Luku 3

Lineaariset operaattorit

3.1. Invariantit aliavaruudet ja ominaisarvot

Tässä kurssin osassa tutkitaan tarkemmin äärellisulotteisten vektoriavaruuksien *endomorfismeja* eli lineaarisia kuvauksia $L\colon V \to V$, missä V on äärellisulotteinen vektoriavaruus. Tällaisia kuvauksia sanotaan myös avaruuden V (lineaarisiksi) *operaattoreiksi*.

Olkoon V äärellisulotteinen vektoriavaruus kunnan K yli ja olkoon $L\colon V \to V$ lineaarinen operaattori. Valitsemalla avaruudelle V erilaisia kantoja $E, E', E'', \ldots$ saadaan kuvaukselle L matriisiesityksiä $[L]_E, [L]_{E'}, [L]_{E''}, \ldots$. Yleisesti ottaen nämä matriisit ovat erilaisia ja riippuvat valitusta kannasta. Monissa sovelluksissa ollaan kiinnostuneita löytämään operaattorille L "mahdollisimman yksinkertaisen" ja "tarpeeksi säännöllisen" matriisiesityksen $[L]_E$. Se, mitä voidaan pitää yksinkertaisena ja säännöllisenä, riippuu tietenkin asiayhteydestä ja sovelluksesta. Esimerkiksi, mitä enemmän nolla-termejä 0_K matriisissa esiintyy, sitä helpompi sitä on yleensä käsitellä, tällaisen matriisin determinantin laskeminen on helpompaa kuin mielivaltaisen matriisin kohdalla. "Yksinkertaisten" matriisien kääntäminen tai kertominen keskenään on helpompaa kuin mielivaltaisten matriisien kohdalla jne. Lisäksi säännöllisten matriisiesitysten avulla voidaan päätellä joitakin mielenkiintoisia teoreettisia tuloksia.

Luvun pääpaino on siis lineaarisen operaattorin matriisiesitysten tutkimisessa. Lisäksi meitä kiinnostavat ainoastaan matriisiesitykset muotoa $[L]_E$ eli avaruuden V yhden kannan E suhteen. Matriisiesityksiä $[L]_{E',E}$ avaruuden *eri* kantojen E', E suhteen ei juuri käsittellä tässä luvussa (koska niiden klassifikaatio on hyvin yksinkertainen eikä kerro operaattorista kovin paljon, kts. kurssin harjoitustehtäviä).

Jatkossa käytetään seuraavaa kätevää merkintätapaa. Olkoot A $(n \times m)$-kokoinen matriisi, B $(n \times p)$-kokoinen matriisi, C $(l \times m)$-kokoinen matriisi ja D $(l \times p)$-kokoinen matriisi. Tällöin voidaan muodostaa $(n + l) \times (m + p)$-kokoinen matriisi

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

jonka n ensimmäistä riviä saadaan yhdistelemällä ilmeisellä tavalla matriisien A ja B vastaavia rivejä ja loput l riviä yhdistelemällä vastaavasti matriisien C ja D rivejä. Tällä

tavalla muodostettua matriisia sanotaan *lohkomatriisiksi*. Nimitys tulee siitä, että ajatellaan tällöin matriiseja A, B, C, D *lohkoina*, joista isompi matriisi ikään kuin "kasataan". Myös isompia kuin neljästä lohkosta muodostettuja lohkomatriisia voi tulla esille.

Olkoon $L\colon V \to V$ äärellisulotteisen K vektoriavaruuden V lineaarinen operaattori. Oletetaan, että jonkun avaruuden V kannan $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$ suhteen matriisi $[L]_E$ on lohkomatriisi muotoa

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}, \tag{3.1}$$

missä A on $(n \times n)$-kokoinen *neliömatriisi*, $n \leq m$, ja 0 on nollamatriisi, jonka kaikki alkiot ovat kunnan K nolla-alkioita. Olkoon $W = \operatorname{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ kannan E n ensimmäisen vektorin $\mathbf{e}_1, \ldots, \mathbf{e}_n$ virittämä aliavaruus. Tällöin esityksestä 3.1 seuraa, että kuvaus L kuvaa aliavaruuden W itselleen, toisin sanoen pätee $L(W) \subset W$. Tällaisia aliavaruuksia sanotaan *invarianteiksi* operaattorin L suhteen.

Avaruuden V aliavaruus W on siis invariantti operaatorin L suhteen, jos $L(W) \subset W$. Tällöin kuvauksen L *rajoittuma* $L|W$ määrittelee aliavaruuden W lineaarisen operaattorin $L|W\colon W \to W$. Yllä ollaan näytetty, että jos operaatorin $L\colon V \to V$ matriisi $[L]_E$ jonkun kannan $E = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ suhteen on lohkomatriisi muotoa

$$[L]_E = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

missä $A \in M(n \times n; K)$ on neliömatriisi, niin avaruudella V on n-ulotteinen aliavaruus $W = \operatorname{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, joka on invariantti operaattorin L suhteen.

Kääntäen, olkoon $W \leq V$ avaruuden V n-ulotteinen aliavaruus, joka on invariantti operaattorin $L\colon V \to V$ suhteen. Valitaan avaruudelle V kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ siten, että $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ on aliavaruuden W kanta. Tämä on mahdollista Lemman 2.48 nojalla. Tällöin matriisi $[L]_E$ on lohkomatriisi muotoa

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

missä $A \in M(n \times n; K)$ on rajottuman $L|W\colon W \to W$ matriisi kannan $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ suhteen.

Avaruuden V triviaalit aliavaruudet $\{\mathbf{0}_V\}$ ja V ovat selvästi invariantteja minkä tahansa operaattorin $L\colon V \to V$ suhteen. Nämä triviaalit invariantit aliavaruudet eivät tietenkään kerro mitään operaattorista L. Voi hyvinkin käydä niin, että nämä ovat operaattorin L ainoat invariaantit aliavaruudet.

Esimerkki 3.2. *Olkoon $m \in \mathbb{N}$. Tarkastellaan $\mathbb{R}$-vektoriavaruutta*

$$P_m = P_m(\mathbb{R}) = \{p\colon \mathbb{R} \to \mathbb{R} \text{ } \textit{on polynomi, jonka aste on korkeintaan } m\}.$$

Tällöin $\dim P_m = m+1$ *ja avaruudella* P_m *on "kanoninen" kanta* $E = (1, x, x^2, \ldots, x^m)$. *Kun* $n \leq m$ *avaruus* P_n *on avaruuden* P_m *aliavaruus, joten saadaan nouseva ketju aliavaruuksia*

$$P_0 \subset P_1 \subset P_2 \subset \ldots \subset P_n \subset \ldots \subset P_m.$$

Lisäksi jokaisella $n = 1, \ldots, n$ *jono* $E_n = (1, x, x^2, \ldots, x^n)$ *on aliavaruuden* P_n *kanta.*

Olkoon $\mathcal{D}\colon P_m \to P_m$, $\mathcal{D}(p) = p'$ derivaatta-*operaattori. Tällöin aliavaruus* P_n *on invariantti operaattorin* $\mathcal{D}$ *suhteen jokaisella* $0 \leq n \leq m$. *Tämä johtuu yksinkertaisesti siitä, että* l*-asteisen polynomin derivaattakuvaus on* $(l-1)$*-asteinen polynomi. Derivaatta-operaattorin matriisi kannan* $E = (1, x, x^2, \ldots, x^m)$ *suhteen on muotoa*

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & m \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Esimerkiksi kun $m = 4$ *derivaatta-operaattorin* $\mathcal{D}\colon P_4 \to P_4$ *matriisi kannan* E *suhteen on*

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Komplementariset invariantit aliavaruudet

Oletetaan, että operaatorilla $L\colon V \to V$ (missä V on äärellisulotteinen) on invariantti aliavaruus W. Aliavaruudella W on Lemman 2.159 nojalla olemassa avaruudessa V komplementti W'. Määritelmän mukaan tämä tarkoittaa sitä, että summa $W + W'$ on suora ja lisäksi pätee $V = W \oplus W'$.

Aliavaruuden komplementti ei ole yleensä yksikäsitteinen, vaan aliavaruudella W on yleensä paljon erilaisia komplementteja. Jos invariantilla aliavaruudella W on olemassa komplementti W', joka on myös invariantti operaatorin L suhteen, operaattorilla L on olemassa matriisiesitys muotoa

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

missä $A \in M(n \times n; K)$ on rajottuman $L|W\colon W \to W$ matriisi (jonkun W:n kannan suhteen) ja $B \in M(l \times l; K)$ on rajottuman $L|W'\colon W' \to W'$ matriisi (jonkun W':n kannan suhteen).

Kääntäen, oletetaan, että operaattorin L matriisi jonkun avaruuden V kannan $E = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ suhteen on lohkomatriisi muotoa

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

missä A on $(n \times n)$-kokoinen matriisi ja B on $(m-n) \times (m-n)$-kokoinen matriisi. Tällöin aliavaruudet

$$W = \mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_n) \text{ ja}$$

$$W' = \mathrm{Span}(\mathbf{e}_{n+1}, \ldots, \mathbf{e}_m)$$

ovat kumpikin invariantteja operaattorin L suhteen ja lisäksi ovat toistensa komplementteja avaruudessa V, $V = W \oplus W'$.

Tämä havainto voidaan helposti yleistää mielivaltaisen monen aliavaruuden tapaukseen. Toisin sanoen jos $V = W_1 \oplus W_2 \oplus \ldots \oplus W_n$, missä jokainen alivaruus W_i, $1 \leq i \leq n$, on invariantti operaattorin L suhteen, niin L voidaan esittää sopivassa kannassa lohkomatriisina

$$\begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \\ 0 & 0 & \ldots & A_n \end{bmatrix},$$

missä A_i on rajoittuman $L|W_i \colon W_i \to W_i$ matriisi (erityisesti neliömatriisi) jokaisella $i = 1, \ldots, n$. Kuvauksen L rajoittumat $L|W_i$ määräävät tällöin kuvauksen L yksikäsitteisesti (suoran summan universaaliominaisuus 2.168). Lisäksi suhteellisen helposti nähdään (harjoitustehtävä), että tällöin

$$\det L = \det A_1 \cdot \det A_2 \cdot \ldots \cdot \det A_n = \prod_{i=1}^{n} \det(L|W_i)$$

eli kuvauksen L determinaantti voidaan laskea rajoittumien $L|W_i$ determinanttien *tulona*. Mitä pienempi neliömatriisi on, sitä helpompi on laskea sen determinantti. Näin ollen on hyödyllistä osata hajoittaa isompi avaruus pienempiin L-invariantteihin palasiin.

Yleisesti ottaen voi kuitenkin hyvinkin käydä niin, että invariantin aliavaruuden W jokainen komplementti ei olekaan invariantti operaattorin L suhteen.

Esimerkkejä 3.3. *1) Olkoon* $L \colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$L(x, y) = (y, x).$$

L on $\mathbb{R}$-vektoriavaruuden $\mathbb{R}^2$ operaattori. Havainnollisesti sen voi ajatella peilauksena suoran $y = x$ suhteen tasossa. Standardikannan suhteen kuvauksen L matriisi on

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Koska L on peilaus suoran $y = x$ suhteen, tämän suoran pisteet pysyvät paikallaan kuvauksessa L, joten aliavaruus

$$W = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

on erityisesti invariantti. Myös aliavaruus

$$W' = \{(x, -x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

on invariantti kuvauksen L suhteen. Aliavaruudet W ja W' ovat komplementaarisia avaruudessa $\mathbb{R}^2$, sillä summa $W + W'$ on suora ja $W \oplus W' = \mathbb{R}^2$. Valitsemalla avaruudessa W kanta $((1, 1))$ ja avaruudessa W' kanta $((1, -1))$, saadaan avaruudelle

$\mathbb{R}^2$ *kanta* $E' = ((1,1),(1,-1))$. *Tämän kannan suhteen kuvauksen matriisi on*

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

erityisesti niin sanottu diagonaalimatriisi, joista puhutaan alla.

2) *Tarkastellaan polynomiavaruutta* $P_m(\mathbb{R})$ *ja sen derivaattaoperaattoria* $\mathcal{D}$. *Esimerkissä 3.2 ollaan todettu, että aliavaruus* $P_n(\mathbb{R})$ *on invariantti operaattorin* $\mathcal{D}$ *suhteen jokaisella* $0 \leq n \leq m$. *Kun* $0 < k < n$ *aliavaruudella* $P_n(\mathbb{R})$ *ei ole olemassa komplementtia* W' *avaruudessa* $P_m(\mathbb{R})$, *joka olisi myös invariantti derivaatta-operaattorin* $\mathcal{D}$ *suhteen. Tämän väitteen todistus jätetään harjoitustehtäväksi.*

Ominaisvektorit ja ominaisarvot.
Olkoon $L\colon V \to V$ K-vektoriavaruuden V endomorfismi ja olkoon W avaruuden V *1-ulotteinen* aliavaruus (eli yksinkertaisin mahdollinen epätriviaali aliavaruus). Tutkitaan millä ehdoilla W on invariantti operaatorin L suhteen.

Koska $\dim W = 1$, aliavaruus W on yhden vektorin $\mathbf{w} \in V$, $\mathbf{w} \neq \mathbf{0}_V$, virittämä, jolloin pätee $W = \{k\mathbf{w} \mid k \in K\}$.

Oletetaan, että W on invariantti operaatorin L suhteen. Tällöin erityisesti $L(\mathbf{w}) \in W$, mikä tarkoittaa täsmälleen sitä, että

$$L(\mathbf{w}) = k'\mathbf{w} \tag{3.4}$$

jollakin $k' \in K$. Kääntäen, oletetaan, että yhtälö (3.4) on voimassa. Olkoon $\mathbf{v} = k\mathbf{w} \in W$, $k \in K$. Tällöin

$$L(\mathbf{v}) = kL(\mathbf{w}) = kk'\mathbf{w} \in W,$$

eli W on invariantti L:n suhteen.

Vektoreita $\mathbf{w} \neq \mathbf{0}_V$, jotka toteutavat yhtälön (3.4) sanotaan operaattorin L *ominaisvektoreiksi.*

Määritelmä 3.5. *Olkoon* $\mathbf{w} \in V, \mathbf{w} \neq \mathbf{0}_V$ *ja olkoon* $L\colon V \to V$ *lineaarinen operaattori. Vektoria* $\mathbf{w}$ *sanotaan operaattorin* L ominaisvektoriksi *jos on olemassa skalaari* $k \in K$ *siten, että*

$$L(\mathbf{w}) = k\mathbf{w}.$$

Skalaari $k \in K$ *on tällöin ominaisvektoriin* $\mathbf{w}$ *liittyvä kuvauksen* L ominaisarvo.

Operaattorin L ominaisarvo *on sellainen* $k \in K$, *joka on jonkun operaattorin* L *ominaisvektoriin* $\mathbf{w}$ *liittyvä ominaisarvo. Toisin sanoen* $k \in K$ *on operaattorin* L *ominaisarvo jos on olemassa* $\mathbf{w} \in V, \mathbf{w} \neq \mathbf{0}_V$, *siten, että*

$$L(\mathbf{w}) = k\mathbf{w}.$$

Huomaa, että nolla-vektori ei kelpaa ominaisvektoriksi. Tämä johtuu siitä, että nolla-vektorille pätee

$$L(\mathbf{0}_V) = \mathbf{0}_V = k\mathbf{0}_V$$

jokaisella $k \in K$. Jos nolla-vektori laskettaisi ominaisvektoriksi, jokaisesta skalaarikunnan alkiosta tulisi jokaisen operaattorin ominaisarvo, eikä ominaisarvon käsitteestä olisi

tällöin mitään hyötyä.

Olkoon $k \in K$. Määritellään

$$V_k = \{\mathbf{v} \in V \mid L(\mathbf{v}) = k\mathbf{v}\}.$$

Toisin sanoen V_k on aliavaruus, joka koostuu kaikista ominaisvektoreista, joita vastaava ominaisarvo on k **sekä** nolla-vektorista. Huomaa, että V_k määritellään kaikilla $k \in K$, eikä ainoastaan ominaisarvoille. Jos k ei ole ominaisarvo, V_k on triviaali aliavaruus $\{\mathbf{0}_V\}$. Jos taas k on ominaisarvo, pätee $\dim V_k \geq 1$ ja aliavaruutta V_k sanotaan tällöin *ominaisarvoon k liittyväksi aliavaruudeksi* tai yksinkertaisesti *ominaisarvoaliavaruudeksi*. Vaikka ominaisarvon käsite oli johdettu tarkastelemalla 1-ulotteisia invariantteja aliavaruuksia, ominaisarvoaliavaruuden V_k ei tarvitse olla 1-ulotteinen. Esimerkiksi kunnan ykkösalkio 1_K on identtisen operaatorin $\mathrm{id}\colon V \to V$ ominaisarvo, johon liittyvä ominaisarvoaliavaruus V_1 on koko avaruus, $V_1 = V$.

Esimerkkejä 3.6. *1) Jatketaan polynomiavaruuden $P_n(\mathbb{R})$ derivaatta-operaatorin $\mathcal{D}$ tutkimista (kts. esim. 3.2). Tämän kuvauksen ainoa ominaisarvo on 0 ja ainoat ominaisvektorit ovat nollasta eroavat vakiopolynomit, eli aliavaruuden $P_0(\mathbb{R})$ vektorit (nollaa lukuunottamatta). Nimittäin, jos f on l-asteinen polynomi $\mathbb{R} \to \mathbb{R}$, niin f' on $(l-1)$-asteinen polynomi. Tästä seuraa, että yhtälö*

$$f' = rf$$

on mahdollinen jos ja vain jos $r = 0$, jolloin $f' = 0$, joten f on vakiopolynomi.

2) Olkoon V vektoriavaruus, jonka muodostavat kaikki kuvaukset $f\colon \mathbb{R} \to \mathbb{R}$, joilla on olemassa n:nnen kertaluvun derivaatta $f^{(n)}$ jokaisella $n \in \mathbb{N}$. Tällöin kaikilla $f \in V$ myös sen derivaattafunktio f' on avaruuden V alkio, jolloin voidaan määritellä lineaarinen operaattori $L\colon V \to V$ kaavalla $L(f) = f'$.

Olkoon $r \in \mathbb{R}$ mielivaltainen. Differentiaaliyhtälöiden kurssilla osoitetaan, että avaruuden V vektori f on yhtälön $f' = rf$ ratkaisu jos ja vain jos $f(x) = ae^{rx}$ kaikilla $x \in \mathbb{R}$, missä $a \in \mathbb{R}$ on vakio. Tästä seuraa, että jokainen reaaliluku on operaatorin L ominaisarvo. Lisäksi ominaisarvoon $r \in \mathbb{R}$ liittyvä aliavaruus V_r on 1-ulotteinen aliavaruus, nimittäin kuvauksen $x \mapsto e^{rx}$ virittämä aliavaruus.

Tarkastellaan samassa avaruudessa V lineaarioperaattoria $L' = L^2 = L \circ L$. Tällöin $L'(f) = f''$ kaikilla $f \in V$. Jälleen kerran voidaan todistaa, että jokainen reaaliluku $r \in \mathbb{R}$ on tämän operaatorin ominaisarvo. Esimerkiksi $(\cos x)'' = -\cos x$ ja $(\sin x)'' = -\sin x$, joten $\cos x$ ja $\sin x$ ovat molemmat ominaisvektoreita, joita vastaava ominaisarvo on (-1). Koska jono $(\cos x, \sin x)$ on vapaa avaruudessa V (vrt. esim. 2.34, (2)), tästä voidaan heti päätellä, että ominaisarvoaliavaruus V_{-1} on ainakin kaksiulotteinen. Differentiaaliyhtälöiden kurssilla itse asiassa osoitetaan, että kuvaukselle L' aliavaruus V_r on täsmälleen kaksiulotteinen kaikilla $r \in \mathbb{R}$.

Tämän esimerkin vektoriavaruus V ei ole äärellisulotteinen. Lemmassa 3.7 alla osoitetaan, että äärellisulotteisen vektoriavaruuden operaattorilla voi olla vain äärellinen määrä ominaisarvoja.

3) Olkoon $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x,y) = (x-y, 2y)$. Tällöin L on $\mathbb{R}$-vektoriavaruuden $\mathbb{R}^2$ operaattori, jonka matriisi standardikannassa on

$$\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Tästä esityksestä nähdään heti , että $L(1,0) = (1,0)$, joten $(1,0)$ on ominaisvektori, jota vastaa ominaisarvo 1. Tutkitaan, onko kuvauksella L muita ominaisarvoja. Olkoon $r \in \mathbb{R}$. Tällöin kaikilla $(x,y) \neq (0,0)$ yhtälö

$$L(x,y) = (x-y, 2y) = r(x,y) = (rx, ry)$$

pätee jos ja vain jos $x - y = rx$ ja $2y = ry$. Jälkimmäisestä yhtälöstä saadaan heti, että $r = 2$ tai $y = 0$. Jos $r = 2$, ensimmäisestä yhtälöstä seuraa, että $y = -x$, joten 2 on eräs ominaisarvo ja sitä vastaava aliavaruus on 1-ulotteinen suora

$$V_2 = \{(x, -x) \mid x \in \mathbb{R}\}.$$

Jos taas $y = 0$, täytyy olla $x \neq 0$ (ominaisvektori ei saa olla nolla-vektori) ja lisäksi ensimmäisestä yhtälöstä tällöin seuraa, että $x = rx$. Tämä on mahdollista jos ja vain jos $r = 1$, jolloin vastaava aliavaruus on 1-ulotteinen suora

$$V_1 = \{(x, 0) \mid x \in \mathbb{R}\}.$$

Ominaisvektoreista koostuvassa kannassa $((1,0),(1,-1))$ kuvauksen matriisi on

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Tällaista matriisia sanotaan diagonaalimatriisiksi*, niistä puhutaan kunnolla alla. Huomaa, että diagonaalimatriisin päälävistäjän alkiot ovat tässä tapauksessa operaattorin ominaisarvot. Tämä ei ole mikään sattuma, yleisemmin niin sanotun yläkolmiomatriisin ominaisarvot ovat samat kuin sen diagonaalialkiot. Tämä väite todistetaan myöhemmin.*

Lemma 3.7. *Olkoon $L\colon V \to V$ K-vektoriavaruuden V operaattori.*

(1) Olkoon $k \in K$. Tällöin V_k on avaruuden V aliavaruus, joka on lisäksi invariantti operaattorin L suhteen. Aliavaruus V_k on epätriviaali jos ja vain jos k on operaattorin L ominaisarvo.

(2) Olkoot $k_1, \ldots, k_n \in K$ eri skalaarikunnan alkioita. Tällöin summa $\sum_{i=1}^{n} V_{k_i}$ on suora.

(3) Olkoon V äärellisulotteinen. Tällöin operaattorilla L on korkeintaan luvun $\dim V$ verran erilaisia ominaisarvoja.

Todistus. (1) Harjoitustehtävä.
(2) Osoitetaan väite induktiolla luvun n suhteen. Kun $n = 1$ ei ole mitään todistettavaa

- yhden aliavaruuden muodostama kokoelma on suora. Oletetaan, että väite on tosi arvolla $(n-1)$. Olkoot $k_1, \ldots, k_n \in K$ eri skalaarikunnan alkiota. Osoitetaan, että summa $\sum_{i=1}^{n} V_{k_i}$ on suora.

Käytetään Lemman 2.155 ehtoa 2. Olkoot $\mathbf{v}_i \in V_{k_i}$, $i = 1, \ldots, n$ ja oletetaan, että

$$\mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n = \mathbf{0}_V. \tag{3.8}$$

Soveltamalla tähän yhtälöön kuvausta L saadaan yhtälöketju

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n = L(\mathbf{v}_1) + L(\mathbf{v}_2) + \ldots + L(\mathbf{v}_n) = L(\mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n) = L(\mathbf{0}_V) = \mathbf{0}_V.$$

Toisaalta, kertomalla yhtälö 3.8 puolittain skalaarilla k_n saadaan yhtälö

$$k_n\mathbf{v}_1 + k_n\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n = \mathbf{0}_V. \tag{3.9}$$

Vähentämällä yhtälö (3.9) yllä todistetusta yhtälöstä

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n = \mathbf{0}_V$$

saadaan yhtälö

$$(k_1 - k_n)\mathbf{v}_1 + \ldots + (k_{n-1} - k_n)\mathbf{v}_{n-1} = \mathbf{0}_V,$$

jossa $\mathbf{v}_n$ on ”eliminoitu pois”. Tämä yhtälö voidaan kirjoittaa muotoon

$$\mathbf{w}_1 + \ldots + \mathbf{w}_{n-1} = \mathbf{0}_V, \tag{3.10}$$

missä $\mathbf{w}_i = (k_i - k_n)\mathbf{v}_i \in V_{k_i}$ kohdan (1) nojalla, $i = 1, \ldots, n-1$. Induktio-oletuksen mukaan summa $\sum_{i=1}^{n-1} V_{k_i}$ on suora. Soveltamalla tätä tietoa ja Lemman 2.155 kohtaa (2), nähdään yhtälöstä (3.10), että

$$(k_i - k_n)\mathbf{v}_i = \mathbf{w}_i = \mathbf{0}_V$$

kaikilla $i = 1, \ldots, n-1$. Koska oletamme, että ominaisarvot k_i ovat kaikki eri alkioita, pätee $k_i - k_n \neq 0_K$. Vektoriavaruuden supistusääntöjen avulla (kts. Lemmaa 2.5) tällöin saadaan, että $\mathbf{v}_i = \mathbf{0}_V$ kaikilla $i = 1, \ldots, n-1$. Silloin yhtälöstä (3.8) seuraa, että myös $\mathbf{v}_n = 0$. On osoitettu, että summa $\sum_{i=1}^{n} V_{k_i}$ toteuttaa Lemman (2.155) kohdan (ii). Näin ollen summa $\sum_{i=1}^{n} V_{k_i}$ on suora.

(iii) Olkoot $k_1, \ldots, k_n$ (joitakin) operaattorin L (eri) *ominaisarvoja*. Tällöin (i) kohdan mukaan V_{k_i} on epätriviaali aliavaruus jokaisella $i = 1, \ldots, n$, erityisesti

$$\dim V_{k_i} \geq 1$$

kaikilla $i = 1, \ldots, n$. Toisaalta, (ii) kohdan nojalla summa $\sum_{i=1}^{n} V_{k_i}$ on suora. Propositiosta 2.160 seuraa tällöin, että

$$\dim V \geq \dim(\oplus_{i=1}^{n} V_{k_i}) \geq \sum_{i=1}^{n} 1 = n.$$

Toisin sanoen operaatorin L mahdollisten ominaisarvojen lukumäärä on korkeintaan $\dim V$. □

Ominaisarvoaliavaruuden V_k dimensiota $\dim V_k$ sanotaan ominaisarvon k *geometriseksi kertaluvuksi.*

Edellisen lemman nojalla ominaisarvoaliavaruus V_k on invariantti operaattorin L suhteen, erityisesti on olemassa rajoittumakuvaus $L|V_k \colon V_k \to V_k$. Itse asiassa ominaisarvon määritelmästä seuraa suoraan, että tämä kuvaus on yksinkertaisesti lineaarinen *skaalaus* eli alkiolla k kertominen, $L|V_k = k\,\mathrm{id}_{V_k}$.

Olkoon V äärellisulotteinen vektoriavaruus ja olkoon $L \colon V \to V$ lineaarinen operaattori. Lemman 3.7 nojalla operaattorilla L on äärellinen määrä ominaisarvoja. Olkoot $k_1, \ldots, k_n$ kaikki operaatorin L (eri) ominaisarvot. Lemman 3.7 nojalla on olemassa avaruuden V aliavaruus

$$V' = \oplus_{i=1}^{n} V_{k_i},$$

joka on kaikkien operaatorin L ominaisvektorien virittää aliavaruus. Yleisesti ottaen voi tietenkin käydä niin, että V' on avaruuden V aito aliavaruus. Itse asiassa operaattorilla ei välttämättä tarvitse olla ominaisvektoreita ollenkaan, jolloin $V' = \{\mathbf{0}_V\}$ on triviaali avaruus. Esimerkiksi tason $\mathbb{R}^2$ epätriviaali kierto (origon ympäri) on operaattori, jolla ei ole ominaisarvoja lainkaan.

Jos yllä tarkastellussa tilanteessä kuitenkin pätee

$$V' = \oplus_{i=1}^{n} V_{k_i} = V,$$

operaattorin L sanotaan olevan *diagonalisoituva*. Syy tämän termin valinnalle on esitetty seuraavassa kappaleessa ja sitä seuraavassa Lemmassa 3.11. Vektoriavaruuden V operaattori on siis diagonalisoituva jos ja vain jos sen ominaisvektorit virittävät koko avaruuden V.

Olkoon $A = (a_{ij}) \in M(n \times n; K)$ neliömatriisi. Matriisia A sanotaan *diagonaalimatriisiksi* jos kaikilla $i \neq j$ pätee $a_{ij} = 0_K$. Toisin sanoen diagonaalimatrisiilla voi olla nollasta eroavia alkioita vain *diagonaalilla* (alkiot muotoa a_{ii}). Tällainen matriisi näyttää seuraavanlaiselta:

$$\begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn}. \end{bmatrix}$$

Lemma 3.11. *Olkoon V äärellisulotteinen K-vektoriavaruus ja olkoon $L \colon V \to V$ lineaarinen operaattori. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(i) L on diagonalisoituva.

(ii) Avaruudella V on olemassa kanta E, jonka suhteen operaattorin L matriisi $[L]_E$ on diagonaalimatriisi.

(iii) Avaruudella V on olemassa kanta, joka koostuu operaattorin L ominaisvektoreista.

Todistus. Olkoon operaattori $L \colon V \to V$ diagonalisoituva. Tämä tarkoittaa sitä, että

$$V = \oplus_{i=1}^{n} V_{k_i},$$

missä $k_1, \ldots, k_n$ ovat operaattorin L eri ominaisarvot. Valitaan jokaisella $i = 1, \ldots, n$ ominaisarvoaliavaruuden V_{k_i} kanta E_i. Koska ominaisarvoaliavaruuksien summa on suora ja sen arvo on koko avaruus V, näiden kantojen yhdiste E on koko avaruuden V kanta. Lisäksi jokaiselle tämän kannan alkiolle $\mathbf{e}$ pätee $L(\mathbf{e}) = k\mathbf{e}$ jollakin $k \in K$, ominaisvektorin määritelmän mukaan. Tästä helposti seuraa, että kuvauksen L matriisi $[L]_E$ kannan E suhteen on diagonaalimatriisi. Lisäksi konstruktion perusteella E koostuu (eräistä) kuvauksen L ominaisvektoreista. Näin ollen ehdosta (i) seuraavat molemmat ehdot (ii) ja (iii).

Oletetaan, että kuvauksen L matriisi $[L]_E = (a_{ij})_{i,j=1}^m$ avaruuden V kannan $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$ suhteen on diagonaalimatriisi. Tämä tarkoittaa täsmälleen siitä, että jokaisella $i = 1, \ldots, n$ pätee

$$L(\mathbf{e}_i) = a_{ii}\mathbf{e}_i.$$

Toisin sanoen kannan E jokainen alkio on operaatorin L ominaisvektori. On osoitettu, että ehdosta (ii) seuraa ehto (iii).

Oletetaan ehto (iii), toisin sanoen oletetaan, että avaruudella V on olemassa kanta $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$, joka koostuu operaatorin L ominaisvektoreista. Olkoon

$$V' = \oplus_{i=1}^n V_{k_i},$$

missä $k_1, \ldots, k_n$ ovat operaatorin L kaikki eri ominaisarvot (joita on äärellinen määrä Lemman 3.7 nojalla). Tällöin $\mathbf{e}_i \in V'$ kaikilla $i = 1, \ldots, n$. Koska avaruus V on kannan E virittämä, tästä seuraa, että $V \subset V'$. Koska triviaalisti pätee $V' \subset V$, ollaan todistettu, että

$$V' = \oplus_{i=1}^n V_{k_i} = V.$$

Näin ollen L on diagonalisoituva. □

Matriisin ominaisarvot ja ominaisvektorit

Olkoon $A \in M(n \times n; K)$ neliömatriisi. Olkoon $L_A \colon K^n \to K^n$ siihen liittyvä kanoninen lineaarinen kuvaus, joka on määritelty kaavalla $L_A(\mathbf{v}) = A\mathbf{v}$ (missä $\mathbf{v}$ tulkitaan pystyvektorina), $\mathbf{v} \in V$. Kuvaus L_A on yksikäsitteinen kuvaus jonka matriisi $[L]_E$ avaruuden K^n standardikannan suhteen on matriisi A.

Skalaarikunnan alkio $k \in K$ sanotaan matriisin A *ominaisarvoksi*, jos k on kuvauksen L_A ominaisarvo. Vastaavasti vektori $\mathbf{v} \in V, \mathbf{v} \neq 0$ on matriisin A *ominaisvektori* jos se on kuvauksen L_A ominaisvektori. Ominaisarvoon liittyvä aliavaruus määritellään vastaavasti kuvauksen L_A avulla.

Määritelmien mukaan $\mathbf{v} \neq \mathbf{0}_V$ on matriisin ominaisvektori jos on olemassa $k \in K$ siten, että $A\mathbf{v} = k\mathbf{v}$. Vastaavasti $k \in K$ on matriisin A ominaisarvo jos on olemassa vektori $\mathbf{v} \neq \mathbf{0}_V$ siten, että $A\mathbf{v} = k\mathbf{v}$. Ominaisarvoon k liittyvä aliavaruus on

$$V_k = \{\mathbf{v} \in V \mid A\mathbf{v} = k\mathbf{v}\}.$$

Neliömatriisia A kutsutaan *diagonalisoituvaksi* jos kuvaus L_A on diagonalisoituva operaattori.

Lemma 3.12. *a) Olkoon E jokin äärellisulotteisen avaruuden V kanta. Olkoon $L \colon V \to V$ operaattori. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(i) *Operaattori L on diagonalisoituva.*

(ii) *Matriisi $[L]_E$ on diagonalisoituva.*

(iii) *On olemassa kääntyvä $(n \times n)$-matriisi J siten, että matriisi $J[L]_E J^{-1}$ on diagonaalimatriisi.*

b) Olkoon $A \in M(n \times n; K)$ neliömatriisi. Tällöin A on diagonalisoituva jos ja vain jos on olemassa kääntyvä $(n \times n)$-matriisi J siten, että matriisi JAJ^{-1} on diagonaalimatriisi.

Todistus. Olkoon E jokin kiinnitetty äärellisulotteisen avaruuden V kanta ja olkoon $L\colon V \to V$ operaattori. Olkoon E' jokin (mahdollisesti toinen) avaruuden V kanta. Tällöin $[L]_{E'} = J[L]_E J^{-1}$, missä $J = [E' \mid E]$ on kannanvaihtomatriisi (kts. 2.87). Toisaalta, kääntäen helposti nähdään (harjoitustehtävä), että kiinteällä kannalla E mikä tahansa *kääntyvä* neliömatriisi J on muotoa $[E' \mid E]$ jollakin (yleensä toisella) avauruuden kannalla E'. Tällöin kannavaihtokaavan 2.87 nojalla seuraa, että $J[L]_E J^{-1} = [L]_{E'}$.

Edellisen Lemman 3.11 nojalla operaattori L on diagonalisoituva jos ja vain jos löytyy sellainen avaruuden V kanta E', jolle matriisi $[L]_{E'}$ on diagonaalimatriisi. Yhdistämällä tätä edellisen kappaleen havaintoihin, nähdään, että kuvaus L on diagonalisoituva jos ja vain jos on olemassa sellainen kääntyvä matriisi J siten, että matriisi $J[L]_E J^{-1}$ on diagonaalimatriisi. Ollaan siis todistettu, että a)-kohdan ehdot (i) ja (iii) ovat yhtäpitäviä.

Sovelletaan tätä tulosta kuvaukseen $L_A\colon K^n \to K^n$ ja avaruuden K^n *standardikantaan* E. Määritelmän mukaan matriisi A on diagonalisoituva jos ja vain jos L_A on diagonalisoituva operaattori. Edellisen nojalla jälkimmäinen ehto on yhtäpitävä sen kanssa, että on olemassa kääntyvä matriisi J jolle $J[L_A]_E J^{-1} = JAJ^{-1}$ on diagonaalimatriisi. Tämä todistaa b)-kohdan väitteen.

Soveltamalla b)-kohdan väitettä matriisiin $A = [L]_E$, nähdään, että $[L]_E$ on diagonalisoituva jos ja vain jos $J[L]_E J^{-1}$ on diagonaalimatriisi jollakin käntyvällä J. Tämä osoittaa sen, että a)-kohdan väitteet (ii) ja (iii) ovat yhtäpitäviä. □

Diagonaalimatriisit ovat käteviä, koska niitä on erityisen helppo käsitellä. Esimerkiksi, olkoon $A = (a_{ij})$ diagonaalimatriisi. Tällöin sen determinantti on

$$\det A = a_{11}a_{22}\dots a_{nn}$$

eli yksinkertaisesti diagonaalialkioiden tulo. Tämä nähdään helposti kehittämällä $\det A$ vaikkapa ensimmäisen sarakkeen suhteen ja jatkamalla induktiolla.

Toinen tunnettu esimerkki diagonaalimatriisien sovelluksista liittyy matriisin potenssien laskemiseen (kts. esimerkki 3.16, 2 alla).

Näistä (ja muistakin) syistä on tärkeää tietää, milloin kuvaus (tai matriisi) on diagonalisoituva. Seuraavaksi esitetään eräs yksinkertainen *riittävä* (mutta ei välttämätön!) diagonalisoituvuuden ehto.

Lemma 3.13. *Olkoon $L\colon V \to V$ lineaarinen endomorfismi ja oletetaan, että avaruuden V dimensio on $n \in \mathbb{N}$. Oletetaan, että operaattorilla L on täsmälleen n eri ominaisarvoa $k_1, \dots, k_n$. Tällöin L on diagonalisoituva.*

Todistus. Jos ominaisarvoja on n, niin aliavaruudelle $V' = \oplus_{i=1}^{n} V_{k_i}$ pätee (Proposition 2.160 nojalla)

$$\dim V' = \sum_{i=1}^{n} \dim V_{k_i} \geq \sum_{i=1}^{n} 1 = n,$$

koska jokaisen ominaisarvoaliavaruuden dimensio on vähintään yksi. Toisaalta triviaalisti pätee $\dim V' \leq \dim V = n$. Näin ollen $\dim V' = n$. Koska n-ulotteisen vektoriavaruuden aidon aliavaruuden dimensio on aina aidosti pienempi kuin n (Lemma 2.48), tästä voidaan päätellä, että $V' = V$. □

Edellisen lemman antama diagonaalisoitumisen ehto on riittävä, mutta ei välttämätön - L voi olla diagonalisoituva vaikka sillä olisi vähemmän kuin luvun $\dim V$ verran erilaisia ominaisarvoa. Esimerkiksi mille tahansa äärellisulotteiselle vektoriavaruudelle V identtisen kuvauksen $\mathrm{id}\colon V \to V$ matriisi minkä tahansa kannan suhteen on diagonaalimatriisi I_n, joten se on varmasti diagonalisoituva, vaikka sillä on vain yksi ominaisarvo $1_K \in K$.

Ominaisarvot polynomiyhtälön juurina

Toistaiseksi ei esitetty mitään yleispäteviä menetelmiä, joilla operaatorin ominaisarvoja voittaisi selvittää. Yksi tällainen menetelmä, jota on hyödyllinen sekä käytännön laskuissa, että erityisesti teorettisissa tarkasteluissa, perustuu determinaanttien käyttöön.

Olkoon $L\colon V \to V$ vektoriavaruuden V operaattori ja olkoon $k \in K$ skalaari. Tällöin jokaisella $\mathbf{v} \in V$ yhtälö

$$L(\mathbf{v}) = k\mathbf{v} = k\,\mathrm{id}_V(\mathbf{v})$$

on voimassa jos ja vain jos on voimassa yhtälö

$$L'(\mathbf{v}) = (k\,\mathrm{id}_V - L)(\mathbf{v}) = 0.$$

Tässä $L' = k\,\mathrm{id}_V - L\colon V \to V$ on myös eräs avaruuden V lineaarinen operaattori. Edellisen nojalla nähdään, että aliavaruus V_k voidaan yhtä hyvin ilmaista tämän operaattorin *ytimenä*,

$$V_k = \mathrm{Ker}(k\,\mathrm{id}_V - L).$$

Tästä seuraa, että $k \in K$ on operaattorin L ominaisarvo *jos ja vain jos* operaattorin $(k\,\mathrm{id}_V - L)$ ydin on epätriviaali, eli toisin sanoen täsmälleen silloin, kun tämä operaattori *ei ole injektio.*

Oletetaan seuraavaksi, että V on äärellisulotteinen. Tällöin Lemman 2.94 nojalla operaattori $(k\,\mathrm{id}_V - L)$ ei ole injektio jos ja vain jos se ei ole isomorfismi. Seurauksen 2.148 mukaan tämä on taas yhtäpitävää sen kanssa, että

$$\det(k\,\mathrm{id}_V - L) = 0_K. \tag{3.14}$$

Olkoon $A = [L]_E$ operaatorin L matriisi jonkun avaruuden V kannan E suhteen. Tällöin ehto (3.14) voidaan kirjoittaa yhtäpitävästi muodossa

$$\det(kI - A) = 0_K. \tag{3.15}$$

Tässä $I = I_n$ on yksikkömatriisi (joka on identtisen kuvauksen matriisi minkä tahansa kannan suhteen), $n = \dim V$.

Kirjoitetaan ehto (3.15) auki. Olkoon

$$A = [L]_E = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}, \end{bmatrix},$$

missä E on jokin kiinnitetty äärellisavaruuden V kanta. Tällöin

$$kI - A = \begin{bmatrix} k - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & k - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & k - a_{nn}. \end{bmatrix}$$

Lasketaan tämän matriisin determinantti $\det(kI - A)$ kaavalla (2.139). Merkitään $B = kI - A$. Kaavan (2.139) nojalla pätee

$$\det(kI - A) = \det B = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{\sigma(i)i}.$$

Toisin sanoen $\det(A - kI)$ on summa termeistä, joista jokainen on (merkkiä $\operatorname{sgn}(\sigma)$ vaille) tulo tasan n:stä alkiosta muotoa b_{ij}. Matriisin $B = kI - A$ määritelmän mukaan pätee

$$b_{ij} = \begin{cases} k - a_{ii}, \ i = j, \\ -a_{ij}, \ i \neq j. \end{cases}$$

Pidetään k muuttujana ja matriisi A (eli myös kertoimia a_{ij}) vakiona. Tällöin $\det(kI - A)$ on summa tuloista muotoa $\operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{\sigma(i)i}$, jossa jokainen tulon tekijä on joko vakio (matriisin A kerroin) tai *ensimmäisen asteen lineaarinen lauseke* muotoa $(k - a_{ii})$ (muuttujan $k \in K$ suhteen). Tästä seuraa, että muuttujan k funktiona lauseke $\det(kI - A)$ on *polynomifunktio*, jonka aste on korkeintaan n, toisin sanoen funktio $p\colon K \to K$, joka voidaan kirjoittaa muodossa

$$p(k) = c_n k^n + c_{n-1} k^{n-1} + \ldots + c_1 k + c_0$$

joillakin vakioilla $c_n, c_{n-1}, \ldots, c_0 \in K$. Lisäksi tämän polynomin aste on *tasan* n ja itse asiassa sen *johtava kerroin* c_n on 1_k. Osoitetaan tämä. Olkoon $\sigma \in S_n$ permutaatio. Jos $\sigma \neq \mathrm{id}_n$ ei ole identtinen kuvaus, tulo $\prod_{i=1}^{n} b_{\sigma(i)i}$ sisältää ainakin yhden vakiotermin $b_{ij} = a_{ij}$, $i \neq j$. Tällöin vastaava lauseke $\operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{\sigma(i)i}$ on korkeintaan $(n-1)$-asteinen polynomi k:n suhteen. Jos taas $\sigma = \mathrm{id}_n$, vastaava termi on n-asteinen polynomi

$$\operatorname{sgn}(\mathrm{id}_n) \prod_{i=1}^{n} (k - a_{ii})^n = k^n + d_{n-1} k^{n-1} + \ldots$$

muuttujan k suhteen, jonka johtava kerroin on 1_K. Tästä seuraa, että myös summa

$$p(k) = \det(kI - A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{\sigma(i)i}$$

on n-asteinen polynomifunktio muuttujan k suhteen ja sen johtava kerroin on 1_K. Tätä polynomia sanotaan matriisin A *karakteristiseksi* polynomiksi.

Jos ollaan ihan tarkkoja, mielivaltaisen kunnan K tapauksessa polynomifunktiolla $p\colon K \to K$ ei välttämättä ole hyvinmääriteltyä astetta, sillä erinäköiset polynomilausekkeet saattavatkin määritellä saman funktion. Esimerkiksi kun $K = \mathbb{Z}_2$ polynomifunktio x^2 ja x ovatkin sama funktio, kuten helposti nähdään laskemalla kummankin arvot jokaisessa kahden alkion kunnan $\mathbb{Z}_2$ alkiossa. Näin ollen ei ole välttämättä korrektia puhua polynomifunktion asteesta ja johtavasta kertoimesta, ainoastaan tietyn polynomilausekkeen esityksen asteesta ja sen johtavasta kertoimesta. Näin ollen edellisessä kappaleessa on oikeastaan näytetty vain, että muuttujan $k \in K$ suhteen funktio $\det(k\,\mathrm{id}_V - A)$ on sellainen funktio, jonka *voidaan esittää* astetta n olevana polynomifunktiona, jonka johtava kerroin on 1_K.

