CALCULUS FENNICUS

TKK:n 1. lukuvuoden laaja matematiikka (2000–2013)

Juhani Pitkäranta

Avoimet oppimateriaalit
Helsinki 2015

Tekijä Juhani Pitkäranta
ISBN 978-952-7010-12-9 (painettu)
978-952-7010-13-6 (pdf)
Kustantaja Avoimet oppimateriaalit ry
Kansi Essi Ruuskanen &
Lilja Tamminen

1. painos

YKL 51.07
UDK 51 (07)

© Juhani Pitkäranta 2015

Painopaikka Picaset, Helsinki, 2015.

Sisältö on lisensoitu vapaalla CC-BY 4.0 -lisenssillä, ehdot ja lisätietoa luettavissa osoitteessa http://creativecommons.org/licenses/by/4.0/deed.fi

Esipuhe

Tämä kirja on syntynyt kirjoittajan TKK:ssa (myöhemmin Aalto-yliopistossa) pitämien matematiikan luentojen pohjalta vuosina 2000–2013. Kirja on aiemmin ilmestynyt kaksiosaisena kattamaan TKK:n ensimmäisen lukuvuoden laajan matematiikan peruskurssit I–II.

Yliopistotasoisen (varmaan muunkin tasoisen) matematiikan perusopetuksen uudistaja joutuu väistämättä monien vaikeiden kysymysten ja valintojen eteen. Toisaalta matematiikka tuntuu olevan kuin kirkko, jossa perustotuudet ovat pysyviä ja liturgisiakin muutoksia vastustetaan kiivaasti. Toisaalta kuitenkin ympäröivä maailma muuttuu koko ajan ja muutospaineet kohdistuvat ennen pitkää myös matematiikkaan. Kirjoittaja on erityisesti joutunut pohtimaan, millainen pitkän aikavälin vaikutus tietokoneilla ja tehokkailla laskimilla mahdollisesti on matematiikan opetukseen yliopistotasolla. — Kun numeerisia ja symbolisia manipulaatioita voi suorittaa kätevästi koneella, niin millainen on se matematiikan taito, jota pitäisi opettaa ihmisille? Olisiko ehkä syytä korostaa aiempaa enemmän matematiikan käytännöllistä puolta ja sovelluksia? Vai päinvastoin matematiikkaa puhtaana abstraktin ajattelun taitona?

Jääköön lukijan pääteltäväksi, onko mainituilla mietteillä ollut lopputulokseen jokin vaikutus (ja jos, niin minkä suuntainen). Mainittakoon kuitenkin yksi tälle oppikirjalle ominainen, matematiikan vakiintuneesta opetusperinteestä poikkeava piirre: Kirjassa esitellään lukujen muodostamat lukujonot ja sarjat heti aluksi, jolloin reaaliluvut voidaan määritellä sellaisina kuin ne laskimien ja tietokoneiden maailmassa 'näkyvät' eli äärettöminä desimaalilukuina.

Kansainvälisenä vertailukohtana tälle kirjalle voi mainita amerikkalaiset Calculus-kirjat. Näihin verrattuna tämä kirja on tavoitteiltaan kunnianhimoisempi. Pyrkimyksenä on esittää aivan perusteista lähtevä, yhtenäinen ja loogisesti etenevä johdatus matematiikan perusideoihin ja moderniin laskutekniikkaan. Loogista aukottomuutta silmällä pitäen abstraktiotasoa on paikoin nostettu tyypillisestä, esim. mainituille Calculus-kirjoille ominaisesta 'katutasosta'. Kirja onkin tarkoitettu melko vaativalle yleisölle, jolle matematiikka on paitsi laskutekniikkaa ja 'kaavoja' myös älyllinen haaste ja aito kiinnostuksen kohde.

Otaniemen kotikentällä tämän kirjan edeltäjiä ovat apulaisprofessori Harri Rikkosen 1960- ja 1970-lukujen vaihteessa ja lehtori Simo K. Kivelän 1980- ja 1990-luvuilla kirjoittamat suomenkieliset oppikirjat tai luentomonisteet. Perinteen vaikutus näkyy myös tässä kirjassa — perinteestä on todella vaikea irrottautua. Perimätiedon ohella hyvin hyödyllisiä ovat olleet ne lukuisat elävät keskustelut, joita kirjoittaja on käynyt TKK:n matematiikan laitoksen henkilökuntaan kuuluvien kanssa. Monista kirjan syntyvaiheissa aktiivisista keskustelukumppaneista on syytä mainita erityisesti lehtori (nyk. emer.) Simo Kivelä, joka ansaitsee 'Latex-guruna' vielä erilliskiitoksen lukuisista matemaattista tekstinkäsittelyä koskevista neuvoista ja teknisestä avusta. Kiitoksen ansaitsee myös lehtori Pekka Alestalo kirjaamistaan yli sadasta kriittisestä kommentista vuodelta 2006, jolloin hän toimi laajojen peruskurssien I-II sijaisluennoitsijana.

Kirjan ensimmäinen, käsin kirjoitettu versio syntyi lukuvuonna 2000-2001. Ko. vuoden yleisöön kuulunut tekn. yo. (nyk. TkT) Antti H. Niemi työsti tekstistä ensimmäisen tietokoneistetun version kesällä 2002. Kirjan muodon teksti sai ensimmäisen kerran lukuvuonna 2005-2006. Uudistettuja painoksia on ilmestynyt vuosina 2007 ja 2009. Nyt ilmestyvässä 'lopullisessa' versiossa on kirjan sisältöä edelleen korjailtu ja virtaviivaistettu, ja myös harjoitustehtäviä on huomattavasti muokattu.

Kirjoittajan ensi kosketus TKK:n laajaan matematiikkaan tapahtui opiskelijana syksyllä 1968. Silloista Matematiikan pitkää peruskurssia luennoi Harri Rikkonen päärakennuksen D-salissa — samassa, jossa kirjoittaja piti ensimmäiset laajan peruskurssin luentonsa 32 vuotta myöhemmin. Noista ajoista on moni asia tätä kirjoitettaessa muuttunut: Enää ei ole TKK:ta eikä (syksyn 2013 opetusuudistuksen jälkeen) enää laajaa matematiikkaakaan. Välitöntä kurssikirjakäyttöä ei tällä kirjalla siis enää ole. Jääköön kirja kuitenkin ainakin verkkojulkaisuna muistuttamaan TKK:ssa omintakeisesti syntyneestä ja puoli vuosisataa jatkuneesta laajan matematiikan opetusperinteestä.

Kiitän Aalto-yliopiston Matematiikan ja systeemianalyysin laitosta taloudellisesta tuesta, joka on kannustanut oppikirjan viimeistelyyn vielä emerituksenakin. Koko luentoyleisöäni syksystä 2000 kevääseen 2013 muistelen lämmöllä ja kiitän tämän kirjan myötä vielä kerran myös kriittisestä palautteesta.

Helsingissä 15.6.2015

Juhani Pitkäranta

Sisältö

I Luvut ja lukujonot **1**

I.1 Katsaus rationaalilukuihin 2
I.2 Kunta 14
I.3 Logiikan ja joukko-opin merkinnöistä 25
I.4 Jonon käsite 35
I.5 Äärettömät desimaaliluvut 43
I.6 Lukujonon raja-arvo 54
I.7 Suppenevien lukujonojen ominaisuuksia 64
I.8 Monotoniset ja rajoitetut lukujonot 75
I.9 Reaaliluvut 86
I.10 Cauchyn jonot 96
I.11 Reaalilukujen ominaisuuksia 106
I.12 Klassinen sarjaoppi. Potenssisarja 113

II Vektorit ja analyyttinen geometria **127**

II.1 Euklidinen taso. Geometriset luvut 128
II.2 Tason vektorit 136
II.3 Skalaaritulo 149
II.4 *Abstrakti skalaaritulo ja normi 157
II.5 Funktion käsite. Trigonometriset funktiot 162
II.6 Avaruuden vektorit. Ristitulo 176
II.7 Suorien, tasojen ja pintojen geometriaa 190
II.8 Käyräviivaiset koordinaatistot 204

III Kompleksiluvut 215

III.1 Osoitinkunta 216

III.2 Kompleksiluvut ja niillä laskeminen 222

III.3 Algebran peruslause 229

III.4 *Kompleksikertoiminen vektoriavaruus 238

IV Reaalimuuttujien funktiot 241

IV.1 Yhden muuttujan funktiot 242

IV.2 Käänteisfunktio. Implisiittifunktiot 253

IV.3 Kahden ja kolmen muuttujan funktiot 265

IV.4 Parametriset käyrät ja pinnat 276

IV.5 *Funktioavaruus 290

V Jatkuvuus ja derivoituvuus 293

V.1 Jatkuvuuden käsite 294

V.2 Funktion raja-arvo 307

V.3 Derivaatta 318

V.4 Trigonometristen funktioiden derivointi 332

V.5 Ääriarvot. Sileys 340

V.6 Differentiaalilaskun väliarvolause 350

V.7 Kiintopisteiteraatio. Newtonin menetelmä 363

V.8 Analyyttiset kompleksifunktiot 376

V.9 *Jatkuvuuden logiikka 380

VI Eksponenttifunktio 391

VI.1 Yleinen eksponenttifunktio $E(x)$ 392

VI.2 Funktiot e^x ja $\ln x$ 399

VI.3 Kompleksinen eksponenttifunktio.
Hyperboliset funktiot 409

VI.4 Eksponenttifunktion sovellusesimerkkejä 417

VII Yhden muuttujan differentiaalilaskenta 421

VII.1 Differentiaali ja muutosnopeus 422
VII.2 Käyrän tangentti ja normaali 430
VII.3 Käyrän kaarevuus 440
VII.4 Taylorin polynomit ja Taylorin lause 450
VII.5 Taylorin polynomien sovelluksia 462
VII.6 Interpolaatiopolynomit 473

VIII Integraali 485

VIII.1 Integraalifunktio 486
VIII.2 Osittaisintegrointi ja sijoitus 493
VIII.3 Osamurtokehitelmät. Sarjamenetelmä 501
VIII.4 Integraalifunktion numeerinen laskeminen. Määrätty integraali 511
VIII.5 Riemannin integraali 523
VIII.6 Analyysin peruslause 534
VIII.7 Riemannin integraalin laajennukset 544
VIII.8 Pinta-ala ja kaarenpituus 556
VIII.9 Numeerinen integrointi 565

IX Differentiaaliyhtälöt 575

IX.1 Differentiaaliyhtälöiden peruskäsitteet 576
IX.2 Separoituva differentiaaliyhtälö 587
IX.3 Palautuvat toisen kertaluvun DY:t 597
IX.4 Ensimmäisen kertaluvun lineaarinen DY 606
IX.5 Lineaariset, vakiokertoimiset DY:t. Eulerin differentiaaliyhtälö 617
IX.6 *Yleinen 2. kertaluvun lineaarinen DY 628
IX.7 Differentiaaliyhtälöiden numeeriset ratkaisumenetelmät . . . 641
IX.8 *Picardin–Lindelöfin lause 649

X Matriisit **657**

X.1 Matriisialgebra 658

X.2 Neliömatriisit. Käänteismatriisi 670

X.3 Gaussin algoritmi 681

X.4 Tuettu Gaussin algoritmi. Singulaariset yhtälöryhmät 693

X.5 Determinantti 703

X.6 Lineaarikuvaukset 714

X.7 Affiinikuvaukset. Geometriset kuvaukset 726

X.8 Lineaaristen yhtälöryhmien sovellusesimerkkejä 736

XI Usean muuttujan differentiaalilaskenta **745**

XI.1 Usean muuttujan funktiot: Jatkuvuus ja raja-arvot 746

XI.2 Osittaisderivaatat 754

XI.3 Gradientti 769

XI.4 Vektorikentät: Divergenssi ja roottori 782

XI.5 Operaattorit grad, div, rot ja Δ käyräviivaisissa koordinaatistoissa 791

XI.6 Epälineaariset yhtälöryhmät: Jacobin matriisi ja Newtonin menetelmä 798

XI.7 Käänteisfunktiolause. Implisiittifunktiolause. Kontraktiokuvauslause 809

XI.8 Usean muuttujan ääriarvotehtävät 822

XI.9 Usean muuttujan Taylorin polynomit 835

XII Matriisin ominaisarvot ja neliömuodot **843**

XII.1 Matriisin ominaisarvot 844

XII.2 Neliömuotojen luokittelu 856

XII.3 Pinnan kaarevuus 870

XII.4 *Tensorit, vektorit ja skalaarit 877

XIII Usean muuttujan integraalilaskenta 883

XIII.1 Pinta-alamitta ja tasointegraalit 884

XIII.2 Tasointegraalien laskutekniikka 896

XIII.3 Avaruusintegraalit 908

XIII.4 Taso- ja avaruusintegraalien muuntaminen 917

XIII.5 Integraalien sovellukset: Tiheys ja kokonaismäärä 927

XIII.6 Viivaintegraalit 942

XIII.7 Pintaintegraalit 951

XIV Gaussin ja Stokesin lauseet 959

XIV.1 Vektorikentät ja polkuintegraalit 960

XIV.2 Gaussin lause 967

XIV.3 Gaussin lauseen sovelluksia 976

XIV.4 Stokesin lause 986

XIV.5 Pyörteetön vektorikenttä 994

Luku I

Luvut ja lukujonot

”Jumala loi luonnolliset luvut — kaikki muu on ihmistyötä” oli saksalaisen matemaatikon **Leopold Kronecker**in (1823-1891) aforismi ja kannanotto koskien matematiikan alkuperää. — Epäilemättä lukumäärää tai järjestystä kuvaavien sanojen ’yksi’, ’kaksi’, jne eriytyminen luonnollisessa kielessä on ollut eräs lähtökohta ja perusedellytys matematiikan syntymiselle. Merkkejä lukujen ja suhteiden tajusta on säilynyt jopa 30 000 vuoden takaa, joten matematiikan voi perustellusti katsoa olevan ihmisen kulttuurissa *sui generis*, omaa lajiaan. Kulttuuri-ilmiönä matematiikkaa voi jossain määrin ymmärtää rinnastamalla se muihin kulttuurin lajeihin. Esimerkiksi jos tarkastellaan matematiikan historiaa ja kumulatiivista rakennetta, tai käyttötapoja, voi nähdä monia yhtymäkohtia muihinkin mahtaviin ihmishengen ilmentymiin, erityisesti luonnollisiin kieliin ja tekniikkaan.

Matematiikkaa voi kuitenkin syvemmin oppia ymmärtämään vain matematiikkaa opiskelemalla. Opiskelu aloitetaan tässä luvussa ’alusta’, eli (matemaattisen) *luvun* käsitteestä. Luvun punaisena lankana on *reaalilukujen* konstruoiminen lähtien ’järjellisistä’ eli *rationaalisista* luvuista. Matematiikan historiassa reaaliluvun käsite on täsmentynyt lopullisesti vasta 1800-luvulla. Historiallisena ongelmana — ja edelleen ongelmana opetuksessa — on käsitteeseen väistämättä (muodossa tai toisessa) sisältyvä ’loputtomuuden’ ajatus. Tässä tekstissä reaalilukuja lähestytään rationaalilukujen muodostamien *lukujonojen*, ja näiden erikoistapausten, *äärettömien desimaalilukujen* kautta. Reaaliluvun käsite tulee tällä tavoin liitetyksi ’oikeaan laskemiseen’ ja myös tavanomaisiin numeromerkintöihin. Päättymättömät lukujonot tulevat jatkossa käyttöön monessa muussakin yhteydessä. Niihin perustuvat viime kädessä sekä monet laskentamenetelmät (algoritmit) että reaalilukuihin perustuvan matemaattisen analyysin eli *reaalianalyysin* käsitteet.

I.1 Katsaus rationaalilukuihin

Luonnollisten lukujen joukkoa merkitsemme

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\}.$$

Yleisemmin *joukko* (engl. set) koostuu *alkioista* (engl. element), joita joukon $\mathbb{N}$ tapauksessa ovat luvut 1, 2, 3 jne. Merkinnät

$$8 \in \mathbb{N}, \quad \infty \notin \mathbb{N}$$

luetaan 'kahdeksan kuuluu $\mathbb{N}$:ään, 'ääretön ei kuulu $\mathbb{N}$:ään', tarkoittaen yksinkertaisesti, että $\mathbb{N}$:ssä on alkio nimeltä 'kahdeksan' mutta ei alkiota nimeltä 'ääretön'.

Joukko $A = \{1, 2, 3\}$ on $\mathbb{N}$:n *osajoukko* (engl. subset). Merkitään

$$A \subset \mathbb{N} \quad \text{tai} \quad \mathbb{N} \supset A$$

ja luetaan 'A kuuluu/sisältyy $\mathbb{N}$:ään' tai '$\mathbb{N}$ sisältää A:n'. Yleisesti merkintä $A \subset B$ tarkoittaa, että jokainen A:n alkio on myös B:n alkio, toisin sanoen pätee

$$x \in A \quad \Rightarrow \quad x \in B.$$

Tässä '$\Rightarrow$' on logiikan symboleihin kuuluva nk. *implikaationuoli*: Merkintä $P \Rightarrow Q$ luetaan 'P:stä seuraa Q', tai 'jos P, niin Q'. Symboli '$\subset$' (kuuluminen osajoukkona) määritellään siis symbolin '$\in$' (kuuluminen alkiona) avulla. Jälkimmäinen symboli on nk. *primitiivi*, jota ei voi enää määritellä muiden symbolien avulla. Joukoista puhuttaessa voi siis merkintöjä $x \in A$ tai $x \notin A$ pitää kaiken loogisen ajattelun lähtökohtana.

Esimerkki 1 Jos $A = \{2, 1, 1, 2, 1\}$ ja $B = \{1, 2\}$, niin $A \subset B$, sillä jos $x \in A$, niin on joko $x = 1$, jolloin $x \in B$, tai $x = 2$, jolloin myöskin $x \in B$. Samalla päättelyllä nähdään, että pätee myös $B \subset A$. $\square$

Jos joukoille A ja B pätee sekä $A \subset B$ että $B \subset A$, kuten esimerkissä, niin sanotaan, että A ja B ovat (joukkoina) *samat*, ja merkitään $A = B$.[†] Logiikan ja joukko-opin merkintöjä käsitellään hieman laajemmin Luvussa I.3. Todettakoon tässä yhteydessä vielä osajoukon määrittelyssä tavallinen merkintä

$$B \;=\; \{\, x \in A \mid E \,\},$$

missä E on jokin *ehto*, joka jokaisen alkion $x \in A$ kohdalla joko toteutuu, jolloin $x \in B$, tai ei toteudu, jolloin $x \notin B$.

[†]Jos $A \subset B$, ja A ei ole sama kuin B, sanotaan että A on B:n *aito* (engl. proper) osajoukko. Joissakin matemaattisissa teksteissä symbolilla '$\subset$' on tämä rajatumpi merkitys, jolloin mahdollisuus $A = B$ varataan merkinnällä $A \subseteq B$. Tässä tekstissä siis pätee $A \subset A$.

Esimerkki 2 *Parilliset* ja *parittomat* luonnolliset luvut määritellään $\mathbb{N}$:n osajoukkoina

$$A = \{\, x \in \mathbb{N} \mid x = 2 \cdot y \text{ jollakin } y \in \mathbb{N} \,\} \;=\; \{\, 2, 4, 6, \;\ldots \},$$
$$B = \{\, x \in \mathbb{N} \mid x \notin A \,\} \;=\; \{\, 1, 3, 5, \;\ldots \} \quad \square$$

Esimerkissä on jo viitattu luonnollisten lukujen joukossa (tunnetulla tavalla) määriteltyihin *laskuoperaatioihin*, joita on kaksi:

$$\begin{array}{rlcll} \text{yhteenlasku:} & x \in \mathbb{N} \;\;\&\;\; y \in \mathbb{N} & \mapsto & x + y \in \mathbb{N} & \text{'}x \text{ plus } y\text{'} \\ \text{kertolasku:} & x \in \mathbb{N} \;\;\&\;\; y \in \mathbb{N} & \mapsto & x \cdot y \;\in \mathbb{N} & \text{'}x \text{ kertaa } y\text{'} \end{array}$$

Tässä '&' on logiikan merkintä, luetaan 'ja' (formaalimpi merkintä '$\wedge$'), ja '$\mapsto$' tarkoittaa *liittämistä*: Lukuihin x ja y, eli $\mathbb{N}$:n *lukupariin*, liitetään (yksikäsitteinen) luku $x + y \in \mathbb{N}$, nimeltään x:n ja y:n *summa*, ja $x \cdot y \in \mathbb{N}$, nimeltään x:n ja y:n *tulo*. Laskuoperaatio tarkoittaa siis yksittäisten lukujen x, y tapauksessa 'liittämissääntöä', yleisemmin 'liittämissäännöstöä'. Säännöt voidaan ajatella joko joistakin yleisemmistä periaatteista (laskusäännöistä) johdetuiksi, tai ne voidaan ymmärtää pelkästään 'luettelona', joka kattaa kaikki lukuparit. Kun yhteen- ja kertolaskun sisältö on yhteisesti sovittu (ja peruskoulun oppikirjoihin painettu), on luonnollisten lukujen perustalle rakennettu *algebra*†, jota merkittäköön $(\mathbb{N}, +, \cdot)$.

Seuraavat hyvin tunnetut peruslait ovat voimassa luonnollisten lukujen laskuoperaatioille:

(L1) $x + y = y + x$

(L2) $x \cdot y = y \cdot x$

(L3) $x + (y + z) = (x + y) + z$

(L4) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

(L5) $x \cdot (y + z) = x \cdot y + x \cdot z$

(L6) $x \cdot 1 = x \;\; \forall x \in \mathbb{N}$

Tässä (L1), (L2) ovat *vaihdantalait*, (L3), (L4) ovat *liitäntälait* ja (L5) on *osittelulaki*. Viimeksi mainitulla on myös rinnakkainen muoto

†Algebra eli 'kirjainlaskenta' on numeroilla laskemisen eli *aritmetiikan* abstraktimpi muoto. Algebran kehitys alkoi arabikulttuurin piirissä 1. vuosituhannella jKr.

(L5') $(x+y)\cdot z = x\cdot z + y\cdot z,$

joka seuraa yhdistämällä (L2) ja (L5). Laissa (L6) '$\forall$' on logiikan symboli, joka luetaan 'kaikille'. Tämä laki ilmaisee, että luku 1 on kertolaskun nk. *ykkösalkio.* Kuten nähdään jäljempänä, laskulait (L1)–(L6) ovat yhteisiä monille algebroille. Yleisemmissä tarkasteluissa nämä lait otetaankin usein oletetuiksi peruslaeiksi eli *aksioomiksi.* †

Laskulaeissa (L1)–(L6) on oletettu normaalit järjestyssäännöt laskuoperaatioiden yhdistelylle, eli ensin lasketaan sulkeiden sisällä olevat lausekkeet, ja kertolaskut aina ennen yhteenlaskuja mikäli sulkeita ei ole. Huomattakoon, että tässä ei ole kyse mistään ylimääräisistä olettamuksista, vaan esimerkiksi sääntö

$$x + y\cdot z \;=\; x + (y\cdot z)$$

on <u>sulkeiden</u> <u>poisjättösääntö</u>, eli kyse on merkintäsopimuksesta. Samaa operaatiota yhdisteltäessä on liitäntälakien (L3), (L4) perusteella tulos aina sama operaatioiden järjestyksestä riippumatta, joten sulkeet voidaan jättää pois (merkintäsopimus!) ja kirjoittaa

$$x + y + x + \ldots, \quad x\cdot y\cdot z\cdot\ldots$$

Tavanomaisen lukujen algebran merkintäsopimuksiin kuuluu myös vapaus jättää kertomerkki merkitsemättä sikäli kuin sekaannuksen vaaraa ei ole:

$$x\cdot y = xy, \quad x\cdot 2 = 2x, \quad 2\cdot(1+2) = 2(1+2) = 2\cdot 3 \neq 23.$$

Peräkkäisiä tuloja laskettaessa on saman luvun n-kertaisesta ($n\in\mathbb{N}$) kertolaskusta tapana käyttää nimitystä *potenssiin korotus* ja käyttää merkintää

$$\underbrace{x\cdot x\cdots x}_{(n\ \text{kpl})} \quad = \quad x^n \qquad \text{'}x \text{ potenssiin } n\text{'} \quad (n\in\mathbb{N}).$$

†Luonnollisten lukujen tapauksessa laskulait (L1)–(L6) ovat johdettavissa luonnollisten lukujen perustana olevista **Peanon aksioomista**. Aksioomat esitti italialainen matemaatikko **Giuseppe Peano** (1858–1932) vuonna 1889. Peanon aksioomat luonnollisille luvuille ovat: (P1) $1\in\mathbb{N}$. (P2) Jokaisella $x\in\mathbb{N}$ on yksikäsitteinen *seuraaja* $x'\in\mathbb{N}$. (P3) Luku 1 ei ole minkään $x\in\mathbb{N}$ seuraaja. (P4) Jokainen $x\in\mathbb{N}$ on korkeintaan yhden luvun seuraaja, ts. jos $x,y\in\mathbb{N}$ ja $x'=y'$, niin $x=y$. (P5) Jos $S\subset\mathbb{N}$ ja pätee (i) $1\in S$ ja (ii) $x'\in S$ aina kun $x\in S$, niin $S=\mathbb{N}$. Aksioomassa (P2) esiintyvä luvun seuraaja on tavanomaisen kielenkäytön mukaan 'seuraava luku'. Aksiooman (P3) mukaan luku 1 on luonnollisista luvuista 'ensimmäinen', sen sijaan 'viimeistä' lukua ei ole, koska jokaisella luvulla on seuraaja (aksiooma (P2)) eikä mikään luku voi esiintyä seuraajaketjussa kahdesti (aksioomat (P3) ja (P4)). Luonnollisia lukuja on siis äärettömän monta. Aksiooma (P5) asettaa nk. *induktioperiaatteen*, ks. Luku I.4 jäljempänä.

Luonnollisten lukujen yhteenlasku ja kertolasku määritellään aksioomiin (P1)–(P5) perustuen asettamalla kummallekin laskuoperaatiolle kaksi aksioomaa. Näiden perusteella minkä tahansa halutun laskuoperaation tulos on määrättävissä lukujen 1, $1'=2$, $2'=3$, jne avulla. Yhteenlaskun aksioomat: (Y1) $x+1=x'$, (Y2) $x+y'=(x+y)'$. Kertolaskun aksioomat: (K1) = (L6), (K2) $x\cdot y' = x\cdot y + x$.

Kokonaisluvut $\mathbb{Z}$

Kokonaislukujen joukko $\mathbb{Z}$ on ensimmäinen askel siinä 'ihmistyössä', jossa lukujärjestelmää pyritään laajentamaan lähtökohtana luonnolliset luvut. Laajennus koostuu kahdesta osa-askeleesta, joista ensimmäinen on luvun *nolla* ('ei mitään') ja toinen *vastaluvun* ('montako puuttuu') käyttöönotto. Näiden ajatusten tuloksena syntyvää lukujoukkoa merkitään

$$\mathbb{Z} = \{\ \ldots, -2, -1, 0, 1, 2,\ \ldots\ \} \quad = \quad \{\ 0, \pm 1, \pm 2,\ \ldots\ \}.$$

Lukualueen laajennus tehdään aksiomaattisesti olettamalla seuraavat lisäaksioomat, jotka siis ovat voimassa $\mathbb{Z}$:ssa mutteivät $\mathbb{N}$:ssä.

(L7) On olemassa *luku* 0, jolle pätee $x + 0 = x \ \ \forall\, x \in \mathbb{Z}$.

(L8) Jokaisella $x \in \mathbb{Z}$ on *vastaluku* $-x$, jolle pätee $x + (-x) = 0$.

Lyhennysmerkintä

$$x + (-y) = x - y \qquad \text{'}x \text{ miinus } y\text{'}$$

tuo $\mathbb{Z}$:aan uuden laskuoperaation, *vähennyslaskun*, jonka tulos on nimeltään lukujen x ja y *erotus*. Tässä on siis kyse yhteenlaskun ja 'merkinvaihdon' $x \mapsto -x$ (jonka voi myös tulkita laskuoperaatioksi) yhdistämisestä.

Kokonaisluvuilla laskettaessa pidetään selviönä, että luku 0 samoin kuin vastaluvut $-1,\ -2, \ldots$ ovat yksikäsitteisiä. Tuttuja ovat myös laskusäännöt $-(-x) = x$, $0 \cdot x = 0$ ja $(-1) \cdot x = -x$. Näitä oletuksia / laskusääntöjä ei tarvitse kuitenkaan sisällyttää kokonaislukujen aksioomiin, sillä ne ovat aksioomien seurauksia:

Lause I.1.1 † Kokonaisluvuille pätee aksioomien (L1)–(L8) perusteella:
(a) Luku 0 on yksikäsitteinen, samoin vastaluku $-x$ jokaisella $x \in \mathbb{Z}$.
(b) $-(-x) = x \ \ \forall x \in \mathbb{Z}$.
(c) $0 \cdot x = 0 \ \ \forall x \in \mathbb{Z}$.
(d) $(-1) \cdot x = -x \ \ \forall x \in \mathbb{Z}$.

†Kun matemaattisessa tekstissä halutaan nostaa esiin jokin selkeästi muotoiltu, tosi väittämä, esim. siksi että väittämään halutaan myöhemmin viitata tai kun halutaan korostaa tuloksen uutuusarvoa, käytetään nimityksiä *Lause* eli *Teoreema* (engl. Theorem), *Propositio* (engl. Proposition), *Apulause* eli *Lemma* (engl. Lemma) ja *Seurauslause* eli *Korollaari* (engl. Corollary). Tässä tekstissä käytetään (hieman epäjohdonmukaisesti) nimityksiä Lause, Propositio, Lemma ja Korollaari. Propositio on lausetta vähäarvoisempi tai sisällöltään teknisempi väittämä — nimitystä käytetään lähinnä, kun halutaan luokitella lauseita niiden painoarvon mukaan. Lemma on yleensä välietappi jonkin lauseen/proposition todistamisessa ja korollaari jo todistetun lauseen tai proposition suoraviivainen seuraamus.

Lause I.1.1 tulee todistetuksi seuraavassa luvussa osana yleisempää lausetta, joten tyydytään tässä todistamaan ainoastaan helpoin osaväittämä (b): Koska aksioomien (L1) ja (L8) mukaan on

$$-x + x = x + (-x) = 0,$$

niin vastaluvun määritelmän mukaisesti on $-(-x) = x$. $\square$

Esimerkki 3 Lauseen I.1.1 laskusäännöt (b) ja (d) yhdistämällä nähdään, että $(-1) \cdot (-1) = -(-1) = 1$. $\square$

Lukujärjestelmät

Kokonaisluvut on käytännössä ilmaistava jossakin *lukujärjestelmässä*. Tavanomaisin, matematiikan 'äidinkieleksi' vakiintunut on *kymmenjärjestelmä* (engl. decimal system), jossa kirjoitetaan

$$\begin{aligned} 1+1 &= 2 \quad \text{'kaksi'}, \\ 2+1 &= 3 \quad \text{'kolme'}, \\ &\vdots \\ 8+1 &= 9 \quad \text{'yhdeksän'}, \\ 9+1 &= 10 \quad \text{'kymmenen'}. \end{aligned}$$

Tässä merkit $0 \ldots 9$ ovat *kymmenjärjestelmän numerot* (engl. digits, lat. digitus = sormi,varvas). Luku 'kymmenen' on luonnollisista luvuista ensimmäinen, jolla ei ole ko. järjestelmässä omaa symbolia, vaan se merkitään yhdistelmänä 'yksi nolla'. Tämä on järjestelmän *kantaluku*. Kymmenjärjestelmän luvut ilmaistaan tämän jälkeen symbolisesti muodossa

$$x \;=\; eabcd\ldots \;\in\; \mathbb{Z},$$

missä e on *etumerkki* (engl. sign), $e = +$ tai $e = -$ ('tyhjä' $= +$), ja $a, b, c, d, \ldots$ ovat kymmenjärjestelmän numeroita. Numeroiden lukumäärästä riippuen tulkinta on

$$ea = \pm a, \quad eab = \pm(a \cdot 10 + b), \quad eabc = \pm(a \cdot 10^2 + b \cdot 10 + c), \quad \text{jne.}$$

Tässä etumerkki on e:n mukainen.[†]

[†]Lukumerkinnän tulkinnan mukaisesti on $\pm 0abc.. = \pm abc..$, ts. 'etunollat' eivät vaikuta luvun tulkintaan.

Yleisemmin jos lukujärjestelmän kantaluku on luku $k \in \mathbb{N}$, $k \neq 1$, niin tämän *k-järjestelmän* luvuilla $0, 1, \ldots, k-1$ on oltava kullakin oma numeromerkki.† Jos $a, b, c, \ldots$ ovat tällaisia merkkeja, niin luku $x = eabcd\ldots$ tulkitaan samoilla säännöillä kuin edellä. On huomioitava ainoastaan, että luku '10' ei nyt ole 'kymmenen', vaan tällä tarkoitetaan lukua $1 \cdot k + 0 = k$. Kantaluvun merkki on siis aina '10', sen sijaan ääntämistapa (esim. 'kymmenen','tusina','tiu') on sovittava erikseen kussakin lukujärjestelmässä (ja kussakin kielessä). Ellei mitään ole sovittu, on luettava 'yksi nolla'. Lukujärjestelmistä kaikkein yksinkertaisin numeromerkistöltään on *binaarijärjestelmä*, jossa kantalukuna on 'ykkösestä seuraava' eli 2. Binaarijärjestelmän numeroita eli *bittejä* ovat vain 0 ja 1.

Esimerkki 4 Kymmenjärjestelmän luku 1253 voidaan lukea 'yksi kaksi viisi kolme', jolloin ainoastaan äännetään luvun 'nimi'. Lukutavassa 'tuhat kaksi sataa viisikymmentä kolme' annetaan jo luvulle tulkinta: 1253 = 'tuhat plus kaksi kertaa sata plus viisi kertaa kymmenen plus kolme'. Tässä 'sata' ja 'tuhat' ovat suomen kielessä sovittuja (ja kouluissa opetettuja) ääntämistapoja luvuille $100 = k^2$ ja $1000 = k^3$, kun $k =$ 'kymmenen'. □

Esimerkki 5 Lausu kymmenjärjestelmän luku 11 binaarijärjestelmässä.

<u>Ratkaisu</u> Kymmenjärjestelmässä voidaan kirjoittaa $11 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1$. Binaarijärjestelmän esitysmuoto saadaan tästä kirjoittamalla merkin '2' tilalle '10':

$$11 \;=\; 1 \cdot 10^3 + 0 \cdot 10^2 + 1 \cdot 10 + 1 \;=\; 1011 \quad \text{(binaarijärjestelmä).} \qquad \square$$

Esimerkki 6 Laske 7-järjestelmän lukujen 145 ja 66 summa ja muunna tulos kymmenjärjestelmään.

<u>Ratkaisu</u> Kun merkitään $k = 7$, niin laskulakeja (L1)–(L6) soveltaen saadaan

$$\begin{aligned} 145 + 66 &= (1 \cdot k^2 + 4 \cdot k + 5) + (6 \cdot k + 6) \\ &= 1 \cdot k^2 + (4 + 6) \cdot k + (5 + 6) \\ &= 1 \cdot k^2 + (1 \cdot k + 3) \cdot k + (1 \cdot k + 4) \\ &= (1 + 1) \cdot k^2 + (3 + 1) \cdot k + 4 \\ &= 2 \cdot k^2 + 4 \cdot k + 4 \\ &= 244. \end{aligned}$$

Kymmenjärjestelmässä ilmaistuna lopputulos on $2 \cdot 7^2 + 4 \cdot 7 + 4 = 130$. □

†Lukujärjestelmien idea keksittiin Kaksoisvirranmaassa noin 4000 vuotta sitten. Varhaisimmissa järjestelmissä kantaluku oli 60. Numeromerkit olivat nuolenpäämerkkien yhdistelmiä.

Rationaaliluvut $\mathbb{Q}$

Siirryttäessä $\mathbb{Z}$:sta *rationaalilukujen* joukkoon $\mathbb{Q}$ tarvitaan enää yksi lisäaksiooma:

(L9) Jokaisella $x \neq 0$ on *käänteisluku* x^{-1} siten, että $x \cdot x^{-1} = 1$.

Lyhennysmerkintä

$$x \cdot y^{-1} = x/y = \frac{x}{y} \qquad \text{'}x \text{ per } y\text{'}$$

määrittelee neljännen laskuoperaation, *jakolaskun*, jonka tulos on nimeltään lukujen x ja y *osamäärä*. Jakolaskussa on siis kyse 'kääntämisoperaation' $y \mapsto y^{-1}$ ja kertolaskun $x, y^{-1} \mapsto x \cdot y^{-1}$ yhdistämisestä. Sanotaan, että luku x on tässä operaatiossa *osoittaja* ja luku y on *nimittäjä* (engl. numerator, denominator). Jakolaskun määritelmän ja aksioomien (L2), (L6) perusteella on

$$x^{-1} = x^{-1} \cdot 1 = 1 \cdot x^{-1} = 1/x.$$

Kun lukujoukkoa $\mathbb{Z}$ laajennetaan siten, että laajennetussa joukossa $\mathbb{Q}$ pätevät aksioomat (L1)-(L9), on tuloksena uusi algebra, jota merkitään $(\mathbb{Q}, +, \cdot)$. Aksiooman (L9) ja jakolaskun määritelmän mukaisesti $\mathbb{Q}$:n alkioita ovat ainakin luvut muotoa $x = p/q$, missä $p \in \mathbb{Z}$ ja $q \in \mathbb{N}$. Osoittautuu (ks. Esimerkki 7 alla ja Harj.teht. 7), että lähdettäessä mainittua muotoa olevista luvuista ovat laskuoperaatioiden $x \mapsto -x$, $x \mapsto x^{-1}$, $x, y \mapsto x + y$ ja $x, y \mapsto x \cdot y$ aina samaa muotoa, eli muun tyyppisiä lukuja ei $\mathbb{Q}$:ssa ole. Rationaalilukujen joukko $\mathbb{Q}$ voidaan siten määritellä

$$\mathbb{Q} = \{\ x = p/q \mid p \in \mathbb{Z} \ \ \& \ \ q \in \mathbb{N}\ \}.$$

Aksioomista (L1)-(L9) voidaan johtaa rationaalilukujen kaikki normaalit laskusäännöt. Seuraavassa kaksi esimerkkiä (ks. myös Harj.teht. 7).

Esimerkki 7 Vedoten aksioomiin (L1)–(L9) ja Lauseeseen I.1.1 perustele rationaalilukujen laskusäännöt

$$\text{a)}\ \frac{0}{q} = 0 \ \ (q \in \mathbb{N}), \qquad \text{b)}\ \frac{1}{q_1} \cdot \frac{1}{q_2} = \frac{1}{q_1 q_2} \ \ (q_1, q_2 \in \mathbb{N}).$$

<u>Ratkaisu</u> a) Lähtien rationaaliluvun määritelmästä ja vedoten Lauseen I.1.1 väittämään (c) sekä aksioomiin (L4), (L9) ja (L6) päätellään:

$$\begin{aligned} 0/q = 0 \cdot q^{-1} &= (0 \cdot q) \cdot q^{-1} \\ &= 0 \cdot (q \cdot q^{-1}) \\ &= 0 \cdot 1 \\ &= 0. \end{aligned}$$

b) Väitetään, että $q_1^{-1}q_2^{-1} = (q_1q_2)^{-1}$, ts. että $y = q_1^{-1}q_2^{-1}$ on luvun $x = q_1q_2$ käänteisluku. Tarkistus:

$$\begin{aligned} x \cdot y = (q_1q_2) \cdot (q_1^{-1}q_2^{-1}) &= (q_1q_2) \cdot (q_2^{-1}q_1^{-1}) \\ &= q_1(q_2q_2^{-1})q_1^{-1} \\ &= q_1 \cdot 1 \cdot q_1^{-1} \\ &= q_1q_1^{-1} = 1. \end{aligned}$$

Tässä vedottiin aksioomiin (L2), (L4), (L9) ja (L6). $\square$

Kun käänteisluvun potenssiin korotuksessa sovitaan merkinnästä

$$(x^{-1})^n = \quad x^{-n}, \quad n \in \mathbb{N},$$

voidaan x^{-1} lukea myös 'x potenssiin miinus 1'. Aksioomat (L2), (L4), (L6) ja (L9) huomioiden seuraa myös

$$x \neq 0 \;\;\&\;\; m, n \in \mathbb{Z} \;\;\&\;\; m, n \neq 0 \quad \Rightarrow \quad x^m \cdot x^n = \begin{cases} x^{m+n}, & \text{kun } m+n \neq 0, \\ 1, & \text{kun } m+n = 0. \end{cases}$$

Tämän säännön yksinkertaistamiseksi on luontevaa <u>määritellä</u>:

$$\boxed{x^0 = 1, \quad \text{kun } x \neq 0 \;\; (x \in \mathbb{Q}).}$$

Tällöin saadaan yleiset laskusäännöt

$$\boxed{x^m x^n = x^{m+n}, \quad x^n y^n = (xy)^n, \quad (x^m)^n = x^{mn}, \quad x, y \neq 0, \; m, n \in \mathbb{Z}.}$$

Myös $0^n = 0 \;\; \forall\, n \in \mathbb{N}$. Sen sijaan määrittelemättä jäävät 0^0 ja 0^{-n}, $n \in \mathbb{N}$, syystä että 0:lla ei ole käänteislukua.

$\mathbb{Q}$:n järjestysrelaatio

Rationaaliluvut voidaan tunnetulla tavalla myös *järjestää*, eli voidaan määritellä *järjestysrelaatio*

$$< : \quad \text{'pienempi kuin'}.$$

Tällöin syntyvää algebraa voidaan merkitä $(\mathbb{Q}, +, \cdot, <)$. Järjestysrelaatioon liitetään seuraavat yleiset aksioomat (eli oletetut peruslait):

(J1) Vaihtoehdoista $\{x < y,\ x = y,\ x > y\}$ on voimassa täsmälleen yksi.

(J2) $x < y \ \ \& \ \ y < z \ \ \Rightarrow \ \ x < z.$

(J3) $x < y \ \ \Rightarrow \ \ x + z < y + z \ \ \forall z.$

(J4) $x > 0 \ \ \& \ \ y > 0 \ \ \Rightarrow \ \ x \cdot y > 0.$

Tässä merkintä $x > y$ ('x suurempi kuin y') on merkinnän $y < x$ vaihtoehtoinen muoto. Käytännössä vertailusta selvitään nopeimmin käyttämällä laskulakeja (L7), (L8) ja (L1), sillä näistä ja aksioomasta (J3) on pääteltävissä, että pätee

$$\begin{cases} x < y & \Leftrightarrow \quad x - y < 0, \\ x = y & \Leftrightarrow \quad x - y = 0, \\ x > y & \Leftrightarrow \quad x - y > 0. \end{cases}$$

Tässä '$\Leftrightarrow$' on logiikan symboleihin kuuluva *ekvivalenssinuoli*, joka voidaan tässä yhteydessä lukea

$$\text{P} \quad \Leftrightarrow \quad \text{Q} \ : \quad \text{'P täsmälleen kun Q'}.$$

Ym. kriteerillä voidaan siis vertailukysymys $x \ ? \ y$ selvittää tutkimalla, onko luku $x - y$ *positiivinen* (> 0), *negatiivinen* (< 0), vai $= 0$. Perusmuotoisesta rationaaliluvusta $x = p/q$, ($p \in \mathbb{Z},\ q \in \mathbb{N}$) puolestaan sovitaan, että vaihtoehdot $x > 0$ ja $x = 0$ vastaavat vaihtoehtoja $p \in \mathbb{N}$ ja $p = 0$, muulloin on $x < 0$. Tämä sopimus on mahdollinen, koska Esimerkin 7 laskusäännön (a) mukaan luvut $0/q$ samastuvat lukuun 0:

$$0 \ = \ \frac{0}{1} \ = \ \frac{0}{2} \ = \ \ldots$$

Toisaalta muita luvun 0 esiintymismuotoja ei $\mathbb{Q}$:ssa ole, sillä jos $x = p/q = 0$, niin seuraa $(p/q) \cdot q = p = 0 \cdot q = 0$ (aksioomat (L4), (L2), (L9), (L6) ja Lauseen I.1.1 väittämä (c)). Aksiooma (J1) on näin ollen voimassa. Myös aksioomien (J2)–(J4) voimassaolo on helposti pääteltävissä (Harj.teht. 8).

Järjestys ja samastus ovat esimerkkejä (joukko–opillisista) *relaatioista*, joissa on kyse joukon kahden alkion välisestä 'suhteesta'. Järjestykseen / samastukseen liittyviä relaatiosymboleja ovat mainittujen lisäksi myös '$\neq$', '$\leq$' ja '$\geq$', jotka määritellään

$$\begin{aligned} x \neq y \quad &\Leftrightarrow \quad \text{ei päde } \ x = y, \\ x \leq y \quad &\Leftrightarrow \quad x < y \ \text{ tai } \ x = y, \\ x \geq y \quad &\Leftrightarrow \quad x > y \ \text{ tai } \ x = y. \end{aligned}$$

Merkintä $P \Leftrightarrow Q$ voidaan tässä lukea 'P tarkoittaa samaa kuin Q'.

Summa- ja tulomerkinnät

Laskettaessa peräkkäin yhteen tai kerrottaessa useita erilaisia lukuja on luvut kätevää numeroida eli *indeksoida*, jolloin voidaan käyttää lyhennysmerkintöjä. Nämä ovat

$$\begin{aligned} &\textit{Summamerkintä:} && x_1 + x_2 + \ \ldots \ + x_n \ = \sum_{i=1}^{n} x_i \\ &\textit{Tulomerkintä:} && x_1 \cdot x_2 \cdot \ \ldots \ \cdot x_n \ = \prod_{i=1}^{n} x_i \end{aligned}$$

Summamerkinnässä symboli i on nimeltään *summausindeksi*. Tämä saa yleensä (ei aina) kokonaislukuarvoja, jolloin symboli valitaan tavallisimmin joukosta $\{i, j, k, l, m, n\}$. Koska kyse on vain lyhennysmerkinnästä, ei symbolin valinta luonnollisesti vaikuta itse laskutoimitukseen:

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j = \sum_{k=1}^{n} a_k = \ldots$$

Peräkkäisissä tuloissa sanotaan kerrottavia lukuja tulon *tekijöiksi* (engl. factor). Kokonaislukujen peräkkäisiin tuloihin liittyy termi *n-kertoma* (engl. n-factorial), joka merkitään $n!$ ja määritellään

$$0! = 1, \quad 1! = 1, \quad n! \ = \ 1 \cdot 2 \ \ldots \ \cdot n \ = \ \prod_{k=1}^{n} k, \quad n = 2, 3, \ \ldots$$

HARJOITUSTEHTÄVIÄ

1. Näytä, että seuraaville joukoille pätee $A = B$, eli A ja B ovat joukkoina samat:

   a) $A = \{3n + 2 \mid n \in \mathbb{Z}\}, \quad B = \{3n - 7 \mid n \in \mathbb{Z}\}$
   b) $A = \{7n + 3 \mid n \in \mathbb{Z}\}, \quad B = \{7n - 32 \mid n \in \mathbb{Z}\}$

2. Vedoten aksioomiin (L1)–(L8) ja Lauseen I.1.1 väittämiin perustele kokonaislukujen laskusäännöt a) $-x + (-y) = -(x + y)$, b) $(-x) \cdot y = -x \cdot y$, c) $(-x) \cdot (-y) = x \cdot y$.

3. Muodosta kantalukuja $k = 2$, $k = 3$ ja $k = 4$ vastaavien lukujärjestelmien kertotaulut.

4. Tusinajärjestelmässä on kymmenjärjestelmän merkkien lisäksi käytössä numeromerkit (vasemmalla kymmenjärjestelmän merkintä)

$$\begin{aligned} 10 &= \diamondsuit \quad \text{'ruutu'} \\ 11 &= \heartsuit \quad \text{'hertta'} \\ 12 &= 10 \quad \text{'tusina' (kantaluku)} \end{aligned}$$

Muunna kymmenjärjestelmän luku 155 tusinajärjestelmään ja tusinajärjestelmän luku 9◇07♡ kymmenjärjestelmään.

5. Suorita kymmenjärjestelmään muuntamatta seuraavat tusinajärjestelmän (ks. edellinen tehtävä) laskuoperaatiot:

$$\text{a)}\ \heartsuit\heartsuit\diamondsuit + \diamondsuit\diamondsuit\heartsuit \qquad \text{b)}\ 247 - 19\diamondsuit \qquad \text{c)}\ 23 \cdot 34$$

6. Rusinajärjestelmässä kymmenjärjestelmän luku 7 äännetään 'rusina', luvut $1 \ldots 6$ merkitään ja äännetään kuten kymmenjärjestelmässä, ja rusinajärjestelmän luvut 100 ja 1000 äännetään 'pulla' ja 'kakku'. Miten merkitään ja lausutaan kymmenjärjestelmän luku 2331 rusinajärjestelmässä?

7. Seuraavissa rationaalilukujen laskusäännöissä on tulos esitetty rationaaliluvun perusmuodossa olettaen, että $p, p_1, p_2 \in \mathbb{Z}$ ja $q, q_1, q_2 \in \mathbb{N}$. Perustele säännöt vedoten aksioomiin (L1)–(L9), Lauseeseen I.1.1 ja Esimerkin 7 laskusääntöihin.

   a) Supistus (lavennus): $(p \cdot m)/(q \cdot m) = p/q, \quad m \in \mathbb{Z},\ m \neq 0.$
   b) Vastaluku: $-(p/q) = (-p)/q.$
   c) Käänteisluku: $(p/q)^{-1} = \begin{cases} q/p, & \text{jos } p \in \mathbb{N}, \\ (-q)/(-p), & \text{jos } p \notin \mathbb{N},\ p \neq 0. \end{cases}$
   d) Yhteenlasku: $p_1/q_1 + p_2/q_2 = (p_1q_2 + p_2q_1)/(q_1q_2).$
   e) Kertolasku: $(p_1/q_1)(p_2/q_2) = (p_1p_2)/(q_1q_2).$

8. Vedoten tehtävän 7 laskusääntöihin ja rationaalilukujen järjestysrelaation määritelmään näytä, että järjestysrelaatiolle ovat voimassa aksioomat (J2), (J3) ja (J4).

9. Ratkaise $\mathbb{Q}$:ssa seuraavat yhtälöt / epäyhtälöt mahdollisimman yksinkertaiseen muotoon. Perustele laskun vaiheet aksioomilla (L1)–(L9), (J1)–(J4)!

   a) $2x + 7 = 9x - 4$ b) $\frac{x}{5} + \frac{2}{3} < \frac{x}{3} + \frac{1}{2}$
   c) $(x + 1)/(x - 2) = 4$ d) $(2x - 1)/(3x + 2) > 3$

10. a) Näytä, että seuraavat päättelysäännöt ovat päteviä sekä kokonaisluvuille ($x_k \in \mathbb{Z}$) että rationaaliluvuille ($x_k \in \mathbb{Q}$). Voit vedota sekä aksioomiin että niistä johdettuihin laskusääntöihin (Lause I.1.1, Tehtävä 7).

$$\prod_{k=1}^{n} x_k = 0 \quad \Leftrightarrow \quad x_k = 0 \text{ jollakin } k \in \{1, \ldots, n\}.$$

$$\prod_{k=1}^{n} x_k < 0 \quad \Leftrightarrow \quad \text{pariton määrä lukuja } x_k \text{ on negatiivisia, muut } > 0.$$

Soveltaen a-kohdan päätelmiä ratkaise $\mathbb{Q}$:ssa:

b) $2x^2 - 3x - 2 \geq 0$ c) $63x^2 - 32x \leq 63$ d) $9x^3 \geq 169x$

e) $\dfrac{x+1}{x-1} \geq \dfrac{x}{x+1}$ f) $\dfrac{5}{6x-7} \leq x$ g) $6x + \dfrac{2x+12}{x^2+x} \geq 13$

11. Sievennä (ol. $n \in \mathbb{N}$)

a) $\displaystyle\sum_{k=1}^{n} n$ b) $\displaystyle\sum_{i=1}^{n^2} \frac{2}{n} - \sum_{j=1}^{n} 1 - \sum_{k=1}^{n} 1$ c) $\displaystyle\prod_{k=1}^{n} \left(1 + \frac{1}{k}\right)$

12. Kaksoissumma $\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}$ ($n, m \in \mathbb{N},\ a_{ij} \in \mathbb{Q}$) määritellään

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} = \sum_{i=1}^{n} \Big(\sum_{j=1}^{m} a_{ij} \Big) = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} a_{ij} \Big) = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}\,.$$

a) Näytä, että tässä sulkeiden poisto ja järjestyksen vaihto ovat todella mahdollisia, ts. että kaksi keskimmäistä lauseketta ovat samat.

b) Sievennä: $\displaystyle\sum_{i=1}^{n} \sum_{j=1}^{n} (a_i + a_j)$. c) Näytä: $\displaystyle\Big(\sum_{i=1}^{n} a_i \Big)^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j\,.$

13. (*) Summaa $\sum_{k=0}^{n} a_k$ sanotaan *teleskooppisummaksi*, jos $a_k = b_{k+1} - b_k\ \forall k$, missä luvut $b_k,\ k = 0 \ldots n+1$, ovat tunnettuja.
a) Määritä teleskooppisumman arvo lukujen b_k avulla.
b) Mistä nimitys teleskooppisumma?
c) Laske teleskooppi-idealla summat

$$\sum_{k=0}^{n} (2k+1) \quad \text{ja} \quad \sum_{k=1}^{n} \frac{1}{k(k+1)}\,, \quad n \in \mathbb{N}.$$

d) Näytä teleskooppisummien avulla: $\sum_{k=1}^{n} k^2 = \frac{1}{6} n(2n^2 + 3n + 1),\ \ n \in \mathbb{N}$.

I.2 Kunta

Rationaaliluvut laskuoperaatioineen ovat esimerkki (toistaiseksi ainoa) algebrallisesta rakennelmasta nimeltä *kunta* (engl. field, ruots. kropp, saks. Körper). Jos $\mathbb{K}$ on jokin lukujoukko (voisi olla yleisempikin joukko), niin sanotaan, että $(\mathbb{K}, +, \cdot)$ on kunta, jos ensinnäkin

- Laskuoperaatiot $x, y \mapsto x + y$ ja $x, y \mapsto x \cdot y$ on määritelty yksikäsitteisesti $\forall x, y \in \mathbb{K}$ ja pätee $x + y \in \mathbb{K}$ ja $x \cdot y \in \mathbb{K}$,

ja lisäksi seuraavat **kunta-aksioomat** ovat voimassa:

- Vaihdanta-, liitäntä- ja osittelulait

(K1) $x + y = y + x$

(K2) $x \cdot y = y \cdot x$

(K3) $x + (y + z) = (x + y) + z$

(K4) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

(K5) $x \cdot (y + z) = x \cdot y + x \cdot z$

- Nolla–alkion ja vasta–alkion olemassaolo

(K6) On olemassa *nolla–alkio* $0 \in \mathbb{K}$, jolle pätee $x + 0 = x \;\; \forall\, x \in \mathbb{K}$.

(K7) Jokaisella $x \in \mathbb{K}$ on *vasta–alkio* $-x \in \mathbb{K}$, jolle pätee $x + (-x) = 0$.

- Ykkösalkion ja käänteisalkion olemassaolo

(K8) On olemassa *ykkösalkio* $1 \in \mathbb{K}$, jolle pätee $x \cdot 1 = x \;\; \forall x \in \mathbb{K}$.

(K9) Jokaisella $x \in \mathbb{K}$, $x \neq 0$ on *käänteisalkio* $x^{-1} \in \mathbb{K}$, jolle pätee $x \cdot x^{-1} = 1$.

- Erotteluaksiooma

(K10) $0 \neq 1$

Aksioomista erillinen alkuoletus katsotaan yleensä $\mathbb{K}$:n laskuoperaatioiden määritelmään sisältyväksi. Erillisenä tämä oletus on huomioitava lähinnä silloin, kun laskuoperaatiot on alunperin määritelty jossakin suuremmassa joukossa $\mathbb{A} \supset \mathbb{K}$, tai kun tarkasteltavaa joukkoa $\mathbb{K}$ halutaan laajentaa. Tällöin on varmistettava, että laskuoperaatioiden tulos pysyy joukossa $\mathbb{K}$, jota tarkastellaan, ks. esimerkit

jäljempänä. Peruslaskuoperaatioiden $(+,\cdot)$ lisäksi kunnassa voidaan aina määritellä yhdistetyt laskuoperaatiot $x,y \mapsto x-y = x+(-y)$ (vähennyslasku) ja $x,y \mapsto x/y = x\cdot y^{-1}$ (jakolasku, $y \neq 0$), jolloin on $-x = 0-x$ ja $x^{-1} = 1/x$. Koska aksiooma (K10) tarkoittaa 'ei päde $0 = 1$', niin kunnassa tarvitaan yleisesti vain samastusrelaatio ($\mathbb{K}$:n alkioiden erotteluperiaate), ei järjestysrelaatiota. †

Esimerkki 1 Yksinkertaisin mahdollinen kunta saadaan, kun valitaan $\mathbb{K} = \{0,1\}$ (missä $0 \neq 1$) ja sovitaan laskusäännöstä $1+1=0$ (!). Tällöin $-1 = 1$ ja muut laskusäännöt ovat pääteltävissä aksioomista (K1), (K2), (K5), (K6) ja (K8):

$$0+0=0, \quad 0+1=1+0=1, \quad 1\cdot 1=1, \quad 0\cdot 1 = 1\cdot 0 = 0,$$
$$0\cdot 0 = 0\cdot(1+1) = 0\cdot 1 + 0\cdot 1 = 0.$$

Näillä säännöillä aksioomat (K1)-(K9) ovat kaikki voimassa (samoin oletus (K0)), joten kyseessä on kunta. □

Lause I.2.1 Jokaisessa kunnassa $(\mathbb{K},+,\cdot)$ pätee

(a) Nolla–alkio, vasta–alkio, ykkösalkio ja käänteisalkio ovat yksikäsitteisiä.

(b) $-(-x) = x \quad \forall x \in \mathbb{K}, \qquad (x^{-1})^{-1} = x \quad \forall x \in \mathbb{K},\ x \neq 0.$

(c) $0\cdot x = 0 \quad \forall x \in \mathbb{K}.$

(d) $-x = (-1)\cdot x \quad \forall x \in \mathbb{K}.$

<u>Todistus</u> (a) Oletetaan, että $0 \in \mathbb{K}$ ja $\theta \in \mathbb{K}$ ovat kaksi nolla–alkiota. Tällöin aksiooman (K6) mukaan on

$$x = x+0 \quad \forall x \in \mathbb{K}, \qquad y+\theta = y \quad \forall y \in \mathbb{K}.$$

Kun valitaan $x = \theta$ ja $y = 0$ ja käytetään aksioomaa (K1), seuraa

$$\theta = \theta + 0 = 0 + \theta = 0,$$

eli $\theta = 0$. Siis nolla–alkio on yksikäsitteinen. Vasta–alkion yksikäsitteisyyden toteamiseksi oletetaan, että $a \in \mathbb{K}$ ja $b \in \mathbb{K}$ ovat saman alkion $x \in \mathbb{K}$ vasta–alkioita. Tällöin aksioomien (K6), (K3) ja (K1) perusteella on

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

†Samastusrelaatiolta '=' edellytetään aina aksioomat (S1) $x = x\ \forall x$, (S2) $x = y \Rightarrow y = x$ ja (S3) $x = y\ \&\ y = z \Rightarrow x = z$. Kunnan samastusrelaatiolta vaaditaan lisäksi yhteensopivuus laskuoperaatioiden kanssa siten, että laskuoperaation tulos on aina yksikäsitteinen. Vaihdantalakien (K1)–(K2) ja aksiooman (S3) perusteella tämä vaatimus toteutuu olettamalla:
(SK1) $x = y \Rightarrow x+z = y+z\ \forall z$ ja (SK2) $x = y \Rightarrow x\cdot z = y\cdot z\ \forall z$.

eli $a = b$. Siis vasta–alkiokin on yksikäsitteinen. Muut väitetyt yksikäsitteisyydet seuraavat samanlaisella päättelyllä.

(b) Väittämän ensimmäinen osa todistettiin edellisessä luvussa (Lause I.1.1(b)). Toisen osan todistamiseksi sovelletaan aksioomia (K2) ja (K9):

$$x^{-1} \cdot x = x \cdot x^{-1} = 1 \quad \Rightarrow \quad (x^{-1})^{-1} = x.$$

(c) Kun merkitään $0 \cdot x = y \in \mathbb{K}$, niin aksioomien (K6), (K2) ja (K5) perusteella

$$y = 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x = y + y. \tag{$\star$}$$

Käyttäen tätä tulosta ja kunta–aksioomia päätellään:

$$\begin{aligned} 0 &= y + (-y) && \text{(K7)} \\ &= (y + y) + (-y) && (\star) \\ &= y + [y + (-y)] && \text{(K3)} \\ &= y + 0 && \text{(K7)} \\ &= y. && \text{(K6)} \end{aligned}$$

(d) Tuloksen (c) ja kunta–aksioomien perusteella

$$\begin{aligned} x + (-1) \cdot x &= x \cdot 1 + x \cdot (-1) && \text{(K8), (K2)} \\ &= x \cdot [1 + (-1)] && \text{(K5)} \\ &= x \cdot 0 && \text{(K7)} \\ &= 0. \quad \square && \text{(K2), (c)} \end{aligned}$$

Yleisessä kunnassa voidaan potenssiin korotus määritellä samalla tavoin kuin rationaalilukujen kunnassa, vrt. edellinen luku.

Esimerkki 2 Johda binomikaavat lausekkeille $(x + y)^2$ ja $(x + y)^3$ sellaisessa kunnassa $(\mathbb{K}, +, \cdot)$, jossa pätee $1 + 1 = \spadesuit$, $\spadesuit + 1 = \clubsuit$.

<u>Ratkaisu</u> Potenssiin korotuksen määritelmän ja kunta-aksioomien perusteella

$$\begin{aligned} (x + y)^2 \;=\; (x + y) \cdot (x + y) \;&=\; x \cdot x + x \cdot y + y \cdot x + y \cdot y \\ &=\; x^2 + (x \cdot y + x \cdot y) + y^2. \end{aligned}$$

Tässä on edelleen kunta–aksioomien ja oletuksen perusteella

$$x \cdot y + x \cdot y \;=\; (x \cdot y) \cdot (1 + 1) \;=\; (x \cdot y) \cdot \spadesuit \;=\; \spadesuit \cdot x \cdot y,$$

joten pyydetty ensimmäinen binomikaava on

$$(x+y)^2 \;=\; x^2 + \spadesuit \cdot x \cdot y + y^2.$$

Vastaavalla päättelyllä saadaan toiseksi kaavaksi

$$(x+y)^3 \;=\; x^3 + \clubsuit \cdot x^2 \cdot y + \clubsuit \cdot x \cdot y^2 + y^3. \quad \square$$

Seuraava kunta–algebran tulos osoittautuu jatkossa hyödylliseksi (Harj.teht. 3a).

Propositio I.2.2 Jokaisessa kunnassa pätee

$$\begin{aligned} x^n - y^n &= (x-y)\left(x^{n-1} + x^{n-2}y + \ldots + y^{n-1}\right) \\ &= (x-y)\sum_{k=0}^{n-1} x^{n-1-k}y^k, \quad n \in \mathbb{N}. \end{aligned}$$

Järjestetty kunta

Jos kunnassa on määritelty järjestysrelaatio edellisen luvun aksioomien (J1)-(J4) mukaisesti, niin sanotaan, että kyseessä on *järjestetty kunta* (engl. ordered field). Järjestetyssä kunnassa jokainen alkio on aksiooman (J1) mukaisesti joko positiivinen ($x > 0$), negatiivinen ($x < 0$), tai $= 0$. Toistaiseksi ainoa esimerkki järjestetystä kunnasta on rationaalilukujen kunta.

Lause I.2.1 (jatko) Jokaisessa järjestetyssä kunnassa $(\mathbb{K}, +, \cdot, <)$ pätee

(e) $x > 0 \;\Rightarrow\; -x < 0, \qquad x > 0 \;\Rightarrow\; x^{-1} > 0.$

(f) $x > 0 \;\&\; y > 0 \;\Rightarrow\; x + y > 0.$

(g) $0 < 1.$

<u>Todistus</u> (e) Oletetaan, että $0 < x$. Tällöin on aksiooman (J3) mukaan myös

$$0 + (-x) \;<\; x + (-x),$$

mikä kunta-aksioomien mukaan pelkistyy ensimmäiseksi väittämäksi $-x < 0$. Väittämän toisen osan todistamiseksi suljetaan pois aksiooman (J1) jättämät muut vaihtoehdot. Jos $x^{-1} = 0$, niin aksioomien (K8), (K2) ja tuloksen (c) mukaan on $1 = x \cdot x^{-1} = x \cdot 0 = 0$ eli $0 = 1$, mikä on ristiriidassa aksiooman (K10) kanssa. Siis mahdollisuus $x^{-1} = 0$ on pois suljettu. Jos $x^{-1} < 0$, niin tuloksen (d)

ja väittämän (e) jo todistetun ensimmäisen osan mukaan on $(-1) \cdot x^{-1} > 0$. Tällöin aksiooman (J4) ja oletuksen $x > 0$ mukaan on myös $(-1) \cdot x^{-1} \cdot x > 0$, mikä sievenee kunta-aksioomien ja mainittujenen tulosten perusteella muotoon $1 < 0$. Tämäkin on mahdotonta aksiooman (J1) ja vielä todistamatta olevan väittämän (g) mukaan, joten sikäli kuin (g) on tosi, jää ainoaksi vaihtoehdoksi $x^{-1} > 0$.

(f) Päätellään ensin aksioomien (J3), (K1) ja (K6) perusteella:

$$0 < y \quad \Rightarrow \quad x < x + y.$$

Jatketaan tästä soveltaen aksioomaa (J3):

$$0 < x \;\; \& \;\; x < x + y \quad \Rightarrow \quad 0 < x + y.$$

(g) Aksioomien (J1) ja (K10) mukaan on oltava joko $0 < 1$ tai $1 < 0$. Jos oletetaan jälkimmäinen vaihtoehto, niin aksiooman (J3) mukaan on siinä tapauksessa

$$1 + (-1) < 0 + (-1),$$

mikä sievenee aksioomien (K7), (K1) ja (K6) perusteella muotoon

$$0 < -1.$$

Tällöin on aksiooman (J4) perusteella oltava myös

$$0 < (-1) \cdot (-1).$$

Mutta väittämien (b), (d) perusteella on $1 = -(-1) = (-1) \cdot (-1)$, joten on päätelty, että oletetussa vaihtoehdossa $1 < 0$ pätee myös $0 < 1$. Aksiooman (J1) mukaan tämä on kuitenkin mahdotonta, joten ainoaksi vaihtoehdoksi jää $0 < 1$. Päättelyssä ei tarvittu vielä avoimena olevaa väittämää (e), joten myös tämän väittämän todistus tuli samalla loppuun viedyksi. $\square$

Järjestetyssä kunnassa voidaan jokaiseen kunnan alkioon liittää 'itseisalkio', eli ko. alkion *itseisarvo* (engl. absolute value) määritelmällä

$$|x| \;=\; \max\{x, -x\} \;=\; \begin{cases} x, & \text{jos } x \geq 0, \\ -x, & \text{jos } x < 0. \end{cases}$$

Itseisarvon avulla on edelleen määriteltävissä kahden luvun välinen *etäisyys*

$$d(x, y) \;=\; |x - y|,$$

jolloin sellaiset sanonnat kuin 'y on lähempänä x:ää kuin z' saavat (algebrallisen) sisällön. Seuraavaa itseisarvoon liittyvää tulosta tarvitaan matemaattisessa analyysissä hyvin usein.

LAUSE I.2.3 (**Kolmioepäyhtälö**) Järjestetyssä kunnassa $(\mathbb{K}, +, \cdot, <)$ pätee

$$\boxed{\,|\,|x| - |y|\,| \;\leq\; |x+y| \;\leq\; |x| + |y|, \quad x, y \in \mathbb{K}.\,}$$

Todistus Päätulos on epäyhtälöistä jälkimmäinen, sillä edellinen seuraa kunta–aksioomista, jälkimmäisestä epäyhtälöstä ja itseisarvon määritelmästä päättelyllä

$$\begin{cases} |x| \;=\; |\,(x+y) + (-y)\,| \;\leq\; |x+y| + |-y| \;=\; |x+y| + |y| \\ |y| \;=\; |\,(x+y) + (-x)\,| \;\leq\; |x+y| + |-x| \;=\; |x+y| + |x| \end{cases}$$
$$\Rightarrow \quad \pm\,(|x| - |y|) \;\leq\; |x+y| \quad \Rightarrow \quad ||x| - |y|| \;\leq\; |x+y|.$$

Kolmioepäyhtälön jälkimmäisessä osassa väitetään itseisarvon määritelmän perusteella, että

$$\pm(x+y) \;\leq\; |x| + |y|.$$

Koska saman määritelmän mukaan on

$$\pm x \leq |x| = a, \quad \pm y \leq |y| = b,$$

niin nähdään, että väitetty epäyhtälö seuraa väittämästä

LEMMA I.2.4 Järjestetyssä kunnassa pätee

$$x \leq a \quad \& \quad y \leq b \quad \Rightarrow \quad x + y \leq a + b.$$

Todistus Väittämässä on neljä vaihtoehtoa: a) $x < a,\ y < b$, b) $x < a,\ y = b$, c) $x = a,\ y < b$, d) $x = a,\ y = b$. Vaihtoehdossa a) on $x < a \;\Leftrightarrow\; a - x > 0$ ja $y < b \;\Leftrightarrow\; b - y > 0$, jolloin Lauseen I.2.1 väittämästä (f) ja kunta-aksioomista seuraa $(a - x) + (b - y) > 0 \;\Leftrightarrow\; x + y < a + b$. Myös vaihtoehdoissa b) ja c) väite toteutuu tässä muodossa ja vaihtoehdossa d) muodossa '=', kuten nähdään helposti. $\square$

Alikunta ja kuntalaajennus

Sanotaan, että kunta $(\mathbb{A}, +, \cdot)$ on kunnan $(\mathbb{K}, +, \cdot)$ *alikunta* (engl. subfield), jos $\mathbb{A} \subset \mathbb{K}$ ja kunnilla on samat laskuoperaatiot, ts. kunnan $(\mathbb{A}, +, \cdot)$ laskuoperaatiot = kunnan $(\mathbb{K}, +, \cdot)$ operaatiot osajoukkoon $\mathbb{A}$ rajoitettuina. Tällöin kunnilla on myös yhteiset nolla- ja ykkösalkiot, eli kunnan $(\mathbb{K}, +, \cdot)$ nolla- ja ykkösalkioille pätee $0, 1 \in \mathbb{A}$. Nimittäin jos $\mathbb{K}$:n nolla–alkio on $0 \in \mathbb{K}$ ja $\mathbb{A}$:n nolla–alkio on $\theta \in \mathbb{A}$, niin jokaisella $x \in \mathbb{A}$ (esim. $x = \theta$) voidaan päätellä

$$x = x + \theta \quad \Rightarrow \quad -x + x = -x + x + \theta \quad \Rightarrow \quad 0 = 0 + \theta = \theta + 0 = \theta.$$

Tässä $-x \in \mathbb{K}$ on x:n vasta–alkio $\mathbb{K}$:ssa, jolloin päättely perustui ensin kunnan $(\mathbb{A}, +, \cdot)$ aksioomaan (K6) ja sen jälkeen kunnan $(\mathbb{K}, +, \cdot)$ aksioomiin (K1), (K3), (K6) ja (K7). Vastaavaan tapaan nähdään, että ykkösalkio on kunnissa sama.

Jos kunta $(\mathbb{A}, +, \cdot)$ on kunnan $(\mathbb{K}, +, \cdot)$ alikunta, sanotaan vastaavasti, että jälkimmäinen on edellisen *kuntalaajennus*. Myöhemmissä luvuissa tehdään useita lukualueen laajennuksia tyyppiä $\mathbb{Q} \hookrightarrow \mathbb{K}$. Näille on yhteistä, että laajennuksissa syntyy uusia kuntia, joiden kaikkien yhteinen alikunta on $(\mathbb{Q}, +, \cdot)$. Seuraavassa hyvin varovainen esimerkki tällaisesta laajennuksesta.

Esimerkki 3 Sovitaan, että on olemassa luku a jolla on ominaisuus

$$a^2 = a \cdot a = 2.$$

(Pidetään tunnettuna, että $a \notin \mathbb{Q}$.) Luvun a ja rationaalilukujen välisistä laskuoperaatioista sovitaan, että $1 \cdot a = a$ ja $0 \cdot a = 0 \in \mathbb{Q}$ ja lisäksi sovitaan, että vaihdanta-, liitäntä- ja osittelulait (K1)–(K5) ovat voimassa myös, kun luku a on operaatioissa mukana. Näiden sopimusten nojalla voidaan muodostaa lukujoukko

$$\mathbb{K} = \{\, z = x + ya \mid x, y \in \mathbb{Q} \,\}.$$

Tämä on $\mathbb{Q}$:n aito laajennus, sillä $x \in \mathbb{Q} \Rightarrow x = x + 0a \in \mathbb{K}$, mutta $0 + 1a = a \notin \mathbb{Q}$. Tehtyjen sopimusten mukaan on erityisesti $0 + 0a = 0$. Luvulla 0 ei ole $\mathbb{K}$:ssa muita esitysmuotoja, sillä jos $x + ya = 0$, niin on oltava $y = 0 \Rightarrow x = 0$, koska muuten olisi $a = -x/y \in \mathbb{Q}$. (Yleisemmin voidaan päätellä, että jokaisen luvun $z \in \mathbb{K}$ esitysmuoto $z = x + ya$ on yksikäsitteinen.)

Jos $z_1 = x_1 + y_1 a \in \mathbb{K}$ ja $z_2 = x_2 + y_2 a \in \mathbb{K}$ $(x_1, x_2, y_1, y_2 \in \mathbb{Q})$, niin oletettujen aksioomien (K1)–(K5) ja laskusäännön $a^2 = 2$ nojalla

$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2) + (y_1 + y_2)a, \\ z_1 z_2 &= (x_1 x_2 + y_1 y_2 a^2) + (x_1 y_2 + x_2 y_1)a = (x_1 x_2 + 2y_2 y_2) + (x_1 y_2 + x_2 y_1)a, \end{aligned}$$

joten $z_1 + z_2 \in \mathbb{K}$ ja $z_1 z_2 \in \mathbb{K}$. Laskuoperaatioita koskeva perusoletus (ensimmäinen oletus edellä) on siis voimassa. Edelleen nähdään, että kunta–aksioomista ovat oletettujen lisäksi voimassa myös (K6), (K8) ja (K10) $(0, 1 \in \mathbb{Q})$ ja (K7) (jos $z = x + ya$, niin $-z = -x - ya$). Lopuksi päätellään, että myös (K9) on voimassa. Nimittäin $zz^{-1} = 1$, kun määritellään

$$z^{-1} = (x^2 - 2y^2)^{-1}(x - ya) = \frac{x}{x^2 - 2y^2} - \frac{y}{x^2 - 2y^2}\, a.$$

Tässä on $x^2 - 2y^2 = 0$ vain kun $x = y = 0$ (koska $(x/y)^2 \neq 2$ kun $x, y \in \mathbb{Q}$ ja $y \neq 0$), joten jokaisella $z \in \mathbb{K}$, $z \neq 0$ on käänteisluku z^{-1}. On päätelty, että $(\mathbb{K}, +, \cdot)$ on kunta ja siis rationaalilukujen kunnan $(Q, +, \cdot)$ aito laajennus.

Kunta $(\mathbb{K}, +, \cdot)$ voidaan myös järjestää. Sovitaan ensinnäkin, että $a > 0$ (mahdollista, koska $a^2 = (-a)^2$). Yleisemmin nähdään, että vertailu $z_1 < z_2$ $(z_1, z_2 \in \mathbb{K})$ palautuu kunta–algebran avulla aina viime kädessä luvun a ja rationaaliluvun vertailuksi. (Esim. $7 - 2a > 1 - a \;\Leftrightarrow\; 6 - 3a > 0 \;\Leftrightarrow\; 3/2 > a$.) Jos $x \in \mathbb{Q}$, niin ilmeisesti on $x < a$ aina kun $x \leq 0$. Jos taas on $x > 0$, niin on myös $x + a > 0 \;\Leftrightarrow\; (x + a)^{-1} > 0$, jolloin vertailukysymys $x \,?\, a$ ratkeaa käyttämällä järjestysrelaation aksioomaa (J4):

$$x > a \quad \Leftrightarrow \quad x - a = \frac{x^2 - a^2}{x + a} > 0 \quad \Leftrightarrow \quad x^2 - a^2 > 0 \quad \Leftrightarrow \quad x^2 > 2.$$

Järjestetyssä kunnassa $(\mathbb{K}, +, \cdot, <)$ luku a sijoittuu siis rationaalilukujen 'väliin': Jokaisen luvun $x \in \mathbb{Q}$ kohdalla voidaan ratkaista, onko $a < x$ vai $a > x$, ja tämä vertailu perustuu vain rationaalilukujen $0, x, 2$ ja x^2 vertailuun. $\square$

Juuriluvut ja murtopotenssit

Esimerkin 3 kuntaa huomattavasti käyttökelpoisempi kuntalaajennus saadaan aikaan, kun rationaalilukujen joukko laajennetaan lukujoukoksi, jossa kuntaoperaatioiden $x, y \mapsto x + y$, $x - y$, xy, x/y lisäksi sallitaan laskuoperaatiot

$$x \mapsto \sqrt[m]{x}, \quad x > 0, \quad m \in \mathbb{N},\ m \geq 2.$$

Tässä luvulla $\sqrt[m]{x}$ tarkoitetaan *juurilukua* y, joka toteuttaa ehdot

$$y^m = x \quad \& \quad y > 0.$$

Sopimuksella $\sqrt[1]{x} = x$ tulee juuriluku $\sqrt[m]{x}$ määritellyksi $\forall m \in \mathbb{N}$. Sanotaan, että $\sqrt[m]{x}$ on luvun x *neliöjuuri*, jos $m = 2$ (lyhennysmerkintä $\sqrt{x}$), *kuutiojuuri*, jos $m = 3$, tai yleisemmin *m:s juuri*, luetaan 'm:s juuri x'. Tällainen luku voidaan yksinkertaisesti sopia 'olemassa olevaksi' samaan tapaan kuin meneteltiin Esimerkissä 3 luvun a $(= \sqrt{2})$ kohdalla. Lisäehto $y > 0$ (joka jo viittaa järjestysrelaatioon) tarvitaan takaamaan luvun y yksikäsitteisyys, sillä m:n ollessa parillinen on $(-y)^m = y^m$. Yksikäsitteisyys mainitulla lisäehdolla seuraa identiteetistä (ks. Propositio I.2.2)

$$(y_1 - y_2)(y_1^{m-1} + y_1^{m-2}y_2 + \ldots + y_2^{m-1}) = y_1^m - y_2^m.$$

Jos tässä oletetetaan, että $y_1^m = y_2^m = x$ ja $y_1, y_2 > 0$, niin oikea puoli $= 0$ ja vasemmalla puolella on jälkimmäinen tekijä > 0, joten on oltava $y_1 - y_2 = 0$.

Laskuoperaatioilla $x \mapsto \sqrt[m]{x}$ laajennettua lukujoukkoa, lähtökohtana rationaaliluvut, merkittäköön symbolilla $\mathbb{J}$. Joukko $\mathbb{J}$ määritellään yksinkertaisesti koostuvaksi luvuista, jotka saadaan äärellisellä määrällä kuntaoperaatioita ja operaatioita $x \mapsto \sqrt[m]{x}$ lähtien luvuista $0, 1 \in \mathbb{Z}$. Tällöin on ilmeistä, että jos $x, y \in \mathbb{J}$,

niin myös $x \pm y$, xy, $x/y \in \mathbb{J}$ $(y \neq 0)$, joten laskuoperaatioita koskeva perusoletus on voimassa. Lukujen $x \in \mathbb{J}$ 'ulkonäkö' voi kylläkin olla konstikas.

Esimerkki 4

$$\left(3+\sqrt{2}-\sqrt[8]{\frac{\sqrt[6]{5-\sqrt[4]{3}}}{\sqrt[4]{3+\sqrt[3]{2}}}+\frac{1}{\sqrt[7]{7}}}\right) \Big/ \left(2+\sqrt{3}+\sqrt[8]{\frac{\sqrt[6]{3-\sqrt[4]{5}}}{\sqrt[4]{2+\sqrt[3]{3}}}+\sqrt[7]{7}}\right) \in \mathbb{J}. \quad \square$$

Juuriluvuilla laskettaessa oletetaan kunta–algebran laskusäännöt päteviksi, ts. oletetaan, että $(\mathbb{J}, +, \cdot)$ on kunta. Tästä oletuksesta sekä juuriluvun määritelmästä voidaan johtaa seuraavat yleiset laskusäännöt (ol. $x, y > 0,\ m, n \in \mathbb{N}$):

$$\text{a)}\ \ \sqrt[m]{x} = \sqrt[mn]{x^n}. \qquad \text{b)}\ \ \sqrt[m]{x}\,\sqrt[m]{y} = \sqrt[m]{xy}. \qquad \text{c)}\ \ \left(\sqrt[m]{x}\right)^{-1} = \sqrt[m]{x^{-1}}.$$

Sääntöjen perustelemiseksi merkitään $a = \sqrt[m]{x}$, $b = \sqrt[m]{y}$, jolloin on $a, b > 0 \ \Rightarrow\ ab, a^{-1} > 0$. Juuriluvun määritelmään ja kunta–algebraan vedoten voidaan tällöin päätellä:

$$\text{a)}\ \ a^m = x \ \Rightarrow\ (a^m)^n = a^{mn} = x^n \ \Rightarrow\ a = \sqrt[mn]{x^n}.$$

$$\text{b)}\ \ a^m = x\ \ \&\ \ b^m = y \ \Rightarrow\ a^m b^m = (ab)^m = xy \ \Rightarrow\ ab = \sqrt[m]{xy}.$$

$$\text{c)}\ \ a^m = x \ \Rightarrow\ (a^m)^{-1} = (a^{-1})^m = x^{-1} \ \Rightarrow\ a^{-1} = \sqrt[m]{x^{-1}}.$$

Juuriluvuilla laskemisen säännöt saadaan kätevästi yhdistetyksi potenssien laskusääntöihin (vrt. edellinen luku), kun määritellään luvun $x > 0$ *murtopotenssi* asettamalla

$$x^{p/q} = \sqrt[q]{x^p}, \quad p \in \mathbb{Z},\ q \in \mathbb{N},\ q \geq 2.$$

Tämä määrittelee yksiselitteisesti luvun x^r jokaisella $r \in \mathbb{Q}$, sillä jos r esitetään perusmuodossa $r = p/q$, $q \in \mathbb{N}$, niin määritelmän ja säännön a) perusteella on $x^{(pn)/(qn)} = \sqrt[qn]{x^{pn}} = \sqrt[nq]{(x^p)^n} = \sqrt[q]{x^p} = x^{p/q}$ $\forall n \in \mathbb{N}$, eli x^r ei riipu r:n esitysmuodosta. Säännöistä a)–c) voidaan nyt johtaa murtopotensseille samat laskusäännöt kuin kokonaislukupotensseille (Harj.teht. 8):

$$\boxed{x^r x^s = x^{r+s}, \quad x^r y^r = (xy)^r, \quad (x^r)^s = x^{rs}, \quad x, y > 0,\ r, s \in \mathbb{Q}.}$$

Luvut $x \in \mathbb{J}$, $x \notin \mathbb{Q}$ on myös mahdollista sijoitella rationaalilukujen 'väleihin' niin, että syntyy järjestetty kunta $(\mathbb{J}, +, \cdot, <)$. Järjestysrelaation määrittely yleisessä tapauksessa vaatisi kuitenkin tähänastista vakavampia lukuteoreettisia pohdiskeluja, siksi asiaan ei toistaiseksi puututa. Todettakoon ainoastaan, että suotuisissa erikoistapauksissa vertailu on mahdollista palauttaa suoraan rationaalilukujen vertailuksi samalla tavoin kuin Esimerkissä 3 edellä. Ideana on tällöin purkaa juurilausekkeet päättelyllä: Jos $x, y > 0$ ja $m \in \mathbb{N}$, niin $x < y \ \Leftrightarrow\ x^m < y^m$ (Harj.teht. 6).

ESIMERKKI 5 Kumpi on suurempi, $x = \sqrt{\sqrt{2}+5}$ vai $y = \sqrt[4]{41}$?

Ratkaisu Koska $x < y \;\Leftrightarrow\; x^4 < y^4$ $(x, y > 0)$, niin päätellään

$$x < y \quad \Leftrightarrow \quad (\sqrt{2}+5)^2 < 41 \quad \Leftrightarrow \quad 10\sqrt{2} + 27 < 41 \quad \Leftrightarrow \quad 10\sqrt{2} < 14.$$

Jatkamalla tästä päättelyllä $a < b \;\Leftrightarrow\; a^2 < b^2$ $(a, b > 0)$ nähdään, että

$$10\sqrt{2} < 14 \quad \Leftrightarrow \quad 200 < 196.$$

Koska viimeinen epäyhtälö ei toteudu, vaihdetaan epäyhtälöiden suunnat ja päätellään: $x > y \;\Leftrightarrow\; 200 > 196$. Siis x on suurempi. $\square$

HARJOITUSTEHTÄVIÄ

1. Lukujoukossa $\mathbb{K}$ on vain luvut 0,1 ja $\diamondsuit = 1 + 1$ (kaikki keskenään eri suuria). Määrittele (jos mahdollista) $\mathbb{K}$:n laskusäännöt (yhteenlasku, kertolasku, vastaluvut, käänteisluvut) siten, että $(\mathbb{K}, +, \cdot)$ on kunta.

2. Vedoten kunta–aksioomiin tai Lauseen I.2.1 väittämiin näytä, että jokaisessa kunnassa pätee: a) $x \cdot y = 0 \;\Leftrightarrow\; x = 0 \;\vee\; y = 0$ (*tulon nollasääntö*), b) $(-x) \cdot (y) = -(x \cdot y)$, c) $(-x) \cdot (-y) = x \cdot y$, d) 0:lla ei ole käänteisalkiota.

3. Näytä, että jokaisessa kunnassa pätee
   a) $x^n - y^n = (x - y) \sum_{k=0}^{n-1} x^{n-1-k} y^k, \;\; n \in \mathbb{N}$
   b) $x^n + y^n = (x + y) \sum_{k=0}^{n-1} (-1)^k x^{n-1-k} y^k,$ jos $n \in \mathbb{N}$ ja n on pariton
   c) $1 + x + \ldots + x^n = (x^{n+1} - 1)(x - 1)^{-1}, \;\; n \in \mathbb{N}, \; x \neq 1$

4. Näytä, että järjestetyn kunnan aksiooma (J3) voidaan korvata aksioomalla (J3') $x < y \;\Leftrightarrow\; x - y < 0$.

5. Näytä, että jokaisessa järjestetyssä kunnassa pätee:
   a) $x > 1 \;\Rightarrow\; 0 < x^{-1} < 1$
   b) $x > 0 \;\&\; y > 1 \;\Rightarrow\; 0 < x/y < x$
   c) $|x \cdot y| = |x| \cdot |y|$
   d) $|x^{-1}| = |x|^{-1}$
   e) $x^2 < y^2 \;\Leftrightarrow\; |x| < |y|$
   f) $0 \leq x \leq a \;\&\; 0 \leq y \leq b \;\Rightarrow\; 0 \leq xy \leq ab$

6. Lähtien tehtävän 3 kaavasta a) näytä, että järjestetyssä kunnassa pätee: Jos $x > 0$, $y > 0$ ja $n \in \mathbb{N}, \; n \geq 2$, niin $x < y \;\Leftrightarrow\; x^n < y^n$.

7. Aseta luvut a, 17/12 ja 72/51 suuruusjärjestykseen Esimerkin 3 kunnassa. Suorita vertailut tarkasti!

8. Perustele murtopotenssien laskusäännöt a) $x^r x^s = x^{r+s}$, b) $x^r y^r = (xy)^r$, c) $(x^r)^s = x^{rs}$ $(x, y > 0,\ r, s \in \mathbb{Q})$.

9. Selvitä ilman laskinta lukujen $x, y \in \mathbb{J}$ suuruusjärjestys palauttamalla vertailu rationaalilukujen vertailuksi:
   a) $x = \sqrt[3]{1020}$, $y = \sqrt{102}$, b) $x = \sqrt[3]{3}$, $y = \sqrt[4]{43/10}$,
   c) $x = 2 - \sqrt{3}$, $y = 1/\sqrt{4\sqrt{3} + 7}$.

10. Olkoon $(\mathbb{K}, +, \cdot, <)$ järjestetty kunta. Näytä: a) $x + 1 > x\ \forall x \in \mathbb{K}$.
    b) Joukossa $\mathbb{K}$ on äärettömän monta eri alkiota.

11. (*) Näytä käsinlaskulla, että pätee
$$\text{a)}\ \frac{577}{408} - \frac{1}{400000} < \sqrt{2} < \frac{577}{408}, \qquad \text{b)}\ \sqrt{2} < \frac{665857}{470832}.$$

12. (*) Näytä, että järjestetyn kunnan aksiooma (J2) voidaan korvata aksioomalla (J2') $x > 0\ \&\ y > 0 \Rightarrow x + y > 0$.

13. (*) Halutaan määritelllä pienin mahdollinen rationaalilukujen kunnan laajennus $(\mathbb{K}, +, \cdot)$ siten, että $\sqrt{2} \in \mathbb{K}$ ja $\sqrt{3} \in \mathbb{K}$. Näytä, että a) $\mathbb{K}$ koostuu luvuista muotoa $x + y\sqrt{2} + z\sqrt{3} + u\sqrt{6}$, $x, y, z, u \in \mathbb{Q}$, b) kunnassa $(\mathbb{K}, +, \cdot)$ on määriteltävissä järjestysrelaatio, joka perustuu vain rationaalilukujen vertailuun.

14. (*) (Big Ben) Tarkastellaan joukkoa $\mathbb{K} =$\{kellon viisarit\}. Jokainen viisari $v \in \mathbb{K}$ on *lukupari* (r, θ), missä $r =$ viisarin pituus (yksikkö m) ja $\theta =$ viisarin suunta asteina, mitattuna klo 12:sta positiivisena myötäpäivään tai negatiivisena vastapäivään. Sovitaan, että $(r_1, \theta_1) = (r_2, \theta_2)$, jos joko (i) $r_1 = r_2 = 0$ tai (ii) $r_1 = r_2 > 0$ ja $\theta_1 - \theta_2 = k \cdot 360$, $k \in \mathbb{Z}$. Jos $r = 0$, sanotaan viisaria (r, θ) *nollaviisariksi*, merkitään 0_v. Määritellään viisareiden kertolasku seuraavasti:
$$(r_1, \theta_1) \cdot (r_2, \theta_2) = (r_1 r_2,\ \theta_1 + \theta_2).$$
    a) Näytä, että kertolaskulle pätee sekä vaihdantalaki että liitäntälaki.
    b) Määrittele ykkösviisari sekä viisarin $(r, \theta) \neq 0_v$ käänteisviisari.
    c) Big Ben, jonka minuuttiviisarin pituus on 4 ja tuntiviisarin pituus on 2, näyttää aikaa noin klo 3. Mikä kellonaika on tarkemmin kyseessä (sekunnin tarkkuus!), kun tiedetään, että myös Big Benin käänteiskello Small Ben näyttää samaan aikaan aivan järkevää (vaikkakin toista) kellonaikaa? — Huomaa, että myös Small Benin minuuttiviisari on pidempi!

I.3 Logiikan ja joukko-opin merkinnöistä

Logiikassa (ja yleisemminkin matematiikassa) tarkastellaan väittämiä eli *propositioita*†, joilla on tietty *totuusarvo*, joko 'tosi' (totuusarvo = 1, tai T = True) tai 'epätosi' (totuusarvo = 0, tai F = False).

Esimerkki 1 Lausumista

A = 'eilen satoi'
B = 'eilen oli pouta'
C = 'huomenna sataa'
D = 'huomenna saisi jo sataa'
E = 'kolme per neljä on pienempi kuin kaksi per kolme'

kolmea ensimmäistä voidaan pitää propositioina, sikäli kuin lausumat rajataan tiettyä päivää ja paikkaa koskeviksi (ja jätetään huomiotta mahdolliset mittausongelmat). Samoin E on propositio. Lausuma D ei ole propositio. □

Esimerkkiväittämistä A ja B ovat toistensa *negaatioita* eli *komplementteja*. Merkitään $B = \neg A$, $A = \neg B$, tai vähemmän muodollisesti $B = ei(A)$, $A = ei(B)$. Väittämän C totuusarvo ei ole (tänään) tiedossa — tämä ei siis ole ongelma logiikassa.

Loogisten operaattorien $\wedge$ ('ja'), $\vee$ ('tai'), $\Rightarrow$ ('seuraa'), $\Leftrightarrow$ ('täsmälleen kun') avulla voidaan propositioista johtaa uusia propositioita. Erityisesti jos A ja B ovat propositioita, niin $A \wedge B$, $A \vee B$, $A \Rightarrow B$ ja $A \Leftrightarrow B$ ovat myös propositioita. Näistä kahden ensimmäisen tulkinta on ilmeinen: $A \wedge B$ on tosi kun A ja B ovat molemmat tosia, muulloin epätosi, ja $A \vee B$ on tosi täsmälleen kun ainakin toinen väittämistä A, B on tosi. Väittämä $A \wedge \neg A$ on *identtisesti epätosi* (epätosi jokaisella A), luonnollisella kielellä 'mahdoton'. Tällainen väittämä on *loogisen ristiriidan* perusmuoto. Väittämä $A \vee \neg A$ on puolestaan *identtisesti tosi* (tosi jokaisella A), eli suhteessa väittämään A 'mitäänsanomaton'.

Esimerkki 2 Jos A on epätosi ja B ja C molemmat tosia, niin $(A \wedge B) \vee C$ on tosi ja $A \wedge (B \vee C)$ on epätosi. Jälkimmäiset kaksi väittämää eivät siis ole yleisesti samanarvoiset (eli sulkeita ei voida poistaa). □

†Proposition sijasta käytetään suomenkielisissä teksteissä usein termiä *väitelause* tai vain 'lause'. Tässä tekstissä termillä 'lause' on erikoismerkitys (= 'teoreema'), termiä 'propositio' sen sijaan käytetään sekä yleis- että erikoismerkityksessä, ks. alaviite Luvussa I.1.

Implikaatio

Implikaatioväittämän $A \Rightarrow B$ looginen tulkinta ei ole aivan ilmeinen. Mahdollisia lukutapoja ovat ensinnäkin:

$A \Rightarrow B$: A:sta seuraa B
A *implikoi* B:n
jos A, niin B
aina kun A, niin B
aina B, kun A
A vain, kun B
A on B:n *riittävä ehto*
B on A:n *välttämätön ehto*

Sovelluksissa, myös luonnollisessa kielessä, voidaan implikaatioväittämä usein tulkita niin, että A on 'syy' ja B on 'seuraus'. Logiikassa ei mitään 'syyllisyyttä' kuitenkaan edellytetä, vaan implikaatio voi olla puhdas sattumakin ('sattumoisin aina B kun A'). Loogisessa kalkyylissä väittämän $A \Rightarrow B$ totuusarvo voidaan laskea tulkinnoista (ks. myös Harj.teht. 4a)

$$A \Rightarrow B \quad \Leftrightarrow \quad \neg\,(A \wedge \neg B) \qquad \text{(a)}$$
$$\Leftrightarrow \quad \neg A \vee (A \wedge B). \qquad \text{(b)}$$

Tässä $P \Leftrightarrow Q$ luetaan 'P:llä ja Q:lla on sama totuusarvo', ks. ekvivalenssinuolen tulkinnat jäljempänä. Implikaation tulkinnan (b) mukaan $A \Rightarrow B$ on tosi täsmälleen, kun joko A on epätosi tai A ja B ovat molemmat tosia. Erityisesti siis 'mahdottomasta seuraa mitä tahansa', eli väittämä $A \Rightarrow B$ on tosi jokaisella B (eli B:n suhteen 'mitäänsanomaton'), jos A on epätosi.

Esimerkki 3 Tulkinnoista (a)–(b) nähdään, että väittämä $A \Rightarrow A$ on identtisesti tosi, eli väittämä 'ei sano mitään' A:sta. □

Esimerkki 4 Jos 0 ja 1 ovat jonkin kunnan nolla- ja ykkösalkiot, niin propositio [$0 = 1 \;\Rightarrow\;$ kirjoittaja on mainio matemaatikko] on (loogisesti) tosi. □

Matematiikan lauseet, propositiot, lemmat ja korollaarit ovat tyypillisesti muotoa 'Jos .. [oletukset], niin .. [väitös]', eli muotoa $A \Rightarrow B$, missä A on <u>oletus</u> (oletukset) ja B on <u>väitös</u>. Loogista päättelyä, joka osoittaa lauseen tai vastaavan todeksi (sillä matematiikan lauseet ovat tosia!) sanotaan *todistukseksi* (engl. proof). Jos lause on muotoa $A \Rightarrow B$, niin todistaminen (eli proposition $A \Rightarrow B$ todeksi näyttäminen) tapahtuu <u>olettamalla</u>, että A on tosi ja <u>näyttämällä</u>, että

tällöin myös B on tosi (riittää, koska $A \Rightarrow B$ on tosi, jos A on epätosi!). Tämän nk. *suoran todistuksen* ohella toinen mahdollinen todistustapa on *epäsuora todistus*. Epäsuoran todistuksen ideana on näyttää, että jos todistettavan väittämän negaatio on tosi, niin seuraa looginen ristiriita (muotoa 'C tosi ja epätosi', missä C on propositio), jolloin päätellään, että negaation on oltava epätosi ja väittämän siis tosi (ks. Harj.teht. 3). Implikaatioväittämän tapauksessa epäsuoran todistuksen rakenne on seuraava, vrt. em. tulkinta (a).

(1) Oletetaan, että A on tosi ja B epätosi (eli $A \Rightarrow B$ epätosi).

(2) Näytetään, että oletus (1) johtaa loogiseen ristiriitaan. Päätellään, että oletus oli väärä ja siis $A \Rightarrow B$ tosi.

Oletusta (1) (tai osaoletusta 'B epätosi') sanotaan *vastaoletukseksi*, ja koko todistustavasta käytetään myös nimitystä *todistus vastaoletuksella* (engl. proof by contradiction). — Huomattakoon, että vastaoletus ei nimestään huolimatta ole oletuksen negaatio vaan pikemminkin 'vastaväitös'.

Epäsuoraan todistustapaan turvauduttiin itse asiassa jo edellä Lauseen I.2.1 kohdassa (g). Tässä väittämä oli muotoa $A \Rightarrow B$, missä

$$\begin{aligned} A &= \text{aksioomat (K1)–(K10) ja (J1)–(J4)}, \\ B &= 0 < 1. \end{aligned}$$

Vastaoletus, että B on epätosi, johti loogiseen ristiriitaan 'C tosi ja epätosi', missä $C = (\text{K10}) \wedge (\text{J1})$.

Ekvivalenssi

Jos A ja B ovat propositioita, niin $A \Leftrightarrow B$ on propositio, joka kertoo, että A ja B ovat samanarvoiset eli *ekvivalentit* väittämät. Tämä tarkoittaa yksinkertaisesti, että väittämien A ja B totuusarvot ovat samat, eli joko molemmat ovat tosia tai molemmat ovat epätosia. Ekvivalenssin yleisiä lukutapoja ovat

$A \Leftrightarrow B$: A ja B ekvivalentit, samanarvoiset
A ja B yhtäpitävät
A *silloin ja vain silloin kun* (engl. if and only if) B
A täsmälleen kun B
A *joss* (engl. iff) B

Ekvivalenssinuolen avulla voidaan esim. ilmaista jokin väittämä toisin sanoin tai merkinnöin, tai vain toisessa muodossa. Kyse voi tällöin olla esim. jonkin merkintätavan *määritelmästä*, tai yleispätevästä toisinnosta eli *tautologiasta*.

Esimerkki 5

$$\begin{aligned} 0 < x < 1 \quad &\Leftrightarrow \quad 0 < x \ \wedge \ x < 1 \\ A \Rightarrow B \quad &\Leftrightarrow \quad \neg B \Rightarrow \neg A \\ A \Rightarrow B \Rightarrow C \quad &\Leftrightarrow \quad A \Rightarrow B \ \wedge \ B \Rightarrow C \\ A \Leftrightarrow B \quad &\Leftrightarrow \quad A \Rightarrow B \ \wedge \ B \Rightarrow A \qquad \square \end{aligned}$$

Ensimmäinen esimerkki määrittelee merkinnän $0 < x < 1$. Toisessa esimerkissä on kyse yleisestä (myös hyödyllisestä!) tautologiasta (Harj.teht. 4b). Kolmannessa esimerkissä määritellään kahden implikaatioväittämän muodostama *päättelyketju*†. Viimeisessä esimerkissä tulkitaan itse väittämä $A \Leftrightarrow B$. Tämän tulkinnan mukaisesti ekvivalenssiväittämä todistetaan osoittamalla erikseen $\boxed{\Rightarrow}$ ($A \Rightarrow B$) ja $\boxed{\Leftarrow}$ ($A \Leftarrow B$) eli näyttämällä implikaatiot tosiksi molempiin suuntiin. Jos kyseessä on yksinkertainen tautologia, voidaan myös muodostaa *totuus*(arvo)*taulu*(kko), jossa käydään läpi kaikki mahdollisuudet.

Esimerkki 6 Näytä identtisesti tosiksi **de Morganin lait**:
a) $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$ b) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$

<u>Ratkaisu</u> Totuustaulussa (alla) on merkitty $P = A \vee B$, $Q = A \wedge B$ ja käyty läpi propositioiden A, B kaikki totuusarvoyhdistelmät (4 kpl). Taulukosta nähdään, että propositioiden $\neg P$ ja $\neg A \wedge \neg B$, samoin propositioiden $\neg Q$ ja $\neg A \vee \neg B$ totuusarvot ovat kaikissa tapauksissa samat. Tämä todistaa väitteet. $\square$

A	B	$\neg A$	$\neg B$	P	$\neg P$	$\neg A \wedge \neg B$	Q	$\neg Q$	$\neg A \vee \neg B$
1	1	0	0	1	0	0	1	0	0
1	0	0	1	1	0	0	0	1	1
0	1	1	0	1	0	0	0	1	1
0	0	1	1	0	1	1	0	1	1

Loogisten operaattorien, samoin järjestys- ja samastusrelaatioiden yms. yhteydessä negaation merkkinä käytetään yleisesti päälleviivausta.

Esimerkki 7

$$\begin{aligned} A \not\Rightarrow B \quad &\Leftrightarrow \quad \neg(A \Rightarrow B) \quad \Leftrightarrow \quad A \wedge \neg B \\ A \not\Rightarrow B \quad &\not\Leftrightarrow \quad A \Rightarrow \neg B \\ x \neq y \quad &\Leftrightarrow \quad \neg(x = y) \quad \square \end{aligned}$$

†Jos matemaattinen lause yms. on implikaatioväittämä muotoa $A \Rightarrow B$, niin todistus on tyypillisesti päättelyketju muotoa $A \Rightarrow C_1 \Rightarrow C_2 \Rightarrow \ldots \Rightarrow C_n \Rightarrow B$, missä osaväittämät $A \Rightarrow C_1$, $C_1 \Rightarrow C_2$, $\ldots$ $C_n \Rightarrow B$ joko ovat ilmeisen tosia, tai tosia muiden tunnettujen lauseiden (tai erikseen todistettavien aputulosten) perusteella. Päättelyketjun määrittely noudattaa tässä matematiikan käytäntöä — formaalissa logiikassa tällaista sopimusta ei ole.

Predikaatti ja kvanttorit

Logiikassa *predikaatti* on sellainen lausuma, jossa on yksi tai useampia vapaita *muuttujia*. Predikaatista tulee propositio, kun muuttujat *sidotaan* — sellaisenaan predikaatti ei ole propositio. Esimerkiksi jos tarkastellaan rationaalilukuja ja kirjoitetaan

$$P(x): \quad x > 2, \qquad Q(x,y): \quad x = y^2,$$

niin $P(x)$ on *yksipaikkainen* (so. yhden muuttujan sisältävä) predikaatti, nimeltään *epäyhtälö* (engl. inequality), ja $Q(x,y)$ on *kaksipaikkainen* predikaatti, nimeltään *yhtälö* (engl. equation). Jokaisella muuttujan x arvolla $P(x)$ on propositio, samoin $Q(x,y)$ kun molemmat muuttujat kiinnitetään. Esim. $P(5/2)$ ja $Q(4,2)$ ovat tosia, $P(1)$ ja $Q(0,1)$ epätosia. Muuttujien sitominen voi tapahtua myös *kvanttorien* avulla. Kvanttoreita ovat symbolit '$\forall$' ja '$\exists$', jotka luetaan

$\forall$ 'kaikille', 'jokaiselle',

$\exists$ 'on olemassa'.

Esimerkiksi jos $X \subset \mathbb{Q}$, niin ym. predikaattiin $P(x)$ liittyviä propostitioita ovat

$$A: \quad \exists x \in X \ (\, P(x)\,), \qquad B: \quad \forall x \in X \ (\, P(x)\,).$$

Tässä kvanttorien ulottuvuus on merkitty sulkeilla. Jos sulkeet halutaan välttää, niin A:n vähemän formaali muoto on

$$A: \quad \exists x \in X \ \text{ siten, että } \ P(x),$$

missä 'siten, että' voidaan haluttaessa lyhentää muotoon 's.e.'. Kummassakin propositiossa tullaan toimeen ilman sulkeita myös, kun kirjoitetaan yksinkertaisemmin

$$A: \quad P(x) \ \text{ jollakin } x \in X, \qquad B: \quad P(x) \ \ \forall x \in X.$$

Esimerkki 8 Jos $X = \mathbb{Q}$, niin predikaatista $P(x): \ x^2 \neq 2$ johdettu propositio $B: \ x^2 \neq 2 \ \forall x \in \mathbb{Q}$ on tosi (Harj.teht. 8). $\square$

Em. propositioiden A, B negaatiot saadaan ymmärrettävämpään muotoon suorittamalla <u>negaation</u> <u>purku</u> seuraavasti:

$$\neg\,[\,\exists x \in X\,(\,P(x)\,)\,] \quad \Leftrightarrow \quad \forall x \in X \ (\,\neg\,P(x)\,),$$
$$\neg\,[\,\forall x \in X \ (\,P(x)\,)\,] \quad \Leftrightarrow \quad \exists x \in X \ (\,\neg\,P(x)\,).$$

Jälkimmäisen säännön mukaisesti proposition $B: \ P(x)\,\forall x \in X$ näyttämiseen <u>epätodeksi</u> riittää löytää yksikin <u>vastaesimerkki</u> $x \in X$, jolle $P(x)$ on epätosi.

Esimerkki 9 Jos $X = \{x \in \mathbb{Q} \mid x < 1\}$, niin propositio

$$A: \quad \forall x \in X \; \exists y \in X \; (y > x)$$

on tosi. Jos $X = \{x \in \mathbb{Q} \mid x \leq 1\}$, niin A:n negaatio

$$\neg A: \quad \exists x \in X \; [\, \nexists y \in X \; (y > x)\,] \;\Leftrightarrow\; \exists x \in X \; [\forall y \in X \, (\, y \leq x \,)\,]$$

on tosi. $\square$

Formaali kvanttorimerkintä $\forall x \in X$ jätetään matemaattisissa teksteissä usein merkitsemättä silloin kun on selvää, että kyse on koko tiettyä joukkoa X (esim. $X = \mathbb{Q}$) koskevasta päättelystä.

Esimerkki 10 Rationaalilukuja koskeva päätelmä

$$2x^2 + x < 1 \;\Leftrightarrow\; -1 < x < \tfrac{1}{2}$$

ratkaisee $\mathbb{Q}$:ssa (oikein!) epäyhtälön $2x^2 + x < 1$. Päätelmä on muodoltaan predikaatti mutta tulkittavissa propositioksi, jossa sidonta $\forall x \in \mathbb{Q}$ on sujuvuussyistä 'näkymätön'. $\square$

Joukko-oppi

Logiikan ohella abstraktin ajattelun perimmäisiä perusteita käsittelee matematiikan laji nimeltä *joukko-oppi* (engl. set theory). Tässä todettakoon ainoastaan lyhyesti eräiden logiikan ja joukko-opin perusideoiden välinen yhteys. Ensinnäkin 'mahdottoman väittämän' muotoa $A \wedge \neg A$ vastine joukko-opissa on *tyhjä joukko*, jonka symboli on '$\emptyset$'. Tyhjässä joukossa ei ole alkioita, eli $x \in \emptyset$ on aina epätosi.

Loogisen negaation $\neg A$ vastine joukko-opissa on joukon A *komplementti*, joka merkitään $\complement(A)$ ja määritellään

$$x \in \complement(A) \quad \Leftrightarrow \quad x \notin A.$$

Käytännössä komplementti on määriteltävä jonkin *universaalijoukon*† U suhteen. Tällä ymmärretään sellaista joukkoa, johon kaikki tarkasteltavat joukot (mukaan lukien joukkojen komplementit) sisältyvät osajoukkoina.

†Termi 'universaalijoukko' saattaa herättää mielikuvan joukosta, joka sisältää kirjaimellisesti 'kaiken'. Tämän tyyppisiä ajatelmia joukko-opissa olikin sen teorian alkuvaiheissa, mutta niiden huomattiin johtavan paradokseihin, ts. loogisiin mahdottomuuksiin. Sittemmin täsmentyneistä joukko-opin aksioomista seuraakin, että 'mikään ei sisällä kaikkea'.

Esimerkki 11 Olkoon $P(x)$ jokin rationaalilukujoukkoon $\mathbb{Q}$ liittyvä predikaatti, esim. $P(x) = x < 2$. Tällöin ehto '$P(x)$ tosi' määrittelee $\mathbb{Q}$:n osajoukon A, jota merkitään $A = \{\, x \in \mathbb{Q} \mid P(x) \,\}$. Universaalijoukoksi on tässä luonnollista ajatella $U = \mathbb{Q}$, jolloin A:n komplementti on

$$\complement(A) \;=\; \{\, x \in \mathbb{Q} \mid \neg P(x) \,\} \;=\; \{\, x \in \mathbb{Q} \mid x \geq 2 \,\}. \quad \square$$

Loogisten yhdistelyjen $A \vee B$, $A \wedge B$, $A \Rightarrow B$ ja $A \Leftrightarrow B$ joukko-opilliset vastineet ovat $A \cup B$, $A \cap B$, $A \subset B$ ja $A = B$, kuten nähdään määritelmistä:

$$\begin{aligned} x \in A \cup B \quad &\Leftrightarrow \quad x \in A \;\vee\; x \in B \\ x \in A \cap B \quad &\Leftrightarrow \quad x \in A \;\wedge\; x \in B \\ A \subset B \quad &\Leftrightarrow \quad x \in A \;\Rightarrow\; x \in B \\ A = B \quad &\Leftrightarrow \quad x \in A \;\Leftrightarrow\; x \in B \end{aligned}$$

Joukko $A \cup B$ on nimeltään A:n ja B:n *yhdiste* eli *unioni* (engl. union) ja $A \cap B$ on A:n ja B:n *leikkaus* (engl. intersection). Jos kahdella joukolla A, B ei ole yhteisiä alkioita, eli $x \in A \wedge x \in B$ on epätosi jokaisella x, niin tämä voidaan ilmaista lyhyesti merkinnällä $A \cap B = \emptyset$. Sanotaan tällöin, että A ja B ovat *erillisiä* eli *pistevieraita* (engl. disjoint).

Esimerkki 12 Esimerkin 10 päättely tulkittiin propositioksi ajattelemalla kvanttorimerkintä $\forall x \in \mathbb{Q}$ lisätyksi. Vielä luontevampi on joukko-opillinen tulkinta: $\{\, x \in \mathbb{Q} \mid 2x^2 + x < 1 \,\} \;=\; \{\, x \in \mathbb{Q} \mid -1 < x < \frac{1}{2} \,\}$. $\square$

Esimerkki 13 Todista joukko-opin de Morganin lait (vrt. Esimerkki 6)

$$\begin{aligned} \complement(A \cup B) \quad &= \quad \complement(A) \cap \complement(B), \\ \complement(A \cap B) \quad &= \quad \complement(A) \cup \complement(B). \end{aligned}$$

<u>Ratkaisu</u> Päättelyketjussa

$$\begin{aligned} x \in \complement(A \cup B) \quad &\Leftrightarrow \quad x \notin (A \cup B) \\ &\Leftrightarrow \quad \neg\,(\, x \in A \;\vee\; x \in B \,) \\ &\Leftrightarrow \quad x \notin A \;\wedge\; x \notin B \qquad\qquad [\,\text{Esim. 6 a)}\,] \\ &\Leftrightarrow \quad x \in \complement(A) \;\wedge\; x \in \complement(B) \\ &\Leftrightarrow \quad x \in \complement(A) \cap \complement(B) \end{aligned}$$

voidaan edetä molempiin suuntiin, joten joukoilla $\complement(A \cup B)$ ja $\complement(A) \cap \complement(B)$ on samat alkiot, ja ensimmäinen väittämä on siis todistettu. Toisen väittämän todistus saadaan tästä vaihdoilla $\cup \hookrightarrow \cap$, $\cap \hookrightarrow \cup$, $\vee \hookrightarrow \wedge$ ja $\wedge \hookrightarrow \vee$. $\square$

Ekvivalenssirelaatio

Relaatio on joukko-opillinen 'suhteen' käsite. Jos R on joukossa A määritelty relaatio ja $x, y \in A$, niin merkintä $x\,R\,y$ luetaan 'x on relaatiossa R y:n kanssa', tai sujuvammin vain 'x [relaation nimi] y'. Relaatioista ovat jo tuttuja käytännössä tärkeimmät, eli järjestys- ja samastusrelaatio. Samastusrelaation on aina oltava *ekvivalenssirelaatio*, jonka aksioomat ovat seuraavat:

(E1) $x\,R\,x \quad \forall x$

(E2) $x\,R\,y \;\Rightarrow\; y\,R\,x$

(E3) $x\,R\,y \;\wedge\; y\,R\,z \;\Rightarrow\; x\,R\,z$

Vaaditut ominaisuudet ovat nimeltään *refleksiivisyys*, *symmetrisyys* ja *transitiivisuus*. Esimerkiksi relaatio '$\leq$' on refleksiivinen ja transitiivinen muttei symmetrinen, ja relaatio '$\neq$' on ainoastaan symmetrinen. Samastusrelaatio 'sama kuin' sen sijaan on mitä ilmeisimmin ekvivalenssirelaatio. Esimerkiksi kun kirjoitetaan $x = y = z$ (tarkoittaen: $x = y$ ja $y = z$), niin on ilmeistä, että $x = z$. Tässä on siis kyse transitiivisuudesta.[†]

Esimerkki 14 Jos $A = \{\text{suomalaiset}\}$, niin seuraavat A:ssa määritellyt relaatiot ovat ekvivalenssirelaatioita:

$$\begin{aligned} x \sim y \quad &\Leftrightarrow \quad x \text{ ja } y \text{ ovat syntyneet samana vuonna,} \\ x \heartsuit y \quad &\Leftrightarrow \quad y = x \quad \text{tai} \quad y \text{ on } x\text{:n puoliso.} \quad \square \end{aligned}$$

Joukossa A määritellyllä ekvivalenssirelaatiolla on se ominaisuus, että se jakaa A:n *ekvivalenssiluokkiin*. Nämä ovat A:n osajoukkoja, joiden sisältämät alkiot ovat kaikki relaatiossa keskenään. Refleksiivisyysominaisuuden (E1) vuoksi jokainen A:n alkio kuuluu ainakin yhteen ekvivalenssiluokkaan (mahdollisesti yksinään). Toisaalta transitiivisuusominaisuudesta (E3) seuraa, että kaksi ekvivalenssiluokkaa ovat joko täysin samat tai ne ovat pistevieraita — siis jokainen A:n alkio kuuluu täsmälleen yhteen ekvivalenssiluokkaan. Jos kyseessä on samastusrelaatio, niin kunkin ekvivalenssiluokan sisältämät alkiot 'luokitellaan samoiksi' eli samastetaan keskenään. Tällöin voidaan puhua myös *samastusluokista*.

Esimerkki 14 (jatko) Relaation '$\sim$' määräämiä ekvivalenssiluokkia sanotaan ikäluokiksi. Relaatio '$\heartsuit$' jakaa A:n ekvivalenssiluokkiin joissa on joko yksi alkio (alaikäiset, sinkut, ym.) tai kaksi alkiota (rekisteröidyt parit). $\square$

[†]Samastusrelaation yleiset aksioomat ovat samat kuin ekvivalenssirelaation, vrt. alaviite edellisessä luvussa.

HARJOITUSTEHTÄVIÄ

1. Olkoon A propositio 'Sataa', B propositio 'Menen lenkille' ja C propositio 'Minulla on aikaa'. Kirjoita näiden ja loogisten operaattoreiden avulla mahdollisimman pelkistetysti:
   a) 'Väsyttää, on pimeää ja TV:stä tulee BB, joten en mene lenkille.'
   b) 'Menen lenkille, satoi tai paistoi!'
   c) 'Sataa, joten en mene lenkille.'
   d) 'Sataa enkä mene lenkille.'
   e) 'Menen lenkille vain, jos minulla on aikaa.'
   f) 'Minulla on aikaa eikä sadakaan, mutta en mene lenkille.'
   g) 'Jos ei sada ja minulla on aikaa, niin menen lenkille.'
   h) 'Minulla on aikaa vain, jos sateen vuoksi en mene lenkille.'

2. Olkoon A = 'Tänään sataa', B = 'Huomenna on pouta' ja C = 'Tänään on pouta'. a) Jos tänään alkanut sade jatkuu huomiseen, niin mitkä ovat propositioiden $A \Rightarrow B$ ja $B \Rightarrow A$ totuusarvot? b) Riippuuko proposition $\neg(\neg A \wedge B) \Rightarrow \neg C$ totuusarvo siitä, mikä päivä on 'tänään'?

3. Osoita joko totuustaulun avulla tai muuten päättelemällä seuraava propositio (nk. *modus tollens*) identtisesti todeksi: $[\,(\neg P \Rightarrow Q) \wedge \neg Q\,] \Rightarrow P$. Miten tulokseen vedotaan epäsuorassa todistuksessa?

4. Näytä totuustaulun avulla tai muuten päättelemällä identtisesti todeksi:
   a) $\neg\,(A \wedge \neg B) \;\Leftrightarrow\; \neg A \vee (A \wedge B)$ b) $(A \Rightarrow B) \;\Leftrightarrow\; (\neg B \Rightarrow \neg A)$
   c) $[\,(A \Rightarrow B) \wedge (B \Rightarrow C)\,] \;\Rightarrow\; (A \Rightarrow C)$
   d) $[\,(A \Rightarrow C) \wedge (B \Rightarrow C)\,] \;\Rightarrow\; [\,(A \vee B) \Rightarrow C\,]$
   e) $[\,(A \wedge B) \Rightarrow C\,] \;\Leftrightarrow\; [\,A \Rightarrow (B \Rightarrow C)\,]$

   Vertaile totuustaulun avulla:

   f) $(A \wedge B) \vee C$ ja $A \wedge (B \vee C)$ g) $A \not\Rightarrow B$ ja $A \Rightarrow \neg B$

   Tutki totuustaulun avulla, voiko seuraavat propositiot ilmaista jollakin yksinkertaisemmalla (ekvivalentilla) tavalla:

   h) $(A \Rightarrow B) \wedge B$ i) $(A \Rightarrow B) \Rightarrow A$ j) $[\,(A \Rightarrow B) \wedge B\,] \;\Rightarrow\; A$

5. Jos seuraavat predikaatit $P(x)$ tulkitaan väittäminä $P(x)\ \forall x \in \mathbb{Q}$, niin mitkä väittämistä ovat tosia ja mitkä epätosia?
   a) $2x > 3 \;\Rightarrow\; x > 4/3$ b) $x < -2 \;\Leftrightarrow\; x^2 > 4$
   c) $3x < 4 \;\Rightarrow\; x < 1 \;\vee\; 2x < 3$ d) $x > 1 \;\not\Leftrightarrow\; x > 0 \;\wedge\; x > 1$
   e) $x = 4/3 \;\wedge\; 3x^2 < x + 4 \;\Rightarrow\; x^2 + 2x + 1 < 0$
   f) $2x^2 + x \;\leq\; 1 \;\leq\; 3x - 2x^2 \;\Leftrightarrow\; x = 4/3$

6. Olkoot x ja y rationaalilukuja. Mitkä seuraavista propositioista ovat tosia?
   a) $\forall x \exists y\ (x \cdot y = 0)$ b) $\forall x \exists y\ (x \cdot y = 1)$
   c) $\exists y \forall x\ (x \cdot y = 0)$ d) $\exists y \forall x\ (x \cdot y = 1)$
   e) $\exists y \forall x\ (x \cdot y = x)$ f) $\exists y \forall x\ (x \cdot y = y)$
   g) $\forall x \exists y\ (y < x)$ h) $\exists x \forall y\ (y \geq x)$
   Miten tilanne muuttuu, jos x ja y ovat positiivisia rationaalilukuja?

7. Olkoon $x \in \mathbb{Q}$. Muodosta seuraavien kahden proposition negaatiot auki purettuina. Mitkä näin syntyvistä neljästä propositiosta ovat tosia?
   a) $\forall(\varepsilon > 0)\exists(x \neq 1)(|x-1| < \varepsilon)$ b) $\exists(x \neq 1)\forall(\varepsilon > 0)(|x-1| < \varepsilon)$

8. Näytä epäsuoralla todistustavalla: a) Ei ole olemassa pienintä positiivista rationaalilukua. b) $x^2 \neq 2\ \forall x \in \mathbb{Q}$. c) $x^2 \neq 3\ \forall x \in \mathbb{Q}$.

9. Universaalijoukko U olkoon kymmenen ensimmäisen luonnollisen luvun muodostama joukko $\{1,2,3,4,5,6,7,8,9,10\}$. Olkoon $A = \{2,5,7,8,10\}$ ja $B = \{2,4,6,8,10\}$. Määritä $A \cup B$, $A \cap B$, $\complement(A)$, $\complement(A \cup B)$ ja $\complement(A) \cap \complement(B)$.

10. Todista seuraavat joukko-opilliset väittämät:
    a) $A \subset B \ \Leftrightarrow\ A \cup B = B \ \Leftrightarrow\ A \cap B = A$
    b) $A \subset B \ \wedge\ A \subset C \ \Leftrightarrow\ A \subset B \cap C$
    c) $A \subset C \ \wedge\ B \subset C \ \Leftrightarrow\ A \cup B \subset C$
    d) $A \subset B \ \Leftrightarrow\ A \cap C \subset B \cap C\ \forall C$
    e) $\emptyset \subset A\ \forall A$

11. Muotoile ja todista Tehtävän 4 väittämien b)–d) joukko-opilliset vastineet.

12. Olkoon $A = \{(p,q) \mid p,q \in \mathbb{Z}\}$ (kokonaislukuparien joukko) ja $Q = \{(p,q) \in A \mid q \neq 0\}$. Määritellään Q:ssa relaatio $(p_1,q_1) \sim (p_2,q_2) \ \Leftrightarrow\ p_1q_2 = p_2q_1$. Näytä, että kyseessä on Q:n ekvivalenssirelaatio. Miten tulos liittyy rationaalilukuihin?

13. Mitä ekvivalenssirelaation ominaisuuksia on seuraavilla, annetuissa joukoissa A määritellyillä relaatioilla?
    a) $A = \mathbb{N}:\ xRy \ \Leftrightarrow\ x + y$ on parillinen.
    b) $A = \mathbb{N}:\ xRy \ \Leftrightarrow\ x + y$ on pariton.
    c) $A = \mathbb{Q}:\ xRy \ \Leftrightarrow\ |x - y| \leq 10^{-100}$.
    d) $A = \{\,\text{suomalaiset}\,\}:\ xRy \ \Leftrightarrow\ x$ ja y ovat toisilleen sukua suoraan ylenevässä tai alenevassa polvessa.

14. (*) Luonnollinen luku $k \in \mathbb{N}$ on luvun $n \in \mathbb{N}$ *tekijä*, jos $n = k \cdot m$ jollakin $m \in \mathbb{N}$. Luku n on *alkuluku*, jos n:llä ei ole muita tekijöitä kuin $k = 1$ ja $k = n$. Näytä, että alkuluvuilla ei ole loppua, ts. ei ole olemassa suurinta alkulukua.

I.4 Jonon käsite

Jono (engl. sequence, ruots. följd) on olio muotoa

$$\{\, a_1, a_2, a_3, \ldots \}.$$

Joukosta jonon erottaa ennen muuta järjestys: Jonon jäsenet, joita kutsutaan jonon *termeiksi* (vrt. joukon alkio) on 'pantu jonoon' eli järjestykseen. Matemaattisessa jonossa järjestys tarkoittaa vastaavuutta luonnollisten lukujen joukon ja jonon välillä:

$$\begin{aligned} 1 &\mapsto a_1 \\ 2 &\mapsto a_2 \\ &\vdots \end{aligned}$$

Tässä '$\mapsto$' tarkoittaa jälleen 'liittämistä': Jokaiseen $n \in \mathbb{N}$ liitetään yksikäsitteinen jonon termi (n:s termi) a_n, jolloin sanotaan, että n on ko. termin järjestysnumero eli *indeksi*. Indeksijoukko on siis koko $\mathbb{N}$, mikä tarkoittaa, että jono on päättymätön[†].

Matemaattisen jonoon, niinkuin joukkoonkin, voi periaatteessa sijoittaa mitä tahansa. Jatkossa tarkastelun kohteena ovat ennen muuta *lukujonot*, joille käytetään lyhennysmerkintöjä

$$\begin{aligned} &\{\, a_n,\ n = 1, 2, \ldots \}, \\ &\{a_n\}_{n=1}^{\infty}, \\ &\{a_n\}. \end{aligned}$$

Jos jonon termit määräytyvät tunnettuna, indeksistä riippuvana lausekkeena, voidaan ko. lauseke kirjoittaa a_n:n paikalle.

Esimerkki 1

$$\begin{aligned} &\{\, 1, -1, 1, -1, \ \ldots \,\} = \{\, (-1)^{n+1},\ \ n = 1, 2, \ \ldots \,\} \\ &\left\{ \frac{1}{9}, \frac{1}{16}, \frac{1}{25}, \ \ldots \right\} = \left\{ \frac{1}{(k+2)^2}\,, \quad k = 1, 2, \ \ldots \right\} \\ &\{\, 1, 1, 2, 6, 24, 120, 720, 5040, 40320, \ \ldots \,\} = \{\, n! \,\}_{n=0}^{\infty} \qquad \square \end{aligned}$$

[†] 'Päättyvästä jonosta' muotoa

$$(\, a_1, a_2, \ldots, a_n \,), \quad n \in \mathbb{N}$$

käytetään tässä tekstissä yleisnimitystä (äärellinen) *järjestetty joukko* (tai 'n alkion järjestetty joukko'), tapauksissa $n = 2, 3$ nimiä *pari* ja *kolmikko*. Jonomerkinnästä poiketen käytetään tässä yhteydessä pääsääntöisesti kaarisulkeita. — Kirjallisuudessa kaarisulkeilla merkitään joskus myös päättymättömiä jonoja, samoin merkintää $< a_n >$ näkee jonoista käytettävän.

Viimeinen esimerkki on myös esimerkki yleisemmästä jonosta muotoa

$$\{\, a_m, a_{m+1}, a_{m+2},\ \ldots \,\} = \{a_n\}_{n=m}^{\infty}, \quad m \in \mathbb{Z}.$$

Indeksin vaihdolla tämä voidaan palauttaa normaalimuotoon:

$$\{a_n\}_{n=m}^{\infty} = \{a_{m+k-1}\}_{k=1}^{\infty}.$$

Esimerkki 2 Lukujono

$$a_0 = \tfrac{1}{2}, \quad a_n = 1 - a_{n-1}^2, \quad n = 1, 2, \ldots$$

on esimerkki *palautuvasta* eli *rekursiivisesta* lukujonosta, joka 'määrittelee itsensä'. Tässä $a_1 = 3/4,\ a_2 = 7/16,\ a_3 = 207/256,\ \ldots$ $\square$

Sarja

Jos $\{\, a_k,\ k = 1, 2,\ \ldots \,\}$ on lukujono ja $\{s_n\}$ toinen lukujono, joka määritellään

$$s_n = a_1 + \ldots + a_n = \sum_{k=1}^{n} a_k\,,$$

niin sanotaan että $\{s_n\}$ on *sarja* (engl. series). Sarjan tavanomainen lyhennetty merkintätapa on

$$\sum_{k=1}^{\infty} a_k \quad \Big(= \ \{\, \sum_{k=1}^{n} a_k\,,\ \ n = 1, 2,\ \ldots \,\} \Big).$$

Lukuja a_k sanotaan *sarjan termeiksi* ja lukuja s_n *sarjan osasummiksi.* Sarja tulkitaan siis lukujonoksi, joka muodostuu sarjan osasummista.†

Esimerkki 3

$$\sum_{k=0}^{\infty} (-1)^k \ = \ \{\, \sum_{k=0}^{n} (-1)^k\,,\ \ n = 0, 1,\ \ldots \,\} \ = \ \{\, 1, 0, 1, 0,\ \ldots \,\} \qquad \square$$

'Kaikkien sarjojen äiti' on (perusmuotoinen) *geometrinen sarja*, jonka termit ovat $a_k = q^k,\ k = 0, 1,\ \ldots$, ts. sarja on muotoa $\sum_{k=0}^{\infty} q^k = \{1, 1+q, 1+q+q^2,\ \ldots\}$‡. Osasummille saadaan tässä tapauksessa laskukaava (ks. Harj.teht. I.2: 3 c)

$$\boxed{s_n = \sum_{k=0}^{n} q^k \ = \ \frac{q^{n+1} - 1}{q - 1}, \quad q \neq 1.}$$

†Sarjan tulkinta tässä tekstissä hieman 'oikoo' sarjan formaalia määritelmää, joka kuuluu: Sarja on jonojen $\{a_k\}$ (sarjan termit) ja $\{s_n\}$ (osasummat) muodostama jonopari.

‡Mukavuussyistä sovittakoon, että geometrisessa sarjassa $\sum_{k=0}^{\infty} q^k$ ensimmäinen termi on $q^0 = 1$ myös kun $q = 0$.

Induktio

Luonnollisiin lukuihin ja jonon käsitteeseen liittyy läheisesti myös matemaattinen todistusperiaate nimeltä *induktio*. Olkoon $P(n)$, $n \in \mathbb{N}$, predikaatti. Tällöin proposition

$$\mathcal{P}: \quad P(n) \ \ \forall n \in \mathbb{N}$$

voi tulkita viittaavan väittämäjonoon $\{\, P(1), P(2), \ldots \}$. Induktion ideana on muuntaa tämä jono palautuvaksi, jolloin jono 'todistaa itsensä' samalla tavoin kuin palautuva lukujono 'laskee itsensä' (vrt. Esimerkki 2 edellä). Idea realisoidaan tarkastelemalla kahta propositiota:

$$\begin{aligned} \mathcal{P}_1 &: \quad P(1), \\ \mathcal{P}_2 &: \quad P(n) \Rightarrow P(n+1) \ \ \forall n \in \mathbb{N}. \end{aligned}$$

Jos $\mathcal{P}_1$ ja $\mathcal{P}_2$ ovat molemmat tosia, niin $\mathcal{P}_1$ käynnistää propositioon $\mathcal{P}_2$ perustuvan 'todistusautomaatin', joten implikaatio

$$\mathcal{P}_1 \wedge \mathcal{P}_2 \quad \Rightarrow \quad \mathcal{P},$$

joka lausuu nk. *induktioperiaatteen*, tuntuu ilmeisen todelta.[†] Induktioperiaatteen mukaan siis väittämän $\mathcal{P}$ todistamiseksi riittää osoittaa, että $\mathcal{P}_1$ ja $\mathcal{P}_2$ ovat molemmat tosia. Proposition $\mathcal{P}_2$ toteennäyttämistä sanotaan *induktioaskeleeksi* ja todistuksen ko. osan lähtöoletusta '$n \in \mathbb{N}$ ja $P(n)$ tosi' *induktio-oletukseksi*.

Näytetään induktion avulla oikeaksi seuraava tuttu tulos:

Propositio I.4.1 (**Binomikaava**) Rationaalilukujen kunnassa tai sen laajennuksissa pätee

$$(x+y)^n \ = \ \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k, \quad n \in \mathbb{N},$$

missä *binomikertoimet* määritellään

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

<u>Todistus</u> Väittämä on muotoa $P(n)$ $\forall n \in \mathbb{N}$. Tässä $P(1)$ on ilmeisen tosi, joten riittää suorittaa induktioaskel. Oletetaan siis, että $P(n)$ on tosi, ts. että binomikaava pätee tietyllä (mutta mielivaltaisesti valitulla) n:n arvolla. Tällöin on

[†]Induktioperiaate seuraa Peanon aksioomasta (P5) asettamalla $S = \{n \in \mathbb{N} \mid P(n) \text{ tosi}\}$, ks. alaviite Luvussa I.1.

kunnan aksioomien ja oletuksen perusteella

$$\begin{aligned}(x+y)^{n+1} &= (x+y)(x+y)^n = (x+y)\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{n-k+1}y^k + \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^{k+1} \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{n+1-k}y^k + \sum_{l=1}^{n+1}\binom{n}{l-1}x^{n+1-l}y^l.\end{aligned}$$

Tässä on jälkimmäisessä summassa tehty indeksin vaihto $l = k+1$. Kun jälleen kirjoitetaan $l = k$, voidaan molemmat summat yhdistää tulokseksi

$$(x+y)^{n+1} = \sum_{k=0}^{n+1} c_k\, x^{n+1-k}y^k,$$

missä

$$c_0 = \binom{n}{0} = 1 = \binom{n+1}{0}, \qquad c_{n+1} = \binom{n}{n} = 1 = \binom{n+1}{n+1},$$

ja indeksin arvoilla $k = 1\ldots n$

$$\begin{aligned}c_k &= \binom{n}{k} + \binom{n}{k-1} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} \\ &= \frac{n!}{(k-1)!(n-k)!}\left(\frac{1}{k} + \frac{1}{n-k+1}\right) \\ &= \frac{n!}{(k-1)!(n-k)!}\cdot\frac{n+1}{k(n-k+1)} \\ &= \frac{(n+1)!}{k!(n-k+1)!} = \binom{n+1}{k}.\end{aligned}$$

Näin ollen $P(n+1)$ on tosi ja induktioaskel siis suoritettu. Todistuksessa tarvittiin yleisten kuntaoperaatioiden lisäksi luonnollisten lukujen välisiä laskuoperaatioita (mukaan lukien jakolasku), joten todistus on pätevä rationaalilukujen kunnassa tai sen laajennuksissa. □

Kuten Proposition I.4.1 todistuksesta käy ilmi, binomikertoimet ovat laskettavissa palautuvasti kaavasta

$$\binom{n+1}{0} = \binom{n+1}{n+1} = 1, \qquad \binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}, \quad k = 1\ldots n.$$

Kaava havainnollistuu *Pascalin kolmiossa*, jonka avulla kertoimet voidaan määrittää kätevästi pienillä n:n arvoilla:

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & 1 & & & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & & & (n=1) \\
 & & & & & 1 & & 2 & & 1 & & & & & & & (n=2) \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & & & (n=3) \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & & (n=4) \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & & (n=5) \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & & (n=6) \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 & & (n=7) \\
1 & 8 & & 28 & & 56 & & 70 & & 56 & & 28 & & 8 & & 1 & (n=8)
\end{array}
$$

Pascalin kolmion avulla binomikertoimet määräytyvät pelkällä luonnollisten lukujen yhteenlaskulla.

Toisena induktion sovelluksena todistetaan usein käytetty epäyhtälö.

PROPOSITIO I.4.2 (**Bernoullin epäyhtälö**) Rationaalilukujen kunnassa tai sen järjestetyissä laajennuksissa pätee

$$\boxed{(1+x)^n \;>\; 1+nx \quad \text{kun} \;\; n \in \mathbb{N},\; n \geq 2 \;\; \text{ja} \;\; x \geq -1,\; x \neq 0.}$$

<u>Todistus</u> Jos $x > 0$, niin epäyhtälö seuraa helpoiten binomikaavasta:

$$(1+x)^n \;=\; 1 + \binom{n}{1} x \;+\; [\ldots] \;=\; 1 + nx + [\ldots],$$

missä $[\ldots] > 0$, kun $x > 0$ ja $n \geq 2$. Yleispätevämpi todistus perustuu induktioon, eikä sekään ole pitkä: Ensinnäkin

$$(1+x)^2 \;=\; 1 + 2x + x^2 \;> 1 + 2x, \quad \text{kun } x \neq 0,$$

joten väittämä on tosi, kun $n = 2$. Toisaalta jos väite on tosi millä tahansa n:n arvolla (induktio-oletus, $n \geq 2$), niin

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)(1+x)^n \\
&\geq (1+x)(1+nx) \\
&= 1 + (n+1)x + nx^2 \\
&> 1 + (n+1)x,
\end{aligned}$$

kun $x \geq -1 \Rightarrow 1 + x \geq 0$ ja $x \neq 0 \Rightarrow x^2 > 0$. Induktioperiaatteen mukaan väittämä on näin todistettu. $\square$

Numeroituvuus, mahtavuus

Sanotaan, että joukot A ja B ovat *yhtä mahtavat* (engl. of the same cardinality), jos joukkojen välillä on olemassa kääntäen yksikäsitteinen vastaavuus siten, että jokaista A:n alkiota vastaa yksikäsitteinen B:n alkio ja kääntäen. Vastaavuutta merkitään jatkossa kaksoisnuolella '$\leftrightarrow$':

$$A \leftrightarrow B \qquad \text{(kääntäen yksikäsitteinen vastaavuus).}$$

Jos erityisesti $A \leftrightarrow \{\,1, 2,\ \ldots\ n\,\}$ jollakin $n \in \mathbb{N}$, niin sanotaan, että A on *äärellinen* (engl. finite) joukko. Tällöin A:ssa on täsmälleen n alkiota, ja A ja B ovat yhtä mahtavat täsmälleen, kun myös B:ssä on n alkiota. Äärellisen joukon tapauksessa vastaavuutta $A \leftrightarrow \{\,1, 2,\ \ldots\ n\}$ sanotaan A:n alkioiden *numeroinniksi* (indeksoinniksi). Numerointi tekee A:sta järjestetyn joukon.

Esimerkki 4

$$\begin{aligned} A &= \{\,1, 2, 3, 1000\,\} \\ B &= \{\,\text{vuohi V, hevonen H, opiskelija O, professori P}\,\} \\ C &= \{\,\text{linnunradan atomit}\,\} \end{aligned}$$

Kaikki kolme ovat äärellisiä joukkoja. Joukot A ja B ovat yhtä mahtavat. $\square$

Jos joukko ei ole äärellinen, niin se on *ääretön* (engl. infinite).

Esimerkki 5 Joukot

$$\begin{aligned} A &= \{\,1, 4, 9, 16, 25, 36, 49,\ \ldots\,\}, \\ B &= \{\,1, 100, 10000, 1000000,\ \ldots\,\} \end{aligned}$$

ovat molemmat äärettömiä. Ne ovat myös yhtä mahtavat, sillä vastaavuuksista

$$\begin{aligned} n \in \mathbb{N} \quad &\leftrightarrow \quad n^2 \in A \\ n \in \mathbb{N} \quad &\leftrightarrow \quad 10^{2n-2} \in B \end{aligned}$$

nähdään, että $A \leftrightarrow \mathbb{N} \leftrightarrow B$. $\square$

Esimerkin perusteella pienemmältä tai suuremmalta 'tuntumiseen' ei voi luottaa vertailtaessa äärettömien joukkojen mahtavuuksia. Sanonnat kuten 'yhtä monta' tai 'sama määrä' on myös selkeintä rajata äärellisten joukkojen vertailuun.

Esimerkin 5 joukkojen A, B alkiot voidaan numeroida esitettyjen vastaavuuksien perusteella. Yleisesti sanotaan, että joukko A on *numeroituva* (tai 'numeroituvasti ääretön', engl. (d)enumerable tai countably infinite), jos $A \leftrightarrow \mathbb{N}$. Numeroituvuus siis tarkoittaa, että joukon alkiot voidaan järjestää jonoksi.

Esimerkki 6 Kokonaislukujoukon $\mathbb{Z} = \{\, 0, \pm 1, \pm 2, \ldots \}$ eräs jonomuoto on

$$\{\, 0, 1, -1, 2, -2, \ \ldots \} \quad = \quad \{\, a_n, \ n = 1, 2, \ \ldots \},$$

joten $\mathbb{Z}$ on numeroituva. Esimerkiksi luku -777 on jonon 1555:s termi. $\square$

Esimerkki 7 Rationaalilukujoukko $\mathbb{Q}$ on esitettävissä muodossa

$$\mathbb{Q} \ = \ A_1 \ \cup \ A_2 \ \cup \ \cdots \ = \ \bigcup_{m=1}^{\infty} A_m,$$

missä

$$\begin{aligned} A_1 \ &= \ \{\, 0 \,\}, \\ A_m \ &= \ \big\{\, x = p/q \in \mathbb{Q} \ \mid \ |p| + |q| = m \ \ \& \ \ x \notin A_k \ \text{ kun } \ k < m \,\big\}, \quad m = 2, 3, \ldots \end{aligned}$$

Osajoukot A_m ovat määritelmän mukaisesti keskenään pistevieraita ($A_m \cap A_k = \emptyset$ kun $k < m$), joten jokainen $x \in \mathbb{Q}$ on enintään yhden osajoukon alkio. Toisaalta määritelmästä seuraa myös, että jos $x \in \mathbb{Q}$, niin $x \in A_m$, missä $m \in \mathbb{N}$ on pienin luku, jolle pätee $m = |p| + |q|$ ja $x = p/q$. Näin ollen jokainen $x \in \mathbb{Q}$ on täsmälleen yhden osajoukon A_m alkio. Joukot A_m ovat myös äärellisiä, joten jokainen niistä voidaan numeroida erikseen. Jos nyt A_m:n alkioiden lukumäärä $= N_m$, niin mielivaltaiselle $x \in \mathbb{Q}$ saadaan yksikäsitteinen järjestysnumero n säännöllä

$$n = \begin{cases} 1, & \text{jos } x = 0, \\ N_1 + \cdots + N_{m-1} + k, & \text{jos } x \text{ on } A_m\text{:n } k\text{:s alkio}, \quad m \geq 2. \end{cases}$$

Kaikki rationaaliluvut on näin järjestetty jonoksi, ja voidaan siis todeta, että $\mathbb{Q}$ on numeroituva. $\square$

Jos joukko on ääretön mutta ei numeroituva, sanotaan että se on *ylinumeroituva* (engl. uncountable). Näinkin 'mahtavia' joukkoja — myös lukujoukkoja — on olemassa, kuten tullaan näkemään.

HARJOITUSTEHTÄVIÄ

1. Näytä induktiolla todeksi summakaavat ($n \in \mathbb{N}$)
   a) $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$ b) $\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$
   c) $\sum_{k=1}^{n} k^3 = \frac{1}{4}n^2(n+1)^2$ d) $\sum_{k=1}^{n} k^4 = \frac{1}{30}n(n+1)(2n+1)(3n^2+3n-1)$

2. Näytä induktiolla seuraavat summakaavat päteviksi jokaisella $n \in \mathbb{N}$:

   a) $1 \cdot 3 + 3 \cdot 5 + \ldots + (2n-1)(2n+1) = \dfrac{n}{3}(4n^2 + 6n - 1)$

   b) $\displaystyle\sum_{k=1}^{n} k^2 2^k = (n^2 - 2n + 3)2^{n+1} - 6$

   c) $\displaystyle\sum_{k=1}^{n} kq^{k-1} = \frac{1 - (n+1)q^n + nq^{n+1}}{(1-q)^2}, \quad q \in \mathbb{Q},\ q \neq 1$

3. Määritellään palautuvat lukujonot

   a) $a_0 = 1, \quad a_{n+1} = qa_n + 1, \quad n = 0, 1 \ldots$

   b) $a_0 = 2, \quad a_{n+1} = \dfrac{a_n}{2} + \dfrac{1}{a_n}, \quad n = 0, 1 \ldots$

   c) $a_1 = 3,\ a_2 = 6,\ a_{n+1} = \dfrac{na_n + a_{n-1} + 3}{n}, \quad n = 2, 3 \ldots$

   d) $a_0 \in \mathbb{Q},\ a_0 \notin \{-1/n \mid n \in \mathbb{N}\}, \quad a_{n+1} = \dfrac{a_n}{1 + a_n}, \quad n = 0, 1 \ldots$

   Näytä induktiolla, että a) $\{a_n\}$ on geometrinen sarja, b) $1 \leq a_n \leq 2\ \forall n$, c) $a_n < 4n\ \forall n$, d) $a_n = a_0/(1 + na_0)\ \forall n$.

4. Laske seuraavien summalausekkeiden arvot ($n \in \mathbb{N}$):

   a) $\displaystyle\sum_{k=0}^{n} \binom{n}{k}$ b) $\displaystyle\sum_{k=0}^{n} (-1)^k \binom{n}{k}$ c) $\displaystyle\sum_{k=0}^{n} 2^k \binom{n}{k}$ d) $\displaystyle\sum_{k=0}^{n} 3^{-k} \binom{n}{k}$

5. (*) Olkoon $x \in \mathbb{Q},\ x > 0$. Näytä, että jos jollakin $a \in \mathbb{Q}$ ja $n \in \mathbb{N},\ n \geq 2$ pätee $1 < x^n \leq a$, niin $1 < x < 1 + (a-1)/n$. *Vihje*: Kirjoita $x = 1 + y$.

6. (*) a) Näytä, että jos $x \in \mathbb{Q},\ x > 0$ ja $n \in \mathbb{N},\ n \geq 3$, niin $(1+x)^n > 1 + nx + \frac{1}{2}n(n-1)x^2$. b) Millaisia vielä parempia arvioita saadaan, jos $n \geq k,\ k = 4, 5, \ldots$? c) Näytä, että on olemassa $n \in \mathbb{N}$ siten, että pätee

$$\frac{(1 + 10^{-100})^n}{n^{100}} > 10^{100}.$$

7. (*) Näytä, että jos joukot $A_1, A_2, A_3, \ldots$ ovat keskenään pistevieraita ja numeroituvia, niin myös joukko

$$A = A_1 \cup A_2 \cup \ldots = \bigcup_{n=1}^{\infty} A_n$$

   on numeroituva. Totea väitteen pätevyys suoraan (määrittämällä A), kun $A_n = \{\, x \in \mathbb{Q} \mid n - 1 \leq |x| < n \,\}$.

I.5 Äärettömät desimaaliluvut

Matematiikan arkipäivässä hyvin yleisesti kohdattava lukujono on *ääretön desimaaliluku*† (engl. infinite decimal). Tämä on ilmiasultaan merkkijono muotoa

$$ed_{-m}d_{-m+1}\dots d_0.d_1d_2d_3\dots \;=\; x_0.d_1d_2d_3\dots\,,$$

missä

- 'e' on *etumerkki*, joko $e = +$ tai $e = -$ ('tyhjä' $= +$)
- '.' on *desimaalipiste* (tai 'desimaalipilkku', d_0:n ja d_1:n välissä)
- $m \in \mathbb{N} \cup \{0\}$
- $x_0 \;=\; ed_{-m}d_{-m+1}\dots d_0 \in \mathbb{Z} \;=$ desimaaliluvun *kokonaislukuosa*
- $d_n,\;\; n = 1, 2, \dots \;=$ desimaaliluvun *desimaalit* ($d_n \in \{0, \dots, 9\}$)

Esimerkki 1

$$\begin{array}{ll} 141.42135623730950488\dots & (\,e = +,\; m = 2,\; x_0 = 141\,) \\ -0.00314159265358979\dots & (\,e = -,\; m = 0,\; x_0 = 0\,) \end{array} \qquad \square$$

Ääretön desimaaliluku on siis pääosin numerojono, jossa numeroiden lisäksi on erillinen etumerkki ja desimaalipiste, jonka sijainti numerojonossa määrittää kokonaislukuosan x_0. Tällaisen merkkijonon <u>tulkitaan</u> <u>tarkoittavan</u> rationaalilukujonoa

$$\{\, x_0, x_1, \dots \} = \{\, x_n \,\}_{n=0}^{\infty}\,,$$

missä x_0 on kokonaislukuosa ja

$$x_n \;=\; x_0 \pm \big(\, d_1 \cdot 10^{-1} + \dots + d_n \cdot 10^{-n} \,\big) \;=\; x_0 \pm \sum_{k=0}^{n} d_k \cdot 10^{-k}, \quad n \in \mathbb{N},$$

missä summalausekkeen etumerkki valitaan e:n mukaan. Tulkinnan mukaisesti ääretön desimaaliluku on siis itse asiassa <u>sarja</u>, jonka ensimmäinen termi on x_0 ja muut $\pm d_k \cdot 10^{-k}$, $k = 1, 2, \dots$ Sarjan osasummat ovat indeksistä $n = 1$ lähtien nk. *äärellisiä* (tai päättyviä, ks. alaviite) *desimaalilukuja*, jotka jatkossa merkitään

$$x_n \;=\; x_0.d_1\dots d_n \quad (n \in \mathbb{N}).$$

†Vaihtoehtoinen nimitys on *päättymätön* desimaaliluku.

Tällä tarkoitetaan siis lukua $\pm(|x_0|+\sum_{k=0}^{n} d_k \cdot 10^{-k})$, missä etumerkki on erikseen asetettu (sama kuin x_0:ssa, jos $x_0 \neq 0$). Kirjoittamalla luku muotoon

$$x_n \;=\; \pm 10^{-n}\left(|x_0|\cdot 10^n + \sum_{k=0}^{n} d_k \cdot 10^{n-k}\right)$$

nähdään, että $x_n = p/10^n$, missä $p = ed_{-m}\ldots d_0 d_1 \ldots d_n \in \mathbb{Z}$ on kokonaisluku. Toisaalta jos lähtökohtana on luku $x = p/10^n,\ p \in \mathbb{Z}$, niin esittämällä p kymmenjärjestelmässä nähdään, että x on saatettavissa muotoon $x = x_0.d_1 \ldots d_n$. Siis ko. muotoa olevat äärelliset desimaaliluvut muodostavat täsmälleen $\mathbb{Q}$:n osajoukon

$$\mathbb{Q}_n = \{\, p/10^n \mid p \in \mathbb{Z} \,\}.$$

Kun tämän mukaisesti asetetaan $\mathbb{Q}_0 = \mathbb{Z}$, niin pätee

$$\mathbb{Z} = \mathbb{Q}_0 \subset \mathbb{Q}_1 \subset \mathbb{Q}_2 \subset \ldots \subset \mathbb{Q}.$$

Lukujoukossa $\mathbb{Q}_n$ luvut ovat tasavälein siten, että peräkkäisten lukujen x, y etäisyys on $|x - y| = 10^{-n}$. Kuvassa on lukujoukkoja $\mathbb{Q}_0, \mathbb{Q}_1, \mathbb{Q}_2$ havainnollistettu geometrisesti.†

Jatkossa äärettömiin desimaalilukuihin viitataan joko tavanomaisilla lukusymboleilla x, y, z jne., tai vaihtoehtoisesti symboleilla $\breve{x}, \breve{y}, \breve{z}$ jne., kun halutaan tehdä ero 'oikeisiin' lukuihin, erityisesti rationaalilukuihin. Äärettömien desimaalilukujen joukkoa merkitään symbolilla $\mathbb{D}$:

$$\mathbb{D} = \{\text{äärettömät desimaaliluvut}\}.$$

Yksittäisen desimaaliluvun tapauksessa pidetään merkkijonoesitystä ko. desimaaliluvun 'nimenä', jolloin merkinnässä

$$\breve{x} \;=\; x_0.d_1 d_2 \ldots \;=\; \{\, x_n \,\}_{n=0}^{\infty}$$

ensimmäinen '=' nimeää ja toinen tulkitsee (antaa merkityssisällön).

†Lukujen 'geometrisointia' tarkastellaan myöhemmin Luvussa II.1. Tässä vaiheessa pidetään geometrisen havainnollistamisen perusideoita tunnettuina.

Esimerkki 1 (jatko) Tehdyin merkintäsopimuksin voidaan kirjoittaa

$$\begin{aligned} x &= 141.421356\ldots \\ &= \{\,141, 141.4, 141.42, \ldots\} \\ &= \{\,141,\ \frac{1414}{10},\ \frac{14142}{100},\ \ldots\} \\ &= \{\,x_0,\ x_1,\ x_2,\ \ldots\} \end{aligned}$$

$$\begin{aligned} y &= -0.00314159\ldots \\ &= \{\,0,\ 0.0,\ 0.00,\ -0.003,\ -0.0031,\ \ldots\} \\ &= \{\,0,\ 0,\ 0,\ -\frac{3}{1000},\ -\frac{31}{10000},\ \ldots\} \\ &= \{\,y_0,\ y_1,\ y_2,\ \ldots\} \qquad \square \end{aligned}$$

Jokainen ääretön desimaaliluku $x \in \mathbb{D}$ on siis merkitykseltään lukujono $\{x_n\}_{n=0}^{\infty}$, missä $x_n \in \mathbb{Q}_n$ $\forall n$. Määritelmänsä mukaisesti lukujono $\{x_n\}$ on määrättävissä asettamalla ensin $x_0 \in \mathbb{Z}$ ja sitten $x_1, x_2, \ldots$ palautuvasti muodossa

$$x_n = \begin{cases} x_{n-1} + d_n \cdot 10^{-n}, & \text{jos } e = +\,, \\ x_{n-1} - d_n \cdot 10^{-n}, & \text{jos } e = -\,, \end{cases}$$

missä e on x:n etumerkki ja $d_n \in \{0, \ldots, 9\}$ $\forall n \in \mathbb{N}$. Tämän mukaan on joko $x_n \geq x_{n-1}$ $\forall n \in \mathbb{N}$ tai $x_n \leq x_{n-1}$ $\forall n \in \mathbb{N}$, riippuen etumerkistä e. Tällaiset 'yksitoikkoiset' jonot ovat huomattavan tärkeä lukujonojen erikoistapaus, joten asetetaan niitä varten

Määritelmä I.5.1 Lukujono $\{\,a_n,\ n = m, m+1, \ldots\}$ on

- *kasvava*, jos $a_{n+1} \geq a_n$ $\forall n \geq m$
- *vähenevä*, jos $a_{n+1} \leq a_n$ $\forall n \geq m$

Lukujono on *aidosti* kasvava, jos $a_{n+1} > a_n$ $\forall n \geq m$ ja aidosti vähenevä, jos $a_{n+1} < a_n$ $\forall n \geq m$. Jos lukujono on (aidosti) kasvava tai vähenevä, se on (aidosti) *monotoninen.*

Ääretön desimaaliluku on siis lukujonona monotoninen: kasvava etumerkin ollessa $e = +$ ja vähenevä, kun etumerkki on $e = -$. Seuraavan tuloksen mukaan tämä kasvu tai väheneminen kuitenkin 'hiipuu' nopeasti n:n kasvaessa.

PROPOSITIO I.5.2 Jos $x = \{ x_n \}_{n=0}^{\infty} \in \mathbb{D}$, niin jokaisella $k \in \mathbb{N} \cup \{0\}$ ja jokaisella $n \in \mathbb{N}$, $n > k$ pätee $|x_n - x_k| < 10^{-k}$, tarkemmin

$$|x_n - x_k| \;\leq\; 10^{-k} - 10^{-n}.$$

<u>Todistus</u> Kun $n > k \geq 0$, on

$$\begin{aligned} |x_n - x_k| \;=\; \sum_{j=k+1}^{n} d_j \cdot 10^{-j} \;&\leq\; \sum_{j=k+1}^{n} 9 \cdot 10^{-j} \\ &=\; 9 \cdot 10^{-k-1} \sum_{i=0}^{n-k-1} \left(\frac{1}{10}\right)^i \\ &=\; 10^{-k} \cdot \frac{9}{10} \cdot \frac{1 - (\frac{1}{10})^{n-k}}{1 - \frac{1}{10}} \\ &=\; 10^{-k} - 10^{-k} \left(\frac{1}{10}\right)^{n-k} \\ &=\; 10^{-k} - 10^{-n}. \quad \square \end{aligned}$$

Kun Proposition I.5.2 tuloksen ohella huomioidaan lukujonon $\{x_n\}$ monotonisuus, niin nähdään, että jokaiselle äärettömälle desimaaliluvulle $\breve{x} = \{x_n\}$ pätee

$$0 \leq k < n \quad \Rightarrow \quad \begin{cases} x_k \leq x_n < x_k + 10^{-k}, & \text{jos etumerkki } e \;=\; +\,, \\ x_k - 10^{-k} < x_n \leq x_k, & \text{jos etumerkki } e \;=\; -\,. \end{cases}$$

Tämän mukaan termien x_n liikkumavara supistuu nopeasti indeksin n kasvaessa.

Jakokulma-algoritmi

Edellä on todettu, että äärettömät desimaaliluvut ovat lukujonoina $\{x_n\}$ sekä monotonisia että 'hiipuvia'. Voi myös käydä niin, että jono $\{x_n\}$ ei ainoastaan 'hiivu', vaan *lähestyy rationaalilukua* $x \in \mathbb{Q}$ siinä mielessä, että $|x_n - x|$ tulee yhä pienemmäksi n:n kasvaessa. Milloin tällainen 'lähestyminen' on yleensä mahdollista, selvitetään myöhemmissä luvuissa. Tässä vaiheessa asetetaan kysymys toisin päin: Jos on annettu $x \in \mathbb{Q}$, niin miten löydetään ääretön desimaaliluku $\breve{x} = \{x_n\}$ siten, että $|x_n - x|$ tulee yhä pienemmäksi n:n kasvaessa? Vastauksen antaa tunnettu *jakokulma-algoritmi*†. Algoritmissa konstruoidaan desimaalivastine luvulle $|x|$ ja varustetaan tämä x:n mukaisella etumerkillä, joten voidaan

† *Algoritmi* on täsmällinen ja toteutuskelpoinen (ohjelmointikelpoinen) toimintaohje jonkin laskennallisen päämäärän saavuttamiseksi. Jos päämääränä on lukujono $\{ a_n,\ n = 1, 2, \ldots \}$, on algoritmia seuraamalla saatava äärellisellä määrällä laskutoimituksia vastaus kysymykseen $a_n = ?$, olipa indeksi $n \in \mathbb{N}$ mikä hyvänsä annettu luku.

rajoittua tapaukseen $x \geq 0$. Valitaan tällöin $x_n \in \mathbb{Q}_n,\ n = 0, 1, 2, \ldots$ siten, että jokaisella n toteutuu ehto

$$x_n \leq x < x_n + 10^{-n}. \tag{1}$$

Jos x ei ole äärellinen desimaaliluku, ts. $x \notin \mathbb{Q}_n\ \forall n$, niin ehdon (1) mukaisesti on $x_n < x < x_n + 10^{-n}\ \forall n$. Jos $x \in \mathbb{Q}_k$ jollakin k, niin $x_n \in \mathbb{Q}_n$ myös kun $n > k$, jolloin ehdosta (1) seuraa, että $x_n = x$ jokaisella $n \geq k$. Molemmissa tapauksissa $x_n + 10^{-n}$ on x:ää lähinnä oleva, x:ää aidosti suurempi luku joukossa $\mathbb{Q}_n$. On ilmeistä, että näin konstruoitu lukujono $\{x_n\}$ on yksikäsitteinen, toteuttaa

$$|x_n - x| \ < \ 10^{-n}, \quad n = 0, 1, \ldots \tag{2}$$

ja määräytyy yksikäsitteisesti myös tästä ehdosta x:n etumerkistä riippumatta.

Jos $x \geq 0$, niin ehdon (1) toteuttavan lukujonon konstruoiminen jakokulma-algoritmilla käy seuraavasti: Olkoon $x = r/s$, missä $r \in \mathbb{N} \cup \{0\}$ ja $s \in \mathbb{N}$. Tällöin voidaan ensinnäkin kirjoittaa

$$\frac{r}{s} = x_0 + \frac{r_1}{s}\,,$$

missä $x_0 \in \mathbb{N} \cup \{0\}$ ja $r_1 \in \{\,0, \ldots, s-1\,\}$. Luvut x_0 ja r_1 ovat ilmeisen yksikäsitteiset. Koska $0 \leq r_1/s < 1 \ \Rightarrow\ 0 \leq 10\,r_1/s < 10$, on edelleen löydettävissä yksikäsitteiset $d_1 \in \{\,0, \ldots, 9\,\}$ ja $r_2 \in \{\,0, \ldots, s-1\,\}$ siten, että

$$\frac{10\,r_1}{s} = d_1 + \frac{r_2}{s}\,.$$

Jatkamalla näin saadaan yleiseksi algoritmiksi

$$\boxed{\frac{10\,r_k}{s} = d_k + \frac{r_{k+1}}{s}, \quad k = 1, 2, \ldots \qquad \text{(jakokulma-algoritmi)}}$$

Tässä $d_k \in \{\,0, \ldots, 9\,\}$ ja luvut $r_{k+1} \in \{\,0, \ldots, s-1\,\}$ ovat nk. *jakojäännöksiä.* Jos algoritmi katkaistaan indeksiin $k = n$, on luvulle x saatu esitysmuoto

$$\begin{aligned} x \ = \ x_0 + \frac{r_1}{s} \ &= \ x_0 + d_1 \cdot 10^{-1} + 10^{-1} \cdot \frac{r_2}{s} \\ &\ \ \vdots \\ &= \ (\,x_0 + d_1 \cdot 10^{-1} + \ldots + d_n \cdot 10^{-n}\,) + 10^{-n} \cdot \frac{r_{n+1}}{s} \\ &= \ x_n + 10^{-n} \cdot \frac{r_{n+1}}{s}\ . \end{aligned}$$

Tässä on $0 \leq 10^{-n} r_{n+1}/s < 10^{-n}$, joten ehto (1) toteutuu algoritmin määrittämälle lukujonolle $\{ x_0, x_0.d_1, x_0.d_1d_2, \ldots \} = \{x_n\}$.

Jos $\breve{x} \in \mathbb{D}$ on jakokulma-algoritmin tuottama desimaalilukuvastine rationaaliluvulle x, on tapana muitta mutkitta kirjoittaa $x = \breve{x}$, eli samastaa x ja $\breve{x}$. Jatkossa noudatetaan tätä sopimusta. Tämän mukaisesti siis jokaiselle rationaaliluvulle luotu 'kopio' äärettömien desimaalilukujen joukossa, eli pätee $\mathbb{D} \supset \mathbb{Q}$.

Esimerkki 2 Tehdyn sopimuksen mukaan

$$\begin{aligned} 50/3 &= 16.6666666666666666666666666666666666666\ldots \in \mathbb{D}, \\ 311/125 &= 2.4880000000000000000000000000000000000\ldots \in \mathbb{D}, \\ 245/17 &= 14.4117647058823529411764705882352941176 4\ldots \in \mathbb{D}. \qquad \square \end{aligned}$$

Esimerkissä luku 311/125 on äärellinen desimaaliluku. Tällaisessa tapauksessa jakojäännös $r_k = 0$ jollakin k (tässä kun $k = 4$), jolloin algoritmi antaa $r_n = d_n = 0$ kun $n \geq k$. Muissakin esimerkeissä algoritmin antama desimaaliluku on *jaksollinen*, eli jollakin $k, m \in \mathbb{N}$ muotoa

$$\breve{x} \;=\; x_0.d_1d_2\ldots d_kd_{k+1}\ldots d_{k+m}d_{k+1}\ldots d_{k+m}\ldots,$$

jolloin jaksollisuus alkaa indeksistä $k+1$ ja jakson pituus $= m$. Jakokulma-algoritmi antaa itse asiassa aina jaksollisen tuloksen, sillä koska jakojäännöksellä r_n on vain $s-1$ erilaista 0:sta poikkeavaa arvoa, niin jollakin $n = k+1 \in \{ 1, \ldots, s \}$ ja $m \in \{ 1, \ldots, s-1 \}$ on oltava $r_{n+m} = r_n$. Tällöin algoritmi antaa $r_{n+m} = r_n$ ja $d_{n+m} = d_n$ myös kun $n > k+1$, jolloin tulos on ym. jaksollista muotoa. Jaksollisuus alkaa siis viimeistään indeksistä $n = s$ ja jakson pituus on enintään $m = s-1$.

Jatkossa merkitään jaksollisten desimaalilukujen joukkoa symbolilla $\mathbb{D}_p$:

$$\mathbb{D}_p = \{\text{jaksolliset desimaaliluvut}\}.$$

Jos $x \in \mathbb{D}$ ja $x \notin \mathbb{D}_p$, niin sanotaan, että x on *jaksoton*.

Äärettömät binaariluvut

Äärettömillä binaariluvuilla tarkoitetaan merkki/lukujonoja muotoa

$$x \;=\; eb_mb_{m+1}\ldots b_0.b_1b_2\ldots \;=\; \{x_n\}_{n=0}^{\infty},$$

missä b_n:t ovat binaarijärjestelmän numeroita eli *bittejä*, $b_n \in \{ 0, 1 \}$, ja

$$x_n \;=\; x_0 \pm \sum_{k=1}^{n} b_k \cdot 2^{-k},$$

missä $x_0 = eb_m b_{m+1} \ldots b_0$ on kokonaislukuosa (binaarimuodossa) ja summalausekkeen etumerkki on sama kuin x:ssä. Jos lukujärjestelmän kantalukuna on 8 tai 16, puhutaan vastaavasti (äärettömistä) *oktaali-* ja *heksadesimaaliluvuista.* Tietokoneiden ja laskinten käyttäminä lukujärjestelminä nämä mainitut ovat tavallisia. — Kaikki mitä edellä on sanottu äärettömistä desimaaliluvuista, on helposti muunnetavissa mainittuja (ja muitakin) lukujärjestelmiä koskevaksi.

Desimaaliluvun skaalaus

Toistaiseksi äärettömille desimaaliluvuille ei määritellä muita laskuoperaatioita kuin *merkin vaihto* ja *skaalaus* = kertominen luvuilla 10^m, $m \in \mathbb{Z}$. Skaalaus luvulla 10^m tarkoittaa yksinkertaisesti desimaalipisteen paikan siirtämistä joko m desimaalia eteenpäin ($m > 0$) tai $|m|$ numeroa taaksepäin ($m < 0$). Jälkimmäisessä tapauksessa lisätään kokonaislukuosaan tarvittaessa riittävä määrä 'etunollia'. Luvulla $10^0 = 1$ skaalattaessa jää desimaaliluku ennalleen. Skaalaukselle pätevät vaihdanta- ja liitäntälait

$$10^m \cdot (10^n \cdot x) \;=\; 10^{m+n} \cdot x \;=\; 10^n \cdot (10^m \cdot x), \quad m, n \in \mathbb{Z}.$$

Kun merkin vaihto tulkitaan samaksi kuin kertominen luvulla -1 (kuten kuntaalgebrassa), niin merkin vaihto ja skaalaus yhdistämällä tulee määritellyksi skaalaus luvuilla $\pm 10^m$, $m \in \mathbb{Z}$.

Esimerkki 1 (jatko)

$$\begin{aligned} 10^{-4} \cdot 141.421356\ldots &= 0.0141421356\ldots \\ -10^4 \cdot (-0.00314159\ldots) &= 31.4159\ldots \end{aligned}$$

Liukuluvut

Liukuluvuiksi (engl. floating point number, ruots. flyttal) sanotaan kymmenjärjestelmässä äärellisiä desimaalilukuja muotoa

$$\pm 10^\beta \cdot 0.d_1 \ldots d_n,$$

missä $\beta \in \mathbb{Z}$ ja $n \in \mathbb{N}$ on <u>kiinteä</u>. Numeerisessa laskennassa liukulukuja käytetään äärettömien desimaalilukujen *approksimaatioina*, eli likimääräisinä vastineina. Approksimointi tapahtuu kirjoittamalla desimaaliluku ensin skaalaustekijän avulla muotoon $x = \pm 10^\beta \cdot 0.d_1 d_2..$, missä $\beta \in \mathbb{Z}$ ja $d_1 \neq 0$ (mahdollista aina kun $x \neq 0$) ja katkaisemalla näin saatu luku desimaaliin d_n. Sanotaan, että näin

saadut numerot $d_1 \dots d_n$ ovat desimaaliluvun n ensimmäistä *merkitsevää numeroa* (engl. significant digits), ja että ko. liukuluku on desimaaliluvun *katkaistu* (engl. truncated) muoto. Katkaisun vaihtoehtona approksimaatiossa voidaan käyttää *pyöristystä* (engl. rounding), yleensä lähimpään samaa tyyppiä olevaan liukulukuun (nk. normaalipyöristys). Tieto desimaalipisteen paikasta suhteessa ensimmäiseen merkitsevään numeroon sisältyy lukuun $\beta \in \mathbb{Z}$. Kun ääretön desimaaliluku ilmoitetaan *n merkitsevän numeron tarkkuudella*, tarkoitetaan luvun (katkaisutua tai pyöristettyä) liukulukuapproksimaatiota.

Esimerkki 3 Yhdeksän ($n = 9$) merkitsevän numeron tarkkuudella on

$$\begin{aligned} 0.00014142135623730 95\dots &\approx 10^{-3} \cdot 0.141421356 \\ -172/3 \;=\; -57.333333333\dots &\approx -10^{2} \cdot 0.573333333 \end{aligned}$$

Tässä katkaisu ja normaalipyöristys antavat saman tuloksen. □

Tietokoneet ja laskimet operoivat numeerisissa laskuissa aina liukuluvuilla, esim. binaarijärjestelmään perustuvilla. Tällöin sekä merkitsevien numeroiden määrä että β:n sisältämien numeroiden määrä on rajoitettu koneelle ominaisesti. Mahdollisia *konelukuja* (engl. machine numbers) on siten aina äärellinen määrä.

Esimerkki 4 Koneluvut ovat muotoa $\pm 2^{\beta} \cdot 0.\, b_1 \dots b_{30}$, missä $\beta = \pm\beta_1 \dots \beta_{10}$, $b_k, \beta_k \in \{\, 0, 1 \,\}$. Yhden luvun tallettamiseen tarvitaan 42 bittiä (numerot ja etumerkit), ja erilaisia lukuja on

$$2^{30} \cdot (2^{11} - 1) + 1 \;=\; 2197949513729 \;\text{ kpl.} \qquad \square$$

Tietokoneen/laskimen suorittamia peruslaskuoperaatioita (yhteen-, vähennys-, kerto- ja jakolasku) koneluvuilla sanotaan *liukulukuoperaatioiksi* (engl. floating point operation, lyh. flop). Nämä poikkeavat hiukan tavanomaisista laskuoperaatioista, koska tulos on yleensä (etenkin kerto- ja jakolaskuissa) katkaistava tai pyöristettävä koneluvuksi. Laskutoimitusten likimääräisyydestä aiheutuvat *pyöristysvirheet* (engl. roundoff errors) voivat vahvistua tai kasautua peräkkäisissä laskutoimituksissa. Yksittäisessäkin laskutoimituksessa voi tapahtua kohtalokas *merkitsevien numeroiden kato*, kuten seuraava yksinkertainen esimerkki osoittaa.

Esimerkki 5 Olkoon a ja b annettuja lukuja. Halutaan laskea

$$x = \frac{1}{a}\left(1 - \frac{1}{1+b}\right).$$

Mikä arvo saadaan x:lle, jos $a = 10^{-20}$, $b = 2 \cdot 10^{-20}$, ja koneluvut ovat muotoa $\pm 10^{\beta} \cdot 0.d_1 \dots d_{15}$ (kymmenjärjestelmä), missä $-99 \le \beta \le 99$?

Ratkaisu Suoraan annettuun lausekkeeseen perustuva algoritmi on:

$$y = 1 + b, \quad z = 1/y, \quad u = 1 - z, \quad x = u/a.$$

Koska y:n tarkka arvo on $y = 1.00000000000000000002$, niin kone sijoittaa muistipaikkaan 'y' tämän luvun katkaistun (tai pyöristetyn) liukulukumuodon $y = 10^1 \cdot 0.100000000000000 = 1$. Laskun muut vaiheet ovat tämän jälkeen virheettömiä, ja lopputulokseksi saadaan $x = 0$. □

Esimerkissä vaarattoman näköinen pyöristys johtaa suureen virheeseen lopputuloksessa. Ongelma poistuu, jos laskut suoritetaan *kaksoistarkkuudella* (engl. double precision), eli käyttämällä liukulukuesitystä, jossa talletetaan 15:n sijasta 30 merkitsevää numeroa. Tällöin saadaan tulos $x = 2$, joka on varsin lähellä tarkkaa arvoa $x = 1.99999999999999999996000000000000000000799999\ldots$ Kaksoistarkkuuteen (tai useampikertaiseen tarkkuuteen) siirtyminen on yleinen menettely pyöristysvirheiden uhatessa. Toinen vaihtoehto on algoritmin parempi suunnittelu.

Esimerkki 5 (jatko) Laskettava lauseke yksinkertaistuu kunta-algebran keinoin muotoon

$$x = \frac{b}{a(1+b)}\,.$$

Tässä approksimaatio $1 + b \approx 1$ ei aiheuta merkittävää virhettä. □

HARJOITUSTEHTÄVIÄ

1. a) Näytä, että jos luvuilla $p, q \in \mathbb{N}$ ei ole yhteisiä tekijöitä, niin $x = p/q \in \mathbb{Q}$ on äärellinen desimaaliluku täsmälleen kun on olemassa $m, n \in \mathbb{N} \cup \{0\}$ siten, että $q = 2^m \cdot 5^n$. b) Ilmoita luku $x = (2/5)^{10}$ äärellisenä desimaalilukuna. c) Millaisia rationaalilukuja ovat äärellisten desimaalilukujen vastineet lukujärjestelmässä, jonka kantaluku on 60? d) Luvulla 1111111/5400 on 60-järjestelmässä äärellistä desimaalilukua vastaava esitysmuoto. Määritä tämä käyttämällä luvuille $n = 0 \ldots 59$ symboleja (n) 60-järjestelmässä.

2. Näytä, että jos lukujonot $\{a_n\}$ ja $\{b_n\}$ $(n \in \mathbb{N})$ ovat kasvavia, niin a) myös lukujono $\{a_n + b_n\}$ on kasvava, b) jos lisäksi $a_n \geq 0\ \forall n$ ja $b_n \geq 0\ \forall n$, niin lukujono $\{a_n b_n\}$ on kasvava. c) Näytä esimerkillä, että jos b)-kohdassa lisäehto $b_n \geq 0\ \forall n$ poistetaan, niin lukujono $\{a_n b_n\}$ voi olla aidosti vähenevä.

3. Desimaaliluvuista $x = \{x_n\} \in \mathbb{D}$ ja $y = \{y_n\} \in \mathbb{D}$ tiedetään, että $x = 23.4569\ldots$ ja $y = 23.4567\ldots$ Kuinka suuri luku $x_n - y_n$ on tämän tiedon perusteella vähintään ja kuinka suuri enintään, kun $n = 10$?

4. Muotoile ja todista Proposition I.5.2 vastine äärettömille binaariluvuille.

5. Laske jakokulma-algoritmin avulla seuraavien rationaalilukujen vastineet jaksollisina desimaalilukuina: a) $1/11$, b) $31/13$, c) $19/17$.

6. Jakokulma-algoritmin mukainen luvun $x \in \mathbb{Q}$ vastine äärettömänä desimaalilukuna olkoon jaksollinen siten, että jakson pituus $= m$ ja jaksollisuus alkaa desimaalista d_k. Anna esimerkki luvusta $x = r/s$, $r, s \in \mathbb{N}$, jolle pätee a) $m = 1$, $k = 5$, b) $m = s - 1$, $k = s$, c) $m = s - 1$, $k = 1$.

7. Kymmenjärjestelmän luku $7/11$ samastuu 2-kantaisessa lukujärjestelmässä äärettömään binaarilukuun ja 3-kantaisessa järjestelmässä äärettömään triaarilukuun. Määritä nämä jakokulma-algoritmilla. *Huomaa*, että jakojäännöstä ei algoritmissa kerrota kymmenellä vaan kantaluvulla, siis 2:lla tai 3:lla. Jatka laskua, kunnes jaksollisuus ilmenee!

8. Halutaan laskea $y = (x + a)^2 - (x + b)^2$ kun $x = 1/7$, $a = 0.1457 \cdot 10^{-6}$ ja $b = 0.6973 \cdot 10^{-6}$. a) Määritä y neljän merkitsevän numeron tarkkuudella muokkaamalla y:n lauseke ensin sellaiseen muotoon, että merkitsevien numeroiden katoa ei tapahdu. b) Jos käytetään alkuperäistä lauseketta, niin millä tarkkuudella laskut on suoritettava, jotta päästään mainittuun tarkkuuteen lopputuloksessa?

9. Jos $x = 7 \cdot 10^{-8}$, niin monenko merkitsevän numeron tarkkuudella on

$$\text{a)} \quad \frac{1}{8 + 3x} \approx 0.125 \qquad \text{b)} \quad (1 + x)^3 \approx 1 + 3x \ ?$$

10. Muokkaa seuraavat lausekkeet niin, että alttius merkitsevien numeroiden kadolle poistuu. Laske sitten lausekkeiden arvot kymmenen merkitsevän numeron tarkkuudella.

$$\text{a)} \quad 10^{20}\left(\frac{1}{1+x} - \frac{1}{1+2x}\right), \quad x = \frac{6}{7} \cdot 10^{-20}$$

$$\text{b)} \quad 10^{40}\left(\frac{1}{1+x} - \frac{1}{1-2x} + 3x\right), \quad x = \frac{2}{7} \cdot 10^{-20}$$

$$\text{c)} \quad 10^{60}\left(\frac{1+x}{1-x} - \frac{1-x}{1+x} - 4x\right), \quad x = \frac{5}{7} \cdot 10^{-20}$$

$$\text{d)} \quad 10^{80}\left(\frac{1+x}{1-x} + \frac{1-x}{1+x} + \frac{2}{1+2x^2} - 4\right), \quad x = \frac{3}{7} \cdot 10^{-20}$$

$$\text{e)} \quad 10^{120}\left(\frac{1+x}{1-x} + \frac{1-x}{1+x} - \frac{2}{1-2x^2+2x^4}\right), \quad x = \frac{4}{7} \cdot 10^{-20}$$

11. (*) (Mummon nelilaskin) Nelilaskin pystyy näyttämään luvun, jossa on kahdeksan numeroa, etumerkki ja mahdollinen desimaalipiste numeroiden välissä, esim. -0.0000999. Montako erilaista rationaalilukua laskin kykenee näyttämään tarkasti?

12. (*) Olkoon $x = \{x_n\} \in \mathbb{D}$ ja $y = \{y_n\} \in \mathbb{D}$. Näytä, että jos $x_k < y_k$ jollakin $k \in \mathbb{N} \cup \{0\}$, niin a) $x_n < y_n$ jokaisella $n \geq k$, b) $x_n \leq y_n$ jokaisella n.

13. (*) Ikuisesti laskeva tietokone tuottaa bittijonon $\{b_n\}_{n=1}^{\infty}$ seuraavalla algoritmilla:

    1. Asetetaan $A = 0$, $B = 1$, $n = 0$.
    2. Asetetaan

$$\begin{aligned}
&n \leftarrow n + 1 \text{ (uusi arvo=vanha arvo + 1)}\\
&C = \frac{1}{2}(A + B) \quad \text{ja}\\
&b_n = 1 \text{ ja } A = C, \quad \text{jos } 3 \cdot C \leq 1\\
&b_n = 0 \text{ ja } B = C, \quad \text{jos } 3 \cdot C > 1
\end{aligned}$$

    Ilmoitetaan b_n ja palataan kohtaan 2.

    a) Millaiset luvut b_n, $n = 1 \ldots 4$, kone ilmoittaa?
    b) Jos laskun lopputulos (iäisen odottelun jälkeen) tulkitaan äärettömänä binaarilukuna $\breve{x} = 0.b_1 b_2 \ldots$, niin minkä rationaaliluvun x binaariesitys on kyseessä? Perustele!

I.6 Lukujonon raja-arvo

Tässä ja seuraavissa luvuissa tarkastelun kohteena ovat järjestetyn kunnan lukujonot $\{a_n\}$. Oletetaan siis, että $a_n \in \mathbb{K}\ \forall n$, missä $(\mathbb{K}, +, \cdot, <)$ on järjestetty kunta; tämä voi olla joko rationaalilukujen kunta tai jokin tämän kuntalaajennus, jossa myös järjestysrelaatio on määritelty. Esimerkiksi voi olla $\mathbb{K} = \mathbb{J}$, vrt. Luku I.2.

Lukujonon $\{a_n\}$ sanotaan *suppenevan* eli *konvergoivan* (engl. converge) kohti lukua a, jos 'n:n kasvaessa a_n tulee yhä lähemmäksi a:ta'. Täsmällisepi määritelmä seuraa hieman tuonnempana; tässä vaiheessa riittäköön toteamus, että suppeneminen tarkoittaa esimerkiksi seuraavaa:

$$\begin{aligned}
|a_n - a| &< 10^{-1} \quad \text{indeksistä } n = N_1 \text{ alkaen,}\\
|a_n - a| &< 10^{-2} \quad \text{indeksistä } n = N_2 \text{ alkaen,}\\
&\vdots\\
|a_n - a| &< 10^{-k} \quad \text{indeksistä } n = N_k \text{ alkaen,}\\
&\vdots
\end{aligned}$$

Tässä jokainen N_k on äärellinen, eli $N_k \in \mathbb{N}$ jokaisella $k \in \mathbb{N}$. Suppenemista on havainnollistettu alla olevissa kuvissa graafisesti†.

Jos lukujono suppenee kohti a:ta, sanotaan lukua a

jonon *raja-arvo*ksi (engl. limit). Myös termiä *limes* käytetään (lat. limes = raja). Merkintätapoja ovat

$$\lim_{n\to\infty} a_n = a \qquad \text{tai} \qquad \lim_{n} a_n = a$$

†Lukujen graafisen (geometrisen) havainnollistamisen menetelmiä pidetään jälleen tuttuina. Asiaan palataan myöhemmin Luvussa II.

tai kuten jatkossa useammin:

$$a_n \to a \quad (\text{kun} \;\; n \to \infty).$$

Merkinnöissä 'lim' luetaan 'limes' ja symboli '$\to$' luetaan 'suppenee kohti', 'menee kohti' tai 'lähestyy'.

Suppeneminen on lukujonolle varsin voimakas vaatimus, ja onkin helppo esittää esimerkkejä jonoista, jotka eivät suppene kohti mitään lukua. Tällainen on vaikkapa *rajatta kasvava* lukujono $\{a_n\} = \{n\}$. Tässä tapuksessa on tapana kirjoittaa

$$a_n \to \infty \;\; \text{kun} \;\; n \to \infty \qquad \text{tai jopa:} \quad \lim_n a_n = \infty.$$

Luvallisia luku- ja puhetapoja ovat 'a_n lähestyy ääretöntä' tai 'raja-arvo on ääretön'. Näiden sanontojen mukaisesti merkintä '$\lim_{n\to\infty}$' luetaan yleensä 'limes n lähestyy ääretöntä', tarkoittaen siis raja-arvoa, kun n kasvaa rajatta. Vastaavasti jos lukujono on *rajatta vähenevä*, voidaan kirjoittaa $a_n \to -\infty$ tai $\lim_n a_n = -\infty$, ja lukea 'raja-arvo on miinus ääretön'. Näistä merkintä- ja puhetavoista huolimatta rajatta kasvava tai rajatta vähenevä lukujono ei ole suppeneva.

Esimerkki 1 Etsi mahdollinen raja-arvo rationaalilukujonolle $\{a_n\}_{n=1}^{\infty}$, kun

a) $a_n = n^2$ b) $a_n = 100 - n$ c) $a_n = (-1)^n$

d) $a_n = (n+1)/n$ e) $a_n = (10^{100} + 2n)/(10^{100} + n)$

<u>Ratkaisu</u> a) Kyseessä on rajatta kasvava jono: $a_n \to \infty$. Raja-arvoa ei siis ole.

b) Tämä jono on rajatta vähenevä: $a_n \to -\infty$. Raja-arvoa ei ole.

c) Jonon termit saavat vuorotellen arvoja $+1$ ja -1, mutta eivät 'lähesty' kumpaakaan näistä tai mitään muutakaan lukua. Päätellään, että tälläkään jonolla ei ole raja-arvoa.

d) Tässä on ensimmäinen suppeneva jono: Kirjoittamalla

$$a_n = 1 + \frac{1}{n}$$

nähdään, että n:n kasvaessa termit tulevat yhä lähemmäksi lukua 1, joten päätellään, että $\lim_n a_n = 1$.

e) Tämä jono on myös suppeneva, mutta oikea raja-arvo paljastuu vasta hyvin suurilla indeksin arvoilla: Jos esimerkiksi tutkitaan jonon termejä indeksin arvoilla $n = N, N+1, \ldots, N+10^{10}$, niin saadaan seuraavat tulokset:

$$N \le n \le N + 10^{10} \quad \Rightarrow \quad \begin{cases} a_n \approx 1, & \text{kun } N = 1, \\ a_n \approx 1, & \text{kun } N = 10^{50}, \\ a_n \approx 3/2, & \text{kun } N = 10^{100}, \\ a_n \approx 2, & \text{kun } N = 10^{120}. \end{cases}$$

Vasta viimeinen testi kertoo totuuden: $\lim_n a_n = 2$. $\square$

Esimerkki 1e) paljastaa, että lukujonon raja-arvoa ei välttämättä heti 'näe' tai löydä kokeilemallakaan. Tämäntyyppinen epävarmuus voidaan poistaa vain täsmällisemmällä raja-arvon määritelmällä.

MÄÄRITELMÄ I.6.1 (**Lukujonon raja-arvo**) Lukujono $\{a_n\}$ suppenee kohti raja-arvoa a, jos jokaisella $\varepsilon > 0$[†] on olemassa $N \in \mathbb{N}$ siten, että pätee

$$|a_n - a| \ < \varepsilon, \quad \text{kun } n > N.$$

Määritelmän mahdollisimman 'pakattu' muoto on

$$\lim_n a_n = a \quad \Leftrightarrow \quad \forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ (\, |a_n - a| < \varepsilon \ \ \forall n > N \,).$$

Jos a:n tilalla on ∞, on määritelmässä ehdon $|a_n - a| < \varepsilon$ tilalla oltava ehto $a_n > M$, missä M on mielivaltaisen suuren luvun symboli. Määritelmän tiivis muoto on tällöin

$$\lim_n a_n = \infty \quad \Leftrightarrow \quad \forall M \ \exists N \in \mathbb{N} \ (\, a_n > M \ \ \forall n > N \,).$$

Määritelmässä I.6.1 N riippuu yleensä (jokseenkin aina!) ε:sta. Riippuvuuden voimakkuudella ei ole väliä, kunhan jokaisella positiivisella ε on löydettävissä äärellinen indeksi N (eli luku $N \in \mathbb{N}$) siten, että asetettu ehto on voimassa. Ehto puolestaan merkitsee, että indeksistä N eteenpäin jonon kaikki termit pysyvät enintään ε:n päässä a:sta. Kun ehtoa kiristetään antamalla ε:lle yhä pienempiä mutta positiivisia arvoja, niin N yleensä kasvaa, mutta pysyy aina äärellisenä. Vastaavasti jos $\lim_n a_n = \infty$, niin indeksistä N eteen päin jonon kaikki termit ovat lukua M suurempia. Luvun M kasvaessa kasvaa yleensä myös N, mutta pysyy äärellisenä jokaisella M:n (äärellisellä) arvolla.

LAUSE I.6.2 Jos lukujono $\{a_n\}$ on suppeneva, niin raja-arvo $a = \lim_n a_n$ on yksikäsitteinen.

Todistus Jos $a_n \to a$ ja $b \neq a$, niin kolmioepäyhtälön (Lause I.2.3) ja suppenemisen määritelmän perusteella jokaisella $\varepsilon > 0$ on olemassa $N \in \mathbb{N}$ siten, että pätee

$$\begin{aligned} |a_n - b| \ &= \ | \, (a_n - a) + (a - b) \, | \\ &\geq \ |a - b| - |a_n - a| \\ &> \ |a - b| - \varepsilon, \quad \text{kun } n > N. \end{aligned}$$

Kun tässä valitaan $\varepsilon = \frac{1}{2}|a - b|$ (mahdollista, koska $a \neq b \ \Rightarrow \ |a - b| > 0$), niin kyseisellä ε:n arvolla pätee

$$|a_n - b| > \varepsilon, \quad \text{kun } n > N,$$

[†]Kreikkalainen kirjain ε (luetaan 'epsilon') on matematiikassa hyvin vakiintunut pienen positiivisen luvun symboli. Toinen usein käytetty symboli on myös kreikkalaiseen kirjaimistoon kuuluva δ (luetaan 'delta').

joten suppenemisen määritelmän mukaan $a_n \not\to b$. On näytetty toteen lauseen väittämä muodossa

$$a_n \to a \quad \& \quad b \neq a \quad \Rightarrow \quad a_n \not\to b. \qquad \square$$

Esimerkki 1 (jatko) c) $a_n = (-1)^n$. Jos valitaan mikä tahansa luku a, niin $\forall n \in \mathbb{N}$ pätee

$$\max\{|a_n - a|,\ |a_{n+1} - a|\} \ = \ \max\{|1 - a|,\ |1 + a|\} \ \geq \ 1.$$

Näin ollen jos valitaan $\varepsilon < 1$, niin suppenevuden määritelmän ehto

$$|a_n - a| \ < \varepsilon, \quad \text{kun } n > N$$

ei ole voimassa millään indeksin N arvolla. Siis Määritelmän I.6.1 mukaista raja-arvoa ei ole olemassa.

d) $a_n = (n+1)/n$. Kun $a = 1$, pätee

$$|a_n - a| = \frac{1}{n} < \varepsilon, \quad \text{kun} \ \ n > \frac{1}{\varepsilon}\,.$$

Näin ollen jos valitaan esimerkiksi $N =$ (lukua $1/\varepsilon$ lähinnä suurempi kokonaisluku), niin raja-arvon määritelmässä asetettu ehto toteutuu ko. N:n arvolla. Siis $a_n \to 1$.

e) $a_n = (10^{100} + 2n)/(10^{100} + n)$. Kun valitaan $a = 2$, niin

$$a_n - a = \frac{10^{100} + 2n}{10^{100} + n} - 2 = -\frac{10^{100}}{10^{100} + n} \quad \Rightarrow \quad |a_n - a| < \frac{10^{100}}{n}\,.$$

Näin ollen

$$|a_n - a| < \varepsilon, \quad \text{kun} \ \ n \geq 10^{100} \cdot \varepsilon^{-1}.$$

Siis jos valitaan N esimerkiksi siten, että

$$10^{100} \cdot \varepsilon^{-1} \ \leq \ N \ < \ 10^{100} \cdot \varepsilon^{-1} + 1,$$

niin suppenevuuden määritelmässä asetettu vaatimus $|a_n - a| < \varepsilon$ on voimassa, kun $n > N$. Vaikka N on hyvin suuri (jo kun $\varepsilon = 1$), on N kuitenkin äärellinen jokaisella $\varepsilon > 0$, joten määritelmän mukaan $a_n \to 2$. $\square$

Seuraavaa raja-arvotulosta tarvitaan käytännössä (ja teoriassakin) usein.

Propositio I.6.3 $\displaystyle\lim_{n\to\infty} q^n = \begin{cases} 0, & \text{jos } -1 < q < 1, \\ \infty, & \text{jos } q > 1. \end{cases}$

Todistus Jos $q = 0$, niin $q^n = 0$ $\forall n \in \mathbb{N}$, jolloin $q^n \to 0$ raja-arvon määritelmän mukaan. Jos $0 < q < 1$, niin $q = 1/(1+x)$, missä $x > 0$. Tällöin on Bernoullin epäyhtälön (Propositio I.4.2) perusteella

$$q^n = \frac{1}{(1+x)^n} < \frac{1}{1+nx} < \frac{1}{nx}, \quad n \geq 2.$$

Tässä on jokaisella $\varepsilon > 0$

$$\frac{1}{nx} < \varepsilon, \quad \text{kun} \;\; n > \frac{1}{x\varepsilon},$$

joten valitsemalla $N \in \mathbb{N}$ siten, että $N \geq (x\varepsilon)^{-1}$ (mahdollista aina kun $x > 0$ ja $\varepsilon > 0$), seuraa

$$0 < q^n < \varepsilon, \quad \text{kun } n > N.$$

Raja-arvon määritelmän perusteella $q^n \to 0$. Jos lopulta $-1 < q < 0$, niin $q^n = (-1)^n(-q)^n$, missä $0 < -q < 1$, joten valisemalla N samalla tavoin kuin edellä seuraa

$$-\varepsilon < q^n < \varepsilon, \quad \text{kun } n > N.$$

Jälleen $q^n \to 0$ raja-arvon määritelmän perusteella, joten ensimmäinen väittämä on todistettu. Tapausta $q > 1$ koskeva toinen väittämä todistetaan vastaavalla tavalla, kirjoittamalla ensin $q = 1 + x$ ja käyttämällä Bernoullin epäyhtälöä. □

Lukujonon suppenemisen määritelmästä todettakoon vielä, että siinä asetettua ehtoa 'jokaisella $\varepsilon > 0$' on mahdollista lieventää määritelmän muuttumatta. Esimerkiksi riittää valita ε jonosta $\{10^{-k},\ k = 0, 1, 2, \ldots\}$, jolloin määritelmäksi tulee (vrt. johdattelu luvun alussa): Jokaisella $k = 0, 1, 2, \ldots$ on olemassa indeksi $N_k \in \mathbb{N}$ siten, että $|a_n - a| < 10^{-k}$ kun $n > N_k$, eli:

$$\lim_n a_n = a \quad \Leftrightarrow \quad \forall k \in \mathbb{N} \cup \{0\}\ \exists N_k \in \mathbb{N}\ (\,|a_n - a| < 10^{-k}\ \ \forall n > N_k\,).$$

Tässä voi 10^{-k}:n tilalla olla yhtä hyvin q^k, $0 < q < 1$, tai vielä yleisemmin, ε on mahdollista poimia mistä tahansa aidosti vähenevästä (ks. Määritelmä I.5.1) lukujonosta $\{b_k\}$, kunhan on ensin varmistettu Määritelmään I.6.1 vedoten, että $\lim_k b_k = 0$. Tulos on täsmällisemmin muotoiltuna seuraava. Todistus jätetään harjoitustehtäväksi (Harj.teht. 7a).

Propositio I.6.4 Olkoon $\{b_k,\ k = 1, 2, \ldots\}$ aidosti vähenevä lukujono, jonka raja-arvo on $\lim_k b_k = 0$. Tällöin lukujonon $\{a_n\}$ raja-arvo $= a$ täsmälleen kun jokaisella $k \in \mathbb{N}$ on olemassa indeksi $N_k \in \mathbb{N}$ siten, että $|a_n - a| < b_k$ kun $n > N_k$.

Sarjan summa

Jos sarja $\sum_{n=m}^{\infty} a_n$ suppenee (osasummiensa lukujonona, vrt. Luku I.4), niin raja-arvoa sanotaan *sarjan summaksi*. Raja-arvo merkitään

$$s \;=\; \lim_{n\to\infty} \sum_{k=m}^{n} a_k \;=\; \sum_{k=m}^{\infty} a_k.^{\dagger}$$

Esimerkki 2 Tutki perusmuotoisen geometrisen sarjan suppenemista.

Ratkaisu Perusmuotoinen geometrinen sarja on lukujono

$$\{s_n\} \;=\; \{\sum_{k=0}^{n} q^k\} \;=\; \begin{cases} \{\, n+1,\ n=0,1,\dots\}, & \text{jos } q=1, \\ \{\, (q^{n+1}-1)/(q-1),\ n=0,1,\dots\}, & \text{jos } q\neq 1. \end{cases}$$

Jos $q=1$, niin sarja ei suppene ($s_n \to \infty$). Jos $q=-1$, niin $\{s_n\} = \{\,1,0,1,0\dots\}$, joten sarja ei suppene tässäkään tapauksessa. Jos $|q|>1$, niin arvioidaan kolmioepäyhtälön avulla

$$|s_n| \;=\; \frac{1}{|q-1|}\,|q^{n+1}-1| \;\geq\; \frac{1}{|q-1|}\,\big(|q|^{n+1}-1\big)\,.$$

Proposition I.6.3 avulla päätellään tästä, että $|s_n| \to \infty$, joten tässäkään tapauksessa sarja ei suppene. Jäljelle jääneessä tapauksessa $-1<q<1$ valitaan raja-arvokandidaatiksi $s=1/(1-q)$, jolloin

$$|s_n - s| \;=\; \frac{|q|^{n+1}}{|1-q|}\,.$$

Olkoon $\varepsilon>0$, jolloin on myös $|1-q|\,\varepsilon>0$. Koska $|q|^{n+1}\to 0$ (Propositio I.6.3) ja $|1-q|\,\varepsilon>0$, niin lukujonon raja-arvon määritelmän perusteella on olemassa indeksi $N\in\mathbb{N}$ siten, että

$$|q|^{n+1} < |1-q|\,\varepsilon \quad\Rightarrow\quad |s_n - s| < \varepsilon, \quad \text{kun } n>N.$$

Tässä $\varepsilon>0$ oli mielivaltainen ja $N\in\mathbb{N}$, joten lukujonon raja-arvon määritelmän perusteella $\lim_n s_n = s$. Siis geometrinen sarja suppenee täsmälleen kun $|q|<1$, ja raja-arvo on tällöin

$$\lim_n s_n \;=\; \lim_n \big(\sum_{k=0}^{n} q^k\big) \;=\; \frac{1}{1-q}\,. \qquad \square$$

†Jos sarja on suppeneva, niin asiayhteydestä on selvitettävä, tarkoittaako $\sum_{n=m}^{\infty} a_n$ itse sarjaa vai sen summaa. Esimerkiksi merkintä $\sum_{k=m}^{\infty} a_k = \sum_{k=m}^{\infty} b_k$ voi tarkoittaa joko, että sarjat ovat samat lukujonoina (jolloin ne ovat samat myös termeittäin: $a_k = b_k\ \forall k$), tai ainoastaan, että sarjojen summat ovat samat.

Sarjan summamerkintää käyttäen esimerkin tulos on

$$\boxed{\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad |q| < 1.}$$

Rationaalilukujono, jonka raja-arvo ei ole rationaalinen

Rationaalilukujonon raja-arvon ei tarvitse olla rationaaliluku, sillä Määritelmän I.6.1 mukaisesti riittää, että raja-arvo on löydettävissä lukujoukosta $\mathbb{K} \supset \mathbb{Q}$, jossa laskuoperaatiot ja järjestysrelaatio ovat määriteltyjä niin, että $(\mathbb{K}, +, \cdot, <)$ on rationaalilukujen kunnan laajennus. Seuraavassa on esimerkki rationaalilukujonosta, joka suppenee, kunhan lukujoukkoon $\mathbb{K} \supset \mathbb{Q}$ sisältyy myös luku $a = \sqrt{2}$ (vrt. Luku I.2, Esimerkki 3).

Esimerkki 3 Tarkastellaan palautuvaa rationaalilukujonoa

$$a_0 = 2, \quad a_{n+1} = \frac{a_n}{2} + \frac{1}{a_n}, \quad n = 0, 1, \ldots$$

Palautuskaavasta seuraa, että $a_n > 0 \; \forall n$ (induktio!) ja että

$$a_{n+1}^2 - 2 = \frac{1}{4a_n^2}\left(a_n^2 - 2\right)^2, \quad n = 0, 1, \ldots$$

Tästä nähdään (induktio!), että on oltava $a_n^2 - 2 > 0 \; \forall n$, jolloin voidaan edelleen päätellä:

$$a_{n+1}^2 - 2 = \frac{a_n^2 - 2}{4a_n^2}(a_n^2 - 2) < \frac{a_n^2}{4a_n^2}(a_n^2 - 2) = \frac{1}{4}(a_n^2 - 2), \quad n = 0, 1, \ldots$$

Koska $a_0^2 - 2 = 2$, niin seuraa (induktio!)

$$0 < a_n^2 - 2 < 2 \cdot 4^{-n}, \quad n = 0, 1, \ldots$$

Propositioiden I.6.3 ja I.6.4 perusteella seuraa tästä, että $a_n^2 \to 2$.

On päätelty, että $a_n > 0 \; \forall n$ ja lisäksi, että lukujono $\{a_n^2\}$ on suppeneva ja $a_n^2 \to 2$. Näytetään nyt, että Määritelmän I.6.1 mukaisesti on oltava $\lim_n a_n = a = \sqrt{2}$. Tätä silmällä pitäen kirjoitetaan ensin

$$a_n^2 - 2 = a_n^2 - a^2 = (a_n + a)(a_n - a).$$

Koska oli $a_n > 0$ ja $a_n^2 > 2 = a^2$ jokaisella n, niin $a_n > a\ \forall n$ (ks. Luku I.2). On myös $a > 1$, joten seuraa

$$\begin{aligned} a_n^2 - 2 \;&=\; (a_n + a)(a_n - a) \;>\; 2a\,(a_n - a) \;>\; 2(a_n - a) \\ \Rightarrow \quad a_n - a \;&<\; 2^{-1}(a_n^2 - 2) \;<\; 4^{-n}, \quad n = 0, 1, \ldots \end{aligned}$$

Siis $0 < a_n - a < 4^{-n}\ \forall n$, joten $a_n \to a$ (vrt. päättely edellä). Lukujonolle $\{a_n\}$ on näin löydetty ei-rationaalinen raja-arvo $a = \sqrt{2}$. $\square$

Esimerkin lukujonolla $\{a_n\}$ ei voi olla rationaalista raja-arvoa, koska $\sqrt{2} \notin \mathbb{Q}$ ja Lauseen I.6.2 mukaan raja-arvo on yksikäsitteinen. Rationaalilukujonojen joukossa on siis sellaisia, jotka 'näyttävät suppenevan', mutta joilla ei kuitenkaan ole raja-arvoa rationaalilukujen joukossa. Milloin rationaalilukujonolle on yleisemmin löydettävissä raja-arvo lukujoukkoa (rationaalilukujen kuntaa) laajentamalla ja milloin ei, on keskeinen kysymys jatkossa.

HARJOITUSTEHTÄVIÄ

1. Seuraavissa esimerkeissä lukujono $\{a_n\}$ joko suppenee kohti lukua a tai kasvaa rajatta. Määritä pienin $N \in \mathbb{N}$ siten, että pätee joko $|a_n - a| < 1/100$ tai $a_n > 100$, kun $n > N$:

   a) $a_n = \dfrac{4n-1}{2n+1}$ b) $a_n = \dfrac{4n+100}{2n+100}$ c) $a_n = \dfrac{3n-10000}{2n+10000}$

   d) $a_n = \dfrac{n^2}{2n+1}$ e) $a_n = \dfrac{n^2}{3n+100}$ f) $a_n = \dfrac{n^2}{4n+10000}$

2. Olkoon

$$a_n = \begin{cases} 1/k, & \text{kun } n = 10^k,\ k \in \mathbb{N}, \\ 0, & \text{muulloin } (n \in \mathbb{N}), \end{cases} \qquad b_n = \begin{cases} 10^{-100}, & \text{kun } n = 10^k,\ k \in \mathbb{N}, \\ 0, & \text{muulloin } (n \in \mathbb{N}). \end{cases}$$

   Näytä, että a) $a_n \to 0$, b) $\{b_n\}$ ei suppene kohti mitään lukua.

3. Todista, että lukujonoille pätee:

   a) $\lim_n a_n = a \;\Rightarrow\; \lim_n |a_n| = |a|$.

   b) $\lim_n a_n = a \;\&\; a_n \geq b_n \geq a\ \forall n \;\Rightarrow\; \lim_n b_n = a$.

   c) $a_n \to a \;\&\; b_n \to b \;\Leftrightarrow\; c_n = \max\{|a_n - a|, |b_n - b|\} \to 0$.

   d) $a_n \to a \;\&\; b_n \to b \;\Leftrightarrow\; (a_n - a)^2 + (b_n - b)^2 \to 0$.

   e) $a_n \to a \;\&\; b_n \to b \;\&\; c_n \to c \;\Leftrightarrow\; (a_n - a)^2 + (b_n - b)^2 + (c_n - c)^2 \to 0$.

4. Lukujonoista $\{a_n\}$ ja $\{b_n\}$ tiedetään, että $a_n \to a$ ja $b_n \to a$ ja lukujonosta $\{c_n\}$, että jokaisella n on joko $c_n = a_n$ tai $c_n = b_n$ (ei tiedetä, kumpi). Näytä, että $c_n \to a$.

5. Olkoon $q \in \mathbb{Q},\ q > 0$. Laske $\lim\limits_n q^{-n} \sum\limits_{k=0}^{n} 2^k$ eri q:n arvoilla.

6. Määritä pienimmät Proposition I.6.4 mukaiset luvut $N_k,\ k = 1, 2, 3$, kun $b_k = 10^{-k}$ ja $a_n = 0.9^n,\ n \in \mathbb{N}$.

7. (*) a) Todista Propositio I.6.4. b) Muotoile ja todista Proposition I.6.4 vastine koskemaan lukujonoa $\{a_n\}$, joka kasvaa rajatta.

8. (*) Todista: $a_n > 0\ \forall n\ \ \&\ \ \lim_n a_n = 0 \ \Rightarrow\ \exists \max_n\{a_n\}$. Näytä myös vastaesimerkillä, että $\lim_n a_n = 0 \ \not\Rightarrow\ \exists \max_n\{a_n\}$.

9. (*) Olkoon $k \in \mathbb{N}$, $\mathbb{Q}_k = \{$ äärelliset desimaaliluvut muotoa $x_0.d_1 \dots d_k \}$, ja $a_n \in \mathbb{Q}_k\ \forall n \in \mathbb{N}$. Näytä, että jos $\lim_n a_n = a$, niin $a \in \mathbb{Q}_k$ ja jollakin $N \in \mathbb{N}$ pätee: $a_n = a$ kun $n > N$.

10. (*) Olkoon $\{\, a_k,\ k = 1, 2, \dots \}$ rationaalilukujono ja $\{\, N_k,\ k = 1, 2, \dots \}$ aidosti kasvava indeksijono (luonnollisten lukujen jono) siten, että $N_1 = 1$. Määritellään jono $\{b_n\}$ asettamalla

$$b_n = a_k, \quad \text{kun } N_k \le n < N_{k+1}, \quad k = 1, 2, \dots$$

Todista: $\lim_k a_k = a \ \Rightarrow\ \lim_n b_n = a$.

11. (*) Tarkastellaan kunnassa $(\mathbb{J}, +, \cdot, <)$ (ks. Luku I.2) palautuvasti määriteltyä lukujonoa

$$a_0 = x, \quad a_{n+1} = \frac{1}{2}\left(a_n + \frac{x}{a_n}\right), \quad n = 0, 1, \dots$$

missä $x \in \mathbb{J},\ x > 0$. Näytä:

a) $a_{n+1}^2 - x = \dfrac{1}{4a_n^2}\left(a_n^2 - x\right)^2, \quad n = 0, 1, \dots$

b) $a_n \ge \sqrt{x}, \quad n = 1, 2, \dots$

c) $\lim_n a_n^2 = x$.

d) $\lim_n a_n = \sqrt{x}$.

I.7 Suppenevien lukujonojen ominaisuuksia

Tässä luvussa tarkastellaan suppenevien lukujonojen ominaisuuksia yleiseltä kannalta — 'luku' (mat.) ymmärretään samoin kuin edellisessä luvussa.

MÄÄRITELMÄ I.7.1 Lukujono $\{a_n\}$ on *rajoitettu* (engl. bounded), jos on olemassa luku C siten, että pätee $|a_n| \leq C \quad \forall n$.

LAUSE I.7.2 Suppeneva lukujono on rajoitettu.

Todistus Olkoon jono yleistä muotoa $\{a_n\}_{n=m}^{\infty}$, $m \in \mathbb{Z}$. Jos pätee $a_n \to a$, niin suppenemisen määritelmän mukaan $\exists N \in \mathbb{N}$ siten, että

$$|a_n - a| < 1, \quad \text{kun } n > N$$

(Määritelmässä I.6.1 valittu $\varepsilon = 1$). Tällöin on kolmioepäyhtälön mukaan

$$\begin{aligned} |a_n| &= |\,(a_n - a) + a\,| \\ &\leq |a_n - a| + |a| \;<\; 1 + |a|, \quad \text{kun } n > N, \end{aligned}$$

joten

$$|a_n| \;\leq\; \max\{\,|a_m|, \ldots, |a_N|, |a| + 1\,\} \;=\; C \quad \forall n \geq m. \quad \square$$

Seuraavat yleiset laskusäännöt ovat huomattavan käyttökelposia sekä raja-arvoja laskettaessa että teoreettisemmissa tarkasteluissa.

LAUSE I.7.3 (**Raja-arvojen yhdistelysäännöt**) Olkoon $\{a_n\}$ ja $\{b_n\}$ lukujonoja ja a, b, λ lukuja. Tällöin jos $\lim_n a_n = a$ ja $\lim_n b_n = b$, niin pätee:

(1) $\lim_n(a_n + b_n) \;=\; a + b.$

(2) $\lim_n(\lambda a_n) \;=\; \lambda a.$

(3) $\lim_n(a_n b_n) \;=\; ab.$

(4) $\lim_n(a_n/b_n) \;=\; a/b$ lisäehdoilla: $b_n \neq 0 \;\forall n$ ja $b \neq 0$.

Todistetaan ainoastaan väittämät (3) ja (4) (muut ovat helpompia: Harj.teht. 1).

Väittämä (3) Kirjoitetaan ensin

$$\begin{aligned} a_n b_n - ab &= (a_n b_n - a b_n) + (a b_n - ab) \\ &= b_n(a_n - a) + a(b_n - b). \end{aligned}$$

Koska $\{b_n\}$ on suppeneva jono, niin Lauseen I.7.2 mukaan on jollakin C

$$|b_n| \leq C \quad \forall n,$$

jolloin kolmioepäyhtälön perusteella

$$|a_n b_n - ab| \;\leq\; C|a_n - a| + |a||b_n - b|.$$

Jatkossa oletetaan, että $C > 0$ ja $|a| > 0$ (jos $C = 0$ tai $a = 0$, niin päättely yksinkertaistuu). Valitaan mielivaltainen $\varepsilon > 0$, jolloin

$$\varepsilon_1 = \frac{\varepsilon}{2C} > 0 \quad \text{ja} \quad \varepsilon_2 = \frac{\varepsilon}{2|a|} > 0.$$

Koska lukujonot $\{a_n\}$ ja $\{b_n\}$ suppenevat, on tällöin löydettävissä indeksit $N_1 \in \mathbb{N}$ ja $N_2 \in \mathbb{N}$, siten että pätee

$$\begin{aligned} |a_n - a| &< \varepsilon_1\,, \quad \text{kun } n > N_1, \\ |b_n - b| &< \varepsilon_2\,, \quad \text{kun } n > N_2. \end{aligned}$$

Tässä molemmat epäyhtälöt ovat voimassa indeksistä $N = \max\{N_1, N_2\} \in \mathbb{N}$ eteenpäin, joten pätee

$$|a_n b_n - ab| \;<\; C\varepsilon_1 + |a|\varepsilon_2 \;=\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\; \varepsilon, \quad \text{kun } n > N.$$

Koska tässä $\varepsilon > 0$ oli mielivaltainen ja $N \in \mathbb{N}$ jokaisella $\varepsilon > 0$, niin lukujonon raja-arvon määritelmän mukaan $a_n b_n \to ab$. $\square$

<u>Väittämä</u> (4) Tässä aloitetaan kirjoittamalla

$$\begin{aligned} \frac{a_n}{b_n} - \frac{a}{b} \;&=\; \frac{a_n b - a b_n}{b_n b} \;=\; \frac{(a_n b - ab) + (ab - a b_n)}{b_n b} \\ &=\; \frac{a_n - a}{b_n} - \frac{a(b_n - b)}{b_n b}\,. \end{aligned}$$

Koska $b_n \to b \neq 0$, niin $\exists N_0 \in \mathbb{N}$ siten, että

$$|b_n - b| < \tfrac{1}{2}|b|, \quad \text{kun } n > N_0$$

(raja-arvon määritelmässä asetettu $\varepsilon = \frac{1}{2}|b|,\ N = N_0$). Tällöin on kolmioepäyhtälön nojalla

$$|b_n| \;\geq\; |b| - |b_n - b| \;>\; \tfrac{1}{2}|b| \quad \forall n > N_0,$$

joten

$$\begin{aligned} \left|\frac{a_n}{b_n} - \frac{a}{b}\right| &\le \frac{|a_n - a|}{|b_n|} + \left|\frac{a}{b_n b}\right| |b_n - b| \\ &< \frac{2}{|b|}|a_n - a| + \frac{2|a|}{b^2}|b_n - b| \quad \forall n > N_0. \end{aligned}$$

Tässä on oikea puoli $< \varepsilon$ ($\varepsilon > 0$), kun

$$\frac{2}{|b|}|a_n - a| < \frac{\varepsilon}{2} \quad \text{ja} \quad \frac{2|a|}{b^2}|b_n - b| < \frac{\varepsilon}{2},$$

eli kun

$$|a_n - a| < \frac{|b|}{4}\varepsilon = \varepsilon_1 \quad \text{ja} \quad |b_n - b| < \frac{b^2}{4|a|}\varepsilon = \varepsilon_2,$$

olettaen että tässä $a \neq 0$. (Jos $a = 0$, päättely yksinkertaistuu.) Koska $a_n \to a$ ja $b_n \to b$ ja koska $\varepsilon_1, \varepsilon_2 > 0$, niin ensimmäinen ehto toteutuu jostakin indeksistä N_1 ja toinen jostakin indeksistä N_2 eteenpäin, jolloin

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| < \varepsilon, \quad \text{kun } n > N = \max\{N_0, N_1, N_2\} \in \mathbb{N}.$$

Tässä oli $\varepsilon > 0$ mielivaltainen, joten määritelmän mukaan $a_n/b_n \to a/b$. $\square$

Esimerkki 1 Laske $\lim_n a_n$, kun

$$a_n = \frac{n^2 + 2n + 2}{3n^2 + n}.$$

<u>Ratkaisu</u> Kirjoittamalla a_n muotoon

$$a_n = \frac{1 + 2 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} \cdot \frac{1}{n}}{3 + \frac{1}{n}}$$

voidaan suoraan soveltaa Lauseen I.7.3 sääntöjä (1)–(4):

$$\begin{aligned} \lim_n a_n &= \frac{\lim_n(1 + 2 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} \cdot \frac{1}{n})}{\lim_n(3 + \frac{1}{n})} \\ &= \frac{\lim_n(1) + 2\lim_n(\frac{1}{n}) + 2\lim_n(\frac{1}{n}) \cdot \lim_n(\frac{1}{n})}{3 + \lim_n(\frac{1}{n})} \\ &= \frac{1 + 2 \cdot 0 + 2 \cdot 0 \cdot 0}{3 + 0} = \frac{1}{3}. \end{aligned}$$

Tässä tarvittiin siis ainoastaan yhtä suoraan määritelmästä todistettavaa rajaarvotulosta $\lim_n(\frac{1}{n}) = 0$ (vrt. Esimerkki 1d edellisessä luvussa). $\square$

Esimerkki 2 Yleinen geometrinen sarja on lukujono

$$\sum_{k=0}^{\infty} aq^k \;=\; \{\sum_{k=0}^{n} aq^k, \quad n = 0, 1, \ldots\}.$$

Kun perusmuotoisen sarjan summa tiedetään (edellisen luvun Esimerkki 2), niin Lauseen I.7.3 säännön (2) mukaan yleisen geometrisen sarjan summa on

$$\sum_{k=0}^{\infty} aq^k \;=\; a\sum_{k=0}^{\infty} q^k \;=\; \frac{a}{1-q}, \quad |q| < 1.$$

Jos $|q| \geq 1$, niin yleinen geometrinen sarja suppenee vain kun $a = 0$. □

Esimerkki 3 Edellisen luvun Esimerkissä 3 tarkasteltiin palautuvaa lukujonoa

$$a_0 = 2, \quad a_{n+1} = \frac{a_n}{2} + \frac{1}{a_n}, \quad n = 0, 1, \ldots$$

Jos oletetaan, että $a_n > 0\ \forall n$ ja $a_n \to a > 0$ (vrt. edellisen luvun tarkastelut), niin Lauseen I.7.3 mukaan on oltava

$$a = \lim_n a_{n+1} = \lim_n \left(\frac{a_n}{2} + \frac{1}{a_n}\right) = \frac{a}{2} + \frac{1}{a}.$$

Siis

$$a = \frac{a}{2} + \frac{1}{a} \quad \Rightarrow \quad a^2 = 2.$$

Koska $a > 0$, niin seuraa $a = \sqrt{2}$. □

Esimerkki 4 Jos palautuva lukujono

$$a_{n+1} = qa_n - 1, \quad n = 0, 1, \ldots$$

suppenee (jollakin a_0) kohti raja-arvoa a, niin tästä oletuksesta seuraa:

$$\begin{aligned} q = 1 \quad &\Rightarrow \quad a = a - 1 \quad \Rightarrow \quad 0 = 1. \\ q \neq 1 \quad &\Rightarrow \quad a = qa - 1 \quad \Rightarrow \quad a = 1/(q-1). \end{aligned}$$

Siis päätellään, että jono <u>ei</u> suppene (millään a_0), jos $q = 1$, ja että muissa tapauksissa ainoa mahdollinen raja-arvo on $a = 1/(q-1)$. Jonoa tarkemmin tutkimalla selviää, että

$$a_n = a_0 q^n - \sum_{k=0}^{n-1} q^k, \quad n = 1, 2, \ldots,$$

joten jono suppenee alkuarvosta a_0 riippumatta täsmälleen kun $|q| < 1$, ja tällöin siis $a_n \to 1/(q-1)$. — Huomattakoon, että päättely tapauksessa $q = 1$ on itse asiassa epäsuora todistus väittämälle: $\{a_n\}$ ei suppene. (Vrt. Luku I.3.) □

Seuraavassa vielä kolme väittämää, jotka perustuvat suoraan lukujonon raja-arvon määritelmään. Väittämistä keskimmäinen tunnetaan '*voileipälauseena*'.

LAUSE I.7.4 Lukujonoille pätee

V1. Jos $a_n \leq b\ \forall n$ ($a_n \geq b\ \forall n$) ja $a_n \to a$, niin $a \leq b$ ($a \geq b$).

V2. Jos $a_n \to c$ & $b_n \to c$ & $a_n \leq c_n \leq b_n\ \forall n$, niin $c_n \to c$.

V3. Jos $a_n \to 0$ ja jono $\{b_n\}$ on rajoitettu, niin $a_n b_n \to 0$.

<u>Todistus</u> V1. Oletetaan, että $a_n \leq b\,\forall n$, ja tehdään vastaoletus: $\lim_n a_n = a > b$. Merkitään $\varepsilon = a - b$, jolloin vastaoletuksen perusteella on $\varepsilon > 0$. Koska $a_n \leq b\,\forall n$ (oletus) ja $b = a - \varepsilon$, $\varepsilon > 0$ (vastaoletuksen seuraamus), niin seuraa $a_n \leq a - \varepsilon\ \forall n$ ja siis $|a_n - a| \geq \varepsilon > 0\ \forall n$. Suppenevuuden määritelmästä seuraa tällöin, että $a_n \not\to a$. Oletuksen mukaan oli kuitenkin $a_n \to a$, joten on päädytty loogiseen ristiriitaan: $a_n \to a$ ja $a_n \not\to a$. Vastaoletus $a > b$ on näin ollen väärä, eli on oltava $a \leq b$. Väittämän ensimmäinen osa on siis todistettu. Toinen osa seuraa, kun jo todistettua väittämää sovelletaan jonoon $\{-a_n\}$.

V2. Koska $a_n \leq c_n \leq b_n\ \forall n$, niin

$$\begin{aligned} a_n - c \leq c_n - c \leq b_n - c \quad \& \quad c - b_n \leq c - c_n \leq c - a_n \\ \Rightarrow \quad |c - c_n| \leq \max\{|a_n - c|, |b_n - c|\}\ \forall n. \end{aligned}$$

Koska $a_n \to c$ ja $b_n \to c$, niin jokaisella $\varepsilon > 0$ on olemassa indeksit $N_1, N_2 \in \mathbb{N}$ siten, että $|a_n - c| < \varepsilon$ kun $n > N_1$ ja $|b_n - c| < \varepsilon$ kun $n > N_2$. Tällöin

$$|c - c_n| \leq \max\{|a_n - c|, |b_n - c|\} < \varepsilon, \quad \text{kun } n > N = \max\{N_1, N_2\}.$$

Määritelmän mukaan $c_n \to c$.

V3. Harjoitustehtävä (Tehtävä 6a). $\square$

ESIMERKKI 5 Väittämän (vrt. Lauseen I.7.4 väittämämä V1)

$$a_n < b\ \forall n \quad \& \quad a_n \to a \quad \Rightarrow \quad a < b$$

osoittaa vääräksi vastaesimerkki $a_n = 1 - 1/n,\ n \in \mathbb{N},\ b = 1$. $\square$

ESIMERKKI 6 Kunnassa $(\mathbb{J}, +, \cdot, <)$ (ks. Luku I.2) määritelty palautuva lukujono

$$a_0 = 0, \quad a_{n+1} = \sqrt{a_n + 1}, \quad n = 0, 1, \ldots$$

ilmoitetaan suppenevaksi. Mikä on raja-arvo?

Ratkaisu Palautuskaavasta seuraa, että $a_n \geq 0$ $\forall n$ ja että

$$a_{n+1}^2 = a_n + 1, \quad n = 0, 1, \ldots$$

Kun merkitään $\lim_n a_n = a$, niin Lauseen I.7.3 perusteella seuraa

$$\lim_n a_{n+1}^2 = a^2 = \lim_n (a_n + 1) = a + 1.$$

Siis $a^2 - a - 1 = 0 \;\Rightarrow\; a = \frac{1}{2}(1 \pm \sqrt{5})$. Koska $a_n \geq 0$ $\forall n$, niin on oltava $a \geq 0$ (Lause I.7.4, väittämä V1). Siis $\lim_n a_n = \frac{1}{2}(1 + \sqrt{5})$. $\square$

Jaksolliset desimaaliluvut

Luvussa I.5 todettiin, että jokaiseen rationaalilukuun $x \in \mathbb{Q}$ voidaan liittää ääretön, jaksollinen desimaaliluku $\breve{x} = \{x_n\} \in \mathbb{D}_p$ kriteerillä

$$|x - x_n| \;<\; 10^{-n} \quad \forall n.$$

Tämä merkitsee (vrt. Propositio I.6.4), että

$$\lim_n x_n = x.$$

Jatkossa on luontevaa hieman väljentää Luvussa I.5 sovittua samastusrelaatiota rationaalilukujen ja äärettömien desimaalilukujen välillä. Asetetaan

MÄÄRITELMÄ I.7.5 Jos $x \in \mathbb{Q}$ ja $\breve{x} = \{x_n\} \in \mathbb{D}$, niin

$$x = \breve{x} \quad \Leftrightarrow \quad \lim_n x_n = x.$$

Määritelmän I.7.5 kriteerillä siis jokainen rationaaliluku samastuu edelleen ainakin yhteen jaksolliseen desimaalilukuun, nimittäin jakokulma-algoritmin antamaan. Entä toisinpäin: Jos lähtökohtana on jaksollinen desimaaliluku, niin onko tämä aina samastettavissa johonkin rationaalilukuun Määritelmän I.7.5 mukaisesti? Asian tutkimiseksi oletetaan, että desimaaliluku $\breve{x}$ on muotoa

$$\breve{x} \;=\; x_0.d_1 \ldots d_k d_{k+1} \ldots d_{k+m} d_{k+1} \ldots d_{k+m} \ldots,$$

eli jaksollisuus alkaa (viimeistään) desimaalista n:o $k+1$ ja jakson pituus on m. Tarkastellaan jonoa $\{x_k, x_{k+m}, x_{k+2m}, \ldots\} = \{x_{k+lm},\ l = 0, 1, \ldots\} = \{y_l\}$. Tälle pätee

$$y_{l+1} - y_l \;=\; \sum_{i=k+1}^{k+m} d_i \cdot 10^{-lm-i} \;=\; 10^{-lm}(y_1 - y_0), \quad l = 0, 1 \ldots$$

joten

$$\begin{aligned} y_{l+1} &= y_0 + (y_1 - y_0) + \ldots + (y_{l+1} - y_l) \\ &= y_0 + (y_1 - y_0) \sum_{j=0}^{l} 10^{-jm}, \quad l = 0, 1 \ldots \end{aligned}$$

Näin ollen jono $\{y_l\}$ suppenee geometrisena sarjana:

$$\lim_{l\to\infty} y_l = y_0 + \frac{y_1 - y_0}{1 - 10^{-m}} = x_k + \frac{x_{k+m} - x_k}{1 - 10^{-m}} = x \in \mathbb{Q}.$$

Tästä ja jonon $\{x_n\}$ monotonisuudesta (vrt. Luku I.5) seuraa, että myös koko jono $\{x_n\}$ suppenee. Nimittäin jos esim. $x > 0$, niin jono $\{x_n\}$ on kasvava, jolloin jokaisella $n > k$ pätee

$$x_{k+lm} = y_l \leq x_n \leq y_{l+1} = x_{k+(l+1)m},$$

missä $l \in \mathbb{N} \cup \{0\}$ on valittu (n:stä riippuen) siten, että

$$k + lm < n \leq k + (l+1)m.$$

Kun tässä $n \to \infty$, niin $l \to \infty$, jolloin $y_l \to x$ ja $y_{l+1} \to x$ ja näin ollen $x_n \to x$ (vrt. Lause I.7.4, väittämä V2). Samastussopimuksen I.7.5 mukaisesti on siis todistettu (ks. myös Lause I.6.2)

Lause I.7.6 Jokainen jaksollinen desimaaliluku samastuu Määritelmän I.7.5 mukaisesti yksikäsitteiseen rationaalilukuun.

Esimerkki 7

$$5.5027027027\ldots = 5.5 + 10^{-1} \cdot \frac{0.027}{1 - 10^{-3}} = \frac{55}{10} + \frac{27}{9990} = \frac{1018}{185}. \quad \square$$

Esimerkki 8 Mitkä desimaaliluvut samastuvat lukuihin 0, 10 ja 3/125 ?

<u>Ratkaisu</u> Vaihtoehtoja on kaikissa tapauksissa kaksi:

$$\begin{aligned} 0 &= +0.0000000\ldots = -0.0000000\ldots \\ 10 &= +10.000000\ldots = +9.9999999\ldots \\ 3/125 &= +0.0080000\ldots = +0.0079999\ldots \qquad \square \end{aligned}$$

Desimaalilukujen samastus

Esimerkin 8 mukaan on mahdollista, että kaksi merkkijonoina erilaista (eli ei-identtistä) desimaalilukua $\breve{x}, \breve{y} \in \mathbb{D}$ samastuu samaan rationaalilukuun x. Tällöin on luonnollista samastaa $\breve{x}$ ja $\breve{y}$ myös keskenään, eli kirjoittaa $\breve{x} = \breve{y}$. Jos $\breve{x} = \{x_n\}$ ja $\breve{y} = \{y_n\}$, niin Määritelmän I.7.5 mukaan mainitussa tilanteessa on $\lim_n x_n = \lim_n y_n = x$, jolloin Lauseen I.7.3 mukaan on $\lim_n(x_n - y_n) = 0$. Otetaan tämä raja-arvoehto yleiseksi samastuskriteeriksi verrattaessa äärettömiä desimaalilukuja suoraan keskenään:

MÄÄRITELMÄ I.7.7 Kaksi desimaalilukua $\breve{x} = \{x_n\} \in \mathbb{D}$ ja $\breve{y} = \{y_n\} \in \mathbb{D}$ samastetaan kriteerillä

$$\breve{x} = \breve{y} \quad \Leftrightarrow \quad \lim_n (x_n - y_n) = 0.$$

Kahden 'eri näköisen' desimaaliluvun samastuminen on siis mahdollista ainakin, jos molemmat samastuvat äärelliseen desimaalilukuun. — Tämä osoittautuu ainoaksi erikoistapaukseksi:

LAUSE I.7.8 Jos desimaaliluvut $\breve{x} = \{x_n\} \in \mathbb{D}$ ja $\breve{y} = \{y_n\} \in \mathbb{D}$ samastuvat keskenään Määritelmän I.7.7 mukaisesti, niin joko $\breve{x}$ ja $\breve{y}$ ovat merkkijonoina identtiset tai $\breve{x} = x = \breve{y}$, missä x on äärellinen desimaaliluku.

<u>Todistus</u> Jos $\{x_n\} = \{0\}$, niin myös $\{y_n\} = \{0\}$, sillä muuten $|x_n - y_n| = |y_n| \not\to 0$, koska jono $\{|y_n|\}$ on kasvava. Tällöin on siis $\breve{x}, \breve{y} = \pm 0.000\ldots$ ja $\breve{x} = 0 = \breve{y}$. Jatkossa oletetetaan, että $\{x_n\} \neq \{0\}$ ja $\{y_n\} \neq \{0\}$. Tällöin on $\breve{x}$:n ja $\breve{y}$:n etumerkkien oltava samat, muuten $|x_n - y_n| = |x_n| + |y_n| \not\to 0$. Koska edelleen $\breve{x}$ ja $\breve{y}$ voidaan kumpikin skaalata tekijällä $\pm 10^m$, $m \in \mathbb{Z}$ samastuskriteerin muuttumatta, niin voidaan olettaa, että $0 \leq x_n < 1$ ja $0 \leq y_n < 1$ jokaisella n, jolloin $\breve{x}$ ja $\breve{y}$ ovat jollakin $k \in \mathbb{N}$ muotoa

$$\begin{aligned} \breve{x} &= 0.d_1 \ldots d_{k-1} d_k \ldots\ldots d_m \ldots\ldots, \\ \breve{y} &= 0.d_1 \ldots d_{k-1} \tilde{d}_k \ldots\ldots \tilde{d}_m \ldots\ldots, \end{aligned}$$

missä $\tilde{d}_k \neq d_k$. Tällöin on $x_n = y_n$ kun $n < k$ ja $x_k - y_k = (d_k - \tilde{d}_k)\, 10^{-k}$. Yleisyyttä rajoittamatta voidaan edelleen olettaa, että $\tilde{d}_k < d_k$ $(d_k \neq 0)$. Tällöin nähdään, että erotus $x_n - y_n$ saa pienimmän mahdollisen arvonsa jokaisella $n > k$, kun $d_n = 0$, $\tilde{d}_n = 9$ $\forall n > k$, eli kun $\breve{x}$ ja $\breve{y}$ ovat muotoa

$$\begin{aligned} \breve{x} &= 0.d_1 \ldots d_{k-1} d_k 000 \ldots \\ \breve{y} &= 0.d_1 \ldots d_{k-1} \tilde{d}_k 999 \ldots \end{aligned}$$

Tässä tapauksessa on (vrt. Propositio I.5.2)

$$\begin{aligned} x_n - y_n &= \left(x_k + 0 \cdot 10^{-k-1} + \ldots + 0 \cdot 10^{-n}\right) - \left(y_k + 9 \cdot 10^{-k-1} + \ldots + 9 \cdot 10^{-n}\right) \\ &= (x_k - y_k) - \left(10^{-k} - 10^{-n}\right) \\ &= (d_k - \tilde{d}_k - 1)\, 10^{-k} + 10^{-n}, \quad n > k. \end{aligned}$$

Siis erotuksen $x_n - y_n$ pienin mahdollinen arvo on 10^{-n} kun $n \geq k$, ja tämä saavutetaan, kun $d_k = \tilde{d}_k + 1$ ja $d_n = 0$ ja $\tilde{d}_n = 9$ jokaisella $n > k$. Tässä erikoistapauksessa on siis $x_n = x_k$ $\forall n \geq k$ ja $\lim_n y_n = x_k$, jolloin $\breve{x} = x_k = \breve{y}$ Määritelmän I.7.5 mukaisesti.

Muissa kuin em. erikoistapauksessa on oltava joko $d_k - \tilde{d}_k \geq 2$ (tapaus 1), $d_m \geq 1$ jollakin $m > k$ (tapaus 2) tai $\tilde{d}_m \leq 8$ jollakin $m > k$ (tapaus 3). Tapauksessa 1 seuraa em. tuloksesta, että $x_n - y_n > 10^{-k} = \varepsilon > 0$ $\forall n > k$, jolloin $x_n - y_n \not\to 0$. Tapauksissa 2 ja 3 on oltava $x_m - y_m \geq 2 \cdot 10^{-m}$ (koska erotuksen pienin arvo oli 10^{-m}). Tällöin koska jono $\{x_n\}$ on kasvava ja koska $y_n < y_m + 10^{-m}$ $\forall n > m$ (Propositio I.5.2), niin seuraa, että $x_n - y_n > 10^{-m} = \varepsilon > 0$, kun $n > m$. Siis myös tässä tapauksessa $x_n - y_n \not\to 0$. On päätelty, että jos $\breve{x}$ ja $\breve{y}$ eivät ole merkkijonoina identtiset, niin oletus $\lim_n (x_n - y_n) = 0$ toteutuu vain edellä kuvatussa erikoistapauksessa, jossa $\breve{x}$ ja $\breve{y}$ samastuvat äärelliseen desimaalilukuun. $\square$

Yhteenvetona rationaalilukujen ja äärettömien desimaalilukujen samastuksesta voidaan todeta:

Korollaari I.7.9 Jokainen rationaaliluku x samastuu vähintään yhteen ja enintään kahteen äärettömään desimaalilukuun, jotka ovat jaksollisia. Samastusvaihtoehtoja on kaksi täsmälleen kun x on äärellinen desimaaliluku.

Esimerkki 9 Mihin äärettömiin binaarilukuihin samastuu kymmenjärjestelmän luku a) $x = 3/4$, b) $y = 4/3$?

<u>Ratkaisu</u> a) Koska $3/4 = 2^{-2}(1 \cdot 2 + 1) = 1 \cdot 2^{-1} + 1 \cdot 2^{-2}$, niin

$$x \;=\; 0.11000000\ldots \;=\; 0.10111111\ldots$$

Vaihtoehtoja on kaksi, koska x on äärellinen binaariluku.

b) Koska

$$\frac{4}{3} \;=\; \frac{1}{1 - \frac{1}{4}} \;=\; 1 + \frac{1}{4} + \frac{1}{4^2} + \ldots \;=\; 1 \cdot 2^0 + 0 \cdot 2^{-1} + 1 \cdot 2^{-2} + 0 \cdot 2^{-3} + \ldots,$$

niin $y = 1.010101\ldots$ Tässä on vain yksi vaihtoehto. $\square$

HARJOITUSTEHTÄVIÄ

1. Todista Lauseen I.7.3 väittämät (1) ja (2).

2. Määritä raja-arvojen yhdistelysääntöjä hyväksi käyttäen

   a) $\lim\limits_n \dfrac{2n^3 - 100n^2 - 5000n}{(n+200)^3}$ b) $\lim\limits_n \dfrac{1}{n^3}\left[(n+3)^4 - (n-2)^4\right]$

   c) $\lim\limits_n \left(\dfrac{1+2+\ldots+n}{n+2} - \dfrac{n}{2}\right)$ d) $\lim\limits_n \left[\dfrac{(n+2)! - n!}{(n+1)! + (n-1)!} - n\right]$

3. Laske lukujonojen (sikäli kuin *luku*jonoja)

$$a_n = \frac{x^n}{1+x^n}\,, \quad b_n = \frac{x^n}{1+x^{2n}}\,, \quad n = 1, 2, \ldots\,,$$

   raja-arvot kaikilla x:n rationaalisilla arvoilla.

4. Määritä seuraavien palautuvien lukujonojen mahdolliset raja-arvot olettaen, että jonot ovat suppenevia:

   a) $a_{n+1} = -\dfrac{1}{3}\,a_n + \dfrac{2n^2}{n^2+1}$ b) $a_{n+1} = \dfrac{na_n - 8n + 3}{2n+5}$

   c) $a_{n+2} = \dfrac{2n}{3n+4} - a_{n+1} - \dfrac{1}{4}\,a_n$ d) $a_{n+1} = \dfrac{1}{2+a_n}$

   e) $a_{n+1} = \sqrt{a_n + 2}$ f) $a_{n+1} = \sqrt[3]{a_n + 6}$

   g) $a_{n+1} = \dfrac{3}{4 - a_n^2}$ h) $a_{n+1} = \dfrac{1}{\sqrt{4 - a_n^2}}$

5. Olkoon $a_n + a \neq 0\ \forall n$ ja $\lim_n (a_n + a)^{-1} = b \neq 0$. Näytä, että jono $\{a_n\}$ suppenee ja määritä $\lim_n a_n$.

6. a) Todista Lauseen I.7.4 väittämä V3.

   b) Olkoon $\{a_n\}$ rajoitettu jono. Todista, että $\lim\limits_n \dfrac{2n + a_n}{n} = 2$.

7. a) Mitkä seuraavista luvuista ovat eri suuria ja mitkä samoja ?
   1.285714285714285714.. , $\frac{11}{9}$, $\frac{1287}{1250}$, $\frac{9}{7}$, 1.029599999.. , 1.02949999.. , 1.029600000.. , a = yksi kokonaista kaksi yhdeksäsosaa.

   b) Mitkä seuraavista desimaaliluvuista ovat kaksittain verrattaessa varmasti eri suuret ja mitkä mahdollisesti samat ?
   $\breve{x} = 2.41789..$, $\breve{y} = 2.41799..$, $\breve{a} = 2.41788..$, $\breve{b} = 2.41798..$, $\breve{c} = 2.4179..$

8. Määritä rationaaliluku x siten, että $x = \breve{x}$, kun
a) $\breve{x} = -0.46127127127\ldots$, b) $\breve{x} = 2015.201520152015\ldots$

9. Eräässä lukujärjestelmässä, jossa kirjoitetaan $1 + 1 = 2$, muodostavat äärettömien desimalilukujen vastineet joukon $\mathbb{L}$. Tässä joukossa on luvulla $\breve{x} = 0.222\ldots \in \mathbb{L}$ toinenkin esitysmuoto. Mikä on ko. lukujärjestelmän kantaluku, ja mihin rationaalilukuun $\breve{x}$ samastuu?

10. Olkoon $x, y \in \mathbb{Q}$ ja $\breve{x}, \breve{y} \in \mathbb{D}$. Lähtien Määritelmistä I.7.5 ja I.7.7 ja käyttäen ainoastaan raja-arvojen yhdistelysääntöjä (Lause I.7.3) näytä, että pätee
a) $\breve{x} = x \;\;\&\;\; x = \breve{y} \;\Rightarrow\; \breve{x} = \breve{y}$, b) $\breve{x} = \breve{y} \;\;\&\;\; \breve{y} = y \;\Rightarrow\; \breve{x} = y$.

11. (*) Todista: $\lim\limits_n a_n = a \;\Rightarrow\; \lim\limits_n \dfrac{a_1 + a_2 + \ldots + a_n}{n} = a$.

12. (*) Määritellään palautuva lukujono

$$a_1 \in \mathbb{Q}, \quad a_{n+1} = \frac{a_n}{1 + 10^{-n} a_n}, \quad n = 1, 2, \ldots$$

Näytä, että $\{a_n\}$ suppenee ja

$$\lim_n a_n = \frac{9a_1}{a_1 + 9},$$

paitsi eräillä poikkeuksillisilla a_1:n arvoilla — millä?
Vihje: Tutki jonoa $\{b_n\} = \{a_n^{-1}\}$.

13. (*) a) Määritä palautuvan lukujonon

$$a_{n+1} = (2 - a_n)^2, \quad n = 0, 1, \ldots$$

mahdolliset raja-arvot $a = \lim_n a_n$. b) Tutkimalla jonoa $\{b_n\} = \{a_n - a\}$ päättele, että $\lim_n a_n = a$ vain kun $a_n = a$ jollakin n. c) Päättele, että jos $a_0 \in \mathbb{Q}$, niin $\{a_n\}$ suppenee täsmälleen, kun a_0 on jokin luvuista $0, 1, 2, 3, 4$.

I.8 Monotoniset ja rajoitetut lukujonot

Tässä luvussa tarkastelun kohteena ovat lukujonot, jotka ovat sekä monotonisia (kasvavia tai väheneviä, Määritelmä I.5.1) että rajoitettuja (Määritelmä I.7.1). Tällaiset lukujonot ovat käytännössä yleisiä, vaikka ne muodostavatkin vain pienen 'vähemmistön' kaikista lukujonoista. Tämän luvun päätulos on, että monotoniselle ja rajoitetulle jonolle löydetään aina raja-arvo — kunhan raja-arvo ja suppeneminen sitä kohti määritellään sopivasti.

Esimerkki 1 Jokainen ääretön desimaaliluku $\breve{x} = \{x_n\} \in \mathbb{D}$ on lukujonona monotoninen, kuten aiemmin on todettu (ks. Luku I.5). Jono $\{x_n\}$ on myös rajoitettu, sillä $|x_n| \leq C = |x_0| + 1 \ \forall n$ (Propsitio I.5.2). $\square$

Esimerkki 2 Näytä, että sarja a) $\sum_{k=0}^{\infty} 1/k!$, b) $\sum_{k=1}^{\infty} 1/k^2$ on lukujonona monotoninen ja rajoitettu.

<u>Ratkaisu</u> Kummassakin tapauksessa kyseeessä on *positiiviterminen* sarja eli sarja muotoa $\sum_k a_k$, missä $a_k > 0 \ \forall k$. Positiiviterminen sarja on (osasummiensa) lukujonona aidostikin kasvava, joten riittää osoittaa, että jonot

$$\begin{aligned}
\text{a)}\ \{s_n\} &= \left\{ \sum_{k=0}^{n} \frac{1}{k!}, \quad n = 0, 1, 2, \ldots \right\} = \left\{ 1, 2, \frac{5}{2}, \ldots \right\}, \\
\text{b)}\ \{s_n\} &= \left\{ \sum_{k=1}^{n} \frac{1}{k^2}, \quad n = 1, 2, 3, \ldots \right\} = \left\{ 1, \frac{5}{4}, \frac{23}{18}, \ldots \right\}
\end{aligned}$$

ovat rajoitettuja. a) Kun $n \geq 4$, voidaan arvioida

$$\begin{aligned}
s_n = \sum_{k=0}^{n} \frac{1}{k!} &= 1 + 1 + \frac{1}{2!}\left(1 + \frac{1}{3} + \frac{1}{3 \cdot 4} + \ldots + \frac{1}{3 \cdot 4 \cdots n}\right) \\
&\leq 2 + \frac{1}{2}\left(1 + \frac{1}{3} + \frac{1}{3^2} + \ldots + \frac{1}{3^{n-2}}\right) \\
&< 2 + \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{3}} = 2.75.
\end{aligned}$$

Siis $0 < s_n < 2.75 \ \forall n$, joten $\{s_n\}$ on rajoitettu.

b) Tässä tarvitaan hiukan enemmän kekseliäisyyttä. Esimerkiksi: Koska

$$\frac{1}{k^2} < \frac{1}{(k-1)k} = \frac{1}{k-1} - \frac{1}{k}, \quad k \geq 2,$$

niin osasummia voidaan arvioida teleskooppisummien (vrt. Harj.teht. I.1:13) avulla seuraavasti:

$$\begin{aligned} s_n &= 1 + \sum_{k=2}^{n} \frac{1}{k^2} \;<\; 1 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &= 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 + 1 + \left(-\frac{1}{2} + \frac{1}{2}\right) + \ldots + \left(-\frac{1}{n-1} + \frac{1}{n-1}\right) - \frac{1}{n} \\ &= 2 - \frac{1}{n} \;<\; 2, \quad n \geq 3. \end{aligned}$$

Siis $0 < s_n < 2 \; \forall n$, joten $\{s_n\}$ on rajoitettu. $\square$

Suppeneminen kohti desimaalilukua

Esimerkin 2 sarjoilla ei kummallakaan ole raja-arvoa toistaiseksi tunnettujen lukujen joukossa, kuten tullaan näkemään. Silti nämä sarjat 'näyttävät suppenevan', kun sarjojen (rationaaliset) osasummat muunnetaan äärettömiksi desimaaliluvuiksi.

Esimerkki 2 (jatko) a) Osasummat $s_1 \ldots s_{12}$ ovat

$$\begin{aligned} s_1 &= 2 &&= 2.0000000000\ldots & s_7 &= \tfrac{685}{252} &&= 2.7182539682\ldots \\ s_2 &= \tfrac{5}{2} &&= 2.5000000000\ldots & s_8 &= \tfrac{109601}{40320} &&= 2.7182787698\ldots \\ s_3 &= \tfrac{8}{3} &&= 2.6666666666\ldots & s_9 &= \tfrac{98641}{36288} &&= 2.7182815255\ldots \\ s_4 &= \tfrac{65}{24} &&= 2.7083333333\ldots & s_{10} &= \tfrac{9864101}{3628800} &&= 2.7182818011\ldots \\ s_5 &= \tfrac{163}{60} &&= 2.7166666666\ldots & s_{11} &= \tfrac{13563139}{4989600} &&= 2.7182818261\ldots \\ s_6 &= \tfrac{1957}{720} &&= 2.7180555555\ldots & s_{12} &= \tfrac{260412269}{95800320} &&= 2.7182818282\ldots \end{aligned}$$

b) Osasummat indeksin arvoilla $n = 10^k$, $k = 1 \ldots 12$ ovat

$$\begin{aligned} s_{10} &= 1.5497677311\ldots & s_{10000000} &= 1.6449339668\ldots \\ s_{100} &= 1.6349839001\ldots & s_{100000000} &= 1.6449340568\ldots \\ s_{1000} &= 1.6439345666\ldots & s_{1000000000} &= 1.6449340658\ldots \\ s_{10000} &= 1.6448340718\ldots & s_{10000000000} &= 1.6449340667\ldots \\ s_{100000} &= 1.6449240668\ldots & s_{100000000000} &= 1.6449340668\ldots \\ s_{1000000} &= 1.6449330668\ldots & s_{1000000000000} &= 1.6449340668\ldots \quad \square \end{aligned}$$

Osasummien perusteella näyttää siltä, että raja-arvo $s = \lim_n s_n$ (eli sarjan summa) on esimerkkitapauksessa mahdollista tulkita äärettömäksi desimaaliluvuksi:

$$\text{a)} \quad \sum_{k=0}^{\infty} \frac{1}{k!} = 2.71828182\ldots \qquad \text{b)} \quad \sum_{k=1}^{\infty} \frac{1}{k^2} = 1.64493406\ldots$$

Mutta mitä täsmällisemmin tarkoittaa, että ääretön desimaaliluku on lukujonon raja-arvo? — Alkuperäiseen raja-arvon määritelmään (Määritelmä I.6.1) ei voida suoraan vedota, koska määritelmän ehdossa $|a_n - a| < \varepsilon$ tarvittavaa vähennyslaskua ja lukujen vertailua ei ole (ainakaan toistaiseksi) määritelty, kun $a \in \mathbb{D}$. Ratkaistaan ongelma asettamalla erillinen määritelmä.

MÄÄRITELMÄ I.8.1 Lukujono $\{a_n\}$ *suppenee kohti* ääretöntä desimaalilukua $\breve{x} = \{x_n\} \in \mathbb{D}$ täsmälleen kun

$$\lim_n (a_n - x_n) = 0.$$

Sanotaan tällöin, että $\breve{x}$ on jonon $\{a_n\}$ *raja-arvo*, ja merkitään $\lim_n a_n = \breve{x}$ tai $a_n \to \breve{x}$. Jos lukujonolla on raja-arvo $\breve{x} \in \mathbb{D}$, niin sanotaan, että lukujono *suppenee*, muulloin lukujono *hajaantuu* eli *divergoi*.

Huomattakoon, että ehdossa $\lim_n(a_n - x_n) = 0$ on kyse lukujonon $\{a_n - x_n\}$ tavanomaisesta (rationaalisesta) raja-arvosta Määritelmän I.6.1 mukaisesti. Jos myös jonolla $\{a_n\}$ on rationaalinen raja-arvo, niin Määritelmät I.8.1 ja I.6.1 ovat sopusoinnussa, sillä jos $a_n \to x \in \mathbb{Q}$ (Määritelmä I.6.1) ja $x = \breve{x} = \{x_n\} \in \mathbb{D}$ (Määritelmä I.7.5), niin $a_n - x_n = (a_n - x) + (x - x_n) \to 0 + 0 = 0$ (Lause I.7.3), eli $a_n \to \breve{x}$ Määritelmän I.8.1 mukaisesti. Myös on voimassa (vrt. Lauseet I.7.2 ja I.6.2)

LAUSE I.8.2 Jos lukujono $\{a_n\}$ suppenee Määritelmän I.8.1 mukaisesti, niin (i) $\{a_n\}$ on rajoitettu lukujono, ja (ii) raja-arvo $\lim_n a_n$ on yksikäsitteinen.

<u>Todistus</u> (i) Kirjoittamalla $a_n = (a_n - x_n) + x_n$ ja soveltamalla kolmioepäyhtälöä päätellään

$$|a_n| \le |a_n - x_n| + |x_n| \le C_1 + C_2 = C \ \forall n,$$

sillä $\{a_n - x_n\}$ on rajoitettu suppenevana jonona (Lause I.7.2) ja myös $\{x_n\}$ on rajoitettu (ks. Esimerkki 1 edellä). Siis $\{a_n\}$ on rajoitettu jono.

(ii) Oletetaan, että $a_n \to \breve{x} = \{x_n\} \in \mathbb{D}$ ja $a_n \to \breve{y} = \{y_n\} \in \mathbb{D}$. Tällöin $a_n - x_n \to 0$ ja $a_n - y_n \to 0$ (Määritelmä I.8.1), jolloin Lauseen I.7.3 perusteella

$$x_n - y_n = (a_n - y_n) - (a_n - x_n) \to 0.$$

Siis $\lim_n(x_n - y_n) = 0$, mikä Määritelmän I.7.7 mukaan tarkoittaa: $\breve{x} = \breve{y}$. $\square$

Esimerkki 3 Määritelmän I.8.1 mukaisesti jokainen ääretön desimaaliluku $\breve{x} = \{x_n\}$ 'suppenee itseensä' lukujonona $\{a_n\} = \{x_n\}$ (!). $\square$

Jatkossa keskeinen kysymys on: Jos lukujonolle $\{a_n\}$ ei ole osoitettavissa mitään ilmeistä (esim. rationaalista) raja-arvoa, niin millaisilla ehdoilla voidaan olla varmoja, että Määritelmän I.8.1 mukainen raja-arvo äärettömänä desimaalilukuna on olemassa? Täsmällinen vastaus tähän kysymykseen saadaan jäljempänä Luvussa I.10; tässä yhteydessä tyydytään osittaiseen vastaukseen. Ensinnäkin, Lauseen I.8.2 mukaan <u>välttämätön</u> ehto raja-arvon olemassaololle on, että $\{a_n\}$ on rajoitettu. Seuraavan lauseen mukaan <u>riittävä</u> ehto on, että $\{a_n\}$ sekä monotoninen että rajoitettu.

*Lause I.8.3 † (**Monotoninen ja rajoitettu lukujono suppenee**) Jokaisella monotonisella ja rajoitetulla lukujonoilla on Määritelmän I.8.1 mukainen raja-arvo äärettömänä desimaalilukuna.‡

<u>Todistus</u> Määritelmästä I.8.1 nähdään, että jos $a_n \to \breve{x}$, niin $-a_n \to -\breve{x}$. Voidaan sen vuoksi rajoittua tapauksiin, joissa joko (i) $\{a_n\}$ on kasvava ja jollakin $N \in \mathbb{N}$ pätee $a_n \geq 0$ $\forall n \geq N$ tai (ii) $\{a_n\}$ on vähenevä ja $a_n \geq 0$ $\forall n \in \mathbb{N}$, sillä muissa tapauksissa jono $\{-a_n\}$ on tyyppiä (i) tai (ii). Olkoon ensin jono tyyppiä (i), ja määritellään jonoon liittyen predikaatti (vrt. Luku I.3)

$$Q(x): \quad a_n < x \ \ \forall n \quad (x \in \mathbb{Q}).$$

Koska $\{a_n\}$ on rajoitettu, on löydettävissä $M \in \mathbb{N}$ siten, että $a_n < M$ $\forall n$. Oletuksien (i) perusteella on silloin $0 \leq a_n < M$ $\forall n \geq N$. Tällöin voidaan valita yksikäsitteisesti $x_0 \in \{0, \dots, M-1\}$ siten, että $Q(x_0)$ on epätosi ja $Q(x_0+1)$ on tosi. Koska $Q(x_0)$ on epätosi, on $a_n \geq x_0$ jollakin n. Olkoon $n = N_0$ pienin tällainen indeksi. Tällöin koska jono $\{a_n\}$ on kasvava ja koska $Q(x_0+1)$ on tosi, on $x_0 \leq a_n < x_0 + 1$ $\forall n \geq N_0$.

Valitaan seuraavaksi $x_1 = x_0 + d_1 \cdot 10^{-1}$, $d_1 \in \{0, 1, \dots, 9\}$ siten, että $Q(x_1)$ on epätosi ja $Q(x_1 + 10^{-1})$ on tosi. Koska $Q(x_1)$ on epätosi, on $a_n \geq x_1$ jollakin n. Olkoon $n = N_1$ pienin tällainen indeksi ($N_1 \geq N_0$). Tällöin koska jono $\{a_n\}$ on kasvava ja koska $Q(x_1 + 10^{-1})$ on tosi, on $x_1 \leq a_n < x_1 + 10^{-1}$ $\forall n \geq N_1$.

Jatkamalla samalla tavoin saadaan konstruoiduksi kasvava jono $\{x_k\}$ ja kasvava indeksijono $\{N_0, N_1, \dots\}$ siten, että jokaisella k pätee

$$x_k \leq a_n < x_k + 10^{-k}, \quad \text{kun } n \geq N_k. \qquad (\star)$$

†Tässä ja jatkossa merkintä (*) lauseen yhteydessä tarkoittaa, että lauseen todistus on tavallista vaativampi looginen konstruktio.

‡Lause I.8.3 on esimerkki olemassaolo- eli *eksistenssilauseesta*, jossa jokin (tässä raja-arvo) väitetään olemassa olevaksi. — Paino lauseessa onkin sanalla 'on'.

Tapauksessa (ii) tullaan samaan tulokseen, kun predikaatti $Q(x)$ määritellään

$$Q(x): \quad a_n < x \text{ jollakin } n \quad (x \in \mathbb{Q}).$$

Kummassakin tapauksessa (i) ja (ii) on siis konstruoitu ääretön desimaaliluku $\breve{x} = \{x_k,\ k = 0, 1, \dots\}$ ja indeksijono $\{N_k,\ k = 0, 1, \dots\}$ siten, että pätee

$$|a_n - x_k| < 10^{-k}, \quad \text{kun } n \geq N_k.$$

Kun huomioidaan, että jonolle $\{x_k\}$ pätee (Propositio I.5.2)

$$|x_n - x_k| < 10^{-k}, \quad \text{kun } n \geq k,$$

niin kirjoittamalla $a_n - x_n = (a_n - x_k) + (x_k - x_n)$ ja soveltamalla kolmioepäyhtälöä seuraa kahden viimeksi kirjoitetun epäyhtälön perusteella

$$|a_n - x_n| \leq |a_n - x_k| + |x_n - x_k| < 2 \cdot 10^{-k}, \quad \text{kun } n \geq \max\{k, N_k\} = N'_k.$$

Proposition I.6.4 mukaan tämä ehto on sama kuin ehto: $\lim_n (a_n - x_n) = 0$. Siis $\lim_n a_n = \breve{x} = \{x_n\}$ Määritelmän I.8.1 mukaisesti. $\square$

Kun lukujono $\{a_n\}$ tunnetaan, niin Lauseen I.8.3 todistuskonstruktiota voidaan seurata algoritmina, joka määrittää raja-arvon alkaen kokonaislukuosasta ja edeten desimaali kerrallaan.

Esimerkki 2 (jatko) Raja-arvo $\lim_n s_n = \{x_k\} \in \mathbb{D}$ määrätään valitsemalla luvut $x_k \in \mathbb{Q}_k,\ k = 0, 1, \dots$ siten, että $s_n \geq x_k$ jollakin n ja $s_n < x_k + 10^{-k}\ \forall n$, jolloin ehto $(\star)$ toteutuu jollakin $N_k \in \mathbb{N}$ (jokaisella k), kun $a_n = s_n$.

a) Koska $s_n < 3\ \forall n$ ja $s_1 = 2 \ \Rightarrow\ s_n \geq 2\ \forall n \geq 1$, niin $x_0 = 2$ ja pienin N_0:n arvo, jolla ehto $(\star)$ toteutuu kun $k = 0$, on $N_0 = 1$. Vastaavasti koska $s_3 < 2.7$, $s_4 > 2.7$ ja ilmeisesti $s_n < 2.8\ \forall n$ (tämä on erikseen varmistettava osasummia arvioimalla, ks. Harj.teht. 1), niin $x_1 = 2.7$ ja indeksin N_1 pienin arvo on $N_1 = 4$. Samalla tavoin nähdään lasketuista osasummista, että $x_2 = 2.71$, $x_3 = 2.718$, $\dots$, $x_8 = 2.71828182$ ja että indeksien $N_k,\ k = 2 \dots 8$ pienimmät arvot ehdossa $(\star)$ ovat $N_2 = 5$, $N_3 = 6$, $N_4 = 7$, $N_5 = N_6 = 9$, $N_7 = 10$ ja $N_8 = 11$.

b) Osasummista

$$\begin{aligned} s_{21} &= 1.59843081\dots & s_{1070} &= 1.64399992\dots \\ s_{22} &= 1.60049693\dots & s_{1071} &= 1.64400079\dots \\ s_{202} &= 1.63999580\dots & s_{29353} &= 1.64489999\dots \\ s_{203} &= 1.64002007\dots & s_{29354} &= 1.64490000\dots \end{aligned}$$

on pääteltävissä (ja varmistettavissa, ks. Harj.teht. 1), että pienimmät indeksin N_k arvot, joilla ehdot $(\star)$ toteutuvat kun $k = 1 \dots 4$, ovat $N_1 = 22$, $N_2 = 203$, $N_3 = 1071$ ja $N_4 = 29354$. $\square$

Esimerkki 4 Luvun I.6 Esimerkissä 3 näytettiin, että palautuvalle rationaalilukujonolle

$$a_0 = 2, \quad a_{n+1} = \frac{a_n}{2} + \frac{1}{a_n}, \quad n = 0, 1, \ldots$$

pätee $a_n > 0 \ \forall n$ ja $a_n^2 > 2 \ \forall n$. Palautuskaavasta ja mainituista tuloksista seuraa myös, että

$$a_{n+1} - a_n \ = \ \frac{a_n}{2} + \frac{1}{a_n} - a_n \ = \ \frac{1}{2a_n}\,(2 - a_n^2) \ < \ 0,$$

joten $\{a_n\}$ on aidosti vähenevä ja myös rajoitettu: $0 < a_n \leq 2 \ \forall n$. Siis $\{a_n\}$ suppenee Määritelmän I.8.1 mukaisesti kohti ääretöntä desimaalilukua. Suppeneminen on itse asiassa hyvin nopeaa:

$$\begin{aligned}
a_0 &= 2.000\ldots \\
a_1 &= 1.5000\ldots \\
a_2 &= 1.41666\ldots \\
a_3 &= 1.41421568\ldots \\
a_4 &= 1.41421356237468\ldots \\
a_5 &= 1.4142135623730950488016896 2\ldots \\
a_6 &= 1.41421356237309504880168872420969807856 96\ldots \\
a_7 &= 1.41421356237309504880168872420969807856 96\ldots \quad \square
\end{aligned}$$

Esimerkin lukujonosta tiedetään, että se suppenee myös Määritelmän I.6.1 mukaisesti kohti lukua $\sqrt{2}$ (ks. Luku I.6). Edellä todettiin myös, että Määritelmien I.6.1 ja I.8.1 mukaiset raja-arvot ovat samat, jos raja-arvo on rationaalinen. On odotettavissa, että näin on laita yleisemminkin, jos molemmat raja-arvot ovat olemassa. Näin olettaen (asia varmistuu seuraavassa luvussa) saa abstrakti luku $\sqrt{2}$ konkreettisen 'ilmiasun' äärettömänä desimaalilukuna:

$$\sqrt{2} = \ 1.41421356237309504880168872420969807856 96\ldots$$

Algoritmina (haluttaessa laskea $\sqrt{2}$:n desimaaleja) esimerkin lukujono $\{a_n\}$ on nopeimpia tunnettuja.†

†Esimerkin algoritmi tunnettiin jo muinaisessa Babyloniassa 3000 vuotta sitten. Yalen yliopistossa säilytettävässä savitaulussa n:o 7289 on algoritmilla laskettu $\sqrt{2}$:n approksimaatio x_3 lähtien alkuarvosta x_1. Laskut on suoritettu 60-kantaisessa lukujärjestelmässä neljän merkitsevän numeron tarkkuudella, jolloin tulokseksi on saatu

$$x_3^* \ = \ 1 + 24 \cdot 60^{-1} + 51 \cdot 60^{-2} + 10 \cdot 60^{-3} \ = \ 1.41421296\,..$$

Katkaisuvirheestä johtuen tuloksen virhe (noin $-6 \cdot 10^{-7}$) sattuu olemaan jopa pienempi kuin x_3:n tarkan arvon virhe (noin $+21 \cdot 10^{-7}$).

Neperin luku

Propositio I.8.4 Lukujonot

$$\{a_n\} \;=\; \left\{\left(1+\frac{1}{n}\right)^n,\; n=1,2,\ldots\right\}, \quad \{b_n\} \;=\; \left\{\left(1-\frac{1}{n}\right)^n,\; n=1,2,\ldots\right\}$$

ovat kasvavia ja rajoitettuja. Lisäksi $\lim_n a_n b_n = 1$.

<u>Todistus</u> Soveltamalla Bernoullin epäyhtälöä (Propositio I.4.2) saadaan

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{(1+\frac{1}{n})^n} \;=\; \left(1+\frac{1}{n}\right)\left(\frac{1+\frac{1}{n+1}}{1+\frac{1}{n}}\right)^{n+1} \\ &= \left(1+\frac{1}{n}\right)\left[1-\frac{1}{(n+1)^2}\right]^{n+1} \;>\; \left(1+\frac{1}{n}\right)\left(1-\frac{1}{n+1}\right) \;=\; 1.\end{aligned}$$

Siis $a_{n+1} > a_n \;\; \forall n$, eli $\{a_n\}$ on (aidosti) kasvava jono. Vastaavalla päättelyllä todetaan (Harj.teht. 4), että myös jono $\{b_n\}$ on aidosti kasvava. Tällöin koska jokaisella $n \geq 2$ pätee

$$1-\frac{1}{n^2} \;<\; 1 \quad\Leftrightarrow\quad 1+\frac{1}{n} \;<\; \frac{1}{1-\frac{1}{n}},$$

päätellään

$$a_n \;<\; \frac{1}{b_n} \;\leq\; \frac{1}{b_2} \;=\; 4 \quad \text{kun } n \geq 2.$$

Siis $1 < a_n < 4$ ja $0 \leq b_n < 1$ kaikilla n, eli jonot $\{a_n\}$ ja $\{b_n\}$ ovat paitsi kasvavia myös rajoitettuja. Bernoullin epäyhtälöstä seuraa lisäksi

$$1 \;>\; a_n b_n \;=\; \left(1-\frac{1}{n^2}\right)^n \;>\; 1-\frac{1}{n} \quad (n \geq 2),$$

joten $0 < 1 - a_n b_n < 1/n \;\; \forall n \geq 2$. Tästä ja lukujonon raja-arvon määritelmästä seuraa väittämän viimeinen osa. $\square$

Koska Proposition I.8.4 jonot ovat kasvavia ja rajoitettuja, niin niille voidaan laskea Määritelmän I.8.1 mukainen raja-arvo äärettömänä desimaalilukuna. Jonon $\{a_n\}$ tapauksessa raja-arvo on nk. *Neperin luku*, jonka vakiintunut symboli on e:

$$\lim_n (1+\tfrac{1}{n})^n = e = \; 2.71828182845905\ldots$$

Luku e on jaksoton, eli se ei samastu mihinkään rationaalilukuun.

Esimerkki 5 Määritä pienimmät indeksit N_k, $k = 0 \ldots 3$ Lauseen I.8.3 todistuskonstruktiossa, kun $a_n = (1 + 1/n)^n$.

Ratkaisu Todistuskonstruktion vertailuluvut ovat

$$x_0 = 2, \quad x_1 = 2.7, \quad x_2 = 2.71, \quad x_3 = 2.718, \quad \ldots$$

Laskemalla jonon $\{a_n\}$ alkupään termejä desimaalilukuina todetaan

$$\begin{aligned} a_1 &= 2.00000000\ldots & a_{163} &= 2.70999015\ldots \\ a_2 &= 2.25000000\ldots & a_{164} &= 2.71004043.. \\ a_{73} &= 2.69989423\ldots & a_{4821} &= 2.71799996\ldots \\ a_{74} &= 2.70013967\ldots & a_{4822} &= 2.71800001\ldots \end{aligned}$$

Siis $N_0 = 1$, $N_1 = 74$, $N_2 = 164$, $N_3 = 4822$. $\square$

Palataan vielä Esimerkin 2 positiivitermiseen sarjaan $\sum_{k=0}^{\infty} 1/k!$. Tämän sarjan edellä lasketuista osasummista voi arvella, että sarja itse asiassa suppenee kohti Neperin lukua (ja vieläpä nopeasti). Tämän arvelun varmistamiseksi riittää näyttää, että jos $s_n = \sum_{k=0}^{n} 1/k!$ ja $a_n = (1 + 1/n)^n$, niin $\lim_n (s_n - a_n) = 0$, vrt. Neperin luvun määritelmä edellä. Todistetaan tämä väittämä, joka samalla antaa nopean algoritmin Neperin luvun desimaalien laskemiseksi.

Propositio I.8.5 $$\lim_n \left[\sum_{k=0}^{n} \frac{1}{k!} - \left(1 + \frac{1}{n}\right)^n \right] = 0.$$

Todistus Binomikaavan mukaan

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &= 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{1 \cdot 2} \left(\frac{1}{n}\right)^2 + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} \left(\frac{1}{n}\right)^3 + \cdots + \left(\frac{1}{n}\right)^n \\ &= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \ldots \\ &\qquad \ldots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right) \\ &= \sum_{k=0}^{n} a_k^{(n)} \frac{1}{k!}, \end{aligned}$$

missä

$$a_k^{(n)} = \begin{cases} 1, & \text{kun } k = 0, 1, \\ \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right), & \text{kun } k = 2 \ldots n. \end{cases}$$

Nähdään, että jokaisella $1 \le m \le n$ pätee

$$1 \;\ge\; a_k^{(n)} \;\ge\; a_m^{(n)} \;\ge\; \left(1 - \frac{m-1}{n}\right)^{m-1}, \quad k = 0 \ldots m$$

$$\Rightarrow \quad 0 \;\le\; 1 - a_k^{(n)} \;\le\; 1 - \left(1 - \frac{m-1}{n}\right)^{m-1}, \quad k = 0 \ldots m.$$

Tämän sekä arvion (ks. Harj.teht. 1)

$$\sum_{k=0}^{n} \frac{1}{k!} \;<\; \sum_{k=0}^{m} \frac{1}{k!} + \frac{2}{(m+1)!}, \quad 0 \le m < n$$

perusteella voidaan arvioida

$$\begin{aligned}
0 \;\le\; \sum_{k=0}^{n} \frac{1}{k!} - \left(1 + \frac{1}{n}\right)^n \;&=\; \sum_{k=0}^{n} \left(1 - a_k^{(n)}\right) \frac{1}{k!} \\
&=\; \sum_{k=0}^{m} \left(1 - a_k^{(n)}\right) \frac{1}{k!} \;+\; \sum_{k=m+1}^{n} \left(1 - a_k^{(n)}\right) \frac{1}{k!} \\
&<\; \left[1 - \left(1 - \frac{m-1}{n}\right)^{m-1}\right] \sum_{k=0}^{m} \frac{1}{k!} \;+\; \sum_{k=m+1}^{n} \frac{1}{k!} \\
&<\; \left[1 - \left(1 - \frac{m-1}{n}\right)^{m-1}\right] \cdot 3 \;+\; \frac{2}{(m+1)!},
\end{aligned}$$

missä $0 \le m < n$. Olkoon nyt $\varepsilon > 0$ ja valitaan $m \in \mathbb{N}$ siten, että

$$\frac{2}{(m+1)!} \;<\; \frac{\varepsilon}{2}$$

(mahdollista, koska $2/(m+1)! \to 0$ kun $m \to \infty$). Kun m on näin kiinnitetty, valitaan edelleen $N \in \mathbb{N}$, $N \ge m$ siten, että

$$1 - \left(1 - \frac{m-1}{n}\right)^{m-1} \;<\; \frac{\varepsilon}{6}, \quad \text{kun } n > N$$

(mahdollista, koska $\left(1 - \frac{m-1}{n}\right)^{m-1} \to 1 \;\Rightarrow\; 1 - \left(1 - \frac{m-1}{n}\right)^{m-1} \to 0$, kun $n \to \infty$). Tällöin seuraa

$$0 \;\le\; \sum_{k=0}^{n} \frac{1}{k!} - \left(1 + \frac{1}{n}\right)^n \;<\; \varepsilon, \quad \text{kun } n > N.$$

Koska tässä $\varepsilon > 0$ oli mielivaltainen ja $N \in \mathbb{N}$ (N riippuu vain ε:sta), niin lukujonon raja-arvon määritelmän perusteella seuraa väite. $\square$

HARJOITUSTEHTÄVIÄ

1. Laskemalla summia a) $s_n = \sum_{k=0}^{n} 1/k!$, b) $s_n = \sum_{k=1}^{n} 1/k^2$ todetaan: a) $s_n = 2.718281826..$ kun $n = 11$, b) $s_n = 1.644934065..$ kun $n = 10^9$. Vedetään tuloksista johtopäätös: $\lim_n s_n = s \in \mathbb{D}$ katkaistuna kahdeksaan merkitsevään numeroon on a) $s = 2.71828182..$ b) $s = 1.64493406..$ Varmista johtopäätös näyttämällä ensin: Jos a) $m \in \mathbb{N} \cup \{0\}$, b) $m \in \mathbb{N}$, niin jokaisella $n > m$ pätee

$$\text{a)} \quad s_n < s_m + \frac{m+2}{m+1} \cdot \frac{1}{(m+1)!}\,, \qquad \text{b)} \quad s_n < s_m + \frac{1}{m}\,.$$

2. Näytä, että positiivitermisen sarjan

$$\sum_{k=0}^{\infty} \frac{10^{-k}}{k+1}$$

osasummien jono on rajoitettu. Laske sarjan summa $x \in \mathbb{D}$ viiden merkitsevän numeron tarkkuudella ja varmista tulos, ts. todista lasketut numerot oikeiksi.

3. Millaisen desimaalimuodon Lauseen I.8.3 todistuskonstruktio antaa raja-arvoille $a = \lim_n(1 - 1/n)$, $b = \lim_n(1 + 1/n)$, $c = \lim_n(-1 + 1/n)$ ja $d = \lim_n(-1 - 1/n)$?

4. Näytä, että Proposition I.8.4 lukujono $\{b_n\}$ on aidosti kasvava.

5. Näytä, että palautuva lukujono

$$a_0 = \frac{1}{10}\,, \quad a_{n+1} = \frac{a_n}{1 + a_n^6}\,, \quad n = 0, 1, \ldots$$

on vähenevä ja rajoitettu. Jonolla on myös rationaalinen raja-arvo — mikä ?

6. Näytä, että palautuva rationaalilukujono

$$a_0 \in \mathbb{Q}, \quad a_{n+1} = \frac{a_n}{1 + 10^{-n} a_n^2}\,, \quad n = 0, 1, \ldots$$

on monotoninen ja rajoitettu. Laske raja-arvo $\lim_n a_n = a \in \mathbb{D}$ kuuden merkitsevän numeron tarkkuudella ja perustele tarkkuus, kun a) $a_0 = 1$, b) $a_0 = 3$.

7. Tarkastellaan palautuvaa rationaalilukujonoa

$$a_0 = 1, \quad a_{n+1} = 1 - \frac{1}{a_n + 3}, \quad n = 0, 1, \ldots$$

a) Näytä induktiolla: $\frac{5}{7} < a_n \leq \frac{3}{4}$, kun $n = 1, 2, \ldots$
b) Näytä: Jonolle $\{b_n\} = \{a_{n+1} - a_n\}$ pätee

$$b_{n+1} = \frac{b_n}{(a_n + 3)(a_{n+1} + 3)}, \quad n = 0, 1, \ldots$$

c) Näytä induktiolla: $\{a_n\}$ on aidosti vähenevä lukujono.
d) Laske raja-arvo $\lim_n a_n$ Määritelmän I.6.1 mukaan.
e) Päättele: Määritelmän I.8.1 mukainen raja-arvo katkaistuna kuuteen merkitsevään numeroon on $\lim_n a_n = 0.732050..$ (voit tukeutua laskimeen!).

8. (*) Eräs algoritmi tunnetun luvun $\pi = 3.14159265358979323846264338 32..$ laskemiseksi perustuu kaavaan

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right).$$

Totea sarja positiivitermiseksi ja arvioi, montako π:n oikeaa desimaalia on poimittavissa sarjan osasumman s_n desimaalilukumuodosta, kun a) $n = 10$, b) $n = 15$, c) $n = 20$. (Kaava on keksitty v. 1995.)

9. (*) Näytä, että seuraavat lukujonot (kunnassa $(\mathbb{J}, +\cdot, <)$, ks. Luku I.2) ovat monotonisia ja rajoitettuja (*vihje*: induktio!). Laske lukujonoille myös raja-arvot (*vihje*: juuriluvun määritelmä ja Lause I.7.3).

a) $a_0 = \sqrt{3}, \quad a_{n+1} = \sqrt{3 + a_n}, \quad n = 0, 1, \ldots$

b) $a_0 = \sqrt{3}, \quad a_{n+1} = \sqrt{2a_n}, \quad n = 0, 1, \ldots$

c) $a_0 = \sqrt{5}, \quad a_{n+1} = \sqrt{2a_n}, \quad n = 0, 1, \ldots$

d) $a_0 = 3, \quad a_{n+1} = \sqrt[4]{6 + a_n^2}, \quad n = 0, 1, \ldots$

10. (*) Näytä, että palautuva lukujono

$$a_{n+1} = \frac{a_n}{5 - a_n}, \quad n = 0, 1, \ldots$$

on rajoitettu ja eräästä indeksistä alkaen monotoninen, sikäli kuin pätee

$$a_0 \neq \frac{4}{1 - 5^{-n}} \quad \forall n \in \mathbb{N}.$$

Määritä myös tällä ehdolla (a_0:sta riippuva) raja-arvo $\lim_n a_n$.
Vihje: Tutki jonoa $\{b_n\} = \{a_n^{-1}\}$.

I.9 Reaaliluvut

Aloitetaan määritelmästä.

MÄÄRITELMÄ I.9.1 (**Reaaliluvut**) *Reaalilukujen* joukko $\mathbb{R}$ koostuu äärettömistä desimaaliluvuista, ts. $\mathbb{R} = \mathbb{D}$.

Määritelmä I.9.1 reaaliluvuille ei ole ainoa mahdollinen, sillä on ilmeistä, että esimerkiksi äärettömät binaariluvut voitaisiin yhtä hyvin ottaa määrittelyn lähtökohdaksi. Monia muitakin lähestymistapoja on, kuten havaitaan tuonnempana. Määritelmä I.9.1 antaa joka tapauksessa reaaliluvuille erään mahdollisen tulkinnan, joka jatkossa otetaan lähtökohdaksi.

Yhdessä aiemmin sovitun samastusrelaation (Määritelmä I.7.7) kanssa Määritelmä I.9.1 kertoo vasta, miltä reaaliluvut 'näyttävät' ja miten niitä erotellaan. Jotta reaaliluvuilla päästäisiin myös laskemaan, on määriteltävä lukujen väliset laskuoperaatiot.

$\mathbb{R}$:n laskuoperaatiot

Viitaten Määritelmään I.8.1 asetetaan

MÄÄRITELMÄ I.9.2 (**Reaalilukujen laskutoimitukset**) Jos $\breve{x} = \{x_n\}$ ja $\breve{y} = \{y_n\}$ ovat reaalilukuja, niin määritellään

$$\begin{aligned} \text{yhteenlasku:} \quad & \breve{x} + \breve{y} = \lim_n (x_n + y_n) = \breve{z} \in \mathbb{R}. \\ \text{kertolasku:} \quad & \breve{x}\breve{y} = \lim_n x_n y_n = \breve{z} \in \mathbb{R}. \\ \text{jakolasku:} \quad & \breve{x}/\breve{y} = \lim_n x_n/y_n = \breve{z} \in \mathbb{R}. \end{aligned}$$

Tässä on jakolaskussa oletettava tavalliseen tapaan $\breve{y} \neq 0$, eli $y_n \not\to 0$ (Määritelmä I.7.5). Ehto merkitsee, että jollakin m on $y_m \neq 0$, jolloin tästä indeksistä alkaen on $|y_n| \geq |y_m| > 0,\ n \geq m$ (koska $\{|y_n|\}$ on kasvava jono). Jakolaskun määritelmä on siis muodollisesti kunnossa, kunhan $\breve{y} \neq 0$ ja jonossa $\{x_n/y_n\}$ rajoitutaan indekseihin $n \geq m$. Yhteenlaskun määritelmästä nähdään, että luvun $\breve{x} = \{x_n\}$ vastaluku on $-\breve{x} = \{-x_n\}$, eli operaatio $\breve{x} \mapsto -\breve{x}$ vastaa desimaaliluvun etumerkin vaihtoa, kuten jo Luvussa I.5 sovittiin. Vähennyslasku määritellään tämän jälkeen normaaliin tapaan, eli $\breve{x} - \breve{y} = \breve{x} + (-\breve{y}) = \lim_n(x_n - y_n) \in \mathbb{R}$.

Määritelmä I.9.2 on käytännön läheinen sikäli, että se perustuu suoraan lukujen $\breve{x}$ ja $\breve{y}$ esitysmuotoihin annetussa lukujärjestelmässä, tässä kymmenjärjestelmässä. Lukujonojen $\{x_n + y_n\}$, $\{x_n y_n\}$ ja $\{x_n/y_n\}$ termien laskemisen voi ajatella

vastaavan likimääräisiä laskuoperaatioita liukuluvuilla (vrt. Luku I.5). Määritelmään sisältyy kuitenkin toistaiseksi ratkaisematon teoreettinen ongelma: Lukujonojen $\{x_n + y_n\}$, $\{x_n y_n\}$ ja $\{x_n/y_n\}$ oletetetaan suppenevan kohti reaalilukua (eli ääretöntä desimaalilukua), mutta toistaiseksi tämä on varmistettu vain monotonisen ja rajoitetun lukujonon osalta (Lause I.8.3). Määritelmän I.9.2 lukujonot ovat kyllä rajoitettuja (mainituin lisäehdoin koskien jonoa $\{x_n/y_n\}$) mutta eivät välttämättä monotonisia. Määritelmän tueksi tarvitaankin seuraava lause.

Lause I.9.3 Jos $\{a_n\}$ ja $\{b_n\}$ ovat monotonisia ja rajoitettuja lukujonoja, niin lukujonot $\{a_n + b_n\}$ ja $\{a_n b_n\}$ suppenevat Määritelmän I.8.1 mukaisesti kohti reaalilukua. Jos edelleen $|b_n| \geq b > 0\ \forall n$, niin myös jono $\{a_n/b_n\}$ suppenee kohti reaalilukua.

Todistus Yhteenlasku. Jos $\{a_n\}$ ja $\{b_n\}$ ovat molemmat kasvavia tai molemmat väheneviä, niin $\{a_n + b_n\}$ on vastaavasti kasvava/vähenevä (Harj.teht. I.5:2), jolloin $\lim_n(a_n + b_n) \in \mathbb{R}$ on olemassa Lauseen I.8.3 perusteella. Olkoon toinen jonoista kasvava ja toinen vähenevä, esim. $\{a_n\}$ kasvava. Valitaan $b \in \mathbb{Q}$ siten, että $b_n + b \geq 0\ \forall n$ (mahdollista, koska $\{b_n\}$ on rajoitettu) ja kirjoitetaan

$$a_n + b_n = (a_n - b) + (b_n + b).$$

Tässä $\{b_n + b\}$ on rajoitettu ja vähenevä jono, joten Lauseen I.8.3 mukaan se suppenee: $\lim_n(b_n + b) = \breve{x} = \{x_n\} \in \mathbb{R}$. Koska $b_n + b \geq 0\ \forall n$, on $\breve{x}$:n etumerkki + (mahdollisesti $\breve{x} = 0$), joten lukujono $\{x_n\}$ on kasvava. Koska myös $\{a_n - b\}$ on kasvava jono, niin samoin on $\{a_n - b + x_n\}$, joten Lauseen I.8.3 mukaan $\lim_n(a_n - b + x_n) = \breve{y} = \{y_n\} \in \mathbb{R}$. Näin määrätty $\breve{y}$ on lukujonon $\{a_n + b_n\}$ raja-arvo Määritelmän I.8.1 mukaisesti, sillä ko. määritelmän ja Lauseen I.7.3 perusteella

$$(a_n + b_n) - y_n = [(a_n - b + x_n) - y_n] + (b_n + b - x_n) \to 0 + 0 = 0.$$

Kertolasku. Olkoon $\lim_n a_n = \breve{x} = \{x_n\} \in \mathbb{R}$ ja $\lim_n b_n = \breve{y} = \{y_n\} \in \mathbb{R}$. Jos $\breve{x}$:llä ja $\breve{y}$:llä on sama etumerkki, niin jono $\{x_n y_n\} = \{|x_n||y_n|\}$ on kasvava (koska $\{|x_n|\}$ ja $\{|y_n|\}$, ovat kasvavia, vrt. Harj.teht. I.5:2), muussa tapauksessa $\{x_n y_n\} = \{-|x_n||y_n|\}$ on vähenevä jono. Lauseen I.8.3 mukaan on siis olemassa raja-arvo $\lim_n x_n y_n = \breve{z} = \{z_n\} \in \mathbb{R}$. Tämä on myös lukujonon $\{a_n b_n\}$ raja-arvo Määritelmän I.8.1 mukaisesti, sillä ko. määritelmän ja Lauseen I.7.4 (V3) mukaan

$$a_n b_n - z_n = (a_n - x_n)b_n + (b_n - y_n)x_n + (x_n y_n - z_n) \to 0 + 0 + 0 = 0.$$

Jakolasku. Koska $\{b_n\}$ on monotoninen lukujono, niin tehdyin lisäoletuksin on jostakin indeksistä $n = m$ alkaen oltava joko $b_n \geq b$ tai $b_n \leq -b$. Tällöin lukujono

$\{1/b_n\}$ on monotoninen indeksistä $n = m$ alkaen. Kirjoitetaan tästä indeksistä eteenpäin $a_n/b_n = a_n(1/b_n)$ ja vedotaan jo käsiteltyyn kertolaskuun. $\square$

Koska Määritelmän I.9.2 lukujonot $\{x_n\}$ ja $\{y_n\}$ ovat monotonisia ja rajoitettuja, niin Lause I.9.3 takaa määritelmän toimivuuden. Havainnollistettakoon lauseen todistuksessa käytettyä päättelyä vielä laskuesimerkillä.

Esimerkki 1 Olkoon

$$\breve{\pi} = 3.141592653\ldots = \{3, 3.1, 3.14, 3.141\ldots\} = \{p_n\},$$
$$\breve{e} = 2.718281828\ldots = \{2, 2.7, 2.71, 2.718\ldots\} = \{e_n\}.$$

Määritä a) $\breve{\pi} + \breve{e}$ ja $\breve{\pi}\breve{e}$, b) $\breve{\pi} - \breve{e}$, c) $\breve{\pi}/\breve{e}$ Määritelmän I.9.2 mukaisesti viiden desimaalin tarkkuudella käyttämällä Lauseen I.8.3 todistusta algoritmina.

<u>Ratkaisu</u> a) Lukujonot

$$\{p_n + e_n\} = \{5, 5.8, 5.85, 5.859\ldots\}, \quad \{p_n e_n\} = \{6, 8.37, 8.5094, 8.537238\ldots\}$$

ovat molemmat kasvavia ja rajoitettuja, joten ne suppenevat kohti reaalilukua (Lause I.8.3). Lauseen I.8.3 todistuskonstruktiota algoritmina käyttäen saadaan

$$\breve{\pi} + \breve{e} = 5.85987\,.., \quad \breve{\pi}\breve{e} = 8.53974\ldots$$

b) Lukujono $\{p_n - e_n\} = \{0, 0.4, 0.43, 0.423, \ldots\}$ ei ole monotoninen, joten Lause I.8.3 ei sovellu suoraan. Kirjoitetaan sen vuoksi ensin

$$p_n - e_n = (p_n - 3) + (3 - e_n).$$

Tässä $\{3 - e_n\}$ on rajoitettu ja vähenevä lukujono. Sen raja-arvoksi saadaan Lauseen I.8.3 todistuskonstruktiolla

$$\lim_n (3 - e_n) = 3 - \breve{e} = 0.28171\,..$$

Kun tämä välitulos tulkitaan lukujonoksi $\{0, 0.2, 0.28, 0.281, \ldots\} = \{x_n\}$, niin $\{p_n - 3\}$ ja $\{x_n\}$ ovat molemmat kasvavia lukujonoja. Siis myös $\{(p_n - 3) + x_n\}$ on kasvava, ja voidaan laskea (algoritmi sama kuin dellä)

$$\breve{\pi} - \breve{e} = \lim_n [(p_n - 3) + x_n] = 0.42331\,..$$

c) Lukujonon $\{p_n/e_n\}$ ei ole monotoninen, joten Lause I.8.3 ei sovellu tähänkään suoraan. Lasketaan ensin välituloksena vähenevän lukujonon $\{1/e_n\}$ raja-arvo:

$$\lim_n (1/e_n) = 1/\breve{e} = 0.36787\,..$$

Kun välitulos tulkitaan lukujonoksi $\{x_n\} = \{0, 0.3, 0.36, 0.367, \ldots\}$, niin $\{p_n x_n\}$ on kasvava lukujono. Tähän Lauseen I.8.3 todistuskonstruktio soveltuu ja antaa

$$\breve{\pi}/\breve{e} = \breve{\pi}(1/\breve{e}) = \lim_n p_n x_n = 1.15572\,.. \quad \square$$

$\mathbb{R}$:n järjestysrelaatio

Myös järjestysrelaation määritelmä reaaliluvuille on suoraviivainen.

MÄÄRITELMÄ I.9.4 (**$\mathbb{R}$:n järjestysrelaatio**) Jos $\breve{x} = \{x_n\} \in \mathbb{R}$, $\breve{y} = \{y_n\} \in \mathbb{R}$, niin $\breve{x} < \breve{y}$ täsmälleen kun $x_n < y_n$ jollakin n ja $\breve{x} \neq \breve{y}$.

Määritelmän I.9.4 mukaisesti voidaan reaalilukujen $\breve{x} = \{x_n\}$ ja $\breve{y} = \{y_n\}$ suuruusjärjestys ratkaista laskennallisesti vertailemalla lukuja x_n, y_n, $n = 0, 1, \ldots$ Vertailua jatketaan, kunnes tavataan ensimmäinen indeksi, jolla $x_n \neq y_n$. (Jollei tällaista indeksiä tavata, on $\breve{x} = \breve{y}$.) Jos $|x_n - y_n| \geq 2 \cdot 10^{-n}$, on suuruusjärjestys ratkennut. Muussa tapauksessa, eli jos $x_n - y_n = \pm 10^{-n}$, on vielä mahdollisuus, että $\breve{x}$ ja $\breve{y}$ samastuvat kumpikin lukuun x_n tai y_n, jolloin on $\breve{x} = \breve{y}$ (ks. Luvun I.7 tarkastelut liittyen Lauseeseen I.7.8).

Määritelmistä I.9.1–I.9.4 on johdettavissa seuraava merkittävä tulos. Todistus esitetään luvun lopussa (harjoitustehtävillä tuettuna).

*LAUSE I.9.5 ($(\mathbb{R}, +, \cdot, <)$ **on järjestetty kunta**) Reaalilukujen joukko varustettuna Määritelmän I.9.2 mukaisilla laskuoperaatioilla ja Määritelmän I.9.4 mukaisella järjestysrelaatiolla on järjestetty kunta.

Määritelmästä I.9.4 seuraa välittömästi, että jos $\breve{x} = \{x_n\} \in \mathbb{R}$, niin

$$\begin{aligned}
\breve{x} = 0 \quad &\Leftrightarrow \quad x_n = 0 \ \ \forall n \\
\breve{x} < 0 \quad &\Leftrightarrow \quad x_n < 0 \ \ \text{jollakin } n \quad (\text{etumerkki } e = -) \\
\breve{x} > 0 \quad &\Leftrightarrow \quad x_n > 0 \ \ \text{jollakin } n \quad (\text{etumerkki } e = +)
\end{aligned}$$

Tämä vertailu siis ratkeaa pelkästään $\breve{x}$:n etumerkin perusteella, ellei ole $\breve{x} = 0$. $\mathbb{R}$:n järjestysrelaation määrittely voidaan vaihtoehtoisesti perustaa tähän vertailuun yhdistettynä vähennyslaskuun (vrt. $\mathbb{Q}$:n järjestysrelaatio, Luku I.1):

$$\breve{x} < \breve{y} \ \Leftrightarrow \ \breve{x} - \breve{y} < 0.$$

On ilmeistä, että järjestysrelaation aksioomista ainakin (J1) (vrt. Luku I.1) on voimassa kummalla tahansa määritelmällä.

Reaaliluvut abstrakteina lukuina

Kun reaaliluvuille on määritelty sekä laskuoperaatiot että järjestysrelaatio, voidaan reaalilukujen olemus jonoina haluttaesssa 'unohtaa' ja käsitellä lukuja vain abstrakteina lukuina, joilla voi laskea ja joita voi vertailla. Tämän ajattelutavan mukaisesti pidetään jatkossa myös reaalilukuja 'oikeina' lukuina, joita ei symbolisissa merkinnöissä enää erotella rationaaliluvuista.

Esimerkki 2 Luvun I.8 tulosten ja Määritelmien I.9.1– I.9.2 perusteella voidaan nyt kirjoittaa

$$\lim_n \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!} = 2.71828182845905\,.. = e \in \mathbb{R},$$
$$\lim_n \left(1-\frac{1}{n}\right)^n = 0.36787944117144\,.. = \frac{1}{e}\,.$$

Ottamalla käyttöön toinen hyvin tunnettu reaaliluku $\pi = 3.1415926535897\,..$ voidaan myös osoittaa, että (vrt. Luku I.8)

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1.64493406684822\,.. = \frac{\pi^2}{6}\,.$$

Reaalilukujen tultua määritellyksi järjestettynä kuntana tulee myös lukujonon raja-arvon alkuperäisestä määritelmästä (Määritelmä I.6.1) pätevä, kun jonon termit ja/tai raja-arvo ovat reaalilukuja. Koska lukujonon suppeneminen kohti reaalista raja-arvoa on aiemmin määritelty erikseen (Määritelmä I.8.1), tarvitaan ympyrän sulkemiseksi seuraava tulos (todistus luvun lopussa).

*Lause I.9.6 Määritelmät I.6.1 ja I.8.1 lukujonon suppenemiselle kohti reaalilukua ovat yhtäpitävät.

Kymmenjako- ja puolituskonstruktiot

Määritelmän I.9.1 mukaan reaalilukua voi laskennallisesti 'vain lähestyä, ei saavuttaa', paitsi jos luku on rationaalinen. Algoritmisesti voi 'lähestyminen' tapahtua esim. suoraan Määritelmään I.9.1 perustuen, jolloin lasketaan ensin luvun kokonaislukuosa ja sen jälkeen desimaaleja yksi kerrallaan. Tarkastellaan nyt hieman yleisemmin tätä laskemisen ongelmaa.

Olkoon reaaliluku a määritelty yksikäsitteisesti, esimerkiksi asettamalla luvulle jokin algebrallinen ehto, määrittelemällä luku jonkin tunnetun jonon raja-arvona, tai jollakin muulla tavalla. Halutaan konstruoida a Määritelmän I.9.1 mukaisena desimaalilukuna, eli muodossa $a = x_0.d_1d_2\ldots$ Oletetaan jatkossa, että a:n etumerkki on $+$ (muuten konstruoidaan luku $-a$). Tällöin a:n kokonaislukosa ja desimaalit luvuissa $x_n = x_0.d_1\ldots d_n$ määrytyvät seuraamalla toimintaohjetta:

Etsi $x_n \in \mathbb{Q}_n$ siten, että pätee $x_n \le a$ ja $x_n + 10^{-n} > a, \quad n = 0, 1, \ldots$

Jos a on rationaaliluku, niin toimintaohje on sama kuin jakokulma-algoritmissa (vrt. Luku I.5). Yleisemmän reaaliluvun $a > 0$ tapauksessa toimintaohjetta voi pitää ajatuskonstruktiona, joka <u>määrittelee</u> luvun a. Annetaan tälle nimi *kymmenjako*konstruktio. — Kymmenjakokonstruktiota on jo aiemmin käytetty Lauseen I.8.3 todistuksessa. Reaaliluvun määritelmänä konstruktiota voi pitää järjestysrelaatioon nojaavana Määritelmän I.9.1 toisintona tai täsmennyksenä.

Kymmenjakokonstruktion on jo aiemmin todettu toimivan myös algoritmina (vrt. edellisen luvun esimerkit ja harjoitustehtävät). Yleisemmin algoritmi toimii, jos oletetaan, että konstruktioon sisältyvät vertailut ovat laskennallisesti toteutettavissa. Näin on ainakin, jos oletetaan laskettava luku $a \in \mathbb{R}$ sellaiseksi, että minkä tahansa <u>rationaaliluvun</u> $x \in \mathbb{Q}$ kohdalla voidaan selvittää <u>äärellisellä</u> määrällä <u>rationaalisia</u> operaatioita (laskuoperaatioita ja vertailuja), mikä vaihtoehdoista $a < x$, $a = x$, $a > x$ on voimassa. Tällöin saadaan ensin a:n etumerkki selville vertaamalla lukuun $x = 0$, minkä jälkeen a:n tai $-a$:n kokonaislukuosa ja desimaalit d_n voidaan laskea ym. toimintaohjetta seuraamalla, periaatteessa mihin tahansa haluttuun indeksiin asti. Tällainen *kymmenjakoalgoritmi* on siis jakokulma-algoritmin yleistys.

Esimerkki 3 Konstruoi reaaliluku $a = \sqrt{2}$ kymmenjakoalgoritmilla.

<u>Ratkaisu</u> Luvun a ja positiivisen rationaaliluvun x vertailu palutuu rationaaliseksi ekvivalenssilla

$$a < x \;\Leftrightarrow\; 2 < x^2.$$

Koska $1^2 < 2$ ja $2^2 > 2$, on ensinnäkin oltava $a = 1.d_1d_2d_3..$ Edelleen koska $1.4^2 = 1.96 < 2$, $1.5^2 = 2.25 > 2$, $1.41^2 = 1.9881 < 2$, $1.42^2 = 2.0164 > 2$, $1.414^2 = 1.999396 < 2$ ja $1.415^2 = 2.002225 > 2$, on $d_1 = 4$, $d_2 = 1$ ja $d_3 = 4$. Jatkamalla tällä tavoin on tuloksena (jaksoton) desimaaliluku $a = 1.4142135623730950488016887..$[†] $\square$

Esimerkki 4 Laske luku $a = \sqrt{\sqrt{2}+1}$ 20 merkitsevän numeron tarkkuudella käyttäen kymmenjakoalgoritmia.

<u>Ratkaisu</u> Verrattaessa lukua a rationaalilukuun $x \geq 1$ pätee

$$a < x \quad\Leftrightarrow\quad \sqrt{2}+1 < x^2 \quad\Leftrightarrow\quad \sqrt{2} < x^2-1 \quad\Leftrightarrow\quad 2 < (x^2-1)^2,$$

joten vertailu palautuu rationaalioperaatioksi. Algoritmia seuraten saadaan

$$a \;=\; 1.55377397403003730 73..\qquad \square$$

[†]Ennen laskimia on neliöjuurien käsinlaskua esimerkin tapaan harjoiteltu kouluissakin. – Nykyisin laskimet ja tietokoneet laskevat neliöjuuria kymmenjakoa paljon tehokkaammilla, palautuviin lukujonoihin perustuvilla algoritmeilla, vrt. edellisen luvun Esimerkki 4.

Esimerkeissä vertailun $a < x$ palauttaminen rationaaliseksi perustuu *symboliseen* (ei-numeeriseen) *laskentaan*, tässä algebraan, joka nojaa viime kädessä reaalilukujen kunta-aksioomiin, järjestysrelaation aksioomiin ja neliöjuuren symboliseen määritelmään $(\sqrt{a})^2 = a$. Kuten näissä esimerkeissä, symbolisen laskennan tehtävänä on yleensä yksinkertaistaa tai selkeyttää laskentatehtävää ennen varsinaisia numeerisia laskuja.

Kymmenjaon vastine voidaan luonnollisesti konstruoida myös muihin lukujärjestelmiin perustuvana. Binaarijärjestelmän tapauksessa käytetään nimitystä *puolituskonstruktio* (tai puolitusmenetelmä, engl. bisection). Puolituskonstruktiossa reaaliluku $a \geq 0$ konstruoidaan binaarimuodossa $a = x_0.b_1b_2b_3..$, missä kokonaislukuosa x_0 ja bitit $b_n \in \{0,1\}$ valitaan siten, että luvut x_0 ja $x_n = x_0 + \sum_{k=1}^{n} b_k \cdot 2^{-k}$, $n \in \mathbb{N}$, toteuttavat

$$x_n \leq a \quad \text{ja} \quad x_n + 2^{-n} > a, \quad n = 0, 1, \ldots$$

Yksinkertaisen logiikkansa vuoksi puolitusmenetelmää käytetään matemaattisten todistusten ajatuskonstruktioissa hyvin usein. (Puolituskonstruktio olisi ollut vaihtoehto myös Lauseen I.8.3 todistuksessa.) Ym. lisäoletusten voimassa ollessa voi konstruktiota käyttää myös toimivana algoritmina.

Reaalilukujen jonot

Reaalilukujen jono-olemuksen 'unohtaminen' on erityisen suositeltavaa silloin, kun tutkitaan *reaalilukujonoja* ja niiden suppenemista. Järjestysrelaatioon sekä (vähäisessä määrin) kunta-algebraan vetoava Määritelmä I.6.1 on sellaisenaan pätevä myös reaalilukujonon suppenemisen määritelmänä. Jos reaalilukujonolla on tämän määritelmän mukainen raja-arvo (reaalilukuna), niin sanotaan, että jono *suppenee*, muussa tapauksessa *hajaantuu* (vrt. Määritelmä I.8.1).

Esimerkki 5 Määritä perusmuotoisen geometrisen sarjan summa, kun $q = 1/\sqrt{2}$.

Ratkaisu Luvussa I.6 johdettu geometrisen sarjan summakaava on pätevä, kun $q \in \mathbb{R}$ ja $|q| < 1$. Tässä on $0 < q < 1$, joten

$$\sum_{n=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^n = \frac{1}{1 - \frac{1}{\sqrt{2}}} = \frac{\sqrt{2}}{\sqrt{2} - 1} = 2 + \sqrt{2} = 3.4142135623730950488\ldots \quad \square$$

Esimerkin lasku on jälleen esimerkki myös symbolisesta laskennasta, jossa lukua $\sqrt{2}$ käsitellään abstraktina lukuna pelkästään kunnan $(\mathbb{R}, +, \cdot)$ yleisiä ominaisuuksia (kunta-aksioomia) ja symbolista määritelmää $(\sqrt{2})^2 = 2$ hyväksi käyttäen.

Todettakoon, että kaikki Luvuissa I.6–I.8 esitetyt lukujonoja koskevat määritelmät ja väittämät ovat päteviä reaalilukujen kunnassa $(\mathbb{R}, +, \cdot, <)$ — syystä, että nämä perustuvat vain järjestetyn kunnan algebraan ja oletukseen, että kunta sisältää rationaaliluvut.

Esimerkkinä reaalilukujonon suppenemistarkastelusta todistettakoon

PROPOSITIO I.9.7 $$\boxed{\lim_n \sqrt[n]{a} = 1 \quad \forall a \in \mathbb{R},\ a > 0.}$$

<u>Todistus</u> Merkitään $b_n = \sqrt[n]{a}$. Tapauksessa $a = 1$ on väittämä ilmeisen tosi ja tapauksessa $0 < a < 1$ pätee: $b_n = 1/(1/a)^{1/n} \to 1$, jos $(1/a)^{1/n} \to 1$ (Lause I.7.3 reaalilukujonoille). Riittää siis tarkastella tapausta $a > 1$. Tällöin on oltava $b_n > 1\ \forall n$ (koska $b_n^n = a > 1$), jolloin seuraa

$$\left(\frac{b_{n+1}}{b_n}\right)^n = \frac{1}{b_{n+1}}\frac{b_{n+1}^{n+1}}{b_n^n} = \frac{1}{b_{n+1}}\frac{a}{a} = \frac{1}{b_{n+1}} < 1.$$

Siis $b_{n+1}/b_n < 1\ \forall n$, joten $\{b_n\}$ on aidosti vähenevä lukujono. Koska $\{b_n\}$ on myös ilmeisen rajoitettu lukujono, niin $b_n \to b \in \mathbb{R}$ (Lause I.8.3), ja koska $b_n > 1\ \forall n$, niin $b \geq 1$ (Propositio I.7.4 (V1) reaalilukujonoille). Tässä vaihtoehto $b > 1$ johtaisi loogiseen ristiriitaan: $a = b_n^n \geq b^n \to \infty$. Siis $\lim_n b_n = 1$. $\square$

*Lauseiden I.9.5 ja I.9.6 todistukset

Lause I.9.5. Kunta-aksioomista todistetaan esimerkkinä ainoastaan aksioomien (K4) (kertolaskun liitäntälaki) ja (K9) (käänteisalkio) voimassaolo, muut jätetään harjoitustehtäväksi (Harj.teht. 1).

$\boxed{\text{K4}}$ Olkoon $\breve{x} = \{x_n\}$, $\breve{y} = \{y_n\}$, $\breve{z} = \{z_n\}$ reaalilukuja ja olkoon edelleen $\breve{x}\breve{y} = \breve{a} = \{a_n\} \in \mathbb{R}$ ja $\breve{y}\breve{z} = \breve{b} = \{b_n\} \in \mathbb{R}$. Tällöin aksiooman (K4) sisältö on Määritelmien I.9.2 ja I.8.1 perusteella

$$\breve{a}\breve{z} = \breve{x}\breve{b} \quad \Leftrightarrow \quad \lim_n (a_n z_n - x_n b_n) = 0.$$

Kirjoitetaan tässä

$$a_n z_n - x_n b_n = (a_n - x_n y_n) z_n + x_n (y_n z_n - b_n).$$

Koska tässä $a_n - x_n y_n \to 0$ ja $y_n z_n - b_n \to 0$ (Määritelmät I.9.2 ja I.8.1) ja jonot $\{x_n\}$ ja $\{z_n\}$ ovat rajoitettuja, niin väite $a_n z_n - x_n b_n \to 0$ seuraa Lauseen I.7.3 ja Proposition I.7.4 (V3) perusteella.

$\boxed{\text{K9}}$ Luvun $\breve{x} = \{x_n\} \in \mathbb{R}$ käänteisluvun $\breve{x}^{-1}$ konstruoimiseksi olkoon $\breve{x} \neq 0$, jolloin jollakin $m \in \mathbb{N}$ on oltava $|x_n| \geq |x_m| > 0 \quad \forall n \geq m$. Tällöin jono $\{\, x_m^{-1}, x_{m+1}^{-1}, \dots \}$ on monotoninen ja rajoitettu, jolloin on olemassa reaaliluku $\breve{y} = \{y_n\}$ siten, että $x_n^{-1} - y_n \to 0$ (Lause I.8.3). Proposition I.7.4 (V3) ja Määritelmän I.9.2 perusteella päätellään tällöin

$$x_n^{-1} - y_n \to 0 \quad \Leftrightarrow \quad 1 - x_n y_n \to 0 \quad \Leftrightarrow \quad \breve{x}\breve{y} = 1.$$

Siis $\breve{y} = \breve{x}^{-1}$ ja aksiooma (K9) on näin ollen voimassa. $\square$

Järjestyrelaation aksiooman (J1) voimassaolo on jo todettu. Aksioomien (J2) ja (J4) todentaminen jätetään harjoitustehtäväksi (Harj.teht. 2cd). Todistetaan siis ainoastaan aksiooman (J3) voimassaolo.

Propositio I.9.8 Reaaliluvuille pätee: Jos $\breve{x} < \breve{y}$, niin $\breve{x} + \breve{z} < \breve{y} + \breve{z} \ \forall \breve{z} \in \mathbb{R}$.

<u>Todistus</u> Olkoon $\breve{x} = \{x_n\} \in \mathbb{R}$, $\breve{y} = \{y_n\} \in \mathbb{R}$, $\breve{z} = \{z_n\} \in \mathbb{R}$ ja merkitään $\breve{x} + \breve{z} = \breve{a} = \{a_n\} \in \mathbb{R}$ ja $\breve{y} + \breve{z} = \breve{b} = \{b_n\} \in \mathbb{R}$, jolloin $x_n + z_n - a_n \to 0$ ja $y_n + z_n - b_n \to 0$, kun $n \to \infty$ (Määritelmä I.9.2). Väitetään: Jos $\breve{x} < \breve{y}$, niin $\breve{a} < \breve{b}$. Tehdään vastaoletus: $\breve{a} = \breve{b}$ tai $\breve{a} > \breve{b}$. Ensimmäisessä vaihtoehdossa on $\lim_n (a_n - b_n) = 0$ (Määritelmä I.7.7), jolloin oletusten ja Lauseen I.7.3 perusteella seuraa

$$x_n - y_n \;=\; (x_n + z_n - a_n) - (y_n + z_n - b_n) + (a_n - b_n) \;\to\; 0.$$

Siis $\breve{x} = \breve{y}$ (Määritelmä I.7.7), mutta tämä on looginen ristiriita, koska (J1) on voimassa ja oletettiin $\breve{x} < \breve{y}$. Toisessa vaihtoehdossa ($\breve{a} > \breve{b}$) on $a_n \geq b_n \ \forall n$ (Harj.teht. 2a) ja samoin $x_n \leq y_n \ \forall n$ (koska $\breve{x} < \breve{y}$), joten seuraa

$$\begin{aligned} c_n \;&=\; (x_n + z_n - a_n) - (y_n + z_n - b_n) \\ &=\; (x_n - y_n) + (b_n - a_n) \;\leq\; x_n - y_n \;\leq\; 0 \ \forall n. \end{aligned}$$

Koska tässä $c_n \to 0$, niin $x_n - y_n \to 0$ (Propositio I.7.4 (V2)), joten on jälleen päädytty loogiseen ristiriitaan: $\breve{x} < \breve{y}$ ja $\breve{x} = \breve{y}$. Vaihtoehdot $\breve{a} = \breve{b}$ ja $\breve{a} > \breve{b}$ on näin muodoin pois suljettu, joten on oltava $\breve{a} < \breve{b}$. $\square$

Lause I.9.6. Perustetaan todistus seuraavaan väittämään, joka seuraa helposti Määritelmästä I.9.4 (Harj.teht. 2b).

Lemma I.9.9 Jos $\breve{x} = \{x_n\} \in \mathbb{R}$, niin $|\breve{x} - x_n| \leq 10^{-n} \ \forall n$.

Oletetaan, että $a_n \to \breve{x}$, $\breve{x} = \{x_n\} \in \mathbb{R}$ Määritelmän I.6.1 mukaisesti ja olkoon $\varepsilon > 0$. Tällöin on olemassa indeksit $N_1, N_2 \in \mathbb{N}$ siten, että $|a_n - \breve{x}| < \varepsilon/2$ kun

$n > N_1$ ja $10^{-n} < \varepsilon/2$ kun $n > N_2$, jolloin (järjestetyn kunnan) kolmioepäyhtälön ja Lemman I.9.9 perusteella

$$|a_n - x_n| \le |a_n - \breve{x}| + |\breve{x} - x_n| < |a_n - \breve{x}| + 10^{-n} < \varepsilon, \quad \text{kun } n > \max\{N_1, N_2\} = N.$$

Tässä $\varepsilon > 0$ oli mielivaltainen, joten Määritelmän I.6.1 mukaan $a_n - x_n \to 0$, eli $a_n \to \breve{x}$ Määritelmän I.8.1 mukaisesti. Tämä todistaa väittämän ensimmäisen osan. Toinen osa todistetaan samankaltaisella päättelyllä. $\square$

HARJOITUSTEHTÄVIÄ

1. Näytä Määritelmään I.9.2 perustuen, että reaaliluvuille ovat voimassa:
   a) yhteen- ja kertolaskun vaihdantalait (K1,K2)
   b) yhteenlaskun liitäntälaki (K3)
   c) yhteen- ja kertolaskun osittelulaki (K5)

2. Olkoon $\breve{x} = \{x_n\} \in \mathbb{R}$, $\breve{y} = \{y_n\} \in \mathbb{R}$ ja $z = \{z_n\} \in \mathbb{R}$. Näytä, että Määritelmien I.9.4 ja I.9.2 perusteella pätee
   a) $\breve{x} < \breve{y} \;\Leftrightarrow\; \breve{x} \ne \breve{y} \;\&\; x_n \le y_n \;\forall n$ b) $|\breve{x} - x_n| \le 10^{-n} \;\forall n$
   c) $\breve{x} < \breve{y} \;\&\; \breve{y} < \breve{z} \;\Rightarrow\; \breve{x} < \breve{z}$ d) $\breve{x} > 0 \;\&\; \breve{y} > 0 \;\Rightarrow\; \breve{x}\breve{y} > 0$

3. Tutki, kuinka suuri virhe tehdään, kun laskettaessa a) $e + \pi$, b) e^2/π, c) $\pi^2 - e^2$ c) $e^5\pi^6$ katkaistaan e ja π ensin 9 merkitsevään numeroon ja laskuoperaatioissa tulos samoin 9 merkitsevään numeroon.

4. Olkoon lukujärjestelmän kantaluku $= k$. Käyttäen kymmenjakoalgoritmia vastaavaa k-jakoalgoritmia konstruoi neljän merkitsevän numeron tarkkuudella a) $\sqrt[3]{11}$ binaarijärjestelmässä, b) $\sqrt{7}$ 3-kantaisessa järjestelmässä.

5. (*) Tiedetään, että $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$. Näytä tämän tiedon perusteella:
   a) $\displaystyle\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$ b) $\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)^2} = \frac{\pi^2}{12}$

6. (*) Olkoon $a_k \to a$ $(k \to \infty)$, missä on desimaalilukumerkinnöin
   $$a_k = x_0^{(k)}.\, d_1^{(k)} d_2^{(k)} \ldots = \{x_n^{(k)}\} \in \mathbb{R}, \quad a = x_0.d_1 d_2 \ldots = \{x_n\} \in \mathbb{R}.$$
   Todista seuraavat väittämät:
   a) Jos a ei ole äärellinen desimaaliluku, niin jokaisella $n \in \mathbb{N}$ on olemassa indeksi $N_n \in \mathbb{N}$ siten, että $a_k = x_0.d_1 \ldots d_n \ldots \;\forall k > N_n$.
   b) a-kohdan väittämä ei ole tosi jokaisella $a \in \mathbb{R}$ (vastaesimerkki!).
   c) Jos a on äärellinen desimaaliluku, niin on olemassa $m \in \mathbb{N}$ ja jokaisella $n \ge m$ indeksi $N_n \in \mathbb{N}$ siten, että $d_n^{(k)} \in \{0, 9\} \;\forall k > N_n$.
   d) Jokaisella $a \in \mathbb{R}$ pätee: $\lim_k x_k^{(k)} = a$.

I.10 Cauchyn jonot

Tässä luvussa 'lukujono' tarkoittaa reaalilukujen jonoa. Tarkastelun kohteena on lukujonojen (myös rationaalilukujonojen) teorian edelleen avoin kysymys, joka kuuluu: Täsmälleen millaisilla, jonoa itseään koskevilla ehdoilla lukujono $\{a_n\}$ suppenee, ts. raja-arvo $\lim_n a_n = a$ on olemassa Määritelmän I.6.1 mukaisesti reaalilukuna? Toistaiseksi tunnetut suppenemisen ehdot Lauseissa I.8.3 ja I.9.3 ovat riittäviä, eivät välttämättömiä. Etsittäessä suppenemiskysymyksen täydellista ratkaisua, ts. sekä välttämättömiä että riittäviä ehtoja suppnemiselle, ratkaisun avaimeksi osoittautuu

MÄÄRITELMÄ I.10.1 (**Cauchyn**[†] **jono**) Lukujono $\{a_n\}$ on *Cauchy* eli *Cauchyn jono*, jos jokaisella $\varepsilon > 0$ on olemassa $N \in \mathbb{N}$ siten, että pätee

$$|a_n - a_m| < \varepsilon \quad \text{kun } n, m > N.$$

ESIMERKKI 1 Näytä, että sarjan $\sum_{k=0}^{\infty}(-1)^k/(k+1)^2$ osasummien jono

$$\{s_n\} = \left\{1,\, 1 - \frac{1}{2^2},\, 1 - \frac{1}{2^2} + \frac{1}{3^2}, \ldots\right\} = \left\{1,\, \frac{3}{4},\, \frac{7}{9}, \ldots\right\}$$

on Cauchyn jono.

<u>Ratkaisu</u> Jos $0 \le m < n$, niin kolmioepäyhtälöä soveltaen voidaan arvioida (ks. Harj.teht. I.8:1)

$$|s_n - s_m| = \left|\sum_{k=m+1}^{n} \frac{(-1)^k}{(k+1)^2}\right| \le \sum_{k=m+1}^{n} \frac{1}{(k+1)^2} < \frac{1}{m+1}\,.$$

Näin ollen jos $\varepsilon > 0$ ja valitaan $N \in \mathbb{N}$ siten, että $1/(N+1) < \varepsilon$ (mahdollista jokaisella $\varepsilon > 0$), niin

$$|s_n - s_m| < \max\left\{\frac{1}{n+1}, \frac{1}{m+1}\right\} < \varepsilon, \quad \text{kun } n, m > N.$$

Siis $\{s_n\}$ on Cauchy. $\square$

ESIMERKKI 2 Jos $a \neq 0$, niin lukujono $\{a_n\} = \{(-1)^n a\}$ ei ole Cauchy. Nimittäin jos $0 < \varepsilon \le 2|a|$, niin Määritelmän I.10.1 ehto ei toteudu millään $N \in \mathbb{N}$, koska $|a_n - a_m| = 2|a| \ge \varepsilon$ aina kun $n - m$ on pariton. $\square$

[†]Ranskalainen matemaatikko **Augustin Louis Cauchy** (1789-1857) kuuluu matemaatikoiden tähtikaartiin kautta aikojen. Cauchy täsmensi merkittävästi matematiikan käsitteistöä ja loi pohjaa uusille tutkimussuunnille. Hänen laaja tuotantonsa ulottui myös fysiikkaan, mm. kiinteän aineen kimmoteoriaan.

Cauchyn jonoilla on samankaltaisia ominaisuuksia kuin suppenevilla jonoilla. Näytetään ensinnäkin:

LAUSE I.10.2 Suppeneva jono on Cauchy.

Todistus Jos $\lim_n a_n = a \in \mathbb{R}$ Määritelmän I.6.1 mukaisesti, niin jokaisella $\varepsilon > 0$ on olemassa $N \in \mathbb{N}$ siten, että pätee $|a_n - a| < \varepsilon/2$ kun $n > N$. Tällöin on kolmioepäyhtälön nojalla

$$\begin{aligned} |a_n - a_m| &= |(a_n - a) + (a - a_m)| \\ &\leq |a_n - a| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \text{kun } n, m > N. \quad \square \end{aligned}$$

Suppenevia lukujonoja koskevilla Lauseilla I.7.2 ja I.7.3 on seuraavat vastineet Cauchyn jonoille:

LAUSE I.10.3 Cauchyn jono on rajoitettu.

LAUSE I.10.4 (**Cauchyn jonojen yhdistelysäännöt**) Jos lukujonot $\{a_n\}$ ja $\{b_n\}$ ovat Cauchyn jonoja ja $\lambda \in \mathbb{R}$, niin myös lukujonot $\{a_n + b_n\}$, $\{\lambda a_n\}$ ja $\{a_n b_n\}$ ovat Cauchyn jonoja. Jos lisäksi $|b_n| \geq \delta > 0 \ \ \forall n$, niin myös $\{a_n/b_n\}$ on Cauchyn jono.

Näiden lauseiden todistukset (jotka sivuutetaan, ks. Harj.teht. 1) ovat hyvin samanlaisia kuin mainittujen vastinlauseiden. Esimerkiksi Lauseen I.10.3 väittämään päädytään, kun vertailukohdaksi otetaan raja-arvon sijasta jonon termi a_m, missä indeksi m valitaan siten, että pätee $|a_n - a_m| < 1$ kun $n > m$, vrt. Määritelmä I.10.1 ja Lauseen I.7.2 todistus.

Päätulos jatkossa on

*LAUSE I.10.5 (**Cauchyn suppenemiskriteeri**) Reaalilukujono $\{a_n\}$ suppenee kohti reaalilukua täsmälleen kun $\{a_n\}$ on Cauchy.

ESIMERKKI 1 (jatko) Lauseen I.10.5 perusteella esimerkin sarja suppenee. Osasummia riitttävän pitkälle laskemalla selviää, että sarjan summa katkaistuna 10 merkitsevään numeroon on

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)^2} = 0.8224670334\ldots \quad \square$$

Koska jo tiedetään, että jokainen suppeneva jono on Cauchy (Lause I.10.2), niin Lauseen I.10.5 todistamiseksi riittää näyttää todeksi implikaatio

$$\{a_n\} \text{ on Cauchy} \quad \Rightarrow \quad \{a_n\} \text{ suppenee.} \qquad (\star)$$

Todistus on useampivaiheinen. Ensin on tutkittava lukujonon raja-arvon määritelmää tarkemmin *osajonon* käsitteen pohjalta. Seuraavassa nämä tarkastelut ja niihin perustuva väittämän ($\star$) todistus erotetaan omaksi osaluvukseen†.

*Osajonot

MÄÄRITELMÄ I.10.6 Lukujono $\{b_k\}_{k=1}^{\infty}$ on jonon $\{a_n\}_{n=m}^{\infty}$ *osajono* (engl. subsequence), jos on olemassa indeksit $m \leq n_1 < n_2 < \dots$ siten, että

$$b_k = a_{n_k}, \quad k = 1, 2, \dots$$

ESIMERKKI 3 Jos $\breve{x} = \{x_n\}_{n=p}^{\infty}$ on jaksollinen desimaaliluku, niin jollakin $m \in \mathbb{N}$ ja riittävän suurella $k \in \mathbb{N}$ osajono

$$\{x_k, x_{k+m}, x_{k+2m}, \dots\} \;=\; \{x_{k+(l-1)m}\}_{l=1}^{\infty}$$

on geometrinen sarja, vrt. Luku I.6. Tässä siis osajonon indeksinä on l ja $n_l = k + (l-1)m$. $\square$

ESIMERKKI 4 Lukujono $\{\, a_n = (-1)^n, \quad n = 0, 1, 2, \dots \}$ ei suppene, mutta sillä on (kohti rationaalilukua) suppenevia osajonoja, esim. $\{\, a_0, a_2, a_4, \dots \}$. $\square$

ESIMERKKI 5 Rajatta kasvavan lukujonon $\{n\} = \{1, 2, 3, \dots\}$ jokainen osajonokin on rajatta kasvava. $\square$

Osajonon käsitettä valaisee hieman seuraava 'lämmittelylause'.

LAUSE I.10.7 $\lim_n a_n = a \in \mathbb{R}$ täsmälleen kun $\lim_k b_k = a$ jokaiselle jonon $\{a_n\}$ osajonolle $\{b_k\}$.

<u>Todistus</u> Väittämän osa $\boxed{\Rightarrow}$ on ilmeinen, sillä Määritelmän I.6.1 perusteella osajono suppenee vähintään yhtä nopeasti kuin itse jono. Tästä huolimatta osa $\boxed{\Leftarrow}$ on vieläkin ilmeisempi, sillä osajonoksi kelpaa myös itse jono $\{a_n\}$. $\square$

Seuraava tulos, jossa osajonon käsite on keskeinen, on avain jatkon kannalta. Tuloksessa tulkitaan proposition $a_n \to a$ negaatio ($a_n \not\to a$) Määritelmään I.6.1 perustuen. — Huomattakoon, että kyse ei ole negaatiosta 'ei suppene' (kohti mitään lukua), koska $a_n \not\to a$ on tosi myös, kun $a_n \to b \neq a$.

*LAUSE I.10.8 Jos $\{a_n\}$ on lukujono ja $a \in \mathbb{R}$, niin $a_n \not\to a$ täsmälleen kun $\exists\, \varepsilon > 0$ ja osajono $\{\, a_{n_1}, a_{n_2}, \dots \}$ siten, että $|a_{n_k} - a| \geq \varepsilon \;\; \forall k \in \mathbb{N}$.

†Merkintä (*) luvun tai osaluvun otsikossa kertoo (tässä ja jatkossa), että kyse on näkökulmaa laajentavasta — usein myös muuta tekstiä haastavammasta — tekstin osasta.

<u>Todistus</u> Lähdetään raja-arvon määritelmästä (Määritelmä I.6.1)

$$a_n \to a \quad \Leftrightarrow \quad \forall \varepsilon > 0 \text{ pätee } [\,\dots\,]$$

ja ryhdytään negaation purkuun: Ensinnäkin (vrt. Luku I.3)

$$a_n \not\to a \quad \Leftrightarrow \quad \text{jollakin } \varepsilon > 0 \text{ ei päde } [\,\dots\,]$$

Tässä on $[\,\dots\,] = [\ \exists N \in \mathbb{N}$ siten, että $\{\,\dots\,\}\ ]$, joten

$$\begin{aligned} \text{ei päde } [\,\dots\,] \quad &\Leftrightarrow \quad \nexists N \in \mathbb{N} \text{ siten, että } \{\,\dots\,\} \\ &\Leftrightarrow \quad ei\,\{\,\dots\,\} \ \ \forall N \in \mathbb{N}. \end{aligned}$$

Tässä on $\{\,\dots\,\} = \{\,|a_n - a| < \varepsilon \ \ \forall n > N\,\}$, joten

$$ei\,\{\,\dots\,\} \quad \Leftrightarrow \quad |a_n - a| \geq \varepsilon \ \text{ jollakin } n > N.$$

Siis on päätelty:

$$a_n \not\to a \quad \Leftrightarrow \quad \text{jollakin } \varepsilon > 0 \text{ pätee: } \ \forall\, N \in \mathbb{N}\ (\,|a_n - a| \geq \varepsilon \text{ jollakin } n > N\,)\,.$$

Kyseisellä ε voidaan nyt menetellä seuraavasti:

1. Valitaan $N = 1$ ja $n_1 > N$ siten, että $|a_{n_1} - a| \geq \varepsilon$.

2. Valitaan $N = n_1$ ja $n_2 > N$ siten, että $|a_{n_2} - a| \geq \varepsilon$.

$\vdots$

Konstruktiota voidaan jatkaa loputtomasti, joten saadaan osajono $\{\,a_{n_1}, a_{n_2}, \dots\,\}$ (ks. kuva alla). Tällä on vaadittu ominaisuus, joten on todistettu väittämän osa $\boxed{\Rightarrow}$. Osa $\boxed{\Leftarrow}$ on ilmeinen raja-arvon määritelmän perusteella. $\square$

Kun Lausetta I.10.8 sovelletaan Cauchyn jonoon, saadaan johdetuksi mielenkiintoinen tulos.

*Lause I.10.9 Jos lukujono $\{a_n\}$ on Cauchyn jono, niin jokaisella $x \in \mathbb{R}$ on voimassa täsmälleen yksi seuraavista vaihtoehdoista:

1. $a_n \to x$.
2. $\exists\, \varepsilon > 0$ ja $N \in \mathbb{N}$ siten, että $a_n < x - \varepsilon$ aina kun $n > N$.
3. $\exists\, \varepsilon > 0$ ja $N \in \mathbb{N}$ siten, että $a_n > x + \varepsilon$ aina kun $n > N$.

<u>Todistus</u> Ensinnäkin on joko $a_n \to x$ (vaihtoehto 1) tai $a_n \not\to x$. Jälkimmäisessä vaihtoehdossa on Lauseen I.10.8 mukaan löydettävissä luku $\varepsilon > 0$ ja osajono $\{\, b_k = a_{n_k},\ k = 1, 2, \dots \}$ siten, että $|b_k - x| \geq 2\varepsilon\ \forall k$. (Tässä on luvun ε tilalle kirjoitettu 2ε mukavuussyistä.) Koska $\{a_n\}$ on Cauchyn jono, on myös olemassa lukua ε vastaava indeksi N siten, että pätee $|a_n - a_m| < \varepsilon$ kun $n, m > N$. Valitaan nyt osajonossa $\{b_k\}$ indeksi k siten, että $n_k > N$, ja kirjoitetaan

$$a_n \;=\; b_k + (a_n - b_k).$$

Tässä on $b_k = a_m$, $m = n_k > N$, joten $|a_n - b_k| < \varepsilon$, kun $n > N$. Myös oli $|b_k - x| \geq 2\varepsilon$, joten on oltava joko $b_k \geq x + 2\varepsilon$ tai $b_k \leq x - 2\varepsilon$. Jos $b_k \geq x + 2\varepsilon$ (kuten kuvassa alla), niin arvioidaan

$$a_n \ \geq x + 2\varepsilon - |a_n - a_k| \ > \ x + 2\varepsilon - \varepsilon \ = \ x + \varepsilon, \quad \text{kun } n > N.$$

Jos $b_k \leq x - 2\varepsilon$, niin arvioidaan

$$a_n \ \leq x - 2\varepsilon + |a_n - a_k| \ < \ x - 2\varepsilon + \varepsilon \ = \ x - \varepsilon, \quad \text{kun } n > N.$$

Lause on näin todistettu. □

Lauseen I.10.9 mukaan Cauchyn jonon on 'valittava puolensa' annetun luvun — esimerkiksi rationaaliluvun — suhteen, ellei ko. luku ole jonon raja-arvo. Käyttäen tätä tulosta 'työhevosena' voidaan nyt todistaa väittämä ($\star$) ja siis myös Lause I.10.5. Todistus on muuten olennaisesti sama kuin Lauseen I.8.3 todistus, paitsi että monotonisuuden sijasta vedotaan Lauseeseen I.10.9.

*Lause I.10.10 Cauchyn jono suppenee.

Todistus Olkoon $\{a_n\}$ Cauchyn jono. Jos $a_n \to x \in \mathbb{Q}$, niin lauseen väittämä on tosi , joten voidaan olettaa, että $\{a_n\}$ ei suppene kohti rationaalilukua. Tällöin

on ensinnäkin jokaisen kokonaisluvun $x \in \mathbb{Z}$ kohdalla voimassa Lauseen I.10.9 vaihtoehdoista joko 2 tai 3. Koska Cauchyn jono on rajoitettu (Lause I.10.3), niin silloin on löydettävissä yksikäsitteinen $x_0 \in \mathbb{Z}$ siten, että $a_n > x_0 + \varepsilon_1$ kun $n > N_1$ ja $a_n < x_0 + 1 - \varepsilon_2$ kun $n > N_2$ (tässä $\varepsilon_1, \varepsilon_2 > 0$ ja $N_1, N_2 \in \mathbb{N}$), jolloin

$$x_0 \;<\; a_n \;<\; x_0 + 1, \quad \text{kun } n > N = \max\{N_1, N_2\}.$$

Oletetaan jatkossa, että $x_0 \geq 0$ (muussa tapauksessa tutkitaan jonoa $\{-a_n\}$). Jatkamalla päättelyä vastaavasti on löydettävissä yksikäsitteiset $d_k \in \{0, \ldots, 9\}$, $k = 1, 2, \ldots$ ja vastaavat äärelliset desimaaliluvut $x_k = x_0.d_1 \ldots d_k$ ja indeksit $N_k \in \mathbb{N}$ siten, että pätee

$$x_k \;<\; a_n \;<\; x_k + 10^{-k}, \quad \text{kun } n > N_k, \quad k = 0, 1, 2, \ldots$$

Tämän mukaan on $\lim_n (a_n - x_n) = 0$ eli $\lim_n a_n = \breve{x} = \{x_n\} \in \mathbb{R}$ Määritelmien I.8.1 ja I.9.1 mukaisesti. Lauseen I.9.6 perusteella $\lim_n a_n = \breve{x}$ myös Määritelmän I.6.1 mukaisesti. $\square$

Reaaliluvut Cauchyn jonoina

Määritelmä I.8.1 tulkitsee äärettömän desimaaliluvun 'suppenevan itseensä' lukujonona. Tällainen näkökulma voidaan ottaa yleisemminkin rationaalilukujen Cauchyn jonoihin, jolloin saadaan seuraava vaihtoehtoinen määritelmä reaaliluvuille (vrt. Määritelmät I.9.1 ja I.7.7).

MÄÄRITELMÄ I.10.11 (**Reaaliluvut Cauchyn jonoina**) Reaaliluvut ovat rationaalilukujen Cauchyn jonoja. Kaksi reaalilukua $a = \{a_n\}$ ja $b = \{b_n\}$ ovat samat täsmälleen kun $\lim_n (a_n - b_n) = 0$.

Määritelmän mukaisesti reaaliluvut ovat rationaalilukujen Cauchyn jonojen muodostamia <u>ekvivalenssiluokkia</u> (tai samastusluokkia, vrt. Luku I.3), joiden sisällä eri jonot samastetaan määritelmän kriteerillä. Jokaiseen samastusluokkaan kuuluu äärettömän monta erilaista jonoa; esim. jonot $\{a_n\}$ ja $\{a_n + n^{-k}\}$ kuuluvat samaan luokkaaan jokaisella $k \in \mathbb{N}$. Ellei kyseessä ole äärellinen desimaaliluku, on samaan samastusluokkaan kuuluvista Cauchyn jonoista täsmälleen yksi Määritelmän I.9.1 mukainen reaaliluvun 'edustaja', ts. ääretön desimaaliluku.

Määritelmään I.10.11 perustuen on laskuoperaatioiden määrittely reaaliluvuille varsin yksinkertaista: Jos $a = \{a_n\}$ ja $b = \{b_n\}$ ovat reaalilukuja (eli $\{a_n\}$ ja $\{b_n\}$ ovat rationaalilukujen Cauchyn jonoja), niin myös jonot $\{a_n + b_n\}$ ja $\{a_n b_n\}$ ovat Cauchyn jonoja (Lause I.10.4) ja edustavat sellaisenaan lukuja $a + b$ ja ab. Vastaavasti jos $b \neq 0$ (eli $b_n \not\to 0$), niin jostakin indeksistä $n = m$ alkaen on

$|b_n| \geq \delta > 0$ (Lause I.10.9), jolloin $\{a_n/b_n\}$ on ko. indeksistä eteenpäin Cauchy (Lause I.10.4) ja edustaa lukua a/b.

Myös järjestysrelaation määrittely käy Cauchyn jonojen avulla helposti, sillä jos $a = \{a_n\}$, niin vaihtoehdot $a = 0$, $a > 0$ ja $a < 0$ määräytyvät asettamalla $x = 0$ Lauseessa I.10.9: $a = 0$ vaihtoehdossa 1, $a < 0$ vaihtoehdossa 2, ja $a > 0$ vaihtoehdossa 3. Yleisemmin lukujen a ja b suuruusjärjestys ratkaistaan vertaamalla lukua $a - b = \{a_n - b_n\}$ lukuun 0 mainitulla tavalla.

Jos Määritelmä I.10.11 tulkitaan laskennallisesti, niin yksittäistä Cauchyn jonoa voi pitää algoritmina, joka tuottaa reaaliluvulle jonon rationaalisia approksimaatioita. Algoritmi on yleisesti sitä parempi, mitä nopeammin jono suppenee (tai 'suippenee') suhteessa tarvittavien (rationaalisten) laskuoperaatioiden määrään. Cauchyn jonoille voidaan siis asettaa myös laatukriteereitä!

Esimerkki 6 Eräs rationaalilukujen Cauchyn jonojen ekvivalenssiluokka on $\sqrt{2}$. Tähän luokkaan kuuluva ääretön desimaaliluku, eli lukujono $\{1, 1.4, 1.414, \ldots\}$ vastaa algoritmina edellisen luvun Esimerkin 3 kymmenjakoalgoritmia. Toinen samaan ekvivalenssiluokkaan kuuluva lukujono on Luvun I.8 Esimerkissä 4 tarkasteltu. Tämä on algoritmina kymmenjakoa selvästi tehokkaampi. □

*Bolzanon–Weierstrassin lause

Asiayhteyden vuoksi esitettäköön vielä osajonon käsitteeseen liittyvä lause, joka kuuluu modernin matemaattisen analyysin peruskiviin.

*Lause I.10.12 (**Bolzano-Weierstrass**[†]) Jokaisella rajoitetulla reaalilukujonolla on suppeneva osajono.

Todistus Todistus on perusidealtaan puolituskonstruktio, jossa looginen päättely nojaa kaksipaikkaiseen predikaattiin

$$Q(x, y): \quad x \leq a_n \leq y \quad \text{äärettömän monella indeksin } n \text{ arvolla.}$$

Koska $\{a_n\}$ on rajoitettu jono, on löydettävissä luvut a ja $b = a + L$ siten, että $Q(a, b)$ on tosi. Valitaan jokin indeksi n_1 siten, että $a \leq a_n \leq b$ kun $n = n_1$, ja tutkitaan seuraavaksi, onko $Q(a, (a+b)/2)$ tosi. Jos on, asetetaan b:n uudeksi arvoksi $(a+b)/2$. Jos ei, on proposition $Q((a+b)/2, b)$ oltava tosi (koska $Q(a, b)$ oli

[†]Bolzanon-Weierstrassin lauseella on monia muotoja. Perusidean esitti ensimmäisenä tšekkiläinen matemaatikko, filosofi ja pappi **Bernard Bolzano** (1781-1848) v. 1817, mutta tulos jäi vähälle huomiolle. Saksalainen **Karl Weierstrass** (1815-1897), jota on pidetty 'modernin matemaattisen analyysin isänä', todisti lauseen myöhemmin tuntematta Bolzanon työtä.

tosi). Tässä tapauksessa asetetaan a:n uudeksi arvoksi $(a+b)/2$. Toisen luvuista a, b tultua näin uudelleen määritellyksi tiedetään, että $Q(a,b)$ on jälleen tosi. Etsitään nyt indeksi $n_2 > n_1$ siten, että $a \leq a_n \leq b$ kun $n = n_2$ (mahdollista, koska $Q(a,b)$ oli tosi). Etenemällä samalla tavoin saadaan konstruoiduksi alkuperäisen jonon osajono $\{\, b_k = a_{n_k},\ k = 1, 2, \dots \}$, jolla on konstruktion perusteella ominaisuus $|b_k - b_l| \ \leq\ 2^{-N}L$ kun $k, l > N$. Näin ollen jos jokaisella $\varepsilon > 0$ valitaan $N \in \mathbb{N}$ siten, että $2^{-N}L < \varepsilon$, niin $|b_k - b_l| < \varepsilon$ kun $k, l > N$. Määritelmän I.10.1 mukaan $\{b_k\}$ on Cauchyn jono ja näin ollen suppee (Lause I.10.10). $\square$

*Epäkonstruktiivinen päättely

Lauseen I.10.12 todistukseen kätkeytyy filosofinen ongelma, joka esiintyy matemaattisessa ajattelussa yleisemminkin: Jos predikaatissa $Q(x,y)$ ajatellaan x ja y kiinnitetyiksi (sidotuiksi), niin proposition $P = Q(x,y)$ totuusarvon selville saamiseksi on käytävä kirjaimellisesti läpi jokainen jonon $\{a_n\}$ termi, ts. on suoritettava äärettömän monta lukujen vartailua. Tämä ei tietenkään ole käytännössä mahdollista. Matemaattisessa analyysissä katsotaan kuitenkin myös tällainen nk. *epäkonstruktiivinen* päättely luvalliseksi. (Viime kädessä 'luvan' antavat joukkoopin perusaksioomat.) Epäkonstruktiiviselle todistukselle on tunnusomaista, että todistusta ei voi seurata 'laskemalla', ellei käytettävissä ole jotakin lisätietoa, ts. ellei tehdä lisäoletuksia. Esimerkiksi aiemmin esitetty Lauseen I.8.3 todistus on myös epäkonstruktiivinen, sillä tämäkin sisältää kuvitelman kaikkien jonon termien vertaamisesta kiinteään lukuun. Epäkonstruktiivisia ovat itse asiassa myös Lauseiden I.10.8 ja I.10.9 todistukset, sillä 'olemassa olevaksi' väitetty luku ε konstruoidaan todistuksissa vain loogisena välttämättömyytenä. Mitään laskennallisesti toimivaa ideaa luvun määrittämiseksi ei anneta — eikä tehtyjen yleisten oletusten puitteissa olisi mahdollistakaan antaa.

Kaikissa mainituissa todistuksissa filosofiset ongelmat poistuvat (tai ainakin vähenevät), jos oletetaan tarkasteltava lukujono sellaiseksi, että todistuksen sisältämät loogiset askeleet voidaan toteuttaa (yksi kerrallaan) äärellisellä määrällä (viime kädessä) rationaalisia laskuoperaatioita tai vertailuja. Tällaista lukujonoa voi kutsua vaikkapa *ennustettava*ksi. Esimerkiksi Lukujen I.4 — I.9 esimerkkijonot, tai yleisemmin kaikki 'ennustettavasti määritellyt' yksittäiset lukujonot, ovat tällaisia. Lisäoletuksen ollessa voimassa voidaan todistuksissa esitettyjä ajatuskonstruktioita seurata myös laskennallisesti. Lauseen I.10.8 tapauksessa tämä tapahtuu esim. hakemalla käypää ε:n arvoa jonosta $\{\, \varepsilon_n = 10^{-n},\ n = 1, 2, \dots \}$. Lisäoletuksen perusteella voidaan millä tahansa kiinteällä $n \in \mathbb{N}$ ratkaista äärellisellä laskutyöllä, onko $\varepsilon = \varepsilon_n$ käypä valinta. Ellei ole, asetetaan $n \leftarrow n+1$, ja jatketaan laskemista. Ennustettavuusoletuksen perusteella laskun tiedetään päättyvän johonkin indeksiin, vaikkei etukäteen tiedetä, mihin.

HARJOITUSTEHTÄVIÄ

1. a) Todista Lause I.10.3.
   b) Todista Lauseen I.10.4 väittämät koskien lukujonoja $\{a_n + b_n\}$ ja $\{\lambda a_n\}$.
   c) Todista Lauseen I.10.4 väittämä koskien lukujonoa $\{a_n b_n\}$.
   d) Todista: Jos $\{a_n\}$ on Cauchy ja $|a_n| \geq \delta > 0\ \forall n$, niin $\{a_n^{-1}\}$ on Cauchy.
   e) Todista Lauseen I.10.4 väittämä koskien lukujonoa $\{a_n/b_n\}$.

2. Olkoon $\{a_k,\ k = 0, 1, \ldots\}$ lukujono ja $b_n = \sum_{k=0}^{n} 2^{-k} a_k,\ n = 0, 1, \ldots$ Näytä, että jos $\{a_k\}$ on rajoitettu, niin $\{b_n\}$ on Cauchy.

3. Olkoon $\{a_n\}$ rationaalilukujono. Näytä todeksi tai epätodeksi:
   a) $a_n^3 \to 4 \ \Rightarrow\ \{a_n\}$ on Cauchy, b) $a_n^4 \to 3 \ \Rightarrow\ \{a_n\}$ on Cauchy.

4. Jos $\{b_k\}$ on jonon $\{a_n,\ n = 1, 2, \ldots\}$ suppeneva osajono, niin mitkä ovat mahdolliset osajonon raja-arvot seuraavissa tapauksissa?
$$\text{a)}\ \ a_n = \frac{(n+1)^n}{(-n)^n} \qquad \text{b)}\ \ a_n = \frac{n^3 + (-1)^n (n-1)^3}{(n+2)^3 + (-1)^n n^3}$$

5. a) Olkoon luvun 1/7 n:s desimaali $= a_n$ ja luvun 1/17 n:s desimaali $= b_n$. Määritä lukujonon $\{a_n + b_n\}$ suppenevien osajonojen mahdolliset raja-arvot. b) Olkoon d_n = tunnetun jaksottoman desimaaliluvun π = 3.14159628.. n:s desimaali. Esitä kymmenen lukujonoa, joista vähintään kaksi on jonon $\{d_n\}$ suppenevia osajonoja. Perustele!

6. a) Näytä Lause I.10.7 päteväksi myös, kun $a = \infty$, b) Näytä, että lukujono ei ole rajoitettu täsmälleen kun jonolla on osajono, joka on aidosti monotoninen ja joko rajatta kasvava tai rajatta vähenevä.

7. (*) Todista: Jos $\{a_n\}$ ja $\{b_n\}$ ovat rajoitettuja lukujonoja, niin on olemassa aidosti kasvava indeksijono $\{n_k\}$ siten, että jonot $\{a_{n_k},\ k = 1, 2, \ldots\}$ ja $\{b_{n_k},\ k = 1, 2, \ldots\}$ ovat molemmat suppenevia.

8. (*) Käyttäen Cauchyn jonoihin perustuvaa reaaliluvun määritelmää näytä, että $(\mathbb{R}, +, \cdot, <)$ on järjestetty kunta.

9. (*) *Satunnaislukujono* on rationaalilukujono $\{a_n,\ n = 1, 2, \ldots\}$, jolle pätee $0 \leq a_n \leq 1\ \forall n$ ja jolla on ominaisuus: Jos $0 \leq x < y \leq 1$ ja N_n = niiden indeksien $k \in \mathbb{N}$ lukumäärä, joille pätee $k \leq n$ ja $x \leq a_k \leq y$ $(n \in \mathbb{N})$, niin $\lim_n N_n/n = y - x$. Bolzanon-Weierstrassin lauseen mukaan jonolla $\{a_n\}$ on suppeneva osajono. Näytä, että pätee vahvempi tulos: Jos $0 \leq x \leq 1$, niin löytyy osajono, jonka raja-arvo $= x$.

10. (*) Näytä, että jokaisella reaalilukujonolla on monotoninen osajono.

I.11 Reaalilukujen ominaisuuksia

Tässä luvussa esitellään eräitä reaalilukujoukkoihin liittyviä matemaattisen analyysin peruskäsitteitä. Samalla tuodaan esiin myös reaalilukujen yleisiä ominaisuuksia, vaihtoehtoisia määritelmiä ja luokittelutapoja.

Supremum ja infimum

MÄÄRITELMÄ I.11.1 Luku $a \in \mathbb{R}$ on joukon $A \subset \mathbb{R}$ *pienin yläraja* eli *supremum*, jos pätee

(1) $x \leq a \quad \forall x \in A.$

(2) $b < a \quad \Rightarrow \quad \exists x \in A \ (x > b).$

Merkitään $a = \sup A$.

Määritelmässä ehto (1) merkitsee, että a on joukon A *yläraja*, ja ehto (2) kertoo, että jos $b < a$, niin b ei ole yläraja. Ehdot merkitsevät siis yhdessä, että a on ylärajoista pienin mahdollinen. Jotta tällainen pienin yläraja voisi olla olemassa, on joukon A oltava ainakin *ylhäältä rajoitettu* (ehto (1) voimassa jollakin $a \in \mathbb{R}$) ja ei-tyhjä (ehto (2) voimassa jollakin $a \in \mathbb{R}$). Sikäli kuin $\sup A$ on olemassa, on määritelmästä varsin ilmeistä, että se on yksikäsitteinen. Jos $a = \sup A$ ja $a \in A$, niin sanotaan, että a on A:n *maksimi* (= suurin luku).

Jos Määritelmässä I.11.1 vaihdetaan epäyhtälöiden suunnat, tulee määritellyksi joukon A *suurin alaraja* eli *infimum*, jota merkitään $\inf A$. Infimum voi olla olemassa vain, jos A on ei-tyhjä ja *alhaalta rajoitettu*, eli jos jollakin $a \in \mathbb{R}$ pätee $x \geq a \quad \forall x \in A$. Jos $a = \inf A$ ja $a \in A$, niin a on A:n *minimi* (= pienin luku).

ESIMERKKI 1 Jos $A = \{x \in \mathbb{R} \mid x \leq 1\}$, niin Määritelmän I.11.1 mukaan $\sup A = 1$. Koska $1 \in A$, niin kyseessä on myös A:n maksimi. Koska A ei ole alhaalta rajoitettu, ei lukua $\inf A$ ole olemassa.

ESIMERKKI 2 Reaalilukujoukolla $A = \{\, x \in \mathbb{R} \mid x^2 < 2 \,\}$ ja rationaalilukujoukolla $B = \{\, x \in \mathbb{Q} \mid x^2 \leq 2 \,\}$ on kummallakin sama pienin yläraja ja suurin alaraja: $\sup A = \sup B = \sqrt{2}$ ja $\inf A = \inf B = -\sqrt{2}$. Maksimia tai minimiä ei kummallakaan joukolla ole. $\square$

Esimerkin 2 mukaan rajoitetun rationaalilukujoukon pienin yläraja tai suurin alaraja ei välttämättä ole rationaalinen. Vastaava ilmiö on entuudestaan tuttu lukujonoista: Rationaalilukujonon raja-arvo ei välttämättä ole rationaalinen.

Lukujonoista tiedetään myös, että kasvavalle ja rajoitetulle rationaalilukujonolle on aina konstruoitavissa raja-arvo reaalilukuna (eli äärettömänä desimaalilukuna, ks. Lauseen I.8.3 todistuskonstruktio). Vastaavalla tavalla on jokaiselle ei-tyhjälle ja ylhäältä rajoitetulle reaalilukujoukolle konstruoitavissa pienin yläraja reaalilukuna.

*Lause I.11.2 (**Supremum–lause**) Jokaisella ei-tyhjällä, ylhäältä rajoitetulla reaalilukujoukolla on supremum.

Todistus Tarkastellaan oletukset täyttävää joukkoa $A \subset \mathbb{R}$. Koska A on ei-tyhjä, on olemasssa $c_1 \in \mathbb{R}$ siten, että $x > c_1$ jollakin $x \in A$, ts. c_1 ei ole A:n yläraja. Koska A on ylhäältä rajoitettu, niin A:lla on yläraja $c_2 = c_1 + L \in \mathbb{R}$. Yleisyyttä rajoittamatta voidaan edelleen olettaa, että c_1 ja c_2 ovat rationaalilukuja, tai jopa kokonaislukuja. Konstruoidaan nyt kasvava rationaalilukujono $\{a_n\}$ ja vähenevä rationaalilukujono $\{b_n\}$ puolituskonstruktiolla seuraavasti:

1. Asetetaan $a_0 = c_1$, $b_0 = c_2$ ja $n = 1$.
2. Asetetaan $c = (c_1 + c_2)/2$. Jos c on A:n yläraja, niin asetetaan $c_2 = c$, muussa tapauksessa asetetaan $c_1 = c$.
3. Asetetaan $a_n = c_1,\ b_n = c_2$.
4. Lisätään indeksin n arvoa yhdellä ($n \leftarrow n + 1$) ja siirrytään kohtaan 2.

Konstruktion mukaisesti jono $\{a_n\}$ on kasvava ja rajoitettu ja jono $\{b_n\}$ on vähenevä ja rajoitettu, ja lisäksi $b_n - a_n = 2^{-n}L$. Siis $\{a_n\}$ ja $\{b_n\}$ suppenevat kohti yhteistä raja arvoa: $a_n \to a$ ja $b_n \to a$ jollakin $a \in \mathbb{R}$. Konstruktion perusteella jokainen b_n on A:n yläraja ja mikään a_n ei ole A:n yläraja, joten ilmeinen kandidaatti A:n pienimmäksi ylärajaksi on yhteinen raja-arvo a. Ensinnäkin tämä on A:n yläraja, sillä jos $x \in A$, niin $x \leq b_n\ \forall n$ (koska jokainen b_n oli yläraja), jolloin seuraa $x \leq \lim_n b_n = a$ (Propositio I.7.4 (V1)). Siis pätee $x \in A \ \Rightarrow\ x \leq a$, eli a on A:n yläraja.

Vielä on näytettävä, että a on ylärajoista pienin, joten olkoon $b < a$. Tällöin koska $a_n \to a$ ja $a > b$, niin jollakin n on $a_n > b$. Silloin b ei voi olla A:n yläraja, koska edes a_n ei ole. Siis mikään lukua a pienempi luku ei ole A:n yläraja, eli $a = \sup A$. $\square$

Jos A on ei-tyhjä, alhaalta rajoitettu reaalilukujoukko, niin on ilmeistä, että $B = \{-x \mid x \in A\}$ on ylhäältä rajoitettu ja että $\inf A = -\sup B$. Tästä ja Lauseesta I.11.2 seuraa välittömästi, että jokaisella ei-tyhjällä, alhaalta rajoitetulla reaalilukujoukolla on infimum.

Aksiomaattiset reaaliluvut

Edellisessä luvussa esitettyjä, desimaalilukuihin tai Cauchyn jonoihin perustuvia reaaliluvun määritelmiä sanotaan *konstruktiivisiksi.* Konstruktiivisille määritelmille on tunnusomaista, että mitään 'ulkopuolista maailmaa' suhteessa rationaalilukuihin ei oleteta. Päätetään vain kutsua tietyn tyyppisiä rationaalilukujen jonoja reaaliluvuiksi, ja määritellään näiden lukujen väliset laskuoperaatiot sekä samastus- ja järjestysrelaatiot palauttamalla kaikki operaatiot viime kädessä rationaalisiksi. Konstruktiivisen määrittelyn hyvä puoli on, että se muistuttaa reaalilukujen todellista, laskennallista 'määrittelyä' laskinten ja tietokoneiden avulla. Toisaalta kun reaaliluvuilla halutaan laskea symbolisesti, on reaaliluvut usein helpompaa mieltää suoraan abstrakteina lukuina kuin konkreettisempien lukujen jonoina (vrt. Esimerkki 5 Luvussa I.9). Abstrakti ajattelu viedään pisimmälle reaalilukujen *aksiomaattisessa* määrittelyssä, jossa yksinkertaisesti sovitaan, että muitakin kuin rationaalilukuja on olemassa. Menettely on tällöin samankaltainen kuin Luvun I.2 Esimerkissä 3, jossa luvun $\sqrt{2}$ 'olemassaoloon' suhtauduttiin sopimuskysymyksenä. Kuten konstruktiivisia määritelmiä, myös aksiomaattisia lähestymistapoja reaalilukuihin on monia samanarvoisia. Määritelmien yhteisenä lähtökohtana on oletus, että reaaliluvut muodostavat järjestetyn kunnan, joka on rationaalilukujen kunnan laajennus. Nämä oletukset jättävät mahdollisuuden, että $\mathbb{R} = \mathbb{Q}$, joten tarvitaan vielä aksiooma, joka erottaa joukot toisistaan. Varsin usein käytetty menettely on tehdä joko Lauseesta I.11.2 tai Lauseesta I.10.5 aksiooma, eli asettaa jompi kumpi seuraavista.

(A) (**Supremum-aksiooma**) Jokaisella ylhäältä rajoitetulla reaalilukujoukolla on pienin yläraja.[†]

(B) (**Täydellisyysaksiooma**) Jokainen reaalilukujen Cauchyn jono suppenee kohti reaalilukua.

Reaalilukujen em. konstruktiivisessa määrittelyssä nämä ovat siis tosia väittämiä. Aksioomina ne ovat myös keskenään vaihdannaiset: Jos oletetaan (A), niin (B) on tosi väittämä, ja päinvastoin.

Kuten konstruktiivisessa, myös reaalilukujen aksiomaattisessa määrittelyssä lähtökohtana ovat siis aluksi rationaaliluvut. Perustuen aksioomaan (A) tai (B), rationaalilukujen joukko *täydennetään* liittämällä joukkoon joko ylhäältä rajoitettujen rationaalilukujoukkojen pienimmät ylärajat tai rationaalisten Cauchyn

[†]Supremum-aksioomaan tai supremum-lauseeseen liittyen reaaliluvut voidaan määritellä myös ylhäältä rajoitettuina rationaalilukujen joukkoina. Tämän nk. *Dedekindin leikkauksiin* nojaavan konstruktiivisen määritelmän esitti saksalainen matemaatikko **Richard Dedekind** (1831-1916). Dedekindin määritelmä on ollut aikanaan suosittu opetuksessakin.

jonojen raja-arvot, jotka siis oletetun aksiooman mukaan ovat olemassa reaalilukuina. Koska täydentäminen on taka-ajatuksena myös konstruktiivisessa määrittelyssä, niin tätä on tapana kutsua *täydennyskonstruktioksi.* Rationaaliluvut voi siis täydentää reaaliluvuiksi joko konstruktiivisesti 'rakentamalla' tai aksiomaattisesti 'olettamalla'.

$\mathbb{R}$ on ylinumeroituva

Jos $a, b \in \mathbb{R}$ ja $a < b$, niin on helposti nähtävissä, että lukujen a, b väliin voidaan aina sijoittaa rationaaliluku, ts. löytyy $x \in \mathbb{Q}$ jolle pätee $a < x < b$. Samantien lukuja löytyy ääretön (numeroituva) määrä, eikä konstruktiossa itse asiassa edes tarvita koko joukkoa $\mathbb{Q}$, vaan esim. äärelliset desimaaliluvut (tai äärelliset binaariluvut) riittävät (Harj.teht. 4). Tuloksen perusteella sanotaan, että $\mathbb{Q}$ (tai mainitun tyyppinen $\mathbb{Q}$:n osajoukko) on $\mathbb{R}$:ssä *tiheä* (engl. dense). Tästä tuloksesta huolimatta on $\mathbb{R}$ joukkoa $\mathbb{Q}$ mahtavampi, sillä pätee (vrt. Luku I.4)

Lause I.11.3 $\mathbb{R}$ on ylinumeroituva joukko.

Todistus Olkoon $\{x^{(n)}\}$ jono reaalilukuja, joille pätee $0 < x^{(n)} < 1 \ \ \forall n$. Kirjoitetaan jonon termit äärettöminä desimaalilukuina:

$$\begin{aligned} x^{(1)} &= 0.d_1^{(1)} d_2^{(1)} d_3^{(1)} \ldots \\ x^{(2)} &= 0.d_1^{(2)} d_2^{(2)} d_3^{(2)} \ldots \\ x^{(3)} &= 0.d_1^{(3)} d_2^{(3)} d_3^{(3)} \ldots \\ &\vdots \end{aligned}$$

Määritellään sitten ääretön desimaaliluku $y = 0.d_1 d_2 \ldots$ valitsemalla

$$d_n \in \{1, \ldots 8\}, \quad d_n \neq d_n^{(n)}, \quad n = 1, 2, \ldots$$

Tällöin $0 < y < 1$ ja y ei ole äärellinen desimaaliluku, joten y ei samastu mihinkään muuhun, merkkijonona vähänkään erilaiseen desimaalilukuun. Konstruktiosta seuraa tällöin, että $y \neq x^{(n)} \ \forall n$. Päätellään, että mikään lukujoukon $A = \{\, x \in \mathbb{R} \mid 0 < x < 1 \,\}$ numerointiyritys ei kata kaikkia joukon A lukuja, joten edes A ei ole numeroituva. $\square$

Olkoon $a, b \in \mathbb{R}$, $a < b$ ja $B = \{\, x \in \mathbb{R} \mid a < x < b \,\}$. Koska

$$B = \{\, x = a + (b - a)t \mid t \in \mathbb{R},\ 0 < t < 1 \,\},$$

niin Lauseen I.11.3 todistuksesta voidaan päätellä, että B on ylinumeroituva. Toisin sanoen: kahden eri suuren reaaliluvun väliin mahtuu aina ylinumeroituva määrä reaalilukuja.

Avoin ja suljettu väli

Joukko $A \subset \mathbb{R}$ on *väli* (engl. interval), jos A:ssa on enemmän kuin yksi alkio ja pätee $x, y \in A \ \& \ x < z < y \ \Rightarrow \ z \in A$. Väli voi olla joko *äärellinen*, eli joukkona ylhäältä ja alhaalta rajoitettu, tai *ääretön* (ylhäältä ja/tai alhaalta rajoittamaton). Äärellisen välin päätyypit ovat *avoin* ja *suljettu* väli, joiden merkinnät ja määritelmät ovat:

$$\begin{aligned} \text{Avoin väli:} \quad (a,b) &= \{\, x \in \mathbb{R} \mid a < x < b \,\}. \\ \text{Suljettu väli:} \quad [\,a,b\,] &= \{\, x \in \mathbb{R} \mid a \le x \le b \,\}. \end{aligned}$$

Tässä $a = \inf A$ ja ja $b = \sup A$ ovat välin *päätepisteet* ($a, b \in \mathbb{R}, \ a < b$). Avoimen välin vaihtoehtoinen merkintätapa on $]a, b[$. Väli voi myös olla muotoa $(a, b]$ tai $[a, b)$ (vaihtoehtoiset merkinnät $]a, b]$ ja $[a, b[\,$), jolloin sanotaan, että väli on *puoliavoin* — määritelmät ovat ilmeiset. Muita kuin välin päätepisteitä sanotaan välin *sisäpisteiksi*. Avoimen välin kaikki pisteet ovat siis sisäpisteitä. Suljettu väli on aina äärellinen. Muun tyyppiset välit voivat olla myös äärettömiä, jolloin päätepisteen puuttuminen merkitään symbolilla $\pm\infty$:

$$\begin{aligned} (a, \infty) &= \{\, x \in \mathbb{R} \mid x > a \,\}, \\ (-\infty, b] &= \{\, x \in \mathbb{R} \mid x \le b \,\}, \\ (-\infty, \infty) &= \mathbb{R}. \end{aligned}$$

Mainittakoon tässä yhteydessä myös yleisesti käytetyt merkinnät

$$\begin{aligned} (0, \infty) &= \mathbb{R}_+ \quad \text{(positiiviset reaaliluvut, 'R plus')}\,, \\ (-\infty, 0) &= \mathbb{R}_- \quad \text{(negatiiviset reaaliluvut, 'R miinus')}\,. \end{aligned}$$

<u>Suljettujen</u> välien $[a_n, b_n]$, $n = 1, 2, \ldots$ muodostamaa <u>jonoa</u> sanotaan *sisäkkäiseksi* (engl. nested = 'pesitetty'), jos $[a_n, b_n] \supset [a_{n+1}, b_{n+1}] \quad \forall n$. Tällaiseen jonoon liittyen reaaliluvuilla on seuraava hauska ominaisuus:

(C) Jos $\{\, [a_n, b_n], \ n = 1, 2, \ldots \}$ on jono sisäkkäisiä suljettuja välejä ja pätee $b_n - a_n \to 0$, niin on olemassa yksikäsitteinen reaaliluku x siten, että $x \in [a_n, b_n] \ \ \forall n \in \mathbb{N}$.

