Luku 4

Sisätuloavaruudet

4.1. Määritelmät ja perusominaisuudet

Tähän asti ollaan tutkittu äärellisulotteisia vektoriavaruuksia mielivaltaisen kunnan K yli. Tässä kurssin osiossa rajoitutaan tarkastelemaan ainoastaan vektoriavaruuksia kuntien $\mathbb{R}$ ja $\mathbb{C}$ yli. Tämä johtuu siitä, että siirrytään käyttämään uutta työkalua nimeltä "sisätulo". Sisätulon käsite on hieman erikoinen ja se nojautuu vahvasti reaalilukujen väliseen järjestysrelaatioon ja sen ominaisuuksiin sekä kompleksilukuihin liittyvään "konjugaatin" käsitteseen. Kummallakin näistä ei ole luonnollista vastinetta muissa kunnassa. Tästä syystä sisätuloa ei voi määritellä mielekkäästi vektoriavaruuksissa mielivaltaisen kunnan K yli.

Vaikka kiinnostuksen kohteen rajoittaminen ainoastaan reaali- ja kompleksikertoimiseen tapaukseen saattaa tässä vaiheessa tuntua liian suppealta erikoistapaukselta, sisätuloavaruudet ovat siitä huolimatta keskeisessa roolissa monissa lineaarialgebran sovelluksissa, sekä matematiikassa, että myös sen ulkopuolella (esimerkiksi fysiikassa).

Reaalikertoimisen vektoriavaruuden sisätulon käsite on lukijalle tödennäköisesti tuttu lineaarialgebran peruskurssilta. Ennen kuin mennään yleiseen kompleksiseen tapaukseen palautetaan ensin mieleen miten reaalikertoiminen sisätuloavaruus määritellään.

Reaalinen tapaus

Olkoon V $\mathbb{R}$-vektoriavaruus. Kuvausta $\langle , \rangle \colon V \times V \to \mathbb{R}$ sanotaan *sisätuloksi* avaruudessa V, jos seuraavat ehdot pätevät kaikilla $\mathbf{v}, \mathbf{v}', \mathbf{w} \in V$ ja kaikilla $r \in \mathbb{R}$.

(1) $\langle \mathbf{v} + \mathbf{v}', \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}', \mathbf{w} \rangle$.

(2) $\langle r\mathbf{v}, \mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle$.

(3) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$.

(4) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$.

(5) $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ jos ja vain jos $\mathbf{v} = 0$.

Tässä parin $(\mathbf{v}, \mathbf{w})$ arvoa kuvauksen $\langle , \rangle$ suhteen merkitään $\langle \mathbf{v}, \mathbf{w} \rangle$.

Sisätulon määritelmän ehdot (1) ja (2) tarkoittavat sitä, että sisätulo on *lineaarinen* ensimmäisen muuttujan suhteen. Kuitenkin tämän sekä ehdon (3) avulla voidaan helposti näyttää, että sisätulo on lineaarinen myös toisen muuttujan suhteen. Nimittäin, olkoot $\mathbf{v}, \mathbf{w}, \mathbf{w}' \in V$, $r \in \mathbb{R}$. Tällöin ehdoista (1)-(3) yllä seuraa, että

$$\langle \mathbf{v}, \mathbf{w} + \mathbf{w}' \rangle = \langle \mathbf{w} + \mathbf{w}', \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}', \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w}' \rangle,$$

$$\langle \mathbf{v}, r\mathbf{w} \rangle = \langle r\mathbf{w}, \mathbf{v} \rangle = r\langle \mathbf{w}, \mathbf{v} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle.$$

Toisin sanoen $\mathbb{R}$-vektoriavaruuden V sisätulo on samalla sen *bilineaarinen muoto*. Lisäksi sisätulon määritelmän ehto (3) sanoo, että tämä muoto on *symmetrinen*.

Sisätulon määritelmän kohdat (4) ja (5) paljastavat sisätulon erikoisen piirteen, nimittäin sen tosiasian, että sisätulon käsite pohjautuu, ainakin osittain, reaalilukujen joukossa $\mathbb{R}$ määriteltyyn järjestysrelaatioon $\geq$. Mielivaltaisessa kunnassa K ei ole välttämättä olemassa mielekkästä järjestysrelaatiota[1]. Tämä on yksi syy siihen, miksi sisätulon käsitettä ei voida yleistää järkevällä tavalla mielivaltaisen skalaarikunnan tapaukseen.

Kompleksinen tapaus

$\mathbb{C}$-kertoimisissa vektoriavaruuksissa yllä annettua sisätulon määritelmää ei voida käyttää sellaisenaan, sillä kompleksilukujen kunnassa $\mathbb{C}$ ei ole määritelty mielekästä järjestysrelaatiota. On kuitenkin keksitty hyödyllinen tapa puhua sisätulosta myös $\mathbb{C}$-vektoriavaruuksien kohdalla. Käsitteen määrittämistä varten palautetaan ensin mieleen kompleksiluvun *konjugaatin* ominaisuuksia.

Olkoon $z = x + iy \in \mathbb{C}$ kompleksiluku, $x, y \in \mathbb{R}$. Kompleksiluvun z *konjugaatti* $\overline{z}$ määritellään kompleksilukuna $\overline{z} = x - iy$. Koska konjugaatin käsite on määritelty kaikille kompleksiluvuille, voidaan puhua *konjugaattikuvauksesta* $\mathbb{C} \to \mathbb{C}, z \mapsto \overline{z}$. Jos joukkoa $\mathbb{C}$ ajattelee perinteisellä tavalla tasona $\mathbb{R}^2$, tämä kuvaus on yksinkertaisesti peilaus x-akselin suhteen.

Konjugaattkuvaus on *kuntaisomorfismi* $\mathbb{C} \to \mathbb{C}$, jonka käänteiskuvaus on konjugaattikuvaus itse. Lisäksi konjugaattikuvaus on jopa $\mathbb{R}$-lineaarinen, jos $\mathbb{C}$ tulkitaan luonnollisella tavalla $\mathbb{R}$-vektoriavaruutena $\mathbb{R}^2$. Näiden väitteiden tarkka osoittaminen jätetään lukijalle harjoitustehtäväksi. Kaikilla $z_1, z_2 \in \mathbb{C}$ ja $r \in \mathbb{R}$ pätee siis

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2},$$

$$\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2},$$

$$\overline{r z_1} = r\overline{z_1}.$$

Lisäksi jokaisella $z \in \mathbb{C}$ kompleksiluvut $z + \overline{z}$ ja $z\overline{z}$ ovatkin reaalilukuja. Osoitetaan tämä. Olkoon $z = x + iy$. Tällöin

$$z + \overline{z} = (x + iy) + (x - iy) = 2x \in \mathbb{R},$$

$$z\overline{z} = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2 \in \mathbb{R}. \tag{4.1}$$

[1] Täsmällisemmin sanottuna: sellaista järjestystä, joka olisi yhteensopiva kunnan algebrallisen struktuurin kanssa.

Tässä jälkimmäisessä yhtälössä esiintyvä lauseke $x^2+y^2 = |z|^2$ on tason pisteen $(x, y) = z$ *normin* eli *itseisarvon* neliö. Kun $z \neq 0$, myös $|z|^2 \neq 0$, jolloin jakamalla yhtälössä (4.1) molemmat puolet normilla $|z|^2$ saadaan yhtälö

$$z \cdot \frac{\overline{z}}{|z|^2} = 1.$$

Ollaan siis johdattu uuden tavan kirjoittaa kompleksiluvun käänteisluku, sillä edellisen nojalla kaikilla $z \in \mathbb{C}$, $z \neq 0$, pätee

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

Tätä voi ajatella myös uutena tapana osoittaa, että jokaisella ei-nolla kompleksiluvulla on käänteisluku. Erityisesti yhtälö $z^{-1} = \overline{z}$ pätee jos ja vain jos $|z| = 1$ eli kun piste $z = (x, y)$ sijaitsee tason yksikköympyrällä S^1.

Olkoot V ja W $\mathbb{C}$-vektoriavaruuksia. Sanomme, että kuvaus $L\colon V \to W$ on *antilineaarinen*, jos kaikilla $\mathbf{v}, \mathbf{w} \in V$ ja kaikilla $z \in \mathbb{C}$ pätee

$$L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w}),\ L(z\mathbf{v}) = \overline{z}L(\mathbf{v}).$$

Antilineaarisista kuvauksista käytetään myös nimityksiä ”semilineaarinen” ja ”konjugaattilineaarinen”.

Olkoot V, W, U $\mathbb{C}$-vektoriavaruuksia. Kuvausta $F\colon V \times W \to U$ sanotaan *seskilineaariseksi* jos se on lineaarinen toisen muuttujan suhteen ja antilineaarinen toisen muuttujan suhteen, eli jos kaikilla $\mathbf{v}, \mathbf{v}' \in V$, $\mathbf{w}, \mathbf{w}' \in W$ ja $z \in \mathbb{C}$ pätee

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

$$F(\mathbf{v}, \mathbf{w} + \mathbf{w}') = F(\mathbf{v}, \mathbf{w}) + F(\mathbf{v}, \mathbf{w}'),$$
$$F(z\mathbf{v}, \mathbf{w}) = zF(\mathbf{v}, \mathbf{w}),$$
$$F(\mathbf{v}, z\mathbf{w}) = \overline{z}F(\mathbf{v}, \mathbf{w}).$$

Englanninkielinen termi ”seskilineaariselle” on *sesquilinear*. Etymologisesti sesqui-etuliite tulee latinasta ja tarkoittaa ”puolitoista”. Siten seskilineaarista kuvausta voisi sanoa suomeksi myös ”puolitoista-lineaariseksi” eli ”3/2-lineaariseksi”. Samalla tavalla antilineaarista kuvausta voidaan ajatella olevan ”puolikas-lineaarinen” (”1/2-lineaarinen”). Nimitys voidaan ajatella viittavan siihen tosiasiaan, että kahden 1/2-lineaarisen kuvauksen yhdiste on lineaarinen kuvaus (jos määritelty).

Jokainen antilineaarinen/seskilineaarinen kuvaus voidaan tarvittaessa tulkita myös lineaarisena/bilineaarisena kuvauksena seuraavassa mielessä.

Olkoon V $\mathbb{C}$-vektoriavaruus. Määritellään $\mathbb{C}$-vektoriavaruus $(\overline{V}, +, \cdot)$ (avaruuden V *konjugaatti*) seuraavasti. Asetetaan $\overline{V} = V$ ja yhteenlaskuna $+$ käytetään samaa yhteenlaskua kuin avaruudessa V. Skalaarikertolasku $\cdot\colon \mathbb{C} \times V \to V$ taas määritellään kaavalla

$$z \cdot \mathbf{v} = \overline{z}\mathbf{v}, z \in \mathbb{C}, \mathbf{v} \in V,$$

missä oikealla puolella esiintyy "alkuperäinen" avaruuden V skalaarikertolasku. Helposti nähdään, että tällä tavalla määritelty algebrallinen struktuuri $(\overline{V}, +, \cdot)$ on todellakin $\mathbb{C}$-vektoriavaruus (tarkka todistus harjoitustehtävänä).

Konjugaattiavaruuden avulla antilineaarinen/seskilineaarinen kuvaus voidaan tulkita lineaarisena/bilineaarisena kuvauksena. Tämä on hyödyllistä tietoa, sillä se antaa mahdollisuuden soveltaa 1/2 ja 3/2-lineaaristen kuvauksiin aikaisemmin kehitettyä lineaaristen ja bilineaaristen kuvausten teoriaa.

Lemma 4.2. *(i) Olkoot V, W $\mathbb{C}$-vektoriavaruuksia. Olkoon $L\colon V \to W$ kuvaus. Tällöin L on antilineaarinen jos ja vain jos se on lineaarinen kuvauksena $L\colon V \to \overline{W}$ tai kuvauksena $L\colon \overline{V} \to W$.*

(ii) Olkoot V, W, U $\mathbb{C}$-vektoriavaruuksia. Olkoon $F\colon V \times W \to U$. Tällöin F on seskilineaarinen jos ja vain jos se on bilineaarinen kuvauksena $F\colon V \times \overline{W} \to U$.

Todistus. Seuraa triviaalisti määritelmistä. □

Kompleksilukujen konstruktion (kts. kappale 1.5 Luvussa 1) yhteydessä on sovittu tulkita reaalilukujen joukko $\mathbb{R}$ kompleksilukujen joukon $\mathbb{C}$ osajoukkona

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

Helposti nähdään, että tämän tulkinnan puitteissa joukko $\mathbb{R}$ koostuu tasan sellaisista kompleksiluvuista $z \in \mathbb{C}$, joille pätee $z = \overline{z}$ eli jotka ovat itsensä konjugaatteja. Käyttämällä tätä sopimusta voidaan myös yleistää joitakin $\mathbb{C}$-vektoriavaruuksiin ja konjugointiin liittyviä määritelmiä myös $\mathbb{R}$-kertoimisessa tapauksessa. Esimerkiksi voidaan puhua reaaliluvun konjugaatista (joka saattuu kyllä olemaan tällöin luku itse). Kuvausta $L\colon V \to W$ voidaan sanoa antilineaariseksi myös silloin kun V ja W ovat $\mathbb{R}$-vektoriavaruuksia, tällöin se tarkoittaa sama asia kuin L on $\mathbb{R}$-lineaarinen, jne. Menettelemällä tällä tavalla säästetään aikaa eikä tarvitse aina käsitellä erikseen $\mathbb{R}$-sisätuloavaruuksien ja $\mathbb{C}$-sisätuloavaruuksien teoriaa, ainakin niin kauan kuin niissä pätevät samat tulokset. Tehdään siis seuraava sopimus.

Symbolilla K tässä luvussa tarkoitetaan tästä lähtien joko reaalilukujen kuntaa $\mathbb{R}$ tai kompleksilukujen kuntaa $\mathbb{C}$. Kunnassa K on tällöin määritelty konjugaattikuvaus $\phi\colon K \to K$, $\phi(z) \mapsto \overline{z}$, joka on $\mathbb{R}$-lineaarinen algebraisomorfismi. Tapauksessa $K = \mathbb{C}$ kyseessä on yllä määritelty kompleksilukujen konjugaattikuvaus ja tapauksessa $K = \mathbb{R}$ kyseessä on tämän konjugaattikuvaukseen rajoittuma osajoukkoon $\mathbb{R}$, joka on tällöin yksinkertaisesti identtinen kuvaus $\mathrm{id}\colon \mathbb{R} \to \mathbb{R}$.

Jokaiseen K-vektoriavaruuteen V voidaan liittää sen *konjugaattiavaruus* $\overline{V}$, joka on määritelty yllä. $\mathbb{R}$-vektoriavaruuksille $\overline{V} = V$. Kun V ja W ovat K-vektoriavaruuksia, voidaan puhua 1/2-lineaarisista kuvauksista $L\colon V \to W$. Samoin voidaan puhua 3/2-lineaarisista kuvauksista $F\colon V \times W \to U$. Edellinen lemma pätee edelleenkin, kun siinä korvataan kunta $\mathbb{C}$ kunnalla K. Kun $K = \mathbb{R}$ antilineaarinen kuvaus on sama asia kuin lineaarinen kuvaus ja seskilineaarinen on sama asia kuin bilineaarinen. Edellisen lemman väite on tapauksessa $K = \mathbb{R}$ vailla varsinaista sisältöä.

Vaikka edellisen kappaleen havaintojen valossa sopimuksemme saattaa tuntua turhalta, se on kuitenkin hyödyllinen, sillä se tarjoaa yhtenäisen näkökulman sisätuloavaruuksien teoriaan. Lisäksi sen avulla voidaan ainakin joissakin tapauksissa säästää aikaa sekä (virtuaalista) paperitilaa, kun tuloksia ja määritelmä ei tarvitse muotoilla ja käsitellä kahteen kertaan erikseen tapauksissa $K = \mathbb{R}$ tai $K = \mathbb{C}$.

Seuraavaksi annetaan yleinen K-sisätulon määritelmä. Tapauksessa $K = \mathbb{R}$ tämä määritelmä on täysin ekvivalentti aikaisemmin määritellyn $\mathbb{R}$-sisätulon määritelmän kanssa.

Sisätulo.
Olkoon V K-vektoriavaruus. Kuvausta $\langle , \rangle : V \times V \to K$ sanotaan *hermiittiseksi muodoksi* avaruudessa V, jos se on 3/2-lineaarinen ja toteuttaa lisäksi ehdon

$$\langle \mathbf{w}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{w} \rangle}$$

kaikilla $\mathbf{v}, \mathbf{w} \in V$. Tätä ehtoa sanotaan *konjugaattisymmetrisyydeksi.* Määritelmän mukaan kuvaus $\langle , \rangle$ on siis hermiittinen muoto jos ja vain jos kaikilla $\mathbf{v}, \mathbf{w}, \mathbf{v}', \mathbf{w}' \in V$ ja $k \in K$ pätevät yhtälöt

$$\langle \mathbf{v} + \mathbf{v}', \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}', \mathbf{w} \rangle, \tag{4.3}$$

$$\langle \mathbf{v}, \mathbf{w} + \mathbf{w}' \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w}' \rangle, \tag{4.4}$$

$$\langle k\mathbf{v}, \mathbf{w} \rangle = k\langle \mathbf{v}, \mathbf{w} \rangle, \tag{4.5}$$

$$\langle \mathbf{v}, k\mathbf{w} \rangle = \overline{k}\langle \mathbf{v}, \mathbf{w} \rangle, \tag{4.6}$$

$$\langle \mathbf{w}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{w} \rangle}. \tag{4.7}$$

Tämä määritelmä on itse asiassa osittain redundantti, sillä voidaan helposti näyttää, että ehdoista (4.3), (4.5) ja (4.7) seuraavat ehdot (4.4) ja (4.6). Hermiittinen muoto voidaan siis määritellä myös kuvauksena $V \times V \to K$, joka on lineaarinen ensimmäisen muuttujan suhteen ja on lisäksi konjugaattisymmetrinen.

Tapauksessa $K = \mathbb{R}$ hermiittinen muoto on sama asia kuin symmetrinen bilineaarinen muoto $V \times V \to \mathbb{R}$.

Olkoon $\langle , \rangle$ hermiittinen muoto K-vektoriavaruudessa V. Tällöin ehdon (4.7) nojalla jokaisella $\mathbf{v} \in V$ pätee $\langle \mathbf{v}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{v} \rangle}$. Koska kompleksiluku on itsensä konjugaatti jos ja vain jos se on reaaliluku, tästä seuraa, että $\langle \mathbf{v}, \mathbf{v} \rangle$ on reaaliluku kaikilla $\mathbf{v} \in V$. Lisäksi ehdosta (4.5) voidaan helposti johtaa, että $\langle \mathbf{0}_V, \mathbf{0}_V \rangle = 0_K$.

Määritelmä 4.8. *Olkoon $\langle , \rangle$ hermiittinen muoto K-vektoriavaruudessa V. Muotoa $\langle , \rangle$ sanotaan avaruuden V* sisätuloksi *jos se on* positiivisesti definiitti *eli jos kaikilla $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}_V$, pätee*

$$\langle \mathbf{v}, \mathbf{v} \rangle > 0.$$

Määritelmä on järkevä, sillä yllä olevan mukaan Hermiittiselle muodolle pätee aina $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$, joten kysymys siitä, onko $\langle \mathbf{v}, \mathbf{v} \rangle$ positiivinen on mielekäs.

Sisätulolla $\langle , \rangle$ varustettua K-vektoriavaruutta V sanotaan *sisätuloavaruudeksi.* Täsmällisesti sanottuna sisätuloavaruus on pari $(V, (\langle , \rangle))$, missä $\langle , \rangle$ on sisätulo avaruudessa V.

Tärkeämmät sisätulon sovellukset liittyvät sen tarjoaman yhteyteen lineaarialgebran väkisen yhteyteen geometriaan ja topologiaan. Sisätulon avulla avaruudessa voidaan määritellä sellaisia käsitteitä kuten etäisyys, vektorien väliset kulmat jne. Olkoon V sisätuloavaruus, joka on varustettu sisätulolla $\langle , \rangle$. Koska jokaisella $\mathbf{v} \in V$ kompleksiluku $\langle \mathbf{v}, \mathbf{v} \rangle$ on määritelmän mukaan ei-negatiivinen reaaliluku, siitä voidaan ottaa (ei-negatiivinen) neliöjuuri. Vektoriavaruudessa V voidaan siis tällöin määritellä *normikuvaus* $|\cdot|$ kaavalla

$$|\mathbf{v}| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Vektorin $\mathbf{v}$ *normi* on tällöin ei-negatiivinen reaaliluku, joka on nolla jos ja vain jos $\mathbf{v} = \mathbf{0}_V$.

Normi toteuttaa yhtälön

$$|k\mathbf{v}| = |k|\,|\mathbf{v}| \tag{4.9}$$

kaikilla $k \in K$, $\mathbf{v} \in V$. Nimittäin sisätulon ominaisuuksista ja normin määritelmästä seuraa, että

$$|k\mathbf{v}|^2 = \langle k\mathbf{v}, k\mathbf{v} \rangle = k\overline{k}\langle \mathbf{v}, \mathbf{v} \rangle = |k|^2\,|\mathbf{v}|^2\,,$$

koska $k\overline{k} = |k|^2$ kaikilla $k \in K$ (kts. yhtälö 4.1). Tässä $|k|$ on kompleksiluvun $k = a + ib$ normi $\sqrt{a^2 + b^2}$. Ottamalla neliöjuuri yhtälön molemmasta puolesta saadaan yhtälö 4.9.

Vektorien $\mathbf{v}, \mathbf{w} \in V$ välinen *etäisyys* $d(\mathbf{v}, \mathbf{w})$ sisätuloavaruudessa V määritellään kaavalla

$$d(\mathbf{v}, \mathbf{w}) = |\mathbf{v} - \mathbf{w}|.$$

Kuuluisasta *Cauchyn-Schwarzin epäyhtälöstä* (jonka todistusta käydään läpi Lemmassa 4.14 alla) seuraa, että tämä etäisyys toteuttaa luonnollisia etäisyyden ominaisuuksia eli on niin sanottu *metriikka*. Myös vektorien välisen *kulman* määrittäminen onnistuu edellä mainitun Cauchyn-Schwarzin epäyhtälön ansiosta. Palaamme tähän sen jälkeen kun kyseinen epäyhtälö on esitetty ja todistettu.

Esimerkkejä 4.10. *1) Klassinen esimerkki $\mathbb{R}$-sisätulosta on avaruuden $\mathbb{R}^n$ ($n \in \mathbb{N}$)* pistetulo $\cdot$, *joka määritellään kaavalla*

$$x \cdot y = \sum_{i=1}^{n} x_i y_i.$$

Yleisemmin olkoon V jokin n-ulotteinen $\mathbb{R}$-vektoriavaruus, jolla on kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Tällöin avaruudessa V voidaan määritellä sisätulo $\langle , \rangle$ kaavalla

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} x_i y_i,$$

missä $\mathbf{v} = \sum_{i=1}^{n} x_i \mathbf{e}_i$ ja $\mathbf{w} = \sum_{i=1}^{n} y_i \mathbf{e}_i$ on vektorien $\mathbf{v}, \mathbf{w}$ esitys kannassa E. Sen osoittaminen, että tämä kaava todellakin määrittelee $\mathbb{R}$-sisätulon jätetään harjoitustehtäväksi. Pistetulo $\mathbb{R}^n$:ssä on erikoistapaus tästä konstruktiosta, jossa kannaksi E otetaan avaruuden $\mathbb{R}^n$ standardikanta.

Erityisesti jokaisessa äärellisulotteisessa $\mathbb{R}$-vektoriavaruudessa voidaan määritellä jokin sisätulo.

2) Tarkastellaan äärellisulotteista $\mathbb{C}$-vektoriavaruutta $\mathbb{C}^n$. Jos yrittää käyttää sisätulona samaa bilineaarista ja symmetrista muotoa $(x, y) \mapsto \sum_{i=1}^n x_i y_i$ kuin $\mathbb{R}$-kertoimisessa tapauksessa, käy niin, että "vektorin x normi" $\sum_{i=1}^n x_i^2$ voi saada negatiivisia arvoja, tai ei ole edes reaaliluku. Näin ollen tämä muoto ei kelpaa sisätuloksi. Itse asiassa helposti nähdään, että muoto $(x, y) \mapsto \sum_{i=1}^n x_i y_i$ ei toteuta sisätulon määritelmää.

"Pistetulo" avaruudessa $\mathbb{C}^n$ määritellään kaavalla

$$x \cdot y = \sum_{i=1}^{n} x_i \overline{y_i}. \tag{4.11}$$

Huomaa konjugaatit. Yleisemmin olkoon V jokin n-ulotteinen $\mathbb{C}$-vektoriavaruus, jolla on kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Tällöin avaruudessa V voidaan määritellä sisätulo $\langle , \rangle$ kaavalla

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} x_i \overline{y}_i,$$

missä $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$ ja $\mathbf{w} = \sum_{i=1}^n y_i \mathbf{e}_i$ on vektorien $\mathbf{v}, \mathbf{w}$ esitys kannassa E. Sen osoittaminen, että tämä kaava todellakin määrittelee $\mathbb{C}$-sisätulon jätetään harjoitustehtäväksi.

Historiallisesti juuri muoto (4.11) (ja sen ominaisuudet) on toiminut $\mathbb{C}$-sisätulon käsitteen määritelmän motivaationa.

Koska reaaliluvun konjugaatti on se itse, kaavaa 4.11 voidaan käyttää pistetulon määritelmänä myös avaruudessa $\mathbb{R}^n$. Näin ollen jatkossa pistetulolla avaruudessa K^n tarkoitetaan kaavan 4.11 antamaa sisätuloa. Tapauksessa $K = \mathbb{R}$ tämä palautuu esimerkissä (1) yllä määriteltyyn $\mathbb{R}^n$:n pistetuloon.