Päädytään siis siihen, että äärellisulotteisen vektoriavaruuden lineaarisen operaatorin $L\colon V \to V$ ominaisarvojen selvittäminen palauttuu *polynomiyhtälön* $\det(kI - [L]_E) = 0$ ratkaisemiseen liittyväksi ongelmaksi. Tässä E on jokin avaruuden V kanta.

Esimerkkejä 3.16. *1) Olkoon K kunta. Tarkastellaan operaattoria $L\colon K^2 \to K^2$, jonka matriisi avaruuden K^2 standardikannan suhteen on*

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Tällöin $L(x, y) = (x + y, y)$ kaikilla $(x, y) \in K^2$. Lasketaan operaattorin L ominaisarvot. Edellisen nojalla kuvauksen L ominaisarvot ovat toisen asteen polynomiyhtälön

$$0 = \det \begin{bmatrix} k-1 & -1 \\ 0 & k-1 \end{bmatrix} = (k-1)^2$$

ratkaisuja. Tällä yhtälöllä on jokaisessa kunnassa K tasan yksi ratkaisu $k = 1_K$. Näin ollen operaattorilla L on vain yksi ominaisarvo. Jos $(x, y) \in V_1$ on ominaisvektori, niin pätee

$$L(x, y) = (x + y, y) = (x, y).$$

Tämä yhtälö selvästi toteutuu tasan silloin kun $x + y = x$, eli silloin kun $y = 0$. Näin ollen ominaisvektorialiavaruus on 1-ulotteinen aliavaruus
$W = \{(x, 0) \mid x = 0\}$ (tason x-akseli). W on L-invariantti aliavaruus.

Ainoat kuvauksen L avaruuden V invariantit aliavaruudet ovat $W = V_1$ sekä triviaalit aliavaruudet $\{0\}$ ja K^2. Tämä nähdään seuraavasti - jos W' on ei-triviaali invariantti aliavaruus, sen on oltava 1-ulotteinen (sillä koko avaruus on 2-ulotteinen). Näin ollen sen jokaisen virittäjän on oltava ominaisvektori ja olla siis johonkin ominaisarvoon liittyvä ominaisvektori. Koska 1 on kuvauksen ainoa ominaisarvo, täytyy olla $W' = W = V_1$.

Näin ollen kuvaus L tarjoaa samalla esimerkin tilanteesta, jossa invariantin aliavaruuden jokainen komplementti ei ole invariantti. Lisäksi operaattori L ei ole diagonalisoituva, sillä sen ominaisvektorien virittämä aliavaruus on aito aliavaruus $V_1 = W$.

2) Eräs sovellus matriisin diagonalisoituvuus-ongelmalle liittyy neliömatriisin A potenssien A^m *laskimiseen, $m \in \mathbb{N}$. Yleisesti ottaen matriisin kaikkien potenssien selvittäminen on vaikea ongelma, eikä ole olemassa yksinkertaista nopeata tapaa laskea matriisin A potensseja A^m kaikilla $m \in \mathbb{N}$, joka olisi kelvollinen sovelluksissa. Kuitenkin, jos A on diagonaalimatriisi, sen potenssien laskeminen on käytännössä triviaalia, sillä jos*

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}, \text{ niin } A^k = \begin{bmatrix} a_{11}^k & 0 & \dots & 0 \\ 0 & a_{22}^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^k \end{bmatrix}$$

kaikilla $k \in \mathbb{N}$ (mieti yksityiskohtia!).

Yleisemmin, olkoon matriisi A diagonalisoituva. *Tällöin on olemassa kääntyvä matriisi J siten, että $A = JDJ^{-1}$, missä D on diagonaalimatriisi (Lemma 3.12). Induktiolla nähdään helposti, että silloin kaikilla $m \in \mathbb{N}$ pätee yhtälö $A^m = JD^mJ^{-1}$. Jos J ja D osataan laskea, tästä saadaan myös potenssi A^m laskettua helposti.*

Näytetään, miten tämän havainnon avulla voidaan johtaa suljettu kaava kuuluisille Fibonaccin luvuille. Fibonaccin lukuja a_m määritellään rekursiivisesti ehdoilla

$$a_0 = 0, a_1 = 1, a_{m+2} = a_{m+1} + a_m, m \geq 0.$$

Fibonaccin lukujono alkaa $(0, 1, 1, 2, 3, 5, 8, 13, 21, \dots)$. Jokaisella $m \in \mathbb{N}$ m:äs Fibonaccin luku a_n voidaan tietysti periaatteessa laskea rekursiivisesti edellisten jonon jäsenten avulla, mutta tämä vaatii aikaa, eikä ole isoilla m kovinkaan käytännöllistä. Lisäksi tämä tapa antaa jonon alkioiden arvoja vain tietyn indeksin asti. Onko mahdollista antaa Fibonaccin luvulle a_m valmis kaava, joka riippuisi vain sen numerosta m ja josta sen arvo voittaisin laskea suoraan?

Olkoon $(x, y) = (a_m, a_{m+1})$ kahden peräkkäisen Fibonaccin luvun muodostama pari. Tällöin seuraavan kahden peräkkäisen Fibonaccin luvun muodostama pari (a_{m+1}, a_{m+2}) on Fibonaccin lukujen määritelmän mukaan pari $(y, x + y)$. Tämä havainto antaa idean tarkastella $\mathbb{R}$-lineaarista kuvausta $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, joka on määritelty kaavalla $L(x, y) = (y, x+y)$. Tämän operaattorin matriisi tason $\mathbb{R}^2$ standardikannan suhteen on

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Tässä $(a_0, a_1) = (0, 1) = L(\mathbf{e}_1)$ ja $(a_1, a_2) = (1, 1) = L(\mathbf{e}_2)$. Induktiolla helposti nähdään, että kaikilla $m \in \mathbb{N}$ pätee

$$A^m = \begin{bmatrix} a_{m-1} & a_m \\ a_m & a_{m+1} \end{bmatrix}.$$

Jos siis osattaisimme laskea potenssin A^m arvoja suoraan, tulisi samalla johdettua kaavan Fibonaccin jonon mielivaltaiselle jäsenelle a_m.

Tätä varten tutkitaan, onko A diagonalisoituva, toisin sanoen tutkitaan sen ominaisarvoja ja ominaisvektoreita. Ominaisarvot ovat polynomiyhtälön

$$\det(kI - A) = \det \begin{bmatrix} k & -1 \\ -1 & k-1 \end{bmatrix} = k(k-1) - 1 = k^2 - k - 1 = 0$$

ratkaisut. Tällä toisen asteen yhtälöllä on $\mathbb{R}$:ssä kaksi juurta, $k_1 = \frac{1}{2}(1+\sqrt{5})$ ja $k_2 = \frac{1}{2}(1-\sqrt{5})$. Lemmasta 3.13 tällöin seuraa, että A on diagonalisoituva (koska $\mathbb{R}^2$ on 2-ulotteinen). Lemman 3.12 nojalla $A = JDJ^{-1}$, missä

$$D = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$$

ja J on eräs kannanvaihtomatriisi. Matriisi J puolestaan saadaan selville laskemalla ominaisvektoreita, sillä tällöin $J = [E \mid E']$, missä E on standardikanta ja E' on jokin ominaisvektoreista koostuva kanta. Sopivan matriisin J selvittäminen jätetään lukijalle harjoitustehtäväksi. Käytämällä hyväksi sitä, että $A^k = JD^kJ^{-1}$, voidaan johtaa Fibonaccin luvulle a_n "suljettu "kaava

$$a_n = \frac{1}{\sqrt{5}}\Big(\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \Big(\frac{1-\sqrt{5}}{2}\Big)^n\Big).$$

Tämä ratkaisutapa yleistyy muidenkin rekurssiivisesti määriteltyjen jonojen tutkimiseen.

3) Tarkastellaan kuvausta $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, jonka matriisi standardikannan suhteen on

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Kyseessä on siis kaavalla $L(x,y) = (-y,x)$ määritelty lineaarinen kuvaus $\mathbb{R}^2 \to \mathbb{R}^2$. Lasketaan sen ominaisarvot. Nämä ovat toisen asteen yhtälön

$$0 = \det \begin{bmatrix} k & 1 \\ -1 & k \end{bmatrix} = k^2 + 1$$

ratkaisut. Reaalilukujen alueessa $\mathbb{R}$ tällä yhtälöllä ei tunnetusti ole juuria, joten operaatorilla L ei ole ominaisarvoja ollenkaan. Tämä ei ole kovin yllättävää, sillä geometrisesti L on itse asiassa tason kierto origon ympäri 90 astetta vastapäivään, joten on selvä, että se ei voi säilyttää minkään nollasta eroavan vektorin suuntaa.

Seuraavaksi tarkastellaan $\mathbb{C}$-vektoriavaruuden $\mathbb{C}^2$ operaattoria $L\colon \mathbb{C}^2 \to \mathbb{C}^2$, joka on määritelty samalla kaavalla $L(x,y) = (-y,x)$. Tällä operaattorilla on avaruuden $\mathbb{C}^2$ standardikannan suhteen sama matriisi

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Ominaisarvot ovat edelleenkin polynomiyhtälön $k^2 + 1 = 0$ ratkaisuja, mutta nyt tämä yhtälö ratkaistaan kompleksilukujen alueessa $\mathbb{C}$. Koska polynomiyhtälöllä $k^2 + 1 = 0$ on kompleksilukujen alueessa $\mathbb{C}$ tasan kaksi juurta, kompleksiluvut i ja $-i$, operaattorilla $L\colon \mathbb{C}^2 \to \mathbb{C}^2$ on jopa kaksi ominaisarvoa. Koska $\dim_{\mathbb{C}} \mathbb{C}^2 = 2$, Lemmasta 3.13 seuraa tämän nojalla, että $L\colon \mathbb{C}^2 \to \mathbb{C}^2$ on diagonalisoituva. Lasketaan vastaavat ominaisvektorit.

Vektori $(x, y) \in \mathbb{C}^2$ *on ominaisarvoa* i *vastaava ominaisvektori jos ja vain jos*

$$L(x, y) = (-y, x) = (ix, iy) = i(x, y).$$

Tästä nähdään, että $x = iy$*, joten*

$$V_i = \{(x, y) \in \mathbb{C}^2 \mid x = iy\} = \mathrm{Span}\{(i, 1)\}.$$

Vastaavalla tavalla nähdään, että

$$V_{-i} = \{(x, y) \in \mathbb{C}^2 \mid x = -iy\} = \mathrm{Span}\{(i, -1)\}.$$

Operaattorin L *matriisi kannan* $((i, 1), (i, -1))$ *suhteen on diagonaalimatriisi*

$$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Viimeinen esimerkki havainnollistaa miten ominaisarvojen olemassaolo ja operaattorin diagonalisoituvuus riippuvat *skalaarikunnan* K ominaisuuksista. Tämä johtuu pohjimmiltaan siitä, että polynomiyhtälöt käyttäytyvät eri tavalla eri kunnissa.

Esimerkissä 3.16, (3) yllä tarkasteltiin operaattoria $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, jolla ei ole ominaisarvoja ollenkaan - juuri siitä syystä, että toisen asteen yhtälöllä ei välttämättä ole ratkaisuja $\mathbb{R}$:ssä. Kun taas samalla tavalla määriteltyä operaattoria tarkastellaan kompleksilukujen yli, ominaisarvoja löytyy, koska vastaavalla toisen asteen yhtälöllä on ratkaisuja $\mathbb{C}$:ssä. Itse asiassa jokaisella $\mathbb{C}$-kertoimisella ei-vakio polynomilla on olemassa kompleksilukualuessa ainakin yksi *juuri*, sillä $\mathbb{C}$ on esimerkki niin sanotusta *algebrallisesti suljetusta kunnasta*.

Algebrallisesti suljetut kunnat

Kuntaa K sanotaan *algebrallisesti suljetuksi*, jos jokaisella polynomifunktiolla $p\colon K \to K$, joka on muotoa

$$p(x) = c_n x^n + c_{n-1} x^{n-1} + \ldots + c_1 x + c_0,$$

missä $n \geq 1, c_n \neq 0, c_0, \ldots, c_n \in K$, on olemassa kunnassa K ainakin yksi *juuri*, eli sellainen $k \in K$ jolle pätee $p(k) = 0$. Yllä siis oletetaan, että p on sellainen funktio, joka voidaan esittää n-asteisena polynomifunktiona, jossa $n \geq 1$. Määritelmä on pakko muotoilla näin, sillä mielivaltaisen kunnan K kohdalla polynomifunktion astetta ei välttämättä voida määritellä yksiselitteisesti. Myöhemmin käsitteelle annetaan myös yhtäpitävä määritelmä, joka perustuu algebrallisten polynomien, ei polynomifunktioiden käyttöön.

Oletetaan, että kunta K on algebrallisesti suljettu kunta, V on epätriviaali äärellisulotteinen K-vektoriavaruus ja $L\colon V \to V$ on lineaarinen operaattori. Tällöin operaattorilla L on ainakin yksi ominaisarvo. Tämä johtuu siitä, että tällöin $\det(k\,\mathrm{id}_V - L) = 0$ on edellisen aliluvun tarkastelujen nojalla polynomiyhtälö. Lisäksi tässä $\det(k\,\mathrm{id}_V - L)$ voidaan esittää polynomifunktiona, jonka aste on $n = \dim V$. Erityisesti algebrallisesti suljetun kunnan määritelmän nojalla tällä yhtälöllä on ainakin yksi ratkaisu, joka on tällöin operaattorin L ominaisarvo.

Seuraavat kaksi tulosta ovat erittäin tärkeitä teorian kannalta, mutta niiden todistukset ovat liaan työläitä ja aikaa vieviä, joten ne esitetään ilman todistuksia. Ensimmäinen väite osoitetaan todeksi esimerkiksi Algebra II-kurssilla, toinen taas millä tahansa kompleksianalyysin peruskurssilla[1]. Kurssin kotisivulta löytyy myös opiskelijan tekemä esitelmä, jossa myös annetaan todistuksia sekä Proposition 3.17 että Proposition 3.18 väitteille.

Propositio 3.17. *Jokainen kunta K voidaan "upottaa" algebrallisesti suljettuun kuntaan. Täsmällisesti sanoen jos K on kunta, on olemassa algebrallisesti suljettu kunta K' ja injektiivinen homomorfismi $f\colon K \to K'$. Tällöin voidaan samaistaa K ja sen kanssa isomorfinen kunta $f(K)$, jolloin K on algebrallisesti suljetun kunnan K' alikunta.*

Propositio 3.18. Algebran peruslause.
Kompleksilukujen kunta $\mathbb{C}$ on algebrallisesti suljettu.

Propositiossa 3.17 riittää olettaa, että on olemassa kunta-homomorfismi $K \to K'$, sillä mikä tahansa homomorfismi kuntien välillä on injektiivinen (Seuraus 1.93). Tämän proposition väitettä voidaan vielä tarkentaa seuraavasti. Nimittäin voidaan todistaa, että jokaiselle kunnalle K löytyy jopa "pienin" algebrallisesti suljettu kunta K', joka sisältää sen alikuntana. "Pienin" tässä yhteydessä tarkoittaa sitä, että jos K'' on toinen algebrallisesti suljettu kunta, joka sisältää K:n alikuntana, se sisältää myös K':n alikuntana (isomorfiaa vaille). Tällaista pienintä algebrallisesti suljettua kuntaa, joka sisältää kunnan K, sanotaan K:n *algebralliseksi sulkeumaksi*. Kaikki nämä väitteet formalisoidaan ja todistetaan Algebra II-kurssilla.

Voidaan osoittaa, että reaalilukujen kunnan $\mathbb{R}$ algebrallinen sulkeuma on kompleksilukujen kunta $\mathbb{C}$. Rationaalilukujen kunnan $\mathbb{Q}$ algebrallinen sulkeuma Q' ei ole kunta $\mathbb{R}$, eihän $\mathbb{R}$ edes ole algebrallisesti suljettu, mutta se ei ole myöskään kunta $\mathbb{C}$. Voidaan osoittaa, että Q' on $\mathbb{C}$:n aito alikunta, joka koostuu niin sanotuista *algebrallisista luvuista*, eli sellaisista kompleksiluvuista, jotka ovat kokonaislukukertoimisten polynomien juuria. Esimerkiksi kompleksiluvut $\sqrt{2}$ tai $\sqrt[3]{7} + i$ ovat algebrallisia (yritä keksiä kokonaislukukertoiminen polynomi, jonka juurena on viimeksi mainittu kompleksiluku). Neperin luku e ja luku π ovat tunnettuja esimerkkejä reaaliluvuista, jotka eivät ole algebrallisia.

Olkoon K algebrallisesti suljettu kunta. Tällöin, kuten yllä on jo todettu, jokaisella äärellisulotteisen K-vektoriavaruuden operaattorilla $L\colon V \to V$ on ainakin yksi ominaisarvo. Tämä johtuu siitä, että polynomiyhtälöllä $\det(k\,\mathrm{id}_V - L)$ täytyy olla ainakin yksi ratkaisu $k \in K$. Kuitenkaan kuvauksen L ei tarvitse olla diagonalisoituva. Paras tulos, joka pätee tässä tapauksessa on se, että L voidaan esittää *yläkolmiomatriisina* (Lemma 3.22 alla).

Olkoon K mielivaltainen kunta (ei välttämättä algebrallisesti suljettu). Neliömatriisia $A = (a_{ij}) \in M(m \times n; K)$ sanotaan *yläkolmiomatriisiksi*, jos kaikki sen alkiot, jotka sijaitsevat "päälävistäjän alapuoleella" ovat nolla-alkioita. Täsmällisemmin sanottuna A on yläkolmiomatriisi jos ja vain jos $a_{ij} = 0_K$ kaikilla $i > j$.

[1]Suhteellisen alkeellinen todistus löytyy seuraavasta Tuomas Hytösen kirjoittamasta artikkelista "Solmu"-lehdessä: `http://solmu.math.helsinki.fi/2011/3/algebran_peruslause.pdf`

Yläkolmiomatriisi siis näyttää seuraavanlaiselta:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Kehittämällä tällainen yläkolmiomatriisi vaikkapa sen ensimmäisen sarakkeen suhteen, nähdään, että $\det A = a_{11} \det B$, missä $B = A_{11}$ on yläkolmiomatriisi, joka saadaan A:sta poistamalla siitä ensimmäinen sarake ja ensimmäinen rivi,

$$B = \begin{bmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Tämä on yläkolmiomatriisi, joka on kooltaan pienempi kuin A. Jatkamalla induktiolla nähdään, että

$$\det A = a_{11} a_{22} \dots a_{nn}.$$

Toisin sanoen *yläkolmiomatriisin determinantti on sen diagonaalialkioiden tulo.*

Olkoon $k \in K$ mielivaltainen skalaari ja $A \in M(m \times n; K)$ yläkolmiomatriisi. Tällöin matriisi $kI_n - A$ on myös yläkolmiomatriisi. Tämän matriisin päälävistäjäalkiot ovat muotoa $k - a_{ii}$, $i = 1, \dots, n$. Edellisen nojalla pätee

$$\det(kI_n - A) = (k - a_{11})(k - a_{22}) \dots (k - a_{nn}).$$

Tästä seuraa, että yläkolmiomatriisin A ominaisarvot ovat täsmälleen sen päälävistäjäalkiot $a_{11}, a_{22}, \dots, a_{nn}$. Erityisesti, jos yläkolmiomatriisin kaikki päälävistäjäalkiot ovat eri alkoita, niin **matriisi on diagonalisoituva** (Lemma 3.13). Jos taas jotkut päälävistäjäalkiot toistuu, matriisi saattaa olla diagonalisoituva tai sitten ei - asia vaatii lisää tutkimusta. Esimerkiksi matriisi

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

on tietysti diagonalisoituva - se on itse asiassa valmiiksi diagonaalimatriisi. Matriisi

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

taas ei ole diagonalisoituva edes algebrallisesti suljetun kunnan $\mathbb{C}$ yli (kts. esimerkki 3.16, 1).

Lemma 3.19. *Olkoon $L\colon V \to V$ äärellisulotteisen K-vektoriavaruuden V operaattori. Olkoon $n = \dim V$. Tällöin on olemassa avaruuden V kanta E, jonka suhteen operaatorin L matriisi $[L]_E$ on yläkolmiomatriisi jos ja vain jos on olemassa nouseva ketju avaruuden V aliavaruuksia*

$$W_1 \subset W_2 \subset \dots \subset W_{n-1},$$

siten, että jokaisella $i = 1, \dots, n-1$ pätee $\dim W_i = i$ ja aliavaruus W_i on invariantti operaattorin L suhteen.

Todistus. Harjoitustehtävä. □

Esimerkki 3.20. *Olkoon P_m $\mathbb{R}$-vektoriavaruus, jonka muodostavat kaikki polynomit $p\colon \mathbb{R} \to \mathbb{R}$, joiden aste on korkeintaan m. Tällöin* $\dim P_m = m+1$. *Olkoon $\mathcal{D}\colon P_m \to P_m$ derivaattaoperaattori, $\mathcal{D}(p) = p'$. Tällöin*

$$P_0 \subset P_1 \subset \ldots \subset P_{m-1}.$$

on nouseva ketju $\mathcal{D}$-invariantteja aliavaruuksia siten, että kaikilla $l = 0, \ldots, m-1$ pätee $\dim P_l = l + 1$ *(esimerkki 3.2). Edellisen lemman mukaan avaruudella P_m on olemassa kanta E, jonka suhteen derivaattaoperaattorin matriisi $[\mathcal{D}]_E$ on yläkolmiomatriisi. Tälläiseksi kannaksi kelpaa itse asiassa "kanoninen" kanta $E = (1, x, x^2, \ldots, x^m)$. Itse asiassa, koska l-asteisen polynomin derivaatan aste on aidosti pienempi kuin l, matriisin $[\mathcal{D}]_E$ päälävistäjäalkiot ovat myös kaikki nolla-alkioita. Esimerkiksi kun $m = 3$, pätee*

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0. \end{bmatrix}$$

Tästä voidaan lisäksi päätellä, että operaatorin $\mathcal{D}$ ainoa ominaisarvo on 0. *Samaan johtopäätökseen ollaan päästy aikaisemmin suoraan esimerkissä 3.6, (1). Tämän esimerkin yhteydessä paljastui, että V_0 on yksiulotteinen aliavaruus, joka koostuu kaikista vakiopolynomeista. Erityisesti $\mathcal{D}$ ei ole diagonalisoituva, kun $m > 1$.*

Esimerkki 3.21. *Operaattorilla $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x, y) = (-y, x)$ ei ole ominaisarvoja (esimerkki 3.16). Tästä seuraa erityisesti, että operaattoria L ei voida esittää yläkolmiomatriisina (sillä yläkolmiomatriisin ensimmäinen sarake edustaa erästä ominaisvektoria).*

Algebrallisesti suljetun kunnan tapauksessa jokainen operaattori voidaan esittää yläkolmiomatriisina.

Propositio 3.22. *Olkoon K algebrallisesti suljettu kunta. Tällöin jokainen lineaarikuvaus $L\colon V \to V$, missä V on äärellisulotteinen K-vektoriavaruus, voidaan esittää yläkolmiomatriisina jonkun kannan suhteen.*

Todistus. Osoitetaan väite induktiolla avaruuden V dimension suhteen. Kun $n = 1$ väite on selvä, sillä jokainen (1×1)-matriisi on triviaalisti yläkolmiomatriisi.

Oletetaan, että väite on tosi avaruuksille, joiden dimensio on korkeintaan $(n-1)$ ja olkoon V n-ulotteinen K-vektoriavaruus. Olkoon $L\colon V \to V$ lineaarinen operaattori. Koska K on algebrallisesti suljettu, operaattorilla L on ainakin yksi ominaisarvo $k_1 \in K$. Olkoon $\mathbf{u} \in V$, $\mathbf{u} \neq \mathbf{0}_V$ jokin ominaisarvoon k liittyvä ominaisvektori, $L(\mathbf{u}) = k_1\mathbf{u}$. Olkoon

$$U = K\mathbf{u} = \{k\mathbf{u} \mid k \in K\} = \mathrm{Span}(\mathbf{u}).$$

Tällöin U on yksiulotteinen aliavaruus, joka on invariantti operaattorin L suhteen. Lisäksi kaikilla $\mathbf{z} \in U$ pätee $L(\mathbf{z}) = k_1\mathbf{z}$.

Lemman 2.159 mukaan aliavaruudella U on olemassa avaruudessa V komplementti W, $V = U \oplus W$. Lemman 2.160 nojalla tällöin pätee $n = \dim V = \dim U + \dim W$, joten

$\dim W = n-1$. Seuraavaksi tekisi mieli soveltaa induktio-oletusta vektoriavaruuteen W. Kuitenkin täytyy muistaa, että mikään ei takaa sitä, että aliavaruus W olisi invariantti operaatorin L suhteen, joten rajoittuma $L|W$ ei välttämättä ole avaruuden W hyvinmääritelty operaattori eikä siihen tästä syystä voi välttämättä soveltaa induktio-oletusta. Täytyy siis ensin keksiä sopiva aliavaruuden W operaattori. Olkoon $p\colon V \to W$ hajotelmaan $U \oplus W$ liittyvä kanoninen projektio. Palautetaan mieleen, miten p on määritelty. Olkoon $\mathbf{v} \in V$. Koska summa $U + W$ on suora, ja $\mathbf{v} \in V = U \oplus W$, on olemassa *yksikäsitteiset* $\mathbf{u} \in U$ ja $\mathbf{w} \in W$ siten, että $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Tällöin asetetaan $p(\mathbf{v}) = \mathbf{w}$. Erityisesti kaikilla $\mathbf{v} \in V$ on voimassa yhtälö

$$\mathbf{v} = \mathbf{u} + p(\mathbf{v}), \tag{3.23}$$

missä $\mathbf{u}$ on jokin aliavaruuden U vektori.

Muodostetaan yhdistetty kuvaus $L' = p \circ L|W\colon W \to W$, $L'(\mathbf{w}) = p(L(\mathbf{w}))$ kaikilla $\mathbf{w} \in W$. Tämä kuvaus on $(n-1)$-ulotteisen K-vektoriavaruuden W lineaarinen operaattori, joten siihen voidaan soveltaa induktio-oletusta. Sen mukaan operaattori L' voidaan esittää jossakin avaruuden W kannassa yläkolmiomatriisina. Lemman 3.19 nojalla tällöin on olemassa nouseva ketju avaruuden W aliavaruuksia

$$W_1' \subset W_2' \subset \ldots \subset W_{n-2}' \subset W_{n-1}' = W,$$

siten, että jokaisella $i = 1, \ldots, n-1$ pätee $\dim W_i' = i$ ja aliavaruus W_i' on invariantti operaattorin L' suhteen.

Olkoon $i = 1, \ldots, n-1$. Merkitään $W_{i+1} = W_i' + U$. Näytetään, että aliavaruus $W_{i+1} = W_i' + U$ on invariantti operaattorin L suhteen. Olkoon

$$\mathbf{v} = \mathbf{w} + \mathbf{z} \in W_{i+1},$$

missä $\mathbf{w} \in W_i'$, $\mathbf{z} \in U$. Kaavasta 3.23 seuraa, että

$$L(\mathbf{w}) = p(L(\mathbf{w})) + \mathbf{z}'$$

jollakin $\mathbf{z}' \in U$. Toisaalta W_i' on invariantti operaattorin $p \circ L$ suhteen, joten tässä $p(L(\mathbf{w})) \in W_i'$. Näin ollen aliavaruuden W_{i+1} vektorille $\mathbf{v}$ pätee

$$L(\mathbf{v}) = L(\mathbf{w}) + L(\mathbf{z}) = p(L(\mathbf{w})) + (L(\mathbf{z}) + \mathbf{z}').$$

Tässä $p(L(\mathbf{w})) \in W_i'$ edellisen nojalla ja $L(\mathbf{z}) + \mathbf{z}' \in U$, koska U on invariantti L:n suhteen. Näin ollen $L(\mathbf{v}) \in W_i' + U = W_{i+1}$. On näytetty, että aliavaruus W_{i+1} on invariantti kuvauksen L suhteen.

Koska summa $W + U$ on suora, seurauksen 2.156 nojalla pätee $W \cap U = \{\mathbf{0}_V\}$. Koska $W_i' \subset W$, pätee myös $W_i' \cap U = \{\mathbf{0}_V\}$. Seurauksen 2.156 nojalla tästä seuraa, että myös summa $W_i' + U$ on suora. Proposition 2.160 nojalla tästä seuraa, että

$$\dim W_{i+1} = \dim W_i' + \dim U = i + 1.$$

Merkitään vielä $W_1 = U$. Nähdään, että

$$U = W_1 \subset W_2 \subset \ldots \subset W_{n-1}$$

on nouseva ketju avaruuden V aliavaruuksia siten, että kaikilla $i = 1, \ldots, n-1$ pätee $\dim W_i = i$. Lisäksi W_i on L-invariantti. Lemmasta 3.19 seuraa nyt, että L voidaan esittää yläkolmiomatriisina. □

On olemassa toinen hauska tapa osoittaa edellisen lemman väite todeksi käytämällä hyväksi duaali-avaruutta V^*. Siihen tutustutaan kurssin harjoitusten yhteydessä.

Seuraavaksi esitetään kaksi edellisen tuloksen sovellusta, joissa mennään analyysin ja topologian puolelle. Siksi esitys ei tule olemaan täysin tarkka. Tarkoitus on näyttää "käsiä heiluttamalla", miten edellä esitettyjä tuloksia voidaan soveltaa käytännössä erityyppisissä tilanteissa.

Esimerkki 3.24. *Olkoon A $\mathbb{R}$-kertoiminen $(n \times n)$-matriisi. Halutaan määritellä sen* eksponenttimatriisi *e^A. Reaaliluvuillehan eksponenttifunktio voidaan tunnetusti määritellä äärettömänä sarjana*

$$e^x = 1 + x + \frac{1}{2!}x^2 + \ldots + \frac{1}{m!}x^m + \ldots.$$

Analyysin peruskursseilla osoitetaan, että tämä sarja suppenee kaikilla $x \in \mathbb{R}$. Neliömatriiseille "matkitaan" samaa määritelmä, eli asetetaan

$$e^A = I + A + \frac{1}{2!}A^2 + \ldots + \frac{1}{m!}A^m + \ldots.$$

Voidaan osoittaa, että tällä tavalla määritelty sarja suppenee jokaisella $A \in M(n \times n; \mathbb{R})$ (matriisien kohdalla suppeneminen tarkoittaa suppenemista koordinaateittain). Näin määritelty matriisin A eksponentti e^A ei ole pelkkä hyödytön analogia, vaan se osoittautuu erittäin tärkeäksi työkaluksi esimerkiksi differentiaaligeometrissa, differentiaaliyhtälöiden teoriassa, Lie-ryhmien teoriassa sekä teoreettisessa fysiikassa.

Osoitetaan, että matriisi e^A on aina kääntyvä eli $\det e^A \neq 0$ (vrt. reaaliluvuille $e^x \neq 0$ kaikilla $x \in \mathbb{R}$). Yläkolmiomatriisille A tai yleisemmin sellaiselle matriisille, joka on similaarinen jonkun yläkolmiomatriisin kanssa, tämä on suhteellisen helppoa. Nimittäin olkoon B jokin yläkolmiomatriisi ja olkoot $b_{11}, \ldots, b_{nn}$ sen päälävistäjäalkiot. Tällöin B^2 on myös yläkolmiomatriisi, jonka diagonaalialkiot ovat $b_{11}^2, \ldots, b_{nn}^2$ ja yleisesti jokaisella $m \in \mathbb{N}$ matriisi B^m on yläkolmiomatriisi, jonka diagonaalialkiot ovat $b_{11}^m, \ldots, b_{nn}^m$. Tästä seuraa, että sarjassa

$$e^B = I + B + \frac{1}{2!}B^2 + \ldots + \frac{1}{m!}B^m + \ldots$$

raja-arvoksi saadaan yläkolmiomatriisi, jonka diagonaalialkiot ovat reaalilukusarjan

$$1 + z + \frac{1}{2!}z^2 + \ldots + \frac{1}{m!}z^m + \ldots = e^z$$

raja-arvoja, missä $z = b_{ii}$, $i = 1, \ldots, n$. Tästä seuraa, että

$$\det e^B = e^{b_{11}} e^{b_{22}} \ldots e^{b_{nn}} \neq 0,$$

sillä $e^x \neq 0$ kun x on reaaliluku. Yleisemmin, olkoon A sellainen matriisi, jota vastaava kuvaus L_A voidaan esittää yläkolmiomuodossa. Tällöin $A = JBJ^{-1}$, missä J on sopiva

kannanvaihtomatriisi ja B on yläkolmiomatriisi. Induktiolla nähdään helposti, että tällöin jokaisella $m \in \mathbb{N}$ pätee $A^m = JB^mJ^{-1}$, joten

$$e^A = I + A + \frac{1}{2!}A^2 + \ldots + \frac{1}{m!}A^m + \ldots = I + JBJ^{-1} + \frac{1}{2!}JB^2J^{-1} + \ldots + \frac{1}{m!}JB^mJ^{-1} + \ldots = Je^BJ^{-1},$$

missä matriisit J ja J^{-1} otetaan "yhteiseksi tekijäksi" viimeisessä välivaiheessa (tällaisten manipulaatioiden "laillisuuden" tarkka perustelu vaatii analyysin tuloksia, jotka sivutetaan tässä). Näin ollen tässä tapauksessa saadaan

$$\det e^A = \det J \det e^B (\det J)^{-1} = \det e^B \neq 0.$$

Näin ollen, jos jokainen $\mathbb{R}$-kertoiminen matriisi olisi similaarinen yläkolmiomatriisin kanssa, väite olisi osoitettu edellisen kappaleen argumentin nojalla. Ongelma on siinä, että $\mathbb{R}$ ei ole algebrallisesti suljettu, eikä siis voida päätellä edellisen proposition avulla, että jokainen $\mathbb{R}$-kertoiminen matriisi on similaarinen yläkolmiomatriisin kanssa. Itse asiassa tämä ei edes pidä paikkaansa, sillä on olemassa $\mathbb{R}$-kertoimisia matriiseja, joita ei voi muuttaa yläkolmiomuotoon - tästä on nähty esimerkkejä.

Kuitenkin $\mathbb{R}$ voidaan upottaa algebrallisesti sujettuun kuntaan, nimittäin kompleksilukujen kuntaan $\mathbb{C}$. Jokainen $\mathbb{R}$-kertoiminen matriisi voidaan ajatella $\mathbb{C}$-kertoimiseksi. Tästä seuraa, että periaatteessa voi aina ainakin yrittää ratkaista ongelmia, jotka koskevat "epätäydellistä" kuntaa $\mathbb{R}$, muuttamalla niitä yleisemmiksi ongelmiksi, jotka koskevat "algebrallisesti parempaa" kuntaa $\mathbb{C}$. Tällöin voidaan käyttää kunnan $\mathbb{C}$ ominaisuuksia, joita kunnalla $\mathbb{R}$ ei ole. Tätä varten määritellään eksponentin käsite myös kompleksikeroimisille neliömatriiseille $A \in M(n \times n; \mathbb{C})$, samalla periaatteella, eli kaavalla

$$e^A = I + A + \frac{1}{2!}A^2 + \ldots + \frac{1}{m!}A^m + \ldots.$$

Samoilla analyyttisillä menetelmillä, joita käytetään $\mathbb{R}$:n tapauksessa, voidaan osoittaa, että tämä sarja suppenee kaikilla kompleksikertoimisilla neliömatriiseilla A.

Edellisen proposition nojalla jokainen $A \in M(n \times n; \mathbb{C})$ voidaan esittää yläkolmiomatriisina B, toisin sanoen muodossa $A = JBJ^{-1}$, missä J on kannanvaihtomatriisi. Tällöin samalla tavalla kuin yllä on näytetty $\mathbb{R}$:n tapauksessa, voidaan osoittaa suoraan, että $\det e^A = \det e^B \neq 0$.

Esimerkki 3.25. *Olkoon A kompleksiarvoinen neliömatriisi. Yleisesti ottaen A ei välttämättä ole diagonalisoituva. Näytetään, että kuitenkin* mielivaltaisen läheltä *A:ta aina löytyy diagonalisoituva matriisi. Tämä tulos on tärkeä analyysissa. Sen avulla voidaan esimerkiksi todistaa matriiseihin liittyviä väitteittä todistamalla ne ensin diagonalisoituville matriiseille (mikä usein voi olla hyvin helppoa) ja sitten yleistää ne mielivaltaisille matriiseille jonkinlaisella "rajankäynnillä".*

Mitä "mielivaltaisen lähellä" tarkoittaa tässä yhteydessä? Se tarkoittaa, että jokaisella $\varepsilon > 0$ muuttamalla matriisin A alkioita korkeintaan luvun ε verran saadaan diagonalisoituvan matriisin. Täsmällisemmin sanottuna väitetään, että jos $A = (a_{ij})$ on neliömatriisi ja $\varepsilon > 0$, niin löytyy (samankokoinen) diagonalisoituva matriisi $B = (b_{ij})$ siten, että kaikilla i, j pätee $|a_{ij} - b_{ij}| < \varepsilon$. Tässä kompleksiluvun itseisarvo $|\cdot|$ on sama asia kuin sen itseisarvo tason $\mathbb{R}^2$ pisteenä (eli sen etäisyys origosta).

Esitetään A muodossa JCJ^{-1}, missä C on yläkolmiomatriisi (tämä on mahdollista, koska $\mathbb{C}$ on algebrallisesti suljettu). Jos löydetään mielivaltaisen lähellä matriisia C diagonalisoituva matriisi B, niin myös JBJ^{-1} on diagonalisoituva matriisi, joka on mielivaltaisen lähellä matriisia A (tässä välivaiheessa tarvitaan matriisien kertolaskun jatkuvuutta *eli analyyttisia käsitteitä ja tuloksia). Näin ollen riittää osoittaa väite yläkolmiomatriiseille. Olkoon A yläkolmiomatriisi. Tällöin matriisin A ominaisarvot ovat sen päälävistäjäalkioita. Jos kaikki nämä päälävistäjäalkiot ovat eri lukuja, Lemmasta 3.13 seuraa, että A on itse asiassa diagonalisoituva ja voidaan valita $A = B$. Yleisesti ottaen matriisin A päälävistäjäalkiot eivät välttämättä ole eri alkioita. Tällöin kuitenkin voidaan helposti muuttaa niiden arvoja "mielivaltaisen vähän" eli korkeintaan epsilonin verran, niin, että päälävistäjäluvuista saadaan eri lukuja. Tällä tavalla saadaan uusi matriisi B, joka on myös yläkolmiomatriisi, on epsilon-lähellä A:ta ja jonka diagonaalialkiot ovat kaikki eri alkioita. Tällainen matriisi on diagonalisoituva (Lemma 3.13).*

3.2. Polynomialgebra

Edellisessä aliluvussa ollaan törmätty tilanteisiin, joissa polynomifunktiot sekä polynomiyhtälöt kunnan yli tulevat lineaarialgebrassa esille luonnollisella tavalla, nimittäin tutkimalla operaattorien ominaisarvoja. Polynomien teorialla on muitakin tärkeitä yhtymäkohtia lineaarialgebran kanssa. Esimerkiksi jos $L\colon V \to V$ on lineaarinen operaattori, voidaan muodostaa sen potensseja L^m kaikilla $m \in \mathbb{N}$. Nämäkin ovat lineaarisia operaattoreita, joten ne voidaan kertoa skalaareilla ja laskea yhteen. Tästä seuraa, että voidaan muodostaa lausekkeita muotoa

$$c_n L^n + c_{n-1} L^{n-1} + \ldots + c_1 L + c_0 \operatorname{id}_V,$$

missä $\operatorname{id}_V = L^0$ on avaruuden V identtinen kuvaus ja $c_0, \ldots, c_n \in K$ ovat skalaareja. Tällainen lauseke voidaan tulkita polynomin $p(x) = c_n x^n + c_{n-1} x^{n-1} + \ldots + c_1 x + c_0$ *arvona* kun muuttujan x paikalle sijoitetaan operaattori L.

Koska polynomit muodostavat tärkeän työkalun lineaarialgebrassa, tässä aliluvussa tutkitaan tarkemmin niiden algebrallista teoriaa. Seuraavassa luvussa opittuja uusia tuloksia taas sovelletaan lineaarisiin operaattoreihin liittyvissä ongelmissa.

Olkoon K kunta. Palautetaan mieleen, että kuvaus $p\colon K \to K$ on *polynomifunktio*, jos jollakin luonnollisella luvulla $n \in \mathbb{N}$ on olemassa kunnan alkiot $a_0, a_1, \ldots, a_n \in K$ siten, että kaikilla $x \in X$ pätee

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0.$$

Tämä voidaan kirjoittaa myös muodossa $p(x) = \sum_{i=0}^{n} a_i x^i$, sillä $x^0 = 1_K$ ja $x^1 = x$ kaikilla $x \in K$.

Polynomifunktiot $p\colon K \to K$ muodostavat *K-algebran* luonnollisella tavalla. Nimittäin olkoot $p, q\colon K \to K$ polynomifunktioita ja olkoot $a_0, \ldots, a_n \in K$ ja $b_0, \ldots, b_m \in K$ sellaisia, että kaikilla $x \in K$ pätee

$$p(x) = \sum_{i=0}^{n} a_i x^i \text{ ja } q(x) = \sum_{i=0}^{m} b_i x^i.$$

Lisäämällä tarvittaessa nolla-termejä voidaan olettaa, että $n = m$. Nimittäin, jos esimerkiksi $n > m$, määritellään $b_i = 0_K$ kaikilla $i = m+1, \ldots, n$, jolloin voidaan kirjoittaa polynomi q muodossa $q(x) = \sum_{i=0}^{n} b_i x^i$. Tällöin kaikilla $x \in K$ pätee

$$(p+q)(x) = \sum_{i=0}^{n} (a_i + b_i) x^i, \tag{3.26}$$

joten $p+q$ on myös polynomifunktio. Lisäksi, jos $k \in K$ on skalaari, kuvaus kp on polynomifunktio, sillä kaikilla $x \in K$ pätee tällöin

$$(kp)(x) = \sum_{i=0}^{n} (k a_i) x^i. \tag{3.27}$$

Näin ollen polynomikuvausten muodostama joukko on K-vektoriavaruus luonnollisella tavalla. Tämän lisäksi polynomikuvauksia voidaan kertoa keskenään (pisteittäin) ja tuloksena saadaan edelleenkin polynomifunktio. Nimittäin olkoot $p, q \colon K \to K$ polynomifunktioita kuten yllä. Tällöin osittelulain ja kunnan kertolaskun vaihdannaisuuden nojalla kaikilla $x \in K$ pätee

$$(pq)(x) = \sum_{i \in [n], j \in [m]} a_i b_j x^i x^j.$$

Potenssisääntöjen mukaanpätee $x^i x^j = x^{i+j}$. Keräämällä yhteen kaikki termit muotoa $a_i b_j x^i x^j = a_i b_j x^{i+j}$, joissa indeksien i, j summa $i+j$ pysyy vakiona, päästään (osittelulain avulla) kirjoittamaan tulo pq muodossa

$$(pq)(x) = \sum_{l=0}^{n+m} c_l x^l, \tag{3.28}$$

missä

$$c_l = \sum_{i+j=l} a_i b_j. \tag{3.29}$$

Polynomikuvausten muodostama joukko $P(K)$ on K-algebra näiden laskutoimitusten suhteen. Tämän näkee helpoiten huomaamalla, että tämä joukko on K-algebran K^K alialgebra (kts. esim. 2.74, (5)).

Tähän asti polynomeja on käsitelty perinteisellä tavalla **funktioina** $K \to K$, joilla on tietty muoto. Nykyalgebrassa kuitenkin suositaan sen sijaan abstraktimpaa ja algebrallisempaa tapaa ajatella polynomeja. Tämä johtuu siitä, että polynomifunktio $p \colon K \to K$ halutaan olevan *yksikäsitteisesti määrätty* sen *kertoimien* $a_0, \ldots, a_n$ muodostamalla jonolla. Toisin sanoen halutaan, että polynomifunktiota ei olisi mahdollista kirjoittaa muodossa

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, x \in K,$$

$a_0, \ldots, a_n \in K, a_n \neq 0$, kahdella eri tavalla. Tällöin polynomi p voidaan samaistaa sen kertoimien muodostaman jonon $(a_0, \ldots, a_n) \in K^{n+1}$ kanssa. Valitettavasti tämä ei aina pidä paikkansa ja on olemassa kuntia, joissa erinäköiset polynomilausekkeet vastaavat

samaa kuvausta. Esimerkiksi kunnassa $K = \mathbb{Z}_2$ polynomit $p(x) = x$ ja $q(x) = x^2$ ovatkin sama kuvaus, mikä nähdään yksinkertaisesti laskemalla sen arvoja kunnan alkioilla,

$$p(0_2) = 0_2 = 0_2^2 = q(0_2),$$

$$p(1_2) = 1_2 = 1_2^2 = q(1_2).$$

Itse asiassa olkoon $K = \{x_0, \ldots, x_n\}$ mikä tahansa *äärellinen* kunta. Tällöin polynomifunktio $p\colon K \to K$,

$$p(x) = (x - x_0)(x - x_1)\ldots(x - x_n)$$

on polynomilausekkeena $(n+1)$-asteinen polynomi, mutta kuvauksena $K \to K$ se on yksinkertaisesti nollakuvaus. Tässä tapauksessa polynomikuvauksen aste ei edes ole hyvinmääritelty - se riippuu siitä, miten kuvaus esitetään polynomilausekkeena. Itse asiassa, kun kunta K on äärellinen, on selvä, että erilaisia kuvauksia $f\colon K \to K$ on vain äärellinen määrä, mutta erinäköisiä *polynomilausekkeita* on pakko olla ääretön määrä - ainakin yksi jokaisella $n \in \mathbb{N}$ (kuvaus $x \mapsto x^n$). Voidaan osoittaa (harjoitustehtävä), että äärettömän kunnan tapauksessa näin ei voi käydä - jos K on ääretön kunta, niin kaksi polynomikuvausta $p, q\colon K \to K$ ovat kuvauksina samat jos ja vain jos niillä on täsmälleen samat kertoimet.