Reaalilukujen konstruktiivisen määrittelyn perusteella (C) on tosi väittämä, sillä x määräytyy jonojen $\{a_n\}$ (kasvava ja rajoitettu) ja $\{b_n\}$ (vähenevä ja rajoitettu) yhteisenä raja-arvona. Reaalilukujen aksiomaattisessa määrittelyssä tämä tulos voidaan myös ottaa perusaksioomaksi (olettaen $a_n, b_n \in \mathbb{Q}$), jolloin edellä mainituista aksioomista (A), (B) tulee tosia väittämiä.

Jos $x \in \mathbb{R}$ ja $\delta \in \mathbb{R},\ \delta > 0$, niin x:n *δ-ympäristö* määritellään avoimena välinä

$$U_\delta(x) \ = \ (x - \delta, x + \delta).$$

Termit 'päätepiste' ja 'ympäristö' viittaavat ajatukseen, että reaaliluku on myös miellettävissä 'pisteeksi' jossakin 'paikassa'. Tämä ajatus ei ole peräisin algebrasta vaan *geometriasta*, toisesta matemaattisen ajattelun suuresta päähaarasta. Geometrian perusteita käsitellään lähemmin Luvussa II.

Algebralliset ja transkendenttiset luvut

Määritelmänsä perusteella reaaliluvut jakautuvat luonnostaan rationaalisiin ja ei-rationaalisiin eli *irrationaalisiin* lukuihin. Toinen, myös yleisesti käytetty luokittelutapa on jakaa reaaliluvut *algebrallisiin* ja ei-algebrallisiin eli *transkendenttisiin* lukuihin. Lukua $x \in \mathbb{R}$ sanotaan algebralliseksi, jos se toteuttaa yhtälön muotoa

$$\sum_{k=0}^{n} a_k x^k \ = \ 0,$$

missä $n \in \mathbb{N}$ ja $a_k \in \mathbb{Z},\ k = 0, 1, \ldots, n$. Luku on siis algebrallinen, jos se on jonkin kokonaislukukertoimisen polynomin nollakohta.

Esimerkki 3 Juuriluvut ja niiden erikoistapauksena rationaaliluvut toteuttavat yhtälön muotoa $px^m - q = 0$, missä $m \in \mathbb{N}$ ja $p, q \in \mathbb{Z}$, joten tällaiset luvut ovat algebrallisia. $\square$

Jos algebrallisten lukujen joukkoa merkitään symbolilla $\mathbb{A}$, niin esimerkin perusteella on siis

$$\mathbb{Q} \ \subset \ \mathbb{A} \ \subset \ \mathbb{R}.$$

Kuten $\mathbb{Q}$, myös $\mathbb{A}$ on osoitettavissa numeroituvaksi joukoksi, joten reaalilukujen 'enemmistö' on transkendenttisia.[†]

[†]Matematiikan lajia, joka tutkii lukujen, kuten reaalilukujen tai luonnollisten lukujen, ominaisuuksia mm. erilaisten luokittelujen näkökulmasta, sanotaan *lukuteoriaksi*. Lukuteorian tunnettuja tuloksia on esimerkiksi, että Neperin luku e on transkendenttinen (**Charles Hermite**, 1873), samoin luku π (**Ferdinand von Lindemann**, 1882). Molempien lukujen irrationaalisuus osoitettiin jo 1700-luvulla. Sen sijaan niinkään yksinkertaisia väittämiä kuin

P1: Luku $e + \pi$ ei ole rationaaliluku.

P2: Luku $e\pi$ ei ole rationaaliluku.

ei ole vielä tätä kirjoitettaessa (2015) pystytty osoittamaan tosiksi (saati epätosiksi). — Lukuteoria on tunnettu monista helposti muotoiltavissa olevista, mutta usein vaikeasti ratkaistavista matemaattisista pähkinöistään.

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien joukkojen supremum, infimum, maksimi ja minimi, sikäli kuin olemassa:
   a) $\{x \in \mathbb{R} \mid (x+1)(x-2)(x+4) > 0\}$ b) $\{x \in \mathbb{R} \mid |x| + |x+2| < 5\}$
   c) $\{x \in \mathbb{R} \mid |x||x+2| < 5\}$ d) $\{n/(n+1) \mid n \in \mathbb{N}\}$
   e) $\{(2-3n)/(5+n) \mid n \in \mathbb{N}\}$ f) $\{n^3 2^{-n} \mid n \in \mathbb{N}\}$

2. a) Millaisille ei-tyhjille joukoille $A \subset \mathbb{R}$ pätee $\inf A = \sup A$?
   c) Olkoon $A \subset \mathbb{R}$ ylhäältä rajoitettu joukko ja $a = \sup A$. Näytä, että jokaisella $\varepsilon > 0$ on olemassa $x \in A$ siten, että $x > a - \varepsilon$.
   c) Näytä, että jos $A, B \subset \mathbb{R}$ ovat molemmat ylhäältä (alhaalta) rajoitettuja ja $A \subset B$, niin $\sup A \leq \sup B$ ($\inf A \geq \inf B$).

3. Olkoon a) $A = \{x \in \mathbb{R} \mid x^3 < 200\}$, b) $A = \{x \in \mathbb{R} \mid x^3 < -400\}$. Seuraa Lauseen I.11.2 todistuskonstruktiota algoritmina indeksiin $n = 3$ asti valitsemalla c_0:n ja c_1:n alkuarvot siten, että algoritmin tuottamille lukujonoille pätee: a) $\{a_n\} = a$, b) $\{b_n\} = a$, missä $a = \sup A$ äärettömänä binaarilukuna.

4. Olkoon $a, b \in \mathbb{R}$ ja $a < b$. Näytä, että on olemassa äärettömän monta eri suurta äärellistä desimaalilukua x, joille pätee $a < x < b$.

5. Määritä kaikki välit, joilla on ominaisuus: Välin päätepisteet — sikäli kuin niitä on — sisältyvät joukkoon $\{-1, 2\}$.

6. Anna seuraaville joukoille yksinkertaisempi määritelmä väleinä, eli muodossa (a, b), $[a, b]$, $(a, b]$ tai $[a, b)$.

$$\text{a)} \quad \bigcup_{n=1}^{\infty} \left(\frac{1}{3n}, \frac{n}{n+1} \right) \qquad \text{b)} \quad \bigcap_{n=1}^{\infty} \left(\frac{1}{3n}, \frac{n+1}{n} \right)$$

$$\text{c)} \quad \bigcap_{n=1}^{\infty} \left[\frac{1}{3n}, \frac{n}{n+1} \right] \qquad \text{d)} \quad \bigcup_{n=1}^{\infty} \left[\frac{1}{3n}, \frac{n+1}{n} \right]$$

7. Näytä: a) Jos x on irrationaalinen, niin samoin on $y = (3+x)/(x-2)$. b) Jos $x \neq 0$ on algebrallinen luku, niin samoin on x^{-1}. c) Jos $x > 0$ on algebrallinen luku, niin samoin on $\sqrt[m]{x}$ $\forall m \in \mathbb{N}$, $m \geq 2$. d) Jokainen rationaalikertoimisen polynomin nollakohta on algebrallinen luku.

8. (*) Olkoon $\{\, [a_n, b_n],\ n = 1, 2, \dots \}$ jono suljettuja välejä. Näytä, että $A = \bigcap_{n=1}^{\infty} [a_n, b_n]$ on joko tyhjä joukko, sisältää täsmälleen yhden alkion, tai on suljettu väli.

I.12 Klassinen sarjaoppi. Potenssisarja

Kerrattakoon aiemmista luvuista, että merkinnällä

$$\sum_{k=1}^{\infty} a_k \;=\; a_1 + a_2 + a_3 + \ldots$$

tarkoitetaan joko sarjaa, eli lukujonoa $\{a_1,\, a_1 + a_2,\, \ldots\} = \{s_1, s_2,\, \ldots\}$, tai sikäli kuin $\{s_n\}$ suppenee (kohti reaalilukua), myös *sarjan summaa* eli raja-arvoa $s = \lim_n s_n$. Kerrattakoon myös, että lukuja a_k sanotaan sarjan *termeiksi* ja lukujonon $\{s_n\}$ termejä sarjan *osasummiksi*. Sarjoilla on paljon käyttöä sovelluksissa ja myös matemaattisessa ajattelussa. Tässä luvussa tarkastellaan sarjojen suppenevuusteoriaa aiempaa yleisemmältä kannalta sekä tutkitaan eräitä sarjojen tavallisia erikoistyyppejä. Sarjan termit a_k oletetaan jatkossa reaaliluvuiksi.

Positiiviterminen sarja

Suppenemisteorian kannalta 'ystävällismielisin', sarja on *positiiviterminen* sarja $\sum_k a_k$, jonka termit ovat positiiviset, tai yleisemmin ei-negatiiviset: $a_k \geq 0\ \forall k$. Esimerkkejä on tarkasteltu jo Luvussa I.8. Positiivitermisen sarjan osasummien jono on kasvava, joten sarja suppenee täsmälleen kun ko. lukujono on rajoitettu (Lause I.8.3). Suppenemistarkastelua voidaan usein helpottaa edelleen vertaamalla sarjan termejä yksinkertaisempaan sarjaan, joka jo tiedetään suppenevaksi tai hajaantuvaksi. Seuraavat vertailuperiaatteet ovat hyödyllisiä.

Lause I.12.1 Jos $0 \leq a_k \leq b_k\ \forall k$, niin pätee

$$\textbf{Majoranttiperiaate:}\quad \sum_k b_k \text{ suppenee} \quad\Rightarrow\quad \sum_k a_k \text{ suppenee}$$

$$\textbf{Minoranttiperiaate:}\quad \sum_k a_k \text{ hajaantuu} \quad\Rightarrow\quad \sum_k b_k \text{ hajaantuu}$$

<u>Todistus</u> Jos sarjojen $\sum_k a_k$ ja $\sum_k b_k$ osasummien jonot ovat vastaavasti $\{s_n\}$ ja $\{t_n\}$, niin oletuksen perusteella on $s_n \leq t_n\ \forall n$. Majoranttiperiaate seuraa tällöin päättelyllä

$$\{t_n\} \text{ suppenee} \;\Rightarrow\; \{t_n\} \text{ rajoitettu} \;\Rightarrow\; \{s_n\} \text{ rajoitettu} \;\Rightarrow\; \{s_n\} \text{ suppenee.}$$

Minoranttiperiaate seuraa vastaavalla päättelyllä – tai vain toteamalla vertailuperiaatteet loogisesti yhtäpitäviksi (vrt. Luku I.3). $\square$

Esimerkki 1 Suppeneeko vai hajaantuuko sarja

$$\text{a)} \quad \sum_{k=1}^{\infty} \frac{(k+3)^2}{k^4}, \qquad \text{b)} \quad \sum_{k=0}^{\infty} \frac{2k+50}{k^2+1} \ ?$$

Ratkaisu a) Koska

$$0 < a_k = \frac{(k+3)^2}{k^4} = \frac{1}{k^2} + \frac{6}{k^3} + \frac{9}{k^4} \leq \frac{16}{k^2}, \quad k \geq 1,$$

ja koska vertailusarja $\sum_{k=1}^{\infty} 16k^{-2} = 16 \sum_{k=1}^{\infty} k^{-2}$ tiedetään suppenevaksi (Luku I.8, Esimerkki 2), niin majoranttiperiaatteen nojalla sarja suppenee.

b) Kun kirjoitetaan

$$a_k \;=\; \frac{2k+50}{k^2+1} \;=\; \frac{1}{k} \cdot \frac{2+50\,k^{-1}}{1+k^{-2}} \;=\; \frac{1}{k} \cdot c_k\,,$$

niin nähdään, että $c_k \to 2$, joten jostakin indeksistä $k = m$ alkaen (itse asiassa kun $k \geq 50$) on $1 \leq c_k \leq 3$. Tällöin on

$$\frac{1}{k} \;\leq\; a_k \;\leq\; \frac{3}{k}, \quad k = m,\, m+1,\, \ldots\,,$$

joten päätellään majorantti- ja minoranttiperiaatteiden avulla, että tarkasteltava sarja suppenee/hajaantuu täsmälleen kun sarja $\sum_{k=m}^{\infty} 1/k$ suppenee/hajaantuu. Tämä vertailusarja osoittautuu hajaantuvaksi (Lause I.12.2 jäljempänä), joten tarkasteltava sarja siis hajaantuu myös. $\square$

Harmoninen, aliharmoninen ja yliharmoninen sarja

Positiivitermistä sarjaa

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \;=\; 1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \ldots\,,$$

missä α on rationaaliluku[†], sanotaan *harmoniseksi*, jos $\alpha = 1$, *aliharmoniseksi*, jos $\alpha < 1$, ja *yliharmoniseksi*, jos $\alpha > 1$. Majorantti- ja minoranttiperiaatteita käytettäessä tämä sarjatyyppi on usein hyvä vertailukohta.

Lause I.12.2 Harmoninen ja aliharmoninen sarja hajaantuvat. Yliharmoninen sarja suppenee.

[†]Toistaiseksi x^α $(x \in \mathbb{R},\ x > 0)$ on määritelty vain, kun $\alpha \in \mathbb{Q}$.

Todistus Tarkastellaan ensin harmonista sarjaa

$$\sum_{k=1}^{\infty} \frac{1}{k} \;=\; 1 + \frac{1}{2} + \frac{1}{3} + \ldots .$$

Arvioidaan osasummia s_4, s_8, s_{16} jne. seuraavasti:

$$\begin{aligned} s_4 &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) \;>\; \frac{3}{2} + 2 \cdot \frac{1}{4} \;=\; 2, \\ s_8 &= s_4 + \left(\frac{1}{5} + \ldots + \frac{1}{8}\right) \;>\; 2 + 4 \cdot \frac{1}{8} \;=\; \frac{5}{2}, \\ &\vdots \end{aligned}$$

Päätellään, että jos $n = 2^m,\ m \in \mathbb{N},\ m \geq 2$, niin $s_n > 1 + m/2$, joten harmonisen sarjan osasummien osajono $\{s_4, s_8, s_{16}, \ldots\}$ on hajaantuva. Siis sarja hajaantuu: $s_n \to \infty$. Tällöin myös aliharmoninen sarja hajaantuu minoranttiperiaatteen nojalla, sillä lukujen $k^\alpha,\ \alpha \in \mathbb{Q}$ määritelmästä (ks. Luku I.2) seuraa helposti, että $k > k^\alpha \;\Leftrightarrow\; k^{-\alpha} > k^{-1}$, kun $\alpha < 1$ ja $k \geq 2$.

Yliharmonisen sarjan tapauksessa merkitään $\alpha = 1 + \delta,\ \delta > 0$ ja arvioidaan osasummia s_3, s_7, jne. seuraavasti:

$$\begin{aligned} s_3 &= 1 + \left(\frac{1}{2^\alpha} + \frac{1}{3^\alpha}\right) \;<\; 1 + 2 \cdot \frac{1}{2^\alpha} \;=\; 1 + \frac{1}{2^\delta}, \\ s_7 &= s_3 + \left(\frac{1}{4^\alpha} + \ldots \frac{1}{7^\alpha}\right) \;<\; 1 + \frac{1}{2^\delta} + 4 \cdot \frac{1}{4^\alpha} \;=\; 1 + \frac{1}{2^\delta} + \left(\frac{1}{2^\delta}\right)^2, \\ &\vdots \end{aligned}$$

(Tässä on käytetty arviota $(k/j)^\alpha > 1 \;\Leftrightarrow\; k^{-\alpha} < j^{-\alpha}$, kun $k/j > 1$ ja $\alpha > 0$.) Päätellään, että jos $n = 2^m - 1,\ m \in \mathbb{N},\ m \geq 2$, niin

$$s_n \;<\; 1 + \frac{1}{2^\delta} + \ldots + \left(\frac{1}{2^\delta}\right)^{m-1} \;<\; \sum_{k=0}^{\infty} \left(\frac{1}{2^\delta}\right)^k \;=\; \frac{2^\delta}{2^\delta - 1}\,.$$

Siis sarjan osasummien jono on rajoitettu, joten sarja suppenee. □

Huomautettakoon, että tapauksessa $\alpha = 2$ Lauseen I.12.2 väittämä on jo aiemmin todistettu muilla keinoin (ks. Luku I.8).

Esimerkki 2 Yliharmoninen sarja $\sum_{k=1}^{\infty} k^{-5/4}$ suppenee, mutta suppeneminen on niin hidasta, että pelkästään osasummia laskemalla ei sarjan summaa saada selville tarkasti. Paremmilla algoritmeilla (ks. Harj.teht. VIII.9:11) tulos on

$$s \;=\; \sum_{k=1}^{\infty} \frac{1}{k^{5/4}} \;=\; 1 + \frac{1}{2^{5/4}} + \frac{1}{3^{5/4}} + \ldots \;=\; 4.5951118258\,..\quad \square$$

Esimerkki 3 Hajaantuvan aliharmonisen sarjan $\sum_{k=1}^{\infty} 1/\sqrt{k}$ osasummille pätee suurilla n:n arvoilla arvio (ks. Harj.teht. 16)

$$s_n = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \approx 2\sqrt{n}. \quad \square$$

Vuorotteleva sarja

Sarjaa sanotaan *vuorottelevaksi* eli *alternoivaksi* (engl. alternating), jos sen termien etumerkki vaihtuu aina indeksistä seuraavaan siirryttäessä. Vuorotteleva sarja on siis muotoa $\sum_k (-1)^k a_k$, missä jonon $\{a_k\}$ termit ovat samanmerkkiset.

Esimerkki 4 Sarja

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{\sqrt{k+1}} = \left\{ 1\,,\ 1 - \frac{1}{\sqrt{2}}\,,\ 1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}}\,, \ldots \right\}$$

on vuorotteleva. $\square$

Esimerkkisarjaan soveltuu seuraava yleisempi väittämä.

Lause I.12.3 Jos vuorottelevalle sarjalle $\sum_{k=0}^{\infty} (-1)^k a_k$ pätee (i) $a_k > 0\ \forall k$, (ii) jono $\{a_k\}$ on vähenevä, ts. $a_0 \geq a_1 \geq a_2 \geq \ldots$ ja (iii) $\lim_k a_k = 0$, niin sarja suppenee. Lisäksi approksimoitaessa sarjan summaa s osasummilla s_n on voimassa arvio

$$|s - s_n| \leq a_{n+1}\,.$$

<u>Todistus</u> Tarkastellaan sarjan osasummia $s_n = \sum_{k=0}^{n} (-1)^k a_k$ erikseen parillisilla ja parittomilla indeksin arvoilla. Koska

$$\begin{aligned} s_{2n} &= (a_0 - a_1) + \ldots + (a_{2n-2} - a_{2n-1}) + a_{2n} \\ &= a_0 - (a_1 - a_2) - \ldots - (a_{2n-1} - a_{2n}), \\ s_{2n+1} &= (a_0 - a_1) + \ldots + (a_{2n} - a_{2n+1}) \\ &= a_0 - (a_1 - a_2) - \ldots - (a_{2n-1} - a_{2n}) - a_{2n+1}. \end{aligned}$$

niin oletuksista (i)–(ii) seuraa, että $0 \leq s_n \leq a_0$ sekä parillisilla että parittomilla n:n arvoilla. Siis $\{s_n\}$ on rajoitettu jono. Samoin nähdään, että jono $\{s_{2n}\}$ on vähenevä ja $\{s_{2n+1}\}$ on kasvava, joten molemmat suppenevat. Koska edelleen $s_{2n} - s_{2n+1} = a_{2n+1} \to 0$ oletuksen (iii) mukaan, niin jonoilla $\{s_{2n}\}$ ja $\{s_{2n+1}\}$ on yhteinen raja-arvo. Tästä on helposti pääteltävissä, että koko jono $\{s_n\}$ suppenee.

Väittämän toisen osan todistamiseksi olkoon ensin n pariton. Tällöin

$$\begin{aligned} s - s_n &= a_{n+1} - (a_{n+2} - a_{n+3}) - \ \ldots \\ &= (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \ \ldots, \end{aligned}$$

joten oletusten perusteella $0 \le s - s_n \le a_{n+1}$. Jos n on parillinen, niin päätellään vastaavasti, että $-a_{n+1} \le s - s_n \le 0$. Siis molemmissa tapauksissa pätee väitetty arvio. $\square$

Esimerkki 4 (jatko) Esimerkkisarja on Lauseen I.12.3 perusteella suppeneva. Summaksi saadaan (numeerisin keinoin) $s = 0.6048986434..$ Tässä tapauksessa arvio $|s - s_n| \le a_{n+1}$ tarkentuu suurilla n:n arvoilla muotoon $|s - s_n| \approx \frac{1}{2}a_{n+1}$, vrt. taulukko alla. (Tarkemmin on osoitettavissa: $\lim_n |s - s_n|/a_{n+1} = \frac{1}{2}$.)

$$\begin{aligned} s - s_{100} &= -0.049628.. & a_{101} &= 0.099503.. \\ s - s_{101} &= +0.049386.. & a_{102} &= 0.099014.. \\ s - s_{1000} &= -0.015799.. & a_{1001} &= 0.031606.. \\ s - s_{1001} &= +0.015791.. & a_{1002} &= 0.031591.. \quad \square \end{aligned}$$

Cauchyn kriteeri sarjoille

Mikäli sarja ei ole positiiviterminen tai vuorotteleva, on suppenemisteoriassa yleensä turvattava Lauseeseen I.10.5. Tämän mukaan sarja $\sum_{k=1}^{\infty} a_k$ suppenee (kohti reaalilukua) täsmälleen, kun sen osasummien $s_n = \sum_{k=1}^{n} a_k$ muodostama jono on Cauchyn jono. Koska $s_m - s_n = \sum_{k=n+1}^{m} a_k$, kun $m > n$, niin Cauchyn jonon määritelmän perusteella pätee siis

Lause I.12.4 (**Cauchyn kriteeri sarjoille**) Sarja $\sum_{k=1}^{n} a_k$ suppenee täsmälleen, kun jokaisella $\varepsilon > 0$ on olemassa $N \in \mathbb{N}$ siten, että pätee

$$|\sum_{k=n+1}^{m} a_k| \ < \ \varepsilon, \quad \text{kun } m > n > N.$$

Kun kriteerissä valitaan $m = n+1$, niin päädytään seuraaviin ilmeisiin (keskenään loogisesti yhtäpitäviin) johtopäätöksiin:

Korollaari I.12.5 Jos sarja $\sum_k a_k$ suppenee, niin $a_k \to 0$.

Korollaari I.12.6 Jos $a_k \not\to 0$, niin sarja $\sum_k a_k$ hajaantuu.

Jos sarjan termien etumerkit vaihtelevat, niin sarjan suppenemiskysymyksen voi yrittää ratkaista yksinkertaisesti vertaamalla positiivitermiseen sarjaan, josta etumerkkien vaihtelu on poistettu. Onnistuessaan tämä yritys perustuu seuraavaan lauseeseen. Ensin määritelmä:

MÄÄRITELMÄ I.12.7 Sarja $\sum_k a_k$ suppenee *itseisesti* (engl. absolutely), jos sarja $\sum_k |a_k|$ suppenee.

LAUSE I.12.8 Jos sarja suppenee itseisesti, niin se suppenee.

Todistus Kolmioepäyhtälön nojalla on

$$|\sum_{k=n+1}^{m} a_k| \leq \sum_{k=n+1}^{m} |a_k| \quad (m > n).$$

Koska sarja $\sum_k |a_k|$ suppenee, on tässä Lauseen I.12.4 mukaan jokaisella $\varepsilon > 0$ löydettävissä $N \in \mathbb{N}$ siten, että oikea puoli $< \varepsilon$ kun $n > N$, jolloin saman kriteerin perusteella myös sarja $\sum_k a_k$ suppenee. □

ESIMERKKI 5 Jos $\{b_k\}$ on lukujono, jolle pätee $b_k \in [-1, 1]$ $\forall k$, niin sarja $\sum_{k=1}^{\infty} b_k/k^2$ suppenee itseisesti (majoranttiperiaate ja Lause I.12.2). Sarja siis suppenee, valittiinpa luvut b_k väliltä $[-1, 1]$ miten tahansa. □

ESIMERKKI 6 Sarja $\sum_{k=1}^{\infty}(-1)^k/\sqrt{k}$ suppenee (Lause I.12.3) mutta ei itseisesti (Lause I.12.2). □

Potenssisarja

Sarjaa muotoa

$$\sum_{k=0}^{\infty} a_k x^k = a_0 + a_1\, x + a_2\, x^2 + \ldots$$

sanotaan *potenssisarja*ksi (engl. power series). Luvut a_k ovat nimeltään *potenssisarjan kertoimet*, ja symbolia x, joka myös edustaa reaalilukua, sanotaan sarjan *muuttuja*ksi. Termi 'muuttuja' kertoo, että x:n lukuarvo voi vaihdella. Kyseessä ei siis ole pelkästään lukujono, vaan pikemminkin joukko lukujonoja, missä kertoimet a_k ajatellaan kiinnitetyiksi ja muuttujalle x sallitaan erilaisia (reaali)arvoja.†

†Tapauksessa $x = 0$ potenssisarja tulkitaan lukujonoksi $\{a_0, a_0, \ldots\}$, eli sarjamerkinnässä sovitaan, että $0^0 = 1$.

Kuten tullaan havaitsemaan, potenssisarja on matematiikassa hyvin keskeinen käsite ja työkalu. Tässä tutkitaan toistaiseksi vain sarjan suppenemiskysymystä, jonka luonteva muotoilu on: Millä x:n arvoilla sarja suppenee? Koska suppenevan sarjan tapauksessa summa yleensä riippuu x:n arvosta, ilmaistaan tämä kirjoittamalla sarjan summaksi $s(x)$, luetaan 's x'.

Esimerkki 7 Perusmuotoinen geometrinen sarja $\sum_{k=0}^{\infty} x^k$ on esimerkki potenssisarjasta. Sarja suppenee täsmälleen kun $x \in (-1, 1)$, ja sarjan summa on tällöin $s(x) = 1/(1-x)$. $\square$

Esimerkin perusteella potenssisarjaa voi pitää geometrisen sarjan yleistyksenä. Osoittautuukin, että suppenevuustarkasteluissa geometrinen sarja on hyvä vertailukohta. Tällaiseen vertailuun perustuu esimerkiksi seuraava, potenssisarjojen suppenemisteorian keskeisin tulos.

Lause I.12.9 **(Potenssisarjan suppeneminen)** Potenssisarjalle $\sum_{k=0}^{\infty} a_k x^k$ on voimassa jokin seuraavista vaihtoehdoista:

(a) Sarja suppenee vain kun $x = 0$.

(b) $\exists \rho \in \mathbb{R}_+$ siten, että sarja suppenee, myös itseisesti, kun $x \in (-\rho, \rho)$ ja hajaantuu aina kun $|x| > \rho$. Tällöin ρ on sarjan *suppenemissäde.*

(c) Sarja suppenee itseisesti $\forall x \in \mathbb{R}$.

Todistus Riittää osoittaa, että (b) tai (c) toteutuu silloin kun (a) ei. Oletetaan siis, että sarja suppenee jollakin $x_0 \in \mathbb{R}$, $x_0 \neq 0$. Tällöin on oltava (ks. Korollaari I.12.5 ja Lause I.7.2)

$$a_n x_0^n \to 0 \;\Rightarrow\; |a_n x_0^n| \leq C \quad \forall n \quad (C \in \mathbb{R}_+).$$

Tästä seuraa, että jos $|x| < |x_0|$, niin

$$|a_n x^n| \leq C|x/x_0|^n,$$

jolloin majoranttiperiaatteen nojalla sarja suppenee itseisesti geometrisen sarjan ($q = |x/x_0| < 1$) tavoin. Siis on todettu: Jos sarja suppenee kun $x = x_0 \neq 0$, niin se suppenee itseisesti $\forall x \in (-|x_0|, |x_0|)$. Tästä seuraa edelleen, että jos sarja hajaantuu kun $x = x_1$, niin se hajaantuu aina kun $|x| > |x_1|$, sillä muutoin se jo todetun perusteella olisi sekä suppeneva että hajaantuva, kun $x = x_1$. On siis päätelty: Jos (a) ei ole voimassa, niin joko (c) sarja suppenee jokaisella $x \in \mathbb{R}$,

jolloin se suppenee jokaisella $x \in \mathbb{R}$ myös itseisesti, tai (b) sarja suppenee kun $|x| < \rho$ ja hajaantuu kun $|x| > \rho$, missä ρ määritellään (vrt. Lause I.11.2)

$$\rho = \sup \left\{ x \in \mathbb{R} \mid \text{sarja} \sum_{k=0}^{\infty} a_k x^k \text{ suppenee} \right\}.$$

Lause on näin todistettu. $\square$

Lauseen vaihtoehdoissa (a), (c) voidaan sopia merkintätavoista

$$\text{(a)} \;\; \rho = 0, \qquad \text{(c)} \;\; \rho = \infty.$$

Sikäli kuin potenssisarjan suppenemissäde pystytään määräämään, ratkaisee siis Lause I.12.9 sarjan suppenemiskysymyksen täydellisesti, lukuun ottamatta muuttujan arvoja $x = \rho$ ja $x = -\rho$, kun $\rho \in \mathbb{R}_+$. Nämä on selvitettävä tapauskohtaisesti, vrt. Esimerkki 9 jäljempänä. Keskeisimmässä laskennallisessa ongelmassa, eli suppenemissäteen määräämisessä, auttaa usein seuraava tulos:

Lause I.12.10 **(Potenssisarjan suppenemissäde)** Jos potenssisarjassa $\sum_{k=0}^{\infty} a_k x^k$ on $a_k \neq 0 \; \forall k$ ja on olemassa raja-arvo

$$\rho = \lim_{k \to \infty} \left| \frac{a_k}{a_{k+1}} \right|,$$

missä $\rho \in \mathbb{R}_+$, $\rho = 0$ tai $\rho = \infty$, niin sarjan suppenemissäde $= \rho$.

<u>Todistus</u> Olkoon ensin $\rho > 0$ ja $|x| < \rho$. Valitaan q siten, että $|x| < q < \rho$ (esim. $q = (|x| + \rho)/2$). Tällöin koska $|a_k|/|a_{k+1}| \to \rho$ ja $q < \rho$, niin jostakin indeksistä $k = m$ alkaen on $|a_k|/|a_{k+1}| \geq q$ (lukujonon suppenemisen määritelmässä valittu $\varepsilon = \rho - q$ ja $m > N$). Tällöin voidaan päätellä

$$|a_{k+1}| \leq q^{-1} |a_k|, \quad k = m, m+1, \ldots \quad \Rightarrow \quad |a_k| \leq q^{m-k} |a_m|, \quad k \geq m,$$

jolloin

$$|a_k x^k| \leq q^{m-k} |a_m| |x|^k = q^m |a_m| \left(\frac{|x|}{q} \right)^k = C \left(\frac{|x|}{q} \right)^k, \quad k \geq m.$$

Tässä on $|x|/q < 1$, joten päätellään majoranttiperiaatteen nojalla, että sarja $\sum_{k=0}^{\infty} a_k x^k$ suppenee itseisesti, indeksistä $k = m$ eteenpäin geometrisen sarjan tavoin. Siis jos $\rho \in \mathbb{R}_+$, niin sarja suppenee itseisesti kun $|x| < \rho$. Jos oletettu raja-arvo on $\rho = \infty$, niin ym. päättelyssä voidaan $\rho \in \mathbb{R}_+$ valita miten tahansa, joten tässä tapauksessa myös suppenemissäde $= \infty$.

Jos $\rho \in \mathbb{R}_+$ ja $|x| > \rho$, niin nähdään samalla tavoin kuin yllä, että indeksin m ollessa riittävän suuri pätee jokaisella $k \geq m$

$$|a_k x^k| \geq q^{m-k}|a_m||x|^k = q^m\,|a_m|\left(\frac{|x|}{q}\right)^k = C\left(\frac{|x|}{q}\right)^k, \quad k \geq m.$$

missä nyt $\rho < q < |x|$. Näin ollen $|a_k x^k| \to \infty$ kun $k \to \infty$, jolloin sarja $\sum_k a_k x^k$ hajaantuu (Korollaari I.12.6). Siis jos $\rho \in \mathbb{R}_+$, niin sarja suppenee kun $|x| < \rho$ ja hajaantuu kun $|x| > \rho$, eli sarjan suppenemissäde $= \rho$. Jos $\rho = 0$, niin sarja todetaan samalla päättelyllä hajaantuvaksi aina kun $x \neq 0$, eli tässä tapauksessa suppenemissäde $= 0$. Lause on näin todistettu. $\square$

Esimerkki 8 Määritä potenssisarjan suppenemissäde, kun kertoimet ovat

a) $a_k = k!$ b) $a_k = (-1)^k(k+1)^{-1}$ c) $a_k = k^2 \cdot 3^{-k}$ d) $a_k = 1/k!$

Ratkaisu Lauseen I.12.10 perusteella

a) $\rho = \lim_k 1/(k+1) = 0$ b) $\rho = \lim_k (k+2)/(k+1) = 1$

c) $\rho = \lim_k 3k^2/(k+1)^2 = 3$ d) $\rho = \lim_k (k+1) = \infty$ $\square$

Jos potenssisarja suppenee, kun $x \in A$, ja hajaantuu, kun $x \notin A$, niin joukkoa A ($A \subset \mathbb{R}$) sanotaan ko. sarjan *suppenemisvälik*si. Lauseen I.12.9 mukaisesti A on jokin vaihtoehdoista $(-\rho, \rho)$, $[-\rho, \rho)$, $(-\rho, \rho]$, $[-\rho, \rho]$, missä $\rho =$ suppenemissäde.

Esimerkki 9 Määritä potenssisarjan $\sum_{k=0}^{\infty} a_k x^k$ suppenemisväli, kun

a) $a_k = 1$ b) $a_k = 1/(k+1)$ c) $a_k = (-1)^k/(k+1)$ d) $a_k = 1/(k+1)^2$

Ratkaisu Lauseen I.12.10 laskukaava antaa suppenemissäteeksi $\rho = 1$ kaikissa tapauksissa, joten sarjat suppenevat, kun $|x| < 1$, ja hajaantuvat, kun $|x| > 1$. Tapauksissa $x = \pm 1$ esimerkkisarjat edustavat jo entuudestaan tuttuja geometrisen, harmonisen, alternoivan ja yliharmonisen sarjan tyyppejä. Tulokset yhdistämällä saadaan vastaukseksi

a) $(-1, 1)$, b) $[-1, 1)$, c) $(-1, 1]$, d) $[-1, 1]$. $\square$

Esimerkki 10 Määritä suppenemisväli potenssisarjalle

$$\sum_{k=0}^{\infty}(-1)^k(k^2+1)\frac{x^{2k}}{2^k} \;=\; 1 - x^2 + \frac{5}{4}x^4 - \ldots$$

Ratkaisu Lauseen I.12.10 laskukaava suppenemissäteelle ei sovellu suoraan, koska sarja on muotoa $\sum_{k=0}^{\infty} a_k x^k$, missä $a_k = 0$ parittomilla indeksin arvoilla. Tehtävä ratkeaa kuitenkin yksinkertaisesti vaihtamalla muuttujaksi $t = x^2$, jolloin sarja saa muodon $\sum_{k=0}^{\infty} b_k t^k$. Lause I.12.10 soveltuu tähän sarjaan, sillä

$$\lim_k \left| \frac{b_k}{b_{k+1}} \right| = \lim_k \frac{(k^2+1)\,2^{-k}}{[(k+1)^2+1]\,2^{-k-1}} = 2.$$

Siis sarja suppenee, kun $|t| = x^2 < 2$, ja hajaantuu, kun $|t| = x^2 > 2$. Suppenemissäde muuttujan x suhteen on näin ollen $\rho = \sqrt{2}$. Muuttujan arvoilla $x = \pm\sqrt{2}$ sarja nähdään hajaantuvaksi, joten suppenemisväli on $(-\sqrt{2}, \sqrt{2})$. $\square$

Esimerkki 11 Millä arvoilla $x \in \mathbb{R}_+$ suppenee sarja

$$\sum_{k=0}^{\infty} \frac{x^{2k-1/2}}{k!} = \frac{1}{\sqrt{x}} + x\sqrt{x} + \frac{x^3\sqrt{x}}{2} + \dots \ ?$$

Ratkaisu Sikäli kuin sarja suppenee, niin sekä osasummista että raja-arvosta $s(x)$ voidaan erottaa skaalaustekijä $1/\sqrt{x}$ (vrt. Lause I.7.3), jolloin

$$s(x) = \frac{1}{\sqrt{x}} \sum_{k=0}^{\infty} \frac{(x^2)^k}{k!}.$$

Tässä $\sum_{k=0}^{\infty} x^{2k}/k!$ on tavanomainen potenssisarja, joka suppenee $\forall x \in \mathbb{R}$, joten kysytty sarja suppenee jokaisella $x \in \mathbb{R}_+$. $\square$

Päätetään potenssisarjojen ensiesittely tulokseen, jolla on myöhempää käyttöä.

Lause I.12.11 Potenssisarjoilla $\sum_{k=1}^{\infty} k^m a_k x^k$ on sama suppenemissäde $\forall m \in \mathbb{Z}$.

Todistus Olkoon sarjan $\sum_k a_k x^k$ $(m = 0)$ suppenemissäde joko $\rho \in \mathbb{R}_+$ tai $\rho = \infty$ ja olkoon $0 < |x| < \rho$. Valitaan $q \in (0,1)$ siten, että $|x|/q < \rho$ (esim. $q = 2|x|/(|x| + \rho)$, jos $\rho \in \mathbb{R}_+$, tai $q = 1/2$, jos $\rho = \infty$), ja kirjoitetaan

$$k^m |a_k||x|^k = (k^m q^k)|a_k| \left(\frac{|x|}{q}\right)^k, \quad k \geq 1,\ m \in \mathbb{Z}.$$

Tässä lukujono $\{k^m q^k,\ k = 1, 2, \dots\}$ on rajoitettu (koska se on suppeneva, ks. Harj.teht. 8), joten jollakin $C \in \mathbb{R}_+$ pätee

$$k^m |a_k||x|^k \leq C|a_k| \left(\frac{|x|}{q}\right)^k, \quad k = 1, 2, \dots$$

Oletuksen mukaan sarja $\sum_k |a_k|(|x|/q)^k$ suppenee (koska oli $|x|/q < \rho$), joten majoranttiperiaatteen nojalla myös sarja $\sum_k k^m|a_k||x|^k$ suppenee. On päätelty, että jos sarjan $\sum_k a_k x^k$ supenemissäde on $\rho \neq 0$, niin myös sarja $\sum_k k^m a_k x^k$ suppenee itseisesti aina kun $|x| < \rho$ ja jokaisella $m \in \mathbb{Z}$.

Kirjoittamalla $k^m a_k = b_k$, $a_k = k^{-m} b_k$ nähdään, että em. päättely on tarkasteltujen sarjojen $\sum_k a_k x^k$ ja $\sum_k k^m a_k x^k$ ($m \in \mathbb{Z}$) suhteen symmetrinen, ts. jos kumman tahansa suppenemissäde on $\rho \neq 0$, niin toinenkin suppenee itseisesti kun $|x| < \rho$. Tästä seuraa, että suppenemissäteiden on oltava kaikissa tapauksissa (myös kun $\rho = 0$) samat, sillä muu johtaisi loogiseen ristiriitaan tehtyjen päätelmien kanssa. $\square$

HARJOITUSTEHTÄVIÄ

1. Tiedetään, että $a_1 = 3$, $a_2 = 5$, $a_3 = 7$, $a_4 = -6$ ja $\sum_{k=1}^{\infty} a_k = 4$. Laske $\sum_{k=1}^{\infty}(a_k + a_{k+1} + a_{k+2} + a_{k+3})$.

2. Luokittele seuraavat sarjat suppeneviksi tai hajaantuviksi. Jos sarjassa on parameri α, niin oleta $\alpha \in \mathbb{Q}$ ja luokittele sarja eri α:n arvoilla.

a) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2 - k + 3}$ b) $\displaystyle\sum_{k=1}^{\infty} \frac{k-1}{k^2+1}$ c) $\displaystyle\sum_{k=1}^{\infty} \frac{100+k}{9901+k-k^3}$

d) $\displaystyle\sum_{k=1}^{\infty} \frac{\sqrt{k+1}-\sqrt{k}}{k}$ e) $\displaystyle\sum_{k=0}^{\infty} \frac{1}{\sqrt{k^2+100}}$ f) $\displaystyle\sum_{k=1}^{\infty} \left(\sqrt{\frac{k+1}{k}} - 1\right)$

g) $\displaystyle\sum_{k=1}^{\infty} \frac{\sqrt{k+1}\sqrt[3]{k+2}}{k^2+k+1}$ h) $\displaystyle\sum_{k=1}^{\infty} \frac{1+k!}{(1+k)!}$ i) $\displaystyle\sum_{k=1}^{\infty} \frac{2+k!}{(2+k)!}$

j) $\displaystyle\sum_{k=0}^{\infty} (-1)^k \left(\sqrt{k+7} - \sqrt{k+1}\right)$ k) $\displaystyle\sum_{k=1}^{\infty} \frac{3(-1)^k+1}{k}$ l) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt[100]{k}}$

m) $\displaystyle\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+|\alpha|}\right)$ n) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^\alpha + k^{2\alpha}}$ o) $\displaystyle\sum_{k=1}^{\infty} \frac{\sqrt{k^\alpha+1}-1}{k}$

p) $\displaystyle\sum_{k=1}^{\infty} (\sqrt{k^{2\alpha}+1} - k^\alpha)$ q) $\displaystyle\sum_{k=1}^{\infty} \frac{k^\alpha+1}{k^{2\alpha}+1}$ r) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k k^{4\alpha+3}+1}{k^{\alpha+2}}$

3. Arvioi, montako termiä on sarjoista otettava seuraavissa laskukaavoissa, jotta π saadaan oikein 10 desimaalin tarkkuudella:

$$\pi = \sum_{k=0}^{\infty} \frac{4(-1)^k}{2k+1} = 2\sqrt{3}\sum_{k=0}^{\infty} \frac{(-1)^k}{3^k(2k+1)} = 2\sqrt{3}\sqrt{\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)^2}}$$

4. a) Näytä vuorotteleva sarja $\sum_{k=0}^{\infty}(-1)^k/\sqrt{k+1}$ suppenevaksi majoranttiperiaatteen avulla ryhmittelemällä sarjan peräkkäiset termit yhteen niin, että sarjasta tulee positiiviterminen.

   b) Tarkastellaan sarjaa $\sum_{k=1}^{\infty} a_k$, missä

$$a_{2p-1} = \frac{1}{\sqrt{p+1}-1}\,, \quad a_{2p} = -\frac{1}{\sqrt{p+1}+1}\,, \quad p = 1, 2, \ldots$$

   Mitkä Lauseen I.12.3 oletuksista ovat voimassa tälle sarjalle? Ratkaise sarjan suppenemiskysymys ryhmittelemällä sarjan peräkkäiset termit parittain yhteen, eli tutkimalla sarjaa $\sum_{p=1}^{\infty} b_p = \sum_{p=1}^{\infty}(a_{2p-1} + a_{2p})$.

5. Osoita oikeaksi tai vääräksi:
   a) $\sum_k a_k$ suppenee $\Rightarrow$ $\{a_k\}$ on rajoitettu jono
   b) $\sum_k a_k$ suppenee ja $\sum_k b_k$ suppenee $\Rightarrow$ $\sum_k(a_k + b_k)$ suppenee
   c) $\sum_k a_k$ hajaantuu ja $\sum_k b_k$ hajaantuu $\Rightarrow$ $\sum_k(a_k + b_k)$ hajaantuu
   d) $\sum_k a_k$ suppenee ja $\sum_k b_k$ hajaantuu $\Rightarrow$ $\sum_k(a_k + b_k)$ hajaantuu
   e) $\sum_k a_k$ suppenee ja $\{b_k\}$ rajoitettu $\Rightarrow$ $\sum_k a_k b_k$ suppenee
   f) $\sum_k |a_k|$ suppenee ja $\{b_k\}$ rajoitettu $\Rightarrow$ $\sum_k |a_k b_k|$ suppenee
   g) $\sum_k a_k$ suppenee $\Rightarrow$ $\sum_k a_k^2$ suppenee
   h) $\sum_k a_k^2$ suppenee $\Rightarrow$ $\sum_k a_k$ suppenee
   i) $\sum_k a_k$ suppenee itseisesti $\Rightarrow$ $\sum_k a_k^4$ suppenee

6. Klassisen sarjaopin *suhdetesti* positiivitermisille sarjoille on väittämä: Jos on olemassa raja-arvo

$$\lim_k \frac{a_{k+1}}{a_k} = L \in \mathbb{R},$$

   niin sarja $\sum_k a_k$ suppenee, jos $L < 1$, ja hajaantuu, jos $L > 1$. Todista!

7. Klassisen sarjaopin *juuritesti* positiivitermisille sarjoille on väittämä: Jos on olemassa raja-arvo

$$\lim_k \sqrt[k]{a_k} = L \in \mathbb{R},$$

   niin sarja $\sum_k a_k$ suppenee, jos $L < 1$, ja hajaantuu, jos $L > 1$. Todista!

8. Todista potenssisarjojen teorian (Lauseet I.12.9 ja I.12.10) ja Korollaarin I.12.5 avulla raja-arvotulos

$$\lim_k k^m q^k = 0, \quad -1 < q < 1,\ m \in \mathbb{N}.$$

9. Olkoon potenssisarjojen $\sum_k a_k x^k$ ja $\sum_k b_k x^k$ suppenemissäteet ρ_1 ja ρ_2. Mitä varmaa voidaan sanoa sarjan $\sum_k (a_k + b_k)\, x^k$ suppenemissäteestä, jos a) $\rho_1 < \rho_2$, b) $\rho_2 = \rho_1 \in \mathbb{R}_+$, c) $\rho_1 = \rho_2 = \infty$, d) $\rho_1 = \rho_2 = 0$?

10. a) Potenssisarjan kertoimista tiedetään, että $k \le a_k \le k^2$ $\forall k \in \mathbb{N}$. Mikä on sarjan suppenemisväli?
b) Potenssisarjan kertoimet ovat $a_k = 1$, kun $k = m!$ ja $m \in \mathbb{N}$, muulloin on $a_k = 0$. Mikä on sarjan summa 20 merkitsevän numeron tarkkuudella, kun $x = 1/10$, ja mikä on sarjan suppenemisväli?

11. Määritä seuraavien potenssisarjojen suppenemisvälit:

$$\text{a)} \sum_{k=0}^{\infty} \frac{k^2}{k+1} x^k \quad \text{b)} \sum_{k=0}^{\infty} (k-2)2^{k+1} x^k \quad \text{c)} \sum_{k=1}^{\infty} (-1)^k (\frac{x}{k})^k$$

$$\text{d)} \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} x^k \quad \text{e)} \sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}} 3^k x^k \quad \text{f)} \sum_{k=1}^{\infty} \frac{\sqrt{k+1}}{k^2} 2^k x^k$$

$$\text{g)} \sum_{k=1}^{\infty} \frac{2^k}{k} x^{2k} \quad \text{h)} \sum_{k=1}^{\infty} \frac{1}{k^3 3^k} x^{3k} \quad \text{i)} \sum_{k=1}^{\infty} \frac{1}{k^{1/3} 3^k} x^{3k}$$

12. Totea seuraavat sarjat joko suppeneviksi tai hajaantuviksi potenssisarjojen teoriaan vedoten:

$$\text{a)} \sum_{k=1}^{\infty} (-1)^k k^{-10} \pi^k e^{-k} \quad \text{b)} \sum_{k=1}^{\infty} \frac{k!}{(3k)^k} \quad \text{c)} \sum_{k=1}^{\infty} \frac{(2k)^k}{k!}$$

13. Määritä muuttujan vaihdolla seuraaville sarjoille joukko $A \subset \mathbb{R}$ siten, että sarja suppenee täsmälleen kun $x \in A$:

$$\text{a)} \sum_{k=1}^{\infty} \frac{1}{k}(2x+1)^k \quad \text{b)} \sum_{k=1}^{\infty} (x^2 - 3x + 1)^k \quad \text{c)} \sum_{k=1}^{\infty} \frac{(-1)^k}{k}(x^2 - 4x + 3)^k$$

14. Todista Lause I.12.11 Lauseen I.12.10 avulla siinä tapuksessa, että raja-arvo $\lim_k |a_k/a_{k+1}|$ on olemassa.

15. Näytä, että sarjoilla $\sum_{k=0}^{\infty} a_k x^k$, $\sum_{k=0}^{\infty} (k-2)^2 a_k x^k$ ja $\sum_{k=1}^{\infty} k^{-2}(k+2)^3 a_k x^k$ on kaikilla sama suppenemissäde.

16. (*) Näytä, että sarjan $\sum_{k=1}^{\infty} 1/\sqrt{k}$ osasummille s_n pätee

$$2\sqrt{n+1} - 2 \;<\; s_n \;\le\; 2\sqrt{n} - 1, \quad n \in \mathbb{N}.$$

Vihje: Vertaa teleskooppisummiin!

17. (*) Todista: Jos positiivitermisen sarjan $\sum_k a_k$ termien jono $\{a_k\}$ on vähenevä ja sarja suppenee, niin $\lim_k k a_k = 0$.

18. (*) Esitä jokin rationaalikertoiminen potenssisarja, jonka suppenemisväli on $(-\sqrt[3]{e}, \sqrt[3]{e}\,]$ (e = Neperin luku).

Luku II

Vektorit ja analyyttinen geometria

Ihmisten luoman matemaattisen kulttuurin peruspilareita on matemaattinen malli nimeltä *euklidinen taso*. Termi viittaa kreikkalaisen **Eukleideen** n. 300 eKr kirjoittamaan teokseen Stoikheia eli 'Alkeet'. Eukleideen pyrkimyksenä oli asettaa selkeät aksiomaattiset perusteet silloiselle matematiikalle, erityisesti *geometrian* nimellä tunnetulle matematiikan suuntaukselle. Geometria on — toisin kuin edellä tarkasteltu lukujen algebra — näköaistimuksiin hyvin voimakkaasti vetoava matematiikan laji. Algebra ja geometria edustavat nykyisessäkin matematiikassa kahta matemaattisen ajattelun suurta päähaaraa.[†]

Eukleideen teos on kaikkien aikojen menestynein matematiikan oppikirja, jonka vaikutus varhaisena aksiomaattisen ajattelun esikuvana on ulottunut matematiikan ulkopuolellekin. Käytännön kannalta 'tyhjästä' teos ei tietenkään syntynyt, vaan sitä edelsi itse asiassa vuosituhantinen laskentamenetelmien perinne sekä Egyptissä että Babyloniassa. Esimerkiksi Egyptissä geometrisia menetelmiä olivat kehittäneet (jos historioitsijoihin on uskominen) sekä maanmittarit että papit.

Tässä luvussa lähdetään euklidisen geometrian perusteista tasossa ja avaruudessa ja paneudutaan tarkastelemaan geometrian pohjalta syntyvää *vektorin* käsitettä ja vektoreilla laskemista eli *vektorialgebraa*. Vektoreiden avulla geometriset ongelmat on mahdollista 'algebralisoida' eli muotoilla ja ratkaista vektorialgebran ongelmina. Tällaista — historiallisesti melko myöhäsyntyistä — lähestymistapaa geometriaan kutsutaan *analyyttiseksi* geometriaksi.

[†]Algebran ja geometrian välinen rajanveto matematiikassa ei ole aina helppoa. Ongelmaa kuvastaa esim. moderni matematiikan laji nimeltä *algebrallinen geometria*.

II.1 Euklidinen taso. Geometriset luvut

Oletetaan lähtökohtaisesti tutuiksi mm. sellaiset geometrian käsitteet kuin

- *piste*, *suora*, *jana*
- *kulma*, *kolmio*, *n-kulmio*, *monikulmio*, *yhdenmuotoisuus*
- *suunta* (*puolisuora*, säde), *yhdensuuntaisuus*, *suunnikas*
- *suora kulma*, *Pythagoraan lause*, *suorakulmio*
- *ympyrä*, janan *pituusmitta* ('mittatikku')

Näiden taustalla ovat euklidisessa geometriassa lähtökohtana oletetut perusaksioomat, jotka tunnetaan *Hilbertin aksioomina*†.

Euklidisen tason perusolioita ovat pisteet, joista muiden olioiden (kuten jana, suora ym.) ajatellaan koostuvan. Euklidista tasoa sanotaankin *pisteavaruudeksi*. Sen symbolinen merkki on E^2, luetaan 'E kaksi'.

†Saksalainen **David Hilbert** (1862-1943) oli modernin matematiikan suunnannäyttäjä ja kaikkien aikojen huomattavimpia matemaatikkoja. Hilbert tunnetaan mm. hänen v. 1900 asettamistaan ratkaisemattomista matematiikan ongelmista, jotka sittemmin on nimetty *Hilbertin ongelmiksi*. Euklidisen geometrian perusaksioomat ovat peräisin Hilbertin pienestä klassikkoteoksesta "Grundlagen der Geometrie" vuodelta 1899. Teos oli ensimmäinen kattava ja systemaattinen esitys geometrian — myös eräiden epäeuklidisten geometrioiden — aksiomaattisista perusteista.

Sellaiset geometriset operaatiot kuin yhdensuuntaisten suorien piirtäminen, suoran kulman konstruoiminen, tai annetun geometrisen olion (pistejoukon) *siirto* ja *kierto* eli nk. *euklidinen liike* (esim. janan siirtely pituus säilyttäen) tapahtuvat euklidisessa geometriassa ympyräviivojen (käytännössä harpin) avulla. Myös suora on konstruoitavissa pelkästään harppia käyttäen, sillä suora on sellaisten pisteiden joukko eli *ura*, jotka ovat yhtä etäällä kahdesta annetusta pisteestä. — Viivoitinta tarvitaan geometriassa siis vain pisteiden 'massatuotantovälineenä'.

Kulma

Kulman määrittelee geometriassa kaksi suuntaa. Konkreettisemmin kulma on ajateltavissa pistejoukoksi, jonka muodostavat kaksi samasta pisteestä lähtevää puolisuoraa, tai vaihtoehtoisesti näiden puolisuorien rajaamaksi euklidisen tason *sektoriksi*. Jälkimmäisessä tulkinnassa (ei edellisessä!) on tehtävä ero *sisäkulman* ja *ulkokulman* välillä. Jos A ja B ovat puolisuorilla olevia pisteitä ja O on puolisuorien yhteinen piste eli kulman *kärki* ($A \neq O$, $B \neq O$), niin kulma merkitään $\measuredangle AOB$. Kuvioon on merkitty sisäkulma (sektoritulkinta!) ympyräkaarella.

Kulmien vertailussa ja laskuoperaatioissa (myös kulman mittauksessa jäljempänä) tulkitaan kulma yleensä sektoriksi. Sanotaan tällöin, että kaksi kulmaa ovat *yhtä suuret*, jos ne voidaan muuntaa toisikseen euklidisella liikkeellä.† Yhtäsuuruus on siis tarkistettavissa geometrisesti. Myös kahden kulman *summa* ja *erotus* voidaan määrätä geometrisesti siirtämällä toista kulmaa euklidisella liikkeellä niin, että kulmat joutuvat 'vierekkäin' (summa) tai 'päällekkäin' (erotus). Geometrisella yhteenlaskulla on esimerkiksi todistettavissa tunnettu kolmion kulmia koskeva väittämä: Kulmien summa = *oikokulma* (Harj.teht. 1b).

Jos kahdessa kolmiossa (vastin)kulmat ovat yhtä suuret, niin kolmiot ovat *yhdenmuotoiset*. Geometrian perusväittämiin kuuluu, että yhdenmuotoisissa kolmioissa vastinsivujen pituuksien suhteet ovat samat. Väittämä näyttelee keskeistä roolia monissa geometrisissa todistuksissa (ks. esim. Harj.teht. 2).

†Kulmien yhtäsuuruus on kulmien joukossa määritelty ekvivalenssirelaatio, vrt. Luku I.3.

Geometriset luvut

Kulmien mittalukuja ei euklidisen geometrian perusoperaatioissa tarvita. Sen sijaan janan pituuden mittaus harpin ja mittajanan avulla on perusgeometriaa. Operaatio tuo geometriaan reaaliluvut (tai ainakin osan niistä) ja myös lukuihin liittyvät laskutoimitukset. Lähtökohtana pidetään lukuja 0 ja 1:

$$0 \leftrightarrow \text{piste}, \quad 1 \leftrightarrow \text{mittajana}.$$

Kun tarkastellaan kiinteää suoraa, jolta valitaan referenssipiste O luvun 0 vastineeksi, niin kahden (janan pituutena) annetun luvun summa ja erotus voidaan määrätä geometrisesti:

Kuvassa summalla ja erotuksella on geometriset vastineet janoina ja suoran pisteinä:

$$x + y \leftrightarrow OQ \leftrightarrow Q,$$
$$x - y \leftrightarrow OR \leftrightarrow R.$$

Kun janan loppupiste on O:sta vasemmalle (suuntahan oli määritelty!), voidaan OR tulkita negatiiviseksi luvuksi $x-y$. Näin on siis lukujärjestelmässä määritelty sekä yhteenlasku että vähennyslasku. Kun sääntöjä $+, -$ sovelletaan peräkkäin lähtien luvusta 1 (= mittajana sijoitettuna suoralle) nähdään, että minkä tahansa kokonaisluvun $m \in \mathbb{Z}$ geometrinen vastine saadaan äärellisellä määrällä geometrisia operaatioita. Siis $\mathbb{Z}$ on konstruoitu geometrisesti. Myös millä tahansa kokonaisluvulla jakaminen onnistuu 'viipaloimalla' luku yhdensuuntaisilla suorilla, joita voidaan <u>tasossa</u> tuottaa harpilla ja viivoittimella (ks. kuvio alla).

Yleisemmin jos lähtökohtana on kaksi mittalukua x, y, niin niiden välinen kerto- ja jakolasku voidaan suoraan geometrisoida yhdenmuotoisten kolmioiden avulla käyttäen suhteita

$$\begin{aligned} \frac{z}{x} &= \frac{y}{1} \quad (z = xy), \\ \frac{z}{1} &= \frac{x}{y} \quad (z = x/y). \end{aligned}$$

Kuvassa (alla) on

$$\begin{aligned} &OA \leftrightarrow 1, \quad OB \leftrightarrow y, \quad BC \leftrightarrow OD \leftrightarrow x, \\ &OE \leftrightarrow x \cdot y, \quad AF \leftrightarrow x/y. \end{aligned}$$

Jatkossa nimetään luvut, jotka ovat konstruoitavissa (janan pituutena) äärellisellä määrällä geometrisia operaatioita, *geometrisiksi luvuiksi*†, ja käytetään tälle lukujoukolle symbolia $\mathbb{G}$:

$$\mathbb{G} \;=\; \{\text{geometriset luvut}\}.$$

Tiedetään jo, että $\mathbb{G}$ sisältää kaikki rationaaliluvut: $\mathbb{G} \supset \mathbb{Q}$. Lisäksi tiedetään, että $\mathbb{G}$ on peruslaskutoimitusten suhteen suljettu:

$$x, y \in \mathbb{G} \quad \Rightarrow \quad x+y,\; x-y,\; xy,\; x/y \in \mathbb{G} \quad (y \neq 0).$$

Tämä merkitsee, että $(\mathbb{G}, +, \cdot)$ on <u>kunta</u>.

Lukujoukon $\mathbb{G}$ muodostamisessa eivät geometrian mahdollisuudet lopu aivan peruslaskutoimituksiin, sillä Pythagoraan lauseen mukaan myös operaatiot

$$\begin{aligned} x, y \;&\mapsto\; \sqrt{x^2+y^2}, \quad x, y \in \mathbb{G}, \\ x, y \;&\mapsto\; \sqrt{x^2-y^2}, \quad x, y \in \mathbb{G},\; x \geq y \end{aligned}$$

ovat geometrisesti toteutettavissa. Erityisesti on siis vuosituhansia päänvaivaa aiheuttanut luku $\sqrt{2}$ geometristen lukujen joukossa:

Itse asiassa neliöjuurioperaatio $x \mapsto \sqrt{x}$ on geometrisesti mahdollinen, olipa mittaluku $x \in \mathbb{G}$ (ajatellaan $x \geq 0$) mikä tahansa:

†Englanninkielinen termi on 'constructible numbers'. Suomennos ei ole vakiintunut.

Tähän geometrian mahdollisuudet kuitenkin loppuvat, joten $\mathbb{G}$ koostuu kaikkiaan luvuista, jotka on saatu mittaluvusta 1 lähtien äärellisellä määrällä viittä eri operaatiota: $+$, $-$, $\cdot$, $/$ ja $\sqrt{\ }$. Kun $\mathbb{G}$ vielä varustetaan geometrisella järjestysrelaatiolla (janojen pituuksien vertailu harpin avulla), niin tuloksena on pelkästään geometrisiin operaatioihin perustuva järjestetty kunta $(\mathbb{G}, +, \cdot, <)$. Tämä on rationaalilukujen kunnan pienin mahdollinen laajennus, joka täyttää ehdon $x \in \mathbb{G}\ \&\ x > 0 \ \Rightarrow\ \sqrt{x} \in \mathbb{G}$ (vrt. lukujoukon $\mathbb{J}$ konstruktio Luvussa I.2). On osoitettavissa, että kaikki geometriset luvut ovat algebrallisia (!), joten pätee

$$\mathbb{Q} \subset \mathbb{G} \subset \mathbb{A} \subset \mathbb{R}.$$

Kun geometriset luvut sijoitetaan em. tavalla kiinteälle suoralle, voidaan näiden lukujen 'väliin' ajatella sijoitelluksi myös muut reaaliluvut. Näin ajatellen jokainen reaaliluku saa geometrisen vastineen *lukusuoran* pisteenä.

Kulman mittaluku

Kulmien käsittelyssä geometrian mahdollisuudet ovat melko rajalliset. Kulmien yhteen- ja vähennyslasku onnistuu geometrisesti, kuten todettiin, ja kulman voi helposti myös puolittaa eli jakaa kahteen yhtä suureen osaan geometrian keinoin. Sen sijaan jo kulman jako kolmeen yhtä suureen osaan (äärellisellä) geometrisella konstruktiolla on mahdotonta, ellei kyse ole erikoisesta (kuten suorasta) kulmasta. Geometrian keinot loppuvat myös, kun kulmalle yritetään määrittää sen suuruutta kuvaava mittaluku. Geometrisesti pystytään konstruoimaan vain lukujono, jonka raja-arvo mittaluku on, eli mittaluku saadaan selville vain päättymättömällä konstruktiolla.

Tarkastellaan kulmaa $\measuredangle AOB$, missä pisteet A ja B on sijoitettu O-keskisen *yksikköympyrän* kehälle, ts. janat OA ja OB ovat valitun mittajanan pituiset (vastaten lukua 1 lukusuoralla).

Kuvion tilanteessa asetetaan kulman $\measuredangle AOB$ mittaluvuksi yksikköympyrän *kaarenpituus* pisteiden A ja B välillä. Koska vaihtoehtoja on kaksi, on päätettävä, mitataanko sisä- vai ulkokulmaa. Seuraavaksi on ratkaistava vakavampi ongelma: Miten määritellä ympyräkaaren pituus? Seuraavassa ongelman ratkaisussa yhdistellään geometrian ja algebran keinoja.

Koska kulman $\measuredangle OAB$ puolitus onnistuu geometrisesti, niin saman tien kulma on jaettavissa $n = 2^k$ yhtä suureen osaan jokaisella $k \in \mathbb{N}$. Olkoon $k \in \mathbb{N}$ annettu ja merkitään jaossa syntyviä osakulmia $\measuredangle P_{i-1}OP_i,\ i = 1 \ldots n = 2^k$, missä $P_0 = A$, $P_n = B$ ja muutkin pisteet P_i ovat samalla O-keskisellä ympyräkaarella. Yhdistetään nyt peräkkäiset pisteet P_i janoilla *murtoviivaksi* $P_0 \ldots P_n$ ja asetetaan $a_k =$ ko. murtoviivan pituus, ts. $a_k =$ janojen $P_{i-1}P_i,\ i = 1 \ldots n$ pituuksien summa. On ilmeistä, että lukujonon $\{a_k\}$ jokainen termi on määrättävissä geometrisesti lähtien luvusta $a_0 =$ janan AB pituus, jolloin on $a_k/a_0 \in \mathbb{G}\ \forall k$ (mahdollisesti $a_0 \notin \mathbb{G}$). Kuviossa on konstruoitu murtoviiva $P_0 \ldots P_n$ indekseillä $k = 0, 1, 2$, kun mitattavana on suora kulma.

Pythagoraan lauseeseen vedoten on helppo näyttää, että lukujono $\{a_k\}$ on aidosti kasvava (Harj.teht. 5a), ja geometrisin perustein on myös osoitettavissa, että $\{a_k\}$ on rajoitettu lukujono (Harj.teht. 5b). Tämän jälkeen on turvattava algebraan: Koska $\{a_k\}$ siis on kasvava ja rajoitettu lukujono, niin on olemassa raja-arvo $\lim_k a_k = a \in \mathbb{R}$ (Lause I.8.3, Määritelmä I.9.1). Määritellään kaaren AB pituus, eli kulman $\measuredangle AOB$ mittaluku, kyseisenä raja-arvona a.

Mittayksikköä kulman kaarenpituusmittauksessa sanotaan *radiaaniksi*. Tunnetusti luku $\pi = 3.14159..$ määritellään yksikköympyrän puolikkaan kaarenpituutena, eli

$$\boxed{\pi \ = \ \text{oikokulman mittaluku.}}$$

Koska π on transkendenttinen luku (ks. Luku I.11), niin myös suoran kulman mittaluku ($= \pi/2$) on transkendenttinen, samoin esimerkiksi tasasivuisen kolmion kulman mittaluku ($= \pi/3$).

Kulmia on perinteisesti mitattu myös *asteina* (°), jolloin oikokulman mitaksi on sovittu 180°. Yleisemmin mittaluvut asteina ja radiaaneina saadaan skaalaamalla toisistaan:

$$\boxed{\text{kulma asteina} \;=\; \frac{180}{\pi} \cdot (\text{kulma radiaaneina}).}$$

HARJOITUSTEHTÄVIÄ

1. a) Suoran leikatessa kaksi yhdensuuntaista suoraa syntyy kahdeksan kulmaa. Näytä, että näistä jokainen on yhtä suuri kolmen muun kanssa.
   b) Todista: Kolmion kulmien summa = oikokulma.

2. a) Jana jakaa suorakulmaisen kolmion kahteen osaan siten, että syntyy kolme yhdenmuotoista kolmiota. Lähtien tästä ajatuksesta ja kolmioiden yhdenmuotoisuusopista todista Pythagoraan lause!
   b) Kolmion *keskijana* yhdistää kolmion kärjen vastakkaisen sivun keskipisteeseen. Lähtien kolmioiden yhdenmuotoisuusopista todista: Kolmion ABC keskijana AD puolittaa jokaisen janan BC suuntaisen janan, jonka päätepisteet ovat sivuilla AB ja AC.

3. Esitä geometrinen konstruktio (aseina harppi ja viivoitin), jonka tulos on
   a) ympyrä, joka kulkee annetun kolmen pisteen kautta,
   b) annetun kolmion sisään piirretty (suurin mahdollinen) ympyrä,
   c) suora, joka kulkee annetun pisteen kautta ja sivuaa annettua ympyrää,
   d) annetun janan AB *kultainen leikkaus* eli janalla oleva piste C siten, että jos a, b, c ovat janojen AB, AC ja CB pituudet, niin $a/b = b/c$.

4. Seuraavista luvuista kaksi ei ole geometrisia. Mitkä kolme ovat?
   a) $\sqrt[1024]{17}$ b) $\sqrt[3]{512}$ c) $\sqrt[6]{343}$ d) $\sqrt[3]{7\sqrt{2}+5}$ e) $\sqrt[3]{5\sqrt{2}+7}$

5. (*) Suoran kulman mittaluku radiaaneina määritellään raja-arvona $a = \lim_k a_k$, missä $4a_k$ = sen säännöllisen 2^{k+2}-kulmion piiri, jonka kärjet ovat yksikköympyrällä.
   a) Näytä, että lukujono $\{a_k\}_{k=0}^{\infty}$ on aidosti kasvava.
   b) Näytä, että $a_k < 2\ \forall k$.
   c) Näytä, että $\{a_k\}$ on laskettavissa palautuvana lukujonona
   $$a_0 = \sqrt{2}, \quad a_{k+1} = \frac{2a_k}{\sqrt{2+\sqrt{4-(2^{-k}a_k)^2}}}, \quad k = 0, 1, \ldots$$
   Laske a_k, $k = 1 \ldots 5$ ja vertaa lukuun $\frac{\pi}{2}$.

II.2 Tason vektorit

Vektori on matemaattisena käsitteenä hieman kaksijakoinen. Se voi olla lähtökohtaisesti geometrinen olio, jolloin se sovelluksissa ajatellaan usein fysikaalisena, 'maailmassa vaikuttavana'. Tällaisia 'fysiikkavektoreita' ovat mm.

- paikka, nopeus, kiihtyvyys, kulmanopeus
- voima, momentti
- X-kentän voimakkuus, X=gravitaatio, sähkö, magneetti...

Toisaalta vektori voidaan tulkita myös algebrallisesti, jolloin päädytään geometris-fysikaalista vektoria yleisempään (myös abstraktimpaan) vektorikäsitteeseen. Jäljempänä nähdään ensimmäisiä esimerkkejä myös tällaisista 'matematiikkavektoreista' (myöhemmin esimerkkejä tulee paljon lisää). Jatkossa lähtökohta vektorin käsitteeseen on kuitenkin aluksi geometrinen.

Euklidisessa tasossa E^2 vektori määritellään geometrisesti *suuntajanana*, ts. janana, jolla on suunta. Ajatus on tällöin, että jana sisältyy sen toisesta päätepisteestä lähtevään puolisuoraan. Puolisuoran ilmaisema suunta kuvataan janan päähän merkityllä nuolenkärjellä:

Vektorilla tarkoitetaan täsmällisemmin vain suuntajanan sisältävää <u>osatietoa</u>

$$\{\text{(janan) pituus, suunta}\}.$$

Tiettyyn suuntajanaan liittyvä vektori merkitään $\overrightarrow{AB}$, yleisemmin käytetään vektorimerkintöjä $\vec{a}, \vec{b}, \vec{v}$ jne.

Huomattakoon, että vektori $\overrightarrow{AB}$ siis <u>ei</u> 'tiedä', missä piste A sijaitsee euklidisessa tasossa. Näin ollen 'vektoria saa siirtää', kunhan 'ei kierrä eikä venytä'. (Fysikaalinen vastine: Voiman vaikutus on sama vaikutuspisteestä riippumatta.)

Jos tunnetaan vektorit $\vec{a} = \overrightarrow{AB}$ ja $\vec{b} = \overrightarrow{BC}$, niin tunnetaan myös $\vec{c} = \overrightarrow{AC}$:

Sanotaan, että $\vec{c}$ on vektoreiden $\vec{a}, \vec{b}$ *summavektori* ja merkitään

$$\vec{c} = \vec{a} + \vec{b}.$$

Summa määrätään siis geometrisesti kolmiodiagrammilla. Näin määritellylle vektorien yhteenlaskulle pätevät tavanomaiset vaihdanta- ja liitäntälait:

(V1) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$

(V2) $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$

Tässä ei ole kyse aksioomista (vrt. vastaavat kunta-aksioomat, Luku I.2) vaan väittämistä, joiden todistus on geometrinen:

Vektorin $\vec{a} = \overrightarrow{AB}$ *itseisarvo* (arkisemmin 'pituus') on

$$|\vec{a}| = |\overrightarrow{AB}| = |AB| = \text{janan } AB \text{ pituus.}$$

Tälle pätee *kolmioepäyhtälö* (vrt. Lause I.2.3)

$$|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}|.$$

Jos $\lambda \in \mathbb{R}$, niin $\lambda\vec{a}$ tarkoittaa *skaalattua* vektoria, jolle pätee:

(1) $|\lambda\vec{a}| = |\lambda||\vec{a}|$

(2) $\lambda\vec{a} \uparrow\uparrow \vec{a}$, jos $\lambda > 0$, ja $\lambda\vec{a} \uparrow\downarrow \vec{a}$, jos $\lambda < 0$

Tässä $\uparrow\uparrow$ tarkoittaa yhdensuuntaisuutta ja $\uparrow\downarrow$ vastakkaissuuntaisuutta. Vektorin skaalausta $\vec{a} \mapsto \lambda\vec{a}$ ($\lambda \in \mathbb{R}$) sanotaan vektorialgebrassa yleisemmin *skalaarilla kertomiseksi.* Skaalaajia eli reaalilukuja kutsutaan siis tässä laskuoperaatiossa *skalaareiksi.* Skalaarilla kertomisen ja yhteenlaskun määritelmistä on pääteltävissä (Harj.teht. 1a), että seuraavat 'luonnonmukaiset' säännöt ovat voimassa (vrt. kunta-aksioomat Luvussa I.2) :

(V3) $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$

(V4) $\lambda(\mu\vec{a}\,) = (\lambda\mu)\vec{a}$

(V5) $\lambda(\vec{a} + \vec{b}\,) = \lambda\vec{a} + \lambda\vec{b}$

(V6) $1\vec{a} = \vec{a}$

Operaation $\vec{a} \to \lambda\vec{a}$ määritelmän mukaisesti on $|0\vec{a}| = 0$. Tällaista vektoria sanotaan *nollavektoriksi* ja merkitään $0\vec{a} = \vec{0}$. Nollavektorin suunta on epämääräinen (ei määritelty). Nollavektori on vektorien yhteenlaskun nolla-alkio, ts. pätee

(V7) $\vec{a} + \vec{0} = \vec{a}\ \forall\vec{a}$

Jokaisella vektorilla $\vec{a}$ on myös *vastavektori* $-\vec{a}$, jolle pätee

(V8) $\vec{a} + (-\vec{a}\,) = \vec{0}$

Nimittäin kun säännössä (V3) valitaan $\lambda = 1, \mu = -1$ ja käytetään nollavektorin määritelmää ja sääntöä (V6), niin seuraa

$$-\vec{a} = (-1)\,\vec{a}.$$

Vastavektorin avulla tulee määritellyksi myös vektorien vähennyslasku:

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b}\,).$$

Säännöt (V1)–(V8), yhdessä nollavektorin ja vastavektorin määritelmien kanssa, muodostavat (tason vektoreiden)) *vektorialgebran* laskulait.

Esimerkki 1 Laskulaeista (V1)–(V8) seuraava identiteetti

$$\vec{a} + \frac{2}{3}\left(\frac{1}{2}\vec{b} - \vec{a}\right) \;=\; \vec{b} + \frac{2}{3}\left(\frac{1}{2}\vec{a} - \vec{b}\right) \;=\; \frac{2}{3}\left[\vec{b} + \frac{1}{2}\left(\vec{a} - \vec{b}\right)\right] \;=\; \frac{1}{3}\left(\vec{a} + \vec{b}\right)$$

voidaan lukea: Kolmion keskijanat leikkaavat toisensa samassa pisteessä, joka jakaa keskijanat suhteessa 2 : 1 (vrt. kuvio). Leikkaupistettä sanotaan kolmion *keskiöksi* (keskipisteeksi). □

Vektoriavaruus

Jos tason vektorien joukkoa merkitään symbolilla V, niin edellä olevilla määritelmillä on syntynyt algebra, joka merkittäköön $(V, \mathbb{R})$. Laskusääntöjen (V1)–(V8) (yleisemmin: aksioomien) ollessa voimassa sanotaan, että kyseessä on (lineaarinen) *vektoriavaruus*. Käsitteeseen siis sisältyvät

- vektorien muodostama joukko V (jossa operoidaan)
- vektorien yhteenlasku $(+)$
- nk. *skalaarien* muodostama *kertojakunta*, tässä $= \mathbb{R}$
- skalaarin ja vektorin kertolasku

Lyhennetty sanonta 'V on vektoriavaruus' tarkoittaa koko tätä algebraa, jolloin kertojakunta on juuri $\mathbb{R}$, tai muuten asiayhteydestä selvä. Sanonnat *reaalikertoiminen* vektoriavaruus tai yleisemmin $\mathbb{K}$-*kertoiminen* vektoriavaruus kiinnittävät kertojakunnan tarkemmin. Kertojakunnan ei siis tarvitse olla $\mathbb{R}$, vaan se voisi olla esim. $\mathbb{K} = \mathbb{Q}$ tai $\mathbb{K} = \mathbb{G} = \{\text{geometriset luvut}\}$. Lineaarinen vektoriavaruus on hyvin keskeinen käsite matematiikassa, ja sillä on keskeinen rooli myös monissa matematiikan sovelluksissa.

Kanta ja koordinaatisto

Olkoon $\vec{v} \in V$ mielivaltainen tason vektori ja $\vec{a}, \vec{b} \in V$ kaksi vektoria, joille pätee

$$\vec{a}, \vec{b} \neq \vec{0}, \quad \vec{a} \neq \lambda \vec{b} \quad \forall\, \lambda \in \mathbb{R}.$$

Tällöin voidaan geometrisella konstruktiolla (ks. kuvio) löytää yksikäsitteiset $x, y \in \mathbb{R}$ siten, että

$$\vec{v} = x\vec{a} + y\vec{b}.$$

Sanotaan tällä perusteella, että $\vec{v}$ on $\vec{a}$:n ja $\vec{b}$:n *lineaariyhdistely* (lineaarinen yhdistely, *lineaarikombinaatio*) ja että $x\vec{a}$ ja $y\vec{b}$ ovat $\vec{v}$:n (vektori)*komponentit* $\vec{a}$:n ja $\vec{b}$:n suuntaan.

Kun siis kaksi em. ehdot täyttävää vektoria on valittu, on koko V esitettävissä muodossa

$$V = \{x\vec{a} + y\vec{b} \mid x, y \in \mathbb{R}\}.$$

Sanotaan, että vektoripari $\{\vec{a}, \vec{b}\}$ on V:n *kanta*† ja että x, y ovat vektorin $\vec{v} = x\vec{a} + y\vec{b} \in V$ *koordinaatit* kannassa $\{\vec{a}, \vec{b}\}$. Jos koordinaatit esitetään lukuparina (x, y) (pari = kahden alkion järjestetty joukko), niin on synnytetty yhteys V:n ja tällaisten lukuparien joukon välille. Merkitään jälkimmäistä joukkoa symbolilla $\mathbb{R}^2$, äännetään 'R kaksi':

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R} \ \& \ y \in \mathbb{R}\}.^{\ddagger}$$

Em. sopimuksilla on itse asiassa luotu *kääntäen yksikäsitteinen* (=molempiin suuntiin yksikäsitteinen) *vastaavuus* V:n ja $\mathbb{R}^2$:n välille. Merkitään tätä:

$$V \leftrightarrow \mathbb{R}^2.$$

Vastaavuuteen liittyen nimetään $\mathbb{R}^2$ kantaan $\{\vec{a}, \vec{b}\}$ liittyväksi V:n *koordinaattiavaruudeksi*. Nimitys jo viittaa siihen, että myös $\mathbb{R}^2$ on tulkittavissa lineaariseksi vektoriavaruudeksi. Nimittäin kaikki vektoriavaruuden laskulait ovat voimassa, kun yhteenlasku ja skalaarilla kertominen $\mathbb{R}^2$:ssa määritellään

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$
$$\lambda(x, y) = (\lambda x, \lambda y).$$

Nämä operaatiot myös vastaavat V:n laskutoimituksia, sillä

$$\vec{v}_1 = x_1\vec{a} + y_1\vec{b},\ \vec{v}_2 = x_2\vec{a} + y_2\vec{b} \ \Rightarrow\ \vec{v}_1 + \vec{v}_2 = (x_1 + x_2)\vec{a} + (y_1 + y_2)\vec{b},$$
$$\vec{v} = x\vec{a} + y\vec{b} \ \Rightarrow\ \lambda\vec{v} = (\lambda x)\vec{a} + (\lambda y)\vec{b}.$$

†Kanta on vektoreiden järjestetty joukko, joten merkintä $(\vec{a}, \vec{b})$ olisi loogisempi. Erinäisten sekaannusten välttämiseksi käytetään tässä yhteydessä kuitenkin yleisemmin aaltosulkeita.

‡Joukko $\mathbb{R}^2$ voidaan myös merkitä $\mathbb{R} \times \mathbb{R}$, jolloin käytetään joukko-opillista *karteesisen tulon* merkintää

$$A \times B = \{(x, y) \mid x \in A \ \& \ y \in B\}.$$

Tämä luetaan 'A:n ja B:n karteesinen tulo', tai vain 'A risti B'. Joukon $A \times B$ alkiot ovat siis pareja. Näiden välinen samastusrelaatio on

$$(x_1, y_1) = (x_2, y_2) \quad \Leftrightarrow \quad x_1 = x_2 \ \wedge \ y_1 = y_2.$$

Vektoriavaruus $(\mathbb{R}^2, \mathbb{R})$, mainituin laskuoperaatioin, on esimerkki (lajissaan ensimmäinen) algebrallisten vektoreiden muodostamasta vektoriavaruudesta. Tämä avaruus on siis geometrisista tulkinnoista vapaa. Toisaalta tulkitsemalla $(\mathbb{R}^2, \mathbb{R})$ tason vektorien koordinaattiavaruudeksi (valitussa kannassa) saadaan vektorien laskuoperaatiot muunnetuksi algebralliseen muotoon. Esimerkiksi yhteenlaskussa tällainen laskukaavio on

$$\begin{array}{ccccc} \vec{v}_1 & \vec{v}_2 & & \vec{v}_1 + \vec{v}_2 \\ \downarrow & \downarrow & & \uparrow \\ (x_1, y_1) & (x_2, y_2) & \mapsto & (x_1 + x_2, y_1 + y_2) \end{array}$$

Kaavion etuna on, että itse laskuoperaatiossa ei tarvita mitään tietoa vektorien 'ulkonäöstä'. Monia geometrian tuloksia voidaan tällä tavoin johtaa algebran keinoin (vrt. Esimerkki 1 edellä).

Koordinaattiavaruus $\mathbb{R}^2$ voidaan yhtä hyvin liittää myös siihen alkuperäiseen pisteavaruuteen E^2, josta koko vektoriajatus oli lähtöisin. Nimittäin jos vektoreita ajatellaan E^2:n suuntajanoina, 'joita voi siirrellä', niin voidaan yhtä hyvin ajatella, että jokainen vektori vastaa yksikäsitteistä suuntajanaa, jonka lähtöpiste on kiinteä piste O. Näin on synnytetty E^2:n ja V:n kääntäen yksikäsitteinen vastaavuus:

$$P \in E^2 \;\leftrightarrow\; \overrightarrow{OP} \in V.$$

Kun nyt tulkitaan myös edellä valitut kantavektorit $\vec{a}, \vec{b}$ pisteestä O alkaviksi suuntajanoiksi, niin euklidiseen tasoon on luotu *koordinaatisto*†, jota merkitään $\{O, \vec{a}, \vec{b}\}$. Pistettä O sanotaan koordinaatiston *origoksi.*

†Ajatuksen koordinaatiston avulla tapahtuvasta geometrian aritmetisoimisesta toi matematiikkaan ranskalainen filosofi-matemaatikko **René Descartes** (1596-1650) tutkielmassaan "La géométrie", joka ilmestyi laajemman filosofisen teoksen liitteenä vuonna 1637. Tutkielma merkitsi *analyyttisen geometrian* alkua, ja enteili yleisemminkin geometrian 'algebralisoitumista' — trendiä, joka erityisesti tietokoneiden aikakaudella on entisestään vahvistunut.

Jos $P \in E^2$ ja $\overrightarrow{OP} = x\vec{a} + y\vec{b}$, niin sanotaan, että *pisteen P koordinaatit* valitussa koordinaatistossa $\{O, \vec{a}, \vec{b}\}$ ovat (x, y). Koordinaatiston ollessa kiinitetty voidaan koordinaattiparia (x, y) haluttaessa pitää P:n 'nimenä', jolloin on lupa kirjoittaa muitta mutkitta

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

Näin on luotu kääntäen yksikäsitteinen vastaavuus $E^2 \leftrightarrow \mathbb{R}^2$.

Esimerkki 2 Olkoon T tason (aito) kolmio, jonka kärkikipisteet ovat $0, A, B$. Merkitään $\vec{a} = \overrightarrow{OA}$, $\vec{b} = \overrightarrow{OB}$. Tällöin T:n kärkipisteet, T:n sivujen keskipisteet ja T:n keskiö koordinaattistossa $\{O, \vec{a}, \vec{b}\}$ ovat (vrt. Esimerkki 1)

kärkipisteet: $O = (0, 0)$, $A = (1, 0)$, $B = (0, 1)$,
sivujen keskipisteet: $O' = (1/2, 1/2)$, $A' = (0, 1/2)$, $B' = (1/2, 0)$,
keskiö: $C = (1/3, 1/3)$. $\square$

Koordinaatistoon tarvitaan siis ensinnäkin referenssipiste O. Tästä nk. origosta valitaan kaksi suuntaa, jotka määräävät vektoreiden $\vec{a}, \vec{b}$ suunnat. Vielä päätetään, että mitattaessa etäisyyttä O:sta on pituusyksikkö α mentäessä $\vec{a}$:n suuntaan ja β mentäessä $\vec{b}$:n suuntaan. Tässä $\alpha, \beta \in \mathbb{R}$ ja $\alpha, \beta > 0$. Kun nyt valitaan $\vec{a}$ ja $\vec{b}$ siten, että $|\vec{a}| = \alpha$ ja $|\vec{b}| = \beta$, ovat $\vec{a}$ ja $\vec{b}$ yksikäsitteisesti määrätyt ja koordinaatisto siis valmis. Origon kautta kulkevia suoria

$$S_1 = \{P \in E^2 \mid \overrightarrow{OP} = \lambda\vec{a},\ \lambda \in \mathbb{R}\},$$
$$S_2 = \{P \in E^2 \mid \overrightarrow{OP} = \mu\vec{b},\ \mu \in \mathbb{R}\},$$

joiden *suuntavektoreina* ovat $\vec{a}, \vec{b}$, sanotaan *koordinaattiakseleiksi.*

Lineaarinen riippumattomuus

Koordinaatiston kantavektoreista $\vec{a}, \vec{b}$ edellä tehdyt oletukset ($\vec{a}, \vec{b} \neq \vec{0}$ ja $\vec{a}, \vec{b}$ eivät saman- tai vastakkaissuuntaiset) voidaan asettaa lyhyemmin ehtona

$$x\vec{a} + y\vec{b} = \vec{0} \quad \Rightarrow \quad x = y = 0. \tag{$\star$}$$

Jos tason vektoreilla on ominaisuus ($\star$), niin sanotaan, että $\{\vec{a}, \vec{b}\}$ on *lineaarisesti riippumaton* (vektori)*systeemi* (= joukko) tai että $\vec{a}$ ja $\vec{b}$ ovat lineaarisesti riippumattomat. Tason vektoriavaruuden V kannaksi kelpaa siis mikä tahansa lineaarisesti riippumaton vektorisysteemi $\{\vec{a}, \vec{b}\}$.

Jos vektorit $\vec{a}, \vec{b}$ eivät ole lineaarisesti riippumattomat, niin ne ovat *lineaarisesti riippuvat.* Ehdon ($\star$) mukaisesti näin on, jos

$$x\vec{a} + y\vec{b} = \vec{0} \quad \text{jollakin } (x, y) \neq (0, 0).$$

Tämä ehto puolestaan toteutuu täsmälleen kun joko (i) $\vec{a} = \vec{0} \Rightarrow x\vec{a} + 0\vec{b} = \vec{0}$ $\forall x \in \mathbb{R}$, (ii) $\vec{b} = \vec{0} \Rightarrow 0\vec{a} + y\vec{b} = \vec{0}$ $\forall y \in \mathbb{R}$, tai (iii) $\vec{a}$ ja $\vec{b}$ ovat yhdensuuntaiset (= saman- tai vastakkaissuuntaiset, ol. $\vec{a} \neq \vec{0}$ ja $\vec{b} \neq \vec{0}$), jolloin $\exists \lambda \in \mathbb{R},\ \lambda \neq 0$ siten, että $\vec{a} + \lambda\vec{b} = \vec{0}$.

Esimerkki 3 Tason vektoreista $\vec{a}, \vec{b}$ tiedetään, että $\vec{a} - \vec{b}$ ja $\vec{a} + 2\vec{b}$ ovat lineaarisesti riippuvat. Voidaanko päätellä, että myös $\vec{a}$ ja $\vec{b}$ ovat lineaarisesti riippuvat?

<u>Ratkaisu</u> Annetun tiedon mukaan on jollakin $(x, y) \neq (0, 0)$

$$x(\vec{a} - \vec{b}) + y(\vec{a} + 2\vec{b}) = \vec{0}.$$

Vektorialgebran säännöillä tämä saadaan muotoon

$$\vec{0} = (x + y)\vec{a} + (-x + 2y)\vec{b} = x'\vec{a} + y'\vec{b}.$$

Koska

$$\begin{cases} x' = 0 \\ y' = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x + y = 0 \\ -x + 2y = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = 0 \\ y = 0 \end{cases}$$

ja tiedetään, että $(x, y) \neq (0, 0)$, niin päätellään, että $(x', y') \neq (0, 0)$. Koska siis jollakin $(x', y') \neq (0, 0)$ on $x'\vec{a} + y'\vec{b} = \vec{0}$, niin vastaus on: Voidaan! $\square$

Koordinaatiston vaihto

Kun tason geometrisia tehtäviä ratkotaan vektorialgebran keinoin, on tehtävään sopivan koordinaatiston valinta yleensä ratkaisemisen ensimmäinen askel (vrt. Esimerkki 1 edellä). Jos valittu koordinaatisto osoittautuu ratkaisun kuluessa epämukavaksi, voidaan suorittaa *koordinaatiston vaihto*. Koordinaatiston vaihto koostuu *origon siirrosta* ja vektoriavaruuden *kannan vaihdosta*, tai pelkästään jommasta kummasta. Kuvassa pisteen P koordinaatit ovat (x, y) koordinaatistossa $\{O, \vec{a}, \vec{b}\}$ ja (x', y') koordinaatistossa $\{O', \vec{c}, \vec{d}\}$.

Vektorien yhteenlaskudiagrammin mukaan on

$$\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P} \quad \Leftrightarrow \quad x\vec{a} + y\vec{b} = \overrightarrow{OO'} + (x'\vec{c} + y'\vec{d}).$$

Olkoon tässä $\overrightarrow{OO'} = \alpha\,\vec{a} + \beta\,\vec{b}$, eli $O' = (\alpha, \beta)$ koordinaatistossa $\{O, \vec{a}, \vec{b}\}$, ja

$$\vec{c} = \lambda_1\vec{a} + \mu_1\vec{b}, \quad \vec{d} = \lambda_2\vec{a} + \mu_2\vec{d},$$

eli vektoreiden $\vec{c}, \vec{d}$ koordinaatit kannassa $\{\vec{a}, \vec{b}\}$ ovat (λ_1, μ_1) ja (λ_2, μ_2). Tällöin em. yhtälö sievenee vektorialgebran säännöillä muotoon

$$(x - \lambda_1 x' - \lambda_2 y' - \alpha)\,\vec{a} + (y - \mu_1 x' - \mu_2 y' - \beta)\,\vec{b} = \vec{0}.$$

Koska vektorit $\vec{a}$ ja $\vec{b}$ ovat lineaarisesti riippumattomat, niin seuraa

$$\begin{cases} x - \lambda_1 x' - \lambda_2 y' - \alpha = 0 \\ y - \mu_1 x' - \mu_2 y' - \beta = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \lambda_1 x' + \lambda_2 y' = x - \alpha \\ \mu_1 x' + \mu_2 y' = y - \beta \end{cases}$$

Ratkaisemalla tästä (x', y') koordinaattien (x, y) avulla — yhtälöryhmä ratkeaa aina kun $\vec{c}$ ja $\vec{d}$ ovat lineaarisesti riippumattomat — saadaan selville *koordinaattimuunnoksen* $(x, y) \hookrightarrow (x', y')$ laskukaavat. Yhtälöryhmästä nähdään myös suoraan, miten käänteinen muunnos $(x', y') \hookrightarrow (x, y)$ on laskettava.

Esimerkki 4 Olkoon

$$\overrightarrow{OO'} = \vec{a} - \vec{b}, \quad \vec{c} = 2\vec{a} + \vec{b}, \quad \vec{d} = \vec{a} - 2\vec{b}.$$

Em. laskun kulku on tällöin

$$\begin{aligned} x\vec{a} + y\vec{b} \;&=\; (\vec{a} - \vec{b}\,) + x'(2\vec{a} + \vec{b}\,) + y'(\vec{a} - 2\vec{b}\,) \\ \Leftrightarrow \quad & (x - 2x' - y' - 1)\,\vec{a} + (y - x' + 2y' + 1)\,\vec{b} = \vec{0} \\ \Leftrightarrow \quad & \begin{cases} 2x' + y' = x - 1 \\ x' - 2y' = y + 1 \end{cases} \end{aligned}$$

Ratkaisemalla saadaan koordinaattimuunnoksen laskukaavoiksi

$$\begin{cases} x = 2x' + y' + 1 \\ y = x' - 2y' - 1 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x' = \frac{1}{5}(2x + y - 1) \\ y' = \frac{1}{5}(x - 2y - 3) \end{cases} \qquad \square$$

Dimensio. Aliavaruus

Koska tason vektoriavaruuden kannassa on oltava kaksi lineaarisesti riippumatonta vektoria, niin sanotaan, että V on 2-*ulotteinen* vektoriavaruus tai että V:n *dimensio* (ulotteisuus) on 2. Merkitään

$$\dim V = 2.$$

Myös koodinaattiavaruus $\mathbb{R}^2$ on vektoriavaruutena 2-ulotteinen. Nimittäin jos merkitään

$$\mathbf{a} = (1, 0), \quad \mathbf{b} = (0, 1),$$

niin $\mathbb{R}^2$:n vektorialgebran mukaan on $(x, y) = x\mathbf{a} + y\mathbf{b}$, eli jokainen $\mathbf{v} = (x, y) \in \mathbb{R}^2$ voidaan esittää yksikäsitteisesti $\mathbf{a}$:n ja $\mathbf{b}$:n lineaariyhdistelynä. Siis $\{\mathbf{a}, \mathbf{b}\}$ on $\mathbb{R}^2$:n kanta, ja koska kannassa on kaksi vektoria, niin $\dim(\mathbb{R}^2) = 2$.

Tason tai $\mathbb{R}^2$:n vektoreista voidaan muodostaa myös 1-*ulotteisia* vektoriavaruuksia. Tason vektoreiden tapauksessa nämä ovat kaikki ilmaistavissa jonkin vektorin $\vec{a} \in V,\ \vec{a} \neq \vec{0}$ avulla muodossa

$$W = \{\vec{v} = x\vec{a},\ x \in \mathbb{R}\}.$$

Koska ilmeisesti pätee

$$\begin{aligned} \vec{v}_1, \vec{v}_2 \in W \quad &\Rightarrow \quad \vec{v}_1 + \vec{v}_2 \in W, \\ \vec{v} \in W \quad &\Rightarrow \quad \lambda\vec{v} \in W \quad \forall\, \lambda \in \mathbb{R}, \end{aligned}$$

on W itsekin vektoriavaruus. Sen kantaan tarvitaan vain yksi vektori, esim $\vec{a}$, joten dim $W = 1$. Koska myös $W \subset V$, sanotaan, että W on V:n (aito) *aliavaruus* (engl. subspace) ja että $\vec{a}$ *virittää* (engl. span) W:n. Aliavaruuden W geometrinen vastine on origon kautta kulkeva suora $S \subset E^2$. Tämä on 1-*ulotteinen pisteavaruus.*

Vastaavuus

$$P \in S \;\leftrightarrow\; \overrightarrow{OP} = x\vec{a} \in W \;\leftrightarrow\; x \in \mathbb{R}$$

synnyttää kääntäen yksikäsitteisen vastaavuuden $\mathbb{R}$:n ja pisteavaruuden S välille. Lukujen geometrisointi lukusuoran pisteiksi (vrt. Luku II.1) perustui juuri tähän vastaavuuteen.

HARJOITUSTEHTÄVIÄ

1. a) Perustele vektorialgebran säännöt (V4)–(V6) geometrisesti.

   b) Vektoreille pätee kolmioepäyhtälö $|\vec{a}+\vec{b}| \leq |\vec{a}|+|\vec{b}|$. Päättele geometriaan enempää turvautumatta, että pätee myös $|\vec{a}| - |\vec{b}| \leq |\vec{a}+\vec{b}|$.

2. Tason vektoreista $\vec{a}, \vec{b}$ oletetaan, että $\vec{a}, \vec{b} \neq \vec{0}$ ja että $\vec{a}$ ja $\vec{b}$ eivät ole yhdensuuntaiset. Millä vektorin $\vec{c}$ arvoilla voidaan vektoreita $\vec{a}+\vec{b}$, $\vec{a}+\vec{c}$ ja $\vec{b}+2\vec{c}$ 'siirtelemällä' muodostaa kolmio?

3. Todista Harjoitustehtävän II.1:2b väittämä vektorialgebran avulla.

4. Nelikulmiossa $ABCD$ on $\overrightarrow{AB} = \vec{a}$, $\overrightarrow{AD} = \vec{b}$ ja $\overrightarrow{BC} = (1/2)(\vec{a}+\vec{b})$. Laske vektorien $\vec{a}$ ja $\vec{b}$ avulla vektori $\overrightarrow{AE}$, missä E on nelikulmion lävistäjien leikkauspiste.

5. Suunnikkaassa ABCD kärki A yhdistetään sivun CD keskipisteeseen P ja kärki B sivun AD keskipisteeseen R. Yhdysjanat leikatkoot pisteessä X. Lausu vektori $\overrightarrow{AX}$ vektoreiden $\vec{u} = \overrightarrow{AB}$ ja $\vec{v} = \overrightarrow{AD}$ avulla.

6. Kolmiossa ABC merkitään $\vec{a} = \overrightarrow{AB}$, $\vec{b} = \overrightarrow{AC}$. a) Päättele geometrisesti, että vektori $\vec{c} = |\vec{b}|\vec{a} + |\vec{a}|\vec{b}$ puolittaa kulman BAC. b) Piste D on janalla BC ja jana AD puolittaa kulman BAC. Todista vektorilaskulla *kulmanpuolittajalause*: Janojen BD ja DC pituuksien suhde $= |\vec{a}|/|\vec{b}|$.

7. Olkoot pisteet M ja N kolmioiden ABC ja DEF keskiöt. Näytä, että $\overrightarrow{AD} + \overrightarrow{BE} + \overrightarrow{CF} = 3\overrightarrow{MN}$.

8. Kolmion ABC sivut BC, CA ja AB jakautuvat pisteissä A^*, B^* ja C^* suhteessa $m : n$. Todista, että kolmioiden ABC ja $A^*B^*C^*$ keskiöt yhtyvät.

9. Olkoon $\{\vec{a}, \vec{b}\}$ tason vektoriavaruuden kanta ja olkoon $\vec{u} = 2\vec{a} + 3\vec{b}$, $\vec{v} = -3\vec{a} + 2\vec{b}$ ja $\vec{w} = -\vec{a} - 2\vec{b}$. Määritä vektorin $2\vec{u} - \vec{v} + \vec{w}$ koordinaatit kannassa $\{\vec{a}, \vec{b}\}$. Piirrä kuva!

10. Vektorit $\vec{a}, \vec{b}$ muodostavat tason vektoriavaruuden V kannan. Näytä, että myös vektorit $\vec{u} = \vec{a} + \vec{b}$, $\vec{v} = \vec{a} - 2\vec{b}$ muodostavat kannan. Laske vektorin $\vec{w} = \vec{a} - \vec{b}$ koordinaatit tässä kannassa.

11. Kolmiossa ABC piste O puolittaa sivun AB, piste E_1 jakaa sivun BC suhteessa $1 : 2$ ja piste E_2 sivun CA suhteessa $1 : 3$. Määritä kolmion kärkipisteiden koordinaatit siinä koordinaatistossa, jonka origo on piste O ja kantavektorit ovat $\overrightarrow{OE_1}$ ja $\overrightarrow{OE_2}$.

12. Jos $\vec{a}, \vec{b}$ ovat lineaarisesti riippumattomat tason vektorit, niin millä t:n arvoilla ($t \in \mathbb{R}$) vektorit $\vec{a} - t\vec{b}$ ja $t\vec{a} - 2\vec{b}$ ovat myös lineaarisesti riippumattomat?

13. Pisteen P koordinaatit ovat (x, y) koordinaatistossa $\{O, \vec{a}, \vec{b}\}$ ja (x', y') koordinaatistossa $\{O', \vec{c}, \vec{d}\}$. Johda koordinaattimuunnoksen laskukaavat molempiin suuntiin, kun
a) $O' = O$, $\vec{c} = \vec{a} + \vec{b}$ ja $\vec{d} = \vec{a} - \vec{b}$.
b) $\overrightarrow{O'O} = \vec{c} + 2\vec{d}$, $\vec{c} = 2\vec{a} + 3\vec{b}$ ja $\vec{d} = -\vec{a} + 2\vec{b}$.

14. Olkoon $\vec{a} = \overrightarrow{OA}$ ja $\vec{b} = \overrightarrow{OB}$. Johda kordinaattimuunnoksen laskukaavat koordinaatistosta $\{O, \vec{a}, \vec{b}\}$ koordinaatistoon $\{O', \vec{c}, \vec{d}\}$, kun O' = kolmion OAB keskiö ja $\vec{c} = \overrightarrow{O'A}$, $\vec{d} = \overrightarrow{O'B}$. Mitkä ovat kolmion kärkipisteiden ja sivujen keskipisteiden koordinaatit jälkimmäisessä koordinaatistossa?

II.3 Skalaaritulo

Olkoot $\vec{a} \neq \vec{0}$ ja $\vec{b} \neq \vec{0}$ kaksi tason vektoria. Näiden kanssa samansuuntaiset *yksikkövektorit* (yksikön pituiset vektorit) ovat

$$\vec{a}^{\,0} = \frac{1}{|\vec{a}|}\,\vec{a} = \frac{\vec{a}}{|\vec{a}|}\,, \qquad \vec{b}^{\,0} = \frac{1}{|\vec{b}|}\,\vec{b} = \frac{\vec{b}}{|\vec{b}|}\,.$$

Liitetään jatkossa vektoreihin $\vec{a}, \vec{b}$ luku $\gamma(\vec{a}, \vec{b}) \in \mathbb{R}$, joka riippuu vain $\vec{a}$:n ja $\vec{b}$:n suunnista, eli vain yksikkövektoreista $\vec{a}^{\,0}, \vec{b}^{\,0}$. Liittäminen tapahtuu geometrisella konstruktiolla seuraavasti: Olkoon $\vec{a}^{\,0} = \overrightarrow{OA}$, $\vec{b}^{\,0} = \overrightarrow{OB}$ ja valitaan pisteiden O ja A kautta kulkevalta suoralta piste B', joka on lähinnä pistettä B. Tällöin kulma $\measuredangle OB'B$ (tai $\measuredangle AB'B$, jos $B' = O$) on suora.

Kuvioon liittyen asetetaan:

$$\gamma(\vec{a}, \vec{b}) = \gamma(\vec{a}^{\,0}, \vec{b}^{\,0}) = \begin{cases} |\overrightarrow{OB'}|, & \text{jos } \overrightarrow{OB'} \uparrow\uparrow \vec{a}^{\,0}, \\ 0, & \text{jos } \overrightarrow{OB'} = \vec{0}, \\ -|\overrightarrow{OB'}|, & \text{jos } \overrightarrow{OB'} \uparrow\downarrow \vec{a}^{\,0}. \end{cases}$$

Luku $\gamma(\vec{a}, \vec{b})$ määräytyy siis geometrisesti janan pituutena tai sen vastalukuna, kun tunnetaan vektorit $\vec{a}$ ja $\vec{b}$. Konstruktio on ilmeisen symmetrinen näiden vektorien suhteen, ts. $\gamma(\vec{a}, \vec{b}) = \gamma(\vec{b}, \vec{a})$. Ilmeistä on myös, että $\gamma(\vec{a}, \vec{b})$ ei muutu, jos konstruktioon liittyvää geometrista kuviota muutetaan euklidisella liikkeellä. Tästä voidaan päätellä, että $\gamma(\vec{a}, \vec{b})$ riippuu vain vektoreiden $\vec{a}, \vec{b}$ suuntien määräämän <u>sisäkulman</u> <u>mittaluvusta</u>. Konstruktiosta nähdään, että sama pätee myös kääntäen: Luku $\gamma(\vec{a}, \vec{b})$ määrää minitun mittaluvun yksikäsitteisesti.

Jatkossa vektorien $\vec{a}, \vec{b}$ suuntien määräämää kulmaa merkitään symbolilla $\measuredangle(\vec{a}, \vec{b})$.[†]

[†]Merkinnässä $\measuredangle(\vec{a}, \vec{b})$ ei ole tarpeen tehdä eroa sisä- ja ulkokulman välillä, ts. kulmaa ei tarvitse tulkita sektoriksi. Vrt. Luku II.1.

Lukua $\gamma(\vec{a},\vec{b})$ sanotaan tästä lähtien *kulman* $\measuredangle(\vec{a},\vec{b})$ *kosiniksi* ja merkitään

$$\boxed{\gamma(\vec{a},\vec{b}) = \cos\measuredangle(\vec{a},\vec{b}).}$$

Koska luku $\cos\measuredangle(\vec{a},\vec{b})$ määrää kulmaan $\measuredangle(\vec{a},\vec{b})$ liittyen sisäkulman mittaluvun, niin se määrää myös kulman geometrisen 'ulkonäön'. Erityisesti pätee

$$\begin{array}{lcll} \cos\measuredangle(\vec{a},\vec{b}) = 1 & \Leftrightarrow & \vec{a}\uparrow\uparrow\vec{b} & \text{(nollakulma)}, \\ \cos\measuredangle(\vec{a},\vec{b}) = -1 & \Leftrightarrow & \vec{a}\uparrow\downarrow\vec{b} & \text{(oikokulma)}, \\ \cos\measuredangle(\vec{a},\vec{b}) = 0 & \Leftrightarrow & \vec{a}\perp\vec{b} & \text{(suora kulma)}. \end{array}$$

Tässä on käytetty merkintää $\vec{a}\perp\vec{b}$ ilmaisemaan, että vektorit $\vec{a}$ ja $\vec{b}$ ovat *kohtisuorat* eli *ortogonaaliset* ($\measuredangle(\vec{a},\vec{b})$ = suora kulma). Luvun $\cos\measuredangle(\vec{a},\vec{b})$ määritelmän ja Pythagoraan lauseen perusteella on kaikissa tapauksissa

$$-1 \le \cos\measuredangle(\vec{a},\vec{b}) \le 1. \tag{1}$$

Määritelmästä nähdään myös, että pätee

$$\cos\measuredangle(\lambda\vec{a},\vec{b}) = \cos\measuredangle(\vec{a},\lambda\vec{b}) = \begin{cases} \cos\measuredangle(\vec{a},\vec{b}), & \text{jos } \lambda > 0, \\ -\cos\measuredangle(\vec{a},\vec{b}), & \text{jos } \lambda < 0. \end{cases} \tag{2}$$

MÄÄRITELMÄ II.3.1 (**Skalaaritulo**) Vektorien $\vec{a},\vec{b}$ *skalaaritulo* eli *pistetulo* on reaaliluku, joka merkitään $\vec{a}\cdot\vec{b}$, luetaan 'a piste b', ja määritellään

$$\vec{a}\cdot\vec{b} = \begin{cases} |\vec{a}||\vec{b}|\cos\measuredangle(\vec{a},\vec{b}), & \text{jos } \vec{a}\neq\vec{0} \text{ ja } \vec{b}\neq\vec{0}, \\ 0, & \text{jos } \vec{a}=\vec{0} \text{ tai } \vec{b}=\vec{0}. \end{cases}$$

Määritelmästä seuraa ensinnäkin

$$\boxed{\vec{a}\cdot\vec{a} = |\vec{a}|^2.} \tag{3}$$

Toiseksi on voimassa symmetrialaki

$$\boxed{\vec{a}\cdot\vec{b} = \vec{b}\cdot\vec{a}.} \tag{4}$$

Kolmanneksi seuraa määritelmästä ja ominaisuudesta (2) skalaarilla kertomisen ja pistetulon välinen osittelulaki

$$\boxed{(\lambda\vec{a})\cdot\vec{b} = \vec{a}\cdot(\lambda\vec{b}) = \lambda(\vec{a}\cdot\vec{b}), \quad \lambda\in\mathbb{R}.} \tag{5}$$

Viimeksi mainitun lain perusteella sulkeiden pois jättäminen merkinnässä $\lambda\vec{a}\cdot\vec{b}$ ei aiheuta sekaannusta.

Hieman vähemmän ilmeinen skalaaritulon määritelmän seuraamus on pistetulon ja vektorien yhteenlaskun välinen osittelulaki

$$\boxed{\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{c}.} \tag{6}$$

Tämän todistamiseksi tarkastellaan alla olevia kuvioita.

Kuvio 1:

Kuvio 2:

$$|\vec{a}||\overrightarrow{OC'}| = |\vec{a}||\overrightarrow{OB'}| + |\vec{a}||\overrightarrow{B'C'}|$$
$$\Rightarrow\ \vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{c}$$

$$|\vec{a}||\overrightarrow{OC'}| = |\vec{a}||\overrightarrow{OB'}| - |\vec{a}||\overrightarrow{B'C'}|$$
$$\Rightarrow\ \vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{c}$$

Kuviossa 1 on $|\overrightarrow{OB'}| = |\cos\measuredangle(\vec{a},\vec{b})||\vec{b}|$ kolmioiden yhdenmuotoisuuden perusteella, jolloin skalaaritulon määritelmästä seuraa

$$\vec{a}\cdot\vec{b} = \begin{cases} |\vec{a}||\overrightarrow{OB'}|, & \text{jos } \overrightarrow{OB'} \uparrow\uparrow \vec{a}, \\ -|\vec{a}||\overrightarrow{OB'}|, & \text{jos } \overrightarrow{OB'} \uparrow\downarrow \vec{a}. \end{cases}$$

Perustuen tähän geometriseen tulkintaan on Kuviossa 2 johdettu osittelulaki (6) tapauksessa $\cos\measuredangle(\vec{a},\vec{b}) > 0$ ja $\cos\measuredangle(\vec{a},\vec{b}+\vec{c}) > 0$. Muissakin tapauksissa ($\cos\measuredangle(\vec{a},\vec{b}) \le 0$ ja/tai $\cos\measuredangle(\vec{a},\vec{b}+\vec{c}) \le 0$) voidaan päättellä vastaavalla tavalla, että osittelulaki (6) on voimassa.

Skalaaritulon ominaisuuksia (4), (5), (6) yhdistelemällä saadaan nk. *bilineaarisuusominaisuudet*

$$\boxed{\begin{aligned} (\lambda\vec{a}+\mu\vec{b})\cdot\vec{c} &= \lambda\,\vec{a}\cdot\vec{c}+\mu\,\vec{b}\cdot\vec{c}, \\ \vec{a}\cdot(\lambda\vec{b}+\mu\vec{c}) &= \lambda\,\vec{a}\cdot\vec{b}+\mu\,\vec{a}\cdot\vec{c}, \quad \lambda,\mu\in\mathbb{R}. \end{aligned}} \tag{7}$$

Skalaaritulo, joka edellä on määritelty geometristen ideoiden pohjalta, on itse asiassa hyvin yleinen vektoriavaruuksiin liittyvä käsite. Yleisemmissä yhteyksissä ei skalaaritulolle oleteta lähtökohtaisesti muita ominaisuuksia kuin (i) *symmetrisyys*, eli symmetrialakia (4) vastaava ominaisuus, (ii) *bilineaarisuus*, eli laskusääntöjen (7) vastineet, ja (iii) *positiividefiniittisyys*, jonka muotoilu tason vektoreille on

$$\vec{a}\cdot\vec{a} \ge 0 \ \forall\vec{a} \quad \text{ja} \quad \vec{a}\cdot\vec{a} = 0 \Rightarrow \vec{a} = \vec{0}.$$

Tämä on ilmeinen em. geometrisen määritelmän perusteella (ominaisuus (3)). Määritelmästä ja ominaisuudesta (1) seuraa myös, että tason vektoreiden skalaaritulolle pätee epäyhtälö

$$\boxed{|\vec{a}\cdot\vec{b}| \le |\vec{a}\,||\vec{b}\,|.} \tag{8}$$

Tällä epäyhtälöllä on yleisempiä — vähemmän ilmeisiä — algerallisia ulottuvuuksia, joita tarkastellaan lähemmin seuraavassa luvussa.

Esimerkki 1 Tason vektoreista $\vec{a},\vec{b}$ ja $\vec{u}$ tiedetään

$$|\vec{a}| = 1, \quad |\vec{b}| = 3, \quad \vec{a}\cdot\vec{b} = 2, \quad \vec{a}\cdot\vec{u} = -2, \quad \vec{b}\cdot\vec{u} = 1.$$

Määritä $\vec{u}$:n koordinaatit kannassa $\{\vec{a},\vec{b}\}$.

<u>Ratkaisu</u> Jos merkitään $\vec{u} = x\vec{a}+y\vec{b}$, niin kertomalla skalaarisesti $\vec{a}$:lla ja $\vec{b}$:lla ja käyttämällä bilineaarisuussääntöjä (7) ja sääntöjä (3),(4) saadaan

$$x\vec{a}+y\vec{b} = \vec{u} \quad \Rightarrow \quad \begin{cases} |\vec{a}|^2x + (\vec{a}\cdot\vec{b})y = \vec{a}\cdot\vec{u} \\ (\vec{a}\cdot\vec{b})x + |\vec{b}|^2y = \vec{b}\cdot\vec{u} \end{cases}$$

Siis

$$\begin{cases} x + 2y = -2 \\ 2x + 9y = 1 \end{cases} \quad \Rightarrow \quad \begin{cases} x = -4 \\ y = 1 \end{cases} \qquad \square$$

Esimerkki 1 (jatko) Jaa esimerkin vektori $\vec{u}$ kahteen komponenttiin siten, että toinen komponentti on $\vec{a}$:n suuntainen ja toinen on kohtisuora a) $\vec{a}$:ta, b) $\vec{b}$:ta vastaan.

<u>Ratkaisu</u> Halutaan laskea $\vec{u}_1, \vec{u}_2$ siten, että $\vec{u} = \vec{u}_1 + \vec{u}_2,\ \vec{u}_1 = t\vec{a},\ t \in \mathbb{R}$ ja

$$\text{a)}\ \ \vec{u}_2 \cdot \vec{a} = (\vec{u} - t\vec{a}) \cdot \vec{a} = 0, \qquad \text{b)}\ \ \vec{u}_2 \cdot \vec{b} = (\vec{u} - t\vec{a}) \cdot \vec{b} = 0.$$

Esimerkin tiedoin saadaan

$$\text{a)}\ \ t = \frac{\vec{a} \cdot \vec{u}}{|\vec{a}|^2} = -2, \qquad \text{b)}\ \ t = \frac{\vec{b} \cdot \vec{u}}{\vec{a} \cdot \vec{b}} = \frac{1}{2}\,.$$

Siis $\vec{u} = \vec{u}_1 + \vec{u}_2$, missä

$$\begin{aligned} &\text{a)}\ \ \vec{u}_1 = -2\vec{a}, \quad \vec{u}_2 = \vec{u} - \vec{u}_1 = (-4\vec{a} + \vec{b}) + 2\vec{a} = -2\vec{a} + \vec{b}, \\ &\text{b)}\ \ \vec{u}_1 = \tfrac{1}{2}\vec{a}, \quad \vec{u}_2 = \vec{u} - \vec{u}_1 = (-4\vec{a} + \vec{b}) - \tfrac{1}{2}\vec{a} = -\tfrac{9}{2}\vec{a} + \vec{b}. \quad \square \end{aligned}$$

Esimerkin a)-kohdassa on kyse perustehtävästä, jossa annettu vektori $\vec{u} \in V$ halutaan jakaa kahteen komponenttiin $\vec{u}_1$, $\vec{u}_2$ siten, että $\vec{u}_1$ on annetun vektorin $\vec{a}$ suuntainen ja $\vec{u}_2 \perp \vec{u}_1$. Tällöin sanotaan, että $\vec{u}_1$ on $\vec{u}$:n *ortogonaaliprojektio* $\vec{a}$:n *suuntaan* eli $\vec{a}$:n virittämään V:n yksiulotteiseen aliavaruuteen

$$W = \{\vec{v} = \lambda\vec{a} \mid \lambda \in \mathbb{R}\}.$$

Euklidisessa tasossa aliavaruutta W vastaa origon kautta kulkeva suora.

Ortonormeerattu kanta

Luvussa II.2 konstruoitiin avaruudelle V kanta käyttäen kahta lineaarisesti riippumatonta vektoria. Skalaaritulon tultua määritellyksi saadaan käytännön laskut huomattavasti yksinkertaistumaan valitsemalla kantavektorit niin, että ne ovat ortogonaaliset, ts. kantavektorien välinen skalaaritulo $= 0$. Jos vielä kantavektorit *normeerataan* yksikkövektoreiksi, niin sanotaan, että näin saatu kanta on *ortonormeerattu* (muussa tapauksessa vain *ortogonaalinen*). Ortonormeerattu kanta $\{\vec{i}, \vec{j}\}$ on siis sellainen, että pätee

$$\begin{cases} \vec{i} \cdot \vec{j} = 0 & \text{(ortogonaalisuusehto)}, \\ |\vec{i}| = |\vec{j}| = 1 & \text{(normeerausehto)}. \end{cases}$$

Nämä ehdot toteuttava $\{\vec{i}, \vec{j}\}$ on lineaarisesti riippumaton, sillä

$$x\vec{i} + y\vec{j} = \vec{0} \quad \Rightarrow \quad \begin{cases} x = \vec{i} \cdot (x\vec{i} + y\vec{j}) = \vec{i} \cdot \vec{0} = 0, \\ y = \vec{j} \cdot (x\vec{i} + y\vec{j}) = \vec{j} \cdot \vec{0} = 0. \end{cases}$$

Vastaavaa euklidisen tason koordinaatistoa sanotaan *karteesiseksi*† koordinaatistoksi.

Ortonormeeratussa kannassa $\{\vec{i}, \vec{j}\}$ annettujen vektorien skalaaritulo on helposti laskettavissa: Jos

$$\vec{v}_1 = x_1\vec{i} + y_1\vec{j}, \quad \vec{v}_2 = x_2\vec{i} + y_2\vec{j},$$

niin skalaaritulon bilineaarisuuden ja symmetrisyyden perusteella

$$\begin{aligned} \vec{v}_1 \cdot \vec{v}_2 &= (x_1\vec{i} + y_1\vec{j}) \cdot (x_2\vec{i} + y_2\vec{j}) \\ &= x_1x_2\,\vec{i} \cdot \vec{i} + (x_1y_2 + x_2y_1)\,\vec{i} \cdot \vec{j} + y_1y_2\,\vec{j} \cdot \vec{j}. \end{aligned}$$

†Termi on muotoutunut ranskalaisen **René Descartes**in latinankielisestä nimestä Cartesius, vrt. alaviite edellisessä luvussa.

Koska tässä on $\vec{i}\cdot\vec{i}=\vec{j}\cdot\vec{j}=1$ ja $\vec{i}\cdot\vec{j}=0$, niin laskukaavaksi tulee

$$\boxed{\vec{v}_1\cdot\vec{v}_2 = x_1x_2+y_1y_2.} \tag{9}$$

Tämän mukaan skalaaritulo määräytyy ortonormeeratussa kannassa laskukaaviolla (vrt. vektorien yhteenlaskun vastaava kaavio edellisessä luvussa)

$$\begin{array}{ccccc} \vec{v}_1 & \vec{v}_2 & & & \\ \downarrow & \downarrow & & & \\ (x_1,y_1) & (x_2,y_2) & \mapsto & (x_1x_2+y_1y_2) & = \vec{v}_1\cdot\vec{v}_2 \end{array}$$

Myös vektorin itseisarvon laskeminen käy ortonormeeratussa kannassa helposti, sillä laskukaavojen (3) ja (9) mukaan

$$\boxed{|\vec{v}\,|^2=\vec{v}\cdot\vec{v}=x^2+y^2, \quad \vec{v}=x\vec{i}+y\vec{j}.} \tag{10}$$

Esimerkki 2 Laske $\cos\measuredangle(\vec{a},\vec{b})$, kun $\vec{a}=4\vec{i}+3\vec{j}$ ja $\vec{b}=2\vec{-}3\vec{j}$.

<u>Ratkaisu</u> Määritelmän II.3.1 ja kaavojen (9)–(10) perusteella

$$\cos\measuredangle(\vec{a},\vec{b})=\frac{\vec{a}\cdot\vec{b}}{|\vec{a}\,||\vec{b}\,|}=\frac{4\cdot 2+3\cdot(-3)}{\sqrt{4^2+3^2}\sqrt{2^2+3^2}}=-\frac{1}{5\sqrt{13}}\,. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Laske
   a) $|4\vec{a}-5\vec{b}|$, kun $|\vec{a}|=1$, $|\vec{b}|=2$ ja $\vec{a}\cdot\vec{b}=-\frac{1}{3}|\vec{a}||\vec{b}|$
   b) $\vec{a}\cdot\vec{b}$, kun $|\vec{a}+3\vec{b}|=16$ ja $|\vec{a}-3\vec{b}|=2\sqrt{58}$

2. a) Kolmiossa ABC ovat sivujen AB, AC ja BC pituudet a, b ja c. Näytä skalaaritulon avulla, että pätee $c^2=a^2+b^2-2ab\cos\measuredangle BAC$.
   b) Nelikulmiossa $ABCD$ on $\cos\measuredangle BAD=\gamma$ ja sivujen AB, AD, BC ja CD pituudet ovat a, b, c ja d. Laske $\cos\measuredangle BCD$.

3. Kun $\vec{a}$ ja $\vec{b}$ ovat tason vektoreita ja $|\vec{a}|=|\vec{b}|$, saa skalaaritulo $(\vec{a}+\vec{b})\cdot(\vec{a}-\vec{b})$ yksinkertaisen muodon. Millaisen? Mitä tulos tarkoittaa geometrisesti, jos $\vec{a}=\overrightarrow{OA}$ ja $\vec{b}=\overrightarrow{OB}$, missä a) OA ja OB ovat suunnikkaan kaksi sivua, b) O on ympyrän keskipiste ja A, B sen kehän pisteitä? Kuva!

4. Suorakulmaisessa kolmiossa ABC on suoran kulman kärjestä lähtevä korkeusjana AD. Sivujen AB ja AC pituudet ovat 5 ja 12. Laske skalaaritulot $\overrightarrow{AB} \cdot \overrightarrow{DC}$, $\overrightarrow{BD} \cdot \overrightarrow{CA}$ ja $\overrightarrow{AC} \cdot \overrightarrow{CD}$.

5. Tason vektoreista $\vec{a}$ ja $\vec{b}$ tiedetään, että $|\vec{a}| = |\vec{b}| = 1$ ja $\vec{a} \cdot \vec{b} = \frac{1}{7}$. Lisäksi tiedetään vektorista $\vec{v}$, että $\vec{a} \cdot \vec{v} = 3$ ja $\vec{b} \cdot \vec{v} = -1$. Määritä $\vec{v}$:n koordinaatit kannassa $\{\vec{a}, \vec{b}\}$.

6. Tason koordinaatistossa $\{O, \vec{a}, \vec{b}\}$ ovat pisteen P koordinaatit $(2, 1)$ ja pisteen O' koordinaatit $(-1, 3)$. Määritä pisteen P koordinaatit koordinaatistossa $\{O', \vec{a} + \vec{b}, \vec{a} - 2\vec{b}\}$, kun tiedetään, että $|\vec{a}| = |\vec{b}| = 1$ ja $\vec{a} \cdot \vec{b} = -\frac{1}{3}$.

7. Laske $|\vec{a}|$, $|\vec{b}|$ ja $\cos \measuredangle(\vec{a}, \vec{b})$, kun
   a) $\vec{a} = \vec{i} - \vec{j}$, $\vec{b} = \vec{i} + 2\vec{j}$
   b) $\vec{a} = -2\vec{i} + 3\vec{j}$, $\vec{b} = 3\vec{i} - \vec{j}$
   c) $\vec{a} = -68\vec{i} + 51\vec{j}$, $\vec{b} = 3\vec{i} + 4\vec{j}$
   d) $\vec{a} = 76\vec{i} - 57\vec{j}$, $\vec{b} = 92\vec{i} + 69\vec{j}$
   e) $\vec{a} = (3\sqrt{2} + 5)\vec{i} + (5\sqrt{2} - 3)\vec{j}$, $\vec{b} = (5\sqrt{2} + 3)\vec{i} + (3\sqrt{2} - 5)\vec{j}$

8. Tason vektoreista $\vec{a}, \vec{b} \in V$ tiedetään, että $|\vec{a}| = 1$, $|\vec{b}| = 3$ ja $\vec{a} \cdot \vec{b} = -2$. Millä kertoimien λ, μ arvoilla $\{\vec{a}, \lambda\vec{a} + \mu\vec{b}\}$ on V:n ortonormeerattu kanta?

9. (*) Näytä skalaarituloa käyttäen, että kolmion korkeusjanat ovat kolmella saman pisteen kautta kulkevalla suoralla, ts. korkeusjanat tai niiden jatkeet leikkaavat samassa pisteessä.

10. (*) Joukko $S \subset E^2$ koostuu pisteistä P, joiden karteesiset koordinaatit toteuttavat ehdon
$$(x - 1)^2 - (y + 2)^2 = 1.$$
Olkoot (ξ, η) pisteen P koordinaatit toisessa ortonormeeratussa koordinaatistossa, jonka origo on pisteessä $(x, y) = (1, -2)$ ja kantavektorit
$$\vec{e}_1 = \frac{1}{\sqrt{2}}(\vec{i} + \vec{j}), \quad \vec{e}_2 = \frac{1}{\sqrt{2}}(\vec{i} - \vec{j}).$$
Jos pisteen $P = (x, y)$ koordinaatit tässä koordinaatistossa ovat (x', y'), niin millä koordinaateille x', y' asetettavalla ehdolla on $P = (x', y') \in S$? Piirrä kuva, jossa näkyvät molemmat koordinaatistot, ja hahmottele kuvaan joukko S.

II.4 *Abstrakti skalaaritulo ja normi

Oletetaan lähtökohdaksi annettu reaalikertoiminen vektoriavaruus $(U, \mathbb{R})$ ja siinä määritelty skalaaritulo eli *sisätulo*, jota merkitään symbolilla $\langle \cdot, \cdot \rangle$. U:n alkioita (vektoreita) merkitään $\mathbf{u}, \mathbf{v}$, jne. — Mitään hypoteeseja näiden 'ulkonäöstä' ei tehdä. Skalaaritulosta oletetaan, että $\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{R}$ on määritelty kaikille pareille $(\mathbf{u}, \mathbf{v})$, $\mathbf{u}, \mathbf{v} \in U$ (eli $(\mathbf{u}, \mathbf{v}) \in U \times U$, vrt. alaviite Luvussa II.2) ja että seuraavat lainalaisuudet ovat voimassa:

1. *Symmetrisyys*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \quad \forall\, \mathbf{u}, \mathbf{v} \in U.$$

2. *Bilineaarisuus*

   (a) $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$,

   (b) $\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{u}, \mathbf{w} \rangle$

   $$\forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in U,\ \alpha, \beta \in \mathbb{R}.$$

3. *Positiividefiniittisyys*

   (a) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \quad \forall\, \mathbf{u} \in U$,

   (b) $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \ \Leftrightarrow\ \mathbf{u} = \mathbf{0}$.

Tässä U:n nollavektoria (vektorien yhteenlaskun nolla-alkiota) on merkitty $\mathbf{0}$:lla.

Oletukset 1–3 ovat yleiset, reaalikertoimisen vektoriavaruuden *skalaaritulon aksioomat.*† Jos $\langle \cdot, \cdot \rangle$ on aksioomat 1–3 täyttävä U:n skalaaritulo, niin sanotaan, että ko. skalaarituloon liittyvä vektorin $\mathbf{u} \in U$ *normi* on

$$|\mathbf{u}| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}.$$

Esimerkki 1 Olkoon $U = (\mathbb{R}^2, \mathbb{R})$ tason vektoriavaruuden $(V, \mathbb{R})$ ortonormeerattua kantaa $\{\vec{i}, \vec{j}\}$ vastaava koordinaattiavaruus. Tällöin jos $\mathbf{u} = (x_1\,, y_1) \in U$ ja $\mathbf{v} = (x_2\,, y_2) \in U$, ja vastaavat tason vektorit ovat $\vec{u} = x_1\vec{i} + y_1\vec{j}$ ja $\vec{v} = x_2\vec{i} + y_2\vec{j}$, niin edellisen luvun päätelmien perusteella aksioomat 1–3 ovat voimassa, kun U:n skalaaritulo määritellään

$$\langle \mathbf{u}, \mathbf{v} \rangle = \vec{u} \cdot \vec{v} = x_1 x_2 + y_1 y_2. \quad \square$$

†Bilineaarisuusvaatimus (b) on aksioomana tarpeeton, koska se seuraa ehdosta (a) sekä symmetriasta. Positiividefiniittisyysaksiooman osa (b) riittää myös asettaa muodossa '$\Rightarrow$', sillä '$\Leftarrow$' seuraa bilineaarisuudesta.

Asetetaan esimerkkiin viitaten

MÄÄRITELMÄ II.4.1 Jos $\mathbf{u} = (x_1\,, y_1) \in \mathbb{R}^2$ ja $\mathbf{v} = (x_2\,, y_2) \in \mathbb{R}^2$, niin sanotaan, että skalaaritulo

$$\langle \mathbf{u}, \mathbf{v} \rangle = x_1 x_2 + y_1 y_2$$

on $\mathbb{R}^2$:n (vektoriavaruuden $(\mathbb{R}^2, \mathbb{R})$) *euklidinen skalaaritulo.* Normi

$$|\mathbf{u}| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = (x^2 + y^2)^{1/2}, \quad \mathbf{u} = (x, y) \in \mathbb{R}^2$$

on $\mathbb{R}^2$:n *euklidinen normi.*

ESIMERKKI 1 (jatko) Tason vektoreiden skalaaritulolle pätevä epäyhtälö $|\vec{u} \cdot \vec{v}\,| \leq |\vec{u}\,||\vec{v}\,|$ voidaan kirjoittaa $\mathbb{R}^2$:n euklidisen skalaaritulon ja normin avulla muodossa

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}| \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^2. \quad \square$$

Esimerkin epäyhtälö tarkoittaa auki kirjoitettunaa väittämää

$$|x_1 x_2 + y_1 y_2| \leq \sqrt{x_1^2 + y_1^2}\,\sqrt{x_2^2 + y_2^2} \quad \forall\, x_1,\, x_2,\, y_1,\, y_2 \in \mathbb{R}.$$

Tämä on siis tosi geometrisin perustein — mutta väittämänä tämä ei näytä lainkaan geometriasta riippuvalta. Seuraava lause osoittaa, että kyseessä ei olekaan geometrinen vaan <u>algebrallinen</u> väittämä: Epäyhtälö on erikoistapaus yleisemmästä, kaikille skalaarituloille pätevästä *Cauchyn* tai *Cauchyn–Schwarzin epäyhtälöstä.* Kyseessä on kolmioepäyhtälön ohella matemaattisen analyysin yleisimmin käytetty epäyhtälö.

LAUSE II.4.2 **(Cauchy–Schwarz)** Jokaiselle aksioomat 1–3 toteuttavalle, reaalikertoimisen vektoriavaruuden U skalaaritulolle pätee epäyhtälö

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|,$$

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

<u>Todistus</u> Jos $\mathbf{u} = \mathbf{0}$, niin $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle = 0$ bilineaarisuusehtojen nojalla, joten tässä tapauksessa väittämä on tosi muodossa $0 \leq 0$. Oletetaan siis, että $\mathbf{u} \neq \mathbf{0}$, jolloin aksiooman 3 mukaan on $\langle \mathbf{u}, \mathbf{u} \rangle > 0$. Tällöin saman aksiooman mukaan

$$\langle \beta \mathbf{u} + \mathbf{v}, \beta \mathbf{u} + \mathbf{v} \rangle \geq 0 \quad \forall \beta \in \mathbb{R}.$$

Aksioomien 1–2 perusteella tämä voidaan kirjoittaa yhtäpitävästi muotoon

$$\beta^2 \langle \mathbf{u}, \mathbf{u} \rangle + 2\beta \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \quad \forall \beta \in \mathbb{R}.$$

Valitsemalla tässä

$$\beta = -\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$$

seuraa

$$\begin{aligned} -\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}^2 + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \quad &\Leftrightarrow \quad \langle \mathbf{u}, \mathbf{v} \rangle^2 \ \leq \ \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \\ &\Leftrightarrow \quad |\langle \mathbf{u}, \mathbf{v} \rangle| \ \leq \ |\mathbf{u}||\mathbf{v}|. \quad \square \end{aligned}$$

Normi

Edellä kutsuttiin jo vektorin $\mathbf{u}$ <u>normiksi</u> skalaaritulon avulla määriteltyä reaalilukua $|\mathbf{u}| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$. Yleisemmin *vekoriavaruuden normi* ei edellytä skalaaritulon määrittelyä, vaan kyse on yleisemmästä tavasta määrätä vektorille 'pituus'. Yleisempää vektorin $\mathbf{u}$ normia merkitään symbolilla $\|\mathbf{u}\|$. Vektorin 'normittamisessa' on kyse laskuoperaatiosta (tai mittausoperaatiosta) $\mathbf{u} \in U \mapsto \|\mathbf{u}\| \in \mathbb{R}$, jonka on täytettävä seuraavat *normin aksioomat*:

1. *Positiividefiniittisyys*
   (a) $\|\mathbf{u}\| \geq 0 \quad \forall\, \mathbf{u} \in U$,
   (b) $\|\mathbf{u}\| = 0 \Rightarrow \mathbf{u} = \mathbf{0}$.
2. *Skaalautuvuus*
   $$\|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\| \quad \forall\, \mathbf{u} \in U,\ \alpha \in \mathbb{R}.$$
3. *Kolmioepäyhtälö*
   $$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| \quad \forall\, \mathbf{u}, \mathbf{v} \in U.$$

Valitsemalla $\alpha = 0$ aksioomassa 2 nähdään, että pätee $\mathbf{u} = \mathbf{0} \Rightarrow \|\mathbf{u}\| = 0$. Aksiooman 1b perusteella tämä imlikaatio pätee myös kääntäen, joten normille on voimassa

$$\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}.$$

Esimerkki 2 Skalaaritulon avulla määritellylle normille $\|\mathbf{u}\| = |\mathbf{u}| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$ aksioomien 1–2 voimassaolo on ilmeistä skalaaritulon positiividefiniittisyyden ja bilineaarisuuden perusteella (skalaaritulon aksioomat 2 ja 3). Myös kolmioepäyhtälö on voimassa (Harj.teht. 5a), joten kyseessä on todella normi. $\square$

Esimerkki 3 Avaruuden $\mathbb{R}^2$ nk. *maksiminormi* määritellään

$$\|\mathbf{u}\| = \|(x, y)\| = \max\{|x|, |y|\}.$$

Aksioomien 1–2 toteutuminen tälle normille on helposti todettavissa. Aksiooman 3 toteen näyttämiseksi olkoon $\mathbf{u} = (x_1\,, y_1) \in \mathbb{R}^2$ ja $\mathbf{v} = (x_2\,, y_2) \in \mathbb{R}^2$, jolloin

$$\|\mathbf{u} + \mathbf{v}\| = \max\{|x_1 + x_2|,\, |y_1 + y_2|\}.$$

Jos tässä on $|x_1 + x_2| \leq |y_1 + y_2|$, niin reaalilukujen kunnassa pätevän kolmioepäyhtälön (Lause I.2.3) perusteella voidaan päätellä:

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\| \;=\; |y_1 + y_2| &\leq |y_1| + |y_2| \\ &\leq \max\{|x_1|,\, |y_1|\} + \max\{|x_2|,\, |y_2|\} \;=\; \|\mathbf{u}\| + \|\mathbf{v}\|. \end{aligned}$$

Jos $|x_1 + x_2| > |y_1 + y_2|$, niin päätellään vastaavasti

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\| \;=\; |x_1 + x_2| &\leq |x_1| + |x_2| \\ &\leq \max\{|x_1|,\, |y_1|\} + \max\{|x_2|,\, |y_2|\} \;=\; \|\mathbf{u}\| + \|\mathbf{v}\|. \end{aligned}$$

Siis kolmioepäyhtälö on voimassa. □

Reaalikertoimista vektoriavaruutta, jossa on määritelty normi, sanotaan (reaalikertoimiseksi) *normiavaruudeksi.* Jos avaruudessa on määritelty myös sisätulo ja normi on siitä johdettu, sanotaan avaruutta *sisätuloavaruudeksi.*

Esimerkki 4 Vektoriavaruus $U = (\mathbb{R}^2, \mathbb{R})$ varustettuna euklidisella skalaaritulolla (Määritelmä II.4.1) on sisätuloavaruus nimeltä *euklidinen avaruus* $\mathbb{R}^2$. Maksiminormilla (Esimerkki 3) varustettuna U on vain normiavaruus, sillä maksiminormi ei ole johdettavissa mistään sisätulosta (Harj.teht. 5b).

Esimerkki 5 Hyvin yksinkertainen esimerkki sisätuloavaruudesta on yksiulotteinen avaruus $(U, \mathbb{R}, \cdot)$, missä $U = \mathbb{R}$. Tässä avaruudessa tulkitaan $\mathbf{u} + \mathbf{v}$ reaalilukujen yhteenlaskuksi ja $\alpha\mathbf{u}$ ja $\langle \mathbf{u}, \mathbf{v} \rangle$ molemmat reaalilukujen kertolaskuksi. Tällöin normi $|\mathbf{u}| = \mathbf{u}$:n itseisarvo, ja Cauchyn–Schwarzin epäyhtälö toteutuu yhtälönä $|\mathbf{uv}| = |\mathbf{u}||\mathbf{v}|$. □

HARJOITUSTEHTÄVIÄ

1. Näytä geometriaan vetoamatta, että $\mathbb{R}^2$:n euklidiselle skalaaritulolle ovat voimassa skalaaritulon aksioomat.

2. Tutki, onko kyseessä $\mathbb{R}^2$:n skalaaritulo, kun $\mathbf{u} = (x_1, y_1)$, $\mathbf{v} = (x_2, y_2)$ ja määritellään
   a) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + x_2 y_2$
   b) $\langle \mathbf{u}, \mathbf{v} \rangle = 2x_1 x_2 + 5y_1 y_2$
   c) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 x_2 - y_1 y_2$
   d) $\langle \mathbf{u}, \mathbf{v} \rangle = (x_1 + y_1)(x_2 + y_2)$
   e) $\langle \mathbf{u}, \mathbf{v} \rangle = (x_1 + y_1)(x_2 + y_2) + y_1 y_2$
   Minkä muodon saa (myönteisessä tapauksessa) Cauchyn–Schwarzin epäyhtälö ?

3. Olkoon $\{\vec{a}, \vec{b}\}$ tason vektoriavaruuden $U = (V, \mathbb{R})$ kanta. Näytä, että jos vektorin $\vec{v} \in V$ koordinaatit ko. kannassa ovat (x, y), niin $\|\vec{v}\| = |x| + |y|$ määrittelee U:n normin.

4. Olkoon $\mathbf{u} = (x, y) \in \mathbb{R}^2$. Mitkä seuraavista ovat vektoriavaruuden $U = (\mathbb{R}^2, \mathbb{R})$ normeja?
   a) $\|\mathbf{u}\| = |x + y|$
   b) $\|\mathbf{u}\| = |x| + |y|$
   b) $\|\mathbf{u}\| = 2|x| - |y|$
   c) $\|\mathbf{u}\| = 0.01|x| + 0.003|y|$
   d) $\|\mathbf{u}\| = x + |y|$
   e) $\|\mathbf{u}\| = |x| + y^2$
   f) $\|\mathbf{u}\| = \sqrt{2x^2 + 3y^2}$
   g) $\|\mathbf{u}\| = \sqrt{|xy|}$

5. a) Todista sisätuloavaruuden normille pätevä kolmioepäyhtälö muodossa $\|\mathbf{u} + \mathbf{v}\|^2 \leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$. b) Näytä, että sisätuloavaruuden normille pätee *suunnikasyhtälö* $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$. Päättele, että $\mathbb{R}^2$:n maksiminormi $\|\mathbf{u}\| = \|(x, y)\| = \max\{|x|, |y|\}$ ei ole johdettavissa mistään sisätulosta.

6. Tutki, millainen geometrinen muodonmuutos tapahtuu $\mathbb{R}^2$:n yksikköympyrässä $S = \{\mathbf{u} = (x, y) \in \mathbb{R}^2 \mid |\mathbf{u}| = 1\}$, kun pisteen etäisyys origosta mitataan euklidisen normin sijasta seuraavilla normeilla:

   a) $\|\mathbf{u}\| = \max\{|x|, |y|\}$ b) $\|\mathbf{u}\| = |x| + |y|$ c) $\|\mathbf{u}\| = \sqrt{4x^2 + y^2}$

7. (*) Olkoon $a, b, c, d \in \mathbb{R}$ ja määritellään

$$\langle \mathbf{u}, \mathbf{v} \rangle = a x_1 x_2 + b x_1 y_2 + c x_2 y_1 + d y_1 y_2 \,,$$

   missä $\mathbf{u} = (x_1, y_1) \in \mathbb{R}^2$ ja $\mathbf{v} = (x_2, y_2) \in \mathbb{R}^2$. Täsmälleen millä luvuille a, b, c, d asetettavilla ehdoilla $\langle \cdot, \cdot \rangle$ on vektoriavaruuden $U = (\mathbb{R}^2, \mathbb{R})$ skalaaritulo?

II.5 Funktion käsite. Trigonometriset funktiot

Tässä luvussa esitellään matematiikassa keskeinen käsite *funktio*, sen johdannainen *reaalifunktio* sekä ensimmäisinä reaalifunktioina *trigonometriset funktiot*. Osoittautuu, että riittävän yleisesti ja abstraktisti ymmärrettynä funktioita on esiintynyt jo aiemmin erinäisillä 'salanimillä'. Reaalifunktioiden osalta ei muita kuin trigonometrisia funktioita toistaiseksi käsitellä; yleisempi reaalifunktioiden teoria esitetään jäljempänä Luvuissa IV-VI.

Funktion käsite

Funktio l. *kuvaus* (engl. function, map, mapping) ymmärretään matematiikassa tavallisimmin kolmikkona muotoa

$$\{\text{joukko } A,\ \text{sääntö } f,\ \text{joukko } B\}.$$

Tällä tarkoitetaan, että on jokin sääntö f, jonka mukaan määräytyy yksikäsitteinen $y \in B$ jokaisella $x \in A$. Yhteys x:n y:n välillä merkitään $y = f(x)$ ja lausutaan 'y on f x'. Voidaan myös käyttää liittämisnuolta ja merkitä $x \mapsto f(x)(= y)$. Sanotaan, että A on f:n *määrittelyjoukko* (tai lähtöjoukko, engl. domain) ja B *maalijoukko* ja merkitään $f : A \to B$. Funktion määrittelyjoukon alkioihin viitataan yleisnimellä *muuttuja* (engl. variable). Jos $y = f(x)$, niin sanotaan, että y on f:n *arvo* (engl. value) *x:ssä* (tai muuttujan arvolla x); lyhyemmin lausutaan 'f x:ssä' tai usein myös 'f pisteessä x'. Maalijoukon osajoukko

$$\{y \in B \mid y = f(x) \text{ jollakin } x \in A\}$$

on nimeltään f:n *arvojoukko* tai A:n *kuva* f:ssä (engl. range, image). Tätä merkitään lyhyesti $f(A)$.

Esimerkki 1 Jos x ja y ovat reaalilukuja ja kirjoitetaan $f(x,y) = x^2y - y^3$, tarkoitetaan (ellei määrittelyjoukosta toisin sovita) funktiota

$$\{\mathbb{R}^2,\ f(x,y) = x^2y - y^3,\ \mathbb{R}\},$$

joka siis liittää mihin tahansa lukupariin $(x,y) \in \mathbb{R}^2$ laskusäännön ilmaiseman reaaliluvun. Funktion arvojen määräämisessä tarvitaan tässä tapauksessa ainoastaan reaalilukujen kerto- ja vähennyslaskuoperaatioita. Esim.

$$\begin{aligned} &f(3,-2) = 3^2 \cdot (-2) - (-2)^3 = -10, \\ &f(2,\sqrt{2}) = 2^2 \cdot \sqrt{2} - (\sqrt{2})^3 = 0, \\ &f(\pi,e) = \pi^2 e - e^3 = 6.74282937.. \quad \square \end{aligned}$$

Funktio joukko-opissa

Hyvin yleisesti ja abstraktisti ajatellen voidaan funktiota pitää 'säännön' sijasta enemmänkin 'luettelona', joka liittää kuhunkin alkioon $x \in A$ yksikäsitteisen (x:stä riippuvan) alkion $y \in B$. Näin ajatellen funktiosta tulee puhtaasti joukko-opillinen käsite: Funktion määrittelee joukon $A \times B = \{(x,y) \mid x \in A \ \&\ y \in B\}$ (= A:n ja B:n karteesinen tulo, vrt. alaviite Luvussa II.2) mikä tahansa osajoukko $F \subset A \times B$, joka toteuttaa ehdon

$$\forall x \in A\ [\,(x,y) \in F \ \text{ täsmälleen yhdellä } \ y \in B\,].$$

Nimittäin tällaista joukkoa vastaa A:ssa määritelty funktio f, kun tulkitaan

$$y = f(x) \ \Leftrightarrow\ (x,y) \in F.$$

Funktion käsitteen kannalta ei siis ole lopulta lainkaan merkitystä sillä, onko funktio tunnettu 'lausekkeena' (laskusääntönä) vai pelkkänä 'datana'.[†]

Esimerkki 2 <u>Reaalilukujono</u> $\{a_1, a_2, \ldots\}$ on tulkittavissa funktioksi tyyppiä $f : \mathbb{N} \to \mathbb{R}$. Nimittäin jono on ajateltavissa 'luettelona' $1 \mapsto a_1$, $2 \mapsto a_2$, ... eli joukkona

$$F = \{(1,a_1),(2,a_2),\ldots\} \subset \mathbb{N} \times \mathbb{Q}. \quad \square$$

Kaikki tähän asti esiintyneet <u>laskuoperaatiot</u> ovat itse asiassa funktioita. Funktion voi tällöin mieltää 'laskukoneeksi', joka antaa laskuoperaation lopputuloksen annetuilla lähtötiedoilla.

[†]Käytännössä 'funktiodata' voi perustua esim. suoriin mittauksiin tai se voi olla tietokonemalleihin perustuva laskettu ennuste, kuten sääennustuksissa.

Esimerkki 3 Reaalilukujen yhteenlasku ja kertolasku ovat funktioita tyyppiä $f:\ \mathbb{R}^2 \mapsto \mathbb{R}$ laskusäännöillä

$$\text{yhteenlasku:} \quad f(x,y) = x + y, \qquad \text{kertolasku:} \quad f(x,y) = xy.$$

Jakolasku on funktio $f : \{(x,y) \in \mathbb{R}^2 \mid y \neq 0\} \to \mathbb{R}$ säännöllä $f(x,y) = x/y$. $\square$

Esimerkki 4 Edellisessä luvussa liitettiin euklidisen tason kulmaan $\measuredangle(\vec{a},\vec{b})$ geometrisin keinoin määräytyvä reaaliluku, jota merkittiin $\cos\measuredangle(\vec{a},\vec{b})$. Koska ko. luku on yksikäsitteinen, niin kyseessä on funktio $\cos:\ A \to \mathbb{R}$, missä $A =$ kaikkien kulmien joukko. Funktion arvojoukko on $[-1,1] \subset \mathbb{R}$. Määrittelyjoukossa voidaan kulmat haluttaessa tulkita tason vektoripareina $(\vec{a},\vec{b})$, missä $\vec{a},\vec{b} \neq \vec{0}$, tai yksikkövektorien pareina ($|\vec{a}| = |\vec{b}| = 1$). $\square$

Reaalifunktio

'Funktioiden äiti' on *reaalifunktio* eli reaalimuuttujan reaaliarvoinen funktio muotoa $f : A \to \mathbb{R}$, missä $A \subset \mathbb{R}$. Myös termiä 'yhden muuttujan funktio' käytetään usein tässä rajatussa merkityksessä. Reaalimuuttujan funktio voidaan kätevästi havainnollistaa *kuvaajan* (engl. graph) avulla. Funktion $f : A \to \mathbb{R}$ kuvaaja joukossa $B \subset A$ ($A \subset \mathbb{R}$) on euklidisen avaruuden E^2 pistejoukko

$$G = \{P = (x,y) \mid y = f(x) \ \ \&\ \ x \in B\},$$

missä P:n koordinaatit (x,y) viittaavat karteesiseen koordinaatistoon (ellei toisin sovita). Huomattakoon, että jos f:n määrittelyjoukko rajataan joukoksi B, niin funktion joukko-opillinen määritelmä on

$$F = \{(x,y) \in \mathbb{R}^2 \mid y = f(x) \ \ \&\ \ x \in B\}.$$

Funktion tulkinta 'käyränä' (eli kuvaajana G) siis yksinkertaisesti geometrisoi funktion abstraktin joukko-opillisen määritelmän.

Kun reaalilukujono tulkitaan funktioksi $f:\ \mathbb{N} \to \mathbb{R}$, niin f:n kuvaaja on E^2:n pistejono. Tätä lukujonon kuvaustapaa on edellä jo käytetty Luvuissa I.6 ja I.10.

Esimerkki 5 Alla on kuvattu funktio, jonka laskusääntö on $f(x) = x^2/(x+1)$, joukoissa $B = [0,4]$ ja $B = [0,4] \cap \mathbb{N}$. Jälkimmäisessä tapauksessa kuvaaja esittää lukujonon $\{a_n\} = \{\, n^2/(n+1),\ n = 1,2,\dots\}$ alkupäätä. $\square$

$$f(x) = \frac{x^2}{x+1}$$

$$a_n = \frac{n^2}{n+1}$$

Jos reaalifunktioon liittyy helposti ilmaistavissa oleva laskusääntö, niin funktio ilmoitetaan tavallisimmin pelkkänä laskusääntönä. Tällöin oletetaan (ellei toisin sovita), että määrittelyjoukko on suurin $\mathbb{R}$:n osajoukko, jossa laskusääntöä voi soveltaa. Yksittäistä laskusäännön käyttöä, eli laskuoperaatiota $x \mapsto f(x)$ sanotaan *funktioevaluaatioksi*. Usein tämä onnistuu vain numeerisesti (eli likimäärin), jolloin käytännössä pystytään vain laskemaan äärellinen määrä termejä lukujonosta, jonka raja-arvo $= f(x)$.

Esimerkki 6 Reaalifunktion $f(x) = x^2/(x^2-2)$ määrittelyjoukko on $A = \{x \in \mathbb{R} \mid x^2 \neq 2\} = (-\infty, -\sqrt{2}) \cup (-\sqrt{2}, \sqrt{2}) \cup (\sqrt{2}, \infty)$. $\square$

Esimerkki 7 Potenssisarja $\sum_k a_k x^k$ voidaan tulkita reaalifunktioksi, jonka arvo x:ssä on sarjan summa ja määrittelyjoukko on sarjan suppenemisväli. Esimerkiksi funktion

$$f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$$

määrittelyjoukko on väli $[-1, 1)$, vrt. Luku I.12. Ellei ole $x = 0$, on funktioevaluaatio $x \mapsto f(x)$ tässä tapauksessa suoritettava numeerisesti, esim. laskemalla sarjan osasummia. $\square$

Funktiot kosini ja sini

Tarkastellaan yksikköympyrää, jonka keskipiste on karteesisen koordinaatiston origossa O ja jolta on valittu referenssipisteeeksi $A = (1, 0)$. Pisteestä A lähtien voidaan yksikköympyrää kiertää joko *positiiviseen kiertosuuntaan* eli vastapäivään tai *negatiiviseen kiertosuuntaan* eli myötäpäivään. Edellisessä tapauksessa asetetaan *kiertokulman* mitaksi kaarenpituus (= kehää pitkin kuljettu matka), jälkimmäisessä sen vastaluku. Tällä tavoin kiertokulman mitta voi saada minkä tahansa reaaliarvon, ja jokaista mittalukua α vastaa yksikäsitteinen kehän piste

$$P(\alpha) = (x(\alpha), y(\alpha)), \quad \alpha \in \mathbb{R}.$$

Tässä siis P ei tarkoita yksittäistä pistettä, vaan yksikäsitteistä riippuvuutta $\alpha \mapsto P(\alpha)$ eli 'pistearvoista' reaalimuuttujan funktiota. Vastaavasti x ja y tarkoittavat edellä funktioita tyyppiä

$$x: \ \mathbb{R} \to [-1, 1], \qquad y: \ \mathbb{R} \to [-1, 1].$$

Kulman kosinin aiemman määritelmän (ks. Luku II.3) mukaisesti on

$$\cos \measuredangle AOP(\alpha) = x(\alpha).$$

Määritellään nyt reaalifunktio cos eli *kosini* vaihtamalla tässä 'kulmamuuttujan' tilalle kiertokulman mittaluku α:

$$\cos \alpha = x(\alpha), \quad \alpha \in \mathbb{R}.$$

Määritellään vastaavasti reaalifunktio sin eli *sini* asettamalla

$$\sin \alpha = y(\alpha), \quad \alpha \in \mathbb{R}.$$

Reaalifunktiot kosini ja sini (määrittelyjoukkona $\mathbb{R}$) ovat *trigonometristen funktioiden* perustyypit. Yhdenmuotoisten kolmioiden periaatteella nähdään, että määritelmät vastaavat tuttuja trigonometrian merkintöjä, kun $0 < \alpha < \pi/2$.

$$\cos\alpha = b/c$$
$$\sin\alpha = a/c$$

Muutamia määritelmistä suoraan seuraavia, tai symmetrian avulla helposti perusteltavia, kosinin ja sinin ominaisuuksia on lueteltu alla. Nämä ovat voimassa $\forall\alpha \in \mathbb{R}$.

$$\cos(\alpha + 2\pi) = \cos\alpha, \quad \sin(\alpha + 2\pi) = \sin\alpha. \tag{1}$$
$$\cos(-\alpha) = \cos\alpha, \quad \sin(-\alpha) = -\sin\alpha. \tag{2}$$
$$\cos(\alpha + \pi) = -\cos\alpha, \quad \sin(\alpha + \pi) = -\sin\alpha. \tag{3}$$
$$\sin\alpha = \cos(\pi/2 - \alpha), \quad \cos\alpha = \sin(\pi/2 - \alpha). \tag{4}$$
$$\cos^2\alpha + \sin^2\alpha = 1. \tag{5}$$

Näistä ominaisuus (1) (joka seuraa myös (3):sta) kertoo, että cos ja sin ovat *jaksollisia* (periodisia) funktioita, jaksona 2π = yksikköympyrän kehän pituus. Ominaisuus (2) kertoo, että kosini on *parillinen* (symmetrinen) ja sini *pariton* (antisymmetrinen) funktio. Ominaisuus (5) on Pythagoraan lause.

Trigonometrisissa funktioevaluaatioissa $x \mapsto \cos x$ ja $x \mapsto \sin x$ on, poikkeukselliset x:n arvoja lukuunottamaatta, turvauduttava laskimiin. Laskinta tarvitaan myös, kun halutaan saada selville, missä kulmassa trigonometrinen funktio saa annetun arvon, ts. halutaan ratkaista α yhtälöstä $\cos\alpha = y$ tai $\sin\alpha = y$, kun tunnetaan $y \in [-1, 1]$. Laskimella saadaan yleensä yhden ratkaisun likiarvo (esim. komennolla $\cos^{-1} y$ tai $\arccos y$). Muut ratkaisut voidaan tämän jälkeen päätellä laskimeen enempää turvautumatta, sillä kosinin ja sinin määritelmien perusteella pätee

$$\boxed{\begin{array}{lcll} \cos\alpha = \cos\beta & \Leftrightarrow & \beta = \alpha + n\cdot 2\pi \ \vee \ \beta = -\alpha + n\cdot 2\pi, & n \in \mathbb{Z}, \\ \sin\alpha = \sin\beta & \Leftrightarrow & \beta = \alpha + n\cdot 2\pi \ \vee \ \beta = \pi - \alpha + n\cdot 2\pi, & n \in \mathbb{Z}. \end{array}}$$

Esimerkki 8 Millä kulman α asteluvuilla välillä $[0°, 360°]$ on $\cos 2\alpha = -2/3$ (yhden desimaalin tarkkuus)?

<u>Ratkaisu</u> Laskin antaa yhtälölle $\cos x = -2/3$ ratkaisun $x = 2.300523.. \widehat{=} 131.8°$, joten mahdolliset α:n arvot ovat

$$\begin{cases} 2\alpha = 131.8° + n\cdot 360° & \Rightarrow \quad \alpha = \ 65.9° + n\cdot 180°, \quad n \in \mathbb{Z}, \\ 2\alpha = -131.8° + m\cdot 360° & \Rightarrow \quad \alpha = -65.9° + m\cdot 180°, \quad m \in \mathbb{Z}. \end{cases}$$

Kysytyt α:n arvot suuruusjärjestyksessä $(n = 0,\ m = 1,\ n = 1,\ m = 2)$:

$$65.9°,\ 114.1°,\ 245.9°,\ 294.1°. \quad \square$$

Perusominaisuuksien (1)–(5) ohella tarvitaan monesti nk. *yhteenlaskukaavoja*

$$\boxed{\begin{array}{l} \cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ \sin(\alpha+\beta) = \ \sin\alpha\cos\beta + \cos\alpha\sin\beta. \end{array}} \tag{6}$$

Nämä ovat voimassa $\forall \alpha, \beta \in \mathbb{R}$. Kaavat on helpointa perustella vektorilaskulla, ks. kuvio.

$$\vec{a} = \cos\alpha\,\vec{i} + \sin\alpha\,\vec{j}$$

$$\vec{b} = \cos\beta\,\vec{i} + \sin\beta\,\vec{j}$$

Kuvion ja skalaaritulon määritelmän perusteella on

$$\vec{a}\cdot\vec{b} = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\beta-\alpha).$$

Vaihtamalla $\alpha \to -\alpha$ ja käyttämällä symmetriaominaisuuksia (2) seuraa kaavoista (6) ensimmäinen. Toinen seuraa tämän jälkeen ensimmäisestä sekä kaavoista (4), (2):

$$\begin{aligned}\sin(\alpha+\beta) &= \cos\left(\tfrac{\pi}{2}-\alpha-\beta\right)\\ &= \cos\left(\tfrac{\pi}{2}-\alpha\right)\cos(-\beta) - \sin\left(\tfrac{\pi}{2}-\alpha\right)\sin(-\beta)\\ &= \sin\alpha\cos\beta + \cos\alpha\sin\beta.\end{aligned}$$

Valitsemalla em. kaavoissa $\beta = \alpha$ saadaan monesti kysytyt *kaksinkertaisen kulman kaavat*:

$$\boxed{\begin{aligned}\cos 2\alpha &= \cos^2\alpha - \sin^2\alpha\\ &= 2\cos^2\alpha - 1\\ &= 1 - 2\sin^2\alpha,\\ \sin 2\alpha &= 2\sin\alpha\cos\alpha.\end{aligned}} \tag{7}$$

Esimerkki 9 Kaavojen (6), (7) ja (5) perusteella

$$\begin{aligned}\sin 3\alpha &= \sin(\alpha+2\alpha)\\ &= \sin\alpha\cos 2\alpha + \cos\alpha\sin 2\alpha\\ &= \sin\alpha(1-2\sin^2\alpha) + 2\cos^2\alpha\sin\alpha\\ &= \sin\alpha - 2\sin^3\alpha + 2(1-\sin^2\alpha)\sin\alpha\\ &= 3\sin\alpha - 4\sin^3\alpha. \quad \square\end{aligned}$$

Esimerkki 10 Laske $x = \sin\frac{\pi}{18} = \sin 10^\circ$.

Ratkaisu Edellisen esimerkin perusteella x toteuttaa kolmannen asteen yhtälön

$$3x - 4x^3 = \sin 30^\circ = \tfrac{1}{2}.$$

Numeerisin keinoin löydetään ratkaisu $x = 0.173648177669303\ldots$† $\square$

Tangentti ja kotangentti

Muista trigonometrisista funktioista tärkeimmät ovat *tangentti* (tan) ja *kotangentti* (cot), jotka määritellään

$$\tan\alpha = \frac{\sin\alpha}{\cos\alpha} \qquad (\alpha \in \mathbb{R},\ \alpha \neq (n+\tfrac{1}{2})\pi,\ n \in \mathbb{Z}),$$

$$\cot\alpha = \frac{\cos\alpha}{\sin\alpha} \qquad (\alpha \in \mathbb{R},\ \alpha \neq n\pi,\ n \in \mathbb{Z}).$$

Tangentin määritelmästä voidaan päätellä, että yhtälöllä $\tan\alpha = y$ on ratkaisu jokaisella $y \in \mathbb{R}$ ja että ratkaisu on yksikäsitteinen esim. lisäehdolla $\alpha \in (-\pi/2,\ \pi/2)$ tai $\alpha \in [0, \pi)$. Poikkeuksellisia y:n arvoja lukuunottamatta tarvitaan tällaisen ratkaisun likiarvon hakemiseen laskin (komento $\tan^{-1} y$ tai $\arctan y$). Muut yhtälön ratkaisut saadaan perusratkaisusta lisäämällä π:n monikertoja, sillä tangentilla on määritelmänsä perusteella ominaisuus

$$\boxed{\tan\alpha = \tan\beta \quad \Leftrightarrow \quad \beta = \alpha + n\cdot\pi, \quad n \in \mathbb{Z}.}$$

†Esimerkin luku x on algebrallinen mutta ei geometrinen, ts. $x \notin \mathbb{G}$ (vrt. Luku II.1). Kulmaa $\frac{\pi}{18} = 10^\circ$ ei siten ole mahdollista konstruoida geometrisesti. Yleisestikin gometrisesti konstruoitavissa olevat kulmat ovat melko harva joukko. Esimerkiksi kulma 3° voidaan konstruoida (ks. Harj.teht. 18), sen sijaan kulmia 2° ja 1° ei voida.

Suoraan kaavasta (5) seuraa tangenttiin liittyvä, usein käyttöön tuleva kaava

$$\boxed{\frac{1}{\cos^2\alpha} = 1 + \tan^2\alpha.} \tag{8}$$

Tangentin avulla saadaan myös joskus tarvittavia *puolen kulman kaavoja*: Kun lähdetään kaavoista (7) ja käytetään kaavaa (8), saadaan

$$\begin{aligned}
\sin\alpha &= 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} = 2\tan\frac{\alpha}{2}\cos^2\frac{\alpha}{2} = \frac{2\tan\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}\,,\\
\cos\alpha &= 2\cos^2\frac{\alpha}{2} - 1 = \frac{2}{1+\tan^2\frac{\alpha}{2}} - 1 = \frac{1-\tan^2\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}\,.
\end{aligned}$$

Näin saatiin kaavat

$$\boxed{\left.\begin{aligned}
\sin\alpha &= \frac{2t}{1+t^2}\\
\cos\alpha &= \frac{1-t^2}{1+t^2}\\
\tan\alpha &= \frac{2t}{1-t^2}
\end{aligned}\right\}\ t = \tan\frac{\alpha}{2}\,.} \tag{9}$$

Näiden mukaan trigonometristen funktioiden arvojen laskemiseen riittää suorittaa ainoastaan yksi 'aidosti trigonometrinen' laskuoperaatio $\alpha \mapsto \tan\frac{\alpha}{2} = t$, minkä jälkeen tarvitaan vain reaalilukujen kunnan laskuoperaatioita.[†]

Trigonometrisista funktioista maininnan arvoisia ovat vielä kosinin ja sinin johdannaisfunktiot *sekantti* (sec) ja *kosekantti* (csc):

$$\sec\alpha = \frac{1}{\cos\alpha} \qquad (\alpha\in\mathbb{R},\ \alpha\neq(n+\tfrac{1}{2})\pi,\ n\in\mathbb{Z}),$$

$$\csc\alpha = \frac{1}{\sin\alpha} \qquad (\alpha\in\mathbb{R},\ \alpha\neq n\pi,\ n\in\mathbb{Z}).$$

[†]Trigonometristen funktioiden laskukaavoista (1)–(9) on syytä huomauttaa, että kaikki näissä kaavoissa esiintyvät laskuoperaatiot, kuten $\alpha\mapsto\sin\alpha$, $\alpha\mapsto\pi/2-\alpha$, $(\alpha,\beta)\mapsto\alpha+\beta$ tai $\alpha\mapsto\alpha/2$, voidaan toteuttaa geometrisesti, jos lähtökohtana olevat kulmat tunnetaan geometrisina olioina. Näin ymmärrettynä kaavat (1)–(9) ovat siis päteviä, vaikkei kulman mittaa lainkaan määriteltäisi.

Sovellusesimerkki: Harmoninen värähtely

Sinin ja kosinin yhdistelmäfunktio $f(x) = A\sin x + B\cos x$ $(A, B \in \mathbb{R})$ on sovelluksissa yleinen. Tyypillisesti kyse on harmonisesta värähtelystä, jossa jokin fysikaalinen suure y (esim. sähköjännite) vaihtelee aikamuuttujan t mukaan siten, että $y(t) = A\sin\omega t + B\sin\omega t$. Suuretta ω (yksikkö = 1/s) sanotaan värähtelyn kulmataajuudeksi. Mainittuun funktioon $f(x)$ päädytään, kun otetaan käyttöön (dimensioton) muuttuja $x = \omega t$.

Tarkasteltava funktio saadaan selvempään muotoon, kun merkitään ensin

$$R = \sqrt{A^2 + B^2}.$$

Tällöin on $(A/R)^2 + (B/R)^2 = 1$, eli piste $P = (A/R, B/R)$ on yksikköympyrällä (ol. $(A, B) \neq (0, 0)$). Näin ollen jollakin $\alpha \in \mathbb{R}$ pätee

$$\frac{A}{R} = \cos\alpha, \quad \frac{B}{R} = \sin\alpha \quad \Rightarrow \quad \tan\alpha = \frac{B}{A}\,.$$

Näistä viimeinen ehto määrää α:n π:n monikertaa vaille yksikäsitteisesti. Huomioimalla myös $\sin\alpha$:n tai $\cos\alpha$:n merkki (kahdesta muusta ehdosta) nähdään, että α on 2π:n monikertaa vaille yksikäsitteinen. Yhteenlaskukaavan (7) perusteella $f(x)$ voidaan nyt esittää muodossa

$$\begin{aligned} f(x) &= R(\cos\alpha\sin x + \sin\alpha\cos x) \\ &= R\sin(x + \alpha) \\ &= R\sin(x - \beta), \quad \beta = -\alpha. \end{aligned}$$

Sanotaan tällöin, että R on (esim. värähtelyn) *amplitudi* ja α (tai $\beta = -\alpha$) on *vaihekulma* (vaihesiirtymä). Funktion f kuvaaja saadaan siis skaalaamalla funktion $\sin x$ kuvaaja amplitudilla R ja siirtämällä skaalattu kuvaaja vaihekulman verran x-akselin suunnassa.

Esimerkki 11 Kirjoita seuraavat funktiot perusmuotoon $f(x) = R\sin(x - \beta)$:

a) $f(x) = \sin x - \sqrt{3}\,\cos x$ b) $f(x) = -3\sin x - 4\cos x$

Ratkaisu a) Tässä on $R = 2$, jolloin α ratkeaa ehdoista

$$\cos\alpha = \frac{1}{2}, \quad \sin\alpha = -\frac{\sqrt{3}}{2} \quad \Rightarrow \quad \alpha = -\frac{\pi}{3} + n \cdot 2\pi, \quad n \in \mathbb{Z}.$$

Siis $f(x) = 2\sin(x - \frac{\pi}{3})$.

b) Tässä on $R = 5$ ja $\tan\alpha = 4/3$. Laskin antaa $\alpha = 0.927295..$ eli $\alpha \approx 53.1°$, mutta ratkaisu ei ole käypä, koska on $\cos\alpha > 0$. Valitaan $\alpha = -\pi + 0.927295..$ $= -2.214297..$ eli $\alpha \approx 53.1° - 180° = -126.9°$, jolloin on edelleen $\tan\alpha = 4/3$ ja $\cos\alpha < 0$. Siis $f(x) = 5\sin(x - \beta)$, missä $\beta = 2.214297.. \approxeq 126.9°$. $\square$

HARJOITUSTEHTÄVIÄ

1. Psykiatrin vastaanottoa voi kuvata matemaattisesti kolmikkona (psykiatri, potilas, diagnoosi). Pohdi seuraavissa tapauksissa (kussakin erikseen), millaisilla oletuksilla kyseessä on funktio:
   a) psykiatri : potilaat $\to$ diagnoosit
   b) potilas : psykiatrit $\to$ diagnoosit
   c) psykiatri : diagnoosit $\to$ potilaat
   d) potilas : diagnoosit $\to$ psykiatrit

2. Funktiot f ja g määritellään laskusäännöillä $f(x, y) = x^5 + 2x^4y - 2x^2y^3$ ja $g(x, y, z) = (x + y + z)/(1 + x^2 + y^2 + z^2)$, missä $x, y, z \in \mathbb{R}$. Laske $f(2, -3)$, $f(\sqrt{2}, \sqrt{2})$, $g(0, 0, 0)$ ja $g(1, -2, 3)$.

3. Mitkä ovat seuraavien funktioiden määrittelyjoukot ($x, y \in \mathbb{R}$)?

   $$\text{a) } f(x) = x^0 \quad \text{b) } f(x) = \frac{x^2}{x} \quad \text{c) } f(x, y) = \frac{x + y}{x^2 - y^2}$$

   Voiko nämä funktiot ilmaista määrittelyjoukossaan jollakin yksinkertaisemmalla laskusäännöllä?

4. Mitä yhteistä ja mitä eroa on seuraavilla reaalifunktioilla?

   $$\text{a) } f(x) = \frac{x^4}{x^2} \text{ ja } g(x) = x^2 \qquad \text{b) } f(x) = \sum_{k=1}^{\infty} \left(\frac{x}{2}\right)^k \text{ ja } g(x) = \frac{x}{2 - x}$$

5. Mitkä seuraavista joukon $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ osajoukoista F ovat funktioita?
   a) $F = \{(1, 2), (2, 1), (3, 3), (\pi, e), (e, \pi)\}$
   b) $F = \{(1, 2), (2, 3), (3, 3), (3, 2), (2, 1)\}$
   c) $F = \{(n, n^2) \mid n \in \mathbb{Z}\}$
   d) $F = \{(n^2, n) \mid n \in \mathbb{Z}\}$
   e) $F = \{(n^2, n) \mid n \in \mathbb{N}\}$

6. Tulkitse funktioina (määrittely- ja arvojoukko!) tason vektorien laskuoperaatiot: yhteenlasku, skalaarilla kertominen ja pistetulo.

7. a) Onko olemassa funktio f, jonka määrittelyjoukko $= \mathbb{R}$ ja $\forall x \in \mathbb{R}$ pätee $f(x) = \sum_{k=0}^{\infty} x^k$? Jos vastaus on myönteinen, niin määrittele f!
b) Yhden reaalimuuttujan sisältävän predikaatin (Luku I.3) voi tulkita funktioksi ja jopa reaalifunktioksi. Miten?

8. Olkoon $\sin\alpha = 7/25$ ja $\cot\beta = -5/12$. Laske lausekkeen $\sin(\alpha - \beta)$ mahdolliset arvot.

9. Nelikulmion sivujen pituudet ovat 1, 2, 3 ja 4, ja yhden kulman mitta asteina on 160°. Laske kaikkien nämä ehdot täyttävien nelikulmioiden kolmen muun kulman mitat 0.1 asteen tarkkuudella. Piirrä kuviot!

10. Ratkaise seuraavat trigonometriset yhtälöt:
a) $\sin 2x = \cos 7x$ b) $\tan 2x = 3\tan x$ c) $4\sin^2 x = \tan x$
d) $|\sin x + |\sin x|| = \cos x + |\cos x|$ e) $\cos 2x = \sin x + \cos x$

11. Ratkaise trigonometrinen epäyhtälö, eli määritä joukko $A \subset \mathbb{R}$ tai $A \subset \mathbb{R}^2$ siten, että epäyhtälö toteutuu täsmälleen kun $x \in A$ tai $(x, y) \in A$:
a) $\sin|x| < |\sin x|$ b) $|\sin 2x| \geq |\sin 3x|$ c) $\sin 4x > \cot x - \tan x$
d) $2\sin(x - y^2) > 1$ e) $\sin(x - y) + \cos x > 0$

12. a) Johda tangentin yhteenlaskukaava $\tan(\alpha + \beta) = \dfrac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}$.
b) Tunnetaan $t = \cos\alpha$. Lausu t:n avulla $\cos n\alpha$, kun $n = 2, 3, 4, 5$.
c) Johda käänteiset yhteenlaskukaavat, joissa $\cos x \cos y$, $\sin x \sin y$ ja $\cos x \sin y$ lausutaan $\cos(x \pm y)$:n ja $\sin(x \pm y)$:n avulla.

13. a) Näytä, että

$$\sin\frac{\alpha}{2} = \pm\sqrt{\tfrac{1}{2}(1 - \cos\alpha)}, \quad \cos\frac{\alpha}{2} = \pm\sqrt{\tfrac{1}{2}(1 + \cos\alpha)}.$$

Millä väleillä on molemmissa kaavoissa voimassa etumerkki + ?
b) Johda kaavat

$$\tan\frac{\alpha}{2} = \frac{\sin\alpha}{1 + \cos\alpha}, \quad \cot\frac{\alpha}{2} = \frac{\sin\alpha}{1 - \cos\alpha}.$$

14. Määritä amplitudi ja vaihekulma:
a) $f(x) = 3\cos x - 4\sin x$ b) $f(x) = -4\sin x + \cos x$
c) $f(x) = 76\cos x + 57\sin x$ d) $f(x) = \sin 2x(\sec x - 2\csc x)$

15. Vaihtovirran kolmen eri vaiheen jännitteet ovat

$$V_i(t) = V_0 \sin(\omega t + \varphi_i), \quad i = 1, 2, 3, \quad \varphi_2 = \varphi_1 + \frac{2\pi}{3}, \ \varphi_3 = \varphi_1 + \frac{4\pi}{3}.$$

Määritä vaiheiden 1 ja 2 välisen jännitteen $V_1 - V_2$ amplitudi ja vaihekulma. Mikä on kaikkien vaiheiden jännitteiden summa?

16. Saata funktio

$$f(x) = \sin x + 2\sin(x + \frac{2\pi}{3}) + 3\sin(x + \frac{4\pi}{3})$$

muotoon $f(x) = R\sin(x + \alpha)$.

17. (*) Funktioista puhuttaessa lausutaan $f(x)$ usein 'f pisteessä x'. Olkoon nyt $x \in \mathbb{R}$ annettu 'piste' ja määritellään 'pistefunktio' x seuraavasti:

$$x(f) = f(x) \ \forall f \in A,$$

missä $A = \{$reaalifunktiot, jotka on määritelty pisteessä $x\}$. Onko tällainen funktion määrittely todella mahdollinen ja jos on, miten $x(f)$ pitäisi lausua?

18. (*) Tasakylkinen kolmio, jonka huippukulma $= 36°$, voidaan jakaa kahteen kolmioon siten, että molemmat osakolmiot ovat tasakylkisiä. Käyttäen tätä ideaa lähtökohtana laske $x = \sin 3°$ tarkasti juurilukujen avulla. Päättele, että kulma $3°$ on konstruoitavissa geometrisesti.

19. (*) Todista: $\sin \dfrac{\pi}{2^{n+1}} \leq \left(\dfrac{1}{\sqrt{2}}\right)^n, \quad n = 0, 1, 2, \ldots$

20. (*) Määritä funktion $f(x) = \cos^2 x + 4\sin x \cos x + 3\sin^2 x$ pienin ja suurin arvo sekä minimi- ja maksimikohdat saattamalla funktio muotoon $f(x) = R\sin(2x + \alpha) + C$.

II.6 Avaruuden vektorit. Ristitulo

Avaruusgeometrisissa tarkasteluissa klassisena lähtökohtana on taaskin *euklidinen pisteavaruus*, tällä kertaa nimeltään E^3, 'E kolme'. Seuraavat avaruudelliset käsitteet oletetaan jatkossa (myös myöhemmissä luvuissa) tunnetuiksi:

- piste, (avaruus)jana, *avaruuskolmio*
- *avaruussuora*, *avaruustaso*, avaruuden puolisuora = *avaruussuunta*
- *tetraedri*, *suuntaissärmiö*, *monitahokas*
- *suorakulmainen särmiö*
- suorakulmaisen särmiön, suuntaissärmiön ja tetraedrin *tilavuus*

Avaruusvektorin määrittelyn lähtökohtana on euklidisen avaruusgeometrian perusoletus (tai perusaksioomien seuraus), että geometria jokaisella E^3:n avaruustasolla on sama kuin E^2:ssa. Tällöin voidaan ensinnäkin mitata avaruusjanan pituus (avaruustasolla, joka sisältää janan). Kun pituuteen liitetään avaruussuunta, tulee määritellyksi avaruusvektori suuntajanana. Vektori on siis jälleen yhdistetty tieto pituudesta (= vektorin itseisarvo) ja suunnasta.

Vektorin kertominen skalaarilla (reaaliluvulla) määritellään kuten tasossa, ts. skaalataan vektorin pituus ja joko säilytetään suunta tai vaihdetaan se vastakkaiseksi, riippuen kertojan etumerkistä. Vektorien yhteenlaskun määrittelemiseksi olkoon O E^3:n referenssipiste = *origo* ja $\vec{a} = \overrightarrow{OA}$ ja $\vec{b} = \overrightarrow{OB}$ kaksi avaruuden vektoria. Tällöin on olemassa avaruustaso T, joka kulkee pisteiden O, A, B kautta (täsmälleen yksi, jos pisteet eivät ole samalla avaruussuoralla). Kun siirrytään tasolle T, voidaan $\vec{a} + \vec{b}$ määritellä T:n kolmiodiagrammilla, eli avaruuskolmion avulla. Myös avaruusvektoreiden *skalaaritulo* määritellään kuten tasossa:

$$\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos \measuredangle(\vec{a}, \vec{b}).$$

Tässä geometrinen konstruktio $\measuredangle(\vec{a}, \vec{b}) \mapsto \cos \measuredangle(\vec{a}, \vec{b})$ toimii tasolla T oletuksen mukaan kuten E^2:ssa, vrt. Luku II.3. Skalaaritulolle ovat myös voimassa samat lait kuin tasossa. Näistä kuitenkin osittelulaki

$$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$$

kaipaa lisäperusteluja, sillä kolmea avaruuden vektoria ei yleisesti voi sijoittaa samaan tasoon.

Kuviossa C ei ole yleisesti samassa tasossa kuin pisteet O, A, B. Suorat l ja l' kuitenkin ovat yhdensuuntaiset (suuntavektori $= \vec{a}$), jolloin ne voidaan leikata kohtisuorasti pisteiden B ja C kautta kulkevilla avaruustasoilla (yhdensuuntaiset katkoviivat kuvassa). Leikkauspisteistä ja pisteistä O, A, B muodostettavat avaruuskolmiot $OB'B$ ja $OC'C$ (ks. kuvio) ovat tällöin suorakulmaiset. Kun nyt skalaaritulon tasogeometrista määritelmää sovelletaan tasoilla, joihin ko. kolmiot sisältyvät, ja huomioidaan, että avaruuden suorakulmiossa $BB'C'C''$ sivujanat BC'' ja $B'C'$ ovat yhtä pitkät, niin seuraa (vrt. Luku II.3)

$$\begin{aligned}\vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{c} &= |\vec{a}|\left(|\overrightarrow{OB'}|+|\overrightarrow{BC''}|\right)\\ &= |\vec{a}|\left(|\overrightarrow{OB'}|+|\overrightarrow{B'C'}|\right) = |\vec{a}||\overrightarrow{OC'}| = \vec{a}\cdot(\vec{b}+\vec{c}).\end{aligned}$$

Siis osittelulaki on pätevä.

Yhteenlaskun, skalaarilla kertomisen ja skalaaritulon tultua määritellyksi joukossa $V = \{\text{avaruuden vektorit}\}$ on $(V, \mathbb{R})$ jälleen vektoriavaruus ja skalaaritulolla (sisätulolla) $\vec{u}, \vec{v} \mapsto \vec{u}\cdot\vec{v}$ varustettuna sisätuloavaruus, vrt. Luvut II.3–II.4. Jos $\vec{a}$, $\vec{b}$ ja $\vec{c}$ ovat kolme avaruuden vektoria, jotka poikkeavat $\vec{0}$:sta eivätkä ole saman avaruustason suuntaiset, niin voidaan päätellä avaruusgeometrisesti (vrt. vastaava tasogeometrinen päättely Luvussa II.2), että jokainen $\vec{v} \in V$ voidaan esittää yksikäsitteisesti muodossa $\vec{v} = x\vec{a} + y\vec{b} + z\vec{c}$, missä $x, y, z \in \mathbb{R}$. Mainitut ehdot

vektoreille $\vec{a}, \vec{b}, \vec{c}$ voidaan pelkistää ehdoksi

$$\boxed{x\vec{a} + y\vec{b} + z\vec{c} = \vec{0} \quad \Rightarrow \quad x = y = z = 0.}$$

Sanotaan tällöin, että vektorit $\vec{a}, \vec{b}, \vec{c}$ ovat *lineaarisesti riippumattomat*, ja että vektorisysteemi $\{\vec{a}, \vec{b}, \vec{c}\}$ on V:n *kanta*. Siis avaruusvektoreista muodostuva vektoriavaruus on 3-ulotteinen: $\dim V = 3$.

Kun jokainen avaruuden vektori $\vec{v} \in V$ lausutaan annettujen kantavektoreiden lineaariyhdistelynä muodossa $\vec{v} = x\vec{a} + y\vec{b} + z\vec{c}$, niin syntyy kääntäen yksikäsitteinen vastaavuus $V \leftrightarrow \mathbb{R}^3$, missä $\mathbb{R}^3$ ('R kolme') on reaalikukukolmikkojen joukko:

$$\mathbb{R}^3 \;=\; \{(x, y, z) \mid x \in \mathbb{R},\; y \in \mathbb{R},\; z \in \mathbb{R}\} \;=\; \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

Myös $\mathbb{R}^3$ on vektoriavaruus, jossa vektorien laskuoperaatiot määritellään (vrt. $\mathbb{R}^2$:n operaatiot)

$$\begin{aligned} (x_1, y_1, z_1) + (x_2, y_2, z_2) &= (x_1 + x_2,\; y_1 + y_2,\; z_1 + z_2), \\ \lambda(x, y, z) &= (\lambda x, \lambda y, \lambda z) \quad (\lambda \in \mathbb{R}). \end{aligned}$$

Avaruusvektoreiden muodostaman vektoriavaruuden *ortonormeerattu kanta* on kolmen vektorin systeemi $\{\vec{i}, \vec{j}, \vec{k}\}$, joka toteuttaa

$$|\vec{i}| = |\vec{j}| = |\vec{k}| = 1, \quad \vec{i} \cdot \vec{j} = \vec{j} \cdot \vec{k} = \vec{i} \cdot \vec{k} = 0.$$

Lisäksi oletetaan yleensä, että $\{\vec{i}, \vec{j}, \vec{k}\}$ on nk. *oikeakätinen systeemi.*† Tällä tarkoitetaan oikeaan käteen liittyvää (funktio)vastaavuutta

$$(\text{ peukalo, etusormi, keskisormi }) \;\leftrightarrow\; (\vec{i}, \vec{j}, \vec{k})$$

tai vaihtoehtoisesti

$$(\text{ peukalo, etusormi, keskisormi }) \;\leftrightarrow\; (\vec{k}, \vec{i}, \vec{j}).$$

†Oikeakätinen vektorisysteemi muuttuu *vasenkätiseksi* (ja vastaavasti vasenkätinen oikeakätiseksi), jos yhden vektorin suunta vaihdetaan vastakkaiseksi.

Em. oletuksin sanotaan E^3:n koordinaatistoa $\{O, \vec{i}, \vec{j}, \vec{k}\}$ *karteesiseksi.* Suoria ja tasoja, joilla koordinaateista kaksi (suora) tai yksi (taso) saa arvon 0, sanotaan *koordinaattiakseleiksi* ja *koordinaattitasoiksi.* Nämä nimetään ko. suorilla tai tasoilla muuttuvien koordinaattien mukaan, esim. x-akseli, xy-taso.

Avaruuden vektoreiden skalaarituloa vastaa $\mathbb{R}^3$:n *euklidinen skalaaritulo* (vrt. Määritelmä II.4.1)

$$\mathbf{u}_1 = (x_1, y_1, z_1),\ \mathbf{u}_2 = (x_2, y_2, z_2)\ : \quad \mathbf{u}_1 \cdot \mathbf{u}_2 = x_1x_2 + y_1y_2 + z_1z_2.$$

Tälle ovat voimassa skalaaritulon aksioomat (ks. Luku II.4), joten pätee myös Cauchyn–Schwarzin epäyhtälö (Lause II.4.2)

$$|\mathbf{u}_1 \cdot \mathbf{u}_2| \leq |\mathbf{u}_1||\mathbf{u}_2|,$$

eli

$$|x_1x_2 + y_1y_2 + z_1z_2| \leq (x_1^2 + y_1^2 + z_1^2)^{1/2}(x_2^2 + y_2^2 + z_2^2)^{1/2},$$

missä

$$|\mathbf{u}| = (x^2 + y^2 + z^2)^{1/2}, \quad \mathbf{u} = (x, y, z) \in \mathbb{R}^3$$

on $\mathbb{R}^3$:n *euklidinen normi.*

Esimerkki 1 Avaruuskolmion kärjet ovat $A = (-1, 1, 2)$, $B = (4, -2, 3)$ ja $C = (-3, 3, 3)$. Laske kolmion kulmien mitat asteina (yhden desimaalin tarkkuus).

<u>Ratkaisu</u> Jos $\vec{a} = \overrightarrow{AB}$ ja $\vec{b} = \overrightarrow{AC}$, niin skalaaritulon määritelmän mukaan

$$\cos \measuredangle BAC = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}.$$

Tässä on $\vec{a} = 5\vec{i} - 3\vec{j} + \vec{k}$ ja $\vec{b} = -2\vec{i} + 2\vec{j} + \vec{k}$, joten saadaan

$$\cos \measuredangle BAC \ = \ \frac{5 \cdot (-2) + (-3) \cdot 2 + 1 \cdot 1}{(5^2 + 3^2 + 1^2)^{1/2}(2^2 + 2^2 + 1^2)^{1/2}} \ = \ -\sqrt{\frac{5}{7}}.$$

Vastaavalla tavalla laskien saadaan

$$\cos \measuredangle ABC \ = \ \sqrt{\frac{250}{259}}, \qquad \cos \measuredangle ACB \ = \ \sqrt{\frac{32}{37}}.$$

Laskimen avulla saadaan vastaukseksi:

$$\measuredangle BAC \approx 147.7^\circ, \quad \measuredangle ABC \approx 10.7^\circ, \quad \measuredangle ACB \approx 21.6^\circ. \quad \square$$

Vektorien ristitulo

Avaruuden vektoreille on määritelty skalaaritulon ohella toinen kertolaskun luonteinen operaatio, jota sanotaan *ristituloksi* tai *vektorituloksi* (engl. cross product, vector product). Jos $V = \{\text{avaruuden vektorit}\}$, niin ristitulo on kuvaus (funktio) tyyppiä $V \times V \to V$, ts. tulos on vektori (tästä nimitys vektoritulo). Ristitulo merkitään $\vec{a} \times \vec{b}$, luetaan 'a risti b'.

MÄÄRITELMÄ II.6.1 (**Ristitulo**) Avaruusvektorien $\vec{a}, \vec{b}$ *ristitulo* eli *vektoritulo* on vektori $\vec{a} \times \vec{b}$, joka toteuttaa

1. $|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}| \sin \measuredangle(\vec{a}, \vec{b})$
2. $\vec{a} \times \vec{b} \perp \vec{a} \quad \& \quad \vec{a} \times \vec{b} \perp \vec{b}$
3. $\{\vec{a},\ \vec{b},\ a \times \vec{b}\}$ on oikeakätinen systeemi.

Säännössä (1) esiintyvä kulman sini tulkitaan ei-negatiiviseksi, ts.

$$\sin \measuredangle(\vec{a}, \vec{b}) = \sin \alpha,$$

missä α on <u>sisäkulman</u> mittaluku ($0 \leq \alpha \leq \pi$). Sääntö (2) jättää ristitulon suunnalle kaksi vaihtoehtoa, joista valinta suoritetaan säännön (3) ilmaisemalla oikean käden säännöllä

$$(\text{ peukalo, etusormi, keskisormi }) \quad \mapsto \quad (\vec{a},\ \vec{b},\ \vec{a} \times \vec{b}).$$

Määritelmästä II.6.1 seuraa vektoritulolle 'vino' vaihdantalaki

$$\boxed{(\vec{b} \times \vec{a}) = -(\vec{a} \times \vec{b}).} \tag{1}$$

Erityisesti on

$$\vec{a} \times \vec{a} = \vec{0}$$

ja yleisemmin

$$\vec{a} \times \vec{b} = \vec{0} \quad \Leftrightarrow \quad \vec{a} \parallel \vec{b} \ \vee \ \vec{a} = \vec{0} \ \vee \ \vec{b} = \vec{0}.$$

(Tässä $\vec{a} \parallel \vec{b}$ tarkoittaa: $\vec{a} \uparrow\uparrow \vec{b}$ tai $\vec{a} \uparrow\downarrow \vec{b}$.) Ristitulosta (jollei muuten) voidaan siis päätellä, ovatko kaksi nollavektorista poikkeavaa vektoria yhdensuuntaiset.

Vektoritulo ei ole liitännäinen:

$$(\vec{a} \times \vec{b}) \times \vec{c} \ \neq \ \vec{a} \times (\vec{b} \times \vec{c}).$$

Esim. jos $\vec{a} = \vec{i}$, $\vec{b} = \vec{i}$ ja $\vec{c} = \vec{j}$, on vasen puoli $= \vec{0}$, mutta oikea puoli $= -\vec{j}$.

Jos $\lambda \in \mathbb{R}$ ja $\lambda \geq 0$, niin ristitulon määritelmästä seuraa

$$(\lambda\vec{a}) \times \vec{b} \ = \ \vec{a} \times (\lambda\vec{b}) \ = \ \lambda(\vec{a} \times \vec{b}).$$

Suuntaussäännöstä seuraa myös että jos $\vec{a}$:n tai $\vec{b}$:n suunta vaihdetaan, niin vaihtuu myös $(\vec{a} \times \vec{b})$:n suunta, ts.

$$(-\vec{a}) \times \vec{b} \ = \ \vec{a} \times (-\vec{b}) \ = \ -(\vec{a} \times \vec{b}).$$

Yhdistämällä nämä tulokset todetaan, että ristitulon ja skalaarilla kertomisen voi yhdistää normaaleilla osittelulaeilla:

$$\boxed{(\lambda\vec{a}) \times \vec{b} \ = \ \vec{a} \times (\lambda\vec{b}) \ = \ \lambda(\vec{a} \times \vec{b}), \quad \lambda \in \mathbb{R}.} \tag{2}$$

Ristitulon ja vektorien yhteenlaskun välillä toimivat myös normaalit osittelulait:

$$\boxed{\begin{aligned} \vec{a} \times (\vec{b} + \vec{c}) &= \vec{a} \times \vec{b} + \vec{a} \times \vec{c}, \\ (\vec{a} + \vec{b}) \times \vec{c} &= \vec{a} \times \vec{c} + \vec{b} \times \vec{c}. \end{aligned}} \tag{3}$$

Nämä lait eivät kuitenkaan ole määritelmän perusteella ilmeisiä, vaan tarvitaan hieman geometrista erittelyä: Olkoon T taso, jonka normaalivektori $= \vec{a}$. Tällöin $\vec{a} \times \vec{b}$ voidaan ymmärtää kolmivaiheisena geometrisena operaationa:

1. Projisioidaan vektori $\vec{b}$ tasolle T (*ortogonaaliprojektio*). Tulos $\vec{b}_T$.
2. Suoritetaan vektorin $\vec{b}_T$ *kierto* tasolla T suunnasta $\vec{a}$ katsottuna vastapäivään kulman $\pi/2$ verran (vektorin pituus säilyy).
3. Skaalataan vaiheen 2 tulos kertomalla luvulla $|\vec{a}|$.

Vaiheita 1–2 on havainnollistettu kahdella kuvalla alla. Ensimmäisessä kuvassa tarkkailija on suunnassa $\vec{a} \times \vec{b}$ pisteestä P, toisessa suunnassa $\vec{a}$. Kuviin on merkitty myös toisen kuvan tarkkailusuunta.

(a) Tarkkailija suunnassa $\vec{a} \times \vec{b}$

(b) Tarkkailija suunnassa $\vec{a}$

Ensimmäisessä kuvassa näkyy projektio-operaatio $\vec{b} \mapsto \vec{b}_T$. Kierto-operaation (vaihe 2) tulos osoittaa katsojan suuntaan. Toisessa kuvassa ei näy vektoria $\vec{a}$, ja $\vec{b}$:stä nähdään vain sen projektio $\vec{b}_T$. Kierto-operaatio näkyy tässä kuvassa.

Olkoon nyt $\vec{b}, \vec{c}$ avaruusvektoreita. Tällöin yhteenlaskudiagrammi $\vec{b} + \vec{c} = \vec{d}$ nähdään alla olevan kuvion mukaisena em. tarkkailutilanteessa (b):

Kuvassa nähdään vain oikean yhteenlaskudiagrammin (kolmio avaruudessa) projektio tasolle T. Kuva kuitenkin kertoo, että projisoitu diagrammi on tason vektoreiden yhteenlaskudiagrammi, ts. $\vec{d}_T = \vec{b}_T + \vec{c}_T$ eli

$$projektio(\vec{b} + \vec{c}) = projektio(\vec{b}) + projektio(\vec{c}).$$

Toisaalta kierto vaiheessa 2 merkitsee vain koko projisoidun yhteenlaskudiagrammin kiertoa, jolloin

$$\mathit{kierto}(\vec{b}_T + \vec{c}_T) = \mathit{kierto}(\vec{b}_T) + \mathit{kierto}(\vec{c}_T).$$

Yhdistämällä nämä kaksi tulosta seuraa:

$$\begin{aligned}\mathit{kierto}\,[\mathit{projektio}(\vec{b} + \vec{c})] &= \mathit{kierto}\,[\mathit{projektio}(\vec{b}\,) + \mathit{projektio}(\vec{c})] \\ &= \mathit{kierto}\,[\mathit{projektio}(\vec{b}\,)] + \mathit{kierto}\,[\mathit{projektio}(\vec{c})].\end{aligned}$$

Kun tässä molemmat puolet kerrotaan luvulla $|\vec{a}|$ ja todetaan, että

$$|\vec{a}|\,\mathit{kierto}\,[\mathit{projektio}(\overrightarrow{\ldots})] \;=\; \vec{a} \times (\overrightarrow{\ldots}),$$

niin on todistettu osittelulain (3) edellinen osa. Jälkimmäinen osa seuraa tästä sekä vaihdantalaista (1):

$$(\vec{a} + \vec{b}) \times \vec{c} \;=\; -\vec{c} \times (\vec{a} + \vec{b}) \;=\; -\vec{c} \times \vec{a} - \vec{c} \times \vec{b} \;=\; \vec{a} \times \vec{c} + \vec{b} \times \vec{c}.$$

Ristitulon determinanttikaava

Ristitulo lasketaan käytännössä ortonormeeratussa kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$. Koska

$$\begin{aligned}&\vec{i} \times \vec{i} = \vec{j} \times \vec{j} = \vec{k} \times \vec{k} = \vec{0}, \\ &\vec{i} \times \vec{j} = \vec{k}, \quad \vec{j} \times \vec{k} = \vec{i}, \quad \vec{k} \times \vec{i} = \vec{j}, \\ &\vec{j} \times \vec{i} = -\vec{k}, \quad \vec{k} \times \vec{j} = -\vec{i}, \quad \vec{i} \times \vec{k} = -\vec{j},\end{aligned}$$

niin sääntöjen (2)–(3) perusteella

$$\begin{aligned}\vec{a} \times \vec{b} &= (x_1\vec{i} + y_1\vec{j} + z_1\vec{k}) \times (x_2\vec{i} + y_2\vec{j} + z_2\vec{k}) \\ &= (y_1z_2 - y_2z_1)\vec{i} - (x_1z_2 - x_2z_1)\vec{j} + (x_1y_2 - x_2y_1)\vec{k}.\end{aligned}$$

Ilmaistaan tämä taulukkomuotoisella muistisäännöllä

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}.$$

Tässä siis toinen ja kolmas vaakarivi muodostuvat $\vec{a}$:n ja $\vec{b}$:n koordinaateista. Kyse on nk. *kolmirivisestä determinantista*, joka purkautuu ensin säännöllä

$$\vec{a} \times \vec{b} = \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\vec{i} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\vec{j} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\vec{k},$$

missä uudet taulukko-oliot ovat *kaksirivisiä determinantteja.*[†] Nämä on saatu alkuperäisestä kolmirivisestä poistamalla ne rivit ja sarakkeet, joiden yhtymäkohdassa on kertoimena oleva $\vec{i}$, $\vec{j}$ tai $\vec{k}$ — siis ensimmäinen rivi sekä sarake, johon ko. vektori sisältyy — huom. myös merkin vaihto kaavan keskimmäisessä termissä! Kaksiriviset determinantit puretaan lopulta säännöllä

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1.$$

Esimerkki 2 Jos $\vec{a} = 3\vec{i} - 6\vec{j} + 9\vec{k}$, $\vec{b} = 2\vec{i} + 5\vec{j} - 3\vec{k}$, $\vec{c} = -\vec{i} + 2\vec{j} - 3\vec{k}$, niin

$$\begin{aligned}
\vec{a} \times \vec{b} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 3 & -6 & 9 \\ 2 & 5 & -3 \end{vmatrix} \\
&= [(-6)\cdot(-3) - 5\cdot 9]\vec{i} - [3\cdot(-3) - 2\cdot 9]\vec{j} + [3\cdot 5 - 2\cdot(-6)]\vec{k} \\
&= -27\vec{i} - 27\vec{j} + 27\vec{k}, \\
\vec{a} \times \vec{c} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 3 & -6 & 9 \\ -1 & 2 & -3 \end{vmatrix} = \vec{0}.
\end{aligned}$$

Jälkimmäinen tulos merkitsee, että vektorit $\vec{a}$ ja $\vec{c}$ ovat yhdensuuntaiset (itse asiassa vastakkaissuuntaiset). □

Skalaarikolmitulo $\vec{a} \times \vec{b} \cdot \vec{c}$

Kolmen avaruusvektorin *skalaarikolmitulo* määritellään

$$\vec{a} \times \vec{b} \cdot \vec{c} = (\vec{a} \times \vec{b}) \cdot \vec{c}.$$

Tässä on $\vec{a} \times \vec{b}$ laskettava ensin, joten sulkeita ei tarvitse merkitä. Karteesisessa koordinaatistossa komponenttimuodossa laskettuna on

$$\vec{a} \times \vec{b} \cdot \vec{c} = \vec{c} \cdot (\vec{a} \times \vec{b}) = \begin{vmatrix} x_3 & y_3 & z_3 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix},$$

missä $\vec{a} \leftrightarrow (x_1, y_1, z_1)$, $\vec{b} \leftrightarrow (x_2, y_2, z_2)$ ja $\vec{c} \leftrightarrow (x_3, y_3, z_3)$. Determinantin määritelmästä seuraa, että rivit voidaan 'kierrättää' siten, että ensimmäinen rivi joutuu

[†]Determinanttioppi on käsinlaskussa ennen paljon käytetty, oma erikoinen matematiikan lajinsa. Tähän yhteyteen lainataan determinanttiopista vain ristitulon käsittelyssä tarvittavat muistisäännöt.

pohjimmaiseksi determinantin arvon muuttumatta:

$$\begin{vmatrix} x_3 & y_3 & z_3 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Näin ollen pätee 'pisteen ja ristin vaihtosääntö'

$$\boxed{\vec{a} \times \vec{b} \cdot \vec{c} = \vec{a} \cdot \vec{b} \times \vec{c}.}$$

Kolmion pinta-ala. Tetraedrin tilavuus

Ristitulolla ja skalaarikolmitulolla on fysikaalisten sovellutusten ohella myös hauskoja geometrisia käyttökohteita. Esimerkiksi avaruuskolmion (myös tasokolmion) pinta-ala voidaan määrätä kätevästi ristitulon avulla, jos tunnetaan kolmion kärkipisteet karteesisessa koordinaatistossa. Esimerkki valaiskoon asiaa.

Esimerkki 3 Avaruuskolmion kärjet ovat pisteissä $A = (1, 0, 0)$, $B = (1, 1, -1)$ ja $C = (0, 2, 2)$. Mikä on kolmion pinta-ala?

Ratkaisu Tunnetun kolmion pinta-alakaavan ja vektoritulon määritelmän perusteella:

$$\begin{aligned} \text{ala} &= \frac{1}{2} \cdot \text{kanta} \cdot \text{korkeus} \\ &= \frac{1}{2} |\overrightarrow{AB}||\overrightarrow{AC}| \sin \measuredangle BAC \\ &= \frac{1}{2} |\overrightarrow{AB} \times \overrightarrow{AC}|. \end{aligned}$$

Tässä on $\overrightarrow{AB} = \vec{j} - \vec{k}$, $\overrightarrow{AC} = -\vec{i} + 2\vec{j} + 2\vec{k}$, joten

$$\begin{aligned} \overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 1 & -1 \\ -1 & 2 & 2 \end{vmatrix} &= \begin{vmatrix} 1 & -1 \\ 2 & 2 \end{vmatrix} \vec{i} - \begin{vmatrix} 0 & -1 \\ -1 & 2 \end{vmatrix} \vec{j} + \begin{vmatrix} 0 & 1 \\ -1 & 2 \end{vmatrix} \vec{k} \\ &= (1 \cdot 2 + 1 \cdot 2)\,\vec{i} - (0 \cdot 2 - 1 \cdot 1)\,\vec{j} + (0 \cdot 2 + 1 \cdot 1)\,\vec{k} \\ &= 4\vec{i} + \vec{j} + \vec{k}. \end{aligned}$$

Siis pinta-ala $= \frac{1}{2}\sqrt{4^2 + 1^2 + 1^2} = \frac{3}{2}\sqrt{2}$. $\square$

Kolmion kolmiulotteinen vastine on tetraedri. Tämän tilavuus saadaan kätevästi lasketuksi skalaarikolmitulon avulla, kun tunnetaan kärkipisteet karteesisessa koordinaatistossa. Esimerkki valaiskoon jälleen asiaa.

Esimerkki 4 Tetraedrin K kärkinä ovat pisteet $O = (0,0,0)$, $A = (1,0,0)$, $B = (1,1,-1)$ ja $C = (0,2,2)$. Määritä tetraedrin tilavuus V.

Ratkaisu Tetraedrin tilavuuden tunnettu laskukaava on

$$V = \frac{1}{3} \cdot \text{pohjan ala} \cdot \text{korkeus}.$$

Jos tässä pohjaksi tulkitaan kolmio ABC, niin

$$V = \frac{1}{6}|\overrightarrow{AO} \cdot \overrightarrow{AB} \times \overrightarrow{AC}|,$$

sillä

$$\text{pohjan ala} = \frac{1}{2}|\overrightarrow{AB} \times \overrightarrow{AC}|, \quad \text{korkeus} = |\overrightarrow{AO'}| = |\overrightarrow{AO}||\cos \measuredangle OAO'|,$$

missä $\overrightarrow{AO'}$ on vektorin $\overrightarrow{AO}$ kohtisuora projektio vektorin $\overrightarrow{AB} \times \overrightarrow{AC}$ suuntaan, ks. kuvio.

Koska

$$\overrightarrow{AO} \cdot \overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ -1 & 2 & 2 \end{vmatrix} = (-1) \cdot \begin{vmatrix} 1 & -1 \\ 2 & 2 \end{vmatrix} = -4,$$

niin $V = \frac{1}{6} \cdot |-4| = \frac{2}{3}$ $\square$

Suunnikkaan ala. Suuntaissärmiön tilavuus

Suunnikas on tason nelikulmio, jonka vastakkaiset sivut ovat yhdensuuntaiset. Avaruudessa suunnikkaan kolmiulotteinen vastine on suuntaissärmiö eli 6-tahokas, jonka sivut ovat suunnikkaita (avaruustasoilla). Tämän erikoistapaus on suorakulmainen särmiö (sivut suorakulmioita). Jos janat OA, OB ja OC ovat suuntaissärmiön kolme särmää, niin sanotaan, että vektorit

$$\vec{a} = \overrightarrow{OA}, \quad \vec{b} = \overrightarrow{OB}, \quad \vec{c} = \overrightarrow{OC}$$

virittävät ko. särmiön. Vastaavasti vektorit $\vec{a}, \vec{b}$, samoin vektorit $\vec{a}, \vec{c}$ ja $\vec{b}, \vec{c}$, virittävät suunnikkaan (avaruustasolla). Suuntaissärmiön tilavuuden laskukaava on: $V =$ pohjan ala $\cdot$ korkeus. Tässä pohja on suunnikas, jonka pinta-ala saadaan ristitulon avulla (ala = kanta $\cdot$ korkeus), joten särmiön tilavuus saadaan skalaarikolmitulon avulla kuten tetraedrin tapauksessa. Laskukaavat, kun virittäjävektorit tunnetaan:

$$\boxed{\begin{aligned} &\text{Sunnikkaan pinta-ala} &&= |\vec{a} \times \vec{b}|. \\ &\text{Suuntaissärmiön tilavuus} &&= |\vec{a} \times \vec{b} \cdot \vec{c}|. \end{aligned}}$$

Vektorikolmitulo $\vec{a} \times (\vec{b} \times \vec{c})$

Jos T on taso, jonka suuntavektoreina ovat $\vec{b}, \vec{c}$ (ol. $\vec{b} \nparallel \vec{c}$), niin vektori $\vec{b} \times \vec{c}$ on tason normaalivektori. Näin ollen $\vec{a} \times (\vec{b} \times \vec{c})$ on jälleen tason suuntainen, ts.

$$\vec{a} \times (\vec{b} \times \vec{c}) = \lambda \vec{b} + \mu \vec{c}$$

joillakin $\lambda, \mu \in \mathbb{R}$. Toisaalta tämä vektori on myös $\vec{a}$:ta vasaan kohtisuora, eli

$$\vec{a} \cdot (\lambda \vec{b} + \mu \vec{c}) = \lambda(\vec{a} \cdot \vec{b}) + \mu(\vec{a} \cdot \vec{c}) = 0.$$

Tämä toteutuu täsmälleen, kun $\lambda = \gamma\, \vec{a} \cdot \vec{c}$ ja $\mu = -\gamma\, \vec{a} \cdot \vec{b}$ jollakin $\gamma \in \mathbb{R}$, joten on päätelty, että

$$\vec{a} \times (\vec{b} \times \vec{c}) = \gamma\, [\, (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}\,].$$

Tapauksessa $\vec{a} = \vec{b} = \vec{i}$, $\vec{c} = \vec{j}$ tämä toteutuu, kun $\gamma = 1$, joten päätellään, että vektorikolmitulolle pätee purkukaava[†]

$$\boxed{\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}.}$$

[†]Purkusääntöä tarkemmin perusteltaessa olisi myös näytettävä, että γ ei riipu vektoreista $\vec{a}$, $\vec{b}$, $\vec{c}$. Perustelut sivuutetaan tässä, sillä purkusääntö on näytettävissä oikeaksi myös suoraan vektorien komponenttimuodosta.

HARJOITUSTEHTÄVIÄ

1. Vektorin $\vec{u}$ koordinaatit kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$ ovat (x, y, z). Mitkä ovat $\vec{u}$:n koordinaatit kannassa $\{\vec{i}+\vec{j},\ \vec{j}+\vec{k},\ \vec{i}+\vec{k}\}$?

2. Laske seuraavien avaruusvektorien muodostamat kulmat $\measuredangle(\vec{a}, \vec{b})$ asteina:
   a) $\vec{a} = \vec{i},\ \vec{b} = \vec{i}+\vec{j}+\vec{k}$ b) $\vec{a} = \vec{i}+\vec{j},\ \vec{b} = \vec{j}+\vec{k}$

3. a) Avaruusvektoreista $\vec{a}, \vec{b}, \vec{c}$ tiedetään, että $|\vec{a}| = |\vec{b}| = 1$, $|\vec{c}| = 2$, $\vec{a} \perp \vec{b}$ ja $\measuredangle(\vec{a}, \vec{c}) = \measuredangle(\vec{b}, \vec{c}) = 60°$. Mikä on vektorin $\vec{u} = \vec{a}+\vec{b}+\vec{c}$ pituus?

   b) Määritä yksikkövektorit, jotka muodostavat yhtä suuret kulmat vektoreiden $\vec{k}$, $\vec{j}+\vec{k}$ ja $\vec{i}+\vec{j}+\vec{k}$ kanssa.

   c) Janasta OP (O = origo) tiedetään, että janan pituus on 5 ja että jana muodostaa positiivisen x-akselin kanssa kulman $\alpha = 32°$ ja positiivisen y-akselin kanssa kulman $\beta = 73°$. Laske P:n koordinaatit.

4. a) Säännöllisen tetraedrin kärjestä lähtevät särmävektorit ovat $\vec{a}$, $\vec{b}$ ja $\vec{c}$. Laske $\cos\measuredangle(\vec{a}, \vec{b}+\vec{c})$ ja $\cos\measuredangle(\vec{a}+\vec{b}+2\vec{c},\ 2\vec{a}-\vec{b})$.

   b) Pöydällä oleva A4-kokoinen paperiarkki on suorakulmio $ABCD$, jonka sivun pituuksien suhde on $|AB|/|BC| = \sqrt{2}$. Arkki taitetaan pitkin lävistäjää AC siten, että kolmio-osa ABC jää pöydälle ja osa ACD kääntyy kolmioksi ACD' pöydän pintaa vastaan vastaan kohtisuoralle tasolle. Laske $\cos\measuredangle D'AB$.

5. a) Määritä kaikki avaruusvektorit $\vec{u}$, jotka täyttävät ehdot (1) $|\vec{u}| = 1$, (2) $\vec{u} = \lambda(\vec{i}+\vec{j}) + \mu(\vec{j}+\vec{k})$ jollakin $(\lambda, \mu) \in \mathbb{R}^2$ ja (3) $\vec{u} \perp \vec{i} - 2\vec{j} - \vec{k}$.

   b) Jaa vektori $\vec{u} = \vec{i}+2\vec{j}-3\vec{k}$ kolmeen komponenttiin, joista yksi on vektorin $\vec{a} = 2\vec{i}+\vec{j}+\vec{k}$ suuntainen, toinen vektorin $\vec{b} = 2\vec{i}+3\vec{j}-3\vec{k}$ suuntainen ja kolmas kohtisuorassa vektoreita $\vec{a}$ ja $\vec{b}$ vastaan. Mikä on luvun $\sin\measuredangle(\vec{u}, \vec{v})$ pienin arvo, kun $\vec{v} = \lambda\vec{a} + \mu\vec{b}$, $(\lambda, \mu) \in \mathbb{R}^2$ $(\vec{v} \neq \vec{0})$?

   c) Muodosta ortonormeerattu, oikeakätinen avaruusvektorien kanta $\{\vec{a}, \vec{b}, \vec{c}\}$ siten, että $\vec{a}$ on vektorin $\vec{i}+\vec{j}+\vec{k}$ suuntainen ja $\vec{b}$ on xy-tason suuntainen. Määräytyykö kanta yksikäsitteisesti näistä ehdoista?

6. Sijoita säännöllinen oktaedri (= kahdeksansivuinen monitahokas, jonka sivutahkot ovat tasasivuisia kolmioita) mukavaan asentoon koordinatistoon siten, että yksi kärki on origossa. Laske tästä kärjestä alkavien särmien vektoriesitykset ja näiden avulla sivutahkojen normaalivektorit. Laske edelleen näiden avulla vierekkäisten sivutahkojen välinen *diedrikulma*, eli kulma, jonka sivutahkot muodostavat yhteisen särmän suunnasta katsottuna.

7. Osoita: $\vec{a}+\vec{b}+\vec{c}=\vec{0} \;\Rightarrow\; \vec{a}\times\vec{b}=\vec{b}\times\vec{c}=\vec{c}\times\vec{a}\,.$

8. a) Pisteet $(-1,-2,4)$, $(5,-1,0)$, $(2,-3,6)$ ja $(1,-1,1)$ ovat sekä tetraedrin että suuntiassärmiön kärkiä. Laske kummankin tilavuus ja pinnan ala.

   b) Tason kolmion kaksi kärkeä ovat pisteissä $(1,-2)$ ja $(3,3)$. Millaisessa E^2:n pistejoukossa on kolmannen kärjen täsmälleen oltava, jotta kolmion pinta-ala $= 10$? Kuva!

9. Päättele, että avaruusvektorit $\vec{a}$, $\vec{b}$, $\vec{c}$ ovat lineaarisesti riippuvat täsmälleen kun vektoreiden virittämän suuntaissärmiön tilavuus $= 0$. Tutki tällä kriteerillä ovatko seuraavat vektorisysteemit lineaarisesti riippumattomat:
   a) $\{\vec{i}-\vec{j}+\vec{k},\ 3\vec{i}+2\vec{j}+\vec{k},\ \vec{i}+\vec{j}-5\vec{k}\}$ b) $\{\vec{i}-2\vec{k},\ 2\vec{i}-\vec{j}+3\vec{k},\ 5\vec{i}-2\vec{j}+4\vec{k}\}$

10. Olkoon $\vec{n}$ avaruuden yksikkövektori. Halutaan esittää avaruusvektori $\vec{F}$ muodossa $\vec{F}=\vec{F}_1+\vec{F}_2$, missä $\vec{F}_1 || \vec{n}$ ja $\vec{F}_2 \perp \vec{n}$. Näytä, että $\vec{F}_1=(\vec{F}\cdot\vec{n})\vec{n}$ ja $\vec{F}_2=-\vec{n}\times(\vec{n}\times\vec{F})$.

11. Todista:
    a) $|\vec{a}\times\vec{b}|^2=|\vec{a}|^2|\vec{b}|^2-(\vec{a}\cdot\vec{b})^2$.
    b) $|\vec{a}\times(\vec{a}\times\vec{b})|^2=|\vec{a}|^2\big[|\vec{a}|^2|\vec{b}|^2-(\vec{a}\cdot\vec{b})^2\big]$.
    c) $(\vec{a}\times\vec{b})\times(\vec{b}\times\vec{c})\cdot(\vec{c}\times\vec{a}) \;=\; (\vec{a}\times\vec{b}\cdot\vec{c})^2$.

12. a) Vektorit $\vec{a}$, $\vec{b}$, $\vec{c}$ ovat lineaarisesti riippumattomat. Sievennä lauseke $[(\vec{a}\times\vec{b})\times(\vec{b}\times\vec{c})]\times(\vec{c}\times\vec{a})$. Millä ehdolla lauseke on $=\vec{0}$?

    b) Määritä vektorin $(\vec{a}\times(\vec{a}\times(\vec{a}\times(\vec{a}\times(\vec{a}\times(\vec{a}\times\vec{b}))))))$ pituus, kun $|\vec{a}|=3$, $|\vec{b}|=1$ ja $\vec{a}\cdot\vec{b}=-2$.

    c) Määritä vektorit $\vec{u}$, jotka toteuttavat $\vec{u}\cdot\vec{j}=0$ ja $\vec{u}\times(\vec{k}\times\vec{u})=\vec{k}$.

13. (*) a) Näytä, että tetraedrin keskijanat, eli kärjen ja vastakkaisen sivun keskiön yhdysjanat, leikkaavat toisensa samassa pisteessä (= tetraedrin keskiö). Missä suhteessa keskijanat jakautuvat leikkauspisteessä?
    c) Anna kaksi esimerkkiä tetraedrista, jonka kaikilla neljällä korkeusjanalla on yhteinen piste.

14. (*) Pisteet $A=(1,1,4)$, $B=(1,-1,3)$, $C=(-1,-1,2)$ ja $D=(0,-2,2)$ ovat eräällä avaruustasolla T. Tutki, millainen ko. tason kuvio syntyy, kun pisteet yhdistetään mainitussa järjestyksessä janoilla suljetuksi murtoviivaksi. Onko kyseessä nelikulmio? Valitse haluamasi kuvakulma (koordinaatisto) tasossa T ja piirrä kuva!

15. (*) Pisteet $(-3,-1,4)$, $(0,-1,-2)$, $(2,5,1)$, $(3,2,7)$ ja $(5,1,-2)$ ovat 5-tahokkaan K kärkipisteet. Laske K:n tilavuus ja pinnan ala.

II.7 Suorien, tasojen ja pintojen geometriaa

Kun geometrian tehtäviä ratkotaan algebran keinoin käyttäen hyväksi (yleensä karteesista) koordinaatistoa, puhutaan *analyyttisestä geometriasta.* Vektorit, skalaari- ja ristituloineen, tarjoavat moniin tehtäviin kätevän apuneuvon.

Jos O on Euklidisen avaruuden (E^2 tai E^3) origo ja P on ko. avaruuden piste, sanotaan vektoria $\overrightarrow{OP}$ tästedes pisteen P *paikkavektoriksi* ja merkitään

$$\vec{r} = \overrightarrow{OP} \quad \text{(paikkavektori)}.$$

Käytetään myös vastaavuusmerkintää

$$\vec{r} \mathrel{\widehat{=}} P \;\Leftrightarrow\; \vec{r} = \overrightarrow{OP}.$$

Suora

Jos $P_0 = (x_0, y_0) \mathrel{\widehat{=}} \vec{r}_0 = x_0\vec{i} + y_0\vec{j}$ on euklidisen tason piste ja $\vec{v} = \alpha\vec{i} + \beta\vec{j}$ on tason vektori, $\vec{v} \neq \vec{0}$, niin pistejoukko

$$S = \{P \in E^2 \mid P \mathrel{\widehat{=}} \vec{r} = \vec{r}_0 + t\vec{v} \quad \text{jollakin } t \in \mathbb{R}\}$$

on suora, jonka *suuntavektori* $= \vec{v}$ ja joka kulkee pisteen P_0 kautta.

Sanotaan, että yhtälö

$$\vec{r} = \vec{r}_0 + t\vec{v} \tag{1}$$

on suoran S *parametrimuotoinen yhtälö* tai *parametrisointi* (parametrisaatio) ja että t on *parametri.* Yleisemmin näillä termeillä tarkoitetaan, että tarkasteltavan pistejoukon (tässä suoran) jokainen piste saadaan annetusta esitysmuodosta

jollakin parametrin arvolla. Tässä tapauksessa jokaista suoran pistettä vastaa yksikäsitteinen parametrin t arvo ja kääntäen, ts. parametrisointi (1) luo kääntäen yksikäsitteisen vastaavuuden $S \leftrightarrow \mathbb{R}$.

Kun kirjoitetaan $\vec{r} = x\vec{i} + y\vec{j}$, niin yhtälölle (1) saadaan *koordinaattimuoto*

$$\begin{cases} x = x_0 + \alpha t, \\ y = y_0 + \beta t. \end{cases} \tag{2}$$

Eliminoimalla t seuraa tästä x:n ja y:n välinen riippuvuus

$$ax + by + c = 0, \tag{3}$$

missä $a = \beta$, $b = -\alpha$ ja $c = -\beta x_0 + \alpha y_0$. Tämä on suoran yhtälön *perusmuoto* (eiparametrinen muoto). Perusmuotoisen yhtälön (3) voi ilmaista myös vektoreiden avulla, sillä kun asetetaan

$$\vec{n} = a\vec{i} + b\vec{j} \quad (= \beta\vec{i} - \alpha\vec{j}\,),$$

niin yhtälö on sama kuin

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0. \tag{4}$$

Tässä $\vec{n}$ on S:n *normaalivektori*, ts. S:n suuntavektoria $\vec{v}$ vastaan kohtisuora ($\vec{0}$:sta poikkeava) vektori. Normaalivektorin avulla suoran yhtälö voidaan siis kirjoittaa suoraan ehtona $\vec{r} - \vec{r}_0 \perp \vec{n}$, kuten geometrisestikin on helppo päätellä.

Esimerkki 1 Suoran yhtälöstä $3x - 2y + 5 = 0$ nähdään heti, että suoran normaalivektori on $\vec{n} = 3\vec{i} - 2\vec{j}$. Siis suoran eräs suuntavektori on $\vec{v} = 2\vec{i} + 3\vec{j}$. Koska esimerkiksi $P_0 = (1, 4)$ on suoralla, saadaan suoralle parametrisoitu esitys (1), missä

$$\vec{r}_0 = \vec{i} + 4\vec{j}, \quad \vec{v} = 2\vec{i} + 3\vec{j}.$$

Parametrisoinnin koordinaattimuoto (2) on tässä tapauksessa

$$\begin{cases} x = 1 + 2t, \\ y = 4 + 3t. \end{cases}$$

Mainittu piste P_0 vastaa siis tässä parametrin arvoa $t = 0$. $\square$

Parametrisointi (1) toimii myös avaruussuoralle. Jos $\vec{r}_0 = x_0\vec{i} + y_0\vec{j} + z_0\vec{k}$ ja $\vec{v} = \alpha\vec{i} + \beta\vec{j} + \gamma\vec{k}$, niin koordinaattimuotoiset suoran yhtälöt ovat

$$\begin{cases} x = x_0 + \alpha t, \\ y = y_0 + \beta t, \\ z = z_0 + \gamma t. \end{cases} \tag{5}$$

Eliminoimalla t jää jäljelle kaksi ei-parametrista yhtälöä.

Esimerkki 2 Suora kulkee pisteiden $A = (1, 1, 1)$ ja $B = (1, -2, 3)$ kautta. Mikä on suoran yhtälö?

<u>Ratkaisu</u> Eräs suoran suuntavektori on

$$\vec{v} = \overrightarrow{AB} = -3\vec{j} + 2\vec{k}$$

ja siis eräs parametriesitys on muotoa (1), missä esimerkiksi

$$\vec{r}_0 = \overrightarrow{OA} = \vec{i} + \vec{j} + \vec{k}.$$

Koordinaattimuodossa

$$\begin{cases} x = 1, \\ y = 1 - 3t, \\ z = 1 + 2t. \end{cases}$$

Eliminoimalla t saadaan saadaan suoran ei-parametrisiksi yhtälöiksi

$$\begin{cases} x & = 1, \\ 2y + 3z & = 5. \end{cases}$$

Nämä voidaan kirjoittaa myös muotoon

$$\begin{cases} \vec{n}_1 \cdot (\vec{r} - \vec{r}_0) = 0, \\ \vec{n}_2 \cdot (\vec{r} - \vec{r}_0) = 0, \end{cases}$$

missä $\vec{n}_1 = \vec{i}$ ja $\vec{n}_2 = 2\vec{j} + 3\vec{k}$ ovat suoran normaalivektoreita. Tämä vastaa tason suoran yhtälöä (4). $\square$

Esimerkki 3 Hävittäjä A nousee kentältä, joka on pisteessä $(5,7,0)$ (yksikkö=km), suuntaan $-\vec{i}-2\vec{j}+\vec{k}$ ja hävittäjä B nousee toiselta kentältä, joka on pisteessä $(1,-9,0)$, suuntaan $-\vec{i}+\vec{j}+2\vec{k}$. Kuinka lähelle toisiaan koneet voivat joutua (pahin mahdollinen skenaario)?

Ratkaisu Tässä on kyse nk. kahden suoran ongelmasta, jossa on etsittävä suorien lyhin etäisyys d. Suorat ovat

$$\begin{aligned} S_1:\ \vec{r} &= 5\vec{i}+7\vec{j}+t(-\vec{i}-2\vec{j}+\vec{k}) = \vec{r}_1+t\vec{v}_1,\\ S_2:\ \vec{r} &= \vec{i}-9\vec{j}+s(-\vec{i}+\vec{j}+2\vec{k}) \ = \vec{r}_2+s\vec{v}_2. \end{aligned}$$

Lyhin etäisyys on $d=|\overrightarrow{PQ}|$, missä $P\in S_1$, $Q\in S_2$ ja vektori $\overrightarrow{PQ}$ on molempia suoria vastaan kohtisuora, ts.

$$\overrightarrow{PQ}\cdot\vec{v}_1=\overrightarrow{PQ}\cdot\vec{v}_2=0.$$

Kun tässä $P\mathrel{\widehat{=}}\vec{r}(t)$ ja $Q\mathrel{\widehat{=}}\vec{r}(s)$ esitetään parametrisointien mukaisesti, niin

$$\begin{aligned} \overrightarrow{PQ} &= \vec{r}_2-\vec{r}_1+s\vec{v}_2-t\vec{v}_1\\ &= (-4+t-s)\vec{i}+(-16+2t+s)\vec{j}+(-t+2s)\vec{k}, \end{aligned}$$

joten saadaan yhtälöryhmä

$$\begin{cases} -1\cdot(-4+t-s)-2\cdot(-16+2t+s)+1\cdot(-t+2s)=0\\ -1\cdot(-4+t-s)+1\cdot(-16+2t+s)+2\cdot(-t+2s)=0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 6t-s=36\\ t-6s=-12 \end{cases}$$

Ratkaisemalla t, s saadaan

$$t = \frac{228}{35}, \quad s = \frac{108}{35}.$$

Näillä arvoilla on

$$\overrightarrow{PQ} = \frac{1}{35}(-20\vec{i} + 4\vec{j} - 12\vec{k}).$$

Koska komponentin $\vec{k}$ kerroin on negatiivinen, lentää kone A kohtaamistilanteessa korkeammalla. Koneiden etäisyys on siis pahimman skenaarion mukaan

$$d = |\overrightarrow{PQ}| = 4/\sqrt{35} \approx 0.676 = 676 \text{ m},$$

ja pisteet P ja Q saadaan erikseenkin suorien parametriesityksistä:

$$S_1 : \ \vec{r}\left(t = \frac{228}{35}\right) \mathrel{\widehat{=}} P \approx (-1.514, -6.029, 6.514),$$
$$S_2 : \ \vec{r}\left(s = \frac{108}{35}\right) \mathrel{\widehat{=}} Q \approx (-2.086, -5.914, 6.171). \quad \square$$

Taso

Tarkastellaan avaruuden tasoa T, joka kulkee pisteiden $P_0 = (x_0, y_0, z_0)$, $P_1 = (x_1, y_1, z_1)$ ja $P_2 = (x_2, y_2, z_2)$ kautta. Oletetaan, että pisteet eivät ole samalla suoralla, jolloin vektorit

$$\vec{v}_1 = \overrightarrow{P_0P_1} = \alpha_1\vec{i} + \beta_1\vec{j} + \gamma_1\vec{k},$$
$$\vec{v}_2 = \overrightarrow{P_0P_2} = \alpha_2\vec{i} + \beta_2\vec{j} + \gamma_2\vec{k}$$

ovat lineaarisesti riippumattomat. Näiden, tason T *suuntavektoreiden* avulla saadaan tasolle parametrisoitu esitysmuoto

$$\vec{r} = \vec{r}_0 + t_1\vec{v}_1 + t_2\vec{v}_2, \quad t_1, t_2 \in \mathbb{R}. \tag{6}$$

Tässä siis reaalisia parametreja on kaksi. Jos parametri halutaan nähdä 'yhtenä', niin parametriksi on tulkittava $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$. Parametrisointi (6) luokin kääntäen yksikäsitteisen vastaavuuden $T \leftrightarrow \mathbb{R}^2$, ja $\{P_0, \vec{v}_1, \vec{v}_2\}$ voidaan tulkita T:n koordinaatistoksi, vrt. Luku II.2.

Yhtälön (6) koordinaattimuoto on

$$\begin{cases} x &= x_0 + \alpha_1 t_1 + \alpha_2 t_2, \\ y &= y_0 + \beta_1 t_1 + \beta_2 t_2, \\ z &= z_0 + \gamma_1 t_1 + \gamma_2 t_2. \end{cases}$$

Parametrien eliminointi käy kuitenkin helpommin vektorimuodosta (6), kun otetaan käyttöön vektori

$$\vec{n} = \vec{v}_1 \times \vec{v}_2.$$

Tämä on molempia vectoreita $\vec{v}_1, \vec{v}_2$ vastaan kohtisuora, eli kyseessä on tason *normaalivektori*. Kertomalla yhtälö (6) skalaarisesti vektorilla $\vec{n}$ ja huomioimalla, että $\vec{n} \cdot \vec{v}_1 = \vec{n} \cdot \vec{v}_2 = 0$, seuraa tasolle yhtälö

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0, \tag{7}$$

joka siis on aivan samaa muotoa kuin tason suoran yhtälö (4). Jos

$$\vec{n} = a\vec{i} + b\vec{j} + c\vec{k}, \quad d = -\vec{n} \cdot \vec{r}_0\,,$$

niin (7) on edelleen sama kuin

$$ax + by + cz + d = 0. \tag{8}$$

Tämä on tason yhtälön *perusmuoto*. Perusmuodosta nähdään siis suoraan, mikä on tason normaalivektori.†

Esimerkki 4 Taso T_1 kulkee pisteiden $A = (1, 1, 2)$, $B = (0, 3, -1)$ ja $C = (-2, -2, -3)$ kautta. Taso T_2 kulkee pisteen $(-1, -1, -1)$ kautta ja sen normaalivektori on $\vec{n}_2 = \vec{i} - \vec{j} - \vec{k}$. Määrää tasojen yhtälöt perusmuodossa (8) ja tasojen leikkaussuoran parametriesitys muodoissa (1) ja (5).

<u>Ratkaisu</u> Tason T_2 yhtälö on

$$\vec{n}_2 \cdot (\vec{r} + \vec{i} + \vec{j} + \vec{k}) = 0,$$

eli

$$T_2: \quad x - y - z - 1 = 0.$$

Tason T_1 normaalivektori on

$$\vec{n}_1 = \overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -1 & 2 & -3 \\ -3 & -3 & -5 \end{vmatrix} = -19\vec{i} + 4\vec{j} + 9\vec{k},$$

joten valitsemalla esimerkiksi $\vec{r}_0 \mathrel{\widehat{=}} A$ saadaan tason T_1 yhtälöksi

$$T_1: \quad \vec{n}_1 \cdot (\vec{r} - \vec{i} - \vec{j} - 2\vec{k}) = 0 \quad \Leftrightarrow \quad -19x + 4y + 9z - 3 = 0.$$

†Tason *normaali* on jokainen avaruussuora, jonka suuntavektori = tason normaalivektori. Termiä käytetään usein myös normaalivektorista puhuttaessa. Esim: 'Yksikkövektori $\vec{n}$ on tason normaali.'

Leikkaussuoran S suuntavektori on

$$\vec{v} = \vec{n}_1 \times \vec{n}_2 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -19 & 4 & 9 \\ 1 & -1 & -1 \end{vmatrix} = 5\vec{i} - 10\vec{j} + 15\vec{k}.$$

Suuntavektoriksi kelpaa myös $\frac{1}{5}\vec{v} = \vec{i} - 2\vec{j} + 3\vec{k}$. Vielä tarvitaan yksi suoran piste $P_0 \mathrel{\widehat{=}} \vec{r}_0$. Tämä löydetään esim. tasojen T_1, T_2 ja (sopivasti valitun) kolmannen tason T_3 leikkauspisteenä. Valitaan kolmanneksi tasoksi yz-taso

$$T_3 : \ x = 0,$$

jolloin $P_0 = (x_0, y_0, z_0)$ ratkeaa (jos ratkeaa) yhtälöryhmästä

$$\begin{cases} -19x_0 + 4y_0 + 9z_0 &= 3 \\ x_0 - y_0 - z_0 &= 1 \\ x_0 &= 0 \end{cases}$$

Ratkaisu on $P_0 = (0, -\frac{12}{5}, \frac{7}{5})$, ja suoran S parametriesitys näin muodoin

$$\vec{r} = \frac{1}{5}(-12\vec{j} + 7\vec{k}) + t(\vec{i} - 2\vec{j} + 3\vec{k}),$$

eli koordinaattimuodossa

$$S : \quad \begin{cases} x &= t, \\ y &= -\frac{12}{5} - 2t, \\ z &= \frac{7}{5} + 3t. \end{cases} \qquad \square$$

Esimerkki 5 Määritä annetun pisteen $P = (x_0, y_0, z_0)$ etäisyys tasosta $T : \ ax + by + cz + d = 0$.

<u>Ratkaisu</u> Jos Q on pistettä P lähinnä oleva piste tasolla T, niin ilmeisesti jollakin $t \in \mathbb{R}$ on $\overrightarrow{PQ} = t\vec{n}$, missä $\vec{n} = a\vec{i} + b\vec{j} + c\vec{k}$ on tason normaalivektori. Siis $Q = (x_0 + ta, y_0 + tb, z_0 + tc)$ jollakin t. Koska $Q \in T$, niin T:n yhtälön perusteella

$$0 \ = \ a(x_0 + ta) + b(y_0 + tb) + c(y_0 + tc) + d \ = \ ax_0 + by_0 + cz_0 + d + t|\vec{n}|^2.$$

Ratkaisemalla tästä t saadaan etäisyyden laskukaavaksi

$$h \ = \ |t||\vec{n}| \ = \ \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Tason pisteen (x_0, y_0) etäisyys suorasta $S : \ ax + by + c = 0$ saadaan samasta kaavasta asettamalla ensin $c = 0$ ja sitten d:n tilalle c. $\square$

Ympyrä ja pallo

Ympyrä on (harpin olomuodossa) euklidisen tason alkuperäisiä olioita. Jos ympyrän keskipiste on $P_0 \mathrel{\widehat{=}} \vec{r}_0$ ja säde R, niin *ympyrän* (ympyräviivan) *yhtälö* on

$$|\vec{r} - \vec{r}_0| = R.$$

Tämä on myös *pallon* (pallopinnan)† yhtälö. Neliöitynä ja koordinaattien avulla kirjoitettuna yhtälöt ovat

$$\begin{aligned} \text{Ympyrä:} \quad & (x - x_0)^2 + (y - y_0)^2 = R^2. \\ \text{Pallo:} \quad & (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2. \end{aligned}$$

Ympyrän tangentti on suora, jolla on ympyrän kanssa täsmälleen yksi yhteinen piste. Sanotaan, että tangentti *sivuaa* ympyrää ko. pisteessä. Jos sivuamispiste on annettu piste Q, niin tangentti konstruoidaan geometrisesti piirtämällä ensin suora S_1 pisteen Q ja ympyrän keskipisteen kautta, jolloin tangentti S_2 on tätä vastaan kohtisuora. Sanotaan, että suora S_1 on *ympyrän normaali* pisteessä Q tai että suora S_1 leikkaa ympyrän *kohtisuorasti*. Yleisemminkin voidaan ympyräviivan ja suoran välinen *leikkauskulma* määritellä ko. suoran ja leikkauspisteeseen asetetun tangentin suuntavektoreiden välisenä kulmana.

Vastaavalla tavalla kuin ympyrän tangentti määritellään *pallon tangenttitaso*. Pallon keskipisteen kautta kulkeva suora on *pallon normaali*, eli se on kohtisuora tangenttitasoa vastaan pallon ja suoran leikkauspisteessä. Jos pallolla ja tasolla T on enemmän kuin yksi yhteinen piste, on yhteisten pisteiden joukko *avaruusympyrä* (ympyrä tasolla T). Jos T kulkee pallon keskipisteen kautta, sanotaan leikkausviivaa pallon *isoympyräksi*. *Avaruusympyrän tangentti* on ympyrän kanssa samassa tasossa oleva, ympyrää sivuava avaruussuora.

Esimerkki 6 Pallon keskipiste on (x_0, y_0, z_0) ja säde $= R$. Millä ehdolla taso $T: \ ax + by + cz + d = 0$ sivuaa palloa?

<u>Ratkaisu</u> Taso T sivuaa palloa täsmälleen, kun pallon keskipisteen etäisyys tasosta $= R$. Esimerkin 5 perusteella ehto on

$$|ax_0 + by_0 + cz_0 + d\,| = R\sqrt{a^2 + b^2 + c^2}. \quad \square$$

†Ympyrällä ja pallolla saatetaan tarkoittaa myös euklidisen avaruuden 'täyteistä' joukkoa

$$K = \{P \mathrel{\widehat{=}} \vec{r} \mid |\vec{r} - \vec{r}_0| \le R\}.$$

Tämän täsmällisempi nimitys tasossa on *kiekko* (engl. disc; ympyrä = circle), avaruudessa *kuula* (engl. ball; pallopinta = sphere).

Esimerkki 7 Yhtälöryhmä

$$\begin{cases} x^2 + y^2 + y^2 = R^2 \\ x + y - z = 1 \end{cases}$$

määrittelee origokeskisen pallon ja tason leikkausviivan S. Koska tason etäisyys origosta on $h = 1/\sqrt{3}$ (Esimerkki 5), niin päätellään, että S on avaruusympyrä jos $R > 1/\sqrt{3}$, piste jos $R = 1/\sqrt{3}$, ja tyhjä joukko jos $R < 1/\sqrt{3}$. Ensiksi mainitussa tapauksessa S:n keskipiste on origon kautta kulkevalla suoralla, jonka suuntavektori on tason normaalivektori $\vec{n} = \vec{i} + \vec{j} - \vec{k}$. $\square$

Lieriö ja kartio

Olkoon S avaruussuora, joka kulkee pisteen $P_0 \mathrel{\widehat{=}} \vec{r}_0$ kautta ja jonka suuntavektori on yksikkövektori $\vec{e}$. Tällöin pisteen $P \mathrel{\widehat{=}} \vec{r}$ etäisyys suorasta S voidaan laskea kaavalla

$$d = |\vec{r} - \vec{r}_0||\sin \measuredangle(\vec{r} - \vec{r}_0, \vec{e})| = |(\vec{r} - \vec{r}_0) \times \vec{e}|.$$

Pinta, jolla tämä etäisyys on vakio $= R$ $(R > 0)$ on nimeltään *lieriö*, tarkemmin *ympyrälieriö*. Suora S on lieriön *akseli* ja $R =$ lieriön *säde*. Asettamalla kaavassa $d = R$ ja neliöimällä saadaan lieriön yhtälölle muoto

$$|(\vec{r} - \vec{r}_0) \times \vec{e}|^2 = R^2. \tag{9}$$

Esimerkki 8 Olkoon lieriön akseli origon kautta kulkeva, yksikkövektorin $\vec{e} = \frac{1}{3}(\vec{i} + 2\vec{j} - 2\vec{k})$ suuntainen suora ja säde $R = 2$. Tällöin on $\vec{r}_0 = \vec{0}$ ja

$$\vec{r} \times \vec{e} = \frac{1}{3}\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x & y & z \\ 1 & 2 & -2 \end{vmatrix} = \frac{1}{3}\left[(-2y - 2z)\vec{i} + (2x + z)\vec{j} + (2x - y)\vec{k}\right],$$

joten lieriön yhtälö on

$$\begin{aligned} &\frac{1}{9}[(2y+2z)^2+(2x+z)^2+(2x-y)^2] \;=\; 4 \\ \Leftrightarrow \quad & 8x^2+5y^2+5z^2-4xy+4xz+8yz=36. \quad \square \end{aligned}$$

Jos yhtälössä (9) korvataan ristitulo pistetulolla ja kirjoitetaan R:n tilalle $\gamma|\vec{r}-\vec{r}_0|$, missä $\gamma \in (0,1)$, niin yhtälö saa muodon

$$|(\vec{r}-\vec{r}_0)\cdot\vec{e}\,|^2=\gamma^2|\vec{r}-\vec{r}_0|^2 \quad \Leftrightarrow \quad |\cos\measuredangle(\vec{r}-\vec{r}_0,\vec{e}\,)|=\gamma. \tag{10}$$

Tämä yhtälö määrittelee *kartion*, tarkemmin *ympyräkartion*. Piste $P_0 \mathrel{\widehat{=}} \vec{r}_0$ on kartion *kärki*. Kartio koostuu kahdesta *puolikartiosta*, joiden yhtälöt ovat

$$\cos\measuredangle(\vec{r}-\vec{r}_0,\vec{e}\,)=\pm\gamma.$$

Koordinaattien x, y, z avulla esitettynä yhtälöt (9) ja (10) voidaan kumpikin kirjoittaa muotoon

$$Ax^2+By^2+Cz^2+Dxy+Exz+Fyz+Gx+Hy+Iz+J=0,$$

missä A, B, jne. ovat reaalilukuja (vrt. esimerkki edellä). Tämä merkitsee, että lieriö ja kartio, myös pallo, ovat nk. *toisen asteen pintoja*. Toisen asteen pinnan ja avaruustason (esim. koordinaattitason) leikkausviiva on ko. tason *toisen asteen käyrä*. Tällainen on esimerkiksi ympyrä. Yleinen toisen asteen käyrän yhtälö xy-tasolla on muotoa

$$Ax^2+By^2+Cxy+Dx+Ey+F=0.$$

Toisen asteen käyrien ja pintojen yleisempi luokittelu on matemaattinen ongelma, johon palataan myöhemmin toisessa asiayhteydessä.

HARJOITUSTEHTÄVIÄ

1. a) Näytä, että yhtälöt esittävät samaa suoraa:

$$\begin{cases} x = 3 + 2t, \\ y = -11 - 6t \end{cases} \quad \text{ja} \quad \begin{cases} x = -1 - t, \\ y = 1 + 3t \end{cases}$$

b) Määritä suorien leikkauspiste:

$$\begin{cases} x = 3 + 2t, \\ y = -1 - 3t \end{cases} \quad \text{ja} \quad \begin{cases} x = -1 - t, \\ y = 2(1 + t) \end{cases}$$

c) Määritä suoran $x = 3+t,\ y = -2-t,\ z = 4-2t$ ja koordinaattitasojen leikkauspisteet.

2. Määritä α siten, että vektori $\vec{i} + 2\vec{j} + \alpha\vec{k}$ sekä avaruussuorat

$$2(x-1) = 1 - y = 2z - 3 \quad \text{ja} \quad \begin{cases} x = 17, \\ y = 7 + 3t, \\ z = 2t \end{cases}$$

ovat saman tason suuntaiset.

3. Määritä pisteet $P_1 \in S_1$ ja $P_2 \in S_2$ suorilla $S_1 :\ x = -y = z$ ja $S_2 :$ $x + y - 1 = 0,\ z = 0$ siten, että vektori $\overrightarrow{P_1P_2}$ on yhdensuuntainen vektorin $2\vec{i} - \vec{j} - \vec{k}$ kanssa.

4. Suora S kulkee pisteen $(-1, 1, 3)$ ja sen janan keskipisteen kautta, jonka xy- ja xz-tasot leikkaavat suorasta $x - 1 = 2(y + 1) = z + 3$. Määritä S:n suuntavektori.

5. Määritä α siten, että suorat $2(x-1) = y+1 = 2\alpha(z-1)$ ja $x+1 = y-1 = z$ leikkaavat. Mikä on suorien lyhin etäisyys, jos $\alpha = 1$?

6. Määritä suorien

$$\vec{r} = 2\vec{i} + 5\vec{k} + t(\vec{i} - 2\vec{j} + 2\vec{k}) \quad \text{ja} \quad \begin{cases} 3x - 2y = 12 \\ x + 2z = 6 \end{cases}$$

lyhin etäisyys ja lähinnä toisiaan olevat pisteet.

7. Määritä sen suoran parametrimuotoinen yhtälö, joka leikkaa kohtisuorasti suorat
$$\begin{cases} x = 1 + t, \\ y = 1 - 2t, \\ z = t \end{cases} \quad \text{ja} \quad \begin{cases} x = -1, \\ y = 2s, \\ z = 1 - 2s. \end{cases}$$

8. Laske pisteen $(2, 3, -1)$ etäisyys suorasta
$$\text{a) } \vec{r} = \vec{i} + 2\vec{j} + 3\vec{k} + t(\vec{i} - 2\vec{j} + 2\vec{k}) \qquad \text{b) } \begin{cases} 3x - y + z = 0 \\ x - 2y + 8 = 0 \end{cases}$$

9. a) Esitä tason T yhtälö perusmuodossa $ax + by + cz + d = 0$, kun parametrimuotoiset yhtälöt ovat
$$T: \begin{cases} x = 2 + 2t_1 + t_2, \\ y = -1 + 3t_1 + 2t_2, \\ z = 3 - t_1 + t_2. \end{cases}$$
b) Taso kulkee pisteen $P = (1, -13, -5)$ kautta ja sen suuntavektorit ovat $\vec{v}_1 = \vec{i} - \vec{j}$ ja $\vec{v}_2 = \vec{i} + \vec{j} + \vec{k}$. Onko piste $Q = (3, -1, 2)$ tasossa?

10. Määritä tason yhtälö (perusmuoto!), kun tiedetään, että taso sisältää suoran $x = 3 + t,\ y = 1 - 2t,\ z = -2 + t$ ja a) kulkee pisteen (0,2,1) kautta, b) on vektorin $3\vec{i} + \vec{j} - 2\vec{k}$ suuntainen

11. Määritä seuraavien tasojen yhteiset pisteet:
a) $x + y - z + 2 = 0,\ 2x + y + 2z - 4 = 0,\ x - y + 3z - 2 = 0$
b) $x + y + z - 6 = 0,\ x + 2y - z - 2 = 0,\ x + 4y - 5z + 5 = 0$
c) $x + 2y - 2z - 1 = 0,\ x - y + z - 2 = 0,\ x + 5y - 5z = 0$

12. Taso sisältää suoran $S_1:\ \vec{r} = (1 + t)\vec{i} + (1 + 2t)\vec{j} + (1 + 3t)\vec{k}$ ja on suoran $S_2:\ \vec{r} = (1 + t)\vec{i} + (-1 + t)\vec{j} + \vec{k}$ suuntainen. Johda tason yhtälö perusmuodossa. Johda samoin sen tason yhtälö, joka sisältää suoran S_2 ja on S_1:n suuntainen.

13. Johda sen tason yhtälö, joka puolittaa pisteen $P_0 = (x_0, y_0, z_0)$ ja tason $T:\ ax + by + cz + d = 0$ väliset janat.

14. Määritä pisteen $(3, 4, -2)$ kohtisuora projektio tasolla, jonka normaalivektori on $\vec{i} - 2\vec{j} + \vec{k}$ ja joka kulkee pisteen $(1, 1, 1)$ kautta.

15. Määritä pisteen $(3, 2, -4)$ peilikuvapiste tason $x + y - 2z + 5 = 0$ suhteen.

16. Määritä origon suurin mahdollinen etäisyys pisteiden $(0, 1, 0)$ ja $(2, 2, -1)$ kautta kulkevasta tasosta. Mikä taso antaa maksimietäisyyden?

17. Määritä tasojen $x + y + z = 3$ ja $3x - 2y - z = 1$ leikkaussuoran kautta kulkevat tasot, jotka puolittavat tasojen välisen kulman.

18. Kheopsin pyramidin (alkuperäinen) korkeus on 147 m ja neliön muotoisen pohjan sivun pituus 230 m. Sijoita pyramidi koordinaatistoon niin, että se tuntee olonsa mahdollisimman mukavaksi, ja määritä tässä koordinaatistossa pyramidin sivutasojen yhtälöt, mittayksikkönä 100 m. Laske myös vierekkäisten sivujen välinen diedrikulma (vrt. Harj.teht II.6:6).

19. Vektorin $2\vec{i}-\vec{j}+3\vec{k}$ suuntaan kulkeva valonsäde heijastuu tasosta T pisteessä $(1, 2, -1)$. Heijastunut säde kulkee pisteen $(2, 5, -3)$ kautta. Mikä on tason yhtälö?

20. Positiivisen z-akselin suunnasta tuleva valonsäde osuu pisteessä $(1, 2, 3)$ tasolla $3x+2y+z = 10$ olevaan peiliin. Määritä heijastuneen säteen suunta. Missä pisteessä heijastunut säde leikkaa (jos leikkaa) xy-tason?

21. Painovoima vaikuttaa negatiivisen z-akselin suuntaan. Pisara putoaa pisteestä $(1, 1, 3)$ tasolle $3x-4y+12z = 12$ ja alkaa valua tasoa pitkin alaspäin. Missä pisteessä pisara kohtaa xy-tason?

22. Tasangolta $z = 0$ kohoaa vuorenrinne pitkin tasoa $x+2y+4z = 0$. Rinteen pisteestä P, joka on korkeudella $h = 10$, lähtee liikkeelle pistemäinen lumivyöry. Se etenee suoraviivaisesti rinnettä alas painovoimalakien mukaisesti. Tasangolle saavuttuaan se jatkaa suoraviivaista liikettään vaakasuoraan, kunnes osuu mökkiin, joka on pisteessä $Q = (10, 40, 0)$. Mikä oli piste P?

23. (Haukka ja kaksi kanaa) Universaalikoordinaatistossa maan pinta on taso $x + 2y - 3z = 0$. Maan pinnalla käyskentelee kaksi pistemäistä kanaa A ja B. Kanoja vaanii pistemäinen haukka, joka lentää maan pinnan suuntaisella tasolla korkeudella $h = 5$. Pistemäinen aurinko loistaa suunnassa $-4\vec{i} + \vec{k}$. Hetkellä H tapahtuu seuraavaa: Kanalta A pääsee säikähtynyt 'kot' sen huomatessa haukan varjon päällään. Kotkotuksen kuulee kana B, joka vilkaisee samassa taivaalle ja näkee haukan suunnassa $\vec{i}-\vec{j}+\vec{k}$. Määritä vektori $\overrightarrow{AB}$ kyseisellä hetkellä H.

24. Missä kulmassa tason suora $y = 4x$ leikkaa ympyräviivan $x^2 + y^2 - 2x - 4y + 4 = 0$?

25. Avaruuskolmion kärjet ovat $(0, 0, 0)$, $(3, 2, 1)$ ja $(2, -1, 3)$. Laske kolmion sisään piirretyn (eli kaikkia sivuja sivuavan) ympyrän keskipiste ja säde.

26. Avaruuden E^3 kolmen pisteen paikkavektorit ovat $\vec{a}$, $\vec{b}$ ja $\vec{c}$. Esitä menettely, jolla voidaan määrittää pisteiden kautta kulkevan ympyrän keskipisteen paikkavektori. Sovella menettelyä, kun pisteet ovat $(1, 2, 3)$, $(2, -5, 3)$ ja $(-1, 3, -6)$. Määritä myös ympyrän säde ja ympyrän tason normaalivektori.

27. Määritä avaruusympyrän

$$\begin{cases} x^2 + y^2 + z^2 = 49 \\ x + 2y - z = 10 \end{cases}$$

tangentti pisteessä $(-2, 3, -6)$.

28. Lieriön säde on $R = 2$ ja akseli on pisteen $P_0 = (1, 1, 1)$ kautta kulkeva suora, jonka suuntavektori on $-\vec{i}+\vec{j}+\vec{k}$. Määritä lieriön yhtälö toisen asteen pinnan yhtälön perusmuodossa. Mitkä ovat lieriön ja suoran $x = y = z$ leikkauspisteet, ja mikä lieriön piste on lähinnä origoa?

29. Lieriöllä ja kartiolla on yhteisenä akselina suora $S: \ x = 2y = -2z$ ja kumpikin kulkee pisteen $(1, 0, 0)$ kautta. Kartion kärkenä on piste $(-2, -1, 1)$. Laske lieriön säde ja kartion (sivuprofiilin) aukeamiskulma sekä saata kummankin pinnan yhtälöt toisen asteen pinnan yhtälön perusmuotoon. Määritä edelleen molempien pintojen ja xy-tason leikkauskäyrien yhtälöt ja hahmottele näiden käyrien muoto.

30. (*) Kuution, jonka sivun pituus = 4, yksi kärki on origossa ja kolme muuta positiivisilla koordinaattiakseleilla. Kuutiota katsotaan kaukaa vektorin $\vec{i}+2\vec{j}+3\vec{k}$ osoittamasta suunnasta kuvakulmassa, jossa z-akseli näkyy pystysuorana. Laske, millaisena kuutio näkyy tästä kuvakulmasta. Kuva!

31. (*) Millä tavoin saadaan selville avaruustasot, jotka sivuavat kolmea annettua palloa? Montako tällaista tasoa on, jos pallot eivät leikkaa tai sivua toisiaan eikä mikään palloista ole toisen sisällä?

32. (*) Näytä, että yhtälö $K: \ xy+yz+xz = 0$ määrittelee kartion, ja määritä se K:n piste, joka on lähinnä pistettä $(-1, 2, 3)$.

33. (*) Kartion kärki on origossa, symmetria-akseli on vektorin $\vec{i}-2\vec{j}+2\vec{k}$ suuntainen ja y-akseli on kartiopinnalla. Taso T kulkee pisteen $(1, 1, 1)$ kautta ja sivuaa kartiota pitkin avaruussuoraa. Määritä T:n yhtälö (kaksi ratkaisua!).

II.8 Käyräviivaiset koordinaatistot

Euklidinen taso ja avaruus ovat matemaattisina käsitteinä aineettomia tai *isotrooppisia* siinä mielessä, että mikään piste tai suunta ei niissä ole erikoisasemassa. Erityisesti voidaan karteesisen koordinaatiston origo, ja myös koordinaattiakselien suunnat ortogonaalisuusehtojen puitteissa, valita mielivaltaisesti, ilman että avaruuden geometriset lait siitä muuttuisivat. Sovellettaessa geometrista ajattelua fysikaaliseen maailmaan törmätään kuitenkin siihen realiteettiin, että maailmassa yleensä on materiaa, joka ei jakaudu tasaisesti. Voi esimerkiksi esiintyä eri materiaalien välisiä rajapintoja. Fysikaalinen ongelma voi olla ei-isotrooppinen myös siinä vaikuttavien ulkoisten voimien (kuten gravitaation) vuoksi. Tällaisissakin tilanteissa ongelmassa voi kuitenkin olla *symmetrioita*, millä tarkoitetaan, että vaikka koko avaruus ei ole isotrooppinen, niin fysikaalisen ongelman geometria kuitenkin on johonkin suuntaan tai joihinkin suuntiin liikuttaessa sama. Esimerkiksi 'pannukakkumaailmassa', jossa puoliavaruus

$$E^3_- = \{P = (x, y, z) \in E^3 \mid z \leq 0\}$$

on homogeenisen materian täyttämä ja muu osa avaruudesta on tyhjää (tai homogeenista toista ainetta), geometria säilyy samana liikuttaessa tasoilla, joilla $z =$ vakio. Toisin sanoen: geometria on koordinaateista x, y riippumaton. Tässä on kyseessä *tasosymmetrinen* tilanne. Hieman realistisemmassa (vaikka edelleen maakeskisessä) 'pallomaailmassa', jossa materian täyttämää palloa ympäröi homogeeninen avaruus, geometria on puolestaan *pallosymmetrinen*. Lopulta 'tekniikan maailmassa', jossa avaruuden homogeenisuuden rikkoo yksinäinen, poikkipinnaltaan pyöreä kappale (maitohinkki, putki, koaksiaalikaapeli ym.) tilanne on *lieriösymmetrinen*.

Lieriösymmetrian vastine E^2:ssa on *ympyräsymmetria* (pyörähdyssymmetria).

Em. tasosymmetrisessä tilanteessa on karteesinen koordinaatisto probleeman geometriaan sopiva, kunhan yksi kantavektoreista (esim. $\vec{k}$) valitaan symmetriatason normaaliksi. Muissa esimerkkitilanteissa sen sijaan tulee kyseeseen siirtyminen sellaiseen *käyräviivaiseen* (engl. curvilinear) koordinaatistoon, jossa jokin koordinaateista on vakio materiaalirajapinnalla. Koordinaatistoa sanotaan käyräviivaiseksi, jos *koordinaattiviivat*, eli E^2:n tai E^3:n pistejoukot, joilla vain yksi koordinaateista on muuttuva, eivät kaikki ole suoria. Seuraavassa esitellään fysiikan kolme tavallisinta käyräviivaista koordinaatistoa, jotka liittyvät em. esimerkkitilanteisiin. Nämä ovat:

- *napa-* eli *polaarikoordinaatisto* (E^2)
- *lieriö-* eli *sylinterikoordinaatisto* (E^3)
- *pallokoordinaatisto* (E^3)

Tällaisia koordinaatistoja sovellettaessa (niinkuin yleensäkin matematiikkaa sovellettaessa) tehdään yleensä käytännön ongelmatilannetta koskevia yksinkertaistavia olettamuksia eli idealisaatioita – joskus voimakkaita.

Napakoordinaatisto

Napakoordinaatistossa euklidisen tason pisteen P koordinaatit ilmoitetaan origon (O) ja P:n välisellä etäisyydellä r ja nk. *napakulmalla* φ, joka mittaa vektorin $\overrightarrow{OP}$ ja positiivisen x-akselin (karteesinen koordinaatisto) välistä kulmaa, positiivisena x-akselista lähtien vastapäivään ja negatiivisena myötäpäivään. Kaikki tason pisteet saadaan, kun napakulma valitaan esim. väliltä $[0, 2\pi)$ tai $(-\pi, \pi]$. Koska origossa napakulma φ ei ole määritelty, niin sovitaan, että $O = (0, \varphi)\ \forall\varphi$.

Muunnoskaavat napakoordinaatistosta karteesiseen ovat

$$\boxed{x = r\cos\varphi, \quad y = r\sin\varphi}$$

ja päinvastaiseen suuntaan esimerkiksi

$$\boxed{r = \sqrt{x^2+y^2}\,, \quad \sin\varphi = y/\sqrt{x^2+y^2}.}$$

Tässä on φ:n määräämiseksi osattava 'kääntää' trigonometrinen sini. Kun huomioidaan myös x:n ja y:n etumerkit, niin kulma φ määräytyy yksikäsitteisesti esim. välillä $[0, 2\pi)$, kun $r > 0$.

Esimerkki 1 Millainen pistejoukko on

$$S = \{P = (r,\varphi) \in E^2 \mid r = 2\cos\varphi,\ \varphi \in [-\tfrac{\pi}{2}\,, \tfrac{\pi}{2}]\}\,?$$

<u>Ratkaisu</u> Jos $r > 0$, niin

$$\begin{aligned} r = 2\cos\varphi \ &\Leftrightarrow\ r^2 = 2r\cos\varphi \\ &\Leftrightarrow x^2 + y^2 = 2x \\ &\Leftrightarrow (x-1)^2 + y^2 = 1. \end{aligned}$$

Myös origo on käyrässä mukana ($r = 0,\ \varphi = \pm\frac{\pi}{2}$), joten kyseessä on ympyrä, jonka säde $= 1$ ja keskipiste $= (1, 0)$. $\square$

Napakoordinaatistossa koordinaattiviivat, joilla $r =$ vakio, ovat origokeskisiä ympyröitä ja viivat, joilla $\varphi =$ vakio, ovat origosta lähteviä puolisuoria. Koska nämä leikkaavat toisensa kohtisuorasti (ks. kuvio), niin voidaan sanoa, että polaarikoordinaatisto on *suorakulmainen*, vaikkakaan ei suoraviivainen koordinaatisto.

Kun napakoordinaatistossa tarkastellaan vektoreita, ilmaistaan nämä yleensä tarkkailupisteen P mukana 'pyörivässä' kannassa $\{\vec{e}_r, \vec{e}_\varphi\}$, missä kantavektorit ovat koordinaattiviivojen yksikkötangenttivektorit P:ssä. Suunnat valitaan siten, että vektorit osoittavat ko. koordinaattiviivalla muuttuvan koordinaatin kasvun suuntaan:

Jos pisteen P napakoordinaatit ovat (r, φ), niin kantavektorit ovat (vrt. kuva)

$$\boxed{\vec{e}_r = \cos\varphi\vec{i} + \sin\varphi\vec{j}, \quad \vec{e}_\varphi = -\sin\varphi\vec{i} + \cos\varphi\vec{j}.}$$

Jos P ajatellaan kiinnitetyksi, voidaan mikä tahansa tason vektori esittää kannassa $\{\vec{e}_r, \vec{e}_\varphi\}$, eli $\{P, \vec{e}_r, \vec{e}_\varphi\}$ on tason koordinaatisto. Toisaalta jos P:tä muutetaan siten, että koordinaatti φ muuttuu, niin muuttuvat myös $\vec{e}_r$ ja $\vec{e}_\varphi$. Koordinaatisto $\{P, \vec{e}_r, \vec{e}_\varphi\}$ poikkeaa siis tässä mielessä vastaavasta suoraviivaisesta koordinaatistosta $\{P, \vec{i}, \vec{j}\}$, jossa kantavektorit $\vec{i}, \vec{j}$ pysyvät vakioina referenssipisteen P muuttuessa.

Tyypillinen fysikaalinen sovellustilanne pisteen P mukana 'pyörivälle' koordinaatistolle $\{P, \vec{e}_r, \vec{e}_\varphi\}$ syntyy, kun P edustaa <u>keskeisvoimakentässä</u> liikkuvaa kappaletta (esim. planeetta, avaruusalus).

Esimerkki 2 **Ceres ja Dawn**. Karteesisen koordinaatiston origossa on kääpiöplaneetta, ja planeetasta suuntaan $3\vec{\imath} + 4\vec{\jmath}$ olevassa pisteessä P on avaruusluotain, jonka massa on $m = 1200$ kg. Luotain kiertää planeettaa tasossa $z = 0$ olevalla ympyräradalla vastapäivään positiivisen z-akselin suunnasta katsottuna. Tarkasteluhetkellä luotaimen ionimoottori käynnistetään, jolloin luotaimeen vaikuttaa planeetan vetovoiman $\vec{G} = -78\vec{e}_r$ lisäksi moottorin työntövoima $\vec{T} = 450(-\vec{\imath} + \vec{\jmath})$ (yksiköt N=kg m/s^2). Mihin suuntaan luotaimen liikerata kaartuu, ja mikä on luotaimen kiihtyvyys liikeradan suuntaan?

Ratkaisu Pisteessä P on $\vec{e}_r = \frac{1}{5}(3\vec{\imath} + 4\vec{\jmath})$, ja luotaimen liikesuunta on $\vec{e}_\varphi = \frac{1}{5}(-4\vec{\imath} + 3\vec{\jmath})$. Koska

$$\begin{cases} (\vec{T} + \vec{G}) \cdot \vec{e}_r = \frac{450}{5}[(-1) \cdot 3 + 1 \cdot 4] - 78 = 12, \\ (\vec{T} + \vec{G}) \cdot \vec{e}_\varphi = \frac{450}{5}[(-1) \cdot (-4) + 1 \cdot 3] + 0 = 630, \end{cases}$$

niin luotaimeen vaikuttava kokonaisvoima koordinaatistossa $\{P, \vec{e}_r, \vec{e}_\varphi\}$ on

$$\vec{F} = \vec{T} + \vec{G} = 12\vec{e}_r + 630\vec{e}_\varphi = F_r\vec{e}_r + F_\varphi\vec{e}_\varphi.$$

Koska $F_r > 0$, niin rata kaartuu oikealle (poispäin planeetasta). Ratakiihtyvyys määräytyy Newtonin liikelaista: $a_\varphi = F_\varphi/m \approx \underline{\underline{0.53 \text{ m/s}^2}}$. $\square$

Lieriökoordinaatisto

Lieriökoordinaatisto on napakoordinaatiston vastine kolmessa dimensiossa. Karteesisista koordinaateista x, y, z jätetään z ennalleen ja muunnetaan $(x, y) \mapsto (r, \varphi)$ samalla tavoin kuin napakoordinaatistoon siirryttäessä. Lieriökoordinaatiston 'pyöriviä' kantavektoreita merkitään $\vec{e}_r, \vec{e}_\varphi, \vec{e}_z$. Tässä $\vec{e}_r, \vec{e}_\varphi$ ovat samat kuin polaarikoordinaatistossa ja $\vec{e}_z = \vec{k}$ kuten karteesisessa koordinaatistossa. Kanta $\{\vec{e}_r, \vec{e}_\varphi, \vec{e}_z\}$ on ortonormeerattu oikeakätinen systeemi.

Esimerkki 3 *Ruuviviiva* on avaruuden E^3 pistejoukko

$$S = \{P = (r, \varphi, z) \mid r = R,\ z = a\varphi\} \quad (a \neq 0). \quad \square$$

Lieriökoordinaatiston koordinaattiviivat ovat E^3:n pistejoukkoja, joissa kaksi koordinaateista r, φ, z saa vakioarvon. Näille saadaan seuraavat geometriset luonnehdinnat (vrt. kuvio edellä):

S_r (φ, z vakioita) : z-akselilta lähtevä, xy-tason suuntainen puolisuora,
S_φ (r, zvakioita) : ympyrä xy-tason suuntaisella tasolla,
S_z (r, φ vakioita) : z-akselin suuntainen suora.

Koordinaattipinnoiksi sanotaan sellaisia E^3:n pistejoukkoja, joilla yksi koordinaateista saa vakioarvon. Lieriökoordinaatistossa näiden geometriset luonnehdinnat ovat

$r =$ vakio : lieriö, akselina z-akseli,
$\varphi =$ vakio : z-akseliin rajoittuva puolitaso,
$z =$ vakio : xy-tason suuntainen taso.

Pallokoordinaatisto

Pisteen $P = (x, y, z)$ *pallokoordinaatit* ovat (ks. kuvio)

- P:n etäisyys origosta: $r = \sqrt{x^2 + y^2 + z^2}$
- paikkavektorin $\vec{r} = \overrightarrow{OP}$ ja positiivisen z-akselin (vektorin $\vec{k}$) muodostama kulma θ, $0 \leq \theta \leq \pi$
- kulma φ, $0 \leq \varphi < 2\pi$, jonka muodostavat positiivinen x-akseli (vektori $\vec{i}$) ja vektorin $\vec{r} = \overrightarrow{OP}$ ortogonaaliprojektio xy-tasolle

Koordinaatteja θ, φ sanotaan *pallonpintakoordinaateiksi.*

Muunnoskaavat ovat

$$\boxed{\begin{aligned} x &= r\sin\theta\cos\varphi, \\ y &= r\sin\theta\sin\varphi, \\ z &= r\cos\theta. \end{aligned}}$$

Pallokoordinaatiston koordinaattiviivat ovat seuraavaa tyyppiä:

S_r (θ, φ vakioita) : origosta lähtevä puolisuora,
S_θ (r, φ vakioita) : origokeskinen, z-akseliin rajoittuva puoliympyrä,
S_φ (r, θ vakioita) : ympyrä xy-tason suuntaisella tasolla.

Koordinaattipinnat voidaan luokitella seuraavasti:

$r =$ vakio : origokeskinen pallo,
$\theta =$ vakio : puolikartio, kärkenä origo,
$\varphi =$ vakio : z-akseliin rajoittuva puolitaso.

Myös pallokoordinaatistossa koordinaattiviivat leikkaavat toisensa kohtisuorasti, ts. koordinaattiviivojen tangentit (ks. edellinen luku) ovat leikkauspisteessä parittain kohtisuorat. Koordinaattiviivojen yksikkötangenttivektoreista (tangenttien suuntavektoreista) muodostettua avaruuden vektorien kantaa merkitään $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$ ja määritellään (vrt. kuvio edellä)

$$\boxed{\begin{aligned} \vec{e}_r &= \sin\theta\cos\varphi\vec{i} + \sin\theta\sin\varphi\vec{j} + \cos\theta\vec{k}, \\ \vec{e}_\theta &= \cos\theta\cos\varphi\vec{i} + \cos\theta\sin\varphi\vec{j} - \sin\theta\vec{k}, \\ \vec{e}_\varphi &= -\sin\varphi\vec{i} + \cos\varphi\vec{j}. \end{aligned}}$$

Tässä $\vec{e}_\varphi$ on sama kuin polaari- ja lieriökoordinaatistoissa ja $\vec{e}_r$:n lauseke seuraa välittömästi edellä esitetyistä koordinaattien muunnoskaavoista:

$$\begin{aligned} \vec{e}_r &= \frac{1}{|\vec{r}|}\vec{r} = \frac{1}{r}(x\vec{i} + y\vec{j} + z\vec{k}) \\ &= \sin\theta\cos\varphi\vec{i} + \sin\theta\sin\varphi\vec{j} + \cos\theta\vec{k}. \end{aligned}$$

Tämän jälkeen $\vec{e}_\theta$ on laskettavissa tiedosta, että $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$ on ortonormeerattu, oikeakätinen systeemi:

$$\vec{e}_\theta = \vec{e}_\varphi \times \vec{e}_r.$$

Esimerkki 4 Pisteen P karteesiset koordinaatit ovat $(x, y, z) = (-2, -3, -6)$. Määritä P:n pallokoordinaatit sekä vektorin $\vec{v} = \vec{i} - \vec{j}$ koordinaatit $(v_r, v_\theta, v_\varphi)$ pallokoordinaatiston kannassa $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$ ko. pisteessä.

Ratkaisu Jos merkitään $\vec{r} = \overrightarrow{OP} = -2\vec{i} - 3\vec{j} - 6\vec{k}$, niin $r = |\vec{r}\,| = 7$ ja

$$\vec{e}_r = \frac{1}{r}\,\vec{r} = \frac{1}{7}(-2\vec{i} - 3\vec{j} - 6\vec{k}).$$

Tällöin

$$\begin{aligned}
\cos\theta &= \vec{e}_r \cdot \vec{k} = -\frac{6}{7}\,, \quad \sin\theta = \frac{\sqrt{13}}{7}\,, \\
\cos\varphi &= \frac{\vec{i}\cdot(-2\vec{i} - 3\vec{j})}{\sqrt{2^2 + 3^2}} = -\frac{2}{\sqrt{13}}\,, \quad \sin\varphi = -\frac{3}{\sqrt{13}}\,,
\end{aligned}$$

joten P:n pallokoordinaatit ovat

$$(r, \theta, \varphi) \approx (7,\ 149.0^\circ,\ 236.3^\circ).$$

Kantavektoreiden $\vec{e}_\theta, \vec{e}_\varphi$ tarkat arvot ovat

$$\begin{aligned}
\vec{e}_\theta &= -\frac{6}{7}\left(-\frac{2}{\sqrt{13}}\right)\vec{i} - \frac{6}{7}\left(-\frac{3}{\sqrt{13}}\right)\vec{j} - \frac{\sqrt{13}}{7}\vec{k} \\
&= \frac{1}{7\sqrt{13}}\,(12\vec{i} + 18\vec{j} - 13\vec{k}), \\
\vec{e}_\varphi &= \frac{1}{\sqrt{13}}\,(3\vec{i} - 2\vec{j}),
\end{aligned}$$

joten kysytyt $\vec{v}$:n koordinaatit ovat

$$\begin{aligned}
v_r &= \vec{v}\cdot\vec{e}_r = \frac{1}{7} \approx 0.143, \\
v_\theta &= \vec{v}\cdot\vec{e}_\theta = -\frac{6}{7\sqrt{13}} \approx -0.238, \\
v_\varphi &= \vec{v}\cdot\vec{e}_\varphi = \frac{5}{\sqrt{13}} \approx 1.387.
\end{aligned}$$

Tarkistus:

$$v_r^2 + v_\theta^2 + v_\varphi^2 = \frac{1274}{13\cdot 49} = 2 = |\vec{v}\,|^2. \quad \text{OK!} \quad \square$$

Esimerkki 5 : **Lentosuunta**. Lentokone lähtee päiväntasaajan pisteestä $A : 90^\circ$ läntistä pituutta ja lentää pitkin isoympyrää, joka kulkeee pisteen $B : 45^\circ$ itäistä pituutta, 60° pohjoista leveyttä, kautta. Määritä koneen lentosuunta pisteessä B sekä kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$ että suhteessa ilmansuuntiin pisteessä B.

<u>Ratkaisu</u> Oletetaan pallon säteeksi $R = 1$, jolloin (ks. kuva)

$$\begin{aligned}\overrightarrow{OA} &= -\vec{j},\\ \overrightarrow{OB} &= \sin 30^\circ \cos 45^\circ \vec{i} + \sin 30^\circ \sin 45^\circ \vec{j} + \cos 30^\circ \vec{k}\\ &= \frac{1}{2\sqrt{2}}\vec{i} + \frac{1}{2\sqrt{2}}\vec{j} + \frac{\sqrt{3}}{2}\vec{k}.\end{aligned}$$

Koneen lentorata on tasossa, joka kulkee pisteiden O, A, B kautta, joten tason normaalivektori on

$$\vec{n} = \overrightarrow{OA} \times \overrightarrow{OB} = -\frac{\sqrt{3}}{2}\vec{i} + \frac{1}{2\sqrt{2}}\vec{k}.$$

Koska lentorata pisteessä B on myös kohtisuorassa vektoria $\overrightarrow{OB}$ vastaan, on lentosuunnan oltava sama kuin vektorilla

$$\vec{n} \times \overrightarrow{OB} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -\frac{\sqrt{3}}{2} & 0 & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{\sqrt{3}}{2} \end{vmatrix} = -\frac{1}{8}\vec{i} + \frac{7}{8}\vec{j} - \frac{1}{4}\sqrt{\frac{3}{2}}\vec{k}.$$

Tämän vektorin pituus on $\sqrt{56}/8$, joten lentosuunnan osoittava yksikkövektori on

$$\underline{\underline{\vec{v}^{\,0} = \frac{1}{\sqrt{56}}(-\vec{i} + 7\vec{j} - \sqrt{6}\vec{k}).}}$$

Itään osoittava yksikkövektori pisteessä B on

$$\vec{e}_\varphi = -\sin 45^\circ \vec{i} + \cos 45^\circ \vec{j} = \frac{1}{\sqrt{2}}(-\vec{i} + \vec{j}),$$

ja etelään osoittava yksikkövektori on

$$\begin{aligned}\vec{e}_\theta &= \cos 30^\circ \cos 45^\circ \vec{i} + \cos 30^\circ \sin 45^\circ \vec{j} - \sin 30^\circ \vec{k} \\ &= \frac{1}{2}\sqrt{\frac{3}{2}}\,(\vec{i} + \vec{j}) - \frac{1}{2}\vec{k}.\end{aligned}$$

Koska

$$\begin{aligned}\vec{v}^{\,0} \cdot \vec{e}_\varphi &= \frac{2}{\sqrt{7}} \approx \cos 41^\circ, \\ \vec{v}^{\,0} \cdot \vec{e}_\theta &= \sqrt{\frac{3}{7}} \approx \cos 49^\circ,\end{aligned}$$

on lentosuunta B:ssä kaakosta hieman itään päin. □

HARJOITUSTEHTÄVIÄ

1. Seuraavassa on annettu piste P joko karteesisessa, lieriö- tai pallokoordinaatistossa. Laske P:n koordinaatit muissa kahdessa koordinaatistossa (tarkasti, jos mahdollista, muuten likiarvoina).

$P = (x, y, z)$:	a) $(1, 1, 1)$	b) $(2, 3, -1)$	c) $(-1, -\sqrt{3}, -5)$
$P = (r, \varphi, z)$:	d) $(\sqrt{2}, \frac{\pi}{2}, 1)$	e) $(1, \frac{\pi}{3}, -1)$	f) $(6\sqrt{2}, \frac{7\pi}{4}, 3)$
$P = (r, \theta, \varphi)$:	g) $(1, \frac{\pi}{2}, \pi)$	h) $(\sqrt{3}, \frac{3\pi}{4}, \frac{3\pi}{4})$	i) $(3, \frac{2\pi}{3}, \frac{5\pi}{6})$

2. a) Tutki, millainen on polaarikoordinaattien avulla ilmaistu tasokäyrä

$$S = \{P = (r, \varphi) \mid r = \cos\varphi + \sin\varphi,\ \varphi \in [-\tfrac{\pi}{4}, \tfrac{3\pi}{4}]\}.$$

b) Tason ympyrän S keskipiste on (x_0, y_0) ja säde $= R$. Näytä, että $P = (x, y) \in S$ täsmälleen kun jollakin $\varphi \in [0, 2\pi)$ pätee

$$\begin{cases} x = x_0 + R\cos\varphi, \\ y = y_0 + R\sin\varphi. \end{cases}$$

Mikä on vastaava väittämä koskien avaruuden pallopintaa, jonka keskipiste on (x_0, y_0, z_0) ja säde R?

3. Pisteiden P_1 ja P_2 pallokoordinaatteja $(r_1, \theta_1, \varphi_1)$ ja $(r_2, \theta_2, \varphi_2)$ verrataan joukon $[0, \infty) \times [0, \pi] \times [0, 2\pi) \subset \mathbb{R}^3$ alkioina. Luettele tapaukset, joissa vertailu antaa tuloksen $P_1 = P_2$, vaikka koordinaatit eivät ole samat.

4. Pisteessä P, jonka karteesiset koordinaatit tunnetaan, vaikuttaa voima $\vec{F}$. Muunna $\vec{F}$ annettuun kantaan ko. pisteessä.
a) $P = (-2, 3, 6)$, $\vec{F} = \vec{i} - \vec{j} + \vec{k}$, $\{\vec{e}_r, \vec{e}_\varphi, \vec{e}_z\}$
b) $P = (-2, 3, 6)$, $\vec{F} = \vec{i} + \vec{j} + \vec{k}$, $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$
c) $P = (6, -6, -3)$, $\vec{F} = -9\vec{e}_r$ (pallok.), $\{\vec{i}, \vec{j}, \vec{k}\}$
d) $P = (-6, -6, 3)$, $\vec{F} = -9\vec{e}_r$ (lieriök.), $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$
e) $P = (-1, \sqrt{3}, -2\sqrt{3})$, $\vec{F} = \vec{e}_\theta - \vec{e}_\varphi$ (pallok.), $\{\vec{i}, \vec{j}, \vec{k}\}$

5. Karteesisen koordinaatiston origo on maapallon keskipisteessä ja vektori $\vec{k}$ osoittaa pohjoisnavalle päin. Maapallon pinnalla eteläisen pallonpuoliskon pisteessä, jonka pallonpintakoordinaatit ovat $\theta = 135°$, $\varphi = 120°$, laiva on matkalla luoteeseen ja lentokone koilliseen. Määritä laivan ja lentokoneen liikesuunnat yksikkövektoreina a) pallokoordinaatistossa vektorien $\{\vec{e}_\theta, \vec{e}_\varphi\}$ avulla, b) karteesisessa koordinaatistossa vektorien $\{\vec{i}, \vec{j}, \vec{k}\}$ avulla.

6. Helsingistä $(60°N, 25°E)$ lennetään lyhintä tietä Tokioon $(36°N, 140°E)$. Miten pitkä on matka ja mihin ilmansuuntaan on Helsingistä lähdettävä? Maapallon säde on 6370 km.

7. (*) Pallokoordinaatin θ avulla ilmaistu puolikartio $\theta = 30°$ leikataan tasolla $T : y + z - 4 = 0$. Ilmaise syntyvä leikkauskäyrä $S \subset E^3$ ensin pallokoordinaattien (r, θ, φ) avulla. Muunna sitten käyrä tason T napakoordinaatistoon, jossa origo on pisteessä $(0, 0, 4)$ ja napakulmaa mitataan sunnasta $\vec{i}$, ja edelleen T:n vastaavaan karteesiseen koordinaatistoon. Hahmottele käyrä viimeksi mainitussa koordinaatistossa.

Luku III

Kompleksiluvut

Siirtyminen reaaliluvista kompleksilukuihin on matemaattisen analyysin merkittävimpiä ja samalla merkillisimpiä aluevaltauksia. Kyse on lukualueen laajennuksesta, ts. siirtymisestä jälleen uudelle 'todellisuuden' tasolle. Vaikka laajennusta voi pitää vain kuvitelmana, niin tämä kuvitelma on yksinkertaistanut matemaattista ajattelua siinä määrin, että sillä on lopulta ollut syvällinen vaikutus kaikkeen matematiikkaan, myös käytännön laskentamenetelmiin. Matemaattisen analyysin perinteessä lukualueen laajennus korostuu käsitteissä *reaalianalyysi* ja *kompleksianalyysi*. Molemmat ovat nykyään hyvin laajoja (ja hieman epämääräisiä) matematiikan alueita. Kompleksianalyysin osa-alueista maininnan arvoinen on kompleksimuuttujan funktioiden teoria eli *funktioteoria*[†].

[†]Funktioteorian tutkimusperinne on Suomessa vahva. Tätä matematiikan suuntausta edusti myös Suomen historian tunnetuin matemaatikko, akateemikko **Rolf Nevanlinna** (1895-1980).

III.1 Osoitinkunta

Lähtökohtana on euklidinen taso ja siihen pystytetty karteesinen koordinaatisto, koordinaatteina (x, y). Kutsutaan tällä kertaa *osoittimeksi* suuntajanaa, jolla on tietty pituus ja suunta. Pituutta merkitään symbolilla r ja suunta mitataan x-akselin suunnasta vastapäivään kiertäen *vaihekulmalla* (napakulmalla) φ. Merkitään

$$\vec{p} = r\angle\varphi\,.$$

Osoitin on siis toistaiseksi täsmälleen samanlainen olio kuin vektori: sillä on pituus ja suunta. Osoittimet myös samastetaan kuten vektorit. On huomioitava ainoastaan, että φ ja $\varphi + 2\pi$ vastaavat samaa suuntaa, joten samastussäännöt ovat

$$r_1\angle\varphi_1 = r_2\angle\varphi_2 \quad \Leftrightarrow \quad \begin{cases} r_1 = r_2 = 0 \ \text{ tai} \\ r_1 = r_2 > 0 \ \wedge \ \varphi_1 - \varphi_2 = k \cdot 2\pi, \quad k \in \mathbb{Z}. \end{cases}$$

Tapauksessa $r = 0$ samastetaan siis kaikki suunnat, kuten nollavektorissa. Osoitinta kutsutaankin tässä tapauksessa *nollaosoittimeksi* ja merkitään

$$0\angle\varphi = \vec{0}\,.$$

Osoittimien joukkoa merkitään jatkossa $\mathbb{P}$:llä:

$$\mathbb{P} = \{\vec{p} = r\angle\varphi \mid r \in [0, \infty),\ \varphi \in \mathbb{R}\}.$$

Osoittimien yhteenlaskuoperaatio $(+)$ määritellään samalla tavoin kuin vektoreille, eli kolmiodiagrammin avulla:

Myös skalaareilla eli reaaliluvuilla kertominen suoritetaan samoin kuin vektoreilla. Erityisesti jos $\lambda \in \mathbb{R}$ ja $\lambda > 0$, niin

$$\lambda(r\angle\varphi) = \lambda r\angle\varphi\,, \quad \lambda > 0.$$

Tapauksessa $\lambda = 0$ on kertolaskun tulos nollaosoitin. Jos $\lambda \in \mathbb{R}$ ja $\lambda < 0$, niin määritellään (vrt. vektorit)

$$\begin{aligned}\lambda(r\angle\varphi) &= |\lambda| \cdot (-1)(r\angle\varphi)\\ &= |\lambda| r\angle(\varphi+\pi)\,, \quad \lambda < 0.\end{aligned}$$

Tässä $r\angle(\varphi+\pi)$ on yhteenlaskun määritelmän mukaisesti osoittimen $\vec{p} = r\angle\varphi$ *vastaosoitin*:

$$-\vec{p} = (-1)\vec{p} = r\angle(\varphi+\pi)\,.$$

Tähänastisen perusteella osoittimien yhteenlaskusta ja skaalauksesta muodostuva algebra $(\mathbb{P}, +, \mathbb{R})$ näyttää yksinkertaisesti tason vektoriavaruudelta. Näin onkin toistaiseksi, ja vektorianalogia voidaan viedä hiukan pidemmällekin: Otetaan käyttöön tason vektoriavaruuden ortonormeerattua kantaa $\{\vec{i}, \vec{j}\}$ vastaava osoitinkanta, jossa kantaosoittimia merkittäköön

$$\begin{aligned}\vec{i} &\leftrightarrow \vec{e} = 1\angle 0,\\ \vec{j} &\leftrightarrow \vec{i} = 1\angle\pi/2.\end{aligned}$$

Jos nyt $\vec{p} = r\angle\varphi \in \mathbb{P}$, niin $\vec{p}$ voidaan ilmaista kantaosoittimien avulla yksikäsitteisesti muodossa

$$\begin{aligned}\vec{p} = r\angle\varphi \;&=\; r\cos\varphi\,\vec{e} + r\,\sin\varphi\,\vec{i}\\ &=\; x\vec{e} + y\,\vec{i}\,. \qquad (1)\end{aligned}$$

Toisaalta jos tunnetaan $\vec{p}$:n koordinaatit x, y osoitinkannassa, niin esitysmuoto $\vec{p} = r\angle\varphi$ saadaan lasketuksi kaavoilla

$$r = \sqrt{x^2 + y^2}\,, \qquad \begin{cases} \cos\varphi = x/r, \\ \sin\varphi = y/r. \end{cases} \tag{2}$$

Tämän mukaan φ määräytyy 2π:n monikertaa vaille yksikäsitteisesti, jos $r > 0$ (eli $(x, y) \neq (0, 0)$), joten osoitin määräytyy sovittujen samastussääntöjen nojalla yksikäsitteisesti. Jatkossa käytetään osoittimen eri esityismuodoista nimityksiä

$\vec{p} = r\angle\varphi$ *polaarimuoto* (polaariesitys),

$\vec{p} = x\vec{e} + y\vec{i}$ *komponenttimuoto*.

Siirtyminen esitysmuodosta toiseen tapahtuu siis säännöillä (1) ja (2). Koska komponenttiesitys luo kääntäen yksikäsitteisen vastaavuuden $\mathbb{P} \leftrightarrow \mathbb{R}^2$, niin osoittimia voidaan ajatella abstraktisti myös reaalilukujen pareina:

$$\vec{p} = x\vec{e} + y\vec{i} \;\leftrightarrow\; (x, y) \in \mathbb{R}^2.$$

Osoittimen komponenttimuoto on erityisen kätevä yhteenlaskussa:

$$\boxed{\begin{aligned} &\vec{p}_1 = r_1\angle\theta_1 \;\;\&\;\; \vec{p}_2 = r_2\angle\theta_2 \\ &\quad\Rightarrow\; \vec{p}_1 + \vec{p}_2 = (r_1\cos\theta_1 + r_2\cos\theta_2)\,\vec{e} + (r_1\sin\theta_1 + r_2\sin\theta_2)\,\vec{i}\,. \end{aligned}}$$

Polaarimuotoon päästään tästä takaisin säännöillä (2) — lopputulosta ei yleisessä muodossa kannata kirjoittaa.

Tähän asti osoitinavaruuden ja tason vektoriavaruuden välinen analogia on täydellinen. Tultaessa osoittimien <u>kertolaskuun</u> tiet kuitenkin eroavat. Osoittimille

ei määritellä skalaarituloa eikä ristituloakaan, vaan toisen tyyppinen tulo, jota jatkossa sanotaan *osoitintuloksi*. Määritelmä on:

$$\boxed{\vec{p}_1 = r_1\angle\varphi_1 \;\;\&\;\; \vec{p}_2 = r_2\angle\varphi \quad\Rightarrow\quad \vec{p}_1\cdot\vec{p}_2 = r_1r_2\angle(\varphi_1+\varphi_2).}$$

Osoitintulo on siis funktio tyyppiä $\mathbb{P}\times\mathbb{P}\to\mathbb{P}$ — kyse on hieman omalaatuisesta reaalilukujen kertolaskun (pituudet) ja yhteenlaskun (suunnat) yhdistelmästä (vrt. Harj.teht. I.2:14).

Osoitintulon määritelmästä seuraa suoraan normaali vaihdantalaki

$$\vec{p}_1\vec{p}_2 = \vec{p}_2\vec{p}_1$$

(tässä jätetty kertomerkki pois tavalliseen tapaan). Myös liitäntälaki on voimassa (Harj.teht. 1). Lisäksi nähdään, että *ykkösosoitin*, eli kertolaskun ykkösalkio, on

$$\vec{1} = \vec{e} = 1\angle 0\,.$$

Jokaisella $\vec{p}\in\mathbb{P},\ \vec{p}\neq\vec{0}$, on myös *käänteisosoitin* $(\vec{p})^{-1}$, joka toteuttaa

$$\vec{p}\cdot(\vec{p})^{-1} = \vec{1} = 1\angle 0\,.$$

Nimittäin jos $\vec{p} = r\angle\varphi,\ r\neq 0$, niin

$$(\vec{p})^{-1} = r^{-1}\angle -\varphi\,.$$

Osoitintulon määritelmästä seuraa välittömästi, että jokaisella $n\in\mathbb{N}$ pätee potenssiin korotuksen laskusääntö

$$\boxed{(\vec{p})^n = r^n\angle n\varphi.}$$

Tapauksessa $\vec{p}\neq\vec{0}$ tämä on pätevä jokaisella $n\in\mathbb{Z}$, kun sovitaan normaaliin tapaan, että

$$(\vec{p})^0 = \vec{1},\qquad (\vec{p})^{-n} = [(\vec{p})^{-1}]^n,\quad n\in\mathbb{N}\quad (\vec{p}\neq\vec{0}).$$

Esimerkki 1 Määritä osoitinyhtälön $(\overrightarrow{p})^3 = -\overrightarrow{p}$ kaikki ratkaisut.

Ratkaisu Osoitintulon määritelmän ja samastussääntöjen nojalla päätellään

$$\begin{aligned}
& (\overrightarrow{p})^3 = -\overrightarrow{p} \\
\Leftrightarrow\quad & r^3\angle 3\varphi \;=\; r\angle(\varphi+\pi) \\
\Leftrightarrow\quad & r^3 = r = 0 \quad \text{tai} \quad r^3 = r > 0 \quad \& \quad 3\varphi = (\varphi+\pi) + k\cdot 2\pi, \quad k\in\mathbb{Z} \\
\Leftrightarrow\quad & r = 0 \quad \text{tai} \quad r = 1 \quad \& \quad \varphi = \frac{\pi}{2} + k\cdot\pi, \quad k\in\mathbb{Z}.
\end{aligned}$$

Tapauksessa $r = 1$ saadaan erilaisia osoittimia vain k:n arvoilla $0, 1$, joten ratkaisut ovat

$$\overrightarrow{p} = \overrightarrow{0}, \;\; \overrightarrow{p} = \pm\overrightarrow{i}. \quad \square$$

Osoitintulon ja osoittimien yhteenlaskun määritelmistä seuraa vielä, että pätee myös osittelulaki

$$\overrightarrow{p}\cdot(\overrightarrow{q}_1 + \overrightarrow{q}_2) = \overrightarrow{p}\cdot\overrightarrow{q}_1 + \overrightarrow{p}\cdot\overrightarrow{q}_2.$$

Tämä on seuraus siitä, että jos $\overrightarrow{p} = r\angle\varphi$, niin kertolasku $\overrightarrow{p}\cdot\overrightarrow{q}$ voidaan kirjoittaa muotoon

$$\overrightarrow{p}\cdot\overrightarrow{q} = r(1\angle\varphi)\cdot\overrightarrow{q}.$$

Tässä osoittimella $(1\angle\varphi)$ kertominen on sama kuin kierto kulman φ verran. Sekä tälle operaatiolle että skalaarilla r kertomiselle pätee yhteenlaskun suhteen osittelulaki (vrt. ristitulon vastaavan osittelulain todistus Luvussa II.6), joten väitetty osittelulaki seuraa.

Kun nyt $\mathbb{P}$:ssä on tullut määritellyksi sekä yhteenlasku että kertolasku, jotka toteuttavat normaalit vaihdanta-, liitäntä- ja osittelulait, ja lisäksi on konstruoitu yhteenlaskun nolla-alkio ja vasta-alkio sekä kertolaskun ykkösalkio ja käänteisalkio, niin on tullut osoitetuksi, että $(\mathbb{P}, +, \cdot)$ on itse asiassa kunta. Nähdään siis, että osoitintulon mukaan ottaminen muuttaa osoitinalgebran varsin radikaalisti. Osoittimia voikin luonnehtia 'pyöriviksi luvuiksi', joita vain lasketaan yhteen kuten vektoreita. — Erillistä skalaarilla kertomisoperaatiota ei osoitintulon määrittelyn jälkeen enää tarvita, sillä skalaarin $\lambda\in\mathbb{R}$ voi tulkita osoittimeksi

$$\overrightarrow{\lambda} = \begin{cases} \lambda\angle 0, & \text{jos } \lambda \geq 0, \\ |\lambda|\angle\pi, & \text{jos } \lambda < 0, \end{cases}$$

jolloin skaalaus on osoitintulon erikoistapaus: $\lambda\overrightarrow{p} = \overrightarrow{\lambda}\overrightarrow{p}$.

Osoitintulolle saadaan varsin yksinkertainen esitystapa myös komponenttimuodossa. Nimittäin jos

$$\vec{p}_1 = r_1 \cos\varphi_1 \vec{e} + r_1 \sin\varphi_1 \vec{i} = x_1 \vec{e} + y_1 \vec{i},$$
$$\vec{p}_2 = r_2 \cos\varphi_2 \vec{e} + r_2 \sin\varphi_2 \vec{i} = x_2 \vec{e} + y_2 \vec{i},$$

niin osoitintulon määritelmästä ja trigonometristen funktioiden yhteenlaskukaavoista (ks. Luku II.5) seuraa

$$\begin{aligned}\vec{p}_1 \cdot \vec{p}_2 &= r_1 r_2 \cos(\varphi_1 + \varphi_2)\, \vec{e} + r_1 r_2 \sin(\varphi_1 + \varphi_2)\, \vec{i} \\ &= r_1 r_2 (\cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2)\, \vec{e} \\ &\qquad + r_1 r_2 (\sin\varphi_1 \cos\varphi_2 + \cos\varphi_1 \sin\varphi_2)\, \vec{i} \\ &= (x_1 x_2 - y_1 y_2) \vec{e} + (x_1 y_2 + y_1 x_2) \vec{i}.\end{aligned}$$

Näin ollen osoitintuloa vastaa $\mathbb{R}^2$:ssä määritelty tulo

$$\boxed{(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2,\ x_1 y_2 + y_1 x_2).} \tag{3}$$

Yhdessä jo aiemmin määritellyn (vektorien) yhteenlaskun kanssa, ts.

$$\boxed{(x_1, y_1) + (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2)} \tag{4}$$

on lukupareista saatu aikaan algebra $(\mathbb{R}^2, +, \cdot)$, joka siis on kunta (!). Kunnan $(\mathbb{R}^2, +, \cdot)$ nolla- ja ykkösalkiot ovat

$$(0,0) \mathrel{\widehat{=}} \vec{0}, \quad (1,0) \mathrel{\widehat{=}} \vec{1}.$$

HARJOITUSTEHTÄVIÄ

1. Todista osoitintulon liitäntälaki $(\vec{p}_1 \vec{p}_2)\vec{p}_3 = \vec{p}_1(\vec{p}_2 \vec{p}_3)$.
2. Näytä suoraan osoitintulon määritelmästä, että pätee:
   a) $-\vec{p} = (-\vec{1})\vec{p}$ b) $\vec{0}\vec{p} = \vec{0}$ c) $\vec{p}\vec{q} = \vec{0} \Rightarrow \vec{p} = \vec{0} \ \vee\ \vec{q} = \vec{0}$
3. Määritä seuraavien osoitinyhtälöiden kaikki ratkaisut:
   a) $(\vec{p})^2 = -\vec{1}$ b) $2(\vec{p})^3 = \vec{1}$ c) $(\vec{p})^4 + 3\vec{p} = \vec{0}$ d) $(\vec{p})^4 = 4\angle\pi$
4. Olkoon $(x, y) \in \mathbb{R}^2$, $(x, y) \neq (0, 0)$. Lähtien lukuparien tulon määritelmästä määritä $(a, b) \in \mathbb{R}^2$ siten, että $(x, y) \cdot (a, b) = (1, 0)$, ts. $(a, b) = (x, y)^{-1}$.

III.2 Kompleksiluvut ja niillä laskeminen

Kompleksilukujen määrittelyn lähtökohdaksi otetaan matematiikassa usein kunta $(\mathbb{R}^2, +, \cdot)$, jossa laskutoimitukset on määritelty edellisen luvun säännöillä (4) ja (3). Suoraviivaisin määritelmä on yksinkertaisesti sopia, että $(\mathbb{R}^2, +, \cdot)$ itse on kompleksilukujen kunta, ts. kompleksiluvut ovat $\mathbb{R}^2$:n lukupareja, joiden väliset laskutoimitukset on määritelty mainitulla tavalla. Käytännön laskurutiineissa käytetään kuitenkin yleensä havainnollisempia kompleksilukujen esitystapoja. Näitä ovat edellä kuvattu polaarimuoto (osoitinmuoto) tai sitäkin yleisempi, komponenttimuodosta johdettu kolmas esitystapa, joka seuraavassa otetaan kompleksilukujen käytännöllisen määrittelyn lähtökohdaksi.

Aloitetaan merkitsemällä $\mathbb{R}^2$:n 'x-akselia' symbolilla $\mathbb{R}^2_0$:

$$\mathbb{R}^2_0 = \{(x, 0) \mid x \in \mathbb{R}\}.$$

Havaitaan, että $\mathbb{R}^2_0$ sisältää $(\mathbb{R}^2, +, \cdot)$:n nolla-alkion $(0, 0)$, samoin ykkösalkion $(1, 0)$. Lisäksi jos $(x, 0)$, $(y, 0)$ ovat mitkä tahansa kaksi $\mathbb{R}^2_0$:n alkiota, niin edellisen luvun laskusääntöjen perusteella näiden summa, erotus, tulo ja osamäärä (jos $y \neq 0$) ovat myös $\mathbb{R}^2_0$:ssa. Nämä tosiasiat yhdessä merkitsevät, että $(\mathbb{R}^2_0, +, \cdot)$ on itsekin kunta ja siis $(\mathbb{R}^2, +, \cdot)$:n <u>alikunta</u> (vrt. Luku I.2). Toisaalta havaitaan, että $\mathbb{R}^2_0$:n ja $\mathbb{R}$:n välinen ilmeinen vastaavuus

$$(x, 0) \in \mathbb{R}^2_0 \;\leftrightarrow\; x \in \mathbb{R}$$

ulottuu samanlaisena laskutoimituksiin: Jos $\mathbb{R}^2_0$:ssa suoritetaan laskutoimitus kunnan $(\mathbb{R}^2, +, \cdot)$ säännöillä, niin tätä vastaa $(\mathbb{R}, +, \cdot)$:n laskutoimitus normaalissa reaalilukujen kunnassa. Esimerkiksi kertolasku

$$(x_1, 0) \cdot (x_2, 0) = (x_1 x_2, 0)$$

(sovellettu edellisen luvun sääntöä (3)) vastaa kertolaskua $\mathbb{R}$:ssä, kun käytetään em. vastaavuusperiaatetta:

$$\begin{array}{rccccc} \mathbb{R}^2_0: & (x_1, 0) & \cdot & (x_2, 0) & = & (x_1 x_2, 0) \\ & \updownarrow & & \updownarrow & & \updownarrow \\ \mathbb{R}: & x_1 & \cdot & x_2 & = & x_1 x_2 \end{array}$$

Sama pätee yhteelaskulle, joten kunnilla $(\mathbb{R}^2_0, +, \cdot)$ ja $(\mathbb{R}, +, \cdot)$ ei käytännön laskennan kannalta ole mitään eroa. Tehdäänkin tämän perusteella kunnassa $(\mathbb{R}^2, +, \cdot)$ samastussopimus

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

Tällä sopimuksella reaalilukujen kunnasta $(\mathbb{R}, +, \cdot)$ tulee kunnan $(\mathbb{R}^2, +, \cdot)$ alikunta. Merkitään vielä $(0, 1) = i$, jolloin pätee

$$\begin{aligned}(x, y) = (x, 0) + (0, y) &= (x, 0) + (y, 0)(0, 1) \\ &= x + yi = x + iy.\end{aligned}$$

Kunnan $(\mathbb{R}^2, +, \cdot)$ kertolaskun määritelmän ja tehdyn samastussopimuksen mukaan on

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

MÄÄRITELMÄ III.2.1 **(Kompleksiluvut)** Kompleksilukujen joukko $\mathbb{C}$ sisältää reaaliluvut ja lisäksi nk. *imaginaariluvun* i, joka toteuttaa

$$i^2 = i \cdot i = -1 \in \mathbb{R}.$$

Kompleksiluvut ovat kunta $(\mathbb{C}, +, \cdot)$, missä jokaisella $z \in \mathbb{C}$ on yksikäsitteinen esitysmuoto

$$z = x + iy, \quad x, y \in \mathbb{R}.$$

Erityisesti on $x + 0i = x \in \mathbb{R}$ ja $0 + 1i = i$.

Määritelmän mukaisesti $(\mathbb{C}, +, \cdot)$ syntyy reaalilukujen kunnan $(\mathbb{R}, +, \cdot)$ laajennuksena, kun lukujoukkoon $\mathbb{R}$ lisätään imaginaariluku i ja huomioidaan kunnan perusaksiooma (K0) — vrt. Luvun I.2 Esimerkki 3, jossa kuntaa $(\mathbb{Q}, +, \cdot)$ laajennettiin vastaavalla tavalla lisäämällä luku $a = \sqrt{2}$. Luvulle i asetettu ehto $i^2 = -1$ ei toteudu millekään $i \in \mathbb{R}$, joten i on 'aidosti imaginaarinen' luku. Osoitinkunnassa $(\mathbb{P}, +, \cdot)$ laskusääntöä $i^2 = -1$ vastaa tulos

$$\overrightarrow{i} \cdot \overrightarrow{i} = 1\angle\pi = -(1\angle 0) = -\overrightarrow{1}.$$

Määritelmän III.2.1 perusteella kompleksiluvuilla voi operoida jokseenkin normaalisti, eli reaalilukujen algebrasta tutulla tavalla. Normaalista poikkeaa vain laskusääntö $i^2 = -1$. Laskennassa kompleksiluvun esitysmuotoja voi vaihdella vapaasti edellisen luvun muunnossääntöjen puitteissa. Tavallisimmin käytetään joko Määritelmän III.2.1 mukaista kompleksiluvun *perusmuotoa*, tai sitten polaarimuotoa. Saman luvun eri esitysmuotojen välillä käytetään jatkossa joko vastaavuusmerkintää '$\widehat{=}$' tai, sikäli kuin luontevaa, yksinkertaisesti samastusta '='.

ESIMERKKI 1 Saata kompleksiluku $(2 + 3i) \cdot (5 - 2i)$ perusmuotoon $x + iy$.

<u>Ratkaisu</u> Määritelmän III.2.1 perusteella

$$(2 + 3i) \cdot (5 - 2i) = 2 \cdot 5 - 2 \cdot 2i + 5 \cdot 3i - 3 \cdot 2 \cdot i^2 = 16 + 11i. \quad \square$$

Esimerkki 2 Olkoon $z = a + ib \neq 0$. Laske z^{-1} perusmuodossa.

Ratkaisu Jos $z^{-1} = x + iy$, niin on oltava

$$\begin{aligned} & (a+ib)(x+iy) = 1 \\ \Leftrightarrow\ & (ax - by) + i(bx + ay) = 1 = 1 + 0i \\ \Leftrightarrow\ & \begin{cases} ax - by = 1 \\ bx + ay = 0 \end{cases} \\ \Leftrightarrow\ & x = \frac{a}{a^2+b^2}\,, \quad y = -\frac{b}{a^2+b^2}\,. \end{aligned}$$

Siis

$$z^{-1} = \frac{1}{a^2+b^2}(a - ib) = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\,i. \quad \square$$

Määritelmä III.2.2 Kompleksiluvun

$$z = x + iy \in \mathbb{C} \ \widehat{=}\ (x, y) \in \mathbb{R}^2 \ \widehat{=}\ r\angle\varphi \in \mathbb{P}$$

- *reaaliosa* on $\operatorname{Re} z = x = r\cos\varphi$.
- *imaginaariosa* on $\operatorname{Im} z = y = r\sin\varphi$.
- *itseisarvo* (moduuli) on $|z| = \sqrt{x^2+y^2} = r$.
- *vaihekulma* (argumentti, napakulma) on $\arg z = \varphi$.
- *liittoluku* eli *konjugaatti* on $\overline{z} = x - iy \ \widehat{=}\ (x, -y) \ \widehat{=}\ r\angle -\varphi\,.$

Kompleksiluvuilla laskettaessa seuraavia Määritelmään III.2.2 liittyviä kaavoja tarvitaan usein:

(1)	$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
(2)	$\overline{z_1 z_2} = \overline{z_1}\ \overline{z_2}$
(3)	$z\overline{z} = \lvert z\rvert^2$
(4)	$\lvert z_1 z_2\rvert = \lvert z_1\rvert\lvert z_2\rvert$
(5)	$\lvert z^{-1}\rvert = \lvert z\rvert^{-1}$
(6)	$\lvert\overline{z}\rvert = \lvert z\rvert$
(7)	$z + \overline{z} = 2\operatorname{Re} z, \quad z - \overline{z} = 2i\operatorname{Im} z$
(8)	$\overline{\overline{z}} = z$

Nämä voi helposti perustella määritelmistä (kaavat (2)–(5) suorimmin polaarimuodosta).

Esimerkki 2 (jatko) Esimerkissä laskettiin kompleksiluvun $z \neq 0$ käänteisluku z^{-1}. Koska $z \neq 0 \Leftrightarrow \overline{z} \neq 0$, niin kaavaa (3) käyttäen voidaan päätellä myös seuraavasti:

$$\begin{aligned} & zz^{-1} = 1 \\ \Leftrightarrow\ & \overline{z}(zz^{-1}) = \overline{z} \\ \Leftrightarrow\ & (\overline{z}z)z^{-1} = \overline{z} \\ \Leftrightarrow\ & |z|^2 z^{-1} = \overline{z} \\ \Leftrightarrow\ & z^{-1} = |z|^{-2}\overline{z}. \end{aligned}$$

Tämä vastaa normaalia lavennusmenettelyä

$$z^{-1} = \frac{1}{a+ib} = \frac{(a-ib)}{(a+ib)(a-ib)} = \frac{1}{a^2+b^2}(a-ib)\,.$$

Vielä nopeampi on kuitenkin osoitinlasku:

$$z = r\angle\varphi \ \Rightarrow\ z^{-1} = r^{-1}\angle -\varphi = r^{-2}(r\angle -\varphi) = |z|^{-2}\overline{z}. \quad \square$$

Esimerkki 3 Saata $z = (1+i)^7$ perusmuotoon $x + iy$.

<u>Ratkaisu</u> Tässäkin polaariesitys on tehokkain: Koska

$$1 + i = r\angle\varphi\,, \quad \text{missä} \quad r = \sqrt{1+1} = \sqrt{2}, \quad \varphi = \frac{\pi}{4}\,,$$

niin

$$\begin{aligned} (1+i)^7 &= (\sqrt{2})^7\angle 7\pi/4 \\ &= 8\sqrt{2}\,\angle -\pi/4 \\ &= 8\sqrt{2}\,(\cos\tfrac{\pi}{4} - i\,\sin\tfrac{\pi}{4}) \\ &= 8(1-i). \quad \square \end{aligned}$$

Kompleksialgebran kaavoista maininnan arvoinen on vielä *de Moivren kaava*

$$\boxed{(9) \qquad (\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi.}$$

Tämä seuraa välittömästi, kun osoitinlaskennan tulos

$$(1\angle\varphi)^n = 1\angle n\varphi$$

esitetään kompleksilukujen perusmuodossa.

Esimerkki 4 de Moivren kaavan ja binomikaavan mukaan

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

$$\Rightarrow \begin{cases}\cos 3\varphi &= \cos^3\varphi - 3\cos\varphi\sin^2\varphi,\\ \sin 3\varphi &= 3\cos^2\varphi\sin\varphi - \sin^3\varphi\end{cases}$$

(vrt. Luvun II.5 Esimerkki 9). □

Kolmioepäyhtälö $\mathbb{C}$:ssä

Kompleksilukujen yhteenlasku on samanlainen operaatio kuin tason vektorien yhteenlasku, ja myös kompleksiluvun itseisarvo vastaa vektorin itseisarvoa. Tästä syystä kompleksiluvuille pätee myös samaa muotoa oleva kolmioepäyhtälö kuin vektoreille:

$$\boxed{||z_1| - |z_2|| \le |z_1 + z_2| \le |z_1| + |z_2|, \quad z_1, z_2 \in \mathbb{C}.}$$

Kolmioepäyhtälö on hyvin keskeinen työkalu kompleksilukuihin perustuvassa matemaattisessa analyysissä eli *kompleksianalyysissä*. Tyypillisenä esimerkkinä kolmioepäyhtälön käytöstä tarkasteltakoon väittämää, joka koskee yleistä *kompleksimuuttujan polynomia* muotoa

$$p(z) = c_0 + c_1 z + \cdots + c_n z^n = \sum_{k=0}^{n} c_k z^k, \quad c_n \neq 0.$$

Tässä siis $z \in \mathbb{C}$ on *kompleksimuuttuja* ja myös luvut c_k, eli polynomin *kertoimet*, ovat kompleksilukuja. Luku $n \in \mathbb{N} \cup \{0\}$ on polynomin *aste*.

Kun kompleksimuuttujan polynomiin sovelletaan kolmioepäyhtälöä sekä em. kaavoja (4), (5), tullaan väittämään, jonka mukaan polynomin itseisarvo $|p(z)|$ kasvaa riittävän suurilla $|z|$:n arvoilla kvalitatiivisesti samalla tavoin kuin polynomin korkeimman asteisen termin itseisarvo, eli riittävän suurilla $|z|$:n arvoilla

$$|p(z)| \sim |c_n z^n| = |c_n||z|^n$$

(tässä on käytetty kaavaa (4)). Väittämä voidaan muotoilla täsmällisemmin esim. seuraavasti (ks. myös Harj.teht. 10):

PROPOSITIO III.2.3 **(Polynomin kasvu)** Jos $p(z) = \sum_{k=0}^{n} c_k z^k$, missä $c_k \in \mathbb{C}$, $n \in \mathbb{N}$ ja $c_n \neq 0$, niin on olemassa $R \in \mathbb{R}_+$ siten, että pätee

$$\frac{1}{2}|c_n||z|^n \leq |p(z)| \leq \frac{3}{2}|c_n||z|^n, \quad \text{kun } z \in \mathbb{C} \text{ ja } |z| \geq R.$$

<u>Todistus</u> Koska $c_n \neq 0$, voidaan kirjoittaa

$$p(z) = c_n z^n (1 + b_{n-1} z^{-1} + \cdots + b_0 z^{-n}),$$

missä $b_k = c_k / c_n$. Käyttämällä kolmioepäyhtälön ensimmäistä ja toista osaa muodoissa

$$|z_1 + z_2| \geq |z_1| - |z_2|, \quad -|z_1 + z_2| \geq -|z_1| - |z_2|$$

sekä kaavoja (4), (5), päätellään

$$\begin{aligned} |p(z)| &= |c_n||z|^n |1 + b_{n-1} z^{-1} + \cdots + b_0 z^{-n}| \\ &\geq |c_n||z|^n (1 - |b_{n-1} z^{-1} + \cdots + b_0 z^{-n}|) \\ &\geq |c_n||z|^n (1 - |b_{n-1} z^{-1}| - \cdots - |b_0 z^{-n}|) \\ &= |c_n||z|^n \left(1 - \frac{|b_{n-1}|}{|z|} - \cdots - \frac{|b_0|}{|z|^n}\right). \end{aligned}$$

Kun

$$|z| \geq \max\{1,\ 2(|b_{n-1}| + \cdots + |b_0|)\} = R,$$

niin pätee

$$\frac{|b_{n-1}|}{|z|} + \cdots + \frac{|b_0|}{|z|^n} \leq \frac{|b_{n-1}| + \cdots + |b_0|}{|z|} \leq \frac{1}{2},$$

joten seuraa väittämän ensimmäinen osa:

$$|p(z)| \geq \frac{1}{2}|c_n||z|^n, \quad \text{kun } |z| \geq R.$$

Jälkimmäinen osa seuraa vastaavasti soveltamalla kolmioepäyhtälön jälkimmäistä osaa. □

ESIMERKKI 5 Jos $p(z) = z^5 - 4i\, z^3 + 10z + (30 + 40i)$, niin em. todistuksen logiikkaa seuraamalla nähdään, että

$$\begin{aligned} |p(z)| &\geq |z|^5 \left(1 - \frac{|5i|}{|z|^2} - \frac{10}{|z|^4} - \frac{|30 + 40i|}{|z|^5}\right) \\ &= |z|^5 \left(1 - \frac{5}{|z|^2} - \frac{10}{|z|^4} - \frac{50}{|z|^5}\right) \\ &\geq |z|^5 \left(1 - \frac{5}{|z|} - \frac{10}{|z|} - \frac{50}{|z|}\right) \\ &= |z|^5 \left(1 - \frac{65}{|z|}\right), \quad \text{kun } |z| \geq 1. \end{aligned}$$

Vastaavasti päätellään, että $|p(z)| \leq |z|^5(1+65|z|^{-1})$, kun $|z| \geq 1$, joten seuraa

$$\frac{1}{2}|z|^5 \;\leq\; |p(z)| \;\leq\; \frac{3}{2}|z|^5, \quad \text{kun } |z| \geq 130 = R. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Muunna seuraavat kompleksiluvut perusmuotoon $x+iy$ annetusta osoitinmuodosta tai osoitinmuotoon annetusta perusmuodosta. Käytä tarvittaessa likiarvoja.

   a) $\sqrt{2}\angle -\frac{5\pi}{4}$ b) $\sqrt{3}-2i$ c) $2\angle 700^\circ$ d) $-3-4i$

2. Saata seuraavat kompleksiluvut perusmuotoon $x+iy$ ja polaarimuotoon:
   a) $(1+i)(1-i)^5$ b) $(1-i)(1+i\sqrt{3})^{-1}$ c) $(\sqrt{2}+1+i)^8$

3. Laske itseisarvo ja vaihekulma käyttäen polaarimuotoa:
   a) $(1+i)^6$ b) $(1-i\sqrt{3})(1-i)^2$ c) $(2-2i)(\sqrt{3}+i)^{-2}$

4. Olkoon $z = \frac{1}{2}(1+i\sqrt{3})$. Millä $n \in \mathbb{Z}$ pätee a) $z^n = z$, b) $z^n = -z$?

5. Kompleksilukujen itseisarvoille pätee $|z_1 z_2|^2 = |z_1|^2|z_2|^2$. Tarkista kaavan pätevyys suoraan kompleksilukujen perusmuodosta, eli kirjoittamalla $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$.

6. Johda de Moivren kaavan avulla:
   a) $\sin 5x$:lle lauseke $\sin x$:n polynomina,
   b) $\sin^5 x$:lle lauseke muodossa $a\sin x + b\sin 3x + c\sin 5x$, $a, b, c \in \mathbb{R}$.

7. Jos kompleksilukujen kunnassa on määritelty järjestysrelaatio, niin ensimmäisen järjestysaksiooman (J1) mukaan on oltava joko $i > 0$, $i = 0$ tai $i < 0$. Päättele, että järjestysrelaatiota ei voi määritellä.

8. Määritä jokin R siten, että pätee

$$|z|^5 \;\leq\; |2z^5 + 1000z^4 - 10^4(3+4i)| \;\leq\; 3|z|^5, \quad \text{kun} |z| \geq R.$$

9. (*) Todista, että kaikilla $z_1, z_2 \in \mathbb{C}$ pätee $(1+|z_1|^2)(1+|z_2|^2) \geq |1+z_1 z_2|^2$.

10. (*) Näytä, että Propositon III.2.3 oletuksin on olemassa $R \in \mathbb{R}_+$ siten, että

$$0.999|c_n||z|^n \;\leq\; |p(z)| \;\leq\; 1.001|c_n||z|^n, \quad \text{kun } z \in \mathbb{C} \text{ ja } |z| \geq R.$$

    Millaisesta vielä yleisemmästä väittämästä tämä on erikoistapaus?

III.3 Algebran peruslause

Kompleksianalyysin alalaji *funktioteoria* tutkii kompleksimuuttujan kompleksiarvoisia funktioita tyyppiä

$$f : \mathbb{C} \to \mathbb{C} \quad \text{tai} \quad f : A \to \mathbb{C}, \quad A \subset \mathbb{C}.$$

Yksinkertaisin (vaan ei vähäisin) esimerkki tällaisesta funktiosta on edellisessä luvussa (Propositio III.2.3) tarkasteltu, koko $\mathbb{C}$:ssä määritelty polynomifunktio

$$p(z) = \sum_{k=0}^{n} c_k z^k, \quad c_k \in \mathbb{C},\ c_n \neq 0. \tag{1}$$

Jokainen kompleksimuuttujan kompleksiarvoinen funktio on esitettävissä muodossa

$$f(z) = \operatorname{Re} f(z) + i \operatorname{Im} f(z) = u(x,y) + iv(x,y), \quad z = x + iy,$$

missä u, v ovat funktioita tyyppiä $u, v : \mathbb{R}^2 \to \mathbb{R}$ (vrt. Luku II.5). Itse f on siis myös tulkittavissa funktioksi tyyppiä $f : \mathbb{R}^2 \to \mathbb{C}$. Kompleksimuuttujan funktioita onkin luontevaa ajatella geometrisesti tasossa määritellyiksi. Tasoa sanotaan tällaisissa tarkasteluissa *kompleksitasoksi*. Kompleksitasossa siis piste (x, y) tarkoittaa kompleksilukua $z = x + iy$.

Kompleksifunktion f *juureksi* (engl. root) tai yksinkertaisesti *nollakohdaksi* sanotaan jokaista $z \in \mathbb{C}$, jolle

$$f(z) = 0.$$

Reaali- ja imaginaariosiin hajotetusta muodosta juuret $z = x + iy$ saadaan yhtälöryhmän

$$\begin{cases} u(x,y) = 0 \\ v(x,y) = 0 \end{cases}$$

ratkaisuista.

Esimerkki 1 $f(z) = z^3 + 1$. Juuret?

<u>Ratkaisu</u> Tässä

$$\begin{aligned} f(x+iy) &= (x+iy)^3 + 1 \\ &= (x^3 - 3xy^2 + 1) + (3x^2y - y^3)i, \end{aligned}$$

joten f:n nollakohdat ratkeavat yhtälöryhmästä

$$\begin{cases} x^3 - 3xy^2 + 1 = 0, \\ 3x^2y - y^3 = 0. \end{cases}$$

Jokaista yhtälöryhmän reaalista ratkaisua $(x,y) \in \mathbb{R}^2$ vastaa f:n juuri $x+iy \in \mathbb{C}$. Juuret löydetään tässä helposti, koska yhtälöryhmän jälkimmäinen yhtälö on kirjoitettavissa muotoon $(3x^2 - y^2)y = 0$. Näin ollen on oltava joko $y = 0$ tai $y^2 = 3x^2 \Leftrightarrow y = \pm\sqrt{3}\,x$. Sijoittamalla nämä edelliseen yhtälöön saadaan juuriksi

$$\begin{aligned} (x_1, y_1) &= (-1, 0) &&\mathrel{\widehat{=}}\ z_1 = -1, \\ (x_2, y_2) &= \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) &&\mathrel{\widehat{=}}\ z_2 = \frac{1}{2}\left(1 + i\sqrt{3}\right), \\ (x_3, y_3) &= \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) &&\mathrel{\widehat{=}}\ z_3 = \frac{1}{2}\left(1 - i\sqrt{3}\right). \end{aligned}$$

Esimerkissä f oli polynomi astetta 3 ja sille löydettiin kolme juurta. Tulos on erikoistapaus lauseesta, joka on sekä laajakantoinen että syvällinen:

*Lause III.3.1 **(Algebran peruslause)** Jos p on polynomi muotoa (1) ja astetta $n \geq 1$, niin on olemassa n kompleksilukua $z_1, \ldots, z_n$ siten, että pätee

$$\begin{aligned} p(z) &= c_n(z - z_1) \cdot \cdots \cdot (z - z_n) \\ &= c_n \prod_{i=1}^{n} (z - z_i). \end{aligned}$$

Lauseen III.3.1 mukaisessa tulohajotelmassa luvut z_i eivät välttämättä ole eri suuret, joten polynomissa voi olla tekijä muotoa

$$(z - z_k)^m, \quad 1 \leq m \leq n,$$

liittyen juureen z_k. Jos p:llä on tämä tekijä mutta ei tekijää $(z - z_k)^{m+1}$, niin sanotaan, että juuri z_k on *m-kertainen* (engl. m-fold) tai että juuren *kertaluku* (engl. order) on m. Jos kertaluku on $m = 1$, sanotaan, että juuri on *yksinkertainen* (engl. simple). Jos samaa juurta edustavat tekijät kootaan yhteen, saadaan Lauseen III.3.1 tulohajotelmalle muoto

$$p(z) = c_n \prod_{i=1}^{\nu} (z - z_i)^{m_i}, \tag{2}$$

missä nyt $z_i \neq z_j$ kun $i \neq j$, $m_i \in \mathbb{N}$ ja

$$\sum_{i=1}^{\nu} m_i = n.$$

Erityisesti jos hajotelmassa (2) on $\nu = n$, eli polynomilla on n eri juurta, on kaikkien juurien oltava yksinkertaisia. Tämä tilanne oli Esimerkissä 1.

Esimerkissä 1 aidosti kompleksiset juuret ($\operatorname{Im} z \neq 0$) esiintyivät *konjugaattiparina*. Näin on yleisemmin silloin, kun (kuten esimerkissä) polynomi on *reaalikertoiminen*, ts. $c_k \in \mathbb{R}$ lausekkeessa (1). Tämä perustuu siihen, että reaalikertoimisen polynomin tapauksessa pätee (vrt. edellisen luvun kaavat (1)–(2))

$$\overline{p(z)} = p(\overline{z}) \quad \text{(reaalikertoiminen polynomi).}$$

Tällöin

$$p(z) = 0 \;\Rightarrow\; p(\overline{z}) = 0,$$

ja juuret siis 'pariutuvat'. Jos konjugaattiparin muodostavat juuret

$$z = a \pm ib,$$

niin polynomissa on tekijänä

$$(z-a-ib)(z-a+ib) = (z-a)^2 + b^2.$$

Reaalikertoiminen polynomi voidaan siis aina hajottaa reaalikertoimisiin tekijöihin, jotka ovat joko muotoa

$$\text{(a)} \qquad (z-a_k)^m, \quad a_k \in \mathbb{R},$$

vastaten m-kertaista reaalijuurta $z_k = a_k$, tai muotoa

$$\text{(b)} \qquad [(z-a_k)^2 + b_k^2\,]^m \quad a_k, b_k \in \mathbb{R},\ b_k \neq 0,$$

vastaten m-kertaista konjugaattijuuriparia $z_k = a_k \pm ib_k$.

Esimerkki 1 (jatko) Esimerkin polynomi hajoaa muotoon

$$\begin{aligned} z^3 + 1 &= (z+1)(z^2 - z + 1) \\ &= (z+1)\left[\left(z - \frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2\right]. \quad \square \end{aligned}$$

'Algebran peruslause' edellä esitetyssä muodossa on itse asiassa oikean *Algebran peruslauseen* tekninen seuraamus. Oikea peruslause muotoillaan:

*Lause III.3.2 **(Algebran peruslause – lyhyt muoto)** Jos $n \geq 1$, niin polynomilla (1) on juuri kompleksitasossa.

Lauseen III.3.2 perusteella polynomi purkautuu heti tekijöihin Lauseen III.3.1 mukaisesti. Nimittäin jos yksi juuri z_1 on löydettävissä, niin $p(z)$ voidaan kirjoittaa muotoon

$$p(z) = p(z) - p(z_1) = \sum_{k=1}^{n} c_k(z^k - z_1^k).$$

Tässä on kunta-algebran perusteella (Propositio I.2.2)

$$z^k - z_1^k = (z - z_1)(z^{k-1} + z_1 z^{k-2} + \ldots + z_1^{k-1}), \quad k = 1 \ldots n,$$

joten nähdään, että $p(z)$ on kirjoitettavissa muotoon

$$p(z) = (z - z_1)\, p_1(z),$$

missä p_1 on polynomi astetta $n-1$ ja muotoa

$$p_1(z) = c_n z^{n-1} + [\,\text{alempiasteisia termejä}\,].$$

Mikäli $n \geq 2$, on p_1:llä puolestaan edelleen juuri Lauseen III.3.2 perusteella, jolloin p_1:stä on erotettavissa tekijä $(z - z_2)$ jne, ja Lauseen III.3.1 väittämä siis seuraa.

Algebran peruslause on seuraamuksiltaan sen verran mittava, että matematiikkaa — edes sovellettua matematiikkaa — ilman sitä on nykyisin vaikea kuvitella. Lause on myös hyvin elegantti em. 'lyhyen kaavan' mukaan esitettynä. Edes todistusperiaatteen taakse kurkistaminen ei eleganssia vähennä, pikemminkin päinvastoin: Algebran peruslauseen voi nykyisin nähdä kompleksifunktioiden yleisemmän teorian seurauksena. Tarkemmin sanoen kyse on nk. *analyyttisten* funktioiden teoriasta. Käsite 'analyyttinen' määritellään myöhemmin Luvussa V.8; tässä yhteydessä todettakoon vain, että kyse on äärimmäisen säännöllisistä kompleksifunktioista, esimerkiksi sellaisista kuin juuri polynomit.

Algebran peruslauseen todistuksen perusidea analyyttisten funktioiden teorian avulla esitettynä on seuraava[†]: Tehdään vastaoletus, että polynomilla p ei ole nollakohtia. Siinä tapauksessa funktio

$$f(z) = \frac{1}{p(z)}$$

on koko kompleksitasossa säännöllinen. Itse asiassa f on 'äärimmäisen säännöllinen' eli analyyttinen. Kun lisäksi käytetään Proposition III.2.3 tulosta, niin voidaan todeta: Jos p:n aste on $n \geq 1$ ja p:llä ei ole nollakohtia, niin funktio $f = 1/p$ toteuttaa

1. f on analyyttinen koko kompleksitasossa.
2. $f(z) \to 0$, kun $|z| \to \infty$.

On olemassa ainakin yksi funktio, joka toteuttaa nämä molemmat ehdot, nimittäin $f(z) = 0$. Kysymys kuulukin: Onko muita? Analyyttisten funktioiden teorian antama — hieman yllättävä — vastaus tähän kysymykseen on:

Ei!

Vastaoletus on näin ollen johtanut loogiseen ristiriitaan ja lause siis on tosi.

[†]Algebran peruslauseen todistamiseen paneutui mm. 'matemaatikkojen ruhtinaaksi' sanottu saksalainen **Carl Friedrich Gauss** (1777-1855). Erään hyvin uskottavan todistuksen Gauss esitti väitöskirjassaan v. 1799. Loogisesti täysin auktoton todistus pystyttiin esittämään vasta myöhemmin 1800-luvulla reaali- ja kompleksianalyysin kehityttyä riittävästi. — Myöhemmin Luvussa V.10 esitetään Algebran peruslauseelle melko suoraviivainen todistus, joka perustuu vain reaalianalyysiin, Propositioon III.2.3 ja toiseen kompleksialgebran väittämään, joka on esitetty harjoitustehtävässä 13.

Kompleksiluvun juuret

Algebran peruslauseen mukaan polynomiyhtälöllä

$$z^n = a, \quad a \in \mathbb{C},\ n \in \mathbb{N},\ n \geq 2 \tag{3}$$

on ainakin yksi juuri. Itse asiassa osoittautuu, että jos $a \neq 0$ (mikä oletetaan), niin juuria on tasan n kappaletta, eli kaikki juuret ovat yksinkertaisia. Merkitään niitä kaikkia symbolilla

$$z = a^{1/n} = \sqrt[n]{a},$$

ja sanotaan, että kyseessä ovat *kompleksiluvun a juuret*. Näitä siis tulee olemaan n kpl, ja halutaan laskea ne perusmuotoon

$$z_k = x_k + iy_k, \quad k = 0 \ldots n-1.$$

Tähän päästään helpoiten polaariesityksen kautta: Kirjoitetaan

$$a = r\angle\varphi = r(\cos\varphi + i\sin\varphi),$$

missä siis $r = |a| \neq 0$. Kun merkitään $z = |z|\angle\psi$, niin

$$\begin{aligned} z^n = a \quad &\Leftrightarrow \quad |z|^n = r \quad \& \quad n\psi = \varphi + k\cdot 2\pi,\ k \in \mathbb{Z} \\ &\Leftrightarrow \quad |z| = \sqrt[n]{r} \quad \& \quad \psi = \frac{\varphi}{n} + k\cdot\frac{2\pi}{n},\ k \in \mathbb{Z}. \end{aligned}$$

Erilaisia ratkaisuja saadaan tästä indeksin arvoilla $k = 0 \ldots n-1$. Nämä voidaan esittää muodossa

$$\begin{aligned} z_k &= \sqrt[n]{r}\angle\tfrac{\varphi}{n} + k\cdot\tfrac{2\pi}{n} \\ &= \left(\sqrt[n]{r}\angle\tfrac{\varphi}{n}\right)\left(1\angle k\cdot\tfrac{2\pi}{n}\right) \\ &= \left(\sqrt[n]{r}\angle\tfrac{\varphi}{n}\right)\left(1\angle\tfrac{2\pi}{n}\right)^k \\ &= z_0\rho^k, \quad k = 0\ldots n-1, \end{aligned}$$

missä siis

$$\begin{aligned} z_0 &= \sqrt[n]{r}\angle\frac{\varphi}{n} = \sqrt[n]{r}\left(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}\right), \quad r = |a},\ \varphi = \arg a, \\ \rho &= 1\angle\frac{2\pi}{n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}. \end{aligned}$$

Perusmuodossa yhtälön (3) ratkaisut ovat

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

Lähtien perusjuuresta z_0 juuret sijaitsevat tasavälein kompleksitason ympyrällä, jonka säde $= \sqrt[n]{|a|}$. Kuvassa on tapaus $n = 8$.

Esimerkki 1 (jatko) Tässä $a = -1 = 1\angle\pi$, siis $r = 1$, $\varphi = \pi$. Juuret ovat

$$z_k = \cos\left(\frac{\pi}{3} + k \cdot \frac{2\pi}{3}\right) + i \sin\left(\frac{\pi}{3} + k \cdot \frac{2\pi}{3}\right), \quad k = 0, 1, 2,$$

eli
$$\begin{cases} z_0 = \dfrac{1}{2} + \dfrac{\sqrt{3}}{2} i, \\ z_1 = -1, \\ z_2 = \dfrac{1}{2} - \dfrac{\sqrt{3}}{2} i. \end{cases} \qquad \square$$

Esimerkki 2 $\sqrt{i} = ?$

Ratkaisu Kyseessä on yhtälön $z^2 = i$ ratkaiseminen. Koska $i = 1\angle\pi/2$, niin ratkaisut ovat

$$z_k = \cos\left(\frac{\pi}{4} + k\pi\right) + i\sin\left(\frac{\pi}{4} + k\pi\right), \quad k = 0, 1$$
$$\Rightarrow \sqrt{i} = \pm\frac{1}{\sqrt{2}}(1+i).$$

Tarkistus:

$$(\sqrt{i})^2 = \frac{1}{2}(1+i)^2 = \frac{1}{2}(1-1+2i) = i. \quad \text{OK!} \quad \square$$

Esimerkki 3 Ratkaise yhtälö $z^2 + (2+2i)z - i = 0$.

Ratkaisu Toisen asteen yhtälön ratkaisukaava perustuu vain kunta-algebraan, joten se on pätevä myös kompleksikertoimiselle polynomille. Ratkaisut siis ovat

$$z_{1,2} = -(1+i) \pm \sqrt{(1+i)^2 + i} = -(1+i) \pm \sqrt{3i},$$

eli edellisen esimerkin perusteella

$$z_1 = -(1+i) + \sqrt{\frac{3}{2}}\,(1+i) = \left(\sqrt{\frac{3}{2}} - 1\right)(1+i),$$
$$z_2 = -(1+i) - \sqrt{\frac{3}{2}}\,(1+i) = -\left(\sqrt{\frac{3}{2}} + 1\right)(1+i). \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien funktioiden kaikki juuret kompleksitasossa jakamalla funktiot reaali- ja imaginaariosiin.
   a) $f(z) = z^2 + i$ b) $f(z) = z^3 + 8i$ c) $f(z) = z^2 + 4\overline{z} - 1$

2. Jos $\operatorname{Re} z = \operatorname{Im} z = a$, niin millä a:n arvoilla pätee $|z - i| < |z - 3|$? Kuvio!

3. Tutki, millaiset kompleksitason pistejoukot tulevat määritellyiksi seuraavilla ehdoilla. Piirrä kuviot.
   a) $|z + 1 + i|^2 = 2$ b) $|z - i| \le 2$ c) $|z + i| = |z - 1 - i|$
   d) $\operatorname{Re}[(z - i)/(z + i)] = 0$

4. Jaa polynomi enintään toisen asteen reaalikertoimisiin tekijöihin:
   a) $z^3 + 2z^2 + 3z + 2$ b) $z^4 + 2$ c) $z^4 + 2z^3 + z^2 - 2z - 2$
   d) $z^8 - 256$ e) $(z^4 + 1)^2 - z^4$

5. Määritä a) kaikki reaalikertoimiset, b) kaikki polynomit p astetta ≤ 5, joilla on ominaisuudet: $p(0) = 1$ ja p:n nollakohdat ovat 1, i ja $-i$.

6. Määritä seuraavien juurien kaikki arvot:
a) $\sqrt[4]{-4}$ b) $\sqrt[6]{-64}$ c) $\sqrt[3]{i-1}$ d) $\sqrt{3+4i}$ e) $\sqrt{-7+24i}$

7. Juuren $\sqrt[6]{2+3i}$ eräs likiarvo on $1.2217 + 0.2019i$. Piirrä kuva, jossa kaikki juuren arvot on sijoitettu kompleksitasoon (laskematta erikseen muiden juurien likiarvoja).

8. Määritä seuraavien yhtälöiden kaikki ratkaisut perusmuodossa $x + iy$.
a) $z^2 + 2iz - i - 1 = 0$ b) $z^2 - 4iz - 4 + i = 0$ c) $z^2 - (3+5i)z = 4 - 7i$
d) $z^4 - 2z^2 + 4 = 0$ e) $z^4 + (1 - 2i\sqrt{3})z^2 - 3 - i\sqrt{3} = 0$

9. Määritä $a \in \mathbb{C}$ siten, että $z = 1 + i$ on polynomin $p(z) = z^3 + az + 1$ juuri, ja laske ja sen jälkeen muut juuret.

10. (*) a) Juuret $\sqrt[4]{3+4i}$ on mahdollista laskea tarkasti perusmuodossa $z_k = x_k + iy_k$ siten, että x_k ja y_k ovat geometrisia lukuja. Laske!

    b) Määritä tarkasti ne kompleksiluvut, joiden viides potenssi = 1. Piirrä kuvio lukujen sijainnista kompleksitasossa.

11. (*) Olkoon $p(z) = c_n z^n + \ldots + c_0$ reaalikertoiminen polynomi, $c_n \neq 0$. Näytä, että polynomilla on ainakin yksi reaalinen nollakohta, jos joko n on pariton tai $c_n c_0 \leq 0$.

12. (*) Todista, että jos kahdella polynomilla astetta $\leq n$ on samat arvot $n+1$ eri kompleksitason pisteessä, niin polynomit ovat samat.

13. (*) Olkoon $p(z)$ kompleksimuuttujan polynomi, joka ei ole vakio, ja olkoon $c \in \mathbb{C}$ ja $p(c) = a \in \mathbb{C}$. Näytä:

    a) On olemassa $m \in \mathbb{N}$ ja $b \in \mathbb{C},\ b \neq 0$ sekä polynomi $q(z)$ siten, että

    $$p(z) = a + b(z-c)^m + (z-c)^{m+1} q(z) \quad \forall z \in \mathbb{C}.$$

    b) Jos a)-kohdan hajotelmassa on $a \neq 0$ ja $q(z) = 0$, niin $|p(z)|$ pienenee johonkin suuntaan c:stä lähdettäessä, t.s. on olemassa $\rho \in \mathbb{C},\ |\rho| = 1$ ja $\delta > 0$ siten, että

    $$|p(c + t\rho)| < |p(c)|, \quad \text{kun } 0 < t < \delta.$$

    c) Jos a)-kohdan hajotelmassa on $a \neq 0$ ja $q(z) \neq 0$, niin b)-kohdan väittämä on edelleen tosi.
    d) Jos $|p(z)|$ saavuttaa paikallisen minimiarvon c:ssä, niin $p(c) = a = 0$.

III.4 *Kompleksikertoiminen vektoriavaruus

Luvussa II.2 esiteltiin algebra nimeltä (lineaarinen) vektoriavaruus, symbolisesti $(U, \mathbb{K})$, missä U on vektorien muodostama joukko ja $\mathbb{K}$ on nk. skalaarien muodostama vektoriavaruuden kertojakunta. Vektoriavaruudessa on määritelty kaksi laskuoperaatiota, vektorien yhteenlasku ja skalaarilla kertominen, jotka ovat funktioita tyyppiä

$$\begin{array}{ll} \text{yhteenlasku:} & U \times U \to U, \\ \text{skalaarilla kertominen:} & \mathbb{K} \times U \to U. \end{array}$$

Sisätuloavaruudeksi sanotaan vektoriavaruutta, jossa on määritelty myös skalaaritulo (eli sisätulo, ks. Luku II.4) funktiona tyyppiä

$$\text{skalaaritulo:} \qquad U \times U \to \mathbb{K}.$$

Tähän asti on käsitelty lähinnä tapausta $\mathbb{K} = \mathbb{R}$, jolloin puhutaan *reaalikertoimisesta* avaruudesta. Laajennus reaaliluvuista kompleksilukuihin mahdollistaa nyt myös *kompleksikertoimisen* vektoriavaruuden, jossa $\mathbb{K} = \mathbb{C}$. Kompleksikertoiminen vektoriavaruus ei algebrana oleellisesti poikkea reaalikertoimisesta niin kauan kuin vain vektorien yhteenlasku ja skalaarilla kertominen on määritelty. Sen sijaan skalaaritulon määrittely funktiona tyyppiä $U \times U \to \mathbb{C}$ aiheuttaa teoriaan lisäpiirteitä, jotka on otettava huomioon. Kompleksikertoimisen sisätuloavaruuden $(U, \mathbb{C})$ sisätulon $\langle \cdot, \cdot \rangle$ on toteutettava seuraavat ehdot (vrt. Luvun II.4 ehdot reaalikertoimiselle tapaukselle) :

1. *Puolisymmetrisyys*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} \quad \forall\ \mathbf{u}, \mathbf{v} \in U.$$

2. *Sekvilineaarisuus*

   (a) $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$,

   (b) $\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \overline{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle + \overline{\beta} \langle \mathbf{u}, \mathbf{w} \rangle$

   $\forall\ \mathbf{u}, \mathbf{v}, \mathbf{w} \in U,\ \alpha, \beta \in \mathbb{C}.$

3. *Positiividefiniittisyys*

   (a) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0\ \forall\ \mathbf{u} \in U$,

   (b) $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \ \Leftrightarrow\ \mathbf{u} = \mathbf{0}$.

Havaitaan, että reaalikertoimisesta tapauksesta poikkeavat vain symmetriaehto (1) sekä sekvilineaarisuusehto (2b) (sekvilineaarinen $= 1\frac{1}{2}-$lineaarinen, engl.

sesquilinear – vrt. bilineaarinen = kaksoislineaarinen, engl. bilinear). Itse asissa vain symmetriaehdon ero on olennainen, sillä (1) & (2a) $\Rightarrow$ (2b) (Harj.teht. 1). Huomattakoon, että symmetriaehto (1) pitää erikoistapauksena sisällään ehdon $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$, kun $\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{R}$, joten aiemmin asetettuja reaalikertoimisen tapauksen ehtoja ei erillisinä enää tarvita. Koska ehto (1) myös takaa, että $\langle \mathbf{u}, \mathbf{u} \rangle \in \mathbb{R}$ $\forall \mathbf{u} \in U$, niin ehto (3a) on mielekäs. Uusista ehdoista aksioomina tarpeellisia ovat (1), (2a), (3a) ja (3b):n osa [$\Rightarrow$], sillä (1) & (2a) $\Rightarrow$ (2b), kuten sanottu, ja (2a) $\Rightarrow$ (3b):n osa [$\Leftarrow$].

Skalaaritulon keskeinen ominaisuus on Cauchyn–Schwarzin epäyhtälö, joka vapautettiin geometriasta Luvussa II.4. Varmistetaan nyt, että epäyhtälö pätee myös kompleksikertoimisessa tapauksessa (vrt. Lause II.4.2).

Lause III.4.1 Jokaiselle aksioomat 1–3 toteuttavalle kompleksikertoimisen vektoriavaruuden U skalaaritulolle pätee Cauchyn–Schwarzin epäyhtälö

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|,$$

missä

$$|\mathbf{u}| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}.$$

<u>Todistus</u> Tapaus $\mathbf{u} = \mathbf{0}$ on jälleen selvä, joten voidaan olettaa $\mathbf{u} \neq \mathbf{0}$, jolloin on $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ aksioomien (3a,b) perusteella. Lähdetään jälleen epäyhtälöstä

$$\langle \beta \mathbf{u} + \mathbf{v}, \beta \mathbf{u} + \mathbf{v} \rangle \geq 0,$$

joka nyt on voimassa $\forall \beta \in \mathbb{C}$ (aksiooma (3a)).

Käyttämällä sekvilineaarisuusehtoja (2a,b), symmetriaehtoa (1) ja Luvun III.2 kaavoja (2), (3), (7) tämä purkautuu muotoon

$$\begin{aligned} &\langle \beta \mathbf{u}, \beta \mathbf{u} \rangle + \langle \beta \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \beta \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \\ \Leftrightarrow\ &|\beta|^2 \langle \mathbf{u}, \mathbf{u} \rangle + 2 \operatorname{Re} \{\beta \langle \mathbf{u}, \mathbf{v} \rangle\} + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0. \end{aligned}$$

Koska tämä on voimassa $\forall \beta \in \mathbb{C}$ ja koska $\langle \mathbf{u}, \mathbf{u} \rangle > 0$, voidaan valita

$$\beta = -\frac{\overline{\langle \mathbf{u}, \mathbf{v} \rangle}}{\langle \mathbf{u}, \mathbf{u} \rangle},$$

jolloin Luvun III.2 kaavojen (3), (6) perusteella seuraa

$$\begin{aligned} &-\frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\langle \mathbf{u}, \mathbf{u} \rangle} + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \\ \Leftrightarrow\ &|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \\ \Leftrightarrow\ &|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|. \qquad \square \end{aligned}$$

Esimerkki 1 Yksinkertaisin esimerkki kompleksikertoimisesta sisätuloavaruudesta on $\mathbb{C}$ itse, ts. $(U, \mathbb{C})$, missä $U = \mathbb{C}$. Nimittäin koska kompleksilukujen yhteenlasku on samanlainen operaatio kuin tason vektoreiden yhteenlasku, ja kertolasku voidaan tulkita myös skalaarilla kertomiseksi, niin $(\mathbb{C}, \mathbb{C})$ on vektoriavaruus. Kun tässä avaruudessa määritellään skalaaritulo

$$\langle u, v\rangle = u\overline{v}, \quad u, v \in \mathbb{C},$$

toteutuvat em. ehdot 1–3 (vrt. Luvun III.2 kaavat). Cauchyn–Schwarzin epäyhtälö pelkistyy tässä erikoistapauksessa yhtälöksi:

$$|\langle u, v\rangle| = |u\overline{v}| = |u||v|, \quad u, v \in \mathbb{C}. \quad \square$$

Esimerkki 2 Kompleksikertoimisia sisätuloavaruuksia ovat myös avaruuksien $(\mathbb{R}^2, \mathbb{R})$ ja $(\mathbb{R}^3, \mathbb{R})$ laajennukset $(\mathbb{C}^2, \mathbb{C})$ ja $(\mathbb{C}^3, \mathbb{C})$. Esimerkiksi $(\mathbb{C}^3, \mathbb{C})$:ssä skalaaritulo määritellään

$$\mathbf{u} = (u_1, u_2, u_3) \in \mathbb{C}^3, \quad \mathbf{v} = (v_1, v_2, v_3) \in \mathbb{C}^3 \ :$$
$$\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{i=1}^{3} u_i \overline{v_i}.$$

Avaruuden $(\mathbb{C}^n, \mathbb{C})$, $(n = 2, 3)$ *euklidinen normi* on

$$|\mathbf{u}| = \langle \mathbf{u}, \mathbf{u}\rangle^{1/2} = (\sum_{i=1}^{n} |u_i|^2)^{1/2}. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Näytä, että skalaaritulon sekvilineaarisuusominaisuus (2b) seuraa ehdoista (1) ja (2a).

2. Ovatko avaruuden $(\mathbb{C}^2, \mathbb{C})$ vektorit

$$\mathbf{u} = (3 - 4i,\ 7 - i), \quad \mathbf{v} = (2 - 11i,\ 13 - 9i)$$

   lineaarisesti riippumattomat?

3. Jaa sisätuloavaruuden $(\mathbb{C}^2, \mathbb{C})$ vektori $\mathbf{u} = (1+i, 2-3i)$ kahteen komponenttiin siten, että toinen komponentti on vektorin $\mathbf{v} = (1-i, 2+i)$ suuntainen ja toinen $\mathbf{v}$:tä vastaan kohtisuora.

Luku IV

Reaalimuuttujien funktiot

Matemaattisten funktioiden päätyypit ovat

- *yhden* reaali*muuttujan* reaaliarvoiset *funktiot*, eli reaalifunktiot
- *useamman* reaali*muuttujan* reaaliarvoiset *funktiot*
- yhden tai useamman reaalimuuttujan *vektoriarvoiset funktiot*
- *kompleksifunktiot*, eli kompleksimuuttujan kompleksiarvoiset funktiot

Tässä luvussa aloitetaan funktoiden tutkimus tarkastelemalla yhden tai useamman, toistaiseksi kahden tai kolmen, reaalimuuttujan reaaliarvoisia funktioita sekä yhden tai kahden muuttujan vektoriarvoisia funktioita. Tarkasteltaville funktiotyypeille on yhteistä niiden saama 'näkyvä' muoto, kun luvut, lukuparit ja lukukolmikot muuttujina tai vektorit funktion arvoina ymmärretään euklidisten pisteavaruuksien tai vastaavien vektoriavaruuksien olioina.

Funktioiden tutkimus on matematiikassa hyvin keskeistä, siksi myös tähän liittyvä käsitteistö ja keinovalikoima on huomattavan laaja. Tässä luvussa ei koko 'teknologiaa' oteta vielä käyttöön, vaan rajoitutaan toistaiseksi kaikkein yksinkertaisimpiin algebran ja geometrian keinoihin. Toisaalta sovelluksia (etenkin fysiikan sovelluksia) ajatellen tässä luvussa tarkasteltavien funktioiden tyyppivalikoima on jo melko edustava. Tarkoituksena on tämän valikoiman puitteissa käydä läpi mm. sellaiset funktioiden algebran käsitteet kuin funktioiden *algebralliset yhdistelyt*, *yhdistetty funktio*, *käänteisfunktio* ja *implisiittifunktio*. Selvimmin 'geometrisia funktioita' ovat Luvussa IV.4 esiteltävät *parametriset käyrät* ja *parametriset pinnat*.

IV.1 Yhden muuttujan funktiot

Yhden (reaali)muuttujan (reaaliarvoisella) funktiolla eli reaalifunktiolla tarkoitetaan funktiota tyyppiä

$$f : \mathcal{D}_f \to \mathbb{R}, \quad \mathcal{D}_f \subset \mathbb{R}.$$

Tässä $\mathcal{D}_f$ on f:n *määrittelyjoukko* (lähtöjoukko, engl. domain) ja joukko

$$\mathcal{R}_f = \{y \in \mathbb{R} \mid y = f(x) \text{ jollakin } x \in \mathcal{D}_f\}$$

on f:n *arvojoukko* (engl. range eli 'kantama').† Funktion f *kuvaaja* (engl. graph) joukossa $A \subset \mathcal{D}_f$ on euklidisen tason (yleensä karteesisen) koordinaatiston avulla määritelty pistejoukko

$$G_{f,A} = \{\, P \mathrel{\widehat{=}} (x, y) \in \mathbb{R}^2 \mid x \in A \ \&\ y = f(x) \,\} \subset E^2.$$

Kuvaajan tarkoituksena on 'geometrisoida' funktio niin, että näköhavainnot tulevat mahdollisiksi.‡ Esimerkiksi arvojoukko $\mathcal{R}_f$ on kuvaajasta helppo hahmottaa, ja sellainen usein luontevalta tuntuva sanonta kuin 'f pisteessä x' ($= f(x)$) sisältää myös geometrisointiajatuksen ($x \mathrel{\widehat{=}} P \in E^1$, vrt. Luku II.2).

Reaalimuuttujan funktiota f tutkitaan useimmiten jollakin avoimella, suljetulla tai puoliavoimella välillä, jolle f:n määrittelyjoukko voidaan haluttaessa ajatella

†Suomenkielisissä teksteissä määrittely- ja arvojoukoille käytetään myös merkintöjä M_f, A_f.

‡Kuvaaja on geometrinen vastine reaalifunktion joukko-opilliselle määritelmälle $\mathbb{R}^2$:n osajoukkona, ks. Luku II.5.

rajatuksi ko. tarkastelussa. Kerrattakoon Luvusta I.9 merkinnät

$$\begin{aligned} \text{avoin väli:} \quad (a,b) &= \{x \in \mathbb{R} \mid a < x < b\}, \\ \text{suljettu väli:} \quad [a,b] &= \{x \in \mathbb{R} \mid a \le x \le b\}, \\ \text{puoliavoin väli:} \quad (a,b] &\text{ tai } [a,b) \end{aligned}$$

ja myös yleisessä käytössä olevat 'äärettömän välin' merkinnät

$$\begin{aligned} (0,\infty) &= \{x \in \mathbb{R} \mid x > 0\} = \mathbb{R}_+, \\ (-\infty,0) &= \{x \in \mathbb{R} \mid x < 0\} = \mathbb{R}_-, \\ (-\infty,\infty) &= \mathbb{R}. \end{aligned}$$

Usein yhden muuttujan reaalifunktioista ilmoitetaan vain laskusääntö (liittämissääntö) muodossa 'funktio $f(x)$'. Tällöin oletetaan (ellei toisin mainita), että määrittelyjoukko $\mathcal{D}_f$ on suurin mahdollinen ts.

$$\mathcal{D}_f = \{x \in \mathbb{R} \mid y = f(x) \text{ määritelty yksikäsitteisesti ja } y \in \mathbb{R}\}.$$

Esimerkki 1

a) $f(x) = x^4 + x^2 + 1$ b) $f(x) = x^2/(x-1)^2$ c) $f(x) = \cot x$

d) $f(x) = x^{3/4}$ e) $f(x) = \begin{cases} \cos x, & \text{kun } x < 0 \\ x - x^2, & \text{kun } 0 \le x \le 1 \\ 2 - \sqrt{x}, & \text{kun } x > 1 \end{cases}$ f) $f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^2}$

a) Polynomi: $\mathcal{D}_f = \mathbb{R}$, $\mathcal{R}_f = [1,\infty)$.

b) *Rationaalifunktio*: $\mathcal{D}_f = \{x \in \mathbb{R} \mid x \neq 1\}$, $\mathcal{R}_f = [0,\infty)$.

c) Trigonometrinen funktio: $\mathcal{D}_f = \{x \in \mathbb{R} \mid x/\pi \notin \mathbb{Z}\}$, $\mathcal{R}_f = \mathbb{R}$.

d) *Potenssifunktio*: $\mathcal{D}_f = \mathcal{R}_f = [0,\infty)$.

e) *Paloittain määritelty* funktio: $\mathcal{D}_f = \mathbb{R}$, $\mathcal{R}_f = (-\infty,1]$.

f) Potenssisarja: Määrittelyjoukko on $\mathcal{D}_f = [-1,1]$ (vrt. Luku I.12). Arvojoukko on vaikeampi määrittää, mutta osoittautuu: $\mathcal{R}_f = [-\frac{\pi^2}{12}, \frac{\pi^2}{6}]$. (Sivuutetaan perustelut.) $\square$

Esimerkin b-kohdan rationaalifunktion yleisempi muoto on $f(x) = p(x)/q(x)$, missä p ja q ovat (reaalikertoimisia) polynomeja. Määrittelyjoukko on tällöin $\mathcal{D}_f = \{x \in \mathbb{R} \mid q(x) \neq 0\}$. Potenssifunktion (esimerkin d-kohta) yleinen muoto on $f(x) = x^\alpha$, missä (toistaiseksi) $\alpha \in \mathbb{Q}$. Määrittelyjoukko on joko $\mathcal{D}_f = [0,\infty)$ ($\alpha > 0,\ \alpha \notin \mathbb{N}$), $\mathcal{D}_f = (0,\infty) = \mathbb{R}_+$ ($\alpha < 0,\ \alpha \notin \mathbb{Z}$), $\mathcal{D}_f = \{x \in \mathbb{R} \mid x \neq 0\}$ ($\alpha \in \mathbb{Z},\ \alpha \le 0$) tai $\mathcal{D}_f = \mathbb{R}$ ($\alpha \in \mathbb{N}$).

Esimerkki 2 Määritä funktion $f(x) = x^2 - 7x + 11$ arvojoukko välillä $[1, 4]$.

Ratkaisu Tehtävän asettelun mukaisesti rajataan funktion määrittelyjoukko tässä väliksi $A = [1, 4]$, jolloin arvojoukolle luonteva merkintä on $f(A)$. Koska

$$f(x) = \left(x - \frac{7}{2}\right)^2 - \frac{49}{4} + 11 = \left(x - \frac{7}{2}\right)^2 - \frac{5}{4},$$

niin $f(A)$:n minimi on

$$f_{\min} = \min\{f(x) \mid x \in [1, 4]\} = f(7/2) = -5/4,$$

ja $f(A)$:n maksimi saavutetaan mahdollisimman etäällä pisteestä $x = 7/2$, eli pisteessä $x = 1$:

$$f_{\max} = \max\{f(x) \mid x \in [1, 4]\} = f(1) = 5.$$

Tähän asti on päätelty: $y \in f(A) \Rightarrow -5/4 \leq y \leq 5$, eli $f(A) \subset [-5/4, 5]$. Toisaalta jos $y \in [-5/4, 5]$, niin yhtälöllä $f(x) = y$ on ratkaisu

$$x = \frac{7}{2} - \sqrt{y + \frac{5}{4}},$$

joka on välillä $[1, 4]$. Siis pätee $y \in [-5/4, 5] \Rightarrow y \in f(A)$, eli $[-5/4, 5] \subset f(A)$. Koska on sekä $f(A) \subset [-5/4, 5]$ että $[-5/4, 5] \subset f(A)$, niin on

$$f(A) = [-5/4, 5]. \quad \square$$

Esimerkissä on kyse tyypillisestä 'funktiotutkimuksesta', jossa on annettu joukko $A \subset \mathcal{D}_f$ (reaalifunktion tapauksessa usein väli) ja on määrättävä $B = f(A)$ eli funktion arvojoukko, kun määrittelyjoukko rajataan A:ksi. Ongelma voi olla asetettu myös käänteisesti niin, että tunnetaan joukko B, ja on määrättävä joukko $A = \{x \in \mathcal{D}_f \mid f(x) \in B\}$. Esimerkiksi funktion nollakohtia määrättäessä on ratkaistava käänteinen ongelma, kun $B = \{0\}$, eli on ratkaistava yhtälö $f(x) = 0$. Jos $B = [0, \infty)$, niin on ratkaistava epäyhtälö $f(x) \geq 0$.

Esimerkki 2 (jatko) Jos esimerkissä asetetaan $\mathcal{D}_f = \mathbb{R}$ ja $B = \{0, 1\}$, niin

$$\begin{aligned} A &= \{x \in \mathbb{R} \mid f(x) \in B\} \\ &= \{x \in \mathbb{R} \mid x^2 - 7x + 11 = 0 \ \vee \ x^2 - 7x + 11 = 1\} \\ &= \{2,\ \tfrac{1}{2}(7 - \sqrt{5}),\ \tfrac{1}{2}(7 + \sqrt{5}),\ 5\}. \quad \square \end{aligned}$$

Monotoniset funktiot

Yhtälöitä ja epäyhtälöitä ratkaistaessa, ja muutenkin funktioita tutkittaessa, on käytännössä suurta hyötyä seuraavasta yhden muuttujan funktioille 'luonnetta' antavasta määritelmästä (vrt. Määritelmä I.5.1 lukujonoille).

