Luku 3

Lineaarikuvaukset äärellisulotteisten vektoriavaruuksien välillä

3.1 Invariantit aliavaruudet ja ominaisarvot

Tässä kurssin osassa käsittelemme vain äärellisulotteisiä vektoriavaruuksia kunnan K yli.
Olkoon $L\colon V \to V$ lineaarikuvaus äärellisulotteisesta vektoriavaruudesta V itselleen. Käytämme siitä myös nimitystä V:n *lineaarinen operaattori*. Valitsemalla V:lle joku kanta $\mathbf{e}$ voimme esittää L matriisina $[L]_{\mathbf{e}}$. Tämä esitys riippuu valitusta kannasta. Monissa sovelluksissa olemme kiinnostuneet löytämään L:lle "mahdollisimman yksinkertainen" ja "säännöllinen" matriisesitys. Se mitä pidetään yksinkertaisena ja säännöllisenä riippuu tietenkin asianyhteydestä ja sovelluksesta. Esimerkiksi mitä enemmän nolla-termejä matriisissa esiintyy, sitä helpompi sitä on yleensä käsitellä - esimerkiksi determinantin laskeminen helpottuu jne.

Käytämme seuraavaa merkintätapaa. Olkoot A $(n \times m)$-matriisi, B $(n \times k)$-matriisi, C $(l \times m)$-matriisi, D $(k \times l)$-matriisi. Tällöin voimme muodostaa $(n+l) \times (m+k)$-matriisi

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

jonka n ensimmäistä riveją saatu yhdistelemällä A:n ja B:n rivit jne.

Olkoon $L\colon V \to V$ äärellisulotteisen vektoriavaruuden V lineaarinen operaattori ja oletetaan, että jonkun V:n kannan $\mathbf{e} = (e_1, \dots, e_n)$ suhteen matriisi $[L]_{\mathbf{e}} = [L]_{\mathbf{e},\mathbf{e}}$ on muotoa

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

missä A on $(k \times k)$-matriisi, $k \leq n$ ja 0 on nollamatriisi, jonka kaikki alkiot ovat kunnan nolla-alkioita. Tällöin, jos merkitään W:llä aliavaruutta, jonka kannan $\mathbf{e}$:n k ensimmäistä vektoria $e_1, \ldots, e_k$ virittävät, niin L kuvaa W:n itselleen, $L(W) \subset W$. Tällaisia aliavaruuksia sanotaan *invarianteiksi* L:n suhteen.

Avaruuden V aliavaruus W on siis invariantti L:n suhteen, jos $L(W) \subset W$. Tällöin L:n rajoittuma $L|W$ määrittelee operaattorin $L|W \colon W \to W$. Olemme näyttäneet, että jos L:n matriisi $[L]_{\mathbf{e}}$ jonkun kannan $\mathbf{e} = (e_1, \ldots, e_n)$ suhteen on muotoa

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

jollakin neliömatriisilla $A \in M(k \times k; \mathbb{R})$, niin L:llä on invariantti aliavaruus $W = \mathrm{Span}\{e_1, \ldots, e_k\}$. Kääntäen, olkoon W invariantti L:n suhteen. Valitaan V:lle kanta $\mathbf{e} = (e_1, \ldots, e_n)$ siten, että $(e_1, \ldots, e_k)$ on W:n kanta (tämä on mahdollista Lemman 2.18 nojalla). Tällöin matriisi $[L]_{\mathbf{e}}$ on muotoa

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

missä A on rajottuman $L|W \colon W \to W$ matriisi kannan $(e_1, \ldots, e_k)$ suhteen.

Avaruuden V triviaalit aliavaruudet $\{0\}$ ja V ovat aina minkä tahansa operaattorin L invariaanttejä aliavaruuksia. Voi hyvinkin käydä niin, että nämä ovat L:n ainoat invariaantit aliavaruudet.

Esimerkki 3.1. *Olkoon $n \in \mathbb{N}$. Tarkastellaan $\mathbb{R}$-vektoriavaruutta*

$$P_n = \{p \colon \mathbb{R} \to \mathbb{R} \text{ on polynomi, jonka aste on korkeintaan } n-1\}.$$

Tällöin, kun $m \leq n$, avaruus P_m on P_n:n aliavaruus eli saadaan nouseva ketju vektoriavaruuksia

$$P_0 = \{0\} \subset P_1 \subset P_2 \subset \ldots \subset P_k \subset \ldots \subset P_n,$$

joissa jokainen avaruus on äärellisulotteinen (harjoitustehtävä).
P_n:ssä on määritelty lineaarinen derivaatta-operaattori $\mathcal{D} \colon P_n \to P_n$, $\mathcal{D}(p) = p'$. Aliavaruus P_k on $\mathcal{D}$:n invariantti aliavaruus jokaisella $0 \leq k \leq n$.

