K)$ karakteristliku polünoomi aste on n ja pealiige on $(-1)^n \lambda^n$.

Lause 10.31. *Sarnaste maatriksite karakteristlikud polünoomid on võrdsed.*

Tõestus. Olgu maatriksid $A, B \in \mathrm{Mat}_n(K)$ sarnased. Siis $B = T^{-1}AT$, kus T on mingi regulaarne maatriks. Kasutades maatriksite ja determinantide omadusi saame, et

$$\begin{aligned} |B - \lambda E| &= |T^{-1}AT - T^{-1}(\lambda E)T| = |T^{-1}(AT - (\lambda E)T)| = |T^{-1}(A - \lambda E)T| \\ &= |T^{-1}||A - \lambda E||T| = |T^{-1}||T||A - \lambda E| = |T^{-1}T||A - \lambda E| = |E||A - \lambda E| \\ &= |A - \lambda E|. \end{aligned}$$

□

Definitsioon 10.32. Maatriksi $A \in \mathrm{Mat}_n(K)$ **omaväärtusteks** nimetatakse selle maatriksi karakteristliku polünoomi juuri.

Näide 10.33. Näites 10.30 vaadeldud maatriksi omaväärtusteks on polünoomi $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$ juured. Seega on sellel maatriksil omaväärtus 1, mille kordsus on kaks.

10.8. Lineaarteisenduse omaväärtused ja omavektorid

Definitsioon 10.34. Olgu V vektorruum üle korpuse K ja olgu φ vektorruumi V lineaarteisendus. Vektorit $0 \neq a \in V$ nimetatakse lineaarteisenduse φ **omavektoriks**, kui leidub selline $\lambda \in K$, et

$$\varphi(a) = \lambda a.$$

Elementi λ nimetatakse sel juhul omavektorile a vastavaks **omaväärtuseks**.

Näide 10.35. Näites 10.3(4) vaadeldud peegeldamisteisenduse omaväärtusteks on 1 ja -1. Omaväärtusele 1 vastavad omavektorid on need nullist erinevad vektorid, mis on paralleelsed y-teljega. Omaväärtusele -1 vastavad omavektorid on need nullist erinevad vektorid, mis on paralleelsed x-teljega.

Vektorruumi samasusteisenduse ja nullteisenduse jaoks on kõik nullist erinevad vektorid omavektorid.

Definitsioon 10.36. Lineaarteisenduse **karakteristlikuks polünoomiks** nimetatakse selle lineaarteisenduse maatriksi karakteristlikku polünoomi.

Paneme tähele, et lause 10.31 põhjal ei sõltu lineaarteisenduse karakteristlik polünoom sellest, millise baasi suhtes me tema maatriksit vaatleme.

Teoreem 10.37. *Lineaarteisenduse omaväärtusteks on selle teisenduse karakteristliku polünoomi juured.*

Tõestus. Olgu V n-mõõtmeline vektorruum üle korpuse K baasiga e ja olgu φ vektorruumi V lineaarteisendus. Olgu $A = A^e_\varphi = (a_{ij}) \in \mathrm{Mat}_n(K)$. Kui $x \in V$, siis järelduse 10.13 tõttu on võrdus $\varphi(x) = \lambda_0 x$ samaväärne koordinaatide veergude võrdusega $\overline{\varphi(x)}_e = \overline{\lambda_0 x}_e$. Lause 10.16 põhjal kehtib iga vektori $x \in V$ korral võrdus

$$\overline{\varphi(x)}_e = A^e_\varphi \overline{x}_e.$$

Tänu lausele 6.46 on iga $\lambda_0 \in K$ korral $\overline{\lambda_0 x}_e = \lambda_0 \overline{x}_e = \lambda_0(E\overline{x}_e) = (\lambda_0 E)\overline{x}_e$. Seega vektor $x \in V \setminus \{0\}$ on lineaarteisenduse φ omavektor parajasti siis, kui

$$A^e_\varphi \overline{x}_e = (\lambda_0 E)\overline{x}_e$$

ehk

$$(A^e_\varphi - \lambda_0 E)\overline{x}_e = \overline{0}_e$$

mingi $\lambda_0 \in K$ korral. Kui

$$\overline{x}_e = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix},$$

siis omavektorite x leidmine on samaväärne homogeense lineaarvõrrandisüsteemi

$$\begin{cases} (a_{11} - \lambda_0)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda_0)x_2 + \ldots + a_{2n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda_0)x_n = 0 \end{cases} \tag{43}$$

nullist erinevate lahendite leidmisega. Element $\lambda_0 \in K$ on omaväärtus parajasti siis, kui sellel süsteemil leidub nullist erinev lahend.

Veendume, et süsteemil (43) leidub nullist erinev lahend parajasti siis, kui λ_0 on karakteristliku polünoomi juur. Kui selle süsteemi maatriksi determinant $|A - \lambda_0 E|$ on nullist erinev, siis on tegemist Crameri peajuhuga ja ainsaks lahendiks on nullvektor, mis ei saa olla omavektor. Seega, kui leidub nullist erinev lahend, siis $|A - \lambda_0 E| = 0$ ja λ_0 on karakteristliku polünoomi

juur. Vastupidi, kui $|A - \lambda_0 E| = 0$, siis süsteemi maatriksi astak $r < n$ ja lahendite fundamentaalsüsteemis leidub $n-r \geqslant 1$ lineaarselt sõltumatut vektorit, mis peavad olema nullist erinevad. Kõik need vektorid on omavektorid. □

Teoreemi 10.37 tõestuse põhjal saame järgmise eeskirja omavektorite leidmiseks.

1. Leiame lineaarteisenduse φ maatriksi A mingi baasi e suhtes.
2. Leiame polünoomi $|A - \lambda E|$ juured $\lambda_1, \ldots, \lambda_m$.
3. Iga λ_i ($i \in \{1, \ldots, m\}$) jaoks lahendame homogeense lineaarvõrrandisüsteemi maatriksiga $A - \lambda_i E$. Selle süsteemi nullist erinevad lahendivektorid on parajasti lineaarteisenduse φ omavektorite koordinaatide vektorid baasi e suhtes.

Võib küsida, et milliste baaside suhtes on lineaarteisenduse maatriks võimalikult lihtne. Ühtedeks lihtsamateks maatriksiteks on diagonaalmaatriksid. Osutub, et kehtib järgmine lause.