MÄÄRITELMÄ IV.1.1 Yhden reaalimuuttujan funktio on *kasvava* (engl. increasing) *välillä* $A \subset \mathcal{D}_f$, jos

$$\forall x_1, x_2 \in A \left[x_1 < x_2 \;\Rightarrow\; f(x_1) \leq f(x_2) \right]$$

ja *aidosti* (engl. strictly) *kasvava*, jos

$$\forall x_1, x_2 \in A \left[x_1 < x_2 \;\Rightarrow\; f(x_1) < f(x_2) \right].$$

Vastaavasti f on *vähenevä* (engl. decreasing) *välillä* $A \subset \mathcal{D}_f$, jos

$$\forall x_1, x_2 \in A \left[x_1 < x_2 \;\Rightarrow\; f(x_1) \geq f(x_2) \right]$$

ja *aidosti vähenevä*, jos

$$\forall x_1, x_2 \in A \left[x_1 < x_2 \;\Rightarrow\; f(x_1) > f(x_2) \right].$$

Jos f on välillä A jompaa kumpaa tyyppiä, niin sanotaan, että f on ko. välillä (aidosti) *monotoninen.*

f kasvava

f aidosti vähenevä

ESIMERKKI 3 Identiteetistä

$$\sqrt{x_1} - \sqrt{x_2} = \frac{x_1 - x_2}{\sqrt{x_1} + \sqrt{x_2}}\,, \quad x_1, x_2 \geq 0,\; x_1 \neq x_2$$

nähdään, että funktio $f(x) = \sqrt{x}$ on välillä $[0, \infty)$ (eli määrittelyjoukossaan) aidosti kasvava. $\square$

Jos f on välillä $A \subset \mathcal{D}_f$ aidosti kasvava ja $c \in A$, niin epäyhtälöllä

$$f(x) \leq c$$

on välille A rajattuna helppo ratkaisu:

$$x \in A \cap (-\infty, a], \quad \text{missä } f(a) = c.$$

Esimerkki 3 (jatko) Jos $a > 0$, niin epäyhtälön $\sqrt{x} \leq a$ ratkaisu on

$$\{x \in \mathbb{R} \mid \sqrt{x} \leq a\} = \{x \in [0, \infty) \mid x \leq a^2\} = [0, a^2\,]. \quad \square$$

Jos funktio ei ole koko tarkasteltavalla välillä monotoninen, on funktiotutkimuksen ensimmäinen askel usein välin jakaminen sellaisiin osaväleihin, joilla monotonisuus toteutuu.

Esimerkki 4 $f(x) = \sqrt{|x|}$ $(\mathcal{D}_f = \mathbb{R})$ on aidosti vähenevä välillä $(-\infty, 0]$ ja aidosti kasvava välillä $[0, \infty)$. $\square$

Esimerkki 5 Jos $f(x) = -x^2 + 6x + 1$, niin kirjoittamalla

$$f(x) = -(x-3)^2 + 10$$

nähdään, että f on aidosti kasvava välillä $(-\infty, 3]$ ja aidosti vähenevä välillä $[3, \infty)$. $\square$

Esimerkki 6 Trigonometristen funktioiden määritelmien (Luku II.5) perusteella seuraavat väittämät ovat uskottavia (myös tosia).

$\sin x$: Aidosti monotoninen väleillä $[(k - \frac{1}{2})\pi, (k + \frac{1}{2})\pi]$, $k \in \mathbb{Z}$: kasvava kun k on parillinen ja vähenevä kun k on pariton.

$\cos x$: Aidosti monotoninen väleillä $[k\pi, (k+1)\pi]$, $k \in \mathbb{Z}$: vähenevä kun k on parillinen ja kasvava kun k on pariton.

$\tan x$: Aidosti kasvava väleillä $((k-\frac{1}{2})\pi, (k+\frac{1}{2})\pi)$, $k \in \mathbb{Z}$.

$\cot x$: Aidosti vähenevä väleillä $((k\pi, (k+1)\pi)$, $k \in \mathbb{Z}$. $\square$

Yhdistetty funktio

Kahden funktion f, g *yhdistetty* (engl. composite) *funktio* merkitään $f \circ g$ ja määritellään laskusäännöllä

$$(f \circ g)(x) = f(g(x)).$$

Tavallinen ääntämistapa on hieman arkinen 'f pallo g'.

Esimerkki 7 Palautuva lukujono muotoa

$$x_0 \in \mathbb{R}, \quad x_n = f(x_{n-1}), \quad n = 1, 2, \ldots$$

voidaan tulkita 'sisäkkäisten', ts. yhdistettyjen funktioiden avulla:

$$\begin{aligned} x_1 &= f(x_0), \\ x_2 &= f(x_1) = f(f(x_0)) = (f \circ f)(x_0), \\ x_3 &= f(f(f(x_0))) = (f \circ f \circ f)(x_0), \quad \text{jne.} \quad \square \end{aligned}$$

Yhdistetyn funktion määrittelyjoukko on aina rajattava niin, että funktiot eivät 'riitele'. Tällöin suurimmaksi mahdolliseksi määrittelyjoukoksi (joka yleensä oletetaan, ellei toisin mainita) tulee

$$\mathcal{D}_{f \circ g} = \{x \in \mathcal{D}_g \mid g(x) \in \mathcal{D}_f\} \subset \mathcal{D}_g.$$

Jos näin määritelty joukko $\mathcal{D}_{f \circ g}$ on tyhjä, ei yhdistettyä funktiota voi määritellä.

Esimerkki 8 Määrittele yhdistetyt funktiot $f \circ g$ ja $g \circ f$, kun $f(x) = \sqrt{a-x}$, $g(x) = \sqrt{x-b}$ $(a, b \in \mathbb{R})$.

Ratkaisu Laskusäännöt ovat

$$(f \circ g)(x) = \sqrt{a - \sqrt{x-b}}, \quad (g \circ f)(x) = \sqrt{\sqrt{a-x} - b}.$$

Koska $\mathcal{D}_f = (-\infty, a]$, $\mathcal{D}_g = [b, \infty)$, on

$$\mathcal{D}_{f \circ g} = \{x \in [b, \infty) \mid \sqrt{x-b} \in (-\infty, a]\}.$$

Jos $a < 0$, niin $\mathcal{D}_{f \circ g} = \emptyset$, muuten

$$\mathcal{D}_{f \circ g} \;=\; \{\, x \in \mathbb{R} \mid x \geq b \;\;\&\;\; x - b \leq a^2 \,\} \;=\; [b, b + a^2] \quad (a \geq 0).$$

Siis

$$\mathcal{D}_{f \circ g} = \begin{cases} \emptyset, & \text{jos } a < 0, \\ [b, b + a^2], & \text{jos } a \geq 0. \end{cases}$$

Vastaavalla tavalla päätellään

$$\begin{aligned} \mathcal{D}_{g \circ f} &= \{\, x \in (-\infty, a] \mid \sqrt{a - x} \in [b, \infty) \,\} \\ &= \begin{cases} (-\infty, a], & \text{jos } b \leq 0, \\ (-\infty, a - b^2], & \text{jos } b > 0. \end{cases} \end{aligned} \qquad \square$$

Muuttujan vaihto

Jos tutkimuskohteena oleva funktio f joukossa $A \subset \mathcal{D}_f$ on esitettävissä yhdistettynä funktiona $f(x) = g(v(x))$, niin usein auttaa *muuttujan vaihto* eli *sijoitus* $t = v(x)$. Tällöin voidaan siirtyä tutkimaan (mahdollisesti yksinkertaisempaa) funktiota $g(t)$ joukossa $B = v(A)$.

Esimerkki 9 Määritä funktion $f(x) = x - 7\sqrt{x} + 11$ arvojoukko $f(A)$ välillä $A = [1, 16]$.

Ratkaisu Sijoituksella $t = \sqrt{x}$ tutkimuskohteeksi tulee funktio $g(t) = t^2 - 7t + 11$ välillä $B = [1, 4]$. Siis $f(A) = g(B) = [-5/4, 5]$ (Esimerkki 2). $\square$

Funktioiden yhdistely laskutoimituksilla

Reaaliarvoisia funktioita on mahdollista yhdistellä peruslaskutoimituksilla samalla tavoin kuin lukujonoja. Tällöin ajatellaan, että kun lukujonoja lasketaan yhteen, kerrotaan ja jaetaan termeittäin, niin funktioita yhdistellään vastaavalla tavalla *pisteittäin*. Näin ajatellen saadaan määritellyksi funktioiden väliset peruslaskutoimitukset. Yhden reaalimuuttujan funktioille täsmällisempi määritelmä on seuraava:

MÄÄRITELMÄ IV.1.2 (**Funktioiden yhdistely**) Jos $f : \mathcal{D}_f \to \mathbb{R}$, $g : \mathcal{D}_g \to \mathbb{R}$, $\mathcal{D}_f, \mathcal{D}_g \subset \mathbb{R}$, niin funktiot λf ($\lambda \in \mathbb{R}$), $f + g$, fg ja f/g määritellään

$$\begin{array}{llll} (1) & (\lambda f)(x) & = \lambda f(x), & x \in \mathcal{D}_f \\ (2) & (f+g)(x) & = f(x) + g(x), & x \in \mathcal{D}_f \cap \mathcal{D}_g \\ (3) & (fg)(x) & = f(x)g(x), & x \in \mathcal{D}_f \cap \mathcal{D}_g \\ (4) & (f/g)(x) & = f(x)/g(x), & x \in \mathcal{D}_f \cap \mathcal{D}_g \ \wedge \ g(x) \neq 0 \end{array}$$

Ainoa uusi piirre lukujonoihin nähden on, että funktioita yhdisteltäessä on määrittelyjoukkoa rajoitettava, ellei ole $\mathcal{D}_f = \mathcal{D}_g$. Perusmuotoisten 'jonofunktioiden' tapauksessa tätä ongelmaa ei ollut, koska määrittelyjoukko oli aina sama ($= \mathbb{N}$). Huomattakoon erityisesti, että funktion f/g määrittelyjoukko on ym. määritelmän mukaisesti

$$\mathcal{D}_{f/g} = \{x \in \mathbb{R} \mid x \in \mathcal{D}_f \ \wedge \ x \in \mathcal{D}_g \ \wedge \ g(x) \neq 0\}.$$

Tässä vaatimus $g(x) \neq 0$ esiintyi jo lukujonojen yhteydessä, vrt. Lause I.7.3.

ESIMERKKI 10 Määritelmän IV.1.2 säännön (4) mukaisesti on

$$\tan = \sin / \cos, \quad \cot = \cos / \sin . \quad \square$$

Parilliset ja parittomat funktiot

MÄÄRITELMÄ IV.1.3 Jos $f : \mathcal{D}_f \to \mathbb{R}$ ja f:n määrittelyjoukko $\mathcal{D}_f \subset \mathbb{R}$ on origon suhteen symmetrinen, ts. pätee $-x \in \mathcal{D}_f \ \forall x \in \mathcal{D}_f$, niin sanotaan, että f on *parillinen* (engl. even), jos

$$f(-x) = f(x) \quad \forall x \in \mathcal{D}_f,$$

ja *pariton* (engl. odd), jos

$$f(-x) = -f(x) \quad \forall x \in \mathcal{D}_f.$$

ESIMERKKI 11 Trigonometrisistä funktioista kosini on parillinen ja sini pariton. Potenssifunktio $f(x) = x^n$, $n \in \mathbb{Z}$, on parillinen/pariton kun n on parillinen/pariton. Funktio $f(x) = 0 \ \forall x \in \mathcal{D}_f$ (esim. $\mathcal{D}_f = \mathbb{R}$) on funktioista ainoa, joka on sekä parillinen että pariton. $\square$

Yhdisteltäessä funktioita Määritelmän IV.1.2 mukaisesti ovat seuraavat säännöt helposti todennettavissa:

(1) f ja g parillisia/parittomia $\Rightarrow$ $f+g$, λf ja $1/f$ parillisia/parittomia

(2a) f ja g parillisia/parittomia $\Rightarrow$ fg parillinen

(2b) f parillinen ja g pariton $\Rightarrow$ fg pariton

Esimerkki 11 (jatko) Ym. sääntöjen perusteella (ja muutenkin) voidaan päätellä: a) Trigonometriset funktiot tan ja cot ovat parittomia. b) Polynomi $f(x) = \sum_{k=0}^{n} a_k x^k$ on parillinen/pariton täsmälleen kun $a_k = 0$ jokaisella parittomalla/parillisella indeksin k arvolla. $\square$

Jos funktio on origon suhteen symmetrisesti määritelty, mutta parillisuuden suhteen 'epäpuhdas', niin se voidaan aina esittää parillisen ja parittoman funktion summana. Nimittäin $f = f_+ + f_-$, missä

$$f_+(x) = \frac{1}{2}[f(x) + f(-x)], \quad f_-(x) = \frac{1}{2}[f(x) - f(-x)].$$

Esimerkki 12

$$f(x) \;=\; \begin{cases} 0, & \text{kun } x \le 0 \\ x, & \text{kun } x > 0 \end{cases} \;=\; \frac{1}{2}\,|x| + \frac{1}{2}x \quad \forall x \in \mathbb{R}. \quad \square$$

Jaksolliset funktiot

Määritelmä IV.1.4 Reaalifunktio f on *jaksollinen* eli *periodinen*, jos jollakin $a \in \mathbb{R}_+$ pätee $x \pm a \in \mathcal{D}_f$ $\forall x \in \mathcal{D}_f$ ja

$$f(x+a) = f(x) \quad \forall x \in \mathcal{D}_f.$$

Tällöin a on f:n *jakso*.

Määritelmän mukaisesti myös jokainen jakson monikerta on jakso. Jaksoista pienin on nimeltään *perusjakso* (usein vain 'jaksoksi' sanottu).

ESIMERKKI 13 Funktiot $|\sin x|$, $|\cos x|$, $\tan x$ ja $\cot x$ ovat jaksollisia, (perus)jaksona $a = \pi$. $\square$

ESIMERKKI 14 Jos $\mathcal{D}_f = \mathbb{R}$, f:n jakso on $a = 1$ ja $f(x) = x$, kun $x \in [0, 1)$, niin

$$f(x) = x - k, \quad \text{kun } x \in [k, k+1),\ k \in \mathbb{Z}.$$

HARJOITUSTEHTÄVIÄ

1. Määritä algebran keinoin seuraavien reaalifunktioiden arvojoukot:
   a) $x^2 + 2x + 8$, b) $1 - x - 4x^2$ c) $1/(2 + x + x^2)$ d) $1/(1 - \sqrt{x})$
   e) $x^2/(1 - x^2)$ f) $(x + 1)/(x + 2)$ g) $|x| + |x + 1|$ h) $|x| - |x + 2|$
   i) $\sum_{k=0}^{\infty} x^k$ j) $\sum_{k=0}^{\infty} (-1)^k 2^k x^k$ k) $\sum_{k=0}^{\infty} x^{2k}$ l) $\sum_{k=0}^{\infty} (-1)^k x^{2k}$

2. Selvitä algebran keinoin, millä väleillä seuraavat funktiot ovat aidosti kasvavia ja millä aidosti väheneviä: a) $1/x^4$, b) $x/(x + 1)$, c) $|x^2 + x - 2|$, d) $1/(x^2 + 2x + 2)$, e) $1/(x^2 + 3x + 2)$, f) $x^4/(2 - x^4)$ g) $\max\{0, \sin x\}$

3. Mitkä ovat seuraavien yhdistettyjen funktioiden määrittely- ja arvojoukot?
   a) $\sqrt{8 - 2x}$ b) $\sqrt{1 - x - x^2}$ c) $1/(1 - \sqrt{2 - x})$ d) $\cos(\sin x)$

4. Määritä yhdistettyjen funktioiden $f \circ g$ ja $g \circ f$ laskusäännöt ja määrittelyjoukot seuraavissa tapauksissa:
   a) $f(x) = 1/\sqrt{x + 1}$, $g(x) = \sqrt{x - 1}$
   b) $f(x) = \sqrt{x + 1}$, $g(x) = x/(1 - 2x)$
   c) $f(x) = \sqrt{2x + 1}$, $g(x) = \sin x$

5. a) Olkoon $f(x) = 1 - 4\sqrt{x}$. Mitkä ovat funktioiden $f \circ f$ ja $f \circ f \circ f$ määrittelyjoukot? b) Olkoon $g(x) = \frac{1}{2}x + 1$. Millainen reaalifunktio on $f(x) = \lim_n g_n(x)$, missä $g_1 = g$, $g_2 = g \circ g$, $g_3 = g \circ g \circ g$, jne.?

6. Määritä algebran keinoin seuraavien funktioiden arvojoukot:
a) $1-\sqrt[4]{17x}+\sqrt{x}$ b) $2+5x^{48}-x^{96}$ c) $x-4\sqrt{x}+3-4|\sqrt{x}-1|$
d) $x^4/(2-x^4)$ e) $\cos x-7\sin x$ f) $\sin x+\cos 2x$ g) $2\sin x-|\cos 2x|$

7. a) Näytä, että jos g on kasvava välillä $A\subset\mathcal{D}_g$ ja f on kasvava välillä $B=g(A)\subset\mathcal{D}_f$, niin $f\circ g$ on kasvava välillä A. Sovella väittämää funktioon $f(x)=\sqrt{x^2+x-2}$ välillä $[1,\infty)$.
b) Näytä, että jos f ja g ovat (aidosti) kasvavia välillä $A\subset\mathcal{D}_f\cap\mathcal{D}_g$, niin samoin on funktio $f+g$. Miten on funktion fg laita?

8. Määrittele $f+g$, fg ja f/g (määrittelyjoukko ja sievennetty laskusääntö), kun a) $f(x)=g(x)=x+1$, b) $f(x)=|x|+x$, $g(x)=|x|-x$,
c) $f(x)=\sin x\sin 2x$, $g(x)=2\cos^3 x$.

9. Näytä: a) Funktion jako parilliseen ja parittomaan osaan on yksikäsitteinen. b) Jos g on parillinen, niin $f\circ g$ on parillinen tai ei määritelty. c) Jos f on parillinen/pariton ja g on pariton, niin $f\circ g$ on parillinen/pariton tai ei määritelty.

10. Jaa seuraavat funktiot parilliseen ja parittomaan osaan:
a) $f(x)=2+x-3x^2+x^4+x^5+\sin x-3\cos x$
b) $f(x)=|x+1|+|x-1|-x+2x^2$
c) $f(x)=\sin(x+\frac{\pi}{4})+\cos(x+\frac{\pi}{3})$

11. a) Näytä, että jos f on (aidosti) kasvava/vähenevä väleillä A_1 ja A_2 ja $A_1\cap A_2\neq\emptyset$, niin f on (aidosti) kasvava/vähenevä välillä $A=A_1\cup A_2$.
b) Olkoon $\mathcal{D}_f=\mathbb{R}$ ja f (aidosti) kasvava välillä $[0,\infty)$. Näytä, että jos lisäksi f on parillinen/pariton, niin f on (aidosti) vähenevä välillä $(-\infty,0]$ / (aidosti) kasvava $\mathbb{R}$:ssä.

12. Funktio $f(x)=\sin\frac{x}{5}+\cos\frac{x}{7}$ on jaksollinen. Mikä on perusjakso?

13. Funktion g jakso on $a=2$ ja $g(x)=1-|x|$, kun $x\in[-1,1]$. Määritä funktion $f(x)=g(x)+cx-5$ nollakohdat, kun a) $c=1$, b) $c=1/2$.

14. Määritellään porrasfunktio
$$f(x)=k,\quad \text{kun}\ \ 2k-2\le x<2k,\ k\in\mathbb{Z}.$$
Määritä $a\in\mathbb{R}$ ja jaksollinen funktio g siten, että $f(x)=ax+g(x)$, $x\in\mathbb{R}$.

15. (*) Todista pelkin algebran keinoin:
a) Funktion $f(x)=(3x^2+3)/(x^2+x+1)$ arvojoukko on $\mathcal{R}_f=[2,6]$.
b) Funktio $f(x)=x^2/(x+1)$ on aidosti kasvava väleillä $(-\infty,-2]$ ja $[0,\infty)$ ja aidosti vähenevä väleillä $[-2,-1)$ ja $(-1,0]$.

IV.2 Käänteisfunktio. Implisiittifunktiot

Sanotaan, että reaalifunktio $f : \mathcal{D}_f \to \mathbb{R}$ on 1-1 ('yksi yhteen', engl. one to one) eli *kääntyvä* (engl. invertible) eli *injektio* eli *injektiivinen*, jos pätee

$$\forall x_1, x_2 \in \mathcal{D}_f \; [\, x_1 \neq x_2 \;\Rightarrow\; f(x_1) \neq f(x_2)\,].$$

Jos f on 1-1 ja $y \in \mathcal{R}_f$, niin yhtälöllä

$$f(x) = y \quad (x \in \mathcal{D}_f)$$

on <u>yksikäsitteinen</u> ratkaisu. Koska jokaiseen $y \in \mathcal{R}_f$ liittyy tällä tavoin yksikäsitteinen $x \in \mathcal{D}_f$, niin kyseessä on funktioriippuvuus $y \mapsto x$, joka merkitään

$$x = f^{-1}(y).$$

Sanotaan, että f^{-1} (luetaan 'f miinus 1', engl. 'f inverse') on f:n *käänteisfunktio*. Käänteisfunktion määrittelyjoukko on siis $\mathcal{D}_{f^{-1}} = \mathcal{R}_f$ ja arvojoukko $\mathcal{R}_{f^{-1}} = \mathcal{D}_f$.

Esimerkki 1 Funktiolle $f(x) = 1/x$ ($\mathcal{D}_f = \{x \in \mathbb{R} \mid x \neq 0\}$) pätee

$$f(x_1) - f(x_2) \;=\; \frac{1}{x_1} - \frac{1}{x_2} \;=\; \frac{x_2 - x_1}{x_1 x_2}\,, \quad x_1\,, x_2 \in \mathcal{D}_f\,.$$

Tämän perusteella on $f(x_1) - f(x_2) \neq 0$ aina kun $x_1, x_2 \neq 0$ ja $x_1 \neq x_2$, joten f on 1-1. Käänteisfunktio löydetään ratkaisemalla yhtälö $f(x) = y$:

$$x^{-1} = y \quad \Rightarrow \quad x = y^{-1}, \quad \text{jos } y \neq 0.$$

Jos $y = 0$, ei yhtälöllä ole ratkaisua, joten käänteisfunktion määrittelyjoukko (= f:n arvojoukko) = $\{y \in \mathbb{R} \mid y \neq 0\}$, ja $f^{-1}(y) = 1/y$. Koska siis $\mathcal{D}_{f^{-1}} = \mathcal{D}_f$ ja $f^{-1}(y) = f(y)\ \forall y \in \mathcal{D}_{f^{-1}} = \mathcal{D}_f$, niin $f^{-1} = f$. $\square$

Määritelmän IV.1.1 nojalla on selviö, että jos funktion f määrittelyjoukko on <u>väli</u> ja f on ko. välillä aidosti kasvava tai aidosti vähenevä (eli aidosti monotoninen), niin f on 1-1. Aito monotonisuus onkin käytännössä tavallisin injektiivisyyden olomuoto silloin kun funktion määrittelyjoukko on väli (tai väliksi rajattu).

Esimerkki 2 Näytä, että funktio $f(x) = x^m$, $m \in \mathbb{N}$ on injektio, kun määrittelyjoukko rajataan väliksi $[0, \infty)$. Määritä käänteisfunktio.

<u>Ratkaisu</u> Kirjoitetaan (ks. Propositio I.2.2)

$$f(x_1) - f(x_2) \;=\; x_1^m - x_2^m \;=\; (x_1 - x_2)(x_1^{m-1} + x_1^{m-2}x_2 + \cdots + x_2^{m-1}).$$

Jos tässä on $x_1, x_2 \geq 0$ ja $x_1 \neq x_2$, niin viimeksi kirjoitetun tulon jälkimmäisessä tekijässä on jokainen yhteenlaskettava ei-negatiivinen ja ainakin yksi on positiivinen (koska joko $x_1 > 0$ tai $x_2 > 0$), joten ko. tekijä on positiivinen. Päätellään, että f on välillä $[0, \infty)$ aidosti kasvava:

$$\forall x_1, x_2 \geq 0 \ [\, x_1 < x_2 \ \Rightarrow \ f(x_1) < f(x_2)\,].$$

Siis f on injektio, joten yhtälöllä $f(x) = y$ ($y \in \mathbb{R}$) on enintään yksi ratkaisu x välillä $[0, \infty)$. Jos $y < 0$, ei ratkaisua ole. Jos $y = 0$, on ratkaisu $x = 0$. Lopulta jos $y > 0$, on ratkaisu myös olemassa ja merkitään $x = \sqrt[m]{y}$. (Luku $\sqrt[m]{y} \in \mathbb{R}_+$ on laskettavissa esim. kymmenjakoalgoritmilla, vrt. Luku I.9.) Kysytty käänteisfunktio on siis $f^{-1}(x) = \sqrt[m]{x}$, määrittelyjoukkona $\mathcal{D}_{f^{-1}} = \mathcal{R}_f = [0, \infty)$. $\square$

Esimerkki 2 (jatko) Jos m on <u>pariton</u>, niin $f(x) = x^m$, $\mathcal{D}_f = \mathbb{R}$, on pariton ja näin muodoin aidosti kasvava koko $\mathbb{R}$:ssä (ks. Harj.teht. IV.1:11b). Siis f on injektio. Yhtälöllä $f(x) = y$ on tällöin (yksikäsitteinen) ratkaisu jokaisella $y \in \mathbb{R}$, sillä jos $y < 0$, niin ratkaisu on $x = -\sqrt[m]{-y}$. Käänteisfunktio merkitään yleensä yksinkertaisesti $f^{-1}(x) = \sqrt[m]{x}$, jolloin siis sovitaan, että

$$\sqrt[m]{-x} = -\sqrt[m]{x}, \quad \text{kun } x > 0 \text{ ja } m \text{ on pariton.}^{\dagger} \quad \square$$

Seuraava yleinen sääntö on Määritelmästä IV.1.1 ja käänteisfunktion määritelmästä helposti johdettavissa (Harj.teht. 1a):

$$\boxed{\begin{array}{l} f \text{ aidosti kasvava/vähenevä välillä } A \subset \mathcal{D}_f \\ \qquad \Rightarrow \quad f^{-1} \text{ aidosti kasvava/vähenevä väleillä } B \subset f(A). \end{array}}$$

Esimerkki 2 (jatko) Koska $f(x) = x^m$ ($m \in \mathbb{N}$) on aidosti kasvava välillä $A = [0, \infty)$, niin $f^{-1}(x) = \sqrt[m]{x}$ on samoin aidosti kasvava välillä $[0, \infty) = f(A)$. $\square$

Käänteisfunktion määritelmästä helposti todennettavissa ovat myös yleiset funktioalgebran lait

$$\boxed{\begin{array}{rl} (f^{-1} \circ f)(x) = x & \forall x \in \mathcal{D}_f, \\ (f \circ f^{-1})(y) = y & \forall y \in \mathcal{R}_f, \\ (f^{-1})^{-1} = f. & \end{array}}$$

†Juurilukujen määrittelyn laajentaminen mainitulla tavalla ei ole aivan ongelmatonta, kuten nähdään laskusta

$$1 = \sqrt[6]{1} = \sqrt[6]{(-1)^2} = (-1)^{2/6} = (-1)^{1/3} = \sqrt[3]{-1} = -1.$$

Tämän tyyppisten ristiriitojen välttämiseksi on selvintä sopia, että reaalinen potenssifunktio $f(x) = x^\alpha$ on määritelty välillä $(-\infty, 0)$ vain kun $\alpha \in \mathbb{Z}$.

Käänteisfunktion f^{-1} kuvaaja saadaan vaihtamalla x ja y f:n kuvaajassa, eli peilaamalla f:n kuvaaja suoran $y = x$ suhteen. Kuvassa $f(x) = x^2$, $\mathcal{D}_f = [0, \infty)$.

Esimerkki 3 Tutki funktion $f(x) = x^5 + 3x$, $\mathcal{D}_f = \mathbb{R}$ (mahdollista) käänteisfunktiota.

<u>Ratkaisu</u> Koska f on pariton ja

$$f(x_1) - f(x_2) = (x_1 - x_2)(x_1^4 + x_1^3 x_2 + x_1^2 x_2^2 + x_1 x_2^3 + x_2^4 + 3),$$

niin päätellän kuten Esimerkissä 2, että f on aidosti kasvava koko $\mathbb{R}$:ssä. Käänteisfunktio $f^{-1} : \mathcal{R}_f \to \mathbb{R}$ on siis olemassa. Kun $y \in \mathcal{R}_f$, niin funktioevaluaatio $y \mapsto x = f^{-1}(y)$ tarkoittaa yhtälön

$$x^5 + 3x = y$$

ratkaisemista. Ratkeavuus jokaisella $y > 0$ on osoitettavissa (Harj.teht. 12), ja koska f on pariton, niin yhtälö ratkeaa myös jokaisella $y \leq 0$, t.s. $\mathcal{R}_f = \mathcal{D}_{f^{-1}} = \mathbb{R}$. Nähdään myös, että esimerkiksi $f^{-1}(0) = 0$, $f^{-1}(4) = 1$, ja $f^{-1}(-38) = -2$. Sen sijaan vaikkapa lukua $a = f^{-1}(1)$ ei voi määrätä 'tarkasti' edes juurilukujen avulla, vaan kyseessä on yleisempi (algebrallinen, vrt. Luku I.11) luku. Tällaisen luvun määrittelyssä on tyydyttävä reaaliluvun yleiseen määritelmään, esim. äärettömänä desimaalilukuna (ks. Harj.teht. 12). Symbolinen laskenta luvulla a on toki myös mahdollista, mutta tällainen laskenta määritelmän $a^5 + 3a = 1$ perusteella on varsin rajoitettua. — Laihan lohdun suokoon tulos

$$a^{-1} = a^4 + 3. \quad \square$$

Injektio, surjektio ja bijektio

Kuten Esimerkissä 2 edellä, on käänteisfunktioita tutkittaessa varsin tavallista, että f ei välttämättä ole injektiivinen koko määrittelyjoukossaan, mutta on kuitenkin injektio, jos määrittelyjoukkoa sopivasti rajoitetaan. Käänteisfunktiotarkastelujen lähtökohdaksi voidaan tällöin ottaa rajoitettu kuvaus

$$f : A \to B,$$

missä $A \subset D_f$ ja $B \supset f(A)$. Yleensä A on D_f:n jokin osaväli. Joukkoon A rajattua funktiota voidaan haluttaessa merkitä $f_{|A}$ ja sanoa, että kyseessä on *f:n rajoittuma A:lle*. Jos tässä B on vielä onnistuttu valitsemaan siten, että $B = f(A)$, niin sanotaan, että $f : A \to B$ on *surjektio B:lle* (ransk. sur jeter = heittää päälle; englanninkielinen 'f onto B' sisältää saman ajatuksen). Termiä käytetään myös silloin kun f ei ole injektio. Siis funktiosta $f : A \to B$ saadaan surjektio yksinkertaisesti rajaamalla maalijoukko $f(A)$:ksi. Jos $f : A \to B$ $(= f_{|A})$ on sekä injektio että surjektio, niin sanotaan, että $f : A \to B$ on *bijektio* (engl. one to one and onto). Tällöin myös käänteisfunktio $f^{-1} : B \to A$ (tässä siis $f^{-1} = (f_{|A})^{-1}$) on bijektio.†

Esimerkki 4 Jos $f(x) = x^m$, niin $f : [0, \infty) \to [0, \infty)$ on bijektio jokaisella $m \in \mathbb{N}$, samoin esim. $f : (0, \infty) \to (0, \infty)$, $f : [0, 1] \to [0, 1]$ ja $f : [1, \infty) \to [1, \infty)$ (vrt. Esimerkki 2). □

Esimerkki 5 Jos $f(x) = x^2$, niin

$f : [-2, 3] \to [0, 10]$ on funktio, mutta ei injektio eikä surjektio.
$f : [-2, 3] \to [0, 9]$ on surjektio, ei injektio.
$f : [0, 3] \to [0, 10]$ on injektio, ei surjektio: käänteisfunktio $f^{-1}(x) = \sqrt{x}$.
$f : [0, 3] \to [0, 9]$ on bijektio: käänteisfunktio $f^{-1}(x) = \sqrt{x}$.
$f : [-2, -1] \to [1, 4]$ on bijektio: käänteisfunktio $f^{-1}(x) = -\sqrt{x}$. □

Trigonometriset käänteisfunktiot

Trigonometriset funktiot eivät ole injektiivisiä koko määrittelyjoukossaan, joten ne eivät ole 'kääntyviä' tavallisessa mielessä. Trigonometrisilla käänteisfunktioilla tarkoitetaankin tavallisimmin funktioita, jotka saadaan rajoittamalla trigonometrinen funktio joko välille $[-\frac{\pi}{2}, \frac{\pi}{2}]$ tai $(-\frac{\pi}{2}, \frac{\pi}{2})$ (sin, tan) tai välille $[0, \pi]$ tai

†Termit injektio, surjektio ja bijektio eivät rajoitu reaalifunktioihin, vaan ne voidaan liittää yhtä hyvin yleiseen joukko-opilliseen funktiokäsitteeseen (ks. Luku II.5). Jos A ja B ovat joukkoja, ja on olemassa bijektio $f : A \to B$, niin sanotaan, että A:n ja B:n välillä on (bijektion f luoma) *kääntäen yksikäsitteinen vastaavuus* ja merkitään $A \leftrightarrow B$. Esim. Luvun II.2 merkinnässä $E^2 \leftrightarrow \mathbb{R}^2$ tarkoitetaan tällaista vastaavuutta annetussa tason koordinaatistossa.

$(0, \pi)$ (cos, cot), jolloin funktio on ko. välillä aidosti monotoninen. Näin saatuja käänteisfunktioita sanotaan *arkusfunktioiksi* tai *syklometrisiksi funktioiksi.* Määritelmät ovat

$$\begin{array}{lll} x = \sin y & \& \quad y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] & \Leftrightarrow \quad y = \operatorname{Arcsin} x, \\ x = \cos y & \& \quad y \in [0, \pi] & \Leftrightarrow \quad y = \operatorname{Arccos} x, \\ x = \tan y & \& \quad y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) & \Leftrightarrow \quad y = \operatorname{Arctan} x, \\ x = \cot y & \& \quad y \in (0, \pi) & \Leftrightarrow \quad y = \operatorname{Arccot} x. \end{array}$$

Nämä luetaan 'arkus sini' jne. Kyseessä ovat trigonometristen käänteisfunktioiden nk. *päähaarat.*[†]

Arkusfunktioiden määrittely- ja arvojoukot ovat määritelmien perusteella seuraavat:

$$\begin{array}{lll} \operatorname{Arcsin} x & : & \mathcal{D}_f = [-1, 1], \quad \mathcal{R}_f = [-\frac{\pi}{2}, \frac{\pi}{2}]. \\ \operatorname{Arccos} x & : & \mathcal{D}_f = [-1, 1], \quad \mathcal{R}_f = [0, \pi]. \\ \operatorname{Arctan} x & : & \mathcal{D}_f = \mathbb{R}, \quad \mathcal{R}_f = (-\frac{\pi}{2}, \frac{\pi}{2}). \\ \operatorname{Arccot} x & : & \mathcal{D}_f = \mathbb{R}, \quad \mathcal{R}_f = (0, \pi). \end{array}$$

Funktiot Arcsin ja Arctan ovat määrittelyvälillään aidosti kasvavia (kuten sin ja tan), muut kaksi aidosti väheneviä.

$y = \operatorname{Arcsin} x$

$y = \operatorname{Arccos} x$

[†]Muille väleille rajoitettujen trigonometristen funktioiden käänteisfunktioista käytetään tässä tekstissä yhteismerkintää arcsin, jne. — ks. tarkempi määrittely luvun lopussa. Kirjallisuudessa trigonometristen käänteisfunktioiden merkintätavat ovat hieman kirjavat. Vanhempia päähaarojen merkintöjä ovat $\overline{\text{arc}}$sin, $\overline{\text{arc}}$cos, $\overline{\text{arc}}$tan, $\overline{\text{arc}}$cot. Laskimissa tavallisia merkintöjä $\sin^{-1}$, $\cos^{-1}$ ja $\tan^{-1}$ näkee myös käytettävän kirjallisuudessa. — Tässä tekstissä määritellään $\sin^n x = (\sin x)^n$ $\forall n \in \mathbb{Z}$, jolloin $\sin^{-1} x = 1/\sin x$.

Funktiopareista (Arcsin, Arccos) ja (Arctan, Arccot) riittää sovelluskäyttöön valita yksi edustaja kummastakin parista (tavallisimmin valitaan Arcsin ja Arctan), sillä parien 'jäsentenvälinen' ratkeaa yksinkertaisiin yhteyksiin

$$\boxed{\begin{aligned} \operatorname{Arcsin} x + \operatorname{Arccos} x &= \frac{\pi}{2} \quad \forall x \in [-1,1], \\ \operatorname{Arctan} x + \operatorname{Arccot} x &= \frac{\pi}{2} \quad \forall x \in \mathbb{R}. \end{aligned}}$$

Näistä esimerkiksi ensimmäinen seuraa päättelyllä (vrt. kuviot edellä)

$$\begin{cases} \dfrac{\pi}{2} - \operatorname{Arccos} x \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right] & \forall x \in [-1,1] \\ \sin\left(\dfrac{\pi}{2} - \operatorname{Arccos} x\right) = \cos(\operatorname{Arccos} x) = x & \forall x \in [-1,1] \end{cases}$$
$$\Rightarrow \quad \frac{\pi}{2} - \operatorname{Arccos} x = \operatorname{Arcsin} x \quad \forall x \in [-1,1].$$

Esimerkki 6 Hahmottele funktion $\operatorname{Arcsin}(\cos x)$ kuvaaja.

<u>Ratkaisu</u> Sievennetään ensin $y = \operatorname{Arcsin}(\cos x)$:

$$\begin{aligned} y &= \operatorname{Arcsin}(\cos x) \\ &\Leftrightarrow \quad y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \& \quad \sin y = \cos x \\ &\Leftrightarrow \quad y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \& \quad \sin y = \sin\left(\frac{\pi}{2} - x\right) \\ &\Leftrightarrow \quad y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \& \quad y = \frac{\pi}{2} \pm x + n \cdot 2\pi, \ n \in \mathbb{Z}. \end{aligned}$$

Tämän perusteella kuvaaja koostuu joukosta suoria rajoitettuna välille $y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ (ks. kuvio). $\square$

Implisiittifunktio $y(x)$

Jos yhtälöstä muotoa

$$F(x, y) = 0$$

on y ratkaistavissa yksikäsitteisesti jokaisella $x \in A$ ($A \subset \mathbb{R}$), niin sanotaan, että yhtälö määrittelee A:ssa *implisiittifunktion* $y = f(x)$. Termillä tarkoitetaan, että funktio on määritelty epäsuorasti eli *implisiittisesti*. Tällöin ei edellytetä, että funktio on kirjoitettavissa suoraan (eli eksplisiittisesti) tunnettuna lausekkeena. Yhtälössä F on kahden reaalimuuttujan reaaliarvoinen funktio — käytännössä lauseke, jossa x ja y esiintyvät. (Kahden muuttujan funktioita itsenäisinä olioina käsitellään seuraavassa luvussa.)

Implisiittifunktiota merkittäessä halutaan usein säilyttää 'paikan tuntu' käyttämällä yleismerkinnän $f(x)$ sijasta merkintää $y(x)$. Tämän mukaisesti siis funktioriippuvuus $x \mapsto y(x)$ määritellään epäsuoralla laskusäännöllä

$$F(x, y(x)) = 0, \quad x \in A.$$

Yksinkertainen esimerkki implisiittifunktiosta on käänteisfunktio, sillä jos käänteisfunktion määrittelyssä vaihdetaan x ja y, niin määritelmä on

$$F(x, y) = x - f(y) = 0, \;\; y \in \mathcal{D}_f \quad \Leftrightarrow \quad y(x) = f^{-1}(x).$$

— Itse asiassa myös tavallinen funktio muuttuu 'implisiittiseksi', kun $y = f(x)$ kirjoitetaan muotoon $F(x, y) = f(x) - y = 0$.

Esimerkki 7 Jos $b \neq 0$, niin tason suoran perusmuotoinen yhtälö

$$ax + by + c = 0$$

voidaan tulkita implisiittifunktion $y(x) = -(ax + c)/b$ määrittelyksi. □

Esimerkki 8 Tarkasteltaessa yhtälöä

$$F(x,y) = x + y + \cos x \sin y = 0$$

osoittautuu (tarkemmat perustelut sivuutetaan), että yhtälöllä on yksikäsitteinen ratkaisu $y(x)$ jokaisella $x \in \mathbb{R}$, t.s. yhtälö määrittelee koko $\mathbb{R}$:ssä implisiittifunktion $y(x)$. Yhtälöstä on pääteltävissä, että $y(x) = -x$ kun $x = (n + \frac{1}{2})\pi$ ja samoin kun $x = n\pi,\ n \in \mathbb{Z}$. Muilla x:n arvoilla funktioevaluaatio $x \mapsto y(x)$ on työläämpi, sillä kyseessä on nk. *transkendenttinen yhtälö*, joka ratkeaa vain numeerisin keinoin. Yhtälöstä on myös x ratkaistavissa yksikäsitteisesti jokaisella $y \in \mathbb{R}$. Näin määräytyy (joskaan ei lausekkeena) funktio $y \mapsto x(y)$, joka on funktion $x \mapsto y(x)$ käänteisfunktio. $\square$

Monihaaraiset implisiittifunktiot

Implisiittifunktion määrittely-yrityksissä varsin tavallinen on tilanne, että yhtälö $F(x,y) = 0$ ratkeaa y:n suhteen annetulla x, mutta ei yksikäsitteisesti. Olkoon tällaisessa tilanteessa $A \subset \mathbb{R}$ väli (suljettu, avoin tai puoliavoin) ja tarkastellaan yhtälöä, kun $x \in A$. Oletetetaan, että on olemassa välille A ominainen indeksijoukko $\Lambda \subset \mathbb{N}$ (voi olla myös $\Lambda = \mathbb{N}$) ja funktio<u>joukko</u>

$$\mathcal{Y} = \{y_i,\ i \in \Lambda\}$$

siten, että pätee

$$\forall x \in A\ [\, F(x,y) = 0 \ \Leftrightarrow\ y \in \{y_i(x),\ i \in \Lambda\}\,].$$

Sanotaan tällöin, että $\mathcal{Y}$ on yhtälön $F(x,y) = 0$ välillä A määrittelemä *monihaarainen* (moniarvoinen) *implisiittifunktio*. Monihaarainen implisiittifunktio ei siis ole funktio vaan sellaisten joukko. Joukon $\mathcal{Y}$ alkioita y_i sanotaan $\mathcal{Y}$:n *haaroiksi* (engl. branch).† Kuviossa haaroja on kolme välillä $A = [a,b]$.‡

†Monihaarainen funktio $\mathcal{Y}$ voidaan myös tulkita *joukkoarvoiseksi* määrittelemällä

$$\mathcal{Y}(x) = \{y_i(x),\ i \in \Lambda\},$$

missä $y_i,\ i \in \Lambda$ ovat $\mathcal{Y}$:n haarat tarkasteltavalla välillä A. Koska $\mathcal{Y}(x)$ on (joukkona) yksikäsitteinen jokaisella $x \in A$, niin $\mathcal{Y}$ on funktio tyyppiä $\mathcal{Y} : A \to \mathcal{B}$, missä maalijoukko $\mathcal{B}$ on 'joukkojen joukko' $\mathcal{B} = \{\, \{t_i,\ i \in \Lambda\} \mid t_i \in \mathbb{R}\ \ \forall i \in \Lambda\,\}$.

‡Jotta monihaaraisen funktion $\mathcal{Y} = \{y_i\}$ haarat y_i olisivat yksiselitteiset, on määrittelyssä estettävä 'hyppiminen haaralta toiselle'. Ehdon voi tulkita geometris–intuitiivisesti niin, että kuvaajien $G_i = \{P = (x,y) \in E^2 \mid y = y_i(x)\ \ \&\ \ x \in A\}$ on oltava *yhtenäisiä* tason pistejoukkoja, ts. kuvaajissa ei sallita 'katkoksia'. (Täsmällisemmin tarkoitetaan, että funktioiden y_i on oltava *jatkuvia*, ks. Luku V.1 jäljempänä.)

Esimerkki 9 Yhtälö $x^2+y^2=1$ määrittelee välillä $[-1,1]$ kaksihaaraisen implisiittifunktion $\mathcal{Y}=\{y_+,y_-\}$, missä $y_\pm=\pm\sqrt{1-x^2}$. $\square$

Esimerkki 10 Millaisen monihaaraisen implisiittifunktion määrittelee yhtälö $y^4-2y^2+x^2-x=0$ eri $\mathbb{R}$:n osaväleillä?

<u>Ratkaisu</u> Jos $x\in\mathbb{R}$, niin yhtälön mahdolliset ratkaisut $y\in\mathbb{R}$ ovat

$$y=\pm\sqrt{1\pm\sqrt{1+x-x^2}}.$$

Tässä on

$$\begin{aligned} 1+x-x^2\geq 0 \quad &\Leftrightarrow \quad \frac{1}{2}(1-\sqrt{5})\leq x\leq\frac{1}{2}(1+\sqrt{5}),\\ 1+x-x^2\geq 1 \quad &\Leftrightarrow \quad 0\leq x\leq 1. \end{aligned}$$

Päätellään, että välillä $[0,1]$ funktio on kaksihaarainen: $\mathcal{Y}=\{y_1\,,\,y_2\}$, missä

$$y_1(x)=\sqrt{1+\sqrt{1+x-x^2}}\,,\quad y_2(x)=-\sqrt{1+\sqrt{1+x-x^2}}\,.$$

Väleillä $[\frac{1}{2}(1-\sqrt{5}),\,0]$ ja $[1,\,\frac{1}{2}(1+\sqrt{5})]$ on $0\leq 1+x-x^2\leq 1$, joten näillä väleillä funktio on nelihaarainen: $\mathcal{Y}=\{y_1\,,\,y_2\,,\,y_3\,,\,y_4\}$, missä y_1 ja y_2 ovat samat kuin edellä ja

$$y_3(x)=\sqrt{1-\sqrt{1+x-x^2}}\,,\quad y_4(x)=-\sqrt{1-\sqrt{1+x-x^2}}\,.$$

Jos $x\in(-\infty,\,\frac{1}{2}(1-\sqrt{5}))$ tai $x\in(\frac{1}{2}(1+\sqrt{5}),\,\infty)$, ei yhtälö toteudu millään $y\in\mathbb{R}$, joten näillä väleillä yhtälö ei määrittele mitään (reaali)funktiota. $\square$

Funktiot arcsin, arccos, arctan

Funktiomerkinnöillä arcsin, arccos ja arctan tarkoitetaan yhtälöiden $x - \sin y = 0$, $x - \cos y = 0$ ja $x - \tan y = 0$ määrittelemiä äärettömän monihaaraisia funktioita, määrittelyjoukkona joko väli $[-1, 1]$ (arcsin, arccos) tai $\mathbb{R}$ (arctan). Esimerkiksi arctan määritellään

$$\arctan = \{y_n,\ n \in \mathbb{Z}\},$$

missä

$$y_n(x) = \operatorname{Arctan} x + n\pi.$$

Tässä $y_0(x) = \operatorname{Arctan} x$ on arctan:n *päähaara* (engl. principal branch).

Esimerkki 11 Jos $\varphi(x, y)$ on karteesisen koordinaatiston pistettä (x, y) vastaava napakulma, $\varphi \in [0, 2\pi)$, niin funktioriippuvuus $(x, y) \mapsto \varphi(x, y)$ on

$$\varphi(x,y) = \begin{cases} \operatorname{Arctan}(y/x), & \text{jos } x > 0 \ \& \ y \geq 0, \\ \operatorname{Arctan}(y/x) + \pi, & \text{jos } x < 0, \\ \operatorname{Arctan}(y/x) + 2\pi, & \text{jos } x > 0 \ \& \ y < 0, \\ \pi/2, & \text{jos } x = 0 \ \& \ y > 0, \\ 3\pi/2, & \text{jos } x = 0 \ \& \ y < 0. \end{cases}$$

Tässä tarvittiin siis peräti kolmea arctan:n haaraa (!). □

HARJOITUSTEHTÄVIÄ

1. Todista: a) Jos f on välillä $A \subset \mathcal{D}_f$ aidosti kasvava tai vähenevä, niin samoin on f^{-1} väleillä $B \subset f(A)$. b) Jos f on 1-1 ja pariton, niin samoin on f^{-1}.

2. a) Funktio $f(x) = x^2 + x + 2$ on kääntyvä, kun se rajoitetaan välille $(-\infty, a]$ tai välille $[a, \infty)$. Määritä a ja käänteisfunktio kummallakin välillä.

   b) Näytä algebran keinoin, että funktio $f(x) = x^3 + 2x$ on koko määrittelyjoukossaan ($= \mathbb{R}$) aidosti kasvava. Määritä $f^{-1}(-1)$, $f^{-1}(3)$, $f^{-1}(10)$, $f^{-1}(12)$ ja $f^{-1}(4\sqrt{2})$ tarkasti, jos mahdollista, muuten likimäärin yhden desimaalin tarkkuudella.

   c) Olkoon $f(x) = (1 - x^n)^{1/n}$, $\mathcal{D}_f = [0, 1]$, $n \in \mathbb{N}$. Näytä, että $f = f^{-1}$.

   d) Olkoon $f(x) = (ax + b)/(cx + d)$. Millä vakioiden a, b, c, d arvoilla on $f = f^{-1}$?

3. Tarkastellaan funktioita $f : A \to R$. Määritä seuraavissa tapauksissa B siten, että $f : A \to B$ on surjektio. Onko $f : A \to B$ tällöin myös bijektio?

   a) $f(x) = -x^5$, $A = (-4, -3]$
   b) $f(x) = x^2$, $A = [\sqrt{2} - \sqrt{3}, \sqrt{2} + \sqrt{3}]$
   c) $f(x) = -6x - x^2$, $A = (-5, -2)$
   d) $f(x) = (x + 5)/(x + 2)$, $A = [-1, 1]$
   e) $f(x) = 3 + 2\sin x$, $A = [3\pi/2, 11\pi/6]$
   f) $f(x) = x + \operatorname{Arcsin} x$, $A = [-1, 1]$

4. Laske tarkka arvo lausekkeelle tai ratkaise yhtälö:

   a) $\sin\left(\operatorname{Arccos}\left(-\frac{3}{5}\right)\right)$ b) $\cos\left(\operatorname{Arcsin}\left(-\frac{3}{5}\right)\right)$

   c) $\sin\left(\operatorname{Arctan}\left(-\sqrt{3}\right)\right)$ d) $\sin\left(\operatorname{Arccot}\left(-\sqrt{3}\right)\right)$

   e) $\sin(\operatorname{Arctan} 2 - \operatorname{Arctan} 3)$ f) $\cos\left(\operatorname{Arccot}\frac{2}{5} + \operatorname{Arctan}\frac{3}{7}\right)$

   g) $\operatorname{Arctan} x = \frac{\pi}{4} - \operatorname{Arctan} 3$ h) $\operatorname{Arccos} x = \operatorname{Arctan} 2 + \operatorname{Arccos}\frac{3}{4}$

5. Saata seuraavat funktiot muotoon $R\sin(x + \alpha)$:

   a) $f(x) = \sin x + \sin\left(x - \operatorname{Arctan}\frac{4}{3}\right)$

   b) $f(x) = \sqrt{5}\sin\left(x + \operatorname{Arctan}\frac{1}{2}\right) + 2\sqrt{2}\sin\left(x + \frac{3\pi}{4}\right) + \sqrt{3}\sin(x - \pi)$

6. Määritä R ja α siten, että

$$2\sin\left(x + \operatorname{Arctan} 2\sqrt{2}\right) + R\sin(x+\alpha) = 6\sin x, \quad x \in \mathbb{R}.$$

7. Sievennä lauseke ja piirrä kuvaaja:

a) $y = \operatorname{Arccos}(\sin x)$ b) $y = \operatorname{Arctan}(\tan x)$

c) $y = \operatorname{Arcsin}\left(2\cos^2 x - 1\right)$ d) $y = \operatorname{Arctan}\dfrac{2\tan x}{1-\tan^2 x}$

8. Näytä, että Esimerkin 8 implisiittifunktio on esitettävissä muodossa $y(x) = u(x) - x$, missä u on jaksollinen. Mikä on perusjakso? Onko $u(x)$ parillinen/pariton? Hahmottele likimäärin funktion $y(x)$ kuvaaja.

9. Seuraavassa on määritelty implisiittifunktioita, mahdollisesti useampihaaraisia. Määritä näiden haarat funktioina $x \mapsto y(x)$ eri $\mathbb{R}$:n osaväleillä. Hahmottele funktiot myös graafisesti!

a) $x^2 + y^2 + 6x = 0$ b) $x^2 - y^2 - 4x = 0$

c) $xy^4 + 3y^2 - x = 0$ d) $|x|^{2/3} + |y|^{2/3} = 1$

e) $x^2 - 4y^2 + y^4 = 0$ f) $xy^4 + 3x^3 + 6y^2 - 12x = 0$

g) $x^2 + \dfrac{\sin y}{3+\sin y} = 0$ h) $\arctan x + \arctan y = \dfrac{\pi}{4}$

10. Mitkä ovat monihaaraisen funktion $\arccos(\cos x)$ arvot, kun $x = 13\pi/4$?

11. a) Määrittele monihaaraisen funktion $\arcsin = \{y_n\,,\ n \in \mathbb{Z}\}$ haarat päähaaran y_0 avulla. b) Ilmaise karteesisen koordinaatiston pisteiden ja napakulman välinen riippuvuus $\varphi(x,y)$ funktioiden y_n avulla.

12. (*) Olkoon $f(x) = x^3 + 3x$ ja $y \in (0,\infty)$. Todista yhtälön $f(x) = y$ ratkeavuus käyttäen kymmenjakoalgoritmia (ks. Luku I.9). *Vihje*: Lähde tuloksista: f on aidosti kasvava välillä $[0,\infty)$ ja $\lim_n f(n) = \infty$.

13. (*) Tutki, millä arvoilla $x \in \mathbb{R}$ tai $(x,y) \in \mathbb{R}^2$ pätee

a) $2\operatorname{Arcsin} x = \operatorname{Arccos}(1-2x^2)$

b) $\operatorname{Arctan} x + \operatorname{Arctan} y = \operatorname{Arctan}\dfrac{x+y}{1-xy}$

14. (*) Tutki, mitkä seuraavista funktioista tyyppiä $f : \mathbb{N} \to \mathbb{Q}$ ovat injektioita: a) $f(n) = n^2 2^{-n}$ b) $f(n) = n^{333} 2^{-n}$ c) $f(n) = n^{243} 3^{-n}$ d) $f(n) = n^{80} 3^{-n}$

IV.3 Kahden ja kolmen muuttujan funktiot

Kahden reaalimuuttujan funktio on tyyppiä

$$f : \mathbb{R}^2 \to \mathbb{R} \quad \text{tai} \quad f : A \to \mathbb{R}, \quad A \subset \mathbb{R}^2.$$

Tällaista funktion *kuvaaja* on Euklidisen avaruuden E^3 pistejoukko

$$G_{f,A} = \{P \mathrel{\widehat{=}} (x, y, z) \in \mathbb{R}^3 \mid z = f(x, y),\ (x, y) \in A\} \subset E^3,$$

jonka yleisnimi on *pinta*. Funktiota voidaan 'katsella' esimerkiksi kehittämällä (yleensä tietokoneen avulla) $G_{f,A}$:n perspektiivikuva. Toinen tapa tehdä kahden muuttujan funktio geometrisesti havainnolliseksi on piirtää funktion *tasa-arvokäyriä* euklidiseen tasoon. Tasa-arvokäyrä on tason pistejoukko

$$S_{f,c} = \{P \mathrel{\widehat{=}} (x, y) \in \mathbb{R}^2 \mid f(x, y) = c\} \subset E^2 \quad (c \in \mathbb{R}).$$

Kartalla tasa-arvokäyriä ovat <u>korkeusviivat</u> (f = maaston korkeusfunktio).[†]

Esimerkki 1 Meren pohjasta kohoavaa saarta kuvaa funktio

$$h(x, y) = -\frac{1}{10}\,(x^2 + 2y^2 + 2xy + 4x + 14y),$$

missä $h(x, y)$ = maaston korkeus merenpinnan tasosta (vedenalaisissa osissa $h(x, y) < 0$). Pituusyksikkö = 100 m. Hahmottele saaren sijainti ja korkeuskäyriä. Missä ja kuinka korkealla on saaren korkein kohta?

<u>Ratkaisu</u> Merkitään $f(x, y) = -10\,h(x, y) = x^2+2y^2+2xy+4x+14y$ ja yritetään ensin kirjoittaa f hieman selkeämpään muotoon

$$f(x, y) = (x - x_0)^2 + 2(y - y_0)^2 + 2(x - x_0)(y - y_0) + c,$$

eli

$$f(x, y) = g(\xi, \eta) = \xi^2 + 2\eta^2 + 2\xi\eta + c, \qquad \begin{cases} \xi = x - x_0, \\ \eta = y - y_0, \end{cases}$$

[†]Geometrian käsite *käyrä* on intuitiivisesti ymmärrettynä 'yhtenäinen viiva vailla leveyttä'. Käsitettä on yllättävän vaikea määritellä täsmällisesti ja samalla yleispätevästi, eikä määrittely edes onnistu pelkin algebran ja geometrian keinoin. Sovellustilanteissa kohdataan tyypillisisesti vain käyrien erikoistapauksia, kuten suoria, toisen asteen käyriä, jne., jolloin käsitteellisiä ongelmia ei yleensä ole. Tässä ja seuraavissa luvuissa käytetään termiä 'käyrä' intuitiivisesti, ajatellen lähinnä ongelmattomia erikoistapauksia. — Termi 'tasa-arvokäyrä' sen sijaan on sovelluksiakin ajatellen tahallisen huolimaton, sillä kyse voi hyvin olla pistejoukosta, jossa on 'läiskiä' tai erillisiä pisteitä, tai joukko voi olla tyhjä. Itse asiassa jos A on mikä tahansa $\mathbb{R}^2$:n osajoukko, niin on helppo määritellä funktio f siten, että $S_{f,c} = A$ annetulla c.

missä $x_0, y_0, c \in \mathbb{R}$ on määrättävä. Muunnos $(x, y) \mapsto (\xi, \eta)$ vastaa siis koordinaatiston origon siirtoa (toistaiseksi tuntemattomaan) pisteeseen (x_0, y_0). Vertaamalla f:n lausekkeita nähdään, että x_0, y_0, c on valittava siten, että pätee

$$\begin{aligned} f(x,y) &= x^2 + 2y^2 + 2xy - (2x_0 + 2y_0)x - (2x_0 + 4y_0)y + (x_0^2 + 2y_0^2 + 2x_0y_0 + c) \\ &= x^2 + 2y^2 + 2xy + 4x + 14y \quad \forall (x,y) \in \mathbb{R}^2, \end{aligned}$$

eli on oltava

$$\begin{cases} 2x_0 + 2y_0 = -4 \\ 2x_0 + 4y_0 = -14 \\ x_0^2 + 2y_0^2 + 2x_0y_0 + c = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_0 = 3 \\ y_0 = -5 \\ c = -29 \end{cases}$$

Siis $\xi = x - 3,\ \eta = y + 5$, ja

$$\begin{aligned} f(x,y) = g(\xi,\eta) &= \xi^2 + 2\eta^2 + 2\xi\eta - 29 \\ &= (\xi + \eta)^2 + \eta^2 - 29. \end{aligned}$$

Viimeksi kirjoitetusta lausekkeesta nähdään, että g:n minimikohta $= h$:n maksimikohta on pisteessä

$$(\xi, \eta) = 0 \ \Leftrightarrow\ (x, y) = (3, -5),$$

ja saaren korkein kohta on 2.9 km meren pinnasta. Saaren rantaviiva on pistejoukko

$$S = \{P \mathrel{\widehat{=}} (x,y) \in \mathbb{R}^2 \mid f(x,y) = 0\}.$$

Hahmottelemalla muitakin korkeusviivoja saadaan saaresta yleiskuva. $\square$

Esimerkki 2 Määritä funktion $f(x,y) = x + 2y + 2$ arvojoukko yksikköneliössä

$$A = \{(x,y) \in \mathbb{R}^2 \mid 0 \le x \le 1,\ 0 \le y \le 1\}.$$

Ratkaisu Tasa-arvokäyrät $S : f(x,y) = c$ ovat suoria

$$x + 2y + 2 = c,$$

joten arvojoukko on (kuva!)

$$f(A) = [2,5]. \quad \square$$

Kolmen muuttujan funktiot

Kolmen reaalimuuttujan funktio on tyyppiä

$$f : \mathbb{R}^3 \to \mathbb{R} \quad \text{tai} \quad f : A \to \mathbb{R}, \quad A \subset \mathbb{R}^3.$$

Kolmen muuttujan funktioita voi havainnollistaa esimerkiksi viipaloimalla: Valitaan äärellinen joukko muuttujan z arvoja z_i ja tutkitaan kahden muuttujan funktioita

$$g_i(x,y) = f(x,y,z_i).$$

Esimerkki 3 : **Tomografia.** Lääketieteessä paljon käytetyllä (tietokone)tomografialla määritetään nk. varjostumafunktio f, jonka määrittelyjoukko on ihmisruumis tai sen osa. Varjostumafunktion arvot ovat reaalilukuja, jotka kertovat varjostuman tummuusasteen. Funktiosta f saadaan mittausten avulla likimäärin selville viipaloitujen funktioiden $g_i(x,y) = f(x,y,z_i)$ arvot valituilla (äärellisen monella) z_i:n arvoilla. Funktiot g_i määrätään yhdistämällä suuri joukko erisuuntaisia röntgen(varjo)kuvia laskennallisin keinoin (tietokoneen avulla).

Olkoon tarkastelun kohteena (idealisoitu) ihmisen pää

$$A = \{P \mathrel{\widehat{=}} (x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \le 7^2\}$$

(pituusyksikkö = cm). Tomografiakuvaus tuottakoon varjostumafunktion

$$f(x, y, z) = 1 + 0.02(2x - 3y + 6z).$$

Missä on kirkkainta ja missä tumminta?

Ratkaisu Funktio f on saatu käytännössä yhdistelemällä joukko viipalefunktioita $g_i(x, y) = f(x, y, z_i)$, missä $-7 < z_i < 7$. Sikäli kuin mittaukset ja niistä lasketut funktiot g_i ja f katsotaan tarkoiksi, on siis oltava

$$g_i(x, y) = 0.02(2x - 3y) + c_i, \quad (x, y) \in A_i\,,$$

missä $c_i = 0.12z_i + 1$, ja A_i on kiekon muotoinen leikkauskuvio

$$A_i = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le r_i^2\}, \quad r_i = \sqrt{7^2 - z_i^2}.$$

Funktion g_i tasa-arvokäyrät ovat suoria

$$2x - 3y = \text{vakio},$$

joten g_i:n maksimi ja minimi löytyvät näiden suorien normaalin $3x + 2y = 0$ ja A_i:n reunaviivan leikkauspisteistä.

Eri viipalekuvia tutkimalla löydetään likimain myös f:n maksimi- ja minimiarvot. Tässä f kuitenkin tunnetaan tarkasti, joten voidaan suoraan tutkia f:n *tasa-arvopintoja*. Nämä ovat tasoja

$$2x - 3y + 6z = \text{vakio},$$

joten voidaan geometrisesti päätellä, että f:n maksimi ja minimi löytyvät origon kautta kulkevan tasa-arvopintojen yhteisen normaalin

$$\begin{cases} x = 2t, \\ y = -3t, \\ z = 6t \end{cases}$$

ja kappaleen reunapinnan

$$x^2 + y^2 + z^2 = 7^2$$

leikkauspisteistä. Nämä pisteet vastaavat t:n arvoja

$$(2t)^2 + (-3t)^2 + (6t)^2 = 7^2 \quad \Leftrightarrow \quad t = \pm 1.$$

Varjostuman maksimi- ja minimiarvoiksi päätellään näin muodoin

$$\begin{aligned} f(2,-3,6) &= 1.98 = f_{\max}\,, \\ f(-2,3,-6) &= 0.02 = f_{\min}\,. \qquad \square \end{aligned}$$

Funktiot käyräviivaisissa koordinaatistoissa

Kahden ja kolmen reaalimuuttujan funktioita tutkittaessa voi olla apua siirtymisestä käyräviivaiseen napa-, lieriö- tai pallokoordinaatistoon silloin kun funktion määrittelyjoukon geometria on sellaiseen muunnokseen sopiva. Esimerkiksi, jos kahden reaalimuuttujan funktio on määritelty pyörähdyssymmetrisessä joukossa $A \subset \mathbb{R}^2$ (voi olla myös $A = \mathbb{R}^2$), voi napakoordinaatteihin siirtyminen auttaa. Siirtyminen tapahtuu muunnoksella (vrt. Luku II.8)

$$\begin{aligned} f(x,y) &= f(r\cos\varphi, r\sin\varphi) \\ &= g(r,\varphi). \end{aligned}$$

Huomattakoon, että muunnettu funktio g on itse asiassa yhdistetty funktio

$$g(r,\varphi) = f(x(r,\varphi), y(r,\varphi)),$$

missä $x(r,\varphi) = r\cos\varphi$, $y(r,\varphi) = r\sin\varphi$. Siirtyminen polaarikoordinaatistosta karteesiseen tapahtuu käänteismuunnoksella

$$f(x,y) = g(r(x,y), \varphi(x,y)),$$

missä $r(x,y) = \sqrt{x^2+y^2}$ ja $\varphi(x,y)$ on pistettä (x,y) vastaava napakulma[†] (φ:n laskukaava on hieman konstikas, ks. edellisen luvun Esimerkki 11).

Kolmen muuttujan funktioita tarkasteltaessa voidaan siirtyä lieriö- tai pallokoordinaatistoon muunnoksilla

$$\begin{aligned} &\text{Lieriö:} \quad f(x,y,z) = f(r\cos\varphi, r\sin\varphi, z) = g(r,\varphi,z). \\ &\text{Pallo:} \quad f(x,y,z) = f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta) = g(r,\theta,\varphi). \end{aligned}$$

Lieriökoordinaatiston tapauksessa muunnos $(x,y) \hookrightarrow (r,\varphi)$, samoin käänteismuunnos $(r,\varphi) \hookrightarrow (x,y)$ on sama kuin napakoordinaatistossa. Pallokoordinaatistossa käänteismuunnos on

$$\begin{aligned} g(r,\theta,\varphi) &= g(r(x,y,z), \theta(x,y,z), \varphi(x,y)) \\ &= f(x,y,z), \end{aligned}$$

missä

$$\begin{aligned} r(x,y,z) &= \sqrt{x^2+y^2+z^2}, \\ \theta(x,y,z) &= \operatorname{Arccos}\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right), \end{aligned}$$

ja $\varphi(x,y)$ on sama kuin napakoordinaatistossa. Suuntakulma θ ei ole määritelty origossa eikä kulma φ z-akselilla.

Esimerkki 4 Vuoristoisen maaston korkeus merenpinnasta on origon ympäristössä

$$f(x,y) = x^2 - 2xy - 3y^2 + 0.2x - 0.1y + 10$$

(yksikkö = 100 m). Origossa on kahden tien risteys. Tiet ovat kartalla suoria ja myös niiden sivuprofiili on suora. Mihin suuntiin tiet kulkevat ja mikä on teiden kaltevuus?

Ratkaisu Siirrytään napakoordinaatistoon:

$$f(x,y) = h(r,\varphi) = r^2(\cos^2\varphi - 2\cos\varphi\sin\varphi - 3\sin^2\varphi) + r(0.2\cos\varphi - 0.1\sin\varphi) + 10.$$

Tiet noudattavat origosta lähdettäessä puolisuoria $\varphi = \varphi_1$ ja $\varphi = \varphi_2$ (koska ovat kartalla suoria). Koska teiden sivuprofiilikin on suora, on oltava

$$\cos^2\varphi - 2\cos\varphi\sin\varphi - 3\sin^2\varphi = 0.$$

[†]Merkinnöissä $x(r,\varphi)$, $y(r,\varphi)$, $r(x,y)$ ja $\varphi(x,y)$ on alunperin riippumattomista muuttujista x,y tai r,φ tehty funktiosymboleja. Koordinaattimuunnoksissa (myös implisiittifunktioissa, vrt. edellinen luku) tällaiset merkinnät ovat tavallisia, koska ne selkeyttävät laskemista.

Tässä ei $\cos\varphi = 0$ ole ratkaisu, joten voidaan jakaa $\cos^2\varphi$:lla:

$$1 - 2\tan\varphi - 3\tan^2\varphi = 0 \quad \Leftrightarrow \quad \begin{cases} \tan\varphi_1 = 1/3 & \Rightarrow \quad \varphi_1 \approx 18^\circ \text{ tai } 198^\circ, \\ \tan\varphi_2 = -1 & \Rightarrow \quad \varphi_2 = 135^\circ \text{ tai } 315^\circ. \end{cases}$$

Kaltevuuskulmat suuntiin $\varphi_1 \approx 18^\circ$ ja $\varphi_2 = 135^\circ$ saadaan ratkaisemalla

$$\tan\alpha_i = 0.2\cos\varphi_i - 0.1\sin\varphi_i \quad \Rightarrow \quad \begin{cases} \alpha_1 \approx 9^\circ, \\ \alpha_2 \approx -12^\circ. \end{cases} \qquad \square$$

Esimerkki 5 Etsi (jos mahdollista) pisteet, joissa funktio

$$f(x,y,z) = xyz(3x^2 + 3y^2 - z^2)/(x^2 + y^2)$$

saavuttaa maksimi- tai minimiarvonsa joukossa

$$A = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \le 96,\ (x,y) \ne (0,0)\}.$$

<u>Ratkaisu</u> Siirrytään pallokoordinaatistoon. Koska pallokoordinaateissa on

$$x^2 + y^2 = r^2\sin^2\theta(\cos^2\varphi + \sin^2\varphi) = r^2\sin^2\theta,$$

niin (vrt. Luku II.5)

$$\begin{aligned} f(x,y,z) \;=\; g(r,\theta,\varphi) &= r^3\cos\theta\,(3\sin^2\theta - \cos^2\theta)\cos\varphi\sin\varphi \\ &= r^3\,(3\cos\theta - 4\cos^3\theta)\cos\varphi\sin\varphi \\ &= -\frac{1}{2}\,r^3\cos 3\theta\sin 2\varphi. \end{aligned}$$

Funktio g saavuttaa maksimiarvonsa $192\sqrt{6}$ pallon pinnalla ($r = \sqrt{96} = 4\sqrt{6}$), kun joko $\cos 3\theta = -1$, $\sin 2\varphi = 1$ tai $\cos 3\theta = 1$, $\sin 2\varphi = -1$. Maksimikohtia on neljä:

$(r,$	$\theta,$	$\varphi)$	$(x,$	$y,$	$z)$
$(4\sqrt{6},$	$\pi/3,$	$\pi/4)$	$(\ 6,$	$6,$	$2\sqrt{6})$
$(4\sqrt{6},$	$\pi/3,$	$5\pi/4)$	$(-6,$	$-6,$	$2\sqrt{6})$
$(4\sqrt{6},$	$2\pi/3,$	$3\pi/4)$	$(-6,$	$6,$	$-2\sqrt{6})$
$(4\sqrt{6},$	$2\pi/3,$	$7\pi/4)$	$(\ 6,$	$-6,$	$-2\sqrt{6})$

Minimikohtia (joissa $g = g_{\min} = -192\sqrt{6}$) on samoin neljä. Nämä saadaan vaihtamalla yhden koordinaatin merkki maksimipisteiden karteesisessa esitysmuodossa. Funktio g saavuttaa maksimi- tai minimiarvonsa myös pallonpintakoordinaattien arvoilla $(\theta,\varphi) \in \{0,\pi\} \times \{\pi/4,\ 3\pi/4,\ 5\pi/4,\ 7\pi/4\}$. Nämä vastaavat kuitenkin karteesisen koordinaatiston pisteitä $(0,0,z)$, jotka ovat alkuperäisen funktion f määrittelyjoukon ulkopuolella. $\square$

Usean muuttujan funktioiden yhdistely

Useamman reaalimuuttujan reaaliarvoisia funktioita voi yhdistellä yhdistetyiksi funktioiksi samoilla periaatteilla kuin yhden muuttujan tapauksessa. Lisäehtona on kuitenkin, että funktiot ovat tyypiltään yhteensopivia. Esimerkiksi funktioiden $f(x)$ ja $g(x,y)$ yhdistetty funktio $F = f \circ g$ voidaan määritellä:

$$F(x,y) = (f \circ g)(x,y) = f(g(x,y)), \quad \mathcal{D}_F = \{(x,y) \in \mathcal{D}_g \mid g(x,y) \in \mathcal{D}_f\}.$$

Voidaan myös määritellä yhdistetyt funktiot

$$\begin{aligned}
&G_1(x) = g(x, f(x)), \quad \mathcal{D}_{G_1} = \{x \in \mathcal{D}_f \mid (x, f(x)) \in \mathcal{D}_g\},\\
&G_2(y) = g(f(y), y), \quad \mathcal{D}_{G_2} = \{y \in \mathcal{D}_f \mid (f(y), y) \in \mathcal{D}_g\},\\
&G_3(x,y) = g(f(x), f(y)), \quad \mathcal{D}_{G_3} = \{(x,y) \in \mathcal{D}_f \times \mathcal{D}_f \mid (f(x), f(y)) \in \mathcal{D}_g\}.
\end{aligned}$$

Funktiota $g \circ f$ sen sijaan ei voi määritellä, ei myöskään funktiota $g \circ g$.

Esimerkki 6 Jos $f(x) = \sqrt{x}$, $\mathcal{D}_f = [0, \infty)$ ja $g(x,y) = x - y^2$, $\mathcal{D}_g = \mathbb{R}^2$, niin

$$\begin{aligned}
&F(x,y) = (f \circ g)(x,y) = \sqrt{x - y^2}, \quad \mathcal{D}_F = \{(x,y) \in \mathbb{R}^2 \mid x \geq y^2\},\\
&G(x) = g(x, f(x)) = 0, \quad \mathcal{D}_G = [0, \infty). \quad \square
\end{aligned}$$

Esimerkki 7 Määritä funktion $f(x,y) = x^2 + 2y^2 + 2xy + 4x + 14y$ minimiarvo suoralla $S:\ y = 2x - 5$. (Vrt. Esimerkki 1.)

<u>Ratkaisu</u> Kun käytetään suoralla S vallitsevalle funktioriippuvuudelle $x \mapsto y$ merkintää $y(x)$, niin kyse on yhdistetystä funktiosta $F(x) = f(x, y(x))$, $\mathcal{D}_F = \mathbb{R}$. Koska

$$\begin{aligned}
F(x) = f(x, 2x-5) &= x^2 + 2(2x-5)^2 + 2x(2x-5) + 4x + 14(2x-5)\\
&= 13x^2 - 18x - 20\\
&= 13\left(x - \frac{9}{13}\right)^2 - \frac{341}{13},
\end{aligned}$$

niin nähdään, että kysytty minimiarvo on

$$F_{\min} = -\frac{341}{13} \approx -26.2.$$

Minimikohta on suoran pisteessä $(9/13, -47/13) \approx (0.69, -3.62)$. $\square$

Useamman muuttujan funktioita voidaan yhdistellä myös laskutoimituksin samalla tavoin kuin yhden muuttujan funktioita (Määritelmä IV.1.2). Ellei yhdisteltävien funktioiden muuttujien lukumäärä täsmää, on ajateltava, että funktioissa on 'näkymättömiä' muuttujia.

Esimerkki 8 Jos $f(x) = \tan x$ ja $g(x,y) = \sin(x+y^2)$, niin kirjoitettaessa

$$F(x,y) = \tan x + \sin(x+y^2)$$

tarkoitetaan tämän laskusäännön ilmaisemaa funktiota F, määrittelyjoukkona

$$\mathcal{D}_F = \{(x,y) \in \mathbb{R}^2 \mid x \neq (n+1/2)\,\pi \ \ \forall n \in \mathbb{Z}\}.$$

Samaan tulokseen tullaan myös funktioiden yhdistelyn kautta ajattelemalla, että $F = \tilde{f} + g$, $\mathcal{D}_F = \mathcal{D}_{\tilde{f}} \cap \mathcal{D}_g$, missä $\tilde{f}$ on saatu f:stä lisäämällä toinen muuttuja:

$$\tilde{f}(x,y) = \tan x, \quad \mathcal{D}_{\tilde{f}} = \mathcal{D}_f \times \mathbb{R}. \quad \square$$

Kahden muuttujan implisiittifunktiot

Jos F on kolmen muuttujan funktio ja yhtälö

$$F(x,y,z) = 0$$

ratkeaa z:n suhteen, kun $(x,y) \in A \subset \mathbb{R}^2$, niin yhtälö määrittelee A:ssa kahden muuttujan (mahdollisesti monihaaraisen) implisiittifunktion. Jos tälle käytetään funktiomerkintää $z(x,y)$, niin määritelmä on siis

$$F(x,y,z(x,y)) = 0, \quad (x,y) \in A.$$

Esimerkki 9 Yhtälö $x^2+y^2+z^2 = R^2$ määrittelee kaksihaaraisen implisiittifunktion $z(x,y) = \pm\sqrt{R^2-x^2-y^2}$ kiekossa $A:\ x^2+y^2 \leq R^2$. $\square$

Esimerkki 10 Jos $m \in \mathbb{N}$, $m \geq 2$, niin yhtälö

$$y, z \in \mathbb{C}: \quad y^m = z$$

määrittelee m-haaraisen kompleksifunktion $y = \sqrt[m]{z}$, vrt. Luku III.3. Jos kirjoitetaan $z = x+iy$ ja $y = u(x,y)+iv(x,y)$, niin u ja v ovat m-haaraisia funktioita tyyppiä $u,v:\ \mathbb{R}^2 \to \mathbb{R}$. $\square$

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien funktioiden arvojoukot annetulla janalla AB:
   a) $f(x,y) = x^2+2xy-y^2$, $A=(1,2)$, $B=(-1,3)$
   b) $f(x,y,z) = x+xy+yz+z^2$, $A=(1,1,1)$, $B=(-1,2,-3)$

2. Määritä seuraavien funktioiden arvojoukot annetussa joukossa A:
   a) $f(x,y) = x + 3y - 1$, $A =$ kolmio, jonka kärjet $(0,1)$, $(4,0)$ ja $(3,4)$
   b) $f(x,y,z) = x+2y-3z$, $A = \{(x,y,z) \mid (x-1)^2+(y+1)^2+(z-2)^2 \leq 9\}$
   c) $f(x,y,z) = x + xy + yz + z^2$, $A =$ suora $2x - 2 = y + 1 = 4 - 2z$

3. Tasangolla $z = 0$ on järvi $A = \{(x,y) \in \mathbb{R}^2 \mid h(x,y) < 0\}$, missä

$$h(x,y) = 4x^2 + y^2 + 24x + 8y$$

   on järven pohjan korkeusprofiili. Tässä h on ilmaistu metreinä ja x, y kilometreinä. Järven poikki kulkee moottoritie suoraa $y = x + 2$ pitkin. Missä on järven syvin kohta ja mikä on syvyys tässä kohdassa? Hahmottele järven rantaviiva ja laske, missä pisteissä moottoritie leikkaa rantaviivan.

4. Muunna seuraavat funktiot napa- tai pallokoordinaatistoon:

   a) $f(x,y) = x + 2y$ b) $f(x,y) = xy^2$ c) $f(x,y) = \dfrac{xy}{x^2+y^2}$

   d) $f(x,y) = \max\{0, x, y\}$ e) $f(x,y) = \begin{cases} x^2/y, & \text{kun } x > 0 \text{ ja } y > 0 \\ 0, & \text{muulloin} \end{cases}$

   f) $f(x,y) = \begin{cases} xy^2 - y^3, & \text{kun } 0 \leq x \leq y \\ 0, & \text{muulloin} \end{cases}$ g) $f(x,y,z) = \dfrac{xyz}{x^2+y^2+z^2}$

   h) $f(x,y,z) = \dfrac{xy^2z^3}{x^2+y^2}$ i) $f(x,y,z) = \begin{cases} z^2 - xyz, & \text{kun } z > 0 \\ 0, & \text{kun } z \leq 0 \end{cases}$

5. Muunna seuraavat napakoordinaateissa ilmaistut funktiot karteesiseen koordinaatistoon:

   a) $g(r,\varphi) = r^3(\cos\varphi - \sin\varphi)$ b) $g(r,\varphi) = \sin 2\varphi - r^2\cos 2\varphi$

   c) $g(r,\varphi) = r^2(\tan\varphi + \cot\varphi)$ c) $g(r,\varphi) = \dfrac{\sin\varphi}{2+\cos\varphi}$

6. Olkoon $f(x,y) = x + y$ ja $g(x,y) = xy$, $(x,y) \in \mathbb{R}^2$. Määrittele sievennettyinä lausekkeina (laskusääntöinä) seuraavat yhdistetyt funktiot:
   a) $f(x, g(x,y))$ b) $f(g(x,y), y)$ c) $f(g(x,y), g(x,y))$
   d) $g(x, f(x,y))$ e) $g(f(x,y), y)$ f) $g(f(x,y), f(x,y))$
   g) $f(f(x,y), f(x,y))$ h) $g(g(x,y), g(x,y))$

7. Olkoon $f(x) = \sqrt{2-x}$ ja $g(x,y) = \sqrt{x-y^2}$ $(x, y \in \mathbb{R})$. Määrittele yhdistetty funktio $F = f \circ g$ (laskusääntö ja määrittelyjoukko).

8. Olkoon $f(x) = \sqrt{25 - x^2}$ ja $g(x, y) = 3x + 4y - 1$ $(x, y \in \mathbb{R})$. Määritä yhdistetyn funktion $F = f \circ g$ pienin ja suurin arvo yksikkökiekon neljänneksessä $A = \{\, (x, y) \in \mathbb{R}^2 \mid x \geq 0 \ \&\ y \geq 0 \ \&\ x^2 + y^2 \leq 1 \,\}$.

9. a) Esitä kompleksifunktion $f(z) = (z+1)^2 + (z+i)^2$ reaaliosa, imaginaariosa ja itseisarvo funktioina tyyppiä $g : \mathbb{R}^2 \to \mathbb{R}$.

   b) Määrittele kaksihaaraiset funktiot $u = \{u_1, u_2\}$ ja $v = \{v_1, v_2\}$ siten, että $u(x, y) + iv(x, y) = \sqrt{z} \ \ \forall z = x + iy \in \mathbb{C}$.

10. Yhtälö $x + 3xyz^2 + z^4 = 0$ $(x, y, z \in \mathbb{R})$ määrittelee implisiittifunktion $z = f(x, y)$. a) Laske f:n arvot pisteissä $(0, 0)$, $(2, -1)$ ja $(1, 1)$. b) Millaisia ovat f:n haarat (määrittelyjoukot ja laskusäännöt) yleisemmin?

11. (*) Vuoristoisen maaston korkeus merenpinnasta on origon O lähellä funktio

$$f(x, y) = \frac{1}{20}(3x^2 - 5xy - 2y^2) + 2.$$

   Origossa on kahden tien S_1, S_2 risteys. Tiet ovat maastoon sovitettuja ja vaakasuoria, ja lisäksi ne ovat kartallakin suoria. Teiden S_1, S_2 poikki kulkee suora rautatie pisteissä $A \in S_1$, $B \in S_2$. Molemmat pisteet ovat O:sta etäisyydellä 1 (yksikkö = km). Mikä on suurin pudotuskorkeus laaksoon rautatiesillalta, jonka päät ovat pisteissä A, B?

12. (*) Halutaan selvittää, missä pisteissä funktio $f(x, y) = x^2 - 2xy + 3y^2$ saavuttaa suurimman ja pienimmän arvonsa a) ympyrällä $S : x^2 + y^2 = 9$, b) kiekossa $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 9\}$. Ratkaise ongelma napakoordinaatteja käyttäen (vrt. Harj.teht. II.5:20).

13. (*) Funktiosta $f(x) = x - x^3$ tiedetään, että välillä $[0, 1]$ f saavuttaa suurimman arvonsa pisteessä $x = 1/\sqrt{3}$. Mihin suuntaan origosta lähdettäessä funktio a) $f(x, y) = xy^2$, b) $f(x, y, z) = xy^2z^3$ kasvaa nopeimmin suhteessa kuljettuun matkaan?

14. (*) (Funktio avaimenreiässä) Olkoon

$$A = \{\, (x, y, z) \in \mathbb{R}^3 \mid (x, y) \in B_1 \cup B_2, \ \ z \in [0, 10] \,\},$$

   missä

$$B_1 = \{\, (x, y) \in \mathbb{R}^2 \mid x^2 + (y - 1)^2 \leq 2 \ \&\ y \geq 0 \,\},$$
$$B_2 = \{\, (x, y) \in \mathbb{R}^2 \mid x \in [-1, 1] \ \&\ y \in [-4, 0] \,\}.$$

   Missä A:n pisteissä funktio $f(x, y, z) = x - 3y + 2z$ saavuttaa suurimman ja missä pienimmän arvonsa?

IV.4 Parametriset käyrät ja pinnat

Tason tai avaruuden *parametriseksi käyräksi* sanotaan funktiota muotoa

$$t \in A \;\mapsto\; P(t) \in E^d,$$

missä $A \subset \mathbb{R}$ on väli (usein suljettu väli) ja $d = 2$ tai $d = 3$. Muuttujaa t sanotaan tässä *parametriksi*. Käyrä on *tasokäyrä* jos $d = 2$, *avaruuskäyrä* jos $d = 3$. Käyttäen jo tutuksi tulleita geometrisia vastaavuuksia voidaan kirjoittaa

$$P(t) \;\widehat{=}\; \overrightarrow{OP}(t) \;=\; \begin{cases} x(t)\vec{i} + y(t)\vec{j} & \widehat{=} \quad (x(t), y(t)), \qquad (d = 2) \\ x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k} & \widehat{=} \quad (x(t), y(t), z(t)), \qquad (d = 3) \end{cases}$$

missä x, y, z ovat funktioita tyyppiä $f : A \to \mathbb{R}$. Tämän mukaisesti parametrinen käyrä voidaan tulkita reaalimuuttujan *vektoriarvoiseksi* funktioksi, jolloin luonteva esitystapa on myös vektorimerkintä

$$\vec{r}(t) \;=\; \begin{cases} x(t)\vec{i} + y(t)\vec{j}, \quad t \in A, & \text{(tasokäyrä)} \\ x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}, \quad t \in A. & \text{(avaruuskäyrä)} \end{cases}$$

Liittyen vastaavuuteen $E^d \leftrightarrow \mathbb{R}^d$ voidaan parametrinen käyrä esittää yhtä hyvin yhtälöryhmänä, esim. tasokäyrän tapauksessa

$$(x, y) \;=\; (x(t), y(t)) \quad \Leftrightarrow \quad \begin{cases} x = x(t), \\ y = y(t). \end{cases} \qquad \text{(tasokäyrä)}^\dagger$$

Esimerkki 1 Reaalimuuttujan funktio $f : [a, b] \to \mathbb{R}$ voidaan tulkita parametriseksi tasokäyräksi

$$\vec{r} = \vec{r}(t) = t\vec{i} + f(t)\vec{j} \quad \Leftrightarrow \quad \begin{cases} x = x(t) = t, \\ y = y(t) = f(t), \quad t \in [a, b]. \end{cases} \qquad \square$$

Esimerkki 2 Luvuista II.7 ja II.8 tuttuja parametrisia avaruuskäyriä ovat

avaruussuora: $\vec{r}(t) \;=\; (x_0 + \alpha t)\vec{i} + (y_0 + \beta t)\vec{j} + (z_0 + \gamma t)\vec{k} \;=\; \vec{r}_0 + t\vec{v}, \quad t \in \mathbb{R},$

ruuviviiva: $\begin{cases} x = x(\varphi) = R\cos\varphi, \\ y = y(\varphi) = R\sin\varphi, \\ z = z(\varphi) = a\varphi, \quad \varphi \in \mathbb{R}. \quad \square \end{cases}$

†Merkinnöissä $x = x(t)$ ja $y = y(t)$ on symboleja x, y käytetty kahdessa eri merkityksessä: oikealla funktion, vasemmalla ko. funktion arvojoukon alkion symbolina. Tämän tyyppisiä epäloogisuuksia pidetään matematiikan käytännössä siedettävinä, syystä että ne yksinkertaistavat merkintöjä.

Parametrisen käyrän euklidiseen avaruuteen jättämä 'geometrinen jälki' on arvojoukko $S = \{P(t) \mathrel{\widehat{=}} \vec{r}(t) \mid t \in A\} \subset E^d$, johon voidaan viitata sellaisilla termeillä kuin *käyrä* (engl. curve) tai (käyrän) *kaari* (engl. arc). Yksinkertaisimmillaan S on jostakin pisteestä A alkava ja toiseen pisteeseen B päättyvä, itseään leikkaamaton ja yhtenäinen viiva, eli nk. *yksinkertainen kaari* (engl. simple arc). Tällaisia ovat esim. jana tai ympyrän kaari. Viiva voi myös olla päättymätön, kuten suora, tai umpinainen *suljettu käyrä* (engl. closed curve), kuten tason tai avaruuden ympyräviiva. Joukko S voi myös olla itseään leikkaava, eli siinä voi olla 'silmukoita'.† Jos lähtökohdaksi otetaan vain tällainen 'viiva' S, eli pelkästään geometrinen objekti, niin funktiota $\vec{r}(t),\ t \in A$, jonka arvojoukko $= S$, sanotaan S:n *parametriesitykseksi* eli *parametrisoinniksi* (parametrisaatioksi). Parametrisoivan funktion $t \in A \mapsto \vec{r}(t) \in S$ ei tarvitse olla injektiivinen, ts. samaan pisteeseen $P \in S$ voidaan päätyä monella (jopa äärettömän monella) parametrin arvolla.

Esimerkki 3 Tason ympyräviivan

$$S = \{P \mathrel{\widehat{=}} (x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = R^2\}$$

luonteva parametrisointi on

$$\begin{cases} x = x_0 + R\cos t, \\ y = y_0 + R\,\sin t, \quad t \in [0, 2\pi). \end{cases}$$

Tässä voi välin $[0, 2\pi)$ tilalla olla myös esim. $A = (-\pi, \pi]$, $A = [0, 2\pi]$ tai $A = \mathbb{R}$. Kahdessa jälkimmäisessä vaihtoehdossa parametrisointi ei ole injektio. □

Esimerkki 4 Jos S on avaruussuora, niin tämän tavanomaisin parametrisaatio on $\vec{r}(t) = \vec{r}_0 + t\vec{v},\ t \in \mathbb{R}$, missä $P_0 \mathrel{\widehat{=}} \vec{r}_0$ on suoran piste ja $\vec{v} \neq \vec{0}$ suoran suuntavektori (vrt. Luku II.7). Tällaisiakin parametrisointeja on jo äärettömän monta, mutta mahdollisuudet eivät lopu tähän: Jos $\vec{r}_0$ ja $\vec{v}$ täyttävät mainitut ehdot, niin parametrisoinniksi voidaan yleisemmin valita

$$\vec{r}(t) \;=\; \vec{r}_0 + f(t)\,\vec{v}, \quad t \in A,$$

missä funktion $f : A \to \mathbb{R}$ valintaa rajoittaa vain ehto $\mathcal{R}_f = \mathbb{R}$. Esimerkiksi voidaan valita $f(t) = \tan t,\ A = (-\pi/2, \pi/2)$ (f injektio), tai $f(t) = t\sin t,\ A = \mathbb{R}$ (f ei injektio). □

†Muodossa $S = \{P(t) \mathrel{\widehat{=}} \vec{r}(t) \mid t \in A\}$ määriteltyjen taso- tai avaruuskäyrien geometrista luokittelua ei todellisuudessa voi täsmentää ilman funktion $\vec{r}$ koordinaattifunktioile x, y, z asetettavia lisäehtoja. Vrt. käyriä koskeva alaviite edellisessä luvussa.

Esimerkki 4 (jatko) Jos halutaan parametrisoida suoralla S oleva jana, jonka päätepisteet ovat $A \mathrel{\widehat{=}} \vec{r}_1$ ja $B \mathrel{\widehat{=}} \vec{r}_2$, niin tämä käy muodossa

$$\vec{r}(t) = f(t)\,\vec{r}_1 + [1 - f(t)]\,\vec{r}_2, \quad t \in A,$$

missä f on funktio tyyppiä $f : A \to \mathbb{R},\ A \subset \mathbb{R}$, ja $\mathcal{R}_f = [0,1]$. Yksinkertaisin parametrisointi saadaan, kun valitaan $A = [0,1]$ ja $f(t) = t$. $\square$

Jos tasokäyrän yhtälöistä $x = x(t),\ y = y(t),\ t \in A$ pystytään eliminoimaan parametri t, on tuloksena yhtälö muotoa

$$F(x,y) = 0,$$

missä siis F on kahden muuttujan funktio, jolle pätee $F(x(t), y(t)) = 0\ \ \forall t \in A$. Jos $S = \{P \in E^2 \mid P \mathrel{\widehat{=}} (x(t), y(t))$ jollakin $t \in A\}$, niin sanotaan tällöin, että ym. yhtälö on *käyrän* S (tai pistejoukon S) *yhtälö*. Avaruuskäyrän tapauksessa johtaa parametrin eliminointi (onnistuessaan) yhtälöryhmään muotoa

$$\begin{cases} F_1(x,y,z) = 0, \\ F_2(x,y,z) = 0. \end{cases}$$

Esimerkki 5 Jos $a, b > 0$, niin parametrinen tasokäyrä

$$S : \begin{cases} x = a\cos t, \\ y = b\sin t, \quad t \in [0, 2\pi] \end{cases}$$

on nimeltään *ellipsi* (tapauksessa $a = b$ ympyrä). Eliminoimalla t saadaan S:n yhtälöksi

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad \square$$

Liikerata

Tyypillisessä parametrisen käyrän fysikaalisessa sovellustilanteessa parametri t on <u>aika</u>muuttuja, A on tarkasteltava <u>aikaväli</u>, ja $P(t) \mathrel{\widehat{=}} \vec{r}(t)$ on liikkuvan pisteen (esim. pistemäiseksi ajatellun partikkelin tai liikkuvan kiinteän kappaleen pisteen) p<u>aikka</u> hetkellä t. Tällöin funktion $P(t),\ t \in A$, arvojoukko S on ko. pisteen <u>liikerata</u> aikavälillä A. Funktio $t \mapsto \vec{r}(t),\ t \in A$ on S:n parametrisointi, joka kertoo koko <u>liikehistorian</u>.

Esimerkki 6 : **Heittoparaabeli**. Kivi heitetään tornista korkeudelta h alkuvauhdilla v_0 ja kulmassa α vaakasuuntaan nähden. Millainen on lentorata, jos ilmanvastusta ei huomioida?

Ratkaisu Tarkastellaan liikettä (avaruustason) koordinaatistossa, jossa x mittaa vaakasuoraa etäisyyttä lähtöpisteestä ja y korkeutta maan pinnan tasosta. Liikelakien mukaan kiven paikka $P(t) = (x(t), y(t))$ on lentoajan t funktiona parametrinen käyrä

$$\begin{cases} x(t) = v_0 t \cos\alpha, \\ y(t) = h + v_0 t \sin\alpha - \frac{1}{2} g t^2, \end{cases}$$

missä g = maan vetovoiman kiihtyvyys. Eliminoimalla t ja huomioimalla, että $1/\cos^2\alpha = 1 + \tan^2\alpha$ saadaan lentoradan yhtälö muotoon

$$y = h + kx - (1 + k^2)\frac{x^2}{2a},$$

missä $k = \tan\alpha$ ja $a = v_0^2/g$. Lentorata on *paraabelin* kaari. Kuvan tapauksessa $\alpha = 0$ kivi törmää maahan hetkellä $t = \sqrt{2h/g}$. $\square$

Esimerkki 7 : **Sykloidi**. R-säteinen pyörä vierii liukumatta pitkin tasoa siten, että pyörän keskipisteen liikenopeus on $v_0\vec{\imath}$, v_0 = vakio. Määritä pyörän ulkokehän pisteen P paikka ajan t funktioina.

Ratkaisu Oletetaan, että pyörä vierii pitkin x-akselia ja että P on origossa, kun $t = 0$. Tällöin ratakäyrän parametriesitys on

$$\begin{cases} x(t) = v_0 t - R\sin\varphi(t), \\ y(t) = R - R\cos\varphi(t), \end{cases}$$

missä $\varphi(t)$ on vierimiskulma.

Koska liukumista ei tapahdu, on oltava $R\varphi(t) = v_0 t$, joten P:n paikkavektori hetkellä t on

$$\vec{r}(t) = \left[v_0 t - R\sin(\frac{v_0 t}{R})\right]\vec{i} + R\left[1 - \cos(\frac{v_0 t}{R})\right]\vec{j}.$$

Jos parametriksi otetaan vierimiskulma φ, niin liikeradan parametriesitys on

$$\begin{cases} x = x(\varphi) = R(\varphi - \sin\varphi), \\ y = y(\varphi) = R(1 - \cos\varphi). \end{cases}$$

Tätä sanotaan *sykloidiksi.* $\square$

Esimerkki 8 Pistemäinen partikkeli on hetkellä $t = 0$ (t:n yksikkö s) pisteessä $(1, 1, 1)$ (yksikkö m) ja liikkuu suoraviivaisesti vakionopeudella (vauhdilla) $v = 10$ (yksikkö m/s) siten, että eräällä ajan hetkellä partikkeli on pisteessä $(2, -1, 0)$. Määritä partikkelin sijainti $(x(t), y(t), z(t))$, kun $t \geq 0$.

Ratkaisu Partikkeli liikkuu suoralla, jonka suuntavektori on $(2\vec{i}-\vec{j})-(\vec{i}+\vec{j}+\vec{k}) = \vec{i} - 2\vec{j} - \vec{k}$. Liikesuuntaan osoittava yksikkövektori on siis

$$\vec{e} = \frac{1}{\sqrt{6}}(\vec{i} - 2\vec{j} - \vec{k}),$$

ja partikkelin paikkavektori hetkellä $t \geq 0$ näin ollen

$$\vec{r}(t) = \vec{i} + \vec{j} + \vec{k} + (vt)\,\vec{e} \quad \Leftrightarrow \quad \begin{cases} x(t) = 1 + \dfrac{10t}{\sqrt{6}}, \\ y(t) = 1 - \dfrac{20t}{\sqrt{6}}, \\ z(t) = 1 - \dfrac{10t}{\sqrt{6}}. \end{cases} \qquad \square$$

Parametriset pinnat

Euklidisen avaruuden E^3 *parametriseksi pinnaksi* sanotaan kuvausta (funktiota) tyyppiä

$$(u, v) \in A \mapsto P(u, v) \in E^3,$$

missä $A \subset \mathbb{R}^2$ ja muuttujia u, v sanotaan parametreiksi. Liittyen vastaavuuksiin $P \in E^3 \leftrightarrow \vec{r} \in V \leftrightarrow (x, y, z) \in \mathbb{R}^3$ ($V = \{\text{avaruuden vektorit}\}$) voidaan kuvauksen maalijoukoksi yhtä hyvin ajatella V tai R^3. Kuvauksesta voidaan tällöin käyttää joko vektorimerkintää

$$\vec{r} = \vec{r}(u, v), \quad (u, v) \in A,$$

tai vastaavaa koordinaattimuotoista esitystä

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \\ z = z(u, v), \quad (u, v) \in A. \end{cases}$$

Funktion $(u, v) \in A \mapsto P(u, v) \in E^3$ arvojoukko $S \subset E^3$ on *pinta* (engl. surface) geometrisena oliona.† Itse funktio on tällöin S:n eräs *parametrisointi.* Jos lähtökohtana on pinta S, niin parametrisointi pyritään usein valitsemaan siten, että lähtöjoukko A on geometrialtaan mahdollisimman yksinkertainen, esim. suorakulmio. *Pinnan S yhtälöksi* sanotaan yhtälöä muotoa

$$F(x, y, z) = 0,$$

joka toteutuu jokaisella $(x, y, z) \mathrel{\widehat{=}} P \in S$. Yhtälöön päädytään, jos parametrit u, v pystytään eliminoidaan ym. koordinaattimuotoisesta esityksestä. Jos alunperin tunnetaan kolmen reaalimuuttujan funktio F, niin sanotaan yleisemmin, että yhtälö $F(x, y, z) = c$ ($c \in \mathbb{R}$) määrittelee F:n *tasa-arvopinnan* (sikäli kuin kyseessä on pinta, ks. alaviite).

Esimerkki 9 'Kaikkien pintojen äiti' on *taso*, jonka yleinen parametriesitys on muotoa (vrt. Luku II.7)

$$\begin{cases} x(u, v) = x_0 + \alpha_1 u + \alpha_2 v, \\ y(u, v) = y_0 + \beta_1 u + \beta_2 v, \\ z(u, v) = z_0 + \gamma_1 u + \gamma_2 v. \end{cases}$$

†Pintojen täsmällisemmässä määrittelyssä on samat ongelmat kuin käyrien, vrt. alaviitteet edellä. Tässä yhteydessä ei mihinkään täsmennysyrityksiin ryhdytä, vaan nojaudutaan geometriseen intuitioon.

Näin määritellen taso kulkee pisteen $\vec{r}_0 \mathrel{\widehat{=}} (x_0, y_0, z_0)$ kautta ja sen suuntavektorit ovat $\vec{v}_1 = \alpha_1 \vec{i} + \beta_1 \vec{j} + \gamma_1 \vec{k}$ ja $\vec{v}_2 = \alpha_2 \vec{i} + \beta_2 \vec{j} + \gamma_2 \vec{k}$. Eliminoimalla parametrit (olettaen $\vec{v}_1$ ja $\vec{v}_2$ lineaarisesti riippumattomiksi) saadaan tasolle johdetuksi yhtälö muotoa $F(x, y, z) = ax + by + cz + d = 0$ (vrt. Luku II.7). $\square$

Esimerkki 10 Jos $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}^2$ on kahden reaalimuuttujan funktio, niin f:n *kuvaaja* joukossa $A \subset \mathcal{D}_f$ on pinta, jonka yhtälö on

$$z = f(x, y), \quad (x, y) \in A.$$

Pinnan luonnollinen parametrisointi on tässä tapauksessa

$$x = u, \quad y = v, \quad z = f(u, v), \quad (u, v) \in A. \quad \square$$

Esimerkki 11 Avaruuden yleisen pallopinnan yhtälö on

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2.$$

Luontevin parametrisointi perustuu pallonpintakoordinaatteihin:

$$\begin{cases} x = x(\theta, \varphi) = x_0 + R \sin\theta \cos\varphi, \\ y = y(\theta, \varphi) = y_0 + R \sin\theta \sin\varphi, \\ z = z(\theta, \varphi) = z_0 + R \cos\theta, \quad (\theta, \varphi) \in [0, \pi] \times [0, 2\pi]. \end{cases}$$

Pallokoordinaatistossa, jonka origo on pisteessä (x_0, y_0, z_0) on pinnan yhtälö kaikkein yksinkertaisin: $r = R$. $\square$

Esimerkki 12 Jos $a, b, c > 0$, niin yhtälö

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

määrittelee pinnan nimeltä *ellipsoidi*. Pallonpintakoordinaatteihin perustuva parametrisointi on

$$\begin{cases} x = a \sin\theta \cos\varphi, \\ y = b \sin\theta \sin\varphi, \\ z = c \cos\theta, \quad (\theta, \varphi) \in [0, \pi] \times [0, 2\pi]. \quad \square \end{cases}$$

Pyörähdyspinnat

Pyörähdyspinta syntyy, kun tasokäyrä pyörähtää tasossa olevan suoran ympäri. Olkoon käyrä annettu muodossa

$$K = \{(x, y) \in \mathbb{R}^2 \mid y = f(x) \ \ \& \ \ x \in [a, b]\},$$

missä $f(x) \geq 0 \ \forall x \in [a, b]$. Tällöin käyrän pyörähtäessä x-akselin ympäri syntyy pinta S, jonka luonnolliset parametrit ovat $u = x$ ja $v = \varphi =$ pyörähdyskulma, jolloin pinnan parametrisoinniksi tulee

$$\begin{cases} x = u, \\ y = f(u) \cos\varphi, \\ z = f(u) \sin\varphi, \end{cases}$$

missä

$$(u, \varphi) \in A = [a, b] \times [0, 2\pi].$$

Eliminoimalla parametrit u, φ saadaan *pyörähdyspinnan yhtälö*

$$\boxed{y^2 + z^2 = [f(x)]^2, \quad x \in [a, b].}$$

Esimerkki 13 Parametrisoi pyörähdyspinta

$$S: \quad x^2 + y^2 = z, \quad z \geq 0.$$

<u>Ratkaisu</u> Pinta S syntyy kun yz-tason käyrä $K = \{(y, z) \in \mathbb{R}^2 \mid z = y^2, \ y \geq 0\}$ pyörähtää z-akselin ympäri. Luonteva parametrisointi saadaan lieriökoordinaattien avulla:

$$\begin{cases} x = r\sin\varphi, \\ y = r\cos\varphi, \\ z = r^2, \end{cases}$$

missä

$$(r,\varphi) \in A = [0,\infty) \times [0,2\pi].$$

Pinta on (pyörähdys)*paraboloidi.* $\square$

Esimerkki on erikoistapaus yleisemmästä pyörähdyspinnasta, joka syntyy, kun yz-tason käyrä $K: \ z = f(y), \ y \in B \subset [0,\infty)$, pyörähtää z-akselin ympäri. Lieriökoordinaatteihin perustuva pinnan (luontevin) parametrisointi on

$$\begin{cases} x = r\cos\varphi, \\ y = r\sin\varphi, \\ z = f(r), \quad (r,\varphi) \in A = B \times [0,2\pi]. \end{cases}$$

Näistä yhtälöistä viimeinen on itse asiassa pinnan yhtälö lierökoordinaateissa (!).

Viivoitinpinnat

Viivoitinpinta syntyy, kun suora tai jana liikkuu avaruudessa siten, että suoran/janan piste $P_0 \mathrel{\widehat{=}} \vec{r}_0$ ja suuntavektori $\vec{t}$ ovat yhdestä parametrista (u) riippuvia. Pinnan luonnollinen parametrisaatio on tällöin

$$\begin{aligned}\vec{r}\,(u,v) &= \vec{r}_0(u) + v\,\vec{t}\,(u)\\ &= x(u,v)\vec{i} + y(u,v)\vec{j} + z(u,v)\vec{k}.\end{aligned}$$

Esimerkki 14 : **Jäähdytystorni**. Jana, jonka päätepisteet ovat $A = (2,0,0)$ ja $B = (0,1,3)$ pyörähtää z-akselin ympäri. Millainen parametrisoitu pinta syntyy? Kyseessä on myös pyörähdyspinta — millainen?

<u>Ratkaisu</u> Kulman u verran (kuvio) pyörähtänyt suuntajana on

$$\begin{aligned}\vec{t}\,(u) &= \overrightarrow{A'B'}\\ &= (-\sin u\,\vec{i} + \cos u\,\vec{j} + 3\vec{k}) - (2\cos u\,\vec{i} + 2\sin u\,\vec{j})\\ &= -(2\cos u + \sin u)\vec{i} + (\cos u - 2\sin u)\vec{j} + 3\vec{k}.\end{aligned}$$

Pinnalle saadaan näin ollen parametrisointi

$$\begin{aligned}\vec{r} = \vec{r}\,(u,v) &= \vec{r}_0(u) + v\vec{t}\,(u)\\ &= 2\cos u\,\vec{i} + 2\sin u\,\vec{j} + v\vec{t}\,(u)\\ &= \ [(2-2v)\cos u - v\sin u]\vec{i} + [v\cos u + (2-2v)\sin u]\vec{j} + 3v\vec{k}\end{aligned}$$

$$\Leftrightarrow \begin{cases} x = (2-2v)\cos u - v\sin u,\\ y = v\cos u + (2-2v)\sin u,\\ z = 3v,\end{cases} \qquad (u,v) \in [0,2\pi]\times[0,1].$$

Parametriesityksestä nähdään, että

$$\begin{aligned}[x(u,v)]^2 + [y(u,v)]^2 &= (2-2v)^2 + v^2\\ &= 5v^2 - 8v + 4.\end{aligned}$$

Koska tässä $v = z(u,v)/3$, niin nähdään, että pinta voidaan esittää lieriökoordinaatistossa muodossa

$$\begin{aligned} r^2 &= \frac{5}{9}\,z^2 - \frac{8}{3}\,z + 4 \\ &= \frac{5}{9}\left(z - \frac{12}{5}\right)^2 + \frac{4}{5}\,. \end{aligned}$$

Tämä on pyörähdyspinta, joka syntyy, kun yz-tason käyrä

$$K: \quad y^2 - \frac{5}{9}\left(z - \frac{12}{5}\right)^2 = \frac{4}{5}\,, \quad z \in [0,3]$$

pyörähtää z-akselin ympäri. Käyrä K on *hyperbelin* kaari, ja pyörähdyspinta on *yksivaippaisen hyperboloidin* osa.

Pinnan kapein kohta on korkeudella $z = 12/5$. $\square$

HARJOITUSTEHTÄVIÄ

1. Hahmottele seuraavien parametristen tasokäyrien kulku. Eliminoimalla parametri johda myös käyrän yhtälö karteesisessa koordinaatistossa.

   a) $x = 2 - t,\ y = t + 1,\ t \in \mathbb{R}$ b) $x = t^2,\ y = 2 - t,\ t \in [0, \infty)$

   c) $x = \dfrac{1}{t},\ y = t - 1,\ t \in (0, 4)$ d) $x = \dfrac{1}{1+t^2},\ y = \dfrac{t}{1+t^2},\ t \in \mathbb{R}$

   e) $\vec{r} = 3\sin \pi t\,\vec{i} + 4\cos \pi t\,\vec{j},\ t \in [-1, 1]$

   f) $x = 1 - \sqrt{4 - t^2},\ y = 2 + t,\ t \in [-2, 2]$

   g) $\vec{r} = t\cos t\,\vec{i} + t\sin t\,\vec{j},\ t \in [0, 4\pi]$

2. a) Tasokäyrän eräs parametrisointi on $\vec{r} = \cos 2t\,\vec{i} + \sin^2 t\,\vec{j},\ t \in \mathbb{R}$. Anna käyrälle vaihtoehtoinen, mahdollisimman yksinkertainen parametrisointi.

   b) Totea, että $\vec{r} = (t-1)\vec{i} + \sqrt{2t - t^2}\,\vec{j},\ t \in [0, 2]$ ja $\vec{r} = t\sqrt{2 - t^2}\,\vec{i} + (1-t^2)\vec{j}$, $t \in [-1, 1]$ ovat saman käyrän parametrisointeja. Mikä käyrä on kyseessä?

3. Tasokäyrä $S : x^3 + y^3 = 3xy$ on nimeltään *Cartesiuksen lehti.* Johda käyrälle parametriesitys kirjoittamalla $y = tx$. Hahmottele käyrän kulku parametrimuodosta ja merkitse kuvioon, mitkä käyrän osat vastaavat parametrin arvoja väleillä $(-\infty, -1)$, $(-1, 0)$ ja $[0, \infty)$. Miksei $t = -1$ vastaa mitään käyrän pistettä?

4. Ympyrä, jonka säde on $R = 1$, vierii liukumatta pitkin positiivista x-akselia. Ympyrän mukana pyörii siihen kiinnitetty jana, jonka toinen päätepiste on ympyrän keskipisteessä ja keskipiste on ympyrän kehällä. Määritä janan toisen (ympyrän ulkopuolella olevan) päätepisteen sijainti parametrisena käyränä $x = x(t),\ y = y(t)$, missä t on ympyrän vierimiskulma mitattuna alkutilanteesta, jossa janan päätepisteet ovat $(0, 1)$ ja $(0, 3)$. Hahmottele käyrä graafisesti. Missä pisteessä käyrä leikkaa ensimmäisen kerran itsensä? (Käyrää sanotaan *venytetyksi sykloidiksi.*)

5. Parametrisoi tason $T : x + y = 4$ ja kartion $K : xy + yz + xz = 0$ leikkauskäyrä ottamalla parametriksi a) $t = x$, b) $t = x - y$.

6. a) Avaruuskäyrän S parametrisointi on $\vec{r} = \vec{r}_0 + \cos t\,\vec{a} + \sin t\,\vec{b},\ t \in [0, 2\pi]$, missä $\vec{a}$, $\vec{b}$ ja $\vec{r}_0$ ovat avaruusvektoreita. Täsmälleen millä ehdoilla S on avaruusympyrä?

   b) Avaruusympyrän keskipiste on $(1, 1, 2)$, säde on $R = 3$ ja ympyrä on tasossa $x - y - 3z + 6 = 0$. Johda ympyrälle jokin parametrisointi muotoa $\vec{r} = \vec{r}_0 + \cos t\,\vec{a} + \sin t\,\vec{b},\ t \in [0, 2\pi]$.

7. Näytä, että yhtälöryhmä $\begin{cases} x^2 + y + z = 2 \\ xy + z = 1 \end{cases}$ määrittelee kaksi leikkaavaa avaruuskäyrää, joista toisen parametrisointi on $\vec{r} = t\vec{i} + (1+t)\vec{j} + (1-t-t^2)\vec{k},\ t \in \mathbb{R}$. Millainen on toinen käyrä?

8. (Kiukkulintu ja kuulantyöntäjät) a) Kiukkulintu lennätetään alkupisteestä $(x, y) = (0, 1)$ (pituusyksikkö = cm) venyttämällä heittoparaabelissa (ks. Esimerkki 6) vakion a arvoksi 8 cm ja tähtäämällä porsaaseen, joka on pisteessä $(4, 3)$. Millä k:n arvoilla tulee osuma?

   b) Teekkarit Yrjölä ja Ståhlberg kisaavat kuulantyönnössä. Ratkaise, kumpi voitti, kun kisaajien parhaissa työnnöissä heittoparaabelin parametrit ovat
   Yrjölä: $h = 2.00$ m, $\alpha = 60.0°$, $a = 8.00$ m
   Ståhlberg: $h = 1.80$ m, $\alpha = 30.0°$, $a = 6.65$ m

9. Esitä jokin parametrisointi seuraavien yhtälöiden määräämille pinnoille:
   a) $x^3y^2z = 5$, b) $(x-z)(x+z) + y + 2z = 0$, c) $x \sin z + xy^5 + y = 1$

10. a) Johda pinnalle S yhtälö muotoa $F(x, y, z) = 0$ parametrisoinnista

$$S: \begin{cases} x = 3 + 2\sin\theta\cos\varphi, \\ y = -1 + \sin\theta\sin\varphi, \\ z = 2 + 3\cos\theta, \quad (\theta, \varphi) \in \mathbb{R}^2. \end{cases}$$

   b) Pallon K keskipiste on $(1, 1, 1)$ ja säde on $R = 2$. Kuvan piirtoa varten halutaan parametrisoida pallon xy-tason yläpuolinen ($z \geq 0$) osa. Esitä parametrisointi!

   c) Pinnan S yhtälö lieriökoordinaateissa on $r = \varphi$, $(\varphi, z) \in A$, missä $A = [0, 4\pi] \times [-5, 5]$. Parametrisoi S viivoitinpintana. Millainen on S:n ja xy-tason leikkauskäyrä?

   d) Pinnan yhtälö lieriökoordinaateissa on $r = z^2|\cos\varphi|$. Esitä pinnan yhtälö karteesisissa koordinaateissa. Millaisia ovat pinnan ja tasojen $z = c$ ($c \in \mathbb{R}$) leikkauskäyrät?

11. a) Määritä sen viivoitinpinnan yhtälö (muodossa $F(x, y, z) = 0$), joka syntyy, kun suora $S:\ x = z,\ y = 1$ pyörähtää x-akselin ympäri. Totea, että sama pinta (nimeltään yksivaippainen hyperboloidi) syntyy myös, kun eräs xy-tason käyrä K pyörähtää x-akselin ympäri. Hahmottele K graafisesti.

   b) Tasokäyrän $K: x^2 - y^2 = 1$ pyörähtäessä x-akselin ympäri syntyy pinta nimeltä *kaksivaippainen hyperboloidi*. Määritä ko. pinnan yhtälö. Missä pisteissä suora $x = y = z$ leikkaa pinnan?

12. a) Puolikartion K kärki on origossa, symmetria-akseli on positiivinen z-akseli ja puolisuora $x = 2y = 3z,\ x \geq 0$ on pinnalla K. Parametrisoi K pyörähdyspintana ja viivoitinpintana. Mikä on K:n yhtälö lieriökoordinaatistossa?

    b) Parametrisoi kartio $K:\ xy + yz + xz = 0$ viivoitinpintana.

13. (*) Ympyrän keskipiste on origossa ja säde on a. Ympyrää pitkin sen sisäpuolella vierii liukumatta toinen ympyrä, jonka säde on $b < a$. Tällöin vierivän ympyrän kiinteä piste P piirtää tasokäyrän nimeltä *hyposykloidi*.
    a) Näytä, että pisteen $(a, 0)$ kautta kulkevan hyposykloidin parametriesitys on

    $$x = (a-b)\cos t + b\cos\left(\frac{a-b}{b}\,t\right), \quad y = (a-b)\sin t - b\sin\left(\frac{a-b}{b}\,t\right).$$

    b) Päättele, että tapauksessa $a = 2b$ piste P liikkuu pitkin janaa.
    c) Näytä, että tapauksessa $a = 4b$ parametriesitys yksinkertaistuu muotoon

    $$x = a\cos^3 t, \quad y = a\sin^3 t.$$

    Hahmottele tämän käyrän — nimeltään *asteroidi* — kulku. Mikä on asteroidin yhtälö karteesisissa koordinaateissa?

14. (*) (Sotaharjoitus 1) Origosta ammutun tykinkuulan lentorata on ajan t funktiona (yksiköt km ja s)

    $$\begin{cases} x(t) = (\sin\theta\cos\varphi + a)\,t, \\ y(t) = (\sin\theta\sin\varphi + b)\,t, \\ z(t) = (\cos\theta)\,t - 0.005\,t^2, \end{cases}$$

    missä θ, φ ovat suuntauskulmat ja a, b ovat tuuliparametreja. Maastoesteet asettavat suuntaukselle rajoituksen $\tan\theta > 0.2$. Miten suuntaus on valittava tuulettomassa säässä ($a = b = 0$), jotta ammus osuisi pisteessä $(10, 20, 0)$ olevaan maaliin? Kuinka korkealla ammus käy? Kuinka kauas maalista ammus osuu tällä suuntauksella, jos $a = 0.002$ ja $b = -0.001$?

15. (*) Jana, jonka pituus on 20, liikkuu seuraavasti: Janan keskipiste liikkuu z-akselia pitkin positiiviseen suuntaan vakionopeudella. Liikkuessaan jana pysyy xy-tason suuntaisena ja pyörii tasaisesti (kulmanopeus vakio) siten, että keskipisteen liikkuessa 30 pituusyksikköä jana pyörii täyden kierroksen positiivisen z-akselin suunnasta katsottuna vastapäivään. Esitä janan avaruuteen piirtämän viivoitinpinnan S parametrisointi, kun tiedetään lisäksi, että piste $(1, 0, 0)$ on tällä pinnalla. Leikkaako suora $z = 25,\ x + y = 4$ pinnan S?

IV.5 *Funktioavaruus

Tarkastellaan samassa joukossa A määriteltyjä yhden, kahden tai kolmen reaalimuuttujan reaaliarvoisia funktioita ja merkitään näiden joukkoa V:llä:

$$V \;=\; \{\,\text{funktiot } f:\; A \to \mathbb{R}\,\}.$$

Joukossa V on määritelty funktioiden yhteenlasku $f, g \mapsto f+g$ ja skalaarilla kertominen $f \mapsto \lambda f$ aiemmin kerrotulla tavalla (Määritelmän IV.1.2 säännöt (1) ja (2), kun $A \subset \mathbb{R}$). Näiden laskuoperaatioiden perusteella V on mahdollista tulkita vektoriavaruudeksi. Samassa joukossa määritellyt funktiot voidaan siis mieltää 'vektoreiksi', jolloin puhutaan *funktioavaruudesta* (engl. function space). Tällaisen avaruuden nolla-alkio on nk. (algebrallinen) *nollafunktio*, joka määritellään

$$\mathbf{0}(x) = 0 \;\; \forall x \in A.$$

Äärellinen funktiojoukko $\{f_i,\; i = 1 \ldots n\} \subset V$ on (albegrallisesti) *lineaarisesti riippumaton*, jos pätee

$$\sum_{i=1}^{n} \lambda_i f_i = \mathbf{0} \quad \Rightarrow \quad \lambda_i = 0,\; i = 1 \ldots n.$$

Esimerkki 1 Enintään kolmannen asteen polynomit (määrittelyjoukko $= \mathbb{R}$) muodostavat funktioalgebran yhdistelysääntöjen perusteella funktioavaruuden

$$V \;=\; \{f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \lambda_4 f_4 \mid \lambda_i \in \mathbb{R}\},$$

missä $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$ ja $f_4(x) = x^3$. Algebran peruslauseesta (ks. Luku III.3) on helposti pääteltävissä, että funktiot f_i ovat lineaarisesti riippumattomia, joten $\{f_i,\; i = 1 \ldots 4\}$ on V:n kanta. Siis V on neliulotteinen vektoriavaruus: dim $V = 4$. $\square$

Esimerkki 2 Funktiot muotoa $f(x) = c_1 \sin x + c_2 \cos x,\; c_1, c_2 \in \mathbb{R}$ muodostavat 2-ulotteisen funktioavaruuden, luonnollisena kantana $\{\sin x, \cos x\}$. $\square$

Esimerkki 3 Funktioavaruus

$$V = \{f(x) = c_1 + c_2 \cos^2 x + c_3 \sin^2 x,\; c_i \in \mathbb{R}\}$$

ei ole 3-ulotteinen, sillä funktiosysteemi $\{f_1\,, f_2\,, f_3\} = \{1, \cos^2 x, \sin^2 x\}$ on lineaarisesti riippuva:

$$f_1 - f_2 - f_3 = \mathbf{0}.$$

Mikä tahansa pari mainitusta kolmesta funktiosta sen sijaan on lineaarisesti riippumaton, joten V on 2-ulotteinen, kantana esim. $\{1, \cos^2 x\}$. Eräs V:n 1-ulotteinen aliavaruus on

$$W = \{\lambda \cos 2x \mid \lambda \in \mathbb{R}\},$$

sillä $\cos 2x = 2f_2 - f_1 \in V$. $\square$

Esimerkki 4 Funktioavaruus

$$V = \{f(x,y) = c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2,\ c_i \in \mathbb{R}\}$$

koostuu kahden muuttujan polynomeista enintään astetta 2 ja tämän aliavaruus

$$W = \{f(x,y) = c_1 + c_2 x + c_3 y,\ c_i \in \mathbb{R}\}$$

enintään ensimmäisen asteen polynomeista. Funktiosysteemi $\{1, x, y, x^2, xy, y^2\}$ on lineaarisesti riippumaton (Harj.teht. 4), joten $\dim V = 6$ ja $\dim W = 3$. $\square$

HARJOITUSTEHTÄVIÄ

1. Näytä, että funktiot $f_1(x) = x$ ja $f_2(x) = |x|$ ovat lineaarisesti riippumattomat, jos määrittelyjoukkona on väli $[-1, 1]$ ja lineaarisesti riippuvat, jos määrittelyjoukko on väli $[0, 1]$.

2. Montako alkiota on joukossa $A \subset \mathbb{R}$ oltava, jotta A:ssa määritellyt funktiot $f_i(x) = x^{i-1},\ i = 1 \ldots n$ ovat lineaarisesti riippumattomat?

3. a) Näytä, että $\{1, \sin x, \sin^2 x, \sin^3 x\}$ on erään 4-ulotteisen funktioavaruuden V kanta (funktioiden määrittelyjoukko $= \mathbb{R}$). b) Sikäli kuin $f(x) = 2 + \cos 2x - \sin 3x \in V$, määrää f:n koordinaatit (kertoimet) mainitussa kannassa.

4. Näytä, että Esimerkin 4 polynomiavaruus V on 6-ulotteinen, ts. näytä, että $c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2 = 0\ \forall (x,y) \in \mathbb{R}^2 \ \Rightarrow\ c_1 = \ldots = c_6 = 0$.

5. (*) Olkoon $V = \{f(x) = c_1 + c_2 x + c_3 x^2 \mid c_i \in \mathbb{R}\}$. Näytä, että

$$\langle f, g \rangle = f(0)g(0) + f(1)g(1) + f(2)g(2), \quad f, g \in V$$

   määrittelee V:n skalaaritulon. Laske funktion $u(x) = 2 + 3x - 2x^2$ ortogonaaliprojektio w (ko. skalaaritulon mielessä) V:n aliavaruuteen W, jonka kanta on $\{1, x\}$. Piirrä samaan kuvaan funktioiden u ja w kuvaajat välillä $[0, 2]$. Tarkista, että $u - w \perp W$, ts. $\langle u - w, v \rangle = 0\ \forall v \in W$.

Luku V

Jatkuvuus ja derivoituvuus

Funktion *jatkuvuus* ja *derivoituvuus* ovat käsitteellisiä peruslähtökohtia matematiikan suuntauksessa, jota kutsutaan väljästi *analyysiksi*. Jatkuvuudesta, tai yleisemmin funktion *säännöllisyydestä* puhuttaessa on kyse funktion arvojen ennustettavuudesta muuttujan (muuttujien) arvojen vaihdellessa. Jatkuvuuden ja derivoituvuuden käsitteet tuovat matemaattisten funktioiden teoriaan kokonaan oman 'makunsa' verrattuna tähän asti tarkasteltuun funktioiden algebraan.

Tässä luvussa tarkastelun kohteena ovat pääosin vain yhden reaalimuuttujan funktiot. Näille funktioille määritellään ensin jatkuvuus peruskäsitteenä ja jatkuvuuteen läheisesti liittyvä *funktion raja-arvon* käsite. Pelkkää jatkuvuutta vahvempina säännöllisyyden lajeina määritellään mm. derivoituvuus (Luku V.3) ja derivoituvuutta vahvemmat *sileyden* lajit (Luku V.5). Näiden käsitteiden pohjalta luonnehditaan funktioita, mm. esitetään kaksi reaalimuuttujan analyysin keskeistä *väliarvolausetta* ja tarkastellaan näiden lauseiden seuraamuksia yhtälöiden ja myös yksinkertaisten *differentiaaliyhtälöiden* ratkeavuusteoriassa.

Derivaatta tuo mukanaan myös derivoimissäännöt eli uuden ulottuvuuden funktioalgebraan. Luvuissa V.3–V.4 johdetaan kaikki keskeiset derivoimissäännöt, mukaanlukien implisiittifunktiot, potenssisarjan summana määritellyt funktiot ja (Luvussa V.4 erikseen) trigonometriset funktiot.

Luvussa V.7 tarkastelun kohteena ovat yhtälöitten likimääräisessä ratkaisussa yleisesti käytettävien algoritmien, *kiintopisteiteraatioiden*, toimintaperiaatteet ja suppenevuusteoria. Luvussa V.8 tarkastellaan lyhyesti jatkuvuuden ja derivoituvuuden käsitteiden laajentamista kompleksifunktioihin ja määritellään tähän liittyen *analyyttisen* kompleksifunktion käsite. Viimeisessä osaluvussa todistetaan jatkuvien funktioiden teorian keskeisimpiä väittämiä kuten *Weierstrassin lause*. Tässä yhteydessä esitetään myös lyhyesti, miten Algebran peruslause on todistettavissa kompleksialgebran ja analyysin tuloksia yhdistelemällä.

V.1 Jatkuvuuden käsite

Funktion jatkuvuuden ongelma tulee eteen niinkin yksinkertaisessa tehtävässä kuin funktion arvon numeerisessa määrittämisessä, eli numeerisessa funktioevaluaatiossa. Tarkastellaan Luvun IV.2 Esimerkin 3 käänteisfunktiota, joka nyt kirjoitettakoon muotoon

$$y = f(x) \;\Leftrightarrow\; y^5 + 3y = x, \quad x \in \mathbb{R}.$$

Tehtävänä olkoon laskea likimäärin luku $a = f(\pi)$, eli evaluoida f numeerisesti π:ssä. Tähän itse asiassa sisältyy kaksi numeerista ongelmaa: Ensinnäkin luvulle π ei ole olemassa 'tarkkaa' numeerista arvoa, ja toiseksi yhtälöä ei yleensä voi ratkaista y:n suhteen tarkasti, vaikka x olisi rationaaliluku. Käytännössä menetellään (laskin/tietokone menettelee) seuraavasti: Valitaan π:lle edustaja x_n rationaalilukujonosta $\{x_n\}$, jolle pätee $x_n \to \pi$. Lasketaan $a_n = f(x_n)$ likimäärin käyttämällä jotakin numeerista algoritmia yhtälön $y^5 + y = x_n$ ratkaisuun. Algoritmi tuottaa käytännössä rationaalilukujonon $\{b_k\}$, jolle pätee $b_k \to a_n$ kun $k \to \infty$. Poimitaan tästä jonosta luvun a_n likiarvoksi b_m jollakin m (esim. $m = 10$), ja ilmoitetaan lopputuloksena tämän luvun likiarvo äärellisenä desimaalilukuna (katkaistuna tai pyöristettynä liukulukuna).

Jos em. laskussa ei huomioida numeerisia pyöristysvirheitä liukulukulaskennassa, niin lopputuloksen b_m virhe koostuu kahdesta osasta:

$$b_m - f(\pi) \;=\; [b_m - f(x_n)] + [f(x_n) - f(\pi)].$$

Tässä ensimmäinen osa on approksimaation $b_m \approx f(x_n)$ virhe, eli kyse on yhtälön $y^5 + 3y = x_n$ ratkaisualgoritmin tarkkuudesta. Virheen jälkimmäisessä osassa sen sijaan on kyse, paitsi approksimaation $x_n \approx x$ tarkkuudesta, myös funktiosta f: Kyse on funktion <u>jatkuvuudesta</u>. Kvalitatiivisesti väittämä 'f on jatkuva x:ssä' tarkoittaa, että funktioevaluaatio $x \mapsto f(x)$ on luotettava, kun muuttujalle sallitaan pieni vaihtelu, ts. pätee

$$x_n \approx x \;\Rightarrow\; f(x_n) \approx f(x).$$

Em. funktion tapauksessa jatkuvuuskysymys ratkeaa seuraavasti: Koska

$$\begin{cases} a^5 + a = \pi, \\ a_n^5 + a_n = x_n, \end{cases}$$

saadaan vähennyslaskulla (vrt. Luvun IV.2 tarkastelut)

$$(a - a_n)(a^4 + a^3 a_n + a^2 a_n^2 + a a_n^3 + a_n^4 + 3) = \pi - x_n.$$

Tässä voidaan turvallisesti olettaa, että $a, a_n \geq 0$, joten seuraa

$$|a_n - a| \leq \tfrac{1}{3}|x_n - \pi| \quad \Leftrightarrow \quad |f(x_n) - f(\pi)| \leq \tfrac{1}{3}|x_n - \pi|.$$

Jatkuvuudelle on näin saatu jopa kvantitatiivinen varmistus: Approksimaation $f(x_n) \approx f(\pi)$ virhe on enintään kolmasosa approksimaation $x_n \approx \pi$ virheestä.

Esimerkki johdattaa seuraavaan funktion jatkuvuuden määritelmään (vaihtoehtoinen määritelmä jäljempänä).

MÄÄRITELMÄ V.1.1 **(Jatkuvuus: jonokriteeri)** Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on *jatkuva* (engl. continuous, ruots. kontinuerlig) *pisteessä* $a \in \mathcal{D}_f$ (tai a:ssa), jos kaikille reaalilukujonoille $\{x_n\}$ pätee

$$x_n \in \mathcal{D}_f \ \forall n \quad \& \quad x_n \to a \quad \Rightarrow \quad f(x_n) \to f(a).$$

Jos f ei ole jatkuva pisteessä $a \in \mathcal{D}_f$, niin f on *epäjatkuva* (engl. discontinuous) pisteesä a.

ESIMERKKI 1 Funktio $f(x) = \begin{cases} x - 1, & \text{jos } x < \pi \\ x, & \text{jos } x \geq \pi \end{cases}$

on epäjatkuva pisteessä $x = \pi$, sillä jos $x_n \to \pi$ ja $x_n < \pi \ \forall n$, niin $f(x_n) = x_n - 1 \to \pi - 1 \neq f(\pi) = \pi$. Muissa pisteissä f on jatkuva, sillä jos esim. $a < \pi$ ja $x_n \to a$, niin jostakin indeksistä N eteenpäin on $|x_n - a| < \pi - a$ (lukujonon suppenemisen määritelmässä valittu $\varepsilon = \pi - a$), jolloin on erityisesti $x_n - a < \pi - a \ \Leftrightarrow \ x_n < \pi$, kun $n > N$. Tällöin on $f(x_n) = x_n - 1, \ n > N$ ja siis $f(x_n) \to a - 1 = f(a)$. $\square$

ESIMERKKI 2 Olkoon $\mathcal{D}_f = \{a\}$ $(a \in \mathbb{R})$ ja $f(a) = c \in \mathbb{R}$. Määritelmän V.1.1 mukaisesti f on jatkuva pisteessä a (!). Yleisemmin reaalifunktio on jatkuva jokaisessa määrittelyjoukkonsa nk. *eristetyssä pisteessä*, ks. Harj.teht. 4. $\square$

Jatkuvuus voidaan määritellä myös suoraan vetoamatta lukujonoihin, jolloin määritelmästä tulee lukujonon suppenemisen määritelmää (Määritelmä I.6.1) muistuttava. Vaihtoehtoinen määritelmä on seuraava.

MÄÄRITELMÄ V.1.2 **(Jatkuvuus: (ε, δ)-kriteeri)** Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on jatkuva pisteessä $a \in \mathcal{D}_f$, jos jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että jokaisella $x \in \mathbb{R}$ pätee

$$x \in \mathcal{D}_f \quad \& \quad |x - a| < \delta \quad \Rightarrow \quad |f(x) - f(a)| < \varepsilon.$$

Määritelmiä V.1.1 ja V.1.2 verrattaessa ei näytä aivan ilmeiseltä, että pätee (ks. todistus luvun lopussa)

*Lause V.1.3 Jatkuvuuden määritelmät V.1.1 ja V.1.2 ovat yhtäpitävät.

Esimerkki 1 (jatko) Jos $a \neq \pi$, niin esimerkin funktiolle pätee

$$|f(x) - f(a)| = |x - a|, \quad \text{kun } |x - a| < |\pi - a|,$$

joten Määritelmän V.1.2 ehto on voimassa, kun valitaan $\delta = \min\{\varepsilon, |\pi - a|\} > 0$. Siis f on jatkuva pisteissä $a \neq \pi$ Määritelmän V.1.2 mukaisesti. $\square$

Esimerkki 3 Näytä, että $f(x) = x^2$ on jatkuva jokaisessa pisteessä $a \in \mathbb{R}$ käyttäen jatkuvuuden a) jonokriteeriä, b) (ε, δ)-kriteeiä.

<u>Ratkaisu</u> a) Määritelmän V.1.1 mukainen jatkuvuus on seuraus Lauseesta I.7.3:

$$x_n \to a \quad \Rightarrow \quad x_n^2 \to a^2 \quad \Rightarrow \quad f(x_n) \to f(a).$$

b) Jos $a \in \mathbb{R}$ ja $|x - a| \leq 1$, niin kunta-algebran ja kolmioepäyhtälön nojalla

$$\begin{aligned} |f(x) - f(a)| \;=\; |x - a||x + a| \;&=\; |x - a||2a + (x - a)| \\ &\leq\; |x - a|(2|a| + |x - a|) \;\leq\; |x - a|(2|a| + 1). \end{aligned}$$

Näin ollen jos $\varepsilon > 0$, niin pätee $|f(x) - f(a)| < \varepsilon$ aina kun $|x - a| < \delta = \min\{1, \varepsilon/(2|a| + 1)\}$ (jolloin myös $|x - a| < 1$). Koska tässä on $\delta > 0$ aina kun $\varepsilon > 0$, niin f on jatkuva a:ssa Määritelmän V.1.2 mukaisesti. $\square$

Jatkuvuuden määrittely jonokriteerillä (Määritelmä V.1.1) on kätevää monissa teoreettisissa tarkasteluissa, jotka tällä tavoin palautuvat suppenevien lukujonojen teoriaan. (Tämä teoria on kylläkin tunnettava, mitä voi pitää myös haittapuolena.) Määritelmä V.1.2 on jatkuvuuden perinteisempi 'koulumääritelmä'. Tämä on usein Määritelmää V.1.1 selkeämpi silloin kun halutaan selvittää, miltä jatkuvat funktiot 'näyttävät'. Esimerkiksi seuraava usein käytetty tulos, joka kertoo jatkuvan funktion 'jäykkyydestä', on tästä määritelmästä helppo johtaa. Todistus jätetään harjoitustehtäväksi (Harj.teht. 3a).

Propositio V.1.4 Jos f on jatkuva pisteessä a ja $f(a) > 0$ $(f(a) < 0)$, niin $\exists \delta > 0$ siten, että $f(x) > 0$ $(f(x) < 0)$ aina kun $x \in (a - \delta, a + \delta) \cap \mathcal{D}_f$.

Jatkuvuus yksittäisessä pisteessä voidaan laajentaa koskemaan joukkoa $A \subset \mathcal{D}_f$: Funktio on jatkuva A:ssa, jos se on jatkuva A:n jokaisessa pisteessä. Jatkuvuustarkasteluille on kuitenkin tyypillistä, että tarkastelu rajoittuu <u>välille</u> $A \subset \mathcal{D}_f$ siten, että funktion ominaisuuksilla välin ulkopuolella ei ole lainkaan merkitystä. Sen vuoksi on tapana asettaa mainitusta yleisestä säännöstä hieman poikkeava

MÄÄRITELMÄ V.1.5 (**Jatkuvuus välillä**) Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on *jatkuva välillä* $A \subset \mathcal{D}_f$, jos jokaisella $x \in A$ ja jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$x_n \in A \ \ \& \ \ x_n \to x \ \ \Rightarrow \ \ f(x_n) \to f(x).$$

Tämän mukaisesti jatkuvuus välillä A tarkoittaa samaa kuin Määritelmän V.1.1 mukainen jatkuvuus A:n jokaisessa pisteessä siinä tapauksessa, että funktion määrittelyjoukko rajataan väliksi A. (Jatkuvuuden vaihtoehtoisessa määritelmässä V.1.2 korvataan ehto $x \in \mathcal{D}_f$ ehdolla $x \in \mathcal{D}_f \cap A$.) Jatkuvuus välillä määritellään siis jatkuvuutena 'sisältä päin' ko. joukossa. Jos väli on avoin, ei tämä rajaus ole tarpeen, sillä jos $x \in (a, b) = A$ ja $x_n \to x$, niin jollakin N pätee $x_n \in A$ $\forall n > N$ (Määritelmä I.6.1, $\varepsilon = \min\{x - a, b - x\} > 0$) eli määritelmän ehto toteutuu indeksistä N eteenpäin joka tapauksessa. Sen sijaan jos väli on suljettu, on ehdolla $x_n \in A$ merkitystä välin päätepisteissä (ei sisäpisteissä).

ESIMERKKI 1 (jatko) Jos $b > \pi$, on esimerkin funktio f jatkuva välillä $[\pi, b]$ (Määritelmä V.1.5) vaikka f ei ole jatkuva pisteessä π (Määritelmä V.1.1). Jos $a < \pi$, on f jatkuva välillä $[a, \pi)$ mutta ei välillä $(a, \pi]$. $\square$

ESIMERKKI 4 Funktio $f(x) = \begin{cases} 1, & \text{jos } x \in \mathbb{Q}, \\ 0, & \text{jos } x \in \mathbb{R},\ x \notin \mathbb{Q} \end{cases}$

on esimerkki 'patologisesta' funktiosta, joka on määritelty koko $\mathbb{R}$:ssä, mutta joka ei ole missään pisteessä jatkuva. Funktion nimi on *Dirichlet'n funktio.* $\square$

Jatkuvien funktioiden yhdistely

Esimerkin 4 vastapainoksi voidaan kysyä, millaiset 'normaalit' funktiot tyypillisesti ovat jatkuvia. Ensimmäinen tuntuma näihin saadaan yhdistelemällä yksinkertaisia funktioita peruslaskutoimituksilla Määritelmän IV.1.2 mukaisesti. Koska funktion jatkuvuudessa on viime kädessä kyse lukujonon $\{f(x_n)\}$ suppenemisesta, on seuraava lause välitön seuraus Lauseesta I.7.3 (Harj.teht. 3b).

LAUSE V.1.6 (**Jatkuvuuden yhdistelysäännöt**) Jos $f : \mathcal{D}_f \to \mathbb{R}$ on jatkuva pisteessä $x \in \mathcal{D}_f$, niin λf on jatkuva x:ssä $\forall \lambda \in \mathbb{R}$. Jos lisäksi $x \in \mathcal{D}_g$ ja $g : \mathcal{D}_g \to \mathbb{R}$ on jatkuva pisteessä x, niin $f + g$ ja fg ovat jatkuvia x:ssä. Jos edelleen $g(x) \neq 0$, niin myös f/g on jatkuva x:ssä.

ESIMERKKI 5 Jokainen polynomi tai rationaalifunktio voidaan määritellä palautuvasti algebrallisena yhdistelmänä perusfunktioista $f_0(x) = 1$ ja $f_1(x) = x$. Koska nämä ovat ilmeisen jatkuvia $\mathbb{R}$:ssä, niin päätellään Lauseesta V.1.6, että

jokainen polynomi on jatkuva $\mathbb{R}$:ssä ja jokainen rationaalifunktio määrittelyjoukossaan, eli muualla kuin nimittäjänsä nollakohdissa. $\square$

Esimerkki 6 Jos pidetään tunnettuna, että trigonometriset funktiot sin ja cos ovat jatkuvia $\mathbb{R}$:ssä, niin Lauseen V.1.6 perusteella funktiot $\tan = \sin / \cos$ ja $\cot = \cos / \sin$ ovat samoin jatkuvia koko määrittelyjoukossaan. $\square$

Lause V.1.7 (**Yhdistetyn funktion jatkuvuus**) Jos g on jatkuva pisteessä $x \in \mathcal{D}_g$, $g(x) \in \mathcal{D}_f$ ja f on jatkuva pisteessä $g(x)$, niin yhdistetty funktio $f \circ g$ on jatkuva x:ssä.

Todistus Koska $x_n \in \mathcal{D}_{f\circ g} \Leftrightarrow x_n \in \mathcal{D}_g \wedge g(x_n) \in \mathcal{D}_f$, niin jonokriteerin avulla päätellään:

$$\begin{aligned} x_n \in \mathcal{D}_{f\circ g} \ \& \ x_n \to x \quad &\Rightarrow \quad g(x_n) \in \mathcal{D}_f \ \& \ x_n \in \mathcal{D}_g \ \& \ x_n \to x \\ &\Rightarrow \quad g(x_n) \in \mathcal{D}_f \ \& \ g(x_n) \to g(x) \\ &\Rightarrow \quad f(g(x_n)) \to f(g(x)). \qquad \square \end{aligned}$$

Esimerkki 7 Näytä, että jos f on jatkuva x:ssä, niin myös $|f|$ on jatkuva x:ssä.

Ratkaisu Funktio $g(x) = |x|$ on helposti osoitettavissa jatkuvaksi $\mathbb{R}$:ssä suoraan jatkuvuuden määritelmistä (tai vedoten Propositioon V.1.8 alla), joten Lauseen V.1.7 nojalla yhdistetty funktio $(g\circ f)(x) = |f(x)|$ on jatkuva jokaisessa pisteessä, jossa f on jatkuva. $\square$

Jatkuvuuden periytyvyyden paloittain (eli eri väleillä erilaisilla laskusäännöillä) määritellyn funktion osalta ratkaisee (Harj.teht. 3c; vrt. myös Esimerkki 1 edellä).

Propositio V.1.8 Olkoon f_1 ja f_2 jatkuvia avoimella välillä (a, b) ja

$$f(x) = \begin{cases} f_1(x), & \text{kun } x \in (a, c), \\ k, & \text{kun } x = c, \\ f_2(x), & \text{kun } x \in (c, b), \end{cases}$$

missä $a < c < b$. Tällöin f on jatkuva c:ssä täsmälleen kun $f_1(c) = f_2(c) = k$.

Esimerkki 8 Funktio

$$f(x) = \begin{cases} x + a, & \text{kun } x \le a, \\ -(x+1)^2, & \text{kun } x > a \end{cases}$$

on polynomi $f_1(x) = x + a$ välillä $(-\infty, a)$ ja polynomi $f_2(x) = -(x+1)^2$ välillä (a, ∞), joten näillä väleillä funktio on jatkuva. Proposition V.1.8 mukaan ehto funktion jatkuvuudelle pisteessä $x = a$ (ja siis ehto jatkuvuudelle koko $\mathbb{R}$:ssä) on

$$f_1(a) = f_2(a) \quad \Leftrightarrow \quad a^2 + 4a + 1 = 0 \quad \Leftrightarrow \quad a = -2 \pm \sqrt{3}. \quad \square$$

Funktio $f(x) = \sqrt[m]{x}$

Osoitetaan seuraavaksi, että funktio $f : [0, \infty) \to [0, \infty)$, joka määritellään

$$f(x) = x^{1/m} = \sqrt[m]{x}, \quad m \in \mathbb{N},\ m \geq 2$$

on koko määrittelyvälillään jatkuva (kuvassa $f(x)$, kun $m = 8$).

Aloitetaan pisteestä $x = 0$, joka on jatkuvuuden kannalta kriittisin (vrt. kuva). Koska f on aidosti kasvava (ks. Luku IV.2, Esimerkki 2), niin jokaisella $a > 0$ pätee

$$0 \leq x < a^m \ \Leftrightarrow\ 0 \leq f(x) < f(a^m) = a.$$

Jos nyt $x_n \geq 0$ ja $x_n \to 0$, niin jokaisella $\varepsilon > 0$ $\exists N \in \mathbb{N}$ siten, että pätee

$$n > N \ \Rightarrow\ 0 \leq x_n < \varepsilon^m \ \Rightarrow\ 0 \leq f(x_n) < \varepsilon.$$

Näin ollen lukujono $\{f(x_n)\}$ on suppeneva ja $f(x_n) \to 0 = f(0)$, joten f on jatkuva 0:ssa.

Muualla kuin origossa f:n jatkuvuus voidaan päätellä kunta-algebran avulla: Jos $x_n \geq 0$ ja $x_n \to x > 0$, niin merkitsemällä $y = f(x)$, $y_n = f(x_n)$ seuraa

$$x - x_n \ = \ y^m - y_n^m \ = \ (y - y_n)(y^{m-1} + y^{m-1}y_n + \cdots + y_n^{m-1}).$$

Tässä on $y_n \geq 0$, $y > 0$, joten päätellään

$$|x - x_n| \ \geq\ |y - y_n| y^{m-1} \ \Leftrightarrow\ |y - y_n| \ \leq\ y^{1-m}|x - x_n|.$$

Näin ollen $x_n \to x \ \Rightarrow\ y_n \to y$, eli f on jatkuva pisteessä x.

Esimerkki 9 Olkoon $\varepsilon = 10^{-6}$. Määritä suurin δ siten, että Määritelmän V.1.2 ehto toteutuu funktiolle $f(x) = \sqrt[8]{x}$ pisteessä $x = 0$.

Ratkaisu Koska f on aidosti kasvava välillä $[0, \infty) = \mathcal{D}_f$, niin $\forall x \in \mathcal{D}_f$ pätee $|f(x) - f(0)| = \sqrt[8]{x} < \varepsilon \;\Leftrightarrow\; 0 \le x < \varepsilon^8 \;\Leftrightarrow\; x \in (-\varepsilon^8, \varepsilon^8) \cap \mathcal{D}_f$. Siis $\delta_{max} = \varepsilon^8 = 10^{-48}$. $\square$

Funktion $f(x) = x^{1/m}$ jatkuvuuden tultua todetuksi seuraa Lauseesta V.1.7 ja funktion $g(x) = |x|$ jatkuvuudesta, että yhdistetty funktio $(f \circ g)(x) = |x|^{1/m}$ on jatkuva koko määrittelyjoukossaan ($= \mathbb{R}$). Kun huomioidaan myös Lause V.1.6, niin seuraa

Propositio V.1.9 Funktio $f(x) = |x|^\alpha$, $\alpha = p/q \in \mathbb{Q}$ on koko määrittelyjoukossaan ($\mathcal{D}_f = \mathbb{R}$ jos $\alpha > 0$, $\mathcal{D}_f = \{x \in \mathbb{R} \mid x \neq 0\}$ jos $\alpha \le 0$) jatkuva.

Lauseiden V.1.6–V.1.7, Proposition V.1.9 ja Esimerkkien 6 ja 7 perusteella voidaan vetää se yleisempi johtopäätös, että kaikki toistaiseksi tunnetut 'yhden lausekkeen' funktiot ovat jatkuvia koko määrittelyjoukossaan (!).

Esimerkki 10 Ilman tapauskohtaista tarkastelua voidaan esimerkiksi seuraavat funktiot päätellä jatkuviksi määrittelyjoukkonsa jokaisessa pisteessä:

$$\sqrt[3]{|1 - \sqrt{x}\,|}\,, \quad \sqrt[6]{\frac{1 + \sqrt{x}}{2 - x^2}}\,, \quad \frac{|\sin x|^{3/4}}{x - \pi}\,, \quad \frac{\cot x}{\sqrt[4]{x + 50\cos(\tan x)}}\,. \quad \square$$

Jatkuvien funktioiden päälauseet

Tässä osaluvussa esitetään kolme matemaattisen analyysin keskeistä lausetta, jotka kaikki koskevat suljetulla välillä jatkuvia funktioita. Ensimmäinen lauseista on myös ensimmäinen yhden reaalimuuttujan funktioita koskevista *väliarvolauseista*, joita on kaikkiaan kolme. (Muut kaksi esitetään myöhemmin.) Tässä esitettävistä päälauseista jälkimmäiset kaksi ovat erikoistapauksia yleisemmistä väittämistä, jotka perustuvat jatkuvuuden syvällisempään logiikkaan. Näiden todistuksia ei vielä esitetä, vaan lauseet muotoillaan ja todistetaan jäljempänä Luvussa V.9 tässä esitettyä yleisemmässä muodossa.

Tarkastellaan suljetulla välillä $[a, b]$ jatkuvaa funktiota f. Olkoon $f(a) \neq f(b)$ ja $c < \eta < d$, missä

$$c = \min\{f(a), f(b)\}, \quad d = \max\{f(a), f(b)\}.$$

Koska f on jatkuva, niin tuntuu luonnolliselta ajatella, että f:n kuvaaja välillä $[a, b]$ on 'jatkuva lanka', joka yhdistää pisteet $A = (a, f(a))$ ja $B = (b, f(b))$.

Tämän intuition mukaisesti tuntuu selvältä, että kuvaajan on leikattava suora $y = \eta$ ainakin kerran välillä (a, b). Toisin sanoen, probleemalla

$$x \in [a, b] : \quad f(x) = \eta$$

on oltava ainakin yksi ratkaisu $x = \xi \in (a, b)$ jokaisella $\eta \in (c, d)$ (kuva).

Ym. geometriselle intuitiolle vahvistuksen antaa

Lause V.1.10 **(Ensimmäinen väliarvolause — Jatkuvien funktioiden väliarvolause)** Jos $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on jatkuva välillä $[a, b] \subset \mathcal{D}_f$ ja $c = \min\{f(a), f(b)\} < d = \max\{f(a), f(b)\}$ (eli $f(a) \neq f(b)$), niin jokaisella $\eta \in (c, d)$ on olemassa $\xi \in (a, b)$ siten, että $f(\xi) = \eta$.†

Jos f ja η täyttävät Lauseen V.1.10 ehdot, niin nähdään, että lauseen väittämä seuraa, kun seuraavaa pelkistetympää tulosta sovelletaan funktioon $g(x) = f(x) - \eta$ tai $g(x) = \eta - f(x)$.

Lause V.1.11 **(Bolzano)** Jos $f(a) < 0$, $f(b) > 0$, ja f on jatkuva välillä $[a, b]$, niin $f(\xi) = 0$ jollakin $\xi \in (a, b)$.

†Lause V.1.10 varmistaa, että välillä $[a, b]$ jatkuvan funktion kuvaaja

$$G_f = \{P = (x, y) \mid x \in [a, b] \ \& \ y = f(x)\}$$

on geometrisen intuition mukainen *käyrä* eli 'yhtenäinen viiva vailla leveyttä'. G_f on 'vailla leveyttä', koska $f(x)$ on yksikäsitteinen $\forall x$ (eli f on funktio), ja 'yhtenäinen', koska f on jatkuva. — Jatkuvuus (yhtenäisyyden takaajana) liitetään käyrän käsitteeseen yleisemminkin. Esimerkiksi yleisessä parametrisessa käyrässä $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$ koordinaattifunktiot $x(t), y(t), z(t)$ oletetaan jatkuviksi tarkasteltavalla välillä. Vrt. alaviitteet Luvuissa IV.2–IV.4.

Todistus on puhtaasti konstruktiivinen ja toimii käytännössäkin algoritmina, joka etsii yhden f:n nollakohdista välillä (a, b). Tavallisin todistustapa on nk. haarukointi eli puolitusmenetelmä (kymmenjakoalgoritmikin toimisi, vrt. Luku I.9). Ensinnäkin voidaan funktioevaluaatioiden (oletetaan tarkoiksi!) perusteella todeta, että joko valinnalla $a_1 = a$, $b_1 = \frac{1}{2}(a + b)$ tai valinnalla $a_1 = \frac{1}{2}(a + b)$, $b_1 = b$ pätee

$$f(a_1) \leq 0 \quad \& \quad f(b_1) \geq 0.$$

Jos tässä on $f(a_1) = 0$ tai $f(b_1) = 0$, on ξ löydetty. Muussa tapauksessa on $f(a_1) < 0$ ja $f(b_1) > 0$, jolloin voidaan jatkaa ξ:n 'haarukointia' välillä (a_1, b_1) samalla periaatteella. Mikäli konstruktio ei katkea ξ:n löytymiseen, syntyy kaksi jonoa $\{a_n\}$, $\{b_n\}$. Konstruktion perusteella $\{a_n\}$ on kasvava, $\{b_n\}$ on vähenevä ja

$$b_n - a_n = \left(\frac{1}{2}\right)^n (b - a) \to 0,$$

joten jonoilla on yhteinen raja-arvo: $a_n \to \xi$ ja $b_n \to \xi$ jollakin $\xi \in \mathbb{R}$. Koska on $a \leq a_n \leq b$ $\forall n$, on myös oltava $a \leq \xi \leq b$ (Propositio I.7.4: V1), eli $\xi \in [a, b]$. Tällöin koska f on välillä $[a, b]$ jatkuva (Määritelmä V.1.5), pätee

$$f(a_n) \to f(\xi) \quad \text{ja} \quad f(b_n) \to f(\xi).$$

Mutta lukujonojen $\{a_n\}$ ja $\{b_n\}$ konstruktion perusteella

$$\begin{aligned} f(a_n) < 0 \ \forall n \ &\Rightarrow \ f(\xi) \leq 0, \\ f(b_n) > 0 \ \forall n \ &\Rightarrow \ f(\xi) \geq 0 \end{aligned}$$

(Propositio I.7.4: V1), joten on oltava $f(\xi) = 0$. Tässä oli $\xi \in [a, b]$, mutta koska $f(a) < 0$ ja $f(b) > 0$, niin $\xi \in (a, b)$. Bolzanon lause on näin todistettu. □

Esimerkki 11 Funktiolle $f(x) = x^5 + 3x$ pätee $f(0) = 0$ ja $f(1) = 4$, joten Lauseen V.1.10 mukaan yhtälöllä

$$x^5 + 3x = 2$$

on ratkaisu $x = \xi$ välillä $(0, 1)$. Ratkaisu on itse asiassa yksikäsitteinen, koska f on 1-1 (vrt. Esimerkki 3 Luvussa IV.2). □

Seuraava lause, jolla on myös yleisempiä muotoja, on reaalianalyysin merkittävimpiä tuloksia (todistus yleisemmässä muodossa Luvussa V.9).

*Lause V.1.12 **(Weierstrass)** Suljetulla välillä $[a, b]$ jatkuva funktio f saavuttaa ko. välillä pienimmän ja suurimman arvonsa, ts. on olemassa $x_1 \in [a, b]$ ja $x_2 \in [a, b]$ siten, että

$$f(x_1) = \min\{f(x) \mid x \in [a, b]\}, \quad f(x_2) = \max\{f(x) \mid x \in [a, b]\}.$$

Esimerkki 12 Funktio

$$f(x) = \begin{cases} x, & \text{kun } x \leq 0, \\ 1 - x, & \text{kun } x > 0 \end{cases}$$

on jatkuva välillä $[a, b]$, jos joko $a < b \leq 0$ tai $0 < a < b$. Tällöin funktio saavuttaa maksimi- ja minimiarvonsa välin päätepisteissä. Muussa tapauksessa, eli jos $a \leq 0$ ja $b > 0$, ei f ole jatkuva välillä $[a, b]$, sillä jos $x_n \in (0, b]$ ja $x_n \to 0$, niin $f(x_n) \to 1 \neq f(0)$. Tässä tapauksessa f ei myöskään saavuta maksimiarvoa välillä $[a, b]$, ainoastaan minimiarvon ($= \min\{f(a), f(b)\}$). $\square$

Yhdistämällä Lauseiden V.1.10 ja V.1.12 väittämät tullaan seuraavaan maininnan arvoiseen päätelmään:

Korollaari V.1.13 Jos f on jatkuva välillä $A = [a, b]$ ja f ei ole vakio, niin $f(A)$ on suljettu väli.

Seuraava lause antaa erään riittävän kriteerin käänteisfunktion jatkuvuudelle. Todistus (hieman yleisemmin muotoillulle lauseelle) esitetään Luvussa V.9.

*Lause V.1.14 **(Käänteisfunktion jatkuvuus)** Jos $f : [a, b] \to [c, d]$ on jatkuva bijektio, niin myös $f^{-1} : [c, d] \to [a, b]$ on jatkuva bijektio.

Esimerkki 13 Funktio $f(x) = x^{1/m}$, $m \in \mathbb{N}$, $m \geq 2$ pääteltiin edellä jatkuvaksi välillä $[0, \infty)$. Tämä seuraa myös Lauseesta V.1.14, sillä $f = g^{-1}$, missä $g(x) = x^m$ on jatkuva bijektio $g : [0, a] \to [0, a^m]$ jokaisella $a > 0$. $\square$

*Lauseen V.1.3 todistus

Lauseen V.1.3 väittämä on kaksiosainen: Väitetään, että jos f on jatkuva pisteessä $a \in \mathcal{D}_f$ Määritelmän V.1.1 jonokriteerillä, niin f on jatkuva a:ssa myös Määritelmän V.1.2 (ε, δ)-kriteerillä ja että tämä implikaatio pätee myös kääntäen.

Todistus $\boxed{\Rightarrow}$ Oletetaan, että f on jatkuva pisteessä $a \in \mathcal{D}_f$ Määritelmän V.1.1 mukaisesti. Tehdään vastaoletus: f ei ole jatkuva a:ssa Määritelmän V.1.2 kriteerillä. Tällöin voidaan samalla tavoin kuin Lauseen I.10.8 todistuksessa päätellä, että jollakin $\varepsilon > 0$ pätee:

$$\forall \delta > 0 \; \exists x \in \mathcal{D}_f \text{ siten, että } \; |x - a| < \delta \;\; \& \;\; |f(x) - f(a)| \geq \varepsilon.$$

Kun nyt δ:n arvoista muodostetaan jono $\{\delta_n\}$, joka suppenee kohti nollaa (esim. $\delta_n = 1/n$), niin vastaavasti on siis löydettävissä jono $\{x_n\}$ siten, että jokaisella n

$$x_n \in \mathcal{D}_f \;\; \& \;\; |x_n - a| < \delta_n \;\; \& \;\; |f(x_n) - f(a)| \geq \varepsilon.$$

Näin on löydetty jono, jolle pätee

$$x_n \in \mathcal{D}_f \;\;\&\;\; x_n \to a \;\;\&\;\; f(x_n) \not\to f(a).$$

Tämä on looginen ristiriita, koska oletettiin, että f on jatkuva a:ssa Määritelmän V.1.1 mukaisesti. Tehty vastaoletus on siis väärä, eli f on jatkuva a:ssa Määritelmän V.1.2 kriteerillä.

$\boxed{\Leftarrow}$ Oletetaan, että f on jatkuva pisteessä $a \in \mathcal{D}_f$ Määritelmän V.1.2 kriteerillä. Olkoon $\varepsilon > 0$ ja valitaan $\delta > 0$ siten, että Määritelmän V.1.2 ehto on täytetty. Edelleen olkoon $\{x_n\}$ lukujono, jolle pätee $x_n \in \mathcal{D}_f\ \forall n$ ja $x_n \to a$. Tällöin jostakin indeksistä $n = N$ eteenpäin on $|x_n - a| < \delta$ (koska oli $\delta > 0$), jolloin tehtyjen oletusten mukaan on myös $|f(x_n) - f(a)| < \varepsilon$. On siis löydetty indeksi $N \in \mathbb{N}$ siten, että pätee

$$|f(x_n) - f(a)| < \varepsilon, \quad \text{kun } n > N.$$

Tässä $\varepsilon > 0$ oli mielivaltainen, joten $f(x_n) \to f(a)$ (Määritelmä I.6.1). Myös jono $\{x_n\}$ oli tehtyjen oletusten puitteissa vapaasti valittu, joten jokaiselle tällaiselle jonolle on voimassa $f(x_n) \to f(a)$. Siis f on jatkuva pisteessä a Määritelmän V.1.1 kriteerillä. $\square$

HARJOITUSTEHTÄVIÄ

1. Todista suoraan Määritelmän V.1.1 ja lukujonoja koskevien tulosten perusteella kukin seuraavista funktioista joko jatkuvaksi tai epäjatkuvaksi pisteessä $x = 0$. Hahmottele myös funktioiden kuvaajat joukossa $A = \mathcal{D}_f \cap [-1, 1]$.

   a) $f(x) = x^7 + 7x + 49$ b) $f(x) = x^5 - 6x^3 + \dfrac{x^3 + 20}{100x^2 - 1}$

   c) $f(x) = \begin{cases} \dfrac{1}{x}\left(\dfrac{1}{1+7x} - \dfrac{1}{1+5x}\right), & \text{kun } x \neq 0, \\ 2, & \text{kun } x = 0 \end{cases}$

   d) $f(x) = \begin{cases} x\cos\dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$ e) $f(x) = \begin{cases} \sin\dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$

2. Näytä funktio $f(x) = 1/x$ jatkuvaksi määrittelyjoukossaan vedoten suoraan a) jatkuvuuden jonokriteeriin, b) jatkuvuuden (ε, δ)-kriteeriin.

3. Käyttäen a)- kohdassa jatkuvuuden (ε, δ)-kriteeriä, b)-kohdassa jonokriteeriä ja c)-kohdassa jompaa kumpaa todista:

   a) Propositio V.1.4. b) Lause V.1.6. c) Propositio V.1.8.

4. Sanotaan, että $a \in A$ on joukon $A \subset \mathbb{R}$ *eristetty piste*, jos jollakin $\delta > 0$ on $(a-\delta, a+\delta) \cap A = \{a\}$. Näytä, että reaalifunktio on jatkuva jokaisessa määrittelyjoukkonsa eristetyssä pisteessä.

5. Hae sellaiset parametrin a arvot, joilla funktiosta tulee jatkuva koko $\mathbb{R}$:ssä, ja hahmottele kuvaaja näillä a:n arvoilla:

   a) $f(x) = \begin{cases} x-1, & \text{kun } x \le a, \\ 1-x^2, & \text{kun } x > a \end{cases}$ b) $f(x) = \begin{cases} a-x, & \text{kun } x \le -2, \\ x^3-4x, & \text{kun } x > -2 \end{cases}$

   c) $f(x) = \begin{cases} \dfrac{x^2-4}{x^3-8}, & \text{kun } x < 2, \\ a, & \text{kun } x \ge 2 \end{cases}$ d) $f(x) = \begin{cases} \sin x, & \text{kun } x \le \pi/3, \\ \cos ax, & \text{kun } x > \pi/3 \end{cases}$

   e) $f(x) = \begin{cases} \dfrac{x^3+4x^2+5x+2}{x^2+3x+2}, & \text{kun } x < a, \\ ax-10, & \text{kun } x \ge a \end{cases}$

6. Totea Lauseiden V.1.6–V.1.7 ja Proposition V.1.9 perusteella seuraavat funktiot jatkuviksi koko määrittelyjoukossaan. — Mikä määrittelyjoukko on?

   a) $f(x) = \sqrt{1 + \dfrac{1}{2-x}}$ b) $f(x) = \sqrt[4]{1 - \dfrac{1}{\sqrt{2-x}}}$

7. Funktio $f(x) = \sin x$ on aidosti kasvava välillä $[-\frac{\pi}{2}, \frac{\pi}{2}]$, aidosti vähenevä välillä $[\frac{\pi}{2}, \frac{3\pi}{2}]$ ja jatkuva koko $\mathbb{R}$:ssä. Täsmälleen montako ratkaisua yhtälöllä $\sin x = y \in [-1, 1]$ on näiden tietojen ja Lauseen V.1.10 perusteella välillä $[-\frac{\pi}{2}, \frac{\pi}{6}]$ eri y:n arvoilla?

8. a) Transkendenttisella yhtälöllä $x = \cos x$ on yksikäsitteinen ratkaisu välillä $[0, 1]$. Seuraten Bolzanon lauseen todistuskonstruktiota laske ratkaisulle lähin ala- ja ylälikiarvo äärellisinä binaarilukuina muotoa $p/64$, $p \in \mathbb{N}$.

   b) Montako nollakohtaa on funktiolla $f(x) = x - \cos 49x$? Jos nollakohtaa etsitään väliltä $[0, 1]$ Bolzanon lauseen todistuksessa käytetyllä algoritmilla, niin mikä nollakohdista löytyy?

9. a) Näytä, että polynomin $p(x) = x^3 - 15x + 1$ kaikki juuret ovat reaalisia ja sijaitsevat välillä $[-4, 4]$.

   b) Näytä, että polynomi $f(x) = (x-a)^2(x-b)^2 + x$ saa jossakin pisteessä arvon $(a+b)/2$.

10. Näytä, että jos Bolzanon lauseen todistuskonstruktiossa oletetaan ainoastaan, että f on määritelty välillä $[a, b]$ ja $f(a) < 0,\ f(b) > 0$ (eli ei oleteta jatkuvuutta), niin konstruktio tuottaa tässä tapauksessa pisteen $\xi \in [a, b]$, jossa on joko $f(\xi) = 0$ tai — toinen vaihtoehto? Anna esimerkki funktiosta, jolle konstruktion tulos on $\xi = a$.

11. Todista Lauseen V.1.10 ja Weierstrassin lauseen avulla väittämä: Jos f on jatkuva välillä $[a, b]$ ja $f(x) \neq 0\ \forall x \in [a, b]$, niin $\exists c > 0$ siten, että joko $f(x) \geq c\ \forall x \in [a, b]$ tai $f(x) \leq -c\ \forall x \in [a, b]$.

12. (*) Funktio f on jatkuva koko $\mathbb{R}$:ssä ja

$$f(x) = \frac{1}{x} \cdot \frac{x - 1 + \sqrt{x^2+1}}{x + 1 + \sqrt{x^2+1}}, \quad \text{kun } x \neq 0.$$

   Näytä, että yksinkertaisempi f:n määritelmä on

$$f(x) = \frac{1}{1 + \sqrt{x^2+1}}, \quad x \in \mathbb{R}.$$

13. (*) Olkoon f on jatkuva välillä $A = [a, b]$. Näytä, että
   a) jos $f(A) \subset A$, niin $f(c) = c$ jollakin $c \in A$.
   b) jos $f(a) = f(b)$, niin $f(c) = f(c + \frac{b-a}{2})$ jollakin $c \in [a, \frac{a+b}{2}]$.

14. (*) Olkoon $\{a_k\,,\ k = 0, 1, 2, \ldots\}$ reaalilukujono, jolle pätee $\lim_k a_k = \infty$ ja olkoon

$$H(x) = \begin{cases} 1 - |x|, & \text{kun } |x| \leq 1 \\ 0, & \text{kun } |x| > 1 \end{cases}$$

   Näytä, että funktio

$$f(x) = \sum_{k=0}^{\infty} H(x - a_k)$$

   on jatkuva koko $\mathbb{R}$:ssä. Mikä on f:n maksimiarvo ja missä se saavutetaan, kun $a_k = (5/4)^k$?

V.2 Funktion raja-arvo

Jatkuvuuteen liittyy läheisesti käsite *funktion raja-arvo*. Raja-arvo kertoo funktion arvojen käyttäytymisestä lähestyttäessä jotakin pistettä $a \in \mathbb{R}$ funktion määrittelyjoukosta käsin. Pisteen a ei tarvitse olla määrittelyjoukossa, riittää että sitä voidaan lähestyä mielivaltaisen lähelle. Tyypillinen sovellustilanne onkin juuri tällainen. Toisaalta, jos piste on määrittelyjoukossa, ei raja-arvo silti riipu funktion arvosta tässä pisteessä.

Funktion raja-arvoja on kahdentyyppisiä, varsinaisia eli raja-arvoja ilman lisämääreitä, ja *toispuolisia* raja-arvoja (engl. one-sided limit). Määritelmät ovat seuraavat.

MÄÄRITELMÄ V.2.1 **(Funktion raja-arvo)** Funktiolla $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on *pisteessä* a (tai a:ssa) *raja-arvo* $A \in \mathbb{R}$, jos jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$\big(x_n \in \mathcal{D}_f \;\;\&\;\; x_n \neq a\big)\,\forall n \;\;\&\;\; x_n \to a \;\;\Rightarrow\;\; f(x_n) \to A,$$

ja oletus on voimassa jollekin jonolle $\{x_n\}$. Funktiolla on pisteessä a *vasemmanpuoleinen raja-arvo* $A_- \in \mathbb{R}$, jos jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$\big(x_n \in \mathcal{D}_f \;\;\&\;\; x_n < a\big)\,\forall n \;\;\&\;\; x_n \to a \;\;\Rightarrow\;\; f(x_n) \to A_-,$$

ja oletus on voimassa jollekin jonolle $\{x_n\}$. Funktiolla on pisteessä a *oikeanpuoleinen raja-arvo* $A_+ \in \mathbb{R}$, jos jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$\big(x_n \in \mathcal{D}_f \;\;\&\;\; x_n > a\big)\,\forall n \;\;\&\;\; x_n \to a \;\;\Rightarrow\;\; f(x_n) \to A_+,$$

ja oletus on voimassa jollekin jonolle $\{x_n\}$.

Raja-arvon määritelmän lisäehto 'oletus voimassa jollekin jonolle $\{x_n\}$' tarkoittaa, että lähestyminen ($x_n \to a$) oletetulla tavalla on mahdollista, ts. että a <u>ei</u> <u>ole</u> <u>eristetty</u> <u>piste</u> joukossa $\mathcal{D}_f \cup \{a\}$ (ks. Harj.teht. V.1:4). Toispuolisissa raja-arvoissa lisäehto tarkoittaa vastaavasti, että a ei ole eristetty piste joukossa $[\mathcal{D}_f \cap (-\infty, a)] \cup \{a\}$ (vasemmanpuoleinen raja-arvo) tai joukossa $[\mathcal{D}_f \cap (a, \infty)] \cup \{a\}$ (oikeanpuoleinen raja-arvo). Eristetyssä pisteessä raja-arvoa ei siis määritellä, olipa piste määrittelyjoukossa $\mathcal{D}_f$ tai ei.

Määritelmässä esiintyville raja-arvoille käytetään merkintöjä

$$A = \lim_{x \to a} f(x), \quad A_\pm = \lim_{x \to a^\pm} f(x),$$

tai merkitään (vrt. lukujonon raja-arvomerkinnät)

$$f(x) \to A, \quad \text{kun } x \to a, \qquad f(x) \to A_\pm, \quad \text{kun } x \to a^\pm.$$

Toispuolisille raja-arvoille kätevä merkintätapa on myös

$$A_+ = f(a^+), \quad A_- = f(a^-).$$

Esimerkki 1 Määritä $\lim_{x\to 1} f(x)$, kun $f(x) = (x-1)/(x^2+x-2)$.

Ratkaisu Määrittelyjoukko on $\mathcal{D}_f = \{x \in \mathbb{R} \mid x \neq 1 \ \& \ x \neq -2\}$ ja $f(x) = 1/(x+2)$ $\forall x \in \mathcal{D}_f$. Näin ollen jos $x_n \in \mathcal{D}_f$ $\forall n$ ($\Rightarrow$ $x_n \neq 1$ $\forall n$) ja $x_n \to 1$, niin $f(x_n) = 1/(x_n+2) \to 1/3$. Siis $\lim_{x\to 1} f(x) = 1/3$ (Määritelmä V.2.1). $\square$

Kuten lukujonon, myös funktion raja-arvomerkinnöissä, voi olla A:n tai $A_\pm$:n tilalla $\infty/-\infty$, jolloin tarkoitetaan Määritelmän V.2.1 mukaisesti, että $f(x)$ kasvaa/vähenee rajatta, kun $x \to a$ tai $x \to a^\pm$.

Esimerkki 1 (jatko) Esimerkin funktiolle pätee

$$\lim_{x\to -2^-} f(x) = -\infty, \quad \lim_{x\to -2^+} f(x) = \infty. \quad \square$$

Sikäli kuin raja-arvo $\lim_{x\to a} f(x)$ on olemassa, yhtyvät myös toispuoliset raja-arvot (sikäli kuin määriteltävissä) tähän. Toisaalta on mahdollista, että molemmat toispuoliset raja-arvot ovat olemassa pisteessä a mutta eri suuret, jolloin raja-arvoa pisteessä a ei ole. Mitä tulee jatkuvuuden ja raja-arvon väliseen yhteyteen, nähdään Määritelmistä V.1.1 ja V.2.1, että sikäli kuin raja-arvo $\lim_{x\to a} f(x)$ on määriteltävissä (eli a ei ole eristetty piste joukossa $\mathcal{D}_f \cup \{a\}$), pätee

$$\boxed{f \text{ jatkuva pisteessä } a \in \mathcal{D}_f \ \Leftrightarrow \ \lim_{x\to a} f(x) = f(a) \quad (a \text{ ei eristetty piste}).}$$

Jos f:llä on toispuolinen raja-arvo $f(a^+)$ tai $f(a^-)$ ja pätee joko $f(a) = f(a+)$ tai $f(a) = f(a^-)$, niin sanotaan vastaavasti, että f on *oikealta jatkuva* tai *vasemmalta jatkuva* pisteessä a. Jos $a \in \mathcal{D}_f$ ja pistettä a voidaan lähestyä molemmista suunnista $\mathcal{D}_f$:stä käsin, niin on ilmeistä, että f on jatkuva a:ssa täsmälleen kun f on sekä vasemmalta että oikealta jatkuva a:ssa. Edelleen nähdään, että f on jatkuva suljetulla välillä $[a,b]$ (Määritelmä V.1.5) täsmälleen kun f on jatkuva avoimella välillä (a,b) ja lisäksi oikealta jatkuva a:ssa ja vasemmalta jatkuva b:ssä.

Kuten jatkuvuus, myös raja-arvo on määriteltävissä vaihtoehtoisella (ε,δ)-kriteerillä, vrt. Määritelmä V.1.2. Vaihtoehtoinen määritelmä muotoillaan tässä lauseena raja-arvolle $\lim_{x\to a} f(x) = A \in \mathbb{R}$. Todistus sivuutetaan, sillä se on hyvin samanlainen kuin Lauseen V.1.3 todistus edellisessä luvussa. (Toispuolisille raja-arvoille on muotoiltavissa vastaava tulos, samoin tapauksille $\lim_{x\to a} f(x) = \pm\infty$.)

*Lause V.2.2 **(Raja-arvon (ε, δ)-kriteeri)** Funktiolla $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on raja-arvo $\lim_{x\to a} f(x) = A \in \mathbb{R}$ täsmälleen kun $\mathcal{D}_f \cap [(a-\delta, a) \cup (a, a+\delta)] \neq \emptyset$ $\forall \delta > 0$ ja jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että jokaisella $x \in \mathbb{R}$ pätee

$$x \in \mathcal{D}_f \;\;\&\;\; 0 < |x-a| < \delta \;\Rightarrow\; |f(x) - A| < \varepsilon.$$

Lukujonoista tiedetään, että suppeneva lukujono on rajoitettu (Lause I.7.2). Funktiolle, jolla on raja-arvo, pätee Lauseen V.2.2 perusteella vastaava tulos:

Lause V.2.3 Jos $\lim_{x\to a} f(x) = A \in \mathbb{R}$, niin f on *rajoitettu* jossakin pisteen ympäristössä, ts. $\exists \delta > 0$ ja $C \in \mathbb{R}_+$ siten, että

$$|f(x)| \leq C \quad \forall\, x \in (a-\delta, a+\delta) \cap \mathcal{D}_f.$$

Todistus Valitaan $\varepsilon = 1$ ja vastaava $\delta > 0$ niin, että Lauseen V.2.2 ehto on voimassa. Tällöin seuraa kolmioepäyhtälöstä, että väittämä on tosi, kun $C = |A| + 1$. $\square$

Funktion approksimointi raja-arvolla

Jos funktiosta tunnetaan raja-arvo $A = \lim_{x\to a} f(x)$, niin Lauseen V.2.2 mukaan voidaan raja-arvopisteen lähellä käyttää approksimaatiota $f(x) \approx A$. Lause ei tosin anna (eikä tehdyin oletuksin voikaan antaa) mitään kvantitatiivista tietoa approksimaation tarkkuudesta, koska δ:n riippuvuutta ε:sta ei tunneta. Tyypillisissä esimerkkitapauksissa funktiosta f kuitenkin yleensä tiedetään lauseessa oletettua enemmän, jolloin approksimaatiolle on ehkä mahdollista johtaa kvantitatiivinen virhearvio tämän lisätiedon perusteella. Likimääräisessä funktioevaluaatiossa raja-arvotieto voi auttaa etenkin silloin, kun funktion laskukaava suoraan käytettynä on altis numeerisille pyöristysvirheille raja-arvopisteen lähellä.

Esimerkki 2 Funktio $f(x) = (1-\cos x)/x^2$ on määritelty, kun $x \neq 0$. Myöhemmin (Luku V.4) osoitetaan, että $\lim_{x\to 0} f(x) = 1/2$. Tähän tulokseen perustuva approksimaatio

$$f(0.000000003) \approx \lim_{x\to 0} f(x) = 0.5$$

on huomattavasti turvallisempi kuin f:n laskukaavan suora käyttö, sillä laskuoperaatiossa $x \mapsto 1 - \cos x$ tapahtuu huomattava merkitsevien numeroiden kato, kun $|x|$ on pieni (vrt. Luku I.5). Paitsi turvallinen, raja-arvoon perustuva approksimaatio on tässä tapauksessa myös hyvin tarkka: virhe on alle 10^{-18}. $\square$

Raja-arvojen yhdistely

Raja-arvojen laskemista helpottavat seuraavat lauseet, jotka ovat Lauseiden V.1.6 ja V.1.7 vastineita. Todistukset ovat Määritelmän V.2.1 perusteella suoraviivaisia (Harj.teht. 1; ks. myös Harj.teht. V.1:4). Lauseet pätevät ilmeisin muutoksin myös toispuolisille raja-arvoille.

LAUSE V.2.4 Jos $\lim_{x\to a} f(x) = A \in \mathbb{R}$, niin

$$\lim_{x\to a}(\lambda f)(x) = \lambda A \quad \forall \lambda \in \mathbb{R}.$$

Jos lisäksi $\lim_{x\to a} g(x) = B \in \mathbb{R}$ ja a ei ole joukon $(\mathcal{D}_f \cap \mathcal{D}_g) \cup \{a\}$ eristetty piste, niin

$$\lim_{x\to a}(f+g)(x) = A + B,$$
$$\lim_{x\to a}(fg)(x) = AB.$$

Jos lisäksi $B \neq 0$ ja a ei ole joukon $\mathcal{D}_{f/g} \cup \{a\} = \{x \in \mathcal{D}_f \cap \mathcal{D}_g \mid g(x) \neq 0\} \cup \{a\}$ eristetty piste, niin

$$\lim_{x\to a}(f/g)(x) = A/B.$$

LAUSE V.2.5 Jos $\lim_{x\to a} g(x) = A \in \mathbb{R}$, f on jatkuva pisteessä $x = A$ ja a ei ole joukon $\mathcal{D}_{f\circ g} \cup \{a\}$ eristetty piste, niin $\lim_{x\to a}(f \circ g)(x) = f(A)$.

ESIMERKKI 2 (jatko) Koska $g(x) = \sqrt{|x|}$ on jatkuva $\mathbb{R}$:ssä (Propositio V.1.9), niin esimerkin raja-arvotuloksesta ja Lauseesta V.2.5 seuraa

$$\lim_{x\to 0} \frac{\sqrt{1-\cos x}}{|x|} = \lim_{x\to 0} \sqrt{\frac{1-\cos x}{x^2}} = \frac{1}{\sqrt{2}}. \quad \square$$

ESIMERKKI 3 Funktioille $f_0(x) = 1$ ja $f_1(x) = x$ ovat voimassa ilmeiset raja-arvotulokset

$$\lim_{x\to a} f_0(x) = 1, \quad \lim_{x\to a} f_1(x) = a, \quad a \in \mathbb{R}.$$

Näiden ja Lauseen V.2.4 perusteella seuraa

$$\lim_{x\to a} \frac{(x+1)^2}{x+2} = \frac{(a+1)^2}{a+2}, \quad \text{kun } a \neq -2.$$

Tulos on selvä myös Lauseen V.1.6 perusteella, sillä tämän mukaan rationaalifunktio f on jatkuva koko määrittelyjoukossaan, jolloin $\lim_{x\to a} f(x) = f(a)$, kun $a \in \mathcal{D}_f$. $\square$

Raja-arvon laskeminen sijoituksella

Raja-arvon $\lim_{x\to a} f(x)$ laskemista on usein mahdollista helpottaa tekemällä *sijoitus* eli *muuttujan vaihto* $x = u(t)$. Muuttujaa vaihdettaessa tukeudutaan seuraavaan lauseeseen, joka on helposti muotoiltavissa myös toispuolista raja-arvoa koskevaksi (Harj.teht. 1c).

Lause V.2.6 Olkoon $u: \ U \to [a-\delta, a+\delta]$ $(\delta > 0)$ jatkuva bijektio, missä U on suljettu väli. Tällöin jos $u(\alpha) = a$, niin pätee

$$\lim_{x\to a} f(x) = \lim_{t\to\alpha} f(u(t)) = \lim_{t\to\alpha} F(t) = A,$$

sikäli kuin raja-arvo oikealla on olemassa ($A \in \mathbb{R}$ tai $A = \pm\infty$).

Todistus Lauseen V.1.14 perusteella myös käänteisfunktio $u^{-1}: \ [a-\delta, a+\delta] \to U$ on jatkuva bijektio. Olkoon nyt $x_n \in \mathcal{D}_f$, $x_n \neq a \ \forall n$ ja $x_n \to a$. Tällöin jostakin indeksistä eteenpäin on $|x_n - a| < \delta$, jolloin voidaan kirjoittaa $x_n = u(t_n)$, $t_n \in U$. Koska $u: \ U \to [a-\delta, a+\delta]$ on injektio, niin $x_n \neq a \ \Rightarrow \ t_n \neq \alpha$. Koska u^{-1} on jatkuva pisteessä a, niin $x_n \to a \ \Rightarrow \ t_n = u^{-1}(x_n) \to u^{-1}(a) = \alpha$. Tällöin oletuksen $\lim_{t\to\alpha} F(t) = A$ perusteella pätee $f(x_n) = f(u(t_n)) = F(t_n) \to A$. On näytetty, että $\lim_{x\to a} f(x) = A$ (Määritelmä V.2.1, kun $A \in \mathbb{R}$; päättely toimii myös, kun $A = \pm\infty$). $\square$

Esimerkki 4 Määritä raja-arvo $A = \lim\limits_{x\to 81} \dfrac{\sqrt{x}-9}{\sqrt[4]{x}-3}$.

Ratkaisu Tässä sopiva sijoitus on $\sqrt[4]{x} = t$, jolloin on $x = t^4 = u(t)$, ja arvoa $x = 81\,(= a)$ vastaa $t = 3\,(= \alpha)$. Koska

$$F(t) = f(t^4) = \frac{t^2-9}{t-3} = t+3 = G(t), \quad \text{kun } t \neq 3,$$

niin funktion G jatkuvuuden perusteella on $\lim_{t\to 3} F(t) = G(3) = 6$. Lauseen V.2.6 oletukset ovat voimassa ($\delta \leq 3$), joten kysytty raja-arvo on $A = 6$. $\square$

Esimerkki 5 Sijoituksella $x = 2t$ saadaan (vrt. Luku II.5)

$$\lim_{x\to 0^+} \frac{\sin\frac{x}{2}}{\sqrt{1-\cos x}} = \lim_{t\to 0^+} \frac{\sin t}{\sqrt{1-\cos 2t}} = \lim_{t\to 0^+} \frac{\sin t}{\sqrt{2\sin^2 t}} = \lim_{t\to 0^+} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

Tässä loppusievennys perustui päättelyyn $t \in (0, \pi] \ \Rightarrow \ \sqrt{\sin^2 t} = \sin t$. $\square$

Esimerkeissä suoritettiin raja-arvolaskuille hyvin tyypillinen nelivaiheinen laskuoperaatio

$$\lim_{x\to a} f(x) = \lim_{t\to\alpha} f(u(t)) = \lim_{t\to\alpha} F(t) = \lim_{t\to\alpha} G(t) = G(\alpha).$$

Tässä

1. Sijoitetaan $x = u(t)$ [tai $v(x) = t \Rightarrow x = u(t)$] ja lasketaan $\alpha = u^{-1}(a)$.
2. Sievennetään $f(u(t))$ lausekkeeksi $F(t)$.
3. Pelkistetään $F(t)$ lausekkeeksi $G(t)$, kun $t \neq \alpha$.
4. Lasketaan raja-arvo vedoten G:n <u>jatkuvuuteen</u> pisteessä $t = \alpha$.

Sikäli kuin muuttujaa ei vaihdeta, lasku supistuu vaiheiksi 3–4 ($F = f$, $t = x$), kuten Esimerkissä 1 edellä.

Funktion jatkaminen

Jos $a \in \mathbb{R}$ ei ole funktion f määrittelyjoukossa, mutta on olemassa (aito) raja-arvo $\lim_{x\to a} f(x) = A \in \mathbb{R}$, niin on luonnollista sisällyttää a määrittelyjoukkoon asettamalla $f(a) = A$. Näin menetellen f:stä tulee Määritelmän V.1.1 mukaisesti jatkuva pisteessä a. Funktion määrittelyjoukon laajentamista tällä tavoin kutsutaan *funktion jatkamiseksi*.

Esimerkki 6 Esimerkissä 1 todettiin, että funktiolla $f(x) = (x-1)/(x^2+x-2)$ on raja-arvo $\lim_{x\to 1} f(x) = 1/3$. Kun piste $x = 1$ sisällytetään f:n määrittelyjoukkoon asettamalla $f(1) = 1/3$, niin f tulee jatketuksi funktioksi $f(x) = 1/(x+2)$. Enempää ei määrittelyjoukkoa voi tässä laajentaa jatkamalla, koska f:llä ei ole reaalista (tai yleisempääkään) raja-arvoa, kun $x \to -2$. □

Esimerkki 7 Esimerkin 2 raja-arvotiedon (ja Esimerkin V.1:6 tiedon) perusteella funktio

$$f(x) = \begin{cases} \dfrac{1-\cos x}{x^2}, & \text{kun } x \neq 0, \\ \dfrac{1}{2}, & \text{kun } x = 0 \end{cases}$$

on jatkuva $\mathbb{R}$:ssä. Funktion arvo 0:ssa on määrätty jatkamalla. □

Paloittainen jatkuvuus

Funktion toispuoliset raja-arvot tulevat käyttöön erityisesti sellaisissa sovellustilanteissa, joissa (usein fysikaalista perua oleva) funktio on jatkuva muualla paitsi erillisissä pisteissä, joissa sillä on yksinkertainen nk. *hyppyepäjatkuvuus* (engl. jump discontinuity). Asetetaan tällaisia käytännön tarpeita silmällä pitäen

MÄÄRITELMÄ V.2.7 Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on välillä $[a, b]$ *paloittain jatkuva* (engl. piecewise continuous), jos $\exists$ pisteet c_k, $k = 0 \ldots n$, $n \in \mathbb{N}$ siten, että

$$a = c_0 < c_1 < \ldots < c_n = b$$

ja pätee

(i) $(c_{k-1}, c_k) \subset \mathcal{D}_f$ ja f on jatkuva välillä (c_{k-1}, c_k), $\quad k = 1 \ldots n$,

(ii) $\exists$ toispuoliset raja-arvot

$$\lim_{x \to c_k^+} f(x) = f(c_k^+), \quad k = 0 \ldots n-1,$$
$$\lim_{x \to c_k^-} f(x) = f(c_k^-), \quad k = 1 \ldots n.$$

Huomattakoon, että raja-arvot $f(c_k^\pm)$ eivät raja-arvon määritelmän mukaisesti riipu f:n mahdollisista arvoista pisteissä c_k. Jos molemmat toispuoliset raja-arvot ovat olemassa, mutta eivät yhdy, on kyseessä hyppyepäjatkuvuus (hyppy$= f(c_k^+) - f(c_k^-)$). Jos yhtyvät, on funktio pisteessä c_k joko jatkuva tai määriteltävissä jatkuvaksi raja-arvon avulla (jatkamismenettely).

ESIMERKKI 8 Olkoon $[x]$ = suurin kokonaisluku, jolle pätee $[x] \leq x$. Funktio $f(x) = x - [x]$, eli

$$f(x) = x - k, \quad \text{kun } x \in [k, k+1), \quad k \in \mathbb{Z},$$

on jokaisella välillä $[a, b] \subset \mathbb{R}$ paloittain jatkuva. Pisteissä $k \in \mathbb{Z}$ funktio voitaisiin määritellä miten tahansa (tai jättää määrittelemättä) ilman, että sillä vaikutettaisiin toispuolisiin raja-arvoihin $f(k^+) = 0$ ja $f(k^-) = 1$. $\square$

Esimerkki 9 **Kiihdytys**. Auto on paikallaan moottorin käydessä. Ajan hetkellä $t = 0$ polkaistaan kaasupoljin pohjaan. Kiihtyvyys $f(t) = ?$

Ratkaisu Idealisoidun matemaattisen mallin mukaan on

$$f(t) = \begin{cases} 0, & \text{kun } t < 0, \\ a = \text{vakio}, & \text{kun } t \geq 0. \end{cases}$$

Todellisuudessa hyppyepäjatkuvuutta ei hetkellä $t = 0$ esiinny. Matemaattinen malli (jos hyvä) riippuukin esimerkissä olennaisesti siitä, millaisessa aikaskaalassa tapahtumia tarkastellaan. Eri skaaloissa funktio f voi näyttää hyvin erilaiselta. □

Raja-arvot $\lim\limits_{x\to\pm\infty} f(x)$

Raja-arvo $\lim_{x\to\infty} f(x)$ määritellään samoin kuin edellä, eli sijoittamalla yleiseen määritelmään $a = \infty$. Tällöin $x_n \to \infty$ tarkoittaa, että $\{x_n\}$ kasvaa rajatta. Raja-arvo $\lim_{x\to-\infty} f(x)$ määritellään vastaavasti. Lause V.2.4 on pätevä myös raja-arvoille $\lim_{x\to\pm\infty} f(x)$, samoin Lause V.2.5, kun määrittelyjoukkoa $\mathcal{D}_{f\circ g}$ koskeva oletus muutetaan joko ehdoksi $\mathcal{D}_{f\circ g} \cap (M, \infty) \neq \emptyset \ \forall M \in \mathbb{R}$ (jos $x \to \infty$) tai ehdoksi $\mathcal{D}_{f\circ g} \cap (-\infty, M) \neq \emptyset \ \forall M \in \mathbb{R}$ (jos $x \to -\infty$).

Raja-arvoja $\lim_{x\to\infty} f(x)$ määrättäessä usein kätevä on sijoitus $x = t^{-1}$, eli siirtyminen tarkastelemaan funktiota $F(t) = f(1/t)$. Koska $x \to \pm\infty \ \Leftrightarrow \ t \to 0^{\pm}$, niin raja-arvojen määritelmistä nähdään, että pätee (vrt. Lause V.2.6)

$$\lim_{x\to\pm\infty} f(x) = \lim_{t\to 0^{\pm}} F(t) = F(0^{\pm}), \quad F(t) = f(1/t).$$

Esimerkki 10 (jatko) Muuttujan vaihdolla $x = t^{-1}$ päätellään Propostitioon V.1.9 ja Lauseisiin V.1.6 ja V.1.7 vedoten:

$$\lim_{x\to\infty} \frac{\sqrt[3]{x} + \sqrt[4]{x}}{\sqrt[3]{8x+3}} = \lim_{t\to 0^+} \frac{1 + \sqrt[12]{t}}{\sqrt[3]{8+3t}} = \frac{1+0}{\sqrt[3]{8+0}} = \frac{1}{2}. \quad \square$$

Lukujonojen teoriasta tiedetään, että monotoninen ja rajoitettu lukujono suppenee. Raja-arvoja $\lim_{x\to\infty} f(x)$ koskeva vastaava väittämä on

LAUSE V.2.8 Jos funktio f on monotoninen ja rajoitettu välillä $[a,\infty)$, niin on olemassa raja-arvo $\lim_{x\to\infty} f(x) \in \mathbb{R}$.

Todistus Olkoon f esim. kasvava välillä $[a,\infty)$. Tällöin jos merkitään $y_k = f(k)$, $k \in \mathbb{N}$ ja $k \geq a$, niin $\{y_k\}$ on kasvava ja rajoitettu lukujono, joten $y_k \to y \in \mathbb{R}$. Olkoon nyt $\{x_n\}$ mikä tahansa lukujono, jolle pätee $x_n \geq a\ \forall n$ ja $x_n \to \infty$. Tällöin jos $x_n \in [k, k+1)$, $k \in \mathbb{N}$, niin $y_k \leq f(x_n) \leq y_{k+1}$, koska f on kasvava. Tässä $k \to \infty$ kun $n \to \infty$ (koska $x_n \to \infty$), joten päätellään (Propositio I.7.4[V2]), että $\lim_n f(x_n) = \lim_k y_k = y$. Koska tämä pätee jokaiselle mainitut ehdot täyttävälle jonolle $\{x_n\}$, niin $\lim_{x\to\infty} f(x) = y$. Jos f on vähenevä välillä $[a,\infty)$, niin päättely on vastaava. $\square$

Asymptootit

Sanotaan, että funktio $g(x)$ on funktion $f(x)$ *asymptootti*, jos

$$\lim_{x\to\infty}[f(x) - g(x)] = 0, \quad \text{tai} \quad \lim_{x\to-\infty}[f(x) - g(x)] = 0.$$

Asymptootin ideana on approksimoida funktiota jollakin (mieluiten) yksinkertaisella funktiolla $g(x)$, kun joko x on hyvin suuri (merkintä $x \gg 1$) tai $-x$ on hyvin suuri (merkintä $x \ll -1$):

$$f(x) \approx g(x), \quad \text{kun } x \gg 1 \text{ tai } x \ll -1.^{\dagger}$$

Kyse on Lauseesta V.2.2, joka on yleistettävissä myös raja-arvoja $\lim_{x\to\pm\infty} f(x)$ koskevaksi: Esimerkiksi jos $\mathcal{D}_f \cap \mathcal{D}_g \supset [a,\infty)$ ja $\lim_{x\to\infty}[f(x) - g(x)] = 0$, niin jokaisella $\varepsilon > 0$ on olemassa $M \in [a,\infty)$ siten, että $|f(x) - g(x)| < \varepsilon\ \forall x > M$.

ESIMERKKI 11 Funktion $f(x) = \sqrt{x^2+4x}$ eräs asymptootti, kun $|x| \gg 1$, on $g(x) = |x+2|$, sillä rajoilla $x \to \pm\infty$ on

$$\begin{aligned} f(x) - g(x) &= \sqrt{x^2+4x} - |x+2| = \frac{x^2+4x-(x+2)^2}{\sqrt{x^2+4x}+|x+2|} \\ &= -\frac{4}{\sqrt{x^2+4x}+|x+2|} \to 0. \quad \square \end{aligned}$$

†Asymptootin perinteisempi ja rajoitetumpi geometrinen määritelmä on tasokäyrään liittyen suora, jota 'käyrä lähestyy äärettömyydessä'. Esimerkiksi käyrän $S : y = x^2/(x+1)$ asymptootteja ovat tämän tulkinnan mukaan suorat $y = x - 1$ ja $x = -1$. Jälkimmäinen on nk. pystysuora asymptootti, jolla ei ole funktiovastinetta.

HARJOITUSTEHTÄVIÄ

1. a) Todista Lause V.2.4. b) Todista Lause V.2.5. c) Muotoile ja todista Lauseen V.2.6 vastine koskien toispuolista raja-arvoa $\lim_{x\to a^+} f(x)$.

2. Funktiosta f tiedetään, että $[-1,1] \subset \mathcal{D}_f$, $f(0) = 0$ ja $\sqrt{2-x^2} \le f(x) \le \sqrt{2+9|x|}$, kun $0 < |x| \le 1$. Näytä, että $\lim_{x\to 0} f(x) = \sqrt{2}$.

3. a) Funktiosta f tiedetään, että $\lim_{x\to 0^+} f(x) = A$. Näytä, että jos f on parillinen, niin $\lim_{x\to 0^-} f(x) = A$, ja jos pariton, niin $\lim_{x\to 0^-} f(x) = -A$. b) Funktiosta f tiedetään, että $\lim_{x\to 0^+} f(x) = A$ ja $\lim_{x\to 0^-} f(x) = B$. Laske raja-arvot $\lim_{x\to 0^+} f(x^2 - x)$ ja $\lim_{x\to 0^-} f(x^2 + x^3)$.

4. Määritä seuraavat raja-arvot, joko reaalilukuna tai muodossa $\pm\infty$, tai totea vaihtoehtoisesti, ettei raja-arvoa ole. Vaihda tarvittaessa muuttujaa.

a) $\displaystyle\lim_{x\to 4}(x^2 - 4x + 1)$ b) $\displaystyle\lim_{x\to 3}\frac{x+3}{x+6}$ c) $\displaystyle\lim_{x\to \pi}\frac{(x+\pi)^2}{\pi x}$

d) $\displaystyle\lim_{x\to -2}\frac{x^2+2x}{x^2-4}$ e) $\displaystyle\lim_{x\to 2}\left(\frac{1}{x-2} - \frac{4}{x^2-4}\right)$ f) $\displaystyle\lim_{x\to \frac{\pi}{2}}\frac{\sin(2x+\pi)}{\cot x}$

g) $\displaystyle\lim_{t\to 0}\frac{(t+1)^2-(t-1)^2}{t}$ h) $\displaystyle\lim_{s\to 0}\frac{s^2+3s}{(s+2)^2-(s-2)^2}$

i) $\displaystyle\lim_{x\to -3}|x-3|$ j) $\displaystyle\lim_{x\to 2}\frac{|x^2-4x+3|}{x^2+2x-3}$ k) $\displaystyle\lim_{x\to 1}\frac{|x^2-4x+3|}{x^2+2x-3}$

l) $\displaystyle\lim_{x\to 0}\frac{\sqrt{1+2x+3x^2}-\sqrt{1-x}}{x}$ m) $\displaystyle\lim_{y\to 1}\frac{y-4\sqrt{y}+3}{y^2-1}$

n) $\displaystyle\lim_{x\to 0}\frac{\sqrt{2-x}-\sqrt{2+x}}{x\sqrt{x}}$ o) $\displaystyle\lim_{x\to 2}\frac{\sqrt{4-4x+x^2}}{x-2}$ p) $\displaystyle\lim_{t\to 8}\frac{t^{2/3}-4}{t^{1/3}-2}$

q) $\displaystyle\lim_{x\to 3^+}\frac{|x-3|}{3-x}$ r) $\displaystyle\lim_{x\to 2^-}\frac{\sqrt{4-4x+x^2}}{x-2}$ s) $\displaystyle\lim_{x\to \pi^+}\frac{\sqrt{1+\cos x}}{\cos\frac{x}{2}}$

t) $\displaystyle\lim_{x\to -0.4^+}\frac{2x+5}{5x+2}$ u) $\displaystyle\lim_{x\to 1^+}\frac{x}{\sqrt{x^2-1}}$ v) $\displaystyle\lim_{x\to 1^+}\frac{\sqrt{x^2-x}}{x-x^2}$

x) $\displaystyle\lim_{x\to \infty}\frac{1+x+x^4}{2+30x+200x^3}$ y) $\displaystyle\lim_{x\to \infty}\left(\frac{x^2}{x+1} - \frac{x^2}{x-1}\right)$

z) $\displaystyle\lim_{x\to -\infty}\frac{2x-1}{\sqrt{3x^2+x+1}}$ å) $\displaystyle\lim_{x\to \infty}\frac{x\sqrt{x+1}\,(1-\sqrt{2x+3})}{7-6x+4x^2}$

ä) $\displaystyle\lim_{x\to \infty}\left(\sqrt{x^2+9x}-\sqrt{x^2-5x}\right)$ ö) $\displaystyle\lim_{x\to -\infty}\left(\sqrt{x^2-4x}-\sqrt{x^2+10x}\right)$

5. Todista Määritelmän V.2.1 ja lukujonojen teorian avulla:
   a) Jos f on monotoninen ja rajoitettu välillä $(a,b) \subset \mathcal{D}_f$, niin on olemassa raja-arvot $f(a^+)$ ja $f(b^-)$.
   b) Jos $f(x) \leq g(x)\ \forall x \in (a,b) \subset \mathcal{D}_f \cap \mathcal{D}_g$, niin $f(a^+) \leq g(a^+)$ ja $f(b^-) \leq g(b^-)$ sikäli kuin raja-arvot ovat olemassa reaalilukuina.

6. Laajenna funktion $f(x) = \dfrac{x^2-1}{\sqrt{x+3}-2}$ määrittelyjoukko väliksi $[-3,\infty)$ jatkamalla. Mikä on jatketun funktion sievennetty laskusääntö?

7. Olkoon f määritelty välillä $[1,\infty)$ ja olkoon $y_n = f(n)$, $n \in \mathbb{N}$. Näytä, että pätee $\lim_{x\to\infty} f(x) = A \ \Rightarrow\ \lim_n y_n = A$. Näytä vastaesimerkillä, että käänteinen implikaatio ei ole tosi.

8. Määritä seuraaville funktioille $f(x)$ tai $y(x)$ asymptootti $g(x)$, joka on annettua muotoa ($a, b \in \mathbb{R}$) ja mahdollisimman tarkka (ellei yksikäsitteinen). Tarkastele erikseen tapaukset $x \to \infty$ ja $x \to -\infty$.
   a) $f(x) = \dfrac{|x+3|}{2x-1}, \quad g(x) = a$
   b) $f(x) = \dfrac{(x+3)^2}{|3x+1|}, \quad g(x) = ax+b$
   c) $f(x) = \sqrt{2x^2+3x+\cos x}, \quad g(x) = ax+b$
   d) $f(x) = \dfrac{1}{\sqrt{3x^2+4x}+\sqrt{x^2+x+1}}, \quad g(x) = \dfrac{a}{x}$
   e) $f(x) = \dfrac{|3x+2|}{2x+3}, \quad g(x) = a + \dfrac{b}{x}$
   f) $9(x-1)^2 - 16(y+2)^2 = 25,\ y(x) \geq -2, \quad g(x) = ax+b$
   g) $(2x^2+1)y + x\cos y = (2x+1)^3, \quad g(x) = ax+b$

9. Funktiolla $f(x) = \sqrt{x^4+8x^3+35x^2+78x+98}$ on asymptootteina toisen asteen polynomeja. Koeta löytää mahdollisimman tarkka tällainen asymptootti $p(x)$ ja arvioi approksimaation $f(x) \approx p(x)$ virhe, kun $|x| \geq 100$.

10. (*) Olkoon $a > 0$ ja $m, n \in \mathbb{N}$. Laske raja-arvo $\displaystyle\lim_{x\to a} \frac{\sqrt[m]{x} - \sqrt[m]{a}}{\sqrt[n]{x} - \sqrt[n]{a}}$.

11. (*) Reaalifunktiosta f tiedetään, että $f(x) = x+1\ \forall x \in \mathbb{R},\ x \notin X$, missä $X \subset \mathbb{R}$ on äärellinen joukko. Joukkoa X ei tunneta, eikä pisteistä $x \in X$ tiedetä muuta kuin että f on näissä pisteissä joko määritelty jollakin tuntemattomilla tavalla (reaaliarvoisena) tai jätetty määrittelemättä. Näytä, että $\lim_{x\to 0} f(x) = 1$.

V.3 Derivaatta

Tarkastellaan funktion f approksimoimista pisteen $a \in \mathcal{D}_f$ lähellä perustuen erilaisiin olettamuksiin funktion ennustettavuudesta. Ensinnäkin jos f jatkuva pisteessä a, niin Määritelmän V.1.2 mukaan f on a:n lähellä likimain vakio: $f(x) \approx f(a)$ kun $x \approx a$. Tämän approksimaation virhe on a:n lähellä tyypillisesti muotoa $f(x) - f(a) \approx L(x-a)$, missä L on f:stä ja a:sta riippuva vakio. Erikoistapauksia (kuten $f(x) =$ vakio tai $f(x) = x^2,\ a = 0$) lukuunottamatta mainittu virhearvio on yleisesti paras mahdollinen.

Esimerkki 1 Jos $f(x) = x^3$ ja $a \neq 0$, niin

$$f(x) - f(a) \;=\; x^3 - a^3 \;=\; (x-a)(x^2 + ax + a^2) \;=\; (x-a)g(x).$$

Tässä $g(x) = x^2 + ax + a^2$ on (polynomina) jatkuva, joten pisteen $x = a$ lähellä on $f(x) - f(a) \approx L(x-a)$, missä $L = g(a) = 3a^2$. $\square$

Jos jatkuvalle funktiolle halutaan pisteen a lähellä olennaisesti tarkempi approksimaatio kuin $f(x) \approx f(a)$, on approksimaation oltava jotakin muuta tyyppiä kuin $f(x) \approx A =$ vakio, sillä näistä vaihtoehdoista valinta $A = f(a)$ on paras. Koska vakio = polynomi astetta $n = 0$, niin luonteva seuraava yritys on käyttää approksimaatiossa polynomia p astetta $n = 1$. Kun ennakkoehdoksi asetetaan $p(a) = f(a)$, niin approksimaatio saa muodon

$$f(x) \approx f(a) + k(x-a).$$

Tässä kerroin $k \in \mathbb{R}$ määrätään (jos mahdollista) niin, että virheelle $g(x) = f(x) - f(a) - k(x-a)$ saadaan a:n lähellä olennaisesti parempi arvio kuin $|g(x)| \leq L|x-a|$. Asetetaan minimiehdoksi

$$\lim_{x\to a} \frac{g(x)}{x-a} = 0 \;\Leftrightarrow\; \lim_{x\to a}\left[\frac{f(x)-f(a)}{x-a} - k\right] = 0 \;\Leftrightarrow\; \lim_{x\to a}\frac{f(x)-f(a)}{x-a} = k.$$

Riippuen funktiosta f kerroin k siis joko määräytyy tästä ehdosta yksikäsitteisesti (raja-arvo oikealla olemassa reaalilukuna) tai sitten kerrointa ei voi määrätä, jolloin ym. approksimaatioyritys katsotaan epäonnistuneeksi.

Esimerkki 1 (jatko) Esimerkin funktiolle pätee

$$\begin{aligned} f(x) &= f(a) + (x-a)(x^2 + ax + a^2) \\ &= f(a) + 3a^2(x-a) + (x-a)(x^2 + ax - 2a^2) \\ &= f(a) + 3a^2(x-a) + (x-a)^2(x+2a). \end{aligned}$$

Siis jos valitaan $k = 3a^2$, niin $g(x)/(x-a) = (x-a)(x+2a) \to 0$ kun $x \to a$, joten approksimaatio onnistui. $\square$

MÄÄRITELMÄ V.3.1 **(Derivaatta)** Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on pisteessä $a \in \mathcal{D}_f$ *derivoituva* (engl. differentiable), jos $(a-\delta, a+\delta) \subset \mathcal{D}_f$ jollakin $\delta > 0$ ja on olemassa raja-arvo

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = k \in \mathbb{R}.$$

Lukua k kutsutaan f:n *derivaataksi* (engl. derivative) pisteessä a, ja merkitään $k = f'(a)$. Jos f on pisteessä a derivoituva, niin polynomia

$$p(x) = f(a) + f'(a)(x - a)$$

sanotaan f:n *linearisoivaksi approksimaatioksi* eli *linearisaatioksi* pisteessä a.

Linearisoivan approksimaation geometrinen vastine on pisteen $P = (a, f(a))$ kautta kulkeva suora, jonka *kulmakerroin* on $k = f'(a)$ $(= p'(a))$. Sanotaan, että ko. suora on käyrän $S : y = f(x)$ *tangentti* pisteessä P. Tämän mukaisesti siis 'derivaatta on tangentin kulmakerroin'.

ESIMERKKI 1 (jatko) Esimerkin tuloksen perusteella funktio $f(x) = x^3$ on derivoituva jokaisessa pisteessä $a \in \mathbb{R}$ ja $f'(a) = 3a^2$. Käyrän $S : y = x^3$ tangentin yhtälö pisteessä $x = a$ on siis $y = p(x) = a^3 + 3a^2(x - a)$. $\square$

ESIMERKKI 2 Funktio $f(x) = |x - a|$ ei ole derivoituva pisteessä a, sillä

$$\frac{f(x) - f(a)}{x - a} = \begin{cases} 1, & \text{kun } x > a, \\ -1, & \text{kun } x < a. \end{cases} \quad \square$$

Jos f on derivoituva pisteessä a, niin Määritelmän V.3.1 ja Lauseen V.2.4 perusteelle pätee

$$\lim_{x\to a} f(x) \;=\; \lim_{x\to a}\left[f(a) + (x-a)\,\frac{f(x)-f(a)}{x-a}\right] \;=\; f(a) + 0\cdot f'(a) = f(a).$$

Siis derivoituvuus on jatkuvuutta vahvempi ominaisuus:

$$\boxed{f \text{ derivoituva pisteessä } a \;\Rightarrow\; f \text{ jatkuva pisteessä } a.}$$

Derivaatta operaattorina

Kun derivaatan määritelmässä kirjoitetaan $x - a = \Delta x$ ja asetetaan a:n tilalle x, niin määritelmä saa muodon

$$f'(x) = \lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}\,.$$

Derivaatan muita merkintöjä ovat

$$\boxed{f'(x) = \frac{df}{dx} = \frac{d}{dx}f = \mathrm{D}f(x).}$$

Kahdessa viimeksi kirjoitetussa merkinnässä tulkitaan derivaatta kohteestaan erillisenä *operaattorina* eli 'funktion funktiona':

$$\mathrm{D} : f \mapsto f', \quad \mathrm{D} = \frac{d}{dx}\,.$$

Operaattoreiksi sanotaan yleisemmin sellaisia funktioita, joiden sekä määrittely- että maalijoukkona ovat (esim. reaalimuuttujan) funktiot. Derivoinnin suorittavaa 'funktion funktiota' D sanotaan *differentiaalioperaattoriksi.*

Differentiaalioperaattori voi määritelmänsä mukaisesti toimia vain sellaisissa pisteissä, joissa f on derivoituva, ts. f':n määrittelyjoukossa

$$\mathcal{D}_{f'} = \{x \in \mathcal{D}_f \mid f \text{ derivoituva pisteessä } x\}.$$

Useat tavalliset funktiot ovat derivoituvia 'melkein kaikkialla' niin, että $\mathcal{D}_{f'}$ saadaan $\mathcal{D}_f$:stä poistamalla enintään äärellinen tai numeroituva määrä pisteitä.

Esimerkki 3 Funktio $f(x) = x - [x]$, missä $[x]$ = suurin kokonaisluku, jolle pätee $[x] \le x$, on derivoituva muualla paitsi pisteissä $k \in \mathbb{Z}$, eli $D_{f'} = \{x \in \mathbb{R} \mid x \notin \mathbb{Z}\}$. Tässä joukossa on $f'(x) = 1$. (Vrt. Esimerkki 8 edellisessä luvussa.) $\square$

Jos f:n derivaatta f' on edelleen derivoituva pisteessä x, voidaan määritellä f:n *toinen derivaatta* pisteessä x:

$$f''(x) = \lim_{\Delta x \to 0} \frac{f'(x+\Delta x) - f'(x)}{\Delta x} = \mathrm{D}f'(x) = \mathrm{D}(\mathrm{D}f(x)) = \mathrm{D}^2 f(x).$$

Pisteissä, joissa f'' on edelleen on derivoituva, voidaan määritellä kolmas derivaatta $f'''(x)$ jne. Yleisesti funktion *n:s derivaatta* pisteessä x on

$$f^{(n)} = \frac{d^n}{dx^n} f(x) = \mathrm{D}^n f(x).$$

Sanotaan, että n on derivaatan $f^{(n)}$ *kertaluku*, ja sovitaan, että $f^{(0)} = f$.

Derivoimissäännöt

Yleisissä *derivoimissäännöissä* esitetään laskukaavat funktioiden algebrallisten johdannaisten

$$f, g, \lambda \ \mapsto \ \lambda f, \ f+g, \ fg, \ f/g, \ f \circ g, \ f^{-1}$$

derivaattojen laskemiseksi f':n ja g':n avulla. Derivoimissäännöistä yksinkertaisimmat ovat

$$\begin{aligned} \mathrm{D}(\lambda f) &= \lambda f', \quad \lambda \in \mathbb{R}, \\ \mathrm{D}(f+g) &= f' + g', \end{aligned}$$

jotka voidaan yhdistää säännöksi

$$\boxed{\mathrm{D}(\alpha f + \beta g) = \alpha \mathrm{D}f + \beta \mathrm{D}g, \quad \alpha, \beta \in \mathbb{R}.} \tag{1}$$

Sääntö (1) merkitsee, että D on *lineaarinen* operaattori. Sääntö on pätevä jokaisessa pisteessä, jossa f ja g ovat molemmat derivoituvia, kuten voidaan helposti todeta raja-arvojen yhdistelysääntöjen (Lause V.2.4) perusteella. Samoin ehdoin pätee tulon derivoimissääntö

$$\boxed{\mathrm{D}(fg) = f'g + fg'.} \tag{2}$$

Tämän perustelemiseksi kirjoitetaan

$$\begin{aligned} &f(x+\Delta x)g(x+\Delta x) - f(x)g(x) \\ &\qquad = [f(x+\Delta x) - f(x)]g(x+\Delta x) + f(x)[g(x+\Delta x) - g(x)], \end{aligned}$$

jaetaan puolittain Δx:llä, ja sovelletaan mainittuja raja-arvojen yhdistelysääntöjä. Huomioidaan myös, että g on (derivoituvana) jatkuva x:ssä.

Esimerkki 4 Lähtemällä ilmeisestä derivoimissäännöstä $\mathrm{D}x = 1$ ja soveltamalla sääntöä (2) funktiojonoon $f_1(x) = x,\ f_n(x) = xf_{n-1}(x),\ n = 2, 3, \ldots$ nähdään induktiolla oikeaksi sääntö

$$\mathrm{D}x^m = mx^{m-1}, \quad m \in \mathbb{N} \cup \{0\}.$$

Toinen tapa johtaa tämä tulos on lähteä suoraan derivaatan määritelmästä ja käyttää binomikaavaan perustuvaa hajotelmaa:

$$\begin{aligned}\frac{(x+\Delta x)^m - x^m}{\Delta x} &= mx^{m-1} + \binom{m}{2} x^{m-2}\Delta x + \ldots + (\Delta x)^{m-1} \\ &\to mx^{m-1}, \quad \text{kun } \Delta x \to 0. \quad \square\end{aligned}$$

Kun esimerkin tuloksen ohella huomioidaan myös sääntö (1), nähdään että polynomin derivaatta on toinen polynomi (astetta alempi).

Esimerkki 5

$$\begin{aligned}f(x) &= x^5 - 4x^4 + 6x^3 + x^2 - 2x + 3 \\ f'(x) &= 5x^4 - 16x^3 + 18x^2 + 2x - 2 \\ f''(x) &= 20x^3 - 48x^2 + 36x + 2 \\ f'''(x) &= 60x^2 - 96x + 36 \\ f^{(4)}(x) &= 120x - 96 \\ f^{(5)}(x) &= 120 \\ f^{(n)}(x) &= 0, \quad n \geq 6 \quad \square\end{aligned}$$

Jos f ja g ovat n kertaa derivoituvia pisteessä x, niin soveltamalla säännön (2) oikealla puolella uudelleen samaa sääntöä yhdessä säännön (1) kanssa saadaan johdetuksi binomikaavaa (Propositio I.4.1) muistuttava *Leibnizin sääntö*

$$\boxed{\mathrm{D}^n(fg) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.}$$

Jos $h(x) = (f/g)(x)$, niin tulon derivoimissääntö (2) sovellettuna identiteettiin $h(x)g(x) = f(x)$ antaa tuloksen $h'g + hg' = f'$. Pisteissä, joissa $g(x) \neq 0$, voidaan tästä ratkaista h:n eli osamäärän derivoimissäännöksi

$$\boxed{\mathrm{D}\left(\frac{f}{g}\right) = \frac{f'}{g} - \frac{fg'}{g^2} = \frac{f'g - fg'}{g^2}.} \tag{3}$$

Säännöstä (3) sekä polynomin derivoimissäännöstä nähdään, että jokainen rationaalifunktio $f = p/q$ (p ja q polynomeja) on koko määrittelyjoukossaan, eli joukossa $\{x \in \mathbb{R} \mid q(x) \neq 0\}$, derivoituva mielivaltaisen monta kertaa ja että derivaatat ovat ovat samassa joukossa määriteltyjä rationaalifunktioita.

Esimerkki 6 Säännön (3) ensimmäisen laskukaavan mukaan

$$\frac{d}{dx}\left(\frac{x}{1-x^2}\right) = \frac{1}{1-x^2} + \frac{2x^2}{(1-x^2)^2} = \frac{1+x^2}{(1-x^2)^2}\,. \quad \square$$

Esimerkki 7 Soveltamalla sääntöä (3) tapaukseen $f(x) = 1$, $g(x) = x^m$, $m \in \mathbb{N}$ nähdään, että Esimerkin 4 sääntö $\mathrm{D}x^m = mx^{m-1}$ on pätevä $\forall m \in \mathbb{Z}$. $\square$

Lause V.3.2 (**Yhdistetyn funktion derivaatta**) Jos g on derivoituva pisteessä x ja f on derivoituva pisteessä $g(x)$, niin $f \circ g$ on derivoituva pisteessä x ja

$$\boxed{\mathrm{D}f(g(x)) = f'(g(x))\,g'(x).}$$

<u>Todistus</u> Merkitään $g(x+\Delta x) - g(x) = \Delta g$. Tällöin koska f on derivoituva pisteessä $g(x)$, niin pätee (vrt. derivaatan määritelmä edellä)

$$f(g(x+\Delta x)) = f(g(x)+\Delta g) = f(g(x)) + f'(g(x))\Delta g + h(\Delta g)\Delta g,$$

missä $h(\Delta g) \to 0$, kun $\Delta g \to 0$. Koska g on derivoituva pisteessä x ja koska $\Delta x \to 0 \;\Rightarrow\; \Delta g \to 0$, niin seuraa

$$\begin{aligned}
\mathrm{D}(f\circ g)(x) &= \lim_{\Delta x\to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x} \\
&= \lim_{\Delta x\to 0}\left[f'(g(x))\frac{\Delta g}{\Delta x} + h(\Delta g)\frac{\Delta g}{\Delta x}\right] \\
&= f'(g(x))g'(x) + 0\cdot g'(x) = f'(g(x))g'(x). \quad \square
\end{aligned}$$

Soveltamalla Lauseen V.3.2 sääntöä ja Esimerkin 7 tulosta funktioon $g\circ f$, missä $g(x) = x^m$, saadaan derivoimissääntö

$$\boxed{\mathrm{D}[f(x)]^m = m[f(x)]^{m-1}f'(x), \quad m \in \mathbb{Z}.} \tag{4}$$

Esimerkki 8 Jatkamalla Esimerkistä 6 säännöillä (2), (4) ja (1) saadaan

$$\begin{aligned}
\frac{d^2}{dx^2}\left(\frac{x}{1-x^2}\right) = \frac{d}{dx}\left(\frac{1+x^2}{(1-x^2)^2}\right) &= \frac{2x}{(1-x^2)^2} + (1+x^2)\,\frac{(-2)(-2x)}{(1-x^2)^3} \\
&= \frac{6x+2x^3}{(1-x^2)^3}\,. \quad \square
\end{aligned}$$

Jos $f(x) = x^{m/n}$, $m \in \mathbb{Z}$, $n \in \mathbb{N}$, $x > 0$, niin $[f(x)]^n = x^m$. Derivoimalla tässä puolittain säännöllä (4) saadaan

$$n[f(x)]^{n-1} f'(x) = mx^{m-1} \quad \Rightarrow \quad f'(x) = (m/n)\, x^{m/n-1}.$$

Näin on johdettu yleisen potenssifunktion derivoimissääntö

$$\boxed{\mathrm{D}x^\alpha = \alpha x^{\alpha-1}, \quad x \in \mathbb{R}_+\,,\ \alpha \in \mathbb{Q}.}$$

Esimerkki 9 Derivoi $f(x) = \sqrt[3]{1+\sqrt{x}}$.

Ratkaisu Potenssifunktion ja yhdistetyn funktion derivoimissääntöjen perusteella

$$\begin{aligned} f'(x) = \frac{1}{3}(1+\sqrt{x})^{-2/3} \cdot \frac{1}{2}\, x^{-1/2} &= \frac{(1+\sqrt{x})^{1/3}}{6\sqrt{x}\,(1+\sqrt{x})} \\ &= \frac{\sqrt[3]{1+\sqrt{x}}}{6(x+\sqrt{x})}\,, \quad x > 0. \quad \square \end{aligned}$$

Yhdistetyn funktion derivoimissäännöstä (Lause V.3.2) voi johtaa myös yleisen käänteisfunktion derivoimissäännön: Oletetaan, että f on derivoituva pisteessä x ja että käänteisfunktio $g = f^{-1}$ on derivoituva pisteessä $f(x)$. Tällöin identiteetistä $g(f(x)) = x$ seuraa derivoimalla puolittain, että $g'(f(x))f'(x) = 1$. Siis tehdyin oletuksin ja lisäoletuksella $f'(x) \neq 0$ pätee

$$\boxed{\mathrm{D}f^{-1}(y) = \frac{1}{f'(x)}, \quad y = f(x).} \tag{5}$$

Käänteisfunktion derivoituvuuden perustelee tarkemmin seuraava lause.

Lause V.3.3 (**Käänteisfunktion derivaatta**) Olkoon $f : [a-\delta, a+\delta] \to U$ jatkuva bijektio jollakin $\delta > 0$ ja olkoon f derivoituvua pisteessä a. Tällöin $f^{-1} : U \to [a-\delta, a+\delta]$ on derivoituva pisteessä $b = f(a)$ täsmälleen kun $f'(a) \neq 0$, jolloin $\mathrm{D}f^{-1}(b) = 1/f'(a)$.

Todistus Koska f on jatkuva ja $1-1$ välillä $[a-\delta, a+\delta]$ ja f:n arvojoukko ko. valillä $= U$, niin Lauseen V.1.10 mukaan pätee $[b-\varepsilon, b+\varepsilon] \subset U$, missä

$$\varepsilon \;=\; \min\{\, |f(a-\delta) - f(a)|,\, |f(a+\delta) - f(a)|\,\} \;>\; 0.$$

Tällöin jos $f'(a) \neq 0$, niin tekemällä välillä $[b-\varepsilon, b+\varepsilon]$ muuttujan vaihto $y = f(x)$ seuraa lauseiden V.2.6 ja V.2.4 perusteella

$$\lim_{y\to b} \frac{f^{-1}(y) - f^{-1}(b)}{y-b} \;=\; \lim_{x\to a} \frac{x-a}{f(x)-f(a)} \;=\; \left[\lim_{x\to a} \frac{f(x)-f(a)}{x-a}\right]^{-1} \;=\; \frac{1}{f'(a)}\,,$$

joten f^{-1} on derivoituva b:ssä ja $\mathrm{D}f^{-1}(b) = 1/f'(a)$ (Määritelmä V.3.1). Jos $f'(a) = 0$, seuraa Lauseesta V.2.6 vastaavasti, että

$$\lim_{y\to b}\left|\frac{f^{-1}(y) - f^{-1}(b)}{y - b}\right| = \lim_{x\to a}\left|\frac{x - a}{f(x) - f(a)}\right| = \infty,$$

joten tässä tapauksessa f^{-1} ei ole derivoituva b:ssä. $\square$

Esimerkki 10 Funktiolle $f(x) = x^5 + 3x$ Lauseen V.3.3 oletukset ovat voimassa jokaisella $a \in \mathbb{R}$ (myös jokaisella $\delta > 0$, ks. Esimerkki 3 Luvussa IV.2), joten sääntö (5) on soveltuva:

$$\mathrm{D}f^{-1}(y) = \frac{1}{5x^4 + 3}, \quad y = x^5 + 3x.$$

Esimerkiksi $\mathrm{D}f^{-1}(0) = 1/3$, koska $y = 0 \Rightarrow x = 0$. Jos halutaan laskea $\mathrm{D}f^{-1}(1)$, on ensin ratkaistava (numeerisesti) yhtälö $x^5 + 3x = 1$. Jos ratkaisu merkitään $x = a$, niin $\mathrm{D}f^{-1}(1) = 1/(5a^4 + 3)$. $\square$

Implisiittinen derivointi

Jos funktio $y(x)$ on määritelty implisiittisesti muodossa (vrt. Luku IV.2)

$$F(x, y) = 0 \quad \Leftrightarrow \quad y = y(x),$$

niin derivaatta $y'(x)$ on mahdollista laskea derivoimalla puolittain yhtälö

$$F(x, y(x)) = 0.$$

Tällaista epäsuoraa menettelyä sanotaan *implisiittiseksi derivoinniksi.* Toistaiseksi tunnetuilla derivointisäännöillä implisiittinen derivointi onnistuu esim. silloin kun funktio F on (kahden muuttujan) polynomi tai rationaalifunktio.

Esimerkki 10 (jatko) Jos esimerkissä merkitään $f^{-1}(x) = y(x)$, niin $y(x)$ voidaan tulkita implisiittifunktioksi, joka määräytyy yhtälöstä $y^5 + 3y = x$, eli

$$[y(x)]^5 + 3y(x) = x, \quad x \in \mathbb{R}.$$

Kun tässä derivoidaan puolittain soveltaen yhdistetyn funktion derivoimissääntöä ja sääntöä (1), niin seuraa

$$\left[5[y(x)]^4 + 3\right] y'(x) = 1 \;\Rightarrow\; y'(x) = \frac{1}{5y^4 + 3} \quad (y = y(x)). \quad \square$$

Esimerkki 11 Yhtälö

$$x^2 + y^2 = 1$$

määrittelee välillä $[-1, 1]$ kaksihaaraisen implisiittifunktion, jonka kumpikin haara on derivoituva välillä $(-1, 1)$. Jos $x \in (-1, 1)$, niin $y'(x)$ lasketaan helposti implisiittisellä dervoinnilla:

$$\mathrm{D}[x^2 + (y(x))^2] = 2x + 2y(x)y'(x) = \mathrm{D}\,1 = 0 \quad \Rightarrow \quad y'(x) = -\frac{x}{y}\,, \quad y = y(x).$$

Tulos on pätevä kummallakin implisiittifunktion haaralla

$$y(x) = \pm\sqrt{1 - x^2}\,,$$

kuten myös suoralla derivoinnilla voi todeta:

$$y(x) = \pm\sqrt{1 - x^2} \;\Rightarrow\; y'(x) = \pm\left(\frac{-x}{\sqrt{1 - x^2}}\right) = -\frac{x}{y(x)}\,, \quad x \in (-1, 1). \quad \square$$

Potenssisarjan derivointi

Tähänastisten tarkastelujen perusteella voidaan pitää pääsääntönä, että tavanomaiset (algebran keinoin lausekkeina määritellyt) funktiot ovat derivoituvia määrittelyjoukossaan, lukuunottamatta mahdollisia erillisiä pisteitä, joita voi olla äärellinen tai joskus numeroituva määrä. Näytetään luvun lopuksi, että pääsäännöstä eivät tee poikkeusta myöskään funktiot, jotka on määritelty potenssisarjojen summina. Tällaisten funktioiden myötä derivoituvien funktioiden joukko laajenee itse asiassa merkittävästi, kuten myöhemmin tullaan näkemään.

Olkoon reaalifunktio f määritelty muodossa

$$f(x) = \sum_{k=0}^{\infty} a_k x^k.$$

Oletetaan, että tässä potenssisarjan suppenemissäde on joko $\rho \in \mathbb{R}_+$ tai $\rho = \infty$. Tällöin suurin avoin väli, joka sisältyy f:n määrittelyjoukkoon, on $(-\rho, \rho)$ (vrt. Luku I.12). Jatkossa näytetään, että f on derivoituva koko tällä välillä ja että derivaatta voidaan laskea yksinkertaisesti derivoimalla sarja termeittäin, ts. pätee

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}, \quad x \in (-\rho, \rho).$$

Derivoinnin tuloksena saadun sarjan suppenemissäde on myös ρ (Lause I.12.11), joten f' on edelleen derivoituva välillä $(-\rho, \rho)$, ja derivoinnin tulos on siis

$$f''(x) = \sum_{k=2}^{\infty} k(k-1)\, a_k x^{k-2}, \quad x \in (-\rho, \rho).$$

Myös tämän sarjan suppenemissäde $= \rho$ (Lause I.12.11), joten f'' on välillä $(-\rho, \rho)$ edelleen derivoituva, jne. Päätellään, että potenssisarjan summana määritelty funktio on avoimella suppenemisvälillään itse asiassa mielivaltaisen monta kertaa derivoituva funktio (!). Päätelmä perustui siis potenssisarjojen suppenemisteoriaan ja seuraavaan väittämään, joka on todistettavissa suoraan derivaatan määritelmästä lähtien (eli aiempiin derivoimissääntöihin vetoamatta).

Lause V.3.4 **(Potenssisarjan derivaatta)** Jos potenssisarjan $\sum_{k=0}^{\infty} a_k x^k$ suppenemissäde on $\rho > 0$, niin sarjan summana määritelty funktio $f(x)$ on derivoituva välillä $(-\rho, \rho)$ ja ko. välillä pätee $f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$.

<u>Todistus</u> Olkoon $x \in (-\rho, \rho)$ ja valitaan $\Delta x \neq 0$ siten, että

$$|x| + |\Delta x| \leq \rho_0 < \rho.$$

Binomikaavan mukaan

$$\begin{aligned}(x + \Delta x)^k &= \sum_{l=0}^{k} \binom{k}{l} (\Delta x)^l x^{k-l} \\ &= x^k + k x^{k-1} \Delta x + \sum_{l=2}^{k} \binom{k}{l} (\Delta x)^l x^{k-l}.\end{aligned}$$

Kirjoitetaan tässä viimeinen termi summausindeksin vaihdolla muotoon

$$\begin{aligned}\sum_{l=2}^{k} \binom{k}{l} (\Delta x)^l x^{k-l} &= \sum_{i=0}^{k-2} \binom{k}{i+2} (\Delta x)^{i+2} x^{k-2-i} \\ &= (\Delta x)^2 \sum_{i=0}^{k-2} \binom{k}{i+2} (\Delta x)^i x^{k-2-i},\end{aligned}$$

ja edelleen

$$\sum_{i=0}^{k-2} \binom{k}{i+2} (\Delta x)^i\, x^{k-2-i} = \sum_{i=0}^{k-2} c_i \binom{k-2}{i} (\Delta x)^i\, x^{k-2-i},$$

missä

$$c_i = \binom{k}{i+2}\binom{k-2}{i}^{-1} = \frac{k!}{(k-i-2)!\,(i+2)!}\cdot\frac{(k-i-2)!\,i!}{(k-2)!} = \frac{k(k-1)}{(i+1)(i+2)}\,.$$

Koska

$$c_i \le \frac{1}{2}\,k(k-1) < \frac{1}{2}\,k^2, \quad i = 0\ldots k-2,$$

ja oli $|x| + |\Delta x| \le \rho_0$, niin saadaan jokaisella $k \ge 2$ arvio

$$\begin{aligned}
\left|\frac{(x+\Delta x)^k - x^k}{\Delta x} - kx^{k-1}\right| &= |\Delta x|\left|\sum_{i=0}^{k-2} c_i \binom{k-2}{i}(\Delta x)^i\, x^{k-2-i}\right| \\
&\le |\Delta x|\sum_{i=0}^{k-2} c_i \binom{k-2}{i}|\Delta x|^i\,|x|^{k-2-i} \\
&\le \frac{1}{2}k^2|\Delta x|\sum_{i=0}^{k-2}\binom{k-2}{i}|\Delta x|^i\,|x|^{k-2-i} \\
&= \frac{1}{2}k^2|\Delta x|(|x| + |\Delta x|)^{k-2} \\
&\le \frac{1}{2}k^2|\Delta x|\rho_0^{k-2}.
\end{aligned}$$

Näin ollen

$$\begin{aligned}
\left|\frac{f(x+\Delta x) - f(x)}{\Delta x} - \sum_{k=1}^{\infty} k a_k x^{k-1}\right| &= \left|\sum_{k=2}^{\infty} a_k\left[\frac{(x+\Delta x)^k - x^k}{\Delta x} - kx^{k-1}\right]\right| \\
&\le \sum_{k=2}^{\infty}|a_k|\left|\frac{(x+\Delta x)^k - x^k}{\Delta x} - kx^{k-1}\right| \\
&\le |\Delta x|\sum_{k=2}^{\infty}\frac{1}{2}k^2|a_k|\rho_0^{k-2}.
\end{aligned}$$

Tässä oikealla oleva sarja suppenee Lauseen I.12.11 perusteella, koska $\rho_0 < \rho$, joten

$$\left|\frac{f(x+\Delta x) - f(x)}{\Delta x} - \sum_{k=1}^{\infty} k a_k x^{k-1}\right| \le C\,|\Delta x|, \qquad C = \sum_{k=2}^{\infty}\frac{1}{2}k^2|a_k|\rho_0^{k-2}.$$

Saatu arvio on pätevä, kun $x \in (-\rho, \rho)$ ja $|\Delta x| < \delta$, missä esim.

$$\delta = \frac{1}{2}\,(\rho - |x|) > 0.$$

Siis f on jokaisessa pisteessä $x \in (-\rho, \rho)$ derivoituva ja

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \sum_{k=1}^{\infty} k a_k x^{k-1}. \quad \square$$

ESIMERKKI 12 Laske $\sum_{k=1}^{\infty} kq^k$, kun $|q| < 1$.

Ratkaisu Derivoidaan funktio

$$f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad x \in (-1, 1)$$

toisaalta Lauseen V.3.4 perusteella ja toisaalta rationaalifunktiona:

$$f'(x) = \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}, \quad x \in (-1, 1).$$

Valitaan $x = q$ ja kerrotaan puolittain q:lla:

$$q \sum_{k=1}^{\infty} kq^{k-1} = \sum_{k=1}^{\infty} kq^k = \frac{q}{(1-q)^2}. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Laske suoraan määritelmästä (derivoimissääntöjä käyttämättä) linearisoiva approksimaatio seuraaville funktioille pisteessä $a = 1$. Piirrä myös kuva!

   a) $f(x) = \dfrac{x}{1+2x}$ b) $f(x) = \dfrac{x}{2+x^2}$ c) $f(x) = \sqrt{10+6x}$

2. Laske seuraavien funktioiden derivaatat pisteessä $a \in \mathcal{D}_f$ suoraan derivaatan määritelmästä:

   a) $f(x) = x^3 + 3x^2$ b) $f(x) = \dfrac{1}{x}$ c) $f(x) = \dfrac{1}{\sqrt{x}}$

3. Olkoon f derivoituva pisteessä a. Määritä seuraavat raja-arvot $f(a)$:n ja $f'(a)$:n avulla:

   a) $\displaystyle\lim_{h \to 0} \frac{f(a+h^2) - f(a-h)}{h}$ b) $\displaystyle\lim_{x \to a} \frac{xf(a) - af(x)}{x-a}$

   c) $\displaystyle\lim_{t \to 0^+} \frac{f(a+\alpha t) - f(a+\beta t)}{t}, \quad \alpha, \beta \in \mathbb{R}$

4. Lähtien derivaatan määritelmästä määritä $f'(0)$ tai näytä, että f ei ole derivoituva 0:ssa:

$$\text{a)}\ f(x) = \begin{cases} x(1+\sqrt{|x|}\cos\dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$$

$$\text{b)}\ f(x) = \begin{cases} x\sin^2\dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$$

$$\text{c)}\ f(x) = \begin{cases} 1 + x\cos\dfrac{1}{x}\sqrt[100]{x\left|\sin\dfrac{1}{x}\right|}, & \text{kun } x > 0, \\ 1 + x^2, & \text{kun } x \leq 0 \end{cases}$$

5. Palauta osamäärän derivoimissääntö tulon derivoimissääntöön johtamalla ensin funktion $1/f$ derivoimissääntö suoraan derivaatan määritelmästä.

6. Laske kohdissa a)–c) $f^{(n)}(x)$ kun $n = 1, 2$ ja kohdissa d)–f) derivaatan $f^{(n)}(x)$, $n \in \mathbb{N}$ yleinen lauseke:

$$\text{a)}\ f(x) = \frac{x}{\sqrt{a^2 - x^2}} \qquad \text{b)}\ f(x) = \frac{1}{\sqrt{x}+1} \qquad \text{c)}\ f(x) = \sqrt{x + \sqrt{x}}$$

$$\text{d)}\ f(x) = \frac{1}{1+x} \qquad \text{e)}\ f(x) = \sqrt{x+1} \qquad \text{f)}\ f(x) = \frac{1-x}{1+x}$$

7. Tiedetään, että $f(2) = 1$, $f'(2) = 3$ ja $f''(2) = -1$. Laske

$$\text{a)}\ \left[\frac{d}{dx}\left(\frac{x^2}{f(x)}\right)\right]_{x=2} \qquad \text{b)}\ \left[\frac{d}{dx}\left(\frac{f(x)}{x^2}\right)\right]_{x=2} \qquad \text{c)}\ \left[\frac{d^2}{dx^2}\left(\frac{f(x)}{x^2}\right)\right]_{x=2}$$

8. Funktio f määritellän funktion $g(x) = x^2 - 3x + 4$ kaksihaaraisena käänteisfunktiona. Laske $f'(2)$ molemmilla haaroilla a) ratkaisemalla ensin yhtälö $g(x) = 2$ ja käyttämällä käänteisfunktion derivoimissääntöä, b) ratkaisemalla $g(y) = x \Rightarrow y = f(x)$, derivoimalla ja sijoittamalla $x = 2$.

9. Seuraavat yhtälöt määrittelevät implisiittifunktion $y(x)$. Laske implisiittisellä derivoinnilla $y'(x)$ annetussa pisteessä $(x, y) = (x, y(x))$:
a) $2x^2 + 3y^2 = 5$, $(x, y) = (1, 1)$
b) $2x^2 + 3y^2 = 5$, $(x, y) = (1, -1)$
c) $x^2y^3 - x^3y^2 = 12$, $(x, y) = (-1, 2)$
d) $(x-1)(x+2y+1) = y^2$, $(x, y) = (2, -1)$

10. Funktio g on funktion f käänteisfunktio eli $y = f(x) \Leftrightarrow x = g(y)$. Johda derivoimissääntö
$$g''(y) = -\frac{f''(x)}{[f'(x)]^3}, \quad y = f(x)$$
ja laske säännöllä $g''(1)$, kun $f(x) = x^5 + 2x + 1$.

11. Laske implisiittisesti derivoimalla y'' x:n ja y:n avulla:
$$\text{a) } xy = x + y \qquad \text{b) } x^2 + y^2 = 1 \qquad \text{c) } x^3 - y^2 + y^3 = x$$

12. a) Näytä, että $xy'' = y$, kun määritellään
$$y(x) = \sum_{k=1}^{\infty} \frac{x^k}{k!(k-1)!}, \quad x \in \mathbb{R}.$$
b) Näytä, että $u' = -v$ ja $v' = u$, kun määritellään
$$u(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \quad v(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \quad x \in \mathbb{R}.$$

13. Seuraavat funktiot ovat rationaalifunktioita välillä $(-1, 1)$. Laske funktioiden lausekkeet potenssisarjaa $\sum_{k=0}^{\infty} x^k$ derivoimalla.
$$\text{a) } f(x) = \sum_{k=2}^{\infty} kx^k \qquad \text{b) } f(x) = \sum_{k=3}^{\infty} (-1)^k kx^k \qquad \text{c) } f(x) = \sum_{k=1}^{\infty} k^2 x^k$$
$$\text{d) } f(x) = \sum_{k=0}^{\infty} (k+2)^2 x^k \qquad \text{e) } f(x) = \sum_{k=1}^{\infty} k^3 x^k \qquad \text{f) } f(x) = \sum_{k=0}^{\infty} (k+1)^3 x^k$$

14. (*) Todista Leibnizin sääntö.

15. (*) Olkoon $a \in \mathbb{R}$ reaalikertoimisen polynomin p nollakohta ja olkoon $m \in \mathbb{N}$, $m \geq 2$. Näytä, että a on m-kertainen nollakohta täsmälleen kun $p^{(k)}(a) = 0$, $k = 1 \ldots m-1$ ja $p^{(m)}(a) \neq 0$.

16. (*) a) Näytä, että jos potenssisarja $\sum_k a_k x^k$ suppenee välillä $(-\rho, \rho)$ ja funktio g on derivoituva välillä $(a, b) \subset (-\rho, \rho)$, niin funktion
$$f(x) = \sum_{k=0}^{\infty} x^k g(x), \quad x \in (a, b)$$
derivaatta välillä (a, b) on laskettavissa derivoimalla sarja termeittäin.
b) Määritä $f'(x)$ ja $f''(x)$ välillä $(0, \infty)$, kun
$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^{k+\frac{1}{2}}, \quad x \geq 0.$$

V.4 Trigonometristen funktioiden derivointi

Trigonometriset funktiot reaalifunktioina perustuvat kaarenpituuden käsitteeseen (yksikköympyrällä, ks. Luku II.5). Käsite on aiemmin määritelty lyhyesti Luvussa II.1. Seuraavassa tarkastellaan aluksi kaarenpituutta hieman yleisemmältä kannalta ja johdetaan näiden tarkastelujen pohjalta raja-arvotulos

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Trigonometristen funktioiden kaikki derivoimissäännöt ovat tästä perustuloksesta ja Luvun V.3 säännöistä johdettavissa.

Kaarenpituus

Olkoon $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, jatkuva suljetulla välillä $[a, b] \subset \mathcal{D}_f$. Merkitään

$$A = (a, f(a)), \quad B = (b, f(b))$$

ja sanotaan, että euklidisen tason käyrä

$$S = \{P = (x, y) \in E^2 \mid x \in [a, b] \ \ \& \ \ y = f(x)\}$$

on pisteitä A, B yhdistävä (yksinkertainen) *kaari* (tai käyrän kaari, engl. arc).

Olkoon edelleen $\{P_0, \ldots, P_n\}$, $n \in \mathbb{N}$, äärellinen, järjestetty pistejoukko kaarella S siten, että

$$P_i = (x_i, f(x_i)), \quad i = 0, \ldots, n,$$

missä

$$a = x_0 < x_1 < \ldots < x_n = b,$$

jolloin siis $P_0 = A$, $P_n = B$. Merkitään tällaista pistejoukkoa symbolisesti $\mathcal{P}$:llä,

$$\mathcal{P} = \{P_0, \ldots, P_n\} \subset S,$$

ja kaikkien tällaisten pistejoukkojen joukkoa $\mathcal{M}$:llä. Joukko $\mathcal{M}$ on ylinumeroituva, sillä jo pisteellä P_1 on ylinumeroituva määrä vaihtoehtoja.

Pistejoukon $\mathcal{P} \in \mathcal{M}$ kautta kulkevaa murtoviivaa, päätepisteinä A ja B, sanottakoon pisteistön $\mathcal{P}$ kautta kulkevaksi kaaren S *oikopoluksi.*

Jokaisella oikopolulla on (geometrinen) *pituus*, jota merkitään symbolilla $s_{\mathcal{P}}$ ja määritellään

$$s_{\mathcal{P}} = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + [f(x_i) - f(x_{i-1})]^2}\,.$$

Sanotaan, että kaari S on *suoristuva* (engl. rectifiable), jos joukko

$$\mathcal{S} = \{s_{\mathcal{P}} \mid \mathcal{P} \in \mathcal{M}\}$$

on (reaalilukujoukkona) rajoitettu. Tässä tapauksessa kaarelle S voidaan määritellä *kaarenpituus.* Nimittäin, koska rajoitetulla joukolla on pienin yläraja eli supremum (vrt. Luku I.11), niin luontevalta tuntuva määritelmä on

$$\boxed{S\text{:n pituus} = s = \sup \mathcal{S}.}$$

Pelkästään f:n jatkuvuudesta ei seuraa, että kaari S on suoristuva (vaikka vastaesimerkkejä ei ole aivan helppo löytää). Jatkossa tehdään funktiosta f sen vuoksi hieman voimakkaampi oletus, joka takaa kaarenpituuden olemassaolon. Oletus on lievennettävissä koskemaan välin $[a, b]$ osavälejä (ks. alaviite jäljempänä), joten tavanomaisia funktioita ajatellen lisäoletus on melko viaton.

LAUSE V.4.1 Jos f on välillä $[a,b]$ jatkuva ja lisäksi monotoninen, niin kaari $S = \{P = (x,y) \in E^2 \mid x \in [a,b] \ \& \ y = f(x)\}$ on suoristuva, ja kaarenpituudelle s pätee $b - a \le s \le b - a + |f(b) - f(a)|$.

Todistus Koska f on monotoninen, niin $s_{\mathcal{P}}$:n lausekkeessa luvut $f(x_i) - f(x_{i-1})$ ovat samanmerkkisiä (tai nollia), jolloin epäyhtälöä $\sqrt{x^2+y^2} \le |x| + |y|$ ensin soveltaen seuraa

$$\begin{aligned} s_{\mathcal{P}} &\le \sum_{i=1}^{n} [\,(x_i - x_{i-1}) + |f(x_i) - f(x_{i-1})|\,] \\ &= \sum_{i=1}^{n} (x_i - x_{i-1}) + \left| \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \right| = b - a + |f(b) - f(a)|. \end{aligned}$$

Todetaan myös, että $s_{\mathcal{P}} \ge b - a \ \forall \ \mathcal{P}$, sillä pienin mahdollinen $s_{\mathcal{P}}$ (jokaisella $\mathcal{P}$) saadaan, kun f on vakio. Tehdyin oletuksin siis pätee

$$b - a \ \le \ s_{\mathcal{P}} \ \le \ b - a + |f(b) - f(a)| \quad \forall \ s_{\mathcal{P}} \in \mathcal{S}.$$

Joukolle $\mathcal{S}$ on näin saatu alaraja ja yläraja, joten luku $s = \sup \mathcal{S}$ on olemassa ja saadut rajat pätevät myös tälle luvulle. $\square$

Ellei S itse satu olemaan murtoviiva (jolloin $s \in \mathcal{S}$), on kaarenpituus määrättävä käytännössä konstruoimalla jono $\mathcal{P}_n \in \mathcal{M}$, $n = 1, 2, \ldots$ niin, että $s_{\mathcal{P}_n} = s_n \to s$, kun $n \to \infty$. Esimerkiksi jos valitaan $\mathcal{P}_n = \{a + k(b-a)/n, \ k = 0 \ldots n\}$, niin algoritmi toimiikin tavanomaisissa tapauksissa (vrt. Harj.teht. 1).†

Jos koko välin $[a,b]$ sijasta tarkastellaan väliä $[a,x]$, $a < x \le b$, ja määritellään

$$S_x = \{P = (t, f(t)) \mid \ a \le t \le x\},$$

niin S_x:n kaarenpituudesta $s(x)$ tulee välillä $[a,b]$ määritelty funktio, kun vielä asetetaan $s(a) = 0$. Koska $s(x_2) - s(x_1)$ tällöin merkitsee kaarenpituutta välillä $[x_1, x_2]$ (ks. alaviite), on Lauseen V.4.1 mukaan

$$s(x_2) - s(x_1) \ \ge \ x_2 - x_1, \quad a \le x_1 < x_2 \le b,$$

joten $s(x)$ on välillä $[a,b]$ aidosti kasvava. Lisäksi nähdään Lauseen V.4.1 ylärajaarviosta, että $s(x)$ on jatkuva välillä $[a,b]$.

†Kaarenpituuden laskemiseen palataan myöhemmin toisessa asiayhteydessä. Tällöin myös näytetään, että kaarenpituus on tarkasteltavan välin suhteen *additiivinen*, ts. jos $a < c < b$, niin $s = s_1 + s_2$, missä s, s_1 ja s_2 ovat vastaavasti kaarenpituudet välillä $[a,b]$, $[a,c]$ ja $[c,b]$. Additiivisuuden ja Lauseen V.4.1 perusteella käyrän $S : y = f(x)$ kaaren suoristuvuuteen välillä $[a,b]$ riittää jatkuvuusoletuksen lisäksi, että väli $[a,b]$ on jaettavissa äärellisen moneen osaväliin, joilla f on monotoninen.

Siirrytään nyt tarkastelemaan jatkon kannalta tärkeää erikoistapausta, missä $[a, b] = [0, 1]$ ja funktio f määritellään

$$f(x) = 1 - \sqrt{1 - x^2}, \quad x \in [0, 1].$$

Lauseen V.4.1 oletukset toteutuvat tällöin, ja S on x-akselia origossa sivuavan 1-säteisen ympyrän kaari.

Kun merkitään $t =$ kaarenpituus välillä $[0, x]$, niin (ks. kuvio)

$$x = \sin t.$$

Kun x:n ja t:n välille oletetaan tämä yhteys välillä $x \in [0, 1] \;\Leftrightarrow\; t \in [0, \frac{\pi}{2}]$, niin ko. välillä voidaan kirjoittaa

$$f(x) \;=\; 1 + \sqrt{1 - x^2} \;=\; 1 - \cos t$$

ja arvioida Lauseen V.4.1 perusteella

$$\begin{cases} t \geq x, \\ t \leq x + f(x) \end{cases} \quad \Leftrightarrow \quad \begin{cases} t \geq \sin t, \\ t \leq \sin t + 1 - \cos t \end{cases}$$

eli

$$t - (1 - \cos t) \;\leq\; \sin t \;\leq\; t, \quad t \in [0, \tfrac{\pi}{2}].$$

Tässä on viitattu aiemmin (Luku II.1) sovittuun merkintään (π:n määritelmä!)

$$S\text{:n pituus} = \tfrac{\pi}{2}.$$

Kun saadun epäyhtälön vasemmalla puolella käytetään trigonometrian kaavaa $1 - \cos t = 2 \sin^2 \frac{t}{2}$ (ks. Luku II.5) ja sovelletaan epäyhtälön jälkimmäistä osaa, seuraa

$$t - (1 - \cos t) \;=\; t - 2 \sin^2 \tfrac{t}{2} \;\geq\; t - 2 \left(\tfrac{t}{2}\right)^2 \;=\; t - \tfrac{1}{2} t^2, \quad t \in [0, \pi].$$

Yhdistämällä epäyhtälöt on tultu päätelmään

$$\begin{aligned} & t - \tfrac{1}{2}t^2 \;\le\; \sin t \;\le\; t, \quad t \in [0, \tfrac{\pi}{2}] \\ \Rightarrow \quad & 1 - \tfrac{1}{2}\,t \;\le\; \frac{\sin t}{t} \;\le\; 1, \quad t \in (0, \tfrac{\pi}{2}] \\ \Rightarrow \quad & 1 - \tfrac{1}{2}|t| \;\le\; \frac{\sin t}{t} \;\le\; 1, \quad 0 < |t| \le \tfrac{\pi}{2}. \end{aligned}$$

Tässä viimeinen päättely perustui funktion $\sin t/t$ parillisuuteen.

Jos nyt $t_n \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $t_n \neq 0$ ja $t_n \to 0$, niin viimeksi kirjoitetun epäyhtälön perusteella

$$1 \;\ge\; \frac{\sin t_n}{t_n} \;\ge\; 1 - \tfrac{1}{2}|t_n| \to 1.$$

Raja-arvon määritelmän (Määritelmä V.2.1) perusteella on johdettu raja-arvotulos

$$\boxed{\lim_{t\to 0} \frac{\sin t}{t} = 1.} \tag{1}$$

Kun sovelletaan tätä tulosta, kaavaa $1 - \cos t = 2\sin^2 \frac{t}{2}$ ja raja-arvojen yhdistelysääntöjä (Lause V.2.4), niin seuraa

$$\boxed{\lim_{t\to 0} \frac{1 - \cos t}{t^2} = \frac{1}{2}\,.} \tag{2}$$

Edellä olennaisesti johdettiin myös seuraavat, kaikilla $t \in \mathbb{R}$ pätevät epäyhtälöt:

$$\boxed{|\sin t| \;\le\; |t|, \quad 0 \;\le\; 1 - \cos t \le \frac{1}{2}\,t^2, \quad t \in \mathbb{R}.}$$

Kun raja-arvotulokset (1), (2) yhdistetään trigonometrisiin yhteenlaskukaavoihin

$$\begin{aligned} \sin(x+t) &= \sin x \cos t + \cos x \sin t, \\ \cos(x+t) &= \cos x \cos t - \sin x \sin t \end{aligned}$$

(ks. Luku II.5) saadaan yleisemmät raja-arvotulokset (Harj.teht. 3)

$$\boxed{\begin{aligned} \lim_{t\to 0} \frac{\sin(x+t) - \sin x}{t} &= \cos x, \\ \lim_{t\to 0} \frac{\cos(x+t) - \cos x}{t} &= -\sin x. \end{aligned}} \tag{3}$$

Trigonometristen funktioiden derivaatat

Raja-arvotulosten (3) perusteella funktioiden sin ja cos derivoimissäännöt ovat

$$\boxed{\begin{aligned} \mathrm{D}\sin x &= \cos x, \quad x \in \mathbb{R}, \\ \mathrm{D}\cos x &= -\sin x, \quad x \in \mathbb{R}. \end{aligned}}$$

Näiden sääntöjen ja Luvun V.3 sääntöjen perusteella on

$$\mathrm{D}\left(\frac{\sin x}{\cos x}\right) = 1 + \frac{\sin^2 x}{\cos^2 x}\,.$$

Tämä ja vastaava cot-funktion derivoimissääntö annetaan yleensä muodossa

$$\boxed{\begin{aligned} \mathrm{D}\tan x &= 1 + \tan^2 x = \frac{1}{\cos^2 x}\,, \\ \mathrm{D}\cot x &= -(1 + \cot^2 x) = -\frac{1}{\sin^2 x}\,. \end{aligned}}$$

Trigonometristen käänteisfunktioiden derivaatat voidaan laskea edellisen luvun säännöllä (5). Esim. jos $y \in (-1, 1)$ ja $\operatorname{Arcsin} y = x \in (-\frac{\pi}{2}, \frac{\pi}{2})$, niin mainittu sääntö yhdessä trigonometrian kaavojen kanssa antaa

$$\frac{d}{dy}\operatorname{Arcsin} y = \frac{1}{\mathrm{D}\sin x} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - \sin^2 x}} = \frac{1}{\sqrt{1 - y^2}}\,.$$

Kun Arccos ja Arctan derivoidaan vastaavalla tavalla on tulos (Harj.teht. 6) :

$$\boxed{\begin{aligned} \mathrm{D}\operatorname{Arcsin} x &= \frac{1}{\sqrt{1 - x^2}}\,, \quad x \in (-1, 1), \\ \mathrm{D}\operatorname{Arccos} x &= -\frac{1}{\sqrt{1 - x^2}}\,, \quad x \in (-1, 1), \\ \mathrm{D}\operatorname{Arctan} x &= \frac{1}{1 + x^2}\,, \quad x \in \mathbb{R}. \end{aligned}}$$

Em. derivoimissääntöjen antamia funktioita edelleen derivoimalla nähdään, että kaikki säännöissä mainitut funktiot ovat mielivaltaisen monta kertaa derivoituvia koko määrittelyjoukossaan, jos tämä koostuu avoimista väleistä, muuten kaikilla määrittelyjoukon avoimilla osaväleillä. Sama pätee myös yleisemmille trigonometrisille funktioille, jotka saadaan yhdistelemällä perusfunktioita laskuoperaatioilla tai yhdistettyinä funtioina.

Esimerkki 1 Funktioiden sin ja cos sekä edellisen luvun dervioimissääntöjen nojalla on

$$\mathrm{D}\,\frac{\sin x}{1+\cos x} = \frac{\cos x}{1+\cos x} + \frac{\sin^2 x}{(1+\cos x)^2} = \frac{\cos x + \cos^2 x + \sin^2 x}{(1+\cos x)^2} = \frac{1}{1+\cos x}\,.$$

Lasku on pätevä jokaisella $x \in \mathbb{R}$, jolla $\cos x \neq -1$, eli derivoitavan funktion koko määrittelyjoukossa. $\square$

Esimerkki 2 Funktiot $u(x) = \cos x$ ja $v(x) = \sin x$ ovat (ainakin eräs) ratkaisu ongelmaan

$$\begin{cases} u' = -v,\ v' = u, & x \in \mathbb{R} \\ u(0) = 1,\ v(0) = 0 \end{cases}$$

(Ongelma on ratkaistavissa toisellakin tavalla, ks. Harj.teht. V.3:12b.) $\square$

HARJOITUSTEHTÄVIÄ

1. Tutki, kuinka tarkasti kaaren

$$S = \{P = (x, y) \mid y = x^2 \ \ \&\ \ x \in [0, 1]\}$$

pituus saadaan lasketuksi käyttämällä pisteistöä $\{(x_i, f(x_i)),\ i = 0 \ldots 10\}$, missä $x_i = i/10$. Tarkka arvo on $s = 1.47894..$

2. Origosta siirrytään käyrää $r = 4\cos\varphi$ (napakoordinaatit) seuraten pisteeseen, jonka x-koordinaatti $= 3$. Mikä on matkan pituus lyhintä reittiä?

3. Johda raja-arvotulokset (3).

4. Näytä, että jos $a, b \in \mathbb{R}$ ja $b \neq 0$, niin $\displaystyle\lim_{x\to 0} \frac{\sin ax}{\sin bx} = \frac{a}{b}\,.$

5. Määritä raja-arvot

a) $\displaystyle\lim_{x\to 0} \frac{x\sin x}{1-\cos x}$ b) $\displaystyle\lim_{x\to 0} \frac{x\sin 2x}{1-\cos 2x}$ c) $\displaystyle\lim_{x\to 0^+} \frac{x-\sqrt{x}}{\sqrt{\sin x}}$

d) $\displaystyle\lim_{x\to \frac{\pi}{2}} (1+\cos 2x)\tan^2 x$ e) $\displaystyle\lim_{x\to\infty} x\sin\frac{1}{x}$

6. Johda funktioiden Arccos ja Arctan derivoimissäännöt.

7. Derivoi ja sievennä lopputulos:

$$\text{a)}\ \frac{1-\cos x}{\sin x} \qquad \text{b)}\ \frac{1+\cos x}{1-\cos x} \qquad \text{c)}\ \frac{\sin x-\cos x}{\sin x+\cos x}$$

$$\text{d)}\ \operatorname{Arcsin}(\cos x) \qquad \text{e)}\ \operatorname{Arccos}(\sin x) \qquad \text{f)}\ \operatorname{Arctan}(\cot x)$$

8. Määritä implisiittifunktion $y(x)$ derivaatta annetussa pisteessä:

$$\text{a)}\ y+2\sin y+\cos y = x, \quad (x,y)=(1,0)$$
$$\text{b)}\ 2x+y-\sqrt{2}\sin(xy) = \frac{\pi}{2}, \quad (x,y)=\left(\frac{\pi}{4},1\right)$$
$$\text{c)}\ x\sin(xy-y^2) = x^2-1, \quad (x,y)=(1,1)$$
$$\text{d)}\ \tan(xy^2) = \frac{2xy}{\pi}, \quad (x,y)=\left(-\pi,\frac{1}{2}\right)$$

9. Näytä induktiolla, että $\mathrm{D}^n \tan x = p(\tan x)$, missä p on polynomi astetta $n+1$ ja muotoa $p(t) = n!\,t^{n+1} + c_{n-1}\,t^{n-1} + c_{n-3}\,t^{n-3} + \ldots$

10. Totea funktio

$$f(x) = \begin{cases} x^2 \sin\dfrac{1}{x^2}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$$

esimerkiksi funktiosta, joka on derivoituva jokaisessa pisteessä $x \in \mathbb{R}$, mutta derivaatta ei ole jokaisessa pisteessä jatkuva. Hahmottele f:n ja f':n kuvaajat!

11. (*) Olkoon s_n käyrän $S : y = x^n$ kaarenpituus välillä $[0,1]$, kun $n \in \mathbb{N}$. Näytä, että $\lim_n s_n = 2$.

12. (*) Näytä sopivalla muuttujan vaihdolla:

$$\text{a)}\ \lim_{x\to\infty} x\left(\frac{\pi}{2} - \operatorname{Arctan} x\right) = 1 \qquad \text{b)}\ \lim_{x\to 1^-} \frac{\pi - 2\operatorname{Arcsin} x}{\sqrt{1-x}} = 2\sqrt{2}$$

V.5 Ääriarvot. Sileys

Derivaatta on mitä mainioin työkalu, kun halutaan luonnehtia tavallisia reaalifunktioita, jotka useimmiten ovat derivoituvia 'melkein kaikkialla'. Tässä ja seuraavassa luvussa tarkastellaan derivaatan käyttöä funktiotutkimuksessa.

Paikalliset ääriarvot

Aloitetaan määritelmästä.

MÄÄRITELMÄ V.5.1 Funktiolla $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on pisteessä $c \in \mathcal{D}_f$ *paikallinen* eli *suhteellinen ääriarvo*, eli c on f:n *paikallinen ääriarvopiste* (ääriarvokohta), jos jollakin $\delta > 0$ ja $\forall x \in \mathbb{R}$ pätee $(c-\delta, c+\delta) \subset \mathcal{D}_f$ ja

$$\begin{aligned} \text{joko:} \quad & 0 < |x-c| < \delta \;\Rightarrow\; f(x) \le f(c), \\ \text{tai:} \quad & 0 < |x-c| < \delta \;\Rightarrow\; f(x) \ge f(c). \end{aligned}$$

Edellisessä tapauksessa on kyseessä *paikallinen maksimi*, jälkimmäisessä *paikallinen minimi*. Jos lisäksi jollakin $\delta > 0$ on $f(x) \neq f(c)$ aina kun $0 < |x-c| < \delta$, niin ääriarvo (ääriarvopiste) on *oleellinen*, muulloin *epäoleellinen*.

ESIMERKKI 1 Funktio

$$f(x) = \begin{cases} \sin^2 \dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$$

saavuttaa pienimmän arvonsa $f_{min} = 0$ pisteessä $x = 0$ sekä pisteissä

$$a_n = \frac{1}{n\pi}, \quad n \in \mathbb{Z},\ n \neq 0.$$

Suurimman arvonsa $f_{max} = 1$ funktio saavuttaa pisteissä

$$b_n = \frac{1}{(n+\frac{1}{2})\pi}, \quad n \in \mathbb{Z}.$$

Pisteet a_n ja b_n ovat myös Määritelmän V.5.1 mukaisia oleellisia paikallisia minimi- ja maksimikohtia, sillä f:n määrittelyn ja sin-funktion tunnettujen ominaisuuksien perusteella on $0 < f(x) < 1$ kaikissa muissa kuin mainituissa pisteissä tai pisteessä $x = 0$. Tällöin jos esim. $c = b_n$, $n \in \mathbb{Z}$, niin Määritelmän V.5.1 paikallisen maksimin ehto on voimassa, kun valitaan δ siten, että $0 < \delta < |b_{n\pm1} - b_n|$. Pisteessä $x = 0$ on f:llä Määritelmän V.5.1 mukainen epäoleellinen paikallinen minimi, sillä jos valitaan mikä tahansa $\delta > 0$, niin $a_n \in (0, \delta)$ kun $n > (\pi\delta)^{-1}$, jolloin välillä $(0, \delta)$ on aina pisteitä, joissa $f(x) = f(0)$. $\square$

Seuraavan lauseen mukaan derivoituvan funktion paikallinen ääriarvokohta on välttämättä myös derivaatan nollakohta. Tähän liittyen sanotaankin derivaatan nollakohtia funktion *kriittisiksi pisteiksi.*

LAUSE V.5.2 Jos $c \in \mathcal{D}_f$ on f:n paikallinen ääriarvopiste ja f on derivoituva c:ssä, niin $f'(c) = 0$.

Todistus Väittämän loogisesti ekvivalentti muoto on: Jos f on c:ssä derivoituva ja $f'(c) \neq 0$, niin c ei ole f:n paikallinen ääriarvopiste. Todistetaan väittämä tässä muodossa, eli oletetaan, että $f'(c) = k \neq 0$. Derivaatan määritelmän perusteella (vrt. Luku V.3) f on tällöin määritelty välillä $(c - \delta, c + \delta)$ jollakin $\delta > 0$ ja ko. välillä pätee

$$f(x) = f(c) + k(x - c) + g(x),$$

missä $\lim_{x \to c} g(x)/(x - c) = 0$. Tällöin koska $k \neq 0$, niin Lauseen V.2.2 mukaan jollakin $\delta > 0$ (vastaten valintaa $\varepsilon = |k|/2 > 0$ ko. lauseessa) pätee myös

$$|g(x)/(x - c)| \leq \tfrac{1}{2}|k| \;\Rightarrow\; |g(x)| \leq \tfrac{1}{2}|k||x - c|, \quad \text{kun } 0 < |x - c| < \delta.$$

Näin ollen jos $k > 0$, niin kolmioepäyhtälön nojalla

$$\begin{cases} f(x) - f(c) \;\geq\; k(x - c) - \frac{k}{2}\,|x - c| \;=\; \frac{k}{2}\,(x - c) \;>\; 0 & \forall x \in (c, c + \delta), \\ f(x) - f(c) \;\leq\; k(x - c) + \frac{k}{2}\,|x - c| \;=\; \frac{k}{2}\,(x - c) \;<\; 0 & \forall x \in (c - \delta, c). \end{cases}$$

Jos $k < 0$, päätellään samalla tavoin, että $f(x) - f(c) < 0$ kun $x \in (c, c + \delta)$ ja $f(x) - f(c) > 0$ kun $x \in (c - \delta, c)$. Määritelmän V.5.1 mukaan kummassakaan tapauksessa c ei ole f:n paikallinen ääriarvopiste. □

ESIMERKKI 1 (jatko) Esimerkin funktiolle pätee

$$f'(x) = -\frac{2}{x^2}\,\sin\frac{1}{x}\cos\frac{1}{x}\,, \quad \text{kun } x \neq 0,$$

joten $f'(a_n) = f'(b_n) = 0$ Lauseen V.5.2 väittämän mukaisesti. Pisteesä $x = 0$ ei f ole derivoituva — tämäkin on sopusoinnussa Lauseen V.5.2 kanssa. □

ESIMERKKI 2 Rationaalifunktio $f(x) = x^2/(x + 1)$ on derivoituva koko määrittelyjoukossaan, joten f:n mahdolliset paikalliset ääriarvokohdat ovat väistämättä kriittisiä pisteitä. Koska

$$f'(x) \;=\; \frac{2x}{1 + x} - \frac{x^2}{(1 + x)^2} \;=\; \frac{x^2 + 2x}{(x + 1)^2}\,, \quad x \neq -1,$$

niin kriittiset pisteet ovat $x = -2$ ja $x = 0$. Osoittautuu, että edellisessä pisteessä f:llä on paikallinen maksimi, jälkimmäisessä paikallinen minimi. (Ks. seuraavan luvun Esimerkki 3.) □

Ääriarvotehtävät

Kysymys funktion paikallisista ääriarvokohdista nousee yleensä esiin osatehtävänä, kun halutaan ratkaista *ääriarvotehtävä*: On annettu funktio f ja joukko $A \subset \mathcal{D}_f$, ja haluttaa määrittää f:n suurin ja/tai pienin arvo A:ssa. Sikäli kuin tehtävä ratkeaa, eli suurin/pienin arvo on löydettävissä, sanotaan ko. arvoa f:n *absoluuttiseksi* maksimiksi/minimiksi A:ssa. Absoluuttinen maksimi ja minimi löytyvät aina, jos joukko A on äärellinen, sillä tällöin kyse on vain valinnasta äärellisessä joukossa. Jos A on ääretön joukko, esim. väli, on tilanne ongelmallisempi. Rajoitutaan tässä sovelluksissa usein esiintyvään perustehtävään, jossa A on suljettu väli ja f on ko. välillä jatkuva Määritelmän V.1.5 mukaisesti. Tällöin ääriarvotehtävän ratkeavuuden takaa kummankin ääriarvon osalta Weierstrassin lause (Lause V.1.12).

Jatkuvaa funktiota f ja suljettua väliä $[a, b]$ koskevan ääriarvotehtävän ratkaiseminen helpottuu käytännössä huomattavasti, jos f on paitsi jatkuva myös derivoituva tai 'melkein derivoituva' välillä (a, b). Tulos on muotoiltavissa Lauseen V.5.2 korollaarina seuraavasti.

Korollaari V.5.3 Olkoon f on jatkuva välillä $[a, b]$ ja olkoon $X \subset [a, b]$ joukko, jolle pätee

1. $a \in X$ ja $b \in X$,
2. $\{c \in (a, b) \mid f \text{ derivoituva } c\text{:ssä ja } f'(c) = 0\} \subset X$,
3. $\{c \in (a, b) \mid f \text{ ei derivoituva } c\text{:ssä}\} \subset X$.

Tällöin $\max_{x\in[a,b]} / \min f(x) = \max_{x\in X} / \min f(x)$.

Korollaarin V.5.3 ehdot täyttävä joukko X siis sisältää välin $[a, b]$ päätepisteet ja lisäksi kaikki välillä (a, b) olevat f:n kriittiset pisteet sekä pisteet, joissa f ei ole derivoituva. Yleensä joukko X voidaan valita niin, että se on äärellinen, jolloin f:n maksimi- ja minimiarvojen haku pelkistyy korollaarin mukaisesti valinnaksi äärellisessä joukossa X. — Huomattakoon, että joukkoon X (sikäli kuin äärellinen) voidaan sen määrittelyn mukaisesti sisällyttää myös sellaisia pisteitä, joissa f:n derivoituvuus on pelkästään epäilyksen alaista. Haluttaessa vain selvittää funktion maksimi- ja minimikohdat ja -arvot välillä $[a, b]$ ei tarkempaa tutkimusta 'epäillyistä' tarvita. Avoimeksi voidaan myös jättää kysymys, ovatko joukkoon X sisällytettävät pisteet f:n paikallisia ääriarvokohtia vai eivät.

Esimerkki 3 Määritä funktion $f(x) = \max\{2 - x,\ 6x - 3x^2\}$ maksimi- ja minimikohdat ja -arvot välillä $[0, 3]$.

Ratkaisu Kysessä on jatkuva ja paloittain polynomiarvoinen funktio. Koska $2 - x = 6x - 3x^2 \;\Leftrightarrow\; 3x^2 - 7x + 2 = 0 \;\Leftrightarrow\; x = \frac{1}{3} \;\vee\; x = 2$, niin päätellään, että

$$f(x) = \begin{cases} 6x - 3x^2, & \text{kun } \frac{1}{3} \le x \le 2, \\ 2 - x, & \text{kun } x \le \frac{1}{3} \text{ tai } x \ge 2. \end{cases}$$

Väleillä $(0, \frac{1}{3})$ ja $(2, 3)$ on $f'(x) = -1 \neq 0$. Välillä $(\frac{1}{3}, 2)$ on $f'(x) = 6 - 6x$, joten tällä välillä on f:llä kriittinen piste $c = 1$. Pisteissä $x = \frac{1}{3}$ ja $x = 2$ ei f mahdollisesti (eikä todellisuudessakaan) ole derivoituva, muissa välin $(0, 3)$ pisteissä on, joten Korollaarissa V.5.3 voidaan valita $X = \{0, \frac{1}{3}, 1, 2, 3\}$. Laskemalla f:n arvot näissä pisteissä todetaan, että $f_{max} = f(1) = 3$, $f_{min} = f(3) = -1$. $\square$

Esimerkissä mahdollisia paikallisia ääriarvokohtia välillä $(0, 3)$ ovat Lauseen V.5.2 mukaan pisteet $\frac{1}{3}, 1, 2$. Funktion lähempi tutkimus näiden pisteiden ympäristössä (algebran keinoin tai jatkossa esitettävin menetelmin) osoittaa, että $c = \frac{1}{3}$ on oleellinen paikallinen minimi, $c = 1$ on oleellinen paikallinen maksimi ja $c = 2$ ei ole paikallinen ääriarvokohta.

Toispuoliset derivaatat

Jos Esimerkissä 3 tarkkaillaan niitä pisteitä, joissa f ei ole derivoituva, nähdään että f:llä on näissäkin pisteissä *toispuoliset* (engl. one-sided) derivaatat seuraavan määritelmän mielessä (vrt. toispuoliset raja-arvot Luvussa V.2).

Määritelmä V.5.4 Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on pisteessä $x \in \mathcal{D}_f$ *vasemmalta derivoituva*, jos $(x - \delta, x] \subset \mathcal{D}_f$ jollakin $\delta > 0$ ja $\exists$ raja-arvo

$$\mathrm{D}_- f(x) = \lim_{\Delta x \to 0^-} \frac{f(x + \Delta x) - f(x)}{\Delta x},$$

ja *oikealta derivoituva*, jos $[x, x + \delta) \subset \mathcal{D}_f$ jollakin $\delta > 0$ ja $\exists$ raja-arvo

$$\mathrm{D}_+ f(x) = \lim_{\Delta x \to 0^+} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Määritelmien V.5.4 ja V.3.1 perusteella f on pisteessä x derivoituva täsmälleen kun f on x:ssä sekä vasemmalta että oikealta derivoituva ja $\mathrm{D}_- f(x) = \mathrm{D}_+ f(x)$ $(= f'(x))$.

Jos f on pisteessä c jatkuva ja lisäksi sekä vasemmalta että oikealta derivoituva, niin pätee (vrt. johdatus derivaattaan Luvussa V.3)

$$\begin{cases} f(x) = f(x) + \mathrm{D}_- f(c)(x - c) + g(x), & \text{kun } x \in (c - \delta, c], \\ f(x) = f(x) + \mathrm{D}_+ f(c)(x - c) + g(x), & \text{kun } x \in [c, c + \delta), \end{cases}$$

missä $g(x)/(x - c) \to 0$, kun $x \to c^-$ tai $x \to c^+$. Tästä nähdään oikeaksi

LAUSE V.5.5 Jos f on pisteessä c jatkuva ja lisäksi sekä vasemmalta että oikealta derivoituva, niin pätee

(a) $\mathrm{D}_-f(c)\,\mathrm{D}_+f(c) < 0 \;\Rightarrow\;$ f:llä on pisteessä c paikallinen ääriarvo, joka on

- oleellinen maksimi, jos $\mathrm{D}_-f(c) > 0$ ja $\mathrm{D}_+f(c) < 0$,
- oleellinen minimi, jos $\mathrm{D}_-f(c) < 0$ ja $\mathrm{D}_+f(c) > 0$.

(b) $\mathrm{D}_-f(c)\,\mathrm{D}_+f(c) > 0 \;\Rightarrow\;$ c ei ole f:n paikallinen ääriarvopiste.

ESIMERKKI 3 (jatko) Esimerkissä on $\mathrm{D}_-f(\frac{1}{3}) = -1$, $\mathrm{D}_+f(\frac{1}{3}) = 4$ ja $\mathrm{D}_-f(2) = -6$, $\mathrm{D}_+f(2) = -1$, joten Lauseen V.5.5 mukaan f:llä on pisteessä $x = \frac{1}{3}$ oleellinen paikallinen minimi ja $x = 2$ ei ole paikallinen ääriarvokohta. $\square$

Jos f on pisteessä c derivoituva ja $f'(c) \neq 0$, niin $\mathrm{D}_-f(c)\,\mathrm{D}_+f(c) = [f'(c)]^2 > 0$, joten Lauseen V.5.5 väittämä (b) sisältää myös Lauseen V.5.2 väittämän muodossa

$$f'(c) \neq 0 \quad \Rightarrow \quad f\text{:llä ei ole paikallista ääriarvoa } c\text{:ssä.}$$

Jos $\mathrm{D}_-f(c) = 0$ tai $\mathrm{D}_+f(c) = 0$ (myös kun $f'(c) = 0$), voi c olla paikallinen ääriarvopiste tai ei, riippuen tapauksesta.

ESIMERKKI 4 Funktioille $f(x) = \begin{cases} x^2, & x < 0, \\ x, & x \geq 0 \end{cases}$ ja $g(x) = \begin{cases} -x^2, & x < 0, \\ x, & x \geq 0 \end{cases}$

pätee $\mathrm{D}_-f(0) = \mathrm{D}_-g(0) = 0$ ja $\mathrm{D}_+f(0) = \mathrm{D}_+g(0) = 1$. Funktiolla f on origossa (oleellinen) paikallinen minimi, g:llä ei ole paikallista ääriarvoa origossa.

$y = f(x)$

$y = g(x)$

Sileys

Seuraava määritelmä yhdistävää suljetulla välillä jatkuvuuden (Määritelmä V.1.5) ja derivoituvuuden käsitteitä.

MÄÄRITELMÄ V.5.6 Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$ on *jatkuvasti derivoituva* (engl. continuously differentiable) *suljetulla välillä* $[a, b] \subset \mathcal{D}_f$, jos f on derivoituva välillä (a, b), oikealta derivoituva pisteessä a, vasemmalta derivoituva pisteessä b, ja derivaatta f', määriteltynä toispuolisena välin $[a, b]$ päätepisteissä, on välillä $[a, b]$ jatkuva funktio. Yleisemmin jos $m \in \mathbb{N}$, niin f on välillä $[a, b]$ m *kertaa jatkuvasti derivoituva*, jos f on m kertaa derivoituva välillä (a, b), m kertaa oikealta derivoituva pisteessä a, m kertaa vasemmalta derivoituva pisteessä b, ja derivaatat $f^{(k)}$, määriteltynä toispuolisina välin päätepisteissä, ovat välillä $[a, b]$ jatkuvia, kun $k = 1 \ldots m$. Edelleen jos $A \subset \mathcal{D}_f$ on puoliavoin tai avoin väli, niin sanotaan, että f on m kertaa jatkuvasti derivoituva A:ssa, jos f on m kertaa jatkuvasti derivoituva jokaisella A:n suljetulla osavälillä.

Sovellettaessa määritelmää suljetun välin tapauksessa, kun $m \geq 2$, ajatellaan derivaatat $f^{(k)}$, $k = 1 \ldots m - 1$ määritellyksi välin päätepisteissä palautuvasti (toispuolisina) alkaen indeksistä $k = 1$. Indeksiä $m = 0$ vastaava termi '0 kertaa jatkuvasti derivoituva' tulkitaan pelkäksi jatkuvuudeksi.

ESIMERKKI 5 Olkoon $n \in \mathbb{N}$ ja tarkastellaan funktiota

$$f(x) = \begin{cases} 0, & \text{kun } x \leq 0, \\ x^n, & \text{kun } x > 0. \end{cases}$$

Derivoimalla pistessä $x = 0$ toispuolisesti todetaan, että f on tässä pistessä (ja siis koko $\mathbb{R}$:ssä) $n - 1$ kertaa derivoituva ja pätee

$$f^{(k)}(x) = \begin{cases} 0, & \text{kun } x \leq 0, \\ \dfrac{n!}{(n-k)!}\, x^{n-k}, & \text{kun } x > 0, \quad k = 1 \ldots n - 1. \end{cases}$$

Koska derivaatat $f^{(k)}(0) = 0$, $k = 1 \ldots n - 1$ ovat $\mathbb{R}$:ssä jatkuvia ja koska $f^{(n-1)}$ ei ole derivoituva pisteessä $x = 0$, niin päätellään, että f on $\mathbb{R}$:ssä täsmälleen $n - 1$ kertaa (ei n kertaa) jatkuvasti derivoituva. $\square$

Jos f on m kertaa mutta ei $m + 1$ kertaa jatkuvasti derivoituva välillä A, niin indeksiä m kutsutaan usein f:n *sileysasteeksi* (engl. degree of smoothness), ja voidaan myös sanoa, että f on *sileä astetta* m ko. välillä. Tällöin 'sileä astetta

nolla' tarkoittaa siis pelkkää jatkuvuutta. Esimerkissä f:n sileysaste ($\mathbb{R}$:ssä) on $m = n - 1$. Jos sileysasteella ei ole ylärajaa, ts. f on 'äärettömän sileä', niin sileysasteeksi voidaan merkitä $m = \infty$. Termillä *sileä* (ilman lisämääreitä) on usein juuri tämä merkitys. Esimerkissä on $m = \infty$ esim. välillä $[0, \infty)$.

Esimerkki 6 Polynomi on sileä (= 'äärettömän sileä') jokaisella välillä, eli $\mathbb{R}$:ssä, samoin trigonometriset funktiot sin, cos ja Arctan. Trigonometrinen funktio tan on sileä väleillä $((n - \frac{1}{2})\pi, (n + \frac{1}{2})\pi)$, $n \in \mathbb{Z}$. $\square$

Esimerkki 7 Potenssisarjan summana määritelty funktio on sileä välillä $(-\rho, \rho)$ (eli väleillä $[a, b] \subset (-\rho, \rho)$), missä ρ = sarjan suppenemissäde, vrt. Luku V.3. $\square$

Edellä Esimerkissä 5 määritelty funktio f on *paloittain sileä* (engl. piecewise smooth) millä tahansa indeksillä $m \in \mathbb{N}$ mitaten. Tämä tarkoittaa, että olipa $m \in \mathbb{N}$ mikä tahansa, niin f täyttää seuraavan määritelmän ehdot.

Määritelmä V.5.7 Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$ on m *kertaa paloittain jatkuvasti derivoituva* välillä $[a, b]$, jos $\exists$ pisteet c_k, $k = 0 \ldots n$, $n \in \mathbb{N}$ ja funktiot f_k, $k = 1 \ldots n$ siten, että pätee

(i) $a = c_0 < c_1 < \ldots < c_n = b$ ja $(c_{k-1}, c_k) \subset \mathcal{D}_f, \quad k = 1 \ldots n,$

(ii) $[c_{k-1}, c_k] \subset \mathcal{D}_{f_k}$ ja f_k on m kertaa jatkuvasti derivoituva välillä $[c_{k-1}, c_k]$, $k = 1 \ldots n$,

(iii) $f(x) = f_k(x)$, kun $x \in (c_{k-1}, c_k), \quad k = 1 \ldots n.$

Esimerkki 8 Jos $a < 0$ ja $b > 0$, niin Esimerkin 5 funktiolle Määritelmän V.5.7 ehdot ovat voimassa jokaisella $m \in \mathbb{N}$, kun asetetaan $n = 1$, $c_1 = 0$, $f_1(x) = 0$ ja $f_2(x) = x^n$. $\square$

Jos f on välillä $[a, b]$ paloittain sileä astetta m millä tahansa $m \in \mathbb{N}$, kuten Esimerkissä 5, niin voidaan sanoa yksinkertaisesti, että f on *paloittain sileä* ilman lisämääreitä. Myös Esimerkin 3 funktio on tätä tyyppiä.

Esimerkki 9 Funktio

$$f(x) = \begin{cases} x^3, & x < 0, \\ x^5\sqrt{x}, & x \geq 0 \end{cases}$$

on välillä $[-1, 1]$ sileä astetta $m = 2$ (täsmälleen) ja paloittain sileä astetta $m = 5$ (täsmälleen). Tässä paloittaista sileyttä rajoittaa, että f on origossa vain viidesti derivoituva oikealta. $\square$

HARJOITUSTEHTÄVIÄ

1. Tutki origon laatu mahdollisena paikallisena ääriarvopisteenä:

   a) $f(x) = 100 + x^{99} - x^{98}$ b) $f(x) = x^4 - x^3 \cos x$

   c) $f(x) = \begin{cases} x^3 \sin \dfrac{1}{x}, & \text{kun } x \neq 0, \\ 0, & \text{kun } x = 0 \end{cases}$ d) $f(x) = \begin{cases} 0.000001, & \text{kun } x = 0, \\ x \cos \dfrac{1}{x}, & \text{kun } x \neq 0 \end{cases}$

2. Määritä seuraavien funktioiden maksimi- ja minimiarvot annetulla välillä sekä pisteet joissa maksimi/minimi saavutetaan:

   a) $f(x) = (x-1)^3(x+1)^3$, väli $[-2, 4]$

   b) $f(x) = |x^5 - 80x + 1|$, väli $[-1, 1]$

   c) $f(x) = \dfrac{2-x}{5-4x+x^2}$, väli $[-100, 2]$

   d) $f(x) = \begin{cases} 8x^2 - 16x - 14, & \text{kun } x \leq 2, \\ 3x^2 - 24x + 22, & \text{kun } x > 2, \end{cases}$ väli $[0, 5]$

   e) $f(x) = \min\{2x^3 - 2x,\ 3 - 2x\}$, väli $[-1, 2]$

   f) $f(x) = |\sin x| - \cos x$, väli $[0, 2\pi]$

   g) $f(x) = x + 5 \sin x$, väli $[0, 2\pi]$

3. Mikä on lausekkeen $\sqrt{x} + 2\sqrt{y}$ pienin arvo, kun $x, y \geq 0$ ja $x + y = 5/6$?

4. Suoran tien varressa, kohtisuorassa tietä vastaan, on mainostaulu, jonka leveys on 10 metriä ja lähin etäisyys tiellä kulkevien ajoradasta on 20 m. Mikä on suurin kulma, jossa tiellä liikkuja taulun näkee?

5. Piste $P = (x, y)$ sijaitsee Cartesiuksen lehdellä $S : x^3 + y^3 = 3xy$. Mikä on suurin mahdollinen x:n arvo, jos $y \geq 0$?

6. Määritä

   a) funktion $f(x, y) = (2x-y)(x+y-1)(x-y-1)$ maksimi- ja minimiarvot janalla, jonka päätepisteet ovat $(-1, 0)$ ja $(3, 1)$,

   b) funktion $f(x, y) = |x^2 + 2y| + 2x^2$ pienin arvo käyrällä $S :\ xy = 1$,

   c) funktion $f(x, y, z) = x + y + z(2x + y + z)(x + 3y + z)(x + y + 4z)$ pienin arvo suoralla $S :\ x = 2y = 3z$.

7. Funktio f on määritelty yksikkökiekon neljänneksessä $A = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1 \ \& \ x \geq 0 \ \& \ y \geq 0\}$ seuraavasti (napakoordinaatit!): $f(r,\phi) = (3r^2 - 2r)(2\phi^2 - 3\phi + 1)$. Missä A:n pisteissä f saavuttaa suurimman ja missä pienimmän arvonsa?

8. Määritä napa- tai pallokoordinaatistoon siirtymällä seuraavien funktioiden maksimi- ja minimiarvot annetussa joukossa A. Anna myös karteesisessa koordinaatistossa pisteet, joissa nämä arvot saavutetaan.

   a) $f(x,y) = x^3y^4, \quad A = \{(x,y) \mid x^2 + y^2 \leq 1\}$
   b) $f(x,y) = x^2y^4, \quad A = \{(x,y) \mid x^2 + y^2 \leq 25\}$
   c) $f(x,y,z) = xyz^3, \quad A = \{(x,y,z) \mid x^2 + y^2 + z^2 \leq 1\}$

9. Mikä on funktion
$$f(x) = |x-3|^{111/11} + (x+1)^3|x+1| + \max\{1, x^2 - 2x + 2\}$$
sileysaste välillä a) $[-2,0]$, b) [0,2], c) [2,4] ?

10. Määritä seuraavien funktioiden sileysasteet välillä $[-1,1]$:

    a) $x^2|\sin x|$ b) $|x|^3(1-\cos x)$ c) $\sin^2 x|\tan x|$

    d) $f(x) = \begin{cases} 1 - \frac{1}{2}x^2, & \text{kun } x < 0, \\ \cos x, & \text{kun } x \geq 0 \end{cases}$ e) $f(x) = \begin{cases} \sin x, & \text{kun } x \leq 0, \\ x - \frac{1}{6}x^3, & \text{kun } x > 0 \end{cases}$

11. Olkoon $f_1(x) = 2x^3 - 7x^2 + 9x$ ja $f_2(x) = ax^2 + bx + c$. Määritä kertoimet a, b, c siten, että funktio
$$f(x) = \begin{cases} f_1(x), & \text{kun } x \leq 1, \\ f_2(x), & \text{kun } x > 1 \end{cases}$$
on kahdesti jatkuvasti derivoituva välillä $[0,2]$. Piirrä samaan kuvaan f:n kuvaaja ko. välillä sekä katkoviivalla f_2:n kuvaaja välillä $[0,1]$ ja f_1:n kuvaaja välillä $[1,2]$.

12. (*) a) Funktiosta $f : \mathbb{R} \to \mathbb{R}$ tiedetään, että f on derivoituva $\mathbb{R}$:ssä ja että
$$\lim_{x\to-\infty} f(x) = A_- \,, \quad \lim_{x\to\infty} f(x) = A_+ \,,$$
missä $A_- \in \mathbb{R}$ ja $A_+ \in \mathbb{R}$. Näytä, että f saavuttaa $\mathbb{R}$:ssä absoluuttisen maksimiarvon täsmälleen kun jossakin f:n kriittisessä pisteessä c on $f(c) \geq \max\{A_-\,, A_+\}$. b) Millä a:n arvoilla funktio
$$f(x) = \frac{x^4 + ax^2}{(x^2+7)^2}$$
saa absoluuttisen maksimiarvon jollakin $x \in \mathbb{R}$?

13. (*) Määritä funktion $f(x) = 4|\cos x| - 3\sin x + 2\cos x$ kriittiset pisteet, derivaatan epäjatkuvuuskohdat, paikalliset ääriarvokohdat ja absoluutiset maksimi- ja minimiarvot välillä $[0, 2\pi]$. Hahmottele käyrän $y = f(x)$ kulku.

14. (*) Määrittele välillä $(0, \infty)$ funktio g siten, että funktion $f(x) = x + a\sin x$ pienin arvo välillä $[0, a]$ on $f_{min} = g(a)$. Hahmottele g:n kuvaaja.

15. (*) a) Olkoon p_1 ja p_2 polynomeja astetta $\leq n$. Näytä, että jos funktio

$$f(x) = \begin{cases} p_1(x), & \text{kun } x < a, \\ p_2(x), & \text{kun } x \geq a \end{cases}$$

on n kertaa jatkuvasti derivoituva välillä $[a-1, a+1]$, niin $p_1 = p_2$.

b) Funktio f olkoon m kertaa jatkuvasti derivoituva välillä $[a, b]$. Näytä, että on olemassa yksikäsitteiset polynomit p_1 ja p_2 astetta $\leq m$ siten, että funktio

$$g(x) = \begin{cases} p_1(x), & \text{kun } x < a, \\ f(x), & \text{kun } x \in [a, b], \\ p_2(x), & \text{kun } x > b \end{cases}$$

on m kertaa jatkuvasti derivoituva $\mathbb{R}$:ssä.

16. (*) (Vasikka-aitaus) Maanviljelijä haluaa aidata navettansa viereen suorakulmion muotoisen aitauksen vasikoiden laidunmaaksi. Aitauksen mitat valitaan aidan kokonaispituuden (x m) funktiona siten, että laitumen pinta-ala on mahdollisimman suuri ja navetan 40 m:n pituinen seinä käytetään hyväksi mahdollisimman hyvin. Näin suunnitellussa aitauksessa olkoon $f(x)$ navetan seinän suuntaisen ja seinää vastaan kohtisuoran sivun pituuksien suhde. Määritä $f(x)$ välillä $(0, \infty)$. Mikä on f:n sileysaste välillä $[0, \infty)$, kun asetetaan $f(0) = f(0^+)$?

17. (*) (Pisin heitto?) Jos ilmanvastusta ei huomioida, niin ilmaan heitetyn kappaleen lentorata noudattaa heittoparaabelia

$$y = h + kx - (1 + k^2)\frac{x^2}{2a},$$

missä x mittaa vaakasuoraa etäisyyttä lähtöpisteestä, y korkeutta maan pinnan tasosta ja vakiot h, k ja a ovat yksittäiselle heitolle ominaisia (ks. Luku IV.4, Esimerkki 6). Jos h ja a kiinnitetään, niin millä k:n arvolla kappale lentää pisimmälle — ja kuinka pitkälle? — ennen kuin törmää maahan?

V.6 Differentiaalilaskun väliarvolause

Seuraava lause kuuluu matemaattisen analyysin huomattavimpiin ja myös hyvin usein käytettyihin perustuloksiin. Lause on väliarvolauseiden sarjan toinen — vrt. Ensimmäinen väliarvolause Luvussa V.1 (Lause V.1.10).

LAUSE V.6.1 **(Toinen väliarvolause – Differentiaalilaskun väliarvolause)** Jos $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on jatkuva välillä $[a, b] \subset \mathcal{D}_f$ ja derivoituva välillä (a, b), niin jollakin $\xi \in (a, b)$ on

$$f(b) - f(a) = f'(\xi)(b - a).$$

Lauseen väittämä on geometrisesti hyvin uskottava (vrt. kuvio), mutta todistuksessa joudutaan kuitenkin tekemisiin jatkuvuuden syvällisemmän logiikan kanssa.

Lauseen V.6.1 todistamiseksi palautetaan väittämä ensinnäkin yksinkertaisemmaksi määrittelemällä

$$g(x) = f(x) - f(a) \cdot \frac{b - x}{b - a} - f(b) \cdot \frac{x - a}{b - a}, \quad x \in [a, b].$$

Tällöin on $g(a) = g(b) = 0$ ja pätee

$$f(b) - f(a) = f'(\xi)(b - a) \quad \Leftrightarrow \quad g'(\xi) = 0.$$

Lauseen V.6.1 väittämä saa näin seuraavan pelkistetymmän muodon (vrt. Ensimmäinen väliarvolause ja sen pelkistetty muoto, Bolzanon lause, Luvussa V.1) :

LAUSE V.6.2 **(Rollen lause)** Jos f on jatkuva välillä $[a, b]$ ja derivoituva välillä (a, b) ja $f(a) = f(b) = 0$, niin $f'(\xi) = 0$ jollakin $\xi \in (a, b)$.

<u>Todistus</u> Koska f on jatkuva välillä $[a, b]$, niin Weierstrassin lauseen (Lause V.1.12) perusteella f saavuttaa minimi- ja maksimiarvonsa välillä $[a, b]$. Kun suljetaan pois ilmeinen tapaus $f(x) = 0 \ \forall x \in [a, b]$ (jolloin $f'(\xi) = 0 \ \forall \xi \in (a, b)$), niin on oltava

$$\begin{aligned} \text{joko:} \quad & f(\xi) = f_{\min} < 0, \quad \xi \in (a, b), \\ \text{tai:} \quad & f(\xi) = f_{\max} > 0, \quad \xi \in (a, b). \end{aligned}$$

Kummassakin tapauksessa on oltava $f'(\xi) = 0$ (Lause V.5.2). Lause V.6.2 on näin todistettu, ja tämän välittömänä seurauksena myös Lause V.6.1. — Huomattakoon, että todistus oli verrattain mutkaton vain siksi, että siinä oli 'kova ydin' (Weierstrassin lause). $\square$

Esimerkki 1 Funktio $f(x) = \sqrt{|x|}$ toteuttaa Lauseen V.6.1 ehdot välillä $[0,1]$, ja väitetty ξ on yksikäsitteinen: $\xi = \frac{1}{4} \in (0,1)$. Välillä $[-1,1]$ Lauseen V.6.1 ehdot eivät toteudu, koska f ei ole derivoituva pisteessä $x = 0$. Väitettyä pistettä ξ (jossa olisi oltava $f'(\xi) = 0$) ei tässä tapauksessa myöskään ole. $\square$

Differentiaalilaskun väliarvolauseella on hyvin monia käyttömuotoja tutkittaessa derivoituvien (tai 'melkein derivoituvien') funktioiden ominaisuuksia. Jatkossa esitellään näistä käyttömuodoista keskeisimmät.

Funktion monotonisuus

Yksinkertaisissa tapauksissa voidaan pelkin algebran keinoin selvittää, millä väleillä annettu funktio on monotoninen (kasvava tai vähenevä, vrt. Luku IV.1). Yleisemmin tehtävä helpottuu huomattavasti, kun otetaan käyttöön seuraava Lauseesta V.6.1 johdettava väittämä.

Lause V.6.3 Olkoon f jatkuva välillä $[a,b]$ ja olkoon $X \subset (a,b)$ äärellinen pistejoukko siten, että pätee

(1) f on derivoituva jokaisessa pisteessä $x \in (a,b),\ x \notin X$.

(2) On voimassa $(\star)$ $\forall x \in (a,b),\ x \notin X$, missä $(\star)$ on jokin seuraavista vaihtoehdoista:

$$\begin{array}{llll} \text{(a)} & f'(x) \geq 0, & \text{(b)} & f'(x) > 0, \\ \text{(c)} & f'(x) \leq 0, & \text{(d)} & f'(x) < 0. \end{array}$$

Tällöin f on (a) kasvava, (b) aidosti kasvava, (c) vähenevä, (d) aidosti vähenevä välillä $[a,b]$.

<u>Todistus</u> Ol. $x_1, x_2 \in [a,b]$, $x_1 < x_2$. Koska joukko X on äärellinen, niin voidaan valita pistejoukko $\{t_1, \ldots, t_{n-1}\} \subset X$ (mahdollisesti tyhjä joukko) siten, että

$$x_1 = t_0 < t_1 < \ldots < t_n = x_2 \ \text{ ja } \ (t_{k-1}, t_k) \cap X = \emptyset,\ k = 1 \ldots n.$$

Tällöin kun kirjoitetaan $f(x_2) - f(x_1)$ teleskooppisummaksi ja sovelletaan Lausetta V.6.1, niin seuraa

$$f(x_2) - f(x_1) = \sum_{k=1}^{n} [f(t_k) - f(t_{k-1})] = \sum_{k=1}^{n} f'(\xi_k)(t_k - t_{k-1}), \quad \xi_k \in (t_{k-1}, t_k),$$

jolloin tehtyjen oletuksien perusteella päätellään

$$f(x_2) - f(x_1) = \begin{cases} \geq 0 & \text{(a)}, \\ > 0 & \text{(b)}, \\ \leq 0 & \text{(c)}, \\ < 0 & \text{(d)}. \end{cases} \qquad \square$$

Sovellettaessa Lausetta V.6.3 voidaan joukkoon X aina lukea f':n nollakohdat, sikäli kuin niitä on äärellinen määrä. Äärellinen määrä f':n nollakohtia ei siis häiritse funktion (aitoakaan) monotonisuutta, kunhan f':n merkki ei nollakohdissa vaihdu.

Esimerkki 2 Funktio $f(x) = \begin{cases} x^3, & \text{kun } -1 \leq x \leq 1, \\ x, & \text{kun } \ 1 < x \leq 2 \end{cases}$

on jatkuva välillä $[-1, 2]$ ja derivoituva välillä $(-1, 2)$ lukuunottamatta pistettä $x = 1$. Kun valitaan $X = \{0, 1\}$, niin Lauseen V.6.3 oletus (b) on voimassa, joten f on välillä $[-1, 2]$ aidosti kasvava. $\square$

Esimerkki 3 Edellisen luvun Esimerkin 2 mukaan funktiolle $f(x) = x^2/(x+1)$ pätee: $f'(x) > 0$ väleillä $(-\infty, -2)$ ja $(0, \infty)$ ja $f'(x) < 0$ väleillä $(-2, -1)$ ja $(-1, 0)$. Lauseen V.6.3 perusteella päätellään, että f on aidosti kasvava väleillä $(-\infty, -2]$ ja $[0, \infty)$ (eli näiden välien suljetuilla osaväleillä) ja vastaavasti aidosti vähenevä väleillä $[-2, -1)$ ja $(-1, 0]$. (Samaan tulokseen tullaan pelkin algebran keinoinkin, mutta työläämmin: Harj.teht. IV.1:15b.) $\square$

Kriittisen pisteen laatu

Lauseen V.6.3 monotonisuuskriteereillä voidaan yleensä selvittää helposti, onko funktion kriittinen piste paikallinen ääriarvokohta (ja minkälaatuinen) vai ei. Nimittäin asia selviää (ellei f ole poikkeuksellisen 'pahantapainen') tutkimalla derivaatan merkkiä kriittisen pisteen lähiympäristössä. Esimerkiksi jos $f'(x) < 0$ välillä $(c - \delta_1, c)$ ja $f'(x) > 0$ välillä $(c,\ c + \delta_2)$ joillakin $\delta_1, \delta_2 > 0$, niin f on välillä $[c - \delta_1, c]$ aidosti vähenevä ja välillä $[c, c + \delta_2]$ aidosti kasvava, jolloin c:n on oltava paikallinen minimikohta. Päättelyssä riittää, että f on pisteessä c jatkuva, ts. derivoituvuutta ei tarvitse olettaa (vrt. Lause V.5.5).

Esimerkki 3 (jatko) Esimerkissä f:n kriittiset pisteet ovat -2 ja 0. Derivaatan merkin perusteella päätellään, että $x = -2$ on f:n paikallinen maksimikohta ja $x = 0$ on paikallinen minimikohta. $\square$

Funktion (käyrän) kaareutuvuus

Jos $f'(x)$:n merkki kertoo, onko f kasvava tai vähenevä, niin $f''(x)$:n merkki puolestaan kertoo funktion f (tai käyrän $y = f(x)$) *kaareutumissuunnan.* Jos f on derivoituva avoimella välillä (a, b), niin sanotaan, että f on ko. välillä *ylös*(päin) *kaareutuva* (egl. concave up), jos f' on ko. välillä aidosti kasvava, ja *alas*(päin) *kaareutuva* (engl. concave down), jos f' on aidosti vähenevä välillä (a, b). Jos f on kahdesti derivoituva, lukuunottamatta mahdollisesti äärellistä pistejoukkoa, niin f:n kaareutumissuunta voidaan päätellä f'':n merkistä Lauseen V.6.3 mukaisesti. Kaareutuvuuden geometrinen tulkinta on ilmeinen, vrt. kuvio.[†]

$f'' > 0$ (ylös kaareutuva) $\qquad$ $f'' < 0$ (alas kaareutuva)

Pistettä, jossa kaareutumissuunta vaihtuu, sanotaan (funktion/käyrän) *käännepisteeksi* (engl. inflection point). Jos f'' on jatkuva käännepisteessä c, niin on oltava $f''(c) = 0$.

Esimerkki 4 Derivoimalla todetaan, että kolmannen asteen polynomilla $f(x) = x^3 + ax^2 + bx + c$ $(a, b, c \in \mathbb{R})$ on täsmälleen yksi käännepiste: $x = -\frac{1}{3}a$. Välillä $(-\infty, -\frac{1}{3}a)$ f on alas ja välillä $(-\frac{1}{3}a, \infty)$ ylös kaareutuva. $\square$

[†]Kaareutuvuuden synonyyminä käytetään matemaattisissa teksteissä myös termiä *kuperuus* (ylös tai alas), mutta tällöin saattaa jäädä epäselväksi, kumpaa kaareutumissuuntaa tarkoitetaan. Kaareutuvuuden (kuperuuden) käsite voidaan määritellä yleisemmin myös derivaatoista riippumatta, jolloin käytetään useammin termejä *konveksi* ja *konkaavi.* Funktiota f sanotaan konveksiksi välillä $[a, b] \in \mathcal{D}_f$, jos f:llä on ominaisuus

$$f\big(tx_1 + (1-t)x_2\big) \le tf(x_1) + (1-t)f(x_2), \quad \text{kun } x_1, x_2 \in [a, b] \text{ ja } t \in [0, 1].$$

Jos tässä ehdossa epäyhtälö toteutuu muodossa '$<$' aina kun $x_1 \neq x_2$ ja $t \in (0, 1)$, niin f on välillä $[a, b]$ *aidosti konveksi.* Geometrisesti aito konveksisuus tarkoittaa, että jos x_1, $x_2 \in [a, b]$ ja $x_1 < x_2$, niin funktion f kuvaaja on pisteiden $(x_1, f(x_1))$ ja $(x_2, f(x_2))$ kautta kulkevan suoran alapuolella avoimella välillä (x_1, x_2). Välillä $[a, b]$ jatkuva ja välillä (a, b) ylös kaareutuva (riittävästi derivoituva) funktio on määritelmien mukaisesti aidosti konveksi. Alaspäin kaareutuvuutta vastaava käsite konkaavius määritellään vastaavasti.

Lipschitz-jatkuvuus

Luvussa V.1 määritelty funktion jatkuvuus voidaan tulkita funktiota koskevaksi minimioletukseksi, kun halutaan taata fuktioevaluaation $x \mapsto f(x)$ luotettavuus. Seuraavassa esitellään 'pelkkää jatkuvuutta' vahvempi jatkuvuuden laji, jolla jatkossa on silloin tällöin (etenkin teoreettista) käyttöä.

MÄÄRITELMÄ V.6.4 Funktio f on *Lipschitz-jatkuva* eli *Lipschitz*† *välillä* $[a, b] \subset \mathcal{D}_f$, jos jollakin $L \in \mathbb{R}_+$ pätee

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \quad \forall x_1, x_2 \in [a, b].$$

Määritelmän lukua L sanotaan f:n *Lipschitz-vakioksi* välillä $[a, b]$. Vakio L ei ole yksikäsitteinen, sillä jos L on f:n Lipschitz-vakio, niin määritelmän mukaan samoin on jokainen $L_1 > L$. Jos f on Lipschitz välillä $[a, b]$, niin f:n pienin Lipschitz-vakion arvo on pienin yläraja (määritelmän perusteella rajoitetulle) reaalilukujoukolle

$$A = \left\{ y = \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} \;\middle|\; x_1, x_2 \in [a, b] \;\;\&\;\; x_1 \neq x_2 \right\},$$

ts. $L_{min} = \sup A$ (vrt. Luku I.11).

Lipschitz-jatkuvuus on käsitteenä lähellä derivoituvuutta, ja Lipschitz-jatkuvia funktioita voikin luonnehtia 'melkein derivoituviksi'.

ESIMERKKI 5 Funktio $f(x) = |x|$ on jokaisella välillä Lipschitz-jatkuva vakiolla $L = 1$ (= pienin L:n arvo), sillä kolmioepäyhtälön nojalla

$$|f(x_1) - f(x_2)| \;=\; ||x_1| - |x_2|| \;\leq\; |x_1 - x_2|, \quad x_1,\, x_2 \in \mathbb{R}. \quad \square$$

Esimerkin funktio on myös paloittain jatkuvasti derivoituva jokaisella välillä (Määritelmä V.5.7, $m = 1$). Yhdessä jatkuvuuden kanssa tämä on yleisemminkin riittävä tae Lipschitz-jatkuvuudelle:

LAUSE V.6.5 Jos f on välillä $[a, b]$ jatkuva ja paloittain jatkuvasti derivoituva, niin f on Lipschitz-jatkuva välillä $[a, b]$.

Todistuksen (Harj.teht. 3) johdannoksi näytettäköön toteen seuraava väittämä, jonka osavättämän (ii) Lause V.6.5 yleistää. Todistus on tyypillinen esimerkki Diferentiaalilaskun väliarvolauseen soveltamisesta.

†**Rudolf Lipschitz** (1832-1903) oli saksalainen matemaatikko.

LAUSE V.6.6 Jos f on välillä $[a,b]$ jatkuvasti derivoituva, niin

(i) $\mathrm{D}_+f(a) = f'(a^+)$, $\mathrm{D}_-f(b) = f'(b^-)$,

(ii) f on välillä $[a,b]$ Lipschitz-jatkuva vakiolla, jonka pienin arvo on $L = \max\limits_{x\in[a,b]} |f'(x)|$.

Todistus (i) Olkoon $\{x_n\}$ jono, jolle pätee $a < x_n \le b$ $\forall n$ ja $x_n \to a^+$. Tällöin Lauseen V.6.1 mukaan jokaisella n on olemassa $\xi_n \in (a, x_n)$ siten, että

$$\frac{f(x_n) - f(a)}{x_n - a} = f'(\xi_n).$$

Tässä $\xi_n \to a^+$ kun $x_n \to a^+$, joten toispuolisen derivaatan $\mathrm{D}_+f(a)$ määritelmän ja f':n oletetun toispuolisen jatkuvuuden perusteella seuraa

$$\mathrm{D}_+f(a) = \lim_{x_n\to a^+} \frac{f(x_n) - f(a)}{x_n - a} = \lim_{x_n\to a^+} f'(\xi_n) = f'(a^+).$$

Väittämän (i) toinen osa todistetaan vastaavasti.

(ii) Jos $x_1, x_2 \in [a,b]$ ja $x_1 < x_2$, niin Lauseen V.6.1 mukaan

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \quad \xi \in (x_1, x_2).$$

Koska f' on jatkuva välillä $[a,b]$, niin Weierstrassin lauseen (Lause V.1.12) mukaan $|f'(x)|$ saavuttaa välillä $[a,b]$ maksimiarvonsa. Kun tämä merkitään $= L$, niin nähdään, että L kelpaa f:n Lipschitz-vakion arvoksi. Valitsemalla x_1 ja x_2 $|f'(x)|$:n maksimikohdan läheltä (vrt. osaväittämän (i) todistus) nähdään myös helposti, että tämä L:n arvo on pienin mahdollinen. $\square$

ESIMERKKI 6 Jos $0 < a < b$, niin funktio $f(x) = \sqrt{x}$ on Lauseen V.6.6 mukaan välillä $[a,b]$ Lipschitz-jatkuva vakiolla $L = 1/(2\sqrt{a})$. Väleillä $[0,b]$ f ei ole Lipschitz-jatkuva (ainoastaan jatkuva), sillä jos valitaan $x_1 = 0$ ja $x_2 > 0$, niin

$$\frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} = \frac{1}{\sqrt{x_2}} \to \infty, \quad \text{kun } x_2 \to 0^+. \quad \square$$

Differentiaaliyhtälö $y' = f(x)$. Integraalifunktio

Derivoimissääntöjen perusteella 'vakion derivaatta on nolla', ts. jos $f(x) = C$, $x \in (a,b)$ jollakin $C \in \mathbb{R}$, niin $f'(x) = 0$, $x \in (a,b)$. Differentiaalilaskun väliarvolauseesta seuraa, että väittämä pätee myös käänteisesti muodossa: vain vakion derivaatta $= 0$. Tämä yksinkertainen tulos osoittautuu seuraamuksiltaan huomattavaksi.

Lause V.6.7 (**Integraalilaskun**[†] **peruslause**) Jos f on derivoituva välillä (a,b) ja $f'(x) = 0\ \forall x \in (a,b)$, niin jollakin $C \in \mathbb{R}$ on $f(x) = C,\ x \in (a,b)$.

Todistus Jos $c \in (a,b)$ ja $x \in (a,b),\ x \neq c$, niin Differentiaalilaskun väliarvolauseen oletukset ovat voimassa välillä $[c,x]$ $(x > c)$ tai välillä $[x,c]$ $(x < c)$. Näin ollen $f(x) - f(c) = f'(\xi)(x-c)$ jollakin $\xi \in (c,x) \subset (a,b)$ tai $\xi \in (x,c) \subset (a,b)$. Koska $f'(\xi) = 0\ \forall \xi \in (a,b)$, niin seuraa $f(x) = f(c) = C,\ x \in (a,b)$. $\square$

Lause V.6.7 voidaan tulkita väittämänä, joka koskee hyvin yksinkertaista (itse asiassa kaikkein yksinkertaisinta) *differentiaaliyhtälöä*

$$y' = 0.$$

Tässä $y(x)$ on tarkasteltavalla välillä (a,b) derivoituva (tuntematon) funktio[‡]. Differentiaaliyhtälön *ratkaisu* on jokainen funktio $y(x)$, joka toteuttaa yhtälön ko. välillä. Lauseen V.6.7 mukaan diferentiaaliyhtälön $y' = 0$ jokainen ratkaisu on vakio, eli ko. differentiaaliyhtälön *yleinen ratkaisu* välillä (a,b) on

$$y(x) = C, \quad x \in (a,b),$$

missä C on nk. *määräämätön vakio* $(C \in \mathbb{R})$.

Em. tulos on helposti yleistettävissä koskemaan differentiaaliyhtälöä

$$y'(x) = f(x), \quad x \in (a,b),$$

missä f on tunnettu funktio välillä (a,b).

Korollaari V.6.8 Jos F on välillä (a,b) derivoituva funktio ja $F'(x) = f(x)$, $x \in (a,b)$, niin differentiaaliyhtälön $y' = f(x)$ yleinen ratkaisu välillä (a,b) on

$$y(x) = F(x) + C, \quad x \in (a,b).$$

Todistus Jos y differentiaaliyhtälön ratkaisu ja merkitään $u(x) = y(x) - F(x)$, niin $u'(x) = f(x) - f(x) = 0$, $x \in (a,b)$, jolloin Lauseen V.6.7 mukaan on oltava

$$u(x) = C \ \Leftrightarrow\ y(x) = F(x) + C, \quad x \in (a,b). \quad \square$$

Esimerkki 7 Differentiaaliyhtälön

$$y' = x^3 - 2x, \quad x \in \mathbb{R}$$

[†] *Integraalilaskenta* on matematiikan laji, jota tarkastellaan perusteellisemmin Luvussa VIII.

[‡] Differentiaaliyhtälöissä käytetään yleensä funktiosymbolia y (ei f), koska kyseeessä on epäsuorasti määritelty funktio (tai funktiojoukko), vrt. implisiittifunktiot (Luku IV.2).

(tässä $(a,b) = (-\infty,\infty)$) ratkaisemiseksi on etsittävä funktio F, jolle pätee $F'(x) = x^3 - 2x$, $x \in \mathbb{R}$. Derivoimissääntöjen perusteella nähdään, että voidaan valita $F(x) = \frac{1}{4}x^4 - x^2$, joten yleinen ratkaisu on

$$y(x) = \frac{1}{4}x^4 - x^2 + C. \quad \square$$

Esimerkki 8 Mikä on differentiaaliyhtälön

$$y'' = 0, \quad x \in \mathbb{R}$$

yleinen ($\mathbb{R}$:ssä kahdesti derivoituva) ratkaisu?

Ratkaisu Jos merkitään $u(x) = y'(x)$, niin differentiaaliyhtälö pelkistyy muotoon $u' = 0$, joten $u(x) = C_1$, $x \in \mathbb{R}$ (LauseV.6.7), ja näin ollen $y' = C_1$, $x \in \mathbb{R}$. Korollaarin V.6.8 ja derivoimissääntöjen mukaan on tällöin oltava

$$y(x) = C_1 x + C_2, \quad x \in \mathbb{R}.$$

Tässä C_1 ja C_2 ovat molemmat määräämättömiä (toisistaan riippumattomia) vakioita, joten yleinen ratkaisu koostuu kaikista polynomeista astetta ≤ 1. $\square$

Differentiaaliyhtälön $y' = f(x)$ ratkaisuja sanotaan funktion f *integraalifunktioiksi.* Integraalifunktio on siis määräämätöntä vakiota, nk. *integroimisvakiota* vaille yksikäsitteinen. Sovelluksissa integroimisvakio määräytyy usein (sovelluksesta peräisin olevasta) *alkuehdosta* muotoa $y(x_0) = y_0$. Kyseessä on tällöin nk. *alkuarvotehtävä* (engl. initial value problem)

$$\begin{cases} y' = f(x), & x \in (a,b), \\ y(x_0) = y_0, \end{cases}$$

missä $x_0 \in (a,b)$ ja $y_0 \in \mathbb{R}$ on annettu. Ratkaisu on

$$y(x) = F(x) - F(x_0) + y_0,$$

missä F on (mikä tahansa) f:n integraalifunktio. Ratkaisu on yksikäsitteinen, sillä $F(x) - F(x_0)$ ei muutu, jos F:ään lisätään vakio.

Esimerkki 8 (jatko) Alkuarvotehtävän

$$\begin{cases} y'' = 0, & x \in \mathbb{R}, \\ y(1) = 0, & y'(1) = 2 \end{cases}$$

yksikäsitteinen ratkaisu on $y(x) = 2x - 2$. $\square$

Annetun funktion f 'integroiminen' eli integraalifunktion etsiminen on hyvin perinteinen matematiikan taitolaji, jonka laajempi esittely kuuluu integraalilaskun yhteyteen (ks. Luvut VIII.1–VIII.3 jäljempänä). Toistaiseksi todettakoon ainoastaan, että derivoimissääntöjen perusteella helposti integroitavissa ovat mm. trigonometriset funktiot sin ja cos, samoin polynomit (vrt. Esimerkki 7). Myös funktio, joka on määritelty välillä $(-\rho, \rho)$ suppenevan potenssisarjan summana on ko. välillä (yllättävänkin) helposti integroitavissa. Esimerkki valaiskoon asiaa.

Esimerkki 9 Funktion $f(x) = x^2/(1-x)$ integraalifunktio ei ole (lausekkeena) esitettävissä toistaiseksi tunnettujen funktioiden avulla. Välillä $(-1, 1)$ on integraalifunktio kuitenkin määrättävissä, sillä tällä välillä pätee

$$f(x) = \frac{x^2}{1-x} = x^2 \sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} x^{k+2}, \quad x \in (-1, 1).$$

Potenssisarjan derivoimissäännön (Lause V.3.4) perusteella päätellään, että f:n integraalifunktio välillä $(-1, 1)$ on

$$F(x) \;=\; C + \sum_{k=0}^{\infty} \frac{x^{k+3}}{k+3} \;=\; C + \frac{x^3}{3} + \frac{x^4}{4} + \ldots \quad \square$$

l'Hospitalin säännöt

l'Hospitalin säännöillä tarkoitetaan derivointiin perustuvia raja-arvojen laskusääntöjä muotoa

$$\lim_{x \to a^{\pm}} \frac{f(x)}{g(x)} = \lim_{x \to a^{\pm}} \frac{f'(x)}{g'(x)}\,.$$

Hankalien raja-arvojen laskemisessa nämä säännöt ovat varteenotettava — myös helposti muistettava — vaihtoehto esim. muuttujan vaihdolle (vrt. Luku V.2). Säännöt pätevät lievin lisäehdoin sellaisissa tapauksissa, joissa $f(a)/g(a)$ on joko muotoa $0/0$ tai ∞/∞.

l'Hospitalin sääntöjen taustalla on jälleen Differentiaalilaskun väliarvolause, tarkemmin seuraava nk. *yleistetty väliarvolause*, joka sekin on Rollen lauseesta johdettavissa (Harj.teht. 9).

Lause V.6.9 Jos f ja g ovat molemmat jatkuvia välillä $[a, b]$ ja derivoituvia välillä (a, b) ja lisäksi $g'(x) \neq 0$ $\forall x \in (a, b)$, niin on olemassa $\xi \in (a, b)$ siten, että

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}\,.$$

LAUSE V.6.10 (**l'Hospitalin**† **säännöt**)

1. Olkoot f ja g derivoituvia välillä $(a, a+\delta)$, $\delta > 0$, ja olkoon $g'(x) \neq 0$ tällä välillä. Edelleen olkoon $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$. Tällöin pätee
$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} = A$$
sikäli kuin raja-arvo oikealla on olemassa ($A \in \mathbb{R}$ tai $A = \pm\infty$).

2. Olkoot f ja g derivoituvia välillä (M, ∞), $M \in \mathbb{R}$, ja olkoon $g'(x) \neq 0$ tällä välillä. Edelleen olkoon $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = 0$. Tällöin pätee
$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = A$$
sikäli kuin raja-arvo oikealla on olemassa ($A \in \mathbb{R}$ tai $A = \pm\infty$).

3. Säännöt 1–2 ovat päteviä myös, kun funktioiden f ja g raja-arvoja koskevat oletukset muutetaan muotoon: $\lim |f(x)| = \lim |g(x)| = \infty$.

<u>Todistus</u> Sääntö 1. Asetetaan $f(a) = g(a) = 0$, jolloin oletusten perusteella f ja g ovat jatkuvia välillä $[a, b]$, kun $a < b < a+\delta$. Näin ollen jos $a < x < a+\delta$, niin Lauseen V.6.9 perusteella pätee jollakin $\xi \in (a, x)$
$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)}\,.$$
Tässä $\xi \to a^+$ kun $x \to a^+$, joten sääntö seuraa.

Sääntö 2. Tehdään muuttujan vaihto $x = 1/t$ ja sovelletaan 1. sääntöä:
$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{t\to 0^+} \frac{f(\frac{1}{t})}{g(\frac{1}{t})} = \lim_{t\to 0^+} \frac{-\frac{1}{t^2} f'(\frac{1}{t})}{-\frac{1}{t^2} g'(\frac{1}{t})} = \lim_{t\to 0^+} \frac{f'(\frac{1}{t})}{g'(\frac{1}{t})} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}\,.$$
Sääntö 3. Myös tämä perustuu Lauseeseen V.6.9, mutta todistus on melko työläs. Todistus sivuutetaan (ks. Harj.teht 14). □

Lauseen V.6.10 säännöillä on ilmeiset vastineensa raja-arvoille $\lim_{x\to a^-} f(x)/g(x)$ ja $\lim_{x\to-\infty} f(x)/g(x)$. Säännöt ovat samoin pätevät raja-arvolle $\lim_{x\to a} f(x)/g(x)$ ($a \in \mathbb{R}$) olettaen, että f ja g ovat derivoituvia väleillä $(a-\delta, a)$ ja $(a, a+\delta)$ ja $g'(x) \neq 0$ näillä väleillä.

†Ranskalainen markiisi ja matemaatikko **Guillaume de l'Hospital** (1661-1704) tunnetaan ennen muuta hänen julkaisemastaan ensimmäisestä differentiaalilaskennan oppikirjasta (1696).

Esimerkki 10 l'Hospitalin 1. säännöllä laskien saadaan (vrt. Esimerkki V.2:4)

$$\lim_{x\to 81} \frac{\sqrt{x}-9}{\sqrt[4]{x}-3} = \lim_{x\to 81} \frac{\frac{1}{2}x^{-1/2}}{\frac{1}{4}x^{-3/4}} = \lim_{x\to 81} 2x^{1/4} = 6.$$

Samaa sääntöä soveltaen seuraa myös

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

Tämä lasku kuitenkin kätkee kehäpäätelmän: Raja-arvoa laskettaessa käytetään derivoimissääntöä, joka perustui ko. raja-arvoon (!) (ks. Luku V.4). □

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien funktioiden kriittiset pisteet ja välit, joilla funktiot ovat aidosti kasvavia tai väheneviä. Määritä myös funktioiden absoluuttiset minimi- ja maksimiarvot määrittelyjoukossaan, sikäli kuin olemassa. Hahmottele funktioiden kuvaajat.

   a) $f(x) = 1 - 6x + 9x^2 - 5x^3$ b) $f(x) = x^4 + x$ c) $f(x) = \dfrac{9+x^2}{1+x}$

   d) $f(x) = \dfrac{x^4+x+1}{x^4+1}$ e) $f(x) = x\sqrt{4-x^2}$ f) $f(x) = \sqrt{3x^2-x^3}$

   g) $f(x) = \dfrac{x^2}{\sqrt{4-x^2}}$ h) $f(x) = \dfrac{x}{\sqrt{x^4+1}}$ i) $f(x) = x - \sin x$

   j) $f(x) = x + 2\cos x$ k) $f(x) = x - \dfrac{2}{\sin x}$ l) $f(x) = x - 2\tan x$

   m) $f(x) = 1 - \dfrac{1}{2}x^2 - \cos x$

2. a) Määritä funktion $f(x) = 3x^4 - 2x^3 + 15x^2 + 10x - 20$ pienin Lipschitz-vakio välillä $[0,6]$. b) Näytä, että jos f on Lipscitz-jatkuva välillä $[a,b]$ vakiolla L, niin käyrän $S : y = f(x)$ kaari välillä $[a,b]$ on suoristuva ja kaarenpituudelle pätee arvio $s \le \sqrt{1+L^2}\,(b-a)$. c) Näytä, että jos f ja g ovat Lipschitz-jatkuvia välillä $[a,b]$ vakioilla L_1 ja L_2, niin $\alpha f + \beta g$ on välillä $[a,b]$ Lipschitz-jatkuva vakiolla $L = |\alpha|L_1 + |\beta|L_2$ $(\alpha, \beta \in \mathbb{R})$.

3. Näytä, että jos f on jatkuva ja paloittain jatkuvasti derivoituva välillä $[a,b]$, niin f on välillä $[a,b]$ Lipschitz-jatkuva vakiolla, jonka pienin arvo on $L = \max_{x\in[a,b]} g(x)$, missä $g(a) = |\mathrm{D}_+ f(a)|$, $g(b) = |\mathrm{D}_- f(b)|$ ja $g(x) =$

$\max\{|\mathrm{D}_- f(x)|, |\mathrm{D}_+ f(x)|\}$, $x \in (a,b)$. Laske tulosta soveltaen funktioiden $f(x) = |x-1| + |x| + |x+1|$ ja $g(x) = \max\{2x^3 - 2x,\ -2x^2 - 4x,\ 3 - 2x\}$ pienin Lipschitz-vakio välillä $[-2, 2]$.

4. Määritä seuraavien funktioiden käännepisteet ja kaareutumissuunnat eri $\mathbb{R}$:n osaväleillä:

   a) $f(x) = 3x^5 + 35x^4 + 100x^2 - 200x$ b) $f(x) = 3x^5 - 10x^4 + 10x^2$
   c) $f(x) = 2x^3 + x + 1 - 1/x$ d) $f(x) = x/(x^2+1)$

5. Johda kaava $\operatorname{Arcsin} x + \operatorname{Arccos} x = \frac{\pi}{2}$, $x \in (-1,1)$ derivoimissäännöistä ja Lauseesta V.6.7.

6. Ratkaise (yleinen ratkaisu tai alkuarvotehtävän ratkaisu) joko lausekkeena tai potenssisarjan avulla:

   a) $y' = 4x^7 - 5x^3 + 3$, $x \in \mathbb{R}$ b) $y' = \dfrac{x^5 - 1}{x - 1}$, $x \in (1, \infty)$, $y(1) = 1$

   c) $y' = \sqrt{x} - \sqrt[3]{x}$, $x \in (0, \infty)$ d) $y' = \dfrac{1}{\sqrt{x}}$, $x \in (0, \infty)$, $y(1) = -1$

   e) $y' = \sin x + \cos x$, $x \in \mathbb{R}$ f) $y' = \dfrac{\cot x}{\sin x}$, $x \in (0, \pi)$, $y(\frac{\pi}{2}) = 1$

   g) $y' = \displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k}$, $x \in (-1,1)$ h) $y' = \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$, $x \in \mathbb{R}$, $y(0) = 1$

   i) $y' = \displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^k$, $x \in \mathbb{R}$ j) $y' = \dfrac{1}{1+x}$, $x \in (-1,1)$, $y(0) = 2$

   k) $y' = \dfrac{1}{1-x^2}$, $x \in (-1,1)$ l) $y' = \dfrac{2+3x}{1-x^2}$, $x \in (-1,1)$, $y(0) = 0$

7. Näytä, että jos $y(x)$ on $\mathbb{R}$:ssä n kertaa derivoituva ($n \in \mathbb{N}$) ja $y^{(n)} = 0$, niin y on polynomi astetta $\leq n$. Mikä on $y(x)$:n lauseke alkuehdoilla $y(x_0) = c_0$, $y'(x_0) = c_1$, $\ldots$, $y^{(n-1)}(x_0) = c_{n-1}$?

8. Näytä Korollaarin V.6.8 avulla, että välillä $(-1,1)$ pätee:

$$\operatorname{Arctan} x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots$$

9. Todista Lause V.6.9 soveltamalla Rollen lausetta funktioon

$$h(x) = [f(b) - f(a)][g(x) - g(a)] - [g(b) - g(a)][f(x) - f(a)].$$

   Miksei voi olla $g(a) = g(b)$?

10. Yritä laskea l'Hospitalin säännöllä raja-arvo $\lim_{x\to\infty} x/\sqrt{x^2+1}$.

11. Laske l'Hospitalin säänöillä:

a) $\displaystyle\lim_{x\to-3}\frac{x^2+3x}{x^2-9}$ b) $\displaystyle\lim_{x\to3^-}\frac{|x^2-4x+3|}{2x^2-5x-3}$ c) $\displaystyle\lim_{t\to8}\frac{t^{2/3}-4}{t^{1/3}-2}$

d) $\displaystyle\lim_{y\to1}\frac{y-4\sqrt{y}+3}{y^2-1}$ e) $\displaystyle\lim_{x\to\infty}\frac{2x^2+x+3}{1-x^2}$ f) $\displaystyle\lim_{x\to0}\frac{\sin 2x}{\sin 3x}$

g) $\displaystyle\lim_{x\to0}\frac{1-\cos 4x}{1-\cos 3x}$ h) $\displaystyle\lim_{x\to0}\frac{x-\sin x}{x^3}$ i) $\displaystyle\lim_{x\to0}\frac{2-x^2-2\cos x}{x^4}$

j) $\displaystyle\lim_{x\to0}\frac{x-\sin x}{x-\tan x}$ k) $\displaystyle\lim_{x\to0^+}\frac{\sin^2 x}{|x-\tan x|}$ l) $\displaystyle\lim_{t\to0}\frac{3\sin t-\sin 3t}{3\tan t-\tan 3t}$

m) $\displaystyle\lim_{x\to1^-}\frac{\operatorname{Arccos} x}{\sqrt{1-x^2}}$ n) $\displaystyle\lim_{x\to\infty} x(2\operatorname{Arctan} x-\pi)$

12. (*) Olkoon f ja g Lipschitz-jatkuvia välillä $[a,b]$. Todista: a) fg on Lipschitz välillä $[a,b]$. b) Jos lisäksi $g(x)\neq 0\ \forall x\in[a,b]$, niin f/g on Lipschitz välillä $[a,b]$.

13. (*) a) Funktio $y(x)$ on alkuarvotehtävän $y'=\frac{1}{1+x^7}$, $x\in(-1,1)$, $y(0)=1$ ratkaisu. Laske $y(\frac{1}{2})$ kahdeksan merkitsevän numeron tarkkuudella.
b) Määritä differentiaaliyhtälön $y'=\frac{1}{\sqrt{x}\,(1+x)}$ yleinen ratkaisu sarjamuotoisena välillä $(0,1)$.

14. (*) Halutaan todistaa, että laskusääntö

$$\lim_{x\to a^+}\frac{f'(x)}{g'(x)}=A\in\mathbb{R}\quad\Rightarrow\quad\lim_{x\to a^+}\frac{f(x)}{g(x)}=A$$

(Lause V.6.10: 1. sääntö, kun $A\in\mathbb{R}$) on pätevä myös, kun funktioiden f ja g raja-arvoja koskeva oletus muutetaan muotoon: $\lim_{x\to a^+}|f(x)|=\infty$ ja $\lim_{x\to a^+}|g(x)|=\infty$. Tarkista ja täydennä todistukseksi päättely:
Jos $a<x<t<a+\delta$, niin jollakin $\xi\in(x,t)$ pätee

$$\frac{f(x)-f(t)}{g(x)-g(t)}=\frac{f'(\xi)}{g'(\xi)}$$

$$\Rightarrow\quad\frac{f(x)}{g(x)}-A=\frac{f'(\xi)}{g'(\xi)}-A+\frac{1}{g(x)}\left(f(t)-g(t)\frac{f'(\xi)}{g'(\xi)}\right)$$

$$\Rightarrow\quad\left|\frac{f(x)}{g(x)}-A\right|\le\left|\frac{f'(\xi)}{g'(\xi)}-A\right|+\frac{1}{|g(x)|}\left(|f(t)|+|g(t)|\left|\frac{f'(\xi)}{g'(\xi)}\right|\right).$$

Jos nyt $\varepsilon>0$, niin viimeksi kirjoitetussa epäyhtälössä voidaan valita ensin t ja sitten x niin, että epäyhtälön oikea puoli on pienempi kuin ε.

V.7 Kiintopisteiteraatio. Newtonin menetelmä

Aiemmin Luvuissa I.4–I.8 on tarkasteltu esimerkkejä palautuvista lukujonoista, jotka määräytyvät alkuarvosta $x_0 \in \mathbb{R}$ ja (johonkin reaalifunktioon f liittyen) palautuskaavasta

$$x_{n+1} = f(x_n), \quad n = 0, 1, \ldots \tag{$\star$}$$

Jatkossa tarkastellaan tällaisia lukujonoja ja niiden suppenemisen ehtoja eiempaa yleisemmältä kannalta.

Oletetaan aluksi, että f on jatkuva suljetulla välillä $[a, b]$ ja että $x_n \in [a, b]$ $\forall n$. Tällöin jos jono $\{x_n\}$ on suppeneva, niin on oltava

$$x_n \to c \in [a, b].$$

Koska f on välillä $[a, b]$ jatkuva, niin

$$x_n \to c \;\Rightarrow\; f(x_n) \to f(c),$$

joten kaavan ($\star$) mukaan

$$c = f(c).$$

Sanotaan, että c on f:n *kiintopiste* (engl. fixed point) ja lukujen x_n laskemista palautuskaavasta ($\star$) sanotaan tämän vuoksi *kiintopisteiteraatioksi* (lat. itero = toistaa, tehdä uudelleen).

Jos halutaan löytää annetun funktion f kiintopiste, niin luonnollinen algoritmi (joskaan ei aina toimiva, ks. tarkastelut jäljempänä) on kiintopisteiteraatio ($\star$) jostakin alkuarvauksesta x_0.

Esimerkki 1 Ratkaise (transkendenttinen) yhtälö $x = \cos x$ kiintopisteiteraatiolla.

<u>Ratkaisu</u> Valitaan alkuarvaukseksi $x_0 = 0$, jolloin saadaan iteraatio

$$x_0 = 0, \quad x_{n+1} = \cos x_n, \quad n = 0, 1, \ldots$$

Tämä suppenee hitaahkosti kohti funktion $f(x) = \cos x$ (ainoaa) kiintopistettä $c = 0.7390851332..$

$$\begin{array}{ll} x_1 = 1 & x_{10} = 0.73140404.. \\ x_2 = 0.5403.. & x_{20} = 0.73893775.. \\ x_3 = 0.8575.. & x_{30} = 0.73908229.. \\ \vdots & \vdots \qquad \square \end{array}$$

Esimerkki herättää kysymyksen, millaisilla ehdoilla kiintopisteiteraatio ($\star$) yleensä suppenee, ja kuinka nopeasti, jos $x_0 \neq c$. (Tapaus $x_0 = c$ on triviaali, koska tällöin $x_n = c\ \forall n$.) Suppenemistarkasteluille suunnan antakoon

Esimerkki 2 Olkoon f ensimmäisen asteen polynomi. Tällöin jos $f(c) = c$, eli c on kiintopiste, niin jollakin $k \in \mathbb{R}$ on $f(x) = c + k(x - c)$, jolloin iteraatiokaavan ($\star$) mukaan on

$$x_{n+1} - c = k(x_n - c), \quad n = 0, 1, \ldots \quad \Rightarrow \quad x_n - c = k^n(x_0 - c), \quad n = 1, 2, \ldots$$

Siis $x_n \to c$ alkuarvosta x_0 riippumatta, jos $|k| < 1$. Jos $|k| \geq 1$, niin $x_n \to c$ vain kun $x_0 = c$. $\square$

Esimerkin mukaan f:n säännöllisyys (edes sileys) ei ole sovelias kriteeri kiintopisteiteraation suppenemiselle, vaan tarvitaan toisen tyyppisiä ehtoja. Seuraavissa kahdessa lauseessa asetetaan, esimerkin tulosta mukaillen, riittävät ehdot sekä kiintopisteiteraation suppenemiselle kohti haluttua kiintopistettä c että iteraation epäonnistumiselle ($x_n \not\to c$).

Lause V.7.1 Jos (i) funktiolla f on kiintopiste c, ja (ii) f on määritelty välillä $[c - a, c + a]$, $a > 0$ ja toteuttaa ehdon

$$|f(x) - f(c)| \leq L|x - c\,|, \quad x \in [c - a, c + a],$$

missä $0 \leq L < 1$[†], niin

(1) c on f:n ainoa kiintopiste välillä $[c - a, c + a]$,

(2) iteraatio ($\star$) suppenee kohti kiintopistettä jokaisella $x_0 \in [c - a, c + a]$ ja pätee
$$|x_n - c| \leq L^n|x_0 - c\,|, \quad n = 1, 2, \ldots$$

<u>Todistus</u> (1) Jos $c_1 \in [c - a, c + a]$ on myös kiintopiste, niin

$$\begin{aligned} c = f(c) \ \ \& \ \ c_1 = f(c_1) \quad &\Rightarrow \quad |c - c_1| = |f(c) - f(c_1)| \ \leq \ L|c - c_1| \\ &\Rightarrow \quad (1 - L)\,|c - c_1| \leq 0. \end{aligned}$$

Koska on $0 \leq L < 1$, niin on oltava $c - c_1 = 0$.

(2) Oletuksien perusteella pätee ensinnäkin

$$|f(x) - c\,| = |f(x) - f(c)| \ \leq \ L|x - c\,| \leq |x - c| \leq a, \quad \text{kun } x \in [c - a, c + a].$$

[†]Ilman lisäehtoa $L < 1$ Lauseen V.7.1 ehtoa (ii) sanotaan *Lipschitz-ehdoksi pisteessä* c, vrt. Lipschitz-jatkuvuuden ehto suljetulla välillä: Määritelmä V.6.4.

Näin ollen jos $x_0 \in [c-a, c+a]$, niin iteraatiolle $(\star)$ pätee $x_n \in [c-a, c+a]$ $\forall n$, joten oletuksien perusteella voidaan arvioida

$$|x_n - c\,| = |f(x_{n-1}) - f(c)| \;\le\; L|x_{n-1} - c\,| \;\le\; \ldots \;\le\; L^n|x_0 - c\,|.$$

Koska $L < 1$, niin $x_n \to c$. $\square$

Lause V.7.2 Jos (i) funktiolla f on kiintopiste c, ja (ii) f on määritelty välillä $[c-a, c+a]$, $a > 0$ ja toteuttaa ehdon

$$|f(x) - f(c)| \ge L|x - c\,|, \quad x \in [c-a, c+a],$$

missä $L > 1$, niin kiintopisteiteraatio $(\star)$ suppenee kohti kiintopistettä c ainoastaan siinä tapauksessa, että $x_n = c$ jollakin n.

Todistus Jos $x_k = c$, niin iteraatiokaavan $(\star)$ mukaan on $x_n = c$ $\forall n \ge k$, jolloin $x_n \to c$. Oletetaan siis, että $x_n \ne c$ $\forall n$, jolloin väittämä on, että $x_n \not\to c$. Jos $|x_n - c\,| > a$ $\forall n$, niin tämä on tosi. Voidaan siis olettaa, että $x_n \in [c-a, c+a]$ jollakin n. Tällöin oletuksien mukaan

$$|x_{n+1} - c\,| = |f(x_n) - f(c)| \ge L|x_n - c\,|.$$

Jos $x_{n+1} \in [c-a, c+a]$, voidaan edelleen arvioida $|x_{n+2} - c\,| \ge L^2|x_n - c\,|$, jne. Koska oletettiin, että $x_n \ne c$ ja $L > 1$, niin päätellään, että jollakin $m \in \mathbb{N}$, $m > n$ pätee

$$|x_m - c\,| \ge L^{m-n}|x_n - c\,| > a.$$

Edellä on päätelty, että jos $x_n \ne c$ $\forall n$, niin mistä tahansa indeksistä N eteenpäin on aina löydettävissä jokin indeksi $n > N$ siten, että $|x_n - c\,| > a$. Tällöin $x_n \not\to c$. $\square$

Jos f on derivoituva kiintopisteen ympäristössä tai ainakin kiintopisteessä (niinkuin usein), niin Lauseissa V.7.1 ja V.7.2 asetettujen ehtojen pätevyyttä voidaan tutkia helposti derivaatan avulla. Ensinnäkin jos oletetaan, että f on jatkuvasti derivoituva välillä $[c-a, c+a]$, niin Differentiaalilaskun väliarvolauseen mukaan on $f(x) - f(c) = f'(\xi)(x-c)$ jollakin $\xi \in (c-a, c+a)$, kun $x \in [c-a, c+a]$. Näin ollen voidaan päätellä:

$$\max_{x\in[c-a,c+a]} |f'(x)| = L < 1 \quad \Rightarrow \quad \text{Lauseen V.7.1 ehdot voimassa.}$$

$$\min_{x\in[c-a,c+a]} |f'(x)| = L > 1 \quad \Rightarrow \quad \text{Lauseen V.7.2 ehdot voimassa.}$$

Esimerkki 1 (jatko) Esimerkissä on $|f'(x)| = |\sin x| \le L < 1$ esim. välillä $[c-0.5, c+0.5]$. Esimerkin iteraatiolle ovat näin ollen voimassa Lauseen V.7.1 ehdot (kun $a = 0.5$) indeksistä $n = 1$ alkaen. $\square$

Jos Lauseiden V.7.1 ja V.7.2 ehdot asetetaan muodossa 'jollakin $a > 0$', ts. väliä $[c-a, c+a]$ ei kiinnitetä etukäteen, niin ehtojen toteutumiselle saadaan seuraava yksinkertainen kriteeri:

PROPOSITIO V.7.3 Jos f on derivoituva kiintopisteessä c, niin pätee:

$$|f'(c)| < 1 \quad \Leftrightarrow \quad \text{Lauseen V.7.1 ehdot voimassa jollakin } a > 0.$$
$$|f'(c)| > 1 \quad \Leftrightarrow \quad \text{Lauseen V.7.2 ehdot voimassa jollakin } a > 0.$$

<u>Todistus</u> Jos $|f'(c)| < 1$, niin derivaatan määritelmän ja raja-arvon (ε, δ)-kriteerin (Lause V.2.2) mukaan jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että pätee

$$\left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \varepsilon, \quad \text{kun } |x - c| < \delta \ \ \& \ \ x \neq c.$$

Kun tässä valitaan $\varepsilon = (1 - |f'(c)|)/2 > 0$ ja käytetään kolmioepäyhtälöä, niin nähdään, että Lauseen V.7.1 oletukset ovat voimassa jokaisella $a \in (0, \delta)$ (esim. $a = \delta/2$), kun valitaan $L = (1 + |f'(c)|)/2 < 1$. Tämä todistaa ensimmäisen väittämän osan $\boxed{\Rightarrow}$. Osa $\boxed{\Leftarrow}$ seuraa, kun Lauseen V.7.1 oletuksen (ii) perusteella päätellään, että on oltava $|f'(c)| \leq L < 1$ (Määritelmät V.3.1 ja V.2.1 sekä Lause I.7.4 [V1]). Toinen väittämä todistetaan vastaavasti. $\square$

Proposition V.7.3 ja Lauseen V.7.1 mukaisesti kiintopisteiteraatio $(\star)$ suppenee kohti kiintopistettä c, jos $|f'(c)| < 1$ ja lisäksi x_0 on <u>riittävän</u> <u>lähellä</u> kiintopistettä. Jos taas $|f'(c)| > 1$, niin $x_n \to c$ on Lauseen V.7.2 mukaisesti tosi vain siinä (melko onnekkaassa) tapauksessa, että $x_n = c$ jollakin n. Näiden tulosten perusteella kiintopistettä sanotaan *attraktiiviseksi* (eli puoleensa vetäväksi), jos $|f'(c)| < 1$ ja *hylkiväksi* eli *repulsiiviseksi*, jos $|f'(c)| > 1$. Luokittelun ulkopuolelle (tapauskohtaisesti tutkittaviksi) jäävät siis ainoastaan sellaiset kiintopisteet, joissa $f'(c) = \pm 1$.

ESIMERKKI 3 Funktion $f(x) = \sqrt{x+1}$ ainoa kiintopiste c on

$$c \geq 0 \ \ \& \ \ c = \sqrt{c+1} \ \Leftrightarrow \ c = \frac{1}{2}(\sqrt{5} + 1).$$

Koska

$$f'(x) = \frac{1}{2\sqrt{1+x}} \ \Rightarrow \ 0 < f'(c) < \frac{1}{2},$$

niin kyseessä on attraktiivinen kiintopiste. Tarkempi tutkimus paljastaa, että kiintopisteiteraatio $x_{n+1} = f(x_n)$ suppenee jokaisella $x_0 \in D_f = [-1, \infty)$. $\square$

Esimerkki 4 Funktiolla $f(x) = 1 - x^2$ on kaksi kiintopistettä:

$$c = 1 - c^2 \;\Leftrightarrow\; c = \frac{1}{2}(-1 \pm \sqrt{5}).$$

Koska $f'(c) = -2c = 1 \pm \sqrt{5}$, niin nähdään, että molemmat kiintopisteet ovat hylkiviä. Näin ollen päätellään (Lause V.7.2), että kiintopisteiteraatio

$$x_0 \in \mathbb{R}, \quad x_{n+1} = 1 - x_n^2, \quad n = 0, 1, \ldots$$

voi olla suppeneva vain jos $x_n = \frac{1}{2}(1 \pm \sqrt{5})$ jollakin n. Tämä mahdollisuus on pois suljettu esim. jos $x_0 \in \mathbb{Q}$, koska tällöin $\{x_n\}$ on rationaalilukujono. $\square$

Asymptoottinen suppenemisnopeus

Tarkastellaan kiintopisteiteraatiota ($\star$) olettaen, että (i) f on derivoituva kiintopisteessä c ja $|f'(c)| < 1$, (ii) $x_n \to c$, ja (3) $x_n \neq c \;\forall n$. Tällöin iteraatiokaavasta ($\star$) ja derivaatan määritelmästä seuraa

$$\lim_{n\to\infty} \frac{x_{n+1} - c}{x_n - c} = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} = f'(c).$$

Tällä perusteella voidaan sanoa, että luku $q = f'(c)$ määrää kiintopisteiteraation *asymptoottisen suppenemisnopeuden*: Suurilla n:n arvoilla on likimain

$$x_n - c \;\sim\; \text{vakio} \times q^n \quad (n \text{ suuri}).$$

Tästä nähdään myös, että jos $q > 0$, niin jono $\{x_n\}$ on asymptoottisesti (eli suurilla n) monotoninen. Jos $q < 0$, niin jono on asymptoottisesti 'hyppelehtivä', vrt. kuvio.

$0 < q < 1$ $\qquad$ $-1 < q < 0$

Esimerkki 5 Esimerkissä 1 oli $f'(c) = -\sin c \approx -0.67$. Suppenemisen verkkaisuus ja 'hyppelehtivyys' sai näin selityksensä. $\square$

Kiintopisteiteraatio $x_{n+1} = f(x_n)$ suppenee asymptoottisesti erityisen nopeasti silloin, kun kiintopisteessä on $f'(c) = 0$. Jotta myös tämä tapaus tulisi tarkemmin tutkituksi, oletettakoon yleisemmin, että jollakin $m \in \mathbb{N}$ ja $A \in \mathbb{R}$, $A \neq 0$ on voimassa

$$\lim_{x \to c} \frac{f(x) - f(c)}{(x-c)^m} = A.$$

Jos oletetaan samoin kuin edellä, että kiintopisteiteraatiolle pätee $x_n \to c$ ja $x_n \neq c \ \forall n$, niin oletuksen perusteella pätee

$$\lim_{n \to \infty} \frac{x_{n+1} - c}{(x_n - c)^m} = \lim_{n \to \infty} \frac{f(x_n) - f(c)}{(x_n - c)^m} = A,$$

jolloin suurilla n:n arvoilla on likimain

$$x_{n+1} - c \approx A(x_n - c)^m \quad (n \text{ suuri}).$$

Tämän perusteella sanotaan, että m on suppenemisnopeuden (asymptoottinen) *kertaluku.* — Huomattakoon, että jos $f'(c) \neq 0$, niin ym. oletus on (derivaatan määritelmän nojalla, vrt. Luku V.3) voimassa kun $m = 1$ ja $A = f'(c)$. Suppenemista tässä tapauksessa (kun $|A| < 1$) sanotaankin *lineaariseksi* (kertaluku = 1), ja muissa tapauksissa *superlineaariseksi.* Superlineaarisista tavallisin on tapaus $m = 2$, jolloin sanotaan, että kiintopisteiteraatio suppenee *kvadraattisesti.*

Esimerkki 6 Luvun I.8 Esimerkissä 4 tarkasteltiin tapauksessa $a = 2$ palautuvaa lukujonoa

$$x_0 = a, \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \quad n = 0, 1, \ldots$$

Olkoon nyt yleisemmin $a > 0, x_0 > 0$ ja tulkitaan lukujono kiintopisteiteraatioksi funktiolle $f(x) = \frac{1}{2}(x + \frac{a}{x})$. Kiintopisteitä on kaksi:

$$c = \frac{1}{2}\left(c + \frac{a}{c}\right) \Leftrightarrow c^2 = a \Leftrightarrow c = \pm\sqrt{a}.$$

Näistä vain $c = \sqrt{a}$ on mahdollinen jonon $\{x_n\}$ raja-arvo, kun $x_0 > 0$, koska tällöin on $x_n > 0 \ \forall n$. Tällä c:n arvolla nähdään, että

$$\begin{aligned} f(x) - f(c) = \frac{1}{2}\left(x + \frac{a}{x}\right) - \sqrt{a} &= \frac{1}{2x}\left(x^2 - 2\sqrt{a}\,x + a\right) \\ &= \frac{1}{2x}(x - c)^2. \end{aligned}$$

Tämän perusteella $(x-c)^{-2}[f(x) - f(c)] \to 1/(2c)$, kun $x \to c = \sqrt{a}$. Siis jos lukujono $\{x_n\}$ suppenee, niin se suppenee kvadraattisesti. — Tarkempi tutkimus osoittaa, että $x_n \to \sqrt{a}$ aina kun $x_0 > 0$. $\square$

Newtonin menetelmä

Jos funktio on derivoituva pisteessä c, niin sitä voidaan approksimoida pisteen c ympäristössä perustuen linearisaatioon (vrt. Luku V.3)

$$f(x) \approx f(c) + f'(c)(x - c).$$

Tähän linearisointiajatukseen perustuu epälineaaristen yhtälöiden ratkaisussa hyvin yleisesti käytetty ja tehokas menetelmä, *Newtonin menetelmä*. Newtonin menetelmässä etsitään yhtälölle

$$f(x) = 0$$

ratkaisua pisteen x_0 (alkuarvaus) lähistöltä. Algoritmissa f linearisoidaan pisteessä x_n (aluksi $n = 0$) ja ratkaistaan linearisoitu yhtälö

$$f(x_n) + f'(x_n)(x - x_n) = 0.$$

Tämä ratkeaa, jos $f'(x_n) \neq 0$. Kun ratkaisua merkitään $x = x_{n+1}$, saadaan algoritmiksi

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \ldots} \qquad (\star\star)$$

Laskimien ja tietokoneohjelmien komentojen 'Solve' tai 'FindRoot' takana on yleensä joko tämä menetelmä tai jokin sen variaatio, kuten *sekanttimenetelmä*, ks. kommentit edempänä.

Esimerkki 7 Jos $f(x) = x^2 - a$, $a > 0$, niin Newtonin algoritmi f:n nollakohdan $c = \sqrt{a}$ määrämiseksi on

$$x_0 > 0, \quad x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \quad n = 0, 1, 2, \ldots$$

Esimerkissä 6 oli siis kyse Newtonin menetelmästä. □

Iteraatiokaavan ($\star\star$) mukaisesti Newtonin algoritmi on kiintepistoiteraatio sovellettuna funktioon

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Algoritmin suosio perustuu siihen, että sikäli kuin iteraatio suppenee, suppeneminen on melko yleisin edellytyksin kvadraattista. Edellytys kvadraattiselle suppenemiselle on, että f on nollakohdan c lähellä riittävän säännöllinen ja että nollakohta on *yksinkertainen*, ts. $f'(c) \neq 0$. Seuraavan täsmällisen suppenemislauseen todistus perusuu differentiaalilaskennan väittämään, jota ei vielä ole käytettävissä. Sen vuoksi todistuksessa rajoitutaan toistaiseksi erikoistapaukseen, jossa f on polynomi. (Yleisempi todistus, ks. Harj.teht. VII.4:16.)

LAUSE V.7.4 Jos f on kahdesti jatkuvasti derivoituva välillä $[c-a, c+a]$ jollakin $a > 0$ ja $f(c) = 0$ ja $f'(c) \neq 0$, niin Newtonin iteraatio ($\star\star$) suppenee c:tä kohti, kun x_0 on c:tä riittävän lähellä, ja pätee

$$\lim_{n\to\infty} \frac{x_{n+1} - c}{(x_n - c)^2} = \frac{f''(c)}{2f'(c)}\,.$$

Todistus, kun f on polynomi: Koska $f(c) = 0$, niin $f(x) = (x - c)g(x)$, missä g on polynomi, ja samalla perusteella $g(x) - g(c) = (x - c)h(x)$ ja $f'(x) - f'(c) = (x - c)r(x)$, missä h ja r ovat polynomeja (Lause III.3.1). Derivoimalla nähdään, että $f'(c) = g(c)$, joten saadaan hajotelmat

$$\begin{aligned} f(x) &= (x - c)[g(c) + (x - c)h(x)] = f'(c)(x - c) + (x - c)^2 h(x), \\ f'(x) &= f'(c) + (x - c)r(x). \end{aligned}$$

Sijoittamalla nämä F:n lausekkeeseen ja huomioimalla, että $F(c) = c$, seuraa

$$\begin{aligned} F(x) - F(c) &= x - c - \frac{f(x)}{f'(x)} \\ &= x - c - \frac{f'(c)(x - c) + (x - c)^2 h(x)}{f'(c) + (x - c)r(x)} \\ &= \frac{(x - c)^2[r(x) - h(x)]}{f'(c) + (x - c)r(x)}\,. \end{aligned}$$

Tämän perusteella

$$\lim_{x\to c} \frac{F(x) - F(c)}{(x - c)^2} = \frac{r(c) - h(c)}{f'(c)}\,.$$

Derivoimalla em. f:n ja f':n hajotelmia nähdään edelleen, että $h(c) = \frac{1}{2}f''(c)$ ja $r(c) = f''(c)$, joten väite seuraa. $\square$

Lauseen V.7.4 perusteella Newtonin iteraation konvergenssi on lauseen oletuksin kvadraattista, tai jopa 'superkvadraattista' (jos $f''(c) = 0$). Jos f täyttää Lauseen V.7.4 säännöllisyysehdon mutta nollakohta c on *kaksinkertainen*, ts.

$$f(c) = f'(c) = 0,\ f''(c) \neq 0,$$

niin Newtonin algoritmi suppenee tässäkin tapauksessa (riittävän läheltä c:tä), mutta suppeneminen hidastuu lineaariseksi. Tarkemmin pätee tässä tapauksessa: $\lim_n (x_{n+1} - c)/(x_n - c) = 1/2$ (Harj.teht. 16; vrt. myös Esimerkki 6, kun $a = 0$).

ESIMERKKI 8 Ratkaise Esimerkin 1 yhtälö $x = \cos x$ Newtonin menetelmällä.

Ratkaisu Kun valitaan $f(x) = x - \cos x$, niin Lauseen V.7.4 ehdot ovat voimassa ja $f''(c) = \cos c \neq 0$, joten Newtonin iteraatio suppenee (sikäli kuin suppenee) kvadraattisesti. Iteraatiokaava on

$$x_{n+1} = x_n - \frac{x_n - \cos x_n}{1 + \sin x_n} = \frac{x_n \sin x_n + \cos x_n}{1 + \sin x_n}, \quad n = 0, 1, 2, \ldots$$

Alkuarvauksella $x_0 = 0$ on tulos (vrt. Esimerkki 1)

$$\begin{aligned} x_0 &= 0 \\ x_1 &= 1 \\ x_2 &= 0.7503638678.. \\ x_3 &= 0.7391128909.. \\ x_4 &= 0.7390851333.. \\ x_5 &= 0.7390851332.. \\ &\vdots \quad \square \end{aligned}$$

Jos derivaatta f' on nopeasti muuttuva f:n nollakohdan lähellä, voi Newtonin menetelmä olla hyvin herkkä alkuarvaukselle, eikä iteratio välttämättä suppene lainkaan. Laskentaohjelma antaa silloin tuloksen 'failed to converge'. Tällöin yleensä yksinkertaisesti vaihdellaan alkuarvoa x_0, kunnes onni kääntyy. Toinen mahdollisuus on käyttää jotakin varmempaa menetelmää hyvän alkuarvauksen hakuun, jolloin Newtonin iteraation tehtäväksi jää 'loppukiihdytys'. Esimerkiksi Bolzanon lauseen (Lause V.1.11) todistuksessa käytetty puolituskonstruktio on aloitusmenetelmänä oivallinen — hidas mutta varma.

Esimerkki 9 Jos funktion $f(x) = x/(1 + x^2)$ nollakohtaa haetaan Newtonin menetelmällä, tulee iteraatiokaavaksi

$$x_{n+1} = F(x_n) = -\frac{2x_n^3}{1 - x_n^2}, \quad n = 0, 1, 2, \ldots$$

Suppenemisalueen rajalle joudutaan, jos valitaan $x_0 = a \neq 0$ siten, että $x_1 = -a$, jolloin iteraatiokaavan mukaan on $x_n = (-1)^n a$. Näin käy siis kun

$$a = \frac{2a^3}{1 - a^2} \quad \& \quad a \neq 0 \quad \Leftrightarrow \quad a = \pm\frac{1}{\sqrt{3}}\,.$$

Jos $|x_0| < 1/\sqrt{3}$, niin iteraatio suppenee: $x_n \to 0$. (Suppenemisnopeuden kertaluku on $m = 3$, sillä $\lim_{x\to 0} x^{-3}F(x) = -2$.) Jos $|x_0| \geq 1/\sqrt{3}$, on tulos 'failed to converge'. Geometrisestikin nähdään, että jos $|x_0| > 1/\sqrt{3}$, niin itse asiassa $|x_n| \to \infty$ (kuvassa $x_n \to -\infty$).