Oletetaan, että L:llä on invariantti aliavaruus W. Aliavaruudella W on Lemman 2.68 nojalla olemassa V:ssä komplementti W', $W \oplus W'$, itse asiassa yleensä paljon erilaisia komplementteja. Yleisesti voi käydä niin, että mikään W:n komplementti ei ole invariantti L:n suhteen. Esimerkiksi, jos edellisessä esimerkissä otetaan $0 < k < n$, niin aliavaruuden P_k mikään komplementti P_n:ssä ei ole enää invariantti derivaatta-operaatorin $\mathcal{D}$ suhteen (harjoitustehtävä).

Jos kuitenkin käy niin onnekkaasti, että on olemassa esitys $V = W \oplus W'$, missä molemmat W ja W' ovat L-invariantteja, niin on olemassa V:n kanta (nimittäin W:n ja W' kantojen yhdiste), jonka suhteen L:n matriisi on muotoa

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

missä A ja B neliömatriisit. Tässä A on silloin rajoittuman $L|W\colon W \to W'$ matriisi ja B on rajoittuman $L|W'\colon W' \to W'$ matriisi.
Kääntäen, jos kannan $(e_1, \ldots, e_n)$ suhteen L:n matriisi on muotoa

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

missä A $(k \times k)$-matriisi ja B $(n-k) \times (n-k)$-matriisi, niin $W = \text{Span}\{e_1, \ldots, e_k\}$ ja $W' = \text{Span}\{e_{k+1}, \ldots, e_n\}$ ovat L-invariantteja ja toistensa komplementteja V:ssä.

Tämä havainto voidaan helposti yleistä mielivaltaisen moneen määrään aliavaruuksia. 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 L-invariantti, niin L voidaan esittää matriisina

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

missä A_i on jonkun kuvauksen $L|W_i\colon W_i \to W_i$ matriisi.
L:n rajoittumat aliavaruksiin W_i määräävät tällöin kuvauksen L yksikäsitteisesti. Lisäksi helposti nähdään, että

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

eli L:n determinaantti voidaan laskea rajoittumien $L|W_i$ determinaanttien tulona.

-**Ominaisvektorit ja ominaisarvot**.
Yksinkertaisin epätriviaali aliavaruus on 1-ulotteinen, eli yhden alkion $x \in V, x \neq 0$ virittämä aliavaruus

$$W = \{kx \mid k \in K\}.$$

Tutkitaan tilanetta, jossa operaattorin L invariantti aliavaruus W on 1-ulotteinen. Olkoon x jokin W:n virittäjä. Tällöin siis $W = \{kx \mid k \in K\}$ ja $x \neq 0$. Koska W on invariantti, niin erityisesti $L(x) = kx$ jollakin $k \in K$. Tällainen k on yksikäsitteinen, sillä vektoriavaruudessa on voimassa supistussäännöt - jos $kx = k'x$, niin $k = k'$, sillä $x \neq 0$. Jos otetaan jokin toinen W:n virittäjä y, niin $y = k'x$ jollakin $k' \in K$ ja $L(y) = L(k'x) = k'L(x) = k'kx = k(k'x) = ky$ (huomaa, miten kunnan vaihdanaisuus on käytetty tässä hyväksi).

Määritelmä 3.2. *Olkoon $x \in V$, $x \neq 0$. Sanomme, että x on operaattorin $L\colon V \to V$* ominaisvektori *jos on olemassa $k \in K$ siten, että $L(x) = kx$. Kerroinkunnan alkio k on tällöin ominaisvektoriin x:n liityvä kuvauksen L* ominaisarvo*.*
Operaattorin L ominaisarvo *on puolestaan johonkin sen ominaisvektoriin liittyvä ominaisarvo.S*

Olkoon $k \in K$ kuvauksen L ominaisarvo. Sitä vastaava *ominaisarvoaliavaruus* määritellään

$$V_k = \{x \in V \mid L(x) = kx\}.$$

Lemma 3.3. *1) Jokainen L:n ominaisarvoaliavaruus on V:n aliavaruus. Se on epätriviaali ja invariantti L:n suhteen.*
2) Olkoot $k_1, \ldots, k_n \in K$ L:n (erilaiset) ominaisarvot. Tällöin summa $\sum_{i=1}^{n} V_{k_i}$ on suora. Erityisesti L:llä voi olla korkeintaan $\dim V$ erilaista ominaisarvoa.

Todistus. 1):n todistus jätetään (helpoksi) harjoitustehtäväksi lukijalle.
2) osoitetaan väite induktiolla n:n suhteen. Sovelletaan Lemman 2.64 ehtoa (ii). Olkoon $x_i \in V_{k_i}$, $i = 1, \ldots, n$ ja

$$x_1 + x_2 + \ldots + x_n = 0.$$

Soveltamalla tähän yhtälön kuvausta L saadaan

$$k_1x_1+k_2x_2+\ldots+k_nx_n = L(x_1)+L(x_2)+\ldots+L(x_n) = L(x_1+x_2+\ldots+x_n) = L(0) = 0.$$