Polynomikuvaukset, jonka kertoimet ovat kunnan K alkioita, yleistyvät luonnollisella ja ja tuottoisalla tavalla *K-algebroihin*. Tämän näkökulman kautta päädytään luonnollisella tavalla algebralliseen abstraktiin polynomin käsitteeseen.

Polynomialgebran konstruuktio.
Ajatuksena on, että polynomikuvaus $p\colon K \to K$,

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, x \in K,$$

korvataan sen kertoimien muodostamalla jonolla $(a_0, \ldots, a_n) \in K^{n+1}$. Tässä luonnollinen luku n vaihtelee polynomin asteen mukaan, joten tällaisen jonon pituus ei ole fiksattu. Jotta algebrallisilla polynomeille saattaisiin yhtenäinen, asteesta riipumaton määritelmä, menetelläänkin formaalilla tasolla seuraavasti.

Olkoon K kunta. Esimerkin 2.4, (2) mukaan on olemassa jonojen $(a_i)_{i\in\mathbb{N}}$ muodostama K-vektoriavaruus $K^{\mathbb{N}}$. Tässä vektoriavaruudessa vektorit lasketaan yhteen ja kerrotaan skalaarilla *koordinaateittain*,

$$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} = (a_i + b_i)_{i\in\mathbb{N}},$$

$$k(a_i)_{i\in\mathbb{N}} = (ka_i)_{i\in\mathbb{N}}.$$

Ajatuksena on, että jono $(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}}$ vastaa abstraktia polynomia, jonka kertoimina ovat luvut a_i. Mielivaltainen jono $(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}}$ ei tällöin kuitenkin välttämättä kelpaa "polynomiksi", sillä polynomilla pitäisi olla äärellinen määrä (nollasta poikkeavia) kertoimia. Tästä syystä esitetään seuraava määritelmä. Jonoa $(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}}$ sanotaan *äärelliskantajaiseksi*, jos on olemassa $n \in \mathbb{N}$ siten, että $a_m = 0_K$ kaikilla $m > n$. Merkitään äärelliskantajaisten jonojen muodostamaa osajoukkoa $K^{(\mathbb{N})}$ (tähän joukkoon ollaan jo törmätty aikaisemmin Esimerkissä 2.47). Seuraavan lemman nojalla $K^{(\mathbb{N})}$ on K-vektoriavaruus.

Lemma 3.30. *Äärelliskantajaisten jonojen osajoukko $K^{(\mathbb{N})}$ on K-vektoriavaruuden $K^{\mathbb{N}}$ aliavaruus.*

Todistus. Tämä seuraa esimerkin 2.47, (5) tuloksesta, mutta osoitetaan tämä väite uudestaan määritelmästä lähtien.

Olkoot $(a_i)_{i\in\mathbb{N}}$ ja $(a'_i)_{i\in\mathbb{N}}$ molemmat äärelliskantajaisia jonoja. Olkoot $n, n' \in \mathbb{N}$ sellaisia, että $a_m = 0_K = a_l$ kaikilla $m > n$, $l > n'$. Tällöin kaikilla $p > \max\{n, n'\}$ pätee $a_p + a'_p = 0_K$. Näin ollen jonojen $(a_i)_{i\in\mathbb{N}}$ ja $(a'_i)_{i\in\mathbb{N}}$ summa eli jono $(a_i + a'_i)_{i\in\mathbb{N}}$ on myös äärelliskantajainen. Näin ollen $K^{(\mathbb{N})}$ on suljettu vektorien yhteenlaskun suhteen.

Samalla tavalla nähdään, että $K^{(\mathbb{N})}$ on suljettu skalaarikertolaskun suhteen. Lisäksi $K^{(\mathbb{N})}$ ei ole tyhjä, sillä esimerkiksi avaruuden $K^{\mathbb{N}}$ nolla-vektori eli jono $(a_i)_{i\in\mathbb{N}}$, jolle $a_i = 0_K$ kaikilla $i \in \mathbb{N}$, on äärelliskantajainen. □

Olkoon $j \in \mathbb{N}$ luonnollinen luku (huom., muista, että tällä kursilla nolla lasketaan luonnolliseksi luvuksi!). Merkitään symbolilla $\mathbf{e}_j$ avaruuden $K^{\mathbb{N}}$ alkiota $(a_i)_{i\in\mathbb{N}}$, jolle pätee $a_j = 1_K$ ja $a_i = 0_K$ kaikilla $i \neq j$. Jono $\mathbf{e}_j$ on tällöin selvästi äärelliskantajainen, eli on avaruuden $K^{(\mathbb{N})}$ alkio kaikilla $j \in \mathbb{N}$.

Lemma 3.31. *Kokoelma $(\mathbf{e}_j)$, $j \in \mathbb{N}$, on vektoriavaruuden $K^{(\mathbb{N})}$ (ääretön)* kanta.

Todistus. Tämä on osoitettu aikaisemmin esimerkissä 2.47, (5). □

Huomaa, että edellisen lemman mukaan äärelliskantajaisten jonojen muodostama vektoriavaruus $K^{(\mathbb{N})}$ on esimerkki "ääretönulotteisesta" vektoriavaruudesta, jolla on kanta, mutta tämä kanta ei ole äärellinen.

Olkoon $\mathbf{v} = (a_i)_{i\in\mathbb{N}} \in K^{(\mathbb{N})}$ äärelliskantajainen jono. Tällöin on olemassa $n \in \mathbb{N}$ siten, että $a_i = 0_K$ kaikilla $i > n$. Jos $\mathbf{v}$ ei ole nolla-jono, voidaan valita pienin tällainen n, jolloin $a_n \neq 0$ ja $\mathbf{v}$ voidaan kirjoittaa kannassa $\{\mathbf{e}_j \mid j \in \mathbb{N}\}$ lineaarisena kombinaationa

$$a_0\mathbf{e}_0 + a_1\mathbf{e}_1 + \ldots + a_n\mathbf{e}_n$$

missä $a_0, \ldots, a_n \in K$ ja $a_n \neq 0_K$. Tällainen esitys on lisäksi yksikäsitteinen (koska $\{\mathbf{e}_j \mid j \in \mathbb{N}\}$ on kanta). Luonnollista lukua $n \in \mathbb{N}$ sanotaan tällöin äärelliskantajaisen jonon $\mathbf{v}$ *asteeksi* ja merkitään $\deg \mathbf{v}$. Toisin sanoen äärelliskantajaisen jonon $\mathbf{v} = (a_n)_{n\in\mathbb{N}}$ aste on suurin $n \in \mathbb{N}$ jolle pätee $a_n \neq 0_K$. Tätä komponenttia a_n sanotaan jonon *johtavaksi kertoimeksi*.

Nolla-vektorilla $\mathbf{0} = (0_K)_{i\in\mathbb{N}}$ ei ole johtavaa kerrointa eikä sen astetta voi määritellä samalla tavalla kuin kaikkien muiden äärelliskantajaisten jonojen kohdalla. Teknisistä syistä sovitaan, että pelkistä nollista koostuvan äärelliskantajaisen perheen aste on $-\infty$ (eli miinus ääretön).

K-vektoriavaruuden $K^{(\mathbb{N})}$ laskutoimitukset ovat määriteltyjä "koordinaateittain", eli kaikilla $\mathbf{v} = (a_i)_{i\in\mathbb{N}}$, $\mathbf{w} = (b_i)_{i\in\mathbb{N}} \in K^{(\mathbb{N})}$, $k \in K$ pätee

$$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} = (a_i + b_i)_{i\in\mathbb{N}},$$

$$k(a_i)_{i\in\mathbb{N}} = (ka_i)_{i\in\mathbb{N}}.$$

Palataan hetkeksi avaruuden $K^{(\mathbb{N})}$ konstruktion intuitiiviseen motivaatioon eli ajatukseen siitä, että äärelliskantajainen perhe $(c_i)_{i\in\mathbb{N}}$, jonka aste on n, vastaa "polynomia"

$$c_nX^n + c_{n-1}X^{n-1} + \ldots + c_1X + c_0,$$

jonka kertoimet ovat $c_0, c_1, \ldots, c_n$. Tässä X tarkoittaa abstraktia "muuttujaa", jolle annetaan täsmällinen algebrallinen tulkinta myöhemmin. Huomataan, että tällä tulkinnalla äärelliskantajaisten jonojen yhteen- ja skalaarikertolaskutoimitukset vastaavat polynomien yhteen- ja skalaarikertolaskutoimituksia, vrt. yhtälöt 3.26 ja 3.27.

Polynomifunktioilla oli vielä yksi laskutoimitus, nimittäin (pisteittäinen) kertolasku. Yhtälöistä 3.28 ja 3.29 seuraa, että jos polynomin p kertoimet ovat (a_i) ja polynomin q kertoimet ovat (b_j), niiden tulo on polynomikuvaus, jonka kertoimet (c_l) määräytyvät yhtälöistä

$$c_l = \sum_{i+j=l} a_ib_j.$$

Lainaamalla tämä kaava, saadaan suoraan määritelmä algebrallisten polynomien kertolaskulle. Täsmällisesti sanottuna olkoot $\mathbf{v} = (a_i)_{i\in\mathbb{N}}$, $\mathbf{w} = (b_j)_{j\in\mathbb{N}} \in K^{(\mathbb{N})}$. Asetetaan jokaisella $l \in \mathbb{N}$

$$c_l = \sum_{i+j=l} a_ib_j. \tag{3.32}$$

Olkoot $n, m \in \mathbb{N}$ sellaisia, että $a_i = 0_K = b_j$ kun $i > n$ ja $j > m$. Olkoon $l > n + m$ ja olkoot $i, j \in \mathbb{N}$ sellaisia, että $i + j = l$. Tällöin joko $i > n$ tai $j > m$, mistä seuraa, että $a_ib_j = 0_K$. Näin ollen $c_l = 0_K$ kun $l > n + m$. Näin ollen kaavalla 3.32 määritelty jono $(c_l)_{l\in\mathbb{N}}$ on äärelliskantajainen. Asetetaan joukon $K^{(\mathbb{N})}$ alkioiden $\mathbf{v} = (a_i)_{i\in\mathbb{N}}$ ja $\mathbf{w} = (b_j)_{j\in\mathbb{N}} \in K^{(\mathbb{N})}$ *tuloksi* $\mathbf{v} \cdot \mathbf{w}$ jono $(c_l)_{l\in\mathbb{N}}$, jonka koordinaatit c_l määritellään kaavalla 3.32. Edellisen nojalla näin saadaan hyvin määritelty kertolaskuoperaatio joukossa $K^{(\mathbb{N})}$. Seuraavaksi voidaan osoittaa suoraan tästä määritelmästä, että tällä kertolaskuoperaatiolla varustettuna $K^{(\mathbb{N})}$ on K-algebra. Suoraan tästä määritelmästä tämä kuitenkin vaatisi raskaita ja pitkiä laskuja. Tästä syystä siirrytään abstraktimpaan lähestymistapaan ja konstruoidaan kertolasku uudestaan lineaarialgebrallisen teorian avulla.

Halutaan siis määritellä kertolaskuoperaatio $\cdot K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$, joka tekisi avaruudesta $K^{(\mathbb{N})}$ (sen K-vektoriavaruuden struktuurin kera) K-algebran. Esimerkissä 2.123 opittiin, että K-vektoriavaruudessa A määritelty laskutoimitus $\cdot\colon A \times A \to A$ tekee A:sta K-algebran jos ja vain jos $\cdot\colon A \times A \to A$ on liitännäinen operaatio, jolla on neutraalialkio 1_K ja joka on kuvauksena *bilineaarinen kuvaus A:n vektoriavaruuden struktuurin suhteen.*

Toisin sanoen on konstruoitava *bilineaarinen* kuvaus $\cdot\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$, joka olisi laskutoimituksena liitännäinen ja olla olisi neutraalialkio. Tähän voidaan käyttää hyväksi multilineaaristen kuvausten teoriaa.

Lemman 3.31 nojalla kokoelma $\{\mathbf{e}_j \mid j \in \mathbb{N}\}$ on vektoriavaruuden $K^{(\mathbb{N})}$ *kanta*. Hyödynnetään tätä tosiasiaa bilineaarisen kuvauksen $\cdot\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$ konstruonnissa. Nimittäin Proposition 2.122 nojalla bilineaarinen kuvaus $\cdot\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$ voidaan määritellä yksikäsitteisesti antamalla sen arvoja kannan alkioiden muodostamilla pareilla. Tämä propositio on tosin muotoiltu ja todistettu ainoastaan äärellisen kannan erikoistapauksessa, mutta se pätee myös yleisesti äärettömän kannan tapauksessa, samalla todistuksella.

Asetetaan kaikilla $n, m \in \mathbb{N}$

$$\mathbf{e}_n \cdot \mathbf{e}_m = \mathbf{e}_{n+m}.$$

Edellisen nojalla tämä kaava määrittelee yksikäsitteisesti erään bilineaarisen kuvauksen $\cdot\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$. Tämä kuvaus voidaan ajatella erääksi laskutoimitukseksi joukossa $K^{(\mathbb{N})}$. Tämä laskutoimitus sanotaan joukon $K^{(\mathbb{N})}$ *kertolaskuksi*.

Propositio 3.33. *Kertolaskulla varustettuna K-vektoriavaruus $K^{(\mathbb{N})}$ on K-algebra. Kertolaskun neutraalialkio on $\mathbf{e}_0$.*

Todistus. Kuvaus $\cdot$ on konstruktion perusteella bilineaarinen kuvaus. Tästä seuraavat osittelulait ja ehto

$$(k\mathbf{a})\mathbf{b} = \mathbf{a}(k\mathbf{b}) = k(\mathbf{ab}).$$

Osoitetaan, että $\mathbf{e}_0$ on kertolaskun neutraalialkio. Määritellään kuvaus $L\colon K^{(\mathbb{N})} \to K^{(\mathbb{N})}$ kaavalla $L(\mathbf{v}) = \mathbf{v} \cdot \mathbf{e}_0$. Koska $\cdot$ on bilineaarinen, L on lineaarinen kuvaus. Lisäksi kaikilla $n \in \mathbb{N}$ pätee määritelmän nojalla

$$L(\mathbf{e}_n) = \mathbf{e}_n \cdot \mathbf{e}_0 = \mathbf{e}_n = \mathrm{id}(\mathbf{e}_n).$$

Lineaariset kuvaukset L ja id siis yhtyvät kannan alkioilla. Koska $\{\mathbf{e}_n \mid n \in \mathbb{N}\}$ on kanta, Propositiosta 2.57 (tai oikeastaan sen yleistyksestä äärettömille kannoille) seuraa, että $L = \mathrm{id}$ on identtinen kuvaus. Toisin sanoen kaikilla $\mathbf{v} \in K^{(\mathbb{N})}$ pätee

$$L(\mathbf{v}) = \mathbf{v} \cdot \mathbf{e}_0 = \mathbf{v} = \mathrm{id}(\mathbf{v}).$$

Yhtälö $\mathbf{e}_0 \cdot \mathbf{v} = \mathbf{v}$ voidaan osoittaa samalla tavalla (tai johtaa kertolaskun vaihdannaisuudesta, joka osoitetaan alla). Näin ollen $\mathbf{e}_0$ on kertolaskun neutraalialkio.

Kertolaskun liitännäisyys ja vaihdannaisuus osoitetaan samalla periaatteella. Tarkastellaan kuvauksia $\Psi_1\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$, $\Psi_2\colon K^{(\mathbb{N})} \times K^{(\mathbb{N})} \times K^{(\mathbb{N})} \to K^{(\mathbb{N})}$,

$$\Psi_1(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{ab})\mathbf{c}, \Psi_2(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a}(\mathbf{bc}).$$

Kertolaskun liitännäisyys tarkoittaa täsmälleen sitä, että nämä kuvaukset ovat samoja, $\Psi_1 = \Psi_2$. Koska kertolasku $\cdot$ on 2-lineaarinen, nähdään helposti (tarkista!), että kumpikin kuvaus Ψ_1, Ψ_2 on 3-lineaarinen. Propositiosta 2.122 (tarkemmin sen yleistyksestä ääarettömille kannoille) seuraa, että nämä kuvaukset ovat samat jos ja vain jos niiden arvot yhtyvät kanta-alkioilla muodestuista kolmikoissa. Koska kaikilla $n, m, p \in \mathbb{N}$ pätee

$$\Psi_1(\mathbf{e}_n, \mathbf{e}_m, \mathbf{e}_p) = (\mathbf{e}_n\mathbf{e}_m)\mathbf{e}_p = \mathbf{e}_{n+m}\mathbf{e}_p =$$

$$\mathbf{e}_{(n+m)+p} = \mathbf{e}_{n+(m+p)} = \mathbf{e}_n\mathbf{e}_{m+p} = \mathbf{e}_n(\mathbf{e}_m\mathbf{e}_p) = \Psi_2(\mathbf{e}_n, \mathbf{e}_m, \mathbf{e}_p),$$

tämä todistaa liitännäisyyden. Vaihdannaisuus osoitetaan samalla tavalla - riittää huomata, että se pätee kanta-alkioille, eli että

$$\mathbf{e}_n \cdot \mathbf{e}_m = \mathbf{e}_{n+m} = \mathbf{e}_{m+n} = \mathbf{e}_m \cdot \mathbf{e}_n.$$

□

On osoitettu, että on olemassa K-algebra $K^{(\mathbb{N})}$, jonka muodostavat äärelliskantajaiset jonot $(a_i)_{i\in\mathbb{N}}$. Koska vektori $\mathbf{e}_0$ on edellisen Proposition mukaan tämän algebran kertolaskun neutraalialkio, merkitään sitä tavalliseen tapaan symbolilla $\mathbf{1}$. Alkiota $\mathbf{e}_1$ taas merkitään "muuttujasymbolilla" $\mathbf{X}$. Induktiolla nähdään helposti, että kaikilla $n \in \mathbb{N}$ pätee

$$\mathbf{e}_n = \underbrace{\mathbf{e}_1 \cdot \ldots \cdot \mathbf{e}_1}_{n \text{ kertaa}} = \mathbf{e}_1^n = \mathbf{X}^n.$$

Tästä seuraa, että jokainen (nolla-jonosta eroava) äärelliskantajainen jono $(a_i)_{i\in\mathbb{N}}$ voidaan kirjoittaa muodossa

$$a_n\mathbf{X}^n + a_{n-1}\mathbf{X}^{n-1} + \ldots + a_1\mathbf{X} + a_0,$$

missä $a_0, \ldots, a_n \in K$, n on tämän jonon aste ja $a_n \neq 0_K$. Lisäksi tällainen esitys on yksikäsitteinen. Tässä viimeinen jäsen $a_0 \cdot 1$ kirjoitetaan muodossa a_0 yksinkertaisuuden vuoksi. Ollaan siis päästy tulkintaan, jossa äärelliskantajainen jono näyttää samalta kuin vanha tuttu "polynomi". Tällä kertaa se ei kuitenkaan ole enää mikään kuvaus ja $\mathbf{X}$ ei ole enää mikään muuttuja, vaan abstrakti algebrallinen olio.

Tästä syystä algebran $K^{(\mathbb{N})}$ alkioita kutsutaan jatkossa *K-kertoimisiksi (algebrallisiksi) polynomeiksi* ja merkitään $K^{(\mathbb{N})} = K[\mathbf{X}]$. Algebraa $K[\mathbf{X}]$ sanotaan kunnan K (algebralliseksi) *polynomialgebraksi*. Jokainen algebrallinen polynomi voidaan kirjoittaa muodossa $\mathbf{v} = \sum a_i\mathbf{X}^i$, missä kontekstista ymmärretään, että summa on äärellinen ja $a_i = 0_K$, kun $i > \deg \mathbf{v}$. Käytännössä algebrallisilla polynomeilla voidaan laskea kuin "tavallisilla" polynomikuvauksilla,

$$\sum a_i\mathbf{X}^i + \sum b_i\mathbf{X}^i = \sum (a_i + b_i)\mathbf{X}^i,$$

$$k(\sum a_i\mathbf{X}^i) = \sum (ka_i)\mathbf{X}^i,$$

$$(\sum a_i\mathbf{X}^i)(\sum b_i\mathbf{X}^i) = \sum (\sum_{p+q=i} a_p b_q)\mathbf{X}^i.$$

Kunta K voidaan "upottaa" sen polynomialgebran $K[X]$ alialgebraksi luonnollisella tavalla. Nimittäin sovitaan samaistamaan kunnan K alkio k ja vastaava algebrallinen polynomi $k\mathbf{e}_0 = k \cdot 1$. Tämä samaistus on järkevä algebrallisesta näkökulmasta, sillä vastaavuus $\phi\colon K \to K[\mathbf{X}]$, $k \mapsto k\mathbf{e}_0$, on *bijektiivinen* kuvaus, joka *säilyttää* K-algebrojen struktuurit, eli on algebraisomorfismi algebran K ja kuva-joukon $\phi(K)$ välillä. Tämän kuva-joukon alkioita (eli 0-asteisia polynomeja ja nollapolynomia) sanotaan *vakiopolynomeiksi*. Edellä tehdyn samastussopimuksen valossa vakiopolynomit $k\mathbf{e}_0$ merkitään jatkossa yksinkertaisesti k ja kohdellaan kuten kunnan K alkioita. Tätä sopimusta on itse asiassa käytetty jo implisiittisesti yllä.

Polynomialgebran rengas-struktuuri
Koska polynomialgebra $K[\mathbf{X}]$ on algebra, se on erityisesti rengas polynomien yhteen- ja kertolaskun suhteen. Koska polynomien kertolasku on vaihdannainen, tämä rengas on jopa *vaihdannainen*. Näin ollen, polynomialgebran tutkimisessa voidaan soveltaa vaihdannaisten renkkaiden teoriaa.

Vaihdannaisen renkaan R alkiota $r \neq 0_R$ sanotaan renkaan R *nollanjakajaksi*, jos on olemassa sellainen $y \in R, y \neq 0$, siten, että $xy = 0$. Esimerkiksi äärellisssä renkaassa $\mathbb{Z}_6$ nollasta eroava alkio 2_6 on nollan jakaja, sillä $2_6 \cdot 3_6 = 6_6 = 0_6$, vaikka $2_6 \neq 0_6 \neq 3_6$.

Vaihdannaista rengasta R sanotaan *kokonaisalueeksi*, jos siinä ei ole nollan jakajia, eli jos kaikilla $x, y \in R, x, y \neq 0_R$, pätee $xy \neq 0_R$. Kokonaisalue on siis sellainen vaihdannainen rengas, jossa pätee koulusta tuttu "nollasääntö". Esimerkiksi jokainen kunta on kokonaisalue. Kokonaislukujen rengas $\mathbb{Z}$ on kanoninen esimerkki kokonaisalueesta, joka ei ole kunta (tästä esimerkistä termi "kokonaisalue" itse asiassa juontuukin). Voidaan osoittaa, että rengas $\mathbb{Z}_n$ on kokonaisalue täsmälleen silloin kun se on kunta, eli silloin kun n on alkuluku.

Kokonaisalueessa R on voimassa tärkeä *supistussääntö*. Tämä sanoo seuraavan. Olkoon R kokonaisalue ja olkoot $a, b, c \in R$. Oletetaan, että $ab = ac$ ja $a \neq 0_R$. Tällöin $b = c$. Supistussäännön todistus ei ole vaikea - jos $ab = ac$, niin osittelulain nojalla

$$a(b - c) = ab - ac = 0_R.$$

Koska $a \neq 0_R$ ja kokonaisalueessa on voimassa nollasääntö, tästä seuraa, että $b - c = 0_R$, toisin sanoen $b = c$.

Lemma 3.34. *Olkoot* $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$. *Tällöin*

- $\deg(\mathbf{p} + \mathbf{q}) \leq \max\{\deg \mathbf{p}, \deg \mathbf{q}\}$,
- $\deg \mathbf{pq} = \deg \mathbf{p} + \deg \mathbf{q}$,
- $K[\mathbf{X}]$ *on renkaana* kokonaisalue.

Tässä tulkitaan $-\infty + n = -\infty$, $(-\infty) \cdot n = -\infty$ *ja* $-\infty \leq n$ *jokaisella* $n \in \mathbb{N} \cup \{-\infty\}$.

Todistus. Harjoitustehtävä. □

Huomautus: Edellisen lemman tulokset paljastavat, miksi nollapolynomin asteeksi on poikkeuksellisesti sovittu miinus ääretömäksi. Nimittäin ilman tätä sopimusta lemman väitteet eivät välttämättä olisi enää voimassa sellaisissa erikoistapauksissa, joissa toinen polynomeista $\mathbf{p}, \mathbf{q}$ on nolla-polynomi.

Olkoot $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$. Sanotaan, että polynomi $\mathbf{p}$ on *jaollinen* polynomilla $\mathbf{q}$, jos on olemassa polynomi $\mathbf{s} \in K[X]$ siten, että $\mathbf{p} = \mathbf{sq}$. Polynomia $\mathbf{q}$ sanotaan tällöin polynomin $\mathbf{p}$ *tekijäksi*. Jos polynomi $\mathbf{p}$ on jaollinen polynomilla $\mathbf{q}$, niin merkitään $\mathbf{q} \mid \mathbf{p}$. Edellisen lemman nojalla tällöin pätee yhtälö $\deg \mathbf{p} = \deg \mathbf{s} + \deg \mathbf{q}$, josta seuraa, että $\deg \mathbf{q} \leq \deg \mathbf{p}$, kun $\mathbf{q} \mid \mathbf{p}$, paitsi jos $\mathbf{p} = \mathbf{0}$ on nollapolynomi. Huomaa, että nollapolynomi on jaollinen millä tahansa polynomilla (valitaan polynomiksi $\mathbf{s}$ nollapolynomi).

Polynomien jaollisuusteorialla on paljon yhteistä kokonaislukujen jaollisuusteorian kanssa. Esimerkiksi seuraavassa propositiossa tarkasteltava tärkeä tulos on täysin analoginen kokonaislukujen joukossa voimassa olevan vastaavan ominaisuuden kanssa.

Propositio 3.35. Polynomien jakoyhtälö
Olkoot $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$, $\mathbf{q} \neq \mathbf{0}$. *Tällöin on olemassa* yksikäsitteiset *polynomit* $\mathbf{r}, \mathbf{s} \in K[\mathbf{X}]$ *siten, että*

$$\mathbf{p} = \mathbf{sq} + \mathbf{r}$$

ja $\deg \mathbf{r} < \deg \mathbf{q}$.

Polynomia $\mathbf{r}$ *sanotaan* jakojäännökseksi *polynomin* $\mathbf{p}$ *jaettaessa polynomilla* $\mathbf{q}$.

Todistus. Aloitetaan osoittamalla polynomien $\mathbf{r}$, $\mathbf{s}$ olemassaoloa. Todistus tulee olemaan formaalisaatio koulusta tutusta jakokulma-algoritmista. Jos $\deg \mathbf{p} < \deg \mathbf{q}$, voidaan valita yksinkertaisesti $\mathbf{p} = \mathbf{r}$, $\mathbf{s} = \mathbf{0}$. Voidaan siis olettaa, että polynomin $\mathbf{p}$ aste on vähintään yhtä suuri kuin polynomin $\mathbf{q}$ aste, $\deg \mathbf{p} \geq \deg \mathbf{q}$.

Menetelmän valaisemiseksi käydään ensin läpi jokin konkreettinen esimerkki. Olkoot esimerkiksi $K = \mathbb{Q}$, $\mathbf{p} = \mathbf{X}^4+1$ ja $\mathbf{q} = 2\mathbf{X}^2+\mathbf{X}+1$. Halutaan jakaa polynomi $\mathbf{p}$ polynomilla $\mathbf{q}$ jakokulmassa. Ideana on "eliminoida" polynomin $\mathbf{p}$ termejä $a_i\mathbf{X}^i$ vähentämällä siitä polynomi $\mathbf{q}$ sopivalla polynomilla kerrottuna. Yritetään ensin päästä eroon polynomin $\mathbf{p}$ korkeinta potenssia olevasta termistä eli termistä $\mathbf{X}^4$, jonka aste on 4. Koska polynomin $\mathbf{q}$ korkeimman asteen termi on $2\mathbf{X}^2$(erityisesti astetta 2 oleva) meidän on nostettava potenssia 2:llä ja samalla mitätöidä kertoimen 2 vaikutusta kertomalla sen käänteisluvulla. Kerrotaan siis polynomi $\mathbf{q}$ polynomilla $\mathbf{h} = \frac{1}{2}\mathbf{X}^2$. Tällöin

$$\mathbf{p} - \mathbf{hq} = (\mathbf{X}^4 + 1) - (\mathbf{X}^4 + \frac{1}{2}\mathbf{X}^3 + \frac{1}{2}\mathbf{X}^2) = -\frac{1}{2}\mathbf{X}^3 - \frac{1}{2}\mathbf{X}^2 + 1.$$

Merkitään $\mathbf{p}' = -\frac{1}{2}\mathbf{X}^3 - \frac{1}{2}\mathbf{X}^2 + 1$ ja sovelletaan siihen samaa *jakoalgoritmia*. Pannaan erityisesti merkille, että polynomin $\mathbf{p}'$ aste on aidosti pienempi kuin polynomin $\mathbf{p}$ aste.

Eliminoidaan polynomin $\mathbf{p}'$ termi $-\frac{1}{2}\mathbf{X}^3$ vähentämällä siitä polynomi $(-\frac{1}{4}\mathbf{X})\mathbf{q}$. Tällöin

$$\mathbf{p}' - (-\frac{1}{4}\mathbf{X})\mathbf{q} = -\frac{1}{4}\mathbf{X}^2 + \frac{1}{4}\mathbf{X} + 1 = \mathbf{p}'',$$

Viimeisenä toimipiteenä eliminoidaan uuden polynomin $\mathbf{p}''$ (jonka aste on nyt kaksi) korkeimman asteen termi, jolloin saadaan

$$\mathbf{p}'' + \frac{1}{8}\mathbf{q} = \frac{3}{8}\mathbf{X} + \frac{9}{8} = \mathbf{r}.$$

Algoritmi päättyy tähän, koska lopputuloksena saatiin polynomi, jonka aste on pienempi kuin jaettavan polynomin $\mathbf{q}$. Tämä polynomi $\mathbf{r}$ kelpaa jakojäännökseksi. Yhdistämällä laskut saadaan

$$\mathbf{p} = \mathbf{p}' + \frac{1}{2}\mathbf{X}^2\mathbf{q} = \mathbf{p}'' - (\frac{1}{4}\mathbf{X} + \frac{1}{2}\mathbf{X}^2)\mathbf{q} = (\frac{1}{2}\mathbf{X}^2 - \frac{1}{4}\mathbf{X} - \frac{1}{8})\mathbf{q} + \mathbf{r} = \mathbf{sq} + \mathbf{r},$$

missä $\mathbf{s} = \frac{1}{2}\mathbf{X}^2 - \frac{1}{4}\mathbf{X} - \frac{1}{8}$ ja $\deg \mathbf{r} < \deg \mathbf{q}$.

Koko lasku jakokulmassa:

$$\begin{array}{rl|l}
 & & \frac{1}{2}X^2 - \frac{1}{4}X - \frac{1}{8} \\
\cline{3-3}
2X^2 + X + 1 & X^4 \qquad\qquad + 1 & \\
 & -X^4 - \frac{1}{2}X^3 - \frac{1}{2}X^2 & \\
\cline{2-2}
 & -\frac{1}{2}X^3 - \frac{1}{2}X^2 & \\
 & \frac{1}{2}X^3 + \frac{1}{4}X^2 + \frac{1}{4}X & \\
\cline{2-2}
 & -\frac{1}{4}X^2 + \frac{1}{4}X + 1 & \\
 & \frac{1}{4}X^2 + \frac{1}{8}X + \frac{1}{8} & \\
\cline{2-2}
 & \frac{3}{8}X + \frac{9}{8} &
\end{array}$$

Yleisellä tasolla formaali todistus etenee samalla periatteella. Oletetaan, että $m = \deg \mathbf{p} \geq \deg \mathbf{q} = n$, sillä muuten väite toteutuu valitsemalla $\mathbf{s} = \mathbf{0}$ ja $\mathbf{r} = \mathbf{p}$. Osoitetaan väite induktiolla luvun $m \leq n$ suhteen. Olkoon polynomin $\mathbf{q}$ johtava kerroin a ja vastaavasti polynomin $\mathbf{p}$ johtava kerroin b. Tällöin

$$\mathbf{p}' = \mathbf{p} - ba^{-1}\mathbf{X}^{m-n}\mathbf{q}$$

on polynomi, jonka aste on aidosti pienempi kuin polynomin $\mathbf{p}$ aste. Lisäksi pätee $\mathbf{p} = \mathbf{p}' + ba^{-1}\mathbf{X}^{m-n}\mathbf{q}$. Tässä $a \neq 0_K$, sillä polynomi $\mathbf{q}$ oletetaan olevan ei-nollapolynomi, joten käänteisalkio a^{-1} on olemassa.

Koska $\deg \mathbf{p}' < \deg \mathbf{p}$, voidaan edetä induktiivisesti. Jos $m = n$ (induktion alkuaskel), ollaan valmiit. Muuten voidaan induktiivisesti olettaa, että on olemassa $\mathbf{r}, \mathbf{s}' \in K[\mathbf{X}]$ siten, että

$$\mathbf{p}' = \mathbf{s}'\mathbf{q} + \mathbf{r}$$

ja $\deg \mathbf{r} < \deg \mathbf{q}$. Tällöin $\mathbf{p} = \mathbf{p}' + ba^{-1}\mathbf{X}^{m-n}\mathbf{q} = \mathbf{s}\mathbf{q} + \mathbf{r}$, missä $\mathbf{s} = \mathbf{s}' + ba^{-1}\mathbf{X}^{m-n}$. Olemassaolo on siten todistettu.

Todistetaan vielä polynomien $\mathbf{s}$ ja $\mathbf{r}$ yksikäsitteisyys. Oletetaan, että

$$\mathbf{s}_1\mathbf{q} + \mathbf{r}_1 = \mathbf{p} = \mathbf{s}_2\mathbf{q} + \mathbf{r}_2, \tag{3.36}$$

missä $\deg \mathbf{r}_1, \deg \mathbf{r}_2 < \deg \mathbf{q}$. Tällöin erityisesti pätee

$$(\mathbf{s}_1 - \mathbf{s}_2)\mathbf{q} = \mathbf{r}_2 - \mathbf{r}_1.$$

Jos $\mathbf{s}_1 - \mathbf{s}_2 \neq \mathbf{0}$ ei ole nolla-polynomi, niin edellisen yhtälön vasemmalla puolella esiintyy polynomi, jonka aste (Lemman 3.34 nojalla) on

$$\deg(\mathbf{s}_1 - \mathbf{s}_2) + \deg \mathbf{q} \geq \deg \mathbf{q}.$$

Oikealla puolella taas esiintyy polynomi $\mathbf{r}_2 - \mathbf{r}_1$, jonka aste on $< \deg \mathbf{q}$. Tämä on ristiriita. Näin olleen täytyy olla $\mathbf{s}_1 - \mathbf{s}_2 = \mathbf{0}$ eli pätee $\mathbf{s}_1 = \mathbf{s}_2$. Yhtälöstä 3.36 seuraa tällöin, että myös $\mathbf{r}_1 = \mathbf{r}_2$. Yksikäsitteisyys on osoitettu. □

Pääideaalirenkaat

Olkoon $(R, +, \cdot)$ *rengas*. Palautetaan mieleen, että renkaan R *ideaali* on sellainen sen osajoukko I, jolle pätevät seuraavat ehdot (i) ja (ii):

(i) $(I, +)$ on Abelin ryhmän $(R, +)$ aliryhmä.

(ii) Kaikilla $x \in I$ ja kaikilla $a \in R$ pätee $xa \in I$ ja $ax \in I$.

Helposti nähdään, että renkaan R mielivaltaisen ideaaleista koostuvan perheen $(I_\alpha)_{\alpha \in \mathbf{A}}$ *leikkaus*

$$\bigcap_{\alpha \in \mathbf{A}} I_\alpha$$

on myös renkaan R ideaali.

Olkoon $A \subset R$ mikä tahansa renkaan R osajoukko. Olkoon $(I_\alpha)_{\alpha \in \mathcal{A}}$ perhe renkaan R ideaaleita, joka koostuu täsmälleen niistä renkaan R ideaaleista J, joille pätee $A \subset J$.

Tämä perhe on epätyhjä, sillä esimerkiksi rengas R on itseesä ideaali, jolle pätee $A \subset R$. Edellisen nojalla tämän perheen leikkaus $I = \bigcap_{\alpha\in\mathcal{A}} I_\alpha$ on myös renkaan R ideaali. Koska $A \subset I_\alpha$ kaikilla $\alpha \in \mathcal{A}$, leikkauksen määritelmän mukaan pätee $A \subset I$, eli I on ideaali, joka sisältää joukon A, toisin sanoen on itse perheen $(I_\alpha)_{\alpha\in\mathcal{A}}$ jäsen. Toisaalta $I \subset I_\alpha$ kaikilla $\alpha \in \mathcal{A}$, joten I on tämän perheen *pienin* jäsen sisältyvyysrelaation suhteen. Ollaan todistettu seuraava lemma. Vertaa ideaalin I konstruktio ja sen olemassaolon todistusta vektoriavaruuden osajoukon virittämän aliavaruuden konstruktioon, Propositio 2.28.

Lemma 3.37. *Olkoon A renkaan R osajoukko. Tällöin on olemassa yksikäsitteinen (sisältyvyysrelaation suhteen) pienin renkaan R ideaali, joka sisältää osajoukon A. Tätä ideaalia sanotaan joukon A* viritetyksi *ideaaliksi ja merkitään symbolilla (A).*

Olkoot $r_1, \ldots, r_n \in R$ renkaan R alkiot (joita on siis äärellinen määrä). Joukon $A = \{r_1, \ldots, r_n\}$ virittämää ideaalia merkitään myös $(r_1, r_2, \ldots, r_n)$. Osoittautuu, että **vaihdannaisen** renkaan tapauksessa on olemassa suhteellisen yksinkertainen tapa karakterisoida tämän ideaalin alkioita.

Lemma 3.38. *Olkoon rengas R* vaihdannainen *ja olkoot $r_1, \ldots, r_n \in R$. Tällöin*

$$(r_1, \ldots, r_n) = \{r_1x_1 + r_2x_2 + \ldots + r_nx_n \mid x_i \in R, i = 1, \ldots, n\}.$$

Todistus. Olkoon

$$J = \{r_1x_1 + r_2x_2 + \ldots + r_nx_n \mid x_i \in R, i = 1, \ldots, n\}.$$

Olkoon $i = 1, \ldots, n$ kiinnitetty. Valitsemalla $x_i = 1_R$, $x_j = 0_R$ kaikilla $j \neq i$, nähdään, että $r_i \in J$. Olkoon I mikä tahansa renkaan R sellainen ideaali, jolle pätee $r_1, \ldots, r_n \in I$. Tällöin ideaalin määritelmän ehdoista (i) ja (ii) seuraa heti, että $J \subset I$. Näin ollen riittää näyttää, että J todellakin on renkaan R ideaali.

Sen todistaminen, että J on ryhmän $(R, +)$ aliryhmä jätetään lukijalle harjoitustehtäväksi. Näytetään, että J toteuttaa myös ehdon (ii) ideaalin määritelmässä. Olkoon

$$y = r_1x_1 + r_2x_2 + \ldots + r_nx_n \in I,$$

$x_i \in R$ kaikilla $i = 1, \ldots, n$. Olkoon $a \in R$ mielivaltainen. Tällöin kertolaskun liitännäisyyden ja osittelulain nojalla

$$ya = (r_1x_1 + r_2x_2 + \ldots + r_nx_n)a = r_1(x_1a) + r_2(x_2a) + \ldots + r_n(x_na) \in J.$$

Lisäksi, koska R on *vaihdannainen* rengas, $ay = ya$, joten myös $ay \in J$. On näytetty, että J toteuttaa ideaalin määritelmän ehdon (ii). □

Huomautus: Edellisen lemman tulos on yleisesti ottaen voimassa ainoastaan *vaihdannaisessa* renkaassa. Jos kertolasku ei ole vaihdannainen, asia on monimutkaisempi.

Tarkastellaan edellisen lemman erikoistapausta, jossa $n = 1$. Yksilön $\{r\}$ virittämää renkaan R ideaalia merkitään (r). Edellisen lemman nojalla vaihdannaisen renkaan R alkiolle x pätee

$$(r) = \{rx \mid x \in R\}.$$

Ideaalia, joka on muotoa (r) jollakin $r \in R$, sanotaan *pääideaaliksi.*

Vaihdannaista rengasta R sanotaan **pääideaalirenkaaksi**, jos sen jokainen ideaali on pääideaali.

Propositio 3.39. *Rengas $K[\mathbf{X}]$ on pääideaalirengas.*

Todistus. Olkoon J polynomialgebran $K[\mathbf{X}]$ ideaali. Jos $J = \{\mathbf{0}_R\}$ on triviaali nollaalkion virittämä ideaali, se on pääideaali, sillä tällöin $J = (\mathbf{0}_R)$. Oletetaan, että ideaali J ei ole triviaali eli sisältää muitakin alkioita kuin nollapolynomin. Valitaan tällaisista eli nollapolynomista eroavista J:n alkioista sellainen $\mathbf{q} \in J$, jonka aste on pienin mahdollinen. Tämä on mahdollista, sillä (nollasta eroavan) polynomin aste on luonnollinen luku ja luonnollisten lukujen jokaisessa epätyhjässä osajoukossa on olemassa pienin luku. Osoitetaan, että ideaali J on tämän alkion $\mathbf{q}$ virittämä, eli $J = (\mathbf{q})$.

Koska $\mathbf{q} \in J$, pätee selvästi $(\mathbf{q}) \subset J$. Riittää siis osoittaa, että jokainen $\mathbf{p} \in J$ voidaan esittää muodossa $\mathbf{p} = \mathbf{sq} \in (\mathbf{q})$ jollakin polynomilla $\mathbf{s}$. Proposition 3.35 nojalla on olemassa polynomit $\mathbf{s}, \mathbf{r} \in K[\mathbf{X}]$ siten, että $\mathbf{p} = \mathbf{sq} + \mathbf{r}$, missä $\deg \mathbf{r} < \deg \mathbf{q}$. Nyt riittää osoittaa, että $\mathbf{r} = \mathbf{0}$.

Koska J on ideaali ja $\mathbf{p}, \mathbf{q} \in J$, pätee $\mathbf{r} = \mathbf{p} - \mathbf{sq} \in J$. Jos $\mathbf{r} \neq \mathbf{0}$ saadaan tästä ristiriita alkion $\mathbf{q}$ valinnan kanssa, sillä $\deg \mathbf{r} < \deg \mathbf{q}$. Näin ollen $\mathbf{r} = \mathbf{0}$. □

Olkoon R *kokonaisalue* (eli vaihdannainen rengas, jossa ei ole nollanjakajia) ja olkoon I jokin sen pääideaali. Olkoot r, r' molemmat ideaalin I virittäjät,

$$(r) = I = (r').$$

Jos $r = 0_R$, niin $I = \{0_R\}$ on triviaali ideaali, jolloin myös $r' = 0_R$. Samanlainen johtopäätös saadaan, jos $r' = 0_R$. Oletetaan, että $r \neq 0_R \neq r'$. Lemman 3.38 nojalla on olemassa $x \in R$ siten, että $r' = rx$. Vastaavasti on olemassa $y \in R$ siten, että $r = r'y$. Tästä saadaan

$$r = r'y = (rx)y = r(xy).$$

Koska $r \neq 0_R$ ja renkaassa R ei ole nollanjakajia, supistamalla r yhtälön molemmalta puolilta saadaan $xy = 1_R$. Koska R on vaihdannainen, tästä seuraa, että myös $yx = 1_R$, toisin sanoen x ja y ovat molemmat renkaan R *kääntyviä* alkioita, $x, y \in R^*$[2]. Lisäksi pätee $y = x^{-1}$. Kääntäen, olkoon $r \in R$ ja olkoon $r' = rx$, missä $x \in R^*$ on renkaan R kääntyvä alkio. Tällöin helposti nähdään (tarkista!), että $(r) = (r')$. Näin ollen, kaksi kokonaisalueen R alkiota r, r' virittävät saman pääideaalin jos ja vain jos on olemassa renkaan kääntyvä alkio $x \in R^*$ jolle pätee $r = r'x$. Tämä on todistetty tapauksessa $r \neq 0_R \neq r'$, mutta pätee yhtä hyvin myös tapauksessa $r = 0_R = r'$, sillä tällöin voidaan valita $x = 1_R$ (renkaan ykkösalkio on aina kääntyvä).