Sekanttimenetelmä

Jos funktion derivoituvuudessa on ongelmia, tai jos derivaattoja on hankala määrätä, voidaan käyttää Newtonin mentelmän lähisukulaista, *sekanttimenetelmää*. Tässä ideana on käyttää pisteiden $(x_{n-1}, f(x_{n-1}))$ ja $(x_n, f(x_n))$ kautta kulkevaa suoraa eli käyrän $y = f(x)$ *sekanttia* funktion approksimointiin määrättäessä seuraavaa pistettä x_{n+1}. Lauseen V.7.4 oletuksilla sekanttimenetelmän iteraatio suppenee lähes yhtä nopeasti kuin Newtonin.† Algoritmin käyntiin saattamiseksi on sekanttimenetelmässä annettava kaksi alkuarvausta x_0, x_1.

Sekanttimenetelmä

Newton

†Sekanttimenetelmän asymptoottinen suppenemisnopeus on lineaarisen ja kvadraattisen suppenemisen välimuoto; tarkemmin on osoitettavissa, että Lauseen V.7.4 oletuksin pätee

$$\lim_n \frac{|x_{n+1} - c|}{|x_n - c|^\alpha} = \left| \frac{f''(c)}{2f'(c)} \right|,$$

missä $\alpha = \frac{1}{2}(\sqrt{5} + 1) \approx 1.62$. (Potenssifunktio $f(x) = x^\alpha$, $\alpha \notin \mathbb{Q}$ määritellään jäljempänä Luvussa VI.1.)

HARJOITUSTEHTÄVIÄ

1. Seuraavat yhtälöt voidaan ratkaista kiintopisteiteraatiolla. Määritä asymptoottiset suppenemisnopeudet tarkasti (jos mahdollista) tai yhden desimaalin tarkkuudella:

$$\text{a)}\ x = \frac{12}{1+x} \qquad \text{b)}\ x = \sqrt{3+x} \qquad \text{c)}\ x = \frac{1}{2+x^2}$$
$$\text{d)}\ x = \sqrt[3]{x+9} \qquad \text{e)}\ x = \cos\frac{x}{3} \qquad \text{f)}\ x = 1 + \frac{1}{4}\sin x$$

2. Tutki, mitkä polynomin $p(x) = x^3 + 8x^2 - 44x - 10$ nollakohdista voidaan tarkentaa kiintopisteiteraatiolla

$$x_{n+1} = \frac{1}{44}(x_n^3 + 8x_n^2 - 10), \quad n = 0, 1, \ldots$$

olettaen, että käytettävissä on riittävän hyvä alkuarvaus x_0. Miten tähän iteraatiomenetelmään on päädytty?

3. Yhtälön $x^3 + x = 1$ reaalista ratkaisua voidaan yrittää hakea kiintopisteiteraatiolla hajottamalla yhtälö muotoon

$$\text{a)}\ x = 1 - x^3 \qquad \text{b)}\ x = x^{-2} - x^{-1} \qquad \text{c)}\ x = \sqrt[3]{1-x} \qquad \text{d)}\ x = \frac{1}{1+x^2}$$

Tutki, miten iteraatiot (asymptoottisesti) suppenevat tai hajaantuvat olettaen, että alkuarvaus on hyvin lähellä kiintopistettä.

4. a) Yhtälö $x = \cos x$ voidaan kirjoittaa muotoon $y = \operatorname{Arccos} y$ ja yrittää ratkaista kiintopisteiteraatiolla $y_{n+1} = \operatorname{Arccos} y_n$. Toimiiko menetelmä?
b) Jos yhtälö ratkaistaan iteraatiolla $x_0 = 0$, $x_{n+1} = \cos x_n$, niin mitä lukua kohti ja kuinka nopeasti iteraatio suppenee, jos funktioevaluaatioissa $x_n \mapsto \cos x_n$ muuttujan yksikkö on aste, ts. $\cos x = \cos x^\circ$? Kokeile valisemalla laskimeen astemoodi ja painelemalla $\boxed{\cos}$ – näppäintä!

5. Näytä, että jos x_0 on rationaaliluku, niin kiintopisteiteraatio

$$x_{n+1} = (x_n - 2)^2, \quad n = 0, 1, \ldots$$

suppenee vain, jos x_0 on jokin luvuista $0, 1, 2, 3, 4$.

6. Johda Newtonin iteraatiokaava luvun $\sqrt[m]{a}$ määräämiseksi funktion $f(x) = x^m - a$ nollakohtana ($a > 0$, $m \in \mathbb{N}$, $m \geq 2$). Päättele suppeneminen kvadraattiseksi. Päättele myös geometrisesti, että iteraatio suppenee aina kun $x_0 > 0$.

7. Etsi seuraavien funktioiden nollakohdat annetulta väliltä neljän desimaalin tarkkuudella käyttäen Newtonin menetelmää:

$$\begin{aligned}
&\text{a)}\ \ f(x) = x^3 + 2x - 1, \quad c \in [0,1] \\
&\text{b)}\ \ f(x) = x^4 - 8x^2 - x + 16, \quad c \in [1,3] \\
&\text{c)}\ \ f(x) = \cos x - x^2, \quad c \in (-\infty, \infty) \\
&\text{d)}\ \ f(x) = 3\sin x - x - 1, \quad c \in [0, \infty)
\end{aligned}$$

8. Laske seuraavien funktioiden maksimi- ja minimiarvot tarkasti, jos mahdollista, muuten kuuden desimaalin tarkkuudella:

$$\text{a)}\ \frac{\sin x}{1+x^2} \qquad \text{b)}\ \frac{\cos x}{1+x^2} \qquad \text{c)}\ f(x) = \begin{cases} \dfrac{\sin x}{x}, & \text{kun } x \neq 0 \\ 1, & \text{kun } x = 0 \end{cases}$$

9. Millä a:n arvoilla yhtälöllä $\cos x = ax$ on täsmälleen kaksi ratkaisua?

10. Laske (likimäärin) a) funktion $f(x,y) = xy^2 + y^4$ maksimiarvo ympyräviivalla $S:\ x^2+y^2 = 1$, b) funktion $f(x,y) = (x-y)(x+y)^2$ pienin ja suurin arvo ympyräviivalla $x = 2\cos t$, $y = 1 + 2\sin t$, $t \in [0, 2\pi)$.

11. Millaisen algoritmisen muodon saa jakolaskuoperaatio $a \mapsto a^{-1}$, kun se suoritetaan soveltamalla Newtonin iteraatiota funktioon $f(x) = x^{-1} - a$? Tarvitaanko algoritmissa jakolaskuja? Kokeile, kun $a = 3$.

12. Funktioevaluaatio $y \mapsto \operatorname{Arctan} y$ halutaan toteuttaa Newtonin menetelmään perustuvalla algoritmilla, joka sisältää $\mathbb{R}$:n kuntaoperaatioiden lisäksi ainoastaan funktioevaluaatioita $x \mapsto \cos x$ ja $x \mapsto \sin x$. Esitä tällainen algoritmi.

13. Laske luvulle $\sqrt{2}$ approksimaatio iteroimalla neljä kertaa sekanttimenetelmällä alkuarvauksista $x_0 = 2$, $x_1 = 1.5$ (funktio $f(x) = x^2 - 2$). Vertaa Newtonin menetelmään.

14. (*) Sanotaan, että funktio f on *kontraktio* välillä $[a,b]$, jos f on välillä $[a,b]$ Lipschitz-jatkuva vakiolla $L < 1$. Todista *Kontraktiokuvauslause*: Jos f on kontraktio välillä $A = [a,b]$ ja lisäksi $f(A) \subset A$, niin pätee:

    (i) f:llä on täsmälleen yksi kiintopiste c välillä A.

    (ii) Kiintopisteiteraatio $x_{n+1} = f(x_n)$, $n = 0, 1, \ldots$ suppenee kohti c:tä jokaisella $x_0 \in A$.

    Vihje: Sovella ensin Bolzanon lausetta funktioon $g(x) = f(x) - x$.

15. (*) Funktiosta f tiedetään, että f on (tuntemattoman) nollakohdan c lähellä jatkuvasti derivoituva ja että $f'(c) \neq 0$. Etsitään nollakohtaa kiintopisteiteraatiolla

$$x_{n+1} = x_n - kf(x_n), \quad n = 0, 1, \ldots$$

Näytä, että jos x_0 on riittävän lähellä c:tä ja valitaan $k = 1/f'(x_0)$, niin iteraatio suppenee kohti c:tä ainakin lineaarisesti. Näytä edelleen, että rajalla $x_0 \to c$ suppeneminen muuttuu superlineaariseksi, ts. suurilla n pätee $x_n - c \sim q^n$, missä $q \to 0$ kun $x_0 \to c$.

16. (*) Näytä, että jos f on polynomi ja c on f:n m-kertainen nollakohta, $m \geq 2$, niin Newtonin iteraatio suppenee c:tä kohti aina kun alkuarvaus on riittävän lähellä c:tä ja suppeneminen on asymptoottisesti lineaarista, tarkemmin

$$q = \lim_n \frac{x_{n+1} - c}{x_n - c} = \frac{1}{m}\,.$$

17. (*) Luku $\sqrt{a}$ voidaan määrätä iteraatioilla

$$\text{a)}\quad x_{n+1} = \frac{x_n^3 + 3ax_n}{3x_n^2 + a}\,, \quad n = 0, 1 \ldots$$

$$\text{b)}\quad x_{n+1} = \frac{3x_n}{8} + \frac{3a}{4x_n} - \frac{a^2}{8x_n^3}\,, \quad n = 0, 1, \ldots$$

Näytä, että jos $x_n \to \sqrt{a}$, niin suppeneminen on kummassakin tapauksessa *kuutiollista* (kertaluku=3). Onko iteraatioilla muita kiintopisteitä ja minkälaatuisia ne ovat? Kokeile menetelmiä käytännössä, kun $a = 2$, $x_0 = 1.5$, ja vertaa Esimerkin 6 kvadraattiseen menetelmään.

18. (*) (Laskiainen, 1. lasku) Lumilautailija haluaa rakentaa mäen, jota pitkin voi laskea xy-tason origosta pisteeseen $(5, -1)$ nopeinta mahdollista reittiä (gravitaation suunta $-\vec{j}$, ei kitkaa). Ryhdy konsultiksi käyttäen vanhaa tietoa[†], jonka mukaan oikea mäen profiili on sykloidin kaari

$$\begin{cases} x = R(t - \sin t), \\ y = -R(1 - \cos t). \end{cases}$$

Laske siis R ja esittele graafisesti ehdotuksesi optimaaliseksi mäeksi.

[†]Lyhimmän ajan käyrän eli *brakistokronin* ongelman ratkaisi sveitsiläinen matemaatikko **Johann Bernoulli** (1667-1748) v. 1697. Ratkaisemisessa kilpaili myös Johann B:n veli **Jakob** (1654-1705), jonka mukaan mm. Bernoullin epäyhtälö (Propositio I.4.2) on nimetty.

V.8 Analyyttiset kompleksifunktiot

Kompleksifunktiolla tarkoitetaan kompleksimuuttujan kompleksiarvoista funktiota eli funktiota tyyppiä $f : \mathcal{D}_f \to \mathbb{C}$, $\mathcal{D}_f \subset \mathbb{C}$. Tällainen on esimerkiksi Luvussa III.3 tarkasteltu (koko $\mathbb{C}$:ssä määritelty) polynomi. Koska z on tulkittavissa tason (kompleksitason) pisteeksi, niin kompleksifunktio voidaan ymmärtää myös kahden reaalimuuttujan kompleksiarvoisena funktiona kirjoittamalla (vrt. Luvun III.3 Esimerkki 1)

$$f(x+iy) = u(x,y) + iv(x,y),$$

missä $u(x,y) = \operatorname{Re} f(x+iy)$ ja $v(x,y) = \operatorname{Im} f(x+iy)$. Huolimatta tästä erosta suhteessa yhden reaalimuuttujan funktioihin voidaan kompleksifunktioille määritellä käsitteet jatkuvuus, raja-arvo ja derivoituvuus aivan samalla tavoin kuin reaalifunktioille. Esimerkiksi f on jatkuva pisteessä $z \in \mathcal{D}_f$ täsmälleen kun kaikille kompleksilukujonoille $\{z_n\}$ pätee (vrt. Määritelmä V.1.1)

$$z_n \in \mathcal{D}_f \;\;\&\;\; z_n \to z \quad \Rightarrow \quad f(z_n) \to f(z).$$

Tässä lukujonojen $\{z_n\}$ ja $\{f(z_n)\}$ suppeneminen viittaa Määritelmään I.6.1, joka toimii sellaisenaan myös kompleksilukujen jonoille, kunhan merkinnän $|\cdot|$ tulkitaan tarkoittavan kompleksiluvun itseisarvoa. Määritelmän mukaan pätee erityisesti (kuten reaalilukujonoillekin)

$$z_n \to z \quad \Leftrightarrow \quad |z_n - z| \to 0.$$

Jos tässä on $z = 0$, niin pätee siis

$$z_n = r_n \angle \varphi_n \to 0 \quad \Leftrightarrow \quad r_n \to 0.$$

Tämän mukaan lähestyminen kohti kompleksitason pistettä (tässä origoa) voi tapahtua äärettömän monesta eri suunnasta ($\varphi_n = \varphi \in [0, 2\pi)\ \forall n$) tai suuntia vaihdellen (esim. spiraalimainen lähestyminen). Jatkossa nähdään, että tämä kvalitatiivinen ero suhteessa reaalilukujonoon (jolla mahdollisia lähestymissuuntia kohti raja-arvoa on vain kaksi) voi tuottaa kompleksifunktoiden raja-arvotarkasteluissa yllätyksiä.

Esimerkki 1 Funktio $f(z) = \overline{z}$ (eli $f(x+iy) = x - iy$) on jatkuva $\mathbb{C}$:ssä, sillä kompleksialgebran (ks. Luku III.2) mukaan

$$|f(z_n) - f(z)| = |\overline{z_n} - \overline{z}| = |\overline{z_n - z}| = |z_n - z| \to 0$$

aina kun $z_n \to z$. $\square$

Myös jatkuvuuden vaihtoehtoinen määritelmä (Määritelmä V.1.2) toimii sellaisenaan kompleksifunktioille, samoin funktion raja-arvon määritelmä. Raja-arvon käsitteeseen perustuva funktion derivaatta määritellään myös samoin kuin reaalifunktioille:

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}.$$

Jos raja-arvo on olemassa, niin sanotaan, että f on *derivoituva* pisteessä z.

Esimerkki 2 Tutki kompleksifunktioiden

$$\text{a) } f(z) = \bar{z}, \quad \text{b) } f(x + iy) = x^2 + iy^2, \quad \text{c) } f(x + iy) = x^2 - y^2 + 2ixy$$

derivoituvuutta.

Ratkaisu a) Jos $\Delta z_n = \Delta r_n \angle \varphi_n = \Delta r_n(\cos \varphi_n + i \sin \varphi_n) \neq 0$ $(\Delta r_n \neq 0)$, niin

$$\frac{f(z + \Delta z_n) - f(z)}{\Delta z_n} = \frac{\cos \varphi_n - i \sin \varphi_n}{\cos \varphi_n + i \sin \varphi_n}.$$

Nähdään, ettei vaadittua raja-arvoa ole, eli f ei ole missään pisteessä derivoituva.

b) Jos tässä merkitään $\Delta z_n = \Delta x_n + i\Delta y_n$, niin

$$\frac{f(z + \Delta z_n) - f(z)}{\Delta z_n} = 2\,\frac{x\Delta x_n + iy\Delta y_n}{\Delta x_n + i\Delta y_n} + \frac{(\Delta x_n)^2 + (\Delta y_n)^2}{\Delta x_n + i\Delta y_n}.$$

Siirtymällä polaariesitykseen nähdään, että jälkimmäinen termi oikealla $\to 0$ aina kun $|\Delta z_n| = r_n \to 0$. Ensimmäisellä termillä sen sijaan on vain lähestymissuunnasta riippuvia suunnattuja raja-arvoja, ellei ole $x = y = t$, jolloin ko. termi yksinkertaistuu muotoon $2t$. Päätellään siis, että f on derivoituva ainoastaan kompleksitason suoralla $z = t(1 + i)$, $t \in \mathbb{R}$, ja tällä suoralla $f'(z) = 2Rez$.

c) Tässä esimerkissä ollaan onnekkaampia, sillä $f(x + iy) = (x + iy)^2$, jolloin

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{(z + \Delta z)^2 - z^2}{\Delta z} = 2z + \Delta z, \quad \Delta z \neq 0.$$

Siis f on derivoituva jokaisessa pisteessä $z \in \mathbb{C}$ ja $f'(z) = 2z$. $\square$

Esimerkki kertoo, että kompleksifunktion derivoituvuus on kaikkea muuta kuin 'läpihuutojuttu' siinäkin tapauksessa, että funktion reaali- ja imaginaariosat ovat säännöllisiä funktioita (kuten esimerkissä polynomeja). Derivoituvuus onkin kompleksifunktiolle paljon vaativampi ominaisuus kuin reaalifunktiolle.

Seuraavassa määritelmässä laajennetaan derivoituvuusehto koskemaan kompleksitason *avointa* osajoukkoa $G \subset \mathbb{C}$.

MÄÄRITELMÄ V.8.1 (**Analyyttinen kompleksifunktio**) Kompleksitason osajoukko $G \subset \mathbb{C}$ on *avoin*, jos jokaisella $z_0 \in G$ on *ympäristö*

$$U_\delta(z_0) = \{z \in \mathbb{C} \mid |z - z_0| < \delta\}, \quad \delta > 0$$

siten, että $U_\delta \subset G$. Kompleksifunktio, joka on derivoituva ei-tyhjässä, avoimessa joukossa $G \subset \mathbb{C}$, on *analyyttinen* G:ssä.

Analyyttiset funktiot muodostavat kompleksifunktioiden tärkeän ja paljon tutkitun 'vähemmistön'.†

ESIMERKKI 2 (jatko) Esimerkissä c-kohdan funktio $f(z) = z^2$ on analyyttinen koko kompleksitasossa ($G = \mathbb{C}$), kun taas a- ja b- kohtien funktiot eivät ole analyyttisiä missään. □

Jos f on analyyttinen koko kompleksitasossa, niin sanotaan, että f on *kokonainen* (engl. entire) funktio. Kaikki polynomit $p(z)$ (myös kompleksikertoimiset) ovat kokonaisia funktioita, sillä polynomin derivoimissääntö perustuu vain kunta-algebraan, joka ei lainkaan muutu siirryttäessä reaalimuuttujasta kompleksimuuttujaan (vrt. Esimerkki 2, c-kohta). Algebraan perustuvat myös funktioiden summan, tulon, osamäärän ja yhdistetyn funktion derivoimissäännöt reaalifunktioille, joten niissäkin voidaan reaalimuuttujan x tilalle vaihtaa kompeksimuuttuja z säännön muuttumatta. Näiden sääntöjen perusteella voidaan erityisesti jokainen (kompleksikertoiminen) *kompleksimuuttujan rationaalifunktio* derivoida samalla tavoin kuin reaalimuuttujan tapauksessa. Rationaalifunktiot ovatkin Määritelmän V.8.1 mukaisesti analyyttisiä koko määrittelyjoukossaan, sillä jos $f(z) = p(z)/q(z)$ (p ja q polynomeja) ja $z_0 \in \mathcal{D}_f$, niin $q(z_0) \neq 0$, jolloin $q(z) \neq 0$ myös jossakin ympäristössä $U_\delta(z_0)$. Tällöin $U_\delta(z_0) \subset \mathcal{D}_f$, eli $\mathcal{D}_f$ on avoin joukko.

ESIMERKKI 3 Rationaalifunktiot

$$f(z) = \frac{1}{z^2+1}, \quad g(z) = \frac{i}{z^2+iz}$$

on määritelty koko kompleksitasossa lukuunottamatta pisteitä $\pm i$ (f) ja $0, -i$ (g). Määrittelyjoukot ovat avoimia, ja molemmat funktiot ovat määrittelyjoukossaan derivoituvia, siis analyyttisiä. Derivaatat lasketaan kuten reaalimuuttujan tapauksessa:

$$f'(z) = -\frac{2z}{(z^2+1)^2}, \quad g'(z) = -\frac{i(2z+i)}{(z^2+iz)^2}. \quad \square$$

†Analyyttisten kompleksifunktioiden teorian tarkempi esittely kuuluu toisiin yhteyksiin. Mainittakoon teorian tuloksista kuitenkin, että jos f on analyyttinen G:ssä, niin samoin on f', jolloin f on itse asiassa mielivaltaisen monta kertaa derivoituva G:ssä. — Tämä tulos kertoo osaltaan, että kompleksifunktion analyytisyys on paljon voimakkaampi ominaisuus kuin reaalifunktion derivoituvuus.

Jos kompleksifunktio on analyyttinen nollakohtansa ympäristössä, niin nollakohtaa voidaan etsiä Newtonin iteraatiolla

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}, \quad n = 0, 1, \ldots$$

Esimerkiksi polynomin yksinkertaista nollakohtaa etsittäessä algoritmi toimii erinomaisesti, kunhan alkuarvaus on riittävän hyvä (ks. Harj.teht. 3–5).

HARJOITUSTEHTÄVIÄ

1. Näytä suoraan kompleksifunktion derivaatan määritelmästä, että
   a) $\mathrm{D}z^{-1} = -z^{-2}$, b) $\mathrm{D}(z+i)^{-2} = -2(z+i)^{-3}$.

2. Missä kompleksitason osajoukossa seuraavat funktiot ovat analyyttisiä?

   a) $f(z) = z^3 + iz$ b) $f(z) = \dfrac{z+1}{z^2+z+1}$ c) $f(z) = \dfrac{1}{z^3-8}$
   d) $f(z) = \dfrac{z}{|z|^2}$ e) $f(z) = \dfrac{\overline{z}}{|z|^2}$ f) $f(z) = (z+\overline{z})^2 - 2|z|^2 - \overline{z}^2$

3. Näytä, että jos $a \in \mathbb{C}$, $a \neq 0$, niin Newtonin iteraatio

   $$z_0 \in \mathbb{C}, \quad z_{n+1} = \frac{1}{2}\left(z_n + \frac{a}{z_n}\right), \quad n = 0, 1, \ldots$$

   (vrt. Luku V.7, Esimerkki 6) suppenee kvadraattisesti kohti funktion $f(z) = z^2 - a$ nollakohtaa, jos alkuarvaus z_0 on nollakohtaa (kumpaa tahansa) riittävän lähellä. Kokeile iteraation toimivuuttaa tapauksessa $a = i$ valinnoilla a) $z_0 = 1$, b) $z_0 = i$, c) $z_0 = -1 - i$.

4. Etsi polynomin $f(z) = z^4 + z + 4$ nollakohdat likimäärin Newtonin menetelmällä. Huomaa, että Newtonin iteraatio ei tässä tapauksessa suppene reaalisilla alkuarvauksilla — miksei?

5. (*) Todista Lauseen V.7.4 vastine analyyttiselle kompleksifunktiolle $f(z)$ tapauksessa, jossa f on polynomi.

V.9 *Jatkuvuuden logiikka

Tässä luvussa todistetaan Luvussa V.1 esitetyt kaksi jatkuvien funktioiden päälausetta, Weierstrassin lause (Lause V.1.12) ja käänteisfunktion jatkuvuutta koskeva Lause V.1.14. Todistukset perustuvat Luvussa I.10 esitettyyn osajonojen teoriaan, ja ne ovat melko vaativia. Erityisesti rajoitettuja reaalilukujonoja koskeva Bolzanon–Weierstrassin lause (Lause I.10.12) on todistuksissa ahkerassa käytössä. Weierstrassin lauseen todistuksen jatkoksi näytetään, että sama todistustekniikka, yhdistettynä eräisiin polynomeja koskeviin teknisempiin väittämiin, johtaa jopa Algebran peruslauseen todistukseen. Luvun lopussa määritellään vielä käsite *tasainen jatkuvuus*, jota voi pitää Lipschitz-jatkuvuuden minimalistisena vastineena. Tasaisen jatkuvuuteen liittyen todistetaan eräs reaalianalyysin hämmästyttävimmistä lauseista.

Koska tarkoituksena on todistaa Lauseet V.1.12 ja V.1.14 hieman yleisemmässä muodossa, määritellään aluksi suljettua väliä yleisempi *kompaktin* joukon käsite. Asiayhteyden vuoksi esitellään samalla muitakin $\mathbb{R}$:n nk. *topologisia* peruskäsitteitä.

Avoimet, suljetut ja kompaktit joukot

Avointa väliä

$$U_\delta(x) = (x - \delta, x + \delta) \quad (\delta > 0)$$

sanotaan pisteen $x \in \mathbb{R}$ (avoimeksi) *ympäristöksi* (tai tarkemmin δ-ympäristöksi, engl. δ-neighbourhood; vrt. vastaava kompleksitason käsite Määritelmässä V.8.1).

MÄÄRITELMÄ V.9.1 Joukko $A \subset \mathbb{R}$ on

- *avoin*, jos $\forall x \in A$ pätee $(x - \delta, x + \delta) \subset A$ jollakin $\delta > 0$,
- *suljettu*, jos kaikille reaalilukujonoille $\{x_n\}$ pätee:
  $x_n \in A \ \forall n \ \ \& \ \ x_n \to x \in \mathbb{R} \ \Rightarrow \ x \in A$,
- *rajoitettu*, jos $\exists C \in \mathbb{R}_+$ siten, että $|x| \le C \ \forall x \in A$,
- *kompakti*, jos A on suljettu ja rajoitettu.

ESIMERKKI 1 Avoin väli (a, b) on avoin joukko, sillä jos $x \in (a, b)$, nii $U_\delta(x) \subset (a, b)$, kun $0 < \delta \le \min\{x - a, b - x\}$. Suljettu väli on vastaavasti suljettu joukko, sillä jos $x_n \in [a, b] \ \forall n$ ja $x_n \to x \in \mathbb{R}$, niin $a \le x \le b$ (Lause I.7.4 [V1]) eli $x \in [a, b]$. Koska suljettu väli on myös rajoitettu ($|x| \le C = \max\{|a|, |b|\} \ \forall x \in [a, b]$), niin suljettu väli on kompakti joukko. $\square$

Esimerkki 2

$[0,1] \cup [2,3]$	on kompakti (ei avoin),
$(0,1) \cup (2,3)$	on avoin ja rajoitettu (ei suljettu),
$\mathbb{R}$	on avoin ja suljettu (ei rajoitettu),
$[0,1)$	on rajoitettu (ei avoin eikä suljettu),
$[0,\infty)$	on suljettu (ei avoin eikä rajoitettu),
$(0,\infty)$	on avoin (ei suljettu eikä rajoitettu),
$\mathbb{Q}$	ei avoin, ei suljettu eikä rajoitettu,
$\emptyset$	(tyhjä) on avoin ja kompakti. $\square$

Määritelmä V.9.2 Joukon $A \subset \mathbb{R}$

- *komplementti* on $\complement(A) = \{x \in \mathbb{R} \mid x \notin A\}$.
- *sulkeuma* (engl. closure) on

$$\overline{A} = \{x \in \mathbb{R} \mid \exists \text{ jono } \{x_n\} \text{ siten, että } x_n \in A \ \forall n \quad \& \quad x_n \to x\}.$$

- *reuna* (engl. boundary) on $\partial A = \overline{A} \cap \overline{\complement(A)}$.
- *sisäpisteiden joukko* on $A_0 = \complement(\overline{\complement(A)})$.

Joukon sulkeuma on nimensä mukaisesti suljettu joukko, ja toinen määritelmä onkin: $\overline{A}$ on pienin suljettu joukko, joka sisältää A:n. Sisäpisteiden joukko A_0 on avoin (ks. Lause V.9.3 jäljempänä); tämän vaihtoehtoisia määritelmiä ovat: A_0 on suurin A:n avoin osajoukko, tai

$$A_0 = \{x \in A \mid x \notin \partial A\}.$$

Pätee myös $\overline{A} = A_0 \cup \partial A$, $A_0 \cap \partial A = \emptyset$, samoin pätee

$$A \text{ suljettu} \ \Leftrightarrow \ A = \overline{A} \ \Leftrightarrow \ \partial A \subset A.$$

Esimerkki 3 Joukkojen

$$A = (0,1], \quad B = (-1,0) \cup (0,1), \quad C = \mathbb{Q}$$

sulkeumat, reunat ja sisäpisteiden joukot ovat

$$\begin{array}{rclrclrcl} \overline{A} & = & [0,1], & \overline{B} & = & [-1,1], & \overline{C} & = & \mathbb{R}, \\ \partial A & = & \{0,1\}, & \partial B & = & \{-1,0,1\}, & \partial C & = & \mathbb{R}, \\ A_0 & = & (0,1), & B_0 & = & B, & C_0 & = & \emptyset. \end{array} \quad \square$$

Haluttaessa selvittää, onko annettu joukko avoin, suljettu tai kompakti, helpottuu tehtävä usein huomattavasti seuraavia loogisia väittämiä hyödyntämällä. Väittämistä ensimmäisen voi tulkita myös suljetun tai avoimen joukon määritelmäksi, jos vain toinen käsitteistä on määritelty erikseen.

Lause V.9.3 Pätee

$$A \text{ avoin} \quad \Leftrightarrow \quad \complement(A) \text{ suljettu}, \qquad A \text{ suljettu} \quad \Leftrightarrow \quad \complement(A) \text{ avoin}.$$

Lause V.9.4 Pätee

$$\begin{aligned} &A, B \text{ avoimia/suljettuja/kompakteja} \\ &\Rightarrow \; A \cup B \text{ ja } A \cap B \text{ avoimia/suljettuja/kompakteja}. \end{aligned}$$

<u>Todistus</u> Väittämiin sisältyy yhteensä kymmenen implikaatioväittämää, joiden todistukset ovat kaikki melko lyhyitä. Todistetaan esimerkkinä ainoastaan väittämä A avoin $\Rightarrow$ $\complement(A)$ suljettu eli: Jos A on avoin, niin jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$x_n \in \complement(A) \; \forall n \quad \& \quad x_n \to x \quad \Rightarrow \quad x \in \complement(A).$$

Tehdään vastaoletus: $x \notin \complement(A) \;\Leftrightarrow\; x \in A$. Tällöin $(x-\delta, x+\delta) \subset A$ jollakin $\delta > 0$, koska A oli avoin. Tällöin koska $x_n \in \complement(A)$ $\forall n$, on oltava

$$|x_n - x| \geq \delta \;\; \forall n \quad \Rightarrow \quad x_n \not\to x.$$

Tämä on looginen ristiriita, koska oletettiin, että $x_n \to x$. $\square$

Esimerkki 4 Joukko $A = (-\infty, a) \cup (b, \infty)$ on Lauseen V.9.4 (tai suoraan Määritelmän V.9.1) mukaan avoin. Lauseen V.9.3 mukaisesti komplementti

$$\complement(A) = \begin{cases} [a, b], & \text{kun } a < b, \\ \{a\}, & \text{kun } a = b, \\ \emptyset, & \text{kun } a > b \end{cases}$$

on suljettu. $\square$

Esimerkki 5 Yhden alkion sisältävä joukko $A = \{a\}$ on Määritelmän V.9.1 mukaan kompakti, joten Lauseen V.9.4 mukaan samoin on jokainen kahden, kolmen, jne. alkion joukko. Siis jokainen äärellinen joukko on kompakti. $\square$

Esimerkki 6 Rationaalifunktion $f(x) = p(x)/q(x)$ (p ja q polynomeja) määrittelyjoukko on $\mathcal{D}_f = \complement(A)$, missä $A = \{q\text{:n nollakohdat}\}$. Äärellisenä joukkona A on suljettu (jopa kompakti), joten Lauseen V.9.3 mukaan $\mathcal{D}_f$ on avoin. $\square$

Weierstrassin lause

Funktion f jatkuvuus suljetulla välillä $[a, b] \subset \mathcal{D}_f$ tarkoitti Määritelmän V.1.5 mukaisesti jatkuvuutta 'sisältä päin' ko. välillä. Asetetaan vastaavalla tavalla yleisempi määritelmä.

MÄÄRITELMÄ V.9.5 Reaaliunktio f on *jatkuva kompaktissa joukossa* $K \subset \mathcal{D}_f$, jos jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$x_n \in K \ \forall n \quad \& \quad x_n \to x \quad \Rightarrow \quad f(x_n) \to f(x).^{\dagger}$$

*LAUSE V.9.6 Jos f on jatkuva kompaktissa joukossa $K \subset \mathcal{D}_f$, niin $f(K)$ on rajoitettu joukko.

Todistus Käytetään epäsuoraa todistustapaa, eli tehdään vastaoletus: $f(K)$ ei ole rajoitettu. Tällöin on olemassa jono $\{x_n\}$ siten, että

$$x_n \in K \ \forall n \quad \& \quad |f(x_n)| \to \infty.$$

Koska A on (kompaktina joukkona) rajoitettu, on jono $\{x_n\}$ rajoitettu. Tällöin jonolla on Bolzanon–Weierstrassin lauseen (Lauseen I.10.12) mukaan suppeneva osajono. Kun indeksoidaan tämä osajono jonoksi $\{x_k\}$, niin pätee siis $x_k \to x \in \mathbb{R}$. Koska K on (kompaktina joukkona) suljettu, on oltava $x \in K$. On siis löydetty jono $\{x_k\}$, jolle pätee $x_k \in K \ \forall k$ ja $x_k \to x \in K$. Tällöin jatkuvuusoletuksen ja Määritelmän V.9.5 mukaan on oltava $f(x_k) \to f(x)$. Toisaalta koska alkuperäiselle jonolle $\{f(x_n)\}$ pätee $|f(x_n)| \to \infty$, niin myös $|f(x_k)| \to \infty$. Oletuksista ja vastaoletuksesta seurasi siis, että on olemassa jono $\{x_k\}$, jolle pätee sekä $f(x_k) \to f(x) \in \mathbb{R}$ että $|f(x_k)| \to \infty$. Tämä on looginen ristiriita, joten lause on todistettu. $\square$

Seuraava tulos sisältää erikoistapauksena Weierstrassin lauseen V.1.12.

*LAUSE V.9.7 Jos reaalifunktio f on jatkuva kompaktissa joukossa $K \subset \mathcal{D}_f$, niin f saavuttaa K:ssa pienimmän ja suurimman arvonsa.

Todistus Koska $f(K)$ on Lauseen V.9.6 mukaan rajoitettu, niin Lauseen I.11.2 mukaan tällä joukolla on supremum: $\sup f(K) = y \in \mathbb{R}$. Supremumin määritelmän (ks. Luku I.11) mukaisesti on olemassa jono $\{y_n\}$, jolle pätee $y_n \in f(K) \ \forall n$ ja $y_n \to y$. Koska $y_n \in f(K)$, on $f(x_n) = y_n$ jollakin $x_n \in K$. Koska K on (kompaktina joukkona) rajoitettu, on jono $\{x_n\}$ on rajoitettu, joten sillä on Lauseen

†Huomautettakoon, että määritelmän ehto toteutuu mille tahansa reaalifunktiolle f, jos $K \subset \mathcal{D}_f$ on äärellinen (ja siis kompakti) joukko. Vrt. alkuperäinen jatkuvuuden määritelmä (Määritelmä V.1.1), jonka mukaan funktio on jatkuva määrittelyjoukkonsa eristetyissä pisteissä.

I.10.12 mukaan suppeneva osajono. Indeksoidaan tämä uudelleen jonoksi $\{x_k\}$, jolloin pätee $x_k \to x \in \mathbb{R}$. Koska K on (kompaktina joukkona) suljettu, niin on oltava $x \in K$. Tällöin oletuksen ja Määritelmän V.9.5 mukaan $f(x_k) \to f(x)$. Mutta jono $\{f(x_k)\}$ on alkuperäisen jonon $\{f(x_n)\}$ osajono, joten pätee myös $f(x_k) \to y$ (Lause I.10.7). Siis $f(x) = y$ (koska lukujonon raja-arvo on yksikäsitteinen). Koska y on $f(K)$:n yläraja, on $f(t) \leq y \ \forall t \in K$ — siis f saavuttaa joukossa K suurimman arvonsa pisteessä x. Lause on näin todistettu maksimiarvon osalta. Minimiarvon osalta lause todistetaan joko vastaavalla päättelyllä tai soveltamlla jo todistettua väittämää funktioon $-f$. $\square$

Esimerkki 7 Funktio

$$f(x) = \begin{cases} 1/x, & \text{kun } x > 0, \\ 0, & \text{kun } x = 0 \end{cases}$$

on jatkuva (kompaktissa) joukossa $K = [a, b]$, jos $0 < a < b$. Tällöin f:n maksimimiarvo K:ssa $= f(a)$ ja minimiarvo $= f(b)$. Kompaktilla välillä $[0, 1] \subset \mathcal{D}_f$ ei f ole Määritelmän V.9.5 mukaisesti jatkuva, eikä f myöskään saavuta maksimiarvoaan tällä välillä (minimiarvo $= 0$). Välin $A = (0, 1]$ (rajoitettu, ei suljettu) jokaisessa pisteessä f on jatkuva, mutta f ei saavuta A:ssa maksimiarvoaan (minimiarvo $= 1$). Välillä A=(1,2) (rajoitettu, ei suljettu) f ei saavuta kumpaakaan ääriarvoaan. Välillä $A = [1, \infty)$ (suljettu, ei rajoitettu) f on jatkuva jokaisessa pisteessä, mutta f ei saavuta A:ssa minimiarvoaan (maksimiarvo $= 1$). $\square$

Algebran peruslause

Käsitteet avoin, suljettu ja kompakti joukko määritellään kompleksitasossa aivan samalla tavoin kuin $\mathbb{R}$:ssä. Erona on ainoastaan, että avoimen joukon määrittelyssä tarvittava ympäristö on $\mathbb{C}$:ssä kiekon muotoinen: $U_\delta(c) = \{z \in \mathbb{C} \mid |z - c| < \delta\}$ (Määritelmä V.8.1). Lauseet V.9.3 ja V.9.4 pätevät myös kompleksitasossa.

Esimerkki 8 Kiekko $K = \{z \in \mathbb{C} \mid |z| \leq R\}$ on suljettu, ts. jokaiselle kompleksilukujonolle $\{z_n\}$ pätee: $z_n \in K \, \forall n \ \& \ z_n \to z \ \Rightarrow \ z \in K$. K on myös rajoitettu ($|z| \leq C = R \ \forall z \in K$), joten K on kompakti joukko. Komplementti $\complement(K) = \{z \in \mathbb{C} \mid |z| > R\}$ on avoin joukko. $\square$

Kompleksifunktion jatkuvuus määriteltiin edellisessä luvussa vastaavalla tavalla kuin reaalifunktioille, ja myös jatkuvuus kompaktissa joukossa (Määritelmä V.9.5) yleistyy vastaavasti koskemaan kompleksifunktioita. Weierstrassin lauseen vastine kompleksifunktioille on seuraava väittämä, jonka todistus noudattaa hyvin tarkoin Lauseen V.9.7 todistuksen logiikkaa. (Todistus sivuutetaan.)

*Lause V.9.8 Jos f on kompaktissa joukossa $K \subset \mathbb{C}$ jatkuva kompleksiarvoinen funktio, niin $|f|$ saavuttaa K:ssa pienimmän ja suurimman arvonsa.

Algebran peruslauseen (Lause III.3.2) todistus voidaan perustaa tähän tulokseen sekä Luvun III.3 tarkasteluihin: Olkoon $p(z)$ polynomi astetta $n \geq 1$ (ei vakio). Tällöin $|p(z)|$ saavuttaa Lauseen V.9.8 mukaan minimiarvonsa kompaktissa joukossa

$$K = \{z \in \mathbb{C} \mid |z| \leq R\}$$

jokaisella $R \in \mathbb{R}_+$. Koska $|p(z)| \sim |z|^n$, kun $|z| \to \infty$ (ks. Propositio III.2.3), niin päätellään, että R:n ollessa riittävän suuri on $|p|$:n minimikohta K:ssa samalla $|p|$:n absoluuttinen minimikohta $\mathbb{C}$:ssä. Siis $|p|$ saavuttaa jossakin pisteessä $c \in \mathbb{C}$ absoluuttisen minimiarvonsa. Tällöin on seuraavan väittämän mukaan oltava $p(c) = 0$, jolloin Algebran peruslause on todistettu.

Lause V.9.9 Jos kompleksimuuttujan polynomi p ei ole vakio, niin $p(c) = 0$ jokaisessa pisteessä $c \in \mathbb{C}$, jossa $|p|$:llä on paikallinen minimi.

Todistus Ks. Harj.teht. III.3:13. $\square$

Käänteisfunktion jatkuvuus

Seuraava käänteisfunktion jatkuvuuden takaava lause on yleistys Lauseesta V.1.14. Todistuksessa on Bolzanon–Weierstrassin lause jälleen keskeisessä roolissa.

*Lause V.9.10 Jos f on jatkuva kompaktissa joukossa $A \subset \mathcal{D}_f$ ja $f : A \to B$ on bijektio, niin B on kompakti ja $f^{-1} : B \to A$ on samoin jatkuva kompaktissa joukossa B.

Todistus Näytetään ensin, että B on kompakti joukko. Lauseen V.9.6 mukaan B on rajoitettu, joten riittää osoittaa, että B on suljettu. Oletetaan siis, että $y_n \in B \; \forall n$ ja että $y_n \to y \in \mathbb{R}$. Koska $y_n \in B = f(A)$, niin jokaisella n on $y_n = f(x_n)$ jollakin $x_n \in A$. Koska A on (kompaktina joukkona) rajoitettu, niin jono $\{x_n\}$ on rajoitettu, jolloin tällä jonolla on osajono $\{x_k\}$, joka suppenee: $x_k \to x \in \mathbb{R}$ (Lause I.10.12). Koska A on (kompaktina joukkona) suljettu, niin $x \in A$. Tällöin koska f on x:ssä jatkuva Määritelmän V.9.5 mukaisesti, niin $f(x_k) \to f(x)$. Mutta $\{f(x_k)\} = \{y_k\}$ on jonon $\{y_n\}$ osajono, joten $f(x_k) = y_k \to y$. Lukujonon rajaarvon yksikäsitteisyyden perusteella on silloin $f(x) = y$, joten $y \in f(A) = B$. On näytetty, että jokaiselle lukujonolle $\{y_n\}$ pätee: $y_n \in B \; \forall n \;\; \& \;\; y_n \to y \;\Rightarrow\; y \in B$. Siis B on suljettu.

Käänteisfunktion f^{-1} väitetyn jatkuvuuden osoittamiseksi on näytettävä: Jos $y_n \in B,\ n = 1, 2, \ldots$ ja $f(x_n) = y_n\,,\ x_n \in A$, niin pätee

$$y_n \to y \in B \quad \Rightarrow \quad x_n \to x \in A \ \ \& \ \ x = f^{-1}(y).$$

Koska $x_n \in A$ ja A on kompakti, niin jono $\{x_n\}$ on rajoitettu, joten jonolla on suppeneva osajono $\{x_k\}:\ x_k \to x \in \mathbb{R}$. Koska A on kompakti, niin $x \in A$, jolloin f:n jatkuvuuden nojalla $f(x_k) \to f(x)$. Mutta jono $\{f(x_k)\}$ on jonon $\{f(x_n)\} = \{y_n\}$ osajono, joten oletuksen $y_n \to y$ mukaan on oltava myös $f(x_k) \to y$. Siis $f(x) = y \ \Leftrightarrow\ x = f^{-1}(y)$. On päätelty, että jonolla $\{x_n\}$ on ainakin osajono $\{x_k\}$, jolle pätee $x_k \to x \in A$ ja $x = f^{-1}(y)$. Näytetään nyt, että myös alkuperäiselle jonolle pätee $x_n \to x$, jolloin lause on todistettu. Tehdään vastaoletus: $x_n \not\to x$. Tällöin Lauseen I.10.8 mukaan on olemassa toinen osajono $\{x_l\}$ ja $\varepsilon > 0$ siten, että pätee

$$|x_l - x| \geq \varepsilon \quad \forall l \in \mathbb{N}.$$

Tälläkin osajonolla on kuitenkin suppeneva osajono $\{x_\nu\}$, jolle siis pätee $x_\nu \to x' \neq x$ kun $\nu \to \infty$. Tällöin on jälleen $x' \in A$, koska A on suljettu, joten f:n jatkuvuuden perusteella $f(x_\nu) \to f(x')$. Toisaalta koska jono $\{f(x_\nu)\}$ on edelleen alkuperäisen jonon $\{f(x_n)\} = \{y_n\}$ osajono, niin oletuksesta $y_n \to y$ seuraa, että myös $f(x_\nu) \to y$. Siis pätee sekä $f(x_\nu) \to f(x')$ että $f(x_\nu) \to y$, jolloin on oltava $f(x') = y = f(x)$. Oletuksista ja vastaoletuksesta on näin johdettu päätelmä: On olemassa $x \in A$ ja $x' \in A$ siten, että $x \neq x'$ ja $f(x) = f(x')$. Mutta (toistaiseksi käyttämättömän) oletuksen mukaan f on injektio, joten $x \neq x' \ \Rightarrow\ f(x) \neq f(x')$. Siis $f(x) = f(x')$ ja $f(x) \neq f(x')$ — looginen ristiriita, joka osoittaa tehdyn vastaoletuksen vääräksi. Siis $x_n \to x = f^{-1}(y)$. $\square$

Esimerkki 9 Määritellään funktio f joukossa

$$D_f = A = [-1, 0] \cup (1, 2]$$

seuraavasti:

$$f(x) = \begin{cases} x & , \text{ kun } x \in [-1, 0] \\ x - 1 & , \text{ kun } x \in (1, 2] \end{cases}$$

Tällöin f on koko määrittelyjoukossaan jatkuva ja $f : A \to [-1, 1]$ on bijektio, joten myös $f^{-1} : [-1, 1] \to A$ on bijektio (vrt. Luku IV.2). Mutta f^{-1} ei ole jatkuva pisteessä $x = 0$. $\square$

Esimerkissä on käänteisfunktion jatkuvuuden kannalta ongelmana, että määrittelyjoukon osaväli $(1, 2]$ ei ole suljettu. Ongelmaa ei voi poistaa ottamalla $x = 1$ mukaan määrittelyjoukkoon, sillä jos asetetaan $f(1) = 0$, niin $f(0) = f(1)$, jolloin f ei ole injektio, ja jos asetetaan $f(1) = c \neq 0$, niin f ei ole (oikealta) jatkuva pisteessä $x = 1$.

Tasainen jatkuvuus

Tarkastellaan vielä uutta jatkuvuuden käsitettä, jota voi pitää Luvussa V.6 määritellyn Lipschitz-jatkuvuuden minimalistisena vastineena (vrt. Määritelmä V.6.4 ja Lause V.9.12 alla).

MÄÄRITELMÄ V.9.11 Reaalifunktio f on joukossa $A \subset \mathcal{D}_f$ *tasaisesti jatkuva* (engl. uniformly continuous), jos kaikille reaalilukujonoille $\{x_n\}$ ja $\{t_n\}$ pätee

$$x_n, t_n \in A \;\;\&\;\; x_n - t_n \to 0 \;\Rightarrow\; f(x_n) - f(t_n) \to 0.$$

Jos f on tasaisesti jatkuva välillä $(c - \delta, c + \delta)$, $\delta > 0$, niin valitsemalla $y_n = c$ ym. määritelmässä nähdään, että f on jatkuva pisteessä c. Tasainen jatkuvuus on siis tässä mielessä vahvempi ominaisuus kuin Määritelmän V.1.1 mukainen jatkuvuus, jota myös tavataan sanoa *pisteittäiseksi* (engl. pointwise).

ESIMERKKI 10 $f(x) = 1/x$ on koko määrittelyjoukossaan (pisteittäin) jatkuva, mutta ei tasaisesti jatkuva: Esim. jos $x_n = 1/n$ ja $t_n = 2/n$, niin $x_n - t_n \to 0$, mutta $f(x_n) - f(t_n) = n/2 \not\to 0$.

Tasaisen jatkuvuuden merkitys näkyy selvemmin seuraavasta tuloksesta, jota voi myös pitää tasaisen jatkuvuuden vaihtoehtoisena määritelmänä, vrt. Lause V.2.2. Todistus (joka sivuutetaan) on idealtaan sama kuin Lauseen V.1.3 todistus.

*LAUSE V.9.12 Funktio $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, on joukossa $A \subset \mathcal{D}_f$ tasaisesti jatkuva täsmälleen kun jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että $\forall x_1, x_2 \in \mathbb{R}$ pätee

$$x_1, x_2 \in A \;\;\&\;\; |x_1 - x_2| < \delta \;\Rightarrow\; |f(x_1) - f(x_2)| < \varepsilon.$$

ESIMERKKI 11 Suljetulla välillä $A = [a, b]$ Lipschitz-jatkuva funktio (Määritelmä V.6.4) on ko. välillä myös tasaisesti jatkuva, sillä Lauseen V.9.12 ehto toteutuu jokaisella $\varepsilon > 0$, kun valitaan $\delta = \varepsilon/L$, missä $L =$ Lipschitz-vakio. $\square$

Seuraava lause — joka jälleen nojaa Bolzanon–Weierstrassin lauseeseen — kuuluu tavanomaisen reaalianalyysin hämmästyttävimpiin tuloksiin.

*Lause V.9.13 Jos reaalifunktio f on jatkuva kompaktissa joukossa $K \subset \mathcal{D}_f$, niin f on K:ssa tasaisesti jatkuva.

Todistus Tehdään vastaoletus: f ei ole K:ssa tasaisesti jatkuva. Tällöin on olemassa jonot $\{x_n\}$ ja $\{t_n\}$ siten, että

$$x_n, t_n \in K \ \forall n \quad \& \quad x_n - t_n \to 0 \quad \& \quad f(x_n) - f(t_n) \not\to 0.$$

Koska $f(x_n) - f(t_n) \not\to 0$, niin jonolla $\{f(x_n) - f(t_n)\}$ on Lauseen I.10.8 mukaan osajono $\{f(x_k) - f(t_k)\}$ siten, että jollakin $\varepsilon > 0$ pätee

$$|f(x_k) - f(t_k)| \geq \varepsilon \quad \forall k \in \mathbb{N}.$$

Mutta jonot $\{x_k\}$ ja $\{t_k\}$ ovat rajoitettuja (koska $x_k, t_k \in K$ ja K on rajoitettu), joten ensinnäkin jonolla $\{x_k\}$ on suppeneva osajono $\{x_l\}$ (Lause I.10.12). Kun nyt siirrytään tarkastelemaan jonon $\{t_k\}$ vastaavaa osajonoa $\{t_l\}$, niin tällä on edelleen suppeneva osajono $\{t_\nu\}$. Mutta tällöin vastaava jonon $\{x_l\}$ osajono $\{x_\nu\}$ suppenee myös (koska on suppenevan jonon osajono). Näin on löydetty lukuparien (x_k, t_k) muodostamasta jonosta osajono $\{(x_\nu, t_\nu)\}$, jossa pätee:

$$x_\nu \to x \in \mathbb{R} \quad \& \quad t_\nu \to t \in \mathbb{R}.$$

Tässä on oltava $x, t \in K$, koska K on suljettu, ja on myös oltava $x = t$, koska $x_n - t_n \to 0 \ \Rightarrow \ x_\nu - t_\nu \to 0$. Koska siis $x_\nu \to x$ ja $t_\nu \to x$, $x \in K$, niin f:n jatkuvuuden perusteella

$$f(x_\nu) - f(t_\nu) \to f(x) - f(x) = 0.$$

On siis päätelty, että $f(x_\nu) - f(t_\nu) \to 0$, ja toisaalta $|f(x_\nu) - f(t_\nu)| \geq \varepsilon > 0 \ \forall \nu$. Tämä on looginen ristiriita, joten vastaoletus on väärä ja lause siis todistettu. $\square$

HARJOITUSTEHTÄVIÄ

1. Mitä ominaisuuksista (avoin, suljettu, rajoitettu, kompakti) on seuraavilla $\mathbb{R}$:n osajoukoilla:
   a) $A = (0, 1) \cup \{\pi\}$
   b) $A = [-1000, 10)$
   c) $A = [-1000, 1000000000000000000000000]$
   d) $A = (-0.00000000002, 0.00000000000000007)$
   e) $A = \mathbb{N}$
   f) $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$
   g) $A = (-\infty, 3]$
   h) $A = \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup [-1, 0]$

2. Näytä, että jos $A \subset \mathbb{R}$, niin pätee

$$A \text{ kompakti} \quad \Rightarrow \quad \sup A = \max A \quad \& \quad \inf A = \min A.$$

Näytä vastaesimerkillä, että implikaatio ei päde kääntäen.

3. Todista seuraavat Lauseiden V.9.3 ja V.9.4 osaväittämät:
   a) A suljettu $\Rightarrow \complement(A)$ avoin
   b) A ja B avoimia $\Rightarrow A \cap B$ avoin
   c) A ja B suljettuja $\Rightarrow A \cup B$ suljettu

4. Määrittele seuraavien joukkojen reuna ∂A:
   a) $A = \{1, 3, 7\}$ b) $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ c) $A = (0, 1) \cap \mathbb{Q}$

5. Näytä, että jos f on $\mathbb{R}$:ssä jatkuva funktio ja $B \subset \mathbb{R}$ on suljettu joukko, niin myös $A = \{x \in \mathbb{R} \mid f(x) \in B\}$ on suljettu. Päättele, että erityisesti f:n nollakohtien joukko on suljettu.

6. Näytä, että yhtälö

$$8 - 63y^5 + 90y^7 - 35y^9 = x$$

määrittelee välillä $[0, 8]$ funktion $y = f(x)$ ja että $f : [0, 8] \to [0, a]$ on jatkuva bijektio eräällä — millä? — a:n arvolla.

7. Funktio f on välillä $[a, b]$ *Hölder–jatkuva indeksillä* α, $\alpha \in (0, 1] \cap \mathbb{Q}$, jos jollakin $C \in \mathbb{R}_+$ pätee

$$|f(x_1) - f(x_2)| \leq C|x_1 - x_2|^\alpha, \quad x_1, x_2 \in [a, b].$$

a) Näytä, että välillä $[a, b]$ Hölder–jatkuva funktio on ko. välillä tasaisesti jatkuva. b) Näytä, että funktio $f(x) = \sqrt[n]{x}$, $n \in \mathbb{N}$ on välillä $[0, 1]$ Hölder–jatkuva täsmälleen indekseillä $0 < \alpha \leq 1/n$.

8. (*) Olkoon f jatkuva välillä $[0, 1]$ ja muodostetaan lukujono $\{y_n\}$ seuraavasti:

$$y_n = \max_i \{f(i/2^n),\ i = 0 \ldots 2^n\}, \quad n = 0, 1, \ldots$$

Näytä, että $\{y_n\}$ on monotonisesti kasvava lukujono ja että

$$\max_{x \in [0,1]} f(x) = \lim_n y_n\,.$$

Löytyykö myös f:n maksimikohta tällä tavoin (algoritmisesti)?

Luku VI

Eksponenttifunktio

Tässä luvussa määriteltävää *eksponenttifunktiota* voi syystä pitää fysiikan yleisimpänä funktiona. Eksponenttifunktio esiintyy mitä moninaisimmissa matemaattisissa malleissa, jotka kuvaavat loputonta kasvua tai vähenemistä/vaimenemista. Fysiikan ohella tällaisia *eksponentiaalisen kasvun* tai *eksponentiaalisen vaimenemisen* malleja on paljon biologiassa, taloustieteissä, ym.

Matemaattisten funktioiden paljoudessakin eksponenttifunktio erottuu siinä määrin erikoislaatuisena, että se ansaitsee oman lukunsa. Jatkossa tarkastellaan tämän funktion ja sen käänteisfunktion, *logaritmifunktion*, matemaattista määrittelyä ja ominaisuuksia ensin reaalifunktioina (Luvut VI.1–VI.2). Sen jälkeen laajennetaan eksponenttifunktio kompleksifunktioksi ja tarkastellaan laajennuksen synnyttämiä johdannaisfunktioita sekä — hieman yllättäviä — yhteyksiä trigonometrisiin funktioihin (Luku VI.3). Viimeisessä osaluvussa käydään lyhyesti läpi eräitä reaalisen eksponenttifunktion sovellusesimerkkejä fysiikassa.

VI.1 Yleinen eksponenttifunktio $E(x)$

MÄÄRITELMÄ VI.1.1 Reaalifunktio E on *eksponenttifunktio*, jos

(E1) $E(x)$ on määritelty $\forall x \in \mathbb{R}$,

(E2) $E(0) \neq 0$,

(E3) E on jatkuva pisteessä $x = 0$,

(E4) $E(x + y) = E(x)E(y) \quad \forall x, y \in \mathbb{R}$.

Eksponenttifunktion keskeisintä aksioomaa E4 voi sovellustilanteissa usein pitää jopa luonnonlakina. Myös aksiooman E3 taustalla voi nähdä fysikaalisia syitä, vrt. sovellusesimerkit jäljempänä Luvussa VI.4.

LAUSE VI.1.2 Jokaisella eksponenttifunktiolla E on seuraavat ominaisuudet:

(a) $E(x) > 0 \quad \forall x \in \mathbb{R}$.

(b) $E(0) = 1$.

(c) E on jatkuva koko $\mathbb{R}$:ssä.

(d) Jos $x \in \mathbb{Q}$, niin $E(x)$ määräytyy yksikäsitteisesti luvusta $E(1) = b$ ja

$$\boxed{E(x) = b^x \quad \forall x \in \mathbb{Q}.}$$

(e) E on $\mathbb{R}$:ssa aidosti kasvava, kun $b = E(1) > 1$, ja aidosti vähenevä, kun $b < 1$. Jos $b = 1$, on $E(x) = 1 \ \forall x \in \mathbb{R}$.

<u>Todistus</u> (a) Koska $E(x) \cdot E(-x) = E(x - x) = E(0) \neq 0$ (E4, E2), on oltava $E(x) \neq 0 \ \forall x \in \mathbb{R}$. Toisaalta on myös $E(x) = E(\frac{x}{2} + \frac{x}{2}) = [E(\frac{x}{2})]^2 \geq 0$ (E4), joten on oltava $E(x) > 0 \ \forall x \in \mathbb{R}$.

(b) Koska $E(0) \neq 0$ (E2) ja $E(0) = E(0+0) = [E(0)]^2$ (E4), on oltava $E(0) = 1$.

(c) Jos $x_n \to x \in \mathbb{R}$, niin $x_n - x \to 0 \ \Rightarrow \ E(x_n - x) \to E(0) = 1$ (E3,(b)). Aksiooman E4 perusteella pätee tällöin

$$E(x_n) = E(x)E(x_n - x) \to E(x) \cdot 1 = E(x).$$

Koska tämä on tosi jokaiselle reaalilukujonolle, jolle $\lim_n x_n = x \in \mathbb{R}$, niin E on jatkuva x:ssä ja siis koko $\mathbb{R}$:ssä.

(d) Jos $E(1) = b$, niin aksiooman E4 mukaan

$$\begin{aligned} b = E\left(\sum_{k=1}^{n} \frac{1}{n}\right) &= \left[E\left(\frac{1}{n}\right)\right]^n \quad \forall n \in \mathbb{N} \\ \Rightarrow\ E\left(\frac{1}{n}\right) &= b^{1/n} \quad \forall n \in \mathbb{N} \\ \Rightarrow\ E\left(\frac{m}{n}\right) = E\left(\sum_{k=1}^{m} \frac{1}{n}\right) &= (b^{1/n})^m = b^{m/n} \quad \forall m, n \in \mathbb{N}. \end{aligned}$$

Koska $E(0) = 1 = b^{0/n}$ ja

$$E\left(\frac{m}{n}\right) \cdot E\left(-\frac{m}{n}\right) = E(0) = 1 \ \Rightarrow\ E\left(-\frac{m}{n}\right) = \left[E\left(\frac{m}{n}\right)\right]^{-1} = b^{-m/n},$$

niin on osoitettu:

$$E\left(\frac{m}{n}\right) = b^{m/n} \quad \forall m \in \mathbb{Z},\ n \in \mathbb{N} \quad \Leftrightarrow \quad \text{väite (d).}$$

(e) Olkoon $b > 1$ ja $x_1 < x_2$. Oletetaan ensin, että $x_1, x_2 \in \mathbb{Q}$, jolloin on $x = x_2 - x_1 \in \mathbb{Q}$ ja $x > 0$, joten $x = p/q$ jollakin $p, q \in \mathbb{N}$. Tällöin on väittämän (d) ja murtopotenssien laskusääntöjen perusteella

$$E(x_1) - E(x_2) \ = \ b^{x_1} - b^{x_2} \ = \ b^{x_1}\,(1 - b^x) \ = \ b^{x_1}\left(1 - \sqrt[q]{b^p}\right).$$

Tässä on $\sqrt[q]{b^p} > 1$, koska $b > 1$, joten $E(x_1) - E(x_2) < 0$. On päätelty, että jos $b > 1$, niin $\forall x_1, x_2 \in \mathbb{Q}$ pätee: $x_1 < x_2 \ \Rightarrow\ E(x_1) < E(x_2)$.

Seuraavaksi olkoon $x_1, x_2 \in \mathbb{R}$ ja $x_1 < x_2$ (edelleen $b > 1$). Valitaan $t_1, t_2 \in \mathbb{Q}$ siten, että $x_1 < t_1 < t_2 < x_2$ ja rationaalilukujonot $\{\alpha_n\}$ ja $\{\beta_n\}$ siten, että $\lim_n \alpha_n = x_1$, $\lim_n \beta_n = x_2$ ja lisäksi $\alpha_n < x_1\ \forall n$ ja $\beta_n > x_2\ \forall n$. Tällöin koska $\alpha_n, t_1, t_2, \beta_n \in \mathbb{Q}$ ja $\alpha_n < t_1 < t_2 < \beta_n$, niin aiemman päättelyn perusteella on $E(\alpha_n) < E(t_1) < E(t_2) < E(\beta_n)\ \forall n$. Toisaalta koska $\alpha_n \to x_1$ ja $\beta_n \to x_2$, niin väittämän (c) perusteella $E(\alpha_n) \to E(x_1)$ ja $E(\beta_n) \to E(x_2)$. Koska tässä on $E(\alpha_n) < E(t_1)\ \forall n$ ja $E(\beta_n) > E(t_2)\ \forall n$, niin seuraa $E(x_1) \le E(t_1)$ ja $E(x_2) \ge E(t_2)$ (Lause I.7.4 [V1]), joten on näytetty, että

$$E(x_1) = \lim_n E(\alpha_n) \le E(t_1) < E(t_2) \le \lim_n E(\beta_n) = E(x_2).$$

Siis $E(x_1) < E(x_2)$. Tämä perustui vain oletukseen, että $x_1, x_2 \in \mathbb{R}$ ja $x_1 < x_2$, joten väittämä (e) on todistettu tapauksessa $b > 1$. Tapauksessa $b < 1$ on todistus

vastaava. Tapauksessa $b = 1$ on väittämien (d) ja (c) mukaan $E(x) = 1\ \forall x \in \mathbb{Q}$ $\Rightarrow\ E(x) = 1\ \forall x \in \mathbb{R}$. $\square$

Lauseesta VI.1.2 on syytä huomauttaa, että siinä ei oteta tarkemmin kantaa, millainen eksponenttifunktioiden joukko on. Tulee ainoastaan näytetyksi, että funktio $E(x) = 1\ \forall x \in \mathbb{R}$ kuuluu joukkoon, jolloin jää jopa se mahdollisuus, että tämä on eksponenttifunktioista ainoa (!). Eksponenttifunktion olemassaolokysymys ehdolla $E(1) = b \neq 1$ onkin oma ongelmansa, jonka ratkaiseminen edellyttää syvällisempiä lukujonoteoreettisia tarkasteluja. Nämä tarkastelut esitetään luvun lopussa, jolloin tulee todistetuksi

*Lause VI.1.3 Aksioomat E1–E4 toteuttava eksponenttifunktio $E(x)$ on olemassa ja määräytyy yksikäsitteisesti ehdosta $E(1) = b$ jokaisella $b \in \mathbb{R}_+$.

Jatkossa käytetään eksponenttifunktiolle merkintää $E(x) = b^x$ myös kun $x \notin \mathbb{Q}$. Tällöin siis tarkoitetaan funktiota, joka toteuttaa aksioomien E1–E4 lisäksi ehdon $E(1) = b$. Mainitun merkinnän myötä tulee määritellyksi myös yleinen potenssifunktio $f(x) = x^\alpha,\ x > 0,\ \alpha \in \mathbb{R}$, eli rajoituksesta $\alpha \in \mathbb{Q}$ voidaan luopua.

Esimerkki 1 Luku π^π on eksponenttifunktion $E(x) = \pi^x$ ja potenssifunktion $f(x) = x^\pi$ yhteinen arvo π:ssä. Lauseen VI.1.2 väittämän (c) perusteella π^π on laskettavissa esimerkiksi raja-arvona $\pi^\pi = \lim_n a_n$, missä

$$\begin{aligned}\{a_n\} &= \{\,\pi^3,\ \pi^{3.1},\ \pi^{3.14},\ \ldots\}\\ &= \{\,\pi^3,\ \sqrt[10]{\pi^{31}},\ \sqrt[100]{\pi^{314}},\ \ldots\}.\end{aligned}$$

Algoritmina tämä on hitaanpuoleinen (tarkka arvo 36.46215960..):

$$\begin{aligned}a_0 &= 31.00627668..\\ a_1 &= 34.76679088..\\ a_2 &= 36.39574388..\\ a_3 &= 36.43743103..\\ a_4 &= 36.45829251..\\ a_5 &= 36.46204884..\qquad \square\end{aligned}$$

Eksponenttifunktion b^x keskeiset laskusäännöt ovat

$$\boxed{b^x b^y = b^{x+y},\quad a^x b^x = (ab)^x,\quad (b^x)^y = b^{xy},\quad a, b \in \mathbb{R}_+,\ x, y \in \mathbb{R}.}$$

Näistä säännöistä ensimmäinen on perusaksiooma E4, joka yleistää murtopotenseille tutun laskusäännön $b^x b^y = b^{x+y}\ \forall x, y \in \mathbb{Q}$. Muutkin säännöt voi tulkita

vastaavien murtopotenssien laskusääntöjen (vrt. Luku I.2) yleistyksiksi. Näiden perustelu jätetään harjoitustehtäviksi (Harj.teht. 4–5).

Koska eksponenttifunktio $E(x)$ on jatkuva ja aidosti monotoninen, kun $E(1) = b \neq 1$, niin Ensimmäisestä väliarvolauseesta (Lause V.1.10) ja ilmeisistä rajaarvotuloksista

$$\lim_{x\to\infty} b^x = \begin{cases} \infty, & \text{jos } b > 1, \\ 0, & \text{jos } b < 1 \end{cases}$$

on pääteltävissä:

$$\boxed{E(x) = b^x \ \text{ on bijektio } \ E : \mathbb{R} \to \mathbb{R}_+ \ \text{ jokaisella } b \in \mathbb{R}_+,\ b \neq 1.}$$

*Eksponenttifunktion konstruktio

Jatkossa esitettävä Lauseen VI.1.3 todistus on konstruktiivinen, ts. todistuksessa konstruoidaan luku $E(x) \in \mathbb{R}$ jokaisella $x \in \mathbb{R}$ niin, että aksioomat E1–E4 sekä lisäehto $E(1) = b \in \mathbb{R}_+$ toteutuvat. Koska Lauseen VI.1.2 perusteella jo tiedetään, että $E(x) = b^x\ \forall x \in \mathbb{Q}$, niin riittää määritellä $E(x)$ myös irrationaalisilla x:n arvoilla ja varmistaa, että näin määritellylle funktiolle aksioomat E3 ja E4 ovat voimassa. Todistetaan ensin aputulos.

Lemma VI.1.4 Jos $b \in \mathbb{R}_+$, niin funktio $F(x) = b^x$, $\mathcal{D}_F = \mathbb{Q}$, on jatkuva pisteessä $x = 0$, ts. jokaiselle rationaalilukujonolle $\{x_n\}$ pätee: $x_n \to 0 \ \Rightarrow\ F(x_n) \to 1$.

<u>Todistus</u> Jos $\{x_n\}$ on rationaalilukujono, jolle pätee $x_n \to 0$, niin $\forall k \in \mathbb{N}$ on olemassa indeksi $N_k \in \mathbb{N}$ siten, että $|x_n| < 1/k\ \forall n > N_k$ (Määritelmä I.6.1,

$\varepsilon = 1/k$). Koska funktio $F(x) = b^x$ on monotoninen $\mathbb{Q}$:ssa (jopa aidosti, jos $b \neq 1$, ks. Lauseen VI.1.2 väittämän (c) todistus), niin tällöin pätee

$$|b^{x_n} - 1| \leq |b^{1/k} - 1|, \quad \text{kun } n > N_k.$$

Tässä $b^{1/k} \to 1$, kun $k \to \infty$ (Propositio I.9.7), joten $\forall \varepsilon > 0$ on olemassa indeksi m siten, että $|b^{1/m} - 1| < \varepsilon$. Valitsemalla $k = m$ seuraa siis

$$|b^{x_n} - 1| < \varepsilon, \quad \text{kun } n > N_m\,.$$

Tässä $\varepsilon > 0$ on mielivaltainen ja $N_m \in \mathbb{N}$, joten lukujonon suppenemisen määritelmän mukaan $b^{x_n} \to 1$. $\square$

Konstruktio. Halutaan määritellä $E(x)$, kun $x \in \mathbb{R}$, $x \notin \mathbb{Q}$. Koska tiedetään, että E (sikäli kuin olemasa) on jatkuva $\mathbb{R}$:ssä (Lause VI.1.2 (c)), niin määrittelyn perustaksi voidaan ottaa jatkaminen (vrt. Luku V.2): Lähtien funktion b^x alkuperäisestä määrittelyjoukosta $\mathbb{Q}$ laajennetaan määrittelyjoukko $\mathbb{R}$:ksi jatkuvuuden perusteella. Tämä merkitsee, että jokaiselle rationaalilukujonolle $\{x_n\}$ on oltava voimassa

$$x_n \to x \in \mathbb{R} \quad \Rightarrow \quad \lim_n b^{x_n} = \lim_n E(x_n) = E(x). \tag{1}$$

Jotta tämä ehto voisi toimia $E(x)$:n määritelmänä, on ensinnäkin varmistettava, että $\forall x \in \mathbb{R}$ ja jokaiselle rationaalilukujonolle $\{x_n\}$ pätee

$$x_n \to x \quad \Rightarrow \quad \{E(x_n)\} \text{ on Cauchy.} \tag{2}$$

Sikäli kuin tämä on tosi, niin Cauchyn suppenemiskriteerin (Lause I.10.5) mukaan $E(x_n) \to y \in \mathbb{R}$. Tämä ei kuitenkaan vielä todista funktioriippuvuutta $x \mapsto y$, sillä on mahdollista, että y riippuu jonosta $\{x_n\}$ eikä vain sen raja-arvosta x. Siksi on vielä näytettävä, että jokaisella $x \in \mathbb{R}$ ja kaikille rationaalilukujonoille $\{x_n\}$ ja $\{x'_n\}$ pätee

$$x_n, x'_n \in \mathbb{Q} \;\;\&\;\; x_n \to x \;\;\&\;\; x'_n \to x \quad \Rightarrow \quad E(x_n) - E(x'_n) \to 0. \tag{3}$$

Jos ehdot (2) ja (3) toteutuvat kaikille rationaalilukujonoille $\{x_n\}$ ja $\{x'_n\}$, niin on näytetty, että jokaisella $x \in \mathbb{R}$ (myös kun $x \in \mathbb{Q}$) on olemassa yksikäsitteinen $y \in \mathbb{R}$ siten, että jokaiselle rationaalilukujonolle $\{x_n\}$ pätee

$$x_n \to x \in \mathbb{R} \quad \Rightarrow \quad E(x_n) \to y.$$

Tällöin y riippuu vain raja-arvosta x, jolloin riippuvuus $x \mapsto y$ voidaan ilmaista funktiosymbolilla E, eli voidaan kirjoittaa $y = E(x)$ jatkamisperiaatteen (1) mukaisesti.

Jatkamisen perustana olevat väittämät (2) ja (3) saadaan todistetuksi vedoten jo todettuun aksiooman E4 pätevyyteen murtopotenssien laskusääntönä (eli rationaalisilla x:n ja y:n arvoilla) ja Lemmaan VI.1.4. Ensinnäkin jos $\{x_n\}$ on rationaalilukujono, niin (E4):n mukaan

$$E(x_n) - E(x_m) = E(x_m)[E(x_n - x_m) - 1].$$

Jos tässä $x_n \to x \in \mathbb{R}$, niin $\{x_n\}$ on Cauchy, joten $x_n - x_m \to 0$, kun $n, m \to \infty$. Tällöin myös $E(x_n - x_m) \to 1$ (Lemma VI.1.4). Koska $\{E(x_n)\} = \{b^{x_n}\}$ on epäilemättä rajoitettu jono, niin päätellään (ks. Lause I.7.4 [V3])

$$\begin{aligned} n, m \to \infty \quad &\Rightarrow \quad [E(x_n - x_m) - 1] \;\to\; 0 \\ &\Rightarrow \quad E(x_m)[E(x_n - x_m) - 1] \;=\; E(x_n) - E(x_m) \;\to\; 0, \end{aligned}$$

mikä todistaa väittämän (2). Väittämä (3) näytetään toteen vastaavalla tavalla lähtien hajotelmasta

$$E(x_n) - E(x'_n) = E(x'_n)[E(x_n - x'_n) - 1].$$

Funktio $E(x)$ on näin määritelty yksikäsitteisesti koko $\mathbb{R}$:ssä tunnetun funktion b^x, $x \in \mathbb{Q}$, $b \in \mathbb{R}_+$ jatkona, jolloin myös ehto $E(1) = b$ toteutuu. Vielä on osoitettava, että näin määritelty E on eksponenttifunktio.

Aksiooma E3. Jatkuvuuden määritelmän mukaisesti on näytettävä, että jokaiselle reaalilukujonolle $\{x_n\}$ pätee

$$x_n \to 0 \quad \Rightarrow \quad E(x_n) \to E(0) = 1.$$

Olkoon siis $\{x_n\}$ reaalilukujono, jolle $\lim_n x_n = 0$. Valitaan rationaalilukujono $\{\alpha_n\}$ siten, että $\lim_n \alpha_n = 0$ ja lisäksi jokaisella n on joko (i) $0 \le x_n \le \alpha_n$ tai (ii) $\alpha_n \le x_n \le 0$. Lauseen VI.1.2 väittämän (e) mukaan edellä konstruoitu funktio E on $\mathbb{R}$:ssä monotoninen, sillä todistus perustui vain funktion $F(x) = b^x$, $x \in \mathbb{Q}$ ominaisuuksiin sekä jatkuvuusehtoon (1). Oletuksien (i) tai (ii) ja monotonisuuden perusteella

$$|E(x_n) - 1| \le |E(\alpha_n) - 1| \;\; \forall n.$$

Tässä on $\alpha_n \in \mathbb{Q}$ $\forall n$ ja $\alpha_n \to 0$, joten $E(\alpha_n) \to 1$ (Lemma VI.1.4). Siis myös $E(x_n) \to 1$, eli E on jatkuva pisteessä $x = 0$.

Aksiooma E4. Olkoon $x, y \in \mathbb{R}$ ja $\{x_n\}, \{y_n\}$ rationaalilukujonoja, joille pätee $x_n \to x$ ja $y_n \to y$. Tällöin jokaisella n pätee $E(x_n + y_n) = E(x_n)E(y_n)$, joten E:n määritelmän (1) ja lukujonojen raja-arvojen yhdistelysääntöjen perusteella

$$\begin{aligned} E(x+y) - E(x)E(y) &= \lim_n E(x_n + y_n) - \lim_n E(x_n) \cdot \lim_n E(y_n) \\ &= \lim_n [E(x_n + y_n) - E(x_n)E(y_n)] \\ &= \lim_n 0 = 0. \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Tarkista, mitkä eksponenttifunktion aksioomat toteuttaa

$$\text{a)}\quad E(x) = 0, \quad x \in \mathbb{R},$$

$$\text{b)}\quad E(x) = \begin{cases} b^x, & \text{jos } x \in \mathbb{Q}, \\ 0, & \text{jos } x \in \mathbb{R},\ x \notin \mathbb{Q}. \end{cases}$$

   Mikä tunnettu funktio sisältyy jälkimmäiseen joukkoon?

2. Olkoon $\alpha \in \mathbb{R}$. Näytä eksponenttifunktion ominaisuuksiin vedoten, että funktio $f(x) = x^\alpha$ on välillä $(0, \infty)$ aidosti kasvava, jos $\alpha > 0$, ja aidosti vähenevä, jos $\alpha < 0$.

3. Näytä, että approksimaation

$$\pi^e \approx \sqrt[100]{\pi^{271}} = a$$

   suhteelliselle virheelle pätee arvio

$$0 < \frac{\pi^e - a}{\pi^e} < \sqrt[100]{\pi} - 1.$$

4. Perustele eksponenttifunktion jatkuvuuteen vedoten laskusääntö $a^x b^x = (ab)^x$, $x \in \mathbb{R}$, $a, b \in \mathbb{R}_+$. Oletetaan sääntö tunnetuksi, kun $x \in \mathbb{Q}$.

5. (*) Halutaan todistaa laskusääntö $(b^x)^y = b^{xy}$ ($b \in \mathbb{R}_+$, $x, y \in \mathbb{R}$) eli sääntö

$$E(x)^y = E(xy), \quad x, y \in \mathbb{R},$$

   kun $E(x) = b^x$. a) Näytä, että jos $y \in \mathbb{Q}$, niin sääntö seuraa aksioomasta E4. b) Näytä eksponenttifunktion jatkuvuuteen vedoten, että sääntö on pätevä myös, kun $y \in \mathbb{R}$, $y \notin \mathbb{Q}$.

VI.2 Funktiot e^x ja $\ln x$

Edellisessä luvussa eksponenttifunktio $E(x)$ konstruoitiin jatkamalla tunnettu funktio b^x, $x \in \mathbb{Q}$. Tällä tavoin tuli varmistetuksi, että eksponenttifunktioita on olemassa, että luku $b = E(1)$ määrää $E(x)$:n yksikäsitteisesti, ja että eksponenttifunktion arvot voidaan laskea numeerisesti suoraan määritelmästä käsin. Mitä ei tullut varmistetuksi on esimerkiksi, onko $E(x)$ mahdollisesti myös derivoituva, eikä vain jatkuva. Yleisemminkin jäi avoimeksi, kuinka säännöllisestä funktiosta on ylipäänsä kyse. Seuraavassa etsitään vastausta näihin kysymyksiin lähestymällä eksponenttifunktiota toisesta suunnasta.

Lähtökohdaksi otetaan eksponenttifunktion aksiooman E3 (ks. edellinen luku) vahvistaminen muotoon

(E3') E on derivoituva pisteessä $x = 0$.

Koska tiedetään, että $E(0) = 1$ (Lause VI.1.2 (b)), niin aksiooman E3' mukaan on olemassa luku $a \in \mathbb{R}$ siten, että

$$E'(0) = \lim_{x \to 0} \frac{E(x) - 1}{x} = a.$$

Kun yhdistetään tämä tulos ja aksiooma E4, niin seuraa

$$\lim_{\Delta x \to 0} \frac{E(x + \Delta x) - E(x)}{\Delta x} = \lim_{\Delta x \to 0} E(x) \frac{E(\Delta x) - 1}{\Delta x} = aE(x), \quad x \in \mathbb{R}.$$

Siis E on derivoituva jokaisessa pisteessä $x \in \mathbb{R}$ ja

$$E'(x) = aE(x), \quad x \in \mathbb{R}.$$

On päätelty, että jos aksioomat E1, E2, E3', E4 toteuttava funktio $E(x)$ on olemassa, niin tämä on derivoituva $\mathbb{R}$:ssä ja jollakin $a \in \mathbb{R}$ ratkaisu probleemalle

$$\text{(P)} \quad \begin{cases} y' = ay, & x \in \mathbb{R}, \\ y(0) = 1. \end{cases}$$

Tässä on kyse alkuarvotehtävästä, jossa on ratkaistava differentiaaliyhtälö $y' = ay$ alkuehdolla $y(0) = 1$ (vrt. Luvussa V.6 tarkasteltu vastaava probleema differentiaaliyhtälölle $y' = f(x)$).

Lause VI.2.1 Probleemalla (P) on ratkaisu $y(x) = E(x)$, missä

$$E(x) = \sum_{k=0}^{\infty} \frac{(ax)^k}{k!} = 1 + ax + \frac{(ax)^2}{2} + \dots$$

Tämä ratkaisu on yksikäsitteinen, ja $E(x)$ on eksponenttifunktio.

Todistus Potenssisarjan derivoimissäännöstä (Lause V.3.4) nähdään välittömästi, että väitetty $y(x)$ on probleeman (P) ratkaisu. Nähdään myös, että funktiolla $E(x)$ on eksponenttifunktiolta vaadittavat ominaisuudet (E1)–(E3), sillä $E(0) = 1$ ja E on määritelty koko $\mathbb{R}$:ssä suppenevan potenssisarjan summana, siis derivoituvana ja jopa sileänä funktiona.

Näytetään seuraavaksi, että $E(x)$ toteuttaa myös eksponenttifunktion aksiooman (E4). Tätä silmällä pitäen tarkastellaan ensin funktiota

$$u(x) = E(x)E(-x), \quad x \in \mathbb{R}.$$

Koska $E' = aE$, niin tulon derivoimissääntöä käyttäen todetaan, että

$$u'(x) = 0, \quad x \in \mathbb{R}. \tag{1}$$

Koska $u(0) = 1$ ja siis $u' = 0$, niin seuraa, että $u(x) = 1 \ \forall x \in \mathbb{R}$ (Lause V.6.7). Siis $E(x)E(-x) = 1 \ \forall x \in \mathbb{R}$, ja näin ollen $E(x) \neq 0 \ \forall x \in \mathbb{R}$ ja

$$E(-x) = [E(x)]^{-1}, \quad x \in \mathbb{R}. \tag{2}$$

Olkoon seuraavaksi $t \in \mathbb{R}$ kiinnitetty ja tarkastellaan funktiota

$$u(x) = E(t + x)E(-t)E(-x), \quad x \in \mathbb{R}.$$

Derivoimalla muuttujan x suhteen ja käyttämällä differentiaaliyhtälöä $E' = aE$ todetaan, että jälleen pätee (1). Koska tuloksen (2) mukaan on $u(0) = 1$, niin päätellään kuten edellä ja käyttämällä uudelleen tulosta (2), että

$$u(x) = E(t + x)E(-t)E(-x) = 1 \ \Leftrightarrow \ E(t + x) = E(t)E(x), \quad x \in \mathbb{R}.$$

Tämä pätee jokaisella $t \in \mathbb{R}$, joten aksiooma E4 on voimassa funktiolle $E(x)$. Siis $E(x)$ on eksponenttifunktio (jokaisella $a \in \mathbb{R}$).

Enää on todistamatta probleeman (P) ratkaisun yksikäsitteisyys. Tätä silmällä pitäen olkoon $y(x)$ (P):n (mikä tahansa) ratkaisu ja tarkastellaan funktiota

$$u(x) = y(x)E(-x), \quad x \in \mathbb{R}.$$

Derivoimalla ja käyttämällä differentiaaliyhtälöitä $y' = ay$ ja $E' = aE$ todetaan, että jälleen pätee (1). Koska $u(0) = 1$, niin seuraa

$$u(x) = y(x)E(-x) = 1 \ \Leftrightarrow \ y(x) = E(x), \quad x \in \mathbb{R}. \quad \square$$

Sanotaan jatkossa *perusekspponenttifunktioksi*, symboli exp, probleeman (P) ratkaisua, kun $a = 1$. Määritelmä potenssisarjana on siis

$$\boxed{\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}\,.}$$

Jos merkitään $e = \exp(1)$, niin edellisen luvun merkinnöin voidaan kirjoittaa

$$\boxed{\exp(x) = e^x.}$$

Tässä e on todellakin Neperin luku, sillä pätee (ks. Propositiot I.8.4 ja I.8.5)

$$\exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e = 2.71828182845905..$$

Tässä jälkimmäinen raja-arvotulos on erikoistapaus yleisemmästä tuloksesta (ks. Harj.teht. 16)

$$\boxed{\lim_n \left(1 + \frac{x}{n}\right)^n = e^x, \quad x \in \mathbb{R}.}$$

Perusekspponenttifunktion e^x määritelmästä seuraa derivoimissääntö

$$\boxed{\mathrm{D}e^x = e^x.}$$

Monet funktion e^x ominaisuuksista voidaan helposti johtaa tästä säännöstä tai potenssisarjaesityksestä. Esimerkiksi potenssisarjaesityksestä seuraa

$$e^x > x^m/m! \;\; \forall m \in \mathbb{N}, \quad \text{kun } x > 0,$$

mistä on edelleen helposti pääteltävissä raja-arvotulokset

$$\boxed{\lim_{x\to\infty} x^{-\alpha} e^x = \infty \quad \Leftrightarrow \quad \lim_{x\to\infty} x^{\alpha} e^{-x} = 0 \quad \forall \alpha \in \mathbb{R}.} \tag{3}$$

Näiden mukaan funktio e^x kasvaa nopeammin, ja funktio e^{-x} vähenee nopeammin, kuin mikään potenssifunktio, kun $x \to \infty$. Potenssisarjan perusteella voidaan myös e^x:n numeroarvoja laskea helposti, sillä pienillä $|x|$:n arvoilla (esim. kun $|x| \leq 1$) sarja suppenee nopeasti ja suuremmilla taas voidaan käyttää kaavaa $e^x = (e^{x/n})^n$. Derivoimissäännöstä (myös potenssisarjasta) nähdään, että funktio e^x on $\mathbb{R}$:ssä mielivaltaisen monta kertaa derivoituva (sileä) funktio.

Esimerkki 1 Jatka funktio $f(x) = (e^x - 1 - x)/x^2$ pisteeseen $x = 0$.

<u>Ratkaisu</u> e^x:n potenssisarjasta nähdään, että

$$f(x) = \frac{1}{2!} + \frac{x}{3!} + \frac{x^2}{4!} + \ldots = \sum_{k=0}^{\infty} \frac{x^k}{(k+2)!}, \quad x \neq 0.$$

Kun asetetaan $f(0) = 1/2$, niin $\forall x \in \mathbb{R}$ pätee siis

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{(k+2)!} = \begin{cases} \dfrac{e^x - 1 - x}{x^2}, & \text{kun } x \neq 0, \\ \frac{1}{2}, & \text{kun } x = 0. \end{cases}$$

Tässä potenssisarjan suppenemissäde on $\rho = \infty$, joten f ei ole vain jatkuva, vaan $\mathbb{R}$:ssä (myös pisteessä $x = 0$!) mielivaltaisen monta kertaa derivoituva. $\square$

Edellisessä luvussa todettiin, että yleinen eksponenttifunktio on muotoa $E(x) = b^x$, missä $b = E(1) \in \mathbb{R}_+$. Laskuissa usein kätevämpi on peruseksponenttifunktiosta johdettu vaihtoehtoinen esitystapa

$$\boxed{E(x) = e^{ax}, \quad E(1) = b = e^a.}$$

Luku $a \in \mathbb{R}$ määräytyy yksikäsitteisesti yhtälöstä $e^a = b$, koska $e^x : \mathbb{R} \to \mathbb{R}_+$ on bijektio. Jokainen eksponenttifunktio on siis muotoa $E(x) = e^{ax}$ jollakin $a \in \mathbb{R}$.

Todettakoon vielä lähtökohtana olleesta probleemasta (P), että sen ratkaisuina saadaan siis kaikki eksponenttifunktiot $y(x) = e^{ax}$, kun a:n arvoa vaihdellaan. Jos ratkaisuksi halutaan tietty eksponenttifunktio $E(x)$, niin a määräytyy joko em. tavalla luvun $b = E(1)$ kautta tai ehdosta $E'(0) = a$. Jos a on kiinnitetty, niin differentiaaliyhtälön $y' = ay$ <u>yleinen</u> <u>ratkaisu</u> (=kaikkien ratkaisujen joukko) on (Harj.teht. 1)

$$y(x) = Ce^{ax}, \quad C \in \mathbb{R}.$$

Alkuehdon $y(0) = y_0$ toteuttava ratkaisu on tämän mukaan yksikäsitteinen ja saadaan valitsemalla $C = y_0$.

Logaritmifunktio

Koska eksponenttifunktio $E : \mathbb{R} \to \mathbb{R}_+$ on bijektio kun $E(1) = b \neq 1$, on sillä käänteisfunktio, jota sanotaan *logaritmifunktioksi*. Käänteisfunktiota merkitään symbolilla $\log_b$:

$$y = b^x \;\Leftrightarrow\; x = \log_b y.$$

Peruseksponenttifunktion e^x käänteisfunktiota sanotaan *luonnolliseksi logaritmiksi*, symbolina ln (toisinaan log):

$$\boxed{y = e^x \;\Leftrightarrow\; x = \ln y.}$$

Luonnollinen logaritmi, samoin kuin muutkin logaritmifunktiot, on koko määrittelyjoukossaan ($= \mathbb{R}_+$) jatkuva ja aidosti kasvava, ja $\ln x \to \infty$ kun $x \to \infty$ ja

$\ln x \to -\infty$ kun $x \to 0^+$. Sijoituksilla $x = \pm \ln t \;\Leftrightarrow\; t = e^{\pm x}$ voidaan rajaarvotuloksista (3) päätellä, että rajoilla $x \to \infty$ ja $x \to 0^+$ funktio $|\ln x|$ kasvaa hitaammin kuin mikään potenssifunktio:

$$\boxed{\lim_{x\to\infty} x^{-\alpha} \ln x \;=\; \lim_{x\to 0^+} x^{\alpha} \ln x = 0 \quad \forall \alpha \in \mathbb{R}_+.} \tag{4}$$

Eksponentti- ja logaritmifunktioilla laskettaessa käytetään yleisimmin peruseksponenttifunktiota e^x ja sen käänteisfunktiota $\ln x$. Useimmin tarvittavat, suoraan määritelmästä ja eksponenttifunktion ominaisuuksista seuraavat kaavat ovat (Harj.teht. 2) :

$$e^{\ln x} = x, \qquad x \in \mathbb{R}_+. \tag{5}$$

$$\ln(xy) = \ln x + \ln y, \quad x, y \in \mathbb{R}_+. \tag{6}$$

$$\ln(1/x) = -\ln x, \qquad x \in \mathbb{R}_+. \tag{7}$$

$$\ln x^y = y \ln x, \qquad x \in \mathbb{R}_+,\; y \in \mathbb{R}. \tag{8}$$

Kaavat (5) ja (8) yhdistämällä saadaan laskukaava

$$\boxed{x^y = e^{y \ln x}, \quad x \in \mathbb{R}_+\,,\; y \in \mathbb{R}.} \tag{9}$$

Kaavasta (5), eksponenttifunktion derivoimissäännöstä ja yhdistetyn funktion derivoimissäännöstä saadaan johdetuksi logaritmifunktion $\ln x$ derivoimissääntö:

$$1 = \mathrm{D}x = \mathrm{D}e^{\ln x} = e^{\ln x}\mathrm{D}\ln x = x\mathrm{D}\ln x \quad \Rightarrow \quad \mathrm{D}\ln x = 1/x, \quad x > 0.$$

Tämän perusteella on $\mathrm{D}\ln(-x) = (-1)/(-x) = 1/x,\; x < 0$, joten saadaan yleisempi sääntö

$$\boxed{\mathrm{D}\ln|x| = \frac{1}{x}\,, \quad x \neq 0.} \tag{10}$$

Esimerkki 2 Kaavan (9) ja säännön (10) perusteella

$$\mathrm{D}x^x = \mathrm{D}e^{x\ln x} = e^{x\ln x}\mathrm{D}(x\ln x) = (\ln x + 1)\,x^x, \quad x > 0.$$

Tuloksesta nähdään, että $f(x) = x^x$ on aidosti vähenevä välillä $(0, 1/e]$ ja aidosti kasvava välillä $[1/e, \infty)$. Siis f:n absoluuttinen minimiarvo on $f_{min} = f(1/e) = e^{-1/e}$. Rajalla $x \to 0^+$ on $f(0^+) = 1$ (Harj.teht. 3e). $\square$

Esimerkki 3 Matemaatikon mökki on kahden tien risteyksessä. Selvitä mökin sijainti, kun tiedetään, että kummallakin tiellä on oheinen viitta.

<u>Ratkaisu</u> Tie 1 on ilmeisesti puolisuora $y = x$, $x > 0$. Tiellä 2 taas ovat esim. pisteet $(2, 4)$ ja $(4, 2)$. Teiden leikkauspisteen selvittämiseksi käytetään ensin kaavoja (5) ja (8) (ol. $x, y \in \mathbb{R}_+$):

$$x^y = y^x \;\Leftrightarrow\; e^{y\ln x} = e^{x\ln y} \;\Leftrightarrow\; y\ln x = x\ln y \;\Leftrightarrow\; \ln y/y = \ln x/x.$$

Säännöstä (10) ja Luvun V.3 derivoimissäännöistä seuraa

$$\mathrm{D}\ln x/x = x^{-2}(1 - \ln x), \quad x > 0,$$

joten funktio $f(x) = \ln x/x$ on aidosti kasvava välillä $(0, e]$ ja aidosti vähenevä välillä $[e, \infty)$. Näin ollen, jos $x \in \mathbb{R}_+$ ja $x > 1$, $x \neq e$, niin yhtälöllä $f(y) = f(x)$ on kaksi ratkaisua y, joista toinen on $y = x$ (ks. kuvio). Kun $x = e$ tai $x \leq 1$, on ainoa ratkaisu $y = x$. Päätellään, että mökki on pisteessä (e, e). $\square$

Esimerkki 4 $f(x) = \ln\left|\dfrac{1 - \cos x}{\sin x}\right|, \quad f'(x) = ?$

Ratkaisu Säännön (10) ja Lukujen V.3–V.4 derivoimissääntöjen perusteella

$$\begin{aligned} \mathrm{D}f(x) &= \frac{\sin x}{1-\cos x} \cdot \left[1 - \frac{(1-\cos x)\cos x}{\sin^2 x}\right] \\ &= \frac{\sin x}{1-\cos x} \cdot \frac{\sin^2 x + \cos^2 x - \cos x}{\sin^2 x} \\ &= \frac{1}{\sin x}. \quad \square \end{aligned}$$

Esimerkin tulos, ja muunnoksella $x \hookrightarrow \pi/2 - x$ saatava vastaava tulos, on syytä panna korvan taakse:

$$\boxed{\begin{aligned} \mathrm{D}\ln\left|\frac{1-\cos x}{\sin x}\right| &= \frac{1}{\sin x}, \qquad x \in \mathbb{R}, \quad \sin x \neq 0, \\ \mathrm{D}\ln\left|\frac{1-\sin x}{\cos x}\right| &= -\frac{1}{\cos x}, \qquad x \in \mathbb{R}, \quad \cos x \neq 0. \end{aligned}}$$

Tässä on itse asiassa $\dfrac{1-\cos x}{\sin x} = \dfrac{\sin x}{1+\cos x} = \tan\dfrac{x}{2}$ (ks. Harj.teht. II.5:13b).

HARJOITUSTEHTÄVIÄ

1. Näytä, että differentiaaliyhtälön $y' = ay$ yleinen ratkaisu $\mathbb{R}$:ssä on $y(x) = Ce^{ax}$. *Vihje*: Tutki funktiota $u(x) = e^{-ax}y(x)$.

2. a)–d) Perustele laskukaavat (5)–(8).

3. Määritä tai näytä oikeaksi raja-arvot

   a) $\displaystyle\lim_{x\to 0} \frac{e^{2x} - 1 - 2x}{x^2}$ b) $\displaystyle\lim_{x\to 0} \frac{1 - 6x + 18x^2 - e^{-6x}}{x^3}$

   c) $\displaystyle\lim_{x\to 0} \frac{3e^x - e^{-x} - 2e^{2x}}{x^2}$ d) $\displaystyle\lim_{x\to\infty} x\left(e^{4/x} - e^{1/x}\right)$

   e) $\displaystyle\lim_{x\to 0^+} x^x = 1$ f) $\displaystyle\lim_{x\to 0^+} x^{x^x} = 0$ g) $\displaystyle\lim_{x\to 0^+} x^{\sqrt{x}} = 1$

4. Derivoi seuraavat funktiot:
   a) 2^x b) $e^{\sqrt{x}}$ c) $\pi^{1/x}$ d) $\ln(\ln x)$ e) x^{2x} f) x^{x^x}

5. Määritä seuraavien funktioiden paikalliset ääriarvokohdat ja absoluuttiset maksimi- ja minimarvot, sikäli kuin olemassa. Hahmottele myös kuvaajat.
   a) $x\ln x$, b) $\sqrt[10]{|x|}\ln|x|$, c) $|x|e^{1/x^2}$, d) $x^{-\pi}\ln x$, e) $x^{1/x}$.

6. a) Millä a:n arvoilla funktio $f(x) = 2\ln x + x^2 - ax + 1$ on aidosti kasvava välillä $(0, \infty)$?

   b) Näytä, että jos $0 < a < b$, niin $1 - \dfrac{a}{b} < \ln\dfrac{b}{a} < \dfrac{b}{a} - 1$.

   c) Todista laskukaavat $\log_a b \cdot \log_b a = 1$ ja $\log_b a \cdot \log_c b = \log_c a$.

7. Määritä funktion $f(t) = e^{-x}\sin x$ paikalliset ääriarvokohdat sekä absoluuttiset maksimi- ja minimiarvot välillä $[0, \infty)$ sekä hahmottele f:n kuvaaja.

8. Päättele, että kiintopisteiteraatio $x_{n+1} = e^{-x_n}$, $n = 0, 1, \ldots$ suppenee jokaisella $x_0 \in \mathbb{R}$. Määritä kiintopiste iteroimalla ensin kolme kertaa alkuarvauksesta $x_0 = 0$ ja kiihdyttämällä sen jälkeen Newtonin menetelmällä.

9. Näytä, että differentiaaliyhtälön $y'' - y' = 0$ yleinen ($\mathbb{R}$:ssä kahdesti derivoituva) ratkaisu on $y(x) = C_1 e^x + C_2$, $C_1, C_2 \in \mathbb{R}$.

10. a) Näytä, että funktiolle

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & \text{kun } x \neq 0 \\ 0, & \text{kun } x = 0 \end{cases}$$

   pätee $f^{(k)}(0) = 0 \ \forall k \in \mathbb{N}$. Hahmottele funktion kuvaaja.

   b) Halutaan määritellä funktio f ehdoilla

   (i) $f(x) > 0$ kun $x \in (0, 1)$
   (ii) $f(x) = 0$ kun $x \leq 0$ tai $x \geq 1$
   (iii) f on koko $\mathbb{R}$:ssä sileä, eli mielivaltaisen monta kertaa derivoituva

   Näytä, että eräs vaatimukset täyttävä funktio on

$$f(x) = \begin{cases} e^{-\frac{1}{x(1-x)}}, & \text{kun } x \in (0, 1) \\ 0, & \text{muulloin} \end{cases}$$

   Hahmottele tämän kuvaaja.

   c) Näytä, että funktio

$$f(x) = \begin{cases} 0, & \text{kun } x/\pi \in \mathbb{Z} \\ e^{-\cot^2 x}, & \text{muulloin} \end{cases}$$

   on koko $\mathbb{R}$:ssä mielivaltaisen monta kertaa derivoituva (sileä) funktio. Hahmottele f:n kuvaaja välillä $[0, \pi]$.

11. (*) Millä k:n arvoilla käyrällä $y = \ln|x|$ ja suoralla $y = k(x - 1) + 1$ on täsmälleen yksi yhteinen piste?

12. (*) Millä a:n arvoilla yhtälöllä $e^x = x + ax^2$ on reaalisia ratkaisuja a) ei yhtään, b) vain yksi, c) täsmälleen kaksi, d) kolme tai enemmän? Määritä ratkaisut numeerisin keinoin tapauksissa $a = 1$ ja $a = 2$.

13. (*) a) Näytä, että

$$1 + x \le e^x \le 1 + x + (e-2)\,x^2 \quad \forall\, x \in [0,1].$$

b) Olkoon $a \ge 1$ ja $n \in \mathbb{N},\ n \ge \ln a$. Näytä, että

$$1 + \frac{\ln a}{n} \le \sqrt[n]{a} \le 1 + \frac{\ln a}{n} + (e-2)\left(\frac{\ln a}{n}\right)^2.$$

Kuinka suuri on approksimaation $\sqrt[n]{a} \approx 1 + \ln a/n$ virhe todellisuudessa, kun $a = \pi$ ja $n = 100$?

14. (*) Näytä, että differentiaaliyhtälön $y' = 2xy$ yleinen ratkaisu $\mathbb{R}$:ssä on $y(x) = Ce^{x^2},\ C \in \mathbb{R}$.

15. (*) *Hermiten polynomi* $H_n(x)$ määritellään derivoimiskaavalla

$$D^n e^{-x^2} = H_n(x) e^{-x^2}, \quad n \in \mathbb{N} \cup \{0\}.$$

Näytä, että pätee palautuskaava

$$H_{n+1}(x) = -2xH_n(x) - 2nH_{n-1}(x), \quad n \in \mathbb{N}.$$

16. (*) Tarkastellaan funktioita

$$f(x) = \left(1 + \frac{1}{x}\right)^x,\ x \in (0,\infty), \qquad g(x) = \left(1 - \frac{1}{x}\right)^x,\ x \in [1,\infty).$$

a) Näytä, että f ja g ovat määrittelyväleillään aidosti kasvavia.
b) Näytä, että

$$\lim_{x\to\infty}\left(1 + \frac{1}{x}\right)^x = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e,$$
$$\lim_{x\to\infty}\left(1 - \frac{1}{x}\right)^x = \lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^n = \frac{1}{e}.$$

c) Näytä, että $\lim\limits_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = e^x\ \forall x \in \mathbb{R}$.

17. (*) (Matemaatikon mökki) Matemaatikon mökki on koordinaatistossa, jonka origo on Helsingissä, positiivinen x-akseli osoittaa itään ja pituusyksikkö = 100 km. Mökki on erään tien varressa kohdassa, jossa tie on itä-länsisuuntainen. Määritä mökin sijainti, kun tiedetään, että tien yhtälö on

$$(2x)^y = y^{3x} \quad (x, y > 0)$$

ja lisäksi tiedetään, että tien päätepiste a) on Helsinki, b) ei ole Helsinki.

18. (*) (Pieniä ihmeitä) Olkoon $A \in \mathbb{R}$. Halutaan löytää $\mathbb{R}$:ssä derivoituvat funktiot u ja v, jotka ratkaisevat alkuarvotehtävän

$$\text{(P)} \quad \begin{cases} u' = -v, \;\; v' = u, \quad x \in \mathbb{R}, \\ u(0) = A, \;\; v(0) = 0. \end{cases}$$

a) Totea: Eräs ratkaisu on $u(x) = A\cos x$, $v(x) = A\sin x$.
b) Todista: Jos u ja v ovat mikä tahansa (P):n ratkaisu, niin

$$[u(x)]^2 + [v(x)]^2 \;=\; A^2 \;\; \forall x \in \mathbb{R}.$$

c) Näytä: Jos $A = 0$, niin (P):n ainoa ratkaisu on $u(x) = v(x) = 0$.
d) Päättele: a-kohdan ratkaisu on ainoa (P):n ratkaisu.
d) Näytä: (P):n ratkaisu on myös

$$u(x) = A\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}\,, \qquad v(x) = A\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}\,.$$

e) Päättele: Jokaisella $x \in \mathbb{R}$ pätee

$$\cos x \;=\; \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}\,, \qquad \sin x \;=\; \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}\,.$$

VI.3 Kompleksinen eksponenttifunktio. Hyperboliset funktiot

Peruseksponenttifunktio voidaan laajentaa kompleksimuuttujan funktioksi $\exp : \mathbb{C} \to \mathbb{C}$ niin, että sen keskeiset ominaisuudet säilyvät. Määritelmä kompleksialueella — joka toimii myös reaalimuuttujaan rajoitettuna — on seuraava.

MÄÄRITELMÄ VI.3.1 Peruseksponenttifunktio $\exp(z) = e^z$ on funktio, jolla on ominaisuudet (aksioomat)

(C1) $\exp(z)$ on määritelty $\forall z \in \mathbb{C}$,

(C2) $\exp(x + i0) = e^x \ \forall x \in \mathbb{R}$,

(C3) $\exp(z)$ on derivoituva origossa,

(C4) $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2) \quad \forall z_1, z_2 \in \mathbb{C}$.

Aksiooman C2 mukaan e^z on reaalisen eksponenttifunktion laajennus kompleksitasoon. Aksiooma C3 (vrt. Luku V.8) vastaa edellisen luvun aksioomaa E3'. Kompleksiselle eksponenttifunktille tämä aksiooman muoto on välttämätön, sillä pelkkä jatkuvuusoletus (vastaten reaalisen eksponenttifunktion aksioomaa E3) ei takaisi funktion $\exp(z)$ yksikäsitteisyyttä kompleksifunktiona (ks. Harj.teht. 1). Aksioomilla C1–C4 sen sijaan $\exp(z)$ tulee yksikäsitteisesti määritellyksi. Osoitetaan tässä ainoastaan, että aksioomien mukainen funktio $\exp(z)$ on olemassa, ja annetaan samalla funktiolle koko $\mathbb{C}$:ssä pätevä laskusääntö.

LAUSE VI.3.2 Funktio

$$\boxed{e^{x+iy} = e^x(\cos y + i \sin y), \quad z = x + iy \in \mathbb{C}} \tag{1}$$

toteuttaa aksioomat C1–C4.

<u>Todistus</u> (C4) Jos $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, niin määritelmän (1) ja kompleksialgebran laskusääntöjen (ks. Luku III.2) perusteella

$$\begin{aligned} e^{z_1+z_2} &= e^{x_1+x_2}[1\angle(y_1 + y_2)] \\ &= e^{x_1} \cdot e^{x_2} \cdot (1\angle y_1) \cdot (1\angle y_2) \\ &= [e^{x_1} \cdot (1\angle y_1)]\,[e^{x_2} \cdot (1\angle y_2)] \\ &= e^{z_1} \cdot e^{z_2}. \end{aligned}$$

(C3) Käytetään hajotelmia

$$\begin{aligned} e^x &= 1 + x + r_1(x), \\ \cos y &= 1 + r_2(y), \\ \sin y &= y + r_3(y), \end{aligned}$$

missä $\lim_{t\to 0} t^{-1} r_i(t) = 0$ (vrt. Luvut V.4 ja VI.2). Näiden ja säännön (1) perusteella on

$$\begin{aligned} e^z &= [1 + x + r_1(x)]\,[1 + iy + r_2(y) + ir_3(y)] \\ &= 1 + x + iy + r(z) \\ &= 1 + z + r(z), \quad z = x + iy, \end{aligned}$$

missä jäännöstermille pätee $\lim_{z\to 0} z^{-1} r(z) = 0$. Näin ollen

$$\lim_{z\to 0} \frac{e^z - e^0}{z} = \lim_{z\to 0} \frac{e^z - 1}{z} = \lim_{z\to 0} \left(1 + \frac{r(z)}{z}\right) = 1.$$

Siis e^z on origossa derivoituva ja derivaatan arvo $= 1$.

(C2) ja (C1) Ilmeisiä. $\square$

Funktio e^z on derivoituva, ei ainoastaan origossa (aksiooma C3), vaan koko kompleksitasossa, sillä ominaisuudet C3–C4 yhdistämällä seuraa

$$\lim_{\Delta z\to 0} \frac{e^{z+\Delta z} - e^z}{\Delta z} = e^z \lim_{\Delta z\to 0} \frac{e^{\Delta z} - 1}{\Delta z} = e^z, \quad z \in \mathbb{C}.$$

Derivoimissääntö on siis sama kuin reaaliselle (perus)eksponenttifunktiolle:

$$\boxed{\mathrm{D}\, e^z = e^z.} \tag{2}$$

Tämän mukaan kompleksifunktio e^z on koko kompleksitasossa analyyttinen eli kokonainen funktio (vrt. Luku V.8). Tähän ja aksioomaan C2 viitaten sanotaankin, että e^z on reaalisen eksponenttifunktion *analyyttinen jatko* koko kompleksitasoon.

Derivoimissäännöstä (2) (joka siis seuraa suoraan aksioomista C3–C4) voidaan myös päätellä, että aksioomien C1–C4 mukainen funktio on yksikäsitteinen. Nimittäin jos funktio $E(z)$ toteuttaa nämä aksioomat, niin on $DE(z) = E(z)$, jolloin derivoimalla funktiota $u(z) = E(z)e^{-z}$ todetaan, että $Du(z) = 0$ (vrt. Lauseen VI.2.1 todistus). Koska aksiooman C2 mukaan on $u(0) = 1$, on pääteltävissä (sivuutetaan yksityiskohdat), että $u(z) = 1 \;\Leftrightarrow\; E(z) = e^z \;\forall z \in \mathbb{C}$.

Määritelmästä (1) nähdään, että

$$|e^z| = e^x > 0 \quad \forall\ z = x + iy \in \mathbb{C},$$

joten kompleksisellakaan eksponenttifunktiolla ei ole nollakohtia. Kun määritelmässä valitaan $x = \operatorname{Re} z = 0$, saadaan *Eulerin kaava*

$$\boxed{e^{iy} = \cos y + i \sin y, \quad y \in \mathbb{R}.} \tag{3}$$

Eulerin kaavaa (3) käyttäen voidaan trigonometriset funktiot cos ja sin lausua kompleksisen eksponenttifunktion avulla:

$$\boxed{\cos x = \tfrac{1}{2}\,(e^{ix} + e^{-ix}), \quad \sin x = \tfrac{1}{2i}\,(e^{ix} - e^{-ix}), \quad x \in \mathbb{R}.} \tag{4}$$

Tällä perusteella voidaan cos ja sin myös laajentaa kompleksimuuttujan funktioiksi. Määritelmät ovat

$$\boxed{\cos z = \tfrac{1}{2}\,(e^{iz} + e^{-iz})\,, \quad \sin z = \tfrac{1}{2i}\,(e^{iz} - e^{-iz})\,, \quad z \in \mathbb{C}.} \tag{5}$$

Kaikki trigonometrisisille funktioille ominaiset laskusäännöt (vrt. Luku II.5) ulottuvat myös kompleksialueelle. Esimerkiksi säännöt

$$\begin{aligned}
&\cos^2 z + \sin^2 z = 1,\\
&\cos 2z = \cos^2 z - \sin^2 z, \quad \sin 2z = 2 \sin z \cos z,\\
&\mathrm{D} \sin z = \cos z, \quad \mathrm{D} \cos z = -\sin z
\end{aligned}$$

ovat todennettavissa suoraan määritelmistä (5) ja e^z:n derivoimissäännöstä (2). Nähdään myös, että $\sin z$ ja $\cos z$ ovat (e^z:n tavoin) analyyttisiä koko kompleksitasossa eli kokonaisia funktioita.

Esimerkki 1 Etsi yhtälön $\sin z = 2$ kaikki ratkaisut kompleksitasosta.

<u>Ratkaisu</u> Kun merkitään $t = e^{iz}$, niin määritelmän (5) mukaan

$$\begin{aligned}
\sin z = 2 \ \Leftrightarrow\ \frac{1}{2i}\,(t - t^{-1}) = 2 \ &\Leftrightarrow\ t^2 - 4it - 1 = 0\\
&\Leftrightarrow\ t = (2 \pm \sqrt{3})i.
\end{aligned}$$

Määritelmän (1) ja aksiooman C3 mukaan

$$t = e^{iz} = e^{-y+ix} = e^{-y} e^{ix}, \quad z = x + iy,$$

joten on oltava

$$\begin{cases} e^{-y} = |t| = 2 \pm \sqrt{3}, \\ e^{ix} = i \end{cases} \quad \Leftrightarrow \quad \begin{cases} y = -\ln(2 \pm \sqrt{3}), \\ x = \frac{\pi}{2} + 2k\pi, \ \ k \in \mathbb{Z}. \end{cases}$$

Tässä on $\ln(2 - \sqrt{3}) = -\ln(2 + \sqrt{3})$, joten ratkaisut ovat

$$z = \frac{\pi}{2} + 2k\pi \ \pm \ \ln(\sqrt{3} + 2)i, \quad k \in \mathbb{Z}. \quad \square$$

Hyperboliset funktiot

Hyberboliset funktiot ovat eksponenttifunktion johdannaisia, joilla on samantyyppisiä ominaisuuksia kuin trigonometrisillä funktioilla. *Hyberbolinen kosini*, symboli cosh (cosinus hyperbolicus), ja *hyberbolinen sini*, symboli sinh (sinus hyperbolicus) määritellään (vrt. trigonometristen funktioiden määritelmä (5))

$$\boxed{\cosh z = \tfrac{1}{2}(e^z + e^{-z}), \quad \sinh z = \tfrac{1}{2}(e^z - e^{-z}), \quad z \in \mathbb{C}.} \tag{6}$$

Esimerkiksi seuraavat laskulait ovat määritelmästä todennettavissa (vrt. trigonometristen funktioiden vastaavat):

$$\begin{aligned} &\cosh^2 z - \sinh^2 z = 1, \\ &\cosh 2z = \cosh^2 z + \sinh^2 z, \quad \sinh 2z = 2 \sinh z \cosh z, \\ &\mathrm{D} \sinh z = \cosh z, \quad \mathrm{D} \cosh z = \sinh z. \end{aligned}$$

Reaalisilla muuttujan arvoilla $\cosh x = \frac{1}{2}(e^x + e^{-x})$ on parillinen ja $\sinh x = \frac{1}{2}(e^x - e^{-x})$ on pariton funktio. Edellinen on aidosti kasvava välillä $[0, \infty)$ ja bijektio kuvauksena $\cosh : [0, \infty) \to [1, \infty)$. Jälkimmäinen on aidosti kasvava koko $\mathbb{R}$:ssä ja bijektio kuvauksena $\sinh : \mathbb{R} \to \mathbb{R}$.

Reaalifunktioiden cosh ja sinh käänteisfunktioita (cosh rajoitettu välille $[0,\infty)$ tai $(-\infty,0]$) sanotaan *area-funktioiksi* ja merkitään arcosh, arsinh. Laskusäännöt saadaan ratkaisemalla

$$\frac{1}{2}(e^y + e^{-y}) = x \;\Leftrightarrow\; y = \operatorname{arcosh} x,$$
$$\frac{1}{2}(e^y - e^{-y}) = x \;\Leftrightarrow\; y = \operatorname{arsinh} x.$$

Nämä ovat toisen asteen yhtälöitä tuntemattoman $t = e^y$ suhteen, joten y on ilmaistavissa logaritmien avulla:

$$\boxed{\begin{aligned} y &= \operatorname{arcosh} x = \pm \ln\left(x + \sqrt{x^2-1}\right), \quad x \geq 1, \\ y &= \operatorname{arsinh} x = \ln\left(x + \sqrt{x^2+1}\right), \quad x \in \mathbb{R}. \end{aligned}}$$

Tässä arcosh:n päähaara Arcosh saadaan etumerkillä $+$. Toisen haaran voi kirjoittaa myös muotoon

$$y \;=\; -\ln\left(x + \sqrt{x^2-1}\right) \;=\; \ln\frac{1}{x+\sqrt{x^2-1}} \;=\; \ln\left(x - \sqrt{x^2-1}\right).$$

Hyberbolinen tangentti tanh (tangens hyperbolicus) määritellään (ol. reaalimuuttuja)

$$\boxed{\tanh x \;=\; \frac{\sinh x}{\cosh x} \;=\; \frac{e^x - e^{-x}}{e^x + e^{-x}} \;=\; \frac{1-e^{-2x}}{1+e^{-2x}}.}$$

Hyperbolinen tangentti on bijektio: $\tanh : \mathbb{R} \to (-1,1)$. Käänteisfunktio on

$$\boxed{\operatorname{artanh} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right), \quad |x| < 1.}$$

Myöhempää käyttöä varten todettakoon vielä hyperbolisten käänteisfunktioiden derivoimiskaavat

$$\boxed{\begin{aligned} \mathrm{D}\ln(x+\sqrt{x^2+1}\,) &= \frac{1}{\sqrt{x^2+1}}\,, \quad x\in\mathbb{R},\\[2ex] \mathrm{D}\left|\ln(x+\sqrt{x^2-1}\,\right| &= \frac{1}{\sqrt{x^2-1}}\,, \quad |x|>1,\\[2ex] \mathrm{D}\,\frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| &= \frac{1}{1-x^2}\,, \quad |x|\neq 1. \end{aligned}}$$

Kompleksinen logaritmifunktio

Kompleksinen (perus)logaritmifunktio määritellään

$$e^w = z \;\Leftrightarrow\; w = \log z.$$

Kun kirjoitetaan $w = u + iv$ ja otetaan z:lle polaariesitys $z = r(\cos\varphi + i\sin\varphi)$, niin yhtälöstä

$$e^w = z \;\Leftrightarrow\; e^u(\cos v + i\sin v) = r(\cos\varphi + i\sin\varphi)$$

nähdään, että on oltava

$$\begin{cases} e^u = r \;\Leftrightarrow\; u = \ln r = \ln|z|,\\ v = \varphi + 2k\pi, \quad k\in\mathbb{Z}. \end{cases}$$

Kyseessä on siis äärettömän monihaarainen funktio. Logaritmifunktion $\log z$ määritelmäksi sovitaan tämän funktion päähaara, jossa $k = 0$:

$$\boxed{\log z = \ln|z| + i\arg z, \quad 0 \le \arg z < 2\pi.}$$

Näin määritelty funktio ei ole jatkuva positiivisella reaaliakselilla, sillä jos $z_n = x + iy_n$, $x > 0$, niin $z_n \to x$, kun $y_n \to 0$, mutta

$$\begin{aligned} y_n \to 0^+ &\;\Rightarrow\; \log z_n \to \ln x,\\ y_n \to 0^- &\;\Rightarrow\; \log z_n \to \ln x + 2\pi i. \end{aligned}$$

Muissa määrittelyjoukon pisteissä $\log z$ on paitsi jatkuva myös derivoituva (ks. Harj.teht. 12), eli $\log z$ on analyyttinen nk. *aukileikatussa kompleksitasossa*, josta origo ja positiivinen reaaliakseli on poistettu.

Johtuen kompleksisen logaritmifunktion haaraisuudesta eivät reaalialueen laskusäännöt yleisty sellaisenaan. Esimerkiksi tulon logaritmi on em. määritelmän perusteella (vrt. Luku VI.2)

$$\log(z_1 z_2) = \begin{cases} \log z_1 + \log z_2\,, & \text{jos } 0 \le \arg z_1 + \arg z_2 < 2\pi, \\ \log z_1 + \log z_2 - 2\pi i, & \text{muulloin.} \end{cases}$$

Esimerkki 2 Funktion log määritelmän mukaisesti

$$\begin{aligned} \log(-1) &= \log(1\angle\pi) = \ln 1 + \pi i = \pi i, \\ \log(1+i) &= \log\left(\sqrt{2}\angle\pi/4\right) = \frac{1}{2}\ln 2 + \frac{\pi}{4}\,i, \\ \log(-1-i) &= \log\left(\sqrt{2}\angle 5\pi/4\right) = \frac{1}{2}\ln 2 + \frac{5\pi}{4}\,i. \quad \square \end{aligned}$$

Kompleksisen logaritmifunktion avulla voidaan määritellä (vrt. edellisen luvun kaava (9))

$$\boxed{z^y = e^{y\log z}, \quad z, y \in \mathbb{C},\ z \neq 0.}$$

Esimerkki 3 Edellisen esimerkin perusteella

$$(-1)^\pi = e^{\pi\log(-1)} = e^{i\pi^2} = \cos\pi^2 + i\sin\pi^2. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Määritellään funktion e^x laajennus kompleksitasoon kaavalla

$$\exp(z) = \exp(x+iy) = e^x(\cos ay + i\sin ay),$$

missä $a \in \mathbb{R},\ a \neq 1$. Mitkä kompleksisen eksponenttifunktion aksioomista ovat voimassa tälle funktiolle?

2. Tarkista seuraavien kaavojen pätevyys, kun $z \in \mathbb{C}$:

a) $\sin^2 z + \cos^2 z = 1, \quad \sin 2z = 2\sin z\cos z, \quad \cos 2z = 2\cos^2 z - 1$

b) $\mathrm{D}\sin z = \cos z, \quad \mathrm{D}\cos z = -\sin z$

3. a) Lausu $\sinh\frac{x}{2}$ ja $\cosh\frac{x}{2}$ $\cosh x$:n avulla.
b) Millaisilla luvuilla n pätee $(\cosh x + \sinh x)^n = \cosh nx + \sinh nx$?

4. a) Määritellään $\coth x = 1/\tanh x$ ($x \in \mathbb{R},\ x \neq 0$). Mikä on funktion $f(x) = 4\tanh x + \coth x$ arvojoukko?
   b) Missä pisteessä funktio $f(x) = \dfrac{\sinh x}{1 - a\cosh x}$, $\quad x \in \mathbb{R}$ saavuttaa suurimman arvonsa, kun $a \in \mathbb{R},\ a > 1$?

5. Näytä, että yhtälöllä $\cos x \cosh x + 1 = 0$ on äärettömän monta reaalista ratkaisua.

6. a) Näytä, että funktioilla Arcosh ja arsinh on asymptoottina funktio $\ln(2x) = \ln x + \ln 2$, kun $x \to \infty$.
   b) Laske funktioiden $\operatorname{arsinh} x$ ja $\operatorname{Arcosh} x$ derivaatat sekä implisiittisesti derivoimalla että suoraan ko. funktioiden lausekkeista.

7. Sievennä:

   a) $\operatorname{Arcosh}(\cosh x)$ b) $\tanh(\operatorname{Arcosh} x)$ c) $\operatorname{artanh} \dfrac{x}{\sqrt{x^2+1}}$

8. Laske seuraavat kompleksiluvut perusmuodossa $x + iy$ (tarkat arvot!):
   a) e^z, $\sinh z$, $\cosh z$ ja $\tanh z$, kun $z = 2 + 3i$
   b) e^z, $\sin z$, $\cos z$ ja $\tan z$, kun $z = -1 - i$
   c) e^z, $\sin z$, $\cos z$ ja $\tan z$, kun $z = 3 - 2i$
   d) $\cosh(\operatorname{arsinh} \frac{4}{3})$, $\tanh(\operatorname{arsinh} \pi)$, $\operatorname{Arcosh} \sqrt{2}$, $\operatorname{artanh} 1$
   e) e^i, i^e, $\log i$, 5^{-i}, $i^{-\pi}$, i^i

9. Tietokoneohjelma laskee luvun $(-\pi)^\pi$ numeroarvoksi $-32.9139 - 15.6897i$. Tarkista lasku!

10. Määritä (tarkasti perusmuodossa $x + iy$) seuraavien yhtälöiden kaikki ratkaisut kompleksitasossa:
    a) $e^z = e$, b) $\cos z = -2$, c) $\cosh z = -1$, d) $\sin z = i$.

11. Määritä seuraaville funktioille suurin kompleksitason avoin osajoukko G, jossa funktio on analyyttinen:

    a) $f(z) = \dfrac{1}{e^z + 1}$ b) $f(z) = \dfrac{z}{\cos z}$ c) $f(z) = \dfrac{1}{z\cosh z + 2z}$

12. (*) Johda suoraan derivaatan määritelmästä derivoimssääntö

$$\mathrm{D}\log z = \frac{1}{z}, \quad z = r\angle\varphi,\ r > 0,\ 0 < \varphi < 2\pi.$$

    Vihje: Kirjoita $\Delta z = \Delta r \angle(\varphi + \psi)$.

VI.4 Eksponenttifunktion sovellusesimerkkejä

Fysiikassa esiintyy koko joukko nk. eksponentiaalisia ilmiöitä, joissa fysikaalisen suureen muuttumista paikan (x) tai ajan (t) funktiona kuvaa eksponenttifunktio. Seuraavassa viisi esimerkkiä.

Esimerkki 1 : **Radioaktiivinen hajoaminen**. Olkoon $A(t)$ tiettyä lajia olevien radioaktiivisten ytimien lukumäärä hetkellä t. Jos merkitään

$$A(t) = A_0 E(t), \quad t \geq 0,$$

missä $A_0 = A(0)$, niin $E(0) = 1$. Jos oletetaan, että sama laki on sovellettavissa jokaisella ajan hetkellä, on oltava

$$A(t_1 + t_2) = A_0 E(t_1 + t_2) = A(t_1)E(t_2) = A_0 E(t_1)E(t_2), \quad t_1, t_2 \geq 0.$$

Siis funktiolla $E(t)$ on ominaisuudet

$$E(0) = 1, \quad E(t_1 + t_2) = E(t_1)E(t_2), \quad t_1, t_2 \geq 0.$$

Päätellään, että E on eksponenttifunktio ja $A(t)$ siis esitettävissä muodossa

$$A(t) = A_0 e^{-at}, \quad t \geq 0,$$

missä a (mittayksikkö 1/s) on ytimille ominainen vakio (dimensiottomana positiivinen). <u>Puoliintumisaika</u> $t_{1/2}$ määritellään ehdosta $A(t_{1/2}) = A_0/2$, jolloin on oltava

$$e^{-a\,t_{1/2}} = \frac{1}{2} \quad \Leftrightarrow \quad a\,t_{1/2} = \ln 2 \quad \Leftrightarrow \quad t_{1/2} = \frac{\ln 2}{a}. \quad \square$$

Esimerkki 2 : **Säteilyn vaimeneminen**. Olkoon $I(x)$ radioaktiivisen säteilyn intensiteetti vaimentavassa (homogeenisessa) väliaineessa kuljetun matkan x funktiona. Jos merkitään

$$I(x) = I_0 E(x), \quad x \geq 0,$$

missä $I_0 = I(0)$, niin päätellään samoin kuin edellisessä esimerkissä, että E on eksponenttifunktio, eli

$$I(x) = I_0 e^{-ax}, \quad x \geq 0,$$

missä a on väliaineelle ominainen vaimennusvakio (mittayksikkö 1/m). Matka jonka kuluessa säteilyn intensiteetti on vaimentunut puoleen alkuintensiteetistä, eli nk. <u>puoliarvomatka</u> on

$$d_{1/2} = \frac{\ln 2}{a}. \quad \square$$

Em. esimerkeissä nojattiin suoraan eksponenttifunktion perusaksioomaan E4. Useammin eksponenttifunktioon päädytään sovellustilanteissa niin, että tarkasteltavan fysikaalisen suureen $y(x)$ (tai $y(t)$) todetaan toteuttavan eksponenttifunktion differentiaaliyhtälön

$$y' = ay, \tag{1}$$

missä a on (dimensiollinen) vakio. Jos oletetaan, että muuttujan (x) fysikaalisesti relevantit arvot ovat (dimensiottomina) ei-negatiivisia ja että tunnetaan arvo $y(0) = y_0$, niin kyseessä on alkuarvotehtävä: On etsittävä funktio $y = y(x)$, joka on jatkuva välillä $[0, \infty)$, derivoituva välillä $(0, \infty)$ ja toteuttaa

$$\begin{cases} y' = ay, & x > 0, \\ y(0) = y_0. \end{cases}$$

Ratkaisu on $y(x) = y_0 e^{ax}$, $x \geq 0$.

Esimerkki 3 : **Sähköpiiri: RC**. Kondensaattorin kapasitanssi on C ja varaus q_0. Hetkellä $t = 0$ (aikayksikkö s) kondensaattoria ryhdytään purkamaan vastuksen R läpi, jolloin varaus hetkellä $t \geq 0$ on $q(t)$. Matemaattinen malli?

Ratkaisu Jos vastuksen läpi kulkeva virta hetkellä t on $i(t)$, niin piiriyhtälöt ovat

$$q(t)/C = Ri(t), \quad i(t) = -q'(t), \quad t > 0.$$

Eliminoimalla $i(t)$ päädytään alkuarvotehtävään

$$\begin{cases} q' = -aq, & t > 0, \\ q(0) = q_0, \end{cases}$$

missä $a = RC$. Ratkaisu on

$$q(t) = q_0 e^{-t/\tau}, \quad t \geq 0,$$

missä $\tau = 1/(RC)$ on piirille ominainen nk. aikavakio (mittayksikkö s). □

Sovelluksissa eksponenttifunktion differentiaaliyhtälöstä (1) esiintyy usein myös variaatio

$$y' = ay + b, \tag{2}$$

missä a ja b ovat vakioita. Olettaen, että $a \neq 0$, tämän eräs ratkaisu on $y_0(x) = -b/a =$ vakio. Kun yleisempää ratkaisua haetaan muodossa $y(x) = y_0(x) + v(x)$, niin todetaan, että funktio v toteuttaa eksponenttifunktion differentiaaliyhtälön (1). Differentiaaliyhtälön (2) yleinen ratkaisu, kun $a \neq 0$, on siis

$$y(x) = -b/a + Ce^{ax}, \quad x \in \mathbb{R}.$$

Sovellustilanteessa vakio C määrätään alkuehdosta.

Esimerkki 4 : **Sähköpiiri: LR**. Induktanssi L ja vastus R on kytketty sarjaan. Hetkellä $t = 0$ s piiri kytketään vakiojännitteeseen E, jolloin piiriin syntyy virta $i(t)$ ($i(0) = 0$). Matemaattinen malli vastuksen yli vaikuttavalle jännitteelle $y(t)$?

Ratkaisu Piiriyhtälöt ovat

$$L\,i'(t) + R\,i(t) = E, \quad R\,i(t) = y(t), \quad t > 0.$$

Eliminoimalla $i(t)$ päädytään differentiaaliyhtälöön (2), missä $a = -R/L$, $b = ER/L$. Alkuarvo on $y(0) = 0$. Alkuarvotehtävän ratkaisu saadaan valitsemalla em. yleisessä ratkaisussa $C = b/a$, jolloin

$$y(t) = (b/a)(e^{at} - 1) = E(1 - e^{-t/\tau}), \quad t \geq 0,$$

missä $\tau = L/R$ (aikavakio, yksikkö s). $\square$

Esimerkki 5 : **Jäähtymislaki**. Hetkellä t (yksikkö s) on kappaleen lämpötila $u(t)$ ja lämpöenergia $U(t) = cu(t)$ ($c =$ lämpökapasiteetti). Miten kappale jäähtyy alkulämpötilasta $u(0) = u_0$, jos oletetaan, että lämpövirta ympäröivään ilmaan on $Q(t) = k[u(t) - u_1]$, missä $u_1 < u_0$ on ulkoilman lämpötila?

Ratkaisu Energian säilymislaki on

$$\begin{aligned} U' = -Q \quad &\Leftrightarrow \quad cu' = -k(u - u_1) \\ &\Leftrightarrow \quad u' = au + b, \quad a = -k/c,\ b = ku_1/c. \end{aligned}$$

Differentiaaliyhtälön yleinen ratkaisu $u(t) = -b/a + Ce^{at} = u_1 + Ce^{at}$ toteuttaa alkuehdon $u(0) = u_0$, kun $C = u_0 - u_1$, joten

$$u(t) = u_1 + (u_0 - u_1)e^{-t/\tau}, \quad t \geq 0,$$

missä $\tau = -a^{-1} = c/k$ on jäähtymisen aikavakio (yksikkö s). $\square$

HARJOITUSTEHTÄVIÄ

1. Plutoniumin Pu^{239} puolintumisaika on 25400 vuotta. Paljonko 1000 kg:sta plutoniumia on jäljellä miljoonan vuoden kuluttua?

2. Positiivisen x akselin suuntaan etenevästä säteilystä pääsee välillä $[0, 3]$ olevan säteilysuojauksen läpi 3% säteilyn intensiteetistä. Suojaus on rakennettu kahdesta materiaalikerroksesta siten, että välillä $[0, 1]$ olevassa materiaalissa 1 on säteilyn vaimennusvakio kolme kertaa niin suuri kuin välillä $[1, 3]$ olevassa materiaalissa 2. Paljonko jälkimmäistä materiaalikerrosta on vahvistettava, jotta säteilystä pääsisi läpi ainoastaan 1% ?

3. Vuotuinen korkoprosentti on 5 ja korko liitetään pääomaan a) jatkuvan koronkoron mukaisesti, b) puolivuosittain. Mikä on eri tavoin saatujen pääomien suhde 50 vuoden kuluttua?

4. Keittokattilan alkulämpötila on 96°C. Kattilan annetaan ensin jäähtyä huoneen lämmössä (20°C), kunnes sen lämpötila on 40°C. Tämän jälkeen kattila sijoitetaan jääkaappiin, jossa lämpötila on 6°C. Jos ensimmäisen jäähdytysvaiheen kesto on 32 min, niin kauanko kattilan on oltava jääkaapissa, jotta se on jäähtynyt lämpötilaan 7°C? Oletetaan, että huoneessa ja jääkaapissa pätee sama eksponentiaalinen jäähtymislaki.

5. Kupillinen kahvia, jonka lämpötila on aluksi 80°C, jäähtyy ulkoilmassa siten, että 5 minuutin kuluttua kahvin lämpötila on 60°C ja 10 minuutin kuluttua 44°C. Mikä on ulkoilman lämpötila?

6. Vaakasuoralla maan pinnalla on kuution muotoinen vesisäiliö, jonka särmän pituus on 10 m. Säiliö on täynnä vettä. Hetkellä $t = 0$ avataan säiliön pohjaventtiili, jolloin säiliö alkaa tyhjentyä nopeudella $kH(t)$, missä $H(t)$ on veden korkeus säiliössä (yksikkö m) hetkellä t (yksikkö h=tunti) ja k on venttiilin asennosta riippuva kerroin (yksikkö m^2/h). Tunnin kuluttua venttiilin avaamisesta havaitaan, että 8% vedestä on virrannut pois. Tällöin venttiiliä avataan lisää, jolloin virtausnopeus hetkellisesti kaksinkertaistuu. Venttiili jätetään tämän jälkeen uuteen asentoonsa. Määritä säiliössä olevan veden määrä $m(t)$ (kuutiometreinä) ajan t funktiona välillä $[0, \infty)$.

7. Kappale on hetkellä $t = 0$ s levossa veden pinnalla ja lähtee vajoamaan noudattaen liikelakia
$$v'(t) = 0.8g - kv(t),$$
missä v on vajoamisnopeus, $g = 10\,\mathrm{m/s}^2$ ja k on vakio. Määritä $v(t)$, $t \geq 0$ (aikayksikkö s), kun tiedetään, että kappale vajoaa hyvin syvässä vedessä (asymptoottisella) nopeudella 3.2 m/s.

8. (*) Edellisessä tehtävässä kappaleen vajoamissyvyys ennen pohjakosketusta toteuttaa
$$\begin{cases} h'(t) = v(t), & t > 0, \\ h(0) = 0. \end{cases}$$
Millä syvyydellä kappaleen vajoamisnopeus on 99.9 % nopeuden asymptoottisesta arvosta?

Luku VII

Yhden muuttujan differentiaalilaskenta

Kun derivaattaa käytetään laskennassa välineenä, puhutaan *differentiaalilaskennasta* (kirjaimellisesti 'pienten erotusten laskennasta', engl. differential calculus). Derivaatan toivat matematiikkaan toisistaan riippumatta englantilainen fyysikko-matemaatikko **Isaac Newton** (1642-1727) ja saksalainen filosofi-matemaatikko **Gottfried Wilhelm Leibniz** (1646-1716) 1600-luvun lopulla. Leibniz on vaikuttanut huomattavasti vielä nykyisinkin käytössä oleviin differentiaalilaskun (myös myöhemmin tarkasteltavan integraalilaskun) merkintöihin.

Koska derivaatta oli alunperinkin fysiikan motivoima käsite (etenkin Newtonin tutkimuksissa), ei derivaatan soveltuvuudessa fysiikkaan ole ihmettelemistä. Pitkälle 1800-luvulle differentiaalilaskennan ja sen pohjalta nousevien matematiikan alojen kehitys olikin voimakkaasti sidoksissa fysiikkaan, ja vielä nykyisinkin on yhteys säilynyt vahvana monissa matematiikan lajeissa.

Tässä luvussa tarkastellaan yhden muuttujan differentiaalilaskennan soveltamista käyräteoriassa, fysiikassa (mm. liikeopissa) ja funktioiden approksimoinnissa. Funktion approksimoinnin keskeinen tulos on Luvussa VII.4 esitettävä ja todistettava *Taylorin lause*. Tässä on kyse linearisoivan approksimaation yleistämisestä *Taylorin polynomeihin* perustuvaksi yleisemmäksi polynomiapproksimaatioksi. Suotuisissa oloissa funktio voidaan esittää myös tarkasti *Taylorin sarjana*. Taylorin sarjat ovat potenssisarjoja — ja potenssisarjojen teorian kauniiksi lopuksi osoittautuukin, että jokainen potenssisarja, jonka suppenemissäde on positiivinen, on itse asiassa sarjan summana määritellyn funktion Taylorin sarja. Viimeisessä osaluvussa esitellään vielä numeerisissa laskentamenetelmissä yleisesti käytettyjen *interpolaatiopolynomien* teoriaa ja käyttötapoja.

VII.1 Differentiaali ja muutosnopeus

Jos f on derivoituva pisteessä x, niin linearisoivan approksimaatioperiaatteen mukaisesti on likimain

$$\Delta f(x) = f(x + \Delta x) - f(x) \approx f'(x)\Delta x,$$

kun $|\Delta x|$ on pieni. Lauseketta $f'(x)\Delta x$ sanotaan f:n *differentiaaliksi* pisteessä x ja merkitään

$$df(x, \Delta x) = f'(x)\Delta x.$$

Differentiaalin avulla voi siis arvioida likimäärin funktion arvon muutoksen Δf, joka vastaa muuttujan pientä muutosta Δx. Jos approksimaation virhettä merkitään

$$r(x, \Delta x) = f(x + \Delta x) - f(x) - f'(x)\Delta x,$$

niin derivaatan määritelmän mukaisesti pätee

$$\lim_{\Delta x \to 0} \frac{r(x, \Delta x)}{\Delta x} = 0.$$

Esimerkki 1 Neliön muotoista mökkiä, joka on ulkomitoiltaan 6 m $\times$ 6 m, kaupataan 36 m^2:n mökkinä. Arvioi mökin todellinen lattiapinta-ala differentiaalin avulla olettaen seinän paksuudeksi 20 cm.

<u>Ratkaisu</u> Jos $f(x) = x^2$, niin lattiapinta-ala on differentiaalin avulla arvioiden

$$A = f(x - \Delta x) \approx f(x) - f'(x)\Delta x = x^2 - 2x\Delta x.$$

Arvoilla $x = 6$ m, $\Delta x = 0.4$ m (huom!) saadaan

$$A \approx 31.2 \text{ m}^2.$$

Approksimaation virhe on tässä tapauksessa tarkasti

$$r(x, \Delta x) = (\Delta x)^2 = 0.16 \text{ m}^2,$$

eli 'todellinen' lattiapinta-ala on $A \approx 31.4$ m^2. $\square$

Jäljempänä Luvussa VII.4 näytetään, että differentiaaliin perustuvan approksimaation virhe on yleisesti suuruusluokkaa $\sim (\Delta x)^2$ silloin kun funktio on riittävän säännöllinen (kuten esimerkissä).

Esimerkki 2 Ideaalikaasun adiabaattisissa (äkillisissä) paineen ja tiheyden vaihteluissa pätee tilanyhtälö

$$p\rho^{-\gamma} = K = \text{vakio},$$

missä γ on kaasulle ominainen vakio ($\gamma > 1$). Jos tiheys muuttuu 2%, niin paljonko muuttuu paine?

<u>Ratkaisu</u> Koska $p = K\rho^\gamma = f(\rho)$, niin

$$\Delta p \approx f'(\rho)\Delta\rho = \gamma K \rho^{\gamma-1}\Delta\rho = \gamma p \rho^{-1}\Delta\rho$$
$$\Rightarrow \frac{\Delta p}{p} \approx \gamma\frac{\Delta\rho}{\rho} = \underline{\underline{2\gamma\,\%}}. \quad \square$$

Muutosnopeus

Derivaatan tavallisin tulkinta fysiikan ym. sovelluksissa on

$$\boxed{\text{Derivaatta} = \text{(hetkellinen) muutosnopeus.}}$$

Jos m on jokin fysikaalinen suure, joka muuttuu ajan (t) (mittayksikkö s) mukana, eli $m = m(t)$, niin hetkellinen muutosnopeus on

$$\lim_{\Delta t\to 0}\frac{m(t+\Delta t)-m(t)}{\Delta t} = \lim_{\Delta t\to 0}\frac{\Delta m}{\Delta t} = m'(t).$$

Jos m:n yksikkö on M, niin muutosnopeuden yksikkö on M/s.

Esimerkki 3 Jos A = kansantalous (yksikkö E), niin talouskasvu on A:n hetkellinen muutosnopeus (yksikkö E/s). Kun sanotaan 'talouskasvu kiihtyy' tai 'talouskasvu hidastuu', tarkoitetaan $A'(t)$:n muutosnopeutta, eli toista derivaattaa $A''(t)$ (yksikkö E/s^2). $\square$

Jos kappale liikkuu siten, että sen sijainti hetkellä t on $x(t)$ (yksiulotteinen liike), niin tunnetusti

$$x'(t) = \underline{\text{nopeus}}, \quad x''(t) = \underline{\text{kiihtyvyys}}.$$

Esimerkki 4 Kaksimetrinen mies kävelee neljän metrin korkeudella olevan katulampun ali hetkellä $t = 0$. Määritä miehen varjon kärjen paikka, nopeus ja kiihtyvyys hetkellä $t \geq 0$, kun mies kävelee vakionopeudella $v_0 = 2$ m/s. (Ol. katu vaakasuora).

<u>Ratkaisu</u> Varjon kärjen paikka määräytyy ehdosta (ks. kuvio alla)

$$\frac{x(t)-v_0 t}{2} = \frac{x(t)}{4},$$

joten

$$x(t) = 2v_0 t,$$
$$x'(t) = 2v_0 = 4\,\text{m/s},$$
$$x''(t) = 0. \quad \square$$

Esimerkki 5 : **Moottori**. Männänvarren, jonka pituus = 2 (mittayksikkö = 10 cm), kampiakseliin kiinnitetty pää liikkuu pitkin yksikköympyrää siten, että napakulma hetkellä t on $\varphi(t) = at$ (a = vakio), ja toinen pää (johon mäntä on kiinnitetty) liukuu pitkin x-akselia. a) Arvioi differentiaalin avulla, paljonko mäntä liikkuu kampiakselin pyöriessä kulmasta $\varphi = 30°$ kulmaan $\varphi + \Delta\varphi = 31°$. b) Mikä on männän nopeuden (vauhdin) maksimiarvo, kun kampiakselin pyörimisnopeus on 3600 kierrosta/min ?

<u>Ratkaisu</u> a) Kampiakselin kiertokulman ollessa φ on männän paikka

$$x(\varphi) = \cos\varphi + \sqrt{4 - \sin^2\varphi} = \cos\varphi + \sqrt{3 + \cos^2\varphi}.$$

Kulman muuttuessa $\Delta\varphi$:n verran on

$$\Delta x(\varphi) \approx x'(\varphi)\Delta\varphi = -\left(\sin\varphi + \frac{\cos\varphi\sin\varphi}{\sqrt{3+\cos^2\varphi}}\right)\Delta\varphi.$$

Arvoilla $\varphi = 30°$, $\Delta\varphi = 1° \mathrel{\widehat{=}} \pi/180$ saadaan

$$|\Delta x| \approx \frac{1}{2}\left(1 + \frac{1}{\sqrt{5}}\right)\frac{\pi}{180} \cdot 10\,\text{cm} \approx \underline{\underline{1.3\ \text{mm}}}.$$

b) Kun kirjoitetaan $x(t) = x(\varphi(t)) = x(at)$, niin

$$x'(t) = a\left(-\sin at - \frac{\cos at \sin at}{\sqrt{3+\cos^2 at}}\right),$$
$$x''(t) = a^2\left(-\cos at + \frac{\sin^2 at - \cos^2 at}{\sqrt{3+\cos^2 at}} - \frac{\cos^2 at\ \sin^2 at}{(3+\cos^2 at)^{3/2}}\right).$$

Nopeuden $x'(t)$ ääriarvokohdissa on oltava

$$x''(t) = 0 \quad \Leftrightarrow \quad \cos at\left(3+\cos^2 at\right)^{3/2} = 3 - 6\cos^2 at - \cos^4 at.$$

Neliöimällä puolittain ja merkitsemällä $u = \cos^2 at$ tämä sievenee polynomiyhtälöksi

$$u^3 + u^2 - 21u + 3 = 0.$$

Välillä $[0,1]$ tällä on yksikäsitteinen ratkaisu

$$u = 0.143987.. \ \Rightarrow\ \cos\varphi(t) = \sqrt{0.143987..} = 0.379456.. \ \Rightarrow\ \varphi(t) \approx 68^\circ.$$

Tätä vastaa vauhdin maksimiarvo

$$|x'(t)| = a\sqrt{1-u}\left(1+\sqrt{\frac{u}{3+u}}\right) \approx 1.12\,a.$$

Annetulla pyörimisnopeudella on $a = 60 \cdot 2\pi/\mathrm{s}$, joten numeroarvoksi saadaan

$$v_{max} \approx 1.12 \cdot 120\pi\,\mathrm{s}^{-1} \cdot 10\,\mathrm{cm} \approx \underline{\underline{42\,\mathrm{m/s}}}. \quad \square$$

Differentiaali ja differentiaaliyhtälöt

Differentiaaleihin perustuva ajattelu on hyvin tavallista silloin, kun luonnonilmiötä (tai muuta ilmiötä) kuvaava matemaattinen malli on differentiaaliyhtälö, ja halutaan johtaa tämä yhtälö. Seuraavissa esimerkeissä päädytään eksponenttifunktion differentiaaliyhtälöön $y' = ay$ fysikaalisesta (tai muusta) lainalaisuudesta muotoa

$$\Delta y(x) \ = \ y(x+\Delta x) - y(x) \ \approx \ ay(x)\Delta x, \quad \text{kun } |\Delta x| \text{ pieni.}$$

Kun tulkitaan tämän tarkoittavan, että

$$\Delta y(x) = ay(x)\Delta x + r(x,\Delta x),$$

missä $\lim_{\Delta x\to 0} r(x,\Delta x)/\Delta x = 0$, niin jakamalla Δx:llä ja antamalla $\Delta x \to 0$ saadaan 'äärettömän pienille' muutoksille pätevä laki

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \frac{dy}{dx} = ay(x).$$

Esimerkki 6 : **Radioaktiivinen hajoaminen**. Radioaktiivisessa aineessa on yksittäisen ytimen todennäköisyys hajota (ja niinmuodoin 'kadota') aikavälillä $[t, t+\Delta t]$ verrannollinen aikavälin pituuteen Δt, kun Δt on pieni. Jos radioaktiivisten (toistaiseksi hajoamattomien) ytimien lukumäärä hetkellä t on $A(t)$, niin saadaan likimääräiseksi hajoamislaiksi

$$\frac{A(t+\Delta t) - A(t)}{A(t)} \;=\; \frac{\Delta A}{A(t)} \;\approx\; -a\Delta t,$$

missä a on ytimille ominainen vakio (dimensiottomana positiivinen, mittayksikkö 1/s). Rajalla $\Delta t \to 0$ saadaan tarkka hajoamislaki (vrt. Esimerkki VI.4: 1)

$$A'(t) = -aA(t) \quad \Rightarrow \quad A(t) = A(0)\, e^{-at}, \quad t \geq 0. \quad \square$$

Esimerkki 7 : **Säteilyn vaimeneminen**. Säteilyn kulkiessa homogeenisessa väliaineessa on yksityisen hiukkasen todennäköisyys törmätä väliaineen atomiytimeen verrannollinen kuljettuun matkaan Δx, kun $|\Delta x|$ on pieni. Jos säteilyn intensiteetti on $I(x)$ (hiukkasta/m^2/s), niin säteily vaimenee tällöin törmäysten johdosta määrällä

$$I(x+\Delta x) - I(x) \;=\; \Delta I \;\approx\; -aI(x)\Delta x,$$

missä a on aineelle ominainen vakio (dimensiottomana positiivinen, mittayksikkö 1/m). Jos tässä oletetaan, että approksimaation virhe on $r(x, \Delta x)$, missä $\lim_{\Delta x \to 0} r(x, \Delta x)/\Delta x = 0$, niin jakamalla Δx:llä ja antamalla $\Delta x \to 0$ saadaan tarkka säteilyn vaimenemislaki (vrt. Esimerkki VI.4: 2)

$$I'(x) = -aI(x) \quad \Rightarrow \quad I(x) = I(0)\, e^{-ax}, \quad x \geq 0. \quad \square$$

Esimerkki 8 : **Koronkorko**. Pääoman kasvaessa jatkuvaa (koron)korkoa noudattaa pääoman määrä $A(t)$ lakia

$$A(t+\Delta t) - A(t) \;=\; \Delta A \;\approx\; aA(t)\Delta t \quad (\Delta t \text{ pieni}),$$

missä $a = \frac{p}{100T}$, korkoprosentin ollessa p ajassa T (esim. $T = 1$ vuosi). Rajalla $\Delta t \to 0$ päädytään pääoman kasvulakiin

$$A'(t) = aA(t) \quad \Rightarrow \quad A(t) = A(0)\, e^{at}, \quad t \geq 0. \quad \square$$

Esimerkki 9 : **Ilmanpaine**. Määritä ilmanpaine korkeudella x maan pinnasta. Oletetaan vakiolämpötila.

Ratkaisu Jos ilman tiheys korkeudella x on $\rho(x)$, niin voimatasapainon perusteella saadaan paineen muutoksen ja tiheyden välille yhteys

$$\Delta p = p(x + \Delta x) - p(x) = -g\rho(x)\Delta x + r(x, \Delta x),$$

missä $g \approx 10$ m/s^2 on maan vetovoiman kiihtyvyys ja $r(x, \Delta x)/\Delta x \to 0$, kun $\Delta x \to 0$.

Kaasun tilanyhtälö vakiolämpötilassa on

$$p/\rho = K = \text{vakio}.$$

Sijoittamalla tästä $\rho(x) = p(x)/K$ tasapainoyhtälöön ja antamalla $\Delta x \to 0$ päädytään paineen vähenemislakiin

$$p' = -p/a \quad \Rightarrow \quad p(x) = p(0)e^{-x/a},$$

missä

$$a = \frac{K}{g} = \frac{p(0)}{\rho(0)g} \approx \frac{10^5 \text{ N/m}^2}{1 \text{ kg/m}^3 \cdot 10 \text{ m/s}^2} = 10 \text{ km}.$$

Paineen puoliarvokorkeus on $h_{1/2} = (\ln 2)a \approx 6.9$ km. $\square$

HARJOITUSTEHTÄVIÄ

1. Laske $df(x, \Delta x) = f'(x)\Delta x$ ja $f(x + \Delta x) - f(x)$ kuudella desimaalilla, kun $x = 1$, $\Delta x = 0.02$ ja a) $f(x) = x^2 + 200x + 700$, b) $f(x) = x^{-2} - x^{-1}$, c) $f(x) = e^x$, d) $f(x) = \ln(1 + x)$, e) $f(x) = \tan \frac{\pi x}{2}$, f) $f(x) = x^x$.

2. Arvioi differentiaalin avulla, paljonko kuution a) tilavuus kasvaa, kun särmä pitenee $p\%$, b) pinta-ala pienenee, kun tilavuus pienenee $p\%$.

3. Lentokone lentää suoraan maassa olevan katsojan yli 10 km:n korkeudella. Lentokoneen näkyessä vaakatasoon nähden kulmassa 60° havaitaan, että kulma muuttuu 4.3° viidessä sekunnissa. Arvioi koneen nopeus käyttäen differentiaalia.

4. Hiihtäjä, jonka paino on 80 kg, laskee mäkeä, jonka kaltevuus on vakio. Hiihtäjän nopeus v nousee arvoon 50 km/h, jolloin painovoima, kitka ja ilmanvastus saavuttavat tasapainon. Painovoima ja kitkavoima ovat verrannollisia hiihtäjän massaan m ja ilmanvastusvoima on $F_v = \text{vakio} \times Av^\gamma$, missä A on hiihtäjän efektiivinen poikkipinta-ala ja $\gamma \geq 1$ on vakio.
Oletetaan lisäksi, että $m = \text{vakio} \times L^{2.4}$ ja $A = \text{vakio} \times L^{1.2}$, missä L on hiihtäjän vatsanympärys. Arvioi tapauksissa a) $\gamma = 1$, b) $\gamma = 2$, kuinka paljon prosentteina hiihtäjän laskuvauhti kasvaa tai pienenee hänen lihottuaan 2 kg. Käytä differentiaalia! (Muut olosuhteet, kuten keli, oletetaan vakioksi.)

5. a) Valmistettaessa x kpl jääkaappeja on tuotantokustannus
$$K(x) = 8000 + 200x + \max\{0,\, 200x - 0.5x^2\}.$$
Arvioi differentiaalin avulla yhden jääkaapin nk. marginaalinen tuotantokustannus, eli luku $K(x+1) - K(x)$, kun $x = 200$.
b) Erään valtion verotuksessa on marginaalinen (eli pienen lisätulon) veroprosentti vuositulon x (yksikkö 10^5 euroa) funktiona
$$p(x) = 20 + 10\,\min\{x + 3x^2,\, 4\}.$$
Millä vuositulon arvolla marginaalinen veroprosentti = 50 ? Entä millä vuositulon arvolla vero on puolet vuositulosta?

6. a) Kuution pinta-ala kasvaa 50 cm^2/s. Mikä on tilavuuden kasvunopeus, kun särmän pituus = 20 cm?
b) Vesisäiliön tilavuus on 400 l. Säiliöstä lasketaan vettä niin, että veden määrä (yksikkö l) säiliössä hetkellä t (yksikkö min) on
$$V(t) = (20 - t)^2, \quad 0 \leq t \leq 20.$$
Mikä on veden hetkellinen virtausnopeus säiliöstä silloin, kun säiliö on vettä puolillaan?

7. Piste P liikkuu pitkin x-akselia negatiiviseen suuntaan vakionopeudella v_1 ja piste Q liikkuu y-akselia pitkin negatiiviseen suuntaan vakionopeudella v_2. Hetkellä $t = 0$ on $P = (a, 0)$ ja $Q = (0, b)$, missä $a, b > 0$. Jos $s(t)$ janan PQ pituus hetkellä t, niin millä hetkellä on $s'(t) = 0$, ja mikä on s:n pienin arvo?

8. (IKEA) Kaupan pysäköintialueelle tulee aikavälillä $[0, T]$ $Q_{in}(t)$ autoa/h ja alueelta lähtee $Q_{out}(t)$ autoa/h (hetkellisiä arvoja). Johda differentiaaliyhtälö (= autojen säilymislaki!) pysäköintialueella olevien autojen määrälle $m(t)$ hetkellä t, ja ratkaise $m(t)$ aikavälillä $[0, T]$, kun $T = 6$h, pysäköintialueella on 120 autoa hetkellä $t = 0$ ja

$$Q_{in}(t) = 200(1 - t/T), \quad Q_{out}(t) = 270(t/T)^2.$$

Jos halutaan vain selvittää, millä ajan hetkellä pysäköintialueella on eniten autoja, niin miten tämä saadaan yksinkertaisimmin selville?

9. (*) (Lusi) Moottoritie kapenee kaksikaistaiseksi tieksi L:ssä. Moottoritietä pitkin tulee L:ään $Q(t)$ autoa/h (hetkellinen arvo), ja L:stä alkavalle kapeammalle tielle mahtuu ajamaan enintään $Q_0 = 1800$ autoa/h. Olkoon $m(t)$ = L:ään ruuhkautuneiden autojen määrä hetkellä t (aikayksikkö =h). Muodosta ruuhkautumisen matemaattinen malli ja ratkaise sen avulla $m(t)$ aikavälillä $t \in [0, 20]$, kun tiedetään, että $m(0) = 0$ ja

$$Q(t) = 540 + \max\{0,\ 1440t - 180t^2\}.$$

10. (*) (Puimakone) Hihnapyörän keskipiste on origossa ja säde on $R = 0.3$ (yksikkö m). Hihnapyörää pyörittää myötäpäivään hihna, joka koskettaa pyörää napakulmissa $\varphi \in [0, \pi]$. Hihnaa liikutetaan toisella samansäteisellä, moottorikäyttöisellä hihnapyörällä, joka sijaitsee negatiivisella y-akselilla.

Olkoon $F(\varphi)$ hihnaa jännittävä voima napakulmassa φ, $\varphi \in [0, \pi]$ ja olkoon μ hihnan ja hihnapyörän välinen kitkakerroin. Tarkastelemalla välillä $[\varphi - \Delta\varphi/2, \varphi + \Delta\varphi/2]$ olevan hihnan palan voimatasapainoa johda differentiaaliyhtälö

$$F'(\varphi) = -\mu F(\varphi), \quad \varphi \in (0, \pi)$$

ja laske hihnaa jännittävät voimat F_1 ja F_2 hihnapyörien välisissä hihnan osissa (ks. kuva), kun $\mu = 0.2$ ja hihnapyörää vääntävä momentti on $M = (F_1 - F_2)R = 1200$ Nm.

VII.2 Käyrän tangentti ja normaali

Derivaatan tavanomainen geometrinen tulkinta euklidisessa tasossa on:

$$\boxed{\text{Derivaatta} = \text{tangentin kulmakerroin.}}$$

Käyrän $y = f(x)$ sivuaja eli *tangentti* pisteessä $(c, f(c))$ on pisteiden $(c, f(c))$ ja $(c + \Delta x, f(c + \Delta x))$ kulkevan suoran eli *sekantin* 'raja-arvo', kun $\Delta x \to 0$. Pisteen $(c, f(c))$ kautta kulkeva suora, joka on tangenttia vastaan kohtisuora, on käyrän $y = f(x)$ *normaali* ko. pisteessä.

Tangentin ja normaalin yhtälöt ovat

$$\begin{aligned} \text{Tangentti}: &\quad y - f(c) = f'(c)(x - c). \\ \text{Normaali}: &\quad x - c = -f'(c)(y - f(c)). \end{aligned}$$

Sanotaan, että kaksi käyrää *leikkaavat kohtisuorasti*, jos leikkauspisteessä tangentit ovat kohtisuorassa toisiaan vastaan.

Esimerkki 1 Suora kulkee pisteen $(4, 0)$ kautta ja leikkaa käyrän $y = x^2$ kohtisuorasti. Suoran yhtälö?

Ratkaisu Suora on käyrän normaali leikkauspisteessä. Jos leikkauspiste on (t, t^2), niin suoran yhtälö on siis

$$x - t = -2t(y - t^2).$$

Tämä kulkee pisteen $(4, 0)$ kautta ehdolla

$$2t^3 + t - 4 = 0 \;\Rightarrow\; t \approx 1.12817390.$$

Suoran yhtälö: $y = k(x - 4)$, $k \approx -0.44319409$. $\square$

Esimerkki 2 Näytä, että *käyräparvet*

$$xy = a, \quad x^2 - y^2 = b,$$

missä $a, b \in \mathbb{R}$, $a \neq 0$, leikkaavat toisensa kohtisuorasti.

Ratkaisu Yhteisissä pisteissä (x, y) on oltava $x \neq 0$, $y \neq 0$, koska $xy = a \neq 0$. Kun merkitään

$$xy = a \;\Leftrightarrow\; y = y(x) \quad (x \neq 0),$$

niin implisiittisesti derivoimalla saadaan

$$\begin{aligned} \frac{d}{dx}[xy(x)] = y(x) + xy'(x) &= \frac{d}{dx}a = 0 \\ \Rightarrow\; y'(x) &= -y(x)/x = -y/x. \end{aligned}$$

Vastaavasti voidaan merkitä

$$x^2 - y^2 = b \;\Leftrightarrow\; y = y(x),$$

jolloin kyse on kaksihaaraisesta implisiittifunktiosta. Molemmille haaroille pätee derivoimissääntö

$$\begin{aligned} \frac{d}{dx}\left(x^2 - [y(x)]^2\right) = 2x - 2y(x)y'(x) &= \frac{d}{dx}b = 0 \\ \Rightarrow\; y'(x) &= x/y(x) = x/y. \end{aligned}$$

Käyrien yhteisissä pisteissä (x, y) tangenttien kulmakertoimien tulo on siis

$$-\frac{y}{x} \cdot \frac{x}{y} = -1,$$

joten tangentit ovat kaikissa leikkauspisteissä kohtisuorat. $\square$

Esimerkki 3 : **Heijastuslaki**. Halutaan määrittää toistaiseksi tuntemattomalla välillä $A \subset \mathbb{R}$ funktio $x \in A \mapsto y(x) \geq 0$ ja vastaava käyränkaari

$$S = \{\, (x,y) \in \mathbb{R}^2 \mid x \in A \;\;\&\;\; y = y(x) \geq 0 \,\}$$

siten, että jokainen pisteestä $(-c, 0)$ lähtenyt ja käyrästä heijastunut valonsäde kulkee pisteen $(c, 0)$ kautta $(c > 0)$. Muotoile ongelma differentiaaliyhtälöksi muotoa

$$F(x, y, y') = 0,$$

ja etsi ratkaisuja *ellipsin* kaarien

$$y = y(x) \;\Leftrightarrow\; \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

joukosta $(a, b \in \mathbb{R}_+)$.

<u>Ratkaisu</u> Valonsäteen heijastuessa pisteessä $P = (x, y)$ on tulokulma sama kuin heijastuskulma käyrän tangenttiin nähden.

Käyrän tangentin suuntainen vektori heijastuspisteessä on (vrt. kuvio)

$$\vec{t} = \vec{i} + y'\vec{j},$$

ja vektorit $\overrightarrow{AP}$ ja $\overrightarrow{PB}$ ovat (kuvio)

$$\overrightarrow{AP} = (x+c)\vec{i} + y\vec{j}, \quad \overrightarrow{PB} = (c-x)\vec{i} - y\vec{j}.$$

Vaadittu heijastusehto on

$$\frac{\overrightarrow{AP} \cdot \vec{t}}{|\overrightarrow{AP}|} = \frac{\overrightarrow{PB} \cdot \vec{t}}{|\overrightarrow{PB}|}$$

$$\Leftrightarrow \frac{(c+x) + yy'}{\sqrt{(c+x)^2 + y^2}} = \frac{(c-x) - yy'}{\sqrt{(c-x)^2 + y^2}}\,.$$

Neliöön korottamalla ja sieventämällä tämä yksinkertaistuu yhtälöksi

$$F(x,y,y') = xy^2(y')^2 - (c^2 - x^2 + y^2)yy' - xy^2 = 0.$$

Kokeillaan, toteutuuko tämä, kun

$$y^2 = b^2 - \frac{b^2}{a^2}x^2 \;\Rightarrow\; y' = -\frac{b^2}{a^2}\,x/y.$$

Näillä sijoituksilla differentiaaliyhtälö pelkistyy yhtälöksi

$$\left(\frac{c^2+b^2}{a^2} - 1\right) b^2 x = 0,$$

joka toteutuu $\forall x$ ehdolla $a^2 = c^2 + b^2$. Tässä $b \in \mathbb{R}_+$ on vapaasti valittavissa, joten ongelman ratkaisuja ovat

$$y = y(x) = b\sqrt{1 - \frac{x^2}{a^2}}, \quad x \in [-a,a] = A, \;\; a = \sqrt{c^2+b^2}, \;\; b \in \mathbb{R}_+.$$

Nämä ovat erimuotoisia ellipsin puolikaaria, joiden *polttopisteinä* ovat $(\pm c, 0)$. Kuvassa on $c = 1$. $\square$

Parametrisen käyrän tangentti

Parametrisen käyrän, eli vektoriarvoisen funktion $t \in A \mapsto \vec{r}(t) \in \mathbb{R}^d$, missä $A \subset \mathbb{R}$ ja $d = 2$ tai $d = 3$ (vrt. Luku IV.4), derivaatta $\vec{r}\,'(t)$ määritellään tavanomaiseen tapaan, eli

$$\vec{r}\,'(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\vec{r}(t+\Delta t) - \vec{r}(t)\right] = \begin{cases} x'(t)\vec{i} + y'(t)\vec{j} & \text{(tasokäyrä)} \\ x'(t)\vec{i} + y'(t)\vec{j} + z'(t)\vec{k} & \text{(avaruuskäyrä)} \end{cases}$$

sellaisissa pisteissä, joissa funktiot $x(t)$, $y(t)$ ja $z(t)$ ovat derivoituvia. Määritelmän mukaan derivaattavektori on pisteiden $P \mathrel{\widehat{=}} \vec{r}(t)$ ja $Q \mathrel{\widehat{=}} \vec{r}(t+\Delta t)$ kulkevan

suoran suuntavektorin raja-arvo, sikäli kuin pisteet P ja Q ovat erillisiä $|\Delta t|$:n ollessa riittävän pieni. Viimeksi mainittu ehto on voimassa ainakin, jos

$$|\vec{r}\,'(t)| \neq 0, \tag{1}$$

sillä linearisoivan approksimaatioperiaatteen mukaisesti pätee

$$|\vec{r}(t+\Delta t) - \vec{r}(t)| \;=\; |\vec{r}\,'(t)||\Delta t| + g(t,\Delta t), \qquad \lim_{\Delta t\to 0} \frac{g(t,\Delta t)}{\Delta t} = 0,$$

jolloin ehdolla (1) on tässä vasen puoli $\neq 0$, kun $\Delta t \in (-\delta,\delta)$ jollakin (riittävän pienellä) $\delta > 0$. Ehdolla (1) voidaan siis vektori $\vec{r}\,'(t) \neq \vec{0}$ tulkita käyrän $S = \{P(t) \mathrel{\widehat{=}} \vec{r}(t) \mid t \in A\}$ *tangenttivektoriksi* pisteessä $P(t) \mathrel{\widehat{=}} \vec{r}(t)$. Vektorin $\vec{r}\,'(t)$ suuntainen yksikkövektori $\vec{\tau}(t)$ on nimeltään *yksikkötangenttivektori*:

$$\boxed{\vec{\tau}(t) = \frac{\vec{r}\,'(t)}{|\vec{r}\,'(t)|} \quad \text{(yksikkötangenttivektori, } |\vec{r}\,'(t)| \neq 0\,).} \tag{2}$$

Huomattakoon, että yksikkötangenttivektori annetussa käyrän pisteessä P on puhtaasti <u>geometrinen</u> käsite: Mahdollisia $\vec{\tau}$:n arvoja (sikäli kuin $\vec{\tau}$ yleensä on määritelty) on vain kaksi. Ehdon (1) ollessa voimassa parametrisointi valitsee näistä toisen kaavan (2) mukaisesti, muuten $\vec{\tau}$ <u>ei</u> <u>riipu</u> <u>parametrisoinnista</u>.

Yksikkötangenttivektori voidaan tulkita selkeämmin geometrisesti, kun merkitään

$$\Delta\vec{r} = \vec{r}(t+\Delta t) - \vec{r}(t).$$

Jos pisteet $P \mathrel{\widehat{=}} \vec{r}(t)$ ja $Q \mathrel{\widehat{=}} \vec{r}(t+\Delta t)$ ovat erillisiä, niin $\Delta\vec{r}$ on pisteiden P ja Q kautta kulkevan käyrän sekantin suuntavektori. Ehdon (1) voimassa ollessa tarkkenee approksimaatio $|\Delta\vec{r}| \approx |\vec{r}\,'(t)||\Delta t|$ rajalla $\Delta t \to 0$, jolloin yksikkötangenttivektorin molemmat arvot saadaan toispuolisina raja-arvoina

$$\vec{\tau}_{\pm} = \lim_{\Delta t\to 0^{\pm}} \frac{\Delta\vec{r}}{|\Delta\vec{r}|}\,.$$

Tutkittaessa käyrän geometriaa parametrisoinnin avulla asetetaan ehto (1) usein lähtökohtaisesti 'mukavuusehtona' koko tarkasteltavalla välillä, jolloin kaava (2) on käytettävissä. Tyypillisesti oletetaan, että $\vec{r}(t)$:n koordinaattifunktiot $x(t)$, $y(t)$, $z(t)$ ovat <u>jatkuvasti</u> <u>derivoituvia</u> suljetulla välillä $A = [a,b]$ ja että ehto (1) on voimassa koko välillä, mukaanlukien päätepisteet, joissa derivaatta $\vec{r}\,'$ tulkitaan toispuoliseksi.

Esimerkki 4 Ruuviviivan $\vec{r}(\varphi) = R\cos\varphi\vec{i} + R\sin\varphi\vec{j} + a\varphi\vec{k}$ $(R > 0,\ a \neq 0)$ yksikkötangenttivektori pisteessä $P(\varphi) = (x(\varphi), y(\varphi), z(\varphi))$ on

$$\vec{\tau}(\varphi) = \pm\frac{-R\sin\varphi\vec{i} + R\cos\varphi\vec{j} + a\vec{k}}{\sqrt{R^2+a^2}} = \pm\frac{1}{\sqrt{R^2+a^2}}(-y\vec{i} + x\vec{j} + a\vec{k}).$$

Tässä on $|\vec{r}\,'(\varphi)| = \sqrt{R^2+a^2} =$ vakio. $\square$

Esimerkki 5 Käyrä $y = x^2$, $z = x^3$ leikkaa tason T kohtisuorasti pisteessä $(1,1,1)$. Tason yhtälö?

Ratkaisu Käyrän eräs parametrisointi on

$$x(t) = t,\ y(t) = t^2,\ z(t) = t^3 \ \Leftrightarrow\ \vec{r}(t) = t\vec{i} + t^2\vec{j} + t^3\vec{k}.$$

Taso T kulkee pisteen $(1,1,1) \mathrel{\widehat{=}} \vec{r}(1)$ kautta ja sen normaalivektori on $\vec{n} = \vec{r}\,'(1) = \vec{i} + 2\vec{j} + 3\vec{k}$, joten tason yhtälö on (vrt. Luku II.7)

$$(x-1) + 2(y-1) + 3(z-1) = 0 \quad\Leftrightarrow\quad x + 2y + 3z - 6 = 0. \quad \square$$

Nopeusvektori

Jos parametrisessa käyrässä $t \mapsto \vec{r}(t)$ parametri t on fysikaalinen aikamuuttuja ja $\vec{r}(t)$ on avaruudessa liikkuvan pisteen (partikkelin yms.) paikka hetkellä t, niin ko. pisteen nopeus(vektori) ja ratanopeus eli vauhti määritellään:

Nopeusvektori:	$\vec{v}(t) = \vec{r}\,'(t) = x'(t)\vec{i} + y'(t)\vec{j} + z'(t)\vec{k}.$
Vauhti:	$v(t) = \lvert\vec{r}\,'(t)\rvert.$

Aikaparametrisoinnissa ehto (1) on siis 'pysähtymiskielto'.

Esimerkki 6 Partikkeli liikkuu pitkin origokeskistä R-säteistä ympyräviivaa siten, että napakulma hetkellä t on $\varphi(t)$. Nopeus ja vauhti hetkellä t?

Ratkaisu Partikkelin paikkavektori hetkellä t on $\vec{r}(t) = R\cos\varphi(t)\vec{i} + R\sin\varphi(t)\vec{j}$, joten

$$\vec{v}(t) = R\varphi'(t)\,[-\sin\varphi(t)\vec{i} + \cos\varphi(t)\vec{j}\,] = \begin{cases} v(t)\vec{\tau}_+(t), & \text{jos } \varphi'(t) \geq 0, \\ v(t)\vec{\tau}_-(t), & \text{jos } \varphi'(t) < 0, \end{cases}$$

missä $v(t) = R|\varphi'(t)|$ on vauhti ja $\vec{\tau}_\pm(t) = \pm[-\sin\varphi(t)\vec{i} + \cos\varphi(t)\vec{j}\,]$ on liikkeen suuntainen yksikkövektori = liikeradan yksikkötangenttivektori. $\square$

Esimerkki 7 : **Sykloidi**. R-säteinen pyörä vierii liukumatta pitkin x-akselia siten, että pyörän keskipisteen liikenopeus on $v_0\vec{i}$, $v_0 =$ vakio. Pyörän ulkokehän piste P on hetkellä $t = 0$ origossa. Mikä on P:n nopeus ajan funktiona?

Ratkaisu

Pisteen P liikerata on sykloidi, jonka aikaparametrisaatio on (vrt. Luku IV.4)

$$x(t) = v_0 t - R\sin\varphi(t),$$
$$y(t) = R - R\cos\varphi(t),$$

missä

$$R\varphi(t) = v_0 t,$$

joten P:n paikkavektori hetkellä t on

$$\vec{r}(t) = \left[v_0 t - R\sin\frac{v_0 t}{R}\right]\vec{i} + R\left[1 - \cos\frac{v_0 t}{R}\right]\vec{j}.$$

Nopeus ja vauhti hetkellä t ovat

$$\vec{v}(t) = \vec{r}\,'(t) = v_0\left[\left(1 - \cos\frac{v_0 t}{R}\right)\vec{i} + \sin\frac{v_0 t}{R}\vec{j}\right],$$
$$v(t) = |\vec{r}\,'(t)| = |v_0|\sqrt{2 - 2\cos(\frac{v_0 t}{R})} = 2|v_0|\left|\sin\frac{v_0 t}{2R}\right|.$$

Vauhdin minimiarvo $v_{\min} = 0$ saavutetaan aina kun piste P koskettaa x-akselia ja maksimiarvo $v_{\max} = 2|v_0|$ aina kun P on korkeudella $2R$. $\square$

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien tasokäyrien tangentin ja normaalin yhtälöt perusmuodossa $ax + by + c = 0$ annetussa pisteessä:
   a) $y = x^3 - 2x^2 + x + 1, \quad (x, y) = (2, 3)$
   b) $y = \ln x, \quad (x, y) = (e, 1)$
   c) $x^2y^3 - x^3y^2 = 12, \quad (x, y) = (-1, 2)$
   d) $x \sin(xy - y^2) = x^2 - 1, \quad (x, y) = (1, 1)$
   e) $x^y = y^x, \quad (x, y) = (2, 4)$

2. a) Määritä a, b, c siten, että käyrät $y = x^2 + ax + b$ ja $y = cx - x^2$ sivuavat toisiaan (eli niillä on yhteinen tangentti) pisteessä $(1, 3)$.

   b) Määritä pisteen $(-4, 23/2)$ kautta kulkevien käyrän $9y = x^2$ normaalien yhtälöt.

   c) Mihin käyrän $y = x^3$ pisteeseen asetettu normaali leikkaa x-akselin pisteessä $(4, 0)$?

   d) Määritä käyrien $y = \sinh x$ ja $y = \cosh x$ pisteseen $x = a$ asetettujen tangenttien ja normaalien leikkauspisteet.

   e) Suora kulkee pisteen $(3, 0)$ kautta ja leikkaa käyrän $y = e^x$ kohtisuorasti. Määritä suoran yhtälö, tarvittaessa numeerisin apukeinoin.

   f) Mikä suora leikkaa kohtisuorasti käyrät $y = e^x$ ja $y = \ln x$? Mikä on käyrien lyhin etäisyys?

3. a) Määritä *logaritmisen spiraalin* $r = e^\varphi$, $\varphi \in \mathbb{R}$ (napakoordinaatit) tangentin ja normaalin yhtälöt pisteessä $(r, \varphi) = (1, 0)$.

   b) Näytä implisiittisellä derivoinnilla, että käyrällä $S: \; x^5 + x^2y^3 + y^5 = 1$ on ainakin yksi vaakasuora (eli x-akselin suuntainen) ja yksi pystysuora (y-akselin suuntainen) tangentti. Missä käyrän pisteissä nämä sijaitsevat?

4. a) Millä ehdolla käyrät $y = e^{ax}$ ja $y = e^{bx}$ leikkaavat kohtisuorasti? Missä kulmassa käyrät leikkaavat, jos $a = 1$ ja $b = 2$?

   b) Laske käyrien $y = \operatorname{Arcsin} x$ ja $y = \operatorname{Arccos} x$ (tangenttien) välinen kulma käyrien leikkauspisteessä.

   c) Näytä, että sikäli kuin käyrät $y = ae^x$ ja $y = \sqrt{b - 2x}$ $(a, b \in \mathbb{R})$ leikkaavat, niin ne leikkaavat kohtisuorasti. Millä ehdolla käyrät leikkaavat?

5. Avaruussuora S on avaruuskäyrän

$$\vec{r}(t) = 7\sqrt{2}\cos t\,\vec{i} + \sqrt{2}\sin t(3\vec{i} - 2\vec{j} + 6\vec{k})$$

   tangentti pisteessä $(4, 2, -6)$. Määritä S:n yhtälö parametrimuodossa.

6. Määritä seuraavien avaruuskäyrien yksikkötangenttivektori annetussa käyrän pisteessä P sekä sen avaruustason yhtälö, jonka käyrä leikkaa kohtisuorasti P:ssä:
a) $\vec{r} = t^3\,\vec{i} + (2t - t^2)\,\vec{j} + (3t - 2t^4)\,\vec{k}, \quad P = (1, 1, 1)$
b) $\vec{r} = e^t\,\vec{i} - \ln(t+1)\,\vec{j} - \cos t\,\vec{k}, \quad P = (1, 0, -1)$
c) $x = 2\sqrt{2}\cos t,\ y = \sqrt{2}\sin t,\ z = 4t, \quad P = (-2, 1, 3\pi)$

7. Näytä, että ruuviviivan $\vec{r}(t) = \cos t\,\vec{i} + \sin t\,\vec{j} + t\,\vec{k},\ t \in \mathbb{R}$ tangenttien ja xy-tason leikkauspisteet muodostavat tasokäyrän

$$x = \cos t + t\sin t, \quad y = \sin t - t\cos t, \quad t \in \mathbb{R}.$$

8. Tasolla liikuu pistemäinen kappale siten, että hetkellä t kappale on pisteessä $P(t) = (x(t), y(t))$, missä $x(t) = \sin f(t)$, $y(t) = 1 - \cos f(t)$, ja edelleen

$$f(t) = \begin{cases} t^2(3-t)^2, & \text{kun } t \in [0,3], \\ 0, & \text{muulloin.} \end{cases}$$

Millä ajan hetkillä kappaleen vauhti (=nopeuden itseisarvo) on suurimmillaan, ja millainen geometrinen jälki kappaleen liikkeestä jää?

9. (*) Määritä (tarvittaessa numeerisin keinoin) suoran S yhtälö tiedoista:
a) S sivuaa käyriä $y = x^2$ ja $y = \ln x$.
b) S leikkaa kohtisuorasti käyrät $y = x^2$ ja $y = \ln x$.

10. (*) Avaruuskäyrä

$$S: \quad \begin{cases} x^3 + 2y^3 = 3x^2yz, \\ y^3 + z^3 = 2xy \end{cases}$$

leikkaa kohtisuorasti avaruustason T pistessä $(1, 1, 1)$. Määritä T:n yhtälö perusmuodossa $ax + by + cz + d = 0$.

11. (*) (Palloaberraatio) Avaruudessa xy-tasoa pitkin negatiivisen y-akselin suuntaan etenevä valonsäde heijastuu pisteessä P pallopeilistä

$$S: \quad x^2 + (y-R)^2 + z^2 = R^2, \quad 0 \le y \le R,$$

jolloin heijastunut säde leikkaa y-akselin pisteessä $y = c$. Laske c:n lauseke P:n x-koordinaatin funktiona. Totea, että pienillä suhteen $|x|/R$ arvoilla on likimain $c(x) \approx R/2$ ja että poikkeama (nk. palloaberraatio) on likimain

$$c(x) - \frac{R}{2} \approx -\frac{3R}{4}\left(\frac{x}{R}\right)^2.$$

12. (*) (Tutka) Origosta lähtevät radioaallot heijastuvat pyörähdyspinnasta, jonka symmetria-akseli on y-kaseli ja profiili xy-tasossa on käyrä $y = y(x)$, $x \in \mathbb{R}$. Profiili on valittu siten, että heijastuneet aallot kulkevat positiivisen y-akselin suuntaan heijastuspisteestä riippumatta. Näytä, että tämä ehto voidaan esittää differentiaaliyhtälönä

$$(xy' - y)^2 = x^2 + y^2.$$

Etsi mahdolliset polynomiratkaisut muotoa $y(x) = ax^2 + b$.

13. (*) (Polttolasi) Lasista on valmistettu linssi, jonka optinen akseli (symmetria-akseli) on y-akseli ja profiili xy-tasossa on

$$A = \{(x, y) \in \mathbb{R}^2 \mid x \in [-a, a] \ \ \& \ \ f(x) \leq y \leq b\},$$

missä f on parillinen funktio ja $f(0) = 0$. Käyrän $y = f(x)$ muoto halutaan sellaiseksi, että linssi kokoaa kaikki negatiivisen y-akselin suuntaan kulkevat, linssin läpäisseet valonsäteet pisteeseen $(0, -c)$. Taittumislaki linssin kaarevalla pinnalla on $k \sin \alpha_1 = \sin \alpha_2$, missä α_1 on säteen tulokulma (lasissa) pinnan normaalin suhteen, α_2 on lähtökulma, ja $k > 1$ on lasin ja ilman välinen taitekerroin.

Näytä, että funktion $y = f(x)$ on toteutettava differentiaaliyhtälö

$$\left[k^2x^2 + (k^2 - 1)(y + c)^2\right](y')^2 - 2x(y + c)y' = x^2.$$

VII.3 Käyrän kaarevuus

Tarkastellaan tason tai avaruuden (parametrisoitua) käyrää

$$S = \{P(t) \in \mathbb{R}^d \mid P(t) \mathrel{\widehat{=}} \vec{r}(t),\ \ t \in [a,b]\},$$

missä

$$\vec{r}(t) = \begin{cases} x(t)\vec{i} + y(t)\vec{j}, & (\mathbb{R}^2) \\ x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}. & (\mathbb{R}^3) \end{cases}$$

Jatkossa oletetaan, että $x(t)$, $y(t)$ ja $z(t)$ ovat kahdesti jatkuvasti derivoituvia välillä $[a,b]$. Merkitään lisäksi $|\vec{r}\,'(t)| = v(t)$ ja asetetaan 'pysähtymiskielto' (vrt. edellinen luku)

$$v(t) = |\vec{r}\,'(t)| > 0 \quad \forall t \in [a,b]. \tag{1}$$

Käyttöön tulevat myös seuraavat derivoimissäännöt vektoriarvoisille funktioille.

Propositio VII.3.1 Jos f on derivoituva pisteessä t ja

$$\vec{u}(t) = u_1(t)\vec{i} + u_2(t)\vec{j} + u_3(t)\vec{k}, \quad \vec{v}(t) = v_1(t)\vec{i} + v_2(t)\vec{j} + v_3(t)\vec{k},$$

missä funktiot u_i ja v_i ovat derivoituvia pisteessä t, niin pätee

$$\boxed{\begin{aligned} \frac{d}{dt}\,[f(t)\vec{u}(t)] &= f'(t)\,\vec{u}(t) + f(t)\,\vec{u}\,'(t), \\ \frac{d}{dt}\,[\vec{u}(t)\cdot\vec{v}(t)] &= \vec{u}\,'(t)\cdot\vec{v}(t) + \vec{u}(t)\cdot\vec{v}\,'(t), \\ \frac{d}{dt}\,[\vec{u}(t)\times\vec{v}(t)] &= \vec{u}\,'(t)\times\vec{v}(t) + \vec{u}(t)\times\vec{v}\,'(t). \end{aligned}}$$

<u>Todistus</u> Kyse on tulon derivoimissäännön (Luku V.3) yleistyksistä. Esimerkkinä johdettakoon säännöistä toinen (muut perustellaan vastaavalla tavalla):

$$\begin{aligned} \frac{d}{dt}\,\vec{u}(t)\cdot\vec{v}(t) &= \frac{d}{dt}\sum_{i=1}^{3} u_i(t)\,v_i(t) = \sum_{i=1}^{3}\frac{d}{dt}\,u_i(t)\,v_i(t) \\ &= \sum_{i=1}^{3}\left[\,u_i'(t)v_i(t) + u_i(t)v_i'(t)\,\right] \\ &= \sum_{i=1}^{3} u_i'(t)\,v_i(t) + \sum_{i=1}^{3} u_i(t)\,v_i'(t) = \vec{u}\,'(t)\cdot\vec{v}(t) + \vec{u}(t)\cdot\vec{v}\,'(t). \quad \square \end{aligned}$$

Kerrattakoon edellisestä luvusta, että kun merkitään

$$\Delta\vec{r} = \vec{r}\,(t + \Delta t) - \vec{r}\,(t) \approx \vec{r}\,'(t)\Delta t,$$

niin ehdon (1) ollessa voimassa voidaan käyrän S yksikkötangenttivektori $\vec{\tau}(t)$ pisteessä $t \in [a, b]$ määrätä toispuolisina raja-arvoina

$$\vec{\tau}_{\pm} = \lim_{\Delta t \to 0^{\pm}} \frac{\Delta\vec{r}}{|\Delta\vec{r}\,|}, \quad t \in [a, b].$$

(Välin päätepisteissä on vain toinen raja-arvoista mahdollinen.) Funktion $t \mapsto \vec{r}\,(t)$ ollessa kahdesti jatkuvasti derivoituva välillä $[a, b]$ voidaan edelleen laskea yksikkötangenttivektorin derivaatta $\vec{\tau}\,'(t)$, kun $t \in [a, b]$ (toispuolinen derivaatta välin päätepisteissä). Jos käyrä S on jana, niin $\vec{\tau}(t)$ on vakio, jolloin $\vec{\tau}\,'(t) = \vec{0}$, $t \in [a, b]$. Yleisemmin itseisarvo $|d\vec{\tau}/dt|$ kertoo, kuinka <u>kaareva</u> käyrä on ko. pisteessä. Kun tangenttivektorin $\vec{\tau}$ muutosta merkitään

$$\Delta\vec{\tau} = \vec{\tau}(t + \Delta t) - \vec{\tau}(t),$$

niin käyrän *kaarevuus* (engl. curvature) määritellään raja-arvona

$$\boxed{\kappa(t) = \frac{1}{R(t)} = \lim_{\Delta t \to 0} \frac{|\Delta\vec{\tau}\,|}{|\Delta\vec{r}\,|} = \frac{1}{v(t)} \left| \frac{d\vec{\tau}}{dt} \right| \quad \text{(kaarevuus).}} \tag{2}$$

Tässä siis $v(t) = |\vec{r}\,'(t)|$, ja $R = 1/\kappa$ on nimeltään *kaarevuussäde* (engl. radius of curvature).

Samoin kuin yksikkötangenttivektori, myös kaarevuus on vain käyrän geometriasta, ei parametrisoinnista riippuva.

Esimerkki 1 Partikkeli liikkuu tasossa pitkin origokeskistä R-säteistä ympyrärataa siten, että napakulma hetkellä t on $\varphi(t)$. Määritä liikeradan kaarevuus pisteessä $P(t) \mathrel{\widehat{=}} \vec{r}\,(t)$, jossa $\varphi'(t) \neq 0$.

Ratkaisu Tässä on (vrt. Esimerkki 6 edellisessä luvussa)

$$\vec{r}\,(t) = R[\cos\varphi(t)\vec{i} + \sin\varphi(t)\vec{j}\,], \quad \vec{\tau}(t) = \pm[-\sin\varphi(t)\vec{i} + \cos\varphi(t)\vec{j}\,],$$

joten $v(t) = |\vec{r}\,'(t)| = R|\varphi'(t)|$ ja $|d\vec{\tau}/dt| = |\varphi'(t)|$. Kaavan (2) mukaan kaarevuus on $\kappa(t) = 1/R =$ vakio, ja kaarevuussäde siis $= R$, kuten odottaa sopikin. $\square$

Vektorin $d\vec{\tau}/dt$ suuntaista yksikkövektoria $\vec{n}$ sanotaan käyrän *päänormaalivektoriksi.* Tämä on todella käyrän normaalivektori, eli tangenttivektoria vastaan kohtisuora. Nimittäin koska $|\vec{\tau}(t)| = 1\ \forall t$, niin Proposition VII.3.1 säännöillä päätellään

$$\vec{\tau}(t)\cdot\vec{\tau}(t) = 1 \quad \Rightarrow \quad \frac{d}{dt}\,\vec{\tau}(t)\cdot\vec{\tau}(t) = 2\,\vec{\tau}(t)\cdot\vec{\tau}\,'(t) = \frac{d}{dt}\,1 = 0.$$

(Yleisemmin on yksikkövektorin derivaatta aina vektoria vastaan kohtisuora.) Kaava (2) huomioiden on siis päädytty vektorimuotoiseen kaarevuuden määritelmään

$$\boxed{\frac{1}{v(t)}\frac{d\vec{\tau}}{dt} = \frac{1}{R}\vec{n}, \qquad \begin{cases} 1/R = \text{kaarevuus}, \\ \vec{n} = \text{päänormaalivektori}. \end{cases}} \tag{3}$$

Esimerkki 1 (jatko) Tässä saadaan kaavasta (3) odotusten mukaisesti

$$\vec{n}(t) = \frac{R}{v(t)}\frac{d\vec{\tau}}{dt} = -\cos\varphi(t)\vec{i} - \sin\varphi(t)\vec{j} = -\vec{r}\,(t)/R. \quad \square$$

Päänormaalivektori $\vec{n}$ on siis käyrän normaalivektoreista (kaksi vaihtoehtoa!) se, joka osoittaa käyrän kaareutumissuuntaan. Kun derivoidaan puolittain yhtälöt $\vec{n}(t)\cdot\vec{n}(t) = 1$ ja $\vec{n}(t)\cdot\vec{\tau}(t) = 0$ ja sovelletaan kavaa (3), niin seuraa

$$\frac{d\vec{n}(t)}{dt}\cdot\vec{n}(t) = 0, \quad \frac{d\vec{n}(t)}{dt}\cdot\vec{\tau}(t) = -\vec{n}(t)\cdot\frac{d\vec{\tau}(t)}{dt} = -\frac{v(t)}{R}\,.$$

Tämän mukaan kaarevuuden määritelmän (3) voi esittää myös muodossa

$$\boxed{\frac{1}{v(t)}\frac{d\vec{n}}{dt} = -\frac{1}{R}\vec{\tau}.} \tag{4}$$

Johdetaan vielä kaarevuudelle lauseke, jonka arvo määräytyy suoraan derivaatoista $\vec{r}\,'$ ja $\vec{r}\,''$. Määritelmän ja Proposition VII.3.1 ensimmäisen säännön mukaan on

$$\frac{1}{R}\vec{n} \;=\; \frac{1}{|\vec{r}\,'|}\frac{d}{dt}\left(\frac{1}{|\vec{r}\,'|}\vec{r}\,'\right) \;=\; \frac{1}{|\vec{r}\,'|^2}\vec{r}\,'' + \frac{d}{dt}\left(\frac{1}{|\vec{r}\,'|}\right)\vec{\tau}.$$

Kertomalla tämä ristiin vektorilla $\vec{\tau}$ ja ottamalla puolittain itseiarvot saadaan

$$\frac{1}{R} \;=\; \frac{|\vec{r}\,'' \times \vec{\tau}\,|}{|\vec{r}\,'|^2}\,.$$

Kun tähän vielä sijoitetaan $\vec{\tau} = \vec{r}\,'/|\vec{r}\,'|$, tulee laskukaavaksi

$$\boxed{\kappa = \frac{1}{R} = \frac{|\vec{r}\,' \times \vec{r}\,''|}{|\vec{r}\,'|^3}\,.} \tag{5}$$

Tulkitaan tämä vielä tasokäyrälle $y = f(x)$. Kun parametrina on $t = x$, niin $\vec{r}(x) = x\vec{i} + f(x)\vec{j}$, $\vec{r}\,' = \vec{i} + f'(x)\vec{j}$, $\vec{r}\,'' = f''(x)\vec{j}$, joten $\vec{r}\,' \times \vec{r}\,'' = f''(x)\vec{k}$. Sijoittamalla tämä kaavaan (5) todetaan, että tasokäyrän kaarevuus pisteessä $(x, f(x))$ on

$$\boxed{\kappa = \frac{1}{R} = \frac{|f''(x)|}{[1 + (f'(x))^2\,]^{3/2}} \quad \text{(tasokäyrä } y = f(x)\text{).}} \tag{6}$$

Ilman itseisarvomerkkejä tätä sanotaan *merkkiseksi* kaarevuudeksi.

Esimerkki 2 Käyrän $S:\ y = x^2$ kaarevuussäde pisteessä (x, x^2) on kaavan (6) mukaan

$$R(x) = \frac{1}{2}\,(1 + 4x^2)^{3/2}. \quad \square$$

Kaarevuuskeskiö. Evoluutta

Kun parametrisen käyrän pisteestä $P(t) \mathrel{\widehat{=}} \vec{r}\,(t)$ kuljetaan kaarevuussäteen R pituinen matka päänormaalivektorin $\vec{n}$ suuntaan, tullaan *kaarevuuskeskiöön.* Tämän paikkavektori on siis

$$\boxed{\vec{r}_0(t) = \vec{r}\,(t) + R(t)\vec{n}(t) \quad \text{(kaarevuuskeskiö).}}$$

Geometrisesti voidaan tulkita niin, että pisteen $P(t) \mathrel{\widehat{=}} \vec{r}\,(t)$ ympäristössä käyrä on likimain ympyräviiva, jonka säde = kaarevuussäde R ja keskipiste = kaarevuuskeskiö. Lisäksi tämä nk. *kaarevuusympyrä* on tasossa, jonka suuntavektoreina ovat $\vec{\tau}$ ja $\vec{n}$. Kaarevuusympyrän 'vieriessä' pitkin käyrää piirtää kaarevuuskeskiö toisen käyrän, jota sanotaan alkuperäisen käyrän *evoluutaksi.*

Esimerkki 3 Määritä käyrän $y = 1/x$, $x > 0$ evoluutta.

Ratkaisu Koska $f'(x) = -1/x^2$, niin käyrän (pää)normaalivektori pisteessä $(x, 1/x)$ on

$$\vec{n} \;=\; (1 + x^{-4})^{-1/2}(x^{-2}\vec{i} + \vec{j}) \;=\; (x^4 + 1)^{-1/2}(\vec{i} + x^2\vec{j}).$$

Kaarevuus ko. pisteessä on

$$\frac{1}{R} = \frac{2x^{-3}}{(1 + x^{-4})^{3/2}} = \frac{2x^3}{(x^4 + 1)^{3/2}}.$$

Käyrän pistettä $(t, 1/t)$ vastaava kaarevuuskeskiö $(x(t), y(t))$ on näin ollen

$$\begin{aligned} x(t) \;&=\; t + \frac{(t^4 + 1)^{3/2}}{2t^3} \cdot (t^4 + 1)^{-1/2} \;=\; \frac{1}{2}(3t + t^{-3}), \\ y(t) \;&=\; t^{-1} + \frac{(t^4 + 1)^{3/2}}{2t^3} \cdot t^2(t^4 + 1)^{-1/2} \;=\; \frac{1}{2}(3t^{-1} + t^3). \end{aligned}$$

Tämä on evoluutan parametriesitys ($t > 0$). Suurilla ja pienillä t:n arvoilla on likimain

$$\begin{cases} t \gg 1 \;\Rightarrow\; y \approx \frac{4}{27}x^3, \\ t \ll 1 \;\Rightarrow\; x \approx \frac{4}{27}y^3. \end{cases}$$

Pisteessä $(x, y) = (2, 2)$ $(t = 1)$ evoluutalla on nk. *kääntymispiste* (engl. turning point). Kääntymispisteessä on $x'(t) = y'(t) = 0$, joten kyseessä on myös 'pysähtymispiste'. $\square$

Kaarevuus fysiikassa: kiihtyvyys

Kun parametrisessa käyrässä on kyse liikkuvan pisteen $P(t) \mathrel{\widehat{=}} \vec{r}(t)$ paikasta ajan funktiona, niin $\vec{v}(t) = \vec{r}\,'(t) =$ nopeusvektori. Tällöin vektoria

$$\vec{a}(t) = \vec{v}\,'(t) = \vec{r}\,''(t) = x''(t)\vec{i} + y''(t)\vec{j} + z''(t)\vec{k}$$

sanotaan <u>kiihtyvyydeksi</u> hetkellä t. Kiihtyvyysvektori voidaan kaartuvalla ratakäyrällä aina esittää muodossa

$$\vec{a}(t) = a_\tau\,\vec{\tau} + a_n\,\vec{n},$$

missä a_τ on liikeradan suuntainen <u>tangentiaalikiihtyvyys</u> ja a_n on päänormaalivektorin suuntainen <u>normaalikiihtyvyys</u>. Tähän tulokseen päädytään, kun kirjoitetaan

$$\vec{v}(t) = v(t)\,\vec{\tau}(t), \quad v(t) = |\vec{v}(t)| = |\vec{r}\,'(t)|$$

ja käytetään Proposition VII.3.1 ensimmäistä sääntöä sekä kaavaa (3):

$$\begin{aligned}\vec{a}(t) &= v'(t)\,\vec{\tau}(t) + v(t)\,\frac{d\vec{\tau}}{dt}\\ &= v'(t)\,\vec{\tau}(t) + \frac{[v(t)]^2}{R}\,\vec{n}(t).\end{aligned}$$

Tämän mukaisesti on siis

$$\boxed{a_\tau(t) = v'(t), \quad a_n(t) = \frac{[v(t)]^2}{R} \quad \text{(tangentiaali- ja normaalikiihtyvyys).}}$$

Kiihtyvyyden kannalta voidaan siis liike ajatella hetkellisesti ympyräliikkeeksi, jossa liikerata yhtyy kaarevuusympyrään.

Esimerkki 4 : **Vapaa putoamisliike**. Maan vetovoimakentässä hetkellä $t = 0$ käynnistyvän vapaan putoamisliikkeen (Newtonin) liikeyhtälö on

$$\vec{a}(t) = \vec{r}\,''(t) = -g\vec{k}, \quad t > 0,$$

missä $g \approx 9.81$ m/s^2. Yhtälö sisältää kolme skalaarista differentiaaliyhtälöä:

$$x''(t) = 0, \quad y''(t) = 0, \quad z''(t) = -g.$$

Näiden mukaisesti on ensinnäkin oltava

$$x'(t) = \alpha, \quad y'(t) = \beta, \quad z'(t) = -gt + \gamma,$$

missä α, β, γ ovat (määräämättömiä) vakioita. Tästä nähdään, että on edelleen oltava (vrt. Luku V.6)

$$x(t) = \alpha\, t + x_0, \quad y(t) = \beta\, t + y_0, \quad z(t) = -\tfrac{1}{2}\, gt^2 + \gamma\, t + z_0,$$

missä x_0, y_0, z_0 ovat jälleen määräämättömiä vakioita. Siis

$$\vec{r}(t) = \vec{r}_0 + \vec{v}_0\, t - \tfrac{1}{2}\, gt^2\, \vec{k}, \quad t > 0, \quad \text{missä}$$
$$\vec{r}_0 = x_0\vec{i} + y_0\vec{j} + z_0\vec{k} = \vec{r}(0^+), \quad \vec{v}_0 = \alpha\vec{i} + \beta\vec{j} + \gamma\vec{k} = \vec{r}\,'(0^+) = \vec{v}(0^+).$$

Jos raja-arvot $\vec{r}(0^+) = \vec{r}_0$ ja $\vec{v}(0^+) = \vec{v}_0$ tunnetaan alkuehtoina, on $\vec{r}(t)$ yksikäsitteisesti määrätty, kun $t \geq 0$. Ratakäyrä on tällöin paraabeli avaruustasossa, jonka suuntavektorit ovat $\vec{v}_0$ ja $\vec{k}$ (vrt. Esimerkki IV.4: 6). $\square$

Esimerkki 5 : **Irtoaminen**. Kappale, johon vaikuttaa painovoima

$$\vec{G} = -mg\vec{j} \quad (m = \text{massa},\ g = \text{maan vetovoiman kiihtyvyys})$$

on levossa origossa ja lähtee siitä kitkattomaan liukuun pitkin käyrää

$$y = -\frac{1}{3}\, x^3, \quad x \geq 0.$$

Määritä kappaleen liikerata muodossa $y = f(x),\ x \geq 0$.

<u>Ratkaisu</u> Rata noudattaa aluksi käyrää $y = -x^3/3$, mutta irtoaa siitä, kun tämän radan mukainen normaalikiihtyvyys ylittää maan vetovoiman kiihtyvyyden normaalin suunnalla. Irtoamisehto on siis

$$\frac{v^2}{R} = -g\vec{j} \cdot \vec{n},$$

missä $\vec{n}$ on (pää)normaalivektori

$$\vec{n} = -\frac{1}{\sqrt{1 + x^4}}(x^2\vec{i} + \vec{j}).$$

Energiaperiaatteen mukaan on

$$\frac{1}{2}m[v(x)]^2 = mg \cdot \frac{1}{3}\, x^3,$$

ja kaarevuus pisteessä x on

$$\frac{1}{R} = \frac{2x}{(1 + x^4)^{3/2}}\,,$$

joten irtoaminen tapahtuu, kun

$$\frac{2x}{(1+x^4)^{3/2}} \cdot \frac{2}{3} g x^3 = \frac{g}{(1+x^4)^{1/2}}$$
$$\Leftrightarrow \; x^4 = 3 \; \Leftrightarrow \; \underline{\underline{x = \sqrt[4]{3} = 1.316074..}}$$

Irtoamisen jälkeen lentorata on paraabeli (vapaa putoamisliike), joten koko liikerata on muotoa

$$y = f(x) = \begin{cases} -x^3/3, & \text{kun } 0 \le x \le \sqrt[4]{3}, \\ Ax^2 + Bx + C, & \text{kun } x > \sqrt[4]{3}. \end{cases}$$

Irtoamiskohdassa on maan vetovoima ainoa kappaleeseen vaikuttava ulkoinen voima. Koska tämä voima on jatkuva, ja myös kappaleeseen vaikuttavan tukivoiman voi olettaa olevan jatkuva irtoamiskohdassa (irtoamisesta eteenpäin tukivoima = 0), niin päätellään, että kiihtyvyysvektori on jatkuva. Tämä merkitsee, että funktio f on kahdesti jatkuvasti derivoituva myös irtoamiskohdassa, eli on oltava

$$f^{(k)}(x_0^+) = f^{(k)}(x_0^-), \quad k = 0, 1, 2 \; , \; x_0 = \sqrt[4]{3}.$$

Näistä ehdoista voidaan ratkaista vakiot A, B, C. Tulos (liikerata) on

$$y = f(x) = \begin{cases} -x^3/3, & x \in [0, \sqrt[4]{3}], \\ -\sqrt[4]{3}\, x^2 + \sqrt{3}\, x - 1/\sqrt[4]{3}, & x \in (\sqrt[4]{3}, \infty). \end{cases}$$

Kuvassa liuku/lentorata on piirretty yhtenäisellä viivalla. Toisen asteen käyrä ennen irtoamiskohtaa ja käyrä $y = -x^3/3$ irtoamiskohdan jälkeen on merkitty katkoviivalla. □

HARJOITUSTEHTÄVIÄ

1. Olkoon $\vec{u}(t)$ ja $\vec{v}(t)$ derivoituvia vektoriarvoisia funktioita. Todista:
   a) $\vec{u}(t) \parallel \vec{v}(t)\ \forall t \quad \Rightarrow \quad \vec{u}\,' \times \vec{v} = \vec{v}\,' \times \vec{u}$
   b) $\vec{u}(t) \perp \vec{v}(t)\ \forall t \quad \Rightarrow \quad \vec{u}\,' \cdot \vec{v} = -\vec{v}\,' \cdot \vec{u}$

2. Määritä seuraavien tasokäyrien karevuusympyrä (säde R sekä kaarevuuskeskiö) annetussa käyrän pisteessä P:
   a) $y = x^2, \quad P = (1,1)$
   b) $y = x^3 - 2x^2, \quad P = (2,0)$
   c) $y = e^x, \quad P = (0,1)$
   d) $\vec{r} = t\cos t\,\vec{i} + t\sin t\,\vec{j}, \quad P = (-\pi, 0)$
   e) $x = t - \sin t,\ y = 1 - \cos t, \quad P = (\pi, 2)$
   f) $2x^2 + 3y^2 = 5, \quad P = (1,-1)$
   g) $x^3 - y^3 + y^2 - 3x + 2 = 0, \quad P = (2,2)$
   h) $r = e^{\varphi}, \quad P = (r,\varphi) = (1,0)$

3. Määritä seuraavien avaruuskäyrien kaarevuusssäde käyrän pisteessä, joka vastaa annettua parametrin arvoa. Määritä myös sen avaruustason yhtälö, jossa kaarevuusympyrä sijaitsee, sekä kaarevuuskeskiö.
   a) $\vec{r} = t\,\vec{i} + t^2\,\vec{j} + t^3\,\vec{k}, \quad t = 1$
   b) $x = \cos t,\ y = \sin t,\ z = (4/\pi)\,t, \quad t = \pi/4$
   c) $x = e^t\cos t,\ y = e^t\sin t,\ z = e^t, \quad t = 0$

4. Sanotaan, että tasokäyrät S_1 ja S_2 *oskuloivat* (suom. suutelevat) pisteessä (x_0, y_0), jos käyrillä on yhteinen kaarevuusympyrä ko. pisteessä. a) Päättele, että jos käyrien yhtälöt ovat $S_1 : y = f(x)$ ja $S_2 : y = g(x)$, niin oskulointiehdot ovat

$$f(x_0) = g(x_0) = y_0\,, \quad f^{(k)}(x_0) = g^{(k)}(x_0),\ \ k = 1, 2.$$

   b) Etsi sellainen toisen asteen polynomikäyrä $S_1 :\ y = ax^2 + bx + c$, joka oskuloi käyrää $S_2 :\ 2x^3 + 6y^3 + xy = 0$ pisteessä $(3,-2)$.

5. Määritä käyrän $y = x^2$ evoluutta. Piirrä kuva!

6. Pistemäinen kappale liikkuu pitkin ruuviviivaa

$$x = \cos\varphi, \quad y = \sin\varphi, \quad z = \varphi, \quad \varphi \in [0, \infty),$$

   siten, että sen ratanopeus (vauhti) on vakio v_0. Määritä kappaleen nopeus $\vec{v}$ ja kiihtyvyys $\vec{a}$ ajan t funktiona ($t \geq 0$), kun kappale on pisteessä $(1,0,0)$ hetkellä $t = 0$. Määritä myös radan kaarevuussäde R ja tarkista, että pätee: $|\vec{a}| = |\vec{v}|^2/R$.

7. Maaston korkeusprofiili tunturimaastossa on

$$h(x, y) = 40 - 0.005xy,$$

missä pituusyksikkö = m. Retkeilijä heittää pisteessä $(x, y) = (0, 0)$ olevasta leiripaikastaan pilaantuneen tomaatin siten, että tomaatin lähtönopeus on

$$\vec{v}_0 = (10\,\mathrm{m/s})(\vec{i} - 2\vec{j} + 2\vec{k}).$$

Missä pisteessä tomaatti törmää maahan ja mikä on tällöin sen vauhti? (Oletetaan $g = 10\mathrm{m/s}^2$, ei ilmanvastusta eikä tuulikorjausta.)

8. (*) Näytä, että sykloidin $S : x = R(t - \sin t),\ y = R(1 - \cos t),\ t \in \mathbb{R}$ evoluutta on toinen, S:n kanssa yhtenevä sykloidi, joka saadaan siirtämällä S vektorin $-\pi R\vec{i} - 2R\vec{j}$ verran. Kuva!

9. (*) Origossa oleva kappale lähtee levosta liukumaan kitkattomasti pitkin käyrää $y = 1 - \cosh x,\ x \geq 0$, painovoiman vaikuttaessa suunnassa $-\vec{j}$. Määritä kappaleen liikerata.

10. (*) Avaruuskäyrän pisteeseen $P(t) \mathrel{\widehat{=}} \vec{r}(t)$ liittyvä *Frenet'n kanta* on vektorisysteemi $\{\vec{\tau}, \vec{n}, \vec{\nu}\}$, missä $\vec{\tau}$ ja $\vec{n}$ ovat yksikkötangentti- ja päänormaalivektorit ko. pisteessä ja $\vec{\nu} = \vec{\tau} \times \vec{n}$ on *sivunormaalivektori*. Avaruuskäyrän *kierevyys* ω (engl. torsion) määritellään tällöin kaavalla

$$\frac{1}{v(t)}\frac{d\vec{\nu}}{dt} = -\omega(t)\vec{n}, \quad v(t) = |\vec{r}\,'(t)|.$$

a) Näytä, että $d\vec{\nu}/dt$ todella on päänormaalivektorin suuntainen.
b) Päättele, että jos avaruuskäyrä on tasokäyrä jollakin avaruustasolla, niin sen kierevyys = 0.
c) Laske ruuviviivan $S :\ x = a\cos t,\ y = a\sin t,\ z = bt$ kierevyys käyrän pisteessä $P(t)$.

11. (*) (Sotaharjoitus 2) Tykinammus laukaistaan origosta lähtönopeudella $\vec{v}_0 = (150\,\mathrm{m/s})(\vec{i} + 2\vec{j} + \vec{k})$. Lentoradalla ammukseen vaikuttaa painovoiman lisäksi tuuli ja nopeuteen verrannollinen vastusvoima siten, että liikeyhtälöt ovat

$$\vec{v}\,' = c\vec{i} - g\vec{k} - k\vec{v}, \quad \vec{r}\,' = \vec{v},$$

missä $\vec{r}(t)$ ja $\vec{v}(t)$ ovat ammuksen paikka- ja nopeusvektorit hetkellä t, $g = 10\ \mathrm{m/s}^2$, $k = 0.01\ \mathrm{s}^{-1}$ ja $c = 0.10\ \mathrm{m/s}^2$. Laske, mihin xy-tason pisteeseen (10 metrin tarkkuus!) ammus putoaa. *Vihje*: Ratkaise liikeyhtälöt erikseen suunnissa $\vec{i}$, $\vec{j}$, $\vec{k}$ (ensin $\vec{v}$, sitten $\vec{r}$). Aloita pystysuunnasta lentoajan selville saamiseksi!

VII.4 Taylorin polynomit ja Taylorin lause

Tässä luvussa tarkastellaan funktioita, jotka ovat annetun pisteen x_0 ympäristössä riittävän sileitä, eli riittävän monta kertaa (jatkuvasti) derivoituvia.

MÄÄRITELMÄ VII.4.1 Funktion $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}$, joka on n kertaa derivoituva pisteessä $x_0 \in \mathcal{D}_f$, *Taylorin polynomi astetta n pisteessä x_0* on

$$\boxed{T_n(x, x_0) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.}$$

Määritelmän mukaiset kolme ensimmäistä Taylorin polynomia ovat

$$\begin{aligned} T_0(x, x_0) &= f(x_0), \\ T_1(x, x_0) &= f(x_0) + f'(x_0)(x - x_0), \\ T_2(x, x_0) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2. \end{aligned}$$

Erityisesti siis $T_1(x, x_0) = f$:n linearisoiva approksimaatio pisteessä x_0.

PROPOSITIO VII.4.2 Funktion f Taylorin polynomi $T_n(x, x_0)$ määräytyy yksikäsitteisesti ehdoista

$$\frac{d^k}{dx^k} T_n(x, x_0)_{|\, x=x_0} = f^{(k)}(x_0), \quad k = 0 \ldots n.$$

<u>Todistus</u> Helposti nähdään, että $T_n(x, x_0)$ toteuttaa mainitut ehdot. Jos jokin toinen polynomi $p(x)$ toteuttaa samat ehdot, eli

$$p^{(k)}(x_0) = f^{(k)}(x_0), \quad k = 0 \ldots n,$$

niin silloin polynomi

$$q(x) = T_n(x, x_0) - p(x) = \sum_{k=0}^{n} a_k x^k$$

toteuttaa

$$q^{(k)}(x_0) = 0, \quad k = 0 \ldots n.$$

Tällöin koska $q^{(n)}(x_0) = n!\, a_n$, seuraa $a_n = 0$, jolloin $q^{(n-1)}(x_0) = (n-1)!\, a_{n-1} \Rightarrow a_{n-1} = 0$, jne. Siis $q = 0$, ja näin ollen em. ehdoista määräytyvä polynomi on yksikäsitteinen. $\square$

Esimerkki 1 Määritä seuraavat Taylorin polynomit pisteessä $x_0 = 0$:

a) $f(x) = \sqrt[3]{1+x}, \quad n = 2$ b) $f(x) = \tan x, \quad n = 5$
c) $f(x) = \ln(1+x), \quad n \in \mathbb{N}$

Ratkaisu a) $f'(x) = \frac{1}{3}(1+x)^{-2/3}, \quad f''(x) = -\frac{2}{9}(1+x)^{-5/3}$

$$\Rightarrow f(0) = 1, \quad f'(0) = \frac{1}{3}, \quad f''(0) = -\frac{2}{9}$$

$$\Rightarrow T_2(x,0) = \underline{\underline{1 + \frac{x}{3} - \frac{1}{9}x^2.}}$$

b) $f'(x) = 1/\cos^2 x, \quad f''(x) = 2\sin x/\cos^3 x,$

$$f'''(x) = 2/\cos^2 x + 6\sin^2 x/\cos^4 x = -4/\cos^2 x + 6/\cos^4 x$$

$$f^{(4)}(x) = -8\sin x/\cos^3 x + 24\sin x/\cos^5 x$$

$$f^{(5)}(x) = -16/\cos^2 x + 120\sin^2 x/\cos^6 x$$

$$\Rightarrow \begin{cases} f(0) = f''(0) = f^{(4)} = 0, \\ f'(0) = 1, \quad f'''(0) = 2, \quad f^{(5)}(0) = 16 \end{cases}$$

$$\Rightarrow T_5(x,0) = \underline{\underline{x + \frac{1}{3}x^3 + \frac{2}{15}x^5.}}$$

c) $f'(x) = (1+x)^{-1}, \quad f''(x) = -(1+x)^{-2}, \quad f'''(x) = 2(1+x)^{-3},$

$$\ldots, \quad f^{(k)} = (-1)^{k-1}(k-1)!(1+x)^{-k}$$

$$\Rightarrow f(0) = 0, \quad f^{(k)}(0) = (-1)^{k-1}(k-1)!, \quad k = 1, 2, \ldots,$$

$$\Rightarrow T_n(x,0) = \underline{\underline{x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots + (-1)^{n-1}\frac{x^n}{n}.}} \quad \square$$

Taylorin polynomin derivaatta on

$$\frac{d}{dx}T_n(x,x_0) = f'(x_0) + f''(x_0)(x-x_0) + \ldots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1}.$$

Derivoinnin tulos = f':n Taylorin polynomi astetta $n-1$ pisteessä x_0, eli lyhyesti: Taylorin polynomin derivaatta = derivaatan Taylorin polynomi (astetta alempi).

Esimerkki 1 (jatko). Esimerkin tuloksista saadaan derivoimalla seuraavat Taylorin polynomit:

a) $f(x) = (1+x)^{-2/3}$: $T_1(x,0) = 1 - \frac{2}{3}x$.
b) $f(x) = 1/\cos^2 x$: $T_4(x,0) = 1 + x^2 + \frac{2}{3}x^4$.
c) $f(x) = 1/(1+x)$: $T_{n-1}(x,0) = 1 - x + \cdots + (-1)^{n-1}x^{n-1}$. □

Taylorin lause

Taylorin polynomeihin perustuu seuraava huomattava approksimaatiolause. Todistus esitetään luvun lopussa.

Lause VII.4.3 (**Taylorin lause**) Jos f on jatkuva välillä $[a,b]$ ja $n+1$ kertaa derivoituva välillä (a,b), niin jokaisella $x_0 \in (a,b)$ ja $x \in [a,b]$, $x \neq x_0$ pätee

$$f(x) = T_n(x,x_0) + R_n(x),$$

missä

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \quad \text{jollakin } \xi \in (x_0,x) \text{ tai } \xi \in (x,x_0).$$

Taylorin lauseen mukaan funktiota, joka on tietyn pisteen x_0 ympäristössä säännöllinen, voi tässä ympäristössä approksimoida polynomilla — nimittäin Taylorin polynomilla — ja approksimaatio on yleisesti ottaen sitä tarkempi, mitä korkeampi on polynomin asteluku, ja mitä lähempänä ollaan pistettä x_0. Tuloksen voi esittää kvalitatiivisesti muodossa:

Sileä funktio $\approx$ polynomi lyhyellä välillä.

Taylorin lauseen tuloksella on perustavaa laatua oleva merkitys lähes kaikessa numeerisessa laskennassa, johon sisältyy funktioiden approksimointia. Virhetermille $R_n(x,x_0)$, eli Taylorin polynomiapproksimaation nk. *jäännöstermille* (engl. remainder), tunnetaan monia muotoja. Lauseen VII.4.3 esittämää sanotaan jäännöstermin *Lagrangen*† muodoksi.

†Italialais-ranskalainen **Joseph Louis** (synt. Giovanni Luigi) **Lagrange** (1736-1813) oli aikansa huomattavimpia matemaatikkoja. Erityisesti differentiaalilaskennan (myös integraalilaskennan) kehittämisessä Lagrangen panos oli merkittävä. Matematiikan ohella Lagrange tutki mekaniikkaa ja saavutti silläkin alalla pysyvän nimen.

Taylorin polynomit, Taylorin lause, ja erityisesti jäljempämä esitettävät *Taylorin sarjat* viittaavat englantilaiseen matemaatikkoon **Brook Taylor**iin (1685-1731). Nimeään kantavaa lausetta ei Taylor todellisuudessa tuntenut.

Esimerkki 2 Soveltamalla derivointisääntöjä

$$\begin{cases} D^k e^x = e^x, \quad k = 0, 1, 2, \ldots \\ D^{2k} \cos x = (-1)^k \cos x, \quad D^{2k+1} \cos x = (-1)^{k+1} \sin x, \quad k = 0, 1, 2, \ldots \\ D^{2k} \sin x = (-1)^k \sin x, \quad D^{2k+1} \sin x = (-1)^k \cos x, \quad k = 0, 1, 2, \ldots \end{cases}$$

ja Lausetta VII.4.3 nähdään, että jos $x \in \mathbb{R}$, $x \neq 0$ ja $n \in \mathbb{N}$, niin funktioille e^x, $\cos x$, $\sin x$ pätee jollakin $\xi \in (0, x)$ $(x > 0)$ tai $\xi \in (x, 0)$ $(x < 0)$

$$\boxed{\begin{aligned} e^x &= \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right) + \frac{e^\xi}{(n+1)!} x^{n+1}, \\ \cos x &= \left(1 - \frac{x^2}{2!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!}\right) + (-1)^{n+1} \frac{\cos \xi}{(2n+2)!} x^{2n+2}, \\ \sin x &= \left(x - \frac{x^3}{3!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}\right) + (-1)^{n+1} \frac{\cos \xi}{(2n+3)!} x^{2n+3}. \end{aligned}}$$

Tässä on sulkeilla ympäröity Taylorin polynomit

$$\begin{aligned} e^x &: \quad T_n(x, 0), \\ \cos x &: \quad T_{2n}(x, 0) = T_{2n+1}(x, 0), \\ \sin x &: \quad T_{2n+1}(x, 0) = T_{2n+2}(x, 0). \quad \square \end{aligned}$$

Taylorin lause on muotoiltavissa myös niin, että funktion $(n+1)$-kertaisen derivoituvuuden sijasta oletetaan ainoastaan n-kertainen derivoituvuus ja derivaatan $f^{(n)}$ jatkuvuus.

Lause VII.4.4 Jos f on n kertaa derivoituva välillä (a, b) ja $f^{(n)}$ on jatkuva ko. välillä, niin jäännöstermille $R_n(x) = f(x) - T_n(x, x_0)$ pätee jokaisella $x_0 \in (a, b)$

$$\lim_{x \to x_0} \frac{R_n^{(k)}(x)}{(x - x_0)^{n-k}} = 0, \quad k = 0 \ldots n.$$

<u>Todistus</u> Tapauksessa $n = 0$ väittämä on tosi jatkuvuuden määritelmän nojalla. Jos $n \geq 1$, niin oletusten ja Taylorin lauseen perusteella on

$$\begin{aligned} f(x) &= T_{n-1}(x, x_0) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n \\ &= \left[T_{n-1}(x, x_0) + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n\right] + \frac{1}{n!}\left[f^{(n)}(\xi) - f^{(n)}(x_0)\right](x - x_0)^n \\ &= T_n(x, x_0) + \frac{1}{n!}\left[f^{(n)}(\xi) - f^{(n)}(x_0)\right](x - x_0)^n, \quad x \in (a, b), \end{aligned}$$

missä $\xi = \xi(x) = x_0$, jos $x = x_0$, muulloin $\xi(x) \in (x_0, x)$ tai $\xi \in (x, x_0)$. Siis

$$R_n(x) \;=\; \frac{1}{n!}\left[f^{(n)}(\xi(x)) - f^{(n)}(x_0)\right](x - x_0)^n,$$

missä $\xi(x) \to x_0$ kun $x \to x_0$. Koska $f^{(n)}$ on jatkuva x_0:ssa, niin seuraa

$$\lim_{x\to x_0} \frac{R_n(x)}{(x - x_0)^n} = \lim_{x\to x_0} \frac{1}{n!}\left[f^{(n)}(\xi(x)) - f^{(n)}(x_0)\right] = 0.$$

Muut väitetyt raja-arvotulokset seuraavat tästä derivoimalla: Koska

$$f^{(k)}(x) = \left(\frac{d}{dx}\right)^k T_n(x, x_0) + R_n^{(k)}(x), \quad x \in (x_0, x_0 + a), \quad k = 1 \ldots n,$$

ja koska tässä $(d/dx)^k\, T_n(x, x_0) = f^{(k)}$:n Taylorin polynomi astetta $n - k$, niin jo todistetun perusteella

$$\lim_{x\to x_0} \frac{R_n^{(k)}(x)}{(x - x_0)^{n-k}} = 0, \quad k = 1 \ldots n. \quad \square$$

Taylorin polynomien nopea laskeminen

Joskus f:n derivaatat ovat niin hankalia laskea, että Taylorin polynomin saa määrätyksi suoremmin muilla menetelmillä, jolloin polynomin avulla voi päinvastoin määrittää derivaatat $f^{(k)}(x_0)$, $k = 0 \ldots n$ (!). Polynomia muilla keinoin määrättäessä riittää, että jäännöstermi saadaan riittävän pieneksi, sillä tälläkin kriteerillä polynomi on yksikäsitteinen:

Propositio VII.4.5 Olkoon funktio f n kertaa derivoituva välillä (a, b) ja olkoon $f^{(n)}$ jatkuva välillä (a, b). Tällöin jos p on polynomi astetta $\leq n$ siten, että jollakin $x_0 \in (a, b)$ pätee

$$\lim_{x\to x_0} \frac{f(x) - p(x)}{(x - x_0)^n} = 0,$$

niin $p(x) = f$:n Taylorin polynomi $T_n(x, x_0)$.

<u>Todistus</u> Kun merkitään $q(x) = p(x) - T_n(x, x_0)$, niin raja-arvojen yhdistelysääntöjen (Lause V.2.4), Lauseen VII.4.4 ja oletuksen mukaan

$$\lim_{x\to x_0} \frac{q(x)}{(x - x_0)^n} \;=\; \lim_{x\to x_0} \left[\frac{f(x) - T_n(x, x_0)}{(x - x_0)^n} - \frac{f(x) - p(x)}{(x - x_0)^n}\right] = 0 - 0 = 0.$$

Tämä on mahdollista vain kun $q(x) = 0$, koska q on polynomi astetta $\leq n$. $\square$

Seuraavissa esimerkeissä käytetään lyhennysmerkintää $[x^m]$ funktiosta muotoa $x^m g(x)$, missä g on rajoitettu pisteen $x = 0$ ympäristössä.

Esimerkki 3 $f(x) = (x+1)/\cos x, \quad T_5(x,0) = ?$

<u>Ratkaisu</u> Koska $\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} + [x^6]$ ja $1/(1-t) = 1 + t + t^2 + [t^3]$, niin

$$\begin{aligned} f(x) &= (x+1)\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)^{-1}(1+[x^6])^{-1} \\ &= (x+1)\left[1 - \left(\frac{x^2}{2} - \frac{x^4}{24}\right)\right]^{-1} + [x^6] \\ &= (x+1)\left[1 + \left(\frac{x^2}{2} - \frac{x^4}{24}\right) + \left(\frac{x^2}{2} - \frac{x^4}{24}\right)^2\right] + [x^6] \\ &= (x+1)\left(1 + \frac{1}{2}x^2 + \frac{5}{24}x^4\right) + [x^6] \\ &= \left(1 + x + \frac{1}{2}x^2 + \frac{1}{2}x^3 + \frac{5}{24}x^4 + \frac{5}{24}x^5\right) + [x^6]. \end{aligned}$$

Proposition VII.4.5 mukaan sulkeissa oleva polynomi = f:n Taylorin polynomi $T_5(x,0)$. $\square$

Esimerkki 4 $f(x) = 1/(1 + x^4 e^{x^2}), \quad f^{(8)}(0) = ?$

<u>Ratkaisu</u>

$$\begin{aligned} f(x) &= 1 - (x^4 e^{x^2}) + (x^4 e^{x^2})^2 + [x^{12}] \\ &= 1 - x^4\left(1 + x^2 + \frac{1}{2}x^4 + [x^6]\right) + x^8\left(1 + [x^2]\right)^2 + [x^{12}] \\ &= 1 - x^4 - x^6 + \frac{1}{2}x^8 + [x^{10}] = T_8(x,0) + [x^{10}] \\ &\Rightarrow \; f^{(8)}(0) = \frac{1}{2}\cdot 8! = \underline{\underline{20160}}. \quad \square \end{aligned}$$

Taylorin sarjat

Kun jäännöstermi Taylorin lauseessa VII.4.3 arvioidaan funktioille $\cos x$ ja $\sin x$, niin nähdään, että (ks. Esimerkki 2 edellä)

$$\begin{aligned} |\cos x - T_{2n}(x,0)| &\le \frac{1}{(2n+2)!}|x|^{2n+2}, \quad x \in \mathbb{R}, \\ |\sin x - T_{2n+1}(x,0)| &\le \frac{1}{(2n+3)!}|x|^{2n+3}, \quad x \in \mathbb{R}. \end{aligned}$$

Koska $|x|^n/n! \to 0 \; \forall x \in \mathbb{R}$, kun $n \to \infty$, niin jokaisella $x \in \mathbb{R}$ pätee

$$\cos x = \lim_{n\to\infty} T_{2n}(x,0), \qquad \sin x = \lim_{n\to\infty} T_{2n+1}(x,0),$$

eli (vrt. Harj.teht. VI.2:18)

$$\boxed{\begin{aligned} \cos x &= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}, \quad x \in \mathbb{R}, \\ \sin x &= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}, \quad x \in \mathbb{R}. \end{aligned}}$$

MÄÄRITELMÄ VII.4.6 Jos f on mielivaltaisen monta kertaa derivoituva x_0:ssa, niin sarja

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

on f:n *Taylorin sarja* x_0:ssa.[†]

Taylorin sarjojen teoriassa aivan ilmeisesti keskeisin kysymys on: Milloin sarja suppenee kohti $f(x)$:ää, ts. milloin pätee

$$f(x) = \lim_{n\to\infty} T_n(x, x_0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \ ?$$

Funktioiden $\sin x$ ja $\cos x$ kohdalla vastaus on: Aina, eli jokaisella $x \in \mathbb{R}$ (myös jokaisella $x_0 \in \mathbb{R}$). Myös eksponenttifunktion e^x kohdalla vastaus on sama; tälle Taylorin lause vahvistaa ennestään tunnetun tuloksen (vrt. Luku VI.2)

$$\boxed{e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad x \in \mathbb{R}.}$$

Tarkasti ottaen ym. kysymys Taylorin sarjan suppnemisesta sisältää kaksi erillistä kysymystä, kuten nähdään seuraavasta esimerkistä.

ESIMERKKI 5 Funktio

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{kun } x \neq 0, \\ 0\ , & \text{kun } x = 0 \end{cases}$$

on mielivaltaisen monta kertaa derivoituva pisteessä $x = 0$ (myös muualla) ja

$$f^{(k)}(0) = 0, \quad k = 0, 1, 2, \ldots ,$$

joten $T_n(x, 0) = 0\ \forall n$. Tässä tapauksessa siis Taylorin sarja suppenee $\forall x \in \mathbb{R}$, mutta $\lim_n T_n(x, 0) = 0 \neq f(x)$, kun $x \neq 0$. $\square$

[†]Tapauksessa $x_0 = 0$ käytetään Taylorin sarjasta myös nimitystä *Maclaurinin sarja.*

Esimerkin mukaan kahdella eri funktiolla voi olla sama Taylorin sarja (esimerkissä funktioilla f ja $g = 0$), joten Taylorin sarjan kertoimista (tai sarjan summasta) ei voi päätellä funktiota, josta sarja on johdettu. Useille 'normaaleille' funktioille f kuitenkin pätee, että f:n Taylorin sarjan summa $= f(x)$ aina kun sarja suppenee. Tällaisia normaalitapauksia ovat esim. rationaalifunktiot.

Esimerkki 6 Funktion $f(x) = 1/(1 + 4x^2)$ Taylorin polynomit origossa ovat

$$T_{2n}(x, 0) = T_{2n+1}(x, 0) = \sum_{k=0}^{n} (-4)^k x^{2k}, \quad n = 0, 1, \ldots$$

Taylorin sarja eli potenssisarja $\{T_n(x, 0),\ n = 0, 1, 2, \ldots\}$ suppenee tässä tapauksessa täsmälleen kun $|x| < 1/2$ (vrt. Luku I.12), ja tällöin summa $= f(x)$:

$$\sum_{k=0}^{\infty} (-4)^k x^{2k} = \frac{1}{1 + 4x^2} = f(x), \quad x \in (-\tfrac{1}{2}, \tfrac{1}{2}). \quad \square$$

Taylorin sarjojen suppenemista tutkittaessa voidaan aina tehdä muuttujan vaihdos $x - x_0 \hookrightarrow x$, jolloin riittää tarkastella yleistä potenssisarjaa muotoa

$$f(x) = \sum_{k=0}^{n} a_k x^k.$$

Tällaisen sarjan suppenemiskysymys on ratkaistu Luvussa I.12: Lauseen I.12.9 mukaan sarja suppenee joko (a) vain kun $x = 0$ tai (b) välillä $(-\rho, \rho)$ (mahdollisesti myös kun $x = \pm\rho$), missä ρ on sarjan suppenemissäde ($\rho \in \mathbb{R}_+$ tai $\rho = \infty$). Luvussa V.3 osoitettiin, että potenssisarjan summana määritelty funktio on mielivaltaisen monta kertaa derivoituva välillä $(-\rho, \rho)$ ja että derivaatat voidaan laskea derivoimalla sarja termeittäin (Lause V.3.4). Näin ollen jos potenssisarjan $\sum_{k=0}^{\infty} a_k x^k$ suppenemissäde on $\rho > 0$ ja $x_0 \in \mathbb{R}$, niin funktio

$$f(x) = \sum_{k=0}^{\infty} a_k \, (x - x_0)^k$$

on määritelty ja mielivaltaisen monta kertaa derivoituva välillä $(x_0 - \rho, x_0 + \rho)$ (koko $\mathbb{R}$:ssä, jos $\rho = \infty$) ja f:n derivaatat voidaan laskea derivoimalla sarja termeittäin. Kun derivoimispisteeksi valitaan erityisesti x_0, saadaan tulos

$$f^{(k)}(x_0) = k!\, a_k \quad \Leftrightarrow \quad a_k = \frac{f^{(k)}(x_0)}{k!}, \quad k = 0, 1, 2, \ldots$$

Siis $f(x)$ on esitettävissä muodossa

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k.$$

On tultu seuraavaan huomionarvoiseen tulokseen:

LAUSE VII.4.7 Jos sarja $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ suppenee välillä $(x_0 - \rho, x_0 + \rho)$, $\rho > 0$, niin ko. sarja = sarjan summana määritellyn funktion Taylorin sarja pisteessä x_0.

ESIMERKKI 7 Funktion $\sin x$ Taylorin sarjasta nähdään, että

$$\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k+1)!} = f(x) = \begin{cases} \sin x/x \ , & x \neq 0, \\ 1 \ , & x = 0. \end{cases}$$

Koska sarja suppenee $\forall x \in \mathbb{R}$, niin kyseessä on sarjan summana määritellyn funktion Taylorin sarja origossa. Funktio f on siis mielivaltaisen monta kertaa derivoituva jokaisessa pisteessä $x \in \mathbb{R}$, origo mukaan lukien (!). $\square$

ESIMERKKI 8 Ratkaise Taylorin sarjoilla alkuarvotehtävä

$$\begin{cases} y' = e^{-x^2}, & x \in \mathbb{R}, \\ y(0) = 0. \end{cases}$$

<u>Ratkaisu</u> Koska eksponenttifunktion e^x Taylorin sarja suppenee kaikkialla, niin

$$e^{-x^2} = \sum_{k=0}^{\infty} \frac{(-x^2)^k}{k!} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^{2k}, \quad x \in \mathbb{R}.$$

Kun valitaan

$$y(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)\,k!} x^{2k+1} = x - \frac{x^3}{18} + \frac{x^5}{600} - \ldots,$$

niin nähdään termeittäin derivoimalla, että $y'(x) = e^{-x^2}$, $x \in \mathbb{R}$. Koska on myös $y(0) = 0$, niin ratkaisu on tässä. $\square$

Taylorin lauseen todistus

Taylorin lauseelle on monia erilaisia todistustapoja. Seuraavassa 'lyhyen kaavan' mukaisessa todistuksessa päättelyn kulmakivi on Rollen lause (Lause V.6.2).

Olkoon $x_0 < x$ (tapaus $x_0 > x$ käsitellään vastaavasti) ja tarkastellaan välillä $[x_0, x]$ funktiota $g(t)$, joka määritellään

$$g(t) = f(t) - T_n(t, x_0) - H(t - x_0)^{n+1}, \quad H = \frac{f(x) - T_n(x, x_0)}{(x - x_0)^{n+1}}.$$

Tälle pätee $g^{(k)}(x_0) = 0,\ k = 0 \ldots n$ ja $g(x) = 0$. Koska siis $g(x_0) = g(x) = 0$, niin Rollen lauseen mukaan on $g'(\xi_1) = 0$ jollakin $\xi_1 \in (x_0, x) \subset (a, b)$. Jos $n \geq 1$, on myös $g'(x_0) = 0$, jolloin saman lauseen mukaan on $g''(\xi_2) = 0$ jollakin $\xi_2 \in (x_0, \xi_1)$. Jos edelleen $n \geq 2$, on myös $g''(x_0) = 0$, joten saman lauseen mukaan on $g'''(\xi_2) = 0$ jollakin $\xi_2 \in (x_0, \xi_1)$. Jatkamalla samalla tavoin päätellään, että yleisesti on $g^{(n)}(x_0) = g^{(n)}(\xi_n) = 0$ jollakin $\xi_n \in (x_0, \xi_{n-1})$, jolloin Rollen lauseen mukaan on $g^{(n+1)}(\xi_{n+1}) = 0$ jollakin $\xi_{n+1} \in (x_0, \xi_n) \subset (x_0, x) \subset (a, b)$. Mutta $g^{(n+1)}(t) = f^{(n+1)}(t) - H(n+1)!$ — Siis $f^{(n+1)}(\xi_{n+1}) - H(n+1)! = 0 \ \Rightarrow$ väite. $\square$

HARJOITUSTEHTÄVIÄ

1. Laske funktion $f(x) = x^4 + 3x^3 + x^2 + 2x + 8$, kaikki Taylorin polynomit $T_n(x, 2),\ n = 0, 1, 2, \ldots$ ja saata ne polynomin perusmuotoon (x:n potenssien mukaan). Piirrä ko. polynomien kuvaajat ja vertaa funktioon f.

2. Laske funktion $f(x) = (x-1)/(x-2)$ Taylorin polynomi $T_3(x, 0)$ ja piirrä funktion, Taylorin polynomin ja jäännöstermin kuvaajat.

3. Määritä seuraaville funktioille Taylorin polynomi $T_n(x, 0)$ annettua astetta n ja arvioi jäännöstermin Lagrangen muodosta mahdollisimman tarkasti luku $r_n = \max\limits_{x \in [-1,1]} |R_n(x)|$:

   a) $f(x) = \cosh x,\ n = 4$ b) $f(x) = e^{-0.2x},\ n = 3$
   c) $f(x) = \sqrt{5 + x},\ n = 3$ d) $f(x) = \sqrt[5]{5 - x},\ n = 3$
   e) $f(x) = 2^x,\ n = 5$ f) $f(x) = \ln(e + x),\ n = 6$
   g) $f(x) = \cot(x + \pi/3),\ n = 4$ h) $f(x) = \sin x - 1/\cos x,\ n = 3$

4. a) Funktioiden $\cosh x$ ja $\sinh x$ Taylorin polynomit $T_n(x, 0)$ voidaan laskea joko suoraan määritelmästä tai funktioiden e^x ja e^{-x} Taylorin polynomien avulla. Varmista, että kummallakin tavalla tulos on sama.

   b) Näytä, että parillisen (vastaavasti parittoman) funktion Taylorin polynomissa $T_n(x, 0)$ on vain parillisia (parittomia) potensseja.

5. Todista Taylorin lauseen avulla:

$$1 - \frac{1}{2}x^2 \le \cos x \le 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4\,, \quad \text{kun } x \in \left[-\frac{\pi}{2}\,,\,\frac{\pi}{2}\right].$$

Ovatko nämä epäyhtälöt tosia myös välin $[-\pi/2,\, \pi/2]$ ulkopuolella?

6. Harjoitustehtävässä V.5:15b oletetaan, että funktio f on m kertaa jatkuvasti derivoituva välillä $[a, b]$. Näytä, että tehtävän ratkaisu on $p_1(x) = T_m^+(a, x)$, $p_2(x) = T_m^-(b, x)$, missä $T_m^+(a, x)$ ja $T_m^-(b, x)$ ovat toispuolisten derivaattojen $\mathrm{D}_+^k f(a)$ ja $\mathrm{D}_-^k f(b)$ avulla määritellyt f:n Taylorin polynomit.

7. Laske seuraavien funktioiden Taylorin polynomi $T_n(x, 0)$ annettua astetta käyttäen mahdollisimman nopeita oikoteitä:

a) $f(x) = 2/(4 + x^3)$, $n = 12$ b) $f(x) = \cos x^4$, $n = 16$
c) $f(x) = \operatorname{Arcsin} x^3$, $n = 6$ d) $f(x) = (x^3 - x^5)\operatorname{Arctan} x^2$, $n = 7$

8. Laske seuraavat derivaatat käyttäen hyväksi Taylorin polynomeja:
a) $f(x) = \sin x^8$, $f^{(40)}(0)$ b) $f(x) = e^{-x^4}$, $f^{(20)}(0)$
c) $f(x) = x^2/(1 + x^4)$, $f^{(100)}(0)$ d) $f(x) = x^3 \ln(2 + x^2)$, $f^{(87)}(0)$
e) $f(x) = e^{x^3}/(1 + e^{x^3})$, $f^{(9)}(0)$ f) $f(x) = \cos(x^2 \sin^2 x)$, $f^{(12)}(0)$

9. Seuraavilla käyrillä on kääntymispiste (vrt. edellisen luvun Esimerkki 3) annetussa pisteessä P. Mihin suuntaan käyrä lähtee pisteestä P?
a) $x(t) = 4t + 1/t$, $y(t) = t^2 + 16/t$, $P = (17/2, 12)$
b) $\vec{r}(t) = t^4\vec{i} + (2 - 2\cos t - t^2)\vec{j} + (t^2 - t\sinh t)\vec{k}$, $P = (0, 0, 0)$

10. Mikä on sarjan

$$\text{a) } 1 + 4 + \frac{16}{2!} + \frac{64}{3!} + \frac{256}{4!} + \dots \qquad \text{b) } 1 + \frac{4}{3!} + \frac{16}{5!} + \frac{64}{7!} + \dots$$

summa?

11. Esitä seuraavien funktioiden funktioiden Taylorin sarjat pisteessä $x_0 = 0$:
a) e^{3x+1} b) $\cos(2x^3)$ c) $\sin(x - \frac{\pi}{4})$ d) $\cos(2x - \pi)$ e) $x^2 \sin 3x$
f) $\sin x \cos x$ g) $(1 + x^3)/(1 + x^2)$ h) $\ln(2 + x^2)$ i) $x^2 \ln(1 + x)$

12. Määritä seuraavien funktioiden Taylorin sarja annetussa pisteessä sekä sarjan suppenemisväli:
a) e^{-2x}, $x_0 = -1$ b) $\sin x$, $x_0 = \frac{\pi}{2}$ c) $\ln x$, $x_0 = 1$
d) $\cos x$, $x_0 = \pi$ e) $\cos^2 x$, $x_0 = \frac{\pi}{8}$ f) $x/(4 + x)$, $x_0 = 3$

13. Olkoon
$$f(x) = \begin{cases} \dfrac{2\cos x - 2}{x^2}, & \text{kun } x \neq 0, \\ -1, & \text{kun } x = 0. \end{cases}$$
Laske $f'(0)$ ja $f''(0)$ a) suoraan derivaatan määritelmästä, b) f:n Taylorin sarjan avulla.

14. Seuraavat funktiot f määritellään kukin jatkuvaksi pisteessä $x = 0$, jolloin funktiot ovat tässä pisteessä (ja siis koko $\mathbb{R}$:ssä) mielivaltaisen monta kertaa derivoituvia. Määritä funktioiden Taylorin sarjat pisteesä $x_0 = 0$ ja näiden avulla $f(0)$, $f'(0)$ ja $f''(0)$.
$$\text{a)}\ \frac{e^x - 1}{x} \qquad \text{b)}\ \frac{e^{-2x} - 1 + 2x - 2x^2}{x^3} \qquad \text{c)}\ \frac{\sinh x - x}{x^3}$$
$$\text{d)}\ \frac{\ln(1+x) - \sin x}{x^2} \qquad \text{e)}\ \frac{e^{2x} - 4e^{-x} - 6x + 3}{x^2}$$

15. Tiedetään, että
$$\sum_{k=0}^{\infty} a_k (x+\pi)^k \;=\; x\cos x, \quad x \in \mathbb{R}.$$
Laske sarjan kertoimet a_0, a_1 ja a_2.

16. (*) Todista Taylorin lauseen avulla Newtonin menetelmän konvergenssia koskeva Lause V.7.4 .

17. (*) Potenssisarja $\sum_{k=0}^{\infty} a_k (x+1)^k$ suppenee pisteen $x = -1$ lähellä, jolloin sarjan summa on
$$\sum_{k=0}^{\infty} a_k (x+1)^k \;=\; \frac{x+2}{x^2+2x+3}\,.$$
Mitkä ovat kertoimien a_0, a_1 ja a_2 arvot ja mikä on sarjan suppenemisväli?

VII.5 Taylorin polynomien sovelluksia

Tässä luvussa tarkastellaan eräitä tavallisia Taylorin polynomien ja Taylorin lauseen käyttötapoja funktiotutkimuksessa ja funktioiden approksimoimisessa. Taylorin polynomeja laskettaessa tai jäännöstermiä arvioitaessa voidaan laskuja usein huomattavasti lyhentää ja selkiyttää käyttämällä nk. *suuruusluokkamerkintöjä*. Esimakua tällaisista merkinnöistä on saatu jo edellisen luvun Esimerkeissä 3 ja 4, joissa käytettiin lyhennysmerkintää $[x^m]$. Tämän tilalla on tavallisempaa käyttää merkintää $\mathcal{O}(|x|^m)$, joka luetaan 'suuruusluokkaa x^m' tai vain 'oo x^m'. Toinen, merkitykseltään hiukan voimakkaampi suuruusluokkamerkintä on $o(|x|^m)$, joka luetaan 'pikku oo x^m'. Näissä merkinnöissä on $|x|^m$ nk. vertailufunktio. Tämän tilalla voi olla yleisempi, ei-negatiivisia arvoja saava funktio, esim. $|x - x_0|^\alpha,\ \alpha \in \mathbb{Q}$.

MÄÄRITELMÄ VII.5.1 **(Suuruusluokkamerkinnät $\mathcal{O}$ ja o)** Jos f ja g on määritelty välillä $[x_0-a, x_0+a]$ $(a > 0)$ ja $g(x) \geq 0\ \forall\, x \in [x_0-a, x_0+a]$, niin käytetään merkintää

$$f(x) = \mathcal{O}(g(x)), \quad x \in [x_0 - a, x_0 + a],$$

jos on olemassa vakio $C \in \mathbb{R}_+$, siten että pätee

$$|f(x)| \leq Cg(x) \quad \forall\ x \in [x_0 - a, x_0 + a].$$

Jos $g(x) > 0\ \forall\, x \neq x_0$ ja pätee $\lim_{x\to x_0}(f/g)(x) = 0$, niin käytetään merkintää

$$f(x) = o(g(x)), \quad \text{kun } x \to x_0.$$

Jos merkinnällä $\mathcal{O}(..)$ halutaan ainoastaan kertoa, että arvio on pätevä jossakin x_0:n ympäristössä, niin tämä voidaan ilmaista kirjoittamalla, kuten vastaavassa merkinnässä $o(..)$,

$$f(x) = \mathcal{O}(g(x)), \quad \text{kun } x \to x_0.$$

Näissä merkinnöissä voi raja-arvon x_0 tilalla olla $\pm\infty$, jolloin kyse on f:n arvioimisesta suurilla $|x|$:n arvoilla. Vertailufunktio on tällöin tyypillisesti $g(x) = |x|^\alpha$ jollakin $\alpha \in \mathbb{Q}$.

Laskuissa suuruusluokkamerkintöjä on kätevä käsitellä kuten funktioita, jolloin niitä voi yhdistellä funktioiden tavoin. Esimerkiksi $f(x) + o(x^2)$ tarkoittaa funktiota $f(x) + g(x)$, missä $g(x) = o(x^2)$, ja $\mathcal{O}(x^2) + \mathcal{O}(y^2)$ tarkoittaa samaa kuin $\mathcal{O}(x^2 + y^2)$. Määritelmästä VII.5.1 on helposti todennettavissa seuraavat suuruusluoka-algebran säännöt (Harj.teht. 1). Säännöissä oletetaan, että $x, y \geq 0$.

$$\boxed{\begin{aligned} \mathcal{O}(x) + \mathcal{O}(y) &= \mathcal{O}(\max\{x, y\}), \\ \mathcal{O}(x)\mathcal{O}(y) &= \mathcal{O}(xy), \\ o(x)\mathcal{O}(y) &= o(xy). \end{aligned}}$$

Sovelluksissa kahta jälkimmäistä sääntöä käytetään tavallisimmin muodossa

$$\mathcal{O}(x^\alpha)\mathcal{O}(x^\beta) = \mathcal{O}(x^{\alpha+\beta}), \quad o(x^\alpha)\mathcal{O}(x^\beta) = o(x^{\alpha+\beta}), \quad x \geq 0,\ \alpha, \beta \in \mathbb{R}.$$

Esimerkki 1 Muuttujan vaihdon $x^2 = t$ ja Taylorin lauseen avulla päätellään, että pisteen $x = 0$ lähellä (esim. välillä $[-1, 1]$) on

$$\sqrt{x^2+1} = \sqrt{1+t} = 1 + \frac{1}{2}t + \mathcal{O}(t^2) = 1 + \frac{1}{2}x^2 + \mathcal{O}(x^4).$$

Suurilla x:n arvoilla ($x \to \infty$) saadaan muuttujan vaihdolla $x^{-2} = t$ vastaavasti arvio

$$\sqrt{x^2+1} = x\sqrt{1+x^{-2}} = x\left(1 + \frac{1}{2x^2} + \mathcal{O}(x^{-4})\right) = x + \frac{1}{2x} + \mathcal{O}(x^{-3}). \quad \square$$

Esimerkki 2 Näytä, että pienillä $|x|$:n arvoilla on

$$f(x) = \frac{1-x}{1+2x-x^2+\mathcal{O}(|x|^3)} = 1 - 3x + 7x^2 + \mathcal{O}(|x|^3).$$

<u>Ratkaisu</u> Koska pienillä $|t|$:n arvoilla on $1/(1+t) = 1 - t + t^2 + \mathcal{O}(|t|^3)$, niin suuruusluokka-algebra antaa (vrt. edellisen luvun Esimerkit 3–4)

$$\begin{aligned} f(x) &= (1-x)\left\{1 - \left[2x - x^2 + \mathcal{O}(|x|^3)\right] + \left[2x - x^2 + \mathcal{O}(|x|^3)\right]^2 + \mathcal{O}(|x|^3)\right\} \\ &= (1-x)[1 - (2x - x^2) + 4x^2] + \mathcal{O}(|x|^3) = 1 - 3x + 7x^2 + \mathcal{O}(|x|^3). \quad \square \end{aligned}$$

Funktion approksimointi

Taylorin polynomien avulla voidaan differentiaaliin perustuva approksimaatio

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x$$

(vrt. Luku VII.1) sekä yleistää että tarkentaa: Jos f on $n+1$ kertaa jatkuvasti derivoituva pisteen x ympäristössä, niin

$$\boxed{\begin{aligned} f(x+\Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2!}f''(x)(\Delta x)^2 \\ + \cdots + \frac{1}{n!}f^{(n)}(x)(\Delta x)^n + \mathcal{O}(|\Delta x|^{n+1}). \end{aligned}}$$

Hyvä arvio approksimaation virheelle, tai ainakin sen suuruusluokalle, on yleensä

$$\text{virhe} = \text{tarkka} - \text{approksimaatio} \approx \frac{1}{(n+1)!}f^{(n+1)}(x)(\Delta x)^{n+1}.$$

Esimerkki 3 $\sqrt[10]{1000} \approx ?$

Ratkaisu

$$\sqrt[10]{1000} \;=\; \sqrt[10]{1024-24} \;=\; 2\sqrt[10]{1-\tfrac{3}{128}}\,.$$

Kun merkitään $f(x) = \sqrt[10]{1-x}$, niin

$$\begin{aligned} f'(x) &= -\tfrac{1}{10}(1-x)^{-9/10}, & f''(x) &= -\tfrac{9}{100}(1-x)^{-19/10}, \\ f'''(x) &= -\tfrac{171}{1000}(1-x)^{-29/10}, & f^{(4)}(x) &= -\tfrac{4959}{10000}(1-x)^{-39/10} \end{aligned}$$

$$\begin{aligned} \Rightarrow\ \sqrt[10]{1000} &= 2f(\tfrac{3}{128}) \\ &\approx 2\left[1 + \left(-\tfrac{1}{10}\right)\cdot\tfrac{3}{128} + \tfrac{1}{2}\cdot\left(-\tfrac{9}{100}\right)\left(\tfrac{3}{128}\right)^2 + \tfrac{1}{6}\cdot\left(-\tfrac{171}{1000}\right)\left(\tfrac{3}{128}\right)^3\right] \\ &\approx 2\,(\,1 - 0.00234275 - 0.0000247192 - 0.000000366926\,) \\ &= 1.995262327671. \end{aligned}$$

Virhe (tarkka arvo − likiarvo) on suuruusluokkaa

$$+\frac{1}{4!}f^{(4)}(0)\cdot(\Delta x)^4 = -\tfrac{1653}{80000}\cdot\left(\tfrac{3}{128}\right)^4 \approx -6\cdot 10^{-9}.$$

Oikea 12-desimaalinen arvo on

$$\sqrt[10]{1000} \approx 1.995262314969. \quad \square$$

Esimerkki 4 Arvioi $\sin 31^\circ$ lähtien arvosta $\sin 30^\circ = \frac{1}{2}$.

Ratkaisu Kun $f(x) = \sin x$, $x = \frac{\pi}{6}$, $\Delta x = \frac{\pi}{180}$, on

$$\begin{aligned} f(x+\Delta x) &\approx \sin x + \cos x\,\Delta x - \tfrac{1}{2}\sin x\,(\Delta x)^2 \\ &= \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}\cdot\tfrac{\pi}{180} - \tfrac{1}{2}\cdot\tfrac{1}{2}\cdot\left(\tfrac{\pi}{180}\right)^2 \;\approx\; 0.51503882. \end{aligned}$$

Virhe (tarkka arvo − likiarvo) on luokkaa

$$-\frac{1}{3!}\cos x\,(\Delta x)^3 \approx -8\cdot 10^{-7}.$$

Oikea 8-desimaalinen arvo on: $\sin 31^\circ \approx 0.51503807$. $\square$

Paikalliset ääriarvot

Lauseen V.5.2 mukaan derivoituvan funktion paikallisessa ääriarvokohdassa on myös derivaatan nollakohta. Jos funktio on derivaatan nollakohdan ympäristössä riittävän säännöllinen, niin Taylorin polynomien avulla voidaan selvittää, onko kyseessä todella ääriarvokohta ja jos, niin minkälaatuinen.

LAUSE VII.5.2 Jos f on pisteen c ympäristössä n kertaa jatkuvasti derivoituva, $n \geq 2$, ja pätee $f^{(k)}(c) = 0, \quad k = 1 \ldots n-1$ ja $f^{(n)}(c) \neq 0$, niin

a) jos n on pariton, niin f:llä ei ole c:ssä paikallista ääriarvoa,

b) jos n on parillinen, niin f:llä on c:ssä

- paikallinen minimi, jos $f^{(n)}(c) > 0$,
- paikallinen maksimi, jos $f^{(n)}(c) < 0$.

<u>Todistus</u> Lauseen VII.4.4 ja oletuksien mukaan

$$f(x) = f(c) + \frac{1}{n!} f^{(n)}(c)(x-c)^n + o(|x-c|^n), \quad \text{kun } x \to c.$$

Koska $f^{(n)}(c) \neq 0$, tämä voidaan kirjoittaa myös muotoon

$$f(x) = f(c) + \frac{1}{n!} f^{(n)}(c)\,[\,1 + o(1)\,]\,(x-c)^n, \quad \text{kun } x \to c,$$

jolloin nähdään, että riittävän pienellä δ ($\delta > 0$) on oltava

$$f(x) = f(c) + \frac{1}{n!} f^{(n)}(c) k(x)(x-c)^n, \quad x \in [c-\delta, c+\delta],$$

missä esimerkiksi

$$\frac{1}{2} \leq k(x) \leq \frac{3}{2}\,.$$

Näin ollen jos n on parillinen ja $f^{n)}(c) > 0$, on $f(x) > f(c)$ kun $x \in [c-\delta, c)$ tai $x \in (c, c+\delta]$, jolloin c on paikallinen minimikohta (Määritelmä V.5.1). Muissa tapauksissa on päättely vastaava (vrt. kuvio). $\square$

n pariton, $f^{(n)}(c) > 0$ | n pariton, $f^{(n)}(c) < 0$ | n parillinen $f^{(n)}(c) > 0$ | n parillinen, $f^{(n)}(c) < 0$

Esimerkki 5 Tutki mahdollisen ääriarvokohdan laatu seuraavissa tapauksissa:

a) $f(x) = \sin x + \cos x, \quad c = \pi/4$
b) $f(x) = e^{-x} + \sin x + \sqrt{1-x^2}, \quad c = 0$
c) $f(x) = 1 - \cos x - \sqrt{1+x^2}, \quad c = 0$

Ratkaisu a) $f'(\pi/4) = 0, \quad f''(\pi/4) = -\sqrt{2} \quad \Rightarrow$ paik. maksimi.

b) $f'(0) = f''(0) = 0, \quad f'''(0) = -2 \quad \Rightarrow$ ei ääriarvoa.

c) $f^{(k)}(0) = 0,\ k = 1\ldots 3,\ f^{(4)}(0) = 5 \quad \Rightarrow$ paik. minimi. □

Funktio $\frac{f(x)}{g(x)} = \frac{0}{0}$, kun $x = a$

Jos funktiot f ja g ovat pisteen a ympäristössä säännöllisiä ja $f(a) = g(a) = 0$, niin Taylorin lauseen avulla voidaan tutkia, millainen funktio $F(x) = f(x)/g(x)$ on pisteen a ympäristössä. Ensinnäkin jos halutaan ratkaista raja-arvokysymys

$$\lim_{x\to a} \frac{f(x)}{g(x)} = ?$$

niin Taylorin lauseeseen perustuva menettely on seuraava: Oletetaan, että f ja g ovat n kertaa jatkuvasti derivoituvia pisteen a ympäristössä ja että pätee

$$\frac{f(a)}{g(a)} = \frac{0}{0}, \quad \frac{f'(a)}{g'(a)} = \frac{0}{0}, \ \ldots, \ \frac{f^{(n-1)}(a)}{g^{(n-1)}(a)} = \frac{0}{0}, \quad \frac{f^{(n)}(a)}{g^{(n)}(a)} = \frac{A}{B},$$

missä $A \neq 0$ tai $B \neq 0$. Tällöin Lauseen VII.4.4 mukaan

$$\frac{f(x)}{g(x)} = \frac{\frac{A}{n!}(x-a)^n + o(|x-a|^n))}{\frac{B}{n!}(x-a)^n + o(|x-a|^n))} = \frac{A + o(1)}{B + o(1)}, \quad \text{kun } x \to a.$$

Päätellään, että jos $A \neq 0$ ja $B = 0$, niin $|f(x)/g(x)| \to \infty$, kun $x \to a$, jolloin raja-arvoa ei ole (reaalilukuna — voi olla $\lim = \pm\infty$). Muussa tapauksessa, eli jos $B \neq 0$, on

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{A}{B} = \frac{f^{(n)}(a)}{g^{(n)}(a)},$$

ja tällöin on yleisemminkin

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} = \ \ldots \ = \lim_{x\to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}.$$

Tämä tulos on siis voimassa (tehtyjen säännöllisyysoletusten puitteissa), kunhan ketjun viimeinen raja-arvo on olemassa. — Tulos on jo ennestään tuttu l'Hospitalin sääntönä (Lause V.6.10).

Esimerkki 6 Määritä raja-arvo $\lim\limits_{x\to 0} \dfrac{\sin 2x - 2x}{\sinh x - x}$.

<u>Ratkaisu</u> Funktiot $f(x) = \sin 2x - 2x$ ja $g(x) = \sinh x - x$ ovat sileitä $\mathbb{R}$:ssä. Derivoimalla todetaan

$$\frac{f'(x)}{g'(x)} = \frac{2\cos 2x - 2}{\cosh x - 1}, \quad \frac{f''(x)}{g''(x)} = \frac{-4\sin 2x}{\sinh x}, \quad \frac{f'''(x)}{g'''(x)} = \frac{-8\cos 2x}{\cosh x}, \ \ldots$$

Havaitaan, että $f^{(k)}(0) = g^{(k)}(0) = 0$, kun $k = 0, 1, 2$, ja $f'''(0) = -8$, $g'''(0) = 1$. Siis

$$\lim_{x\to 0} \frac{\sin 2x - 2x}{\sinh x - x} = \frac{-8}{1} = -8. \quad \square$$

Taylorin lauseen perusteella saadaan siis raja-arvolle $\lim_{x\to a} f(x)/g(x)$ sama laskukaava kuin l'Hospitalin säännöllä. Funktiosta $f(x)/g(x)$ saadaan kuitenkin Taylorin polynomien avulla selville paljon muutakin. Seuraavassa esimerkki vaativammasta tehtävän asettelusta.

Esimerkki 7 Määritä polynomit p ja q siten, että väleillä $[-1, 0)$ ja $(0, 1]$ pätee

$$F(x) = \frac{\sin 2x - 2x}{\sinh x - x} = p(x) + \mathcal{O}(x^4), \qquad G(x) = \frac{\sin x}{2x^2} + \frac{\cos x - 1}{x^3} = q(x) + \mathcal{O}(|x|^5).$$

<u>Ratkaisu</u>

$$\begin{aligned}
\sin 2x - 2x &= \Big(2x - \frac{1}{6}(2x)^3 + \frac{1}{120}(2x)^5 + \mathcal{O}(|x|^7)\Big) - 2x \\
&= -\frac{4}{3}x^3 + \frac{4}{15}x^5 + \mathcal{O}(|x|^7), \\
\sinh x - x &= \Big(x + \frac{1}{6}x^3 + \frac{1}{120}x^5 + \mathcal{O}(|x|^7)\Big) - x \\
&= \frac{1}{6}x^3 + \frac{1}{120}x^5 + \mathcal{O}(|x|^7) \\
\Rightarrow \quad \frac{\sin 2x - 2x}{\sinh x - x} &= \frac{-8 + \frac{8}{5}x^2 + \mathcal{O}(x^4)}{1 + \frac{1}{20}x^2 + \mathcal{O}(x^4)} \\
&= -8 + 2x^2 + \mathcal{O}(x^4) = p(x) + \mathcal{O}(x^4).
\end{aligned}$$

$$\begin{aligned}
\frac{\sin x}{2x^2} + \frac{\cos x - 1}{x^3} &= \frac{1}{2x^2}\Big(x - \frac{x^3}{6} + \frac{x^5}{120} + \mathcal{O}(|x|^7)\Big) \\
&\quad + \frac{1}{x^3}\Big(-\frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \mathcal{O}(x^8)\Big) \\
&= -\frac{1}{24}x + \frac{1}{360}x^3 + \mathcal{O}(|x^5|) = q(x) + \mathcal{O}(|x|^5).
\end{aligned}$$

Tuloksista voi päätellä esimerkiksi raja-arvot

$$\lim_{x\to 0}\left(\frac{\sin 2x - 2x}{x^2\sinh x - x^3} + \frac{8}{x^2}\right) = 2, \qquad \lim_{x\to 0}\left(\frac{\sin x}{2x^3} + \frac{\cos x - 1}{x^4}\right) = -\frac{1}{24}\,.$$

Nähdään myös, että jos asetetaan (jatkamisperiaatteella) $F(0) = -8$ ja $G(0) = 0$, niin F:llä on pisteessä $x = 0$ paikallinen minimi ja G on origon ympäristössä aidosti vähenevä. $\square$

Differenssiapproksimaatiot

Taylorin lauseella on usein keskeinen rooli, kun halutaan johtaa virhearvioita numeerisille (likimääräisille) laskentamenetelmille, jotka perustuvat oletettuun funktion säännöllisyyteen. Tarkastellaan esimerkkinä derivaattojen numeerisessa laskemisessa käytettäviä nk. *differenssiapproksimaatioita.* Näissä ei funktiota oleteta tunnetuksi lausekkeena vaan riittää tuntea funktion arvo äärellisen monessa (tyypillisesti vain muutamassa) pisteessä tarkasteltavan pisteen lähellä. Seuraavat kolme ovat differenssiapproksimaatioista tavallisimmat (ol. $h > 0$):

$$f'(a) \approx \frac{f(a+h) - f(a)}{h}\,, \tag{a}$$

$$f'(a) \approx \frac{f(a+h) - f(a-h)}{2h}\,, \tag{b}$$

$$f''(a) \approx \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}\,. \tag{c}$$

Näistä (a) perustuu suoraan derivaatan määritelmään. Approksimaatioita (b) ja (c) sanotaan *keskeis*differenssiapproksimaatioiksi, syystä että funktio evaluoidaan näissä a:n suhteen symmetrisessä pisteistössä.

Em. approksimaatioille voidaan johtaa virhearvio Taylorin lauseesta olettaen, että f on riittävän säännöllinen. Tarkastellaan esimerkkinä approksimaatiota (b), muut jätetään harjoitustehtäviksi (Harj.teht. 14,15).

Propositio VII.5.3 Jos f on kolmesti jatkuvasti derivoituva välilä $[a-h, a+h]$, niin differenssiapproksimaatiolle (b) pätee virhearvio

$$\left| f'(a) - \frac{f(a+h) - f(a-h)}{2h} \right| \le \frac{1}{6} M\, h^2, \quad M = \max_{x\in[a-h,a+h]} |f'''(x)|.$$

<u>Todistus</u> Oletuksien perusteella on joillakin $\xi_1 \in (a, a+h)$ ja $\xi_2 \in (a-h, a)$

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}f''(a)\,h^2 + \frac{1}{6}f'''(\xi_1)\,h^3,$$

$$f(a-h) = f(a) - f'(a)h + \frac{1}{2}f''(a)\,h^2 - \frac{1}{6}f'''(\xi_2)\,h^3$$

(Lause VII.4.3: välin $[a, b]$ tilalla väli $[a-h, a+h]$, $x_0 = a$, $x = a \pm h$, $n = 2$). Näistä seuraa vähennyslaskulla

$$f'(a) - \frac{f(a+h) - f(a-h)}{2h} = -\frac{1}{12}\left[f'''(\xi_1) + f'''(\xi_2)\right] h^2.$$

Käyttämällä oikealla arviota $|f'''(\xi_1) + f'''(\xi_2)| \leq |f'''(\xi_1)| + |f'''(\xi_2)| \leq 2M$ seuraa väite. $\square$

Proposition VII.5.3 perusteella approksimaatio (b) on tarkka toisen asteen polynomeille (joille $M = 0$). Valitsemalla $f(x) = \frac{1}{6}M(x-a)^3$ nähdään myös, että virhearvio on tehdyin oletuksin tarkin mahdollinen.

Yleisesti sanotaan, että differenssiapproksimaation (tarkkuuden) *kertaluku* on r, jos sen virhe on $\mathcal{O}(h^r)$ mutta ei $o(h^r)$ yleiselle, riittävän säännölliselle funktiolle. Approksimaation (b) kertaluku on siis $r = 2$, eli kyseessä on *toisen kertaluvun* approksimaatio. Approksimaation (c) kertaluku on samoin $r = 2$ (Harj.teht. 15), sen sijaan (a) on ensimmäisen kertaluvun ($r = 1$) approksimaatio (Harj.teht. 14).

HARJOITUSTEHTÄVIÄ

1. Perustele suuruusluokka-algebran säännöt (ol. $x, y \geq 0$)

$$\mathcal{O}(x) + \mathcal{O}(y) = \mathcal{O}(\max\{x, y\}), \quad \mathcal{O}(x)\mathcal{O}(y) = \mathcal{O}(xy), \quad o(x)\mathcal{O}(y) = o(xy).$$

2. Mitä funktion f ominaisuuksia tarkoitetaan seuraavilla merkinnöillä?
   a) $f(x) = f(a) + o(1)$, kun $x \to a$.
   b) $f(x_1) - f(x_2) = \mathcal{O}(|x_1 - x_2|)$, kun $x_1, x_2 \in [a, b]$.
   c) $f(x) = f(a) + k(x-a) + o(|x-a|)$, kun $x \to a$.

3. Anna suuruusluokka-arviot seuraavien approksimaatioiden virheille:
   a) $\sqrt{x^4+2} \approx \sqrt{2}$, kun $x \to 0$
   b) $\sqrt{x^4+3x} \approx \sqrt{3x}$, kun $x \to 0^+$
   c) $\sqrt{x^4+4x} \approx x^2$, kun $x \to \pm\infty$
   d) $\sqrt{x^4+4x} \approx x^2 + 2x^{-1}$ kun $x \to \infty$
   e) $x/\sin x \approx 1$, kun $x \to 0$
   f) $\ln(1+e^x) \approx x$, kun $x \to \infty$

4. Arvioi approksimaation $f(x) \approx T_2(x, 0)$ virheen itseisarvo välillä $[-1/2,\ 1/2]$ seuraaville funktioille:
   a) $\sqrt{1+x}$ b) $\sqrt[10]{1+x}$ c) $\tan x$ d) $\ln(1+x)$ e) $e^{\sin x}$

5. Arvioi seuraavien funktioapproksimaatioiden virhe:

a) $\sqrt[3]{8+x} \approx 2 + \dfrac{x}{12} - \dfrac{x^2}{288}\,, \quad \text{kun } |x| \le 1$

b) $\sin x \approx x - \dfrac{1}{6}x^3, \quad \text{kun } x \mathrel{\widehat{=}} \alpha \in [0°, 10°]$

c) $\dfrac{1}{x^4}\left[\dfrac{1}{\sqrt{1+x^2}} - \sqrt{1-x^2}\right] \approx \dfrac{1}{2}\,, \quad \text{kun } |x| \le 0.15 \text{ ja } x \ne 0$

6. Jos $T_n(x,0)$ on funktion $f(x) = \cos x$ Taylorin polynomi, niin millä a:n ja n:n arvoilla voidaan taata, että
a) $|\cos x - T_2(x,0)| \le 10^{-4}$ välillä $[-a, a]$,
b) $|\cos x - T_n(x,0)| \le 10^{-4}$ välillä $[-\pi/2,\, \pi/2]$?

7. a) Laske luvulle $1/\sqrt[4]{e}$ rationaalinen likiarvo approksimoimalla funktiota e^x toisen asteen Taylorin polynomilla. Arvioi approksimaation virhe Taylorin lauseen avulla ja vertaa virheen tarkkaan arvoon.
b) Laske luvulle $a = \sqrt[12]{4000}$ rationaalinen likiarvo kirjoittamalla $4000 = 4096(1-x)$ ja approksimoimalla funktiota $f(x) = \sqrt[12]{1-x}$ toisen asteen Taylorin polynomilla. Arvioi approksimaation virhe Taylorin lauseen avulla ja vertaa virheen tarkkaan arvoon.

8. Suhteellisuusteorian mukaan vauhdilla v liikkuvan partikkelin liike-energia on

$$E_k = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} - mc^2,$$

missä m on partikkelin massa ja $c \approx 3 \cdot 10^8$ m/s on valon nopeus. Millaisilla vauhdin v arvoilla approksimaation $E_k \approx \frac{1}{2}mv^2$ suhteellinen virhe on enintään 0.01% ?

9. Seuraavilla funktioilla piste $x = 0$ on mahdollinen paikallinen minimi- tai maksimikohta. Tutki asia derivoimalla!

a) $(1-x)e^x$ b) $(2-x^2)\cos x$ c) $(x-x^4)\sin x$ d) $\sqrt{1-x^2} - 1/\sqrt{1+x^2}$

10. Millä a:n ja b:n arvoilla funktiolla

a) $e^x + ax + bx^2$ b) $x\sin x + ax^2 + bx^4$ c) $\sin x + \cos x + ax + bx^2$

on paikallinen minimi pisteessä $x = 0$?

11. Yhtälö $y(\cos y - \sin y) = 2\sin x + \cos x + ax + b$ määritelee pisteen $x = 0$ ympäristössä funktion $y(x)$. Määritä vakiot a ja b, kun tiedetään, että $y(x)$ saavuttaa pisteessä $x = 0$ paikallisen ääriarvon $y(0) = 0$. Onko kyseessä maksimi vai minimi?

12. Määritä seuraavat raja-arvot:

a) $\displaystyle\lim_{x\to 0} \frac{\sqrt[7]{1+x}-1}{x}$ b) $\displaystyle\lim_{x\to 0} \frac{3x}{\tan 4x}$ c) $\displaystyle\lim_{x\to 0} \frac{2\cos x - 2 + x^2}{x^4}$

d) $\displaystyle\lim_{x\to 0} \frac{x - \sin x}{x - \tan x}$ e) $\displaystyle\lim_{x\to 0} \frac{10^x - e^x}{x}$ f) $\displaystyle\lim_{x\to 0} \frac{1-\cos ax}{1-\cos bx}$

g) $\displaystyle\lim_{x\to \pi/2} \frac{\cos 3x}{\pi - 2x}$ h) $\displaystyle\lim_{t\to \pi/2} \frac{\ln \sin t}{\cos t}$ i) $\displaystyle\lim_{x\to 0^+} \frac{\sin^2 x}{\tan x - x}$

j) $\displaystyle\lim_{x\to 1} \frac{\ln(ex) - 1}{\sin \pi x}$ k) $\displaystyle\lim_{x\to -\infty} x \sin \frac{1}{x}$ l) $\displaystyle\lim_{x\to \infty} x^2 \ln(\cos \frac{\pi}{x})$

m) $\displaystyle\lim_{x\to \infty} \left[\sqrt{x^2 + 154x} - \ln(5 + e^x)\right]$ n) $\displaystyle\lim_{x\to \infty} \left(\sqrt[6]{x^6 + 42x^5 + 77x} - x\right)$

13. Määritä seuraaville funktioille f mahdollisimman alhaista astetta oleva polynomi p siten, että annetulla m ja jollakin $\delta > 0$ on $f(x) = p(x) + \mathcal{O}(|x|^m)$, kun $0 < |x| < \delta$. Tutki myös pisteen $x = 0$ laatu mahdollisena f:n paikallisena ääriarvokohtana, kun asetetaan $f(0) = \lim_{x\to 0} f(x)$.

a) $\displaystyle\frac{1-\cos x}{x^2}$, $m = 7$ b) $\displaystyle\frac{\sin x - x}{x^3}$, $m = 8$ c) $\displaystyle\frac{x^3}{\sin x - x}$, $m = 3$

d) $\displaystyle\frac{x}{e^x - 1}$, $m = 3$ e) $\displaystyle\frac{\ln(1+x)}{x}$, $m = 5$ f) $\displaystyle\frac{\sqrt{1+x}-1}{\sqrt[3]{1-x}-1}$, $m = 3$

14. Näytä, että jos f on jatkuva välillä $[a, a+h]$ ja kahdesti derivoituva välillä $(a, a+h)$, niin jollakin $\xi \in (a, a+h)$ pätee

$$f'(a) - \frac{f(a+h) - f(a)}{h} = -\frac{1}{2} f''(\xi)\, h.$$

15. (*) Näytä Taylorin lauseen avulla, että jos f on neljä kertaa jatkuvasti derivoituva välillä $[a-h, a+h]$, niin pätee

$$\left| f''(a) - \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} \right| \le \frac{1}{12} \left(\max_{x\in[a-h,a+h]} |f^{(4)}(x)| \right) h^2.$$

Miten arvio toteutuu funktiolle $f(x) = (x-a)^4$?

16. (*) a) Näytä, että eräillä (millä?) A:n ja B:n arvoilla pätee

$$\operatorname{Arccos} x \;=\; A\sqrt{1-x} + B(1-x)^{3/2} + \mathcal{O}\left((1-x)^{5/2}\right), \quad \text{kun } x \to 1^-.$$

b) Näytä, että jos sykloidin yhtälöt $x = R(t - \sin t)$, $y = R(1 - \cos t)$ eliminoidaan muotoon $y = y(x)$ (ratkaisemalla $t = t(x)$ ensimmäisestä yhtälöstä), niin origon ympäristössä pätee

$$y(x) = 3\sqrt[3]{\frac{Rx^2}{6}} + \mathcal{O}\left(\sqrt[3]{\frac{x^4}{R}}\right).$$

17. (*) Olkoon c funktion f yksinkertainen nollakohta ja olkoon f kolme kertaa jatkuvasti derivoituva c:n ympäristössä. Halutaan määrätä c Newtonin menetelmästä muunnellulla iteraatiolla $x_{n+1} = F(x_n)$, missä

$$F(x) = x - \frac{f(x)}{f'(x)} + [f(x)]^2 g(x).$$

Miten $g(x)$ on valittava, jotta iteraation suppeneminen kohti c:tä on vähintään kuutiollista? Millä ehdolla suppeneminen on täsmälleen kuutiollista? Sovella menetelmää funktioon $f(x) = x^2 - a$, $a > 0$ (vrt. Harj.teht. V.7:17b).

VII.6 Interpolaatiopolynomit

Tavallisessa *polynomi-interpolaatiossa* on lähtöajatuksena, että funktiosta f tunnetaan vain äärellinen määrä pistearvoja:

$$f(x_i) = f_i, \quad i = 0 \dots n.$$

Näiden tietojen perusteella halutaan esittää f likimäärin polynomina muuallakin kuin pisteissä x_i, esim. jollakin välillä. Interpolaatio on kyseessä erityisesti silloin, kun $f(x)$ halutaan laskea 'välipisteissä' $x \in [\min\{x_i\}, \max\{x_i\}]$, muussa tapauksessa (eli kun $x < \min\{x_i\}$ tai $x > \max\{x_i\}$) sanotaan, että kyseessä on *ekstrapolaatio*. Annetut tiedot voivat olla esim. mitattua 'dataa'. Toinen yleinen sovellustilanne on sellainen, jossa funktio f tunnetaan epäsuorasti, esim. differentiaaliyhtälön ratkaisuna, ja halutaan laskea arvot f_i valituissa pisteissä x_i. Tällöin polynomi-interpolaatioista on hyötyä itse laskenta-algoritmin suunnittelussa.

Funktion polynomiapproksimaatiossa on yleensä perusoletuksena (tai ainakin toivomuksena), että funktio on riittävän säännöllinen, jolloin Taylorin lauseen perusteella tiedetään, että funktiota voi (ainakin lyhyellä välillä) approksimoida hyvin polynomilla, nimittäin Taylorin polynomilla. Koska itse Taylorin polynomia ei ym. tiedoista voi suoraan määrätä, on luontevaa valita approksimoivaksi polynomiksi p sellainen, joka sopii annettuihin tietoihin, eli toteuttaa

$$p(x_i) = f(x_i), \quad i = 0 \dots n. \tag{1}$$

Koska tässä on ehtoja $n+1$ kpl, on edelleen luonnollista valita p:n asteluvuksi n, jolloin p:ssä on vapaita kertoimia myöskin $n+1$ kpl. Näin määriteltyä polynomia p sanotaan f:n *Lagrangen interpolaatiopolynomiksi* pisteissä x_i. Pisteitä x_i sanotaan tässä yhteydessä *interpolaatiopisteiksi* ja ehtoja (1) *interpolaatioehdoiksi*. Nämä ehdot todella määrittelevät yksikäsitteisen polynomin astetta n:

Propositio VII.6.1 Lagrangen interpolaatiopolynomi on yksikäsitteinen.

<u>Todistus</u> Ehdot (1) toteuttavan interpolaatiopolynomin olemassaolo seuraa jäljempänä esitettävästä laskukaavasta (2), joten riittää näyttää yksikäsitteisyys. Olkoot siis p_1 ja p_2 molemmat ehdot (1) täyttäviä polynomeja astetta n. Tällöin $q(x) = p_1(x) - p_2(x)$ on polynomi enintään astetta n ja pätee

$$q(x_i) = 0, \quad i = 0 \dots n,$$

eli q:lla on $n+1$ reaalista nollakohtaa. Algebran peruslauseen mukaan on tällöin oltava $q = 0$. Siis $p_1 = p_2$, eli ehdot (1) täyttävä polynomi (sikäli kuin olemassa) on yksikäsitteinen. $\square$

Lagrangen interpolaatiopolynomin approksimaatiovirheelle pätee seuraava tulos, joka muistuttaa Taylorin polynomien virhekaavaa (Lause VII.4.3). Jäljempänä nähdäänkin (Lause VII.6.4), että nämä kaksi tulosta ovat erikoistapauksia yleisemmästä interpolaatiopolynomien virhekaavasta. Jatkossa sanotaan äärellisen pisteistön X *virittämäksi* väliksi suljettua väliä $[a, b]$, missä $a = \min\{x \mid x \in X\}$ ja $b = \max\{x \mid x \in X\}$.

Lause VII.6.2 Olkoon f jatkuva välillä $[a, b]$ ja $n + 1$ kertaa derivoituva välillä (a, b), missä $[a, b]$ on pisteen x ja erillisten pisteiden $x_0, \ldots, x_n$ virittämä väli, $n \in \mathbb{N}$. Tällöin jos p on f:n Lagrangen interpolaatiopolynomi pisteissä $x_0, \ldots, x_n$, niin jollakin $\xi \in (a, b)$ pätee virhekaava

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) \prod_{i=0}^{n} (x - x_i).$$

Todistus perustuu samaan ideaan kuin Taylorin lauseen todistus Luvussa VII.4. Ensinnäkin jos $x \in \{x_0, \ldots, x_n\}$, on väittämä tosi jokaisella $\xi \in (a, b)$, joten voidaan olettaa, että $x \notin \{x_0, \ldots, x_n\}$. Merkitään

$$w(t) = \prod_{i=0}^{n} (t - x_i)$$

ja tutkitaan funktiota

$$g(t) = f(t) - p(t) - Hw(t), \quad H = [f(x) - p(x)]/w(x).$$

Funktion g nollakohtia ovat interpolointipisteet $x_0, \ldots, x_n$ ja lisäksi piste x. Koska nollakohtia välillä $[a, b]$ on siis ainakin $n + 2$ kpl, ja koska näiden välissä on aina derivaatan nollakohta (Lause V.6.2), on $g'(t)$:llä ainakin $n + 1$ nollakohtaa avoimella välillä (a, b). Tällöin g'':lla on ainakin n nollakohtaa tällä välillä, ja lopulta $g^{(n+1)}$:lla ainakin yksi nollakohta $\xi \in (a, b)$. Mutta tällöin

$$\begin{aligned} 0 = g^{(n+1)}(\xi) &= f^{(n+1)}(\xi) - p^{(n+1)}(\xi) - Hw^{(n+1)}(\xi) \\ &= f^{(n+1)}(\xi) - H(n+1)! \end{aligned}$$

$$\Rightarrow \quad H = \frac{f(x) - p(x)}{w(x)} = \frac{1}{(n+1)!} f^{(n+1)}(\xi). \quad \square$$

Interpolaatioperiaatteista on astelukuun $n = 1$ perustuva *lineaarinen interpolaatio* yksinkertaisuutensa vuoksi hyvin yleisesti käytetty. Esimerkiksi käyrän $y = f(x)$ approksimointi pisteiden $(x_i, f(x_i))$ kautta kulkevalla murtoviivalla

(vaikkapa kuvan piirtämiseksi tai kaarenpituuden arvioimiseksi) tarkoittaa f:n *paloittain lineaarista* interpolaatiota.

Jos f on kahdesti jatkuvasti derivoituva interpolointivälillä $[x_0, x_1]$, niin lineaarisen interpolaation virhe ko. välillä on Lauseen VII.6.2 mukaan enintään

$$\begin{aligned}\max_{x\in[x_0,x_1]} |f(x) - p(x)| &\le \max_{x\in[x_0,x_1]} \left|\frac{1}{2}(x-x_0)(x-x_1)\right| \max_{\xi\in[x_0,x_1]} |f''(\xi)| \\ &= \frac{1}{8}\,(x_1-x_0)^2 \max_{x\in[x_0,x_1]} |f''(x)|.\end{aligned}$$

Jos interpolointi tapahtuu paloittain välillä $[a, b]$ ja interpolointipisteiden väli on enintään h, niin ko. approksimaation maksimivirhe on siis enintään

$$\max_{x\in[a,b]} |f(x) - \tilde{f}(x)| \le \frac{1}{8}\,h^2 \max_{x\in[a,b]} |f''(x)|.$$

Tämä arvio ei ole parannettavissa, sillä jos $f''(x) =$ vakio (eli f on toisen asteen polynomi) ja interpolointipisteet ovat tasaväliset, niin virhearvion yläraja totetuu peräkkäisten interpolointipisteiden puolivälissä.

Jos *kvadraattisessa* (= toisen asteen) interpolaatiossa pisteet x_0, x_1, x_2 ovat tasavälein ja väli $= h$, niin

$$\max_{x\in[x_0,x_2]} |(x-x_0)(x-x_1)(x-x_2)| = \max_{x\in[-h,h]} |x|(h^2-x^2) = \frac{2}{3\sqrt{3}}\,h^3,$$

joten kvadraattinen interpolaatiovirhe on tasavälisten interpolointipisteiden virittämällä välillä enintään

$$\max_{x\in[x_0,x_2]} |f(x) - p(x)| \le \frac{1}{9\sqrt{3}}\,h^3 \max_{x\in[x_0,x_2]} |f'''(x)| \quad (h = x_1 - x_0 = x_2 - x_1).$$

Esimerkki 1 Välillä $(0,\infty)$ määritelty funktio E toteuttaa ehdot

$$E'(x) = -\frac{e^{-x}}{x}, \quad x > 0, \qquad \lim_{x\to\infty} E(x) = 0.^{\dagger}$$

Halutaan määrittää E:n arvot likimäärin välillä $[1,2]$ laskemalla ensin E pisteissä $x_i = 1 + (i-1)h,\ i = 0\ldots n,\ h = 1/n$ (oletetetaan, että tämä onnistuu hyvin tarkasti) ja käyttämällä interpolaatiota muissa pisteissä. Kuinka suuri on n:n oltava, jos käytetään a) lineaarista, b) kvadraattista interpolaatiota ja halutaan, että interpolaatiovirhe on enintään $5\cdot 10^{-9}$?

<u>Ratkaisu</u> Koska $E'(x) = -e^{-x}/x,\ x > 0$, niin

$$\begin{aligned} E''(x) &= e^{-x}(\tfrac{1}{x} + \tfrac{1}{x^2}) \quad \Rightarrow \quad |E''(x)| \le 2/e, \quad x \in [1,2], \\ E'''(x) &= -e^{-x}(\tfrac{1}{x} + \tfrac{2}{x^2} + \tfrac{2}{x^3}) \quad \Rightarrow \quad |E'''(x)| \le 5/e, \quad x \in [1,2]. \end{aligned}$$

Näin ollen riittää valita $N = n+1$ siten, että pätee

$$\begin{aligned} &\text{a)} \quad \frac{1}{8}\cdot\frac{2}{e}\cdot\left(\frac{1}{N-1}\right)^2 \le 5\cdot 10^{-9} \quad \Leftrightarrow \quad \underline{\underline{N \ge 4290}}, \\ &\text{b)} \quad \frac{1}{9\sqrt{3}}\cdot\frac{5}{e}\cdot\left(\frac{1}{N-1}\right)^3 \le 5\cdot 10^{-9} \quad \Leftrightarrow \quad \underline{\underline{N \ge 288}}. \quad \square \end{aligned}$$

Lagrangen kantapolynomit

Korkeampiasteisissa polynomi-interpolaatioissa on usein kyse *polynomisovituksesta* pisteisiin $(x_i, f(x_i))$, jolloin arvot $f(x_i)$ voivat olla esimerkiksi mitattuja. Kehittyneissä numeerisen ja symbolisen analyysin ohjelmistoissa on tähän tehtävään omat komentonsa (esim. Mathematica: `Fit`, Matlab: `Polyfit`). Käsin laskettaessa, tai etenkin haluttaessa tutkia interpolaatiopolynomin ominaisuuksia teoreettiselta kannalta, voidaan käyttää esitysmuotoa

$$p(x) = \sum_{i=0}^{n} f(x_i)L_i(x), \tag{2}$$

missä nk. (Lagrangen) *kantapolynomit* $L_i(x)$ (astetta n) määräytyvät interpolaatioehdoista

$$L_i(x_j) = \begin{cases} 1, & \text{kun } j = i, \\ 0, & \text{kun } j \ne i \quad (j \in \{0,\ldots,n\}). \end{cases} \tag{3}$$

†Funktio on nimeltään *eksponentiaali-integraalifunktio.*

Helposti on tarkistettavissa, että nämä ehdot toteuttava polynomi on (vrt. kuvio)

$$L_i(x) = \frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)} = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)}.$$

Ehdoista (3) nähdään myös välittömästi, että polynomi (2) täyttää asetetut interpolaatioehdot (1). Lagrangen interpolaatio-ongelman ratkeavuus on siis näin tullut todetuksi.

Esimerkki 2 Säännöllisestä funktiosta tiedetään mittaustuloksina

$$f_1 = f(0.1) = 1.4491, \quad f_2 = f(0.2) = 1.4832, \quad f_3 = f(0.3) = 1.5166.$$

Arvioi $f(0)$.

Ratkaisu Sovitetaan mittaustuloksiin toisen asteen Lagrangen interpolaatiopolynomi p ja arvioidaan $f(0) \approx p(0)$. Tässä on $\{x_0, x_1, x_2\} = \{0.1, 0.2, 0.3\}$, joten

$$\begin{aligned} p(0) &= 1.4491 \cdot \frac{(0-0.2)\cdot(0-0.3)}{(0.1-0.2)\cdot(0.1-0.3)} \\ &\quad + 1.4832 \cdot \frac{(0-0.1)\cdot(0-0.3)}{(0.2-0.1)\cdot(0.2-0.3)} \\ &\quad + 1.5166 \cdot \frac{(0-0.1)\cdot(0-0.2)}{(0.3-0.1)\cdot(0.3-0.2)} \\ &= 3\cdot 1.4491 - 3\cdot 1.4832 + 1\cdot 1.5166 = \underline{\underline{1.4143}}. \end{aligned}$$

Tässä 'mittaustulokset' on itse asiassa saatu funktiosta $f(x) = \sqrt{2+x}$, jolle $f(0) = 1.4142$. $\square$

Kantapolynomeihin perustuva esitysmuoto (2) on erityisen kätevä silloin, kun halutaan arvioida funktion evaluoinnin virheiden (esim. mittausvirheiden) vaikutus interpolointitulokseen. Nimittäin jos virheellisten arvojen $\tilde{f}_i$ (oikea arvo $= f_i$) perusteella lasketaan interpolaatiopolynomi $\tilde{p}$, niin kaavan (2) mukaan

$$p(x) - \tilde{p}(x) = \sum_{i=0}^{n} (f_i - \tilde{f}_i) L_i(x).$$

Jos erityisesti tiedetään, että $|f_i - \tilde{f}_i| \leq \delta,\ i = 0, \ldots n$, niin

$$|p(x) - \tilde{p}(x)| \leq \delta \sum_{i=0}^{n} |L_i(x)| = \delta K(x).$$

Tässä määritelty virheen vahvistuskerroin $K(x) = \sum_{i=0}^{n} |L_i(x)|$ siis kertoo, kuinka paljon evaluointivirheet voivat pahimmillaan vahvistua pisteessä x.

Esimerkki 2 (jatko) Esimerkissä on $L_1(0) = 3$, $L_2(0) = -3$ ja $L_3(0) = 1$. Siis $K(0) = 3 + 3 + 1 = 7$, eli virheiden vaikutus pisteessä $x = 0$ on pahimmassa tapauksessa 7-kertainen yksittäisiin evaluointivirheisiin verrattuna. Jos oletetaan, että virheet ovat enintään 10^{-4} itseisarvoltaan, niin pahin vaihtoehto toteutuu, kun $f_1 - \tilde{f}_1 = -(f_2 - \tilde{f}_2) = f_3 - \tilde{f}_3 = \pm 10^{-4}$. $\square$

Ekstrapolaatio

Esimerkissä 2 funktiota approksimoitiin interpolaatiopisteiden virittämän välin ulkopuolella, jolloin sanotaan että kyse on *ekstrapolaatiosta*. Ekstrapolaatio on vanhastaan hyvin suosittu ja melko yleispätevä tapa parantaa numeeristen laskujen tarkkuutta. Ekstrapolaatiota voidaan käyttää aina, kun laskettavan suureen voidaan otaksua riippuva säännöllisellä (eli sileällä) tavalla jostakin laskentaan liittyvästä parametrista. Olkoon esimerkiksi laskettava suure reaaliluku a, joka määräytyy raja-arvona

$$a = \lim_{h \to 0^+} f(h),$$

missä jokainen $f(h)$, $h > 0$, on laskettavissa, mutta laskenta tulee yhä työläämmäksi h:n pienetessä. Jos nyt voidaan olettaa, että funktio $f(x)$ on säännöllinen jollakin välillä $[0, b]$, $b > 0$, voidaan numeerisen algoritmin antamia approksimaatioita

$$a \approx a_n = f(x_n), \quad n = 1, 2, \ldots \quad (x_n \to 0^+)$$

parantaa ekstrapolaatiolla. Näin syntyy nk. *ekstrapolaatiotaulukko*, jossa lasketuihin tuloksiin sovitetaan yhä korkeampiasteisia polynomeja $p(x)$, ja arvioidaan kunkin polynomin avulla $a \approx p(0)$:

	aste = 0	aste = 1	aste = 2	aste = 3
x_1	$f(x_1)$			
x_2	$f(x_2)$	$p^{(1,2)}(0)$		
x_3	$f(x_3)$	$p^{(2,3)}(0)$	$p^{(1,3)}(0)$	
x_4	$f(x_4)$	$p^{(3,4)}(0)$	$p^{(2,4)}(0)$	$p^{(1,4)}(0)$

Tässä $p^{(i,j)}(x)$ tarkoittaa pisteisiin $x_i \dots x_j$ sovitettua interpolaatiopolynomia astetta $j-i$ ($p^{(i,i)}(x) = f(x_i) =$ vakio). Osoittautuu, että taulukon sarakkeet määräytyvät palautuvasti edellisen sarakkeen avulla. Nimittäin

$$\boxed{p^{(i,j)}(x) = \frac{(x_j - x)p^{(i,j-1)}(x) + (x - x_i)p^{(i+1,j)}(x)}{x_j - x_i}\,.} \tag{4}$$

Tämä palautuskaava (perustelu induktiolla: Harj.teht. 10) helpottaa taulukon muodostamista huomattavasti.† Seuraavassa esimerkki kaavaan (4) perustuvasta 'laskemisen taiteesta'.

Esimerkki 3 *Stirlingin kaavan* mukaan $\phi(n) = n!/(\sqrt{2\pi n}\, e^{-n} n^n) \approx 1$ suurilla n:n arvoilla. Ekstrapoloi $\phi(100)$ arvoista $\phi(n)$, $n = 5 \dots 9$, kun tiedetään lisäksi, että $\phi(n) = f(1/n)$, missä $f(x)$ on sileä funktio välillä $[0, 1]$.

<u>Ratkaisu</u> Muodostetaan ekstrapolaatiotaulukko

x_i	$f(x_i)$	aste = 1	aste = 2	aste = 3	aste=4
1/5	1.0167..				
1/6	1.0139..	1.00076..			
1/7	1.0119..	1.00078..	1.000823..		
1/8	1.0104..	1.00079..	1.000827..	1.000833565..	
1/9	1.0092..	1.00080..	1.000830..	1.000833632..	1.000833708..

Oletettavasti neljännen asteen interpolaatio antaa tarkimman tuloksen, joten $\phi(100) \approx \underline{\underline{1.00083371}}$. (Oikea arvo on on 1.000833677..) □

Esimerkissä olennaista oli lisätieto, joka mahdollisti oikean muuttujan valinnan ($x_i = 1/n_i$) ekstrapoloinnissa.

†Kaavaan (4) perustuvaa ekstrapolaatiotaulukkoa sanotaan *Nevillen kaavioksi.* Ennen tietokoneiden aikaa tällaisilla (käsinlaskua helpottavilla) algoritmisilla keksinnöillä oli huomattava käytännön merkitys.

*Yleistetty polynomi-interpolaatio

Jos p on polynomi astetta n ja toteuttaa ehdot

$$\begin{aligned} p^{(k)}(x_i) = f^{(k)}(x_i), \quad k = 0 \ldots \nu_i - 1, \quad i = 1 \ldots m, \\ \text{missä} \quad \nu_i \in \mathbb{N} \quad \text{ja} \quad \sum_{i=1}^{m} \nu_i = n + 1, \end{aligned} \tag{5}$$

niin sanotaan, että p on funktion f *yleistetty interpolaatiopolynomi* pisteissä x_i. Tässä ja jatkossa oletetaan, että ehdoissa (5) esiintyvät f:n derivaatat (jos $\nu_i \geq 2$) ovat olemassa pisteissä x_i. Ehtojen (5) mukaisesti interpolaatiopisteiden lukumäärä m voi yleistetyssä polynomi-interpolaatiossa astetta n olla mikä tahansa välillä $1 \leq m \leq n + 1$. Tapauksessa $m = n + 1$ on oltava $\nu_i = 1$ $\forall i$, jolloin kyseessä on Lagrangen interpolaatio. Toisessa ääripäässä ($m = 1$) on taas oltava $\nu_1 = n + 1$, jolloin $p(x) = f$:n Taylorin polynomi $T_n(x, x_1)$.

Propositio VII.6.3 Jos p on polynomi astetta n, niin p määräytyy ehdoista (5) yksikäsitteisesti.

Todistus Interpolaatio-ongelman (5) ratkeavuus voidaan todeta samaan tapaan kuin Lagrangen interpolaation tapauksessa, ks. Harj.teht. 12. Yksikäsitteisyyden toteamiseksi riittää osoittaa, että jos $f^{(k)}(x_i) = 0$ $\forall i, k$, niin on oltava $p = 0$ (vrt. Proposition VII.6.1 todistus). Tässä tapauksessa on x_i polynomin p ν_i-kertainen nollakohta interpolaatioehtojen (5) mukaan, joten on oltava

$$p(x) = q(x) \prod_{i=1}^{m} (x - x_i)^{\nu_i} = q(x)\, w(x),$$

missä q on polynomi. Mutta p on astetta n, ja ehtojen (5) perusteella w on astetta $n + 1$, joten ainoa mahdollisuus on $q = 0$, jolloin myös $p = 0$. $\square$

Seuraava yleinen polynomiapproksimaatiotulos, jonka täydellistä todistusta ei esitetä, sisältää erikoistapauksina sekä Lauseen VII.6.2 että Taylorin lauseen VII.4.3.

Lause VII.6.4 **(Usean pisteen Taylorin lause)** Olkoon f jatkuva välillä $[a, b]$ ja $n + 1$ kertaa derivoituva välillä (a, b), missä $[a, b]$ on pisteen x ja erillisten pisteiden $x_0, \ldots, x_m$ virittämä väli. Tällöin jos p on ehdoilla (5) määritelty yleistetty interpolaatiopolynomi astetta n, niin jollakin $\xi \in (a, b)$ pätee virhekaava

$$f(x) - p(x) = \frac{1}{(n+1)!}\, f^{(n+1)}(\xi) \prod_{i=1}^{m} (x - x_i)^{\nu_i}.$$

Todistus (idea) Jos $x \in \{x_1, \ldots, x_m\}$, on väite tosi $\forall \xi$. Olkoon siis $x \neq x_i \ \forall i$, ja otetaan tarkastelun kohteeksi funktio

$$g(t) = f(t) - p(t) - Hw(t), \quad H = [f(x) - p(x)]/w(x), \quad w(x) = \prod_{i=1}^{m} (x - x_i)^{\nu_i}.$$

Koska $w(t) = t^{n+1} +$ (polynomi astetta n), niin $w^{(n+1)}(t) = (n+1)! \ \forall t$, joten virhekaava väittää, että $g^{(n+1)}(\xi) = 0$ jollakin $\xi \in (a, b)$. Aiemmin tämä on osoitettu tapauksissa $m = 1, \ \nu_1 = n + 1$ (Taylorin lause) ja $m = n + 1, \ \nu_i = 1$ (Lause VII.6.2). Yleisemmässäkin tapauksessa on päättely vastaava. Esimerkiksi olkoon

$$\nu_i = 2, \ i = 1 \ldots m, \quad m \geq 2, \quad n = 2m - 1.$$

Tällöin koska $g(x_i) = 0, \ i = 1 \ldots m$, ja $g(x) = 0$, on pisteiden x ja $x_i, \ i = 1 \ldots m$, välisillä avoimilla väleillä kullakin g':n nollakohta (yhteensä m kpl). Toisaalta on myös $g'(x_i) = 0, \ i = 1 \ldots m$, joten g':lla on välillä $[a, b]$ ainakin $2m$ nollakohtaa. Tästä seuraa (vrt. Lauseen VII.6.2 todistus), että ainakin yhdessä pisteessä $\xi \in (a, b)$ on $g^{(2m)}(\xi) = g^{(n+1)}(\xi) = 0$, eli väite on tosi oletetussa tapauksessa. (Yleinen tapaus sivuutetaan.) $\square$

Edellä tarkastellussa tapauksessa, jossa $\nu_i = 2, \ i = 1 \ldots m$, sanotaan polynomia p funktion f *Hermiten* interpolaatiopolynomiksi pisteissä x_i.

Esimerkki 1 (jatko) Montako tasavälistä jakopistettä x_i tarvitaan, jos pisteiden välissä käytetään kolmannen asteen Hermiten interpolaatiota?

Ratkaisu Lauseen VII.6.4 mukaan interpolaatiovirhe välillä $[x_i, x_{i+1}]$ on enintään

$$\begin{aligned} |E(x) - p(x)| &\leq \max_{x \in [x_i, x_{i+1}]} \frac{1}{4!} (x - x_i)^2 (x - x_{i+1})^2 \max_{\xi \in [x_i, x_{i+1}]} |E^{(4)}(\xi)| \\ &= \frac{1}{384} h^4 \max_{x \in [x_i, x_{i+1}]} |E^{(4)}(x)|. \end{aligned}$$

Välillä $[1, 2]$ on $|E^{(4)}(x)| \leq 16/e$, joten vaadittuun tarkkuuteen riittää:

$$\frac{1}{384} \cdot \frac{16}{e} \cdot \left(\frac{1}{N - 1} \right)^4 \leq 5 \cdot 10^{-9} \quad \Leftrightarrow \quad \underline{\underline{N \geq 43}}. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Funktiosta $f(x)$ tiedetään: $f(-0.1) = 1.70 \pm 0.05$, $f(0.2) = 1.80 \pm 0.03$ ja $-1 \leq f''(x) \leq 0$ välillä $[-0.1, 0.2]$. Määritä mahdollisimman ahdas väli, jolla $f(0)$ varmasti sijaitsee.

2. Funktion e^x arvot on laskettu viiden merkitsevän numeron tarkkuudella (normaalipyöristys) välin $[0,2]$ pisteissä $x_i = i/100,\ i = 0\ldots 200$. Arvioi, kuinka suuri on näistä arvoista lasketun a) lineaarisen, b) kvadraattisen interpolaation virhe enintään välillä $[0,2]$. Arvioi erikseen pyöristysvirheiden vaikutus.

3. Eräästä välillä $(0,\infty)$ määritellystä, säännöllisestä funktiosta F tiedetään, että F saavuttaa absoluuttisen minimiarvonsa välillä $[1,2]$. Lisäksi tiedetään, että $F(0.5) = \sqrt{\pi}$, $F(1) = 1$, $F(1.5) = \frac{1}{2}\sqrt{\pi}$ ja $F(2) = 1$. Laske F:n minimikohta ja -arvo likimäärin käyttäen a) kvadraattista interpolaatiota pisteissä $\{0.5, 1, 1.5\}$, b) kvadraaattista interpolaatiota pisteissä $\{1, 1.5, 2\}$, c) kolmannen asteen interpolaatiota kaikissa neljässä pisteessä. (Oletukset täyttää Γ-funktio, ks. Harj.teht. VIII.7:5 ja Propositio XIII.4.1.)

4. Parametrista käyrää $\vec{r} = \vec{u}(t) = x(t)\vec{i} + y(t)\vec{j},\ t \in [a,b]$, approksimoidaan pisteiden $(x(t_i), y(t_i)),\ i = 0\ldots n$ kautta kulkevalla murtoviivalla $\vec{r} = \vec{v}(t) = \hat{x}(t)\vec{i} + \hat{y}(t)\vec{j}$ $(a = t_0 < t_1 < \cdots < t_n = b)$. Jos $x(t)$ ja $y(t)$ ovat välillä $[a,b]$ kahdesti jatkuvasti derivoituvia, $|x''(t)| \le M$, $|y''(t)| \le M$ ja $t_i - t_{i-1} \le h$, niin kuinka suuri on enintään $\delta_h = \max_{t\in[a,b]} |\vec{u}(t) - \vec{v}(t)|$? Vertaa arviota todellisuuteen tapauksessa $x(t) = R\cos t,\ y(t) = R\sin t,\ t \in [0,\pi]$.

5. Heilurilla varustettu seinäkello jätättää vuorokaudessa 5 min 24 s. Kelloa rukataan kiertämällä heilurin päässä olevaa ruuvia kiinni 5 täyttä kierrosta (jolloin heilurin varsi hieman lyhenee). Rukkauksen jälkeen havaitaan kellon edistävän 3 min 36 s vuorokaudessa. a) Miten kelloa kannattaa seuraavaksi rukata? b) Olkoon kellon suhteellinen edistämä vuorokaudessa $f(x)$, missä $x =$ ruuvin kiertymä (kierroksina) oikeasta säätöasennosta kiinni päin. Arvioi, montako sekuntia kello edistää tai jätättää vuorokaudessa a-kohdan rukkauksen jälkeen, jos oletetaan, että $f''(0) = +2\cdot 10^{-5}$.

6. Ekstrapoloi luvuista $7!\ldots 10!$ raja-arvo $a = \lim_{n\to\infty} n!e^n n^{-n-\frac{1}{2}}$ ja vertaa tarkkaan arvoon $a = \sqrt{2\pi}$. (Vrt. Esimerkki 3.)

7. Todista Lause VII.6.4 tapauksessa $m = n = 3,\ \nu_1 = \nu_3 = 1,\ \nu_2 = 2$.

8. Funktioiden $\sin x$ ja $\cos x$ arvot halutaan määrätä välillä $[0,\pi/4]$ siten, että funktioiden arvot lasketaan ensin riittävän tarkasti ko. välin tasavälisessä pisteistössä ja sen jälkeen käytetään kolmannen asteen Hermiten interpolaatiota pisteiden välillä. Montako interpolointipistettä tarvitaan, jos virheen sallitaan olevan enintään $5\cdot 10^{-9}$?

9. Halutaan laskea numeerisesti funktion f derivaatta $f^{(k)}(a)$ approksimaatiolla $f^{(k)}(a) \approx p^{(k)}(a)$, missä $p(x)$ on f:n interpolaatiopolynomi. Näytä, että seuraavissa tapauksissa vaihtoehtoiset interpolaatiot johtavat samaan differenssiapproksimaatioon — millaiseen? (Vrt. edellinen luku.)

   a) $k = 1$: Lineaarinen interpolaatio pisteissä $a \pm h$ tai kvadraattinen interpolaatio pisteissä a ja $a \pm h$.
   b) $k = 2$: Kvadraattinen interpolaatio pisteissä a ja $a \pm h$ tai yleistetty kolmannen asteen interpolaatio samoissa pisteissä lisäehdolla $p'(a) = f'(a)$.

10. (*) Todista palautuskaava (4) induktiolla.

11. (*) (Rajatieto) Lukujonosta $\{a_n\}$, missä $a_n = \sum_{k=1}^{n} k^{-5/4}$, $n = 1, 2, \ldots$ tiedetään, että on olemassa (tuntemattomat) $a, b, c \in \mathbb{R}$ siten, että suurilla n:n arvoilla pätee

$$a_n = a + n^{-1/4}(b + cn^{-1}) + \mathcal{O}(n^{-9/4}).$$

   Ekstrapoloi raja-arvo $\lim_n a_n = a$ tämän tiedon perusteella mahdollisimman tarkasti tuloksista $a_{100} = 3.331779$, $a_{200} = 3.532117$, $a_{400} = 3.700964$. (Tarkka arvo kuudella desimaalilla: $a = 4.595112$.)

12. (*) Näytä, että jos interpolaatio-ongelmassa (5) on $f^{(k)}(x_i) = 1$ kun $i = j$ ja $k = l$ $(1 \le j \le m,\ 0 \le l \le \nu_j - 1)$ ja $f^{(k)}(x_i) = 0$ muulloin, niin ongelmalla on ratkaisu

$$p(x) = L_{j,l}(x) = q(x) \prod_{\substack{i=1 \\ i \neq j}}^{m} (x - x_i)^{\nu_i},$$

   missä q on polynomi muotoa $q(x) = \sum_{r=l}^{\nu_j - 1} c_r (x - x_j)^r$. Päättele edelleen, että interpolaatio-ongelman (5) ratkaisu yleisessä tapauksessa on

$$p(x) = \sum_{i=1}^{m} \sum_{k=0}^{\nu_i - 1} f^{(k)}(x_i) L_{i,k}(x).$$

Luku VIII

Integraali

Yhden reaalimuutujan analyysissä on *integraalin* käsitteellä kaksi olomuotoa: *määräämätön* ja *määrätty* integraali. Edellisellä tarkoitetaan annetun funktion *integraalifunktiota*, eli kyse on derivoinnin käänteisoperaatiosta, johon on jo alustavasti tutustuttu Luvussa V.6. Tässä luvussa, tarkemmin luvuissa VIII.1–VIII.3, esitellään aiempaa systemaattisemmin integraalifunktioiden etsimisessä käytetyt menetelmät ja funktiotyypit, jotka näillä menetelmillä ovat hallittavissa. Kyseessä on hyvin perinteinen matematiikan taitolaji, jossa keskeistä roolia näyttelevät erilaiset funktioalgebran keinot, kuten *osittaisintegrointi*, muuttujan vaihto eli *sijoitus* ja *osamurtokehitelmät*.

Luvuissa VIII.4–VIII.6 tarkastellaan määrätyn integraalin, tarkemmin *Riemann-integraalin*, eri määrittelytapoja ja ominaisuuksia sekä liittymistä määräämättömään integraaliin *Analyysin peruslauseen* kautta. Luvussa VIII.7 tarkastellaan Riemannin integraalin käsitteen laajennuksia, *epäoleellisia* integraaleja, ja näihin liittyen integraalien ja sarjojen vertailua. Kahdessa viimeisessä osaluvussa tarkastelun kohteena ovat määrättyyn integraaliin perustuvat pinta-alan ja kaarenpituuden yksinkertaisimmat laskukaavat sovelluksineen (Luku VIII.8) ja lopuksi *numeerisen integroinnin* menetelmät (Luku VIII.9).

VIII.1 Integraalifunktio

Palautettakoon mieliin Luvusta V.6, että funktiota $F(x)$ sanotaan funktion $f(x)$ *integraalifunktioksi* avoimella välillä (a, b), jos F on derivoituva välillä (a, b) ja

$$F'(x) = f(x), \quad x \in (a, b).$$

Integraalifunktiota merkitään

$$F(x) = \int f(x)\, dx.$$

Hieman erikoinen merkintä viittaa integraalifunktion laskennalliseen määritelmään, joka esitetään myöhemmin (Luku VIII.4). Integraalifunktiota sanotaan myös *määräämättömäksi integraaliksi* (engl. indefinite integral, antiderivative). Tässä 'määräämätön' viittaa siihen, että integraalifunktio ei ole yksikäsitteinen, vaan siihen voidaan lisätä mielivaltainen vakio: Jos yksi integraalifunktio F tunnetaan, niin kaikki integraalifunktiot ovat esitettävissä nk. *integroimisvakion* C avulla muodossa

$$\int f(x)\, dx = F(x) + C.$$

Määrämätön integraali on siis itse asiassa funktiojoukko. Kyseessä on differentiaaliyhtälön $y' = f(x)$ yleinen ratkaisu välillä (a, b) (ks. Korollaari V.6.8).

Esimerkki 1 Määritä funktion

$$f(x) = \begin{cases} 1, & \text{kun } x \le 2, \\ \dfrac{10}{x} - 2x, & \text{kun } x > 2 \end{cases}$$

integraalifunktio $\mathbb{R}$:ssä.

<u>Ratkaisu</u> Koska $f(2^-) = f(2^+) = 1 = f(2)$, niin f on jatkuva $\mathbb{R}$:ssä. Tunnettujen derivoimissääntöjen (Luku V.3) perusteella on oltava

$$F(x) = \begin{cases} x + C_1\,, & \text{kun } x \in (-\infty, 2), \\ 10 \ln x - x^2 + C_2\,, & \text{kun } x \in (2, \infty), \end{cases}$$

jolloin on $F'(x) = f(x)$, kun $x \neq 2$. Jotta F olisi derivoituva myös pisteessä $x = 2$, on F:n oltava tässä pisteessä ainakin jatkuva, eli on oltava

$$2 + C_1 = 10 \ln 2 - 4 + C_2 \quad \Rightarrow \quad C_2 = C_1 - 10 \ln 2 + 6.$$

Kirjoittamalla $C_1 = C$ on saatu

$$F(x) = \begin{cases} x + C, & \text{kun } x \leq 2, \\ 10\ln(x/2) - x^2 + 6 + C, & \text{kun } x \geq 2. \end{cases}$$

Funktion F toispuoliset derivaatat pisteessä $x = 2$ ovat

$$D_-F(2) = f(2^-), \quad D_+F(2) = f(2^+).$$

Koska $f(2^-) = f(2^+)$, niin F on derivoituva myös pisteessä $x = 2$ ja $F'(2) = f(2)$. Integraalifunktio on siis löydetty. Kuvassa f ja F, kun $C = 0$. $\square$

'Integrointi', ymmärrettynä integraalifunktion etsimisenä, on siis derivoinnin käänteistoimi. Sekä yksittäisten funktioiden integroimiskaavat että integraalifunktioiden yleisemmät laskusäännöt ovatkin derivoimissääntöjä 'nurinpäin' luettuna.

Esimerkki 2 Koska (vrt. Luku VI.3)

$$D\ln(x + \sqrt{x^2+1}) = \frac{1}{\sqrt{x^2+1}}, \quad x \in \mathbb{R},$$

niin saadaan (koko $\mathbb{R}$:ssä pätevä) integroimissääntö

$$\int \frac{1}{\sqrt{x^2+1}}\,dx \;=\; \ln(x + \sqrt{x^2+1}) + C. \quad \square$$

Integrointimenetelmistä yksinkertaisin on 'taulukkomenetelmä' eli kaavakokoelma, joka perustuu suoraan tunnettuihin derivoimissääntöihin. Seuraavaan taulukkoon on koottu tällaisista integroimiskaavoista keskeisimmät (ml. esimerkin tulos). Nämä voidaan suoraan lukea Lukujen V.3, V.4, VI.2 ja VI.3 derivoimissäännöistä.

$$
\begin{aligned}
&(1) \quad \int x^\alpha \, dx &&= \frac{1}{\alpha+1} x^{\alpha+1} + C, \quad \alpha \in \mathbb{R},\ \alpha \neq -1.\\
&(2) \quad \int \frac{1}{x} \, dx &&= \ln|x| + C.\\
&(3) \quad \int \cos x \, dx &&= \sin x + C.\\
&(4) \quad \int \sin x \, dx &&= -\cos x + C.\\
&(5) \quad \int e^x \, dx &&= e^x + C.\\
&(6) \quad \int \frac{1}{\cos^2 x} \, dx &&= \tan x + C.\\
&(7) \quad \int \frac{1}{\sin^2 x} \, dx &&= -\cot x + C.\\
&(8) \quad \int \frac{1}{\sqrt{1-x^2}} \, dx &&= \operatorname{Arcsin} x + C.\\
&(9) \quad \int \frac{1}{\sqrt{x^2+1}} \, dx &&= \ln(x + \sqrt{x^2+1}) + C.\\
&(10) \quad \int \frac{1}{\sqrt{x^2-1}} \, dx &&= \ln|x + \sqrt{x^2-1}| + C.\\
&(11) \quad \int \frac{1}{1+x^2} \, dx &&= \operatorname{Arctan} x + C.\\
&(12) \quad \int \frac{1}{1-x^2} \, dx &&= \frac{1}{2} \ln\left|\frac{1+x}{1-x}\right| + C.
\end{aligned}
$$

Taulukkoa voitaisiin helposti jatkaa tunnettujen derivoimissääntöjen perusteella. Esimerkiksi pätee (ks. Luku VI.2) :

$$
\begin{aligned}
&(13) \quad \int \frac{1}{\sin x} \, dx = \ln\left|\frac{1-\cos x}{\sin x}\right| + C.\\
&(14) \quad \int \frac{1}{\cos x} \, dx = -\ln\left|\frac{1-\sin x}{\cos x}\right| + C.
\end{aligned}
$$

Sääntöjä (1)–(14) voidaan tarkemmin käyttää sellaisilla <u>avoimilla</u> <u>väleillä</u>, joilla annettu funktio f ja integraalifunktio F ovat molemmat määriteltyjä ja jälkimmäinen on derivoituva. Esimerkiksi säännöt (3)–(5) ovat päteviä koko $\mathbb{R}$:ssä,

sääntö (2) on pätevä väleillä $(-\infty, 0)$ ja $(0, \infty)$, sääntö (8) välillä $(-1, 1)$, ja sääntö (7) väleillä $(k\pi, (k+1)\pi)$, $k \in \mathbb{Z}$. Silloin kun integroimissäännön pätevyys on rajoitettu useammalle pistevieraalle $\mathbb{R}$:n osavälille (kuten sännöissä (2) ja (7)), ei integroimisvakion tarvitse olla eri osaväleillä sama.

Taulukkokaavat (9), (10) ja (12) voidaan ilmaista myös hyperbolisten käänteisfunktioiden arsinh, Arcosh ja artanh avulla, sillä pätee (ks. Luku VI.3)

$$\begin{aligned}
\ln(x + \sqrt{x^2+1}) &= \operatorname{arsinh} x, \quad x \in \mathbb{R}, \\
\ln|x + \sqrt{x^2-1}| &= \begin{cases} \operatorname{Arcosh} x, & \text{kun } x > 1, \\ -\operatorname{Arcosh}|x|, & \text{kun } x < -1, \end{cases} \\
\frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| &= \begin{cases} \operatorname{artanh} x, & \text{kun } |x| < 1, \\ \operatorname{artanh}(1/x), & \text{kun } |x| > 1. \end{cases}
\end{aligned}$$

Integraalifunktion etsiminen — silloin kun funktio ei löydy suoraan yksinkertaisimpien derivoimissääntöjen perusteella — on taitolaji, jota perinteisessä matematiikassa on harjoiteltu varsin paljon. Lajilla on urheilullista mielenkiintoa mm. siksi, että monien varsin yksinkertaisten funktioiden, kuten

$$\frac{1}{x^4+1}, \quad \frac{\sqrt{x^2+1}}{x}$$

integraalifunktioiden määrittäminen on työlästä. On myös paljon funktioita, esim.

$$e^{x^2}, \quad \cos x^2, \quad \frac{e^x}{x}, \quad \frac{\sin x}{x}, \quad \frac{1}{\ln x}, \quad \sqrt{x}\sin x, \quad \sqrt{x^4+1}, \quad \sqrt{1+\sin^2 x},$$

joiden integraalifunktiot eivät ole nk. *alkeisfunktioita*, eli ne eivät ole polynomien tai rationaalifunktioiden, juurilausekkeiden, eksponentti- tai logaritmifunktioiden tai trigonometristen funktioiden tai niiden käänteisfuntioiden avulla ilmaistavissa olevia lausekkeita. Sanotaan tällöin, että funktio ei ole integroitavissa *suljetussa muodossa* (engl. in closed form). Tällaisista integraalifunktioista on ennen puhuttu kunnioittavaan sävyyn, kutsuen niitä mm. 'korkeammiksi transkendenttifunktioiksi'. Nykyisin laskimet ja tietokoneet, ja niiden myötä numeerisen laskennan helppous, ovat pudottaneet tällaisetkin funktiot tavallisten kuolevaisten joukkoon.

Tietokoneiden myötä on tullut myös symbolisen laskennan mahdollisuus, mikä on vähentänyt perinteisen käsityötaidon merkitystä integraalifunktioiden etsinnässä. Tavallisimmat integroimissäännöt ja -menetelmät on silti edelleen syytä tuntea, mm. koska tällä laskutekniikalla on yleisempääkin käyttöä. Tässä ja kahdessa seuraavassa luvussa käydään läpi integroimistekniikan yleisimmät menetelmät.

Kolme yleistä integroimissääntöä

Yleisiä derivoimissääntöjä 'nurinpäin' lukemalla saadaan yleisiä integroimissääntöjä. Seuraavista kolmesta säännöstä ensimäinen kertoo, että integrointi on derivoinnin tavoin <u>lineaarinen</u> toimitus. Toinen ja kolmas sääntö ovat yhdistetyn funktion derivoimissääntöön suoraan perustuvia.

$$\int f(x)\,dx = F(x) + C \quad \& \quad \int g(x)\,dx = G(x) + C$$

$$\Rightarrow \quad \begin{cases} \int [\alpha f(x) + \beta g(x)]\,dx = \alpha F(x) + \beta G(x) + C, \quad \alpha, \beta \in \mathbb{R}, & \text{(I-1)} \\ \int f(ax+b)\,dx = \dfrac{1}{a} F(ax+b) + C, \quad a, b \in \mathbb{R},\ a \neq 0, & \text{(I-2)} \\ \int f(g(x))g'(x)\,dx = F(g(x)) + C. & \text{(I-3)} \end{cases}$$

Näissä säännöissä (kuten säännöissä (1)–(14)) on huomioitava, että integroimisvakio C on nk. *geneerinen* vakio, joka voi eri yhteyksissä (kuten lausekkeissa $F(x) + C$ ja $G(x) + C$) aina saada erilaisia arvoja, ellei toisin sovita.

Esimerkki 3 $\displaystyle\int \frac{1}{x^2 + x}\,dx = ?$

<u>Ratkaisu</u> Hajotetaan integroitava funktio ensin kahden yksinkertaisemman funktion summaksi:

$$\frac{1}{x^2+x} = \frac{1}{x(x+1)} = \frac{A}{x} + \frac{B}{x+1} = \frac{(A+B)x + A}{x(x+1)}$$

$$\Rightarrow \quad \begin{cases} A + B = 0 \\ A = 1 \end{cases} \quad \Leftrightarrow \quad \begin{cases} A = 1 \\ B = -1 \end{cases} \quad \Rightarrow \quad \frac{1}{x^2+x} = \frac{1}{x} - \frac{1}{x+1}$$

$$\Rightarrow \quad \int \frac{1}{x^2+x}\,dx \overset{\text{(I-1)}}{=} \int \frac{1}{x}\,dx - \int \frac{1}{x+1}\,dx$$

$$\overset{(2),\,\text{(I-2)}}{=} \ln|x| - \ln|x+1| + C = \underline{\underline{\ln\left|\frac{x}{x+1}\right| + C.}} \quad \square$$

Esimerkki 4 Johda taulukkokaava (12).

<u>Ratkaisu</u> Edellistä esimerkkiä mukaillen lasketaan

$$\frac{1}{1-x^2} = \frac{1}{(1+x)(1-x)} = \frac{1}{2}\frac{1}{1+x} + \frac{1}{2}\frac{1}{1-x}$$

$$\Rightarrow \quad \int \frac{1}{1-x^2}\,dx = \frac{1}{2}\ln|1+x| - \frac{1}{2}\ln|1-x| + C = \frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| + C. \quad \square$$

Esimerkeissä integraalifunktio löydettiin nk. *osamurtokehitelmän* avulla. Tämä on yleisempiinkin rationaalifunktiohin soveltuva menetelmä, kuten nähdään jäljempänä Luvussa VIII.3.

Esimerkki 5 Säännön (I-3), taulukkokaavan (1) ja derivoimissääntöjen $\mathrm{D}\ln|x| = 1/x$ ja $\mathrm{D}\sin x = \cos x$ perusteella

$$\int \frac{1}{x}\,\ln|x|\,dx = \frac{1}{2}\,(\ln|x|)^2 + C,$$
$$\int \sin^4 x\cos x\,dx = \frac{1}{5}\,\sin^5 x + C. \quad \square$$

Esimerkki 6 $\displaystyle\int \frac{1}{x^2+x+1}\,dx = ?$

<u>Ratkaisu</u> Koska

$$\frac{1}{x^2+x+1} = \frac{1}{(x+\frac{1}{2})^2+\frac{3}{4}} = \frac{4}{3}\,\frac{1}{(\frac{2x+1}{\sqrt{3}})^2+1},$$

niin

$$\begin{aligned}\int \frac{1}{x^2+x+1}\,dx &\overset{\text{(I-1)}}{=} \frac{4}{3}\int \frac{1}{(\frac{2x+1}{\sqrt{3}})^2+1}\,dx \\ &\overset{\text{(I-2),(11)}}{=} \frac{4}{3}\cdot\frac{\sqrt{3}}{2}\,\mathrm{Arctan}\left(\frac{2x+1}{\sqrt{3}}\right)+C \\ &= \underline{\underline{\frac{2}{\sqrt{3}}\,\mathrm{Arctan}\left(\frac{2x+1}{\sqrt{3}}\right)+C.}} \quad \square\end{aligned}$$

Esimerkki 7 a) $\displaystyle\int \cos^2 x\,dx = ?$ b) $\displaystyle\int \cos^3 x\,dx = ?$

<u>Ratkaisu</u> Koska

$$\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos 2x, \quad \cos^3 x = (1-\sin^2 x)\cos x,$$

niin

$$\text{a)}\ \int \cos^2 x\,dx = \frac{1}{2}\,x + \frac{1}{4}\sin 2x + C = \underline{\underline{\frac{1}{2}\,(x+\cos x\sin x)+C,}}$$
$$\text{b)}\ \int \cos^3 x\,dx = \int \left(\cos x - \sin^2 x\cos x\right)\,dx = \underline{\underline{\sin x - \frac{1}{3}\sin^3 x + C.}} \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Määritä se funktion $2x-3$ integraalifunktio, jonka kuvaaja sivuaa suoraa $x+y=0$.

2. Ratkaise seuraavat alkuarvotehtävät välillä $(-\infty,\infty)$:
   a) $F'(x)=2x^2-|x|,\ F(0)=1.$
   b) $F'(x)=|x^2+4x|,\ F(-4)=0.$
   c) $F'(x)=\max\{x,\,8x-x^2\},\ F(4)=-1.$
   d) $F'(x)=\min\{4-4x+x^2,\,40+2x-x^2\},\ F(0)=180.$

3. Määritä seuraavien funktioiden integraalifunktiot taulukkokaavoihin (1)–(12) ja sääntöihin (I-1)–(I-3) vedoten. Muunna funktio tarvittaessa ensin integroinnin kannalta soveliaampaan muotoon.

   a) $(2x^2-5x+7)(4x-4)$ b) $\dfrac{1}{\sqrt{2x+3}}$ c) $\dfrac{3}{3-\pi x}$ d) $\dfrac{x^2}{4x^3-1}$

   e) $\dfrac{2x^3+x}{(x^4+x^2+1)^3}$ f) $\dfrac{x}{(x^2+1)\sqrt{x^2+1}}$ g) $\dfrac{1}{x\ln|x|}$ h) $\dfrac{e^x}{e^x+1}$

   i) $\dfrac{1}{x(\ln|x|)^2}$ j) $\dfrac{1}{\sqrt{x^2-3}}$ k) $\dfrac{1}{\sqrt{x^2+3}}$ l) $\dfrac{x+2}{\sqrt{x^2-1}}$

   m) $\dfrac{3x-2}{\sqrt{2x^2+5}}$ n) $\dfrac{1}{x^2+14x+50}$ o) $\dfrac{1}{x^2+14x+48}$

   p) $\dfrac{1}{x^2+2x+10}$ q) $\dfrac{1}{x^2+3x-10}$ r) $\dfrac{2x+1}{x^2+3x-10}$

   s) $\sin 2x\,\cos^4 x$ t) $\dfrac{\sin x}{\cos^3 x}$ u) $\tan x$ v) $\tan^2 x$ x) $\tan^3 x$

   y) $\dfrac{e^{\tan x}}{\cos^2 x}$ z) $\sin^5 x$ å) $\cosh^2 x$ ä) $\tanh x$ ö) $\dfrac{\operatorname{arsinh} x}{\sqrt{x^2+1}}$

4. (*) Määritä $a\in\mathbb{R}$ ja $\mathbb{R}$:ssä jaksollinen funktio u siten, että $y(x)=ax+u(x)$ on alkuarvotehtävän $y'=|\sin x|,\ y(0)=0$ ratkaisu.

5. (*) Olkoon $f(0)=0$ ja f:n määritelmä muualla kuin origossa

   a) $f(x)=\dfrac{x}{\sqrt{|x|}}\,,$ b) $f(x)=\dfrac{x}{|x|}\,,$ c) $f(x)=\dfrac{1}{x}\left(\sin\dfrac{1}{x^2}+x^2\cos\dfrac{1}{x^2}\right).$

   Konstruoi f:n integraalifunktio välillä $(-1,1)$ tai näytä, että ko. välillä f:llä ei ole integraalifunktiota.

VIII.2 Osittaisintegrointi ja sijoitus

Integroimistekniikan, eli integraalifunktioiden etsimisen, kaksi keskeisintä yleistä metodia ovat *osittaisintegrointi* (engl. integration by parts, partial integration) ja *sijoitus*(menettely) (engl. substitution). Näillä menetelmillä on 'matematiikan kaavoina' yleisempääkin käyttöä.

Osittaisintegroinnin kaava on yksinkertaisesti tulon derivoimissääntö toisin kirjoitettuna:

$$\frac{d}{dx}(fg) = f'g + fg' \quad \Leftrightarrow \quad f'g = \frac{d}{dx}(fg) - fg'$$

$$\Rightarrow \quad \boxed{\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx.}$$

Esimerkki 1 a) $\displaystyle\int xe^x\,dx = ?$ b) $\displaystyle\int e^x \sin x\,dx = ?$

<u>Ratkaisu</u>

a) $$\int xe^x\,dx = \int \underbrace{e^x}_{f'(x)}\underbrace{x}_{g(x)}\,dx = e^x \cdot x - \int e^x\,dx = \underline{\underline{(x-1)e^x + C}}.$$

<u>Tarkistus</u>: $\displaystyle\frac{d}{dx}((x-1)e^x) = e^x + (x-1)e^x = xe^x.$ OK! $\square$

b) Valitaan $f'(x) = f(x) = e^x$ ja integroidaan kahdesti osittain:

$$\begin{aligned} F(x) &= \int e^x \sin x\,dx = e^x \sin x - \int e^x \cos x\,dx \\ &= e^x \sin x - e^x \cos x - \int e^x \sin x\,dx \\ &= e^x(\sin x - \cos x) - F(x) \end{aligned}$$

$$\Rightarrow \quad F(x) = \underline{\underline{\frac{1}{2}e^x(\sin x - \cos x) + C}}. \quad \square$$

Esimerkki 2 Määritä integraalifunktio $\displaystyle\int \sqrt{a^2+x^2}\,dx$, kun $a > 0$.

<u>Ratkaisu</u> Osittain integroimalla $(f'(x) = 1,\ g(x) = \sqrt{a^2+x^2}\,)$ saadaan ensin

$$\int \sqrt{a^2+x^2}\,dx = x\sqrt{a^2+x^2} - \int \frac{x^2}{\sqrt{a^2+x^2}}\,dx.$$

Tässä on

$$\begin{aligned} -\int \frac{x^2}{\sqrt{a^2+x^2}}\,dx &= -\int \frac{(x^2+a^2)-a^2}{\sqrt{a^2+x^2}}\,dx \\ &= -\int \sqrt{a^2+x^2}\,dx + \int \frac{a^2}{\sqrt{a^2+x^2}}\,dx, \end{aligned}$$

joten seuraa (ks. edellisen luvun kaavat (9) ja (I-2))

$$\begin{aligned} 2\int \sqrt{a^2+x^2}\,dx &= x\sqrt{a^2+x^2} + \int \frac{a}{\sqrt{(x/a)^2+1}}\,dx \\ &= x\sqrt{a^2+x^2} + a^2 \ln\bigl[(x/a) + \sqrt{(x/a)^2+1}\,\bigr] + C. \end{aligned}$$

Kun tässä kirjoitettaan vakion C tilalle $a^2 \ln a + C$ (mahdollista, koska C on joka tapauksessa määräämätön), saadaan tulos muotoon

$$\int \sqrt{a^2+x^2}\,dx = \underline{\underline{\frac{1}{2}\bigl[x\sqrt{a^2+x^2} + a^2 \ln\bigl(x+\sqrt{a^2+x^2}\bigr)\bigr] + C.}} \quad \square$$

Reduktiokaavat

Osittain integroimalla voidaan johtaa *reduktiokaavoja* (palautuskaavoja) koko joukolle integraaleja, joissa on kokonaislukuparametri. Esimerkiksi seuraavat integraalit saadaan tällä tavoin lasketuksi suljetussa muodossa:

$$\int x^n e^x\,dx, \quad \int x^n \cos x\,dx, \quad x^n \sin x\,dx, \quad \int (1+x^2)^{-n}\,dx, \quad n \in \mathbb{N},$$
$$\int \cos^n x\,dx, \quad \int \sin^n x\,dx, \quad n \in \mathbb{Z}.$$

Tarkastellaan esimerkkinä integraalia

$$I_n(x) = \int \cos^n x\,dx.$$

Osittain integroimalla saadaan

$$\begin{aligned} I_n(x) &= \int \cos x \cdot \cos^{n-1} x\,dx \qquad [\,f'(x) = \cos x,\ g(x) = \cos^{n-1} x\,] \\ &= \sin x \cos^{n-1} x + (n-1)\int \sin^2 x \cos^{n-2} x\,dx \\ &= \sin x \cos^{n-1} x + (n-1)\int (1-\cos^2 x)\cos^{n-2} x\,dx \\ &= \sin x \cos^{n-1} x + (n-1)I_{n-2}(x) - (n-1)I_n(x). \end{aligned}$$

Tämän perusteella integraalille $I_n(x)$ pätee reduktiokaava

$$nI_n(x) - (n-1)I_{n-2}(x) = \sin x \cos^{n-1} x. \tag{1}$$

Kaava pätee itse asiassa kun $n \hookrightarrow \alpha$, $\alpha \in \mathbb{R}$, mutta siitä on hyötyä lähinnä kun $n \in \mathbb{Z}$, jolloin I_n voidaan palauttaa kaavan avulla tapauksiin $n = 0, \pm 1$ ja näin ollen integroida suljetussa muodossa (tapaukseen $n = -1$ soveltuu edellisen luvun taulukkokaava (14)).

Esimerkki 3 a) $\displaystyle\int \cos^3 x\,dx = ?$ b) $\displaystyle\int \cos^{-3} x\,dx = ?$

<u>Ratkaisu</u> Indekseillä $n = 3$ ja $n = -1$ reduktioaava (1) antaa

$$\begin{aligned} I_3(x) &= \frac{1}{3}\sin x \cos^2 x + \frac{2}{3} I_1(x), \\ I_{-3}(x) &= \frac{1}{2}\sin x \cos^{-2} x + \frac{1}{2} I_{-1}(x), \end{aligned}$$

joten (vrt. edellisen luvun Esimerkki 7)

$$\begin{aligned} \text{a)} \quad \int \cos^3 x\,dx &= \frac{1}{3}\sin x \cos^2 x + \frac{2}{3}\int \cos x\,dx \\ &= \underline{\underline{-\frac{1}{3}\sin^3 x + \sin x + C}} \end{aligned}$$

ja edellisen luvun kaavan (14) perusteella

$$\begin{aligned} \text{b)} \quad \int \cos^{-3} x\,dx &= \frac{1}{2}\sin x \cos^{-2} x + \frac{1}{2}\int \cos^{-1} x\,dx \\ &= \underline{\underline{\frac{1}{2}\ln\left|\frac{\cos x}{1-\sin x}\right| + \frac{1}{2}\,\frac{\sin x}{\cos^2 x} + C.}} \quad \square \end{aligned}$$

Toisena esimerkkinä tarkasteltakoon integraalia

$$I_n(x) = \int \frac{1}{(1+x^2)^n}\,dx, \quad n \geq 2.$$

Pyritään palauttamaan tämä tapaukseen $n = 1$, jossa integraalifunktio tunnetaan (edellisen luvun taulukkokaava (11)). Valitaan $f'(x) = 1$, $g(x) = (1+x^2)^{-n}$ ja integroidaan osittain:

$$\begin{aligned} I_n(x) &= \frac{x}{(1+x^2)^n} + 2n\int \frac{x^2}{(1+x^2)^{n+1}}\,dx \quad [x^2 = (x^2+1) - 1] \\ &= \frac{x}{(1+x^2)^n} + 2nI_n(x) - 2nI_{n+1}(x). \end{aligned}$$

Kirjoittamalla n:n tilalle $n-1$ päädytään reduktiokaavaan

$$I_n(x) = \frac{1}{2n-2}\,\frac{x}{(1+x^2)^{n-1}} + \frac{2n-3}{2n-2}\,I_{n-1}(x). \tag{2}$$

Esimerkki 4 Kaavaa (2) kahdesti soveltamalla saadaan

$$\begin{aligned}\int \frac{1}{(1+x^2)^3}\,dx &= \frac{1}{4}\frac{x}{(1+x^2)^2} + \frac{3}{4}\int \frac{1}{(1+x^2)^2}\,dx \\ &= \frac{1}{4}\,\frac{x}{(1+x^2)^2} + \frac{3}{4}\left(\frac{1}{2}\,\frac{x}{1+x^2} + \frac{1}{2}\int \frac{1}{1+x^2}\,dx\right) \\ &= \underline{\underline{\frac{1}{4}\,\frac{x}{(1+x^2)^2} + \frac{3}{8}\,\frac{x}{1+x^2} + \frac{3}{8}\operatorname{Arctan} x + C.}} \quad \square\end{aligned}$$

Sijoitusmenettely

Sijoituksella tarkoitetaan integroinnissa <u>muuttujan</u> <u>vaihtoa</u> (kuten raja-arvoja laskettaessa, vrt. Luku V.2). Integroinnin sijoitusmenettely perustuu seuraavaan tulokseen, joka puolestaan perustuu yhdistetyn funktion ja käänteisfunktion derivoimissääntöihin.

Propositio VIII.2.1 Olkoon $u : (c,d) \longrightarrow (a,b)$ bijektio ja olkoon u derivoituva välillä (c,d) ja käänteisfunktio $v = u^{-1}$ derivoituva välillä (a,b). Tällöin, jos $G(t)$ on funktion $g(t) = f(u(t))u'(t)$ integraalifunktio välillä (c,d), ts.

$$\int f(u(t))u'(t)\,dt = G(t) + C, \quad t \in (c,d),$$

niin funktion $f(x)$ integraalifunktio välillä (a,b) on

$$\int f(x)\,dx = G(v(x)) + C, \quad x \in (a,b).$$

<u>Todistus</u> Jos $x \in (a,b)$ ja $t = v(x) \;\Leftrightarrow\; x = u(t)$, niin yhdistetyn funktion derivoimissäännön ja oletuksien perusteella

$$\begin{aligned}\frac{d}{dx}G(v(x)) &= G'(v(x))v'(x) = G'(t)v'(x) \\ &= f(u(t))u'(t)v'(x).\end{aligned}$$

Tässä on $u(t) = x$, joten $f(u(t)) = f(x)$ ja käänteisfunktion derivoimissäännön (Luku V.3, kaava (5)) mukaan

$$u'(t)v'(x) = 1 \quad \Rightarrow \quad \text{väite.} \quad \square$$

Sovelluksissa sijoitus tehdään usein (ehkä useammin) muodossa $v(x) = t$. Mekaanisesti sijoitusmenettely toimii siis seuraavalla tavalla:

Ongelma: $\displaystyle\int f(x)\,dx = ?$

1. Tehdään sijoitus: $x = u(t)$ tai $v(x) = t$.
2. Ratkaistaan
$$x = u(t) \;\Rightarrow\; t = v(x) \quad \text{tai} \quad v(x) = t \;\Rightarrow\; x = u(t).$$
3. Sijoitetaan integraaliin
$$x = u(t) \;\; (\text{tai } v(x) = t) \quad \text{ja} \quad dx = u'(t)\,dt.$$
4. Kirjoitetaan $f(u(t))u'(t) = g(t)$ (mahdollinen sievennys) ja ratkaistaan muunnettu ongelma:
$$\int g(t)\,dt = G(t) + C.$$
5. Sijoitetaan ratkaisuun $t = v(x)$, jolloin saadaan alkuperäisen ongelman ratkaisu:
$$\int f(x)\,dx = G(v(x)) + C.$$

Vaiheessa 2 ei haittaa, vaikka käänteisfunktio $v = u^{-1}$ tai $u = v^{-1}$ olisi monihaarainen (yhtälön $u(t) = x$ tai $v(x) = t$ ratkaisu monikäsitteinen), kunhan Proposition VIII.2.1 oletukset toteutuvat valitulla haaralla. On ainoastaan pidettävä huolta, ettei laskun eri vaiheissa hypellä haaralta toiselle.

Esimerkki 5 $\displaystyle\int \frac{1}{\sqrt{x}+1}\,dx = ?$

<u>Ratkaisu</u> Tässä on $(a, b) = (0, \infty)$.

1. Tehdään sijoitus: $v(x) = \sqrt{x} = t \in (0, \infty)$.
2. Ratkaistaan: $\sqrt{x} = t \;\Rightarrow\; x = u(t) = t^2$.
3. Sijoitetaan integraaliin: $\sqrt{x} = t,\; dx = 2t\,dt$.
4. Ratkaistaan muunnettu ongelma:
$$\int \frac{2t}{t+1}\,dt \;=\; \int \left(2 - \frac{2}{t+1}\right) dt \;=\; 2t - 2\ln(t+1) + C.$$

5. Alkuperäisen ongelman ratkaisu sijoituksella $t = \sqrt{x}$:

$$\int \frac{1}{\sqrt{x}+1}\,dx \;=\; \underline{\underline{2\sqrt{x} - 2\ln(\sqrt{x}+1) + C.}}$$

Esimerkki 6 $\displaystyle\int \frac{1}{e^x+1}\,dx = ?$

<u>Ratkaisu</u> Sijoittamalla $e^x = t \in (0,\infty) \;\Rightarrow\; x = \ln t,\; dx = \dfrac{1}{t}\,dt$ saadaan

$$\int \frac{1}{e^x+1}\,dx \;=\; \int \frac{1}{(t+1)t}\,dt \;=\; \ln\left|\frac{t}{t+1}\right| + C \quad \text{(Esimerkki VIII.1: 3)}$$

$$\Rightarrow \int \frac{1}{e^x+1}\,dx \;=\; \ln\left(\frac{e^x}{e^x+1}\right) + C \;=\; \underline{\underline{x - \ln(e^x+1) + C.}} \quad \square$$

Esimerkki 7 $\displaystyle\int \frac{1}{(1+x^2)^2}\,dx = ?$

<u>Ratkaisu</u> Tämän voi laskea reduktiokaavalla (2). Toinen vaihtoehto on trigonometrinen sijoitus

$$x = \tan t, \quad dx = \frac{1}{\cos^2 t}\,dt, \quad t \in (-\pi/2, \pi/2).$$

Koska $1 + \tan^2 t = 1/\cos^2 t$, saadaan muunnetuksi integraaliksi

$$\int \frac{1}{(1+\tan^2 t)^2} \cdot \frac{1}{\cos^2 t}\,dt = \int \cos^2 t\,dt = \frac{1}{2}(t + \sin t \cos t) + C$$

$$= \frac{1}{2}\left(t + \frac{\sin t}{\cos t} \cdot \cos^2 t\right) + C = \frac{1}{2}\left(t + \frac{\tan t}{1+\tan^2 t}\right) + C.$$

Ratkaisemalla lopuksi

$$x = \tan t \;\;\&\;\; t \in (-\pi/2, \pi/2) \quad \Leftrightarrow \quad t = \operatorname{Arctan} x$$

saadaan tulos

$$\int \frac{1}{(1+x^2)^2}\,dx = \underline{\underline{\frac{1}{2}\left(\operatorname{Arctan} x + \frac{x}{1+x^2}\right) + C.}}$$

<u>Vertailu</u>: Reduktiokaava (2) on menetelmänä suoraviivaisempi. $\square$

Esimerkki 8 $\displaystyle\int \frac{1}{x+1+\sqrt[3]{x+1}}\,dx = ?$

Ratkaisu Sijoituksella

$$\sqrt[3]{x+1} = t \;\Leftrightarrow\; x = t^3 - 1, \quad dx = 3t^2\,dt \quad (x > -1,\; t > 0)$$

saadaan

$$\begin{aligned}\int \frac{1}{x+1+\sqrt[3]{x+1}}\,dx = \int \frac{3t^2}{t^3+t}\,dt &= \int \frac{3t}{t^2+1}\,dt \\ &= \frac{3}{2}\ln(t^2+1) + C \\ &= \underline{\underline{\frac{3}{2}\ln[(x+1)^{2/3}+1] + C.}}\end{aligned}$$

Jos välillä $x \in (-\infty, -1)$ tulkitaan $\sqrt[3]{x+1} = -\sqrt[3]{|x+1|}$, niin tulos on pätevä myös tällä välillä. $\square$

Esimerkki 9 $\displaystyle\int \frac{e^x}{\sqrt{x}}\,dx \;= ?$

Ratkaisu Sijoituksella $\sqrt{x} = t,\; x = t^2,\; dx = 2t\,dt$ tehtävä muuntuu muotoon

$$\int 2e^{t^2}\,dt \;= ?$$

Tällä ei ole alkeisfunktioratkaisua (ainoastaan sarjaratkaisu, ks. seuraava luku), joten tyydytään tulokseen

$$\int \frac{e^x}{\sqrt{x}}\,dx = 2G(\sqrt{x}), \quad G(t) = \int e^{t^2}\,dt. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Laske osittain integroimalla

a) $\displaystyle\int x\cos x\,dx$ b) $\displaystyle\int x^3 e^{-x^2}\,dx$ c) $\displaystyle\int x^3\cos(x^2)\,dx$

d) $\displaystyle\int x^2 e^{ax}\,dx,\;\; a \in \mathbb{R}$ e) $\displaystyle\int x^\alpha \ln x\,dx,\;\; \alpha \in \mathbb{R}$

f) $\displaystyle\int \operatorname{Arctan} x\,dx$ g) $\displaystyle\int \operatorname{Arcsin} x\,dx$ h) $\displaystyle\int x\operatorname{Arctan} x\,dx$

2. Johda osittain integroimalla reduktiokaava annetulle integraalille $I_n(x)$ ja laske kaavan avulla $I_n(x)$ annetuilla n:n arvoilla.
   a) $\int x^n e^x\,dx,\ n \in \mathbb{N} \cup \{0\},\quad n = 1, 2, 3.$
   b) $\int x^n \sin x\,dx,\ n \in \mathbb{N} \cup \{0\},\quad n = 1, 2, 3.$
   c) $\int x^n \cos x\,dx,\ n \in \mathbb{N} \cup \{0\},\quad n = 1, 2, 3.$
   d) $\int \sin^n x\,dx,\ n \in \mathbb{Z},\quad n = -3, 4.$
   e) $\int (\ln x)^n\,dx,\ n \in \mathbb{N} \cup \{0\},\quad n = 1, 2.$
   f) $\int (x^2+1)^{n-1/2}\,dx,\ n \in \mathbb{Z},\quad n = -1, 2.$
   g) $\int (x^2-1)^{n-1/2}\,dx,\ n \in \mathbb{Z},\quad n = -1, 2.$
   h) $\int (1-x^2)^{n-1/2}\,dx,\ n \in \mathbb{Z},\quad n = -1, 2.$

3. Laske annettua sijoitusta käyttäen:

   a) $\displaystyle\int \frac{1}{\sqrt{5-x^2}}\,dx,\quad x = \sqrt{5}\,t$ b) $\displaystyle\int \frac{1}{x^2+2x+3}\,dx,\quad x+1 = \sqrt{2}\,t$

   c) $\displaystyle\int e^{\sqrt{x}}\,dx,\quad x = t^2$ d) $\displaystyle\int e^{\sqrt[3]{x}}\,dx,\quad \sqrt[3]{x} = t$

4. Muunna seuraavat integraalit toiseen muotoon annetulla sijoituksella.

   a) $\displaystyle\int \frac{1}{\sqrt{1+x^4}}\,dx,\quad x^4 = t$ b) $\displaystyle\int e^{x^2}\,dx,\quad x^2 = t$

   c) $\displaystyle\int \sin(\ln x)\,dx,\quad x = e^t$ d) $\displaystyle\int \ln(\tan x)\,dx,\quad \tan x = t$

5. Funktioiden $x^\alpha \sin \beta x,\ x^\alpha \cos \beta x,\ x^\alpha e^{\beta x}$ $(\alpha, \beta \in \mathbb{R})$ integraalifunktiot ovat alkeisfunktioita vain kun joko $\beta = 0$ tai $\alpha \in \mathbb{N} \cup \{0\}$. Mitkä seuraavista ovat tämän tiedon perusteella alkeisfunktioita ja mitkä eivät?

   a) $\displaystyle\int \cos(x^2)\,dx$ b) $\displaystyle\int x^{13}\sqrt{x}\cos\sqrt{x}\,dx$ c) $\displaystyle\int x^5 e^{x^4}\,dx$

   d) $\displaystyle\int x^7 e^{x^4}\,dx$ e) $\displaystyle\int \sqrt{\ln x}\,dx$ f) $\displaystyle\int \frac{1}{\ln x}\,dx$

   g) $\displaystyle\int \frac{1}{x\sqrt{\ln x}}\,dx$ h) $\displaystyle\int \sin(e^x)\,dx$ i) $\displaystyle\int \cos x\,\ln x\,dx$

6. (*) Olkoon $n \in \mathbb{N} \cup \{0\}$ ja $a, b \in \mathbb{R}$. Johda reduktiokaavat integraaleille

$$I_n(x) = \int x^n e^{ax} \cos bx, \quad J_n(x) = \int x^n e^{ax} \sin bx.$$

VIII.3 Osamurtokehitelmät. Sarjamenetelmä

Yleisen rationaalifunktion integraalifunktio on löydettävissä käyttäen nk. *osamurtokehitelmää* yhdessä edellisen luvun tekniikoiden kanssa. Tarkastellaan yleistä (reaalikertoimista) rationaalifunktiota

$$f(x) = \frac{p(x)}{q(x)} = \frac{a_m x^m + \cdots + a_0}{b_n x^n + \cdots b_0},$$

missä $m \geq 0$, $n \geq 1$ ja $a_m, b_n \neq 0$. Jos $m \geq n$, suoritetaan ensiksi jakolasku:

$$\frac{p(x)}{q(x)} = p_0(x) + \frac{p_1(x)}{q(x)}\,,$$

missä p_0 on polynomi astetta $m - n$,

$$p_0(x) = c_{m-n}x^{m-n} + \cdots + c_0,$$

ja jakojäännöksen $p_1(x)$ aste on $n - 1$:

$$p_1(x) = d_{n-1}x^{n-1} + \cdots + d_0.$$

Jakolasku on suoritettavissa esimerkiksi yksinkertaisesti identifioimalla polynomien kertoimet yhtälössä

$$p(x) = p_0(x)q(x) + p_1(x),$$

jolloin tuntemattomille kertoimille $c_{m-n}, \ldots, c_0$ ja $d_{n-1}, \ldots, d_0$ saadaan (ratkeava) lineaarinen yhtälöryhmä. Seuraavassa esimerkissä käytetään vaihtoehtoista (myös yleispätevää) menetelmää, *polynomien jakoalgoritmia*, joka on etenkin käsinlaskussa kätevä. Algoritmin yleisperiaate on esimerkistä arvattavissa.

Esimerkki 1 Laske p_0 ja p_1, kun $f(x) = \dfrac{x^3 + 2x^2}{x^2 - 2x + 3}$.

<u>Ratkaisu</u>

$$\begin{aligned} x^3 + 2x^2 &= [x(x^2 - 2x + 3) + 2x^2 - 3x] + 2x^2 \\ &= x(x^2 - 2x + 3) + 4x^2 - 3x \\ &= x(x^2 - 2x + 3) + [4(x^2 - 2x + 3) + 8x - 12] - 3x \\ &= (x + 4)(x^2 - 2x + 3) + 5x - 12 \end{aligned}$$

$$\Rightarrow \frac{x^3 + 2x^2}{x^2 - 2x + 3} = x + 4 + \frac{5x - 12}{x^2 - 2x + 3} = p_0(x) + \frac{p_1(x)}{q(x)}\,. \quad \square$$

Kun jakolasku on suoritettu, saadaan integraalin lineaarisuuden nojalla

$$\int \frac{p(x)}{q(x)}\,dx = \int p_0(x)\,dx + \int \frac{p_1(x)}{q(x)}\,dx.$$

Tässä polynomiosan integraalifunktio on heti määrättävissä (polynomina), joten jatkossa voidaan rajoittua tutkimaan jälkimmäistä termiä, jossa osoittajan aste on pienempi kuin nimittäjän aste. Oletetaan, että tunnetaan kaikki q:n (reaaliset ja kompleksiset) nollakohdat — nämähän ovat ainakin numeerisesti laskettavissa. Jos nollakohdat $x_1 \dots x_n$ ovat kaikki reaalisia ja yksinkertaisia, niin saadaan osamurtokehitelmä (ks. Lause VIII.3.1 alla)

$$\frac{p_1(x)}{q(x)} = \frac{A_1}{x - x_1} + \cdots + \frac{A_n}{x - x_n}\,.$$

Kertoimet A_i voidaan määrätä lineaarisesta yhtälöryhmästä, joka syntyy, kun tämä yhtälö kerrotaan puolittian $q(x)$:llä ja identifioidaan polynomien kertoimet vasemmalla ja oikealla. Helpommin kertoimet kuitenkin lasketaan kirjoittamalla $q(x) = (x - x_i)q_i(x)$, jolloin kertomalla yhtälö puolittain $(x - x_i)$:llä seuraa

$$A_i = \lim_{x \to x_i} \left[(x - x_i)\frac{p_1(x)}{q(x)} \right] = \frac{p(x_i)}{q_i(x_i)}\,.$$

Jos $q(x)$:llä on tekijänä $(x - a)^k$ vastaten k-kertaista reaalijuurta $x_i = a$, niin osamurtokehitelmässä tätä tekijää vastaavat yleisemmin termit

$$(x - a)^k \;\hookrightarrow\; \frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \cdots + \frac{A_k}{(x - a)^k}\,.$$

Näistä kukin termi on suoraan integroitavissa. Jos lopuksi $q(x)$:llä on tekijänä $(x^2 + ax + b)^k$ vastaten k-kertaista kompleksista konjugaattijuuriparia (jolloin on $b - a^2/4 > 0$), niin tämä tekijä tuottaa osamurtokehitelmään termit

$$(x^2 + ax + b)^k \;\hookrightarrow\; \frac{B_1 x + C_1}{x^2 + ax + b} + \frac{B_2 x + C_2}{(x^2 + ax + b)^2} + \cdots + \frac{B_k x + C_k}{(x^2 + ax + b)^k}\,.$$

Näiden integroimiseksi huomataan ensinnäkin, että

$$\frac{Ax + B}{(x^2 + ax + b)^l} = \frac{A}{2}\,\frac{2x + a}{(x^2 + ax + b)^l} + \frac{B - \frac{1}{2}\,aA}{(x^2 + ax + b)^l}\,,$$

joten

$$\int \frac{Ax + B}{(x^2 + ax + b)^l}\,dx = \frac{A}{2} \cdot \begin{cases} \ln(x^2 + ax + b), & \text{jos } l = 1, \\ -\dfrac{1}{l - 1}\,\dfrac{1}{(x^2 + ax + b)^{l-1}}\,, & \text{jos } l > 1 \end{cases}$$
$$+\ (B - \frac{1}{2}\,aA) \int \frac{1}{(x^2 + ax + b)^l}\,dx.$$

Tässä jäljelle jäänyt integraali on käsiteltävissä edellisen luvun menetelmin, sillä kun kirjoitetaan

$$x^2 + ax + b = \left(x + \frac{a}{2}\right)^2 + D^2, \quad D = \sqrt{b - \frac{a^2}{4}},$$

niin sijoituksella

$$x + \frac{a}{2} = Dt, \quad dx = Ddt$$

saadaan

$$\int \frac{1}{(x^2 + ax + b)^l}\, dx = D^{-2l+1} \int \frac{1}{(t^2+1)^l}\, dt, \quad t = D^{-1}\left(x + \frac{a}{2}\right).$$

Kun em. tulokset yhdistetään ja huomioidaan edellisen luvun tulokset, voidaan päätellä, että jokaisen rationaalifunktion integraalifunktio on muotoa

$$\int \frac{p(x)}{q(x)}\, dx = R(x) + \sum A_i F_i(x) + C,$$

missä R on rationaalifunktio (tai polynomi), kertoimet A_i ovat reaalilukuja ja kukin F_i (enintään n kpl, $n = q$:n aste) on jokin seuraavista funktiomuodoista:

$$\ln|x + a_i|, \quad \ln(x^2 + a_i x + b_i), \quad \mathrm{Arctan}(a_i x + b_i).$$

On siis päätelty, että rationaalifunktio voidaan aina integroida suljetussa muodossa, kun nimittäjäpolynomin kaikki nollakohdat (kertalukuineen) tunnetaan. Päätelmä perustui oletukseen, että osamurtokehitelmä edellä kuvatulla tavalla on aina mahdollinen, ts. että pätee

Lause VIII.3.1 (**Rationaalifunktion osamurtohajotelma**) Olkoon $f(x) = p(x)/q(x)$ reaalinen rationaalifunktio, jossa osoittajapolynomin p aste $< q$:n aste. Olkoon nimittäjäpolynomin q reaalijuuret x_i, $i = 1 \ldots n_1$ ja kompleksijuuret $c_j = a_j \pm ib_j$, $j = 1 \ldots n_2$ $(a_j, b_j \in \mathbb{R},\ b_j \neq 0)$. Edelleen olkoon näiden juurien kertaluvut m_i, $i = 1 \ldots n_1$ ja ν_j, $j = 1 \ldots n_2$. Tällöin on olemassa kertoimet $A_{ik}, B_{jk}, C_{jk} \in \mathbb{R}$ siten, että jokaisella $x \in \mathcal{D}_f$ pätee

$$f(x) = \sum_{i=1}^{n_1} \sum_{k=1}^{m_i} \frac{A_{ik}}{(x - x_i)^k} + \sum_{j=1}^{n_2} \sum_{k=1}^{\nu_j} \frac{B_{jk}x + C_{jk}}{[(x - a_j)^2 + b_j^2]^k}.$$

<u>Todistus</u> Sivuutetaan, ks. Harj.teht. 8.