Lause 10.38. *Lineaarteisenduse φ maatriks baasi $e = \{e_1, \ldots, e_n\}$ suhtes on diagonaalmaatriks parajasti siis, kui see baas koosneb teisenduse φ omavektoritest.*

Tõestus. Olgu φ vektorruumi V (üle korpuse K) lineaarteisendus ja olgu $e = \{e_1, \ldots, e_n\}$ vektorruumi V baas.
Tarvilikkus. Kui

$$A_\varphi^e = \begin{pmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k_n \end{pmatrix},$$

siis lineaarteisenduse maatriksi definitsiooni tõttu iga $i \in \{1, \ldots, n\}$ korral $\varphi(e_i) = k_i e_i$, mis tähendab, et $e_1, \ldots, e_n$ on φ omavektorid.
Piisavus. Olgu $e_1, \ldots, e_n$ lineaarteisenduse φ omavektorid. Siis leiduvad $k_1, \ldots, k_n \in K$ nii, et $\varphi(e_i) = k_i e_i$ iga $i \in \{1, \ldots, n\}$ korral. Seega A_φ^e on diagonaalmaatriks, mille peadiagonaalil on elemendid $k_1, \ldots, k_n$. □

11. Eukleidiline ruum

11.1. Eukleidilise ruumi mõiste ja põhiomadused

Eukleidilise ruumi mõiste on tekkinud geomeetriliste vektorite vektorruumide üldistamisel.

Definitsioon 11.1. Olgu E vektorruum üle korpuse $\mathbb{R}$. Kujutust

$$E \times E \to \mathbb{R}, \quad (a, b) \mapsto \langle a, b \rangle$$

(vektorite järjestatud paarile (a, b) seatakse vastavusse reaalarv, mida tähistatakse $\langle a, b \rangle$) nimetatakse **skalaarkorrutamiseks**, kui tal on järgmised omadused:

SK1. $\langle a, b \rangle = \langle b, a \rangle$ iga $a, b \in E$ korral,

SK2. $\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$ iga $a, b, c \in E$ korral,

SK3. $\langle ka, b \rangle = k\langle a, b \rangle$ iga $a, b \in E$ ja $k \in \mathbb{R}$ korral,

SK4. $\langle a, a \rangle > 0$ iga $a \in E \setminus \{0\}$ korral.

Reaalarvu $\langle a, b \rangle$ nimetatakse vektorite a ja b **skalaarkorrutiseks**. Skalaarkorrutist $\langle a, a \rangle$ nimetatakse vektori a **skalaarruuduks**.

Definitsioon 11.2. Eukleidiline ruum on vektorruum üle korpuse $\mathbb{R}$ koos sellel defineeritud skalaarkorrutamisega.

Definitsioon 11.3. Eukleidilist ruumi nimetame **triviaalseks**, kui ta sisaldab vaid nullvektorit.

Näide 11.4. Tasandi vabavektorite vektorruumil $\mathbb{E}_2$ ja kolmemõõtmelise ruumi vabavektorite vektorruumil $\mathbb{E}_3$ saab skalaarkorrutamise defineerida valemiga

$$\langle \vec{a}, \vec{b} \rangle := |\vec{a}| \cdot |\vec{b}| \cdot \cos \angle(\vec{a}, \vec{b}),$$

kus $|\vec{a}|$ ja $|\vec{b}|$ on vektorite $\vec{a}$ ja $\vec{b}$ pikkused ning $\angle(\vec{a}, \vec{b})$ on nende vektorite vaheline nurk.

Näide 11.5. Vektorruumis $\mathbb{R}^n$ (üle korpuse $\mathbb{R}$) saab skalaarkorrutamise defineerida võrdusega

$$\langle (k_1, k_2, \ldots, k_n), (l_1, l_2, \ldots, l_n) \rangle := k_1 l_1 + k_2 l_2 + \ldots + k_n l_n,$$

s.t. kahe lõpliku jada skalaarkorrutis on nende jadade vastavate komponentide korrutiste summa. Sellist skalaarkorrutamist nimetame vektorruumi $\mathbb{R}^n$ **standardseks skalaarkorrutamiseks**. Tegelikult on vektorruumil $\mathbb{R}^n$ võimalik defineerida lõpmata palju erinevaid skalaarkorrutamisi, neist igaüks tekitab erineva eukleidilise ruumi.

Märkus 11.6. Standardse skalaarkorrutamise abil saab defineerida reaalarvuliste elementidega maatriksite korrutamise. Olgu $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ maatriks, mille reavektorid on $A_1, \ldots, A_m \in \mathbb{R}^n$ ja $B \in \mathrm{Mat}_{np}(\mathbb{R})$ maatriks, mille veeruvektorid on $B^1, \ldots, B^p \in \mathbb{R}^n$. Siis

$$AB = \begin{pmatrix} \langle A_1, B^1 \rangle & \langle A_1, B^2 \rangle & \ldots & \langle A_1, B^p \rangle \\ \langle A_2, B^1 \rangle & \langle A_2, B^2 \rangle & \ldots & \langle A_2, B^p \rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle A_m, B^1 \rangle & \langle A_m, B^2 \rangle & \ldots & \langle A_m, B^p \rangle \end{pmatrix} \in \mathrm{Mat}_{mp}(\mathbb{R}).$$

Loetleme nüüd mõned omadused, mis lihtsasti järelduvad skalaarkorrutamise definitsioonist.

Lause 11.7. *Eukleidilise ruumi E mistahes vektorite $a, b, c, a_1, \ldots, a_n$ ning mistahes reaalarvu k korral*

1. $\langle a, b+c\rangle = \langle a, b\rangle + \langle a, c\rangle$,
2. $\langle a, kb\rangle = k\langle a, b\rangle$,
3. $\langle a-b, c\rangle = \langle a, c\rangle - \langle b, c\rangle$,
4. $\langle a_1 + \ldots + a_n, b\rangle = \langle a_1, b\rangle + \ldots + \langle a_n, b\rangle$,
5. $\langle 0, b\rangle = 0$ *ja* $\langle a, 0\rangle = 0$.

Tõestus. Võrdused 1 ja 2 järelduvad definitsiooni 11.1 tingimustest SK1, SK2 ja SK3. Võrduse 3 tõestuseks märgime, et

$$\begin{aligned}\langle a-b, c\rangle &= \langle a+(-b), c\rangle = \langle a, c\rangle + \langle -b, c\rangle = \langle a, c\rangle + \langle (-1)b, c\rangle = \langle a, c\rangle + (-1)\langle b, c\rangle \\ &= \langle a, c\rangle + (-\langle b, c\rangle) = \langle a, c\rangle - \langle b, c\rangle.\end{aligned}$$

Võrduse 4 saame tingimuse SK2 korduval rakendamisel.