Toisaalta kertomalla sama yhtälö skalaarilla k_n saadaan

$$k_nx_1 + k_nx_2 + \ldots + k_nx_n = 0.$$

Vähentämällä toinen yhtälö toisesta, saadaan yhtälö

$$(k_1 - k_n)x_1 + \ldots + (k_{n-1} - k_n)x_{n-1} = 0,$$

joka voidaan kirjoittaa muotoon $y_1 + y_2 + \ldots + y_{n-1} = 0$, missä $y_i \in V_{k_i}$, $i = 1, \ldots, n-1$. Soveltamalla induktio-oletusta saadaan $(k_i - k_n)x_i = y_i = 0$, $i = 1, \ldots, n-1$. Koska oletamme, että ominaisarvot k_i ovat kaikki eri luvut, $k_i - k_n \neq 0$. Koska K on kunta, tästä seuraa, että $x_i = 0$ kaikilla $i = 1, \ldots, n-1$. Tällöin myös $x_n = 0$. Väite on todistettu.
Jokainen ominaisarvoaliavaruus on epätriviaali, erityisesti sen dimensio on vähintään 1. Proposition 2.69 nojalla saadaan

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

□

Ominaisarvoaliavaruuden V_k dimensiota sanotaan ominaisarvon k *geometriseksi kertaluvuksi.*

Olkoon $L\colon V \to V$ operaattori ja olkoon $k_1, \ldots, k_n$ sen kaikki erilaiset ominaisarvot. Edellisen tuloksen nojalla on olemassa V:n alimoduli $V' = \oplus_{i=1}^{n} V_{k_i}$, jonka kaikki ominaisvektorit virittää. Voi tietenkin käydä niin, että V' on aito alimoduli. Itse asiassa eihän operaattorilla edes tarvitse olla ominaisvektoreita ollenkaan. Esimerkiksi tason aito kierto on operaattori, jolla ei ole ominaisarvoja.
Jos kuitenkin pätee $V' = \oplus_{i=1}^{n} V_{k_i} = V$, sanomme operaatorin L olevan *diagonalisoituva*. Syy tähän termiin on seuraava.
Neliömatriisi $A = (a_{ij}) \in M(n \times n; K)$ on *diagonaalimatriisi* jos $a_{ij} = 0$ kaikilla $i \neq j$. Toisin sanoen diagonaalimatrisiilla on nollasta eroavia alkioita vain *diagonaalilla* (eli alkiot jotka ovat muotoa a_{ii}).
Olkoon operaattori L diagonalisoituva. Valitsemme jokaiselle ominaisarvoaliavaruudelle V_{k_i} kannan. Koska ominaisarvoaliavaruuksien summa on suora ja yhtyy koko avaruuteen, näiden kantojen yhdiste on koko avaruuden V kanta. Lisäksi jokaiselle tämän kannan alkiolle e pätee $L(e) = ke$ jollakin $k \in K$. Tästä seuraa, että L:n matriisi tämän kannan suhteen on diagonaalimatriisi. Kääntäen, jos L:n matriisi jonkun kannan suhteen on diagonaalimatriisi, niin tämän kannan jokainen alkio on ominaisvektori, joten L on diagonalisoituva. Olemme osoittaneet seuraavan tuloksen.

Lemma 3.4. *Operaattori $L\colon V \to V$ on diagonalisoituva jos ja vain jos V:llä on olemassa kanta, jonka suhteen L:n matriisi on diagonaalimatriisi.*

Olkoon $A \in M(n\times n; K)$ neliömatriisi. Liitämme siihen lineaarikuvauksen $L_A\colon K^n \to K^n$, joka on määritely kaavalla $L_A(x) = Ax$ (missä x tulkitaan pystyvektorina). Tämä on itse asiassa yksikäsitteinen kuvaus jonka matriisi K^n:n standardikannan suhteen on A.
Sanomme, että matriisi A on *diagonalisoituva* jos kuvaus L_A on diagonalisoituva operaattori.

Lemma 3.5. *a) Olkoon $A \in M(n \times n; K)$. Tällöin A on diagonalisoituva jos ja vain jos on olemassa kääntyvä $(n \times n)$-matriisi J siten, että matriisi JAJ^{-1} on diagonaalimatriisi.*
b) Olkoon $L\colon V \to V$ operaattori. Olkoon $\mathbf{e}$ jokin V:n kanta. Tällöin L on diagonalisoituva jos ja vain jos matriisi $[L]_\mathbf{e}$ on diagonalisoituva.

Todistus. Kiinnitetään avaruuden kanta $\mathbf{e}$. Olkoon $\mathbf{f}$ jokin toinen V:n kanta. Tällöin