Sovelletaan edellisen kappaleen tulosta polynomirenkaaseen $K[\mathbf{X}]$. Lemman 3.34 nojalla tämä rengas on kokonaisalue. Edellisestä seuraa, että polynomit $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ virittävät saman renkaan $K[\mathbf{X}]$ ideaalin, $(\mathbf{p}) = (\mathbf{q})$, jos ja vain jos $\mathbf{p} = \mathbf{sq}$, missä $\mathbf{s}$ on *kääntyvä* polynomi. Tutkitaan, mitkä polynomirenkaan $K[\mathbf{X}]$ alkiot ovat kääntyviä tässä renkaassa. Osoittautuu, että tällaisia ovat täsmälleen nollasta eroavat vakiopolynomit.

[2]Renkaan kohdalla "tähti"-merkintä R^* tarkoittaa renkaan kääntyvien alkioden muodostamaa joukkoa. Vektoriavaruuksien kohdalla merkintää V^* taas tarkoittaa avaruuden duaali-avaruutta. Kyseessä on siis ikävä "tupla-notaatio" - samalla merkinnällä tarkoitetaan eri asioita kontekstista riippuen. Valitettavasti molemmat merkintätavat ovat varsin vakiintuneita ja yleisessä käytössä, joten meidänkin on pakko käyttää niitä.

Koska vakiopolynomi on sovittu olevan sama asia kuin vastaavan kunnan K alkio, voidaan kirjoittaa

$$K[\mathbf{X}]^* = K^* = K \setminus \{0_K\}.$$

Lemma 3.40. *Olkoon* $\mathbf{p} \in K[\mathbf{X}]$*. Tällöin* $\mathbf{p}$ *on kääntyvä renkaassa* $K[\mathbf{X}]$ *(polynomien kertolaskun suhteen) jos ja vain jos* $\deg \mathbf{p} = 0$ *eli jos ja vain jos* $\mathbf{p}$ *on nollasta eroava vakiopolynomi.*

Todistus. Harjoitustehtävä. □

Näin ollen kaksi polynomia $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ virittävät saman renkaan $K[\mathbf{X}]$ ideaalin, $(\mathbf{p}) = (\mathbf{q})$, jos ja vain jos on olemassa $k \in K, k \neq 0_K$ siten, että $\mathbf{p} = k\mathbf{q}$. Tällöin polynomeilla $\mathbf{p}, \mathbf{q}$ on erityisesti sama aste ja, jos kyseessä ei ole triviaali nollan viritämä ideaali, niiden johtaville kertoimille a_n, b_n pätee $a_n = kb_n$. Valitsemalla $b_n = k^{-1}$, nähdään, että epätriviaalille pääideaalille $(\mathbf{p})$ voidaan aina valita virittäjä $\mathbf{p}$, jonka johtava kerroin on 1_k. Tällaisia polynomeja sanomme *pääpolynomeiksi*. Lisäksi edellisestä seuraa, että tällainen pääideaalin pääpolynomi-virittäjä on *yksikäsitteinen*. Koska Lemman 3.39 nojalla jokainen renkaan $K[\mathbf{X}]$ ideaali on pääideaali, tästä seuraa, että jokainen polynomirenkaan $K[\mathbf{X}]$ epätriviaali ideaali $I \neq \{\mathbf{0}\}$ voidaan esittää muodossa $I = (\mathbf{p})$, missä $\mathbf{p}$ on pääpolynomi, *yksikäsitteisellä tavalla*.

Polynomien jaollisuusteoria

Palautetaan mieleen, että polynomi $\mathbf{p} \in K[\mathbf{X}]$ on *jaollinen* polynomilla $\mathbf{q} \in K[\mathbf{X}]$ jos on olemassa sellainen polynomi $\mathbf{s} \in K[\mathbf{X}]$, jolle pätee $\mathbf{p} = \mathbf{sq}$. Tällöin merkitään $\mathbf{q} \mid \mathbf{p}$ ja sanotaan, että polynomi $\mathbf{q}$ *jakaa* polynomin $\mathbf{p}$ tai on sen *tekijä*. Pääideaalien kielellä jaollisuusrelaatio voidaan muotoilla seuraavasti (todistus jätetään harjoitustehtäväksi).

Lemma 3.41. *Olkoot* $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ *polynomeja. Tällöin* $\mathbf{q} \mid \mathbf{p}$ *jos ja vain jos polynomien* $\mathbf{p}$ *ja* $\mathbf{q}$ *virittämille pääideaaleille pätee* $(\mathbf{p}) \subset (\mathbf{q})$*.*

Jokainen polynomi $\mathbf{p} \in K[\mathbf{X}]$ on selvästi jaollinen jokaisella nollasta eroavalla nolla-asteisella polynomilla $k \in K, k \neq 0_K$ (muista, että nolla-steiset polynomit samaistetaan kunnan K alkioiden kanssa), sekä jokaisella muotoa $k\mathbf{p}$ olevalla polynomilla, missä $k \in K$, $k \neq 0_K$. Jos nämä ovat polynomin $\mathbf{p}$ **ainoat** tekijät ja lisäksi pätee $\deg \mathbf{p} > 0$ (eli polynomi $\mathbf{p}$ ei ole kunnan K alkio), polynomia $\mathbf{p}$ sanotaan **jaottomiksi**. Tästä määritelmästä seuraa helposti, että polynomi $\mathbf{p} \neq \mathbf{0}$ ei ole jaoton jos ja vain jos $\deg \mathbf{p} = 0$ (eli polynomi on kunnan K alkio) tai $\mathbf{p} = \mathbf{qr}$ joillakin $\mathbf{q}, \mathbf{r} \in K[X]$, $\deg \mathbf{q} < \mathbf{p}, \deg \mathbf{r} < \mathbf{p}$. Toisin sanoen ei-vakiopolynomi $\mathbf{p}$ ei ole jaoton jos ja vain jos se voidaan esittää kahden *aidosti pienempää astetta olevan polynomin* tulona. Tällöin kumpikin näistä polynomin $\mathbf{p}$ tekijöistä ei myöskään ole vakiopolynomi. Tästä karakterisaatiosta erityisesti seuraa, että jokainen 1-asteinen polynomi on jaoton.

Jaottomilla polynomeilla on polynomien jaollisuusteoriassa sama rooli kuin *alkuluvuilla* on luonnollisten lukujen jaollisuusteoriassa. Muun muassa pätee tärkeä *hajotelmalause* (joka kokonaislukujen teoriassa tunnetaan nimellä "aritmetiikan peruslause"), jonka mukaan jokainen polynomi voidaan esittää jaottomien polynomien tulona, vieläpä "oleellisesti" yksikäsitteisellä tavalla. Osoitetaan ensin, että tällainen esitys on olemasaa.

Propositio 3.42. *Jokainen ei-vakio polynomi* $\mathbf{p} \in K[\mathbf{X}]$ *voidaan esittää jaottomien polynomien tulona. Tarkemmin sanottuna, jos* $\mathbf{p} \in K[\mathbf{X}]$ *on ei-vakio polynomi, niin on olemassa jaottomat polynomit* $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m$ *siten, että*

$$\mathbf{p} = \mathbf{q}_1\mathbf{q}_2 \ldots \mathbf{q}_m.$$

Todistus. Osoitetaan väite induktiolla polynomin asteen $n = \deg \mathbf{p}$ suhteen. Jos $n = 1$, polynomi on jaoton, jolloin väite pätee triviaalisti.

Olkoon $\deg \mathbf{p} = n > 1$. Oletetaan, että väite on osoitettu kaikille pienempiasteisille polynomeille. Jos $\mathbf{p}$ on jaoton, asia on selvä. Muuten $\mathbf{p} = \mathbf{qr}$ joillakin $\mathbf{q}, \mathbf{r} \in K[X]$, $\deg \mathbf{q} < \mathbf{p}, \deg \mathbf{r} < \mathbf{p}$. Lemmasta 3.34 seuraa tällöin, että $\deg \mathbf{q} + \deg \mathbf{r} = \mathbf{p}$, joten kumpikin polynomeista $\mathbf{q}, \mathbf{r}$ ei voi olla nolla-asteinen (sillä muuten toisen aste olisi sama kuin polynomin $\mathbf{p}$ aste). Induktio-oletuksen nojalla kumpikin polynomeista $\mathbf{q}, \mathbf{r}$ voidaan esittää jaottomien polynomien tulona. Yhdistämällä näitä tuloja saadaan polynomille $\mathbf{p}$ esitys jaottomien polynomien tulona. □

Seuraavaksi osoitetaan että polynomin esitys jaottomien polynomien tulona

$$\mathbf{p} = \mathbf{q}_1\mathbf{q}_2 \ldots \mathbf{q}_m \tag{3.43}$$

on "oleellisesti" yksikäsitteinen. Selvitetään ensin, mitä "oleellinen yksikäsitteisyys" tarkoittaa tässä yhteydessä. Ei voida vaatia, että tällainen esitys olisi kirjaimellisesti yksikäsitteinen kahdesta syystä. Ensinnäkin esityksessä 3.43 oikealla puolella esiintyvät polynomit voidaan kirjoittaa missä tahansa järjestyksessä, koska polynomien kertolasku on vaihdannainen. Toinen syy on seuraava. Ei ole vaikeata huomata, että jos $\mathbf{q} \in K[\mathbf{X}]$ on jaoton polynomi ja $k \in K, k \neq 0$, on mikä tahansa skalaari, niin myös polynomi $k\mathbf{q}$ on jaoton. Näin ollen tulo 3.43 voidaan yhtä hyvin kirjoittaa myös esimerkiksi muodossa

$$\mathbf{p} = (k_1\mathbf{q}_1)(k_2\mathbf{q}_2) \ldots (k_m\mathbf{q}_m), \tag{3.44}$$

missä skalaarit $k_1, k_2, \ldots, k_m$ on valittu niin, että $k_1k_2 \ldots k_m = 1_K$. Nämä havainnot hankaloittavat yksikäsitteisyys-väitteen tarkkaa formulointia. Tästä syystä polynoimien teorian "Aritmetiikan peruslause"[3] eli Propositiolle 3.48 on valittu hieman erilainen muotoilu kuin vastaavalle tuloksele luonnollisille luvuille.

Polynomeja $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ sanotaan *keskenään jaottomiksi*, jos ainoat niiden *yhteiset tekijät* ovat vakiopolynomeja. Täsmällisemmin sanottuna $\mathbf{p}, \mathbf{q}$ ovat keskenään jaottomia jos ehdoista $\mathbf{r} \in K[\mathbf{X}]$, $\mathbf{r} \mid \mathbf{p}$, $\mathbf{r} \mid \mathbf{q}$ seuraa, että $\mathbf{r}$ on vakiopolynomi (eli sen aste on nolla tai miinus ääretön).

Lemma 3.45. *Polynomit* $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ *ovat keskenään jaottomia jos ja vain jos on olemassa sellaiset polynomit* $\mathbf{s}, \mathbf{t} \in K[\mathbf{X}]$ *joille pätee*

$$\mathbf{ps} + \mathbf{qt} = 1_K.$$

Tässä 1_K *on* 0*-asteinen vakiopolynomi* $1_K\mathbf{X}^0$.

[3]Tulosta, jonka mukaan jokainen positiivinen kokonaisluku voidaan kirjoittaa alkulukujen tulona yksiksäsitteisellä tavalla (tekijöiden järjestystä vaille) sanotaan "Aritmetiikan peruslauseeksi". Polynomien teoriassa vastaava tulos muotoillaan Propositiossa 3.48.

Todistus. Ehdon riittävyyden osoittaminen jätetään harjoitustehtäväksi.

Oletetaan, että polynomit $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ ovat keskenään jaottomia. Koska nollapolynomi on jaollinen millä tahansa polynomilla, se ei voi olla keskenään jaoton minkään polynomin kanssa. Erityisesti $\mathbf{p}$ ja $\mathbf{q}$ eivät ole nollapolynomeja. Lemman 3.38 nojalla polynomien $\mathbf{p}, \mathbf{q}$ *virittämälle ideaalille* $I = (\mathbf{p}, \mathbf{q})$ pätee

$$I = \{\mathbf{ps} + \mathbf{qt} \mid \mathbf{s}, \mathbf{t} \in K[\mathbf{X}]\}.$$

Tämä joukko ei ole triviaali ideaali $\{\mathbf{0}\}$, sillä se sisältää ainakin polynomin $\mathbf{p}$, joka ei voi olla nollapolynomi. Proposition 3.39 nojalla I on pääideaali, toisin sanoen on olemassa sellainen $\mathbf{r}$ jolle pätee

$$I = (\mathbf{r}) = \{\mathbf{rs} \mid \mathbf{s} \in K[\mathbf{X}]\}.$$

Lisäksi voidaan olettaa, että virittäjäpolynomi $\mathbf{r}$ on *pääpolynomi* (jokaisella ei-triviaalilla polynomirenkaan ideaalilla on yksikäsitteinen pääpolynomi-virittäjä).

Koska $\mathbf{p} \in I$, tästä seuraa, että on olemassa sellainen polynomi $\mathbf{s} \in K[\mathbf{X}]$ jolle pätee $\mathbf{p} = \mathbf{rs}$. Määritelmän mukaan tämä tarkoittaa sitä, että $\mathbf{r}$ on polynomin $\mathbf{p}$ *tekijä*. Samalla tavalla nähdään, että $\mathbf{r}$ on polynomin $\mathbf{q}$ tekijä. Koska polynomit $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ ovat keskenään jaottomia, polynomin $\mathbf{r}$ on oltava vakiopolynomi, eli muotoa $k \in K$. Koska $\mathbf{r}$ on pääpolynomi, täytyy olla $k = 1_K$. Toisin sanoen $\mathbf{r} = 1_K$.

Lemman 3.38 nojalla on olemassa polynomit $\mathbf{s}, \mathbf{t} \in K[\mathbf{X}]$ joille pätee

$$\mathbf{ps} + \mathbf{qt} = \mathbf{r} = 1_K.$$

Väite on todistettu. □

Lemma 3.46. *Olkoot* $\mathbf{p}, \mathbf{q}, \mathbf{r} \in K[\mathbf{X}]$. *Oletetaan, että tulo* $\mathbf{qr}$ *on jaollinen polynomilla* $\mathbf{p}$. *Oletetaan lisäksi, että polynomit* $\mathbf{p}$ *ja* $\mathbf{q}$ *ovat keskenään jaottomia. Tällöin polynomi* $\mathbf{r}$ *on jaollinen polynomilla* $\mathbf{p}$.

Todistus. Edellisen lemman nojalla on olemassa polynomit $\mathbf{s}, \mathbf{t} \in K[\mathbf{X}]$ joille pätee yhtälö $\mathbf{ps} + \mathbf{qt} = 1_K$. Kertomalla tämä yhtälö polynomilla $\mathbf{r}$ saadaan

$$\mathbf{r} = \mathbf{p}(\mathbf{rs}) + (\mathbf{rq})\mathbf{t},$$

missä $\mathbf{rq} = \mathbf{pu}$ jollakin $\mathbf{u} \in K[\mathbf{X}]$ oletuksen nojalla. Näin ollen $\mathbf{r} = \mathbf{p}(\mathbf{rs} + \mathbf{ut})$, mistä väite seuraa. □

Seuraus 3.47. *Olkoon* $\mathbf{p}$ jaoton *polynomi ja olkoot* $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m$ *polynomeja. Oletetaan, että* $\mathbf{p}$ *jakaa tulon* $\mathbf{q}_1\mathbf{q}_2 \ldots \mathbf{q}_m$. *Tällöin on olemassa* $i = 1, \ldots, m$ *siten, että* $\mathbf{p}$ *jakaa polynomin* $\mathbf{q}_i$.

Todistus. Pannaan ensin merkille seuraava tosiasia. Olkoon $\mathbf{p}$ *jaoton* polynomi ja olkoon $\mathbf{q}$ *mielivaltainen* polynomi. Tällöin joko $\mathbf{p}$ on polynomin $\mathbf{q}$ tekijä tai polynomit $\mathbf{p}, \mathbf{q}$ ovat *keskenään jaottomat*. Tämä seuraa suoraan jaottoman polynomin määritelmästä.

Todistetaan seurauksen väite todeksi. Jos $\mathbf{p}$ on polynomin $\mathbf{q}_1$ tekijä, ollaan valmiit. Muuten polynomit $\mathbf{p}$ ja $\mathbf{q}_1$ ovat (edellisen kappaleen nojalla) keskenään jaottomia. Koska $\mathbf{p}$ jakaa tulon $\mathbf{q}_1(\mathbf{q}_2 \ldots \mathbf{q}_m)$, edellisestä Lemmasta seuraa, että $\mathbf{p}$ on tulon $\mathbf{q}_2 \ldots \mathbf{q}_m$ tekijä. Väite saadaan jatkamalla tällä tavalla induktiolla. □

Nyt voidaan muotoilla ja todistaa ”polynomialgebran aritmetiikan peruslause”. Huomaa, että vaikka Propositio 3.42 oli muotoiltu ja todistettu ainoastaan ei-vakio polynomeille, seuraava tulos on muotoiltu sillä tavalla, että se pätee myös nollasta eroaville vakio-polynomeille.

Propositio 3.48. *Jokainen polynomi* $\mathbf{p} \neq \mathbf{0} \in K[X]$ *voidaan esittää muodossa*

$$\mathbf{p} = k \cdot \mathbf{q}_1\mathbf{q}_2 \dots \mathbf{q}_m \tag{3.49}$$

missä $k \in K^*$ *ja* $\mathbf{q}_i$ *on jaoton* pääpolynomi *jokaisella* $i = 1, \dots, m$, $m \in \mathbb{N}$. *Tällainen esitys on polynomien* $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m$ *järjestystä vaille yksikäsitteinen.*

Todistus. Jos $\mathbf{p} = k'$ on nollasta eroava vakio-polynomi, se voidaan kirjoittaa muodossa 3.49 ainoastaan jos valitaan $k = k'$ ja $m = 0$ (eli oikealla puolella ei esiinny jaottomia polynomeja lainkaan), sillä muuten oikealla puolella esiintyisi polynomi, jonka aste on ainakin 1 (Lemma 3.34). Näin ollen väite on tosi vakio-polynomeille ja voimme jatkossa olettaa, että $\deg \mathbf{p} \geq 1$.

Proposition 3.42 nojalla $\mathbf{p}$ voidaan esittää muodossa

$$\mathbf{p} = \mathbf{q}_1'\mathbf{q}_2' \dots \mathbf{q}_m',$$

missä $\mathbf{q}_i'$ on jaoton polynomi jokaisella $i = 1, \dots, m$. On selvä, että kaikilla $i = 1, \dots, m$ voidaan kirjoittaa $\mathbf{q}_i' = k_i\mathbf{q}_i$, missä $k_i \neq 0_K$ on polynomin $\mathbf{q}_i'$ johtava kerroin ja $\mathbf{q}_i = k^{-1}\mathbf{q}_i'$ on jaoton pääpolynomi. Tästä saadaan polynomille $\mathbf{p}$ esitys

$$\mathbf{p} = k \cdot \mathbf{q}_1\mathbf{q}_2 \dots \mathbf{q}_m,$$

missä $k = k_1k_2 \dots k_n \in K^*$ ja $\mathbf{q}_i$ on jaoton pääpolynomi jokaisella $i = 1, \dots, m$, $m \in \mathbb{N}$.

Osoitetaan esityksen yksikäsitteisyys. Oletetaan, että

$$k \cdot \mathbf{q}_1\mathbf{q}_2 \dots \mathbf{q}_m = \mathbf{p} = k' \cdot \mathbf{q}_1'\mathbf{q}_2 \dots \mathbf{q}_n',$$

missä $k, k' \in K^*$ ja $\mathbf{q}_i, \mathbf{q}_j'$ ovat jaottomia pääpolynomeja jokaisella $i = 1, \dots, m$ ja $j = 1, \dots, n$. Vasemmanpuoleisesta esityksestä saadaan polynomin $\mathbf{p}$ johtavaksi kertoimeksi skalaari k ja oikeanpuoleisesta vastaavasti k'. Näin ollen $k = k' \neq 0_K$. Koska K on kunta, voidaan jakaa alkiolla k, jolloin saadaan yhtälö

$$\mathbf{q}_1\mathbf{q}_2 \dots \mathbf{q}_m = \mathbf{q}_1'\mathbf{q}_2 \dots \mathbf{q}_n'. \tag{3.50}$$

Polynomi $\mathbf{q}_m$ on jaoton ja se jakaa oikeanpuoleisen tulon $\mathbf{q}_1'\mathbf{q}_2 \dots \mathbf{q}_n'$. Edellisen seurauksen mukaan tällöin $\mathbf{q}_m$ jakaa ainakin yhden polynomeista $\mathbf{q}_1', \mathbf{q}_2, \dots, \mathbf{q}_n'$. Koska näiden polynomien järjestyksellä ei ole merkitystä, voidaan esimerkiksi olettaa, että polynomi $\mathbf{q}_m$ jakaa polynomin $\mathbf{q}_n'$. Kuitenkin kumpikin näistä polynomeista on jaoton, joten tämä on mahdollista jos ja vain jos $\mathbf{q}_n' = k\mathbf{q}_m$ jollakin skalaarilla k. Koska $\mathbf{q}_m$ on pääpolynomi, polynomin $\mathbf{q}_n'$ johtava kerroin on tällöin k. Koska $\mathbf{q}_n'$ on myös pääpolynomi, tästä seuraa, että $k = 1$ eli $\mathbf{q}_n' = \mathbf{q}_m$.

Koska $K[\mathbf{X}]$ on kokonaisalue, yhtälön 3.50 molemmilta puolelta voidaan supistaa polynomi $\mathbf{q}_n' = \mathbf{q}_m$, jolloin saadaan yhtälö

$$\mathbf{q}_1\mathbf{q}_2 \dots \mathbf{q}_{m-1} = \mathbf{q}_1'\mathbf{q}_2 \dots \mathbf{q}_{n-1}'.$$

Jatketaan induktiivisesti samalla tavalla. Jokaisessa välivaiheessa yhtälön molemmalta puolilta supistuu yksi polynomi. Loppujen lopuksi äärellisen monen askeleen jälkeen päädytään tilanteeseen, jossa yhtälön toiselta puoleta supistuvat kaikki sen tekijät, eli sinne jää vain vakiopolynomi 1_K. Tällöin ei toisellekaan puolelle voi jäädä polynomeja, sillä muuten yhtälön toisella puolella esiintyisi vakio, eli nollaasteinen polynomi, kun taas toisella esiintyisi polynomi, jonka aste on vähintään yksi. Näin ollen pitää olla $m = n$. Lisäksi yllä olevasta seuraa, että yhtälön 3.50 molemmilla puolilla esiintyvät samat polynomit (mahdollisesti eri järjestyksissä). □

Polynomialgebran univeraaliominaisuus
Olkoon A jokin K-algebra ja olkoon $\mathbf{v} \in A$ tämän algebran alkio. Olkoon

$$\mathbf{p} = (a_i)_{i\in\mathbb{N}} = \sum_{i=0}^{n} a_i \mathbf{X}^i$$

kunnan K *polynomialgebran* $K[X]$ alkio (tässä $n = \deg \mathbf{p}$). Halutaan pystyä "sijoittamaan" algebrallisessa polynomissa $\mathbf{p}$ abstraktin "muuttujasymbolin" X paikalle algebran A vektori $\mathbf{v}$. Käytännössä tämä tarkoittaa sitä, että muodostetaan "polynomilauseke"

$$\mathbf{p}(\mathbf{v}) = \sum_{i=0}^{n} a_i \mathbf{v}^i.$$

Tämä lauseke on järkevä, sillä A on algebra, joten sen alkiolle $\mathbf{v}$ on määritelty algebrassa potenssit $\mathbf{v}^i$, $i \in \mathbb{N}$, nämä voidaan kertoa vasemmalta kunnan K alkioilla, eli skalaareilla a_i, ja lopuksi voidaan laskea mielivaltainen (äärellinen) määrä tällaisia termejä yhteen. Kun $\mathbf{v}$ käy läpi algebran A vektoreita, tällainen "sijoitusoperaatio" määrittelee algebran A *K-kertoimisen polynomikuvauksen* $p\colon A \to A$, $p(\mathbf{v}) = \sum_{i=0}^{n} a_i \mathbf{v}^i$. Pannaan erityisesti merkille, että abstraktia polynomia $\mathbf{p} \in K[\mathbf{X}]$ merkitään tässä materiaalissa yleensä lihavoidulla fontilla, kun taas sitä vastaavaa kuvausta $p\colon A \to A$ tavallisella fontilla. Näin voidaan erottaa kummasta on missäkin tapauksessa kyse. Jokainen abstrakti polynomialgebran $K[\mathbf{X}]$ alkio $\mathbf{p}$ siis määrittelee jokaisessa K-algebrassa A erään K-kertoimisen polynomikuvauksen p. Tämä kuvaus riippuu tietysti paitsi polynomista $\mathbf{p}$ myös algebrasta A.

Edellä tarkasteltu algebran alkion sijoituksen idea voidaan muotoilla hieman eri näkökulmasta myös tärkeän polynomialgebran $K[\mathbf{X}]$ *universaaliominaisuuden* muodossa.

Propositio 3.51. *Olkoon A K-algebra ja olkoon $\mathbf{v} \in A$ mielivaltainen. Tällöin on olemassa* tasan yksi *K-algebrojen välinen homomorfismi $S_{\mathbf{v}}\colon K[\mathbf{X}] \to A$ jolle pätee $S_{\mathbf{v}}(\mathbf{X}) = \mathbf{v}$. Tätä kuvausta kutsutaan* vektorin $\mathbf{v}$ määrittelemäksi sijoitushomomorfismiksi.

Lisäksi kaikilla $\mathbf{p} \in K[\mathbf{X}]$ pätee $S_{\mathbf{v}}(\mathbf{p}) = p(\mathbf{v})$.

Todistus. **Kuvauksen yksikäsitteisyys**
Olkoon $S_{\mathbf{v}}\colon K[\mathbf{X}] \to A$ algebrojen välinen homomorfismi, jolle pätee $S_{\mathbf{v}}(\mathbf{X}) = \mathbf{v}$. Osoitetaan induktiolla luvun $n \in \mathbb{N}$ suhteen, että kaikilla $n \in \mathbb{N}$ pätee

$$S_{\mathbf{v}}(\mathbf{X}^n) = \mathbf{v}^n. \tag{3.52}$$

Tapauksessa $n = 0$ yhtälö

$$S_{\mathbf{v}}(\mathbf{X}^0) = S_{\mathbf{v}}(\mathbf{e}_0) = \mathbf{1}_A = \mathbf{v}^0,$$

pätee, sillä S on algebrahomomorfismi, joten se kuvaa ykkösalkion ykkösalkioksi.

Tapauksessa $n = 1$ yhtälö

$$S_{\mathbf{v}}(\mathbf{X}^1) = S_{\mathbf{v}}(\mathbf{X}) = \mathbf{v}$$

pätee oletuksen nojalla. Oletetaan, että $S_{\mathbf{v}}(\mathbf{X}^n) = \mathbf{v}^n$ jollakin $n \geq 1$. Tällöin

$$S_{\mathbf{v}}(\mathbf{X}^{n+1}) = S_{\mathbf{v}}(\mathbf{X}^n)S_{\mathbf{v}}(\mathbf{X}) = \mathbf{v}^n\mathbf{v} = \mathbf{v}^{n+1}.$$

Väite (3.52) on todistettu induktiolla. Oletuksen nojalla $S_{\mathbf{v}}$ on erityisesti lineaarinen kuvaus K-vektoriavaruuksien $K[\mathbf{X}]$ ja A välillä. Lisäksi $\{\mathbf{X}^n \mid n \in \mathbb{N}\}$ on vektoriavaruuden $K[\mathbf{X}]$ kanta. Koska lineaarinen kuvaus määräytyy yksikäsitteisesti kannan alkioiden kuvista, tämä osoittaa sen, että kuvaus $S_{\mathbf{v}}$ on yksikäsitteinen. Lisäksi, jos

$$\mathbf{p} = (a_i)_{i\in\mathbb{N}} = \sum_{i=0}^{n} a_i\mathbf{X}^i$$

on polynomialgebran $K[\mathbf{X}]$ alkio, kaavan

$$S_{\mathbf{v}}(\mathbf{p}) = \sum_{i=0}^{n} a_i\mathbf{v}^i = p(\mathbf{v}) \tag{3.53}$$

on annettava kuvauksen $S_{\mathbf{v}}$ arvo pisteessä $\mathbf{p} \in K[\mathbf{X}]$. Tässä $p\colon A \to A$ on vastaava polynomikuvaus algebrassa A. Erityisesti $S_{\mathbf{v}}$ on yksikäsitteisesti määrätty.

Kuvauksen olemassaolo

Osoitetaan, että kaavalla 3.53 määritelty kuvaus $S_{\mathbf{v}}\colon K[\mathbf{X}] \to A$ on todellakin algebrojen välinen homomorfismi. Kuvaus $S_{\mathbf{v}}$ on siis yksikäsitteinen lineaarinen kuvaus K-vektoriavaruuksien $K[\mathbf{X}]$ ja A välillä, joka kuvaa kannan $\{\mathbf{X}^n \mid n \in \mathbb{N}\}$ alkioita säännöllä $S_{\mathbf{v}}(\mathbf{X}^n) = \mathbf{v}^n$. Lemman 2.57 yleistys äärettömän kannan tapauksessa implikoi, että tällainen kuvaus on olemassa (ja yksikäsitteinen). Lisäksi kuvauksen määritelmän nojalla pätee erityisesti

$$S_{\mathbf{v}}(\mathbf{e}_0) = S_{\mathbf{v}}(\mathbf{X}^0) = \mathbf{v}^0 = \mathbf{1}_A,$$

toisin sanoen $S_{\mathbf{v}}$ kuvaa algebran $K[\mathbf{X}]$ ykkösalkio algebran A ykkösalkiolle. Lisäksi

$$S_{\mathbf{v}}(\mathbf{X}) = S_{\mathbf{v}}(\mathbf{X}^1) = \mathbf{v}^1 = \mathbf{v}$$

eli $S_{\mathbf{v}}$ toteuttaa siltä vaadittavan ehdon.

Osoitetaan vielä että kuvaus $S_{\mathbf{v}}$ on myös yhteensopiva algebrojen $K[\mathbf{X}]$ ja A kertolaskuoperaatioiden kanssa. Toisin sanoen on näytettävä, että kaikilla $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ pätee

$$S_{\mathbf{v}}(\mathbf{pq}) = S_{\mathbf{v}}(\mathbf{p})S_{\mathbf{v}}(\mathbf{q}).$$

Tämän yhtälön kummallakin puolella esiintyviä lausekkeita voidaan ajatella kahden muuttujan kuvauksina $K[\mathbf{X}] \times K[\mathbf{X}] \to A$. Vasemmalla puolella kyse on kuvauksesta

$(\mathbf{p}, \mathbf{q}) \mapsto S_{\mathbf{v}}(\mathbf{pq})$ ja oikealla kuvauksesta $(\mathbf{p}, \mathbf{q}) \mapsto S_{\mathbf{v}}(\mathbf{p})S_{\mathbf{v}}(\mathbf{q})$. Koska $S_{\mathbf{v}}$ on lineaarinen ja kertolaskuoperaatiot kummassakin algebrassa ovat bilineaarisia, nähdään helposti, että kumpikin yllä mainittu kuvaus on bilineaarinen (tarkista yksityiskohdat). Lemman 2.122 nojalla riittää osoittaa, että nämä kuvaukset saavat samat arvot vektoriavaruuden $K[\mathbf{X}]$ kannan alkioilla. Olkoot $n, m \in \mathbb{N}$. Tällöin polynomialgebran kertolaskun määritelmän ja potenssisääntöjen nojalla saadaan

$$S_{\mathbf{v}}(\mathbf{e}_n\mathbf{e}_m) = S_{\mathbf{v}}(\mathbf{e}_{n+m}) = \mathbf{v}^{n+m} = \mathbf{v}^n\mathbf{v}^m = S_{\mathbf{v}}(\mathbf{e}_n)S_{\mathbf{v}}(\mathbf{e}_m).$$

Näin ollen

$$S_{\mathbf{v}}(\mathbf{pq}) = S_{\mathbf{v}}(\mathbf{p})S_{\mathbf{v}}(\mathbf{q})$$

kaikilla $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$. □

Kuten universaaliominaisuuksien kohdalla yleensäkin, voidaan osoittaa, että edellisessä propositiossa mainittu polynomialgebran $K[\mathbf{X}]$ universaaliominaisuus määrää parin $(K[\mathbf{X}], \mathbf{X})$ yksikäsitteisesti algebraisomorfismia vaille. Tätä väitettä ei formuloida eikä todisteta tässä sen tarkemmin.

Pannaan merkille, että sijoitushomomorfismissa $S_{\mathbf{v}} \colon K[\mathbf{X}] \to A$ kunnan alkio $\mathbf{v}$ pidetään vakiona ja muuttujana on polynomi $\mathbf{p} \in K[\mathbf{X}]$. Edellä tarkastellussa polynomin $\mathbf{p}$ määrämässä K-kertoimisessa polynomikuvauksessa $p \colon A \to A$ on taas käytössä päinvastainen näkökulma - siinä polynomi on kiinnitetty ja muuttujana on algebran alkio.

Sijoitushomomorfismin avulla polynomialgebraan liittyviä käsitteitä ja tuloksia voidaan soveltaa mielivaltaisessa K-algebrassa A.

Olkoon $\mathbf{p} \in K[\mathbf{X}]$ polynomi ja olkoon A jokin K-algebra. Olkoon $\mathbf{v} \in A$. Tällöin vektoria $\mathbf{v}$ sanotaan polynomin $\mathbf{p}$ *juureksi* (algebrassa A) jos

$$S_{\mathbf{v}}(\mathbf{p}) = p(\mathbf{v}) = \mathbf{0}_A.$$

Havainnollisesti katsoen $\mathbf{v}$ on siis polynomin juuri, jos polynomin "arvoksi" tulee nolla, kun siihen sijoitetaan symbolin $\mathbf{X}$ paikalle vektori $\mathbf{v}$. Jos $p = \sum_{i=0}^{n} a_i\mathbf{X}^i$, niin $\mathbf{v} \in A$ on polynomin $\mathbf{p}$ juuri jos ja vain jos algebrassa A pätee yhtälö $\sum_{i=0}^{n} a_i\mathbf{v}^i = \mathbf{0}_A$.

Koska kunta K voidaan aina ajatella K-algebrana luonnollisella tavalla, voidaan erityisesti puhua polynomialgebran alkion $\mathbf{p}$ juurista kunnassa K. Polynomin juuret ja niiden ominaisuudet riippuvat yleisesti ottaen siitä, missä algebrassa (tai kunnassa) niitä tarkastellaan. Seuraavassa tuloksessa tarkastellaan nimenomaan polynomin $\mathbf{p} \in K[\mathbf{X}]$ käyttäytymistä *kerroinkunnan* K suhteen.

Propositio 3.54. *Olkoon K kunta ja olkoon $\mathbf{p} \in K[\mathbf{X}]$, $\mathbf{p} \neq \mathbf{0}$.*

(1) Olkoon $k \in K$. Tällöin k on polynomin $\mathbf{p}$ juuri jos ja vain jos $\mathbf{p}$ on jaollinen polynomilla $(\mathbf{X} - k)$ (renkaassa $K[\mathbf{X}]$).

(2) Polynomilla $\mathbf{p}$ on korkeintaan *polynomin $\mathbf{p}$ asteen $\deg \mathbf{p}$ verran eri juurta kunnassa K.*

Todistus. (1) Sovelletaan polynomien jakoyhtälöä (Propositio 3.35) polynomeihin $\mathbf{p}$ ja $\mathbf{q} = X - k$. Sen nojalla on olemassa polynomit $\mathbf{s}, \mathbf{r} \in K[\mathbf{X}]$ siten, että $\mathbf{p} = \mathbf{sq} + \mathbf{r}$ ja $\deg \mathbf{r} < \deg \mathbf{q} = 1$. Viimeksi mainitusta ehdosta seuraa, että jakojäännös $\mathbf{r}$ on itse asiassa vakiopolynomi, $\mathbf{r} = k'$ jollakin $k' \in K$. Sijoitamalla yhtälöön $\mathbf{p} = \mathbf{sq} + k'$ kunnan K alkio k (eli, täsmällisesti sanottuna, laskemalla yhtälön kummankin puolen sijoitushomomorfismin $S_k \colon K[\mathbf{X}] \to K$ arvo) saadaan

$$S_k(\mathbf{p}) = p(k) = s(k)q(k) + k'.$$

Tässä

$$q(k) = S_k(\mathbf{X} - k) = S_k(X) - S_k(k) = k - k = 0_K.$$

Näin ollen päädytään yhtälöön $p(k) = k'$. Tästä seuraa, että $p(k) = 0$ (eli k on polynomin $\mathbf{p}$ juuri) jos ja vain jos jakojäännös $\mathbf{r} = k'$ on kunnan nolla-alkio. Jälkimmäinen ehto on taas ekvivalentti sen kanssa, että $\mathbf{p}$ on jaollinen polynomilla $\mathbf{q} = \mathbf{X} - k$.

(2) Väite voidaan osoittaa induktiolla polynomin asteen $\deg \mathbf{p}$ suhteen. Jos $\mathbf{p}$ on nollasta eroava vakiopolynomi, sillä ei ole juuria lainkaan (eli juurten lukumäärä on nolla) ja sen aste on 0. Väite siis pätee tässä tapauksessa.

Oletetaan, että väite pätee polynomeille, joiden aste on korkeintaan $n \geq 0$. Olkoon $\mathbf{p}$ polynomi, jonka aste on $(n + 1)$. Jos tällä polynomilla ei ole juuria, ei ole mitään todistettavaa. Muuten on olemassa $k \in K$ jolle pätee $p(k) = 0_K$. Kohdan (1) nojalla $\mathbf{p}$ on jaollinen polynomilla $\mathbf{q} = \mathbf{X} - k$, toisin sanoen $\mathbf{p} = (\mathbf{X} - k)\mathbf{s}$ jollakin polynomilla $\mathbf{s}$. Lemmasta 3.34 seuraa, että $\deg \mathbf{s} = \deg \mathbf{p} - 1$. Lisäksi $\mathbf{s}$ ei voi olla nolla-polynomi, sillä tällöin $\mathbf{p}$ olisi nolla-polynomi. Näin ollen, induktio-oletuksen nojalla, polynomilla $\mathbf{s}$ on korkeintaan $\deg \mathbf{s} = \deg \mathbf{p} - 1$ eilaista juurta. Olkoon k' mikä tahansa polynomin $\mathbf{p}$ juuri. Sijoittamalla se (sijoitushomomorfismin avulla) polynomiyhtälöön $\mathbf{p} = (\mathbf{X} - k)\mathbf{s}$ saadaan kunnassa K yhtälö

$$0_k = p(k') = (k' - k)s(k').$$

Koska K on kuntana erityisesti kokonaisalue, tästä seuraa (nollan sääntö), että joko $k' = k$ tai $s(k) = 0$. Toisin sanoen $k = k'$ tai k on polynomin $\mathbf{s}$ juuri. Koska induktioletuksen mukaan polynomilla $\mathbf{s}$ on korkeintaan $\deg \mathbf{p} - 1$ juurta, tästä seuraa, että polynomilla $\mathbf{p}$ on korkeintaan $\deg \mathbf{p}$ juurta. □

Olkoon $a \in K$ polynomin $\mathbf{p} \in K[X]$ juuri, $\mathbf{p} \neq 0$. Edellisen proposition mukaan on olemassa polynomi $\mathbf{q} \in K[\mathbf{X}]$ siten, että

$$\mathbf{p} = (\mathbf{X} - k)\mathbf{p}_1.$$

Tällöin välttämättä pätee $\deg \mathbf{p}_1 = \deg \mathbf{p} - 1$.

Voi käydä niin, että sama skalaari k on myös polynomin $\mathbf{s}$ juuri. Tällöin $\mathbf{p}_1 = (\mathbf{X} - k)\mathbf{p}_2$, joten

$$\mathbf{p} = (\mathbf{X} - k)^2\mathbf{p}_2.$$

Jatkamalla samalla tavalla niin kauan kunnes se ei ole enää mahdollista, lopulta päädytään esitykseen

$$\mathbf{p} = (\mathbf{X} - k)^l\mathbf{p}_l,$$

missä k ei enää ole polynomin $\mathbf{p}_l$ juuri. Luonnollista lukua $l \in \mathbb{N}$ sanotaan tällöin juuren k *kertaluvuksi* polynomissa $\mathbf{p}$. Juuren kertaluku on siis suurin luonnollinen luku l, jolla $\mathbf{p}$ on jaollinen polynomilla $(\mathbf{X}-k)^l$. On selvä, että juuren kertaluku on korkeintaan polynomin $\mathbf{p}$ aste.

Jos juuren kertaluku on 1, juurta sanotaan polynomin *yksinkertaiseksi* juureksi. Juurta, joka ei ole yksinkertainen, sanotaan *monikertaiseksi* juureksi.

Jos polynomialgebran $K[\mathbf{X}]$ jokaisella ei-vakio polynomilla on ainakin yksi juuri kunnassa K, kuntaa K sanotaan **algebrallisesti suljetuksi**. Tätä käsitetty ollaan käytetty aikaisemmin, aliluvussa 3.1, mutta silloin sen määrittelemiseksi käytettiin polynomikuvauksen käsitettä, ei abstraktia algebrallista polynomia. On kuitenkin selvää, että nämä kaksi määritelmää ovat ekvivalentteja. Aliluvussa 3.1 todettiin myös, että esimerkiksi kunnat $\mathbb{Q}$ ja $\mathbb{R}$ eivät ole algebrallisesti suljettuja, mutta kompleksilukujen kunta $\mathbb{C}$ on algebrallisesti suljettu. Viimeksi mainittua tärkeätä tulosta (Algebran peruslause) ei tällä kursilla todisteta, vaan sitä pidetään tunnettuna. Samoin pidetään tunnettuna Lauseen 3.17 tulosta, jonka mukaan jokainen kunta voidaan tarvittaessa "täydentää" algebrallisesti suljetuksi kunnaksi "lisäämällä" siihen "puuttuvia juuria". Näin saatua kunnan laajennusta sanotaan tämän kunnan algebralliseksi sulkeamaksi. Esimerkiksi kunnan $\mathbb{R}$ algebrallinen sulkeama on kompleksilukujen kunta $\mathbb{C}$.

Iteroimalla edellisen lemman tulosta (sekä käyttämällä kunnan ominaisuuksia) saadaan seuraava vahvempi versio edellisestä lemmasta (yksityiskohdat harjoitustehtävänä).

Seuraus 3.55. *Olkoon K kunta ja olkoon $\mathbf{p}$ polynomialgebran $K[\mathbf{X}]$ alkio. Olkoot $k_1, \ldots, k_m$ tämän polynomin kaikki eri juuret kunnassa K ja olkoon l_i juuren k_i kertaluku jokaisella $i = 1, \ldots, m$. Tällöin*

$$\mathbf{p} = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \ldots (\mathbf{X} - k_m)^{l_m}\mathbf{q},$$

missä $\mathbf{q}$ on sellainen polynomialgebran $K[\mathbf{X}]$ alkio, jolla ei ole lainkaan juuria kunnassa K.

Jos edellisessä tuloksessa K on algebrallisesti suljettu, polynomin $\mathbf{q}$ on oltava vakiopolynomi. Tästä saadaan seuraava tulos.

Seuraus 3.56. *Olkoon K* algebrallisesti suljettu *kunta ja olkoon $\mathbf{p}$ polynomialgebran $K[\mathbf{X}]$ alkio. Olkoot $k_1, k_2, \ldots, k_m$ tämän polynomin* kaikki *juuret kunnassa K ja olkoon l_i juuren k_i kertaluku jokaisella $i = 1, \ldots, m$. Tällöin*

$$\mathbf{p} = a(\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \ldots (\mathbf{X} - k_m)^{l_m}$$

missä $a \in K$ on polynomin $\mathbf{p}$ johtava kerroin.

Sama tulos pätee, vaikka K ei olisi algebrallisesti suljettu, jos polynomin juurten lukumäärä (kertalukuineen laskettuna) on sama kuin polynomin aste. Tarkemmin sanottuna pätee seuraava tulos, jonka voi helposti johtaa Seurauksesta 3.55.

Seuraus 3.57. *Olkoon K kunta ja olkoon $\mathbf{p}$ polynomialgebran $K[\mathbf{X}]$ alkio. Olkoot $k_1, k_2, \ldots, k_m$ tämän polynomin* kaikki *juuret kunnassa K ja olkoon l_i juuren k_i kertaluku jokaisella $i = 1, \ldots, m$. Oletetaan, että*

$$l_1 + l_2 + \ldots + l_m = \deg \mathbf{p}.$$

Tällöin

$$\mathbf{p} = a(\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \dots (\mathbf{X} - k_m)^{l_m}$$

missä $a \in K$ *on polynomin* $\mathbf{p}$ *johtava kerroin.*

Esimerkkejä 3.58. *1) Mikä tahansa ensimmäisen asteen polynomi* $\mathbf{p} = a\mathbf{X} + b \in K[\mathbf{X}]$, $a, b \in K$, $a \neq 0_K$, *on jaoton. Lisäksi sillä on aina tasan yksi juuri* $k = -ba^{-1}$.