Tõestame võrduse $\langle 0, b\rangle = 0$. Selleks võtame tingimuses SK3 $k = 0 \in \mathbb{R}$ ja $a = 0 \in E$. Siis

$$\langle 0, b\rangle = \langle 0 \cdot 0, b\rangle = \langle ka, b\rangle = k\langle a, b\rangle = 0 \cdot \langle 0, b\rangle = 0,$$

sest korrutades reaalarvu $\langle 0, b\rangle$ reaalarvuga 0 saame nulli. Võrduse $\langle a, 0\rangle = 0$ saab tõestada analoogiliselt. □

Eukleidilistes ruumides on võimalik rääkida sellistest geomeetrilistest mõistetest nagu vektori pikkus, vektorite vaheline nurk ja vektorite ortogonaalsus, isegi siis, kui need eukleidilised ruumid ise ei ole $\mathbb{E}_2$ ega $\mathbb{E}_3$.

Definitsioon 11.8. Eukleidilise ruumi E vektori a **pikkus** $|a|$ defineeritakse võrdusega

$$\boxed{|a| := \sqrt{\langle a, a\rangle},}$$

s.t. vektori pikkus on ruutjuur tema skalaarruudust. **Ühikvektor** on vektor, mille pikkus on 1.

Tänu lause 11.7 viimasele väitele võime öelda, et nullvektori pikkus on null. Tingimusest SK4 järeldub, et kõigi ülejäänud vektorite pikkused on positiivsed reaalarvud.

Järgnevas lauses tähistab sümbol $\mathrm{abs}(k)$ reaalarvu k absoluutväärtust.

Lause 11.9. *Eukleidilise ruumi E mistahes vektorite a ja b korral*

$$\boxed{\mathrm{abs}(\langle a, b\rangle) \leqslant |a| \cdot |b|,} \tag{44}$$

s.t. skalaarkorrutise $\langle a, b\rangle$ absoluutväärtus ei ületa vektorite a ja b pikkuste korrutist.

TÕESTUS. Kui $a = 0$ või $b = 0$, siis lause 11.7(5) tõttu on tõestatava võrratuse mõlemad pooled võrdsed nulliga ning seega võrratus kehtib. Eeldame edasises, et $a \neq 0$ ja $b \neq 0$. Olgu k suvaline reaalarv. Siis kasutades skalaarkorrutamise definitsiooni ja lauset 11.7 saame, et

$$0 \leqslant \langle a - kb, a - kb\rangle = \langle a, a\rangle - \langle a, kb\rangle - \langle kb, a\rangle + \langle kb, kb\rangle = \langle a, a\rangle - 2k\langle a, b\rangle + k^2\langle b, b\rangle.$$

Võttes

$$k := \frac{\langle a, b\rangle}{\langle b, b\rangle}$$

ja asendades selle arvu eelmisse võrratusse võime öelda, et

$$\langle a, a\rangle - 2\frac{\langle a, b\rangle}{\langle b, b\rangle}\langle a, b\rangle + \frac{\langle a, b\rangle^2}{\langle b, b\rangle} \geqslant 0.$$

Korrutades selle võrratuse mõlemaid pooli positiivse reaalarvuga $\langle b, b\rangle$ saame võrratuse

$$\langle a, a\rangle\langle b, b\rangle - 2\langle a, b\rangle^2 + \langle a, b\rangle^2 \geqslant 0,$$

kust $\langle a, a\rangle\langle b, b\rangle - \langle a, b\rangle^2 \geqslant 0$ ehk $\langle a, a\rangle\langle b, b\rangle \geqslant \langle a, b\rangle^2$. Võttes mõlemast poolest ruutjuure ja kasutades pikkuse definitsiooni saamegi võrratuse $\operatorname{abs}(\langle a, b\rangle) \leqslant |a| \cdot |b|$. □

Võrratust (44) nimetatakse **Cauchy-Bunjakovski**[11] **võrratuseks**. Tänu sellele võrratusele $-|a| \cdot |b| \leqslant \langle a, b\rangle \leqslant |a| \cdot |b|$, järelikult

$$-1 \leqslant \frac{\langle a, b\rangle}{|a| \cdot |b|} \leqslant 1$$

ja seega omab mõtet järgmine definitsioon.

Definitsioon 11.10. Eukleidilise ruumi E nullist erinevate vektorite a ja b vaheliseks **nurgaks** loetakse sellist nurka $\varphi \in [0, \pi]$, mille korral

$$\boxed{\cos\varphi = \frac{\langle a, b\rangle}{|a| \cdot |b|}.}$$

Kui $a = 0$ või $b = 0$, siis a ja b vaheline nurk ei ole määratud.

11.2. Ortogonaalsed vektorite süsteemid

Definitsioonist 11.10 järeldub, et kui eukleidilises ruumis $\langle a, b\rangle = 0$, siis nurk vektorite a ja b vahel on $\frac{\pi}{2}$.

Definitsioon 11.11. Öeldakse, et eukleidilise ruumi E vektorid a ja b on **ortogonaalsed** ehk **risti** (tähistus $a \perp b$), kui $\langle a, b\rangle = 0$.

Kuna $\langle a, 0\rangle = 0$, siis nullvektor on ortogonaalne kõigi vektoritega.

Definitsioon 11.12. Eukleidilise ruumi vektorite süsteemi nimetatakse **ortogonaalseks**, kui selle süsteemi vektorid on paarikaupa ortogonaalsed. Teiste sõnadega öeldes: vektorite süsteem $a_1, \ldots, a_m$ on ortogonaalne, kui

$$(\forall i, j \in \{1, \ldots, m\})(i \neq j \Longrightarrow \langle a_i, a_j\rangle = 0).$$

[11] Prantsuse matemaatiku Augustin Louis Cauchy (1789–1857) ja vene matemaatiku Viktor Bunjakovski (1804–1889) auks.

Näide 11.13. Eukleidilise ruumi $\mathbb{R}^4$ vektorite süsteem

$$\begin{aligned} a_1 &= (1,1,1,1), \\ a_2 &= (1,1,-1,-1), \\ a_3 &= (-3,3,0,0) \end{aligned}$$

on ortogonaalne.