$$[L]_\mathbf{f} = J[L]_\mathbf{e}J^{-1},$$

missä $J = [\mathbf{f}|\mathbf{e}]$ on kannanvaihtomatriisi. Toisaalta mikä tahansa kääntyvä neliömatriisi J on muotoa $[\mathbf{f}|\mathbf{e}]$ jollakin kannalla bff (harjoitustehtävä) ja tällöin $J[L]_\mathbf{e}J^{-1} = [L]_\mathbf{f}$. Edellisen lemman nojalla L on diagonalisoituva jos ja vain jos löytyy kanta $\mathbf{f}$ jolle $[L]_\mathbf{f}$ on diagonaalimatriisi. Näin ollen olemme näyttäneet, että L on diagonalisoituva jos ja vain jos on olemassa kääntyvä matriisi J siten, että $J[L]_\mathbf{e}J^{-1}$ on diagonaalimatriisi.
Sovelletaan tämä tulos ensin kuvaukseen L_A ja R^n:n standardikantaan $\mathbf{e}$. Koska määritelmämme mukaan matriisi A on diagonalisoituva jos ja vain jos $[L]_\mathbf{f}$ on diagonaalimatriisi jonkun kannan $\{\mathbf{f}\}$ suhteen, olemme todistaneet a)-kohdan ehdon.
Yhdistämällä a)-kohdan tulos ja edellä todistettu L:n diagonaalisaatin karakterisaatio, saadaan b)-kohdan tulos osoitettua. $\square$

Diagonaalimatriisit on erityisen helppoa käsitellä. Esimerkiksi jos $A = (a_{ij})$ on diagonaalimatriisi, sen determinantti on $\det A = a_{11}a_{22}\dots a_{nn}$ eli diagonaalialkioiden tulo, mitä nähdään helposti kehittämällä $\det A$ ensimmäisen sarakkeen suhteen ja jatkamalla induktiolla.
Siksi on tärkeätä tiedä, milloin kuvaus (tai matriisi) on diagonalisoituva. Yksi helppo riittävä ehto annetaan seuraavassa lemmassa.

Lemma 3.6. *Olkoon $L\colon V \to V$ lineaarinen endomorfismi, missä $\dim V = n$. Oletetaan, että L:llä on n erilaista ominaisarvoa $k_1, \dots, k_n$. Tällöin L on diagonalisoituva.*

Todistus. Jos ominaisarvoja on n, niin aliavaruuden $V' = \oplus_{i=1}^n V_{k_i}$ dimensio on ainakin n. Toisaalta se ei voi olla enemmän kuin n, jos $n = \dim V$. Mutta

n-ulotteisen vektoriavaruuden aidon aliavaruuden dimension on aidosti pienempi kuin n (Lemma 2.18), joten $V' = V$. $\square$

Edellisen lemman antama diagonaalisoitumisen ehto on riittäävä, mutta ei välttämätön - L voi olla diagonalisoituva vaikka sillä olisi vähemmän kuin $\dim V$ ominaisarvoa. Tämä on selvä - voihan joku ominaisarvoaliavaruus olla enemmän kuin yksiulotteinen. Esimerkiksi identtisen kuvauksen $\mathrm{id}\colon V \to V$ matriisi on aina diagonaalimatriisi, joten se on varmasti diagonalisoituva, vaikka sillä on vain yksi ominaisarvo $1 \in K$.

Ennen kuin voimme mennä käytännön esimerkkeihin ominaisarvoista, meidän on kehitettävää menetelmiä, jolla voimme laskea annetun kuvauksen ominaisarvot. Yksi tällainen menetelmä löytyy determinantin avulla. Nimittäin olkoon $\lambda \in K$ mielivaltainen. Tällöin λ on L:n ominaisarvo jos ja vain jos $L(x) = \lambda x$ jollakin $x \neq 0$, eli jos ja vain jos $(L - \lambda\,\mathrm{id})(x) = 0$ jollakin $x \neq 0$. Koska $L - \lambda\,\mathrm{id}$ on lineaarinen kuvaus, tämä on yhtäpitävä sen kanssa, että $L - \lambda\,\mathrm{id}$ ei ole injektio. Koska $L - \lambda\,\mathrm{id}$ on äärellisulotteisten vektoriavaruuksien välinen lineaarinen kuvaus, se ei ole inejektio jos ja vain jos se ei ole isomorfismi, eli jos ja vain jos $\det(L - \lambda\,\mathrm{id}) = 0$.
Olkoon A L:n matriisi (jonkun kannan suhteen). Tällöin ehto on yhtäpitävä ehdon $\det(A - \lambda I) = 0$ kanssa, missä I on yksikkömatriisi (identtisen kuvauksen matriisi).
Pidetään $\lambda \in K$ muuttujana, ja tarkastellaan matriisia $A - \lambda I$, joka on muotoa

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda. \end{pmatrix}$$

Lasketaan sen determinantti esimerkiksi kaavalla (2.92). Saadaan lopputuloksessa summa termeistä, joista jokainen on tulo n:stä matriisin $A - \lambda I$ alkioista. Jokaisessa tällaisessa tulossa voi esiintyä termi $(a_{ii} - \lambda)$ korkeintaan n kertaa. Näin ollen $\det(A - \lambda I)$ on λ:n suhteen *polynomifunktio*, eli funktio $p\colon K \times K$, jonka voi kirjoittaa muotoon