Esimerkki 2 $\displaystyle\int \frac{x^2}{x^3 + x^2 + x - 3}\, dx = ?$

Ratkaisu Nimittäjäpolynomin eräs juuri on $x_1 = 1$: $q(x) = x^3 + x^2 + x - 3 = (x-1)(x^2+2x+3)$. Muut kaksi juurta ovat kompleksisia (konjugaattipari), joten osamurtokehitelmä on muotoa

$$\frac{x^2}{x^3+x^2+x-3} = \frac{A}{x-1} + \frac{Bx+C}{x^2+2x+3}\,.$$

Kertomalla tässä molemmat puolet $q(x)$:llä saadaan ensin määrätyksi kertoimet A, B, C, minkä jälkeen integrointi onnistuu:

$$\begin{aligned}
& x^2 = A(x^2+2x+3) + (Bx+C)(x-1) \quad \forall x \in \mathbb{R} \\
\Leftrightarrow \quad & (A+B-1)x^2 + (2A-B+C)x + (3A-C) = 0 \quad \forall x \in \mathbb{R} \\
\Leftrightarrow \quad & \begin{cases} A+B-1 & = 0 \\ 2A-B+C & = 0 \\ 3A-C & = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} A & = 1/6 \\ B & = 5/6 \\ C & = 1/2 \end{cases} \\
\Rightarrow \quad & \int \frac{x^2}{x^3+x^2+x-3}\,dx = \frac{1}{6}\int \frac{1}{x-1}\,dx + \frac{1}{6}\int \frac{5x+3}{x^2+2x+3}\,dx \\
& = \frac{1}{6}\ln|x-1| + \frac{1}{6}\int \left(\frac{5}{2}\,\frac{2x+2}{x^2+2x+3} - \frac{2}{x^2+2x+3}\right) dx \\
& = \frac{1}{6}\ln|x-1| + \frac{5}{12}\int \frac{2x+2}{x^2+2x+3}\,dx - \frac{1}{6}\int \frac{1}{(\frac{x+1}{\sqrt{2}})^2+1}\,dx \\
& = \underline{\underline{\frac{1}{6}\ln|x-1| + \frac{5}{12}\ln(x^2+2x+3) - \frac{\sqrt{2}}{6}\operatorname{Arctan}\left(\frac{x+1}{\sqrt{2}}\right) + C.}}
\end{aligned}$$

Kerroin A olisi voitu määrätä myös suoremmin laskemalla

$$A = \lim_{x\to 1}\left[(x-1)\frac{x^2}{x^3+x^2+x-3}\right] = \lim_{x\to 1}\frac{x^2}{x^2+2x+3} = \frac{1}{6}\,. \quad \square$$

Rationaalisiksi palautuvia integraaleja

Jos $R(x,y)$ on kahden muuttujan rationaalifunktio, niin mm. seuraavat integraalit ovat palautettavissa rationaalisiksi sopivilla sijoituksilla:

a) $\int R(\sin x, \cos x)\,dx$, b) $\int R(e^x)\,dx$,

c) $\int R(x, \sqrt{1-x^2}\,)\,dx$, d) $\int R(x, \sqrt{x^2+1}\,)\,dx$, e) $\int R(x, \sqrt{x^2-1}\,)\,dx$,

f) $\int R(x, \sqrt[m]{ax+b}\,)\,dx \quad (a \neq 0,\ m \in \mathbb{N})$.

Tapaus a) Tässä toimii trigonometrinen 'gurusijoitus' (vrt. Luku II.5)

$$\tan\frac{1}{2}x = t \quad \Rightarrow \quad \begin{cases} \sin x = \dfrac{2t}{1+t^2}, \quad \cos x = \dfrac{1-t^2}{1+t^2}, \\ \dfrac{dt}{dx} = \dfrac{1}{2}\left(1+\tan^2\dfrac{1}{2}x\right) \quad \Rightarrow \quad dx = \dfrac{2\,dt}{1+t^2} \end{cases}$$

$$\begin{aligned} \Rightarrow \quad \int R(\sin x, \cos x)\,dx &= \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2}{1+t^2}\,dt \\ &= \int Q(t)\,dt, \end{aligned}$$

missä Q on rationaalifunktio.

Tapaus b) Sijoituksella (vrt. edellisen luvun Esimerkki 6)

$$e^x = t \quad \Rightarrow \quad x = \ln t, \quad dx = \frac{1}{t}\,dt$$

saadaan integraali muotoon

$$\int R(e^x)\,dx = \int R(t)\frac{1}{t}\,dt = \int Q(t)\,dt,$$

missä Q on jälleen rationaalifunktio.

Tapaus c) Tämä palautuu tapaukseen a) sijoituksilla

$$x = \cos t \quad \text{tai} \quad x = \sin t.$$

Tapaukset d) ja e) Nämä palautuvat tapaukseen b) sijoituksilla

$$\text{d) } x = \sinh t, \qquad \text{e) } x = \cosh t.$$

Tapaus f) Tässä toimii sijoitus (vrt. edellisen luvun Esimerkki 8)

$$\sqrt[m]{ax+b} = t \quad \Rightarrow \quad x = \frac{1}{a}(t^m - b), \quad dx = \frac{m}{a}\,t^{m-1}\,dt$$

$$\Rightarrow \int R(x, \sqrt[m]{ax+b}\,)\,dx = \int R\left(\frac{t^m-b}{a}, t\right)\frac{m}{a}\,t^{m-1}\,dt = \int Q(t)\,dt.$$

Esimerkki 3 $\displaystyle\int \frac{1}{\sin x}\,dx = ?$

<u>Ratkaisu</u> Sijoituksella $\tan\frac{1}{2}x = t$ saadaan

$$\begin{aligned}\int \frac{1}{\sin x}\,dx &= \int \frac{1+t^2}{2t}\cdot\frac{2}{1+t^2}\,dt = \int \frac{1}{t}\,dt\\ &= \ln|t| + C\\ &= \ln\left|\tan\frac{1}{2}x\right| + C = \underline{\underline{\ln\left|\frac{1-\cos x}{\sin x}\right| + C.}} \quad \square\end{aligned}$$

Esimerkki 4 $\displaystyle\int \frac{\sqrt{x^2+1}}{x}\,dx \;= ?$

<u>Ratkaisu</u> Tehdään ensin sijoitus

$$x = \sinh t, \quad dx = \cosh t\,dt, \quad \sqrt{x^2+1} = \cosh t$$

$$\begin{aligned}\Rightarrow \int \frac{\sqrt{x^2+1}}{x}\,dx &= \int \frac{\cosh^2 t}{\sinh t}\,dt = \int \frac{\sinh^2 t + 1}{\sinh t}\,dt\\ &= \int \sinh t\,dt + \int \frac{1}{\sinh t}\,dt = \cosh t + \int \frac{2}{e^t - e^{-t}}\,dt.\end{aligned}$$

Tehdään uusi sijoitus

$$e^t = u, \quad t = \ln u, \quad dt = \frac{1}{u}\,du$$

$$\begin{aligned}\Rightarrow \int \frac{2}{e^t - e^{-t}}\,dt &= \int \frac{2}{u - u^{-1}}\cdot\frac{1}{u}\,du = \int \frac{2}{u^2-1}\,du\\ &= \ln\left|\frac{u-1}{u+1}\right| + C = \ln\left|\frac{e^t-1}{e^t+1}\right| + C.\end{aligned}$$

Koska $x = \sinh t \;\Leftrightarrow\; e^t = x + \sqrt{x^2+1}$, niin todetaan, että

$$\int \frac{\sqrt{x^2+1}}{x}\,dx = \sqrt{x^2+1} + \ln\left|\frac{x-1+\sqrt{x^2+1}}{x+1+\sqrt{x^2+1}}\right| + C.$$

Huomioimalla vielä, että (vrt. Harj.teht. V.1:12)

$$\frac{x-1+\sqrt{x^2+1}}{x+1+\sqrt{x^2+1}} = \frac{x}{1+\sqrt{x^2+1}}, \quad x \in \mathbb{R},$$

saadaan lopputulokselle hiukan sievempi muoto:

$$\int \frac{\sqrt{x^2+1}}{x}\,dx = \underline{\underline{\sqrt{x^2+1} + \ln\frac{|x|}{1+\sqrt{x^2+1}} + C.}} \quad \square$$

Integroinnin sarjamenetelmä

Silloin kun funktio on esitettävissä suppenevan potenssisarjan (Taylorin sarjan, ks. Luku VII.4) summana, voidaan myös integraalifunktio esittää tässä muodossa. Oletetaan, että

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k, \quad |x - x_0| < \rho \quad (\rho > 0).$$

Tällöin myös sarja

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x - x_0)^{k+1}$$

suppenee, kun $|x - x_0| < \rho$ (Lause I.12.11) ja sarja termeittäin derivoimalla (mikä Lauseen V.3.4 mukaan on sallittua) todetaan, että F on f:n integraalifunktio välillä $(x_0 - \rho, x_0 + \rho)$.

Esimerkki 5 $\displaystyle\int e^{x^2}\, dx \; = ?$

<u>Ratkaisu</u>

$$e^{x^2} = \sum_{k=0}^{\infty} \frac{1}{k!} x^{2k}, \quad x \in (-\infty, \infty),$$

joten

$$\begin{aligned} \int e^{x^2}\, dx &= \sum_{k=0}^{\infty} \frac{1}{(2k+1)k!} x^{2k+1} + C \\ &= C + x + \frac{1}{3} x^3 + \frac{1}{10} x^5 + \frac{1}{42} x^7 + .., \quad x \in \mathbb{R}. \quad \square \end{aligned}$$

Esimerkki 6 Määritä sarjoja hyväksi käyttäen

$$\text{a)} \int \frac{e^{-x}}{x}\, dx, \qquad \text{b)} \int \frac{\sin x}{x^4}\, dx.$$

<u>Ratkaisu</u> Lähtien Taylorin sarjoista

$$e^{-x} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^k, \qquad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

saadaan

$$\begin{aligned}\int \frac{e^{-x}}{x}\,dx &= \int \frac{1}{x}\left(1 - x + \frac{1}{2!}\,x^2 - \frac{1}{3!}\,x^3 + \dots\right)\,dx \\ &= \int \left(x^{-1} - 1 + \frac{1}{2!}\,x - \frac{1}{3!}\,x^2 + \dots\right)\,dx \\ &= \left(\ln|x| - x + \frac{1}{2\cdot 2!}\,x^2 - \frac{1}{3\cdot 3!}\,x^3 + \dots\right) + C \\ &= \ln|x| + C + \sum_{k=1}^{\infty} \frac{(-1)^k}{k\cdot k!}\,x^k,\end{aligned}$$

$$\begin{aligned}\int \frac{\sin x}{x^4}\,dx &= \int \frac{1}{x^4}\left(x - \frac{1}{3!}\,x^3 + \frac{1}{5!}\,x^5 - \dots\right)\,dx \\ &= \int \left(x^{-3} - \frac{1}{3!}\,x^{-1} + \frac{1}{5!}\,x - \dots\right)\,dx \\ &= \left(-\frac{1}{2}\,x^{-2} - \frac{1}{6}\ln|x| + \frac{1}{2\cdot 5!}\,x^2 - \dots\right) + C \\ &= -\frac{1}{2x^2} - \frac{1}{6}\ln|x| + C + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{2k(2k+3)!}\,x^{2k}.\end{aligned}$$

Molemmat tulokset ovat päteviä väleillä $(-\infty, 0)$ ja $(0, \infty)$. $\square$

HARJOITUSTEHTÄVIÄ

1. Integroi seuraavat rationaalifunktiot:

   a) $\dfrac{x^3 + 2x^2 - x + 5}{x + 2}$ b) $\dfrac{x^2}{x^2 + x - 2}$ c) $\dfrac{1}{x^3 - 4x^2 + 3x}$

   d) $\dfrac{24}{x(x^2 - 1)(x^2 - 4)}$ e) $\dfrac{1}{(1 - x)^2(1 + x)}$ f) $\dfrac{3 + x - 2x^3}{1 - x^3}$

   g) $\dfrac{x^8 + 1}{x^6 + x^4}$ h) $\dfrac{3x^2 + 1}{(x^2 - 1)^2}$ i) $\dfrac{1}{x^4 + 3x^2 + 2}$ j) $\dfrac{1}{x^4 + 1}$

   k) $\dfrac{1}{x(x^2 + a^2)}$ l) $\dfrac{1}{x^2(x^2 - a^2)}$ m) $\dfrac{x^3}{x^3 - a^3}$ n) $\dfrac{1}{x^4 - a^4}$

2. Muunna integraali

$$\int \frac{1}{x^4}\sqrt{x^2 + a^2}\,dx, \quad a > 0$$

   sijoituksilla a) $x = a\sinh t$, b) $x = a\tan t$, c) $x = a/t$. Valitse vaihtoehdoista helpoin ja laske integraali suljetussa muodossa.

3. Integroi sopivalla sijoituksella:

a) $\dfrac{e^{3x}}{e^x+2}$ b) $\dfrac{1}{\cosh x}$ c) $\dfrac{1}{1+\sqrt[3]{x}}$ d) $\dfrac{1}{\sqrt{x+1}+\sqrt[4]{x+1}}$

e) $x^2\sqrt{a^2-x^2}$ f) $\dfrac{1}{x\sqrt{1-x^2}}$ g) $\dfrac{1}{x(x^2+1)^{3/2}}$

h) $\dfrac{1}{5+4\sin x}$ i) $\dfrac{2-\sin x}{2+\cos x}$ j) $\dfrac{1}{5-4\sin x+3\cos x}$

4. Laske seuraavat integraalit sijoituksella $\tan x = t$:

a) $\displaystyle\int \frac{\tan x}{1+2\tan x}\,dx$ b) $\displaystyle\int \frac{1}{a\cos^2 x + b\sin^2 x}\,dx, \quad (a,b)\neq(0,0)$

5. Jos $R(x,y)$ on kahden muuttujan rationaalinen lauseke, niin millaisella sijoituksella integraali

$$\int R\left(x,\ \sqrt[m]{\frac{x+a}{x+b}}\right)dx, \quad m\in\mathbb{N},\ m\geq 2,\quad a,b\in\mathbb{R}$$

muuntuu rationaaliseksi? Laske tällä tavoin $\displaystyle\int \frac{1}{x}\sqrt{\frac{x+1}{x-1}}\,dx$.

6. Laske seuraavien funktioiden integraalifunktiot potenssisarjojen avulla:

a) $\dfrac{\sin x}{x}$ b) $\dfrac{\cosh x}{x}$ c) $\dfrac{1-\cos x}{x^3}$ d) $\dfrac{\sin x}{x^6}$

e) $x^3e^{x^3}$ f) $\sin x^2$ g) $\dfrac{\sin x^2}{x^3}$ h) $\dfrac{\cos x^3}{\sqrt{x}}$

7. (*) Esitä alkeisfunktioina (jos mahdollista) tai potenssisarjoina funktiot F ja G siten, että pätee

a) $\displaystyle\int e^x \ln x\,dx = F(x)\ln x + G(x) + C, \quad x\in(0,\infty),$

b) $\displaystyle\int \frac{1}{\sqrt{x}}\ \ln x\,\sin x\,dx = x\sqrt{x}\,[\,F(x)\ln x + G(x)\,] + C, \quad x\in(0,\infty),$

c) $\displaystyle\int \frac{\cos x}{x^2+1}\,dx = F(x)\,\mathrm{Arctan}\,x + G(x), \quad x\in\mathbb{R},$

d) $\displaystyle\int \frac{\sin x}{x^2+1}\,dx = F(x)\ln(x^2+1) + G(x), \quad x\in\mathbb{R}.$

8. (*) Lauseen VIII.3.1 todistamiseksi tarkastellaan kompleksimuuttujan rationaalifunktiota $f(z) = p(z)/q(z)$, missä q on reaalikertoiminen polynomi astetta $n \in \mathbb{N}$ ja p on reaalikertoiminen polynomi astetta $\leq n-1$.

a) Olkoon $c \in \mathbb{C}$ polynomin q yksinkertainen nollakohta ja olkoon $q(z) = (z-c)\,q_1(z)$ ja $g(z) = p(z)/q_1(z)$. Näytä, että

$$f(z) \;=\; \frac{g(c)}{z-c} + \frac{r(z)}{q_1(z)}\,, \quad z \in \mathcal{D}_f,$$

missä r on polynomi astetta $\leq n-2$. Päättele, että jos $c = a \in \mathbb{R}$, niin r on reaalikertoiminen.

b) Olkoon $c = a+ib$ ja $\overline{c} = a-ib$ polynomin q yksinkertaiset kompleksijuuret $(a, b \in \mathbb{R},\ b \neq 0)$ ja kirjoitetaan $q(z) = (z-c)(z-\overline{c})\,q_1(z)$. Näytä, että on olemassa $A \in \mathbb{C}$ ja $B, C \in \mathbb{R}$ siten, että pätee

$$f(z) \;=\; \frac{A}{z-c} + \frac{\overline{A}}{z-\overline{c}} + \frac{r(z)}{q_1(z)} \;=\; \frac{Bz+C}{(z-a)^2+b^2} + \frac{r(z)}{p_1(z)}, \quad z \in \mathcal{D}_f,$$

missä r on reaalikertoiminen polynomi astetta $\leq n-3$.

c) Olkoon $c \in \mathbb{C}$ q:n m-kertainen juuri ja olkoon $q(z) = (z-c)^m q_1(z)$ ja $g(z) = p(z)/q_1(z)$. Näytä, että pätee

$$f(z) \;=\; \sum_{k=1}^{m} \frac{A_k}{(z-c)^k} + \frac{r(z)}{p_1(z)}, \quad z \in \mathcal{D}_f,$$

missä $A_k = g^{(m-k)}(c)/(m-k)!$ ja r on polynomi astetta $\leq n-m-1$. Päättele, että jos $c = a \in \mathbb{R}$, niin kertoimet A_k ovat reaaliset ja r on reaalikertoiminen.

d) Olkoon $c = a+ib$ ja $\overline{c} = a-ib$ polynomin q m-kertaiset kompleksijuuret $(a, b \in \mathbb{R},\ b \neq 0)$ ja kirjoitetaan $q(z) = (z-c)^m(z-\overline{c})^m q_1(z)$. Näytä, että on olemassa kertoimet $A_k \in \mathbb{C}$ ja $B_k, C_k \in \mathbb{R}$ siten, että pätee

$$\begin{aligned} f(z) \;&=\; \sum_{k=1}^{m} \frac{A_k}{(z-c)^k} + \sum_{k=1}^{m} \frac{\overline{A}_k}{(z-\overline{c})^k} + \frac{r(z)}{q_1(z)} \\ &=\; \sum_{k=1}^{m} \frac{B_k z + C_k}{[(z-a)^2+b^2]^k} + \frac{r(z)}{q_1(z)}, \quad z \in \mathcal{D}_f, \end{aligned}$$

missä r on reaalikertoiminen polynomi astetta $\leq n-2m-1$.

e) Perustuen kohtien a) ja b) tuloksiin ja Algebran peruslauseeseen, johda Lauseen VIII.3.1 väittämä siinä tapauksessa, että q:n kaikki juuret ovat yksinkertaisia.

f) Todista Lause VIII.3.1 yleisessä tapauksessa perustuen kohtien c) ja d) tuloksiin ja Algebran peruslauseeseen.

VIII.4 Integraalifunktion numeerinen laskeminen. Määrätty integraali

Annetun funktion integraalifunktion määräämiseksi on toistaiseksi tarkasteltu erilaisia funktioalgebran keinoja. Näissä funktio on tunnettava laskusääntönä eli jonakin algebrallisena lausekkeena tarkasteltavalla välillä tai sen osaväleillä. Sovelluksia ajatellen tällainen oletus on melko rajoittava, sillä funktio voi yhtä hyvin olla pelkkä pisteittäin määritelty 'luettelo'. Näin on esimerkiksi, jos funktion arvoja voidaan käytännössä vain mitata tai jos funktion määritelmä on epäsuora niin, että funktioevaluaatio $x \mapsto f(x)$ onnistuu vain numeerisesti. Jos tällaisella funktiolla on integraalifunktio, niin luonnollisesti integraalifunktionkaan arvoja ei voida käytännössä laskea muuten kuin numeerisesti. Mutta miten? — Jatkossa ratkaistaan tämä ongelma johtamalla yleispätevä laskukaava numeeriselle funktioevaluaatiolle $x \mapsto F(x)$, missä F on annetun funktion integraalifunktio. Laskukaava johdattaa samalla integraalifunktiosta riippumattomaan *määrätyn integraalin* käsitteeseen.

Oletetaan, että f on määritelty (reaalifunktiona) välillä $[a, b]$, eli jokaiseen $x \in [a, b]$ liittyy yksikäsitteinen luku $f(x) \in \mathbb{R}$. Täsmennetään f:n integraalifunktion määrittämisen ongelma välillä (a, b) seuraavasti:

(P) Etsi funktio $y(x)$, joka on jatkuva välillä $[a, b]$, derivoituva välillä (a, b) ja toteuttaa

$$\begin{cases} y'(x) = f(x), & x \in (a, b), \\ y(a) = 0. \end{cases}$$

Tämän mukaan etsitään siis funktiota, joka on f:n integraalifunktio välillä (a, b) ja toteuttaa lisäksi asetetun jatkuvuusehdon ja alkuehdon. Koska derivoituvuus jo takaa jatkuvuuden välillä (a, b), niin jatkuvuusehto merkitsee ainoastaan lisäehtoja: $y(x)$ on oikealta jatkuva a:ssa ja vasemmalta jatkuva b:sä.

Sikäli kuin probleemalla (P) on ratkaisu, takaa asetettu alkuehto ratkaisun yksikäsitteisyyden (vrt. Korollaari V.6.8). Vaikeampaan kysymykseen, joka koskee ratkaisun olemassoloa, antaa erään vastauksen seuraava huomattava lause.

*Lause VIII.4.1 Jos f on jatkuva välillä $[a, b]$, niin probleemalla (P) on ratkaisu.

Lause VIII.4.1 on keskeinen osaväittämä yleisemmässä lauseessa, joka tunnetaan *Analyysin peruslauseen* nimellä. Tämä muotoillaan ja (osittain) todistetaan jäljempänä Luvussa VIII.6. Tässä luvussa asetetaan kysymys toisin: Jos yksinkertaisesti <u>oletetaan</u>, että (P) ratkeaa, niin miten luku $y(x) \in \mathbb{R}$ voidaan käytännössä <u>laskea</u>, kun $x \in (a, b]$ on annettu? Jatkossa tarkastellaan tätä kysymystä

ensin tapauksessa $x = b$, jolloin tehtävänä on siis laskea luku $y(b)$, kun $y(x)$ on määritelty epäsuorasti (P):n ratkaisuna. Koska $y(b)$ on reaaliluku, niin ko. lukua ei yleisesti voi laskea 'tarkasti', vaan käytännössä on tyydyttävä konstruoimaan $y(b)$ jonkin (viime kädessä rationaalisen) lukujonon $\{Y_n\}$ raja-arvona. Tehtävänä on siis määrittää algoritmi, jolla on mahdollista laskea jokaisella $n \in \mathbb{N}$ luku Y_n siten, että $\lim_n Y_n = y(b)$.

Olkoon $n \in \mathbb{N}$ annettu. Lukua Y_n määrättäessä otetaan lähtökohdaksi välin $[a, b]$ *jako osaväleihin* $[x_{k-1}, x_k]$, $k = 1 \ldots n$, missä $a = x_0 < x_1 < \ldots < x_n = b$. Jako voi olla esim. *tasavälinen*, jolloin on $x_k = a + kh_n$, missä $h_n = (b - a)/n$. Yleisemmin osavälijakoja rajoitetaan ainoastaan ehdolla

$$\lim_n h_n = 0, \quad \text{missä } h_n = \max_{k=1\ldots n}(x_k - x_{k-1}), \ \ n \in \mathbb{N}. \tag{$\star$}$$

Tässä h_n on osavälijaon nk. *tiheysparametri*. Tasavälisille jaoille raja-arvoehto ($\star$) luonnollisesti toteutuu.

Kun pisteet x_k, $k = 0 \ldots n$ on valittu em. tavalla, niin (P):n ratkaisulle $y(x)$ voidaan ensinnäkin kirjoittaa

$$y(b) \ = \ [y(x_n) - y(x_{n-1})] + \ldots + [y(x_1) - y(x_0)] + y(x_0)$$

ja alkuehdon $y(x_0) = y(a) = 0$ perusteella siis

$$y(b) = \sum_{k=0}^{n} [y(x_k) - y(x_{k-1})].$$

Funktio $y(x)$ toteuttaa Differentiaalilaskun väliarvolauseen ehdot osaväleillä $[x_{k-1}, x_k]$ ja lisäksi $y'(x) = f(x)$, $x \in (x_{k-1}, x_k)$. Väliarvolauseen mukaan on tällöin olemassa pisteet $\xi_k \in (x_{k-1}, x_k)$, $k = 1 \ldots n$ siten, että

$$y(x_k) - y(x_{k-1}) = f(\xi_k)(x_k - x_{k-1}), \quad k = 1 \ldots n.$$

Yhdistämällä kaksi viimeistä tulosta on päätelty, että (P):n ratkaisulle pätee

$$y(b) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Tämä ei toimi suoraan laskukaavana, koska pisteitä ξ_k ei tunneta, mutta kaava on luonteva lähtökohta approksimatiolle: Korvataan tuntemattomat pisteet ξ_k joillakin <u>valituilla</u> pisteillä $\xi_k \in [x_{k-1}, x_k]$, esim. $\xi_k = x_{k-1}$ tai $\xi_k = (x_{k-1} + x_k)/2$. Kun valinta on jokaisella osavälillä tehty, niin lasketaan

$$Y_n = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Näin muodostuu lukujono $\{Y_n\}$, jonka n:s termi siis lasketaan seuraavasti:

A1. Valitaan *jakopisteet* $x_k,\ k = 0 \ldots n$ niin, että $a = x_0 < x_1 < \ldots < x_n = b$.

A2. Valitaan *välipisteet* $\xi_k \in [x_{k-1}, x_k],\ k = 1 \ldots n$.

A3. Lasketaan $Y_n = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$.

Tämän mukaisesti Y_n:n laskemiseksi riittää suorittaa äärellinen määrä laskuoperaatioita: n kpl (käytännössä yleensä likimääräisiä) funktioevaluaatioita $x \mapsto f(x)$ ja lisäksi n kertolaskua ja $n-1$ yhteenlaskua.

Verrattaessa algoritmilla A1–A3 laskettua lukua Y_n lukuun $y(b)$ on em. tarkkaa $y(b)$:n lauseketta verrattava Y_n:n vastaavaan lausekkeeseen (A3). Näissä välipisteet ξ_k eivät ole samat, joten kirjoitettakoon lausekkeessa (A3) ξ_k:n tilalle η_k. Approksimaatiolle $Y_n \approx y(b)$ saadaan tällöin virhekaava

$$y(b) - Y_n \;=\; \sum_{k=1}^{n} [f(\xi_k) - f(\eta_k)](x_k - x_{k-1}).$$

Virheen arvioimiseksi on siis pystyttävä arvioimaan erotuksia $f(\xi_k) - f(\eta_k)$, kun pisteistä ξ_k ja η_k tiedetään ainoastaan, että ne ovat samalla osavälillä $[x_{k-1}, x_k]$. Arvion onnistumiseksi toivotulla tavalla on ilmeisesti tehtävä jokin säännöllisyysoletus funktiosta f. — Yleisesti jos f on riittävän säännöllinen, on $|f(\xi) - f(\eta)|$ enintään verrannollinen lukuun $|\xi - \eta|$, ts. $f(\xi) - f(\eta) = \mathcal{O}(|\xi - \eta|)$. Tämän takaava minimioletus on, että f on välillä $[a, b]$ Lipschitz-jatkuva, ts. $\exists\, L \in \mathbb{R}_+$ (= f:n Lipschitz-vakio) siten, että

$$|f(\xi) - f(\eta)| \le L|\xi - \eta| \quad \forall\, \xi, \eta \in [a, b]$$

(ks. Määritelmä V.6.4 ja myös Lause V.6.5).

Kun oletetaan f:n Lipschitz-jatkuvuus, niin pätee

$$|f(\xi_k) - f(\eta_k)| \le L|\xi_k - \eta_k|, \quad k = 1 \ldots n.$$

Tässä on $|\xi_k - \eta_k| \le x_k - x_{k-1}$, koska $\xi_k, \eta_k \in [x_{k-1}, x_k]$ ja edelleen $x_k - x_{k-1} \le h_n$, joten $|f(\xi_k) - f(\eta_k)| \le Lh_n,\ k = 1 \ldots n$. Kun käytetään näitä arvioita yhdessä kolmioepäyhtälön kanssa em. virhekaavassa, niin seuraa

$$\begin{aligned} |y(b) - Y_n)| \;&\le\; \sum_{k=1}^{n} |f(\xi_k) - f(\eta_k)|(x_k - x_{k-1}) \\ &\le\; Lh_n \sum_{k=1}^{n} (x_k - x_{k-1}) \;=\; L(b-a)h_n. \end{aligned}$$

On päädytty seuraavaan tulokseen.

Lause VIII.4.2 Olkoon f Lipschitz-jatkuva välillä $[a,b]$ vakiolla L ja olkoon $y(x)$ probleeman (P) ratkaisu. Tällöin jos $\{Y_n\}$ on mikä tahansa algoritmilla A1–A3 laskettu lukujono raja-arvoehdolla $(\star)$, niin $\lim_n Y_n = y(b)$, tarkemmin $|y(b) - Y_n| \le L(b-a)h_n \ \forall n \in \mathbb{N}$.†

Edellä laskettiin toistaiseksi (P):n ratkaisu vain pisteessä $x = b$. Tulos on kuitenkin helposti yleistettävissä. Nimittäin jos $y(x)$ on (P):n ratkaisu ja kiinnitetään $x \in (a,b]$, niin funktio $y(t),\ t \in [a,x]$ on jatkuva välillä $[a,x]$ (koska on jatkuva välillä $[a,b]$) ja ratkaisu alkuarvotehtävälle

$$\begin{cases} y'(t) = f(t), \ \ t \in (a,x), \\ y(a) = 0. \end{cases}$$

Kyseessä on siis probleema (P), missä muuttujana on x:n sijasta t ja b:n tilalla on x. Näin muodoin kun tehdään samat vaihdokset algoritmissa A1–A3, niin saadaan lasketuksi $y(x)$ missä tahansa halutussa pisteessä $x \in (a,b]$.

Esimerkki 1 Ratkaise alkuarvotehtävä $y'(x) = x,\ x > 0,\ y(0) = 0$ algoritmilla A1–A3.

<u>Ratkaisu</u> Kiinnitetään $x > 0$ ja tarkastellaan väliä $[0,x]$. Asetetaan jakopisteet t_k tasavälisesti, eli $t_k = kx/n,\ k = 0 \ldots n$, jolloin laskukaavan (A3) mukaan on

$$Y_n = \sum_{k=1}^{n} f(\xi_k)(t_k - t_{k-1}) = \sum_{k=1}^{n} \xi_k \frac{x}{n}.$$

Välipisteiden valinnassa kokeiltakoon vaihtoehtoja a) $\xi_k = t_k$, b) $\xi_k = t_{k-1}$ ja c) $\xi_k = (t_{k-1} + t_k)/2,\ k = 1 \ldots n$, jolloin saadaan:

$$\text{a)}\ \ Y_n = \sum_{k=1}^{n} k \frac{x}{n} \cdot \frac{x}{n} = \frac{x^2}{n^2} \sum_{k=1}^{n} k = \frac{x^2}{n^2} \cdot \frac{1}{2} n(n+1) = \frac{x^2}{2} + \frac{x^2}{2n}.$$

$$\text{b)}\ \ Y_n = \sum_{k=1}^{n} (k-1) \frac{x}{n} \cdot \frac{x}{n} = \frac{x^2}{n^2} \sum_{k=1}^{n} (k-1) = \frac{x^2}{n^2} \cdot \frac{1}{2} n(n-1) = \frac{x^2}{2} - \frac{x^2}{2n}.$$

$$\text{c)}\ \ Y_n = \sum_{k=1}^{n} \frac{1}{2} \left[(k-1) \frac{x}{n} + k \frac{x}{n} \right] \cdot \frac{x}{n} = \frac{x^2}{n^2} \sum_{k=1}^{n} \left(k - \frac{1}{2} \right) = \frac{x^2}{n^2} \cdot \frac{1}{2} n^2 = \frac{x^2}{2}.$$

†Kuten nähdään jäljempänä Luvussa VIII.6, Lauseen VIII.4.2 raja-arvoväittämä (ilman tarkennusta) on tosi myös, jos f oletetetaan ainoastaan jatkuvaksi välillä $[a,b]$. Mitään kvantitatiivista virhearviota approksimaatiolle $Y_n \approx y(b)$ ei pelkän jatkuvuusoletuksen perusteella saada.

Havaitaan, että kaikissa tapauksissa on $\lim_n Y_n = \frac{1}{2}x^2$. Lause VIII.4.2 soveltuu ($L = 1$), joten $y(x) = \frac{1}{2}x^2,\ x > 0$. $\square$

Esimerkissä pätee approksimaatiolle $Y_n \approx y(x)$ Lauseen VIII.4.2 mukaan virhearvio $|y(x) - Y_n| \leq x^2/n$ ($L = 1,\ b - a = x,\ h_n = x/n$). Havaitut virheet ovat a) $y(x) - Y_n = -\frac{x^2}{2n}$, b) $y(x) - Y_n = \frac{x^2}{2n}$, $y(x) - Y_n = 0$, joten nämä ovat arvion kanssa sopusoinnussa.

Jos (P):n ratkaisu $y(x)$ halutaan laskea useammassa pisteessä, voidaan algoritmi A1–A3 käynnistää jokaista laskentapistettä varten erikseen, kuten esimerkissä. Numeerisesti laskettaessa algoritmi kannattaa kuitenkin ottaa tehokkaampaan käyttöön huomioimalla, että kaavan (A3) mukaan luku Y_n tullaan laskeneeksi käytännössä palautuvasti muodossa

$$Y_0 = 0, \quad Y_k = Y_{k-1} + f(\xi_k)(x_k - x_{k-1}), \quad k = 1 \ldots n.$$

Luvut $Y_k = \sum_{i=1}^{k} f(\xi_i)(x_i - x_{i-1}),\ k = 1 \ldots n-1$ saadaan tällöin summauksen välituloksina. Nämä ovat algoritmin A1–A3 mukaisia approksimaatioita luvuille $y(x_k)$, kun jakopisteinä ovat $x_i,\ i = 0 \ldots k$ välillä $[a, x_k]$. Samoin oletuksin kuin edellä myös näiden virhe on $\mathcal{O}(h_n)$, tarkemmin $|y(x_k) - Y_k| \leq L(x_k - a)h_n,\ k = 1 \ldots n-1$. Luvut Y_k kannattaa siis kaikki huomioida, jolloin saadaan käsitys funktiosta $y(x)$ koko tarkasteltavalla välillä ilman lisätyötä.

Esimerkki 1 (jatko) Kuviossa on valittu $x = 2,\ n = 4$ ja yhdistetty esimerkin valinnoilla lasketut pisteet (x_k, Y_k) murtoviivaksi. Tämä esittää funktion $y(x)$ kuvaajaa likimäärin välillä $[0, 2]$. Tulos on tarkin tapauksessa (c), mutta muissakin tapauksissa tarkkuus kasvaa kiinteällä x, kun $n \to \infty$.

Määrätty integraali

Algoritmin A1–A3 mukaan lasketulla luvulla $y(b) = \lim_n Y_n$, missä $y(x)$ on probleeman (P) ratkaisu, on matematiikassa oma nimensä ja merkintänsä: Lukua kutsutaan funktion f *määrätyksi integraaliksi yli välin* $[a, b]$, merkitään

$$y(b) = \int_a^b f(x)\,dx$$

ja luetaan 'integraali a:sta b:hen $f(x)\,dx$'. Sanotaan edelleen, että a on integraalin *alaraja*, b on *yläraja* ja $[a, b]$ on *integroimisväli*. Muuttuja x, eli *integroimismuuttuja*, on 'dummy' (kuten summausindeksi), ts. muuttuja voidaan vaihtaa integraalin merkityksen muuttumatta.

Määrätty integraali on siis integroitavasta funktiosta f ja integroimisvälistä riippuva reaaliluku, joka on käytännösä laskettavissa lukujonon raja-arvona algoritmin (A1)–(A3) mukaisesti. Kun Y_n:n laskukaavassa (A3) käytetään erotuksille $x_k - x_{k-1}$ lyhenysmerkintää Δx_k, niin määrätyn integraalin laskeminen edellä esitetyllä tavalla voidaan tiivistää *summakaavaksi*

$$\boxed{\int_a^b f(x)\,dx = \operatorname*{Lim}_n \sum_{k=1}^{n} f(\xi_k)\Delta x_k.} \tag{1}$$

Tässä raja-arvomerkintään 'Lim_n' on sisällytetty ensinnäkin algoritmissa A1–A3 asetetut rajoitukset koskien pisteiden x_k ja ξ_k valintaa, toiseksi ehto ($\star$) ja kolmanneksi vaatimus, että <u>kaikki</u> algoritmin (A1)–(A3) mukaiset lukujonot $\{Y_n\}$ suppenevat kohti <u>samaa</u> raja-arvoa (= määrätyn integraalin arvo). Edellä esitetyn perusteella viimeksi mainittu vaatimus toteutuu (ja summakaava on siis pätevä) ainakin oletuksin, että probleema (P) on ratkeava (perusoletus toistaiseksi!) ja f on Lipschitz-jatkuva välillä $[a, b]$.

Em. oletuksilla probleeman (P) ratkaisulle johdettiin edellä myös yleisempi laskukaava, joka voidaan nyt kirjoittaa määrättynä integraalina:

$$y(x) = \int_a^x f(t)\,dt, \quad a < x \le b. \tag{2}$$

Koska (P):n ratkaisu on (eräs) f:n integraalifunktio välillä (a, b), niin kaava (2) yhdessä summakaavan (1) kanssa antaa tähänastisesta integroimistekniikasta (myös derivoimissäännöistä!) riippumattoman menetelmän integraalifunktion määräämiseksi.†

†Integraalifunktion merkinnän '$\int f(x)\,dx$' taustalla on määrätyn integraalin summakaava. Merkinnän otti käyttöön **G.W. Leibniz** ajatellen summakaavan 'rajankäyntejä' $\sum \hookrightarrow \int$ ja $\Delta x_k \hookrightarrow dx$, kun $n \to \infty$. Leibnizin omien päiväkirjamerkintöjen mukaan integraalimerkinnän tarkka keksimispäivä oli 29.10.1675.

Toisaalta jos välillä (a, b) tunnetaan f:n integraalifunktio F (esim. alkeisfunktiona tai sarjana) ja F on jatkuva välillä $[a, b]$ (välttämäntön lisäoletus!), niin luvun $\int_a^b f(x)\,dx$ laskeminen käy päinsä suoraan F:n avulla. Nimittäin näillä oletuksilla probleeman (P) ratkaisu on $y(x) = F(x) - F(a)$, jolloin

$$\int_a^b f(x)\,dx \;=\; y(b) \;=\; F(b) - F(a).$$

Tämä tunnettuun integraalifunktioon F perustuva määrätyn integraalin laskukaava esitetään yleensä *sijoituskaavana*

$$\boxed{\int_a^b f(x)\,dx = \Big|_{x=a}^{x=b} F(t) = \Big|_a^b F(x).} \tag{3}$$

Tässä oikea puoli luetaan 'sijoitus a:sta b:hen $F(x)$'. Kaava (3) saadaan päteväksi myös tapauksessa $a > b$, kun <u>sovitaan</u>, että määrätylle integraalille on voimassa *vaihtosääntö*

$$\boxed{\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.} \tag{4}$$

Mukavuussyistä oletetaan tämä päteväksi myös kun $a = b$, jolloin tulee sovituksi, että

$$\int_a^a f(x)\,dx = 0.$$

Huomautettakoon määrätyn integraalin käsitteestä, että toistaiseksi kyse ei ole muusta kuin erikoisesta merkinnästä luvulle $y(b)$ siinä tapauksessa, että probleemalla (P) on ratkaisu $y(x)$. Samaa ajatusta kertaavat myös sijoituskaava (3) ja vaihtosääntö (4). Seuraavassa luvussa nähdään, että määrätystä integraalista tulee (P):n ratkeavuudesta (ja yleisemminkin integraalifunktion olemassaolosta) riippumaton käsite, kun määritelmäksi otetaan suoraan summakaava (1), johon edellä päädyttiin numeerisena laskukaavana luvulle $y(b)$. Myöhemmissä luvuissa määrätyn integraalin käsite saa edelleen lisää 'eloa' erilaisista sovelluksista. Kuten tullaan näkemään, sovelluksissa määrättyyn integraaliin päädytään suoraan summakaavan kautta. Tällöin sijoituskaava (3) näyttäytyy oikotienä integraalin arvon 'tarkkaan' laskemiseen silloin, kun funktiolle f on löydettävissä (edellisten lukujen menetelmin) välillä $[a, b]$ jatkuva integraalifunktio F.

Seuraavassa esitetään määrätyn integraalin kolme keskeistä ominaisuutta. Nämä pysyvät voimassa, kun määrittelyn perustaksi jatkossa otetaan summakaava (1), mutta toistaiseksi ominaisuudet perustellaan vedoten oletettuun probleeman (P) ratkeavuuteen.

Additiivisuus. Lineaarisuus. Vertailuperiaate

Jos probleema (P) on ratkeava, niin jokaisella $c \in (a, b)$ voidaan kirjoittaa $y(b) = y(c) + [y(b) - y(c)]$ eli määrätyn integraalin avulla ilmaistuna

$$\boxed{\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.} \tag{5}$$

Tämän säännön perusteella sanotaan, että määrätty integraali on *additiivinen integroimisvälin suhteen.* Kun huomiodaan myös vaihtosääntö (4), niin todetaan, että additiivisuussääntö (5) on pätevä lukujen a, b, c suuruusjärjestyksestä riippumatta edellyttäen, että probleema (P) on ratkaistavissa, kun a:n tilalla on $\min\{a, b, c\}$ ja b:n tilalla $\max\{a, b, c\}$.

Olkoon (P):n ratkaisu $F(x)$ ja lisäksi olkoon (P):llä ratkaisu $G(x)$, kun f:n tilalla on funktio g. Tällöin jos $\alpha, \beta \in \mathbb{R}$, niin $y(x) = \alpha F(x) + \beta G(x)$ on (P):n ratkaisu, kun f:n tilalla on $\alpha f + \beta g$. Erityisesti on siis $y(b) = \alpha F(b) + \beta G(b)$, eli määrätylle integraalille pätee *lineaarisuussääntö* (vrt. määräämättömän integraalin vastaava sääntö VIII.1: (I-1))

$$\boxed{\int_a^b [\alpha f(x) + \beta g(x)]\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx, \quad \alpha, \beta \in \mathbb{R}.} \tag{6}$$

Jos mainittujen oletusten lisäksi oletetetaan, että että $f(x) \le g(x)\ \forall x \in [a, b]$, niin $F'(x) - G'(x) = f(x) - g(x) \le 0,\ x \in (a, b)$, jolloin $F(x) - G(x)$ on vähenevä välillä $[a, b]$ (Lause V.6.3). Erityisesti on $F(b) - G(b) \le F(0) - G(0) = 0$ eli $F(b) \le G(b)$. Näin ollen määrätyille integraaleille pätee *vertailuperiaate*

$$\boxed{f(x) \le g(x) \quad \forall x \in [a, b] \ \Rightarrow \ \int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.} \tag{7}$$

Koska $\pm f(x) \le |f(x)|$ ja $\int_a^b [\pm f(x)]\,dx = \pm \int_a^b f(x)\,dx$ (lineaarisuussääntö!), niin vertailuperiaattetta soveltaen seuraa erityisesti

$$\boxed{\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.} \tag{8}$$

Tätä sanotaan *integraalien kolmioepäyhtälöksi.*[†]

[†]Epäyhtälö (8) on sukua järjestetyn kunnan kolmioepäyhtälölle, joka pätee summakaavan (1) äärellisille summille.

Esimerkki 2 Sijoituskaavan (3) perusteella on

$$\int_0^2 \sin x\,dx = \Big|_0^2 (-\cos x) = \Big|_2^0 \cos x = 1 - \cos 2,$$

$$\int_x^{x^2} t\,dt = \Big|_x^{x^2} \frac{1}{2}t^2 = \frac{1}{2}(x^4 - x^2), \quad x \in \mathbb{R}.$$

Esimerkki 3 Laske $\int_0^2 f(x)\,dx$, kun $f(x) = \max\{\sqrt{x}, x^2\}$.

<u>Ratkaisu</u> Todetaan ensin, että $f(x) = \sqrt{x}$ välillä $[0,1]$ ja $f(x) = x^2$ välillä $[1,2]$. Tällöin käyttämällä ensin additiivisuussääntöä (5) ja sitten sijoituskaavaa (3) saadaan

$$\int_0^2 f(x)\,dx \;=\; \int_0^1 \sqrt{x}\,dx + \int_1^2 x^2\,dx \;=\; \Big|_0^1 \frac{2}{3}\,x^{3/2} + \Big|_1^2 \frac{1}{3}\,x^3 \;=\; 3. \qquad \square$$

Esimerkki 4 Määritä funktion $f(x) = \min\{3x, 4 - x^2\}$ integraalifunktio $\mathbb{R}$:ssä käyttäen hyväksi määrättyä integraalia.

<u>Ratkaisu</u> Funktio $y(x) = \int_0^x f(t)\,dt$ on kysytty integraalifunktio lisäehdolla $y(0) = 0$. Koska $f(x) = 3x$, kun $x \in [-4,1]$, ja $f(x) = 4 - x^2$, kun $x \geq 1$ tai $x \leq -4$, niin pätee

$$\begin{aligned}
x \in [-4,1] \quad &\Rightarrow \quad y(x) = \int_0^x 3t\,dt = \Big|_0^x \frac{3}{2}\,t^2 = \frac{3}{2}\,x^2, \\
x \in [1,\infty) \quad &\Rightarrow \quad y(x) = y(1) + [y(x) - y(1)] \\
&\qquad\qquad = \frac{3}{2} + \int_1^x (4 - t^2)\,dt \\
&\qquad\qquad = \frac{3}{2} + \Big|_1^x \left(4t - \frac{1}{3}\,t^3\right) = -\frac{1}{3}\,x^3 + 4x - \frac{13}{6}\,, \\
x \in (-\infty,-4] \quad &\Rightarrow \quad y(x) = y(-4) + [y(x) - y(-4)] \\
&\qquad\qquad = 24 + \int_{-4}^x (4 - t^2)\,dt \\
&\qquad\qquad = 24 + \Big|_{-4}^x \left(4t - \frac{1}{3}\,t^3\right) = -\frac{1}{3}\,x^3 + 4x + \frac{56}{3}\,.
\end{aligned}$$

Tässä $y(1) = 3/2$ ja $y(-4) = 24$ saatiin ensimmäisestä lausekkeesta. Yleinen integraalifunktio on $F(x) = y(x) + C,\ C \in \mathbb{R}$, joten

$$\int f(x)\,dx = \begin{cases} -\frac{1}{3}x^3 + 4x + \frac{56}{3} + C, & \text{kun } x < -4, \\ \frac{3}{2}x^2 + C, & \text{kun } -4 \le x \le 1, \\ -\frac{1}{3}x^3 + 4x - \frac{13}{6} + C, & \text{kun } x > 1. \end{cases} \qquad \square$$

Esimerkki 5 Sievennä lauseke $\dfrac{d}{dx}\displaystyle\int_{\sqrt{x}}^{1/\sqrt{x}} e^{-t^2},\ x > 0.$

<u>Ratkaisu</u> Funktiolla e^{-t^2} on $\mathbb{R}$:ssä sarjamuotoinen integraalifunktio $F(t)$, mutta sievennyksessä riittää tieto, että $F(t)$ on olemassa:

$$\begin{aligned} \frac{d}{dx}\int_{\sqrt{x}}^{1/\sqrt{x}} e^{-t^2}\,dt &= \frac{d}{dx}[F(\frac{1}{\sqrt{x}}) - F(\sqrt{x})] \\ &= -\frac{1}{2x\sqrt{x}}f(\frac{1}{\sqrt{x}}) - \frac{1}{2\sqrt{x}}f(\sqrt{x}) \\ &= -\frac{1}{2\sqrt{x}}\left(\frac{e^{-1/x}}{x} + e^{-x}\right). \quad \square \end{aligned}$$

Esimerkki 6 Arvioi virhe approksimaatiossa $\displaystyle\int_{10}^{20} \frac{x^4}{x^5+1}\,dx \approx \ln 2.$

<u>Ratkaisu</u> Integroimisvälillä on

$$\frac{x^4}{x^5+1} = \frac{1}{x} - \frac{1}{x(x^5+1)},$$

joten lineaarisuussäännön (6) ja sijoituskaavan (3) nojalla on

$$\int_{10}^{20} \frac{x^4}{x^5+1}\,dx = \int_{10}^{20} \frac{1}{x}\,dx - \delta = \Big|_{10}^{20} \ln x - \delta = \ln 2 - \delta,$$

missä on edelleen vertailuperiaatteen (7) ja sijoituskaavan nojalla

$$0 \le \delta = \int_{10}^{20} \frac{1}{x(x^5+1)}\,dx \le \int_{10}^{20} \frac{1}{x^6}\,dx = \Big|_{10}^{20} -\frac{1}{5x^5} < 2 \cdot 10^{-6}.$$

Siis approksimaatio on ylälikiarvo, jonka virhe on alle $2 \cdot 10^{-6}$. $\square$

HARJOITUSTEHTÄVIÄ

1. Ratkaise probleema (P), tai osoita ratkeamattomuus, kun $[a,b]=[0,1]$ ja

a) $f(x)=\begin{cases}\ln x, & \text{kun } x>0,\\ 0, & \text{kun } x=0,\end{cases}$ b) $f(x)=\begin{cases}\frac{1}{x}, & \text{kun } x>0,\\ 0, & \text{kun } x=0,\end{cases}$

c) $f(x)=\begin{cases}\frac{1}{\sqrt{x}}, & \text{kun } x>0,\\ 1, & \text{kun } x=0,\end{cases}$ d) $f(x)=\begin{cases}\frac{1}{\sqrt{|2x-1|}}, & \text{kun } x\neq\frac{1}{2},\\ 0, & \text{kun } x=\frac{1}{2}.\end{cases}$

2. Ratkaise alkuarvotehtävä $y'(x)=f(x)$, $x>0$, $y(0)=0$ käyttämällä algoritmia A1–A3 ja tasavälisiä jakoja välillä $[0,x]$, kun a) $f(x)=x^2$, b) $f(x)=x^3$, c) $f(x)=e^{-x}$, d) $f(x)=2^x$.
Lisätietoja: $\sum_{k=1}^n k^2=\frac{1}{6}n(n+1)(2n+1)$, $\sum_{k=1}^n k^3=\frac{1}{4}n^2(n+1)^2$.

3. Laske seuraavat raja-arvot tulkitsemalla ne määrätyiksi integraaleiksi.

a) $\displaystyle\lim_{n\to\infty}\frac{1}{n^4}\left[1^3+2^3+\ldots+(4n-1)^3\right]$ b) $\displaystyle\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n+k}$

c) $\displaystyle\lim_{n\to\infty}\frac{\pi}{n}\sum_{k=1}^{n-1}\sin\frac{k\pi}{n}$ d) $\displaystyle\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=0}^{n-1}\sqrt{n^2-k^2}$

e) $\displaystyle\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=0}^{n}\sqrt{n^2+k^2}$ f) $\displaystyle\lim_{n\to\infty}\sum_{k=0}^{n}\frac{n}{n^2+k^2}$

4. Laske sääntöjen (3)–(6) avulla tarkasti (jos mahdollista) tai numeerisena likiarvona:

a) $\displaystyle\int_{-2}^{2}(x^2+3)^2\,dx$ b) $\displaystyle\int_{4}^{9}\left(\sqrt{x}-\frac{1}{\sqrt{x}}\right)dx$ c) $\displaystyle\int_{1}^{10}\frac{1}{x^3+x}\,dx$

d) $\displaystyle\int_{-1}^{1}2^x\,dx$ e) $\displaystyle\int_{0}^{4}|\sin\theta|\,d\theta$ f) $\displaystyle\int_{0}^{\pi}\max\{\cos x,\,\sin 2x\}\,dx$

g) $\displaystyle\int_{1}^{2}|x^3+x^2-3|\,dx$ h) $\displaystyle\int_{0}^{\pi}|x-\cos x|\,dx$ i) $\displaystyle\int_{0}^{4}\min\{4x,\,e^x\}\,dx$

5. Laske seuraavien funktioiden derivaatat vapaan muuttujan suhteen.

a) $\displaystyle\int_{\pi}^{x}\sin^3 t\,dt$ b) $\displaystyle\int_{x}^{2\pi}(\sin^2 u-u^4e^{-u})\,du$ c) $\displaystyle\int_{x}^{5x}\frac{e^s}{s^2+1}\,ds$

d) $\displaystyle\int_{0}^{3\sinh 2x}\sqrt{9+t^2}\,dt$ e) $\displaystyle\int_{-\pi}^{t}\frac{\cos y}{1+y^2}\,dy$ f) $\displaystyle\int_{\sin\theta}^{\cos\theta}\sqrt{1-x^2}\,dx$

6. Määritä seuraavien funktioiden integraalifunktiot $\mathbb{R}$:ssä käyttäen hyväksi määrättyä integraalia:

$$\text{a)}\ \ f(x) = |x| - |x-2| \qquad \text{b)}\ \ f(x) = |x^2 - 7x + 10|$$
$$\text{c)}\ \ f(x) = \max\{\, x^2 + 2x + 3,\ 9 - 2x - x^2\}$$

7. Määritä seuraavien funktioiden pienimmät arvot ja piirrä funktioiden kuvaajat.

$$\text{a)}\ \ f(x) = \int_0^1 |x-t|\,dt \qquad \text{b)}\ \ f(x) = \int_0^\pi (x\cos t - t\cos x)^2\,dt$$

8. Todista:

$$\text{a)}\ \ 1 \le \int_0^1 \frac{1+x^{20}}{1+x^{21}} \le \frac{22}{21} \qquad \text{b)}\ \ \frac{2}{\sqrt[4]{e}} \le \int_0^2 e^{x^2-x}\,dx \le 2e^2$$
$$\text{c)}\ \ \int_3^5 \frac{x}{e\ln x}\,dx > 2 \qquad \text{d)}\ \ 0 < \int_{50}^{100} \frac{x^3}{x^6+8x+9}\,dx < 1.5\cdot 10^{-4}$$
$$\text{e)}\ \ \int_{100}^{300} \frac{x^5}{x^6+x-1}\,dx = \ln 3 - \delta, \quad 0 < \delta < 2\cdot 10^{-11}$$

9. (*) Perustele likimääräinen laskukaava

$$\int_1^2 \frac{e^{-x}}{x}\,dx \approx \ln 2 + \sum_{k=1}^{n} \frac{(-1)^k}{k\cdot k!}(2^k - 1), \quad n \in \mathbb{N},\ \ n \gg 1$$

ja arvioi tämän virhe, kun $n = 10$.

10. (*) Olkoon $F(x) = \displaystyle\int_0^{2x-x^2} \cos\left(\frac{1}{1+t^2}\right)dt, \quad G(x) = \int_4^{x^2} e^{t^2}\,dt.$

a) Tutki, saavuttaako F jollakin x pienimmän tai suurimman arvonsa.
b) Laske raja-arvo $\lim_{x\to 2} G(x)/(x^3-8)$.

11. (*) Probleemassa (P) olkoon $[a,b] = [-1,1]$, $f(0) = 0$ ja f:n määritelmä muualla kuin origossa

$$\text{a)}\ \ f(x) = \frac{1}{x}\left(\sin\frac{1}{x} + x\cos\frac{1}{x}\right), \quad \text{b)}\ \ f(x) = \frac{1}{x}\left(\sin\frac{1}{x^2} + x^2\cos\frac{1}{x^2}\right).$$

Näytä, että probleema (P) a) ei ratkea, b) ratkeaa.

VIII.5 Riemannin integraali

Tarkastellaan suljetulla välillä $[a, b]$ määriteltyä funktiota f, joka olkoon *rajoitettu*, ts. on olemassa $M \in \mathbb{R}_+$ siten, että pätee

$$|f(x)| \leq M \quad \forall x \in [a, b].$$

Välin $[a, b]$ yleistä *jakoa* merkitään jatkossa symbolilla X. Kuten edellisessä luvussa, jako tarkoittaa äärellistä järjestettyä joukkoa $X = \{x_0, x_1, \ldots, x_n\}$, missä $n \in \mathbb{N}$ ja

$$a = x_0 < x_1 < \ldots < x_n = b.$$

Jaon X keskeisin parametri on *tiheysparametri*, joka määritellään

$$h_X = \max_{k=1\ldots n} (x_k - x_{k-1}).$$

Jakoon X liittyen otetaan vielä käyttöön *välipisteistö* $\Xi_X = \{\xi_1, \ldots, \xi_n\}$, missä $\xi_k \in [x_{k-1}, x_k]$, $k = 1 \ldots n$ (muuten Ξ_X on vapaasti valittavissa).

Em. merkinnöin liitetään jokaiseen pariin (X, Ξ_X) reaaliluku $\sigma(f, X, \Xi_X)$, joka määritellään

$$\sigma(f, X, \Xi_X) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Edellisen luvun tapaan käytetään raja-arvomerkintää $\mathrm{Lim}_{h_X \to 0}\, \sigma(f, X, \Xi_X) = A$ kuvaamaan sellaista tilannetta, jossa summa lähestyy aina samaa raja-arvoa A ($A \in \mathbb{R}$), kun $h_X \to 0$. Kuten funktion raja-arvo, myös raja-arvo 'Lim' voidaan määritellä kahdella tavalla, joko lukujonojen avulla tai '(ε, δ)-määritelmänä', vrt. Määritelmä V.2.1 ja Lause V.2.2. Jälkimmäinen määrittelytapa muotoiltakoon jälleen lauseena. (Todistus sivuutetaan, vrt. Lauseen V.2.2 todistus.)

MÄÄRITELMÄ VIII.5.1 $\mathrm{Lim}_{h_X \to 0}\, \sigma(f, X, \Xi_X) = A \in \mathbb{R}$, jos jokaiselle jonolle $\{\sigma(f, X_n, \Xi_{X_n})\}$, jolle pätee $h_{X_n} \to 0$, on voimassa $\lim_{n\to\infty} \sigma(f, X_n, \Xi_{X_n}) = A$.

*LAUSE VIII.5.2 **(Lim: (ε, δ)-kriteeri)** $\mathrm{Lim}_{h_X \to 0}\, \sigma(f, X, \Xi_X) = A$ täsmälleen kun jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että jokaiselle parille (X, Ξ_X), jolle $h_X < \delta$, pätee

$$|\sigma(f, X, \Xi_X) - A| < \varepsilon.$$

Määritelmän VIII.5.1 mukaisesti raja-arvo 'Lim' tarkoittaa kaikista mahdollisista summista $\sigma(f, X, \Xi)$ poimittujen, ehdon $h_{X_n} \to 0$ täyttävien (ei muita ehtoja!) <u>lukujonojen</u> $\{\sigma(f, X_n, \Xi_{X_n})\}$ <u>yhteistä</u> <u>raja-arvoa</u>. Sikäli kuin tällainen yhteinen

raja-arvo on olemassa, Määritelmä VIII.5.1 tarjoaa myös algoritmin sen laskemiseksi: Valitaan mikä tahansa jono tiheneviä (esim. tasavälisiä) jakoja X_n, liitetään kuhunkin X_n jokin välipisteistö Ξ_n (esim. osavälien keskipisteet tai toinen päätepisteistä) ja lasketaan $A = \lim_n \sigma(f, X_n, \Xi_{X_n}) = \lim_n A_n$.

Edellisessä luvussa todettiin, että jos f:llä on välillä (a, b) integraalifunktio F, ja F on lisäksi jatkuva välillä $[a, b]$, niin jokaisella X on olemassa välipisteistö Ξ_X siten, että pätee

$$\sigma(f, X, \Xi_X) = F(b) - F(a) = \Big|_a^b F(x) = \int_a^b f(x)\,dx.$$

Edelleen näytettiin, että sikäli kuin määrätty integraali $\int_a^b f(x)\,dx$ määritellään tällä tavoin (mainituin oletuksin) ja lisäksi oletetaan, että f on Lipschitz-jatkuva välillä $[a, b]$, niin sekä jokaiselle jaolle X että jokaiselle välipisteistölle Ξ_X pätee (Lause VIII.4.2)

$$\sigma(f, X, \Xi_X) = \int_a^b f(x)\,dx + \mathcal{O}(h_X).$$

Määritelmän VIII.5.1 (tai Lauseen VIII.5.2) perusteella todetaan, että mainituilla (melko voimakkailla) oletuksilla pätee

$$\operatorname*{Lim}_{h_X \to 0} \sigma(f, X, \Xi_X) = \int_a^b f(x)\,dx.$$

Yleisemmin, jos f mahdollisesti ei täytä mainittuja ehtoja, niin tehdään tästä laskukaavasta integraalifunktiosta riippumaton <u>määrätyn</u> <u>integraalin</u> <u>määritelmä</u>.

MÄÄRITELMÄ VIII.5.3 **(Riemannin† integraali)** Olkoon f määritelty ja rajoitettu välillä $[a, b]$. Jos

$$\operatorname*{Lim}_{h_X \to 0} \sigma(f, X, \Xi_X) = A \quad (A \in \mathbb{R}),$$

niin sanotaan, että f on *Riemann-integroituva* (tai integroituva Riemannin mielessä) välillä $[a, b]$. Lukua A sanotaan f:n *Riemannin integraaliksi* (Riemann-integraaliksi) välillä $[a, b]$ ja merkitään

$$A = \int_a^b f(x)\,dx.$$

†Saksalainen matemaatikko **Georg Friedrich Bernhard Riemann** (1826-1866) on 1800-luvun (ja kaikkienkin aikojen) matematiikan suuria nimiä. Riemann oli 'puhdas matemaatikko' selvemmin kuin edeltäjänsä, joista huomattavimmatkin (esim. Euler, Lagrange, Cauchy, Gauss) tutkivat matematiikan ohella fysiikkaa tai muita matematiikan sovelluksia. Riemannin työt koskivat integraalien lisäksi mm. kompleksifunktioiden teoriaa (*Riemannin pinnat*), alkulukujen jakautumista, ja geometrian matemaattisia perusteita. Monet Riemannin tuloksista olivat uraa uurtavia ja näyttivät suuntaa myöhemmälle matematiikan tutkimukselle, joka 1800-luvulta lähtien erkaantui yhä selvemmin fysiikasta.

Määritelmään liittyen sanotaan summia $\sigma(f, X, \Xi_X)$ *Riemannin summiksi.* Jatkossa sanotaan yksinkertaisesti, että f on *integroituva* välillä $[a, b]$, jos f on sekä määritelty, rajoitettu että Riemann-integroituva ko. välillä .

Esimerkki 1 Olkoon $c \in \mathbb{R}$ ja määritellään

$$f(x) = \begin{cases} 0, & \text{kun } x < 1/2 \\ c, & \text{kun } x = 1/2 \\ 1, & \text{kun } x > 1/2 \end{cases}$$

Tutki f:n integroituvuutta välillä $[0, 1]$.

<u>Ratkaisu</u> Jos $X = \{x_0, \ldots, x_n\}$, $n \geq 2$, on välin $[0, 1]$ jako, niin jollakin $j \in \mathbb{N}$ on $x_{j-1} < 1/2 < x_{j+1}$. Tällöin on $f(\xi_k) = 0$ kun $k \leq j - 1$ ja $f(\xi_k) = 1$ kun $k \geq j + 2$, joten

$$\begin{aligned} \sigma(f, X, \Xi_X) &= \sum_{k=j}^{j+1} f(\xi_k)(x_k - x_{k-1}) + \sum_{k=j+2}^{n} (x_k - x_{k-1}) \\ &= f(\xi_j)(x_j - x_{j-1}) + f(\xi_{j+1})(x_{j+1} - x_j) + (1 - x_{j+1}). \end{aligned}$$

Kun $h_X \to 0$, niin $1 - x_{j+1} \to 1/2$, koska $0 < x_{j+1} - 1/2 < x_{j+1} - x_{j-1} < 2h_X$, ja kolmioepäyhtälön nojalla

$$\begin{aligned} |f(\xi_j)(x_j - x_{j-1}) + f(\xi_{j+1})(x_{j+1} - x_j)| &\leq \max\{|c|, 1\}(x_{j+1} - x_{j-1}) \\ &\leq 2\max\{|c|, 1\}h_X \to 0. \end{aligned}$$

Siis f on integroituva välillä $[0, 1]$ ja $\displaystyle\int_0^1 f(x)\,dx = \frac{1}{2}$. $\square$

Esimerkki 2 Dirichlet'n funktio

$$f(x) = \begin{cases} 1, & \text{jos } x \in \mathbb{Q} \\ 0, & \text{muulloin} \end{cases}$$

ei ole integroituva välillä $[0, 1]$, sillä jos valitaan $\Xi_X \subset \mathbb{Q}$, niin $\sigma(f, X, \Xi_X) = 1$ ja jos valitaan $\Xi_X \cap \mathbb{Q} = \emptyset$, niin $\sigma(f, X, \Xi_X) = 0$, olipa X mikä hyvänsä. $\square$

Esimerkeistä nähdään, että kaikki integroituvat funktiot eivät ole jatkuvia eivätkä kaikki rajoitetut funktiot ole integroituvia. Esimerkissä 1 integraalin arvo ei riipu funktion arvosta epäjatkuvuuspisteessä. — Yleisemminkin pätee, että jos f on välillä $[a, b]$ integroituva, niin integraalin arvo ei muutu, jos f määritellään

uudelleen äärellisen monessa pisteessä (tai jopa suppenevassa jonossa pisteitä, ks. Harj.teht. 8). Integraali on siis 'tunnoton' funktion yksittäisille pistearvoille samaan tapaan kuin funktion raja-arvo.

Näytetään nyt, että edellisessä luvussa todetut määrätyn integraalin keskeiset ominaisuudet, eli additiivisuus integroimisvälin suhteen, lineaarisuus integroitavan funktion suhteen ja integraalien vertialuperiaate, pysyvät voimassa myös Määritelmän VIII.5.3 mukaisille integraaleille.

LAUSE VIII.5.4 **(Integraalin additiivisuus)** Jos f on integroituva väleillä $[a, b]$ ja $[b, c]$ $(a < b < c)$, niin f on integroituva välillä $[a, c]$, ja

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

Todistus Tarkastellaan Riemannin summia $\sigma(f, X, \Xi_X)$ välillä $[a, c]$. Oletetaan aluksi, että $b \in X$, ja jaetaan summa kahteen osaan:

$$\sigma(f, X, \Xi_X) = \sum_{k:\, x_k \le b} f(\xi_k)(x_k - x_{k-1}) + \sum_{k:\, x_k > b} f(\xi_k)(x_k - x_{k-1}).$$

Määritelmien VIII.5.3 ja VIII.5.1 mukaan

$$\sum_{k:\, x_k \le b} f(\xi_k)(x_k - x_{k-1}) \to \int_a^b f(x)\,dx, \quad \text{kun } h_X \to 0,$$
$$\sum_{k:\, x_k > b} f(\xi_k)(x_k - x_{k-1}) \to \int_b^c f(x)\,dx, \quad \text{kun } h_X \to 0,$$

koska f oli integroituva väleillä $[a, b]$ ja $[b, c]$. Päätellään siis, että rajoituksen $b \in X$ ollessa voimassa pätee

$$h_X \to 0 \;\Rightarrow\; \sigma(f, X, \Xi_X) \to \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

Tarkastellaan seuraavaksi Riemannin summaa $\sigma(f, X, \Xi_X)$, missä $X = \{x_k\}$ ja $x_{m-1} < b < x_m$ jollakin m (jolloin $b \notin X$). Verrataan tätä summaan $\sigma(f, X', \Xi_{X'})$, missä X' on saatu X:stä lisäämällä vain jakopiste b ja $\Xi_{X'}$ on valittu siten, että välipisteet $\xi_k \in \Xi_X$ ja $\xi'_l \in \Xi_{X'}$ ovat samat jaoille yhteisillä osaväleillä, eli jos $[x_{k-1}, x_k] = [x'_{l-1}, x'_l]$. Tällöin on $x'_{m-1} = x_{m-1}$, $x'_m = b$ ja $x'_{m+1} = x_m$, joten

$$\begin{aligned}\sigma(f, X, \Xi_X) &- \sigma(f, X', \Xi_{X'}) \\ &= f(\xi_m)(x_m - x_{m-1}) - f(\xi'_m)(b - x_m) - f(\xi'_{m+1})(x_{m+1} - b).\end{aligned}$$

Jos $|f(x)| \leq M$ $\forall x \in [a,c]$, niin kolmioepäyhtälöä käyttäen päätellään

$$\begin{aligned} |\sigma(f,X,\Xi_X) - \sigma(f,X',\Xi_{X'})| &\leq 2M(x_m - x_{m-1}) \\ &\leq 2Mh_X \to 0, \quad \text{kun } h_X \to 0. \end{aligned}$$

Summalla $\sigma(f,X,\Xi_X)$ on siis sama raja-arvo riippumatta siitä, onko $b \in X$ tosi vai ei. $\square$

Kun käytetään edellisessä luvussa sovittua integroimisrajojen vaihtosääntöä

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx,$$

niin nähdään, että Lauseen VIII.5.4 väittämä on tosi lukujen a,b,c suuruussuhteista rippumatta, kunhan f on integroituva väleillä $[\min\{a,b\},\max\{a,b\}]$ ja $[\min\{b,c\},\max\{b,c\}]$.

Lause VIII.5.5 **(Integraalin lineaarisuus)** Jos f ja g ovat integroituvia välillä $[a,b]$, niin $\alpha f + \beta g$ on integroituva välillä $[a,b]$ jokaisella $\alpha,\beta \in \mathbb{R}$ ja

$$\int_a^b [\alpha f(x) + \beta g(x)]\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx.$$

Todistus Jos $\{X_n\}$ ja $\{\Xi_n\}$ ovat jako- ja välipisteistöt välillä $[a,b]$, niin vastaaville Riemannin summille pätee

$$\sigma(\alpha f + \beta g, X_n, \Xi_n) \;=\; \alpha\,\sigma(f,X_n,\Xi_n) + \beta\,\sigma(g,X_n,\Xi_n), \quad \alpha,\beta \in \mathbb{R}.$$

Väitetty lineaarisuussääntö seuraa tästä Määritelmän VIII.5.3 ja lukujonojen raja-arvojen yhdistelysääntöjen (Lause I.7.3) nojalla, kun oletetaan $\lim_n h_{X_n} = 0$. $\square$

Lause VIII.5.6 **(Integraalien vertailuperiaate)** Jos f ja g ovat integroituvia välillä $[a,b]$, niin pätee

$$f(x) \leq g(x) \quad \forall x \in [a,b] \;\Rightarrow\; \int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Todistus Myös tämä väittämä palautuu lukujonojen teoriaan: Vertailuperiaate on ilmeisen pätevä Riemannin summille $\sigma(f,X_n,\Xi_n)$ ja $\sigma(g,X_n,\Xi_n)$, joten se pätee myös raja-arvoille (Propositio I.7.4 [V1]). $\square$

*Riemannin ylä- ja alaintegraalit

Määritelmä VIII.5.3 Riemann-integroituvuudelle on sikäli konkreettiinen, että se antaa suoraan myös integraalin numeeriseen laskemiseen soveltuvan algoritmin. Seuraavassa lähestytään Riemannin integraalia toisella, hieman abstraktimmalla tavalla, mikä johtaa vaihtoehtoiseen Riemann-integroituvuuden kriteeriin (Lause VIII.5.8). Jatkossa tämä kriteeri osoittautuu käteväksi erilaisissa teoreettisissa tarkasteluissa (ks. myös Harj.teht. 6–10).

Olkoon f välillä $[a,b]$ määritelty ja rajoitettu funktio ja $X = \{x_k,\ k = 0\dots n\}$ välin $[a,b]$ jako. Koska f on rajoitettu, niin joukot $Y_k = \{f(x) \mid x \in [x_{k-1}, x_k]\}$ ovat rajoitettuja. Siis on olemassa luvut (vrt. Luku I.11)

$$\overline{f}_k = \sup Y_k, \quad \underline{f}_k = \inf Y_k, \quad k = 1\dots n.$$

Funktion f jakoon X liittyviksi *Riemannin ylä- ja alasummiksi* sanotaan summia

$$\begin{aligned} \overline{\sigma}(f,X) &= \sum_{k=1}^{n} \overline{f}_k (x_k - x_{k-1}) \quad \text{(yläsumma)}, \\ \underline{\sigma}(f,X) &= \sum_{k=1}^{n} \underline{f}_k (x_k - x_{k-1}) \quad \text{(alasumma)}. \end{aligned}$$

Huomattakoon, että jos f on jatkuva välillä $[a,b]$, niin f saavuttaa maksimi- ja minimiarvonsa jokaisella osavälillä $[x_{k-1}, x_k] \subset [a,b]$ (Lause V.1.12), jolloin $\overline{f}_k = f(\xi_k)$ ja $\underline{f}_k = f(\eta_k)$ joillakin $\xi_k, \eta_k \in [x_{k-1}, x_k]$. Tässä tapauksessa siis Riemannin ylä- ja alasummat ovat Riemannin summien erikoistapauksia.

Olkoon nyt $\mathcal{X}$ välin $[a,b]$ kaikkien mahdollisten jakojen X joukko. Tällöin reaalilukujoukot $\{\overline{\sigma}(f,X) \mid X \in \mathcal{X}\}$ ja $\{\underline{\sigma}(f,X) \mid X \in \mathcal{X}\}$ ovat rajoitettuja, sillä jos $|f(x)| \le M,\ x \in [a,b]$, niin $|\overline{f}_k| \le M$ ja $|\underline{f}_k| \le M \ \ \forall k$, jolloin

$$|\overline{\sigma}(f,X)|,\ |\underline{\sigma}(f,X)| \ \le\ M \sum_{k=1}^{n} (x_k - x_{k-1}) \ =\ M(b-a).$$

Mainituilla lukujoukoilla on siis sekä supremum että infimum.

Määritelmä VIII.5.7 Jos f on määritelty ja rajoitettu välillä $[a,b]$ ja $\mathcal{X}$ on välin $[a,b]$ kaikkien mahdollisten jakojen X joukko, niin funktion f *Riemannin ylä- ja alaintegraalit* välillä $[a,b]$ määritellään

$$\begin{aligned} \overline{\int_a^b} f(x)\,dx &= \inf_{X \in \mathcal{X}} \overline{\sigma}(f,X) \quad \text{(yläintegraali)}, \\ \underline{\int_a^b} f(x)\,dx &= \sup_{X \in \mathcal{X}} \underline{\sigma}(f,X) \quad \text{(alaintegraali)}. \end{aligned}$$

Määritelmän VIII.5.7 mukaisesti siis jokainen välillä $[a, b]$ määritelty ja rajoitettu funktio on ko. välillä sekä 'yläintegroituva' että 'alaintegroituva'.

ESIMERKKI 3 Jos $f(x) = 1$, kun $x \in \mathbb{Q}$, ja $f(x) = 0$ muulloin, niin

$$\overline{\int_0^1} f(x)\,dx = 1, \quad \underline{\int_0^1} f(x)\,dx = 0. \quad \square$$

Esimerkki johdattelee seuraavaan hyvin eleganttiin Riemann-integroituvuuden kriteeriin. Todistus esitetään luvun lopussa.

*LAUSE VIII.5.8 **(Riemann-integroituvuus)** Välillä $[a, b]$ määritelty ja rajoitettu funktio f on ko. välillä Riemann-integroituva täsmälleen kun f:n ylä- ja alaintegraalit ko. välillä ovat samat, jolloin f:n Riemannin integraali = ylä- ja alaintegraalien yhteinen arvo.

Seuraavassa esitetään kaksi Riemannin ylä- ja alasummia koskevaa väittämää, joilla on teoreettisissa tarkasteluissa käyttöä yhdessä Lauseen VIII.5.8 kanssa. Ensinnäkin näytetään, että jaon *tihennyksessä* yläsumma pienenee tai pysyy samana ja vastaavasti alasumma suurenee tai pysyy samana.

LAUSE VIII.5.9 Jos X ja X' ovat välin $[a, b]$ jakoja ja $X' \supset X$ (eli X' on X:n tihennys), niin

$$\underline{\sigma}(f, X) \le \underline{\sigma}(f, X') \le \overline{\sigma}(f, X') \le \overline{\sigma}(f, X).$$

<u>Todistus</u> Tihennys voidaan ajatella suoritetuksi lisäämällä jakoon X yksi piste kerrallaan, kunnes päädytään jakoon X'. Väittämä on tosi, jos se on tosi jokaiselle tällaiselle osatihennykselle. Olkoon siis $X = \{x_k,\ k = 0 \ldots n\}$ ja olkoon $X' = X \cup \{x'_k\}$, missä $x'_k \in (x_{k-1}, x_k)$. Tällöin jos $\overline{f}_k$ ja $\underline{f}_k$ määritellään kuten edellä (jakoon X liittyen) ja merkitään

$$\alpha_k = \sup_{x \in [x_{k-1},\, x'_k]} f(x), \quad \beta_k = \sup_{x \in [x'_k,\, x_k]} f(x),$$

niin $\alpha_k \le \overline{f}_k$ ja $\beta_k \le \overline{f}_k$ (vrt. Harj.teht. I.11:2c), joten

$$\alpha_k(x'_k - x_{k-1}) + \beta_k(x_k - x'_k) \ \le\ \overline{f}_k(x_k - x_{k-1}).$$

Koska jaot ovat samat välin $[x_{k-1}, x_k]$ ulkopuolella, niin päätellään, että on oltava $\overline{\sigma}(f, X') \le \overline{\sigma}(f, X)$. Vastaavalla tavalla näytetään, että $\underline{\sigma}(f, X) \le \underline{\sigma}(f, X')$. Ylä- ja alasummien määritelmän perusteella on myös $\underline{\sigma}(f, X') \le \overline{\sigma}(f, X')$, joten väite seuraa. $\square$

Seuraava väittämä konkretisoi ylä- ja alaintegraalit lukujonojen avulla. Väittämä nojaa oleellisesti edelliseen.

*Lause VIII.5.10 Olkoon f määritelty ja rajoitettu välillä $[a,b]$. Tällöin jokaiselle jonolle $\{X_n\}$ välin $[a,b]$ jakoja, jolle $h_{X_n} \to 0$, pätee

$$\lim_n \overline{\sigma}(f, X_n) = \overline{\int_a^b} f(x)\,dx, \qquad \lim_n \underline{\sigma}(f, X_n) = \underline{\int_a^b} f(x)\,dx.$$

Todistus Olkoon $\{X_n\}$ oletusten mukainen jono ja olkoon $\varepsilon > 0$. Yläintegraalin määritelmän nojalla on olemassa välin $[a,b]$ jako $X = \{x_k,\ k = 0 \ldots K\}$ siten, että

$$0 \ \le\ \overline{\sigma}(f, X) - \overline{\int_a^b} f(x)\,dx \ <\ \frac{\varepsilon}{2}\,.$$

Merktitään $X_n' = X_n \cup X$, jolloin X_n' on jakojen X_n ja X tihennys. Tihennyksessä $X_n \hookrightarrow X_n'$ jakautuvat jaon X_n synnyttämistä osaväleistä täsmälleen ne, joiden sisäpisteenä on ainakin yksi joukon X piste. Tällaisia osavälejä on enintään $K-1$, ja näistä kunkin pituus on enintään h_{X_n}. Koska vain nämä osavälit vaikuttavat erotukseen $\overline{\sigma}(f, X_n) - \overline{\sigma}(f, X_n')$, niin merkitsemällä mainittujen osavälien osuuksia summissa $\overline{\sigma}(f, X_n)$ ja $\overline{\sigma}(f, X_n')$ symboleilla $\overline{\sigma}_0(f, X_n)$ ja $\overline{\sigma}_0(f, X_n')$ ja olettaen, että $|f(x)| \le M,\ x \in [a,b]$, voidaan Lauseen VIII.5.9 ja kolmioepäyhtälöm perusteella arvioida

$$\begin{aligned} 0 \le \overline{\sigma}(f, X_n) - \overline{\sigma}(f, X_n') \ &= \ |\overline{\sigma}_0(f, X_n) - \overline{\sigma}_0(f, X_n')| \\ &\le \ |\overline{\sigma}_0(f, X_n)| + |\overline{\sigma}_0(f, X_n')| \ \le\ (K-1)\cdot 2M \cdot h_{X_n}\,. \end{aligned}$$

Toisaalta koska X_n' on myös X:n tihennys, niin pätee (Lause VIII.5.9)

$$\overline{\sigma}(f, X_n') \le \overline{\sigma}(f, X).$$

Yhdistämällä epäyhtälöt seuraa

$$\begin{aligned} \overline{\int_a^b} f(x)\,dx \ \le\ \overline{\sigma}(f, X_n) \ &\le\ \overline{\sigma}(f, X_n') + 2M(K-1)\,h_{X_n} \\ &\le\ \overline{\sigma}(f, X) + 2M(K-1)\,h_{X_n} \\ &<\ \overline{\int_a^b} f(x)\,dx + \frac{\varepsilon}{2} + 2M(K-1)\,h_{X_n}\,. \end{aligned}$$

Koska $h_{X_n} \to 0$ kun $n \to \infty$ ja koska K on n:stä riippumaton, niin tässä on $2M(K-1\,h_{X_n} < \varepsilon/2$ jostakin indeksistä N eteenpäin, jolloin seuraa

$$\overline{\int_a^b} f(x)\,dx \ \le\ \overline{\sigma}(f, X_n) \ <\ \overline{\int_a^b} f(x)\,dx + \varepsilon, \quad n > N.$$

Tässä $\varepsilon > 0$ oli mielivaltainen ja $N \in \mathbb{N}$, joten lukujonon suppenemisen määritelmän nojalla on todistettu väittämän yläsummia koskeva osa. Alasummia koskeva osaväittämä todistetaan vastaavalla tavalla. $\square$

Esimerkkinä Lauseiden VIII.5.8 ja VIII.5.10 soveltamisesta todistettakoon

LAUSE VIII.5.11 **(Integroituvuus osavälillä)** Jos välillä $[a, b]$ määritelty ja rajoitettu funktio f on ko. välillä Riemann-integroituva, niin f on Riemann-integroituva myös jokaisella osavälillä $[c, d] \subset [a, b]$.

<u>Todistus</u> Olkoon $\{X_n\}$ jono välin jakoja siten, että $h_{X_n} \to 0$ ja $c, d \in X_n$ $\forall n$. Tällöin X_n sisältää jokaisella n välin $[c, d]$ jaon X'_n, jolle $h_{X'_n} \leq h_{X_n} \to 0$. Kun merkitään $X_n = \{x_k,\ k = 0 \ldots n\}$ ja $X'_n = \{x_k,\ k = l-1 \ldots m\}$, niin

$$\begin{aligned}\overline{\sigma}(f, X_n) - \underline{\sigma}(f, X_n) &= \sum_{k=1}^{n} (\overline{f}_k - \underline{f}_k)(x_k - x_{k-1}) \geq \sum_{k=l}^{m} (\overline{f}_k - \underline{f}_k)(x_k - x_{k-1}) \\ &= \overline{\sigma}(f, X'_n) - \underline{\sigma}(f, X'_n) \geq 0.\end{aligned}$$

Koska tässä pätee oletusten ja Lauseiden VIII.5.8 ja VIII.5.10 perusteella

$$\overline{\sigma}(f, X_n) - \underline{\sigma}(f, X_n) \to \overline{\int_a^b} f(x)\, dx - \underline{\int_a^b} f(x)\, dx = 0,$$

niin $\overline{\sigma}(f, X'_n) - \underline{\sigma}(f, X'_n) \to 0$ (Propositio I.7.4 [V2]). Koska myös $h_{X'_n} \to 0$, niin Lauseen VIII.5.10 perusteella f:n ylä- ja alaintegraalit välillä $[c, d]$ ovat samat, eli f on Riemann-integroituva ko. välillä (Lause VIII.5.8). $\square$

Lauseen VIII.5.8 todistus. $\boxed{\Rightarrow}$ Oletetaan, että f on Riemann-integroituva välillä $[a, b]$ Määritelmän VIII.5.3 mukaisesti. Olkoon $\{X_n\}$ jono välin $[a, b]$ jakoja, jolle $h_{X_n} \to 0$. Valitaan indeksi n ja merkitään $X_n = \{x_k,\ k = 0 \ldots K\}$. Nyt voidaan jokaisella $k = 1 \ldots K$ valita $\xi_k, \xi'_k \in [x_{k-1}, x_k]$ siten, että pätee

$$\begin{aligned} 0 &\leq \overline{f}_k - f(\xi_k) < 2^{-n}, \qquad \overline{f}_k = \sup_{x \in [x_{k-1}, x_k]} f(x), \\ 0 &\leq f(\xi'_k) - \underline{f}_k < 2^{-n}, \qquad \underline{f}_k = \inf_{x \in [x_{k-1}, x_k]} f(x). \end{aligned}$$

Tällöin

$$\begin{aligned} 0 &\leq \overline{\sigma}(f, X_n) - \sigma(f, \Xi_n, X_n) < 2^{-n}(b-a), \qquad \Xi_n = \{\xi_k\}, \\ 0 &\leq \sigma(f, \Xi'_n, X_n) - \underline{\sigma}(f, X_n) < 2^{-n}(b-a), \qquad \Xi'_n = \{\xi'_k\}. \end{aligned}$$

Koska tässä $\sigma(f, \Xi_n, X_n) \to A \in \mathbb{R}$ ja $\sigma(f, \Xi'_n, X_n) \to A$ oletuksen mukaan, niin seuraa, että myös $\overline{\sigma}(f, X_n) \to A$ ja $\underline{\sigma}(f, X_n) \to A$. Lauseen VIII.5.10 mukaan A on tällöin f:n ylä- ja alaintegraalien yhteinen arvo.

$\boxed{\Leftarrow}$ Oletetetaan, että f:n ylä- ja alaintegraaleilla välillä $[a,b]$ on yhteinen arvo A. Tällöin jos $\{X_n\}$ on jono välin $[a,b]$ jakoja ja $h_{X_n} \to 0$, niin Lauseen VIII.5.10 perusteella $\overline{\sigma}(f,X_n) \to A$ ja $\underline{\sigma}(f,X_n) \to A$. Tällöin myös $\sigma(f,\Xi_{X_n},X_n) \to A \;\forall\, \Xi_{X_n}$, koska $\underline{\sigma}(f,X_n) \le \sigma(f,\Xi_{X_n},X_n) \le \overline{\sigma}(f,X_n)$ (Propositio I.7.4 [V2]). Siis f on integroituva Määritelmän VIII.5.3 mukaisesti. $\square$

HARJOITUSTEHTÄVIÄ

1. a) Olkoon $f(x) = x^2$, kun $x \in \mathbb{Z}$, ja $f(x) = 0$, kun $x \in \mathbb{R},\ x \notin \mathbb{Z}$. Näytä, että f on integroituva jokaisella välillä $[a,b] \subset \mathbb{R}$ ja että $\int_a^b f(x)\,dx = 0$.
   b) Olkoon $f(x) = 1$, kun $x \in \{\,2^{-k},\ k = 0,1,\dots\}$ ja $f(x) = 0$ muulloin. Laskemalla yläsummat $\overline{\sigma}(f,X_n)$, missä X_n on välin $[0,1]$ tasavälinen jako 2^n osaväliin, näytä, että $\int_0^1 f(x)\,dx = 0$.

2. a) Todista: Jos f on integroituva välillä $[-a,a]$, niin pätee

$$\begin{aligned} f \text{ parillinen} \quad &\Rightarrow \quad \int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx, \\ f \text{ pariton} \quad &\Rightarrow \quad \int_{-a}^{a} f(x)\,dx = 0. \end{aligned}$$

   b) Todista: Jos f on L-jaksoinen ja interoituva välillä $[0,L]$, niin f on integroituva välillä $[a,a+L]$ jokaisella $a \in \mathbb{R}$ ja

$$\int_a^{a+L} f(x)\,dx = \int_0^L f(x)\,dx.$$

3. Olkoon $f(x) = \begin{cases} 1/x, & \text{kun } x > 0 \\ 0, & \text{kun } x = 0 \end{cases}$
   a) Totea, että f on määritelty välillä $[0,1]$ mutta ei rajoitettu.
   b) Näytä, että f ei ole Riemann-integroituva välillä $[0,1]$ konstruoimalla jono Riemannin summia, jolle pätee $h_{X_n} \to 0$ ja $\sigma(X_n,\Xi_{X_n}) \to \infty$.

4. Näytä, että jos f on integroituva jokaisella välillä $[a,b] \subset \mathbb{R}$ ja integraali jokaisen välin yli $= 0$, niin $f(x) = 0$ jokaisessa pisteessä, jossa f on joko vasemmalta tai oikealta jatkuva.

5. Näytä Proposition VIII.5.9 avulla, että mikään Riemannin alasumma ei voi olla suurempi kuin yläsumma, ts. kaikille välin $[a,b]$ jaoille X, X' pätee $\underline{\sigma}(f,X) \le \overline{\sigma}(f,X')$.

6. (*) Näytä, että Määritelmä VIII.5.3 ei muutu, jos välin $[a, b]$ jaot X_n rajaarvossa 'Lim' (Määritelmä VIII.5.1) rajoitetaan tasavälisiksi. Voidaanko myös välipisteistöjen Ξ_{X_n} valintaa rajoittaa (esim. $\Xi_{X_n} \subset X_n$) Riemann-integroituvuuden määritelmän muuttumatta?

7. (*) Näytä, että jos f on määritelty välillä $[a, b]$ mutta ei ole rajoitettu, niin on olemassa jono Riemannin summia, jolle pätee $h_{X_n} \to 0$ ja $|\sigma(X_n, \Xi_{X_n})| \to \infty$ (vrt. Harj.teht. 3). — Mikä Määritelmän VIII.5.3 oletus on siis tarpeeton?

8. (*) Olkoon X suppenevan reaalilukujonon termeistä muodostettu (äärellinen tai numeroituva) joukko, olkoon f määritelty ja rajoitettu välillä $[a, b]$ ja olkoon g määritelty välillä $[a, b]$ siten, että $g(x) = f(x)$, kun $x \notin X$. Joukosta $\{g(x) \mid x \in X \cap [a, b]\} \subset \mathbb{R}$ tiedetään ainoastaan, että se on rajoitettu. Näytä, että jos f on välillä $[a, b]$ Riemann-integroituva, niin samoin on g ja $\int_a^b g(x)\,dx = \int_a^b f(x)\,dx$.

9. (*) a) Olkoon f määritelty ja rajoitettu välillä $[a, b]$ ja $\mathcal{X}$ välin $[a, b]$ jakojen joukko. Liitetään jokaiseen $X = \{x_k,\ k = 0 \ldots n\} \in \mathcal{X}$ luku

$$\delta(f, X) = \sum_{k=1}^{n} (\overline{f}_k - \underline{f}_k),$$

missä $\overline{f}_k$ ja $\underline{f}_k$ ovat f:n supremum ja infimum osavälillä $[x_{k-1}, x_k]$. Näytä, että jos joukko $\{\delta(f, X) \mid X \in \mathcal{X}\}$ on rajoitettu, niin f on Riemann-integroituva välillä $[a, b]$.

b) Näytä että jos f on määritelty ja monotoninen (kasvava tai vähenevä) välillä $[a, b]$, niin f on Riemann-integroituva välillä $[a, b]$.

10. (*) a) Näytä, että jos f on määritelty, rajoitettu ja Riemann-integroituva välillä $[a, b]$, niin myös seuraavat funktiot ovat Riemann-integroituvia välillä $[a, b]$:

$$f_+(x) = \max\{f(x), 0\}, \quad f_-(x) = \min\{f(x), 0\}.$$

b) Näytä, että seuraavista väittämistä ensimmäinen on tosi, toinen epätosi:

$$\begin{aligned} f &\text{ Riemann-integroituva} &\Rightarrow\ & |f| \text{ Riemann-integroituva,} \\ |f| &\text{ Riemann-integroituva} &\Rightarrow\ & f \text{ Riemann-integroituva.} \end{aligned}$$

c) Näytä, että jos f ja g ovat välillä $[a, b]$ määriteltyjä, rajoitettuja ja Riemann-integroituvia, niin samoin on tulo fg. *Vihje*: Kirjoita $f = f_+ + f_-$, $g = g_+ + g_-$ (ks. a)-kohta).

VIII.6 Analyysin peruslause

Palataan Luvussa VIII.4 tarkasteltuun probleemaan (P), jossa etsittiin annetun funktion f integraalifunktiota $y(x)$ välillä (a, b) lisäehdoilla, että $y(x)$ on jatkuva välillä $[a, b]$ ja $y(a) = 0$. Kysymys probleeman ratkeavuudesta jätettiin tuolloin avoimeksi. Näytetään nyt, että riittävä ehto ratkeavuudelle on, että funktio f on jatkuva välillä $[a, b]$. Kyseessä on eräs matemaattisen analyysin perustavimmista tuloksista, ja se tunnetaankin nimellä *Analyysin peruslause*†.

*Lause VIII.6.1 **(Analyysin peruslause)** Jos f on jatkuva välillä $[a, b]$, niin

1. f on Riemann-integroituva välillä $[a, b]$.
2. Funktio $F(x) = \int_a^x f(t)\,dt$ on jatkuvasti derivoituva välillä $[a, b]$ ja $F'(x) = f(x)$ ko. välillä.

Seuraavassa todistetaan Lauseen VIII.6.1 ensimmäinen (vaativampi) osaväittämä heikennetyssä muodossa, uudistamalla Luvussa VIII.4 tehty Lipschitz-oletus

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \quad \forall x_1, x_2 \in [a, b].$$

Lause VIII.6.2 Välillä $[a, b]$ Lipschitz-jatkuva funktio on Riemann-integroituva.

Todistus Olkoon $X_n = \{x_k,\ k = 0 \ldots n\}$ välin $[a, b]$ jako. Koska f on (ainakin) jatkuva välillä $[a, b]$, niin f saavuttaa väleillä $[x_{k-1}, x_k]$ maksimi- ja minimiarvonsa (LauseV.1.12), joten jakoon X_n liittyvät Riemannin ylä- ja alasummat (ks. edellinen luku) voidaan kirjoittaa

$$\overline{\sigma}(f, X_n) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}), \qquad \underline{\sigma}(f, X_n) = \sum_{k=1}^{n} f(\eta_k)(x_k - x_{k-1}),$$

missä $\xi_k, \eta_k \in [x_{k-1}, x_k]$. Tämän ja oletetun Lipschitz-jatkuvuuden perusteella on

$$\begin{aligned} 0 \leq \overline{\sigma}(f, X_n) - \underline{\sigma}(f, X_n) &= \sum_{k=1}^{n} [f(\xi_k) - f(\eta_k)](x_k - x_{k-1}) \\ &\leq \sum_{k=1}^{n} L|\xi_k - \eta_k|(x_k - x_{k-1}) \\ &\leq \sum_{k=1}^{n} L h_{X_n}(x_k - x_{k-1}) = L(b - a)h_{X_n}. \end{aligned}$$

†Englanninkielinen nimi on 'The Fundamental Theorem of Calculus'. Lauseen muotoilut kirjallisuudessa ovat vaihtelevia. — Usein peruslauseena esitetään vain Lauseen VIII.6.1 toinen osaväittämä, jolloin lause jää sisällöltään melko kevyeksi, ks. Lause VIII.6.3 jäljempänä.

On päätelty, että jos $\{X_n\}$ on jono jakoja, jolle $h_{X_n} \to 0$, niin

$$\overline{\sigma}(f, X_n) - \underline{\sigma}(f, X_n) \;=\; \mathcal{O}(h_{X_n}) \;\to\; 0.$$

Tämän ja Lauseiden VIII.5.10 ja VIII.5.8 perusteella seuraa väite. □

Analyysin peruslauseen 1. väittämän mukaisesti Lause VIII.6.2 on tosi myös, jos f on pelkästään jatkuva välillä $[a, b]$. Todistuksen pääidea on tässäkin tapauksessa sama, mutta todistuksen on nojattava syvällisempään jatkuvuuden logiikkaan, tarkemmin sanoen <u>tasaiseen</u> <u>jatkuvuuteen</u> (ks. Luku V.9). Todettakoon tässä ainoastaan, että pelkän jatkuvuusoletuksen perusteella em. todistuskonstruktion loppupäätelmäksi tulee

$$\overline{\sigma}(f, X_n) - \underline{\sigma}(f, X_n) = o(1) \;\to\; 0, \quad \text{kun } h_{X_n} \to 0.$$

Väitetty integroituvuus (Analyysin peruslauseen 1. väittämä) tulee siis näinkin todistetuksi.

Analyysin peruslauseen 2. väittämän todistamiseksi tarkastellaan (sovelluksiakin silmällä pitäen, ks. esimerkki jäljempänä) yleisempää tilannetta, jossa funktio f on integroituva välillä $[a, b]$ mutta ei välttämättä jatkuva koko välillä $[a, b]$. Tällöinkin funktio $F(x) = \int_a^x f(t)\,dt$ on määritelty koko välillä $[a, b]$ (Lause VIII.5.11). Seuraavan lauseen toinen väittämä todistaa Lauseen VIII.6.1 toisen osaväittämän, joten Analyysin peruslause tulee samalla kokonaaan todistetuksi.

Lause VIII.6.3 Olkoon f määritelty, rajoitettu ja Riemann-integroituva välillä $[a, b]$ ja $F(x) = \int_a^x f(t)\,dt,\ x \in [a, b]$. Tällöin pätee:

1. Jos $|f(x)| \le M\ \forall x \in [a, b]$, niin F on välillä $[a, b]$ Lipschitz-jatkuva vakiolla $L = M$.

2. Jos f:llä on oikeanpuoleinen raja-arvo $f(x^+)$ pisteessä $x \in [a, b)$, niin F on pisteessä x oikealta derivoituva ja $D_+F(x) = f(x^+)$. Vastaavasti jos f:llä on vasemmanpuoleinen raja-arvo $f(x^-)$ pisteessä $x \in (a, b]$, niin F on vasemmalta derivoituva pisteessä x ja $D_-F(x) = f(x^-)$.

<u>Todistus</u> 1. Jos $-M \le f(x) \le M,\ x \in [a, b]$, niin integraalien vertailuperiaatteen (Lause VIII.5.6) nojalla pätee jokaisella osavälillä $[x_1, x_2] \subset [a, b]$ $(x_1 < x_2)$

$$-M(x_2 - x_1) = \int_{x_1}^{x_2} (-M)\,dt \le \int_{x_1}^{x_2} f(t)\,dt \le \int_{x_1}^{x_2} M\,dt = M(x_2 - x_1).$$

Integraalin additiivisuuden (Lause VIII.5.4) perusteella seuraa

$$|F(x_2) - F(x_1)| = \left| \int_{x_1}^{x_2} f(t)\,dt \right| \le M|x_1 - x_2| \quad \forall x_1, x_2 \in [a, b] \;\Rightarrow\; \text{väite 1.}$$

2. Olkoon oikeanpuoleinen raja-arvo $f(x^+)$ olemassa pisteessä $x \in [a, b)$ ja olkoon $\varepsilon > 0$. Tällöin on olemassa $\delta \in (0, b - x]$ siten, että $|f(t) - f(x^+)| < \varepsilon$, kun $t \in [x, x + \delta)$ (Lause V.2.2). Tällöin jos $0 < \Delta x < \delta$, niin integraalin lineaarisuuden, additiivisuuden ja vertailuperiaatteen nojalla

$$\left|\frac{F(x+\Delta x) - F(x)}{\Delta x} - f(x^+)\right| = \left|\frac{1}{\Delta x}\int_x^{x+\Delta x} [f(t) - f(x^+)]\,dt\right|$$
$$\leq \frac{1}{\Delta x}\int_x^{x+\Delta x} |f(t) - f(x^+)|\,dt \leq \frac{1}{\Delta x}\int_x^{x+\Delta x} \varepsilon\,dt = \varepsilon.$$

Koska tässä $\varepsilon > 0$ oli mielivaltainen ja epäyhtälö on pätevä jokaisella $\Delta x \in (0, \delta)$, missä $\delta > 0$, niin F on määritelmän mukaan oikealta derivoituva pisteessä x ja $D_+F(x) = f(x^+)$. Väittämän toinen osa todistetaan vastaavasti. $\square$

Esimerkki 1 : **Liikelaki**. Jos kappaleen (massa $= m$) suoraviivaisessa liikkeessä vaikuttaa liikesuuntaan voima $f(t)$ (t=aika), niin <u>liikemäärän</u> <u>säilymislaki</u> aikavälillä $[0, \infty)$ on

$$mv(t) - mv(0) = \int_0^t f(t')dt',$$

missä $v(t) =$ kappaleen nopeus hetkellä t ja oikealla puolella oleva (Riemannin) integraali on voiman f <u>impulssi</u> aikavälillä $[0, t]$. Jos f on jatkuva pisteessä $t > 0$, niin Lauseen VIII.6.3 perusteella liikemäärän säilymislaista seuraa (puolittain derivoimalla) <u>liikeyhtälö</u>

$$mv'(t) = f(t).$$

Jos f on pisteessä t epäjatkuva (fysikaalisesti mahdollista!), ei liikeyhtälö ole (yleensä) voimassa ko. hetkellä. Liikemäärän säilymislaki sen sijaan säilyttää pätevyytensä, edellyttäen ainoastaan, että f on (Riemann-)integroituva aikaväleillä $[0, t]$. Liikemäärän säilymislakia voidaan näin ollen pitää 'alkuperäisenä' fysiikan lakina ja liikeyhtälöä pikemminkin tämän seuraamuksena mainitun lisäoletuksen (f jatkuva pisteessä t) vallitessa. Esimerkiksi jos $v(0) = v_0$ ja

$$f(t) = \begin{cases} F, & \text{kun } t \in [t_1, t_2], \\ 0, & \text{muuten,} \end{cases}$$

missä $F \neq 0$ on vakio ja $0 < t_1 < t_2$, niin liikemäärän säilymislain mukainen (fysikaalisesti oikea!) ratkaisu on

$$v(t) = \begin{cases} v_0, & \text{kun } t \in [0, t_1], \\ v_0 + \dfrac{F}{m}(t - t_1), & \text{kun } t \in [t_1, t_2], \\ v_0 + \dfrac{F}{m}(t_2 - t_1), & \text{kun } t \in [t_2, \infty). \end{cases}$$

Liikeyhtälö $mv'(t) = f(t)$ on tässä tapauksessa voimassa väleillä $(0, t_1)$, (t_1, t_2) ja (t_2, ∞). Pisteissä t_1 ja t_2 nopeus $v(t)$ on Lauseen VIII.6.3 mukaisesti jatkuva (koska f on rajoitettu) ja sekä vasemmalta että oikealta derivoituva. Koska f:llä on näissä pisteissä hyppyepäjatkuvuus, ovat toispuoliset derivaatat erisuuret, ja näin ollen $v(t)$ ei ole derivoituva pisteissä t_1 ja t_2. $\square$

Funktion $\int_{g_1(x)}^{g_2(x)} f(t)\,dt$ derivaatta

Integraalin additiivisuuden nojalla otsikon funktio voidaan esittää muodossa $F(g_2(x)) - F(g_1(x))$, missä $F(x) = \int_a^x f(t)\,dt$. Yhdistetyn funktion derivoimissäännön ja Lauseen VIII.6.3 perusteella päädytään derivoimissääntöön

$$\boxed{\frac{d}{dx}\int_{g_1(x)}^{g_2(x)} f(t)\,dt = f(g_2(x))g_2'(x) - f(g_1(x))g_1'(x)}$$

seuraavin oletuksin:

(i) $g_1(t)$ ja $g_2(t)$ ovat derivoituvia pisteessä $t = x$.

(ii) $f(t)$ on määritelty, rajoitettu ja Riemann-integroituva jollakin välillä $[a, b]$, missä $a < \min\{g_1(x), g_2(x)\}$ ja $b > \max\{g_1(x), g_2(x)\}$.

(iii) $f(t)$ on jatkuva pisteissä $g_1(x)$ ja $g_2(x)$.

Esimerkki 2 Jokaisella $x \neq 0$ (myös kun $x = 1$ (!)) pätee

$$\frac{d}{dx}\int_x^{1/x} \frac{e^t}{t}\,dt = -\frac{1}{x^2}\left(x\,e^{1/x}\right) - \frac{e^x}{x} = -\frac{1}{x}(e^x + e^{1/x}),$$

sillä oletukset (i)–(iii) ovat voimassa. Samasta syystä pätee jokaisella $x > 0$

$$\frac{d}{dx}\int_x^{x^2} \frac{e^t}{t}\,dt = 2x \cdot \frac{e^{x^2}}{x^2} - \frac{e^x}{x} = \frac{1}{x}(2e^{x^2} - e^x).$$

Sen sijaan jos $x \leq 0$, niin tämä lasku ei ole pätevä, koska tällöin ei oletus (ii) (tapauksessa $x = 0$ ei myöskään (iii)) ole voimassa. $\square$

Osittaisintegrointi ja sijoitus määrätyssä integraalissa

Määrätylle integraalille pätevät riittävin säännöllisyysoletuksin seuraavat laskukaavat, vrt. vastaavat määräämättömän integraalin kaavat Luvussa VIII.2.

1. Osittaisintegrointikaava

$$\int_a^b f'(x)g(x)\,dx = \Big|_a^b f(x)g(x) - \int_a^b f(x)g'(x)\,dx.$$

2. Sijoituskaava

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(u(t))u'(t)\,dt, \quad u(\alpha) = a,\ u(\beta) = b.$$

Osittaisintegrointikaava on pätevä silloin kun f ja g ovat välillä $[a, b]$ jatkuvasti derivoituvia. Tällöin $f'g$ ja fg' ovat jatkuvia välillä $[a, b]$, jolloin Analyysin peruslauseen ja integraalin lineaarisuuden perusteella pätee

$$\begin{aligned}\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx &= \int_a^b [f'(x)g(x) + f(x)g'(x)]\,dx \\ &= \int_a^b \frac{d}{dx} f(x)g(x)\,dx = \Big|_a^b f(x)g(x).\end{aligned}$$

Määrätyn integraalin sijoituskaava vaatii hieman pitemmät perustelut. Olkoon esimerkiksi $\alpha < \beta$ (kaavan pätevyys ei tätä edellytä). Tällöin on oletettava

(i) u on jatkuvasti derivoituva välillä $[\alpha, \beta]$,

(ii) f on jatkuva välillä $[A, B]$, jolle pätee $t \in [\alpha, \beta] \ \Rightarrow\ u(t) \in [A, B]$.

Kun merkitään

$$F(x) = \int_a^x f(s)ds, \quad x \in [A, B],$$

niin Analyysin peruslauseen mukaan on

$$F'(x) = f(x), \quad x \in [A, B].$$

Tällöin on oletuksien (i)-(ii) perusteella

$$\frac{d}{dt}F(u(t)) = f(u(t))u'(t), \quad t \in [\alpha, \beta],$$

joten

$$\begin{aligned}\int_\alpha^\beta f(u(t))u'(t)\,dt &= \Big|_\alpha^\beta F(u(t)) \\ &= F(u(\beta)) - F(u(\alpha)) \\ &= F(b) - F(a) = \int_a^b f(x)\,dx.\end{aligned}$$

Tapauksessa $\alpha > \beta$ on puhuttava välistä $[\beta, \alpha]$, muuten perustelut ovat samat. Huomattakoon, että koska funktion u injektiviisyttä ei vaadittu, eivät α ja β välttämättä ole yksikäsitteiset, vrt. kuvio.

Esimerkki 3 Kun integraalissa

$$\int_0^2 \frac{e^{-x}}{1+\sqrt{x}}\,dx$$

tehdään sijoitus

$$\sqrt{x} = t \geq 0 \quad \Leftrightarrow \quad x = u(t) = t^2,$$

niin

$$dx = u'(t)dt = 2tdt$$

ja

$$u(\alpha) = 0 \;\Rightarrow \alpha = 0, \quad u(\beta) = 2 \;\Rightarrow\; \beta = \sqrt{2},$$

joten integraali saa muodon

$$\int_0^2 \frac{e^{-x}}{1+\sqrt{x}}\,dx = \int_0^{\sqrt{2}} \frac{e^{-t^2}}{1+t}\,2t\,dt = 2\int_0^{\sqrt{2}} \frac{te^{-t^2}}{1+t}\,dt. \quad \square$$

Esimerkin muunnetussa integraalissa (toisin kuin alkuperäisessä) integroitava funktio on säännöllinen (sileä) koko integroimisvälillä. Integroitavan funktion säännöllisyys on yleisesti eduksi silloin, kun integraalin arvo lasketaan numeerisilla menetelmillä, ks. Luku VIII.9 jäljempänä.

Integraalilaskun väliarvolause

Seuraava lause on väliarvolauseiden sarjan kolmas ja viimeinen, vrt. Lauseet V.1.10 ja V.6.1. Verrattuna aiempiin väliarvolauseisiin tämä lause ei ole kovin itsenäinen, sillä se seuraa helposti Analyysin peruslauseesta ja Differentiaalilaskun väliarvolauseesta (Harj.teht. 8a).

Lause VIII.6.4 **(Integraalilaskun väliarvolause)** Jos f on jatkuva välillä $[a, b]$, niin jollakin $\xi \in (a, b)$ on

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

Seuraavaa Lauseen VIII.6.4 yleisempää (myös itsenäisempää) muotoa sanotaan integraalilaskun *yleistetyksi* väliarvolauseeksi.

Lause VIII.6.5 Jos f on jatkuva välillä $[a, b]$, g on Riemann–integroituva välillä $[a, b]$ ja joko $g(x) \geq 0$ tai $g(x) \leq 0$ koko välillä $[a, b]$, niin jollakin $\xi \in [a, b]$ on

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx.$$

<u>Todistus</u> Jos $g(x) \geq 0$ ja $f_{\min}$ ja $f_{\max}$ ovat f:n minimi- ja maksimiarvot välillä $[a, b]$, niin päätellään

$$\begin{aligned}
&f_{\min}\,g(x) \leq f(x)g(x) \leq f_{\max}\,g(x), \quad x \in [a,b] \\
&\Rightarrow\ f_{\min}\int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq f_{\max}\int_a^b g(x)\,dx \\
&\Rightarrow\ \int_a^b f(x)g(x)\,dx = \eta\int_a^b g(x)\,dx, \quad \eta \in [f_{\min}, f_{\max}].
\end{aligned}$$

Ensimmäisen väliarvolauseen (Lause V.1.10) mukaan on tässä $\eta = f(\xi)$ jollakin $\xi \in [a, b]$, joten väite seuraa. Tapauksessa $g(t) \leq 0$ voidaan käyttää jo todistettua väittämää, kun g:n tilalla on $-g$. $\square$

Funktion f *keskiarvoksi* (engl. average value) välillä $[a, b]$ sanotaan lukua

$$K(f) = \frac{1}{b-a} \int_a^b f(x)\, dx.$$

Jos $g(x) \geq 0 \; \forall x \in [a, b]$ ja $\displaystyle\int_a^b g(x)\, dx > 0$, niin lukua

$$K_g(f) = \frac{\int_a^b f(x)g(x)\, dx}{\int_a^b g(x)\, dx}$$

sanotaan f:n *painotetuksi keskiarvoksi* (engl. weighted average) välillä $[a, b]$ ja funktiota g tällöin *painofunktioksi.* Lauseen VIII.6.5 mukaan välillä $[a, b]$ jatkuvalle funktiolle pätee $K_g(f) = f(\xi)$ jollakin $\xi \in [a, b]$.

Taylorin lauseen integraalimuoto

Monissa sovelluksissa (esim. numeeristen menetelmien virhettä arvioitaessa, ks. Harj.teht. 11) seuraava Taylorin lauseen vaihtoehtoinen *integraalimuoto* on hyvin kätevä — vrt. Lause VII.4.3, jota kutsutaan Taylorin lauseen *väliarvomuodoksi.*

Lause VIII.6.6 **(Taylorin lause – integraalimuoto)** Jos f on $n+1$ kertaa jatkuvasti derivoituva välillä $[x_0, x]$ $(x > x_0)$ tai $[x, x_0]$ $(x < x_0)$ ja $T_n(x, x_0)$ on f:n Taylorin polynomi astetta n pisteessä x_0, niin

$$f(x) - T_n(x, x_0) = \int_{x_0}^{x} \frac{1}{n!}(x-t)^n f^{(n+1)}(t)\, dt.$$

<u>Todistus</u> Lähdetään ilmeisestä identiteetistä

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\, dt,$$

joka on sama kuin lauseen väittämä tapauksessa $n = 0$. Jos tässä f on kahdesti jatkuvasti derivoituva integroimisvälillä (oletus tapauksessa $n = 1$), voidaan integroida osittain seuraavasti:

$$\begin{aligned} f(x) &= f(x_0) - \Big|_{x_0}^{x} (x-t)f'(t) + \int_{x_0}^{x} (x-t)f''(t)\, dt \\ &= f(x_0) + f'(x_0)(x - x_0) + \int_0^x (x-t)f''(t)\, dt. \end{aligned}$$

Näin on todistettu väittämä tapauksessa $n = 1$. Jatkamalla osittaisintegrointia nähdään vastaavasti, että väittämä on tosi, kun $n = 2$, jne. (Yleinen todistus: Harj.teht. 8b.) $\square$

HARJOITUSTEHTÄVIÄ

1. Jos f jatkuva välillä $[a, b]$, niin mikä on funktion

$$g(x) = \int_a^b [f(t) - x]^2 \, dt$$

pienin arvo ja missä se saavutetaan?

2. Laske:

a) $\dfrac{d}{dx} \displaystyle\int_2^x \frac{\sin t}{t} \, dt$ b) $\dfrac{d}{dt} \displaystyle\int_t^3 \frac{\sin x}{x} \, dx$ c) $\dfrac{d}{dx} \displaystyle\int_{x^2}^x \frac{\sin t}{t} \, dt$

d) $\dfrac{d}{dx} x^2 \displaystyle\int_0^{x^2} \frac{\sin u}{u} \, du$ e) $\dfrac{d}{dt} \displaystyle\int_{-t^{-1}}^t \frac{\cos y}{1 + y^2} \, dy$

f) $\dfrac{d}{d\theta} \displaystyle\int_{\sin\theta}^{\cos\theta} \frac{1}{1 - x^2} \, dx$ g) $\dfrac{d}{dx} F(\sqrt{x}), \quad F(t) = \displaystyle\int_0^t \cos(x^2) \, dx$

h) $H'(2), \quad H(x) = 3x \displaystyle\int_4^{x^2} e^{-\sqrt{t}} \, dt$

3. Ratkaise $y(x)$:

a) $y(x) = 1 - \displaystyle\int_0^x y(t) \, dt$ b) $y(x) = \pi \left(1 + \displaystyle\int_1^x y(t) \, dt \right)$

4. Laske integroimalla osittain kohdissa a)–f) ja sijoituksella kohdissa g)–l):

a) $\displaystyle\int_0^1 x e^{-x} \, dx$ b) $\displaystyle\int_0^\pi x \cos x \, dx$ c) $\displaystyle\int_0^\pi x \sin x \, dx$

d) $\displaystyle\int_0^2 \sqrt{x} \ln x \, dx$ e) $\displaystyle\int_0^1 \operatorname{Arctan} x \, dx$ f) $\displaystyle\int_0^\pi e^{-x} \sin x \, dx$

g) $\displaystyle\int_{-2}^2 \frac{1}{2x^2 + 4x + 3} \, dx$ h) $\displaystyle\int_0^{\ln 2} \frac{e^x}{1 + e^x} \, dx$ i) $\displaystyle\int_1^4 \frac{\sqrt{x}}{\sqrt[4]{x} + 1} \, dx$

j) $\displaystyle\int_0^1 \sqrt{\frac{1 - x}{1 + x}} \, dx$ k) $\displaystyle\int_0^{\pi/4} \frac{1}{1 + \tan x} \, dx$ l) $\displaystyle\int_0^\pi \frac{\sin x}{2 - \cos x} \, dx$

5. Olkoon $R > 0$ ja $ab \neq 0$. Laske

$$\text{a)} \quad \int_0^R x^2\sqrt{R^2 - x^2}\,dx, \qquad \text{b)} \quad \int_0^{\pi/2} \frac{1}{a^2\cos^2 x + b^2 \sin^2 x}\,dx.$$

6. Yksinkertaista seuraavat funktiolausekkeet sijoitusta käyttäen ($x > 0$).

$$\text{a)} \quad f(x) = \int_0^x \sqrt{t}\,\sqrt{x^2 - t^2}\,dx \qquad \text{b)} \quad f(x) = \int_{-x/2}^{x/2} \frac{1}{(x^2 - t^2)^{3/2}}\,dt$$

7. Laske funktion $f(x) = \sin^2 x$ a) keskiarvo, b) funktiolla $g(x) = x$ painotettu keskiarvo välillä $[0, \pi]$.

8. a) Todista Integraalilaskun väliarvolause soveltamalla Differentiaalilaskun väliarvolausetta funktioon $F(x) = \int_a^x f(t)\,dt$. b) Todista Lause VIII.6.6 induktiolla. c) Johda Lauseesta VIII.6.6 Taylorin lauseen väliarvomuoto (Lause VII.4.3) käyttämällä hyväksi Lausetta VIII.6.5.

9. (*) Onko funktiolla $F(x) = \displaystyle\int_0^{2x-x^2} \cos\left(\frac{1}{1+t^2}\right) dt$ suurin tai pienin arvo? Perustele!

10. (*) Näytä oikeaksi *osittaissummauksen* kaava

$$\sum_{k=1}^{n} f_k g_k = \Big|_{k=1}^{k=n} F_k g_k - \sum_{k=1}^{n-1} F_k\,(g_{k+1} - g_k), \qquad F_k = \sum_{i=1}^{k} f_i.$$

Mikä on vastaava kaava määrätylle integraalille?

11. (*) Olkoon f kolmesti jatkuvasti derivoituva välillä $[a - h, a + h]$. Johda Lauseen VIII.6.6 avulla keskeisdifferenssiapproksimaation virhekaava

$$f'(a) - \frac{f(a+h) - f(a-h)}{2h} = -\frac{1}{h}\int_{a-h}^{a+h} k(t) f'''(t)\,dt,$$

$$\text{missä} \quad k(t) = \begin{cases} \frac{1}{2}(a - h - t)^2, & \text{kun } t \in [\,a - h, a\,], \\ \frac{1}{2}(a + h - t)^2, & \text{kun } t \in [\,a, a + h\,]. \end{cases}$$

Johda tästä edelleen virhekaavalle väliarvomuoto (vrt. Propositio VII.5.3)

$$f'(a) - \frac{f(a+h) - f(a-h)}{2h} = -\frac{1}{6}\,h^2 f'''(\xi), \qquad \xi \in [a - h, a + h].$$

VIII.7 Riemannin integraalin laajennukset

Riemannin integraalin määritelmässä on perusoletuksena, että integroitava funktio f on sekä <u>määritelty</u> että <u>rajoitettu</u> koko integroimisvälillä, jonka on oltava (äärellinen) <u>suljettu</u> väli. Ilman näitä rajoituksia ei integraalia voi yleisesti määritellä Riemannin summien (tai ylä- ja alasummien) raja-arvona, ts. rajoitukset ovat välttämättömiä, jotta f olisi Riemann-integroituva. Monien sovellusten kannalta kuitenkin mainitut rajoitukset ovat turhan voimakkaita tai jopa keinotekoisia. Esimerkiksi koska tiedetään, että integraalin arvo ei riipu f:n arvoista yksittäisessä (tai äärellisen monessa) pisteessä, niin tuntuu turhalta ylipäänsä vaatia, että f on määritelty jokaisessa integroimisvälin pisteessä. Tällaisten turhien rajoitusten poistamiseksi on tullut tavaksi konstruoida Riemannin integraalin määritelmälle erilaisia laajennuksia, joita sanotaan *epäoleellisiksi* (engl. improper, kirjaimellisesti 'sopimaton' tai 'hyvien tapojen vastainen') Riemannin integraaleiksi. Laajennukset eivät johda kovin selkeään yleiseen integraalin määritelmään, vaan niiden tarkoituksena on lähinnä 'paikata' alkuperäistä määritelmää niin, että tavallisimmat sovelluksissa esiintyvät tapaukset tulevat katetuiksi. Jatkossa esitetään näistä laajennuksista tavallisimmat.†

Jos f on rajoitettu ja Riemann-integroituva välillä $[a, b]$, niin funktion

$$F(x) = \int_a^x f(t)\,dt$$

Lipschitz-jatkuvuuden (Lause VIII.6.3) perusteella

$$\int_{a+\varepsilon}^b f(t)\,dt = F(b) - F(a+\varepsilon) = \int_a^b f(t)\,dt + \mathcal{O}(\varepsilon),$$

kun $0 < \varepsilon < b - a$, joten

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b f(x)\,dx.$$

Oikealla oleva raja-arvo ei riipu f:n arvoista pisteessä a — itse asiassa raja-arvo ei riipu edes siitä, onko f määritelty tässä pisteessä. Tämän perusteella on varsin luontevaa määritellä integraali ko. raja-arvona (mikäli olemassa) silloinkin, kun f ei ole pisteessä a määritelty ja/tai f ei ole rajoitettu ko. pistettä lähestyttäessä.

†Riemannin integraalin ongelmat havaittiin jo 1800-luvun jälkipuoliskolla, jolloin mittojen ja integraalien teoria kehittyi voimakkaasti. Lopullisen ratkaisun ongelmaan toi ranskalaisen **Henri Lebesgue**n (1875-1941) vuonna 1906 esittämä määritelmä, joka on sittemmin tunnettu *Lebesguen integraalina*. Lebesguen integraali poistaa Riemannin integraalin kauneusvirheet, mutta sen määrittely vaatii syvällisempiä mittateoreettisia pohdiskeluja.

Esimerkki 1 Jos $f(x) = 1/\sqrt{x-a}$, niin f ei ole välillä $[a,b]$ rajoitettu (eikä pisteessä a määritelty), mutta em. raja-arvo on olemassa:

$$\int_{a+\varepsilon}^{b} \frac{1}{\sqrt{x-a}}\,dx \;=\; \Big|_{x=a+\varepsilon}^{x=b} 2\sqrt{x-a} \;=\; 2\sqrt{b-a} - 2\sqrt{\varepsilon}$$
$$\to\; 2\sqrt{b-a}, \quad \text{kun } \varepsilon \to 0^+. \quad \square$$

Kun em. raja-arvomenettelyä sovelletaan integroimisvälin kummassakin päätepisteessä, tullaan seuraavaan määritelmään.

Määritelmä VIII.7.1 **(Riemannin integraalin 1. laajennus)** Olkoon $a < c < b$. Jos f on rajoitettu ja Riemann-integroituva väleillä $[a+\varepsilon, c]$ ja $[c, b-\varepsilon]$ aina kun $\varepsilon > 0$ ($\varepsilon < \min\{c-a, b-c\}$), niin määritellään

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^{c} f(x)\,dx + \lim_{\varepsilon\to 0^+} \int_c^{b-\varepsilon} f(x)\,dx,$$

sikäli kuin oikealla puolella olevat raja-arvot ovat olemassa. Sanotaan tällöin, että f on integroituva välillä $[a,b]$.

Määritelmän voi esittää hieman lyhyemmin muodossa

$$\int_a^b f(x)\,dx = \lim_{\varepsilon_1,\varepsilon_2\to 0^+} \int_{a+\varepsilon_1}^{b-\varepsilon_2} f(x)\,dx,$$

missä raja-arvomerkintä oikealla tarkoittaa, että $\varepsilon_1 \to 0^+$ ja $\varepsilon_2 \to 0^+$ toisistaan riippumatta. Jos oletetaan, että f:llä on välillä (a,b) integraalifunktio F, niin noudattaen lyhennettyä merkintätapaa on Määritelmän VIII.7.1 mukaisesti

$$\int_a^b f(x)\,dx = \lim_{\varepsilon_1,\varepsilon_2\to 0^+} \Big|_{a+\varepsilon_1}^{b-\varepsilon_2} F(x) = F(a^+) - F(b^-).$$

Siis f on tässä tapauksessa integroituva täsmälleen kun raja-arvot $F(a^+)$ ja $F(b^-)$ ovat olemassa eli kun F on jatkuva välillä $[a,b]$.

Esimerkki 1 (jatko) Funktiolla $f(x) = 1/\sqrt{x-a}$ on välillä $[a,b]$ jatkuva integraalifunktio $F(x) = 2\sqrt{x-a}$. Integraalin arvoa laskettaessa on käytännössä turvallista (ja hyväksyttyä) ohittaa Määritelmän VIII.7.1 raja-arvoprosessi eli laskea suoraan

$$\int_a^b \frac{1}{\sqrt{x-a}}\,dx \;=\; \Big|_{x=a}^{x=b} 2\sqrt{x-a} \;=\; 2\sqrt{b-a}. \quad \square$$

Koska integraalin additiivisuutta on luonnollista pitää myös em. tavalla laajennettujen (epäoleellisten) integraalien ominaisuutena, saadaan välittömästi seuraava laajennus.

MÄÄRITELMÄ VIII.7.2 **(Riemannin integraalin 2. laajennus)** Jos $a = c_0 < c_1 < \ldots < c_n = b$ ja f on integroituva väleillä $[c_{k-1}, c_k]$, $k = 1 \ldots n$, niin f on integroituva välillä $[a, b]$ ja

$$\int_a^b f(x)\,dx = \sum_{k=1}^{m} \int_{c_{k-1}}^{c_k} f(x)\,dx.$$

Tämän määritelmän mukaisesti on esimerkiksi välillä $[a, b]$ paloittain jatkuva funktio ko. välillä integroituva. Määritelmän mukaisesti integraalin arvo on tällöin riippumaton siitä, miten f on määritelty (tai onko lainkaan määritelty) epäjatkuvuuspisteissä.

ESIMERKKI 2 Jos $f(x) = k$, kun $k-1 < x < k$, $k = 1 \ldots n$ ($f(k)$ joko määritelty tai ei, kun $k = 0 \ldots n$), niin Määritelmän VIII.7.2 mukaisesti

$$\int_0^n f(x)\,dx \;=\; \sum_{k=1}^{n} \int_{k-1}^{k} f(x)\,dx \;=\; \sum_{k=1}^{n} k \;=\; \frac{1}{2}n(n+1). \quad \square$$

Lopuksi laajennetaan integraalin käsite koskemaan äärettömiä välejä $[a, \infty)$, $(-\infty, b]$ ja $(-\infty, \infty)$:

MÄÄRITELMÄ VIII.7.3 **(Riemannin integraalin 3. laajennus)** Jos f on integroituva väleillä $[a, M]$ jokaisella $M > a$ ja on olemassa raja-arvo

$$A = \lim_{M\to\infty} \int_a^M f(x)\,dx,$$

niin sanotaan, että f on integroituva välillä $[a, \infty)$, lukua A sanotaan f:n integraaliksi ko. välillä, ja merkitään

$$A = \int_a^\infty f(x)\,dx.$$

Vastaavasti määritellään

$$\int_{-\infty}^b f(x)\,dx = \lim_{M\to\infty} \int_{-M}^b f(x)\,dx,$$
$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx, \quad c \in \mathbb{R}.$$

Esimerkki 3 Funktio $f(x) = 1/(1+x^2)$ on integroituva välillä $(-\infty, \infty)$, sillä Määritelmän VIII.7.3 mukaisesti

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{1+x^2} &= \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx + \int_{0}^{\infty} \frac{1}{1+x^2}\,dx \\ &= \lim_{M\to\infty} \int_{-M}^{0} \frac{1}{1+x^2}\,dx + \lim_{M\to\infty} \int_{0}^{M} \frac{1}{1+x^2}\,dx \\ &= \lim_{M\to\infty} \Big|_{-M}^{0} \operatorname{Arctan} x + \lim_{M\to\infty} \Big|_{0}^{M} \operatorname{Arctan} x = \frac{\pi}{2} + \frac{\pi}{2} = \pi. \quad \square\end{aligned}$$

Jos f on tarkasteltavalla (äärellisellä tai äärettömällä) välillä integroituva joko tavanomaisessa tai laajennetussa mielessä, niin sanotaan, että f:n integraali ko. välin yli *suppenee*. Jos f ei ole integroituva, niin sanotaan, että integraali *hajaantuu*. Kuten esimerkissä, integraalin suppeneminen selviää helposti integraalifunktion avulla, sikäli kuin sellainen on integroimsvälillä löydettävissä. Nimittäin Määritelmän VIII.7.3 mukaan myös äärettömillä integroimisväleillä $[a, \infty)$, $(-\infty, b]$ ja $(-\infty, \infty)$ pätee: jos f:n integraalifunktio välillä (a, b) on F, niin integraalin $\int_a^b f(x)\,dx$ arvo (sikäli kuin olemassa) on

$$\int_a^b f(x)\,dx = \lim_{x\to b^+} F(x) - \lim_{x\to a^-} F(x),$$

eli integraali suppenee täsmälleen kun molemmat raja-arvot oikealla puolella ovat olemassa. Mikäli koko integroimisvälillä toimivaa integraalifunktiota ei löydy, voidaan väli aina jakaa äärelliseen määrään osavälejä, käyttää integraalin additiivisuutta ja tutkia integraalin suppeneminen kullakin osavälillä erikseen.

Esimerkki 4 Raja-arvoissa hieman 'oikaisten' voidaan laskea

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \Big|_{-\infty}^{\infty} \operatorname{Arctan} x = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi \quad \text{(suppenee)},$$

$$\int_0^1 \frac{1}{x} = \Big|_0^1 \ln x = \infty \quad \text{(hajaantuu)}.$$

Varomaton on sen sijaan lasku

$$\int_{-\infty}^{\infty} \frac{1}{x^2}\,dx = \Big|_{-\infty}^{\infty} -\frac{1}{x} = 0,$$

sillä huomaamatta jäi, että $F(x) = -1/x$ ei ole funktion $f(x) = 1/x^2$ integraalifunktio koko välillä $(-\infty, \infty)$. Jakamalla integraali kahteen osaan väleille $(-\infty, 0]$ ja $[0, \infty)$ nähdään, että se hajaantuu. $\square$

Epäoleellisten integraalien vertailu

Silloin kun epäoleellisen integraalin suppenevuutta ei voida suoraan tutkia integraalifunktion avulla, tai se on hankalaa, voidaan suppenevuus yleensä selvittää vertaamalla integraalia yksinkertaisempaan, laskettavissa olevaan integraaliin. Vertailussa lähdetään ennestään tutusta vertailuperiaatteesta (Lause VIII.5.6)

$$f(x) \leq g(x) \quad \forall x \in [a,b] \;\Rightarrow\; \int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Sovellettaessa tätä epäoleellisiin integraaleihin äärellisellä välillä $[a,b]$ riittää olettaa, että $f(x) \leq g(x)$ on voimassa avoimella välillä (a,b) (Määritelmä VIII.7.1) tai avoimilla osaväleillä $(c_{k-1}, c_k) \subset (a,b),\ k = 1 \ldots n$ (Määritelmä VIII.7.2), sillä koska vertailuperiaate on pätevä tavallisille Riemannin integraaleille, se pysyy voimassa myös määritelmien raja-arvoille. Vertailtaessa integraaleja $\int_a^\infty f(x)\,dx$ ja $\int_a^\infty g(x)\,dx$ (Määritelmä VIII.7.3) riittää vastaavasti olettaa, että $f(x) \leq g(x)$ väleillä $(c_{k-1}, c_k),\ k = 1, 2, \ldots$ missä $a = c_0 < c_1 < \ldots$ ja $\lim_k c_k = \infty$.

Tarkastellaan esimerkkinä integraalia $\int_a^\infty f(x)\,dx$. (Integraali $\int_{-\infty}^b f(x)\,dx$ palautuu tähän muuttujan vaihdolla $t = -x$.) Funktioille, jotka eivät vaihda merkkiään integroimisvälillä (lukuunottamatta mahdollisia erillisiä pisteitä, jotka eivät vaikuta integraalin arvoon), voi integraalien suppenemistarkasteluissa käyttää seuraavia vertailuperiaatteita, vrt. vastaavat periaatteet positiivitermisille sarjoille (Lause I.12.1). Yksinkertaisuuden vuoksi oletetetaan tässä funktiot määritellyksi ja epäyhtälöt voimassa oleviksi koko välillä $[a, \infty)$.

Lause VIII.7.4 (**Majorantti- ja minoranttiperiaatteet integraaleille**)
Jos f ja g ovat määriteltyjä välillä $[a, \infty)$, integroituvia välillä $[a, M]$ jokaisella $M > a$ ja $0 \leq f(x) \leq g(x)\ \forall x \in [a, \infty)$, niin pätee

$$\begin{aligned} \int_a^\infty g(x)\,dx \text{ suppenee} \quad &\Rightarrow \quad \int_a^\infty f(x)\,dx \text{ suppenee,} \\ \int_a^\infty f(x)\,dx \text{ hajaantuu} \quad &\Rightarrow \quad \int_a^\infty g(x)\,dx \text{ hajaantuu.} \end{aligned}$$

<u>Todistus</u> Väittämät ovat loogisesti ekvivalentteja, joten riittää todistaa ensimmäinen. Oletetaan siis, että $\int_a^\infty g(x)\,dx$ suppenee. Koska $f(x) \geq 0$ välillä $[a, \infty)$, niin funktio $F(x) = \int_a^x f(t)\,dt$ on ko. välillä kasvava:

$$F(x_2) - F(x_1) \;=\; \int_{x_1}^{x_2} f(t)\,dt \;\geq\; 0, \quad \text{kun } a \leq x_1 < x_2.$$

Samasta syystä $G(x) = \int_a^x g(t)\,dt$ on kasvava välillä $[a,\infty)$, ja koska $f(x) \le g(x),\ x \in [a,\infty)$ ja integraali $\int_a^\infty g(t)\,dt$ suppenee, niin voidaan päätellä:

$$F(x) \le G(x) \le \lim_{x\to\infty} G(x) = \int_a^\infty g(t)\,dt, \quad x \in [a,\infty).$$

Näin ollen $F(x)$ on välillä $[0,\infty)$ sekä kasvava että rajoitettu. Lauseen V.2.8 mukaan on tällöin olemassa raja-arvo $\lim_{x\to\infty} F(x) = \int_a^\infty f(x)\,dx$. $\square$

Esimerkki 5 Lauseen VIII.7.4 perusteella integraali $\int_1^\infty e^{-x}/x\,dx$ suppenee, sillä välillä $[1,\infty)$ on $e^{-x}/x \le e^{-x}$ ja integraali $\int_0^\infty e^{-x}\,dx$ suppenee:

$$\int_1^\infty e^{-x}\,dx = \Big|_1^\infty (-e^{-x}) = 1. \quad \square$$

Lauseen VIII.7.4 majorantti- ja minoranttiperiaatteet yleistyvät sopivin oletuksin koskemaan muitakin epäoleellisia integraaleja. Esimerkiksi olettaen että $0 \le f(x) \le g(x)\ \forall x \in (a,b)$, Määritelmän VIII.7.1 mukaisille integraaleille pätee

$$\int_a^b g(x)\,dx \text{ suppenee} \quad \Rightarrow \quad \int_a^b f(x)\,dx \text{ suppenee.}$$

Esimerkki 6 Integraalia $\int_0^1 e^{-x}/x\,dx$ voi tutkia muuttujan vaihdolla $t = 1/x$, jolloin Lause VIII.7.4 soveltuu. Suorempi päättely kuitenkin riittää: Integraali hajaantuu, koska $e^{-x}/x \ge e^{-1}x^{-1} > 0,\ x \in (0,1]$ ja $\int_0^1 x^{-1}\,dx$ hajaantuu. $\square$

Cauchyn kriteeri integraaleille

Seuraava yleinen suppenemiskriteeri toimii kaikille epäoleellisille integraaleille muotoa $\int_a^\infty f(x)\,dx$, vrt. vastaava Cauchyn kriteeri sarjoille (Lause I.12.4).

Lause VIII.7.5 (**Cauchyn kriteeri integraaleille**) Olkoon f on integroituva välillä $[a,M]$ jokaisella $M > a$. Tällöin integraali $\int_a^\infty f(x)\,dx$ suppenee täsmälleen kun jokaisella $\varepsilon > 0$ on olemassa $M > a$ siten, että pätee

$$\Big|\int_{M_1}^{M_2} f(x)\,dx\Big| < \varepsilon, \quad \text{kun } M_1, M_2 > M.$$

<u>Todistus</u> $\boxed{\Rightarrow}$ Jos integraali suppenee ja $M_1, M_2 > a$, niin

$$\begin{aligned}\int_{M_1}^{M_2} f(x)\,dx &= \int_a^{M_2} f(x)\,dx - \int_a^{M_1} f(x)\,dx \\ &\to \int_a^\infty f(x)\,dx - \int_a^\infty f(x)\,dx = 0, \quad \text{kun } M_1, M_2 \to \infty,\end{aligned}$$

joten Cauchyn kriteeri on täytetty.

$\boxed{\Leftarrow}$ Oletetaan, että Cauchyn kriteeri on täytetty ja merkitään

$$A_n = \int_a^n f(x)\,dx, \quad n \in \mathbb{N},\ n > a.$$

Oletuksen perusteella

$$A_n - A_m = \int_m^n f(x)\,dx \to 0, \quad \text{kun } n, m \to \infty,$$

joten $\{A_n\}$ on Cauchyn jono, ja siis (Lause I.10.5)

$$A_n \to A \in \mathbb{R}, \quad \text{kun } n \to \infty.$$

Käyttämällä tätä tulosta ja vetoamalla uudelleen oletukseen todetaan, että jos $\varepsilon > 0$, niin valitsemalla $M \in \mathbb{R}$ ja $n \in \mathbb{N}$ riittävän suuriksi on

$$\begin{aligned} \Big|\int_a^M f(x)\,dx - A\Big| &\le \Big|\int_a^M f(x)\,dx - A_n\Big| + |A_n - A| \\ &= \Big|\int_n^M f(x)\,dx\Big| + |A_n - A| < \varepsilon. \end{aligned}$$

Näin ollen f on integroituva välillä $[a, \infty)$:

$$\lim_{M\to\infty} \int_a^M f(x)\,dx = A. \quad \square$$

Esimerkki 7 Suppeneeko vai hajaantuuko integraali $\displaystyle\int_0^\infty \frac{\cos x}{\sqrt{x}}\,dx$?

<u>Ratkaisu</u> Jaetaan integraali kahteen osaan:

$$\int_0^\infty \frac{\cos x}{\sqrt{x}}\,dx = \int_0^{\pi/2} \frac{\cos x}{\sqrt{x}}\,dx + \int_{\pi/2}^\infty \frac{\cos x}{\sqrt{x}}\,dx.$$

Integroimalla osittain ensimmäisessä osassa todetaan

$$\int_0^{\pi/2} \frac{\cos x}{\sqrt{x}}\,dx = \Big|_0^{\pi/2} 2\sqrt{x}\cos x + \int_0^{\pi/2} 2\sqrt{x}\sin x\,dx = 2\int_0^{\pi/2} \sqrt{x}\sin x\,dx.$$

Funktio $f(x) = \sqrt{x}\sin x$ on jatkuva välillä $[0, \pi/2]$, joten tämä osa integraalista suppenee. Integraalin loppuosassa integroidaan osittain toisin päin:

$$\int_{\pi/2}^M \frac{1}{\sqrt{x}}\cos x\,dx = \Big|_{\pi/2}^M \frac{1}{\sqrt{x}}\sin x + \frac{1}{2}\int_{\pi/2}^M x^{-3/2}\sin x\,dx.$$

Tässä

$$\Big|_{\pi/2}^{M} \frac{1}{\sqrt{x}} \sin x \to -\sqrt{\frac{2}{\pi}}, \quad \text{kun } M \to \infty,$$

joten suppenemiskysymys siirtyy koskemaan integraalia

$$\int_{\pi/2}^{\infty} x^{-3/2} \sin x \, dx.$$

Kolmioepäyhtälön (Luku VIII.4, epäyhtälö (8)) avulla arvioiden todetaan, että jos $\pi/2 < M_1 < M_2$, niin

$$\begin{aligned} \left| \int_{M_1}^{M_2} x^{-3/2} \sin x \, dx \right| &\le \int_{M_1}^{M_2} x^{-3/2} |\sin x| \, dx \\ &\le \int_{M_1}^{M_2} x^{-3/2} \, dx \\ &= -2 \Big|_{M_1}^{M_2} x^{-1/2} \to 0, \quad \text{kun } M_1, M_2 \to \infty, \end{aligned}$$

joten Cauchyn kriteerin perusteella myös tämä osa integraalista suppenee. Siis vastaus on: Suppenee! $\square$

Integraalit ja sarjat

Edellä on jo nähty, että integraalien $\int_a^\infty f(x)\,dx$ ja sarjojen suppenemisteoriat muistuttavat toisiaan. Laskennallisemmaltakin kannalta integraalien ja sarjojen yhteys on merkittävä, sillä osoittautuu, että integraaalien avulla voidaan sekä yksinkertaistaa sarjojen suppenemistarkasteluja että tehostaa sarjojen summien numeerista laskemista.

Tarkastellaan esimerkkinä sarjaa $\sum_{k=1}^\infty f(k)$, missä oletetaan, että f on määritelty välillä $[1, \infty)$ ja lisäksi jollakin $n \in \mathbb{N}$ pätee:

(i) f on vähenevä välillä $[n, \infty)$ ja $\lim_{x\to\infty} f(x) = 0$.

Oletuksesta (i) seuraa, että $f(x) \ge 0$ välillä $[n, \infty)$ ja jokaisella $k \ge n$ on voimassa

$$f(k+1) \le f(x) \le f(k), \quad \text{kun } x \in [k, k+1].$$

Näin ollen

$$\int_k^{k+1} f(k+1)\,dx \;\le\; \int_k^{k+1} f(x)\,dx \;\le\; \int_k^{k+1} f(k)\,dx,$$

eli

$$f(k+1) \;\le\; \int_k^{k+1} f(x)\,dx \;\le\; f(k).$$

Summaamalla em. tulos yli indeksien $k = n \ldots N-1$ $(N > n)$ ja käyttämällä integraalin additiivisuutta saadaan

$$\sum_{k=n+1}^{N} f(k) \;\le\; \int_n^N f(x)\,dx \;\le\; \sum_{k=n}^{N-1} f(k) \;=\; \sum_{k=n+1}^{N-1} f(k) + f(n).$$

Koska tässä $\lim_n f(n) = 0$ oletuksen (i) mukaan, niin Cauchyn sarja- ja integraalikriteerien (Lauseet I.12.4 ja VIII.7.5) perusteella päätellään, että sarja $\sum_{k=1}^{\infty} f(k)$ suppenee täsmälleen kun integraali $\int_n^{\infty} f(x)\,dx$ suppenee. Lisäksi rajalla $N \to \infty$ saadaan arviot

$$\begin{aligned} &\sum_{k=n+1}^{\infty} f(k) \;\le\; \int_n^{\infty} f(x)\,dx \;\le\; \sum_{k=n+1}^{\infty} f(k) + f(n) \\ \Leftrightarrow\quad &\int_n^{\infty} f(x)\,dx - f(n) \;\le\; \sum_{k=n+1}^{\infty} f(k) \;\le\; \int_n^{\infty} f(x)\,dx. \end{aligned}$$

Tästä seuraa sarjan summan laskemisen kannalta mielenkiintoinen tulos:

$$\begin{aligned} \sum_{k=1}^{\infty} f(k) \;&=\; \sum_{k=1}^{n} f(k) + \int_n^{\infty} f(x)\,dx \;-\varepsilon_n \\ &=\; \sum_{k=1}^{n-1} f(k) + \int_n^{\infty} f(x)\,dx \;+\delta_n, \end{aligned}$$

missä $0 \le \varepsilon_n \le f(n)$ ja $0 \le \delta_n \le f(n)$ ($\delta_n = f(n) - \varepsilon_n$). Tämä tarkoittaa, että kun sarjan 'häntä' (indeksistä n tai $n+1$ alkaen) lasketaan integraalina, niin tällä tavoin saatavista approksimaatioista

$$\sum_{k=1}^{\infty} f(k) \approx \sum_{k=1}^{n-1} f(k) + \int_n^{\infty} f(x)\,dx \tag{a}$$

$$\approx \sum_{k=1}^{n} f(k) + \int_n^{\infty} f(x)\,dx \tag{b}$$

(a) antaa sarjan summalle alalikiarvon ja (b) ylälikiarvon, ja kummankin virhe on enintään $f(n)$ = sarjan n:s termi (= approksimaatioiden erotus). Tulos on siis pätevä, mikäli em. oletus (i) on voimassa.

Esimerkki 8 Millä α:n arvoilla ($\alpha \in \mathbb{R}$) sarja $\sum_{k=1}^{\infty} 1/k^{\alpha}$ suppenee?

Ratkaisu Sarja epäilemättä hajaantuu, jos $\alpha \le 0$. Jos $\alpha > 0$, niin oletus (i) on voimassa, kun $n = 1$. Integraali $\int_1^{\infty} x^{-\alpha}\,dx$ suppenee täsmälleen kun $\alpha > 1$, joten sarja suppenee täsmälleen samalla ehdolla. □

Esimerkin kysymys ratkaistiin jo aiemmin (Lause I.12.2) mutta vaivalloisemmin. Paitsi yksinkertaistaa suppenemistarkastelun, integraaliin vertaaminen helpottaa usein myös sarjan summan numeerista laskemista.

Esimerkki 8 (jatko) Esimerkin suppenevan (yliharmonisen) sarjan summalle pätee approksimaation (b) perusteella

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} \approx \sum_{k=1}^{n} \frac{1}{k^{\alpha}} + \frac{1}{\alpha - 1}\, n^{1-\alpha}, \quad \alpha > 1.$$

Tämä on ylälikiarvo, jota voi verrata pelkkään sarjan katkaisuun eli alalikiarvoon

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} \approx \sum_{k=1}^{n} \frac{1}{k^{\alpha}}\,.$$

Edellisen virhe on enintään $n^{-\alpha}$, jälkimmäisen $n^{1-\alpha}/(\alpha-1) + \mathcal{O}(n^{-\alpha})$. Esim. jos $\alpha = 5/4$, on tarkkuusero käytännössä dramaattinen, ks. Harj.teht. 7. □

Huomautettakoon, että esimerkissä vieläkin tehokkaampi algoritmivaihtoehto on approksimaatioiden (a) ja (b) keskiarvo eli

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} \approx \sum_{k=1}^{n} \frac{1}{k^{\alpha}} + \frac{1}{\alpha - 1}\, n^{1-\alpha} - \frac{1}{2}\, n^{-\alpha}\,.$$

Tämän virhe on suuruusluokkaa $\mathcal{O}(n^{-1-\alpha})$ (ks. Harj.teht. VIII.9:11).

HARJOITUSTEHTÄVIÄ

1. Laske tai osoita hajaantuvaksi kohdissa a)–o). Kohdissa p)–ö) määritä k:n arvot ($k \in \mathbb{R}$), joilla integraali suppenee.

a) $\displaystyle\int_0^6 \frac{2x}{x^2-4}\,dx$ b) $\displaystyle\int_0^6 \frac{2x}{|x^2-4|^{2/3}}\,dx$ c) $\displaystyle\int_0^1 \frac{x^5}{(1-x^2)^{3/2}}\,dx$

d) $\displaystyle\int_0^1 x^{-4/5}\ln x\,dx$ e) $\displaystyle\int_0^{\pi/2} \tan x\,dx$ f) $\displaystyle\int_0^{\pi} \frac{x}{\cos^2 x}\,dx$

g) $\displaystyle\int_0^2 \frac{x^2}{\sqrt{2x-x^2}}\,dx$ h) $\displaystyle\int_{-\infty}^{\infty} \frac{1}{x^2+7x+12}\,dx$ i) $\displaystyle\int_0^{\infty} \frac{1}{x^3+1}\,dx$

j) $\displaystyle\int_{-\infty}^{\infty} \frac{1}{x^2+7x+13}\,dx$ k) $\displaystyle\int_0^{\infty} \frac{1}{x^3-1}\,dx$ l) $\displaystyle\int_1^{\infty} \frac{1+\sqrt{x}}{x^2+x}\,dx$

m) $\displaystyle\int_1^{\infty} \frac{1+\sqrt{x}}{x\sqrt{x}+2x}\,dx$ n) $\displaystyle\int_0^{\infty} \frac{1}{\sqrt{e^x-1}}\,dx$ o) $\displaystyle\int_{-\infty}^{\infty} e^{-|x|}\cos x\,dx$

p) $\displaystyle\int_0^1 \frac{1+x+x^2}{x^k}\,dx$ q) $\displaystyle\int_1^{\infty} \frac{1+x+x^2}{x^k}\,dx$ r) $\displaystyle\int_0^{\infty} \frac{1}{kx^2+1}\,dx$

s) $\displaystyle\int_0^1 \frac{x^{k-1}+x^{-k}}{1+x}\,dx$ t) $\displaystyle\int_0^{\pi/2} \sin^k x\,dx$ u) $\displaystyle\int_0^{\infty} x^k e^{-x}\,dx$

v) $\displaystyle\int_0^{\infty} x^k \sin x\,dx$ x) $\displaystyle\int_0^{\infty} \frac{|k|^x}{\sqrt{e^x+1}}\,dx$ y) $\displaystyle\int_0^{\infty} x^k \ln x\,dx$

z) $\displaystyle\int_0^1 \frac{e^x}{(1-x^2)^k}\,dx$ å) $\displaystyle\int_0^{\infty} |x^2-1|^k e^{-x}\,dx$ ä) $\displaystyle\int_{-\infty}^{\infty} \frac{\ln|x|}{|x^3+1|^k}\,dx$

ö) $\displaystyle\int_{-\infty}^{\infty} \left[\sqrt[4]{|\sin x\cos x|}\;e^{|x|}\,(\ln|x|)^8\right]^k dx$

2. Johda palautuskaava seuraaville integraaleille, kun $n \in \mathbb{N} \cup \{0\}$.

a) $\displaystyle\int_0^{\infty} x^n e^{-x}\,dx$ b) $\displaystyle\int_0^{\infty} \frac{1}{(x^2+1)^{n+1}}\,dx$ c) $\displaystyle\int_0^1 (\ln x)^n\,dx$

3. Todista suppeneminen sijoituksella $t = 1/x$:

a) $\displaystyle\int_0^1 e^{-1/x}\,dx$ b) $\displaystyle\int_0^1 \cos\frac{1}{x}\,dx$ c) $\displaystyle\int_0^{\infty} \frac{1}{x}\sin\frac{1}{x}\,dx$

4. Sievennä ja piirrä kuvaaja:

$$f(x) = \int_0^{\pi} \frac{x\sin t}{\sqrt{1+2x\cos t+x^2}}\,dx.$$

5. *Gammafunktio* (Γ-funktio) määritellään

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt.$$

Näytä tosiksi seuraavat väittämät ja hahmottele näiden perusteella Γ-funktion kulku. (Kohdassa e) on $\Gamma(1/2) = \sqrt{\pi} \approx 1.77$, ks. Propositio XIII.4.1.)

a) Γ on määritelty välillä $(0, \infty)$.
b) $\Gamma(x+1) = x\Gamma(x)\ \forall x > 0$.
c) $\Gamma(n) = (n-1)!\ \forall n \in \mathbb{N}$.
d) $\lim_{x\to 0^+} \Gamma(x) = \lim_{x\to\infty} \Gamma(x) = \infty$.
e) $\Gamma(1/2) = 2\,\Gamma(3/2) = \int_{-\infty}^{\infty} e^{-x^2}\, dx$.

6. a) Halutaan selvittää, millä α:n arvoilla ($\alpha \in \mathbb{R}$) sarja

$$\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^\alpha}$$

suppenee. Ratkaise kysymys integraaliin vertaamalla.

b) Olkoon $n \in \mathbb{N}$ suuri luku. Perustele integraalin avulla arviot

$$\sum_{k=1}^{n} k^\alpha = \begin{cases} \frac{1}{\alpha+1}\, n^{\alpha+1} + \mathcal{O}(n^\alpha), & \text{kun } \alpha > -1, \\ \ln n + \mathcal{O}(n^{-1}), & \text{kun } \alpha = -1. \end{cases}$$

Mikä on jäännöstermi $\mathcal{O}(n^\alpha)$ tarkasti tapauksissa $\alpha = 0$ ja $\alpha = 1$?

7. Käytettävissä on tietokone, joka laskee summan $s_n = \sum_{k=1}^{n} k^{-5/4}$ ajassa $10^{-10} n$ sekuntia. Summien s_n avulla halutaan laskea raja-arvo $s = \lim_n s_n$ kuuden desimaalin tarkkuudella (virhe $< 5 \cdot 10^{-7}$). Arvioi laskenta-aika
a) vuosimiljardeina approksimaatioilla $s \approx s_n$,
b) mikrosekunteina approksimaatioilla $s \approx s_n + 4/\sqrt[4]{n}$.

8. (*) *Eulerin vakio* $\gamma = 0.5772156649..$ määritellään

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln n \right) = \sum_{k=1}^{\infty} a_k, \quad a_k = \frac{1}{k} - \ln\left(1 + \frac{1}{k}\right).$$

a) Perustele jälkimmäinen laskukaava sekä ko. sarjan suppeneminen.
b) Näytä, että suurilla $n \in \mathbb{N}$ pätee $\sum_{k=1}^{n} a_k + \frac{1}{2n} = \gamma + \mathcal{O}(n^{-2})$.

9. (*) Arvioi luku $n!$ sekä ylhäältä että alhaalta vertaamalla lukua $\ln(n!) = \ln 1 + \ln 2 + \cdots + \ln n$ integraaliin. Kuinka moninumeroinen luku on 1000! (kymmenjärjestelmässä) ?

VIII.8 Pinta-ala ja kaarenpituus

Tässä luvussa tarkastellaan määrätyn integraalin käyttöä kahdessa tasogeometrisessa tehtävässä: käyrän viivan rajoittaman *tasoalueen* ($\mathbb{R}^2$:n tai E^2:n osajoukon) pinta-alan määrittämisessä ja yksinkertaisen käyränkaaren kaarenpituuden laskemisessa.

Olkoon $f(x) \geq 0$ välillä $[a, b]$. Tarkastellaan käyrän $y = f(x)$ ja suorien $y = 0$, $x = a$ ja $x = b$ rajaamaa $\mathbb{R}^2$:n osajoukkoa

$$A = \{\, (x, y) \in \mathbb{R}^2 \mid x \in [a, b] \ \ \& \ \ 0 \leq y \leq f(x) \,\}.$$

Halutaan määrittää A:n pinta-ala. Tämä on ei-negatiivinen reaaliluku, jota merkitään jatkossa symbolilla $\mu(A)$. — Kyseessä on itse asiassa funktio tyyppiä $\mu : \ \mathcal{M} \to [0, \infty)$, missä $\mathcal{M}$ koostuu $\mathbb{R}^2$:n osajoukoista, tarkemmin sanoen sellaisista nk. *mitallisista* joukoista, joiden pinta-ala on määriteltävissä. Lukua $\mu(A)$ sanotaan tämän mukaisesti myös A:n (pinta-ala)*mitaksi* (engl. measure)[†]. Ajatellen toistaiseksi vain oletettua muotoa olevia tasoalueita asetetaan pinta-alamitalle μ seuraavat kaksi aksioomaa:

A1. *Vertailuperiaate*: Jos $0 \leq m \leq f(x) \leq M \ \forall x \in [a, b]$, niin

$$m(b - a) \leq \mu(A) \leq M(b - a).$$

[†]Käsitteet 'mitta' ja 'mitallisuus' viittaavat matematiikan lajiin nimeltä *mittateoria*. Yleisemmän mittateorian perusideoita ei tässä yhteydessä tarkastella, vaan asiaan palataan myöhemmin usean muuttujan analyysin yhteydessä.

A2. *Additiivisuus*: Jos $A_1 = \{\, (x,y) \in A \mid x \in [a,c] \,\}$ ja $A_2 = \{\, (x,y) \in A \mid x \in [b,c] \,\}$, missä $a < c < b$, niin A on mitallinen täsmälleen kun A_1 ja A_2 ovat molemmat mitallisia, ja pätee

$$\mu(A) = \mu(A_1) + \mu(A_2).$$

Oletetaan nyt A sellaiseksi, että aksioomien A1–A2 mukainen pinta-alamitta $\mu(A)$ on määriteltävissä, ja tutkitaan, mitä tästä seuraa. Ensinnäkin todetaan additiivisuusaksioomaa A2 toistuvasti soveltamalla, että jos $X = \{x_k,\ k = 0 \ldots n\}$ on välin $[a,b]$ jako (ts. $a = x_0 < x_1 < \ldots < x_n = b$), ja merkitään

$$A_k = \{\, (x,y) \in A \mid x_{k-1} \le x \le x_k \,\}, \quad k = 1 \ldots n,$$

niin jokainen A_k on mitallinen ja

$$\mu(A) = \sum_{k=1}^{n} \mu(A_k).$$

Vertailuperiaatteen A1 mukaan pätee tässä

$$\begin{aligned} m_k \le f(x) \le M_k \ \forall x \in [x_{k-1}, x_k] \\ \Rightarrow \quad & m_k(x_k - x_{k-1}) \le \mu(A_k) \le M_k(x_k - x_{k-1}) \\ \Rightarrow \quad & \mu(A_k) = \eta_k(x_k - x_{k-1}), \quad \eta_k \in [m_k, M_k]. \end{aligned}$$

Jos f on jatkuva välillä$[a,b]$, niin Weierstrassin lauseen (Lause V.1.12) mukaan voidaan tässä valita

$$m_k = \min_{x \in [x_{k-1}, x_k]} \{f(x)\}, \quad M_k = \max_{x \in [x_{k-1}, x_k]} \{f(x)\},$$

jolloin Ensimmäisen väliarvolauseen (Lause V.1.10) mukaan on olemassa $\xi_k \in [x_{k-1}, x_k]$ siten, että $f(\xi_k) = \eta_k$.

Yhdistämällä em. päätelmät todetaan, että jos f on jatkuva välillä $[a, b]$, niin jokaiseen välin $[a, b]$ jakoon X liittyen $\mu(A)$ on ilmaistavissa eräänä ko. jakoon liittyvänä Riemannin summana:

$$\mu(A) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}), \quad \xi_k \in [x_{k-1}, x_k], \quad k = 1 \ldots n.$$

Koska Analyysin peruslauseen ja Riemannin integraalin määritelmän mukaan tämän kaavan oikealla puolella olevat summat lähestyvät määrättyä integraalia $\int_a^b f(x)\,dx$ jaon tihetessä, riippumatta pisteiden ξ_k valinnasta, niin seuraa, että $\mu(A)$ on määriteltävä kaavalla

$$\boxed{\mu(A) = \int_a^b f(x)\,dx.} \tag{1}$$

Näin määritellylle mitalle aksioomat A1–A2 myös toteutuvat integraalin vastaavien ominaisuuksien (Lauseet VIII.5.6 ja VIII.5.4) perusteella. On siis päätelty, että jos f on jatkuva välillä $[a, b]$, niin A on aksioomien A1–A2 mukaisesti mitallinen jos (ja vain jos!) mitta määritellään kaavalla (1).

Esimerkki 1 *Paraabelin segmentin*

$$A = \{\, (x, y) \in \mathbb{R}^2 \mid x \in [0, 2] \ \ \& \ \ 0 \le y \le 2x - x^2 \,\}$$

pinta-ala on kaavan (1) mukaisesti†

$$\mu(A) = \int_0^2 (2x - x^2)\,dx = \Big|_0^2 \left(x^2 - \frac{1}{3}x^3\right) = \underline{\underline{\frac{4}{3}}}. \quad \square$$

Laskukaava (1) seuraa myös pelkästään olettamalla, että f on välillä $[a, b]$ rajoitettu ja Riemann-integroituva. Nimittäin jos X on välin $[a, b]$ jako, niin em. päätelmien mukaisesti voidaan $\mu(A)$ (sikäli kuin olemassa) arvioida Riemannin ylä- ja alasummilla (ks. Luku VIII.5):

$$\underline{\sigma}(f, X) \le \mu(A) \le \overline{\sigma}(f, X).$$

Tällöin seuraa Lauseesta VIII.5.8, että mitta on määriteltävä kaavalla (1), jolloin aksioomat A1–A2 myös toteutuvat. On päädytty seuraavaan tulokseen.

†Paraabelin segmentin pinta-alan laski ensimmäisenä antiikin huomattavin matemaatikko **Arkhimedes** (287-212 eKr). Arkhimedeen menetelmä perustui segmenttiä approksimoiviin monikulmioihin, ks. Harj.teht. VIII.9:7.

LAUSE VIII.8.1 Jos f on määritelty, rajoitettu ja Riemann-integroituva välillä $[a,b]$ ja $f(x) \geq 0,\ x \in [a,b]$, niin tasoalueen

$$A = \{\, (x,y) \in \mathbb{R}^2 \mid x \in [a,b] \ \ \&\ \ 0 \leq y \leq f(x) \,\}$$

pinta-ala $\mu(A)$ määräytyy aksioomista A1–A2 yksikäsitteisesti kaavalla (1).

ESIMERKKI 2 Laske ympyräneljänneksen

$$A = \{\, (x,y) \in \mathbb{R}^2 \mid x,y \geq 0 \ \ \&\ \ x^2 + y^2 \leq R^2 \,\}$$

pinta-ala.

<u>Ratkaisu</u> Kaavan (1) mukaan laskien saadaan

$$\begin{aligned}
\mu(A) &= \int_0^R \sqrt{R^2 - x^2}\, dx \\
&\qquad [\ \text{sijoitus}\ \ x = R\sin t,\ \ dx = R\cos t\, dt\ ] \\
&= \int_0^{\pi/2} R^2 \cos^2 t\, dt \\
&= \Big|_0^{\pi/2} \frac{1}{2} R^2 (t + \sin t \cos t) = \underline{\underline{\frac{1}{4}\pi R^2}}. \quad \square
\end{aligned}$$

ESIMERKKI 3 Määritä $\mu(A)$, kun

$$A = \{\, (x,y) \in \mathbb{R}^2 \mid x \in [0,2] \ \ \&\ \ 0 \leq y \leq \max\{x, 2e^{-x}\} \,\}.$$

<u>Ratkaisu</u> Tässä on ensin ratkaistava

$$x = 2e^{-x} \ \Rightarrow\ x = c \approx 0.852606,$$

jolloin on (vrt. kuva)

$$f(x) = \begin{cases} 2e^{-x}, & \text{kun } x \in [0,c], \\ x, & \text{kun } x \in [c,2]. \end{cases}$$

Kaavan (1) ja integraalin additiivisuuden perusteella saadaan

$$\begin{aligned}\mu(A) = \int_0^2 f(x)\,dx &= \int_0^c 2e^{-x}\,dx + \int_c^2 x\,dx \\ &= \Big|_0^c (-2e^{-x}) + \Big|_c^2 \tfrac{1}{2}x^2 \\ &= 4 - 2e^{-c} - \tfrac{1}{2}c^2 \approx \underline{\underline{2.78393}}. \quad \square\end{aligned}$$

Sopimalla pinta-alamitalle lisää ominaisuuksia voidaan integraalien avulla laskea myös yleisempien tasoalueiden pinta-aloja. Tarkasteltakoon tässä ainoastaan esimerkkinä kahden käyrän väliin jäävän alueen pinta-alaa. Olkoon $0 \le g(x) \le f(x)\ \forall x \in [a,b]$. Halutaan märätä $\mu(A)$, kun

$$A = \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \ \ \& \ \ g(x) \le y \le f(x)\}.$$

Kun merkitään

$$\begin{aligned}A_1 &= \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \ \ \& \ \ 0 \le y \le g(x)\}, \\ A_2 &= \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \ \ \& \ \ 0 \le y \le f(x)\},\end{aligned}$$

niin olettamalla additiivisuuslaski $\mu(A_2) = \mu(A_1) + \mu(A)$ (pinta-alamitan uusi aksiooma!) seuraa laskukaava

$$\mu(A) = \mu(A_2) - \mu(A_1) = \int_a^b [f(x) - g(x)]\,dx.$$

Tätä voidaan pitää pätevänä aina kun f ja g ovat integroituvia välillä $[a,b]$ ja $f(x) \geq g(x)$ $\forall x \in [a,b]$. Ellei jälkimmäinen oletus ole voimassa, sovelletaan kaavaa asettamalla f:n tilalle $\max\{f,g\}$ ja g:n tilalle $\min\{f,g\}$, jolloin yleiseksi laskukaavaksi kahden käyrän väliin jäävän alueen pinta-alalle välillä $[a,b]$ tulee

$$\boxed{\mu(A) = \int_a^b |f(x) - g(x)|\, dx.} \tag{2}$$

Esimerkki 4 Laske käyrien $y = x$ ja $y = 2e^{-x}$ väliin jäävän alueen pinta-ala välillä $[0,2]$.

<u>Ratkaisu</u> Kaavan (2) ja edellisen esimerkin perusteella

$$\begin{aligned}\mu(A) = \int_0^2 |2e^{-x} - x|\, dx &= \int_0^c (2e^{-x} - x)\, dx + \int_c^2 (x - 2e^{-x})\, dx \\ &= \Big|_0^c \left(-2e^{-x} - \frac{1}{2}x^2\right) + \Big|_c^2 \left(\frac{1}{2}x^2 + 2e^{-x}\right) \\ &= 4 + 2e^{-2} - 4e^{-c} - c^2 \approx \underline{\underline{1.83852}}. \quad \square\end{aligned}$$

Kaarenpituus

Luvussa V.4 esitetyn määritelmän mukaisesti voidaan kaaren

$$S = \{\, (x,y) \in \mathbb{R}^2 \mid x \in [a,b] \ \ \& \ \ y = f(x) \,\}$$

kaarenpituusmitta $\mu(S)$ määritellä välin $[a,b]$ jakoihin X liittyvien summien

$$\sigma(f,X) = \sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + [f(x_k) - f(x_{k-1})]^2}$$

pienimpänä ylärajana

$$\mu(S) = \sup_X \sigma(f,X),$$

sikäli kuin summien joukko on rajoitettu. Jos oletetaan, että f on jatkuva välillä $[a,b]$ ja derivoituva välillä (a,b), niin Differentiaalilaskun väliarvolauseen mukaan on olemassa $\xi_k \in (x_{k-1}, x_k)$, $k = 1 \ldots n$ siten, että

$$f(x_k) - f(x_{k-1}) = f'(\xi_k)(x_k - x_{k-1}), \quad k = 1 \ldots n,$$

jolloin

$$\sigma(f,X) = \sum_{k=1}^{n} \sqrt{1 + [f'(\xi_k)]^2}\, (x_k - x_{k-1}).$$

Mainituin oletuksin summa $\sigma(f, X)$ on siis eräs välin $[a, b]$ jakoon X ja funktioon $g(x) = \sqrt{1 + [f'(x)]^2}$ liittyvä Riemannin summa. Päätellään, että sikäli kuin g on Riemann-integroituva välillä $[a, b]$ (Analyysin peruslauseen mukaan riittää, että f on jatkuvasti derivoituva välillä $[a, b]$), on kaarenpituusmitta määriteltävissä ja pätee integraalikaava

$$\boxed{\mu(S) = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.} \tag{3}$$

Esimerkki 5 Laske kaarenpituus puoliympyrän kaarelle

$$S = \{\, (x, y) \in \mathbb{R}^2 \mid -R \le x \le R \ \ \&\ \ y = \sqrt{R^2 - x^2}\,\}.$$

<u>Ratkaisu</u> Tässä on $f(x) = \sqrt{R^2 - x^2}$, joten kaavan (3) mukaan on

$$\mu(S) \ = \int_{-R}^{R} \sqrt{1 + \frac{x^2}{R^2 - x^2}}\, dx \ = \int_{-R}^{R} \frac{R}{\sqrt{R^2 - x^2}}\, dx.$$

Tämä on suppeneva (epäoleellinen) integraali. Sijoituksella $x = R\cos t,\ t \in [0, \pi]$ saadaan

$$\mu(S) = \int_{\pi}^{0} (-R)\, dt = \int_0^{\pi} R\, dt = \underline{\underline{\pi R}}. \quad \square$$

Esimerkki 6 Laske käyrän $y = x^2$ kaarenpituus välillä $[0, 1]$.

<u>Ratkaisu</u> Kaavan (3) ja Luvun VIII.2 Esimerkin 2 perusteella (sijoitus $2x = t$)

$$\begin{aligned}\mu(S) = \int_0^1 \sqrt{1 + 4x^2}\, dx &= \Big|_{t=0}^{t=2} \frac{1}{4}\left[t\sqrt{1 + t^2} + \ln\left(t + \sqrt{1 + t^2}\right)\right] \\ &= \frac{1}{2}\sqrt{5} + \frac{1}{4}\ln(2 + \sqrt{5}) \approx \underline{\underline{1.47894}}. \quad \square\end{aligned}$$

Esimerkki 7 Käyrän $y = \sin x$ kaarenpituus välillä $[0, \pi]$ on

$$\mu(S) = \int_0^{\pi} \sqrt{1 + \cos^2 x}\, dx.$$

Tämä on nk. *elliptinen integraali*, jota ei voi laskea suljetussa muodossa. $\square$

Kuten esimerkeistä nähdään, kaarenpituusintegraali voi olla melko yksinkertaisissakin tapauksissa hankala tai peräti mahdoton suljetussa muodossa integroimisen kannalta. Vaihtoehtona on tällöin integraalin laskeminen suoraan numeerisin menetelmin (ks. seuraava luku).

HARJOITUSTEHTÄVIÄ

1. Laske seuraavien tasoalueiden pinta-alat (tarkasti tai likiarvona).

   a) $0 \le y \le |\cos 3x|, \ \ x \in [0, \pi]$ b) $0 \le \min\{x, \cos x\}, \ \ x \in [0, \pi/2]$

   c) $\min\{\cos x, \sin 2x\} \le y \le \max\{\cos x, \sin 2x\}, \ \ x \in [0, \pi]$

   d) $|x| \le y \le \sqrt{x^2+1}\,, \ \ |x| \le \sqrt{a^2-1}\,, \ \ a > 1$

2. Laske seuraavien käyrien väliin jäävän alueen pinta-ala annetulla välillä (tarkasti tai likiarvona).

   a) $y = \sinh x$, $y = \cosh x$, $[-\ln 2, \ln 2]$ b) $y = \ln x$, $y = e^x$, $[1/2, 2]$

   c) $y = \cos x$, $y = x$, $[0, \pi]$ d) $y = \sin x$, $y = x^2$, $[0, 1]$

3. Laske likiarvo käyrien $y = e^x$ ja $y = 3 - x^2$ väliin jäävän rajoitetun tasoalueen pinta-alalle.

4. Näytä, että ellipsin

$$S: \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a, b > 0)$$

   sisään jäävän alueen pinta-ala on $\mu(A) = \pi ab$.

5. Laske seuraavien käyrien kaarenpituudet annetulla välillä.

   a) $y = 2\sqrt{x}$, $[0, 4]$ b) $y = \ln x$, $[1, e]$

   c) $y = \ln \cos x$, $[0, \pi/4]$ d) $y = (1 - x/3)\sqrt{x}\,,$ $[0, 3]$

6. Näytä, että ellipsin puolikaaren $S: \ (x/a)^2 + (y/b)^2 = 1 \ \ \& \ \ y \ge 0$ pituus saadaan (elliptisenä) integraalina

$$\mu(S) = \int_{-\pi/2}^{\pi/2} \sqrt{a^2 \cos^2 t + b^2 \sin^2 t}\, dt.$$

7. (*) Laske käyrien $y = e^{-|x|} \cos x$ ja $y = e^{-|x|} \sin x$ väliin jäävän tasoalueen pinta-ala.

8. (*) Laske seuraavien alueiden pinta-alat ($a > 0$):

   a) Sykloidin $x = a(t - \sin t)$, $y = a(1 - \cos t)$ ja x-akselin väliin jäävä alue välillä $x \in [0, 2\pi a]$.

   b) Asteroidin $|x|^{2/3} + |y|^{2/3} = a^{2/3}$ sisään jäävä alue.

9. (*) Laske seuraavien kaarien kaarenpituudet ($a > 0$):

   a) Sykloidin kaari $S:\ x = a(t - \sin t),\ y = a(1 - \cos t),\ t \in [0, 2\pi]$.

   b) Asteroidin kaari $S:\ |x|^{2/3} + |y|^{2/3} = a^{2/3},\ x, y \geq 0$.

10. (*) Painovoiman vaikuttaessa suunnassa $-\vec{j}$ noudattaa xy-tasolla vapaasti riippuva ketju *ketjuviivaa*

$$y = a \cosh\left(\frac{x-b}{a}\right) + c,$$

missä $a > 0$ ja $b, c \in \mathbb{R}$.

a) Laske vakiot a, b, c ja hahmottele kejuviivan kulku, kun ketjun päät ovat pisteissä $(-2, 0)$ ja $(2, 0)$ ja ketjun pituus $= 8$.

b) Jos ketjun päät ovat pisteissä $(-2, 4)$ ja $(4, 4)$ ja ketju kulkee origon kautta, niin mikä on ketjun pituus?

VIII.9 Numeerinen integrointi

Numeerisella integroinnilla tarkoitetaan määrätyn integraalin, eli reaaliluvun

$$I(f,a,b) = \int_a^b f(x)\,dx$$

laskemista numeerisin keinoin (likimäärin). Jatkossa integroitavalle funktiolle f asetetaan melko voimakkaita säännöllisyysvaatimuksia. Vähimmäisvaatimus on, että f jatkuva välillä $[a,b]$, jolloin f on ko. välillä myös Riemann-integroituva (Lause VIII.6.1).

Numeerisen integroinnin menetelmät ovat yleensä nk. *numeerisia kvadratuureja* (tai kvadratuurisääntöjä, engl. quadrature rule)[†], joissa integraalia approksimoidaan äärellisenä summana muotoa

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{N} w_i f(x_i). \tag{1}$$

Tässä pisteitä $x_i,\ i = 1\ldots N$ (yleensä $x_i \in [a,b]$) sanotaan *kvadratuuripisteiksi* ja lukuja w_i *kvadratuuripainoiksi*. Kvadratuuripisteitä ja -painoja valittaessa integroimisväli jaetaan usein ensin osaväleihin $[x_{k-1}, x_k],\ k = 1\ldots n$ (kuten määrätyn integraalin määrittelyssä alunperin). Tällöin integraalin additiivisuuden nojalla on

$$I(f,a,b) = \int_a^b f(x)\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)\,dx = \sum_{k=1}^{n} I_k\,. \tag{2}$$

Jos integroitava funktio on säännöllinen ja osavälit $[x_{k-1}, x_k]$ riittävän lyhyitä, voidaan osaintegraalit

$$I_k = \int_{x_{k-1}}^{x_k} f(x)\,dx$$

laskea likimäärin käyttämällä suhteellisen yksinkertaista kvadratuuria muotoa (1), missä $a = x_{k-1}$ ja $b = x_k$. Perustana kvadratuurin yksinkertaistumiselle lyhyellä välillä on Taylorin lause, jonka mukaan säännöllinen funktio on likimain (matala-asteinen) polynomi lyhyellä välillä (ks. Luku VII.4). Numeerisen integroinnin virheanalyysin lähtöajatuksena onkin juuri vertailu polynomeihin.

Jatkossa ajatellaan väli $[a,b]$ alkuperäisen integroimisvälin lyhyeksi osaväliksi ja f ko. välillä (riittävän) säännölliseksi. Tällaisten lähtöoletusten vallitessa yksinkertaisimmat numeeriset kvadratuurit ovat:

[†]Kvadratuuri tarkoittaa kirjaimellisesti 'neliöimistä'. Termi viittaa integraalin ja pinta-alan väliseen yhteyteen.

Keskipistesääntö: $$\int_a^b f(x)\,dx \approx (b-a)f(\frac{a+b}{2}).$$

Puolisuunnikassääntö: $$\int_a^b f(x)\,dx \approx \frac{1}{2}(b-a)[f(a)+f(b)].$$

Simpsonin sääntö: $$\int_a^b f(x)\,dx \approx \frac{1}{6}(b-a)\left[f(a)+4f(\frac{a+b}{2})+f(b)\right].$$

Näistä etenkin Simpsonin sääntö (osaväleillä käytettynä) on laskinten ja tietokoneiden yleisesti käyttämä numeerisen integroinnin menetelmä.† Puolisuunnikassääntöön viitataan usein myös nimellä *Trapetsi*. Nimensä mukaisesti sääntö antaa integraalin arvoksi *puolisuunnikkaan* (trapetsin) pinta-alan:

Kun ym. sääntöjä käytetään osaväleillä hajotelmassa (2), niin tuloksena on alkuperäisen integraalin approksimaatio muotoa (1).

Sanotaan tällöin, että kyseessä on *yhdistetty* (engl. composite) kvadratuuri. Esimerkiksi *yhdistetty keskipistesääntö* on kvadratuuri muotoa (1), missä valitaan

$$x_i = \frac{1}{2}(x_{i-1}+x_i), \quad w_i = x_i - x_{i-1}\,, \quad i = 1\ldots n = N.$$

Yhdistetty keskipistesääntö on siis eräs integraalia approksimoivista Riemannin summista. Suoraan Riemannin summiin perustuvista approksimaatioista tämä on yleensä tarkin (ks. virhearviot jäljempänä).

†Simpsonin sääntö on numeerisen integroinnin klassikko, jota on aikanaan käytetty paljon käsinlaskussakin. Säännön keksi englantilainen matemaatikko **Thomas Simpson** (1710-1761).

Jos hajotelmassa (2) jako on tasavälinen ja vastaavassa yhdistetyssä kvadratuurissa (1) merkitään $h = x_i - x_{i-1}$ (= peräkkäisten kvadratuuripisteiden väli) ja $f_i = f(x_i)$, niin *yhdistetty trapetsi* saa muodon

$$\int_a^b f(x)\,dx \approx h\left[\frac{1}{2}f_0 + f_1 + \cdots + f_{n-1} + \frac{1}{2}f_n\right]$$

ja *yhdistetty Simpsonin sääntö* muodon

$$\int_a^b f(x)\,dx \approx \frac{h}{3}\left[\,f_0 + 4f_1 + 2f_2 + \cdots + 2f_{2n-2} + 4f_{2n-1} + f_{2n}\,\right].$$

Yhdistetyssä Simpsonin säännössä joka toinen kvadratuuripiste on osavälijaon jakopiste ja joka toinen (suurimmalla painokertoimella varustettu) on osavälin keskipiste.

Esimerkki 1 Kun integraalia

$$\int_0^1 xe^{-x^2}/(x+1)\,dx$$

approksimoidaan käyttäen tasavälistä yhdistettyä trapetsia ja Simpsonia, saadaan seuraavat tulokset:

pisteitä	Trapetsi	Simpson
3	0.17578506065833	0.20372346078015
21	0.20230210857337	0.20256800959472
201	0.20256528212086	0.20256794027440
2001	0.20256791368608	0.20256794026753

Tässä Simpson on Trapetsia selvästi nopeampi. Tämä on odotettavissa, koska integroitava funktio on sileä, ks. virhearviot jäljempänä. □

Esimerkki 2 Sama vertailuasetelma kuin edellisessä esimerkissä. Lasketaan

$$\int_0^1 e^{-x}/(\sqrt{x}+1)\,dx.$$

Tulokset ovat tässä tapauksessa:

pisteitä	Trapetsi	Simpson
3	0.47363365	0.43418824
21	0.40741212	0.40604377
201	0.40520890	0.40516458
2001	0.40513820	0.40513679

Konvergenssi on nyt selvästi hitaampaa kuin edellisessä esimerkissä, eikä eri menetelmillä ole tässä merkittävää eroa. Syynä on ilmeisesti se, että integroitava funktio on vähemmän säännöllinen kuin edellisessä esimerkissä (jatkuva, mutta ei jatkuvasti derivoituva integroimisvälillä). □

Em. esimerkeissä on merkille pantavaa, että kun Esimerkin 2 integraalissa tehdään sijoitus $\sqrt{x} = t$, on tuloksena Esimerkin 1 integraali (kertoimella 2). Sijoitus siis kannattaa tässä tehdä, vaikka se ei muuten (suljetussa muodossa integroimisen kannalta) tee tehtävää helpommaksi. Myös osittaisintegroinnilla, sarjakehitelmillä ym. voi usein manipuloida tehtävää niin, että integroitavasta funktiosta tulee säännöllisempi, jolloin numeeriset kvadratuurit toimivat paremmin.

Virhearvioista

Osavälijakoon perustuvan yhdistetyn kvadratuurin virhe arvioidaan ensin erikseen kullakin osavälillä. Tarkastellaan jatkossa edellä esitettyä kolmea esimerkkiä: Keskipistesääntö, Trapetsi ja Simpson.

Arvioitaessa virhettä yksittäisellä osavälillä on ensimmäisenä tehtävänä tutkia, kuinka korkea-asteisille polynomeille kvadratuuri on tarkka, ts. on määrättävä indeksi m siten, että kvadratuuri on tarkka polynomille astetta m mutta ei astetta $m + 1$. Mainituissa esimerkkitapauksissa on tulos seuraava:

Sääntö	Tarkka polynomeille astetta
Keskipistesääntö	$m = 1$
Trapetsi	$m = 1$
Simpson	$m = 3$

Aloitetaan keskipistesäännöstä. Oletetetaan, että funktio f on tarkasteltavalla (osa)välillä $[a, b]$ kahdesti jatkuvasti derivoituva. Merkitään $c = (a + b)/2$ ja käytetään integroimisvirheelle symbolia $E(f)$:

$$E(f) = \int_a^b f(x)\,dx - (b - a)f(c).$$

Integroimisvirheen arvioiminen perustuu kahteen perushavaintoon: Ensinnäkin nähdään, että pätee

$$E(f + g) = E(f) + E(g). \tag{a}$$

(Yleisemmin $E(f)$ on lineaarinen f:n suhteen, eli $E(\alpha f + \beta g) = \alpha E(f) + \beta E(g)$, $\alpha, \beta \in \mathbb{R}$.) Toinen perushavainto on em. taulukkoon perustuva:

$$f(x) = A + Bx \quad (A, B \in \mathbb{R}) \quad \Rightarrow \quad E(f) = 0. \tag{b}$$

Olkoon nyt $p(x) = f$:n Taylorin polynomi astetta $n = 1$ pisteessä c, eli

$$p(x) = f(c) + f'(c)(x - c).$$

Tällöin on havaintojen (a)–(b) perusteella, ja koska $(f - p)(c) = 0$,

$$E(f) = E(f - p + p) = E(f - p) + E(p) = E(f - p) = \int_a^b (f - p)(x)\,dx.$$

Taylorin lauseen (Lause VII.4.3) perusteella on

$$f(x) - p(x) = \frac{1}{2} f''(\xi)(x - c)^2 = f''(\xi)\,\omega(x), \quad x \in [a, b],$$

missä $\xi = \xi(x) \in (a, b)$ ja $\omega(x) = \frac{1}{2}(x - c)^2 \geq 0$. Näin ollen jos

$$M_1 = \min_{x \in [a,b]} f''(x), \quad M_2 = \max_{x \in [a,b]} f''(x),$$

niin voidaan päätellä

$$\begin{aligned} & M_1\,\omega(x) \;\leq\; (f - p)(x) \;\leq\; M_2\,\omega(x), \quad x \in [a, b] \\ \Rightarrow \quad & M_1 \int_a^b \omega(x)\,dx \;\leq\; \int_a^b (f - p)(x)\,dx \;\leq\; M_2 \int_a^b \omega(x)\,dx \\ \Leftrightarrow \quad & \frac{1}{24} M_1\,(b - a)^3 \;\leq\; \int_a^b (f - p)(x)\,dx \;\leq\; \frac{1}{24} M_2\,(b - a)^3. \end{aligned}$$

Siis jollakin $\eta \in [M_1, M_2]$ pätee

$$E(f) = \frac{\eta}{24}\,(b - a)^3.$$

Koska f'' on jatkuva välillä $[a, b]$, niin $M_1 = f''(x_1)$ ja $M_2 = f''(x_2)$ joillakin $x_1, x_2 \in [a, b]$ (Lause V.1.12), jolloin $\eta = f''(\xi)$ jollakin $\xi \in [a, b]$ (Lause V.1.10). Keskipistesäännölle on näin johdettu virhekaava

$$\boxed{E(f) = \frac{1}{24}\,(b - a)^3 f''(\xi), \quad \xi \in [a, b] \quad \text{(Keskipistesääntö).}}$$

Trapetsille voidaan johtaa virhekaava vastaavaan tapaan. Tässä tapauksessa on vertailupolynomiksi p kuitenkin syytä valita Taylorin polynomin sijasta ensimmäisen asteen interpolaatiopolypainnomi, jolle pätee $p(a) = f(a)$ ja $p(b) = f(b)$ (ks. Luku VII.6). Virhekaavaksi saadaan (Harj.teht. 4)

$$\boxed{E(f) = -\frac{1}{12}\,(b - a)^3 f''(\xi), \quad \xi \in [a, b] \quad \text{(Trapetsi).}}$$

Keskipistesäännön ja trapetsin virhekaavoista nähdään, että jos f on ylöspäin kaareutuva ($f''(x) > 0$ välillä $[a, b]$), niin keskipistesääntö antaa integraalille liian pienen ja Trapetsi liian suuren arvon. Em. virheanalyysiin liittyen ilmiön voi selittää niin, että keskipistesääntö integroi oikein f:n Taylorin polynomin $T_1(x, c)$, kun taas Trapetsi integroi oikein f:n interpolaatiopolynomin, vrt. kuvio.

Myös Simpsonin säännölle voidaan johtaa virhekaava samantyyppisellä päättelyllä kuin edellä. Olettaen, että f on neljä kertaa jatkuvasti derivoituva välillä $[a, b]$, saadaan virhekaavaksi (Harj.teht. 8)

$$\boxed{E(f) = -\frac{1}{90}\, h^5 f^{(4)}(\xi), \quad h = (b-a)/2, \quad \xi \in [a, b] \quad \text{(Simpson)}.}$$

Em. virhekaavoista voidaan edelleen johtaa virhearvioita yhdistetyille kvadratuureille. Tarkastellaan jälleen esimerkkinä keskipistesääntöä, rajoittuen tasaväliseen jakoon eli olettaen väli $[a, b]$ jaetuksi osaväleihin $[x_{k-1}, x_k]$, $k = 1 \ldots n$, missä $x_k - x_{k-1} = h = (b-a)/n$ $\forall k$. Olkoon funktio f kahdesti jatkuvasti derivoituva välillä $[a, b]$, ja olkoon M_1 ja M_2 f'':n minimi- ja maksimiarvot välillä $[a, b]$. Tällöin jos $E_k(f)$ = keskipistesäännön integrointivirhe osavälillä $[x_{k-1}, x_k]$, niin em. virhekaavan perusteella

$$\frac{1}{24}\, M_1 h^3 \le E_k(f) \le \frac{1}{24}\, M_2 h^3.$$

Summaamalla yli k:n ja huomioimalla, että $nh = b - a$ saadaan kokonaisvirheelle $E(f) = \sum_{k=1}^{n} E_k(f)$ arviot

$$\frac{1}{24}\,(b-a) M_1 h^2 \le E(f) \le \frac{1}{24}\,(b-a) M_2 h^2.$$

Tässä on $M_1 = f(x_1)$, $M_2 = f(x_2)$ joillakin $x_1, x_2 \in [a, b]$, joten päädytään väliarvomuotoiseen virhekaavaan kuten edellä. Tasavälisen trapetsin ja Simpsonin tapauksessa menetellään vastaavasti, jolloin tuloksena ovat seuraavat virhekaavat:

Tasavälinen, yhdistetty kp.-sääntö:	$E(f) = \frac{1}{24}(b-a)f''(\xi)\,h^2$
Tasavälinen, yhdistetty trapetsi:	$E(f) = -\frac{1}{12}(b-a)f''(\xi)\,h^2$
Tasavälinen, yhdistetty Simpson:	$E(f) = -\frac{1}{180}(b-a)f^{(4)}(\xi)\,h^4$

Tässä on kaikissa tapauksissa $\xi \in [a,b]$, ja h on lähimpien kvadratuuripisteiden väli, eli joko h = osavälin pituus (Keskipistesääntö, Trapetsi), tai h = puolet osavälin pituudesta (Simpson).

Muun kuin tasavälisen jaon tapauksessa voidaan em. yhdistettyjen kvadratuurien virhe arvioida muodossa $|E(f)| \leq \ldots$, missä oikealla puolella on $\pm f^{(k)}(\xi)$:n sijasta $|f^{(k)}(x)|$:n maksimiarvo välillä $[a,b]$ ja h on peräkkäisten kvadratuuripisteiden suurin väli. Yleisesti jos osavälijakoon perustuvan yhdistetyn numeerisen kvadratuurin virhe on suuruusluokkaa $\mathcal{O}(h^r)$ (integroitavan funktion f ollessa riittävän säännöllinen), mutta ei luokkaa $\mathcal{O}(h^{r+1})$ (vaikka f olisi kuinka säännöllinen), niin sanotaan, että ko. menetelmän (tarkkuuden) *kertaluku* (engl. order of accuracy) on r. Kuten em. virhetarkastelusta voi päätellä, on yleisesti $r = m+1$, jos numeerinen kvadratuuri integroi osaväleillä tarkasti polynomin astetta m mutta ei polynomia astetta $m+1$. Yhdistetty keskipistesääntö ja trapetsi ovat siis toisen kertaluvun menetelmiä, ja yhdistetyn Simpsonin kertaluku on $r = 4$. Esimerkkejä ensimmäisen kertaluvun menetelmistä ovat Riemannin summakaavat, joissa kvadratuuripisteet (eli välipisteet ξ_k, vrt. Luku VIII.4) eivät ole osavälien keskipisteitä (ks. Harj.teht 6). Riemannin summiin perustuvista approksimaatioista yhdistetty keskipistesääntö on siis tarkkuutensa puolesta omassa luokassaan.

Sovellusesimerkki: Adaptiivinen Simpson

Laskimissa ja numeeris–symbolisissa tietokoneohjelmistoissa numeerisen integroinnin komentojen (esim. Mathematica: `NIntegrate`) takana on usein yhdistetty Simpson, mahdollisesti myös korkeamman kertaluvun yhdistettyjä kvadratuureja. Jakoa osaväleihin ei näissä ohjelmissa yleensä suoriteta tasavälisesti, vaan integroitavaan funktioon sopeutuen *adaptiivisesti*. Adaptiivisen Simpsonin menetelmän ideana on selvittää 'nuuskimalla', kuinka suuri on integroitavan funktion neljäs derivaatta $f^{(4)}(x)$ tarkasteltavalla osavälillä $[x_{k-1}, x_k]$. Perusajatus on yksinkertainen: Jos osaväli on lyhyt (niinkuin laskun kuluessa ennen pitkää on), niin voidaan olettaa, että $f^{(4)}(x)$ on ko. välillä likimain vakio $= M_k$. Tällöin vakio M_k saadaan selville tulkitsemalla itse algoritmin antamia tuloksia *a posteriori* (eli laskemisen jälkeen) seuraavasti: Sovelletaan ensin Simpsonin sääntöä välillä $[x_{k-1}, x_k]$ – tulos I_1. Jaetaan sitten väli puoliksi ja sovelletaan Simpsonin sääntöä

kummallakin osavälillä erikseen – tulos I_2. Jos nyt integraalin tarkka arvo $= I$ välillä $[x_{k-1}, x_k]$, niin Simpsonin virhekaavan mukaan

$$\begin{cases} I - I_1 = -\frac{1}{90} M_k h^5, \\ I - I_2 = -\frac{1}{90} M_k \cdot 2 \cdot \left(\frac{h}{2}\right)^5 \end{cases} \quad \Rightarrow \; M_k = 96 h^{-5}(I_1 - I_2), \quad h = \tfrac{1}{2}(x_k - x_{k-1}).$$

Todellisuudessa $f^{(4)}$ ei ole aivan vakio edes lyhyellä välillä (jos olisi, saataisiin myös integraalin tarkka arvo I selville!), mutta saatu arvio on yleensä adaptiivisiin tarkoituksiin riittävä: Sen avulla voidaan ohjata algoritmia tihentämään jakoa siellä, missä laskun antama luku M_k on itseisarvoltaan suuri. Algoritmi pyrkii tarkemmin tihentämään jakoa niin, että virhetiheys, eli integroimisvirhe osavälillä jaettuna osavälin pituudella, on suunnilleen sama jokaisella osavälillä. (Tällainen jako on osoitettavissa laskutyön kannalta optimaaliseksi.)

Adaptiivinen algoritmi toimii käytännössä hämmästyttävän hyvin myös useimmissa sellaisissa tilanteissa, joissa funktio ei ole lainkaan niin säännöllinen kuin em. laskussa oletetaan (eli neljä kertaa jatkuvasti derivoituva). Näin käy vaikkapa Esimerkin 2 integroimistehtävässä, jossa integroitava funktio ei ole edes jatkuvasti derivoituva ($f(x) \approx 1 - \sqrt{x}$ pisteen $x = 0$ lähellä). Laskemalla lukuja M_k ym. tavalla algoritmi päätyy tihentämään jakoa voimakkaasti origon lähellä ja pystyy tämän 'koneälyn' ansiosta laskemaan integraalin arvon vaaditulla tarkkuudella lähes yhtä nopeasti kuin Esimerkin 1 tilanteessa (ks. Harj.teht. 10c).

HARJOITUSTEHTÄVIÄ

1. Integraali $\int_0^1 f(x)\,dx$ lasketaan tasavälisellä, yhdistetyllä Trapetsilla jakamalla integroimisväli n osaväliin. Laske näin saadun likiarvon virhe tarkasti, kun a) $f(x) = x(1-x)$, b) $f(x) = e^x$.

2. Laske seuraaville integraalille likiarvot käyttämällä 11 pisteen tasavälistä, yhdistettyä Trapetsia (10 osaväliä) ja Simpsonin sääntöä (5 osaväliä) ja vertaa tarkkaan arvoon.

$$\text{a)} \int_0^1 \frac{1}{1+x^2}\,dx \qquad \text{b)} \int_0^1 \frac{4x^3}{1+x}\,dx \qquad \text{c)} \int_0^1 \frac{1}{1+\sqrt[4]{x}}\,dx$$

3. Numeerisen integroinnin nk. *korjattu trapetsikaava* on

$$\int_a^b f(x)\,dx \approx \frac{1}{2}(b-a)[f(a)+f(b)] - \frac{1}{12}(b-a)^2[f'(b) - f'(a)].$$

Näytä, että kaava on tarkka polynomeille astetta $m = 3$. Millainen on vastaava yhdistetty kvadratuuri, jos jako osaväleihin on tasavälinen?

4. Olkoon f kahdesti jatkuvasti derivoituva välillä $[a, b]$ ja $g = f - p$, missä p on ensimmäisen asteen (Lagrangen) interpolaatiopolynomi, jolle pätee $p(a) = f(a)$ ja $p(b) = f(b)$. Näytä osittain integroimalla, että pätee

$$-\frac{1}{2}\int_a^b (x-a)(b-x)g''(x)\,dx = \int_a^b g(x)\,dx.$$

Johda Trapetsin virhekaava tästä tuloksesta.

5. a) Olkoon p funktion f toisen asteen interpolaatiopolynomi, joka määritellään ehdoilla $p(a-h) = f(a-h)$, $p(a) = f(a)$ ja $p(a+h) = f(a+h)$ välillä $[a-h, a+h]$. Näytä, että Simpsonin sääntö ko. välillä on sama kuin approksimaatio

$$\int_{a-h}^{a+h} f(x)\,dx \approx \int_{a-h}^{a+h} p(x)\,dx.$$

b) Todista Simpsonin säännön virhekaava siinä tapauksessa, että $f^{(4)}$ on vakio välillä $[a-h, a+h]$ (eli f on polynomi astetta 4).

6. Integraalin approksimoiminen Riemannin summalla, jossa välipisteiksi osaväleillä $[x_{k-1}, x]$ valitaan $\xi_k = x_{k-1}$, vastaa osaväleillä tehtyä approksimaatiota $\int_a^b f(x)\,dx \approx (b-a)f(a)$. Johda tälle virhekaava

$$E(f) = \frac{1}{2}(b-a)^2 f'(\xi), \quad \xi \in [a, b].$$

Mikä on virhekaava koko välillä, jos jako osaväleihin on tasavälinen?

7. (*) (Arkhimedeen algoritmi) Suora leikkaa paraabelin $K:\ y = x^2$ pisteissä (a, a^2) ja (b, b^2), jolloin suora erottaa paraabelista segmentin A välillä $[a, b]$ $(a < b)$. Olkoon A_n segmentin pinta-alan $\mu(A)$ likiarvo, joka saadaan jakamalla väli $[a, b]$ tasavälisesti 2^n osaväliin $(n \in \mathbb{N})$ ja käyttämällä yhdistettyä trapetsia. Näytä pelkin algebran keinoin, että jono $\{A_n\}$ on geometrinen sarja. Laske $\mu(A)$ $(= \lim_n A_n)$ tällä perusteella.

8. (*) a) (Vrt. Tehtävä 5a.) Olkoon p funktion f yleistetty kolmannen asteen interpolaatiopolynomi, joka määritellään ehdoilla $p(a-h) = f(a-h)$, $p(a+h) = f(a+h)$, $p(a) = f(a)$ ja $p'(a) = f'(a)$ välillä $[a-h, a+h]$. Näytä, että Simpsonin sääntö ko. välillä on sama kuin approksimaatio

$$\int_{a-h}^{a+h} f(x)\,dx \approx \int_{a-h}^{a+h} p(x)\,dx.$$

b) Johda Simpsonin virhekaava. *Vihje*: Lause VII.6.4.

9. (*) Olkoon f neljä kertaa jatkuvasti derivoituva välillä $[a,b]$. Johda korjatulle trapetsikaavalle (Harj.teht. 3) virhekaava

$$E(f) = \frac{1}{720}\,(b-a)^5 f^{(4)}(\xi), \quad \xi \in [a,b].$$

Vihje: Hermiten interpolointi ja Lause VII.6.4!

10. (*) Tarkastellaan numeerisen integroinnin virhettä laskettaessa integraali $\int_0^1 \sqrt{x}\,dx$ yhdistetyillä kvadratuureilla, joissa väli $[0,1]$ jaetaan eri tavoilla n osaväliin.

a) Näytä, että jos jako on tasavälinen, niin sekä yhdistetyn trapetsin että yhdistetyn Simpsonin virhe on $\sim n^{-3/2}$ suurilla n:n arvoilla.
b) Valitaan välin $[0,1]$ jakopisteiksi $x_k = (k/n)^{4/3}$, $k = 0\ldots n$ ja käytetään yhdistettyä trapetsia. Näytä, että virhe $= \mathcal{O}(n^{-2}\ln n)$.
c) Valitaan välin $[0,1]$ jakopisteiksi $x_k = (k/n)^{8/3}$, $k = 0\ldots n$ ja käytetään yhdistettyä Simpsonia. Näytä, että virhe $= \mathcal{O}(n^{-4}\ln n)$. (Adaptiivinen Simpsonin algoritmi päätyy likimain tähän jakoon suurilla n:n arvoilla.)

11. (*) (Hitaan sarjan kiihdytys) Suppeneva integraali $\int_n^\infty f(x)\,dx$, missä $n \in \mathbb{N}$ on suuri luku, voidaan laskea likimäärin käyttämällä joko a) yhdistettyä trapetsia tai b) yhdistettyä, korjattua trapetsia (Harj.teht. 3) välin $[n,\infty)$ jaossa osaväleihin $[k,k+1]$, $k \in \mathbb{N}$, $k \geq n$. Lähtien näistä ajatuksista ja mainittujen kvadratuurien virhearvioista (ks. Harj.teht. 9) johda seuraavat tulokset, kun $f(x) = 1/x^\alpha$, $\alpha > 1$:

$$\text{a)}\quad \sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \sum_{k=1}^{n} \frac{1}{k^\alpha} + \frac{1}{\alpha-1}\,n^{1-\alpha} - \frac{1}{2}\,n^{-\alpha} + \mathcal{O}(n^{-\alpha-1})\,.$$

$$\text{b)}\quad \sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \sum_{k=1}^{n} \frac{1}{k^\alpha} + \frac{1}{\alpha-1}\,n^{1-\alpha} - \frac{1}{2}\,n^{-\alpha} + \frac{\alpha}{12}\,n^{-\alpha-1} + \mathcal{O}(n^{-\alpha-3})\,.$$

Vihje: Huomioi, että

$$\text{a)}\quad \sum_{k=n+1}^{\infty} f(k) = \sum_{k=n}^{\infty} \frac{1}{2}\,\big[\,f(k)+f(k+1)\,\big] - \frac{1}{2}\,f(n),$$

$$\text{b)}\quad \sum_{k=n+1}^{\infty} f(k) = \sum_{k=n}^{\infty} \left(\frac{1}{2}\,\big[\,f(k)+f(k+1)\,\big] - \frac{1}{12}\,\big[\,f'(k+1)-f'(k)\,\big]\right) - \frac{1}{2}\,f(n) - \frac{1}{12}\,f'(n).$$

Luku IX

Differentiaaliyhtälöt

”Luonnonlait on kirjoitettu matematiikan kielellä”, lausui fyysikko ja tähtitieteilijä **Galileo Galilei** (1564-1642). Galilein yhä ajankohtainen lausuma tuo erityisesti mieleen *differentiaaliyhtälöt*, sillä melkeinpä kaikki luonnonlait ovat sellaisia. Paitsi luonnonlaeissa, differentiaaliyhtälöitä tavataan nykyisin mitä erilaisimmissa (ihmisen luomissa) fysiikan, biologian, talousieteen ym. matemaattisissa malleissa. — Tällaisten ’kantaäitinä’ voi pitää aiemmin Luvuissa VI.2 ja VI.4 tarkasteltua eksponentiaalisen kasvun tai vaimenemisen mallia.

Tässä luvussa käydään ensin läpi differentiaaliyhtälöihin liittyvät peruskäsitteet (Luku IX.1), minkä jälkeen tarkastellaan differentiaaliyhtälöiden klassisia ratkaisumenetelmiä ja näihin liittyviä sovellusesimerkkejä (Luvut IX.2–IX.6). Differentiaaliyhtälöiden perinteinen ratkaisemistekniikka on Luvuissa VIII.1–VIII.3 läpikäytyyn integroimistekniikkaan pitkälti perustuva (ja tähän tekniikkaan verrattavissa oleva) matematiikan taitolaji. Joillekin differentiaaliyhtälöiden erikoistyypeille ratkaisut ovat löydettävissä pelkällä ’sivistyneellä arvauksella’, muissa tapauksissa ratkaiseminen pyritään palauttamaan ’integroimisiin’ eli *kvadratuureihin*.

Differentiaaliyhtälöiden matemaattisessa paljoudessa perinteisin menetelmin ratkeavia voi pitää harvinaisuuksina, mutta sovelluksissa tällaiset erikoistapaukset ovat kuitenkin melko yleisiä. Luvuissa IX.2–IX.6 käydään läpi näistä erikoistapauksista tavallisimmat. Luvussa IX.7 tarkastellaan perinteisiä ratkaisumenetelmiä yleispätevämpiä *numeerisia* menetelmiä, joilla differentiaaliyhtälöiden ja myös useamman yhtälön muodostamien *differentiaaliyhtälösysteemien* ratkaisuja voidaan laskea likimäärin. Likimääräisen ratkaisemisen ideoihin perustuu myös differentiaaliyhtälöiden teorian päälause, *Picardin–Lindelöfin lause*, joka esitellään Luvussa IX.8.

IX.1 Differentiaaliyhtälöiden peruskäsitteet

Differentiaaliyhtälön (jatkossa lyhennys DY), tai tarkemmin *tavallisen differentiaaliyhtälön* (engl. ordinary differential equation eli ODE) yleinen muoto on

$$F(x, y, y', \ldots, y^{(n)}) = 0,$$

missä x on reaalimuuttuja, $y = y(x)$ on tuntematon reaaliarvoinen funktio, ja F jokin (tunnettu) $n + 2$ reaalimuuttujan lauseke (kyseessä on funktio tyyppiä $F : \mathbb{R}^{n+2} \to \mathbb{R}$, vrt. Luku IV.3). Kun yhtälö kirjoitetaan tarkemmin muodossa

$$F(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0,$$

niin nähdään selvemmin, että yhtälössä on vain yksi vapaa muuttuja (x). Nimitys 'tavallinen' DY tulee juuri tästä ominaisuudesta. ('Epätavallisista' differentiaaliyhtälöistä ei toistaiseksi puhuta.)

Ym. differentiaaliyhtälössä indeksi $n \in \mathbb{N}$ on yhtälön *kertaluku* (engl. order). Jos yhtälö kirjoitetaan muotoon

$$y^{(n)} = f(x, y, \ldots, y^{(n-1)}),$$

niin tällaista muotoa sanotaan differentiaaliyhtälön *normaalimuodoksi*. Ensimmäisen kertaluvun differentiaaliyhtälön normaalimuoto on siis

$$y' = f(x, y).$$

Normaalimuoto on matemaattisen teorian kannalta sikäli edullinen, että siihen perustuen voidaan differentiaaliyhtälön ratkeavuusehdot asettaa paljon helpommin kuin yleisemmässä muodossa.

Yksittäisratkaisu. Yleinen ratkaisu

Jos differentiaaliyhtälön kertaluku on n, niin sen ratkaisu, tarkemmin *yksittäisratkaisu*, on jokainen funktio $y(x)$, joka on jollakin <u>avoimella</u> <u>välillä</u> (a, b) (voi olla $a = -\infty$ ja/tai $b = \infty$) <u>n</u> <u>kertaa</u> <u>derivoituva</u> ja toteuttaa yhtälön ko. välillä. Differentiaaliyhtälön *yleisellä ratkaisulla* tarkoitetaan sellaista <u>funktiojoukkoa</u>, joka sisältää kaikki (tai 'melkein kaikki', ks. huomatukset jäljempänä) ratkaisut. Jos yhtälön kertaluku on n, niin yleinen ratkaisu on pääsääntöisesti muotoa

$$y = Y(x, C_1, \ldots, C_n),$$

missä $C_1, \ldots, C_n$ ovat vapaasti (tai esim. joiltakin väleiltä vapaasti) valittavia *vakioita*, ja Y on jokin $n+1$:n muuttujan funktio. Sikäli kuin vakiot ovat täysin vapaasti valittavissa, yleinen ratkaisu on siis funktiojoukko

$$\mathcal{Y} = \{\, y(x) = Y(x, C_1, \ldots, C_n) \mid C_1, \ldots, C_n \in \mathbb{R} \,\}.$$

Differentiaaliyhtälön 'ratkaisemisella' tarkoitetaan yleensä yleisen ratkaisun määrittämistä.

Esimerkki 1 Differentiaaliyhtälöiden

a) $y' = e^{2x}$, b) $y' = 2y$

yleiset ratkaisut $\mathbb{R}$:ssä (välillä $(-\infty, \infty)$) ovat

a) $y(x) = \frac{1}{2}\, e^{2x} + C \;\; (C \in \mathbb{R})$, b) $y(x) = Ce^{2x} \;\; (C \in \mathbb{R})$. $\square$

Esimerkki 2 Ratkaise: a) $y''' = 0$, b) $y''' + 2y'' = 0$.

<u>Ratkaisu</u> a) Integroimalla saadaan

$$\begin{aligned} y'''(x) = 0 \quad &\Rightarrow \quad y''(x) = C_1 \\ &\Rightarrow \quad y' = C_1 x + C_2 \\ &\Rightarrow \quad y(x) = \frac{1}{2} C_1 x^2 + C_2 x + C_3. \end{aligned}$$

Tässä voidaan C_1:n tilalle yhtä hyvin kirjoittaa $2C_1$, jolloin yleiselle ratkaisulle saadaan luontevampi muoto

$$y(x) = C_1 x^2 + C_2 x + C_3 \quad (C_1, C_2, C_3 \in \mathbb{R}).$$

b) Ratkaisu voidaan tässä keksiä kirjoittamalla ensin $y''(x) = u(x)$, jolloin yhtälö yksinkertaistuu muotoon $u' + 2u = 0$. Tämän yleinen ratkaisu on $u(x) = C_1 e^{-2x}$. Koska $u = y''$, niin integroimalla seuraa

$$\begin{aligned} y''(x) = C_1 e^{-2x} \quad &\Rightarrow \quad y'(x) = \int C_1 e^{-2x}\, dx = -\frac{1}{2} C_1 e^{-2x} + C_2 \\ &\Rightarrow \quad y(x) = \int y'(x)\, dx = \frac{1}{4} C_1 e^{-2x} + C_2 x + C_3. \end{aligned}$$

Kirjoittamalla C_1:n tilalle $4C_1$ saadaan yleiselle ratkaisulle muoto

$$y(x) = C_1 e^{-2x} + C_2 x + C_3 \quad (C_1, C_2, C_3 \in \mathbb{R}). \quad \square$$

Erikoisratkaisut

Jos differentiaaliyhtälöllä on muitakin ratkaisuja kuin yleiseen ratkaisulausekkeeseen $y = Y(x, C_1, \ldots, C_n)$ sisältyvät, niin tällaisia 'yllätysratkaisuja' sanotaan *erikoisratkaisuiksi*. Milloin ratkaisu on 'erikoinen' ja milloin ei, voi riippua yleisen ratkaisun esitysmuodosta.

Esimerkki 3 Differentiaaliyhtälöiden

$$\text{a) } y' = y^2, \quad \text{b) } (y')^2 = y$$

yleiset ratkaisut ovat

$$\text{a) } y(x) = \frac{1}{C - x}, \quad \text{b) } y(x) = \frac{1}{4}(x - C)^2,$$

kuten saatetaan helposti tarkistaa ($C \in \mathbb{R}$). Yhtälöillä on myös ilmeinen ratkaisu, jota ei saada yleisen ratkaisun lausekkeesta, nimittäin

$$y(x) = 0.$$

Tätä on siis pidettävä erikoisratkaisuna. Tapauksessa a) voi kuitenkin yleisen ratkaisun esittää myös muodossa

$$\text{a) } y(x) = \frac{C}{1 - Cx}$$

(aiemmassa ratkaisussa kirjoitettu C:n tilalle $1/C$), jolloin $y(x) = 0$ sisältyy tähän ($C = 0$). Tapauksessa b) vieläkin 'erikoisempi' ratkaisu on

$$y(x) = \begin{cases} 0, & \text{kun } x \leq C, \\ \frac{1}{4}(x - C)^2, & \text{kun } x > C. \end{cases} \qquad \square$$

Käyräparven differentiaaliyhtälö

Differentiaaliyhtälön yleinen ratkaisu voidaan tulkita geometrisesti *käyräparveksi* (eli käyrien joukoksi). Jos tunnetaan käyräparvi, niin sen differentiaaliyhtälö on johdettavissa derivoimalla. Nimittäin jos käyräparven funktiot ovat muotoa $y = Y(x, C_1, \ldots, C_n)$, niin derivoimalla n kertaa saadaan $n+1$:n yhtälön ryhmä, josta vakiot $C_1, \ldots, C_n$ ovat (ainakin periaatteessa) eliminoitavissa. Tällöin saadaan ko. käyräparvelle differentiaaliyhtälö kertalukua n.

Esimerkki 4 Minkä differentiaaliyhtälön yleinen ratkaisu on

$$\text{a)}\ \ y = (C + x)e^x, \qquad \text{b)}\ \ y = \frac{C_1}{x + C_2}\ ?$$

Ratkaisu a) Derivoimalla kerran saadaan yhtälöryhmä

$$\begin{cases} y = Ce^x + xe^x \\ y' = Ce^x + xe^x + e^x \end{cases}$$

Vähennyslaskulla saadaan differentiaaliyhtälöksi

$$y' - y = e^x.$$

b) Derivoidaan kahdesti:

$$\begin{cases} y \ = C_1(x + C_2)^{-1} \\ y' \ = -C_1(x + C_2)^{-2} \\ y'' = 2C_1(x + C_2)^{-3} \end{cases}$$

Eliminoimalla C_1 ja C_2 saadaan differentiaaliyhtälöksi

$$yy'' = 2(y')^2. \quad \square$$

Kohtisuorat leikkaajat

Jos yksiparametrisen käyräparven $y = Y(x, C)$ differentiaaliyhtälö on

$$y' = f(x, y),$$

niin ratkaisemalla differentiaaliyhtälö

$$y' = -\frac{1}{f(x, y)}$$

löydetään käyräparven *kohtisuorat leikkaajat* (vrt. Esimerkki VII.2:2).

Esimerkki 5 Ympyräparven $x^2+y^2 = C^2$ differentiaaliyhtälöksi saadaan implisiittisesti derivoimalla

$$x + yy' = 0.$$

Kohtisuorien leikkaajien differentiaaliyhtälön

$$x - y/y' = 0$$

yleinen ratkaisu on $y = Cx$, kuten saattoi (geometrisesti) arvata. $\square$

Alku- ja reunaehdot

Koska differentiaaliyhtälön ratkaisu sisältää määräämättömiä vakioita, tarvitaan sovellustilanteissa lisäehtoja, jotta ratkaisu olisi yksikäsitteinen. Lisäehdot on pääteltävä sovellustilanteesta, eli ne kuuluvat matemaattiseen malliin samoin kuin itse differentiaaliyhtälökin.

Jos differentiaaliyhtälön kertaluku on n, niin ratkaisussa on yleensä n määrämätöntä vakiota, jolloin tarvitaan n lisäehtoa. Yksinkertaisin tapa asettaa lisäehdot on kiinnittää jossakin pisteessä x_0 derivaattojen $y^{(k)}(x_0)$ arvot, kun $k = 0 \dots n-1$. Näin saadaan *alkuarvotehtävä* (engl. initial value problem)

$$\begin{cases} F(x, y', \dots, y^{(n)}) = 0, & x \in (a, b), \\ y(x_0) \qquad = A_0, \\ \quad \vdots \qquad\quad \vdots \;\; \vdots \\ y^{(n-1)}(x_0) = A_{n-1}. \end{cases}$$

Tässä voi olla $x_0 \in (a, b)$ tai myös $x_0 = a$ tai $x_0 = b$. Jos x_0 on välin päätepiste, niin alkuehdot on tulkittava derivaattojen $y^{(k)}(x)$ <u>toispuolisina</u> <u>raja-arvoina</u> (tai toispuolisina derivaattoina) kun $x \to a^+$ $(x_0 = a)$ tai $x \to b^-$ $(x_0 = b)$. Edellytys on tällöin, että differentiaaliyhtälön ratkaisuille nämä ovat olemassa.

Esimerkki 6 Jos kappale (massa $= m$) on hetkellä $t = 0$ levossa pisteessä x_0, ja kappaleeseen vaikuttaa voima $f(t)$ kun $t > 0$, niin kappaleen sijainti $x(t)$ hetkellä t saadaan selville ratkaisemalla alkuarvoprobleema

$$\begin{cases} mx''(t) = f(t), & t > 0, \\ x(0) \;= x_0, \\ x'(0) = 0. \end{cases}$$

Alkuehdot voi asettaa täsmällisemmin muodossa $x(0^+) = x_0,\ D_+x(0) = 0.$ $\square$

Lisäehtoja voidaan myös asettaa useammassa pisteessä. Alkuarvotehtävän ohella tyypillisin on *kahden pisteen reuna-arvotehtävä* (engl. two-point boundary value problem), jossa ehdot asetetaan tarkasteltavan välin päätepisteissä. Jos $n = 2$, niin kahden pisteen reuna-arvotehtävän normaalimuoto on

$$\begin{cases} y'' = f(x, y, y'), & x \in (a, b), \\ y(a) = A,\ y(b) = B. \end{cases}$$

Reunaehtojen asettelussa on tässä oletettava, että ratkaisu on oikealta jatkuva a:ssa ja vasemmalta jatkuva b:ssä. Koska ratkaisu on joka tapauksesa (kahdestikin) derivoituvana jatkuva välillä (a, b), niin lisäoletukset tarkoittavat samaa kuin jatkuvuus välillä $[a, b]$.

Esimerkki 7 Ratkaise kahden pisteen reuna-arvotehtävä

$$\begin{cases} y''' + 2y'' = 0, \quad x \in (0,1), \\ y(0) = 1,\ y'(0) = -1,\ y(1) = 0. \end{cases}$$

Ratkaisu Yleinen ratkaisu on (ks. Esimerkki 2)

$$y(x) = C_1 e^{-2x} + C_2 x + C_3,$$

joten saadaan yhtälöryhmä

$$\begin{cases} y(0) &= C_1 & & & + & C_3 & = & 1 \\ y'(0) &= -2\,C_1 & + & C_2 & & & = & -1 \\ y(1) &= e^{-2}\,C_1 & + & C_2 & + & C_3 & = & 0 \end{cases} \quad \Rightarrow \quad \begin{cases} C_1 = 0 \\ C_2 = -1 \\ C_3 = 1 \end{cases}$$

Ratkaisu on siis $y(x) = -x + 1$. $\square$

Differentiaaliyhtälösysteemit

Differentiaaliyhtälösysteemillä tarkoitetaan useamman differentiaaliyhtälön muodostamaa yhtälöryhmää. *Normaalimuotoinen* tavallinen differentiaaliyhtälösysteemi on jollakin $n \in \mathbb{N},\ n \geq 2$ muotoa

$$\begin{cases} y_1' = f_1(x, y_1, \ldots, y_n), \\ y_2' = f_2(x, y_1, \ldots, y_n), \\ \quad \vdots \\ y_n' = f_n(x, y_1, \ldots, y_n). \end{cases}$$

Tässä x on riippumaton muuttuja ja funktiot $y_i(x),\ i = 1 \ldots n$, ovat tuntemattomia. Ratkaisu (yksittäinen tai yleinen) on ko. systeemin jollakin avoimella välillä toteuttavien funktioiden y_i muodostama (järjestetty) joukko

$$\mathbf{y}(x) = (y_1(x), \ldots y_n(x)).$$

Tämä on itse asiassa funktioiden muodostama vektori eli vektoriarvoinen funktio. Tapauksissa $n = 2, 3$ ratkaisu on haluttaessa tulkittavissa tason tai avaruuden parametriseksi käyräksi (parametrina tässä x, vrt. Luku IV.4). *Alkuarvotehtävässä* vaaditaan, että ratkaisu toteuttaa differentiaaliyhtälöiden lisäksi n lisäehtoa muotoa

$$y_i(x_0) = A_i, \quad i = 1 \ldots n.$$

Normaalimuotoinen korkeamman kertaluvun differentiaaliyhtälö

$$y^{(n)} = f(x, y', \ldots, y^{(n-1)})$$

voidaan aina kirjoittaa normaalimuotoiseksi differentiaaliyhtälösysteemiksi. Nimittäin kun kirjoitetaan

$$y_1 = y,\ y_2 = y',\ \ldots,\ y_n = y^{(n-1)},$$

niin nämä yhtälöt yhdessä differentiaaliyhtälön kanssa muodostavat systeemin

$$\begin{cases} y_1' &= y_2, \\ &\vdots \\ y_{n-1}' &= y_n, \\ y_n' &= f(x, y_1, \ldots y_n). \end{cases}$$

Tämä on em. normaalimuotoa.

Esimerkki 8 Differentiaaliyhtälön $y''' = x^2y^3 + x(y')^2 + y''$ systeemimuoto on

$$\begin{cases} y_1' = y_2, \\ y_2' = y_3, \\ y_3' = x^2y_1^3 + xy_2^2 + y_3. \end{cases} \qquad \square$$

Korkeamman kertaluvun differentiaaliyhtälön kirjoittaminen systeemimuotoon auttaa sekä teoreettisissa tarkasteluissa että numeerisissa ratkaisumenetelmissä (ks. Luvut IX.5, IX.7, IX.8). Joskus systeemimuoto on edullinen myös klassisissa ratkaisumenetelmissä (ks. Luku IX.3).

Ratkaisujen säännöllisyys

Alkuarvotehtävää voi pitää implisiittisenä funktion $y(x)$ määritelmänä, jolloin implisiittisen derivoinnin avulla on mahdollista laskea ratkaisufunktion korkeampia derivaattoja alkuarvopisteessä x_0. Tällä tavoin voidaan usein myös selvittää, kuinka säännöllinen ratkaisu on sellaisella välillä, jolla se on (n kertaa derivoituvana) olemassa.

Esimerkki 9 Olkoon $a > 0$ ja tarkastellaan alkuarvotehtävää

$$\begin{cases} y' = x + y^2, \quad x \in (-a, a), \\ y(0) = 0. \end{cases}$$

Jos oletetaan tehtävä ratkeavaksi (kysymystä tarkastellaan myöhemmin Luvussa IX.8), niin differentiaaliyhtälöstä voi päätellä, että ratkaisu on ko. välillä mielivaltaisen monta kertaa derivoituva (sileä). Nimittäin ensinnäkin, koska ratkaisu $y(x)$ on välillä $(-a, a)$ derivoituva (perusoletus), niin differentiaaliyhtälön mukaan se on myös kahdesti derivoituva:

$$y''(x) = \frac{d}{dx}[x + (y(x))^2] = 1 + 2y(x)y'(x) = 1 + 2y(x)[x + (y(x))^2], \quad x \in (-a, a).$$

Tämän mukaan y'' on edelleen derivoituva välillä $(-a, a)$, eli y on kolmesti derivoituva, jne. Päätellään, että $y(x)$ on mielivaltaisen monta kertaa derivoituva välillä $(-a, a)$. Ratkaisufunktion derivaatat alkuarvopisteessä $x = 0$ ovat myös suoraan laskettavissa, sillä koska $y(0) = 0$ (alkuehto), niin differentiaaliyhtälön mukaan on oltava $y'(0) = 0$, jolloin em. lauseke $y''(x)$:lle antaa $y''(0) = 1$. Jatkamalla implisiittistä derivointia nähdään, että ratkaisufunktion <u>kaikki</u> derivaatat pisteessä $x = 0$ määräytyvät yksikäsitteisesti (!). $\square$

Ratkaiseminen kvadratuureilla

Jatkossa tarkastellaan lähinnä sellaisia diferentiaaliyhtälöiden erkoistapauksia, joille on mahdollista laskea 'tarkka' ratkaisu palauttamalla tehtävä tunnettujen funktioiden integraalifunktioiden etsimiseksi. Sanotaan tällöin, että differentiaaliyhtälö ratkeaa *kvadratuureilla* eli 'integroimisilla'. (Kvadratuuri tarkoittaa sananmukaisesti 'neliöimistä', vrt. alaviite Luvussa VIII.9.) Ratkeaminen kvadratuureilla <u>ei</u> edellytä, että integraalifunktiot ovat alkeisfunktioita.

Esimerkki 10 Alkuarvotehtävä

$$\begin{cases} y''' = \sin x/x = 1 - \dfrac{x^2}{6} + \dfrac{x^4}{120} - \dots, & x > 0, \\ y(0) = 1,\ y'(0) = y''(0) = 0 \end{cases}$$

ratkeaa kolmella peräkkeisellä kvadratuurilla:

$$\begin{aligned} y''(x) &= y''(x) - y''(0) = \int_0^x y'''(t)dt = \int_0^x (\sin t/t)dt \\ &= x - \frac{x^3}{18} + \frac{x^5}{600} - \dots \\ y'(x) &= y'(x) - y'(0) = \int_0^x y''(t)dt \\ &= \frac{x^2}{2} - \frac{x^4}{72} + \frac{x^6}{3600} - \dots \end{aligned}$$

$$y(x) = y(0) + [y(x) - y(0)] = y(0) + \int_0^x y'(t)dt$$
$$= 1 + \frac{x^3}{6} - \frac{x^5}{360} + \frac{x^7}{25200} - \dots \qquad \square$$

Esimerkki 9 (jatko) Esimerkin differentiaaliyhtälö $y' = x + y^2$ ei ratkea kvadratuureilla. (Yhtälö on *Riccatin* tyyppiä, ks. Harj.teht. IX.4:10.) $\square$

Viimeisessä esimerkissä ei tarkoiteta, ettei ratkaisuja ole, vaan ainoastaan, että ratkaiseminen ei palaudu 'integroimisiin'. Ratkaisu on tällöin määrättävä muilla keinoin. Esim. jos alkuarvo $y(x_0)$ on tunnettu, voidaan käyttää alkuarvotehtävien numeerisia ratkaisumenetelmiä (ks. Luku IX.7). Myös Taylorin polynomit antavat likimääräistä tietoa ratkaisun kulusta (vrt. Harj.teht. 10).

HARJOITUSTEHTÄVIÄ

1. Tarkista, että $y = 2x + Ce^x$ on differentiaaliyhtälön $y' = y + 2(1-x)$ yleinen ratkaisu. Piirrä pisteiden $(0,1)$ ja $(0,-1)$ kautta kulkevat ratkaisukäyrät.

2. Ratkaise kvadratuureilla (yleinen ratkaisu tai alkuarvotehtävän ratkaisu):

   a) $y'' = \sin x$ b) $y''' = 24x + \cos x$ c) $y'' = \ln x$

   d) $y'' = \dfrac{1}{x^2+1}$, $x \in \mathbb{R}$, $y(0) = 1$, $y'(0) = 0$

   e) $y'' = \dfrac{1}{x^2-2x}$, $x \in (0,2)$, $y(1) = y'(1) = 0$

3. Määritä seuraavien differentiaaliyhtälöiden yleiset ratkaisut palauttamalla yhtälöt ensimmäiseen kertalukuun (sijoitus $u(x) = y^{(k)}(x)$).

   a) $y'' - y' = 0$ b) $2y''' + 3y'' = 0$ c) $y^{(5)} - 5y^{(4)} = 0$

4. Määritä seuraavien käyräparvien differentiaaliyhtälöt

   a) $y = \sin(x + C)$ b) $y = C_1 + C_2 \ln|x|$ c) $y = C_1 + \dfrac{1}{x + C_2}$

   d) $y = (1 + C_1)\ln|x + C_2| - C_1 x + C_2$ e) $x = C_1 e^y + C_2 e^{-y} + 3$

   f) x-kaselia sivuavat ympyrät g) suoraa $y = x$ sivuavat ympyrät

5. Käyrän $y = F(x)$ liukuessa pitkin y-akselia muodostuu käyräparvi. Määritä ko. parven kohtisuorat leikkaajat, kun a) $F(x) = e^x$, b) $F(x) = \ln|x|$.

6. Käyrä $y = u(x)$ leikkaa kohtisuorasti differentiaaliyhtälön $y' = x + y^2$ ratkaisukäyrät. Minkä differentialiyhtälön ratkaisu u on?

7. Esitä normaalimuotoisena differentiaaliyhtälösysteeminä:

a) $yy'' + xy' = 0$ b) $y'' = (x + y')^2 + y'''$ c) $y^{(4)} = \dfrac{y'y''}{1 + x + y'''}$

d) $\begin{cases} u' = (u + v)^2, \\ v'' = x + uv' \end{cases}$ e) $\begin{cases} u'' = uv, \\ v'' = -xuv \end{cases}$ f) $\begin{cases} u^{(4)} = u''v'' + v, \\ v'' = x + u''' + 2v' \end{cases}$

8. Alkuarvotehtävällä

$$\begin{cases} y' = xy + \sin y, \quad x \in \mathbb{R}, \\ y(x_0) = y_0 \end{cases}$$

on yksikäsitteinen ratkaisu jokaisella $(x_0, y_0) \in \mathbb{R}^2$. a) Päättele, että ratkaisu on $\mathbb{R}$:ssä mielivaltaisen monta kertaa derivoituva (sileä). b) Näytä, että ratkaisukäyrä joko sivuaa x-akselia tai ei kosketa sitä lainkaan.

9. Alkuarvotehtävä $xy' = x + y$, $y(1) = 1$ määrittelee pisteen $P = (1, 1)$ kauttaa kulkevan käyrän S. Määritä S:n kaarevuuskeskiö pisteessä P.

10. Määritä alkuarvotehtävän ratkaisufunktion $y(x)$ Taylorin polynomi $T_n(x, 0)$ yleisellä tai annetulla n:n arvolla:
a) $y'' = y$, $y(0) = y'(0) = 1$
b) $y'' = -y$, $y(0) = 1$, $y'(0) = 0$
c) $y' = x + y^2$, $y(0) = 1$; $n = 3$
d) $y' = xy + \sin y$, $y(0) = 1$; $n = 3$
e) $yy'' + y' + y = 0$, $y(0) = 1$, $y'(0) = 0$; $n = 4$
f) $y'' = yy' - x^2$, $y(0) = y'(0) = 1$; $n = 4$

11. (*) (Koirakäyrä) Rekan perävaunu on kiinnitetty vetoautoon akselitapilla, joka on origossa. Vetoauton nokka osoittaa postiivisen y-akselin suuntaan. Perävaunun keskiviiva on x-akselilla, ja perävaunun akselin (siis sen jolla pyörät ovat) keskipiste on pisteessä $(a, 0)$, $a > 0$. Vetoauton liikkuessa pitkin positiivista y-akselia piirtää perävaunun akselikeskiö erään käyrän $y = y(x)$ välillä $0 < x \leq a$. Johda tälle "koirakäyrälle" differentiaaliyhtälö

$$y' = -\frac{\sqrt{a^2 - x^2}}{x},$$

ja määritä käyrä $y = y(x)$ tämän ratkaisuna. Huomioi myös alkuehto.

12. (*) Seuraavat käyräparvet on annettu joko parametrimuodossa tai napakoordinaattien avulla. Johda käyräparvien ja niiden kohtisuorien leikkaajien differentiaaliyhtälöt normaalimuodossa $y' = f(x, y)$.

a) $\begin{cases} x = t + \cos t + C, \\ y = 1 + \sin t \end{cases}$ b) $\begin{cases} x = e^t + t + C, \\ y = 2e^t - t + 2C \end{cases}$

c) $r = C\cos\varphi$ d) $r = C\varphi$ e) $r = Ce^{\varphi}$

13. (*) Käyrää $S : y = y_0(x)$ sanotaan yksiparametrisen käyräparven $y = Y(x, C)$ *verhokäyräksi* (engl. envelope), jos S sivuaa jokaista parven käyrää (eli S:llä ja jokaisella parven käyrällä on yhteinen piste ja siinä yhteinen tangentti).

a) Näytä, että jos käyräparven differentiaaliyhtälö on $F(x, y, y') = 0$, niin myös $y = y_0(x)$ on tämän differentiaaliyhtälön (erikois)ratkaisu.
b) Suoraparvella $y = Cx + 2C^2$, $C \in \mathbb{R}$ on verhokäyränä eräs toisen asteen polynomikäyrä (paraabeli). Määritä tämä ja varmista piirtämällä kuvio!

14. (*) (Sotaharjoitus 3) Tykinkuulan lentorata on parametrinen käyrä $\vec{r} = \vec{r}(t)$ (t = aika). Lentoradan pisteessä $(x, y, z) \mathrel{\widehat{=}} \vec{r}$ kuulaan vaikuttavat voimat ovat

$$\begin{aligned} \vec{G} &= -mg\vec{k}, \\ \vec{T} &= T_1(x, y, z, t)\vec{i} + T_2(x, y, z, t)\vec{j}, \\ \vec{F} &= -k|\vec{v}|\vec{v}, \end{aligned}$$

missä $\vec{G}$ on painovoima (m = massa, g = maan vetovoiman kiihtyvyys), $\vec{T}$ edustaa tuulta ja $\vec{F}$ vauhdin neliöön verrannollista ilmanvastusvoimaa ($\vec{v} = \vec{r}\,'$, k = vakio). Esitä ammuksen liikeyhtälö $m\vec{r}\,'' = \vec{G} + \vec{T} + \vec{F}$ normaalimuotoisena differentiaaliyhtälösysteeminä kokoa $n = 6$ kirjoittamalla

$$(x, y, z, x', y', z') = (y_1, \ldots, y_6).$$

IX.2 Separoituva differentiaaliyhtälö

Ensimmäisen kertaluvun normaalimuotoinen differentiaaliyhtälö on *separoituva*, jos se on muotoa

$$y' = \frac{f(x)}{g(y)},$$

eli jos muuttujat x ja y erottuvat (separoituvat) oikealla puolella. Kirjoittamalla $g(y)$:n tilalle $1/g(y)$ ja/tai $f(x)$:n tilalle $1/f(x)$ saadaan separoituvuudelle muita ilmenemismuotoja. Oletettu muoto on kätevä lähinnä ratkeavuusoletuksien muotoilun kannalta. Jatkossa oletetaan funktioista f ja g, että joillakin $x_1 < x_2$ ja $y_1 < y_2$ pätee

(i) f on jatkuva välillä $[x_1, x_2]$ ja g on jatkuva välillä $[y_1, y_2]$,

(ii) $g(y) \neq 0 \quad \forall y \in [y_1, y_2]$.

Valitaan nyt $x_0 \in (x_1, x_2)$, $y_0 \in (y_1, y_2)$ (avoimet välit!) ja tarkastellaan alkuarvotehtävää

$$\begin{cases} y' = f(x)/g(y), & x \in (x_0 - \delta, x_0 + \delta) \\ y(x_0) = y_0 \end{cases} \tag{$\star$}$$

Tässä valitaan δ ensinnäkin niin, että toteutuu

$$0 < \delta \leq \delta_1 = \min\{x_0 - x_1, x_2 - x_0\},$$

jolloin $[x_0 - \delta, x_0 + \delta] \subset [x_1, x_2]$. Jatkossa tehdään myös toinen oletus, joka koskee alkuarvotehtävän ($\star$) ratkeavuutta. Tämäkin ehto rajoittaa δ:n valintaa, mutta toistaiseksi jätetään avoimeksi, miten. Oletus muotoillaan seuraavasti:

(iii) Alkuarvotehtävällä ($\star$) on välillä $(x_0 - \delta, x_0 + \delta)$ derivoituva ratkaisu $y(x)$, ja pätee $y(x) \in [y_1, y_2] \ \forall x \in (x_0 - \delta, x_0 + \delta)$.

Tehtyjen oletuksien (i)–(iii) perusteella funktiot $f(x)$, $y(x)$ ja $g(y(x))$ ovat jatkuvia välillä $(x_0 - \delta, x_0 + \delta)$ ja $g(y(x)) \neq 0$ ko. välillä. Tähän ja Analyysin peruslauseeseen vedoten voidaan päätellä

$$\begin{aligned} y'(x) = \frac{f(x)}{g(y(x))} \quad &\Leftrightarrow \quad g(y(x))y'(x) = f(x) \\ &\Leftrightarrow \quad \frac{d}{dx}\int_{y_0}^{y(x)} g(t)\,dt = \frac{d}{dx}\int_{x_0}^{x} f(t)\,dt, \quad x \in (x_0 - \delta, x_0 + \delta) \end{aligned}$$

eli

$$y'(x) = \frac{f(x)}{g(y(x))} \quad \Leftrightarrow \quad \frac{d}{dx}[G(y(x)) - F(x)] = 0, \quad x \in (x_0 - \delta, x_0 + \delta),$$
$$G(y) = \int_{y_0}^{y} g(t)dt, \quad F(x) = \int_{x_0}^{x} f(t)dt.$$

Tämän mukaan on oltava $G(y(x)) - F(x) = C =$ vakio välillä $(x_0 - \delta, x_0 + \delta)$ (Lause V.6.7). Toisaalta koska $G(y(x_0)) = G(y_0) = 0$ ja $F(x_0) = 0$, niin $C = 0$, joten on päätelty:

$$G(y(x)) = F(x), \quad x \in (x_0 - \delta, x_0 + \delta).$$

Varmistetaan nyt, että tämä yhtälö määrittelee implisiittisesti funktion $y(x)$. Ensinnäkin oletuksista (i)–(ii) funktiolle g ja Lauseesta V.1.12 seuraa, että

$$\text{joko: } g(y) \geq c > 0 \ \ \forall y \in [y_1, y_2], \qquad \text{tai: } g(y) \leq -c < 0 \ \ \forall y \in [y_1, y_2].$$

(Tässä c:n suurin arvo $= |g(y)|$:n minimiarvo välillä $[y_1, y_2]$.) Tämän perusteella funktio $G(y)$ on joko aidosti kasvava tai aidosti vähenevä välillä $[y_1, y_2]$, siis G on 1–1 ko. välillä.

Koska siis G on 1–1 välillä $[y_1, y_2]$ ja koska oletuksen (iii) mukaan $y(x) \in [y_1, y_2]$ kun $x \in (x_0 - \delta, x_0 + \delta)$, niin voidaan kirjoittaa

$$G(y(x)) = F(x) \ \Leftrightarrow \ y(x) = G^{-1}(F(x)), \quad x \in (x_0 - \delta, x_0 + \delta).$$

Oletuksien (i)–(iii) voimassa ollessa on näin muodoin päätelty:

$$y'(x) = \frac{f(x)}{g(y(x))} \quad \Leftrightarrow \quad y(x) = G^{-1}(F(x)), \quad x \in (x_0 - \delta, x_0 + \delta).$$

Lähtemällä saadusta ratkaisukaavasta nähdään em. päättelyketjusta, että kaavan mukainen $y(x)$ on alkuarvotehtävän $(\star)$ yksikäsitteinen (derivoituva) ratkaistu,

sikäli kuin oletukseen (iii) sisältyvä ehto $y(x) \in [y_1, y_2]$ $\forall x \in (x_0 - \delta, x_0 + \delta)$ on voimassa. On siis enää varmistettava, että myös tämä ehto toteutuu saadulle (ainoalle mahdolliselle) ratkaisulle $y(x) = G^{-1}(F(x))$.

Todetaan ensinnäkin, että em. arvioista funktiolle $g(y)$ sekä integraalien vertailuperiaatteesta (Lause VIII.5.6) seuraa

$$\Big|\int_{y_0}^{y(x)} g(t)dt\Big| \geq c\,|y(x) - y_0| \quad \forall x \in (x_0 - \delta, x_0 + \delta).$$

Toisaalta oletuksen (i) ja Lauseiden V.1.12 ja VIII.5.6 perusteella pätee

$$\Big|\int_{x_0}^{x} f(t)dt\Big| \leq M|x - x_0|, \quad x \in [x_1, x_2],$$

missä $M = \max_{x\in[x_1,x_2]} |f(x)|$. Yhdistämällä arviot seuraa

$$\begin{aligned} c\,|y(x) - y_0| &\leq |G(y(x))| = |F(x)| \leq M|x - x_0| \\ \Rightarrow\ |y(x) - y_0| &\leq \frac{M}{c}|x - x_0| < \frac{M\delta}{c} \quad \forall x \in (x_0 - \delta, x_0 + \delta). \end{aligned}$$

Näin ollen kun valitaan δ siten, että toteutuu

$$0 < \delta \leq \min\{\delta_1, \delta_2\}, \quad \delta_2 = \frac{c}{M}\min\{y_0 - y_1, y_2 - y_0\},$$

niin oletuksen (iii) toteutuminen on varmistettu ja näin todistettu

Lause IX.2.1 Jos $x_0 \in (x_1, x_2)$ ja $y_0 \in (y_1, y_2)$, niin oletuksien (i), (ii) ollessa voimassa on olemassa $\delta > 0$ siten, että alkuarvotehtävällä $(\star)$ on yksikäsitteinen ratkaisu $y(x)$ välillä $(x_0 - \delta, x_0 + \delta) \subset [x_1, x_2]$. Ratkaisukaava ko. välillä on

$$\int_{y_0}^{y(x)} g(t)dt = \int_{x_0}^{x} f(t)dt.$$

Helposti muistettava, muodollinen menettely separoituvan differentiaaliyhtälön yleisen ratkaisun hakemiseksi on kaavio

$$\begin{aligned} \frac{dy}{dx} = \frac{f(x)}{g(y)} \quad &\Rightarrow \quad g(y)\,dy = f(x)\,dx \\ &\Rightarrow \quad \int g(y)\,dy = \int f(x)\,dx. \end{aligned}$$

Tässä ensimmäinen vaihe on 'separointi' ja toinen 'integrointi'. Menettely on pätevä Lauseen IX.2.1 ehdoin, ja alkuehdon $y(x_0) = y_0$ toteuttava ratkaisu saadaan siis muuttamalla määräämättömät integraalit määrätyiksi.

Esimerkki 1 Aiemmin Luvussa VI.2 ratkaistussa alkuarvotehtävässä

$$\begin{cases} y' = y, \quad x \in \mathbb{R}, \\ y(0) = y_0 \end{cases}$$

differentiaaliyhtälö on em. separoituvaa tyyppiä ($f(x) = 1,\ g(y) = 1/y$), joten ratkaisukaava on Lauseen IX.2.1 mukaan

$$\int_{y_0}^{y(x)} \frac{1}{t}\, dt = \int_0^x dt.$$

Kaava on pätevä, jos joko $y_0, y(x) \in (0, \infty)$ tai $y_0, y(x) \in (-\infty, 0)$, jolloin ratkaisuksi saadaan

$$\ln(y/y_0) = x \ \Leftrightarrow\ y(x) = y_0 e^x, \quad x \in \mathbb{R}.$$

Saatu ratkaisu on käypä myös tapauksessa $y_0 = 0$, johon separointimenettely ei sovellu. $\square$

Esimerkki 2 Määritä differentiaaliyhtälön $1 + y^2 + xyy' = 0$ yleinen ratkaisu.

<u>Ratkaisu</u> Separoimalla ja integroimalla saadaan

$$\begin{aligned} \frac{ydy}{1+y^2} = -\frac{dx}{x} \ &\Rightarrow \int \frac{ydy}{1+y^2} = -\int \frac{dx}{x} \\ &\Rightarrow \frac{1}{2}\ln(1+y^2) = -\ln|x| + C. \end{aligned}$$

Kun kirjoitetaan C:n tilalle $\ln C,\ C > 0$, saadaan sieventämällä

$$x^2(1+y^2) = C^2.$$

Ratkaisemalla tästä $y = y(x)$ saadaan kaksihaarainen ratkaisu

$$y(x) = \pm\frac{1}{x}\sqrt{C^2 - x^2}, \quad C > 0.$$

Tämä on pätevä väleillä $(-C, 0)$ ja $(0, C)$. $\square$

Esimerkissä vaihto $C \hookrightarrow \ln C$ perustui siihen, että $\ln C$ saa kaikki mahdolliset reaaliarvot, kun $C \in \mathbb{R}_+$. Tällaisilla manipulaatioilla voidaan ratkaisu usein saattaa yksinkertaisempaan muotoon.

Esimerkki 3 Ratkaise differentiaaliyhtälö $y' = \sin y$.

<u>Ratkaisu</u>
$$\begin{aligned}
\frac{dy}{\sin y} = dx \quad &\Rightarrow \quad \int \frac{dy}{\sin y} = \int dx \\
&\Rightarrow \quad \ln\left|\tan\frac{1}{2}y\right| = x + \ln C, \quad C > 0 \\
&\Leftrightarrow \quad \left|\tan\frac{1}{2}y\right| = Ce^x, \quad C > 0 \\
&\Leftrightarrow \quad \tan\frac{1}{2}y = Ce^x, \quad C \neq 0 \quad (\pm C \hookrightarrow C) \\
&\Leftrightarrow \quad y(x) = 2\arctan(Ce^x), \quad C \neq 0 \\
&\Leftrightarrow \quad y(x) = 2\,\mathrm{Arctan}(Ce^x) + 2n\pi, \quad C \neq 0,\ n \in \mathbb{Z}.
\end{aligned}$$

Saatu ratkaisu on käypä myös kun $C = 0$, sillä $y(x) = 2n\pi$ on alkuperäisen differentiaaliyhtälön ratkaisu jokaisella $n \in \mathbb{Z}$. Nämä ratkaisut voidaan lukea yleiseen ratkaisuun kuuluviksi, vaikka separointimenettely ei niihin suoraan ulotu. Yleiseen ratkaisuun kuulumattomia erikoisratkaisuja ovat lisäksi

$$y(x) = (2n+1)\pi, \quad n \in \mathbb{Z}.$$

Nämä ovat em. ratkaisufunktioiden raja-arvoja, kun $C \to \pm\infty$. $\square$

Autonominen differentiaaliyhtälö

Separoituvaa differentiaaliyhtälöä muotoa

$$y' = f(y),$$

missä siis f ei riipu vapaasta muuttujasta x, sanotaan *autonomiseksi*. Kuten Esimerkissä 3 edellä, jokaista f:n nollakohtaa y_i vastaa vakioratkaisu

$$y(x) = y_i.$$

Tämä on erikoisratkaisu tai mahdollisesti yleiseen ratkaisulausekkeeseen sisällytettävissä oleva, vrt. esimerkit edellä. Separointimenettelyllä ei tällaisia ratkaisuja saada suoraan, koska funktio $g(y) = 1/f(y)$ ei ole jatkuva f:n nollakohtien ympäristössä (vrt. Lause IX.2.1).

Esimerkki 4 Ratkaise differentiaaliyhtälöt $y' = y(1-y)$.

<u>Ratkaisu</u> Separoimalla ja integroimalla saadaan

$$\begin{aligned}
dx = \frac{dy}{y(1-y)} \quad &\Rightarrow \quad \int dx = x = \int \frac{1}{y(1-y)}\,dy = \int \left(\frac{1}{y} + \frac{1}{1-y}\right) dy \\
&\Rightarrow \quad x = \ln\left|\frac{y}{1-y}\right| + \ln C, \quad C > 0.
\end{aligned}$$

Tämän mukaan on

$$\frac{y}{1-y} = \pm C^{-1}e^x \quad \Leftrightarrow \quad y(x) = \frac{e^x}{e^x \pm C}, \quad C > 0$$
$$\Leftrightarrow \quad y(x) = \frac{e^x}{e^x + C}, \quad C \neq 0.$$

Rajoitus $C \neq 0$ voidaan poistaa saadusta yleisen ratkaisun lausekkeesta, koska myös $y(x) = 1$ on ratkaisu. Tämä on kaikkien ratkaisujen yhteinen raja-arvo (asymptoottinen ratkaisu), kun $x \to \infty$. Ratkaisu on edelleen myös $y(x) = 0$, joka saadaan yleisen ratkaisulausekkeen raja-arvona, kun $C \to \infty$. □

Esimerkki 5 : **Logistinen kasvumalli**. Jos fysikaalisen suureen (esim. väkiluvun) kasvua ajan t funktiona kuvaa autonominen differentiaaliyhtälö

$$y' = ay - by^2,$$

missä a ja b ovat (dimensiottomina) positiivisia vakioita, niin kasvumallia (myös differentiaaliyhtälöä) sanotaan *logistiseksi*. Sijoituksella $y(t) = (a/b)\,u(t)$ logistinen DY muuntuu muotoon

$$u' = a(u - u^2).$$

Ratkaisu separoimalla (vrt. edellinen esimerkki):

$$a\,dt = \frac{du}{u(1-u)} \quad \Rightarrow \quad u(t) = \frac{e^{at}}{e^{at} + C}, \quad C \in \mathbb{R}.$$

Ratkaisu on fysikaalisesti järkevä (kasvava) vain kun $C > 0$. Tässä tapauksessa voidaan kirjoittaa $C = e^{at_0}$, $t_0 \in \mathbb{R}$. Kun vielä merkitään $\tau = 1/a$ (aikavakio), niin ratkaisut saadaan muotoon

$$y(t) = A\,Y\left(\frac{t - t_0}{\tau}\right), \quad A = a/b, \quad \tau = 1/a, \quad Y(t) = \frac{e^t}{e^t + 1} \quad (t_0 \in \mathbb{R}). \quad \square$$

Differentiaaliyhtälö $y' = f(y/x)$

Tasa-asteinen differentiaaliyhtälö $y' = f(y/x)$ muuntuu sijoituksella

$$u(x) = y(x)/x \;\Rightarrow\; y(x) = xu(x) \;\Rightarrow\; y' = xu' + u$$

yhtälöksi

$$xu' = f(u) - u.$$

Tämä on separoituva, joten tasa-asteinen yhtälö on separoituvaksi palautuva. Jos $f(u_0) - u_0 = 0$, niin erikoisratkaisu on

$$u(x) = u_0 \quad \Rightarrow \quad y(x) = u_0 x.$$

Yleinen ratkaisu löydetään separoimalla.

Esimerkki 6 Ratkaise differentiaaliyhtälö

$$y' = \frac{x+y}{x-y}\,.$$

Ratkaisu Tämä on tasa-asteinen, joten tehdään sijoitus $u = y/x$:

$$\Rightarrow\; x\frac{du}{dx} = \frac{1+u}{1-u} - u = \frac{1+u^2}{1-u}\,.$$

Separoimalla ja integroimalla seuraa

$$\begin{aligned}
\int \frac{dx}{x} = \ln|x| &= \int \frac{1-u}{1+u^2}\,du \\
&= \operatorname{Arctan} u - \frac{1}{2}\ln(1+u^2) + \ln C \\
\Leftrightarrow\; \operatorname{Arctan} u &= \ln\left(\frac{|x|\sqrt{1+u^2}}{C}\right) \\
\Leftrightarrow\; \operatorname{Arctan}\frac{y}{x} &= \ln\left(\frac{\sqrt{x^2+y^2}}{C}\right) \quad (C>0).
\end{aligned}$$

Siirtymällä napakoordinaatistoon saadaan ratkaisulle helpompi muoto

$$\varphi = \ln\frac{r}{C} \;\Leftrightarrow\; r = Ce^{\varphi} \quad (C>0).$$

Ratkaisut ovat nk. *logaritmisia spiraaleja.* □

Differentiaaliyhtälö $y' = f(ax + by + c)$

Differentiaaliyhtälö $y' = f(ax + by + c)$, missä $a, b, c \in \mathbb{R}$ ja $b \neq 0$, on myös separoituvaksi palautuva. Tässä luonteva sijoitus on

$$u = ax + by + c \;\Rightarrow\; y = \frac{1}{b}(u - ax - c),$$

jolloin

$$\frac{dy}{dx} = \frac{1}{b}\left(\frac{du}{dx} - a\right),$$

ja näin ollen u:lle saadaan separoituva (itse asiassa autonominen) DY

$$u' = bf(u) + a.$$

Esimerkki 7 Ratkaise differentiaaliyhtälö

$$y' = (x - y)^2 + 1.$$

<u>Ratkaisu</u> Sijoituksella

$$x - y = u \;\Leftrightarrow\; y = x - u \;\Rightarrow\; y' = 1 - u'$$

saadaan differentiaaliyhtälöksi

$$u' = -u^2.$$

Separoimalla saadaan yleiseksi ratkaisuksi $u(x) = 1/(x + C)$, joten alkuperäisen yhtälön yleinen ratkaisu on

$$y(x) = x - \frac{1}{x + C}\,.$$

Tämän lisäksi on muunnetun yhtälön ratkaisua $u(x) = 0$ vastaava erikoisratkaisu

$$y(x) = x. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Määritä seuraavien separoituvien tai sellaisiksi palautuvien differentiaaliyhtälöiden yleiset ratkaisut tai ratkaise alkuarvotehtävä. Alkuarvotehtävän tapauksessa selvitä myös, millä välillä ratkaisu on pätevä.

   a) $x^2y' = y^2$ b) $y' = -2xy$ c) $y' = (1-y)^2$ d) $y' = y^2 - 1$

   e) $(1+x)y' = 1+y$ f) $1 + y^2 - xy' = 0$ g) $(1+x^3)y' = x^2y$

   h) $(1-x^2)y' = 1-y^2$ i) $y' = \tan y$ j) $xy' = \cot y$

   k) $(x-3)y' = -y,\ y(-1) = 1$ l) $y' = y^2,\ y(0) = 4$

   m) $y' + 5x^4y^2 = 0,\ y(0) = 1$ n) $y' - 5x^4y^2 = 0,\ y(0) = 1$

   o) $y' \operatorname{Arctan} y = 1,\ y(1) = 1$ p) $y' \sin x = y \ln y,\ y(\pi/2) = 1$

   q) $(1+e^x)yy' = e^x,\ y(1) = 1$ r) $\cos^2 x \cos(\ln y)y' = y,\ y(\pi/4) = 2$

   s) $y' = \dfrac{y}{x-y}$ t) $(3x+y)y' = y$ u) $xyy' = x^2 + y^2$

   v) $(2x^2+y^2)y' = 2xy$ x) $xy' = y + \tan\dfrac{y-x}{x}$ y) $y' = e^{x-2y}$

   z) $xy' = xe^{y/x} + y,\ y(1) = 0$ å) $xy' = y \ln\dfrac{y}{x},\ y(1) = 1$

   ä) $(x+y)y' = 1 - x - y,\ y(0) = 1$ ö) $y' = (2x+y+3)^2,\ y(0) = 0$

2. Ratkaise alkuarvotehtävä $e^y y' = x + e^y - 1,\ y(0) = y_0$ sijoituksella $u = x + e^y$. Millä y_0:n arvoilla ratkaisu on pätevä koko $\mathbb{R}$:ssä?

3. Määritä differentiaaliyhtälön $y' = 2x|y-1|$ ratkaisukäyrät, jotka sivuavat x-akselia.

4. Määritä seuraavien käyräparvien kohtisuorat leikkaajat.

   a) $y = Cx^2$ b) $y = e^x + C$ c) $x^2 + 2y^2 = C^2$ d) $e^{x+y} = (x+C)^2$

5. Määritä käyrät, joilla on ominaisuus: käyrän ja sen normaalin leikkauspisteen etäisyys ko. normaalin ja x-akselin leikkauspisteestä $= a =$ vakio.

6. Määritä välillä $[0, \infty)$ jatkuva funktio $y(x)$, joka toteuttaa *integraaliyhtälön*

$$1 + \int_0^x \frac{y(t)}{t^2+1}\,dt = y(x), \quad x \geq 0.$$

7. (Hypetia, Utopia ja Apatia) H:n, U:n ja A:n valtakunnissa oli elintaso = 1 vuonna 2000 ja = 1.01 vuonna 2001. Ennusta ko. valtakuntien elintasot vuonna 2100, kun tiedetään, että elintason kasvunopeus on H:ssa suoraan verrannollinen elintason neliöön, U:ssa suoraan verrannollinen elintasoon ja A:ssa kääntäen verrannollinen elintasoon.

8. Vuonna 1960 oli maapallon väkiluku 3.0 mrd ja ja kasvunopeus tuolloin 50 milj./v. Jos oletetaan, että väkiluku $y(t) \to 12$ mrd kun $t \to \infty$, niin mikä olisi väkiluvun pitänyt olla vuonna 2000 logistisen kasvumallin mukaan? (Todellinen väkiluku oli 6.1 mrd.)

9. Laskuvarjohyppääjän putoamisnopeus noudattaa varjon auettua liikelakia $mv' = mg - av - bv^2$, missä $m = 80$ kg, $g = 9.8$ m/s^2, $a = 120$ kg/s ja $b = 4.0$ kg/m. Määritä $v(t)$, $t \geq 0$ (aikayksikkö = s), kun varjon aukeamishetkellä $t = 0$ putoamisnopeus on 20 m/s. Miten rajanopeus $v_\infty = \lim_{t\to\infty} v(t)$ saadaan helpoimmin selville? Hahmottele $v(t)$ aikaväleillä $[0, 1]$ ja $[0, 60]$.

10. Avaruudessa suuntaan $-\vec{k}$ etenevät valonsäteet heijastuvat pyörähdyspinnasta $S : z = u(r)$, $r = \sqrt{x^2 + y^2}$ siten, että jokainen heijastunut säde kulkee origon kautta. Funktion $u(r)$ on tällöin toteutettava tasa-asteinen differentiaaliyhtälö $ru' = \sqrt{r^2 + u^2} + u$ (vrt. Harj.teht. VII.2:12). Ratkaise!

11. (*) Differentiaaliyhtälö
$$y' = f\left(\frac{2x - y + 1}{x - 2y + 1}\right)$$
palautuu tasa-asteiseksi sijoituksilla $x = x_0 + t$, $y = y_0 + u$, kun x_0 ja y_0 valitaan sopivasti. Ratkaise yhtälö tällä tavoin, kun $f(x) = x$.

12. (*) Seuraavat differentiaaliyhtälöt palautuvat tasa-asteisiksi sijoituksilla $x = t^\alpha$, $y = z^\beta$, kun α ja β valitaan sopivasti. Ratkaise!
a) $42(x^2 - xy^2)y' + y^3 = 0$ b) $(x^2y^2 - 1)y' + 2xy^3 = 0$

13. (*) Lieriön muotoisessa astiassa on vettä 200 litraa. Astian pohjaan avataan aukko, jonka pinta-ala = 20 cm^2. Vesi purkautuu aukosta nopeudella $v = \sqrt{2gh}$, missä $g = 9.8$ m/s^2 ja h = veden korkeus astiassa kyseisellä hetkellä. (Purkausnopeus oletetaan samaksi aukon eri kohdissa.) Laske, missä ajassa astia tyhjenee, kun veden korkeus on aluksi $h = 1$ m.

14. (*) Käyrällä $S : y = y(x)$, $x \geq 0$, missä $y(x)$ on jatkuvasti derivoituva välillä $[0, \infty)$, on ominaisuus: Käyrän kaarenpituus välillä $[0, x]$ on sama kuin käyrän ja x-akselin välisen alueen pinta-ala ko. välillä. Määritä kaikki käyrät, jotka toteuttavat tämän ehdon ja lisäksi alkuehdon $y(0) = 2$.

IX.3 Palautuvat toisen kertaluvun DY:t

Yleinen toisen kertaluvun normaalimuotoinen differentiaaliyhtälö $y'' = f(x, y, y')$ ratkeaa kvadratuureilla eräissä erikoistapauksissa. Näistä tarkastellaan tässä vain kahta (sovelluksissa melko yleistä) tapausta, joissa f ei riipu joko y:stä tai x:stä, jolloin differentiaaliyhtälö on joko muotoa $y'' = f(x, y')$ tai $y'' = f(y, y')$. Kummassakin tapauksessa on kyse <u>ensimmäiseen</u> <u>kertalukuun</u> <u>palautuvasta</u> differentiaaliyhtälöstä. Sikäli (ja vain sikäli) kuin ko. 1. kertaluvun DY on separoituva, sellaiseksi palautuva tai muuten kvadratuureilla ratkaistavissa, ratkeaa alkuperäinen DY kvadratuureilla.

Differentiaaliyhtälö $y'' = f(x, y')$ palautuu välittömästi 1. kertalukuun, sillä sijoituksella $y' = u$ tämä voidaan kirjoittaa ryhmäksi

$$\begin{cases} u' = f(x, u), \\ y' = u. \end{cases}$$

Jos tässä ensimmäinen yhtälö on ratkaistavissa muotoon $u = U(x, C_1)$, ratkeaa jälkimmäinen yhtälö yhdellä kvadratuurilla muotoon $y(x) = Y(x, C_1, C_2)$. Kvadratuureilla ratkeava on esim. muotoa $y'' = f(y')$ oleva DY, ks. sovellusesimerkki luvun lopussa. Siirrytään tässä tarkastelemaan mainittua toista erikoistapausta.

Differentiaaliyhtälö $y'' = f(y, y')$

Tässäkin tapauksessa muunnetaan ratkaistava DY ensin ryhmäksi sijoituksella $y' = u$:

$$\begin{cases} u' = f(y, u), \\ y' = u. \end{cases}$$

Esitetään tässä u yhdistettynä funktiona muodossa $u = u(x) = U(y(x)) = U(y)$. (Jos ratkaisu on $y = Y(x)$, niin $U(y) = (u \circ Y^{-1})(y)$ väleillä, joilla Y on kääntyvä.) Mainituin merkinnöin ja em. differentiaaliyhtälöiden perusteella seuraa

$$\begin{aligned} f(y, U(y)) = \frac{du}{dx} = \frac{d}{dx}U(y(x)) = \frac{dU}{dy} \cdot \frac{dy}{dx} = \frac{dU}{dy} \cdot u(x) = \frac{dU}{dy} \cdot U(y) \\ \Rightarrow \quad \frac{dU}{dy} = \frac{f(y, U)}{U}. \end{aligned}$$

Sikäli kuin saatu ensimmäisen kertaluvun DY on ratkaistavissa muotoon $U(y) = V(y, C_1)$, voidaan edelleen y ratkaista separoituvasta (autonomisesta) yhtälöstä

$$y' = U(y) = V(y, C_1).$$

Alkuperäinen toisen kertaluvun DY on näin purettu kahdeksi peräkkäiseksi ensimmäisen kertaluvun DY:ksi, joista jälkimmäinen on separoituva.

Em. ratkaisutavasta saadaan käytännössä helpommin muistettava, kun kirjoitetaan yksinkertaisesti $u(x) = u(y)$ (eli ajatellaan y muuttujaksi x:n sijasta), ja suoritetaan differentiaaliyhtälöryhmässä muodollinen jakolasku puolittain:

$$\left.\begin{aligned} \frac{du}{dx} &= f(y,u) \\ \frac{dy}{dx} &= u \end{aligned}\right\} \quad \Rightarrow \quad \frac{du}{dy} = \frac{f(y,u)}{u}\,.$$

Esimerkki 1 Ratkaise diferentiaaliyhtälö $y'' = yy'$.

<u>Ratkaisu</u> Tässä on $f(y,u) = yu$, joten em. ratkaisutavalla saadaan

$$\begin{aligned} \frac{du}{dy} = y \quad &\Rightarrow \quad u = \int y\,dy = \frac{1}{2}\,y^2 + \frac{C_1}{2} \\ &\Leftrightarrow \quad \frac{dy}{dx} = \frac{1}{2}\,y^2 + \frac{C_1}{2} \quad \Rightarrow \quad x = \int \frac{2}{y^2 + C_1}\,dy, \quad C_1 \in \mathbb{R}. \end{aligned}$$

Tästä eteenpäin on eroteltava tapaukset $C_1 > 0$, $C_1 = 0$ ja $C_1 < 0$. Ensinnäkin jos $C_1 > 0$, niin kirjoitetaan C_1:n tilalle C_1^2 $(C_1 > 0)$, jolloin saadaan

$$\begin{aligned} x &= \frac{2}{C_1^2} \int \frac{1}{(y/C_1)^2 + 1}\,dy = \frac{2}{C_1}\,\mathrm{Arctan}\,\frac{y}{C_1} - \frac{2C_2}{C_1} \\ &\Rightarrow \qquad \underline{\underline{y(x) = C_1 \tan\left(\frac{C_1 x}{2} + C_2\right), \quad C_1 > 0,\ C_2 \in \mathbb{R}.}} \end{aligned}$$

Tapauksessa $C_1 < 0$ saadaan vastaavasti kirjoittamalla C_1:n tilalle $-C_1^2$

$$\begin{aligned} x &= \frac{2}{C_1^2} \int \frac{1}{(y/C_1)^2 - 1}\,dy = -\frac{2}{C_1}\,\mathrm{artanh}\,\frac{y}{C_1} - \frac{2C_2}{C_1} \\ &\Rightarrow \qquad \underline{\underline{y(x) = -C_1 \tanh\left(\frac{C_1 x}{2} + C_2\right), \quad C_1 > 0,\ C_2 \in \mathbb{R}.}} \end{aligned}$$

Lopulta tapauksessa $C_1 = 0$ ovat ratkaisut

$$x = \int \frac{2}{y^2}\,dy = -\frac{2}{y} - C \quad \Rightarrow \quad \underline{\underline{y(x) = -\frac{2}{x + C}\,, \quad C \in \mathbb{R}.}}$$

Näiden lisäksi on vielä otettava mukaan autonomisten differentiaaliyhtälöiden $y' = \frac{1}{2}(y^2 + C_1)$ erikoisratkaisut, kun $C_1 = -C^2 \le 0$:

$$\underline{\underline{y(x) = C, \quad C \in \mathbb{R}.}} \quad \square$$

Differentiaaliyhtälö $y'' = f(y)$

Myös edellä tarkastellun DY-tyypin erikoistapaus $y'' = f(y)$ on sovelluksissa yleinen. Em. ratkaisukaavion mukaisesti tämä palautuu separoituviksi 1. kertaluvun yhtälöiksi, joten yhtälö ratkeaa aina kvadratuureilla.

Esimerkki 2 : **Harmoninen värähtely**. Kappale, jonka massa $= m$, on kiinnitetty jouseen, jonka jousivakio $= k$. Hetkellä t kappaleen etäisyys tasapainoasemasta on $y(t)$, jolloin liikeyhtälö on

$$my'' = -ky.$$

Ratkaise $y(t)$ alkuehdoilla $y(0) = 0$, $y'(0) = v_0$.

<u>Ratkaisu</u> Tässä on $f(y) = -(k/m)y$, joten saadaan

$$\begin{aligned} \frac{du}{dy} &= \frac{f(y)}{u} = -\frac{ky}{mu} \\ &\Rightarrow\ mu\frac{du}{dy} + ky = 0 \\ &\Leftrightarrow\ \frac{d}{dy}\left(\frac{1}{2}mu^2 + \frac{1}{2}ky^2\right) = 0 \\ &\Leftrightarrow\ \frac{1}{2}mu^2 + \frac{1}{2}ky^2 = C. \end{aligned}$$

Tässä välituloksessa on $u = y' =$ kappaleen nopeus, $E_k = \frac{1}{2}mu^2 =$ <u>liike-energia</u>, ja $E_p = \frac{1}{2}ky^2 =$ jousen <u>potentiaalienergia</u>. Tulos tunnetaan <u>energiaperiaatteena</u>: Kokonaisenergia $=$ vakio. Koska alkuhetkellä on $y = 0,\ u = v_0$, niin on oltava

$$C = \frac{1}{2}mv_0^2.$$

Jatkamalla saadusta energian säilymislaista tällä C:n arvolla, kirjoittamalla

$$T = \sqrt{\frac{m}{k}}, \quad L = v_0 T,$$

olettamalla, että $|y| < L$, ja käyttämällä alkuehtoa $y(0) = 0$ saadaan

$$T\frac{dy}{dt} = \sqrt{L^2 - y^2} \quad \Rightarrow \quad \int_0^{y(t)} \frac{1}{\sqrt{L^2 - y^2}}\,dy = \int_0^t \frac{1}{T}\,dt = \frac{t}{T}\,.$$

Välillä $t/T \in (-\pi/2, \pi/2)$ ratkaisu on kirjoitettavissa muodossa

$$\operatorname{Arcsin}(y/L) = t/T \ \Leftrightarrow\ y(t) = L\sin(t/T), \quad t/T \in (-\pi/2, \pi/2).$$

Lopputuloksesta nähdään, että rajoitus $t/T \in (-\pi/2, \pi/2)$ voidaan poistaa. Saatiin siis kaikkilla t:n arvoilla pätevä ratkaisu $y(t) = v_0 T \sin(t/T)$, missä $T = \sqrt{m/k}$. $\square$

Esimerkki 3 : **Räjähdys**. Nopeaa kemiallista rekatiota kuvaa DY-systeemi

$$u' = 3y^2, \;\; y' = u,$$

missä $u = u(t)$ ja $y = y(t)$ (t = aika). Oletetaan alkuehdot $y(0) = 1$ ja $u(0) = 0$. Laske reaktioon kuluva aika T, ts. määritä suurin T siten, että alkuarvotehtävällä on ratkaisu välillä $(0, T)$.

Ratkaisu Derivoimalla jälkimmäinen yhtälö ja eliminoimalla u nähdään, että y toteuttaa differentiaaliyhtälön $y'' = 3y^2$. Ratkaisussa lähdetään kuitenkin alkuperäisestä systeemimuodosta, sillä se on valmiiksi ratkaisukaavioon sopiva:

$$\begin{aligned} \frac{du}{dy} = \frac{3y^2}{u} \;&\Rightarrow\; u\frac{du}{dy} - 3y^2 = 0 \\ &\Leftrightarrow\; \frac{d}{dy}\left(\frac{1}{2}u^2 - y^3\right) = 0 \\ &\Leftrightarrow\; \frac{1}{2}u^2 - y^3 = C. \end{aligned}$$

Koska $y = 1$ ja $u = 0$, kun $t = 0$, niin $C = -1$, joten

$$u = y' = \sqrt{2}\sqrt{y^3 - 1} \quad (y \geq 1).$$

Separoimalla ja integroimalla, ja käyttämällä alkuehtoa $y(0) = 1$, saadaan

$$\frac{1}{\sqrt{2}} \int_1^{y(t)} \frac{dx}{\sqrt{x^3 - 1}} = \int_0^t dx = t \;\Leftrightarrow\; y(t) = \Phi^{-1}(t),$$

missä

$$\Phi(s) = \frac{1}{\sqrt{2}} \int_1^s \frac{dx}{\sqrt{x^3 - 1}}, \quad s \geq 1.$$

Funktio Φ ei ole alkeisfunktio, mutta koska $x^3 - 1 = (x-1)(x^2+x+1) > 3(x-1)$, kun $x > 1$, niin

$$\frac{1}{\sqrt{x^3 - 1}} < \frac{1}{\sqrt{3}} \frac{1}{\sqrt{x - 1}}, \quad \text{kun } x > 1,$$

joten (majoranttiperiaatteen nojalla, ks. Luku VIII.7) $\Phi(s)$ on määritelty suppenevana integraalina jokaisella $s > 1$. Edelleen koska

$$\frac{1}{\sqrt{x^3 - 1}} \sim \frac{1}{x^{3/2}}, \quad \text{kun } x \gg 1,$$

niin Φ:llä on raja-arvo, kun $s \to \infty$, ja myös tämä on laskettavissa suppenevana integraalina:

$$\lim_{s\to\infty} \Phi(s) = \int_1^\infty \frac{dx}{\sqrt{x^3-1}} = \Phi_\infty.$$

Koska $\Phi'(s) = 1/\sqrt{s^3-1} > 0$ kun $s > 1$, niin Φ on välillä $[1,\infty)$ aidosti kasvava. Siis $\Phi : [1,\infty) \mapsto [0,\Phi_\infty)$ on bijektio, ja näin ollen $y(t) = \Phi^{-1}(t)$ on myös bijektio:

$$y : [0,T) \longrightarrow [1,\infty),$$

missä

$$T = \Phi_\infty = \frac{1}{\sqrt{2}} \int_1^\infty \frac{dx}{\sqrt{x^3-1}}$$
$$\approx \underline{\underline{1.71732}}. \quad \square$$

Sovellusesimerkki: Ketjuviiva

Tehtävän kuvaus

Ketju on asetettu riippumaan siten, että ketjun päät ovat pisteissä (b,b) ja $(-b,b)$, $b > 0$, ja ketju kulkee origon kautta. Gravitaation suunta $= -\vec{j}$ ja ketjun paino pituusyksikköä kohti $= \rho$.

Tehtävä Määritä ketjun muoto, eli *ketjuviiva*

$$y = y(x), \quad x \in [-b,b].$$

Ratkaisu

Koska tehtävän kuvaus on fysikaalinen, tarvitaan ensin fysikaalisista periaatteista johdettu matemaattinen malli. Tämän johtamiseksi tarkastellaan pientä ketjun palaa välillä $[x, x+\Delta x]$. Kuvaan on merkitty ketjunpalaan vaikuttavat vaakasuorat sidosvoimat (T), pystysuorat sidosvoimat (F) ja gravitaatiovoima (ΔG). Koska kyseessä on ketju, ei sidosmomentteja (taivutusmomentteja) ole.

Tasapainoyhtälöt ovat

$$\begin{aligned} T(x+\Delta x) - T(x) &= 0, && \text{(vaakasuunta)} \\ F(x+\Delta x) - F(x) &= \Delta G, && \text{(pystysuunta)} \\ F(x)\Delta x - T(x)\Delta y &= \mathcal{O}(\Delta G \Delta x). && \text{(momenttitasapaino)} \end{aligned}$$

Ketjun palan pituus on likimain (vrt. Luku VIII.8)

$$\begin{aligned} \Delta s \cong \sqrt{(\Delta x)^2 + (\Delta y)^2} &= \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2}\,\Delta x \\ &\approx \sqrt{1 + [y'(x)]^2}\,\Delta x, \end{aligned}$$

joten gravitaatiovoima ΔG on likimain

$$\Delta G = \rho g \Delta s \approx \rho g \sqrt{1 + [y'(x)]^2} \Delta x.$$

Jakamalla (Δx):llä saadaan rajalla $\Delta x \to 0$ seuraava matemaattinen malli:

$$\begin{cases} T' = 0, \\ F' = \rho g \sqrt{1 + (y')^2}, \\ F - Ty' = 0. \end{cases}$$

Ensimmäisen ja kolmannen yhtälön perusteella

$$T(x) = T_0 = \text{vakio}, \quad F(x) = T_0 y'(x).$$

Sijoittamalla keskimmäiseen yhtälöön saadaan ketjuviivan differentiaaliyhtälö:

$$ay'' = \sqrt{1 + (y')^2}, \quad a = \frac{T_0}{\rho g}.$$

Tästä tulee separoituva sijoituksella $u(x) = y'(x)$:

$$\begin{aligned} au' = \sqrt{1+u^2} \;&\Rightarrow\; \int \frac{du}{\sqrt{1+u^2}} = \int \frac{dx}{a} \\ &\Rightarrow\; \operatorname{arsinh} u = \frac{x}{a} - \frac{c}{a} \quad (C = -\frac{c}{a}) \\ &\Leftrightarrow\; u(x) = \sinh\left(\frac{x-c}{a}\right). \end{aligned}$$

Tästä ja ehdosta $y(0) = 0$ seuraa

$$y(x) = \int_0^x u(t)\,dt = a\cosh\left(\frac{x-c}{a}\right) - a\cosh\frac{c}{a}.$$

Koska $y(b) = y(-b)$, niin on oltava $c = 0$, eli

$$y(x) = a\left(\cosh\frac{x}{a} - 1\right).$$

Ehto $y(b) = b$ (joka enää on käyttämättä) johtaa transkendenttiseen yhtälöön

$$\cosh\alpha = \alpha + 1, \quad \alpha = b/a.$$

Ratkaisu (ketjuviiva) on siis

$$y(x) = \frac{b}{\alpha}\left(\cosh\frac{\alpha x}{b} - 1\right), \quad \cosh\alpha = \alpha + 1, \quad \alpha > 0.$$

Vakion α numeerinen arvo on

$$\alpha = 1.6161375137743..$$

Em. ratkaisusta nähdään, että <u>yleinen</u> <u>ketjuviiva</u> (gravitaation vaikuttaessa suunnassa $-\vec{j}$) on muotoa

$$y(x) = a\cosh\left(\frac{x-c}{a}\right) + C,$$

missä $a \in \mathbb{R}_+$ ja $c, C \in \mathbb{R}$ ovat määräämättömiä vakioita. Vakiot määräytyvät esimerkiksi antamalla (kuten esimerkissä) kolme pistettä, joiden kautta ketju kulkee. Tunnettu voi myös olla ripustuspisteessä vaikuttava vaakasuora tukivoima T_0 ja/tai ketjun pituus, joka välillä $[x_1, x_2]$ on (vrt. Luku VIII.8)

$$L = \int_{x_1}^{x_2} \sqrt{1 + [y'(x)]^2}\, dx = \int_{x_1}^{x_2} \sqrt{1 + \sinh^2\left(\frac{x-c}{a}\right)}\, dx$$
$$= \int_{x_1}^{x_2} \cosh\left(\frac{x-c}{a}\right) dx = a \Big|_{x_1}^{x_2} \sinh\left(\frac{x-c}{a}\right). \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Ratkaise kvadratuureilla (yleinen ratkaisu tai alkuarvotehtävän ratkaisu). Alkuarvotehtävissä määritä myös suurin väli, jolla ratkaisu on pätevä.

   a) $y'' = (y')^2$ b) $2yy'' = 1 + (y')^2$ c) $y^2y'' = y'$ d) $y'' = (y')^3 e^y$
   e) $y'y''' = 2(y'')^2$ f) $x^2y'' = (y')^2$ g) $y'' = 2x(y')^2$ h) $x^3y'' = 2(y')^2$
   i) $(1 + y^2)y'' = y' + (y')^3,\ y(0) = 1,\ y'(0) = -1$
   j) $y'' = 1 - (y')^2,\ y(0) = 0,\ y'(0) = -1$
   k) $y^3y'' = -1,\ y(0) = 1,\ y'(0) = -1$
   l) $\sqrt{y}y'' = 1,\ y(0) = 1,\ y'(0) = -2$
   m) $3y^2y'y'' = -1,\ y(0) = y'(0) = 1$
   n) $2yy'' - 3(y')^2 = 4y^2,\ y(0) = 2,\ y'(0) = 4$
   o) $2x^2y''y''' = -1,\ y(1) = 0,\ y'(1) = 1,\ y''(1) = 2$
   p) $2y^{(4)} + 3(y'')^2 = 0,\ y(0) = y'(0) = 0,\ y''(0) = y'''(0) = -1$

2. Määritä se yhtälön $yy'' + 2(y')^2 = 0$ ratkaisukäyrä, joka sivuaa suoraa $y = x$ pisteessä $(1, 1)$.

3. Millaisille käyrille $y = y(x)$ pätee: Käyrän kaarenpituus välillä $[0, x]$ on verrannollinen lukuun $y'(x)$?

4. Määritä käyrät $S : y = y(x)$, joiden (merkkiselle) kaarevuudelle pisteessä $P = (x, y) \in S$ pätee laskukaava

$$\kappa = \frac{y''}{[1 + (y')^2]^{3/2}} = y^{-1} \sin\alpha \cos\alpha,$$

   missä $\alpha \in [0, \pi)$ on käyrän tangentin suuntakulma pisteessä P.

5. Pitkin x-akselia liikkuvaan kappaleeseen, jonka massa $= m$, vaikuttaa x-akselin suuntainen voima $\vec{F} = f(x)\vec{i}$, missä $x = x(t)$ on kappaleen paikka hetkellä t. Johda liikeyhtälöstä $mx'' = f(x)$ energiaperiaate

$$\frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 = \int_{x_1}^{x_2} f(x)\,dx,$$

missä $v_i =$ kappaleen vauhti pisteessä x_i $(i = 1, 2)$.

6. (*) Matemaattisessa heilurissa pistemäinen massa heiluu painottoman sauvan (pituus R) varassa, jolloin liikeyhtälö on $R\theta'' = -g\sin\theta$, missä $\theta = \theta(t)$ on heilahduskulma mitattuna tasapainoasemasta.

   a) Kertomalla liikeyhtälö θ':lla johda energiaperiaate muotoa $F(\theta, \theta')$=vakio.
   b) Käyttäen a-kohdan tulosta määrää $\theta(t)$, $t \geq 0$ alkuehdoilla

$$\theta(0) = 0,\ \theta'(0) = \frac{1}{T}, \quad T = \frac{1}{2}\sqrt{\frac{R}{g}}.$$

   (Tässä erikoistapauksessa $\theta(t)$ on alkeisfunktio!)

7. (*) (Jänis ja huuhkaja) Hetkellä H jänis lähtee origosta juoksemaan pitkin positiivista y-akselia vauhdilla v (vakio). Pisteessä $(4, 0)$ oleva huuhkaja huomaa samalla hetkellä jäniksen ja lähtee lentämään matalalla siten, että lentosuunta on koko ajan jänistä kohti. Johda huuhkajan lentokäyrälle $y = y(x)$ differentiaaliyhtälö kertalukua 2 olettaen, että huuhkajan lentovauhti on a) v, b) $2v$. Huomioiden myös alkuehdot määritä lentokäyrä. Missä pisteessä (jos missään) huuhkaja tavoittaa jäniksen?

8. (*) (Laskiainen, 2. lasku) Lumilautailija laskee origosta pisteeseen $P = (a, -b)$ pitkin käyrää $y = -y(x)$ $(a > 0,\ b \geq 0$, gravitaation suunta $-\vec{j}\,)$. Jos laskijan alkuvauhti $= 0$ eikä kitkaa ole, niin energiaperiaatteen mukaan laskijan vauhti pisteessä $(x, -y(x))$ määräytyy yhtälöstä $\frac{1}{2}[v(x)]^2 = gy(x)$, missä g on maan vetovoiman kiihtyvyys. a) Näytä, että laskuaika pisteeseen P on

$$t = \int_0^a \sqrt{\frac{1 + [y'(x)]^2}{2gy(x)}}\,dx.$$

   b) Tästä lausekkeesta on osoitettavissa, että lasku on nopein mahdollinen, jos $y(x)$ toteuttaa differentiaaliyhtälön $2yy'' + 1 + (y')^2 = 0$. Näytä, että tämä nopeimman laskun käyrä (*brakistokroni*) on sykloidin kaari.

IX.4 Ensimmäisen kertaluvun lineaarinen DY

Separoituvan (tai sellaiseksi palautuvan) differentiaaliyhtälön ohella toinen sovelluksissa hyvin yleinen differentiaaliyhtälön tyyppi on *lineaarinen* DY. Tässä luvussa tarkastelun kohteena on 1. kertaluvun lineaarinen differentiaaliyhtälö, jonka yleinen muoto on

$$y' + P(x)y = R(x). \tag{1}$$

Tässä P on nk. *kerroinfunktio*. Sekä P että yhtälön *oikea puoli* eli R oletetaan jatkuviksi välillä (a, b), jolla ratkaisua haetaan. Kuten yleensä, ratkaisulta y edellytetään derivoituvuus välillä (a, b). Koska derivoituvuudesta seuraa jatkuvuus, niin differentiaaliyhtälöstä nähdään, että edelleen myös y' on jatkuva välillä (a, b).

Kun otetaan käyttöön *operaattori*merkintä

$$\mathrm{L}y = \frac{dy}{dx} + P(x)y^{\dagger},$$

niin yhtälö (1) voidaan kirjoittaa lyhyesti

$$\mathrm{L}y = R(x).$$

Yhtälöä sanotaan lineaariseksi sen vuoksi, että pätee

$$\mathrm{L}(c_1y_1 + c_2y_2) = c_1\mathrm{L}y_1 + c_2\mathrm{L}y_2\,, \quad c_1, c_2 \in \mathbb{R},$$

eli L on *lineaarinen* operaattori.[‡].

Jos tunnetaan yksikin lineaarisen yhtälön (1) yksittäisratkaisu $y = y_0(x)$, niin yhtälön yleisen ratkaisun etsiminen helpottuu huomattavasti. Nimittäin jos $y = y(x)$ on mikä tahansa toinen ratkaisu, niin yhtälöistä

$$\begin{cases} \mathrm{L}y = R(x), \\ \mathrm{L}y_0 = R(x) \end{cases}$$

seuraa L:n lineaarisuuden nojalla

$$0 = \mathrm{L}y - \mathrm{L}y_0 = \mathrm{L}(y - y_0),$$

eli erotus $u(x) = y(x) - y_0(x)$ toteuttaa yksinkertaisemman yhtälön

$$y' + P(x)y = 0. \tag{2}$$

[†]Tapauksessa $P = 0$ on $\mathrm{L} = \mathrm{D}$, vrt. Luku V.3. Yleisemmin voidaan kirjoitaa $\mathrm{L} = \mathrm{D} + P(x)$, jolloin kyse on operaattorien $f \mapsto \mathrm{D}f$ ja $f \mapsto P(x)f$ yhteenlaskusta 'funktion funktioina'.

[‡]Jos operaattori ei ole lineaarinen, niin se on *epälineaarinen*. Epälineaarinen on vaikkapa operaattori $\mathrm{L}(y) = y' + y^2$.

Tätä sanotaan (lineaariseksi 1. kertaluvun) *homogeeniseksi* yhtälöksi. Jos $R \neq 0$, niin yhtälöä (1) sanotaan vastaavasti *epähomogeeniseksi* tai, etenkin ratkaisemisen yhteydessä, *täydelliseksi* yhtälöksi. On päätelty:

Lineaarisen yhtälön (1) yleinen ratkaisu = täydellisen yhtälön (1) yksittäisratkaisu + homogeenisen yhtälön (2) yleinen ratkaisu.

Sikäli kuin yksittäisratkaisu $y_0(x)$ tunnetaan, on jäljellä siis homogeenisen yhtälön (2) yleisen ratkaisun etsiminen. Tämä käy helposti, koska yhtälö on separoituva: Luvun IX.2 menetelmin saadaan ratkaisuksi

$$y(x) = Ce^{-\int P(x)dx}.$$

(Tässä ei integraalifunktioon $\int P(x)\,dx$ tarvitse sisällyttää toista määräämätöntä vakiota C_1, koska tämän vaikutus sisältyy jo vakioon C kertoimena e^{-C_1}.) Yhtälö (1) on näin ratkaistu yhdellä kvadratuurilla:

$$y(x) = y_0(x) + Ce^{-\int P(x)dx}.$$

Jos yksityisratkaisua ei tunnetta, ratkeaa yhtälö <u>kahdella</u> kvadratuurilla. Nimittäin yksittäisratkaisu saadaan aina selville käyttäen *vakion varioinnin* nimellä tunnettua (yhden kvadratuurin sisältävää) menettelyä. Vakion variointi tarkoittaa homogeenisen yhtälön yleisessä ratkaisussa tehtävää muutosta $C \hookrightarrow C(x)$ ja näin saatavan funktion

$$y(x) = C(x)e^{-\int P(x)dx}$$

kokeilua täydellisen yhtälön mahdollisena ratkaisuna. Sijoitus yhtälöön antaa

$$\mathrm{L}y = C'(x)e^{-\int P(x)dx} + C(x)\mathrm{L}[e^{-\int P(x)dx}] = R(x).$$

Tässä on $\mathrm{L}[e^{-\int P(x)dx}] = 0$, joten saadaan

$$\begin{aligned} C'(x)e^{-\int P(x)dx} = R(x) \quad &\Leftrightarrow \quad C'(x) = e^{\int P(x)dx}R(x) \\ &\Leftrightarrow \quad C(x) = \int e^{\int P(x)dx}R(x)\,dx + C. \end{aligned}$$

Analyysin peruslauseen nojalla tämä ratkaisu on pätevä välillä (a, b), jolla P ja R ovat jatkuvia. Kun vakio C jätetään yksittäisratkaisussa kiinnittämättä, voidaan yhtälön (1) yleiseksi ratkaisuksi kirjoittaa suoraan

$$y(x) = e^{-\int P(x)dx}C(x) = e^{-\int P(x)dx}\left(C + \int e^{\int P(x)dx}R(x)\,dx\right).$$

Alkuehdon $y(x_0) = Y_0$ ($x_0 \in (a,b)$) toteuttava ratkaisu on yksikäsitteinen ja kirjoitettavissa suoraan määrättyjen integraalien avulla:

$$y(x) = e^{-\int_{x_0}^{x} P(t)dt} \left(Y_0 + \int_{x_0}^{x} e^{\int_{x_0}^{s} P(t)dt} R(s)\, ds \right). \tag{3}$$

Jos tässä kiinnitetään vain x_0 ja annetaan Y_0:n olla muuttuva ($Y_0 \hookrightarrow C$), niin saadaan jälleen yleinen ratkaisu.

Esimerkki 1 Ratkaise differentiaaliyhtälö

$$y' + \frac{2y}{1-x^2} = \frac{1+x}{(1-x)^3}\,.$$

<u>Ratkaisu</u> Ratkaistaan ensin homogeeninen yhtälö:

$$\begin{aligned} \frac{dy}{dx} + \frac{2y}{1-x^2} = 0 \;&\Rightarrow\; \int \frac{dy}{y} = \int \frac{2dx}{x^2-1} \\ &\Rightarrow\; \ln|y| = \ln\left|\frac{x-1}{x+1}\right| + \ln|C| \\ &\Rightarrow\; y = C\,\frac{x-1}{x+1}\,. \end{aligned}$$

Yleinen ratkaisu vakion varioinnilla:

$$\begin{aligned} y = C(x)\,\frac{x-1}{x+1} \;&\Rightarrow\; C'(x)\,\frac{x-1}{x+1} = \frac{1+x}{(1-x)^3} \\ \Rightarrow\; C'(x) &= -\frac{(x+1)^2}{(x-1)^4} = -\frac{[(x-1)+2]^2}{(x-1)^4} \\ &= -\frac{1}{(x-1)^2} - \frac{4}{(x-1)^3} - \frac{4}{(x-1)^4} \\ \Rightarrow\; C(x) &= \frac{1}{x-1} + \frac{2}{(x-1)^2} + \frac{4}{3(x-1)^3} + C \\ &= \frac{3x^2+1}{3(x-1)^3} + C \\ \Rightarrow\; y(x) = C(x)\,\frac{x-1}{x+1} &= \frac{3x^2+1}{3(x-1)^2(x+1)} + C\,\frac{x-1}{x+1}\,. \end{aligned}$$

Ratkaisu on pätevä väleillä $(-\infty,-1)$, $(-1,1)$ ja $(1,\infty)$. $\square$

Integroivan tekijän menettely

Vakion variointia suoraviivaisempi tapa ratkaista yhtälö (1) on kertoa yhtälö ensin nk. *integroivalla tekijällä*, joka määritellään

$$H(x) = e^{\int P(x)dx}.$$

Koska $H' = P(x)H$, niin

$$H(x)[y' + P(x)y\,] = \frac{d}{dx}[H(x)y\,].$$

Näin ollen, ja koska $H(x) > 0\ \forall x$, niin saadaan ratkaisukaavio

$$\begin{aligned} y' + P(x)y = R(x) \quad &\Leftrightarrow \quad H(x)[y' + P(x)y] = H(x)R(x) \\ &\Leftrightarrow \quad [H(x)y\,]' = H(x)R(x) \\ &\Leftrightarrow \quad H(x)y(x) = \int H(x)R(x)\,dx \\ &\Leftrightarrow \quad y(x) = H(x)^{-1}\int H(x)R(x)\,dx, \quad x \in (a,b). \end{aligned}$$

Päättely nojaa jälleen Analyysin peruslauseeseen ja tehtyihin jatkuvuusoletuksiin. Lopputulos on sama kuin edellä, ja alkuarvotehtävän ratkaisu saadaan jälleen vaihtamalla määräämättömät integraalit määrätyiksi.

Esimerkki 2 Ratkaise differentiaaliyhtälö $y' + y\cot x = e^{\cos x}$.

<u>Ratkaisu</u> Koska $\int \cot x\,dx = \ln|\sin x| + C$, niin esim. välillä $(0,\pi)$ (jolla $\cot x$ on jatkuva) voidaan integroivaksi tekijäksi valita

$$H(x) = e^{\ln\sin x} = \sin x.$$

Em. ratkaisukaaviota seuraten päätellään

$$\begin{aligned} y' + y\cot x = e^{\cos x} \quad &\Leftrightarrow \quad y'\sin x + y\cos x = \sin x e^{\cos x} \\ &\Leftrightarrow \quad (y\sin x)' = \sin x\, e^{\cos x} \\ &\Leftrightarrow \quad y\sin x = \int \sin x\, e^{\cos x}\,dx = -e^{\cos x} + C \\ &\Leftrightarrow \quad y(x) = \frac{C - e^{\cos x}}{\sin x}\,. \end{aligned}$$

Ratkaisu on pätevä väleillä $(n\pi, (n+1)\pi),\ n \in \mathbb{Z}$. $\square$

Vakiokertoiminen yhtälö

Sovelluksissa hyvin yleinen yhtälön (1) erikoistapaus on *vakiokertoiminen* 1. kertaluvun lineaarinen DY, joka on muotoa

$$y' + ay = f(x) \quad (a \in \mathbb{R},\ a \neq 0). \tag{4}$$

(Tapaus $a = 0$ sivuutetaan entuudestaan tuttuna.) Tämän ratkaisu alkuehdolla $y(x_0) = Y_0$ on ratkaisukaavan (3) mukaisesti

$$\begin{aligned} y(x) &= e^{-a(x-x_0)} \left(Y_0 + \int_{x_0}^{x} e^{a(t-x_0)} f(t)\, dt \right) \\ &= Y_0 e^{-a(x-x_0)} + \int_{x_0}^{x} e^{a(t-x)} f(t)\, dt. \end{aligned}$$

Ellei haluta käyttää suoraan tätä kaavaa (yleisessä ratkaisussa $Y_0 \hookrightarrow C$), niin yksittäisratkaisu on usein löydettävissä helposti myös 'sivistyneellä arvauksella'. Edellytyksenä on, että $f(x)$ on riittävän yksinkertaista muotoa.

Esimerkki 3 Etsi yhtälölle $y' + y = f(x)$ yksittäisratkaisu $y_0(x)$, kun

a) $f(x) = x^2$, b) $f(x) = e^x$, c) $f(x) = e^{-x}$, d) $f(x) = \cos 2x$.

<u>Ratkaisu</u> a) Kokeillaan, olisiko yksittäisratkaisu toisen asteen polynomi muotoa $y_0(x) = Ax^2 + Bx + C$. Sijoitus yhtälöön antaa

$$Ax^2 + (2A + B)x + (B + C) = x^2.$$

Tämä toteutuu jokaisella x, kun

$$\begin{cases} A = 1 \\ 2A + B = 0 \\ B + C = 0 \end{cases} \Leftrightarrow \begin{cases} A = 1 \\ B = -2 \\ C = 2 \end{cases}$$

Siis yksittäisratkaisu on $y_0(x) = x^2 - 2x + 2$.

<u>Ratkaisu</u> b) Kokeillaan yksittäisratkaisua muotoa $y_0(x) = Ae^x$:

$$2Ae^x = e^x \quad \forall x \ \Leftrightarrow\ A = \frac{1}{2}.$$

Siis $y_0(x) = \frac{1}{2}e^x$ toimii.

Ratkaisu c) Tässä yritys $y_0(x) = Ae^{-x}$ ei onnistu, sillä tämä sattuu olemaan homogeeniyhtälön ratkaisu. Yritetään

$$y_0(x) = Axe^{-x} : \quad Ae^{-x} - Axe^{-x} + Axe^{-x} = e^{-x} \;\Rightarrow\; A = 1.$$

Siis $y_0(x) = xe^{-x}$ toimii.

Ratkaisu d) Yritys: $y_0(x) = A\cos 2x + B\sin 2x$

$$\begin{aligned} &\Rightarrow\; (A+2B)\cos 2x + (-2A+B)\sin 2x = \cos 2x \\ &\Rightarrow \begin{cases} A + 2B = 1 \\ -2A \;\; + B = 0 \end{cases} \Leftrightarrow \begin{cases} A = 1/5 \\ B = 2/5 \end{cases} \end{aligned}$$

Siis $y_0(x) = \frac{1}{5}(\cos 2x + 2\sin 2x)$ on yksittäisratkaisu. □

Yksittäisratkaisua etsittäessä on syytä huomioida myös lineaarisuussääntö:

$$\begin{aligned} &\mathrm{L}y_1 = f_1(x) \;\;\&\;\; \mathrm{L}y_2 = f_2(x) \\ &\quad\Rightarrow\; \mathrm{L}(c_1y_1 + c_2y_2) = c_1f_1(x) + c_2f_2(x). \end{aligned}$$

Esimerkki 3 (jatko) Etsi yksittäisratkaisu, kun

$$\text{e)}\;\; f(x) = 4\sin^2 x - 5x^2 - 2, \quad \text{f)}\;\; f(x) = 6\sinh x - 10\cos 2x.$$

Ratkaisu Koska

$$\text{e)}\;\; f(x) = -2\cos 2x - 5x^2, \quad \text{f)}\;\; f(x) = 3e^x - 3e^{-x} - 10\cos 2x,$$

niin lineaarisuussääntöä ja kohtien a)-d) tuloksia käyttäen saadaan

$$\begin{aligned} \text{e)}\;\; y_0(x) &= -2\cdot\frac{1}{5}(\cos 2x + 2\sin 2x) - 5(x^2 - 2x + 2) \\ &= -\frac{2}{5}\cos 2x - \frac{4}{5}\sin 2x - 5x^2 + 10x - 10, \\ \text{f)}\;\; y_0(x) &= 3\cdot\frac{1}{2}\,e^x - 3xe^{-x} - 10\cdot\frac{1}{5}(\cos 2x + 2\sin 2x) \\ &= \frac{3}{2}\,e^x - 3xe^{-x} - 2\cos 2x - 4\sin 2x. \quad \square \end{aligned}$$

Seuraavan taulukon säännöt ovat em. esimerkin yleistyksiä. Taulukossa oletetaan differentiaaliyhtälö (4), missä $a \neq 0$. Lisäksi $n \in \mathbb{N} \cup \{0\}$ ja $\omega \neq 0$.

$f(x)$	$y_0(x)$ muotoa
x^n	$A_n x^n + \cdots + A_0$
$x^n e^{bx}, \quad b \neq -a$	$(A_n x^n + \cdots + A_0)e^{bx}$
$x^n e^{-ax}$	$A x^{n+1} e^{-ax}$
$x^n(A\cos\omega x + B\sin\omega x)$	$(A_n x^n + \cdots + A_0)\cos\omega x + (B_n x^n + \cdots + B_0)\sin\omega x$

Jaksolliset ratkaisut

Sovelluksissa yleinen vakikertoimisen yhtälön (4) erikoistapaus on sellainen, jossa yhtälön oikea puoli f on jaksollinen. Tällöin yhtälölle on löydettävissä yksittäisratkaisu y_p, joka on samoin jaksollinen, ja y_p:n jakso $= f$:n jakso. (Tapauksessa $a = 0$ ei jaksollista ratkaisua yleisesti ole, ks. Harj.teht. 7.) Olkoon f koko $\mathbb{R}$:ssä määritelty, jatkuva ja L-jaksoinen funktio. Etsitään L-jaksoinen ratkaisu määräämällä tämä yhden jakson pituisella välillä, esim. välillä $[0, L]$. Tällä välillä riittää asettaa jaksollisuusehto $y(0) = y(L)$, joten on siis löydettävä funktio y_p, joka on välillä $[0, L]$ jatkuva, välillä $(0, L)$ derivoituva ja toteuttaa

$$\begin{cases} y' + ay = f(x), & x \in (0, L), \\ y(0) = y(L). \end{cases}$$

Ratkaisu on muotoa $y(x) = Ce^{-ax} + y_0(x)$, missä y_0 on jokin differentiaaliyhtälön yksittäisratkaisu, esim. (vrt. ratkaisukaava edellä)

$$y_0(x) = \int_0^x e^{a(t-x)} f(t)\, dt.$$

Jaksollisuusehto määrää vakion C, jolloin ratkaisuksi saadaan (olettaen $a \neq 0$)

$$y_p(x) = \frac{y_0(L) - y_0(0)}{1 - e^{-aL}}\, e^{-ax} + y_0(x), \quad x \in [0, L].$$

Kun jatketaan y_p koko $\mathbb{R}$:ään jaksollisuusehdolla $y_p(x \pm L) = y_p(x)$, niin on pääteltävissä (Harj.teht. 7a), että näin määritelty y_p on differentiaaliyhtälön ratkaisu $\mathbb{R}$:ssä. Koska y_p siis on yksittäisratkaisu, niin yleinen ratkaisu $\mathbb{R}$:ssä on

$$y(x) = Ce^{-ax} + y_p(x).$$

Alkuehto $y(x_0) = Y_0$ määrää C:n arvoksi $C = (Y_0 - y_p(x_0))e^{ax_0}$.

Sovelluksissa on yleensä $a > 0$ ja muuttujan x tilalla aika (t), jolloin ym. ratkaisussa on siis kaksi osaa, jaksollinen osa $y_p(t)$ ja nk. transientti, joka 'kuolee pois', kun $t \to \infty$. Seuraavassa esimerkki sähkötekniikasta.

Sovellusesimerkki: Sähköpiiri RC

Tehtävän kuvaus

Oheisessa sähköpiirissä on sarjaan kytketty vastus (R) ja kondensaattori, jonka kapasitanssi $= C$. Virta piirissä hetkellä t on $i(t)$, kondensaatorin varaus $= y(t)$ ja jännite kondensaattorin yli $= u(t)$.

Piiriä syötetään kokoaaltotasasuunnatulla jännitteellä

$$e(t) = E|\sin \omega t|.$$

Matemaattinen malli (Ol. kytkin suljettu)

$$\begin{aligned} Ri(t) + u(t) &= e(t), \\ y(t) &= Cu(t), \\ y'(t) &= i(t). \end{aligned}$$

Alkuoletukset

Ajanhetkillä $t \le t_0$ on $i(t) = q(t) = 0$. Kytkin suljetaan hetkellä $t = t_0$, ja pidetään sen jälkeen suljettuna.

Tehtävä

Määritä kondensaattorin varaus $y(t)$, kun $t \ge t_0$.

Ratkaisu

Merkitään $Q = EC$ ja $\tau = RC$ ($\tau =$ aikavakio). Eliminoimalla $i(t)$ ja $u(t)$ saadaan ratkaistavaksi alkuarvotehtävä

$$\begin{cases} \tau y' + y = Q|\sin \omega t|, & t > t_0, \\ y(t_0) = 0. \end{cases}$$

Koska differentiaaliyhtälön oikea puoli on jaksollinen, jaksona

$$T = \pi/\omega,$$

on yhtälölle löydettävissä yksittäisratkaisu y_p, joka on samoin T-jaksoinen. Tämä on välillä $[0,T]$ jatkuva, välillä $(0,T)$ derivoituva ja toteuttaa

$$\begin{cases} \tau y_p' + y_p = Q\sin\omega t, & t \in (0,T), \\ y_p(0) = y_p(T). \end{cases}$$

Etsitään ensin differentiaaliyhtälön yksittäisratkaisu y_0 välillä $(0,T)$:

$$\begin{aligned} & y_0(t) = A\sin\omega t + B\cos\omega t \\ & \Rightarrow\ (A - \tau\omega B)\sin\omega t + (\tau\omega A + B)\cos\omega t = Q\sin\omega t, \quad 0 < t < T \\ & \Leftrightarrow \begin{cases} A - \tau\omega B = Q \\ \tau\omega A + \quad B = 0 \end{cases} \\ & \Leftrightarrow\ A = \frac{Q}{k^2+1}, \quad B = -\frac{kQ}{k^2+1}, \qquad k = \tau\omega. \end{aligned}$$

(Tässä uusi parametri k on dimensioton). Etsityn T-jaksoisen ratkaisun on siis oltava välillä $[0,T]$ muotoa

$$y_p(t) = Ce^{-t/\tau} + \frac{Q}{k^2+1}(\sin\omega t - k\cos\omega t).$$

Vakio C määräytyy jaksollisuusehdosta:

$$\begin{aligned} y_p(0) = y_p(T) \ &\Leftrightarrow\ C - \frac{kQ}{k^2+1} = Ce^{-T/\tau} + \frac{kQ}{k^2+1} \\ &\Leftrightarrow\ C = \frac{2kQ}{(k^2+1)(1-e^{-T/\tau})}\,. \end{aligned}$$

Siis välillä $[0,T]$ pätee

$$y_p(t) = \frac{Q}{k^2+1}\left(\frac{2k}{1-e^{-T/\tau}}\,e^{-t/\tau} + \sin\omega t - k\cos\omega t\right),$$

missä

$$Q = EC, \quad \tau = RC, \quad T = \frac{\pi}{\omega}\,, \quad k = \tau\omega = \pi\frac{\tau}{T}\,.$$

Alkuehdon $y(t_0) = 0$ toteuttava ratkaisu on

$$y(t) = -y_p(t_0)e^{-(t-t_0)/\tau} + y_p(t), \quad t \geq t_0.$$

Jos sattuu olemaan $y_p(t_0) = 0$, ei transientti 'herää'. $\square$

HARJOITUSTEHTÄVIÄ

1. Ratkaise (yleinen ratkaisu tai alkuarvotehtävän ratkaisu):

   a) $y' + 2xy = 2xe^{-x^2}$ b) $(1+x^2)y' - 2xy = (1+x^2)^2$
   c) $\cos xy' - \sin xy = xe^x$ d) $y' - y = \cosh x$
   e) $xy' + 2y = x^3,\ y(1) = 1$ f) $y' + y\cos x = \sin x \cos x,\ y(0) = 1$
   g) $y' + y\tan x = \sin^3 x,\ y(0) = 1$ h) $y' + |x - |x||y = x,\ y(0) = 0$
   i) $y' + 2y = x^3 - x,\ y(0) = 1$ j) $y' - y = e^x - \sin x,\ y(0) = 0$

2. a) Differentiaaliyhtälöllä $y' + P(x)y = (x+1)^2 e^x$ on ratkaisu $y = (x^2-1)e^x$. Määritä yleinen ratkaisu.
   b) Olkoon funktiot f ja g derivoituvia ja f' ja g' jatkuvia välillä (a,b), ja olkoon $f(x) \neq 0\ \forall x \in (a,b)$. Minkä lineaarisen differentiaaliyhtälön yleinen ratkaisu välillä (a,b) on $y(x) = Cf(x) + g(x),\ C \in \mathbb{R}$?

3. Määritä $\mathbb{R}$:sssä jatkuva funktio $y(x)$, joka toteuttaa yhtälön

$$2\int_0^x ty(t)\,dt = x^2 + y(x), \quad x \in \mathbb{R}.$$

4. Näytä, että *Bernoullin* DY $y' = A(x)y + B(x)y^k$ $(k \in \mathbb{R},\ k \neq 0,\ k \neq 1)$ palautuu lineaariseksi sijoituksella $u = y^{1-k}$. Ratkaise tällä periaatteella

   a) $xy' + y = x^3y^2$ b) $y' + y = y^2(\cos x - \sin x)$
   c) $3y' + y = (1-2x)y^4$ d) $y' + 2y/(1-x) = 4(x^2-x)\sqrt{y}$

5. Käyrän $y = y(x)$ pisteeseen P asetetaan tangentti, joka leikkaa y-akselin pisteessä Q. Määritä kaikki käyrät, joilla on ominaisuus: kolmion OPQ (O = origo) pinta-ala $= a^2$ = vakio.

6. Funktio $R(x)$ on jatkuva $\mathbb{R}$:ssä ja $-2x^2 \leq R(x) \leq x^2\ \forall x \in \mathbb{R}$. Funktio $y(x)$ on alkuarvotehtävän $y' + 3x^2y = R(x),\ y(0) = 0$ ratkaisu. Mitä arvoja $y(-1)$ voi saada?

7. Olkoon f ja y_p koko $\mathbb{R}$:ssä määriteltyjä, jatkuvia ja L-jaksoisia funktioita, ja lisäksi olkoon y_p differentiaaliyhtälön $y' + ay = f(x)$ $(a \in \mathbb{R})$ ratkaisu välillä $(0,L)$.
   a) Näytä, että $y_p' + ay_p = f(x)\ \forall x \in \mathbb{R}$.
   b) Olkoon F ja Y f:n ja y_p:n keskiarvot välillä $[0,L]$. Näytä, että $F = aY$. — Johtopäätös, jos $a = 0$?

8. Olkoon
$$f(x) = \begin{cases} 0, & \text{kun } x = k \in \mathbb{Z}, \\ 2x - 2k - 1, & \text{kun } x \in (k, k+1),\ k \in \mathbb{Z}. \end{cases}$$
Määritä funktio y_p, joka on (1) jaksollinen, (2) koko $\mathbb{R}$:ssä määritelty ja jatkuva ja (3) on differentiaaliyhtälön $y' + y = f(x)$ ratkaisu väleillä $(k, k+1)$, $k \in \mathbb{Z}$. Hahmottele y_p graafisesti välillä $[0, 2]$. Toteutuuko differentiaaliyhtälö pisteessä $x = 1$?

9. (*) a) Näytä, että alkuarvotehtävän $y' = x + |y|$, $y(-2) = 1$ ratkaisulla on minimi $y(\ln 2 - 1) = \ln 2 - 1$.
b) Määritä differentiaaliyhtälön $y' = |y - x|$ yleinen ratkaisu. Laske myös $y_1(1)$ ja $y_2(1)$ yksittäisratkaisuille, jotka toteuttavat ehdot $y_1(-1) = -1/2$ ja $y_2(-1) = 1/2$.

10. (*) *Riccatin* DY on muotoa $y' = A(x) + B(x)y + C(x)y^2$. Jos tälle tunnetaan yksittäisratkaisu y_0, niin yhtälö palautuu Bernoullin DY:ksi (ks. Tehtävä 4) sijoituksella $y = y_0 + u$. Ratkaise tällä periaatteella seuraavat differentiaaliyhtälöt annettua lisätietoa käyttäen.
a) $y' = 1 + x + x^2 - (2x + 1)y + y^2$, $y_0(x) =$ polynomi astetta 1
b) $y' = y^2 - x^2 y - (x - 1)^2$, $y_0(x) =$ polynomi astetta 2
c) $x^2 y' + (xy - 2)^2 = 0$, $y_0(x) = a/x$ $(a \in \mathbb{R})$

11. (*) Funktio $y(x)$ on derivoituva välillä $(0, \infty)$, oikealta jatkuva pisteessä $x = 0$ ja $y(0^+) = 1$. Lisäksi tiedetään, että $y'(x) + y(x) \geq f(x)$ kun $x > 0$, missä f on välillä $[0, \infty)$ jatkuva funktio. Näytä, että jokaisella $x > 0$ pätee
$$y(x) \geq e^{-x} + \int_0^x e^{t-x} f(t)\, dt.$$

12. (*) Olkoon $\tau > 0$. Ratkaise differentiaaliyhtälö
$$\tau y' + y = |2k - t|, \quad t \in (0, \infty) \cap [2k - 1, 2k + 1], \quad k \in \mathbb{N} \cup \{0\}$$
muodossa $y(t) = Ce^{-t/\tau} + y_p(t)$, missä y_p on jaksollinen, jaksona $T = 2$. Hahmottele graafisesti alkuehdon $y(0) = 2$ toteuttava ratkaisu τ:n arvoilla 10, 1 ja 0.1.

IX.5 Lineaariset, vakiokertoimiset DY:t. Eulerin differentiaaliyhtälö

Edellisessä luvussa ratkaistiin jo ensimmäisen kertaluvun lineaarinen ja vakiokertoiminen differentiaaliyhtälö $y' + ay = f(x)$. Sovelluksissa yleinen on myös vastaava toisen kertaluvun differentiaaliyhtälö, jonka yleinen muoto on

$$y'' + ay' + by = f(x), \quad a, b \in \mathbb{R}. \tag{1}$$

Tämän ratkaisemiseksi tarkastellaan ensin vastaavaa homogeenista yhtälöä

$$y'' + ay' + by = 0. \tag{2}$$

Kun tämän ratkaisua etsitään yritteellä

$$y(x) = e^{rx},$$

niin sijoittamalla yhtälöön saadaan vakion r määrämiseksi *karakteristinen yhtälö*

$$\boxed{r^2 + ar + b = 0 \quad \text{(karakteristinen yhtälö).}}$$

Karakteristisen yhtälön juuret ovat

$$r = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}\,,$$

jolloin on kolme mahdollisuutta:

1. $a^2 - 4b > 0$. Tässä tapauksessa karakteristisella yhtälöllä on kaksi erisuurta reaalijuurta r_1 ja r_2 ja homogeeniyhtälöllä (2) siis ratkaisut

$$y_1(x) = e^{r_1 x}, \quad y_2(x) = e^{r_2 x}.$$

Ratkaisu on myös $y(x) = C_1 y_1(x) + C_2 y_2(x)$ $(C_1, C_2 \in \mathbb{R})$, sillä jos yhtälö (2) kirjoitetaan $\mathrm{L}y = 0$, niin L on lineaarinen operaattori, jolloin on $\mathrm{L}(C_1 y_1 + C_2 y_2) = C_1 \mathrm{L}y_1 + C_2 \mathrm{L}y_2 = 0$. Siis jokainen $y(x)$ muotoa

$$y(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}, \quad C_1, C_2 \in \mathbb{R}$$

on homogeeniyhtälön (2) ratkaisu. Näytetään nyt, että kyseessä on yhtälön (2) yleinen ratkaisu eli mikä tahansa ratkaisu on tätä muotoa. Tätä silmällä pitäen kirjoitetaan ensin karakteristinen yhtälö muotoon

$$r^2 + ar + b \;=\; (r - r_1)(r - r_2) \;=\; r^2 - (r_1 + r_2)r + r_1 r_2 \;=\; 0,$$

missä siis $a = -r_1 - r_2$ ja $b = r_1 r_2$. Olkoon $y(x)$ homogeeniyhtälön (2) ratkaisu. Tällöin derivoimalla funktio $u(x) = y'(x) - r_2 y(x)$ todetaan, että

$$u' - r_1 u = y'' - (r_1 + r_2)y' + r_1 r_2 y = y'' + ay' + by = 0.$$

Siis $u(x)$ ja $y(x)$ ovat ratkaistavissa differentiaaliyhtälöryhmästä

$$\begin{cases} u' - r_1 u = 0, \\ y' - r_2 y = u(x). \end{cases}$$

Kun tässä ensimmäisen yhtälön yleinen ratkaisu $u(x) = C_1 e^{r_1 x}$ ($C_1 \in \mathbb{R}$) sijoitetaan jälkimmäiseen yhtälöön, niin todetaan tämän yksittäisratkaisuksi $y_0(x) = (r_1 - r_2)^{-1} C_1 e^{r_1 x}$ (vrt. edellinen luku) ja yleiseksi ratkaisuksi siis

$$y(x) = \frac{C_1}{r_1 - r_2} e^{r_1 x} + C_2 e^{r_2 x}.$$

Tässä voidaan C_1:n tilalle kirjoittaa yhtä hyvin $(r_1 - r_2)C_1$, joten todetaan, että jokainen yhtälön (2) ratkaisu on muotoa $y(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}$, $C_1, C_2 \in \mathbb{R}$. Jokainen tällainen funktio oli myös ratkaisu, joten kyseessä on yleinen ratkaisu.

2. $a^2 - 4b = 0$. Tässä tapauksessa karakteristisella yhtälöllä on kaksoisjuuri $r = -a/2$, joten em. yhtälöryhmässä on $r_1 = r_2 = r$. Kun ensimmäisen yhtälön yleinen ratkaisu $u(x) = C_1 e^{rx}$ ($C_1 \in \mathbb{R}$) sijoitetaan jälkimmäiseen, niin tämän yksittäisratkaisu on $y_0(x) = C_1 x e^{rx}$ (vrt. edellinen luku) ja yleinen ratkaisu siis

$$y(x) = (C_1 x + C_2)e^{rx}, \quad r = -\frac{a}{2}.$$

Tämä on myös yhtälön (2) yleinen ratkaisu.

3. $a^2 - 4b < 0$. Tässä tapauksessa karakteristisen yhtälön juuret muodostavat konjugaattiparin

$$r_{1,2} = \alpha \pm i\beta, \quad \alpha = -\frac{a}{2},\ \beta = \sqrt{b - \frac{a^2}{4}}.$$

Sijoituksella $y(x) = e^{\alpha x} u(x)$ yhtälö (2) muuntuu muotoon

$$u'' + \beta^2 u = 0.$$

Tämä on ratkaistavissa 1. kertalukuun palutuvana DY:nä (ks. Luku IX.3), mutta suoremminkin voi päätellä, että ratkaisuja ovat $u_1(x) = \cos\beta x$ ja $u_2(x) = \sin\beta x$ ja yleinen ratkaisu siis ilmeisesti $u(x) = C_1 \cos\beta x + C_2 \sin\beta x$, $C_1, C_2 \in \mathbb{R}$. Näin saadaan yhtälön (2) yleiseksi ratkaisuksi

$$y(x) = e^{\alpha x}(C_1 \cos\beta x + C_2 \sin\beta x).$$

Esimerkki 1 Ratkaise reuna-arvotehtävä

$$\begin{cases} y'' - y' + 2y = 0, & x \in (0,1), \\ y(0) = 1, \ y(1) = 0. \end{cases}$$

Ratkaisu Karakteristisen yhtälön $r^2 - r - 2$ juuret ovat $r_1 = 2$, $r_2 = -1$, joten differentiaaliyhtälön yleinen ratkaisu on

$$y(x) = C_1 e^{2x} + C_2 e^{-x}.$$

Reunaehdot toteutuvat, kun

$$\begin{cases} C_1 + \quad C_2 = 1 \\ e^2 C_1 + e^{-1} C_2 = 0 \end{cases} \Leftrightarrow \begin{cases} C_1 = -1/(e^3 - 1), \\ C_2 = \ e^3/(e^3 - 1). \end{cases}$$

Siis reuna-arvotehtävän ratkaisu on

$$y(x) = \frac{1}{e^3 - 1}(-e^{2x} + e^{3-x}). \quad \square$$

Täydellinen vakiokertoiminen yhtälö

Koska differentiaaliyhtälö (1) on lineaarinen, niin sen ratkaisulle pätee sama yleisperiaate kuin ensimmäisen kertaluvun lineaariselle DY:lle: Täydellisen yhtälön yleinen ratkaisu = homogeenisen yhtälön yleinen ratkaisu + täydellisen yhtälön yksittäisratkaisu $y_0(x)$. Yksittäisratkaisun määräämiseksi tarkastellaan tässä yhteydessä vain 'sivistyneen arvauksen' menetelmiä, jotka toimivat silloin, kun $f(x)$ yhtälössä (1) on riittävän yksinkertaista muotoa. (Yleisempi menetelmä esitetään seuraavassa luvussa; ks. myös Harj.teht. 9.)

Esimerkki 2 Ratkaise $y'' + y = x^2 + 2e^{2x}$.

Ratkaisu Yritetään yksittäisratkaisua muodossa

$$y(x) = Ax^2 + Bx + C + De^{2x}.$$

Sijoittamalla yhtälöön todetaan tämä ratkaisuksi kun $A = 1$, $B = 0$, $C = -2$ ja $D = 2/5$. Homogeenisen yhtälön yleinen ratkaisu on

$$y(x) = C_1 \cos x + C_2 \sin x,$$

joten täydellisen yhtälön yleinen ratkaisu on

$$y(x) = C_1 \cos x + C_2 \sin x + x^2 - 2 + \frac{2}{5}e^{2x}. \quad \square$$

Jos yleisemmin yhtälön (1) oikea puoli $R(x)$ on muotoa

$$\text{a)}\ \ R(x) = x^n e^{\alpha x}, \quad \text{tai} \quad \text{b)}\ \ R(x) = x^n(A\cos\omega x + B\sin\omega x),$$

missä $n \in \{0, 1, 2, \ldots\}$, $\alpha \in \mathbb{R}$, ja $\omega \in \mathbb{R}$, $\omega \neq 0$, niin yhtälön (1) yksittäisratkaisu on löydettävissä vastaavasti muodossa

a) $y(x) = p(x)e^{\alpha x}$,

b) $y(x) = p(x)\cos\omega x + q(x)\sin\omega x$,

missä p ja q ovat polynomeja. Pääsääntöisesti p ja q ovat astetta n. Poikkeuksen muodostavat ne tapaukset, joissa pääsäännön mukainen yrite sattuu olemaan homogeenisen yhtälön ratkaisu. Tapauksessa a) tämä on mahdollista kun $n = 0$ tai $n = 1$, tapauksessa b) kun $n = 0$. Tällöin polynomin astetta on nostettava yhdellä, tapauksessa a) mahdollisesti kahdella, jotta yksittäisratkaisu löytyisi.

Esimerkki 3 Ratkaise alkuarvotehtävä

$$\begin{cases} y'' + 2y' + y = e^{-x}, & x \in \mathbb{R}, \\ y(1) = y'(1) = 0. \end{cases}$$

<u>Ratkaisu</u> Karaktristisella yhtälöllä on kaksoisjuuri $r = -1$, joten homogeenisen yhtälön yleinen ratkaisu on $y(x) = (C_1 + C_2 x)e^{-x}$. Koska sekä $f(x) = e^{-x}$ että xe^{-x} ovat homogeenisen yhtälön ratkaisuja, niin yksittäisratkaisua on etsittävä muodossa

$$y(x) = (Ax^2 + Bx + C)\, e^{-x}.$$

Tämä osoittautuu ratkaisuksi, kun valitaan $A = 1/2$ ja $B, C \in \mathbb{R}$, eli saatiin suoraan yleinen ratkaisu

$$y(x) = \Big(\frac{1}{2}x^2 + Bx + C\Big)e^{-x}, \quad B, C \in \mathbb{R}.$$

(Samaan tulokseen olisi tultu, jos yritteessä olisi valittu 'viisaammin' $B = C = 0$ ja lisätty homogeenisen yhtälön yleinen ratkaisu vasta jälkikäteen.) Alkuehdot toteutuvat, kun $A = -1$ ja $B = 1/2$, joten alkuarvotehtävän ratkaisu on

$$y(x) = \frac{1}{2}(x-1)^2 e^{-x}. \quad \square$$

Kompleksiarvoiset ratkaisut

Vakiokertoimisia differentiaaliyhtälöitä ratkaistaessa on usein kätevää suorittaa laskut kompleksiarvoisia funktioita käyttäen silloinkin, kun pyritään reaaliseen lopputulokseen. Menetelmä on kätevä erityisesti silloin, kun karakteristisen yhtälön juuret ovat kompleksilukuja, tai kun täydellisen yhtälön (1) oikealla puolella esiintyy trigonometrisia funktioita. Kompleksifunktioilla laskettaessa hyväksytään homogeenisen yhtälön ratkaisuyritteessä

$$y(x) = e^{rx}$$

suoraan myös kompleksiset r:n arvot. Kyseessä on tällöin reaalimuuttujan kompleksiarvoinen funktio, jonka derivaatta määritellään normaaliin tapaan eli erotusosamäärän raja-arvona. Kun derivaatan määritelmässä merkitään po. funktion tapauksessa

$$z = rx, \quad \Delta z = r\Delta x,$$

niin nähdään, että

$$\frac{y(x+\Delta x) - y(x)}{\Delta x} = r\,\frac{e^{z+\Delta z} - e^{z}}{\Delta z}\,.$$

Tässä $\Delta x \to 0 \;\Rightarrow\; \Delta z \to 0$, joten kompleksifunktion e^z derivoimissäännön perusteella (ks. Luku VI.3) voidaan todeta, että pätee odotetusti

$$\boxed{\frac{d}{dx}\,e^{rx} = re^{rx}, \quad r \in \mathbb{C}.}$$

Esimerkki 4 Ratkaise differentiaaliyhtälö (2) kompleksifunktioiden avulla, kun $a^2 - 4b < 0$.

<u>Ratkaisu</u> Karakteristisen yhtälön juuret ovat $r_{1,2} = \alpha \pm i\beta$, joten yleinen ratkaisu saadaan aiempaan tapaan ratkaisemalla yhtälöryhmä

$$\begin{cases} u' - r_1 u = 0, \\ y' - r_2 y = u(x). \end{cases}$$

Kuten aiemmin (vrt. tapaus $r_1, r_2 \in \mathbb{R}$ edellä) saadaan yleiseksi ratkaisuksi

$$\begin{aligned} y(x) &= C_1 e^{r_1 x} + C_2 e^{r_2 x} \\ &= e^{\alpha x}(C_1 e^{i\beta x} + C_2 e^{-i\beta x}), \quad C_1, C_2 \in \mathbb{C}. \end{aligned}$$

— Huomattakoon, että tässä ei kertoimia C_1, C_2 ole syytä rajoittaa reaalisiksi. Kun huomioidaan Eulerin kaava

$$e^{\pm i\beta x} = \cos\beta x \pm i\sin\beta x,$$

niin nähdään, että ratkaisu on esitettävissä yhtäpitävästi muodossa

$$y(x) = e^{\alpha x}(D_1\cos\beta x + D_2\sin\beta x),$$

$$\text{missä} \quad \begin{cases} D_1 = C_1 + C_2, \\ D_2 = i(C_1 - C_2). \end{cases}$$

Tästä nähdään, että valitsemalla $C_2 = \overline{C}_1$ ($= C_1$:n konjugaatti) saadaan yleinen reaalinen ratkaisu aiemmin esitetyssä muodossa. $\square$

Esimerkin menettelyllä voidaan hakea homogeenisen yhtälön (2) yleinen ratkaisu myös kompleksikertoimisessa tapauksessa ($a, b \in \mathbb{C}$), mikäli sellainen tilanne eteen tulisi. Tutkitaan sen sijaan sovelluksissa hyvin yleistä laskentatapaa, jossa reaalikertoimiselle täydelliselle yhtälölle

$$\mathrm{L}y = y'' + ay' + by = \begin{cases} \sin\omega x \\ \cos\omega x \end{cases}$$

etsitään yksittäisratkaisu käyttäen hyväksi kompleksifunktioita. Menetelmä perustuu seuraavaan yksinkertaiseen havaintoon: Jos $f(x) = f_1(x) + if_2(x)$, missä f_1 ja f_2 ovat reaaliarvoisia, niin y on yhtälön $\mathrm{L}y = f$ yksittäisratkaisu täsmälleen kun $y = y_1 + iy_2$, missä $\mathrm{L}y_1 = f_1$ ja $\mathrm{L}y_2 = f_2$. (Tämä on helposti todettavissa L:n lineaarisuuden ja kertoimien a, b reaalisuuden perusteella — kompleksikertoimisessa tapauksessa sääntö ei päde.)

Esimerkki 5 Etsi yksittäisratkaisu differentiaaliyhtälölle $y'' - y' + y = \sin\omega x$ ($\omega \neq 0$) käyttäen kompleksifunktioita.

<u>Ratkaisu</u> Koska $\sin\omega x = \mathrm{Im}\,(e^{i\omega x})$, niin probleema voidaan ratkaista etsimällä kompleksiarvoinen yksittäisratkaisu differentiaaliyhtälölle

$$y'' - y' + y = e^{i\omega x},$$

ja ottamalla ratkaisusta imaginaariosa. Kompleksinen ratkaisu löytyy helposti sijottamalla yrite $y(x) = Ae^{i\omega x}$ ($A \in \mathbb{C}$) yhtälöön:

$$\begin{aligned} &(-\omega^2 - i\omega + 1)Ae^{i\omega x} = e^{i\omega x} \quad \forall x \\ &\Rightarrow \quad A = \frac{1}{1-\omega^2 - i\omega} = \frac{1-\omega^2+i\omega}{\omega^4-\omega^2+1} \\ &\Rightarrow \quad y(x) = \frac{1-\omega^2+i\omega}{\omega^4-\omega^2+1}\,e^{i\omega x} = \frac{1-\omega^2+i\omega}{\omega^4-\omega^2+1}\,(\cos\omega x + i\sin\omega x). \end{aligned}$$

Kysytty ratkaisu on tämän imaginaariosa, eli

$$y(x) = \frac{1}{\omega^4 - \omega^2 + 1}\,[\,\omega\cos\omega x + (1-\omega^2)\sin\omega x\,]. \quad \square$$

Esimerkissä olisi luonnollisesti tultu toimeen myös reaalisella yritteellä $y(x) = A\cos\omega x + B\sin\omega x$ $(A, B \in \mathbb{R})$, mutta laskusta olisi tullut ikävämpi.

Eulerin differentiaaliyhtälö

Differentiaaliyhtälöä muotoa

$$x^2y'' + axy' + by = f(x)$$

sanotaan (toisen kertaluvun) *Eulerin* DY-tyypiksi. Perusmuodossa

$$y'' + ax^{-1}y' + bx^{-2}y = x^{-2}f(x)$$

on $x = 0$ kertoimien epäjatkuvuuspiste, joten Eulerin differentiaaliyhtälöä on tarkasteltava erikseen väleillä $(-\infty, 0)$ ja $(0, \infty)$ — sovelluksissa yleensä välillä $(0, \infty)$. Kummallakin välillä yhtälö palautuu vakiokertoimiseksi sijoituksella

$$|x| = e^t \;\Leftrightarrow\; t = \ln|x|,$$

sillä kun esim. välillä $(0, \infty)$ kirjoitetaan

$$y(x) = y(e^t) = u(t) = u(\ln x),$$

niin saadaan

$$\begin{aligned} y'(x) &= \frac{1}{x}u'(\ln x) = \frac{1}{x}u'(t), \\ y''(x) &= \frac{1}{x^2}u''(\ln x) - \frac{1}{x^2}u'(\ln x) = \frac{1}{x^2}[u''(t) - u'(t)], \end{aligned}$$

joten yhtälö saadaan vakiokertoimiseen muotoon

$$u''(t) + (a-1)u'(t) + bu(t) = f(e^t).$$

Välillä $(-\infty, 0)$ tulee oikealle puolelle $f(-e^t)$, muuten tulos on sama.

Homogeenista Eulerin yhtälöä ei käytännössä tarvitse muuntaa vakiokertoimiseksi, sillä ratkaisua voi etsiä suoraan yritteellä

$$y(x) = x^r.$$

Tällöin r:n määräämiseksi saadaan toisen asteen karakteristinen yhtälö. Riippuen siitä, millaisia juuret ovat, yleiseksi ratkaisuksi tulee jokin seuraavista:

a) $y(x) = C_1 x^{r_1} + C_2 x^{r_2}$,

b) $y(x) = x^r (C_1 + C_2 \ln|x|)$,

c) $y(x) = x^\alpha [\, C_1 \cos(\beta \ln|x|) + C_2 \sin(\beta \ln|x|)\,]$.

Tapauksessa a) juuret ovat reaaliset ja erisuuruiset, tapauksessa b) on reaalinen kaksoisjuuri, ja tapauksessa c) juurina on konjugaattipari $\alpha \pm i\beta$. Väitetyt yleisen ratkaisun muodot voi päätellä em. muunnoksen avulla.

Täydellinen Eulerin yhtälö ratkaistaan suotuisissa tapauksissa 'sivistyneellä arvauksella'. Arvauksen muodon voi johtaa ajatellen muunnosta vakiokertoimiseen tilanteeseen.

Esimerkki 6 Ratkaise alkuarvotehtävä

$$\begin{cases} x^2 y'' + 3xy' + y = 1/x, \;\; x > 0, \\ y(1) = 1, \;\; y'(1) = 0. \end{cases}$$

<u>Ratkaisu</u> Homogeenisessa yhtälössä sijoitus $y(x) = x^r$ johtaa karakteristiseen yhtälöön

$$r(r-1) + 3r + 1 = r^2 + 2r + 1 = 0.$$

Tällä on kaksoisjuuri $r = -1$, joten homogeenisen yhtälön yleinen ratkaisu on

$$y(x) = \frac{1}{x}(C_1 + C_2 \ln x).$$

Koska oikea puoli $f(x) = x^{-1}$ on homogeenisen yhtälön ratkaisu, samoin kuin $x^{-1} \ln|x|$, niin täydellisen yhtälön yksittäisratkaisu on etsittävä muodossa

$$y_0(x) = Ax^{-1}(\ln x)^2.$$

Kun sijoitetaan yhtälöön tämä sekä derivaatat

$$\begin{aligned} y_0'(x) &= Ax^{-2}[\,-(\ln x)^2 + 2\ln x\,], \\ y_0''(x) &= Ax^{-3}[\,2(\ln x)^2 - 6\ln x + 2\,], \end{aligned}$$

niin saadaan ratkaisu, kun $A = 1/2$. Siis yleinen ratkaisu on

$$y(x) = \frac{1}{x}\left(C_1 + C_2 \ln x + \frac{1}{2}(\ln x)^2\right).$$

Alkuehdot toteutuvat, kun valitaan $C_1 = C_2 = 1$. $\square$

Korkeamman kertaluvun vakiokertoimiset DY:t

Yleinen lineaarinen, vakiokertoiminen, homogeeninen, n:nnen kertaluvun differentiaaliyhtälö

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0$$

voidaan ratkaista yritteellä $y(x) = e^{rx}$, kuten tapauksissa $n = 1, 2$. Saadaan karakteristinen yhtälö

$$r^n + a_{n-1}r^{n-1} + \cdots + a_0 = 0,$$

jonka kutakin juurta vastaa juuren kertaluvun mukainen määrä lineaarisesti riippumattomia ratkaisuja. Jos kyseessä on m-kertainen reaalijuuri, niin nämä ratkaisut ovat

$$x^k e^{rx}, \quad k = 0 \ldots m-1.$$

Jos kyseessä on m-kertainen konjugaattipari $\alpha \pm i\beta$, niin vastaavat ratkaisut ovat

$$x^k e^{\alpha x} \cos \beta x, \quad x^k e^{\alpha x} \sin \beta x, \quad k = 0 \ldots m-1.$$

Homogeenisen yhtälön yleinen ratkaisu saadaan kaikkien näiden ratkaisujen (yhteensä n kpl) lineaarisena yhdistelynä.

Esimerkki 7 Ratkaise $y''' - y = x^2, \ x \in \mathbb{R}$.

<u>Ratkaisu</u> Yksittäisratkaisu on $y(x) = -x^2$. Homogeenista yhtälöä ratkaistaessa tulee karakteristiseksi yhtälöksi $r^3 - 1 = 0$. Juuret ovat

$$r_1 = 1, \quad r_{2,3} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2},$$

joten yleinen ratkaisu on

$$y(x) = C_1 e^x + e^{-x/2}\left(C_2 \cos\frac{\sqrt{3}x}{2} + C_3 \sin\frac{\sqrt{3}x}{2}\right) - x^2. \quad \square$$

Kuten esimerkissä, voidaan täydellisen yhtälön ratkaisu usein löytää kokeilemalla. Yleisempiin menetelmiin ei juuri käytännön tarvetta olekaan.

Todettakoon lopuksi, että yleinen n:nnen kertaluvun homogeeninen Eulerin DY

$$x^n y^{(n)} + a_{n-1}x^{n-1}y^{(n-1)} + \ldots + a_0 y = 0$$

ratkaistaan yritteellä $y(x) = x^r$ samaan tapaan kuin edellä.

HARJOITUSTEHTÄVIÄ

1. Ratkaise (yleinen ratkaisu kohdissa a)–n), yleinen ratkaisu kaikilla lukupareilla $(a, b) \in \mathbb{R}^2$ kohdissa o)–t), alku- tai reuna-arvotehtävän ratkaisu kohdissa u)–ö)):

   a) $y'' + y' - 30y = 0$ b) $y'' + 4y' + y = 0$
   c) $y'' + 6y' + 10y = 0$ d) $y'' + 4y' + 6y = 0$
   e) $y'' + 4y' + 5y = 3x - 2$ f) $y'' - y' - 2y = x^2$
   g) $y'' - 7y' + 6y = \sin x$ h) $y'' + 4y = \sin 3x$
   i) $y'' + 2y' + 5y = e^{-x}$ j) $y'' + y' - 2y = x + e^x$
   k) $y'' - 4y = xe^{2x}$ l) $y'' + 4y' + 4y = (x-1)^2 e^{-2x}$
   m) $y'' - 6y' + 9y = (xe^x)^3$ n) $y'' + 4y' + 4y = e^{-2x} + \sin x$
   o) $y'' + ay' = e^{bx}$ p) $y'' - a^2 y = e^{bx}$
   q) $y'' + a^2 y = e^{bx}$ r) $y'' + a^2 y = \sin bx$
   s) $y'' + a^2 y = x \sin bx$ t) $y'' + 2y' + (1 + a^2)y = e^{bx}$
   u) $y'' + 2y' + y = 0,\ y(0) = 0,\ y'(0) = 1$
   v) $y'' + 2y' + 2y = 0,\ y(0) = 1,\ y'(0) = 0$
   x) $y'' - 4y' - 5y = 0,\ y(0) = 1,\ \lim_{x\to\infty} y(x) = 0$
   y) $y'' + 3y' + 2y = 0,\ y(0) = y(1) = 1$
   x) $y'' - 4y' + 5y = \sin x,\ y(0) = 0,\ y'(0) = -1$
   å) $y'' - 3y' + x^2 - 1 = 0,\ y(0) = 1,\ y'(0) = 0$
   ä) $y'' - 5y' + 6y = e^x,\ y(0) = y(1) = 0$
   ö) $4y'' + 8y' + 5y = x^2,\ y(0) = y(\pi) = 1$

2. Etsi seuraaville differentiaaliyhtälöille ensin kompleksiarvoinen yksittäisratkaisu, kun yhtälön oikea puoli on $f(x) = e^{i\omega x}$ ($\omega \in \mathbb{R},\ \omega \neq 0$), ja määritä sen avulla reaalinen yksittäisratkaisu.

   a) $y'' - 2y' + 2y = \sin \omega x$ b) $y'' - 3y' + 2y = \sin \omega x$
   c) $y'' + 4y' + 4y = \cos \omega x$ d) $y'' + y' + y = \sin \omega x - 2\cos \omega x$

3. (Resonanssi) Kappaleeseen, jonka massa $= m$ ja paikka $= y(t)$ hetkellä t, kohdistuu jousivoima ja lisäksi ulkoinen kuorma $f(t)$, jolloin $y(t)$ toteuttaa liikeyhälön $my'' = -ky + f(t)$. Määritä $y(t),\ t \geq 0$ alkuehdoilla $y(0) = y'(0) = 0$, kun $f(t) = F \sin \omega t$ ($F =$ vakio). Tarkastele erikseen tapaukset a) $\omega \neq \omega_0$ ja b) $\omega = \omega_0$, missä $\omega_0 = \sqrt{k/m}$ (nk. resonanssitaajuus).

4. (Töyssy) Auton etupyörä osuu töyssyyn, jolloin pyörän liikeyhtälö pystysuunnassa on $my'' = F_1 + F_2$, missä $m =$ pyörän massa, $F_1 = -ky$ on jousivoima ja $F_2 = -cy'$ on iskunvaimentimen vastusvoima. Hahmottele ratkaisu $y(t)$ (aikayksikkö s) alkuehdoilla $y(0) = 0$ m, $y'(0) = 60$ m/s, kun $m = 15$ kg, $k = 6 \cdot 10^5$ kg/s^2 ja iskunvaimennin on a) uusi: $c = 7500$ kg/s, b) pian vaihdettava: $c = 6000$ kg/s, c) heti vaihdettava: $c = 4500$ kg/s.

5. Ratkaise (kohdassa i) kaikilla $(a, b) \in \mathbb{R}^2$):
   a) $x^2y'' - xy' + y = 0$ b) $x^2y'' - 6xy' + 7y = 0$
   c) $x^2y'' - 2xy' + 2y = x$ d) $x^2y'' + xy' - y = (x+1)^2/x^2$
   e) $x^2y'' + xy' + y = x^3 - 2x$ f) $x^2y'' - 3xy' + 4y = x^2 \ln|x|$
   g) $(x^2 + 2x + 1)y'' + (x+1)y' + y = x^2$
   h) $(3x+2)^2y'' + (9x+6)y' - 36y = 81x + 18$
   i) $x^2y'' + (2a+1)xy' + by = 0$

6. a) Määritä kaikki välillä $(0, \infty)$ derivoituvat funktiot, jotka toteuttavat yhtälön $x^2y'(x) = 2\int_0^x y(t)\,dt,\ x > 0$.
   b) Ratkaise differentiaaliyhtälö $x^2y''' + 2(x^2 - x)y'' + (x^2 - 2x + 2)y' = x^3$ tekemällä sijoitus $u = y'/x$.

7. Ratkaise (yleinen ratkaisu tai alkuarvotehtävän ratkaisu):
   a) $y^{(5)} + 2y''' + y' = 0$ b) $y^{(7)} + 3y^{(6)} + 3y^{(5)} + y^{(4)} = 0$
   c) $y''' + 3y'' - 2y = \sin x$ d) $y^{(4)} + 3y''' + 3y'' + y' = 2e^{-2x} - 2x$
   e) $y^{(4)} + a^4y = x^2,\ a \geq 0$
   f) $y''' - (a+2)y'' + (2a+1)y' - ay = x + 1,\ a \in \mathbb{R}$
   g) $x^3y''' + 2ax^2y'' - 4a(xy' - y) = 0,\ a \in \mathbb{R}$
   h) $y''' - 5y'' + 17y' - 13y = 0,\ y(0) = y'(0) = 0,\ y''(0) = 24$
   i) $y''' - y'' - y' + y = e^x,\ y(0) = y'(0) = y''(0) = 0$
   j) $y''' - ay' = 0,\ y(0) = 0,\ y'(0) = 1,\ y''(0) = 2,\ a \in \mathbb{R}$
   k) $x^3y''' + xy' - y = \sqrt{x},\ y(1) = 1,\ y'(1) = y''(1) = 0$

8. (*) (Pystyammunta) Luoti ammutaan suoraan ylöspäin lähtönopeudella 800 m/s. Lennon aikana luotiin vaikuttaa nopeuteen verrannollinen vastusvoima, jolloin nousukorkeus $y(t)$ toteuttaa differentiaaliyhtälön

$$y'' + ay' + g = 0,$$

   missä $g = 9.8$ m/s^2, $a = 0.10$ s^{-1} luodin nousuvaiheessa ja $a = 0.20$ s^{-1} putoamisvaiheessa. Määritä luodin lennon lakikorkeus, lentoon kuluva aika ja luodin paluunopeus sen pudotessa maahan.

9. (*) Vakiokertoimisessa differentiaaliyhtälössä $y'' + ay' + by = f(x)$ olkoon karakteristisen yhtälön juuret $r_1, r_2 \in \mathbb{C}$. Kirjoittamalla differentiaaliyhtälö systeemiksi funktioille y ja $u = y' - r_2y$ näytä, että yhtälön yksittäisratkaisu alkuehdoilla $y(x_0) = y'(x_0) = 0$ on

$$y_0(x) = \int_{x_0}^{x} e^{r_1(x-t)} \left[\int_{x_0}^{t} e^{r_2(t-s)} f(s)\,ds \right] dt.$$

   Tarkista kaavan toimivuus, kun $a = -2$, $b = 1$, $x_0 = 0$ ja $f(x) = x^2 - 4x + 2$.

IX.6 *Yleinen 2. kertaluvun lineaarinen DY

Yleinen toisen kertaluvun lineaarinen differentiaaliyhtälö on muotoa

$$y'' + P(x)y' + Q(x)y = R(x). \tag{1}$$

Perusoletus on, että sekä *kerroinfunktiot* P ja Q että yhtälön *oikea puoli* R ovat jatkuvia tarkasteltavalla välillä (a, b). Ratkaisuna pidetään tällöin jokaista ko. välillä kahdesti derivoituvaa funktiota $y(x)$, joka toteuttaa yhtälön. — Yhtälöstä on tällöin luettavissa, että myös y'' on jatkuva välillä (a, b). Jatkossa tarkastellaan aluksi yhtälön (1) homogeenista erikoistapausta, jossa $R(x) = 0$.

Homogeeninen yhtälö: Ratkaisuavaruus

Kun homogeeninen differentiaaliyhtälö kirjoitetaan

$$\mathrm{L}y = y'' + P(x)y' + Q(x)y = 0, \tag{2}$$

niin operaattori L on jälleen lineaarinen:

$$\mathrm{L}(c_1y_1 + c_2y_2) = c_1\mathrm{L}y_1 + c_2\mathrm{L}y_2, \quad c_1, c_2 \in \mathbb{R}.$$

Siis jos y_1 ja y_2 ovat yhtälön (2) ratkaisuja, niin ratkaisu on myös mikä tahansa näiden lineaarinen yhdistely

$$y(x) = C_1y_1(x) + C_2y_2(x), \quad C_1, C_2 \in \mathbb{R}.$$

Ratkeavuusteorian keskeisin tulos on, että tämä itse asiassa on yhtälön (2) yleinen ratkaisu edellyttäen ainoastaan, että funktiot y_1 ja y_2 ovat *lineaarisesti riippumattomat* välillä (a, b). Tällä tarkoitetaan, että pätee (vrt. Luku IV.5)

$$c_1y_1(x) + c_2y_2(x) = 0 \quad \forall x \in (a, b) \;\Rightarrow\; c_1 = c_2 = 0.$$

Jos tämä ei päde, niin sanotaan, että y_1 ja y_2 ovat *lineaarisesti riippuvat* välillä (a, b). Tässä tapauksessa on joko $y_2 = cy_1$ tai $y_1 = cy_2$ ko. välillä jollakin $c \in \mathbb{R}$.

*Lause IX.6.1 Jos differentiaaliyhtälössä (2) kerroinfunktiot P ja Q ovat jatkuvia välillä (a, b), niin yhtälöllä on ko. välillä kaksi lineaarisesti riippumatonta ratkaisua. Edelleen jos y_1 ja y_2 ovat mitkä tahansa kaksi yhtälön (2) lineaarisesti riippumatonta ratkaisua välillä (a, b), niin yleinen ratkaisu ko. välillä on

$$y(x) = C_1y_1(x) + C_2y_2(x), \quad C_1, C_2 \in \mathbb{R}.$$

Esimerkki 1 Eulerin differentiaaliyhtälön

$$x^2y'' - 3xy' - 3y = 0$$

perusmuodossa (2) kerroinfunktiot ovat $P(x) = -3/x$ ja $Q(x) = -3/x^2$. Nämä ovat jatkuvia väleillä $(-\infty, 0)$ ja $(0, \infty)$, ja näillä väleillä yleinen ratkaisu on (ks. edellinen luku)

$$y(x) = C_1x + C_2x^3, \quad C_1, C_2 \in \mathbb{R}.$$

Tulos on sopusoinnussa Lauseen IX.6.1 kanssa, sillä $y_1(x) = x$ ja $y_2(x) = x^3$ ovat lineaarisesti riippumattomat jokaisella välillä (a, b):

$$c_1x + c_2x^3 = 0 \;\; \forall x \in (a, b) \quad \Leftrightarrow \quad c_1 = c_2 = 0. \quad \square$$

Esimerkki 1 (jatko) Jos esimerkin differentiaaliyhtälöä $x^2y'' - 3xy' - 3y = 0$ tarkastellaan välillä $(-\infty, \infty)$, niin ratkaisuksi havaitaan myös

$$y_3(x) = |x|^3 = \begin{cases} x^3, & \text{kun } x \geq 0, \\ -x^3, & \text{kun } x < 0, \end{cases}$$

sillä tämä on kahdesti derivoituva ja toteuttaa yhtälön koko $\mathbb{R}$:ssä. Tätä funktiota ei kuitenkaan voi ilmaista muodossa $C_1y_1(x) + C_2y_2(x)$, sillä $y_1(x) = x$ ja $y_2(x) = x^3$ (samoin niiden lineaariset yhdistelyt) ovat parittomia, kun taas y_3 on parillinen. — Tulos ei ole ristiriidassa Lauseen IX.6.1 kanssa, sillä lauseen oletus kerroinfunktioiden P ja Q jatkuvuudesta ei toteudu välillä $(-\infty, \infty)$. Huomattakoon myös, että väleillä $(0, \infty)$ ja $(-\infty, 0)$ on $y_3(x) = \pm y_2(x)$, joten näillä väleillä y_3 ei tuo mitään uutta yleiseen ratkaisuun. $\square$

Lauseen IX.6.1 tulos on tulkittavissa niin, että yhtälön (2) ratkaisujen joukko $\mathcal{Y}$ on tehdyin oletuksin 2–ulotteinen vektoriavaruus (funktioavaruus, vrt. Luku IV.5). Tämän avaruuden kanta on mikä tahansa lineaarisesti riippumaton ratkaisupari $\{y_1, y_2\}$, jolloin koko ratkaisuavaruus voidaan esittää muodossa

$$\mathcal{Y} = \{y = c_1y_1 + c_2y_2 \mid c_1, c_2 \in \mathbb{R}\}.$$

Lauseen IX.6.1 (osittainen) todistus esitetään jäljempänä. Tätä ennen tarkastellaan lauseen seuraamuksia ajatellen yhtälön (2) ratkaisemista kvadratuureilla.

Homogeeninen yhtälö: Ratkaiseminen kvadratuureilla

Ensinnäkin joudutaan toteamaan, että yhtälö (2) ei aina ratkea kvadratuureilla. Kvadratuureihin palautumaton on esimerkksi niinkin yksinkertainen yhtälö kuin

$$y'' = xy.$$

Tämän huonon uutisen jälkeen todettakoon, että jos yhtälölle (2) on keksittävissä edes yksi ei-triviaali (eli nollasta poikkeava) ratkaisu y_1, niin toinen, y_1:stä lineaarisesti riippumaton ratkaisu — ja niin muodoin yleinen ratkaisu — on konstruoitavissa kvadratuureilla. Jatkossa rajoitutaan tähän tapaukseen, eli oletetaan yksi ratkaisu $y_1 \neq 0$ tunnetuksi.

Jos tunnetaan yhtälön (2) ratkaisu $y_1 \neq 0$, niin toista ratkaisua voidaan etsiä tuttuun tapaan vakion varioinnilla, eli muodossa

$$y(x) = C(x)y_1(x).$$

Tällöin on

$$y' = y_1C' + y_1'C, \quad y'' = y_1C'' + 2y_1'C' + y_1''C.$$

Koska $\mathrm{L}y_1 = 0$, seuraa C:lle differentiaaliyhtälö

$$y_1(x)C'' + [2y_1'(x) + P(x)y_1(x)]C' = 0.$$

Sijoituksella $u = C'$ tästä tulee separoituva:

$$u' + K(x)u = 0, \quad K(x) = 2\,\frac{y_1'(x)}{y_1(x)} + P(x) = 2\frac{d}{dx}\ln|y_1(x)| + P(x).$$

Separoimalla ja integroimalla saadaan

$$\begin{aligned} &\int \frac{du}{u} = -\int K(x)\,dx \\ \Leftrightarrow\ &\ln|u| = -2\ln|y_1(x)| - \int P(x)\,dx + A \quad (A \in \mathbb{R}). \end{aligned}$$

Valitsemalla $A = 0$ (yksi ratkaisu $u \neq 0$ riittää!) saadaan

$$u(x) = \frac{e^{-\int P(x)dx}}{[y_1(x)]^2} = C'(x).$$

Näin ollen eräs ratkaisu on

$$\begin{aligned} y_2(x) &= y_1(x)\int u(x)\,dx \\ &= y_1(x)\int \frac{e^{-\int P(x)dx}}{[y_1(x)]^2}\,dx. \end{aligned}$$

Ratkaisu on pätevä ainakin y_1:n nollakohtien välisillä avoimilla väleillä. Tällaisilla väleillä suhde $u = y_2/y_1$ ei selvästikään ole vakio (koska on positiivisen funktion integraalifunktio), joten löydetty ratkaisu on y_1:stä lineaarisesti riippumaton. Homogeenisen yhtälön (2) yleinen ratkaisu on tällöin (ainakin mainituilla väleillä) kirjoitettavissa

$$y(x) = C_1y_1(x) + C_2y_2(x).$$

Esimerkki 2 Differentiaaliyhtälöllä

$$y'' + xy' - y = 0$$

on ilmeinen ratkaisu $y_1(x) = x$, joten etsitään toista muodossa

$$y(x) = xC(x) \quad \Rightarrow \quad y' = xC' + C, \quad y'' = xC'' + 2C'.$$

Sijoitus yhtälöön antaa

$$xu' + (x^2 + 2)u = 0, \quad u = C'.$$

Ratkaisu separoimalla (ol. $x > 0$):

$$\int \frac{du}{u} = \ln u = -\int \left(x + \frac{2}{x}\right) dx = \ln\left(x^{-2}e^{-x^2/2}\right) + \ln A$$
$$\Rightarrow \quad u = C' = Ax^{-2}e^{-x^2/2}.$$

Valitsemalla $A = 1$, asettamalle ehto $\lim_{x\to\infty} C(x) = 0$ ja integroimalla osittain saadaan

$$C(x) = -\int_x^\infty u(t)\,dt = -x^{-1}e^{-x^2/2} + \int_x^\infty e^{-t^2/2}\,dt.$$

Yleinen ratkaisu on näin ollen

$$y(x) = C_1 x + C_2\, xC(x) = C_1 x + C_2\left(-e^{-x^2/2} + x\int_x^\infty e^{-t^2/2}\,dt\right), \quad C_1, C_2 \in \mathbb{R}.$$

Vaikka laskun välivaiheissa oletettiin $x > 0$, on ratkaisu pätevä koko $\mathbb{R}$:ssä. □

Homogeeninen yhtälö: ratkeavuusteoria

Palataan Lauseeseen IX.6.1, jota ei ole todistettu. Lauseen väittämistä syvällisin koskee — kuten tavallista — ratkaisujen y_1 ja y_2 olemassaoloa. Jatkossa johdetaan tämä tulos, samoin kuin Lauseen IX.6.1 muut väittämät, seuraavasta peruslauseesta, joka puolestaan on erikoistapaus paljon yleisemmästä alkuarvotehtävän ratkeavuutta koskevasta väittämästä. Väittämää ei tässä vaiheessa todisteta, vaan asiaan palataan myöhemmin luvussa IX.8. Puheena olevaan erikoistapaukseen sovellettuna ratkeavuusväittämä on seuraava:

*Lause IX.6.2 Jos P ja Q ovat jatkuvia välillä (a, b) ja $x_0 \in (a, b)$, niin alkuarvotehtävällä

$$\begin{cases} y'' + P(x)y' + Q(x)y = 0, & x \in (a, b), \\ y(x_0) = A, \\ y'(x_0) = B \end{cases}$$

on yksikäsitteinen ratkaisu jokaisella $A, B \in \mathbb{R}$.

Jatkossa siis todistetaan väittämä: Lause IX.6.2 $\Rightarrow$ Lause IX.6.1. Todistuksessa näyttelee keskeistä roolia seuraava käsite.

MÄÄRITELMÄ IX.6.3 Välillä (a,b) derivoituvien funktioiden y_1 ja y_2 *Wronskin determinantti* on

$$W_{12}(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = (y_1y_2' - y_2y_1')(x), \quad x \in (a,b).$$

PROPOSITIO IX.6.4 Jos y_1 ja y_2 ovat derivoituvia välillä (a,b) ja $W_{12}(x_0) = (y_1y_2' - y_2y_1')(x_0) \neq 0$ jollakin $x_0 \in (a,b)$, niin y_1 ja y_2 ovat lineaarisesti riippumattomat välillä (a,b).

<u>Todistus</u> Jos joillakin $c_1, c_2 \in \mathbb{R}$ pätee

$$c_1y_1(x) + c_2y_2(x) = 0, \quad x \in (a,b),$$

niin derivoimalla seuraa, että myös

$$c_1y_1'(x) + c_2y_2'(x) = 0, \quad x \in (a,b).$$

Kun nämä yhtälöt kirjoitetaan pisteessä x_0, niin saadaan yhtälöryhmä

$$\begin{cases} y_1(x_0)c_1 + y_2(x_0)c_2 = 0, \\ y_1'(x_0)c_1 + y_2'(x_0)c_2 = 0. \end{cases}$$

Koska $W_{12}(x_0) \neq 0$, niin yhtälöryhmän ainoa ratkaisu on $c_1 = c_2 = 0$.[†] Siis oletuksesta $c_1y_1(x) + c_2y_2(x) = 0$, $x \in (a,b)$ seuraa, että $c_1 = c_2 = 0$, joten y_1 ja y_2 ovat lineaarisesti riippumattomat välillä (a,b). $\square$

ESIMERKKI 3 Funktiot $y_1(x) = x^3$ ja $y_2(x) = |x|^3$ ovat (kehdestikin) derivoituvia ja lineaarisesti riippumattomia $\mathbb{R}$:ssä (vrt. Esimerkki 1 edellä). Laskemalla näiden funktioiden Wronskin determinantti todetaan, että $W_{12}(x) = 0$ $\forall x \in \mathbb{R}$. Tämän (vasta)esimerkin perusteella päätellään:

$$\begin{aligned} &y_1 \text{ ja } y_2 \text{ lineaarisesti riippumattomat välillä } (a,b) \\ &\qquad \not\Rightarrow \quad W_{12}(x_0) \neq 0 \quad \text{jollakin } x_0 \in (a,b). \quad \square \end{aligned}$$

[†]Jos $a_{ij}, x_i, b_i \in \mathbb{R}$, $i,j = 1,2$ ja $W = a_{11}a_{22} - a_{21}a_{12} \neq 0$, niin pätee

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = W^{-1}(a_{22}b_1 - a_{12}b_2), \\ x_2 = W^{-1}(a_{11}b_2 - a_{21}b_1). \end{cases}$$

Esimerkin perusteella Proposition IX.6.4 (implikaatio)väittämä ei päde kääntäen. Jos sen sijaan oletuksia vahvistetaan niin, että y_1 ja y_2 oletetaan differentiaaliyhtälön (2) ratkaisuiksi välillä (a, b), niin käänteinenkin väittämä on tosi. Tällaisella oletuksella saadaan seuraava, Propositiota IX.6.4 huomattavasti vahvempi tulos.

LAUSE IX.6.5 Jos y_1 ja y_2 ovat differentiaaliyhtälön (2) ratkaisuja välillä (a, b) (P ja Q jatkuvia välillä (a, b)) ja $W_{12}(x)$ on funktioiden y_1, y_2 Wronskin determinantti välillä (a, b), niin pätee

$$\begin{aligned} & W_{12}(x_0) = 0 \quad \text{jollakin } x_0 \in (a, b) \\ \Leftrightarrow \quad & W_{12}(x) \ = 0 \quad \text{jokaisella } x \in (a, b) \\ \Leftrightarrow \quad & y_1 \text{ ja } y_2 \text{ lineaarisesti riippuvat välillä } (a, b). \end{aligned}$$

<u>Todistus</u> Oletuksien mukaan

$$\begin{cases} y_1'' + P(x)y_1' + Q(x)y_1 = 0, & x \in (a, b), \\ y_2'' + P(x)y_2' + Q(x)y_2 = 0, & x \in (a, b). \end{cases}$$

Kertomalla ensimmäinen yhtälö y_2:lla ja toinen y_1:llä ja vähentämällä seuraa

$$y_1 y_2'' - y_2 y_1'' + P(x)(y_1 y_2' - y_2 y_1') = 0, \quad x \in (a, b).$$

Mutta

$$y_1 y_2'' - y_2 y_1'' = \frac{d}{dx}(y_1 y_2' - y_2 y_1').$$

Siis on päätelty, että $y(x) = W_{12}(x)$ on jokaisella $x_0 \in (a, b)$ ratkaisu alkuarvotehtävälle

$$\begin{cases} y' + P(x)y = 0, & x \in (a, b) \\ y(x_0) = W_{12}(x_0) \end{cases} \tag{3}$$

Koska ratkaisu on (vrt. Luku IX.4)

$$W_{12}(x) = W_{12}(x_0)\, e^{-\int_{x_0}^{x} P(t)dt}, \quad x \in (a, b)$$

ja koska $e^t \neq 0 \ \forall t \in \mathbb{R}$, niin seuraa ensimmäinen osaväittämä: $W_{12}(x_0) = 0 \ \Leftrightarrow \ W_{12}(x) = 0 \ \forall x \in (a, b)$.

Tapauksessa $y_1(x) = 0 \ \forall x \in (a, b)$ on toinenkin osaväittämä tosi, joten oletetaan, että $y_1(x_0) \neq 0$ jollakin $x_0 \in (a, b)$. Tällöin on y_1:n jatkuvuuden nojalla $y_1(x) \neq 0$ jollakin välillä $(x_0 - \delta, x_0 + \delta) \subset (a, b)$, $\delta > 0$. Tällä välillä y_2/y_1 on derivoituva ja

$$\frac{d}{dx}\left[\frac{y_2(x)}{y_1(x)}\right] = \frac{W_{12}(x)}{[y_1(x)]^2}.$$

Näin ollen jos $W_{12}(x) = 0$ välillä (a, b), niin jollakin $C \in \mathbb{R}$ pätee

$$\frac{y_2(x)}{y_1(x)} = C \;\Leftrightarrow\; -Cy_1(x) + y_2(x) = 0 \quad \text{välillä } (x_0 - \delta, x_0 + \delta).$$

Kun merkitään $y(x) = -Cy_1(x) + y_2(x),\ x \in (a, b)$, niin on siis $y(x) = 0$ mainitulla osavälillä, jolloin on erityisesti $y(x_0) = y'(x_0) = 0$. Koska y on myös differentiaaliyhtälön (2) ratkaisu välillä (a, b), niin y ratkaisee siis Lauseen IX.6.2 alkuarvotehtävän, kun $A = B = 0$. Koska ilmeinen ratkaisu on myös $y = 0$ ja koska ratkaisu on mainitun lauseen mukaan yksikäsitteinen, niin on siis oltava $y(x) = -Cy_1(x) + y_2(x) = 0,\ x \in (a, b)$. Tällöin y_1 ja y_2 ovat lineaarisesti riippuvat välillä (a, b). Päättely perustui oletukseen, että $W_{12}(x) = 0$ välillä (a, b), joten toisen osaväittämän osa $\boxed{\Rightarrow}$ tuli todistetuksi. Jäljelle jäävä osa $\boxed{\Leftarrow}$ on jo todistettu, sillä tämä on Proposition IX.6.4 väittämän loogisesti ekvivalenttinen muoto. $\square$

Nyt ollaan valmiita esittämään (Lauseeseen IX.6.2 nojaava)

<u>Lauseen IX.6.1 todistus</u> Olkoot y_1 ja y_2 Lauseen IX.6.2 alkuarvotehtävän ratkaisut, kun a) $A = 1,\ B = 0$, b) $A = 0,\ B = 1$. Tällöin funktioiden y_1 ja y_2 Wronskin determinantti x_0:ssa on

$$W_{12}(x_0) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

Proposition IX.6.4 mukaan y_1 ja y_2 ovat lineaarisesti riippumattomat välillä (a, b), joten ensimmäinen osaväittämä on todistettu.