Lause 11.14. *Kui eukleidilise ruumi vektorid a ja b on ortogonaalsed ning k ja l on reaalarvud, siis on ka vektorid ka ja lb ortogonaalsed.*

Tõestus. Kui a ja b on ortogonaalsed vektorid siis $\langle ka, lb\rangle = kl\langle a, b\rangle = kl \cdot 0 = 0$, mis tähendab, et ka ka ja lb on ortogonaalsed. □

Osutub, et ortogonaalsus on üsna tugev omadus, sellest järeldub lineaarne sõltumatus. Võiks isegi öelda, et lineaarselt sõltumatud süsteemid on head, aga ortogonaalsed on veel paremad.

Lause 11.15. *Nullist erinevate vektorite ortogonaalne süsteem on lineaarselt sõltumatu.*

Tõestus. Olgu $a_1, \ldots, a_m$ nullist erinevate vektorite ortogonaalne süsteem. Oletame, et

$$k_1a_1 + \ldots + k_ia_i + \ldots + k_ma_m = 0.$$

Siis iga $i \in \{1, \ldots, m\}$ korral

$$\langle k_1a_1 + \ldots + k_ia_i + \ldots + k_ma_m, a_i\rangle = \langle 0, a_i\rangle = 0.$$

Kasutades lauset 11.7 saame võrduse

$$k_1\langle a_1, a_i\rangle + \ldots + k_i\langle a_i, a_i\rangle + \ldots + k_m\langle a_m, a_i\rangle = 0.$$

Ortogonaalsuse tõttu annab viimane võrdus, et $k_i\langle a_i, a_i\rangle = 0$, samas SK4 põhjal $\langle a_i, a_i\rangle \neq 0$. Järelikult $k_i = 0$ iga $i \in \{1, \ldots, m\}$ korral, mis tähendab, et süsteem $a_1, \ldots, a_m$ on lineaarselt sõltumatu. □

Järgnevalt vaatleme ühte algoritmi, mida kutsutakse **Grami-Schmidti**[12] **ortogonaliseerimisprotsessiks**. Selle protsessi eesmärk on lähtudes vektorite süsteemist $a_1, a_2, \ldots, a_m$, kus ei ole nullvektoreid, konstrueerida uus süsteem $b_1, b_2, \ldots, b_r$ nii, et

1. süsteem $b_1, b_2, \ldots, b_r$ on ortogonaalne,
2. $L(a_1, a_2, \ldots, a_m) = L(b_1, b_2, \ldots, b_r)$.

Uue süsteemi esimeseks vektoriks võetakse vana süsteemi esimene vektor: $b_1 := a_1$. Hakkame vektorile b_1 lisama teisi vektoreid. Uue süsteemi teist vektorit b_2 otsime kujul

$$b_2 = k_1b_1 + a_2,$$

kus k_1 on mingi reaalarv. Me soovime, et b_1 ja b_2 oleks ortogonaalsed, s.t. $0 = \langle b_1, b_2\rangle = \langle b_1, k_1b_1 + a_2\rangle$, kust

$$k_1\langle b_1, b_1\rangle + \langle b_1, a_2\rangle = 0$$

[12] Taani matemaatiku Jørgen Pedersen Grami (1850–1916) ja saksa matemaatiku Erhard Schmidti (1876–1959) auks. Schmidt õppis aastatel 1893–1899 Tartu Ülikoolis ja lõpetas selle kandidaaditööga. Hiljem oli ta Berliini Humboldti ülikooli rektor ja Saksa DV Teaduste Akadeemia liige.

Kuna b_1 ei ole nullvektor, siis $\langle b_1, b_1\rangle \neq 0$ ja saame arvutada

$$k_1 := -\frac{\langle b_1, a_2\rangle}{\langle b_1, b_1\rangle}.$$

Sellise k_1 väärtuse korral on $b_2 \perp b_1$. Kuna $b_2 = k_1a_1+a_2 \in L(a_1, a_2)$ ja $a_2 = b_2-k_1b_1 \in L(b_1, b_2)$, siis lause 6.15 põhjal võime öelda, et $L(a_1, a_2) = L(b_1, b_2)$. Võib juhtuda, et $b_2 = 0$. Siis jätame vektori b_2 konstrueeritavast süsteemist välja. See ei muuda ei uue süsteemi ortogonaalsust ega lineaarset katet.

Oletame nüüd, et lähtudes vektoritest $a_1, \ldots, a_s$, $s \in \{2, \ldots, m-1\}$, oleme konstrueerinud vektorid $b_1, \ldots, b_t$ $(t \leq s)$ nii, et

- süsteem $b_1, b_2, \ldots, b_t$ on ortogonaalne,
- $L(a_1, a_2, \ldots, a_s) = L(b_1, b_2, \ldots, b_t)$.

Uue süsteemi järgmist vektorit b_{t+1} otsime kujul

$$b_{t+1} = k_1b_1 + \ldots + k_tb_t + a_{s+1},$$

kus $k_1, \ldots, k_t$ on mingid reaalarvud. Me soovime, et iga $i \in \{1, \ldots, t\}$ korral $\langle b_i, b_{t+1}\rangle = 0$, s.t.

$$k_1\langle b_i, b_1\rangle + \ldots + k_t\langle b_i, b_t\rangle + \langle b_i, a_{s+1}\rangle = 0.$$

Kuna süsteem $b_1, b_2, \ldots, b_t$ on ortogonaalne, siis saame, et

$$k_i\langle b_i, b_i\rangle + \langle b_i, a_{s+1}\rangle = 0,$$

kust

$$k_i = -\frac{\langle b_i, a_{s+1}\rangle}{\langle b_i, b_i\rangle}.$$

Leides arvud $k_1, \ldots, k_t$ selle valemi järgi saame arvutada vektori b_{t+1}, kusjuures vektorite süsteem $b_1, \ldots, b_t, b_{t+1}$ on ortogonaalne. Et $b_1, \ldots, b_t \in L(a_1, a_2, \ldots, a_s)$, siis $b_{t+1} \in L(a_1, a_2, \ldots, a_{s+1})$ ja seega kehtib sisalduvus $L(b_1, b_2, \ldots, b_t, b_{t+1}) \subseteq L(a_1, a_2, \ldots, a_{s+1})$. Teisest küljest, me näeme, et $a_{s+1} = b_{t+1} - k_1b_1 - \ldots - k_tb_t \in L(b_1, b_2, \ldots, b_t, b_{t+1})$ ning järelikult kehtib ka sisalduvus $L(a_1, a_2, \ldots, a_{s+1}) \subseteq L(b_1, b_2, \ldots, b_t, b_{t+1})$. Kokkuvõttes oleme saanud, et kehtib võrdus

$$L(a_1, a_2, \ldots, a_{s+1}) = L(b_1, b_2, \ldots, b_t, b_{t+1}).$$

Kui $b_{t+1} = 0$, siis me teda uude süsteemi ei võta.