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

joillakin $c_n, c_{n-1}, \ldots, c_0 \in K$ (polynomin *kertoimet*), $c_n \neq 0$.Tällaisen polynomin *aste* on n. Kuvauksen $\det(A - \lambda I)$ tapauksessa "johtava" kerroin $c_n = (-1)^n$, joten se on tasan n-asteinen polynomi.
Näin ollen, kun etsimme L:n ominaisarvot joudumme ratkaisemaan *polynomiyhtälön*

$$\det(A - \lambda I) = 0.$$

Esimerkkejä 3.7. *1) Tarkastellaan kuvausta $L\colon \mathbb{R}^2 \to R^2$, jonka matriisi standardin kannan suhteen on*

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

Lasketaan sen ominaisarvot. Saadaan toisen asteen yhtälön

$$0 = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2,$$

jolla on tasan yksi ratkaisu $\lambda = 1$. Näin ollen kuvauksella on vain yksi ominaisarvo. Jos $(x, y) \in V_1$ on ominaisvektori, niin $L(x, y) = (x + y, y) = (x, y)$ jos ja vain jos $x + y = x$ eli $y = 0$. Näin ollen ominaisvektorialiavaruus on 1-ulotteinen aliavaruus $W = \{(x, 0) \mid x = 0\}$ (x-akseli). W on L-invariantti aliavaruus.
Kuvauksella L ei ole muita invariantteja aliavaruuksia kuin W ja triviaalit aliavaruudet $\{0\}$ ja $\mathbb{R}^2$. 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 eli kuulua avaruuteen W yllä.
Tässä saadaan siis esimerkki tilanteesta, jossa invariantin aliavaruuden jokainen komplementti ei ole invariantti.
Erityisesti L ei ole diagonalisoituva.

2) Tarkastellaan kuvaus $L\colon \mathbb{R}^2 \to R^2$, jonka matriisi standardin kannan suhteen on

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

Lasketaan sen ominaisarvot. Saadaan toisen asteen yhtälö

$$0 = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = (-\lambda)^2 + 1 = \lambda^2 + 1.$$

$\mathbb{R}$:ssä tällä yhtälöllä ei tunnetusti ole juuria, joten L:llä ei ole ominaisarvoja ollenkaan. Tämä ei ole kovin yllättävää, sillä L on itse asiassa tason kierto origon ympäri 90 vastapäivään.

Jos taas tarkastellaan kuvausta $L\colon \mathbb{C}^2 \to \mathbb{C}^2$, jolla on standardin kannan suhteen sama matriisi, niin tällä kuvauksella on jopa kaksi ominaisarvoja - sillä polynomiyhtälöllä $\lambda^2 + 1 = 0$ on $\mathbb{C}$:ssä kaksi juurta - $i = (0, 1)$ ja $-i = (0, -1)$. Koska $\dim_{\mathbb{C}} \mathbb{C}^2 = 2$, Korollaarin 3.6 nojalla, L on diagonalisoituva. Lasketaan vastaavat ominaisvektorit.

$L(x, y) = (-y, x) = i(x, y)$ jos ja vain jos $x = iy$ eli esimerkiksi vektori $v_1 = (i, 1)$ on ominaisvektori. Vastaavasti nähdään, että $v_2 = (-i, 1)$ on arvoa $-i$ vastaava ominaisvektori. Kannassa (v_1, v_2) kuvauksen matriisi on diagonaalimatriisi

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

Edellisessä esimerkissä nähtiin, että ominasivektorien olemassaolo ja erityisesti diagonalisaation mahdollisuus riippuvat kerroinkunnasta, jonka suhteen asia tarkastellaan. Tämä johtuu pohjimmiltaan siitä, että polynomiyhtälöt käyttäytyvät eri tavalla eri kunnissa.
Esimerkissä 2) yllä tarkasteltiin lineaarikuvausta, jolla ei ole ominaisarvoja ollenkaan - koska toisen asteen yhtälöllä ei välttämättä ole $\mathbb{R}$:ssä juuria. Kompleksilukujen suhteen taas ominaisarvo löytyi, koska $\mathbb{C}$:ssä jokaisella polynomilla on juuri, eli $\mathbb{C}$ on esimerkki niin sanotusta *algebrallisesti suljetusta kunnasta.*

Olkoon K kunta. Sanomme, että K on *algebrallisesti suljettu*, jos jokaisella ei-vakio polynomiyhtälöllä $p\colon K \to K$,

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

$n \geq 1, c_n \neq 0, c_0, \ldots, c_n \in K$ on olemassa ainakin yksi *juuri*, eli sellainen $k \in K$ jolle $p(k) = 0$.
Edellisestä seuraa, että jos K on algebrallisesti suljettu kunta, V on epätriviaali K-vektoriavaruus ja $L\colon V \to V$ on lineaarinen operaattori, niin L:llä on ainakin yksi ominaisarvo.

Seuraavat kaksi tulosta ovat erittäin tärkeitä, mutta niiden todistukset ovat liaan työläitä ja aikaa vieviä, joten esitämme ne ilman todistuksia. Ensimmäinen väite todistetaan esimerkiksi Algebra II-kurssilla, toinen taas millä tahansa kompleksianalyysin peruskurssilla.