Jos K *on algebrallisesti suljettu kunta, ainoat jaottomat polynomialgebran alkiot ovat ensimmäisen asteen polynomeja. Tämä seuraa suoraan Seurauksesta 3.56. Myös käänteinen väite pätee - jos ainoastaan ensimmäisen asteen polynomit ovat jaottomia, kunta on algebrallisesti suljettu.*

2) Polynomi $\mathbf{X}^2 + 1 \in \mathbb{R}[\mathbf{X}]$ *on jaoton. Tämä nähdään seuraavasti. Oletetaan, että* $\mathbf{X}^2 + 1 = \mathbf{pq}$, *missä* $\deg \mathbf{p}, \deg \mathbf{q} < 2, \deg \mathbf{p} + \deg \mathbf{p} = 2$. *Tällöin ainoa mahdollisuus on* $\deg \mathbf{p} = \deg \mathbf{q} = 1$. *Edellisen esimerkin nojalla kummallakin polynomilla* $\mathbf{p}$, $\mathbf{q}$ *on juuri. Selvästi kummankin polynomin* $\mathbf{p}$ *ja* $\mathbf{q}$ *juuri on myös polynomin* $\mathbf{X}^2 + 1$ *juuri. Kuitenkin polynomilla* $\mathbf{X}^2 + 1$ *ei ole juurta reaalilukukunnassa, sillä ei ole olemassa reaalilukua* $x \in \mathbb{R}$, *jolle pätee* $x^2 + 1 = 0$. *Näin ollen polynomi* $\mathbf{X}^2 + 1 \in \mathbb{R}[\mathbf{X}]$ *on jaoton.*

Kompleksilukujen joukko $\mathbb{C}$ *on* $\mathbb{R}$*-algebra. Samalla polynomilla* $\mathbf{X}^2 + 1 \in \mathbb{R}[\mathbf{X}]$ *on* $\mathbb{R}$*-algebrassa* $\mathbb{C}$ *tasan kaksi juurta, kompleksiluvut* i *ja* $-i$.

Kompleksilukujen joukko $\mathbb{C}$ *on myös kunta itse. Sen polynomialgebra* $\mathbb{C}[\mathbf{X}]$ *sisältää "samannäköisen alkion"* $\mathbf{X}^2 + 1 \in \mathbb{C}[\mathbf{X}]$. *Tällä polynomilla on kunnassa* $\mathbb{C}$ *tasan kaksi juurta* i *ja* $-i$. *Koska* $\mathbb{C}$ *on algebrallisesti suljettu, Seurauksen 3.56 nojalla saadaan hajotelma* $\mathbf{X}^2 + 1 = (\mathbf{X} - i)(\mathbf{X} + i)$. *Sama tulos voidaan helposti johtaa myös Seurauksesta 3.57, jos ei halua turhaan turvautua siihen tosiasiaan, että kunta* $\mathbb{C}$ *on algebrallisesti suljettu.*

3) Samalla tavalla kuin edellisessä esimerkissä on tehty polynomin $\mathbf{X}^2 + 1 \in \mathbb{R}[\mathbf{X}]$ *kohdalla, voidaan osoittaa, että astetta* 2 *tai* 3 *oleva polynomi on jaoton jos ja vain jos sillä ei ole juuria.*

Sen sijaan kun $\deg \mathbf{p} > 3$ *tämä ekvivalenssi ei enää pidä paikaansa. On totta, että jos polynomilla on juuri ja sen aste on suurempi kuin kaksi, niin polynomi ei ole jaoton. Tämä seuraa suoraan sitä, että polynomi on tällöin jaollinen ensimmäistä astetta olevalla polynomilla* $\mathbf{X} - k$ *(missä* k *on polynomin juuri). Kääntäinen väite ei välttämättä kuitenkaan päde yleisesti. Esimerkiksi kunnan* $\mathbb{R}$ *yli polynomi* $\mathbf{X}^4 + 1$ *ei ole jaoton, sillä*

$$\mathbf{X}^4 + 1 = (\mathbf{X}^2 + \sqrt{2}\mathbf{X} + 1)(\mathbf{X}^2 - \sqrt{2}\mathbf{X} + 1).$$

Kuitenkin polynomilla $\mathbf{X}^4 + 1$ *ei ole juuria kunnassa* $\mathbb{R}$.

4) Käytämällä hyväksi sitä, että kunta $\mathbb{C}$ *on algebrallisesti suljettu, voidaan osoittaa, että polynomialgebrassa* $\mathbb{R}[\mathbf{X}]$ *jaottomat polynomit ovat tasan kaikki ensimmäisen asteen polynomit ja sellaiset toisen asteen polynomit, joilla ei ole reaalijuuria (todistus harjoitustehtävänä). Tästä seuraa (Lemma 3.42), että mikä tahansa ei-vakio*

polynomi algebrassa $\mathbb{R}[\mathbf{X}]$ *voidaan kirjoittaa ensimmäisen ja toisen asteen polynomien tulona.*

Tästä voidaan helposti johtaa, että paritonta astetta olevalla polynomialgebran $\mathbb{R}[X]$ *polynomilla on ainakin yksi reaalinen juuri (koska sillä täytyy tällöin olla ainakin yksi ensimmäisen asteen tekijä). Sama tulos voidaan todistaa myös käytämällä klassisen reaalianalyysin menetelmiä. Nimittäin olkoon* $\deg \mathbf{p}$ *pariton,* $\mathbf{p} \in \mathbb{R}[\mathbf{X}]$*. Tällöin (kuten analyysin kurssilla todistetaan)* $\mathbf{p}$ *kasvaa rajatta, kun lähestytään toista äärettömyyttä, ja pienenee rajatta, kun lähestytään toista äärettömyyttä. Erityisesti polynomi* $\mathbf{p}$ *saa sekä positiivisia, että negatiivisia arvoja. Toisaalta polynomi on jatkuva funktio, joten Bolzanon lauseesta seuraa tällöin, että jossakin pisteessä se saa arvon* 0*.*

5) *Esimerkissä 2.74, (4) on tarkasteltu* kvaternioiden agebraa $\mathbb{H}$*. Tämä on* 4*-ulotteinen* $\mathbb{R}$*-algebra, jolla on kanta* $(1, i, j, k)$*, missä* 1 *on kertolaskun neutraalialkio (joka voidaan samaistaa reaaliluvun* 1 *kanssa) ja imaginääriyksiköt* i, j, k *toteuttavat (muun muassa) yhtälöt* $i^2 = j^2 = k^2 = -1$*. Tästä seuraa, että* $\mathbb{R}$*-kertoimisella kaksiasteisella polynomilla* $\mathbf{X}^2 + 1 \in \mathbb{R}[\mathbf{X}]$ *on* $\mathbb{R}$*-algebrassa* $\mathbb{H}$ *ainakin kolme ratkaisua. Itse asiassa voidaan osoittaa, että tällä polynomilla on kvarternioiden algebrassa jopa äärettömän monta ratkaisua. Näin ollen Lemman 3.54 väite (2) ei ole yleisesti voimassa* K*-algebralle. Huomaa, että kvaternionien algebra on renkaana jopa vinokunta, eli se toteuttaa kaikki kunnan aksioomat, paitsi kertolaskun vaihdannaisuuden. Näin ollen Lemman 3.54 väitteet eivät päde jopa vinokunnissa.*

Tekijäalgebrat ja algebrojen isomorfialause

Poiketaan hetkellisesti juonen yleisestä kulusta puhumaan lyhyesti tekijäalgebran käsitteestä, sillä sitä tarvitaan algebrallisten alkioiden teorian yhteydessä.

Puhuttaessa jostakin algebrallisesta struktuurista, nousevat luonnollisiksi kysymykset siitä, miten määritellään tämäntyyppisten struktuurien väliset (homo)morfismit, alistruktuurit sekä miten saadaan konstruoitua tekijästruktuurit. Olkoot A ja B K-algebrat. Kuvaus $f\colon A \to B$ on K-algebrojen välinen homomorfismi, jos se on K-lineaarinen kuvaus A:n ja B:n K-vektoriavaruusstruktuurien suhteen ja säilyttää kertolaskun, toisin sanoen, jos f on lisäksi rengashomomorfismi $f\colon (A, +, \cdot) \to (B, +, \cdot)$. Bijektiivistä algebrojen välistä homomorfismia sanotaan algebrojen isomorfismiksi. K-algebran A osajoukkoa A' sanotaan algebran A alialgebraksi, jos se on K-vektoriavaruuden A aliavaruus ja renkaan A alirengas. Algebran A *ideaaliksi* sanotaan sellaista sen osajoukkoa, joka on renkaan A ideaali. Helposti nähdään, että mikä tahansa algebran A ideaali on myös sen *aliavaruus*, kun A ajatellaan vektoriavaruutena. Käänteinen väite ei päde - aliavaruuten ei tarvitse olla ideaali.

Esimerkki 3.59. *Proposition 3.39 nojalla polynomirenkaan* $K[\mathbf{X}]$ *jokainen rengasideaali on pääideaali muotoa* $(\mathbf{p})$ *jollakin* $\mathbf{p} \in K[\mathbf{X}]$*. Helposti nähdään laskemalla suoraan, että joukko* $(\mathbf{p})$ *on myös vektoriavaruuden* $K[\mathbf{X}]$ *aliavaruus. Näin pitääkin olla yleisen teorian mukaan.*

Olkoon I algebran A ideaali. Tällöin, jos I ajatellaan vektoriavaruuden A aliavaruutena, voidaan muodostaa *tekijäavaruus* A/I (kts. osio 2.1). Tämän vektoriavaruuden alkiot ovat ekvivalenssiluokkia muotoa $\mathbf{a} + I$, $\mathbf{a} \in A$. Kaksi algebran A alkiota $\mathbf{a}, \mathbf{b} \in A$

määritelevät saman ekvivalenssiluokan, eli

$$\mathbf{a} + I = \mathbf{b} + I$$

pätee jos ja vain jos $\mathbf{a} - \mathbf{b} \in I$. Kanoninen projektio $p\colon A \to A/I$, $p(\mathbf{v}) = \mathbf{v} + I$ on (surjektiivinen) lineaarinen kuvaus.

Toisaalta, jos algebran ideaali I ajatellaan renkaan A ideaalina, voidaan muodostaa *tekijärengas* A/I (kts. osio 1.8). Määritelmän mukaan tämä rengas on joukkona sama kuin edellisessä kappaleessa tarkasteltu tekijavaruus A/I, eli koostuu ekvivalenssiluokista $\mathbf{a} + I$, $\mathbf{a} \in A$. Kanoninen projektio $p\colon A \to A/I$, $p(\mathbf{v}) = \mathbf{v} + I$, on (surjektiivinen) rengashomormofismi.

Yhdistämällä nämä tulokset, nähdään, että A/I on K-algebra (kertolaskun bilineaarisuuden tarkistus jätetään lukijalle). Ekvivalenssiluokilla lasketaan algebrassa A/I kuin A:n alkioilla, toisin sanoen tekijäalgebrassa A/I pätevät laskusäännöt

$$(\mathbf{v} + I) + (\mathbf{w} + I) = (\mathbf{v} + \mathbf{w}) + I,$$

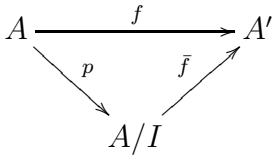
$$k(\mathbf{v} + I) = (k\mathbf{v}) + I \text{ ja}$$

$$(\mathbf{v} + I) \cdot (\mathbf{w} + I) = \mathbf{v}\mathbf{w} + I,$$

joita voi pitää tekijäalgebran A/I laskutoimitusten määritelminä. Kanoninen projektio $p\colon A \to A/I$, $p(\mathbf{v}) = \mathbf{v} + I$ on (surjektiivinen) algebrojen välinen homomorfismi.

Tekijäalgebroille pätevät samantyyppiset hajotelma-ja isomorfialauseet kuin muidenkin tekijästruktuurien kohdalla. Ne voidaan todistaa yhdistämällä vastaavat tulokset tekijärenkaille (Lauseet 1.101 ja 1.102) ja tekijäavaruuksille (Lauseet 2.24 ja 2.25).

Lause 3.60. Algebrahomomorfismien hajotelmalause
Olkoon $f\colon A \to A'$ algebrojen välinen homomorfismi. Olkoon I algebran A ideaali ja olkoon $p\colon A \to A/I$ kanoninen projektio tekijäalgebralle. Tällöin on olemassa indusoitu algebrahomomorfismi $\bar{f}\colon A/I \to A'$ jolle pätee $f = \bar{f} \circ p$,



jos ja vain jos $I \subset \operatorname{Ker} f$.
Jos tällainen kuvaus $\bar{f}$ on olemassa, niin se on yksikäsitteinen. Lisäksi pätee $\operatorname{Im} \bar{f} = \operatorname{Im} f$. Erityisesti $\bar{f}$ on surjektio jos ja vain jos f on surjektio.

Lisäksi $\bar{f}$ on injektio jos ja vain jos $I = \operatorname{Ker} f$.

Lause 3.61. Algebrojen isomorfialause *Olkoon $f\colon A \to A'$ algebrojen välinen homomorfismi. Tällöin $\operatorname{Ker} f$ on algebran A ideaali ja indusoitu kuvaus $\bar{f}\colon A/\operatorname{Ker} f \to \operatorname{Im} f$ on algebrojen välinen isomorfismi.*

Algebralliset alkiot.
Olkoon A K-algebra ja olkoon $\mathbf{a} \in A$. Alkiota $\mathbf{a}$ sanotaan algebran A *algebralliseksi* alkioksi, jos on olemassa polynomi $\mathbf{p} \in K[\mathbf{X}], \mathbf{p} \neq \mathbf{0}$, siten, että $p(\mathbf{a}) = \mathbf{0}_A$. Alkiota $\mathbf{a} \in A$,

joka ei ole algebrallinen, sanotaan *traskendentiksi*.

Historiallisesti algebrallisen ja transkedentin alkion käsitteet otettiin ensimmäisenä käyttöön $\mathbb{Q}$-algebran tapauksessa $\mathbb{C}$. Tämä erikoistapaus on edelleenkin hyvin tärkeä. Kompleksilukua z sanotaan algebralliseksi (kunnan $\mathbb{Q}$ suhteen), jos z on jonkun nollasta eroavan $\mathbb{Q}$-kertoimisen polynomin $\mathbf{p}$ juuri. Helposti nähdään (laaventamalla kaikki polynomin kertoimet samannimisiksi), että tällöin z on jopa jonkun *kokonaislukukertoimisen* nollasta eroavan polynomin $\mathbf{q}$ juuri. Jokainen rationaaliluvuista ja imaginaariyksiköstä i neljän peruslaskutoimituksen ja juurten $\sqrt[n]{}$ avulla rakenettu kompleksiluku on algebrallinen, mutta on olemassa myös algebrallisia lukuja, joita ei voi esittää tällaisessa muodossa. Tämä on seuraus kuuluisasta Galois'n teoriasta, jota tarkastellaan Algebra II kurssilla. Tunnettuja esimerkkejä transkendenteista kompleksiluvuista ovat luku π ja Neperin luku e. Jonkun tietyn luvun osoittaminen transkendentiksi on yleensä suhteellisen vaikeaa, joten transkendentteista luvuista on hankala antaa konkreettisia esimerkkejä. Sen sijaan on helppoa näyttää, että tällaisia lukuja on varmasti olemassa. Nimittäin $\mathbb{Q}$-kertoimisia polynomeja on vain numeroituva määrä ja jokaisella sellaisella polynomilla on (Lemma 3.54) äärellinen määrä juuria kompleksilukujen kunnassa. Tästä seuraa, että algebrallisia kompleksilukuja on vain numeroituvan paljon. Koska kompleksilukujen joukko on ylinumeroituva, tästä seuraa, että "suurin osa" kompleksilukuja on transkendentteja.

Puetaan algebrallisen/transkendentin alkion määritelmä abstraktimpaan muotoon. Olkoon A K-algebra ja olkoon $\mathbf{a} \in A$. Olkoon $S_{\mathbf{a}} \colon K[\mathbf{X}] \to A$ sijoitushomomorfismi, joka kuvaa muuttujasymbolin $\mathbf{X}$ alkiolle $\mathbf{a}$ (Propositio 3.51). Määritelmän mukaan kuvaus $S_{\mathbf{a}}$ on algebrojen välinen homomorfismi, joten siihen voidaan soveltaa algebrojen isomorfialausetta 3.61. Tämä tulos sanoo, että sijoitushomomorfismi $S_{\mathbf{a}}$ indusoi *algebrojen välisen isomorfismin*

$$\widetilde{S_{\mathbf{a}}} \colon K[\mathbf{X}]/\operatorname{Ker} S_{\mathbf{a}} \cong \operatorname{Im} S_{\mathbf{a}}.$$

Tässä sijoitushomorfismin ydin

$$\operatorname{Ker} S_{\mathbf{a}} = \{\mathbf{p} \in K[\mathbf{X}] \mid p(\mathbf{a}) = \mathbf{0}_A\}$$

koostuu täsmälleen niistä polynomeista $\mathbf{p} \in K[\mathbf{X}]$, joiden eräänä juurena on alkio $\mathbf{a}$. Sijoitushomomorfismin kuva

$$\operatorname{Im} S_{\mathbf{a}} = \{p(\mathbf{a}) \mid \mathbf{p} \in K[\mathbf{X}]\}$$

taas on algebran A alialgebra ja koostuu täsmälleen kaikista K-kertoimisista "polynomilausekkeista", joissa "muuttujan roolissa" on $\mathbf{a}$. Toisin sanoen

$$\operatorname{Im} S_{\mathbf{a}} = \{c_n\mathbf{a}^n + c_{n-1}\mathbf{a}^{n-1} + \ldots + c_1\mathbf{a} + c_0 \mid c_0, c_1, \ldots, c_n \in K, n \in \mathbb{N}\}.$$

Merkitsemme tätä joukkoa myös symbolilla $K[\mathbf{a}]$. Ei ole vaikeata nähdä, että $K[\mathbf{a}]$ on (sisältyvyysrelaation suhteen) *pienin* algebran A alialgebra, joka sisältää alkion $\mathbf{a}$. Tästä syystä alialgebraa $K[\mathbf{a}]$ sanotaan myös alkion $\mathbf{a}$ *virittämäksi* algebran A alialgebraksi.

Vaihtoehto 1: Alkio $\mathbf{a} \in A$ on transkendentti. Tämä on yhtäpitävää sen kanssa, että sijoitushomomorfismin $S_{\mathbf{a}}$ ydin $\operatorname{Ker} S_{\mathbf{a}}$ koostuu ainoastaan nolla-polynomista. Sijoitushomorfismi $S_{\mathbf{a}}$ on tällöin algebraisomorfismi polynomialgebran $K[\mathbf{X}]$ ja algebran $K[\mathbf{a}]$

välillä. Algebran transkendentti alkio siis käyttäytyy algebrallisesti täsmälleen samalla tavalla kuin "geneerinen muuttujasymboli" X.

Vaihtoehto 2: Alkio $\mathbf{a} \in A$ on algebrallinen. Tällöin sijoitushomomorfismin $S_{\mathbf{a}}$ ydin $\operatorname{Ker} S_{\mathbf{a}}$ sisältää muutakin kuin nollapolynomin. Tämä ydin on algebran $K[\mathbf{X}]$ ideaali. Proposition 3.39 nojalla tämä ideaali on pääideaali, eli on jonkun polynomin $\mathbf{q} \in K[\mathbf{X}]$ virittämä, $\operatorname{Ker} S_{\mathbf{a}} = (\mathbf{q})$. Isomorfialauseen 3.61 nojalla algebra $K[\mathbf{a}]$ on algebrana isomorfinen *tekijäalgebran* $K[\mathbf{X}]/(\mathbf{q})$ kanssa. Koska ideaali $(\mathbf{q})$ ei ole tässä tapauksessa triviaali, $\mathbf{q} \neq \mathbf{0}$, voidaan jopa vaatia, että $\mathbf{q}$ on *pääpolynomi*. Tällöin $\mathbf{q}$ määräytyy yksikäsitteisesti ja sitä sanotaan algebran A algebrallisen alkion $\mathbf{a}$ *minimipolynomiksi*. Minimipolynomi on siis asteeltaan *pienin* nollasta eroava polynomi $\mathbf{q}$, jolle pätee $q(\mathbf{a}) = \mathbf{0}_A$. Lisäksi, jos $\mathbf{p} \in K[\mathbf{X}]$ on mikä tahansa polynomi, niin yhtälö $p(\mathbf{a}) = \mathbf{0}_A$ pätee jos ja vain jos $\mathbf{p}$ on *jaollinen* alkion $\mathbf{a}$ minimipolynomilla $\mathbf{q}$.

Esimerkkejä 3.62. *1) Tarkastellaan $\mathbb{Q}$-algebran $\mathbb{C}$ alkiota i. Tämä on algebrallinen, koska $i^2 + 1 = 0$. Itse asiassa polynomi $\mathbf{X}^2 + 1 \in \mathbb{Q}[\mathbf{X}]$ on alkion i* minimipolynomi. *Edellisen tarkastelun nojalla algebra $\mathbb{Q}[i]$ on isomorfinen tekijäalgebran $\mathbb{Q}[\mathbf{X}]/(\mathbf{X}^2+1)$ kanssa. Propositiosta 3.63 alla seuraa, että algebra $\mathbb{Q}[i]$ on 2-ulotteinen ($\mathbb{Q}$-vektoriavaruutena) ja*

$$\mathbb{Q}[i] = \{a + ib \mid a, b \in \mathbb{Q}\}.$$

(2) Korvataan edellisessä esimerkissä kunta $\mathbb{Q}$ reaalilukujen kunnalla $\mathbb{R}$. Tällöin analogiset tulokset pätevät, alialgebra $\mathbb{R}[i]$ on sellainen 2-ulotteinen $\mathbb{R}$-vektoriavaruus, joka on isomorfinen tekijäalgebran $\mathbb{R}[\mathbf{X}]/(\mathbf{X}^2 + 1)$ kanssa, ja

$$\mathbb{R}[i] = \{x + iy \mid x, y \in \mathbb{R}\}.$$

Tästä yhtälöstä kuitenkin seuraa heti, että $\mathbb{R}[i] = \mathbb{C}$. Näin ollen $\mathbb{R}$-algebrana (erityisesti renkaana) kompleksilukujen kunta $\mathbb{C}$ on isomorfinen polynomialgebran tekijäalgebran $\mathbb{R}[\mathbf{X}]/(\mathbf{X}^2 + 1)$ kanssa. Tästä saadaan uusi tapa konstruoida/määritellä kompleksiluvut!

(3) Edellisessä esimerkissä näytetään, miten polynomialgebran tekijäalgebrojen avulla voidaan konstruoida kuntalaajennuksia. *Tarkemmin sanottuna olkoon K kunta ja olkoon $\mathbf{p} \in K[\mathbf{X}]$* jaoton *polynomi, jonka aste on suurempi kuin yksi. Tällöin sillä ei ole juuria kunnassa K, mutta tekijäalgebrassa $K[\mathbf{X}]/(\mathbf{p})$ muuttujasymboli $\mathbf{X}$ on polynomin $\mathbf{p}$ juuri. Lisäksi, koska $\mathbf{p}$ on jaoton, Propositiosta 3.63 alla seuraa, että tekijäalgebra $K[\mathbf{X}]/(\mathbf{p})$ on silloin kunta. Kunta K voidaan ajatella tämän kunnan alikuntana. Näin ollen tämän konstruktion avulla voidaan laajentaa kunta K kunnaksi, jossa polynomilla $\mathbf{p}$ on juuri. Iteroimalla ja yleistämällä tämä konstruktio tietyllä ovelalla tavalla voidaan konstruoida sellainen kunnan K kuntalaajennus, jossa jokaisella polynomilla on juuri, toisin sanoen osoittaa Propositiossa 3.17 mainittu tulos todeksi. Yksityiskohtiin tutustutaan esimerkiksi Algebra II-kurssilla.*

Algebran algebrallisen alkion virittämä alialgebra $K[\mathbf{a}]$ on edellisten tarkastelujen nojalla isomorfinen sellaisen polynomialgebran tekijäalgebran kanssa, joka on muotoa $K[\mathbf{X}]/(\mathbf{q})$. Seuraavaksi tutkitaan tällaista muotoa olevan tekijäalgebran rakennetta tarkemmin.

Propositio 3.63. *Olkoon K kunta, ja olkoon $\mathbf{p}$ polynomialgebran $K[\mathbf{X}]$ nollapolynomista eroava alkio. Olkoon $n = \deg \mathbf{p} \in \mathbb{N}$. Tällöin tekijäalgebra $K[\mathbf{X}]/(\mathbf{p})$ on K-vektoriavaruutena n-ulotteinen vektoriavaruus. Sen eräs kanta on*

$$(\bar{\mathbf{1}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}^2, \ldots, \bar{\mathbf{X}}^{n-1}).$$

Lisäksi seuraavat ehdot ovat yhtäpitäviä.

(1) Tekijäalgebra $K[\mathbf{X}]/(\mathbf{p})$ on renkaana kunta.

(2) Tekijäalgebra $K[\mathbf{X}]/(\mathbf{p})$ on renkaana kokonaisalue.

(3) Polynomi $\mathbf{p}$ on jaoton.

Todistus. Olkoon $\mathbf{p} = \sum_{i=0}^{n} a_i \mathbf{X}^i$ n-asteinen polynomi, $\mathbf{p} \neq 0$. Osoitetaan, että jono $(\bar{\mathbf{1}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}, \ldots, \bar{\mathbf{X}}^{n-1})$ on K-vektoriavaruuden $K[\mathbf{X}]/(\mathbf{p})$ kanta.

Aloitetaan näyttämällä, että kyseinen jono on vapaa. Olkoot $b_0, b_1, \ldots, b_{n-1} \in K$ sellaisia, että

$$b_{n-1}\bar{\mathbf{X}}^{n-1} + \ldots + b_1\bar{\mathbf{X}} + b_0 = \overline{\mathbf{0}} \in K[\mathbf{X}]/(\mathbf{p}). \tag{3.64}$$

Määritellään polynomi $\mathbf{q} \in K[\mathbf{X}]$ kaavalla

$$\mathbf{q} = \sum_{i=0}^{n-1} b_i \mathbf{X}^i.$$

Tällöin ehto 3.64 tarkoittaa sitä, että tekijäavaruudessa $K[\mathbf{X}]/(\mathbf{p})$ pätee $\overline{\mathbf{q}} = \overline{\mathbf{0}}$. Tekijäavaruuden määritelmän mukaan tämä tarkoittaa sitä, että polynomi $\mathbf{q}$ kuuluu pääideaaliin $(\mathbf{p})$. Tämän pääideaalin määritelmän mukaan tämä taas tarkoittaa sitä, että on olemassa polynomi $\mathbf{s} \in K[\mathbf{X}]$ siten, että $\mathbf{q} = \mathbf{sp}$. Oletetaan, että $\mathbf{s} \neq \mathbf{0}$. Tällöin tämän yhtälön vasemmalla puolella on polynomi, jonka aste on korkeintaan $(n-1)$, kun taas oikealla puolella on polynomi, jonka aste on ainakin $\mathbf{s} + \mathbf{p} \geq n$. Saadaan siis ristiriita. Näin ollen $\mathbf{s} = \mathbf{0}$, joten myös $\mathbf{q} = \mathbf{0}$. Tämä on puolestaan yhtäpitävä sen kanssa, että $b_i = 0_K$ kaikilla $i = 0, \ldots, n-1$. Tästä seuraa, että jono $(\bar{\mathbf{1}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}^2, \ldots, \bar{\mathbf{X}}^{n-1})$ on vapaa.

Seuraavaksi osoitetaan, että tämä jono virittää tekijäavaruuden $K[\mathbf{X}]/(\mathbf{p})$. Olkoon $\mathbf{q}$ mielivaltainen polynomialgebran $K[\mathbf{X}]$ alkio. Jakoyhtälön (Proposition 3.35) nojalla mukaan on olemassa polynomit $\mathbf{s}, \mathbf{r} \in K[\mathbf{X}]$, siten, että

$$\mathbf{q} = \mathbf{sp} + \mathbf{r} \text{ ja } \deg \mathbf{r} \leq n - 1.$$

Tästä seuraa, että tekijäavaruudessa pätee

$$\overline{\mathbf{q}} = \overline{\mathbf{s}}\ \overline{\mathbf{p}} + \overline{\mathbf{r}} = \overline{\mathbf{r}}.$$

Koska $\deg \mathbf{r} \leq n - 1$, on olemassa $b_0, b_1, \ldots, b_{n-1}$ siten, että

$$\mathbf{r} = b_{n-1}\mathbf{X}^{n-1} + \ldots + b_1\mathbf{X} + b_0.$$

Näin ollen

$$\overline{\mathbf{q}} = \overline{\mathbf{r}} = b_{n-1}\bar{\mathbf{X}}^{n-1} + \ldots + b_1\bar{\mathbf{X}} + b_0.$$

Koska $\overline{\mathbf{q}}$ on mielivaltainen tekijäavaruuden $K[\mathbf{X}]/(\mathbf{p})$ alkio, on osoitettu, että jono $(\bar{\mathbf{1}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}^2, \ldots, \bar{\mathbf{X}}^{n-1})$ virittää tekijäavaruuden $K[\mathbf{X}]/(\mathbf{p})$.

Muiden Proposition 3.63 väitteiden todistaminen jätetään harjoitustehtäväksi. □

Esimerkki 3.65. *Olkoot* $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ *kaksi erilaista* n*-asteista polynomia. Edellisen proposition nojalla tekijäavaruudet* $K[\mathbf{X}]/(\mathbf{p})$ *ja* $K[\mathbf{X}]/(\mathbf{q})$ *ovat molemmat* n*-ulotteisia* K*-vektoriavaruuksia, erityisesti isomorfisia vektoriavaruuksina. Kuitenkin täytyy muistaa, että nämä avaruudet ovat myös renkaita ja renkkaina ne saattavat olla hyvin erilaisia!*

Esimerkiksi olkoot $\mathbf{p} = \mathbf{X}^2 + 1$ *ja* $\mathbf{q} = \mathbf{X}^2 \in \mathbb{R}[\mathbf{X}]$*. Tällöin* $K[\mathbf{X}]/(\mathbf{p})$ *ja* $K[\mathbf{X}]/(\mathbf{q})$ *ovat molemmat* 2*-ulotteisia* $\mathbb{R}$*-vektoriavaruuksia. Tekijärengas* $K[\mathbf{X}]/(\mathbf{p})$ *on isomorfinen kompleksilukujen kunnan* $\mathbb{C}$ *kanssa, erityisesti se on kokonaisalue. Tekijärengas* $K[\mathbf{X}]/(\mathbf{q})$ *taas ei ole edes kokonaisalue, sillä siinä pätee* $\overline{\mathbf{X}}^2 = \mathbf{0}$*.*

Seuraus 3.66. *Olkoon* A K*-algebra ja olkoon* $\mathbf{a} \in A$*. Tällöin* $\mathbf{a}$ *on algebrallinen jos ja vain jos sen virittämä alialgebra* $K[\mathbf{a}]$ *on* K*-vektoriavaruutena äärellisulotteinen.*

Jos algebra A *on* K*-vektoriavaruutena äärellisulotteinen, niin jokainen sen alkio on algebrallinen.*

Todistus. Jälkimmäinen väite seuraa edellisestä, sillä $K[\mathbf{a}]$ on vektoriavaruuden A aliavaruus.

Olkoon $\mathbf{a}$ algebrallinen. Tällöin $K[\mathbf{a}]$ on isomorfinen tekijäavaruuden $K[\mathbf{X}]/(\mathbf{q})$ kanssa, missä $\mathbf{q} \neq \mathbf{0}$ on alkion $\mathbf{a}$ minimipolynomi. Edellisen proposition nojalla $K[\mathbf{X}]/(\mathbf{q})$ on K-vektoriavaruutena äärellisulotteinen, joten myös $K[\mathbf{a}]$ on äärellisulotteinen.

Olkoon $\mathbf{a}$ transkendentti. Tällöin $K[\mathbf{a}]$ on isomorfinen polynomialgebran $K[\mathbf{X}]$ kanssa. Tässä jälkimmäisen tiedetään olevan ääretönulotteinen. Näin ollen myös $K[\mathbf{a}]$ on ääretönulotteinen. □

3.3. Polynomit lineaarialgebrassa

Olkoon V n-ulotteinen K-vektoriavaruus, $n \in \mathbb{N}$. Tällöin sen endomorfismien muodostama joukko

$$L(V) = \{L \colon V \to V \mid L \text{ on } K\text{-lineaarinen}\}$$

on K-algebra, jossa yhteenlasku ja skalaarikertolasku ovat määriteltyjä pisteittäin ja kertolaskuna käytetään kuvausten yhdistämisoperaatiota. Edellisen aliluvun nojalla jokaisella $L \in L(V)$ on olemassa algebrojen välinen *sijoitushomomorfismi* $S_L \colon K[\mathbf{X}] \to L(V)$. Käytännössä tämä tarkoittaa sitä, että jokaisella algebrallisella polynomilla $\mathbf{p} \in K[\mathbf{X}]$, $\mathbf{p} = \sum_{i=0}^{n} a_i \mathbf{X}^i$ voidaan muodostaa polynomilauseke

$$S_L(\mathbf{p}) = p(L) = \sum_{i=0}^{n} a_i \mathbf{X}^i = a_0 \operatorname{id}_V + a_1 L + \ldots + a_n L^n,$$

jonka arvo on myös algebran $L(V)$ alkio, eli eräs lineaarinen operaattori $V \to V$. On tärkeätä muistaa, että sijoitushomomorfismissa polynomin vakiotermi a_0 kuvautuu alkioksi $a_0 1_A$, missä 1_A on algebran kertolaskun neutraalialkio. Näin ollen kun $A = L(V)$ polynomin vakiotermiä a_0 vastaa algebrassa $L(V)$ operaattori $a_0 \operatorname{id}_V \colon V \to V$. Tämä on siis lineaarinen kuvaus, joka on määritelty kaavalla $a_0 \operatorname{id}_V(\mathbf{v}) = a_0 \mathbf{v}$ kaikilla $\mathbf{v}$.

K-vektoriavaruutena algebra $L(V)$ on äärellisulotteinen, itse asiassa Proposition 2.64 nojalla pätee

$$\dim L(V) = (\dim V)^2 = n^2.$$

Seurauksesta (3.66) seuraa tällöin, että jokainen sen alkio on *algebrallinen*. Toisin sanoen jokaista operaattoria $L \in L(V)$ vastaa sen *minimipolynomi*, jota merkitään symbolilla $\mathbf{m}_L$. Määritelmänsä mukaan $\mathbf{m}_L$ on pääpolynomi, $m_L(L) = 0$ ja pääideaali $(\mathbf{m}_L)$ koostuu tasan niistä polynomeista $\mathbf{p} \in K[\mathbf{X}]$, joille pätee $p(L) = 0$,

$$(\mathbf{m}_L) = \{\mathbf{p} \in K[\mathbf{X}] \mid p(L) = 0\}.$$

Operaattorin L minimipolynomi on asteltaan pienin pääpolynomi, jonka arvo operaattorissa L on nolla-operaattori. Operaattorin $L\colon V \to V$ virittämä alialgebra

$$K[L] = \{\sum_{i=0}^{n} a_i L^i\}$$

on algebrana isomorfinen tekijäalgebran $K[\mathbf{X}]/(\mathbf{m}_L)$ kanssa. Koska $\dim_K L(V) = n^2$, $\dim_K K[L] \leq n^2$, joten myös tekijäalgebran $K[\mathbf{X}]/(\mathbf{m}_L)$ dimensio K-vektoriavaruutena on korkeintaan n^2. Toisaalta Proposition 3.63 nojalla

$$\dim K[\mathbf{X}]/(\mathbf{m}_L) = \deg \mathbf{m}_L.$$

Näin ollen minimipolynomin $\mathbf{m}_L$ aste on korkeintaan n^2. Itse asiassa, osoittautuu, että aina pätee $\deg \mathbf{m}_L \leq n$. Tämä on seuraus Cayley-Hamiltonin Lauseesta, joka todistaan tässä aliluvussa myöhemmin.

Koska lineaarisia kuvauksia $L\colon V \to V$ voidaan ajatella myös $(n \times n)$-kokoisina matriiseina, samanlaisia tarkasteluja voidaan suorittaa matriiseille. Itse asiassa tällöin ei tarvitse puhua mistään konkreettisesta vektoriavaruudesta V.

Tarkemmin sanottuna, olkoon K kunta ja olkoon $n \in \mathbb{N}$. Tällöin $(n \times n)$-kokoisten neliömatriisien joukko $M(n \times n; K)$ on K-algebra. Jokaisella $A \in M(n \times n; K)$ voidaan muodostaa sijoitushomorfismi $S_A\colon K[\mathbf{X}] \to M(n \times n; K)$. Käytännössä tämä tarkoittaa sitä, että jokaisella $\mathbf{p} = \sum_{i=0}^{n} a_i \mathbf{X}^i \in K[\mathbf{X}]$ voidaan muodostaa neliömatriisi

$$S_A(\mathbf{p}) = p(A) = a_0 I_n + a_1 A + \ldots + a_n A^n \in M(n \times n; K).$$

Tässä I_n on yksikkömatriisi. Koska $\dim_K M(n \times n; K) = n^2$, jokainen matriisi $A \in M(n \times n; K)$ on algebrallinen K:n suhteen, joten sillä on olemassa nollasta eroava *minimipolynomi* $\mathbf{m}_A \in K[\mathbf{X}]$. Tämä on pääpolynomi, jolle pätee $m_A(A) = 0$, missä 0 on $(n \times n)$-kokoinen n-matriisi. Minimipolynomi virittää pääideaalin

$$(\mathbf{m}_A) = \{\mathbf{p} \in K[\mathbf{X}] \mid p(A) = 0\}.$$

Kuten yllä, nähdään, että matriisin A virittämä alialgebra

$$K[A] = \{\sum_{i=0}^{n} a_i A^i\}$$

on algebrana isomorfinen tekijäalgebran $K[\mathbf{X}]/(\mathbf{m}_A)$ kanssa ja

$$\dim K[A] = \dim K[\mathbf{X}]/(\mathbf{m}_A) = \deg \mathbf{m}_A \leq n^2.$$

Cayley-Hamiltonin Lauseesta 3.84 alla seuraa, että tämän arvion oikealla puolella luku n^2 voidaan korvata luvulla n.

Olkoon V n-ulotteinen K-vektoriavaruus ja olkoon E sen kanta. Tällöin vastaavuus $\Phi_E\colon L(V) \to M(n\times n; K)$, $\Phi_E(L) \mapsto [L]_E$ on K- algebrojen välinen isomorfismi (Seuraus 2.86). Helposti nähdään, että jokaisella $L \in L(V)$ sijoitushomomorfismeille S_L ja $S_{\Phi(L)}$ pätee (tarkista!) yhtälö $S_{\Phi(L)} = \Phi_E \circ S_L$. Koska Φ_E on isomorfismi ja minimipolynomit määräytyvät sijoitushomomorfismien ytimistä, tästä voidaan helposti johtaa seuraava yksinkertainen tulos (tarkka todistus harjoitustehtävänä). Tulos on melkein itsestään selvää, sillä operaattori ja sen matriisi voidaan ajatella olevan "sama asia".

Lemma 3.67. *Olkoon V n-ulotteinen K-vektoriavaruus ja olkoon E sen kanta. Olkoon $L\colon V \to V$ lineaarinen operaatorilla. Tällöin*

$$\mathbf{m}_L = \mathbf{m}_{[L]_E}.$$

Esimerkki 3.68. *Lineaarioperaattoria $L\colon V \to V$ sanotaan nilpotentiksi, jos on olemassa luonnollinen luku $m \in \mathbb{N}$, jolle pätee $L^m = 0$. Samoin neliömatriisia A sanotaan nilpotentiksi, jos $A^m = 0$ jollakin $m \in \mathbb{N}$. Pienintä luonnollista lukua $m \in \mathbb{N}$ jolle pätee $L^m = 0$ ($A^m = 0$) sanotaan nilpotentin operaattorin L (nilpotentin matriisin A) asteeksi.*

Olkoon m nilpotentin operaattorin L (matriisin A) aste. Tällöin polynomille $\mathbf{p} = \mathbf{X}^m$ pätee $\mathbf{p}(L) = 0$ ($\mathbf{p}(A) = 0$). Tästä seuraa, että operaattorin L (matriisin A) minipolynomi on polynomin $\mathbf{X}^m$ tekijä, eli muotoa $\mathbf{X}^k$, $0 \leq k \leq m$ (tämän todistamiseksi tarvitaan hajotelma jaottomiin tekijöihin, kts. Lause 3.48). Tällöin minimipolynomin määritelmän mukaan $X^k(L) = 0$. Asteen määritelmän mukaan m on kuitenkin pienin luonnollinen luku k, jolla on tämä ominaisuus. Näin ollen täytyy olla $k = m$ ja operaatiorin L (matriisin A) minipolynomi on polynomi $\mathbf{X}^m$.

Tyypillinen esimerkki nilpotentista operaattorista on esimerkissä 3.2 tarkasteltu derivaattaoperaatori $\mathcal{D}\colon P_n(\mathbb{R}) \to P_n(\mathbb{R})$ korkeintaan n-asteisten reaalilukukertoimisten polynomien muodostamassa vektoriavaruudessa $P_n(\mathbb{R})$. Tämän operaatorin aste nilpotentina operaattorina on $n+1$ (HT).

Karakteristinen polynomi

Minimipolynomin lisäksi jokaiseen operaattoriin tai neliömatriisiin voidaan liittää niin sanottu *karakteristinen polynomi* χ_L (χ_A). Tämä on sama polynomi, josta puhuttiin jo aikaisemmin ominaisarvojen yhteydessä, eli polynomi $k \mapsto \det(kI - A)$. Näin ajateltuna karakteristinen polynomi on kuitenkin *polynomifunktio*, ei abstrakti algebrallinen polynomi. Algebrallisen hengen mukaisesti halutaan kuitenkin määritellä se myös algebrallisesti eli polynomialgebran $K[\mathbf{X}]$ alkiona.

Määritellään karakteristinen polynomi ensin matriiseille. Olkoon A $(n \times n)$-matriisi. Tavoitteena olisi määritellä algebrallinen polynomi $\chi_A \in K[\mathbf{X}]$ siten, että ainakin kaikilla $k \in K$ pätisi

$$\chi_A(k) = \det(kI - A).$$

Yksi tapa olisi huomata, että polynomifunktion määritelmän $\chi_A\colon K \to K$ mukaan pätee

$$\chi_A(k) = \det(kI - A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)(k\delta_{\sigma(1)1} - a_{\sigma(1)1})(k\delta_{\sigma(2)2} - a_{\sigma(2)2}) \dots (k\delta_{\sigma(n)n} - a_{\sigma(n)n}),$$

missä δ_{ij} on Kronickerin delta (jonka arvo on yksi, kun $i = j$, ja nolla muuten). Tämä seuraa suoraan kaavasta 2.139. Matkimalla tätä kaavaa eli sijoittamalla muuttujan k paikalle abstrakti algebrallinen muuttujasymboli $\mathbf{X}$ voidaan määritellä

$$\chi_A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(\mathbf{X}\delta_{\sigma(1)1} - a_{\sigma(1)1})(\mathbf{X}\delta_{\sigma(2)2} - a_{\sigma(2)2}) \dots (\mathbf{X}\delta_{\sigma(n)n} - a_{\sigma(n)n}).$$

Tällä kaavalla saadaan hyvin määritelty polynomialgebran $K[\mathbf{X}]$ alkio, mutta tällainen konstruktio on kömpelö ja sitä on hyvin vaikeata soveltaa käytännössä.

Toinen tapa on käyttää hyväksi abstraktia teoriaa. Se on elegantimpi, mutta valitettavasti sisältää suhteellisen työläitä teknisiä yksityiskohtia.

Ajatus on seuraava. Koska K on polynomialgebran $K[\mathbf{X}]$ osajoukko (kunnan alkiothan vastaavat nolla-asteisia polynomeja), jokainen neliömatriisi $A \in M(n \times n; K)$ voidaan ajatella matriisina, jonka alkiot ovat polynomit , eli joukon $K[\mathbf{X}]$ alkioita. Tällaisten matriisien muodostamaa joukkoa voidaan merkitä $M(n \times n; K[\mathbf{X}])$. Tässä joukossa voidaan muodostaa matriisi $\mathbf{X}I_n - A$, jonka diagonaalialkiot ovat polynomeja $\mathbf{X} - a_{ii} \in K[\mathbf{X}]$, ja muut alkiot ovat kunnan alkioita (eli nolla-asteisia polynomeja) $-a_{ij}$. Karakteristinen polynomi $\chi(A)$ määritellään sen jälkeen *tämän matriisin determinanttina*.

Ongelma on kuitenkin siinä, että aikaisemmin ollaan käsitelty ainoastaan sellaisia matriiseja, **joiden kertoimet ovat jonkun kunnan alkioita**. Ainoastaan sellaisille matriiseille olemme määritelleet determinantin käsitteen ja tutkineet determinantin ominaisuuksia. Polynomialgebra $K[\mathbf{X}]$ taas ei ole kunta, se on vaihdannainen rengas, jossa vakiopolynomit ovat ainoat kääntyvät alkiot. Näin ollen, yleisesti ottaen, meidän olisi pitänyt kehittää determinanttien teoriaa sellaisille matriiseille, joiden alkiot ovat mielivaltaisen vaihdannaisen renkaan alkioita. Tämä on tietysti työläs, joskin suhteellisen suoraviivainen, yleistys kuntakertoimisten matriisien determinanttien teoriasta.