3) Olkoon $I = [a, b]$ suljettu väli $\mathbb{R}$:ssä, $a < b$. Olkoon

$$V = \{f \colon [a, b] \to \mathbb{C} \mid f \text{ on jatkuva } \}.$$

Tällöin V on $\mathbb{C}$-vektoriavaruus (laskutoimitukset pisteittäin) ja kaava

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)} dx$$

määrittelee avaruudessa V sisätulon. Avaruuden V määritelmässä voidaan korvata vaatimus "f on jatkuva" vaatimuksella f on Riemann-integroituva" tai jopa "Lebesguen-integroituva", jos mitta-teoria on tuttu. Tässä siis kompleksilukuarvoisen funktion $f = u + iv$, missä $u, v \colon [a, b] \to \mathbb{R}$, integraali määritellään "komponenteittain" eli kaavalla

$$\int f = \int u + i \int v.$$

Kaikkien Lebesguen-integroituvien funktioiden avaruutta tällä sisätulolla varustettuna merkitään symbolilla $L^2([a, b]; \mathbb{C})$. Tämä ääretönulotteinen sisätulo-avaruus on *erittäin tärkeä esimerkki* täydellisestä *sisätuloavaruudesta eli niin sanotusta Hilbertin avaruudessa. "Täydellisyys" tässä viittaa tämän avaruuden metrisen topologian ominaisuuteen, jonka mukaan jokainen jono, joka "näyttää suppenevan" suppenee oikeasti*[2]*. Jos taas avaruuteen V otetaan mukaan vain kaikki jatkuvat funktiot tai jopa kaikki Riemann-integroituvat funktiot, niin näin saatu sisätuloavaruus ei ole enää täydellinen. Tämä on yksi tärkeimpiä syitä siihen miksi Lebesguen integraali on tarpeellinen ja miksi se on teoreettisesta näkökulmasta "parempi" kuin Riemann-integraali. Tarkemmin näistä asioista ja L^p-avaruuksista yleisemmin puhutaan esimerkiksi kursseilla "Mitta ja integraali", "Reaalianalyysi" ja "Funktionaalianalyysin peruskurssi".*

Olkoon V sisätuloavaruus ja olkoot $\mathbf{v}, \mathbf{w} \in V$. Vektoreita $\mathbf{v}, \mathbf{w}$ sanotaan olevan *kohtisuorassa (toisiaan vastaan)* eli *ortogonaalisia* jos $\langle \mathbf{v}, \mathbf{w} \rangle = 0_K$. Huomaa, että kyseessä on symmetrinen relaatio, sillä jos $\langle \mathbf{v}, \mathbf{w} \rangle = 0_K$, niin myös $\langle \mathbf{w}, \mathbf{v} \rangle = \overline{0_K} = 0_K$. Kun vektorit $\mathbf{v}$ ja $\mathbf{w}$ ovat ortogonaalisia merkitään myös $\mathbf{v} \perp \mathbf{w}$.

Olkoot $\mathbf{v}, \mathbf{w} \in V$ mielivaltaisia. Tutkitaan voidaanko vektori $\mathbf{v}$ kirjoittaa muodossa

$$\mathbf{v} = k\mathbf{w} + \mathbf{u},$$

missä $k \in K$ on skalaari ja $\mathbf{u} \perp \mathbf{w}$. Havainnollisesti sanoen tutkitaan onko mahdollista jakaa vektori $\mathbf{v}$ kahden "komponentin" summaksi, jossa toinen komponentti on yhdensuuntainen vektorin $\mathbf{w}$ kanssa ja toinen komponentti on kohtisuorassa vektoria $\mathbf{w}$ vastaan. Vektoria $k\mathbf{w}$ on tällöin luonnollista kutsua vektorin $\mathbf{v}$ *kohtisuoraksi projektioksi* vektorin $\mathbf{w}$ suuntaan.

Jos $\mathbf{w} = \mathbf{0}_V$, asia on selvä - asetetaan $\mathbf{u} = \mathbf{v}$ ja $k \in K$ voidaan valita vapaasti. Olkoon $\mathbf{w} \neq \mathbf{0}_V$ ja oletetaan, että $\mathbf{v} = k\mathbf{w} + \mathbf{u}$. Tällöin $\mathbf{u} = \mathbf{v} - k\mathbf{w}$, joten ehto $\langle \mathbf{u}, \mathbf{w} \rangle = 0_K$ on yhtäpitävä sen kanssa, että

$$0_K = \langle \mathbf{v}, \mathbf{w} \rangle - k\langle \mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle - k\, |\mathbf{w}|^2 .$$

Tästä nähdään, että ehto $\langle \mathbf{u}, \mathbf{w} \rangle = 0_K$ pätee jos ja vain jos valitaan

$$k = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{w}|^2}.$$

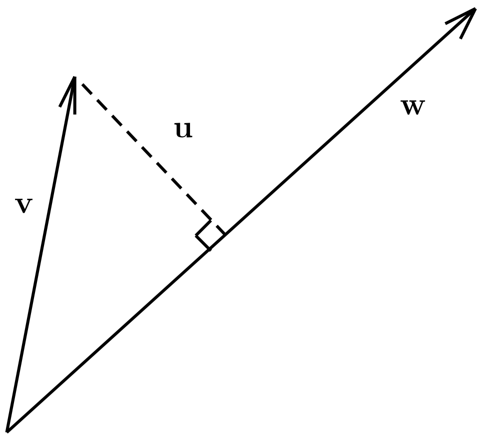
On osoitettu seuraava tulos.

Lemma 4.12. *Olkoon V sisätuloavaruus ja olkoot $\mathbf{v}, \mathbf{w} \in V$. Oletetaan, että $\mathbf{w} \neq \mathbf{0}_V$. Tällöin*

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{w}|^2}\mathbf{w} + \mathbf{u},$$

missä $\mathbf{u} \perp \mathbf{w}$.

[2] Täsmällisesti: jossa jokainen Cauchy-jono suppenee.



Normin ominaisuuksia

Nyt voidaan palata sisätulon määrittelemän normin ominaisuuksiin. Aloitetaan yksinkertaisesta tuloksesta, jonka tunnetaan myös nimellä *Pythagoran lause.*

Lemma 4.13. *Olkoon V sisätuloavaruus. Oletetaan, että vektorit $\mathbf{v}, \mathbf{w}$ ovat kohtisuorassa toisiaan vastaan. Tällöin*

$$|\mathbf{v} + \mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 .$$

Todistus. Normin määritelmän, oletuksen sekä sisätulon ominaisuuksien nojalla pätee

$$|\mathbf{v} + \mathbf{w}|^2 = \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle = |\mathbf{v}|^2 + |\mathbf{w}|^2 .$$

□

Seuraavaksi todistetaan *Cauchy-Schwarzin epäyhtälö.* Kirjallisuudessa sille annetaan usein monimutkainen ja tekninen todistus, jossa käytetään toisen asteen yhtälön diskriminanttia. Tässä materiaalissa sille tarjotaan yksinkertaisempi ja intuitiivisempi todistus, joka perustuu ortogonaalisen projektion käsitteeseen.

Lemma 4.14. Cauchy-Schwarzin epäyhtälö.
Olkoon $\langle , \rangle$ sisätulo K-vektoriavaruudessa V. Olkoot $\mathbf{v}, \mathbf{w} \in V$. Tällöin

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq |\mathbf{v}||\mathbf{w}| \tag{4.15}$$

Tässä vasemmalla puolella $|\cdot|$ tarkoittaa kompleksilukujen itseisarvoa ja oikealla puolella sisätulon määritelemää normia $|\mathbf{v}| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. Lisäksi Cauchy-Schwarzin epäyhtälö pätee yhtälönä

$$|\langle \mathbf{v}, \mathbf{w} \rangle| = |\mathbf{v}||\mathbf{w}|$$

jos ja vain jos joukko $\{\mathbf{v}, \mathbf{w}\}$ on sidottu vektoriavaruudessa V.

Todistus. Jos $\mathbf{w} = \mathbf{0}_V$, epäyhtälö 4.15 selvästi pätee jopa yhtälönä. Huomaa, että tällöin joukko $\{\mathbf{v}, \mathbf{w}\}$ on sidottu, sillä se sisältää nolla-vektorin.

Oletetaan, että $\mathbf{w} \neq \mathbf{0}_V$. Tällöin Lemman 4.12 nojalla

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{w}|^2}\mathbf{w} + \mathbf{u},$$

missä $\mathbf{u} \perp \mathbf{w}$. Vektorit $\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{w}|^2} \mathbf{w}$ ja $\mathbf{u}$ ovat tällöin ortogonaalisia, joten Pythagoran lauseen sekä yhtälön 4.9 nojalla pätee

$$|\mathbf{v}|^2 = \left(\frac{|\langle \mathbf{v}, \mathbf{w} \rangle|}{|\mathbf{w}|^2} \right)^2 |\mathbf{w}|^2 + |\mathbf{u}|^2 = \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{|\mathbf{w}|^2} + |\mathbf{u}|^2 .$$

Koska $|\mathbf{u}|^2 \geq 0$, tästä nähdään, että

$$|\mathbf{v}|^2 \geq \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{|\mathbf{w}|^2},$$

mistä Cauchy-Schwarzin epäyhtälö seuraa suoraan. Huomaa itseisarvoja - niitä tarvitaan erityisesti kompleksisessä tapauksessa.

Edellisestä laskusta seuraa myös suoraan, että Cauchy-Schwarzin epäyhtälö pätee yhtälönä jos ja vain jos $\mathbf{u} = \mathbf{0}_V$. Tämä on puolestaan mahdollista jos ja vain jos pätee $\mathbf{v} \in \mathrm{Span}(\mathbf{w})$. Tämä on taas yhtäpitävä sen kanssa, että joukko $\{\mathbf{v}, \mathbf{w}\}$ on sidottu vektoriavaruudessa V. □

Seuraus 4.16. *Olkoon V K-sisätuloavaruus V ja olkoon $|\cdot|$ sen sisätulon määrämä normi. Olkoot $\mathbf{v}, \mathbf{w} \in V$, $k \in K$. Tällöin*

(1) $|\mathbf{v} + \mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}|$, *(kolmioepäyhtälö),*

(2) $|k\mathbf{v}| = |k|\,|\mathbf{v}|$,

(3) $|\mathbf{v}| \geq 0$. *Yhtälö* $|\mathbf{v}| = 0$ *pätee jos ja vain jos* $\mathbf{v} = \mathbf{0}_V$.

Todistus. Kohdat (2) ja (3) on osoitettu aikaisemmin. Kohdan (1) todistaminen jätetään harjoitustehtäväksi. □

Mikä tahansa K-vektoriavaruudessa V määriteltyä kuvausta $|\cdot| : V \to \mathbb{R}$, joka toteuttaa edellisen seurauksen ehdot (1)-(3), sanotaan vektoriavaruuden V *normiksi.* On siis osoitettu, että jokainen sisätulo määrittelee tietyllä kanonisella tavalla erään normin. Käänteinen väite ei päde - on olemassa vektoriavaruuden normeja, jotka eivät ole minkään sisätulon "määräämiä". Voidaan osoittaa, että vektoriavaruuden V normi $|\cdot|$ on jonkun sisätulon määräämä jos ja vain jos se kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee niin sanottu *suunnikassäänntö*

$$|\mathbf{v} + \mathbf{w}|^2 + |\mathbf{v} - \mathbf{w}|^2 = 2(|\mathbf{v}|^2 + |\mathbf{w}|^2).$$

Tason $\mathbb{R}^2$ standardinormin $|(x, y)| = \sqrt{x^2 + y^2}$ tapauksessa tämä sääntö ilmaisee sen geometrisen tosiasian, että suunnikaan lävistäjien pituuksien summa on sama kuin suunnikaan piiri eli kaikkien sivujen pituuksien summa. Tästä nimitys "suunnikassääntö" tuleekin.

Vektoriavaruutta, joka on varustettu jollakin normilla, sanotaan *normiavaruudeksi.* Jokaisessa normiavaruudessa, erityisesti siis jokaisessa sisätuloavaruudessa, on olemassa luonnollinen pisteiden välinen etäisyys, eli *metriikka.* Tämä määritellään kaavalla

$$d(\mathbf{v}, \mathbf{w}) = |\mathbf{v} - \mathbf{w}| .$$

Tämä kuvaus toteuttaa tällöin kaikki metriikan[3] aksioomat eli

$$d(\mathbf{v},\mathbf{u}) \leq d(\mathbf{v},\mathbf{w}) + d(\mathbf{w},\mathbf{u}) \text{ kaikilla } \mathbf{v},\mathbf{u},\mathbf{w} \in V, \text{ (kolmioepäyhtälö)},$$

$$d(\mathbf{v},\mathbf{w}) = d(\mathbf{w},\mathbf{v}) \text{ kaikilla } \mathbf{v},\mathbf{w} \in V, \text{ (symmetrisyys)},$$

$$d(\mathbf{v},\mathbf{v}) \geq 0 \text{ kaikilla } \mathbf{v} \in V,$$

$$d(\mathbf{v},\mathbf{v}) = 0 \text{ jos ja vain jos } \mathbf{v} = \mathbf{0}_V.$$

Näin ollen, normiavaruudessa voidaan tutkia sen *topologisia* eli "geometrisia" ominaisuuksia. Tällä kurssilla emme ole niistä erityisen kinnostuneita, sillä mielenkiinnon kohteena on algebra, ei topologia.

Cauchy-Schwarzin epäyhtälön avulla $\mathbb{R}$-sisätuloavaruuksissa voidaan määritellä myös vektorien välisiä *kulmia*. Olkoot $\mathbf{v},\mathbf{w}$ nollavektorista eroavia $\mathbb{R}$-sisätuloavaruuden V vektoreita. Tällöin Cauchy-Schwarzin epäyhtälöstä seuraa, että

$$-1 \leq \frac{\langle \mathbf{v},\mathbf{w}\rangle}{|\mathbf{v}|\cdot|\mathbf{w}|} \leq 1.$$

Trigonometristen funktioiden teoriasta seuraa tällöin, että on olemassa yksikäsitteinen $\alpha \in [0,\pi]$, jolle pätee

$$\cos\alpha = \frac{\langle \mathbf{v},\mathbf{w}\rangle}{|\mathbf{v}|\cdot|\mathbf{w}|}.$$

Tätä kulmaa α sanotaan vektorien $\mathbf{v}$, $\mathbf{w}$ väliseksi kulmaksi.

$\mathbb{C}$-kertoimisissa sisätuloavaruuksissa suure $\frac{\langle \mathbf{v},\mathbf{w}\rangle}{|\mathbf{v}|\cdot|\mathbf{w}|}$ on yleisesti ottaen kompleksiluku, joten ei ole mielekästä sanoa, että se sijaitsee lukujen -1 ja 1 välissä. Tällöin voidaan ainoastaan määritellä "vektorien välisen terävän kulman" kaavalla

$$\cos\alpha = \frac{|\langle \mathbf{v},\mathbf{w}\rangle|}{|\mathbf{v}|\cdot|\mathbf{w}|}.$$

Ortogonaalisuus

Avaruuden K^n standardikanta $E = (\mathbf{e}_1,\ldots,\mathbf{e}_n)$ toteuttaa tunnetusti avaruuden K^n "standardin" pistetulon $\cdot$ suhteen yhtälön

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij},$$

missä δ_{ij} on Kroneckerin delta. Näin ollen standardikanta E on esimerkki niin sanotusta *ortonormaalista* kannasta.

Määritelmä 4.17. *Olkoon V sisätuloavaruus. Osajoukon $A \subset V$ sanotaan olevan ortogonaalinen, jos kaikilla $\mathbf{v},\mathbf{w} \in A, \mathbf{v} \neq \mathbf{w}$, pätee $\langle \mathbf{v},\mathbf{w}\rangle = 0_K$.*

Ortogonaalisen osajoukon sanotaan olevaan ortonormaali jos lisäksi kaikilla $\mathbf{v} \in A$ pätee $\langle \mathbf{v},\mathbf{v}\rangle = 1_K$.

[3] "Metriikka" on topologinen käsite. Siitä puhutaan kurssilla "Topologia I".

Toisin sanoen osajoukko on ortogonaalinen jos kaikki sen vektorit ovat pareittain kohtisuorassa toisiaan vastaan. Osajoukko on lisäksi ortonormaali jos kaikki sen alkiot ovat "normeerattuja" eli niiden normi on tasan yksi.

Voidaan myös puhua ortogonaalisesta tai ortonormaalista jonosta. Tarkemmin sanottuna jono $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, joka koostuu sisätulon avaruuden V vektoreista, on ortogonaalinen, jos $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0_K$ pätee aina kun $i \neq j$. Se on ortonormaali jos lisäksi $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1_K$ kaikilla $i = 1, \ldots, n$.

Sisätulon avaruuden V *ortogonaalinen/ortonormaali kanta* on sellainen sen ortogonaalinen/ortonormaali osajoukko/jono, joka on lisäksi sen kanta. Seuraavan tuloksen mukaan jokainen ortonormaali osajoukko/jono on vapaa. Näin ollen se on kanta jos ja vain jos se virittää koko avaruuden V.

Lemma 4.18. *a) Sisätuloavaruuden ortonormaali osajoukko/jono on vapaa.*
b) Olkoon $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ ortonormaali jono sisätuloavaruudessa V. Oletetaan, että

$$\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{v}_i.$$

Tällöin kaikilla $i = 1, \ldots, n$ pätee $k_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$.

Todistus. Oletetaan, että A on ortonormaali. Olkoot $\mathbf{v}_1, \ldots, \mathbf{v}_n \in A$ eri vektorit. Riittää osoittaa, että jono $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ on vapaa, toisin sanoen, että jokaisen aliavaruuden $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ vektorin $\mathbf{v}$ esitys muodossa $\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{v}_i$ on yksikäsitteinen.

Oletetaan siis, että $\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{v}_i$. Tällöin jokaisella $j = 1, \ldots, n$ saadaan suoralla laskulla (sisätulon ominaisuuksien ja ortonormaalin jonon määritelmän nojalla), että

$$\langle \mathbf{v}, \mathbf{v}_j \rangle = \langle \sum_{i=1}^{n} k_i \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} k_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = k_j.$$

Lineaarisen kombinaation $\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{v}_i$ kertoimet ovat siis yksikäsitteisesti määrättyä. Koska tämä pätee kaikille osajoukon A alkiosta muodostetuille äärellisille jonoille $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, joukko A on vapaa. Sama todistus kelpaa myös äärellisen jonon tapauksessa. Näin ollen a)-kohdan väite on tosi. Lisäksi edellinen lasku osoittaa myös suoraan b)-kohdan väiteen todeksi. □

Huomautus: Yleisemmin nähdään helposti, että ortogonaalinen osajoukko on vapaa jos ja vain jos se ei sisällä nolla-vektoria.

Propositio 4.19. Gram-Schmidtin ortogonalisaatio.
Olkoon $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ sisätuloavaruuden V vapaa jono. Tällöin on olemassa ortonormaali jono $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ siten, että

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$$

kaikilla $j = 1, \ldots, n$.

Todistus. Haluttu jono $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ konstruoidaan induktiivisesti.

Alkuaskel. Määritellään $\mathbf{e}_1 = \mathbf{v}_1/|\mathbf{v}_1|$. Tällöin $\mathbf{e}_1$ on hyvin määritelty, sillä vapaa jono ei voi sisältää nolla-vektoria, joten erityisesti $|\mathbf{v}_1| \neq 0_K$.

Selvästi $|\mathbf{e}_1| = 1$ ja $\mathrm{Span}(\mathbf{e}_1) = \mathrm{Span}(\mathbf{v}_1)$.

Induktio-askel. Oletetaan, että $j < n$, ja oletetaan, että on konstruoitu ortonormaali jono $(\mathbf{e}_1, \ldots, \mathbf{e}_j)$ siten, että $\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$. Osoitetaan, että on olemassa $\mathbf{e}_{j+1} \in V$ siten, että jono $(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{e}_{j+1})$ on ortonormaali ja

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{e}_{j+1}) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \mathbf{v}_{j+1}).$$

Etsitään ensin sellainen $\mathbf{w} \in V$, $\mathbf{w} \neq \mathbf{0}_V$, jolle pätee

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w}) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \mathbf{v}_{j+1})$$

ja lisäksi jono $(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w})$ on ortogonaalinen. Nimittäin, jos sellainen $\mathbf{w}$ löytyy, riittää sen jälkeen normeerata se, eli asettaa

$$\mathbf{e}_{j+1} = \mathbf{w}/|\mathbf{w}|.$$

Koska jono $(\mathbf{e}_1, \ldots, \mathbf{e}_j)$ on ortogonaalinen induktio-oletuksen nojalla, jonosta $(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w})$ tulee ortogonaalinen jos ja vain jos uusi vektori $\mathbf{w}$ on kohtisuorassa vektoria $\mathbf{e}_i$ vastaan kaikilla $i = 1, \ldots, j$.

Koska induktio-oletuksen nojalla pätee $\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$, nähdään helposti, että jos yhtälö

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w}) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \mathbf{v}_{j+1})$$

pätee, niin myös pätee $\mathbf{w} \in \mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{v}_{j+1})$ (miksi?). Tästä syystä etsitään sopiva vektori $\mathbf{w}$ muodossa

$$\mathbf{w} = \mathbf{v}_{j+1} + k_1\mathbf{e}_1 + \ldots + k_j\mathbf{e}_j$$

joillakin skalaareilla $k_1, \ldots, k_j$. Huomataan heti, että jos $\mathbf{w}$ on tätä muotoa, niin pätee (yksityiskohtien verifiointi lukijalle mietittäväksi)

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w}) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \mathbf{v}_{j+1}).$$

Näin ollen, riittää osoittaa, että skalaarit $k_1, \ldots, k_j$ voidaan valita niin, että vektori muotoa $\mathbf{w} = \mathbf{v}_{j+1} + k_1\mathbf{e}_1 + \ldots + k_j\mathbf{e}_j$ toteuttaa vaaditun ortogonaalisuus-ehdon, eli kaikilla $i = 1, \ldots, j$ pätee yhtälö $\langle \mathbf{w}, \mathbf{e}_i \rangle = 0_K$.

Olkoon siis $\mathbf{w} = \mathbf{v}_{j+1} + k_1\mathbf{e}_1 + \ldots + k_j\mathbf{e}_j$. Tällöin jonon $(\mathbf{e}_1, \ldots, \mathbf{e}_j)$ ortonormaalisuuden nojalla pätee

$$\langle \mathbf{w}, \mathbf{e}_i \rangle = \langle \mathbf{v}_{j+1}, \mathbf{e}_i \rangle + \sum_{l=1}^{j} k_l \langle \mathbf{e}_l, \mathbf{e}_i \rangle = \langle \mathbf{v}_{j+1}, \mathbf{e}_i \rangle + k_i.$$

Tästä nähdään, että $\langle \mathbf{w}, \mathbf{e}_i \rangle = 0_K$, jos ja vain jos valitaan $k_i = -\langle \mathbf{v}_{j+1}, \mathbf{e}_i \rangle$. Näin ollen valitsemalla vektoriksi $\mathbf{w}$ vektori

$$\mathbf{w} = \mathbf{v}_{j+1} - \sum_{i=1}^{j} \langle \mathbf{v}_{j+1}, \mathbf{e}_i \rangle \mathbf{e}_i,$$

saadaan sellainen vektori $\mathbf{w} \neq \mathbf{0}_V$, jolle jono $(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w})$ on ortogonaalinen ja lisäksi pätee

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{w}) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j, \mathbf{v}_{j+1}).$$

Lukijalle jää sen verifiointi, että todellakin $\mathbf{w} \neq \mathbf{0}_V$. Asettamalla tämän jälkeen $\mathbf{e}_{j+1} = \mathbf{w}/|\mathbf{w}|$ saadaan todistettua väite tapauksessa $j+1$. Induktiivista konstruktiota jatketaan tällä tavalla kunnes päästään arvoon $j = n$. □

Seuraus 4.20. *Jokaisella äärellisulotteisella sisätuloavaruudella on ortonormaali kanta.*

Todistus. Valitaan ensin äärellisulotteiselle vektoriavaruudelle V jokin äärellinen kanta $E = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Koska E on tällöin erityisesti vapaa, edellisen proposition nojalla avaruudessa V on olemassa ortonormaali jono $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ siten, että kaikilla $j = 1, \ldots, n$ pätee

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_j) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_j).$$

Erityisesti tällöin

$$\mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V,$$

joten jono $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ virittää avaruuden V. Koska jokainen ortonormaali jono on vapaa (Lemma 4.18), jono $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ on kanta. Toinen tapa nähdä, että $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ on kanta, on huomata, että se on virittävä jono, jonka pituus on sama kuin avaruuden dimensio, jolloin johtopäätös perustuu Seuraukseen 2.49. □

Olkoon V äärellisulotteinen K-vektoriavaruus ja olkoon $\langle , \rangle$ sisätulo avaruudessa V. Edellisen lemman nojalla avaruudella V on olemassa ortonormaali kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Olkoot $\mathbf{v}, \mathbf{w} \in V$. Koska E on kanta, on olemassa yksikäsitteiset kerroinkunnan K alkiot $k_1, \ldots, k_n$, $k'_1, \ldots, k'_n$, siten, että

$$\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{v}_i \text{ ja } \mathbf{w} = \sum_{i=1}^{n} k'_i \mathbf{v}_i.$$

Itse asiassa Lemman 4.18 nojalla pätee $k_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$, $k'_i = \langle \mathbf{w}, \mathbf{v}_i \rangle$, mutta tätä tietoa ei tarvita nyt.

Sisätulon ominaisuuksien nojalla saadaan suoralla laskulla

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i,j}^{n} k_i \overline{k'_j} \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} k_i \overline{k'_i}. \tag{4.21}$$

Toisin sanoen ortonormaalissa kannassa n-ulotteisen sisätuloavaruuden sisätulo "näyttää" samanlaiselta kuin avaruuden K^n "tavallinen" pistetulo (kts. esim. 4.10). Formaalimmin sama havainto voidaan ilmaistaa, sanoa, että n-ulotteinen sisätuloavaruus V on *isomorfinen* (sisätuloavaruutena) avaruuden K^n kanssa, jossa jälkimmäinen on varustettu pistetulolla. Isomorfismi tässä yhteydessä tarkoittaa sellaista lineaarista bijektiota $L \colon V \to K^n$, joka säilyttää sisätulot. Tällaisia kuvauksia sanotaan myös *unitaarisiksi*, unitaarisia kuvauksia tutkitaan myöhemmin tarkemmin adjungaattien teorian yhteydessä.