Propositio 3.8. *Jokainen kunta voidaan "upottaa" algebrallisesti suljettuun kuntaan. Toisin sanoen jos K on kunta, on olemassa algebrallisesti suljettu kunta K' ja injektiivinen homomorfismi $f\colon K \to K'$. Samastamalla K ja isomorfinen sen kanssa kunta $f(K)$ voimme ajatella, että K on K':n alikunta.*

Propositio 3.9. *Kompleksilukujen kunta $\mathbb{C}$ on algebrallisesti suljettu.*

Propositiossa 3.8 riittää olettaa, että on olemassa kunta-homomorfismi $K \to K'$, sillä mikä tahansa kunta-homomorfismi on injektiivinen. Proposition väitettä voidaan tarkentaa - voidaan todistaa, että jokaiselle kunnalle K on olemassa jopa "pienin" algebrallisesti suljettu kunta K', joka sisältää sen alikuntana. "Pienin" tässä yhteydessä tarkoittaa, että jos K'' on toinen algebrallisesti suljettu kunta, joka sisältää K alikuntana, se sisältää myös K':n alikuntana. Tällainen pieni algebrallisesti suljettu kunta, joka sisältää K:n, sanotaan K:n *algebralliseksi sulkeumaksi*. Kaikki nämä väitteet formalisoidaan ja todistetaan Algebra II-kurssilla.
Helposti nähdään, että (jos Proposition 3.9 väite hyväksytään todeksi) $\mathbb{R}$:n algebrallinen sulkeuma on kompleksilukujen kunta $\mathbb{C}$.
Rationaalilukujen kunnan $\mathbb{Q}$ algebrallinen sulkeuma Q' ei ole $\mathbb{R}$ - sillä eihän $\mathbb{R}$ edes ole algebrallisesti suljettu, mutta ei ole myöskään kunta $\mathbb{C}$, vaan on $\mathbb{C}$:n aito alikunta, joka koostuu niin sanotuista *algebrallisista luvuista*, eli sellaisista kompleksiluvuista, jotka ovat jonkun kokonaisluku-kertoimisen polynomin juureja. Esimerkiksi $\sqrt{2}$ tai $\sqrt[3]{7}+i$ ovat algebrallisia (yritä keksiä konkonaiskertoiminen polynomi, jonka juurena on viimeksi mainittu kompleksiluku). Kuuluisa Neperin luku e sekä luku π ovat esimerkkejä reaaliluvuista, jotka eivät ole algebrallisia.

Olkoon K algebrallisesti suljettu kunta ja $L\colon V \to V$ erään äärellisulotteisen K-vektoriavaruuden V operaattori. Tällöin L:llä on ainakin yksi ominaisarvo ja ainakin yksi ominaisvektori, mutta L ei ole vältämättä diagonalisoituva. Parasta mitä tässä tapauksessa voidaan tehdä, on esittää L *yläkolmiomatriisina.*

Matriisi $A = (a_{ij}) \in M(m \times n; R)$, missä R on mielivaltainen rengas sanomme *yläkolmiomatriisiksi*, jos kaikki sen alkiot jotka ovat "diagonalin alapuoleella" ovat nolleja. Täsmällisemmin sanottuna A on yläkolmiomatriisi jos ja vain jos $a_{ij} = 0$ kaikilla $i > j$.
Yläkolmiomatrisii siis näyttää seuraavanalaiselta,

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

Kehittämällä tällainen matriisi vaikkapa ensimmäisen sarakkeen suhteen ja jatkamalla näin induktiolla nähdään, että

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

on matriisin diagonaalialkioiden tulo. Jos $\lambda \in K$, niin matriisi $A - \lambda I$ on myös yläkolmiomatriisi ja edellisen nojalla

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

Tästä seuraa, että yläkolmionmatriisin A ominaisarvot ovat tasan sen diagonaalialkiot. Erityisesti jos yläkolmiomatriisin kaikki diagonaalialkiot ovat eri alkiot, **se on diagonalisoituva** (Korolaari 3.6).

Lemma 3.10. *Olkoon $L\colon V \to V$ operaattori, missä V on (äärellisulotteinen) K-vektoriavaruus (K mielivaltainen kunta, ei välttämättä algebrallisesti suljettu). Olkoon $n = \dim V$. Tällöin on olemassa V:n kanta, jonka suhteen L:n matriisi on yläkolmiomatriisi jos ja vain jos on olemassa kasvava ketju V:n aliavaruuksia*

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

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

Todistus. Harjoitustehtävä. □

Esimerkki 3.11. *Olkoon P_n $\mathbb{R}$-vektoriavaruus, jonka muodostavat kaikki polynomit $p\colon \mathbb{R} \to \mathbb{R}$, joiden aste on korkeintaan $n-1$. Olkoon $D\colon P_n \to P_n$ derivaatta-kuvaus, $D(p) = p'$. Tällöin*

$$P_1 \subset P_2 \subset \dots \subset P_{n-1}$$

on ketju D-invariantteja aliavaruuksia joille $\dim P_i = i$. Kannan $\{1, t, t^2, \dots, t^{n-1}\}$ suhteen kuvauksen D matriisi on yläkolmiomatriisi. Itse asiassa, koska k-asteisen polynomin derivaatan aste on aidosti pienimpi kuin k, tämän matriisin diagonaalialkiot ovat myös kaikki nolleja. Näin ollen derivaattakuvauksen ainoa ominaisarvo on 0, mikä on muutenkin helppo nähdä suoraan. Koska vastaava ominaisarvoaliavaruus P_n^0 koostuu tasan vakiopolynomeista, ei tämä kuvaus ole diagonalisoiva, kun $n \geq 2$.