Nii saamegi konstrueerida nõutavate omadustega süsteemi $b_1, b_2, \ldots, b_r$. Vastavalt konstruktsioonile on $r \leqslant m$, s.t. uues süsteemis võib olla vähem vektoreid kui esialgses.

Märgime veel, et kui esialgne süsteem $a_1, a_2, \ldots, a_m$ on lineaarselt sõltumatu, siis võrdus $b_{t+1} = 0$ ei ole võimalik. Tõepoolest, kui see nii oleks, siis $k_1b_1 + \ldots + k_tb_t + a_{s+1} = 0$, aga kuna $k_1b_1 + \ldots + k_tb_t \in L(a_1, a_2, \ldots, a_s)$, siis näeme, et meil oleks mittetriviaalne lineaarkombinatsioon vektoritest $a_1, \ldots, a_s, a_{s+1}$, mis võrduks nullvektoriga, see on aga vastuolus lineaarse sõltumatusega. Niisiis lineaarselt sõltumatu süsteemi $a_1, a_2, \ldots, a_m$ korral me ühtegi vektorit b_{t+1} välja jätma ei pea ning järelikult $r = m$.

Definitsioon 11.16. Eukleidilise ruumi baasi, mis on samal ajal ortogonaalne vektorite süsteem, nimetatakse **ortogonaalseks baasiks**.

Ortogonaalse baasi leidmiseks eukleidilises ruumis E tuleks võtta selles ruumis mingi moodustajate süsteem $a_1, \ldots, a_m$ ja rakendada sellele ortogonaliseerimisprotsessi. Kuna saadud süsteem $b_1, b_2, \ldots, b_r$ on ortogonaalne, siis tänu lausele 11.15 on ta ka lineaarselt sõltumatu. Võrduse $E = L(a_1, a_2, \ldots, a_m) = L(b_1, b_2, \ldots, b_r)$ tõttu on ta ka moodustajate süsteem.

Definitsioon 11.17. Eukleidilise ruumi vektorite süsteemi nimetatakse **ortonormeerituks**, kui see süsteem on ortogonaalne ja selle süsteemi kõik vektorid on ühikvektorid. Eukleidilise ruumi baasi, mis on samal ajal ortonormeeritud vektorite süsteem, nimetatakse **ortonormeeritud baasiks**.

Näide 11.18. Vektorruumide $\mathbb{E}_2$ ja $\mathbb{E}_3$ ortonormeeritud baase nimetatakse harilikult ristbaasideks ja tähistatakse vastavalt $\{\vec{i}, \vec{j}\}$ ja $\{\vec{i}, \vec{j}, \vec{k}\}$. Selliste baasidega on lugeja kindlasti varem kokku puutunud. Geomeetrias on ristbaasidel väga tähtis koht. Näiteks enamkasutatud geomeetriliste joonte (ringjoon, ellips, parabool jne.) või pindade (sfäär, koonus, silinder, hüperboloid) võrrandid on ristbaaside suhtes märksa lihtsamad kui suvaliste baaside suhtes. Seetõttu kasutatakse nende uurimisel just ristbaase sisaldavaid koordinaatide süsteeme ehk ristreepereid.

On selge, et vektori a korral $|a| = 1$ parajasti siis, kui $\langle a, a\rangle = 1$. Tänu sellele saab vektorite süsteemi ortonormeeritust väljendada Kroneckeri delta abil.

Lemma 11.19. *Eukleidilise ruumi vektorite süsteem $a_1, a_2, \ldots, a_m$ on ortonormeeritud parajasti siis, kui $\langle a_i, a_j\rangle = \delta_{ij}$ iga $i, j \in \{1, \ldots, m\}$ korral.*

Olgu $b_1, b_2, \ldots, b_r$ eukleidilise ruumi E ortogonaalne baas. Võtame

$$c_i := \frac{1}{|b_i|} b_i,$$

$i = 1, \ldots, r$. Siis

$$|c_i| = \left|\frac{1}{|b_i|} b_i\right| = \sqrt{\left\langle \frac{1}{|b_i|} b_i, \frac{1}{|b_i|} b_i \right\rangle} = \sqrt{\frac{1}{|b_i|^2} \langle b_i, b_i\rangle} = \sqrt{\frac{1}{\langle b_i, b_i\rangle} \langle b_i, b_i\rangle} = 1,$$

s.t. $c_1, \ldots, c_r$ on ühikvektorid. Tänu lausele 11.14 on nad paarikaupa ortogonaalsed ja lihtne on aru saada, et ka $L(c_1, \ldots, c_r) = L(b_1, \ldots, b_r) = E$. Seega on $c_1, \ldots, c_r$ eukleidilise ruumi E ortonormeeritud baas.

Viimati kirjeldatud protsessi nimetatakse süsteemi $b_1, b_2, \ldots, b_r$ **ortonormeerimiseks**.

Ortonormeeritud baasid on kasulikud ka selle poolest, et kui teame vektorite koordinaate ortonormeeritud baasi suhtes, siis on hästi lihtne arvutada nende vektorite skalaarkorrutist. Selleks on koordinaatvektorite (mis kuuluvad vektorruumi $\mathbb{R}^n$) standardne skalaarkorrutis.

Lause 11.20. *Olgu $e = \{e_1, \ldots, e_n\}$ ortonormeeritud baas eukleidilises ruumis E ning olgu vektorite $x, y \in E$ koordinaadid baasi e suhtes vastavalt $x_1, \ldots, x_n$ ja $y_1, \ldots, y_n$. Siis*

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i.$$

Tõestus. Niisiis $x = \sum_{i=1}^{n} x_i e_i$ ja $y = \sum_{j=1}^{n} y_j e_j$. Kasutades skalaarkorrutamise omadusi ja baasi e ortonormeeritust saame, et

$$\langle x, y\rangle = \left\langle \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j \right\rangle = \sum_{i,j=1}^{n} \langle x_i e_i, y_j e_j\rangle = \sum_{i,j=1}^{n} x_i y_j \langle e_i, e_j\rangle = \sum_{i,j=1}^{n} x_i y_j \delta_{ij} = \sum_{i=1}^{n} x_i y_i.$$

□

Osutub, et skalaarkorrutise abil saab avaldada ka vektori koordinaate.