Ortogonaalinen komplementti

Olkoon $A \subset V$ sisätuloavaruuden V mielivaltainen osajoukko. Sen *ortogonaalinen komplementti* $A^\perp$ määritellään kaavalla

$$A^\perp = \{\mathbf{w} \in V \mid \langle \mathbf{a}, \mathbf{w} \rangle = 0_K \text{ kaikilla } \mathbf{a} \in A\}.$$

Toisin sanoen ortogonaaliseen komplementtiin $A^\perp$ kuuluvat täsmälleen ne avaruuden V vektorit, jotka ovat kohtisuorassa *jokaista* joukon A vektoria vastaan.

Lemmassa 4.22 alla todetaan, että ortogonaalinen komplementti $A^\perp$ on avaruuden V *aliavaruus* kaikilla $A \subset V$.

Olkoon W äärellisulotteisen vektoriavaruuden V aliavaruus. Tällöin, kuten tiedämme yleisestä vektoriavaruuksien teoriasta, aliavaruudella W on olemassa avaruudessa V *komplementti* U, eli sellainen aliavaruus, jolle summa $W + U$ on suora ja pätee yhtälö $W \oplus U = V$ (Lemma 2.159). Yleensä tällaisia avaruuden W komplementteja on olemassa paljon erilaisia, eikä ole mitään "kanonista" tapa valita niistä yhtä "oikeata", joka olisi jotenkin erityisessä asemassa. Sisätuloavaruuksissa tällainen tapa kuitenkin löytyy, sillä osoittautuu, että aliavaruuden W ortogonaalinen komplementti $W^\perp$ on aina aliavaruuden W komplementti vektoriavaruuksien teorian mielessä. Lisäksi tällä komplementilla on mielenkiintoisia teoreettisia ominaisuuksia, jotka erottavat sen kaikista muista aliavaruuden W komplementeista (esimerkki tällaisesta annetaan Lemmassa 4.24 alla).

Lemma 4.22. *Olkoon V sisätuloavaruus. Tällöin seuraavat väitteet ovat voimassa.*

1) $A^\perp$ on avaruuden V aliavaruus jokaisella $A \subset V$.

2) Olkoon $W \leq V$ avaruuden V äärellisulotteinen aliavaruus, niin summa $W + W^\perp$ on suora ja

$$W \oplus W^\perp = V.$$

Toisin sanoen $W^\perp$ on tällöin aliavaruuden W komplementti vektoriavaruudessa V. (Huom, itse avaruutta V ei oleteta äärellisulotteiseksi).

Todistus. 1) Harjoitustehtävä.
2) Osoitetaan, että summa $W + W^\perp$ on suora ja että sen arvo on koko avaruus V.

Tarkastelemalla avaruuden V sisätulon rajoittumaa joukkoon $W \times W$, nähdään, että W on äärellisulotteinen sisätuloavaruus. Lemman 4.19 nojalla sisätuloavaruudella W on olemassa ortonormaali kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Todetaan ensin, että

$$W^\perp = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}^\perp. \tag{4.23}$$

On selvä, että $W^\perp \subset \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}^\perp$. Kääntäen, oletetaan, että $\mathbf{u} \in \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}^\perp$ ja olkoon $\mathbf{w} \in W$. Koska E on avaruuden W kanta, on olemassa $k_1, \ldots, k_n$ siten, että $\mathbf{w} = \sum_{i=1}^n k_i \mathbf{e}_i$. Koska sisätulo on lineaarinen ensimmäisen muuttujan suhteen, tästä seuraa, että

$$\langle \mathbf{w}, \mathbf{u} \rangle = \sum_{i=1}^n k_i \langle \mathbf{e}_i, \mathbf{u} \rangle = 0_K,$$

joten $\mathbf{u} \in W^\perp$. On todistettu, että yhtälö 4.23 pitää paikkansa.

Olkoon $\mathbf{v} \in V$ mielivaltainen. On osoitettava, että on olemassa yksikäsitteiset $\mathbf{w} \in W$ ja $\mathbf{u} \in W^\perp$ siten, että $\mathbf{v} = \mathbf{w} + \mathbf{u}$. Edellisen nojalla ehto $\mathbf{u} = \mathbf{v} - \mathbf{w} \in W^\perp$ pätee jos ja vain jos kaikilla $j = 1, \ldots, n$ pätee $\langle \mathbf{v} - \mathbf{w}, \mathbf{e}_j \rangle = 0_K$. Toisaalta vektori $\mathbf{w}$ on aliavaruuden W vektori jos ja vain jos $\mathbf{w} = \sum_{i=1}^n k_i \mathbf{e}_i$ joillakin skalaareilla $k = 1, \ldots, n$. Tällöin jokaisella $j = 1, \ldots, n$ jonon E ortonormaalisuuden nojalla pätee

$$\langle \mathbf{v} - \mathbf{w}, \mathbf{e}_j \rangle = \langle \mathbf{v}, \mathbf{e}_j \rangle - k_j.$$

Tästä nähdään, että ehto $\langle \mathbf{v} - \mathbf{w}, \mathbf{e}_j \rangle = 0_K$ toteutuu kaikilla $j = 1, \ldots, n$ vektorille $\mathbf{w} = \sum_{i=1}^n k_i \mathbf{e}_i$ jos ja vain jos asetetaan
$k_j = \langle \mathbf{v}, \mathbf{e}_j \rangle$, $j = 1, \ldots, n$. Erityisesti siis esitys muodossa $\mathbf{v} = \mathbf{w} + \mathbf{u}$, missä $\mathbf{w} \in W$, $\mathbf{u} \in W^\perp$, löytyy jokaiselle $\mathbf{v} \in V$. Lisäksi edellisen tarkastelun nojalla tällainen esitys on *yksikäsitteinen*. Suoran summan määritelmän nojalla tämä implikoi sen, että summa $W + W^\perp$ on suora ja lisäksi $W \oplus W^\perp = V$. $\square$

Huomautus: Tapauksessa $W = \mathrm{Span}(\mathbf{w})$ on 1-ulotteinen, tulos $W \oplus W^\perp = V$ on meille tuttu ennestään - kyse on samasta asiasta kuin yhtälössä 4.12 eli vektorin $\mathbf{v}$ jakamisesta kahden komponentin summaksi, jossa toinen komponentti on vektorin $\mathbf{w}$ suuntainen ja toinen kohtisuorassa vektoria $\mathbf{w}$ vastaan.

Ortogonaalinen projektio

Olkoon W sisätuloavaruuden V äärellisulotteinen aliavaruus. Koska tällöin edellisen nojalla pätee $V = W \oplus W^\perp$, on olemassa tähän suoraan summaan liittyvä kanoninen projektiokuvaus $p_W \colon V \to W$, jota sanotaan tässä yhteydessä myös *ortogonaaliseksi projektioksi aliavaruudelle* W. Alkioiden tasolla projektio p_U on määritelty seuraavasti. Esitetään vektori $\mathbf{v} \in V$ muodossa $\mathbf{v} = \mathbf{w} + \mathbf{u}$, missä $\mathbf{w} \in W$ ja $\mathbf{u} \in W^\perp$. Edellisen tuloksen nojalla tällainen esitys on olemassa ja yksikäsitteinen. Projektiokuvauksen määritelmän mukaan tällöin pätee $p_W(\mathbf{v}) = \mathbf{w}$. Projektiokuvaus on lineaarinen.

Ortogonaalisen projektion tärkeys piilee seuraavassa tuloksessa. Sen mukaan $p_W(\mathbf{v})$ on kaikista aliavaruuden W vektoreista *lähimpänä* vektoria $\mathbf{v}$.

Propositio 4.24. *Olkoon V sisätuloavaruus ja $W \leq V$. Olkoon $\mathbf{v} \in V$. Tällöin kaikilla $\mathbf{w} \in W$ pätee*

$$|\mathbf{v} - p_W(\mathbf{v})| \leq |\mathbf{v} - \mathbf{w}|.$$

Lisäksi tämä epäyhtälö pätee yhtälönä jos ja vain jos $\mathbf{w} = p_W(\mathbf{v})$.

Todistus. Olkoon $\mathbf{w} \in W$. Tällöin $\mathbf{x} = p_W(\mathbf{v}) - \mathbf{w}$ on aliavaruuden W vektori, kun taas $\mathbf{y} = \mathbf{v} - p_W(\mathbf{v})$ on ortogonaalisen projektio määritelmän mukaan ortogonaalisen komplementin $W^\perp$ vektori. Toisin sanoen vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat kohtisuorassa toisiaan vastaan. Pythagoran lauseen nojalla tästä seuraa, että

$$|\mathbf{v} - \mathbf{w}|^2 = |\mathbf{x} + \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 = |\mathbf{v} - p_W(\mathbf{v})|^2 + |p_W(\mathbf{v}) - \mathbf{w}|^2 \geq |\mathbf{v} - p_W(\mathbf{v})|^2.$$

Väite saadaan ottamalla epäyhtälön kummastakin puolesta neliöjuuri. Lisäksi yllä olevasta seuraa, että yhtälö

$$|\mathbf{v} - p_W(\mathbf{v})| = |\mathbf{v} - \mathbf{w}|$$

pätee jos ja vain jos $|p_W(\mathbf{v}) - \mathbf{w}| = 0_K$ eli silloin kun $\mathbf{w} = p_W(\mathbf{v})$. $\square$

Esimerkki 4.25. *Edellistä tulosta käytetään hyväksi sovelluksissa sekä numeeristen menetelmien teoriassa. Nimittäin Propositio 4.24 voidaan tulkita tarkoittavan sitä, että vektorin* $\mathbf{v} \in V$ *"paras approksimaatio" aliavaruuden* W *vektorilla on vektori* $p_W(\mathbf{v})$*. Näiden etäisyys* $|\mathbf{v} - p_W(\mathbf{v})|$ *tällöin mittaa sitä, kuinka hyvin ortogonaalinen projektio approksimoi vektoria* $\mathbf{v}$*. Jos tämä etäisyys on "pieni", approksimaatio on "hyvä".*

Esimerkiksi olkoon V *kaikkien jatkuvien kuvausten* $f\colon [-\pi, \pi] \to \mathbb{R}$ *muodostama vektoriavaruus. Tässä (ääretönulotteisessa)* $\mathbb{R}$*-vektoriavaruudessa (laskutoimitukset pisteittäisiä) on olemassa integraalin määrämä sisätulo* $\langle , \rangle$,

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg \; dx$$

(kts. esimerkki 4.10).

Olkoon $W = \mathrm{Span}(1, x, x^2, x^3, x^4, x^5)$ *kaikkien korkeintaan viidettä astetta olevien polynomikuvausten* $f\colon \mathbb{R} \to \mathbb{R}$ *muodostama avaruuden* V *aliavaruus. Tunnetusti* W *on äärellisulotteinen ja* $E = (1, x, x^2, x^3, x^4, x^5)$ *on sen kanta. Olkoon* $g \in V$, $g(x) = \sin x$*. Tällöin ortogonaalista projektiota* $p_W(g)$ *voidaan ajatella sellaisena korkeintaan viidennen asteen polynomina, joka on kaikista tällaisista polynomeista lähimpänä sinifunktiota eli approksimoi sen mahdollisimman hyvin integraalin mielessä.*

Polynomin $p_W(g)$ *määrittämiseksi pitää ensin keksiä aliavaruudelle* W *jokin ortonormaali kanta. Tämä onnistuu esimerkiksi soveltamalla Gram-Schmidtin ortogonaalisaatioalgoritmia kantaan* $E = (1, x, x^2, x^3, x^4, x^5)$*. Näin saadaan avaruudelle* W *jokin ortonormaali kanta* $E' = (f_1, \ldots, f_6)$*. Tämän jälkeen huomataan, että Proposition 4.22 todistuksen nojalla vektorin* g *ortonormaali projektio* $P_W(g)$ *saadaan kaavalla*

$$P_W(g) = \sum_{i=1}^{6} k_i f_i, \ \textit{missä}$$

$$k_j = \langle g, f_j\rangle = \int_{-\pi}^{\pi} f_i g dx.$$

Käymällä läpi teknisiä laskuja, voidaan näyttää, että sinifunktion paras approksimaatio viidennen asteen polynomilla (integraalin mielessä) on polynomi

$$0,987862x - 0,155271x^3 + 0,00564312x^5,$$

missä tarkat kertoimet on korvattu niiden desimaaliapproksimaatiolla tietyllä tarkuudella.

Onko saatu approksimaatio "hyvä" selviää kunhan lasketaan etäisyys $|g - P_W(g)|$*, joka on tässä tapauksessa integraalin*

$$\int_{-\pi}^{\pi} (g - P_W(g))^2 dx$$

neliöjuuri. Osoittautuu, että tämä on hyvin pieni luku, joten approksimaatio on "tarpeeksi hyvä" käytännön sovellusten kannalta. Sen avulla voidaan esimerkiksi laskea sinifunktion arvoja välillä $[-\pi, \pi]$ *laskemalla tarkkojen arvojen sijaan polynomifunktion* $P_W(g)$ *arvoja. Koska polynomin arvoja on helpompi laskea kuin trigonometristen funktioiden arvoja* *(riittää osata laskea yhteen ja kertoa keskenään reaalilukuja), tästä saadaan eräs tapa laskea sinifunktion likiarvoja käytännössä. Itse asiassa juuri tämäntyyppisillä menetelmillä niitä laskettiinkin ennen kuin laskimet ja tietokoneet keksittiin. Toisaalta myös laskimet ja tietokoneet käyttävät tämäntyyppisiä numeerisia menetelmiä kun ne laskevat trigonometristen funktioiden arvoja.*

4.2. Adjungaatti ja unitaarisuus

Olkoon V äärellisulotteinen vektoriavaruus. Kuten Luvun 2 sisällön perusteella tiedetään, duaaliavaruus V^* on tällöin isomorfinen avaruuden V kanssa (Lause 2.105). Yleisesti ottaen kuitenkin isomorfismi $V \to V^*$ ei ole "kanoninen" eli se riippuu kantojen valinnasta eikä mikään isomorfismeista $V \to V^*$ ole erikoisasemassa.

Osoittautuu, että kun äärellisulotteinen avaruus V varustetaan jollakin sisätulolla, on olemassa kanoninen ("luonnollinenä") bijektio $\Phi\colon V \to V^*$. Valitettavasti yleisessä tapauksessa tämä bijektio Φ ei ole lineaarinen, vaan ainoastaan antilineaarinen. Toisin sanoen tämä kanoninen bijektio on vektoriavaruuksien välinen isomorfismi ainostaan kun kerroinkunta on reaalilukujen kunta $\mathbb{R}$.

Esitetään tämän kanonisen kuvauksen Φ konstruktio. Olkoon V sisätuloavaruus (ei välttämättä äärellisulotteinen) ja olkoon $\mathbf{w} \in V$. Määritellään kuvaus $L_w\colon V \to K$ kaavalla $L_{\mathbf{w}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle$. Koska sisätulo on lineaarinen ensimmäisen muuttujan suhteen, kuvaus $L_{\mathbf{w}}$ on tällöin lineaarinen jokaisella $\mathbf{w} \in V$. Toisin sanoen $L_{\mathbf{w}} \in V^*$ kaikilla $\mathbf{w} \in V$ ja voidaan määritellä kuvaus $\Phi\colon V \to V^*$ kaavalla

$$\Phi(\mathbf{w}) = L_{\mathbf{w}}, \mathbf{w} \in V.$$

Propositio 4.26. *Kuvaus $\Phi\colon V \to V^*$ on antilineaarinen injektio.*
Jos avaruus V on äärellisulotteinen, Φ on bijektio, erityisesti lineaarinen isomorfismi $V \to \overline{V^*}$.

Todistus. Osoitetaan, että kuvaus Φ on antilineaarinen. Kaikilla $\mathbf{w}, \mathbf{w}', \mathbf{v} \in V$ ja $k \in K$ pätee sisätulon määritelmän nojalla

$$\Phi(\mathbf{w} + \mathbf{w}')(\mathbf{v}) = L_{\mathbf{w}+\mathbf{w}'}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} + \mathbf{w}' \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w}' \rangle =$$

$$\Phi(\mathbf{w})(\mathbf{v}) + \Phi(\mathbf{w}')(\mathbf{v}) = (\Phi(\mathbf{w}) + \Phi(\mathbf{w}'))(\mathbf{v}) \text{ ja}$$
$$\Phi(k\mathbf{w})(\mathbf{v}) = \langle \mathbf{v}, k\mathbf{w} \rangle = \overline{k}\langle \mathbf{v}, \mathbf{w} \rangle = (\overline{k}\Phi(\mathbf{w}))(\mathbf{v}).$$

Näin ollen kuvaus $\Phi\colon V \to V^*$ on antilineaarinen. Lemman 4.2 nojalla Φ on lineaarinen, kun se ajatellaan kuvauksena $V \to \overline{V^*}$ (tai yhtä hyvin kuvauksena $\overline{V} \to V^*$).

Osoitetaan seuraavaksi, että Φ on injektio. Olkoon $\mathbf{w} \in W$ sellainen, että $\Phi(\mathbf{w}) = \mathbf{0}_{V^*}$. Tällöin kaikilla $\mathbf{v} \in V$ pätee

$$\langle \mathbf{v}, \mathbf{w} \rangle = L_{\mathbf{w}}(\mathbf{v}) = \Phi(\mathbf{w})(\mathbf{v}) = 0_K.$$

Erityisesti valitsemalla $\mathbf{v} = \mathbf{w}$ saadaan, että $\langle \mathbf{w}, \mathbf{w} \rangle = 0$, mistä sisätulon määritelmän nojalla seuraa, että $\mathbf{w} = \mathbf{0}_V$.

Edellisessä kappaleessa on näytetty, että lineaarisen kuvauksen $\Phi\colon V \to \overline{V^*}$ ydin $\operatorname{Ker}\Phi$ on triviaali, joten Φ on injektio.

Oletetaan, että V on äärellisulotteinen. Helposti nähdään (mieti!), että tällöin myös sen konjugaattiavaruus $\overline{V}$ on äärellisulotteinen ja itse asiassa $\dim \overline{V} = \dim V$. Koska Φ on lineaarinen injektio $\overline{V} \to V^*$ ja koska

$$\dim \overline{V} = \dim V = \dim V^*,$$

se on jopa bijektio (Seuraus 2.94). $\square$

Seuraus 4.27. *Olkoon V äärellisulotteinen sisätuloavaruus ja olkoon $L \in V^*$ lineaarinen muoto. Tällöin on olemassa yksikäsitteinen $\mathbf{w} \in V$ siten, että*

$$L(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle \textit{ kaikilla } \mathbf{v} \in V.$$

Todistus. Tämä on vain toinen tapa ilmaistaa se tosiasia, että kuvaus $\Phi\colon V \to V^*$ on bijektio, kun V on äärellisulotteinen. $\square$

Lineaarisen kuvauksen adjungaatti

Olkoot V, W äärellisulotteisia sisätuloavaruuksia ja olkoon $L\colon V \to W$ lineaarinen kuvaus. Tällöin on olemassa duaalikuvaus $L^*\colon W^* \to V^*$, joka on erityisesti lineaarinen kuvaus. Proposition 4.26 nojalla on olemassa kanoniset antilineaariset bijektiot $\Phi_V\colon V \to V^*$ ja $\Phi_W\colon W \to W^*$. Voidaan muodostaa yhdistetty kuvaus

$$L' = \Phi_V^{-1} \circ L^* \circ \Phi_W\colon W \to V.$$

Antilineaarisen bijektion käänteiskuvaus on selvästi myös antilineaarinen (yksinkertainen Lemman 4.2 sovellus) , joten $\Phi_V^{-1}\colon V^* \to V$ on antilineaarinen. Samoin nähdään helposti, että antilineaarisen ja lineaarisen kuvauksen yhdiste on antilineaarinen, kun taas kahden antilineaarisen kuvauksen yhdiste on lineaarinen kuvaus. Erityisesti yhdistetty kuvaus $\Phi_V^{-1} \circ L^*$ on antilineaarinen, joten yllä konstruoitu kuvaus $L' = \Phi_V^{-1} \circ L^* \circ \Phi_W$ on lineaarinen kuvaus $W \to V$. Tässä duaalikuvaus L^* ajatellaan kuvauksena $\overline{W^*} \to \overline{V^*}$, mutta näin tulkittuna se on edellenkin lineaarinen.

Karakterisoidaan yllä konstruoitu kuvaus $L'\colon W \to V$ suoraan alkioiden tasolla ilman viittauksia duaaliavaruuksiin. Olkoot $\mathbf{v} \in V$ ja $\mathbf{w} \in W$. Tällöin duaalikuvauksen ja kuvauksen Φ määritelmien nojalla pätee

$$(L^* \circ \Phi_W)(\mathbf{w}) = L^*(\Phi_W(\mathbf{w})) = \Phi_W(\mathbf{w}) \circ L,$$

joten

$$(L^* \circ \Phi_W)(\mathbf{w})(\mathbf{v}) = (\Phi_W(\mathbf{w}) \circ L)(\mathbf{v}) = \Phi_W(\mathbf{w})(L(\mathbf{v})) = \langle L(\mathbf{v}), \mathbf{w} \rangle.$$

Toisaalta pätee $L^* \circ \Phi_W = \Phi_V \circ L'$ ja lisäksi

$$(\Phi_V \circ L')(\mathbf{w})(\mathbf{v}) = \Phi_V(L'(\mathbf{w}))(\mathbf{v}) = \langle \mathbf{v}, L'(\mathbf{w}) \rangle.$$

Vertaamalla tuloksia nähdään, että kaikilla $\mathbf{v} \in V$ ja $\mathbf{w} \in W$ pätee yhtälö

$$\langle L\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, L'\mathbf{w} \rangle. \tag{4.28}$$

Kääntäen, oletetaan, että $L' \colon W \to V$ on sellainen kuvaus, jolle yhtälö 4.28 pätee kaikilla $\mathbf{v} \in V$ ja $\mathbf{w} \in W$. Tällöin käymällä läpi samoja välivaiheita käänteisessä järjestyksessä nähdään, että kaikilla $\mathbf{v} \in V$ ja $\mathbf{w} \in W$ pätee

$$(L^* \circ \Phi_W)(\mathbf{w})(\mathbf{v}) = (\Phi_V \circ L')(\mathbf{w})(\mathbf{v}),$$

joten $L^* \circ \Phi_W = \Phi_V \circ L'$. Koska $\Phi_V : V \to V^*$ on bijektio tästä seuraa, että

$$L' = \Phi_V^{-1} \circ L^* \circ \Phi_W$$

on yllä konstruoitu kuvaus, erityisesti yksikäsitteisesti määrätty ehdolla 4.28.

Koska kuvausta L' voidaan havainnollisesti ajatella " (antilineaarista) kanonista isomorfismia vaille" olevan sama asia kuin duaalikuvaus L^*, sitä merkitään yleensä kirjallisuudessa samalla symbolilla L^* ja sanotaan kuvauksen $L \colon W \to V$ **adjungaatiksi** (engl. adjoint). Kyseessä on siis jälleen kerran esimerkki tupla-notaatiosta, joka on kuitenkin hyvin vakiintunut, joten mekin noudatamme tätä käytäntöä ja merkitsemme adjungaattikuvausta symbolilla L^* tästä lähtien. Yksi tapa erottaa duaalikuvaus adjungaatista on huomata, että duaalikuvaus on kuvaus $L^* \colon W^* \to V^*$ kun taas adjungaattikuvaus on kuvaus $L^* \colon W \to V$. Adjungaatti voidaan määritellä ainoastaan kun vektoriavaruudet V ja W ovat sisätuloilla varustettuja ja se riippuu näiden sisätulojen valinnoista.

Seuraava tulos on todistettu tarkastelussa yllä.

Lemma 4.29. *Olkoot V ja W äärellisulotteisia sisätuloavaruuksia ja olkoon $L \colon V \to W$ lineaarinen kuvaus. Tällöin on olemassa yksikäsitteinen lineaarinen kuvaus $L^* \colon W \to V$ siten, että kaikilla $\mathbf{v} \in V$ ja $\mathbf{w} \in W$ pätee yhtälö*

$$\langle L(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, L^*(\mathbf{w}) \rangle. \tag{4.30}$$

Kuvausta L^* sanotaan kuvauksen L adjungaatiksi.

Tutkitaan seuraavaksi mikä yhteys on olemassa kuvauksen L ja sen adjungaatin matriisiesityksien välillä. Olkoon $E = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ avaruuden V **ortonormaali** kanta ja olkoon $F = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ avaruuden W **ortonormaali** kanta. Olkoon $A = (a_{ij})_{i=1,\ldots,m, j=1,\ldots,n}$ kuvauksen L matriisi $[L]_{F,E}$ kantojen E ja F suhteen. Määritelmän mukaan tämä tarkoittaa sitä, että

$$L(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{w}_i.$$

kaikilla $j = 1, \ldots, n$. Lemman 4.18 nojalla kaikilla $j = 1, \ldots, n$, $i = 1, \ldots, m$ pätee tällöin

$$a_{ij} = \langle L(\mathbf{v}_j), \mathbf{w}_i \rangle,$$

koska kanta F on ortonormaali. Samalla tavalla nähdään, että matriisin $[L^*]_{E,F} = (b_{ij})_{i=1,\ldots,m, j=1,\ldots,n}$ kertoimet määräytyvät yhtälöistä

$$b_{ij} = \langle L^*(\mathbf{w}_j), \mathbf{v}_i \rangle.$$

Sisätulon ominaisuuksien sekä yhtälön 4.30 nojalla tästä seuraa, että

$$\overline{b_{ij}} = \langle \mathbf{v}_i, L^*(\mathbf{w}_j) \rangle = \langle L(\mathbf{v}_i), \mathbf{w}_j \rangle = a_{ji},$$

kaikilla $i = 1, \ldots, n$, $j = 1, \ldots, m$. Tämä voidaan kirjoittaa myös muodossa $b_{ij} = \overline{a_{ji}}$. Toisin sanoen matriisi $[L^*]_{E,F}$ saadaan ottamalla matriisista $[L]_{F,E}$ transpoosi ja muuttamalla sen jälkeen kaikki matriisin alkiot konjugaatteikseen.