Propositio 3.12. *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 V:n dimension suhteen. Kun $n = 1$ väite on selvä, sillä jokainen (1×1)-matriisi on triviaalisti yläkolmiomatriisi. Oletetaan, että väite on tosi avaruuksille, joiden dimensio korkeintaan $n-1$ ja olkoon V:n dimensio n. Olkoon $L\colon V \to V$ operaattori. Koska K on algebrallisesti suljettu, L:llä on ominaisarvo λ ja siihen liittyvä ominaisvektori

x. Olkoon $U = Kx$ x:n virittämä aliavaruus. Lemman 2.68 mukaan U:llä on olemassa V:ssä komplementti W, $V = U \oplus W$. Tällöin $\dim W = n - 1$. Olkoon $p\colon V \to W$ kanoninen projektio. Kuvaus $L'\colon W \to W$, $L'(x) = p(L(x))$ on $n-1$-ulotteisen K-vektoriavaruuden operaattori, joten induktio-oletuksen mukaan W:llä on kanta $(e_1, \ldots, e_{n-1})$, jonka suhteen L':n matriisi on yläkolmiomatriisi. Tämä tarkoittaa sitä, että jos merkitään W_i:llä aliavaruutta, jonka jono $(e_1, \ldots, e_i)$ virittää, niin $L'(W_i) \subset W_i$. Jono $(x, e_1, \ldots, e_{n-1})$ on V:n kanta. Jos merkitään V_i:llä alivaaruutta, jonka osajono $(x, e_1, \ldots, e_i)$ virittää, niin $L(V_i) \subset V_i$ kaikilla $i = 0, \ldots, n-1$. Tästä seuraa, että kannan $(x, e_1, \ldots, e_{n-1})$ suhteen L:n matriisi on yläkolmiomatriisi.

□

Toinen tapa osoittaa edellisen Lemman väitettä todeksi on käyttää hyväksi duaali-avaruutta. Nimittäin duaalikuvaus $L^*\colon V^* \to V^*$ on n-ulotteisen avaruuden V^* operaattori, joten, K:n olleessa algebrallisesti suljettu, sillä on ominaisvektori $f\colon V \to K$. Tällöin $W = \operatorname{Ker} L$ on $(n-1)$-ulotteinen (niin sanottu *hyperavaruus*) L-invariantti aliavaruus V:ssä (harjoitustehtävä). Soveltamalla sama tulos kuvauksen L rajoittumaan $L|\colon W \to W$ löydetään $n - 2$ ulotteinen L-invariantti aliavaruus jne. Jatkamalla näin saadaan osoitettua, että L:lle pätee Lemman 3.10 ehto.

Esitetään kaksi edellisen tuloksen sovellusta, joissa mennään analyysin ja topologian puoleella. Siksi esitämme joitakin vaiheita ilman todistuksia. Tarkoitus on näyttää "käsiä heiluttamalla", miten edellä saadut tulokset sovelletaan käytännössä erityyppisiin tilanteisiin.

Esimerkki 3.13. *Olkoon A $\mathbb{R}$-kertoiminen $(n \times n)$-matriisi. Haluamme määritellä sen eksponentti 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 kursseilla osoitetaan, että tämä sarja suppenee kaikilla $x \in \mathbb{R}$. Neliömatriiseille "apinoidaan" samaa määritelmää, eli asetetaan

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

Voidaan osoittaa (me ei mennä siihen tällä kurssilla), että näin määritelty sarja suppenee jokaisella matriisilla $A \in M(n \times n; \mathbb{R})$. Näin määritelty matriisin eksponentti ei ole pelkäästään hyödytön analogia, vaan erittäin tärkeä otus esimerkiksi differentiaaligeometrissa, lie ryhmien teoriassa tai jopa fyysikassa. Esimerkiksi kuvaus $t \mapsto e^{tA}$ on differentiaaliyhtälön $y' = Ay$ ratkaisu

matriisien avaruudessa.
Haluamme osoittaa, että matriisi e^A *on aina kääntyvä eli* $\det e^A \neq 0$ *(vrt. reaalilivuille* $e^x \neq 0$ *kaikilla* $x \in \mathbb{R}$*). Emme voi tähän käyttää suoraan edellista tulosta, sillä eihän* $\mathbb{R}$ *edes ole algebrallisesti suljettu.*
Tätä varten määritellään eksponenttikuvaus kaikille kompleksiarvoisille neliömatriiseille samalla kaavalla, eli jos $A \in M(n \times n; \mathbb{C})$*, asetetamme*