Lause 11.21. *Eukleidilise ruumi E vektori x i-s koordinaat ortonormeeritud baasi $e = \{e_1, \ldots, e_n\}$ suhtes on $\langle x, e_i \rangle$.*

TÕESTUS. Olgu $x = \sum_{j=1}^{n} x_j e_j$. Siis

$$\langle x, e_i \rangle = \left\langle \sum_{j=1}^{n} x_j e_j, e_i \right\rangle = \sum_{j=1}^{n} \langle x_j e_j, e_i \rangle = \sum_{j=1}^{n} x_j \langle e_j, e_i \rangle = x_i \langle e_i, e_i \rangle = x_i.$$

□

Kasutades lineaarteisenduse maatriksi definitsiooni ja lauset 11.21 saame selle maatriksi esitada skalaarkorrutiste abil.

Järeldus 11.22. *Kui $e = \{e_1, \ldots, e_n\}$ on eukleidilise ruumi E ortonormeeritud baas ja φ on selle eukleidilise ruumi lineaarteisendus, siis selle teisenduse maatriks baasi e suhtes on*

$$A_\varphi^e = \begin{pmatrix} \langle \varphi(e_1), e_1 \rangle & \langle \varphi(e_2), e_1 \rangle & \ldots & \langle \varphi(e_n), e_1 \rangle \\ \langle \varphi(e_1), e_2 \rangle & \langle \varphi(e_2), e_2 \rangle & \ldots & \langle \varphi(e_n), e_2 \rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle \varphi(e_1), e_n \rangle & \langle \varphi(e_2), e_n \rangle & \ldots & \langle \varphi(e_n), e_n \rangle \end{pmatrix}.$$

11.3. Ortogonaalsed maatriksid ja ortogonaalsed teisendused

Definitsioon 11.23. Reaalarvuliste elementidega ruutmaatriksit A nimetatakse **ortogonaalseks**, kui tema transponeerimisel saame pöördmaatriksi, s.t. $A^T = A^{-1}$.

Definitsioonist järeldub, et ortogonaalsed maatriksid peavad olema pööratavad. Kui eespool nägime, et pöördmaatriksi leidmine suvalise maatriksi korral on suhteliselt töömahukas, siis ortogonaalsete maatriksite pöördmaatriksi leidmine on väga lihtne — piisab vaid transponeerimisest.

Lause 11.24. *Ruutmaatriksi $A \in \mathrm{Mat}_n(\mathbb{R})$ jaoks on järgmised väited samaväärsed:*

1. *A on ortogonaalne,*
2. *$A^T A = E$ ja $AA^T = E$,*
3. *A reavektorite süsteem on ortonormeeritud,*
4. *A veeruvektorite süsteem on ortonormeeritud.*

TÕESTUS. Väidete 1 ja 2 samaväärsus tuleneb otse ortogonaalse maatriksi ja pöördmaatriksi definitsioonist.

$2 \Rightarrow 3$. Olgu $AA^T = E$, kus $A \in \mathrm{Mat}_n(\mathbb{R})$. Maatriksi A^T j-s veeruvektor on maatriksi A j-s reavektor. Seega märkuse 11.6 põhjal

$$\begin{pmatrix} \langle A_1, A_1 \rangle & \langle A_1, A_2 \rangle & \ldots & \langle A_1, A_n \rangle \\ \langle A_2, A_1 \rangle & \langle A_2, A_2 \rangle & \ldots & \langle A_2, A_n \rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle A_n, A_1 \rangle & \langle A_n, A_2 \rangle & \ldots & \langle A_n, A_n \rangle \end{pmatrix} = AA^T = E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Võrreldes esimest ja viimast maatriksit saame, et $\langle A_i, A_j\rangle = \delta_{ij}$, mis tähendab, et vektorite süsteem $A_1, A_2, \ldots, A_n$ vektoruumis $\mathbb{R}^n$ on ortonormeeritud (vt. lemmat 11.19). Analoogiliselt saab tõestada implikatsiooni $2 \Rightarrow 4$.

$3 \Rightarrow 1$. Olgu A reavektorite süsteem ortonormeeritud. Siis arvutades välja korrutise AA^T näeme, et $AA^T = E$. Kuna $|A||A^T| = |AA^T| = |E| = 1$, siis $|A| \neq 0$ ning järelikult A on pööratav. Korrutades võrduse $AA^T = E$ mõlemaid pooli vasakult maatriksiga A^{-1} saame, et $A^{-1}(AA^T) = A^{-1}E$ ja seega

$$A^T = EA^T = (A^{-1}A)A^T = A^{-1}(AA^T) = A^{-1}E = A^{-1},$$

mis tähendabki, et maatriks A on ortogonaalne. Analoogiliselt saab tõestada, et $4 \Rightarrow 1$. □

Lause 11.25. *Üleminekumaatriks ortonormeeritud baasilt ortonormeeritud baasile on ortogonaalne maatriks.*

Tõestus. Olgu $e = \{e_1, \ldots, e_n\}$ ja $e' = \{e'_1, \ldots, e'_n\}$ ortonormeeritud baasid eukleidilises ruumis E ning olgu $T = (t_{ij})$ üleminekumaatriks baasilt e baasile e'. Siis $i, j \in \{1, \ldots, n\}$ korral $e'_i = t_{1i}e_1 + \ldots + t_{ni}e_n$ ja $e'_j = t_{1j}e_1 + \ldots + t_{nj}e_n$. Maatriksi T i-nda ja j-nda veeruvektori skalaarkorrutis on

$$\langle T^i, T^j\rangle = t_{1i}t_{1j} + t_{2i}t_{2j} + \ldots + t_{ni}t_{nj} = \langle e'_i, e'_j\rangle = \delta_{ij},$$

kus keskmise võrduse juures oleme kasutanud lauset 11.20 ja viimase võrduse juures baasi e' ortonormeeritust. Seega veeruvektorite süsteem on ortonormeeritud ning lausest 11.24 järeldub, et maatriks T on ortogonaalne. □

Lause 11.26. *Ortogonaalse maatriksi determinant on kas* 1 *või* -1.