Määritelmä 4.31. *Olkoon $A = (a_{ij}) \in M(n \times m; K)$ matriisi. Sen konjugaatti $\overline{A}$ määritellään $(n \times m)$-kokoisena matriisina, jolle pätee $\overline{A}(i,j) = \overline{a_{ij}}$.*
Matriisin A adjungaatti A^* määritellään $(m \times n)$-matriisina $\overline{(A^T)} = (\overline{A})^T$.

Yllä ollaan todistettu seuraava tulos.

Lemma 4.32. *Olkoon $L\colon V \to W$ sisätuloavaruuksien V ja W välinen lineaarinen kuvaus. Olkoot $E = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ avaruuden V ortonormaali kanta ja olkoon $F = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ avaruuden W ortonormaali kanta. Tällöin adjungaattikuvauksen L^* matriisi kantojen F ja E suhteen on matriisin $[L]_{F,E}$ adjungaatti,*

$$[L^*]_{E,F} = [L]^*_{F,E}.$$

Korostetaan vielä, että kaava $[L^*]_{E,F} = [L]^*_{F,E}$ on voimassa ainoastaan kun kumpikin kanta E, F on *ortonormaali*. Jos kannat eivät ole ortonormaaleja, yhtälö ei välttämättä päde.

Helposti nähdään (tarkista!), että konjugaattikuvaus $A \mapsto \overline{A}$ on *antilineaarinen bijektio* $M(n \times m; K) \to M(n \times m; K)$. Tämän kuvauksen käänteiskuvaus on konjugaattikuvaus itse. Lisäksi tämä kuvaus säilyttää matriisien kertolaskun eli jos B on $(m \times p)$-kokoinen matriisi, niin pätee

$$\overline{AB} = \overline{A} \cdot \overline{B}.$$

Erityisesti, jos tätä operaatiota tarkastellaan K-algebrassa $M(n \times n; K)$, niin se säilyttää neliömatriisien kertolaskun ja lisäksi pätee $\overline{I_n} = I_n$.

Koska adjungaatti on yhdistelmä transpoosista (joka on lineaarinen operaatio) ja konjugaatista, operaatio $A \mapsto A^*$ on antilineaarinen bijektio $M(n \times m; K) \to M(m \times n; K)$. Lisäksi, koska transpoosi vaihtaa kertolaskun järjestystä (kts. kaava 2.112 aliluvussa 2.4), kaikille $A \in M(n \times m; K)$ ja $B \in M(m \times p; K)$ pätee yhtälö

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

Tapauksessa $K = \mathbb{R}$, matriisin adjungaatti on sama asia kuin matriisin transpoosi, $A^* = A^T$.

Karakterisoidaan vielä matriisin adjungaatti standardin avaruuden K^n pistetulon avulla. Olkoon $A \in M(n \times m; K)$ matriisi. Tällöin on olemassa kanoninen matriisiin A liittyvä kuvaus $L_A\colon K^m \to K^n$. Tämä kuvaus on itse asiassa yksikäsitteinen lineaarinen kuvaus $K^m \to K^n$, jolle pätee $[L_A]_{E',E} = A$, missä E ja E' ovat vastaavasti avaruuksien K^m ja K^n *standardikannat*. Koska nämä kannat ovat ortonormaaleja avaruuksien K^m ja K^n *pistetulon* (kts. esim. 4.10) suhteen, Lemmasta 4.32 yllä seuraa, että

$$A^* = ([L_A]_{E',E})^* = [L_A^*]_{E,E'}.$$

Toisaalta adjungaattikuvaus $L^*\colon K^n \to K^m$ on (Lemma 4.29) ainoa kuvaus, jolle pätee

$$L_A(\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot L_A^*(\mathbf{w})$$

kaikilla $\mathbf{v} \in K^n$, $\mathbf{w} \in K^m$. Lisäksi pätee $L_A(\mathbf{v}) = A\mathbf{v}$ ja vastaavasti $L_A^*(\mathbf{w}) = A^*\mathbf{w}$. Tässä avaruuden K^n vektorit tutkitaan $(n \times 1)$-matriisina eli "pystyvektorina" ja vastaavasti avaruudelle K^m. Näin ollen kaikilla $\mathbf{v}, \mathbf{w} \in K^n$ pätee

$$(A\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (A^*\mathbf{w}). \tag{4.33}$$

Tämä yhtälö karakterisoi adjungaattimatriisin yksikäsitteisesti.

Unitaariset kuvaukset

Sisätuloavaruus V voidaan ajatella uudentyyppisen matemaattisen struktuurin edustajana. Kyseessä on siis "sisätulostruktuurin tyyppi", joka on formaalisti pari $(V, (\langle , \rangle))$, missä $V = (V, +, \cdot)$ on vektoriavaruus ja $\langle , \rangle$ on eräs sisätulo V:ssä.

Aina kun tarkastelun kohteeksi tulee jokin uusi abstrakti matemaattisen struktuurin tyyppi, nykymatematiikan "hyviin tapoihin" kuuluu myös, että mietitään mitkä ovat tämän tyypin edustajien väliset *morfismit* eli struktuuria säilyttävät kuvaukset.

Määritelmä 4.34. *Olkoot V ja W sisätuloavaruuksia. Lineaarista kuvausta $L\colon V \to W$ sanotaan unitaariseksi, jos kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee*

$$\langle L(\mathbf{v}), L(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle.$$

Propositio 4.35. *Olkoot V ja W äärellisulotteisia sisätuloavaruuksia ja olkoon $L\colon V \to W$ lineaarinen kuvaus. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) L on unitaarinen.

(2) L säilyttää normit eli jokaisella $\mathbf{v} \in V$ pätee

$$|L(\mathbf{v})| = |\mathbf{v}| .$$

(3) L kuvaa ortonormaalit joukot ortonormaaleksi joukoiksi. Toisin sanoen jos $A \subset V$ on ortonormaali, niin myös $L(A) \subset W$ on ortonormaali.

(4) On olemassa sellainen avaruuden V ortonormaali kanta $E = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, jolle jono $F = (L(\mathbf{e}_1), \dots, L(\mathbf{e}_n))$ on myös ortonormaali.

Todistus. (1)$\Rightarrow$ (2). Asetetaan unitaarisen kuvauksen määritelmässä $\mathbf{v} = \mathbf{w}$. Tällöin saadaan

$$|L(\mathbf{v})|^2 = \langle L(\mathbf{v}), L(\mathbf{v}) \rangle = \langle \mathbf{v}, \mathbf{v} \rangle = |\mathbf{v}|^2 .$$

Ottamalla yhtälön kummastakin puolesta neliöjuuri saadaan väite (2) osoitettua.

(2)$\Rightarrow$ (1). Käänteisen väiteen todistusta varten tarvitsemme tavan esittää sisätulo normin avulla. Helposti nähdään laskemalla suoraan (takista laskut), että kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{1}{4}(|\mathbf{v} + \mathbf{w}|^2 - |\mathbf{v} - \mathbf{w}|^2),$$

jos V on $\mathbb{R}$-kertoiminen ja

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{1}{2}(|\mathbf{v} + \mathbf{w}|^2 + i\,|\mathbf{v} + i\mathbf{w}|^2 - (1+i)(|\mathbf{v}|^2 + |\mathbf{w}|^2)),$$

jos V on $\mathbb{C}$-kertoiminen. Jos oletetaan, että väite (2) on tosi, näiden yhtälöiden avulla saadaan helposti (1) todistettua. Yksityiskohtien verifiointi jätetään harjoitustehtäväksi.

(1)$\Rightarrow$(3). Oletetaan, että L on unitaarinen ja olkoon $A \subset V$ ortonormaali. Tällöin kaikilla $\mathbf{v}, \mathbf{w} \in A, \mathbf{v} \neq \mathbf{w}$, pätee

$$\langle L(\mathbf{v}), L(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = 0_K.$$

Lisäksi, koska tässä vaiheessa tiedetään jo, että ehto (1) implikoi ehdon (2), jokaisella $\mathbf{v} \in A$ pätee

$$|L(\mathbf{v})| = |\mathbf{v}| = 1_K.$$

Toisin sanoen $L(A)$ on ortonormaali.

(3)$\Rightarrow$(4). Tämä on melkein triviaalia.
(4)$\Rightarrow$(1). Olkoon $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ sellainen avaruuden V ortonormaali kanta jolle jono $F = (L(\mathbf{e}_1), \ldots, L(\mathbf{e}_n))$ on myös ortonormaali. Osoitetaan, että kuvaus L on unitaarinen. Olkoot $\mathbf{v}, \mathbf{w} \in V$. Esitetään kumpikin vektori kannassa E,

$$\mathbf{v} = \sum_{i=1}^{n} k_i \mathbf{e}_i \text{ ja } \mathbf{w} = \sum_{i=1}^{n} k_i' \mathbf{e}_i.$$

Tällöin, koska jono E on ortonormaali, pätee (kts. yhtälö 4.21)

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} k_i \overline{k_i'}.$$

Toisaalta lineaarisuuden nojalla pätee

$$L(\mathbf{v}) = \sum_{i=1}^{n} k_i L(\mathbf{e}_i) \text{ ja } L(\mathbf{w}) = \sum_{i=1}^{n} k_i' L(\mathbf{e}_i),$$

joten, koska F oletetaan ortonormaaliksi, yhtä hyvin pätee samalla perustelulla, että

$$\langle L(\mathbf{v}), L(\mathbf{w}) \rangle = k_i \overline{k_i'}.$$

Toisin sanoen kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee

$$\langle L(\mathbf{v}), L(\mathbf{w}) \rangle = k_i \overline{k_i'} = \langle \mathbf{v}, \mathbf{w} \rangle,$$

joten L on unitaarinen. □

Seuraus 4.36. *Olkoot V ja W äärellisulotteisia sisätuloavaruuksia ja olkoon $L \colon V \to W$ unitaarinen. Tällöin L on injektio. Erityisesti, jos V on äärellisulotteinen ja $L \colon V \to V$ on unitaarinen operaattori, L on isomorfismi.*

Todistus. Osoitetaan, että unitaarisen kuvauksen $L \colon V \to W$ ydin $\operatorname{Ker} L$ on triviaali. Olkoon $\mathbf{v} \in \operatorname{Ker} L$. Tällöin edellisen proposition nojalla

$$|\mathbf{v}| = |L(\mathbf{v})| = |\mathbf{0}_V| = 0_K,$$

mistä seuraa, että $\mathbf{v} = \mathbf{0}_V$.

Toinen väite seuraa suoraan siitä, että jokainen lineaarinen injektio $L \colon V \to V$ on isomorfismi, kun V on äärellisulotteinen (Seuraus 2.94). □

Jatkossa tutkitaan ainoastaan sellaisia unitaarisia kuvauksia, jotka ovat äärellisulotteisen sisätuloavaruuden operaattoreita eli unitaarisia kuvauksia $L \colon V \to V$, missä V on äärellisulotteinen. Osoittautuu, että tällainen unitaarinen operaattori voidaan karakterisoida sellaisena kääntyvänä operaattorina, jonka käänteiskuvaus on sen adjungaatti.

Lemma 4.37. *Olkoon $L\colon V \to V$ operaattori äärellisulotteisessa sisätuloavaruudessa V. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) L on unitaarinen.

(2) $L^*L = \mathrm{id}_V$.

(3) $LL^* = \mathrm{id}_V$.

(4) L on isomorfismi ja
$$L^{-1} = L^*.$$

Todistus. Koska V on äärellisulotteinen, ehdot (2)-(4) ovat yhtäpitäviä Seurauksen 2.95 nojalla (kun se käännetään lineaaristen kuvausten kielelle).

Olkoon $L\colon V \to V$ unitaarinen. Olkoot $\mathbf{v}, \mathbf{w} \in V$. Tällöin adjungaatin ominaisuuden 4.30 nojalla pätee
$$\langle \mathbf{v}, L^*L(\mathbf{w})\rangle = \langle L(\mathbf{v}), L(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle.$$
Tästä seuraa, että kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee
$$\langle \mathbf{v}, (L^*L(\mathbf{w}) - \mathbf{w})\rangle = 0_K.$$
Sijoittamalla tähän erityisesti $\mathbf{v} = L^*L(\mathbf{w}) - \mathbf{w}$, saadaan
$$|\mathbf{v}|^2 = \langle \mathbf{v}, \mathbf{v}\rangle = 0_K,$$
jolloin $\mathbf{v} = L^*L(\mathbf{w}) - \mathbf{w} = \mathbf{0}_V$. Tämä pätee jokaisella $\mathbf{w} \in V$. Näin ollen $L^*L = \mathrm{id}_V$.

Kääntäen oletetaan, että $L^*L = \mathrm{id}_V$ ja osoitetaan, että L on unitaarinen. Olkoot $\mathbf{v}, \mathbf{w} \in V$. Tällöin
$$\langle L(\mathbf{v}), L(\mathbf{w})\rangle = \langle \mathbf{v}, L^*L(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle.$$
Näin ollen L on unitaarinen. □

Olkoon $E = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ sisätuloavaruuden V *ortonormaali* kanta. Olkoon $L\colon V \to V$ operaattori ja olkoon $A = [L]_E$ kuvauksen L matriisi tämän kannan suhteen. Tällöin Lemman 4.32 nojalla, matriisin A adjungaatti A^* on operaattorin L^* matriisi $[L^*]_E$ kannan E suhteen. Tiedetään, että vastaavuus "operaattori$\mapsto$matriisi tietyssä kannassa" on algebraisomorfismi eli erityisesti yhteensopiva matriisien kertolaskun ja kuvausten yhdistämisoperaation kanssa. Tämän ja edellisen lemman nojalla päätellään helposti, että operaattori L on unitaarinen jos ja vain jos sen matriisille A pätevät yhtälöt
$$A^*A = A^*A = I_n.$$
Tämä havainto motivoi seuraavan määritelmän.

Määritelmä 4.38. *Neliömatriisia $U \in M(n \times n; K)$ sanotaan unitaariseksi jos*
$$UU^* = I_n = U^*U,$$
toisin sanoen jos U on kääntyvä ja sen käänteismatriisi on sen adjungaattimatriisi U^*.

Edellisen tarkastelun nojalla saadaan heti seuraava tulos.

Lemma 4.39. *Olkoon $L\colon V \to V$ äärellisulotteisen sisätuloavaruuden V operaattori ja olkoon E jokin avaruuden V ortonormaali kanta. Tällöin L on unitaarinen jos ja vain jos matriisi $[L]_E$ on unitaarinen.*

Kun kerroinkunta K on reaalilukujen kunta $\mathbb{R}$ unitaarisia kuvauksia/matriiseja sanotaan yleensä *ortogonaalisiksi*, kun taas termi "unitaarinen" usein varataan ainoastaan kompleksista tapausta varten. Tämä on meidän kannalta hieman ikävä, mutta historiallisesti vakiintunut käytäntö.

Tässä materiaalissa "unitaarinen kuvaus/matriisi" tarkoittaa aina yleistä tapausta, jossa K voi olla $\mathbb{R}$ tai $\mathbb{C}$, ellei muuta sanota. Jos taas puhutaan "ortogonaalisista kuvauksista/matriisesista", sen ymmärretään tarkoittavan "unitaarinen JA reaalikertoiminen".

Koska reaalikertoimisen matriisin konjugaatti on matriisi itse, reaaliarvoinen neliömatriisi O on ortogonaalinen jos ja vain $OO^T = I_n$.

Ortogonaaliset ja unitaariset ryhmät

Olkoon $A = (a_{ij}) \in M(n \times n; K)$ neliömatriisi. Tällöin sen determinaantti on määritelty ja voidaan ilmaista kaavalla

$$\det(A) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}.$$

Sijoittamalla tähän matriisin A kertoimien paikalle konjugaattimatriisin $\overline{A}$ kertoimia sekä käyttämällä hyväksi sitä, että konjugaattikuvaus on kunnan K kuntahomomorfismi, nähdään helposti, että

$$\det \overline{A} = \overline{\det A}.$$

Koska neliömatriisin A transpoosilla A^T on sama determinantti kuin matriisilla A (Lemma 2.140), saadaan tästä, että jokaiselle neliömatriisille A pätee

$$\det A^* = \det(\overline{A})^T = \det(\overline{A}) = \overline{\det A}.$$

Sovelletaan tätä tulosta unitaarisen matriisin tapaukseen. Olkoon U unitaarinen. Tällöin $UU^* = I_n$, joten, ottamalla tämän yhtälön kummaltakin puolelta determinantti, sekä käytämällä hyväksi sitä, että $|z|^2 = z\overline{z}$ kaikilla $z \in \mathbb{C}$, saadaan

$$|\det U|^2 = \det U \overline{\det U} = \det U \det U^* = \det UU^* = \det I_n = 1_K.$$

Tästä seuraa, että $|\det U| = 1$. Toisin sanoen unitaarisen matriisin U determinantti on sellainen kompleksiluku, jonka normi on yksi, eli yksikköympyrän S^1 alkio. Erityisesti tapauksessa $K = \mathbb{R}$ saadaan, että ortogonaalisen matriisin O determinantti on ± 1.

Unitaaristen ja ortogonaalisten matriisien avulla voidaan määritellä tärkeitä esimerkkejä klassisista kompakteista matriisiryhmistä, niin sanotuista unitaarisista ja ortogonaalisista ryhmistä. Näillä ryhmillä on erittäin tärkeä rooli transformaatio- ja Lie-ryhmien teoriassa, differentiaaligeometriassa, esitysteoriassa ja kvanttifysiikassa.

Olkoon $n \in \mathbb{N}$ ja $K = \mathbb{R}, \mathbb{C}$. Määritellään

$$U(n) = \{U \in M(n \times n; \mathbb{C}) \mid UU^* = I_n\},$$
$$O(n) = \{O \in M(n \times n; \mathbb{R}) \mid OO^T = I_n\},$$
$$SL(n; K) = \{A \in M(n \times n; K) \mid \det A = 1\},$$
$$SU(n) = U(n) \cap SL(n; \mathbb{C}),$$
$$SO(n) = O(n) \cap SL(n; \mathbb{R}).$$

Kaikki nämä joukot ovat *ryhmiä* matriisin kertolaskun suhteen. Ryhmää $U(n)$ sanotaan *unitaariseksi* ryhmäksi, ryhmää $O(n)$ sanotaan *ortogonaaliseksi* ryhmäksi. Erikoinen lineaarinen ryhmä (engl. special linear ryhmä) $SL(n; K)$ ei ole kompakti ryhmä, eikä liity mitenkään unitaarisiin matriiseihin, se on mainittu tässä yhteydessä yleissivistyksen takia ja koska sitä tarvitaan *erikoisen unitaarisen ryhmän* $SU(n)$ ja *erikoisen ortogonaalisen ryhmän* $SO(n)$ määritelmässä. Itse asiassa erikoinen unitaarinen ryhmä $SL(n; K)$ ei ole mitään muuta kun determinanttikuvauksen $\det \colon GL(n; K) \to K \setminus \{0\}$ ydin. Huomaa, että tämä kuvaus on ryhmien välinen homomorfismi, sillä

$$\det(AB) = \det(A)\det(B).$$

Tässä siis $GL(n; K)$ on varustettu matriisien kertolaskulla ja $K \setminus \{0\}$ on varustettu skaalarikunnan kertolaskulla.

Helposti nähdään, että homomorfismi $\det \colon GL(n; K) \to K \setminus \{0\}$ on surjektio, joten ryhmäteorian isomorfialauseen (Lause 1.99) nojalla pätee

$$GL(n; K)/SL(n; K) \cong K \setminus \{0\}.$$

Kuten yllä huomattiin, jokaisen ortogonaalisen matriisin $O \in O(n)$ determinantille pätee $\det O \in \{1, -1\} = \mathbb{Z}_2$ (tässä identifioidaan ryhmä $\mathbb{Z}_2$ ja kertolaskulla varustettu ryhmä $\{1, -1\}$, sillä ne ovat isomorfisia). Näin ollen $\det \colon O(n) \to \mathbb{Z}_2$ on ryhmähomomorfismi ja helposti nähdään, että näin määriteltynä se on surjektio. Koska tämän kuvauksen ydin on $SO(n)$, isomorfialauseen nojalla pätee

$$O(n)/SO(n) \cong \mathbb{Z}_2.$$

Näin ollen $O(n)$ on "kaksi kertaa suurempi" kuin $SO(n)$. Ryhmällä $O(n)$ on kaksi sivuluokkaa aliryhmän $SO(n)$ suhteen, toinen on aliryhmä $SO(n)$ itse ja toinen on osajoukko

$$SO(n)_- = \{O \in O(n) \mid \det O = -1\}.$$

Seuraavaksi tarkastellaan samalla tavalla determinanttikuvausta unitaarisessa ryhmässä $U(n)$. Yllä on osoitettu, että kaikilla $U \in U(n)$ pätee $\det U \in S^1$, missä

$$S^1 = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Voidaan näytää, että kuvaus $\det \colon U(n) \to S^1$ on surjektiivinen ryhmähomomorfismi. Koska sen ydin on tässä tapauksessa aliryhmä $SU(n)$, tästä saadaan, että pätee

$$U(n)/SU(n) \cong S^1.$$

Seuraavaksi selvitetään ortogonaalisen ryhmän $O(2)$ sekä sen aliryhmän $SO(2)$ alkioiden rakennetta. Ennen kuin mennään varsinaisiin laskuihin, käydään läpi seuraava kätevä tekninen aputulos.

Lemma 4.40. *Olkoon $U \in M(n \times n; K)$ neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) U on unitaarinen.

(2) U^T on unitaarinen

(3) Matriisin U sarakkeiden muodostama jono $(c_1(U), \ldots, c_n(U))$ on ortonormaali avaruuden K^n standardin pistetulon suhteen.

(4) Matriisin U rivien muodostama jono $(r_1(U), \ldots, r_n(U))$ on ortonormaali avaruuden K^n standardin pistetulon suhteen.

Todistus. Olkoon $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ avaruuden K^n standardikanta.

Jos U ajatellaan lineaarikuvauksena $L_U \colon K^n \to K^n$, niin sen sarakkeet $c_1(U), \ldots, c_n(U)$ ovat standardikannan alkioiden kuvia tämän kuvauksen suhteen, täsmällisemmin ilmaistuna kaikilla $i = 1, \ldots, n$ pätee

$$L_U(\mathbf{e}_i) = c_i(U).$$

Koska standardikanta E on ortonormaali avaruuden K^n pistetulon suhteen, kuvaus L_U on unitaarinen (pistetulon mielessä) jos ja vain jos sen matriisi $[L_U]_E$ ortonormaalin kannan E suhteen on unitaarinen (Lemma 4.39). Koska $[L_U]_E = U$, nähdään siis, että U on unitaarinen matriisi jos ja vain jos kuvaus L_U on unitaarinen.

Toisaalta Proposition 4.35 nojalla L_U on unitaarinen jos ja vain jos se kuvaa ortonormaalin kannan E ortonormaaliksi jonoksi, toisin sanoen jos ja vain jos jono $(c_1(U), \ldots, c_n(U))$ on ortonormaali. Yhdistämällä tuloksia nähdään, että U on unitaarinen jos ja vain jos jono $(c_1(U), \ldots, c_n(U))$ on ortonormaali pistetulon suhteen. Toisin sanoen väitteet (1) ja (3) ovat yhtäpitäviä.

Osoitetaan seuraavaksi, että väitteet (1) ja (2) ovat yhtäpitäviä. Koska $(U^T)^T = U$, symmetrian vuoksi riittää osoittaa, että ehto (1) implikoi ehdon (2). Oletetaan siis, että U on unitaarinen eli että $UU^* = I_n = U^*U$. Ottamalla transpoosit tämän yhtälöketjun kaikilta puolilta saadaan yhtälöketju

$$(U^*)^T U^T = I_n^T = U^T (U^*)^T.$$

Mutta $U^* = \overline{U}^T$, joten $(U^*)^T = \overline{U} = \overline{U^T}^T = (U^T)^*$. Lisäksi selvästi $I_n^T = I_n$. Näin ollen

$$(U^T)^* U^T = I_n = U^T (U^T)^*,$$

toisin sanoen U^T on unitaarinen eli (2) pätee. Symmetrian vuoksi voidaan päätellä, että ehdot (1) ja (2) ovat yhtäpitäviä.