Toinen tapa kiertää tämä ongelma on **upottaa** polynomialgebra $K[\mathbf{X}]$ tietynlaisen **kunnan** $K(\mathbf{X})$ osajoukoksi. Tällöin voidaan puhua matriisista $\mathbf{X}I_n - A$ ja laskea sen determinantti, sillä se voidaan tällöin tulkita $(n \times n)$-matriisina, jonka kertoimet ovat kunnassa $K(\mathbf{X})$. Kunta $K(\mathbf{X})$ on niin sanottu *osamääräkunta*, jonka voidaan määritellä mille tahansa *kokonaisalue*-renkaalle.

Kokonaisalueen osamääräkunta

Olkoon R *kokonaisalue*. Tämä tarkoittaa sitä, että R on *vaihdannainen rengas*, jossa on voimassa "nollasääntö" eli kaikilla $a, b \in R$ ehto $ab = 0_R$ pätee jos ja vain jos $a = 0_R$ tai $b = 0_R$. Tästä seuraa, että vaihdannainen rengas on kokonaisalue jos ja vain jos sen nollasta eroavien alkioiden osajoukko $R \setminus \{0_R\}$ on vakaa renkaan kertolaskun suhteen. Kokonaisalueessa on voimassa "supistussääntö" - jos $a, b, c \in R$ ja $ab = ac$, niin joko $a = 0$ tai $b = c$. Kääntäen, jos vaihdannaisessa renkaassa pätee tämä supistussääntö, rengas on kokonaisalue.

Klassinen esimerkki kokonaisrenkaasta on kokonaislukujen rengas $\mathbb{Z}$. Itse asiassa juuri tämä rengas ja sen ominaisuudet johtivat aikoinaan kokonaisalueen käsitteen syntyyn (mikä näkyy termissä). Proposition 3.34 mukaan K-polynomialgebra $K[\mathbf{X}]$ on kokonai-

salue, kun K on mielivaltainen kunta. Kokonaislukujen modulo n muodostama äärellinen rengas $\mathbb{Z}_n$ on kokonaisalue jos ja vain jos n on alkuluku.

Kokonaislukujen rengas $\mathbb{Z}$ ei ole kunta, koska kaikilla kokonaisluvuilla, paitsi luvuilla ± 1, ei ole joukossa $\mathbb{Z}$ käänteislukua. Luonnollinen yritys "korjata" tämä ongelma on laajentaa kokonaislukujen joukko *rationaalilukujen joukkoon* $\mathbb{Q}$. Tämän joukon alkiot ovat murtolukuja muotoa a/b, missä $a, b \in \mathbb{Z}$ ovat kokonaislukuja ja $b \neq 0$. Murtoluku a/b on täysin määrätty kun tiedämme sen osoittajan a ja nimittäjän b, toisin sanoen, kun *järjestetty pari* (a, b) on annettu. Jokaista paria (a, b), missä $a, b \in \mathbb{Z}$, $b \neq 0$, vastaa murtoluku $a/b \in \mathbb{Q}$, mutta tämä vastavuus *ei ole* bijektiivinen, sillä eri pareja saattaa vastata sama murtoluku. Esimerkiksi pareja $(2, 3)$ ja $(4, 6)$ vastaa sama reaaliluku, $2/3 = 4/6$. Itse asiassa, kuten koulumatematiikastakin tiedämme, murtoluvut a/b ja c/d ovat samoja jos ja vain jos $ad = bc$.

Rationaaliluvuilla lasketaan seuraavilla säännöillä. Olkoot $a/b, c/d \in \mathbb{Q}$. Tällöin, tunnetusti,

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$
$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Mielivaltaisessa kokonaisalueessa voidaan menetellä samalla tavalla "lainaamalla" rationaalilukujen konstruktion välivaiheita. Olkoon R mielivaltainen *kokonaisalue*. Olkoon

$$X = \{(a, b) \in R^2 \mid b \neq 0_R\}.$$

Joukossa X määritellään laskutoimitukset $+$ ja $\cdot$ kaavoilla

$$(a, b) + (c, d) = (ad + bc, bd),$$

$$(a, b) \cdot (c, d) = (ac, bd).$$

Kumpikin laskutoimitus on hyvinmääritelty algebrallinen operaatio joukossa X, sillä R on kokonaisalue. Nimittäin, olkoot (a, b), $(c, d) \in R$. Tällöin $b \neq 0_R \neq d$, joten, kokonaisalueen määritelmän mukaan myös $bd \neq 0_R$. Tästä seuraa, että sekä pari $(ad + bc, bd)$, että pari (ac, bd) ovat joukossa X.

Lemma 3.69. *Yllä määritellyt laskutoimitukset $+$ ja $\cdot$ joukossa X ovat vaihdannaisia ja liitännäisia. Pari $(0, 1)$ on laskutoimituksen $+$ neutraalialkio ja pari $(1, 1)$ on laskutoimituksen $\cdot$ neutraalialkio.*

Kaikilla $a, b, c, d, e, f \in R$, $d, e, f \neq 0_R$ pätee

$$(a, b) \cdot (c, d) + (a, b) \cdot (e, f) = (b, b) \cdot (a, b) \cdot ((c, d) + (e, f)). \tag{3.70}$$

Todistus. Harjoitustehtävä. □

Määritellään joukossa X relaatio $\sim$ ehdolla $(a, b) \sim (a', b')$ jos ja vain jos $ab' = a'b$ (kerrotaan koordinaatteja "ristiin").

Lemma 3.71. *Relaatio $\sim$ on ekvivalenssirelaatio joukossa X. Tämä relaatio on yhteensopiva joukon X laskutoimitusten $+$ ja $\cdot$ kanssa.*

Todistus. Harjoitustehtävä. □

Edellisen lemman nojalla on olemassa *tekijäjoukko* $X/\sim$, lisäksi tässä joukossa voidaan määritellä indusoituja laskutoimituksia $+$ ja $\cdot$ (merkitään niitä tässä yhteydessä samoilla symboleilla kuin joukon X laskutoimituksia). Tekijäjoukkoa $X/\sim$ merkitään $Q(R)$. Joukon $Q(R)$ alkiot ovat ekvivalenssiluokkia $\overline{(a,b)}$, missä $a, b \in R$, $b \neq 0_R$. Tälle ekvivalenssiluokalle käytämme myös merkintää a/b tai $\frac{a}{b}$. Relaation $\sim$ määritelmän mukaan tällöin

$$\frac{a}{b} = \frac{c}{d}$$

jos ja vain jos renkaassa R pätee yhtälö $ad = bc$. Erityisesti kaikilla $a, b, c \in R, b, c \neq 0_R$ pätee

$$\frac{ac}{bc} = \frac{a}{b}. \tag{3.72}$$

Kun tämä yhtälö luetaan vasemmalta oikealle, se voidaan tulkita "yhteisen tekijän c supistuksena". Kun se luetaan taas oikealta vasemmalle, kyseessä on "laventaminen alkiolla c. " Laskutoimitukset $+$ ja $\cdot$ voidaan näillä merkinnöillä kirjoittaa yhtälöinä

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Laskutoimitusta $+$ sanotaan luonnollisesti joukon $Q(R)$ yhteenlaskuksi, laskutoimitusta $\cdot$ vastaavasti kertolaskuksi.

Propositio 3.73. *Kolmikko* $(Q(R), +, \cdot)$ *on kunta.*

Todistus. Yhteen- ja kertolaskun vaihdannaisuus ja liitännäisyys periytyvät joukon X vastaavien laskutoimitusten vastaavista ominaisuuksista (Lemma 3.69). Samalla tavalla Lemmasta 3.69 seuraa, että laskutoimituksella $+$ on neutraalialkiona alkio $0/1$ ja laskutoimituksella $+$ on neutraalialkiona alkio $1/1$. Huomataan, että yhtälön 3.72 nojalla kaikilla $b \in R, b \neq 0_R$, pätee

$$\frac{0}{1} = \frac{0}{b} \text{ ja } \frac{1}{1} = \frac{b}{b}.$$

Olkoot $a, b \in R, b \neq 0_R$. Tällöin edellisen nojalla

$$\frac{a}{b} + \frac{-a}{b} = \frac{ab + (-a)b}{b^2} = \frac{0}{b^2} = \frac{0}{1}.$$

Koska $0/1$ on yhteenlaskun $+$ neutraalialkio, tästä seuraa, että jokaisella $a/b \in Q(R)$ on olemassa vasta-alkio, joka on lisäksi yllä olevan laskun nojalla alkio $(-a)/b$.

Olkoot $a, b \in R, b \neq 0_R$. Tällöin relaation $\sim$ määritelmän nojalla $a/b = 0/1$ jos ja vain jos $a = a \cdot 1 = b \neq 0 = 0$. Näin ollen, jos $a/b \neq 0/1$, pätee $a \neq 0_R$. Tällöin myös b/a on hyvin määritelty $Q(R)$:n alkio ja

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ab} = \frac{1}{1}.$$

Koska $1/1$ on kertolaskun $\cdot$ neutraalialkio, tästä seuraa, että jokaisella $a/b \in Q(R)$, $a/b \neq 0 = 0/1$ on olemassa käänteisalkio kertolaskun suhteen. Lisäksi yllä olevan laskun nojalla $(a/b)^{-1} = b/a$.

Jäljellä on osittelulaki. Lemmasta 3.69 seuraa, että kaikilla $a, b, c, d, e, f \in R$, $d, e, f \neq 0_R$ pätee

$$\frac{a}{b} \cdot \frac{c}{d} + \frac{a}{b} \cdot \frac{e}{f} = \frac{b}{b} \cdot \frac{a}{b} \cdot (\frac{c}{d} + \frac{e}{f}).$$

Kuitenkin tässä vaiheessa tiedetään jo, että $\frac{b}{b} = 1/1$ on kertolaskun neutraalialio. Näin ollen myös osittelulaki

$$\frac{a}{b} \cdot \frac{c}{d} + \frac{a}{b} \cdot \frac{e}{f} = \frac{a}{b} \cdot (\frac{c}{d} + \frac{e}{f})$$

on voimassa. On osoitettu, että kolmikko $(Q(R), +, \cdot)$ on kunta. □

Kuntaa $Q(R)$ sanotaan kokonaisalueen R *osamääräkunnaksi*. Korostetaan vielä kerran, että osamääräkunnan konstruktio sellaisena kuin se on annettu yllä on mahdollista suorittaa vain kokonaisalueen tapauksessa, ei yleisen (jopa vaihdannaisen) renkaan kohdalla.

Näytetään vielä, että kokonaisalue R voidaan ”upottaa” jakokuntaansa $Q(R)$. Olkoot $a, b \in R$ mielivaltaiset. Tällöin

$$\frac{a}{1} + \frac{b}{1} = \frac{a1 + 1b}{1 \cdot 1} = \frac{a+b}{1}, \text{ ja } \frac{a}{1} \cdot \frac{b}{1} = \frac{ab}{1}.$$

Lisäksi $1/1$ on renkaan $Q(R)$ kertolaskun neutraalialkio. Nämä havainnot tarkoittavat täsmälleen sitä, että kuvaus $f\colon R \to Q(R)$, $f(a) = a/1$ on *rengashomomorfismi*. Osoitetaan vielä, että f on injektio. Olkoot $a, b \in R$. Oletetaan, että

$$f(a) = \frac{a}{1} = \frac{b}{1} = f(b).$$

Tällöin ekvivalenssirelaation $\sim$ nojalla $a = a \cdot 1 = 1 \cdot b = b$. Näin ollen f on injektiivinen rengashomomorfismi, eli isomorfismi $R \to f(R)$. Rengas R on siis kanonisella tavalla isomorfinen erään kunnan $Q(R)$ alirenkaan $f(R)$ kanssa. Tästä johtuen voidaan ajatella, että kokonaisalue R on osamääräkuntansa alirengas.

Tapauksessa $R = \mathbb{Z}$ osamääräkunnasta $Q(R)$ tulee täsmälleen samannäköinen kuin rationaalilukujen kunta $\mathbb{Q}$. Näin ollen isomorfiaa vaille $Q(\mathbb{Z}) = \mathbb{Q}$. Itse asiassa osamääräkunnan $Q(\mathbb{Z})$ konstruktiota voidaan pitää rationaalilukujen kunnan $\mathbb{Q}$ konstruktiona. Toisin sanoen, jos oletetaan, että kokonaislukujen rengas $\mathbb{Z}$ on annettu, tällä tavalla saadaan osoitettua, että on olemassa rationaalilukujen kunta $\mathbb{Q}$.

Olkoon K kunta. Polynomialgebra $K[\mathbf{X}]$ on tällöin kokonaisalue Lemman 3.34 nojalla. Näin ollen on olemassa sen osamääräkunta $Q(K[\mathbf{X}])$, jota merkitään jatkossa symbolilla $K(\mathbf{X})$. Havainnollisesti ajatellen jakokunta $K(X)$ koostuu abstakteista *murtolauseekkeista* muotoa

$$\frac{a_n\mathbf{X}^n + a_{n-1}\mathbf{X}^{n-1} + \ldots + a_1\mathbf{X} + a_0}{b_m\mathbf{X}^m + b_{m-1}\mathbf{X}^{m-1} + \ldots + b_1\mathbf{X} + b_0},$$

missä $a_0, \ldots, a_n, b_0, \ldots, b_m \in K$ ja $\mathbf{X}$ on abstrakti algebrallinen ”muuttujasymboli” (joka on renkaan $K[\mathbf{X}]$ alkio).

Karakteristinen polynomi

Nyt voidaan palata matriisin ja lineaarisen operaattorin karakteristisen polynomin määritelmään formaalilla tasolla. Olkoon K kunta, olkoon $n \in \mathbb{N}$ ja olkoon $A \in M(n \times n; K)$ $(n \times n)$-kokoinen K-kertoiminen neliömatriisi. Kuntaa K voidaan ajatella renkaan $K[\mathbf{X}]$ alikuntana (samastetaan kunnan alkiot ja vakiopolynomit). Rengas $K[\mathbf{X}]$ puolestaan voidaan edellisen nojalla ajatella osamääräkuntansa $K(\mathbf{X})$ alirenkaana.

Muodostetaan $K(\mathbf{X})$-kertoiminen $(n \times n)$-kokoinen matriisi $B = \mathbf{X}I_n - A$. Tämä on siis matriisijoukon $M(n \times n; K(\mathbf{X}))$ alkio. Matriisin $B = (b_{ij})$ alkiot b_{ij}, $i, j = 1, \ldots, n$ ovat määritelty seuraavasti,

$$b_{ij} = \begin{cases} \mathbf{X} - a_{ii}, \text{ kun } i = j, \\ -a_{ij}, \text{ muuten.} \end{cases}$$

Koska matriisin B kertoimet sijaitsevat eräässä kunnassa $K(\mathbf{X})$, siihen voidaan soveltaa matriisien determinanttien teoriaa, joka on kehitetty aliluvussa 2.5. Erityisesti voidaan puhua matriisin B determinantista $\det B$, joka on tällöin, määritelmän mukaan, jakokunnan $K(\mathbf{X})$ alkio, toisin sanoen eräs *murtolauseke* muotoa P/Q, missä $P, Q \in K[\mathbf{X}]$ ovat K-kertoimisia polynomeja. Kuitenkin, kaavan 2.139 nojalla voidaan kirjoittaa matriisin B determinantti auki muodossa

$$\det(\mathbf{X}I_n - A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(\mathbf{X}\delta_{\sigma(1)1} - a_{\sigma(1)1})(\mathbf{X}\delta_{\sigma(2)2} - a_{\sigma(2)2}) \ldots (\mathbf{X}\delta_{\sigma(n)n} - a_{\sigma(n)n}). \tag{3.74}$$

Tästä kaavasta nähdään, että $\det(\mathbf{X}I_n - A)$ on polynomien tulona ja summana itse polynomi, eli polynomialgebran $K[\mathbf{X}]$ alkio. Tätä polynomia $\det(\mathbf{X}I_n - A) \in K[\mathbf{X}]$ sanotaan *matriisin A karakteristiseksi polynomiksi*, sitä merkitään symbolilla χ_A.

Sijoitushomomorfismin välityksellä voidaan jokainen skalaarikunnan alkio $k \in K$ "sijoittaa" karakteristiseen polynomiin χ_A muuttujasymbolin paikalle. Näin saadaan muodostettua vastaava *karakteristinen polynomifunktio* $\chi_A \colon K \to K$. Yhtälöstä 3.74 seuraa, että jokaisella $k \in K$ pätee

$$\chi_A(k) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(k\delta_{\sigma(1)1} - a_{\sigma(1)1})(k\delta_{\sigma(2)2} - a_{\sigma(2)2}) \ldots (k\delta_{\sigma(n)n} - a_{\sigma(n)n}).$$

Toisaalta kaavan 2.139 nojalla tämän yhtälön toisella puolella on laskettu auki K-kertoimisen $(n \times n)$-kokoisen matriisin $(kI_n - A) \in M_n(n \times n; K)$ determinantti. Toisin sanoen kaikilla $k \in K$ pätee $\chi_A(k) = \det(kI_n - A)$. Tämä on sama polynomifunktio, jota on käytetty aliluvussa 3.1 matriisin A ominaisarvojen laskemiseen.

Lineaarisen operaattorin karakteristinen polynomi

Seuraavaksi määritellään lineaarisen operaattorin $L \colon V \to V$ karakteristinen polynomi $\chi_L \in K[\mathbf{X}]$ abstraktina algebrallisena polynomina. Olkoon V äärellisulotteinen K-vektoriavaruus ja olkoon E jokin sen kanta. Tällöin $A = [L]_E$ on $(n \times n)$-kokoinen K-kertoiminen matriisi, joten sen karakteristinen polynomi $\chi_A \in K[\mathbf{X}]$ on jo määritelty yllä. Asetetaan $\chi_L = \chi_A$.

Tietysti täytyy tarkistaa, että tällainen määritelmä ei riipu kannan valinnasta. Olkoon E' jokin toinen avaruuden V kanta. Tällöin kannanvaihtokaavan mukaan pätee $[L]_{E'} = Y[L]_E Y^{-1}$, missä $Y = [E' \mid E]$ on kannanvaihtomatriisi. Halutaan osoittaa, että $\chi_{[L]_E} = \chi_{[L]_{E'}}$. Edellisen nojalla tämä seuraa suoraan seuraavasta lemmasta.

Lemma 3.75. *Olkoot $A, Y \in M(n \times n; K)$, missä Y on kääntyvä matriisi. Olkoon $B = YAY^{-1}$. Tällöin*

$$\chi_A = \chi_B.$$

Todistus. Ajatellaan matriiseja A, B ja Y matriiseina osamääräkunnan $K(\mathbf{X})$ yli eli joukon $M(n \times n; K(\mathbf{X}))$ alkioina. Tällöin joukossa $M(n \times n; K(\mathbf{X}))$ voidaan laskea

$$\mathbf{X}I_n - B = \mathbf{X}I_n - YAY^{-1} = \mathbf{X}YY^{-1} - YAY^{-1} = Y(\mathbf{X}I_n)Y^{-1} - YAY^{-1} = Y(\mathbf{X}I_n - A)Y^{-1}.$$

Tässä muuttujasymboli $\mathbf{X}$ käsitellään *skalaarina*, sillä nyt lasketaan kerroinkunnan $K(\mathbf{X})$ yli ja $\mathbf{X}$ on tämän kerroinkunnan alkio. Tämän perusteella voidaan kirjoittaa

$$\mathbf{X}YY^{-1} = Y(\mathbf{X}Y^{-1}) = Y(\mathbf{X}I_n)Y^{-1},$$

sillä $K(\mathbf{X})$-algebrassa $M(n \times n; K(\mathbf{X}))$ pätee, kuten jokaisessa algebrassa, laskusääntö $k(\mathbf{ab}) = \mathbf{a}(k\mathbf{b})$ (kertolaskun bilineaarisuus), kun k on skalaari ja $\mathbf{a}, \mathbf{b}$ ovat agebran alkioita. Näin ollen

$$\chi_B = \det(\mathbf{X}I_n - B) = \det Y \det(\mathbf{X}I_n - A)(\det Y)^{-1} = \det(\mathbf{X}I_n - A) = \chi_A.$$

Tässä laskussa käytetään hyväksi determinantin ominaisuuksia, tarkemmin sanottuna Proposition 2.144 tulosta ja sen seurauksia. Huomaa erityisesti, että tämä ei olisi mahdollista, jos emme olisi varmoja siitä, että kaikki edellä tarkastellut matriisit ovat matriiseja jonkun kunnan yli, sillä kaikki determinantin teoreettiset ominaisuudet (kuten Proposition 2.144) todistettiin ainoastaan kuntakertoimiselle matriiseille. Juuri tästä syystä joudutaan turvautumaan osamääräkuntaan. □

Keskustelu polynomeista oli alunperin motivoitu ominaisarvojen tarkastelulla. Ominaisarvoihin liittyvien ongelmien ratkaiseminen onkin yksi (mutta ei suinkaan ainoa) polynomialgebran sovelluksista äärellisulotteisessa lineaarialgebrassa.

Propositio 3.76. *Olkoon $A \in M(n \times n; K)$ neliömatriisi.*

- *Karakteristisen polynomin χ_A aste on tasan n ja se on pääpolynomi.*
- *Kunnan K alkio $k \in K$ on matriisin A ominaisarvo jos ja vain jos k on karakteristisen polynomin χ_A juuri.*

Todistus. Ensimmäinen väite osoitetaan samalla tavalla kuin se oli osoitettu polynomifunktiolle $\chi_A \colon K \to K$ aliluvussa 3.1 eli yhtälön

$$\chi_A = \sum_{\sigma \in S_n} (\operatorname{sgn}(\sigma))(\mathbf{X}\delta_{\sigma(1)1} - a_{\sigma(1)1})(\mathbf{X}\delta_{\sigma(2)2} - a_{\sigma(2)2}) \dots (\mathbf{X}\delta_{\sigma(n)n} - a_{\sigma(n)n})$$

avulla. Toinen väite oli osoitettu jo aikaisemmin samassa aliluvussa 3.1, sillä karakteristisen polynomin χ_A juuret kunnassa K ovat täsmälleen sama asia kuin vastaavan polynomifunktion $\chi_A \colon K \to K$ nollakohdat. □

Kääntämällä edellisen proposition tuloksia lineaaristen endomorfismien kielelle saadaan samalla osoitettua samanlaisia tuloksia lineaarisille operaattoreille.

Propositio 3.77. *Olkoon V n-ulotteinen K-vektoriavaruus ja olkoon $L\colon V \to V$ lineaarinen operaattori. Tällöin seuraavat väitteet ovat tosia.*

- *Karakteristisen polynomin χ_L aste on tasan n.*
- *Kunnan K alkio $k \in K$ on operaattorin L ominaisarvo jos ja vain jos k on karakteristisen polynomin χ_L juuri.*

Käytännössä karakteristisia polynomeja lasketaan samalla tavalla kuin ennenkin. Esimerkiksi jos $L\colon V \to V$ on 2-ulotteisen $\mathbb{R}$-vektoriavaruuden lineaarinen operaattori, jonka matriisi $[L]_E$ jonkun V:n kannan E suhteen on

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$

niin L:n karakteristinen polynomi on

$$\chi_L = \det(\mathbf{X}I_n - [L]_E) = \det \begin{bmatrix} \mathbf{X}-1 & 1 \\ 0 & \mathbf{X}-1 \end{bmatrix} = (\mathbf{X}-1)^2 - 0 \cdot 1 = \mathbf{X}^2 - 2\mathbf{X} + 1.$$

Koska $\chi_L = (\mathbf{X}-1)^2$, kuvauksen L ainoa ominaisarvo r on polynomin χ_L ainoa juuri $r = 1$.

Ominaisarvon geometriset ja algebralliset kertaluvut
Olkoon $L\colon V \to V$ äärellisulotteisen K-vektoriavaruuden lineaarinen operaattori ja olkoon $k \in K$ sen ominaisarvo. Palautetaan mieleen, että avaruuden V aliavaruutta

$$V_k = \{\mathbf{v} \in V \mid L(\mathbf{v}) = k\mathbf{v}\}$$

sanotaan ominaisarvoon liittyväksi aliavaruudeksi. Tämän avaruuden dimensiota $\dim V_k$ sanotaan ominaisarvon k *geometriseksi kertaluvuksi* ja merkitään symbolilla $\dim^g(k; L)$.

Jokainen operaattorin L ominaisarvo k on, toisaalta, sen karakteristisen polynomin χ_L juuri. Tämän juuren *kertalukua* l sanotaan ominaisarvon k *algebralliseksi kertaluvuksi* ja merkitään symbolilla $\dim^a(k; L)$. Tällöin $\chi_L = (\mathbf{X}-k)^l\mathbf{p}$, missä k ei ole polynomin $\mathbf{p}$ juuri.

Neliömatriisin $A \in M(n \times n; K)$ ominaisarvon k geometrinen ja algebrallinen kertaluvut määritellään samalla tavalla. Täsmällisesti sanottuna olkoon $L_A\colon K^n \to K^n$ kanoninen kuvaus, joka liittyy matriisiin A. Tällöin matriisin ominaisarvon k geometrinen kertaluku on $\dim^g(k; A) = \dim V_k$, missä

$$V_k = \{\mathbf{v} \in K^n \mid A\mathbf{v} = k\mathbf{v}\}$$

on ominaisarvoa k vastaava ominaisarvoaliavaruus. Ominaisarvon k algebrallinen kertaluku $\dim^a(k; L)$ on vastaavasti sen kertaluku karakteristisen polynomin χ_A juurena.

Lemma 3.78. *Äärellisulotteisen K-vektoriavaruuden V lineaarisen operaattorin $L \in L(V)$ ominaisarvon k geometrinen kertaluku on pienempi tai yhtä suuri kuin sen algebrallinen kertaluku,*

$$\dim^g(k; L) \leq \dim^a(k; L).$$

Todistus. Olkoon

$$V_k = \{\mathbf{v} \in V \mid L(\mathbf{v}) = k\mathbf{v}\}$$

ominaisarvoa k vastaava ominaisarvoaliavaruus. Valitaan aliavaruudelle V_k jokin kanta $E' = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$, missä $n = \dim V_k$, ja täydennetään se koko avaruuden kannaksi $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots, \mathbf{e}_m)$. Tällöin (koska V_k on invariantti aliavaruus) operaattorin L matriisi $[L]_E$ kannan E suhteen on lohkomatriisi muotoa

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

missä $A \in M(n \times n; K)$ on rajoittuman $L|V_k \colon V_k \to V_k$ matriisi kannan E' suhteen. Tästä seuraa, että

$$\chi_L = \chi_{[L]_E} = \det \begin{bmatrix} \mathbf{X}I_n - A & -B \\ 0 & \mathbf{X}I_n - D \end{bmatrix} = \det(\mathbf{X}I_{m-n} - A) \det(\mathbf{X}I_n - D). \tag{3.79}$$

Tarkat perustelut sille, että tässä tapauksessa determinanteilla voidaan laskea näin, jätetään harjoitustehtävänä.

Ominaisarvon määritelmän nojalla rajoittuma $L|V_k$ on kaavalla $\mathbf{w} \mapsto k\mathbf{w}$ määritely kuvaus, joten A on *diagonaalimatriisi*, jonka jokainen diagonaalialkio on k (ja muut alkiot nollia). Tästä seuraa, että $\mathbf{X}I_n - A$ on puolestaan diagonaalimatriisi, jonka jokainen diagonaalialkio on $\mathbf{X} - k$. Näin ollen $\det(\mathbf{X}I_n - A) = (\mathbf{X} - k)^n$, missä

$$n = \dim V_k = \dim^g(k; L)$$

on ominaisarvon k geometrinen kertaluku. Yhtälöketjusta 3.79 saadaan tällöin, että $\chi_L = (\mathbf{X} - k)^n \mathbf{p}$, missä $\mathbf{p} = \det(\mathbf{X}I_{m-n} - D)$ on eräs polynomi (itse asiassa matriisin D karakteristinen polynomi). Tästä nähdään, että polynomin χ_L juuren k *algebrallinen kertaluku* $\dim^a(k; L)$ on *vähintään* n (huom., se voi olla suurempi kuin n, jos k saattuu olemaan myös polynomin $\mathbf{p}$ juuri). Koska $n = \dim^g(k; L)$, lemman väite on osoitettu. □

Esimerkki 3.80. *Esimerkissä 3.16, 1) on näytetty, että $\mathbb{R}$-kertoimisen matriisin*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

karakteristinen polynomi on $(\mathbf{X} - 1)^2$. Tällä polynomilla on vain yksi juuri 1, jonka kertaluku (juurena) on 2. Ajatellaan A lineaarisena kuvauksena $L_A \colon \mathbb{R}^2 \to \mathbb{R}^2$, jonka matriisi standardikannan suhteen on A. Tällöin 1 on tämän operaattorin ainoa ominaisarvo. Vastaava ominaisarvoaliavaruus

$$V_1 = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

on yksiulotteinen, joten tämän ominaisarvon geometrinen kertaluku on 1. Tämä on siis esimerkki tilanteesta, jossa ominaisarvon algebrallinen kertaluku on aidosti suurempi kuin sen geometrinen kertaluku.

Seuraavassa tuloksessa operaattorin diagonalisoituvuutta luonnehditaan karakteristisen polynomien juurten ominaisuuksien avulla.

Lemma 3.81. *Olkoon $L\colon V \to V$ lineaarinen operaattori äärellisulotteisessa K-vektoriavaruudessa V. Tällöin L on diagonaalisoituva jos ja vain jos seuraavat ehdot pätevät.*

(1) Operaattorin L jokaisen ominaisarvon geometrinen kertaluku on sama kuin sen algebrallinen kertaluku.

(2) Operaattorin L karakteristinen polynomi χ_L voidaan jakaa polynomialgebrassa $K[\mathbf{X}]$ ensimmäisen asteen tekijöihin, toisin sanoen

$$\chi_L = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \dots (\mathbf{X} - k_m)^{l_m}$$

joillakin $k_1, \dots, k_n \in K$, $l_1, \dots, l_m \geq 1$.

Todistus. Olkoon $n = \dim V$. Tällöin karakteristisen polynomin χ_L aste on tasan n ja sillä on korkeintaan n juurta kunnassa K. Olkoot $k_1, \dots, k_m$ kaikki karakteristisen polynomin χ_L (eri) juuret kunnassa K. Tällöin k_i on kuvauksen L ominaisarvo kaikilla $i = 1, \dots, m$ ja kääntäen kaikki operaattorin L ominaisarvot ovat tätä muotoa. Proposition 3.55 nojalla

$$\chi_L = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \dots (\mathbf{X} - k_m)^{l_m}\mathbf{q},$$

missä $l_i = \dim^a(k_i; L)$ on ominaisarvon k_i algebrallinen kertaluku ja $\mathbf{q} \neq \mathbf{0}$ on sellainen polynomialgebran $K[\mathbf{X}]$ alkio, jolla ei ole lainkaan juuria kunnassa K (mahdollisesti vakiopolynomi 1). Tässä $\mathbf{q}$ on välttämättä pääpolynomi (koska χ_L on pääpolynomi). Tällöin (Lemma 3.34)

$$n = l_1 + l_2 + \dots + l_n + \deg \mathbf{q}. \tag{3.82}$$

Olkoon $W = \oplus_{i=1}^m V_{k_i}$ (tämä summa on suora Lemman 3.7 nojalla), tällöin L on diagonalisoituva jos ja vain jos $W = V$. Koska W on avaruuden V aliavaruus, tämä on mahdollista jos ja vain jos $\dim W = \dim V = n$ (Propositio 2.48). Toisaalta Proposition 2.160 nojalla

$$\dim W = \sum_{i=1}^{n} \dim V_{k_i} = \sum_{i=1}^{n} \dim^g(k_i; L).$$

Näin ollen, edellisen lemman nojalla,

$$\dim W = \sum_{i=1}^{n} \dim^g(k_i; L) \overset{(i)}{\leq} \sum_{i=1}^{n} \dim^a(k_i; L) = \sum_{i=1}^{n} l_i \overset{(ii)}{\leq} n.$$

Tästä epäyhtälöketjusta nähdään, että $\dim W = n$ jos ja vain jos kohdissa (i) ja (ii) epäyhtälöt ovatkin yhtälöitä. Yhtälö kohdassa (i) pätee jos ja vain jos $\dim^g(k_i; L) = \dim^a(k_i; L)$ kaikilla $i = 1, \dots, m$, toisin sanoen jos ja vain jos jokaisen ominaisarvon geometrinen kertaluku on sama kuin sen algebrallinen kertaluku. Tämä on ehto (1). Yhtälö kohdassa (ii) taas pätee jos ja vain jos $\deg \mathbf{q} = 0$ (kts. yhtälö 3.82) eli jos ja vain jos $\mathbf{q} = 1$. Toisin sanoen yhtälö kohdassa (ii) on yhtäpitävä ehdon (2) kanssa. □

Koska algebrallisesti suljetun kunnan K tapauksessa edellisen Lemman ehto (2) on voimassa mille tahansa polynomialgebran $K[\mathbf{X}]$ polynomille (Seuraus 3.56), saadaan algebrallisesti suljetun kunnan erikoistapauksessa seuraava tulos.

Seuraus 3.83. *Olkoon $L\colon V \to V$ lineaarinen operaattori äärellisulotteisessa K-vektoriavaruudessa V, missä kunta K on algebrallisesti suljettu. Tällöin L on diagonalisoituva jos ja vain jos operaattorin L jokaisen ominaisarvon geometrinen kertaluku on sama kuin sen algebrallinen kertaluku.*

Cayley-Hamiltonin lause.
Tällä nimellä tunnetaan seuraava Lause 3.84, jota pidetään yhtenä lineaarialgebran tärkeämpinä perustuloksina.

Olkoon V K-vektoriavaruus ja olkoon $L\colon V \to V$ lineaarinen operaattori. Koska *kaikkien* lineaaristen operaattorien $V \to V$ muodostama joukko $L(V)$ on K-*algebra*, on olemassa *sijoitushomomorfismi* $S_L\colon K[\mathbf{X}] \to L(V)$. Käytännössä tämä tarkoittaa sitä, että jokaista polynomia $\mathbf{p} = \sum_{i=0}^{n} a_i \mathbf{X}^i \in K[\mathbf{X}]$ vastaa algebran $L(V)$ alkio $p(L) = \sum_{i=0}^{n} a_i L^i$, joka saadaan sijoittamalla muuttujasymbolin $\mathbf{X}$ paikalle operaattori L. Tässä $L^0 = \mathrm{id}_V\colon V \to V$ on identtinen kuvaus (algebran $L(V)$ neutraalialkio).

Erityisesti operaattori L voidaan sijoittaa *sen karakteristiseen polynomiin* χ_L, jolloin saadaan eräs lineaarinen operaattori $\chi_L(L)$. Cayley-Hamiltonin lause sanoo, että tämä operaattori on itse asiassa aina nollakuvaus.

Lause 3.84. *Olkoon $L\colon V \to V$ lineaarinen kuvaus. Tällöin*

$$\chi_L(L) = 0.$$

Todistus. Korvaamalla L matriisilla $A = [L]_E$ (jonkun avaruuden V kannan E suhteen) nähdään, että riittää todistaa samanlainen väite matriisille A, toisin sanoen riittää näyttää, että jokaiselle $(n \times n)$-kokoiselle K-kertoimiselle matriisille $A \in M(n \times n; K)$ pätee

$$\chi_A(A) = 0.$$

Olkoon K' jokin kunnan K sisältävä algebrallisesti suljettu kunta (joka on olemassa Proposition 3.17 mukaan). Ajatelemme siis kunta K jonkun algebrallisesti suljetun kunnan K' alikuntana (esimerkiksi tapauksessa $K = \mathbb{R}$ voidaan ottaa $K' = \mathbb{C}$). Jokainen K-kertoiminen matriisi voidaan yhtä hyvin ajatella K'-kertoimisena matriisina eli algebran $M(n \times n; K')$ alkiona. Selvästi matriisin A karakterteristinen polynomi χ_A ei riipu siitä, ajatellaanko A K-kertoimisena matriisina vai K'-kertoimisena matriisina. Edellisessä tapauksessa χ_A on formaalisti K-kertoiminen polynomi (eli polynomiagebran $K[\mathbf{X}]$ alkio), jälkimmäisessä se on K'-kertoiminen polynomi (eli polynomiagebran $K'[\mathbf{X}]$ alkio), mutta kummassakin tapauksessa sillä on täsmälleen samat kertoimet (jotka ovat kunnan K alkioita). Tällöin yhtälö $\chi_A(A) = 0$ pätee algebrassa $M(n \times n; K)$ jos ja vain jos se pätee algebrassa $M(n \times n; K')$, joten, voidaan olettaa yhtä hyvin alusta alkaen, että $K = K'$ on algebrallisesti suljettu kunta.

Huomaa, että samantyyppista temppua ei oltaisi voitu suorittaa yhtä luonnollisella tavalla lineaariselle operaattorille $L\colon V \to V$ (kuin sitä vastaaville matriisille). Tämä on esimerkki tilanteesta jossa matriisi-näkökulma lineaarisiin kuvauksiin on kätevämpi kuin niiden tarkastelu varsinaisina kuvauksina. Seuraavaksi näkökulma vaihdetaan kuitenkin takaisin ja tarkastellaan lineaarisia operaattoreita matriisien sijaan.

Olkoon K algebrallisesti suljettu kunta, olkoon V äärellisulotteinen K-vektoriavaruus ja olkoon $L\colon V \to V$ lineaarinen operaattori. Proposition 3.22 nojalla avaruudella V on olemassa kanta $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ jonka suhteen kuvauksen L matriisi $[L]_E$ on *yläkolmiomatriisi*

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ 0 & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}$$

jokaisella $i = 1, \ldots, n$. Olkoon $W_i = \operatorname{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_i)$. Tällöin

$$W_1 \subset W_2 \subset \ldots \subset W_{n-1} \subset W_n = V,$$

$\dim W_i = i$ jokaisella $i = 1, \ldots, n$ ja jokainen aliavaruus W_i on L-invariantti. Operaattorin L karakteristiselle polynomille χ_L pätee

$$\chi_L = \chi_A = \det(\mathbf{X}I_n - A).$$

Tässä $\mathbf{X}I_n - A$ on myös *yläkolmiomatriisi*, jonka diagonaalialkiot ovat $\mathbf{X} - a_{ii}$, $i = 1, \ldots, n$. Koska yläkolmiomatriisin determinantti on sen diagonaalialkioiden tulo, tästä saadaan suoraan, että

$$\chi_L = \prod_{i=1}^{n} (\mathbf{X} - a_{ii}),$$

joten $\chi_L(L) = \prod_{i=1}^{n} (L - a_{ii} \operatorname{id}_V)$.

Koska matriisi $A = [L]_E$ on yläkolmiomatriisi, pätee

$$L(\mathbf{e}_i) = \mathbf{v}_i + a_{ii}\mathbf{e}_i = \mathbf{v}_i + k_i\mathbf{e}_i,$$

missä $\mathbf{v}_i \in W_{i-1}$, jokaisella $i = 1, \ldots, n$. Tässä tulkitaan $W_0 = \{0\}$ sekä merkitään $a_{ii} = k_i$. Olkoon $i = 1, \ldots, n$, tällöin

$$(L - k_i \operatorname{id}_V)(\mathbf{e}_i) = \mathbf{v}_i + k_i\mathbf{e}_i - k_i\mathbf{e}_i = \mathbf{v}_i \in W_{i-1}. \tag{3.85}$$

Toisaalta, koska W_{i-1} on invariantti sekä operaattorin L, että identtisen operaattorin id_V suhteen, kaikilla $j = 1, \ldots, i-1$ pätee

$$(L - k_i \operatorname{id}_V)(\mathbf{e}_j) \in W_{i-1}. \tag{3.86}$$

Yhtälöistä 3.85 ja 3.86 seuraa, että $(L - k_i \operatorname{id}_V)(W_i) \subset W_{i-1}$ kaikilla $i = 1, \ldots, n$. Näin ollen saadaan induktiivisesti

$$\chi_L(L)(V) = \prod_{i=1}^{n} (L - k_i)(V) \subset \prod_{i=1}^{n-1} (L - k_i) W_{n-1} \subset \ldots \subset (L - k_1)(W_1) \subset W_0 = \{0\}.$$

Toisin sanoen $\chi_L(L)$ on avaruuden V nolla-operaattori, mikä pitikin todistaa. □

Cayley-Hamiltonin lause pätee tietysti yhtä hyvin myös matriiseille. Toisin sanoen jos $A \in M(n \times n; K)$ on K-kertoiminen $(n \times n)$-kokoinen matriisi (missä K on mielivaltainen

kunta), niin $\chi_A(A) = 0$ (nolla-matriisi).

Huomatus: Joskus Cayley-Hamiltonin Lauseelle esitetään seuraava houkuttelevan yksinkertaiselta näyttävä, mutta täysin virheellinen "todistus". Määritelmän mukaan pätee $\chi_L = \det(\mathbf{X}I - A)$. Sijoittamalla tässä muuttujasymbolin $\mathbf{X}$ paikalle matriisi A, saadaan

$$\chi_A(A) = \det(AI - A) = \det(A - A) = \det 0 = 0.$$

Tässä "todistuksessa" ei kuitenkaan ole mitään järkeä. Jätetään lukijalle pohdittavaksi miksi näin on.

Seuraus 3.87. *Olkoon $L\colon V \to V$ äärellisulotteisen K-vektoriavaruuden operaattori. Tällöin sen* minimipolynomi $\mathbf{m}_L$ *on sen karakteristisen polynomin χ_L tekijä, $\mathbf{m}_L \mid \chi_L$.*

Erityisesti minimipolynomin $\mathbf{m}_L$ aste on korkeintaan avaruuden V dimensio $n = \deg \chi_L$, $\deg \mathbf{m}_L \leq n$.

Lisäksi minimipolynomilla ja karakteristisella polynomilla on kunnassa K täsmälleen samat juuret.

Todistus. Määritelmän mukaan minimipolynomi $\mathbf{m}_L$ on ideaalin

$$I = \{\mathbf{p} \in K[\mathbf{X}] \mid p(L) = 0\}$$

virittäjä, jolloin $\mathbf{m}_L$ on jokaisen tämän ideaalin alkion $\mathbf{p}$ tekijä. Koska Cayley-Hamiltonin lauseen nojalla $\chi_L \in I$, minimipolynomi $\mathbf{m}_L$ on polynomin χ_L tekijä.

Erityisesi jokainen minimipolynomin juuri on myös karakteristisen polynomin juuri.

Kääntäen olkoon k jokin karakteristisen yhtälön χ_L juuri. Tällöin k on operaattorin L ominaisarvo, joten on olemassa ominaisvektori $\mathbf{v} \neq \mathbf{0}_V$, $L(\mathbf{v}) = k\mathbf{v}$. Tästä seuraa induktiolla, että jokaisella $n \in \mathbb{N}$ pätee $L^n(\mathbf{v}) = k^n\mathbf{v}$. Tästä puolestaan seuraa yleisemmin, että jokaiselle polynomialgebran $K[\mathbf{X}]$ alkiolle $\mathbf{p} = c_n\mathbf{X}^n + c_{n-1}\mathbf{X}^{n-1} + \ldots + c_1\mathbf{X} + c_0$ pätee

$$\begin{gathered} p(L)(\mathbf{v}) = (c_nL^n + c_{n-1}L^{n-1} + \ldots + c_1L + c_0\,\mathrm{id}_V)(\mathbf{v}) = \\ c_nL^n(\mathbf{v}) + c_{n-1}L^{n-1}(\mathbf{v}) + \ldots + c_1L(\mathbf{v}) + c_0(\mathbf{v}) = \\ c_nk^n(\mathbf{v}) + c_{n-1}k^n(\mathbf{v}) + \ldots + c_1k(\mathbf{v}) + c_0(\mathbf{v}) = p(k)\mathbf{v}. \end{gathered}$$

Koska $m_L(L) = 0$ minimipolynomin määritelmän mukaan, pätee erityisesti

$$\mathbf{0}_V = m_L(L)(\mathbf{v}) = m_L(k)\mathbf{v}.$$

Koska $\mathbf{v} \neq \mathbf{0}_V$, tästä seuraa, että $m_L(k) = 0$, eli k on minimipolynomin juuri. □

Tämäkin tulos voidaan ilmaista myös matriisien kielellä. Olkoon $A \in M(n \times n; K)$ K-kertoiminen neliömatriisi. Tällöin matriisin A minimipolynomi $\mathbf{m}_A$ on karakteristisen polynomin χ_A tekijä. Lisäksi kummallakin polynomilla on kunnassa K täsmälleen samat juuret.