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

Voidaan edelleenkin osoittaa, että tämä sarja suppenee aina. Lisäksi edellä tarkasteltu reaaliarvoisten matriisien tapaus on tämän määritelmän erikoistapaus, sillä onhan jokainen reaaliarvoinen matriisi myös erityisesti kompleksiarvoinen (tämä on tyyppillinen esimerkki siitä, miten voimme käyttää kompleksilukuja, vaikka ratkaisemme reaalilukuhin liittyvää ongelmaa).
Edellisen proposition nojalla voimme jokainen $A \in M(n \times n; \mathbb{C})$ *esittää yläkolmiomatriisina* B*, vaihtamalla kannan standardikannasta johonkin toiseen kantaan. Tällöin siis* $A = JBJ^{-1}$*, missä* J *on kannanvaihtomatriisi. Induktiolla helposti nähdään, että* $A^n = JB^nJ^{-1}$ *kaikilla* $n \in \mathbb{N}$*, 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 otamme J *ja* J^{-1} *"yhteisksi tekijöiksi" viimeisessä sarjassa, koska ne esiintyvät jokaisessa termissa (näiden manipulaatioiden tarkka perustelu vaatii analyysin tuloksia, sivutamme yksityiskohdat). Koska* $\det X = \det JXJ^{-1}$ *jokaisella matriisille* X*, riittää siis osoittaa, että* $\det e^B \neq 0$ *kun* B *on yläkolmiomatriisi.*
Huomaa, että jos olisimme pysyneet $\mathbb{R}$*:ssä, emme olisi voineet käyttää tällaista tekniikkaa, sillä kun yllä* A *on reaaliarvoinen,* B *ja* J *eivät enää yleisesti ole.*

Jos B *on yläkolmiomatriisi, jonka diagonaalialkiot ovat* $b_{11}, \ldots, b_{nn}$*, niin* B^2 *on myös yläkolmiomatriisi, jonka diagonaalialkiot ovat* $b_{11}^2, \ldots, b_{nn}^2$ *(helppo lasku). Induktiolla nähdään, että jokaisella* $m \in \mathbb{N}$ *matriisi* B^n *on yläkolmiomatriisi, jonka diagonaalialkiot ovat* $b_{11}^n, \ldots, b_{nn}^m$*. Näin ollen sarjassa*

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

rajana saadaan yläkolmiomatriisi, jonka diagonaalialkiot ovat sarjoja

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

missä $z = b_{ii}$. Näin ollen

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

sillä $e^z \neq 0$ jokaisella kompleksiluvulla $z \in \mathbb{C}$ (jos et ole aikaisemmin törmännyt kompleksuluvun eskponenttiin, joudut vain uskomaan tätä, mutta mielummin ota asiasta selvää). Väite on todistettu.

Esimerkki 3.14. *Olkoon A kompleksiarvoinen neliömatrisi. Osoitetaan, että* mielivaltaisen lähellä *A:ta löytyy matriisi, joka on diagonalisoituva. Tämä tulos on tärkeä analyysissa. Sen avulla voidaan esimerkiksi todistaa matriisien liittyviä väiteittä todistamalla ne ensin diagonaalisoituville matriisille (mikä usein voi olla hyvin helppoa) ja sitten yleistämällä se mielivaltaiselle matriisille jonkinlaisella "raja-käynnillä".*
Mitä mielivaltaisen lähellä tarkoittaa tässä? Se tarkoittaa, että jokaisella $\varepsilon > 0$ voimme muuttaa A:ssa esiintyvät alkiot korkeintaan ε:n verran, niin, että saamme diagonalisoituvan matriisin. Täsmällisemmin jos $A = (a_{ij})$, niin väitämme, että jokaisella $\varepsilon > 0$ löytyy diagonalisoituva matriisi $B = (b_{ij})$ siten, että $|a_{ij} - b_{ij}| < \varepsilon$ kaikilla i, j. Kompleksiluvun itseisarvo $|\cdot|$ on sama asia kuin sen itseisarvo tason $\mathbb{R}^2$ pisteenä (eli etäisyys origoon).

Esitetään A muodossa JCJ^{-1}, missä C on yläkolmiomatriisi. Jos löydämme mielivaltaisen lähellä C:tä diagonalisoituvan matriisin B, niin JBJ^{-1} on diagonalisoituva matriisi, joka on mielivaltaisen lähellä A:ta (koska matriisien kertolasku on niin sanottu "jatkuva operaatio", sivuutamme tarkka perustelu). Näin ollen riittää olettaa, että A on yläkolmiomatriisi. Sen diagonaalialkiot eivät vältämättä ole eri alkiot, mutta selvästi voimme muuttaa niiden arvot hyvin vähän, niin, että saadaan eri luvut diagonaalilla. Tällöin saadaan uusi matriisi B, joka on myös yläkolmiomatriisi, joka on hyvin lähellä A:ta ja jonka diagonaalialkiot ovat kaikki eri alkiot. Mutta tällainen matriisi on aina diagonalisoituva (kts. yllä).