Soveltamalla tämän jälkeen yllä jo todistettua kohtien (1) ja (3) yhtäpitävyyttä matriisiin U^T ja huomaamalla, että tämän matriisin sarakkeet ovat alkuperäisen matriisin U rivejä, nähdään, että matriisi U^T on unitaarinen jos ja vain jos $(r_1(U), \ldots, r_n(U))$ on ortonormaali avaruuden K^n standardin pistetulon suhteen. Toisin sanoen (2) ja (4) ovat ekvivalentteja. $\square$

Esimerkki 4.41. *Olkoon* $U = [z]$ (1×1)*-matriisi,* $z \in \mathbb{C}$*. Tällöin* $U^T = U$*, joten* $U^* = [\overline{z}]$ *ja* U *on unitaarinen jos ja vain jos* $[z\overline{z}] = UU^* = I_1 = [1]$*. Näin ollen* U *on unitaarinen jos ja vain jos* $z \in S^1$*. Erityisesti unitaarinen ryhmä* $U(1)$ *on isomorfinen ryhmän* S^1 *kanssa (missä jälkimmäinen on varustettu kompleksilukujen kertolaskulla). Kun rajoitutaan tarkastelu reaaliarvoisiin eli ortogonaalisiin matriiseihin, nähdään samalla tavalla, että* $O(1) = \{1, -1\} \cong \mathbb{Z}_2$*.*

Tutkitaan seuraavaksi minkälaisia ortogonaaliset (2×2)*-matriisit eli ryhmän* $O(2)$ *alkiot ovat.*

Olkoon

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

ortogonaalinen (2×2)*-matriisi,* $a, b, c, d \in \mathbb{R}$*. Matriisin* A *sarakkeet ovat* $c_1(A) = (a, b)$ *ja* $c_2(A) = (c, d)$*. Edellisen lemman nojalla nämä sarakkeet muodostavat ortonormaalin jonon* $\mathbb{R}^2$*:n pistetulon suhteen. Erityisesti täytyy päteä*

$$a^2 + b^2 = (a, b) \cdot (a, b) = 1.$$

Tarkastelemalla samalla tavalla matriisin A *rivejä saadaan edellisen lemman nojalla, että* $a^2 + c^2 = 1$*. Näin ollen*

$$b^2 = 1 - a^2 = c^2, \text{ eli}$$

$$b = \pm c.$$

Samalla tavalla nähdään, että $d = \pm a$*.*

Koska A *on ortogonaalinen, pätee* $\det A = \pm 1$*. Tutkitaan ensin tapausta* $\det A = 1$ *ja osoitetaan, että tällöin* $a = d$ *ja* $b = -c$*. Edelisen nojalla* $a = \pm d$*. Jos pätee* $a = -d$*, saadaan suoraan laskemalla, että*

$$1 = \det A = ad - bc = -a^2 \pm b^2.$$

Yhtälö $1 = -a^2 - b^2$*, joka saadaan tapauksessa* $b = c$ *on selvästi mahdottomuus (ollaan nyt reaalialueessa), joten tästä seuraa, että täytyy olla* $b = -c$*, jolloin* $1 = b^2 - a^2$*. Koska toisaalta pätee* $a^2 + b^2 = 1$*, vähentämällä yhtälö* $b^2 - a^2 = 1$ *yhtälöstä* $a^2 + b^2 = 1$ *nähdään helposti, että* $a = 0$*. Tällöin* $d = \pm a = 0$*, jolloin erityisesti yhtä hyvin päte* $a = d$*. On osoitettu, että tapauksessa* $\det A = 1$ *välttämättä pätee* $a = d$ *ja* $b = -c$*.*

On näytetty, että jokainen ortogonaalinen (2×2)*-matriisi* $A \in SO(2)$ *on muotoa*

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

missä $a, b \in \mathbb{R}$*,* $a^2 + b^2 = 1$*. Kääntäen nähdään helposti, että tällainen matriisi toteuttaa edellisen lemman ehdon (3) ja sen determinantti on tasan* 1*, joten se on erikoisen ortogonaalisen ryhmän* $SO(2)$ *alkio. Näin ollen*

$$SO(2) = \{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R}, (a, b) \in S^1\}.$$

Itse asiassa kuvaus $\phi\colon SO(2) \to S^1$,

$$\phi(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}) = (a, b),$$

on ryhmien välinen isomorfismi (tarkistus HT). Näin ollen

$$SO(2) \cong S^1 \cong U(1).$$

Ryhmän $SO(2)$ *alkiot kirjoitetaan usein myös niin sanotussa geometrisessa muodossa*

$$A_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix},$$

missä $\alpha \in \mathbb{R}$. *Tähän muotoon päästään kun huomataan, että jokainen ympyrän* S^1 *piste* (a, b) *voidaan kirjoittaa muodossa* $(\cos\alpha, \sin\alpha)$, *missä* α *on itse asiassa tämän pisteen "vaihekulma", eli kulma jonka se (paikkavektorina) muodostaa* x*-akselin kanssa. Tämä vaihekulma ei ole yksikäsitteinen vaan on tunnetusti määritelty luvun* 2π *monikerran verran.*

Matriisin esityksellä muodossa A_α *on luonnollinen geometrinen tulkinta, sillä matriisia* A_α *vastaava lineaarinen kuvaus* $L_{A_\alpha}\colon \mathbb{R}^2 \to \mathbb{R}^2$ *on tason kierto origon ympäri kulman* α *verran (vastapäivään). Näin ollen voidaan sanoa, että erikoinen ortogonaalinen ryhmä* $S0(2)$ *koostuu täsmälleen kaikista tason* $\mathbb{R}^2$ *kierroista.*

Tutkitaan vielä matriisin A_α *ominaisarvoja. Tämän matriisin karakteristinen polynomi on*

$$(X - \cos\alpha)^2 + \sin^2\alpha,$$

mistä nähdään, että jos $b = \sin\alpha \neq 0$, *ominaisarvoja ei ole. Jos taas* $b = 0$ *eli* $\alpha = 0$ *tai* π, *matriisi* A_α *on matriisi* $\pm I_2$ *ja jokainen nollasta eroava vektori on sen ominaisvektori. Geometrisesti tämä on selvä - kierto muuttaa jokaisen nollasta eroavan vektorin suunta, paitsi jos kierto on triviaali nollakierto tai kierto puoli täyskierrosta (*π *radiaania eli 180 astetta), jolloin vektori kuvautuu vasta-vektorikseen.*

Tutkimatta on vielä tapaus $\det A = -1$. *Mutta jos*

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

on ortogonaalinen matriisi, jonka determinantti on -1, *niin matriisi*

$$A' = \begin{bmatrix} a & -c \\ b & -d \end{bmatrix},$$

on ortogonaalinen matriisi, jonka determinantti on 1 *eli ryhmän* $SO(2)$ *alkio. Koska tällaisten muoto on jo selvitetty yllä, tästä saadaan helposti, että*

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \tag{4.42}$$

missä $a^2 + b^2 = 1$, $a, b \in \mathbb{R}$. Kääntäen minkä tahansa muotoa 4.42 olevan matriisin helposti nähdään olevan ortogonaalinen matriisi, jonka determinantti on (-1). Muotoa 4.42 olevan matriisin karakteristinen polynomi on polynomi

$$(X - a)(X + a) - b^2 = X^2 - (a^2 + b^2) = X^2 - 1,$$

jonka reaaliset juuret ovat 1 ja -1. Näin ollen ortogonaalinen (2×2)-matriisi, jonka determinantti on miinus yksi, on aina diagonalisoituva. Lemman 3.12 nojalla nähdään, että A on similaarinen matriisin

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

kanssa. Tämä matriisi B on kuvauksena $L_B \colon \mathbb{R}^2 \to \mathbb{R}^2$ peilaus x-akselin suhteen, eli itse asiassa konjugaattikuvaus, jos tulkitaan $\mathbb{C} = \mathbb{R}^2$.

Jos ortogonaalinen (2×2)-matriisi A, jonka determinantti on (-1) ajatellaan kuvauksena $L_A \colon \mathbb{R}^2 \to \mathbb{R}^2$, tällä kuvauksella on siis edellisen nojalla matriisiesitys muotoa

$$[L_A]_E = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

missä $E = (\mathbf{v}_1, \mathbf{v}_2)$ on jokin matriisin A ominaisavektoreista koostuva avaruuden $\mathbb{R}^2$ kanta. Näytetään, että tällainen kanta on aina ortogonaalinen ja se voidaan valita jopa ortonormaaliksi. Nimittäin, koska E on ominaisvektoreista koostuva kanta, pätee

$$L_A(\mathbf{v}_1) = \mathbf{v}_1,$$

$$L_A(\mathbf{v}_2) = -\mathbf{v}_2.$$

Toisaalta, koska A on ortogonaalinen, myös kuvaus L_A on ortogonaalinen. Tästä seuraa, että

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = L_A(\mathbf{v}_1) \cdot L_A(\mathbf{v}_2) = \mathbf{v}_1 \cdot (-\mathbf{v}_2) = -\mathbf{v}_1 \cdot \mathbf{v}_2.$$

Toisin sanoen $x = \mathbf{v}_1 \cdot \mathbf{v}_2$ on sellainen reaaliluku, jolle pätee $x = -x$. Tämä on mahdollista ainoastaan jos $x = 0$. Näin ollen $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, joten kanta E ortogonaalinen. Normeramalla sen vektoreita yksikköpituisiksi saadaan ortonormaali kanta E, jonka suhteen kuvauksen L_A matriisi on matriisi

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

Näin ollen, jokainen tason $\mathbb{R}^2$ ortogonaalinen kuvaus, jonka determinantti on (-1), on peilaus jossakin tason ortogonaalisen koordinaatiston suhteen. Kääntäen mikä tahansa tällainen kuvaus on ortogonaalinen ja sen determinantti on -1.

Jokainen tason $\mathbb{R}^2$ ortogonaalinen kuvaus (pistetuln suhteen) on siis joko kierto tai peilaus.

Edellisen esimerkin tulos voidaan yleistää korkeampiin ulottuvuuksiin. Itse asiassa osoittautuu, että mikä tahansa n-ulotteisen $\mathbb{R}$-sisätuloavaruuden ortogonaalinen operaattori on tietyssä m,ielessä ”suora summa” kierroista ja peilauksista. Tämä väite formuloidaan täsmällisesti ja todistetaan seuraavassa aliluvussa yleisimmän *normaalien operaattorien* teorian avulla.

4.3. Normaalit operaattorit ja diagonalisointi

Tässä luvussa palataan lineaariseen kuvaukseen diagonaalisointiongelmaan – tällä kertaa uuden työkalun eli sisätulon varustettuna.

Kuten lukija on jo varmasti huomannut tässä vaiheessa, sisätuloavaruuksien kohdalla ei yleisesti ottaen olla kiinnostuneita avaruuden mielivaltaisista kannoista, vaan ainoastaan *ortonormaaleista kannoista*, koska ne ovat "yhteensopivia" sisätulon kanssa. Tästä syystä diagonaalisaation ongelmakin tutkitaan tässä yhteydessä ainoastaan ortonormaalin kannan suhteen.

Olkoon V äärellisulotteinen sisätuloavaruus ja olkoon $L\colon V \to V$ operaattori. Operaattorin L sanotaan olevan *diagonalisoituva ortonormaalissa kannassa*, jos on olemassa avaruuden V *ortonormaali* kanta E, joka koostuu operaattorin L ominaisavektoreista. Tällöin voidaan myös sanoa, että operaattori L on diagonalisoituva ortonormaalissa kannassa E.

Vastaavasti K-kertoimisen $(n \times n)$-neliömatriisin A sanotaan olevan diagonalisoituva ortonormaalissa kannassa, jos sitä vastaava lineaarinen kuvaus $L_A\colon K^n \to K^n$ on diagonalisoituva jossakin avaruuden K^n ortonormaalissa kannassa (missä K^n varustetaan standardilla pistetulollaan).

Lemma 4.43. *a) Olkoon $L\colon V \to V$ operaattori äärellisulotteisessa sisätuloavaruudessa V. Olkoon E jokin avaruuden V kiinnitetty ortonormaali kanta. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(i) L on diagonalisoituva jossakin (mahdollisesti toisessa) avaruuden V ortonormaalissa kannassa E'.

(ii) Matriisi $[L]_E$ on diagonalisoituva (mahdollisesti toisessa) avaruuden V ortonormaalissa kannassa E'.

(iii) On olemassa unitaarinen $(n \times n)$-matriisi U siten, että matriisi $U[L]_E U^*$ on diagonaalimatriisi.

b) Olkoon $A \in M(n \times n; K)$ mielivaltainen neliömatriisi. Tällöin A on diagonalisoituva ortonormaalissa kannassa jos ja vain jos on olemassa unitaarinen $(n \times n)$-matriisi U siten, että $UAU^{-1} = UAU^*$ on diagonaalimatriisi.

Todistus. Tämä on oleellisesti sisätuloversio Lemmasta 3.12 ja todistus on analoginen. Huomataan, että mielivaltainen kääntyvä matriisi J Lemman 3.12 muotoilussa vaihtuu tässä yhteydessä unitaariseeksi matriisiksi. Tämä johtuu siitä, että kannanvaihtomatriisi $[E' \mid E]$ on unitaarinen matriisi kun E ja E' ovat ortonormaaleja kantoja. Yksityiskohtien läpikäynti jätetään harjoitustehtäväksi. $\square$

Olkoon operaattori $L\colon V \to V$ diagonalisoituva ortonormaalissa kannassa E. Tällöin matriisi $D = [L]_E$ on diagonaalimatriisi eli muotoa

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

Koska kannan E oletetaan olevan ortonormaali, operaattorin L adjungaatin L^* matriisi $[L^*]_E$ saman kannan E suhteen on matriisi D^* (Lemma 4.32) eli toisin sanoen diagonaalimatriisi

$$D^* = \begin{bmatrix} \overline{k_1} & 0 & \dots & 0 \\ 0 & \overline{k_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \overline{k_n} \end{bmatrix}.$$

Kaikki samankokoiset diagonaalimatriisit selvästi kommutoivat keskenään, joten erityisesti pätee

$$[LL^*]_E = [L]_E[L^*]_E = DD^* = D^*D = [L^*]_E[L]_E = [L^*L]_E.$$

Tästä seuraa, että myös kuvaukset L ja L^* kommutoivat keskenään, toisin sanoen pätee $LL^* = L^*L$. Tämä havainto motivoi seuraavan määritelmän.

Määritelmä 4.44. *Olkoon V sisätuloavaruus. Lineaarista operaattoria $L\colon V \to V$ sanotaan normaaliksi, jos pätee $LL^* = L^*L$.*

Neliömatriisia $A \in M(n \times n; K)$ sanotaan normaaliksi jos pätee $AA^* = A^*A$.

Kun E on äärellisulotteisen sisätuloavaruuden V ortonormaali kanta, operaattori $L\colon V \to V$ on normaali jos ja vain jos matriisi $[L]_E$ on normaali. Tämä seuraa helposti määritelmistä.

Edellä on osoitettu, että jokainen ortonormaalissa kannassa diagonalisoituva operaattori on erityisesti normaali. Osoittautuu, että tapauksessa $K = \mathbb{C}$ myös käänteinen väite pätee. Tämä todistetaan Propositiossa 4.50 alla. Tapauksessa $K = \mathbb{R}$ vastaava tulos ei päde. Diagonalisoituvuutta ortonormaalissa kannassa $\mathbb{R}$-sisätuloavaruuksien maailmassa karakterisoidaan Seurauksessa 4.56.

Tutkitaan tarkemmin normaalien operaattorien ominaisuuksia.

Lemma 4.45. *Olkoon $L\colon V \to V$, missä V on äärellisulotteinen sisätuloavaruus. Tällöin L on normaali jos ja vain jos kaikilla $\mathbf{v} \in V$ pätee*

$$|L(\mathbf{v})| = |L^*(\mathbf{v})|.$$

Todistus. Osoitetaan, että normaali kuvaus toteuttaa yhtälön

$$|L(\mathbf{v})| = |L^*(\mathbf{v})|$$

kaikilla $\mathbf{v} \in V$. Käänteistä väitettä ei tarvita jatkossa, joten sen todistus jätetään harjoitustehtäväksi.

Olkoon L normaali. Jokaisella $\mathbf{v} \in V$ pätee tällöin oletuksen ja adjungaatin ominaisuuksien nojalla

$$|L(\mathbf{v})|^2 = \langle L(\mathbf{v}), L(\mathbf{v})\rangle = \langle \mathbf{v}, L^*L(\mathbf{v})\rangle =$$
$$\langle \mathbf{v}, LL^*(\mathbf{v})\rangle = \langle L^*(\mathbf{v}), L^*(\mathbf{v})\rangle = |L^*(\mathbf{v})|^2.$$

Toiseksi viimeisessä välivaiheessa on käytetty hyväksi sitä, että kuvauksen L^* adjungaatti on kuvaus L, $(L^*)^* = L$.

Väite on todistettu. □

Lemma 4.46. *Olkoon V äärellisulotteinen K-sisätuloavaruus ja olkoon $L\colon V \to V$ normaali operaattori.*

(1) Olkoon $k \in K$ operaattorin L ominaisarvo ja olkoon $\mathbf{v}$ ominaisarvoon k liittyvä ominaisvektori. Tällöin $L^*(\mathbf{v}) = \overline{k}\mathbf{v}$. Erityisesti $\overline{k}$ on operaattorin L^* ominaisarvo.

(2) Olkoot $k, k' \in K$ operaattorin L eri ominaisarvoja, $k \neq k'$. Olkoon $\mathbf{v}$ ominaisarvoon k liittyvä ominaisvektori ja olkoon $\mathbf{w}$ ominaisarvoon k' liittyvä ominaisvektori. Tällöin $\mathbf{v} \perp \mathbf{w}$.

Todistus. Väitteen (1) todistus jätetään harjoitustehtäväksi.

Olkoot $k, k' \in K$ operaattorin L eri ominaisarvoja, $k \neq k'$. Olkoon $\mathbf{v}$ ominaisarvoon k liittyvä ominaisvektori ja olkoon $\mathbf{w}$ ominaisarvoon k' liittyvä ominaisvektori. Tällöin $L(\mathbf{v}) = k\mathbf{v}$ ja kohdan (1) nojalla $L^*(\mathbf{w}) = \overline{k'}\mathbf{w}$. Suoralla laskulla saadaan, että

$$k\langle \mathbf{v}, \mathbf{w}\rangle = \langle k\mathbf{v}, \mathbf{w}\rangle = \langle L(\mathbf{v}), \mathbf{w}\rangle =$$

$$\langle \mathbf{v}, L^*(\mathbf{w})\rangle = \langle \mathbf{v}, \overline{k'}\mathbf{w}\rangle = \langle k'\mathbf{v}, \mathbf{w}\rangle = k'\langle \mathbf{v}, \mathbf{w}\rangle.$$

Koska $k \neq k'$, tästä seuraa, että $\langle \mathbf{v}, \mathbf{w}\rangle = 0_K$. □

Propositio 4.47. *Olkoon V äärellisulotteinen sisätuloavaruus. Olkoon $L\colon V \to V$ normaali operaattori. Olkoon $W \subset V$ aliavaruus, joka on invariantti kuvauksen L suhteen. Tällöin sen ortogonaalinen komplementti $W^\perp$ on myös invariantti kuvauksen L suhteen.*

Todistus. Sisätuloavaruuden V sisätulon rajoittuma joukkoon $W \times W$ määrittelee luonnollisella tavalla sisätulon struktuurin myös aliavaruuteen W. Samoin $W^\perp$ voidaan ajatella sisätuloavaruutena luonnollisella tavalla. Koska W ja $W^\perp$ ovat tällöin kumpikin äärellisulotteisia sisätuloavaruuksia, Seurauksen 4.20 nojalla avaruudessa W voidaan valita ortonormaali kanta $E_1 = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ ja vastaavasti avaruudessa $W^\perp$ voidaan valita ortonormaali kanta $E_2 = (\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$. Lisäksi Lemman 4.22 nojalla pätee $W \oplus W^\perp = V$, joten, kun nämä kannat yhdistetään, saadaan koko avaruuden V kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$. Lisäksi tämä kanta on ortonormaali, sillä kannat E_1 ja E_2 ovat ortonormaaleja, lisäksi jokainen kannan E_1 alkio on W:n alkiona kohtisuorassa jokaisen kannan E_2 alkiota vastaan (sillä ne ovat ortogonaalisen komplementin $W^\perp$ alkioita).

Esitetään L matriisina kannassa E. Koska W on invariantti L:n suhteen, matriisi $[L]_E$ voidaan kirjoittaa lohkomatriisina

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

missä A on rajoittuman $L|W\colon W \to W$ matriisi kannassa E_1. Aliavaruuden $W^\perp$ kannan E_2 vektorien $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ kuvien komponentit kannan E_1 suhteen muodostavat matriisin B. $W^\perp$ on invariantti L:n suhteen jos ja vain jos kaikki nämä komponentit ovat nollia. Näin ollen täytyy osoittaa, että yllä pätee $B = 0$.

Koska kanta E on ortonormaali, Lemman 4.32 avulla nähdään, että adjungaatin L^* matriisi kannan E on matriisi

$$[L^*]_E = [L]_E^* = \begin{bmatrix} A^* & 0 \\ B^* & C^* \end{bmatrix}.$$

Olkoon $A = (a_{ij})_{i,j=1,\ldots,k}$. Tällöin kaikilla $l = 1, \ldots, k$ pätee

$$L(\mathbf{e}_l) = \sum_{i=1}^{k} a_{il}\mathbf{e}_i.$$

Koska jono $E_1 = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ on ortonormaali, tästä seuraa Pythagoraan lauseen (Lemma 4.13) nojalla, että

$$|L(\mathbf{e}_l)|^2 = \sum_{i=1}^{k} |a_{il}|^2.$$

Tästä puolestaan seuraa, että

$$\sum_{l=1}^{k} |L(\mathbf{e}_l)|^2 = \sum_{i,l=1}^{k} |a_{il}|^2 = a,$$

missä a on siis matriisin A alkioiden itseisarvojen neliöiden summa.

Lasketaan samalla tavalla vastaava summa matriisille A^*. Täsmällisemmin sanottuna ensin todetaan, että jokaisella $l = 1, \ldots, k$ pätee

$$L^*(\mathbf{e}_l) = \sum_{i=1}^{k} \overline{a_{li}} e_i + \sum_{i=k+1}^{n} \overline{b_{li}} e_i.$$

Koska jono E on ortonormaali, pätee Pythagoran lauseen nojalla

$$|L^*(\mathbf{e}_l)| = \sum_{i=1}^{n} |\overline{a_{li}}|^2 + \sum_{i=k+1}^{n} \left|\overline{b_{li}}\right|^2 =$$

$$= \sum_{i=1}^{n} |a_{li}|^2 + \sum_{i=k+1}^{n} |b_{li}|^2.$$

Tässä viimeisessä välivaiheessa käytetään hyväksi sitä, että $|\overline{z}| = |z|$ kaikilla $z \in \mathbb{C}$.

Näin ollen

$$\sum_{l=1}^{k} |L^*(\mathbf{e}_l)| = \sum_{i,l=1}^{k} |a_{li}|^2 + \sum_{i=k+1,l=1}^{n} |b_{li}|^2 = a + b,$$

missä a on kuten yllä matriisin A alkioiden itseisarvojen neliöiden summa ja b vastaavasti matriisin B alkioiden itseisarvojen neliöiden summa. Toisaalta Lemman 4.45 nojalla jokaisella $l = 1, \ldots, k$ pätee

$$|L^*(\mathbf{e}_l)| = |L(\mathbf{e}_l)|,$$

joten

$$a = \sum_{l=1}^{k} |L(\mathbf{e}_l)|^2 = \sum_{l=1}^{k} |L^*(\mathbf{e}_l)| = a + b.$$

Tästä seuraa, että $b = 0$. Toisaalta b on summa ei-negatiivisista reaaliluvuista $\left|\overline{b_{li}}\right|^2$, mistä seuraa, että $\left|\overline{b_{li}}\right|^2 = 0$ kaikilla $l = 1, \ldots, k$, $i = k+1, \ldots, n$. Koska jokainen matriisin B alkion itseisarvon neliö on tätä muotoa, matriisin B on oltava nolla-matriisi. Väite on todistettu. $\square$

Lemma 4.48. *Oletetaan, että $L\colon V \to V$ on lineaarinen operaattori äärellisulotteisessa sisätuloavaruudessa V. Olkoon W L-invariantti aliavaruus. Tällöin $W^\perp$ on invariantti operaattorin L^* suhteen.*

Todistus. Harjoitustehtävä. □

Seuraus 4.49. *Olkoon $L\colon V \to V$ normaali operaattori äärellisulotteisessa sisätuloavaruudessa V ja olkoon $W \subset V$ L-invariantti aliavaruus. Tällöin rajoittuma $L|W : W \to W$ on myös normaali.*

Todistus. Osoitetaan ensin, että W on myös L^*-invariantti. Koska L on normaali, Proposition 4.47 nojalla ortogonaalinen komplimentti $W^\perp$ on myös L-invariantti. Kun edellistä lemmaa sovelletaan kuvaukseen L ja invarianttiin aliavaruuteen $W^\perp$, saadaan, että $W = (W^\perp)^\perp$ on L^*-invariantti. Yhtälön $W = (W^\perp)^\perp$ osoittaminen jätetään harjoitustehtäväksi.

Erityisesti rajoittuma $L' = L^*|W : W \to W$ on hyvin määritelty aliavaruuden W lineaarinen operaattori. Koska kaikilla $\mathbf{v}, \mathbf{w} \in W$ adjungaatin määritelmän nojalla pätee

$$\langle L(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, L'(\mathbf{w})\rangle,$$

kuvaus $L'\colon W \to W$ on operaattorin $L|W$ adjungaattikuvaus (Lemma 4.29, yksikäsitteisyys). Toisin sanoen

$$(L|W)^* = L^*|W.$$

Koska yhtälö $LL^* = L^*L$ pätee koko avaruudessa V, se pätee erityisesti aliavaruudessa W. Näin ollen $L|W$ on normaali. □

Nyt voidaan karakterisoida operaattorit, jotka ovat diagonalisoituvia ortonormaalissa kannassa tapauksessa $K = \mathbb{C}$.

Propositio 4.50. *Olkoon V $\mathbb{C}$-kertoiminen sisätuloavaruus. Olkoon $L\colon V \to V$ operaattori. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) L on normaali.

(2) L on diagonalisoituva ortonormaalissa kannassa.

Todistus. Aikaisemmin on näytetty, että jokainen ortonormaalissa kannassa diagonalisoituva operaattori on välttämättä normaali (tämä pätee myös kun $K = \mathbb{R}$). Näin ollen riittää todistaa käänteinen suunta.

Oletetaan, että $L\colon V \to V$ on normaali operaattori, missä V on äärellisulotteinen $\mathbb{C}$-sisätuloavaruus. Todistetaan induktiolla luvun $n = \dim V$ suhteen, että L on diagonalisoituva ortonormaalissa kannassa.

Tapaus $n = 1$ on triviaali - jokainen lineaarinen kuvaus $L\colon K \to K$ on sekä normaali, että diagonalisoituva ortonormaalissa kannassa.

Oletetaan, että väite on tosi tapauksessa $\dim V = n - 1$ ja tarkastellaan tapausta $\dim V = n$. Koska kompleksilukujen kunta $\mathbb{C}$ on algebrallisesti suljettu, operaattorilla L on olemassa ominaisvektori $\mathbf{e}_1$ (tämä on todistuksen ainoa osa joka ei menisi läpi reaalilukujen kerroinkunnan $\mathbb{R}$ tapauksessa). Normeeramalla tämä vektori tarvittaessa, voidaan olettaa, että $|\mathbf{e}_1| = 1$. Tämän ominaisvektorin virittämä aliavaruus $U = \mathrm{Span}(\mathbf{e}_1)$ on tällöin L-invariantti, joten sen ortogonaalinen komplementti $W = U^{\perp}$ on myös L-invariantti (Lemma 4.47). Lisäksi edellisen seurauksen nojalla rajoittuma $L|W \colon W \to W$ on avaruuden W normaali operaattori. Koska $\dim W = \dim V - 1 = n-1$, induktioletuksen nojalla aliavaruudessa W on olemassa ortonormaali kanta $(\mathbf{e}_2, \ldots, \mathbf{e}_n)$, joka koostuu operaattorin L ominaisvektoreista. Tällöin jono $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ on avaruuden V ortonormaali kanta, joka koostuu operaattorin L ominaisvektoreista. Toisin sanoen L on diagonalisoituva ortonormaalissa kannassa. $\square$

Siirtymällä edellisessä tuloksessa kuvausten kielestä matriisien kielelle nähdään, että $\mathbb{C}$-kertoiminen neliömatriisi $A \in M(n \times n; \mathbb{C})$ voidaan kirjoittaa muodossa $A = UDU^*$ jollakin unitaarisella $U \in U(n)$ ja diagonaalimatriisila D jos ja vain jos A on normaali, eli jos ja vain jos pätee yhtälö $AA^* = A^*A$.