Edellisen seurauksen tulos voidaan muotoilla konkreettisemmin seuraavasti. Olkoon $L\colon V \to V$ äärellisulotteisen K-vektoriavaruuden lineaarinen operaattori. Proposition 3.55 nojalla

$$\chi_L = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \ldots (\mathbf{X} - k_m)^{l_m}\mathbf{q},$$

missä $k_1, \ldots, k_m$ ovat operaattorin L kaikki (eri) ominaisarvot, $l_i \geq 1$ on ominaisarvon k_i algebrallinen kertaluku ja $\mathbf{q}$ on sellainen polynomialgebran $K[\mathbf{X}]$ alkio, jolla ei ole lainkaan juuria kunnassa K. Edellisestä tuloksesta tällöin seuraa, että

$$\mathbf{m}_L = (\mathbf{X} - k_1)^{l'_1}(\mathbf{X} - k_2)^{l'_2} \ldots (\mathbf{X} - k_m)^{l'_m}\mathbf{r},$$

missä $1 \leq l'_i \leq l_i$ jokaisella $i = 1, \ldots, m$ ja $\mathbf{r}$ on jokin polynomin $\mathbf{q}$ tekijä.

Yleisemmin voidaan osoittaa (sivutetaan tässä materiaalissa, todistettu kurssin harjoitustehtävien yhteydessä), että polynomeilla χ_L ja $\mathbf{m}_L$ on *täsmälleen samat jaottomat tekijät.*

Koska yleisesti ottaen karakteristinen polynomin laskeminen on suhteellisen ”helppoa” (ainakin suoraviivaista - lasketaan vain eräs determinantti), edellisen tuloksen avulla voidaan määrittää operaattorin (tai matriisin) minimipolynomi.

Esimerkkejä 3.88. *Olkoon*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Tällöin $\chi_A = (\mathbf{X} - 1)^2$, joten edellisen nojalla matriisin A minimipolynomi $\mathbf{m}_A$ on joko ensimmäisen asteen polynomi $(\mathbf{X} - 1)$ tai karakteristinen polynomi χ_A itse. Koska

$$(X - 1)(A) = A - I_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

ei ole nollamatriisi, polynomi $(\mathbf{X}-1)$ ei voi olla matriisin A minimipolynomi. Näin ollen täytyy olla

$$\mathbf{m}_A = \chi_A = (\mathbf{X} - 1)^2 = \mathbf{X}^2 - 2\mathbf{X} + 1.$$

Laskemalla voidaan tarkistaa (jätetään varsinainen lasku lukjalle), että tosiaankin

$$A^2 - 2A + 1 = 0.$$

Matriisin

$$B = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

karakteristinen polynomi χ_B on myös polynomi $(\mathbf{X} - 1)^2$, joten voidaan päätellä kuten yllä, että sen minimipolynomi on joko polynomi $\mathbf{X} - 1$ tai polynomi $\chi_B = (\mathbf{X} - 1)^2$. Tällä kertaa

$$(X - 1)(B) = B - I_2 = 0,$$

joten minimipolynomi on $\mathbf{X} - 1$. Minimipolynomi voi siis todellakin olla karakteristisen polynomin aito tekijä.

Matriisin

$$C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

karakteristinen polynomi χ_C on polynomi $\mathbf{X}^2+1$. Tarkastellaan matriisia C ensin kunnan $\mathbb{R}$ yli, eli matriisialgebran $M(2 \times 2; \mathbb{R})$ alkiona. Tällöin sen karakteristinen polynomi

$\chi_C = \mathbf{X}^2 + 1$ *on jaoton. Koska matriisin* C *minimipolynomi on ei-vakio pääpolynomi, joka on karakteristisen polynomin tekijä, täytyy olla*

$$\mathbf{m}_C = \chi_C = \mathbf{X}^2 + 1.$$

Jos matriisia C *ajatellaan sen sijaan matriisina kunnan* $\mathbb{C}$ *yli, eli matriisialgebran* $M(2 \times 2; \mathbb{C})$ *alkiona, johtopäätös on sama, mutta täysin eri perusteella. Nimittäin kunnan* $\mathbb{C}$ *yli karakteristinen polynomi* $\mathbf{X}^2 + 1 = (\mathbf{X} - i)(\mathbf{X} + i)$ *ei ole enää jaoton. Sillä on tasan kaksi juurta,* i *ja* $-i$*. Koska edellisen tuloksen mukaan minimipolynomilla on samat juuret, tämän täytyy olla jaollinen ainakin polynomilla* $(\mathbf{X} - i)(\mathbf{X} + i) = \mathbf{X}^2 + 1$*. Toisaalta minimipolynomi on saman tuloksen nojalla tämän polynomin tekijä. Koska minimipolynomi on pääpolynomi, myös kunnan* $\mathbb{C}$ *yli täytyy päteä*

$$\mathbf{m}_C = \chi_C = \mathbf{X}^2 + 1.$$

3.4. Jordanin normaalimuoto

Tässä aliluvussa palataan Luvun 3 alussa esitettyyn ongelmaan eli tutkitaan lineaaristen operaattorien *matriisiesityksiä*. Nyt ollaan kuitenkin varustettuja uudella työkalulla, kunnan K *polynomialgebralla* $K[\mathbf{X}]$.

Olkoon V K-vektoriavaruus ja olkoon $L\colon V \to V$ operaattori. Olkoon $\mathbf{p} = \sum_{i=0}^{n} k_i \mathbf{X}^i$ polynomialgebran $K[\mathbf{X}]$ alkio. Tällöin $p(L) = \sum_{i=0}^{n} k_i L^i$ on eräs vektoriavaruuden V lineaarinen operaattori $p(L)\colon V \to V$, joten voidaan puhua sen *ytimestä* $\operatorname{Ker} p(L)$. Kun operaattori L on kiinnitetty, merkitään jatkossa tätä ydintä myös symbolilla $V_{\mathbf{p}}$,

$$V_{\mathbf{p}} = \{\mathbf{v} \in V \mid p(L)(\mathbf{v}) = \mathbf{0}_V\} = \operatorname{Ker} p(L).$$

Kun $\mathbf{p} = \mathbf{X} - k$ on ensimmäisen asteen pääpolynomi, $k \in K$, avaruus $V_{\mathbf{p}}$ ei ole mitään muuta kuin alkiota k vastaava ominaisarvoaliavaruus V_k (joka on triviaali, jos k ei olekaan operaattorin ominaisarvo). Näin ollen, avaruutta $V_{\mathbf{p}}$ voi ajatella ominaisarvoaliavaruuden *yleistyksenä*.

Koska lineaarisen operaattorin ydin on aina aliavaruus, $V_{\mathbf{p}}$ on avaruuden V aliavaruus (mahdollisesti triviaali). Lisäksi tämä aliavaruus on *invariantti* operaattorin L suhteen. Nimittäin olkoon $\mathbf{v} \in V_{\mathbf{p}}$, eli $p(L)(\mathbf{v}) = \mathbf{0}_V$. Helposti nähdään, että operaattorit $p(L)$ ja L *kommutoivat*, eli $p(L)L = Lp(L)$ (tarkka todistus harjoitustehtävänä, huom., yleensä erilaiset operaattorit eivät kommutoi). Tästä seuraa, että

$$p(L)(L(\mathbf{v})) = L(p(L)(\mathbf{v})) = L(\mathbf{0}_V) = \mathbf{0}_V,$$

joten myös $L(\mathbf{v}) \in V_{\mathbf{p}}$.

Lemma 3.89. *Oletetaan, että polynomit* $\mathbf{p}, \mathbf{q} \in K[\mathbf{X}]$ *ovat keskenään jaottomia. Tällöin summa* $V_{\mathbf{p}} + V_{\mathbf{q}}$ *on* suora *ja* $V_{\mathbf{p}} \oplus V_{\mathbf{q}} = V_{\mathbf{pq}}$.

Todistus. Osoitetaan ensin, että $V_{\mathbf{p}} + V_{\mathbf{q}} = V_{\mathbf{pq}}$. Olkoon $\mathbf{v} \in V_{\mathbf{p}}$, jolloin $p(L)(\mathbf{v}) = \mathbf{0}_V$. Tällöin

$$(pq)(L)(\mathbf{v}) = (qp)(L)(\mathbf{v}) = (q(L)p(L))(\mathbf{v}) = q(L)(p(L)(\mathbf{v})) = q(L)(\mathbf{0}_V) = \mathbf{0}_V.$$

Näin ollen $V_{\mathbf{p}} \subset V_{\mathbf{pq}}$. Samalla tavalla nähdään, että $V_{\mathbf{q}} \subset V_{\mathbf{pq}}$. Koska $V_{\mathbf{pq}}$ on aliavaruus, erityisesti suljettu vektorien yhteenlaskun suhteen, näistä seuraa, että

$$V_{\mathbf{p}} + V_{\mathbf{q}} \subset V_{\mathbf{pq}}.$$

Koska $\mathbf{p}$ ja $\mathbf{q}$ ovat keskenään jaottomia, Lemman 3.45 nojalla on olemassa sellaiset polynomit $\mathbf{s}, \mathbf{t} \in K[\mathbf{X}]$ joille pätee

$$\mathbf{sq} + \mathbf{tp} = 1_K.$$

Soveltamalla tähän yhtälöön sijoitushomomorfismia S_L eli "sijoittamalla muuttujan $\mathbf{X}$ paikalle operaattori L", saadaan avaruuden V operaattorien välinen yhtälö

$$s(L)q(L) + t(L)p(L) = \mathrm{id}_V\,.$$

Alkioiden tasolla tämä yhtälö tarkoittaa sitä, että kaikilla $\mathbf{v} \in V$ pätee

$$s(L)q(L)(\mathbf{v}) + t(L)p(L)(\mathbf{v}) = \mathbf{v}.$$

Oletetaan, että $\mathbf{v} \in V_{\mathbf{pq}}$. Tällöin edellisen nojalla $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, missä $\mathbf{w}_1 = s(L)q(L)(\mathbf{v})$ ja $\mathbf{w}_2 = t(L)p(L)(\mathbf{v})$. Osoitetaan, että $\mathbf{w}_1 \in V_{\mathbf{p}}$ ja $\mathbf{w}_2 \in V_{\mathbf{q}}$. Tämä on suora lasku,

$$p(L)(\mathbf{w}_1) = p(L)s(L)q(L)(\mathbf{v}) = (psq)(L)(\mathbf{v}) = s(L)(pq)(L)(\mathbf{v}) = \mathbf{0}_V,$$

sillä $\mathbf{v} \in V_{\mathbf{pq}}$. Samalla tavalla nähdään, että $\mathbf{w}_2 \in V_{\mathbf{q}}$. Koska $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, on osoitettu, että $V_{\mathbf{pq}} \subset V_{\mathbf{p}} + V_{\mathbf{q}}$. Näin ollen

$$V_{\mathbf{p}} + V_{\mathbf{q}} = V_{\mathbf{pq}}.$$

Näytetään vielä, että summa $V_{\mathbf{p}} + V_{\mathbf{q}}$ on suora. Seurauksen 2.156 nojalla riittää osoittaa, että $V_{\mathbf{p}} \cap V_{\mathbf{q}} = \{\mathbf{0}_V\}$. Olkoon $\mathbf{v} \in V_{\mathbf{p}} \cap V_{\mathbf{q}}$, jolloin

$$p(L)(\mathbf{v}) = \mathbf{0}_V = q(L)(\mathbf{v}).$$

Edellä on osoitettu, että on olemassa polynomit $\mathbf{s}, \mathbf{t} \in K[\mathbf{X}]$ joille pätee

$$s(L)q(L) + t(L)p(L) = \mathrm{id}_V\,.$$

Laskemalla tämän yhtälön molemman puolen arvo vektorissa $\mathbf{v}$ saadaan

$$\mathbf{0}_V = (s(L)p(L) + t(L)q(L))(\mathbf{v}) = \mathbf{v}.$$

Näin ollen summa $V_{\mathbf{p}} + V_{\mathbf{q}}$ on suora. □

Tarkastellaan operaattorin $L\colon V \to V$ karakteristista polynomia χ_L. Kuten jokainen polynomi, tämä polynomi voidaan esittää (oleellisesti yksikäsitteisellä tavalla) jaottomien pääpolynomien tulona (Prop 3.42) eli muodossa

$$\chi_L = \mathbf{p}_1^{l_1}\mathbf{p}_2^{l_2}\dots\mathbf{p}_m^{l_m},$$

missä $\mathbf{p}_i$, $i = 1, \dots, m$, ovat (eri) jaottomat pääpolynomit ja $l_i \geq 1$ kaikilla i.

Propositio 3.90. *Olkoot $\mathbf{p}_i$ ja l_i, $i = 1, \dots, m$ kuten yllä. Tällöin seuraavat väitteet pitävät paikkansa.*

(1) $V = \oplus_{i=1}^{m} V_{\mathbf{p}_i^{l_i}}$.

(2) *Kaikilla* $i, j \in \{1, \ldots, m\}$ *aliavaruus* $V_{\mathbf{p}_i^{l_i}}$ *on* $p_j(L)$*-invariantti.*

(3) *Jos* $j \neq i$, *operaattorin* $p_j(L)$ *rajoittuma aliavaruuteen* $V_{\mathbf{p}_i^{l_i}}$ *on tämän aliavaruuden* isomorfismi.

(4) *Jokaisella* $i = 1, \ldots, m$ *pätee*

$$V_{\mathbf{p}_i^{l_i}} = \bigcup_{n=1}^{\infty} V_{\mathbf{p}_i^n}.$$

Todistus. Koska jokainen polynomi $\mathbf{p}_i$ on jaoton, nähdään helposti, että polynomit $\mathbf{p}_1^{l_1}$ ja $\mathbf{q}_1 = \mathbf{p}_2^{l_2} \ldots \mathbf{p}_m^{l_m}$ ovat keskenään jaottomia. Edellisen lemman nojalla tästä seuraa, että $V_{\chi_L} = V_{\mathbf{p}_1^{l_1}} \oplus V_{\mathbf{q}_1}$. Vastaavalla tavalla osoitetaan, että $V_{\mathbf{q}_1} = V_{\mathbf{p}_2^{l_1}} \oplus V_{\mathbf{q}_2}$, missä $\mathbf{q}_2 = \mathbf{p}_3^{l_3} \ldots \mathbf{p}_m^{l_3}$. Jatkamalla tällä tavalla *rekursiivisesti* saadaan osoitettua, että

$$V_{\chi_L} = \oplus_{i=1}^{m} V_{\mathbf{p}_i^{l_i}}.$$

Ensimmäinen väite seuraa, kunhan näytetään että $V_{\chi_L} = V$. Koska $V_{\chi_L} = \mathrm{Ker}(\chi_L(L))$, tämä on yhtäpitävää sen kanssa, että $\chi_L(L) \colon V \to V$ on nolla-operaattori. Tämä on puolestaan täsmälleen Cayley-Hamiltonin lauseen 3.84 väite. Kohta (1) on todistettu.

Olkoon W avaruuden V aliavaruus, joka on invariantti operaattorin L suhteen ja olkoon $\mathbf{q} \in K[\mathbf{X}]$ mikä tahansa polynomi. Tällöin W on myös invariantti operaattorin $q(L)$ suhteen (tarkka todistus harjoitustehtävänä). Erityisesti aliavaruus $V_{\mathbf{p}_i^{l_i}}$ on invariantti aliavaruuden $p_j(L)$ kaikilla $i, j = 1, \ldots, m$ ja rajoittuma $L_{i,j} = p_j(L)|\colon V_{\mathbf{p}_i^{l_i}} \to V_{\mathbf{p}_i^{l_i}}$ on avaruuden $V_{\mathbf{p}_i^{l_i}}$ hyvinmääritelty lineaarinen operaattori. Kohta (2) on osoitettu. Näytetään, että tämä operaattori on isomorfismi kun $i \neq j$. Koska äärellisulotteisen vektoriavaruuden operaattori on isomorfismi jos ja vain jos se on injektio, riittää näyttää, että operaattorin $L_{i,j}$ ydin on triviaali. Olkoon $\mathbf{v} \in V_{\mathbf{p}_i^{l_i}}$ sellainen, että $p_j(L)(\mathbf{v}) = \mathbf{0}_V$. Tällöin myös $p_j^{l_j}(L)(\mathbf{v}) = \mathbf{0}_V$ joten aliavaruuden $V_{\mathbf{p}_j^{l_j}}$ määritelmän nojalla pätee myös $\mathbf{v} \in V_{\mathbf{p}_j^{l_j}}$. Toisaalta, todistuksen ensimmäisen osan nojalla aliavaruudet $V_{\mathbf{p}_i^{l_i}}$ ja $V_{\mathbf{p}_j^{l_j}}$ *leikkaavat vain origossa*, sillä ne muodostavat (muiden avaruuksien $V_{\mathbf{p}_e^{l_e}}$ kera) suoran summan. Näin ollen $\mathbf{v} = \mathbf{0}_V$, toisin sanoen operaattorin $L_{i,j}$ ydin on triviaali. Väite (3) on todistettu.

Osoitetaan vielä, että

$$V_{\mathbf{p}_i^{l_i}} = \bigcup_{n=1}^{\infty} V_{\mathbf{p}_i^n} = \{\mathbf{v} \in V \mid \text{ on olemassa } n \in \mathbb{N}_+ \text{ siten, että } p_i^n(L)(\mathbf{v}) = \mathbf{0}_V\}$$

kaikilla $i = 1, \ldots, m$. Selvästi

$$V_{\mathbf{p}_i^{l_i}} \subset \bigcup_{k=1}^{\infty} V_{\mathbf{p}_i^k}.$$

Kääntäen olkoon $\mathbf{v} \in V$ sellainen, että $p_i^n(L)(\mathbf{v}) = \mathbf{0}_V$ jollakin $n \in \mathbb{N}_+$. Edellä todistetun nojalla $\mathbf{v} = \sum_{j=1}^m \mathbf{w}_j$, missä $\mathbf{w}_j \in V_{\mathbf{p}_j^{l_j}}$ kaikilla $j = 1, \ldots, m$. Tästä seuraa, että

$$\mathbf{0}_V = p_i^n(L)(\mathbf{v}) = \sum_{j=1}^m p_i^n(L)(\mathbf{w}_j),$$

missä edellä todistetun kohdan (2) nojalla pätee $p_i^n(L)(\mathbf{w}_j) \in V_{\mathbf{p}_j^{l_j}}$. Koska summa $\oplus_{i=1}^m V_{\mathbf{p}_i^{l_i}}$ on suora, tästä seuraa, että $p_i^n(L)(\mathbf{w}_j) = \mathbf{0}_V$ kaikilla $j = 1, \ldots, m$. Jos $j \neq i$, edellä todistetun nojalla $p_i(L)$ on isomorfismi, erityisesti injektio, aliavaruudessa $V_{\mathbf{p}_j^{l_j}}$. Tästä seuraa, että myös $p_i^n(L)$ on injektio aliavaruudessa $V_{\mathbf{p}_j^{l_j}}$, kun $j \neq i$, eli sen ydin on triviaali. Koska $p_i^n(L)(\mathbf{w}_j) = \mathbf{0}_V$, tästä saadaan, että $\mathbf{w}_j = \mathbf{0}_V$ kaikilla $j \neq i$. Näin ollen $\mathbf{v} = \mathbf{w}_i$ kuuluu aliavaruuteen $V_{\mathbf{p}_i^{l_i}}$, mitä pitikin todistaa. □

Koska jokainen aliavaruus $V_{\mathbf{p}_i}^{l_i}$ on L-invariantti, edellisestä propositiosta seuraa, että riittää tutkia operaattorin L käyttäytymistä aliavaruudessa $W_i = V_{\mathbf{p}_i^{l_i}}$ jokaisella $i = 1, \ldots, n$, eli operaattorin $L|W_i \colon W_i \to W_i$ käyttäytymistä.

Yleisesti voidaan osoittaa todeksi (sivutetaan tässä materiaalissa, katso kurssin harjoitustehtäviä) seuraavat edellä määriteltyjen aliavaruuksien $W_i = V_{\mathbf{p}_i^{l_i}}$ ominaisuudet.

- Aliavaruus $W_i = V_{\mathbf{p}_i^{l_i}}$ on epätriviaali jokaisella $i = 1, \ldots, n$. Itse asiassa pätee $\dim V_{\mathbf{p}_i^{l_i}} = \deg \mathbf{p}_i^{l_i} = l_i \deg \mathbf{p}_i$.
- Operaattorin $L_i = L|W_i$ karakteristinen polynomi on polynomi $\mathbf{p}_i^{l_i}$.
- Operaattorin $L_i = L|W_i$ minimipolynomi on polynomi $\mathbf{p}_i^{l_i'}$ jollakin $1 \leq l_i' \leq l_i$.
- Operaattorin L minimipolynomi on tällöin polynomi $\prod_{i=1}^n \mathbf{p}_i^{l_i'}$.

Emme siis tutki näitä väitteitä yleisesti, vaan tyydytään tarkastelemaan tarkemmin ainoastaan erikoistapausta, jossa jokainen karakteristisen polynomin χ_L jaoton tekijä $\mathbf{p}_i$ on *ensimmäisen asteen polynomi* eli polynomi muotoa $\mathbf{X} - k$ jollakin $k \in K$. Oletetaan siis, että

$$\chi_L = (\mathbf{X} - k_1)^{l_1} (\mathbf{X} - k_2)^{l_2} \ldots (\mathbf{X} - k_m)^{l_m},$$

missä $k_1, \ldots, k_m \in K$ ovat tällöin tunnetusti operaattorin L kaikki (eri) ominaisarvot. Edellisen proposition nojalla avaruus V on tällöin suora summa aliavaruuksista

$$V_{(\mathbf{X}-k_i)^{l_i}} = \operatorname{Ker}(L - k_i \operatorname{id}_V)^{l_i}.$$

Aliavaruutta $V_{(\mathbf{X}-k_i)^{l_i}}$ merkitään myös symbolilla V^{k_i} ja sitä sanotaan *yleistetyksi ominaisarvoaliavaruudeksi*. Vastaavalle ominaisarvoaliavaruudelle V_{k_i} selvästi pätee

$$V_{k_i} \subset V^{k_i}.$$

Propositiosta 3.90 saadaan heti seuraava tulos.

Propositio 3.91. *Olkoon $L\colon V \to V$ operaattori, jonka karakteristinen polynomi on jaettavissa ensimmäisen asteen tekijöihin. Olkoot $k_1, \ldots, k_m$ operattorin L kaikki eri ominaisarvot. Tällöin*

- $V = \oplus_{i=1}^{m} V^{k_i}$.
- *Kaikilla $i, j = 1, \ldots, m$ aliavaruus V^{k_j} on invariantti operaattorin $(L - k_i \operatorname{id}_V)$ suhteen.*
- *Jos $j \neq i$, operaattorin $(L - k_i \operatorname{id}_V)$ rajoittuma aliavaruuteen V^{k_j} on tämän aliavaruuden* isomorfismi.
- *Jokaisella $i = 1, \ldots, m$ pätee*

$$V^{k_i} = \{\mathbf{v} \in V \mid \text{ on olemassa } n \in \mathbb{N}_+ \text{ siten, että } (L - k_i)^n(\mathbf{v}) = \mathbf{0}_V\}.$$

"Oikeat" ominaisarvoaliavaruudet V_{k_i} muodostavat suoran summan (Lemma 3.7), mutta tämän suoran summan arvo, eli aliavaruus $\oplus_{i=1}^{m} V_{k_i}$, on koko avaruus V jos ja vain jos operaattori L on diagonalisoituva. Proposition 3.91 nojalla myös yleistetyt ominaisarvoaliavaruudet muodostavat aina suoran summan, mutta *tämän summan arvo on aina koko avaruus V*. Tästä seuraa, että operaattorin L käyttäytyminen palautuu sen rajoittuman käyttäyttymiseen jokaisessa yleistetyssä ominaisarvoaliavaruudessa. Näin ollen yleistetyt ominaisarvoaliavaruudet ikään kuin korjaavat ominaisarvoaliavaruuksiin liityviä edellä mainittuja "puutteitä" silloin, kun L ei olekaan diagonalisoituva. Korostetaan vielä kuitenkin, että nämä yleistetyt ominaisarvoaliavaruudet "kertovat operaattorista L kaiken" ainoastaan silloin, kun sen karakteristinen polynomi on jaettavissa ensimmäisen asteen tekijöihin. Tämä oletus pitää aina paikkaansa, jos kunta K on *algebrallisesti suljettu*, mutta ei yleisesti.

Siirrytään siis tutkimaan operaattorin L käyttäytymistä yleistetyssä ominaisarvoaliavaruudessa V^k, missä k on operaattorin L jokin ominaisarvo. Huomataan, että yleistetyssä ominaisarvoaliavaruudessa $V^k = \operatorname{Ker}(L - k \operatorname{id}_V)^l$ operaattorin $L' = L - k \operatorname{id}_V$ rajoittuma on tämän aliavaruuden **nilpotentti operaattori**, kts. esimerkki 3.68. Tästä seuraa, että seuraavaksi on luonnollista tarkastella tarkemmin nilpotenttin operaattorien teoriaa.

Nilpotentit operaattorit

Olkoon $L\colon V \to V$ *nilpotentti* operaattori ja olkoon m sen aste (kts. esimerkkiä 3.68). Palautetaan mieleen, että tällöin $L^m = 0$, mutta $L^i \neq 0$ kaikilla $i < m$. Tästä seuraa, että operaattorin L minimipolynomin täytyy olla polynomi $\mathbf{X}^m$. Korollarin 3.87 nojalla pätee $m \leq n = \dim V$.

Sama tulos seuraa myös seuraavasta Lemmasta 3.92, kun valitaan sen muotoilusta vektoriksi $\mathbf{v}$ jokin avaruuden V vektori, jolle pätee $L^{m-1}\mathbf{v} \neq \mathbf{0}_V$ (tällainen on olemassa nilpotentin operaattorin asteen määritelmän nojalla).

Olkoon $L\colon V \to V$ nilpotentti operaattori, jonka aste on $m \in \mathbb{N}$ ja olkoon $\mathbf{v} \in V$ mielivaltainen vektori. Koska pätee $L^m(\mathbf{v}) = \mathbf{0}_V$, on erityisesti olemassa sellaisia luonnollisia lukuja m', joille pätee $L^{m'}(\mathbf{v}) = \mathbf{0}_V$. **Pienintä** tällaista luonnollista lukua $m(\mathbf{v})$ sanotaan vektorin $\mathbf{v}$ **asteeksi** nilpotentin operaattorin L suhteen. Vektorin aste riippuu, tietysti,

tästä vektorista, eikä välttämättä ole sama kuin edellä määritelty operaattorin L aste m. Yleisesti selvästi $m(\mathbf{v}) \leq m$, mutta tämä epäyhtälö voi olla aito.

Lemma 3.92. *Olkoon $L\colon V \to V$ nilpotentti operaattori, missä V on äärellisulotteinen K-vektoriavaruus ja olkoon $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}_V$. Olkoon $m(\mathbf{v})$ vektorin $\mathbf{v}$ aste operaattorin L suhteen. Tällöin jono*

$$(\mathbf{v}, L\mathbf{v}, L^2\mathbf{v}, \ldots, L^{m(\mathbf{v})-1}\mathbf{v})$$

on vapaa. Sen virittämää $m(\mathbf{v})$-ulotteista avaruuden V aliavaruutta sanotaan vektorin $\mathbf{v}$ (ja tietysti operaattorin L) määräämäksi sykliseksi aliavaruudeksi. *Lisäksi pätee*

$$\mathrm{Span}(\mathbf{v}, L\mathbf{v}, L^2\mathbf{v}, \ldots, L^{m(\mathbf{v})-1}\mathbf{v}) = K[\mathbf{X}](L)\mathbf{v} = \{p(L)(\mathbf{v}) \mid \mathbf{p} \in K[\mathbf{X}]\}.$$

Todistus. Oletetaan, että

$$k_0\mathbf{v} + k_1L(\mathbf{v}) + \ldots + k_{m(\mathbf{v})-1}L^{m(\mathbf{v})-1}(\mathbf{v}) = \mathbf{0}_V. \tag{3.93}$$

On osoitettava, että $k_0 = k_1 = \ldots = k_{m(\mathbf{v})-1} = 0_K$. Tehdään vasta-oletus:esitys 3.93 on epätriviaali. Valitaan pienin luonnollinen luku $i \leq m(\mathbf{v}) - 1$ jolle pätee $k_i \neq 0_K$. Tällöin yhtälö 3.93 voidaan kirjoittaa muotoon

$$k_iL^i(\mathbf{v}) + \ldots + k_{m(\mathbf{v})-1}L^{m(\mathbf{v})-1}(\mathbf{v}) = \mathbf{0}_V.$$

Jakamalla tämä yhtälö nollasta eroavalla kunnan alkiolla k_i saadaan yhtälö muotoa

$$L^i(\mathbf{v}) + k'_{i+1}L^{i+1}(\mathbf{v}) + \ldots + k'_{m(\mathbf{v})-1}L^{m(\mathbf{v})-1}(\mathbf{v}) = \mathbf{0}_V.$$

Sovelletaan tämän yhtälön molempiin puoliin kuvausta $L^{m(\mathbf{v})-1-i}$ (joka on hyvin määritelty, sillä $i \leq m(\mathbf{v}) - 1$). Koska selvästi kaikilla $n \geq m(\mathbf{v})$ pätee

$$L^n(\mathbf{v}) = L^{n-m(\mathbf{v})}L^{m(\mathbf{v})}(\mathbf{v}) = \mathbf{0}_V,$$

tällöin saadaan yhtälö

$$L^{m(\mathbf{v})-1}(\mathbf{v}) = \mathbf{0}_V.$$

Tämä on kuitenkin ristiriidassa luvun $m(\mathbf{v})$ määritelmän kanssa. Näin ollen vasta-oletus johti ristiriitaan, joten esitys 3.93 on triviaali. Väite on todistettu.

Yhtälön

$$\mathrm{Span}(\mathbf{v}, L\mathbf{v}, L^2\mathbf{v}, \ldots, L^{m(\mathbf{v})-1}\mathbf{v}) = K[\mathbf{X}](L)\mathbf{v} = \{p(L)(\mathbf{v}) \mid \mathbf{p} \in K[\mathbf{X}]\}$$

todistaminen jätetään harjoitustehtäväksi. □

Esimerkki 3.94. *Olkoon $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ lineaarinen operaattori, jonka matriisi avaruuden $\mathbb{R}^3$ standardikannan $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ suhteen on*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Olkoon $\mathbf{v} = (x_1, x_2, x_3) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \neq \mathbf{0}$ *avaruuden* $\mathbb{R}^3$ *mielivaltainen nolla-vektorista eroava vektori. Tällöin* $L(\mathbf{v}) = x_3\mathbf{e}_2$ *ja* $L^2(\mathbf{v}) = \mathbf{0}_V$*, joten* $L^2 = 0$ *ja* $L \neq 0$*. Toisin sanoen* L *on nilpotentti operaattori, jonka minimipolynomi on* $\mathbf{X}^2$*. Jos* $x_3 = 0$*, vektorin* $\mathbf{v}$ *aste operaattorin* L *suhteen on yksi ja yhden alkion jono* $(\mathbf{v})$ *on edellisen lemman nojalla vapaa, mikä on muutenkin tunnettua (yksiö on vapaa jos ja vain jos se ei sisällä nolla-vektoria). Jos* $x_3 \neq 0$*, vektorin* $\mathbf{v}$ *aste operaattorin* L *suhteen on kaksi ja edellisen lemman nojalla jono* $(\mathbf{v}, L(\mathbf{v}))$ *on vapaa. Se virittää siis kaksiulotteisen syklisen aliavaruuden, joka on aito aliavaruus (koska koko avaruus on kolmeulotteinen).*

Olkoon $L\colon V \to V$ (nolla-operaattorista eroava) nilpotentti operaattori ja olkoon $\mathbf{v} \in V$ nolla-vektorista eroava vektori. Edellisen lemman nojalla on olemassa vektorin $\mathbf{v}$ määräämä *syklinen aliavaruus*

$$W = \operatorname{Span}(\mathbf{v}, L\mathbf{v}, L^2\mathbf{v}, \ldots, L^{m(\mathbf{v})-1}\mathbf{v}),$$

missä $m(\mathbf{v})$ on vektorin $\mathbf{v}$ aste operaattorin L suhteen. Aliavaruuden W dimensio on tällöin tasan $m(\mathbf{v})$ eli vektorin $\mathbf{v}$ aste.

Syklinen aliavaruus W on L-invariantti. Tämä johtuu siitä, että avaruudella W on kanta

$$(\mathbf{v}, L\mathbf{v}, L^2\mathbf{v}, \ldots, L^{m(\mathbf{v})-1}\mathbf{v}) = (\mathbf{e}_1, \ldots, \mathbf{e}_{m'}),$$

jonka alkioille pätee $L(\mathbf{e}_i) = \mathbf{e}_{i+1} \in W$, kun $i < m'$, ja $L(\mathbf{e}_{m'}) = \mathbf{0}_V \in W$. Koska W:n kannan alkioiden L-kuvat pysyvät aliavaruudessa W, myös kaikkien W:n alkioiden L-kuvat pysyvät aliavaruudessa W, toisin sanoen W on L-invariantti.

Jos avaruudessa V saattuu olemaan vektori $\mathbf{v}$, jonka aste $m(\mathbf{v})$ on sama kuin avaruuden V dimensio, tällöin sen määräämä syklinen aliavaruus W on välttämättä koko avaruus V. Yleensä näin onnekkaasti ei kuitenkaan käy, esimerkiksi esimerkissä 3.94 esiintyvän kuvauksen L tapauksessa jokaisen vektorin aste on korkeintaan kaksi, vaikka koko avaruuden dimensio on kolme.

Osoittautuu, että kuitenkin myös tällaisessa tapauksessa avaruus voidaan aina hajottaa syklisten aliavaruuksien *suoraksi summaksi*.

Propositio 3.95. *Olkoon* $L\colon V \to V$ *nilpotentti operaattori. Tällöin on olemassa sellaiset* $\mathbf{v}_1, \ldots, \mathbf{v}_l \in V$*, joilla avaruus* V *on suora summa* $\oplus_{i=1}^{l} V_i$*, missä* V_i *on vektorin* $\mathbf{v}_i$ *määrämä syklinen aliavaruus* $V_i = K[\mathbf{X}](L)(\mathbf{v}_i)$*,* $i = 1, \ldots, l$*.*

Todistus. Osoitetaan väite induktiolla avaruuden V dimension $\dim V = n$ suhteen. Jos $n = 0$ tai $n = 1$, väite on selvä, sillä tällöin L on välttämättä nolla-operaattori. Nimittäin nilpotentin operaattorin L aste on korkeintaan avaruuden V dimensio, eli tässä tapauksessa nolla tai yksi, joten pätee $L^1 = 0$.

Oletetaan, että väite on jo todistettu kaikissa tapauksissa, joissa avaruuden dimensio on korkeintaan $(n-1)$ ja oletetaan, että $\dim V = n > 1$. Olkoon $L\colon V \to V$ on nilpotentti operaattori. Koska L on nilpotentti, se ei voi olla surjektio. Nimittäin, jos se olisi surjektio, niin se olisi dimensiosyistä jopa isomorfismi (Seuraus 2.94). Tällöin myös sen jokainen potenssi L^m, $m \in \mathbb{N}$, olisi isomorfismi, erityisesti mikään kuvauksista L^m ei olisi voinut olla nollakuvaus.

Koska L ei ole surjektio, sen kuva $U = L(V)$ on avaruuden V *aito* aliavaruus, $\dim U < \dim V = n$. Selvästi on olemassa sellainen avaruuden V aliavaruus W, jolle pätee $U \subset W \subset V$ ja $\dim W = n - 1$. Tällöin

$$L(W) \subset L(V) = U \subset W,$$

joten aliavaruus W on invariantti operaattorin L suhteen ja rajoittuma $L' = L|W$ on hyvinmääritelty avaruuden W operaattori. Selvästi L' on nilpotentti, joten, koska $\dim W = n - 1$, voidaan soveltaa induktio-oletusta operaattoriin L'. On siis olemassa vektorit $\mathbf{v}_1, \ldots, \mathbf{v}_p \in W$ siten, että

$$W = \oplus_{j=1}^{p} W_j, \text{ missä} \tag{3.96}$$

$$W_j = K[\mathbf{X}](L)(\mathbf{v}_j) = \mathrm{Span}(\mathbf{v}_j, L(\mathbf{v}_j), \ldots, L^{m(\mathbf{v}_j)-1}(\mathbf{v}_j))$$

kaikilla $j = 1, \ldots, p$. Merkitään yksinkertaisuuden vuoksi $m(\mathbf{v}_j) = m_j$ kaikilla $j = 1, \ldots, p$. Permutoimalla vektoreiden järjestystä tarvittaessa, voidaan olettaa, että

$$m_1 \geq m_2 \geq \ldots \geq m_p. \tag{3.97}$$

Syy siihen, miksi haluamme asettaa luvut m_j tällaiseen järjestykseen selviää todistuksessa myöhemmin.

Koska avaruus W on $(n-1)$-ulotteinen, on olemassa sellainen vektori $\mathbf{u} \in V$ jolle pätee

$$W \oplus \mathrm{Span}(\mathbf{u}) = V.$$

Tämä nähdään esimerkiksi täydentämällä avaruuden W (jokin) kanta avaruuden V kannaksi. Itse asiassa vektoriksi $\mathbf{u}$ kelpaa mikä tahansa avaruuden V vektori, joka ei ole aliavaruudessa W. Koska $L(V) \subset W$, erityisesti pätee $L(\mathbf{u}) \in W$. Esityksestä 3.96 tällöin seuraa, että jokaisella $j = 1, \ldots, p$ on olemassa (yksikäsitteiset) $\mathbf{w}_j \in W_j$ siten, että

$$L(\mathbf{u}) = \mathbf{w}_1 + \mathbf{w}_2 + \ldots + \mathbf{w}_p. \tag{3.98}$$

Aliavaruuden W_j määritelmän nojalla jokaisella $j = 1, \ldots, p$ on olemassa lineaarinen kombinaatio

$$\mathbf{w}_j = k_{0j}\mathbf{v}_j + k_{1j}L(\mathbf{v}_j) + \ldots + k_{m_j-1,j}L^{m_j-1}(\mathbf{v}_j).$$

Tämä yhtälö voidaan kirjoittaa muotoon

$$\mathbf{w}_j = k_j\mathbf{v}_j + L(\mathbf{w}'_j),$$

missä $k_j = k_{0j}$ ja

$$\mathbf{w}'_j = k_{1j}\mathbf{v}_j + \ldots + k_{m_j-1,j}L^{m_j-2}(\mathbf{v}_j) \in W_j.$$

Sijoittamalla nämä yhtälöt yhtälöön 3.98 saadaan yhtälö

$$L(\mathbf{u}) = \sum_{j=1}^{p} k_j\mathbf{v}_j + L(\mathbf{w}),$$

missä $\mathbf{w} = \sum_{j=1}^{p} \mathbf{w}'_j \in W$ ja $k_1, \ldots, k_p \in K$. Tästä puolestaan seuraa, että

$$L(\mathbf{u} - \mathbf{w}) = \sum_{j=1}^{p} k_j\mathbf{v}_j,$$

joten, jos korvataan vektori $\mathbf{u}$ vektorilla $\mathbf{v} = \mathbf{u} - \mathbf{w}$, saadaan avaruudelle V esitys suorana summana $V = W \oplus \mathrm{Span}(\mathbf{v})$, missä $\mathbf{v}$ ei ole aliavaruuden W alkio, mutta pätee yhtälö muotoa

$$L(\mathbf{v}) = \sum_{j=1}^{p} k_j \mathbf{v}_j.$$

Jos kaikki kertoimet k_j, $i = 1, \ldots, n$ ovat tässä nolla-alkioita, pätee $L(\mathbf{v}) = \mathbf{0}_V$, jolloin 1-ulotteinen aliavaruus $\mathrm{Span}(\mathbf{v})$ on triviaalisti syklinen. Koska $V = W \oplus \mathrm{Span}(\mathbf{v})$ on tällöin suora summa syklisistä aliavaruuksista, väite on todistettu tässä tapauksessa.

Jos taas $L(\mathbf{v}) \neq \mathbf{0}_V$, niin välttämättä on olemassa ainakin yksi indeksi l jolle pätee $k_l \neq 0_K$. Valitaan pienin tällainen indeksi l, jolloin

$$L(\mathbf{v}) = k_l \mathbf{v}_l + \sum_{j=l+1}^{p} k_j \mathbf{v}_j.$$

Jakamalla tämä yhtälö nollasta eroavalla skalaarikunnan alkiolla k_l saadaan yhtälö

$$L(\mathbf{e}) = \mathbf{v}_l + \sum_{j=l+1}^{p} k'_j \mathbf{v}_j = \mathbf{v}_l + \mathbf{f},$$

missä $\mathbf{e} = k_l^{-1}\mathbf{v}$, $k'_j = k_j/k_l$ ja

$$\mathbf{f} = \sum_{j=l+1}^{p} k'_j \mathbf{v}_j \in \oplus_{j=l+1}^{p} W_j.$$

Koska vektorit $\mathbf{e}$ ja $\mathbf{v}$ eroavat toisistaan vain nollasta eroavalla skalaarikertoimella, pätee $\mathrm{Span}(\mathbf{e}) = \mathrm{Span}(\mathbf{v})$, joten

$$V = W \oplus \mathrm{Span}(\mathbf{v}) = W \oplus \mathrm{Span}(\mathbf{e}).$$

Koska $L(\mathbf{e}) = \mathbf{v}_l + \mathbf{f}$, seuraava idea on korvata vektori $\mathbf{v}_l$ vektorilla $\mathbf{v}'_l = \mathbf{v}_l + \mathbf{f}$ sekä aliavaruus W_l tämän vektorin määräämällä syklisellä aliavarudella $W'_l = K[\mathbf{X}](L)(\mathbf{v}'_l)$. Halutaan näyttää, että tämän korvauksen jälkeen alkuasetelmat pysyvät samoina. Täsmällisesti sanottuna väitetään seuraavaksi, että summa $\sum_{j\neq l} W_j + W'_l$ on suora ja sen arvo on aliavaruus W, eli

$$W = \oplus_{j\neq l} W_j \oplus W'_l.$$

Aloitetaan osoittamalla, että summan $\sum_{j\neq l} W_j + W'_l$ arvo on aliavaruus W. Koska $\mathbf{v}'_l = \mathbf{v}_l + \mathbf{f} \in W$ ja W on L-invariantti, on selvää, että $W'_l \subset W$. Koska $W_j \subset W$ kaikilla $j \neq p$, tästä seuraa, että $\sum_{j\neq l} W_j + W'_l$ on avaruuden W aliavaruus. Kääntäen jokaisella $i = 0, \ldots, m_l - 1$ pätee

$$L^i(\mathbf{v}_l) = L^i(\mathbf{v}'_l) - L^i(\mathbf{f}) \in \sum_{j\neq l} W_j + W'_l.$$

Koska vektorit $L^i(\mathbf{v}_l)$, $i = 0, \ldots, m_l - 1$ muodostavat aliavaruuden W_i kannan, tästä seuraa, että $W_l \subset \sum_{j\neq l} W_j + W'_l$. Tästä voidaan helposti päätellä, että

$$W = \sum_{j\neq l} W_j + W_l \subset \sum_{j\neq l} W_j + W'_l.$$

On siis näytetty, että $W = \sum_{j\neq l} W_j + W_l'$. Seuraavaksi näytetään, että tämän yhtälön toisella puolella esiintyvä summa onkin suora. Proposition 2.160 nojalla riittää osoittaa, että

$$\dim W = \sum_{j\neq l} \dim W_j + \dim W_l'. \tag{3.99}$$

Propositiosta 2.160 ja yhtälöstä 3.96 seuraa, että

$$\dim W = \sum_{j=1}^{p} \dim W_j.$$

Tästä seuraa, että yhtälö 3.99 on tosi jos ja vain jos $\dim W_l = \dim W_l'$. Koska toisaalta $\dim W_l = m_l = m(\mathbf{v}_l)$ on vektorin $\mathbf{v}_l$ aste operaatorin L suhteen ja vastaavasti $\dim W_l'$ on vektorin $\mathbf{v}_l'$ aste operaatorin L suhteen, riittää osoittaa, että vektorin $\mathbf{v}_l' = \mathbf{v}_l + \mathbf{f}$ aste operaattorin L suhteen on tasan m_l.

Koska $\mathbf{f} \in \oplus_{j=l+1}^{p} W_j$, epäyhtälökejusta 3.97 seuraa, että $L^{m_l}(\mathbf{f}) = \mathbf{0}_V$ (vasta tässä paljastuu, mistä syystä aikaisemmin ollaan järjestetty vektoreita $\mathbf{v}_i$ epäyhtälöketjun 3.97 mukaan). Toisaalta yhtä hyvin pätee $L^{m_l}(\mathbf{v}_l) = \mathbf{0}_V$, joten erityisesti pätee myös $L^{m_l}(\mathbf{v}_l + \mathbf{f}) = \mathbf{0}_V$. Toisaalta

$$L^{m_l-1}(\mathbf{v}_l + \mathbf{f}) = L^{m_l-1}(\mathbf{v}_l) + L^{m_l-1}(\mathbf{f}) \in W_l \oplus (\oplus_{j=l+1}^{p} W_j).$$

Koska tämä summa on suora ja koska $L^{m_l-1}(\mathbf{v}_l) \neq \mathbf{0}_V$, tästä seuraa, että $L^{m_l-1}(\mathbf{v}_l + \mathbf{f}) \neq \mathbf{0}_V$. On näytetty, että vektorin $\mathbf{v}_l' = \mathbf{v}_l + \mathbf{f}$ määräämän syklisen aliavaruuden W_l' dimensio on tasan m_l.