Toisen osaväittämän todistamiseksi oletetaan, että y_1 ja y_2 ovat mitkä tahansa kaksi differentiaaliyhtälön (2) lineaarisesti riippumatonta ratkaisua välillä (a, b) ja että u on kolmas saman DY:n ratkaisu. Tällöin jos voidaan määritellä kertoimet c_1 ja c_2 siten, että funktiolle $y(x) = c_1y_1(x) + c_2y_2(x)$ pätee $y(x_0) = u(x_0),\ y'(x_0) = u'(x_0)$, niin u ja y ovat saman alkuarvotehtävän ratkaisuja, jolloin Lauseen IX.6.2 mukaan on oltava $u(x) = y(x),\ x \in (a, b)$. Kertoimet määräytyvät yhtälöryhmästä

$$\begin{cases} y_1(x_0)c_1 + y_2(x_0)c_2 = u(x_0), \\ y_1'(x_0)c_1 + y_2'(x_0)c_2 = u'(x_0). \end{cases}$$

Tämä ratkeaa yksikäsitteisesti ehdolla $y_1(x_0)y_2'(x_0) - y_1'(x_0)y_2(x_0) \neq 0$, eli ehdolla $W_{12}(x_0) \neq 0$, missä W_{12} on funktioiden y_1 ja y_2 Wronskin determinantti. Mutta oletuksen ja Lauseen IX.6.5 perusteella $W_{12}(x_0) \neq 0$ jokaisella $x_0 \in (a, b)$. Siis u on esitettävissä funktioiden y_1 ja y_2 lineaarisena yhdistelynä. $\square$

Täydellinen yhtälö: vakioiden variointi

Jos homogeenisen yhtälön (2) yleinen ratkaisu tunnetaan, niin täydellisen yhtälön (1) ratkaisemiseksi riittää jälleen löytää tälle yksittäisratkaisu. Yleinen menetelmä perustuu vakioiden variointiin: Jos homogeenisen yhtälön yleinen ratkaisu on $y(x) = C_1y_1(x) + C_2y_2(x)$, niin täydelliselle yhtälölle löydetään yksittäisratkaisu muodossa

$$y(x) = C_1(x)y_1(x) + C_2(x)y_2(x).$$

Tällöin

$$y' = C_1(x)y_1'(x) + C_2(x)y_2'(x) + C_1'(x)y_1(x) + C_2'(x)y_2(x).$$

Tämän lausekkeen yksinkertaistamiseksi asetetaan lisäehto

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0$$

(osoittautuu mahdolliseksi!), jolloin

$$\begin{aligned} y' &= C_1(x)y_1'(x) + C_2(x)y_2'(x), \\ y'' &= C_1(x)y_1''(x) + C_2(x)y_2''(x) + C_1'(x)y_1'(x) + C_2'(x)y_2'(x) \end{aligned}$$

ja näin ollen

$$\begin{aligned} \mathrm{L}y &= y'' + P(x)y' + Q(x)y \\ &= C_1(x)\mathrm{L}y_1 + C_2(x)\mathrm{L}y_2 + C_1'(x)y_1'(x) + C_2'(x)y_2'(x). \end{aligned}$$

Tässä on $\mathrm{L}y_1 = \mathrm{L}y_2 = 0$, joten oletettu lisäehto huomioiden on saatu yhtälöryhmä

$$\begin{cases} y_1(x)C_1' + y_2(x)C_2' = 0, \\ y_1'(x)C_1' + y_2'(x)C_2' = R(x). \end{cases}$$

Koska tässä on (Lause IX.6.5)

$$(y_1y_2' - y_2y_1')(x) = W_{12}(x) \neq 0,$$

niin yhtälöryhmä ratkeaa:

$$C_1' = -\frac{y_2(x)R(x)}{W_{12}(x)}, \quad C_2' = \frac{y_1(x)R(x)}{W_{12}(x)}.$$

Integroimalla $C_1(x)$ ja $C_2(x)$ näistä yhtälöistä on löydetty täydellisen yhtälön (1) yksittäisratkaisu. Yhtälön (1) yleiseksi ratkaisuksi tulee näin muodoin

$$y(x) = C_1y_1(x) + C_2y_2(x) - y_1(x)\int \frac{y_2(x)R(x)}{W_{12}(x)}\,dx + y_2(x)\int \frac{y_1(x)R(x)}{W_{12}(x)}\,dx.$$

Esimerkki 4 Ratkaise Eulerin differentiaaliyhtälö

$$x^2y'' + 4xy' + 2y = f(x), \quad x > 0.$$

Ratkaisu Homogeenisen yhtälön yleinen ratkaisu on (ks. edellinen luku)

$$y(x) = \frac{C_1}{x} + \frac{C_2}{x^2} = C_1y_1(x) + C_2y_2(x).$$

Em. laskukaavoissa on

$$R(x) = \frac{f(x)}{x^2}, \quad W_{12}(x) = (y_1y_2' - y_1'y_2)(x) = -\frac{1}{x^4},$$

joten vakioiden variointi antaa

$$C_1' = f(x) \quad \Rightarrow \quad C_1(x) = \int f(x)\,dx,$$

$$C_2' = -xf(x) \quad \Rightarrow \quad C_2(x) = -\int xf(x)\,dx.$$

Yksittäisratkaisu on siis

$$y(x) = \frac{1}{x}\int f(x)\,dx - \frac{1}{x^2}\int xf(x)\,dx.$$

Tämä on myös yleinen ratkaisu, kun oikealla puolella integraaleihin sisällytetään määräämättömät integroimisvakiot. □

Tunnettuja differentiaaliyhtälöitä

Toisen kertaluvun lineaarinen, homogeeninen differentiaaliyhtälö on sovelluksissa yleinen differentiaaliyhtälön tyyppi. Seuraavassa muutamia tunnettuja differentiaaliyhtälöitä, jotka on nimetty niitä tutkineiden matemaatikkojen mukaan.

Legendre: $(x^2 - 1)y'' + 2xy' - n(n+1)y = 0$

Tšebyšev: $(1 - x^2)y'' - xy' + n^2y = 0$

Hermite: $y'' - 2xy' + 2ny = 0$

Laguerre: $xy'' + (1 - x)y' + ny = 0$

Bessel: $x^2y'' + xy' + (x^2 - p^2)y = 0$

Näistä muilla paitsi Besselin differentiaaliyhtälöllä on ratkaisuna polynomi astetta n, kun $n \in \mathbb{N} \cup \{0\}$ (sovelluksien kannalta kiinnostavin tapaus). Ko. polynomit on nimetty samoin kuin differentiaaliyhtälöt. Besselin differentiaaliyhtälössä $p \geq 0$ on reaalinen paramatetri

Sarjaratkaisut

Jos toisen kertaluvun lineaarisella, homogeenisella differentiaaliyhtälöllä ei ole yhtään 'tunnistettavaa' ratkaisua, niin ratkaisemista voidaan yrittää *sarjamenetelmällä*. Tällöin ratkaisua etsitään joko potenssisarjana tai potenssisarjoja sisältävänä lausekkeena. Menetelmän perusidea on sama kuin integroinnin sarjamenetelmässä (ks. Luku VIII.3).

Esimerkkinä sarjamenetelmän soveltamisesta olkoon Besselin differentiaaliyhtälö, jonka perusmuoto on

$$y'' + \frac{1}{x}y' + (1 - \frac{p^2}{x^2})y = 0.$$

Yleinen ratkaisu (sovelluksissa yleensä välillä $(0, \infty)$) kirjoitetaan muodossa

$$y(x) = C_1 J_p(x) + C_2 Y_p(x),$$

missä J_p ja Y_p ovat *Besselin funktioita*, tarkemmin *ensimmäisen lajin* (J_p) ja *toisen lajin* (Y_p) Besselin funktioita. (Tapauksessa $p = 1/2$ nämä ovat poikkeuksellisesti alkeisfunktioita, ks. Harj.teht. 7.) Funktio J_p on muotoa

$$J_p(x) = \sum_{k=0}^{\infty} a_k x^{2k+p} = x^p F(x),$$

missä potenssisarja $F(x) = \sum_{k=0}^{\infty} a_k x^{2k}$ suppenee kaikkialla. Muilla kuin kokonaislukuarvoilla on vastaavasti

$$Y_p(x) = \sum_{k=0}^{\infty} a_k x^{2k-p} = x^{-p} G(x), \quad p \notin \{0, 1, 2, \ldots\}$$

missä G:n sarja jälleen suppenee kaikkialla. Jos p on kokonaisluku, niin Y_p sisältää myös logaritmisen termin muotoa $x^{-p} H(x) \ln |x|$, missä $H(x)$ on ilmaistavissa potenssisarjana.

Olkoon $p = 0$, ja yritetään sarjaratkaisua $y(x) = \sum_{k=0}^{\infty} a_k x^k$. Sijoitus yhtälöön antaa

$$\sum_{k=2}^{\infty} k(k-1) a_k x^{k-2} + \sum_{k=1}^{\infty} k a_k x^{k-2} + \sum_{k=0}^{\infty} a_k x^k = 0.$$

Kun kahdessa ensimmäisessä summassa vaihdetaan summausindeksiksi k:n tilalle $k + 2$, saadaan

$$\sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k + \left[a_1 x^{-1} + \sum_{k=0}^{\infty} (k+2) a_{k+2} x^k \right] + \sum_{k=0}^{\infty} a_k x^k = 0 \quad \forall x$$

eli yhdistämällä summat

$$a_1 x^{-1} + \sum_{k=0}^{\infty} [(k+2)^2 a_{k+2} + a_k] x^k = 0 \quad \forall x.$$

Nähdään, että tämä toteutuu, kun valitaan

$$\begin{aligned} &a_1 = a_3 = a_5 = \ldots = 0 \\ &a_0 = 1, \quad a_{k+2} = -\frac{1}{(k+2)^2}\, a_k, \quad k = 0, 2, 4 \ldots \end{aligned}$$

Kertoimille a_{2k} saadaan (pienen mietiskelyn jälkeen) lauseke

$$a_{2k} = (-1)^k \frac{1}{(k!)^2 \cdot 4^k}, \quad k = 0, 1, \ldots$$

Ratkaisu, nimeltään J_0, on siis

$$J_0(x) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{(k!)^2} \left(\frac{x}{2}\right)^{2k}.$$

Tämä sarja suppenee (ja on termeittäin derivoitavissa) kaikkialla, joten on löydetty koko $\mathbb{R}$:ssä pätevä ratkaisu differentiaaliyhtölölle $xy'' + y' + xy = 0$. Toisen lajin ratkaisu $Y_0(x)$ on pätevä vain erikseen väleillä $(-\infty, 0)$ ja $(0, \infty)$. Tämäkin on ilmaistavissa potenssisarjojen avulla (ks. Harj.teht. 10).

HARJOITUSTEHTÄVIÄ

1. Minkä toisen kertaluvun lineaarisen ja homogeenisen differentiaaliyhtälön ratkaisuja (jollakin avoimella välillä) ovat a) $y_1(x) = x + 1$ ja $y_2(x) = x^2$, b) $y_1(x) = x$ ja $y_2(x) = e^x$, c) $y_1(x) = x^2$ ja $y_2(x) = \sin x$?

2. Ratkaise käyttäen annettua, yksittäisratkaisua koskevaa lisätietoa:
   a) $xy'' - (x+3)y' + y = 0$; polynomi
   b) $x^2(\ln x - 1)y'' - xy' + y = 0$; polynomi
   c) $(x^2 - 2x - 1)y'' - (2x - 1)y' + 2y = 0$; polynomi
   d) $y'' + (\tan x - 2\cot x)y' + 2(\cot^2 x)y = 0$; $\sin x$
   e) $y'' + 2(1 - \tan^2 x)y = 0$; $\cos^2 x$
   f) $(x-2)y'' - (4x-7)y' + (4x-6)y = 0$; e^{ax}

3. Differentiaaliyhtälöllä $y'' + P(x)y' + Q(x)y = 0$ on helposti arvattava eitriviaali ratkaisu, jos $1 + P(x) + Q(x) = 0$. Ratkaise tämän (ja tarvittaessa toisenkin) arvauksen perusteella:

   a) $(x-1)y'' - xy' + y = 0$
   b) $(2x - x^2)y'' + (x^2 - 2)y' + (2 - 2x)y = 0$
   c) $y'' - (2 + 2x + x^{-1})y' + (1 + 2x + x^{-1})y = 0$
   d) $(x+1)y'' - xy' - y = (x+1)^2$

4. Oletetaan, että differentiaaliyhtälön $y'' + P(x)y' + Q(x)y = 0$ kerroinfunktioille pätee: P on derivoituva ja P' ja Q ovat jatkuvia tarkasteltavalla välillä (a, b). Tällöin voidaan sijoituksella $y(x) = K(x)u(x)$ muuntaa DY muotoon
$$u'' + I(x)u = 0.$$
   Miten $K(x)$ on valittava ja mikä on $I(x)$:n lauseke?

5. Ratkaise vakioiden varioinnilla:

   a) $y'' + y = \tan x$
   b) $y'' - y = 1/(e^x + 1)$
   c) $y'' + 2y' + y = 1/(x^2 + 1)$
   d) $x^2y'' + 2xy' - 2y = x^2e^x$
   e) $y'' + 5y' + 4y = f(x)$
   f) $y'' + 4y' + 4y = f(x)$
   g) $y'' + 4y' + 5y = f(x)$
   h) $x^2y'' + 5xy' + 3y = f(x)$
   i) $x^2y'' + 5xy' + 4y = f(x)$
   j) $x^2y'' - 2y = f(x)$

6. a) *Legendren polynomi* astetta n määritellään
$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n.$$
   Totea, että P_n on polynomi astetta n ja näytä, että tämä on Legendren differentiaaliyhtälön ratkaisu n:n arvoilla $0, 1, 2, 3$.

   b) Määritä Hermiten ja Laguerren polynomit astetta $n \leq 2$, ts. kyseisten differentiaaliyhtälöiden polynomiratkaisut, kun $n = 0, 1, 2$.

   Määritä c) Legendren, d) Tšebyševin, e) Hermiten, f) Laguerren DY:n yleinen ratkaisu, kun $n = 1$.

7. Tapauksessa $p = 1/2$ Besselin differentiaaliyhtälö on ratkaistavissa alkeisfunktioilla. Määritä ratkaisu sijoituksella $y(x) = u(x)/\sqrt{x}$.

8. Seuraavien differentiaaliyhtälöiden yleinen ratkaisu on esitettävissä pisteen $x = 0$ lähellä muodossa $y(x) = C_1 \sum_{k=0}^{\infty} a_k x^k + C_2 \sum_{k=0}^{\infty} b_k x^k$, missä $a_0 = b_1 = 1$ ja $a_1 = b_0 = 0$. Määritä molempien potenssisarjojen neljä ensimmäistä nollasta poikkeavaa termiä.

   a) $y'' + 2xy' + 5y = 0$ b) $y'' + 2xy + 6y = 0$ c) $y'' + xy' - x^2y = 0$
   d) $y'' + (x+1)y' + (x-1)y = 0$ e) $(1 - x^2)y'' - xy' + ay = 0,\ a \in \mathbb{R}$

9. (*) *Tšebyševin polynomi* astetta n määritellään välillä $[-1,1]$ kaavalla

$$T_n(x) = \cos(n \operatorname{Arccos} x), \quad x \in [-1,1].$$

a) Näytä, että T_n on polynomi astetta n (sijoita $x = \cos\theta$). b) Näytä, että T_n on Tšebyševin differentiaaliyhtälön ratkaisu välillä $(-1,1)$. c) Laske polynomin T_5 kertoimet ja hahmottele $T_5(x)$ graafisesti välillä $[-2,2]$. Mitkä ovat T_5:n nollakohdat ja missä pisteissä on $T_5(x) = \pm 1$?

10. (*) Näytä, että jos $p = 0$, niin Besselin differentiaaliyhtälöllä on välillä $(0,\infty)$ (toisen lajin) ratkaisu muotoa $Y_0(x) = J_0(x)\ln x + F(x)$, missä F on ilmaistavissa kaikkialla suppenevana potenssisarjana.

11. (*) Näytä, että differentiaaliyhtälön $y'' = xy$ yleinen ratkaisu on esitettävissä potenssisarjana muodossa

$$y(x) = C_1 \sum_{k=0}^{\infty} a_k x^{3k} + C_2 \sum_{k=0}^{\infty} b_k x^{3k+1}, \quad a_0 = b_0 = 1.$$

Määrittele kerroinjonot $\{a_k\}$ ja $\{b_k\}$ palautuvina lukujonoina ja päättele, että sarja suppenee kaikkialla. Millainen on yleisemmin differentiaaliyhtälön $y'' = x^m y$ yleinen sarjaratkaisu, kun $m \in \mathbb{N}$?

IX.7 Differentiaaliyhtälöiden numeeriset ratkaisumenetelmät

Edellisissä luvuissa on kehitelty differentiaaliyhtälöiden ratkaisemistaitoa lähinnä nk. 'tarkan' ratkaisemisen näkökulmasta. Tämä ratkaisuoppi on hyvin perinteistä, ja sillä on oma pysyvä arvonsa. — Etenkin silloin kun ratkaisun voi esittää alkeisfunktioden avulla suhteellisen yksinkertaisena lausekkeena, on perinteistä ratkaisutapaa vaikea sivuuttaa. Sovelluksissa törmätään kuitenkin hyvin usein differentiaaliyhtälöihin (ja etenkin -systeemeihin), joihin mikään tähän asti tarjotuista resepteistä ei tehoa. Silloin on tyydyttävä etsimään ratkaisu likimäärin *numeerisesti*, käytännössä tietokoneen tai tehokkaan laskimen avustuksella.[†] Numeerinen ratkaiseminen siis poikkeaa melkoisesti perinteisistä menetelmistä, joilla ratkaisu yleensä saadaan (jos saadaan) 'käsipelillä'. Mitään periaatteellista eroa ei numeerisen ja 'tarkan' ratkaisemisen välillä silti ole, sillä myös numeerisin keinoin saadaan tarkka ratkaisu, kun ajatellaan algoritmin tuottama likimääräisratkaisujen jono 'loppuun asti' lasketuksi. — Itse asiassa juuri näin meneteltiin Luvussa VIII.4, kun etsittiin yleispätevää ratkaisutapaa differentiaaliyhtälölle $y' = f(x)$.

Jatkossa tarkastellaan lyhyesti alkuarvotehtävien numeerisessa ratkaisussa käytettävien nk. *askelmenetelmien* perusiedoita. Otetaan tarkastelun kohteeksi jälleen ensimmäisen kertaluvun normaalimuotoinen alkuarvotehtävä

$$\begin{cases} y' = f(x, y), & x > x_0, \\ y(x_0) = y_0. \end{cases} \tag{1}$$

Kaikissa askelmenetelmissä on ideana lähteä alkuarvopisteestä x_0 ja edetä lyhyin askelin, ensin pisteeseen x_1, sitten pisteeseen x_2, jne. Jokaisella askeleella $x_k \to x_{k+1}$ lasketaan luku y_{k+1}, joka pyrkii olemaan likiarvo tarkan ratkaisun $y(x)$ arvolle pisteessä x_{k+1}:

$$y_{k+1} \approx y(x_{k+1}), \quad k = 0, 1, \ldots$$

Lukua y_{k+1} laskettaessa tunnetaan jo luvut $y_0, y_1, \ldots y_k$ (alkuehdosta ja aikaisemmilta askelilta), joten y_{k+1} voidaan laskea palautuvasti. Jos laskukaavassa esiintyy vain edeltävä luku y_k, on kyseessä *yksiaskelmenetetelmä*, muussa tapauksessa *moniaskelmenetelmä*. Erilaiset askelmenetelmät erottaa siis lopulta toisistaan vain se, millaista palautuskaavaa tai yleisempää algoritmia luvun y_{k+1} laskemisessa

[†]Differentiaaliyhtälöitä ratkottiin numeerisesti jo ennen tietokoneiden aikaa, mutta ratkaiseminen rajoittui vain hyvin tärkeinä pidettyihin kohteisiin, kuten tykinammusten lentoratojen määrittämiseen. Tietokoneiden aikakaudella differentiaaliyhtälöiden numeerisesta ratkaisemisesta on tullut arkipäivää mitä erilaisimmilla tieteen ja tekniikan aloilla. Samalla on ratkaisumenetelmistä kehittynyt laaja menetelmätiede, josta on kirjoitettu paksuja kirjoja.

käytetään. Tarkkuutta voidaan säädellä sekä menetelmän valinnalla että *askelpituuksilla* $x_{k+1} - x_k$ (vrt. numeerisen integroinnin menetelmät Luvussa VIII.9). Seuraavassa tarkastellaan ensin askelmenetelmistä kaikkein yksinkertaisinta — myös vanhinta.

Eulerin menetelmä

Eulerin menetelmä† on yksiaskelmenetelmä, jossa luku y_{k+1} lasketaan palautuskaavasta

$$\boxed{y_{k+1} = y_k + hf(x_k, y_k), \quad h = x_{k+1} - x_k, \quad k = 0, 1, \ldots}$$

Kaavan voi johtaa esim. seuraavasti: Integroidaan differentiaaliyhtälö $y' = f(x, y)$ puolittain välin $[x_k,\, x_{k+1}]$ yli, jolloin seuraa

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(x, y(x))dx.$$

Oletetaan, että $f(x, y(x)) = y'(x)$ on likimain vakio välillä $[x_k, x_{k+1}]$ ja tehdään tähän nojaten approksimaatio

$$\int_{x_k}^{x_{k+1}} f(x, y(x))dx \approx hf(x_k, y(x_k)), \quad h = x_{k+1} - x_k.$$

(Kyse on yksinkertaisesta numeerisesta integroinnista, vrt. Luku VIII.9.) Eulerin menetelmään päädytään, kun tulos

$$y_{k+1} \approx y_k + hf(x_k, y_k), \quad y_k = y(x_k),\ y_{k+1} = y(x_{k+1})$$

'murjotaan' muotoon

$$y_{k+1} = y_k + hf(x_k, y_k), \quad y_k \approx y(x_k),\ y_{k+1} \approx y(x_{k+1}).$$

Jos askelpituus h oletetaan tunnetuksi (esim. vakio), niin Eulerin menetelmän askel koostuu yhdestä funktioevaluaatiosta, yhdestä kertolaskusta ja kahdesta yhteenlaskusta:

$$x_k = x_{k-1} + h, \quad t_k = f(x_k, y_k), \quad y_{k+1} = y_k + ht_k.$$

†Sveitsiläinen matemaatikko **Leonhard Euler** (1707-1783) on entuudestaan tuttu jo kompleksisesta eksponenttifunktiosta (Eulerin kaava) ja hänen mukaansa nimetystä differentiaaliyhtälöstä. *Eulerin yhtälöt* (vallan toisessa merkityksessä) ovat keskeisiä matematiikan lajissa nimeltä *variaatiolaskenta* ja (jälleen toisessa merkityksessä) virtausmekaniikassa. Euler tutki aikansa matematiikkaa ja fysiikkaa hyvin laajasti, myös tähtitiedettä, musiikkia, merenkulkua, ym. Euler oli 1700-luvun merkittävimpiä ja kaikkien aikojen monipuolisimpia matemaatikkoja.

Esimerkki 1 Laske Neperin luvulle e likiarvo y_n ratkaisemalla alkuarvotehtävä $y' = y,\ y(0) = 1$ Eulerin menetelmällä ja vakioaskelpituudella $h = 1/n,\ n \in \mathbb{N}$. Arvioi virhe.

Ratkaisu Eulerin menetelmällä laskien saadaan

$$\begin{aligned} y_0 = 1, \quad y_k &= y_{k-1} + hy_{k-1} = (1+h)y_{k-1}, \quad k = 1, 2 \ldots \\ \Rightarrow \quad y_n &= (1+h)^n = \underline{\underline{\left(1 + \frac{1}{n}\right)^n}}. \end{aligned}$$

Virheen arvioimiseksi tarkastellaan lauseketta

$$\ln \frac{y_n}{e} = n \ln\left(1 + \frac{1}{n}\right) - 1.$$

Taylorin lauseen (ks. Luku VII.4) mukaan on $\ln(1+x) = x - \frac{1}{2}x^2 + \mathcal{O}(|x|^3)$, joten seuraa

$$\begin{aligned} \ln \frac{y_n}{e} = -\frac{1}{2n} + \mathcal{O}(n^{-2}) \quad &\Rightarrow \quad \frac{y_n}{e} = e^{-\frac{1}{2n} + \mathcal{O}(n^{-2})} = 1 - \frac{1}{2n} + \mathcal{O}(n^{-2}) \\ &\Rightarrow \quad e - y_n = \underline{\underline{\frac{e}{2n} + \mathcal{O}(n^{-2})}}. \quad \square \end{aligned}$$

Eulerin menetelmän virhe

Edellä esimerkissä saatu virhearvio $y_n - y(x_n) = \mathcal{O}(h)$ pätee Eulerin menetelmälle melko yleisin edellytyksin. Yleisemmän virhearvion johtamiseksi olkoon askelpituus $h =$ vakio ja kirjoitetaan $Y_k = y(x_k) = y(x_0 + kh)$, missä $y(x)$ on alkuarvotehtävän (1) ratkaisu. Arvioidaan ensin Eulerin menetelmän nk. *konsistenssivirhe*, eli luku δ_k yhtälössä

$$Y_{k+1} = Y_k + hf(x_k, Y_k) + \delta_k.$$

Kun merkitään $u(x) = f(x, y(x)) = y'(x)$, niin osittain integroimalla seuraa (ks. myös Eulerin menetelmän johto edellä)

$$\begin{aligned} \delta_k &= \int_{x_k}^{x_{k+1}} u(x)\,dx - hu(x_k) \\ &= \Big|_{x_k}^{x_{k+1}} (x - x_{k+1})u(x) - \int_{x_k}^{x_{k+1}} (x - x_{k+1})u'(x)\,dx - hu(x_k) \\ &= \int_{x_k}^{x_{k+1}} (x_{k+1} - x)y''(x)\,dx. \end{aligned}$$

Oltetetaan nyt (Oletus #1), että y on kahdesti jatkuvasti derivoituva ja pätee arvio $|y''(x)| \leq M$ (kyseeseen tulevilla x:n arvoilla). Tällöin konsistenssivirhe on enintään

$$|\delta_k| \leq \int_{x_k}^{x_{k+1}} (x_{k+1} - x)\, M\, dx = \frac{1}{2}\, Mh^2.$$

Käyttäen tätä arviota ja kolmioepäyhtälöä voidaan nyt päätellä:

$$\begin{cases} Y_{k+1} = Y_k + hf(x_k, Y_k) + \delta_k, \\ y_{k+1} = y_k + hf(x_k, y_k) \end{cases}$$

$$\Rightarrow \quad |Y_{k+1} - y_{k+1}| \leq |Y_k - y_k| + h|f(x_k, Y_k) - f(x_k, y_k)| + \frac{1}{2}\, Mh^2.$$

Oletetaan (Oletus #2), että oikealla puolella voidaan käyttää arviota

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

($x = x_k$, $y_1 = Y_k$, $y_2 = y_k$), missä L on vakio. Tällöin seuraa

$$|Y_{k+1} - y_{k+1}| \leq (1 + hL)\, |Y_k - y_k| + \frac{1}{2}\, Mh^2.$$

Soveltamalla tätä epäyhtälöä palautuvasti, kun $k = 0, 1 \ldots$ seuraa (induktio!)

$$|Y_n - y_n| \;\leq\; \frac{1}{2}\, Mh^2 \sum_{k=0}^{n-1} (1 + hL)^k \;=\; \frac{M}{2L}\big[(1 + hL)^n - 1\big]\, h, \quad n = 1, 2 \ldots$$

Käyttämällä tässä vielä arviota $(1 + hL)^n \leq \big(e^{hL}\big)^n = e^{Lnh} = e^{L(x_n - x_0)}$ saadaan Eulerin menetelmän virhearvioksi tehdyin oletuksin

$$\boxed{|y(x_n) - y_n| \leq \frac{M}{2L}\big(e^{L(x_n - x_0)} - 1\big)\, h, \quad x_n = x_0 + nh.} \tag{2}$$

Esimerkki 2 Alkuarvotehtävä

$$\begin{cases} y' = e^{-x} + x \sin y, \quad x > 0, \\ y(0) = 1 \end{cases}$$

ei ratkea perinteisin menetelmin, joten on tyytyminen numeeriseen ratkaisuun. Differentiaaliyhtälöstä nähdään, että ratkaisulle pätee $|y'(x)| \leq 1 + x$, kun $x \geq 0$. Derivoimalla implisiittisesti ja käyttämällä tätä arviota seuraa

$$y''(x) = -e^{-x} + \sin y + xy' \cos y \quad \Rightarrow \quad |y''(x)| \leq 2 + x + x^2, \;\; x \geq 0.$$

Koska edelleen on

$$|f(x,y_1) - f(x,y_2)| \;=\; x|\sin y_1 - \sin y_2| \;\le\; x|y_1 - y_2|, \quad x \ge 0,$$

niin nähdään, että Eulerin menetelmän virhearvio (2) on pätevä, kun arviossa asetetaan

$$L = x_n = nh, \quad M = 2 + x_n + x_n^2 \quad (n \ge 1).$$

Esimerkiksi jos halutaan ratkaisu välillä $[0,2]$, niin voidaan valita $L = 2,\ M = 8$, jolloin virhe on enintään

$$|y(x_n) - y_n| < 2e^4 h < 110h \quad (x_n \le 2\,).$$

Todennäköisesti virhe on huomattavasti tätä arviota pienempi. □

Differentiaaliyhtälösysteemit

Eulerin menetelmän, samoin kuin muidenkin askelmenetetelmien (ks. esimerkit jäljempänä) vahva puoli sovellusten kannalta on, että menetelmät soveltuvat sellaisenaan myös normaalimuotoisen differentiaaliyhtälösysteemin ratkaisuun, sikäli kuin kyseessä on alkuarvotehtävä. Myös korkeamman kertaluvun differentiaaliyhtälöiden alkuarvotehtävät ratkeavat tällä tavoin kirjoittamalla tehtävä ensin systeemimuotoon, vrt. Luku IX.1. Esimerkki valaiskoon asiaa.

Esimerkki 3 Esitä Eulerin menetelmään perustuva algoritmi, jolla voidaan ratkaista numeerisesti alkuarvotehtävä

$$\begin{cases} y''' = x + yy'y'', \quad x > 0, \\ y(0) = 1,\ y'(0) = y''(0) = 0. \end{cases}$$

<u>Ratkaisu</u> Alkuehdoista ja ratkaistavan differentiaaliyhtälön systeemimuodosta

$$\begin{cases} y' = u, \\ u' = v, \\ v' = x + yuv \end{cases}$$

saadaan algoritmiksi

$$y_0 = 1,\ u_0 = v_0 = 0, \quad \begin{cases} y_{k+1} = y_k + hu_k, \\ u_{k+1} = u_k + hv_k, \\ v_{k+1} = v_k + h(x_k + y_k u_k v_k), \quad k = 0, 1 \ldots \end{cases} \quad \square$$

*Muita askelmenetelmiä

Tavallisten differentiaaliyhtälöiden alkuarvotehtävissä käytettävien 'mahdollisten menetelmien luettelo' on nykyisin tavattoman pitkä. Seuraavassa vain muutamia esimerkkejä usein käytetyistä tai muuten hyvin tunnetuista askelmenetelmistä.

Implisiittinen Euler

$y_{k+1} = y_k + hf(x_{k+1}, y_{k+1})$

Trapetsi

$y_{k+1} = y_k + \frac{h}{2}\big[f(x_k, y_k) + f(x_{k+1}, y_{k+1})\big]$

Keskipistesääntö

$y_{k+1} = y_k + hf\left(x_k + \frac{1}{2}h,\ \frac{1}{2}y_k + \frac{1}{2}y_{k+1}\right)$

BDF-2

$y_{k+1} = \frac{4}{3}y_k - \frac{1}{3}y_{k-1} + \frac{2}{3}hf(x_{k+1}, y_{k+1})$

BDF-3

$y_{k+1} = \frac{18}{11}y_k - \frac{9}{11}y_{k-1} + \frac{2}{11}y_{k-2} + \frac{6}{11}hf(x_{k+1}, y_{k+1})$

BDF-4

$y_{k+1} = \frac{48}{25}y_k - \frac{36}{25}y_{k-1} + \frac{16}{25}y_{k-2} - \frac{3}{25}y_{k-3} + \frac{12}{25}hf(x_{k+1}, y_{k+1})$

Klassinen Runge-Kutta

$$\begin{aligned} y_{k+1} &= y_k + \tfrac{h}{6}\big(V_1 + 2V_2 + 2V_3 + V_4\big), \\ &\quad V_1 = f(x_k, y_k), \quad V_2 = f\left(x_k + \tfrac{1}{2}h, y_k + \tfrac{1}{2}hV_1\right) \\ &\quad V_3 = f\left(x_k + \tfrac{1}{2}h, y_k + \tfrac{1}{2}hV_2\right), \quad V_4 = f\left(x_k + h, y_k + hV_3\right) \end{aligned}$$

Näistä implisiittinen (samoin kuin tavallinen) Eulerin menetelmä on *ensimmäisen kertaluvun* menetelmä, eli virhe on suotuisissa oloissa luokkaa $\mathcal{O}(h)$. Trapetsi, keskipistesääntö ja BDF-2 ovat *toisen kertaluvun* menetelmiä (virhe $\mathcal{O}(h^2)$ riittävän säännöllisille ratkaisuille), BDF-3 on kertalukua 3 ja BDF-4 ja klassinen Runge-Kutta molemmat kertalukua 4. Esimerkeistä BDF-2, BDF-3 ja BDF-4 edustavat moniaskelmenetelmiä (askelpituus oletettu vakioksi). Näiden käynnistämiseksi on laskettava ensin erillisellä algoritmilla $y_1 \ldots y_{m-1}$ (BDF-m). — Nimi

BDF on lyhenne sanoista Backward Differentiation Formula (ks. Harj.Teht. 5).

Kaikissa esimerkeissä, viimeistä lukuun ottamatta, on kyse *implisiittisestä* menetelmästä. Tällä tarkoitetaan, että laskettava luku y_{k+1} esiintyy laskukaavassa myös funktion f 'alla'. Tällöin lukua ei saada selville pelkästään funktioevaluaatioilla, vaan on ratkaistava yhtälö (systeemin tapauksessa yhtälöryhmä) muotoa $y_{k+1} = F(y_{k+1})$. Käytännössä tämä ratkaistaan likimäärin, esimerkiksi muutamalla (1–3) Newtonin iteraatiolla tai yksinkertaisemmin kiintopisteiteraatiolla:

$$t_{j+1} = F(t_j), \quad j = 0 \ldots m\,; \quad y_{k+1} = t_{m+1}.$$

(Vrt. Luku V.7 — yhtälöryhmien iteratiivisista ratkaisumenetelmistä tulee puhe myöhemmin.) Iteraation alkuarvoksi voidaan ottaa esim. Eulerin menetelmän antama $t_0 = y_k + hf(x_k, y_k)$.

Implisiittiset menetelmät kuten implisiittinen Euler ja BDF-m ovat differentiaaliyhtälöiden 'black box'-ratkaisijoina hyvin suosittuja, syystä että ne ovat vaihtelevissa tilanteissa suhteellisen varmoja. — Todettakoon tässä ainoastaan, että varmuus tulee esiin erityisesti sellaisissa nk. *kankeissa* (engl. stiff) tehtävissä, joissa ratkaisu sisältää nopeasti (eksponentiaalisesti) vaimenevia transientteja. Tällaisessa tehtävässä esimerkiksi implisiittinen Euler on tavallista Eulerin menetelmää huomattavasti varmempi vaihtoehto, vaikka menetelmien muodollisessa tarkkuudessa ei ole eroa (ks. Harj.teht. 6). Sen sijaan tehtävissä, joissa mainitun tyyppisiä erityisongelmia ei esiinny, ovat yksinkertaisemmat *eksplisiittiset* menetelmät (kuten Euler ja klassinen Runge-Kutta) edelleen varteenotettavia vaihtoehtoja ja myös käytössä.

HARJOITUSTEHTÄVIÄ

1. Laske seuraavissa alkuarvotehtävissä Eulerin menetelmällä likiarvot luvuille $y(x_k)$, $k = 1 \ldots 5$, missä $x_k = k/5$ (askelpituus $h = 0.2$). Vertaa tarkkaan ratkaisuun!

   a) $\begin{cases} y' = -y^2, \\ y(0) = 1 \end{cases}$ b) $\begin{cases} y' = y^2, \\ y(0) = 1 \end{cases}$ c) $\begin{cases} y' = -2xy^2, \\ y(0) = 1 \end{cases}$ d) $\begin{cases} y' = 2xy^2, \\ y(0) = 1 \end{cases}$

2. Alkuarvotehtävä

$$y(0) = 1, \quad y' = \frac{x}{\sqrt{y^2+1}} + \frac{2x^2}{x^2+1}$$

   ratkaistaan numeerisesti välillä $[0, 10]$ käyttäen Eulerin menetelmää ja vakioaskelpituutta h. Arvioi vakiot L ja M Eulerin menetelmän virhearviossa ko. välillä. Millä askelpituuksilla voidaan taata, että numeerisen ratkaisun suhteellinen virhe ei ylitä yhdessäkään laskentapisteessä arvoa 10^{-4}?

3. Alkuarvotehtävä $y' = -1/y,\ y(0) = 3$ ratkaistaan numeerisesti välillä $[0, 1]$ käyttäen askelpituutta $h = 0.2$. Laske seuraavilla menetelmillä approksimaatio y_5 luvulle $y(1)$: Euler, Trapetsi, BDF-2, BDF-3, Klassinen Runge-Kutta. Vertaa tarkkaan arvoon!

4. Alkuarvotehtävässä $y' = f(x),\ y(0) = 0$ klassinen Runge-Kutta-menetelmä pelkistyy erääksi tunnetuksi numeerisen integroinnin säännöksi. Mistä säännöstä on kyse?

5. Differentiaaliyhtälön $y' = f(x, y)$ ratkaisumenetelmä BDF-m johdetaan kirjoittamalla yhtälö pisteessä $x = x_{k+1}$, tekemällä approksimaatio
$$y'(x_{k+1}) \approx \frac{1}{h}\sum_{j=0}^{m} \alpha_j y_{k+1-j}, \quad y_i = y(x_i)$$
ja vaatimalla, että tämä on tarkka polynomeille astetta $\leq m$. Tarkista laskemalla kertoimet α_j, että tuloksena on todella menetelmä BDF-2, kun $m = 2$. Millainen menetelmä on BDF-1 ?

6. Olkoon $\tau = 0.02$. Halutaan ratkaista numeerisesti alkuarvotehtävä
$$\begin{cases} \tau y' + y = \sin x, & x > 0, \\ y(0) = 0.1. \end{cases}$$
Vertaa tavallisen ja implisiittisen Eulerin menetelmän antamia ratkaisuja laskemalla molemmilla menetelmillä $y_k \approx y(kh),\ k = 1 \ldots 5$ askelpituuksilla $h = 0.01$, $h = 0.04$ ja $h = 0.08$. Hahmottele kaikki ratkaisut graafisesti ja vertaa myös tarkkaan ratkaisuun! (Ks. sovellusesimerkki Luvussa IX.4.)

7. Tasossa liikkuvaan avaruusalukseen vaikuttaa origossa olevan maapallon vetovoima $\vec{F} = -(k/r^2)\,\vec{e}_r$, missä k on vakio (napakoordinaatisto). Hetkellä t alus on paikassa $(x(t), y(t))$ ja sen nopeus on $u(t)\vec{i} + v(t)\vec{j}$. Kirjoita aluksen liikeyhtälö $m\vec{r}'' = \vec{F}$ karteesisessa koordinaatistossa normaalimuotoiseksi differentiaaliyhtälösysteemiksi ja esitä Eulerin menetelmään perustuva algoritmi yhtälöiden numeeriseksi ratkaisemiseksi. Oletetaan tunnetuksi alkuarvot x_0, y_0, u_0, v_0 hetkellä $t = 0$.

8. (*) Ratkaistaessa Eulerin menetelmällä alkuarvotehtävää
$$\begin{cases} y' = \dfrac{y^2}{x + y + 1} + 4x + 2, & x > 0, \\ y(0) = A \end{cases}$$
askelpituudella $h = 0.1$ havaitaan yllättäen, että saadaan tarkka ratkaisu, ts. $y_k = y(x_k)\ \forall k\ (x_k = kh)$. Mikä on A:n arvo?

IX.8 *Picardin–Lindelöfin lause

Palataan tutkimaan ensimmäisen kertaluvun alkuarvotehtävää

$$\begin{cases} y' = f(x,y), \\ y(x_0) = y_0 \end{cases} \tag{1}$$

tällä kertaa teoreettisemmalta kannalta. Asetetaan kysymys: Millä, funktiolle f ja luvuille x_0, y_0, asetettavilla ehdoilla voidaan taata, että alkuarvotehtävällä on yksikäsitteinen ratkaisu pisteen x_0 ympäristössä? Vastauksen antaa seuraava kuuluisa lause.

*Lause IX.8.1 **(Picard–Lindelöf†)** Olkoon $T = [a,b] \times [c,d] \subset \mathbb{R}^2$ ja oletetaan, että $f : T \to \mathbb{R}$ toteuttaa ehdot

(i) Funktio $x \mapsto f(x,y)$ on jatkuva välillä $[a,b]$ jokaisella $y \in [c,d]$.

(ii) Funktio $y \mapsto f(x,y)$ on jokaisella $x \in [a,b]$ Lipschitz-jatkuva välillä $[c,d]$ vakiolla, joka ei riipu x:stä, ts. $\exists L \in \mathbb{R}_+$ siten, että pätee

$$|f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2| \quad \forall (x,y_1) \in T,\ (x,y_2) \in T.$$

Tällöin, jos $(x_0,y_0) \in (a,b) \times (c,d)$, niin on olemassa $\delta > 0$ siten, että alkuarvotehtävällä

$$\begin{cases} y' = f(x,y), \quad x \in (x_0 - \delta, x_0 + \delta), \\ y(x_0) = y_0 \end{cases}$$

on yksikäsitteinen ratkaisu. Edelleen jos oletukset ovat voimassa, kun välin $[c,d]$ tilalla on $\mathbb{R}$ ($T = [a,b] \times \mathbb{R}$), niin ratkaisu on olemassa ja yksikäsitteinen koko avoimella välillä (a,b), kun $(x_0,y_0) \in (a,b) \times \mathbb{R}$.

Seuraavassa käydään lyhyesti läpi Lauseen IX.8.1 todistukset ideat ja päävaiheet. (Todistuksen tarkkaan läpiviemiseen ovat käsitteelliset eväämme hieman vajavaiset.) Lähtökohtana on ajatus, että alkuarvotehtävän ratkaisu konstruoidaan nk. *Picardin iteraatiolla.* Tässä puolestaan idea on seuraava: Kun ollaan lähellä pistettä x_0, niin voidaan olettaa, että $y(x)$ on lähellä y_0:aa. Ratkaisun ensimmäinen approksimaatio on tämän mukaan vakio:

$$y(x) \approx y_0(x) = y_0.$$

†Picardin–Lindelöfin lauseen alkuperäisversion esitti ja todisti tässä esitettyä yleisemmässä muodossa ranskalainen matemaatikko **Emile Picard** (1853-1941) vuonna 1893. Suomalainen matemaatikko **Ernst Lindelöf** (1870-1946) tarkensi tulosta hiukan myöhemmin. Lauseen väittämän jälkimmäinen osa (koskien ratkeavuutta koko tarkasteltavalla välillä) on Lindelöfin käsialaa.

Parannetaan tätä nyt asteittain ratkaisemalla palautuvasti alkuarvotehtävät

$$\begin{cases} y_k'(x) = f(x, y_{k-1}(x)), \\ y_k(x_0) = y_0, \quad k = 1, 2, \ldots \end{cases}$$

Koska tässä y_{k-1} tunnetaan aikaisemmilta iteraatiokierroksilta (tai alkuehdosta), saadaan jokainen y_k määrätyksi palautuvasti tunnetun funktion integraalina:

$$y_0(x) = y_0, \quad y_k(x) = y_0 + \int_{x_0}^{x} f(t, y_{k-1}(t))dt, \quad k = 1, 2, \ldots \tag{2}$$

Näin on määritelty Picardin iteraatio. — Huomattakoon, että tässä on itse asiassa kyse <u>kiintopisteiteraatiosta</u>, missä 'pisteen' sijasta etsitään funktiota (vrt. Luku V.7). Iteraatiota sovelletaan alkuarvotehtävän integroituun muotoon

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt,$$

eli Picardin iteraatiossa on kyse tämän *integraaliyhtälön* iteratiivisesta ratkaisemisesta. — Katkaisemalla iteraatio saadaan likimääräinen ratkaisu, kuten kiintopisteiteraatioissa yleensä.

Picardin iteraatiota voi siis pitää alkuarvotehtävän likimääräisen ratkaisemisen menetelmänä. Toisin kuin edellisessä luvussa tarkastellut numeeriset menetelmät, Picardin iteraatio on kuitenkin lähinnä 'ajattelumenetelmä', sillä käytännössä iteraatioon sisältyvät integraalit tulevat yleensä nopeasti hankaliksi.

Esimerkki 1 Laske kaksi ensimmäistä Picardin iteraattia ($y_k(x)$, $k = 1, 2$) alkuarvotehtävälle

$$\begin{cases} y' = x + y^2, \\ y(0) = 1. \end{cases}$$

<u>Ratkaisu</u> Tässä on $f(x, y) = x + y^2$, $x_0 = 0$, $y_0 = 1$, joten $y_0(x) = 1$ ja

$$\begin{aligned} y_1(x) &= 1 + \int_0^x (t+1)\, dt = \underline{\underline{1 + x + \frac{1}{2}x^2}}, \\ y_2(x) &= 1 + \int_0^x \left[t + \left(1 + t + \frac{1}{2}t^2\right)^2 \right] dt \\ &= \underline{\underline{1 + x + \frac{3}{2}x^2 + \frac{2}{3}x^3 + \frac{1}{4}x^4 + \frac{1}{20}x^5}}. \end{aligned}$$

Seuraavaa iteraatti $y_3(x)$ olisi jo polynomi astetta 11. □

Palataan Lauseen IX.8.1 todistukseen, jossa siis ideana on konstruoida alkuarvotehtävän ratkaisu Picardin iteraatiolla (2). Tarkoituksena on tällöin osoittaa, että iteraatio suppenee kohti välillä $[x_0 - \delta, x_0 + \delta]$ jatkuvaa funktiota y, joka toteuttaa

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt, \quad x \in [x_0 - \delta, x_0 + \delta]. \tag{3}$$

Parametrin δ valinnalla (δ riittävän pieni) taataan, että $[x_0 - \delta, x_0 + \delta] \subset [a, b]$ ja että $y(x)$ ei 'karkaa' väliltä $[c, d]$, ts. $(x, y(x)) \in T \quad \forall x \in [x_0 - \delta, x_0 + \delta]$ (vrt. vastaava ajatus Lauseen IX.2.1 todistuksessa). Tällöin oletuksen (i) mukaan yhdistetty funktio $x \mapsto f(x, y(x))$ on jatkuva välillä $[x_0 - \delta, x_0 + \delta]$, jolloin yhtälöstä (3) ja Analyysin peruslauseesta seuraa, että y on jatkuvasti derivoituva välillä $[x_0 - \delta, x_0 + \delta]$. Tällöin y on myös alkuarvotehtävän (1) ratkaisu välillä $(x_0 - \delta, x_0 + \delta)$.

Iteraation (2) suppenemisen toteamiseksi otetaan käyttöön välillä $[x_0 - \delta, x_0 + \delta]$ määriteltyä jatkuvaa funktiota mittaava *maksiminormi*

$$\|y\| = \max \{ \, |y(x)| \mid x_0 - \delta \leq x \leq x_0 + \delta \, \}.$$

Tälle (niinkuin normeille yleensä, vrt. Luku II.4) pätee <u>kolmioepäyhtälö</u>

$$\|y_1 + y_2\| \leq \|y_1\| + \|y_2\|.$$

Lähdetään nyt Lauseen IX.8.1 oletuksesta (ii) ja arvioidaan

$$\begin{aligned} |y_{k+1}(x) - y_k(x)| &= \left| \int_{x_0}^{x} [f(t, y_k(t)) - f(t, y_{k-1}(t))] \, dt \right| \\ &\leq \left| \int_{x_0}^{x} L|y_k(t) - y_{k-1}(t)| \, dt \right| \\ &\leq L|x - x_0| \|y_k - y_{k-1}\| \\ &\leq L\delta \, \|y_k - y_{k-1}\|, \quad x \in [x_0 - \delta, x_0 + \delta], \quad k = 1, 2, \ldots \end{aligned}$$

Tulos on kirjoitettavissa

$$\|y_{k+1} - y_k\| \;\leq\; L\delta \, \|y_k - y_{k-1}\|, \quad k = 1, 2, \ldots$$

Merkitään nyt $q = L\delta$, ja oletetaan jatkossa, että $q < 1$ (järjestettävissä valitsemalla $\delta < 1/L$). Käyttämällä saatua arviota palautuvasti seuraa

$$\|y_{k+1} - y_k\| \;\leq\; q^k \|y_1 - y_0\|, \quad k = 1, 2, \ldots$$

Oletuksen (i) mukaan funtkio $g(x) = f(x, y_0)$ on jatkuva välillä $[a, b]$, joten iteraatiokaavan (2) perusteella

$$\|y_1 - y_0\| = \max_{x\in[x_0-\delta,x_0+\delta]} \left| \int_{x_0}^{x} f(t, y_0)\, dt \right| \le C\delta,$$

missä C on $|g|$:n maksimiarvo välillä $[a, b]$. Yhdistämällä saadut arviot ja käyttämällä kolmioepäyhtälöä seuraa

$$\begin{aligned} \|y_n - y_k\| &= \|(y_n - y_{n-1}) + (y_{n-1} - y_{n-2}) + \ldots + (y_{k+1} - y_k)\| \\ &\le \|y_n - y_{n-1}\| + \|y_{n-1} - y_{n-2}\| + \ldots + \|y_{k+1} - y_k\| \\ &\le (q^{n-1} + q^{n-2} + \ldots + q^k)\, C\delta \\ &< C\delta q^k \sum_{i=0}^{\infty} q^i = \frac{C\delta}{1-q}\, q^k, \quad 0 \le k < n. \end{aligned}$$

Tästä voidaan päätellä ensinnäkin, että jos δ on riittävän pieni, niin $\|y_n - y_0\| \le C\delta/(1-\delta) \le \min\{y_0 - c,\ d - y_0\}$ $\forall n$, jolloin $y_n(x)$:n pako väliltä $[c, d]$ on estetty. Toiseksi päätellään, että jono $\{y_k(x)\}$ on Cauchy jokaisella $x \in [x_0 - \delta, x_0 + \delta]$ ja siis suppenee:

$$y_k(x) \to y(x) \in [c, d], \quad x \in [x_0 - \delta, x_0 + \delta].$$

On edelleen pääteltävissä, että tämä suppeneminen on itse asiassa *tasaista* välillä $[x_0 - \delta, x_0 + \delta]$, ts. pätee

$$\|y_k - y\| = \max_{x\in[x_0-\delta,\, x_0+\delta]} |y_k(x) - y(x)| \to 0, \quad \text{kun } k \to \infty,$$

ja että raja-arvona konstruoitu $y(x)$ on jatkuva välillä $[x_0 - \delta, x_0 + \delta]$. (Tässä on todistuksen käsitteellisesti vaativin kohta!) Ottamalla lopuksi iteraatiokaavassa (2) puolittain raja-arvo, kun $k \to \infty$, päätellään, että $y(x)$ ratkaisee probleeman (3), jolloin Lauseen IX.8.1 ensimmäisestä väittämästä on todistettu ratkaisun olemassaoloa koskeva osa. Ratkaisun yksikäsitteisyys on todettavissa suoraviivaisemmin, ks. Harj.teht. 6.

Lauseen IX.8.1 toista väittämää todistettaessa ei jonon $\{y_k(x)\}$ pako-ongelmaa ole, joten väli $[x_0 - \delta, x_0 + \delta]$ voidaan korvata välillä $[a, b]$. Tarkentamalla em. laskua osoittautuu, että voidaan myös arvioida

$$|y_{k+1}(x) - y_k(x)| \le C\, \frac{L^k |x - x_0|^k}{k!}, \quad x \in [a, b].$$

Tutkittaessa jonon $\{y_k(x)\}$ suppenemista tulee siis vertailukohdaksi nyt geometrisen sarjan sijasta eksponenttifunktion sarja

$$e^{L|x-x_0|} = \sum_{k=0}^{\infty} \frac{L^k |x - x_0|^k}{k!}.$$

Tämä suppenee kaikkialla, joten voidaan päätellä samalla tavoin kuin edellä, että funktio $y(x) = \lim_k y_k(x)$ on integraaliyhtälön (3) ratkaisu, tällä kertaa koko välillä $[a, b]$. Alkuarvotehtävä (1) on näin ratkaistu yksikäsitteisesti välillä (a, b).

Esimerkki 2 Alkuarvotehtävä

$$\begin{cases} y' = \sin y, \\ y(x_0) = y_0 \end{cases}$$

toteuttaa Lauseen IX.8.1 ehdot, kun $T = \mathbb{R}^2$, sillä $f(y) = \sin y$ on kaikilla väleillä Lipschitz-jatkuva vakiolla $L = 1$. Alkuarvotehtävällä on siis yksikäsitteinen ratkaisu välillä $(-\infty, \infty)$ jokaisella $(x_0, y_0) \in \mathbb{R}^2$. Jos valitaan esim. $x_0 = 0$, niin nähdään, että y_0:n eri arvoja vastaavat ratkaisut löytyvät kaikki Esimerkissä IX.2:3 laskettujen ratkaisujen joukosta. Näin on tullut varmistetuksi, että mainitussa esimerkissä löydettiin kaikki differentiaaliyhtälön $y' = \sin y$ ratkaisut. □

Esimerkki 3 Alkuarvotehtävälle

$$\begin{cases} y' = \sqrt{|y|}, \\ y(0) = 0 \end{cases}$$

Lauseen IX.8.1 ehdot eivät toteudu millään joukon $T = (a, b) \times (c, d)$ valinnalla, sillä on oltava $0 \in (a, b)$, jolloin $f(y) = \sqrt{|y|}$ ei ole Lipschitz-jatkuva välillä $[a, b]$. Ratkaisu ei olekaan yksikäsitteinen (ks. Esimerkki IX.1: 3b). □

Esimerkki 4 Jos lineaarisessa alkuarvotehtävässä

$$\begin{cases} y' + P(x)y = R(x), \\ y(x_0) = y_0 \end{cases}$$

P ja Q ovat jatkuvia välillä $[a, b]$, niin Lauseen IX.8.1 ehdot ovat voimassa, kun valitaan $T = [a, b] \times \mathbb{R}$. Nimittäin

(i) Funktio $x \mapsto f(x, y) = -P(x)y + R(x)$ on jatkuva välillä $[a, b]$ jokaisella $y \in \mathbb{R}$.

(ii) $|f(x, y_1) - f(x, y_2)| = |P(x)||y_1 - y_2| \leq L|y_1 - y_2|, \quad (x, y_1),\ (x, y_2) \in T,$ missä $L = \max\limits_{x \in [a,b]} |P(x)|$.

Alkuarvotehtävä siis ratkeaa yksikäsitteisesti välillä (a, b), kun $x_0 \in (a, b)$. (Sama pääteltiin Luvussa IX.4 toisin keinoin.) □

Differentiaaliyhtälösysteemin ratkeavuus

Picardin-Lindelöfin lauseen todellinen hienous piilee siinä, että lauseen väittämä voidaan yleistää hyvin suoraviivaisesti koskemaan yleistä normaalimuotoista differentiaaliyhtälösysteemiä (vrt. Luku IX.1)

$$\begin{cases} y_1'(x) = f_1(x, y_1, \ldots, y_n), \\ y_2'(x) = f_2(x, y_1, \ldots, y_n), \\ \qquad \vdots \\ y_n'(x) = f_n(x, y_1 \ldots, y_n). \end{cases}$$

Alkuehdoiksi oletetaan

$$y_i(x_0) = A_i, \quad i = 1 \ldots n,$$

missä $x_0 \in (a, b)$. Tarkastellaan esimerkkinä alkuarvotehtävää kokoa $n = 2$:

$$\begin{cases} y' = f_1(x, y, u), \\ u' = f_2(x, y, u), \quad x \in (a, b), \\ y(x_0) = y_0, \ u(x_0) = u_0. \end{cases} \tag{4}$$

Picardin-Lindelöfin lauseen muotoilussa tälle ongelmalle asetetaan

$$T = [a, b] \times [c_1, d_1] \times [c_2, d_2] \subset \mathbb{R}^3 \quad \text{tai} \quad T = [a, b] \times \mathbb{R}^2 \subset \mathbb{R}^3$$

ja oletetaan

(i) Funktiot $x \mapsto f_1(x, y, u)$ ja $x \mapsto f_2(x, y, u)$ ovat jatkuvia välillä $[a, b]$ jokaisella $(y, u) \in [c_1, d_1] \times [c_2, d_2]$ tai jokaisella $(y, u) \in \mathbb{R}^2$.

(ii) f_1 ja f_2 toteuttavat muuttujien y, u suhteen Lipschitz-ehdon: Jollakin $L \in \mathbb{R}_+$ pätee

$$|f_i(x, y_1, u_1) - f_i(x, y_2, u_2)| \le L \max\{|y_1 - y_2|, |u_1 - u_2|\}, \quad i = 1, 2$$
$$\forall (x, y_1, u_1) \in T, \ (x, y_2, u_2) \in T.$$

Näillä oletuksilla päädytään vastaavaan väittämään kuin edellä, ts. alkuarvotehtävä (4) ratkeaa yksikäsitteisesti, joko x_0:n lähellä välillä $(x_0 - \delta, x_0 + \delta)$, tai vahvemmin oletuksin koko välillä (a, b). Myös todistus on perusidealtaan sama kuin edellä: Lähtökohtana on ratkaisun konstruoiminen Picardin iteraatiolla

$$\begin{cases} y_k(x) = y_0 + \displaystyle\int_{x_0}^{x} f_1(t, y_{k-1}(t), u_{k-1}(t))\, dt, \\ u_k(x) = u_0 + \displaystyle\int_{x_0}^{x} f_2(t, y_{k-1}(t), u_{k-1}(t))\, dt, \quad k = 1, 2, \ldots \end{cases}$$

Palataan lopuksi Luvussa IX.6 tarkasteltuun toisen kertaluvun lineaariseen ja homogeeniseen alkuarvotehtävään

$$\begin{cases} y'' + P(x)y' + Q(x)y = 0, \\ y(0) = A, \ y'(0) = B. \end{cases} \tag{5}$$

Kun tässä kirjoitetaan $u = y'$, niin päästään systeemimuotoon (4), missä

$$f_1(x, y, u) = u, \quad f_2(x, y, u) = -Q(x)y - P(x)u.$$

Jos P ja Q ovat jatkuvia välillä (a, b) (kuten oletettiin Luvussa IX.6), niin $x \mapsto f_1(x, y, u)$ ja $x \mapsto f_2(x, y, u)$ ovat jatkuvia välillä $[a_1, b_1]$ jokaisella $(y, u) \in \mathbb{R}^2$, kun $[a_1, b_1] \subset (a, b)$. Koska

$$\begin{aligned} f_1(x, y_1, u_1) - f_1(x, y_2, u_2) &= u_1 - u_2, \\ f_2(x, y_1, u_1) - f_2(x, y_2, u_2) &= -Q(x)(y_1 - y_2) - P(x)(u_1 - u_2), \end{aligned}$$

niin nähdään, että myös ehto (ii) toteutuu, kun asetetaan

$$L = \max\{1, L_1\}, \quad L_1 = \max_{x \in [a_1, b_1]} \big\{ |P(x)| + |Q(x)| \big\}.$$

Näin ollen Picardin–Lindelöfin lauseen (vahvemman väittämän) mukaan alkuarvotehtävä (5) ratkeaa yksikäsitteisesti välillä (a_1, b_1). Koska tämä on totta aina kun $[a_1, b_1] \subset (a, b)$, niin tehtävällä on yksikäsitteinen ratkaisu koko avoimella välillä (a, b). Lause IX.6.2 on näin tullut todistetuksi Picardin–Lindelöfin lauseen seuraamuksena.

HARJOITUSTEHTÄVIÄ

1. Mitä voidaan Picardin–Lindelöfin lauseen valossa sanoa seuraavien alkuarvotehtävien ratkeavuudesta ja ratkaisujen yksikäsitteisyydestä $\mathbb{R}$:ssä tai $\mathbb{R}$:n osaväleillä?

   a) $y' = \cos y + \dfrac{x}{y^2 + 1}, \ \ y(0) = A \in \mathbb{R}$

   b) $y' = \dfrac{y}{\cosh x} + e^{x - y^2}, \ \ y(0) = A \in \mathbb{R}$

   c) $y' = \dfrac{y}{\cos x} + e^{x - y^2}, \ \ y(0) = A \in \mathbb{R}$

   d) $y' = 8xy + \dfrac{\sin(xy)}{x^2 + 2x - 35}, \ \ y(0) = A \in \mathbb{R}$

   e) $y'' = \cos(xy), \ \ y(0) = A \in \mathbb{R}, \ y'(0) = B \in \mathbb{R}$

2. Määritä seuraaville alkuarvotehtäville kaikki alkupisteen lähiympäristössä pätevät ratkaisut. Vertaa Picardin–Lindelöfin lauseen väittämään.

   a) $y' = y^2,\ y(0) = 1$ b) $y' = y^2,\ y(0) = 1000$
   c) $y' = y\sqrt{|y|},\ y(0) = 1$ d) $y' = \sqrt[4]{|y|},\ y(0) = 1$
   e) $y' = \sqrt[4]{|y|},\ y(0) = 0.0001$ f) $y' = \sqrt[4]{|y|},\ y(0) = 0$

3. Halutaan laskea Neperin luku e lähtien tiedosta, että $e = y(1)$, missä $y(x)$ ratkaisee alkuarvotehtävän $y' = y,\ y(0) = 1$. Millainen laskukaava saadaan luvulle e, kun tehtävä ratkaistaan Picardin iteraatiolla?

4. Laske seuraavissa alkuarvotehtävissä Picardin iteraatit $y_1(x)$ ja $y_2(x)$ ja vertaa tarkkaan ratkaisuun. (Vrt. Harj.teht. IX.7:1.)

   a) $\begin{cases} y' = -y^2, \\ y(0) = 1 \end{cases}$ b) $\begin{cases} y' = y^2, \\ y(0) = 1 \end{cases}$ c) $\begin{cases} y' = -2xy^2, \\ y(0) = 1 \end{cases}$ d) $\begin{cases} y' = 2xy^2, \\ y(0) = 1 \end{cases}$

5. Laske alkuarvotehtävälle

   a) $\begin{cases} y'' = -y, \\ y(0) = 1,\ y'(0) = 0 \end{cases}$ b) $\begin{cases} y'' = -y, \\ y(0) = 0,\ y'(0) = 1 \end{cases}$

   likimääräinen ratkaisu iteroimalla $2n$ kertaa ($n \in \mathbb{N}$) Picardin iteraatiolla tehtävän systeemimuodosta.

6. (*) Näytä, että Lauseen IX.8.1 oletuksin alkuarvotehtävän (1) ratkaisu (jos olemassa) on yksikäsitteinen välillä $(x_0 - \delta, x_0 + \delta)$, kun δ on riittävän pieni. *Vihje*: Oleta $y_1(x)$ ja $y_2(x)$ ratkaisuiksi, lähde tehtävän integraalimuodosta (3), käytä oletusta (ii) ja valitse $\delta < 1/L$.

Luku X

Matriisit

> ”Jos yhtälöryhmässä on yli 50 tuntematonta, on ratkaisemisessa syytä käyttää tietokonetta”.[†]

Monilla insinöörialoilla on ollut pitkä *lineaaristen yhtälöryhmien* ratkaisemisen perinne jo ennen tietokoneiden aikaa. Myös lineaaristen yhtälöryhmien matematiikassa, *lineaarialgebrassa*, perinteet ovat pitkät. Ne ulottuvat 1700-luvulle, jolloin kehitettiin *determinantin* käsitteeseen perustuva yhtälöryhmien ratkeavuusteoria ja ratkaisukaavat. Determinanteilla on vielä nykyäänkin käyttöä teoreettisissa tarkasteluissa, erinäisissä laskukaavoissa (vrt. Luku II.6) ja yleisemminkin käsinlaskussa silloin, kun yhtälöryhmän koko on pieni. Muuten determinanttioppia on pidettävä vanhentuneena, syystä että tämän opin mukaiset laskukaavat eivät sovellu laajamittaiseen numeeriseen käyttöön. Tietokoneiden aikakaudella lineaarialgebra onkin pitkälti irtaantunut determinanttiperinteestä ja vanhoista käsinlaskun menetelmistä.

Tässä luvussa tutkimuksen kohteena on lineaarialgebran avainkäsite, *matriisi*. Matriiseilla laskemista eli *matriisialgebraa* tarkastellaan ensin Luvussa X.1. Tämän jälkeen Luvuissa X.2–X.4 tutkitaan lineaarisia yhtälöryhmiä, niiden ratkaisemista *Gaussin algoritmilla* ja ratkaisemiseen keskeisesti liittyvää *käänteismatriisin* käsitettä. Luvussa X.5 käydään lyhyesti läpi vanha determinanttioppi laskukaavoineen. Luvuissa X.6—X.7 ovat tutkimuskohteena matriisialgebraan perustuvat *lineaarikuvaukset* ja *affiinikuvaukset* sekä näiden geometriset sovellutukset, *geometriset kuvaukset*. Viimeisessä osaluvussa esitellään joitakin lineaaristen yhtälöryhmien sovellusesimerkkejä perinteisestä insinöörimatematiikasta.

[†]Lentokoneen siiven lujuuslaskentaan perehtyneen insinöörin toteamus alan julkaisussa 1950-luvulla.

X.1 Matriisialgebra

Yleinen m yhtälön ja n tuntemattoman *lineaarinen yhtälöryhmä* (-systeemi) on muotoa

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \quad\vdots \qquad\quad \vdots \qquad\qquad \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m. \end{cases} \tag{1}$$

Tässä $m, n \in \mathbb{N}$, a_{ij}:t ovat yhtälöryhmän *kertoimet*, ja myös luvut b_i, $i = 1 \ldots m$, jotka muodostavat yhtälöryhmän nk. 'oikean puolen' (tai 'vakiotermin'), oletetaan tunnetuiksi. Tuntemattomia ovat siis x_i, $i = 1 \ldots n$, ja ongelmana yleensä näiden löytäminen, eli yhtälöryhmän *ratkaiseminen*. Jatkossa oletetaan pääsääntöisesti, että $a_{ij}, x_i, b_i \in \mathbb{R}$. Yleisemmin voitaisiin olettaa, että $a_{ij}, x_i, b_i \in \mathbb{K}$, missä $\mathbb{K}$ on mikä tahansa kunta, sillä yhtälöryhmää kirjoitettaessa ja ratkaistaessa tarvitaan vain kunnan laskuoperaatioita. Esimerkiksi voisi olla $\mathbb{K} = \mathbb{C}$ tai (kuten yksinkertaisissa esimerkeissä usein) $\mathbb{K} = \mathbb{Q}$.

Yhtälöryhmän (1) sanotaan olevan *kokoa* $m \times n$ ('m kertaa n'). Yhtälöryhmällä ei välttämättä ole ratkaisua lainkaan, tai ratkaisu voi olla monikäsitteinen. Jos yhtälöryhmällä on yksikäsitteinen ratkaisu, valittiinpa oikealla puolella olevat luvut b_i miten tahansa, niin sanotaan, että yhtälöryhmä on *säännöllinen*. Muulloin, eli jos ratkaisua ei ole jollakin (b_i) tai ratkaisu on monikäsitteinen, yhtälöryhmä on epäsäännöllinen eli *singulaarinen*. Yhtälöryhmän säännöllisyys tai singulaarisuus ei siis riipu luvuista b_i vaan on ainoastaan kerrointaulukon (a_{ij}) ominaisuus. Tullaan mäkemään, että ehto $m = n$ on välttämätön ehto säännöllisyydelle, ts. säännöllisessä ryhmässä on yhtälöiden lukumäärän (m) ja tuntemattomien lukumäärän (n) täsmättävä. Tapauksessa $m > n$ sanotaan yhtälöryhmää (1) *ylimääräytyväksi* (engl. overdetermined), tapauksessa $m < n$ *alimääräytyväksi* (engl. underdetermined). Nämä ovat siis aina singulaarisia systeemejä.

Esimerkki 1 Singulaarisia systeemejä tyyppiä $m = n = 2$ ovat esimerkiksi

$$\begin{cases} x_1 + x_2 = b_1, \\ x_1 + x_2 = b_2 \end{cases} \quad \text{ja} \quad \begin{cases} x_1 + 2x_2 = b_1, \\ 2x_1 + 4x_2 = b_2. \end{cases} \qquad \square$$

Jatkossa suoritettavien algebrallisten tarkastelujen lähtökohtana on yleinen lineaarinen yhtälöryhmä (1). Kirjoitetaan yhtälöryhmä ensinnäkin taulukkomuotoon (1. abstraktiovaihe)

$$\begin{bmatrix} a_{11}x_1 + & \ldots & + a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 + & \ldots & + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

eli

$$\mathbf{y} = \mathbf{b}, \tag{2}$$

missä

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Tässä $\mathbf{b}$, $\mathbf{y}$ ovat nk. *pystyvektoreita* (engl. column vector). Jatkossa käytetään pystyvektoreiden symboleina lihavoituja pieniä kirjaimia $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}$, jne.

Ym. pystyvektoreissa $\mathbf{y}$, $\mathbf{b}$ on m vektorin *koko* (tai tyyppi), mikä voidaan ilmaista myös termillä *m-vektori*. Luvut b_i, y_i ovat ko. vektoreiden *alkiot* (tai komponentit). Voidaan myös käyttää merkintöjä

$$[\mathbf{b}]_i = b_i, \quad \mathbf{b} = (b_i), \quad \mathbf{b} = (b_i)_{i=1}^{m}.$$

Termi 'vektori' viittaa siihen, että pystyvektoreille voidaan määritellä vektoreiden peruslaskuoperaatiot, eli vektorien yhteenlasku ja skalaarilla kertominen. Määritelmät ovat

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}, \qquad \lambda \mathbf{x} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} \qquad (\lambda \in \mathbb{R}).$$

Yhteenlaskussa vektoreiden $\mathbf{x}$, $\mathbf{y}$ on oltava <u>samaa</u> <u>kokoa</u>, muuten yhteenlaskua ei voi määritellä. Laskuoperaatioissa siis lasketaan yhteen ja kerrotaan alkioittain, kuten aiemmin tehtiin Luvuissa II.2 ja II.6 käsiteltäessä $\mathbb{R}^2$:n lukupareja ja $\mathbb{R}^3$:n lukukolmikkoja vektoreina. Pystyvektoreita kokoa n voidaankin pitää näiden entuudestaan tuttujen 'algebravektoreiden' yleistyksenä. (Lukuparien tai -kolmikoiden käsittely pystyvektorina ei mitenkään vaikuta vektorioperaatioihin.) Yhdenmukaisesti aiempien merkintöjen kanssa sovitaan, että pystyvektorit kokoa n muodostavat vektoriavaruuden nimeltä $\mathbb{R}^n$. Avaruuden $\mathbb{R}^n$ *nollavektori* on aiempaan tapaan vektori, jonka kaikki alkiot ovat nollia. Tätä merkitään symbolilla $\mathbf{0}$. Avaruuden $\mathbb{R}^n$ luonnollinen *kanta* saadaan, kun kirjoitetaan

$$\mathbf{x} = x_1 \mathbf{e}_1 + \ldots + x_n \mathbf{e}_n = \sum_{i=1}^{n} x_i \mathbf{e}_i,$$

jolloin

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

eli

$$[\mathbf{e}_i]_j \;=\; \delta_{ij} = \begin{cases} 1, & \text{jos } i = j, \\ 0, & \text{jos } i \neq j. \end{cases}$$

Tässä symboli δ_{ij} on nk. *Kroneckerin delta* (-symboli). Vektoreita $\mathbf{e}_i$ sanotaan $\mathbb{R}^n$:n *yksikkövektoreiksi.* Jokainen $\mathbf{x} \in \mathbb{R}^n$ voidaan siis kirjoittaa (ilmeisen yksikäsitteisesti) vektorien $\mathbf{e}_i$ *lineaariyhdistelynä*, eli $\{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ on $\mathbb{R}^n$:n kanta. Koska kannassa on n vektoria, sanotaan, että R^n:n *dimensio* on n:

$$\dim \mathbb{R}^n = n.$$

Toistaiseksi ei siis ole väliä, ovatko $\mathbb{R}^n$:n vektorit 'pystyssä' vai 'kumossa', kunhan yhtälössä (2) vain tulkitaan yhtäsuuruusmerkki normaaliin tapaan, eli

$$\mathbf{y} = \mathbf{b} \quad \Leftrightarrow \quad y_i = b_i, \quad i = 1 \ldots n.$$

Seuraavssa (toisessa) abstraktion vaiheessa kuitenkin tulee ero, kun yhtälössä (2) kertoimet a_{ij} erotetaan erilliseksi olioksi kirjoittamalla

$$\mathbf{y} \;=\; \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \;=\; \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \\ \vdots \\ \\ x_n \end{bmatrix} \;=\; \mathbf{A}\mathbf{x},$$

jolloin alkuperäinen yhtälöryhmä (1) tulee kirjoitetuksi muotoon

$$\mathbf{A}\mathbf{x} \;=\; \mathbf{b}. \tag{3}$$

Olio $\mathbf{A}$, joka siis näyttää kaksiulotteiselta lukutaulukolta, on nimeltään *matriisi* (engl. matrix). Indeksit m, n määräävät, että matriisi on *tyyppiä* tai *kokoa* $m \times n$ ('m kertaa n'). Tässä m on matriisin (vaaka)*rivien* (engl. row) ja n *sarakkeiden* (tai pysyrivien, engl. column) lukumäärä. Luvut a_{ij} ovat nimeltään matriisin *alkiot* tai *elementit.* Näille voidaan myös käyttää merkintää $[\mathbf{A}]_{ij}$, ja voidaan myös kirjoittaa

$$\mathbf{A} \;=\; (a_{ij}) \qquad \text{tai} \qquad \mathbf{A} \;=\; (a_{ij},\; i = 1 \ldots m,\; j = 1 \ldots n).$$

Määritelmän mukaisesti pystyvektori kokoa m (m-vektori) on matriisi tyyppiä $m \times 1$. Jatkossa erotetaan matriisit kuitenkin yleisemmin vektoreista käyttämällä matriisien symboleina lihavoituja isoja kirjaimia $\mathbf{A}$, $\mathbf{B}$, jne.

Kuten vektori, matriisikin 'elää' vasta sille määriteltyjen laskuoperaatioiden kautta. Yhtälössä (3) esiityy jo eräs kaikkein keskeisimmistä operaatioista: <u>matriisin</u>

ja vektorin kertolasku. Vertaamalla yhtälöitä (1)–(3) nähdään heti, mikä on tämän kertolaskun määritelmä: $\mathbf{Ax}$ on pystyvektori kokoa m, jonka alkiot $[\mathbf{Ax}]_i$ määritellään

$$\boxed{\begin{aligned} [\mathbf{Ax}]_i &= \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1\ldots m, \\ \mathbf{x} &= (\,x_i,\ i = 1\ldots n\,), \\ \mathbf{A} &= (\,a_{ij},\ i = 1\ldots m,\ j = 1\ldots n\,). \end{aligned}}$$

Määritelmän mukaisesti on $\mathbf{A}$:n sarakkeiden lukumäärän täsmättävä vektorin $\mathbf{x}$ kokoon, jotta $\mathbf{Ax}$ olisi määritelty. Tämäntyyppiset yhteensopivuusehdot rajoittavat kaikkia matriisien välisiä laskuoperaatioita.

Toinen matriisin ja vektorin tulon esitysmuoto saadaan, kun merkitään $\mathbf{a}_j = \mathbf{A}$:n j:s sarake, ts.

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Kun $\mathbf{A}$ esitetään sarakkeittensa avulla, niin em. määritelmästä nähdään, että pätee

$$\boxed{\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \quad \Rightarrow \quad \mathbf{Ax} = \sum_{j=1}^{n} x_j\mathbf{a}_j.}$$

Siis matriisin $\mathbf{A}$ ja vektorin $\mathbf{x} = (x_i)$ tulo on $\mathbf{A}$:n sarakkeiden $\mathbf{a}_j$ lineaarinen yhdistely, kertoimina x_j. Erityisesti siis pätee

$$\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \quad \Rightarrow \quad \mathbf{Ae}_k = \mathbf{a}_k, \quad k = 1\ldots n.$$

Tästä on puolestaan helppo päätellä seuraava tulos, jolla on käyttöä jatkossa.

Propositio X.1.1 Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat matriiseja kokoa $m \times n$, niin pätee

$$\mathbf{Ax} = \mathbf{Bx} \ \ \forall \mathbf{x} \in \mathbb{R}^n \quad \Rightarrow \quad \mathbf{A} = \mathbf{B}.$$

Esimerkki 2 Laske $\mathbf{Ax}$ ja $\mathbf{Bx}$, kun

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Ratkaisu

$$\mathbf{Ax} = \begin{bmatrix} 1\cdot 1 + 2\cdot(-1) \\ 0\cdot 1 + 1\cdot(-1) \end{bmatrix} = (1)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

$$\mathbf{Bx} = \begin{bmatrix} 1\cdot 1 + 3\cdot(-1) \\ 2\cdot 1 + (-1)\cdot(-1) \\ 1\cdot 1 + 1\cdot(-1) \end{bmatrix} = (1)\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}. \quad \square$$

Ym. määritelmästä nähdään, että matriisilla kertominen on kerrottavan vektorin suhteen lineaarinen laskutoimitus:

$$\mathbf{A}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathbf{Ax} + \mu\mathbf{Ay} \quad \forall\, \lambda, \mu \in \mathbb{R}.$$

(Tässä luonnollisesti edellytetään, että $\mathbf{x}$ ja $\mathbf{y}$ ovat samaa kokoa ja $\mathbf{A}$:n kanssa yhteensopivia).

Matriiseille määritellään skalaarilla kertominen ja samankokoisten matriisien yhteenlasku samalla tavoin kuin vektoreille, eli suoritetaan laskuoperaatiot alkioittain. Siis jos $\mathbf{A} = (a_{ij})$ ja $\mathbf{B} = (b_{ij})$ ovat samaa kokoa, niin

$$[\lambda\mathbf{A}]_{ij} = \lambda a_{ij}, \qquad [\mathbf{A} + \mathbf{B}]_{ij} = a_{ij} + b_{ij}.$$

Vektorioperaatiot $\mathbb{R}^n$:ssä voidaan tulkita näiden operaatioiden erikoistapauksiksi ($\mathbf{A}$ ja $\mathbf{B}$ kokoa $n \times 1$). Kuten samankokoiset vektorit, myös samankokoiset matriisit muodostavat näillä laskuoperaatioilla varustettuna vektoriavaruuden. Tämän avaruuden nolla-alkio on nk. *nollamatriisi*, jonka kaikki alkiot ovat nollia (symboli $\mathbf{0}$).

Matriisitulo

Kahden matriisin $\mathbf{A}$, $\mathbf{B}$ tulo eli *matriisitulo* voidaan määritellä, edellyttäen että matriisit ovat kooltaan yhteensopivat. Olkoon $\mathbf{x}$ pystyvektori kokoa n ja olkoon $\mathbf{A} = (a_{ij})$ kokoa $m \times p$ ja $\mathbf{B} = (b_{ij})$ kokoa $p \times n$ ($m, n, p \in \mathbb{N}$). Tällöin yhdistetty matriisi-vektoritulo $\mathbf{A}(\mathbf{Bx})$ on määritelty, kun $\mathbf{x} \in \mathbb{R}^n$. Halutaan kirjoittaa tämä muodossa

$$\mathbf{A}(\mathbf{Bx}) = (\mathbf{AB})\,\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \tag{4}$$

ja ottaa tämä matriisitulon $\mathbf{AB}$ määritelmäksi. Tällöin on oltava

$$\begin{aligned}[(\mathbf{AB})\mathbf{x}]_i = [\mathbf{A}(\mathbf{Bx})]_i &= \sum_{k=1}^{p} a_{ik}[\mathbf{Bx}]_k = \sum_{k=1}^{p} a_{ik} \sum_{j=1}^{n} b_{kj}x_j \\ &= \sum_{j=1}^{n} (\sum_{k=1}^{p} a_{ik}b_{kj})x_j = \sum_{j=1}^{n} [\mathbf{AB}]_{ij}x_j\,.\end{aligned}$$

Siis nähdään, että (4) toteutuu, kun määritellään

$$\boxed{\begin{aligned} [\mathbf{AB}]_{ij} &= \sum_{k=1}^{p} a_{ik} b_{kj}, \quad i = 1 \ldots m, \quad j = 1 \ldots n, \\ &\mathbf{A} = (\, a_{ij},\ i = 1 \ldots m,\ j = 1 \ldots p\,), \\ &\mathbf{B} = (\, b_{ij},\ i = 1 \ldots p,\ j = 1 \ldots n\,). \end{aligned}}$$

Matriisitulo on näin määritelty yksikäsitteisesti, sillä Proposition X.1.1 mukaan (4) voi toteutua vain yhdelle matriisille $\mathbf{AB}$. Määritelmästä nähdään, että $(\mathbf{AB})$:n j:s sarake $= \mathbf{Ab}_j$, missä $\mathbf{b}_j = (b_{kj},\ k = 1 \ldots p)$ on $\mathbf{B}$:n j:s sarake. Näin ollen jos $\mathbf{B}$ esitetään sarakkeittensa avulla, niin matriisitulolle $\mathbf{AB}$ saadaan myös esitysmuoto

$$\boxed{\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{bmatrix} \quad \Rightarrow \quad \mathbf{AB} = \begin{bmatrix} \mathbf{Ab}_1 & \ldots & \mathbf{Ab}_n \end{bmatrix}.}$$

Matriisitulo palautuu näin matriisin ja vektorin tuloiksi: Kerrotaan $\mathbf{B}$:n kukin sarakevektori $\mathbf{A}$:lla ja sijoitetaan tulokset $(\mathbf{AB})$:n sarakkeiksi.

Matriisitulon taustalla olevasta sopimuksesta (4) voidaan päätellä, että matriisitulo on liitännäinen. Nimittäin jos $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{x}$ ovat kooltaan yhteensopivia, niin sopimuksen (4) mukaan pätee jokaisella $\mathbf{x} \in \mathbb{R}^n$

$$[\mathbf{A}(\mathbf{BC})]\mathbf{x} \;=\; \mathbf{A}[(\mathbf{BC})\mathbf{x}] \;=\; \mathbf{A}[\mathbf{B}(\mathbf{Cx})] = \;(\mathbf{AB})(\mathbf{Cx}) \;=\; [(\mathbf{AB})\mathbf{C}]\mathbf{x},$$

jolloin on oltava (Propositio X.1.1)

$$\boxed{\mathbf{A}(\mathbf{BC}) \;=\; (\mathbf{AB})\mathbf{C}.}$$

Vaihdannainen matriisitulo ei sen sijaan ole. Nimittäin ensinnäkin tuloon liittyvät yhteensopivuussäännöt ovat varsin rajoittavia: Tulot $\mathbf{AB}$ ja $\mathbf{BA}$ eivät välttämättä ole molemmat määriteltyjä, ja vaikka olisivat, eivät välttämättä samaa tyyppiä. Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat molemmat kokoa $n \times n$, eli samankokoisia *neliömatriiseja* (engl. square matrix), niin asia on näiltä osin kunnossa, mutta tällöinkin on yleisesti $\mathbf{AB} \neq \mathbf{BA}$.

Neliömatriisit ovat lineaaristen yhtälöryhmien algebran kannalta huomattavan tärkeä matriisien erikoisluokka, josta puhutaan enemmän seuraavassa luvussa. Todettakoon tässä yhteydessä kuitenkin, että jos samankokoisten neliömatriisien tapauksessa sattuu olemaan voimassa $\mathbf{AB} = \mathbf{BA}$, niin sanotaan, että matriisit $\mathbf{A}$ ja $\mathbf{B}$ *kommutoivat*:

$$\mathbf{AB} \;= \mathbf{BA} \quad \Leftrightarrow \quad \mathbf{A} \text{ ja } \mathbf{B} \text{ kommutoivat.}$$

Jokainen neliömatriisi kommutoi triviaalisti ainakin itsensä kanssa.

Esimerkki 3 Matriisien

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

yhdeksästä mahdollisesta keskinäisestä tulosta $\mathbf{AC}$, $\mathbf{BC}$ ja $\mathbf{CC}$ eivät ole määriteltyjä, muut ovat:

$$\mathbf{AB} = \left[\mathbf{A}\begin{bmatrix} 0 \\ 0 \end{bmatrix} \ \mathbf{A}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right] = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{BA} = \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{CA} = \left[\mathbf{C}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \ \mathbf{C}\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right] = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{CB} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{AA} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}, \quad \mathbf{BB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \qquad \square$$

Esimerkissä neliömatriisit $\mathbf{A}$ ja $\mathbf{B}$ siis eivät kommutoi. Neliömatriisille määritelty tulo $\mathbf{AA}$ voidaan kirjoittaa $\mathbf{A}^2$, ja määritellä yleisemminkin neliömatriisin potenssiin korotus:

$$\mathbf{A}^k = \underbrace{\mathbf{AA}\cdots\mathbf{A}}_{k \text{ kpl}}, \quad k \in \mathbb{N} \quad (\mathbf{A} \text{ neliömatriisi}).$$

Esimerkistä nähdään, että neliömatriisin tapauksessa voi olla $\mathbf{B}^2 = \mathbf{0}$ (= nollamatriisi), vaikka $\mathbf{B} \neq \mathbf{0}$.

Matriisin transpoosi

Vielä on määrittelemättä yksi keskeinen matriisialgebran operaatio, *transponointi*. Tällä tarkoitetaan yksinkertaisesti matriisin $\mathbf{A}$ rivien ja sarakkeiden vaihtoa keskenään. Tulosta merkitään $\mathbf{A}^T$ ja kutsutaan $\mathbf{A}$:n *transpoosiksi* :

$$\boxed{[\mathbf{A}^T]_{ij} = [\mathbf{A}]_{ji}.}$$

Esimerkki 3 (jatko) Esimerkin matriiseille

$$\mathbf{A}^T = \mathbf{A}, \quad \mathbf{B}^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{C}^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \qquad \square$$

Kuten esimerkissä, neliömatriisin tapauksessa (ei muussa) on mahdollista, että $\mathbf{A}^T = \mathbf{A}$, jolloin sanotaan, että $\mathbf{A}$ on *symmetrinen*:

$$\mathbf{A}^T = \mathbf{A} \quad \Leftrightarrow \quad \mathbf{A} \text{ symmetrinen.}$$

Transponoinnin laskusäännöistä tärkeimmät ovat

$$\boxed{(\mathbf{A}^T)^T = \mathbf{A}, \quad (\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T.}$$

Näistä ensimmäinen on ilmeinen, ja myös jälkimmäisen (tulon transponointisäännön) voi johtaa suoraviivaisesti matriisitulon määritelmästä (Harj.teht. 11a). Kun tulon transponontisääntöä sovelletaan useampikertaiseen matriisituloon, saadaan yleisempi sääntö

$$(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_n)^T \;=\; \mathbf{A}_n^T\mathbf{A}_{n-1}^T\ldots\mathbf{A}_1^T.$$

Siis tulon transpoosi = transpoosien tulo käänteisessä järjestyksessä.

Muista matriisialgebran laskusäännöistä mainittakoon vielä määritelmien perusteella ilmeiset (oletetaan matriisien yhteensopivuus)

$$\begin{aligned}
&\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}, \qquad (\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C}),\\
&\mathbf{A}(\lambda\mathbf{B}) = (\lambda\mathbf{A})\mathbf{B} = \lambda(\mathbf{AB}), \quad \lambda\in\mathbb{R},\\
&\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}, \qquad (\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC},\\
&(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T+\mathbf{B}^T, \qquad (\lambda\mathbf{A})^T = \lambda\mathbf{A}^T, \quad \lambda\in\mathbb{R}.
\end{aligned}$$

Vaakavektorit — euklidinen skalaaritulo ja normi

Kun pystyvektori $\mathbf{x} = (x_i)_{i=1}^n$ tulkitaan matriisiksi kokoa $n\times 1$, niin tällä matriisilla on transpoosi, jota merkitään $\mathbf{x}^T$ ja sanotaan *vaakavektoriksi* (engl. row vector). Vektorialgebran kannalta pysty- ja vaakavektoreilla ei ole eroa, joten vektoriavaruus $\mathbb{R}^n$ voidaan yhtä hyvin ajatella vaakavektoreista koostuvaksi. Matriisialgebran kannalta sen sijaan pysty- ja vaakavektorit ovat erilaisia olioita. Erityisen mielenkiintoinen on matriisitulo $\mathbf{x}^T\mathbf{y}$, missä $\mathbf{y} = (y_i)_{i=1}^n$ on samaa kokoa oleva pystyvektori kuin $\mathbf{x}$. Matriisitulon määritelmän mukaan $\mathbf{x}^T\mathbf{y}$ on matriisi kokoa 1×1 eli skalaari:

$$\mathbf{x}^T\mathbf{y} = \sum_{i=1}^n x_iy_i \;\in\mathbb{R}.$$

Tässä on itse asiassa määritelty matriisialgebran keinoin $\mathbb{R}^n$:n *euklidinen skalaaritulo*

$$\boxed{\langle\mathbf{x},\mathbf{y}\rangle = \mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x} = \sum_{i=1}^n x_iy_i, \quad \mathbf{x},\mathbf{y}\in\mathbb{R}^n.}$$

Tämä on aiemmin Luvuissa II.2 ja II.6 määriteltyjen $\mathbb{R}^2$:n ja $\mathbb{R}^3$:n skalaaritulojen yleistys. Cauchyn-Schwarzin epäyhtälö $\mathbb{R}^n$:n euklidiselle skalaaritulolle on (ks. Lause II.4.2)

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leq |\mathbf{x}||\mathbf{y}|, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

missä

$$|\mathbf{x}| = \langle \mathbf{x}, \mathbf{x}\rangle^{1/2} = (\mathbf{x}^T\mathbf{x})^{1/2} = \big(\sum_{i=1}^{n} x_i^2\big)^{1/2}$$

on $\mathbb{R}^n$:n *euklidinen normi*. Viitaten näin määriteltyyn skalaarituloon ja normiin käytetään avaruudesta $\mathbb{R}^n$ yleisesti nimitystä *euklidinen avaruus* $\mathbb{R}^n$. Edellä määritelty $\mathbb{R}^n$:n luonnollinen kanta $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ on euklidisen avaruuden kantana *ortonormeerattu*:

$$\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij} \quad \text{(Kroneckerin } \delta\text{)}.$$

Jos $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ ja $\mathbf{A}$ on matriisi kokoa $m \times n$, niin skalaaritulo $\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle = (\mathbf{A}\mathbf{x})^T\mathbf{y}$ on määritelty. Tässä on tulon transponointisäännön mukaan $(\mathbf{A}\mathbf{x})^T = \mathbf{x}^T\mathbf{A}^T$, joten pätee

$$\boxed{\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{A}^T\mathbf{y}\rangle.}$$

Huomautettakoon vielä, että jos $\mathbf{A}$ on kokoa $m\times p$, $\mathbf{B}$ on kokoa $p\times n$, ja merkitään $\mathbf{A}^T = [\mathbf{a}_1 \ldots \mathbf{a}_m]$ (eli $\mathbf{A}$:n i:s rivi $= \mathbf{a}_i^T$) ja $\mathbf{B} = [\mathbf{b}_1 \ldots \mathbf{b}_n]$, niin matriisitulon määritelmän perusteella

$$[\mathbf{AB}]_{ij} = \mathbf{a}_i^T\mathbf{b}_j = \langle \mathbf{a}_i, \mathbf{b}_j\rangle.$$

Tulo $\mathbf{AB}$ on siis taulukko, joka muodostuu $\mathbf{A}$:n rivien $\mathbf{B}$:n sarakkeiden välisistä ($\mathbb{R}^p$:n) skalaarituloista.

Kompleksiset vektorit ja matriisit

Koska matriisialgebrassa on perimmältään kyse vain matriisialkioiden välisistä peruslaskutoimituksista ja niiden yhdistelystä, voidaan matriisialkioiden ajatella kuuluvan $\mathbb{R}$:n sijasta mihin tahansa kuntaan $\mathbb{K}$. Yksinkertaisissa laskuesimerkeissä (kuten esimerkit edellä) on usein $\mathbb{K} = \mathbb{Q}$. Yleisemmisssä matemaattisissa tarkasteluissa, samoin monissa sovelluksissa (esim. sähkötekniikassa) on sallittava matriisialkioiden kompleksiarvoisuus, jolloin $\mathbb{K} = \mathbb{C}$. Tässä tapauksessa pystyvektorit kokoa n muodostavat euklidisen avaruuden $\mathbb{C}^n$, jonka skalaaritulo on

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n} x_i\bar{y}_i = \mathbf{x}^T\bar{\mathbf{y}}.$$

Tässä $\bar{\mathbf{y}}$, yleisemmin $\bar{\mathbf{A}}$, tarkoittaa kompleksista konjugointia alkioittain.

Esimerkki 4 Jos

$$\mathbf{x} = [\,i,\ 1+i,\ 2-i\,]^T \in \mathbb{C}^3, \quad \mathbf{y} = [\,1-2i,\ 2i,\ -1-i\,]^T \in \mathbb{C}^3,$$

niin

$$\langle \mathbf{x}, \mathbf{x} \rangle = |\mathbf{x}|^2 = 1 + (1+1) + (4+1) = 8, \quad \langle \mathbf{y}, \mathbf{y} \rangle = (1+4) + 4 + (1+1) = 11,$$
$$\langle \mathbf{x}, \mathbf{y} \rangle = i(1+2i) + (1+i)(-2i) + (2-i)(-1+i) = -1+3i. \qquad \square$$

Matriisia $\mathbf{A}^* = \bar{\mathbf{A}}^T$ sanotaan $\mathbf{A}$:n *liittomatriisiksi* (engl. adjoint). Jos $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{y} \in \mathbb{C}^m$ ja $\mathbf{A}$ on kompleksinen matriisi kokoa $m \times n$, niin liittomatriisin ja $\mathbb{C}^m$:n skalaaritulon määritelmistä seuraa helposti (vrt. reaalinen tapaus edellä)

$$\boxed{\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle.}$$

Jos $\mathbf{A} = \{a_{ij}\}_{i,j=1}^n$ on neliömatriisi ja $\mathbf{A}^* = \mathbf{A}$, ts. $a_{ji} = \bar{a}_{ij}$ $\forall i, j$, niin sanotaan, että $\mathbf{A}$ on *hermiittinen* (engl. Hermitean):

$$\mathbf{A}^* = \mathbf{A} \quad \Leftrightarrow \quad \mathbf{A} \text{ hermiittinen.}$$

Jokainen reaalinen ja symmetrinen matriisi on määritelmän mukaan myös hermiittinen.

Esimerkki 5 Yleinen hermiittinen matriisi kokoa 2×2 on muotoa

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

missä $a, b \in \mathbb{R}$ ja $c \in \mathbb{C}$. $\square$

HARJOITUSTEHTÄVIÄ

1. Laske $2\mathbf{A} + 3\mathbf{B}$, $\mathbf{A} - \mathbf{B}^T$, $\mathbf{AB}$ ja $\mathbf{BA}$, kun

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix}.$$

2. Olkoon

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

Laske $\mathbf{AB}$, $\mathbf{AC}$, $\mathbf{BC}$, $\mathbf{CA}$ ja $\mathbf{BA}^T$.

3. Millä matriisien tyyppiä koskevilla oletuksilla tulot $\mathbf{AB}$ ja $\mathbf{BA}$ ovat
a) molemmat määriteltyjä, b) samaa tyyppiä?

4. Olkoon

$$\mathbf{a} = [1, 3, 5, 2], \quad \mathbf{b} = [-1, 3, 2, 4]^T, \quad \mathbf{A} = ((-1)^{i+j},\ i = 1, 2,\ j = 1, 2, 3, 4).$$

Laske vektorien/matriisien $\mathbf{a}$, $\mathbf{b}$, $\mathbf{A}$ keskinäisistä (kaksittaisista) tuloista kaikki, jotka ovat määriteltyjä.

5. Matriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat kokoa 10×10 ja niiden alkiot ovat $a_{ij} = i + j$ ja $b_{ij} = i - j$. Laske tulomatriisin $\mathbf{C} = \mathbf{AB}$ alkio c_{ij}.

6. Olkoon

$$\mathbf{A} = a \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b & -c & -c \\ -c & d & d \\ -c & d & d \end{bmatrix}.$$

Määritä luvut a, b, c, d siten, että $\mathbf{A} \neq \mathbf{0}$, $\mathbf{B} \neq \mathbf{0}$, $\mathbf{A}^2 = \mathbf{A}$, $\mathbf{B}^2 = \mathbf{B}$ ja $\mathbf{AB} = \mathbf{0}$.

7. Määritä matriisi $\mathbf{A}$, kun tiedetään, että

$$\mathbf{A} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 2 \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}.$$

8. Hae kaikki matriisit $\mathbf{B}$, jotka kommutoivat matriisin

$$\text{a)}\ \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \qquad \text{b)}\ \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

kanssa.

9. Olkoon $\mathbf{A}$ ja $\mathbf{B}$ samaa kokoa olevia neliömatriiseja. Todista:
a) $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 \ \Leftrightarrow\ \mathbf{A}$ ja $\mathbf{B}$ kommutoivat
b) $\mathbf{A}$ ja $\mathbf{B}$ symmetriset $\not\Rightarrow$ $\mathbf{AB}$ symmetrinen
c) $\mathbf{A}$ ja $\mathbf{B}$ symmetriset ja kommutoivat $\Rightarrow$ $\mathbf{AB}$ symmetrinen

10. Laske a) $\mathbf{A}^k\ \forall k \in \mathbb{N}$, b) $\mathbf{B} = (\mathbf{I} + \mathbf{A})^{100}$, kun

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

11. a) Todista matriisitulon transponointisääntö $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

    b) Näytä, että jokainen neliömatriisi on esitettävissä yksikäsitteisesti muodossa $\mathbf{A} = \mathbf{B} + \mathbf{C}$, missä $\mathbf{B}$ on symmetrinen ja $\mathbf{C}$ on *vinosymmetrinen*: $\mathbf{C} + \mathbf{C}^T = \mathbf{0}$.

12. Olkoon

$$\mathbf{A} = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

    Sievennä lauseke $f(x, y, z) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x}$ toisen asteen polynomiksi.

13. Laske $\mathbf{Ax}$, $\mathbf{Bx}$, $\mathbf{AB}$, $(\mathbf{AB})^*$ ja $\mathbf{B}^*\mathbf{A}^*$, kun

$$\mathbf{A} = \begin{bmatrix} i & 1-i & 2+i \\ 1-i & -2i & 3+2i \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2-i & i & -i \\ -i & 2 & 1+i \\ 3i & 2+2i & 1-i \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 3-i \\ 2+i \end{bmatrix}.$$

14. (*) Jos reaalinen matriisi $\mathbf{A} = (a_{ij})$ kokoa $m \times n$ tulkitaan avaruuden $\mathbb{R}^{mn}$ alkiona, niin ko. avaruuden euklidista normia vastaa *matriisinormi*

$$\|\mathbf{A}\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}.$$

    Näytä, että pätee $|\mathbf{Ax}| \leq \|\mathbf{A}\||\mathbf{x}|$ $\forall \mathbf{x} \in \mathbb{R}^n$.

15. (*) Olkoon $x \in \mathbb{R}$ ja määritellään

$$\mathbf{A} = \begin{bmatrix} x & 1 & 0 & 0 \\ 0 & x & 1 & 0 \\ 0 & 0 & x & 1 \\ 0 & 0 & 0 & x \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \exp(\mathbf{A}) = \mathbf{I} + \sum_{k=1}^{\infty} \frac{1}{k!} \mathbf{A}^k.$$

    Laske matriisin $\exp(\mathbf{A})$ alkiot x:n funktiona.

16. (*) (Naudat laitumella) Nautalaumassa on on 83 täysin ruskeata eläintä, 77 sarvipäätä, 36 sukupuoleltaan sonnia, 22 ruskeata sarvipäätä, 15 ruskeata sonnia, 25 sarvipäistä sonnia ja 7 ruskeata sarvipäistä sonnia. Kaikki lauman eläimet kuuluvat johonkin mainituista ryhmistä. Montako eläintä laumassa on? *Vihje*: Jaa eläimet pistevieraisiin ryhmiin, esim. ruskeita sarvipäisiä sonneja x_1 kpl ... ei-ruskeita sarvettomia lehmiä x_8 kpl. Kirjoita lineaarinen yhtälöryhmä ja ratkaise!

X.2 Neliömatriisit. Käänteismatriisi

Tässä ja seuraavassa luvussa tutkimuskohteena on lineaarinen yhtälöryhmä

$$\mathbf{Ax} = \mathbf{b},$$

missä yhtälöitä ja tuntemattomia on yhtä monta, eli $\mathbf{A}$ on neliömatriisi kokoa $n \times n$ ja $\mathbf{b} \in \mathbb{R}^n$ ($n \in \mathbb{N}$). (Mukavuussyistä ajatellaan $\mathbf{A}$ ja $\mathbf{b}$ reaalisiksi — tulokset jatkossa pätevät sellaisinaan myös kompleksialueella.)

Sekä matriisitulon että matriisi-vektoritulon kannalta tärkeä neliömatriisien erikoistapaus on *yksikkömatriisi* (engl. unit matrix) eli *identiteettimatriisi*, jonka symboli on $\mathbf{I}$ ja määritelmä

$$[\mathbf{I}]_{ij} = (\delta_{ij}) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = [\mathbf{e}_1 \dots \mathbf{e}_n].$$

Jos $\mathbf{A}$ ja $\mathbf{I}$ ovat kokoa $n \times n$ ja $\mathbf{x}$ on pystyvektori kokoa n, niin pätee

$$\text{(a)} \quad \mathbf{AI} = \mathbf{IA} = \mathbf{A}, \qquad \text{(b)} \quad \mathbf{Ix} = \mathbf{x}.$$

Ominaisuuden (a) mukaan $\mathbf{I}$ on matriisikertolaskun ykkösalkio samankokoisten neliömatriisien välisissä operaatioissa — tästä nimi 'yksikkömatriisi'. Ominaisuuden (b) mukaan kuvaus $\mathbf{x} \mapsto \mathbf{Ix}$ on $\mathbb{R}^n$:n *identiteettikuvaus* — tästä nimi 'identiteettimatriisi'.

Jos matriisien kertolaskuun liittyy 'ykkösmatriisi', niin liittyykö myös käänteismatriisi? — Tässä tullaankin neliömatriisien teorian keskeisimpään kysymykseen.

MÄÄRITELMÄ X.2.1 Neliömatriisi $\mathbf{A}$ on *säännöllinen* eli *ei-singulaarinen* eli *kääntyvä*, jos on olemassa matriisi $\mathbf{B}$ siten, että pätee

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

Sanotaan, että $\mathbf{B}$ on $\mathbf{A}$:n *käänteismatriisi* (engl. inverse matrix) ja merkitään

$$\mathbf{B} = \mathbf{A}^{-1}.$$

Jos $\mathbf{A}$ ei ole säännöllinen, se on epäsäännöllinen eli *singulaarinen*.

Käänteismatriisin symboli $\mathbf{A}^{-1}$ luetaan yleensä 'A miinus yksi'. Jos $\mathbf{A}^{-1}$ on olemassa, niin se on yksikäsitteinen. Nimittäin jos $\mathbf{B}$ ja $\mathbf{C}$ ovat molemmat $\mathbf{A}$:n käänteismatriiseja, niin seuraaan väittämän mukaan on oltava $\mathbf{B} = \mathbf{C}$.

PROPOSITIO X.2.2 Jos on olemassa matriisit $\mathbf{B}$ ja $\mathbf{C}$ kokoa $n \times n$ siten, että $\mathbf{AB} = \mathbf{I}$ ja $\mathbf{CA} = \mathbf{I}$, niin $\mathbf{B} = \mathbf{C}$.

Todistus Oletuksien ja matriisitulon liitännäisyyden perusteella

$$\mathbf{B} = \mathbf{IB} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}. \quad \square$$

Jos $\mathbf{A}$ on säännöllinen, niin Määritelmästä X.2.1 nähdään välittömästi, että myös $\mathbf{B} = \mathbf{A}^{-1}$ on säännöllinen ja $\mathbf{B}^{-1} = \mathbf{A}$. Suorittamalla määritelmässä transponointi nähdään myös, että

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I} \;\Leftrightarrow\; \mathbf{B}^T\mathbf{A}^T = \mathbf{A}^T\mathbf{B}^T = \mathbf{I}^T = \mathbf{I}.$$

Siis: Jos $\mathbf{A}$ on säännöllinen, niin sekä $\mathbf{A}^{-1}$ että $\mathbf{A}^T$ ovat säännöllisiä, ja pätee

$$\boxed{(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \qquad (\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.}$$

Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat säännöllisiä ja samaa kokoa, niin myös tulo $\mathbf{AB}$ on säännöllinen. Nimittäin matriisitulon säännöin ja käänteismatriisin määritelmän perusteella

$$\begin{aligned}(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) &= \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I},\\ (\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) &= \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{IB} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I},\end{aligned}$$

joten

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

Tulos on helposti yleistettävissä seuraavasti (vrt. vastaava transponointisääntö edellisessä luvussa):

PROPOSITIO X.2.3 Jos $\mathbf{A}_k$, $k = 1 \ldots m$ ovat säännöllisiä ja samaa kokoa olevia neliömatriiseja, niin tulo $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_m$ on myös säännöllinen ja $\mathbf{A}^{-1} = \mathbf{A}_m^{-1}\mathbf{A}_{m-1}^{-1} \cdots \mathbf{A}^{-1}$.

Lineaarista yhtälöryhmää ratkaistaessa kerroinmatriisin säännöllisyys 'ratkaisee' ongelman periaatteelliselta kannalta seuraavasti:

PROPOSITIO X.2.4 Jos yhtälöryhmässä $\mathbf{Ax} = \mathbf{b}$ kerroinmatriisi $\mathbf{A}$ on säännöllinen neliömatriisi kokoa $n \times n$, niin yhtälöryhmällä on jokaisella $\mathbf{b} \in \mathbb{R}^n$ yksikäsitteinen ratkaisu

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Todistus Merkitään $\mathbf{B} = \mathbf{A}^{-1}$ ja päätellään:

$$\begin{aligned} &\text{a)}\ \mathbf{AB} = \mathbf{I} \quad \Rightarrow \quad \mathbf{A}(\mathbf{Bb}) = (\mathbf{AB})\mathbf{b} = \mathbf{Ib} = \mathbf{b}. \\ &\text{b)}\ \mathbf{BA} = \mathbf{I} \quad \Rightarrow \quad \mathbf{x} = \mathbf{Ix} = (\mathbf{BA})\mathbf{x} = \mathbf{B}(\mathbf{Ax}). \end{aligned}$$

Tämän perusteella todetaan: a) $\mathbf{x} = \mathbf{Bb}$ on yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ ratkaisu. b) Jos $\mathbf{Ax} = \mathbf{b}$, niin $\mathbf{x} = \mathbf{Bb}$, eli tämä on ainoa mahdollinen ratkaisu. □

Proposition X.2.4 mukaan yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ kerroinmatriisin mahdollinen singulaarisuus paljastuu yhälöryhmää ratkaistaessa:

Korollaari X.2.5 Jos yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ ($\mathbf{A}$ kokoa $n \times n$) joko ei ratkea jollakin $\mathbf{b} \in \mathbb{R}^n$ tai ratkaisu ei ole yksikäsitteinen, niin $\mathbf{A}$ on singulaarinen.

Neliömatriisin säännöllisyyskriteerit

Proposition X.2.4 mukaan kerroinmatriisin $\mathbf{A}$ säännöllisyys takaa yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ säännöllisyyden eli yksikäsitteisen ratkeavuuden jokaisella $\mathbf{b} \in \mathbb{R}^n$. Tämä väittämä pätee myös kääntäen, ja itse asiassa pätee paljon vahvempi tulos: Jos yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ joko ratkeaa yksikäsitteisesti kun $\mathbf{b} = \mathbf{0}$ (eli ainoa ratkaisu on $\mathbf{x} = \mathbf{0}$) tai ratkeaa jokaisella $\mathbf{b} \in \mathbb{R}^n$, niin kummassakin tapauksessa $\mathbf{A}$ on säännöllinen matriisi. Muotoillaan tulos seuraavasti:

*Lause X.2.6 (**Neliömatriisin säännöllisyys**) Jos $\mathbf{A}$ on neliömatriisi kokoa $n \times n$, niin seuraavat väittämät ovat keskenään yhtäpitävät.

E0. $\mathbf{A}$ on säännöllinen matriisi.

E1. Yhtälöryhmän $\mathbf{Ax} = \mathbf{0}$ ainoa ratkaisu on $\mathbf{x} = \mathbf{0}$.

E2. Yhtälöryhmällä $\mathbf{Ax} = \mathbf{b}$ on ratkaisu jokaisella $\mathbf{b} \in \mathbb{R}^n$.

Lause X.2.6 sisältää kuusi implikaatioväittämää, joista riippumattomia ovat esim.: E0 $\Rightarrow$ E1, E0 $\Rightarrow$ E2, E1 $\Rightarrow$ E0 ja E2 $\Rightarrow$ E0. Näistä kaksi ensimmäistä ovat Propositioon X.2.4 sisältyviä ja siis jo selviä. Kaksi muuta sen sijaan ovat syvällisempiä. Nämä sisältyvät erikoistapauksina yleisempään lineaarisia yhtälöryhmiä koskevaan väittämään, joka tunnetaan nimellä *Lineaarialgebran peruslause*. Yleisempää lausetta (joka koskee myös systeemejä kokoa $m \times n$, $m \neq n$) ei tässä muotoilla. Todetaan sen sijaan Lauseen X.2.6 mielenkiintoinen seuraamus:

Korollaari X.2.7 Samaa kokoa oleville neliömatriiseille pätee

$$\mathbf{AB} = \mathbf{I} \quad \Rightarrow \quad \mathbf{BA} = \mathbf{I}.$$

Todistus Jos $\mathbf{AB} = \mathbf{I}$, niin säännöllisyyskriteeri E2 on täytetty (ks. Proposition X.2.4 todistus, osa a)), joten Lauseen X.2.6 mukaan $\mathbf{A}$ on säännöllinen. Tällöin jos $\mathbf{C} = \mathbf{A}^{-1}$, niin $\mathbf{AB} = \mathbf{CA} = \mathbf{I}$, jolloin on oltava $\mathbf{C} = \mathbf{B}$ (Propositio X.2.2). Siis $\mathbf{BA} = \mathbf{I}$. $\square$

Korollaarin X.2.7 mukaan siis jo ehto $\mathbf{AB} = \mathbf{I}$ riittää takaamaan, että $\mathbf{B} = \mathbf{A}^{-1}$ ja $\mathbf{A} = \mathbf{B}^{-1}$, ts. toinen Määritelmän X.2.1 ehdoista on turha (!).

Jatkossa ei Lausetta X.2.6 pyritä heti todistamaan, vaan seuraavissa kahdessa luvussa tarkastellaan ensin lineaarisen yhtälöryhmän ratkaisemista algoritmisin keinoin. Osoittautuu, että ratkaisualgoritmin sivutuotteena sadaan myös Lause X.2.6 todistetuksi.†

Tämän luvun loppuosassa kohdisteaan huomio kahteen neliömatriisien erikoisluokkaan, *ortogonaali-* ja *kolmio*matriiseihin ja edellisten alaluokkaan *permutaatio*matriiseihin. Permutaatio- ja kolmiomatriiseilla on jatkossa keskeinen rooli lineaarisen yhtälöryhmän ratkaisualgoritmissa. Kuten nähdään, näiden matriisien osalta säännöllisyyskysymys voidaan ratkaista tukeutumatta Lauseeseen X.2.6. Kolmiomatriisin tapauksessa perustana on seuraavista kahdesta väittämästä jälkimmäinen (joka nojaa edelliseen). Tämä muistuttaa Lauseen X.2.6 väittämää E2 $\Rightarrow$ E0 mutta perustuu tätä vahvempiin oletuksiin. Muotoillaan ensin alkuperäinen oletusväittämä E2 toisella tavalla:

Propositio X.2.8 Pätee: E2 $\Leftrightarrow$ $\mathbf{AB} = \mathbf{I}$ jollakin $\mathbf{B}$.

Todistus $\boxed{\Rightarrow}$ Jos E2 on tosi, niin on olemassa vektorit $\mathbf{b}_k \in \mathbb{R}^n$ siten, että $\mathbf{Ab}_k = \mathbf{e}_k,\ k = 1 \ldots n$. Tällöin jos $\mathbf{B} = [\mathbf{b}_1 \ldots \mathbf{b}_n]$, niin $\mathbf{AB} = [\mathbf{Ab}_1 \ \ldots \ \mathbf{Ab}_n] = [\mathbf{e}_1 \ldots \mathbf{e}_n] = \mathbf{I}$. $\boxed{\Leftarrow}$ Ks. Proposition X.2.4 todistus, osa a). $\square$

Propositio X.2.9 Jos $\mathbf{A}$ on kokoa $n \times n$ ja yhtälöryhmät $\mathbf{Ax} = \mathbf{b}$ ja $\mathbf{A}^T\mathbf{x} = \mathbf{b}$ ovat molemmat ratkeavia jokaisella $\mathbf{b} \in \mathbb{R}^n$, niin $\mathbf{A}$ on säännöllinen matriisi.

Todistus Oletuksien ja Propostition X.2.8 perusteella on olemassa matriisit $\mathbf{B}$ ja $\mathbf{C}$ siten, että $\mathbf{AB} = \mathbf{I}$ ja $\mathbf{A}^T\mathbf{C} = \mathbf{I}$. Tällöin on myös $(\mathbf{A}^T\mathbf{C})^T = \mathbf{C}^T\mathbf{A} = \mathbf{I}^T = \mathbf{I}$. Siis $\mathbf{AB} = \mathbf{C}^T\mathbf{A} = \mathbf{I}$, mistä seuraa (Propositio X.2.2), että $\mathbf{B} = \mathbf{C}^T (= \mathbf{A}^{-1})$, eli $\mathbf{A}$ on säännöllinen. $\square$

†Lineaarialgebraa käsittelevässä kirjallisuudessa todistetaan Lineaarialgebran peruslause (ja siihen sisältyen Lause X.2.6) yleensä ei-algoritmisesti, abstraktin lineaarialgebran keinoin.

Ortogonaalimatriisit

Erikoisen säännöllisten matriisien luokan muodostavat *ortogonaaliset* matriisit. Jos neliömatriisi $\mathbf{A}$ esitetään sarakkeidensa avulla muodossa

$$\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n],$$

niin sanotaan, että $\mathbf{A}$ on ortogonaalinen, jos $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ on ortonormeerattu systeemi $\mathbb{R}^n$:ssä, ts.

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \mathbf{a}_i^T \cdot \mathbf{a}_j = \delta_{ij}, \quad i, j = 1 \dots n.$$

Matriisitulon määritelmän perusteella (ks. edellinen luku) tämä on sama kuin ehto

$$\mathbf{A}^T\mathbf{A} = \mathbf{I}.$$

Tällöin Korollaarin X.2.7 mukaan on myös $\mathbf{A}\mathbf{A}^T = \mathbf{I}$, eli ortogonaalisessa matriisissa myös rivit muodostavat ortonormeeratun systeemin (!) ja $\mathbf{A}^{-1} = \mathbf{A}^T$. Toisaalta, jos $\mathbf{A}^T\mathbf{A} = \mathbf{I}$, niin

$$\delta_{ij} = [\mathbf{A}^T\mathbf{A}]_{ij} = \mathbf{a}_i^T\mathbf{a}_j,$$

eli $\mathbf{A}$ on ortogonaalinen. On siis päädytty (Korollaariin X.2.7 ja siis Lauseeseen X.2.6 osaksi vedoten) tulokseen

$$\boxed{\mathbf{A} \text{ ortogonaalinen } \Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T.}$$

Jos $\mathbf{A}$ on sekä ortogonaalinen että <u>symmetrinen</u>, niin $\mathbf{A}\mathbf{A} = \mathbf{I}$. Voidaan siis todeta (tällä kertaa Lauseeseen X.2.6 vetoamatta), että pätee:

Propositio X.2.10 Jos $\mathbf{A}$ on ortogonaalinen ja symmetrinen matriisi, niin $\mathbf{A}$ on säännöllinen ja $\mathbf{A}^{-1} = \mathbf{A}$.

Permutaatiomatriisi $\mathbf{I}_p$

Jatkon kannalta erityisen kiinnostava ortogonaalisten matriisien luokka koostuu *permutaatiomatriiseista*, joissa on samat sarakkeet kuin yksikkömatriisissa $\mathbf{I}$ mutta vaihdetussa (permutoidussa) järjestyksessä. Permutaatiomatriisi merkitään

$$\mathbf{I}_p = [\mathbf{e}_{i_1} \ \mathbf{e}_{i_2} \ \dots \mathbf{e}_{i_n}],$$

missä $p = (i_1, i_2, \dots, i_n)$ on järjestetyn lukujoukon $(1, 2, \dots, n)$ *permutaatio* (eli samatluvut eri järjestyksessä). Permutaatio-opin tunnettu (helppo) tulos on, että

lukujoukon $(1, 2, \ldots, n)$ erilaisia permutaatioita — ja siis myös erilaisia permutaatiomatriiseja $\mathbf{I}_p$ kokoa $n \times n$ — on $n!$ kpl. Jokainen permutaatio p on saavutettavissa suorittamalla (p:stä riippuva) määrä m peräkkäisiä *parivaihtoja*, joissa kaksi lukua vaihdetaan keskenään. Luku m ei ole yksikäsitteinen, koska saman parivaihdon toisto palauttaa alkuperäisen järjestyksen. Sen sijaan on osoitettavissa (vaikka ei aivan helposti), että luvun m parillisuus/parittomuus on permutaatiolle ominainen, ts. jokainen permutaatio on joko *parillinen* tai *pariton*.

Jokainen permutaatiomatriisi $\mathbf{I}_p$ on luonnollisesti ortogonaalinen, ts. $\mathbf{I}_p^T\mathbf{I}_p = \mathbf{I}$. Helposti on nähtävissä, että myös $\mathbf{I}_p$:n rivit saadaan permutoimalla $\mathbf{I}$:n rivit, eli $\mathbf{I}_p^T = \mathbf{I}_q$, missä q on toinen lukujen $(1, 2, \ldots, n)$ permutaatio ($q = p$, jos $\mathbf{I}_p$ on symmetrinen). Koska $\mathbf{I} = \mathbf{I}_q^T\mathbf{I}_q = \mathbf{I}_p\mathbf{I}_p^T$, niin $\mathbf{I}_p^T\mathbf{I}_p = \mathbf{I}_p\mathbf{I}_p^T = \mathbf{I}$. Siis on päätelty (Lauseeseen X.2.6 vetoamatta), että $\mathbf{I}_p^{-1} = \mathbf{I}_p^T$.

Esimerkki 1 Tapauksessa $n = 3$ erilaisia permutaatiomatriiseja on $3! = 6$ kpl, vastaten permutaatioita $p = (1, 2, 3), (2, 1, 3), (1, 3, 2), (3, 2, 1), (2, 3, 1), (3, 1, 2)$. Näistä ensimmäinen ja kaksi viimeistä ovat parillisia (0 tai 2 parivaihtoa), muut parittomia (1 parivaihto). Kuviosta (merkitty $\bullet = 1$ ja $\cdot = 0$) nähdään, että jos $p = (2, 3, 1)$, niin $\mathbf{I}_p^T = \mathbf{I}_q$, missä $q = (3, 1, 2)$.

$$\begin{bmatrix} \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \end{bmatrix} \begin{bmatrix} \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \end{bmatrix} \begin{bmatrix} \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \end{bmatrix} \qquad \square$$

Jos $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ on matriisi kokoa $n \times n$ ja $\mathbf{I}_p = [\mathbf{e}_{i_1} \ldots \mathbf{e}_{i_n}]$, niin

$$\mathbf{A}\mathbf{I}_p = [\mathbf{A}\mathbf{e}_{i_1} \ldots \mathbf{A}\mathbf{e}_{i_n}] = [\mathbf{a}_{i_1} \ldots \mathbf{a}_{i_n}].$$

Operaatio $\mathbf{A} \mapsto \mathbf{A}\mathbf{I}_p$ siis permutoi $\mathbf{A}$:n sarakkeet $\mathbf{I}_p$:n sarakkeiden mukaiseen järjestykseen p. Vastaavasti koska $(\mathbf{I}_p\mathbf{A})^T = \mathbf{A}^T\mathbf{I}_p^T$, niin päätellään, että operaatio $\mathbf{A} \mapsto \mathbf{I}_p\mathbf{A}$ permutoi $\mathbf{A}^T$:n sarakkeet (eli $\mathbf{A}$:n rivit) $\mathbf{I}_p^T$:n sarakkeiden (eli $\mathbf{I}_p$:n rivien) mukaiseen järjestykseen. Yhteeneveto:

$\mathbf{A} \mapsto \mathbf{A}\mathbf{I}_p$:	$\mathbf{A}$:n sarakkeiden permutointi $\mathbf{I}_p$:n sarakkeiden mukaisesti.
$\mathbf{A} \mapsto \mathbf{I}_p\mathbf{A}$:	$\mathbf{A}$:n rivien permutointi $\mathbf{I}_p$:n rivien mukaisesti.

Esimerkki 1 (jatko)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 2 \\ 6 & 4 & 5 \\ 9 & 7 & 8 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{bmatrix}. \quad \square$$

Edellisen luvun matriisialgebran säännöistä voidaan päätellä, että matriisin sarakkeiden permutointi operaatiolla $\mathbf{A} \mapsto \mathbf{A}\mathbf{I}_p$ toimii yhtä hyvin matriiseille kokoa $m \times n$, $m \neq n$, ja vastaavasti rivien permutointi matriiseille kokoa $n \times m$. Neliömatriisin tapauksessa nähdään myös, että sarakkeiden tai rivien permutoinnissa säännöllinen matriisi säilyy säännöllisenä (koska säännöllisten matriisien tulo on säännöllinen). Samoin singulaarinen matriisi säilyy singulaarisena (Harj.teht. 1d).

Permutaatiomatriisin erikoistapaus on *vaihtomatriisi*, jossa vain kaksi $\mathbf{I}$:n saraketta on vaihdettu keskenään. Jos vaihdetut sarakkeet ovat $\mathbf{e}_k$ ja $\mathbf{e}_m$ $(k < m)$, niin vaihtomatriisi on $\mathbf{V} = [\,\mathbf{e}_1 \ldots \mathbf{e}_{k-1}\ \mathbf{e}_m\ \mathbf{e}_{k+1} \ldots \mathbf{e}_{m-1}\,\mathbf{e}_k\ \mathbf{e}_{m+1} \ldots \mathbf{e}_n\,]$. Tämä on paitsi ortogonaalinen myös symmetrinen:

$$\mathbf{V} = \begin{bmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 0 & \cdots\cdots & & 1 & & & & \\ & & & 1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 1 & & & & \\ & & 1 & \cdots\cdots & & 0 & & & & \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & 1 \end{bmatrix}$$

Koska permutaatioon p päästään peräkkäisillä parivaihdoilla, niin vastaavasti $\mathbf{I}_p$ saadaan kertomalla $\mathbf{I}$ oikealta vastaavilla vaihtomatriiseilla $\mathbf{V}_i$, $i = 1 \ldots m$, eli

$$\mathbf{I}_p = \mathbf{I}\mathbf{V}_1 \ldots \mathbf{V}_m = \mathbf{V}_1 \ldots \mathbf{V}_m.$$

Koska jokainen $\mathbf{V}_i$ on symmetrinen, niin tulon transponointisäännön perusteella

$$\mathbf{I}_p^{-1} = \mathbf{I}_p^T = \mathbf{V}_m \ldots \mathbf{V}_1.$$

Kolmiomatriisit

MÄÄRITELMÄ X.2.11 Neliömatriisi $\mathbf{A} = (a_{ij})$ on

- *yläkolmiomatriisi*, jos $a_{ij} = 0$, kun $i > j$,
- *alakolmiomatriisi*, jos $a_{ij} = 0$, kun $i < j$,
- *diagonaalinen* l. *diagonaalimatriisi*, jos $a_{ij} = 0$, kun $i \neq j$.

$$\begin{bmatrix} \# & \# & \# & \# \\ 0 & \# & \# & \# \\ 0 & 0 & \# & \# \\ 0 & 0 & 0 & \# \end{bmatrix} \qquad \begin{bmatrix} \# & 0 & 0 & 0 \\ \# & \# & 0 & 0 \\ \# & \# & \# & 0 \\ \# & \# & \# & \# \end{bmatrix} \qquad \begin{bmatrix} \# & 0 & 0 & 0 \\ 0 & \# & 0 & 0 \\ 0 & 0 & \# & 0 \\ 0 & 0 & 0 & \# \end{bmatrix}$$

yläkolmio alakolmio diagonaalinen

Määritelmän mukaisesti diagonaalimatriisi on molempien kolmiomatriisien erikoistapaus. Määritelmän matriisityypeille käytetään usein erikoissymboleja $\mathbf{U}$ (yläkolmio, engl. Upper triangular), $\mathbf{L}$ (alakolmio, engl. Lower triangular) ja $\mathbf{D}$ (diagonaalimatriisi). Neliömatriisin $\mathbf{A}$ alkioita a_{ii} sanotaan yleisesti $\mathbf{A}$:n *lävistäjäalkioiksi* ja (järjestettyä) joukkoa $(a_{ii},\ i = 1 \ldots n)$ $\mathbf{A}$:n *lävistäjäksi* eli *diagonaaliksi*. Joukko $(a_{i,i+k},\ i = 1 \ldots n-k)$ on vastaavasti k:s *ylädiagonaali* ja $(a_{i-k,i},\ i = k+1 \ldots n)$ k:s *aladiagonaali*. Diagonaalimatriisi esitetään usein lävistäjäalkioidensa avulla käyttäen merkintää

$$\mathbf{D} = \operatorname{diag}(d_i), \quad d_i = [\mathbf{D}]_{ii}\,.$$

Kolmiomatriisin säännöllisyyskysymyksen ratkaisee

Lause X.2.12 **(Kolmiomatriisin säännöllisyys)** Jos $\mathbf{A} = (a_{ij})$ on kolmiomatriisi kokoa $n \times n$, niin $\mathbf{A}$ on säännöllinen täsmälleen kun lävistäjäalkiot a_{ii} ovat kaikki nollasta poikkeavat, ja tällöin pätee

(i) $\mathbf{A}$ diagonaalinen/yläkolmio/alakolmio
$\Rightarrow \mathbf{A}^{-1}$ diagonaalinen/yläkolmio/alakolmio

(ii) $[\mathbf{A}^{-1}]_{ii} = 1/a_{ii}, \quad i = 1 \ldots n.$

<u>Todistus</u> Diagonaalimatriisin tapauksessa nähdään, että jos $a_{kk} = 0$ jollakin $k \in \{1, \ldots, n\}$, niin yhtälöryhmällä $\mathbf{Ax} = \mathbf{0}$ on monikäsitteinen ratkaisu $\mathbf{x} = x_k\mathbf{e}_k,\ x_k \in \mathbb{R}$, joten Korollaarin X.2.5 mukaan $\mathbf{A}$ on singulaarinen. Jos $a_{ii} \neq 0$, $i = 1 \ldots n$, niin $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ toteutuu väitteen mukaisella valinnalla, eli

$$\mathbf{A}^{-1} = \operatorname{diag}(1/a_{ii}).$$

Oletetaan seuraavaksi, että $\mathbf{A}$ on alakolmiomatriisi, jolloin yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ on auki kirjoitettuna

$$\begin{cases} a_{11}x_1 & = b_1, \\ a_{21}x_1 + a_{22}x_2 & = b_2, \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n & = b_n. \end{cases}$$

Jos $a_{ii} \neq 0\ \forall i$, niin yhtälöryhmä ratkeaa purkamalla se palautuvasti alusta:

$$\begin{aligned} x_1 &= a_{11}^{-1} b_1 = b_{11} b_1, \\ x_2 &= -a_{22}^{-1} a_{21} x_1 + a_{22}^{-1} b_2 = -a_{22}^{-1} a_{21} b_{11} b_1 + a_{22}^{-1} b_2 = b_{21} b_1 + b_{22} b_2, \end{aligned}$$

ja yleisesti (induktio!)

$$x_i = \sum_{j=1}^{i} b_{ij} b_j, \quad i = 1 \ldots n,$$

missä $b_{ii} = a_{ii}^{-1}$. Kun asetetaan $b_{ij} = 0,\ j > i$, ja $\mathbf{B} = \{b_{ij}\}$, niin $\mathbf{B}$ on siis alakolmiomatriisi, lävistäjäalkioin $b_{ii} = a_{ii}^{-1}$, ja yhtälöryhmän ratkaisu on

$$\mathbf{x} = \mathbf{B}\mathbf{b}.$$

Tähän siis päädyttiin olettaen, että $a_{ii} \neq 0\ \forall i$. Jos tämä oletus ei toteudu, niin jollakin $k \in \{1, \ldots, n\}$ pätee: $a_{kk} = 0$ ja $a_{ii} \neq 0,\ i = 1 \ldots k-1$. Silloin nähdään em. algoritmista, että jos valitaan $b_i = 0,\ i = 1 \ldots k-1$, niin on oltava $x_i = 0,\ i = 1 \ldots k-1$, jolloin k:s yhtälö saa muodon $0 = b_k$. Näin ollen yhtälöryhmä $\mathbf{A}\mathbf{x} = \mathbf{b}$ ei yleisesti ratkea, joten Korollaarin X.2.5 perusteella $\mathbf{A}$ on singulaarinen.

Jos $\mathbf{A}$ on yläkolmiomatriisi, niin yhtälöryhmä $\mathbf{A}\mathbf{x} = \mathbf{b}$ purkautuu lopusta lukien: Ratkaistaan ensin n:s yhtälö $a_{nn} x_n = b_n$, sitten $(n-1)$:s yhtälö, jne. Tässäkin tapauksessa päätellään, että $\mathbf{A}$ on singulaarinen, jos $a_{kk} = 0$ jollakin k, muuten löytyy matriisi $\mathbf{B}$, jolle pätee $\mathbf{A}\mathbf{B} = \mathbf{I}$. Matriisi $\mathbf{B}$ on jälleen samaa tyyppiä kuin $\mathbf{A}$ (yläkolmio), ja lävistäjäalkiot ovat $b_{ii} = a_{ii}^{-1}$.

Olkoon nyt $\mathbf{A}$ yleisemmin kolmiomatriisi (ylä- tai alakolmio) ja $\mathbf{A}$:n lävistäjäalkiot nollasta poikkeavat. Tällöin $\mathbf{A}^T$ on myös kolmiomatriisi, jonka lävistäjäalkiot ovat samat kuin $\mathbf{A}$:n. Em. päättelyn mukaan yhtälöryhmät $\mathbf{A}\mathbf{x} = \mathbf{b}$ ja $\mathbf{A}^T\mathbf{x} = \mathbf{b}$ ovat molemmat ratkeavia jokaisella $\mathbf{b} \in \mathbb{R}^n$, joten $\mathbf{A}$ on säännöllinen matriisi (Propositio X.2.9). Kolmiomatriisin säännöllisyyskysymys on näin ratkaistu Lauseesta X.2.6 riippumatta. $\square$

Em. todistuksen sivutuotteena saatiin myös algoritmi kolmiomatriisin $\mathbf{A}$ käänteismatriisin laskemiseksi: Ratkaistaan (todistuksessa esitetyllä tavalla) lineaarinen yhtälöryhmä $\mathbf{A}\mathbf{x} = \mathbf{b}$ <u>yleisellä</u> $\mathbf{b} \in \mathbb{R}^n$. Kun ratkaisu esitetään muodossa

$$x_i = \sum_{j=1}^{n} b_{ij} b_j, \quad i = 1 \ldots n \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{B}\mathbf{b},$$

niin $\mathbf{B} = \mathbf{A}^{-1}$, eli kertoimet b_{ij} ovat käänteismatriisin alkiot. Nämä tulevat algoritmin kuluesssa lasketuksi palautuvasti riveittäin. Vain diagonaalisen (tai muulla tavoin erikoisen, ks. Harj.teht. 7b) matriisin tapauksessa on käänteismatriisin $\mathbf{A}^{-1}$ alkioille mahdollista laskea yksinkertaiset lausekkeet $\mathbf{A}$:n alkioiden avulla.

HARJOITUSTEHTÄVIÄ

1. Olkoon $\mathbf{A}$, $\mathbf{B}$ ja $\mathbf{C}$ samaa kokoa olevia neliömatriiseja. Todista:
   a) $\mathbf{A}$ säännöllinen ja symmetrinen $\Rightarrow$ $\mathbf{A}^{-1}$ symmetrinen.
   b) $\mathbf{CA} = \mathbf{CB}$ ja $\mathbf{C}$ säännöllinen $\Rightarrow$ $\mathbf{A} = \mathbf{B}$.
   c) $\mathbf{AB}$ singulaarinen $\Rightarrow$ $\mathbf{A}$ tai $\mathbf{B}$ singulaarinen.
   d) $\mathbf{A}$ singulaarinen ja $\mathbf{B}$ säännöllinen $\Rightarrow$ $\mathbf{AB}$ ja $\mathbf{BA}$ singulaariset.
   e) $\mathbf{A}$ ja $\mathbf{B}$ ortogonaaliset $\Rightarrow$ $\mathbf{AB}$ ortogonaalinen.

2. Tarkista kokoa 2×2 olevan matriisin käänteismatriisin laskusääntö
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{D}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad D = ad - bc \neq 0.$$

3. a) Näytä, että jokainen ortogonaalinen 2×2-matriisi voidaan kirjoittaa jollakin $\theta \in \mathbb{R}$ jompaan kumpaan seuraavista muodoista:
$$\mathbf{A} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{tai} \quad \mathbf{A} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$
   b) Olkoon $\mathbf{A}$ kokoa 3×3 oleva ortogonaalimatriisi, jonka alkioista tiedetään: $a_{11} = \frac{3}{7}$, $a_{12} = -\frac{2}{7}$, $a_{22} = \frac{6}{7}$, $a_{21} < 0$, $a_{31} > 0$, $a_{13} < 0$. Laske $\mathbf{A}$ ja $\mathbf{A}^{-1}$.
   c) Totea, että yhtälöryhmässä
$$\begin{cases} \sqrt{\frac{2}{3}}\,x_1 + \frac{1}{2}\,x_2 - \frac{1}{2\sqrt{3}}\,x_3 & = 1 \\ \frac{1}{\sqrt{3}}\,x_1 - \frac{1}{\sqrt{2}}\,x_2 + \frac{1}{\sqrt{6}}\,x_3 & = 2 \\ -\frac{1}{2}\,x_2 - \frac{\sqrt{3}}{2}\,x_3 & = 3 \end{cases}$$
   kerroinmatriisi on ortogonaalinen. Ratkaise tätä tietoa käyttäen!
   d) Matriisilla
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$
   on ominaisuus: $\lambda\mathbf{A}$ on ortogonaalinen eräällä $\lambda \in \mathbb{R}$. Määritä tätä tietoa hyväksi käyttäen vaakavektori $\mathbf{x}^T$ siten, että $\mathbf{x}^T\mathbf{A} = [7, 13, -3, -9]$.

4. a) Olkoon $\mathbf{H} = \mathbf{I} - \mathbf{x}\mathbf{x}^T$, missä $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ ja $|\mathbf{x}| \neq 1$. Näytä, että eräällä $\lambda \in \mathbb{R}$ pätee $\mathbf{H}^{-1} = \mathbf{I} + \lambda\mathbf{x}\mathbf{x}^T$.
   b) Olkoon $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ ja $r = 2|\mathbf{x}|^{-2}$. Näytä, että $\mathbf{H} = \mathbf{I} - r\,\mathbf{x}\mathbf{x}^T$ on symmetrinen ja ortogonaalinen matriisi.
   c) Näytä, että jos matriisille $\mathbf{A}$ pätee $\mathbf{A} + \mathbf{A}^T = \mathbf{0}$ ja $\mathbf{I} + \mathbf{A}$ on säännöllinen matriisi, niin $\mathbf{B} = (\mathbf{I} + \mathbf{A})^{-1}(\mathbf{I} - \mathbf{A})$ on ortogonaalinen.

5. Onko joukon $(1, 2, \ldots, 10)$ permutaatio $(4, 6, 2, 7, 9, 1, 3, 10, 5, 8)$ parillinen vai pariton?

6. Matriiseista $\mathbf{V}_1$ ja $\mathbf{V}_2$ kokoa 3×3 tiedetään, että kertoessaan vasemmalta matriisin $\mathbf{A}$ matriisi $\mathbf{V}_1$ vaihtaa $\mathbf{A}$:n rivien 1 ja 2 järjestyksen ja $\mathbf{V}_2$ rivien 2 ja 3 järjestyksen. Millaisen rivien permutoinnin silloin tuottavat $\mathbf{V}_1\mathbf{V}_2$ ja $\mathbf{V}_2\mathbf{V}_1$? Määritä myös $\mathbf{V}_1$ ja $\mathbf{V}_2$ sekä mainitut tulot.

7. a) Olkoon $\mathbf{A} = (a_{ij})$ kokoa $n \times n$ ja $\mathbf{D} = \operatorname{diag}\{d_i,\ i = 1 \ldots n\}$. Määritä matriisien $\mathbf{AD}$ ja $\mathbf{DA}$ alkiot.
   b) Olkoon $n, k \in \mathbb{N},\ n \geq 2,\ k \leq n,\ \mathbf{A} = (a_{ij})$ kokoa $n \times n$, $a_{ii} = a \neq 0$ kun $i \neq k$, $a_{kk} = b \neq 0$, $a_{ik} = c \neq 0$ kun $i > k$ ja $a_{ij} = 0$ muulloin. Määritä käänteismatriisin $\mathbf{A}^{-1}$ alkiot.

8. Määritä seuraavien kolmiomatriisien käänteismatriisit ratkaisemalla yleinen lineaarinen yhtälöryhmä, jonka kerroinmatriisina on ko. matriisi.

   a) $\begin{bmatrix} 3 & 0 \\ -5 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 4 & 0 & 0 \\ -2 & 2 & 0 \\ 3 & 1 & -1 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 0 & -1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

   d) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ f) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

9. (*) Näytä Lauseeseen X.2.6 vedoten, että neliömatriiseille pätee:
   a) $\mathbf{AB}$ säännöllinen $\Rightarrow$ $\mathbf{A}$ ja $\mathbf{B}$ säännölliset.
   b) $\mathbf{A}^2 + 3\mathbf{A} + 2\mathbf{I} = \mathbf{0}$ $\Rightarrow$ $\mathbf{A}$ säännöllinen.

10. (*) Näytä, että on olemassa permutaatiomatriisit $\mathbf{U}$ ja $\mathbf{V}$ siten, että pätee

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 0 \end{bmatrix} = \mathbf{U} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{V}.$$

    Laske käänteismatriisi $\mathbf{A}^{-1}$ tämän tiedon avulla. Tarkista, että $\mathbf{AA}^{-1} = \mathbf{I}$.

11. (*) Olkoon $a, b \in \mathbb{R}$ ja $b \neq 0$. Näytä matriisialgebran avulla, että pätee: Differentiaaliyhtälöllä $y'' + ay' + by = x^{100}$ on yksittäisratkaisuna polynomi astetta 100.

X.3 Gaussin algoritmi

Lineaarisia yhtälöryhmiä, jotka eivät ole kooltaan aivan hirmuisia (esim. $n < 10^4$) ratkotaan käytännössä usein *Gaussin algoritmilla* eli *Gaussin eliminaatiolla*. Algoritmilla on monia variaatioita, ja niillä monia nimiä, mutta pääidea on kaikissa sama.

Lähdetään ratkaisemaan yhtälöryhmää

$$\begin{cases} a_{11}x_1 + a_{12}\,x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}\,x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\ a_{i1}x_1 + a_{i2}\;x_2 + \ldots + a_{in}x_n = b_i, \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{cases}$$

Gaussin algoritmissa oletetaan, että yhtälöryhmällä on ratkaisu (x_i) ja toimitaan ikäänkuin ratkaisu olisi sijoitettu tuntemattomien paikalle. Ratkaisun ei tarvitse olla yksikäsitteinen — tämä selviää algoritmin kuluessa. Samoin selviää, onko ratkaisua ylipäänsä olemassa.

Gaussin algoritmissa yhtälöryhmä muunnetaan askelittain yksinkertaisempaan muotoon. Muunnetut yhtälöryhmät ovat alkuperäisen kanssa ekvivalenttisia (eli muunnokset voidaan tehdä molempiin suuntiin), joten mitään informaatiota ei laskentaprosessin kuluessa menetetä.

Olettaen, että $a_{11} \neq 0$ (tapaukseen $a_{11} = 0$ palataan myöhemmin), Gaussin algoritmin ensimmäinen nk. *eliminaatioaskel* on: Ratkaistaan 1. yhtälöstä

$$x_1 = -\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}}\,x_j + \frac{b_1}{a_{11}}$$

ja sijoitetaan tämä lauseke muihin yhtälöihin x_1:n paikalle. Sievennysten jälkeen on tuloksena alkuperäisen kanssa ekvivalenttinen yhtälöryhmä

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}\,x_n = b_1, \\ \qquad\qquad a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n = b_2^{(1)}, \\ \qquad\qquad \quad\vdots \qquad\qquad\quad \vdots \qquad \vdots \\ \qquad\qquad a_{i2}^{(1)}x_2 + \ldots + a_{in}^{(1)}x_n = b_i^{(1)}, \\ \qquad\qquad \quad\vdots \qquad\qquad\quad \vdots \qquad \vdots \\ \qquad\qquad a_{n2}^{(1)}x_2 + \ldots + a_{nn}^{(1)}x_n = b_n^{(1)}. \end{cases}$$

Tässä ensimmäinen yhtälö on jätetty alkuperäiseen muotoonsa ja on merkitty

$$\begin{aligned} a_{ij}^{(1)} &= a_{ij} - a_{i1}a_{1j}/a_{11}, \quad i,j = 2\ldots n, \\ b_i^{(1)} &= b_i - a_{i1}b_1/a_{11}, \quad i = 2\ldots n. \end{aligned}$$

Taulukkomuodossa laskien tämä muunnos toteutuu, kun kerrotaan 1. yhtälö puolittain luvulla $-a_{i1}/a_{11}$, lasketaan tulos puolittain yhteen i:nnen yhtälön kanssa, ja korvataan tuloksella aiempi i:s yhtälö, $i = 2\ldots n$.

Ensimmäisen eliminaatioaskeleen jälkeinen yhtälöryhmä on matriisimuodossa

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix}.$$

Koska tässä yhtälöryhmässä x_1 esiintyy vain ensimmäisessä yhtälössä, voidaan tämä yhtälö jättää jatkossa omaan rauhaansa ja tarkastella supistettua yhtälöryhmää

$$\begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & \ldots & a_{2n}^{(1)} \\ a_{33}^{(1)} & a_{33}^{(1)} & \ldots & a_{3n}^{(1)} \\ \vdots & \vdots & & \vdots \\ a_{n2}^{(1)} & a_{n3}^{(1)} & \ldots & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_2^{(1)} \\ b_3^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix}.$$

Jos $a_{22}^{(1)} \neq 0$, suoritetaan tähän yhtälöryhmään samanlainen eliminaatiomuunnos kuin edellä, jolloin se muuntuu muotoon

$$\begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & \ldots & a_{2n}^{(1)} \\ 0 & a_{33}^{(2)} & \ldots & a_{3n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n3}^{(2)} & \ldots & a_{nn}^{(2)} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(2)} \end{bmatrix}.$$

Tässä x_2 esiintyy vain ensimmäisessä yhtälössä, joten yhtälöryhmä supistuu jälleen. Yleisesti on $k-1$ eliminaatioaskeleen jälkeen koko yhtälöryhmä saatu muotoon

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1k} & \ldots & a_{1n} \\ 0 & a_{22}^{(1)} & \ldots & a_{2k}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & & & \\ 0 & 0 & \ldots & a_{kk}^{(k-1)} & \ldots & a_{kn}^{(k-1)} \\ \vdots & \vdots & & \vdots & & \\ 0 & 0 & \ldots & a_{nk}^{(k-1)} & \ldots & a_{nn}^{(k-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_k^{(k-1)} \\ \vdots \\ b_n^{(k-1)} \end{bmatrix}.$$

Edellytyksenä tähän muotoon pääsemiseksi on siis, että

$$a_{ll}^{(l-1)} \neq 0, \quad l = 1 \ldots k-1,$$

missä $a_{11}^{(0)} = a_{11}$. Alkiota $a_{kk}^{(k-1)}$ sanotaan seuraavan eli k:nnen eliminaatioaskeleen *tukialkioksi* (engl. pivot element). Mikäli kaikki tukialkiot ovat nollasta poikkeavia ja myös $a_{nn}^{(n-1)} \neq 0$, päädytään $n-1$ askeleen jälkeen yhtälöryhmään, jonka kerroinmatriisi on säännöllinen kolmiomatriisi:

$$\begin{bmatrix} a_{11} & a_{22} & \ldots & a_{1n} \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn}^{(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(k-1)} \end{bmatrix}.$$

Kuten pian nähdään, algoritmin läpimeno tällä tavoin osoittaa, että alkuperäisen yhtälöryhmän kerroinmatriisi $\mathbf{A}$ on säännöllinen. Joka tapauksessa muunnettu yhtälöryhmä ratkeaa helposti purkamalla se lopusta lukien, vrt. Lauseen X.2.12 todistus edellä. Tätä algoritmin vaihetta kutsutaan *takaisinsijoitukseksi.*

Esimerkki 1 (Ks. Luku X.8, Esimerkki 2) Ratkaise Gaussin eliminaatiolla yhtälöryhmä

$$\begin{bmatrix} 3 & -1 & 0 & 0 & -1 & 1 \\ 1 & -3 & 0 & 2 & -1 & 1 \\ 0 & 0 & 3 & 1 & -1 & -1 \\ 0 & -2 & 1 & 3 & -1 & -1 \\ 1 & -1 & 1 & 1 & -4 & 0 \\ -1 & 1 & 1 & 1 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = F \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + G \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

missä $F, G \in \mathbb{R}$.

<u>Ratkaisu</u> Ratkaisu on $\mathbf{x} = F\mathbf{x}_1 + G\mathbf{x}_2$, missä $\mathbf{A}\mathbf{x}_1 = \mathbf{b}_1 = [0\ 1\ 0\ 0\ 0\ 0]^T$ ja $\mathbf{A}\mathbf{x}_2 = \mathbf{b}_2 = [0\ 0\ 0\ 1\ 0\ 0]^T$, joten on ratkaistava kaksi lineaarista yhtälöryhmää. Gaussin algoritmilla voidaan molemmat yhtälöryhmät käsitellä samanaikaisesti: Kirjoitetaan yhtälöryhmät taulukkomuotoon

$$\left[\begin{array}{rrrrrr|rr} 3 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & -3 & 0 & 2 & -1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & -1 & 0 & 0 \\ 0 & -2 & 1 & 3 & -1 & -1 & 0 & 1 \\ 1 & -1 & 1 & 1 & -4 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 & -4 & 0 & 0 \end{array}\right]$$

ja ulotetaan eliminaatio koko taulukkoon. Seuraavassa muunnettu taulukko kunkin eliminaatioaskeleen jälkeen.

$$(k=1)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & \frac{2}{3} & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & -1 & 0 & 0 \\ 0 & -2 & 1 & 3 & -1 & -11 & 0 & 1 \\ 0 & -\frac{2}{3} & 1 & 1 & -\frac{11}{3} & -\frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 1 & 1 & -\frac{1}{3} & -\frac{11}{3} & 0 & 0 \end{array}\right]$$

$$(k=2)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & \frac{2}{3} & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{4} & 1 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{7}{2} & -\frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{2} & -\frac{7}{2} & \frac{1}{4} & 0 \end{array}\right]$$

$$(k=3)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & \frac{2}{3} & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{1}{6} & -\frac{7}{6} & -\frac{3}{4} & 1 \\ 0 & 0 & 0 & \frac{1}{6} & -\frac{19}{6} & -\frac{1}{6} & -\frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{1}{6} & -\frac{19}{6} & \frac{1}{4} & 0 \end{array}\right]$$

$$(k=4)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & \frac{2}{3} & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{1}{6} & -\frac{7}{6} & -\frac{3}{4} & 1 \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & -1 \end{array}\right]$$

Askelta $k=5$ ei tässä tapauksessa tarvita, koska on $a_{65}^{(4)}=0$.

Ratkaisu saadaan tästä takaisinsijoituksella. — Käytetään tässä kuitenkin toista, laskutyön kannalta samanarvoista mahdollisuutta, jossa eliminaatioalgoritmia käytetään uudelleen 'takaperin'. Koska kerroinmatriisi on yläkolmio, niin takaisin päin eliminoitaessa muuttuu matriisissa kullakin eliminaatioaskeleella vain tukialkion kohdalla oleva sarake, tai tarkemmin tämän lävistäjän yläpuolinen

osa. Eliminaatiovaiheet takaisin päin ovat seuraavat.

$$(k=5)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & -1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & 0 & \frac{4}{3} & -\frac{1}{3} \\ 0 & 0 & 3 & 1 & -1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{1}{6} & 0 & -\frac{4}{3} & \frac{19}{12} \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & -1 \end{array}\right]$$

$$(k=4)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & 0 & 0 & \frac{6}{11} & -\frac{5}{11} \\ 0 & -\frac{8}{3} & 0 & 2 & 0 & 0 & \frac{15}{11} & -\frac{10}{33} \\ 0 & 0 & 3 & 1 & 0 & 0 & -\frac{5}{11} & \frac{6}{11} \\ 0 & 0 & 0 & \frac{7}{6} & 0 & 0 & -\frac{175}{132} & \frac{35}{22} \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & -1 \end{array}\right]$$

$$(k=3)\qquad \left[\begin{array}{cccccc|cc} 3 & -1 & 0 & 0 & 0 & 0 & \frac{6}{11} & -\frac{5}{11} \\ 0 & -\frac{8}{3} & 0 & 0 & 0 & 0 & \frac{40}{11} & -\frac{100}{33} \\ 0 & 0 & 3 & 0 & 0 & 0 & \frac{15}{22} & -\frac{9}{11} \\ 0 & 0 & 0 & \frac{7}{6} & 0 & 0 & -\frac{175}{132} & \frac{35}{22} \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & -1 \end{array}\right]$$

$$(k=2)\qquad \left[\begin{array}{cccccc|cc} 3 & 0 & 0 & 0 & 0 & 0 & -\frac{9}{11} & \frac{15}{22} \\ 0 & -\frac{8}{3} & 0 & 0 & 0 & 0 & \frac{40}{11} & -\frac{100}{33} \\ 0 & 0 & 3 & 0 & 0 & 0 & \frac{15}{22} & -\frac{9}{11} \\ 0 & 0 & 0 & \frac{7}{6} & 0 & 0 & -\frac{175}{132} & \frac{35}{22} \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & -1 \end{array}\right]$$

Lopuksi diagonaalisen systeemin ratkaisu:

$$\left[\begin{array}{cccccc|cc} 1 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{11} & \frac{5}{22} \\ 0 & 1 & 0 & 0 & 0 & 0 & -\frac{15}{11} & \frac{25}{22} \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{5}{22} & -\frac{3}{11} \\ 0 & 0 & 0 & 1 & 0 & 0 & -\frac{25}{22} & \frac{30}{22} \\ 0 & 0 & 0 & 0 & 1 & 0 & \frac{1}{22} & \frac{1}{22} \\ 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{array}\right].$$

Kysytty yhtälöryhmän ratkaisu on näin muodoin

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \frac{F}{22} \begin{bmatrix} -6 \\ -30 \\ 5 \\ -25 \\ 1 \\ -11 \end{bmatrix} + \frac{G}{22} \begin{bmatrix} 5 \\ 25 \\ -6 \\ 30 \\ 1 \\ 11 \end{bmatrix}. \qquad \square$$

Gaussin algoritmin työmäärä

Gaussin algoritmi koostuu siis kahdesta vaiheesta, eliminaatiovaiheesta ja takaisinsijoituksesta. Englannin kielessä käytetään myös termejä 'forward sweep' ja 'backward sweep'. Näistä ensimmäinen 'pyyhkäisy' on työmäärää (laskenta-aikaa) ajatellen huomattavasti raskaampi. Tämä nähdään eliminaatiovaiheen yleisestä algoritmimuodosta, joka on

$$\begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} - a_{ik}^{(k-1)} a_{kj}^{(k-1)} / a_{kk}^{(k-1)}, \quad i,j = k+1 \ldots n, \\ b_i^{(k)} &= b_i^{(k-1)} - a_{ik}^{(k-1)} b_k^{(k-1)} / a_{kk}^{(k-1)}, \quad i = k+1 \ldots n. \end{aligned}$$

Tässä on jakolaskut $a_{ik}^{(k-1)}/a_{kk}^{(k-1)}$, $i = k+1 \ldots n$ syytä laskea ensin, jolloin algoritmi on ohjelmointikielellä

$$\begin{aligned} l_{ik} &\leftarrow a_{ik}/a_{kk}\,, & i &= k+1 \ldots n, \\ a_{ij} &\leftarrow a_{ij} - l_{ik} a_{kj}\,, & i,j &= k+1 \ldots n, \\ b_i &\leftarrow b_i - l_{ik} b_k\,, & i &= k+1 \ldots n. \end{aligned}$$

Työmäärää arvioitaessa pidettäköön työyksikkönä laskuoperaatiota, joka koostuu yhdestä kertolaskusta (tai jakolaskusta) ja yhdestä yhteen- tai vähennyslaskusta. Algoritmista nähdään, että algoritmin työläin osa koostuu muunnoksista $a_{ij}^{(k)} \leftarrow a_{ij}^{(k-1)}$. Näiden vaatima työmäärä koko eliminaatiovaiheessa on mainittuina operaatioina

$$W = (n-1)^2 + (n-2)^2 + \cdots + 1 = \frac{1}{6}(n-1)n(2n-1).$$

Muut algoritmin osat, kuten muunnokset $b_i^{(k)} \leftarrow b_i^{(k-1)}$ ja takaisinsijoitukset ovat työmäärältään suuruusluokkaa $\mathcal{O}(n^2)$, joten Gaussin algoritmin työmäärä on

$$\boxed{W = \frac{1}{3} n^3 + \mathcal{O}(n^2).}$$

Tästä nähdään, että esimerkiksi tuhannen yhtälön yhtälöryhmän ($W \approx 0.3 \cdot 10^9$) ratkaiseminen tietokoneella ei ole ongelma laskenta-ajan kannalta. Sen sijaan yhtälöryhmä kokoa $n = 10^6$ ($W \approx 0.3 \cdot 10^{18}$) saattaa jo olla ylivoimainen tehtävä. — Näinkin suuria yhtälöryhmiä ratkotaan nykyisin, mutta silloin käytetään yleensä hyväksi kerroinmatriisin erikoisominaisuuksia. Toinen, perinteisesti suosittu vaihtoehto on käyttää tehtävään sopivia *iteratiivisia* (likimääräisiä) ratkaisutapoja. Suhteessa iteratiivisiin menetelmiin sanotaan Gaussin algoritmiin perustuvaa ratkaisua *suoraksi* ratkaisutavaksi. Kohtuullisilla n:n arvoilla yhtälöryhmän suora ratkaisu Gaussin algoritmilla on edelleen kilpailukykyinen menetelmä, ja tietokoneiden suorituskyvyn kasvaessa se on myös voittanut alaa iteratiivisilta menetelmiltä. Etenkin silloin kun algoritmi menee läpi em. suoraviivaisella tavalla se on myös hyvin helppo ohjelmoida.

Neliömatriisin LU-hajotelma

Gaussin eliminaatiolla, sikäli kuin se onnistuu edellä kuvatulla tavalla, saadaan yhtälöryhmälle $\mathbf{Ax} = \mathbf{b}$ ratkaisu jokaisella $\mathbf{b} \in \mathbb{R}^n$, joten Lauseen X.2.6 perusteella kerroinmatriisin on oltava säännöllinen. Ko. lauseeseen (jota ei vielä ole todistettu) ei kuitenkaan tarvitse vedota, sillä kerroinmatriisin säännöllisyys tulee Gaussin algoritmilla suoraan todistetuksi. Päättely on seuraava: Ensinnäkin todetaan, että Gaussin algoritmin eliminaatiovaihe muuntaa alkuperäisen yhtälöryhmän

$$\mathbf{A}\ \mathbf{x} = \mathbf{b}$$

muotoon

$$\mathbf{Ux} = \mathbf{c} = \mathbf{Rb},$$

missä $\mathbf{U}$ on muunnettu kerroinmatriisi, eli yläkolmiomatriisi, jonka lävistäjäalkiot ovat $a_{kk}^{(k-1)} \neq 0,\ k = 1 \ldots n$. Matriisi $\mathbf{R}$ ei ole eliminaatiovaiheen jälkeen suoraan nähtävissä, mutta se saadaan selville tarkastelemalla yhtälöryhmää

$$\mathbf{Ly} = \mathbf{b},$$

missä $\mathbf{L}$ on alakolmiomatriisi, joka määritellään

$$[\mathbf{L}]_{ik} = \begin{cases} 0, & \text{jos } i < k, \\ 1, & \text{jos } i = k, \\ l_{ik} = a_{ik}^{(k-1)} / a_{kk}^{(k-1)}, & \text{jos } i > k. \end{cases}$$

Tässä siis $a_{ik}^{(k-1)}$, $i = k \dots n$, ovat muunnetun kerroinmatriisin alkioita $k-1$ eliminaatioaskeleen jälkeen:

$$\begin{array}{ll} a_{k-1,k-1}^{(k-2)} & \\ 0 & a_{kk}^{(k-1)} \\ \vdots & \vdots \\ 0 & a_{ik}^{(k-1)} \quad \to \; l_{ik} = a_{ik}^{(k-1)}/a_{kk}^{(k-1)} \\ \vdots & \vdots \\ 0 & a_{nk}^{(k-1)} \end{array}$$

Matriisin $\mathbf{L}$ muodosta nähdään, että jos em. yhtälöryhmään sovelletaan Gaussin algoritmia, niin jo eliminaatiovaihe antaa ratkaisun $\mathbf{y} = \mathbf{L}^{-1}\mathbf{b}$. Toisaalta nähdään vertaamalla Gaussin algoritmeja sovellettuna yhtälöryhmiin $\mathbf{Ly} = \mathbf{b}$ ja $\mathbf{Ax} = \mathbf{b}$, että eliminaatiovaiheen muunnokset $b_i^{(k)} \leftarrow b_i^{(k-1)}$ ovat kummassakin tapauksessa täsmälleen samat. Siis on päätelty:

$$\mathbf{y} = \mathbf{L}^{-1}\mathbf{b} = \mathbf{Rb}.$$

Koska tämä pätee $\forall \mathbf{b} \in \mathbb{R}^n$, on $\mathbf{R} = \mathbf{L}^{-1}$ (Propositio X.1.1), ja näin ollen

$$\mathbf{Ax} = \mathbf{b} \;\Leftrightarrow\; \mathbf{LUx} = \mathbf{b}$$

Siis $\mathbf{Ax} = \mathbf{LUx}$ $\forall \mathbf{x} \in \mathbb{R}^n$, joten

$$\boxed{\mathbf{A} = \mathbf{LU}.}$$

Tätä sanotaan matriisin *LU-hajotelmaksi* (engl. *LU*-decomposition). Sen laskemiseksi ei siis tarvita muuta kuin Gaussin algoritmin eliminaatiovaihe, kunhan muistetaan tallettaa matriisi $\mathbf{L}$, eli luvut

$$[\mathbf{L}]_{ik} = l_{ik} = a_{ik}^{(k-1)}/a_{kk}^{(k-1)}, \quad i = k+1 \dots n, \; k = 1 \dots n-1.$$

Nämä tulevat eliminaation kuluessa joka tapauksessa lasketuksi, mutta ilman talletusta $\mathbf{L}$ ei jää näkyviin (koska yhtälöryhmän ratkaisualgoritmi laskee valmiiksi vektorin $\mathbf{L}^{-1}\mathbf{b} = \mathbf{Rb}$).

Esimerkki 1 (jatko) Seuraamalla eliminaatiovaiheita ($k = 1 \dots 4$) nähdään, että esimerkin yhtälöryhmässä kerroinmatriisin *LU*-hajotelma on

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{3} & \frac{1}{7} & 1 & 0 \\ -\frac{1}{3} & -\frac{1}{4} & \frac{1}{3} & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 0 & 0 & -1 & 1 \\ 0 & -\frac{8}{3} & 0 & 2 & -\frac{2}{3} & \frac{2}{3} \\ 0 & 0 & 3 & 1 & -1 & -1 \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{1}{6} & -\frac{7}{6} \\ 0 & 0 & 0 & 0 & -\frac{22}{7} & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix} = \mathbf{LU}. \quad \square$$

Matriisin LU-hajotelma siis onnistuu Gaussin eliminaatiolla, jos $a_{kk}^{(k-1)} \neq 0,\ k = 1 \ldots n$. Tällöin $\mathbf{L}$ ja $\mathbf{U}$ ovat säännöllisiä (Lause X.2.12), joten $\mathbf{A}$ on säännöllinen ja

$$\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}.$$

Mainittu tukialkioita koskeva ehto on siis riittävä ehto matriisin säännöllisyydelle. Seuraavassa luvussa nähdään, että kun Gaussin algoritmia hieman muunnellaan, se ratkaisee neliömatriisin säännöllisyyskysymyksen yleisessäkin tapauksessa.

Käänteismatriisin laskeminen

Sikäli kuin Gaussin eliminaatio yhtälöryhmälle $\mathbf{Ax} = \mathbf{b}$ onnistuu edellä kuvatulla tavalla (tukialkiot $\neq 0$), niin kerroinmatriisi on siis säännöllinen, jolloin yhtälöryhmän ratkaisu voidaan kirjoittaa $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Jos halutaan laskea myös käänteismatriisi $\mathbf{A}^{-1}$ eikä vain ratkaisua $\mathbf{x}$, niin eräs (etenkin käsinlaskussa luonteva) menettely on soveltaa Gaussin algoritmia <u>symbolisesti</u> siten, että $\mathbf{b}$:n alkioille b_i ei anneta numeroarvoja. Algoritmia käyttäen ratkaisu saadaan tällöin muotoon $\mathbf{x} = \mathbf{Bb}$, jolloin on $\mathbf{A}^{-1} = \mathbf{B}$.

Esimerkki 2 Laske matriisin

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

käänteismatriisi.

<u>Ratkaisu</u> Eliminaatiolla ja takaisinsijoituksella saadaan

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\mapsto \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ -b_1 + b_2 \\ 3b_1 - 2b_2 + b_3 \end{bmatrix}$$

$$\mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} b_1 + b_2 + b_3 \\ 2b_1 - b_2 - b_3 \\ -b_1 + 2b_2 - b_3 \end{bmatrix}.$$

Tuloksesta voidaan lukea käänteismatriisi:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{A}^{-1}\mathbf{b}. \quad \square$$

Edellistä suoraviivaisempi tapa laskea käänteismatriisi on käyttää hyväksi matriisialgebran identieettiä (vrt. Luku X.1)

$$\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}[\mathbf{e}_1 \dots \mathbf{e}_n] = [\mathbf{A}^{-1}\mathbf{e}_1 \dots \mathbf{A}^{-1}\mathbf{e}_n].$$

Tämän mukaisesti $\mathbf{A}^{-1}$:n sarakkeet voidaan laskea ratkaisemalla lineaarinen yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ kun $\mathbf{b} = \mathbf{e}_i,\ i = 1 \dots n$. Koska Gaussin algoritmissa voidaan käsitellä yhtä aikaa useita eri vektoreita $\mathbf{b}$ (vrt. Esimerkki 1), voidaan myös $\mathbf{A}^{-1}$:n sarakkeet määrätä kaikki yhdellä kertaa. Käänteismatriisin laskeminen voidaan tällöin kuvata kaksivaiheisena operaationa (eliminaatio ja takaisinsijoitus)

$$[\,\mathbf{A} \mid \mathbf{I}\,] \mapsto [\,\mathbf{U} \mid \mathbf{L}^{-1}\,] \mapsto [\,\mathbf{I} \mid \mathbf{A}^{-1}\,].$$

Vastaavalla tavalla menetellään itse asiassa em. symbolisessa laskussa, sillä kun kirjoitetaan $\mathbf{b} = \sum_{i=1}^{n} b_i\mathbf{e}_i$, niin pätee (vrt. Luku X.1)

$$\mathbf{A}^{-1}\mathbf{b} = \sum_{i=1}^{n} b_i\mathbf{A}^{-1}\mathbf{e}_i = [\mathbf{A}^{-1}\mathbf{e}_1 \dots \mathbf{A}^{-1}\mathbf{e}_n]\,\mathbf{b}.$$

Esimerkki 2 (jatko) Esimerkin lasku toisin organisoituna:

$$[\,\mathbf{A} \mid \mathbf{I}\,] = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -3 & 1 & -2 & 1 \end{array}\right]$$

$$\mapsto \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{array}\right] = [\,\mathbf{I} \mid \mathbf{A}^{-1}\,]. \quad \square$$

Käänteismatriisi voidaan myös laskea käyttäen LU-hajotelmaan perustuvaa kaavaa $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$, eli laskemalla ensin ym. tavalla erikseen käänteismatriisit $\mathbf{L}^{-1}$ ja $\mathbf{U}^{-1}$ ja sitten näiden tulo. Sekä tällä tavoin että ym. algoritmin mukaisesti laskien tulee työmääräksi $W = n^3 + \mathcal{O}(n^2)$ (Harj.teht. 6) — siis noin kolminkertainen työmäärä verrattuna LU-hajotelman laskemiseen tai yksittäisen yhtälöryhmän ratkaisuun. Osoittautuukin, että yhtälöryhmiä numeerisesti ratkaistaessa ei käänteismatriisin laskeminen yksinkertaisesti kannata. Nimittäin vaikka ratkaistaisiin useita yhtälöryhmiä kerroinmatriisin pysyessä samana (esim. vaihteleva

kuormitus ristikkorakenteessa, vrt. Luku X.8), niin Gaussin algoritmin eliminaatiovaiheen toistaminen pystytään välttämään pelkän LU-hajotelman avulla: Kun hajotelman matriisit $\mathbf{L}$ ja $\mathbf{U}$ tallennetaan, niin yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ voidaan purkaa kahdeksi peräkkäiseksi yhtälöryhmäksi:

$$\mathbf{LUx} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{Ly} = \mathbf{b}, \quad \mathbf{Ux} = \mathbf{y}.$$

Nämä ratkeavat eteen- ja takaisinsijoituksilla, jolloin yhtälöryhmän ratkaisemisen työmääräksi tulee $W \sim n^2$ (ks. Harj.teht. 5b). Jos $\mathbf{A}^{-1}$ olisi tallennettu, olisi työmäärä sama, joten käänteismatriisin laskemisella ei voiteta mitään (!).

HARJOITUSTEHTÄVIÄ

1. Kirjoita seuraavat yhtälöryhmät taulukkomuotoon ja ratkaise Gaussin algoritmilla. Laske myös kerroinmatriisien LU-hajotelmat.

a) $\begin{cases} x + y = 33 \\ 2x + 4y = 100 \end{cases}$ b) $\begin{cases} x + ay = 1 \\ ax + y = 2 \end{cases}$ $(a \neq \pm 1)$

c) $\begin{cases} x_1 + 3x_2 - x_3 = 4 \\ -x_1 + x_2 + 2x_3 = 7 \\ 2x_1 + 6x_2 + 2x_3 = 20 \end{cases}$ d) $\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + x_2 - x_3 = 2 \\ x_1 - 6x_3 = 4 \end{cases}$

e) $\begin{cases} x_1 + x_2 + 2x_3 = -1 \\ -x_1 + x_2 = 1 \\ 3x_1 - x_2 + 3x_3 = -3 \end{cases}$ f) $\begin{cases} x_1 - x_2 + 4x_3 = 8 \\ x_1 - 2x_2 + 3x_3 = 0 \\ -2x_1 + 4x_2 - 5x_3 = 8 \end{cases}$

g) $\begin{cases} 2x_1 - x_2 = 1 \\ -x_1 + 2x_2 - x_3 = 2 \\ -x_2 + 2x_3 - x_4 = -1 \\ -x_3 + 2x_4 = 1 \end{cases}$ h) $\begin{cases} 2x_1 - x_2 - x_4 = -4 \\ -x_1 + 2x_2 - x_3 = 1 \\ -x_2 + 2x_3 - x_4 = 4 \\ -x_1 - x_3 + 3x_4 = 10 \end{cases}$

2. Olkoon

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 5 \\ 3 \\ 4 \\ 0 \end{bmatrix}$$

Ratkaise Gaussin algoritmilla yhtälöryhmät $\mathbf{Ax} = \mathbf{b}_1$ ja $\mathbf{Ax} = \mathbf{b}_2$ ja laske matriisin $\mathbf{A}$ LU-hajotelma.

3. Olkoon
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 1 \end{bmatrix}.$$
Laske käänteismatriisi $\mathbf{A}^{-1}$ Gaussin algoritmilla
a) ratkaisemalla yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ yleisellä $\mathbf{b} \in \mathbb{R}^4$ (symbolinen lasku),
b) ratkaisemalla taulukkomuodossa yhtälöryhmät $\mathbf{Ax} = \mathbf{e}_i,\ i = 1 \ldots 4$

4. Näytä, että jos lineaarisen yhtälöryhmän kerroimatriisi $\mathbf{A}$ on symmetrinen, niin Gaussin algoritmi säilyttää symmetrisenä sen matriisin osan, johon eliminaatio ei ole vielä edennyt, ts. $(k-1)$:n eliminaatioaskeleen jälkeen pätee
$$a_{ji}^{(k-1)} = a_{ij}^{(k-1)} \quad \forall\, i, j \geq k.$$
Päättele, että Gaussin algoritmin vaatima laskutyö on symmetrisen kerroinmatriisin tapauksessa $W = \frac{1}{6}\, n^3 + \mathcal{O}(n^2)$.

5. Olkoon $\mathbf{A}, \mathbf{B}$ matriiseja kokoa $n \times n$, $\mathbf{L}, \mathbf{U}$ ala- ja yläkolmiomatriiseja samaa kokoa ja $\mathbf{b} \in \mathbb{R}^n$. Näytä oikeaksi seuraavat laskuoperaatioiden työmääriä koskevat arviot (työyksikkö = kertolasku + yhteenlasku).

a)	$\mathbf{A}, \mathbf{b} \mapsto \mathbf{Ab}$	$n^2 + \mathcal{O}(n)$
b)	$\mathbf{L}, \mathbf{b} \mapsto \mathbf{L}^{-1}\mathbf{b}$	$\frac{1}{2}\, n^2 + \mathcal{O}(n)$
c)	$\mathbf{A}, \mathbf{B} \mapsto \mathbf{AB}$	$n^3 + \mathcal{O}(n^2)$
d)	$\mathbf{L}, \mathbf{U} \mapsto \mathbf{LU}$	$\frac{1}{3}\, n^3 + \mathcal{O}(n^2)$

6. (*) Näytä, että laskuoperaation $\mathbf{A} \mapsto \mathbf{A}^{-1}$ työmäärä Gaussin algoritmin eri vaihtoehdoissa on ($\mathbf{A}$ kokoa $n \times n$, työyksikkö = kertolasku + yhteenlasku)

a)	$[\,\mathbf{A} \mid \mathbf{I}\,] \mapsto [\,\mathbf{I} \mid \mathbf{A}^{-1}\,]$	$n^3 + \mathcal{O}(n^2)$
b)	$\mathbf{A} \mapsto \mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$	$n^3 + \mathcal{O}(n^2)$

X.4 Tuettu Gaussin algoritmi. Singulaariset yhtälöryhmät

Jos Gaussin algoritmissa tullaan $k-1$ eliminaatioaskeleen jälkeen tilanteeseen, jossa tukialkio $a_{kk}^{(k-1)} = 0$, niin algoritmissa päästään eteenpäin käyttämällä nk. *tuentaa* (engl. pivoting). Tuennassa yksinkertaisesti vaihdetaan matriisin rivien ja/tai sarakkeiden järjestystä (vastaten yhtälöiden tai tuntemattomien järjestyksen vaihtoa yhtälöryhmässä). Pyrkimyksenä on löytää jokin sellainen järjestys, jossa tukialkiosta tulee nollasta poikkeava, jolloin Gaussin algoritmia voidaan jatkaa. Tarkastellaan aluksi nk. *osittaista* eli *rivituentaa*, jossa menettely on seuraava: Käydään läpi k:nnessa sarakkeessa lävistäjän alapuolella olevat alkiot

$$a_{ik}^{(k-1)}, \quad i = k+1 \ldots n.$$

Valitaan (jos mahdollista) indeksi $i = l$ siten, että $a_{lk}^{(k-1)} \neq 0$, ja suoritetaan yhtälöryhmässä (eli matriisissa $\mathbf{A}^{(k-1)}$ ja vektorissa $\mathbf{b}^{(k-1)}$) rivinvaihto

$$k\text{:s rivi} \;\leftrightarrows\; l\text{:s rivi}.$$

Tämä vastaa yhtälöryhmän yhtälöiden järjestyksen vaihtoa. Gaussin algoritmista nähdään, että sikäli kuin rivien uudelleen järjestely suoritetaan riveillä $i = k \ldots n$, ei järjestyksen vaihto häiritse itse algoritmia. Toisin sanoen, lopputulos on sama, jos rivit vaihdetaan jo alkuperäisessä yhtälöryhmässä ja sen jälkeen suoritetaan $k-1$ eliminaatioaskelta (Harj.teht. 1a). Tukioperaation jälkeen uusi tukialkio on nollasta poikkeava, jolloin Gaussin algoritmia voidaan jatkaa.[†]

Em. tukioperaatio epäonnistuu vain siinä tapauksessa, että on

$$a_{ik}^{(k-1)} = 0, \quad i = k \ldots n.$$

Gaussin algoritmi on tällöin osoittanut alkuperäisen matriisin $\mathbf{A}$ singulaariseksi. Nimittäin koska tällöin tuntematon x_k ei esiinny lainkaan muunnetun yhtälöryhmän riveillä $i = k \ldots n$, niin yhtälöryhmälle $\mathbf{Ax} = \mathbf{0}$ saadaan monikäsitteinen ratkaisu antamalla x_k:lle mielivaltainen arvo, asettamalla $x_{k+1} = \ldots = x_n = 0$, ja ratkaisemalla x_i, $i = 1 \ldots k-1$ muunnetusta yhtälöryhmästä takaisinsijoituksella (x_k:n avulla). Siis $\mathbf{A}$ on singulaarinen (Korollaari X.2.5).

[†] *Automaattisessa* rivituennassa suoritetaan rivien vaihto ennen jokaista eliminaatioaskelta niin, että tukialkioksi tulee alkioista $a_{ik}^{(k-1)}$, $i = k \ldots n$ itseisarvoltaan suurin. Tämä vähentää pyöristysvirheiden kasautumista ratkaisulagoritmin kuluessa (liukulukulaskennassa). Monissa tietokoneohjelmistojen 'black box'–ratkaisijoissa käytetään automaattista rivituentaa (esim. Mathematica: `LinearSolve`).

Osittaisen tukioperaation epäonnistuttua (ja siis matriisin osoittauduttua singulaariseksi) voidaan Gaussin algoritmia vielä jatkaa käyttämällä nk. *täydellistä tuentaa*, jossa käydään läpi kaikki matriisialkiot rivistä k ja sarakkeesta k alkaen, eli alkiot

$$a_{ij}^{(k-1)}, \quad i,j = k \dots n.$$

Oletetaan, että jollakin (l,m) on $a_{lm}^{(k-1)} \neq 0$. (Mahdollisesti valitaan alkioista itseisarvoltaan suurin kuten automaattisessa rivituennassa, vrt. alaviite edellä.)

$$\begin{array}{ccccc} 0 & \dots & a_{km}^{(k-1)} & \dots & a_{kn}^{(k-1)} \\ \vdots & & \vdots & & \\ a_{lk}^{(k-1)} & \dots & a_{lm}^{(k-1)} & & \\ \vdots & & & & \\ a_{nk}^{(k-1)} & & & & \end{array}\Bigg]$$

Suoritetaan tällöin rivien ja sarakkeiden vaihto

$$k\text{:s rivi} \leftrightarrows l\text{:s rivi}, \quad k\text{:s sarake} \leftrightarrows m\text{:s sarake}.$$

Sarakkeiden vaihto vastaa yhtälöryhmässä tuntemattomien uudelleen indeksointia. Indekseihin $j \geq k$ sovellettuna tämäkään tukioperaatio ei häiritse itse algoritmia, ts. operaatio voidaan ajatella suoritetuksi jo yhtälöryhmän alkuperäisessä muodossa (Harj.teht. 1a). Sikäli kuin tukioperaatio onnistuu oletetulla tavalla, on operaation jälkeen uusi tukialkio $a_{lm}^{(k-1)} \neq 0$, jolloin Gaussin algoritmia voidaan jatkaa. Muussa tapauksessa on tultu tilanteeseen, jossa $a_{ij}^{(k-1)} = 0$ kun $i = k \dots n$ ja $j = 1 \dots n$. Gaussin algoritmi on tällöin päättynyt *perusmuotoiseen singulaariseen yhtälöryhmään*, joka on siis jollakin $k \in \{0, \dots, n-1\}$ muotoa

$$\begin{bmatrix} u_{11} & \dots & & & u_{1n} \\ 0 & \ddots & & & \\ \vdots & & u_{kk} & \dots & u_{kn} \\ 0 & & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ \vdots \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \\ \vdots \\ \vdots \\ c_n \end{bmatrix}.$$

Tässä $\mathbf{U} = \{u_{ij}\}$ on yläkolmiomatriisi, jonka $n-k$ viimeistä riviä ovat nollarivejä ja $u_{ii} \neq 0$, kun $i = 1 \dots k$ (jos $k = 0$, niin $\mathbf{U} = \mathbf{A} = \mathbf{0}$). Kertoimet u_{ij} ja c_i on saatu alkuperäisestä yhtälöryhmästä $\mathbf{Ax} = \mathbf{b}$ tuetulla Gaussin algoritmilla, eli eliminaatioita ja tukioperaatioita (rivien ja sarakkeiden vaihtoja) yhdistelemällä. Vektori $\mathbf{y}$ sisältää alkuperäiset tuntemattomat x_i uudessa, sarakkeiden vaihtojen

määrämässä järjestyksessä. Jos algoritmin kuluessa on käytetty vain rivituentaa, on $\mathbf{y} = \mathbf{x}$.

Yllämainitusta perusmuotoisesta yhtälöryhmästä nähdään, että sillä on ratkaisu täsmälleen kun

$$c_{k+1} = \ldots = c_n = 0.$$

Tällöin ratkaisu on monikäsitteinen: $y_i,\ i = k+1 \ldots n$, voidaan valita mielivaltaisesti, minkä jälkeen $y_1, \ldots, y_k$ määräytyvät takaisinsijoituksella.

Esimerkki 1 Seuraavassa sovelletaan täysin tuettua Gaussin algoritmia singulaariseen yhtälöryhmään.

$$\begin{bmatrix} 0 & 0 & 3 \\ 1 & 2 & 0 \\ 2 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_3 \\ b_2 \\ b_1 \end{bmatrix} \qquad (\text{rivien vaihto } 1 \leftrightarrows 3, \quad \text{sarakkeiden vaihto } 1 \leftrightarrows 2\,)$$

$$\begin{bmatrix} 4 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_3 \\ b_2 - b_3/2 \\ b_1 \end{bmatrix} \qquad (\text{eliminaatio})$$

$$\begin{bmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} b_3 \\ b_1 \\ b_2 - b_3/2 \end{bmatrix} \qquad (\text{rivien ja sarakkeiden vaihto } 2 \leftrightarrows 3\,)$$

Nähdään, että saadulla perusmuotoisella yhtälöryhmällä on ratkaisu täsmälleen, kun

$$b_2 - b_3/2 = 0 \quad \Leftrightarrow \quad b_3 = 2b_2,$$

ja yleinen ratkaisu on tällöin

$$\begin{cases} x_1 &= t, \\ x_2 &= -t/2 + b_3/4 \;=\; (b_2 - t)/2, \\ x_3 &= b_1/3 \end{cases} \qquad (t \in \mathbb{R}). \qquad \square$$

Esimerkin tuloksen olisi voinut nähdä suoremminkin. Yleisen algoritmin etuna on kuitenkin, että se toimii aina, myös silloin kun laskija ei 'näe'.

Neliöatriisin yleinen tulohajotelma

Tuetunkin Gaussin algoritmin avulla voidaan määrätä matriisin $\mathbf{A}$ käänteismatriisi, sikäli kuin $\mathbf{A}$ on säännöllinen. Kun tukioperaatiot ajatellaan tehdyksi ennen Gaussin algoritmin soveltamista, niin ensinnäkin rivituennat vastaavat yhtälöryhmän muunnosta

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{VAx} = \mathbf{Vb},$$

missä $\mathbf{V} = \mathbf{I}_p$ on rivien lopullista järjestystä p vastaava permutaatiomatriisi. — Kerrattakoon Luvusta X.2, että jos rivinvaihdot vastaavat vaihtomatriiseja $\mathbf{V}_i,\ i = 1 \ldots m$, niin

$$\mathbf{V} = \mathbf{V}_m \cdots \mathbf{V}_2 \mathbf{V}_1$$

(käänteinen järjestys!) ja $\mathbf{V}^{-1} = \mathbf{V}^T = \mathbf{V}_1 \mathbf{V}_2 \cdots \mathbf{V}_m$. Sarakkeiden vaihtojen lopputulos taas vastaa muunnosta $\mathbf{y} = \mathbf{W}^T\mathbf{x} \ \Leftrightarrow\ \mathbf{x} = \mathbf{Wy}$, missä permutaatiomatriisi $\mathbf{W} = \mathbf{I}_q$ vastaa sarakkeiden lopullista järjestystä q. Jos sarakkeiden vaihtoja vastaavat vaihtomatriisit ovat $\mathbf{W}_i,\ i = 1 \ldots l$, niin

$$\mathbf{W} = \mathbf{W}_1 \mathbf{W}_2 \ldots \mathbf{W}_l$$

(sama järjestys kuin vaihdoissa!) ja $\mathbf{W}^{-1} = \mathbf{W}^T = \mathbf{W}_l \mathbf{W}_{l-1} \ldots \mathbf{W}_1$. Kun kaikki tuennat huomioidaan, niin yhtälöryhmä muuntuu muotoon

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{VAx} = \mathbf{Vb} \quad \Leftrightarrow \quad (\mathbf{VAW})(\mathbf{W}^T\mathbf{x}) = \mathbf{Vb} \quad \Leftrightarrow \quad \mathbf{By} = \mathbf{c},$$

missä $\mathbf{c} = \mathbf{Vb}$ (yhtälöt vaihdetussa järjestyksessä p), $\mathbf{y} = \mathbf{W}^T\mathbf{x}$ (tuntemattomat vaihdetussa järjestyksessä q) ja

$$\mathbf{B} = \mathbf{VAW} \ \Leftrightarrow\ \mathbf{A} = \mathbf{V}^T\mathbf{BW}^T.$$

Kun muunnettuun yhtälöryhmään $\mathbf{By} = \mathbf{c}$ sovelletaan Gaussin algoritmia, ei tuentaa enää tarvita. Eliminaation lopputuloksena $\mathbf{B}$ muuntuu yläkolmiomatriisiksi $\mathbf{U}$, jolloin on kaksi vaihtoehtoa:

a) $[\mathbf{U}]_{kk} \neq 0,\ k = 1 \ldots n$. Tällöin eliminaatiovaihe menee läpi kokonaisuudessaan, $\mathbf{U}$ on säännöllinen, $\mathbf{B}$:lle saadaan LU-hajotelma

$$\mathbf{B} = \mathbf{LU},$$

ja $\mathbf{A}$:lle vastaavasti tulohajotelma

$$\boxed{\mathbf{A} = \mathbf{V}^T\mathbf{LUW}^T.} \qquad (\star)$$

Tällöin $\mathbf{A}$ on säännöllinen ja

$$\mathbf{A}^{-1} = \mathbf{WU}^{-1}\mathbf{L}^{-1}\mathbf{V} = \mathbf{WB}^{-1}\mathbf{V}.$$

b) $[\mathbf{U}]_{ii} \neq 0,\ i = 1 \ldots k < n,\ [\mathbf{U}]_{i,i,} = 0,\ i = k \ldots n$ (ol. $\mathbf{A} \neq \mathbf{0}$). Tällöin $\mathbf{U}$ on em. singulaarista perusmuotoa, eli $[\mathbf{U}]_{ij} = 0$, kun $i > k,\ j = 1 \ldots n$, ja eliminaatioalgoritmi pysähtyy riville k. Pysähtyminen vastaa sitä, että matriisissa $\mathbf{L}$ astetetaan

$$[\mathbf{L}]_{ij} = \delta_{ij}, \quad \text{kun} \quad i \in \{k+1, \ldots, n\},\ j \in \{1, \ldots, n\}.$$

Kun $\mathbf{L}$ määrätään muilta osin samalla tavoin kuin säännöllisen matriisin tapauksessa, nähdään, että tulohajotelma ($\star$) on edelleen pätevä.

Kysymys neliömatriisin säännöllisyydestä tai singulaarisuudesta on näin muodoin ratkaistu Gaussin algoritmilla seuraavasti: Jokainen neliömatriisi voidaan esittää tulomuodossa ($\star$), missä $\mathbf{U}$ on yläkolmiomatriisi ja $\mathbf{L}, \mathbf{V}, \mathbf{W}$ ovat säännöllisiä matriiseja. Matriisi $\mathbf{U}$ on joko säännöllinen tai singulaarinen, mikä ratkaisee myös $\mathbf{A}$:n laadun.

Esimerkki 1 (jatko) Tässä on

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 2 & 0 \\ 2 & 4 & 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \mathbf{VAW} = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{bmatrix},$$

$$\mathbf{y} = \mathbf{W}^T\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix}, \quad \mathbf{c} = \mathbf{Vb} = \begin{bmatrix} b_3 \\ b_1 \\ b_2 \end{bmatrix}.$$

Muunnetusta yhtälöryhmästä

$$\begin{bmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} b_3 \\ b_1 \\ b_2 \end{bmatrix}$$

päästään yläkolmiomuotoon yhdellä eliminaatioaskeleella. Koska tukioperaatiot on suoritettu jo etukäteen, ei niitä enää eliminaation yhteydessä tarvita. □

Säännöllisyysaste

Jos $\mathbf{A}$ on singulaarinen, niin em. tuetun Gaussin algoritmin lopetusindeksiä k sanotaan $\mathbf{A}$:n *säännöllisyysasteeksi* eli *rangiksi* (engl. rank), merkitään $k =$

$r(\mathbf{A})$. Luku $n - r(\mathbf{A})$ siis kertoo, kuinka monta nollariviä on $\mathbf{A}$:n tulohajotelman perusmuotoisessa singulaarisessa yläkolmiomatriisissa $\mathbf{U}$. Jos $\mathbf{A}$ on säännöllinen, niin $r(\mathbf{A}) = n$, jolloin myös $r(\mathbf{A}^T) = n$. Yleisemminkin pätee

$$\boxed{r(\mathbf{A}^T) = r(\mathbf{A}).}$$

Myös tämän voi perustella Gaussin algoritmin antaman tulohajotelman avulla (sivuutetaan perustelut).

Yhtälöryhmä kokoa $m \times n$

Gaussin algoritmi soveltuu myös sellaiseen lineaariseen yhtälöryhmään $\mathbf{Ax} = \mathbf{b}$, jossa yhtälöiden ja tuntemattomien lukumäärät eivät ole samat, ts. $\mathbf{A}$ on kokoa $m \times n,\ m \neq n$ ($\mathbf{b} \in \mathbb{R}^m$). Kun tuettua Gaussin algoritmia sovelletaan tällaiseen yhtälöryhmään, on lopputulos jälleen muunnettu perusmuotoinen yhtälöryhmä $\mathbf{Uy} = \mathbf{c}$, missä $\mathbf{U} = \{u_{ij},\ i = 1 \ldots m,\ j = 1 \ldots n\}$ on yläkolmiomatriisi (ts. $u_{ij} = 0$ kun $i > j$), ja jollakin $k,\ 0 \leq k \leq \min\{m, n\}$, pätee

$$u_{ii} \neq 0,\ \text{kun } i = 1 \ldots k, \quad u_{ij} = 0,\ \text{kun } i = k+1 \ldots m,\ j = 1 \ldots n.$$

Indeksiä k sanotaan jälleen $\mathbf{A}$:n *säännöllisyysasteeksi* eli *rangiksi.* Jos k on suurin mahdollinen, eli jos $k = \min\{m, n\}$, niin sanotaan, että $\mathbf{A}$:lla on *täysi* säännöllisyysaste (engl. full rank). Tässä tapauksessa (täydellisesti tuettua) Gaussin eliminaatiota voidaan jatkaa, kunnes algoritmi törmää matriisin $\mathbf{A}$ 'laitaan' ($m > n$) tai 'pohjaan' ($m < n$). Muunnettu yhtälöryhmä $\mathbf{Uy} = \mathbf{c}$ voidaan esittää *lohkomatriisi*muodossa

$$\begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix},$$

missä $\mathbf{U}_{11} = \{u_{ij},\ i, j = 1 \ldots k\}$ on säännöllinen yläkolmiomatriisi, $\mathbf{U}_{12}$ on kokoa $k \times (n-k)$, $\mathbf{U}_{21}$ on nollamatriisi kokoa $(m-k) \times k$, $\mathbf{U}_{22}$ on nollamatriisi kokoa $(m-k) \times (n-k)$, ja $\mathbf{c}_1$ on k-vektori. Koska $\mathbf{U}_{21}$ ja $\mathbf{U}_{22}$ ovat nollamatriiseja, niin yhtälöryhmän avautuu muotoon

$$\begin{cases} \mathbf{U}_{11}\,\mathbf{y}_1 + \mathbf{U}_{12}\,\mathbf{y}_2 &= \mathbf{c}_1, \\ \mathbf{0} &= \mathbf{c}_2. \end{cases}$$

Tästä nähdään, että sikäli kuin $k < m$ (näin on aina kun $m > n$), on yhtälöryhmällä sekä välttämätön että riittävä

$$\textit{ratkeavuusehto}: \quad \mathbf{0} = \mathbf{c}_2.$$

Jos $k < n$ (näin on aina kun $m < n$), ei ratkaisu ole tälläkään ehdolla yksikäsitteinen, sillä ratkaisussa voidaan asettaa $\mathbf{y}_1 = \mathbf{U}_{11}^{-1}(\mathbf{c}_1 - \mathbf{U}_{12}\,\mathbf{y}_2)$, olipa $\mathbf{y}_2 \in \mathbb{R}^{n-k}$ mikä tahansa. Päätellään siis erityisesti, että yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ kokoa $m \times n$, $m \neq n$, on aina singulaarinen: Joko ratkaisua ei ole jokaisella $\mathbf{b} \in \mathbb{R}^m$, tai ratkaisu ei ole yksikäsitteinen. Edellinen tilanne vallitsee aina kun $m > n$, jälkimmäinen aina kun $m < n$.

Esimerkki 2 Määritä singulaarisen yhtälöryhmän

$$\left\{\begin{array}{rrrcl} x_1 & +x_2 & +x_3 & = & b_1 \\ 2x_1 & +x_2 & & = & b_2 \\ -x_1 & & +x_3 & = & b_3 \\ & +x_2 & +2x_3 & = & b_4 \\ x_1 & -x_2 & -3x_3 & = & b_5 \end{array}\right.$$

ratkeavuusehdot, yleinen ratkaisu ja kerroinmatriisin säännöllisyysaste.

<u>Ratkaisu</u> Tässä on $m = 5$ ja $n = 3$. Gaussin algoritmissa selvitään ilman tuentaa:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} \longmapsto \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{bmatrix} \begin{bmatrix} b_1 \\ -2b_1 + b_2 \\ b_1 + b_3 \\ b_4 \\ -b_1 + b_5 \end{bmatrix}$$

$$\longmapsto \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ -2b_1 + b_2 \\ -b_1 + b_2 + b_3 \\ -2b_1 + b_2 + b_4 \\ 3b_1 - 2b_2 + b_5 \end{bmatrix}$$

Saadun singulaarisen perusmuodon mukaan kerroinmatriisin säännöllisyysaste on $k = 2$. Ratkeavuusehdot voidaan lukea yhtälöryhmän kolmelta viimeiseltä riviltä:

$$\left\{\begin{array}{rrrrrcl} -b_1 & +b_2 & +b_3 & & & = & 0 \\ -2b_1 & +b_2 & & +b_4 & & = & 0 \\ 3b_1 & -2b_2 & & & +b_5 & = & 0 \end{array}\right.$$

eli

$$\begin{bmatrix} 1 & 0 & 0 & -2 & 3 \\ 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} b_5 \\ b_4 \\ b_3 \\ b_2 \\ b_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Koska tässä vasemmalla oleva matriisi on perusmuotoinen yläkolmio (ellei olisi, sovellettaisiin ensin Gaussin algoritmia), niin nähdään, että b_1 ja b_2 voidaan valita vapaasti, minkä jälkeen b_3, b_4 ja b_5 määräytyvät:

$$\begin{cases} b_3 = b_1 - b_2, \\ b_4 = 2b_1 - b_2, \\ b_5 = -3b_1 + 2b_2. \end{cases}$$

Näiden ehtojen voimassa ollessa saadaan yleinen ratkaisu takaisinsijoituksella:

$$\begin{cases} x_1 = -b_1 + b_2 + t, \\ x_2 = 2b_1 - b_2 - 2t, \qquad (t \in \mathbb{R}). \\ x_3 = t \end{cases}$$

Ratkaisu ei ole yksikäsitteinen, koska kerroinmatriisin säännöllisyysaste ei ole täysi: $r(\mathbf{A}) = 2 < \min\{m, n\} = 3$. $\square$

Lauseen X.2.6 todistus

Tulohajotelman ($\star$) perusteella pätee

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{Uy} = \mathbf{Cb}, \qquad \mathbf{y} = \mathbf{Wx}, \;\; \mathbf{C} = \mathbf{L}^{-1}\mathbf{V}.$$

Jos $\mathbf{A}$ on singulaarinen, niin $\mathbf{U}$:n riveistä ainakin n:s on nollarivi, jolloin yhtälöryhmän $\mathbf{Uy} = \mathbf{Cb}$ n:s yhtälö saa muodon $0 = [\mathbf{Cb}]_n$. Jos on $\mathbf{b} = \mathbf{0}$, niin yhtälöryhmällä on monikäsitteinen ratkaisu, missä y_n on vapaasti valittavissa. Jokaista ratkaisua $\mathbf{y} \neq \mathbf{0}$ ($y_n \neq 0$) vastaa yhtälöryhmän $\mathbf{Ax} = \mathbf{0}$ ratkaisu $\mathbf{x} = \mathbf{W}^T\mathbf{y} \neq \mathbf{0}$. Jos taas $\mathbf{b} \in \mathbb{R}^n$ valitaan siten, että $[\mathbf{Cb}]_n \neq 0$, niin mainittu yhtälö ei toteudu, jolloin on $\mathbf{Uy} \neq \mathbf{Cb}$ $\forall \mathbf{y} \in \mathbb{R}^n$. Tällöin on myös $\mathbf{Ax} \neq \mathbf{b}$, $\mathbf{x} = \mathbf{W}^T\mathbf{y}$, $\mathbf{y} \in \mathbb{R}^n$ (koska $\mathbf{Uy} = \mathbf{Cb} \;\Leftrightarrow\; \mathbf{Ax} = \mathbf{b}$). Siis $\mathbf{Ax} \neq \mathbf{b}$ $\forall \mathbf{x} \in \mathbb{R}^n$. On päätelty:

$$\mathbf{A} \text{ singulaarinen} \quad \Rightarrow \quad \begin{cases} \text{jollakin } \mathbf{x} \neq \mathbf{0} \text{ on} & \mathbf{Ax} = \mathbf{0}, \\ \text{jollakin } \mathbf{b} \in \mathbb{R}^n \text{ on} & \mathbf{Ax} \neq \mathbf{b} \;\; \forall \mathbf{x} \in \mathbb{R}^n. \end{cases}$$

Nämä väittämät ovat samanarvoisia kuin Lauseen X.2.6 osaväittämät E1 $\Rightarrow$ E0 ja E2 $\Rightarrow$ E0, jotka siis ovat tosia. Koska E0 $\Rightarrow$ E1 ja E0 $\Rightarrow$ E2 olivat tosia jo Proposition X.2.4 mukaan, niin lause on todistettu. $\square$

HARJOITUSTEHTÄVIÄ

1. a) Näytä, että tuetun Gaussin algoritmin kuluessa suoritetut tukioperaatiot (rivien ja sarkkeiden vaihdot) voidaan suorittaa ennen eliminaatioita lopputuloksen muuttumatta.

   b) Olkoon $\mathbf{A}$ matriisi kokoa $n \times n$ ja olkoon $\mathbf{A}_k$ matriisi kokoa $k \times k$, joka koostuu $\mathbf{A}$:n alkioista $a_{ij},\ i, j = 1 \ldots k$. Näytä, että yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$ ratkeaa Gaussin algoritmilla ilman tukioperaatioita täsmälleen kun $\mathbf{A}_k$ on säännöllinen matriisi jokaisella $k = 1 \ldots n$.

2. Muunna singulaariseen perusmuotoon tuetulla Gaussin algoritmilla, määritä ratkaisut tai totea ratkeamattomuus:

   a) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 3 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 12 \end{bmatrix}$

   c) $\begin{bmatrix} 0 & 1 & -1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 1 & 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

   d) $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 1 & 1 & 2 & 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

3. Määritä $\mathbf{A}$:n säännöllisyysaste, kaikki yhtälöryhmän $\mathbf{Ax} = \mathbf{0}$ ratkaisut sekä tulohajotelma $\mathbf{A} = \mathbf{V}^T\mathbf{LUW}^T$, kun $\mathbf{A} =$

   a) $\begin{bmatrix} 1 & 2 & -1 \\ -2 & -4 & 2 \\ -1 & -2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 3 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 2 & 1 \\ 3 & 3 & 2 \\ 4 & 4 & 3 \end{bmatrix}$

   d) $\begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 0 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$ f) $\begin{bmatrix} 1 & -2 & 1 & 0 \\ -1 & 2 & -1 & 1 \\ 2 & -4 & 2 & 4 \\ 1 & 1 & -1 & -1 \end{bmatrix}$

4. Saata seuraavat yhtälöryhmät tuetulla Gaussin algoritmilla singulaariseen perusmuotoon. Määritä ratkeavuusehdot ja yleinen ratkaisu sekä edelleen kerroinmatriisin tulohajotelma $\mathbf{A} = \mathbf{V}^T\mathbf{L}\mathbf{U}\mathbf{W}^T$.

a) $\begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & 1 \\ 5 & -10 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & -4 & -3 \\ 1 & 0 & 2 & 1 \\ 3 & 2 & -2 & -3 \\ 2 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 2 & -1 & 2 & 2 \\ 2 & 4 & -2 & 0 & 4 \\ -1 & -2 & 1 & -2 & -2 \\ -1 & -2 & 0 & -3 & -3 \\ -2 & -4 & 4 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$

5. Muunna singulaariseen perusmuotoon ja ratkaise, mikäli mahdollista:

a) $\begin{bmatrix} 1 & 0 & 4 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 1 \\ 4 & 6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 6 \\ 3 & 6 & 8 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 1 \end{bmatrix}$ d) $\begin{bmatrix} 2 & 1 & 4 \\ 6 & 3 & 8 \\ 4 & 2 & 7 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 1 \\ 1 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 4 & 7 & -3 \\ -2 & 3 & -6 & 1 \\ 0 & 11 & 8 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

6. Olkoon a) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix}$, b) $\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 & -2 \\ 3 & -2 & 7 & 5 \\ 2 & -5 & 2 & 7 \end{bmatrix}$.
Määritä Gaussin algoritmilla ratkeavuusehdot ja yleinen ratkaisu yhtälöryhmille $\mathbf{A}\mathbf{x} = \mathbf{b}$ ($\mathbf{b} \in \mathbb{R}^3$) ja $\mathbf{A}^T\mathbf{x} = \mathbf{b}$ ($\mathbf{b} \in \mathbb{R}^4$).

X.5 Determinantti

Olkoon $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ n kappaletta pystyvektoreita kokoa n. Asetetaan

MÄÄRITELMÄ X.5.1 Reaaliarvoinen funktio $V(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $\mathbf{a}_i \in \mathbb{R}^n$, on (normeerattu) *determinanttifunktio*, jos pätee:

(i) V on lineaarinen jokaisen vektorin $\mathbf{a}_i$ suhteen.

(ii) Jos järjestetty indeksijoukko $(i_1, \ldots, i_n)$ on saatu joukosta $(1, 2, \ldots, n)$ vaihtamalla kahden alkion paikkaa, niin

$$V(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_n}) = -V(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

(iii) Jos $[\mathbf{e}_i]_j = \delta_{ij}$, niin $V(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$.

Selityksiä:

(i) Lineaarisuus tarkoittaa, että jokaisella $k \in \{1, \ldots, n\}$ ja $\forall \lambda, \mu \in \mathbb{R}$ pätee

$$\begin{aligned} &V(\mathbf{a}_1, \ldots, \lambda \mathbf{a}_k + \mu \mathbf{b}_k, \ldots, \mathbf{a}_n) \\ &\qquad = \lambda V(\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n) + \mu V(\mathbf{a}_1, \ldots, \mathbf{b}_k, \ldots, \mathbf{a}_n). \end{aligned}$$

(ii) Tämä *vaihtosääntö* tarkoittaa, että determinantin arvo vaihtuu vastaluvukseen, kun järjestetyssä joukossa $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ kahden vektorin paikka vaihdetaan (Luvusta X.2 tuttu parivaihto). Jos vaihdettavat vektorit ovat samat, ei determinantin arvo luonnollisesti muutu, joten vaihtosäännöstä seuraa *nollasääntö*

$$\mathbf{a}_i = \mathbf{a}_j\,, \;\; i \neq j \quad \Rightarrow \quad V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0.$$

Jos yleisemmin $p = (i_1, \ldots, i_n)$ on indeksijoukon $(1, \ldots, n)$ permutaatio, niin p on joko parillinen tai pariton, eli p saadaan joko parilllisella tai parittomalla määrällä parivaihtoja (vrt. Luku X.2). Vaihtosääntö yleistyy siis säännöksi: Parillisessa permutaatiossa determinantin arvo säilyy, parittomassa vaihtuu vastaluvukseen.

(iii) Tämä on normeerausehto. Ilman tätä ehtoa determinanttifunktio voitaisiin kertoa mielivaltaisella vakiolla, jolloin seuraava väittämä ei olisi tosi.

PROPOSITIO X.5.2 Determinanttifunktio on funktio, ts. $V(\mathbf{a}_1, \ldots \mathbf{a}_n) \in \mathbb{R}$ on yksikäsitteisesti määrätty, kun $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$.

Todistus Lasketaan determinanttifunktion arvo annetuilla säännöillä (i)–(iii). Merkitään $[\mathbf{a}_j]_i = a_{ij}$, jolloin

$$\mathbf{a}_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i$$

ja siis

$$V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = V\Big(\sum_{i=1}^{n} a_{i1}\mathbf{e}_i, \ldots, \sum_{i=1}^{n} a_{in}\mathbf{e}_i\Big).$$

Käyttämällä toistuvasti lineaarisuusominaisuutta (i) nähdään, että tämä purkautuu summaksi

$$V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} \cdots a_{i_n n} V(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}).$$

Tässä on n^n termiä, mutta nollasäännön perusteella summasta voidaan jättää pois kaikki ne termit, joissa sama indeksi i_k esiintyy kahdesti. Näin muodoin

$$V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{p} a_{i_1 1} \cdots a_{i_n n} V(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}),$$

missä summaus käy läpi kaikki joukon $(1, \ldots, n)$ eri permutaatiot $p = (i_1, \ldots, i_n)$ ($n!$ kpl). Säännöistä (ii) ja (iii) seuraa edelleen, että $V(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = \sigma_p = \pm 1$ riippuen siitä, onko permutaatio p parillinen vai pariton. Siis

$$\boxed{V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{p} \sigma_p \, a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}} \tag{1}$$

ja näin muodoin $V(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ on yksikäsitteisesti määrätty. □

Determinanttifunktion ominaisuudet (i)–(iii) voi päätellä käänteisesti laskukaavasta (1), joten tämä kaava on käypä myös determinanttifunktion määritelmänä. Kaavalle saadaan hieman selkeämpi muoto, kun merkitään (vrt. Luku X.2)

$$\mathbf{I}_p = [\mathbf{e}_{i_1} \ldots \mathbf{e}_{i_n}], \quad \Lambda_p = \{(i,j) \mid [\mathbf{I}_p]_{ij} = 1\}, \quad p = (i_1, \ldots, i_n).$$

Tällöin nähdään, että kaava (1) on sama kuin

$$V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{p} \sigma_p \prod_{(i,j)\in\Lambda_p} a_{ij}. \tag{2}$$

Determinantin arvoa käytännössä laskettaessa kaavat (1)–(2) ovat petollisia, sillä pelkkiä kertolaskuja tarvitaan niitä käyttäen peräti $(n-1)n!$ kpl. — Kaavoja käytetäänkin algoritmeina yleensä vain tapauksissa $n = 2$ ja $n = 3$. Teoreettista käyttöä kaavoilla (1)–(2) on sen sijaan yleisemminkin, kuten nähdään jatkossa.

Matriisin determinantti

Matriisin determinantti määritellään yksinkertaisesti ajattelemalla, että determinanttifunktiossa $V(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ vektorit $\mathbf{a}_j$ ovat (neliö)matriisin $\mathbf{A}$ sarakkeet. Determinanttia merkitään $\det \mathbf{A}$ (toisinaan $|\mathbf{A}|$), ja määritelmä on siis

$$\det \mathbf{A} = V(\mathbf{a}_1, \ldots, \mathbf{a}_n), \quad \mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n].$$

Tämän mukaisesti on kaavassa (2) $\sigma_p = \det \mathbf{I}_p$.

Determinantti on siis määritelty vain neliömatriiseille. Jos $\mathbf{A}$ on annettu taulukkona, niin determinanttia merkitään taulukkoa rajoittavin pystyviivoin:

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}, \quad \mathbf{A} = (a_{ij}).$$

Normeeraussäännön (iii) mukaan yksikkömatriisin determinantti on

$$\det \mathbf{I} = 1.$$

Kaavasta (1) seuraa myös helposti sääntö

$$\det(\lambda \mathbf{A}) = \lambda^n \det \mathbf{A}, \quad \lambda \in \mathbb{R}.$$

Myös seuraavat kaksi determinantin ominaisuutta ovat määritelmästä johdettavissa, mutta nämä ovat vähemmän ilmeisiä. — Kyseessä ovat determinanttiopin keskeisimmät väittämät.

Lause X.5.3 Matriisin determinantille pätee

$$\boxed{\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}, \qquad \det \mathbf{A}^T = \det \mathbf{A}.}$$

Todistus Olkoon $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ ja $\mathbf{B} = [\mathbf{b}_1 \ldots \mathbf{b}_n] = (b_{ij})$. Käyttämällä Luvun X.1 purkukaavoja voidaan kirjoittaa

$$\mathbf{AB} = [\mathbf{Ab}_1 \ldots \mathbf{Ab}_n] = \Big[\sum_{i_1=1}^{n} b_{i_1 1} \mathbf{a}_{i_1} \ \ldots \ \sum_{i_n=1}^{n} b_{i_n n} \mathbf{a}_{i_n} \Big],$$

jolloin determinantin lineaarisuussäännön (i) perusteella seuraa

$$\det \mathbf{AB} = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1 1} \ldots b_{i_n n} \det [\mathbf{a}_{i_1} \ldots \mathbf{a}_{i_n}].$$

Tästä summasta voidaan nollasäännön perusteella jälleen jättää pois kaikki termit, joissa sama indeksi i_k esiintyy kahdesti. Jäljelle jäävissä termeissä on $p = (i_1, \ldots, i_n)$ joukon $(1, \ldots, n)$ permutaatio, jolloin vaihtosäännön (ii) mukaan

$$\det [\mathbf{a}_{i_1} \ldots \mathbf{a}_{i_n}] = \sigma_p \det [\mathbf{a}_1 \ldots \mathbf{a}_n] = \sigma_p \det \mathbf{A}, \quad p = (i_1, \ldots, i_n).$$

Tässä σ_p on määritelty kuten kaavassa (1), joten

$$\det \mathbf{AB} = \det \mathbf{A} \sum_p \sigma_p \, b_{i_1 1} \ldots b_{i_n n} = \det \mathbf{A} \det \mathbf{B}.$$

Determinantin transponointisäännön todistamiseksi toetaan ensinnäkin, että tämä pätee kaikille permutaatiomatriiseille $\mathbf{I}_p$. Nimittäin koska $\mathbf{I}_p^T \mathbf{I}_p = \mathbf{I}$, niin jo todistetun tulon determinanttisäännön perusteella on

$$\det \mathbf{I}_p^T \det \mathbf{I}_p = 1.$$

Koska tässä on $\det \mathbf{I}_p = \pm 1$ ja $\det \mathbf{I}_p^T = \det \mathbf{I}_q = \pm 1$, niin päätellään, että on oltava $\det \mathbf{I}_p^T = \det \mathbf{I}_p$. Kun huomioidaan tämä ja kirjoitetaan $\mathbf{I}_p^T = \mathbf{I}_q$, niin laskukaavan (2) perusteella on

$$\det \mathbf{A}^T = \sum_p \det \mathbf{I}_p \prod_{(i,j) \in \Lambda_p} a_{ji} = \sum_q \det \mathbf{I}_q \prod_{(i,j) \in \Lambda_q} a_{ij} = \det \mathbf{A}. \quad \square$$

Säännöstä $\det \mathbf{A}^T = \det \mathbf{A}$ voidaan päätellä, että determinantin vaihtösääntö ja nollasääntö pätevät myös muodossa: $\det \mathbf{A}$ vaihtuu vastaluvukseen, jos $\mathbf{A}$:n kaksi riviä vaihdetaan keskenään, ja $\det \mathbf{A} = 0$, jos $\mathbf{A}$:n kaksi riviä ovat samat.

Jos $\mathbf{A}$ on säännöllinen matriisi, niin soveltamalla tulon determinanttisääntöä tuloon $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ nähdään, että

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

Nähdään myös, että pätee

$$\mathbf{A} \text{ säännöllinen} \;\Rightarrow\; \det \mathbf{A} \neq 0,$$

mikä on loogisesti sama kuin

$$\det \mathbf{A} = 0 \;\Rightarrow\; \mathbf{A} \text{ singulaarinen}.$$

Tämä implikaatio pätee käänteiseenkin suuntaan, ja tässä onkin determinanttiteorian tärkein tulos sovellutusten kannalta. Itse tulos on suora seuraus edellisen luvun tulohajotelmista, tulon determinanttisäännöstä ja kolmiomatriisin determinanttisäännöstä, joka esitetään jäljempänä (Propositio X.5.6).

Lause X.5.4 **(Determinanttikriteeri)** Neliömatriisille pätee

$$\mathbf{A} \text{ singulaarinen} \;\Leftrightarrow\; \det \mathbf{A} = 0.$$

Determinantin laskeminen

Laskukaavaa (1) käytetään yleensä vain tapauksissa $n = 2, 3$:

$$n = 2: \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}\,.$$

$$n = 3: \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{31}a_{12}a_{23} + a_{21}a_{32}a_{13} \\ -a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}\,. \end{array}$$

Nämä voi helposti johtaa myös suoraan Määritelmästä X.5.1. Esimerkiksi:

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= V\left(\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \right) \\ &= V\left(a_{11} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_{21} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, a_{12} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \\ &= a_{11}a_{12} \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{21}a_{12} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + a_{21}a_{22} \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} \\ &= a_{11}a_{12} \cdot 0 + a_{11}a_{22} \cdot 1 + a_{21}a_{12} \cdot (-1) + a_{21}a_{22} \cdot 0 = a_{11}a_{22} - a_{21}a_{12}\,. \end{aligned}$$

Tapauksessa $n = 3$ voi laskukaavan muistaa ryhmittelemällä yhteenlaskettavat termit kaavan (2) mukaisesti seuraavasti (nk. *Sarrus'n sääntö*):

$$\text{etumerkki } + \qquad \begin{vmatrix} \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \end{vmatrix} \quad \begin{vmatrix} \cdot & \bullet & \cdot \\ \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \end{vmatrix} \quad \begin{vmatrix} \cdot & \cdot & \bullet \\ \bullet & \cdot & \cdot \\ \cdot & \bullet & \cdot \end{vmatrix}$$

$$\text{etumerkki } - \qquad \begin{vmatrix} \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \end{vmatrix} \quad \begin{vmatrix} \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \\ \cdot & \bullet & \cdot \end{vmatrix} \quad \begin{vmatrix} \cdot & \bullet & \cdot \\ \bullet & \cdot & \cdot \\ \cdot & \cdot & \bullet \end{vmatrix}$$

Esimerkki 1

$$\begin{vmatrix} 1 & 1 & 3 \\ 2 & 0 & 2 \\ 3 & -1 & 1 \end{vmatrix} = \begin{array}{l} 1 \cdot 0 \cdot 1 + 1 \cdot 2 \cdot 3 + 2 \cdot (-1) \cdot 3 \\ -3 \cdot 0 \cdot 3 - 1 \cdot (-1) \cdot 2 - 2 \cdot 1 \cdot 1 = 0. \end{array}$$

Matriisina tämä on siis singulaarinen. $\square$

Sarrus'n sääntö rajoittuu determinantteihin kokoa 3×3. Hieman isompia determinantteja käsivoimin purettaessa on seuraava nk. *alideterminanttisääntö* laskukaavoja (1)–(2) huomattavasti helppokäyttöisempi. Sääntö on yleispätevä, ja se on myös käypä determinanttifunktion määritelmänä (ks. Harj.teht. 10).

LAUSE X.5.5 **(Alideterminanttisääntö)** Jos $\mathbf{A}$ on neliömatriisi kokoa $n \times n$, niin jokaisella $k \in \{1, \ldots, n\}$ pätee

$$\begin{aligned} \det \mathbf{A} &= \sum_{i=1}^{n} (-1)^{k+i} a_{ik} \det \mathbf{A}^{(i,k)} \\ &= \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det \mathbf{A}^{(k,j)}, \end{aligned}$$

missä $\mathbf{A}^{(i,j)}$ on matriisi kokoa $(n-1) \times (n-1)$, joka saadaan poistamalla $\mathbf{A}$:sta i:s rivi ja j:s sarake.

<u>Todistus</u> Tarkastellaan väitetyistä purkusäännöistä ensimmäistä, kun $k = 1$. Determinantin lineaarisuussäännön (i) perusteella voidaan kirjoittaa

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{i1} \det \mathbf{A}_i,$$

missä $\mathbf{A}_i$ on saatu matriisista $\mathbf{A}$ korvaamalla ensimmäinen sarake yksikkövektorilla $\mathbf{e}_i$. Vaihdetaan matriisissa $\mathbf{A}_i$ ensin ensimmäinen ja i:s rivi keskenään ja permutoidaan tämän jälkeen rivit n:o $2 \ldots i$ siten, että i:s rivi siirtyy toiseksi riviksi muiden rivien järjestyksen muuttumatta. Jälkimmäisessä operaatiossa tarvitaan $i-2$ parittaista rivin vaihtoa, joten vaihtoja kertyy kaikkiaan $i-1$ kpl. Jos vaihtojen jälkeen saatavaa matriisia merkitään $\tilde{\mathbf{A}}_i$, niin on siis $\det \mathbf{A}_i = (-1)^{i-1} \det \tilde{\mathbf{A}}_i$. Vaihtojen seurauksena on $\tilde{\mathbf{A}}_i$:n ensimmäinen sarake $= \mathbf{e}_1$ ja $\tilde{\mathbf{A}}_i^{(1,1)} = \mathbf{A}^{(i,1)}$. Determinantin laskukaavasta (1) nähdään tällöin, että $\det \tilde{\mathbf{A}}_i = \det \tilde{\mathbf{A}}_i^{(1,1)} = \det \mathbf{A}^{(i,1)}$. Yhdistämällä päätelmät seuraa

$$\det \mathbf{A} = \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det \mathbf{A}^{(i,1)} = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det \mathbf{A}^{(i,1)}.$$

Väitetyistä purkusäännöistä ensimmäinen on näin todistettu tapauksessa $k = 1$. Muut väitetyt säännöt palautuvat tähän tapaukseen vaihto- ja transponointisääntöjen avulla. $\square$

Determinantteja $\det \mathbf{A}^{(i,j)}$ sanotaan determinantin $\det \mathbf{A}$ *alideterminanteiksi* — siitä Lauseen X.5.5 purkusääntöjen nimi.

Esimerkki 2 Kun valitaan $k = 1$ ja kehitetään determinantti ensimmäisen sarakkeen mukaan, niin saadaan

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

Alideterminantit voidaan kehittää Sarrus'n säännöllä tai alideterminanttisäännöllä. Sovelletaan jälkimmäistä siten, että kaksi ensimmäistä alideterminanttia kehitetään ensimmäisen sarakkeen mukaan ($k = 1$) ja viimeinen kolmannen rivin mukaan ($k = 3$):

$$\begin{vmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 2 \end{vmatrix} = (-1) \cdot \begin{vmatrix} 0 & 1 \\ 2 & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = (-1) \cdot (-2) - 1 \cdot 0 = 2,$$

$$\begin{vmatrix} 2 & -2 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 0 & 1 \\ 2 & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} -2 & 1 \\ 2 & 2 \end{vmatrix} = 2 \cdot (-2) - 1 \cdot (-6) = 2,$$

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

Determinantin arvoksi tulee $1 \cdot 2 - 2 \cdot 2 - 1 \cdot (-3) = 1$. $\square$

Alideterminanttisäännöllä voidaan yleinen, kokoa $n \times n$ oleva determinantti purkaa lopulta alideterminanteiksi kokoa 2×2. Tarvittava laskuoperaatioiden määrä on kuitenkin edelleen suuruusluokkaa $W \sim n!$ (ks. Harj.teht. 9), joten suurilla n:n arvoilla on tämäkin menetelmä kelvoton — Miten siis ylipäänsä on mahdollista laskea esimerkiksi determinantti kokoa $n = 100$? Vastaukseen johtaa seuraava tulos, joka on helposti todistettavissa alideterminanttisäännöllä.

Propositio X.5.6 Kolmiomatriisin $\mathbf{A} = \{a_{ij}\}$ determinantti on matriisin lävistäjäalkioiden tulo:

$$\det \mathbf{A} = \prod_{i=1}^{n} a_{ii}.$$

<u>Todistus</u> Kun alideterminanttisäännössä valitaan $k = 1$ ja kehitetään determinantti joko ensimmäisen rivin (alakolmio) tai ensimmäisen sarakkeen (yläkolmio)

mukaan, niin a_{11} on ko. rivin/sarakkeen ainoa nollasta poikkeava alkio, joten Lauseen X.5.5 mukaan

$$\det \mathbf{A} = a_{11} \det \mathbf{A}^{(1,1)}.$$

Tässä $\mathbf{A}^{(1,1)}$ on jälleen kolmiomatriisi, joten sama sääntö soveltuu yhä uudelleen ja johtaa väitettyyn lopputulokseen. $\square$

Propositioon X.5.6 perustuen voidaan determinantin arvo laskea — kuinkas muuten — Gaussin algoritmilla. Nimittäin jos algoritmi menee läpi ilman tukioperaatioita (rivien/sarakkeiden vaihtoja), niin se johtaa LU-hajotelmaan

$$\mathbf{A} = \mathbf{L}\mathbf{U},$$

missä $[\mathbf{L}]_{kk} = 1$ ja $[\mathbf{U}]_{kk} = a_{kk}^{(k-1)}$, $k = 1 \ldots n$ (vrt. edellinen luku). Tällöin on tulon determinanttisäännön ja Proposition X.5.6 mukaan

$$\det \mathbf{A} = \det \mathbf{L} \det \mathbf{U} = \prod_{k=1}^{n} a_{kk}^{(k-1)}.$$

Determinantin arvo on siis yksinkertaisesti tukialkioiden tulo Gaussin algoritmin eliminaatiovaiheessa (!). Jos eliminaatiovaiheessa tehdään tukioperaatioita, tulee determinantin arvoksi

$$\det \mathbf{A} = (-1)^m \prod_{k=1}^{n} [\mathbf{U}]_{kk},$$

missä m on tehtyjen rivien/sarakkeiden vaihtojen kokonaismäärä.

Gaussin algoritmia käyttäen tulee siis determinantin laskemisen työmääräksi $W \sim \frac{1}{3}n^3$ (kerto- ja yhteenlaskua) — eli tämä on suurilla n:n arvoilla huomattavasti tehokkaampi menetelmä kuin mikään edellä esitetyistä vaihtoehdoista. Puhtaasti numeerisissa laskuissa, ja suurilla n, determinantilla ei olekaan juuri käytännön merkitystä — selvittäähän Gaussin algoritmi matriisin säännöllisyyskysymyksen muutenkin. Determinantin käsite on kuitenkin käyttökelpoinen erinäisissä lineaarialgebran teoreettisissa tarkasteluissa, ja myös symbolisessa laskennassa determinantilla on käyttöä. Symbolisen laskennan ongelma voi olla esimerkiksi sellainen, että tarkasteltava matriisi riippuu jostakin parametrista (muuttujasta) s, ts. $\mathbf{A} = (a_{ij}(s)) = \mathbf{A}(s)$, ja halutaan vaikkapa selvittää, millä s:n arvoilla $\mathbf{A}(s)$ on singulaarinen. Jos $\mathbf{A}$ ei ole kooltaan kovin suuri, niin determinanttiehto

$$\det \mathbf{A}(s) = 0$$

on ratkaisun lähtökohtana luonnollinen ja (etenkin käsinlaskussa) usein käytetty.

Cramerin sääntö

Determinantin muista käyttömuodoista on syytä vielä mainita (ilman todistusta, ks. Harj.teht. 11) seuraavat kaksi laskusääntöä.

LAUSE X.5.7 (**Cramerin**† **sääntö**) Jos $\mathbf{A}$ on säännöllinen neliömatriisi kokoa $n \times n$, niin yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ ratkaisu on

$$\mathbf{x} = (x_i), \quad x_i = \frac{\det \mathbf{A}^{(i)}}{\det \mathbf{A}}, \quad i = 1 \ldots n,$$

missä $\mathbf{A}^{(i)}$ saadaan $\mathbf{A}$:sta korvaamalla i:s sarake $\mathbf{b}$:llä.

LAUSE X.5.8 Säännöllisen matriisin $\mathbf{A}$ käänteismatriisi on

$$\mathbf{A}^{-1} = \mathbf{B}^T, \quad [\mathbf{B}]_{ij} = (-1)^{i+j} \frac{\det \mathbf{A}^{(ij)}}{\det \mathbf{A}}.$$

Cramerin säännöllä on pienikokoisia yhtälöryhmiä ratkaistaessa edelleen jonkin verran käyttöä käsinlaskussa. Etenkin jos kerroinmatriisin alkiot ovat kokonaislukuja, voidaan säännöllä minimoida (käsinlaskussa vaivalloiset) jakolaskut. Myös symbolisessa (käsin)laskennassa Cramerin säännöllä on käyttöä samaan tapaan kuin determinantilla yleensä. — Numeerisessa matriisilaskennassa sen sijaan Cramerin säännöllä on kyseenalainen kunnia esiintyä 'maailman huonoimpana' lineaarisen yhtälöryhmän ratkaisualgoritmina.

HARJOITUSTEHTÄVIÄ

1. Näytä, että ortogonaalisen matriisin determinantilla on vain kaksi mahdollista arvoa: Joko $\det \mathbf{A} = 1$ tai $\det \mathbf{A} = -1$.

2. Laske Sarrus'n säännöllä:

$$\text{a)} \begin{vmatrix} 2 & 3 & -1 \\ -1 & 2 & 0 \\ 1 & 4 & 3 \end{vmatrix} \qquad \text{b)} \begin{vmatrix} 0 & -2 & -3 \\ 1 & 0 & -2 \\ 3 & 4 & 0 \end{vmatrix} \qquad \text{c)} \begin{vmatrix} -3 & 7 & 2 \\ -5 & 4 & 0 \\ 9 & -1 & -6 \end{vmatrix}$$

3. Olkoon $\alpha, \beta, \gamma \in \mathbb{R}$. Laske

$$\text{a)} \begin{vmatrix} 1 & \alpha & \alpha^2 \\ 1 & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{vmatrix} \qquad \text{b)} \begin{vmatrix} \alpha+\beta & \alpha+2\beta & \alpha+3\beta \\ \alpha+3\beta & \alpha+\beta & \alpha+2\beta \\ \alpha+2\beta & \alpha+3\beta & \alpha+\beta \end{vmatrix} \qquad \text{c)} \begin{vmatrix} 2 & \alpha & \alpha^2 \\ 1 & \alpha^2 & \alpha \\ \alpha & \alpha^3 & 1 \end{vmatrix}$$

†Sveitsiläinen matemaatikko **Gabriel Cramer** (1704–1752) oli determinanttiteorian uranuurtajia teoksellaan "Introduction à l'analyse des lignes courbes algébrigues" (1750).

4. Laske a) alideterminanttikehitelmällä, b) ja c) Gaussin algortimilla, d) molemmilla:

$$\text{a)} \begin{vmatrix} 1 & -1 & 0 & 2 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \end{vmatrix} \qquad \text{b)} \begin{vmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{vmatrix}$$

$$\text{c)} \begin{vmatrix} 2 & 1 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 5 \end{vmatrix} \qquad \text{d)} \begin{vmatrix} 1 & 1 & 4 & 5 \\ 1 & 1 & 5 & 4 \\ 2 & 4 & 1 & 1 \\ 4 & 2 & 1 & 1 \end{vmatrix}$$

5. Laske $\det \mathbf{B}$, kun $\mathbf{B} = (-5\mathbf{A}\mathbf{A}^T)^7$ ja $\mathbf{A} = \begin{bmatrix} 13 & 8 & 6 \\ -13 & -8 & -4 \\ 8 & 5 & 5 \end{bmatrix}$.

6. Määritä kaikki reaaliset tai kompleksiset λ:n arvot, joilla seuraavat matriisit ovat singulaarisia.

$$\text{a)} \begin{bmatrix} 2-\lambda & -4 \\ 2 & 6-\lambda \end{bmatrix} \qquad \text{b)} \begin{bmatrix} \lambda-2 & 3 & 1 \\ \lambda-4 & 3 & 2 \\ \lambda-6 & \lambda & 3 \end{bmatrix}$$

$$\text{c)} \begin{bmatrix} 1 & 3-\lambda & 4 \\ 4-\lambda & 2 & -1 \\ 1 & \lambda-6 & 2 \end{bmatrix} \qquad \text{d)} \begin{bmatrix} 1-\lambda & -1 & 2 \\ 1 & 2-\lambda & -13 \\ -2 & 1 & 1+\lambda \end{bmatrix}$$

7. Olkoon $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$, ja määritellään

$$\mathbf{b}_k = \mathbf{a}_k, \quad \mathbf{b}_j = \mathbf{a}_j + \beta_j \mathbf{a}_k, \quad j = 1 \ldots n,\ j \neq k,$$

missä $1 \leq k \leq n$ ja $\beta_j \in \mathbb{R}$. Näytä determinanttifunktion V ominaisuuksien perusteella, että pätee $V(\mathbf{b}_1, \ldots, \mathbf{b}_n) = V(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Miten tämä tulos liittyy Gaussin algoritmiin sovellettuna matriisiin $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]^T$?

8. Neliömatriisi $\mathbf{A} = (a_{ij})$ olkoon kokoa $n \times n$ ja *tridiagonaalinen*, ts. $a_{ij} = 0$, kun $|i-j| \geq 2$. Edelleen olkoon $a_{ii} = 1$, $i = 1 \ldots n$, ja $a_{ij} = \lambda$, kun $|i-j| = 1$. a) Näytä, että determinantille $D_n = \det \mathbf{A}$ pätee palautuskaava $D_n = D_{n-1} - \lambda^2 D_{n-2}$. b) Laske D_n, $n = 2 \ldots 10$, kun $\lambda = 2$. c) Jos $\lambda = 1$, niin millä n:n arvoilla $\mathbf{A}$ on singulaarinen?

9. (*) Näytä, että jos determinantti kokoa $n \geq 3$ lasketaan alideterminanttisäännöllä, niin tarvittava kertolaskujen lukumäärä on

$$W_n = n! \sum_{k=1}^{n-1} \frac{1}{k!} \,.$$

10. (*) Määritellään determinanttifunktio $V(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \det[\mathbf{a}_1 \ldots \mathbf{a}_n]$ palautuvasti käyttäen Lauseen X.5.5 ensimmäistä purkusääntöä ($k = 1$) sekä sääntöä $\det a = a$ determinantille kokoa $n = 1$.

    a) Näytä induktiolla, että mainitulla tavalla määritellyllä funktiolla on determinanttifunktion ominaisuudet (i)–(iii).

    b) Näytä, että jos $p = (i_1, \ldots, i_n)$ on joukon $(1, \ldots, n)$ permutaatio ja $\mathbf{I}_p = [\mathbf{e}_{i_1} \ldots \mathbf{e}_{i_n}]$, niin a-kohdan määritelmän mukaisesti on joko $\det \mathbf{I}_p = 1$ tai $\det \mathbf{I}_p = -1$. Päättele, että edellisessä tapauksessa permutaatio p on parillinen (eli saavutettavissa parillisella määrällä parivaihtoja) ja jälkimmäisessä pariton — siis jokainen permutaatio on jompaa kumpaa tyyppiä.

11. (*) Valitsemalla $\mathbf{b} = b_j \mathbf{e}_j,\ j = 1 \ldots n$ johda Cramerin sääntö alideterminanttisäännöstä. Todista edelleen Lause X.5.8 lähtien Cramerin säännöstä.

X.6 Lineaarikuvaukset

Avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ väliseksi *lineaarikuvaukseksi* sanotaan funktiota tyyppiä

$$\mathbf{f}: \ \mathbb{R}^n \to \mathbb{R}^m, \quad m, n \in \mathbb{N},$$

($\mathcal{D}_\mathbf{f} = \mathbb{R}^n, \ \mathcal{R}_\mathbf{f} \subset \mathbb{R}^m$), jolle on voimassa ehto

$$\mathbf{f}(\alpha\mathbf{x} + \beta\mathbf{y}) \ = \ \alpha\mathbf{f}(\mathbf{x}) + \beta\mathbf{f}(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \ \alpha, \beta \in \mathbb{R}. \tag{1}$$

Lineaarikuvauksia tarkasteltaessa on tapana poiketa hieman tavallisista funktiomerkinnöistä: Lineaarikuvausta merkitään

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} \quad \text{(lineaarikuvaus)}.$$

Oikealla siis jätetään sulkeet pois muuttujan ympäriltä, ja tämä ilmenee myös lukutavassa: Luetaan '$A\,x$' mieluummin kuin 'A x:ssä'. Lukutavan mukaisesti ajatellaan kuvaamisen (eli funktion arvon määräämisen) tapahtuvan siten, että A *operoi* $\mathbf{x}$:ään. Operoinnin voi ymmärtää abstraktina kertolaskuna (= matriisin ja vektorin tulo), joka on vektorien yhteenlaskun että skalaarilla kertomisen suhteen *lineaarinen* eli toteutaa tavanomaiset osittelulait näiden laskuoperaatioiden suhteen:

$$\boxed{A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y} \quad \text{(lineaarikuvaus)}.}$$

Lineaarikuvausten symboleina käytetään yleensä isoja kirjaimia A, B jne.

Jos $\mathbf{A}$ on matriisi kokoa $m \times n$, niin matriisialgebran sääntöjen perusteella on ilmeistä, että funktio $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ on lineaarikuvaus tyyppiä $\mathbf{f}: \ \mathbb{R}^n \to \mathbb{R}^m$. Osoittautuu, ettei muunlaisia, tätä tyyppiä olevia lineaarikuvauksia olekaan:

PROPOSITIO X.6.1 Jos $\mathbf{f}: \ \mathbb{R}^n \to \mathbb{R}^m$ on lineaarikuvaus, niin on olemassa yksikäsitteinen matriisi $\mathbf{A}$ kokoa $m \times n$, siten että $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}, \ \mathbf{x} \in \mathbb{R}^n$.

<u>Todistus</u> Kun $\mathbf{x} \in \mathbb{R}^n$ esitetään kannan $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ avulla muodossa

$$\mathbf{x} \ = \ \sum_{i=1}^{n} x_i \mathbf{e}_i,$$

niin $\mathbf{f}$:n lineaarisuuden ja matriisialgebran sääntöjen (Luku X.1) nojalla

$$\mathbf{f}(\mathbf{x}) \ = \ \sum_{i=1}^{n} x_i \mathbf{f}(\mathbf{e}_i) \ = \ \mathbf{A}\mathbf{x}, \qquad \mathbf{A} = [\mathbf{f}(\mathbf{e}_1) \ldots \mathbf{f}(\mathbf{e}_n)]. \quad \square$$

Proposition X.6.1 perusteella lineaarikuvausten ja matriisien välillä on kääntäen yksikäsitteinen vastaavuus:

$$\text{lineaarikuvaus } A: \ \mathbb{R}^n \to \mathbb{R}^m \quad \longleftrightarrow \quad \text{matriisi } \mathbf{A} \text{ kokoa } m \times n \ \ (A\mathbf{x} = \mathbf{A}\mathbf{x}).$$

Jos matriisi $\mathbf{A}$ vastaa lineaarikuvausta A, niin vastaavuuteen viitataan sanomalla, että $\mathbf{A}$ on ko. *lineaarikuvauksen matriisi* (tai esitysmatriisi). Esitysmatriisin kautta voidaan siis lineaarikuvauksilla laskeminen aina palauttaa matriisialgebraan. Toisaalta voidaan matriisin ja vektorin kertolasku aina haluttaessa nähdä lineaarisena kuvauksena, ts. erään lineaarikuvauksen laskusäännön soveltamisena. Erityisesti kahden matriisin tulo vastaa kahden lineaarikuvauksen yhdistelyä:

$$AB \ \leftrightarrow \ \mathbf{AB}.$$

Yhdistelyä koskevat rajoitukset ovat samat kuin matriisitulossa: On oletettava, että $B: \ \mathbb{R}^n \to \mathbb{R}^p$ ja $A: \ \mathbb{R}^p \to \mathbb{R}^m$ joillakin $m, n, p \in \mathbb{N}$, jolloin myös tulo $\mathbf{AB}$ on määritelty.

Lineaarikuvauksen $A: \ \mathbb{R}^n \to \mathbb{R}^m$ *käänteiskuvauksen* A^{-1} muodollinen määritelmä on

$$A\mathbf{x} = \mathbf{y} \quad \Leftrightarrow \quad \mathbf{x} = A^{-1}\mathbf{y}.$$

Käänteiskuvaus on tällä tavoin määriteltävissä, mikäli A on 1-1 ('yksi yhteen'), eli toteuttaa ehdon

$$A\mathbf{x}_1 = A\mathbf{x}_2 \quad \Rightarrow \quad \mathbf{x}_1 = \mathbf{x}_2.$$

Koska A on lineaarinen, tämä on sama kuin ehto

$$A\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}.$$

Aiemmista tuloksista (Luku X.4) voidaan päätellä, että tämä ehto voi toteutua vain jos $m \geq n$ (välttämätön ehto). Jos ehto toteutuu, on A *kääntyvä*, ja käänteiskuvaus on tällöin määritelty kuvauksena tyyppiä $A^{-1}: \ \mathbb{R}^m \to \mathbb{R}^n$. Määrittelyjoukko ei kuitenkaan ole välttämättä koko $\mathbb{R}^m$, sillä käänteiskuvauksen määritelmän mukaan $A^{-1}\mathbf{y}$ on määritelty vain sellaisille vektoreille $\mathbf{y} \in \mathbb{R}^m$, joille $\mathbf{y} = \mathbf{A}\mathbf{x}$ jollakin $\mathbf{x} \in \mathbb{R}^n$, ts. yhtälöryhmällä $\mathbf{A}\mathbf{x} = \mathbf{y}$ on oltava ratkaisu. Aiempien tulosten perustella yhtälöryhmän $\mathbf{A}\mathbf{x} = \mathbf{y}$ ratkeavuus jokaisella $\mathbf{y} \in \mathbb{R}^m$ on mahdollinen vain jos $m \leq n$ (välttämätön ehto). Päätellään siis, että käänteiskuvaus A^{-1} on on olemassa ja määritelty koko $\mathbb{R}^m$:ssä täsmälleen kun $m = n$ ja A:n matriisi $\mathbf{A}$ (neliömatriisi) on säännöllinen. Käänteiskuvaus on tällöin lineaarikuvaus tyyppiä $A^{-1}: \mathbb{R}^n \to \mathbb{R}^n$ (samaa tyyppiä kuin A), ja käänteiskuvauksen matriisi on luonnollisesti $\mathbf{A}^{-1}$.

Lineaarikuvauksen $A: \ \mathbb{R}^n \to \mathbb{R}^n$ käänteiskuvauksen lineaarisuutta käytetään usein hyödyksi lineaarisia yhtälöryhmiä ratkaistaessa. Tyypillinen sovellutustilanne on sellainen, että yhtälöryhmä $\mathbf{A}\mathbf{x} = \mathbf{b}$ halutaan ratkaista useilla eri $\mathbf{b}$:n

arvoilla $\mathbf{A}$:n pysyessä samana. Jos ratkaisua merkitään kuvauksena $\mathbf{b} \mapsto \mathbf{x}$ (kuvauksen $\mathbf{x} \mapsto \mathbf{b} = \mathbf{A}\mathbf{x}$ käänteiskuvaus) ja oletetaan, että tunnetaan ratkaisut $\mathbf{b}_1 \mapsto \mathbf{x}_1$ ja $\mathbf{b}_2 \mapsto \mathbf{x}_2$, niin lineaarisuuden perusteella tiedetään (vrt. Luku X.3, Esimerkki 1)

$$\alpha\mathbf{b}_1 + \beta\mathbf{b}_2 \longmapsto \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 \quad (\alpha, \beta \in \mathbb{R}).$$

Kannan vaihto $\mathbb{R}^n$:ssä

$\mathbb{R}^n$:n vektorisysteemiä (järjestettyä joukkoa) $\{\mathbf{a}_1, \ldots \mathbf{a}_m\}$ sanotaan *lineaarisesti riippumattomaksi*, jos pätee

$$x_1\mathbf{a}_1 + \ldots + x_m\mathbf{a}_m = \mathbf{0} \quad \Rightarrow \quad x_i = 0, \quad i = 1 \ldots m.$$

Jos vektorit $\mathbf{a}_i$ kootaan matriisin $\mathbf{A}$ sarakkeiksi ($\mathbf{A}$ kokoa $n \times m$), niin voidaan kirjoittaa $x_1\mathbf{a}_1 + \ldots + x_m\mathbf{a}_m = \mathbf{A}\mathbf{x}$ (vrt. Luku X.1), joten lineaarisen riippumattomuuden ehto saa muodon

$$\mathbf{x} \in \mathbb{R}^m \ \ \& \ \ \mathbf{A}\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}, \qquad \mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m].$$

Luvun X.4 tulosten mukaan tämä ehto voi olla voimassa vain kun $m \leq n$. Siis $\mathbb{R}^n$:ssä on jokainen $n+1$ vektorin (tai useamman vektorin) muodostama systeemi *lineaarisesti riippuva*. Toisaalta jos $m = n$, niin lineaarisesti riippumattoman systeemin muodostaa ainakin $\mathbb{R}^n$:n luonnollinen kanta $\{\mathbf{e}_1, \ldots \mathbf{e}_n\}$, jota ym. ehdossa vastaa matriisi $\mathbf{A} = \mathbf{I}$. Luku n on siis lineaarisen riippumattomuuden kannalta kriittinen luku, ja tämä onkin jo aiemmin nimetty avaruuden $\mathbb{R}^n$ *dimensioksi* (=ulotteisuus). Jos $m = n$, niin ym. ehto toteutuu täsmälleen kun $\mathbf{A}$ on säännöllinen matriisi (Lause X.2.6). Näin ollen, ja koska $\mathbf{A}$ on säännöllinen täsmälleen kun $\mathbf{A}^T$ on säännöllinen (ks. Luku X.2), saadaan neliömatriisin säännöllisyydelle seuraavat uudet (geometriset) tulkinnat:

$\mathbf{A}$ säännöllinen	$\Leftrightarrow$	$\mathbf{A}$:n sarakkeet lineaarisesti riippumattomat
	$\Leftrightarrow$	$\mathbf{A}$:n rivit lineaarisesti riippumattomat.

Jos $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ on $\mathbb{R}^n$:n lineaarisesti riippumaton vektorisysteemi, eli matriisi $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$ on säännöllinen, niin mielivaltainen vektori $\mathbf{b} \in \mathbb{R}^n$ voidaan lausua yksikäsitteisesti vektorien $\mathbf{a}_i$ lineaarisena yhdistelynä. Nimittäin kertoimet x_i ratkeavat yksikäsitteisesti yhtälöryhmästä

$$\sum_{i=1}^{n} x_i\mathbf{a}_i = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{A}\mathbf{x} = \mathbf{b}.$$

Tällä perusteella voidaan sanoa, että jokainen lineaarisesti riippumaton n vektorin systeemi $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ on $\mathbb{R}^n$:n *kanta*. Jos siis $\mathbf{A}$ on mikä tahansa säännöllinen matriisi kokoa $n \times n$, niin $\mathbf{A}$:n sarakkeet (myös rivit) muodostavat $\mathbb{R}^n$:n kannan. — Huomattakoon, että tällainen kanta on *ortonormeerattu*, eli $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \delta_{ij}$ (Kroneckerin δ), täsmälleen kun $\mathbf{A}$ on ortogonaalinen (vrt. Luku X.2). Päätellään siis, että jokainen säännöllinen matriisi kokoa $n \times n$ edustaa sarakkeittensa (tai riviensä) kautta $\mathbb{R}^n$:n kantaa, ja jokainen ortogonaalinen matriisi edustaa ortonormeerattua kantaa.

Jos halutaan suorittaa *kannan vaihto*, lähtökohtana $\mathbb{R}^n$:n luonnollinen kanta $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, niin vektorin $\mathbf{x}$ koordinaatit x_i' kannassa $\{\mathbf{a}_1, \ldots \mathbf{a}_m\}$ ovat siis ratkaistavissa yhtälöryhmästä

$$\mathbf{A}\mathbf{x}' = \mathbf{x} \quad \Rightarrow \quad \mathbf{x}' = \mathbf{A}^{-1}\mathbf{x}, \qquad \mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n].$$

Kannan vaihdossa, eli koordinaattimuunnoksessa $\mathbf{x} \mapsto \mathbf{x}'$, on siis kyse lineaarikuvauksesta. Jos uusikin kanta on ortonormeerattu, niin muunnoksen matriisi on helppo laskea: $\mathbf{A}^{-1} = \mathbf{A}^T$.

Esimerkki 1 Määritä vektorin $x\vec{i}+y\vec{j}+z\vec{k}$ koordinaatit x', y', z' kannassa, jonka muodostavat vektorit $\vec{a}_1 = \vec{i} + \vec{j}$, $\vec{a}_2 = \vec{j} + \vec{k}$ ja $\vec{a}_3 = \vec{i} + \vec{k}$.

<u>Ratkaisu</u>

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ x \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad \square$$

Esimerkki 2 Pallokoordinaatiston pisteessä $P = (r, \theta, \varphi)$ käytetään avaruusvektoreille kantaa $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$, missä (vrt. Luku II.8)

$$\begin{cases} \vec{e}_r &= \sin\theta\cos\varphi\vec{i} + \sin\theta\sin\varphi\vec{j} + \cos\theta\vec{k}, \\ \vec{e}_\theta &= \cos\theta\cos\varphi\vec{i} + \cos\theta\sin\varphi\vec{j} - \sin\theta\vec{k}, \\ \vec{e}_\varphi &= -\sin\varphi\vec{i} + \cos\varphi\vec{j}. \end{cases}$$

Kun kirjoitetaan

$$\vec{F} = F_1\vec{i} + F_2\vec{j} + F_3\vec{k} = F_r\vec{e}_r + F_\theta\vec{e}_\theta + F_\varphi\vec{e}_\varphi\,,$$

niin muunnoskaava $(F_r, F_\theta, F_\varphi) \mapsto (F_1, F_2, F_3)$ on matriisimuodossa

$$\begin{bmatrix} \sin\theta\cos\varphi & \cos\theta\cos\varphi & -\sin\varphi \\ \sin\theta\sin\varphi & \cos\theta\sin\varphi & \cos\varphi \\ \cos\theta & -\sin\theta & 0 \end{bmatrix} \begin{bmatrix} F_r \\ F_\theta \\ F_\varphi \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}.$$

Koska $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi\}$ on ortonormeerattu systeemi, niin käänteismuunnos saadaan transponoimalla:

$$\begin{bmatrix} F_r \\ F_\theta \\ F_\varphi \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}. \quad \square$$

Koordinaatiston kierto

Edellä todettiin, että jos $\mathbf{A}$ on ortogonaalinen matriisi kokoa $n \times n$, niin $\mathbf{A}$:n sarakkeet (myös rivit) määrittelevät $\mathbb{R}^n$:n ortonormeeratun kannan. Sanotaan, että ko. kanta on *positiivisesti suunnistettu* (tai 'oikeakätinen'), jos $\det \mathbf{A} = 1$ ja *negatiivisesti suunnistettu* ('vasenkätinen'), jos $\det \mathbf{A} = -1$. (Muita mahdollisia arvoja ei $\det \mathbf{A}$:lla ole, ks. Harj.teht. X.5: 1). Jos yhden kantavektorin suunta vaihdetaan, niin determinanttiopin mukaisesti $\det \mathbf{A}$:n merkki vaihtuu, eli kannan suunnistus vaihtuu.

Esimerkki 3 Jos $\{\vec{a}, \vec{b}, \vec{c}\}$ on avaruusvektoreiden ortonormeerattu kanta, niin ko. vektoreiden koordinaattivektorit kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$ muodostavat $\mathbb{R}^3$:n ortonormeeratun kannan $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, eli matriisi $\mathbf{A} = [\mathbf{a}\,\mathbf{b}\,\mathbf{c}]$ on ortogonaalinen (kuten edellisessä esimerkissä). Geometrisin vektorioperaatioin laskettuna $\mathbf{A}$:n determinantti on (vrt. Luku II.6)

$$\det \mathbf{A} = \det \mathbf{A}^T = (\vec{a} \times \vec{b}) \cdot \vec{c} = [\,\vec{a}\,\vec{b}\,\vec{c}\,].$$

Koska $\{\vec{a}, \vec{b}, \vec{c}\}$ on ortonormeerattu systeemi, niin on joko $\vec{c} = \vec{a} \times \vec{b}$, jolloin $\det \mathbf{A} = |\vec{c}|^2 = 1$, tai $\vec{c} = -(\vec{a} \times \vec{b})$, jolloin $\det \mathbf{A} = -1$. Edellisessä tapauksessa kanta on siis positiivisesti, jälkimmäisessä negatiivisesti suunnistettu. $\square$

Esimerkki 4 Napakoordinaatiston kannat $\{\vec{e}_r, \vec{e}_\varphi\}$ (vrt. Luku II.8) kattavat eri φ:n arvoilla kaikki mahdolliset tason vektoreiden ortonormeeratut, positiivisesti suunnistetut systeemit. Näin ollen kaikki ortogonaaliset matriisit $\mathbf{A}$ kokoa 2×2, jotka täyttävät ehdon $\det \mathbf{A} = 1$, ovat muotoa

$$\mathbf{A} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}, \quad \varphi \in [0, 2\pi). \quad \square$$

Jos $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ on ortogonaalinen matriisi ja $\det \mathbf{A} = 1$, niin kannan vaihtoa $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \hookrightarrow \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ sanotaan *koordinaatiston kierroksi.* Nimittäin on osoitettavissa, että kannan vaihto on tällöin toteutettavissa <u>jatkuvana</u> muunnoksena muotoa

$$t \in [0, 1] \;\mapsto\; \{\mathbf{a}_1(t), \ldots\, \mathbf{a}_n(t)\}$$

siten, että seuraavat ehdot toteutuvat:

(i) $\mathbf{a}_i(0) = \mathbf{e}_i$ ja $\mathbf{a}_i(1) = \mathbf{a}_i, \;\; i = 1 \ldots n.$

(ii) Vektorisysteemi $\{\mathbf{a}_1(t), \ldots, \mathbf{a}_n(t)\}$ on ortonormeerattu ja positiivisesti suunnistettu jokaisella $t \in [0, 1]$.

(iii) Jatkuvuusehto: Funktiot $t \in [0, 1] \mapsto \mathbf{a}_j(t) \in \mathbb{R}^n$ ovat jatkuvia, ts. reaalifunktiot $t \mapsto a_{ij}(t) = [\mathbf{a}_j(t)]_i, \;\; i, j = 1 \ldots n$ ovat jatkuvia välillä $[0, 1]$.

Ehdot täyttävä muunnos voidaan kuvitella toteutettavan niin, että vektorisysteemiä $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ kierretään 'jäykkänä kappaleena', jolloin ortogonaalisuus ja suunnistus säilyvät. Perussysteemistä $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ lähtien on siis tällaisella jatkuvalla kierrolla mahdollista päätyä mihin tahansa annettuun, ortonormeerattuun ja positiivisesti suunnistettuun systeemiin $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. Jos ko. systeemi on negatiivisesti suunnistettu, niin jossakin on tapahduttava 'hyppy' (yhden vektorin suunnan vaihto), jolloin muunnos ei ole jatkuva. Tämän väittämän mukaisesti siis kaikki $\mathbb{R}^n$:n positiivisesti suunnistetut ortonormeeratut systeemit ovat keskenään (samoin negatiivisesti suunnistetut keskenään) *kiertoekvivalentteja*. Tapauksessa $n = 2$ väittämän voi todistaa Esimerkin 4 tuloksen perusteella (Harj.teht. 5). Yleinen tapaus on haastavampi — tarkemmat perustelut sivuutetaan.

$\mathbb{R}^n$:n aliavaruudet

Olkoon annettu m vektoria $\mathbf{a}_j \in \mathbb{R}^n, \;\; j = 1 \ldots m$ (voi olla $m \leq n$ tai $m > n$). Tällöin vektorijoukko, joka määritellään

$$W = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} = \sum_{j=1}^{m} y_j \mathbf{a}_j \;\text{ jollakin }\; [y_1, \ldots, y_m]^T \in \mathbb{R}^m\}$$

on $\mathbb{R}^n$:n *aliavaruus*, sillä määritelyn perusteella pätee $W \subset \mathbb{R}^n$ ja

$$\mathbf{v}_1, \mathbf{v}_2 \in W \quad \Rightarrow \quad \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \in W \quad \forall\, \alpha_1, \alpha_2 \in \mathbb{R}.$$

Sanotaan, että W on vektoreiden $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *virittämä* avaruus. Jos vektorit $\mathbf{a}_i$ ovat lineaarisesti riippumattomat (jolloin on oltava $m \leq n$), niin dim W $= m$ ja $\{\mathbf{a}_1, \ldots \mathbf{a}_m\}$ on W:n kanta. Ovatko vektorit $\mathbf{a}_i$ lineaarisesti riippumattomat vai eivät, selviää Gaussin algoritmilla: Muodostetaan vektorit $\mathbf{a}_j$ sarakkeina matriisi $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_m]$ ja etsitään yhtälöryhmän $\mathbf{A}\mathbf{x} = \mathbf{0}$ kaikki ratkaisut $\mathbf{x} \in \mathbb{R}^m$ (vrt. Luku X.4).

Esimerkki 5 Ovatko $\mathbb{R}^4$:n vektorit $\mathbf{a}_1 = [1, 2, 1, -1]^T$, $\mathbf{a}_2 = [2, -1, 3, -1]^T$ ja $\mathbf{a}_3 = [0, 5, -1, -1]^T$ lineaarisesti riippumattomat?

Ratkaisu Muunnetaan yhtälöryhmä $\mathbf{Ax} = \mathbf{0}$ Gaussin algoritmilla singulaariseen perusmuotoon $\mathbf{Ux} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 5 \\ 1 & 3 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 2 & 0 \\ 0 & -5 & -5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Tämän perusteella

$$\mathbf{Ax} = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{x} = t\,[-2, 1, 1]^T,\ t \in \mathbb{R}.$$

Vastaus: Eivät, sillä $-2\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$. $\square$

Gaussin algoritmilla selviää myös yleisemmin, mikä on annettujen vektorien $\mathbf{a}_j \in \mathbb{R}^n,\ j = 1 \ldots m$ virittämän aliavaruuden W dimensio ja mikä ko. vektorisysteemin lineaarisesti riippumaton osasysteemi kelpaa W:n kannaksi silloin, kun koko systeemi on lineaarisesti riippuva.

Esimerkki 5 (jatko) Esimerkin matriisin $\mathbf{A}$ sarakkeet virittävät $\mathbb{R}^4$:n aliavaruuden W ja rivit $\mathbb{R}^3$:n aliavaruuden U. Määritä näiden aliavaruuksien dimensiot ja kummallekin jokin kanta.

Ratkaisu Esimerkin algoritmissa ei tehty rivien eikä sarakkeiden vaihtoja, joten tuloksesta on suoraan luettavissa: dim W = dim $U = 2$, W:n eräs kanta = $\mathbf{A}$:n sarakkeet 1 ja 2 ja U:n eräs kanta = $\mathbf{A}$:n rivit 1 ja 2, ts.

$$W = \{\mathbf{v} = \alpha_1[1, 2, 1, -1]^T + \alpha_2[2, -1, 3, -1]^T,\ \ \alpha_1, \alpha_2 \in \mathbb{R}\},$$
$$U = \{\mathbf{v} = \alpha_1[1, 2, 0] + \alpha_2[2, -1, 5],\ \ \alpha_1, \alpha_2 \in \mathbb{R}\}. \quad \square$$

Esimerkki 6 Matriisin

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 0 & 3 & 4 & 1 \\ 1 & -3 & 2 & -1 & -1 & 2 \\ -3 & 3 & -3 & 0 & 2 & -1 \\ 6 & -2 & 4 & 2 & 1 & 3 \end{bmatrix}$$

sarakkeet virittävät $\mathbb{R}^4$:n aliavaruuden W ja rivit $\mathbb{R}^6$:n aliavaruuden U. Määritä näiden aliavaruuksien dimensiot ja kummallekin jokin kanta.

Ratkaisu Yhtälöryhmän

$$\begin{bmatrix} 3 & 3 & 0 & 3 & 4 & 1 \\ 1 & -3 & 2 & -1 & -1 & 2 \\ -3 & 3 & -3 & 0 & 2 & -1 \\ 6 & -2 & 4 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

singulaariseksi perusmuodoksi saadaan tuetulla Gaussin algoritmilla

$$\begin{bmatrix} 3 & 3 & 4 & 1 & 0 & 3 \\ 0 & -4 & -\frac{7}{3} & \frac{5}{3} & 2 & -2 \\ 0 & 0 & \frac{5}{2} & \frac{5}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_5 \\ x_6 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Tähän muotoon pääsemiseksi ei ole tehty rivinvaihtoja. Sarakkeita sen sijaan on vaihdettu niin, että uusi järjestys on 125634, kuten näkyy lopputuloksesta. Päätellään siis: dim W = dim $U = 3$, U:n erään kannan muodostavat $\mathbf{A}$:n kolme ensimmäistä riviä ja W:n erään kannan $\mathbf{A}$:n sarakkeet 1, 2 ja 5. $\square$

Em. esimerkeissä matriisin sarakkeiden ja rivien virittämien avaruuksien dimensiot ovat samat. Kyse ei ole sattumasta, vaan dimensiot ovat <u>aina</u> samat: dimensioiden yhteinen arvo = matriisin säännöllisyysaste (ks. Luku X.4).

*Ortogonaaliprojektio $\mathbb{R}^n$:ssä

Olkoon $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, $1 \leq m < n$, lineaarisesti riippumaton vektorisysteemi $\mathbb{R}^n$:ssä, ja olkoon W ko. vektorien virittämä R^n:n (aito) aliavaruus:

$$W = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} = \sum_{i=1}^{m} x_i \mathbf{a}_i, \quad x_i \in \mathbb{R}\}.$$

Tällöin jos $\mathbf{u} \in \mathbb{R}^n$, niin $\mathbf{u}$:n *ortogonaaliprojektio* aliavaruuteen W määritellään vektorina $\mathbf{w}$, joka toteuttaa

$$\mathbf{w} \in W \quad \& \quad \langle \mathbf{u} - \mathbf{w}, \mathbf{v} \rangle = 0 \;\; \forall\, \mathbf{v} \in W.$$

Kirjoittamalla tässä $\mathbf{v}$ vektoreiden $\mathbf{a}_1, \ldots, \mathbf{a}_m$ lineaarikombinaationa (mahdollista jokaiselle $\mathbf{v} \in W$) ja käyttämällä skalaaritulon bilineaarisuutta nähdään, että ortogonaalisuusehto riittää asettaa vektoreille $\mathbf{v} = \mathbf{a}_i$, $i = 1 \ldots m$:

$$\langle \mathbf{u} - \mathbf{w}, \mathbf{v} \rangle = 0 \;\; \forall\, \mathbf{v} \in W \quad \Leftrightarrow \quad \langle \mathbf{u} - \mathbf{w}, \mathbf{a}_i \rangle = 0, \quad i = 1 \ldots m.$$

Ortogonaaliprojektion haku on näin pelkistetty yhtälöryhmäksi. Kun tässä kirjoitetaan vielä $\mathbf{w} = \sum_{j=1}^{m} x_j \mathbf{a}_j$ (mahdollista, koska $\mathbf{w} \in W$), ja käytetään edelleen skalaaritulon bilineaarisuutta ja symmetrisyyttä, niin nähdään, että kyseessä on seuraava lineaarinen yhtälöryhmä kerroinvektorille $\mathbf{x} = [x_1, \ldots, x_m]^T \in \mathbb{R}^m$:

$$\boxed{\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad [\mathbf{A}]_{ij} = \langle \mathbf{a}_i, \mathbf{a}_j \rangle, \quad [\mathbf{b}]_i = \langle \mathbf{u}, \mathbf{a}_i \rangle, \quad i, j = 1 \ldots m.}$$

Tätä muotoilua sanotaan projektio-ongelman *normaalimuodoksi.*

Projektio-ongelmaan päädytään lineaarialgebran sovelluksissa usein minimointiongelman kautta: Halutaan etsiä annettua vektoria $\mathbf{u} \in \mathbb{R}^n$ jonkin normin $\|\cdot\|$ mielessä lähin vektori aliavaruudesta W, ts. halutaan ratkaista minimointiongelma

$$\mathbf{v} \in W : \quad \|\mathbf{u} - \mathbf{v}\| = \min!$$

Jos normi liittyy skalaarituloon $\langle \cdot, \cdot \rangle$ siten, että $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$ (ks. Luku II.4), niin ongelman ratkaisu saadaan $\mathbf{u}$:n ortogonaaliprojektiona $\mathbf{w} \in W$ ko. skalaaritulon suhteen. Nimittäin jos $\mathbf{v} \in W$ on mielivaltainen ja merkitään $\mathbf{v}_0 = \mathbf{w} - \mathbf{v}$, niin $\langle \mathbf{u} - \mathbf{w}, \mathbf{v}_0 \rangle = 0$ (koska $\mathbf{v}_0 \in W$), joten seuraa

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\|^2 &= \|(\mathbf{u} - \mathbf{w}) + \mathbf{v}_0)\|^2 \;=\; \langle\, (\mathbf{u} - \mathbf{w}) + \mathbf{v}_0\,,\, (\mathbf{u} - \mathbf{w}) + \mathbf{v}_0\,\rangle \\ &= |\mathbf{u} - \mathbf{w}|^2 + 2\,\langle \mathbf{u} - \mathbf{w}, \mathbf{v}_0 \rangle + |\mathbf{v}_0|^2 \\ &= |\mathbf{u} - \mathbf{w}|^2 + |\mathbf{v}_0|^2. \end{aligned}$$

Siis $\|\mathbf{u} - \mathbf{v}\|$:n minimiarvo saavutetaan, kun $\mathbf{v}_0 = \mathbf{0} \;\Leftrightarrow\; \mathbf{v} = \mathbf{w}$.

Projektio-ongelman yksikäsitteisen ratkeavuuden takaa

Lause X.6.2 ($\mathbb{R}^n$**:n projektiolause**) Jos $W \subset \mathbb{R}^n$ on lineaarisesti riippumattomien vektorien $\mathbf{a}_1, \ldots, \mathbf{a}_m$ virittämä $\mathbb{R}^n$:n aliavaruus ($m < n$) ja $\mathbf{u} \in \mathbb{R}^n$, niin on olemassa yksikäsitteinen $\mathbf{w} \in W$ siten, että $\langle \mathbf{u} - \mathbf{w}, \mathbf{v} \rangle = 0 \;\; \forall\, \mathbf{v} \in W$.

Todistus Lause väittää, että matriisi $\mathbf{A} = (\langle \mathbf{a}_i, \mathbf{a}_j \rangle)$ (neliömatriisi kokoa $m \times m$) on säännöllinen. Tämän näyttämiseksi olkoon $\mathbf{w} = \sum_{i=1}^m x_i \mathbf{a}_i$ ja lasketaan ensin

$$|\mathbf{w}|^2 \;=\; \langle \mathbf{w}, \mathbf{w} \rangle \;=\; \langle \sum_{i=1}^m x_i \mathbf{a}_i, \sum_{j=1}^m x_j \mathbf{a}_j \,\rangle.$$

Skalaaritulon bilineaarisuuden ja matriisialgebran nojalla tämä purkautuu muotoon

$$|\mathbf{w}|^2 \;=\; \sum_{i=1}^m \sum_{j=1}^m \langle \mathbf{a}_i, \mathbf{a}_j \rangle x_i x_j \;=\; \mathbf{x}^T \mathbf{A} \mathbf{x}.$$

Tämän perusteella päätellään seuraavasti:

$$\mathbf{A}\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \quad \Rightarrow \quad |\mathbf{w}|^2 = 0 \quad \Rightarrow \quad \mathbf{w} = \mathbf{0}.$$

Koska $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ oli lineaarisesti riippumaton systeemi, niin päätellään edelleen: $\mathbf{w} = \mathbf{0} \;\Rightarrow\; \mathbf{x} = \mathbf{0}$. On siis päätelty: $\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x} = \mathbf{0}$, joten $\mathbf{A}$ on säännöllinen. $\square$

Projektio-ongelman normaalimuodon ratkaisu $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ antaa siis ortogonaaliprojektion $\mathbf{w}$ koordinaatit W:n kannassa $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$. Ratkaisua varten on ensin laskettava matriisi $\mathbf{A}$ ja vektori $\mathbf{b}$. Nämä riippuvat vektoreista $\mathbf{u}$ ja $\{\mathbf{a}_i\}$ sekä valitusta $\mathbb{R}^n$:n skalaaritulosta, joka siis voi olla muukin kuin euklidinen. Vektorin $\mathbf{b}$ riippuvuus vektorista $\mathbf{u}$ voidaan purkaa matriisialgebraksi:

$$\mathbf{u} = \sum_{j=1}^{n} u_j \mathbf{e}_j \quad \Rightarrow \quad [\mathbf{b}]_i \;=\; \langle \mathbf{u}, \mathbf{a}_i \rangle \;=\; \langle \mathbf{a}_i, \mathbf{u} \rangle \;=\; \sum_{j=1}^{n} \langle \mathbf{a}_i, \mathbf{e}_j \rangle u_j.$$

Siis $\mathbf{b} = \mathbf{B}\mathbf{u}$, missä

$$[\mathbf{B}]_{ij} \;=\; \langle \mathbf{a}_i, \mathbf{e}_j \rangle, \quad i = 1 \ldots m,\; j = 1 \ldots n.$$

(Euklidisen skalaaritulon tapauksessa on $[\mathbf{B}]_{ij} = [\mathbf{a}_i]_j$). Tämän tulkinnan mukaisesti ortogonaaliprojektio $\mathbf{u} \mapsto \mathbf{x}$ on lineaarikuvaus tyyppiä $P:\ \mathbb{R}^n \to \mathbb{R}^m$ ja P:n matriisi on $\mathbf{A}^{-1}\mathbf{B}$.

Esimerkki 7 Laske vektorin $\mathbf{u} = [0, 1, 0, 0]^T$ ortogonaaliprojektio euklidisen skalaaritulon suhteen aliavaruuteen $W \subset \mathbb{R}^4$, jonka virittävät vektorit $\mathbf{a}_1 = [1, 1, 1, 1]^T$, $\mathbf{a}_2 = [1, 1, 1, 0]^T$ ja $\mathbf{a}_3 = [0, 1, 1, 1]^T$. Määritä myös yksikkövektori $\mathbf{n} \in \mathbb{R}^4$, joka on ortogonaalinen aliavaruutta W vastaan.

<u>Ratkaisu</u> Projektioperiaatteen mukaisesti vektorin $\mathbf{w} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3$ koordinaatit x_i määräytyvät ratkaisemalla yhtälöryhmä $\mathbf{A}\mathbf{x} = \mathbf{b}$, missä

$$\mathbf{A} = (\,\langle \mathbf{a}_i, \mathbf{a}_j \rangle\,) = \begin{bmatrix} 4 & 3 & 3 \\ 3 & 3 & 2 \\ 3 & 2 & 3 \end{bmatrix}, \quad \mathbf{b} = \mathbf{B}\mathbf{u} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Ratkaisu on $\mathbf{x} = \frac{1}{2}\,[-1, 1, 1]^T$, joten kysytty projektio on

$$\mathbf{w} \;= \frac{1}{2}(-\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)^T \;=\; \frac{1}{2}[0, 1, 1, 0]^T.$$

Vektori

$$\mathbf{n} \;=\; \lambda\,(\mathbf{u} - \mathbf{w}) \;=\; \frac{1}{2}\lambda\,[0, 1, -1, 0]^T$$

on ortogonaalinen W:tä vastaan, joten kysyttyjä yksikkövektoreita ovat

$$\mathbf{n} = \pm\frac{1}{\sqrt{2}}\,[0, 1, -1, 0]^T. \qquad \square$$

HARJOITUSTEHTÄVIÄ

1. a) Lineaarikuvaukselle $A : \mathbb{R}^3 \mapsto \mathbb{R}^3$ pätee: $A\mathbf{a} = (0, 2, 1)$, $A\mathbf{b} = (1, 2, 3)$ ja $A\mathbf{c} = (1, -1, 0)$. Laske $A(3\mathbf{a} - 2\mathbf{b} + \mathbf{c})$.
   b) Lineaarikuvaukselle $A : \mathbb{R}^2 \mapsto \mathbb{R}^2$ pätee: $A(1, 1) = (3, -1)$ ja $A(2, -1) = (1, 2)$. Laske A:n matriisi.
   c) Funktiolle $\mathbf{f} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ pätee: $\mathbf{f}(0, 1, 1) = (1, 0, 0)$, $\mathbf{f}(1, 0, 1) = (1, 1, 0)$ ja $\mathbf{f}(1, -1, 0) = (1, 1, 1)$. Voiko $\mathbf{f}$ olla lineaarikuvaus?

2. Olkoon $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ lineaarikuvaus ja $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$.
   a) Näytä, että jos $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ on lineaarisesti riippuva niin samoin on $\{A\mathbf{a}_1, \ldots, A\mathbf{a}_m\}$.
   b) Näytä, että jos $\{A\mathbf{a}_1, \ldots, A\mathbf{a}_m\}$ on lineaarisesti riippumaton, siin samoin on $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$.
   c) Oletetaan lisäksi, että A on kääntyvä. Näytä, että tällöin $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ on lineaarisesti riippuva/riippumaton täsmälleen kun $\{A\mathbf{a}_1, \ldots, A\mathbf{a}_m\}$ on lineaarisesti riippuva/riippumaton.

3. Laske matriisimuotoiset koordinaattien muunnoskaavat $\mathbf{x} = \mathbf{A}\mathbf{x}'$ ja $\mathbf{x}' = \mathbf{B}\mathbf{x}$ seuraaville $\mathbb{R}^n$:n tai vastaavan geometrisen vektoriavaruuden kannan vaihdoille.
   a) $\{\mathbf{e}_1, \mathbf{e}_2\} \hookrightarrow \{\mathbf{e}_1 + 9\mathbf{e}_2, 6\mathbf{e}_1 + 8\mathbf{e}_2\}$
   b) $\{\vec{i}, \vec{j}\} \hookrightarrow \{\vec{i} - 2\vec{j}, 3\vec{i} + \vec{j}\}$
   c) $\{\mathbf{e}_1, \mathbf{e}_2\} \hookrightarrow \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{a} = [5, -1]^T$, $\mathbf{b} = [1, 5]^T$
   d) $\{\mathbf{a}, \mathbf{b}\} \hookrightarrow \{\mathbf{c}, \mathbf{d}\}$, $\mathbf{a} = [1, 1]^T$, $\mathbf{b} = [1, 2]^T$, $\mathbf{c} = [2, 1]^T$, $\mathbf{d} = [1, -1]^T$
   e) $\{\vec{i}, \vec{j}, \vec{k}\} \hookrightarrow \{\vec{i} + 2\vec{j}, 3\vec{j} + 4\vec{k}, 6\vec{i} + 5\vec{k}\}$
   f) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \hookrightarrow \{\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3, -\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3\}$
   g) $\{\mathbf{e}_1, \ldots, \mathbf{e}_4\} \hookrightarrow \{\mathbf{a}_1, \ldots, \mathbf{a}_4\}$, $\mathbf{a}_i = \mathbf{e}_1 + \ldots + \mathbf{e}_i$, $i = 1 \ldots 4$
   h) $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \hookrightarrow \{\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2 + \mathbf{e}_3, \ldots, \mathbf{e}_{n-1} + \mathbf{e}_n, \mathbf{e}_1 + \mathbf{e}_n\}$, $n = 7$

4. Tutki, ovatko seuraavat vektorit lineaarisesti riippumattomat.
   a) $[0, 0, 4, 1]$, $[2, 1, -1, 1]$, $[1, -1, 2, 1]$, $[1, 2, 1, 1]$
   b) $[1, -1, 1, -1, 1]$, $[1, 1, 1, 1, 1]$, $[1, 1, -1, 1, 1]$, $[1, 1, 1, -1, 1]$
   c) $[1, 1, 1, 1, 1, 1, 1, 1, 1]$, $[0, 1, 1, 1, 1, 1, 1, 1, 1, 1]$, $\ldots$, $[0, 0, 0, 0, 0, 0, 0, 0, 1]$

5. Konstruoi jatkuva, kannan ortonormeerauksen ja suunnistuksen säilyttävä muunnos $t \in [0, 1] \mapsto \{\mathbf{a}(t), \mathbf{b}(t)\}$ $\mathbb{R}^2$:n kannasta $\{\mathbf{e}_1, \mathbf{e}_2\}$ annettuun, ortonormeerattuun ja positiivisesti suunnistettuun kantaan $\{\mathbf{a}, \mathbf{b}\}$.
   Vihje: Esimerkki 4.

6. Näytä, että seuraavissa $\mathbb{R}^3$:n koordinaattien muunnoskaavoissa on kyse koordinaatiston kierrosta. Mikä on käänteinen muunnoskaava? Minkä pisteiden koordinaatit pysyvät kierrossa ennallaan?

$$\text{a)} \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 3 & -2 & 6 \\ 2 & -6 & -3 \\ 6 & 3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\text{b)} \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 10 & -5 & 10 \\ -11 & -2 & 10 \\ -2 & -14 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

7. Seuraavien matriisien sarakket ja rivit virittävät erään $\mathbb{R}^n$:n aliavaruuden ($n = 3, 4, 5$ tai 7). Määritä tuetulla Gaussin algoritmilla ko. avaruuden dimensio ja jokin kanta.

$$\text{a)} \quad \begin{bmatrix} 1 & 5 & 0 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \qquad \text{b)} \quad \begin{bmatrix} 3 & 4 & -2 & 1 \\ 2 & 1 & 3 & -2 \\ 1 & -2 & 8 & -5 \end{bmatrix}$$

$$\text{c)} \quad \begin{bmatrix} 1 & 2 & -3 & 1 & 5 \\ 3 & 0 & 3 & -3 & 3 \\ 2 & -1 & 4 & -3 & 1 \\ 0 & 2 & -4 & 2 & 4 \end{bmatrix} \qquad \text{d)} \quad \begin{bmatrix} 1 & 0 & -1 & -1 & 3 & 3 & 5 \\ 1 & 1 & 1 & 2 & 2 & -1 & 3 \\ 1 & 2 & 3 & 5 & 1 & -5 & 1 \\ 3 & 2 & 1 & 3 & 7 & 2 & 11 \end{bmatrix}$$

8. (*) $\mathbb{R}^3$:n yleinen kierretty (ortonormeerattu ja positiivisesti suunnistettu) kanta $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ syntyy kiertämällä kantaa $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ x_1-akselin ympäri kulma α, kierron tulosta x_2-akselin ympäri kulma β ja lopuksi näiden kiertojen tulosta x_3-akselin ympäri kulma γ. Määritä vektorit $\mathbf{a}, \mathbf{b}, \mathbf{c}$ kulmien α, β, γ avulla. Millä kulmien arvoilla syntyy a) pallokoordinaattikanta $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$ vastaten pallonpintakoordinaatteja θ, φ, b) tehtävän 6a kierretty kanta?

9. (*) Olkoon $W \subset \mathbb{R}^4$ aliavaruus, jonka virittävät vektorit $\mathbf{a} = [1, 1, 1, 1]$ ja $\mathbf{b} = [1, 1, -1, -1]$. Määritä vektorin $\mathbf{u} = [1, 2, -1, 3]$ ortogonaaliprojektio aliavaruuteen W a) euklidisen skalaaritulon, b) skalaaritulon $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + 2x_3y_3 + 3x_4y_4$ mielessä. Millaisen minimointiongelman muotoa

$$\mathbf{x} \in W : \; f(\mathbf{x}) \; = \; \min!$$

nämä ortogonaaliprojektiot ratkaisevat?

X.7 Affiinikuvaukset. Geometriset kuvaukset

Affiinikuvaukset ovat lineaarikuvauksia hieman yleisempiä kuvauksia tyyppiä

$$\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^m, \quad \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b},$$

missä $\mathbf{A}$ on matriisi kokoa $m \times n$ ja $\mathbf{b} \in \mathbb{R}^m$. Affiinikuvaukset ovat oma tärkeä luokkansa lineaaristen ($\mathbf{b} = \mathbf{0}$) ja yleisempien *epälineaaristen* (= ei-lineaaristen) kuvausten välimaastossa. Ominaisuuksiltaan affiinikuvaukset ovat lähellä lineaarikuvauksia. Esim. jos $\mathbf{f}_1(\mathbf{x}) = \mathbf{A}_1\mathbf{x} + \mathbf{b}_1$ on affiinikuvaus tyyppiä $\mathbf{f}_1 : \mathbb{R}^n \to \mathbb{R}^p$ ja $\mathbf{f}_2(\mathbf{x}) = \mathbf{A}_2\mathbf{x} + \mathbf{b}_2$ on affiinikuvaus typpiä $\mathbf{f}_2 : \mathbb{R}^p \to \mathbb{R}^m$, niin yhdistetty kuvaus $\mathbf{f}_2 \circ \mathbf{f}_1$ on määritelty ja myös affiininen:

$$(\mathbf{f}_2 \circ \mathbf{f}_1)(\mathbf{x}) = \mathbf{A}_2(\mathbf{A}_1\mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2 = \mathbf{A}_2\mathbf{A}_1\mathbf{x} + (\mathbf{A}_2\mathbf{b}_1 + \mathbf{b}_2), \quad \mathbf{x} \in \mathbb{R}^n.$$

Samoin nähdään, että jos $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ on tyyppiä $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$, niin $\mathbf{f}$ on kääntyvä täsmälleen kun $\mathbf{A}$ on säännöllinen matriisi, jolloin myös käänteiskuvaus $\mathbf{f}^{-1}$ on affiininen.

Edellisessä luvussa todettiin, että $\mathbb{R}^n$:n kannan vaihtoa vastaava koordinaattimuunnos on lineaarikuvaus. Jos kannan vaihtoon yhdistetään myös *origon siirto*, niin koordinaattimuunnoksesta tulee affiinikuvaus.

Esimerkki 1 Esitä affiinikuvauksina koordinaattien väliset muunnoskaavat E^3:n koordinaatistojen $(O, \vec{i}, \vec{j}, \vec{k})$ ja $(O', \vec{a}, \vec{b}, \vec{c})$ välillä, kun ensin mainitussa koordinaatistossa on $O' = (1, -2, 2)$ ja

$$\vec{a} = \vec{i} + \vec{j} - \vec{k}, \quad \vec{b} = \vec{j} + 2\vec{k}, \quad \vec{c} = \vec{i} + \vec{k}.$$

<u>Ratkaisu</u> Jos pisteen P koordinaatit ovat (x, y, z) ja (x', y', z'), niin on oltava voimassa

$$x\vec{i} + y\vec{j} + z\vec{k} = \overrightarrow{OO'} + x'\vec{a} + y'\vec{b} + z'\vec{c}.$$

Tämä on tulkittavissa vektorimuotoiseksi muunnoskaavaksi $(x', y', z') \mapsto (x, y, z)$, jonka matriisimuoto on affiinikuvaus:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

Ratkaisemalla tästä $[x', y', z']^T$ saadaan käänteismuunnokseksi

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 1 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 5 \\ 3 \\ -9 \end{bmatrix}. \quad \square$$

Geometriset kuvaukset

Geometrisilla kuvauksilla ymmärretään yleensä euklidisen pisteavaruuden E^2 tai E^3 kuvauksia, joilla on jokin (verrattain yksinkertainen) geometrinen merkitys. Koordinaattiavaruuteen palauttamalla voidaan geometriset kuvaukset aina tulkita myös $\mathbb{R}^2$:n tai $\mathbb{R}^3$:n kuvauksina. Jatkossa rajoitutaan yksinkertaisimpiin geometrisiin kuvauksiin, jotka ovat affiinisia. Tällaisia ovat mm. seuraavat kuvaukset $P \mapsto Q$ tyyppiä $f:\ E^2 \to E^2$ tai $f:\ E^3 \to E^3$ tai näiden yhdistelyt.

— *siirto* eli *translaatio*: $\overrightarrow{OQ} = \overrightarrow{OP} + \vec{b}$ ($\vec{b}$ annettu)

— *kierto* tasossa origon ympäri: $\overrightarrow{OQ} = \overrightarrow{OP}$ kierrettynä kulman φ verran vastapäivään ($\varphi \in \mathbb{R}$ annettu)

— *yhdensuuntaisprojektio* avaruustasolle T (tason suoralle S) suoran L, tai vektorin $\vec{v} = L$:n suuntavektori, suunnassa: $\mathbb{Q} \in T$ tai $Q \in S$ ja $\overrightarrow{PQ} = \lambda\vec{v},\ \lambda \in \mathbb{R}$ ($\vec{v}$ ei T:n/S:n suuntainen)

— *suuntaisprojektio avaruussuoralle* S tason T, tai vektorien $\vec{v}_1$, $\vec{v}_2 = T$:n suuntavektorit, suunnassa : $Q \in S$ ja $\overrightarrow{PQ} = \lambda\vec{a}_1 + \mu\vec{a}_2,\ (\lambda, \mu) \in \mathbb{R}^2$ ($\{\vec{v}, \vec{a}_1, \vec{a}_2\}$ lineaarisesti riippumaton, $\vec{v} = S$:n suuntavektori)

— *peilaus* annetun tason (avaruudessa) tai suoran (tasossa) suhteen

— *skaalaus* eli *homotetia* origon suhteen: $\overrightarrow{OQ} = \lambda\,\overrightarrow{OP}$ ($\lambda \in \mathbb{R}$ annettu)

Geometrirista kuvauksista yksinkertaisimmat ovat siirto ja homotetia, joiden affiinimuodot ovat $\mathbf{f}(\mathbf{x}) = \mathbf{I}\mathbf{x} + \mathbf{b}$ ja $\mathbf{f}(\mathbf{x}) = \lambda\mathbf{I}\mathbf{x}$ ($\mathbf{I}$ yksikkömatriisi). Yleinen affinikuvaus $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ on lineaarikuvauksen ja siirron yhdistelmä:

$$\mathbf{x} \longmapsto \mathbf{A}\mathbf{x} \longmapsto \mathbf{A}\mathbf{x} + \mathbf{b}.$$

Jos yhdensuuntaisprojektiossa taso T/suora S tai suuntaisprojektiossa suora S kulkee origon kautta, on $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, jolloin projektio on lineaarikuvuas. Näiden projektioiden erikoistapauksia ovat ortogonaaliprojektiot, joissa projisoidaan tasoa/suoraa vastaan kohtisuorasti.

Esimerkki 2 Taso T kulkee origon kautta ja sen suuntavektorit ovat $\vec{v}_1 = \vec{i} + \vec{j}$ ja $\vec{v}_2 = \vec{j} + \vec{k}$. Piste $P \mathrel{\widehat{=}} x_1\vec{i} + x_2\vec{j} + x_3\vec{k} \mathrel{\widehat{=}} [x_1, x_2, x_3]^T = \mathbf{x}$ projisoidaan kohtisuorasti tason T pisteeksi

$$Q \mathrel{\widehat{=}} y_1\vec{i} + y_2\vec{j} + y_3\vec{k} \;=\; z_1\vec{v}_1 + z_2\vec{v}_2.$$

Määritä vastaavien lineaarikuvausten

$$A: \ \mathbf{x} \longmapsto \mathbf{y} = [y_1, y_2, y_3]^T, \qquad B: \ \mathbf{x} \longmapsto \mathbf{z} = [z_1, z_2]^T$$

matriisit $\mathbf{A}, \mathbf{B}$.

<u>Ratkaisu</u> Projektioehto on

$$\begin{aligned}
&(x_1\vec{i} + x_2\vec{j} + x_3\vec{k} - z_1\vec{v}_1 - z_2\vec{v}_2) \cdot \vec{v}_i = 0, \quad i = 1, 2 \\
\Leftrightarrow \quad & \begin{cases} x_1 + x_2 - 2z_1 - z_2 = 0 \\ x_2 + x_3 - z_1 - 2z_2 = 0 \end{cases} \\
\Leftrightarrow \quad & \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
\Leftrightarrow \quad & \mathbf{Cz} = \mathbf{Dx}.
\end{aligned}$$

Siis $\mathbf{z} = \mathbf{Bx}$, missä

$$\mathbf{B} \ = \ \mathbf{C}^{-1}\mathbf{D} \ = \ \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \ = \ \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \end{bmatrix}.$$

Koska

$$y_1\vec{i} + y_2\vec{j} + y_3\vec{k} \ = \ z_1\vec{v}_1 + z_2\vec{v}_2 \ = \ z_1\vec{i} + (z_1 + z_2)\vec{j} + z_2\vec{k},$$

niin

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{B} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Siis $\mathbf{y} = \mathbf{Ax}$, missä

$$\mathbf{A} \ = \ \frac{1}{3}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \end{bmatrix} \ = \ \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

<u>Tarkistus</u>: Soveltamalla yhtälöryhmään $\mathbf{Ax} = \mathbf{0}$ Gaussin algoritmia saadaan ratkaisuksi $\mathbf{x} = t[1, -1, 1]^T$, $t \in \mathbb{R}$. Tämä sopii yhteen sen tiedon kanssa, että tason T normaalivektori on $\vec{n} = \vec{i} - \vec{j} + \vec{k}$. $\square$

Ortogonaaliprojektion avulla voidaan myös peilauskuvaus määritellä helposti: Jos P_0 on P:n projektiopiste, niin $\overrightarrow{OQ} = \overrightarrow{OP} + 2\,\overrightarrow{PP_0}$ määrittää peilikuvan Q.

Esimerkki 2 (jatko) Taso T kulkee pisteen $(1,-1,2)$ kautta ja tason suuntavektoreita ovat $\vec{v}_1 = \vec{i} + \vec{j}$ ja $\vec{v}_2 = \vec{j} + \vec{k}$. Määritä peilauskuvaus $(x_1, x_2, x_3) \triangleq P \longmapsto Q \triangleq (y_1, y_2, y_3)$ affiinikuvauksena $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$.

Ratkaisu Merkitään $\mathbf{a} = [1,-1,2]^T$ ja olkoon projektiopiste $P_0 \triangleq (z_1, z_2, z_3)$. Tällöin $\mathbf{z} - \mathbf{a} = A(\mathbf{x} - \mathbf{a})$, missä A on edellisessä esimerkissä laskettu lineaarikuvaus. Kun tämän matriisia merkitään nyt symbolilla $\mathbf{A}_0$, niin on siis

$$\mathbf{z} - \mathbf{a} = \mathbf{A}_0(\mathbf{x} - \mathbf{a}), \quad \mathbf{A}_0 \;=\; \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

Näin ollen

$$\begin{aligned} \mathbf{y} \;&=\; \mathbf{x} + 2(\mathbf{z} - \mathbf{x}) \;=\; (2\mathbf{A}_0 - \mathbf{I})\mathbf{x} + 2(\mathbf{I} - \mathbf{A}_0)\mathbf{a} \\ &=\; \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \frac{8}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \\ &=\; \mathbf{A}\mathbf{x} + \mathbf{b}. \qquad \square \end{aligned}$$

Esimerkki 3 Tulkitse kierto tasossa lineaarikuvauksina $\mathbf{x} \mapsto \mathbf{y} = \mathbf{A}\mathbf{x}$.

Ratkaisu Vektorin kierto tasossa annetun kulman φ verran on tulkittavissa joko lineaarikuvauksena tyyppiä $A : \mathbb{R}^2 \to \mathbb{R}^2$ tai kompleksitason kuvauksena $z = x_1 + ix_2 \longmapsto w = y_1 + iy_2$. Jälkimmäisen tulkinnan perusteella on

$$\begin{aligned} y_1 + iy_2 &= (\cos\varphi + i\sin\varphi)(x_1 + ix_2) \\ &= (x_1\cos\varphi - x_2\sin\varphi) + i(x_1\sin\varphi + x_2\cos\varphi), \end{aligned}$$

joten A:n matriisi on

$$\mathbf{A} \;=\; \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \qquad \square$$

Myös pisteen *kierto avaruussuoran ympäri* on tulkittavissa affiinikuvauksena: Jos P_0 on P:n kohtisuora projektio suoralle S (erikoistapaus yhdensuuntaisprojektiosta avaruussuoralle), niin $\overrightarrow{OQ} = \overrightarrow{OP_0} + \overrightarrow{P_0Q}$, missä $\overrightarrow{P_0Q}$ = vektori $\overrightarrow{P_0P}$ kierrettynä pisteen P_0 kautta kulkevassa tasossa, jonka normaali on S.

Esimerkki 3 (jatko) Vektorin $\vec{r} = x_1\vec{i} + x_2\vec{j} + x_3\vec{k}$ kiertyessä x_1-akselin ympäri pysyy komponentti x_1 muuttumattomana ja $[x_2, x_3]^T$ muuntuu kuten tasokierrossa. Näin ollen kierto vastaa lineaarikuvausta $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, missä

$$\mathbf{A} \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}. \qquad \square$$

Siirron, taso- tai avaruuskierron, peilauksen ja skaalauksen yhdistelmiä E^n:ssä ($n = 2$ tai $n = 3$) kutsutaan *euklidisiksi kuvauksiksi*†. Euklidisen tason kahta pistejoukkoa $K_1 \subset E^2$ ja $K_2 \subset E^2$ (nämä voivat olla geometrisia kuvioita kuten kolmioita) sanotaan *yhdenmuotoisiksi*, jos on olemassa euklidinen kuvaus $\mathbf{f} : E^2 \to E^2$ siten, että $K_2 = \mathbf{f}(K_1)$. Jos $\mathbf{f}$ koostuu vain siirroista, kierroista ja peilauksista, niin sanotaan, että K_1 ja $\mathbf{f}(K_2)$ ovat *yhtenevät*.

*$\mathbb{R}^n$:n geometriaa

Geometrinen ajattelu voidaan ulottaa yleisesti avaruuteen $\mathbb{R}^n$, jolloin $\mathbb{R}^n$:n alkiot ovat (kuvitellun) euklidisen pisteavaruuden E^n pisteiden vastineita. Sanonta 'pisteessä $\mathbf{x}$' on tällöin tulkittava niin, että piste $P \in E^n$ ja sen vastine $\mathbf{x} \in \mathbb{R}^n$ (= P:n koordinaatit $\mathbb{R}^n$:n valitussa ortonormeeratussa kannassa) samastetaan, tai jälkimmäinen tulkitaan edellisen 'nimeksi'. Kun omaksutaan tämä ajattelutapa, niin geometrian pisteet, janat, kolmiot ym. voidaan tuoda luontevasti suoraan avaruuteen $\mathbb{R}^n$.

Seuraavassa luetellaan eräitä tavallisimpia geometrian käsitteitä avaruuteen $\mathbb{R}^n$ tuotuina ja yleistettyinä:

— *Piste* on alkio $\mathbf{a} \in \mathbb{R}^n$.

— *Suora* on pistejoukko $S = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + t\mathbf{v},\ t \in \mathbb{R}\}$, missä $\mathbf{x}_0, \mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{v} \neq \mathbf{0}$.

— *Hypertaso* (*taso* tapauksessa $n = 3$) on pistejoukko $T = \{\mathbf{x} \in \mathbb{R} \mid \mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{n-1} t_i\mathbf{v}_i,\ \mathbf{t} = (t_i) \in \mathbb{R}^{n-1}\}$, missä $\mathbf{x}_0, \mathbf{v}_i \in \mathbb{R}^n$ ja $\{\mathbf{v}_1, \ldots \mathbf{v}_{n-1}\}$ on lineaarisesti riippumaton.

— *m-taso* (m-ulotteinen taso) on pistejoukko $T = \{\mathbf{x} \in \mathbb{R} \mid \mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{m} t_i\mathbf{v}_i,\ \mathbf{t} = (t_i) \in \mathbb{R}^m\}$, missä $2 \leq m \leq n-1$, $\mathbf{x}_0, \mathbf{v}_i \in \mathbb{R}^n$ ja $\{\mathbf{v}_1, \ldots \mathbf{v}_m\}$ on lineaarisesti riippumaton.

— *Jana* on pistejoukko $S = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \lambda\mathbf{a}_1 + (1-\lambda)\mathbf{a}_2,\ \lambda \in [0,1]\}$, missä $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^n$ ja $\mathbf{a}_1 \neq \mathbf{a}_2$.

— *Kolmio* on pistejoukko $K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^{3} \lambda_i\mathbf{a}_i,\ \lambda_i \in [0,1],\ \ \sum_{i=1}^{3} \lambda_i = 1\}$, missä $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^n$ ja $\{\mathbf{a}_2 - \mathbf{a}_1, \mathbf{a}_3 - \mathbf{a}_1\}$ on lineaarisesti riippumaton.

†Kuvausta $K \longmapsto \mathbf{f}(K)$, missä $\mathbf{f}$ on pelkästään siirtojen ja kiertojen yhdistelmä, sanotaan *euklidiseksi liikkeeksi*, mekaniikassa (kun $K \subset E^3$ edustaa kiinteää kappaletta) jäykän kappaleen liikkeeksi.

— *m-simpleksi* ($2 \leq m \leq n+1$) on pistejoukko
$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i,\ \lambda_i \in [0,1],\ \ \sum_{i=1}^m \lambda_i = 1\}$, missä
$\mathbf{a}_i \in \mathbb{R}^n$ ja $\{\mathbf{a}_2 - \mathbf{a}_1, \ldots, \mathbf{a}_m - \mathbf{a}_1\}$ on lineaarisesti riippumaton.

— *m-suuntaissärmiö* ($2 \leq m \leq n$) on pistejoukko
$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^m \lambda_i \mathbf{a}_i,\ \lambda_i \in [0,1]\}$, missä $\mathbf{x}_0, \mathbf{a}_i \in \mathbb{R}^n$ ja $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ on lineaarisesti riippumaton

— *Pallo (kuula)* on pistejoukko $B(\mathbf{x}_0, R) = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{x}_0| \leq R\}$, missä $\mathbf{x}_0 \in \mathbb{R}^n$ ja $R > 0$.

— *Pallopinta* on pistejoukko $K = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{x}_0| = R\}$, missä $\mathbf{x}_0 \in \mathbb{R}^n$ ja $R > 0$.

— *n-suorakulmio* (koordinaatiakselien mukaan suunnattu) on pistejoukko
$K = \{\mathbf{x} = (x_i) \in \mathbb{R}^n \mid x_i \in [a_i, b_i],\ i = 1 \ldots n\}$, missä
$a_i, b_i \in \mathbb{R},\ a_i < b_i,\ i = 1 \ldots n$.

Määritelmien mukaisesti hypertaso = $(n-1)$–ulotteinen taso, 2-simpleksi = jana, 3-simpleksi = kolmio ja $\mathbb{R}^3$:n 4-simpleksi = tetraedri. Hypertason ja m-tason määritelmissä sanotaan vektoreita $\mathbf{v}_i$ ko. tason *suuntavektoreiksi*. Simpleksin määritelmässä pisteet $\mathbf{a}_i$ ovat simpleksin *kärkipisteet* (engl. vertex). Jos n-suorakulmiossa on $b_i - a_i = a,\ i = 1 \ldots n$, niin kyseessä on *n-kuutio* (sivun pituus a).

Ym. määritelmässä on hypertaso ja m-taso esitetty *parametrisessa* muodossa suuntavektoreittensa $\mathbf{v}_i$ avulla. Hypertasolle voidaan myös kirjoittaa *hypertason yhtälö* muodossa

$$T: \quad \langle \mathbf{x} - \mathbf{x}_0, \mathbf{n} \rangle = \mathbf{n}^T(\mathbf{x} - \mathbf{x}_0) = 0,$$

missä vektori $\mathbf{n}$ on T:n *normaali*, eli kohtisuorassa vektoreita $\mathbf{v}_i$ vastaan. Tällainen $\mathbf{n}$ on löydettävissä esim. ortogonaaliprojektion avulla, vrt. Esimerkki 7 edellisessä luvussa. Yleisemmin voidaan m-taso määritellä yhtälöryhmällä

$$T: \quad \langle \mathbf{x} - \mathbf{x}_0, \mathbf{n}_i \rangle = 0, \quad i = 1 \ldots n - m,$$

missä vektorit $\mathbf{n}_i$ ovat keskenään lineaarisesti riippumattomia T:n normaalivektoreita. Hypertasoihin ja m-tasoihin liittyviä geometrisia ongelmia voidaan tyypillisesti ratkaista matriisialgebran avulla.

Esimerkki 4 Millä ehdolla hypertasot

$$T_i: \quad \langle \mathbf{x} - \mathbf{a}_i, \mathbf{n}_i \rangle = 0, \quad i = 1, \ldots, n$$

leikkaavat täsmälleen yhdessä pisteessä?

Ratkaisu Matriisialgebran mukaan tasojen T_i yhtälöt muodostavat lineaarisen yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$, missä $\mathbf{A}$:n i:s rivi $= \mathbf{n}_i^T$ ja $[\mathbf{b}]_i = \langle \mathbf{a}_i, \mathbf{n}_i \rangle$. Siis vastaus on: Täsmälleen sillä ehdolla, että $\mathbf{A}$ on säännöllinen eli että normaalivektorit $\mathbf{n}_i$ muodostavat lineaarisesti riippumattoman systeemin ($\mathbb{R}^n$:n kannan). $\square$

*Lineaarinen optimointi $\mathbb{R}^n$:ssä

Reaaliarvoisen $\mathbb{R}^n$:n lineaarikuvauksen $L: \mathbb{R}^n \to \mathbb{R}$ yleinen muoto on

$$L\mathbf{x} = \mathbf{a}^T\mathbf{x},$$

missä $\mathbf{a} \in \mathbb{R}^n$ (oletetaan jatkossa: $\mathbf{a} \neq \mathbf{0}$). Tällaisen funktion maksimi- tai minimiarvon etsimistä annetussa joukossa $K \subset \mathbb{R}^n$ sanotaan *lineaarisen optimoinnin* tehtäväksi. Jos T on hypertaso, jonka yhtälö on

$$T: \quad \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle = \mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) = 0,$$

niin $L\mathbf{x} = \mathbf{a}^T\mathbf{x}_0 =$ vakio, kun $\mathbf{x} \in T$, eli tällaiset hypertasot ovat L:n *tasa-arvopintoja*. Toisaalta jos tarkastellaan L:n arvoja pisteen $\mathbf{c}$ kautta kulkevalla, tasa-arvopintoja vastaan kohtisuoralla suoralla

$$S: \quad \mathbf{x} = \mathbf{c} + t\mathbf{a}, \quad t \in \mathbb{R},$$

niin nähdään, että tällä suoralla on

$$L\mathbf{x} = L(\mathbf{c} + t\mathbf{a}) = \alpha t + \beta, \quad \alpha = |\mathbf{a}|^2 > 0, \ \beta = \mathbf{a}^T\mathbf{c}.$$

Jos nyt L:n tasa-arvopintojen yhtälössä valitaan $\mathbf{x}_0 = \mathbf{c} + t\mathbf{a} \in S$, ja merkitään tätä tasa-arvopintaa (hypertasoa) T_t:llä, niin parametria t vaihtelemalla tulee koko avaruus $\mathbb{R}^n$ 'pyyhityksi' hypertasoilla T_t. Tällöin nähdään heti, että jos L:llä on pistejoukossa K maksimi- tai minimiarvo, niin tämän on oltava vastaavasti pienin / suurin luku reaalilukujoukossa

$$A = \{t \in \mathbb{R} \mid T_t \cap K \neq \emptyset\}.$$

Optimointiongelma on näin 'ratkaistu' geometrisesti. — Idean soveltaminen käytäntöön kylläkin edellyttää, että joukko A on määrättävissä. Seuraavassa esimerkissä K on geometrisesti yksinkertainen ja n kohtuullinen, jolloin ongelma ratkeaa 'käsipelillä'.

Esimerkki 5 Määritä funktion $f(x_1, x_2, x_3, x_4) = 2x_1 - 4x_2 + 5x_3 - 6x_4$ pienin ja suurin arvo $\mathbb{R}^4$:n pallossa

$$K = \{\mathbf{x} \in \mathbb{R}^4 \mid |\mathbf{x} - \mathbf{x}_0| \leq R\}, \quad \text{missä } \mathbf{x}_0 = (1, -1, 3, 2) \text{ ja } R = 3.$$

Ratkaisu Tässä on $f(\mathbf{x}) = L\mathbf{x} = \mathbf{a}^T\mathbf{x}$, $\mathbf{a} = [2, -4, 5, -6]^T$, $\mathbf{x} = [x_1, x_2, x_3, x_4]^T$. Jos T_t on $\mathbb{R}^4$:n hypertaso, joka kulkee pisteen $P_t \mathrel{\widehat{=}} \mathbf{x}_0 + t\mathbf{a}$ kautta, niin pätee

$$T_t \cap K \neq \emptyset \quad \Leftrightarrow \quad |t\mathbf{a}| = |t||\mathbf{a}| \leq 3 \quad \Leftrightarrow \quad |t| \;\leq\; \frac{3}{\sqrt{2^2+4^2+5^2+6^2}} \;=\; \frac{1}{3}\,.$$

Siis $\{t \in \mathbb{R} \mid T_t \cap K \neq \emptyset\} = [-1/3,\, 1/3]$. Koska hypertasolla T_t on

$$f(\mathbf{x}) \;=\; \mathbf{a}^T(\mathbf{x}_0 + t\mathbf{a}) \;=\; \mathbf{a}^T\mathbf{x}_0 + |\mathbf{a}|^2 t \;=\; 9 + 81t, \quad \mathbf{x} \in T_t\,,$$

niin

$$f_{min} = 9 + 81 \cdot \left(-\frac{1}{3}\right) = \underline{\underline{-18}}, \quad f_{max} = 9 + 81 \cdot \frac{1}{3} = \underline{\underline{36}}. \quad \square$$

Sovelluksissa hyvin yleinen on lineaarisen optimoinnin ongelma, jossa K on määritelty *lineaaristen epäyhtälöiden* avulla muodossa

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T\mathbf{x} \leq c_i\,,\; i = 1 \ldots m\},$$

missä $\mathbf{a}_i \in \mathbb{R}^n$, $\mathbf{a}_i \neq \mathbf{0}$, ja $c_i \in \mathbb{R}$. Joukkoa K voi luonnehtia geometrisesti *monitahokkaaksi*, jota hypertasot $T_i \;:\; \mathbf{a}_i^T\mathbf{x} = c_i$ rajoittavat (yleensä on $m > n$). Tällaista ongelmaa on perinteisesti kutsuttu *lineaarisen ohjelmoinnin* (engl. linear programming) ongelmaksi. Ongelman sekä suoraan (periaatteessa tarkkaan) että likimääräiseen iteratiiviseen ratkaisuun on kehitetty algoritmeja.

HARJOITUSTEHTÄVIÄ

1. Millaiseksi muuntuu a) suorakulmio, jonka kärjet ovat pisteissä $(1,1)$, $(3,1)$, $(3,2)$ ja $(1,2)$, b) yksikkympyrä $S \;:\; x^2 + y^2 = 1$ affiinikuvauksessa
$$\mathbf{f}(x,y) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 5 \\ 6 \end{bmatrix} \;?$$
Piirrä kuva! Määritä myös käyrän $S' = \mathbf{f}(S)$ yhtälö.

2. Esitä affiinikuvuaksena: $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ tai $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$:
   a) Kohtisuora projektio tasossa suoralle $x + 2y + 1 = 0$
   b) Projektio suoralle $x - 3y - 2 = 0$ suoran $y = 2x$ suunnassa
   c) Kierto tasossa pisteen $(2,3)$ ympäri $60°$ vastapäivään
   d) Projektio tasolle $T : x + 2z - 3y + 3 = 0$ vektorin $\vec{i} + \vec{j} - \vec{k}$ suunnassa
   e) Peilaus tason $T : 3x - 2y + z - 5 = 0$ suhteen
   f) Peilaus avaruussuoran $S : \vec{r} = (1+t)\vec{i} + (t-2)\vec{j} + (2t-3)\vec{k}$ suhteen

3. Taso T kulkee origon kautta ja sen suuntavektoreita ovat $\vec{a}_1 = \vec{i} + \vec{j} + \vec{k}$ ja $\vec{a}_2 = \vec{j} - 2\vec{k}$. Suora S kulkee pisteen $(0, 1, 2)$ kautta ja sen suuntavektori on $\vec{a}_3 = -\vec{i} + \vec{j} + \vec{k}$.
   a) Olkoon $\mathbf{y} = [y_1, y_2, y_3]^T$ vektorin $x_i\vec{i} + x_2\vec{j} + x_3\vec{k}$ koordinaatit kannassa $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$. Määritä lineaarikuvaus $\mathbf{x} \mapsto \mathbf{y}$, ts. matriisi $\mathbf{C}$ siten, että $\mathbf{y} = \mathbf{C}\mathbf{x}$.
   b) Käyttäen a-kohdan tulosta ja matriisialgebraa määritä affiinikuvaus $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, joka suorittaa pisteen $P = (x_1, x_2, x_3)$ suuntaisprojisoinnin suoralle S tason T suunnassa ($\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$).
   c) Määrittele affiinikuvaus tyyppiä $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$, joka suorittaa pisteen $P = (x_1, x_2, x_3)$ projisoinnin suoran S sunnassa tason T pisteeksi Q, joka on ilmoitettu koordinaatteina kannassa $\{\mathbf{a}_1, \mathbf{a}_2\}$.

4. Vektori $\vec{b}$ saadaan vektorista $\vec{a} = 2\vec{i} + \vec{j} + 2\vec{k}$ kiertämällä $\vec{a}$:ta ensin x-akselin ympäri kulma $30°$ ja näin saatua vektoria edelleen y-akselin ympäri $60°$ (kierrot positiivisten koordinaattiakselien suunnasta katsoen vastapäivään). Määritä kummankin kierto-operaation (lineaarikuvauksia) matriisit, näiden avulla koko operaation matriisi $\mathbf{A}$ ja $\mathbf{A}$:n avulla vektori $\vec{b}$. Mitä vektoria mainituilla tavoilla kierrettäessä saadaan lopputulokseksi $\vec{b} = \vec{k}$?

5. Osoita lineaarikuvaus

$$\mathbf{f}(x, y) = \frac{1}{25}\begin{bmatrix} 7 & -24 \\ -24 & -7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

   peilaukseksi erään suoran suhteen ja määritä suoran yhtälö.

6. Kolmion A kärjet ovat pisteissä $(0, 0)$, $(1, 0)$ ja $(0, 1)$ ja kolmion B kärjet ovat pisteissä $(-1, 1)$, $(-2, 2)$ ja $(-3, 0)$. Määritä kaikki kolmiot C, joille pätee $C = \mathbf{f}(B)$ ja $B = \mathbf{f}(A)$, missä $\mathbf{f}$ on tason affiinikuvaus.

7. Seuraavat lineaarikuvaukset määrittelevät yhdensuuntaisprojektion origon kautta kulkevalle tasolle T suoran S sunnassa. Määritä T:n yhtälö ja S:n suuntavektori.

$$\text{a)} \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4 & 1 & -3 \\ 1 & 4 & 3 \\ -1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\text{b)} \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 7 & -6 & -9 \\ -2 & 6 & -6 \\ -1 & -2 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

8. Näytä, että jos affiinimuunnos $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$, missä $n = 2$ tai $n = 3$, on kääntyvä, niin $\mathbf{f}$

   a) kuvaa suoran suoraksi ($n = 2, 3$) ja tason tasoksi ($n = 3$)
   b) kuvaa janan janaksi ja kolmion kolmioksi
   c) kuvaa tetraedrin tetraedriksi ($n = 3$),
   d) säilyttää suorien ja tasojen yhdensuuntaisuuden,
   e) säilyttää kahden yhdensuuntaisen janan pituuksien suhteen.
   f) Miten ominaisuus c) muuttuu, jos $\mathbf{f}$ ei ole kääntyvä?

9. Missä joukon $A \subset \mathbb{R}^4$ pisteessä funktio

$$f(\mathbf{x}) = x_1 - 2x_2 - x_3 + 2x_4$$

   saavuttaa suurimman arvonsa, kun A on

   a) yksikkökkuutio: $A = \{\mathbf{x} \in \mathbb{R}^4 \mid |x_i| \leq 1,\ i = 1 \ldots 4\}$,
   b) yksikköpallo: $A = \{\mathbf{x} \in \mathbb{R}^4 \mid x_1^2 + \ldots + x_4^2 \leq 1\}$?

10. (*) Määritä affiinikuvaus $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$, joka suorittaa kierron suoran S : $x = 1+6t,\ y = 1-3t,\ z = -2t$ ympäri, suunnasta $\vec{v} = 6\vec{i} - 3\vec{j} - 2\vec{j}$ katsoen kulman φ verran vastapäivään. *Vihje*: Määritä $\mathbf{f}$ ensin koordinaatistossa, jossa S on koordinaattiakseli — ks. Harj.teht. X.6: 6a!

11. (*) Ratkaise lineaarisen optimoinnin tehtävä $f(x, y, z) = x+2y+3x =$max! ehdoilla

$$\begin{cases} x \geq 0,\ y \geq 0,\ z \geq 0, \\ 2x + y - 2z \leq 6, \\ -x + 5y + z \leq 8, \\ 4x + 2y + 7z \leq 23. \end{cases}$$

   Vihje: Riittää tutkia monitahokkaan kärkipisteet!

12. (*) (Illaksi kotiin) Liikemies näkee lentokoneesta kotitalonsa suunnassa $-2\vec{i} - 3\vec{j} - \vec{k}$. Talon perusosa on suorakulmainen särmiö, jonka korkeus $= 1$ ja jonka maan pinnalla olevat nurkat ovat xy-tason pisteissä $(0, 0)$, $(3, 0)$, $(3, 2)$ ja $(0, 2)$. Talossa on symmetrinen harjakatto, harjan korkeus $= 2$ ja suunta $= \vec{i}$. Räystäät talon sivuilla ja päädyissä ulottuvat eäisyydelle 0.2 seinistä. Määrittele affiinikuvaus $\mathbf{f} : \mathbb{R}^3 \mapsto \mathbb{R}^2$, joka muuntaa kolmiulotteisen todellisuuden liikemiehen näkemäksi projektiokuvaksi (oleta luonteva pään asento). Piirrä kuva talosta sellaisena kuin liikemies sen näkee.

X.8 *Lineaaristen yhtälöryhmien perinteisiä sovellusesimerkkejä

Tietokoneiden avulla suoritettavalle laajamittaiselle tieteellis-tekniselle laskennalle — esimerkiksi numeeriselle sään ennustamiselle — on tyypillistä, että huomattava osa tietokoneen käyttämästä laskenta-ajasta kuluu suurten lineaaristen yhtälöryhmien ratkaisuun. Yhtälöryhmiin päädytään, kun tarkastelun kohteena olevat luonnonlait (säätä ennustettaessa virtausmekaniikan perusyhtälöt) *diskretoidaan*, eli muunnetaan numeerisen laskennan kannalta soveliaaseen likimääräiseen muotoon. Laskemisen palauttaminen nimenomaan lineaarisiksi yhtälöryhmiksi on malleissa tavallista senkin vuoksi, että lähinnä vain tämän tyyppisten perustehtävien numeerisen ratkaisemisen tietokone viime kädessä 'osaa'.

Modernien tieteellis-teknisten laskentatehtävien ohella suuren lineaarisen yhtälöryhmän ongelma on keskeinen myös monissa perinteisemmissä — usein tietokonetta paljon vanhemmissa — matemaattisissa malleissa, joita esiintyy erityisesti insinööritieteissä. Seuraavassa esitetään kaksi tällaista klassista sovellusesimerkkiä, toinen sähkötekniikasta ja toinen rakenteiden mekaniikasta.

Sähköpiiri: Vastusverkko

Tarkastellaan klassista sähköpiiriä, joka on pelkkien vastusten muodostama verkko. Olkoot verkon solmupisteet P_i, $i = 1 \ldots n$, ja olkoot pisteet P_i, P_j yhdistetty vastuksella R_{ij} (jollei suoraa yhteyttä ole, asetetaan $1/R_{ij} = 0$ yhtälöryhmässä (3) jäljempänä). Oheisessa kuvassa on yleinen neljän solmun verkko.

Jos merkitään I_{ij} = virta vastuksen R_{ij} läpi, positiivisena P_i:stä poispäin (vrt. kuvio), niin Kirchhoffin lain (**G.R. Kirchhoff**, 1845) mukaan

$$\sum_{j=1}^{n} I_{ij} = I_i, \quad i = 1 \ldots n, \tag{1}$$

missä I_i = solmuun P_i ulkoapäin syötetty virta. Toisaalta Ohmin lain (**G.S. Ohm**, 1826) mukaan

$$R_{ij}I_{ij} \;=\; V_i - V_j, \quad i,j = 1\ldots n, \tag{2}$$

missä V_i = jännite solmussa P_i. Ohmin lakia (2) käyttäen voidaan virrat I_{ij} yhtälöryhmässä (1) lausua jännitteiden avulla, jolloin tuloksena on systeemi

$$\sum_{j=1}^{n} a_{ij}V_j \;=\; I_i, \quad i = 1\ldots n, \tag{3}$$

missä

$$a_{ii} \;=\; \sum_{j=1}^{n}(1/R_{ij}), \quad a_{ij} \;=\; -1/R_{ij}, \;\; i \neq j.$$

Tämä on lineaarinen yhtälöryhmä kokoa $n \times n$, missä $x_i = V_i$, $b_i = I_i$.

Esimerkki 1 Jos nelisolmuisessa verkossa (vrt. kuvio edellä) on $R_{ij} = R$, $i,j = 1\ldots 4$, $i \neq j$, niin jännitteille V_i saadaan yhtälöryhmä

$$\begin{cases} 3V_1 - V_2 - V_3 - V_4 = RI_1 \\ -V_1 + 3V_2 - V_3 - V_4 = RI_2 \\ -V_1 - V_2 + 3V_3 - V_4 = RI_3 \\ -V_1 - V_2 - V_3 + 3V_4 = RI_4 \end{cases}$$

Tämä on singulaarinen systeemi: Jos esim. $I_1 = I \neq 0$ ja $I_2 = I_3 = I_4 = 0$, niin ratkaisua ei ole. Jos myös $I_1 = 0$, niin $V_i = V$, $i = 1\ldots 4$ on ratkaisu millä tahansa $V \in \mathbb{R}$, ts. ratkaisuja on äärettömän monta. $\square$

Esimerkin singulaarisuusongelma on yhtälöryhmän (3) ominaisuus yleisemminkin. Nimittän kertoimien a_{ij} ym. lausekkeista nähdään, että

$$\sum_{j=1}^{n} a_{kj} \;=\; \sum_{i=1}^{n} a_{ik} \;=\; 0, \quad k = 1\ldots n.$$

Tästä on kaksi seuraamusta: Ensinnäkin jos yhtälöryhmän (3) yhtälöt lasketaan puolittain yhteen saadaan

$$(\sum_{i=1}^{n} a_{i1})\, V_1 + \ldots + (\sum_{i=1}^{n} a_{in})\, V_n \;=\; \sum_{i=1}^{n} I_i,$$

eli

$$0 \;=\; \sum_{i=1}^{n} I_i. \tag{4}$$

Toiseksi nähdään, että jos $\{V_i\}$ on systeemin (4) ratkaisu, niin myös $\{V_i + V\}$ on ratkaisu millä tahansa $V \in \mathbb{R}$, sillä

$$\sum_{j=1}^{n} a_{ij}(V_j + V) \;=\; \sum_{j=1}^{n} a_{ij}V_j \;+\; (\sum_{j=1}^{n} a_{ij})\, V \;=\; I_i + 0 \;=\; I_i, \quad i = 1 \ldots n.$$

On siis päätelty: (a) Jos yhtälöryhmällä (3) on ratkaisu, niin ehto (4) toteutuu, ts. tämä on <u>välttämätön</u> <u>ehto</u> ratkeavuudelle. (b) Jos yhtälöryhmä (3) on ratkeava, niin ratkaisuja on äärettömän monta. Yhtälöryhmä (3) on siis yleisesti singulaarinen.

Singulaarisuusongelma on merkki siitä, että matemaattista mallia ei ole ajateltu (fysikaaliselta kannalta) aivan loppuun asti. — Ongelmaan onkin yksinkertainen 'sähkömiehen ratkaisu': <u>Maadoitetaan</u> verkon yksi solmu P_k. Matemaattisessa mallissa tämä merkitsee solmussa P_k asetettavia ehtoja

$$V_k \;=\; V_{ref}, \quad I_k \;=\; -\sum_{i \neq k} I_i. \tag{5}$$

Ensimmäinen ehto asettaa jännitteelle V_k jonkin valinnaisen referenssiarvon V_{ref}, esim. $V_{ref} = 0$. Tämä on luvallista, koska fysikaalisesti merkitseviä ovat vain jännite-erot. Toisessa ehdossa voidaan I_k tulkita <u>maadoitusvirraksi</u>, joka määräytyy fysikaalisesti siten, että syöttövirtojen tasapainoehto (4) (myös fysikaalinen ehto!) toteutuu. Ehto $V_k = V_{ref}$ poistaa V_k:n tuntemattomien joukosta, jolloin yhtälöryhmä (3) voidaan kirjoittaa muotoon

$$\sum_{j \neq k} a_{ij}V_j \;=\; I_i - a_{ik}V_{ref} \;= J_i, \quad i = 1 \ldots n.$$

Tästä voidaan edelleen k:s yhtälö poistaa tarpeettomana. Nimittäin koska

$$\sum_{i=1}^{n} a_{ij} = 0 \quad \Rightarrow \quad a_{kj} = -\sum_{i \neq k} a_{ij}, \qquad j = 1 \ldots n,$$

niin olettamalla yhtälöryhmän (3) muut yhtälöt ($i \neq k$) sekä ehdon (5) jälkimmäinen osa toteutuviksi päätellään

$$\begin{aligned} \sum_{j=1}^{n} a_{kj}V_j \;&=\; -\sum_{j=1}^{n}(\sum_{i \neq k} a_{ij})\, V_j \\ &=\; -\sum_{i \neq k}(\sum_{j=1}^{n} a_{ij}V_j) \;=\; -\sum_{i \neq k} I_i \;=\; I_k. \end{aligned}$$

(Tässä tehty summeerausjärjestyksen vaihto perustuu yhteenlaskun vaihdantalakiin.) Tuloksen perusteella k:s yhtälö on muiden yhtälöiden ja lisäehdon (5) seuraus, siis tarpeeton.

Em. toimenpiteiden jälkeen on sähköpiirin matemaattinen malli supistunut lineaariseksi yhtälöryhmäksi kokoa $(n-1)\times(n-1)$. Tapauksessa $V_{ref}=0$ redusoitu malli saadaan yksinkertaisesti poistamalla yhtälöryhmän taulukkomuodosta taulukon k:s sarake (vastaten ehtoa $V_k=0$) ja k:s rivi (k:nnen yhtälön poisto). Malli saa tällöin muodon

$$\sum_{j\neq k} a_{ij}V_j \;=\; I_i, \quad i=1\ldots n, \quad i\neq k. \tag{6}$$

Osoittautuu (tarkemmat perustelut sivuutetaan), että yhtälöryhmä (6) on aina säännöllinen.

Esimerkki 1 (jatko) Jos asetetaan solmussa P_4 maadoitusehto $V_4=0$, niin tapauksessa $I_1=I,\ I_2=I_3=0$ on redusoidulla systeemillä

$$\begin{cases} 3V_1 - V_2 - V_3 = RI \\ -V_1 + 3V_2 - V_3 = 0 \\ -V_1 - V_2 + 3V_3 = 0 \end{cases}$$

(yksikäsitteinen) ratkaisu

$$V_1=\frac{1}{2}RI, \quad V_2=\frac{1}{4}RI, \quad V_3=\frac{1}{4}RI.$$

Ratkaisun perusteella solmujen P_1 ja P_4 välinen efektiivinen vastus (kokonaisvastus) on

$$R_{eff} \;=\; \frac{V_1-V_4}{I} \;=\; \frac{1}{2}R. \qquad \square$$

Ristikkorakenne

Tarkastellaan kimmoisten <u>sauvojen</u> muodostamaa ristikkorakennetta, jossa sauvat on liitetty toisiinsa solmupisteissä $P_i,\ i=1\ldots p$. Oletetaan, että liitokset välittävät sauvasta toiseen vain vetoa/puristusta, eivät vääntöä (esim. löysä pulttiliitos). Sauvassa, jonka päätepisteet ovat P_i, P_j, olkoon sauvaa venyttävä <u>voima</u>

$$\vec{F}_{ij} \;=\; F_{ij}\vec{t}_{ij},$$

missä $\vec{t}_{ij}$ on vektorin $\overrightarrow{P_iP_j}$ suuntainen yksikkövektori (kyseessä on puristus, jos $F_{ij}<0$). Tällöin <u>voimatasapaino</u> pisteessä P_i edellyttää, että

$$\sum_{j\in\Lambda_i}\vec{F}_{ij} + \vec{G}_i = \vec{0}, \quad i=1\ldots p, \tag{7}$$

missä $\{ P_j \mid j \in \Lambda_i \}$ on niiden solmupisteiden joukko, joihin P_i on yhdistetty sauvalla, ja $\vec{G}_i$ = ulkoinen kuorma solmussa P_i.

Kuvassa kuormitustilaa on purettu kahdeksi nk. vapaakappalekuvioksi, joissa kummasakin vallitsee voimatasapaino. — Huomattakoon, että liitoksia koskeva tasapainoehto (7) on hyvin samankaltainen kuin sähköpiirejä koskeva Kirchhoffin laki (1), joka myös on tasapainoehto (virtoja koskeva).

Kun yhtälöryhmässä (7) vektoriyhtälöt puretaan koordinaattimuotoon karteesisessa koordinaatistossa, niin skalaaristen yhtälöiden kokonaismääräksi tulee $2p$ tai $3p$ riippuen siitä, onko kyseessä taso- vai avaruusristikkorakenne. Tuntemattomien F_{ij} lukumäärä on luonnollisesti sama kuin sauvojen lukumäärä rakenteessa. Yhtälöryhmä (7) voi olla suoraan yksikäsitteisesti ratkeava, jolloin sanotaan,

että rakenne on staattisesti määräytyvä. Tavallisempi tilanne on kuitenkin, että systeemi (7) on alimääräytyvä, eli staattisesti määräämätön. Tällöin tarvitaan tasapainolakien lisäksi toinen fysiikan laki, Hooken laki (**R. Hooke**, 1678). Hooken lain mukaan jokainen sauva käyttäytyy ristikkoa kuormitettaessa kuten jousi, ts. voima F_{ij} on suoraan verrannollinen sauvan venymään.

Oletetaan, että ristikkoa kuormitettaessa solmu P_i (paikkavektori $\vec{r}_i$) siirtyy paikkaan Q_i, ja merkitään siirtymää $\vec{u}_i = \overrightarrow{P_iQ_i}$. Olettaen, että siirtymät ovat pieniä verrattuna sauvan pituuteen, voidaan kunkin sauvan venymä (negatiivinen venymä tarkoittaa puristumaa) laskea likimäärin kaavasta (vrt. kuvio)

$$|(\vec{r}_i + \vec{u}_i) - (\vec{r}_j + \vec{u}_j)| \;-\; |\vec{r}_i - \vec{r}_j| \;\approx\; (\vec{u}_j - \vec{u}_i) \cdot \vec{t}_{ij}.$$

(Approksimaatiossa oletetaan, että siirtymien $\vec{u}_i$ ja $\vec{u}_j$ sauvaa vastaan kohtisuorat komponentit eivät vaikuta sauvan pituuteen.) Hooken lain mukaan venymä on suoraan verrannollinen voimaan, eli

$$F_{ij} \;=\; K_{ij}(\vec{u}_j - \vec{u}_i) \cdot \vec{t}_{ij}, \tag{8}$$

missä K_{ij} on sauvalle ominainen kerroin (jousivakio). — Huomattakon jälleen analogia sähköpiireihin: Siirtymät $\vec{u}_i$ vastaavat jännitteitä, voimat F_{ij} virtoja, ja Hooken laki (8) vertautuu vastuksia koskevaan Ohmin lakiin. Jos sauvat ovat tasapaksuisia ja tehty homogeenisesta materiaalista, on jousivakio K_{ij} suoraan verrannollinen sauvan poikkipinta-alaan ja kääntäen verrannollinen sauvan pituuteen. Jousivakio riippuu myös materiaalin laadusta nk. kimmokertoimen (materiaalivakio) kautta.

Hooken lakia (8) käyttäen tulee yhtälöryhmästä (7) lineaarinen, tuntemattomina siirtymävektorien $\vec{u}_i$ koordinaatit. Tuntemattomien määrä on siis joko $2p$ (tasoristikko) tai $3p$ (avaruusristikko). Jotta yhtälöryhmästä saataisiin säännöllinen, on ristikko vielä 'maadoitettava', eli tuettava riittävästi. Tuentaehdot voivat olla esim. muotoa $\vec{u}_k = \vec{0}$, jolloin solmun P_k siirtyminen rakennetta kuormitettaessa on estetty. Tällaisessa solmussa ei voimatasapainoehtoja tarvitse kirjoittaa, sillä tasapainosta huolehtivat (tuntemattomat) tukivoimat, jotka vastaavat sähköpiirin maadoitusvirtoja. Tuentaehdoilla on yleisesti estettävä sellaiset siirtymätilat,

jotka eivät aiheuta rakenteessa mitään kuormituksia ($F_{ij} = 0 \ \forall \ i, j$). Erityisesti koko rakenteen liikkuminen jäykkänä kappaleena on matemaattisessa mallissa estettävä. Kun tällaiset riittävät (fysikaalisesti usein ilmeiset) lisäehdot on asetettu, tulee yhtälöryhmästä (7)–(8) säännöllinen. Siirtymät $\vec{u}_i$ voidaan tällöin ratkaista ensin, ja näiden avulla edelleen (fysikaalisesti ehkä kiinnostavammat) kuormitukset F_{ij} Hooken laeista (8).

Esimerkki 2

Oheisessa (taso)ristikossa on sauvoja kahta tyyppiä, jousivakiot K_1 ja K_2. Sauvoja on 7 kpl ja yhtälöryhmässä (7) on $3 \times 2 = 6$ yhtälöä (tuetuissa solmuissa A, B voimatasapainosta huolehtivat tukivoimat), joten rakenne on staattisesti määräämätön. Kun siirtymiä solmuissa P_i merkitään

$$\vec{u}_i \;=\; u_i\vec{i} + v_i\vec{j}, \quad i = 1, 2, 3,$$

niin lyhennysmerkinnöin

$$a = 1 + 2K_2/K_1, \quad b = 2K_2/K_1, \quad g_1 = G_1/K_2, \quad g_2 = -G_2/K_2$$

yhtälöryhmä (7)–(8) saa muodon

$$\left\{\begin{array}{rcrcrcrcrcrcl}
au_1 & - & v_1 & & & & & - & u_3 & + & v_3 & = & 0 \\
u_1 & - & av_1 & & & + & bv_2 & - & u_3 & + & v_3 & = & g_1 \\
 & & & & au_2 & + & v_2 & - & u_3 & - & v_3 & = & 0 \\
 & - & bv_1 & + & u_2 & + & av_2 & - & u_3 & - & v_3 & = & g_2 \\
u_1 & - & v_1 & + & u_2 & + & v_2 & - & 4u_3 & & & = & 0 \\
-u_1 & + & v_1 & + & u_2 & + & v_2 & & & - & 4v_3 & = & 0
\end{array}\right.$$

Tämä on säännöllinen ryhmä, josta siirtymät u_i, v_i ovat ratkaistavissa. Sauvoja venyttävät tai puristavat voimat ovat tämän jälkeen laskettavissa Hooken laeista (8). Esimerkiksi sauvaa P_1P_3 venyttää (negatiivisena puristaa) voima

$$F_{13} \;=\; K_2\,[(u_3 - u_1)\vec{i} + (v_3 - v_1)\vec{j}\,] \cdot \frac{1}{\sqrt{2}}\,(-\vec{i} + \vec{j}) = \frac{K_2}{\sqrt{2}}\,(u_1 - v_1 - u_3 + v_3). \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Vastuspiirin solmut ovat tasasivuisen kolmion ABC kärjet ja kolmion keskipiste D. Sivuilla AB, BC ja CA on vastukset $R_1 = 1$, $R_2 = 2$ ja $R_3 = 3$. Kolmion kärjet A, B ja C on yhdistetty keskipisteeseen vastuksilla $R_4 = R_5 = 2$ ja $R_6 = 3$. Olettaen virtasyöttö I solmuun A ja maadoitus solmuun C, muodosta lineaarinen yhtälöryhmä solmujen A, B, D jännitteille ja laske solmujen A ja C välinen vastus.

2. Kuution kärjet $P_i,\ i = 0 \ldots 7$, on numeroitu siten, että P_0 ja P_7 ovat vastakkaiset kärjet kuution yhdellä sivulla. Kärjet ovat solmuja virtapiirissä, jonka johdot kulkevat pitkin kuution särmiä ja jokaisella särmällä vastus $= R$. Solmujen jännitteistä (x_i) tiedetään, että $x_0 = E$ (jännitelähde) ja $x_7 = 0$. Muodosta lineaarinen yhtälöryhmä jännitteille $x_i,\ i = 1 \ldots 6$. Ratkaise ja laske solmujen P_0 ja P_7 välinen vastus.

3. (*) (Silta) Tason ristikkorakenteen solmut $P_1, \ldots, P_7$ ovat pisteissä $(0,0)$, $(1,1)$, $(2,0)$, $(3,1)$, $(4,0)$, $(5,1)$ ja $(6,0)$. Solmut on yhdistetty sauvoilla siten, että muodostuu viisi vierekkäistä tasakylkistä ja suorakulmaista kolmiota. Solmu P_1 on jäykästi tuettu ja solmun P_7 siirtyminen suuntaan $\vec{j}$ on estetty, muita tukia ei ole. Solmuihin P_3 ja P_5 vaikuttavat kuormat ovat $\vec{F}_3 = -aG\vec{j}$ ja $\vec{F}_5 = -bG\vec{j}$, missä $a, b \geq 0$ ovat dimensiottomia vakioita. Muita kuormia ei ole. Merkitse sauvoja venyttäviä (negatiivisena puristavia) voimia symboleilla $T_i = x_i G$ (11 sauvaa!) ja kirjoita näiden avulla solmujen tasapainoyhtälöt. Ratkaise systeemi (mahdollista, koska rakenne on staattisesti määräytyvä), ja vastaa ratkaisun perusteella kysymykseen: Jos jokainen sauva kestää puristusta G:n verran ja vetoa rajattomasti, niin millaiset kuormat ovat mahdollisia rakenteen romahtamatta? Piirrä kuva ab-tasoon!

Luku XI

Usean muuttujan differentiaalilaskenta

Tässä luvussa tarkastelun kohteena ovat *vektorimuuttujan*, eli useamman reaalimuuttujan funktiot tyyppiä $f : \mathbb{R}^n \to \mathbb{R}$ ja vektorimuuttujan *vektoriarvoiset* funktiot tyyppiä $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Edellisten erikoistapaukset, kahden ja kolmen reaalimuuttujan funktiot, ovat ennestään tuttuja Luvusta IV.3. Myös vektoriarvoisia funktioita on tavattu jo aiemmin, sillä sekä Luvussa IV.4 esitellyt parametriset käyrät ja parametriset pinnat että Luvuissa X.6–X.7 käsitellyt lineaari- ja affiinikuvaukset ovat tällaisten erikoistapauksia. Vektoriarvoisia funktioita ovat myös tason ja avaruuden *vektorikentät*, joissa vektori ymmärretään geometris–fysikaalisena 'nuolivektorina'.

Luvussa XI.1 yleistetään Luvuista V.1–V.2 tutut jatkuvuuden ja raja-arvon käsitteet useamman muuttujan funktioille. Tämän jälkeen luvun keskeisen sisällön muodostavat erilaiset derivaatan käsitteen laajennukset useamman muuttujan tilanteisiin ja näihin laajennuksiin liittyvä laskutekniikka (differentiaalilaskenta) ja sovellukset. Yleisempinä derivaatan käsitteinä esitellään tässä luvussa *osittaisderivaatat* ja niistä muodostettu *gradientti* (Luvut XI.2–XI.3), vektorikenttiin liittyvät *divergenssi* ja *roottori* (Luvut XI.4–XI.5) ja lopulta yleisempiin vektoriarvoisiin funktioihin eli *epälineaarisiin kuvauksiin* liittyen *Jacobin matriisi* (Luku XI.6). Sovelluksina tarkastellaan mm. geometrisia tehtäviä kuten *pinnan tangenttitason* määrittämistä (Luku XI.3), fysiikan *osittaisdifferentiaaliyhtälöitä* (Luvut XI.4–XI.5), *epälineaarisia yhtälöryhmiä* ja niiden numeerista ratkaisemista (Luku XI.6) ja *optimoinnin* ongelmia (Luku XI.8). Luvussa XI.7 esitellään kolme epälineaarisiin yhtälöryhmiin liittyvää *epälineaarisen analyysin* vahvaa lausetta. Viimeisessä osaluvussa (Luku XI.9) määritellään *Taylorin polynomi* useamman muuttujan funktioille.

XI.1 Usean muuttujan funktiot: Jatkuvuus ja raja-arvot

Kahden ja kolmen reaalimuuttujan funktioiden algebraa on tarkasteltu aiemmin Luvussa IV.3. Tässä ja seuraavissa luvuissa kohteena ovat kahden ja kolmen muuttujan funktioiden lisäksi yleisemmät n reaalimuuttujan reaaliarvoiset funktiot muotoa

$$y = f(x_1, \ldots, x_n).$$

Tässä voi käyttää myös matriisilaskun merkintää

$$y = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n$$

($\mathbf{x}$ pysty- tai vaakavektori) ja puhua *vektorimuuttujan* funktiosta. Mahdollista (ja matemaattisissa teksteissä tavallistakin) on matriisilaskun merkinnän sijasta käyttää myös vektorimuuttujalle yksinkertaista symbolia x, eli merkitä

$$y = f(x), \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Jatkossa käytetään eri merkintätapoja rinnakkain, jolloin matriisialgebran merkinnöissä $\mathbf{x}$ tarkoittaa pystyvektoria. Kahden ja kolmen muuttujan tapauksissa merkitään vektorimuuttuja vanhaan tapaan (x, y) tai (x, y, z).

Funktion jatkuvuuden ja raja-arvon käsitteet, jotka toistaiseksi on liitetty vain yhden muuttujan funktioihin (ks. Luvut V.1–V.2), ovat yleistettävissä melko suoraviivaisesti useamman muuttujan funktioita koskeviksi. Aloitetaan kahden muuttujan funktioista.

Funktio $f(x, y)$

Kahden muuttujan funktion jatkuvuuden ja raja-arvon määritelmien alustukseksi tarvitaan

MÄÄRITELMÄ XI.1.1 Jono $\{(x_n, y_n)\}$, missä $(x_n, y_n) \in \mathbb{R}^2$, $n \in \mathbb{N}$, *suppenee kohti* (lähestyy) *lukuparia* (pistettä) (x, y) täsmälleen kun $x_n \to x$ ja $y_n \to y$.

Kun lukuparien jonon suppenemiselle käytetään aiempaan tapaan merkintää '$\to$', niin on siis sovittu:

$$(x_n, y_n) \to (x, y) \quad \Leftrightarrow \quad x_n \to x \quad \& \quad y_n \to y.$$

Tämän kanssa yhtäpitävä sopimus on (ks. Harj.teht. I.6: 3d)

$$(x_n, y_n) \to (x, y) \quad \Leftrightarrow \quad (x_n - x)^2 + (y_n - y)^2 \to 0.$$

Tämän mukaan siis $(x_n, y_n) \to (x, y)$ tarkoittaa yksinkertaisesti, että pisteen (x_n, y_n) (geometrinen) etäisyys pisteestä (x, y) lähestyy (lukujonona) 0:aa, kun $n \to \infty$. Jos erityisesti $(x, y) = (0, 0)$, niin polaarimuunnoksen $x_n = r_n \cos \varphi_n$, $y_n = r_n \sin \varphi_n$ perusteella on

$$(x_n, y_n) \to (0, 0) \quad \Leftrightarrow \quad r_n \to 0 \qquad \text{(polaarikoordinaatisto)}.$$

Koska tämä ei aseta mitään rajoituksia jonolle $\{\varphi_n\}$, niin lähestyminen voi tapahtua esim. pitkin puolisuoraa, jolla $\varphi_n = \varphi =$ vakio, tai se voi olla suunnaltaan 'hyppelehtivää', esim. spiraalimaista. Joka tapauksessa mahdollisia lähestymissuuntia (vastaten puolisuoria) on äärettömän monta. — Tämä on olennainen ero verrattuna yhden muuttujan tilanteeseen, jossa pistettä voi lähestyä vain kahdesta eri suunnasta.

Kun kaksiulotteinen 'lähestyminen' $(x_n, y_n) \to (x, y)$ on näin määritelty, on Määritelmien V.1.1 (jatkuvuus) ja V.2.1 (funktion raja-arvo) yleistäminen kahden muuttujan tilanteeseen suoraviivaista:

Määritelmä XI.1.2 Funktio $f : \mathcal{D}_f \subset \mathbb{R}^2 \to \mathbb{R}$ on *jatkuva* pisteessä $(x, y) \in \mathcal{D}_f$ täsmälleen kun kaikille reaalilukuparien jonoille $\{(x_n, y_n)\}$ pätee

$$(x_n, y_n) \in \mathcal{D}_f \ \ \forall n \quad \& \quad (x_n, y_n) \to (x, y) \quad \Rightarrow \quad f(x_n, y_n) \to f(x, y).$$

Määritelmä XI.1.3 Funktiolla $f(x, y)$ on pisteessä (a, b) *raja-arvo* $A \in \mathbb{R}$, jos jokaiselle reaalilukuparien jonolle $\{(x_n, y_n)\}$ pätee

$$(a, b) \neq (x_n, y_n) \in \mathcal{D}_f \ \ \forall n \quad \& \quad (x_n, y_n) \to (a, b) \quad \Rightarrow \quad f(x_n, y_n) \to A,$$

ja oletus on voimassa jollekin jonolle $\{(x_n, y_n)\}$. Raja-arvo merkitään

$$\lim_{(x,y)\to(a,b)} f(x, y) = A.$$

Kuten yhden muuttujan tapauksessa, Määritelmä XI.1.3 estää raja-arvon määrittelyn sellaisessa (eristetyssä) pistessä (a, b), jota kohti lähestyminen $(x_n, y_n) \to (a, b)$ joukosta $\mathcal{D}_f$ käsin ei ole mahdollista lisäehdolla $(x_n, y_n) \neq (a, b)$ $\forall n$. Määritelmässä XI.1.2 tätä lisäehtoa ei ole, joten jos $(a, b) \in \mathcal{D}_f$ on $\mathcal{D}_f$:n eristetty piste, niin f on jatkuva pisteessä (a, b) (vrt. Harj.teht. V.1: 4).

Esimerkki 1 Määritellään funktiot

$$f_1(x, y) = \frac{x^2 y}{x^2 + y^2}, \quad f_2(x, y) = \frac{xy}{x^2 + y^2}, \quad (x, y) \in \mathbb{R}^2, \ (x, y) \neq (0, 0).$$

Tutki, ovatko funktiot jatkuvia origossa, kun asetetaan $f_1(0, 0) = f_2(0, 0) = 0$.

Ratkaisu Polaarimuunnoksilla

$$\begin{aligned} g_1(r,\varphi) &= f_1(r\cos\varphi, r\sin\varphi) = r\cos^2\varphi\sin\varphi, \\ g_2(r,\varphi) &= f_2(r\cos\varphi, r\sin\varphi) = \cos\varphi\sin\varphi \end{aligned}$$

päätellään: Jos $(x_n, y_n) \mathrel{\widehat{=}} (r_n, \varphi_n)$, niin

$$r_n \to 0 \quad \Rightarrow \quad |f_1(x_n, y_n)| = |g_1(r_n, \varphi_n)| \le r_n \to 0 = f(0,0),$$

joten f_1 on jatkuva pisteessä $(0,0)$. Sen sijaan f_2 on epäjatkuva origossa, sillä $r_n \to 0 \not\Rightarrow g_2(r_n, \varphi_n) \to 0$. $\square$

Raja-arvoille ja jatkuvuudelle pätevät samanlaiset yhdistelysäännöt kuin yhden muuttujan tapauksessa, sillä näiden tulosten taustalla oleva logiikka ja algebra ei olennaisesti riipu muuttujien lukumäärästä. Jatkuvuuden yhdistelytulokset ovat seuraavat:

Lause XI.1.4 Jos $f: D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^2$, ja $g: D_g \to \mathbb{R}$, $D_g \subset \mathbb{R}^2$, ovat jatkuvia pisteessä $(x,y) \in D_f \cap D_g$, niin myös λf ($\lambda \in \mathbb{R}$), $f+g$ ja fg ovat jatkuvia pisteessä (x,y). Jos lisäksi $g(x,y) \neq 0$, niin myös f/g on jatkuva pisteessä (x,y).

Lause XI.1.5 Jos $f: D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^2$, on jatkuva pisteessä $(x,y) \in D_f$ ja $g: D_g \to \mathbb{R}$, $D_g \subset \mathbb{R}$, on jatkuva pisteessä $f(x,y) \in \mathcal{D}_g$, niin yhdistetty funktio $g \circ f$ on jatkuva pisteessä (x,y).

Määritelmän XI.1.2 mukainen jatkuvuus yksittäisessä pisteessä laajenee luonnollisella tavalla jatkuvuudeksi joukossa: f on jatkuva joukossa $A \subset \mathcal{D}_f$, jos f on jatkuva A:n jokaisessa pisteessä. Kuten yhden muuttujan tapauksessa, tavanomaiset 'yhden lausekkeen funktiot' ovat jatkuvia koko määrittelyjoukossaan.

Esimerkki 1 (jatko) Jos esimerkissä jätetään $f_1(0,0)$ ja $f_2(0,0)$ erikseen määrittelemättä, niin esimerkin funktiot ovat Lauseen XI.1.4 perusteella jatkuvia koko yhteisessä määritelyjoukossaan

$$\mathcal{D}_{f_1} = \mathcal{D}_{f_2} = \{(x,y) \in \mathbb{R}^2 \mid (x,y) \neq (0,0)\}.$$

Funktiolla f_1 on origossa raja-arvo

$$\lim_{(x,y)\to(0,0)} f_1(x,y) = 0$$

(Määritelmä XI.1.3), joten f_1 saadaan origossa (ja siis koko $\mathbb{R}^2$:ssa) jatkuvaksi asettamalla $f_1(0,0) = 0$ (funktion jatkaminen!). Funktiolla f_2 ei tätä raja-arvoa ole, joten f_2:n jatkaminen origoon ei ole mahdollista. $\square$

Esimerkki 2 Rationaalifunktion $f(x,y) = p(x,y)/q(x,y)$ (p ja q polynomeja) määrittelyjoukko on $\mathcal{D}_f = \{(x,y) \in \mathbb{R}^2 \mid q(x,y) \neq 0\}$. Lauseen XI.1.4 mukaan p ja q ovat jatkuvia koko määrittelyjoukossaan ($= \mathbb{R}^2$), samoin f. $\square$

Esimerkki 3 Missä pisteissä funktio $f(x,y) = \sqrt{y - x^2}/(x - y^2)$ on a) määritelty, b) jatkuva?

<u>Ratkaisu</u> a) Määrittelyjoukko on $D_f = \{\, (x,y) \in \mathbb{R}^2 \mid y \geq x^2 \;\; \& \;\; x \neq y^2 \,\}$.
b) Kyseessä on yhdistetty ja yhdistelty funktio $f = (g \circ f_1)/f_2$, missä

$$f_1(x,y) = y - x^2,\ D_{f_1} = \mathbb{R}^2, \quad f_2(x,y) = x - y^2,\ D_{f_2} = \mathbb{R}^2,$$
$$g(t) = \sqrt{t},\ D_g = [0,\infty) \subset \mathbb{R}.$$

Lauseiden XI.1.4 ja XI.1.5 perusteella f on jatkuva koko määrittelyjoukossaan. $\square$

Kuten yhden muuttujan tapauksessa, myös kahden muuttujan funktion jatkuvuus ja raja-arvo voidaan määritellä vaihtoehtoisesti vetoamatta luku(pari)jonoihin. Vaihtoehtoisessa '(ε,δ)-määritelmässä' tarvitaan avointa väliä $(x-\delta, x+\delta)$ vastaava *ympäristö*, tarkemmin *pisteen avoin δ-ympäristö*, jota merkitään $U_\delta(x,y)$ ($\delta > 0$). Kuten yhdessä ulottuvuudessa, tällä tarkoitetaan joukkoa, joka ympäröi pistettä δ:n verran, tai δ:aan verrannollisen matkan, <u>joka</u> <u>suuntaan</u>. Tyypillisesti $U_\delta(x,y)$ ajatellaan joko neliön tai kiekon muotoiseksi, eli

$$\begin{aligned} \text{joko}: \quad & U_\delta(x,y) = (x-\delta, x+\delta) \times (y-\delta, y+\delta), \\ \text{tai}: \quad & U_\delta(x,y) = \{(x',y') \in \mathbb{R}^2 \mid (x'-x)^2 + (y'-y)^2 < \delta^2\}. \end{aligned}$$

Jatkuvuuden vaihtoehtoinen määritelmä on (vrt. Määritelmä V.1.2)

Määritelmä XI.1.6 Funktio $f:\ \mathcal{D}_f \subset \mathbb{R}^2 \to \mathbb{R}$ on jatkuva pistessä $(x,y) \in \mathcal{D}_f$ täsmälleen kun jokaisella $\varepsilon > 0$ on olemassa $\delta > 0$ siten, että

$$|f(x',y') - f(x,y)| < \varepsilon \quad \forall (x',y') \in U_\delta(x,y) \cap \mathcal{D}_f.$$

Määritelmät XI.1.2 ja XI.1.6 voidaan osoittaa yhtäpitäviksi samaan tapaan kuin yhden muuttujan tapauksessa (vrt. Lause V.1.3).

Ympäristö $U_\delta(x,y)$ on nimensä mukaisesti esimerkki *avoimesta* joukosta, jonka yleisempi määritelmä on (vrt. Määritelmä V.8.1 joukoille $A \in \mathbb{C}$)

Määritelmä XI.1.7 Joukko $A \subset \mathbb{R}^2$ on *avoin*, jos jokaisella$(x,y) \in A$ on olemassa $\delta > 0$ siten, että $U_\delta(x,y) \subset A$.

Funktio $f(x, y, z)$

Määritelmillä XI.1.2 ja XI.1.3 on ilmeiset vastineensa kolmen muuttujan funktioille $f(x, y, z)$. Esimerkiksi jatkuvuusehto pisteessä $(x, y, z) \in \mathcal{D}_f$ on

$$(x_n, y_n, z_n) \in \mathcal{D}_f \ \ \forall n \quad \& \quad (x_n, y_n, z_n) \to (x, y, z) \quad \Rightarrow \quad f(x_n, y_n, z_n) \to f(x, y, z),$$

missä (vrt. Harj.teht. I.6: 3e)

$$\begin{aligned} (x_n, y_n, z_n) \to (x, y, z) \quad &\Leftrightarrow \quad x_n \to x \ \ \& \ \ y_n \to y \ \ \& \ \ z_n \to z \\ &\Leftrightarrow \quad (x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2 \to 0. \end{aligned}$$

Erityisesti jos $(x, y, z) = (0, 0, 0)$, niin pallokoordinaattimuunnoksen $(x_n, y_n, z_n) \mathrel{\widehat{=}} (r_n, \theta_n \varphi_n)$ perusteella on

$$(x_n, y_n, z_n) \to (0, 0, 0) \quad \Leftrightarrow \quad r_n \to 0 \qquad \text{(pallokoordinaatisto).}$$

Esimerkki 4 Täsmälleen millä ehdoilla luvuille $\alpha, \beta, \gamma \in \mathbb{R}$ on

$$\lim_{(x,y,z)\to(0,0,0)} f(x, y, z) = 0, \quad \text{kun} \quad f(x, y, z) = \frac{|x|^\alpha |y|^\beta |z|^\gamma}{x^2 + y^2 + z^2} \ ?$$

<u>Ratkaisu</u> Tehdään pallokoordinaatimuunnos:

$$f(x, y, z) = g(r, \theta, \varphi) = r^{\alpha+\beta+\gamma-2} |\sin\theta|^{\alpha+\beta} |\cos\theta|^\gamma |\cos\varphi|^\alpha |\sin\varphi|^\beta.$$

Päätellään, että $r_n \to 0 \ \Rightarrow \ g(r_n, \theta_n, \varphi_n) \to 0$ on tosi täsmälleen ehdoilla

$$\alpha + \beta + \gamma > 2, \ \ \alpha \geq 0, \ \ \beta \geq 0, \ \ \gamma \geq 0. \quad \square$$

Funktio $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$

Yleiselle n muuttujan reaaliarvoiselle funktiolle f jatkuvuuseehto pisteessä $\mathbf{x} \in \mathcal{D}_f$ on

$$\mathbf{x}_k \in \mathcal{D}_f \ \ \forall k \ \ \mathbf{x}_k \to \mathbf{x} \ \ \Rightarrow \ \ f(\mathbf{x}_k) \to f(\mathbf{x}),$$

missä $\{\mathbf{x}_k\}$ on $\mathbb{R}^n$:n vektorijono ja *suppeneminen* $\mathbf{x}_k \to \mathbf{x}$ kohti *raja-arvoa* $\mathbf{x} \in \mathbb{R}^n$ tarkoittaa:

$$\mathbf{x}_k \to \mathbf{x} \quad \Leftrightarrow \quad (\mathbf{x}_k)_i \to (\mathbf{x})_i = x_i \,, \ \ i = 1 \ldots n \quad \Leftrightarrow \quad |\mathbf{x}_k - \mathbf{x}| \to 0,$$

missä $|\cdot|$ on $\mathbb{R}^n$:n euklidinen normi. Määritelmien XI.1.6 ja XI.1.7 n-ulotteisissa vastineissa ympäristö $U_\delta(\mathbf{x})$ tulkitaan joko n-ulotteiseksi kuutioksi, jonka särmän pituus $= 2\delta$, tai n-ulotteiseksi kuulaksi, jonka säde $= \delta$.

*Jatkuvuus kompaktissa joukossa

Asetetaan (vrt. Määritelmät V.9.1 ja V.9.5)

MÄÄRITELMÄ XI.1.8 Joukko $K \subset \mathbb{R}^n$ on

- *suljettu*, jos kaikille $\mathbb{R}^n$:n vektorijonoille $\{\mathbf{x}_k\}$ pätee: $\mathbf{x}_k \in K \ \forall k \ \ \& \ \ \mathbf{x}_k \to \mathbf{x} \in \mathbb{R}^n \ \Rightarrow \ \mathbf{x} \in K$,
- *rajoitettu*, jos $\exists C \in \mathbb{R}_+$ siten, että $|\mathbf{x}| \leq C \ \forall \mathbf{x} \in K$,
- *kompakti*, jos K on suljettu ja rajoitettu.

Suljetun joukon vaihtoehtoinen määritelmä on (vrt. Lause V.9.3)

LAUSE XI.1.9 Joukko $A \subset \mathbb{R}^n$ on suljettu täsmälleen kun A:n *komplementti* $\complement(A) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \notin A\}$ on avoin.

<u>Todistus</u> Harj.teht. 4.

ESIMERKKI 5 Kompakteja joukkoja $K \subset \mathbb{R}^n$ ovat esimerkiksi (suljettu) n-ulotteinen suorakulmainen särmiö

$$K = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$$

ja (suljettu) n-ulotteinen kuula

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{x}_0| \leq R\}.$$

Myös jokainen äärellinen $\mathbb{R}^n$:n osajoukko on kompakti (Harj.teht. 6). $\square$

Kompaktin joukon käsite on keskeinen seuraavassa Weierstrassin lauseen V.1.12 yleistyksessä, joka tapuksessa $n = 1$ on todistettu aiemmin Lauseena V.9.7. Yleisempää todistusta ei esitetä; todettakoon ainoastaan, että todistuksen logiikka on hyvin samanlainen kuin yhden muuttujan tapauksessa, ks. Luku V.9. Asetetaan ensin (vrt. Määritelmä V.9.5)

MÄÄRITELMÄ XI.1.10 Funktio $f : \mathcal{D}_f \subset \mathbb{R}^n \to \mathbb{R}$ on *jatkuva kompaktissa joukossa* $K \subset \mathcal{D}_f$, jos jokaiselle $\mathbb{R}^n$:n vektorijonolle $\{\mathbf{x}_k\}$ pätee

$$\mathbf{x}_k \in K \ \forall k \quad \& \quad \mathbf{x}_k \to \mathbf{x} \quad \Rightarrow \quad f(\mathbf{x}_k) \to f(\mathbf{x}).$$

*LAUSE XI.1.11 Jos funktio $f : \mathcal{D}_f \subset \mathbb{R}^n \to \mathbb{R}$ on jatkuva kompaktissa joukossa $K \subset \mathcal{D}_f$, niin f saavuttaa K:ssa pienimmän ja suurimman arvonsa.

HARJOITUSTEHTÄVIÄ

1. Kohdissa a)–h) määritä raja-arvo tai päättele, ettei raja-arvoa ole. Kohdissa i)–j) määritä kaikki kertoimien a, b, c arvot, joilla raja-arvo on olemassa. Siirtyminen napa- tai pallokoordinaatistoon voi auttaa.

a) $\displaystyle\lim_{(x,y)\to(2,-1)} (xy + y^2)$ b) $\displaystyle\lim_{(x,y)\to(0,0)} \sqrt{x^2 + 2y^2}$ c) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 + y^2}{y}$

d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x}{x^2 + y^2}$ e) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{y^3}{x^2 + y^2}$ f) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2 + y^4}$

g) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2 + y^2}$ h) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{\sin(xyz)}{x^2 + y^2 + z^2}$

i) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{ax^2 + bxy + cy^2}$ j) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{x^2 + y^2 - z^2}{ax^2 + by^2 + cz^2}$

2. Kohdissa a)–i) jatka funktion f määritelmä niin, että funktiosta tulee jatkuva mahdollisimman suuressa $\mathbb{R}^2$:n tai $\mathbb{R}^3$:n osajoukossa. Kohdissa j)–m) määritä kaikki pisteet, joissa f on epäjatkuva.

a) $f(x,y) = \dfrac{x\sqrt[4]{|y|}}{|x| + |y|}$ b) $f(x,y) = \dfrac{\sqrt{(1+x^2)(1+y^2)} - 1}{x^2 + y^2}$

c) $f(x,y) = \dfrac{x^3 - 8y^3}{x - 2y}$ d) $f(x,y) = \dfrac{x + y}{x^3 + y^3}$ e) $f(x,y) = \dfrac{x^2 + y}{x^4 - y^2}$

f) $f(x,y) = \dfrac{\sin(x+y)}{x^2 - y^2}$ g) $f(x,y) = \dfrac{\sin(x^3 + y^3)}{x + y}$

h) $f(x,y,z) = \dfrac{1 - \cos(xyz)}{x^2y^2z^2}$ i) $f(x,y,z) = \dfrac{x^2 - y^2 + z^2 + 2xz}{x + y + z}$

j) $f(x,y) = \begin{cases} 4x^2 + y^2 - 47, & \text{kun } x^2 + y^2 < 25 \\ x^2 - 2y^2 + 28, & \text{kun } x^2 + y^2 \geq 25 \end{cases}$

k) $f(x,y) = \begin{cases} x^2 + xy - 2y^2, & \text{kun } 0 < |y| < x \\ 0, & \text{muulloin} \end{cases}$

l) $f(x,y) = \begin{cases} y\sin\left(\dfrac{1}{x} + \dfrac{1}{y}\right), & \text{kun } x \neq 0 \text{ ja } y \neq 0 \\ 0, & \text{kun } x = 0 \text{ tai } y = 0 \end{cases}$

m) $f(x,y,z) = \begin{cases} \dfrac{\sin(xyz)}{\sin(xy)}, & \text{kun } xy \neq 0 \\ z, & \text{kun } xy = 0 \end{cases}$

3. (*) Määritellään $f : \mathbb{R}^2 \mapsto \mathbb{R}$ seuraavasti:

$$f(x,y) = \begin{cases} 1, & \text{kun } x^2 < y < 2x^2 \\ 0, & \text{muulloin} \end{cases}$$

Osoita, että funktiolla on sama raja-arvo origossa lähestyttäessä origoa mitä tahansa suoraa $S: \ ax + by = 0$ $(a, b \in \mathbb{R})$ pitkin, mutta siitä huolimatta ei ole olemassa raja-arvoa $\lim_{(x,y)\to(0,0)} f(x,y)$.

4. (*) Todista Lause XI.1.9.

5. (*) Olkoon $f(x,y) = p(x,y)/q(x,y)$, missä p ja q ovat määriteltyjä ja jatkuvia koko $\mathbb{R}^2$:ssa (esim. polynomeja).

   a) Näytä suoraan Lauseen XI.1.6 avulla, että $\mathcal{D}_f \subset \mathbb{R}^2$ on avoin.

   b) Päättele $\mathcal{D}_f$ avoimeksi osoittamalla ensin Määritelmään XI.1.8 nojaten, että $\{(x,y) \in \mathbb{R}^2 \mid q(x,y) = 0\}$ on suljettu joukko ja vetoamalla Lauseeseen XI.1.9.

6. (*) Olkoon $\mathbf{a}_i \in \mathbb{R}^n$, $i = 1 \ldots m$ $(m \in \mathbb{N})$, $\mathbf{a}_i \neq \mathbf{a}_j$ kun $i \neq j$, ja $K = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$.

   a) Näytä, että $K \subset \mathbb{R}^n$ on kompakti joukko.

   b) Näytä, että jos $f : \mathcal{D}_f \subset \mathbb{R}^n \to \mathbb{R}$ on mikä tahansa funktio, jolle pätee $\mathbf{a}_i \in \mathcal{D}_f$, $i = 1 \ldots m$, niin f on jatkuva K:ssa Määritelmän XI.1.10 mukaisesti.

   c) Tarkista, että Lauseen XI.1.11 väittämä on sopusoinnussa b)-kohdan tuloksen kanssa.

XI.2 Osittaisderivaatat

Kahden tai useamman muuttujan funktion *osittaisderivaatalla* (engl. partial derivative) tarkoitetaan yksinkertaisesti funktion derivaattaa jonkin muuttujan suhteen muiden muuttujien pysyessä vakioina raja-arvoprosessissa. Osittaisderivaatan symboli on ∂, joka luetaan samoin kuin tavallinen derivaatta eli 'dee' (engl. 'partial'). Esimerkiksi kahden muuttujan funktion $f(x,y)$ osittaisderivaattoja merkitään

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x,y)-f(x,y)}{\Delta x}\,,\\ \frac{\partial f}{\partial y}(x,y) &= \lim_{\Delta y\to 0}\frac{f(x,y+\Delta y)-f(x,y)}{\Delta y}\,.\end{aligned}$$

Vähemmän tilaa vieviä (ja laskentaakin nopeuttavia) merkintätapoja ovat

$$\partial_x f,\ \partial_y f, \quad \text{tai} \quad f_x,\ f_y.$$

Näistä ensimmäinen merkintätapa viittaa siihen tosiasiaan, että kuten tavallinen derivaatta, myös osittaisderivaatta on 'funktion funktio' eli operaattori. Alaindeksoitu merkintä on tavallinen etenkin silloin, kun muuttujat ovat fysikaalisia paikka- ja aikamuuttujia.