Unitaaristen kuvausten ja matriisien rakenne

Olkoon V äärellisulotteinen K-sisätuloavaruus (missä K on $\mathbb{R}$ tai $\mathbb{C}$ kuten yleensä). Olkoon $L \colon V \to V$ *unitaarinen* kuvaus (eli ortogonaalinen jos $K = \mathbb{R}$). Tällöin $L^* = L^{-1}$ ja koska $LL^{-1} = I_n = L^{-1}L$, unitaarinen operaattori L on erityisesti normaali. Samalla tavalla nähdään, että jokainen unitaarinen matriisi on normaali.

Karakterisoidaan edellisen proposition avulla kompleksikertoimiset unitaariset kuvaukset/matriisit.

Seuraus 4.51. *Olkoon $L \colon V \to V$ operaattori, missä V on äärellisulotteinen $\mathbb{C}$-sisätuloavaruus. Tällöin L on unitaarinen jos ja vain jos L on diagonalisoituva ortonormaalissa kannassa ja kaikki sen ominaisavot $z \in \mathbb{C}$ toteuttavat ehdon $|z| = 1$. Vastaavasti matriisi $A \in M(n \times n; \mathbb{C})$ on unitaarinen jos ja vain jos $A = UDU^{-1} = UDU^*$, missä $U \in U(n)$ ja*

$$D = \begin{bmatrix} z_1 & 0 & \ldots & 0 \\ 0 & z_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & z_n \end{bmatrix},$$

on diagonaalimatriisi, jonka kaikki diagonaalialkiot toteuttavat ehdon $|z_i| = 1$.

Todistus. Kaikki väitteet ovat samojen väitteiden kuvaus/matriisi tulkinnat, joten riittää osoittaa jokainen väite vain kerran kuvaukselle tai matriisille.

Kuten yllä todettiin, jokainen unitaarinen operaattori $L \colon V \to V$ on erityisesti normaali, joten, koska skaalarikunta oletetaan olevan $K = \mathbb{C}$, L on diagonalisoituva ortonormaalissa kannassa edellisen proposition nojalla. Osoitetaan vielä, että unitaarisen operaattorin L jokaisen ominaisarvon itseisarvo on 1. Tämä itse asiassa pätee myös kun $K = \mathbb{R}$. Nimittäin olkoon $z \in K, \mathbf{v} \in V, \mathbf{v} \neq 0$ siten, että $L(\mathbf{v}) = z\mathbf{v}$. Koska unitaarinen kuvaus L säilyttää normin (Propositio 4.35), pätee

$$1 \cdot |\mathbf{v}| = |\mathbf{v}| = |L(\mathbf{v})| = |z\mathbf{v}| = |z|\,|\mathbf{v}|.$$

Koska $|\mathbf{v}| \neq 0$, vektoriavaruuksien supsitussäännön nojalla saadaan $|z| = 1$, kuten pitikin. On osoitettu, että tapauksessa $K = \mathbb{C}$ unitaarinen operaattori on diagonalisoituva ja jokaisen sen ominaisarvon itseisarvo on yksi.

Kääntäen olkoon D diagonaalinen matriisi

$$\begin{bmatrix} z_1 & 0 & \dots & 0 \\ 0 & z_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & z_n \end{bmatrix},$$

jossa kaikki diagonaalialkiot toteuttavat ehdon $|z_i| = 1$. Osoitetaan, että tällöin matriisi $A = UDU^{-1} = UDU^*$, missä $U \in U(n)$ on unitaarinen, on myös unitaarinen.

Koska $|z_i| = 1$ kaikilla $i = 1, \dots, n$, saadaan erityisesti, että $\overline{z_i} = z_i^{-1}$. Näin ollen

$$D^* = \begin{bmatrix} \overline{z_1} & 0 & \dots & 0 \\ 0 & \overline{z_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \overline{z_n} \end{bmatrix} = \begin{bmatrix} z_1^{-1} & 0 & \dots & 0 \\ 0 & z_2^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & z_n^{-1} \end{bmatrix}.$$

Tästä nähdään helposti suoralla laskulla, että $DD^* = I_n$, joten D on unitaarinen. Tästä puolestaan seuraa, että

$$A^{-1} = (UDU^{-1})^{-1} = UD^{-1}U^{-1} = UD^*U^{-1} = UD^*U^* = (UDU^*)^* = A^*.$$

Toisin sanoen A on unitaarinen. □

Edellinen propositio sekä Propositio 4.50 eivät päde $\mathbb{R}$-sisätuloavaruuksille. Esimerkiksi ortogonaalinen (2×2)-matriisi

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

ei ole diagonalisoituva $\mathbb{R}$-kertoimisena matriisina kun $\alpha \in \mathbb{R}$, $\alpha \neq n\pi$ jollakin $n \in \mathbb{Z}$ (kts. esim. 4.41).

Tutkitaan tarkemmin normaaleja $\mathbb{R}$-kertoimisia (2×2)-matriiseja.
Olkoon

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

mielivaltainen matriisijoukon $M(2 \times 2; \mathbb{R})$ alkio. Tällöin

$$A^* = A^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Jos A on normaali, myös L_A on normaali, joten Lemman 4.45 nojalla erityisesti pätee

$$a^2 + b^2 = |L_A(\mathbf{e}_1)|^2 = |L_A^*(\mathbf{e}_1)|^2 = a^2 + c^2,$$

mistä seuraa, että $b^2 = c^2$ eli $b = \pm c$. Jos $b = c$, matriisi A on niin sanottu *symmetrinen matriisi* eli sille pätee $A^T = A$. Tällainen reaalikertoiminen matriisi on selvästi normaali (sillä jokainen matriisi kommutoi itseensä kanssa).

Toinen mahdollisuus on $b = -c$. Tällöin

$$A^T A = \begin{bmatrix} a^2 + b^2 & -ab + bd \\ -ab + bd & d^2 - b^2 \end{bmatrix} \text{ ja}$$

$$AA^T = \begin{bmatrix} a^2 + b^2 & ab - bd \\ ab - bd & d^2 - b^2 \end{bmatrix}.$$

Näistä kaavoista nähdään, että A on normaali (eli $AA^T = A^T A$) jos ja vain jos $b(a-d) = 0$ eli jos $b = 0$ tai $a = d$. Jos $b = 0$ matriisi on edellenkin symmetrinen. Muuten matriisi on muotoa

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Tämä on samannäköinen kuin esitys, joka johdettiin esimerkissä 4.41 ortogonaalisille (2×2)-matriiseille, joiden determinantti on 1, paitsi, että tässä ei vaadita, että pätisi $a^2 + b^2 = 1$. Jos tässä merkitään $r = a^2 + b^2$ ja oletetaan, että $r > 0$ (muuten matriisi on nollamatriisi, erityisesti symmetrinen), nähdään, että $A = rO$, missä $O \in SO(2)$ on ortogonaalinen matriisi, jonka determinantti on 1. Tällainen matriisi on normaali (koska se on vakiota vaille jopa ortogonaalinen). Näin ollen ollaan todistettu seuraava tulos.

Lemma 4.52. *Matriisi $A \in M(2 \times 2; \mathbb{R})$ on normaali jos ja vain jos se on symmetrinen matriisi tai muotoa $A = rO$, missä $r > 0$ ja $O \in SO(2)$.*

Tämän tuloksen avulla johdetaan hajotelmalause normaaleille operaattoreille $\mathbb{R}$-sisätuloavaruuksissa. Sitä varten tarvitaan ensin aputuloksia. Yllä tarkasteltu reaalinen (2×2)-kokoinen symmetrinen matriisi

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

on aina diagonalisoituva, mikä nähdään helposti laskemalla sen ominaisarvoja (harjotustehtävä). Tämä ei ole mikään sattuma - osoittautuu, että jokainen symmetrinen $\mathbb{R}$-kertoiminen matriisi on diagonalisoituva kunnan $\mathbb{R}$ suhteen. Tämä osoitetaan yleisemmin Propositiossa 4.55 alla.

Määritelmä 4.53. *Olkoon $L\colon V \to V$ operaattori K-sisätuloavaruudessa V. Operaattoria L sanotaan* itseadjungoiduksi *jos L on itsensä adjungaatti, toisin sanoen jos $L^* = L$. Samoin matriisia $A \in M(n \times n; K)$ sanotaan itseadjungoiduksi jos $A^* = A$.*

Adjungaatin määrittelevän ominaisuuden (Lemma 4.29) nojalla seuraa, että lineaarinen operaattori $L\colon V \to V$ on itseadjungoitu jos ja vain jos kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee

$$\langle L(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, L(\mathbf{w}) \rangle. \tag{4.54}$$

On selvä, että äärellisulotteisen sisätuloavaruuden operaattori on itseadjungoitu jos ja vain jos sen matriisi jonkun avaruuden ortonormaalin kannan suhteen on itseadjungoitu matriisi. Tapauksessa $K = \mathbb{R}$ matriisi A on itseadjungoitu jos ja vain jos se on symmetrinen, $A^T = A$. Tästä syystä tapauksessa $K = \mathbb{R}$ itseadjungoittuja operaattoreita sanotaan usein myös *symmetrisiksi.*

Määritelmän mukaan neliömatriisi $A = (a_{ij})$ on itseadjungoitu jos ja vain jos kaikilla $i, j = 1, \ldots, n$ pätee $a_{ij} = \overline{a_{ji}}$. Erityisesti tällaisen matriisin diagonaalialkiot ovat aina reaalisia, sillä niille pätee $a_{ii} = \overline{a_{ii}}$.

Propositio 4.55. *Olkoon $L\colon V \to V$ operaattori äärellisulotteisessa K-sisätuloavaruudessa, missä $K = \mathbb{R}$ tai $\mathbb{C}$. Tällöin L on itseadjungoitu jos ja vain jos se on diagonalisoituva ortonormaalissa kannassa ja kaikki sen ominaisarvot ovat reaalisia.*

Neliömatriisi $A \in M(n \times n; K)$ on itseadjungoitu jos ja vain jos $A = UDU^{-1}$, missä U on unitaarinen ja D on diagonaalimatriisi, jonka kaikki alkiot ovat reaalisia.

Todistus. Oletetaan, että operaattori L on diagonalisoituva jossakin ortonormaalissa kannassa $E = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ ja lisäksi L:n kaikki ominaisarvot ovat reaalilukuja (tämä lisäoletus on tietenkin turha tapauksessa $K = \mathbb{R}$). Tällöin siis kaikilla $i = 1, \ldots, n$ on olemassa reaaliluku $r_i \in \mathbb{R}$ siten, että

$$L(\mathbf{v}_i) = r_i \mathbf{v}_i.$$

Tästä seuraa, että matriisi $[L]_E$ on diagonaalimatriisi

$$D = \begin{bmatrix} r_1 & 0 & \ldots & 0 \\ 0 & r_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & r_n \end{bmatrix},$$

jonka kaikki alkiot ovat reaalilukuja. Tällaiselle matriisille selvästi pätee $D^* = D$. Koska kanta E on ortonormaali, Lemman 4.32 nojalla pätee

$$[L^*]_E = D^* = D = [L]_E.$$

Koska vastaavuus matriisien ja lineaaristen kuvausten välillä on bijektiivinen, tästä seuraa, että $L^* = L$. Toisin sanoen L on itseadjungoitu.

Käänteisen suunnan todistusta varten on tarkasteltava reaalinen ja kompleksinen tapaus erikseen. Oletetaan siis, että $L\colon V \to V$ on itseadjungoitu kuvaus ja osoitetaan, että se on diagonalisoituva ortonormaalissa kannassa ja lisäksi sen kaikki ominaisarvot ovat reaalilukuja.

Tarkastellaan ensin yksinkertaisempaa tapausta $K = \mathbb{C}$. Tällöin, koska itseadjungoitu operaattori L on selvästi normaali, Proposition 4.50 nojalla se on diagonalisoituva ortonormaalissa kannassa. Nyt riittää siis vain osoittaa, että L:n jokainen ominaisarvo on reaaliluku. Olkoon E sellainen ortonormaali kanta, jonka suhteen $D = [L]_E$ on diagonaalimatriisi,

$$D = \begin{bmatrix} r_1 & 0 & \ldots & 0 \\ 0 & r_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & r_n \end{bmatrix}.$$

Tällöin tämän matriisin lävistäjäalkiot $r_1, \ldots, r_n$ ovat täsmälleen kaikki operaattorin L ominaisarvot. Koska $L^* = L$, matriisille D pätee $D^* = D$. Mutta toisaalta

$$D^* = \begin{bmatrix} \overline{r}_1 & 0 & \ldots & 0 \\ 0 & \overline{r}_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \overline{r}_n \end{bmatrix},$$

joten $\overline{r}_i = r_i$ kaikilla $i = 1, \ldots, n$. Tunnetusti tämä on yhtäpitävää sen kanssa, että $r_i \in \mathbb{R}$ kaikilla $i = 1, \ldots, n$. Tämä on se mitä pitää todistaa.

Oletetaan seuraavaksi, että $K = \mathbb{R}$ ja operaattori $L\colon V \to V$ on symmetrinen. Koska tällöin jokainen L:n ominaisarvo on reaalinen määritelmän mukaan, riittää osoittaa, että L on diagonalisoituva ortonormaalissa kannassa.

Aloitetaan osoittamalla, että symmetrisellä operaattorilla $L\colon V \to V$, missä V on $\mathbb{R}$-sisätuloavaruus on ainakin yksi (reaalinen) ominaisarvo. Tämä on epätriviaalia, sillä tunnetusti $\mathbb{R}$-vektoriavaruuden operaattorilla ei tarvitse olla ominaisarvoja.

Siirtymällä matriisien kielelle riittää osoittaa, että jokaisella symmetrisellä neliömatriisilla $A \in M(n \times n; \mathbb{R})$ on ainakin yksi (reaalinen) ominaisarvo. Tosiin sanoen on löydettävä $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq \mathbf{0}$, ja $r \in \mathbb{R}$ siten, että $A\mathbf{v} = r\mathbf{v}$. Koska matriisi A on reaalikertoiminen, se voidaan yhtä hyvin tulkita erityisesti kompleksilukukertoimisena, toisin sanoen joukon $M(n \times n; \mathbb{C})$ alkiona. Tällaista matriisia taas vastaa lineaarinen kuvaus $L_A\colon \mathbb{C}^n \to \mathbb{C}^n$. Tämän kuvauksen matriisi avaruuden $\mathbb{C}^n$ standardin kannan E suhteen on matriisi A. Koska matriisi A on itseadjungoitu, myös kuvaus $L_A\colon \mathbb{C} \to \mathbb{C}$ on itseadjungoitu. Kuitenkin edellä on todistettu, että jokainen itseadjungoidu operaattori $\mathbb{C}$-sisätuloavaruudessa on diagonalisoituva ortonormaalissa kannassa ja jokainen sen ominaisarvo on lisäksi reaalinen. Matriisilla A on siis kompleksikertoimisena matriisina ajateltuna ainakin yksi reaalinen ominaisarvo r. Olkoon $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$ tätä ominaisarvoa vastaava ominaisvektori. Koska jokaisella $i = 1, \ldots, n$ on olemassa yksikäsitteiset $x_i, y_i \in \mathbb{R}$ siten, että $z_i = x_i + iy_i$, nähdään helposti, että vektori $\mathbf{z}$ voidaan kirjoittaa yksikäsitteisellä tavalla muodossa $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, missä $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Tässä siis samastetaan avaruuden $\mathbb{R}^n$ vektori $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ samannäköisen avaruuden $\mathbb{C}^n$ vektorin $(x_1, \ldots, x_n)$ kanssa.

Koska ominaisvektori $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ on nollavektorista eroava, ainakin toinen vektoreista $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ on nollavektorista eroava. Lisäksi

$$A\mathbf{x} + iA\mathbf{y} = L_A(\mathbf{x} + i\mathbf{y}) = L_A(\mathbf{z}) = r\mathbf{z} = r\mathbf{x} + ir\mathbf{y}.$$

Koska A on reaalikertoiminen, $A\mathbf{x}, A\mathbf{y} \in \mathbb{R}^n$. Koska avaruuden $\mathbb{C}^n$ vektorin esitys muodossa $\mathbf{v} + i\mathbf{u}$, missä $\mathbf{v}$, $\mathbf{u} \in \mathbb{R}^n$, on yksikäsitteinen, voidaan päätellä, että $A\mathbf{x} = r\mathbf{x}$ ja $A\mathbf{y} = r\mathbf{y}$. Koska ainakin toinen vektoreista $\mathbf{x}, \mathbf{y}$ on nollavektorista eroava, ainakin toinen niistä kelpaa matriisin A ominaisvektoriksi. Näin ollen symmetrisellä reaalisella matriisilla A on ainakin yksi ominaisarvo, joten myös jokaisella symmetrisellä operaattorilla $L\colon V \to V$, missä V on äärellisulotteinen $\mathbb{R}$-sisätuloavaruus, on reaalinen ominaisarvo.

Nyt voidaan osoittaa, että symmetrinen operaattori $L\colon V \to V$ on diagonalisoituva ortonormaalissa kannassa, kun V on äärellisulotteinen $\mathbb{R}$-sisätuloavaruus. Todistus etenee induktiolla luvun $n = \dim V$ suhteen. Alkuaskel $n = 1$ on selvä. Oletetaan, että väite pätee avaruuksille, joiden dimensio on $(n-1)$ ja oletetaan, että $\dim V = n$. Olkoon $L\colon V \to V$ symmetrinen operaattori. Edellä on osoitettu, että operaattorilla L on olemassa ominaisarvo r. Olkoon $\mathbf{v}_n \in V$ jokin tähän ominaisarvoon liittyvä ominaisvektori. Normeeramalla se tarvittaessa, voidaan olettaa, että $|\mathbf{v}_n| = 1$. Ominaisvektorin $\mathbf{v}$ virittämä aliavaruus $W = \mathrm{Span}(\mathbf{v}_n)$ on 1-ulotteinen ja L-invariantti. Lisäksi pätee $W \oplus W^\perp = V$, joten $\dim W^\perp = n-1$. Koska L on symmetrinen, se on erityisesti normaali, joten Lemmojen 4.47 ja 4.49 nojalla aliavaruus $U = W^\perp$ on L-invariantti. Osoitetaan, että rajoittuma $L|U : U \to U$ on avaruuden U symmetrinen operaattori. Koska L on symmetrinen, kaikilla

$\mathbf{v}, \mathbf{w} \in V$ pätee (kts. yhtälö 4.54)

$$\langle L(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, L(\mathbf{w}) \rangle.$$

Erityisesti tällöin kaikilla $\mathbf{v}, \mathbf{w} \in U$ pätee

$$\langle L(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, L(\mathbf{w}) \rangle.$$

Näin ollen kuvaus $L| \colon U \to U$ toteuttaa adjungaattinsa määrittelevän ominaisuuden (Lemma 4.29), joten se on itsensä adjungaatti. Toisin sanoen $L|U$ on sisätuloavaruuden U symmetrinen operaattori. Induktio-oletuksen nojalla avaruudella $U = W^{\perp}$ on olemassa ortonormaali kanta $(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$, joka koostuu operaattorin L ominaisvektoreista. Tällöin $E = (\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n)$ on avaruuden V ortonormaali kanta, joka koostuu ominaisvektoreista. Todistus on valmis. □

Tarkastelemalla edellisessä propositiossa tapausta $K = \mathbb{R}$, saadaan täydellinen karakterisaatio sellaisille äärellisulotteisen $\mathbb{R}$-sisätuloavaruuden V operaattoreille, jotka ovat diagonalisoituvia ortonormaalissa kannassa.

Seuraus 4.56. *Olkoon V äärellisulotteinen $\mathbb{R}$-sisätuloavaruus. Olkoon $L \colon V \to V$. Tällöin L on diagonalisoituva ortonormaalissa kannassa jos ja vain jos se on symmetrinen eli jos ja vain jos $L^* = L$.*

Vastaus siihen kysymyksen milloin äärellisulotteisen K-sisätuloavaruuden operaattori $L \colon V \to V$ on diagonalisoituva ortonormaalissa kannassa on siis osoittautunut riippuvan kerroinkunnasta K. Kun $K = \mathbb{C}$ tämä ehto on yhtäpitävä sen kanssa, että L on normaali eli kommutoi oman adjungaatin kanssa, $L^*L = LL^*$ (Propositio 4.50). Tapauksessa $K = \mathbb{R}$ taas tämä ehto on yhtäpitävä sen kanssa, että L on itsensä adjungaatti, $L^* = L$ (edellinen seuraus). Tulosten eron syy on ainakin osittain siinä, että $\mathbb{C}$ on algebrallisesti suljettu (jolloin jokaisella operaattorilla on automaattisesti ainakin yksi ominaisarvo), kun taas $\mathbb{R}$ ei ole.

Reaalikertoimisten normaalien operaattorien rakenne

Normaali kuvaus $L \colon V \to V$, missä V on äärellisulotteinen $\mathbb{C}$-sisätuloavaruus, on diagonalisoituva ortonormaalissa kannassa. Tämä on todistettu Seurauksessa 4.51. Tulos voidaan muotoilla uudestaan suoran summan kielellä seuraavasti. Olkoon $L \colon V \to V$ unitaarinen kuvaus äärellisulotteisessä $\mathbb{C}$-sisätuloavaruudessa V. Tällöin $V = W_1 \oplus W_2 \oplus \ldots \oplus W_n$, missä W_i on L-invariantti 1-ulotteinen aliavaruus jokaisella $i = 1, \ldots, n$. Lisäksi aliavaruudet W_i ovat "pareittain kohtisuorassa" toisiaan vastaan, eli $\mathbf{v} \perp \mathbf{w}$ kun $\mathbf{v} \in W_i$, $\mathbf{w} \in W_j$, $i \neq j$. Tähän tulkintaan päästään kun valitaan ortonormaali kanta $E = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, joka koostuu operaattorin ominaisvektoreista ja asetetaan $W_i = \mathrm{Span}(\mathbf{v}_i)$ kaikilla $i = 1, \ldots, n$.

Normaaleille operaattoreille $\mathbb{R}$-sisätuloavaruudessa V (eli ortogonaalisille kuvauksille) tämäntyyppinen tulos ei päde, sillä tällainen operaattori ei välttämättä ole edes diagonalisoituva. Osoitetaan kuitenkin, että V tällöin voidaan esittää suorana summana L-invarianteista aliavaruuksista, jotka ovat kohtisuorassa pareittain ja joiden dimensio on korkeintaan kaksi.

Aloitetaan aputuloksella, jonka mukaan jokaiselle äärellisulotteiselle $\mathbb{R}$-vektoriavaruuden V operaattorille löytyy korkeintaan kaksiulotteinen aliavaruus, joka on invariantti tämän operaattorin suhteen. Tämä tulos muotoillaan ja todistetaan vektoriavaruuksille, ei sisätuloavaruuksille, joten siinänsä se kuuluu pikemminkin yleiseen invarianttialiavaruuksien teoriaan, jota tarkasteltiin Luvussa 3. Tästä syystä tämä väite olisi voitu yhtä hyvin käydä läpi aikaisemminkin edellisessä luvussa. Toisaalta tämä on ensimmäinen kerta kun sitä tarvitaan tällä kurssilla.

Lemma 4.57. *Olkoon V epätriviaali äärellisulotteinen $\mathbb{R}$-vektoriavaruus ja olkoon L sen lineaarinen operaattori. Tällöin avaruudella V on olemassa L-invariantti aliavaruus W, jonka dimensio on 1 tai 2.*

Todistus. Koska jokainen äärellisulotteinen $\mathbb{R}$-vektoriavaruus on isomorfinen avaruuden $\mathbb{R}^n$ kanssa jollakin $n \in \mathbb{N}$, riittää osoittaa, että väite pätee kun $L = L_A \colon \mathbb{R}^n \to \mathbb{R}^n$, missä A on jokin $(n \times n)$-kokoinen reaalilukukertoiminen neliömatriisi.

Käytetään samaa temppua kuin Proposition 4.55 todistuksessa. Ajatellaan A kompleksilukukertoimisena matriisina, jolloin se määrittelee lineaarisen kuvauksen $L_A^{\mathbb{C}} \colon \mathbb{C}^n \to \mathbb{C}^n$. Jokainen avaruuden $\mathbb{C}^n$ alkio voidaan yksikäsitteisellä tavalla kirjoittaa muodossa $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, missä $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Helposti nähdään, että tällöin

$$L_A^{\mathbb{C}}(\mathbf{z}) = L_A(\mathbf{x}) + iL_A(\mathbf{y}) = A\mathbf{x} + iA\mathbf{y}.$$

Koska kunta $\mathbb{C}$ on algebrallisesti suljettu, operaattorilla $L_A^{\mathbb{C}} \colon \mathbb{C}^n \to \mathbb{C}^n$ on olemassa (kompleksinen) ominaisarvo $c = a + ib$. Olkoon $\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathbb{C}^n$ vastaava ominaisvektori. Tällöin ainakin toinen vektoreista $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ on nollavektorista eroava, joten näiden vektorien virittämä avaruuden $\mathbb{R}^n$ aliavaruus $W = \mathrm{Span}(\mathbf{x}, \mathbf{y})$ on 1- tai 2-ulotteinen. Osoitetaan, että W on invariantti kuvauksen L_A suhteen. Koska $\mathbf{z}$ on ominaisvektori, pätee

$$L_A(\mathbf{x}) + iL_A(\mathbf{y}) = L_A(\mathbf{z}) = c\mathbf{z} = (a + ib)(\mathbf{x} + i\mathbf{y}) = (a\mathbf{x} - b\mathbf{y}) + i(a\mathbf{y} + b\mathbf{x}).$$

Tästä seuraa, että

$$\begin{aligned} L_A(\mathbf{x}) &= a\mathbf{x} - b\mathbf{y} \in W, \text{ ja} \\ L_A(\mathbf{y}) &= a\mathbf{y} + b\mathbf{x} \in W. \end{aligned}$$

Koska aliavaruuden W jokainen alkio on virittäjien $\mathbf{x}$ ja $\mathbf{y}$ lineaarinen kombinaatio, tästä seuraa, että W on L_A-invariantti. □

Olkoot $W_1, \ldots, W_n$ sisätuloavaruuden V aliavaruuksia. Sanotaan, että aliavaruudet $W_1, \ldots, W_n$ ovat *pareittain kohtisuorassa toisiaan vastaan* jos $\mathbf{v} \perp \mathbf{w}$ aina kun $\mathbf{v} \in W_i$, $\mathbf{w} \in W_j$, $i \neq j$.