Tõestus. Kui A on ortogonaalne maatriks, siis $A^TA = E$. Kuna $|A^T| = |A|$, siis

$$1 = |E| = |A^TA| = |A^T||A| = |A||A| = |A|^2.$$

Ainsad reaalarvud, mille ruut on 1, on arvud 1 ja -1, seega $|A| \in \{1, -1\}$. □

Järeldus 11.27. *Ortogonaalne maatriks on regulaarne.*

Tähistame kõigi n-ndat järku ortogonaalsete maatriksite hulga sümboliga O_n. Siis on selge, et

$$O_n \subseteq \mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R}).$$

Lause 11.28. *Hulk* O_n *on rühm maatriksite korrutamise suhtes.*

Tõestus. Olgu $A, B \in O_n$. Siis $A^T = A^{-1}$ ja $B^T = B^{-1}$. Kasutades korrutise transponeerimise ja korrutise pöördmaatriksi leidmise omadusi saame, et

$$(AB)^T = B^TA^T = B^{-1}A^{-1} = (AB)^{-1},$$

mis tähendab, et $AB \in O_n$. Seega maatriksite korrutamine on algebraline tehe hulgal O_n. On selge, et see tehe on assotsiatiivne ja $E \in O_n$ on selle tehte suhtes ühikelement.

Kui $A \in O_n$, siis

$$(A^{-1})^T = (A^T)^T = A = (A^{-1})^{-1}.$$

Seega $A^{-1} \in O_n$ ja hulga O_n igal elemendil leidub korrutamise suhtes pöördelement. Sellega on tõestatud, et O_n on rühm. □

Vaatleme nüüd eukleidiliste ruumide lineaarteisendusi.

Definitsioon 11.29. Eukleidilise ruumi E lineaarteisendust φ nimetatakse **ortogonaalseks teisenduseks** ehk **ortogonaalteisenduseks**, kui ta säilitab vektorite skalaarruudud, see tähendab, et

$$\langle \varphi(x), \varphi(x) \rangle = \langle x, x \rangle$$

iga $x \in E$ korral.

Lemma 11.30. *Ortogonaalteisendus säilitab kõik skalaarkorrutised.*

Tõestus. Olgu φ eukleidilise ruumi E ortogonaalteisendus ja $x, y \in E$. Siis

$$\langle \varphi(x+y), \varphi(x+y) \rangle = \langle x+y, x+y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle.$$

Kasutades seda, et φ on lineaarteisendus saame, et

$$\begin{aligned} \langle \varphi(x+y), \varphi(x+y) \rangle &= \langle \varphi(x) + \varphi(y), \varphi(x) + \varphi(y) \rangle \\ &= \langle \varphi(x), \varphi(x) \rangle + 2\langle \varphi(x), \varphi(y) \rangle + \langle \varphi(y), \varphi(y) \rangle \\ &= \langle x, x \rangle + 2\langle \varphi(x), \varphi(y) \rangle + \langle y, y \rangle. \end{aligned}$$

Seega $2\langle x, y \rangle = 2\langle \varphi(x), \varphi(y) \rangle$, kust reaalarvuga 2 jagades saame võrduse

$$\langle x, y \rangle = \langle \varphi(x), \varphi(y) \rangle.$$

□

Järeldus 11.31. *Ortogonaalteisendus säilitab vektorite pikkused ja vektorite vahelised nurgad.*

Lause 11.32. *Kui eukleidilise ruumi lineaarteisendus viib mingi ortonormeeritud baasi ortonormeeritud baasiks, siis see teisendus on ortogonaalteisendus.*

Tõestus. Olgu E eukleidiline ruum ortonormeeritud baasiga $e = \{e_1, \ldots, e_n\}$ ja olgu φ selle eukleidilise ruumi lineaarteisendus. Eeldame, et teisendus φ viib baasi e ortonormeeritud baasiks $\varphi(e_1), \ldots, \varphi(e_n)$. Vaatleme suvalist vektorit $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n \in E$. Kuna φ on lineaarteisendus, siis $\varphi(x) = x_1\varphi(e_1) + x_2\varphi(e_2) + \ldots + x_n\varphi(e_n)$. Kasutades lauset 11.20 kaks korda saame arvutada

$$\langle \varphi(x), \varphi(x) \rangle = x_1^2 + x_2^2 + \ldots + x_n^2 = \langle x, x \rangle.$$

Seega on φ ortogonaalteisendus. □

Osutub, et lineaarteisenduse ortogonaalsuse üle saab otsustada tema maatriksi põhjal.

Teoreem 11.33. *Olgu φ eukleidilise ruumi E lineaarteisendus ja olgu $e = \{e_1, \ldots, e_n\}$ eukleidilise ruumi E ortonormeeritud baas. Siis φ on ortogonaalteisendus parajasti siis kui A_φ^e on ortogonaalne maatriks.*

Tõestus. Olgu $A = A_\varphi^e = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ lineaarteisenduse $\varphi : E \to E$ maatriks baasi e suhtes. Siis $\varphi(e_i) = a_{1i}e_1 + \ldots + a_{ni}e_n$ ja $\varphi(e_j) = a_{1j}e_1 + \ldots + a_{nj}e_n$ iga $i, j \in \{1, \ldots, n\}$ korral. Tarvilikkus. Olgu φ ortogonaalteisendus. Kasutades lauset 11.20, lemmat 11.30 ja baasi e ortonormeeritust võime öelda, et

$$\langle A^i, A^j \rangle = \sum_{k=1}^{n} a_{ki}a_{kj} = \langle \varphi(e_i), \varphi(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{ij}$$

mistahes $i, j \in \{1, \ldots, n\}$ korral. Tänu lemmale 11.19 on A veeruvektorite süsteem ortonormeeritud ja lause 11.24 põhjal on maatriks A ortogonaalne.
Piisavus. Olgu nüüd maatriks A^e_φ ortogonaalne. Siis tema veeruvektorite süsteem on ortonormeeritud. Järelikult

$$\delta_{ij} = \langle A^i, A^j \rangle = \sum_{k=1}^{n} a_{ki} a_{kj} = \langle \varphi(e_i), \varphi(e_j) \rangle,$$

mis tähendab, et süsteem $\varphi(e_1), \ldots, \varphi(e_n)$ on ortonormeeritud ja muuhulgas ei ole selles süsteemis nullvektoreid. Lause 11.15 põhjal on see süsteem lineaarselt sõltumatu ja lause 6.43 tõttu on see süsteem eukleidilise ruumi E baas. Kuna φ viib ortonormeeritud baasi $\{e_1, \ldots, e_n\}$ ortonormeeritud baasiks $\{\varphi(e_1), \ldots, \varphi(e_n)\}$, siis lausest 11.32 saame järeldada, et φ on ortogonaalteisendus. □

11.4. Sümmeetrilised maatriksid ja sümmeetrilised teisendused

Meenutame, et maatriksit A nimetatakse sümmeetriliseks, kui $A^T = A$. Kui $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$, siis tema sümmeetrilisus tähendab seda, et $a_{ij} = a_{ji}$ iga $i, j \in \{1, \ldots, n\}$ korral.