On osoitettu, että vektorin $\mathbf{v}_l$ korvaaminen vektorilla $\mathbf{v}_l'$ ei muuta alkuasetelmia, joten merkitään tätä vektoria jatkossa yksinkertaisesti $\mathbf{v}_l$. Tähän menneessä on siis osoitettu, että on olemassa sellainen $\mathbf{e} \in V$ ja sellainen indeksi $l = 1, \ldots, n$, että $(n-1)$-ulotteinen L-invariantti aliavaruus W voidaan esittää suorana summana $W = \oplus_{j=1}^{p} W_j$, missä W_j on erään vektorin $\mathbf{v}_j$ määräämä syklinen aliavaruus jokaisella $j = 1, \ldots, n$ ja lisäksi pätee $V = W \oplus \mathrm{Span}(\mathbf{e})$ sekä yhtälö

$$L(\mathbf{e}) = \mathbf{v}_l. \tag{3.100}$$

Täydennetään aliavaruus W_l aliavaruudeksi $V_l = W_l \oplus \mathrm{Span}(\mathbf{e})$, joka on vektorin $\mathbf{e}$ määrämä syklinen aliavaruus, sillä $(\mathbf{e}, L(\mathbf{e}), \ldots, L^{m(l)}(\mathbf{e}))$ on (yhtälön 3.100 nojalla) sen eräs kanta. Jos merkitään vielä $V_j = W_j$ kaikilla $j = 1, \ldots, p$, $j \neq l$, päädytään siihen, että $V = \oplus_{j=1}^{p} V_j$, missä jokainen aliavaruus V_j on syklinen aliavaruus. Proposition väite on todistettu. □

Määritelmä 3.101. *Olkoon $L\colon W \to W$ äärellisulotteisen vektoriavaruuden operaattori. Avaruuden W kantaa $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ sanotaan sykliseksi (operaattorin L suhteen) jos kaikilla $i = 1, \ldots, n$ pätee $L(\mathbf{e}_i) = \mathbf{e}_{i-1}$. Tässä sovitaan merkitsemään $\mathbf{e}_0 = \mathbf{0}_V$.*

Olkoon $L\colon V \to V$ nilpotentti operaattori, missä V on äärellisulotteinen vektoriavaruus. Ollaan osoitettu, että tällöin avaruus V voidaan esittää suorana summana $V = \oplus_{i=1}^{n} V_i$, missä jokainen aliavaruus V_i on L-invariantti syklinen aliavaruus. Tämä tarkoittaa puolestaan sitä, että jokaisella $i = 1, \ldots, n$ on olemassa sellainen vektori $\mathbf{v}_i \in V_i$

siten, että jono $(\mathbf{v}_i, L\mathbf{v}_i, L^2\mathbf{v}_i, \ldots, L^{m_i-1}\mathbf{v}_i)$ on avaruuden V_i kanta (missä $m_i = \dim V_i$). Tällä kannalla on seuraava ominaisuus - jokaisen sen alkion L-kuva on kannan seuraava alkio, lukuunottamatta kannan viimeistä alkiota, joka kuvautuu operaattorissa L nollavektoriksi. Tämä kanta ei ole syklinen edellä annetun määritelmän (3.101) mielessä, mutta jos sen alkioita kirjoitetaan käänteisessä järjestyksessä, eli kantana

$$(L^{m_i-1}\mathbf{v}_i, \ldots, L^2\mathbf{v}_i, L\mathbf{v}_i, \mathbf{v}_i),$$

tämä kanta on jo syklinen määritelmän (3.101) mielessä. Formaalilla tasolla on tapana käyttää syklisiä kantoja, koska matriisi tällaisen kannan suhteen on *yläkolmiomatriisi*[4]. Nimittäin nilpotentin operaattorin $L\colon W_i \to W_i$ matriisi syklisen kannan suhteen on matriisi muotoa

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Tämän matriisin diagonaalialkiot ovat nollia, alkiot heti niiden ”yläpuolella” ykkösiä ja muut alkiot myös nollia. Tällainen matriisi on esimerkki niin sanotusta *Jordanin solusta*.

Määritelmä 3.102. *Olkoon K kunta, olkoon $k \in K$ ja olkoon $n \geq 1$ kokonaisluku. Tällöin $(n \times n)$-kokoinen* Jordanin k-solu *$N(k,n)$ on $(n \times n)$-kokoinen matriisi*

$$N(k,n) = \begin{bmatrix} k & 1 & 0 & \ldots & 0 \\ 0 & k & 1 & \ldots & 0 \\ 0 & 0 & k & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 1 \\ 0 & 0 & 0 & \ldots & k \end{bmatrix}.$$

Täsmällisesti määriteltynä Jordanin solu $N(k,n)$ on siis sellainen $(n \times n)$-kokoinen matriisi $A = (a_{ij})_{i,j=1}^n$ jolle pätee

$$a_{ii} = k, \ j = 1, \ldots, n,$$

$$a_{i,i+1} = 1, \ i = 1, \ldots, n-1,$$

ja $a_{ij} = 0$ kun $j \neq i, i+1$. Tässä $1 = 1_K$ on kunnan ykkösalkio ja $0 = 0_K$ on kunnan nolla-alkio.

Esimerkkejä 3.103. *Matriisi*

$$\begin{bmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{bmatrix}$$

on (3×3) kokoinen Jordanin (-3)-solu. Matriisi

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{bmatrix}$$

ei ole Jordanin solu, koska sen $(1,2)$-alkio on nolla, eikä ykkönen.

[4]Jos olisimme pysyneet kannan järjestyksessä $(\mathbf{v}_i, L\mathbf{v}_i, L^2\mathbf{v}_i, \ldots, L^{m_i-1}\mathbf{v}_i)$, matriisi tämän kannan suhteen olisi ”alakolmiomatriisi”. On vakiintunut tapa käyttää tässä yhteydessä yläkolmiomatriiseja, mutta tietysti loppujen lopuksi kyse on ”makuasiasta”.

Määritelmä 3.104. *Neliömatriisin A sanotaan olevan* **Jordanin normaalimuodossa** *jos se on "suora summa" Jordanin soluista, eli jos se on lohkomatriisi muotoa*

$$A = \begin{bmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_l \end{bmatrix},$$

missä $A_i = N(k_i, m_i)$ on Jordanin k_i-solu, $k_i \in K$, $m_i \geq 1, l \in \mathbb{N}$, $i = 1, \dots, l$. Tällaista matriisia merkitään jatkossa myös jonona $(N(k_1, m_1), N(k_2, m_2), \dots, N(k_l, m_l))$.

Esimerkkejä 3.105. *Matriisi*

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

ei ole Jordanin solu, mutta se on Jordanin normaalimuodossa. Tämä johtuu siitä, että se voidaan kirjoittaa lohkomatriisina

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

missä $A_1 = \begin{bmatrix} 2 \end{bmatrix}$ on (1×1)-kokoinen Jordanin 2-solu ja

$$A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

on (2×2)-kokoinen Jordanin 1-solu. Pannaan erityisesti merkille, että Jordanin normaalimuodossa voi hyvinkin esiintyä (1×1)-kokoisia soluja muotoa $\begin{bmatrix} k \end{bmatrix}$. Esimerkiksi diagonaalimatriisi

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

on Jordanin normaalissa muodossa ja koostuu kolmesta (-3)-solusta, joista jokainen on (1×1)-solu. Itse asiassa samalla tavalla nähdään, että jokainen diagonaalimatriisi on Jordanin normaalimuodossa ja koostuu tällöin Jordanin (1×1)-soluista.

Tutkitaan seuraavaksi milloin annettu lineaarinen operaattori $L \colon V \to V$ voidaan esittää Jordanin normaalimuodossa jossakin avaruuden V kannassa. Jordanin normaalimuodon määritelmästä saadaan heti seuraava tapa luonnehtia tällaisia operaattoria.

Lemma 3.106. *Oletetaan, että V on äärellisulotteinen K-vektoriavaruus ja $L \in L(V)$ on lineaarinen operaattori. Tällöin avaruudella V on olemassa kanta E siten, että matriisi $[L]_E$ on Jordanin normaalimuodossa jos ja vain jos on olemassa avaruuden V aliavaruudet W_i, $i = 1, \dots, n$, siten, että*

- *W_i on L-invariantti jokaisella $i = 1, \dots, n$,*
- *$V = \oplus_{i=1}^{n} W_i$,*

- *operaattorin L rajoittuma $L|W_i\colon W_i \to W_i$ voidaan esittää jossakin avaruuden W_i kannassa Jordanin soluna.*

Propositiosta 3.95, Lemmasta 3.92 sekä edellisestä lemmasta saadaan seuraava tulos.

Propositio 3.107. *Olkoon $L\colon V \to V$ äärellisulotteisen K-vektoriavaruuden nilpotentti operaattori. Tällöin on olemassa avaruuden V kanta E, jonka suhteen operaattorin matriisi $[L]_E$ on Jordanin normaalimuodossa*

$$A = \begin{bmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_l \end{bmatrix},$$

missä jokainen Jordanin solu A_i on Jordanin 0_K-solu, $i = 1, \dots, l$.

Todistus. Proposition 3.95 nojalla $V = \oplus_{i=1}^{l} V_i$, missä V_i on jonkun vektorin $\mathbf{v}_i$ määrämä syklinen aliavaruus $V_i = K[\mathbf{X}](L)(\mathbf{v}_i)$ kaikilla $i = 1, \dots, l$. Tällaisella avaruudella on Lemman 3.92 nojalla olemassa syklinen kanta

$$(L^{m_i-1}(\mathbf{v}), \dots, L(\mathbf{v}_i), \mathbf{v}_i) = (\mathbf{e}_1, \dots, \mathbf{e}_{m_i}).$$

Tässä kannan alkioit kirjoitetaan toisessa järjestyksessä kuin Lemman 3.92 muotoilussa. Tämä johtuu siitä, että halutaan erityisesti määritelmän 3.101 mukainen kanta. Luonnollinen luku m_i on vektorin $\mathbf{v}$ aste nilpotentin operaattorin L suhteen. Syklisellä kannalla $(\mathbf{e}_1, \dots, \mathbf{e}_{m_i})$ on konstruktion perusteella ominaisuus

$$L(\mathbf{e}_i) = \begin{cases} \mathbf{0}_V, \text{ jos } i = 1, \\ \mathbf{e}_{i-1}, \text{ muuten.} \end{cases}$$

Tämä tarkoittaa sitä, että rajoittumalla $L|W_i\colon W_i \to W_i$ on tämän kannan suhteen matriisiesitys muotoa

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

eli Jordanin 0_K-solu. Väite seuraa tämän nojalla edellisestä lemmasta. □

Seuraus 3.108. *Olkoon $L\colon V \to V$ lineaarinen operaattori, missä V äärellisulotteinen K-vektoriavaruus. Tällöin L voidaan esittää jossakin avaruuden V kannassa E Jordanin normaalimuodossa jos ja vain jos operaattorin L karakteristinen polynomi on jaettavissa (polynomialgebrassa $K[\mathbf{X}]$) ensimmäisen asteen tekijöihin.*

Erityisesti, jos K on algebrallisesti suljettu kunta (esim. $K = \mathbb{C}$), niin jokainen äärellisulotteisen K-vektoriavaruuden operaattori voidaan esittää Jordanin normaalimuodossa.

Todistus. Oletetaan, että operaattorin $L\colon V \to V$ matriisi $A = [L]_E$ jonkun avaruuden V kannan E suhteen on Jordanin normaalimuodossa, eli muodossa

$$A = \begin{bmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_l \end{bmatrix}, \tag{3.109}$$

missä A_i on Jordanin k_i-solu jokaisella $i = 1, \dots, l$. Lasketaan matriisin A (ja siis myös operaattorin L) karakteristinen polynomi. Esityksestä 3.109 seuraa, että

$$\chi_L = \chi_A = \chi_{A_1}\chi_{A_2}\cdots\chi_{A_l},$$

joten riittää osata laskea Jordanin solun A_i karakteristinen polynomi jokaisella $i = 1, \dots, l$.

Olkoon B $(m \times m)$-kokoinen Jordanin k-solu $N(k, m)$, missä $k \in K$. Tällöin

$$B = \begin{bmatrix} k & 1 & 0 & \dots & 0 \\ 0 & k & 1 & \dots & 0 \\ 0 & 0 & k & \dots & 0 \\ \dots & \dots & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & k \end{bmatrix}$$

on *yläkolmiomatriisi*, jonka jokainen diagonaalialkio on k. Tästä seuraa, että matriisi $\mathbf{X}I_n - B$ on myös $(m \times m)$-kokoinen yläkolmiomatriisi, jonka jokainen diagonaalialkio on $\mathbf{X} - k$. Näin ollen

$$\chi_B = \det(\mathbf{X}I_n - B) = (\mathbf{X} - k)^m.$$

Edellisen nojalla Jordanin normaalissa muodossa olevan matriisin

$$A = (N(k_1, m_1), N(k_2, m_2), \dots, N(k_l, m_l))$$

karakteristinen polynomi on

$$\chi_A = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \dots (\mathbf{X} - k_m)^{l_m},$$

erityisesti se on *jaettavissa ensimmäisen asteen tekijöihin* polynomialgebrassa $K[\mathbf{X}]$. Ehdon *välttämättömyys* on osoitettu.

Oletetaan kääntäen, että operaattorin L karakteristinen polynomi χ_L voidaan esittää ensimmäisen asteen tekijöiden tulona eli muodossa

$$\chi_L = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \dots (\mathbf{X} - k_m)^{l_m}.$$

Proposition 3.90 nojalla pätee $V = \oplus_{i=1}^m W_i$, missä $W_i = V_{(\mathbf{X}-k_i)^{l_i}} = \operatorname{Ker}(L - k_i \operatorname{id}_V)^{l_i}$ on ominaisarvoon k_i liittyvä *yleistetty ominaisarvoaliavaruus*. Koska summa $\oplus_{i=1}^m W_i$ on suora ja jokainen aliavaruus W_i on L-invariantti, riittää osoittaa, että rajoittuma $L_i = L|W_i$ voidaan esittää Jordanin normaalimuodossa jokaisella $i = 1, \dots, n$. Avaruuden W_i määritelmän nojalla pätee $(L_i - k_i \operatorname{id}_{W_i})^{l_i} = 0$, joten operaattori $L_i' = L_i - k_i \operatorname{id}_{W_i}$ on avaruuden W_i *nilpotentti* operaattori. Edellisen proposition 3.107 nojalla L_i' voidaan esittää jossakin

avaruuden W_i kannassa E_i Jordanin normaalimuodossa, missä lisäksi jokainen solu on 0-solu, toisin sanoen

$$A = [L_i']_{E_i} = \begin{bmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_n \end{bmatrix},$$

missä

$$A_j = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

on Jordanin 0-solu kaikilla $j = 1, \dots, n$.

Operaattorin $k_i \operatorname{id}_{W_i}$ (joka on skalaarikerrointa vaille identtinen operaattori) esitys kannassa E_i on diagonaalimatriisi B, jonka jokainen diagonaalialkio on k_i. Tästä seuraa, että operaattorin $L_i \colon W_i \to W_i$ esitys kannassa E_i on matriisi $A + B$. Tämä matriisi on tällöin selvästi Jordanin normaalimuodossa oleva matriisi, jonka jokainen solu on k_i-solu. □

Seuraavaksi tutkitaan Jordanin normaalimuodon *yksikäsitteisyyttä*. Oletetaan, että lineaarisen operaattorin $L \colon V \to V$ karakteristinen polynomi on jaettavissa polynomialgebrassa $K[\mathbf{X}]$ ensimmäisen asteen tekijöihin, jolloin operaattori L voidaan esittää jossakin kannassa Jordanin normaalimuodossa. Tällainen kanta ei välttämättä ole yksikäsitteinen. Voiko operaattorin esitys Jordanin normaalissa muodossa olla kuitenkin kannasta riippumaton?

On selvä, että aivan kirjaimellisesti otettuna Jordanin normaalimuoto ei yleensä voi olla yksikäsitteinen siitä yksinkertaisesta syystä, että siinä esiintyviä Jordanin soluja voidaan *permutoida* eli esittää missä tahansa järjestyksessä, jolloin saadaan erinäköisiä matriiseja, jotka ovat kaikki kuitenkin Jordanin normaalissa muodossa. Tarkastellaan esimerkkinä operaattoria $L \colon \mathbb{C}^4 \to \mathbb{C}^4$, jonka matriisi standarikannan $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ suhteen on matriisi

$$A = \begin{bmatrix} i & 1 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Tämä on Jordanin normaalimuodossa. Kannan $E' = (\mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_1, \mathbf{e}_2)$ suhteen tämän operaattorin matriisi on

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & i & 1 \\ 0 & 0 & 0 & i \end{bmatrix}.$$

Tämäkin on Jordanin normaalimuodossa, mutta solut ovat eri järjestyksessä kuin matriisissa A.

Osoittautuu kuitenkin, että solujen permutaatiota vaille Jordanin normaalimuodossa oleva esitys lineaariselle operaattorille on yksikäsitteinen.

Propositio 3.110. *Olkoon $L\colon V \to V$ lineaarinen operaattori, missä V on äärellisulotteinen K-vektoriavaruus. Oletetaan, että operaattorin L karakteristinen polynomi on jaettavissa ensimmäisen asteen tekijöihin, jolloin se voidaan esittää Jordanin normaalimuodossa. Tällainen esitys on yksikäsitteinen siinä esiintyvien Jordanin solujen permutaatiota vaille.*

Toisin sanoen, jokaisessa operaattorin L Jordanin normaalimuodossa olevassa matriisiesityksessä on jokaisella $n \in \mathbb{N}_+$ ja $k \in K$ täsmälleen sama määrä $(n \times n)$-kokoisia k-soluja.

Todistus. Olkoon $L\colon V \to V$ sellainen operaattori, joka voidaan esittää Jordanin normaalimuodossa olevana matriisina $A = (N(k_1, l_1), N(k_2, l_2), \ldots, N(k_l, l_m))$ eli sellainen operaattori, jonka karakteristinen polynomi on jaettavissa ensimmäisen asteen tekijöihin. Huomataan, että tässä voi käydä niin, että $k_i = k_j$ joillakin $i \neq j$ (esimerkiksi nilpotentin operaattorin Jordanin esityksen kaikki solut ovat 0_K-soluja). Laskemalla esityksen A avulla operaattorin L karakteristinen polynomi, saadaan

$$\chi_L = (\mathbf{X} - k_1)^{l_1}(\mathbf{X} - k_2)^{l_2} \ldots (\mathbf{X} - k_l)^{l_m}. \tag{3.111}$$

Tästä seuraa, että kunnan alkiot k_i, $i = 1, \ldots, n$, ja ainoastaan ne, ovat operaattorin L *ominaisarvoja*. Nämä tunnetusti eivät riipu operaattorin matriisiesityksestä. Korostetaan vielä kerran kuitenkin, että esityksessä (3.111) ensimmäisen asteen tekijät $(\mathbf{X} - k_i)$ *eivät* välttämättä ole eri tekijöitä eri indekseillä $i = 1, \ldots, n$, joten me emme voi tehdä tästä automaattisesti johtopäätöstä, että myös kertoimet l_i riippuvat ainoastaan polynomista χ_L (joka tunnetusti riippuu ainoastaan operaattorista L, eikä sen matriisiesityksestä). Näin olisi mahdollista menetellä Proposition 3.48 (polynomien hajotelma jaottomiin tekijöihin) avulla, **jos** alkiot k_i olisivat varmasti erilaisia. Koska näin ei kuitenkaan välttämättä ole, ainoa mitä tässä vaiheessa voidaan päätellä Propositiosta 3.48 on se, että summa $l_{i_1} + l_{i_2} + \ldots + l_{i_t}$, missä $k_{i_1}, k_{i_2}, k_{i_t}$ ovat ominaisarvon k kaikki esiintymiset alkioiden $k_1, \ldots, k_l$ kesken, riippuu vain operaattorista L. Jordanin normaalissa muodossa olevan matriisin muoto riippuu kuitenkin kaikista luvuista $l_{i_1}, l_{i_2}, \ldots, l_{i_t}$, eikä ainoastaan niiden summasta. Esimerkiksi Jordanin normaalimuodossa olevat matriisit

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ovat kaikki oleellisesti erilaisia, vaikka kaikilla kolmella on täsmälleen sama karakteristinen polynomi $(\mathbf{X} - 1)^3$. Ero on 1-solujen *määrässä* sekä *koossa*, nimittäin ensimmäisellä matriisilla on yksi (3×3)-kokoinen 1-solu, toisella yksi (1×1)-kokoinen 1-solu ja yksi (2×2)-kokoinen 1-solu, kolmannella kolme (1×1)-kokoista 1-solua.

Näin ollen väitteen todistamiseksi täytyy vielä näyttää, että jokaiselle operaattorin L ominaisarvolle k ja jokaiselle positiiviselle kokonaisluvulle $n \in \mathbb{N}_+$ $(n \times n)$-kokoisten k-solujen *lukumäärä* operaattorin L Jordanin normaalimuodossa olevassa matriisiesityksessä *ei riipu* kannan valinnasta.

Olkoon k operaattorin L ominaisarvo ja olkoon $m \in \mathbb{N}_+$. Merkitään symbolilla s_m $(m \times m)$-kokoisten k-solujen lukumäärä jossakin operaattorin L Jordanin normaalismuodossa olevassa matriisiesityksessä $[L]_E = A$. Tarkoitus on "sitoa" luku s_m operaattoriin L

invariantilla tavalla, eli sellaisella, joka riippuu vain kuvauksesta L, eikä kannan E valinnasta. Määritelmän mukaan $s_m \geq 0$ (tapaus $s_m = 0$ esiintyy silloin kun A ei sisällä yhtäkään $(m \times m)$-kokoisia k-soluja lainkaan).

Huomataan ensin seuraava. Olkoot $k'_1, \ldots, k'_s$ kaikki operaattorin L *eri* ominaisarvot. Proposition 3.90 nojalla pätee $V = \oplus_{i=1}^{s} V^{k'_i}$, missä $V^{k'_i} = V_{(\mathbf{X}-k'_i)^{l(k'_i)}}$ on operaattorin L ominaisarvoon k'_i liittyvä *yleistetty ominaisarvoaliavaruus*. Permutoimalla tarvittaessa Jordanin soluja, voidaan olettaa, että operaattorin L Jordanin normaalimuodossa oleva matriisiesitys A on muotoa

$$(N(k'_1, l'_{1,1}), N(k'_1, l'_{1,2}), \ldots, N(k'_1, l'_{1,m_1}), N(k'_2, l'_{2,1}), \ldots, N(k'_2, l'_{2,m_2}), \ldots, N(k'_s, l'_{s,1}), \ldots, N(k'_s, l'_{s,m_s})).$$

Tässä jonossa siis ensin luetellaan kaikki k'_1-solut, sitten kaikki k'_2 solut jne. kunnes viimeisenä tulevat kaikki k_s-solut. Luku m_i ilmaisee k'_i-solujen (mahdollisesti erikokoisten) lukumäärän. Lemman 3.106 nojalla voidaan kirjoittaa

$$V = \oplus_{i=1}^{s}(\oplus_{j_i=1}^{m_i} W_{k'_i, l'_{i,j_i}}) = \oplus_{i=1}^{s} W_{k'_i},$$

missä $W_{k'_i, l'_{i,j_i}}$ on L-invariantti aliavaruus, joka vastaa yhtä solua ja jossa operaattorin L (rajoittuman) matriisesitys on eräs $(l'_{i,j_i} \times l'_{i,j_i})$-kokoinen k'_i-solu. Suoraa summaa $\oplus_{j_i=1}^{m_i} W_{k'_i, l'_{i,j_i}}$, jossa kootaan yhteen kaikki samaa ominaisarvoa vastaavat aliavaruudet $W_{k'_i, l'_{i,j_i}}$ on yllä merkitty symbolilla $W_{k'_i}$, $i = 1, \ldots, s$.

Jordanin solun määritelmästä seuraa, että operaattori $(L - k'_i \operatorname{id}_V)$ on nilpotentti aliavaruudessa $W_{k'_i, l'_{i,j_i}}$. Toisin sanoen erityisesti jokaisella $\mathbf{v} \in W_{k'_i, l'_{i,j_i}}$ on olemassa $n \in \mathbb{N}_+$ siten, että $(L - k'_i \operatorname{id}_V)^n(\mathbf{v})$. Propositiosta 3.90 seuraa tällöin, että $\mathbf{v} \in V^{k'_i}$. Näin ollen $W_{k'_i, l'_{i,j_i}} \subset V^{k_i}$ kaikilla $j_i = 1, \ldots, m_i$. Tästä seuraa, että myös aliavaruudelle $W_{k'_i} = \oplus_{j_i=1}^{m_i} W_{k'_i, l'_{i,j_i}}$ pätee $W'_{k'_i} \subset V^{k'_i}$. Koska toisaalta

$$\oplus_{i=1}^{s} W_{k'_i} = V = \oplus_{i=1}^{s} V^{k'_i},$$

täytyy olla $W_{k'_i} = V^{k_i}$ kaikilla $i = 1, \ldots, s$ (mieti miksi). Toisin sanoen missä tahansa Jordanin normaalimuodossa olevassa matriisiesityksessä aliavaruus, joka vastaa kaikkia k'_i-soluja on täsmälleen yleistetty ominaisarvoalivaruus $V^{k'_i}$. Erityisesti tämä aliavaruus ei riipu esityksestä (eli kannan valinnasta). Lisäksi tähän aliavaruuteen rajoitettuna operaattorin L Jordanin normaalimuodossa oleva esitys sisältää ainoastaan k'_i-soluja.

Näin ollen voidaan rajoittua tapaukseen, jossa tämän proposition oletukset täyttävällä operaattorilla on vain yksi ominaisarvo k. Tällöin täytyy osoittaa, että jokaisella $n \in \mathbb{N}_+$ $(n \times n)$-kokoisten k-solujen lukumäärä s_n ei riipu kannan valinnasta. Olkoon

$$A = \begin{bmatrix} A_1 & 0 & 0 & \ldots & 0 \\ 0 & A_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & A_l \end{bmatrix}$$

operaattorin L esitys Jordanin normaalimuodossa, missä nyt jokainen matriisi A_i on Jordanin k-solu. Tämän esityksen olemassaolo tarkoittaa sitä, että avaruus V voidaan

esittää suorana summana $\oplus_{i=1}^{l} V_i$, missä V_i on L-invariantti aliavaruus, jossa rajoittuma-operaattori $L|V_i$ voidaan esittää Jordanin k-soluna. Tämän solun koko on tällöin $(n_i \times n_i)$, missä $n_i = \dim V_i$. Tämä puolestaan tarkoittaa sitä, että operaattorin $S = (L - k\,\mathrm{id})$ rajoittuma aliavaruuteen V_i on nilpotentti operaattori, joka voidaan esittää Jordanin 0-soluna. Tämä taas implikoi, että aliavaruudella V_i on olemassa syklinen kanta muotoa $(S^{m_i-1}(\mathbf{v}), \ldots, S(\mathbf{v}), \mathbf{v})$, missä m_i on sellainen, että $S^{m_i}(\mathbf{v}) = \mathbf{0}_V$.

Olkoon $t \geq 0$ kiinteä luonnollinen luku. Lasketaan edellisten havaintojen avulla avaruuden V operaattorin $S^t = (L - k\,\mathrm{id}_V)^t$ kuvan dimensio $r_t = \dim \mathrm{Im}(L - k\,\mathrm{id}_V)^t$. Jokaisella $i = 1, \ldots, n$ operaattori $S^t = (L - k\,\mathrm{id}_V)^t$ selvästi kuvaa kannan $(S^{m_i-1}(\mathbf{v}), \ldots, S(\mathbf{v}), \mathbf{v})$ virittämän aliavaruuden V_i aliavaruudeksi, jonka virittää jono

$$(S^{m_i-1}(\mathbf{v}), \ldots, S^t(\mathbf{v})).$$

Jos $t \geq m_i$ tämä jono tulkitaan tyhjäksi, mikä on yhtäpitävää sen kanssa, että tällaisessa tapauksessa S^t kuvaa aliavaruuden V_i jokaisen vektorin nolla-vektoriksi. Jokaisessa tapauksessa jono $(S^{m_i-1}(\mathbf{v}), \ldots, S^t(\mathbf{v}))$ on vapaan jonon $(S^{m_i-1}(\mathbf{v}), \ldots, S(\mathbf{v}), \mathbf{v})$ osajono, joten se on myös vapaa. Tästä seuraa, että jono $(S^{m_i-1}(\mathbf{v}), \ldots, S^t(\mathbf{v}))$ on kuvan $S^t(V_i)$ kanta. Tästä voidaan päätellä, että

$$\dim S^t(V_i) = m_i - t = \dim V_i - t,$$

jos $t < \dim V_i$, muuten $\dim S^t(V_i) = 0$. Koska jokainen aliavaruus V_i on S^t-invariantti ja summa $\oplus_{i=1}^{m} V_i$ on suora, summa $\oplus_{i=1}^{m} S^t(V_i) = S^t(V)$ on myös suora, joten

$$r_t = \dim \mathrm{Im}\, S^t = \sum_{i=1}^{m} \dim S^t(V_i).$$

Edellisestä seuraa toisaalta, että $\dim S^t(V_i) = 0$ jos V_i vastaa Jordanin solua, jonka koolle $(n_i \times n_i)$ pätee $n_i \leq t$. Toisin sanoen

$$r_t = \sum_{n_i > t} (n_i - t),$$

missä summataan yli matriisin A kaikkien Jordanin $(n_i \times n_i)$ solujen, joille $n_i > t$. Tästä seuraa, että kaikilla $t \geq 0$ pätee

$$r_t - r_{t+1} = \sum_{n_i > t} (n_i - t) - \sum_{n_i > t+1} (n_i - t - 1) =$$

$$\sum_{n_i > t} (n_i - t) - \sum_{n_i > t+1} (n_i - t) + \sum_{n_i > t+1} 1 =$$

$$\sum_{n_i = t+1} 1 + \sum_{n_i > t+1} 1 = \sum_{n_i \geq t+1} 1 = s_{t+1} + s_{t+2} + \ldots,$$

missä s_n on matriisin A $(n \times n)$-kokoisten Jordanin solujen lukumäärä jokaisella $n \in \mathbb{N}$. Tästä saadaan kaikilla $t \geq 1$ yhtälö

$$s_t = (r_{t-1} - r_t) - (r_t - r_{t+1}) = r_{t-1} - 2r_t + r_{t+1}.$$

Tässä $r_0 = \dim V = \dim \mathrm{Im}\, S^0 = \dim \mathrm{id}_V$. Koska suureet tämän yhtälön oikealla puolella riippuvat vain operaattorista L, ei tietystä tavasta esittää se Jordanin matriisina, sama pätee yhtälön vasemalle puoleelle, eli luvulle s_t. Tämä luku taas ilmaisee matriisin Jordanin $(t \times t)$-solujen lukumäärän. □

Minimipolynomi Jordanin normaalimuodon avulla
Olkoon

$$\chi_L = \prod_{i=1}^{k} (\mathbf{X} - k_i)^{m_i}$$

operaattorin $L\colon V \to V$ karakteristinen polynomi, joka voidaan jakaa ensimmäisen asteen tekijöihin. Tällöin L voidaan esittää Jordanin normaalimuodossa olevana matriisina A. Jos $n_1^i, \ldots, n_{s_i}^i$ ovat erilaisten k_i-solujen koot tässä matriisissa, niin pätee $n_1^i + \ldots + n_{s_i}^i = m_i$ (mikä nähdään helposti laskemalla χ_L matriisin A avulla). Minimaalinen polynomi, joka nollaa $(n \times n)$-kokoisen Jordanin k-solun on polynomi $(\mathbf{X} - k)^n$. Tästä nähdään helposti, että operaattorin L minimipolynomi on itse asiassa polynomi

$$\mathbf{m}_L = \prod_{i=1}^{k} (\mathbf{X} - k_i)^{m_i'}, \tag{3.112}$$

missä $m_i' = \max\{n_1^i, n_2^i, \ldots, n_{s_i}^i\}$ on suurin kaikista luvuista n, joilla matriisissa A esiintyy $(n \times n)$-kokoinen Jordanin k_i-solu.

Jokainen diagonaalimatriisi on Jordanin normaalimuodossa, jossa jokaisen solun koko on (1×1). Kääntäen jokainen operaattori, jonka Jordanin normaalimuodossa olevassa matriisiesityksessä esiintyy vain (1×1)-soluja on diagonalisoituva. Tämän sekä Lemman 3.89 avulla voidaan todistaa seuraava kätevä karakterisaatio operaattorin diagonaalisoituvuudelle (todistus harjoitustehtävänä).

Propositio 3.113. *Olkoon $L\colon V \to V$ operaattori. Tällöin se on diagonalisoituva jos ja vain jos sen minimipolynomi on muotoa*

$$\mathbf{m}_L = \prod_{i=1}^{m} (\mathbf{X} - k_i),$$

missä k_i ovat kunnan eri alkiot. Toisin sanoen operaattori L on diagonalisoituva, jos sen minimipolynomi on jaettavissa ensimmäisen asteen tekijöihin ja minimipolynomin jokaisen juuren kertaluku on yksi.

Tulos on käytännössä hyödyllinen, sillä joskus minimipolynomi voidaan laskea suoraan, jolloin sen avulla voidaan päätellä operaattorin olevan diagonalisoituva ilman, että varsinainen diagonalisaatio löydetään.

Esimerkki 3.114. *Olkoon $L\colon \mathbb{Z}_7^6 \to \mathbb{Z}_7^6$ lineaarinen operaattori äärellisulotteisessa $\mathbb{Z}_7$-vektoriavaruudessa $\mathbb{Z}_7^6$. Oletetaan, että*

$$\chi_L = (\mathbf{X} - 3_7)^2 (\mathbf{X} - 2_7)^4,$$

$$\mathbf{m}_L = (\mathbf{X} - 3_7)^2 (\mathbf{X} - 2_7)^2.$$

Tutkitaan minkälainen operaattorin L Jordanin normaalimuodossa oleva matriisiesitys A voisi olla. Operaattorin karaktersistinen polynomi on jaettavissa ensimmäisen asteen tekijöihin ja operaattorin ominaisarvot ovat 2_7 ja 3_7. Tästä seuraa, että kaikki Jordanin solut matriisissa A ovat joko 2_7-soluja tai 3_7-soluja. Karakteristisesta polynomista

nähdään, että matriisissa A luku 2_7 esiintyy diagonaalilla tasan neljä kertaa ja luku 3_7 esiintyy diagonaalilla tasan kaksi kertaa.

Tästä seuraa, että A sisältää tasan yhden (2×2)-kokoisen 3_7-solun tai tasan kaksi (1×1)-kokoista 3_7-solua. Jälkimmäisessä tapauksessa yhtälöstä 3.112 seuraisi, että operaattorin L minimipolynomi sisältäisi tekijän $(\mathbf{X} - 3_7)$ vain kerran, mikä on vastoin oletuksiamme. Näin ollen A sisältää tasan yhden (2×2)-kokoisen 3_7-solun.

Mietitään seuraavaksi 2_7-soluja. Koska luku 2_7 esiintyy diagonaalilla tasan neljä kertaa, vaihtoehdot ovat:

(1) Neljä (1×1)-kokoista 2_7-solua.

(2) Kaksi (1×1)-kokoista 2_7-solua ja yksi (2×2)-kokoinen 2_7-solua.

(3) Kaksi (2×2)-kokoista 2_7-solua.

(4) Yksi (1×1)-kokoinen 2_7-solu ja yksi (3×3)-kokoinen 2_7-solu.

(5) Yksi (4×4)-kokoinen 2_7-solu.

Tarkistetaan yhtälön (3.112) avulla mitkä vaihtoehdot ovat mahdolliset. Vaihtoehdossa (1) tekijä $(\mathbf{X} - 2_7)$ esiintyisi minimipolynomissa tasan kerran. Vaihtoehdossa (4) tekijä $(\mathbf{X} - 2_7)$ esiintyisi minimipolynomissa kertoimella kolme. Vaihtoehdossa (5) tekijä $(\mathbf{X} - 2_7)$ esiintyisi minimipolynomissa kertoimella neljä. Ainoastaan vaihtoehdot (2) ja (3) ovat yhteensopivia oletuksen kanssa - kummassakin minimipolynomissa täytyy tekijän $(\mathbf{X} - 2_7)$ esiintyä yhtälön (3.112) nojalla tasan kertoimella kaksi. Näin ollen L:n esitys Jordanin normaalissa muodossa on (solujen permutaatioita vaille) joko matriisi

$$A_1 = \begin{bmatrix} 2_7 & 1_7 & 0 & 0 & 0 & 0 \\ 0 & 2_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2_7 & 1_7 & 0 & 0 \\ 0 & 0 & 0 & 2_7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3_7 & 1_7 \\ 0 & 0 & 0 & 0 & 0 & 3_7 \end{bmatrix}$$

tai matriisi

$$A_2 = \begin{bmatrix} 2_7 & 1_7 & 0 & 0 & 0 & 0 \\ 0 & 2_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2_7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2_7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3_7 & 1_7 \\ 0 & 0 & 0 & 0 & 0 & 3_7 \end{bmatrix}.$$

Kumpikin matriisi toteuttaa oletukset ja voisi olla kuvauksen L matriisiesitys. Jos operaattorista ei ole mitään muuta informaatiota, emme voi päätellä kumpi on operaattorin Jordanin normaalimuodossa oleva matriisiesitys.

Esimerkki 3.115. *Tarkastellaan lineaarista operaattoria $L\colon \mathbb{R}^3 \to \mathbb{R}^3$, jonka matriisi standardikannan suhteen on*

$$\begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}.$$

Tutkitaan onko operaattori esittävissä Jordanin normaalimuodossa ja jos on, niin mikä on sen esitys Jordanin normaalimuodossa. Aloitetaan laskemalla operaattorin karakteristinen polynomi. Määritelmän mukaan

$$\chi_{\mathbf{L}} = \det \begin{bmatrix} \mathbf{X}-3 & 0 & -8 \\ -3 & \mathbf{X}+1 & -6 \\ 2 & 0 & \mathbf{X}+5 \end{bmatrix}.$$

Koska tämän matriisin toisessa sarakkeessa on jopa kaksi nollaa, kannattaa laskea tämä determinantti kehittämällä se toisen sarakkeen suhteen, jolloin saadaan

$$\chi_{\mathbf{L}} = (\mathbf{X}+1)\det \begin{bmatrix} \mathbf{X}-3 & -8 \\ 2 & \mathbf{X}+5 \end{bmatrix} = (\mathbf{X}+1)((\mathbf{X}-3)(\mathbf{X}+5)+16) = (\mathbf{X}+1)(\mathbf{X}^2+2\mathbf{X}+1) = (\mathbf{X}+1)^3.$$

Näin ollen operaattorin ainoa ominaisarvo on (-1)*. Koska karakteristinen polynomi on jaettavissa ensimmäisen asteen tekijöihin, operaattori voidaan esittää Jordanin normaalissa muodossa ja sen jokainen solu on* (-1)*-solu. Koska minimipolynomi on karakteristisen polynomin tekijä, se on muotoa* $(\mathbf{X}+1)^i$*,* $i = 1, 2, 3$*. Jos* $\mathbf{m}_L = \mathbf{X}+1$*, yhtälön 3.112 nojalla voidaan päätellä, että operaattorin matriisiesitys Jordanin normaalimuodossa on diagonaalinen matriisi*

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(kaikki solut (1×1)*-soluja). Vastaavasti, jos minimipolynomi on* $(\mathbf{X}+1)^2$*, samalla tavalla voidaan päätellä, että matriisissa täytyy tällöin olla yksi* (1×1)*-solu ja yksi* (2×2) *solu, jolloin se on matriisi*

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

ja jos minimipolynomi on $(\mathbf{X}+1)^3$ *matriisissa täytyy olla yksi* (3×3)*-solu, jolloin se on matriisi*

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

Sen selvittämiseksi mikä kolmesta matriisista on tässä tapauskessa oikea, riittää siis vain selvittää minimipolynomi, tai oikeastaan riittää ysinkertaisesti tarkistaa suoraan laskemalla mikä kolmesta minimipolynomin kandidaatista on minimipolynomi. Jos minimipolynomi olisi $\mathbf{X}+1$*,* L *toteuttaisi yhtälön* $L + \mathrm{id}_V = 0$*, eli* $L = -\mathrm{id}_V$*. Selvästi tämä ei pidä paikkaansa (operaattorin* $-\mathrm{id}_V$ *matriisi minkä tahansa kannan suhteen on matriisi* $-I_3$*). Seuraavaksi tarkistetaan onko* $(\mathbf{X}+1)^2$ *minimipolynomi, jolloin riittää laskea toteuttaako* L *yhtälön* $(L+\mathrm{id}_V)^2 = 0$*. Suoraan laskemalla nähdään, että näin on (tarkista yksitysikohdat!), sillä*

$$\left(\begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Näin ollen minimipolynomi on $(\mathbf{X}+1)^2$*, jolloin Jordanin normaalimuodossa olevan esityksen on oltava (solujen permutaatiota vaille) matriisi*

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Sen selvittäminen, minkä kannan suhteen tämä matriisi olisi operaattorin L matriisi vaatisi lisää laskuja. Itse matriisiesityksen selvittämiseksi tällaisen kannan määrittäminen ei ollut tarpeellista - asia voidaan päätellä teoreettisten tulosten avulla.

Lisätietoa - yleistyksiä.
Tässä luvussa on annettu rikkaasta lineaaristen operaattoreiden teoriasta vain esimakua. Jordanin normaalimuodon lisäksi on olemassa paljon muitakin hyödyllisiä tapoja hajoittaa avaruus invarianttien aliavaruusten suoraksi summaksi sekä esittää operaattori "säännöllisessä", yksinkertaisessa matriisimuodossa. Mainitaan ilman todistusta muutama luonnollinen yleistys tämän aliluvun tuloksista.

Olkoon $L\colon V \to V$ operaattori äärellisulotteisessa K-vektoriavaruudessa. Jokaisella $\mathbf{v} \in V$ polynomialgebran $K[\mathbf{X}]$ osajoukko

$$I = \{\mathbf{p} \in K[X] \mid p(L)(\mathbf{v}) = \mathbf{0}_V\}$$

osoittautuu sen epätriviaaliksi ideaaliksi. Koska polynomialgebran jokainen ideaali on pääideaali, on olemassa yksikäsitteinen polynomi $\mathbf{m}_{L,v}$, joka virittää ideaalin I. Tätä polynomia sanotaan vektorin $\mathbf{v}$ minimipolynomiksi operaattorin L suhteen.

Seuraavassa lemmassa kohta (3) on Lemman 3.92 yleistys.

Lemma 3.116. *Olkoon $L\colon V \to V$ operaattori äärellisulotteisessa K-vektoriavaruudessa. Tällöin seuraavat väitteet pitävät paikkansa.*

(1) Polynomi $\mathbf{m}_{L,v}$ on minimipolynomin $\mathbf{m}_L$ tekijä jokaisella $\mathbf{v} \in V$.

(2) On olemassa sellainen $\mathbf{v} \in V$ jolle pätee $\mathbf{m}_{L,v} = \mathbf{m}_L$.

(3) Olkoon $\mathbf{v} \in V$ ja olkoon $n = \deg \mathbf{m}_{L,v}$. Tällöin jono

$$(\mathbf{v}, L(\mathbf{v}), L^2(\mathbf{v}), \ldots, L^{n-1}(\mathbf{v}))$$

on vapaa ja virittää aliavaruuden

$$\{p(L)(\mathbf{v}) \mid \mathbf{p} \in K[\mathbf{X}]\}.$$

Tämä avaruus on L-invariantti ja sitä sanotaan sykliseksi.

Kun $W = \{p(L)(\mathbf{v}) \mid \mathbf{p} \in K[\mathbf{X}]\}$ on syklinen aliavaruus kuten yllä, niin rajoittuman $L|W\colon W \to W$ matriisi *syklisessä* kannassa $(L^{n-1}(\mathbf{v}), \ldots, L^2(\mathbf{v}), \mathbf{v})$ on niin sanottu "syklinen solu" muotoa

$$\begin{bmatrix} -a_{n-1} & 1 & 0 & \ldots & 0 \\ -a_{n-2} & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \\ -a_1 & 0 & 0 & \ldots & 1 \\ -a_0 & 0 & 0 & \ldots & 0. \end{bmatrix} \tag{3.117}$$

Tätä matriisia voidaan pitää Jordanin 0-solun yleistyksenä. Sen ensimmäisen sarakkeen alkiot a_i vastaavat pääpolynomin $\mathbf{m}_{L,v} = \mathbf{X}^n + a_{n-1}\mathbf{X}^{n-1} + \ldots + a_1\mathbf{X} + a_0$ kertoimia.

Myös seuraava Proposition 3.95 yleistys pitää paikkaansa.

Propositio 3.118. *Olkoon $L\colon V \to V$ operaattori. Tällöin on olemassa avaruuden V vektorit $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ siten, että V on suora summa $\oplus_{i=1}^n V_i$, missä V_i on vektorin $\mathbf{v}_i$ määräämä syklinen aliavaruus*

$$V_i = K[\mathbf{X}](L)(\mathbf{v}_i), i = 1, \ldots, k.$$

Toisin sanoen jokainen operaattori voidaan esittää lohkomatriisina, joka on suora summa syklisistä soluista jotka ovat muotoa 3.117. Tällaista esitystä sanotaan *Frobeniuksen normaaliksi muodoksi.*