Propositio 4.58. *Olkoon V äärellisulotteinen $\mathbb{R}$-sisätuloavaruus ja olkoon $L \colon V \to V$ operaattori. Tällöin operaattori L on normaali jos ja vain jos on olemassa avaruuden V L-invariantit aliavaruudet $W, V_1, \ldots, V_m$ siten, että seuraavat ehdot ovat voimassa.*

- *Aliavaruudet $W, V_1, \ldots, V_m$ ovat pareittain kohtisuorassa toisiaan vastaan.*

- *Aliavaruudet* $W, V_1, \ldots, V_m$ *muodostavat suoran summan, jonka arvo on koko avaruus* V.

- *Operaattori* $L|W\colon W \to W$ *on diagonalisoituva ortonormaalissa kannassa.*

- $\dim V_i = 2$ *kaikilla* $i = 1, \ldots, m$,

- *Operaattori* $L|V_i\colon V_i \to V_i$ *on muotoa* rO, *missä* $r > 0$ *ja* O *on sellainen kaksiulotteisen aliavaruuden* V_i *ortogonaalinen operaattori, jolla ei ole ominaisarvoja ja jonka determinantti on* 1.

Todistus. Oletetaan, että L on normaali. Olkoon W kaikkien sen ominaisvektorien virittämä aliavaruus. Olkoot $r_1, \ldots, r_n$ kaikki operaattorin L (eri) ominaisarvot. Valitaan jokaisella $i = 1, \ldots, n$ ominaisarvoaliavaruudessa V_{r_i} jokin ortonormaali kanta E_i. Koska W on suora summa aliavaruuksista V_{r_i}, näiden kantojen yhdiste E on avaruuden W kanta. Koska L on normaali, Lemmasta 4.46 seuraa, että aliavaruudet V_{r_i} ovat pareittain kohtisuorassa toisiaan vastaan. Tästä seuraa, että kanta E on itse asiassa ortonormaali. Koska tämä kanta koostuu operaattorin $L|W$ ominaisvektoreista, operaattori $L|W\colon W \to W$ on diagonalisoituva ortonormaalissa kannassa.

Koska W on L-invariantti, Proposition 4.47 nojalla sen ortogonaalinen komplementti $W^\perp$ on myös L-invariantti. Seurauksen 4.49 nojalla rajoittuma $L|W^\perp\colon W^\perp \to W^\perp$ on avaruuden $W^\perp$ normaali operaattori. Lisäksi tällä rajoittumalla $L|W^\perp$ ei ole ominaisarvoja eikä ominaisvektoreita, sillä tämän operaattorin ominaisvektori olisi erityisesti operaattorin L ominaisvektori, joten kuuluisi aliavaruuteen W. Tämä on kuitenkin mahdotonta, sillä $W \cap W^\perp = \{0\}$.

Korvaamalla V avaruudella $W^\perp$ nähdään edellisen nojalla, että nyt riittää näyttää normaalille operaattorelle $L\colon V \to V$, jolla ei ole ominaisarvoja, että V on suora summa 2-ulotteisista L-invarianteista aliavaruuksista $W_1, \ldots, W_n$, jotka ovat kohtisuorassa toisia vastaan. Nimittäin tällöin Seurauksen 4.49 nojalla rajoittuma $L|W_i$ on kaksiulotteisen $\mathbb{R}$-sisätuloavaruuden normaali operaattori ja Lemman 4.52 mukaan tällainen operaattori on joko symmetrinen tai muotoa rO missä $r > 0$ ja $O \in SO(2)$ on ortogonaalinen matriisi, jolle $\det O = 1$. Tässä tapauksessa $L|W_i$ ei kuitenkaan voi olla symmetrinen, sillä tällöin se olisi myös diagonalisoituva, mikä on ristiriita, sillä oletamme, että L:llä ei ole ominaisarvoja. Näin ollen operaattorin $L|W_i$ on pakko olla muotoa rO joillakin $r > 0$ ja $O \in SO(2)$.

Oletetaan siis, että L on $\mathbb{R}$-sisätuloavaruuden V normaali operaattori, jolla ei ole ominaisarvoja. Edellisen lemman nojalla operaattorilla L on 2-ulotteinen invariantti aliavaruus U. Koska L on normaali, myös ortogonaalinen komplementti $U^\perp$ on L-invariantti ja $L|U^\perp$ on myös normaali operaattori, jolla ei ole ominaisarvoja. Jatkamalla induktiolla, saadaan osoitettua haluttu väite.

Oletetaan kääntäen, että kuvaus L toteuttaa lemman ehdot. Tällöin sen rajoittuma jokaiseen aliavaruuteen $W, V_1, \ldots, V_n$ on selvästi normaali. Koska näiden aliavaruuksien oletetaan olevan kohtisuorassa toisiaan vastaan, nähdään helposti suoralla laskulla, että L on normaali (mieti yksityiskohtia läpi). □

Seuraus 4.59. *Olkoon $L\colon V \to V$ operaattori, missä V on äärellisulotteinen $\mathbb{R}$-sisätuloavaruus. Tällöin L on ortogonaalinen jos ja vain jos avaruudella V on ortonormaali kanta E, jossa L voidaan esittää lohkomatriisina muotoa*

$$\begin{bmatrix} A_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & A_2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_l & 0 & 0 \\ 0 & 0 & \dots & 0 & -I_m & 0 \\ 0 & 0 & \dots & 0 & 0 & I_n \end{bmatrix},$$

missä $A_k \in SO(2 \times 2; \mathbb{R})$, $k = 1, \dots, l$. Lisäksi tällöin $\det L = (-1)^m$. Surkastuneet tapaukset joissa $l = 0$, $m = 0$ tai $n = 0$ ovat myös mahdollisia.

Jokainen ortogonaalinen operaattori on siis "ortogonaalinen" yhdistelmä peilauksista ja kaksiulotteisista kierroista.

Todistus. Seuraa edellisestä propositiosta, kunhan huomataan, että ortogonaalisen operaattorin ainoat mahdolliset ominaisarvot ovat ± 1 ja lisäksi, että sen rajoittuma invarianttiin aliavaruuteen on myös ortogonaalinen operaattori. □

Erityisesti kolmiulotteisessa avaruudessa $\mathbb{R}^3$ saatiin johdettua edellisen seurauksen erikoistapauksena merkittävä tulos, joka oli tuttu jo Eulerille - jokainen etäisyydet ja avaruuden orientaation säilyttävä kolmeulotteisen avaruuden lineaarinen operaattori $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ voidaan esittää jossakin avaruuden ortonormaalissa kannassa matriisina

$$\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

eli havainnollisesti ajattelen se on kaksiulotteinen kierto paikallaan pysyvän akselin ympäri. Tässä "orientaation säilyttäminen" tarkoittaa, että $\det L > 0$.

4.4. Hermiittiset muodot ja polaarinen hajotelma

Olkoon V K-sisätuloavaruus. Palautetaan mieleen, että hermiittiseksi muodoksi sanotaan seskilineaarista muotoa $H\colon V \times V \to K$, joka toteuttaa lisäksi ehdon

$$H(\mathbf{v}, \mathbf{w}) = \overline{H(\mathbf{w}, \mathbf{v})}$$

kaikilla $\mathbf{v}, \mathbf{w} \in V$. Tarkastelemalla tässä määritelmässä erikoistapausta $\mathbf{v} = \mathbf{w}$ nähdään erityisesti, että tällöin $H(\mathbf{v}, \mathbf{v})$ on reaaliluku kaikilla $\mathbf{v} \in V$ (sillä se on itsensä konjugaatti). Tapauksessa $K = \mathbb{R}$ hermiittinen muoto on sama asia kuin bilineaarinen ja symmetrinen muoto.

Sisätulo $\langle , \rangle$ K-vektoriavaruudessa V on määritelmän mukaan eräs hermiittinen muoto (joka toteuttaa myös lisäehtoja). Kuin yksi tällainen muoto vektoriavaruudessa V on kiinnitetty, puhutaan sisätuloavaruudesta. Seuraavaksi tutkitaan mitä voidaan sanoa tällöin mielivaltaisesta, mahdollisesti toisesta, saman avaruuden hermiittisestä muodosta.

Oletetaan, että K-vektoriavaruus on varustettu sisätulolla $\langle , \rangle$. Olkoon $F\colon V \times V \to K$ jokin *seskilineaarinen* muoto avaruudessa V ja olkoon $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jokin avaruuden V ortonormaali kanta. Muodon F matriisi $A = [F]_E$ kannan E suhteen määritellään $(n \times n)$-kokoisena matriisina $(a_{ij}) \in M(n \times n; K)$ jolle pätee

$$a_{ij} = F(\mathbf{e}_i, \mathbf{e}_j), i, j = 1, \ldots, n.$$

Olkoot $\mathbf{v}, \mathbf{w} \in V$ mielivaltaiset vektorit. Tällöin ne voidaan esittää ortonormaalissa kannassa E lineaarisina kombinaatioina

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i \text{ ja } \mathbf{w} = \sum_{i=1}^{n} w_i \mathbf{e}_i.$$

Tällöin, koska kanta E on ortonormaali, pätee

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} v_i \overline{w}_i. \tag{4.60}$$

Tämä tulos voidaan muotoilla myös seuraavasti. Merkitään symboleilla $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$ avaruuden K^n vektoreita, joiden komponentit ovat vektorien $\mathbf{v}, \mathbf{w}$ komponentit kannassa E. Täsmällisemmin sanottuna asetetaan

$$\vec{\mathbf{v}} = (v_1, \ldots, v_n) \text{ ja } \vec{\mathbf{w}} = (w_1, \ldots, w_n),$$

jolloin $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in K^n$. Tällä merkinnällä kaava 4.60 voidaan kirjoittaa muodossa

$$\langle \mathbf{v}, \mathbf{w} \rangle = \vec{\mathbf{v}} \cdot \vec{\mathbf{w}},$$

missä $\cdot$ tarkoittaa avaruuden K^n ”tavallista” pistetuloa (kts. esim. 4.10). Tällä tavalla mielivaltainen sisätulo äärellisulotteisessa vektoriavaruudessa voidaan ilmaista pistetulona. Huomaa, että itse asiassa $\vec{\mathbf{v}} = L(\mathbf{v})$, missä $L\colon V \to K^n$ on yksikäsitteinen lineaarinen isomorfismi, joka kuvaa kannan E avaruuden K^n standardikannaksi.

Palautetaan myös mieleen, että merkintä $A\vec{\mathbf{v}}$, missä A on $(n \times n)$-kokoinen matriisi ja $\vec{\mathbf{v}} \in K^n$, tarkoittaa matriisien A ja $\vec{\mathbf{v}}$ kertolaskua, missä jälkimmäinen tulkitaan ”pystyvektorina” eli $(n \times 1)$-kokoisena matriisina. Käyttämällä kanonista kuvausta $L_A\colon K^n \to K^n$ voidaan myös kirjoittaa $A\vec{\mathbf{v}} = L_A(\vec{\mathbf{v}})$.

Seuraavassa lemmassa näytetään, miten seskilineaarinen muoto voidaan ilmaista sen matriisin ja pistetulon avulla.

Lemma 4.61. *Olkoon V äärellisulotteinen K-sisätuloavaruus ja olkoon E sen ortonormaali kanta. Olkoon $F\colon V \times V \to K$ seskilineaarinen muoto ja olkoon $A = [F]_E$ sen matriisi kannassa E. Tällöin kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee*

$$F(\mathbf{v}, \mathbf{w}) = (A\vec{\mathbf{v}}) \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}} \cdot (A^* \vec{\mathbf{w}}).$$

Todistus. Esitetään vektorit $\mathbf{v}, \mathbf{w} \in V$ kannassa E,

$$\mathbf{v} = \sum v_i \mathbf{e}_i, \ \ \mathbf{w} = \sum w_i \mathbf{e}_i.$$

Tällöin, koska F on seskilineaarinen, pätee

$$F(\mathbf{v}, \mathbf{w}) = \sum_{i,j} v_i \overline{w_j} F(\mathbf{e}_i, \mathbf{e}_j) = \sum_{i,j} v_i \overline{w_j} a_{ij}.$$

Toisaalta määritelmän mukaan $A\vec{\mathbf{v}}$ on sellainen avaruuden K^n vektori jonka j:nnes komponentti on $c_j = \sum_{i=1}^n a_{ij} v_i, j = 1, \ldots, n$. Tästä seuraa pistetulon määritelmän nojalla, että

$$A\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \sum_{j=1}^n c_j \overline{w_j} = \sum_{i,j} a_{ij} v_i \overline{w_j}.$$

Vertaamalla tuloksia nähdään, että kaikilla $\mathbf{v}, \mathbf{w} \in V$ pätee $F(\mathbf{v}, \mathbf{w}) = A\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. Yhtälö $F(\mathbf{v}, \mathbf{w}) = \vec{\mathbf{v}} \cdot (A^* \vec{\mathbf{w}})$ saadaan tästä yhtälön 4.33 nojalla. □

Propositio 4.62. *Olkoon V äärellisulotteinen K-sisätuloavaruus ja olkoon E sen ortonormaali kanta. Olkoon $H \colon V \times V \to K$ seskilineaarinen muoto ja olkoon $A = [H]_E$ sen matriisi kannassa E. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) Muoto H on hermiittiinen.

(2) Matriisi A on itseadjungoitu.

(3) Avaruudella V on olemassa ortonormaali kanta E' jonka suhteen muoto H voidaan ilmaista kaavalla

$$H(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^n a_i x_i \overline{y_i},$$

missä $a_1, \ldots, a_n$ ovat kiinteitä reaalilukuja ja x_i (vastaavasti y_i) ovat vektorin $\mathbf{v} \in V$ (vastaavasti vektorin $\mathbf{w} \in V$) koordinaatteja kannassa E'. (Huomaa, että kanta E' voi siis olla eri kanta kuin E).

Todistus. (1)⇒(2) Olkoon $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ Oletetaan, että H on hermiittinen. Tällöin erityisesti kaikilla $i, j = 1, \ldots, n$ pätee

$$a_{ji} = H(\mathbf{e}_j, \mathbf{e}_i) = \overline{H(\mathbf{e}_i, \mathbf{e}_j)} = \overline{a_{ij}},$$

joten matriisi A on itseadjungoitu.

(2)⇒(3). Oletetaan, että matriisi A on itseadjungoitu, $A = A^*$. Proposition 4.55 nojalla itseadjungoitu matriisi A on diagonalisoituva ortonormaalissa kannassa ja lisäksi kaikki sen ominaisarvot ovat reaalilukuja. Olkoon $F = (\vec{\mathbf{f}}_1, \ldots, \vec{\mathbf{f}}_n)$ avaruuden K^n ortonormaali kanta (pistetulon suhteen), joka koostuu matriisin A ominaisvektoreista. Tällöin jokaisella $i = 1, \ldots, n$ pätee $A(\vec{\mathbf{f}}_i) = a_i \vec{\mathbf{f}}_i$ jollakin $a_i \in \mathbb{R}$. Olkoon $\vec{\mathbf{v}} \in K^n$ mielivaltainen. Esitetään se kannassa F eli muodossa $\vec{\mathbf{v}} = \sum x_i \vec{\mathbf{f}}_i$. Tällöin

$$A\vec{\mathbf{v}} = \sum_{i=1}^n a_i x_i \vec{\mathbf{f}}_i,$$

joten, koska kanta F on ortonormaali pistetulon suhteen ja pätee

$$A\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \sum_{i=1}^{n} a_i x_i \overline{y_i}. \tag{4.63}$$

Olkoon $L\colon V \to K^n$ lineaarinen isomorfismi, joka kuvaa avaruuden V kannan E avaruuden K^n standardikannaksi, tällöin $L(\mathbf{v}) = \vec{\mathbf{v}}$ kaikilla $\mathbf{v} \in V$. Lisäksi L on unitaarinen kuvaus, sillä se kuvaa erään ortonormaalin kannan ortonormaaliksi kannaksi (Propositio 4.35). Koska F on avaruuden K^n kanta ja $L^{-1}\colon K^n \to V$ on isomorfismi, se kuvaa kannan F erääksi avaruuden V kannaksi $E' = (L^{-1}(\vec{\mathbf{f}_1}), \ldots, L^{-1}(\vec{\mathbf{f}_n}))$. Koska F on ortonormaali ja L^{-1} on unitaarisen kuvauksen käänteiskuvauksena myös unitaarinen, kanta E' on ortonormaali (Propositio 4.35). Lisäksi kaikilla $\mathbf{v} \in V$ pätee yllä käytetyillä merkinnöillä

$$\mathbf{v} = L^{-1}(\vec{\mathbf{v}}) = L^{-1}(\sum x_i \vec{\mathbf{f}_i}) = \sum x_i (L^{-1}\vec{\mathbf{f}_i}),$$

mistä nähdään, että vektorin $\mathbf{v}$ koordinaatit kannassa E' ovat samat kuin vektorin $L(\mathbf{v}) = \vec{\mathbf{v}}$ koordinaatit kannassa F. Edellisessä lemmassa osoitetun yhtälön

$$H(\mathbf{v}, \mathbf{w}) = (A\vec{\mathbf{v}}) \cdot \vec{\mathbf{w}}$$

sekä yllä osoitetun yhtälön 4.63 nojalla saadaan

$$H(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} a_i x_i \overline{y_i},$$

missä $a_1, \ldots, a_n$ ovat kiinteitä reaalilukuja ja x_i (vastaavasti y_i) ovat vektorin $\mathbf{v} \in V$ (vastaavasti vektorin $\mathbf{w} \in V$) koordinaatteja kannassa E'. Ehto (3) on osoitettu.

(2)$\Rightarrow$(3). Yksinkertainen suoraviivainen lasku (harjoitustehtävä). $\square$

Edellisen tuloksen kohdassa (3) esiintyvän kannan E' alkioiden virittämiä yksiulotteisia aliavaruuksia sanotaan hermiittisen muodon H *pääakseleiksi*. Nämä akselit muodostavat siis avaruuden V ”suorakulmaisen koordinaatiston”, jossa hermiittisellä muodolla H on mahdollisimman yksinkertainen esitys.

Esimerkki 4.64. *Olkoon $H\colon \mathbb{R}^2 \to \mathbb{R}^2$ kaavalla*

$$H((x_1, y_1), (x_2, y_2)) = 3x_1y_1 + 4x_1y_2 + 4x_2y_1 - 3x_2y_2$$

määritelty kuvaus. Helposti nähdään, että tämä kuvaus on bilineaarinen ja symmetrinen, toisin sanoen (koska nyt $K = \mathbb{R}$) se on hermiittinen muoto $\mathbb{R}$-vektoriavaruudessa $\mathbb{R}^2$. Varustetaan $\mathbb{R}^2$ pistetulolla. Standardikannan E suhteen muodon H matriisi on symmetrinen matriisi

$$\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}.$$

Suoralla laskulla nähdään, että tällä matriisilla on kaksi ominaisarvoa, 5 ja -5. Olkoon E jokin ortonormaali kanta tasossa $\mathbb{R}^2$, joka koostuu tämän matriisin ominaisvektoreista. Kannassa E muoto H voidaan ilmaista kaavalla

$$H(\mathbf{x}, \mathbf{y}) = 5x_1'y_1' - 5x_2'y_2' = 5(x_1'y_1' - x_2'y_2'),$$

missä x_1', x_2' ovat vektorin $\mathbf{x}$ koordinaatit kannassa E ja vastaavasti y_1', y_2' ovat vektorin $\mathbf{y}$ koordinaatit kannassa E.

Positiivisesti semidefiniitit muodot

Vektoriavaruuden V hermittistä muotoa sanotaan *positiivisesti semidefiniitiksi* jos $H(\mathbf{v}, \mathbf{v}) \geq 0$ kaikilla $\mathbf{v} \in V$. Hermittista muotoa sanotaan *positiivisesti definiitiksi* jos $H(\mathbf{v}, \mathbf{v}) > 0$ kaikilla $\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}_V$.

Määritelmän mukaan vektoriavaruuden sisätulo on sama asia kuin tämän avaruuden positiivisesti definiitti hermiittinen muoto. Sisätuloavaruudessa siis asetetaan erikoisasemaan yksi tietty positiivisesti definiitti hermiittinen muoto. Seuraavaksi tutkitaan mitä tällöin voidaan sanoa avaruuden muista positiivisesti definiiteista hermiittisistä muodoista.

Lemma 4.65. *Olkoon H äärellisulotteisen sisätuloavaruuden V hermiittinen muoto. Olkoon E avaruuden V ortonormaali kanta ja olkoon $A = [H]_E$ muodon H matriisi kannan E suhteen. Tällöin seuraavat väitteet ovat yhtäpitäviä.*

(i) Muoto H on positiivisesti semidefiniitti.

(ii) Matriisin A kaikki ominaisarvot ovat ei-negatiivisia reaalilukuja.

(iii) Avaruudella V on olemassa ortonormaali kanta E' jonka suhteen muoto H voidaan ilmaista kaavalla

$$H(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} a_i x_i \overline{y_i},$$

missä $a_1, \ldots, a_n$ ovat kiinteitä ei-negatiivisia reaalilukuja ja x_i (vastaavasti y_i) ovat vektorin $\mathbf{v} \in V$ (vastaavasti vektorin $\mathbf{w} \in V$) koordinaatteja kannassa E'. (Tässä kanta E' voi siis olla eri kanta kuin E).

Todistus. Proposition 4.62 todistuksen nojalla kohdassa (iii) esiintyvät luvut a_i ovat joka tapauksessa täsmälleen hermiittisen muodon H matriisin A ominaisarvoja. Näin ollen Proposition 4.62 nojalla riittää osoittaa, että hermiittinen muoto H on positiivisesti semidefiniitti jos ja vain jos sen matriisin A kaikki ominaisarvot ovat ei-negatiivisia.

Olkoon H hermiittinen muoto. Proposition 4.62 (sekä sen todistuksen) nojalla avaruudella V on olemassa ortonormaali kanta E' jonka suhteen muoto H voidaan ilmaista kaavalla

$$H(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} a_i x_i \overline{y_i},$$

missä $a_1, \ldots, a_n$ ovat täsmälleen kaikki matriisin A ominaisarvot ja x_i (vastaavasti y_i) ovat vektorin $\mathbf{v} \in V$ (vastaavasti vektorin $\mathbf{w} \in V$) koordinaatteja kannassa E'. Erityisesti kaikilla $\mathbf{v} \in V$ pätee tällöin

$$H(\mathbf{v}, \mathbf{v}) = \sum_{i=1}^{n} a_i x_i \overline{x_i} = \sum_{i=1}^{n} a_i |x_i|^2 .$$

Koska koordinaatit x_i voivat saada mielivaltaisia arvoja skalaarikunnasta K, tästä nähdään, että epäyhtälö $H(\mathbf{v}, \mathbf{v}) \geq 0$ pätee kaikilla $\mathbf{v} \in V$ jos ja vain jos $a_i \geq 0$ kaikilla $i = 1, \ldots, n$. $\square$

Täysin samanlaisella todistuksella saadaan osoitettua seuraava karakterisaatio positiivisesti definiitille hermiittisille muodoille.

Lemma 4.66. *Olkoon H äärellisulotteisen sisätuloavaruuden V hermiittinen muoto. Olkoon E avaruuden V ortonormaali kanta ja olkoon $A = [H]_E$ muodon H matriisi kannan E suhteen. Tällöin seuraavat väitteet ovat yhtäpitäviä.*

(i) Muoto H on positiivisesti definiitti.

(ii) Matriisin A kaikki ominaisarvot ovat positiivisia reaalilukuja.

(iii) Avaruudella V on olemassa ortonormaali kanta E' jonka suhteen muoto H voidaan ilmaista kaavalla

$$H(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} a_i x_i \overline{y_i},$$

missä $a_1, \ldots, a_n$ ovat kiinteitä positiivisia reaalilukuja ja x_i (vastaavasti y_i) ovat vektorin $\mathbf{v} \in V$ (vastaavasti vektorin $\mathbf{w} \in V$) koordinaatteja kannassa E'. (Tässä kanta E' voi siis olla eri kanta kuin E).

Esimerkki 4.67. *Soveltamalla edellisen lemman sekä Lemman 4.61 tuloksia pistetulolla varustetussa vektoriavaruudessa K^n, nähdään, että mielivaltainen sisätulo $\langle , \rangle$ (eli positiivisesti definiitti hermiittinen muoto!) avaruudessa K^n on annettu kaavalla*

$$\langle \mathbf{v}, \mathbf{w} \rangle = A\mathbf{v} \cdot \mathbf{w},$$

missä A on jokin itseadjungoitu $(n \times n)$-kokoinen matriisi, jonka ominaisarvot ovat positiivisia reaalilukuja. Sisätulot äärellisulotteisessa vektoriavaruudessa siis vastaavat yksikäsitteisesti itseadjungoittuja matriiseja, joiden ominaisarvot ovat positiivisia.