Osittaisderivaatta käyttäytyy monessa suhteessa kuten tavallinen derivaatta. Se on esimerkiksi derivoitavan funktion suhteen *lineaarinen*:

$$\partial_x\big[\lambda f(x,y)+\mu g(x,y)\big] = \lambda\partial_x f(x,y)+\mu\partial_x g(x,y), \quad \lambda,\mu\in\mathbb{R}.$$

Osittaisderivoinnille pätee myös mm. tulon derivoimissääntö

$$\frac{\partial}{\partial x}(fg) = \frac{\partial f}{\partial x}g + f\frac{\partial g}{\partial x}\,,$$

sillä tämänkin säännön kannalta kyse on tavallisesta derivoinnista yhden valitun muuttujan suhteen.

Esimerkki 1 Funktion

$$f(x,y) = \begin{cases} 0, & \text{kun } (x,y)=(0,0),\\ xy/(x^2+y^2)\,, & \text{muulloin}\end{cases}$$

osittaisderivaatat muualla kuin origossa ovat

$$f_x(x,y) = \frac{y^3-x^2y}{(x^2+y^2)^2}\,, \quad f_y(x,y) = \frac{x^3-xy^2}{(x^2+y^2)^2}\,, \quad (x,y)\neq(0,0).$$

Tässä on käytetty tavallisia (yhden muuttujan) rationaalifunktion derivoimissääntöjä. Koska $f(x,0) = f(0,y) = 0,\ x,y \in \mathbb{R}$, niin f:llä on osittaisderivaatat myös origossa:

$$f_x(0,0) = f_y(0,0) = 0. \quad \square$$

Esimerkin funktio on origossa epäjatkuva (ks. edellinen luku, Esimerkki 1). — Siis funktio voi olla epäjatkuva yksittäisessä pisteessä, vaikka osittaisderivaatat ovat olemassa kaikkialla. Tämä ero yhden muuttujan funktioihin (joille derivaatan olemassaolo pisteessä takaa jatkuvuuden ko. pisteessä) selittyy sillä, että osittaisderivaattojen raja-arvoissa pistettä lähestytään vain koordinaattiakselien suunnissa, kun usean muuttujan jatkuvuuden määritelmässä (ks. edellinen luku) mahdollisia lähestymissuuntia on äärettömän monta. Usean muuttujan funktion derivoituvuus eli nk. *differentioituvuus* onkin määriteltävä koordinaatistosta riippumattomana 'joka suuntaan derivoituvuutena', jotta se vastaa yhden muuttujan derivoituvuuden käsitettä. Määritelmä asetetaan seuraavassa luvussa.

Funktion osittaisderivaattoja voidaan (funktioina) derivoida edelleen eri muuttujien suhteen, jolloin saadaan *toisen kertaluvun* osittaisderivaattoja, näitä derivoimalla *kolmannen kertaluvun* osittaisderivaattoja, jne. Korkeamman kertaluvun osittaisderivaattoja merkitään

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial x^2} = \partial_x(\partial_x f) = (f_x)_x = f_{xx}, \\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial x \partial y} = \partial_x(\partial_y f) = (f_y)_x = f_{yx}, \\
\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) &= \frac{\partial^3 f}{\partial x^2 \partial y} = \partial_x^2(\partial_y f) = (f_y)_{xx} = f_{yxx}, \\
\frac{\partial^2}{\partial y^2}\left(\frac{\partial^2 f}{\partial x^2}\right) &= \frac{\partial^4 f}{\partial y^2 \partial x^2} = \partial_y^2(\partial_x^2 f) = (f_{xx})_{yy} = f_{xxyy}\,, \quad \text{jne.}
\end{aligned}$$

Esimerkki 2 Funktion $f(x,y) = \ln(x^2+y^2)$ osittaiderivaatat toiseen kertalukuun asti pisteissä $(x,y) \neq (0,0)$ (eli f:n määrittelyjoukossa) ovat

$$\begin{aligned}
f_x &= \frac{2x}{x^2+y^2}\,, \quad f_y = \frac{2y}{x^2+y^2} \\
f_{xx} &= \frac{2(y^2-x^2)}{(x^2+y^2)^2}\,, \quad f_{yy} = -f_{xx}, \quad f_{xy} = f_{yx} = -\frac{4xy}{(x^2+y^2)^2}\,. \quad \square
\end{aligned}$$

Esimerkin tuloksessa $f_{xy} = f_{yx}$ on kyse yleisemmästä osittaisderivoinnin *vaihtosäännöstä*

$$\boxed{f_{xy} = f_{yx} \quad \text{(vaihtosääntö).}} \tag{1}$$

Sääntö on pätevä lievin säännöllisyysehdoin, jotka yleensä voidaan olettaa. Ko. ehdot sekä säännön perustelu esitetään luvun lopussa (Lause XI.2.1). Vaihtosäännön mukaan derivointijärjestys on vapaa myös korkeamman kertaluvun osittaisderivaatoissa; esim. $f_{xxy} = f_{xyx} = f_{yxx}$.

Monen muuttujan osittaisderivaatat

Vaihtosääntö (1) pätee myös useamman kuin kahden muuttujan tapauksessa, koska kahden muuttujan suhteen derivoitaessa funktio voidaan ajatella vain ko. muuttujista riippuvaksi. Siis jos funktion $f(x_1, \ldots, x_n)$ osittaiderivaatat $\partial f/\partial x_k$ ja $\partial^2 f/(\partial x_k \partial x_l)$ ovat olemassa ja jatkuvia, kun $k, l = 1 \ldots n,\ k \neq l$, niin

$$\frac{\partial^2 f}{\partial x_k \partial x_l} = \frac{\partial^2 f}{\partial x_l \partial x_k}, \quad k, l \in \{1, \ldots, n\},\ \ k \neq l.$$

Koska derivointijärjestys näin muodoin on vapaa (lievin ehdoin, jotka yleensä oletetaan), riittää korkeammissa osittaisderivaatoissa vain tietää, kuinka monta kertaa kunkin muuttujan suhteen derivoidaan. Yleistä osittaisderivaattaa merkitään tällöin käyttäen nk. *moni-indeksiä* (engl. multi-index), eli järjestettyä indeksijoukkoa (indeksivektoria) muotoa

$$\alpha = (\alpha_1, \ldots, \alpha_n), \quad \alpha_i \in \mathbb{N} \cup \{0\}.$$

Tällöin $\partial^\alpha f$ tarkoittaa osittaisderivaattaa

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f,$$

missä $\partial_i = \partial/\partial x_i$, ja osittaisderivaatan *kertalukua* on merkitty

$$|\alpha| = \sum_{i=1}^{n} \alpha_i.$$

Jos moni-indeksissä α on $\alpha_k = 0$, ei ko. muuttujan suhteen derivoida, ts. $\partial_k^0 f = f$.

Esimerkki 3 Jos $x_1 = x$ ja $x_2 = y$, niin vaihtosääntö huomioiden

$$\begin{aligned}
\partial^{(1,1)} f(x,y) &= \partial_x \partial_y f(x,y) = f_{xy}(x,y),\\
\partial^{(0,2)} f(x,y) &= (\partial_y)^2 f(x,y) = f_{yy}(x,y),\\
\partial^{(2,1)} f(x,y) &= (\partial_x)^2 \partial_y f(x,y) = f_{xxy}(x,y). \quad \square
\end{aligned}$$

Esimerkki 4 Tulon derivoimissääntöä ja lineaarisuussääntöä soveltaen

$$\begin{aligned}\partial^{(1,1)}(fg)(x,y) &= \partial_x\partial_y(fg)(x,y)\\ &= \partial_x\left(f_y g + f g_y\right)(x,y)\\ &= \left(f_{xy}g + f_x g_y + f_y g_x + f g_{xy}\right)(x,y). \quad \square\end{aligned}$$

Esimerkki 5 **Neliömuoto**. Yleinen $\mathbb{R}^n$:n nk. *neliömuoto* (engl. quadratic form) on funktio muotoa

$$f(x_1,\ldots,x_n) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}x_ix_j, \quad a_{ij}\in\mathbb{R},$$

tai matriisialgebran merkinnöin

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}, \quad \mathbf{x}=(x_i),\ \mathbf{A}=(a_{ij}).$$

Tässä voidaan olettaa, että $a_{ij} = a_{ji}$ ($\mathbf{A}$ symmetrinen), koska f riippuu vain summista $a_{ij}+a_{ji}$ ($a_{ij}x_ix_j + a_{ji}x_jx_i = (a_{ij}+a_{ji})x_ix_j$). Kun tehdään tämä oletus, niin tietystä muuttujasta x_k riippuvat termit erottuvat muodossa

$$f(x_1,\ldots,x_n) = \frac{1}{2}a_{kk}x_k^2 + \sum_{j\neq k} a_{kj}x_kx_j + [\ldots],$$

missä $[\ldots]$ ei sisällä muuttujaa x_k. Näin ollen

$$\frac{\partial f}{\partial x_k} = \sum_{j=1}^{n} a_{kj}x_j = [\mathbf{A}\mathbf{x}]_k,\ \ k=1\ldots n \quad\Rightarrow\quad \left(\frac{\partial f}{\partial x_i}\right) = \mathbf{A}\mathbf{x}.$$

Toisen kertaluvun osittaisderivaatat ovat

$$\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right) = (a_{ij}) = \mathbf{A}.$$

Korkeamman kertaluvun osittaisderivaatat häviävät. $\square$

Osittaisdifferentiaalioperaattorit

Kun osittaisderivaatoilla lasketaan, on usein kätevää ajatella symbolit ∂_x, ∂_y, jne. derivoinnin yksittäisistä kohteista irrotetuiksi operaattoreiksi. Esimerkiksi jos ∂^α ja ∂^β ovat n muuttujan differentaalioperaattoreita, niin riittävän säännöllisille funktioille f pätee vaihtosäännön perusteella

$$\partial^\alpha\left[\partial^\beta f(x)\right] = \partial^\beta\left[\partial^\alpha f(x)\right] = \partial^{\alpha+\beta}f(x).$$

(Tässä $\alpha+\beta$ on moni-indeksien summa vektoreina.) Tämän voi ilmaista lyhyemmin pelkästään operaattoreita koskevana laskusääntönä

$$\partial^\alpha\partial^\beta = \partial^\beta\partial^\alpha = \partial^{\alpha+\beta}.$$

Tällöin $\partial^\alpha\partial^\beta$ tarkoitaa *operaattorituloa* eli kaksivaiheista (yhdistettyä) operaatiota $f \mapsto \partial^\beta f \mapsto \partial^\alpha\partial^\beta f$. Vaihtosäännön perusteella differentiaalioperaattorit ∂^α ja ∂^β siis *kommutoivat*, eli pätee vaihdantalaki $\partial^\alpha\partial^\beta = \partial^\beta\partial^\alpha$. — Huomattakoon, että operaattori ∂^α on määritelmänsä mukaisesti itsekin $(|\alpha|-1)$-kertainen operaattoritulo, jonka tekijöitä ovat yksittäiset differentiaalioperaattorit $\partial_k = \partial/\partial x_k$.

Operaattoritulon (eli yhdistetyn operaattorin) ohella toinen yleinen differentiaalioperaattorien yhdistelytapa on *lineaariyhdistely*. Tässä periaate on sama kuin funktioilla yleensä:

$$(\lambda\partial^\alpha + \mu\partial^\beta)f = \lambda\partial^\alpha f + \mu\partial^\beta f, \quad \lambda,\mu \in \mathbb{R}.$$

Esimerkki 6 Kun määritellään kahden muuttujan funktioille differentiaalioperaattori $A = \partial_x + \partial_y$, niin

$$\begin{aligned} A^2 &= (\partial_x + \partial_y)(\partial_x + \partial_y) \\ &= \partial_x(\partial_x + \partial_y) + \partial_y(\partial_x + \partial_y) \\ &= (\partial_x)^2 + \partial_x\partial_y + \partial_y\partial_x + (\partial_y)^2 \\ &= (\partial_x)^2 + 2\partial_x\partial_y + (\partial_y)^2. \end{aligned}$$

Tämän mukaisesti on $A^2 f = A(Af) = f_{xx} + 2f_{xy} + f_{yy}$. Yleisemminkin A^k, $k \in \mathbb{N}$ purkautuu binomikaavalla, esim.

$$A^3 f = f_{xxx} + 3f_{xxy} + 3f_{xyy} + f_{yyy}. \quad \square$$

Lineaarisen yhdistelyn kaavassa voivat kertoimet λ, μ olla myös muuttuvia, ts. funktioita. Muodostettaessa tällaisten *muuttuvakertoimisten* operaattorien tuloja derivointi kohdistuu myös kertoimiin, jolloin tulos muuttuu vakiokertoimisesta tapauksesta. Operaattorituloja purettaessa tarvitaan tällöin myös tulon derivoimissääntöä.

Esimerkki 7 Määritellään kahden muttujan funktioille differentiaalioperaattorit $A = x\partial_x + y\partial_y$ ja $B = y^2\partial_x - x^2\partial_y$. Tällöin

$$\begin{aligned} AB &= (x\partial_x + y\partial_y)(y^2\partial_x - x^2\partial_y) \\ &= x\partial_x(y^2\partial_x) - x\partial_x(x^2\partial_y) + y\partial_y(y^2\partial_x) - y\partial_y(x^2\partial_y) \\ &= xy^2(\partial_x)^2 - 2x^2\partial_y - x^3\partial_x\partial_y + 2y^2\partial_x + y^3\partial_y\partial_x - x^2y(\partial_y)^2 \\ &= xy^2(\partial_x)^2 - x^2y(\partial_y)^2 + (y^3 - x^3)\partial_x\partial_y + 2y^2\partial_x - 2x^2\partial_y, \end{aligned}$$

$$BA = AB - y^2\partial_x + x^2\partial_y. \quad \square$$

Esimerkissä differentiaalioperaattorit eivät kommutoi, ts. $AB \neq BA$. Muuttuvakertoimisessa tapauksessa tämä on pääsääntö. Vakiokertoimiset operaattorit sen sijaan kommutoivat aina (vaihtosäännön ehdoin).

Ketjusäännöt

Jos funktiossa $f(x_1, \ldots, x_n)$ muuttujista x_i tehdään muuttujan t funktioita, ts. $x_i = x_i(t)$, niin saadaan yhdistetty funktio $F(t) = f(x_1(t), \ldots, x_n(t))$. Tälle pätee seuraava yhdistetyn funktion derivoimissäännön yleistys, josta käytetään nimeä *ketjusääntö* (engl. chain rule).

$$\boxed{\begin{aligned} &F(t) = f\big(x_1(t), \ldots, x_n(t)\big) \\ &\Rightarrow\ F'(t) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_i}\, x_i'(t) \quad \text{(ketjusääntö).} \end{aligned}} \tag{2a}$$

Jos t:n tilalla on vektorimuuttuja $(u_1, \ldots, u_m)$, niin ketjusäännön yleisempi muoto on

$$\boxed{\begin{aligned} &F(u_1, \ldots, u_m) = f\big(x_1(u_1, \ldots, u_m), \ldots, x_n(u_1, \ldots, u_m)\big) \\ &\Rightarrow\ \frac{\partial F}{\partial u_k} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial u_k} \quad \text{(yleinen ketjusääntö).} \end{aligned}} \tag{2b}$$

Lauseessa XI.2.2 jäljempänä esitetään riittävät sännöllisyysehdot ketjusäännön (2a) pätevyydelle tapauksessa $n = 2$. Vastaavat ehdot ja todistus yleiselle säännölle (2b) ($n, m \in \mathbb{N}$) ovat tästä tapauksesta helposti yleistettävissä.

Esimerkki 8 Funktio $y(x)$ on alkuarvotehtävän

$$\begin{cases} y' = f(x, y), \\ y(0) = 1 \end{cases}$$

ratkaisu. Määritä y:n Taylorin polynomi $T_2(x, 0)$, kun tiedetään, että $f(0, 1) = 2$, $f_x(0, 1) = -2$ ja $f_y(0, 1) = 4$.

<u>Ratkaisu</u> Differentiaaliyhtälön ja alkuehdon perusteella on $y'(0) = f(0, y(0)) = f(0, 1) = 2$, jolloin differentiaaliyhtälön ja ketjusäännön (2a) perusteella on edelleen

$$y''(x) = \frac{d}{dx} f(x, y(x)) = f_x(x, y(x)) + f_y(x, y(x)) y'(x),$$

joten $y''(0) = -2 + 4 \cdot 2 = 6$. Siis kysytty Taylorin polynomi on

$$T_2(x, 0) = y(0) + y'(0)\, x + \tfrac{1}{2}\, y''(0)\, x^2 = 1 + 2x + 3x^2. \quad \square$$

Implisiittinen (osittais)derivointi

Eräs ketjusääntöjen (2a)–(2b) sovellus on yleinen implisiittisen derivoinnin tai osittaisderivoinnin periaate. Olkoon esimerkiksi F kahden muuttujan funktio ja oletetaan, että yhtälö

$$F(x, y) = 0$$

määrittelee derivoituvan funktion $y = y(x)$. Silloin on $F(x, y(x)) = 0$, joten säännön (2a) perusteella (oletetaan F:n riittävä säännöllisyys)

$$0 = \frac{d}{dx}F(x, y(x)) = F_x(x, y(x)) + F_y(x, y(x))y'(x).$$

Sikäli kuin $F_y(x, y(x)) \neq 0$, voidaan tästä ratkaista $y'(x)$:

$$y'(x) = -\frac{F_x(x, y)}{F_y(x, y)}, \quad y = y(x).$$

Esimerkki 9 Yhtälö

$$F(x, y) = xy^3 - e^{-xy^2} + y + \sin(y \cos x) = 0$$

määrittelee origon ympäristössä funktion $y(x)$. Jos halutaan laskea $y'(0)$, on ensin laskettava $y(0) = y_0$ ratkaisemalla (numeerisin keinoin)

$$y + \sin y = 1 \;\Rightarrow\; y = y_0.$$

Tämän jälkeen seuraa ketjusäännöstä (2a)

$$F_x(0, y_0) + F_y(0, y_0)y'(0) = 0,$$

missä

$$\begin{aligned} F_x &= y^3 + y^2 e^{-xy^2} - y \sin x \cos(y \cos x) \\ F_y &= 3xy^2 + 2xye^{-xy^2} + 1 + \cos x \cos(y \cos x) \\ &\Rightarrow\; y'(0) = -\frac{y_0^3 + y_0^2}{1 + \cos y_0}. \quad \square \end{aligned}$$

Kun implisiittifunktiossa on useampia muuttujia, voidaan ketjusääntöä (2b) käyttää *implisiittiseen osittaisderivointiin.* Esimerkiksi jos yhtälö

$$F(x, y, z) = 0$$

määrittelee funktion $z = z(x, y)$, niin säännön (2b) mukaan pätee

$$0 = \partial_x F(x, y, z(z, y)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x},$$
$$0 = \partial_y F(x, y, z(z, y)) = \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial y}.$$

Sikäli kuin on $F_z(x, y, z) \neq 0$, voidaan tästä ratkaista $\partial z/\partial x$ ja $\partial z/\partial y$.

Esimerkki 10 Pisteen $(x, y, z) = (1, 2, 0)$ ympäristössä määritellään

$$z = f(x, y) \;\Leftrightarrow\; 4x^2 - y^2 + z^2 + e^z = 1.$$

Implisiittisesti derivoimalla saadaan

$$8x + (2z + e^z)\frac{\partial z}{\partial x} = 0, \quad -2y + (2z + e^z)\frac{\partial z}{\partial y} = 0.$$

Tässä on $\partial z/\partial x = f_x(x, y)$ ja $\partial z/\partial y = f_y(x, y)$, joten sijoittamalla $x = 1$, $y = 2$, $z = 0$ saadaan

$$f_x(1, 2) = -8, \quad f_y(1, 2) = 4.$$

Derivoimalla edelleen implisiittisesti voidaan laskea f:n korkeamman kertaluvun osittaisderivaattoja. Esimerkiksi

$$\begin{aligned} 0 &= \partial_y \left(8x + (2z + e^z)\frac{\partial z}{\partial x}\right) = (2 + e^z)\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + (2z + e^z)\frac{\partial^2 z}{\partial x \partial y} \\ \Rightarrow\; 0 &= 3 \cdot (-8) \cdot 4 + 1 \cdot f_{xy}(1, 2) \;\Rightarrow\; f_{xy}(1, 2) = 96. \quad \square \end{aligned}$$

Derivointi integraalin alla

Halutaan johtaa derivoimissääntö funktiolle F, joka on määritelty integraalin avulla muodossa

$$F(x) = \int_a^b f(x, t)\, dt.$$

Koska integraalin lineaarisuuden nojalla pätee

$$\frac{F(x + \Delta x) - F(x)}{\Delta x} = \int_a^b \frac{f(x + \Delta x, t) - f(x, t)}{\Delta x}\, dt,$$

niin näyttäisi, että derivointi on vietävissä integraalin alle osittaisderivoinniksi muutettuna:

$$\boxed{F(x) = \int_a^b f(x, t)\, dt \;\Rightarrow\; F'(x) = \int_a^b \partial_x f(x, t)\, dt.} \tag{3a}$$

Useamman muuttujan tapauksessa tämä vastaa sääntöä

$$\boxed{\frac{\partial}{\partial x_k}\int_a^b f(x_1,\ldots,x_n,t)\,dt = \int_a^b \frac{\partial}{\partial x_k} f(x_1,\ldots,x_n,t)\,dt.} \tag{3b}$$

Säännöt (3a)–(3b) ovat todella voimassa melko yleispätevästi, joten niitä voidaan pitää (kuten vaihtosääntöä (1)) käytännössä sovellettavina pääsääntöinä, joista on vain harvinaisia poikkeuksia (ks. Harj.teht. 13a). — Epäselvissä tapauksissa on sääntöjen (3a)–(3b) perustelussa viime kädessä nojattava suoraan derivaatan määritelmään (ks. Harj.teht. 21). Siinä tapauksessa, että kyseessä on tavanomainen (Riemannin) integraali välillä $[a,b]$, rittävät takeet säännön (3a) pätevyydelle antaa Lause XI.2.3 alla.

Esimerkki 11 $F(x) = \displaystyle\int_1^2 \frac{e^{-xt}}{t}\,dt, \quad F'(0) = ?$

Ratkaisu Sääntö (3a) on pätevä (ks. Lause XI.2.3), joten

$$F'(0) = \left[\int_1^2 \partial_x\left(\frac{e^{-xt}}{t}\right)dt\right]_{x=0} = \left[\int_1^2 (-e^{-xt})\,dt\right]_{x=0} = \int_1^2 (-1)\,dt = -1. \quad \square$$

Osittaisderivoinnin sääntöjen perustelut

Seuraavassa perustellaan osittaisderivoinnin vaihtosääntö (1), ketjusääntö (2a) ja derivoinnin ja integroinnin vaihtosääntö (3a), rajoittuen viimeksi mainitussa tapauksessa tavalliseen Riemannin integraaliin välillä $[a,b]$. Jatkuvuusoletuksissa viitataan edellisen luvun määritelmiin.

Lause XI.2.1 **(Osittaisderivoinnin vaihtosääntö)** Jos kahden muuttujan funktio f on jatkuva ja osittaisderivaatat f_x ja f_y olemassa ja jatkuvia pisteen (x,y) ympäristössä $U_\delta(x,y) \subset \mathcal{D}_f$ ja lisäksi osittaisderivaatat f_{xy} ja f_{yx} ovat olemassa ko. ympäristössä ja jatkuvia pisteessä (x,y), niin $f_{xy}(x,y) = f_{yx}(x,y)$.

Todistus Tarkastellaan jonoa $\{(\Delta x_n, \Delta y_n)\}$, jolle pätee $(\Delta x_n, \Delta y_n) \to (0,0)$ ja $|\Delta x_n|^2 + |\Delta y_n|^2 < \delta^2\ \forall n$, jolloin $(x+\Delta x_n, y+\Delta y_n) \in U_\delta(x,y)\ \forall n$. Tutkitaan lauseketta

$$\delta_n[f] = f(x+\Delta x_n, y+\Delta y_n) - f(x, y+\Delta y_n) - f(x+\Delta x_n, y) + f(x,y).$$

Ryhmittelemällä tämän lausekkeen termejä eri tavoin, merkitsemällä

$$F_n(t) = f(t, y+\Delta y_n) - f(t,y), \quad G_n(t) = f(x+\Delta x_n, t) - f(x,t)$$

ja käyttämällä Differentiaalilaskun väliarvolausetta (ks. Harj.teht. 15) seuraa

$$\begin{aligned}
\delta_n[f] &= F_n(x+\Delta x_n) - F_n(x) \\
&= F_n'(\xi_n)\Delta x_n \\
&= \partial_x[f(t,y+\Delta y_n) - f(t,y)]_{t=\xi_n}\Delta x_n \\
&= [f_x(\xi_n, y+\Delta y_n) - f_x(\xi_n, y)]\Delta x_n \\
&= \big[\partial_y[f_x(\xi_n,y)]_{y=\eta_n}\Delta y_n\big]\Delta x_n \\
&= f_{xy}(\xi_n,\eta_n)\Delta x_n \Delta y_n, \\
\delta_n[f] &= G_n(y+\Delta y_n) - G_n(y) \\
&= G_n'(\eta_n')\Delta y_n \\
&= \partial_y[f(x+\Delta x_n, t) - f(x,t)]_{t=\eta_n'}\Delta y_n \\
&= [f_y(x+\Delta x_n, \eta_n') - f_y(x,\eta_n')]\Delta y_n \\
&= \big[\partial_x[f_y(x,\eta_n')]_{x=\xi_n'}\Delta x_n\big]\Delta y_n \\
&= f_{yx}(\xi_n',\eta_n')\Delta x_n \Delta y_n.
\end{aligned}$$

Tässä on $|\xi_n - x|, |\xi_n' - x| \le |\Delta x_n|$ ja $|\eta_n - y|, |\eta_n' - y| \le |\Delta y_n|$ (Harj.teht. 15), joten $(\xi_n,\eta_n) \to (x,y)$ ja $(\xi_n',\eta_n') \to (x,y)$, koska $(\Delta x_n, \Delta y_n) \to (0,0)$. Näin ollen ja koska f_{xy} ja f_{yx} ovat oletuksen mukaan jatkuvia pisteessä (x,y), niin seuraa

$$f_{xy}(\xi_n,\eta_n) \to f_{xy}(x,y) \quad \text{ja} \quad f_{yx}(\xi_n',\eta_n') \to f_{yx}(x,y), \quad \text{kun } n \to \infty.$$

Mutta em. tulosten mukaan on tässä $f_{xy}(\xi_n,\eta_n) = f_{yx}(\xi_n',\eta_n')$ $\forall n$, joten on oltava myös $f_{xy}(x,y) = f_{yx}(x,y)$. $\square$

Lause XI.2.2 **(Ketjusääntö)** Jos yhdistetyssä funktiossa $F(t) = f(x(t),y(t))$ $f(x,y)$ on jatkuva pisteen $(x(t),y(t))$ ympäristössä $U_\delta(x(t),y(t))$, osittaisderivaatat f_x ja f_y ovat olemassa ko. ympäristössä ja jatkuvia pisteessä $(x(t),y(t))$, ja derivaatat $x'(t)$ ja $y'(t)$ ovat olemassa, niin pätee derivoimissääntö

$$F'(t) = f_x(x(t),y(t))x'(t) + f_y(x(t),y(t))y'(t).$$

<u>Todistus</u> Olkoon $\varepsilon > 0$ riittävän pieni (tarkempi ehto hetimiten) ja tarkastellaan jonoa $\{\Delta t_n\}$, jolle pätee $0 < |\Delta t_n| < \varepsilon$ $\forall n$ ja $\Delta t_n \to 0$ $(n \to \infty)$. Kirjoitetaan $x(t) = x$, $y(t) = y$ ja $x(t+\Delta t_n) = x + \Delta x_n$, $y(t+\Delta t_n) = y + \Delta y_n$, $n = 1,2,\ldots$ Koska $x'(t)$ ja $y'(t)$ ovat olemassa, niin

$$\Delta x_n = x'(t)\Delta t_n + o(|\Delta t_n|), \quad \Delta y_n = y'(t)\Delta t_n + o(|\Delta t_n|).$$

Näin ollen $\varepsilon > 0$ voidaan valita niin, että toteutuu $(\Delta x_n)^2 + (\Delta y_n)^2 < \delta^2$ $\forall n$ kun $|\Delta t_n| < \varepsilon$ $\forall n$, jolloin $(x+\Delta x_n, y+\Delta y_n) \in U_\delta(x,y)$ $\forall n$. Olettaen näin pätee

Differentiaalilaskun väliarvolauseen perusteella

$$\begin{aligned} F(t+\Delta t_n) - F(t) &= f(x+\Delta x_n, y+\Delta y_n) - f(x,y) \\ &= [f(x+\Delta x_n, y+\Delta y_n) - f(x, y+\Delta y_n)] + [f(x, y+\Delta y_n) - f(x,y)] \\ &= f_x(\xi_n, y+\Delta y_n)\Delta x_n + f_y(x,\eta_n)\Delta y_n, \end{aligned}$$

missä $\xi_n \in [x, x+\Delta x_n]$ tai $\xi_n \in [x+\Delta x_n, x]$ ja vastaavasti $\eta_n \in [y, y+\Delta y_n]$ tai $\eta_n \in [y+\Delta y_n, y]$. (Jos $\Delta x_n = 0$, asetetaan $\xi_n = x$, vast. $\eta_n = y$ jos $\Delta y_n = 0$.) Jakamalla puolittain Δt_n:llä seuraa $\Delta x_n/\Delta t_n \to x'(t)$ ja $\Delta y_n/\Delta t_n \to y'(t)$, kun $n \to \infty$, ja jatkuvuusoletuksien nojalla $f_x(\xi_n, y+\Delta y_n) \to f_x(x,y)$ ja $f_y(x,\eta_n) \to f_y(x,y)$. Näin ollen raja-arvo $\lim_{\Delta t\to 0}[F(t+\Delta t) - F(t)]/\Delta t = F'(t)$ on olemassa ja väitetty derivoimissääntö on pätevä. $\square$

Seuraavan lauseen jatkuvuusoletuksissa viittataan Määritelmään XI.1.10. Todistus nojaa jatkuvuuden syvällisempään logiikkaan (tasaiseen jatkuvuuteen, vrt. Luku V.9). Esitetään todistuksesta vain helpotettu versio, joka perustuu hieman vahvennettuihin oletuksiin.

*Lause XI.2.3 (**Derivointi integraalin alla**) Jos $f(x,t)$ on jatkuva ja $\partial_x f = f_x(x,t)$ olemassa ja jatkuva suorakulmiossa $K_\delta = [c-\delta, c+\delta] \times [a,b]$ jollakin $\delta > 0$, niin derivoimissääntö (3a) on pätevä, kun $x = c$.

<u>Todistus</u> (helpotettu) Oletetaan lisäksi, että f_x toteuttaa Lipschitz-ehdon

$$|f_x(x_1,t) - f_x(x_2,t)| \le L|x_1 - x_2| \quad \forall\, (x_1,t),\, (x_2,t) \in K_\delta. \qquad (\star)$$

Jos $0 < |\Delta x| \le \delta$, niin oletusten ($f$ jatkuva, f_x olemassa) ja Differentiaalilaskun väliarvolauseen nojalla

$$\begin{aligned} f(c+\Delta x, t) - f(c,t) - f_x(c,t)\Delta x \;&=\; [f_x(\xi,t) - f_x(c,t)]\Delta x, \\ &\qquad \text{missä } \xi \in (c - |\Delta x|,\, c + |\Delta x|) \subset (c-\delta, c+\delta). \end{aligned}$$

Tämän ja Lipschitz-ehdon ($\star$) perusteella voidaan arvioida

$$\left| \frac{f(c+\Delta x, t) - f(c,t)}{\Delta x} - f_x(c,t) \right| \le L|\xi - c| \le L|\Delta x|, \quad t \in [a,b].$$

Näin ollen, ja koska $f(c+\Delta x, t)$, $f(c,t)$ ja $f_x(c,t)$ ovat (t:n suhteen) Riemann-integroituvia välillä $[a,b]$ oletusten ja Analyysin peruslauseen nojalla, päätellään

$$\begin{aligned} &\int_a^b \frac{f(c+\Delta x, t) - f(c,t)}{\Delta x}\, dt = \int_a^b f_x(c,t)\, dt + \mathcal{O}(|\Delta x|) \\ &\Rightarrow \quad F'(c) = \lim_{\Delta x \to 0} \int_a^b \frac{f(c+\Delta x, t) - f(c,t)}{\Delta x}\, dt = \int_a^b f_x(c,t)\, dt. \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske osittaisderivaatat kaikkien muuttujien suhteen:

   a) $xy^3 + x\sin(xy)$ b) $\ln\sin(x-2y)$ c) $\operatorname{Arcsin}\dfrac{y}{x}$

   d) x^y e) $\log_y x$ f) z^{xy} g) z^{x^y} h) $e^{xy(s^2+st)}$

2. Missä pisteissä seuraavilla funktioilla on molemmat osittaisderivaatat?

   a) $\sqrt{x^2+y^2}$ b) $(x+y)|x+y|$ c) $\sqrt{|x^2-y^2|}$

3. Laske

   a) $(3\partial_x - 2\partial_y)^2(x^2+xy), \quad (x,y)=(0,0)$
   b) $(\partial_x+\partial_y)(\partial_y+\partial_z)\,xy^2z^3, \quad (x,y,z)=(1,1,1)$
   c) $\partial_x xy\partial_y xy^2, \quad (x,y)=(2,1)$
   d) $(x\partial_y - y\partial_x)^2(xy-y^2), \quad (x,y)=(1,1)$
   e) $(\partial_1+\ldots+\partial_n)(x_1^2+\ldots+x_n^2), \quad (x_1,\ldots,x_n)=(1,2,\ldots,n)$
   f) $\partial_1\cdots\partial_n \ln(x_1^2+\ldots+x_n^2), \quad (x_1,\ldots,x_n)=(1,1,\ldots,1)$

4. Laske funktion f osittaisderivaatat annetussa pisteessä P. Jos on annettu m, laske myös korkeamman kertaluvun osittaisderivaatat kertalukuun m asti. Kohdissa j)–m) käytä implisiittistä osittaisderivointia.

   a) $f(x,y)=x-y+2,\ \ P=(3,2),\ \ m=2$
   b) $f(x,y)=xy+x^2,\ \ P=(2,0),\ \ m=2$
   c) $f(x,y)=\operatorname{Arctan}(y/x),\ \ P=(-1,1)$
   d) $f(x,y)=\sin(x\sqrt{y}),\ \ P=(\pi/3,4)$
   e) $f(x,y)=1/\sqrt{x^2+y^2},\ \ P=(-3,4)$
   f) $f(x,y,z)=xyz,\ \ P=(1,-1,2),\ \ m=3$
   g) $f(x,y,z)=\ln(1+e^{xyz}),\ \ P=(2,0,-1)$
   h) $f(x,y,z)=x^{y\ln z},\ \ P=(e,2,e)$
   i) $f(x_1,x_2,x_3,x_4)=(x_1-x_2^2)/(x_3+x_4^2),\ \ P=(3,1,-1,-2)$
   j) $z=f(x,y):\ x^2-2y^2-3yz^3-z=0,\ \ P:\ x=2,\ y=z=-1$
   k) $z=f(x,y):\ \cos xyz=\sin(x-y+2z),\ \ P:\ x=y=1,\ z=\pi/2$
   l) $z=f(x,y):\ 2xz^3-xyz=1,\ \ P:\ x=y=z=1,\ m=2$
   m) $u=f(x,y,z):\ xyzu+xyu^2+xu^3=1,\ \ P:\ x=u=1,\ z=y=-1$

5. Laske funktion $f(x,y)$ osittaisderivaattojen avulla lausekkeet A^2f, B^2f ja $(AB-BA)f$, kun

   a) $A=\partial_x+\partial_y\,,\ \ B=\partial_x-\partial_y$ b) $A=2\partial_x+\partial_y\,,\ \ B=\partial_x+3\partial_y$
   c) $A=y\partial_x\,,\ \ B=x\partial_y$ d) $A=x\partial_x+y\partial_y\,,\ \ B=y\partial_x-x\partial_y$

6. Näytä, että pätee:

   a) $\big[(\partial_x)^2+(\partial_y)^2\big]\ln(x^2+y^2)=0,\quad (x,y)\neq(0,0)$
   b) $\big[(\partial_x)^2+(\partial_y)^2+(\partial_z)^2\big]\dfrac{1}{\sqrt{x^2+y^2+z^2}}=0,\quad (x,y,z)\neq(0,0,0)$

7. Laske sekä suoraan että ketjusäännön avulla:
   a) $\partial u/\partial t$, kun $u=\sqrt{x^2+y^2}\,,\ x=e^{st},\ y=1+s^2\cos t$
   b) $\partial z/\partial x$, kun $z=\mathrm{Arctan}(u/v),\ u=2x+y,\ v=2x-y$
   c) dz/dt, kun $z=txy^2,\ x=t+\ln(y+t^2),\ y=e^t$

8. Oletetaan, että tunnetaan funktion $f(x,y)$ osittaisderivaattojen lausekkeet $f_x(x,y)$, $f_y(x,y)$, $f_{xx}(x,y)$, jne. Laske näiden avulla:

   a) $\dfrac{\partial}{\partial x}f(2x,3y)$ b) $\dfrac{\partial}{\partial x}f(2y,3x)$ c) $\dfrac{\partial}{\partial y}f(yf(x,t),f(y,t))$
   d) $\dfrac{\partial^2}{\partial s\partial t}f(t\sin s,t\cos s)$ e) $\dfrac{\partial^3}{\partial t^2\partial s}f(s^2-t,s+t^2)$

9. Funktiosta $y(x)$ tiedetään, että y toteuttaa välillä $(-a,a)$ differentiaaliyhtälön $y'=f(x,y)$ ja että $y(0)=0$. Funktiosta $f(x,y)$ tiedetään, että f on origon ympäristössä säännöllinen ja $f(0,0)=1$, $f_x(0,0)=-1$, $f_y(0,0)=2$, $f_{xx}(0,0)=4$, $f_{xy}(0,0)=-2$, $f_{yy}(0,0)=-6$. Määritä funktion $y(x)$ kolmannen asteen Taylorin polynomi $T_3(x,0)$.

10. Sanotaan, että differentiaaliyhtälö $f(x,y)+g(x,y)y'=0$ on *eksakti*, jos $f(x,y)=F_x(x,y)$ ja $g(x,y)=F_y(x,y)$ jollakin F. Päättele, että eksaktin differentiaaliyhtälön yleinen ratkaisu on $F(x,y)=C$. Sovella ratkaisuideaa:
    a) $e^{x^2}(y'+2xy)=0$ b) $3x^2+6xy^2+(6x^2y+4y^3)y'=0$
    c) $e^{-y}+(1-xe^{-y})y'=0$ d) $2x(x+\cos^2 y)+(2y-x^2\sin 2y)y'=0$

11. Laske $f'(0)$:

    a) $f(x)=\displaystyle\int_0^1\frac{e^{xt}}{1+t^2}\,dt$ b) $f(x)=\displaystyle\int_1^2\ln t\,e^{-xt^2}\,dt$
    c) $f(x)=\displaystyle\int_0^\pi\sin xt\cos t\,dt$ d) $f(x)=\displaystyle\int_0^1\ln(1+xt+t^2)e^{-xt}\,dt$

12. Funktio f määritellään origon ymäristössä kaavalla

$$f(x,y) = \int_0^1 \frac{\sin xt \cos yt}{1+x+y+t}\,dt.$$

Laske f:n osittaisderivaatat toiseen kertalukuun asti origossa.

13. a) Olkoon $F(x) = \int_0^\infty x^3 e^{-x^2 t}\,dt,\ x \in \mathbb{R}$. Näytä, että F on kaikkialla derivoituva, mutta $F'(0)$ ei ole laskettavissa derivoimalla integraalin alla.
b) Totea, että derivoimissäännön (3a) mukaan

$$\frac{d}{dx}\int_1^\infty \frac{e^{-xt}}{t}\,dt = -\frac{e^{-x}}{x}, \quad x > 0.$$

Varmista tulos vaihtamalla integroimismuuttujaksi $s = xt$.

14. Laske seuraavat integraalit derivoimalla annettua funktiota $F(x),\ x > 0$.

a) $I_n = \displaystyle\int_0^\infty t^n e^{-t}\,dt, \quad n \in \mathbb{N}, \quad F(x) = \int_0^\infty e^{-xt}\,dt$

b) $I_n = \displaystyle\int_0^\infty \frac{1}{(t^2+1)^n}\,dt, \quad n = 2,3,4, \quad F(x) = \int_0^\infty \frac{1}{t^2+x^2}\,dt$

c) $I_n(x) = \displaystyle\int_0^1 t^{x-1}(\ln t)^n\,dt, \quad n \in \mathbb{N}, \quad F(x) = \int_0^1 t^{x-1}\,dt$

15. Mihin (yhden muuttujan) funktioihin Lauseen XI.2.1 todistuksessa sovelletaan Differentiaalilaskun väliarvolausetta ja millä väleillä? — Tarkista väliarvolauseen soveltuvuus lauseen oletusten perusteella, ja tarkista myös päättelyn toimivuus siinä tapauksessa, että $\Delta x_n = 0$ tai $\Delta y_n = 0$.

16. (*) Tutki, mitkä Lauseen XI.2.1 oletuksista ovat voimassa, kun $(x,y) = (0,0)$ ja

$$f(x,y) = \begin{cases} 0, & \text{kun } (x,y) = (0,0) \\ \dfrac{x^3y - xy^3}{x^2+y^2}, & \text{muulloin} \end{cases}$$

Päteekö vaihtosääntö?

17. (*) Näytä, että jos funktio $f(x,y)$ ja osittaisderivaatat f_x ja f_y ovat jatkuvia avoimessa suorakulmiossa $A = (a,b) \times (c,d)$ ja $f_{xx} = f_{xy} = f_{yx} = f_{yy} = 0$ A:ssa, niin $f(x,y) = a + bx + cy,\ (x,y) \in A$ jollakin $(a,b,c) \in \mathbb{R}^3$ (eli f on ensimmäisen asteen polynomi).

18. (*) Lentokone lentää ylöspäin pitkin avaruuskäyrää

$$S: \; y = x^2, \; z = \frac{1}{3}(2x + y^2),$$

missä $z > 0$ on korkeus maan pinnasta (yksiköt km). Ilman lämpötila lentoradan lähellä on (yksikkö °C)

$$T(x, y, z) = -10(z^2 - z + 1) + (2x^2 + 3y)/(1 + z^2).$$

Ulkoilman lämpötilaa mitataan myös koneessa — olkoon mittaustulos $T(t)$ hetkellä t (min). Eräällä hetkellä kone on pisteessä $(1, 1, 1)$ ja sen vauhti on 6 km/min. Mikä on kyseisellä hetkellä mittarilukemasta $T(t)$ laskettu ulkolämpötilan hetkellinen muuttumisnopeus $T'(t)$ (yksikkö °C/min)?

19. (*) Olkoon

$$F(x, y) = \int_0^\infty \frac{e^{-xt} - e^{-yt}}{t}\, dt, \quad x > 0, \; y > 0.$$

Laskemalla $\partial F/\partial x$ ja $\partial F/\partial y$ näytä, että $F(x, y) = \ln(y/x)$.

20. (*) Laske $y(0)$ ja $y'(0)$, kun $y(x)$ määritellään pisteen $x = 0$ ympäristössä kaavalla

$$\int_0^1 \frac{e^{xyt}}{x + y + t}\, dt = \ln 2.$$

21. (*) Totea, että laskusäännön (3a) mukaan Γ-funktion $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$ derivaatta on

$$\Gamma'(x) = \int_0^\infty t^{x-1} \ln t\, e^{-t}\, dt, \quad x > 0.$$

Varmista tämä derivoimissääntö suoraan derivaatan määritelmästä näyttämällä, että on olemassa vain x:stä riippuva vakio $C(x)$ siten, että pätee

$$\left| \frac{\Gamma(x + \Delta x) - \Gamma(x)}{\Delta x} - \int_0^\infty t^{x-1} \ln t\, e^{-t}\, dt \right| \le C(x)|\Delta x|,$$

kun $x > 0$ ja $|\Delta x| \le x/2$. *Vihje*: Todista ensin aputulos:

$$|e^x - 1 - x| \le \frac{1}{2} x^2 e^{|x|}, \quad x \in \mathbb{R}.$$

XI.3 Gradientti

Palautettakoon mieliin Luvusta VII.1, että jos yhden muuttujan funktio f on derivoituva pisteessä x, niin pätee

$$f(x + \Delta x) = f(x) + df(x, \Delta x) + o(|\Delta x|),$$

missä $df(x, \Delta x) = f'(x)\Delta x$ on f:n differentiaali. Vastaava tulos kahden muuttujan funktiolle $f(x, y)$ on kehitelmä muotoa

$$f(x + \Delta x, y + \Delta y) = f(x, y) + df(x, y, \Delta x, \Delta y) + o(h), \quad h = \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

Sikäli kuin tämä on pätevä jossakin pisteen (x, y) ympäristössä, sanotaan jälleen, että $df(x, y, \Delta x, \Delta y)$ on f:n *differentiaali* pisteessä (x, y). Differentiaalin lausekkeen johtamiseksi oletetaan, kuten ketjusäännössä edellä (Lause XI.2.2), että f on jatkuva pisteen (x, y) ympäristössä $U_\delta(x, y)$ ja osittaisderivaatat f_x ja f_y ovat olemassa ko. ympäristössä ja jatkuvia pisteessä (x, y). Tällöin pätee (vrt. Lauseen XI.2.2 todistus)

$$\begin{aligned} f(x + \Delta x, &y + \Delta y) - f(x, y) \\ &= [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] + [f(x, y + \Delta y) - f(x, y)] \\ &= f_x(\xi, y + \Delta y)\Delta x + f_y(x, \eta)\Delta y \\ &= f_x(x, y)\Delta x + f_y(x, y)\Delta y + r(x, y, \Delta x, \Delta y), \end{aligned}$$

missä

$$r(x, y, \Delta x, \Delta y) = [f_x(\xi, y + \Delta y) - f_x(x, y)]\Delta x + [f_y(x, \eta) - f_y(x, y)]\Delta y.$$

Koska tässä on $|\xi - x| \le |\Delta x|$ ja $|\eta - y| \le |\Delta y|$ ja f_x ja f_y ovat jatkuvia pisteessä (x, y), niin $f_x(\xi, y + \Delta y) - f_x(x, y) = o(1)$ ja $f_y(x, \eta) - f_y(x, y) = o(1)$, kun $h \to 0$. Siis (vrt. suuruusluokka-algebran säännöt Luvussa VII.5)

$$r(x, y, \Delta x, \Delta y) = o(1)\Delta x + o(1)\Delta y = o(h), \quad h = \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

On päätelty, että pätee kehitelmä

$$\boxed{f(x + \Delta x, y + \Delta y) = f(x, y) + f_x(x, y)\Delta x + f_y(x, y)\Delta y + o(h).} \tag{1}$$

Etsitty differentiaalin lauseke (tehdyin oletuksin) on näin muodoin

$$df(x, y, \Delta x, \Delta y)) = f_x(x, y)\Delta x + f_y(x, y)\Delta y.$$

Esimerkki 1 Arvioi $f(1.01, 1.98)$ differentiaalin $df(1, 2, 0.01, -0.02)$ avulla, kun $f(x, y) = xy - 2y^3$.

Ratkaisu Tässä on $f_x(x, y) = y$ ja $f_y(x, y) = x - 6y^2$, joten

$$\begin{aligned} f(1.01, 1.98) &= f(1 + 0.01, 2 - 0.02) \\ &\approx f(1, 2) + f_x(1, 2) \cdot 0.01 + f_y(1, 2) \cdot (-0.02) \\ &= -14 + 2 \cdot 0.01 + (-23) \cdot (-0.02) = -13.52. \end{aligned}$$

Tarkka arvo on $f(1.01, 1.98) = -13.524984$. $\square$

Kun differentiaalikehitelmässä (1) otetaan käyttöön vektorimerkintä

$$\Delta\vec{r} = \Delta x\,\vec{i} + \Delta y\,\vec{j},$$

ja käytetään myös muuttujien x, y tilalla vektorimerkintää $\vec{r} = x\vec{i} + y\vec{j}$, niin kehitelmä voidaan kirjoittaa muodossa

$$f(\vec{r} + \Delta\vec{r}) = f(\vec{r}) + \nabla f(\vec{r}) \cdot \Delta\vec{r} + o(|\Delta\vec{r}|), \tag{2}$$

missä ∇f on f:n *gradientti*, joka määritellään

$$\boxed{\nabla f(x, y) = f_x(x, y)\vec{i} + f_y(x, y)\vec{j}.}$$

Myös kolmen muuttujan funktiolle on johdettavissa differentiaalikehitelmä muotoa (2) samalla tavoin kuin edellä. Tällöin on $\Delta\vec{r} = \Delta x\,\vec{i} + \Delta y\,\vec{j} + \Delta z\,\vec{k}$, ja gradientti määritellään

$$\boxed{\nabla f(x, y, z) = f_x(x, y, z)\vec{i} + f_y(x, y, z)\vec{j} + f_z(x, y, x)\vec{k}.}$$

Ym. tulosten perusteella gradienttia voi pitää 'yleistettynä derivaattana'. Yhden muuttujan derivaattaoperaattoria $\mathrm{D} = d/dx$ vastaa siis kahden ja kolmen muuttujan tapauksessa vektorimuotoinen differentiaalioperaattori

$$\nabla = \begin{cases} \vec{i}\,\partial_x + \vec{j}\,\partial_y & (d = 2), \\ \vec{i}\,\partial_x + \vec{j}\,\partial_y + \vec{k}\,\partial_z & (d = 3). \end{cases}$$

Gradientille käytetään myös symbolia 'grad'. Lukutapoja ovat 'gradientti', 'grad' ja 'nabla'.

Esimerkki 2 Funktion $f(x, y, z) = x^2 + 3y^2 + xyz$ gradientti on

$$\begin{aligned} \nabla f &= (\vec{i}\,\partial_x + \vec{j}\,\partial_y + \vec{k}\,\partial_z)(x^2 + 3y^2 + xyz) \\ &= (2x + yz)\vec{i} + (6y + xz)\vec{j} + xy\vec{k}. \quad \square \end{aligned}$$

Monen muuttujan gradientti

Edellä johdetut differentiaalikehitelmät ovat helposti yleistettävissä koskemaan yleisempää n muuttujan funktioita. Tällöin on kätevintä siirtyä matriisialgebran merkintätapoihin, eli indeksoidaan muuttujat,

$$x \hookrightarrow x_1, \quad y \hookrightarrow x_2, \quad \ldots,$$

ja kirjoitetaan $\vec{r}$:n paikalle pystyvektori $\mathbf{x}$:

$$\vec{r} \hookrightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}, \quad \Delta\vec{r} \hookrightarrow \Delta\mathbf{x}.$$

Näillä merkinnöillä differentiaalikehitelmä (2) saa muodon

$$f(\mathbf{x} + \Delta\mathbf{x}) = f(\mathbf{x}) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{x})\Delta x_k + o(|\Delta\mathbf{x}|),$$

missä $|\cdot|$ tarkoittaa $\mathbb{R}^n$:n euklidista normia. Tulos voidaan kirjoittaa matriisialgebran kielellä muotoon

$$\boxed{\begin{aligned} f(\mathbf{x} + \Delta\mathbf{x}) &= f(\mathbf{x}) + \langle \boldsymbol{\nabla} f(\mathbf{x}), \Delta\mathbf{x} \rangle + o(|\Delta\mathbf{x}|), \\ \boldsymbol{\nabla} f(\mathbf{x}) &= \left[\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right]^T. \end{aligned}} \tag{3}$$

Kehitelmässä (3) on termi $df(\mathbf{x}, \Delta\mathbf{x}) = \langle \boldsymbol{\nabla} f(\mathbf{x}), \Delta\mathbf{x} \rangle = (\boldsymbol{\nabla} f(\mathbf{x}))^T \Delta\mathbf{x}$ jälleen *differentiaali*. Muuttujien lukumäärästä riippumatta voi differentiaalia käyttää funktion arvon muutosten arviointiin samaan tapaan kuin Esimerkissä 1 edellä.

Esimerkki 3 Ristikkorakenteen (ks. Luku X.8) kuormitus aiheuttaa yksittäisen sauvan päätepisteiden P ja Q siirtymät

$$\begin{aligned} \vec{u} &= \overrightarrow{PP'} = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}, \\ \vec{v} &= \overrightarrow{QQ'} = v_1\vec{i} + v_2\vec{j} + v_3\vec{k}. \end{aligned}$$

Olettaen, että $|\vec{u}|, |\vec{v}| \ll |\vec{a}|$, $\vec{a} = \overrightarrow{QP}$, määritä sauvan venymä likimäärin differentiaalin avulla.

Ratkaisu Venymä on kuuden muuttujan funktio

$$f(u_1, u_2, u_3, v_1, v_2, v_3) = |\vec{a} + \vec{u} - \vec{v}| - |\vec{a}|.$$

Tässä on

$$\begin{aligned} |\vec{a} + \vec{u} - \vec{v}\,|^2 &= (\vec{a} + \vec{u} - \vec{v}) \cdot (\vec{a} + \vec{u} - \vec{v}) \\ &= |\vec{a}\,|^2 + 2\vec{a} \cdot (\vec{u} - \vec{v}) + |\vec{u} - \vec{v}\,|^2, \end{aligned}$$

joten

$$f(u_1, \ldots, v_3) = \left[|\vec{a}|^2 + \sum_{k=1}^{3} 2a_k(u_k - v_k) + \sum_{k=1}^{3} (u_k - v_k)^2 \right]^{1/2} - |\vec{a}|.$$

Tällöin on $f(\mathbf{0}) = 0$ ja

$$\frac{\partial f}{\partial u_k}(\mathbf{0}) = \frac{a_k}{|\vec{a}|}, \quad \frac{\partial f}{\partial v_k}(\mathbf{0}) = -\frac{a_k}{|\vec{a}|}, \quad k = 1, 2, 3,$$

joten differentiaalikehitelmän (3) mukaan

$$f(u_1, \ldots, v_3) \approx \sum_{k=1}^{3} \frac{a_k}{|\vec{a}|}\,(u_k - v_k) = \underline{\underline{\vec{t} \cdot (\vec{u} - \vec{v})}},$$

missä $\vec{t} = \vec{a}/|\vec{a}\,|$ on $\vec{a}$:n suuntainen yksikkövektori. $\square$

Differentioituvuus

Kehitelmä (3) on pätevä olettaen, että f on jatkuva $\mathbf{x}$:n ympäristössä

$$U_\delta(\mathbf{x}) = \{\mathbf{x}' \in \mathbb{R}^n \mid |\mathbf{x}' - \mathbf{x}| < \delta\}, \quad \delta > 0$$

ja lisäksi osittaisderivaatat $\partial f/\partial x_k,\ k = 1 \ldots n$ ovat olemassa ko. ympäristössä ja jatkuvia pisteessä $\mathbf{x}$. Tällaisista taustaoletuksista päästään eroon, kun asetetaan seuraava yleisempi *differentioituvuuden* ja samalla gradientin määritelmä.

Määritelmä XI.3.1 Funktio $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, on *differentioituva* pisteessä $\mathbf{x} \in D_f$, jos f on määritelty pisteen $\mathbf{x}$ jossakin ympäristössä ja on olemassa pystyvektori $\mathbf{h} \in \mathbb{R}^n$ siten, että pätee

$$f(\mathbf{x} + \Delta\mathbf{x}) = f(\mathbf{x}) + \mathbf{h}^T \Delta\mathbf{x} + o(|\Delta\mathbf{x}|), \quad \text{kun } |\Delta\mathbf{x}| \to 0.$$

Tällöin $\mathbf{h}$ on f:n *gradientti* pisteessä $\mathbf{x}$, merkitään $\mathbf{h} = \boldsymbol{\nabla} f(\mathbf{x})$.

Jos f on differentioituva, niin gradientti on laskettavissa osittaisderivaattojen avulla kaavan (3) mukaisesti:

PROPOSITIO XI.3.2 Jos f on differentioituva pisteessä $\mathbf{x} \in \mathbb{R}^n$, niin osittaisderivaatat $\partial f/\partial x_k,\ k = 1 \ldots n$ ovat olemassa pisteesä $\mathbf{x}$, ja pätee $[\boldsymbol{\nabla} f(\mathbf{x})]_k = \partial f(\mathbf{x})/\partial x_k,\ k = 1 \ldots n$.

Todistus Kun kehitelmässä (3) asetetaan $\Delta\mathbf{x} = \Delta x_k \mathbf{e}_k$ ($\mathbf{e}_k$ = euklidinen yksikkövektori), niin kehitelmä voidaan kirjoittaa muotoon

$$f(x_1, \ldots, x_{k-1}, x_k + \Delta x_k, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_n) = [\boldsymbol{\nabla} f(\mathbf{x})]_k \Delta x_k + o(|\Delta x_k|).$$

Jakamalla (Δx_k):lla ja antamalla $\Delta x_k \to 0$ seuraa $[\boldsymbol{\nabla} f(\mathbf{x})]_k = \partial f(\mathbf{x})/\partial x_k$. $\square$

Yhteenvetona edellä esitetystä voidaan todeta, että funktion jatkuvuus pisteen $\mathbf{x}$ ympäristössä, osittaisderivaattojen olemassaolo ko. ympäristössä ja jatkuvuus pisteessä $\mathbf{x}$ rittävät yhdessä takaamaan differentioituvuuden; toisaalta differentioituvuus pisteessä $\mathbf{x}$ takaa vain osittaisderivaattojen olemassaolon kyseisessä pisteessä, ei muualla. Vertaamalla edellisen luvun tarkasteluihin nähdäänkin nyt, että differentioituvuus on usean muuttujan derivoituvuuden luontevin määritelmä. Nimittäin ensinnäkin kehitelmän (3) perusteella on ilmeistä, että differentioituvuudesta seuraa välittömästi jatkuvuus ko. pisteessä, kuten yhden muuttujan tapauksessa (vrt. Luku V.3)). Toiseksi nähdään, että myös osittaisderivoinnin ketjusääntöjä perusteltaessa voidaan ulkofunktiota koskevat jatkuvuus- ym. oletukset korvata differentioituvuudella. Nimittäin jos tarkastellaan Lauseen XI.2.2 funktiota $F(t) = f(x(t), y(t))$, merkitään $x(t) = x$, $y(t) = y$, $x(t + \Delta t) = x + \Delta x$, $y(t + \Delta t) = y + \Delta y$ ja oletetaan, että f on differentioituva pisteessä (x, y), niin kehitelmän (1) perusteella

$$\frac{F(t + \Delta t) - F(t)}{\Delta t} = f_x(x, y)\frac{\Delta x}{\Delta t} + f_y(x, y)\frac{\Delta y}{\Delta t} + (\Delta t)^{-1} o(h).$$

Jos edelleen derivaatat $x'(t)$ ja $y'(t)$ ovat olemassa, niin tässä on

$$h = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{[x'(t)]^2 + [y'(t)]^2}\Delta t + o(|\Delta t|),$$

joten

$$(\Delta t)^{-1} o(h) = (\Delta t)^{-1} o(|\Delta t|) \to 0, \quad \text{kun } \Delta t \to 0.$$

Näin muodoin seuraa ketjusääntö $F'(t) = f_x(x, y)x'(t) + f_y(x, y)y'(t)$.

Yhden muuttujan derivoituvuuden ja useamman muuttujan differentioituvuuden vastaavuudesta kertoo myös seuraava Differentiaalilaskun väliarvolauseen yleistys. Todistus jätetään harjoitustehtäväksi (Harj.teht. 7c).

LAUSE XI.3.3 ($\mathbb{R}^n$**:n väliarvolause**) Olkoon $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} \neq \mathbf{b}$ ja olkoon $f(\mathbf{x})$ jatkuva pisteessä $\mathbf{x}(t) = t\mathbf{a} + (1-t)\mathbf{b}$ jokaisella $t \in [0,1]$ ja differentioituva pisteessä $\mathbf{x}(t)$ jokaisella $t \in (0,1)$. Tällöin jollakin $\xi \in (0,1)$ pätee

$$f(\mathbf{b}) - f(\mathbf{a}) = \langle \boldsymbol{\nabla} f\big(\mathbf{x}(\xi)\big), \mathbf{b} - \mathbf{a} \rangle.$$

Suunnatut derivaatat

Gradienttia käytetään usean muuttujan funktioita tutkittaessa samaan tapaan kuin derivaattaa yhden muuttujan tapauksessa. Gradientin avulla voidaan erityisesti laskea, millä tavoin funktion arvot muuttuvat lähellä annettua pistettä $\mathbf{x}$, kun ko. pisteestä siirrytään jonkin $\mathbb{R}^n$:n yksikkövektorin osoittamaan suuntaan. Nimittäin jos $\mathbf{e}$ on tällainen vektori, eli

$$\mathbf{e} = (e_i), \quad |\mathbf{e}|^2 = \sum_{i=1}^n e_i^2 = 1,$$

niin asettamalla $\Delta\mathbf{x} = \Delta t\, \mathbf{e}$ kehitelmässä (3) saadaan tulos

$$f(\mathbf{x} + \Delta t\, \mathbf{e}) = f(\mathbf{x}) + \partial_{\mathbf{e}} f(\mathbf{x}) \Delta t + o(|\Delta t|), \tag{4}$$

missä

$$\boxed{\partial_{\mathbf{e}} f(\mathbf{x}) = \sum_{k=1}^n e_k \frac{\partial f}{\partial x_k}(\mathbf{x}) = \langle \boldsymbol{\nabla} f(\mathbf{x}), \mathbf{e} \rangle.}$$

Näin määritelty $\partial_{\mathbf{e}} f$ voidaan nimetä f:n *suunnatuksi derivaataksi* suuntaan $\mathbf{e}$, sillä määritelmän (4) mukaan on

$$\partial_{\mathbf{e}} f(\mathbf{x}) = \lim_{\Delta t \to 0} \frac{f(\mathbf{x} + \Delta t\, \mathbf{e}) - f(\mathbf{x})}{\Delta t} = g_{\mathbf{e}}'(0), \quad g_{\mathbf{e}}(t) = f(\mathbf{x} + t\mathbf{e}).$$

Erityisesti jos valitaan $\mathbf{e} = \mathbf{e}_k$ (euklidinen yksikkövektori), niin $\partial_{\mathbf{e}} f = \partial f / \partial x_k$. Siis osittaisderivaatat ovat suunnattujen derivaattojen erikoistapauksia.

Kahden tai kolmen muuttujan tapauksessa voidaan kehitelmä (4) kirjoittaa haluttaessa muotoon

$$f(\vec{r} + \Delta t\, \vec{e}) = f(\vec{r}) + \vec{e} \cdot \nabla f(\vec{r}) \Delta t + o(|\Delta t|),$$

jolloin suunnattu derivaatta on differentiaalioperaattori

$$\partial_e = \vec{e} \cdot \nabla = \begin{cases} e_x \partial_x + e_y \partial y & (d = 2), \\ e_x \partial_x + e_y \partial y + e_z \partial z & (d = 3). \end{cases}$$

Jos monen muuttujan tapauksessa on $\boldsymbol{\nabla} f(\mathbf{x}) \neq \mathbf{0}$, niin suunnaksi $\mathbf{e}$ voidaan valita $\mathbf{e} = \pm[\boldsymbol{\nabla} f(\mathbf{x})]/|\boldsymbol{\nabla} f(\mathbf{x})|$, jolloin on vastaavasti $\partial_{\mathbf{e}} f(\mathbf{x}) = \pm|\boldsymbol{\nabla} f(\mathbf{x})|$. Koska toisaalta Cauchyn-Schwarzin epäyhtälön perusteella on

$$|\partial_{\mathbf{e}} f(\mathbf{x})| = \left| \sum_{k=1}^{n} e_k \frac{\partial f}{x_k} \right| \leq |\mathbf{e}||\boldsymbol{\nabla} f(\mathbf{x})| = |\boldsymbol{\nabla} f(\mathbf{x})|,$$

niin on päätelty, että suunnatun derivaatan suurin arvo $= |\boldsymbol{\nabla} f(\mathbf{x})|$ gradientin suuntaan ja pienin arvo $= -|\boldsymbol{\nabla} f(\mathbf{x})|$ vastakkaiseen suuntaan. Siis:

> Funktio kasvaa voimakkaimmin gradientin suuntaan, vähenee voimakkaimmin negatiivisen gradientin suuntaan, ja muuttuu vähiten suuntiin $\mathbf{e}$, joille $\langle \boldsymbol{\nabla} f, \mathbf{e} \rangle = 0$.

Esimerkki 4 Laske funktion $f(x, y, z) = x^3 - 2x^2yz + 4yz^2 - 2x - y + z$ derivaatta suuntaan $\vec{e} = (\vec{i} - 2\vec{j} - \vec{k})/\sqrt{6}$ pisteessä $(1, -1, 2)$ sekä määritä suunta $\vec{n}$, johon funktio kasvaa nopeimmin ko. pisteessä.

<u>Ratkaisu</u> Gradientti on

$$\begin{aligned} &\nabla f(x, y, z) = (3x^2 - 4xyz - 2)\vec{i} + (-2x^2z + 4z^2 - 1)\vec{j} + (-2x^2y + 8yz + 1)\vec{k} \\ &\Rightarrow \quad \nabla f(1, -1, 2) = 9\vec{i} + 11\vec{j} - 13\vec{k}. \end{aligned}$$

Derivaatta suuntaan $\vec{e}$ on

$$\vec{e} \cdot \nabla f(1, -1, 2) = \frac{1}{\sqrt{6}} [1 \cdot 9 - 2 \cdot 11 - 1 \cdot (-13)] = \underline{\underline{0}}.$$

Nopeimman kasvun suunta on

$$\begin{aligned} \vec{n} = \frac{\nabla f(1, -1, 2)}{|\nabla f(1, -1, 2)|} &= \frac{1}{\sqrt{371}} (9\vec{i} + 11\vec{j} - 13\vec{k}) \\ &\approx \underline{\underline{0.47\vec{i} + 0.57\vec{j} - 0.67\vec{k}}}. \quad \square \end{aligned}$$

Esimerkki 5 Kolmion kärjet ovat pisteissä $P_1 = (1, 0)$, $P_2 = (0, 2)$ ja $P_3 = (2, 3)$. Kolmion kutakin kärkeä liikutetaan sama (pieni) matka Δs siten, että P_1 siirtyy suuntaan $\vec{j}$, P_2 suuntaan $-\vec{i}$ ja P_3 suuntaan $\vec{e}$. Miten suunta $\vec{e}$ on valittava, jotta kolmion pinta-ala a) muuttuisi mahdollisimman vähän, b) kasvaisi mahdollisimman paljon, c) pienenisi mahdollisimman paljon?

Ratkaisu Jos kolmion kärjet muutoksen jälkeen ovat $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ ja $P_3 = (x_3, y_3)$, niin pinta-ala on (vrt. Luku II.6)

$$\begin{aligned} f(\mathbf{x}) = f(x_1, y_1, x_2, y_2, x_3, y_3) &= \frac{1}{2}\begin{vmatrix} x_3 - x_1 & y_3 - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix} \\ &= \frac{1}{2}(x_3 - x_1)(y_2 - y_1) - \frac{1}{2}(x_2 - x_1)(y_3 - y_1). \end{aligned}$$

Jos $\vec{e} = \cos\varphi\vec{i} + \sin\varphi\vec{j}$, niin kärkien siirtymää kuvaa vektori

$$\Delta\mathbf{x} = \Delta s\,[0, 1, -1, 0, \cos\varphi, \sin\varphi]^T.$$

Funktion f gradientti on

$$\begin{aligned} &\boldsymbol{\nabla} f(\mathbf{x}) = \frac{1}{2}[y_3 - y_2, x_2 - x_3, y_1 - y_3, x_3 - x_1, y_2 - y_1, x_1 - x_2]^T \\ \Rightarrow\quad &\boldsymbol{\nabla} f(1, 0, 0, 2, 2, 3) = \frac{1}{2}[1, -2, -3, 1, 2, 1]^T. \end{aligned}$$

Differentiaalin avulla arvioitu pinta-alan muutos on siis

$$\Delta f \approx \langle \boldsymbol{\nabla} f(1, 0, 0, 2, 2, 3), \Delta\mathbf{x}\rangle = \frac{\Delta s}{2}(1 + 2\cos\varphi + \sin\varphi).$$

Pinta-ala muuttuu vähiten, kun $1 + 2\cos\varphi + \sin\varphi = 0$, eli kun joko $\vec{e} = -\vec{j}$ ($\varphi = 3\pi/2$) tai $\vec{e} = (-4\vec{i} + 3\vec{j})/5$. Pinta-ala muuttuu eniten, kun

$$D_\varphi(1 + 2\cos\varphi + \sin\varphi) = -2\sin\varphi + \cos\varphi = 0,$$

eli kun $\vec{e} = \pm(2\vec{i} + \vec{j})/\sqrt{5}$. Näihin suuntiin on vastaavasti $\Delta f \approx \frac{\Delta s}{2}(1 \pm \sqrt{5})$.

Vastaus:

a) $\vec{e} = -\vec{j}$ tai $\vec{e} = \dfrac{1}{5}(-4\vec{i} + 3\vec{j})$, b) $\vec{e} = \dfrac{1}{\sqrt{5}}(2\vec{i} + \vec{j})$, c) $\vec{e} = -\dfrac{1}{\sqrt{5}}(2\vec{i} + \vec{j})$. □

Tasa-arvopinnat

Samaan tapaan kuin derivaatan avulla voidaan määrätä käyrän tangentti ja normaali (vrt. Luku VII.2), voidaan osittaisderivaattojen avulla määrätä *pinnan tangenttitaso* ja tämän tason normaali, eli *pinnan normaali* (suora tai vektori). Jos pinnan $S \subset E^3$ yhtälö on

$$S:\ F(x, y, z) = 0,$$

niin pinnan normaali (ja normaalin avulla tangenttitaso) annetussa pisteessä voidaan laskea nopeimmin ajattelemalla, että pinta on funktion F tasa-arvopinta. On ilmeistä, että yksikkövektori $\vec{t}$ on F:n tasa-arvopinnan tangenttivektori pisteessä (x_0, y_0, z_0) täsmälleen, kun F:n suunnattu derivaatta $\vec{t}$:n suuntaan häviää ko. pisteessä, eli kun $\vec{t} \cdot \nabla F(x_0, y_0, z_0) = 0$. Tämän perusteella on $\vec{n} = \nabla F(x_0, y_0, z_0)$ pinnan normaalivektori ko. pisteessä. Samaan tulokseen tullaan, jos tarkasteltava pinta on pisteen (x_0, y_0, z_0) kautta kulkeva F:n tasa-arvopinta, jonka yhtälö on $F(x, y, z) = c = F(x_0, y_0, z_0)$. On siis päätelty, että

> $\vec{n} = \nabla F(x_0, y_0, x_0)$ on funktion F tasa-arvopinnan normaalivektori pisteessä (x_0, y_0, z_0).

Esimerkki 6 Määritä pinnan

$$S: \ F(x, y, z) = x^2 + y^2 + z^2 - xy - yz - xz = 1.$$

tangenttitason yhtälö pisteessä $(1, 1, 2)$.

Ratkaisu Funktion F gradientti on

$$\nabla F(x, y, z) = (2x - y - z)\vec{i} + (2y - x - z)\vec{j} + (2z - y - x)\vec{k},$$

joten pinnan normaalivektori pisteessä $(1, 1, 2)$ on

$$\vec{n} = \nabla F(1, 1, 2) = -\vec{i} - \vec{j} + 2\vec{k}.$$

Tangenttitason yhtälö pisteessä $(1, 1, 2) \mathrel{\widehat{=}} \vec{r}_0$ on

$$T: \ (\vec{r} - \vec{r}_0) \cdot \vec{n} = 0 \ \Leftrightarrow \ \underline{\underline{x + y - 2z + 2 = 0}}. \quad \square$$

Esimerkki 7 Jos pinnan yhtälö on annettu muodossa $S: \ z = f(x, y)$, niin kirjoittamalla yhtälö muotoon $F(x, y, z) = z - f(x, y) = 0$ saadaan normaalivektoriksi pisteessä $(x, y, z) \in S$:

$$\vec{n} = \nabla F(x, y, z) = -f_x(x, y)\vec{i} - f_y(x, y)\vec{j} + \vec{k}. \quad \square$$

Esimerkki 8 Määritä käyrän

$$\begin{cases} x^2 - y^2 + z^2 = 1, \\ x^2 + y^2 - 2z^2 = 0 \end{cases}$$

tangenttivektori pisteessä $(1, 1, 1)$.

Ratkaisu Kyseessä on kahden pinnan leikkauskäyrä. Pinnat ovat funktioiden $F_1(x,y,z) = x^2 - y^2 + z^2$ ja $F_2(x,y,z) = x^2 + y^2 - 2z^2$ tasa-arvopintoja, joiden normaalivektorit pintojen yhteisessä pisteessä $(x,y,z) = (1,1,1)$ ovat

$$\begin{aligned}\vec{n}_1 &= \nabla F_1(x,y,z) = 2x\,\vec{i} - 2y\,\vec{j} + 2z\,\vec{k} = 2(\vec{i} - \vec{j} + \vec{k}),\\ \vec{n}_2 &= \nabla F_2(x,y,z) = 2x\,\vec{i} + 2y\,\vec{j} - 4z\,\vec{k} = 2(\vec{i} + \vec{j} - 2\vec{k}).\end{aligned}$$

Kysytty tangenttivektori $\vec{t}$ on näitä vastaan kohtisuora, eli esimerkiksi

$$\vec{t} = \frac{1}{4}\,\vec{n}_1 \times \vec{n}_2 = \underline{\underline{\vec{i} + 3\vec{j} + 2\vec{k}}}. \quad \square$$

Parametrisen pinnan normaali

Tarkastellaan $\mathbb{R}^3$:n parametrisoitua pintaa (vrt. Luku IV.4)

$$S:\ \vec{r} = \vec{r}(u,v) \ \Leftrightarrow\ \begin{cases} x = x(u,v),\\ y = y(u,v),\\ z = z(u,v).\end{cases}$$

Jos oletetaan, että tässä funktiot x, y ja z ovat differentioituvia pisteessä (u,v), niin on ilmeistä, että vektorit

$$\begin{aligned}\vec{t}_u &= \lim_{\Delta u \to 0} \frac{\vec{r}(u+\Delta u, v) - \vec{r}(u,v)}{\Delta u} = \frac{\partial \vec{r}}{\partial u}(u,v),\\ \vec{t}_v &= \lim_{\Delta v \to 0} \frac{\vec{r}(u, v+\Delta v) - \vec{r}(u,v)}{\Delta v} = \frac{\partial \vec{r}}{\partial v}(u,v)\end{aligned}$$

ovat pinnan tangenttitason suuntaisia eli *pinnan tangenttivektoreita* pisteessä $P \mathrel{\widehat{=}} \vec{r}(u,v)$. Parametrisaatiolta on sopivaa odottaa, että nämä vektorit ovat lineaarisesti riippumattomat, jolloin pinnan normaalivektori pisteessä P on

$$\boxed{\vec{n} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \quad \text{(parametrisen pinnan normaali).}}$$

Esimerkki 9 (Vrt. Esimerkki 7.) Jos pinnan yhtälö on $S:\ z = f(x,y)$, niin parametrisoidusta esityksestä

$$S:\ \vec{r}(x,y) = x\,\vec{i} + y\,\vec{j} + f(x,y)\,\vec{k}$$

saadaan S:n tangentti- ja normaalivektoreiksi pisteessä $(x,y,f(x,y))$:

$$\begin{aligned}\vec{t}_x &= \partial_x \vec{r}(x,y) = \vec{i} + f_x(x,y)\,\vec{k},\\ \vec{t}_y &= \partial_y \vec{r}(x,y) = \vec{j} + f_y(x,y)\,\vec{k},\\ \vec{n} &= \vec{t}_x \times \vec{t}_y = -f_x(x,y)\,\vec{i} - f_y(x,y)\,\vec{j} + \vec{k}. \quad \square\end{aligned}$$

Esimerkki 10 (Vrt. Esimerkki IV.4: 14.) Määritä parametrisen pinnan

$$S: \begin{cases} x = (2-2v)\cos u - v\sin u, \\ y = v\cos u + (2-2v)\sin u, \\ z = 3v \end{cases}$$

normaali pisteessä $(x, y, z) = (2, 0, 0)$.

<u>Ratkaisu</u> Koska $(2,0,0) \mathrel{\widehat{=}} \vec{r} = 2\vec{i} = \vec{r}(0,0)$, niin lasketaan

$$\begin{aligned} \vec{t}_u(0,0) &= \frac{\partial \vec{r}}{\partial u}(0,0) = \frac{\partial x}{\partial u}(0,0)\vec{i} + \frac{\partial y}{\partial u}(0,0)\vec{j} + \frac{\partial z}{\partial u}(0,0)\vec{k} = 2\vec{j}, \\ \vec{t}_v(0,0) &= \frac{\partial \vec{r}}{\partial v}(0,0) = -2\vec{i} + \vec{j} + 3\vec{k} \\ \Rightarrow\ \vec{n} &= \vec{t}_u \times \vec{t}_v = 6\vec{i} + 4\vec{k} \uparrow\uparrow \underline{\underline{3\vec{i} + 2\vec{k}}}. \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske funktion arvolle likiarvo annetussa pisteessä käyttäen differentiaalia.

   a) x^2y^3, $(3.1, 0.9)$ b) $\operatorname{Arctan}(y/x)$, $(3.01, 2.98)$
   c) $24\,(x^2+xy+y^2)^{-1}$, $(2.1, 1.9)$ d) $\sin(\pi xy + \ln y)$, $(0.01, 1.05)$
   e) $\sqrt{x+2y+3z}$, $(1.9, 1.8, 1.1)$ f) xe^{y+z^2}, $(2.05, -3.92, 1.97)$

2. Suorakulmaisen särmiön muotoisen laatikon särmät on mitattu 1%:n tarkkuudella. Arvioi, kuinka suuri vaikutus (%) mittausvirheillä on enintään mittausten perusteella laskettuun laatikon a) tilavuuteen, b) pohjan pinta-alaan, c) lävistäjään.

3. Pisteen A etäisyys tarkkailupisteestä B on pyritty selvittämään kolmiomittauksella valitsemalla toinen tarkkailupiste C ja mittaamalla kulmat ABC ja ACB. Mittaustulokset ovat: $\measuredangle ACB = 45 \pm 0.1°$, $\measuredangle ABC = 105 \pm 0.1°$ ja pisteiden B ja C välimatkaksi on mitattu 265 ± 0.5 m. Laske kysytty etäisyys d ja arvioi virhe Δd differentiaalin avulla (maksimivirhe).

4. Tornin korkeus mitataan kulmamittauksella kahdesta pisteestä A ja B, jotka ovat samassa suunnassa tornista katsoen. Mitatut kulmat ovat $50 \pm 1°$ ja $35 \pm 1°$. Lisäksi pisteiden A ja B välimatkaksi on mitattu 100 ± 1 m. Mikä on tornin korkeuden laskettu arvo ja kuinka suuri voi mittausvirhe korkeintaan olla differentiaalin perusteella arvioiden?

5. Suunnistaja, joka on kartan origossa, on matkalla rastille, joka on kartan mukaan suunnassa $\varphi = 30°$ (polaarikoordinaatti). Sateen ja hikoilun vuoksi karttapaperi on kutistunut x-suunnassa 1% ja y-suunnassa 0.6%. Arvioi, kuinka monta % suunnistajan on syytä korjata rastin suuntaa ja kartalta mitattua eäisyyttä päästäkseen oikeisiin arvoihin.

6. Nelikulmion pinta-ala on funktio $f(\mathbf{x}) = f(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$, missä (x_i, y_i), $i = 1, \ldots, 4$ ovat nelikulmion kärkipisteet. Laske ∇f pisteessä $\mathbf{x} = (0, 0, 1, 0, 1, 1, 0, 1)$ ja differentiaalin avulla likimäärin sen nelikulmion pinta-ala, jonka kärjet ovat pisteissä $(0.03, -0.05)$, $(1.02, -0.02)$, $(0.96, 1.01)$, $(-0.01, 0.98)$.

7. a) Olkoon $f(0,0) = 0$ ja $f(x,y) = xy\sin(1/\sqrt{x^2+y^2}\,)$ muulloin. Näytä, että f on origossa differentioituva mutta f_x ja f_y eivät ole origossa jatkuvia.
   b) Olkoon $f(x,y) = 0$ kun $x = 0$ tai $y = 0$ ja $f(x,y) = |xy|^\alpha (x^2+y^2)^\beta$ muulloin ($\alpha, \beta \in \mathbb{R}$). Määritä $A \subset \mathbb{R}^2$ ja $B \subset \mathbb{R}^2$ siten, että f on differentioituva origossa täsmälleen kun $(\alpha, \beta) \in A$ ja differentioituva jokaisessa pisteessä $(x,y) \in \mathbb{R}^2$ täsmälleen kun $(\alpha, \beta) \in B$.
   c) Todista Lause XI.3.3 tarkastelemalla funktiota $F(t) = f\big(t\mathbf{a} + (1-t)\mathbf{b}\big)$.

8. Laske suunnattu derivaatta annetussa pisteessä sekä annettuun että funktion nopeimman kasvun suuntaan:
   a) e^{xy+y}, $(0,0)$, $\vec{i} + \vec{j}$
   b) $\sin(\pi xy)\cos(\pi y^2)$, $(1,2)$, $\vec{i} - 2\vec{j}$
   c) xy^2z^3, $(-3,2,1)$, $6\vec{i} - 2\vec{j} + 3\vec{k}$
   d) $xy^2 + yz^3$, $(3,-2,-1)$, $\vec{2}i + \vec{2}j - 3\vec{k}$

9. Seuraavassa on annettu funktio $f(x,y)$ ja piste (x_0, y_0). Määritä kussakin tapauksessa (i) gradientti $\nabla f(x_0, y_0)$, (ii) käyrän $f(x,y) = f(x_0,y_0)$ tangentin yhtälö pisteessä (x_0, y_0), (iii) pinnan $z = f(x,y)$ tangenttitason yhtälö pisteessä (x_0, y_0, z_0), $z_0 = f(x_0, y_0)$.
   a) $f(x,y) = x^2 - y^2$, $(2,-1)$ b) $f(x,y) = x^2/(x+y)$, $(1,1)$
   c) $f(x,y) = e^{x^2y+x^2}$, $(2,0)$ d) $f(x,y) = \ln(x^2+y^2)$, $(1,-2)$

10. Määritä seuraavien pintojen tangenttitaso annetussa pisteessä P sekä ne pintojen pisteet, joissa tangenttitaso on xy-tason suuntainen.
    a) $x^2 - 2xy - y^2 + z^2 - 2x - 2y - 3z = 0$, $P = (0,0,0)$
    b) $xy + yz - 2xz + z^3 = 1$, $P = (1,1,1)$
    c) $e^{x+2y-z} = x + y^2 + 2z + 1$, $P = (0,0,0)$

11. Määritä seuraavien funktioiden tasa-arvopinnan normaali ja ko. pinnan tangenttitason yhtälö annetussa pisteessä.

    a) $f(x,y,z) = x^2y + y^2z + z^2x, \quad (1,-1,1)$
    b) $f(x,y,z) = \cos(x+2y+3z), \quad (\pi/2, \pi, \pi)$
    c) $f(x,y,z) = ye^{-x^2}, \quad (0,1,1)$
    d) $f(x,y,z) = \sin x \sin 2y \sin 3z \sin(3x-2y+z), \quad (\pi/6, \pi/6, \pi/6)$

12. Määritä avaruuskäyrän tangenttivektori annetussa pisteessä P:

$$\text{a)} \quad \begin{cases} x^2 - y^2 + 2z^2 = 13, \\ 3x + 2y - z = 3, \end{cases} \qquad P = (-2,3,-3)$$

$$\text{b)} \quad \begin{cases} x^3 - xyz^3 + yz^5 = 1, \\ xy^2 + yz + 2xz^3 = 0, \end{cases} \qquad P = (1,2,-1)$$

13. (*) Kepin päät ovat pisteissä $A = (1,1,-2)$ ja $B = (-1,-2,4)$. Keppiä liikutetaan niin, että sen päät liukuvat pitkin käyriä

$$S_1 : \begin{cases} x^3 + y^2 + z = 0, \\ x + y^2 + z^3 = -5 \end{cases} \quad \text{ja} \quad S_2 : \begin{cases} x^3 + y^2 + z = 7, \\ x + y^2 + z^3 = 67 \end{cases}$$

(kepin pituus säilyy). Olkoon kepin päät pienen siirron jälkeen pisteissä $A' \in S_1$ ja $B' \in S_2$. Laske suhteen $|AA'|/|BB'|$ raja-arvo, kun $|AA'| \to 0$.

14. (*) Suora, jonka eräs piste on A, liikkuu siten, että suora on koko ajan xy-tason suuntainen, piste A liikkuu vakionopeudella z-akselia pitkin ylöspäin, ja samalla suora pyörii vakiokulmanopeudella z-akselin ympäri vastapäivään, positiivisen z-akselin suunnasta katsoen. Suora on x akselin suuntainen täsmälleen kun A on jossakin pisteistä $(0,0,3n)$, $n \in \mathbb{Z}$. Määritä suoran liikkuessaan piirtämän (viivoitin)pinnan parametrinen esitys sekä ko. pinnan normaali ja tangenttitaso siinä pinnan ja suoran $x = 1$, $y = 2$ leikkauspisteessä, joka on lähinnä xy-tasoa.

15. (*) (Jyrkin lasku) Tasosta kohoaa tunturi. Sektorissa $A : 0 \le y \le 2x$ on tunturin korkeus tasosta (yksiköt km)

$$h(x,y) = e^{(-x^2-2xy+y^2)/100}, \quad (x,y) \in A.$$

Laskettelija Jyrki lähtee pisteestä $(3,6)$ ja laskee sukset suunnattuina joka hetki niin, että lasku on jyrkin mahdollinen. Laske Jyrkin laskureitti!

XI.4 Vektorikentät: Divergenssi ja roottori

Tässä luvussa tarkastellaan funktioita, jotka riippuvat kahdesta tai kolmesta fysikaalisesta paikkamuuttujasta (x, y tai x, y, z), mahdollisesti lisäksi fysikaalisesta aikamuuttujasta t. Funktiot voivat olla joko skalaarifunktioita, eli reaaliarvoisia (tai skalaariarvoisia, kirjaimellisesti 'skaalaajafunktioita') tai vektoriarvoisia, eli *vektorikenttiä* (engl. vector field). Jatkossa ajatellaan ainoastaan fysikaalisia vektorikenttiä, joissa on joko kaksi (tasokenttä) tai kolme (avaruuskenttä) komponenttia. Skalaarifunktiot, joita kutsutaan myös *skalaarikentiksi*, voivat siis olla tyyppiä $f(x, y)$, $f(x, y, z)$, $f(x, y, t)$ tai $f(x, y, z, t)$ (joskus aikamuuttuja halutaan kirjoittaa ensimmäiseksi, esim. $f(t, x, y, z)$) ja vektorikentät esim. tyyppiä

$$\vec{F}(x, y, z, t) = F_1(x, y, z, t)\,\vec{i} + F_2(x, y, z, t)\,\vec{j} + F_3(x, y, z, t)\,\vec{k}.$$

Fysikaalisia vektorikenttiä ovat esim. virtaavan nesteen, kaasun, yms. nopeusvektoreiden muodostama virtauskenttä, (esim. gravitaatioon liittyvä) voimakenttä, sähkömagneettisiin ilmiöihin liittyvät sähkökenttä ja magneettikenttä, ym. Tällaiset fysikaaliset kentät ovat joko staattisia tai dynaamisia, riippuen siitä onko muuttujien joukossa myös aika (t) vai ei. Vektorikenttä kuvataan usein graafisesti 'nuolikenttänä' (kuvassa virtauskenttä).

Vektorikenttään kohdistuvina differentiaalioperaattorit $\partial_x, \partial_y, \partial_z, \partial_t$ ymmärretään samalla tavoin kuin skalaarin ja vektorin kertolaskussa, eli operaattorit kohdistuvat vektorikentän jokaiseen komponenttiin. Esim.

$$\partial_x \vec{F}(x, y, z, t) = \frac{\partial F_1}{\partial x}\,\vec{i} + \frac{\partial F_2}{\partial x}\,\vec{j} + \frac{\partial F_3}{\partial x}\,\vec{k}.$$

Gradienttikentät

Jos f on skalaarifunktio (skalaarikenttä) edellä mainittua tyyppiä, niin funktion gradientti ymmärretään vain paikkamuuttujiin x, y tai x, y, z kohdistuvana. Gradientti toimii tällöin differentiaalioperaattorina, joka muuntaa skalaarifunktion vektoriarvoiseksi funktioksi eli vektorikentäksi. Monet fysikaaliset vektorikentät ovat tällaisia *gradienttikenttiä*. Jos $\vec{F}$ on gradienttikenttä ja

$$\vec{F} = -\nabla u$$

(miinusmerkki fysikaalisista mukavuussyistä), niin sanotaan, että u on kentän $\vec{F}$ *(skalaari)potentiaali*.

Esimerkki 1 Fysikaalisia esimerkkejä potentiaaleista ja gradienttikentistä ovat esimerkiksi:

Potentiaali	Gradienttikenttä
$u =$ gravitaatiopotentiaali	$\vec{G} = -\nabla u =$ gravitaatiokenttä
$u =$ sähköpotentiaali	$\vec{E} = -\nabla u =$ sähkökenttä
$u =$ lämpötila	$\vec{J} = -\lambda \nabla u =$ lämpövirran tiheys

Yhteyttä $\vec{J} = -\lambda \nabla u$ sanotaan lämpöopissa Fourier'n laiksi. Sen mukaan siis lämpöenergia virtaa lämpötilan negatiivisen gradientin $-\nabla u$ suuntaan ja virtatiheys (yksikkö $\mathrm{W/m^2}$) on gradientin itseisarvoon verrannollinen. Verrannollisuuskerrointa sanotaan materiaalin lämmönjohtavuudeksi. $\square$

Divergenssi ja roottori. Laplacen operaattori

Differentiaalioperaattori $\nabla = \partial_x \vec{i} + \partial_y \vec{i} + \partial_z \vec{k}$ voidaan myös ymmärtää operaattorina, jonka kohteena ovat vektoriarvoiset funktiot eli vektorikentät. Tällöin operaattori ∇ operoi kohdefunktioon $\vec{F}$ joko pistetulon (skalaaritulon) tai ristitulon (vektoritulon) välityksellä. Kun kohdefunktioksi oletetaan

$$\vec{F}(x, y, z) = F_1(x, y, z)\vec{i} + F_2(x, y, z)\vec{j} + F_3(x, y, z)\vec{k},$$

niin tulevat näin määritellyksi kentän $\vec{F}$ *divergenssi* (engl. divergence)

$$\boxed{\nabla \cdot \vec{F} = \partial_x F_1 + \partial_y F_2 + \partial_z F_3 \quad \text{(divergenssi)}}$$

ja *roottori* (engl. curl)

$$\boxed{\nabla \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial_x & \partial_y & \partial_z \\ F_1 & F_2 & F_3 \end{vmatrix} \quad \text{(roottori).}}$$

Roottorin laskukaavan purettu muoto on

$$\nabla \times \vec{F} = (\partial_y F_3 - \partial_z F_2)\vec{i} - (\partial_x F_3 - \partial_z F_1)\vec{j} + (\partial_x F_2 - \partial_y F_1)\vec{k}.$$

Divergenssiä ja roottoria merkitään myös symboleilla div ja rot (tai $\overrightarrow{\text{rot}}$). Tavallisia lukutapoja ovat 'nabla piste' ja 'nabla risti', vastaten gradientin ∇f lukutapaa 'nabla f'.

Esimerkki 2 Funktion $\vec{F}(x,y,z) = x^2\vec{i} + y^2 z\vec{j} - (2xz + yz^2)\vec{k}$ divergenssi ja roottori ovat

$$\begin{aligned} \nabla \cdot \vec{F} &= 2x + 2yz - 2x - 2yz = 0, \\ \nabla \times \vec{F} &= (\partial_y F_3 - \partial_z F_2)\vec{i} + (\partial_z F_1 - \partial_x F_3)\vec{j} + (\partial_x F_2 - \partial_y F_1)\vec{k} \\ &= -(y^2 + z^2)\vec{i} + 2z\vec{j}. \quad \square \end{aligned}$$

Piste- ja ristitulojen kautta tulevat määritellyksi myös mm. operaattoritulot

(1) $\nabla \cdot \nabla$ (2) $\nabla(\nabla\cdot\,)$ (3) $\nabla \times (\nabla\times\,)$ (4) $\nabla \times \nabla$

Näistä viimeinen on nollaoperaattori:

$$\boxed{\nabla \times \nabla = \vec{0}.}$$

Nollaoperaattori on myös $\nabla \cdot (\nabla\times\,)$, ts. pätee (vrt. Luku II.6)

$$\boxed{\nabla \cdot \nabla \times \vec{F} = (\nabla \times \nabla) \cdot \vec{F} = 0.}$$

Operaattorien (1), (2) ja (3) välillä on yhteys (Harj.teht. 3f — vrt. Luku II.6)

$$\boxed{\nabla \times (\nabla \times \vec{F}) = \nabla(\nabla \cdot \vec{F}) - (\nabla \cdot \nabla)\vec{F}.}$$

Operaattoria $\nabla \cdot \nabla$, joka voi operoida sekä skalaariin että vektoriin, sanotaan *Laplacen operaattoriksi.* Tämä on hyvin yleinen differentiaalioperaattori erilaisissa fysiikan laeissa (ks. huomautukset jäljempänä), siksi sille on vakiintunut erillinen symboli Δ. Myös merkintää ∇^2 käytetään. Lukutapoja ovat 'Laplace' tai 'nabla toiseen'. Määritelmä on siis

$$\begin{aligned} \nabla \cdot \nabla &= (\vec{i}\,\partial_x + \vec{j}\,\partial_y + \vec{k}\,\partial_z) \cdot (\vec{i}\,\partial_x + \vec{j}\,\partial_y + \vec{k}\,\partial_z) \\ &= \partial_x^2 + \partial_y^2 + \partial_z^2 = \Delta. \end{aligned}$$

Operaattorit $\nabla\star$ ja $\vec{u}\cdot\nabla$

Yleisiä 'nablaamissääntöjä' voidaan usein yksinkertaistaa käyttämällä operaattorimerkintää $\nabla\star$, missä '$\star$' voi olla 'piste' ($\nabla\cdot$), 'risti' ($\nabla\times$) tai 'tyhjä' (∇). Viimeksi mainitussa tapauksessa tulkitaan kohde aina skalaarikentäksi, ts. $\nabla\vec{F} = \nabla F$. Operaattorimerkintä $\nabla\star$ on kätevä esimerkiksi tulon derivoimissäännön yleistyksessä: Jos $f = f(x,y,z)$ on skalaarikenttä ja $\vec{F}$ kolmiulotteinen vektorikenttä, niin pätee

$$\boxed{\nabla\star(f\vec{F}) = \nabla f\star\vec{F} + f\nabla\star\vec{F}.}$$

(Tapauksessa $\star$ = 'tyhjä' tämä tulkitaan: $\nabla(fF) = F\nabla f + f\nabla F$.) Perustelu jätetään harjoitustehtäväksi (Harj.teht. 3abc).

Muuttuvakertoimisista vektoridifferentiaalioperaattoreista on syytä mainita sovelluksissa yleinen operaattori

$$\vec{u}\cdot\nabla = u_1\partial_x + u_2\partial_y + u_3\partial_z,$$

missä $\vec{u} = \vec{u}(x,y,z)$ (tai $\vec{u} = \vec{u}(x,y,z,t)$) on vektorikenttä. Tämä on skalaarinen operaattori, joka voi kohdistua skalaari- tai vektorikenttään. Edellisessä (ei jälkimmäisessä) tapauksessa $\vec{u}\cdot\nabla f$ on laskettavissa myös kaksivaiheisesti:

$$f \mapsto \nabla f \mapsto \vec{u}\cdot\nabla f.$$

Muuttuvakertoimisia operaattoreita yhdisteltäessä on huomioitava, että derivointi kohdistuu tällöin myös kertoimiin (vrt. Luku XI.2).

Vectorikentän lähde. Poissonin ja Laplacen yhtälöt

Kun $\vec{F}$ on jokin fysikaalinen kenttä, kuten sähkö-, magneetti-, virtaus- ym. kenttä, sanotaan divergenssiä

$$\nabla\cdot\vec{F} = \rho$$

kentän $\vec{F}$ *lähteeksi*. Kentän ja lähteen ρ välinen yhteys on tällaisissa sovellutuksissa luonnonlaki, jossa lähde tulkitaan kentän $\vec{F}$ aiheuttajaksi — siitä nimi 'lähde'. Tyypillinen lähtökohta onkin, että lähde on tunnettu ja kenttä halutaan määrätä. Esimerkiksi gravitaatiokenttää määrättäessä on ρ verrannollinen massatiheyteen ja sähkökentän tapauksessa varaustiheyteen.

Jos vektorikentän $\vec{F}$ lähde ρ on tunnettu ja lisäksi tiedetään, että $\vec{F}$ on gradienttikenttä, eli on olemassa potentiaali u siten, että $\vec{F} = -\nabla u$ (näin on esim.

gravitaatio- ja sähkökentän tapauksessa), niin kentän ja lähteen välinen yhteys $\nabla\vec{F} = \rho$ voidaan kirjoittaa muotoon

$$-\Delta u = \rho.$$

Tätä sanotaan *Poissonin yhtälöksi.* Sellaisissa pisteissä, joissa kenttä $\vec{F}$ on *lähteetön*, eli $\nabla \cdot \vec{F} = 0$, toteuttaa kentän potentiaali *Laplacen yhtälön*†

$$\Delta u = 0.$$

Funktiota, joka toteuttaa Laplacen yhtälön, sanotaan *harmoniseksi.*

Sovellustilanteissa Poissonin tai Laplacen yhtälö esiintyy tyypillisesti *reuna-arvotehtävässä* (engl. boundary value problem), jossa tarkasteltavan (avoimen) joukon A — tai niinkuin tällaisissa yhteyksissä useammin sanotaan, *alueen* A ($A \subset \mathbb{R}^2$ tai $A \subset \mathbb{R}^3$) — *reunalla* ∂A asetetaan jokin *reunaehto.* Sovellustilanteessa reunaehto määräytyy fysikaalisin perustein. Esimerkiksi jos reunalla asetetaan nk. (homogeeninen) *Dirichlet'n ehto*, niin saadaan reuna-arvotehtävän perusmuoto:

$$\begin{cases} -\Delta u = \rho & A\text{:n sisäpisteissä,} \\ u = 0 & \text{reunalla } \partial A. \end{cases}$$

Tämä on osoitettavissa tietyin (funktiota ρ ja reunaviivaa/pintaa $S = \partial A$ koskevin) säännöllisyysehdoin yksikäsitteisesti ratkeavaksi. Itse ratkaisemisen on yleensä tapahduttava numeerisin (likimääräisin) keinoin, muutamia poikkeustilanteita lukuun ottamatta.

Esimerkki 3 Jos A on R-säteinen kiekko ($A \subset \mathbb{R}^2$) tai kuula ($A \subset \mathbb{R}^3$) ja $\rho = Q =$ vakio, niin reuna-arvotehtävän

$$\begin{cases} -\Delta u = Q & A\text{:n sisäpisteissä,} \\ u = 0 & \text{reunalla } \partial A \end{cases}$$

ratkaisu on kummassakin tapauksessa toisen asteen polynomi:

$$A \subset \mathbb{R}^2: \qquad u(x,y) = \frac{Q}{4}(R^2 - x^2 - y^2).$$
$$A \subset \mathbb{R}^3: \qquad u(x,y,z) = \frac{Q}{6}(R^2 - x^2 - y^2 - z^2). \quad \square$$

†Laplacen ja Poissonin yhtälöiden nimet viittaavat ranskalaisiin matemaatikko-fyysikkoihin **P.S. de Laplace** (1749-1827) ja **S.D. Poisson** (1781-1840), jotka tutkivat tämän tyyppisten yhtälöiden teoriaa. Yhtälöiden käytännöllisen ratkaisemisen perustan loi samoihin aikoihin vaikuttanut **J.B.J. Fourier** (1768-1830). Myös Fourier oli matemaatikko-fysiikko, joka tutki erityisesti lämmönjohtumisilmiöitä. Näistä tutkimuksista alkunsa saaneet *Fourier-sarjat* ja *Fourier-muunnokset* ovat tehneet Fourier'n nimestä kuolemattoman.

Poissonin ja Laplacen yhtälöt ovat esimerkkejä yhtälöistä, joista käytetään yleisnimeä *osittaisdifferentiaaliyhtälö* (engl. partial differential equation, lyh. PDE). Paitsi Poissonin ja Laplacen yhtälöissä, Laplacen operaattori esiintyy monissa muissakin fysiikan osittaisdifferentiaaliyhtälöissä. Mainittakoon tässä ainoastaan ajasta riippuvia ilmiöitä kuvaavat *diffuusioyhtälö* ja *aaltoyhtälö*, joiden perusmuodot ovat (sovelluksissa näistä esiintyy erilaisia variaatioita):

$$\begin{aligned} &\text{Diffuusioyhtälö}: \quad u_t = \Delta u. \\ &\text{Aaltoyhtälö}: \quad u_{tt} = \Delta u. \end{aligned}$$

Diffuusioyhtälön tyyppiä on esimerkiksi (aikariippuva) lämmönjohtumisyhtälö.

Pyörrekenttä. Maxwellin yhtälöt

Fysikaalisen vektorikentän roottoria

$$\nabla \times \vec{F} = \vec{\omega}$$

sanotaan kentän $\vec{F}$ *pyörrekentäksi.* Pyörrekentillä on fysikaalinen merkitys esimerkiksi virtausmekaniikassa, mutta erityisen keskeisellä sijalla ne ovat sähkömagneettisten kenttien teoriassa. Tämä nähdään jo sähkömagneettisten kenttien perusyhtälöistä, *Maxwellin yhtälöistä* (**J.C. Maxwell**, 1864):

$$\begin{cases} \nabla \times \vec{E} = -\dfrac{\partial \vec{B}}{\partial t}, \\ \nabla \times \vec{H} = \dfrac{\partial \vec{D}}{\partial t} + \vec{J}, \\ \nabla \cdot \vec{D} = \rho, \\ \nabla \cdot \vec{B} = 0. \end{cases}$$

Tässä esiintyvien kenttien nimet ovat:

$\vec{E}$:	sähkökenttä	$\vec{D}$:	sähkövuon tiheys
$\vec{H}$:	magneettikenttä	$\vec{B}$:	magneettivuon tiheys
ρ :	varaustiheys	$\vec{J}$:	virtatiheys

Maxwellin yhtälöitä ratkaistaessa on huomioitava kenttien $\vec{E}, \vec{D}$ ja $\vec{H}, \vec{B}$ keskinäinen riippuvuus, joka on väliaineelle ominainen. Yksinkertisimmillaan riippuvuudet ovat ilmaistavissa kahden materiaalivakion ϵ, μ avulla muodossa

$$\vec{D} = \epsilon \vec{E}, \quad \vec{B} = \mu \vec{H}.$$

Vektorikenttää $\vec{F}$ sanotaan *pyörteettömäksi* sellaisessa alueessa, jossa $\nabla \times \vec{F} = \vec{0}$. Kaikki gradienttikentät, eli skalaaripotentiaalin avulla ilmaistavat kentät, ovat pyörteettömiä (koska $\nabla \times \nabla = \vec{0}$). Maxwellin yhtälöiden mukaan sähkökenttä $\vec{E}$ on pyörteetön stationaarisessa tilanteessa, jossa kentät ovat ajasta riippumattomia. Sekä lähteetön että pyörteetön on esimerkiksi gravitaatiokenttä tyhjässä (massattomassa) avaruudessa, ja yleisemminkin gradienttikenttä lähteettömässä alueessa.

Roottori tasossa

Jos tason vektorikenttä tulkitaan (z-koordinaatista riippumattomaksi ja xy-tason suuntaiseksi) avaruuden vektorikentäksi, niin kentän roottori on

$$\nabla \times [F_1(x,y)\,\vec{i} + F_2(x,y)\,\vec{j}\,] = (\partial_x F_2 - \partial_y F_1)\,\vec{k}.$$

Tässä on tapana kirjoittaa

$$\operatorname{rot} \vec{F} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \quad \text{(tason vektorikenttä)},$$

jolloin roottori tasossa tulee määritellyksi skalaarisena operaattorina (kuten div).

Esimerkki 4 Maxwellin yhtälöiden mukaan stationaarisessa tilanteessa magneettikentän $\vec{H}$ pyörrekenttä on sähkövirran tiheys $\vec{J}$. Jos oletetetaan tyypillinen sovellustilanne, jossa virta kulkee suorassa, z-akselin suuntaisessa johtimessa, niin $\vec{J}$ on muotoa

$$\vec{J}(x,y,z) = \begin{cases} I(x,y)\vec{k}, & \text{kun } (x,y) \in A, \\ \vec{0}, & \text{muulloin.} \end{cases}$$

Tässä $A \subset \mathbb{R}^2$ edustaa johtimen poikkileikkausta ja I on (skalaarinen) virran tiheys johtimessa. Magneettikenttä on xy-tason suuntainen ja z-koordinaatista riippumaton, ts. muotoa $\vec{H} = H_1(x,y)\vec{i} + H_2(x,y)\vec{j}$, jolloin Maxwellin yhtälö $\nabla \times \vec{H} = \vec{J}$ pelkistyy johtimen poikkipinnalla skalaariseksi yhtälöksi

$$\operatorname{rot} \vec{H} = \partial_x H_2 - \partial_y H_1 = I(x,y).$$

Johtimen ulkopuolella magneettikenttä on pyörteetön: $\operatorname{rot} \vec{H} = 0$. $\square$

*Navier–Stokesin virtausyhtälöt

Nesteen tai kaasun (esim. veden tai ilman) virtausta hallitsevat perusyhtälöt ovat nk. *säilymislakeja* (engl. conservation laws), jotka ilmaisevat matemaattisesti (osittaisdifferentiaaliyhtälöinä) massan, liikemäärän ja energian säilymisen

virtauksessa. Massan ja liikemäärän säilymislaeissa esiintyvät vektori- ja skalaarikentät ovat

$$\begin{aligned} \vec{u}: &\quad \text{virtausnopeus} \\ p: &\quad \text{paine} \\ \rho: &\quad \text{(massa)tiheys} \end{aligned}$$

Jos virtaavan aineen massatiheys voidaan olettaa vakioksi (esim. vesi, monissa virtaustilanteissa myös ilma) ja lisäksi virtaus on ulkoisista voimista vapaa, niin liikemäärän ja massan säilymislait voidaan kirjoittaa *Navier–Stokesin yhtälöinä* (**C.-L. Navier, G.G. Stokes**, 1822)

$$\begin{cases} \vec{u}_t - \nu\Delta\vec{u} + (\vec{u}\cdot\nabla)\vec{u} + \rho^{-1}\nabla p = \vec{0}, \\ \nabla\cdot\vec{u} = 0. \end{cases}$$

Tässä ν on virtaavan aineen (kinemaattinen) viskositeetti.

HARJOITUSTEHTÄVIÄ

1. Laske divergenssi ja roottori:

   a) $x\vec{i} + y\vec{j} + z\vec{k}$ b) $yz\vec{i} + xz\vec{j} + xy\vec{k}$ c) $xy^2\vec{i} - yz^2\vec{j} + zx^2\vec{k}$

   d) $f(x)\vec{i} + g(y)\vec{j} + h(z)\vec{k}$ e) $f(y,z)\vec{i} + g(x,z)\vec{j} + h(x,y)\vec{k}$

2. Olkoon $\vec{F}(x,y,z) = (xy - z^2)\,\vec{i} + xyz\,\vec{j} + (x - y^2 - z^2)\,\vec{k}$.

   a) Laske $\nabla\cdot\vec{F}$, $\nabla(\nabla\cdot\vec{F})$, $\nabla\times\vec{F}$, $\nabla\times(\nabla\times\vec{F})$ ja $\Delta\vec{F}$ ja tarkista, että pätee $\nabla\times(\nabla\times\vec{F}) = \nabla(\nabla\cdot\vec{F}) - \Delta\vec{F}$.

   b) Laske $(\vec{F}\cdot\nabla)\vec{F}$ ja $(\vec{F}\times\nabla)\cdot\vec{F}$.

3. Olkoon $\vec{F}$ ja $\vec{G}$ tason tai avaruuden (mahdollisesti ajasta riippuvia) vektorikenttiä ja f ja g skalaarikenttiä. Perustele seuraavat säännöt (oletetaan riittävä derivoituvuus):

   a) $\nabla(fg) = g\nabla f + f\nabla g$ b) $\nabla\cdot(f\vec{F}) = \vec{F}\cdot\nabla f + f\,\nabla\cdot\vec{F}$

   c) $\nabla\times(f\vec{F}) = \nabla f\times\vec{F} + f\,\nabla\times\vec{F}$ d) $\partial_t(\nabla\star\vec{F}) = \nabla\star\vec{F}_t$

   e) $\nabla\cdot(\vec{F}\times\vec{G}) = (\nabla\times\vec{F})\cdot\vec{G} - (\nabla\times\vec{G})\cdot\vec{F}$

   f) $\nabla\times(\nabla\times\vec{F}) = \nabla(\nabla\cdot\vec{F}) - \Delta\vec{F}$

   g) $\Box(\vec{F}\circ\vec{G}) = (\Box\vec{F})\circ\vec{G} + \vec{F}\circ\Box\vec{G}$ ja $\Box(f\vec{F}) = (\Box f)\vec{F} + f\Box\vec{F}$, kun $\Box\in\{\partial_x,\partial_y,\partial_z,\partial_t\}$ ja $\circ\in\{\,\cdot\,,\times\,\}$

   h) $(\vec{u}\,\Box)\circ f\vec{F} = \Box f\;\vec{u}\circ\vec{F} + f\vec{u}\circ\Box\vec{F}$, kun $\Box\in\{\partial_x,\partial_y,\partial_z,\partial_t\}$ ja $\circ\in\{\,\cdot\,,\times\,\}$

4. a) Näytä, että seuraavat funktiot ovat määrittelyjoukossaan harmonisia:

$$xy, \quad x^2 - y^2, \quad x^3 - 3xy^2, \quad x^3y - xy^3, \quad \ln(x^2 + y^2), \quad \operatorname{Arctan} \frac{y}{x}$$
$$e^{ax}(A \sin ay + B \cos ay) \quad (a, A, B \in \mathbb{R})$$

b) Monen muuttujan Laplacen operaattori määritellään $\Delta = \partial_1^2 + \ldots + \partial_n^2$. Tutki, millaisille reaalifunktioille f pätee: $f\big(|\mathbf{x}|\big) = f(\sqrt{x_1^2 + \ldots + x_n^2}\,\big)$ on harmoninen muualla kuin origossa.

5. Seuraavilla osittaisdifferentiaaliyhtälöillä väitetään olevan annettua muotoa olevia ratkaisuja. Tarkista väittämät (funktiot f ja g oletetaan kahdesti derivoituviksi).

a) $u_{tt} = c^2 u_{xx}\,, \quad u(x,t) = f(x + ct) + g(x - ct)$

b) $u_{xx} + u_{yy} = 2u_{xy}\,, \quad u(x,y) = xf(x + y) + yg(x + y)$

c) $u_t = u_{xx}\,, \quad u(x,t) = e^{-\omega^2 t}(A \sin \omega x + B \sin \omega x) \quad (\omega, A, B \in \mathbb{R})$

d) $u_t = u_{xx}\,, \quad u(x,t) = t^{-1/2} e^{-x^2/4t}$

e) $u_t = u_{xx} + u_{yy}\,, \quad u(x,y,t) = t^{-1} e^{-(x^2+y^2)/4t}$

6. Näytä, että osittaisdifferentiaaliyhtälön $u_{xy} = 0$ yleinen ratkaisu $\mathbb{R}^2$:ssa on $u(x,y) = f(x) + g(y)$, missä $f : \mathbb{R} \to \mathbb{R}$ ja $g : \mathbb{R} \to \mathbb{R}$ ovat mitä tahansa $\mathbb{R}$:ssä derivoituvia funktioita.

7. (*) Todista (olettaen riittävä derivoituvuuus):

$$\nabla \times (\vec{F} \times \vec{G}) = (\nabla \cdot \vec{G})\vec{F} + (\vec{G} \cdot \nabla)\vec{F} - (\nabla \cdot \vec{F})\vec{G} - (\vec{F} \cdot \nabla)\vec{G}.$$

8. (*) Olkoon $A \subset \mathbb{R}^2$ tasasivuinen kolmio, jonka kärjet ovat $(0,0)$, $(a,0)$ ja $(a/2, \sqrt{3}a/2)$ $(a > 0)$. Näytä, että reuna-arvotehtävällä

$$\begin{cases} -\Delta u = Q & A\text{:n sisäpisteissä,} \\ u = 0 & \text{reunalla } \partial A \end{cases}$$

($Q =$ vakio) on polynomiratkaisu.

9. (*) Tarkastellaan Maxwellin yhtälöitä homogeenisessa väliaineessa, jossa $\rho = 0$, $\vec{J} = \vec{0}$ ja kenttien $\vec{D}, \vec{E}$ ja $\vec{B}, \vec{H}$ välillä on yhteydet $\vec{D} = \epsilon \vec{E}$ ja $\vec{B} = \mu \vec{H}$ (ϵ ja μ materiaalivakioita). Näytä, että kaikki mainitut kentät, sikäli kuin riittävän säännöllisiä, toteuttavat vektorimuotoisen aaltoyhtälön

$$\vec{F}_{tt} = c^2 \Delta \vec{F},$$

missä $c = (\epsilon\mu)^{-1/2}$ (= valon nopeus väliaineessa).

XI.5 Operaattorit grad, div, rot ja Δ käyräviivaisissa koordinaatistoissa

Fysikaalisten symmetrioiden vuoksi vektorikenttä/skaalaarikenttä halutaan usein esittää käyräviivaisessa polaari-, lieriö- tai pallokoordinaatistossa. Tällöin myös vektoridifferentiaalioperaattorit on muunnettava karteesisesta ko. käyräviivaiseen koordinaatistoon. Tarkastellaan esimerkkinä muunnosta tason polaarikoordinaatistoon.

Ensinnäkin on tutkittava, miten operaattorit ∂_x ja ∂_y ilmaistaan polaarikoordinaatiston vastaavien operaattorien ∂_r ja ∂_φ avulla. Lähtökohtana on ketjusääntö (ks. Luku XI.2), jonka mukaan

$$u(x,y) = v(r,\varphi) = v(r(x,y),\varphi(x,y))$$
$$\Rightarrow \begin{cases} \partial_x u = \dfrac{\partial r}{\partial x}\,\partial_r v + \dfrac{\partial \varphi}{\partial x}\,\partial_\varphi v, \\ \partial_y u = \dfrac{\partial r}{\partial y}\,\partial_r v + \dfrac{\partial \varphi}{\partial y}\,\partial_\varphi v. \end{cases}$$

Tässä on $r = r(x,y) = \sqrt{x^2+y^2}$, joten

$$\frac{\partial r}{\partial x} = \frac{x}{r} = \cos\varphi, \quad \frac{\partial r}{\partial y} = \frac{y}{r} = \sin\varphi.$$

Derivoimalla implisiittisesti x:n suhteen (implisiittinen osittaisderivointi!) saadaan tällöin

$$\begin{aligned} x = r\cos\varphi \;&\Rightarrow\; 1 = \frac{\partial r}{\partial x}\cos\varphi - r\sin\varphi\frac{\partial\varphi}{\partial x} \\ &\Rightarrow\; 1 = \cos^2\varphi - r\sin\varphi\frac{\partial\varphi}{\partial x} \\ &\Rightarrow\; \frac{\partial\varphi}{\partial x} = -\frac{1}{r}\sin\varphi. \end{aligned}$$

Vastaavasti derivoimalla y:n suhteen saadaan

$$\begin{aligned} 0 = \frac{\partial r}{\partial y}\cos\varphi - r\sin\varphi\frac{\partial\varphi}{\partial y} &= \sin\varphi\cos\varphi - r\sin\varphi\frac{\partial\varphi}{\partial y} \\ \Rightarrow\; \frac{\partial\varphi}{\partial y} &= \frac{1}{r}\cos\varphi. \end{aligned}$$

Näin on saatu muunnoskaavat

$$\boxed{\begin{aligned} \partial_x &= \cos\varphi\,\partial_r - \frac{1}{r}\sin\varphi\,\partial_\varphi, \\ \partial_y &= \sin\varphi\,\partial_r + \frac{1}{r}\cos\varphi\,\partial_\varphi. \end{aligned}}$$

Kun näiden lisäksi huomioidaan kantavektorien $\vec{i}, \vec{j}$ ja $\vec{e}_r, \vec{e}_\varphi$ väliset muunnoskaavat

$$\begin{cases} \vec{i} = \cos\varphi\,\vec{e}_r - \sin\varphi\,\vec{e}_\varphi, \\ \vec{j} = \sin\varphi\,\vec{e}_r + \cos\varphi\,\vec{e}_\varphi, \end{cases}$$

niin operaattorille ∇ saadaan ensin esitysmuoto

$$\begin{aligned} \nabla \;=\; \vec{i}\,\partial_x + \vec{j}\,\partial_y \;=\; & (\cos\varphi\,\vec{e}_r - \sin\varphi\,\vec{e}_\varphi)(\cos\varphi\,\partial_r - \frac{1}{r}\sin\varphi\,\partial_\varphi) \\ + \; & (\sin\varphi\,\vec{e}_r + \cos\varphi\,\vec{e}_\varphi)(\sin\varphi\,\partial_r + \frac{1}{r}\cos\varphi\,\partial_\varphi) \end{aligned}$$

ja sievennysten jälkeen

$$\boxed{\nabla = \vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi \quad \text{(polaarikoordinaatisto).}}$$

Esimerkki 1 Laske funktion $u(x,y) = x/(x^2+y^2)$ gradientti polaarikoordinaatistossa.

Ratkaisu Koska

$$u(x,y) \;=\; \frac{x}{x^2+y^2} \;=\; \frac{r\cos\varphi}{r^2} \;=\; \frac{\cos\varphi}{r} \;=\; v(r,\varphi),$$

niin em. muunnoskaavan mukaan

$$\nabla u \;=\; \frac{\partial v}{\partial r}\,\vec{e}_r + \frac{1}{r}\frac{\partial v}{\partial \varphi}\,\vec{e}_\varphi \;=\; \underline{\underline{-\frac{1}{r^2}(\cos\varphi\,\vec{e}_r + \sin\varphi\,\vec{e}_\varphi)}}. \quad \square$$

Annetun vektorikentän $\vec{F}$ divergenssin laskemiseksi oletetaan, että $\vec{F}$ tunnetaan muodossa

$$\vec{F} = F_r(r,\varphi,z)\,\vec{e}_r + F_\varphi(r,\varphi,z)\,\vec{e}_\varphi.$$

Tällöin on

$$\begin{aligned} \nabla\cdot\vec{F} \;=\; & \left(\vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi\right)\cdot(F_r\vec{e}_r + F_\varphi\vec{e}_\varphi) \\ =\; & (\vec{e}_r\,\partial_r)\cdot(F_r\vec{e}_r) \;+\; (\vec{e}_r\,\partial_r)\cdot(F_\varphi\vec{e}_\varphi) \\ +\; & \frac{1}{r}\,(\vec{e}_\varphi\,\partial_\varphi)\cdot(F_r\vec{e}_r) + \frac{1}{r}\,(\vec{e}_\varphi\,\partial_\varphi)\cdot(F_\varphi\vec{e}_\varphi). \end{aligned}$$

Tässä on huomioitava, että vektorit $\vec{e}_r$ ja $\vec{e}_\varphi$ eivät ole vakiovektoreita vaan riippuvat koordinaatista φ. Tarvitaan siis myös derivaatat $\partial_\varphi \vec{e}_r$ ja $\partial_\varphi \vec{e}_\varphi$:

$$\begin{aligned}\partial_\varphi \vec{e}_r &= \partial_\varphi(\cos\varphi\,\vec{i} + \sin\varphi\,\vec{j}) = -\sin\varphi\,\vec{i} + \cos\varphi\,\vec{j},\\ \partial_\varphi \vec{e}_\varphi &= \partial_\varphi(-\sin\varphi\,\vec{i} + \cos\varphi\,\vec{j}) = -\cos\varphi\,\vec{i} - \sin\varphi\,\vec{j},\end{aligned}$$

eli

$$\boxed{\partial_\varphi \vec{e}_r = \vec{e}_\varphi, \quad \partial_\varphi \vec{e}_\varphi = -\vec{e}_r.}$$

Käyttämällä näitä kaavoja ja tulon derivoimissääntöjä (ks. Harj.teht. XI.4: 3h) saadaan ym. lauseke puretuksi termeittäin:

$$\begin{aligned}(\vec{e}_r\,\partial_r)\cdot(F_r\vec{e}_r) &= \partial_r F_r\,\vec{e}_r\cdot\vec{e}_r + F_r\,\vec{e}_r\cdot\partial_\varphi\vec{e}_r = \partial_r F_r + F_r\,\vec{e}_r\cdot\vec{e}_\varphi = \partial_r F_r,\\ (\vec{e}_r\,\partial_r)\cdot(F_\varphi\vec{e}_\varphi) &= \partial_r F_\varphi\,\vec{e}_r\cdot\vec{e}_\varphi + F_\varphi\,\vec{e}_r\cdot\partial_r\vec{e}_\varphi = 0,\\ \frac{1}{r}(\vec{e}_\varphi\,\partial_\varphi)\cdot(F_r\vec{e}_r) &= \frac{1}{r}\left[\partial_\varphi F_r\,\vec{e}_\varphi\cdot\vec{e}_r + F_r\,\vec{e}_\varphi\cdot\partial_\varphi\,\vec{e}_r\right] = \frac{1}{r}F_r\,\vec{e}_\varphi\cdot\vec{e}_\varphi = \frac{1}{r}F_r\,,\\ \frac{1}{r}(\vec{e}_\varphi\,\partial_\varphi)\cdot(F_\varphi\vec{e}_\varphi) &= \frac{1}{r}\left[\partial_\varphi F_\varphi\,\vec{e}_\varphi\cdot\vec{e}_\varphi + F_\varphi\,\vec{e}_\varphi\cdot\partial_\varphi\,\vec{e}_\varphi\right]\\ &= \frac{1}{r}\left[\partial_\varphi F_\varphi - F_\varphi\,\vec{e}_\varphi\cdot\vec{e}_r\right] = \frac{1}{r}\partial_\varphi F_\varphi.\end{aligned}$$

Lopputulos:

$$\boxed{\nabla\cdot\vec{F} = \partial_r F_r + \frac{1}{r}F_r + \frac{1}{r}\partial_\varphi F_\varphi \quad \text{(polaarikoordinaatisto).}}$$

Kaksi ensimmäistä termiä yhdistämällä saadaa tälle vaihtoehtoinen nk. *säilymislakimuoto* (engl. conservation form)

$$\boxed{\nabla\cdot\vec{F} = \frac{1}{r}\partial_r(rF_r) + \frac{1}{r}\partial_\varphi F_\varphi \quad \text{(säilymislakimuoto).}}$$

Esimerkki 2 Vektorikentän $\vec{F} = r^{-1}\,\vec{e}_r$ divergenssi muualla kuin origossa on

$$\nabla\cdot\vec{F} = \frac{1}{r}\partial_r\left(r\cdot\frac{1}{r}\right) = 0,$$

joten kenttä on lähteetön aluessa $A = \{(x,y)\in\mathbb{R}^2 \mid (x,y)\neq(0,0)\}$. $\square$

Seuraavassa taulukossa esitetään koottuna operaattorien ∇, $\nabla\cdot$, $\nabla\times$ ja Δ (säilymislakimuotoiset) laskusäännöt lieriö- ja pallokoordinaatistoissa.

Lieriökoordinaatisto

$$\nabla = \vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi + \vec{e}_z\,\partial_z$$

$$\nabla\cdot\vec{F} = \frac{1}{r}\,\partial_r(rF_r) + \frac{1}{r}\,\partial_\varphi F_\varphi + \partial_z F_z$$

$$\nabla\times\vec{F} = \frac{1}{r}\begin{vmatrix}\vec{e}_r & r\vec{e}_\varphi & \vec{e}_z\\ \partial_r & \partial_\varphi & \partial_z\\ F_r & rF_\varphi & F_z\end{vmatrix}$$

$$\Delta = \frac{1}{r}\,\partial_r(r\partial_r) + \frac{1}{r^2}\partial_\varphi^2 + \partial_z^2$$

Pallokoordinaatisto

$$\nabla = \vec{e}_r\,\partial_r + \frac{1}{r}\vec{e}_\theta\,\partial_\theta + \frac{1}{r\sin\theta}\vec{e}_\varphi\,\partial_\varphi$$

$$\nabla\cdot\vec{F} = \frac{1}{r^2}\partial_r(r^2F_r) + \frac{1}{r\sin\theta}\partial_\theta(\sin\theta\,F_\theta) + \frac{1}{r\sin\theta}\partial_\varphi F_\varphi$$

$$\nabla\times\vec{F} = \frac{1}{r^2\sin\theta}\begin{vmatrix}\vec{e}_r & r\vec{e}_\theta & r\sin\theta\,\vec{e}_\varphi\\ \partial_r & \partial_\theta & \partial_\varphi\\ F_r & rF_\theta & r\sin\theta\,F_\varphi\end{vmatrix}$$

$$\Delta = \frac{1}{r^2}\partial_r(r^2\partial_r) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\,\partial_\theta) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2$$

Taulukon mukaan Laplacen operaattorin muunnos tason karteesisesta koordinaatistosta polaarikoordinaatistoon on

$$\Delta = \partial_x^2 + \partial_y^2 = \frac{1}{r}\,\partial_r(r\partial_r) + \frac{1}{r^2}\,\partial_\varphi^2 = \partial_r^2 + \frac{1}{r}\,\partial_r + \frac{1}{r^2}\,\partial_\varphi^2.$$

Tähän päädytään sieventämällä edellä kuvatulla tavalla operaattorilauseke

$$\Delta = \nabla\cdot\nabla = \left(\vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi\right)\cdot\left(\vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi\right)$$

(Harj.teht. 4).

Esimerkki 3 Olkoon kenttä $\vec{F}$ on pallosymmetrinen, eli pallokoordinaatistossa muotoa $\vec{F} = F(r)\vec{e}_r$. Millainen on funktio $F(r)$, jos kenttä on origon ulkopuolella lähteetön?

Ratkaisu Taulukon mukaan kentän divergenssi on $\nabla \cdot \vec{F} = r^{-2}\partial_r(r^2F(r))$, joten origon ulkopuolella ($r > 0$) on oltava

$$\partial_r(r^2F(r)) = 0 \;\Rightarrow\; r^2F(r) = A \;\Rightarrow\; F(r) = \frac{A}{r^2}, \quad A \in \mathbb{R}. \quad \square$$

Polaari- tai pallosymmetrisessä tilanteessa, jossa lähde ρ ja vektorikentän potentiaali u riippuvat vain radiaalikoordinaatista r, voidaan Poissonin yhtälö kirjoittaa tavallisena differentiaaliyhtälönä (vrt. taulukko edellä):

$$\begin{aligned} &\text{Polaarisymmetria}: && \frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) &&= u'' + \frac{1}{r}u' = -\rho(r). \\ &\text{Pallosymmetria}: && \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{du}{dr}\right) &&= u'' + \frac{2}{r}u' = -\rho(r). \end{aligned}$$

Nämä ovat molemmat Eulerin tyyppiä (vrt. Luku IX.5), joten ne ovat ratkaistavissa kvadratuureilla. Nopeimmin ratkaiseminen käy suoraan säilymislakimuodosta.

Esimerkki 4 Lämpötila pyöreän vastuslangan poikkipinnalla toteuttaa yhtälön

$$-\Delta u = Q = \text{vakio}.$$

Määrää u, kun tiedetään, että ulkopinnalla $u(R) = 0$.

Ratkaisu Lähdetään säilymislakimuodosta (polaarisymmetria):

$$\begin{aligned} &-\frac{1}{r}\frac{d}{dr}(ru') = Q \;\Rightarrow\; \frac{d}{dr}(ru') = -Qr \\ &\Rightarrow\; ru' = -\frac{1}{2}Qr^2 + A \;\Rightarrow\; u' = -\frac{1}{2}Qr + \frac{A}{r} \\ &\Rightarrow\; u(r) = -\frac{1}{4}Qr^2 + A\ln r + B \quad (A, B \in \mathbb{R}) \\ &\Rightarrow\; A = 0,\; B = \frac{1}{4}QR^2 \\ &\Rightarrow\; u(r) = \underline{\underline{\frac{1}{4}Q(R^2 - r^2)}}. \end{aligned}$$

Karteesisessa koordinaatistossa esitettynä ratkaisu on

$$u(x,y) = \frac{Q}{4}\,(R^2 - x^2 - y^2).$$

Ratkaisusta oli jätettävä pois logaritminen termi $A \ln r = \frac{1}{2} A \ln(x^2 + y^2)$, syystä että yhtälö $-\Delta u = Q$ ei muuten toteutuisi origossa (muualla kylläkin). $\square$

HARJOITUSTEHTÄVIÄ

1. Laske vektorikentän

$$\vec{F} = (x^2 + y^2 + z^2)(x\vec{i} + y\vec{j} + z\vec{k})$$

divergenssi $\nabla \cdot \vec{F}$ a) karteesisessa koordinaatistossa, b) muuntamalla $\vec{F}$ ensin pallokoordinaatistoon.

2. Näytä, että seuraavat polaari- tai pallokoordinaatistossa määritellyt funktiot ovat määrittelyjoukossaan harmonisia.

a) $u(r,\varphi) = \dfrac{\cos\varphi}{r}$ b) $u(r,\varphi) = \dfrac{\sin 2\varphi}{r^2}$ c) $u(r,\theta,\varphi) = \dfrac{\sin\theta\cos\varphi}{r^2}$

3. Vektorikentästä $\vec{F} = F_1\vec{i} + F_2\vec{j} + F_3\vec{k}$ tiedetään, että funktiot F_i ovat $\mathbb{R}^3$:ssa differentioituvia ja että kenttä on $\mathbb{R}^3$:ssa pyörteetön. Määritä kentän karteesiset komponentit F_i, kun tiedetään lisäksi, että kenttä on pallokoordinaatistossa muotoa

$$\vec{F} = r^3 \sin^2\theta\,\vec{e}_r + F_\theta(r,\theta,\varphi)\,\vec{e}_\theta\,.$$

4. Sievennä Lapalacen operaattorin lauseke polaarikoordinaatistossa, ts. näytä, että

$$\left(\vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi\right)\cdot\left(\vec{e}_r\,\partial_r + \frac{1}{r}\,\vec{e}_\varphi\,\partial_\varphi\right) = \frac{1}{r}\,\partial_r(r\partial_r) + \frac{1}{r^2}\partial_\varphi^2\,.$$

5. Sähköjohtimessa kulkeva virta, jonka tiheys on $\vec{J} = J(r)\vec{e}_z$ (lieriökoordinaatisto), aiheuttaa magneettikentän muotoa $\vec{H} = H(r)\vec{e}_\varphi$, missä $H(0) = 0$. Määrää $H(r)$ $J(r)$:n avulla Maxwellin yhtälöstä $\nabla \times \vec{H} = \vec{J}$.

6. Homogeenisesta materiaalista valmistetussa lieriön tai lieriökuoren muotoisen kappaleen poikkipinnalla

$$A = \{(x, y) \in \mathbb{R}^2 \mid R_1^2 \leq x^2 + y^2 \leq R_2^2\}$$

lämpötila u toteuttaa A:n sisäpisteissä Poissonin yhtälön $-u_{xx} - u_{yy} = q$, missä q riippuu vain polaarikoordinaatista r, sekä reunalla ∂A annetut reunaehdot. Määrää u seuraavissa tapauksissa:

a) $R_1 = 0,\ R_2 = R,\ q(r) = Q =$ vakio, $u(R) = u_0$
b) $R_1 = 0,\ R_2 = R,\ q(r) = Q =$ vakio, $u'(R) = -ku(R)$
c) $R_1 = 0,\ R_2 = R,\ q(r) = Q(1 - r^2/R^2),\ u'(R) = 0$
d) $R_1 = R,\ R_2 = 2R,\ q(r) = 0, u(R_1) = u_0,\ u'(R_2) = 0$
e) $R_1 = R,\ R_2 = 2R,\ q(r) = Q =$ vakio, $u(R_1) = u(R_2) = 0$

7. Homogeenisesta materiaalista valmistetussa pallon tai pallokuoren muotoisessa kapplaeessa

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid R_1^2 \leq x^2 + y^2 + z^2 \leq R_2^2\}$$

lämpötila u toteuttaa A:n sisäpisteissä Poissonin yhtälön $-\Delta u = q$, missä q riippuu vain pallokoordinaatista r, sekä reunalla ∂A annetut reunaehdot. Määrää u edellisen tehtävän tapauksissa.

8. (*) Johda pallokoordinaatiston kantavektorien derivoimiskaavat

$$\begin{aligned} &\partial_\theta \vec{e}_r = \vec{e}_\theta\,, && \partial_\theta \vec{e}_\theta = -\vec{e}_r\,, && \partial_\theta \vec{e}_\varphi = \vec{0}\,, \\ &\partial_\varphi \vec{e}_r = \sin\theta \vec{e}_\varphi\,, && \partial_\varphi \vec{e}_\theta = \cos\theta \vec{e}_\varphi\,, && \partial_\varphi \vec{e}_\varphi = -\sin\theta \vec{e}_r - \cos\theta \vec{e}_\theta \end{aligned}$$

ja näiden avulla gradientin ja Laplacen operaattorin esitysmuodot pallokoordinaatistossa.

9. (*) (Keplerin laki) Avaruusalus, jonka massa $= m$, liikkuu tason keskeisvoimakentässä $\vec{F} = \nabla(k/r)$ ($k =$ vakio, $r =$ etäisyys origosta) siten, että aluksen polaarikoordinaatit ovat $r(t)$ ja $\varphi(t)$ hetkellä t. Kirjoita aluksen liikeyhtälö $m\vec{r}'' = \vec{F}$ differentiaaliyhtälösysteeminä funktioille $r(t)$ ja $\varphi(t)$ ja näytä, että pätee <u>Keplerin toinen laki</u>

$$[r(t)]^2 \varphi'(t) = \text{vakio}.$$

XI.6 Epälineaariset yhtälöryhmät: Jacobin matriisi ja Newtonin menetelmä

Tarkastellaan yleistä m yhtälön ja n tuntemattoman yhtälöryhmää muotoa

$$\begin{cases} f_1(x_1,\ldots,x_n) &= 0 \\ \quad\vdots \\ f_m(x_1,\ldots,x_n) &= 0 \end{cases}$$

missä funktiot f_i, $i = 1\ldots m$, ovat n reaalimuuttujan reaaliarvoisia funktioita, eli funktioita tyyppiä $f_i : \mathcal{D}_{f_i} \to \mathbb{R}$, $\mathcal{D}_{f_i} \subset \mathbb{R}^n$. Yhtälöryhmä on *epälineaarinen*, jos se ei ole lineaarinen, eli jos ainakin yksi funktioista f_i on ei-affiininen kuvaus (vrt. Luku X.7). Kun määritellään

$$\mathbf{f} = (f_1,\ldots,f_m), \quad \mathcal{D}_{\mathbf{f}} = \mathcal{D}_{f_1} \cap \mathcal{D}_{f_2} \cap \cdots \cap \mathcal{D}_{f_m},$$

niin ym. yhtälöryhmän voi kirjoittaa kätevästi vektorimuodossa:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

Tässä siis $\mathbf{f}$ on funktio (kuvaus) tyyppiä $\mathbf{f} : \mathcal{D}_{\mathbf{f}} \to \mathbb{R}^m$, $\mathcal{D}_{\mathbf{f}} \subset \mathbb{R}^n$. Tämä on *vektorimuuttujan vektoriarvoinen* funktio, jonka voi tulkita aiemmin Luvuissa X.6–X.7 määriteltyjen lineaari- ja affiinikuvausten yleistykseksi. Myös termiä *vektorikenttä* (vrt. Luku XI.4) voi käyttää. Yhtälöryhmiä matriisialgebran keinoin käsiteltäessä tulkitaan $\mathbf{f}$ tavallisesti pystyvektoriksi. Kuten nähdään jatkossa, matriisialgebraa tarvitaan muun muassa, kun halutaan määrittää algoritmisin keinoin ko. yhtälöryhmän *ratkaisut*, eli joukko

$$X = \{\mathbf{x} \in \mathcal{D}_{\mathbf{f}} \mid \mathbf{f}(\mathbf{x}) = \mathbf{0}\}.$$

Yhtälöryhmiä käytännössä (numeerisesti) ratkaistaessa on tavallisimmin $m = n$.

Esimerkki 1 Epälineaaristen yhtälöryhmien

$$\text{a)} \begin{cases} 3x + y + z + 9 = 0 \\ 7x + y + 3z + 27 = 0 \\ x^2 + y^2 + z^2 - 81 = 0 \end{cases} \qquad \text{b)} \begin{cases} x + y + z - 16 = 0 \\ x^2 + y^2 + z^2 - 81 = 0 \end{cases}$$

vektorimuotoisia esitystapoja ovat

$$\begin{aligned} \text{a) } \mathbf{f}(x,y,z) &= (\,3x + y + z + 9,\ 7x + y + 3x + 27,\ x^2 + y^2 + z^2 - 81\,) \\ &= (0,0,0), \\ \text{b) } \mathbf{f}(x,y,z) &= (x + y + z - 16,\ x^2 + y^2 + z^2 - 81\,) \\ &= (0,0), \end{aligned}$$

tai pystyvektoreiden avulla

$$\text{a) } \mathbf{f}(x,y,z) = \begin{bmatrix} 3x+y+z+9 \\ 7x+y+3z+27 \\ x^2+y^2+z^2-81 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\text{b) } \mathbf{f}(x,y,z) = \begin{bmatrix} x+y+z-16 \\ x^2+y^2+z^2-81 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Algebrallisin (tai b-kohdassa geometrisin) keinoin voidaan todeta ratkaisuiksi

$$\text{a) } X = \{\, (0,0,-9),\ (6,-6,3)\,\}, \quad \text{b) } X = \emptyset. \quad \square$$

Vektoriarvoista funktiota $\mathbf{f}$ sanotaan *differentioituvaksi* pisteessä $\mathbf{x} \in \mathcal{D}_{\mathbf{f}}$, jos jokainen komponenttifunktio f_i on ko. pisteessä differentioituva (Määritelmän XI.3.1 mielessä). Jos $\mathbf{f}$ on differentioituva, niin yhdistämällä komponenttifunktioiden differentiaalikehitelmät saadaan tulos

$$\mathbf{f}(\mathbf{x}+\Delta\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{J}\,\mathbf{f}(\mathbf{x})\Delta\mathbf{x} + \mathbf{r}(\mathbf{x}),$$

missä $\mathbf{J}$ on $\mathbf{f}$:n *Jacobin*† *matriisi* (engl. Jacobian)

$$\boxed{\mathbf{J}\,\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{x}) & \dots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \\ \dfrac{\partial f_m}{\partial x_1}(\mathbf{x}) & \dots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{bmatrix} \qquad \text{(Jacobin matriisi)}}$$

ja jäännösfunktiolle $\mathbf{r}$ pätee arvio

$$|\mathbf{r}(\mathbf{x})| = o(|\Delta\mathbf{x}|), \quad \text{kun } |\Delta x| \to 0.$$

Tapauksessa $m = 1$ Jacobin matriisi on sama kuin transponoitu gradientti, eli $\mathbf{J}f = (\boldsymbol{\nabla} f)^T$. Tapauksessa $n = 1$ Jacobin matriisi puolestaan palautuu vektoriarvoisen yhden muuttujan funktion (tavalliseksi) derivaataksi. Tapauksissa $m = 2, 3$ tämä on ennestään tuttu parametrisen käyrän derivaattana (Luku VII.2). Jacobin matriisi on siis jälleen derivaattakäsitteen yleistys kahdellakin eri tavalla. Määritelmästä nähdään yleisemmin, että Jacobin matriisi on esitettävissä joko gradienttien ∇f_i tai osittaisderivaattojen $\partial_j \mathbf{f}$ avulla muodoissa

$$\mathbf{Jf} = [\nabla f_1 \dots \nabla f_m]^T = [\partial_1 \mathbf{f} \dots \partial_n \mathbf{f}].$$

†Saksalainen matemaatikko **K.G.J. Jacobi** eli vuosina 1804-1851.

Esimerkki 2 Kun yhtälöryhmä

$$\begin{cases} x^3 - 2xy^5 - x - 2 = 0 \\ -x^3y + 2xy^2 + y^5 + 3 = 0 \end{cases}$$

kirjoitetaan muotoon $\mathbf{f}(x,y) = \mathbf{0}$, $\mathbf{f} = [f_1, f_2]^T$, niin Jacobin matriisi on

$$\mathbf{J}\,\mathbf{f}(x,y) = \begin{bmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{bmatrix} = \begin{bmatrix} 3x^2 - 2y^5 - 1 & -10xy^4 \\ -3x^2 + 2y^2 & -x^3 + 4xy + 5y^4 \end{bmatrix}.$$

Yhtälöryhmän eräs ratkaisu on $(x,y) = (2,1)$. Tässä pisteessä on

$$\mathbf{J}\,\mathbf{f}(2,1) = \begin{bmatrix} 9 & -20 \\ -10 & 5 \end{bmatrix}. \quad \square$$

Jos $\mathbf{f}$ on differentioituva pisteessä $\mathbf{a}$, niin approksimaatiota

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{a}) + \mathbf{J}\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

sanotaan aiempaan tapaan (vrt. Luku V.3) $\mathbf{f}$:n *linearisaatioksi* pisteessä $\mathbf{a}$. (Oikeammin kyse on 'affinisaatiosta'.) Koska approksimaation virhe on suuruusluokkaa $o(|\mathbf{x} - \mathbf{a}|)$ (differentioituvuuden määritelmän mukaisesti), niin johtopäätös on:

$$\boxed{\text{Differentioituva kuvaus} \;\approx\; \text{affiinikuvaus } \underline{\text{paikallisesti}}.}$$

Esimerkki 2 (jatko) Pisteen $(-1,1) \mathrel{\widehat{=}} \mathbf{a} = [-1,1]^T$ ympäristössä on likimain

$$\mathbf{f}(x,y) \;\approx\; \mathbf{f}(-1,1) + \mathbf{J}(-1,1)(\mathbf{x} - \mathbf{a}) \;=\; \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 & 10 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x+1 \\ y-1 \end{bmatrix}. \quad \square$$

Newtonin menetelmä yhtälöryhmille

Tarkastellaan yhtälöryhmää $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, missä $\mathbf{f}$ on tyyppiä $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$. Newtonin menetelmässä tehdään yhtälöryhmässä linearisoiva approksimaatio pisteessä $\mathbf{x}_k$, eli ratkaistaan

$$\mathbf{f}(\mathbf{x}_k) + \mathbf{J}(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) = \mathbf{0}.$$

Kun ratkaisua merkitään $\mathbf{x} = \mathbf{x}_{k+1}$, on tuloksena iteraatiokaava

$$\mathbf{J}(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) = -\mathbf{f}(\mathbf{x}_k)$$

eli

$$\boxed{\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}(\mathbf{x}_k)^{-1}\,\mathbf{f}(\mathbf{x}_k) \quad \text{(Newton).}}$$

Perusajatus tässä on täsmälleen sama kuin yhden yhtälön ($n = 1$) tapauksessa (vrt. Luku V.7): kyseessä on *kiintopisteiteraatio* muotoa

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k), \quad k = 0, 1, \ldots,$$

missä $\mathbf{F}(\mathbf{x}) = \mathbf{x} - \mathbf{J}(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x})$.

Newtonin menetelmä suppenee myös yhtälöryhmien ratkaisussa samantapaisin oletuksin kuin yhden yhtälön tapauksessa. — Kerrattakoon Luvusta XI.1, että vektorijonon $\{\mathbf{x}_k\}$ suppeneminen kohti vektoria $\mathbf{c} \in \mathbb{R}^n$ tarkoittaa:

$$\mathbf{x}_k \to \mathbf{c} \quad \Leftrightarrow \quad |\mathbf{x}_k - \mathbf{c}| \to 0 \quad (k \to \infty),$$

missä $|\cdot|$ on $\mathbb{R}^n$:n euklidinen normi. Kuten tapauksessa $n = 1$, Newtonin menetelmän suppeneminen on myös (riittävin oletuksin) *kvadraattista* seuraavassa konvergenssilauseessa lähemmin määriteltävällä tavalla. Lause on todistuksineen varsin suoraviivainen yleistys Lauseesta V.7.4, mutta todistuksen yksityiskohdat ovat teknisempiä. (Todistus nojaa usean muuttujan Taylorin lauseeseen, ks. Luku XI.9 jäljempänä, vrt. myös Harj.teht. VII.4:16.) Sivuutetaan todistus.

Lause XI.6.1 Olkoon $\mathbf{f} : \mathcal{D}_\mathbf{f} \to \mathbb{R}^n$, $\mathcal{D}_\mathbf{f} \subset \mathbb{R}^n$, $\mathbf{f} = (f_i)$, ja olkoon $\mathbf{f}(\mathbf{c}) = \mathbf{0}$, $\mathbf{c} \in \mathbb{R}^n$. Olkoot edelleen osittaisderivaatat $\partial^\alpha f_i(\mathbf{x})$ olemassa ja jatkuvia pisteen $\mathbf{c}$ ympäristössä, kun $i = 1 \ldots n$ ja $|\alpha| \leq 2$, ja olkoon Jacobin matriisi $\mathbf{J}\,\mathbf{f}(\mathbf{c})$ säännöllinen. Tällöin Newtonin iteraatio $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}(\mathbf{x}_k)^{-1}\,\mathbf{f}(\mathbf{x}_k)$ suppenee kohti $\mathbf{c}$:tä, mikäli $|\mathbf{x}_0 - \mathbf{c}|$ on riittävän pieni, ja suppeneminen on kvadraattista:

$$|\mathbf{x}_{k+1} - \mathbf{c}| \leq C|\mathbf{x}_k - \mathbf{c}|^2 \quad (C = \text{vakio}).$$

Esimerkki 3 Etsi Newtonin menetelmällä ratkaisu yhtälöryhmälle

$$\begin{cases} x^3 - 2xy^5 - x = 2.04 \\ -x^3y + 2xy^2 + y^5 = -0.05 \end{cases}$$

pisteen $(-1, 1)$ läheltä.

<u>Ratkaisu</u> Tässä on

$$\mathbf{f}(x, y) = \begin{bmatrix} x^3 - 2xy^5 - x - 2.04 \\ -x^3y + 2xy^2 + y^5 + 0.05 \end{bmatrix}$$

ja Jacobin matriisi $\mathbf{J}(x, y)$ on sama kuin Esimerkissä 2. Iteraatiokaavan

$$[x_{k+1},\, y_{k+1}]^T = [x_k,\, y_k]^T - \mathbf{J}(x_k, y_k)^{-1}\mathbf{f}(x_k, y_k)$$

mukaan laskien saadaan

$$\begin{aligned}
(x_0, y_0) &= (-1,\ 1) \\
(x_1, y_1) &= (-0.942, 1.004) \\
(x_2, y_2) &= (-0.918940, 1.006156) \\
(x_3, y_3) &= (-0.914247, 1.006622) \\
(x_4, y_4) &= (-0.914214, 1.006633) \\
(x_5, y_5) &= (-0.914216, 1.006633) \\
&\vdots \qquad \square
\end{aligned}$$

Esimerkissä Newtonin iteraatio suppeni nopeasti, koska alkuarvaus oli lähellä ratkaisua. Ellei hyvää alkuarvausta ole käytettävissä, voidaan ensin koettaa paikallistaa ratkaisut jollakin yksinkertaisemmalla menetelmällä. Esimerkiksi jos yhtälöryhmä ei ole kooltaan suuri, voidaan kokeilemalla tutkia, missä funktio

$$F(\mathbf{x}) = \sum_{i=1}^{n} [f_i(\mathbf{x})]^2$$

saa pieniä arvoja. Vieläkin suoraviivaisempi on 'yrityksen ja erehdyksen' menetelmä: Iteroidaan vaihtelevilla alkuarvauksilla, kunnes sattumalta osutaan nk. *suppenemisaltaaseen*, eli riittävän lähelle jotakin ratkisua. Mikäli iteraatio ei suppene esimerkiksi kymmenen kierroksen jälkeen, vaihdetaan alkuarvausta.

Esimerkki 4 Ratkaise yhtälöryhmä

$$\begin{cases} x^2 + y^2 = 1 \\ e^x + \sin y = 1 \end{cases}$$

<u>Ratkaisu</u> Tässä on

$$\mathbf{f}(x, y) = \begin{bmatrix} x^2 + y^2 - 1 \\ e^x + \sin y - 1 \end{bmatrix}, \qquad \mathbf{J}\,\mathbf{f}(x, y) = \begin{bmatrix} 2x & 2y \\ e^x & \cos y \end{bmatrix},$$

joten Newtonin iteraatio saa muodon

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \begin{bmatrix} 2x_k & 2y_k \\ e^{x_k} & \cos y^k \end{bmatrix}^{-1} \begin{bmatrix} x_k^2 + y_k^2 - 1 \\ e^{x_k} + \sin y_k - 1 \end{bmatrix}.$$

Ratkaisuja on kaksi:

$$\begin{aligned}
A: &\quad (x, y) = (\ \ 0.5538..,\ -0.8327..\,) \\
B: &\quad (x, y) = (-0.8089..,\ \ \ 0.5879..\,)
\end{aligned}$$

Seuraavassa tuloksia alkuarvauksista riippuen. Merkintä * tarkoittaa epäonnistumista (singulaarinen Jacobin matriisi tai iteraation hajaantuminen).

$$\begin{array}{lll}
(\ 1,\ \ 0) \to A & (\ 0\ ,\ \ 0) \to * & (-\ 5,\ \ 0\ ) \to B \\
(\ 1,\ \ 1) \to B & (\ 0.1,\ \ 0) \to A & (-10,\ \ 0\ ) \to B \\
(-1,\ \ 1) \to B & (-0.1,\ \ 0) \to * & (\ \ 3,\ \ 3\ ) \to B \\
(-1,-1) \to B & (\ 5\ ,\ \ 0) \to * & (\ \ 3,-3\ ) \to A \\
(\ 0,-1) \to A & (\ 0\ ,\ \ 5) \to B & (\ \ 3,-1.5) \to A \\
(\ 1,-1) \to A & (\ 0\ ,-5) \to A & (\ \ 3,-1.2) \to B
\end{array} \qquad \square$$

Suuret yhtälöryhmät. Jatkamismenettely

Newtonin menetelmää käytetään yleisesti myös hyvin suurien epälineaaristen yhtälöryhmien ratkaisuun. Tällaiset ongelmat syntyvät usein osana laajempaa laskentaprosessia, jolloin on myös tavallista, että hyvä alkuarvaus on tarjolla laskennan aikaisemmista vaiheista. Näin on esimerkiksi ratkaistaessa differentiaaliyhtälösysteemeitä numeerisilla askelmenetelmillä (ks. Luku IX.7).

Jos halutaan vain ratkaista yksittäinen suurikokoinen yhtälöryhmä ilman ennakkotietoa ratkaisun (ratkaisujen) sijainnista, voi toimivan alkuarvauksen löytäminen Newtonin iteraatiolle olla hyvin vaikeaa. Tällaisessa tilanteessa varsin yleisesti käytetty on nk. *jatkamismenettely* (engl. continuation method), jossa ratkaistava ongelma parametrisoidaan muotoon

$$P(t), \quad 0 \le t < 1.$$

Oletetaan, että $P(1)$ on ongelma, joka varsinaisesti halutaan ratkaista, ja että $P(0)$ on helppo ratkaista. Tällöin jatkamismenettelyn ideana on ratkaista peräkkäin

$$P(t_0), P(t_1), \ldots, P(t_N), \quad t_N = 1,$$

missä (esimerkiksi) $t_k = k\Delta t$. Jos Δt on pieni ja parametrisointi suoritettu hyvin, on probleeman $P(t_{k-1})$ ratkaisu hyvä alkuarvaus Newtonin iteraatiolle, kun ratkaistaan probleemaa $P(t_k)$. Näin voidaan askelittain edetä lopputilaan $t_N = 1$.[†]

[†]Jatkamismenettelyyn perustuvan algoritmin suunnittelussa on yleensä apua probleeman fysikaalisesta taustasta. Esimerkiksi rakenteiden mekaniikan ongelmassa, jossa halutaan simuloida rakenteen suuria muodonmuutoksia kuormituksen alaisena, on luontevaa menetelllä kuten koejärjestelyssä: Parametrisoidaan kuormituksen voimakkuus, eli lähdetään kuormittamattomasta tilasta ja lisätään kuormitusta pienin askelin, kunnes päädytään haluttuun lopputilaan. Parametrin t voi tällöin kuvitella koejärjestelyyn liittyväksi aikamuuttujaksi. Aikamuuttujan lisääminen (aikaparametrisointi) on luonnollista yleisemminkin silloin, kun fysiikasta peräisin olevaan ongelmaan etsitään ajasta riippumatonta (tasapaino)ratkaisua.

Ellei parempaa ideaa ole käytettävissä, voi yleisen yhtälöryhmän parametrisoida vaikkapa muunnoksella muotoa

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \;\hookrightarrow \mathbf{g}(\mathbf{x}) + t\,[\,\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x})\,] = \mathbf{0},$$

missä $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ on jokin helposti ratkeava (esim. lineaarinen) yhtälöryhmä.

Esimerkki 5 Yhtälöryhmälle

$$\begin{cases} x^3 - 2y^3 + x + 2y = 10 \\ x^3 + y^3 - 2xy^2 - 3x - y = 10 \end{cases}$$

antaa Newtonin algoritmi alkuarvauksilla $(x_0, y_0) = (1, 1)$ ja $(x_0, y_0) = (-1, -1)$ ratkaisun $(x, y) = (-2.25872, -2.42834)$. Kokeillaan, löytyykö jokin muu ratkaisu jatkamismenettelyllä: Parametrisoidaan yhtälöryhmä muotoon

$$\mathbf{f}(t, \mathbf{x}) = \begin{bmatrix} t(x^3 - 2y^3) + x + 2y \\ t(x^3 + 2y^3 - 2xy^2) - 3x - y \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

ja ratkaistaan tämä parametrin arvoilla $t_k = 0.2k,\ k = 0 \ldots 5$. Kun $k = 0$, on ratkaisu $(x, y) = (-6, 8)$. Tästä eteenpäin $(k = 1 \ldots 5)$ käytetään edellisellä k:n arvolla saatua ratkaisua alkuarvauksena Newtonin iteraatiolle. Tulos:

t	Ratkaisu
0.2	$(10.5223, 8.55714)$
0.4	$(7.60605, 6.14808)$
0.6	$(6.30833, 5.07754)$
0.8	$(5.53177, 4.43769)$
1.0	$(5.00000, 4.00000)$

Ratkaisu $(x, y) = (5, 4)$ löytyy myös alkuarvauksia vaihtelemalla. Esim. $(x_0, y_0) = (3, 3)$ tai jopa $(x_0, y_0) = (0, 0)$ johtavat tähän. $\square$

Yksinkertaistettu Newtonin iteraatio

Epälineearista yhtälöryhmää Newtonin menetelmällä ratkaistaessa on jokaisella iteraatiokierroksella suoritettava seuraavat laskuoperaatiot:

- Funktioevaluaatio: $\mathbf{x}_k \;\mapsto\; \mathbf{f}(\mathbf{x}_k) = \mathbf{y}_k$.
- Jacobin matriisin *päivitys* (engl. updating) $\mathbf{x}_k \;\mapsto\; \mathbf{Jf}(\mathbf{x}_k) = \mathbf{A}_k$.

- Lineaarisen yhtälöryhmän ratkaisu: $\mathbf{y}_k \mapsto \mathbf{u}_k = \mathbf{A}_k^{-1}\mathbf{y}_k$ $(\mathbf{A}_k\mathbf{u}_k = \mathbf{y}_k)$.

- Vähennyslasku: $\mathbf{x}_k\,, \mathbf{u}_k \;\mapsto\; \mathbf{x}_k - \mathbf{u}_k = \mathbf{x}_{k+1}$.

Sekä Jacobin matriisin päivittäminen että etenkin lineaarisen yhtälöryhmän ratkaiseminen ovat suhteellisen raskaita operaatioita, jos yhtälöryhmän koko on suuri. Hyvin suuria yhtälöryhmiä ratkaistaessa saatetaankin päivittämisessä hieman 'laiskotella', ts. päivitystä ei suoriteta jokaisella iteraatiokierroksella. Tämä helpottaa myös lineaaristen yhtälöryhmien (suoraa) ratkaisemista, koska yhtälöryhmien kerroinmatriisi pysyy samana päivityksien välissä. Riittää silloin laskea Jacobin matriisin LU-hajotelma jokaisen päivityksen jälkeen ja pitää se tallessa seuraavaan päivitykseen asti (vrt. Luku X.3). Riippuen siitä, kuinka usein päivitys suoritetaan, saadaan Newtonin menetelmälle erilaisia variaatioita, joista voidaan valita ongelmakohtaisesti tehokkain.

Em. variaatioista yksinkertaisimmassa lasketaan Jacobin matriisi vain kerran, alkuarvauspisteessä. Kun merkitään $\mathbf{A} = \mathbf{J}(\mathbf{x}_0)$, niin iteraatiokaavaksi tulee

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{A}^{-1}\mathbf{f}(\mathbf{x}_k), \quad k = 0, 1 \ldots$$

Tämä Newtonin menetelmän äärimmäinen yksinkertaistus ei ole aidon Newtonin menetelmän veroinen käytännössä, sillä se suppenee (sikäli kuin suppenee) vain lineaarisesti (vrt. Harj.teht. V.7: 15, kun $n = 1$). — Sen sijaan teoreettisten tarkastelujen 'ajattelumenetelmäksi' tämä menetelmä sopii hyvin, ks. seuraava luku.

Esimerkki 6 Esimerkin 3 tilanteessa on

$$\mathbf{A} = \mathbf{Jf}(-1,1) = \begin{bmatrix} 0 & 10 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{A}^{-1} = \frac{1}{10}\begin{bmatrix} 2 & -10 \\ 1 & 0 \end{bmatrix},$$

joten ym. tavalla yksinkertaistetuksi Newtonin iteraatiokaavaksi tulee

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \frac{1}{10}\begin{bmatrix} 2 & -10 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_k^3 - 2x_k y_k^5 - x_k - 2.04 \\ -x_k^3 y_k + 2x_k y_k^2 + y_k^5 - 0.05 \end{bmatrix}.$$

Iteraatio suppenee, mutta selvästi aitoa Newtonin iteraatiota hitaammin (vrt.

Esimerkki 3):

$$\begin{aligned}
(x_0, y_0) &= (-1,\ 1) \\
(x_1, y_1) &= (-0.942, 1.004) \\
(x_2, y_2) &= (-0.929318, 1.005191) \\
(x_3, y_3) &= (-0.923142, 1.005780) \\
(x_4, y_4) &= (-0.919696, 1.006109) \\
(x_5, y_5) &= (-0.917651, 1.006304) \\
&\vdots \quad \square
\end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske seuraavissa tapauksissa $\mathbf{f}$:n Jakobin matriisi ensimmäisessä annetussa pisteessä, approksimoi $\mathbf{f}$ affiinikuvauksella ja laske tämän avulla $\mathbf{f}$:n arvo likimäärin toisessa annetussa pisteessä. Vertaa tarkkaan arvoon.
   a) $\mathbf{f}(x, y) = (x^3 + y^3, x^3 - y^3)$, $(1, 1)$, $(0.9, 1.1)$
   b) $\mathbf{f}(x, y) = (e^x \sin \pi y, e^{-x} \cos \pi y)$, $(0, 1/3)$, $(-0.04, 0.30)$
   c) $\mathbf{f}(x, y, z) = (x^2 y, x^2 z, y^2 - z^2)$, $(1, 3, 3)$ $(0.99, , 3.02, 2.97)$
   d) $\mathbf{f}(x, y, z) = (x^2 + yz, y^2 - x \ln z)$, $(2, 2, 1)$, $(1.98, 2.01, 1.03)$
   e) $\mathbf{f}(\mathbf{x}) = (x_1 x_2,\ x_1^2 - x_3^2,\ x_2^2 + x_3 x_4,\ x_2 x_4 + x_4^2)$, $(1, 2, -2, -1)$, $(1.01, 2.02, -2.02, -1.01)$

2. Yhtälöryhmällä

$$[\mathbf{f}(x, y, z)]^T = [\, x^3 - x^2 y + x z^3,\ xyz + z^2,\ x^3 + y^3 + z^3 \,] = \mathbf{b}^T$$

   on muuan helppo ratkaisu, kun $\mathbf{b}^T = [1, 2, 3]$. Approksimoimalla $\mathbf{f}$ affiinikuvauksella ko. ratkaisupisteen lähellä laske yhtälöryhmän ratkaisu likimäärin, kun $\mathbf{b} = [1.04, 1.98, 3.02]^T$.

3. Laske koordinaattimuunnoksien $(r, \varphi, z) \mapsto (x, y, z)$ (lieriök. $\hookrightarrow$ karteesinen) ja $(r, \theta, \varphi) \mapsto (x, y, z)$ (pallok. $\hookrightarrow$ karteesinen) Jacobin matriisit $\mathbf{J}(r, \varphi, z)$ ja $\mathbf{J}(r, \theta, \varphi)$. Missä pisteissä $\mathbf{J}$ on singulaarinen matriisi?

4. Olkoon $\mathbf{f}$ tyyppiä $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^p$ ja $\mathbf{g}$ tyyppiä $\mathbf{g} : \mathbb{R}^p \to \mathbb{R}^m$. Johda yhdistetyn funktion $\mathbf{F}(\mathbf{x}) = (\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}[\mathbf{f}(\mathbf{x})]$ Jacobin matriisin laskusääntö $\mathbf{JF}(\mathbf{x}) = \mathbf{Jg}[\mathbf{f}(\mathbf{x})]\mathbf{Jf}(\mathbf{x})$. *Vihje*: Ketjusääntö!

5. Etsi seuraaville yhtälöryhmille ratkaisu Newtonin menetelmällä lähtien alkuarvauksesta $(x_0, y_0) = (1, 1)$.

$$\text{a)} \begin{cases} x^4 + y^4 = 2xy^5 \\ x + x^2 + y^4 = 4 \end{cases} \qquad \text{b)} \begin{cases} x^3 + 3x = y^4 + 4y \\ \cos \pi x + xye^{-y} = 0 \end{cases}$$

6. Etsi seuraavien yhtälöryhmien ratkaisut neljän merkitsevän numeron tarkkuudella käyttäen kaksiulotteista Newtonin menetelmää. Laske a)- ja b)-kohdissa ratkaisu myös vaihtoehtoisella tavalla, jossa eliminoidaan ensin toinen tuntemattomista ja käytetään 1-ulotteista Newtonin menetelmää.
   a) $y - e^x = 0, \;\; x - \sin y = 0$
   b) $x^2 + y^2 = 1, \;\; y = e^x$ (kaksi ratkaisua)
   c) $y - \sin x = 0, \;\; x^2 + (y+1)^2 = 2$ (kaksi ratkaisua)
   d) $x^2 - xy + 2y^2 = 10, \;\; x^3y^2 = 2$ (neljä ratkaisua)
   e) $\sin x + \sin y = 1, \;\; x^3 = y^2 < 30$ (neljä ratkaisua)

7. Selvitä (kaksiulotteista) Newtonin menetelmää käyttäen, missä pisteissä Cartesiuksen lehden $x^3 + y^3 = 3xy$ tangentin ja jonkin koordinaatiakselin välinen kulma on 45°.

8. Määritä likimäärin seuraavien yhtälöryhmien ratkaisut eliminoimalla ensin yksi tuntemattomista ja käyttämällä sen jälkeen Newtonin menetelmää.

$$\text{a)} \begin{cases} x^2 + y^2 + z^2 = 1 \\ z = xy \\ 6xz = 1 \end{cases} \qquad \text{b)} \begin{cases} x^2 + y^2 + z^2 = 1 \\ y = \sin z \\ z + z^2 + z^3 = x + y \end{cases} \qquad \text{c)} \begin{cases} x^2 + y^4 = 1 \\ z = x^3 y \\ e^x = 2y - z \end{cases}$$

9. Yhtälöryhmällä

$$x + yze^x \;=\; x^3 + y^3 z \;=\; y + e^{xyz} \;=\; 2.1$$

   on ratkaisu pisteen $(0, 1, 2)$ lähellä. Iteroi tästä alkuarvauksesta kolme kertaa Newtonin menetelmän yksinkertaistetulla versiolla, jossa Jacobin matriisia ei päivitetä. Vertaa oikean Newtonin iteraation antamiin tuloksiin.

10. (*) Etsi jatkamismenettelyä ja Newtonin menetelmää käyttäen jokin ratkaisu yhtälöryhmälle

$$\begin{cases} 3x^3 - y^3 + z^3 + x - 2z + 40 = 0 \\ 4x^3 - 3y^3 - z^3 - 2x - 3y - 40 = 0 \\ x^3 + y^3 + z^3 - 5x - 5y + 5z - 40 = 0 \end{cases}$$

11. (*) Määritä suora, joka sivuaa kahdessa eri pisteessä käyrää

$$K: \; x^2y^2 - 2xy^3 + y^4 - 4x^2y + 6xy^2 - 2y^3 + 4x^2 - 3y^2 - 12x + y = 20.$$

12. (*) (Avaruuspysäköinti) Tähtien välisen avaruden pisteessä $P = (x, y, z)$ (pituusyksikkö = valovuosi) on avaruusasema. Avaruusasemaan vaikuttaa lähimpien tähtien vetovoima $\vec{G} = -k\nabla u$, missä k on vakio ja

$$u(x, y, z) = \sum_{i=1}^{n} \frac{m_i}{|P - P_i|},$$

missä m_i on tähden T_i massa ja P_i sijainti. Määritä avaruuspysäköintiä varten kaikki pisteet P, joissa $\vec{G} = \vec{0}$, kun $m_i = iM$, $i = 1 \ldots n$ (M = vakio) ja

a) $n = 3$, $P_1 = (0, 0, 0)$, $P_2 = (1, 0, 0)$, $P_3 = (0, 1, 0)$,
b) $n = 4$, $P_1 = (0, 0, 0)$, $P_2 = (1, 0, 0)$, $P_3 = (0, 1, 0)$, $P_4 = (0, 0, 1)$.

XI.7 Käänteisfunktiolause. Implisiittifunktiolause. Kontraktiokuvauslause

Tässä luvussa tarkastellaan peruskysymystä epälineaarisen yhtälöryhmän ratkeavuudesta, kun yhtälöitä ja tuntemattomia on yhtä monta. Olkoon $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$, $\mathbf{x} \in \mathcal{D}_\mathbf{f} \subset \mathbb{R}^n$ ja tarkastellaan yhtälöryhmää

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}, \tag{$\star$}$$

missä $\mathbf{y} \in \mathbb{R}^n$. Sikäli kuin yhtälöryhmällä on yksikäsitteinen ratkaisu $\mathbf{x} \in \mathcal{D}_\mathbf{f}$ jollakin $\mathbf{y}$, voidaan ratkaisu muodollisesti kirjoittaa

$$\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y}),$$

missä $\mathbf{f}^{-1}(\mathbf{y}) = (g_1(\mathbf{y}), \ldots, g_n(\mathbf{y}))$ on $\mathbf{f}$:n *käänteisfunktio.* Jos yhtälöryhmällä on yksikäsitteinen jokaisella $\mathbf{y} \in B$ ($B \subset \mathbb{R}^n$) ja $A = \mathbf{f}^{-1}(B) = \{\mathbf{x} \in \mathcal{D}_\mathbf{f} \mid \mathbf{f}(\mathbf{x}) \in B\}$, niin $\mathbf{f} : A \to B$ on bijektio.

Esimerkki 1 Olkoon $\mathbf{f}(x, y) = (x+y, x^2-y^2)$, $\mathbf{x} = (x, y) \in \mathbb{R}^2$. Kun merkitään $\mathbf{y} = (u, v)$, niin yhtälöryhmä ($\star$) on

$$\begin{cases} x + y = u, \\ x^2 - y^2 = v. \end{cases}$$

Jälkimmäisen yhtälön muodosta $u(x - y) = v$ nähdään, että jos $u \neq 0$, niin yhtälöryhmällä on yksikäsitteinen ratkaisu $x = \frac{1}{2}(u^2 + v)/u$, $y = \frac{1}{2}(u^2 - v)/u$. Suoralla $S : x + y = 0$ on $\mathbf{f}(x, y) = (0, 0)$, joten tällä suoralla $\mathbf{f}$ ei ole 1-1. Päätellään, että jos $A = \{(x, y) \in \mathbb{R}^2 \mid x + y \neq 0\}$ ja $B = \{(u, v) \in \mathbb{R}^2 \mid u \neq 0\}$, niin $\mathbf{f} : A \to B$ on bijektio ja

$$\mathbf{f}^{-1}(\mathbf{y}) = \left(\frac{u^2+v}{2u}, \frac{u^2-v}{2u}\right), \quad \mathbf{y} = (u, v) \in B. \quad \square$$

Paikallinen käänteisfunktio. Käänteisfunktiolause

Esimerkissä 1 yhtälöryhmä ($\star$) ratkesi täydellisesti. Tavallisemmin epälineaarisen yhtälöryhmän täydellinen ratkaiseminen on vaikeaa ellei mahdotonta, eikä myöskään mitään lineaaristen yhtälöryhmien teoriaan verrattavaa yleisempää ratkeavuusteoriaa ole. Sen sijaan voidaan suhteellisen yleisin ehdoin selvittää kysymys yhtälöryhmän paikallisesta ratkeavuudesta seuraavan (sovelluksissa tyypillisen, vrt. edellinen luku) ongelman asettelun mukaisesti: Jos $\mathbf{a} \in \mathcal{D}_\mathbf{f}$ ja $\mathbf{b} = \mathbf{f}(\mathbf{a})$, niin onko (ja millä ehdoilla) yhtälöryhmällä ($\star$) yksikäsitteinen ratkaisu lähellä $\mathbf{a}$:ta, kun $\mathbf{y}$ on samoin lähellä $\mathbf{b}$:tä?

MÄÄRITELMÄ XI.7.1 Funktio $\mathbf{f} : \mathcal{D}_f \to \mathbb{R}^n$, $\mathcal{D}_\mathbf{f} \subset \mathbb{R}^n$, on *paikallisesti kääntyvä* (engl. locally invertible) pisteessä $\mathbf{a} \in D_\mathbf{f}$, jos $\exists\delta > 0$ siten, että $\mathbf{f}$ on määritelty ympäristössä $U_\delta(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x}-\mathbf{a}| < \delta\}$ ja on tähän ympäristöön rajoitettuna kääntäen yksikäsitteinen kuvaus. Käänteisfunktiota

$$\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y}) \;\Leftrightarrow\; \mathbf{y} = \mathbf{f}(\mathbf{x}) \;\;\&\;\; \mathbf{x} \in U_\delta(\mathbf{a})$$

sanotaan $\mathbf{f}$:n *paikalliseksi* (lokaaliksi) *käänteisfunktioksi.*

ESIMERKKI 1 (jatko) Ratkaisemalla yhtälöryhmä $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ todettiin esimerkin funktio $\mathbf{f}$ kääntyväksi rajoitettuna joukkoon $A = \{(x, y) \in \mathbb{R}^2 \mid x + y \neq 0\}$. Myös Määritelmän XI.7.1 mukaisesti $\mathbf{f}$ on paikallisesti kääntyvä jokaisessa pisteessä $\mathbf{a} = (a, b) \in A$, sillä A on avoin joukko, jolloin $\forall(a, b) \in A\ \exists\delta > 0$ siten, että $U_\delta(a, b) \subset A$, jolloin $\mathbf{f}$ on tähän ympäristöön rajoitettuna 1–1. (Tarkemmin pätee tässä: $U_\delta(a, b) \subset A$ täsmälleen kun $\delta \leq d$, missä $d = |a + b|/\sqrt{2}$, on pisteen $P = (a, b)$ etäisyys suorasta $S : x + y = 0$.) $\square$

Yhden reaalimuuttujan funktion tapauksessa paikalliselle kääntyvyydelle voidaan helposti antaa riittävät ehdot derivaatan avulla. Oletetaan:

(i) f on derivoituva välillä $(a - \delta, a + \delta)$ jollakin $\delta > 0$.

(ii) $f'(x)$ on jatkuva pisteesä $x = a$.

(iii) $f'(a) \neq 0$.

Näiden oletuksien perusteella on joko $f'(x) > 0$ tai $f'(x) < 0$ välillä $[a - \rho, a + \rho]$ jollakin $\rho \in (0, \delta)$ (Propositio V.1.4), jolloin f on ko. välillä aidosti monotoninen (Lause V.6.3) ja siis 1-1. Ehdot (i)–(iii) ovat siis riittävät f:n paikalliselle kääntyvyydelle pisteessä a. Koska f on näillä ehdoilla sekä aidosti monotoninen että jatkuva välillä $[a - \rho, a + \rho]$, niin päätellään myös (Lause V.1.10): Yhtälöllä $f(x) = y$ on yksikäsitteinen ratkaisu välillä $[a - \rho, a + \rho]$ aina kun $|y - f(a)| \leq \varepsilon$, missä $\varepsilon = \min\{\,|f(a) - f(a - \rho)|, |f(a) - f(a + \rho)|\,\} > 0$. Edelleen ehdoilla (i)—(iii) pätee myös (Lause V.3.3, $\delta = \rho$): f:n paikallinen käänteisfunktio f^{-1} on derivoituva pisteessä $b = f(a)$ ja $\mathrm{D}f^{-1}(b) = 1/f'(a)$.

ESIMERKKI 2 Funktio $f(x) = x^2$ täyttää oletukset (i)–(iii) pisteissä $a \neq 0$, joten näissä pisteissä f on paikallisesti kääntyvä. — Tämä on selvää muutenkin, sillä f tiedetään kääntyväksi väleillä $(-\infty, 0]$ ja $[0, \infty)$. Välillä $(-\delta, \delta)$ ei f ole 1-1 millään $\delta > 0$, joten f ei ole paikallisesti kääntyvä pisteessä $a = 0$. Funktio $f(x) = x^3$ on sen sijaan 1-1 välillä $(-\infty, \infty)$, joten f on paikallisesti kääntyvä jokaisessa pisteessä, myös pisteessä $a = 0$, jossa $f'(a) = 0$. $\square$

Esimerkistä nähdään, että tapaus $f'(a) = 0$ on f:n paikallisen kääntyvyyden kannalta avoin tapaus, eli f voi olla paikallisesti kääntyvä pisteessä a tai ei.

Edellä mainituilla väittämillä, perustuen oletuksiin (i)–(iii), on myös useamman muuttujan funktioita koskevat vastineet, jotka seuraavassa kootaan yhdeksi lauseeksi. Tämän *Käänteisfunktiolauseen* todistus nojaa keskeisesti toiseen huomattavaan lauseeseen, *Kontraktiokuvauslauseeseen*, joka esitetään ja todistetaan jäljempänä. Käänteiskuvauslauseen todistus esitetään vasta luvun lopussa.

*Lause XI.7.2 **(Käänteisfunktiolause)** Oletetaan, että funktio $\mathbf{f} : \mathcal{D}_\mathbf{f} \to \mathbb{R}^n$, $\mathcal{D}_\mathbf{f} \subset \mathbb{R}^n$ täyttää ehdot: (i) $\exists\delta > 0$ siten, että f on differentioituva pisteen $\mathbf{a}$ ympäristössä $U_\delta(\mathbf{a}) \subset \mathcal{D}_\mathbf{f}$. (ii) Osittaisderivaatat $\partial f_i(\mathbf{x})/\partial x_j$, $i, j = 1 \ldots n$ ovat jatkuvia pisteessä $\mathbf{x} = \mathbf{a}$. (iii) Jacobin matriisi $\mathbf{Jf}(\mathbf{x}) = (\partial f_i(\mathbf{x})/\partial x_j)$ on säännöllinen pisteessä $\mathbf{x} = \mathbf{a}$. Tällöin on olemassa $\rho \in (0, \delta)$ ja $\varepsilon > 0$ siten, että $\mathbf{f}$ on 1-1 joukossa $K = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{a}| \le \rho\}$ ja yhtälöryhmällä $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ on ratkaisu $\mathbf{x} \in K$ aina kun $|\mathbf{y} - \mathbf{b}| \le \varepsilon$, missä $\mathbf{b} = f(\mathbf{a})$. Lisäksi $\mathbf{f}$:n paikallinen käänteisfunktio $\mathbf{f}^{-1}$ on differentioituva pisteessä $\mathbf{b}$ ja $\mathbf{J}\,\mathbf{f}^{-1}(\mathbf{b}) = [\mathbf{J}\,\mathbf{f}(\mathbf{a})]^{-1}$.

Kuviossa on joukot $B = \{\mathbf{y} \in \mathbb{R}^n \mid |\mathbf{y} - \mathbf{b}| \le \varepsilon\}$ ja $\mathbf{f}^{-1}(B)$ rajattu yhtenäisellä viivalla ja joukot K ja $\mathbf{f}(K)$ pisteviivalla. Lauseen väittämän mukaisesti paikallinen käänteisfunktio $\mathbf{f}^{-1}$ on määritelty B:ssä ja $\mathbf{f}^{-1}(B) \subset K$.

Esimerkki 1 (jatko) Esimerkissä $\mathbf{f}$:n paikallinen kääntyvyys voidaan selvittää myös ratkaisematta yhtälöryhmää $(\star)$: Lauseen XI.7.2 mukaan $\mathbf{f}$ on paikallisesti kääntyvä pisteessä (a, b) aina kun $a + b \neq 0$, sillä

$$\mathbf{Jf}(a,b) = \begin{bmatrix} 1 & 1 \\ 2a & -2b \end{bmatrix} \Rightarrow \quad \det \mathbf{J}(a,b) = -2(a+b) \neq 0, \quad \text{kun } a + b \neq 0. \quad \square$$

Esimerkki 3 (Vrt. Esimerkki 2 edellisessä luvussa.) Käänteisfunktiolauseen mukaan yhtälöryhmällä

$$\mathbf{f}(x,y) = \mathbf{y} \;\Leftrightarrow\; \begin{cases} x^3 - 2xy^5 - x = u \\ -x^3y + 2xy^2 + y^5 = v \end{cases}$$

on pisteen $\mathbf{a} = (2,1)$ lähiympäristössä yksikäsitteinen ratkaisu, kun $|\mathbf{y} - \mathbf{b}| \le \varepsilon$, missä $\mathbf{b} = (2,3)$ ja $\varepsilon > 0$ on riittävän pieni. Pisteen $\mathbf{x} = \mathbf{a}$ lähiympäristössä on likimain

$$\mathbf{f}(\mathbf{x}) \approx \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 9 & -20 \\ -10 & 5 \end{bmatrix} \begin{bmatrix} x-2 \\ y-1 \end{bmatrix},$$

joten ehdot $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ ja $\mathbf{y} \in B : |\mathbf{y} - \mathbf{b}| \le \varepsilon \Leftrightarrow (u-2)^2 + (v-3)^2 \le \varepsilon^2$ vastaavat pisteen $\mathbf{x} = \mathbf{a}$ lähellä likimain ehtoa

$$\begin{aligned} &[9(x-2) - 20(y-1)]^2 + [-10(x-2) + 5(y-1)]^2 \le \varepsilon^2 \\ \Leftrightarrow\quad &181(x-2)^2 - 460(x-2)(y-1) + 425(y-1)^2 \le \varepsilon^2. \end{aligned}$$

Tämän mukaan joukkoa $\mathbf{f}^{-1}(B)$ rajaa likimain toisen asteen käyrä — Kyseessä on (vino) *ellipsi*, jonka keskipiste on $\mathbf{a} = (2,1)$, vrt. kuvio edellä. $\square$

Implisiittifunktiolause

Käänteisfunktiolauseen hieman yleisempi muoto on *Implisiittifunktiolause*, jossa tarkastellaan yhtälöryhmää muotoa

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{0},$$

missä $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^p$ ja $\mathbf{F} : D_{\mathbf{F}} \to \mathbb{R}^p$, $D_{\mathbf{F}} \subset \mathbb{R}^{n+p}$. Olkoon annettu piste $(\mathbf{x},\mathbf{y}) = (\mathbf{a},\mathbf{b})$, jossa $\mathbf{F}(\mathbf{a},\mathbf{b}) = \mathbf{0}$. Kysytään: Voidaanko $\mathbf{y}$ ratkaista yhtälöryhmästä muodossa $\mathbf{y} = \mathbf{g}(\mathbf{x})$ pisteen $(\mathbf{a},\mathbf{b})$ ympäristössä? — Riittävä ehto ratkeavuudelle saadaan jälleen tutkimalla $\mathbf{F}$:n Jacobin matriisia

$$\mathbf{J}\,\mathbf{F}(\mathbf{x},\mathbf{y}) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} & \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_p} \\ \vdots & & & & & \\ \dfrac{\partial F_p}{\partial x_1} & \cdots & \dfrac{\partial F_p}{\partial x_n} & \dfrac{\partial F_p}{\partial y_1} & \cdots & \dfrac{\partial F_p}{\partial y_p} \end{bmatrix}.$$

Kun kirjoitetaan tämä muotoon

$$\mathbf{J}\,\mathbf{F}(\mathbf{x},\mathbf{y}) = [\mathbf{J}_{\mathbf{x}}\mathbf{F}(\mathbf{x},\mathbf{y}), \mathbf{J}_{\mathbf{y}}\mathbf{F}(\mathbf{x},\mathbf{y})],$$

missä $\mathbf{J_x F}$ sisältää $\mathbf{J}\,\mathbf{F}$:n ensimmäiset n saraketta, niin ratkeavuuden kannalta kriittinen on matriisi $\mathbf{J_y F}(\mathbf{a},\mathbf{b})$ (kokoa $p\times p$). Jos tämä on säännöllinen, niin ratkaiseminen yleensä onnistuu. Tarkemmin muotoiltuna tämä väittämä on mukaelma Lauseesta XI.7.2, ja myös todistus (jota ei esitetä) noudattaa tämän lauseen todistuksen ajatuskulkua, ks. luvun loppu

*Lause XI.7.3 **(Implisiittifunktiolause)** Oletetaan, että funktio $\mathbf{F}(\mathbf{x},\mathbf{y})$, missä $\mathbf{x}\in\mathbb{R}^n$, $\mathbf{y}\in\mathbb{R}^p$ ja $\mathbf{F}:\mathcal{D}_\mathbf{F}\to\mathbb{R}^p$, $\mathcal{D}_\mathbf{F}\subset\mathbb{R}^{n+p}$, täyttää ehdot: (i) $\exists\delta>0$ siten, että $\mathbf{F}$ on differentioituva pisteen $(\mathbf{a},\mathbf{b})$ ympäristössä $U_\delta(\mathbf{a},\mathbf{b})\subset D_\mathbf{F}$. (ii) Osittaisderivaatat $[\mathbf{J_x F}(\mathbf{x},\mathbf{y})]_{ij}=\partial F_i(\mathbf{x},\mathbf{y})/\partial x_j,\ i=1\ldots p,\ j=1\ldots n$ ja $[\mathbf{J_y F}(\mathbf{x},\mathbf{y})]_{ij}=\partial F_i(\mathbf{x},\mathbf{y})/\partial y_j,\ i,j=1\ldots p$ ovat jatkuvia pisteessä $(\mathbf{x},\mathbf{y})=(\mathbf{a},\mathbf{b})$. (iii) Matriisi $\mathbf{J_y F}(\mathbf{a},\mathbf{b})$ on säännöllinen. Tällöin on olemassa $\rho\in(0,\delta)$ ja $\varepsilon>0$ siten, että yhtälöryhmä $\mathbf{F}(\mathbf{x},\mathbf{y})=\mathbf{0}$ ratkeaa ehdolla $|\mathbf{y}-\mathbf{b}|\le\rho$ yksikäsitteisesti muotoon $\mathbf{y}=\mathbf{g}(\mathbf{x})$ aina kun $|\mathbf{x}-\mathbf{a}|\le\varepsilon$. Lisäksi $\mathbf{g}$ on differentioituva pisteessä $\mathbf{a}$ ja $\mathbf{Jg}(\mathbf{a})=-[\mathbf{J_y F}(\mathbf{a},\mathbf{b})]^{-1}\mathbf{J_x F}(\mathbf{a},\mathbf{b})$.

Lauseen XI.7.3 erikoistapaaus on Lause XI.7.2, kun $\mathbf{F}(\mathbf{x},\mathbf{y})=\mathbf{x}-\mathbf{f}(\mathbf{y})$ $(p=n)$. Jos $\mathbf{g}$:n differentioituvuus oletetaan, niin Jacobin matriisin $\mathbf{Jg}(\mathbf{a})$ laskusäännön voi johtaa implisiittisellä osittaisderivoinnilla (Harj.teht. 7).

Esimerkki 4 Missä tasokäyrän $S:\ F(x,y)=x^2-2xy-y^2-1=0$ pisteissä voidaan käyrän yhtälöstä ratkaista paikallisesti a) y x:n avulla, b) x y:n avulla?

<u>Ratkaisu</u> a) Ratkeavuusehto on $F_y(x,y)=-2x-2y\neq 0\ \Leftrightarrow\ x+y\neq 0$. Koska $x+y=0\ \Rightarrow\ F(x,y)=2x^2-1$, niin poikkeuspisteitä ovat ainoastaan $(1/\sqrt{2},-1/\sqrt{2})$ ja $(-1/\sqrt{2},1/\sqrt{2})$. Muissa pisteissä ratkaiseminen onnistuu.

b) Ratkeavuusehto on $F_x(x,y)=2x-2y\neq 0\ \Leftrightarrow\ y\neq x$. Tämä ehto toteutuu kaikissa käyrän pisteissä (koska $y=x\ \Rightarrow\ F(x,y)=-2x^2-1<0$), joten myös ratkaiseminen onnistuu kaikissa pisteissä. $\square$

Esimerkki 5 Yhtälöryhmä

$$\begin{cases} 2x^2+xy-yz=0\\ 2xz-2y^2+z^2=0\end{cases}$$

ratkaistaan pisteen $(x,y,z)=(1,2,2)$ ympäristössä muotoon $y=y(x),\ z=z(x)$. Laske $y'(1)$ ja $z'(1)$.

<u>Ratkaisu</u> Pisteessä $(x,y,z)=(1,2,2)$ on

$$\mathbf{J}_{y,z}\mathbf{F}(x,y,z)=\begin{bmatrix}\partial_y F_1 & \partial_z F_1\\ \partial_y F_2 & \partial_z F_2\end{bmatrix}=\begin{bmatrix}x-z & -y\\ -4y & 2x+2z\end{bmatrix}=\begin{bmatrix}-1 & -2\\ -8 & 6\end{bmatrix}.$$

Tämä on säännöllinen matriisi, joten ratkaiseminen onnistuu väitetyllä tavalla. Koska $\partial_x F_1 = 4x + y = 6$ ja $\partial_x F_2 = 2z = 4$, kun $(x, y, z) = (1, 2, 2)$, niin saadaan

$$\mathbf{Jg}(1) = \begin{bmatrix} y'(1) \\ z'(1) \end{bmatrix} = -\begin{bmatrix} -1 & -2 \\ -8 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Samaan tulokseen tullaan, kun derivoidaan implisiittisesti yhtälöt

$$2x^2 + xy(x) - y(x)z(x) = 0, \quad 2xz(x) - 2[y(x)]^2 + [z(x)]^2 = 0,$$

asetetaan $(x, y, z) = (1, 2, 2)$ ja ratkaistaan derivaatat $y'(1)$ ja $z'(1)$ syntyvästä (lineaarisesta) yhtälöryhmästä. $\square$

*Kontraktiokuvauslause

Yhtälöryhmän $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ paikallisessa ratkeavuusteoriassa lähtökohta on sama kuin ratkaisua käytännössä etsittäessä, eli kirjoitetaan yhtälöryhmä muotoon $\mathbf{x} = \mathbf{F}(\mathbf{x})$, jolloin kyse on siitä, onko $\mathbf{F}$:llä yksikäsitteinen *kiintopiste* $\mathbf{a}$:n lähellä. Kuten arvata saattaa, $\mathbf{F}$ pyritään valitsemaan niin, että kiintopiste löytyy kiintopisteiteraatiolla. Iteraation suppenemiselle, ja samalla kiintopisteen yksikäsitteisyydelle annetun pisteen $\mathbf{a}$ lähiympäristössä, antaa takeet

*Lause XI.7.4 (**Kontraktiokuvauslause**[†]) Olkoon $\mathbf{F} : \mathcal{D}_{\mathbf{F}} \to \mathbb{R}^n$, $\mathcal{D}_{\mathbf{F}} \subset \mathbb{R}^n$ ja $K = \{\, \mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{a}| \le \rho \,\} \subset \mathcal{D}_{\mathbf{F}}$ jollakin $\mathbf{a} \in \mathbb{R}^n$ ja $\rho > 0$. Tällöin jos

(i) $\mathbf{F}$ on joukkoon K rajoitettuna *kontraktio(kuvaus)*, eli

$$|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})| \le L|\mathbf{x} - \mathbf{y}|, \quad \mathbf{x}, \mathbf{y} \in K, \quad \text{missä } L < 1,$$

(ii) $\mathbf{F}(K) \subset K$,

niin

(1) $\mathbf{F}$:llä on K:ssa täsmälleen yksi kiintopiste $\mathbf{c}$.

(2) Kiintopisteiteraatio $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$, $k = 0, 1, \ldots$ suppenee kohti $\mathbf{c}$:tä jokaisella $\mathbf{x}_0 \in K$.

[†]Kontraktiokuvauslause tunnetaan myös nimellä *Banachin kiintopistelause*, syystä että lause on pätevä euklidisia avaruuksia $\mathbb{R}^n$ yleisemmissä *Banach-avaruuksissa*. Hyvin monet matemaattisten yhtälöiden ratkavuutta koskevat väittämät nojaavat Kontraktiokuvauslauseeseen sen yleisemmissä muodoissa. Tällainen on esimerkiksi differentiaaliyhtälöiden ratkeavuutta koskeva Picardin-Lindelöfin lause (Luku IX.8).

Todistus Ensinnäkin oletuksesta (ii) seuraa, että kiintopisteiteraatiolle $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$ pätee

$$\mathbf{x}_0 \in K \;\Rightarrow\; \mathbf{x}_k \in K \quad \forall k,$$

joten oletuksen (i) perusteella

$$\begin{aligned} |\mathbf{x}_{k+1} - \mathbf{x}_k| &= |\mathbf{F}(\mathbf{x}_k) - \mathbf{F}(\mathbf{x}_{k-1})| \\ &\le L|\mathbf{x}_k - \mathbf{x}_{k-1}| \\ &\;\;\vdots \\ &\le L^k|\mathbf{x}_1 - \mathbf{x}_0|. \end{aligned}$$

Tästä seuraa, että jos $l > k$, niin

$$\begin{aligned} |\mathbf{x}_l - \mathbf{x}_k| &\le |\mathbf{x}_l - \mathbf{x}_{l-1}| + \ldots + |\mathbf{x}_{k+1} - \mathbf{x}_k| \\ &\le \sum_{i=k}^{l-1} L^i|\mathbf{x}_1 - \mathbf{x}_0| \\ &\le |\mathbf{x}_1 - \mathbf{x}_0| \sum_{i=k}^{\infty} L^i \\ &= L^k(1-L)^{-1}|\mathbf{x}_1 - \mathbf{x}_0|. \end{aligned}$$

Siis yleisemmin

$$|\mathbf{x}_k - \mathbf{x}_l| < L^{\min\{k,l\}}(1-L)^{-1}|\mathbf{x}_1 - \mathbf{x}_0|.$$

Tämän mukaan $\{\mathbf{x}_k\}$ on $\mathbb{R}^n$*:n Cauchyn jono*, ts.

$$|\mathbf{x}_k - \mathbf{x}_l| \to 0, \quad \text{kun} \;\; \min\{k,l\} \to \infty.$$

Kun tässä merkitään $\mathbf{x}_k = ((\mathbf{x}_k)_i,\; i = 1 \ldots n)$, niin seuraa

$$|(\mathbf{x}_k)_i - (\mathbf{x}_l)_i| \le |\mathbf{x}_k - \mathbf{x}_l| \to 0, \quad \text{kun} \;\; \min\{k,l\} \to \infty, \quad i = 1 \ldots n.$$

Siis reaalilukujono $\{(\mathbf{x}_k)_i,\; k = 0, 1, \ldots\}$ on Cauchyn jono jokaisella i. Cauchyn kriteerin (Lause I.10.5) mukaan on tällöin olemassa reaaliluvut $c_i,\; i = 1 \ldots n$ siten, että $\lim_k (\mathbf{x}_k)_i = c_i,\; i = 1 \ldots n$, jolloin on myös $\lim_k |\mathbf{c}_k - \mathbf{c}| = 0$, $\mathbf{c} = (c_i)$ (vrt. Luku XI.1). On siis päätelty, että jono $\{\mathbf{x}_k\}$ (ja yleisemminkin jokainen $\mathbb{R}^n$:n Cauchyn jono) suppenee:

$$\mathbf{x}_k \to \mathbf{c} \in \mathbb{R}^n.$$

Mutta koska K on määritelmänsä perusteella suljettu (itse asiassa kompakti, ks. Määritelmä XI.1.8), niin $\mathbf{c} \in K$. Tällöin koska $\mathbf{F}$ on oletuksen (i) perusteella K:ssa jatkuva, niin seuraa $\mathbf{F}(x_k) \to \mathbf{F}(\mathbf{c})$, jolloin iteraatiokaavan $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$

perusteella on $\mathbf{c} = \mathbf{F}(\mathbf{c})$, eli $\mathbf{c}$ on kiintopiste. On siis osoitettu, että ainakin yksi kiintopiste $\mathbf{c} \in K$ on olemassa, ja että kiintopisteiteraatio suppenee jokaisella $\mathbf{x}_0 \in K$ kohti jotakin K:ssa olevaa kiintopistettä. Lopuksi näytetään, että kiintopisteitä on K:ssa vain yksi: Olkoot $\mathbf{c}_1$ ja $\mathbf{c}_2$ $\mathbf{F}$:n kiintopisteitä K:ssa, eli olkoon $\mathbf{c}_1 = \mathbf{F}(\mathbf{c}_1)$, $\mathbf{c}_2 = \mathbf{F}(\mathbf{c}_2)$ ja $\mathbf{c}_1, \mathbf{c}_2 \in K$. Oletuksesta (i) seuraa silloin

$$|\mathbf{c}_1 - \mathbf{c}_2| = |\mathbf{F}(\mathbf{c}_1) - \mathbf{F}(\mathbf{c}_2)| \le L|\mathbf{c}_1 - \mathbf{c}_2| \quad \Leftrightarrow \quad (1-L)|\mathbf{c}_1 - \mathbf{c}_2| \le 0.$$

Koska tässä on $1 - L > 0$ (oletus (i)), niin on oltava $|\mathbf{c}_1 - \mathbf{c}_2| = 0 \;\Leftrightarrow\; \mathbf{c}_1 = \mathbf{c}_2$. Kontraktiokuvauslause on näin todistettu. $\square$

*Käänteisfunktiolauseen todistus

Lähdetään differentiaalikehitelmästä

$$\mathbf{f}(\mathbf{x}) = \mathbf{b} + \mathbf{J}\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}) + \mathbf{r}(\mathbf{x}), \quad \mathbf{b} = \mathbf{f}(\mathbf{a}), \quad \mathbf{x} \in K, \tag{a}$$

missä $K = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x} - \mathbf{a}| \le \rho\}$, $\rho \in (0, \delta)$. Tämän perusteella on

$$\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2) = \mathbf{J}\mathbf{f}(\mathbf{a})(\mathbf{x}_1 - \mathbf{x}_2) + \mathbf{r}(\mathbf{x}_1) - \mathbf{r}(\mathbf{x}_2), \quad \mathbf{x}_1, \mathbf{x}_2 \in K. \tag{b}$$

Kun tässä sovelletaan vasemmalla puolella $\mathbb{R}^n$ väliarvolausetta (Lause XI.3.3) funktioihin $f_i = (\mathbf{f})_i$, niin seuraa

$$f_i(\mathbf{x}_1) - f_i(\mathbf{x}_2) = [\nabla f_i(\boldsymbol{\xi}_i)]^T(\mathbf{x}_1 - \mathbf{x}_2), \quad i = 1 \ \ldots n, \tag{c}$$

missä $\boldsymbol{\xi}_i \in K$ (tarkemmin $\boldsymbol{\xi}_i = t_i(\mathbf{x}_1 - \mathbf{x}_2)$, $t_i \in (0,1)$). Jatkuvuusoletuksien perusteella on tässä edelleen

$$\nabla f_i(\boldsymbol{\xi}_i) = \nabla f_i(\mathbf{a}) + \mathbf{d}_i, \quad |\mathbf{d}_i| \to 0 \text{ kun } \rho \to 0, \quad i = 1 \ldots n. \tag{d}$$

(Tässä sivuutetaan hieman teknisiä yksityiskohtia.) Yhdistämällä (b), (c) ja (d) nähdään, että

$$\mathbf{r}(\mathbf{x}_1) - \mathbf{r}(\mathbf{x}_2) \;=\; [\mathbf{d}_1 \ldots \mathbf{d}_n]^T(\mathbf{x}_1 - \mathbf{x}_2),$$

joten Cauchyn–Schwarzin epäyhtälön perusteella

$$|\mathbf{r}(\mathbf{x}_1 - \mathbf{r}(\mathbf{x}_2)| \;=\; \Big(\sum_{i=1}^n \langle \mathbf{d}_i, \mathbf{x}_1 - \mathbf{x}_2\rangle^2\Big)^{1/2} \;\le\; L|\mathbf{x}_1 - \mathbf{x}_2|, \tag{e}$$

missä

$$L = \big(\sum_{i=1}^n |\mathbf{d}_i|^2\big)^{1/2} \to 0, \quad \text{kun } \rho \to 0.$$

Koska $\mathbf{Jf}(\mathbf{a})$ on säännöllinen, niin yhtälö (b) on toisin kirjoitettuna

$$\mathbf{x}_1 - \mathbf{x}_2 = [\mathbf{Jf}(\mathbf{a})]^{-1}[\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)] - [\mathbf{Jf}(\mathbf{a})]^{-1}[\mathbf{r}(\mathbf{x}_1) - \mathbf{r}(\mathbf{x}_2)]. \tag{f}$$

Kun tässä sovelletaan kolmioepäyhtälöä ja arviota (e), niin seuraa

$$\begin{aligned} |\mathbf{x}_1 - \mathbf{x}_2| &\le |[\mathbf{Jf}(\mathbf{a})]^{-1}[\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)]| + |[\mathbf{Jf}(\mathbf{a})]^{-1}[\mathbf{r}(\mathbf{x}_1) - \mathbf{r}(\mathbf{x}_2)]| \\ &\le C|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)| + CL|\mathbf{x}_1 - \mathbf{x}_2|, \quad \mathbf{x}_1, \mathbf{x}_2 \in K, \end{aligned} \tag{g}$$

missä $C = \|[\mathbf{Jf}(\mathbf{a})]^{-1}\|$ (ks. Harj.teht. X.1:14). Koska C on vain matriisista $[\mathbf{Jf}(\mathbf{a})]^{-1}$ riippuva vakio ja $L \to 0$ kun $\rho \to 0$, niin voidaan valita $\rho \in (0, \delta)$ siten, että pätee

$$CL \le \frac{1}{2}, \tag{h}$$

jolloin epäyhtälöistä (g) ja (h) seuraa

$$|\mathbf{x}_1 - \mathbf{x}_2| \;\le\; 2C|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)| \quad \forall\, \mathbf{x}_1, \mathbf{x}_2 \in K. \tag{i}$$

Tämän mukaan $\forall\, \mathbf{x}_1, \mathbf{x}_2 \in K$ pätee $\mathbf{f}(\mathbf{x}_1) = \mathbf{f}(\mathbf{x}_2) \;\Rightarrow\; \mathbf{x}_1 = \mathbf{x}_2$, joten $\mathbf{f}$ on K:ssa 1-1, kun $\rho \in (0, \delta)$ on valittu niin, että ehto (h) toteutuu. Käänteisfunktiolauseen 1. osaväittämä on näin todistettu.

Käänteisfunktiolauseen toisen väittämän todistamiseksi osoitetaan, että kun ρ on valittu em. tavalla, niin $\exists\, \varepsilon > 0$ siten, että Kontraktiokuvauslauseen ehdot toteutuvat, kun $|\mathbf{y} - \mathbf{b}| \le \varepsilon$ ja ratkaistava yhtälöryhmä kirjoitetaan muotoon

$$\mathbf{f}(\mathbf{x}) = \mathbf{y} \;\Leftrightarrow\; \mathbf{x} = \mathbf{F}(\mathbf{x}), \quad \mathbf{F}(\mathbf{x}) = \mathbf{x} - [\mathbf{Jf}(\mathbf{a})]^{-1}[\,\mathbf{f}(\mathbf{x}) - \mathbf{y}\,], \quad \mathbf{x} \in K.$$

— Huomattakoon, että kiintopisteiteraatio $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$ on tällöin sama kuin edellisessä luvussa tarkasteltu yksinkertaistettu Newtonin iteraatio sovellettuna yhtälöryhmään $\mathbf{f}(\mathbf{x}) - \mathbf{y} = \mathbf{0}$ (alkuarvauksella $\mathbf{x}_0 = \mathbf{a}$).

Deifferentiaalikehitelmän (a) perusteella $\mathbf{F}$:n lauseke saadaan muotoon

$$\mathbf{F}(\mathbf{x}) = \mathbf{a} + [\mathbf{Jf}(\mathbf{a})]^{-1}(\mathbf{y} - \mathbf{b}) - [\mathbf{Jf}(\mathbf{a})]^{-1}\mathbf{r}(\mathbf{x}), \quad \mathbf{x} \in K.$$

Kun käytetään arviota (e) tässä lausekkeessa, niin seuraa

$$\begin{aligned} |\mathbf{F}(\mathbf{x}_1) - \mathbf{F}(\mathbf{x}_2)| &= \big|[\mathbf{Jf}(\mathbf{a})]^{-1}[\,\mathbf{r}(\mathbf{x}_1) - \mathbf{r}(\mathbf{x}_2)\,]\big| \\ &\le CL|\mathbf{x}_1 - \mathbf{x}_2|, \quad \mathbf{x}_1, \mathbf{x}_2 \in K \end{aligned} \tag{j}$$

ja koska $\mathbf{r}(\mathbf{a}) = \mathbf{0}$, niin seuraa myös

$$\begin{aligned} |\mathbf{F}(\mathbf{x}) - \mathbf{a}| &= \big|[\mathbf{Jf}(\mathbf{a})]^{-1}(\mathbf{y} - \mathbf{b}) - [\mathbf{Jf}(\mathbf{a})]^{-1}[\,\mathbf{r}(\mathbf{x}) - \mathbf{r}(\mathbf{a})\,]\big| \\ &\le C|\mathbf{y} - \mathbf{b}| + CL|\mathbf{x} - \mathbf{a}|, \quad \mathbf{x} \in K \end{aligned} \tag{k}$$

(C kuten edellä). Valitaan nyt ε siten, että toteutuu

$$0 < \varepsilon \leq \frac{\rho}{2C}\,. \tag{l}$$

Tällöin jos $|\mathbf{y}-\mathbf{b}| \leq \varepsilon$, niin epäyhtälöistä (j), (k), (h) ja (l) seuraa

$$\begin{cases} |\mathbf{F}(\mathbf{x}_1)-\mathbf{F}(\mathbf{x}_2)| \;\leq\; \frac{1}{2}|\mathbf{x}_1-\mathbf{x}_2|, & \mathbf{x}_1\,,\mathbf{x}_2 \in K, \\ |\mathbf{F}(\mathbf{x})-\mathbf{a}| \;\leq \frac{\rho}{2}+\frac{\rho}{2}=\rho, & \mathbf{x}\in K. \end{cases}$$

Näistä ensimmäisen ehdon mukaan $\mathbf{F}$ on kontraktiokuvaus K:ssa ja toisen ehdon mukaan $\mathbf{F}(K)\subset K$, joten Kontraktiokuvauslauseen ehdot ovat täytetyt. Tämän lauseen mukaan yhtälöryhmällä $\mathbf{f}(\mathbf{x})=\mathbf{y} \Leftrightarrow \mathbf{x}=\mathbf{F}(\mathbf{x})$ on K:ssa yksikäsitteinen ratkaisu, kun ρ ja ε on valittu em. tavalla ja $|\mathbf{y}-\mathbf{b}|\leq\varepsilon$. Käänteisfunktiolauseen toinen väittämä on näin todistettu.

Viimeisen osaväittämän todistamiseksi valitaan yhtälössä (f) $\mathbf{x}_1=\mathbf{x}=\mathbf{f}^{-1}(\mathbf{y})$, missä $0<|\mathbf{y}-\mathbf{b}|\leq\varepsilon$ (jolloin $\mathbf{x}\neq\mathbf{a}$) ja $\mathbf{x}_2=\mathbf{a}$, jolloin yhtälö saa muodon

$$\mathbf{f}^{-1}(\mathbf{y})-\mathbf{a}=[\mathbf{Jf}(\mathbf{a})]^{-1}(\mathbf{y}-\mathbf{b})-[\mathbf{Jf}(\mathbf{a})]^{-1}\mathbf{r}(\mathbf{x}). \tag{m}$$

Epäyhtälön (i) mukaan on (samoilla $\mathbf{x}$ ja $\mathbf{y}$)

$$|\mathbf{x}-\mathbf{a}|\leq 2C|\mathbf{y}-\mathbf{b}|.$$

Tämän sekä arvion $|[\mathbf{Jf}(\mathbf{a})]^{-1}\mathbf{r}(\mathbf{x})|\leq C|\mathbf{r}(\mathbf{x})|$ perusteella

$$\frac{|[\mathbf{Jf}(\mathbf{a})]^{-1}\mathbf{r}(\mathbf{x})|}{|\mathbf{y}-\mathbf{b}|} \leq \frac{C|\mathbf{r}(\mathbf{x})|}{|\mathbf{x}-\mathbf{a}|}\cdot\frac{|\mathbf{x}-\mathbf{a}|}{|\mathbf{y}-\mathbf{b}|} \leq 2C^2\,\frac{|\mathbf{r}(\mathbf{x})|}{|\mathbf{x}-\mathbf{a}|}\,.$$

Koska tässä $|\mathbf{y}-\mathbf{b}|\to 0 \Rightarrow |\mathbf{x}-\mathbf{a}|\to 0$ (edellisen epäyhtälön mukaan) ja koska $|\mathbf{r}(\mathbf{x})|/|\mathbf{x}-\mathbf{a}|\to 0$, kun $|\mathbf{x}-\mathbf{a}|\to 0$, niin $|[\mathbf{Jf}(\mathbf{a})]^{-1}\mathbf{r}(\mathbf{x})| = o(|\mathbf{y}-\mathbf{b}|)$. Yhtälö (m) voidaan siis kirjoittaa

$$\mathbf{f}^{-1}(\mathbf{y})=\mathbf{f}^{-1}(\mathbf{b})+[\mathbf{Jf}(\mathbf{a})]^{-1}(\mathbf{y}-\mathbf{b})+\mathbf{r}(\mathbf{y}), \quad |\mathbf{r}(\mathbf{y})|=o(|\mathbf{y}-\mathbf{b}|).$$

Näin ollen $\mathbf{f}^{-1}$ on differentioituva pisteessä $\mathbf{b}$ ja $\mathbf{Jf}^{-1}(\mathbf{b})=[\mathbf{Jf}(\mathbf{a})]^{-1}$. $\square$

HARJOITUSTEHTÄVIÄ

1. Tarkastellaan funktiota $\mathbf{f}(x,y)=(2x^2-y^2,-x^2+2y^2)$. a) Määritä $\mathbf{f}$:n arvojoukko $\mathbf{f}(\mathbb{R}^2)$. b) Millaisille joukoille $A\subset\mathbb{R}^2$ pätee: $f: A\to f(A)$ on bijektio? c) Näytä, että $\mathbf{f}$ on paikallisesti kääntyvä pisteessä (a,b), jos $ab\neq 0$. Onko tämä ehto myös välttämätön paikalliselle kääntyvyydelle? d) Määritä $\mathbf{f}$:n paikallisten käänteisfunktioiden lausekkeet pisteissä $(1,1)$, $(-1,1)$, $(1,-1)$ ja $(-1,-1)$.

2. Olkoon $\mathbf{f}(x,y) = (x-y+e^{x+2y}, x+y+e^{x-2y})$ ja olkoon $\mathbf{f}^{-1} = \mathbf{f}$:n paikallinen käänteisfunktio origon ympäristössä. Edelleen olkoon
   a) $K_a(x_0,y_0) = \{(x,y) \in \mathbb{R}^2 \mid |x-x_0| \le a \ \ \& \ \ |y-y_0| \le a\}$,
   b) $K_a(x_0,y_0) = \{(x,y) \in \mathbb{R}^2 \mid (x-x_0)^2 + (y-y_0)^2 \le a^2\}$,
   missä $a > 0$ on pieni luku. Approksimoimalla $\mathbf{f}$ affiinikuvauksella määritä likimäärin joukot $\mathbf{f}\big(K_a(0,0)\big)$ ja $\mathbf{f}^{-1}\big(K_a(1,1)\big)$. Piirrä kuviot!

3. Näytä, että seuraavat funktiot ovat bijektioita kuvauksina $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$. Määritä käänteisfunktion lauseke, mikäli mahdollista. Määritä myös pisteet, joissa $\mathbf{Jf}$ on singulaarinen.
   a) $\mathbf{f}(x,y) = (x^3+y^3, x^3-y^3)$ b) $\mathbf{f}(x,y) = ((x+y)^3, (x-y)^3)$
   c) $\mathbf{f}(x,y) = (x-y, x+y^3)$ d) $\mathbf{f}(x,y) = (e^x - y, x + x^3 + y + y^3)$

4. Olkoon $\mathbf{f} = (e^x \cos y, e^x \sin y)$. a) Näytä, että $\mathbf{f}$ on kaikkialla paikallisesti kääntyvä. b) Määritä $\mathbf{f}(A)$, $A = \mathbb{R} \times [0, 2\pi]$. Onko $\mathbf{f} : A \to \mathbf{f}(A)$ 1-1 ? c) Olkoon $A = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < R^2\}$. Mikä on suurin R:n arvo, jolla $\mathbf{f} : A \to \mathbf{f}(A)$ on bijektio?

5. Missä pisteissä (x,y) toteutuvat Käänteisfunktiolauseen oletukset funktiolle $\mathbf{f}(x,y) = (x^2+2y, y^2-2x)$? Määritä Jacobin matriisi $\mathbf{Jf}^{-1}(u,v)$, $(u,v) = \mathbf{f}(x,y)$ näissä pisteissä.

6. Tasossa liikkuvan robotin käsivarsi koostuu origoon kiinnitetystä olkavarresta (janasta) OP ja pisteeseen P kiinnitetystä kyynärvarresta PQ. Kummankin varren pituus = 1. Kättä (Q) ohjataan säätämällä olkavarren ja x-akselin välistä napakulmaa φ ja kyynärvarren ja olkavarren välistä kulmaa θ. Kulmat ovat vaihdeltavissa väleillä $\varphi \in [0,\pi]$ ja $\theta \in [-5\pi/6, 5\pi/6]$ ($\theta > 0$ vastapäivään).
   a) Määritä robotin käden sijainti funktiona $\mathbf{f}(\varphi,\theta)$.
   b) Mihin tason pisteisiin käsi ulottuu, eli mikä on $\mathbf{f}$:n arvojoukko? Kuvio!
   c) Totea, että $\mathbf{f}$ ei ole 1–1. Missä pisteissä $(\varphi,\theta) \in (0,\pi) \times (-5\pi/6, 5\pi/6)$ toteutuvat Käänteisfunktiolauseen ehdot?
   d) Robotin käsi halutaan siirtää pisteestä $(1.4, 0.2)$ pisteeseen $(1.42, 0.22)$. Arvioi differentiaalin avulla, millaisilla kulman muutoksilla $\Delta\varphi$ ja $\Delta\theta$ tämä liike saadaan aikaan, kun käsivarsi on asennossa, jossa $\theta < 0$.

7. Johda implisiittifunktion $\mathbf{y} = \mathbf{g}(\mathbf{x})$ (Lause XI.7.3) Jacobin matriisin laskusääntö implisiittisellä osittaisderivoinnilla yhtälöryhmästä

$$\mathbf{F}(\mathbf{x}, \mathbf{g}(\mathbf{x})) = \mathbf{0} \quad \Leftrightarrow \quad F_i\big(\mathbf{x}, g_1(\mathbf{x}), \ldots, g_p(\mathbf{x})\big) = 0, \quad i = 1 \ldots p.$$

8. Oletetaan, että yhtälöryhmä $F(x,y,u,v)=0,\ G(x,y,u,v)=0$ toteutuu, kun $(x,y,u,v)=(a,b,c,d)$, ja että ko. pisteen lähellä yhtälöryhmä ratkeaa sekä muotoon $(u,v)=\mathbf{g}(x,y)$ että muotoon $(x,y)=\mathbf{h}(u,v)$. Johda Implisiittifunktiolauseen viimeisestä osaväittämästä (ko. lauseen oletuksin) tulos $\mathbf{Jg}(a,b)\mathbf{Jh}(c,d)=\mathbf{I}$. Millä muulla tavalla tämä on pääteltävissä?

9. Olkoon f säännöllinen funktio tyyppiä $f:\mathbb{R}^2\to\mathbb{R}$. Jos yhtälöryhmä

$$\begin{cases} f(f(x,y),z)=0 \\ f(x,f(y,z))=0 \end{cases}$$

toteutuu, kun $(x,y,z)=(a,b,c)$, niin millaisella f:n osittaisderivaattoja f_x ja f_y koskevalla ehdolla voidaan taata, että pisteen (a,b,c) ympäristössä yhtälöryhmä ratkeaa muotoon $x=x(z),\ y=y(z)$?

10. Olkoon $a,b,c>0$. Näytä, että yhtälöryhmä

$$\begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1 \\ \dfrac{(2x-a)^2}{a^2}+\dfrac{4y^2}{b^2}=1 \end{cases}$$

määrittelee pisteen $P=(a,b,c)$ ympäristössä parametrisen käyrän $S:\ x=f(z)\ \&\ y=g(z)$. Mikä on S:n tangenttivektori pisteessä P?

11. Näytä, että yhtälö tai yhtälöryhmä määrittelee annetun pisteen P ympäristössä implisiittifunktion annettua muotoa sekä ratkaise lisätehtävä:

a) $x^6+2y^4-3x^2y=0,\ P=(1,1),\ y=f(x)$. Käyrän $y=f(x)$ tangentin yhtälö pisteessä P?

b) $\sin(yz)+y^2e^z=x,\ P=(e\pi^2,\pi,1),\ z=f(x,y)$. Arvioi $f(27,3)$ differentiaalin avulla.

c) $ye^{xz}+\sin(x-y+z)=0,\ P=(0,0,0),\ z=f(x,y)$. Pinnan $z=f(x,y)$ tangenttitason yhtälö origossa?

d) $x=z+y\sin z,\ P=(0,0,0),\ z=f(x,y)$. Laske $f_{xy}(0,0)$.

e) $x-y-z+1=0\ \&\ x^2+y^2-2z=0,\ P=(1,1,1),\ (x,y)=\mathbf{f}(z)$. Geometrinen tulkinta?

f) $2x^2+3y^2+z^2=47\ \&\ x^2+2y^2-z=0,\ P=(-2,1,6),\ (y,z)=(f(x),g(x))$. Laske $f'(-2)$ ja $g'(-2)$.

g) $y^5+xy+z^2=4\ \&\ e^{xz}=y^2,\ (3,1,0),\ (x,y)=(f(z),g(z))$. Laske $f'(0)$ ja $g'(0)$.

h) $2x^2+3y^2+z^2=47\ \&\ x^2+2y^2-z=0,\ P=(-2,1,6),\ (y,z)=(f(x),g(x))$. Laske $f'(-2)$ ja $g'(-2)$.

12. Näytä, että seuraavat yhtälöryhmät määrittelevät annettua muotoa olevan implisiittifunktion $\mathbf{g}$ annetun pisteen lähellä, ja laske Jacobin matriisi $\mathbf{Jg}$ ko. pisteessä.

a) $\begin{cases} x - u^3 - v^3 = 0 \\ y - uv + v^2 = 0, \end{cases}$ $\quad (u, v) = \mathbf{g}(x, y), \quad (x, y, u, v) = (2, 0, 1, 1)$

b) $\begin{cases} xe^y - u^3 z + \sin v = 0 \\ x^3 u^2 + yv - uz \cos v = 0, \end{cases}$ $\quad (u, v) = \mathbf{g}(x, y, z), \quad (x, y, z, u, v) = (1, 0, 1, 1, 0)$

c) $\begin{cases} xy^2 + zu + v^2 = 3 \\ x^3 z + 2y - uv = 2 \\ xu + yv - xyz = 1, \end{cases}$ $\quad (x, y, z) = \mathbf{g}(u, v), \quad (x, y, z, u, v) = (1, 1, 1, 1, 1)$

13. (*) Näytä, että funktio $\mathbf{f}(x, y) = (2x - xy, 4y - xy)$ on paikallisesti kääntyvä pisteessä (a, b) täsmälleen kun $a + 2b \neq 4$.

14. (*) Oletetaan, että yhtälö $F(x, y, z) = 0$ määrittelee (Implisiittifunktiolauseen oletuksin) implisiittifunktiot $x = x(y, z)$, $y = y(z, x)$ ja $z = z(x, y)$. Näytä, että pätee:

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1.$$

Tarkista säännön pätevyys, kun $F(x, y, z) = x^3 y^2 z - 1$. Muotoile ja todista vastaava yleisempi väittämä koskien yhtälön $F(x_1, \ldots, x_n) = 0$ määrittelemiä implisiittifunktioita (n kpl, $n \geq 2$).

XI.8 Usean muuttujan ääriarvotehtävät

Aloitetaan määritelmästä.

MÄÄRITELMÄ XI.8.1 Funktiolla $f : \mathcal{D}_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, on pisteesä $\mathbf{c} \in \mathcal{D}_f$ *paikallinen maksimi*, jos jollakin $\delta > 0$ pätee

$$0 < |\mathbf{x} - \mathbf{c}| < \delta \;\Rightarrow\; x \in \mathcal{D}_f \text{ ja } f(\mathbf{x}) < f(\mathbf{c}),$$

ja *paikallinen minimi*, jos jollakin $\delta > 0$ pätee

$$0 < |\mathbf{x} - \mathbf{c}| < \delta \;\Rightarrow\; x \in \mathcal{D}_f \text{ ja } f(\mathbf{x}) > f(\mathbf{c}).$$

Jos f:llä on pisteessä $\mathbf{c}$ paikallinen maksimi tai minimi, niin f:llä on *paikallinen ääriarvo* pisteessä $\mathbf{c}$. Jos $\mathbf{c} \in \mathcal{D}_f$ ei ole funktion f paikallinen ääriarvokohta, mutta jollakin $\delta > 0$ pätee

$$\begin{aligned} \text{joko}: \quad & 0 < |\mathbf{x} - \mathbf{c}| < \delta \;\Rightarrow\; x \in \mathcal{D}_f \text{ ja } f(\mathbf{x}) \le f(\mathbf{c}), \\ \text{tai}: \quad & 0 < |\mathbf{x} - \mathbf{c}| < \delta \;\Rightarrow\; x \in \mathcal{D}_f \text{ ja } f(\mathbf{x}) \ge f(\mathbf{c}), \end{aligned}$$

niin sanotaan, että $\mathbf{c}$ on f:n *laakapiste*.

ESIMERKKI 1 Funktiolla $f(x, y) = ax^2 + by^2$ on origossa paikallinen ääriarvokohta, jos $ab > 0$. Kyseessä on paikallinen minimi, jos $a > 0$ ja $b > 0$, ja paikallinen maksimi, jos $a < 0$ ja $b < 0$. Jos on $ab = 0$, niin origo on laakapiste. Tapauksessa $ab < 0$ ei origo ole ääriarvokohta eikä laakapiste. □

Funktion paikallista ääriarvoa sanotaan myös *suhteelliseksi*, jolloin funktion maksimia/minimiä koko tarkastelun kohteena olevassa joukossa (esim. koko määrittelyjoukossa) sanotaan *absoluuttiseksi*. Myös termiä *globaali* ääriarvo käytetään, etenkin jos kyse on funktion koko määrittelyjoukosta.

ESIMERKKI 1 (jatko) Jos joukko $A \subset \mathbb{R}^2$ sisältää origon, niin tapauksessa $ab > 0$ on $f(0, 0) = 0$ funktion absoluuttinen minimi/maksimi A:ssa. □

Gradientin avulla voidaan helposti paikallistaa reaaliarvoisen funktion paikalliset minimit ja maksimit, jos funktio on differentioituva. Nimittäin Luvun XI.3 tarkastelujen perusteella seuraava tulos on varsin ilmeinen:

PROPOSITIO XI.8.2 Jos $f : \mathcal{D}_f \to \mathbb{R}$, $\mathcal{D}_f \subset \mathbb{R}^n$, on differentioituva pisteessä $\mathbf{c} \in \mathcal{D}_f$ ja $\mathbf{c}$ on$f$:n paikallinen ääriarvo- tai laakapiste, niin $\boldsymbol{\nabla} f(\mathbf{c}) = \mathbf{0}$.

Gradientin nollakohtia sanotaan yleisemmin funktion *kriittisiksi pisteiksi*.

Esimerkki 2 Etsi funktion $f(x,y) = x^2 + 2y^2 + 2xy + 4x + 14y$ kriittiset pisteet. (Vrt. Luku IV.3, Esimerkki 1.)

Ratkaisu

$$\nabla f(x,y) = \vec{0} \quad \Leftrightarrow \quad \begin{cases} \partial_x f(x,y) = 2x + 2y + 4 = 0, \\ \partial_y f(x,y) = 2x + 4y + 14 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = 3, \\ y = -5. \end{cases}$$

Ainoa kriittinen piste on siis $(3, -5)$. $\square$

Esimerkissä kriittinen piste on funktion absoluuttinen minimi, kuten selvitettiin Luvussa IV.3 pelkin algebran keinoin. — Kriitisen pisteen ei yleisemmin tarvitse olla paikallinen ääriarvokohta tai laakapiste. Esimerkiksi funktion

$$f(x,y) = x^2 - y^2$$

kriittinen (origo) on nk. *satulapiste* (engl. saddle point) — termi ei kaivanne lähempää määrittelyä. Funktion kriittisten pisteiden luokittelu ääriarvokohdiksi, laakapisteiksi tai satulapisteiksi on ongelma, jonka (algebrallisella) ratkaisulla on yleisempääkin mielenkiintoa. Asiaan palataan hieman myöhemmin. Tässä yhteydessä rajoitutaan ääriarvotehtäviin, joissa kriittisten pisteiden luokittelulla ei ole keskeistä merkitystä.

Rajoitettu optimointi

Sovelluksissa esiintyvät usean muuttujan ääriarvotehtävät voidaan yleensä muotoilla *rajoitetun optimoinnin* (engl. constrained optimization) ongelmina. Tyypillisessä rajoitetun optimoinnin ongelmassa halutaan etsiä reaaliarvoisen funktion $f(\mathbf{x})$, $\mathcal{D}_f \subset \mathbb{R}^n$ minimi tai maksimi joukossa A, jonka määritelmä asetetaan muodossa

$$\mathbf{x} \in A \quad \Leftrightarrow \quad g_i(\mathbf{x}) \geq 0, \quad i = 1 \ldots m.$$

Sovelluksissa kyse on yleensä jonkin käytännön kannalta tärkeän (kuten taloudellisen hyödyn tai tappion) maksimointi tai minimointi, eli *optimointi*, tehtävän asettelussa väistämättömien *rajoitusehtojen* voimassa ollessa. — Optimointialgoritmien suunnittelun (tai käytön) kannalta on huomion arvoista, että epäyhtälörajoitteisiin voi sisältyä myös yhtälörajoitteita muodossa

$$g(\mathbf{x}) = 0 \quad \Leftrightarrow \quad \begin{cases} g(\mathbf{x}) & \geq 0, \\ -g(\mathbf{x}) & \geq 0. \end{cases}$$

Rajoitteita sanotaan *lineaarisiksi*, jos funktiot g_i ovat ensimmäisen asteen polynomeja (affiinikuvauksia). Jos myös f on tätä tyyppiä, on kyse *lineaarisen ohjelmoinnin* ongelmasta, vrt. luku X.7.

Tarkastellaan esimerkkinä kahden muuttujan tilannetta, jossa optimointiongelman asettelu on

$$f(x,y) = \min / \max\,! \quad \text{ehdoilla} \quad g_i(x,y) \geq 0, \quad i = 1 \ldots m.$$

Oletetaan jatkossa, että sekä f että rajoitefunktiot g_i ovat määriteltyjä ja differentioituvia $\mathbb{R}^2$:ssa. Ongelman voi purkaa kahteen osaan: Ratkaistaan ensin *rajoittamaton* ongelma käymällä läpi f:n kaikki kriittiset pisteet. Kriittisistä pisteistä *käypiä* ovat ne, joissa asetetut rajoitusehdot ovat voimassa. Oletetaan, että näitä on äärellinen määrä, jolloin voidaan suorittaa valinta: Valitaan pisteistä se, jossa f saa pienimmän/suurimman arvonsa.

Optimointitehtävää ei ole vielä ratkaistu, vaan hankalampi vaihe on vasta edessä: On määrättävä erikseen funktion minimi/maksimi tarkasteltavan joukon A *reunalla* ∂A, joka määritellään

$$(x,y) \in \partial A \quad \Leftrightarrow \quad g_i(x,y) = 0 \quad \text{jollakin } i.$$

Kuvassa on koko optimointiongelma kuvattuna yhden rajoitusehdon ($m = 1$) tapauksessa.

Optimointiongelmaa yhden tai useamman yhtälörajoitteen alaisena sanotaan *sidotuksi* ääriarvotehtäväksi. Kahden muuttujan tapauksessa ongelma ratkeaa luontevasti, jos oletetaan, että jokainen rajoitusehto määrittelee parametrisen käyrän

muotoa

$$g_i(x,y) = 0 \;\Leftrightarrow\; x = x_i(t),\; y = y_i(t),\; t \in \mathbb{R}, \quad i = 1 \ldots m,$$

ja oletetaan vielä, että reuna ∂A voidaan pilkkoa käyränkaariksi

$$\partial A_i = \{(x,y) \in \mathbb{R}^2 \mid x = x_i(t),\; y = y_i(t),\; t \in [a_i, b_i]\},$$

jotka yhdessä peittävät ∂A:n ja joiden yhteisiä pisteitä ovat enintään päätepisteet $(x_i(a_i), y_i(a_i))$ ja $(x_i(b_i), y_i(b_i))$. Optimointiongelman jäljellä oleva osa on näillä oletuksilla palautettu yhden muuttujan rajoitetuiksi optimointiongelmiksi:

$$F_i(t) = f\big(x_i(t), y_i(t)\big) = \min/\max\,! \quad \text{kun} \quad a_i \le t \le b_i\,, \quad i = 1 \ldots m.$$

Nämä ongelmat ratkeavat normaaliin tapaan, eli etsimällä funktioiden F_i kriittiset pisteet (= derivaatan nollakohdat) avoimilta väleiltä (a_i, b_i) ja tutkimalla erikseen päätepisteet a_i, b_i. Sikäli kuin funktioilla F_i on käypiä kriittisiä pisteitä äärellinen määrä, on löydetty äärellinen pistejoukko, josta optimoitavan funktion ääriarvot välttämättä löytyvät.

Esimerkki 3 Ratkaise rajoitetut optimointitehtävät

$$f(x,y) = 6x^2 - 4xy + 9y^2 + 16x - 22y = \min\,! \quad \text{ja} \quad f(x,y) = \max\,!$$
$$\text{kun} \quad x \ge 0, \quad x^2 + (y-1)^2 \le 1.$$

<u>Ratkaisu</u> Kriittisiä pisteitä on yksi:

$$\nabla f(x,y) = \vec{0} \quad \Leftrightarrow \quad \begin{cases} 12x - 4y + 16 = 0, \\ -4x + 18y - 22 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = -1, \\ y = 1. \end{cases}$$

Tämä ei ole käypä, joten siirrytään tutkimaan funktiota tarkasteltavan joukon A reunalla ∂A:

$$(x,y) \in \partial A \quad \Leftrightarrow \quad x = 0 \quad \text{tai} \quad x^2 + (y-1)^2 = 1.$$

Tämä jakautuu kahteen osaan:

$$\partial A_1 : \quad x = 0 \;\;\&\;\; x^2 + (y-1)^2 \le 1 \quad \Leftrightarrow \quad x = 0 \;\;\&\;\; y \in [0,2].$$
$$\partial A_2 : \quad x^2 + (y-1)^2 = 1 \;\;\&\;\; x \ge 0.$$

Reunan osa ∂A_1 on jana, jolla

$$F_1(y) = f(0,y) = 9y^2 - 22y.$$

Funktion F_1 kriittinen piste $y = 11/9$ on käypä, joten merkitään muistiin f:n arvo tässä pisteessä sekä janan ∂A_1 päätepisteissä:

$$f(0, 11/9) = -121/9 = -13.4444.., \quad f(0,0) = 0, \quad f(0,2) = -8.$$

Siirrytään reunan osalle ∂A_2. Tämä on puoliympyrän kaari, joka parametrisoituu muodossa

$$x = \sin t, \quad y = 1 - \cos t, \quad t \in [0, \pi].$$

Tutkittava funktio on näin ollen

$$\begin{aligned} F_2(t) &= f(\sin t, 1 - \cos t) \\ &= 6\sin^2 t + 4\sin t\cos t + 9\cos^2 t + 12\sin t + 4\cos t - 13, \quad t \in [0, \pi]. \end{aligned}$$

Funktion F_2 kriittiset pisteet löytyvät ratkaisemalla yhtälö

$$F_2'(t) = \ldots = -3\sin 2t + 4\cos 2t + 12\cos t - 4\sin t = 0.$$

Numeerisin keinoin saadaan ainoaksi käyväksi ratkaisuksi

$$t = 0.927295.., \quad (x(t), y(t)) = (0.800000.., 0.400000..) = (4/5, 2/5).$$

Tässä pisteessä on $F_2(0.927295..) = 8$. Koska reunan osien ∂A_1 ja ∂A_2 päätepisteet ovat samat, niin on päätelty, että f:n pienin ja suurin arvo löytyvät joukosta $\{-121/9,\ 0,\ -8,\ 8\}$. Optimointitehtävien ratkaisut ovat siis

$$\begin{aligned} &f_{\min} = -121/9 = -13.4444.. \quad \text{pisteessä } (x, y) = (0, 11/9) = (0,\ 1.1111..), \\ &f_{\max} = 8 \quad \text{pisteessä } (x, y) = (4/5,\ 2/5) = (0.8,\ 0.4). \quad \square \end{aligned}$$

Lagrangen kertojien menetelmä

Edellä esitettyä ratkaisutapaa voi periaatteessa soveltaa useammankin muuttujan rajoitettuun optimointitehtävään. Ongelmaksi kuitenkin muodostuvat sidotut ääriarvotehtävät, joita joudutaan ratkaisemaan optimoitaessa kohdefunktiota tarkasteltavan joukon reunalla. Useamman kuin kahden muuttujan tapauksessa tällaisten tehtävien ratkaiseminen parametrisoimalla, eli vapaiden muuttujien lukumäärää vähentämällä, on usein hankalaa. Kolmen tai useamman muuttujan sidotuissa ääriarvotehtävissä käytetäänkin yleensä toista, paljon elegantimpaa ja yleispätevämpää menetelmää, jota keksijänsä **J. L. Lagrangen** (1736-1813) mukaan sanotaan *Lagrangen kertojien* (engl. Lagrange multipliers) *menetelmäksi.*

Tarkastellaan aluksi jälleen kahden muuttujan tilannetta. Olkoon etsittävä funktion $f(x, y)$ paikalliset ääriarvokohdat käyrällä S, jonka määrittelee sidosehto $g(x, y) = 0$. Olkoon $P = (x_0, y_0) \in S$ eräs tällainen ääriarvokohta ja olkoon $\vec{t}$ S:n tangenttivektori pisteessä P. Tällöin on oltava

$$\vec{t} \cdot \nabla f(x_0, y_0) = 0.$$

Toisaalta koska käyrä $g(x, y) = 0$ funktion g tasa-arvokäyrä, niin pätee myös (vrt. Luku XI.3)

$$\vec{t} \cdot \nabla g(x_0, y_0) = 0.$$

Näin ollen, sikäli kuin $\nabla g(x_0, y_0) \neq \vec{0}$ (mikä jatkossa oletetaan — pisteet, joissa $g(x_0, y_0) = 0$ ja $\nabla g(x_0, y_0) = \vec{0}$, on tutkittava erikseen), on oltava jollakin $\lambda \in \mathbb{R}$

$$\nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) = \vec{0}.$$

Kysyttyjä paikallisia ääriarvopisteitä voidaan siis hakea ratkaisemalla yhtälöryhmä

$$\begin{cases} f_x(x, y) + \lambda g_x(x, y) = 0, \\ f_y(x, y) + \lambda g_y(x, y) = 0, \\ g(x, y) = 0. \end{cases}$$

Tässä on kolme tuntematonta ja kolme yhtälöä, joten ainakin muodollinen puoli on kunnossa.

Em. yhtälöryhmään päädytään helpommin muistettavalla tavalla, kun määritellään funktio

$$F(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

Tässä on siis rajoitusehdossa esiintyvä funktio g tuotu mukaan kertoimella λ — *Lagrangen kertojalla* — painotettuna. Ym. yhtälöryhmään päädytään nyt yksinkertaisesti, kun haetaan F:n kriittiset pisteet ehdoista

$$\begin{cases} \partial_x F(x, y, \lambda) = 0, \\ \partial_y F(x, y, \lambda) = 0, \\ \partial_\lambda F(x, y, \lambda) = 0. \end{cases}$$

Em. geometrinen perustelu toimii myös kolmen muuttujan tapauksessa. Nimittäin jos $P = (x_0, y_0, z_0)$ on funktion $f(x, y, z)$ paikallinen ääriarvopiste pinnalla $S: \ g(x, y, z) = 0$, niin jokaiselle pinnan tangenttivektorille $\vec{t}$ pisteessä P on oltava voimassa

$$\vec{t} \cdot \nabla f(x_0, y_0, z_0) = 0.$$

Tällöin $\nabla f(x_0, y_0, z_0)$ on pinnan S normaali, joten jollakin $\lambda \in \mathbb{R}$

$$\nabla f(x_0, y_0, z_0) + \lambda \nabla g(x_0, y_0, z_0) = \vec{0}$$

(olettaen, että $\nabla g(x_0, y_0, z_0) \neq \vec{0}$). Ääriarvopisteet löytyvät tällöin funktion

$$F(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$$

kriittisten pisteiden joukosta.

Esimerkki 4 Hae funktion $f(x, y, z) = xy + yz - xz$ ääriarvot yksikköpallon pinnalla.

<u>Ratkaisu</u> Rajoitusehto on $x^2 + y^2 + z^2 = 1$, joten haetaan funktion

$$F(x, y, z, \lambda) = xy + yz - xz + \lambda(x^2 + y^2 + z^2 - 1)$$

kriittisiä pisteitä:

$$\begin{cases} \partial_x F = y - z + 2\lambda x = 0, \\ \partial_y F = x + z + 2\lambda y = 0, \\ \partial_z F = y - x + 2\lambda z = 0, \\ \partial_\lambda F = x^2 + y^2 + z^2 - 1 = 0. \end{cases}$$

Kolme ensimmäistä yhtälöä on matriisimuodossa

$$\begin{bmatrix} 2\lambda & 1 & -1 \\ 1 & 2\lambda & 1 \\ -1 & 1 & 2\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}.$$

Tälle haetaan ei-triviaalia ratkaisua, joten kerroinmatriisin on oltava singulaarinen. Kerroinmatriisin determinantti on

$$\det(\mathbf{A}) = 8\lambda^3 - 6\lambda - 2 = 2(\lambda - 1)(2\lambda + 1)^2,$$

joten on oltava joko $\lambda = 1$ tai $\lambda = -1/2$. Kun $\lambda = 1$, tulee yhtälöryhmäksi

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \underset{\text{(Gauss)}}{\Leftrightarrow} \begin{bmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}.$$

Tämän ratkaisu yhdessä rajoitusehdon $x^2 + y^2 + z^2 = 1$ kanssa on

$$(x, y, z) = t(1, -1, 1), \quad t = \pm 1/\sqrt{3},$$

ja funktion f arvo näissä pisteissä on

$$f\left(\pm\frac{1}{\sqrt{3}}, \mp\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right) = \underline{\underline{-1}}.$$

Jos $\lambda = -1/2$, niin yhtälöryhmäksi tulee

$$\begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \;\Leftrightarrow\; \begin{bmatrix} -1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}.$$

Rajoitusehdon toteuttavia ratkaisuja ovat $(x, y, z) = (t - s, t, s)$, missä $t, s \in \mathbb{R}$ toteuttavat

$$(t - s)^2 + t^2 + s^2 = 1 \;\Leftrightarrow\; 2(t^2 + s^2 - ts) = 1.$$

Funktion f arvoksi näissä pisteissä tulee

$$\begin{aligned} f(x, y, z) &= (t - s)t + ts - (t - s)s \\ &= t^2 + s^2 - ts = \underline{\underline{1/2}}. \end{aligned}$$

Siis

$$\begin{aligned} &f_{\min} = -1 \quad \text{pisteissä} \quad \pm \frac{1}{\sqrt{3}}(1, -1, 1), \\ &f_{\max} = 1/2 \quad \text{käyrällä } S: \ (x, y, z) = (t - s, t, s), \quad (t, s) \in \mathbb{R}^2, \ t^2 + s^2 - ts = 1/2. \end{aligned}$$

Maksimipisteet muodostavat yksikköpallon isoympyrän tasolla $T : x - y + z = 0$. $\square$

Lagrangen kertojien menetelmä toimii, poikkeustilanteita lukuunottamatta, myös useamman rajoitusehdon tapauksessa. Jos on haettava funktion $f(x_1, \ldots, x_n)$ ääriarvot rajoitusehdoilla

$$g_i(x_1, \ldots, x_n) = 0, \quad i = 1 \ldots m < n,$$

niin nämä löydetään funktion

$$F(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

kriittisten pisteiden joukosta. Esimerkiksi tapauksessa $n = 3$, $m = 2$ on tutkittava funktiota

$$F(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g_1(x, y, z) + \mu g_2(x, y, z).$$

Menetelmän idean voi tässäkin tapauksessa perustella geometrisesti: Rajoitusehtojen voi olettaa määrittelevän avaruuskäyrän K, joka on pintojen

$$S_1 : \ g_1(x, y, z) = 0, \quad S_2 : \ g_2(x, y, z) = 0$$

leikkaus. Jos pintojen normaalit käyrän pisteessä (x_0, y_0, z_0) ovat

$$\vec{n}_i = \nabla g_i(x_0, y_0, z_0), \quad i = 1, 2,$$

niin käyrän K tangentti $\vec{t}$ on näitä kumpaakin vastaan kohtisuora. Ääriarvoehto on samaa muotoa kuin yhden rajoitusehdon tapauksessa, eli on oltava

$$\vec{t} \cdot \nabla f(x_0, y_0, z_0) = 0.$$

Koska tässä $\vec{t}$ siis on kohtisuorassa vektoreita $\vec{n}_1$ ja $\vec{n}_2$ vastaan, niin sikäli kuin $\vec{n}_1$ ja $\vec{n}_2$ ovat lineaarisesti riippumattomat, on oltava $\nabla f(x_0, y_0, z_0) = \lambda \vec{n}_1 + \mu \vec{n}_2$ jollakin $(\lambda, \mu) \in \mathbb{R}^2$ eli

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g_1(x_0, y_0, z_0) + \mu \nabla g_2(x_0, y_0, z_0).$$

Esimerkki 5 Minimoi $f(x, y, z) = x + 2y - z$ ehdoilla $x^2 + y^2 + z^2 = 12$ ja $x + y + z = 1$.

<u>Ratkaisu</u> Etsitään funktion

$$F(x, y, z, \lambda, \mu) = x + 2y - z + \lambda(x^2 + y^2 + z^2 - 12) + \mu(x + y + z - 1)$$

kriittiset pisteet:

$$\begin{cases} 2\lambda x + \mu = -1 \\ 2\lambda y + \mu = -2 \\ 2\lambda z + \mu = 1 \\ z^2 + y^2 + z^2 = 12 \\ x + y + z = 1 \end{cases} \Rightarrow \begin{cases} 2\lambda(x + y + z) + 3\mu = -2 \\ x + y + z = 1 \end{cases} \Rightarrow \quad 2\lambda = -3\mu - 2.$$

Yhtälöryhmän kolmesta ensimmäisestä ja neljännestä yhtälöstä seuraa myös

$$4\lambda^2(x^2 + y^2 + z^2) = (\mu + 1)^2 + (\mu + 2)^2 + (\mu - 1)^2 = 48\lambda^2.$$

Tässä on edellisen tuloksen perusteella $48\lambda^2 = 12(3\mu + 2)^2$, joten seuraa

$$105\mu^2 + 140\mu + 42 = 0 \quad \Rightarrow \quad (\mu, \lambda) = \begin{cases} (-0.877485, \quad 0.316228), \\ (-0.455848, \ -0.316228) \end{cases}$$

$$\Rightarrow \quad (x, y, z) = \begin{cases} (-0.19371, \ -1.77485, \quad 2.96856), \\ (\ \ 0.86038, \quad 2.44152, \ -2.30190). \end{cases}$$

Näistä ensimmäinen piste antaa minimiarvon $f_{\min} = \underline{\underline{-6.71197}}$. $\square$

Lagrangen kertojien menetelmää käytettäessä ei syntyvän epälineaarisen yhtälöryhmän ratkaiseminen luonnollisesti aina onnistu käsinlaskulla. Ratkaisut etsitään silloin numeerisin keinoin, tavallisimmin Newtonin menetelmällä. Numeeriseen ratkaisemiseen yhdistettynä Lagrangen kertojien menetelmä näyttääkin itse asiassa parhaat puolensa, sillä ratkaisualgoritmista tulee tällä tavoin hyvin suoraviivainen ja helposti (konevoimin) toteutettava.

Esimerkki 6 Esimerkissä 3 ratkaistiin sidosehdon parametrisoinnilla ääriarvotehtävä

$$\begin{aligned} &f(x,y) = 6x^2 - 4xy + 9y^2 + 16x - 22y = \max\,! \\ &\text{ehdolla} \quad g(x,y) = x^2 + (y-1)^2 - 1 = 0. \end{aligned}$$

Lagrangen kertojaa käytettäessä etsitään funktion

$$F(x,y,\lambda) = 6x^2 - 4xy + 9y^2 + 16x - 22y + \lambda[x^2 + (y-1)^2 - 1]$$

kriittiset pisteet eli ratkaistaan yhtälöryhmä

$$\begin{cases} 6x - 2y + x\lambda - 8 = 0, \\ -2x + 9y + (y-1)\lambda - 11 = 0, \\ x^2 + (y-1)^2 - 1 = 0. \end{cases}$$

Kun ratkaisussa käytetään Newtonin menetelmää, saadaan iteraatiokaavaksi

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \\ \lambda_k \end{bmatrix} - \begin{bmatrix} \lambda_k + 6 & -2 & x_k \\ -2 & \lambda_k + 9 & y_k - 1 \\ 2x_k & 2y_k - 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 6x_k - 2y_k + x_k\lambda_k - 8 \\ -2x_k + 9y_k + (y_k - 1)\lambda_k - 11 \\ x_k^2 + (y_k - 1)^2 - 1 \end{bmatrix}.$$

Alkuarvauksella $(x_0, y_0, \lambda_0) = (1, 0.5, 0)$ saadaan

k	x_k	y_k	λ_k
0	1.000000	0.500000	0.0000
1	1.000000	0.750000	−12.5000
2	0.815287	0.136146	−14.9283
3	0.818006	0.376563	−14.8590
4	0.800355	0.399765	−14.9999
5	0.800000	0.400000	−15.0000

Suppeneminen oli hieman onnekasta, sillä alkuarvaus osoittautui Lagrangen kertojan osalta kehnoksi. Parempi alkuarvaus olisi voitu laskea esim. seuraavasti (yleisempi menettely: ks. Harj.teht. 12):

$$\vec{a} = \nabla f(x_0, y_0) = 26\vec{i} - 17\vec{j}, \quad \vec{b} = \nabla g(x_0, y_0) = 2\vec{i} - \vec{j},$$

$$\vec{a} + \lambda\vec{b} \approx \vec{0} \;\Rightarrow\; \lambda \approx -\frac{\vec{a}\cdot\vec{b}}{|\vec{b}|^2} = -13.8 = \lambda_0. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Määritä funktion suurin ja pienin arvo annetuilla rajoitusehdoilla. Käytä sidosehdoissa (mikäli niitä on) parametrisointia.

a) $x - x^2 + y^2, \quad 0 \le x \le 2, \;\; 0 \le y \le 1$
b) $xy - 2x, \quad -1 \le x \le 1, \;\; 0 \le y \le 1$
c) $xy - x^3y^2, \quad 0 \le x, y \le 1$
d) $xy(a - 2x - y), \quad |x| \le a, \;\; |y| \le a$
e) $xy(1 - x - 2y), \quad x, y \ge 0, \;\; x + y \le 1$
f) $x^2 - 2y^2 - xy - x, \quad x, y \ge -1, \;\; x + y \le 1$
g) $xy - x^2, \quad x^2 + y^2 \le 1$
h) $3 + x - x^2 - y^2, \quad 2x^2 + y^2 \le 1$
i) $y^2 - 2x^2, \quad x^2 - y \le 1, \;\; x + y \le 1$
j) $xy + x - y, \quad x^2 + y^2 \le 1, \;\; x + y + 1 \ge 0, \;\; x - y - 1 \le 0$
k) $x^3 + 3y^3, \quad 0 \le xy \le 1, \;\; x^2 + y^2 \le 4$
l) $\sin x + \sin y - \sin(x + y), \quad 0 \le x, y \le \pi$
m) $\sin x \cos y, \quad x, y \ge 0, \;\; x + y \le 2\pi$
n) $\sin x \sin y \sin(x + y), \quad x, y \ge 0, \;\; x + y \le \pi$
o) $(x - 3y)e^y, \quad |x| \le 1, \;\; |y| \le 1$
p) $xe^{-x^2-y^2}, \quad x^2 + y^2 \le R^2$
q) $(x + 2y)e^{-x^2-y^2}, \quad (x, y) \in \mathbb{R}^2$
r) $xy^2e^{-xy}, \quad x, y \ge 0$
s) $(2x - y)(1 + x^2 + y^2)^{-1}, \quad (x, y) \in \mathbb{R}^2$
t) $(1 + 2x + 3y)(1 + x^2 + y^2)^{-1}, \quad (x, y) \in \mathbb{R}^2$
u) $xy^2 + yz^2, \quad x^2 + y^2 + z^2 \le 1$
v) $xy + yz, \quad x^2 + y^2 + z^2 \le 1$
x) $xy^2z^3, \quad x^2 + y^2 + z^2 \le 1$
y) $xyze^{-x^2-2y^2-3z^2}, \quad (x, y, z) \in \mathbb{R}^3$
z) $(x + y + z)e^{-x^2-2y^2-3z^2}, \quad (x, y, z) \in \mathbb{R}^3$
å) $(x + 2y + 3z + 4u)e^{-x^2y^2z^2u^2}, \quad (x, y, z, u) \in \mathbb{R}^4$

2. Halutaan määrätä funktion $f(x,y,z) = xy + z^2$ minimi- ja maksimiarvo pallopinnalla $S: x^2+y^2+z^2 = 1$. Ratkaise tehtävä a) parametrisoimalla S pallonpintakoordinaateilla, b) Lagrangen kertojien avulla. c) Mitkä ovat f:n minimi- ja maksimiarvot ehdolla $x^2+y^2+z^2 \leq 1$?

3. Määritä Lagrangen kertojien avulla funktion

$$\text{a)}\ \ f(x,y,z) = x^2+y^2+z^2 \qquad \text{b)}\ \ f(x,y,z) = xy$$

pienin ja suurin arvo pallon $S_1 : x^2+y^2+z^2+4x+4y = 0$ ja lieriön $S_2 : x^2+y^2+2x+2y = 6$ leikkauskäyrällä.

4. Ratkaise seuraavat sidotut ääriarvotehtävät Lagrangen kertojien avulla.
   a) $x^2+y^2 = \text{min!/max!}$ käyrällä $S: x^2+8xy+7y^2 = 45$
   b) $x^2+y^2 = \text{min!/max!}$ käyrällä $S: x^4+x^2y^2+5y^4 = 95$
   c) $x^3y^5 = \text{max!}$ ehdolla $x+y = 8$
   d) $x+2y-3x = \text{min!/max!}$ pallopinnalla $x^2+y^2+z^2 = 1$
   e) $x^2+y^2+z^2 = \text{min!}$ ehdolla $xyz^2 = 2$
   f) $xyz = \text{min!/max!}$ pallopinnalla $x^2+y^2+z^2 = 12$
   g) $x = \text{min!/max!}$ käyrällä $S: z = x+y\ \ \&\ \ x^2+2y^2+2z^2 = 8$
   h) $x^2+y^2+z^2 = \text{min!/max!}$ käyrällä $S: z^2 = x^2+y^2\ \ \&\ \ x-2z = 3$
   i) $z = \text{min!/max!}$ käyrällä $S: x^2+y^2 = 8\ \ \&\ \ x+y+z = 1$
   j) $x^3+y^3+z^3 = \text{min!/max!}$ käyrällä $S: x+y+z = 3\ \ \&\ \ x^2+y^2+z^2 = 27$
   k) $\mathbf{a}^T\mathbf{x} = \text{min!/max!}$ ehdolla $|\mathbf{x}| = 1\ \ (\mathbf{a},\mathbf{x} \in \mathbb{R}^n)$
   l) $x_1^{2m}+\ldots+x_n^{2m} = \text{min!}$ ehdolla $x_1+\ldots+x_n = a \neq 0\ \ (m \in \mathbb{N})$

5. Ratkaise Lagrangen kertojaa käyttäen:
   a) Mikä on origon lyhin etäisyys käyrästä $S: x^2+8xy+7y^2 = 45$?
   b) Mikä pinnan $S: xy+2yz+3xz = 0$ piste Q on lähinnä pistettä $P = (-2,2,-2)$?

6. Tetraedrin yksi kärki on origossa ja muut kolme kärkeä ovat origosta suuntiin $\vec{i}$, $\vec{i}-\vec{j}+\vec{k}$ ja $-\vec{i}+\vec{j}+\vec{k}$. Lisäksi origon etäisyys tetraedrin vastakkaisesta sivutahkosta $= 3$. Mikä on tetraedrin pienin tilavuus näillä ehdoilla?

7. a) Näytä, että suorakulmaisista särmöistä, joilla on sama sivutahkojen yhteenlaskettu pinta-ala, kuutio on tilavuudeltaan suurin.

   b) Kanneton suorakulmaisen särmiön muotoinen laatikko valmistetaan kahdesta eri materiaalista siten, että pohjamateriaali on (pintayksikköä kohti) viisi kertaa kalliimpaa kuin laatikon seinämien materiaali. Miten laatikon mitat on valittava, kun halutaan minimoida materiaalikustannukset ehdolla, että laatikon tilavuus $= a^3$?

8. a) Ellipsoidin muotoinen jalokivi

$$A: \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1$$

halutaan hioa suorakulmaisen särmiön muotoiseksi. Miten särmiön mitat on valittava, jotta kiveä menisi hukkaan mahdollisimman vähän?

b) Näytä, että jos α, β ja γ ovat kolmion kulmat, niin

$$\sin\frac{\alpha}{2}\sin\frac{\beta}{2}\sin\frac{\gamma}{2} \leq \frac{1}{8}\,.$$

9. Laske (tarvittaessa numeerisin keinoin) funktion $f(x,y) = (x-y^2)e^{xy}$ pienin ja suurin arvo joukossa a) $A: x^2+y^2 \leq 4$, b) $A: x^4+y^4 \leq 16$.

10. (*) Todista väittämä: Lukujen $a_k \geq 0$, $k = 1\ldots n$ *geometrinen keskiarvo* on enintään yhtä suuri kuin *aritmeettinen keskiarvo*, ts.

$$\sqrt[n]{a_1\cdots a_n} \leq \frac{1}{n}(a_1+\ldots+a_n).$$

Vihje: Aseta sopiva sidosehto!

11. (*) Etsi funktion $f(x,y,z) = (x+2y)e^{xz} + xyz + y^3$ ääriarvokohdat pallopinnalla $S: (x-3)^2+(y+3)^2+(z+4)^2 = 16$. Menetelmä: Lagrangen kertojat ja Newton!

12. (*) Yleisessä Lagrangen kertojien menetelmässä etsitään funktion $F(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x})$, $\mathbf{x}\in\mathbb{R}^n$, $m<n$ kriittistä pistettä $(\mathbf{c},\boldsymbol{\mu})$. Oletetaan, että eräässä $\mathbf{c}$:n ympäristössä $U_\delta(\mathbf{c})$ vektorit $\boldsymbol{\nabla} g_i(\mathbf{x})$, $i = 1\ldots m$ ovat lineaarisesti riippumattomat ja että on käytettävissä alkuarvaus $\mathbf{x}_0 \in U_\delta(\mathbf{c})$. Näytä, että jos alkuarvaus $\boldsymbol{\lambda}_0$ määrätään $\mathbf{x}_0$:n avulla laskemalla

$$|\boldsymbol{\nabla} f(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i \boldsymbol{\nabla} g_i(\mathbf{x}_0)|^2 = \text{min!} \quad \Rightarrow \quad (\lambda_i) = \boldsymbol{\lambda}_0,$$

($|\cdot| = \mathbb{R}^n$:n euklidinen normi), niin pätee: $\mathbf{x}_0 = \mathbf{c} \;\Rightarrow\; \boldsymbol{\lambda}_0 = \boldsymbol{\mu}$. Millaiseen $\boldsymbol{\lambda}_0$:n laskukaavaan tämä menettely johtaa tapauksessa $m = 1$? Sovella menettelyä sidottuun ääriarvotehtävään

$$\begin{aligned} &f(x,y,z) = x^2+y^2+z^2 = \text{min!} \\ &\text{ehdoilla} \quad x^3y + xy^3 + 2 = 0, \quad x^3y + y^3z + xz^3 + 1 = 0, \end{aligned}$$

kun alkuarvaus on $(x_0,y_0,z_0) = (1,-1,1)$. Määritä F:n kriittinen piste likimäärin yhdellä Newtonin iteraation askeleella.

XI.9 Usean muuttujan Taylorin polynomit

Tarkastellaan kahden muuttujan funktiota $f = f(x, y)$ pisteitä (x_0, y_0) ja (x, y) yhdistävällä janalla S. Tällä janalla f voidaan tulkita yhden muuttujan funktiona

$$g(t) = f(x_0 + h_x t, y_0 + h_y t), \quad t \in [0, 1],$$

missä on merkitty

$$h_x = x - x_0, \quad h_y = y - y_0.$$

Oletetaan, että funktion f osittaisderivaatat ∂^α ovat kertalukuun $|\alpha| = n+1$ asti olemassa ja jatkuvia janalla S. Tällöin voidaan osittaisderivoinnin ketjusäännöistä päätellä, että funktio g on $n+1$ kertaa jatkuvasti derivoituva välillä $[0, 1]$, jolloin Taylorin lauseen mukaan pätee jokaisella $t \in (0, 1]$

$$g(t) = \sum_{k=0}^{n} \frac{1}{k!} g^{(k)}(0)\, t^k + \frac{1}{(n+1)!} g^{(n+1)}(\xi)\, t^{n+1},$$

missä $\xi \in (0, t)$. Kun erityisesti valitaan $t = 1$, on tulos

$$f(x, y) = g(1) = \sum_{k=0}^{n} \frac{1}{k!} g^{(k)}(0) + \frac{1}{(n+1)!} g^{(n+1)}(\xi). \qquad (\star)$$

Tässä voidaan derivaatta $g'(t)$ laskea ketjusäännön perusteella muodossa

$$\begin{aligned} g'(t) &= h_x f_x(x_0 + h_x t, y_0 + h_y t) + h_y f_y(x_0 + h_x t, y_0 + h_y t) \\ &= [\,(h_x \partial_x + h_y \partial_y) f\,](x_0 + h_x t, y_0 + h_y t). \end{aligned}$$

Soveltamalla tätä sääntöä uudelleen nähdään, että

$$g^{(k)}(t) = [\,(h_x \partial_x + h_y \partial_y)^k f\,](x_0 + h_x t, y_0 + h_y t).$$

Tässä h_x ja h_y tulkitaan derivoitaessa vakioiksi (koska derivointi kohdistuu muuttujaan t). Tuloksen perusteella $f(x,y)$:n lauseke ($\star$) voidaan tulkita *kahden muuttujan Taylorin kaavana*

$$\boxed{\begin{aligned} f(x,y) &= T_n(x,y,x_0,y_0) + R_n(x,y), \quad \text{missä} \\ T_n(x,y,x_0,y_0) &= \sum_{k=0}^{n} \frac{1}{k!}\,[\,(h_x\partial_x + h_y\partial_y)^k f\,](x_0,y_0), \\ R_n(x,y) &= \frac{1}{(n+1)!}\,[\,(h_x\partial_x + h_y\partial_y)^{n+1} f\,](x(\xi),y(\xi)), \\ h_x &= x - x_0, \quad h_y = y - y_0, \\ x(\xi) &= x_0 + h_x\xi, \quad y(\xi) = y_0 + h_y\xi, \quad \xi \in (0,1). \end{aligned}}$$

Tässä $T_n(x,y,x_0,y_0)$ on f:n *Taylorin polynomi* astetta n pisteessä (x_0,y_0) ja R_n on *jäännöstermi.* Huomioimalla, että (binomikaavan perusteella, vrt. Luku XI.2)

$$(h_x\partial_x + h_y\partial_y)^k f = \sum_{l=0}^{k} \binom{k}{l} h_x^{k-l} h_y^l \partial_x^{k-l}\partial_y^l f, \qquad \binom{k}{l} = \frac{k!}{(k-l)!\,l!}\,,$$

saadaan polynomille $T_n(x,y,x_0,y_0)$ konkreettisempi esitysmuoto

$$\boxed{T_n(x,y,x_0,y_0) = \sum_{k=0}^{n}\sum_{l=0}^{k} \frac{1}{(k-l)!\,l!}\,\frac{\partial^k f}{\partial x^{k-l}\partial y^l}(x_0,y_0)\,(x-x_0)^{k-l}(y-y_0)^l.}$$

Tästä ja helposti todennettavasta derivointikaavasta

$$\left[\frac{\partial^{i+j}}{\partial x^i \partial y^j}(x-x_0)^m (y-y_0)^n\right]\Bigg|_{\substack{x=x_0\\ y=y_0}} = \begin{cases} m!\,n!, & \text{jos } i=m \text{ ja } j=n, \\ 0, & \text{muulloin} \end{cases}$$

nähdään, että polynomi $p(x,y) = T_n(x,y,x_0,y_0)$ toteuttaa ehdot

$$\partial^\alpha p(x_0,y_0) = \partial^\alpha f(x_0,y_0), \quad |\alpha| \le n.$$

Taylorin polynomi myös määräytyy yksikäsitteisesti näistä ehdoista, kuten yhden muuttujan tapauksessa (vrt. Luku VII.4).

Esimerkki 1 Määrää funktion

$$f(x,y) = e^{-x} + xy + y^2 + \cos x\,\cos y$$

Taylorin polynomi $T_2(x,y,0,0)$.

Ratkaisu Derivoimalla todetaan

$$f(0,0) = 2, \quad f_x(0,0) = -1, \quad f_y(0,0) = 0,$$
$$f_{xx}(0,0) = 0, \quad f_{xy}(0,0) = f_{yx}(0,0) = f_{yy}(0,0) = 1,$$

joten

$$\begin{aligned} T_2(x,y,0,0) &= 2 + (-1)\cdot x + 0\cdot y + \frac{1}{2!0!}\cdot 0\cdot x^2 + 2\cdot\frac{1}{1!1!}\cdot 1\cdot xy + \frac{1}{0!2!}\cdot 1\cdot y^2 \\ &= 2 - x + xy + \frac{1}{2}y^2. \quad \square \end{aligned}$$

Taylorin kaava siis pätee, kunhan f:n osittaisderivaatat ovat kertalukuun $n+1$ asti jatkuvia tarkastelupisteen (x,y) ja pisteen (x_0,y_0) välisellä yhdysjanalla. Taylorin kaava on siten kahdenkin muuttujan tapauksessa johtoajatukseltaan yksiulotteinen. Jos halutaan, että kaava pätee tarkastelupisteen liikkuessa jossakin joukossa A pisteen (x_0,y_0) ympäristössä, niin onkin oletettava, f:n riittävän säännöllisyyden lisäksi, että A on pisteen (x_0,y_0) suhteen nk. *tähden muotoinen* (engl. star-shaped). Tämä tarkoittaa, että jokaiseen pisteeseen $(x,y) \in A$ on pisteestä (x_0,y_0) 'näköyhteys', eli pisteiden välinen yhdysjana on kokonaisuudessaan joukossa A. Näin ajatellen voidaan Taylorin kaavasta johtaa seuraava yleinen polynomiapproksimaatiotulos kahden muuttujan funktiolle.

Lause XI.9.1 Jos funktion $f = f(x,y)$ osittaisderivaatat ovat kertalukuun $n+1$ asti jatkuvia pisteen (x_0,y_0) suhteen tähden muotoisessa ja kompaktissa joukossa A, niin jokaisella $(x,y) \in A$ pätee

$$f(x,y) = T_n(x,y,x_0,y_0) + \mathcal{O}(h^{n+1}), \quad h = \sqrt{(x-x_0)^2 + (y-y_0)^2}.$$

Jos f:n osittaisderivaatat ovat A:ssa jatkuvia kertalukuun n asti, niin

$$f(x,y) = T_n(x,y,x_0,y_0) + o(h^n), \quad (x,y) \in A.$$

Todistus Jos f:n osittaisderivaatat ovat kertalukuun $n+1$ jatkuvia kompaktissa joukossa A, niin ne ovat myös rajoitettuja (ks. Lause XI.1.11):

$$|\partial^\alpha f(x,y)| \le C_\alpha \quad \forall\,(x,y) \in A, \quad |\alpha| \le n+1.$$

Valitsemalla $C = \max_{|\alpha|=n+1} C_\alpha$ seuraa Taylorin kaavan jäännöstermille arvio

$$\begin{aligned} |R_n(x,y)| &\le \frac{C}{(n+1)!}\sum_{l=0}^{n+1}\binom{n+1}{l}|x-x_0|^{n+1-l}|y-y_0|^l \\ &= \frac{C}{(n+1)!}\,(|x-x_0| + |y-y_0|)^{n+1} \\ &\le \frac{C(\sqrt{2})^{n+1}}{(n+1)!}\,h^{n+1} = \mathcal{O}(h^{n+1}), \quad (x,y) \in A. \end{aligned}$$

Astetta heikompien säännöllisyysoletusten ollessa voimassa päätellään vastaavasti (vrt. Lause VII.4.4)

$$f(x,y) = T_{n-1}(x,y,x_0,y_0) + R_{n-1}(x,y) = T_n(x,y,x_0,y_0) + o(h^n). \quad \square$$

Tapauksessa $n = 1$ Lauseen XI.9.1 jälkimmäinen väittämä on: Jos f, f_x, f_y ovat jatkuvia pisteen (x_0, y_0) (tähden muotoisessa) ympäristössä, niin

$$\begin{aligned} f(x,y) &= T_1(x,y,x_0,y_0) + o(h) \\ &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) + o(h). \end{aligned}$$

Tulos on ennestään tuttu, sillä tehdyin oletuksin f on differentioituva pisteessä (x_0, y_0) (vrt. Luku XI.3). Lauseen XI.9.1 jälkimmäistä väittämää voidaan käyttää hyväksi myös määrättäessä f:n Taylorin polynomia: Väittämän mukaan riittää löytää mikä tahansa polynomi $p(x,y)$, jolle pätee

$$f(x,y) - p(x,y) = o(h^n), \quad h = \sqrt{(x-x_0)^2 + (y-y_0)^2}.$$

Tällainen polynomi on välttämättä yksikäsitteinen — siis $p(x,y) = T_n(x,y,x_0,y_0)$ (vrt. Propositio VII.4.5).

Esimerkki 1 (jatko) Esimerkin funktiolle pätee

$$\begin{aligned} f(x,y) &= 1 - x + \frac{1}{2}x^2 + \mathcal{O}(|x|^3) + xy + y^2 + (1 - \frac{1}{2}x^2)(1 - \frac{1}{2}y^2) + \mathcal{O}(x^4 + y^4) \\ &= 2 - x + xy + \frac{1}{2}y^2 + \mathcal{O}(|x|^3 + |y|^3) \\ &= 2 - x + xy + \frac{1}{2}y^2 + \mathcal{O}(h^3), \end{aligned}$$

joten myös tällä perusteella on

$$T_2(x,y,0,0) = 2 - x + xy + \frac{1}{2}y^2. \quad \square$$

Monen muuttujan Taylorin polynomit

Em. tarkastelut yleistyvät helposti kolmen ja useamman muuttujan funktioihin. Kolmen muuttujan funktion $f = f(x,y,z)$ Taylorin polynomi määritellään

$$\begin{gathered} T_n(x,y,z,x_0,y_0,z_0) = \sum_{k=0}^{n} \frac{1}{k!}(h_x\partial_x + h_y\partial_y + h_z\partial_z)^k f(x_0,y_0,z_0), \\ h_x = x - x_0, \quad h_y = y - y_0, \quad h_z = z - z_0. \end{gathered}$$

Yleisemmin jos $f = f(\mathbf{x}) = f(x_1, \ldots, x_d)$, niin f:n Taylorin polynomi astetta n pisteessä $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$ on

$$\boxed{T_n(\mathbf{x}, \mathbf{a}) = \sum_{k=0}^{n} \frac{1}{k!} \big[\big(\sum_{i=1}^{d} h_i \partial_i \big)^k f \big](\mathbf{a}), \quad h_i = x_i - a_i, \quad i = 1 \ldots d.}$$

Mainittakoon ilman tarkempia perusteluja, että tämä lauseke on purettavissa seuraavaan muotoon, joka ei ulkonäöltään juuri poikkea yhden muuttujan tilanteesta (vrt. myös tapaus $d = 2$ edellä):

$$\boxed{\begin{aligned} &T_n(\mathbf{x}, \mathbf{a}) = \sum_{|\alpha| \le n} \frac{1}{\alpha!} \partial^\alpha f(\mathbf{a}) \, (\mathbf{x} - \mathbf{a})^\alpha, \\ &\alpha! = \prod_{i=1}^{d} \alpha_i \,!, \quad \mathbf{x}^\alpha = \prod_{i=1}^{d} x_i^{\alpha_i}, \\ &\alpha = (\alpha_1, \ldots, \alpha_d), \quad \mathbf{x} = (x_1, \ldots, x_d). \end{aligned}}$$

Kaavassa summaus käy yli kaikkien erilaisten moni-indeksien (eli järjestettyjen indeksijoukkojen), joille $|\alpha| = \alpha_1 + \ldots \alpha_d \le n$.

Taylorin polynomi $p(\mathbf{x}) = T_n(\mathbf{x}, \mathbf{a})$ määräytyy aina myös yksikäsitteisesti polynomina astetta n, joka toteuttaa ehdot

$$\partial^\alpha p(\mathbf{a}) = \partial^\alpha f(\mathbf{a}), \quad |\alpha| \le n.$$

Sikäli kuin f ja f:n osittaisderivaatat kertalukuun n asti ovat jatkuvia pisteen $\mathbf{a}$ ympäristössä, niin riittää myös löytää (keinolla millä hyvänsä) polynomi p astetta n, jolle pätee

$$f(\mathbf{x}) = p(\mathbf{x}) + o(|\mathbf{x} - \mathbf{a}|^n).$$

Tällöin on $p(\mathbf{x}) = T_n(\mathbf{x}, \mathbf{a})$.

Polynomi $T_2(\mathbf{x}, \mathbf{a})$ – Hessen matriisi

Monen muuttujan Taylorin polynomeja lasketaan harvemmin astelukua $n = 2$ pidemmälle. Em. kaavan mukaan f:n toisen asteen Taylorin polynomi on

$$T_2(\mathbf{x}, \mathbf{a}) = f(\mathbf{a}) + \sum_{i=1}^{d} \partial_i f(\mathbf{a})(x_i - a_i) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_i \partial_j f(\mathbf{a})(x_i - a_i)(x_j - a_j).$$

Tämän voi esittää matriisimuodossa

$$T_2(\mathbf{x},\mathbf{a}) = f(\mathbf{a}) + [\boldsymbol{\nabla} f(\mathbf{a})]^T(\mathbf{x}-\mathbf{a}) + \frac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{H}(\mathbf{a})(\mathbf{x}-\mathbf{a}),$$

missä $\boldsymbol{\nabla} f(\mathbf{a})$ on gradientti ja $\mathbf{H}(\mathbf{a})$ on f:n toisen kertaluvun osittaisderivaatoista koostuva nk. *Hessen matriisi* (engl. Hessian) :

$$[\,\mathbf{H}(\mathbf{a})\,]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}), \quad i,j = 1\ldots d.$$

Esimerkki 2 Määrää funktion

$$f(x_1,\ldots,x_d) = \prod_{i=1}^{d} e^{c_i x_i}$$

Taylorin polynomi $T_2(\mathbf{x},\mathbf{0})$.

Ratkaisu Koska

$$f(\mathbf{0}) = 1, \quad \frac{\partial f}{\partial x_i}(\mathbf{0}) = c_i\,, \quad i = 1\ldots d, \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = c_i c_j\,, \quad i,j = 1\ldots d,$$

niin

$$\begin{aligned} T_2(\mathbf{x},\mathbf{0}) &= 1 + \left(\sum_{i=1}^{d} x_i \frac{\partial}{\partial x_i}\right) f(\mathbf{0}) + \frac{1}{2}\left(\sum_{i=1}^{d} x_i \frac{\partial}{\partial x_i}\right)^2 f(\mathbf{0}) \\ &= 1 + \sum_{i=1}^{d} c_i x_i + \frac{1}{2}\sum_{i,j=1}^{d} c_i c_j x_i x_j\,. \end{aligned}$$

Kun merkitään $\mathbf{c} = [c_1,\ldots,c_d]^T$ ja $\mathbf{H} = (c_i c_j) = \mathbf{c}\mathbf{c}^T$ (= f:n Hessen matriisi origossa), niin tulos saadaan muotoon

$$\begin{aligned} T_2(\mathbf{x},\mathbf{0}) &= 1 + \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} \\ &= 1 + \mathbf{c}^T\mathbf{x} + \frac{1}{2}(\mathbf{c}^T\mathbf{x})^2. \end{aligned}$$

Lyhyempää tietä samaan tulokseen tullaan, kun huomataan käyttää reaalifunktion $\exp(x)$ ominaisuuksia:

$$\begin{aligned} f(\mathbf{x}) = \exp(\mathbf{c}^T\mathbf{x}) &= 1 + \mathbf{c}^T\mathbf{x} + \frac{1}{2}(\mathbf{c}^T\mathbf{x})^2 + \mathcal{O}(|\mathbf{c}^T\mathbf{x}|^3) \\ &= 1 + \mathbf{c}^T\mathbf{x} + \frac{1}{2}(\mathbf{c}^T\mathbf{x})^2 + \mathcal{O}(|\mathbf{x}|^3). \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske seuraavien funktioiden Taylorin polynomit annettua astetta n annetussa pisteessä. Käytä tilaisuuden tullen apuna yhden muuttujan Taylorin polynomeja.

   a) $f(x,y) = 2x^4 - 5y^3 + 2xy^2,\ \ (0,0),\ \ n = 3$

   b) $f(x,y) = \sqrt{x^3 y}\,,\ \ (1,4),\ \ n = 4$

   c) $f(x,y) = \dfrac{1}{2 + x - 2y}\,,\ \ (2,1),\ \ n = 3$

   d) $f(x,y) = y^2 \ln x,\ \ (1,0),\ \ n = 4$

   e) $f(x,y) = \sin(xy),\ \ (0,0),\ \ n \in \mathbb{N}$

   f) $f(x,y) = e^{xy} \sin y,\ \ (0,0),\ \ n = 3$

   g) $f(x,y) = \cos(x + \sin y),\ \ (0,0),\ \ n = 4$

   h) $f(x,y) = \displaystyle\int_0^{x+y^2-1} e^{-t^2}\,dt + \int_0^{\pi/2} \cos(xt^2 + yt)\,dt,\ \ (0,1),\ \ n = 2$

   i) $x^2 + 2y^2 + 3z^2 = 6 \ \Leftrightarrow\ z = f(x,y),\ \ (1,1,1),\ \ n = 2$

   j) $x + xy^2 + 2z - \sin z = 0 \ \Leftrightarrow\ z = f(x,y),\ \ (0,0,0),\ \ n = 2$

   k) $f(x,y,z) = e^{xyz},\ \ (0,0,0),\ \ n \in \mathbb{N}$

   l) $f(x,y,z) = \dfrac{1}{x + y + z}\,,\ \ (1,0,0),\ \ n = 3$

   m) $f(\mathbf{x}) = \cos(x_1 + \ldots + x_n),\ \ \mathbf{0},\ \ n = 2$

   n) $f(\mathbf{x}) = \dfrac{1}{x_1 + \ldots + x_n}\,,\ \ (1,\ldots,1),\ \ n = 2$

2. Laske funktion $f(x,y) = \ln(1 + x^2 + y^2)$ osittaisderivaatta $f_{xxxxxxyy}(0,0)$ Taylorin polynomien avulla.

3. Arvioi, kuinka suuri on seuraavien approksimaatioiden virhe enintään kiekossa $A:\ x^2 + y^2 \le a^2\ (a > 0)$.

   a) $e^{x+2y} \approx 1 + x + 2y$

   b) $e^{x+2y} \approx 1 + x + 2y + \dfrac{1}{2}x^2 + 2xy + 2y^2$

   b) $\cos(xy) \approx 1 - \dfrac{1}{2}x^2y^2$

   d) $2xy + 2\cos(x - y) \approx 2 + x^2 + y^2$

4. Määritä funktion toisen asteen Taylorin polynomi ja sen avulla Hessen matriisi origossa:

a) $x\sin(x+y)$	b) $xy\cos(x-y)$	c) $e^{x+y}+e^{x-y}$
d) $\dfrac{1-x+x^2}{1+y-y^2}$	e) $\cos(x+y-2z)$	f) $\dfrac{e^{xy}+e^{yz}+e^{xz}}{1+x+y+z}$

5. Yhtälö $x^y = y^x \;\Leftrightarrow\; y\ln x = x\ln y$ määrittelee käyrän, mahdollisesti useampia, pisteen (a,a) $(a>0)$ ympäristössä. Tutki asiaa approksimoimalla funktiota $f(x,y) = y\ln x - x\ln y$ ensimmäisen, toisen ja kolmannen asteen Taylorin polynomilla pisteessä (a,a).

Luku XII

Matriisin ominaisarvot ja neliömuodot

Matriiseja ja matriisialgebraa on aiemmin (Luku X) tarkasteltu lähinnä lineaaristen yhtälöryhmien teorian ja ratkaisemisen näkökulmasta. Tässä luvussa avataan matriisilaskuun ja sen sovelluksiin kokonaan uusi näkökulma, kun tarkastelun kohteeksi otetaan matriisiin liitettävät erityiset luvut ja vektorit, nimeltään *ominaisarvot* ja *ominaisvektorit*.

Luvussa XII.1 määritellään matriisin ominaisarvot ja -vektorit, eli yhdistettynä *ominaisparit*, sekä esitetään ominaisarvoteorian keskeisimmät väittämät. Osoittautuu, että yleisen reaalisen matriisin ominaisarvot ja -vektorit voivat olla kompleksisia. Symmetrisen matriisin ominaisarvo-ongelma sen sijaan ratkeaa kokonaan $\mathbb{R}^n$:ssä, ja osoittautuu, että tässä tapauksessa ominaisvektoreista voidaan myös muodostaa $\mathbb{R}^n$:n ortonormeerattu kanta. Tähän perustuva matriisin *diagonalisointi* on keskeisellä sijalla Luvussa XII.2, jossa symmetrisen matriisin ominaisarvoteoriaa sovelletaan $\mathbb{R}^n$:n neliömuotoihin. Tällä tavoin ensinnäkin ratkeaa (ainakin osittain) Luvussa XI.8 avoimeksi jätetty ongelma: Onko funktion $f(\mathbf{x})$ kriittinen piste paikallinen minimi, maksimi vai ei kumpikaan? Neliömuotoihin liittyen ratkaistaan Luvussa XII.2 myös toisen asteen käyrien ja ja toisen asteen pintojen geometrisen luokittelun ongelmat matriisien ominaisarvoteorian avulla.

Luvussa XII.3 osoitetaan, että symmetriset matriisit ja niiden ominaisarvoteoria ovat avaimia myös *pinnan kaarevuuden* käsitteeseen. Viimeisessä osaluvussa esitellään vielä lyhyesti käsite *tensori*, jollaiseksi esim. pinnan kaarevuus on ymmärrettävissä. Tensori ja sen erilaiset esiintymismuodot johdattavat pohtimaan uudelleen myös vektorin ja skalaarin käsitteitä.

XII.1 Matriisin ominaisarvot

Aloitetaan määritelmästä.

MÄÄRITELMÄ XII.1.1 Jos $\mathbf{A}$ on neliömatriisi kokoa $n \times n$, niin kompleksiluku $\lambda \in \mathbb{C}$ on $\mathbf{A}$:n *ominaisarvo* (engl. eigenvalue, ruots. egenvärde, saks. Eigenwert), ja vektori $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x} \neq \mathbf{0}$, on ominaisarvoon λ liittyvä *ominaisvektori* (engl. eigenvector), jos pätee

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Määritelmässä voi $\mathbf{A}$ olla yhtä hyvin reaalinen kuin kompleksinen matriisi. Määritelmä siis <u>ei</u> yksinkertaistu muotoon $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ siinäkään tapauksessa, että $\mathbf{A}$ on reaalinen.

Määritelmästä nähdään heti, että ominaisvektorit eivät ole yksikäsitteiset: Jos $\mathbf{x}$ on ominaisvektori, niin samoin on $\alpha\mathbf{x}$ jokaisella $\alpha \in \mathbb{C}$, $\alpha \neq 0$. Ominaisvektorit voidaan siis *normeerata* mielivaltaisella tavalla, esimerkiksi ehdolla $|\mathbf{x}| = 1$.

Matriisin ominaisarvot ja -vektorit, eli *ominaisparit* (engl. eigenpair) määrittelevää yhtälöä $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ sanotaan *ominaisarvoyhtälöksi*. Ominaisparien laskeminen aloitetaan määrämällä tästä yhtälöstä ensin kaikki ominaisarvot. Ominaisarvojen joukkoa, jota sanotaan $\mathbf{A}$:n *spektriksi*, merkitään jatkossa symbolilla $\sigma(\mathbf{A})$:

$$\sigma(\mathbf{A}) = \{\mathbf{A}\text{:n ominaisarvot}\} \subset \mathbb{C}.$$

Spektrin avulla saadaan matriisin säännöllisyydelle jälleen uusi luonnehdinta:

$$\boxed{\mathbf{A} \text{ säännöllinen} \quad \Leftrightarrow \quad 0 \notin \sigma(\mathbf{A}).}$$

Kun ominaisarvoyhtälö kirjoitetaan muotoon

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

($\mathbf{I}$ = yksikkömatriisi), niin matriisialgebran ja determinanttiopin mukaan tällä on ei-triviaali ratkaisu, eli $\lambda \in \sigma(\mathbf{A})$, täsmällen kun

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

Tämä on siis *ominaisarvoehto*. Koska λ esiintyy tässä ainoastaan $(\mathbf{A} - \lambda\mathbf{I})$:n diagonaalilla, niin alideterminanttisäännön (Lause X.5.5) perusteella nähdään, että ominaisarvoehto purkautuu yhtälöksi muotoa

$$p(\lambda) = 0,$$

missä $p(\lambda)$ on polynomi astetta n, tarkemmin

$$p(\lambda) = (-1)^n \lambda^n + \{\text{alempiasteisia termejä}\}.$$

Ominaisarvoehtoa $p(\lambda) = 0$ sanotaan $\mathbf{A}$:n *karakteristiseksi yhtälöksi* ja polynomia $p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$ $\mathbf{A}$:n *karakteristiseksi polynomiksi.* Matriisin $\mathbf{A}$ ominaisarvot ovat siis $\mathbf{A}$:n karakteristisen polynomin juuret, ja näin ollen $\mathbf{A}$:n spektri $\sigma(\mathbf{A})$ koostuu vähintään yhdestä ja enintään n:stä kompleksiluvusta (Algebran peruslause). Jos $\mathbf{A}$ on reaalinen matriisi, niin karakteristinen polynomi $p(\lambda)$ on reaalikertoiminen, jolloin juuret ovat joko reaalisia tai muodostavat konjugaattipareja. Jos $\lambda \in \sigma(\mathbf{A})$ on $p(\lambda)$:n m-kertainen juuri, niin lukua $m_a(\lambda) = m$ sanotaan ko. ominaisarvon *algebralliseksi kertaluvuksi.* Algebran peruslauseen mukaan on

$$\sum_{\lambda \in \sigma(\mathbf{A})} m_a(\lambda) = n.$$

Jos $m_a(\lambda) = 1$, on kyseessä *yksinkertainen* (engl. simple) ominaisarvo.

Kun ominaisarvo $\lambda \in \sigma(\mathbf{A})$ on löydetty (mahdollisesti numeerisin apukeinoin), saadaan vastaavat ominaisvektorit selville ominaisarvoyhtälöstä $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, joka ominaisarvon määritelmän mukaisesti on singulaarinen yhtälöryhmä $\mathbb{C}^n$:ssä. Kun tämä yhtälöryhmä saatetaan Gaussin algoritmilla singulaariseen perusmuotoonsa, saadaan kaikki ratkaisut selville. Yleisesti jos $\mathbf{x}_1 \in \mathbb{C}^n$ ja $\mathbf{x}_2 \in \mathbb{C}^n$ ovat ominaisarvoyhtälön ratkaisuja, niin samoin on $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ $\forall \alpha_1, \alpha_2 \in \mathbb{C}$, joten tiettyyn ominaisarvoon liittyvät ominaisvektorit muodostavat $\mathbb{C}^n$:n aliavaruuden. Tätä aliavaruutta sanotaan ko. ominaisarvoon liittyväksi *ominaisavaruudeksi* (engl. eigenspace). Jos ominaisavaruuden dimensio on ν, eli jos samaan ominaisarvoon λ voidaan liittää täsmälleen ν lineaarisesti riippumatonta ominaisvektoria, niin sanotaan, että luku ν (aina $\nu \geq 1$) on ko. ominaisarvon *geometrinen kertaluku*, ja merkitään

$$m_g(\lambda) = \nu.$$

Jokaiseen ominaisarvoon $\lambda \in \sigma(\mathbf{A})$ liittyy siis kaksi kertalukua, algebrallinen kertaluku $m_a(\lambda)$ ja geometrinen kertaluku $m_g(\lambda)$, jotka molemmat ovat luonnollisia lukuja. Voidaan myös osoittaa, että aina pätee

$$m_g(\lambda) \leq m_a(\lambda), \quad \lambda \in \sigma(\mathbf{A}).$$

Esimerkki 1 Määritä matriisin

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}$$

ominaisarvot ja ominaisvektorit.

<u>Ratkaisu</u> Ominaisarvojen determinanttiehto on

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 6-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{vmatrix}$$
$$= (6-\lambda)(\lambda^2-4\lambda-12) = -(6-\lambda)^2(\lambda+2) = 0$$

(determinantti purettu toisen vaaka- tai pystyrivin mukaan), joten ominaisarvot ja niiden algebralliset kertaluvut ovat

$$\lambda_1 = 6, \quad m_a(\lambda_1) = 2, \qquad \lambda_2 = -2, \quad m_a(\lambda_2) = 1.$$

Ominaisvektorit saadaan ratkaisemalla ominaisarvoyhtälö kummallekin ominaisarvolle erikseen:

$$(\mathbf{A}-\lambda_1\mathbf{I})\mathbf{x} = \mathbf{0} \;\Leftrightarrow\; \begin{bmatrix} -4 & 0 & 4 \\ 0 & 0 & 0 \\ 4 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\Leftrightarrow\; \begin{bmatrix} -4 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}.$$

$$(\mathbf{A}-\lambda_2\mathbf{I})\mathbf{x} = \mathbf{0} \;\Leftrightarrow\; \begin{bmatrix} 4 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\Leftrightarrow\; \begin{bmatrix} 4 & 0 & 4 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}.$$

Nähdään, että ominaisarvoon λ_1 liittyy kaksi lineaarisesti riippumatonta ominaisvektoria, esimerkiksi

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

ja ominaisarvoon λ_2 liittyy yksi lineaarisesti riippumaton ominaisvektori, esimerkiksi

$$\mathbf{x}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Ominaisarvojen geometriset kertaluvut ovat näin ollen

$$m_g(\lambda_1) = 2, \quad m_g(\lambda_2) = 1. \quad \square$$

Esimerkki 2 Laske matriisin

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 1 & -1 \end{bmatrix}$$

ominaisarvot ja ominaisvektorit.

Ratkaisu Karakteristinen polynomi on

$$p(\lambda) = (1 - \lambda)(\lambda^2 + 1),$$

joten ominaisarvot ovat

$$\lambda_1 = 1, \quad \lambda_2 = i, \quad \lambda_3 = -i.$$

Samaan tapaan laskien kuin edellisessä esimerkeissä saadaan ominaisvektoreiksi

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1+i \\ 1+i \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1-i \\ 1-i \\ 1 \end{bmatrix}.$$

Ominaisarvojen kertaluvut ovat

$$m_a(\lambda_i) = m_g(\lambda_i) = 1, \quad i = 1, 2, 3. \quad \square$$

Esimerkki 3 Laske diagonaalimatriisin $\mathbf{A} = \operatorname{diag}(d_i,\ i = 1 \ldots n)$ ominaisarvot ja ominaisvektorit.

Ratkaisu Karakteristinen yhtälö on

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \prod_{i=1}^{n} (d_i - \lambda) = 0.$$

Jos $d_i \neq d_j$, kun $i \neq j$, niin ominaisarvot ja niiden algebrallisest kertaluvut ovat

$$\lambda_i = d_i, \quad m_a(\lambda_i) = 1, \quad i = 1 \ldots n.$$

Jos $d_i = d\ \forall i$ (jolloin $\mathbf{A} = d\mathbf{I}$), niin $\mathbf{A}$:lla on yksi n-kertainen ominaisarvo:

$$\lambda = d, \quad m_a(\lambda) = n.$$

Jos yleisemmin luku d esiintyy $\mathbf{A}$:n diagonaalilla täsmälleen m kertaa, niin ominaisarvon $\lambda = d$ algebrallinen kertaluku on $m_a(\lambda) = m$. Myös geometrinen kertaluku on $m_g(\lambda) = m$, sillä ominaisarvoyhtälöstä nähdään, että ominaisvektoreita ovat $\mathbb{R}^n$:n peruskantavektorit $\mathbf{e}_i$ jokaisella i, jolla $d_i = d$. Erityisesti jos $d_i = d\ \forall i$, niin jokainen vektori $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$, on ominaisvektori. $\square$

Tähänastisissa esimerkeissä ovat omainaisarvojen algebralliset ja geometriset kertaluvut olleet samat. Seuraavasta esimerkistä nähdään, että näin ei ole aina.

Esimerkki 4 Määrää $\mathbf{A}$:n ominaisarvot ja ominaisvektorit, kun

$$\mathbf{A} = \begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & a & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{bmatrix}$$

eli

$$A = (a_{ij})_{i,j=1}^n, \quad a_{ij} = \begin{cases} a, & \text{kun } i = j, \\ 1, & \text{kun } j = i+1, \\ 0, & \text{muulloin.} \end{cases}$$

<u>Ratkaisu</u> Kyseessä on kolmiomatriisi, jonka karakteristinen yhtälö on

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (a - \lambda)^n = 0.$$

Tämän mukaan $\lambda = a$ on n-kertainen ominaisarvo ($m_a(\lambda) = n$). Kun $\lambda = a$, niin ominaisarvoyhtälöt ovat

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \ \Leftrightarrow \ \begin{cases} x_2 = 0 \\ x_3 = 0 \\ \quad\vdots \\ x_n = 0 \end{cases}$$

joten $\mathbf{A}$:lla on vain yksi lineaarisesti riippumaton ominaisvektori $\mathbf{x} = [1, 0, \ldots, 0]^T$. Siis ainoan ominaisarvon geometrinen kertaluku on $m_g(\lambda) = 1$. $\square$

Todetaan tässä yhteydessä vielä matriisin ominaisvektoreiden lineaarista riippuvuutta rajoittava yleinen tulos.

Lause XII.1.2 Matriisin eri ominaisarvoihin liittyvät ominaisvektorit ovat keskenään lineaarisesti riippumattomat.

Todistus Olkooon $\mathbf{A}\mathbf{a}_i = \lambda_i \mathbf{a}_i$, $i = 1 \ldots m$, missä $\mathbf{a}_i \neq \mathbf{0}$ ja $\lambda_i \neq \lambda_j$, kun $i \neq j$. Kun merkitään

$$\mathbf{B} = (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{A} - \lambda_2 \mathbf{I}) \cdots (\mathbf{A} - \lambda_{m-1} \mathbf{I}) = \prod_{k=1}^{m-1} (\mathbf{A} - \lambda_k \mathbf{I}),$$

niin oletusten perusteella

$$\mathbf{B}\mathbf{a}_j = \prod_{k=1}^{m-1} (\lambda_j - \lambda_k)\, \mathbf{a}_j = \begin{cases} \mathbf{0}, & j = 1 \ldots m-1, \\ \prod_{k=1}^{m-1} (\lambda_m - \lambda_k)\, \mathbf{a}_m, & j = m. \end{cases}$$

Oletetaan nyt, että

$$\sum_{j=1}^{m} x_j \mathbf{a}_j = \mathbf{0}.$$

Tällöin seuraa

$$\mathbf{0} = \mathbf{B} \sum_{j=1}^{m} x_j \mathbf{a}_j \;=\; \sum_{j=1}^{m} x_j \mathbf{B}\mathbf{a}_j \;=\; \prod_{k=1}^{m-1} (\lambda_m - \lambda_k) x_m \mathbf{a}_m\,,$$

joten on oltava $x_m = 0$. Siis

$$\sum_{j=1}^{m-1} x_j \mathbf{a}_j = \mathbf{0}.$$

Jos $m = 2$, niin tämä yhtälö toteutuu vain kun $x_1 = 0$, jolloin on päätelty, että $x_1 = x_2 = 0$. Jos $m > 2$, niin kertomalla yhtälö puolittain matriisilla

$$\mathbf{B} = \prod_{k=1}^{m-2} (\mathbf{A} - \lambda_k \mathbf{I})$$

päätellään samalla tavoin kuin edellä, että $x_{m-1} = 0$, jne. Siis vektorit $\mathbf{a}_1, \ldots, \mathbf{a}_m$ ovat lineaarisesti riippumattomat:

$$\sum_{j=1}^{m} x_j \mathbf{a}_j = \mathbf{0} \quad \Rightarrow \quad x_m = x_{m-1} = \ldots = x_1 = 0. \quad \square$$

Lauseella XII.1.2 on seuraava ilmeinen seuraamus:

Korollaari XII.1.3 Jos $\mathbf{A}$ on kokoa $n \times n$ ja kaikki $\mathbf{A}$:n ominaisarvot ovat yksinkertaisia, niin $\mathbf{A}$:n ominaisvektoreista voidaan muodostaa $\mathbb{C}^n$:n kanta.

Symmetrinen matriisi

Reaalisen ja symmetrisen matriisin ominaisarvoteoria voidaan pelkistää seuraavaksi kauniiksi lauseeksi. Lauseen täydellinen todistus sopii paremmin lineaarialgebran laajemman esittelyn yhteyteen, joten rajoitutaan tässä todistamaan vain kaksi lauseeseen sisältyvää (helpompaa) osaväittämää.

LAUSE XII.1.4 Olkoon $\mathbf{A}$ reaalinen ja symmetrinen matriisi kokoa $n \times n$. Tällöin $\mathbf{A}$:n ominaisarvot ovat reaaliset, ominaisarvojen algebralliset ja geometriset kertaluvut ovat samat, ja $\mathbf{A}$:n ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n ortonormeerattu kanta.

Todistus (osittainen) Jos

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \lambda \in \mathbb{C},\ \mathbf{x} \in \mathbb{C}^n,\ \mathbf{x} \neq \mathbf{0},$$

niin

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \lambda \langle \mathbf{x}, \mathbf{x} \rangle = \lambda |\mathbf{x}|^2,$$

missä $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^n$:n euklidinen skalaaritulo ja $|\cdot|$ on euklidinen normi:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i \overline{y}_i, \quad |\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Konjugoimalla saatu yhtälö puolittain saadaan

$$\overline{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle} = \overline{\lambda |\mathbf{x}|^2} = \overline{\lambda} |\mathbf{x}|^2.$$

Mutta $\mathbf{A}$:n oletetun reaalisuuden ja symmetrian perusteella on (vrt. Luku X.1)

$$\overline{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle} = \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle,$$

joten $\lambda |\mathbf{x}|^2 = \overline{\lambda} |\mathbf{x}|^2$. Tässä on $|\mathbf{x}|^2 \in \mathbb{R}_+$, joten $\lambda = \overline{\lambda}$, eli $\lambda \in \mathbb{R}$.

Olkoot λ ja μ $\mathbf{A}$:n kaksi eri suurta ominaisarvoa ja $\mathbf{x}$ ja $\mathbf{y}$ vastaavat ominaisvektorit:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad \& \quad \mathbf{A}\mathbf{y} = \mu\mathbf{y}.$$

Jo todetun perusteella on $\lambda, \mu \in \mathbb{R}$, joten seuraa

$$\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \lambda\mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle,$$
$$\langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{x}, \mu\mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle.$$

Mutta koska $\mathbf{A}$ on reaalinen ja symmetrinen, niin

$$\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle,$$

joten seuraa

$$(\lambda - \mu)\langle \mathbf{x}, \mathbf{y} \rangle = 0 \;\Rightarrow\; \langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

Yhdistämällä tähänastiset päätelmät on todistettu, että $\mathbf{A}$:n ominaisarvot ovat reaaliset, ja että eri ominaisarvoihin liittyvät ominaisvektorit ovat ortogonaaliset. Todistuksen loppuosa sivuutetaan. $\square$

Lauseen XII.1.4 tuloksesta seuraa, että reaalisen ja symmetrisen matriisin tapauksessa ei kompleksilukuja tarvita myöskään ominaisvektoreissa. Nimittäin jos $\mathbf{x} \in \mathbb{C}^n$ on ominaisvektori ja $\mathbf{x} = \mathbf{u} + i\mathbf{v}$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, niin $\mathbf{A}$:n, λ:n, $\mathbf{u}$:n ja $\mathbf{v}$:n reaalisuuden perusteella

$$\begin{aligned} \mathbf{A}\mathbf{x} = \lambda\mathbf{x} \;&\Leftrightarrow\; \mathbf{A}\mathbf{u} + i\mathbf{A}\mathbf{v} = \lambda\mathbf{u} + i\lambda\mathbf{v} \\ &\Leftrightarrow\; \mathbf{A}\mathbf{u} = \lambda\mathbf{u} \;\;\&\;\; \mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \end{aligned}$$

Reaalisen ja symmetrisen matriisin ominaisarvoyhtälö $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ voidaan siis ratkaista kokonaan $\mathbb{R}^n$:ssä.

Esimerkki 5 Symmetrisen matriisin

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}$$

ortonormeeratut ominaisvektorit ovat (vrt. Esimerkki 1)

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}}[1, 0, 1]^T, \quad \mathbf{x}_2 = [0, 1, 0]^T, \quad \mathbf{x}_3 = \frac{1}{\sqrt{2}}[-1, 0, 1]^T. \quad \square$$

Diagonalisoituva matriisi

Olkoon $\mathbf{A}$ reaalinen ja symmetrinen matriisi kokoa $n \times n$ ja olkoon $\mathbf{C}$ reaalinen matriisi, jonka sarakkeina ovat $\mathbf{A}$:n ortonormeeratut ominaisvektorit:

$$\mathbf{C} = [\mathbf{c}_1 \ldots \mathbf{c}_n], \quad \mathbf{A}\mathbf{c}_i = \lambda_i \mathbf{c}_i, \quad i = 1 \ldots n, \quad \langle \mathbf{c}_i, \mathbf{c}_j \rangle = \delta_{ij}.$$

Tässä voidaaan ominaisarvoyhtälöt $\mathbf{A}\mathbf{c}_i = \lambda_i \mathbf{c}_i$ koota matriisiyhtälöksi

$$[\mathbf{A}\mathbf{c}_1 \ldots \mathbf{A}\mathbf{c}_n] \;=\; [\lambda_1 \mathbf{c}_1 \ldots \lambda_n \mathbf{c}_n] \;=\; [\mathbf{C}(\lambda_1 \mathbf{e}_1) \ldots \mathbf{C}(\lambda_n \mathbf{e}_n)]$$

eli (vrt. Luku X.1)

$$\mathbf{A}\mathbf{C} = \mathbf{C}\boldsymbol{\Lambda},$$

missä $\mathbf{\Lambda} = [\lambda_1 \mathbf{e}_1 \dots \lambda_n \mathbf{e}_n] = \mathrm{diag}(\lambda_i)$. Koska $\mathbf{C}$ on ortogonaalinen, on $\mathbf{C}^{-1} = \mathbf{C}^T$, joten seuraa

$$\boxed{\mathbf{C}^T\mathbf{A}\mathbf{C} = \mathbf{\Lambda} \quad \Leftrightarrow \quad \mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^T.}$$

Tässä tuloksessa siis $\mathbf{C}$:n sarakkeina ovat $\mathbf{A}$:n <u>ortonormeeratut</u> ominaisvektorit jossakin järjestyksessä ja $\mathbf{\Lambda}$ on diagonaalimatriisi, jonka lävistäjällä ovat $\mathbf{A}$:n ominaisarvot $\mathbf{C}$:n määräämässä järjestyksessä siten, että $\mathbf{A}\mathbf{c}_i = \lambda \mathbf{c}_i$, $\lambda_i = [\mathbf{\Lambda}]_{ii}$. Annetaan tässä yhteydessä hieman yleisempi määritelmä.

Määritelmä XII.1.5 (**Similariteettimuunnos**) Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat samankokoisia neliömatriiseja, ja on olemassa säännöllinen matriisi $\mathbf{C}$ siten, että $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$, niin sanotaan, että $\mathbf{A}$ ja $\mathbf{B}$ ovat *similaariset*, ja kuvausta $\mathbf{A} \mapsto \mathbf{B}$ sanotaan *similariteettimuunnokseksi*. Similariteettimuunnos on *ortogonaalinen*, jos $\mathbf{C}$ on ortogonaalinen matriisi. Jos $\mathbf{A}$ on similaarinen diagonaalimatriisin kanssa, niin sanotaan, että $\mathbf{A}$ on *diagonalisoituva*.

Reaalinen ja symmetrinen matriisi on siis määritelmän mukaisesti diagonalisoituva, ja tässä tapauksessa matriisi $\mathbf{C}$ on myös valittavissa niin, että diagonalisoiva similariteettimuunnos on ortogonaalinen. Yleisemmin voidaan $\mathbf{C}$:n sarakkeiksi valita mitkä tahansa n lineaarisesti riippumatonta ominaisvektoria. Tällaisen ominaisvektorisysteemin olemassaolo on myös välttämätön ehto sille, että matriisi on diagonalisoituva:

Lause XII.1.6 Matriisi $\mathbf{A}$ kokoa $n \times n$ on diagonalisoituva täsmälleen kun $\mathbf{A}$:lla on n lineaarisesti riippumatonta ominaisvektoria. Tällöin jos matriisin $\mathbf{C}$ sarakkeina ovat mitkä tahansa $\mathbf{A}$:n lineaarisesti riippumattomat ominaisvektorit ja diagonaalimatriisin $\mathbf{\Lambda}$ lävistäjäalkioina ovat vastaavat ominaisarvot, niin $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$.

<u>Todistus</u> Jos $\mathbf{A}$:lla on lineaarisesti riippumattomat ominaisvektorit $\mathbf{c}_i$, $i = 1 \dots n$, niin muodostamalla nämä sarakkeina matriisi $\mathbf{C} = [\mathbf{c}_1 \dots \mathbf{c}_n]$ todetaan kuten edellä, että $\mathbf{A}\mathbf{C} = \mathbf{C}\mathbf{\Lambda} \;\Leftrightarrow\; \mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$, missä $\mathbf{\Lambda} = \mathrm{diag}\,(\lambda_i)$ ja $\lambda_i =$ ominaisvektoriin $\mathbf{c}_i$ liittyvä ominaisarvo. Tämä todistaa väittämän osan $\boxed{\Leftarrow}$. Osan $\boxed{\Rightarrow}$ todistamiseksi olkoon $\mathbf{A} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$, missä $\mathbf{C}$ on säännöllinen ja $\mathbf{D} = \mathrm{diag}\,(d_i)$ on diagonaalimatriisi. Tällöin pätee

$$\mathbf{D}\mathbf{e}_i = d_i\mathbf{e}_i \quad \Leftrightarrow \quad (\mathbf{C}^{-1}\mathbf{D}\mathbf{C})(\mathbf{C}^{-1}\mathbf{e}_i) = d_i\mathbf{C}^{-1}\mathbf{e}_i, \quad i = 1 \dots n,$$

joten matriisilla $\mathbf{A} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$ on ominaisvektorit $\mathbf{c}_i = \mathbf{C}^{-1}\mathbf{e}_i$ ja vastaavat ominaisarvot d_i. Vektorit $\mathbf{c}_i$ ovat lineaarisesti riippumattomat, sillä

$$\sum_{i=1}^{n} x_i\mathbf{c}_i = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{C}\sum_{i=1}^{n} x_i\mathbf{c}_i = \sum_{i=1}^{n} x_i\mathbf{e}_i = \mathbf{0} \quad \Leftrightarrow \quad x_i = 0,\ i = 1 \dots n.$$

Siis $\{d_i,\ i = 1 \ldots n\} \subset \sigma(\mathbf{A})$ ja $\mathbf{A}$:lla on n lineaarisesti riippumatonta ominaisvektoria vastaten näitä ominaisarvoja. Tällöin on myös oltava $\{d_i,\ i = 1 \ldots n\} = \sigma(\mathbf{A})$, sillä muu johtaisi ristiriitaan Lauseen XII.1.2 kanssa. $\square$

Esimerkki 6 Matriisilla

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

on vain yksi ominaisvektori (Esimerkki 4), joten $\mathbf{A}$ ei ole diagonalisoituva (Lause XII.1.6). $\square$

Esimerkki 7 Määrittele diagonalisoiva muunnos $\mathbf{A} \mapsto \mathbf{D} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ matriisille

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}.$$

<u>Ratkaisu</u> saadaan suoraan edellisen luvun Esimerkeistä 1 ja 5:

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Tässä on valittu ortogonaalinen muunnos. Mahdollista on myös valita $\mathbf{C}$:n sarakkeiksi esim. vektorit $\mathbf{c}_1 = [1, 1, 1]^T$, $\mathbf{c}_2 = [1, 0, -1]^T$ ja $\mathbf{c}_3 = [0, 1, 0]^T$, sillä nämäkin ovat $\mathbf{A}$:n lineaarisesti riippumattomia ominaisvektoreita (vastaten ominaisarvoja $\lambda_1 = 6$, $\lambda_2 = -2$ ja $\lambda_3 = 6$, ks. edellisen luvun Esimerkki 1). Kysytyksi muunnokseksi saadaan tällä valinnalla

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \qquad \square$$

HARJOITUSTEHTÄVIÄ

1. Todista:
   a) Jos $\lambda \in \sigma(\mathbf{A})$, niin $\lambda^p \in \sigma(\mathbf{A}^p)\ \forall p \in \mathbb{N}$.
   b) Jos $\mathbf{A}$ on säännöllinen ja $\lambda \in \sigma(\mathbf{A})$, niin $\lambda^{-1} \in \sigma(\mathbf{A}^{-1})$.
   c) Jos $\mathbf{A}$ on reaalinen, niin $\sigma(\mathbf{A}^T) = \sigma(\mathbf{A})$.
   d) Jos $1 \in \sigma(\mathbf{A})$ tai $2 \in \sigma(\mathbf{A})$, niin $\mathbf{B} = 2\mathbf{I} - 3\mathbf{A} + \mathbf{A}^2$ on singulaarinen.

2. Määritä seuraavien matriisien ominaisarvot ja -vektorit sekä ominaisarvojen algebralliset ja geometriset kertaluvut.

a) $\begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 4 & 4 \\ -1 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ e) $\begin{bmatrix} -2 & 1 & 1 \\ 0 & 0 & 4 \\ 0 & -1 & -4 \end{bmatrix}$ f) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 1 & 0 & 3 \end{bmatrix}$

g) $\begin{bmatrix} 2 & 2 & 3 \\ 4 & 1 & 0 \\ 1 & -2 & 0 \end{bmatrix}$ h) $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 9 & 0 & 3 & 0 \\ 1 & 0 & 0 & 4 \end{bmatrix}$ i) $\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 3 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

3. Olkoon

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & t & t \\ t & 1 & t \\ t & t & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \alpha & 1 & 1 & \beta \\ 4 & \alpha & \beta & 2 \\ 4 & \beta & \alpha & 2 \\ \beta & 3 & 3 & \alpha \end{bmatrix}.$$

a) Matriisin $\mathbf{A}$ ominaisvektoreita ovat $[1,1,1]^T$, $[1,0,1]^T$ ja $[0,0,1]^T$. Laske tätä tietoa hyväksi käyttäen $\mathbf{A}^{11}\mathbf{x}$, kun $\mathbf{x} = [2,1,1]^T$.

b) Matriisin $\mathbf{B}$ ominaisarvot ovat reaalifunktioita $t \mapsto \lambda_i(t)$. Hahmottele näiden kuvaajat!

c) Matriisilla $\mathbf{C}$ on ominaisarvo λ ja vastaava ominaisvektori $[0,1,-1,0]^T$. Millainen ehto tästä seuraa luvuille α, β ja λ?

4. a) Näytä, että jos neliömatriisille pätee $\mathbf{A}^2 = \mathbf{A}$, niin $\sigma(\mathbf{A}) \subset \{0,1\}$. Näytä esimerkillä, että vaihtoehdot $\sigma(\mathbf{A}) = \{0\}$, $\sigma(\mathbf{A}) = \{1\}$ ja $\sigma(\mathbf{A}) = \{0,1\}$ ovat kaikki mahdollisia.

b) Lineaarikuvaus $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$, missä $\mathbf{A}$ on kokoa 3×3, vastaa projektiota tasolle $T: x + 3y - 2z = 0$ suoran $S: x = y = z$ sunnassa. Näytä, että $\sigma(\mathbf{A}) = \{0,1\}$ ja että $\mathbf{A}$:lla on kolme lineaarisesti riippumatonta ominaisvektoria. Mitkä ovat ominaisarvojen kertaluvut? Päteekö myös $\mathbf{A}^2 = \mathbf{A}$?

c) Olkoon $\mathbf{A} = \mathbf{a}\mathbf{a}^T$, missä $\mathbf{a} \in \mathbb{R}^n$, $n \geq 2$ ja $|\mathbf{a}| = 1$. Näytä, että $\sigma(\mathbf{A}) = \{0,1\}$. Mitkä ovat ominaisarvojen geometriset kertaluvut?

5. Matriisilla kokoa 3×3 on ominaisuus: Matriisin ominaisarvot ovat $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 2$ ja vastaavat ominaisvektorit ovat $[1,1,1]^T$, $[1,1,-1]^T$ ja $[0,1,2]^T$. Montako erilaista nämä ehdot täyttävää matriisia on olemassa? Määritä yksi!

6. Tutki, ovatko seuraavat matriisit $\mathbf{A}$ diagonalisoituvia. Myönteisessä tapauksessa laske matriisille jokin (reaalinen tai kompleksinen) tulohajotelma muotoa $\mathbf{A} = \mathbf{CDC}^{-1}$, missä $\mathbf{D}$ on diagonaalinen.

a) $\begin{bmatrix} 1 & 0 \\ 4 & -3 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}$ c) $\begin{bmatrix} -3 & 2 \\ -2 & 1 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 2 \\ -4 & -3 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ -2 & 5 & -1 \end{bmatrix}$ f) $\begin{bmatrix} -2 & -1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ g) $\begin{bmatrix} 0 & 2 & 0 \\ -2 & 2 & 2 \\ 1 & 1 & 0 \end{bmatrix}$

h) $\begin{bmatrix} -1 & -1 & 0 \\ 1 & -2 & 3 \\ 1 & -2 & -6 \end{bmatrix}$ i) $\begin{bmatrix} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ j) $\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

7. a) Näytä, että similaarisuus on samaa kokoa olevien neliömatriisien välinen ekvivalenssirelaatio. Millainen on ekvivalenssiluokka, joka sisältää yksikkömatriisin $\mathbf{I}$?

   b) Olkoon $\mathbf{A}$ ja $\mathbf{B}$ lineaarikuvauksen $A : \mathbb{R}^n \to \mathbb{R}^n$ matriisit avaruuden $\mathbb{R}^n$ kahdessa eri kannassa. Näytä, että $\mathbf{A}$ ja $\mathbf{B}$ ovat similaariset.

8. Määritä seuraavien matriisien ominaisarvot ja muodosta ominaisvektoreista $\mathbb{R}^2$:n, $\mathbb{R}^3$:n tai $\mathbb{R}^4$:n ortonormeerattu, positiivisesti suunnistettu kanta.

a) $\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}$ d) $\begin{bmatrix} 0 & -2 \\ -2 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$ f) $\begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$ g) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

h) $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ i) $\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}$ j) $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}$

9. (*) *Cayleyn–Hamiltonin lauseen* mukaan neliömatriisi toteuttaa oman karakteristisen yhtälönsä, ts. jos $p(\lambda)$ on matriisin karakteristinen polynomi, niin $p(\mathbf{A}) = \mathbf{0}$. a) Todista lause siinä tapauksessa, että $\mathbf{A}$ on diagonalisoituva. b) Näytä, että lause on tosi myös Esimerkin 4 ei-diagonalisoituvalle matriisille

XII.2 Neliömuotojen luokittelu

Yleinen $\mathbb{R}^n$:n *neliömuoto* on funktio $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x},\ \mathbf{x} \in \mathbb{R}^n$, missä $\mathbf{A}$ on reaalinen ja symmetrinen matriisi kokoa $n \times n$ (ks. Esimerkki XI.2: 5). Tässä luvussa tarkastellaan neliömuodon f luokittelua sen mukaan, minkä merkkisiä arvoja f saa muualla kuin origossa. Yleisessä tapauksessa luokitteluongelman ratkaisu nojaa edellisessä luvussa esitettyyn symmetrisen matriisin ominaisarvoteoriaan ja tähän perustuvaan matriisin diagonalisointiin (Lauseet XII.1.4 ja XII.1.6). Tämä ratkaisu myös tuottaa neliömuodon *pääakselikoordinaatiston*, jossa neliömuoto (eli matriisi $\mathbf{A}$) diagonalisoituu.

Jatkossa luokitellaan ensin neliömuodot ja ratkaistaan luokitteluongelma, rajoittuen ei-triviaaleihin tapauksiin $\mathbf{A} \neq \mathbf{0}$ ja $n \geq 2$. Tämän jälkeen tarkastellaan kahta neliömuotojen teorian sovellusta: n muuttujan funktion kriittisen pisteen luokittelua mahdollisena paikallisena ääriarvokohtana (Sovellus #1) ja toisen asteen käyrien ja pintojen geometrista luokittelua (Sovellus #2).

Jos $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ ja kirjoitetaan $\mathbf{e} = \mathbf{x}/|\mathbf{x}|$, niin neliömuodolle $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ pätee

$$f(\mathbf{x}) = f(|\mathbf{x}|\mathbf{e}) = |\mathbf{x}|^2\mathbf{e}^T\mathbf{A}\mathbf{e} = |\mathbf{x}|^2 f(\mathbf{e}). \qquad (\star)$$

Tämän mukaan nelömuodon arvojen etumerkkejä tutkittaessa riittää määrittää f:n arvojoukko $\mathbb{R}^n$:n yksikköpallolla

$$B = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x}| = 1\}.$$

Arvojoukon $f(B)$ mukaisesti neliömuoto luokitellaan seuraavasti.

MÄÄRITELMÄ XII.2.1 Neliömuoto $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ on

- *positiivisesti definiitti*, jos $f(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in B$,
- *negatiivisesti definiitti*, jos $f(\mathbf{x}) < 0 \quad \forall \mathbf{x} \in B$,
- *positiivisisesti semidefiniitti* (puolidefiniitti), jos $f(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in B$ ja $f(\mathbf{x}) = 0$ jollakin $\mathbf{x} \in B$,
- *negatiivisisesti semidefiniitti*, jos $f(\mathbf{x}) \leq 0 \quad \forall \mathbf{x} \in B$ ja $f(\mathbf{x}) = 0$ jollakin $\mathbf{x} \in B$,
- *indefiniitti* (ei-definiitti), jos $f(\mathbf{x}) > 0$ ja $f(\mathbf{y}) < 0$ joillakin $\mathbf{x}, \mathbf{y} \in B$.

Yksinkertaisimmassa eli kahden muuttujan tapauksessa neliömuodon definiitisyysongelma ratkeaa helposti. Kun merkitään $\mathbf{x} = (x, y)$ ja $\mathbf{A} = (a_{ij})$, niin yleinen kahden muuttujan neliömuoto on funktio

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2, \quad a_{ij} \in \mathbb{R}.$$

Definiittisyyslaji saadaan määrätyksi esimerkiksi laskemalla suoraan f:n arvojoukko yksikköympyrällä $B = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Tämä käy parametrisoimalla B muodossa $x = \cos t,\ y = \sin t,\ t \in [0, 2\pi]$, jolloin tutkimuskohteeksi tulee funktio

$$f(\cos t, \sin t) = F(t) = a_{11}\cos^2 t + 2a_{12}\cos t \sin t + a_{22}\sin^2 t.$$

F:n arvojouko $[F_{\min}, F_{\max}]$ välillä $[0, 2\pi]$ määrää f:n definiittisyyden lajin.

Toinen laskutapa on määrätä f:n nollakohdat origon kautta kulkevilla suorilla $S_t:\ y = tx,\ t \in \mathbb{R}$ ja (erikseen) suoralla $S_y:\ x = 0$:

$$\begin{aligned} y = tx \ &: \quad f(x, tx) = (a_{11} + 2a_{12}t + a_{22}t^2)x^2 = F(t)x^2 \\ x = 0 \ &: \quad f(0, y) = a_{22}y^2. \end{aligned}$$

Sen mukaan, onko $f(x, y) = 0$ vain origossa, pitkin yhtä suoraa, vai pitkin kahta origossa leikkaavaa suoraa, on pääteltävissä, että f on vastaavasti definiitti, semidefiniitti tai indefiniitti.

Esimerkki 1 Luokittele neliömuoto $f(x, y) = x^2 + axy + 4y^2$, kun
a) $a = 3$, b) $a = 4$, c) $a = 5$.

<u>Ratkaisu</u>

a) $$\begin{aligned} f(\cos t, \sin t) &= \cos^2 t + 3\cos t \sin t + \sin^2 t \\ &= (\cos t + \tfrac{3}{2}\sin t)^2 + \tfrac{1}{4}\sin^2 t > 0 \quad \forall t \in [0, 2\pi] \\ &\Rightarrow\ f \text{ on positiivisesti definiitti.} \end{aligned}$$

b) $$\begin{aligned} f(x, y) &= (x + 2y)^2 \\ &\Rightarrow\ f \text{ on positiivisesti semidefiniitti.} \end{aligned}$$

c) $$\begin{aligned} f(0, y) &= 4y^2 > 0, \text{ kun } y \neq 0, \\ f(x, tx) &= (1 + 5t + 4t^2)x^2 = 0, \text{ kun } t = -1 \text{ tai } t = -1/4 \\ &\Rightarrow\ f \text{ on indefiniitti.} \quad \square \end{aligned}$$

Neliömuodon diagonalisointi

Useamman kuin kahden muuttujan tapauksessa nelömuodon luokitteluongelma ratkeaa luontevimmin symmetrisen matriisin ominaisarvoteorian avulla: Suoritetaan matriisin $\mathbf{A}$ diagonalisaatio

$$\mathbf{D} = \mathbf{C}^T\mathbf{A}\mathbf{C} \;\Leftrightarrow\; \mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}^T,$$

missä $\mathbf{D} = \operatorname{diag}(\lambda_i)$ koostuu $\mathbf{A}$:n ominaisarvoista ja $\mathbf{C}$ ortonormeeratuista ominaisvektoreista. Tällöin $f(\mathbf{x})$ saa muodon

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{C}\mathbf{D}\mathbf{C}^T\mathbf{x} = (\mathbf{C}^T\mathbf{x})^T\mathbf{D}(\mathbf{C}^T\mathbf{x}).$$

Siis jos tehdään $\mathbb{R}^n$:ssä muuttujan vaihto

$$\mathbf{y} = \mathbf{C}^T\mathbf{x} \;\Leftrightarrow\; \mathbf{x} = \mathbf{C}\mathbf{y},$$

niin f saa muodon

$$f(\mathbf{x}) = g(\mathbf{y}) = \mathbf{y}^T\mathbf{D}\mathbf{y} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

Neliömuotoa $g(\mathbf{y})$ sanotaan — varsin hyvällä syyllä — f:n *diagonalisoiduksi muodoksi*. Huomattakoon, että $\mathbf{C}$:n ollessa ortogonaalinen säilyttää muunnos $\mathbf{x} \mapsto \mathbf{y} = \mathbf{C}^T\mathbf{x}$ vektorin pituuden:

$$|\mathbf{y}|^2 = \mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{C}\mathbf{C}^T\mathbf{x} = \mathbf{x}^T\mathbf{x} = |\mathbf{x}|^2.$$

Määritelmän XII.2.1 mukaisesti neliömuodon luokitteluongelma ratkeaa näin $\mathbf{A}$:n ominaisarvojen avulla seuraavasti:

$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$:

f positiivisesti definiitti $\Leftrightarrow\; \lambda > 0 \quad \forall\lambda \in \sigma(\mathbf{A})$.

f negatiivisesti definiitti $\Leftrightarrow\; \lambda < 0 \quad \forall\lambda \in \sigma(\mathbf{A})$.

f positiivisesti semidefiniitti $\Leftrightarrow\; \lambda \geq 0 \quad \forall\lambda \in \sigma(\mathbf{A})$ ja $0 \in \sigma(\mathbf{A})$.

f negatiivisesti semidefiniitti $\Leftrightarrow\; \lambda \leq 0 \quad \forall\lambda \in \sigma(\mathbf{A})$ ja $0 \in \sigma(\mathbf{A})$.

f indefiniitti $\Leftrightarrow\; \lambda, \mu \in \sigma(\mathbf{A})$ joillakin $\lambda > 0$ ja $\mu < 0$.

Kun huomioidaan myös skaalaustulos ($\star$), niin em. luokittelujen nojalla pätee erityisesti

- f positiivisesti definiitti $\Rightarrow$ $f(\mathbf{x}) \geq \lambda|\mathbf{x}|^2 \ \forall \mathbf{x} \in \mathbb{R}^n$,
- f negatiivisesti definiitti $\Rightarrow$ $f(\mathbf{x}) \leq -\lambda|\mathbf{x}|^2 \ \forall \mathbf{x} \in \mathbb{R}^n$,

missä $\lambda > 0$ on pienin luvuista $|\lambda_i|$, $\lambda_i \in \sigma(\mathbf{A})$.

Esimerkki 1 (jatko) Esimerkissä on

$$f(x,y) = \mathbf{x}^T\mathbf{A}\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & \frac{a}{2} \\ \frac{a}{2} & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Esimerkin luokitteluihin tapauksissa a) $a = 3$, b) $a = 4$, c) $a = 5$ päädytään myös laskemalla

$$\text{a) } \sigma(\mathbf{A}) = \{\frac{1}{2}(5 \pm 3\sqrt{2})\}, \quad \text{b) } \sigma(\mathbf{A}) = \{0, 5\}, \quad \text{c) } \sigma(\mathbf{A}) = \{\frac{1}{2}(5 \pm \sqrt{34}\}. \quad \square$$

Esimerkki 2 Luokittele $\mathbb{R}^3$:n neliömuoto $f(x,y,z) = x^2 + 3y^2 + z^2 + 4xz$.

<u>Ratkaisu</u>

$$\begin{aligned} f(x,y,z) &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \mathbf{x}^T\mathbf{A}\mathbf{x}, \quad \mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix}^T. \end{aligned}$$

Tässä on $\sigma(2\mathbf{A}) = \{-2, 6\}$ (edellisen luvun Esimerkki 1), joten $\sigma(\mathbf{A}) = \{-1, 3\}$ ja f on siis indefiniitti. $\square$

Neliömuodon pääakselikoordinaatisto

Neliömuodon $\mathbf{x}^T\mathbf{A}\mathbf{x}$ definiittisyyslaji voidaan päätellä pelkästään $\mathbf{A}$:n ominaisarvojen avulla, mutta kiinnostava on myös tieto, millaisessa koordinaatistossa neliömuoto diagonalisoituu. Tämän tiedon välittävät $\mathbf{A}$:n ortonormeeratut ominaisvektorit, jotka määräävät diagonalisoivan koordinaattimuunnoksen $\mathbf{y} = \mathbf{C}^T\mathbf{x}$. Tämän mukaisesti vektorin $\mathbf{y}$ alkiot ovat vektorin $\mathbf{x}$ koordinaatit kannassa $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$, jonka muodostavat $\mathbf{C}$:n sarakkeet. Jatkossa oletetaan, että kanta $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ on myös <u>positiivisesti</u> <u>suunnistettu</u> (tarvittaessa yhden vektorin suunnan vaihto vastakkaiseksi), jolloin on $\det \mathbf{C} = 1$ ja kannan vaihdossa $\{\mathbf{e}_i\} \hookrightarrow \{\mathbf{c}_i\}$ on kyse <u>koordinaatiston</u> <u>kierrosta</u> (ks. Luku X.6).

Kierrettyä koordinaatistoa, jossa neliömuoto $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ diagonalisoituu, sanotaan ko. neliömuodon *pääakselikoordinaatistoksi.* Kyseisen koordinaatiston kantana toimivat siis $\mathbf{A}$:n ortonormeeratut ja positiivisesti suunnistetut ominaisvektorit.

Esimerkki 2 (jatko) Esimerkissä pääakselikoordinaatiston (ortonormeeratut ja positiivisesti suunnistetut) kantavektorit ovat

$$\mathbf{c}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0 & 1\end{bmatrix}^T, \quad \mathbf{c}_2 = \begin{bmatrix}0 & 1 & 0\end{bmatrix}^T, \quad \mathbf{c}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 0 & 1\end{bmatrix}^T$$

vastaten $\mathbf{A}$:n ominaisarvoja $\lambda_1 = \lambda_2 = 3$ ja $\lambda_3 = -1$ (vrt. edellinen luku). Kun muunnettuja koordinaatteja merkitään $\mathbf{y} = (\xi, \eta, \zeta)$, niin koordinaatiston kiertomuunnos on

$$\mathbf{y} = \mathbf{C}^T\mathbf{x} \quad \Leftrightarrow \quad \begin{bmatrix}\xi\\ \eta\\ \zeta\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix},$$

ja tässä pääakselikoordinaatistossa on siis

$$f(x,y,z) = g(\xi,\eta,\zeta) = 3\xi^2 + 3\eta^2 - \zeta^2. \quad \square$$

Sovellus 1: Funktion kriittisen pisteen luokittelu

Funktion $f(\mathbf{x})$ kriittisen pisteen luokittelu mahdollisena paikallisena ääriarvokohtana (vrt. Propositio XI.8.2) on kysymys, jote ei toistaiseksi ole tarkastelu. Palataan nyt tähän kysymykseen neliömuotijen teorian valossa. Olkoon $\mathbf{c}$ funktion f kriittinen piste, ts. $\boldsymbol{\nabla} f(\mathbf{c}) = \mathbf{0}$. Tällöin jos f on $\mathbf{c}$:n ympäristössä riittävän säännöllinen, niin usean muuttujan Taylorin kaavan (ks. Luku XI.9) perusteella pätee

$$f(\mathbf{x}) = f(\mathbf{c}) + \tfrac{1}{2}(\mathbf{x}-\mathbf{c})^T\mathbf{H}(\mathbf{c})(\mathbf{x}-\mathbf{c}) + o(|\mathbf{x}-\mathbf{c}|^2),$$

missä $\mathbf{H}(\mathbf{c})$ on funktion Hessen matriisi pisteessä $\mathbf{c}$:

$$[\,\mathbf{H}(\mathbf{c})\,]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{c}), \quad i,j = 1\ldots n.$$

Sen seikan selvittämiseksi, onko $\mathbf{c}$ f:n paikallinen ääriarvokohta vai ei, on siis ilmeisesti ensisijaisesti tutkittava, millaisia arvoja funktio

$$g(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{c})^T\mathbf{H}(\mathbf{c})(\mathbf{x}-\mathbf{c})$$

saa pisteen $\mathbf{c}$ lähiympäristössä. Kun siirretään origo pisteeseen $\mathbf{c}$ koordinaattimuunnoksella $\mathbf{x}' = \mathbf{x} - \mathbf{c}$, niin tutkimuskohteeksi tulee $\mathbb{R}^n$:n neliömuoto $h(\mathbf{x}') = g(\mathbf{x}' + \mathbf{c}) = (\mathbf{x}')^T \mathbf{A}\mathbf{x}'$, $\mathbf{A} = \frac{1}{2}\mathbf{H}(\mathbf{c})$.

Olkoon neliömuoto $h(\mathbf{x}')$ positiivisesti definiitti. Tällöin pätee mainituin oletuksin jollakin $\delta > 0$

$$f(\mathbf{x}) \geq f(\mathbf{c}) + \lambda|\mathbf{x} - \mathbf{c}|^2 - c_\delta|\mathbf{x} - \mathbf{c}|^2, \quad \text{kun } 0 < |\mathbf{x} - \mathbf{c}| < \delta,$$

missä $\lambda > 0$ ja $c_\delta \to 0$, kun $\delta \to 0$. Siis jos valitaan $\delta > 0$ siten, että $c_\delta \leq \lambda/2$, niin ko. δ:n arvolla pätee

$$f(\mathbf{x}) \geq f(\mathbf{c}) + \frac{1}{2}\lambda|\mathbf{x} - \mathbf{c}|^2, \quad 0 < |\mathbf{x} - \mathbf{c}| < \delta.$$

Näin ollen $\mathbf{c}$:ssä on f:n paikallinen minimi. Vastaavalla tavalla päätellään, että $\mathbf{c}$ on f:n paikallinen maksimipiste, jos $h(\mathbf{x}')$ on negatiivisesti definiitti, ja että $\mathbf{c}$ ei ole f:n paikallinen ääriarvo- eikä laakapiste, jos $h(\mathbf{x}')$ on indefiniitti. Päätelmät yhdistettyinä ovat siis seuraavat.

Propositio XII.2.2 Jos funktiolle $f : A \to \mathbb{R}$, $A \subset \mathbb{R}^n$, on pisteen $\mathbf{c} \in A$ jossakin ympäristössä voimassa approksimaatio

$$f(\mathbf{x}) = f(\mathbf{c}) + g(\mathbf{x} - \mathbf{c}) + o(|\mathbf{x} - \mathbf{c}|^2), \quad \mathbf{x} \in U_\delta(\mathbf{a}) \subset A,$$

missä g on neliömuoto, niin pätee

1. g positiivisesti definiitti $\Rightarrow$ $\mathbf{c}$ on f:n paikallinen minimipiste.
2. g negatiivisesti definiitti $\Rightarrow$ $\mathbf{c}$ on f:n paikallinen maksimipiste.
3. g indefiniitti $\Rightarrow$ $\mathbf{c}$ ei ole f:n paikallinen ääriarvo- eikä laakapiste.

Huomattakoon, että neliömuodon indefiniittisyys on mahdollinen vain kahden tai useamman muuttujan tilanteessa, sillä jos $n = 1$, niin $\mathbf{x}^T\mathbf{A}\mathbf{x} = Ax^2$, joten neliömuoto on joko positiivisesti definiitti ($A > 0$), negatiivisesti definiitti ($A < 0$) tai $= 0$ ($A = 0$). Ehto $f''(c) > 0$ on yhden muuttujan analyysista tuttu paikallisen minimin (riittävä) ehto. Tämän vastine n muuttujan tilanteessa on siis ehto

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x} \neq \mathbf{0}\ ; \quad \mathbf{A} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{c})\right),$$

eli toisen derivaatan $f''(c)$ vastine n muuttujan ääriarvotarkastelussa on kaikkien toisen kertaluvun osittaisderivaattojen muodostama Hessen matriisi.

Esimerkki 3 Tutki, onko seuraavilla funktioilla paikallista ääriarvoa origossa:
a) $f(x,y) = \cos x + 2y - 2e^y$, b) $f(x,y) = e^{-x^2}e^{-y^2} + 3\sin xy$.

Ratkaisu a) $f(x,y) = -1 + (-\frac{1}{2}x^2 - y^2) + \ldots$ Neliömuoto-osa on negatiivisesti definiitti. Siis f:llä on origossa paikallinen maksimi.
b) $f(x,y) = 1 + (-x^2 - y^2 + 3xy) + \ldots$ Neliömuoto-osa on indefiniitti. Siis origo ei ole f:n paikallinen ääriarvopiste eikä myöskään laakapiste. □

Esimerkki 4 Selvitä funktion $f(x,y) = x^2 + 2y^2 + 2xy + 4x + 14y$ kriittisen pisteen $(3,-5)$ laatu (vrt. Luvun XI.8 Esimerkki 2).

Ratkaisu Siirretään kriittinen piste ensin origoon:

$$\begin{aligned} &x = 3 + x',\ y = -5 + y' \\ &g(x',y') = f(3+x', -5+y') = (x')^2 + 2x'y' + 2(y')^2 - 29. \end{aligned}$$

Koska g:n nelömuoto-osa on positiivisesti definiitti, on f:llä on pisteessä $(3,-5)$ paikallinen minimi. Funktion muodosta nähdään, että tämä on myös absoluuttinen minimi. □

Propositio XII.2.2 jättää avoimeksi tapauksen, jossa neliömuoto $g(\mathbf{x}')$ on puolidefiniitti. Kuten tapauksessa $n = 1$ ($f''(c) = 0$), on mahdollisen ääriarvokohdan laatu ratkaistava tapauskohtaisesti.

Esimerkki 5 Määrittele kriittisen pisteen $(0,0)$ laatu, kun

$$f(x,y) = x^2 + 2xy + y^2 + a(x+y)^3 + bx^n, \quad a,b \in \mathbb{R}, \quad n \in \{3,4,\ldots\}.$$

Ratkaisu Koska f:n neliömuoto-osa

$$g(x,y) = x^2 + 2xy + y^2 = (x+y)^2$$

on positiivisesti puolidefiniiti, on tutkimuksia jatkettava ainakin suoralla $y = -x$:

$$f(x,-x) = bx^n.$$

Päätellään: Jos $b = 0$, on origo laakapiste. Jos $b > 0$ ja n on parillinen, on origo f:n paikallinen minimipiste. Muulloin, eli jos $b < 0$ ja n on parillinen, tai jos $b \neq 0$ ja n on pariton, saa f origon jokaisessa ympäristössä sekä positiivisia että negatiivisia arvoja, joten origo ei ole paikallinen ääriarvo- eikä laakapiste. □

Sovellus 2: Toisen asteen käyrät ja pinnat

Palautettakoon mieliin Luvusta II.7, että yleisen toisen asteen käyrän yhtälö tasossa on muotoa

$$Ax^2 + By^2 + 2Cxy + Dx + Ey + F = 0,$$

missä $A, \ldots, F \in \mathbb{R}$ ja ainakin yksi luvuista A, B, C on nollasta poikkeava, ja yleisen toisen asteen pinnan yhtälö avaruudessa on muotoa

$$Ax^2 + By^2 + Cz^2 + 2Dxy + 2Eyz + 2Fxz + Gx + Hy + Iz + J = 0,$$

missä $A, \ldots, J \in \mathbb{R}$ ja ainakin yksi luvuista $A, \ldots, F$ on nollasta poikkeava. Jos merkitään

$$\text{Taso:} \qquad \mathbf{A} = \begin{bmatrix} A & C \\ C & B \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} D \\ E \end{bmatrix}, \quad c = F,$$

$$\text{Avaruus:} \quad \mathbf{A} = \begin{bmatrix} A & D & F \\ D & B & E \\ F & E & C \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} G \\ H \\ I \end{bmatrix}, \quad c = J,$$

niin toisen asteen käyrän ja pinnan yhtälöt saadaan muotoon

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{a}^T\mathbf{x} + c = 0.$$

Suoritetaan tässä koordinaattimuunnos

$$\mathbf{x} = \mathbf{C}\mathbf{y} + \mathbf{b},$$

missä $\mathbf{y} = [\xi, \eta]^T$ (taso) tai $\mathbf{y} = [\xi, \eta, \zeta]^T$ (avaruus), ja $\mathbf{C}$ ja $\mathbf{b}$ valitaan niin, että yhtälö saa muunnetuissa koordinaateissa mahdollisimman yksinkertaisen muodon. Ensinnäkin valitaan $\mathbf{C}$ niin, että neliömuoto $\mathbf{x}^T\mathbf{A}\mathbf{x}$ diagonalisoituu, eli valitaan $\mathbf{C}$:n sarakkeiksi $\mathbf{A}$:n ortonormeeratut ja positiivisesti suunnistetut ominaisvektorit. Tällöin yhtälö saadaan muotoon

$$\mathbf{y}^T\mathbf{D}\mathbf{y} + \mathbf{d}^T\mathbf{y} + d = 0$$

missä $\mathbf{D}$ on diagonaalinen matriisi (diagonaalilla $\mathbf{A}$:n ominaisarvot) ja

$$\mathbf{d} = \mathbf{C}^T(2\mathbf{A}\mathbf{b} + \mathbf{a}), \quad d = \mathbf{b}^T(\mathbf{A}\mathbf{b} + \mathbf{a}) + c.$$

Jos $\mathbf{A}$ on säännöllinen matriisi, niin määrätään $\mathbf{b}$ siten, että $\mathbf{d} = \mathbf{0}$, muussa tapauksessa saatetaan $\mathbf{b}$:n valinnalla termi $\mathbf{d}^T\mathbf{y} + d$ mahdollisimman yksinkertaiseen muotoon. (Jos kyseessä on avaruuden pinta ja $\lambda = 0$ on $\mathbf{A}$:n kaksinkertainen

ominaisarvo, voi yhtälön yksinkertaistumiseen vaikuttaa myös $\mathbf{C}$:n valinnalla, ks. Harj.teht. 10.) Yhtälön näin pelkistyttyä voidaan sen määräämä käyrä tai pinta luokitella geometrisesti.

Jos kyseessä on tasokäyrä, niin em. muunnokset johtavat yhteen seuraavista perusmuodoista:

1. $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} = q, \quad q \in \{1, 0, -1\}, \quad a, b > 0.$

2. $\dfrac{\xi^2}{a^2} - \dfrac{\eta^2}{b^2} = q, \quad q \in \{1, 0, -1\}, \quad a, b > 0.$

3. $\eta = a\xi^2 \quad \text{tai} \quad \xi = a\eta^2, \quad a \in \mathbb{R},\ a \neq 0.$

4. $\xi^2 = a \quad \text{tai} \quad \eta^2 = a, \quad a \in \mathbb{R}.$

Tapauksessa 1 käyrä on *ellipsi*, jos $q = 1$, muulloin kyseessä on piste ($q = 0$) tai tyhjä joukko ($q = -1$). Tapauksessa 2 käyrä on *hyperbeli*, jos $q = \pm 1$, muuten kysessä on kaksi toisensa leikkaavaa suoraa ($q = 0$). Tapauksessa 3 käyrä on *paraabeli*, ja tapauksessa 4 on kyseessä joko kaksi yhdensuuntaista suoraa ($a > 0$), yksi suora ($a = 0$), tai tyhjä joukko ($a < 0$). Tapaukseen 1 päädytään kun $\mathbf{A}$ on definiitti (ominaisarvot samanmerkkiset), tapaukseen 2 kun $\mathbf{A}$ on indefiniitti (ominaisarvot erimerkkiset) ja tapauksiin 3 ja 4 kun $\mathbf{A}$ on puolidefiniitti (yksi ominaisarvo $= 0$).

Esimerkki 6 Etsi sellainen koordinaatisto, jossa funktion

$$f(x, y) = 13x^2 - 8xy + 7y^2 - 42x + 36y$$

tasa-arvokäyrät saavat toisen asteen käyrän perusmuodon. Kuvio!

<u>Ratkaisu</u> Kirjoitetaan tasa-arvokäyrän $S:\ f(x, y) = c$ yhtälö ensin muotoon

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -42 & 36 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - c = 0.$$

Matriisin $\mathbf{A} = \begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix}$ ominaisarvot ja ortonormeeratut, positiivisesti suunnistetut ominaisvektroit ovat

$$\lambda_1 = 15,\ \mathbf{c}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \lambda_2 = 5,\ \mathbf{c}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Suoritetaan koordinaattimuunnos kahdessa vaiheessa. Ensin kierto neliömuodon pääakselikoordinaatistoon (x', y') (edellä $\mathbf{b} = \mathbf{0}$):

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Tulos:

$$\begin{aligned} f(x,y) = c \quad &\Leftrightarrow \quad 15(x')^2 + 5(y')^2 - \frac{120}{\sqrt{5}}\, x' + \frac{30}{\sqrt{5}}\, y' - c = 0 \\ &\Leftrightarrow \quad 15\left(x' - \frac{4}{\sqrt{5}}\right)^2 + 5\left(y' + \frac{3}{\sqrt{5}}\right)^2 - 57 - c = 0. \end{aligned}$$

Toinen vaihe on origon siirto, koordinaateiksi

$$\xi = x' - \frac{4}{\sqrt{5}}\,, \quad \eta = y' + \frac{3}{\sqrt{5}}\,.$$

Lopputulos: $\quad f(x,y) = c \quad \Leftrightarrow \quad 3\xi^2 + \eta^2 = \frac{1}{5}(57 + c).$

Tämän mukaan tasa-arvokäyrä on joko ellipsi ($c > -57$), piste ($c = -57$) tai tyhjä joukko ($c < -57$). Ellipsien yhteisiä pääakseleita ovat suorat $\xi = 0$ ja $\eta = 0$, ja niiden yhteisessä keskipisteessä

$$(\xi, \eta) = (0,0) \;\Leftrightarrow\; (x', y') = (4/\sqrt{5}, -3/\sqrt{5}) \;\Leftrightarrow\; (x, y) = (1, -2)$$

f saavuttaa absoluuttisen minimiarvonsa $f_{\min} = -57$. $\square$

Toisen asteen pinnat voidaan em. koordinaattimuunnoksen antaman perusmuodon avulla luokitella seuraaviin yhdeksään päätyyppiin (yhtälöissä voi koordinaattien järjestystä vaihdella).

1. *Ellipsoidi:* $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} + \dfrac{\zeta^2}{c^2} = 1.$
2. *Yksivaippainen hyperboloidi:* $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} - \dfrac{\zeta^2}{c^2} = 1.$
3. *Kaksivaippainen hyperboloidi:* $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} - \dfrac{\zeta^2}{c^2} = -1.$
4. *Kartio:* $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} - \dfrac{\zeta^2}{c^2} = 0.$
5. *Elliptinen paraboloidi:* $\zeta = \dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2}\,.$
6. *Hyperbolinen paraboloidi:* $\zeta = \dfrac{\xi^2}{a^2} - \dfrac{\eta^2}{b^2}\,.$
7. *Elliptinen lieriö:* $\dfrac{\xi^2}{a^2} + \dfrac{\eta^2}{b^2} = 1.$
8. *Hyperbolinen lieriö:* $\dfrac{\xi^2}{a^2} - \dfrac{\eta^2}{b^2} = 1.$
9. *Parabolinen lieriö:* $\eta = a\xi^2.$

Hyperbolinen paraboloidi on toiselta nimeltään *satulapinta*. Tapauksissa 1–4 $\mathbf{A}$ on säännöllinen matriisi, ja tapauksessa 1 ominaisarvot ovat lisäksi samanmerkkiset. Tapauksissa 5–8 on $\mathbf{A}$:lla yksinkertainen ja tapauksessa 9 kaksinkertainen ominaisarvo $\lambda = 0$. Näiden päätyyppien lisäksi toisen asteen pinnan yhtälö voi esittää kahta toisensa leikkaavaa tasoa, kahta yhdensuuntaista tasoa, yhtä tasoa, avaruussuoraa, pistettä, tai ei mitään.

Esimerkki 7 Luokittele eri c:n arvoilla toisen asteen pinta

$$5x^2 + 2y^2 + 2z^2 + 2xy - 4yz - 2xz + 10x + 2y - 2z + c = 0.$$

Ratkaisu Matriisin

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & -1 \\ 1 & 2 & -2 \\ -1 & -2 & 2 \end{bmatrix}$$

karakteristinen yhtälö on

$$\begin{vmatrix} 5-\lambda & 1 & -1 \\ 1 & 2-\lambda & -2 \\ -1 & -2 & 2-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 18\lambda = 0.$$

Ominaisarvot ovat $\lambda_1 = 0$, $\lambda_2 = 3$ ja $\lambda_3 = 6$. Näitä vastaavat, ortonormeeratut ja positiivisesti suunnistetut ominaisvektorit ovat matriisin $\mathbf{C}$ sarakkeiksi koottuna:

$$\mathbf{C} = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

Tehdään ensin koordinaatiston kiertomuunnos $\mathbf{C}\mathbf{x}' = \mathbf{x} \;\Leftrightarrow\; \mathbf{x}' = \mathbf{C}^T\mathbf{x}$, missä $\mathbf{x} = [x, y, z]^T$, $\mathbf{x}' = [x', y', z']^T$, jolloin pinnan yhtälö saadaan muotoon

$$\begin{aligned} 0 &= 3(y')^2 + 6(z')^2 + \frac{5}{\sqrt{3}}\,y' + \frac{22}{\sqrt{6}}\,z' + c \\ &= 3\left(y' + \frac{1}{\sqrt{3}}\right)^2 + 6\left(z' + \frac{2}{\sqrt{6}}\right)^2 + c - 5. \end{aligned}$$

Tästä päästään perusmuotoon origon siirrolla:

$$\xi = x', \quad \eta = y' + \frac{1}{\sqrt{3}}, \quad \zeta = z' + \frac{2}{\sqrt{6}} \quad \Rightarrow \quad 3\eta^2 + 6\zeta^2 = 5 - c.$$

Tästä nähdään, että jos $c < 5$, niin pinta on elliptinen lieriö. Lieriön symmetria-akseli on suora $S : \eta = \zeta = 0$. Suora S kulkee pisteen $(\xi, \eta, \zeta) = (0, 0, 0)$ eli pisteen $(x, y, z) = (-1, 0, 0)$ kautta ja sen suuntavektori alkuperäisessä koordinaatistossa on $\mathbf{t} = [0, 1, 1]^T \mathrel{\widehat{=}} \vec{j} + \vec{k}$ (ominaisarvoa $\lambda_1 = 0$ vastaava $\mathbf{A}$:n ominaisvektori, joka määrää ξ-akselin suunnan). Tapauksessa $c = 5$ yhtälö määrittelee vain suoran S ja tapauksessa $c > 5$ ei mitään. — Suora S on itse asiassa tasojen $T_1 : x - y + z + 1 = 0$ ja $T_2 : 2x + y - z + 2 = 0$ leikkaussuora, sillä alkuperäinen yhtälö on sama kuin

$$(x - y + z + 1)^2 + (2x + y - z + 2)^2 = 5 - c. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Luokittele seuraavat neliömuodot määrittämällä a)-kohdassa arvojoukko yksiköympyrällä ja b)- ja c)-kohdissa sijoituksella $y = tx$.
   a) $3x^2 - 16xy + 16y^2$ b) $4x^2 - 16xy + 16y^2$ c) $5x^2 - 16xy + 16y^2$

2. Määritä nelömuodon pääakselikoordinaatisto ja diagonalisoitu muoto:
   a) $f(x, y) = xy$ b) $f(x, y) = 4x^2 - 16xy + 16y^2$
   c) $f(x, y) = -3x^2 - 3y^2 + 4xy$ d) $f(x, y, z) = 2x^2 + 3y^2 + 2z^2 - 6xz$
   e) $f(x, y, z, u) = x^2 + y^2 + z^2 + u^2 + 2yz + 2xu$

3. Määritä neliömuoto $f(x, y, z)$, jonka pääakselikoordinaatiston kantavektorit ovat
$$\mathbf{c}_1 = \frac{1}{7}\begin{bmatrix} 3 \\ -2 \\ 6 \end{bmatrix}, \quad \mathbf{c}_2 = \frac{1}{7}\begin{bmatrix} -2 \\ 6 \\ 3 \end{bmatrix}, \quad \mathbf{c}_3 = \frac{1}{7}\begin{bmatrix} -6 \\ -3 \\ 2 \end{bmatrix}$$
   ja jonka diagonalisoitu muoto pääakselikoordinaatistossa on
$$g(\xi, \eta, \zeta) = 49\xi^2 - 98\eta^2 + 147\zeta^2.$$

4. Tutki, onko origo funktion paikallinen ääriarvokohta:
   a) $f(x, y) = \cos(x - 2y) - \sin xy$ b) $f(x, y) = e^x y - xe^y + \sin(x - y)$
   c) $f(x, y) = 1 - x + 2y + e^{\frac{1}{2}xy} + \sin(x - 2y) + 2\cos(x + y)$
   d) $f(x, y) = 1 + x + 2y + (1 + x + 2y)^{-1} + 2\sin(xy) - 12\cos y$
   e) $f(x, y) = e^{x^2+y^2}\sqrt[3]{7 + 42xy}$
   f) $f(x, y, z) = x + y + z + e^{-x-y-z} + (x + y)\sin(x + y) + \cos(y + z)$
   g) $f(x, y, z) = x + y + z - e^{x+y+z} - (x + y)\sin(x + y) + \cos(y + z)$

5. Luokittele origo funktion kriittisenä pisteenä eri parametrin a arvoilla:
   a) $f(x, y) = \dfrac{1 + x^2 + 2y^2}{1 + axy}$ b) $f(x, y) = \sin xy + a\cos(x + y)$

6. Etsi seuraavien funktioiden kaikki kriittiset pisteet ja luokittele ne.
   a) $x^2 + 2y^2 - 4x + 4y$ b) $xy - x + y$ c) $x^3 + y^3 - 3xy$
   d) $x^4 + y^4 - 4xy$ e) $\cos(x + y)$ f) $\dfrac{x}{y} + \dfrac{8}{x} - y$
   g) $\dfrac{1}{x - y + x^2 + y^2}$ h) $\dfrac{xy}{1 + x + y}$ i) $(x - 3y)e^{x^2+y^2}$

7. Etsi karteesinen koordinaatisto, jossa seuraavien funktioiden tasa-arvokäyrät ovat toisen asteen käyrän yksinkertaisinta perusmuotoa. Hahmottele tasa-arvokäyrien kulku.
   a) $16x^2 + 9y^2 + 24xy$ b) $2x^2 + 3y^2 + 2xy$ c) $xy + y^2$
   d) $xy + x + 2y + 1$ e) $x^2 + 4xy + 4y^2 + 6x + 8y$
   f) $2x^2 + 3y^2 + 2xy - 4x - 4y$ g) $3x^2 + 4xy + y^2 + 2x$

8. Etsi seuraaville toisen asteen pinnoille koordinaatisto, jossa pinnan yhtälö pelkistyy perusmuotoonsa. Luokittele pinnat.
   a) $xy+yz+xz = 0$ b) $xy-yz+xz = 0$ c) $xy+yz+x+y+z = 0$
   d) $xy + yx - xz + 2x - 3y + z + 1 = 0$
   e) $x^2 + y^2 + z^2 + 8xy + 6yz - 24y + 8z = 0$
   f) $x^2 + y^2 + 3z^2 + 2xy + 4yz + 4xz + x + y + z - 18 = 0$

9. (*) Näytä, että eräällä $m \in \mathbb{N}$ pätee: Jos origon läpi kuljetaan pitkin käyriä

$$y = kx^n,\ k \in \mathbb{R},\ n \in \mathbb{N},\ n \neq m \quad \text{tai} \quad x = ky^n,\ k \in \mathbb{R},\ n \in \mathbb{N},$$

   niin funktiolla $f(x, y) = y^2 + 3x^{50}y + 2x^{100}$ on kullakin käyrällä origossa paikallinen minimi. Näytä, että origo ei kuitenkaan ole f:n paikallinen minimipiste eikä edes laakapiste.

10. Toisen asteen pinnan yhtälö $\mathbf{x}^T\mathbf{A}\mathbf{x}+\mathbf{b}^T\mathbf{x}+c = 0$, $\mathbf{x}^T = [x, y, z]$, on saatettu koordinatiston kierrolla muotoon

$$\xi^2 + G\xi + H\eta + I\zeta + J = 0$$

   ($\mathbf{A}$:lla kaksinkertainen ominaisarvo $\lambda = 0$). Näytä, että uudella koordinaatiston kierrolla ja sen jälkeen origon siirolla yhtälö saadaan perusmuotoon $\eta' = a(\xi')^2$ tai $(\xi')^2 = a$ ($a \in \mathbb{R}$). Sovella menettelyä:
   a) $x^2 + x + 2y - 2z - 4 = 0$ b) $(x + y + 2z)^2 + x + y - z = 0$
   c) $x^2 + y^2 + 4z^2 - 2xy + 4xz - 4yz - 1 = 0$
   d) $x^2 + 4y^2 + z^2 + 4xy + 2xz + 4yz - 4x - 3y - 4z + 4 = 0$
   e) $9x^2 + 4y^2 + z^2 - 12xy - 6xz + 4yz + 6x - 4y + 2z + 2 = 0$

11. (*) Määritä pääakselikoordinaatisto ja luokittele pinta:
   a) $17x^2 + 76y^2 + 54z^2 - 60xy + 12xz + 48yz + 4x + 16y + 36z + 6 = 0$
   b) $x^2 - 68y^2 + 18z^2 + 36xy + 60xz - 96yz - 22x - 53y + 26z + 121 = 0$
   c) $13x^2 + 40y^2 + 45z^2 - 36xy + 24xz + 12yz - 34x - 17y + 44z + 124 = 0$
   d) $32x^2 - 20y^2 + 135z^2 - 24xy + 156xz - 132yz + 52x - 44y + 90z + 15 = 0$
   e) $48x^2 + 117y^2 + 31z^2 - 36xy - 60xz + 96yz - 216x - 66y + 86z + 300 = 0$

XII.3 Pinnan kaarevuus

Palautettakoon mieliin käyräteoriasta, että käyrän kaarevuus kertoo, miten nopeasti käyrän yksikkötangenttivektori, tai vaihtoehtoisesti kaaretumissuuntaan osoittava päänormaalivektori, muuttuu käyrää pitkin kuljettaessa, ks. Luku VII.3, kaavat (3)–(4).

Jos halutaan määrätä tasokäyrän kaarevuus annetussa käyrän pisteessä P, niin laskukaavojen kannalta mukavin on koordinaatisto, jonka origo on P:ssä ja x-akseli käyrän tangentin suuntainen (kuvio). Olkoon käyrän yhtälö tässä koordinaatistossa $y = f(x)$, missä f on kahdesti jatkuvasti derivoituva pisteen $x = 0$ lähellä. Tällöin käyrän yksikkönormaalivektori P:n lähellä on

$$\vec{n}(x) = \frac{1}{\sqrt{1 + [f'(x)]^2}}\,[-f'(x)\vec{i} + \vec{j}\,].$$

Kuvion tilanteessa käyrä on koordinaatistossa (x, y) alaspäin kaartuva P:n lähellä ($f''(0) < 0$), joten tässä tapauksessa käyrän päänormaalivektori P:n lähellä $= -\vec{n}(x)$. Kun huomioidaan, että $f'(0) = 0$, seuraa derivoimalla

$$\frac{d\vec{n}}{dx}(0) = -f''(0)\,\vec{i}.$$

Valitussa koordinaatistossa siis käyrän (merkkinen) kaarevuus P:ssä on $\kappa = f''(0)$ (vrt. Luvun VII.3 kaavat (4) ja (6) päänormaalivektorille $\vec{n}$).

Tarkastellaan nyt pintaa, jonka yhtälö on $z = f(x, y)$. Vastaavasti kuin edellä oletetaan tässä koordinaatisto valituksi siten, että xy-taso on pinnan tangenttitaso origossa (kuvio). Oletetaan myös, että f:n osittaisderivaatat ovat toiseen kertalukuun asti jatkuvia pisteen $(x, y) = (0, 0)$ lähellä.

Oletettuun tilanteeseen voidaan ajatella päästyn siten, että on valittu annetun pinnan jokin piste P, siirretty koordinaatiston origo ko. pisteeseen, ja sen jälkeen valittu tässä pisteessä (kierretty) koordinaatisto, jossa xy-taso on pinnan tangenttitaso P:ssä.

Pinnan normaali pisteessä $(x, y, f(x, y))$ on (vrt. Luku XI.3)

$$\vec{n}(x, y) = \frac{-f_x\vec{i} - f_y\vec{j} + \vec{k}}{\sqrt{1 + f_x^2 + f_y^2}}\,.$$

Kordinaatiston valinnan perusteella on $f_x(0, 0) = f_y(0, 0) = 0$, joten seuraa

$$\begin{aligned}\frac{\partial\vec{n}}{\partial x}(0, 0) &= -a_{11}\vec{i} - a_{12}\vec{j},\\ \frac{\partial\vec{n}}{\partial y}(0, 0) &= -a_{12}\vec{i} - a_{22}\vec{j},\end{aligned}$$

missä

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} f_{xx}(0, 0) & f_{xy}(0, 0) \\ f_{xy}(0, 0) & f_{yy}(0, 0) \end{bmatrix}.$$

Matriisia $\mathbf{A}$ sanotaan (tehdyin oletuksin) pinnan *kaarevuusmatriisiksi* tarkastellussa pisteessä ja mainitulla tavalla valituissa koordinaateissa (x, y) pinnan tangenttitasossa. Luvut a_{11} ja a_{22} ilmaisevat pinnan kaarevuuden koordinaattiakselien suunnassa: Origon lähellä käyrät $S_1 :\ y = 0,\ z = f(x, 0)$ ja $S_2 :\ x = 0,\ z = f(0, y)$ ovat likimain ympyräviivoja, joiden keskipisteet ovat pisteissä $(0, 0, 1/a_{11})$ ja $(0, 0, 1/a_{22})$. Luku a_{12} on pinnan *kierevyys* origossa. Kierevyys kertoo, kuinka nopeasti (ja mihin suuntaan) pinnan normaali 'kaatuu sivulle' koordinaattiviivojen suuntaan kuljettaessa (vrt. avaruuskäyrän kierevyys: Harj.teht. VII.3:10).

Kaarevuusmatriisi luonnollisesti riippuu tarkastelun kohteena olevasta pinnan pisteestä, mutta se riippuu myös valitusta koordinaatistosta (x, y) pinnan tangenttitasossa. Onkin syytä tutkia, miten kaarevuusmatriisi muuttuu, kun koordinaattiakseleita x, y kierretään z-akselin ympäri, eli kun tehdään koordinaattimuunnos

$$\mathbf{C}\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \;\Leftrightarrow\; \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \mathbf{C}^T\begin{bmatrix} x \\ y \end{bmatrix},$$

missä $\mathbf{C}$ on ortogonaalinen matriisi (kokoa 2×2). Vastaus tähän kysymykseen saadaan, kun ajatellaan funktio $f(x, y)$ approksimoiduksi origon lähellä Taylorin kaavan mukaisesti:

$$f(x, y) = \tfrac{1}{2}\,a_{11}\,x^2 + \tfrac{1}{2}\,a_{22}\,y^2 + a_{12}\,xy + \ldots$$

Vain kehitelmän näkyvillä termeillä on merkitystä kaarevuuden kannalta, sillä jäännöstermin toisen kertaluvun osittaisderivaatat häviävät origossa. Päätellään siis, että kaarevuusmatriisi muuntuu koordinaatiston kierrossa samalla tavoin kuin neliömuoto. Toisin sanoen, kierretyssä (ξ, η)-koordinaatistossa kaarevuusmatriisi on

$$\mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}.$$

Jos tässä erityisesti valitaan **C**:n pystyriveiksi **A**:n ortonormeeratut ja positiivisesti suunnistetut ominaisvektorit, niin **B** on diagonaalinen:

$$\mathbf{B} = \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix}.$$

Sanotaan, että **B**:n lävistäjäalkiot κ_1, κ_2 ovat pinnan *pääkaarevuudet* (engl. principal curvatures) tarkasteltavassa pisteessä, ja että koordinaatisto (ξ, η) on ko. pisteessä *pääkaarevuuskoordinaatisto*. Kun valitaan $\zeta = z$, niin koordinaatistossa (ξ, η, ζ) pinnan yhtälö on siis origon ympäristössä muotoa

$$\zeta = \frac{1}{2}\kappa_1\xi^2 + \frac{1}{2}\kappa_2\eta^2 + \ldots$$

Riittävän säännöllinen pinta on siis jokaisen pisteensä P lähiympäristössä likimain joko (a) <u>elliptinen</u> <u>paraboloidi</u> ($\kappa_1\kappa_2 > 0$), (b) <u>parabolinen</u> <u>lieriö</u> tai <u>taso</u> ($\kappa_1\kappa_2 = 0$; taso jos $\kappa_1 = \kappa_2 = 0$) tai (c) <u>hyperbolinen</u> <u>parabolidi</u> ($\kappa_1\kappa_2 < 0$), vrt. toisen asteen pintojen luokittelu edellä. Tämän jaottelun mukaisesti sanotaan, että P on pinnan (a) *elliptinen*, (b) *parabolinen*, tai (c) *hyperbolinen* piste.