Osutub, et on võimalik vaadelda eukleidiliste ruumide teatud lineaarteisenduste klassi, mis on sümmeetriliste maatriksitega seotud sarnaselt sellega, kuidas ortogonaalteisendused on seotud ortogonaalmaatriksitega.

Definitsioon 11.34. Eukleidilise ruumi E lineaarteisendust φ nimetatakse **sümmeetriliseks teisenduseks**, kui

$$\langle \varphi(x), y \rangle = \langle x, \varphi(y) \rangle$$

iga $x, y \in E$ korral.

Teoreem 11.35. *Olgu φ eukleidilise ruumi E lineaarteisendus ja olgu $e = \{e_1, \ldots, e_n\}$ eukleidilise ruumi E ortonormeeritud baas. Siis φ on sümmeetriline teisendus parajasti siis kui A^e_φ on sümmeetriline maatriks.*

Tõestus. Tarvilikkus. Olgu φ sümmeetriline teisendus ja olgu $A = A^e_\varphi = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ teisenduse φ maatriks baasi $e = \{e_1, \ldots, e_n\}$ suhtes. Kasutades lineaarteisenduse maatriksi definitsiooni, lauset 11.21, teisenduse φ sümmeetrilisust, skalaarkorrutamise kommutatiivsust ja veelkord lauset 11.21 saame, et

$$a_{ij} = \langle \varphi(e_j), e_i \rangle = \langle e_j, \varphi(e_i) \rangle = \langle \varphi(e_i), e_j \rangle = a_{ji}$$

mistahes $i, j \in \{1, \ldots, n\}$ korral. Seega A on sümmeetriline maatriks.
Piisavus. Olgu nüüd maatriks $A = A^e_\varphi = (a_{ij})$ sümmeetriline. Kui vektori $x \in E$ koordinaadid baasi e suhtes on $x_1, \ldots, x_n$ (s.t. $x = x_1 e_1 + \ldots + x_n e_n$), siis tähistame

$$x_e := (x_1\ x_2\ \ldots\ x_n) \in \mathrm{Mat}_{1,n}(\mathbb{R}).$$

(Siis muuhulgas $x_e^T = \overline{x}_e$, vt. tähistusi enne lauset 10.16.) Lause 11.20 ütleb seda, et mistahes vektorite $x, y \in E$ korral maatriksite x_e ja y_e^T korrutis $x_e y_e^T$ on (1×1)-maatriks, mille ainus element on $\langle x, y \rangle$. Samastades selle maatriksi tema ainsa elemendiga võime kirjutada, et

$$\langle x, y \rangle = x_e y_e^T.$$

Lausest 10.16 järeldub, et $\varphi(x)_e^T = Ax_e^T$ ja $\varphi(y)_e^T = Ay_e^T$. Transponeerides saame, et

$$\varphi(x)_e = \left(\varphi(x)_e^T\right)^T = \left(Ax_e^T\right)^T = x_e A^T.$$

Seega

$$\langle \varphi(x), y\rangle = \varphi(x)_e y_e^T = x_e A^T y_e^T = x_e A y_e^T = x_e \varphi(y)_e^T = \langle x, \varphi(y)\rangle.$$

Sellega oleme näidanud, et φ on sümmeetriline teisendus. □

Näide 11.36. Vaatleme jälle lineaarteisendust $\varphi : \mathbb{E}_2 \to \mathbb{E}_2$, mis seisneb vektorite peegeldamises y-telje suhtes (vt. näidet 10.15(3)). Kuna selle teisenduse maatriks ristbaasi $\{\vec{i}, \vec{j}\}$ suhtes on nii ortogonaalne kui ka sümmeetriline, siis teoreemi 11.33 põhjal on see teisendus ortogonaalne ja teoreemi 11.35 põhjal on ta sümmeetriline.

Kui φ on n-mõõtmelise eukleidilise ruumi sümmeetriline teisendus, siis tema karakteristlik polünoom kuulub ringi $\mathbb{R}[\lambda]$. Kui vaatleksime seda polünoomina üle korpuse $\mathbb{C}$, siis algebra põhiteoreem ütleb, et kordsusi arvestades on karakteristlikul polünoomil n kompleksarvulist juurt. Osutub, et sümmeetrilise teisenduse puhul on kõik need juured tegelikult reaalarvud.

Teoreem 11.37. *Mittetriviaalse eukleidilise ruumi sümmeetrilise teisenduse karakteristliku polünoomi juured on reaalarvud.*

Järeldus 11.38. *Mittetriviaalse eukleidilise ruumi sümmeetrilisel teisendusel leidub vähemalt üks omavektor.*

Teoreem 11.39. *Mittetriviaalse eukleidilise ruumi lineaarteisendus on sümmeetriline parajasti siis, kui leidub selle lineaarteisenduse omavektoritest koosnev ortonormeeritud baas.*

Teoreemidest 10.21, 11.35, 11.39, lausest 10.38 ja järeldusest 10.28 on lihtsasti tuletatav järgmine tulemus.

Järeldus 11.40. *Kui $A \in \mathrm{Mat}_n(\mathbb{R})$ on sümmeetriline maatriks, siis A on sarnane diagonaalmaatriksiga ehk A on* **diagonaliseeritav**.

Kasutatud kirjandus

1. M. Kilp, Algebra I, Eesti Matemaatika Selts, Tartu, 2005.
2. K. Kaarli, Algebra I loengute slaidid,
http://math.ut.ee/pmi/kursused/algebraI/algebra1_ slides.pdf .
3. A. Parring, Algebra ja Geomeetria loengukonspekt,
http://math.ut.ee/pmi/kursused/ag/parring/.
4. G. Kangro, Kõrgem algebra I, RK "Teaduslik kirjandus", Tartu, 1948.
5. A. I. Kostrikin, Vvedenie v algebru, Nauka, Moskva, 1977 (vene keeles).