Edelliset tulokset motivoivat seuraavan määritelmän. Itseadjungoitua matriisia $A \in M(n \times n; K)$ sanotaan *positiivisesti semidefiniitiksi* jos kaikki sen ominaisarvot (jotka ovat aina reaalisia) ovat ei-negatiivisia ja vastaavasti *positiivisesti definiitiksi* jos kaikki sen ominaisarvot ovat positiivisia. Puhutaan myös negatiivisesti definiitista (kaikki ominaisarvot negatiivisia) ja negatiivisesti semi-definiitista (kaikki ominaisarvot ei-positiivisia) itseadjunoiduista matriiseista.

Koska diagonalisoituva matriisi on kääntyvä jos ja vain jos nolla ei ole sen ominaisarvo, on positiivisesti semidefiiniitti matriisi positiivisesti definiitti jos ja vain jos se on kääntyvä.

Itseadjungoitua matriisia A sanotaan *ei-definiitiksi* jos se ei ole positiivisesti semidefiniitti eikä negatiivisesti semidefiniitti. Tämä ehto on ekvivalentti sen kanssa, että lauseke $A\mathbf{v} \cdot \mathbf{v}$ saa sekä aidosti positiivisia, että aidosti negatiivisia arvoja, kun $\mathbf{v}$ käy läpi avaruuden K^n vektoreita.

Esimerkki 4.68. *Olkoon $f : \mathbb{R} \to \mathbb{R}$ kahdesti derivoituva funktio. Analyysista tiedetään, että funktion f* ääriarvot *sijaitsevat derivaatan f' nollakohdissa eli niin sanotuissa* kriittisissä pisteissä. *Toisaalta käänteinen väite ei päde eli kriittisessä pisteessä ei välttämättä ole ääriarvoa. Silloinkin, kun kriittisestä pisteestä löytyy ääriarvo, olisi hyvä vielä osata* *selvittää onko kyseessä lokaali maksimi vai minimi. Tämän selvittämiseksi on kehitetty kaikenlaisia menetelmiä, joista yksi perustuu toisen derivatan f'' tutkimiseen. Nimittäin pätee seuraavas - jos kriittisessä pisteessä a on voimassa $f''(a) > 0$, niin siinä pisteessä funktiolla on (aito) lokaali minimi. Jos taas pätee $f''(a) < 0$ niin funktiolla on pisteessä aito lokaali maksimi. Jos $f''(a) = 0$ tämä "toisen derivaatan testi" ei kerro mitään.*

Tämä lähestymistapa voidaan yleistää funktioihin $f \colon \mathbb{R}^n \to \mathbb{R}$, jolloin toinen derivaatta korvaataan niin sanotulla Hessen matriisilla, *jonka kertoimet ovat f:n toisen kertaluvun osittaisderivaatat (annetussa pisteessä x_0) $D_{ij}f(x_0) = D_iD_jf(x_0)$. Tämä on siis matriisi*

$$H = \begin{bmatrix} D_{11}f(x_0) & D_{12}f(x_0) & \dots & D_{1n}f(x_0) \\ D_{21}f(x_0) & D_{22}f(x_0) & \dots & D_{2n}f(x_0) \\ \dots & \dots & \dots & \dots \\ D_{n1}f(x_0) & D_{n2}f(x_0) & \dots & D_{nn}f(x_0) \end{bmatrix}.$$

Oletetaan, että funktio f on vähintään C^2-luokkaa, eli sillä on olemassa kaikki toisen kertalun jatkuvat osittaisderivaatat. Analyysin kurssilla todistetaan, että tällöin funktion Hessen matriisi on symmetrinen. Tämä johtuu siitä, että peräkkäisessä osittaiderivoinnissa tulos ei riipu siitä, missä järjestyksessä osittaisderivaattoja otetaan, eli kaikilla i, j pätee $D_iD_jf = D_jD_if$.

C^2-luokkaan kuuluville funktiolle $f \colon \mathbb{R}^n \to \mathbb{R}$ on voimassa niin sanottu toisen kertaluvun Taylorin kaava

$$f(x) = f(a) + \sum_{i=1}^{n} D_if(x_0)(x_i - a_i) + \frac{1}{2}(\sum_{i,j} D_{ij}f(x_0)(x_i - a_i)(x_j - a_j)) + h(x)|x - x_0|^2,$$

missä $h(x) \to 0$ kun $x \to x_0$. Välttämätön (mutta ei riittävä) ehto sille, että C^2-funktiolla $f \colon \mathbb{R}^n \to \mathbb{R}$ on pisteessä x_0 ääriarvo on se, että piste x_0 on kriittinen, eli kaikki funktion ensimmäisen kertaluvun osittaisderivat häviävät pisteessä x_0. Toisin sanoen täytyy olla $D_i(x_0) = 0$ kaikilla $i = 1, \dots, n$. Näin ollen krittisessä pisteessä x_0 toisen kertaluvun Taylorin kaavasta tulee yhtälö

$$f(x) - f(x_0) = \frac{1}{2}(H(x - x_0) \cdot (x - x_0)) + h(x)|x - x_0|^2,$$

missä h on "pieni" kun ollaan lähellä pistettä x_0 ja H on Hessen matriisi.

Näin ollen, jos Hessen matriisi H on pisteessä x_0 positiivisesti definiitti, edellisen kaavan avulla nähdään, että epäyhtälö $f(x) - f(x_0) > 0$ pätee kun $x \neq x_0$ on "tarpeeksi lähellä" pistettä x_0. Tällöin pisteessä x_0 on määritelmän mukaan funktion f lokaali minimikohta. Jos taas H on negatiivisesti definiitti, samalla tavalla nähdään, että pisteessä x_0 funktiolla on lokaali maksimi. Jos H on ei-definiitti, pisteessä x_0 ei ole ääriarvoa. Jäljellä olevassa tapauksessa, jossa H on positiivisesti tai negatiivisesti semi-definiitti, ei voida sanoa mitään.

Hessen matriisin H definiittisyyttä voi tutkia lineaarialgebrallisin keinoin. Esimerkiksi voidaan osoittaa, että $\mathbb{R}$-kertoiminen symmetrinen matriisi A on positiivisesti definiitti *jos ja vain jos* $\det A_1 = a_{11} > 0, \det A_2 > 0, \ldots, \det A_{n-1} > 0, \det A_n = \det A > 0$. *Tässä* A_i *on matriisin* A $(i \times i)$*-kokoinen alimatriisi, jonka sarakkeet ovat* A*:n sarakkeiden* $c_1(A), \ldots, c_i(A)$ *"rajoittumat" riviin* i *asti.*

Tämä karakterisaatio on käytännössä erittäin hyödyllinen, sillä determinaanttien laskiminen on periaatteessa helpompaa kuin esimerkiksi ominaisarvojen selvittäminen. Tuloksella on myös teoreettista merkitystä.

Polaarinen hajotelma

Propositio 4.69. *Olkoon* $A \in M(n \times n; K)$*. Tällöin seuraavat ehdot ovat yhtäpitäviä.*

(1) A *on positiivisesti semidefiniitti itseadjungoitu matriisi.*

(2) On olemassa matriisi $B \in M(n \times n; K)$ *siten, että* $A = B^*B$*.*

(3) On olemassa yksikäsitteinen positiivisesti semidefiniitti itseadjungoitu matriisi $B \in M(n \times n; K)$ *siten, että* $A = B^2$*.*

Todistus. Osoitetaan, että (1)⇒(3)⇒(2)⇒(1).

Olkoon itseadjungoitu matriisi A positiivesesti semidefiniitti. Osoitetaan, että on olemassa yksikäsitteinen itseadjungoitu positiivisesti semidefiniitti matriisi B, jolle pätee $B^2 = A$.

Olkoon B itseadjungoitu positiivisesti semidefiniitti matriisi, jolle pätee $B^2 = A$. Ajatellaan sekä matriisia A, että matriisia B, vastaavina kuvauksina $L_A, L_B \colon K^n \to K^n$. Koska B on itseadjungoituna matriisina diagonalisoituva ortonormaalissa kannassa, avaruudella K^n on olemassa ortonormaali kanta E siten, että matriisi $[L_B]_E$ on diagonaalimatriisi

$$D = \begin{bmatrix} r_1 & 0 & \ldots & 0 \\ 0 & r_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & r_n \end{bmatrix},$$

missä diagonaalialkiot $r_1, \ldots, r_n$ ovat reaalilukuja (Propositio 4.55). Lisäksi, koska B oletetaan positiivisesti semidefiniitiksi, matriisin A ominaisarvot $r_1, \ldots, r_n$ ovat itse asiassa *ei-negatiivisia* reaalilukuja. Koska $L_A = L_B^2$, kuvauksen L_A matriisi kannan E suhteen on diagonaalimatriisi

$$D^2 = \begin{bmatrix} r_1^2 & 0 & \ldots & 0 \\ 0 & r_2^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & r_n^2 \end{bmatrix}.$$

Tästä nähdään, että matriisin A ominaisarvot ovat täsmälleen luvut $s_i = r_i^2$.

Olkoon V_r matriisin B ominaisarvon r määrämä aliavaruus

$$V_r = \{\mathbf{v} \in K^n \mid B^2(\mathbf{v}) = r\mathbf{v}\}$$

ja olkoon W_s vastaavasti matriisin A ominaisarvon s määrämä aliavaruus. Jokaiselle $\mathbf{v} \in V_r$ pätee

$$A(\mathbf{v}) = B^2(\mathbf{v}) = r_i^2\mathbf{v} = s_i\mathbf{s},$$

joten $V_r \subset W_{r^2}$. Koska sekä A, että B ovat diagonalisoituvia, pätee

$$\oplus V_r = V = \oplus W_{r^2},$$

missä summataan yli kaikkien matriisin B ominaisarvojen r. Tästä ja siitä, että $V_r \subset W_{r^2}$, seuraa helposti (mieti!), että itse asiassa $V_{r_i} = W_{r_i^2}$.

Matriisi B on siis yksikäsitteinen, sillä se edellisen nojalla se on määritelty (lineaarisena kuvauksena) jokaisessa matriisin A ominaisarvoaliavaruudessa W_s kaavalla $B(\mathbf{v}) = \sqrt{s}\mathbf{v}$. Kääntäen, koska aliavaruudet W_s muodostavat suoran summa, tällä tavalla voidaan määritellä yksikäsitteinen lineaarinen kuvaus $B\colon K^n \to K^n$ (suoran summan universaaliominaisuus). Helposti nähdään, että näin määritelty kuvaus on itseadjungoitu ja positiivisesti semidefiniitti, sillä se on silloin diagonalisoituva ortonormaalissa kannassa ja jokainen sen ominaisarvo on ei-negatiivinen reaaliluku. Lisäksi konstruktion perusteella pätee $B^2 = A$. On osoitettu, että ehdosta (1) seuraa ehto (3).

Ehto (3) implikoi ehdon (2) melkein triviaalisti, sillä itseadjungoidulle matriisille B pätee määritelmän perusteella $B^2 = B^*B$.

Implikaation (2)⇒(1) osoittamiseksi on näytettävä, että matriisi $A = B^*B$ on itseadjungoitu ja positiivisesti semidefiniitti jokaisella $B \in M(n \times n; K)$. Ehdosta $A = B^*B$ seuraa, että

$$A^* = (B^*B)^* = (B^*)^*B = B^*B = A.$$

Näin ollen A on itseadjungoitu. Olkoon $r \in \mathbb{R}$ jokin matriisin A ominaisarvo, tällöin on olemassa $\mathbf{v} \in K^n, \mathbf{v} \neq \mathbf{0}_V$, siten, että $B^*B(\mathbf{v}) = r\mathbf{v}$. Koska B^* on matriisin B adjungaatti, tästä seuraa, että

$$0 \leq B(\mathbf{v}) \cdot B(\mathbf{v}) = B^*B(\mathbf{v}) \cdot \mathbf{v} = r\mathbf{v} \cdot \mathbf{v} = r(\mathbf{v} \cdot \mathbf{v}),$$

missä $\cdot$ on kanoninen pistetulo K^n:ssä. Koska $\mathbf{v} \cdot \mathbf{v} > 0$, tästä seuraa, että $r \geq 0$. □

Edellisen tuloksen kohdan (3) voidaan ajatella tarkoittavan, että jokaisella itseadjungoidulla positiivisesti semidefiniitilla matriisilla A on olemassa yksikäsitteinen ”neliöjuuri” B itseadjungoitujen positiivisesti semidefiniittien matriisien joukossa. Kun A on positiivisesti definiitti (eli myös kääntyvä), myös sen yksikäsitteinen positiivisesti semidefiniitti neliöjuuri B on kääntyvä, joten se on jopa positiivisesti definiitti.

Propositio 4.70. Polaarinen hajotelma.
Jokainen neliömatriisi $A \in M(n \times n; K)$ voidaan esittää tulona US, missä matriisi $S \in M(n \times n; K)$ on itseadjungoitu positiivisesti semidefiniitti ja matriisi $U \in M(n \times n; K)$ on unitaarinen. Lisäksi matriisi S on aina yksikäsitteinen. Matriisi U on yksikäsitteinen jos A on kääntyvä. Tällöin S on jopa positiivisesti definiitti.

Todistus. Jos $A = US$, missä S on positiivisesti semidefiniitti ja U on unitaarinen, niin

$$A^*A = S^*U^*US = S^*S = S^2,$$

sillä $U^*U = I_n$ unitaarisen matriisin määritelmän nojalla ja $S^* = S$ itseadjungoidun matriisin määritelmän nojalla.

Edellisen proposition nojalla matriisi $B = A^*A$ on positiivisesti semidefiniitti. Saman proposition nojalla matriisilla B on yksikäsitteinen neliöjuuri itseadjungoitujen positiivisesti semidefiniittien matriisien joukossa. Koska $S^2 = B$, missä kumpikin S ja B ovat positiivisesti semidefiniitteja, matriisi S on yksikäsitteisesti määritelty - se on välttämättä yksikäsitteinen positiivisesti semidefiniitti neliöjuuri matriisista A^*A.

Jos matriisi A on kääntyvä, todistuksen loppuosa on suhteellisen helppo. Ensinnäkin tällöin pätee

$$(\det S)^2 = \det(S^2) = \det(A^*A) = \det A^* \det A = \overline{\det A} \det A = |\det A|^2 > 0,$$

mikä osoittaa sen, että myös matriisi S on kääntyvä. Erityisesti nolla ei voi olla sen ominaisarvo, joten kaikki matriisin A ominaisarvot ovat aidosti positiivisia. Toisin sanoen S on tällöin positiivisesti definiitti. Jos pätee $A = US$, niin välttämättä $U = AS^{-1}$. Jäljellä on tällöin vain sen osoittaminen, että tällä tavalla määritelty matriisi U on unitaarinen. Tämä on suoraviivainen lasku,

$$U^*U = (S^{-1})^*A^*AS^{-1} = S^{-1}S^2S^{-1} = \mathrm{id}\,.$$

Tässä on käytetty tietoa siitä, että kääntyvän itseadjungoidun matriisin käänteismatriisi on myös itseadjungoitu (tarkista).

Jos matriisi A ei ole kääntyvä, ei ole myöskään matriisi S, joten on keksittävä toinen todistustapa, erityisesti toinen tapa konstruoida sopiva matriisi U. Idea on seuraava - jos on olemassa sellainen unitaarinen matriisi U, jolle pätee $A = US$, niin jokaiselle $\mathbf{v} \in V$ välttämättä pätee

$$|A\mathbf{v}| = |U(S\mathbf{v})| = |S\mathbf{v}|\,,$$

sillä U säilyttää normit. Todistus etenee käänteisessä järjestyksessä - aloitetaan näyttämällä, että yhtälö $|A\mathbf{v}| = |S\mathbf{v}|$ on voimassa kaikilla $\mathbf{v} \in V$. Ajatellaan kaikki matriisit lineaarisina kuvauksina $K^n \to K^n$ (missä avaruus K^n varustetaan tavallisella pistetulolla). Olkoon $\mathbf{v} \in K^n$. Koska $S^2 = A^*A$ ja S on itseadjungoitu, pätee

$$|A(\mathbf{v})|^2 = \langle A(\mathbf{v}), A(\mathbf{v})\rangle = \langle A^*A(\mathbf{v}), \mathbf{v}\rangle = \langle S^2(\mathbf{v}), \mathbf{v}\rangle = \langle S(\mathbf{v}), S(\mathbf{v})\rangle = |S(\mathbf{v})|^2\,.$$

Näin ollen yhtälö $|A(\mathbf{v})| = |S(\mathbf{v})|$ pätee kaikilla $\mathbf{v} \in K^n$. Koska halutaan unitarinen kuvaus (matriisi) U, jolle pätee $US = A$, S:n kuvajoukossa $\operatorname{Im} S$ kuvauksen U täytyy toteuttaa ehto $U(\mathbf{w}) = A(\mathbf{v})$, missä $\mathbf{w} = S(\mathbf{v})$. Otetaan tämäkin havainto huomioon konstruktiota silmällä pitäen. Nimittäin, määritellään U ensin aliavaruudessa $W = \operatorname{Im} S$ kaavalla $U(\mathbf{w}) = A(\mathbf{v})$, missä $\mathbf{v}$ on sellainen vektori, jolle pätee $\mathbf{w} = S(\mathbf{v})$. Koska S ei välttämättä ole kääntyvä, tällainen $\mathbf{v}$ ei yleisesti ottaen kuitenkaan ole yksikäsitteinen, joten on ensin osoitettava, että tällä tavalla saadaan hyvinmääritelty kuvaus U. Toisin sanoen on näytettävä, että ehdosta $S(\mathbf{v}) = S(\mathbf{u})$ seuraa, että $A(\mathbf{v}) = A(\mathbf{u})$. Oletetaan, että $S(\mathbf{v}) = S(\mathbf{u})$. Tällöin

$$0_K = |S(\mathbf{v}) - S(\mathbf{u})| = |S(\mathbf{v} - \mathbf{u})| = |A(\mathbf{v} - \mathbf{u})|,$$

joten $A(\mathbf{v}) = A(\mathbf{u})$.

Näin ollen edellä konstruoitu kuvaus $U\colon W \to K^n$, missä $W = \operatorname{Im} S$, on hyvin määritelty. Helposti nähdään, että U on lineaarinen. Lisäksi määritelmänsä perusteella se toteuttaa ehdon $US = A$.

Tarkistetaan, että U säilyttää normit. Olkoon $\mathbf{w} = S(\mathbf{v}) \in W = \operatorname{Im} S$, $\mathbf{v} \in K^n$. Tällöin
$$|U(\mathbf{w})|^2 = |US(\mathbf{v})|^2 = |A(\mathbf{v})|^2 = |S(\mathbf{v})|^2 = |\mathbf{w}|^2 .$$
Kuvaus U on näin ollen unitarinen (Lemma 4.35), erityisesti injektiivinen (Lemman 4.35 seuraus). Tästä seuraa, että avaruuden $\operatorname{Im} U = W'$ dimensio on sama kuin avaruuden W dimensio. Todistus on valmis, kunhan osoitetaan vielä, että $U\colon W \to K^n$ voidaan laajentaa koko avaruuden K^n unitaariseksi kuvaukseksi $K^n \to K^n$. Huomaa, että tällöin ehto $US = A$ on edelleenkin voimassa.

Esitetään avaruus K^n suorana summana $K^n = W \oplus W^\perp$ sekä suorana summana $W' \oplus W'^\perp$. Tässä siis $W' = \operatorname{Im} U$. Koska $\dim W = \dim W'$, myös
$$\dim W^\perp = n - \dim W = n - \dim W' = \dim W'^\perp.$$
Valitaan aliavaruudelle $W^\perp$ ortonormaali kanta $E = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$. Valitaan vastaavalla tavalla aliavaruudelle $(W')^\perp$ ortonormaali kanta $E' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_k)$. Huomaa, että kummankin kannan pituus k on välttämättä sama, sillä avaruudet ovat samaa dimensiota. Määritellään lineaarinen kuvaus $U'\colon W^\perp \to W'^\perp$ ehdolla $U'(\mathbf{e}_i) = \mathbf{e}'_i$, $i = 1, \ldots, k$. Lemman 4.35 nojalla näin määritelty kuvaus U' on unitaarinen (kuvaa ortonormaalin kannan ortonormaaliksi kannaksi). Lopuksi laitetaan kuvaukset $U\colon W \to W'$ ja $U'\colon W^\perp \to W'^\perp$ "yhteen" määritelemällä kuvaus $U\colon K^n \to K^n$ kaavalla
$$U(\mathbf{w} + \mathbf{w}') = U(\mathbf{w}) + U'(\mathbf{w}'), \mathbf{w} \in W, \mathbf{w}' \in W^\perp.$$
Tämä on hyvin määritelty ja yksikäsitteinen, koska $W \oplus W^\perp = K^n$ (Propositio 4.22). Helposti nähdään, että näin määritelty kuvaus on unitaarinen. Propositio on todistettu. $\square$

Kun edellistä tulosta sovelletaan matriisin A sijaan sen adjungaattiin A^*, saadaan esitys $A^* = US$, missä U on unitaarinen ja S on positiivisesti semidefiniitti. Soveltamalla tämän yhtälön kumpaankin puoleen adjungaatti-operaatiota, saadaan matriisille A esitys muotoa $A = SU'$, missä S on itseadjungoitu positiivisesti semidefiniitti ja $U' = U^*$ on unitaarinen. Tätäkin tulosta sanotaan polaariseksi hajotelmaksi.

Nimitys "polaarinen hajotelma" tulee hyödyllisestä analogiasta matriisien ja kompleksilukujen välillä. Voidaan ajatella, että itseadjungoidut matriisit vastaavat reaalilukuja, positiivisesti semidefiniitit matriisit vastaavat ei-negatiivisia reaalilukuja ja unitaariset matriisit vastaavat yksikköympyrän alkioita $z \in S^1, |z| = 1$. Tällöin edellä johdettu matriisien polaarinen hajotelma $A = SU$ vastaa tuttua kompleksilukujen polaarista esitystä muodossa $z = rt$, missä $r = \sqrt{|z|} \geq 0$ on ei-negatiivinen reaaliluku ja $t \in S^1$.

Sovellus integraalilaskennassa.
Näytetään "käsiä heiluttamalla" miten edellisten tulosten avulla voidaan johtaa tärkeitä integrointiteorian tuloksia.

Analyysin peruskurssilta tuttu yksiulotteisen (Riemann)-integraalin "muuttujavaihtokaava" on yhtälö

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

Kun tätä tulosta yleistetään moniulotteiseen avaruuteen eli funktiolle $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \to \mathbb{R}^n$ derivaatta $g'(x)$ pitää korvata lineaarikuvauksen $g'(x)\colon \mathbb{R}^n \to \mathbb{R}^n$ determinantilla, jolloin vastaava moniulotteinen integraalilaskennan muuttujanvaihtokaava näyttää seuraavalta:

$$\int_{g(E)} f(x)dx = \int_E (f \circ g)(x) \left|\det g'(x)\right| dx. \tag{4.71}$$

Jos ollaan tarkkoja, tämän kaavan voimassaoloa varten pitää olettaa, että integroituva alue E sekä funktiot f, g ovat tarpeeksi "säännöllisiä". Teknisiä yksityiskohtia täsmennetään mitta-teorian ja modernin analyysin kursseilla. Esimerkiksi kaava on voimassa kun f ja g ovat jatkuvia kuvauksia ja E on suljettu kompakti osajoukko (esimerkiksi n-ulotteinen suorakulmio).

Muuttujavaihtokaava 4.71 voidaan johtaa seuraavien välivaiheiden kautta. Ensin käytetään hyväksi sitä, että derivaatta $g'(x)$ on kuvauksen g "lineaarinen approksimaatio pisteen x lähellä", joten lokaalisti g voidaan korvata derivaatallaan, jolloin nähdään, että kaava riittää osoittaa erikoistapauksessa, kun $g = L$ on lineaarinen kuvaus. Tämä taas palautuu yksinkertaisempaan väitteeseen, jonka mukaan

$$m(L(E)) = \left|\det L\right| m(E), \tag{4.72}$$

missä m on "mitta" eli geometrisen joukon "tilavuus" (pinta-ala) $\mathbb{R}^n$:ssä, E tarpeeksi säännöllinen integroituva alue sekä $L\colon \mathbb{R}^n \to \mathbb{R}^n$ lineaarinen kuvaus. Näiden välivaiheiden täsmällinen perustelu ja läpikäynti on pitkä prosessi täynnä epätriviaaleja yksityiskohtia, joihin tällä kurssilla ei ole mitään mahdollisuutta mennä sen enempää.

Todistuksen ydin pilee kuitenkin kaavan 4.72 todistamisessa. Se taas onnistuu polaarisen hajotelman avulla. Nimittäin lineaarisen kuvauksen L matriisi (standardin kannan suhteen) voidaan kirjoittaa muodossa $A = OS$, missä S on symmetrinen matriisi ja O ortogonaalinen. Lisäksi S on symmetrisenä matriisina diagonalisoituva ortonormaalissa kannassa, joten $S = O'DO'^{-1}$ jollakin diagonaalimatriisilla D ja ortogonaalisella matriisilla O'. Näin ollen A voidaan kirjoittaa muodossa $A = OO'DO'^{-1}$. Tästä seuraa, että kaava 4.72 riittää todistaa sellaisissa erikoistapauksissa, missä kuvauksen matriisi A on ortogonaalinen matriisi tai A on diagonaalinen matriisi.

Kumpikin väite on suhteellisen yksinkertainen. Nimittäin diagonaalimatriisi vain venyttää avaruuden $\mathbb{R}^n$ standardikoordinaattisuoria eikä ole vaikea laskea suoraan miten kappaleen mitta muuttuu tällaisessa muunnoksessa (vrt. suorakulmion tilavuuden laskemiseen). Ortogonaalinen kuvaus taas säilyttää pisteiden väliset etäisyydet, joten se kuvaa minkä tahansa pallon samankokoiseksi palloksi. Koska jokaisen joukon mittaa voidaan approksimoida mielivaltaisen tarkasti pallojen mitoilla (tämä osoitetaan esim. Mitta ja integraali-kurssilla), tästä seuraa, että $m(O(A)) = m(A) = |\det O| m(A)$, sillä ortogonaalisen kuvauksen determinaantin itseisarvo on aina 1. Näin ollen kaava 4.72 pätee myös ortogonaalisille matriiseille. Polaarinen hajotelman avulla tällöin helposti nähdään, että se pätee kaikille matriiseille.