Jos pinnan piste P on pääkaarevuuskoordinaatiston (ξ, η, ζ) origo em. tavalla ja pintaa tarkastellaan positiivisen ζ-akselin (eli pinnan normaalin) suunnasta, niin elliptisessä tapauksessa pinta kaartuu joka suuntaan joko tarkkailijaa kohti (jos $\kappa_1 > 0$ ja $\kappa_2 > 0$) tai tarkkailijasta poispäin (jos $\kappa_1 < 0$ ja $\kappa_2 < 0$). Hyperbolisessa tapauksessa pinta kaartuu toiseen pääkaarevuussuuntaan tarkkailijaa kohti ja toiseen tarkkailijasta poispäin, eli P on nk. *satulapiste*. Parabolisessa tapauksessa pinta taas kaartuu oleellisesti vain toiseen pääkaarevuussuuntaan. Kaikissa tapauksissa kaartuminen tapahtuu pääkaarevuussuuntiin siten, että $\xi\zeta$-tason ja pinnan leikkausviiva on origon lähellä likimain ympyräviiva, jonka keskipiste on $Q_1 = (0, 0, 1/\kappa_1)$, ja vastaavasti $\eta\zeta$-tason ja pinnan leikkausviiva on likimain ympyräviiva, jonka keskipiste on $Q_2 = (0, 0, 1/\kappa_2)$. Lukuja $R_i = 1/|\kappa_i|$ (voi olla myös $R_i = \infty$) sanotaan tämän mukaisesti *pääkaarevuussäteiksi* ja pisteitä Q_i *kaarevuuskeskiöiksi*. Kaarevuuskeskiöt voi mieltää geometrisesti niin, että pinnan normaali kiertyy likimain kaarevuuskeskiön ympäri, kun normaalin ja pinnan leikkauspiste siirtyy tarkasteltavasta pisteestä P vastaavaan pääkaarevuussuuntaan. Tämän mukaisesti piste P on siis elliptinen täsmälleen kun kaarevuuskeskiöt ovat pinnan samalla puolella ja äärellisen matkan päässä ja hyperbolinen täsmälleen kun kaarevuuskeskiöt ovat eri puolilla pintaa ja äärellisen matkan päässä.

Esimerkki 1 Liikuttaessa pallopinnalla

$$S: \ (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2$$

mihin tahansa suuntaan kiertyy pallon keskipisteestä $Q = (x_0, y_0, z_0)$ poispäin osoittava normaalivektori $\vec{n}$ keskipisteen ympäri. Molemmat kaarevuuskeskiöt

ovat siis pisteessä Q. Päätellään, että jos karteesisen koordinaatiston (ξ, η, ζ) origo on pallopinnalla ja $\xi\eta$-taso on pallopinnan tangenttitaso, niin jokainen tällainen koordinaatisto on pääkaarevuuskoordinaatisto. Jos normaalivektori $\vec{n}$ on valittu mainitulla tavalla, niin kaarevuusmatriisi on

$$\mathbf{A} = \begin{bmatrix} -\frac{1}{R} & 0 \\ 0 & -\frac{1}{R} \end{bmatrix}.$$

Pinnan jokainen piste on elliptinen. $\square$

Esimerkki 2 **Torus**. xy-tason ympyräviiva

$$K = \{(x, y) \in \mathbb{R}^2 \mid (x - R)^2 + y^2 = a^2\}, \quad \text{missä} \quad 0 < a < R,$$

pyörähtää avaruudessa y-akselin ympäri. Laske näin syntyvän pinnan, *toruksen*, pääkaarevuudet sen xy-tasolla olevassa pisteessä $P = (x, y, 0)$.

<u>Ratkaisu</u> Geometrisesti voidaan päätellä, että pääkaarevuussuunnat pisteessä P ovat $\vec{\tau}$ ja $\vec{k}$, missä $\vec{\tau}$ on xy-tason suuntainen S:n yksikötangenttivektori ja $\vec{k}$ on xy-tason normaali. Nimittäin kun pisteestä P liikutaan suuntaan $\vec{\tau}$, kiertyy pinnan normaali pisteen $Q_1 = (R, 0, 0)$ ympäri. Suuntaan $\vec{k}$ liikuttaessa normaali taas kiertyy y-akselilla olevan pisteen $Q_2 = (0, c, 0)$ ympäri (ks. kuvio), koska y-akseli on pinnan symmetria-akseli. Pisteet Q_1, Q_2 ovat siis kaarevuuskeksiöt. Kun pinnan normaali $\vec{n}$ valitaan osoittamaan pisteestä Q_1 poispäin, niin näin valitussa pääkaarevuuskoordinaatistossa (ξ, η, ζ) on $Q_1 = (0, 0, -a)$ ja (ks. kuvio)

$$Q_2 = \begin{cases} (0, 0, d), & \text{jos } x < R, \\ (0, 0, -d), & \text{jos } x > R, \end{cases}$$

missä d on pisteen P etäisyys y-akselista suuntaan $\pm\vec{n}$. Tapauksessa $x < R$ päätellään yhdenmuotoisten kolmioiden avulla että (ks. kuvio)

$$\frac{a+d}{R} = \frac{a}{R-x} \quad \Rightarrow \quad d = \frac{ax}{R-x},$$

joten pääkaarevuudet ovat tässä tapauksessa

$$\underline{\underline{\kappa_1 = -\frac{1}{a}, \quad \kappa_2 = \frac{1}{a}\left(\frac{R}{x} - 1\right).}}$$

Nämä todetaan päteviksi myös tapauksessa $x \geq R$. Piste $P = (x, y, 0)$ on tämän mukaisesti elliptinen kun $x > R$, parabolinen kun $x = R$, ja hyperbolinen kun $x < R$. $\square$

Esimerkki 3 Määritä pinnan $S: \ z = \sin xy$ pääkaarevuudet ja pääkaarevuuskoordinaatisto origossa.

Ratkaisu Pinnan tangenttitaso origossa on xy-taso, ja kaarevuusmatriisi origossa on

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

Tämän ominaisarvot ovat $\lambda_1 = +1$ ja $\lambda_2 = -1$ sekä vastaavat ominaisvektorit $[1, 1]^T$ ja $[-1, 1]^T$. Siis pääkaarevuudet ovat $\kappa_1 = +1$ ja $\kappa_2 = -1$, ja pääkaarevuuskoordinaatiston kantavektorit ovat

$$\vec{e}_1 = (\vec{i} + \vec{j})/\sqrt{2}, \quad \vec{e}_2 = (-\vec{i} + \vec{j})/\sqrt{2}.$$

Tässä koordinaatistossa pinnan yhtälö on sarjamuodossa

$$\begin{aligned} z &= xy - \frac{(xy)^3}{3!} + \dots \\ &= \frac{1}{2}(\xi^2 - \eta^2) - \frac{1}{48}(\xi^2 - \eta^2)^3 + \dots \end{aligned}$$

Pinta on siis origon lähellä likimain hyperbolinen paraboloidi (satulapinta), eli origo on pinnan hyperbolinen piste. $\square$

Esimerkki 4 Määritä pinnan $S: \ z = xy - y^2$ pääkaarevuudet ja -suunnat pisteessä $P = (1, 1, 0)$.

Ratkaisu Jos $z(x, y) = xy - y^2$, niin $z_x(1, 1) = 1$ ja $z_y(1, 1) = -1$, joten pinnan yksikkönormaalivektori pisteessä $(1, 1, 0)$ on

$$\vec{n} = \frac{1}{\sqrt{3}}(-\vec{i} + \vec{j} + \vec{k}).$$

Yksikkötangenttivektoreita ovat

$$\vec{t}_1 = \frac{1}{\sqrt{2}}(\vec{i} + \vec{j}), \quad \vec{t}_2 = \vec{n} \times \vec{t}_1 = \frac{1}{\sqrt{6}}(-\vec{i} + \vec{j} - 2\vec{k}),$$

ja $\{\vec{t}_1, \vec{t}_2, \vec{n}\}$ on ortonormeerattu, positiivisesti suunnistettu systeemi. Suoritetaan koordinaattimuunnos $(x, y, z) \hookrightarrow (\xi, \eta, \zeta)$ koordinaatistoon $\{P, \vec{t}_1, \vec{t}_2, \vec{n}\}$:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = \begin{bmatrix} x-1 \\ y-1 \\ z \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} x &= \frac{\xi}{\sqrt{2}} - \frac{\eta}{\sqrt{6}} - \frac{\zeta}{\sqrt{3}} + 1, \\ y &= \frac{\xi}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} + \frac{\zeta}{\sqrt{3}} + 1, \\ z &= \quad\; - \frac{2\eta}{\sqrt{6}} + \frac{\zeta}{\sqrt{3}}\,. \end{cases}$$

Sijoitus yhtälöön $z = y(x - y)$ antaa

$$\begin{aligned} -\frac{2}{\sqrt{6}}\,\eta + \frac{1}{\sqrt{3}}\,\zeta &= \left(\frac{1}{\sqrt{2}}\,\xi + \frac{1}{\sqrt{6}}\,\eta + \frac{1}{\sqrt{3}}\,\zeta + 1\right)\left(-\frac{2}{\sqrt{6}}\,\eta - \frac{2}{\sqrt{3}}\,\zeta\right) \\ \Leftrightarrow \quad 3\sqrt{3}\,\zeta + \sqrt{6}\,\xi\zeta + 2\sqrt{2}\,\eta\zeta + \zeta^2 + \sqrt{3}\,\xi\eta + \eta^2 &= 0. \end{aligned}$$

Tämä määrää implisiittisesti funktion $\zeta = \zeta(\xi, \eta)$, jolle $\zeta(0,0) = \zeta_\xi(0,0) = \zeta_\eta(0,0) = 0$. Implisiittisesti derivoimalla saadaan lasketuksi toisen kertaluvun osittaisderivaatat (vrt. Harj.teht. 3):

$$\zeta_{\xi\xi}(0,0) = 0, \quad \zeta_{\xi\eta}(0,0) = -\frac{1}{3}\,, \quad \zeta_{\eta\eta}(0,0) = -\frac{2}{3\sqrt{3}}\,.$$

Kaarevuusmatriisi valitussa koordinaatistossa on siis

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3\sqrt{3}} \end{bmatrix}.$$

A:n ominaisarvot ovat

$$\lambda_1 = \frac{1}{3\sqrt{3}}\,, \quad \lambda_2 = -\frac{1}{\sqrt{3}}\,,$$

ja ominaisvektorit ovat suorilla

$$\lambda_1 : \ \xi + \sqrt{3}\,\eta = 0, \quad \lambda_2 : \ \sqrt{3}\,\xi - \eta = 0.$$

Näiden suorien suuntavektoreita ovat

$$\begin{aligned} \vec{v}_1 &= \sqrt{3}\,\vec{t}_1 - \vec{t}_2 = \frac{2}{\sqrt{6}}(2\vec{i} + \vec{j} + \vec{k}), \\ \vec{v}_2 &= \vec{t}_1 + \sqrt{3}\,\vec{t}_2 = \sqrt{2}\,(\vec{j} - \vec{k}). \end{aligned}$$

Siis pääkaarevuudet ja -suunnat alkuperäisessä koordinaatistossa ovat

$$\kappa_1 = \underline{\underline{\frac{1}{3\sqrt{3}}}}, \quad \text{suunta } \underline{\underline{2\vec{i} + \vec{j} + \vec{k}}},$$
$$\kappa_2 = \underline{\underline{-\frac{1}{\sqrt{3}}}}, \quad \text{suunta } \underline{\underline{\vec{j} - \vec{k}}}.$$

Piste P on pinnan hyperbolinen piste. □

HARJOITUSTEHTÄVIÄ

1. Määritä pinnan pääkaarevuudet ja pääkaarevuuskoordinaatisto annetussa pisteessä:
   a) $z = (x+y)(x-2y)$, $(0,0,0)$ b) $z = 2\sin xy + \cos(x+y)$, $(0,0,1)$
   c) $z = 1 - x + 2y + e^{x-2y}$, $(0,0,2)$ d) $x^2 + 2y^2 - 2xy + z^2 = 1$, $(0,0,-1)$
   e) $\cos(x+y) + \sin(x+y+z) - \cos z = 0$, $(0,0,0)$

2. Käyrä $K : y = 5\sin x$ pyörähtää avaruudessa x-akselin ympäri. Määritä syntyvän pyörähdyspinnan pääkaarevuussäteet ja -suunnat pisteessä
   a) $(\pi/6, 5/2, 0)$, b) $(\pi/2, 4, -3)$, c) $(3\pi/4, -5/2, -5/2)$.

3. Tarkastellaan pintaa $S : F(x,y,z) = 0$, missä F:n osittaisderivaatat ovat jatkuvia toiseen kertalukuun asti. Näytä, että jos taso $T : z = z_0$ on S:n tangenttitaso pisteessä $P = (x_0, y_0, z_0) \in S$ ja $F_z(P) \neq 0$, niin S:n kaarevuusmatriisi pisteessä P on
$$\mathbf{A} = -\frac{1}{F_z(P)} \begin{bmatrix} F_{xx}(P) & F_{xy}(P) \\ F_{xy}(P) & F_{zz}(P) \end{bmatrix}.$$

4. (*) Määritä pinnan pääkaarevuudet ja pääkaarevuuskoordinaatisto annetussa pisteessä:
   a) $9x^2 + 4y^2 + z^2 = 14$, $(1,-1,1)$ b) $2x^2 - 2y^2 - z^2 = 1$, $(1,0,-1)$
   c) $xy + 2yz - 3xz = 0$, $(1,1,1)$ d) $xy^2z^3 = 1$, $(1,1,1)$

5. (*) Suora S_t liikkuu avaruudessa parametrin $t \in \mathbb{R}$ mukaan siten, että suora kulkee pisteen $(0,0,2t)$ kautta ja suoran suuntavektori on $\vec{a}(t) = \cos t\,\vec{i} + \sin t\,\vec{j}$. Määritä suoran synnyttämän viivoitinpinnan pääkaarevuudet ja pääkaarevuussuunnat origossa.

XII.4 *Tensorit, vektorit ja skalaarit

Kuten edellä on todettu, on neliömuodolla $\mathbf{x}^T\mathbf{A}\mathbf{x}$ se ominaisuus, että koordinaatistoa kierrettäessä neliömuodon matriisi muuntuu similaarisesti:

$$\mathbf{x} \hookrightarrow \mathbf{y} = \mathbf{C}^T\mathbf{x} \;\Rightarrow\; \mathbf{A} \hookrightarrow \mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}.$$

On myös nähty, että pinnan kaarevuusmatriisilla on tämä sama ominaisuus, kun koordinaatiston kierto tapahtuu pinnan normaalin ympäri. Yleisesti sanotaan sellaista oliota, jolla on jokaisessa koordinaatistossa matriisin olomuoto, ja joka muuntuu koordinaatistoa kierrettäessä ym. säännön mukaan, *tensoriksi*.† Neliömuoto, matriisinsa kautta nähtynä, on siis tensori, itse asiassa *symmetrinen tensori*. Myös pinnan kaarevuus on symmetrinen tensori, ja kaarevuusmatriisin sijasta puhutaankin yleensä *kaarevuustensorista*. Näin puhuen kaarevuus tulee selvemmin ymmärretyksi pinnan paikalliseksi geometriseksi ominaisuudeksi, joka vain näyttäytyy erilaisena eri koordinaatistoissa.

Jos $f = f(x, y, z)$ on kahdesti jatkuvasti derivoituva, niin tietyssä pisteessä laskettu Hessen matriisi

$$\mathbf{A} = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{bmatrix}$$

muuntuu koordinaatistoa kierrettäessä ym. similaarisuussäännön mukaisesti. — Tämä nähdään approksimoimalla f:ää toisen asteen Taylorin polynomilla. Silloin voidaan ajatella, että f:n 'toinen derivaatta' ei ole mihinkään yksittäiseen koordinaatistoon sidottu toisten osittaisderivaattojen matriisi, vaan pikemmin <u>tensori</u>, jonka ilmiasuja nämä matriisit ovat (!).

Myös <u>lineaarikuvaus</u> $A : \mathbb{R}^n \to \mathbb{R}^n$ voidaan ymmärtää tensorina, sillä jos

$$A: \; \mathbf{x} \mapsto \mathbf{y} \;\Leftrightarrow\; \mathbf{y} = \mathbf{A}\mathbf{x},$$

niin nähdään, että koordinaattimuunnoksessa

$$\mathbf{x} = \mathbf{C}\mathbf{x}', \quad \mathbf{y} = \mathbf{C}\mathbf{y}' \quad (\mathbf{C}^{-1} = \mathbf{C}^T)$$

lineaarikuvaus saa muodon

$$\mathbf{y}' = \mathbf{B}\mathbf{x}', \quad \mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}.$$

†Tensori on matematiikassa tässä esitettyä yleisempi käsite. Tensorit liitetään, ei ainoastaan karteesisen koordinaatiston kiertoihin, vaan yleisempien käyräviivaisten kordinaatistojen välisiin muunnoksiin. Muunnoksiin liittyvä differentiaalilaskenta, nk. *tensorilaskenta*, on oma matematiikan lajinsa. Tensoreilla ja tensorilaskennalla on paljon käyttöä fysiikassa.

Esimerkki 1 **Magneettinen permeabiliteetti**. Sähkömagnetiikassa magneettivuon tiheyden ($\vec{B}$) ja magneettikentän voimakkuuden ($\vec{H}$) välinen yksinkertaisin materiaalilaki on

$$\vec{B} = \mu \vec{H},$$

missä materiaalivakio μ on väliaineen magneettinen permeabiliteetti. Tämä laki on sama kaikissa koordinaatistoissa. Yleisemmin, jos materiaali on epäisotrooppinen, voi $\vec{B}$:n ja $\vec{H}$:n riippuvuus olla yleisempi lineaarikuvaus muotoa

$$\mathbf{B} = \boldsymbol{\mu}\mathbf{H} \quad \Leftrightarrow \quad \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}.$$

Tällöin $\boldsymbol{\mu}$ on tensori, eli matriisi (μ_{ij}) muuntuu koordinaatistoa kierrettäessä similaarisesti. $\square$

Esimerkki 2 Kahdesti jatkuvasti derivoituvasta funktiosta $f = f(x, y)$ tiedetään, että $f_{xx}(0,0) = 24$, $f_{yy}(0,0) = -24$ ja $f_{xy}(0,0) = -7$. Laske $f_{\xi\xi}(0,0)$, $f_{\xi\eta}(0,0)$ ja $f_{\eta\eta}(0,0)$ koordinaatistossa (ξ, η), jonka kantavektorit ovat

$$\vec{e}_\xi = \frac{1}{5}(4\vec{i} + 3\vec{j}), \quad \vec{e}_\eta = \frac{1}{5}(-3\vec{i} + 4\vec{j}).$$

Ratkaisu Koordinaattimuunnos $(x, y) \hookrightarrow (\xi, \eta)$ määräytyy ehdosta

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \mathbf{C} \begin{bmatrix} \xi \\ \eta \end{bmatrix}.$$

Koska f:n toisen kertaluvun osittaisderivaatat muodostavat tensorin, on

$$\begin{aligned} \begin{bmatrix} f_{\xi\xi}(0,0) & f_{\xi\eta}(0,0) \\ f_{\xi\eta}(0,0) & f_{\eta\eta}(0,0) \end{bmatrix} &= \mathbf{C}^T \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{xy}(0,0) & f_{yy}(0,0) \end{bmatrix} \mathbf{C} \\ &= \frac{1}{25} \begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 24 & -7 \\ -7 & -24 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -25 \\ -25 & 0 \end{bmatrix}. \end{aligned}$$

Siis $f_{\xi\xi}(0,0) = f_{\eta\eta}(0,0) = 0$, $f_{\xi\eta} = -25$. — Eräs oletukset toteuttava funktio on

$$f(x, y) = (4x + 3y)(3x - 4y) = -25\xi\eta. \quad \square$$

Vektorit

Jos tensori on olio, joka 'näyttää matriisilta' jokaisessa (karteesisessa) koordinaatistossa, niin tason tai avaruuden vektori voidaan vastaavasti tulkita olioksi, jonka olomuoto on lukupari tai lukukolmikkko jokaisessa koordinaatistossa. Kuten tensori, tämä pari tai kolmikko muuntuu koordinaatistoa kierrettäessä systemaattisella tavalla. Esimerkiksi jos $\vec{v} = x\vec{i} + y\vec{j} + z\vec{k}$ on avaruusvektori, niin muunnossääntö $\mathbf{x} \mathrel{\widehat{=}} (x, y, z) \mapsto (x', y', z') \mathrel{\widehat{=}} \mathbf{x}'$ siirryttäessä kierrettyyn koordinaatistoon, jonka kantavektorien koordinaatit kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$ ovat ortogonaalimatriisin $\mathbf{C}$ sarakkeet, on (vrt. Luku X.6)

$$\mathbf{C}\mathbf{x}' = \mathbf{x} \quad \Leftrightarrow \quad \mathbf{x}' = \mathbf{C}^T\mathbf{x}.$$

Siis ei ainoastaan tensorin, vaan myös vektorin 'perimmäinen olemus' paljastuu vasta koordinaatistoa kierrettäessä (!). — Havaitaan myös, että symmetrisellä tensorilla ja vektorilla on se yhteinen piirre, että löytyy koordinaatisto, tai koordinaatistoja, joissa tensorin/vektorin ilmiasu matriisina/$\mathbb{R}^n$:n alkiona saa mahdollisimman yksinkertaisen muodon. Symmetrisen tensorin tapauksessa tämä yksinkertaisin muoto on diagonaalimatriisi. Tason tai avaruuden vektorin yksinkertaisin muoto on lukupari $(a, 0)$ tai lukukolmikko $(a, 0, 0)$, missä $a > 0$ on vektorin pituus. Tällaiseen koordinaatistoon siirryttäessä palataan itse asiassa vektorin alkuperäiseen geometriseen ideaan — vektorihan oli alunperin suunnattu jana. Vektorin suunnan pysyvyys koordinaatistoa kierrettäessä varmistetaan em. muunnossäännöllä. Sen mukaan siis vektorin ilmiasu $\mathbb{R}^n$:n alkiona <u>ei</u> ole pysyvä, vaan nimenomaan muuttuu koordinaatistoa kierrettäessä, muunnossäännön ilmaisemalla systemaattisella tavalla.

Esimerkki 3 Voidaan kuvitella (etenkin tapauksissa $n = 2$ ja $n = 3$), että euklidinen avaruus $\mathbb{R}^n$ on pisteavaruuden E^n koordinaattiavaruus jossakin valitussa (karteesisessa) koordinaatistossa. Jos nyt valittua pistettä $P \in E^n$ edustaa vektori $\mathbf{x}$ (paikkavektori), niin koordinaatistoa kierrettäessä P:n koordinaatit muuntuvat kaavan $\mathbf{x}' = \mathbf{C}^T\mathbf{x}$ mukaisesti. Muunnoskaava varmistaa, että piste 'pysyy paikallaan', eli pisteen paikkavektori on vektori juuri tämän vaatimuksen vuoksi. Vastaavasti jos tarkastellaan funktion $f(\mathbf{x})$ muuntumista koordinaatiston kierrossa, niin muunnossäännöllä $g(\mathbf{x}') = f(\mathbf{C}\mathbf{x}')$ tarkoitetaan, että 'funktio pysyy paikallaan', ts. funktion arvo tietyssä (paikallaan pysyvässä) pisteessä ei muutu. $\square$

Esimerkki 4 Olkoon $f = f(x, y, z)$ on differentioituva funktio. Kierretään koordinaatistoa muunnoksella $\mathbf{C}\mathbf{x}' = \mathbf{x} \Leftrightarrow \mathbf{x}' = \mathbf{C}^T\mathbf{x}$ ($\mathbf{x} = [x, y, z]^T$, $\mathbf{x}' = [x', y', z']^T$) ja merkitään $g(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{C}\mathbf{x}')$. Olkoon $\nabla f(0, 0, 0) = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$. Tällöin

jos $\mathbf{a} = [a_1, a_2, a_3]^T$, niin differentioituvuuden määritelmän mukaisesti on

$$f(x, y, z) = f(0, 0, 0) + \mathbf{a}^T\mathbf{x} + o(|\mathbf{x}|).$$

Ketjusääntöjen perusteella myös g on differentioituva origossa, joten jollakin $\mathbf{b} = [b_1, b_2, b_2]^T$ on

$$g(x', y', z') = g(0, 0, 0) + \mathbf{b}^T\mathbf{x}' + o(|\mathbf{x}'|) = f(0, 0, 0) + \mathbf{b}^T\mathbf{x}' + o(|\mathbf{x}'|).$$

Kun näissä hajotelmissa oletetaan, että $\mathbf{C}\mathbf{x}' = \mathbf{x}$, niin $|\mathbf{x}'| = |\mathbf{x}|$ ja $g(x', y', z') = f(x, y, z)$, joten on oltava

$$\mathbf{a}^T\mathbf{x} = \mathbf{b}^T\mathbf{x}' = \mathbf{b}^T\mathbf{C}^T\mathbf{x} = (\mathbf{C}\mathbf{b})^T\mathbf{x} \quad \Rightarrow \quad \mathbf{C}\mathbf{b} = \mathbf{a}.$$

Tulos merkitsee, että funktion gradientti (tulkittuna pystyvektorina) muuntuu koordinaatistoa kierrettäessä kuten vektori. Samaan tulokseen tullaan myös laskemalla g:n gradientti suoraan ketjusääntöjen avulla kaavasta $g(\mathbf{x}') = f(\mathbf{C}\mathbf{x}')$. □

Funktion gradientti on jo aiemminkin tulkittu vektoriksi, mutta tällöin tulkinta perustui vain gradientin 'ulkonäköön' kiinteässä koordinaatistossa. Nyt voidaan siis vahvistaa, että tulkinta kestää myös koordinaatiston kierron:

Gradientti *on* vektori.

Skalaarit

Jos tensorin ja vektorin todellinen luonne paljastuu vasta koordinaatiston kierrossa, niin samoin on skalaarin laita. Skalaari on olio, jonka ilmiasu jokaisessa koordinaatistossa on luku (tässä reaaliluku), ja nimenomaan aina sama luku. Esimerkiksi jos $f = f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, on reaaliarvoinen funktio, niin f:n arvo tietyssä (kiinteäksi ajatellussa) pisteessä on skalaari. Vektoriin liittyvä skalaari on vektorin pituus, koska pituus on reaaliluku, joka ei muutu koordinaatiston kierrossa. Kahteen vektoriin $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ liittyvä skalaari on (nimensä mukaisesti) skalaaritulo, sillä jos $\mathbf{x}' = \mathbf{C}^T\mathbf{x}$ ja $\mathbf{y}' = \mathbf{C}^T\mathbf{y}$, niin

$$(\mathbf{x}')^T\mathbf{y}' = (\mathbf{C}^T\mathbf{x})^T\mathbf{C}^T\mathbf{y} = \mathbf{x}^T\mathbf{C}\mathbf{C}^T\mathbf{y} = \mathbf{x}^T\mathbf{I}\mathbf{y} = \mathbf{x}^T\mathbf{y}.$$

Tensoreihin liittyvä mielenkiintoinen skalaari on tensorin *jälki*:

MÄÄRITELMÄ XII.4.1 Jos T on $\mathbb{R}^n$:n tensori, jonka matriisi annetussa ortonormeeratussa koordinaatistossa on $\mathbf{A} = (a_{ij})$, niin T:n *jälki* (engl. trace) on

$$tr\, T = \sum_{i=1}^{n} a_{ii}\,.$$

LAUSE XII.4.2 Tensorin jälki on skalaari.

Todistus Olkoon $\mathbf{C}$ ortogonaalimatriisi kokoa $n \times n$ ja $\mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}$. Lause väittää, että

$$\sum_{i=1}^{n} [\mathbf{B}]_{ii} = \sum_{i=1}^{n} [\mathbf{A}]_{ii} \,.$$

Lähdetään matriisitulon määritelmästä ja vaihdetaan summausjärjestystä:

$$\begin{aligned}
\sum_{i=1}^{n} [\mathbf{B}]_{ii} &= \sum_{i=1}^{n} \sum_{j=1}^{n} [\mathbf{C}^T]_{ij} \, [\mathbf{A}\mathbf{C}]_{ji} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} [\mathbf{C}^T]_{ij} \sum_{k=1}^{n} [\mathbf{A}]_{jk} \, [\mathbf{C}]_{ki} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} [\mathbf{A}]_{jk} \left(\sum_{i=1}^{n} [\mathbf{C}]_{ki} \, [\mathbf{C}^T]_{ij} \right) .
\end{aligned}$$

Tässä on $\mathbf{C}$:n ortogonaalisuuden perusteella

$$\sum_{i=1}^{n} [\mathbf{C}]_{ki} \, [\mathbf{C}^T]_{ij} = [\mathbf{C}\mathbf{C}^T]_{kj} = [\mathbf{I}]_{kj} = \delta_{kj} \,,$$

joten

$$\sum_{i=1}^{n} [\mathbf{B}]_{ii} = \sum_{j=1}^{n} \sum_{k=1}^{n} [\mathbf{A}]_{jk} \delta_{kj} = \sum_{j=1}^{n} [\mathbf{A}]_{jj} \,. \quad \square$$

ESIMERKKI 5 Tarkastellaan funktion $u = u(x, y)$ 'toista derivaattaa', eli tensoria T, jonka matriisi on

$$\mathbf{A} = \begin{bmatrix} u_{xx}(x, y) & u_{xy}(x, y) \\ u_{xy}(x, y) & u_{yy}(x, y) \end{bmatrix} .$$

Kun koordinaatistoa kierretään muunnoksella $\mathbf{C}[\xi, \eta]^T = [x, y]^T$ ja merkitään $v(\xi, \eta) = u(x(\xi, \eta), y(\xi, \eta))$, niin T:n matriisi kierretyssä koordinaatistossa on

$$\begin{bmatrix} v_{\xi\xi}(\xi, \eta) & v_{\xi\eta}(\xi, \eta) \\ v_{\xi\eta}(\xi, \eta) & v_{\eta\eta}(\xi, \eta) \end{bmatrix} = \mathbf{C}^T \begin{bmatrix} u_{xx}(x, y) & u_{xy}(x, y) \\ u_{xy}(x, y) & u_{yy}(x, y) \end{bmatrix} \mathbf{C}.$$

Lauseen XII.4.2 mukaan pätee

$$(\partial_\xi^2 + \partial_\eta^2) v(\xi, \eta) \;=\; (\partial_x^2 + \partial_y^2) u(x, y). \quad \square$$

Esimerkistä voidaan vetää se (yleisemminkin $\mathbb{R}^n$:ssä pätevä) merkittävä johtopäätös, että Laplacen operaattori Δ 'näyttää samalta' kaikissa karteesisissa koordinaatistoissa. Koska Δ myös operoi sekä skalaareihin että vektoreihin kuten skalaari, niin voidaan kahdessa eri merkityksessä sanoa:

Δ *on skalaarinen operaattori.*

HARJOITUSTEHTÄVIÄ

1. Kahdesti jatkuvasti derivoituvasta funktiosta $f(x,y)$ tiedetään: $f_{xx}(0,0) = 0$, $f_{yy}(0,0) = -2$ ja $f_{xy}(0,0) = 2$. Laske $f_{\xi\xi}(0,0)$ ja $f_{\eta\eta}(0,0)$ sellaisessa kierretyssä koordinaatistossa, jossa $f_{\xi\eta}(0,0) = 0$.

2. Eräässä materiaalissa magneettikentän tiheyden ja magneettikentän voimakkuuden välillä on riippuvuus $\mathbf{B} = \boldsymbol{\mu}\mathbf{H}$, missä $\boldsymbol{\mu}$ on tensori, jonka matriisiesitys on
$$\boldsymbol{\mu} = \begin{bmatrix} \mu & 0 & 0 \\ 0 & 2\mu & 0 \\ 0 & 0 & 3\mu \end{bmatrix}$$
koordinaatistossa, jonka kantavektorit ovat
$$\vec{e}_1 = \frac{1}{5}(3\vec{i} + 4\vec{j}), \quad \vec{e}_2 = \frac{1}{25}(12\vec{i} - 9\vec{j} + 20\vec{k}), \quad \vec{e}_3 = \frac{1}{25}(16\vec{i} - 12\vec{j} - 15\vec{k}).$$
Millainen on $\boldsymbol{\mu}$:n esitysmuoto peruskoordinaatiston kannassa $\{\vec{i}, \vec{j}, \vec{k}\}$?

3. Näytä, että jos $\vec{a}$ ja $\vec{F}$ ovat vektorikenttiä (eli vektoriarvoisia funktioita), niin a) $\nabla \cdot \vec{F}$ on skalaarikenttä, b) $(\vec{a} \cdot \nabla)\vec{F}$ on vektorikenttä.

4. (*) a) Näytä, että differentiaalioperaattori
$$A = a\partial_x^2 + b\partial_y^2 + c\partial_x\partial_y,$$
missä a, b ja c ovat skalaareita ($a, b, c \in \mathbb{R}$), säilyy muuttumattomana $\mathbb{R}^2$:n koordinaatiston kierrossa täsmälleen kun $a = b$ ja $c = 0$, eli kun $A = a\Delta$.
b) Millaiseksi osittaisdifferentialiyhtälö $u_{xx} - u_{yy} = 0$ muuntuu kierretyssä koordinaatistossa, jonka kantavektorit ovat
$$\vec{e}_1 = \frac{1}{\sqrt{2}}(\vec{i} + \vec{j}), \quad \vec{e}_2 = \frac{1}{\sqrt{2}}(-\vec{i} + \vec{j})\,?$$
c) Jos $\vec{a} = \vec{i} + \vec{j}$, niin millaisen muodon differentiaalioperaattori
$$(\vec{a} \cdot \nabla)^2 = (\partial_x)^2 + (\partial_y)^2 + 2\partial_x\partial_y$$
saa b-kohdan kierretyssä koordinaatistossa?

Luku XIII

Usean muuttujan integraalilaskenta

Tämän luvun sisällön muodostavat erilaiset <u>määrätyn</u> <u>integraalin</u> käsitteen laajennukset useampiin ulottuvuuksiin ja erilaisiin geometrisiin tilanteisiin. Tarkasteltavia integraalien lajeja ovat kahden tai useamman muuttujan funktioihin liitettävät *taso-* ja *avasuusintegraalit*, *viiva-* eli *käyräintegraalit* ja *pintaintegraalit*. Sovelluksina käsitellään mm. pinta-alojen, tilavuuksien, kaarenpituuksien, *painopisteiden*, *hitausmomenttien*, ym. laskemista.

Luvussa XIII.1 tarkastellaan aluksi tasointegraalien *mittateoreettisia* perusteita. Osoittautuu, että tasointegraalissa, samoin muissakin integraalien lajeissa, on kyse integroimisesta jonkin *mitan* suhteen. Tasointegraaliin liittyvä mitta on $\mathbb{R}^2$:n osajoukkoihin liitettävä *pinta-alamitta*, joka Luvussa XIII.1 määritellään tarkemmin *Jordanin* mittana. Paitsi tämän mitan ominaisuuksia ja suhdetta integraaliin, Luvussa XIII.1 tarkastellaan myös tasointegraalin laskemista numeerisesti 'suoraan määritelmästä'.

Usean muuttujan integraalilaskun keskeistä sisältöä on väistämättä integraalien klassinen laskutekniikka, jonka perusta on *Fubinin lause* (Luku XIII.2). Lause palauttaa taso- ja avaruusintegraalit peräkkäisiksi eli *iteroiduiksi* 1-ulotteisiksi integraaleiksi, jotka voidaan suotuisissa oloissa laskea suljetussa muodossa. Fubinin lauseen käyttöä tarkastellaan Luvuissa XIII.2–XIII.3.

Integraalien klassisessa laskutekniikkassa auttaa usein myös *muuttujan vaihto*, esimerkiksi siirtyminen tason tai avaruuden käyräviivaisiin koordinaatistoihin. Muuttujan vaihdon laskutekniikkaa tarkastellaan Luvussa XIII.4.

Taso- ja avaruusintegraalien moninaisia sovelluksia käydään läpi Luvussa XIII.5. Lopuksi Luvuissa XIII.6–XIII.7 tarkastellaan taso- ja avaruuskäyriin sekä $\mathbb{R}^3$:n kaareviin pintoihin liittyviä viiva- ja pintaintegraaleja sovelluksineen.

XIII.1 Pinta-alamitta ja tasointegraalit

Tasoalueen pinta-alan määräämistä määrätyn integraalin avulla on tarkasteltu aiemmin Luvussa VIII.8. Tässä luvussa tarkastelun kohteena ovat yleisemmät *tasointegraalit* muotoa

$$\int_A f\,d\mu = I(f, A, \mu), \qquad (\star)$$

missä $A \subset \mathbb{R}^2$, f on A:ssa määritelty (kahden muuttujan) funktio ja μ on $\mathbb{R}^2$:n *pinta-alamitta*. Integraali $I(f, A, \mu)$ on tarkemmin reaaliluku, joka riippuu funktiosta f, joukosta A ja mitasta μ, kuten ilmenee lukutavasta: f:n integraali yli A:n mitan μ suhteen. — Kyseessä on ennestään tutun määrätyn integraalin yleistys; nimittäin kuten jäljempänä havaitaan, merkintä $(\star)$ on pätevä myös, kun f on välillä $A = [a, b]$ määritelty (integroituva) funktio, $I(f, A, \mu) = \int_a^b f(x)\,dx$ ja μ on $\mathbb{R}$:n *pituusmitta*.

Jordanin pinta-alamitta

Koska tasointegraali määritellään suhteessa pinta-alamittaan, on tämä ensin määriteltävä. *Mitta* (engl. measure) on yleisemmin *joukkofunktio*, joka liittää *mitalliseen* (engl. measurable) joukkoon (tässä $\mathbb{R}^2$:n osajoukkoon) reaaliluvun:

$$\mu : \mathcal{M} \longrightarrow \mathbb{R}, \quad \mathcal{M} = \{\text{mitalliset joukot } A \subset \mathbb{R}^2\}.$$

Pinta-alamitaksi oletetaan jatkossa *Jordan-mitta*†, joka jatkossa määritellään niin, että seuraavat kolme aksioomaa toteutuvat. Aksioomista ensimmäisessä asetetaan koordinattiakselien mukaan suunnatun nk. *perussuorakulmion* mitta. Muut kaksi ovat yleisempienkin *positiivisten* mittojen aksioomia (tai aksioomien seurauksia).

A1. *Perussuorakulmion mitta*: $T = \{(x, y) \in \mathbb{R}^2 \mid x \in [a_1, b_1] \ \ \& \ \ y \in [a_2, b_2]\}$:
$\mu(T) = (b_1 - a_1)(b_2 - a_2)$.

A2. *Additiivisuus*: $A, B \in \mathcal{M} \ \ \& \ \ A \cap B = \emptyset$
$\Rightarrow \ A \cup B \in \mathcal{M} \ \ \& \ \ \mu(A \cup B) = \mu(A) + \mu(B)$.

A3. *Positiivisuus*: $\mu(A) \geq 0 \quad \forall A \in \mathcal{M}$.

†Jordan-mittaa sanotaan myös *Peanon–Jordanin* mitaksi. Termit viittaavat matemaatikoihin **Camille Jordan** (ransk. 1838-1922) ja **Giuseppe Peano** (ital. 1858-1932).

Huomattakoon, että aksiooman A1 mukaisesti mitta μ on lähtökohtaisesti valitusta (karteesisesta) koordinaatistosta riippuva. Mittoja käsittelevälle matematiikan lajille, *mittateorialle*, tällainen lähtökohta on tyypillinen, jolloin mitan riippumattomuus koordinaatistosta (mikäli tosi) on aina erikseen osoitettava.

Tyhjä joukko $\emptyset$ katsotaan aina mitalliseksi, jolloin aksiooman A2 mukaan on oltava $\mu(\emptyset) = 0$. Yleisemmin jos $A \in \mathcal{M}$ ja $\mu(A) = 0$, sanotaan että A on *nollamittainen*. Nollamittaiset joukot ovat mittateoriassa tärkeällä sijalla.†

Esimerkki 1 Koordinaattiakselien suuntaiset janat $A = \{a\} \times [b, c]$ ja $B = [a, b] \times \{c\}$ voidaan tulkita suorakulmion erikoistapauksiksi, jolloin aksiooman A1 mukaan $\mu(A) = \mu(B) = 0$. $\square$

Esimerkki 2 Jos $A = [a_1, b_1] \times [a_2, b_2]$ ja $B = [b_1, c_1] \times [a_2, b_2]$, niin $A \cup B = [a_1, c_1] \times [a_2, b_2]$, joten aksiooman A1 mukaan on $\mu(A \cup B) = \mu(A) + \mu(B)$. $\square$

Jälkimmäisessä esimerkissä additiivisuusaksiooma A2 toteutuu, vaikka $A \cap B \neq \emptyset$. Tämä johtuu siitä, että (edellisen esimerkin mukaan) $A \cap B$ on nollamittainen. — Yleisemminkin voi aksiooomassa A2 ehdon $A \cap B = \emptyset$ lieventää ehdoksi $\mu(A \cap B) = 0$ (ks. Jordan-mitan määrittely jäljempänä ja Harj.teht. 3c).

Integraali yli rajoitetun joukon

Olkoon $A \subset \mathbb{R}^2$ rajoitettu joukko, jolloin se sisältyy johonkin perussuorakulmioon:

$$A \subset T = [a_1, b_1] \times [a_2, b_2]$$

(ks. kuvio alla). Olkoon f A:ssa määritelty rajoitettu funktio ja olkoon $f_0 = f$:n *nollajatko* A:n ulkopuolelle:

$$f_0(x, y) = \begin{cases} f(x, y), & \text{kun } (x, y) \in A, \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin sovitaan ensinnäkin, että

$$\boxed{\int_A f \, d\mu = \int_T f_0 \, d\mu, \quad A \subset T.}$$

†Nollamittaisuuteen liittyy erikoinen matematiikan termi *melkein kaikkialla* (engl. almost everywhere), joka tarkoittaa: muualla kuin nollamittaisessa (osa)joukossa.

Kuten määrätyssä integraalissa, integraalin $\int_T f_0 \, d\mu$ määrittelyn lähtökohtana on T:n *jako* (ositus), tässä tapauksessa perussuorakulmion muotoisiin osajoukkoihin:

$$\mathcal{T}_h = \{T_{kl},\ k = 1, \ldots, m,\ l = 1, \ldots, n\},$$

$$T_{kl} = [x_{k-1}, x_k] \times [y_{l-1}, y_l], \qquad \begin{cases} a_1 = x_0 < x_1 < \ldots x_m = b_1, \\ a_2 = y_0 < y_1 < \ldots < y_n = b_2. \end{cases}$$

Jaon (eli suorakulmiojoukon $\mathcal{T}_h$) indeksoinnissa käytetty parametri h on jälleen *tiheysparametri*. Tämän määritelmä kahdessa ulottuvuudessa on

$$h = \max\{h_x,\, h_y\}, \quad h_x = \max_k\{x_k - x_{k-1}\}, \quad h_y = \max_l\{y_l - y_{l-1}\}.$$

Integraali $\int_T f_0\,d\mu$ määritellään nyt määrätyn integraalin tapaan eli *Riemannin summien* avulla raja-arvoprosessilla (vrt. Määritelmät VIII.5.3 ja VIII.5.1)

$$\int_T f_0\,d\mu = \operatorname*{Lim}_{h\to 0} \sum_{k=1}^{m}\sum_{l=1}^{n} f_0(\xi_{kl},\eta_{kl})\mu(T_{kl}),$$

missä $(\xi_{kl},\eta_{kl}) \in T_{kl}$. Vaihtoehtoinen, etenkin teoreettisissa tarkasteluissa kätevä integroituvuuden kriteeri saadaan, kun määritellään jakoon $\mathcal{T}_h$ liittyen

$$-\ \textit{yläsumma}: \quad \overline{\sigma}_h(f_0,\mathcal{T}_h) = \sum_{k=1}^{m}\sum_{l=1}^{n} M_{kl}\,\mu(T_{kl}), \quad M_{kl} = \sup_{(x,y)\in T_{kl}} f_0(x,y),$$

$$-\ \textit{alasumma}: \quad \underline{\sigma}_h(f_0,\mathcal{T}_h) = \sum_{k=1}^{m}\sum_{l=1}^{n} m_{kl}\,\mu(T_{kl}), \quad m_{kl} = \inf_{(x,y)\in T_{kl}} f_0(x,y).$$

Tällöin f on integroituva yli A:n täsmälleen kun pätee (vrt. Lause VIII.5.8)

$$\sup_{\mathcal{T}_h} \underline{\sigma}_h(f_0,\mathcal{T}_h) = \underline{I}(f,\mu,A) = \overline{I}(f,\mu,A) = \inf_{\mathcal{T}_h} \overline{\sigma}_h(f_0,\mathcal{T}_h),$$

ja integraalin arvo = näiden *ala-* ja *yläintegraalien* yhteinen arvo $I(f,A,\mu)$.

Molemmista määrittelytavoista nähdään, että $\int_T f_0\,d\mu$ ei riipu suorakulmion T valinnasta sikäli kuin $T \supset A$ (Harj.teht. 1). — Tämä tulos ennakoitiin jo edellä, kun sovittiin: $\int_A f\,d\mu = \int_T f_0\,d\mu,\ T \supset A$.

Lineaarisuus. Additiivisuus. Vertailuperiaate

Jos merkitään rajoitetun joukon A yli integroituvien funktioiden joukkoa symbolilla $\mathcal{R}_A$, niin seuraavat yleiset ominaisuudet ovat yhteisiä tasointegraaleille ja yksiulotteisille määrätyille integraaleille (ks. Lauseet VIII.5.4–VIII.5.6). Nämä ovat voimassa myös jatkossa määriteltäville muille integraalityypeille.

Lineaarisuus : $f,g \in \mathcal{R}_A \;\Rightarrow\; \alpha f + \beta g \in \mathcal{R}_A \quad \forall\, \alpha,\beta \in \mathbb{R}$ ja

$$\int_A (\alpha f + \beta g)\,d\mu = \alpha \int_A f\,d\mu + \beta \int_A g\,d\mu.$$

Additiivisuus : $f \in \mathcal{R}_A \;\&\; f \in \mathcal{R}_B \;\&\; A \cap B = \emptyset$

$$\Rightarrow \int_{A\cup B} f\,d\mu = \int_A f\,d\mu + \int_B f\,d\mu.$$

Vertailuperiaate : $f,g \in \mathcal{R}_A \;\&\; f \le g$ A:ssa $\Rightarrow \displaystyle\int_A f\,d\mu \le \int_A g\,d\mu.$

Lineaarisuus ja vertailuperiaate seuraavat suoraan määritelmästä, kuten yksiulotteisessa tapauksessa (ks. Luku VIII.5). Additiivisuus on johdettavissa yksinkertaisimmin lineaarisuudesta: Kun merkitään f:n nollajatkoja joukkojen A, B ja $A \cup B$ suhteen vastaavasti symboleilla f_A, f_B ja f_0, niin pätee

$$A \cap B = \emptyset \quad \Rightarrow \quad f_0 = f_A + f_B.$$

Näin ollen jos $T = [a_1, b_1] \times [a_2, b_2] \supset A \cup B$, niin lineaarisuuden nojalla

$$\int_T f_0 \, d\mu = \int_T f_A \, d\mu + \int_T f_B \, d\mu \quad \Leftrightarrow \quad \int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu.$$

Rajoitetun joukon mitta

Em. määritelmissä on integraalin riippuvuus mitasta varsin yksinkertainen, sillä määritelmissä tarvitaan vain suorakulmioiden T_{kl} mittoja aksiooman A1 mukaisesti. Mutta kun integraali on kerran määritelty käyttäen hyväksi mitan yksinkertaisimpia ominaisuuksia, voidaankin määritellä yleisemmän joukon A mitta käyttäen hyväksi integraalia (!). Menettely on seuraava: Tarkastellaan funktiota

$$f(x, y) = 1, \quad (x, y) \in A.$$

Tämän nollajatkoa sanotaan A:n *karakteristiseksi funktioksi* ja merkitään

$$\chi_A(x, y) = \begin{cases} 1, & \text{kun } (x, y) \in A, \\ 0, & \text{kun } (x, y) \notin A. \end{cases}$$

Sikäli kuin χ_A on Riemann-integroituva yli suorakulmion $T \supset A$, on em. määritelmien mukaan

$$\int_A d\mu = \int_T \chi_A \, d\mu, \quad T = [a_1, b_1] \times [a_2, b_2] \supset A.$$

MÄÄRITELMÄ XIII.1.1 (**Jordan-mitta**) Rajoitettu joukko $A \subset \mathbb{R}^2$ on *Jordan-mitallinen* täsmälleen kun A:n karakteristinen funktio χ_A on Riemann-integroituva yli suorakulmioiden $T = [a_1, b_1] \times [a_2, b_2] \supset A$ ja A:n *Jordan-mitta* on tällöin

$$\boxed{\mu(A) = \int_A d\mu.}$$

Mitta-aksioomien A1–A3 toteutuminen on tämän määritelmän ja em. yleisten integraalin ominaisuuksien perusteella ilmeistä.

Kaikki rajoitetut joukot eivät ole mitallisia (ks. Harj.teht. 2a), mutta jokaiselle rajoitetulle joukolle $A \subset \mathbb{R}^2$ voidaan määritellä

$$\begin{aligned} &\textit{ulkomitta}: \quad \overline{\mu}(A) = \inf_{\mathcal{T}_h} \overline{\sigma}_h(\chi_A, \mathcal{T}_h), \\ &\textit{sisämitta}: \quad \underline{\mu}(A) = \sup_{\mathcal{T}_h} \underline{\sigma}_h(\chi_A, \mathcal{T}_h), \end{aligned}$$

missä inf ja sup lasketaan yli kaikkien perussuorakulmion $T \supset A$ jakojen. Joukko A on mitallinen täsmälleen kun $\overline{\mu}(A) = \underline{\mu}(A)$.

Kuviossa on

$$\begin{aligned} \mu(\text{▒}) &= \underline{\sigma}_h(\chi_A, \mathcal{T}_h) = \sum_{T_{kl} \in \mathcal{T}_h : T_{kl} \subset A} \mu(T_{kl}), \\ \mu(\text{///}) &= \overline{\sigma}_h(\chi_A, \mathcal{T}_h) - \underline{\sigma}_h(\chi_A, \mathcal{T}_h). \end{aligned}$$

Esimerkki 3 $A =$ jana, jonka päätepisteet ovat $(0,0)$ ja $(2,1)$. Näytä: $\mu(A) = 0$.

<u>Ratkaisu</u> Jaetaan $T = [0,2] \times [0,1] \supset A$ suorakulmioihin kokoa $h \times h/2$, missä $2/h = n \in \mathbb{N}$. Tällöin on (vrt. kuvio)

$$\overline{\sigma}_h(\chi_A, \mathcal{T}_h) = \frac{2}{h} \cdot \frac{h^2}{2} = h.$$

Näin ollen $\overline{\mu}(A) \leq h = 2/n \ \forall n \in \mathbb{N}$, joten $\overline{\mu}(A) = 0$ ja siis $\mu(A) = 0$. $\square$

Mainittakoon vielä seuraava lause, jonka osittainen todistus jätetään harjoitustehtäväksi (Harj.teht. 11). Lauseen jatkuvuusoletus (jota harjoitustehtävässä on vahvistettu) viittaa Määritelmään XI.1.10.

*Lause XIII.1.2 Jos f on jatkuva suorakulmiossa $T = [a, b] \times [c, d]$ ja $A \subset T$ on mitallinen, niin f on Riemann-integroituva yli A:n.

Siirto-, peilaus- ja kiertoinvarianssi

Jotta Jordan-mitta vastaisi tavanomaista geometriasta tunnettua pinta-alaa, on mitan oltava *siirto-*, *peilaus* ja *kiertoinvariantti.* Tällä tarkoitetaan, että jos A on mitallinen ja A' on A:n kanssa (geometrisesti) *yhtenevä*, eli $A' = \mathbf{f}(A)$, missä joillakin $a, b, \theta \in \mathbb{R}$ on

$$\mathbf{f}(x, y) = (a + x\cos\theta \mp y\sin\theta,\ b + x\sin\theta \pm y\cos\theta),$$

niin $\mu(A') = \mu(A)$. Kuvaus $\mathbf{f} : A \to A'$ on affiinimuunnoksen erikoistapaus, joka koostuu *siirrosta* (siirtovektori $\vec{r}_0 = a\vec{i} + b\vec{j}$), *kierrosta* (kiertokulma θ) ja mahdollisesta *peilauksesta* (merkin vaihtelu). Yhdistettyä siirtoa ja kiertoa (kuva) sanotaan *euklidiseksi liikkeeksi.*

Siirto- ja peilausinvarianssi ovat mitan μ määritelmästä varsin ilmeisiä. Kierron suhteen sen sijaan voisi määritelmän koordinaattiriippuvuuden epäillä aiheuttavan ongelmia. Tällaisia ei todellisuudessa ole, vaan pätee (ks. Harj.teht. 6,13)

Lause XIII.1.3 Jordanin pinta-alamitta on siirto-, peilaus- ja kiertoinvariantti, ts. geometrisesti yhtenevien joukkojen mitat ovat samat.

Keskipistesääntö

Kuten yhden muuttujan integraaleja, myös tasointegraaleja voi laskea numeerisesti suoraan määritelmästä käsin, eli approksimoimalla integraalia Riemannin summalla:

$$\int_A f\,d\mu = \int_T f_0\,d\mu \approx \sum_{k=1}^{m}\sum_{l=1}^{n} f_0(\xi_{kl},\eta_{kl})\mu(T_{kl}), \quad T \supset A.$$

Jos tässä valitaan $(\xi_{kl},\eta_{kl}) = T_{kl}$:n keskipiste, niin approksimaatiota sanotaan *yhdistetyksi keskipistesäännöksi*. Kuten yhdessä dimensiossa (vrt. Luku VIII.9), tämä on Riemannin summiin perustuvista approksimaatioista yleensä tarkin.

Esimerkki 4 Jos $A = [a,b]\times[c,d]$, niin integraali $\int_A x\,d\mu$ saadaan lasketuksi jakamalla A $m\times n$ samankokoiseen suorakulmioon ja käyttämällä yhdistettyä keskipistesääntöä:

$$\begin{aligned}\int_A x\,d\mu &\approx \sum_{k=1}^{m}\sum_{l=1}^{n}\left[a+\frac{(k-\frac{1}{2})(b-a)}{m}\right]\frac{b-a}{m}\cdot\frac{c-d}{n}\\ &= (b-a)(c-d)\left[a+\frac{b-a}{m^2}\sum_{k=1}^{m}(k-\tfrac{1}{2})\right]\\ &= (b-a)(c-d)[a+\tfrac{1}{2}(b-a)] = \tfrac{1}{2}(a+b)\mu(A).\end{aligned}$$

Koska tulos ei riipu parametreista m,n, on oltava $\int_A x\,d\mu = \frac{1}{2}(a+b)\mu(A)$. Vastaavalla tavalla laskien saadaan $\int_A y\,d\mu = \frac{1}{2}(c+d)\mu(A)$. $\square$

Esimerkissä saatiin jo yhdellä suorakulmiolla ($m=n=1$, $T_{11}=A$) tarkka tulos. Yleisemmin keskipistesääntö integroi suorakulmiossa tarkasti ensimmäisen asteen polynomin:

$$f(x,y) = a_0 + a_1x + a_2y \quad (a_0,a_1,a_2\in\mathbb{R})$$
$$\Rightarrow \int_T f\,d\mu = f(x_0,y_0)\mu(T).$$

Tämän ominaisuuden ja kahden muuttujan Taylorin kaavan (Luku XI.9) perusteella keskipistesäännölle on johdettavissa virhearvio samaan tapaan kuin yhdessä dimensiossa (Harj.teht. 15).

Esimerkki 5 Olkoon $A = [0,1]\times[0,1]$ ja laskettavana integraali $\int_A x^2y^2\,d\mu$. Kun A jaetaan neliöihin kokoa $h\times h$ ja kussakin neliössä integroimispisteeksi valitaan (a) vasen alanurkka, (b) keskipiste, niin integraalille saadaan seuraavat likiarvot (tarkka arvo $= 1/9$; taulukkoon on merkitty myös summattavien termien lukumäärä $N = h^{-2}$):

h	N	Likiarvo (a)	Likiarvo (b)
0.1	10^2	0.081225000..	0.110556250..
0.01	10^4	0.107813722..	0.111105555..
0.001	10^6	0.110778138..	0.111111055..
0.0001	10^8	0.111077781..	0.111111110..

□

Jos joukko A on yleisempi joukko kuin suorakulmio, niin Riemannin summiin perustuva numeerinen approksimaatio

$$\int_A f\,d\mu = \int_T f_0\,d\mu \approx \sum_{k=1}^{m}\sum_{l=1}^{n} f_0(\xi_{kl},\eta_{kl})\mu(T_{kl})$$

ei välttämättä ole kovin hyvä, syystä että f:n nollajatko f_0 on (yleensä) epäjatkuva sellaisissa suorakulmioissa T_{kl}, jotka leikkaavat reunaviivaa ∂A (kuvio). Tällöin arvio

$$\int_{T_{kl}} f_0\,d\mu \approx f_0(\xi_{kl},\eta_{kl})\mu(T_{kl})$$

on pääsääntöisesti kehno, valittiinpa $(\xi_{kl},\eta_{kl}) \in T_{kl}$ miten hyvänsä.

Jos halutaan parantaa algoritmia, niin on syytä käyttää tarkempia approksimaatioita integraaleille

$$\int_{T_{kl}} f_0\,d\mu = \int_{T_{kl}\cap A} f\,d\mu$$

silloin kun $T_{kl}\cap\partial A \neq \emptyset$. Esimerkiksi jakoa voidaan paikallisesti tihentää ja käyttää samaa algoritmia uudelleen (T:n tilalla T_{kl}, A:n tilalla $A\cap T_{kl}$).

$\mathbb{R}$:n pituusmitta

Määriteltäessä $\mathbb{R}$:n (Jordanin) *pituusmitta* kirjoitetaan aksiooman A1 tilalle

A1. *Suljetun välin mitta*: $A = [a, b]: \ \mu(A) = b - a.$

Muut aksioomat säilyvät ennallaan. Jos $A \subset \mathbb{R}$ on rajoitettu joukko, niin määritellään

$$\int_A f\, d\mu = \int_T f_0\, d\mu, \quad T = [a, b] \supset A,$$

missä f_0 on f:n nollajatko, ja tässä edelleen

$$\int_T f_0\, d\mu = \operatorname*{Lim}_{h \to 0} \sum_{k=1}^{n} f_0(\xi_k)\mu(T_k),$$

missä $T_k = [x_{k-1}, x_k]$, $\xi_k \in T_k$ ja $a = x_0 < x_1 < \ldots < x_n = b$. Koska $\mu(T_k) = x_k - x_{k-1}$, niin tässä $\int_T f_0\, d\mu = \int_a^b f_0(x)\, dx$ (Määritelmä VIII.5.3). Määrätyssä (Riemannin) integraalissa on siis kyse integroimisesta $\mathbb{R}$:n pituusmitan suhteen. Rajoitetun joukon $A \subset \mathbb{R}$ Jordan-mitta määritellään kuten edellä, eli integraalin avulla:

$$\mu(A) = \int_A d\mu.$$

Esimerkki 6 Jos $A = [0, 1] \cup [2, 4]$ niin integraalin additiivisuuden (tai suoraan määritelmän) perusteella

$$\int_A f\, d\mu = \int_{[0,1]} f\, d\mu + \int_{[2,4]} f\, d\mu = \int_0^1 f(x)\, dx + \int_2^4 f(x)\, dx.$$

Erityisesti on $\mu(A) = \int_A d\mu = 1 + 2 = 3$. $\square$

HARJOITUSTEHTÄVIÄ

1. Näytä, että jos T_1 ja T_2 ovat joukon $A \subset \mathbb{R}^2$ sisältäviä perussuorakulmioita, niin samoin on $T_1 \cap T_2$. Päättele tästä, että määritelmä $\int_A f\, d\mu = \int_T f_0\, d\mu$ antaa saman tuloksen perussuorakulmion $T \supset A$ valinnasta riippumatta.

2. a) Näytä, että joukko $A = ([0, 1] \cap \mathbb{Q}) \times [0, 1]$ ei ole Jordan-mitallinen.
   b) Anna esimerkki joukosta A, jolle pätee $A \subset [0, 1] \times [0, 1]$, $\overline{\mu}(A) = 1$ ja $\underline{\mu}(A) = 0.99$.

3. Todista seuraavat rajoitettuja joukkoja $A, B \subset \mathbb{R}^2$ koskevat väittämät:
   a) Jos $A \subset B$, niin $\underline{\mu}(A) \leq \underline{\mu}(B)$ ja $\overline{\mu}(A) \leq \overline{\mu}(B)$.
   b) Jos $\mu(B) = 0$, niin $\underline{\mu}(A \cup B) = \underline{\mu}(A)$ ja $\overline{\mu}(A \cup B) = \overline{\mu}(A)$.
   c) Jos $\mu(A \cap B) = 0$, niin $\underline{\mu}(A \cup B) = \underline{\mu}(A) + \underline{\mu}(B)$ ja $\overline{\mu}(A \cup B) = \overline{\mu}(A) + \overline{\mu}(B)$.

4. Laske joukon $A = \{(x, y) \in \mathbb{R}^2 \mid x \in [0, 1] \ \& \ 0 \leq y \leq x^2\}$ Jordanin ulko- ja sisämitalle approksimaatiot jakamalla $T = [0, 1] \times [0, 1]$ suorakulmioihin kokoa $h \times h^2$ ($h^{-1} \in \mathbb{N}$) ja näiden perusteella mitta $\mu(A)$.

5. Olkoon $A \subset \mathbb{R}^2$ rajoitettu joukko ja f ja g määriteltyjä ja rajoitettuja A:ssa. Näytä, että jos $f(x, y) \leq g(x, y)$ $\forall (x, y) \in A$, niin $\overline{I}(f, \mu, A) \leq \underline{I}(g, \mu, A)$.

6. a) Olkoon A = suorakulmainen kolmio, jonka kärjet ovat pisteissä $(0, 0)$, $(a, 0)$ ja $(0, b)$. Näytä Jordan-mitan määritelmästä, että $\mu(A) = \frac{1}{2}|a||b|$.
   b) Vedoten a-kohdan tulokseen ja mitan additiivisuuteen päättele, että suorakulmion mitta on kiertoinvariantti.

7. Laske yhdistettyä keskipistesääntöä käyttäen integraali $\int_A (1 + x - 2y)\, d\mu$, kun $A \subset \mathbb{R}^2$ on T:n muotoinen alue, jonka nurkkapisteet ovat $(2, 0)$, $(3, 0)$, $(3, 3)$, $(5, 3)$, $(5, 4)$, $(0, 4)$, $(0, 3)$ ja $(2, 3)$.

8. Olkoon $A = [-a/2, a/2] \times [-b/2, b/2]$. Jakamalla A samankokoisiin suorakulmioihin ja käyttämällä yhdistettyä keskipistesääntöä näytä oikeaksi:
$$\int_A x^2\, d\mu = \frac{1}{12}\, a^3 b, \quad \int_A y^2\, d\mu = \frac{1}{12}\, ab^3, \quad \int_A |xy|\, d\mu = \frac{1}{16}\, a^2 b^2.$$

9. Olkoon $A = [0, 1] \times [0, 1]$. Laske seuraavat integraalit jakamalla A neliöihin kokoa $h \times h$ ($h^{-1} \in \mathbb{N}$), käyttämällä yhdistettyä keskipistesääntöä, ja laskemalla tuloksen raja-arvo, kun $h \to 0$.
$$\text{a)} \int_A xy^2\, d\mu \qquad \text{b)} \int_A x^2 y^2\, d\mu \qquad \text{c)} \int_A e^{x+y}\, d\mu$$

10. (*) Luvussa VIII.8 väitettiin, että jos
$$A = \{\, (x, y) \in \mathbb{R}^2 \mid x \in [a, b] \ \& \ 0 \leq y \leq f(x) \,\},$$
missä f on rajoitettu, ei-negatiivinen ja Riemann-integroituva välillä $[a, b]$, niin $\mu(A) = \int_a^b f(x)\, dx$. Todista väittämä uudelleen näyttämällä, että
$$\underline{\mu}(A) = \underline{\int_a^b} f(x)\, dx, \quad \overline{\mu}(A) = \overline{\int_a^b} f(x)\, dx$$
ja vetoamalla Lauseeseen VIII.5.8.

11. (*) Todista Lauseen XIII.1.2 väittämä, kun jatkuvuuden sijasta oletetaan, että f toteuttaa Lipschitz-jatkuvuusehdon

$$|f(x_1, y_1) - f(x_2, y_2)| \leq L(|x_1 - x_2| + |y_1 - y_2|), \quad (x_1, y_1), (x_2, y_2) \in T.$$

Vihje: Vertaa ylä- ja alasummia $\overline{\sigma}(f_0, \mathcal{T}_h)$ ja $\underline{\sigma}(f_0, \mathcal{T}_h)$ ($\mathcal{T}_h = T$:n ositus).

12. (*) Olkoon $A \subset \mathbb{R}^2$ rajoitettu joukko ja f määritelty A:ssa. Määritellään

$$f_+(x, y) = \begin{cases} f(x, y), & \text{kun } f(x, y) > 0, \\ 0, & \text{muulloin.} \end{cases}$$

a) Näytä: $\overline{I}(f_+, A) - \underline{I}(f_+, A) \leq \overline{I}(f, A) - \underline{I}(f, A)$.
b) Päättele: Jos f on Riemann-integroituva yli A:n, niin samoin ovat f_+, $f_- = f - f_+$ ja $|f|$.

13. (*) Lähtien tehtävän 6 tuloksesta todista: Jos $A \subset \mathbb{R}^2$ on rajoitettu, niin $\underline{\mu}(A)$ ja $\overline{\mu}(A)$ ovat kiertoinvariantteja. Päättele, että jos A on Jordan-mitallinen, niin $\mu(A)$ on kiertoinvariantti.

14. (*) Integraalille $\int_A f \, d\mu$, missä A on kolmio, jonka kärjet ovat $(0, 0)$, $(1, 0)$ ja $(0, 1)$, ja $f(x, y) = x + 2y$, lasketaan likiarvo jakamalla $T = [0, 1] \times [0, 1] \supset A$ neliöihin kokoa $h \times h$ ja laskemalla $\int_T f_0 \, d\mu$ yhdistetyllä keskipistesäännöllä.
a) Näytä, että tuloksena on ylälikiarvo, jonka virhe $= \frac{3}{2}h + \mathcal{O}(h^2)$.
b) Mikä on integraalin tarkka arvo?

15. (*) Olkoon f:n osittaisderivaatat toiseen kertalukuun asti jatkuvia suorakulmiossa $A = [-a/2, a/2] \times [-b/2, b/2]$. Käyttämällä Taylorin lausetta, integraalien vertailuperiaatetta ja tehtävän 8 tuloksia näytä oikeaksi keskipistesäännön virhearvio

$$\left| \int_A f \, d\mu - f(0, 0) \, ab \right| \leq \frac{ab}{24}(M_{11}a^2 + M_{22}b^2) + \frac{1}{16} M_{12} \, a^2 b^2,$$

missä M_{11}, M_{22} ja M_{12} ovat osittaisderivaattojen f_{xx}, f_{yy} ja f_{xy} itseisarvojen maksimiarvot A:ssa. Näytä edelleen, että jos A jaetaan suorakulmioihin kokoa $h_1 \times h_2$ ($a/h_1 \in \mathbb{N}$, $b/h_2 \in \mathbb{N}$), niin yhdistetylle keskipistesäännölle pätee virhearvio

$$|E(f)| \leq \frac{ab}{24} \left(M_{11}h_1^2 + M_{22}h_2^2 + \frac{3}{2} M_{12} \, h_1 h_2 \right).$$

Mikä $E(f)$ on tarkasti, jos f on toisen asteen polynomi ja tunnetaan osittaisderivaattojen f_{xx}, f_{yy} ja f_{xy} (vakio)arvot F_{11}, F_{22} ja F_{12}?

XIII.2 Tasointegraalien laskutekniikka

Tasointegraali on mahdollistaa palauttaa yhden muuttujan integrointitehtäviksi, jolloin integraali voidaan suotuisissa oloissa laskea käyttäen hyväksi integraalifunktioita, eli suljetussa muodossa. Tässä luvussa tarkastellaan tällaista laskutekniikkaa.

Tarkastellaan joukkoa $A \subset \mathbb{R}^2$ joka on nk. *x-projisoituva*, eli muotoa

$$A = \{(x, y) \in \mathbb{R}^2 \mid x \in [a, b],\ g_1(x) \leq y \leq g_2(x)\},$$

missä oletetaan, että g_1 ja g_2 ovat välillä $[a, b]$ rajoitettuja. Olkoon $T = [a, b] \times [c, d]$ suorakulmio, joka sisältää A:n, ja olkoon $\mathcal{T}_h$ jokin T:n jako.

Olkoon f määritelty T:ssä ja rajoitettu, ja tarkastellaan integraalin $\int_A f\, d\mu$ määrittävää Riemannin summaa. Olkoon $\mathcal{T}_h = \{T_{kl},\ k = 1 \ldots m,\ l = 1 \ldots n\}$ ja valitaan

$$(\xi_{kl}, \eta_{kl}) = (\xi_k, \eta_l) \in T_{kl}, \quad k = 1 \ldots m,\ l = 1 \ldots n.$$

Tällöin summa muodon

$$\sum_{T_{kl} \in \mathcal{T}_h} f_0(\xi_{kl}, \eta_{kl}) \mu(T_{kl}) = \sum_{k=1}^{m} \sum_{l=1}^{n} f_0(\xi_k, \eta_l)(x_k - x_{k-1})(y_l - y_{l-1}).$$

Sikäli kuin funktio $y \mapsto f_0(\xi_k, y)$ on Riemann-integroituva välillä $[c, d]$, on sisemmällä summalla raja-arvo

$$\sum_{l=1}^{n} f_0(\xi_k, \eta_l)(y_l - y_{l-1}) \underset{h \to 0}{\longrightarrow} \int_c^d f_0(\xi_k, y)\, dy = \int_{g_1(\xi_k)}^{g_2(\xi_k)} f(\xi_k, y)\, dy.$$

Kun merkitään

$$F(x) = \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy, \quad x \in [a,b],$$

niin päätellään, että sikäli kuin F on edelleen Riemann-integroituva välillä $[a,b]$, niin jaon $\mathcal{T}_h$ ollessa riittävän tiheä on likimäärin

$$\sum_{T_{kl}\in\mathcal{T}_h} f_0(\xi_k,\eta_l)\mu(T_{kl}) \approx \sum_{l=1}^{n} F(\xi_k)(x_k - x_{k-1}) \approx \int_a^b F(x)\,dx.$$

Näyttää siis ilmeiseltä, että funktioiden g_1, g_2 ja f ollessa riittävän säännöllisiä pätee

$$h \to 0 \;\Rightarrow\; \sum_{T_{kl}\in\mathcal{T}_h} f_0(\xi_k,\eta_l)\mu(T_{kl}) \to \int_a^b F(x)\,dx.$$

Näin on johdettu (tai ainakin tehty uskottavaksi) laskukaava

$$\boxed{\begin{aligned} &A = \{(x,y)\in\mathbb{R}^2 \mid x\in[a,b],\ g_1(x)\le y\le g_2(x)\} \\ &\Rightarrow \quad \int_A f\,d\mu = \int_a^b \left[\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\right] dx. \end{aligned}} \tag{1}$$

Tämän mukaan siis tasointegraali yli x-projisoituvan joukon on laskettavissa peräkkäisinä yksiulotteisina, nk. *iteroituina integraaleina*. Kyseessä on *Fubinin lauseen* nimellä tunnetun väittämän eräs muoto. Lause muotoillaan täsmällisemmin ja todistetaan luvun lopussa.

Jos $A \subset \mathbb{R}^2$ on *y-projisoituva*, ts. (vrt. kuvio vasemmalla)

$$A = \{(x,y)\in\mathbb{R}^2 \mid y\in[a,b],\ g_1(y)\le x\le g_2(y)\},$$

niin integrointikaavassa (1) on vaihdettava x ja y, jolloin kaava saa muodon

$$\int_A f\,d\mu = \int_a^b \left[\int_{g_1(y)}^{g_2(y)} f(x,y)\,dx \right] dy. \tag{2}$$

Jos A ei ole kumpaakaan projisoituvaa tyyppiä, niin kaavoja (1) ja (2) päästään yleensä käyttämään, kun A ensin jaetaan sopiviin osiin (vrt. kuvio edellä) ja käytetään integraalin additiivisuutta:

$$\int_A f\,d\mu = \sum_{i=1}^m \int_{A_i} f\,d\mu, \quad A = \left(\bigcup_{i=1}^m A_i \right) \cup B, \quad \mu(A_i \cap A_j) = \mu(B) = 0,\ i \neq j.$$

Iteraatiokaavojen (1)–(2) perusteella voidaan tasointegraalikin laskea suotuisissa oloissa integraalifunktioiden avulla. Olkoon esimerkiksi A x-projisoituva ja oletetaan, että on löydettävissä funktion $f(x,y)$ integraalifunktio muuttujan y suhteen, ts. funktio $F(x,y)$, jolle pätee

$$\frac{\partial}{\partial y} F(x,y) = f(x,y), \quad (x,y) \in A.$$

Tällöin kaavan (1) mukaan

$$\int_A f(x,y)\,d\mu = \int_a^b \left[\Big|_{y=g_1(x)}^{y=g_2(x)} F(x,y) \right] dx = \int_a^b G(x)\,dx,$$

missä

$$G(x) = F(x, g_2(x)) - F(x, g_1(x)).$$

Jos edelleen on löydettävissä funktion G integraalifunktio H, niin integraali saadaan lasketuksi suljetussa muodossa:

$$\int f\,d\mu = \Big|_a^b H(x) = H(b) - H(a).$$

Laskun onnistuminen tällä tavoin riippuu siis paitsi funktiosta f, myös funktioista g_1 ja g_2, eli A:n reunaviivan muodosta. Onnistuminen on taattu esimerkiksi silloin, kun f, g_1 ja g_2 ovat polynomeja, sillä tällöin myös F, G ja H ovat polynomeja.

Ajatellen iteraatiokaavoja (1)–(2) käytetään tasointegraaleille merkintätapoja

$$\int_A f\,dxdy \quad \text{tai} \quad \iint_A f\,dxdy.$$

Näissä siis kirjoitetaan $d\mu = dxdy$, mikä merkintä viittaa 'differentiaalisen suorakulmion' $[x, x+dx] \times [y, y+dy]$ pinta-alaan. Merkintää käytetään jatkossa.

Esimerkki 1 Olkoon $A = [0,1] \times [0,1]$. Laske $\int_A \sqrt{|x-y|}\,dxdy$.

Ratkaisu Kaavan (1) mukaan

$$\begin{aligned}
\int_A \sqrt{|x-y|}\,dxdy &= \int_0^1 \left[\int_0^1 \sqrt{|x-y|}\,dy\right] dx \\
&= \int_0^1 \left[\int_0^x \sqrt{x-y}\,dy + \int_x^1 \sqrt{y-x}\,dy\right] dx \\
&= \int_0^1 \left[-\frac{2}{3}\Big|_{y=0}^{y=x} (x-y)^{3/2} + \frac{2}{3}\Big|_{y=x}^{y=1} (y-x)^{3/2}\right] dx \\
&= \int_0^1 \frac{2}{3}\left[x^{3/2} + (1-x)^{3/2}\right] dx \\
&= \Big|_0^1 \frac{4}{15}\left[x^{5/2} - (1-x)^{5/2}\right] = \underline{\underline{\frac{8}{15}}}. \quad \square
\end{aligned}$$

Esimerkki 2 $A =$ kolmio, jonka kärjet ovat $(0,0)$, $(1,0)$ ja $(0,1)$. Laske

$$I = \int_A xy^2\,dxdy.$$

Ratkaisu

$$\begin{aligned}
I &= \int_0^1 \left[\int_0^{1-x} xy^2\,dy\right] dx \\
&= \int_0^1 \left[\Big|_{y=0}^{y=1-x} \frac{1}{3}xy^3\right] dx \\
&= \int_0^1 \frac{1}{3}x(1-x)^3\,dx \\
&= \Big|_0^1 -\frac{1}{12}x(1-x)^4 + \int_0^1 \frac{1}{12}(1-x)^4\,dx \qquad \text{(osittaisintegrointi)} \\
&= \int_0^1 \frac{1}{12}(1-x)^4\,dx = \Big|_0^1 -\frac{1}{60}(1-x)^5 = \underline{\underline{\frac{1}{60}}}. \quad \square
\end{aligned}$$

Esimerkki 3 Laske $\mu(A)$, kun $A = R$-säteinen kiekko.

Ratkaisu Kaavan (1) mukaan

$$\mu(A) = \int_A d\mu = \int_{-R}^{R} \left[\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \right] dx$$
$$= \int_{-R}^{R} \left[\Big|_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y \right] dx = 2\int_{-R}^{R} \sqrt{R^2-x^2}\, dx.$$

Sijoitus $x = R\sin t$, $dx = R\cos t\, dt$, $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ antaa

$$2\int_{-R}^{R} \sqrt{R^2-x^2}\, dx = 2R^2 \int_{-\pi/2}^{\pi/2} \cos^2 t\, dt$$
$$= R^2 \Big|_{-\pi/2}^{\pi/2} (t + \sin t \cos t) = \pi R^2. \quad \square$$

Esimerkin 3 yleistyksenä nähdään kaavasta (1), että x-projisoituvan joukon pinta-alamitta on laskettavissa kaavalla

$$\mu(A) = \int_a^b [g_2(x) - g_1(x)]\, dx,$$

joka on tuttu jo Luvusta VIII.8. Kaavan käyttö vastaa pinta-alan laskemista 'siivuttamalla' A osiin muotoa $\Delta A = \{(x,y) \in A \mid x \in [x, x+\Delta x]\}$, jolloin on $\mu(\Delta A) \approx [g_2(x) - g_1(x)]\Delta x$ pienillä Δx:n arvoilla ja olettaen, että g_1 ja g_2 ovat jatkuvia pisteessä x. (Jatkuvuusoletus on lievennettävissä, vrt. Luku VIII.8.)

Epäoleelliset tasointegraalit

Jos tasointegraalissa $\int_A f\,dxdy$ joko A ei ole rajoitettu tai f ei ole A:ssa rajoitettu, voidaan tasointegraalin määrittelyä laajentaa vastaavaan tapaan kuin yksiulotteisen Riemannin integraalin tapauksessa, vrt. Luku VIII.7. Jos määritelmän laajennus tuottaa integraalille yksikäsitteisen (reaali)arvon, sanotaan jälleen, että näin määritelty *epäoleellinen* integraali *suppenee*. Valaistakoon asiaa ensin esimerkeillä.

Esimerkki 4 Olkoon $A = [0,1] \times [0,1]$ ja $B = [0,\infty) \times [0,\infty)$. Laske

$$\text{a)} \int_A \frac{y^2}{\sqrt{x}}\,dxdy, \qquad \text{b)} \int_B \frac{x}{1+x^2+y^2}\,dxdy.$$

<u>Ratkaisu</u> a) Koska f ei ole A:ssa rajoitettu (ei edes määritelty, kun $x = 0$), niin kyseessä on epäoleellinen integraali. Tämä voidaan (ainakin yrittää) laskea raja-arvona

$$\int_A f\,dxdy = \lim_{\varepsilon\to 0^+} \int_{A_\varepsilon} f\,dxdy,$$

missä $A_\varepsilon = [\varepsilon,1]\times[0,1]$. Tässä $\int_{A_\varepsilon} f\,dxdy$ on tavanomainen Riemannin integraali, joten iteraatiokaava (1) soveltuu:

$$\begin{aligned}\int_{A_\varepsilon} f\,dxdy &= \int_\varepsilon^1 \left[\int_0^1 \frac{y^2}{\sqrt{x}}\,dy\right] dx = \int_\varepsilon^1 \frac{1}{\sqrt{x}}\left[\int_0^1 y^2\,dy\right] dx \\ &= \int_\varepsilon^1 \frac{1}{\sqrt{x}}\left[\Big|_{y=0}^{y=1} \frac{y^3}{3}\right] dx = \int_\varepsilon^1 \frac{1}{3\sqrt{x}}\,dx = \Big|_\varepsilon^1 \frac{2}{3}\sqrt{x} \to \underline{\underline{\frac{2}{3}}}, \quad \text{kun } \varepsilon \to 0^+.\end{aligned}$$

Integraali siis suppenee. Oikea integraalin arvo olisi saatu myös luottaen suoraan kaavaan (1) eli ohittamalla raja-arvoprosessi:

$$\int_A f\,dxdy = \int_0^1 \left[\int_0^1 \frac{y^2}{\sqrt{x}}\,dy\right] dx = \int_0^1 \frac{1}{\sqrt{x}}\,dx \cdot \int_0^1 y^2\,dy = 2\cdot\frac{1}{3} = \frac{2}{3}\,.$$

b) Myös tämä integraali suppenee, ja suoraan kaavaan (2) perustuva lasku

$$\begin{aligned}\int_A \frac{x}{(1+x^2+y^2)^2}\,dxdy &= \int_0^\infty \left[\int_0^\infty \frac{x}{(1+x^2+y^2)^2}\,dx\right] dy \\ &= \int_0^\infty \left[\Big|_{x=0}^{x=\infty} -\frac{1}{2(1+x^2+y^2)}\right] dy \\ &= \int_0^\infty \frac{1}{2(1+y^2)}\,dy = \Big|_{y=0}^{y=\infty} \frac{1}{2}\operatorname{Arctan} y = \frac{\pi}{4}\end{aligned}$$

antaa oikean tuloksen. Asian voi varmistaa laskemalla ensin

$$\int_{T_{ab}} \frac{x}{(1+x^2+y^2)^2}\,dxdy = \int_0^b \left[\int_0^a \frac{x}{(1+x^2+y^2)^2}\,dx\right] dy,$$

missä $T_{ab} = [0,a] \times [0,b]$, ja toteamalla, että tämän integraalin arvo $\to \pi/4$ aina kun $a \to \infty$ ja $b \to \infty$. $\square$

Esimerkin perusteella kaavat (1)–(2) näyttävät soveltuvan suoraan myös epäoleellisiin integraaleihin. Useimmiten näin onkin käytännössä, mutta kyse on vain pääsäännöstä, kuten seuraava esimerkki osoittaa.

Esimerkki 5 Laske säännöillä (1) ja (2) integraali

$$I = \int_A f(x,y)\,dxdy, \quad A = [0,\infty) \times [0,1], \quad f(x,y) = xy(2-xy)e^{-xy}.$$

<u>Ratkaisu</u>

$$\text{Sääntö (1)}: \quad I = \int_0^\infty \left[\Big|_{y=0}^{y=1} xy^2 e^{-xy}\right] dx = \int_0^\infty xe^{-x} = 1.$$

$$\text{Sääntö (2)}: \quad I = \int_0^1 \left[\lim_{a\to\infty} \Big|_{x=0}^{x=a} x^2 y e^{-xy}\right] dy = \int_0^1 \left[\lim_{a\to\infty} a^2 y e^{-ay}\right] dy = 0. \quad \square$$

Esimerkin ristiriidan voi pelkistää seuraavaan päättelyketjuun, jossa kysymysmerkillä varustettu päätelmä (ja vain se!) on väärä:

$$\begin{aligned}
1 &= \int_0^\infty \left[\int_0^1 f(x,y)\,dy\right] dx \\
&= \lim_{a\to\infty} \int_0^a \left[\int_0^1 f(x,y)\,dy\right] dx \\
&= \lim_{a\to\infty} \int_0^1 \left[\int_0^a f(x,y)\,dx\right] dy \\
&= \int_0^1 \lim_{a\to\infty} \left[\int_0^a f(x,y)\,dx\right] dy \qquad (?) \\
&= \int_0^1 \left[\int_0^\infty f(x,y)\,dx\right] dy = 0.
\end{aligned}$$

Esimerkissä ongelmaksi osoittautuu integroitavan funktion <u>merkinvaihtelu</u>: Jos tämä sallitaan, eivät iteraatiokaavat (1)–(2) välttämättä anna ristiriidattomia

tuloksia. Ongelma ratkaistaankin poistamalla merkinvaihtelun mahdollisuus epäoleellisen integraalin määritelmässä seuraavasti:

Määritellään funktion $f(x,y)$ *positiivinen osa* f_+ ja *negatiivinen osa* f_- asettamalla f:n määrittelyjoukossa D_f

$$f_+(x,y) = \begin{cases} f(x,y), & \text{kun } f(x,y) > 0, \\ 0, & \text{muulloin,} \end{cases}$$

$$f_-(x,y) = \begin{cases} -f(x,y), & \text{kun } f(x,y) < 0, \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin jos $A \subset D_f$ on rajoitettu joukko ja f on Riemann-integroituva yli A:n, niin samoin ovat f_+ ja f_- (Harj.teht. XIII.1:12). Koska on myös $f = f_+ - f_-$, niin seuraa

$$\int_A f\,d\mu = \int_A f_+\,d\mu - \int_A f_-\,d\mu.$$

Jos A ei ole rajoitettu tai f ei ole A:ssa rajoitettu, niin otetaan tämä integraalin <u>määritelmäksi</u>, eli asetetaan vaatimus, että f_+ ja f_- ovat molemmat integroituvia yli A:n. Funktion merkinvaihtelun mahdollisuus on näin poistettu integraalin laajennetusta määritelmästä.

Esimerkki 5 (jatko). Tässä on

$$f_+(x,y) = \begin{cases} xy(2-xy)e^{-xy}, & \text{kun } 0 \le xy \le 2, \\ 0, & \text{muulloin,} \end{cases}$$

joten kun $A = [0,\infty) \times [0,1]$, saadaan

$$\begin{aligned} \int_A f_+\,dxdy &= \int_0^1 \left[\int_0^\infty f_+(x,y)\,dx \right] dy \\ &= \int_0^1 \left[\int_0^{2/y} xy(2-xy)e^{-xy}\,dx \right] dy \\ &= \int_0^1 \left[\Big|_{x=0}^{x=2/y} x^2 y e^{-xy} \right] dy \;=\; \int_0^1 4e^{-2}y^{-1}\,dy. \end{aligned}$$

Saatu integraali ei suppene, joten f_+ ei ole integroituva yli A:n, eikä määritelmän mukaan siis myöskään f. $\square$

Todettakoon lopuksi, että em. määritelmänkään nojalla iteraatiokaavat (1)–(2) eivät sovellu epäoleellisiin integraaleihin aivan yleispätevästi, mutta poikkeukset näistä (pää)säännöistä ovat melko eksoottisia, ks. Harj.teht. 10.

Fubinin lause

Tarkoituksena on perustella tasointegraalin iteraatiokaava (1). Tässä riittää rajoittua erikoistapaukseen, jossa A on suorakulmio, eli laskukaavaan

$$\int_T f\,d\mu = \int_a^b \left[\int_c^d f(x,y)\,dy\right] dx, \quad T = [a,b]\times[c,d].$$

Nimittäin jos tässä on

$$T \supset A = \{(x,y)\in\mathbb{R}^2 \mid x\in[a,b] \ \ \& \ \ g_1(x)\le y\le g_2(x)\},$$

niin kyse on kaavasta (1), kun valitaan $f = f_0 = f$:n nollajatko A:n ulkopuolelle.

Lause XIII.2.1 **(Tasointegraalin iteraatiokaava – Fubinin† lause)** Olkoon f määritelty, rajoitettu ja Riemann-integroituva suorakulmiossa $T = [a,b]\times[c,d]$. Olkoon edelleen $f(x,y)$ muuttujan y suhteen Riemann-integroituva välillä $[c,d]$ jokaisella $x\in[a,b]$ ja muuttujan x suhteen Riemann-integroituva välillä $[a,b]$ jokaisella $y\in[c,d]$. Tällöin funktiot

$$F(x) = \int_c^d f(x,y)\,dy, \quad G(y) = \int_a^b f(x,y)\,dx$$

ovat Riemann-integroituvia väleillä $[a,b]$ ja $[c,d]$, ja pätee

$$\int_T f(x,y)\,dxdy = \int_a^b F(x)\,dx = \int_c^d G(y)\,dy.$$

<u>Todistus</u> Olkoon $\mathcal{T}_h = \{T_{kl},\ k=1\ldots m,\ l=1\ldots n\}$, $T_{kl} = [x_{k-1},x_k]\times[y_{l-1},y_l]$, suorakulmion T jako, ja olkoon

$$m_{kl} = \inf_{(x,y)\in T_{kl}} f(x,y), \quad M_{kl} = \sup_{(x,y)\in T_{kl}} f(x,y).$$

Tällöin $m_{kl}\le f(x,y)\le M_{kl}\ \forall (x,y)\in T_{kl}$, joten integraalien vertailuperiaatteen nojalla

$$m_{kl}(y_l - y_{l-1}) \le \int_{y_{l-1}}^{y_l} f(x,y)\,dy \le M_{kl}(y_l - y_{l-1}) \quad \forall x\in[x_{k-1},x_k].$$

Summaamalla tämä yli indeksin l seuraa

$$m_k \;=\; \sum_{l=1}^n m_{kl}(y_l - y_{l-1}) \le F(x) \le \sum_{l=1}^n M_{kl}(y_l - y_{l-1}) \;=\; M_k \quad \forall x\in[x_{k-1},x_k],$$

†Italialainen matemaatikko **Guido Fubini** eli vuosina 1897-1943.

mistä puolestaan seuraa

$$\begin{aligned}\sum_{k=1}^{m}\sum_{l=1}^{n} m_{kl}(x_k - x_{k-1})(y_l - y_{l-1}) &= \sum_{k=1}^{m} m_k(x_k - x_{k-1}) \\ &\le \underline{\int_a^b} F(x)\,dx \le \overline{\int_a^b} F(x)\,dx \\ &\le \sum_{k=1}^{m} M_k(x_k - x_{k-1}) \\ &= \sum_{k=1}^{m}\sum_{l=1}^{n} M_{kl}(x_k - x_{k-1})(y_l - y_{l-1}),\end{aligned}$$

eli

$$\underline{\sigma}(f, \mathcal{T}_h) \le \underline{\int_a^b} F(x)\,dx \le \overline{\int_a^b} F(x)\,dx \le \overline{\sigma}(f, \mathcal{T}_h),$$

missä $\underline{\sigma}(f, \mathcal{T}_h)$ ja $\overline{\sigma}(f, \mathcal{T}_h)$ ovat jakoon $\mathcal{T}_h$ liittyvät Riemannin ala- ja yläsummat. Ottamalla tässä vasemmalla supremum ja oikealla infimum kaikkien jakojen $\mathcal{T}_h$ suhteen seuraa tasointegraalin määritelmän nojalla (vrt. edellinen luku)

$$\int_T f\,d\mu \le \underline{\int_a^b} F(x)\,dx \le \overline{\int_a^b} F(x)\,dx \le \int_T f\,d\mu.$$

Tämän mukaan F:n ylä- ja alaintegraalit välillä $[a, b]$ ovat samat, joten F on välillä $[a, b]$ Riemann-integroituva (Lause VIII.5.8), ja epäyhtälöketjusta seuraa myös väitetty (ensimmäinen) laskukaava. Toinen laskukaava perustellaan vastaavasti.

Mainittakoon, että Fubinin lauseessa ei välttämättä tarvita oletuksia funktion f integroituvuudesta erikseen muuttujien x ja y suhteen, vaan riittää, että f on integroituva yli T:n. Tällöin on väittämässä funktiot F ja G määriteltävä joko ylä- tai alaintegraaleina (mikä tahansa neljästä vaihtoehdosta), muuten lauseen väittämä säilyttää pätevyytensä, ja todistuskin pysyy olennaisesti samana (!). Jos siis oletetaan ainoastaan, että f on Riemann-integroituva yli T:n, niin pätee esimerkiksi laskukaava

$$\int_T f\,dxdy = \int_a^b \left[\overline{\int_c^d} f(x, y)\,dy\right] dx.$$

Huomautettakoon lopuksi, että Lauseiden VIII.6.1 ja XIII.1.2 perusteella Fubinin lauseen ehdot toteutuvat oletetussa muodossa, jos f on jatkuva suorakulmiossa $T = [a, b] \times [c, d]$.

HARJOITUSTEHTÄVIÄ

1. Olkoon $f(x,y) = e^{x^4-2xy-y^2}$ ja $A = [-1,1] \times [0,2]$. Laske $\int_A f_{xy}(x,y)\,dxdy$.

2. Laske annetun funktion $f(x,y)$ integraali yli annetun joukon A:
   a) $xy + y^2$, $A:\ 0 \le x \le 1\ \ \&\ \ 0 \le y \le 1 - x$
   b) $(x+y)e^{x+y}$, $A:\ x, y \ge 0\ \ \&\ \ x + y \le 1$
   c) $\cos y$, $A:\ 0 \le x \le \pi\ \ \&\ \ |y| \le x$
   d) $y^2 e^{xy}$ $A:\ 0 \le x \le 2\ \ \&\ \ x \le y \le 2$
   e) xy^2, $A:\ 0 \le x \le 1\ \ \&\ \ x^2 \le y \le \sqrt{x}$
   f) $x \cos y$, $A:\ 0 \le x \le 1\ \ \&\ \ 0 \le y \le 1 - x^2$
   g) $\sqrt{a^2 - y^2}$, $A =$ kolmio, jonka kärjet $(0,0)$, $(0,a)$ ja (a,a) $(a > 0)$
   h) $xy(1+x^4)^{-1}$, $A =$ kolmio, jonka kärjet $(0,0)$, $(1,0)$ ja $(1,1)$

3. Määritä pinta-ala $\mu(A)$, kun A määritellään annetuilla ehdoilla $(a > 0)$.
   a) $ax \ge y^2$, $x^3 \le ay^2$ b) $ax \ge y^2$, $x^2 + y^2 \le a^2$
   c) $a^2 y \le x^3$, $y \ge 3x - 2a$ d) $x^2 - y^2 \le a^2$, $|y| \le x/2$
   e) $x, y \ge 0$, $ax \ge y^2$, $ay \ge x^2$, $8xy^2 \le a^3$

4. Määritä ja luokittele funktion
$$f(x,y) = \int_A (10 - xu - yv^2)^2\,dudv, \quad A = [0,1] \times [0,1]$$
   kriittiset pisteet.

5. Suppeneeko vai hajaantuuko integraali $\int_A (x-y)^{-1}\,dx$, kun
   a) $A = [0,1] \times [0,1]$, b) $A = [0,1] \times [-1,0]$?

6. Laske seuraavat integraalit kaikilla arvoila $\alpha \in \mathbb{R}$, joilla integraali suppenee.
   a) $\int_A x^\alpha y^\alpha\,dxdy$, $A = [1,\infty) \times [1,\infty)$
   b) $\int_A |x-2y|^\alpha\,dxdy$, $A = [1,2] \times [0,1]$
   c) $\int_A |x-2y|^\alpha\,dxdy$, $A = [1,2] \times [1,2]$
   d) $\int_A x^\alpha y e^{-xy}\,dxdy$, $A = [0,1] \times [0,\infty)$

7. Seuraavat integraalit ovat muotoa $\int_A f(x,y)\,dxdy$. Määritä ensin A ja laske sitten integraalin arvo valitsemalla sopiva integroimisjärjestys!

   a) $\displaystyle\int_0^1 \left[\int_x^1 e^{-y^2}\,dy\right] dx$ b) $\displaystyle\int_0^{\pi/2} \left[\int_y^{\pi/2} \frac{\sin x}{x}\,dx\right] dy$

   c) $\displaystyle\int_0^1 \left[\int_{x^2}^x xy^{-1}e^y\,dy\right] dx$ d) $\displaystyle\int_0^1 \left[\int_x^1 \frac{y^\alpha}{x^2+y^2}\,dy\right] dx$ $(\alpha > 0)$

8. Tiedetään, että $\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}$.† Laske tällä perusteella

a) $\displaystyle\int_{\mathbb{R}^2} |x| e^{-x^2-y^2}\,dxdy$ b) $\displaystyle\int_{\mathbb{R}^2} (x^2+y^2) e^{-x^2-y^2}\,dxdy$

c) $\displaystyle\int_{\mathbb{R}^2} |x+y| e^{-x^2-y^2}\,dxdy$ d) $\displaystyle\int_{\mathbb{R}^2} |x-y^2| e^{-|x|-y^2}\,dxdy$

9. (*) a) Näytä integraaliin vertaamalla, että 2-ulotteinen sarja

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{1}{(i+j)^\alpha}$$

suppenee täsmälleen kun $\alpha > 2$. b) Näytä, että approksimaation

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{1}{(i+j)^\alpha} \approx \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{(i+j)^\alpha} \quad (\alpha > 2,\ N \in \mathbb{N})$$

virhe on luokkaa $\mathcal{O}(N^{2-\alpha})$. c) Näytä, että tarkemmin pätee

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{1}{(i+j)^\alpha} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{(i+j)^\alpha} + \frac{2-2^{2-\alpha}}{(\alpha-2)(\alpha-1)} N^{2-\alpha} + \mathcal{O}(N^{1-\alpha}).$$

10. (*) (Riemann vastaan Fubini) Järjestetään rationaaliluvut välillä $[0,1]$ jonoksi $\{x_n,\ n = 0,1,2,\ldots\}$ ja määritellään joukossa $A = [0,1]\times[0,\infty)$ funktio f seuraavasti:

$$f(x,y) = \begin{cases} e^{n-y}, & \text{kun } x = x_n \text{ ja } y \geq n, \\ 0, & \text{muulloin.} \end{cases}$$

a) Näytä, että f on Riemann-integroituva yli A:n (laajennettu määritelmä).
b) Näytä, että laskukaavoista (1)–(2) jälkimmäinen toimii, edellinen ei, ts.

$$\int_A f\,dxdy = \int_0^\infty \left[\int_0^1 f(x,y)\,dx\right] dy \neq \int_0^1 \left[\int_0^\infty f(x,y)\,dy\right] dx.$$

†Ks. Propositio XIII.4.1 jäljempänä.

XIII.3 Avaruusintegraalit

Integraalia muotoa

$$\int_A f\,d\mu, \quad A \subset \mathbb{R}^3, \quad f = f(x,y,z),$$

missä μ on $\mathbb{R}^3$:n *tilavuusmitta*, sanotaan *avaruusintegraaliksi* (myös tilavuusintegraaliksi, engl. volume integral). Tilavuusmitta ja -integraali määritellään samalla periaattella kuin tasossa, vain sillä erotuksella, että perussuorakulmion tilalla on suorakulmainen perussärmiö, jonka mitaksi oletetaan (tilavuusmitan aksiooma)

$$T = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]: \quad \mu(T) = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3).$$

Jordan-mittana (ulko- ja sisämittojen avulla, vrt. Luku XIII.1) määritelty tilavuusmitta on siirto-, peilaus- ja kiertoinvariantti, ts. se on valitusta (karteesisesta) koordinaatistosta riippumaton.

Kuten tasointegraalit, avaruusintegraalitkin voidaan laskea suoraan numeerisesti integraalin määritelmästä. Jos $\mathcal{T}_h$ on särmiön $T \supset A$ jako osasärmiöihin, joiden särmät ovat enintään h:n pituiset, vaatii jakoon liittyvän Riemannin summan laskeminen yleisesti $N \sim h^{-3}$ laskuoperaatiota, kun vastaava luku tasossa on $N \sim h^{-2}$ ja yhdessä dimensiossa $N \sim h^{-1}$. Tarkkuus riippuu kaikissa tapauksissa tiheysparametrista h olennaisesti samalla tavalla (esimerkiksi virhe $\sim h^2$), joten samaan tarkkuuteen pyrittäessä työmäärä kasvaa voimakkaasti integraalin dimension kasvaessa. (Sama ilmiö vaivaa kaikkia numeerisen integroinnin menetelmiä — ja numeerisia laskentamenetelmiä yleisemminkin.) Riemannin summiin perustuvista menetelmistä tarkin on jälleen (yhdistetty) *keskipistesääntö*, joka integroi kussakin osasärmiössä tarkasti ensimmäisen asteen polynomin.

Esimerkki 1 Olkoon $A = [0,1] \times [0,1] \times [0,1]$ ja laskettava $\int_A 3x^2y^2z^2\,d\mu$. Kun A jaetaan suorakulmaisiin särmiöihin kokoa $h \times h \times h$ ja kunkin särmiön yli integroidaan keskipistesäännöllä, niin laskettavaksi tulee summa, jossa on $N = h^{-3}$ termiä. Tällä tavoin saadaan integraalille seuraavat likiarvot (tarkka arvo $= 1/9$, vrt. Esimerkki XIII.1: 5):

h	N	Likiarvo
0.1	10^3	0.110279859..
0.01	10^6	0.111102777..
0.001	10^9	0.111111027..
0.0001	10^{12}	0.111111110..

□

Myös avaruusintegraalin voi palauttaa peräkkäisiksi yksiulotteisiksi integraaleiksi, jolloin integraalin voi suotuisissa oloissa laskea suljetussa muodossa. Ensinnäkin jos $T = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ on suorakulmainen särmiö, ja merkitään $Q = [a_1, b_1] \times [a_2, b_2]$, niin seuraten Fubinin lauseen todistuksessa käytettyä ajatusta saadaan iteraatiokaava

$$\int_T f\,d\mu = \int_Q F(x,y)\,dxdy, \quad F(x,y) = \int_{a_3}^{b_3} f(x,y,z)\,dz.$$

Kun tässä edelleen käytetään Fubinin lauseen laskukaavaa, on tuloksena kolminkertainen iteraatiokaava

$$\int_T f\,d\mu = \int_{a_1}^{b_1} \left\{ \int_{a_2}^{b_2} \left[\int_{a_3}^{b_3} f(x,y,z)\,dz \right] dy \right\} dx.$$

Esimerkki 1 (jatko) Esimerkin integraali purkautuu kolmen integraalin tuloksi:

$$\begin{aligned}\int_A f\,d\mu &= \int_0^1 \left\{ \int_0^1 \left[\int_0^1 3x^2y^2z^2\,dz \right] dy \right\} dx \\ &= 3\int_0^1 x^2 \left\{ \int_0^1 y^2 \left[\int_0^1 z^2\,dz \right] dy \right\} dx \\ &= 3\int_0^1 x^2\,dx \int_0^1 y^2\,dy \int_0^1 z^2\,dz = 3\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3} = \frac{1}{9}. \quad \square\end{aligned}$$

Jos $A \subset \mathbb{R}^3$ on yleisempi joukko kuin suorakulmainen särmiö, niin määritellään kuten tasossa

$$\int_A f\,d\mu = \int_T f_0\,d\mu,$$

missä f_0 on f:n nollajatko suorakulmaiseen särmiöön $T \supset A$. Soveltamalla tässä ym. iteraatiokaavaa saadaan erilaisia A:n muotoon sovitettuja integraalin purkukaavoja. Esimerkiksi jos A esitetään muodossa

$$A = \{(x,y,z) \in \mathbb{R}^3 \mid x \in B \subset \mathbb{R} \ \ \& \ \ (y,z) \in C(x) \subset \mathbb{R}^2\},$$

niin saadaan purkukaava

$$\int_A f\,d\mu = \int_B \left[\int_{C(x)} f(x,y,z)\,dydz \right] dx, \qquad (\star)$$

missä sisemmän tasointegraalin voi edelleen purkaa edellisen luvun menetelmin. Purkusäännön $(\star)$ voi havainnollistaa geometrisesti 'siivutusperiaatteena', jossa A jaetaan osiin

$$\Delta A = \{(x,y,z) \in A \mid x \in \Delta B\},$$

jolloin on likimäärin (vrt. kuvio)

$$\int_{\Delta A} f\,d\mu \approx \left[\int_{C(x)} f(x,y,z)\,dydz\right]\mu(\Delta B) \quad (\mu = \text{pituusmitta}).$$

Ajatellen, että avaruusintegraali viime kädessä palautuu peräkkäisiksi yksiulotteisiksi integraaleiksi, käytetään avaruusintegraaleille usein merkintöjä

$$\int_A f\,dxdydz \quad \text{tai} \quad \iiint f\,dxdydz,$$

missä siis on merkitty $d\mu = dxdydz$.†

Esimerkki 2 A = tetraedri, jonka kärjet ovat $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ ja $(0,0,1)$. Laske $\int_A f\,dxdydz$, kun $f(x,y,z) = xyz$.

<u>Ratkaisu</u> Purkukaavan ($\star$) nojalla on ensinnäkin

$$\int_A f\,dxdydz = \int_0^1 \left[\int_{C(x)} xyz\,dydz\right] dx,$$

missä yz-tason joukko $C(x)$ määräytyy ehdoista $y \geq 0$, $z \geq 0$ ja

$$x + y + z \leq 1 \quad \Leftrightarrow \quad y + z \leq 1 - x,$$

†Sovelluksissa (etenkin fysiikassa) avaruusintegraalin mittamerkintä on usein dV.

eli $C(x)$ on kolmio, jonka kärjet ovat pisteissä $(0,0,0)$, $(0,1-x,0)$ ja $(0,0,1-x)$. Soveltaen sisempään integraaliin edellisen luvun purkusääntöjä saadaan

$$\begin{aligned}
\int_A f\,dxdydz &= \int_0^1 \left\{ \int_0^{1-x} \left[\int_0^{1-x-y} xyz\,dz \right] dy \right\} dx \\
&= \int_0^1 \left\{ \int_0^{1-x} \left[\Big|_{z=0}^{z=1-x-y} \frac{1}{2}xyz^2 \right] dy \right\} dx \\
&= \int_0^1 \left[\int_0^{1-x} \frac{1}{2}xy(1-x-y)^2\,dy \right] dx \\
\text{(os. int.)} \quad &= \int_0^1 \left[\Big|_{y=0}^{y=1-x} -\frac{1}{6}xy(1-x-y)^3 + \int_0^{1-x} \frac{1}{6}x(1-x-y)^3\,dy \right] dx \\
&= \int_0^1 \left[\Big|_{y=0}^{y=1-x} -\frac{1}{24}x(1-x-y)^4 \right] dx \\
&= \int_0^1 \frac{1}{24}x(1-x)^4\,dx \\
\text{(os. int.)} \quad &= \Big|_0^1 -\frac{1}{120}x(1-x)^5 + \int_0^1 \frac{1}{120}(1-x)^5\,dx \\
&= \Big|_0^1 -\frac{1}{720}(1-x)^6 = \underline{\underline{\frac{1}{720}}}. \quad \square
\end{aligned}$$

Purkusääntö $(\star)$ on erityisen kätevä silloin kun f ei riipu muuttujista y, z. Tällöin integraali palautuu suoraan 1-ulotteiseksi edellyttäen, että joukkojen $C(x)$ pinta-alamitta on helposti määrättävissä:

$$\int_A f\,dxdydz = \int_B f(x)\,\mu(C(x))\,dx, \quad f = f(x) \quad (\mu = \text{pinta-alamitta}).$$

Esimerkki 3 Laske R-säteisen pallon tilavuus.

<u>Ratkaisu</u> Käytetään purkukaavaa $(\star)$, jolloin

$$B = [-R, R], \quad C(x) = \{(y,z) \in \mathbb{R}^2 \mid y^2 + z^2 \le R^2 - x^2\}.$$

Koska $C(x)$ on kiekko, jonka säde $= \sqrt{R^2 - x^2}$, niin $\mu(C(x)) = \pi(R^2 - x^2)$. Siis

$$\mu(A) = \int_A d\mu = \int_{-R}^{R} \pi(R^2 - x^2)\,dx = \Big|_{-R}^{R} \pi\left(R^2 x - \frac{1}{3}x^3\right) = \frac{4}{3}\pi R^3. \quad \square$$

Esimerkki 4 Laske $\int_A x\,dxdydz$, kun $A = \{(x,y,z) \mid x^2+y^2+z^2 \le R^2,\ x \ge 0\}$ (puolipallo).

Ratkaisu Purkukaavassa $(\star)$ on tässä $B = [0,R]$ ja $C(x)$ on sama kuin edellisessä esimerkissä, joten saadaan

$$\begin{aligned}\int_A f\,dxdydz &= \int_B f(x)\,\mu(C(x))\,dx \\ &= \int_0^R x\cdot\pi(R^2-x^2)\,dx \\ &= \pi\Big|_0^R\left(\frac{1}{2}R^2x^2-\frac{1}{4}x^4\right) = \underline{\underline{\frac{1}{4}\pi R^4}}. \quad \square\end{aligned}$$

Esimerkki 5 Suorien $x=0$, $y=0$ ja käyrän $y=e^{-x}$ rajaama tasoalue pyörähtää x-akselin ympäri, jolloin syntyy avaruuden $\mathbb{R}^3$ joukko A. Laske tilavuus $\mu(A)$.

Ratkaisu A ei ole rajoitettu, joten kyseessä on epäoleellinen integraali. Kaavassa $(\star)$ on $B=[0,\infty)$ ja $C(x)$ on kiekko, jonka säde $= e^{-x}$, joten

$$\mu(A) = \int_A dxdydz = \int_0^\infty \pi(e^{-x})^2\,dx = \int_0^\infty \pi e^{-2x}\,dx = \Big|_0^\infty -\frac{\pi}{2}e^{-2x} = \underline{\underline{\frac{\pi}{2}}}. \quad \square$$

Esimerkki 6 Kaksi R-säteistä lieriötä leikkaa kohtisuorasti toisensa. Mikä on molempien lieriöiden sisään jäävän joukon A tilavuusmitta?

Ratkaisu Olkoon toisen lieriön akseli z-akseli ja toisen y-akseli, jolloin

$$A = \{(x,y,z)\in\mathbb{R}^3 \mid x^2+y^2\le R^2 \ \ \&\ \ x^2+z^2\le R^2\}.$$

Kun leikataan tämä yz-tason suuntaisilla tasoilla, niin todetaan, että

$$A = \{(x,y,z)\in\mathbb{R}^3 \mid x\in[-R,R] \ \ \&\ \ (y,z)\in C(x)\}, \quad \text{missä}$$

$$\begin{aligned}C(x) &= \{(y,z)\in\mathbb{R}^2 \mid x^2+y^2\le R^2 \ \ \&\ \ x^2+z^2\le R^2\} \\ &= \{(y,z)\in\mathbb{R}^2 \mid y^2\le R^2-x^2 \ \ \&\ \ z^2\le R^2-x^2\}.\end{aligned}$$

Poikkileikkaus $C(x)$ on tämän mukaan neliö, jonka sivun pituus $= 2\sqrt{R^2-x^2}$. Kaavan $(\star)$ mukaan on siis

$$\mu(A) = \int_{-R}^R \mu(C(x))\,dx = \int_{-R}^R 4(R^2-x^2)\,dx = 4\Big|_{-R}^R (R^2x-\frac{1}{3}x^3) = \underline{\underline{\frac{16}{3}R^3}}. \quad \square$$

Avaruusintegraalit $\mathbb{R}^n$:ssä

Yleinen *n-ulotteinen avaruusintegraali* on muotoa $I(f, A, \mu) = \int_A f\, d\mu$, missä $A \subset \mathbb{R}^n$, $f = f(\mathbf{x}) = f(x_1, \ldots, x_n)$, ja μ on *n-ulotteinen tilavuusmitta* (Jordan-mitta). Tämän ominaisuuksiin kuuluu, että n-ulotteisen suorakulmaisen särmiön mitta on (aksiooma)

$$\mu(T) = \prod_{i=1}^{n}(b_i - a_i), \quad T = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

Integraalin vaihtoehtoiset merkintätavat

$$\int_A f\, dx_1 \cdots dx_n, \quad \iint \cdots \int f\, dx_1 \cdots dx_n.$$

viittaavat jälleen siihen tosiseikkaan, että integraali on palautettavissa peräkkäisiksi yksiulotteisiksi integraaleiksi. Iteraatiokaavan perusmuoto tapauksessa $A = T$ on

$$\int_T f\, d\mu = \int_{a_1}^{b_1} \left[\int_{a_2}^{b_2} \cdots \left[\int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \right] \cdots \right] dx_1.$$

Esimerkki 7 Laske neliulotteisen R-säteisen pallon (kuulan) tilavuus $\mu(A)$.

<u>Ratkaisu</u> Kun kirjoitetaan

$$A = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 \in [-R, R] \ \ \& \ \ (x_2, x_3, x_4) \in C(x_1)\},$$

niin (vrt. purkukaava ($\star$) edellä)

$$\int_A d\mu = \int_{-R}^{R} \mu(C(x_1))\, dx_1,$$

missä μ on $\mathbb{R}^3$:n tilavuusmitta. Tässä $C(x_1) \subset \mathbb{R}^3$ on pallo, jonka säde on $\sqrt{R^2 - x_1^2}$, joten

$$\begin{aligned} \mu(A) &= \int_{-R}^{R} \frac{4}{3}\pi(R^2 - x_1^2)^{3/2}\, dx_1 \quad (\text{sij. } x_1 = R\sin t) \\ &= \frac{4\pi}{3}R^4 \int_{-\pi/2}^{\pi/2} \cos^4 t\, dt = \frac{4\pi}{3}R^4 \cdot \frac{3\pi}{8} = \underline{\underline{\frac{1}{2}\pi^2 R^4}}. \quad \square \end{aligned}$$

Mitä tulee n-ulotteisen Jordan-mitan teoreettisiin ominaisuuksiin, todettakoon ainoastaan (ilman todistusta), että nämä ominaisuudet ovat vastaavat kuin tasossa. Esimerkiksi mitta on siirto- ja kiertoinvariantti ja määriteltävissä ylä- ja alaintegraalien avulla, vrt. Luku XIII.1. Pätee myös Lauseen XIII.1.2 yleistys:

*Lause XIII.3.1 Jos f on jatkuva särmiössä $T = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ ja $A \subset T$ on mitallinen, niin f on Riemann-integroituva yli A:n.

*Suuntaissärmiön $K \subset \mathbb{R}^n$ tilavuus

Palautettakoon mieliin Luvusta X.7, että avaruuden $\mathbb{R}^n$ n-ulotteinen suuntaissärmiö, jonka kärki on pisteessä $\mathbf{x}_0$ ja kärjestä lähtevät särmävektorit ovat $\mathbf{a}_i$, on joukko

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{n} t_i \mathbf{a}_i,\ t_i \in [0,1],\ i = 1 \dots n\}.$$

Jos kirjoitetaan $\mathbf{a}_i = \sum_{j=1}^{n} a_{ij}\mathbf{e}_j$ ja muodostetaan kertoimista a_{ij} matriisi $\mathbf{A}$, niin tapauksissa $n = 2$ ja $n = 3$ tiedetään, että K:n pinta-ala ($n = 2$) tai tilavuus ($n = 3$) on laskettavissa kaavalla $\mu(K) = |\det\mathbf{A}|$ (ks. Luku II.6). Näytetään, että tämä pätee yleisesti.

Propositio XIII.3.2 Avaruuden $\mathbb{R}^n$ n-ulotteisen suuntaissärmiön K tilavuusmitta on $\mu(K) = |\det\mathbf{A}|$, missä matriisin $\mathbf{A}$ riveinä ovat K:n särmävektoreiden $\mathbf{a}_i$ kertoimet a_{ij} kannassa $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

<u>Todistus</u> Pidetään tunnettuna, että $\mathbb{R}^n$:n tilavuusmitta on kiertoinvariantti. Tällöin kantavektori $\mathbf{e}_n$ voidaan valita vektoreiden $\mathbf{a}_i$, $i = 1 \dots n-1$ virittämää hypertasoa (aliavaruutta) vastaan kohtisuoraksi, jolloin on $a_{in} = 0$, kun $i = 1 \dots n-1$. Kierto ei vaikuta myöskään determinantin $\det\mathbf{A}$ arvoon, sillä kierto vastaa muunnosta $\mathbf{A}^T = [\mathbf{a}_1, \dots, \mathbf{a}_n] \hookrightarrow [\mathbf{C}\mathbf{a}_1, \dots, \mathbf{C}\mathbf{a}_n] = \mathbf{C}\mathbf{A}^T$ eli $\mathbf{A} \hookrightarrow \mathbf{A}\mathbf{C}^T$, missä $\det\mathbf{C} = 1$. Oletetussa koordinaatistossa voidaan suuntaissärmiö esittää muodossa

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + t_n a_{nn}\mathbf{e}_n + \mathbf{y},\ \ t_n \in [0,1]\ \ \&\ \ \mathbf{y} \in A(t_n)\},$$

missä $A(t_n) \subset \mathbb{R}^{n-1}$ on $(n-1)$-ulotteinen suuntaissärmiö, jonka särmävektorit ovat $\mathbf{b}_i = \sum_{j=1}^{n-1} a_{ij}\mathbf{e}_j$, $i = 1 \dots n-1$, ja kärki on pisteessä $\mathbf{y}_0(t_n) = t_n \sum_{j=1}^{n-1} a_{nj}\mathbf{e}_j$. Koska $\mathbb{R}^{n-1}$:n tilavuusmitta on siirtoinvariantti, niin joukon $A(t_n)$ mitta on t_n:stä riippumaton; merkitään $\mu(A(t_n)) = V_{n-1}$. Purkukaavaa ($\star$) vastaten K:n mitta on tällöin

$$\mu(K) = \int_0^{|a_{nn}|} \left[\int_{A(t_n)} dx_1 \dots dx_{n-1}\right] dx_n = |a_{nn}| V_{n-1}\,.$$

Jos merkitään $\mu(K) = V_n$, niin on siis saatu palautuskaava

$$V_n = |a_{nn}| V_{n-1}\,.$$

Toisaalta jos merkitään $D_n = |\det\mathbf{A}|$ ja huomioidaan, että $a_{in} = 0$ kun $i \neq n$, niin alideterminanttisäännön (Lause X.5.5) perusteella saadaan vastaava palautuskaava

$$D_n = |a_{nn}| D_{n-1}\,.$$

Koska palautuskaavat ovat samaa muotoa ja koska $V_n = D_n$ kun $n = 2, 3$, niin päätellään, että $V_n = D_n$ $\forall n$. $\square$

HARJOITUSTEHTÄVIÄ

1. Laske tilavuus $\mu(A)$, kun $A \subset \mathbb{R}^3$ on määritelty annetuilla ehdoilla.
   a) $0 \le x \le 1$ & $0 \le y \le x$ & $0 \le z \le 1 - x^2$
   b) $0 \le y \le 1$ & $0 \le x \le y$ & $0 \le z \le 1 - x^2$
   c) $x, y \ge 0$ & $x + y \le 1$ & $0 \le z \le 1 - x^2 - y^2$
   d) $1 \le x \le 2$ & $0 \le y \le x$ & $0 \le z \le 1/(x + y)$
   e) $0 \le z \le 1 - x^2 - 2y^2$
   f) $0 \le x \le \sqrt[4]{\pi}$ & $0 \le y \le x$ & $0 \le z \le x^2 \sin(y^4)$
   g) $x^2 \le z \le 1 - y^2$
   h) $x^2 + y^2 \le 8$ & $y - 4 \le z \le 8 - x$

2. a)–e) Laske edellisen tehtävän tapauksissa a)–e) integraalit $\int_A x\, d\mu$, $\int_A y\, d\mu$ ja $\int_A z\, d\mu$.

3. Laske $\int_A f(x, y, z)\, dxdydz$ annetuilla f ja A:
   a) $f:\ xyz,\quad A:\ 0 \le x \le 1$ & $-2 \le y \le 0$ & $1 \le z \le 4$
   b) $f:\ x^2 + y^2 + z^2,\quad A:\ 0 \le x, y, z \le a\ \ (a > 0)$
   c) $f:\ (1 - x - y)^5,\quad A:\ x, y, z \ge 0$ & $x + y + z \le 1$
   d) $f:\ xyz^4,\quad A:\ x, y, z \ge 0$ & $x + y + z \le 1$
   e) $f:\ 3 + 2xy,\quad A:\ z \ge 0$ & $x^2 + y^2 + z^2 \le 4$
   f) $f:\ x,\quad A:\ x, y, z \ge 0$ & $(x/a) + (y/b) + (z/c) \le 1\ \ (a, b, c > 0)$
   g) $f:\ xy + z^2,\quad A:\ 0 \le z \le 1 - |x| - |y|$
   h) $f:\ yz^2 e^{-xyz},\quad A:\ 0 \le x, y, z \le 1$
   i) $f:\ y,\quad A:\ 0 \le x, y, z \le 1$ & $1 - y \le z \le 2 - x - y$
   j) $f:\ (x + y + z)^{-3},\quad A:\ 1 \le z \le 2$ & $0 \le y \le z$ & $0 \le x \le y + z$
   k) $f:\ \cos x \cos y \cos z,\quad A:\ x, y, z \ge 0$ & $x + y + z \le \pi$

4. Perustuen joukon $A \subset \mathbb{R}^3$ esitysmuotoon

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in B \subset \mathbb{R}^2 \ \ \&\ \ z \in C(x, y) \subset \mathbb{R}\}$$

   johda purkukaava

$$\int_A f\, d\mu = \int_B \left[\int_{C(x,y)} f(x, y, z)\, dz \right] dxdy.$$

   Sovella kaavaa, kun $A = \{\, (x, y, z) \in \mathbb{R}^3 \mid x, y, z \ge 0 \ \ \&\ \ x^2 + y + z \le 1 \}$ ja $f(x, y, z) = xy^2$.

5. Tetredria $K \subset \mathbb{R}^3$ rajoittavat tasot $z = 0$, $x = 2y$, $x = -y$ ja $y + z = a$ $(a > 0)$. Laske $\int_K z^2\, dxdydz$.

6. Teltan pohja on xy-tasolla oleva kiekko, jonka säde on $R = 2$ m, ja teltan jokainen xz-tason suuntainen poikkileikkaus on tasakylkinen kolmio, jonka korkeus on $h = 2$ m. Laske teltan tilavuus.

7. Näytä, että ellipsoidin $S : x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ sisään jäävän joukon $A \subset \mathbb{R}^3$ tilavuus on $\mu(A) = 4\pi abc/3$ (vrt. Harj.teht. VIII.8:4).

8. Olkoon $a, b > 0$, $a \neq b$. Laske lieriöiden $S_1 : x^2+y^2 = a^2$ ja $S_2 : x^2+z^2 = b^2$ sisään jäävän joukon $A \subset \mathbb{R}^3$ tilavuus $\mu(A)$ yksiulotteisena integraalina. Voiko integraalin laskea suljetussa muodossa?

9. Joukko $\{(x, y, z) \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2} \leq z \leq a\}$ edustaa täydessä viinilasissa olevaa viiniä. Lasiin upotetaan varovasti R-säteinen kuula, jolloin viiniä valuu ulos, kunnes kuula uppoaa kokonaan tai pysähtyy lasiin. Millä suhteen R/a arvolla viiniä valuu ulos eniten?

10. Olkoon $A = [0, 1] \times [0, 1] \times \ldots \times [0, 1] \subset \mathbb{R}^n$. Laske $\int_A f(\mathbf{x})\, dx_1 \ldots dx_n$, kun
a) $f(\mathbf{x}) = \sum_{i=1}^n x_i$, b) $f(\mathbf{x}) = \prod_{i=1}^n x_i$, c) $f(\mathbf{x}) = e^{-(x_1+x_2+\cdots+x_n)}$.

11. Olkoon $K \subset \mathbb{R}^n$ n-simpleksi, jonka kärjet ovat origo ja pisteet $(a, 0, \ldots, 0)$, $(0, a, 0, \ldots, 0)$, $\ldots, (0, \ldots, 0, a)$ $(a > 0)$. Näytä, että K:n n-ulotteinen tilavuus on $\mu(K) = a^n/n!$.

12. Suuntaissärmiön $K \subset \mathbb{R}^4$ särmävektorit ovat $[1, 1, 1, 1]^T$, $[0, 1, 1, 0]^T$, $[0, 0, 1, 1]^T$ ja $[1, 1, 0, 1]^T$. Laske K:n neliulotteinen tilavuus $\mu(K)$.

13. (*) Seuraavissa integraaleissa integroidaan erään joukon $A \subset \mathbb{R}^3$ yli. Määritä ensin A ja laske sitten integraalin arvo valitsemalla sopiva integroimisjärjestys.

$$\text{a)} \int_0^1 \left[\int_z^1 \left(\int_0^x e^{x^3}\, dy\right) dx\right] dz \quad \text{b)} \int_0^1 \left[\int_0^{1-x} \left(\int_y^1 \frac{\sin(\pi z)}{z(2-z)}\, dz\right) dy\right] dx$$

14. (*) Kolmen R-säteisen lieriön akselit leikkaavat toisensa kohtisuorasti samassa pisteessä. Näytä, että kaikkien kolmen lieriön sisään jäävän joukon A tilavuusmitta on $\mu(A) = 8(2 - \sqrt{2})R^3$.

15. (*) a) Laske $\mathbb{R}^n$:n R-säteisen pallon (kuulan) n-ulotteinen tilavuus, kun $n = 5$ ja $n = 6$. b) Millä n:n arvolla yksikköpallon $(R = 1)$ tilavuus on suurin? Perustele!

XIII.4 Taso- ja avaruusintegraalien muuntaminen

Olkoon $A, B \subset \mathbb{R}^n$, $n \in \mathbb{N}$ ja olkoon $\mathbf{u} : B \to A$ bijektio (tai 'melkein bijektio', ks. huomautukset jäljempänä).

Halutaan laskea integraali $\int_A f\,d\mu = \int_A f\,dx_1 \ldots dx_n$ muunnetussa muodossa

$$\int_A f\,d\mu = \int_B g\,d\mu', \quad g(\mathbf{t}) = f(\mathbf{u}(\mathbf{t})), \quad \mathbf{t} \in B.$$

Tässä g on kuvauksen $\mathbf{u}$ välittämä f:n vastine B:ssä, ja μ' on toinen, toistaiseksi tuntematon $\mathbb{R}^n$:n mitta. Kyse on siis muuttujan vaihdosta eli <u>sijoituksesta</u>

$$\mathbf{x} = \mathbf{u}(\mathbf{t}), \quad \mathbf{t} \in B.$$

Muunnoksessa on ensinnäkin määrättävä funktio g. Sikäli kuin lähtökohtana on (niinkuin yleensä) tunnettu kuvaus $\mathbf{u}$, tämä on suoraviivainen toimenpide. Myös mitan μ' formaali määrittely on helppoa. Nimittäin jos $\mathbf{u}(\Delta B) = \Delta A \subset A$ ja valitaan $f(\mathbf{x}) = 1,\ \mathbf{x} \in \Delta A$ (jolloin $g(\mathbf{t}) = 1,\ \mathbf{t} \in \Delta B$), on mitan ja integraalin yhteyden (vrt. Luku XIII.1) ja oletetun integraalien välisen yhteyden perusteella

$$\mu'(\Delta B) = \int_{\Delta B} d\mu' = \int_{\Delta A} d\mu = \mu(\Delta A).$$

Siis $\mu'(\Delta B) = \mu(\Delta A)$, eli ΔB:n mitta määräytyy vastinjoukon $\Delta A = \mathbf{u}(\Delta B)$ Jordan-mittana. — Tässä on huomattava, että kuvauksen $\mathbf{u}$ on oltava (ainakin lähes) injektio, jotta μ' olisi todella mitta. Nimittäin koska mitta on additiivinen, niin B:n mitallisille osajoukoille on oltava voimassa

$$\Delta B_1 \cap \Delta B_2 = \emptyset \quad \Rightarrow \quad \mu'(\Delta B_1 \cup \Delta B_2) = \mu'(\Delta B_1) + \mu'(\Delta B_2).$$

Jos $\mathbf{u}$ ei olisi injektio, niin joillakin $\mathbf{x}_1, \mathbf{x}_2 \in B$, $\mathbf{x}_1 \neq \mathbf{x}_2$ olisi $\mathbf{u}(\mathbf{x}_1) = \mathbf{u}(\mathbf{x}_2)$, tai yleisemmin $\mathbf{u}(\Delta B_1) = \mathbf{u}(\Delta B_2) = \Delta A$, missä $\Delta B_1, \Delta B_2 \subset B$ ja $\Delta B_1 \cap \Delta B_2 = \emptyset$.

Tällöin $\mu'(\Delta B_1 \cup \Delta B_2) = \mu(\Delta A)$ ja $\mu'(\Delta B_1) + \mu'(\Delta B_2) = 2\mu(\Delta A) \neq \mu(\Delta A)$, ellei ole $\mu(\Delta A) = 0$. Päätellään siis, että $\mathbf{u}$:n epäinjektiivisyys voidaan sallia enintään (mitan μ' suhteen) nollamittaisessa joukossa.

Sikäli kuin $\mathbf{u}$ on injektio, tai mainitulla tavalla 'melkein', on integraalin muuntamiskysymys siis periaatteessa ratkaistu. Toistaiseksi ei kuitenkaan ole selvää, millä tavoin mitan μ' avulla voidaan käytännössä laskea (muuten kuin palaamalla alkuperäisiin muuttujiin). Tämä laskutekninen kysymys onkin kaikkein keskeisin, sikäli kuin muunnoksesta halutaan jotakin hyötyä.

Muuntosuhde

Tutkitaan muunnettua integraalia $\int_B g\, d\mu'$. Oletetaan, että B on rajoitettu ja $g(\mathbf{t})$ rajoitettu B:ssä, jolloin on (vrt. Luvut XIII.1 ja XIII.3) $\int_B g\, d\mu' = \int_T g_0\, d\mu'$, missä $T \supset B$ on n-ulotteinen suorakulmainen perussärmiö ja $g_0 = g$:n nollajatko B:n ulkopuolelle. Oletetaan jatkossa, että $\mathbf{u} : T \to \mathbf{u}(T)$ on (lähes) injektio ja että $\mathbf{u}(T')$ on Jordan-mitallinen aina kun $T' \subset T$ on Jordan-mitallinen. Tällöin jos $\mathcal{T}_h = \{T_k\}$ on T:n jako osasärmiöihin, joiden särmien pituus on enintään h, niin vastaavasti kuin Jordan-mitan suhteen integroitaessa pätee summakaava

$$\int_T g_0\, d\mu' \;=\; \operatorname*{Lim}_{h\to 0} \sum_k g_0(\mathbf{t}_k)\mu'(T_k) \;=\; \operatorname*{Lim}_{h\to 0} \sum_k g_0(\mathbf{t}_k)\mu(\mathbf{u}(T_k)),$$

missä $\mathbf{t}_k \in T_k$.

Oletetaan nyt, että kuvaukselle $\mathbf{u}$ on määriteltävissä T:ssä jatkuvana funktiona $J(\mathbf{t})$ nk. *muuntosuhde* (mittasuhde, suurennussuhde) siten, että jokaisella $T_k \in \mathcal{T}_h$ ja jokaisella $\mathbf{t}_k \in T_k$ pätee[†]

$$\left| \frac{\mu(\mathbf{u}(T_k))}{\mu(T_k)} - J(\mathbf{t}_k) \right| \;\leq\; \varepsilon(h) \to 0 \quad \text{kun } \; h \to 0,$$

missä $\varepsilon(h)$ riippuu vain tiheysparametrista h (ei k:sta eikä yleisemmin $\mathcal{T}_h$:sta).

[†]Muuntosuhteen määrittely voidaan yleisemmin rajoittaa särmiöihin, joille pätee $T_k \cap U = \emptyset$, missä $\mu'(U) = 0$. Tällöin muuntosuhde ei välttämättä ole koko T:ssä jatkuva tai edes määritelty.

Tämän perusteella voidaan em. summakaavassa kirjoittaa likimäärin $\mu(\mathbf{u}(T_k) \approx J(\mathbf{t}_k)\mu(T_k)$. Olettaen, että $|g_0(\mathbf{t})| \leq C,\ \mathbf{t} \in T$ (g oli rajoitettu), saadaan tälle approksimatiolle tehtyjen oletusten perusteella virhearvio

$$\left|\sum_k g_0(\mathbf{t}_k)\mu(\mathbf{u}(T_k)) - \sum_k g_0(\mathbf{t}_k)J(\mathbf{t}_k)\mu(T_k)\right| \leq C\varepsilon(h)\sum_k \mu(T_k) = C\varepsilon(h)\mu(T).$$

Oletukseen $\lim_{h\to 0}\varepsilon(h) = 0$ perustuen ja integraalin summakaavaa uudelleen soveltaen seuraa

$$\operatorname*{Lim}_{h\to 0}\sum_k g_0(\mathbf{t}_k)\mu(\mathbf{u}(T_k)) \;=\; \operatorname*{Lim}_{h\to 0}\sum_k g_0(\mathbf{t}_k)J(\mathbf{t}_k)\mu(T_k) \;=\; \int_T g_0(\mathbf{t})J(\mathbf{t})\,d\mu.$$

Tässä on edelleen $\int_T g_0(\mathbf{t})J(\mathbf{t})\,d\mu = \int_B g(\mathbf{t})J(\mathbf{t})\,d\mu$ (koska $g_0J = (gJ)_0$), joten on päätelty:

$$\int_A f\,d\mu = \int_B g\,d\mu' = \int_B g(\mathbf{t})J(\mathbf{t})\,d\mu.$$

Integraalia $\int_A f\,d\mu$ muunnettaessa on siis muunnosten $f(\mathbf{x}) = f(\mathbf{u}(\mathbf{t})) = g(\mathbf{t})$ ja $A \hookrightarrow B$, $\mathbf{u}(B) = A$, lisäksi suoritettava paikalliseen muuntosuhteeseen $J(\mathbf{t})$ perustuva <u>mittamuunnos</u> $d\mu' = J(\mathbf{t})\,d\mu$, eli muunnoskaava on

$$\boxed{\int_A f(\mathbf{x})\,dx_1\ldots dx_n = \int_B g(\mathbf{t})J(\mathbf{t})\,dt_1\ldots dt_n.} \qquad (\star)$$

Esimerkki 1 Olkoon $A = \{(x,y) \in \mathbb{R}^2 \mid x \in [1,2] \ \ \&\ \ x \leq y \leq 2x\}$. Laske $\int_A xy\,dxdy$ käyttäen sijoitusta $(x,y) = \mathbf{u}(\mathbf{t}) = \mathbf{u}(t,s) = (t,st)$.

<u>Ratkaisu</u> Ensinnäkin todetaan, että $A = \mathbf{u}(B)$, missä $B = [1,2]\times[1,2]$, ja että $\mathbf{u}: B \to A$ on bijektio. Muuntosuhteen laskemiseksi tarkastellaan suorakulmiota $\Delta T = [t, t+\Delta t]\times[s, s+\Delta s]$, missä $t, s \geq 1$ ja $0 < \Delta t, \Delta s \leq h$. Tällöin on

$$\begin{aligned}
\Delta A \;=\; \mathbf{u}(\Delta T) \;&= \{(x,y) \in \mathbb{R}^2 \mid t \leq x \leq t+\Delta t \ \ \&\ \ sx \leq y \leq (s+\Delta s)x\}\\
\Rightarrow \quad \mu(\Delta A) &= \int_t^{t+\Delta t}\left(\int_{sx}^{(s+\Delta s)x} dy\right)dx \;=\; \int_t^{t+\Delta t}(\Delta s)x\,dx\\
&= \frac{1}{2}\Delta s\left[(t+\Delta t)^2 - t^2\right] \;=\; t\,\Delta t\Delta s + \frac{1}{2}(\Delta t)^2\Delta s.
\end{aligned}$$

Koska $\mu(\Delta A)/\mu(\Delta T) = t + \frac{1}{2}\Delta t = t + \mathcal{O}(h)$, niin muuntosuhteen määritelmän perusteella on $J(t,s) = t$. Siis kaavan $(\star)$ mukaan

$$\int_A f(x,y)\,dxdy = \int_B f(t,st)\,t\,dtds = \int_1^2 t^3\,dt\cdot\int_1^2 s\,ds = \frac{15}{4}\cdot\frac{3}{2} = \frac{45}{8}\,.$$

Tarkistus (ilman muuttujan vaihtoa):

$$\int_A xy\,dxdy = \int_1^2 \left[\int_x^{2x} xy\,dy\right] dx = \int_1^2 \left[\Big|_{y=x}^{y=2x} \frac{1}{2}xy^2\right] dx = \int_1^2 \frac{3}{2}x^3\,dx = \frac{45}{8}. \quad \square$$

Kuten esimerkissä, integraalin muuntamisen keskeisin laskutekninen ongelma on muuntosuhteen määrääminen. Jatkossa ratkaistaan tämä ongelma differentiaalilaskennan keinoin. Aloitetaan yksiulotteisesta integraalista.

Muuntosuhde $\mathbb{R}$:ssä

Olkoon $A = [a, b]$ ja $B = [c, d]$ suljettuja välejä ja $u : B \to A$ bijektio. Oletetaan, että u on jatkuvasti derivoituva välillä B. Tällöin jos $T_k = [t_{k-1}, t_k] \subset B$, missä $t_k - t_{k-1} \le h$, niin jollakin $\xi_k \in T_k$ on (Lause V.6.1).

$$\mu(u(T_k)) = |u(t_k) - u(t_{k-1})| = |u'(\xi_k)|(t_k - t_{k-1}) = |u'(\xi_k)|\mu(T_k).$$

Tässä u' (ja näin ollen myös $|u'|$) on jatkuva, joten $|u'(\xi_k)| - |u'(t)| = o(1)\ \forall t \in T_k$, kun $h \to 0$.† Näin ollen jokaisella $t \in T_k$

$$\mu(u(T_k)) \;=\; |u'(t)|\mu(T_k) + o(1)\mu(T_k), \quad \text{kun } h \to 0.$$

Muuntosuhteen määritelmän mukaan on siis $J(t) = |u'(t)|$, jolloin muunnoskaava ($\star$) saa muodon

$$\int_a^b f(x)\,dx = \int_B f(u(t))|u'(t)|\,dt, \quad u(B) = [a, b]. \tag{a}$$

Koska oletettiin, että $u :\ B \to [a, b]$ on jatkuva bijektio, niin u on välillä B joko aidosti kasvava tai aidosti vähenevä. Tällöin jos $u(\alpha) = a$ ja $u(\beta) = b$, niin

$$\begin{aligned} &u \text{ kasvava } (u'(x) \ge 0) \quad \Rightarrow\ B = [\alpha, \beta], \\ &u \text{ vähenevä } (u'(x) \le 0) \ \Rightarrow\ B = [\beta, \alpha]. \end{aligned}$$

Kummassakin tapauksessa kaava (a) voidaan kirjoittaa muotoon

$$\int_a^b f(x)\,dx = \int_B f\big(u(t)\big)|u'(t)|\,dt = \int_\alpha^\beta f(u(t))u'(t)\,dt. \tag{b}$$

Tässä (määrätyn integraalin vaihtosääntöön perustuvassa) laskukaavassa on kyse <u>sijoituksesta</u> <u>määrätyssä</u> <u>integraalissa</u>, vrt. Luku VIII.6. Kaava (b) on (tietyin edellytyksin, ks. mainittu luku) pätevä, vaikka u ei olisikaan bijektio. Sen sijaan kaava (a) <u>ei</u> ole pätevä, jos u' vaihtaa merkkinsä välillä B.

†Päättely nojaa tässä jatkuvuuden syvällisempään logiikkaan: Tehdyin oletuksin $|u'|$ on tasaisesti jatkuva välillä $[a, b]$ (Lause V.9.13, ks. myös Lause V.9.12).

Esimerkki 2 Jos integraalissa $\int_0^1 x\,dx$ tehdään sijoitus $x = u(t) = 3t - 2t^2$, niin $u'(t) = 3-4t$, $u(0) = 0$, $u(1) = 1$, ja määrätyn integraalin sijoituskaava (b) on pätevä:

$$\begin{aligned}\int_0^1 x\,dx &= \int_0^1 f(u(t))u'(t)\,dt \\ &= \int_0^1 (9t - 18t^2 + 8t^3)\,dt \\ &= \Big|_0^1 (\tfrac{9}{2}t^2 - 6t^3 + 2t^4) = \frac{1}{2}\,.\end{aligned}$$

Sen sijaan kaava (a) antaa väärän tuloksen:

$$\int_0^1 f(u(t))|u'(t)|\,dt = \int_0^{3/4} (9t - 18t^2 + 8t^3)\,dt - \int_{3/4}^1 (9t - 18t^2 + 8t^3)\,dt = \frac{49}{64}\,.$$

Kummallakin tavalla saadaan oikea tulos, jos muunnetuksi integroimisväliksi valitaan $B = [0, \frac{1}{2}]$, jolloin $u : B \to [0, 1]$ on bijektio (vrt. kuvio). $\square$

Esimerkistä nähdään, että määrätyn integraalin muunnoskaava ($\star$) antaa yleisesti väärän tuloksen, jos muunnoskuvaus ei ole injektiivinen. Ongelma on hankalampi useammassa dimensiossa, missä epäinjektiivisyyttä ei aina ole helppo havaita. Useammassa dimensiossa ei ongelman poistamiseksi myöskään ole mitään yksinkertaista, määrätyn integraalin sijoituskaavaan verrattavaa oikotietä.

Muuntosuhde $\mathbb{R}^n$:ssä

Tarkastellaan ensin kaksiulotteista muunnosta

$$\begin{cases} x = u(t,s), \\ y = v(t,s), \quad \mathbf{t} = (t,s) \in T,\ \mathbf{u} = (u,v), \end{cases}$$

missä T on perussuorakulmio. Oletetaan, että u ja v ovat jatkuvasti derivoituvia T:ssä (osittaisderivaatat olemassa ja jatkuvia T:ssä, Määritelmä XI.1.10). Tällöin suorakulmion $\Delta T = [t_0 + \Delta t] \times [s_0 + \Delta s] \subset T$ kuvautumista voi tutkia likimäärin linearisoivalla approksimaatiolla (vrt. Luku XI.6)

$$\begin{cases} x(t,s) \approx u(t_0,s_0) + u_t(t_0,s_0)(t-t_0) + u_s(t_0,s_0)(s-s_0), \\ y(t,s) \approx v(t_0,s_0) + v_t(t_0,s_0)(t-t_0) + v_s(t_0,s_0)(s-s_0). \end{cases}$$

Tämä on affiinikuvaus, jonka mukaisesti suorakulmio kuvautuu suunnikkaaksi. Suunnikkaan virittävät vektorit

$$\vec{a} = \Delta t\big[u_t(t_0, s_0)\vec{i} + v_t(t_0, s_0)\vec{j}\,\big],$$
$$\vec{b} = \Delta s\big[u_s(t_0, s_0)\vec{i} + v_s(t_0, s_0)\vec{j}\,\big].$$

Kun määritellään kuvauksen $\mathbf{u}$ Jacobin matriisi (vrt. Luku XI.6)

$$\mathbf{Ju}(t, s) = \begin{bmatrix} u_t & u_s \\ v_t & v_s \end{bmatrix},$$

niin nähdään, että suunnikkaan pinta-ala on

$$\mu(\Delta A) = |\vec{a} \times \vec{b}| = |\det[\mathbf{Ju}(t_0, s_0)]|\,|\Delta t||\Delta s|.$$

Koska $\mu(\Delta T) = |\Delta t||\Delta s|$, niin muuntosuhde on tämän perusteella $J(t, s) = |\det[\mathbf{Ju}(t, s)]|$. Oletetuin säännöllisyyehdoin tämä osoittautuu oikeaksi tulokseksi (sivuutetaan tarkemmat perustelut), joten tasointegraalin muuntosuhde saadaan *Jacobin determinantin* $\det[\mathbf{Ju}(t, s)]$ avulla:

Muuntosuhde = kuvauksen $\mathbf{u}$ Jacobin determinantin itseisarvo.

Em. tulos on pätevä myös yleisemmin n-ulotteiselle avaruusintegraalille. Nimittäin jos $\mathbf{u} : T \to \mathbf{u}(T)$ on epälineaarinen kuvaus ($T \subset \mathbb{R}^n$ perussärmiö), niin tämän linearisaatio pisteessä $\mathbf{t}$ on $\mathbb{R}^n$:n affiinikuvaus, joka kuvaa suorakulmaisen särmiön suuntaissärmiöksi. Ko. suuntaissärmiön särmävektoreista muodostettu determinantti = $\mathbf{u}$:n Jacobin (matriisin) determinantti pisteessä $\mathbf{t}$. Koska determinantin itseisarvo on särmiön tilavuus (Propositio XIII.3.2), niin ym. sääntö on pätevä. Siis yleistä n-ulotteista integraalia muunnettaessa mitan muunnoskaava on

$$\boxed{d\mu = |\det(\mathbf{Ju})|\, dt_1 \cdots dt_n.}$$

Esimerkki 3 A = suunnikas, jonka kärjet ovat $(0, 0)$, $(1, 0)$, $(2, 1)$ ja $(3, 1)$. Laske $\int_A xy\, dxdy$.

Ratkaisu Tehdään muuttujan vaihto

$$\begin{cases} x = u(t,s) = t + 2s, \\ y = v(t,s) = s, \end{cases}$$

jolloin $\mathbf{u} : B \to A$, $B = [0,1] \times [0,1]$ on bijektio ja

$$\mathbf{J} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \;\Rightarrow\; J = |\det \mathbf{J}| = 1$$

$$\Rightarrow \int_A xy\,dxdy = \int_0^1 \left[\int_0^1 (t+2s)s\,ds \right] dt$$

$$= \int_0^1 \left(\frac{1}{2}t + \frac{2}{3} \right) dt = \underline{\underline{\frac{11}{12}}}. \quad \square$$

Esimerkki 4 Laske integraali

$$\int_A e^{-x-y-z}\,dxdydz, \quad A = \{(x,y,z) \in \mathbb{R}^3 \mid x+y \geq 0 \;\&\; y+z \geq 0 \;\&\; x+z \geq 0\}.$$

Ratkaisu Tehdään A:n muotoon sopiva muuttujan vaihdos

$$\begin{cases} u = x + y \\ v = y + z \\ w = x + z \end{cases} \quad \Leftrightarrow \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix},$$

jolloin muuntosuhde on

$$J = \left(\frac{1}{2}\right)^3 \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{vmatrix} = \frac{1}{2}.$$

Kuvaus $(u,v,w) \in B \mapsto (x,y,z) \in A$ on ilmeinen bijektio, kun $B = [0,\infty) \times [0,\infty) \times [0,\infty)$, joten muunnoskaavaa $(\star)$ ja Fubinin sääntöä soveltaen on tulos

$$\int_A e^{-x-y-z}\,dxdydz = \int_B e^{-\frac{1}{2}(u+v+w)}\,\tfrac{1}{2}\,dudvdw$$

$$= \frac{1}{2} \int_0^\infty \int_0^\infty \int_0^\infty e^{-\frac{1}{2}u} e^{-\frac{1}{2}v} e^{-\frac{1}{2}w}\,dudvdw = \frac{1}{2} \left[\int_0^\infty e^{-\frac{1}{2}u}\,du \right]^3 = \underline{\underline{4}}. \quad \square$$

Integraalit käyräviivaisissa koordinaatistoissa

Integraalien muunnoskaavat tulevat fysikaalisissa sovelluksissa käyttöön useimmin silloin, kun karteesisesta koordinaatistosta halutaan siirtyä käyräviivaiseen napa-, lieriö- tai pallokoordinaatistoon. Jos tasossa siirrytään napakoordinaatteihin, niin muunnoksen

$$x = r\cos\varphi = u(r,\varphi), \quad y = r\sin\varphi = v(r,\varphi)$$

Jacobin matriisi on

$$\mathbf{J} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}.$$

Tämän determinantin itseisarvo on $J = r$, joten muunnoskaavaksi tulee

$$\boxed{\int_A f(x,y)\,dxdy = \int_B g(r,\varphi)\,rdrd\varphi.}$$

Esimerkiksi jos $A =$ origokeskinen, R-säteinen kiekko, niin kaava on pätevä, kun $B = [0,R]\times[0,2\pi]$. Muunnoksen lievästä epäinjektiivisyydestä ei tässä ole haittaa, koska se rajoittuu nollamittaiseen osajoukkoon ($r = 0$, $\varphi = 0$ tai $\varphi = 2\pi$).

Esimerkki 5 Laske $\displaystyle\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dxdy$.

<u>Ratkaisu</u> Siirrytään napkoordinaatistoon. Koordinaattimuunnos $\mathbf{u} : B \to \mathbb{R}^2$ on (melkein) bijektio, kun valitaan $B = \{(r,\varphi)\in\mathbb{R}^2 \mid r \geq 0,\ 0 \leq \varphi \leq 2\pi\}$, joten

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dxdy &= \int_B e^{-r^2} r\,drd\varphi \\ &= \int_0^{2\pi}\left[\int_0^\infty e^{-r^2} r\,dr\right]d\varphi = 2\pi\Big|_0^\infty\left(-\frac{1}{2}e^{-r^2}\right) = \underline{\underline{\pi}}. \quad \square\end{aligned}$$

Esimerkin integraali laskettuna karteesisessa koordinaatistossa on

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dxdy &= \int_{\mathbb{R}^2} e^{-x^2}\cdot e^{-y^2}\,dxdy \\ &= \left[\int_{-\infty}^\infty e^{-x^2}\,dx\right]\left[\int_{-\infty}^\infty e^{-y^2}\,dy\right] = \left[\int_{-\infty}^\infty e^{-x^2}\,dx\right]^2.\end{aligned}$$

Saatiin siis hauska tulos:

Propositio XIII.4.1 $\int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}$.

Siirryttäessä kolmessa dimensiossa lierökoordinaatistoon on muuntosuhde sama kuin tasossa siirryttäessä polaarikoordinaatistoon, eli $J(r,\varphi,z) = r$. Pallokoordinaatistoon siirryttäessä saadaan muuntosuhteeksi

$$J = |\det \mathbf{J}| = \begin{vmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} = r^2\sin\theta.$$

Integraalin muunnoskaavat lieriö- ja pallokoordinaatistoon ovat näin ollen

$$\boxed{\begin{array}{ll} \text{Lieriökoord.:} & \displaystyle\int_A f(x,y,z)\,dxdydz = \int_B g(r,\varphi,z)\,r\,drd\varphi dz \\ \text{Pallokoord.:} & \displaystyle\int_A f(x,y,z)\,dxdydz = \int_B g(r,\theta,\varphi)\,r^2\sin\theta\,drd\theta d\varphi \end{array}}$$

Esimerkki 6 Laske $I = \int_A (x^2+y^2)\,dxdydz$, kun A on *pallokuori* eli kahden pallopinnan väliin jäävä alue:

$$A = \{\,(x,y,z) \in \mathbb{R}^3 \mid R_1^2 \le x^2+y^2+z^2 \le R_2^2\,\}.$$

<u>Ratkaisu</u> Lasku käy helpoiten pallokoordinaattien avulla:

$$\begin{aligned} I &= \int_0^{2\pi}\left\{\int_0^{\pi}\left[\int_{R_1}^{R_2} r^2\sin^2\theta\cdot r^2\sin\theta\,dr\right]d\theta\right\}d\varphi \\ &= \int_0^{2\pi} d\varphi\cdot\int_0^{\pi}\sin^3\theta\,d\theta\cdot\int_{R_1}^{R_2} r^4\,dr = 2\pi\cdot\frac{4}{3}\cdot\frac{1}{5}(R_2^5-R_1^5) = \underline{\underline{\frac{8\pi}{15}(R_2^5-R_1^5)}}. \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Muunna integraali $\int_A f(x,y)\,dxdy$ annettua muunnosta $\mathbf{u}$ käyttäen laskemalla ensin tarkasti suhde $\mu(\Delta A)/\mu(\Delta T)$, missä $\Delta T = [t, t+\Delta t, s, s+\Delta s]$ ja $\Delta A = \mathbf{u}(\Delta T)$, ja muuntosuhde tämän raja-arvona, kun $\max\{\Delta t, \Delta s\} \to 0$:
   a) $A:\ x \in [1,2]\ \ \&\ \ -x^2 \le y \le 3x^2,\ \ \mathbf{u}(t,s) = (t, st^2)$
   b) $A:\ x \in [1,\infty)\ \ \&\ \ 1/x^2 \le y \le 2/x^2,\ \ \mathbf{u}(t,s) = (t, s/t^2)$

2. Laske sopivalla muunnoksella:
   a) $\int_A (x+2y)^4(x-2y)^6\,dxdy,\ \ A:\ |x+2y| \le 1\ \ \&\ \ |x-2y| \le 2$
   b) $\int_A (2x+3y)^2(x-5y)^2\,dxdy,\ \ A:\ |2x+3y| \le 4\ \ \&\ \ |x-5y| \le 3$

3. Laske seuraavat integraalit sijoituksella muotoa $x = as\,\cos t,\ y = bs\,\sin t$.

$$\text{a)}\ \int_{\mathbb{R}^2} e^{-3x^2-4y^2}\,dxdy \qquad \text{b)}\ \int_{\mathbb{R}^2} \frac{1}{(1+x^2+4y^2)^2}\,dxdy$$

$$\text{c)}\ \int_A \ln\left(1+\frac{x^2}{4}+\frac{y^2}{9}\right)dxdy, \quad A:\ x,y \geq 0\ \ \&\ \ 9x^2+4y^2 \leq 36$$

4. a) Joukko $A \subset \mathbb{R}^2$ sijaitsee koordinaattineljänneksessä $x, y > 0$ ja rajoittuu käyriin $y^3 = ax^2$, $y^3 = bx^2$, $x^4 = cy^3$ ja $x^4 = dy^3$, missä $0 < a < b$ ja $0 < c < d$. Laske pinta-ala $\mu(A)$ sijoituksella $y^3 = ux^2,\ x^4 = vy^3$.

   b) Ratkaise Harj.teht. XIII.3:7 muuttujan vaihdolla $(x, y, z) = (au, bv, cw)$.

5. Laske integraali $\int_A f\,dxdydz$ lieriö- tai pallokoordinaatteihin siirtymällä:

   a) $A:\ x^2+y^2 \leq z \leq \sqrt{2-x^2-y^2}, \quad f(x,y,z) = z$
   b) $A:\ x^2+y^2+z^2 \leq 12,\ z \geq x^2+y^2, \quad f(x,y,z) = 1$
   c) $A:\ x^2+y^2+z^2 \leq R^2,\ x,y,z \geq 0, \quad f(x,y,z) = x$
   d) $A:\ x^2+y^2+z^2 \leq R^2,\ x,y,z \geq 0, \quad f(x,y,z = xyz$
   e) $A:\ 1 \leq x^2+y^2 \leq 4,\ 0 \leq y \leq x,\ 0 \leq z \leq 1, \quad f(x,y,z) = x^2+y^2$
   f) $A:\ 4 \leq x^2+y^2+z^2 \leq 9,\ z \geq 0, \quad f(x,y,z) = z$
   g) $A:\ x^2+y^2 \geq R^2,\ x^2+y^2+z^2 \leq R^2, \quad f(x,y,z) = x^2+y^2$

6. Puolikartio $z = \sqrt{x^2+y^2}$ jakaa origokeskiset pallot kahteen osaan. Laske näiden osien tilavuuksien suhde pallokoordinaattien avulla.

7. (*) Olkoon $f(x,y) = e^{(y-x)/(y+x)}$ ja A kolmio, jonka kärjet ovat $(0,0)$, $(1,0)$ ja $(0,1)$. Laske integraali $\int_A f\,dxdy$ a) polaarikoordinaattien avulla, b) sijoituksella $u = y-x,\ v = y+x$.

8. (*) Laske tasointegraali $\displaystyle\int_0^\infty \left[\int_0^x \frac{e^{-x-2y}}{x+2y}\,dy\right]dx$.

9. (*) Laske tilavuus $\mu(A)$, kun $A \subset \mathbb{R}^3$ määritellään ehdoilla

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \leq 1\ \ \&\ \ z+y \geq b \quad (a,b,c>0).$$

10. (*) Olkoon K suuntaissärmiö, jonka yksi kärki on origossa ja origosta lähtevien särmien toiset päätepisteet ovat $(2,1,1)$, $(-1,2,2)$ ja $(0,-2,1)$. Laske $\int_K xye^z\,dxdydz$ sopivalla muuttujan vaihdolla.

XIII.5 Integraalien sovellukset: Tiheys ja kokonaismäärä

Fysikaalisissa yhteyksissä käytetään tasointegraaleista usein nimeä pintaintegraali ja $\mathbb{R}^3$:n avasuusintegraalista nimeä tilavuusintegraali.

Tasointegraalista tulee 'pintaintegraali', kun taso ajatellaan $\mathbb{R}^3$:n (tai E^3:n) avaruustasoksi ja pinta-alamitta ko. tasolla määritellyksi Jordan-mitaksi. Fysikaalisesti A voi olla esimerkiksi levymäisen kappaleen sivupinta tai yleisemmän kolmiulotteisen kappaleen ulkopinnan tasomainen osa.

Ym. tavalla ymmärrettynä tasointegraalista tulee myös matemattisena käsitteenä erikoistapaus yleisemmästä *pintaintegraalista* (engl. surface integral), jossa pinta voi olla kaareva. Pintaintegraaleja tässä yleisemmässä merkityksessä käsitellään edempänä Luvussa XIII.7. Myös yksiulotteinen integraali voidaan tulkita 'avaruudellisesti' ajattelemalla, että kyseessä on avaruussuoralla määritelty pituusmitta. Näin ymmärrettynä yksiulotteinen integraali on erikoistapaus avaruuden (myös tason) käyrään liitettävästä *viiva-* eli *käyräintegraalista.* Näitä käsitellään seuraavassa luvussa.

Fysikaalisissa sovelluksissa sanotaan yleensä alueeksi (tilavuusintegraalin tapauksessa joskus 'tilavuudeksi') joukkoa, jonka yli integroidaan. Termi *alue* (engl. domain tai region) on myös matematiikassa tietyn tyyppisistä joukoista käytetty termi. Jatkossa ymmärettäköön alue kuitenkin 'fysikaalisena joukkona'.

Tiheys ja kokonaismäärä

Olkoon $A \subset \mathbb{R}^d$ joukko tai fysikaalisemmin alue, missä $d \in \{1, 2, 3\}$. Tyypillisessä integraalien sovellustilanteessa tunnetaan fysikaalisen suureen *tiheys* (tiheysfunktio, jakauma) alueessa A ja tehtävänä on laskea suureen *kokonaismäärä* kaavasta

$$\boxed{\text{Kokonaismäärä } A\text{:ssa} = \text{tiheyden integraali yli } A\text{:n.}}$$

Jos tiheysfunktio $= f$ ja kokonaismäärää merkitään symbolilla F, niin laskukaava on siis

$$F = \int_A f \, d\mu,$$

missä μ on $\mathbb{R}^d$:n Jordan-mitta. Tiheysfunktio ja kokonaismäärä voivat olla myös vektoriarvoisia: Esimerkiksi jos $A \subset \mathbb{R}^3$ ja tiheysfunktio on A:ssa määritelty vektorikettä $\vec{f}(x,y,z) = f_1(x,y,z)\vec{i}+f_2(x,y,z)\vec{j}+f_3(x,y,z)\vec{k}$, niin kokonaismäärä on

$$\vec{F} = \int_A \vec{f}\,d\mu = \vec{i}\int_A f_1\,dxdydz + \vec{j}\int_A f_2\,dxdydz + \vec{k}\int_A f_3\,dxdydz.$$

Sovellusesimerkkejä em. ajattelusta ovat vaikkapa seuraavat:

tiheys	kokonaismäärä
ρ = massatiheys [kg/m^d]	m = kokonaismassa [kg]
σ = varaustiheys [C/m^d]	Q = kokonaisvaraus [C]
$\vec{f}$ = voimatiheys [N/m^d]	$\vec{F}$ = kokonaisvoima [N]

Jos fysikaalinen suure on jakautunut tasomaiselle (tai yleisemmälle, ks. Luku XIII.7) pinnalle, niin tiheyttä sanotaan *pintatiheydeksi* (esim. pintavaraustiheys, yksikkö C/m^2). Suoralle (tai yleisemmälle käyrälle, ks. Luku XIII.6) jakautuneen suureen yhteydessä puhutaan vastaavasti *viivatiheydestä* (esim. massan viivatiheys langassa, yksikkö kg/m).

Esimerkki 1 Ilman tiheys korkeudella x maanpinnasta olkoon $\rho(x) = \rho_0 e^{-x/a}$, missä $\rho_0 = 1$ kg/m^3 ja $a = 10$ km. Kuinka paljon ilmaa (yksikkö = kg) on kuvitellussa, pystysuorassa ja 100 km korkeassa putkessa, jonka poikkipinta-ala $= 1$ m^2 ?

<u>Ratkaisu</u> Ilman massatiheys (viivatiheys) kuvitellussa putkessa on

$$f(x) = 1\ \mathrm{m}^2 \cdot \rho(x) = e^{-x/a}\,\frac{\mathrm{kg}}{\mathrm{m}},$$

joten kokonaismassa on

$$\begin{aligned} m &= 1\,\frac{\mathrm{kg}}{\mathrm{m}} \cdot \int_0^{100\,\mathrm{km}} e^{-x/a}\,dx \\ &\approx 1\,\frac{\mathrm{kg}}{\mathrm{m}} \cdot \int_0^{\infty} e^{-x/a}\,dx = 1\,\frac{\mathrm{kg}}{\mathrm{m}} \cdot a = 1\,\frac{\mathrm{kg}}{\mathrm{m}} \cdot 10000\,\mathrm{m} = \underline{\underline{10000\,\mathrm{kg}}}. \quad \square \end{aligned}$$

Esimerkki 2 Neliön muotoisella metsäaukiolla $A = [0,a] \times [0,a]$ on lumi jakautunut epätasaisesti siten, että massan pintatiheys on

$$\rho(x,y) = 36\rho_0 \left[1 + \left(\frac{x+y}{a}\right)^2\right], \quad \rho_0 = 1 \text{ kg/m}^2.$$

Suuriko on lumen kumulatiivinen sademäärä [cm], jos tasan jakautuneelle lumelle pätee 1 kg/m$^2 \mathrel{\widehat{=}}$ 1 cm?

<u>Ratkaisu</u> Lumen kokonaismassa on

$$\begin{aligned} m &= \int_A \rho\, dxdy = 36\rho_0 \int_0^a \int_0^a \left[1 + \left(\frac{x+y}{a}\right)^2\right] dxdy \\ &= 36\rho_0 \int_0^a \left\{ \Big|_{y=0}^{y=a} \left[y + \frac{a}{3}\left(\frac{x+y}{a}\right)^3\right] \right\} dx \\ &= 36\rho_0 \int_0^a \left[a + \frac{1}{3a^2}(x+a)^3 - \frac{1}{3a^2}x^3\right] dx \\ &= 36\rho_0 \Big|_0^a \left[ax + \frac{1}{12a^2}(x+a)^4 - \frac{1}{12a^2}x^4\right] \\ &= 78\rho_0 a^2 \mathrel{\widehat{=}} \underline{\underline{78 \text{ cm}}} \quad (\rho_0 a^2 \mathrel{\widehat{=}} 1 \text{ cm}). \quad \square \end{aligned}$$

Esimerkki 3 Puolipallon muotoisessa hiekkakasassa $A:\ x^2+y^2+z^2 \le R^2,\ z \ge 0$ on hiekan massatiheys

$$\rho(x,y,z) = \rho_0 + \rho_1(x,y,z) = \rho_0 + 0.036\rho_0(1 - z/R),$$

missä ρ_0 (= vakio) on kuivan hiekan ja ρ_1 on hiekkaan sitoutuneen veden tiheys. Montako prosenttia hiekkakasan koko massasta on vettä?

<u>Ratkaisu</u> Kuivan hiekan massa on

$$m_0 = \int_A \rho_0\, dxdydz = \rho_0 \mu(A) = \frac{2}{3}\pi\rho_0 R^3 \approx 0.667\pi\rho_0 R^3$$

ja veden (vrt. Esimerkki XIII.3: 4)

$$m_1 = \int_A \rho_1\, dxdydz = 0.036\rho_0 \left(\frac{2}{3}\pi R^3 - \frac{1}{4}\pi R^3\right) = 0.015\pi\rho_0 R^3.$$

Veden suhteellinen osuus on siis $0.015/0.682 \approx 0.022$. Vastaus: 2.2 %. $\square$

Kokonaismäärän aksioomat

Em. esimerkeissä pidettiin tiheyden ja kokonaismäärän välistä integraalikaavaa 'annettuna' eli fysikaalisena lähtöoletuksena. Integraalikaava on kuitenkin mahdollista johtaa paljon yksinkertaisemmista oletuksista, joita voidaan pitää edellä käytettyjen matemaattisten mallien yhteisinä perusaksioomina. Olkoon tiheysfunktio f ja merkitään kokonaismäärää A:ssa ($A \subset \mathbb{R}^d$) symbolilla $F(A)$. Oletetaan:

A1. Additiivisuus: Jos $\mu(A \cap B) = 0$, niin $F(A \cup B) = F(A) + F(B)$.

A2. Vertailuperiaate: Jos jokaisella $\mathbf{x} \in A$ pätee $m \leq f(\mathbf{x}) \leq M$, niin $m\,\mu(A) \leq F(A) \leq M\mu(A)$.

Jos nyt $A \subset \mathbb{R}^d$ ja halutaan määrätä kokonaismäärä $F(A)$, niin tulkitaan tiheysfunktio f määritellyksi A:n ulkopuolella nollajatkona f_0. Tällöin jos $T \supset A$ on suljettu väli / perussuorakulmio / suorakulmainen perussärmiö ja $\mathcal{T}_h$ on T:n jako, niin oletuksista A1–A2 (kun $f = f_0$) seuraa

$$\underline{\sigma}_h(f_0, \mathcal{T}_h) \leq F(A) \leq \overline{\sigma}_h(f_0, \mathcal{T}_h),$$

missä $\underline{\sigma}_h(f_0, \mathcal{T}_h)$ ja $\overline{\sigma}_h(f_0, \mathcal{T}_h)$ ovat jakoon $\mathcal{T}_h$ liittyvät ala- ja yläsummat (vrt. Luvut VIII.5 ja XIII.1). Koska tämä arvio on pätevä jokaisella $\mathcal{T}_h$, niin sikäli kuin f on integroituva yli A:n, on oltava $F(A) = \int_A f\,d\mu$. Integraalikaava siis seuraa oletetuista (sovelluksissa yleensä helposti hyväksyttävistä) aksioomista A1–A2.

Esimerkki 4 Olkoon $A = [a, b]$ ja $F(A)$ käyrän $y = f(x)$ ja x-akselin väliin jäävän tasoalueen pinta-ala. Tällöin jos $f(x) \geq 0$ $\forall x \in A$ ja f on integroituva välillä A, niin aksioomista A1–A2 seuraa integraalikaava $F(A) = \int_A f\,dx$. — Vrt. Luvun VIII.8 tarkastelut.

Momentti

Jos $\vec{f} = f_1\vec{i} + f_2\vec{j} + f_3\vec{k}$ on voimatiheys $\mathbb{R}^3$:ssa, niin ko. voimien momenttitiheys avaruuden pisteen $P_0 \mathrel{\widehat{=}} \vec{r}_0 = x_0\vec{i} + y_0\vec{j} + z_0\vec{k}$ suhteen pisteessä (x, y, z) on $\vec{m} = (\vec{r} - \vec{r}_0) \times \vec{f}$, missä $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$. Tämän mukaisesti alueeseen $A \subset \mathbb{R}^3$ jakautuneiden voimien kokonaismomentti pisteen P_0 suhteen on

$$\vec{M} = \int_A \vec{m}\,d\mu = \int_A (\vec{r} - \vec{r}_0) \times \vec{f}\,d\mu,$$

missä μ on $\mathbb{R}^3$:n tilavuusmitta. Kaava on pätevä myös, jos $\vec{f}$ on pintatiheys avaruustasolla T ($A \subset T$) tai viivatiheys avaruussuoralla S ($A \subset S$), jolloin μ on vastaavasti 2-ulotteinen tai 1-ulotteinen Jordan-mitta. Esimerkiksi jos $\vec{f}$ on pintatiheys xy-tason alueessa A ja $P_0 =$ origo, niin kokonaismomentin laskukaava on

$$\begin{aligned}\vec{M} &= \int_A (x\vec{i} + y\vec{j}\,) \times (f_1\vec{i} + f_2\vec{j} + f_3\vec{k}\,)\,dxdy \\ &= \vec{i}\int_A y f_3(x,y)\,dxdy - \vec{j}\int_A x f_3(x,y)\,dxdy + \vec{k}\int_A [x f_2(x,y) - y f_1(x,y)]\,dxdy.\end{aligned}$$

Esimerkki 5 **Sulkuportti**. Kanavan sulkuportti on neliön muotoinen, sivun pituus 4 metriä. Portti aukeaa oven tavoin. Portin toisella puolen on vettä koko portin korkeudella (4 m), toisella puolella ei ole vettä. Määritä porttiin kohdistuva kokonaisvoima ja porttia auki vääntävä momentti M.

Ratkaisu Porttiin kohdistuu normaalin suuntainen paine (ks. kuvio)

$$\begin{aligned}&\vec{f}(x,y) = \rho_0 g(a-y)\vec{k}, \quad (x,y) \in A,\\ &A = [0,a]\times[0,a], \quad a = 4\,\mathrm{m},\\ &\rho_0 g = 1000\,G/\mathrm{m}^3, \quad G \approx 9.8\,\mathrm{N}.\end{aligned}$$

Porttiin kohdistuva kokonaisvoima on paineen (pintatiheyden) integraali:

$$\begin{aligned}\vec{F} = \int_A \vec{f}\,dxdy &= 1000\,G\,\mathrm{m}^{-3}\vec{k}\int_0^a \left[\int_0^a (a-y)\,dy\right] dx \\ &= 500\,G\,\mathrm{m}^{-3}a^3\vec{k} = \underline{\underline{32000\,G\vec{k}}}.\end{aligned}$$

Momentin (porttia kiinni vääntävänä positiivinen) $\vec{j}$-komponentti on

$$\begin{aligned}M_y &= -\int_A x f_3(x,y)\,dxdy = -1000\,G\,\mathrm{m}^{-3}\int_0^a \left[\int_0^a x(a-y)\,dy\right] dx \\ &= -1000\,G\,\mathrm{m}^{-3}\cdot\frac{1}{4}a^4 = -64000\,G\,\mathrm{m}.\end{aligned}$$

Siis luukkua auki vääntävä momentti on

$$M = -M_y \approx \underline{\underline{6.3\cdot 10^5\,\mathrm{Nm}}}. \quad \square$$

Painopiste. Keskiö

Tavallisin esimerkki avaruuteen jakautuneesta voimasta on gravitaatio, jonka tiheys on

$$\vec{f}(x,y,z) = \rho(x,y,z)g\vec{e}, \quad (x,y,z) \in A.$$

Tässä A voi edustaa esimerkiksi kiinteää kappaletta, $\rho =$ massatiheys [kg/m^3], g on gravitaation kiihtyvyys (maan pinnalla $g \approx 9.8\,\mathrm{m/s^2}$), ja $\vec{e} =$ gravitaatiovoiman suuntavektori (yksikkövektori). Kokonaisvoiman

$$\vec{G} = \int_A \vec{f}\,dxdydz = mg\vec{e}, \quad m = \int_A \rho\,dxdydz$$

ohella kiinnostava on gravitaatiovoimien kokonaismomentti, joka pisteen $P_0 \mathrel{\widehat{=}} \vec{r}_0 = x_0\vec{i} + y_0\vec{j} + z_0\vec{k}$ suhteen on

$$\vec{M} = \int_A (\vec{r} - \vec{r}_0) \times \vec{e}\,\rho g\,dxdydz = g\left(\int_A (\vec{r} - \vec{r}_0)\rho\,dxdydz\right) \times \vec{e}.$$

Pistettä P_0 sanotaan kappaleen painopisteeksi (engl. center of gravity), jos painovoimien momentti P_0:n suhteen $= \vec{0}$ riippumatta vektorista $\vec{e}$ (eli riippumatta kappaleen asennosta suhteessa painovoimakenttään). Näin on täsmälleen kun

$$\vec{0} = \int_A (\vec{r} - \vec{r}_0)\rho\,dxdydz = \int_A \rho\vec{r}\,dxdydz - \vec{r}_0 \int_A \rho\,dxdydz,$$

joten painopisteen paikkavektori on

$$\boxed{\vec{r}_0 = \frac{1}{m(A)} \int_A \rho\vec{r}\,dxdydz, \quad m(A) = \int_A \rho\,dxdydz.}$$

Tässä $m(A) =$ kappaleen massa. Painopisteen koordinaatit ovat siis

$$x_0 = \frac{1}{m(A)} \int_A x\rho(x,y,z)\,dxdydz, \quad \text{jne.}$$

Jos $\rho = \rho_0 =$ vakio, niin painopisteen paikkavektori on

$$\vec{r}_0 = \frac{1}{\mu(A)} \int_A \vec{r}\,dxdydz \quad (\rho = \text{vakio}).$$

Tämän matemaattinen yleistys on joukon $A \subset \mathbb{R}^n$ *keskiö*, joka määritellään

$$\mathbf{x}_0 = \frac{1}{\mu(A)} \int_A \mathbf{x}\,d\mu = \frac{1}{\mu(A)} \sum_{i=1}^{n} \left(\int_A x_i\,d\mu\right) \mathbf{e}_i,$$

missä μ on n-ulotteinen Jordan-mitta. Koska funktiot $f_i(\mathbf{x}) = x_i$ epäilemättä ovat integroituvia jokaisen mitallisen joukon yli (Lause XIII.3.1), niin joukon $A \subset \mathbb{R}^n$ keskiö on määritelty aina kun A on mitallinen ja $\mu(A) \neq 0$.

Esimerkki 6 Määritä seuraavien joukkojen keskiöt.

a) $A = \{(x,y,z) \in \mathbb{R}^3 \mid \ x^2+y^2 \le R^2 \ \ \& \ \ x \ge 0 \ \ \& \ \ z \in [-H/2, H/2]\,\}$
b) $A = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \le R^2 \ \ \& \ \ x \ge 0\,\}$

Ratkaisu a) Tässä on $\mu(A) = \frac{1}{2}\pi R^2 H$ (kuten integroimalla helposti selviää), ja symmetriasyistä $\int_A y\,dxdydz = \int_A z\,dxdydz = 0$, joten $y_0 = z_0 = 0$. Keskiön x-koordinaatti on lieriökoordinaateilla laskien

$$\begin{aligned} x_0 = \frac{1}{\mu(A)}\int_A x\,dxdydz &= \frac{2}{\pi R^2 H}\int_0^R \left\{ \int_{-\pi/2}^{\pi/2} \left[\int_{-H/2}^{H/2} r\cos\varphi \cdot r\,dz \right] d\varphi \right\} dr \\ &= \frac{2}{\pi R^2 H} \cdot \int_0^R r^2\,dr \cdot \int_{-\pi/2}^{\pi/2} \cos\varphi\,d\varphi \cdot \int_{-H/2}^{H/2} dz \\ &= \frac{2}{\pi R^2 H} \cdot \frac{1}{3}R^3 \cdot 2 \cdot H = \underline{\underline{\frac{4}{3\pi}R}}. \end{aligned}$$

b) Tässäkin on $y_0 = z_0 = 0$, ja (ks. Esimerkit XIII.3: 3–4)

$$x_0 = \frac{1}{\mu(A)}\int_A x\,dxdydz = \frac{3}{2\pi R^3} \cdot \frac{\pi R^4}{4} = \underline{\underline{\frac{3}{8}R}}. \quad \square$$

Pappuksen sääntö

Olkoon A xy-tason alue, jolle pätee $y \ge 0$ $\forall (x,y) \in A$. Halutaan laskea sen pyörähdyskappaleen V tilavuus, joka syntyy A:n pyörähtäessä x-akselin ympäri. Olkoon $B = \{x \in \mathbb{R} \mid (x,y) \in A$ jollakin $y \in \mathbb{R}\}$. Tällöin V voidaan esittää muodossa

$$\begin{aligned} &V = \{(x,y,z) \in \mathbb{R}^3 \mid x \in B \ \ \& \ \ (y,z) \in C(x)\}, \\ &\text{missä} \ \ C(x) = \{(y,z) \in \mathbb{R}^2 \mid \sqrt{y^2+z^2} \in D(x)\}, \\ &\text{missä} \ \ D(x) = \{y \in \mathbb{R} \mid (x,y) \in A\} \subset [0,\infty). \end{aligned}$$

Tämän perusteella on ensinnäkin (vrt. Luku XIII.3)

$$\mu(V) = \int_B \mu\big(C(x)\big)\,dx.$$

Koska tässä $C(x)$ on ympyräsymmetrinen yz-tason origon suhteen, niin napakoordinaattimuunnoksella $y = r\cos\varphi$, $z = r\sin\varphi$ saadaan pinta-ala $\mu\big(C(x)\big)$ lasketuksi muodossa

$$\mu\big(C(x)\big) = \int_{r\in D(x)} \int_{\varphi=0}^{2\pi} r\,drd\varphi = 2\pi \int_{D(x)} r\,dr = 2\pi \int_{D(x)} y\,dy.$$

Näin ollen

$$\mu(V) = 2\pi \int_B \left[\int_{D(x)} y\, dy \right] dx = 2\pi \int_A y\, dxdy.$$

Tässä voidaan vielä kirjoittaa $\int_A y\, dxdy = y_0 \mu(A)$, missä $\mu(A) = A$:n pinta-ala ja $y_0 = A$:n keskiön y-koordinaatti. Näin muodoin saadaan *Pappuksen* (Pappoksen) *sääntönä*† tunnettu laskukaava

$$\boxed{\mu(V) = \mu(A) \cdot s, \quad s = 2\pi y_0 \quad \text{(Pappuksen sääntö).}}$$

Tässä $s =$ A:n keskiön pyörähdyksessä kulkema matka.

Esimerkki 7 Kun ympyräviiva $K:\ x^2 + (y-R)^2 = a^2$, missä $R \geq a$, pyörähtää x-akselin ympäri, niin syntyvän pyörähdyspinnan (toruksen) sisään jäävän alueen tilavuus on Pappuksen säännön mukaan $\mu(V) = \pi a^2 \cdot 2\pi R = 2\pi^2 a^2 R$. $\square$

Massamitta. Jordan-mitan yleistykset

Jos $A \subset \mathbb{R}^3$ edustaa kolmiulotteista kappaletta, jonka massatiheys ρ ei ole vakio, niin kappaleen painopiste määritellään

$$\vec{r}_0 = \frac{1}{m(A)} \int_A \vec{r}\, \rho\, dxdydz, \quad m(A) = \int_A \rho\, dxdydz.$$

Kun tässä kirjoitetaan $dm = \rho\, dxdydz$, niin painopisteen lauseke voidaan esittää hieman elegantimmin muodossa

$$\vec{r}_0 = \frac{1}{m(A)} \int_A \vec{r}\, dm, \quad m(A) = \int_A dm.$$

Näin tulee määritellyksi massatiheyteen ρ liitettävä <u>massamitta</u> m. Kyseessä on todellakin (additiivinen ja dimensiottomana positiivinen) mitta, joka siis A:n tilavuuden sijasta mittaa A:n sisältämää kokonaismassaa.

Yleisemmin jos $\rho(\mathbf{x})$ on $\mathbb{R}^n$:ssä määritelty ei-negatiivinen funktio, joka on integroituva yli joukon $A \subset \mathbb{R}^n$, niin A:lle voidaan määritellä mitta

$$\mu_\rho(A) = \int_A \rho\, d\mu = \int_A d\mu_\rho.$$

Funktiota ρ sanotaan tällöin *mitan* μ_ρ *tiheysfunktioksi.* Rajoitettu joukko $A \subset \mathbb{R}^n$ katsotaan μ_ρ-mitalliseksi, kun ρ on integroituva (tavallisessa tai laajennetussa mielessä) yli A:n. — Jordanin pituus-, pinta-ala- ja tilavuusmitat voidaan siis nähdä yleisempien mittojen erikoistapauksina: Jordanin mitan mukainen joukon 'massasisältö' saadaan asettamalla mitan tiheysfunktioksi $\rho = 1$.

†Pappuksen sääntöä on sanottu myös *Guldinin säännöksi* sveitsiläisen matemaatikon **Paul Guldin**in (1577-1643) mukaan. Säännön keksi kuitenkin kreikkalainen **Pappos** jo 300-luvulla.

Hitausmomentti

Jos $A \subset \mathbb{R}^3$ edustaa kiinteää kappaletta ja S on avaruussuora, niin kappaleen hitausmomentti suoran S suhteen määritellään

$$I_S = \int_A r^2 \rho \, dxdydz = \int_A r^2 \, dm,$$

missä m on edellä määritelty massamitta (ρ = massatiheys) ja $r = r(x, y, z)$ on pisteen (x, y, z) etäisyys suorasta S. Hitaustiheys kappaleessa on siis $f = r^2\rho$.†

Jos suora kulkee pisteen $P_0 \mathrel{\widehat{=}} \vec{r}_0$ kautta ja sen suuntavektori on yksikkövektori $\vec{e}$, niin (vrt. kuvio)

$$r(x, y, z) = |\vec{e} \times (\vec{r} - \vec{r}_0)|.$$

Olkoon edellä origo kappaleen painopiste ja merkitään R = origon etäisyys suorasta S, jolloin

$$\vec{e} \times (\vec{r} - \vec{r}_0) = \vec{e} \times \vec{r} - R\vec{n}, \quad |\vec{n}\,| = 1$$

($\vec{n}$ = vakiovektori), ja näin ollen

$$r^2 \;=\; (\vec{e} \times \vec{r} - R\vec{n}) \cdot (\vec{e} \times \vec{r} - R\vec{n}) \;=\; |\vec{e} \times \vec{r}\,|^2 - 2R\,\vec{n} \cdot \vec{e} \times \vec{r} + R^2.$$

Kun sijoitetaan tämä em. integraalikaavaan, niin saadaan

$$I_S = \int_A |\vec{e} \times \vec{r}\,|^2 \, dm - 2R\vec{n} \cdot \vec{e} \times \int_A \vec{r}\, dm + R^2 \int_A dm.$$

Tässä ensimmäinen termi = kappaleen hitausmomentti origon kautta kulkevan, suoran S suuntaisen suoran S' suhteen, toinen termi = 0, koska origo on kappaleen painopiste, ja kolmas termi = mR^2, missä $m = m(A)$ on kappaleen massa. Hitausmomentille on näin saatu *Steinerin sääntönä* tunnettu palautuskaava

$$\boxed{I_S = I_{S'} + mR^2 \quad \text{(Steinerin sääntö).}}$$

Tässä siis S ja S' ovat yhdensuuntaiset suorat, S' kulkee kappaleen painopisteen kautta, R = suorien välinen etäisyys, ja m = kappaleen massa.

†Hitausmomentti määrittelee fysikaalisesti kappaleen pyörimishitauden pyörimisakselin S suhteen. Jos $\theta(t)$ on pyörimiskulma ajan hetkellä t, niin pyörimisliikkeen (Newtonin) yhtälö on $I_S\theta'' = M_S$, missä M_S on kappaleeseen vaikuttava momentti S:n suuntaan.

Esimerkki 8 Homogeenisen, R-säteisen teräskuulan massa $= m$. Laske kuulan hitausmomentti a) kuulan keskipisteen kautta kulkevan suoran S, b) kuulaa sivuavan suoran S' suhteen.

Ratkaisu a) Symmetriasyistä hitausmomentti on S:n suuntavektorista $\vec{e}$ riippumaton. Kun valitaan $\vec{e} = \vec{k}$ ja huomioidaan, että $m = \rho_0 \cdot \frac{4}{3}\pi R^3$, missä ρ_0 on kuulan massatiheys, niin (ks. edellisen luvun Esimerkki 6)

$$I_S = \int_A (x^2 + z^2)\rho_0 \, dxdydz = \rho_0 \cdot \frac{8}{15}\pi R^5 = \underline{\underline{\frac{2}{5} mR^2}}.$$

b) Origo on kappaleen painopiste, joten a-kohdan ja Steinerin säännön mukaan

$$I_{S'} = \frac{2}{5} mR^2 + mR^2 = \underline{\underline{\frac{7}{5} mR^2}}. \quad \square$$

*Hitaustensori

Tarkastellaan vielä kappaleen hitausmomentin I_S laskemista matriisialgebran keinoin. Jos suoran S suuntavektori (yksikkövektori) on $\vec{e} = e_1\vec{i} + e_2\vec{j} + e_3\vec{k}$, niin

$$\begin{aligned} \vec{e} \times \vec{r} &= \vec{e} \times (x\vec{i} + y\vec{j} + z\vec{k}) \\ &= (e_2 z - e_3 y)\vec{i} + (e_3 x - e_1 z)\vec{j} + (e_1 y - e_2 x)\vec{k}, \end{aligned}$$

joten

$$\begin{aligned} I_S = \int_A |\vec{e} \times \vec{r}|^2 \, dm &= \int_A \left[(e_2 z - e_3 y)^2 + (e_3 x - e_1 z)^2 + (e_1 y - e_2 x)^2 \right] dm \\ &= \sum_{i=1}^{3} \sum_{j=1}^{3} I_{ij} e_i e_j = \mathbf{e}^T \mathbf{I} \, \mathbf{e}, \end{aligned}$$

missä $\mathbf{e}^T = [e_1, e_2, e_3]$ ja $\mathbf{I} = (I_{ij})$ on symmetrinen nk. hitausmatriisi, jonka alkiot ovat

$$I_{11} = \int_A (y^2 + z^2) \, dm, \quad I_{22} = \int_A (x^2 + z^2) \, dm, \quad I_{33} = \int_A (x^2 + y^2) \, dm,$$

$$I_{12} = I_{21} = -\int_A xy \, dm, \quad I_{13} = I_{31} = -\int_A xz \, dm, \quad I_{23} = I_{32} = -\int_A yz \, dm.$$

Hitausmomentin määritelmän mukaan lävistäjäalkiot I_{ii} ovat hitausmomentteja koordinaattiakselien suhteen. Kun näissä alkioissa kirjoitetaan $y^2 + z^2 = -x^2 + (x^2 + y^2 + z^2)$ jne, niin nähdään, että hitausmomentille I_S pätee myös laskukaava

$$\boxed{I_S = J_0 - \mathbf{e}^T \mathbf{J} \, \mathbf{e},} \qquad (\star)$$

missä

$$J_0 = \int_A (x^2 + y^2 + z^2)\, dm, \quad \mathbf{J} = \int_A \begin{bmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{bmatrix} dm.$$

Matriisin $\mathbf{J}$ alkioita sanotaan <u>hitaustuloiksi</u>. Jos siis halutaan laskea kappaleen hitausmomentti annetun suoran S suhteen, niin ensin on laskettava jossakin koordinaatistossa hitaustulomatriisi $\mathbf{J}$ eli integraalit

$$J_{11} = \int_A x^2\, dm, \quad J_{12} = \int_A xy\, dm, \quad J_{13} = \int_A xz\, dm, \quad \text{jne}$$

(6 erilaista). Tämän jälkeen hitausmomentti määräytyy kaavasta ($\star$), missä $\mathbf{e}$ on suoran S suuntainen yksikkövektori ja $J_0 = J_{11} + J_{22} + J_{33}$.

Jos edellä yksikkövektori $\vec{e}$ ilmaistaan kierretyssä koordinaatistossa $\{\vec{i}', \vec{j}', \vec{k}'\}$ koordinaativektorina $\mathbf{e}'$, niin $\mathbf{e} = \mathbf{C}\mathbf{e}'$, missä $\mathbf{C}$ on ortogonaalinen matriisi (vrt. Luku X.6). Kaavassa ($\star$) tämä muunnos merkitsee, että $\mathbf{e}$:n tilalle tulee $\mathbf{e}'$ ja $\mathbf{J}$:n tilalle $\mathbf{C}^T\mathbf{J}\mathbf{C}$ (J_0 ei muutu). Hitaustulomatriisi, samoin hitausmatriisi, muuntuu siis koordinaatistoa kierrettäessä kuten (symmetrinen) tensori, vrt. Luku XII.4. Sanotaankin, että kyseessä on *hitaustensori*, jolloin <u>hitaus</u> (pyörimishitaus, vrt. alaviite edellä) tulee ymmärretyksi koordinaatiston kierrosta riippumattomana kappaleen ominaisuutena suhteessa valittuun koordinaatiston origoon (pyörimiskeskus). Päätellään edelleen, että koordinaatiston kierrolla löytyy aina *päähitauskoordinaatisto*, jossa hitausmatriisi on diagonaalinen (!). Nimittäin tämä löytyy ratkaisemalla hitausmatriisin ominaisarvo-ongelma (vrt. Luku XII.2). Ominaisarvoja, eli hitausmatriisin lävistäjäalkioita päähitauskoordinaatistossa, sanotaan *päähitausmomenteiksi.*

Esimerkki 9 Määritä homogeenisen kuution (sivun pituus a, massatiheys ρ_0, massa $m = \rho_0 a^3$) hitausmomentti lävistäjän suhteen.

<u>Ratkaisu</u> Olkoon $A = [-a/2, a/2] \times [-a/2, a/2] \times [-a/2, a/2]$, jolloin symmetrian perusteella on $I_{ij} = 0$ kun $i \neq j$ (päähitauskoordinaatisto). Päähitausmomentit ovat myös symmetrian perusteella samansuuruiset, eli jokaisella $i = 1, 2, 3$ on

$$I_{ii} = \rho_0 \int_A (y^2 + z^2)\, dxdydz = 2\rho_0 \int_A z^2\, dxdydx = 2\rho_0 a^2 \int_{-a/2}^{a/2} z^2\, dz = \frac{1}{6}\, ma^2.$$

Koska hitausmatriisi on diagonaalinen ja lävistäjäalkiot ovat samansuuruiset, niin hitausmatriisi on sama kaikissa koordinaatistoissa, joiden origo on kuution keskipisteessä (painopisteessä). Siis vastaus: $I_S = \underline{\underline{\frac{1}{6}ma^2}}$. $\square$

*Todennäköisyysmitta

Matematiikan lajissa nimeltä *todennäköisyyslaskenta* tarkastellaan nk. *satunnaismuuttujia* $\mathbf{x} \in \mathbb{R}^n$. Jokaiseen satunnaismuuttujaan liittyy ko. muuttujalle ominainen *todennäköisyysmitta* P, jonka tiheysfunktiota f ($f(\mathbf{x}) \geq 0\ \forall \mathbf{x}$) sanotaan *todennäköisyystiheydeksi* tai *-jakaumaksi.* Todennäköisyysmitalta edellytetään, että koko $\mathbb{R}^n$ on P-mitallinen ja

$$P(\mathbb{R}^n) = \int_{R^n} dP = \int_{\mathbb{R}^n} f\, d\mu = 1.$$

Joukon $A \subset \mathbb{R}^n$ todennäköisyysmittaa sanotaan A:n *todennäköisyydeksi.* Todennäköisyysmitan avulla määriteltyä $\mathbb{R}^n$:n painopistettä

$$\mathbf{x}_0 = \int_{\mathbb{R}^n} \mathbf{x}\, dP = \int_{R^n} \mathbf{x}\, f(\mathbf{x})\, d\mu$$

sanotaan (tarkasteltavan satunnaismuuttujan) *odotusarvoksi.*

Esimerkki 10 Olkoon $A \subset \mathbb{R}^n$ mitallinen ja $\mu(A) > 0$ ($\mu = \mathbb{R}^n$:n tilavuusmitta) ja olkoon satunnaismuuttujan $\mathbf{x}$ todennäköisyystiheys *tasajakauma*

$$f(\mathbf{x}) = \begin{cases} 1/\mu(A), & \text{jos } \mathbf{x} \in A, \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin on

$$P(B) = \int_B dP = \int_B f(\mathbf{x})\, d\mu = \frac{\mu(A \cap B)}{\mu(A)}, \quad B \subset \mathbb{R}^n.$$

Odotusarvo = A:n keskiö. □

HARJOITUSTEHTÄVIÄ

1. Pullapitkossa, joka sijaitsee x-akselilla välillä $[0, a]$, $a = 44$ cm, ovat rusinat jauhautuneet siten, että rusinoiden viivatiheys on

$$\rho(x) = 1.5\rho_0 \left[1 + \frac{x}{a}\left(1 - \frac{x}{a}\right)\right], \quad x \in [0, a], \quad \rho_0 = \frac{1}{\text{cm}}.$$

Montako (saman kokoista) rusinaa pullaan on pantu?

2. Pisteen $P_0 \mathrel{\widehat{=}} \vec{r}_0$ keskietäisyys joukon $A \subset R^2$ pisteistä voidaan määritellä integraalina $d = \frac{1}{\mu(A)} \int_A |\vec{r} - \vec{r}_0| \, d\mu$. Laske pisteen $(2,0)$ keskietäisyys neliöstä $A = [0,1] \times [0,1]$.

3. Kolmion muotoisen levyn pinnalla $A: \ x \geq 0 \ \ \& \ \ x + |y| \leq a$ vaikuttaa pinnan normaalin suuntainen paine $p = p_0(1 - x^2/a^2)$, $p_0 =$ vakio. Laske paineesta aiheutuva kokonaisvoima sekä momentti origon suhteen.

4. Rakennuksen nurkkauksessa on tetraedrin muotoinen hiekkakasa

$$A = \{(x,y,z) \in \mathbb{R}^3 \mid x,y,z \geq 0 \ \ \& \ \ x + y + z \leq a\}.$$

Kasassa massatiheys on

$$\rho(x,y,z) = \rho_0 + \frac{\rho_0}{30}\left(1 - \frac{x+y+z}{a}\right)^2,$$

missä ρ_0 (= vakio) on kuivan hiekan massatiheys ja loppuosa tiheydestä edustaa hiekkaan sitoutunutta vettä. Jos kasan koko massa on 100 kg, niin montako kiloa kasassa on vettä?

5. Kohdissa a)–k) määritä joukon A keskiö, muissa kohdissa kappaleen $A \subset \mathbb{R}^3$ painopiste, kun massatiheys ρ on annettu ($a > 0$, $\rho_0 =$ vakio).

a) $A \subset \mathbb{R}^2: \ x,y \geq 0 \ \ \& \ \ x + y \leq 1$
b) $A \subset \mathbb{R}^2: \ 0 \leq x \leq 1 \ \ \& \ \ -x^2 \leq y \leq x^3$
c) $A \subset \mathbb{R}^2: \ 0 \leq x \leq 1 \ \ \& \ \ x^m \leq y \leq \sqrt[m]{x} \ \ (m \in \mathbb{N}, \ m \geq 2)$
d) $A \subset \mathbb{R}^2: \ x = t - \sin t \ \ \& \ \ 0 \leq y \leq 1 - \cos t, \ \ t \in [0, 2\pi]$
e) $A \subset \mathbb{R}^2: \ x^2 + y^2 \leq a^2 \ \ \& \ \ 0 \leq y \leq x$
f) $A \subset \mathbb{R}^3: \ x,y,z \geq 0 \ \ \& \ \ x + 2x + 3y \leq 6$
g) $A \subset \mathbb{R}^3: \ 0 \leq x \leq 1 \ \ \& \ \ 0 \leq y \leq x^2 \ \ \& \ \ 0 \leq z \leq xy^2$
h) $A \subset \mathbb{R}^3: \ 0 \leq x,y,z \leq 1 \ \ \& \ \ x + y + z \leq 2$
i) $A \subset \mathbb{R}^3: \ 0 \leq x,y,z \leq 4 \ \ \& \ \ (x-1)^2 + (y-2)^2 + (z-3)^2 \geq 1$
j) $A \subset \mathbb{R}^3: \ x,y,z \geq 0 \ \ \& \ \ x^2 + y^2 + z^2 \leq a^2$
k) $A \subset \mathbb{R}^3: \ 0 \leq y \leq x \ \ \& \ \ z \geq 0 \ \ \& \ \ x^2 + y^2 + z^2 \leq a^2$
l) $A: \ x,y,z \geq 0 \ \ \& \ \ x + y + z \leq a, \ \ \rho = \rho_0 z/a$
m) $A: \ 0 \leq x \leq a \ \ \& \ \ a|y| \leq x^2 \ \ \& \ |z| \leq a, \ \ \rho = \rho_0 x/a$
n) $A: \ 0 \leq x,y,z \leq a, \ \ \rho = \rho_0(x^2 + y^2 + z^2)/a^2$
o) $A: \ 0 \leq x \leq a \ \ \& \ \ y^2 + z^2 \leq 4x^2, \ \ \rho = \rho_0 x/a$
p) $A: \ 0 \leq x \leq a\pi/2 \ \ \& \ \ y^2 + z^2 \leq a^2 \cos^2(x/a), \ \ \rho = \rho_0 (x/a)^2$
q) $A: \ x^2 + y^2 \leq a^2 \ \ \& \ \ 0 \leq z \leq 2a, \ \ \rho = \rho_0 z$
r) $A: \ x,y,z \geq 0 \ \ \& \ \ x^2 + y^2 + z^2 \leq a^2, \ \ \rho = \rho_0(x^2 + y^2)/a^2$
s) $A: \ x \geq 0 \ \ \& \ \ y^2 + z^2 \leq a^6/(a^2 + x^2)^2, \ \ \rho = \rho_0 x/a$
t) $A: \ x,y \geq 0 \ \ \& \ \ 0 \leq z \leq e^{-(x+y)/a}, \ \ \rho = \rho_0(x^2 + y^2)/a^2$

6. Kolmion muotoiseen patoluukkuun, jonka kärjet ovat pisteissä $A = (0, 0, 0)$, $B = (0, 0, -L)$ ja $C = (0, L, 0)$, vaikuttaa veden paine $\vec{p}(y, z) = \rho g z \vec{i}$. Luukkua ollaan juuri avaamassa (hallitusti), jolloin luukkuun vaikuttavat painekuorman lisäksi aukeamista hillitsevä momentti $\vec{M} = M\vec{k}$ ja saranoiden tukivoimat $\vec{F}_A = F_A\vec{i}$ ja $\vec{F}_B = F_B\vec{i}$. Laske painekuormituksen kokonaisvoima ja momentti saranan A suhteen ja näiden avulla edelleen (voima- ja momenttitasapainosta) tuntemattomat suureet M, F_A ja F_B.

7. Seuraavassa on annettu tasoalue A. Määritä Pappuksen sääntöä hyväksi käyttäen a-kohdassa A:n keskiö ja muissa kohdissa tilavuus $\mu(V)$, missä V on pyörähdyskappale, joka syntyy A:n pyörähtäessä x-akselin ympäri.
   a) $x^2 + y^2 \leq R^2, \;\; y \geq 0$.
   b) Kolmio, jonka kärjet ovat $(2, 1)$, $(1, 2)$ ja $(4, 2)$.
   c) Suunnikas, jonka kärjet ovat $(3, 2)$, $(1, 3)$, $(4, 6)$ ja $(6, 5)$.
   d) Käyrien $y = x^2$ ja $y = \sqrt{x}$ väliin jäävä alue välillä $x \in [0, 1]$.
   e) Käyrien $x = \sin y$ ja $x = -\sin y$ väliin jäävä alue välillä $y \in [0, \pi]$.
   f) Positiivisen y-akselin ja käyrän $x = e^{-y}$ väliin jäävä alue.

8. Koordinaattitasojen ja tasojen $z = a$ ja $x + y + z = 3a$ $(a > 0)$ rajaaman kappaleen hitausmomentti z-akselin suhteen on $I_z = kma^2$, missä $m =$ kappaleen massa. Laske kerroin k, kun kappaleen massatiheys on a) $\rho = \rho_0$, b) $\rho = \rho_0(x + y + z)/a$ $(\rho_0 =$ vakio$)$.

9. a) Homogeenista kappaletta rajaavat taso $z = a$ ja paraboloidi $x^2+y^2 = az$, $a > 0$. Laske lieriökoordinaateilla kertoimet k_x, k_y ja k_z kaavoissa $I_x = k_x ma^2$, $I_y = k_y ma^2$ ja $I_z = k_z ma^2$, missä I_x, I_y, I_z ovat kappaleen hitausmomentit koordinaattiakselien suhteen ja $m =$ kappaleen massa.

   b) Millä α:n arvolla $(\alpha \in \mathbb{R})$ homogeenisen kappaleen $A: \; 0 \leq z/a \leq (r/a)^\alpha$ (lieriökoordinaatit, $a > 0$) hitausmomentit kaikkien kolmen karteesisen koordinaatiakselin suhteen ovat samat?

   c) R-säteisen kuulan massatiheys pallokoordinaatistossa on $\rho(r) = \rho_0(r/R)^\alpha$ $(\rho_0 =$ vakio$)$ ja massa $= m$. Millä α:n arvolla kuulan hitausmomentti kuulan keskipisteen kautta kulkevan suoran suhteen on $I_S = \frac{1}{2}mR^2$?

10. Koivupuusta valmistettu kappale (homogeeninen, tiheys $= 500$ kg/m^3) on muodoltaan suorakulmainen särmiö, jonka särmien pituudet ovat 20, 30 ja 60 cm. Laske kappaleen hitausmomentit a) särmien, b) lävistäjien suhteen.

11. R-säteisen homogeenisen kuulan keskipiste $= (3a, -2a, 6a)$ ja massa $= m$. Määritä jokin päähitauskoordinaatisto ja päähitausmomentit, kun pyörimiskeskus = origo. *Vihje*: Ks. Harj.teht. X.6: 6a.

12. Jos x ja y ovat riippumattomia satunnaislukuja, niin todennäköisyys sille, että (x, y) osuu joukkoon $A \subset \mathbb{R}^2$ on $P(A) = \mu(A \cap B)$ missä μ on $\mathbb{R}^2$:n pinta-alamitta ja $B = [0, 1] \times [0, 1]$. Millä todennäköisyydellä on
a) $x^2 + y^2 < 1$, b) $x + y \geq 3/2$, c) $x + y = 1$?

13. (*) a) Koordinaattitasot ja taso $x + y + z = a > 0$ rajaavat homogeenisen kappaleen, jonka massa $= m$. Määritä kappaleen päähitauskoordinaatisto ja päähitausmomentit, kun pyörimiskeskus = origo.

    b) Kappaleen massa $= m$ ja painopiste on origossa. Näytä, että jos kappaleen hitaustulomatriisi origon suhteen $= \mathbf{J}$, niin hitaustulomatriisi pisteen $P = (x_0, y_0, z_0)$ suhteen on $\mathbf{J} + \mathbf{a}\mathbf{a}^T$, missä $\mathbf{a}^T = [x_0, y_0, z_0]$.

    c) Ratkaise a)-kohdan ongelma, kun pyörimiskeskus on painopiste.

14. (*) (Monumentti) Tuntemattoman matemaatikon muistomerkki on tehty umpiraudasta noudattaen seuraavia ohjeita koordinaatistossa, jossa positiivinen x-akseli osoittaa itään ja positiivinen y-akseli pohjoiseen. Maan pinnan taso on xy-taso ja pituusyksikkö on metri.

    1. Idästä katsoen monumentin profiili on suorien $z = 0$, $y = 2$ ja käyrän $z = y^2$ rajaama alue.
    2. Päältä katsoen monumentin profiili on vinoneliö, jonka kärjet ovat pisteissä $A = (-\frac{3}{2}, 0)$, $B = (1, 0)$, $C = (0, 2)$ ja $D = (\frac{5}{2}, 2)$.
    3. Monumentin jokainen pystysuora, itä-länsisuuntainen poikkileikkaus on kolmio, jonka kärki on maan pinnalla janalla BC ja tämän kärjen vastainen sivu on vaakasuora.

    a) Paljonko monumentti painaa? ($\rho = 7800$ kg/m^3)
    b) Määritä monumentin painopiste.
    c) Monumentti tuetaan kiinnittämällä kärki A maahan upotettuun betonipainoon. Kuinka suuri on tämän vastapainon massan vähintään oltava, jotta monumentti ei kaadu?

XIII.6 Viivaintegraalit

Tässä luvussa tarkastellaan integraaleja muotoa

$$I(f, S, \mu) = \int_S f \, d\mu,$$

missä $S \subset \mathbb{R}^2$ tai $S \subset \mathbb{R}^3$ on (tason tai avaruuden) *käyrä* ja μ on käyrään liitettävä *kaarenpituusmitta*. Tällaisia integraaleja sanotaan *viivaintegraaleiksi* (tai käyräintegraaleiksi). Kaarenpituusmitta pitkin suoraa on tavallinen Jordanin pituusmitta. Tasokäyrän (kaaren) $S : \ y = f(x) \ \ \& \ \ x \in [a, b]$ tapauksessa kaarenpituusmitta on määritelty aiemmin (Luku V.4) ja tulkittu myös integraalin avulla (Luku VIII.8). Seuraavassa yleistetään tämä käsite ja tutkitaan, miten viivaintegraaleja — ja itse kaarenpituusmittoja — käytännössä lasketaan.

Otetaan sama lähtökohta kuin Luvussa IV.4, eli oletetaan käyrä S parametrisoiduksi reaalimuuttujan t avulla välillä $[a, b]$. Tällöin esimerkiksi avaruuskäyrä on määritelty $\mathbb{R}^3$:n (tai E^3:n) pistejoukkona

$$S = \{(x, y, z) \mid x = x(t), \ y = y(t), \ z = z(t), \ t \in [a, b]\}.$$

Funktiot $x(t)$, $y(t)$, $z(t)$ (tasokäyrän tapauksessa $x(t)$, $y(t)$) oletetaan jatkossa ainakin jatkuviksi välillä $[a, b]$, jolloin käyrässä ei ole 'katkoksia'. Funktioiden säännöllisyysoletuksia lisätään jatkossa tarpeen mukaan.[†]

Jos käyrän parametrisointi $t \in [a, b] \mapsto \vec{r}(t)$ on injektio puoliavoimella välillä $[a, b)$, niin kyseessä on *yksinkertainen käyrä* eli *Jordan-käyrä*. Jos lisäksi $\vec{r}(a) = \vec{r}(b)$, niin käyrä on *suljettu*.

Jatkossa esitettävässä laskutekniikassa käyrän parametrisoinnilta edellytetään vain riittävä säännöllisyys. Käyrän ei siis tarvitse olla Jordan-käyrä vaan se voi olla itseään leikkaava tai jopa itsensä päälle kiertyvä.

[†]Huomautettakoon, että mikäli funktiot $x(t)$, $y(t)$ ja $z(t)$ ovat pelkästään jatkuvia, ei käyrän S 'ulkonäöstä' voi tehdä pitkälle meneviä johtopäätöksiä. Esimerkiksi on konstruoitavissa välillä $[0, 1]$ jatkuvat funktiot $x(t)$ ja $y(t)$ siten, että $S = \{(x(t), y(t)) \mid t \in [0, 1]\} = [0, 1] \times [0, 1]$ (!). Ensimmäisenä tällaisen 'täyteiskäyrän' kostruoi **G. Peano** vuonna 1890.

Esimerkki 1 Yksikköympyrä tai sen kaari voidaan parametrisoida muodossa

$$S = \{\, (x, y) \in \mathbb{R}^2 \mid x = \cos t,\ y = \sin t,\ t \in [a, b] \,\}.$$

Tämä on Jordan-käyrä (ympyräkaari), jos $b - a < 2\pi$, ja suljettu Jordan-käyrä, jos $b - a = 2\pi$. Yleisemmin kyseessä on parametrinen käyrä, jonka voi tulkita esim. edustavan tasaista ympyräliikettä (aikaparametrisointi). $\square$

Esimerkki 2 Ruuviviiva on avaruuden (ei-suljettu) Jordan-käyrä muotoa

$$S = \{\, (x, y, z) \in \mathbb{R}^3 \mid x = R\cos t,\ y = R\sin t,\ z = ct,\ t \in [a, b] \,\},$$

missä $R > 0$ ja $c \neq 0$. $\square$

Kaarenpituusmitta

Yleistetään aluksi kaarenpituusmitan käsite Luvusta V.4 parametriselle avaruuskäyrälle $S = \{(x(t), y(t), z(t)) \mid t \in [a, b]\}$. (Tasokäyrä tulee käsitellyksi erikoistapauksena $z(t) = 0$.) Olkoon $\{t_k,\ k = 0, \ldots n\}$ välin $[a, b]$ jako ($a = t_0 < t_1 < \ldots < t_n = b,\ n \in \mathbb{N}$), olkoon $\mathcal{P} = \{P_k = (x(t_k), y(t_k), z(t_k)),\ k = 0, \ldots, n\}$, ja olkoon edelleen $S_{\mathcal{P}}$ janoista $P_{k-1}P_k$, $k = 1 \ldots n$ koostuva murtoviiva. Tällöin jos kaarenpituusmitalle oletetaan perusaksioomat

- aksiooma 1: janan AB mitta $= |\overrightarrow{AB}|$,
- aksiooma 2: mitan additiivisuus murtoviivalla,

niin $S_{\mathcal{P}}$:n pituusmitta on

$$\mu(S_{\mathcal{P}}) = \sum_{k=1}^{n} |\vec{r}\,(t_k) - \vec{r}\,(t_{k-1})|.$$

Kuten aiemmin Luvussa V.4, käyrän S mitta määritellään tällöin

$$\mu(S) = \sup_{\mathcal{P}} \mu(S_{\mathcal{P}}).$$

Mitta on määritelty täsmälleen kun reaalilukujoukko $\{\mu(S_{\mathcal{P}})\}$ on rajoitettu, eli S on mitallinen eli *suoristuva* täsmälleen tällä ehdolla.

Ym. määritelmästä nähdään, että kaarenpituusmitta muistuttaa Jordan-mittoja sikäli, että perusaksioomissa tarvitaan ainoastaan yksinkertaisen perusjoukon (tässä janan) mitta ja additiivisuusperiaate. Jordan-mittojen konstruktiota muistuttaa myös kaarenpituuden määrittelyssä käytetty approksimaatioperiaate, jonka mukaan käyrää approksimoidaan perusjoukkojen äärellisillä yhdistelmillä.

Kaarenpituusparametrisaatio

Olkoon S suoristuva Jordan-käyrä. Tällöin jos $S(t)$, $t \in [a,b]$ on käyrän S osa, joka vastaa parametrin arvoja välillä $[a,t]$, niin kaarenpituusmitan määritelmästä ja kuvauksen $t \mapsto \vec{r}(t)$ oletetusta injektiivisyydestä seuraa

$$t_1, t_2 \in [a,b] \;\;\&\;\; t_1 < t_2 \;\Rightarrow\; \mu(S(t_2)) - \mu(S(t_1)) \geq |\vec{r}(t_2) - \vec{r}(t_1)| > 0.$$

Näin ollen jos merkitään

$$s(t) = \mu(S(t)),$$

niin $s(t)$ on välillä $[a,b]$ aidosti kasvava.

Siis jokaista $s \in [0, \mu(S)]$ vastaa yksikäsitteinen $t \in [a,b]$ siten, että $\mu(S(t)) = s$. Tämän kääntäen yksikäsitteisen vastaavuuden perusteella käyrä S voidaan yhtä hyvin parametrisoida kaarenpituuden s avulla, eli voidaan kirjoittaa

$$S = \{P \in E^d \mid P \mathrel{\widehat{=}} \vec{r} = \vec{r}(s),\; 0 \leq s \leq \mu(S)\}.$$

Tätä sanotaan *kaarenpituusparametrisaatioksi.* Sen laskeminen käytännössä on ongelma sinänsä, eikä siitä olekaan hyötyä viivaintegraaleja käytännössä laskettaessa. Sen sijaan käyräteorettisten tarkastelujen 'ajatteluparametrisaationa' se on luonnollinen, syystä että määritelmän perusteella pätee

$$|\vec{r}\,'(s)| = 1.$$

Tämän ominaisuuden vuoksi kaarenpituusparametriaatio on 'helpoin' parametrisaatio silloin kun halutaan tutkia käyrän (parametrisaatiosta riippumattomia) geometrisia ominaisuuksia, vrt. Lukujen VII.2 ja VII.3 tarkastelut. Kaarenpituusparametrisaatiota ajatellaan myös yleisessä kaarenpituusmitan merkinnässä

$$d\mu = ds \quad \text{(kaarenpituusmitta).}$$

Tämä merkintä omaksutaan jatkossa.

Viivaintegraalien laskutekniikka

Viivaintegraaleja käytännössä laskettaessa lähdetään käyrän tunnetuksi oletetusta parametrisaatiosta $t \mapsto \vec{r}(t)$ (joka harvoin on kaarenpituusparametrisaatio) ja pyritään tämän avulla 'oikomaan' integraali tavalliseksi määrätyksi integraaliksi yli välin $[a, b]$. Tämä tarkoittaa muunnosta

$$\int_S f\, ds = \int_{[a,b]} g\, d\mu',$$

missä $g(t) = f(x(t), y(t), z(t))$ ja μ' on kaarenpituusmitan vastine välillä $[a, b]$. Mitan μ' määrittämiseksi menetellään samalla tavoin kuin muuttujan vaihdossa (vrt. Luku XIII.4). Jos $\Delta S(t)$ on parametriväliä $[t, t+\Delta t] \subset [a, b]$ vastaava käyrän S osa, niin sikäli kun on määritettävissä *muuntosuhde*

$$J(t) = \lim_{\Delta t \to 0^+} \frac{\mu(\Delta S(t))}{\Delta t},$$

ja lisäksi $J(t)$ on riittävän säännöllinen (esim. jatkuva tai paloittain jatkuva), niin

$$\int_S f\, ds = \int_a^b g(t) J(t)\, dt.$$

Viivaintegraali on näin palautettu tavalliseksi määrätyksi integraaliksi Jordanmitan suhteen.

Muuntosuhteen $J(t)$ määräämiseksi oletetaan, että $x(t)$, $y(t)$ ja $z(t)$ ovat jatkuvasti derivoituvia välillä $[a, b]$. (Oletus on hiukan lievennettävissä, ks. kommentit jäljempänä). Olkoon ΔS käyrän osa välillä $[t, t + \Delta t] \subset [a, b]$ ($\Delta t > 0$) ja olkoon $\{t_j,\ j = 0 \ldots n\}$ välin $[t, t + \Delta t]$ jako, ts. $t = t_0 < t_1 < \ldots < t_n = t + \Delta t$. Tällöin Differentiaalilaskun väliarvolauseen mukaan

$$x(t_j) - x(t_{j-1}) = x'(\xi_j)(t_j - t_{j-1}), \quad \xi_j \in (t_{j-1}, t_j).$$

Koska x' on jatkuva välillä $[a, b]$ ja $\xi_j \in [t, t + \Delta t]$, niin tässä

$$x'(\xi_j) = x'(t) + o(1), \quad \text{kun } \Delta t \to 0.$$

[Tässä nojataan jatkuvuuden syvällisempään logiikkaan (tasaiseen jatkuvuuteen), vrt. vastaava päättely Luvussa XIII.4.] Arvioimalla samalla tavoin $y(t_j) - y(t_{j-1})$ ja $z(t_j) - z(t_{j-1})$ päätellään, että

$$\begin{aligned}&\sqrt{[x(t_j) - x(t_{j-1})]^2 + [y(t_j) - y(t_{j-1})]^2 + [z(t_j) - z(t_{j-1})]^2}\\ &\quad = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\,(t_j - t_{j-1}) + o(1)(t_j - t_{j-1}), \quad \text{kun } \Delta t \to 0.\end{aligned}$$

Summaamalla yli j:n ja käyttämällä kaarenpituuden määritelmää seuraa

$$\mu(\Delta S) = |\vec{r}\,'(t)|\Delta t + o(\Delta t), \quad \text{kun } \Delta t \to 0.$$

Muuntosuhteen määritelmän nojalla on $J(t) = |\vec{r}\,'(t)|$. Siis kaarenpituusmitan muunnoskaava avaruuskäyrälle on

$$ds = |\vec{r}\,'(t)|\,dt = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\,dt,$$

ja viivaintegraalille on saatu laskukaava

$$\boxed{\int_S f\,ds = \int_a^b g(t)|\vec{r}\,'(t)|\,dt, \quad g(t) = f(\vec{r}\,(t)).} \qquad (\star)$$

Kaava on pätevä hieman heikomminkin kuin oletetuin edellytyksin. Esimerkiksi seuraavat oletukset ovat yhdessä riittävät:

(i) $x(t)$, $y(t)$, $z(t)$ ovat jatkuvia välillä $[a, b]$.

(ii) $x'(t)$, $y'(t)$ $z'(t)$ ovat olemassa ja jatkuvia avoimilla väleillä (c_{i-1}, c_i), $i = 1 \ldots m$, missä $a = c_0 < c_1 < \ldots < c_m = b$ $(m \in \mathbb{N})$.

(iii) $g(t)|\vec{r}\,'(t)|$ on (laajennetussa mielessä) Riemann-integroituva välillä $[a, b]$.

Sovelluksissa ei näitä lievennyksiä pitemmälle ole juuri tarvetta mennä.

Jos kaavassa $(\star)$ asetetaan $f = 1$ $(\Rightarrow g = 1)$, on tuloksena kaarenpituuden laskukaava

$$\boxed{\mu(S) = \int_a^b |\vec{r}\,'(t)|\,dt.}$$

Huomattakoon, että jos käytetään kaarenpituusparametrisaatiota $(t = s)$, niin $a = 0$, $b = \mu(S)$ ja $|\vec{r}\,'(s)| = 1$, jolloin kaava saa muodon

$$\mu(S) = \int_0^{\mu(S)} ds = \Big|_0^{\mu(S)} s = \mu(S).$$

Tämä on luonnollisesti vain tautologia, josta ei ole käytännön hyötyä kaarenpituuden laskemisessa.

Esimerkki 3 Ohuesta, tasapaksusta ja homogeenisesta langasta valmistettu kierrejousi on avaruuskäyrä $x = R\cos t$, $y = R\sin t$, $z = Ht/(12\pi)$, $0 \le t \le 12\pi$. Laske a) jousen pituus, b) hitausmomentti suoran $x = R$, $y = 0$ suhteen, kun jousen massa $= m$.

Ratkaisu a)

$$\begin{aligned}
\mu(S) &= \int_0^{12\pi} \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\, dt \\
&= \int_0^{12\pi} \sqrt{R^2 + \left(\frac{H}{12\pi}\right)^2}\, dt \\
&= \underline{\underline{\sqrt{(12\pi R)^2 + H^2}}}.
\end{aligned}$$

b) Olkoon langan massatiheys pituusyksikköä kohti $= \rho_0$, jolloin a-kohdan mukaan $m = \rho_0\sqrt{(12\pi R)^2 + H^2}$. Hitausmomentti suoran $S:\ x = R,\ y = 0$ suhteen on

$$\begin{aligned}
I_S &= \int_S \left[(x-R)^2 + y^2\right]\ \rho_0\, ds \\
&= \int_0^{12\pi} \rho_0 R^2 \left[(1-\cos t)^2 + \sin^2 t\right] \sqrt{R^2 + \left(\frac{H}{12\pi}\right)^2}\, dt \\
&= \frac{mR^2}{12\pi} \int_0^{12\pi} (2 - 2\cos t)\, dt \\
&= \underline{\underline{2mR^2}}.
\end{aligned}$$

Tähän tulokseen olisi voitu tulla ilman integrointiakin: Koska langan koko massa on vakioetäisyydellä R z-akselista, niin hitausmomentti z-akselin suhteen on $I_z = mR^2$, jolloin Steinerin säännöstä (ks. edellinen luku) seuraa $I_S = mR^2 + mR^2 = 2mR^2$. $\square$

Tasokäyrät

Tasokäyrän tapauksessa on $|\vec{r}\,'(t)| = \sqrt{[x'(t)]^2 + [y'(t)]^2}$. Jos käyrä on annettu muodossa $y = f(x)$ ja valitaan $t = x$, niin saadaan (ennestään tuttu, vrt. Luku VIII.8) tulos $|\vec{r}\,'(x)| = \sqrt{1 + [f'(x)]^2}$. Jos taas käyrä on annettu polaarimuodossa $r = f(\varphi)$, niin on luonnollista valita parametriksi $t = \varphi$, jolloin

$$x'(\varphi) = \frac{d}{d\varphi}[f(\varphi)\cos\varphi] = f'(\varphi)\cos\varphi - f(\varphi)\sin\varphi,$$
$$y'(\varphi) = \frac{d}{d\varphi}[f(\varphi)\sin\varphi] = f'(\varphi)\sin\varphi + f(\varphi)\cos\varphi,$$

ja näin ollen

$$|\vec{r}\,'(\varphi)| = \sqrt{[x'(\varphi)]^2 + [y'(\varphi)]^2} = \sqrt{[f(\varphi)]^2 + [f'(\varphi)]^2}.$$

Laskettaessa viivaintegraaleja tasokäyrien yli voidaan siis käyrän esitysmuodosta riippuen käyttää mittamuunnoksia

$$\boxed{ds = \begin{cases} \sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt & (\,x = x(t),\ y = y(t)\,) \\ \sqrt{1 + [f'(x)]^2}\,dx & (\,y = f(x)\,) \\ \sqrt{[f(\varphi)]^2 + [f'(\varphi)]^2}\,d\varphi & (\,r = f(\varphi)\,) \end{cases}}$$

Esimerkki 4 Laske puoliympyrän kaaren $S:\ x^2 + y^2 = R^2,\ y \geq 0$ pituus $\mu(S)$ kolmella eri tavalla.

<u>Ratkaisu</u> (a) $x = R\cos t,\ y = R\sin t,\ t \in [0, \pi]$

$$\Rightarrow \mu(S) = \int_0^\pi \sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt = \int_0^\pi R\,dt = \underline{\underline{\pi R}}.$$

(b) $y = \sqrt{R^2 - x^2} = f(x),\ x \in [-R, R]$

$$\Rightarrow\ \mu(S) = \int_{-R}^{R} \sqrt{1 + [f'(x)]^2}\,dx = \int_{-R}^{R} \frac{R}{\sqrt{R^2 - x^2}}\,dx$$
$$= \int_{-\pi/2}^{\pi/2} R\,dt = \underline{\underline{\pi R}}. \qquad [\,\text{sijoitus } x = R\sin t\,]$$

(c) $r = R = f(\varphi),\ \varphi \in [0, \pi]$

$$\Rightarrow\ \mu(S) = \int_0^\pi \sqrt{[f(\varphi)]^2 + [f'(\varphi)]^2}\,d\varphi = \int_0^\pi R\,d\varphi = \underline{\underline{\pi R}}. \quad \square$$

Esimerkki 5 Logaritmisen spiraalin $r = e^{-\varphi}$ kaarenpituus välillä $\varphi \in [0, \infty)$ on

$$\begin{aligned}\mu(S) &= \int_0^\infty \sqrt{[f(\varphi)]^2 + [f'(\varphi)]^2}\, d\varphi \\ &= \int_0^\infty \sqrt{2}\, e^{-\varphi}\, d\varphi \;=\; \underline{\underline{\sqrt{2}}}. \quad \square\end{aligned}$$

Esimerkki 6 Laske viivaintegraali $\int_S f\, ds$, kun a) $f(x,y) = y$, b) $f(x,y) = x$ ja $S = \{\, (x,y) \in \mathbb{R}^2 \mid y = x^3/3 \;\;\&\;\; x \in [0,1] \,\}$.

<u>Ratkaisu</u>

a)
$$\begin{aligned}\int_S y\, ds &= \int_0^1 \frac{1}{3} x^3 \sqrt{1+x^4}\, dx \\ &= \Big|_0^1 \frac{1}{18}(1+x^4)^{3/2} = \underline{\underline{\frac{1}{18}(2\sqrt{2}-1)}}.\end{aligned}$$

b)
$$\begin{aligned}\int_S x\, ds &= \int_0^1 x\sqrt{1+x^4}\, dx \qquad [\,\text{sijoitus } x^2 = t,\; x\,dx = \tfrac{1}{2}\,dt\,] \\ &= \frac{1}{2}\int_0^1 \sqrt{1+t^2}\, dt \qquad [\,\text{sijoitus } t = \sinh u,\; dt = \cosh u\, du\,] \\ &= \frac{1}{2}\int_0^{\ln(\sqrt{2}+1)} \cosh^2 u\, du \\ &= \frac{1}{8}\int_0^{\ln(\sqrt{2}+1)} (e^{2u} + e^{-2u} + 2)\, du \\ &= \frac{1}{16}\Big|_0^{\ln(\sqrt{2}+1)} (e^{2u} - e^{-2u} + 4u) = \underline{\underline{\frac{1}{4}\big[\sqrt{2} + \ln(\sqrt{2}+1)\big]}}. \quad \square\end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Rautalanka, jonka pituus $= 10\pi$, on taivuteltu noudattamaan pisteestä $(3,0,0)$ lähtien a) ruuviviivaa $S: x = 3\cos t,\; y = 3\sin t,\; z = 4t,\; t \geq 0$, b) käyrää $S: x = 3\cos^2 t,\; y = 4\sin^2 t,\; z = 5\sin t\cos t,\; t \geq 0$. Laske langan toinen päätepiste.

2. Laske viivaintegraali $\int_S f\,ds$ annetuilla f ja S:
   a) $f(x,y) = xy, \quad S:\ x = 2\cos t,\ y = \sin t,\ t \in [0, \frac{\pi}{2}]$
   b) $f(r,\varphi) = r^2\varphi, \quad S:\ r = \varphi,\ \varphi \in [0, 4\pi]$ (polaarik.)
   c) $f(x,y,z) = x^2 z, \quad S:\ r = 1,\ z = \varphi,\ \varphi \in [0, 4\pi]$ (lieriök.)
   d) $f(x,y,z) = x^2 z^3, \quad S:\ r = 1,\ \theta = 2\pi/3,\ \varphi \in [0, 3\pi]$ (pallok.)

3. Käyrän kaaren $S : t \in [a,b] \mapsto \vec{r}(t)$ (matemaattinen) keskiö lasketaan kaavalla $\vec{r}_0 = \frac{1}{\mu(S)} \int_S \vec{r}\,ds$. Sovella kaavaa:
   a) $\vec{r} = R\cos t\vec{i} + R\sin t\vec{j},\ t \in [0, \pi]$ (puoliympyrä)
   b) $\vec{r} = R(t - \sin t)\vec{i} + R(1 - \cos t)\vec{j},\ t \in [0, 2\pi]$ (sykloidin kaari)
   c) $\vec{r} = R\cos^3 t\,\vec{i} + R\sin^3 t\,\vec{j},\ t \in [0, \pi/2]$ (asteroidin kaari)

4. Suoran langan pätepisteet ovat $(0,0)$ ja $(2a,0)$ ja langan massatiheys pituusyksikköä kohti on $\rho(x) = \rho_0(1 + x/a)$ (ρ_0 = vakio, $a > 0$). Lanka taivutetaan noudattamaan origosta lähtien ympyräviivaa $x^2 + y^2 = 2ay,\ x \geq 0$. Määritä taivutetun langan painopiste.

5. Lenkkeilijä lähtee pisteestä $(1,0)$ (pituusyksikkö = km) kiertämään puru-rataa $S : x^2 + y^2 = 1$ vastapäivään. Missä pisteessä lenkkeilijä on tunnin juostuaan, jos hänen juoksuvauhtinsa on $v(s) = v_0 e^{-0.05s}$, missä $v_0 = 12$ km/h ja s = juostu matka kilometreina?

6. (*) Esitä seuraavat käyrät kaarenpituusparametrisaation avulla.
   a) $x = R(t - \sin t),\ y = R(1 - \cos t),\ t \in [0, 2\pi]$
   b) $y = x^2,\ x \geq 0$

7. (*) (Laskiainen, 3. lasku) Näytä, että harjoitustehtävässä IX.3:8 lumilautailijan lyhin laskuaika pisteeseen $P = (a,0)$ $(a > 0)$ on $t_{min} = \sqrt{2\pi a/g}$.

8. (*) Lieriöt $L_1 : x^2 + y^2 = R^2$, $L_2 : (y - R)^2 + z^2 = R^2$ ja pallo $K : (x - R)^2 + y^2 + z^2 = R^2$ leikkaavat toisensa pitkin kolmea avaruuskäyrää. Näytä, että näiden leikkauskäyrien pituudet ovat
   a) $\mu(L_1 \cap L_2) = \int_0^{\pi/2} \sqrt{f(\sin t)}\,dt, \quad f(x) = [1 - x^2(1-x)^2]/(2x - x^2)$,
   b) $\mu(L_1 \cap K) = \mu(L_2 \cap K) = \int_{\pi/6}^{\pi/2} \sqrt{g(\sin t)}\,dt, \quad g(x) = (2x - x^2)/(2x - 1)$.

9. (*) (Köysi katolla) xy-tasolle on pystytetty pitkä rakennus, jonka vesikatto on pinnalla $z = 4 - x^2/2$ ja räystäät suorilla $x = \pm 2,\ z = 2$. Rakennuksen yli on heitetty (hyvin taipuisa) köysi siten, että köyden kumpikin pää ulottuu xy-tasolle. a) Mikä on tällaisen köyden minimipituus? b) Mikä on köyden minimipituus sillä lisäehdolla, että köysi tulee räystäiden yli pisteissä $(2,2,2)$ ja $(-2,-2,2)$? c) Millaisena tasokäyränä köysi näkyy b-kohdassa, kun sitä katsellaan kaukaa positiivisen z-akselin suunnasta?

XIII.7 Pintaintegraalit

Pintaintegraalilla tarkoitetaan integraalia muotoa

$$I(f, A, \mu) = \int_A f \, d\mu,$$

missä A on pinnan $S \subset \mathbb{R}^3$ osa ja μ *pinta-alamitta pinnalla* S. Jatkossa käytetään kaarevan pinnan pinta-alamitan yhteydessä merkintää $d\mu = dS$. Tämän mitan suhde tason (Jordanin) pinta-alamittaan on samantyyppinen kuin kaarenpituusmitan suhde $\mathbb{R}$:n pituusmittaan. Pintaintegraali on näin ajatellen tasointegraalin yleistys, ja etenkin sovelluksissa termiä saatetaan käyttää tasointegraaleistakin puhuttaessa (vrt. Luku XIII.5).

Kuten käyrän kaarenpituusmitta, voidaan myös kaarevan pinnan pinta-alamitta määritellä pintaa 'oikaisevien' approksimaatioiden kautta. Luonteva menettely on approksimoida pintaa <u>kolmion</u> <u>muotoisilla</u> <u>tasopinnoilla</u>. Kun kolmiot sijoitetaan verkoksi, jonka solmupisteet ovat pinnalla, niin verkkoa tihennettäessä tulee kolmioiden yhteenlasketun pinta-alan lähestyä A:n pinta-alamittaa $\mu(A)$. Jos tämä sovitaan mitallisuuden määritelmäksi, niin pinnan kolmioapproksimaatioita voi käyttää pinta-alan numeeriseen laskemiseen 'suoraan määritelmästä'. Tällaisilla approksimaatioilla on muutakin käyttöä, esimerkiksi esitettäessä pintoja graafisesti (tietokonegrafiikka) tai jopa konstruoitaessa todellisia pintoja (esim. kaarevat peilipinnat).

Suoran 'numeronmurskauksen' vaihtoehtona pinta-alaa ja pintaintegraaleja on jälleen mahdollista käsitellä myös differentiaalilaskennan keinoin, jolloin pintaintegraalille saadaan laskukaava tasointegraalina. Fubinin lauseen avulla tämä palautuu edelleen yksiulotteisiksi integraaleiksi, jotka suotuisissa (käytännössä kylläkin harvinaisissa) oloissa voidaan laskea suljetussa muodossa. Pintaintegraalin laskukaavaa johdettaessa on luonnollinen lähtökohta pinnan parametrisointi eli esittäminen *parametrisena pintana* (vrt. Luku IV.4). Jos parametrisointi esitetään vektorimuodossa

$$\vec{r} = \vec{r}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$$

ja oletetaan, että $(x(u,v), y(u,v), z(u,v)) \in A \;\Leftrightarrow\; (u,v) \in B \subset \mathbb{R}^2$, niin pintaintegraali on muunnettavissa muotoon

$$\int_A f\,dS = \int_B g\,d\mu', \quad g(u,v) = f(x(u,v), y(u,v), z(u,v)),$$

missä μ' on pinnan S pinta-alamitan vastine parametritasossa. Muunnos parametritasoon on jälleen muotoa

$$d\mu' = J\,dudv,$$

missä $J = J(u,v)$ on (yleensä laskettavissa oleva) muuntosuhde. Kun muuntosuhde tunnetaan, on laskukaava valmis:

$$\int_A f\,dS = \int_B gJ\,dudv.$$

Muuntosuhteen laskemiseksi tutkitaan suorakulmion

$$\Delta B = [u_0, u_0 + \Delta u] \times [v_0, v_0 + \Delta v] \subset B$$

muuntumista. Oletetaan funktiot $x(u,v), y(u,v), z(u,v)$ riittävän säännöllisiksi niin, että voidaan käyttää linearisoivaa approksimaatiota

$$\vec{r}(u,v) \approx \vec{r}(u_0,v_0) + (u-u_0)\frac{\partial \vec{r}}{\partial u}(u_0,v_0) + (v-v_0)\frac{\partial \vec{r}}{\partial v}(u_0,v_0).$$

Tämän mukaisesti ΔB kuvautuu pintaa pisteessä $P_0 \mathrel{\widehat{=}} \vec{r}(u_0,v_0)$ sivuavalle tangenttitasolle. Kuva on tämän tason suunnikas, jonka virittävät vektorit

$$\vec{v}_1 = \Delta u\,\partial_u \vec{r}, \quad \vec{v}_2 = \Delta v\,\partial_v \vec{r}.$$

Suunnikaan pinta-ala on $|\vec{v}_1 \times \vec{v}_2| = |\partial_u \vec{r} \times \partial_v \vec{r}||\Delta u||\Delta v|$, joten päätellään, että muuntosuhde on[†]

$$J = \left|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right|.$$

Jos pinta on annettu muodossa $z = f(x,y)$, eli $\vec{r} = \vec{r}(x,y) = x\vec{i} + y\vec{j} + f(x,y)\vec{k}$, niin

$$\partial_x \vec{r} \times \partial_y \vec{r} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = -f_x\vec{i} - f_y\vec{j} + \vec{k}$$

$$\Rightarrow\; J(x,y) = \sqrt{1 + [f_x(x,y)]^2 + [f_y(x,y)]^2}.$$

[†]Muuntosuhteen laskukaava on tarkemmin perusteltavissa olettaen, että funktioiden $x(u,v), y(u,v), z(u,v)$ osittaisderivaatat ovat jatkuvia perussuorakulmiossa $T \supset B$,

Koska tässä $\vec{n} = (-f_x\vec{i} - f_y\vec{j} + \vec{k})/J = n_x\vec{i} + n_y\vec{j} + n_z\vec{k}$ on pinnan yksikkönormaalivektori, niin tuloksen voi esittää myös muodossa $J(x,y) = 1/|n_z(x,y)|$.

R-säteisellä pallopinnalla on

$$\begin{aligned} &\vec{r} = \vec{r}(\theta,\varphi) = R\sin\theta\cos\varphi\,\vec{i} + R\sin\theta\sin\varphi\,\vec{j} + R\cos\theta\,\vec{k} \\ &\Rightarrow\ J(\theta,\varphi) = |\partial_\theta\vec{r}\times\partial_\varphi\vec{r}\,| = R^2\sin\theta. \end{aligned}$$

Yleisemmällä pyörähdyspinnalla, joka syntyy käyrän $K : y = f(x),\ x \in [a,b]$ pyörähtäessä x-akselin tai käyrän $K : y = f(z),\ z \in [a,b]$ z-akselin ympäri (ks. kuviot alla) voidaan muuntosuhde laskea vastaavaan tapaan (Harj.teht. 1). Tulokset koottuna:

Pinta	Muuttujat	Muuntosuhde
$\vec{r} = \vec{r}(u,v)$	u, v	$\lvert\partial_u\vec{r}\times\partial_v\vec{r}\,\rvert$
$z = f(x,y)$	x, y	$\sqrt{1+[f_x(x,y)]^2+[f_y(x,y)]^2}$
pallopinta $r = R$	θ, φ	$R^2\sin\theta$
pyörähdyspinta: $y = f(x) \hookrightarrow \curvearrowright\!\!\longrightarrow_x$	x, φ	$\lvert f(x)\rvert\sqrt{1+[f'(x)]^2}$
pyörähdyspinta: $z = f(r) \hookrightarrow \uparrow^z\!\circlearrowleft$	r, φ	$r\sqrt{1+[f'(r)]^2}$

Esimerkki 1 R-säteiselle pallopinnalle on jakautunut tasaisesti massa m. Laske hitausmomentti z-akselin suhteen.

Ratkaisu

$$\begin{aligned} I &= \int_A \rho_0(x^2+y^2)\,dS \\ &= \int_B \rho_0 R^2 \sin^2\theta \cdot R^2 \sin\theta\, d\theta d\varphi \qquad (\,B=[0,\pi]\times[0,2\pi]\,) \\ &= \rho_0 R^4 \int_0^{2\pi} d\varphi \cdot \int_0^{\pi} \sin^3\theta\, d\theta \\ &= 2\pi\rho_0 R^4 \Big|_{\theta=0}^{\theta=\pi} (-\cos\theta + \frac{1}{3}\cos^3\theta) \\ &= \frac{8\pi}{3}\rho_0 R^4. \end{aligned}$$

Koska

$$\begin{aligned} m = \int_A \rho_0\, dS &= \rho_0 R^2 \int_0^{2\pi} d\varphi \cdot \int_0^{\pi} \sin\theta\, d\theta \\ &= \rho_0 \cdot 4\pi R^2, \end{aligned}$$

niin $I = \underline{\underline{\frac{2}{3}mR^2}}$. $\square$

Esimerkki 2 Laske $\int_A f\,dS$, kun $f(x,y,z)=xy$ ja

$$A = \left\{ (x,y,z)\in\mathbb{R}^3 \mid z = x^2 + \frac{1}{2}y^2 \ \ \&\ \ (x,y)\in B \right\}, \quad B=[0,1]\times[0,1].$$

Ratkaisu

$$\begin{aligned} \int_A f\,dS &= \int_B xy\sqrt{1+4x^2+y^2}\,dxdy \\ &= \int_0^1 x\left(\int_0^1 y\sqrt{1+4x^2+y^2}\,dy\right)dx \\ &= \int_0^1 x\left[\Big|_{y=0}^{y=1} \frac{1}{3}(1+4x^2+y^2)^{3/2}\right]dx \\ &= \int_0^1 \left[\frac{1}{3}x(2+4x^2)^{3/2} - \frac{1}{3}x(1+4x^2)^{3/2}\right]dx \end{aligned}$$

$$= \Big|_0^1 \left[\frac{1}{60}(2+4x^2)^{5/2} - \frac{1}{60}(1+4x^2)^{5/2}\right]$$
$$= \underline{\underline{\frac{1}{60}\big(36\sqrt{6} - 25\sqrt{5} - 4\sqrt{2} + 1\big).}} \quad \square$$

Esimerkki 3 Katkaistun suoran ympyräkartion korkeus $= h$, pohjan säde $= R$ ja katkaisukohdassa säde $= r$. Laske vaipan ala $\mu(A)$.

<u>Ratkaisu</u> Tässä on kyseessä pyörähdyspinta, joka syntyy, kun jana

$$K:\ y = f(x) = r + (R-r)\frac{x}{h}, \quad x \in [0,h]$$

pyörähtää x-akselin ympäri. Parametreilla $(x,\varphi) \in B = [0,1] \times [0,2\pi]$ laskien saadaan

$$\begin{aligned}
\mu(A) &= \int_B f(x)\sqrt{1+[f'(x)]^2}\,dxd\varphi \\
&= 2\pi \int_0^h \sqrt{1+\left(\frac{R-r}{h}\right)^2}\,\left[r + (R-r)\frac{x}{h}\right]dx \\
&= \pi(R+r)h\sqrt{1+\left(\frac{R-r}{h}\right)^2} \\
&= \underline{\underline{\pi(R+r)\sqrt{h^2+(R-r)^2}.}} \quad \square
\end{aligned}$$

Esimerkki 4 Laske viivoitinpinnan

$$\vec{r}(u,v) = v[\cos(\omega u)\vec{i} + \sin(\omega u)\vec{j}\,] + u\vec{k}, \quad (u,v) \in [0,H] \times [0,L]$$

pinta-ala.

Ratkaisu

$$\partial_u \vec{r} \times \partial_v \vec{r} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -\omega v \sin(\omega u) & \omega v \cos(\omega u) & 1 \\ \cos(\omega u) & \sin(\omega u) & 0 \end{vmatrix}$$
$$= -\sin(\omega u)\vec{i} + \cos(\omega u)\vec{j} - \omega v \vec{k}$$
$$\Rightarrow \quad J = |\partial_u \vec{r} \times \partial_v \vec{r}| = \sqrt{1+\omega^2 v^2}.$$

$$\begin{aligned}
\mu(A) &= \int_0^H \left(\int_0^L \sqrt{1+\omega^2 v^2}\, dv \right) du \\
&= H \int_0^L \sqrt{1+\omega^2 v^2}\, dv \\
&\qquad [\text{ sijoitus } \ \omega v = \sinh t, \quad dv = \omega^{-1} \cosh t\, dt, \quad \alpha = \operatorname{arsinh}(\omega L) \ ] \\
&= H\omega^{-1} \int_0^\alpha \cosh^2 t\, dt \\
&= H\omega^{-1} \Big|_0^\alpha \left(\frac{1}{2} \sinh t \cosh t + \frac{1}{2} t \right) \\
&= \underline{\underline{\frac{1}{2} \left[\ln(\omega L + \sqrt{\omega^2 L^2 + 1}) + \omega L \sqrt{\omega^2 L^2 + 1} \right] H\omega^{-1}.}} \quad \square
\end{aligned}$$

Avaruuskulma

Olkoon $V \subset \mathbb{R}^3$ avaruuden osajoukko, käytännössä esim. kiinteä kappale, pinta tai pinnan osa. Määritellään pallokoordinaatistossa joukko $B \subset [0,\pi] \times [0,2\pi]$ ehdolla

$$B = \{ (\theta, \varphi) \mid P = (r, \theta, \varphi) \in V \text{ jollakin } r > 0 \}.$$

Tällöin V näkyy origosta suuntaan $\vec{e}_r(\theta,\varphi)$ täsmälleen kun $(\theta,\varphi) \in B$. Olkoon A yksikköpallon S osa, joka vastaa pallonpintakoordinaattien (θ,φ) joukkoa B.

Tällöin sanotaan, että V näkyy origosta *avaruuskulmassa* (engl. solid angle)

$$\Omega = \int_A dS = \int_B \sin\theta \, d\theta d\varphi.$$

Tämän suurin arvo (kun $A = S$) on 4π.

Esimerkki 5 Missä avaruuskulmassa lieriöpinta

$$A = \{(x, y, z) \in \mathbb{R}^2 \ \mid x^2 + y^2 = R^2, \ 0 \le z \le H\}$$

näkyy origosta?

Ratkaisu Kyseisellä pinnalla pallonpintakoordinaatit saavat arvot

$$(\theta, \varphi) \in B = [\theta_0, \pi/2] \times [0, 2\pi], \quad \cos\theta_0 = \frac{R}{\sqrt{R^2 + H^2}},$$

joten

$$\Omega = \int_B \sin\theta \, d\theta d\varphi = \int_0^{2\pi} d\varphi \cdot \int_{\theta_0}^{\pi/2} \sin\theta \, d\theta = 2\pi \cos\theta_0 = \underline{\underline{2\pi \frac{H}{\sqrt{R^2 + H^2}}}}. \quad \square$$

HARJOITUSTEHTÄVIÄ

1. Halutaan laskea pinta-alamitan muuntosuhde J pyörähdyspinnalle, joka syntyy, kun a) xy-tason käyrä $K : y = f(x)$ pyörähtää x-akselin ympäri, b) yz-tason käyrä $K : y = f(z)$ pyörähtää z-akselin ympäri. Laske muuntosuhde parametrisaatioille
   a) $y = f(x)\cos\varphi, \ \ z = f(x)\sin\varphi,$
   b) $x = r\cos\varphi, \ \ y = r\sin\varphi, \ \ z = f(r).$

2. Laske sen pinnan ala, jonka lieriö $L : \ x^2 + y^2 = a^2 \ (a > 0)$ leikkaa
   a) kartiosta $z^2 = x^2 + y^2$,
   b) satulapinnasta $az = xy$,
   c) paraboloidista $az = x^2 + y^2$.

3. Laske pinta-ala $\mu(A)$, kun A on
   a) joukon $V = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \le z \le \sqrt{x^2 + y^2}\}$ reunapinta ∂V,
   b) parametrinen pinta $x = 8u^2, \ y = v^2, \ z = 4uv, \ (u, v) \in [0, 1] \times [0, 3]$.

4. a) Laske puolipallon (pinnan) keskiö.
   b) Pinnalle $S:\ x^2+y^2=R^2,\ 0\le z\le H$ on jakautunut tasaisesti massa m. Laske lieriökoordinaateilla hitausmomentit koordinaattiakselien suhteen.

5. Olkoon $f(x)\ge 0$, kun $x\in[a,b]$ ja S pinta, joka syntyy, kun käyrä $K:\ y=f(x)\ \ \&\ \ x\in[a,b]$ pyörähtää x-akselin ympäri. Todista *Guldinin sääntö*: S:n pinta-ala $=$ K:n pituus kertaa K:n keskiön pyörähdyksessä kulkema matka.

6. Laske, missä avaruuskulmassa kohde näkyy origosta:
   a) R-säteinen kiekko tasolla $z=a>0$, keskipiste z-akselilla.
   b) R-säteinen kuula, jonka keskipisteen etäisyys origosta $=a\ge R$.
   c) Puolikartion rajaama joukko $A:\ z\ge\sqrt{x^2+y^2}+a,\ a\ge 0$.
   d) Pararaboloidi $z=x^2+y^2+a,\ a\in\mathbb{R}$.

7. (*) *Vivian'in ikkunaksi* sanotaan sitä pintaa A, jonka lieriö $x^2+y^2=Rx$ erottaa pallosta $x^2+y^2+z^2=R^2$. Laske ko. pinnan ala.

8. (*) Olkoon $A\subset\mathbb{R}^3$ lieriöiden
$$S_1:\ x^2+y^2=R^2,\quad S_2:\ y^2+z^2=R^2,\quad S_3:\ x^2+z^2=R^2$$
   sisään jäävä joukko. Laske A:n reunapinnan ∂A pinta-ala.

9. (*) (N:s ydinvoimala) Voimalaitoksen jäähdytystornin vaipan ulkopinnan parametriesitys on (vrt. Esimerkki IV.4: 14)
$$\begin{cases} x=a[(2-2v)\cos u-v\sin u],\\ y=a[v\cos u+(2-2v)\sin u],\\ z=3av,\end{cases}$$
   missä a=50 m ja $(u,v)\in B=[0,2\pi]\times[0,1]$. Vaippa on valmistettu betonista ja sen paksuus on 20 cm. Laske tarvittavan betonin määrälle likiarvo käyttäen pintaintegraalia.

10. (*) a) Levy
$$A=\{\,(x,y,z)\mid \frac{x^2}{a^2}+\frac{y^2}{b^2}\le 1,\ z=c\,\}\quad (a,b,c>0)$$
   näkyy origosta avaruuskulmassa Ω. Johda Ω:lle laskukaava muotoa $\Omega=\int_0^{2\pi}f(\varphi)\,d\varphi$. b) Laske, missä avaruuskulmassa avaruusneliö
$$K=\{\,(x,y,z)\mid 0\le x,y\le 1,\ z=1\,\}$$
   näkyy origosta.

Luku XIV

Gaussin ja Stokesin lauseet

Gaussin lause ja *Stokesin lause* ovat usean muuttujan differentiaali- ja integraalilaskun keskeisiä tuloksia. Lauseet tunnetaan enemmän käyttökelpoisuutensa kuin matemaattisen suuruutensa vuoksi, ja niihin viitataankin usein arkisissa yhteyksissä nimillä 'Gaussin kaava' ja 'Stokesin kaava'. Hieman yksinkertaistaen näissä kaavoissa on kyse integraalikaavan

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

yleistämisestä koskemaan useamman (käytännössä yleensä kahden tai kolmen) muuttujan vektoriarvoisia funktioita eli vektorikenttiä. Gaussin lauseeseen (kaavaan) vedotaan hyvin usein silloin, kun erilaiset fysiikan *säilymislait* halutaan kirjoittaa osittaisdifferentiaaliyhtälöitten muotoon. Myös Stokesin lauseella on tällaista käyttöä etenkin sähkömagnetiikassa.

Luvussa XIV.1 tarkastellaan ensin viivaintegraaleille sukua olevia *polkuintegraaleja*. Näillä on käyttöä Gaussin ja Stokesin lauseiden yhteydessä ja yleisemminkin fysikaalisten vektorikenttien sovelluksissa. Luvussa XIV.2 johdetaan Gaussin lause tasossa ja avaruudessa ja esitetään lauseen yleistetty muoto. Lähtökohtana ovat taso- ja avaruusintegraaleja koskevat *Greenin kaavat*. Luvussa XIV.3 tarkastellaan esimerkkien valossa Gaussin lauseen käyttöä, kun halutaan johtaa fysiikan osittaisdifferentiaaliyhtälöitä tai fysikaalisten vektorikenttien jatkuvuusehtoja materiaalirajapinnoilla.

Luvussa XIV.4 johdetaan Stokesin lause ensin tasoon rajoittuen. Yleistettäessä tulos koskemaan avaruuden pintoja tarvitaan pinnan *suunnistuvuuden* käsite. Luvussa XIV.5 ratkaistaan Stokesin lauseen avulla fysiikassa keskeinen vektorikenttiä koskeva kysymys: Millä ehdoilla pyörteetön kenttä on gradienttikenttä eli lausuttavissa skalaaripotentiaalin avulla?

XIV.1 Vektorikentät ja polkuintegraalit

Luvussa XIII.6 tarkasteltiin viivaintegraaleja, joissa tason tai avaruuden käyrän yli integroidaan kaarenpituusmitan suhteen. Tässä luvussa tarkastelun kohteena on toinen viivaintegraalien luokka, jolle käytetään jatkossa nimeä *polkuintegraalit* (engl. path integral, suom. myös *tieintegraali*). Polkuintegraaleille on ominaista, että integrointi käyrää pitkin tapahtuu tiettyyn suuntaan, siksi nimitys *polku*, jonka voi tulkita suunnatuksi käyräksi. Suunnan ohella polkuintegraaleille on tyypillistä, että integrointiin liittyvä mitta ei ole kaarenpituusmitta vaan muu yksiulotteinen mitta, joka on tapauskohtainen. Polun suunta vaikuttaa polkuintegraaliin niin, että jos vain suunta vaihtuu, eli polku pysyy muuten (käyränä) samana, niin polkuintegraalin arvo vaihtuu vastaluvukseen. Polun parametrisoinnin kautta tämä vastaa määrätyn integraalin vaihtosääntöä, ks. esimerkit jäljempänä.

Jatkossa rajoitutaan sellaisiin polkuintegraaleihin, jotka sovelluksissa liitetään yleensä fysikaalisiin vektorikenttiin (kuten voima-, sähkö- ja magneettikenttiin). Vektorikenttiin liittyvillä polkuintegraaleilla on jatkossa käyttöä myös Gaussin ja Stokesin lauseiden yhteydessä.[†]

Erotukseksi käyrästä polku merkitään jatkossa symbolilla p tai tarkemmin

$$p : A \to B,$$

jolloin merkintä kertoo sekä polun (käyrän) päätepisteet että polun suunnan. Luontevasti polun suunnan määrittää parametrisointi: suunta on joko parametrin kasvusuunta tai vastakkainen suunta. Jos polku on parametrisoitu välillä $t \in [a, b]$, niin polku voidaan merkitä tarkemmin kuten parametrinen käyrä:

$$p: \ t \in [a, b] \ \mapsto \ \vec{r}(t).$$

Parametrisoinnin ei tarvitse olla 1–1, joten polku (kuten parametrinen käyrä) voi leikata itsensä tai kiertyä itsensä päälle.

[†]Matemaattisissa teksteissä erilaisten viivaintegraalien nimet eivät ole täysin vakiintuneet. Esim. saatetaan puhua 'vektorikenttien viivantegraaleista', kun tarkoitetaan polkuintegraaleja tämän tekstin merkityksessä.

Jatkossa tarkastelun kohteena ovat tason tai avaruuden polut muotoa $t \in [a,b] \mapsto (x(t), y(t))$ tai $t \in [a,b] \mapsto (x(t), y(t), z(t))$. Funktioiden $x(t)$, $y(t)$, $z(t)$ oletetaan olevan joko jatkuvasti derivoituvia välillä $[a,b]$ tai toteuttavan vastaavat, hieman heikommat säännöllisyysehdot, vrt. Luku XIII.6.

Kuten aiemmin, voidaan laskea *polun pituus* integraalina

$$\mu(p) = \int_a^b |\vec{r}\,'(t)|\, dt.$$

Tässä on kuitenkin kyse jo ennestään tutusta viivaintegraalista, jossa mitta on kaarenpituusmitta eikä polun suunnalla ole väliä.

Polkuintegraaleja (jatkon kannalta myös merkittävimpiä) ovat

$$\int_p f\, dx, \quad \int_p f\, dy, \quad \int_p f\, dz, \tag{$\star$}$$

missä $f(x,y,z)$ (tasossa $f(x,y)$) on tunnettu funktio. Näihin integraaleihin liittyvä mitta on tavallinen (1-ulotteinen) Jordan-mitta. Jos tunnetaan polun parametrisointi välillä $t \in [a,b]$, niin esimerkiksi $\int_p f\, dx$ lasketaan yksinkertaisesti kirjoittamalla $dx = x'(t)dt$ (kuten muuttujan vaihdossa). Jos vielä oletetaan, että $t = a$ vastaa polun alkupistettä ja $t = b$ loppupistettä, niin saadaan laskukaava

$$\int_p f\, dx = \int_a^b f(x(t), y(t), z(t)\, x'(t)\, dt.$$

Parametrisoinnin vaihto vastaa tässä kaavassa (toista) muuttujan vaihtoa, joten parametrisointi ei vaikuta integraalin arvoon.

Esimerkki 1 Tason polku p kulkee pisteestä $(0,0)$ pisteeseen $(1,1)$ pitkin käyrää $y = x^2$ ja polku $-p$ pitkin samaa käyrää pisteesät $(1,1)$ pisteeseen $(0,0)$. Laske $\int_p xy\, dy$ ja $\int_{-p} xy\, dy$.

<u>Ratkaisu</u> Valitaan molemmissa integraaleissa parametriksi $t = x$:

$$\int_p xy\, dy = \int_0^1 x \cdot x^2 \cdot 2x\, dx = \int_0^1 2x^4\, dx = \underline{\underline{\frac{2}{5}}},$$
$$\int_{-p} xy\, dy = \int_1^0 2x^4\, dx = -\int_0^1 2x^4\, dx = \underline{\underline{-\frac{2}{5}}}. \quad \square$$

Esimerkin jälkimmäisessä integraalissa käytettiin määrätyn integraalin vaihtosääntöä. — Itse asiassa kun vaihtosääntö huomioidaan, niin määrätty integraali on itsekin tulkittavissa polkuintegraaliksi: $\int_a^b f(x)\, dx = \int_p f\, dx$, missä p on $\mathbb{R}$:n polku, jonka alkupiste on a ja loppupiste b (!).

Polkuintegraali $\int_p \vec{F} \cdot d\vec{r}$

Fysiikan sovelluksissa polkuintegraalit liittyvät usein vektorikenttiin. Tyypillinen esimerkki on voimakentässä $\vec{F}(x, y, z)$ liikkuva (pistemäinen) kappale, jonka liikerata tunnetaan jollakin aikavälillä $t \in [a, b]$. Tällöin voimakentän kappaleeseen tekemä työ lasketaan polkuintegraalina pitkin kappaleen kulkemaa liikerataa. Liikerata tulkitaan siis poluksi $p : t \mapsto \vec{r}(t)$, $t \in [a, b]$. Jos $\vec{F}$ = vakio, niin fysiikan lakien mukaan työ $= |\vec{F}| \cdot s$, missä s = polulla kuljettu matka voiman vaikutussuunnassa, eli

$$W = \vec{F} \cdot [\vec{r}(b) - \vec{r}(a)] \quad (\vec{F} \text{ vakio}).$$

Vakiovoimakentän tekemän työn kannalta polkua siis 'mittaa' vektori

$$\vec{\mu}(p) = \vec{r}(b) - \vec{r}(a).$$

Jos tämä tulkitaan polun (vektoriarvoiseksi) mitaksi, niin nähdään, että tämäkin mitta on additiivinen: Jos p_1 ja p_2 ovat p:n osapolkuja vastaten parametrin arvoja väleillä $[a, t_0]$ ja $[t_0, b]$, $t_0 \in (a, b)$, niin $\vec{\mu}(p) = \vec{\mu}(p_1) + \vec{\mu}(p_2)$, vrt. kuvio.

Entä jos voimakenttä $\vec{F}$ ei ole vakio, mutta on jatkuva? Tällöin menetellään niinkuin integraaleissa yleensä: Otetaan käyttöön välin $[a, b]$ jako $\{t_k,\ k = 0, \ldots, n\}$, $a = t_0 < t_1 < \ldots t_n = b$, jolloin polku p jakautuu peräkkäisiksi osapoluiksi Δp_k, $k = 1 \ldots n$ vastaten parametrin arvoja väleillä $[t_{k-1}, t_k]$. Koska funktio $t \mapsto \vec{r}(t)$ on jatkuva, niin osapolkujen Δp_k päätepisteet $\vec{r}(t_{k-1}) = \vec{r}_{k-1}$ ja $\vec{r}(t_k) = \vec{r}_k$ tulevat yhä lähemmäksi toisiaan jaon tihetessä. Tällöin voima $\vec{F}$ on osapolulla Δp_k likimain vakio (koska $\vec{F}$ oli jatkuva), joten voiman tekemä työ tällä osapolulla on likimain

$$\Delta W_k \approx \vec{F}(\vec{r}_{k-1}) \cdot \vec{\mu}(\Delta p_k), \quad \vec{\mu}(\Delta p_k) = \vec{r}_k - \vec{r}_{k-1}\,.$$

Kun jaon tiheysparametri $h = \max_k |\vec{r}_k - \vec{r}_{k-1}| \to 0$, saadaan voimakentän tekemälle kokonaistyölle integraalilauseke

$$W = \lim_{h \to 0} \sum_{k=1}^{n} \Delta W_k = \int_p \vec{F} \cdot d\vec{\mu}.$$

Tämä on siis tulkittava polkuintegraaliksi vektorimitan $\vec{\mu}$ suhteen (!). Integraali saa hieman konkreettisen muodon, kun käytetään merkintää $d\vec{\mu} = d\vec{r}$, jolloin työn integraalikaava siis on

$$\boxed{W = \int_p \vec{F} \cdot d\vec{r} \quad \text{(työintegraali).}}$$

Määritelmän mukaisesti W:lle saadaan likiarvoja summien avulla. Esimerkiksi

$$W \approx \sum_{k=1}^{n} \vec{F}(\vec{r}_{k-1}) \cdot (\vec{r}_k - \vec{r}_{k-1}).$$

Näin laskettaessa polusta p ei tarvitse tehdä voimakkaita säännöllisyysoletuksia. Esimerkiksi työ W voidaan laskea, vaikka p ei olisi suoristuva (ts. pituusmitta ei määritelty). Tämä johtuu siitä, että työintegraali mittaa vain siirtymää voiman vaikutussuunnassa, ei kaarenpituutta.

Jos oletetaan parametrisointi $p : t \in [a, b] \mapsto \vec{r}(t)$, niin työintegraalissa voidaan kirjoitta $d\vec{r} = \vec{r}\,'(t)dt$, jolloin saadaan laskukaava

$$W = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt.$$

Jos taas voimakenttä esitetään koordinaattimuodossa

$$\vec{F}(x, y, z) = F_1(x, y, z)\vec{i} + F_2(x, y, z)\vec{j} + F_3(x, y, z)\vec{k},$$

niin kirjoittamall $d\vec{r} = dx\,\vec{i} + dy\,\vec{j} + dz\,\vec{k}$ työintegraali purkautuu polkuintegraalien ($\star$) summaksi:

$$W = \int_p (F_1\, dx + F_2\, dy + F_3\, dz).$$

Riittävän säännöllisellä polulla työintegraalin voi ilmaista kolmannellakin tavalla, sillä

$$d\vec{r} = \frac{\vec{r}\,'(t)}{|\vec{r}\,'(t)|}|\vec{r}\,'(t)|dt = \vec{t}\, ds,$$

missä $\vec{t}$ on polun suuntainen yksikkötangenttivektori. Tämän mukaan siis työintegraali voidaan haluttaessa liittää myös kaarenpituusmittaan laskukaavalla

$$W = \int_p \vec{F} \cdot \vec{t}\, ds.$$

Kuten kaavan johdosta ilmenee, tässä on $\vec{t}\, ds = \vec{r}\,'(t)\, dt$, joten riippuvuus kaarenpituusmitasta on näennäinen.

Esimerkki 2 Määritä voimakentän

$$\vec{F} = (x^2 - y)\vec{i} - 2xy\vec{j}$$

tekemä työ kappaleen liikkuessa polulla

$$p: \ x(t) = 2\cos t, \ y(t) = \sin t, \ t \in [0, 2\pi].$$

Ratkaisu $\quad dx = -2\sin t\, dt, \quad dy = \cos t\, dt, \quad t \in [0, 2\pi]$

$$\begin{aligned} \Rightarrow \quad W &= \int_0^{2\pi} [\,(4\cos^2 t - \sin t)(-2\sin t) - (4\cos t \sin t)\cos t\,]\, dt \\ &= \int_0^{2\pi} (-12\cos^2 t \sin t + 2\sin^2 t)\, dt \\ &= \Big|_0^{2\pi} (4\cos^3 t - \cos t \sin t + t) = \underline{\underline{2\pi}}. \quad \square \end{aligned}$$

Gradienttikenttä ja työintegraali

Jos voimakenttä $\vec{F}$ on gradienttikenttä, eli lausuttavissa skalaaripotentiaalin u avulla muodossa

$$\vec{F} = -\nabla u,$$

niin työintegraali voidaan laskea hyvin yksinkertaisesti. Nimittäin tässä tapauksessa on derivoinnin ketjusäännön (Luku XI.2) perusteella

$$\vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t) = -\nabla u(\vec{r}(t)) \cdot \vec{r}\,'(t) = -\frac{d}{dt} u(\vec{r}(t)),$$

joten

$$W = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt = -\Big|_a^b u(\vec{r}(t)) = u(\vec{r}(a)) - u(\vec{r}(b)).$$

Siis gradienttikentän työintegraali määräytyy pelkästään polun päätepisteistä:

Gradienttikentän tekemä työ = potentiaaliero polun alku- ja loppupisteiden välillä.

Esimerkki 2 (jatko). Esimerkissä polun alku- ja loppupisteet ovat samat. Koska $W \neq 0$, ei esimerkin kenttä $\vec{F}$ ole gradienttikenttä. Sen sijaan jos esimerkiksi

$$\vec{F} = (x^2 - y^2)\vec{i} - 2xy\vec{j},$$

niin ilman enempää laskemista selviää, että $W = 0$, sillä

$$\vec{F} = -\nabla(-\frac{1}{3}x^3 + xy^2). \quad \square$$

Vektoriarvoiset polkuintegraalit

Fysiikan sovelluksissa (esimerkiksi sähkömagnetiikassa) esiintyy myös vektoriarvoisia polkuintegraaleja muotoa

$$\int_p f\,d\vec{r} \quad \text{tai} \quad \int_p \vec{F} \times d\vec{r}.$$

Nämä voidaan laskea parametrisoinnin avulla samalla periaatteella kuin työintegraalikin.

Esimerkki 2 (jatko) Laske $\int_p x\,d\vec{r}$ ja $\int_p \vec{r} \times d\vec{r}$, kun p on esimerkin polku.

Ratkaisu

$$\begin{aligned}
\int_p x\,d\vec{r} &= \int_0^{2\pi} x(t)[x'(t)\vec{i} + y'(t)\vec{j}\,]\,dt \\
&= \int_0^{2\pi} 2\cos t\,(-2\sin t\,\vec{i} + \cos t\,\vec{j}\,)\,dt \\
&= -\vec{i}\int_0^{2\pi} 4\cos t\sin t\,dt + \vec{j}\int_0^{2\pi} 2\cos^2 t\,dt \\
&= -\vec{i}\,\Big|_0^{2\pi} 2\sin^2 t + \vec{j}\,\Big|_0^{2\pi} (t + \cos t\sin t) = \underline{\underline{2\pi\vec{j}}}, \\
\int_p \vec{r} \times d\vec{r} &= \int_0^{2\pi} [x(t)\vec{i} + y(t)\vec{j}\,] \times [x'(t)\vec{i} + y'(t)\vec{j}\,]\,dt \\
&= \int_0^{2\pi} (2\cos t\,\vec{i} + \sin t\,\vec{j}\,) \times (-2\sin t\,\vec{i} + \cos t\,\vec{j}\,)\,dt \\
&= \int_0^{2\pi} (2\cos^2 t + 2\sin^2 t)\vec{k}\,dt = \vec{k}\int_0^{2\pi} 2\,dt = \underline{\underline{4\pi\vec{k}}}. \quad \square
\end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske polkuintegraali $\int_p (9x^2y\,dx - 11xy^2\,dy)$, kun polku p kulkee origosta pisteeseen $(1,1)$ a) pitkin käyrää $x(t) = t^2,\ y(t) = t^3$, b) pitkin suoraa, c) pitkin käyrää $\vec{r}(t) = t\vec{i} + t^\alpha\vec{j},\ \alpha > 0$.

2. Laske polkuintegraali $\int_p (xdy - ydx)$, kun polku p kulkee pisteestä $(1,0)$ pitkin logaritmista spiraalia $r = e^{-\varphi}$ origoon.

3. Laske polkuintegraali $\int_p [(y-x)\,dx + xy\,dy]$, kun polku polku p on määritelty seuraavasti:

   a) Pisteestä $(1,0)$ pisteeseen $(-1,0)$ yksikköympyrää pitkin vastapäivään

   b) Pisteestä $(1,0)$ pisteeseen $(-1,0)$ yksikköympyrää pitkin myötäpäivään

   c) Murtoviiva $ABCD$, missä $A = (1,0)$, $B = (1,1)$, $C = (-1,1)$ ja $D = (-1,0)$

   d) Pisteestä $(1,0)$ yksikköympyrää pitkin takaisin lähtöpisteeseen vastapäivään kiertäen

   e) Origosta pitkin x-akselia pisteeseen $(\pi,0)$ ja takaisin origoon pitkin käyrää $y = \sin x$

   f) Pisteestä $(a,0)$ takaisin lähtöpisteeseen kiertäen vastapäivään ellipsiä $x^2/a^2 + y^2/b^2 = 1$ $(a,b>0)$.

4. Laske seuraavat polkuintegraalit.

   a) $\int_p \vec{F}\cdot d\vec{r}$, kun $\vec{F}(x,y,z) = \sqrt{y}\,\vec{i} + 2x\vec{j} + 3y\vec{k}$ ja polku kulkee origosta pisteeseen $(3,9,27)$ pitkin käyrää $\vec{r}(t) = t\vec{i} + t^2\vec{j} + t^3\vec{k}$.

   b) $\int_p \vec{F}\cdot d\vec{r}$, kun $\vec{F}(x,y,z) = x^3\vec{i} + y^2\vec{j} + z\vec{k}$ ja polku kulkee origosta pisteeseen $(1,1,2)$ pitkin käyrää $S:\ x = y,\ z = x^2 + y^2$.

   c) $\int_p \vec{F}\times d\vec{r}$, kun $\vec{F}(x,y,z) = xyz\vec{i} + y^2\vec{k}$ ja polku p seuraa tason $x = y$ ja pinnan $z = x^2$ leikkauskäyrää origosta pisteeseen $(2,2,4)$.

5. Polun p alkupiste on $(-1,1,-1)$ ja loppupiste $(1,2,3)$. Laske näillä tiedoilla seuraavat polkuintegraalit (työintegraalit) kirjoittamalla integraalit ensin muotoon $\int_p \nabla u\cdot d\vec{r}$.

   a) $\int_p (yz\,dx + zx\,dy + xy\,dz)$ b) $\int_p (yz^2\,dx + xz^2\,dy + 2xyz\,dz)$

   c) $\int_p [e^x y\,dx + (e^x + z^2)\,dy + 2yz\,dz]$ d) $\int_p \sin\frac{\pi(x+y+z)}{6}\,(dx + dy + dz)$

6. Määritellään $g(u,v) = \int_p (y\,dx + 2x\,dy)$, missä p kulkee origosta pisteeseen (u,v) suoraa pitkin. Mikä on g:n maksimiarvo yksikköympyrällä?

7. (*) Kappale, jonka massa $= m$, liikkuu aikavälillä $[t_1,t_2]$ pitkin polkua $p: t \mapsto \vec{r}(t)$ pisteestä P_1 pisteeseen P_2. Liikkeen aikana kappaleeseen vaikuttaa voimakenttä $\vec{F}$. Johda liikeyhtälöstä $m\vec{r}\,'' = \vec{F}$ energiaperiaate

$$\frac{1}{2}m(v_2^2 - v_1^2) = \int_p \vec{F}\cdot d\vec{r}, \quad \text{missä } v_i = |\vec{r}\,'(t_i)|,\ i = 1,2.$$

XIV.2 Gaussin lause

Lähdetään tarkastelemaan integraalikaavan

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

yleistämistä, ensin yhdessä dimensiossa. Olkoon $A \subset \mathbb{R}$ äärellinen yhdistelmä pistevieraita, suljettuja välejä:

$$A = \bigcup_{i=1}^{n} A_i, \quad A_i = [a_i, b_i], \quad A_i \cap A_j = \emptyset,\ i \neq j.$$

Tällöin ym. kaava voidaan kirjoittaa muotoon

$$\begin{aligned}\int_A f'\,dx &= \sum_{i=1}^{n} [f(b_i) - f(a_i)] \\ &= \sum_{x \in \partial A} \omega(x) f(x),\end{aligned}$$

missä ∂A on A:n reuna (koostuu pisteistä a_i, b_i, $i = 1 \ldots n$) ja $w(x)$ saa arvoja ± 1 seuraavan säännön mukaan:

$$\omega(x) = \begin{cases} +1, & \text{jos } (x-\delta, x) \subset A \text{ jollakin } \delta > 0, \\ -1, & \text{jos } (x, x+\delta) \subset A \text{ jollakin } \delta > 0. \end{cases}$$

$\omega(x) = +1$ $\qquad$ $\omega(x) = -1$

Siirrytään nyt kahteen dimensioon. Olkoon $f = f(x, y)$ määritelty ja jatkuvasti derivoituva joukossa $A \subset \mathbb{R}^2$ (josta tehdään hetimiten yksinkertaistavia olettamuksia) ja tarkastellaan integraalia

$$\int_A \frac{\partial f}{\partial x}\,dxdy.$$

Jatkossa oletetaan, että joukko A on suljettu ja jaettavissa äärellisen moneen yksinkertaista muotoa olevaan osaan A_i siten, että osat koskettavat toisiaan enintään reunoillaan, ts.

$$A = \bigcup_{i=1}^{n} A_i, \quad A_i \cap A_j = \partial A_i \cap \partial A_j,\ i \neq j.$$

Osat A_i oletetaan edelleen kaikki y-projisoituviksi. Tarkemmin sanoen, oletetaan, että jokainen A_i on muotoa

$$A_i = \{\, (x,y) \in \mathbb{R}^2 \mid a_i(y) \le x \le b_i(y) \;\; \& \;\; y \in [c_i, d_i] \,\},$$

missä funktiot $a_i(y)$, $b_i(y)$ ovat välillä $[c_i, d_i]$ jatkuvia.

Joukkoa A koskeva oletus siis on, että A:n ositus ym. tavalla on mahdollinen. Ajatellen sovelluksissa kohdattavia 'käytännön joukkoja' ei oletus ole kovin rajoittava, vrt. kuvio.

Kun A:n ositus em. tavalla on tehty, voidaan tarkastelun kohteena oleva integraali purkaa osiin additiivisuusperiaatteella (sillä $\mu(A_i \cap A_j) = 0$, $i \neq j$) :

$$\int_A \frac{\partial f}{\partial x}\, dxdy = \sum_{i=1}^{n} \int_{A_i} \frac{\partial f}{\partial x}\, dxdy.$$

Tässä on Fubinin lauseen mukaan

$$\int_{A_i} \frac{\partial f}{\partial x}\, dxdy = \int_{c_i}^{d_i} \left(\int_{a_i(y)}^{b_i(y)} \frac{\partial f}{\partial x}\, dx \right) dy$$
$$= \int_{c_i}^{d_i} [f(b_i(y), y) - f(a_i(y), y)]\, dy = \int_{S_i} \omega_i(x,y) f(x,y)\, dy,$$

missä $S_i \subset \partial A_i$ ja funktio w_i määritellään kuten kuviossa.

Kun saatu tulos

$$\int_{A_i} \frac{\partial f}{\partial x}\, dxdy = \int_{S_i} \omega_i(x,y) f(x,y)\, dy$$

summataan yli i:n ja huomataan, että

$$i \neq j \;\; \& \;\; (x,y) \in S_i \cap S_j \;\Rightarrow\; \omega_i(x,y) + \omega_j(x,y) = 0,$$

niin nähdään, että oikealla puolella viivaintegraalit yli osajoukkojen A_i yhteisten (eli A:n sisään jäävien) reunaviivojen kuomoutuvat. Käyttäen kummallakin puolella integraalin additiivisuusperiaatetta saadaan tulos näin ollen muotoon

$$\int_A \frac{\partial f}{\partial x}\, dxdy = \int_{\partial A} \omega^{(x)} f\, dy,$$

missä $\omega^{(x)}(x,y) = 0$ x-akselin suuntaisilla reunaviivan ∂A osilla ja muuten $\omega^{(x)}(x,y)$ saa arvoja ± 1 kuvion mukaisesti.

Olettaen, että A on jaettavissa samalla tavoin x-projisoituviin osiin, saadaan vastaavasti integraalikaava

$$\int_A \frac{\partial f}{\partial y}\, dxdy = \int_{\partial A} \omega^{(y)} f\, dx,$$

missä $\omega^{(y)}(x,y) = 0$ y-akselin suuntaisilla reunaviivan ∂A osilla ja muuten $\omega^{(y)}(x,y)$ saa arvoja ± 1 kuvion mukaisesti.

Huomattakoon, että saaduissa integraalikaavoissa reunaviivan ∂A yli laskettavat integraalit ovat polkuintegraaleja (vrt. edellinen luku). Nämä purkautuvat äärellisiksi summiksi, joissa kukin termi on määrätty integraali yli suljetun välin, mitan ollessa tavallinen $\mathbb{R}$:n pituusmitta. Kaavoissa ei siis edellytetä edes reunaviivan ∂A suoristuvuutta. Olettamalla reunaviivalle lisää säännöllisyyttä saadaan tulokset kuitenkin helpommin muistettavaan muotoon.

Em. ositukset jakavat reunaviivan ∂A osiin, jotka ovat joko muotoa

$$S_i = \{\, (x,y) \in \mathbb{R}^2 \mid x = f_i(y) \;\;\&\;\; y \in [a_i, b_i] \,\}$$

tai muotoa

$$S_i = \{\, (x,y) \in \mathbb{R}^2 \mid y = g_i(x) \;\;\&\;\; x \in [c_i, d_i] \,\}.$$

Oletetaan tässä, että funktiot f_i ovat jatkuvia suljetuilla väleillä $[a_i, b_i]$ ja että derivaatat f_i' ovat jatkuvia avoimilla väleillä (a_i, b_i). Vastaavasti oletetaan, että funktiot g_i ovat jatkuvia suljetuilla väleillä $[c_i, d_i]$ ja derivaatat g_i' jatkuvia avoimilla väleillä (c_i, d_i). (Nämäkään lisäoletukset eivät ole käytännön kannalta kovin rajoittava, vrt. osituskuvio edellä.) Jatkossa sanottakoon joukkoa A, joka totetuttaa kaikki tehdyt oletukset *perusalueeksi*. Perusalue on siis joukko, joka on jaettavissa äärellisen moneen x-projisoituvaan osaan, samoin äärellisen moneen y-projisoituvaan osaan siten, että näiden osien reunaviivalla ∂A sijaitsevat reunakäyrät toteuttavat tehdyt säännöllisyysoletukset. Perusalue on kompakti joukko (vrt. Luku XI.1).

Jos A on perusalue, niin reunalla ∂A on, erillisiä pisteitä lukuunottamatta, määritelty reunan *ulkonormaali*, eli reunaa vastaan kohtisuora, joukosta A poispäin osoittava yksikkövektori, jota merkittäköön

$$\vec{n}(x,y) = n_x(x,y)\vec{i} + n_y(x,y)\vec{j}, \quad |\vec{n}| = 1.$$

Niissä pisteissä $(x_i, y_i) \in \partial A$, joissa ulkonormaali $\vec{n}$ on määritelty, on funktion $\omega^{(x)}$ määritelmän mukaisesti

$$\omega^{(x)}(x,y) = \begin{cases} +1, & \text{jos } n_x(x,y) > 0, \\ 0, & \text{jos } n_x(x,y) = 0, \\ -1, & \text{jos } n_x(x,y) < 0, \end{cases}$$

ja $\omega^{(y)}(x,y)$ määräytyy vastaavasti $n_y(x,y)$:n perusteella. Kun nämä yhteydet otetaan lukuun, niin voidaan kirjoittaa (vrt. kuvio),

$$\omega^{(x)} dy = n_x\, ds, \quad \omega^{(y)} dx = n_y\, ds,$$

$$|n_x| = \sin\theta = \frac{dy}{ds}$$
$$|n_y| = \cos\theta = \frac{dx}{ds}$$

jolloin saadaan helposti muistettavat *Greenin tasokaavat*

$$\boxed{\begin{aligned} \int_A \frac{\partial f}{\partial x}\, dxdy &= \int_{\partial A} n_x f\, ds, \\ \int_A \frac{\partial f}{\partial y}\, dxdy &= \int_{\partial A} n_y f\, ds. \end{aligned} \qquad \text{(Greenin tasokaavat)}}$$

Korostettakoon vielä, että näissä kaavoissa kaarenpituusmitta on otettu käyttöön vain muistisäännön vuoksi. Todellisuudessa integraalit kaavojen oikealla puolella ovat polkuintegraaleja.

Tarkastellaan seuraavaksi A:ssa määriteltyä vektorikenttää

$$\vec{F}(x,y) = F_1(x,y)\vec{i} + F_2(x,y)\vec{j},$$

joka olkoon jatkuvasti derivoituva. Soveltaen Greenin tasokaavoja saadaan

$$\begin{aligned}\int_A \nabla \cdot \vec{F}\, dxdy &= \int_A \frac{\partial F_1}{\partial x}\, dxdy + \int_A \frac{\partial F_2}{\partial y}\, dxdy \\ &= \int_{\partial A} (n_x F_1 + n_y F_2)\, ds = \int_{\partial A} \vec{n} \cdot \vec{F}\, ds.\end{aligned}$$

Näin on johdettu

Lause XIV.2.1 **(Gaussin lause tasossa)** Jos $A \subset \mathbb{R}^2$ on perusalue ja $\vec{F}$ on A:ssa määritelty, jatkuvasti derivoituva vektorikenttä, niin pätee

$$\boxed{\int_A \nabla \cdot \vec{F}\, dxdy = \int_{\partial A} \vec{n} \cdot \vec{F}\, ds.}$$

Lauseen XIV.2.1 laskukaavalla, jota sanotaan jatkossa *Gaussin tasokaavaksi*, on vastine myös kolmessa (ja useammassakin) dimensiossa. Olkoon $V \subset \mathbb{R}^3$ ja oletetaan, että V on jaettavissa äärellisen moneen xy-projisoituvaan, yz-projisoituvaan, ja xz-projisoituvaan osaan samaan tapaan kuin edellä. Esimerkiksi xy-projisoituvat osat V_i ovat tällöin muotoa

$$V_i = \{\, (x,y,z) \in \mathbb{R}^3 \mid (x,y) \in B_i \ \ \&\ \ a_i(x,y) \le z \le b_i(x,y) \,\}.$$

Tässä joukot $B_i \subset \mathbb{R}^2$ oletetaan edelleen tason perusalueiksi, ja lisäksi oletetaan, että funktiot a_i ja b_i ovat B_i:ssa jatkuvia ja B_i:n sisäpisteissä jatkuvasti derivoituvia (osittaisderivaatat jatkuvia). Joukkoa V, joka toteuttaa nämä ja vastaavat oletukset koskien yz- ja xz-projisoituvia osituksia, sanotaan $\mathbb{R}^3$:n perusalueeksi. Jos V on nämä ehdot täyttävä, niin vastaavaan tapaan kuin tasossa nähdään oikeaksi integraalikaava

$$\int_V \frac{\partial f}{\partial z}\, dxdydz = \int_{\partial V} \omega^{(z)} f\, dxdy,$$

missä $\omega^{(z)}$ saa arvoja $0, \pm 1$ vastaavalla periaatteella kuin tasossa. Tässä voidaan kirjoittaa $dxdy = |n_z| dS$, missä n_z on ∂V:n yksikkönormaalivektorin z-komponentti ja dS viittaa ∂V:n pinta-alamittaan (vrt. Luku XIII.7). Kun huomioidaan myös $\omega^{(z)}$:n merkinvaihtelu, niin integraalikaavalle saadaan muoto

$$\int_V \frac{\partial f}{\partial z}\, dxdydz = \int_{\partial V} n_z f\, dS,$$

missä $\vec{n}$ on ∂V:n ulkonormaali (yksikkövektori). Tämä on yksi kolmesta *Greenin avaruuskaavasta* — muut kaksi ovat ilmeisiä. Yhdistämällä nämä kaavat seuraa

LAUSE XIV.2.2 **(Gaussin lause avaruudessa)** Jos $V \subset \mathbb{R}^3$ on perusalue ja $\vec{F}$ on V:ssä määritelty, jatkuvasti derivoituva vektorikenttä, niin pätee

$$\boxed{\int_V \nabla \cdot \vec{F}\, dxdydz = \int_{\partial V} \vec{n} \cdot \vec{F}\, dS.}$$

Myös tässä *Gaussin avaruuskaavassa* pinnan ∂A pinta-alamitta palvelee vain muistisääntönä. Todellisuudessa kaavan oikea puoli koostuu tasointegraaleista, kuten kaavan johdosta ilmenee.

Gaussin avaruuskaavassa kirjoitetaan oikea puoli usein muotoon

$$\int_{\partial A} \vec{n} \cdot \vec{F}\, dS = \int_{\partial A} \vec{F} \cdot d\vec{a},$$

missä $d\vec{a} = \vec{n}\, dS$ on 'vektoroitu' mitta. Vastaavasti voidaan tasokaavassa kirjoittaa

$$\vec{n} \cdot \vec{F}\, ds = \vec{F} \cdot d\vec{n}, \quad d\vec{n} = \vec{n}\, ds.$$

ESIMERKKI 1 Olkoon $V = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2\}$ ja $\vec{F}(x, y, z) = x\vec{i} + 2y\vec{j} + 3z\vec{k}$. Laske integraali $\int_{\partial V} \vec{n} \cdot \vec{F}\, dS$ kahdella eri tavalla.

<u>Ratkaisu</u> a) Pinnalla ∂V on $\vec{n} = (x\vec{i}+y\vec{j}+z\vec{k})/R$, joten pallonpintakoordinaattien avulla suoraan laskien saadaan

$$\begin{aligned}
\int_{\partial V} \vec{n}\cdot\vec{F}\, dS &= \int_{\partial V} R^{-1}(x^2 + 2y^2 + 3z^2)\, dS \\
&= \int_0^{\pi} \int_0^{2\pi} R^{-1}(R^2 \sin^2\theta \cos^2\varphi + 2R^2 \sin^2\theta \sin^2\varphi + 3\cos^2\theta)\, R^2 \sin\theta\, d\theta d\varphi \\
&= R^3 \int_0^{\pi} \sin^3\theta\, d\theta \int_0^{2\pi} \cos^2\varphi\, d\varphi + 2R^3 \int_0^{\pi} \sin^3\theta\, d\theta \int_0^{2\pi} \sin^2\varphi\, d\varphi \\
&\qquad + 3R^3 \int_0^{\pi} \cos^2\theta \sin\theta\, d\theta \int_0^{2\pi} d\varphi \\
&= R^3 \left(\frac{4}{3} \cdot \pi + 2 \cdot \frac{4}{3} \cdot \pi + 3 \cdot \frac{2}{3} \cdot 2\pi\right) = \underline{\underline{8\pi\, R^3}}.
\end{aligned}$$

b) Gaussin (avaruus)kaavan mukaan

$$\begin{aligned}
\int_{\partial V} \vec{n} \cdot \vec{F}\, dS = \int_V \nabla \cdot \vec{F}\, dxdydz &= \int_V 6\, dxdydz \\
&= 6\mu(V) = 6 \cdot \frac{4}{3}\pi\, R^3 = \underline{\underline{8\pi\, R^3}}. \quad \square
\end{aligned}$$

Yleistetty Gaussin lause

Samaan tapaan kuin edellä johdettaessa Gaussin tasokaava Greenin kaavoista voidaan päätellä, että Gaussin tasokaava on voimassa myös, jos pistetulon paikalla on kaavan kummallakin puolella ristitulo, eli kaava pysyy voimassa muunnoksin

$$\nabla \cdot \vec{F} \hookrightarrow \nabla \times \vec{F}, \quad \vec{n} \cdot \vec{F} \hookrightarrow \vec{n} \times \vec{F}.$$

Edelleen kaava pätee myös muunnoksin

$$\nabla \cdot \vec{F} \hookrightarrow \nabla f, \quad \vec{n} \cdot \vec{F} \hookrightarrow f\vec{n}.$$

Gaussin avaruuskaava (Lause XIV.2.2) pätee samoin muunnoksin. Käyttämällä merkintöjä $d\vec{n} = \vec{n}\,ds$ (taso) $d\vec{a} = \vec{n}\,dS$ (avaruus) ja yleiskertomerkkiä $*$ voi tulokset yhdistää muotoon

$$\boxed{\begin{aligned} \int_A \nabla * \vec{F}\,dxdy &= \int_{\partial A} d\vec{n} * \vec{F} \quad \text{(taso)}, \\ \int_V \nabla * \vec{F}\,dxdydz &= \int_{\partial V} d\vec{a} * \vec{F} \quad \text{(avaruus)}. \end{aligned}}$$

Tapauksessa $*$ = 'tyhjä' (skalaarin ja vekorin kertolasku) on tässä vektorikenttä $\vec{F}$ tulkittava skalaarikentäksi: $\vec{F} \hookrightarrow f$. Tulokset, jotka siis pätevät Gaussin lauseen oletuksin, tunnetaan *yleistettynä Gaussin lauseena.*

Esimerkki 2 Veteen upotetun kappaleen $V \subset \mathbb{R}^3$ reunapinnalla ∂V vaikuttaa paine

$$\vec{f}(x, y, z) = -\rho_0 g z\,\vec{n}$$

(ρ_0 = veden tiheys). Mikä on kappaleeseen kohdistuva kokonaisvoima $\vec{F}$?

<u>Ratkaisu</u> Yleistetyn Gaussin lauseen ($*$ = 'tyhjä') mukaan

$$\begin{aligned} \vec{F} &= -\rho_0 g \int_{\partial V} z\,\vec{n}\,dS = -\rho_0 g \int_V \nabla z\,dV \\ &= -\rho_0 g\,\vec{e}_z \int_V dV = \underline{\underline{-m(V) g\,\vec{e}_z}}, \end{aligned}$$

missä $m(V)$ = kappaleen syrjäyttämän vesimäärän massa. Tulos tunnetaan <u>Arkhimedeen</u> <u>lakina</u>. □

HARJOITUSTEHTÄVIÄ

1. a) Laske $\int_A (\partial f/\partial y)\, dxdy$, kun A on origokeskinen yksikkökiekko ja $f(x,y) = 2x - 3y^2 + (x^2+y^2-1)\sin(1+xy)$.

   b) Laske $\int_{\partial A} \vec{F}\cdot d\vec{n}$, kun A on ellipsin $S:\ x^2/a^2 + y^2/b^2 = 1$ sisään jäävä alue ja $\vec{F} = (x+e^y)\vec{i} + (\sin x + 2y)\vec{j}$.

2. Olkoon $V = \{(x,y,z)\in\mathbb{R}^3 \mid |x|\le 1\ \&\ |y|\le 1\ \&\ |z|\le 1\}$ ja $\vec{F}(x,y,z) = x\,\vec{i} + y^2\,\vec{j} - \vec{k}$. Laske integraalit $\int_{\partial V} d\vec{a}\cdot\vec{F}$ ja $\int_{\partial V} d\vec{a}\times\vec{F}$ a) suoraan pintaintegraaleina, b) avaruusintegraaleina käyttäen yleistettyä Gaussin lausetta.

3. Olkoon $V = \{(x,y,z)\in\mathbb{R}^3 \mid x^2+y^2+z^2\le 1\ \&\ x\ge 0\ \&\ y\ge 0\}$. Laske $\int_{\partial V}[(x+y)\,\vec{i} - 2xz\,\vec{j} + (y-z)\,\vec{k}]\times d\vec{a}$ a) suoraan, b) tilavuusintegraaliksi muuntamalla.

4. Laske $\int_{\partial V}\vec{F}\cdot d\vec{a}$ ja $\int_{\partial V}\vec{F}\times d\vec{a}$, kun $\vec{F}(x,y,z) = 3xz^2\vec{i} - x\vec{j} - y\vec{k}$ ja $V = \{(x,y,z)\in\mathbb{R}^3 \mid 0\le x\le 1\ \&\ y\ge 0\ \&\ y^2+z^2\le 1\}$.

5. Olkoon $\vec{F}$ säännöllinen vektorikenttä ja u säännöllinen skalaarikenttä perusalueessa $V\subset\mathbb{R}^3$. Näytä, että $\int_{\partial V}\nabla\times\vec{F}\cdot d\vec{a} = 0$ ja $\int_{\partial V}\nabla u\times d\vec{a} = \vec{0}$.

6. Jos $A\subset\mathbb{R}^2$ ja $V\subset\mathbb{R}^3$ ovat perusalueita, niin mikä geometrinen merkitys on seuraavilla integraaleilla?

   a) $\dfrac{1}{2}\displaystyle\int_{\partial A}\vec{r}\cdot d\vec{n}$ b) $\dfrac{1}{3}\displaystyle\int_{\partial V}\vec{r}\cdot d\vec{a}$ c) $\dfrac{1}{2\mu(V)}\displaystyle\int_{\partial V}(x^2+y^2+z^2)\, d\vec{a}$

7. Olkoon u, f ja $\vec{F}$ riittävän säännöllisiä skalaari- ja vektorikenttiä perusalueessa $V\subset\mathbb{R}^3$. Näytä, että pätee

$$\int_V \nabla u * \vec{F}\, dxdydz = \int_{\partial V} u\, d\vec{a} * \vec{F} - \int_V u\nabla * \vec{F}\, dxdydz,$$

   missä $\vec{F} = f$, jos $* =$ 'tyhjä'. Mikä on kaavan 2-ulotteinen vastine? Entä 1-ulotteinen vastine, kun $V = [a,b]$?

8. (*) (Zeppeliini) Tyynellä säällä ilman paine $p(x,y,z)$ ja tiheys $\rho(x,y,z)$ toteuttavat tasapainolain

$$\frac{\partial p}{\partial z} = -g\rho,$$

   missä $z =$ korkeus maan pinnasta ja $g =$ maan vetovoiman kiihtyvyys. Jos $V\subset\mathbb{R}^3$ on ilmassa leijuva ilmalaiva, niin päteekö Arkhimedeen laki?

XIV.3 Gaussin lauseen sovelluksia

Fysiikassa Gaussin avaruuskaava (Lause XIV.2.2) esitetään usein muodossa

$$\int_V \nabla \cdot \vec{F}\, dV = \int_{\partial V} \vec{F} \cdot d\vec{a}, \quad V \subset \mathbb{R}^3,$$

missä on merkitty $dV = dxdydz$. Jos $A \subset S$, missä S on avaruuden pinta, niin pintaintegraali

$$\phi = \int_A \vec{F} \cdot d\vec{a} = \int_A \vec{F} \cdot \vec{n}\, dS \qquad (\star)$$

on vektorikentän $\vec{F}$ *vuo* (engl. flux) A:n läpi. Ajatellen tätä yhteyttä sanotaan itse vektorikenttää fysikaalisissa sovelluksissa usein vuontiheydeksi. Seuraavassa kahdessa sovellusesimerkissä johdetaan Gaussin kaavan avulla fysikaalista ilmiötä kuvaava säilymislaki osittaisdifferentiaaliyhtälön muodossa. Ensimmäinen esimerkki on virtausmekaniikasta ja toinen lämpöopista.

Sovellusesimerkki: Massan säilymislaki virtauksessa

Tarkastellaan virtaavaa nestettä tai kaasua, jonka nopeus hetkellä t pisteessä (x, y, z) on $\vec{v} = \vec{v}(t, x, y, z)$ (vektorikenttä, yksikkö m/s) ja tiheys on $\rho = \rho(t, x, y, z)$ (skalaarikenttä, yksikkö kg/m^3). Olkoon $A \subset S$, missä S on avaruuden säännöllinen (tai ainakin paloittain säännöllinen) pinta. Tarkastellaan A:n pientä, lähes tasomaista palaa ΔA, jonka pinta-ala $= \Delta S$. Jos oletetaan, että ρ ja $\vec{v}$ ovat ΔA:n ympäristössä lähes vakioita (jatkuvuusoletus!), niin voidaan päätellä, että (lyhyellä) aikavälillä $[t, t + \Delta t]$ pinnanpalan ΔA läpi menevät ne hiukkaset, jotka ovat hetkellä t joukossa

$$\Delta V = \{\, (x, y, z) \mathrel{\widehat{=}} \vec{r} \mid \vec{r} = \vec{r}_0 - \tau\vec{v},\ r_0 \in A,\ \tau \in [0, \Delta t]\,\}.$$

Olkoon $\vec{n}$ virtaussuuntaan osoittava S:n yksikkönormaalivektori, ts. $\vec{n} \cdot \vec{v} \geq 0$. Jos oletetaan $\vec{v}$ vakioksi ja samoin $\vec{n}$ vakioksi ΔA:ssa (eli pinta tasoksi), niin ΔV on näillä oletuksilla suuntaissärmiö, jonka poikkipinta-ala virtausta vastaan

kohtisuorassa tasossa on $\Delta S\,\vec{v}\cdot\vec{n}/|\vec{v}\,|$. Jos edellen myös ρ oletetaan vakioksi, niin ΔV:ssä olevan nesteen/kaasun massa on tehdyin oletuksin

$$\begin{aligned}\Delta m = \rho\mu(\Delta V) &= \rho\,[\Delta S\,\vec{v}\cdot\vec{n}/|\vec{v}\,|]\,|\vec{v}\,|\Delta t = \rho\vec{v}\cdot\vec{n}\,\Delta S\Delta t\\ \Rightarrow\quad \frac{\Delta m}{\Delta S\Delta t} &= \rho\vec{v}\cdot\vec{n}.\end{aligned}$$

Voidaan olettaa, että tehdyt likmääräistykset ovat voimassa lähes kaikissa A:n pisteissä, lukuun ottamatta mahdollista (pinta-alamitan suhteen) nollamittaista osajoukkoa. Muissa kuin näissä poikkeuspisteissä voidaan mainittujen oletusten myös olettaa toteutuvan tarkasti raja-arvoina, kun $\Delta t \to 0$ ja ΔA kutistuu pisteeksi. Näin olettaen voidaan <u>massavuo</u> pinnan A läpi (yksikkö kg/s) laskea integraalikaavalla ($\star$), missä <u>massavuon</u> <u>tiheys</u> (yksikkö kg/m^2/s) määritellään

$$\vec{F} = \rho\vec{v}.$$

Oletetaan nyt, että S on suljettu pinta, joka sulkee sisäänsä perusalueen $V \subset \mathbb{R}^3$ ($S = \partial V$). Tällöin nesteen/kaasun kokonaismassa V:ssä hetkellä t on

$$m(t) = \int_V \rho(t,x,y,z)\,dV.$$

Oletetaan, että $m(t)$ voi muuttua vain V:n reunapinnan läpi tapahtuvan virtauksen vuoksi (eli muilla tavoilla massaa ei 'häviä' tai 'synny'). Tällöin on voimassa tasapainolaki

$$m'(t) = -\int_{\partial V} \rho\vec{v}\cdot d\vec{a},$$

missä $d\vec{a} = \vec{n}dS$ osoittaa V:stä poispäin, eli $\vec{n} = V$:n ulkonormaali. Kun tässä oikealla käyttetään Gaussin kaavaa ja vasemmalla kirjoitetaan (vrt. Luku XI.2)

$$m'(t) = \frac{d}{dt}\int_V \rho(t,x,y,z)\,dV = \int_V \frac{\partial\rho}{\partial t}(t,x,y,z)\,dV,$$

niin tasapainolaki saa muodon

$$\int_V \big[\rho_t + \nabla\cdot(\rho\vec{v})\big]\,dV = 0. \tag{a}$$

Olkoon nyt $t \in \mathbb{R}$ kiinteä ajanhetki, $(x,y,z) \in \mathbb{R}^3$ kiinteä piste, ja oletetaan, että ρ ja $\vec{v}$ ovat jatkuvasti derivoituvia (osittaisderivaatat jatkuvia) pisteen $(t,x,y,z) \in \mathbb{R}^4$ ympäristössä. Tällöin jos tasapainolaissa (a) valitaan $V = B_\varepsilon = \varepsilon$-säteinen pallo (kuula), jonka keskipiste on (x,y,z), niin

$$\begin{aligned}0 &= \int_{B_\varepsilon} \big[\rho_t + \nabla\cdot(\rho\vec{v})\big]\,dV\\ &= \big[\rho_t + \nabla\cdot(\rho\vec{v})\big](t,x,y,z)\,\mu(B_\varepsilon) + o(1)\mu(B_\varepsilon), \quad \text{kun } \varepsilon \to 0.\end{aligned}$$

Tämän mukaan on tarkasteltavassa pisteessä (t, x, y, z) oltava voimassa

$$\boxed{\rho_t + \nabla \cdot (\rho \vec{v}) = 0.} \tag{b}$$

Tämä tunnetaan virtausmekaniikan (differentiaalisena) massan säilymislakina. — Huomattakoon, että koska tätä johdettaessa tehdyt jatkuvuusoletukset eivät ole fysiikan sanelemia, niin säilymislain integraalimuoto (a) (joka pätee heikommin säännöllisyysehdoin jokaiselle perusalueelle V) on luonnonlakina alkuperäisempi kuin differentiaalinen muoto (b). Vrt. vastaava asetelma Esimerkissä VIII.6: 1.

Sovellusesimerkki: Lämmön johtuminen

Tarkastellaan lämmön johtumisen ongelmaa, kun avaruuden $\mathbb{R}^3$ täyttää fysikaalisilta ominaisuuksiltaan tunnettu materiaali. Ongelmaan liittyvät perussuureet ovat lämpötila $u = u(t, x, y, z)$, lämpövuon tiheys $\vec{J} = \vec{J}(t, x, y, z)$ (yksikkö $\mathrm{W/m^2}$), ja lämpölähteen tiheys $\rho(t, x, y, z)$ (yksikkö $\mathrm{W/m^3}$). Materiaalikertoimina oletetaan vielä tunnetuiksi materiaalin ominaislämpö $c = c(x, y, z, u)$ ja lämmönjohtavuus $\lambda = \lambda(x, y, z, u)$. (Tässä oletetaan, että materiaalikertoimet voivat riippua paikkamuuttujien lisäksi lämpötilasta.)

Jos $V \subset \mathbb{R}^3$, niin V:n sisältämän lämpöenergian kokonaismäärä hetkellä t on (ominaislämmön määritelmä)

$$E(t) = \int_V cu \, dV.$$

Jos oletetaan, että V:n energiataseeseen vaikuttavat vain lämpölähde V:ssä ja lämmön johtuminen V:n reunapinnan läpi, niin energian tasapainoyhtälö V:ssä on

$$E'(t) = \int_V \rho \, dV - \int_{\partial V} \vec{J} \cdot d\vec{a}.$$

Kun tässä kirjoitetaan

$$E'(t) = \frac{d}{dt} \int_V cu \, dV = \int_V cu_t \, dV, \qquad \int_{\partial V} \vec{J} \cdot d\vec{a} = \int_V \nabla \cdot \vec{J} \, dV,$$

niin tasapainoyhtälö saa muodon

$$\int_V (cu_t + \nabla \cdot \vec{J} - \rho) \, dV = 0.$$

Samanlaisella päättelyllä (ja oletuksilla) kuin edellisessä esimerkissä seuraa tästä differentiaalinen energian säilymislaki

$$cu_t + \nabla \cdot \vec{J} = \rho.$$

Jos vielä oletetaan lämmönjohtumisen Fourier'n laki

$$\vec{J} = -\lambda \nabla u,$$

niin energian säilymislaki saa lämmönjohtumisyhtälönä tunnetun muodon

$$\boxed{cu_t - \nabla \cdot (\lambda \nabla u) = \rho.}$$

Jos $\rho = 0$ ja c ja λ ovat vakioita (homogeeninen materiaali, c:llä ja λ:lla ei lämpötilariippuvuutta), niin lämmönjohtumisyhtälö yksinkertaistuu muotoon

$$u_t = k\Delta u, \quad k = \lambda/c.$$

*Vektorikenttien epäjatkuvuudet

Olkoon vektorikenttä $\vec{F}$ määritelty $\mathbb{R}^3$:ssa ja jatkuvasti derivoituva. Tällöin kentän lähde $\rho = \nabla \cdot \vec{F}$ on $\mathbb{R}^3$:ssa jatkuva, ja Gaussin lauseen mukaan pätee

$$\int_V \rho \, dV = \int_{\partial V} \vec{F} \cdot d\vec{a} \quad (V \subset \mathbb{R}^3 \text{ perusalue}). \tag{1}$$

Toisaalta, jos mainittujen säännöllisyysoletusten lisäksi oletetaan ainoastaan (1), niin aiemman päättelyn mukaisesti seuraa, että on oltava $\nabla \cdot \vec{F} = \rho$. Siis jos $\vec{F}$ on $\mathbb{R}^3$:ssa jatkuvasti derivoituva ja ρ jatkuva, niin pätee

$$\text{Ehto (1) voimassa} \quad \Leftrightarrow \quad \nabla \cdot \vec{F} = \rho \ \ \mathbb{R}^3 : \text{ssa}.$$

Entä jos $\vec{F}$ ei ole jatkuvasti derivoituva, tai ρ ei ole jatkuva, mutta $\vec{F}$:n ja ρ:n välillä on ehdon (1) mukainen yhteys? — Tällöin on luontevaa sopia, että ehto (1) määrittelee ρ:n kentän $\vec{F}$ lähteeksi. Fysikaalisissa sovelluksissa ajatellaan tällöin, että ehto (1) itse asiassa ilmaisee luonnonlain sen alkuperäisessä muodossa (vrt. massan säilymislain johto edellä).

Vektorikentän ja lähteen välisen riippuvuuden esittäminen integraalimuotoisena 'Gaussin lakina' (1) on erityisen hyödyllistä silloin, kun halutaan johtaa fysikaalisten vektorikenttien jatkuvuusehtoja materiaalirajapinnoilla. Tarkastellaan esimerkkinä tällaisten ehtojen asettamista tasolla (tasomaisella materiaalirajapinnalla). Olkoon T avaruustaso, joka jakaa $\mathbb{R}^3$:n kahteen avoimeen osaan V_1 ja V_2 (eli $\mathbb{R}^3$ jakautuu pistevieraisiin osiin V_1, V_2 ja T), ja olkoon $\vec{F}$ paloittain jatkuvasti derivoituva vektorikenttä muotoa

$$\vec{F}(x,y,z) = \begin{cases} \vec{F}_1(x,y,z), & \text{kun } (x,y,z) \in V_1, \\ \vec{F}_2(x,y,z), & \text{kun } (x,y,z) \in V_2, \end{cases}$$

missä $\vec{F}_1$ ja $\vec{F}_2$ ovat molemmat koko $\mathbb{R}^3$:ssa määriteltyjä ja jatkuvasti derivoituvia. Olkoon vastaavasti ρ paloittain jatkuva funktio muotoa

$$\rho(x,y,z) = \begin{cases} \rho_1(x,y,z), & \text{kun } (x,y,z) \in V_1, \\ \rho_2(x,y,z), & \text{kun } (x,y,z) \in V_2, \end{cases}$$

missä ρ_1 ja ρ_2 ovat molemmat koko $\mathbb{R}^3$:ssa määriteltyjä ja jatkuvia.

Em. säännöllisyysoletuksien lisäksi oletetaan vielä, että kenttää $\vec{F}$ ja funktiota ρ sitoo ehto (1), ts. ρ on kentän $\vec{F}$ lähde. Tällöin jos ko. ehdossa valitaan $V \subset V_1$ tai $V \subset V_2$ niin seuraa, että on oltava $\nabla \cdot \vec{F}_i = \rho_i$ V_i:ssä, $i = 1, 2$, eli pätee

$$\nabla \cdot \vec{F}(x,y,z) = \rho(x,y,z), \quad (x,y,z) \in V_1 \cup V_2. \tag{2}$$

Tutkitaan seuraavaksi, mitä ehdosta (1) seuraa, kun joukko V valitaan siten, että $V \cap T \neq \emptyset$. Tätä silmällä pitäen tarkastellaan pistettä $(x,y,z) \in T$ ja määritellään tätä pistettä ympäröivä 'pillerirasia' $B_{h,d}$ ehdoilla: (i) $B_{h,d}$ on suorakulmainen särmiö, jonka keskipiste $= (x,y,z)$ ja kaksi sivutahkoa ovat tason T suuntaiset. (ii) $B_{h,d}$:n tasoa T vastaan kohtisuoran särmän pituus $= d$. (iii) Suorakulmio $T_h = B_{h,d} \cap T$ on d:stä riippumaton ja T_h:n suurimman sivun pituus $= h$.

Valitaan ehdossa (1) $V = B_{h,d}$, missä $T_h = B_{h,d} \cap T$ on kiinteä ja $d \to 0$. Tällöin seuraa

$$\lim_{d\to 0} \int_{\partial B_{h,d}} \vec{F} \cdot d\vec{a} = \int_{T_h} \vec{n} \cdot (\vec{F}_2 - \vec{F}_1)\, d\mu = \lim_{d\to 0} \int_{B_{h,d}} \rho\, dV = 0,$$

missä μ on tason T pinta-alamitta ja $\vec{n}$ on T:n yksikkönormaalivektori, joka osoittaa osajoukon V_2 suuntaan (ks. kuviot). Kun tässä tuloksessa annetaan edelleen h:n lähestyä 0:aa ja huomioidaan kenttien $\vec{F}_1$ ja $\vec{F}_2$ jatkuvuus pisteessä (x, y, z), niin seuraa, että on oltava $\vec{n} \cdot (\vec{F}_1 - \vec{F}_2)(x, y, z) = 0$. Tämä on voimassa jokaisessa T:n pisteessä, joten on päätelty, että tasolla T on voimassa jatkuvuusehto

$$\vec{n} \cdot \vec{F}_1(x, y, z) = \vec{n} \cdot \vec{F}_2(x, y, z), \quad (x, y, z) \in T. \tag{3}$$

Sekä tämä että (2) ovat siis ehdon (1) seurauksia (tehtyjen säännöllisyysoletusten puitteissa).

Toisaalta jos lähdetään samoista säännöllisyysoletuksista ja oletetaan lisäksi (2) ja (3), niin näistä yhdessä seuraa (1). Nimittäin jos V on perusalue ja $A_i = V \cap V_i \neq \emptyset$, $i = 1, 2$, niin oletuksen (2) ja Gaussin lauseen perusteella voidaan päätellä

$$\begin{aligned} (2) \quad &\Rightarrow \quad \sum_{i=1}^{2} \int_{A_i} \nabla \cdot \vec{F}_i\, dV = \sum_{i=1}^{2} \int_{A_i} \rho_i\, dV = \int_V \rho\, dV \\ &\Rightarrow \quad \sum_{i=1}^{2} \int_{\partial A_i} \vec{F}_i \cdot d\vec{a} = \int_V \rho\, dV. \end{aligned}$$

Koska tässä on

$$\partial A_1 \cup \partial A_2 = \partial V \cup (\partial A_1 \cap \partial A_2), \quad \partial A_1 \cap \partial A_2 \subset T,$$

niin oletuksen (3) perusteella päätellään edelleen

$$\sum_{i=1}^{2} \int_{\partial A_i} \vec{F}_i \cdot d\vec{a} = \int_{\partial V} \vec{F} \cdot d\vec{a} + \int_{\partial A_1 \cap \partial A_2} \vec{n} \cdot (\vec{F}_1 - \vec{F}_2)\, d\mu = \int_{\partial V} \vec{F} \cdot d\vec{a}.$$

Siis ehto (1) on voimassa. Näin on päätelty, että tehtyjen säännöllisyysoletusten puitteissa pätee

$$\boxed{\text{Ehto (1)} \quad \Leftrightarrow \quad \begin{cases} \nabla \cdot \vec{F} = \rho & \text{joukossa } V_1 \cup V_2, \\ \vec{n} \cdot (\vec{F}_1 - \vec{F}_2) = 0 & \text{rajapinnalla } T. \end{cases}}$$

Johtopäätös on sama myös, jos rajapinta $\partial V_1 \cap \partial V_2$ on osa kaarevaa, riittävän säännöllistä pintaa S. Koskien fysikaalisen vektorikentän jatkuvuusehtoja materiaalirajapinnalla on siis päätelty, että jos vektorikentällä on materiaalirajapinnan läheisyydessä paloittain jatkuva lähde, niin kentän normaalikomponentti on jatkuva materiaalirajapinnalla. Sen sijaan kentän tangentiaalikomponentin ei tarvitse olla jatkuva.

Samaan tapaan kuin ehdossa (1) voidaan myös vektorikentän ja sen pyörrekentän välinen yhteys esittää integraalimuotoisena määritelmänä. Jos $\vec{F}$ on jatkuvasti derivoituva $\mathbb{R}^3$:ssa, niin $\vec{F}$:n pyörrekenttä on $\vec{\omega} = \nabla \times \vec{F}$ (vrt. Luku XI.4). Tällöin pätee yleistetyn Gaussin lauseen perusteella

$$\int_V \vec{\omega}\, dV = \int_{\partial V} d\vec{a} \times \vec{F} \quad (V \subset \mathbb{R}^3 \text{ perusalue}). \tag{4}$$

Jos $\vec{F}$ ei ole koko $\mathbb{R}^3$:ssa jatkuvasti derivoituva, tai $\vec{\omega}$ ei ole jatkuva, niin katsotaan ehto (4) pyörrekentän $\vec{\omega}$ määritelmäksi. Jos nyt $\vec{F}$ toteuttaa samat ehdot kuin edellä ja oletetaan, että määritelmän (4) mukainen pyörrekenttä on muotoa

$$\vec{\omega}(x,y,z) = \begin{cases} \vec{\omega}_1(x,y,z), & \text{kun } (x,y,z) \in V_1, \\ \vec{\omega}_2(x,y,z), & \text{kun } (x,y,z) \in V_2, \end{cases}$$

missä $\vec{\omega}_1$ ja $\vec{\omega}_2$ ovat molemmat koko $\mathbb{R}^3$:ssa määriteltyjä ja jatkuvia, niin samalla tavoin kuin edellä päätellään, että pätee

$$\boxed{\text{Ehto (4)} \quad \Leftrightarrow \quad \begin{cases} \nabla \times \vec{F} = \vec{\omega} & \text{joukossa } V_1 \cup V_2, \\ \vec{n} \times (\vec{F}_1 - \vec{F}_2) = \vec{0} & \text{rajapinnalla } T. \end{cases}}$$

Tässä tapauksessa kentän tangentiaalikomponentti on jatkuva materiaalirajapinnalla, sillä tangentiaalikomponentti on (vrt. Luku II.6)

$$\vec{F}_t = \vec{F} - (\vec{n} \cdot \vec{F})\, \vec{n} = -\vec{n} \times (\vec{n} \times \vec{F}).$$

Kentän normaalikomponentin ei tarvitse olla jatkuva.

Esimerkki 1 Staattinen sähkökenttä $\vec{E}$ ja sähkövuon tiheys $\vec{D}$ toteuttavat Maxwellin yhtälöt (vrt. Luku XI.4)

$$\nabla \times \vec{E} = \vec{0}, \quad \nabla \cdot \vec{D} = \rho,$$

missä ρ on varaustiheys. Oletetaan, että edellä V_1 ja V_2 edustavat kahta eri materiaalia, joissa vallitsevat materiaalilait ovat

$$\vec{D} = \epsilon_i \vec{E} \quad V_i \text{:ssä}, \quad i = 1, 2,$$

missä ϵ_i (= materiaalin i sähköinen permittiivisyys) oletetaan V_i:ssa vakioksi. Jos oletetaan ρ paloittain jatkuvaksi kuten edellä, niin jatkuvuusehdot materiaalirajapinnalla ovat

$$\begin{cases} \vec{n} \times (\vec{E}_1 - \vec{E}_2) = \vec{0}, \\ \vec{n} \cdot (\vec{D}_1 - \vec{D}_2) = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \vec{n} \times \vec{E}_1 = \vec{n} \times \vec{E}_2, \\ \epsilon_1 \, \vec{n} \cdot \vec{E}_1 = \epsilon_2 \, \vec{n} \cdot \vec{E}_2. \end{cases}$$

Sähkökentän tangentiaalikomponentti on siis materiaalirajapinnalla jatkuva. Normaalikomponentti sen sijaan on epäjatkuva, ellei ole joko $\epsilon_1 = \epsilon_2$ tai $\vec{n} \cdot \vec{E}_i = 0$. $\square$

Esimerkki 1 (jatko) Esimerkissä olkoon

$$V_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z < 0\}, \quad V_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z > 0\},$$

ja $\epsilon_1/\epsilon_2 = 2$. Laske sähkökenttä $\vec{E}_2(0, 0, 0)$, kun tiedetään, että $\vec{E}_1(0, 0, 0) = E(\vec{i} + 2\vec{j} - \vec{k})$.

<u>Ratkaisu</u> Materiaalirajapinta on taso $T: \; x + y + z = 0$. Tämän materiaalia 2 kohti osoittava yksikkönormaalivektori on $\vec{n} = (\vec{i} + \vec{j} + \vec{k})/\sqrt{3}$, joten

$$\vec{n} \cdot \vec{E}_1(0, 0, 0) = \frac{2E}{\sqrt{3}}, \quad \vec{n} \times \vec{E}_1(0, 0, 0) = \frac{E}{\sqrt{3}}(-3\vec{i} + 2\vec{j} + \vec{k}).$$

Jatkuvuusehtojen perusteella on

$$\vec{n} \cdot \vec{E}_2(0, 0, 0) = \frac{\epsilon_1}{\epsilon_2} \, \vec{n} \cdot \vec{E}_1(0, 0, 0), \quad \vec{n} \times \vec{E}_2(0, 0, 0) = \vec{n} \times \vec{E}_1(0, 0, 0),$$

joten

$$\begin{aligned} \vec{E}_2(0, 0, 0) &= [\vec{n} \cdot \vec{E}_2(0, 0, 0)] \, \vec{n} - \vec{n} \times [\vec{n} \times \vec{E}_2(0, 0, 0)] \\ &= 2 \, [\vec{n} \cdot \vec{E}_1(0, 0, 0)] \, \vec{n} - \vec{n} \times [\vec{n} \times \vec{E}_1(0, 0, 0)] \\ &= \frac{4E}{3}(\vec{i} + \vec{j} + \vec{k}) + \frac{E}{3} \, (\vec{i} + 4\vec{j} - 5\vec{k}) \\ &= \underline{\underline{\frac{E}{3} \, (5\vec{i} + 8\vec{j} - \vec{k}\,).}} \quad \square \end{aligned}$$

HARJOITUSTEHTÄVIÄ

1. Laske vektorikentän $\vec{F}$ vuo origokeskisen, R-säteisen pallopinnan läpi:
   a) $\vec{F} = x^2y^4z\vec{k}$ b) $\vec{F} = x\vec{i} - 2y\vec{j} + 4z\vec{k}$ c) $\vec{F} = ye^z\vec{i} + x^2e^z\vec{j} + xy\vec{k}$
   d) $\vec{F} = (x^2+y^2)\vec{i} + (y^2-z^2)\vec{j} + z\vec{k}$ e) $\vec{F} = x^2\vec{i} + 3yz^2\vec{j} + (3y^2z + x^2)\vec{k}$

2. Laske vektorikentän $\vec{F} = x^2\vec{i} + y^2\vec{j} + z^2\vec{k}$ vuo annetun alueen $V \subset \mathbb{R}^3$ reunapinnan ∂V läpi:
   a) $V:\ (x-2)^2 + y^2 + (z-3)^2 \le 9$ b) $V:\ x^2 + y^2 + 4(z-1)^2 \le 4$
   c) $V:\ x, y, z \ge 0\ \&\ x + y + z \le 3$ d) $V:\ x^2 + y^2 \le 2y\ \&\ 0 \le z \le 4$

3. Tetraedrin muotoista aluetta $V \subset \mathbb{R}^3$ rajoittavat tasot $z = 0$, $x = 2y$, $x = -y$ ja $y + z = a$ $(a > 0)$. Laske vektorikentän
$$\vec{F} = (3xyz + e^{yz})\vec{i} + (x^2 + y^2z)\vec{j} + (y^2 - 2yz^2)\vec{k}$$
   vuo V:n reunapinnan ∂V läpi sisältä ulospäin.

4. Olkoon $S:\ x^2 + y^2 + z^2 = 1$, $B \subset S$ ja
$$V = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) = t(u, v, w),\ t \in [0, R]\ \&\ (u, v, w) \in B\}.$$
   Laske V:n tilavuus $\mu(V)$ a) pallokoordinaattien avulla, b) soveltamalla Gaussin lausetta vektorikenttään $\vec{F} = x\vec{i} + y\vec{j} + z\vec{k}$.

5. Pyörivässä virtauskentässä massavuon tiheys pisteessä $(x, y, z) \mathrel{\widehat{=}} \vec{r}$ on $\vec{F} = Q\,\vec{r} \times (2\vec{i} - \vec{j} + 3\vec{k})$, missä $Q = 24$ kg/m^3/s ja $\vec{r}$:n yksikkö = m. Olkoon V kuutio, jonka yksi kärki on origossa, origosta lähtevät särmät ovat positiivisilla koordinaattiakseleilla ja sivun pituus on $a = 2$ m. Laske massavuo (yksikkö kg/s) V:n kunkin sivutahkon läpi laskettuna positiivisena kuution sisältä ulospäin.

6. Vektorikenttä $\vec{F}$ ja sen lähde ρ ovat a) pallokoordinaatistossa, b) lieriökoordinaatistossa muotoa $\vec{F} = F(r)\vec{e}_r$, $\rho = \rho(r)$. Laske $F(r)$ lähteen $\rho(r)$ avulla käyttämällä Gaussin kaavaa sopivasti valitussa alueessa V.

7. Olkoot u ja v säännöllisiä funktioita perusalueessa $A \subset \mathbb{R}^2$ tai $V \subset \mathbb{R}^3$. Soveltamalla Gaussin lausetta vektorikenttään $\vec{F} = u\nabla v - v\nabla u$ näytä, että
$$\int_A (u\Delta v - v\Delta u)\, dxdy = \int_{\partial V} \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) ds,$$
$$\int_V (u\Delta v - v\Delta u)\, dxdydz = \int_{\partial V} \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) dS,$$
   missä $\partial/\partial n$ tarkoittaa suunnattua derivaattaa ∂A:n tai ∂V:n normaalin suuntaan.

8. (Autojen säilyminen) Yksiulotteisessa liikennevirrassa on $\rho(t,x)$ autotiheys tiellä (yksikkö autoa/km) ja $v(t,x)$ nopeus (yksikkö km/h). a) Määrittele autovuo tiellä ja johda autojen säilymislaki. b) Pisteessä $x = 0$ tie 1 kapenee tieksi 2. Tiellä 1 ($x < 0$) liikennevirta on tasainen: $\rho = 26$ autoa/km ja $v = 120$ km/h. Myös tiellä 2 liikennevirta on tasainen (ρ ja v vakioita) siten, että tien maksimikapasiteetti 3000 autoa/h ei ylity. Voiko tällaisessa liikennetilanteessa piste $x = 0$ olla ruuhkaton? Jos ei, niin monellako autolla vähintään ruuhka kasvaa tunnissa?

9. Lämmönjohtumisen perusyhtälöt ovat

$$\nabla \cdot \vec{J} = \rho, \quad \vec{J} = -\lambda \nabla u,$$

missä $\vec{J}$ on lämpövirran tiheys, ρ lämpölähteen tiheys, u lämpötila ja λ materiaalin lämmönjohtavuus. Olkoon $\lambda = \lambda_i =$ vakio V_i:ssa, $i = 1, 2$, missä $\lambda_2 = 3\lambda_1$ ja

$$V_{1,2} = \{\, (x,y,z) \in \mathbb{R}^3 \mid x + 2y + 2z \lesseqgtr 0 \,\}.$$

Määritä lämpövirtavektorin $\vec{J}$ raja-arvo $\vec{J}_2(0,0,0)$ materiaalin 2 (V_2) puolelta origoa lähestyttäessä, kun tiedetään, että ko. raja-arvo materiaalin 1 puolelta on $\vec{J}_1(0,0,0) = J(-\vec{i}+\vec{j}+3\vec{k})$ (J vakio). Lähde ρ on kummassakin materiaalissa vakio ja lämpötila u on materiaalirajapinnalla jatkuva.

10. (*) Olkoon $A \subset \mathbb{R}^2$ tai $V \subset \mathbb{R}^3$ perusalue. Näytä, että jos u_1 ja u_2 ovat reuna-arvotehtävän

$$\begin{cases} \Delta u = \rho & A\text{:ssa } (V\text{:ssä}), \\ u = 0 & \text{reunalla } \partial A \text{ (reunalla } \partial V) \end{cases}$$

ratkaisuja ja lisäksi riittävän säännöllisiä, niin funktiolle $u = u_1 - u_2$ pätee

$$\int_A |\nabla u|^2 \, dxdy = 0 \quad \left(\int_V |\nabla u|^2 \, dxdydz = 0 \right).$$

Päättele, että reuna-arvotehtävän ratkaisu — sikäli kuin olemassa ja riittävän säännöllinen — on yksikäsitteinen.

XIV.4 Stokesin lause

Tarkastellaan aluksi tason perusaluetta A (ks. Luku XIV.2) ja A:ssa määriteltyä vektorikenttää $\vec{F}(x,y) = F_1(x,y)\vec{i} + F_2(x,y)\vec{j}$. Kun tulkitaan $\vec{F}$ avaruuden vektorikentäksi, niin $\vec{F}$:n roottori on (vrt. Luku XI.4)

$$\nabla \times \vec{F} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\vec{k} = (\operatorname{rot}\vec{F}\,)\,\vec{k},$$

joten Greenin tasokaavojen (ks. Luku XIV.2) perusteella

$$\begin{aligned}\int_A \operatorname{rot}\vec{F}\,dxdy &= \int_{\partial A} (-n_y F_1 + n_x F_2)\,ds \\ &= \int_{\partial A} \vec{t}\cdot\vec{F}\,ds,\end{aligned}$$

missä $\vec{t}$ on reunan tangenttivektori suunnistettuna kuvan mukaisesti, eli $\vec{t}$ osoittaa oikealle reunan ulkonormaalin suunnalta katsottuna. Tätä sanotaan reunan *positiiviseksi suunnistukseksi.*

Integraalia yli ∂A:n, em. tavalla laskettuna, sanotaan *kiertointegraaliksi* ja merkitään symbolilla $\oint_{\partial A}$. Huomioiden, että (vrt. Luku XIV.1)

$$\vec{t}\,ds = d\vec{r}$$

nähdään, että kiertointegraali on yhdistelmä polkuintegraaleja (itse asiassa työintegraaleja), joissa kukin reunan ∂A erillinen osa kierretään suunnistuksen määräämällä tavalla. Tulos mainituin merkinnöin on siis

$$\int_A \operatorname{rot}\vec{F}\,dxdy = \oint_{\partial A} \vec{F}\cdot d\vec{r}. \tag{1}$$

Olkoon seuraavaksi $\vec{F}$ avaruuden vektorikenttä ja A perusalue avaruustasolla T. Valitaan karteesinen koordinaatisto siten, että T on xy-taso ja kirjoitetaan tässä koordinaatistossa $\vec{F} = F_1\vec{i} + F_2\vec{j} + F_3\vec{k}$ ja $\mathrm{rot}\vec{F} = \partial_x F_2 - \partial_y F_1$. Tällöin nähdään, että integraalikaava (1) säilyttää pätevyytensä, koska kaavassa oikealla puolella on $d\vec{r} \cdot \vec{k} = 0$, jolloin $\vec{F} \cdot d\vec{r} = (F_1\vec{i} + F_2\vec{j}) \cdot d\vec{r}$. Koska kaavassa vasemmalla puolella voidaan myös kirjoittaa $\mathrm{rot}\vec{F} = (\nabla \times \vec{F}) \cdot \vec{k}$, niin voidaan vetää yleisempi johtopäätös: Jos $\vec{F}$ on jatkuvasti derivoituva avaruuden vektorikenttä ja A on perusalue avaruustasolla, jonka yksikkönormaalivektori on $\vec{n}$, niin pätee integraalikaava

$$\int_A (\nabla \times \vec{F}) \cdot \vec{n}\, dS = \oint_{\partial A} \vec{F} \cdot d\vec{r}. \tag{2}$$

Tässä on normaalin $\vec{n}$ suunta ja reunan ∂A suunnistus sidottava toisiinsa siten, että kun reunaviivaa katsellaan normaalin $\vec{n}$ osoittamalta puolelta, niin tilanne on kuvan mukainen.

Kaava (2) on erikoistapaus vieläkin yleisemmästä tuloksesta, joka tunnetaan *Stokesin*† *lauseen* nimellä. Stokesin lauseen mukaan kaava (2) on tietyin edellytyksin yleistettävissä koskemaan myös avaruuden kaarevia pintoja. Tämä yleisempi tulos on integraalikaava (Stokesin kaava)

$$\boxed{\int_A (\nabla \times \vec{F}) \cdot d\vec{a} = \oint_{\partial A} \vec{F} \cdot d\vec{r}.} \tag{3}$$

Tässä $\vec{F}$ on avaruuden vektorikenttä, A on (riittävän säännöllisen muotoinen) alue avaruuden (riittävän säännöllisellä) pinnalla S ja $d\vec{a} = \vec{n}\, dS$, missä dS viittaa S:n pinta-alamittaan ja $\vec{n}$ = pinnan yksikkönormaalivektori.

Kaavassa (3) edellytetään, että pinnan normaali $\vec{n}$ ja reunaviivan ∂A suunnistus on sidottu toisiinsa edellä kuvatulla tavalla. Tällaisen suunnistuksen järjestäminen ei ole ongelma siinä tapauksessa, että A on avaruustasolla. Avaruuden kaarevilla pinnoilla voi kuitenkin ongelmia tulla, ja yleisessä Stokesin lauseessa pintaa A onkin rajoitettava geometrisella ehdolla: Pinnan on oltava oletetulla tavalla *suunnistuva*. Tarkastellaan asiaa esimerkkien valossa.

†**Sir George Gabriel Stokes** (1819-1903) oli englantilainen fyysikko–matemaatikko.

Esimerkki 1 Puolipallo on suunnistuva (vrt. kuvio), joten Stokesin kaava (3) on sille pätevä. Kaavan voi myös todistaa pallokoordinaatteja käyttäen: Olkoon pallo R-säteinen ja

$$\vec{F} = F(r,\theta,\varphi)\vec{e}_r + F_\theta(r,\theta,\varphi)\vec{e}_\theta + F_\varphi(r,\theta,\varphi)\vec{e}_\varphi.$$

Tällöin (vrt. Luku XI.4)

$$\nabla\times\vec{F} = \frac{1}{r\sin\theta}[\partial_\theta(\sin\theta\, F_\varphi) - \partial_\varphi F_\theta]\vec{e}_r + [\ldots]\vec{e}_\theta + [\ldots]\vec{e}_\varphi,$$

joten kun valitaan $\vec{n} = \vec{e}_r$ (toinen vaihtoehto olisi $\vec{n} = -\vec{e}_r$), niin

$$\begin{aligned}
\int_A (\nabla\times\vec{F})\cdot d\vec{a} &= \int_0^{\pi/2}\int_0^{2\pi} \frac{1}{R\sin\theta}[\partial_\theta(\sin\theta F_\varphi) - \partial_\varphi F_\theta]\, R^2\sin\theta\, d\theta d\varphi \\
&= R\int_0^{2\pi}\left(\int_0^{\pi/2}\partial_\theta(\sin\theta F_\varphi)\, d\theta\right) d\varphi - R\int_0^{\pi/2}\left(\int_0^{2\pi}\partial_\varphi F_\theta\, d\varphi\right) d\theta \\
&= R\int_0^{2\pi}\left[\Big|_{\theta=0}^{\theta=\pi/2}\big[\sin\theta\, F_\varphi(\theta,\varphi)\big]\right] d\varphi - R\int_0^{\pi/2}\left[\Big|_{\varphi=0}^{\varphi=2\pi} F_\theta(\theta,\varphi)\right] d\theta \\
&= R\int_0^{2\pi} F_\varphi(\tfrac{\pi}{2},\varphi)\, d\varphi.
\end{aligned}$$

Reunaviivalla ∂A on $\theta = \frac{\pi}{2}$ ja $\vec{F}\cdot d\vec{r} = \vec{F}\cdot(R\, d\varphi\, \vec{e}_\varphi) = RF_\varphi\, d\varphi$, joten todetaan kaava (3) päteväksi puolipallolle. $\square$

Esimerkki 2 Toisessa esimerkkitapauksessa tarkastellaan pintaa, joka saadaan poistamalla suorakulmaisen särmiön ulkopinnasta kaksi vierekkäistä sivutahkoa, vrt. kuvio.

Tässä tapauksessa pinta koostuu tasopinnoista A_i, joille kullekin pätee Stokesin kaava (3), eli

$$\int_{A_i} (\nabla \times \vec{F}) \cdot d\vec{a} = \oint_{\partial A_i} \vec{F} \cdot d\vec{r}.$$

Kun normaalin suunta jokaisella osalla valitaan osoittamaan särmiön ulkopuolelle, niin nähdään, että

$$\begin{aligned}\int_A (\nabla \times \vec{F}) \cdot d\vec{a} &= \sum_i \int_{A_i} (\nabla \times \vec{F}) \cdot d\vec{a} \\ &= \sum_i \oint_{\partial A_i} \vec{F} \cdot d\vec{r} \\ &= \int_{\partial A} \vec{F} \cdot d\vec{r},\end{aligned}$$

sillä viivaintegraalit osien yhteisten särmien $\partial A_i \cap \partial A_j$ yli kumoutuvat summauksessa, vrt. kuvio. Siis kaava (3) on pätevä tässäkin tapauksessa. Samaan tulokseen tultaisiin myös poistamalla särmiöstä mitkä sivutahkot tahansa (1–5 kpl). □

Esimerkki 3 Kolmantena esimerkkinä tarkastellaan pintaa, joka saadaan lieriövaipasta leikkaamalla vaippa poikki, kiertämällä toista päätä 180°, ja liittämällä päät jälleen yhteen.

Tuloksena oleva pinta on nk. *Möbiuksen nauha*†. Möbiuksen nauhalle Stokesin kaava (3) ei ole pätevä. Sen sijaan kaava on kyllä voimassa, jos liitossauma jätetään avoimeksi ja viivaintegraalit sauman kummallakin puolella huomioidaan kaavassa (3). Möbiuksen nauhan tapauksessa nämä saumaintegraalit lasketaan samansuuntaisina, jolloin ne eivät saumaa yhteen liitettäessä kumoudu, toisin kuin lieriöpinnan tapauksessa, vrt. kuvio.

Lieriö Möbius

Möbiuksen nauha on siis erimerkki ei–suunnistuvasta pinnasta, jolle Stokesin kaava (3) ei päde. Möbiuksen nauha on itse asiassa *yksipuolinen* pinta, jolla normaalin $\vec{n}$ suuntaa ei voida valita ristiriidattomasti niin, että se muuttuisi jatkuvasti pintaa pitkin kuljettaessa. (Jos pinta leikataan auki, niin liimauskohdassa normaali on epäjatkuva.) $\square$

Stokesin lause voidaan yksinkertaisen pinnan tapauksessa todistaa pinnan parametrisaatiota käyttäen, mutta todistus on tällöin melko tekninen ja vaatii suhteellisen voimakkaita säännöllisyysoletuksia. Läpinäkyvämpi ja yleispätevämpi todistus saadaan, kun pintaa approksimoidaan kolmion muotoisilla tasopinnoilla, samalla tavoin kuin pinta-alamitan määrittelyssä (vrt. Luku XIII.7). Koska Stokesin tasokaava pätee jokaiselle tasokolmiopinnalle, seuraa summaamalla, että Stokesin kaava (3) pätee myös kolmioista muodostetulle pinnan A approksimaatiolle — edellyttäen, että summauksessa viivaintegraalit yli vierekkäisten kolmioiden yhteisten sivujen kumoutuvat. Suunnistuvuusehto on juuri tässä. Sikäli kuin pinta on suunnistuva, saadaan kolmioverkkoa tihentämällä raja-arvotulos (3) edellyttäen, että vektorikenttä $\vec{F}$, pinta A (pinnan parametrisaatio), ja reunaviiva ∂A ovat riittävän säännöllisiä. Vektorikentän osalta riittää, että se on jatkuvasti derivoituva (todellisuudessa riittää hieman vähempikin), eivätkä pintaa A koskevat olettamuksetkaan ole käytännön kannalta kovin rajoittavia, vrt. esimerkit edellä.

†**August Möbius** (1790–1868) oli saksalainen matemaatikko.

Stokesin lauseella, kuten Gaussin lauseellakin, on useampiulotteiset vastineensa. Nämä kuuluvat matematiikan alaan nimeltä *differentiaaligeometria*. Tällaisiin laajempiin yhteyksiin sijoittuu luontevimmin myös Stokesin lauseen tarkempi todistus, joka tässä sivuutetaan.

Esimerkki 4 Kappale liikkuu voimakentässä

$$\vec{F} = \frac{F}{a}\,(y\,\vec{i} - 2z\,\vec{j} + 3x\,\vec{k}\,)$$

pitkin ympyrärataa, joka on pallopinnan $x^2 + y^2 + z^2 = a^2$ ja tason $x + y + z = 0$ leikkausviiva. Määritä Stokesin lauseen avulla voimakentän tekemä työ yhden kierroksen aikana, kun kiertosuunta on pisteestä $(10, 10, 10)$ katsottuna myötäpäivään.

Ratkaisu Kappaleen liikerata on ∂A, missä A on origokeskinen, a-säteinen kiekko tasolla $x + y + z = 0$. Stokesin lauseen mukaan

$$W = \oint_{\partial A} \vec{F} \cdot d\vec{r} = \int_A (\nabla \times \vec{F}) \cdot \vec{n}\, dS.$$

Tässä on

$$\nabla \times \vec{F} = \frac{F}{a} \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial_x & \partial_y & \partial_z \\ y & -2z & 3x \end{vmatrix} = \frac{F}{a}(2\vec{i} - 3\vec{j} - \vec{k}).$$

Oletettua kiertosuuntaa vastaa $\vec{n} = -\frac{1}{\sqrt{3}}(\vec{i} + \vec{j} + \vec{k})$, joten $(\nabla \times \vec{F}) \cdot \vec{n} = \frac{2F}{\sqrt{3}\,a}$. Siis

$$W = \int_A \frac{2F}{\sqrt{3}\,a}\, dS = \frac{2F}{\sqrt{3}\,a}\mu(A) = \frac{2F}{\sqrt{3}\,a}\,\pi a^2 = \underline{\underline{(2\pi/\sqrt{3})\,Fa}}. \quad \square$$

Yleistetty Stokesin lause

Kuten Gaussin lause, voidaan myös Stokesin lause yleistää niin, että pistetulon tilalla voi olla yleisempi vektoritulo, eli joko ristitulo tai skalaarin ja vektorin kertolasku. Yleistetty Stokesin kaava on

$$\boxed{\oint_{\partial A} d\vec{r} * \vec{F} = \int_A (d\vec{a} \times \nabla) * \vec{F}.}$$

Tässä on jälleen asetettava $\vec{F} \hookrightarrow f$, kun $*$ = 'tyhjä'. Yleistetty kaava pätee samanlaisin edellytyksin kuin peruskaava (3), ja sen voi myös perustella samaan tapaan kuin edellä, eli lähtemällä kaavan tasomuodosta.

HARJOITUSTEHTÄVIÄ

1. Olkoon $A = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \le 1 \ \& \ x \ge 0\}$. Laske seuraavat kiertointegraalit yli ∂A:n (vastapäivään kiertäen) sekä suoraan viivaintegraaleina että muuntamalla ne ensin tasointegraaleiksi (yleistettyä) Stokesin kaavaa käyttäen.
   a) $\oint_{\partial A} [\, y^2\, dx + (x+y^2)\, dy\,]$ b) $\oint_{\partial A} (x+y^2)\, d\vec{r}$
   c) $\oint_{\partial A} (xy^2\vec{i} + |y|\vec{j}\,) \cdot d\vec{r}$ d) $\oint_{\partial_A} (y^3\vec{i} - x\vec{j}\,) \times d\vec{r}$

2. Laske $\oint_{\partial A} \vec{F} \cdot d\vec{r}$ a) suoraan viivaintegraalina, b) Stokesin lauseen avulla, kun $\vec{F} = x\vec{k}$ ja $A \subset S$ vastaa parametrien arvoja $(u,v) \in [0,1] \times [0,3]$ pinnalla $S:\ x = 8u^2,\ y = v^2,\ z = 4uv$.

3. Näytä Stokesin lauseen avulla, että jos S on pallon $K:\ x^2+y^2+z^2 = R^2$ ja tason $T:\ x+y+z = 0$ leikkauskäyrä, niin $\oint_S (y\, dx + z\, dy + x\, dz) = \pm\sqrt{3}\pi R^2$.

4. Laske pintaintegraali $\int_A \nabla \times (\vec{k} \times \vec{r}\,) \cdot d\vec{a}$, kun S, $A \subset S$ ja S:n yksikkönormaalivektorin $\vec{n} = n_x\vec{i} + n_y\vec{j} + n_z\vec{k}$ suunta on annettu seuraavilla ehdoilla.
   a) $S:\ z = 0, \quad A:\ x^2+y^2 \le 1, \quad n_z \ge 0$
   b) $S:\ x^2+y^2+z^2 = 1, \quad A:\ z \ge 0, \quad n_z \ge 0$
   c) $S:\ x^2+y^2+z^2 = 1, \quad A:\ z \le 0, \quad n_z \le 0$

5. Halutaan laskea kiertointegraali $\oint_S \vec{F} \cdot d\vec{r}$ yli suljetun käyrän $S:\ \vec{r} = \cos t\,\vec{i} + \sin t\,\vec{j} + \sin 2t\,\vec{k},\ t \in [0, 2\pi]$, kun $\vec{F} = (e^x - y^3)\vec{i} + (e^y + x^3)\vec{j} + e^z\vec{k}$. Laske integraali Stokesin kaavan avulla huomioimalla, että S on pinnalla $z = 2xy$.

6. Olkoon S lieriön $K:\ (x-1)^2 + 4y^2 = 16$ ja tason $T:\ 2x+y+z = 3$ leikkauskäyrä ja $\vec{F} = [\sin(x^2) + y^2 + z^2]\vec{i} + (2xy+z)\vec{j} + (xz+2yz)\vec{k}$. Laske $\oint_S \vec{F} \cdot d\vec{r}$, kun S:n suunnistus on vastapäivään kaukaa positiiviselta z-akselilta katsoen.

7. Valitse pinnalle $A:\ z = 9 - x^2 + y^2,\ x^2+y^2 \le 9$ suunnistus ja laske $\int_A \nabla \times \vec{F} \cdot d\vec{a}$, kun $\vec{F} = -y\vec{i} + x^2\vec{j} + z\vec{k}$.

8. Laske Stokesin kaavan avulla pintaintegraali
$$\int_A \nabla \times [(x - z^2)\vec{i} + (x^3 + z)\vec{j} + xy\vec{k}] \cdot d\vec{a},$$
   missä A on kartiopinnan $S:\ x^2+y^2 = (z-1)^2$ tasojen $z = 0$ ja $z = 1$ väliin jäävä osa.

9. Sähkövirran tiheys $\vec{J}$ ja magneettikentän voimakkuus $\vec{H}$ ovat lieriökoordinaatistossa muotoa $\vec{J} = J(r)\vec{e}_z$, $\vec{H} = H(r)\vec{e}_\varphi$. Määrää $H(r)$ funktion $J(r)$ avulla käyttämällä Maxwellin yhtälöä $\nabla \times \vec{H} = \vec{J}$ ja Stokesin kaavaa sopivasti valitulla pinnalla A.

10. Näytä yleistetty Stokesin kaava oikeaksi siinä tapauksessa, että $\vec{F}$ on avaruuden vektorikenttä ja $A \subset T$, missä T on avaruustaso.

11. Olkoon S yksinkertainen suljettu käyrä avaruustasolla T, jonka yksikkönormaalivektori on $\vec{n} = a\vec{i} + b\vec{j} + c\vec{k}$. Näytä, että S:n sisään jäävän pinnan A ala on

$$\mu(A) = \frac{1}{2}\left|\oint_S \big[(bz - cy)dx + (cx - az)dy + (ay - bx)dz\big]\right|.$$

12. Määritellään pinta A parametrisaatiolla

$$A: \quad \begin{cases} x = \cos 2u(a + bv\sin nu), \\ y = \sin 2u(a + bv\sin nu), \\ z = bv\cos nu, \end{cases} \qquad (u,v) \in \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right) \times \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right],$$

missä $0 < b < a$ ja $n \in \mathbb{N}$. Millainen on A:n reuna ∂A? Millä n:n arvoilla A on suunnistuva pinta?

13. (*) Laske kiertointegraali $\oint_S \vec{F} \cdot d\vec{r}$, kun $\vec{F} = ye^x\vec{i} + (x^2 + e^x)\vec{j} + e^z\vec{k}$ ja S on suljettu parametrinen käyrä

$$\vec{r} = (2\cos t - 3\sin t)\vec{i} + (2 + 3\cos t + 6\sin t)\vec{j} + (1 + 6\cos t - 2\sin t)\vec{k},$$
$$t \in [0, 2\pi].$$

XIV.5 Pyörteetön vektorikenttä

Jos $\vec{F}$ on fysikaalinen virtaus-, sähkö-, magneetti- ym. kenttä, sanotaan kenttää $\vec{\omega} = \nabla \times \vec{F}$ kentän pyörrekentäksi (vrt. Luku XI.4). Stokesin lause ilmaisee yhteyden kentän $\vec{F}$ ja pyörrekentän $\vec{\omega}$ välillä:

$$\oint_{\partial A} \vec{F} \cdot d\vec{r} = \int_A \vec{\omega} \cdot d\vec{a}.$$

Tällä yhteydellä on paljon käyttöä esimerkiksi sähkömagneetikassa, jossa pyörrekentät ovat näkyvässä roolissa jo perusyhtälöissä (Maxwellin yhtälöissä, vrt. Luku XI.4). Jos $\vec{F}$ on pyörteetön kenttä, eli jos $\vec{\omega} = \vec{0}$, antaa Stokesin lause tuloksen

$$\oint_{\partial A} \vec{F} \cdot d\vec{r} = 0. \tag{1}$$

Luvun XIV.1 termein tämä voidaan lukea: Pyörteettömän vektorikentän tekemä työ suljettua polkua pitkin $= 0$. Sanotaan, että kenttä $\vec{F}$ on tällöin *konservatiivinen* (energian säilyttävä).

Jos $\vec{F}$ on gradienttikenttä, ts. $\vec{F} = -\nabla u$, niin $\vec{F}$ on pyörteetön. Tulos (1) on tällöin voimassa — ja tiedettiin ilman Stokesin lausettakin, sillä gradienttikentän tekemä työ $= 0$, jos polun alku- ja loppupisteet ovat samat, vrt. Luku XIV.1. — Mutta entä jos kysytäänkin toisin päin: Millainen on riittävän säännöllinen, mutta muuten mahdollisimman yleinen vektorikenttä $\vec{F}$, joka on joko (a) konservatiivinen tai (b) pyörteetön? Seuraavassa keskitytään vastaamaan näihin kysymyksiin. Tarkastellaan aluksi tason vektorikenttiä.

Olkoon $A \subset \mathbb{R}^2$ <u>avoin</u> joukko. Sanotaan, että tällainen joukko on *yhtenäinen* (engl. connected), jos mielivaltaiset kaksi A:n pistettä ovat yhdistettävissä jatkuvalla, suoristuvalla parametrisella käyrällä, jonka päätepisteet ovat mainitut pisteet ja joka on kokonaisuudessaan A:ssa. Jatkossa sanotaan avointa ja yhtenäistä joukkoa *alueeksi* (engl. domain)[†].

Lause XIV.5.1 Jos tason vektorikenttä $\vec{F}$ on jatkuva ja konservatiivinen alueessa $A \subset \mathbb{R}^2$, niin $\vec{F}$ on gradienttikenttä, ts. $\vec{F} = -\nabla u$, missä u on A:ssa jatkuvasti derivoituva.

<u>Todistus</u> Olkoon $\vec{r}_0 \mathrel{\widehat{=}} (x_0, y_0) \in A$ kiinteä ja $\vec{r} \mathrel{\widehat{=}} (x, y) \in A$ muuttuva A:n piste. Olkoon edelleen p yksinkertainen, suoristuva parametrinen käyrä (polku) $\vec{r} =$

[†]Alue määritellään hieman yleisemmin joukkona $B = A \cup S$, missä A on avoin ja yhtenäinen ja $S \subset \partial A$. Jos $S = \emptyset$, niin $B = A$ on *avoin alue*. Jos $S = \partial A$, niin $B = \overline{A}$ on *suljettu alue*.

$\vec{r}(t)$, $t \in [a,b]$, jolle $\vec{r}(a) = \vec{r}_0$ ja $\vec{r}_b = \vec{r} \mathrel{\widehat{=}} (x,y)$. Määritellään

$$u(p,x,y) = -\int_p \vec{F} \cdot d\vec{r}.$$

Jos u ei riipu polun p valinnasta, ts. $u(p,x,y) = u(x,y)$, niin jatkamalla polkua pisteestä $\vec{r}$ seuraa jatkuvuusoletuksesta, että

$$u(\vec{r} + \Delta\vec{r}) = -\vec{F}(\vec{r}) \cdot \Delta\vec{r} + o(|\Delta\vec{r}|).$$

Tämän mukaan u on differentioituva ja $\nabla u = -\vec{F}$. Riittää siis osoittaa, että $u(p,x,y)$ riippuu vain polun päätepisteistä. Olkoon p_1 ja p_2 kaksi polkua joilla on sama alkupiste (x_0,y_0) ja sama loppupiste (x,y). Jos vastaavat geometriset käyrät ovat S_1 ja S_2, niin

$$S_1 \cup S_2 = \bigcup_i S_i, \quad S_i \subset A,$$

missä S_i:t ovat suljettuja käyriä (vrt. kuvio).

Tällöin

$$\int_{p_1} \vec{F} \cdot d\vec{r} - \int_{p_2} \vec{F} \cdot d\vec{r} = \sum_i \oint_{S_i} \vec{F} \cdot d\vec{r}.$$

Koska $\vec{F}$ on konservatiivinen, on $\oint_{S_i} \vec{F} \cdot d\vec{r} = 0$, joten

$$\int_{p_1} \vec{F} \cdot d\vec{r} = \int_{p_2} \vec{F} \cdot d\vec{r}.$$

Siis $u(p,x,y) = u(x,y)$ on polun valinnasta riippumaton. $\square$

Sanotaan, että alue $A \subset \mathbb{R}^2$ on *yhdesti yhtenäinen* (engl. simply connected), jos pätee $B \subset A$ aina kun B on alue, jolle $\partial B \subset A$. Kun kirjoitetaan $S = \partial B$, niin ehdon hieman havainnollisempi muotoilu on: $A \subset \mathbb{R}^2$ on yhdesti yhtenäinen, jos jokainen suljettu käyrä $S \subset A$ voidaan kutistaa pisteeksi niin, että käyrä pysyy kutistuessaan koko ajan A:ssa. Yhdesti yhteinäinen joukko on 'järvi ilman saaria'.

yhdesti yhtenäinen

ei yhdesti yhtenäinen

Seuraava Stokesin lauseen seurannaistulos on vektorikenttien teorian keskeisimpiä. Tuloksella on paljon käyttöä fysiikassa.

Lause XIV.5.2 Jos tason vektorikenttä $\vec{F}$ on jatkuvasti derivoituva ja pyörteetön ($\operatorname{rot}\vec{F} = 0$) yhdesti yhtenäisessä alueessa A, niin $\vec{F}$ on gradienttikenttä.

Todistus Kentän potentiaali määritetään polkuintegraalina kuten edellä, jolloin riittää jälleen osoittaa, että integraalin arvo riippuu vain polun päätepisteistä. Tämän todistamiseksi valitaan jälleen kaksi polkua, joilla on samat päätepisteet, jolloin (vrt. todistus edellä)

$$\int_{p_1} \vec{F} \cdot d\vec{r} - \int_{p_2} \vec{F} \cdot d\vec{r} = \sum_i \oint_{S_i} \vec{F} \cdot d\vec{r}.$$

Nytkin on oikea puoli $= 0$, sillä oletetun yhdesti yhtenäisyyden perusteella käyrien S_i sisäänsä sulkemat alueet A_i sisältyvät joukkoon A, jolloin Stokesin tasokaavan (edellinen luvun kaava (1)) ja $\vec{F}$:n pyörteettömyyden perusteella

$$\oint_{S_i} \vec{F} \cdot d\vec{r} = \int_{A_i} \operatorname{rot}\vec{F}\, dxdy = 0. \quad \square$$

Lauseiden XIV.5.1–XIV.5.2 todistuksien perusteella gradienttikentän potentiaali saadaan lasketuksi työintegraalina:

$$u(\vec{r}) = -\int_{p:\,\vec{r}_0 \to \vec{r}} \vec{F} \cdot d\vec{r}.$$

Laskukaavassa voidaan polku p voidaan valita vapaasti, kunhan se pysyy alueessa A (jossa $\vec{F}$ on määritelty). Jos A:n geometria ei aseta rajoituksia, voidaan polku valita esimerkiksi seuraamaan koordinaattiakselien suuntaisia suoria, jolloin potentiaalin laskukaavaksi tulee

$$\begin{aligned}\vec{F}(x,y) &= F_1(x,y)\vec{i} + F_2(x,y)\vec{j}\ :\\ u(x,y) &= -\int_{x_0}^{x} F_1(t,y_0)\,dt - \int_{y_0}^{y} F_2(x,t)\,dt.\end{aligned}$$

Esimerkki 1 Onko vektorikenttä

$$\vec{F}(x,y) = (x^3 - y^3 + \cos x)\vec{i} - (3xy^2 + e^{-y})\vec{j}$$

gradienttikenttä?

Ratkaisu Koska

$$\operatorname{rot}\vec{F} = -\partial_x(3xy^2 + e^{-y}) - \partial_y(x^3 - y^3 + \cos x) = 0,$$

niin $\vec{F}$ on gradienttikenttä $\mathbb{R}^2$:ssa (Lause XIV.5.2). Valinnalla $(x_0, y_0) = (0,0)$ potentiaaliksi saadaan

$$\begin{aligned}u(x,y) &= -\int_0^x (t^3 + \cos t)\,dt + \int_0^y (3xt^2 + e^{-t})\,dt\\ &= -\frac{1}{4}x^4 - \sin x + xy^3 - e^{-y} + 2.\end{aligned}$$

Yleinen potentiaalifunktio on

$$u(x,y) = -\frac{1}{4}x^4 + xy^3 - \sin x - e^{-y} + C \quad (C = \text{vakio}). \quad \square$$

Esimerkki 2 Polaarikoordinaatistossa on määritelty vektorikenttä $\vec{F} = \frac{1}{r}\,\vec{e}_\varphi$. Tutki, onko $\vec{F}$ gradienttikenttä alueessa A, kun
a) $A = \{\,(x,y) \in \mathbb{R}^2 \mid x > 0 \ \ \&\ \ y > 0\,\}$,
b) $A = \{\,(x,y) \in \mathbb{R}^2 \mid (x,y) \neq (0,0)\,\}$.

Ratkaisu Polaarikoordinaatistossa on (vrt. Luku XI.4)

$$\vec{F} = F_r\vec{e}_r + F_\varphi\vec{e}_\varphi\ : \quad \operatorname{rot}\vec{F} = \frac{1}{r}\,\partial_r(rF_\varphi) - \frac{1}{r}\,\partial_\varphi F_r.$$

Tässä on $F_r = 0$ ja $F_\varphi = 1/r$, joten $\vec{F}$ on pyörteetön jokaisessa alueessa, joka ei sisällä origoa.

a) Tässä A on yhdesti yhtenäinen, joten $\vec{F}$ on gradienttikenttä. Valitsemalla $r_0 > 0$ voidaan potentiaali $u(x,y) = v(r,\varphi)$ laskea polkuintegraalina

$$v(r,\varphi) = -\int_{r_0}^{r} F_r(t,0)\,dt - \int_0^{\varphi} F_\varphi(r,t)\,r\,dt \;=\; -\varphi$$

$$\Rightarrow \quad u(x,y) \;=\; \underline{\underline{-\mathrm{Arcsin}\left(\frac{y}{\sqrt{x^2+y^2}}\right)}}.$$

b) Tässä A ei ole yhdesti yhtenäinen, joten tutkitaan erikseen, onko $\vec{F}$ konservatiivinen. Valitaan polku $p : \vec{r} = \vec{r}(\varphi)$, $0 \le \varphi \le 2\pi$ siten, että $\vec{r}(\varphi)$ kiertää r_0-säteisen ympyräviivan:

$$\vec{r}(\varphi) = r_0\,\vec{e}_r, \quad r_0 > 0,\ 0 \le \varphi \le 2\pi.$$

Tällöin polun alku- ja loppupisteet ovat samat, $\vec{r}(\varphi) \mathrel{\widehat{=}} (x,y) \in A$ jokaisella φ ja

$$\int_p \vec{F}\cdot d\vec{r} = \int_0^{2\pi} \frac{1}{r_0}\cdot r_0\,d\varphi = 2\pi \neq 0,$$

Siis $\vec{F}$ ei ole A:ssa konservatiivinen eikä näin muodoin myöskään gradienttikenttä A:ssa. $\square$

Avaruuden pyörteetön vektorikenttä

Vertaamalla Lauseen XIV.5.1 todistusta Stokesin lauseen kolmiulotteiseen versioon nähdään, että todistus (ja siis myös Lauseen XIV.5.1 väittämä) on sellaisenaan pätevä myös avaruuden vektorikentälle. Sikäli kuin alueen geometria ei aseta rajoituksia, voidaan kolmiulotteisen, konservatiivisen vektorikentän $\vec{F} = F_1(x,y,z)\vec{i} + F_2(x,y,z)\vec{j} + F_3(x,y,z)\vec{k}$ potentiaali laskea esimerkiksi polkuintegraalina

$$u(x,y,z) = -\int_{x_0}^{x} F_1(t,y_0,z_0)\,dt - \int_{y_0}^{y} F_2(x,t,z_0)\,dt - \int_{z_0}^{z} F_3(x,y,t)\,dt.$$

Myös Lause XIV.5.2 yleistyy kolmeen ulottuvuuteen, kunhan yhdesti yhtenäisyys määritellään sopivasti. Tarkastellaan yksinkertaisia, suljettuja, suoristuvia käyriä $S \subset A$, joille on löydettävissä Stokesin lauseen ehdot täyttävä pinta $B \subset \mathbb{R}^3$ siten, että $\partial B = S$. Jos jokaiselle tällaiselle käyrälle on pinta B vielä määrättävissä siten, että $B \subset A$, niin aluetta A sanotaan yhdesti yhtenäiseksi. Esimerkiksi

$$A = \{\,(x,y,z) \in \mathbb{R}^3 \mid a^2 < x^2+y^2+z^2 < R^2\,\} \quad (0 \le a < R)$$

on yhdesti yhtenäinen (!). Sen sijaan

$$A = \{\, (x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 < R^2 \ \ \&\ \ x^2+y^2 > a^2 \,\} \quad (0 \le a < R)$$

ei ole yhdesti yhtenäinen, sillä jos

$$S = \{\, (x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 = b^2 \ \ \&\ \ z = 0 \,\},$$

niin $S \subset A$, kun $a < b < R$, mutta jos $B \subset \mathbb{R}^3$ on pinta, niin ehdot $S = \partial B$ ja $B \subset A$ eivät voi yhtä aikaa toteutua.

Esimerkki 3 Pallokoordinaatistossa on määritelty vektorikenttä

$$\vec{F}(r,\theta,\varphi) = \frac{1}{r^3}(2\cos\theta\,\vec{e}_r + \sin\theta\,\vec{e}_\theta), \quad r \neq 0.$$

Määrää kentän potentiaali, jos sellainen on.

<u>Ratkaisu</u> Tässä on $\vec{F} = F_r\vec{e}_r + F_\theta\vec{e}_\theta + F_\varphi\vec{e}_\varphi$, missä

$$F_r = 2r^{-3}\cos\theta, \quad F_\theta = r^{-3}\sin\theta, \quad F_\varphi = 0,$$

joten (ks. Luku XI.5)

$$\nabla \times \vec{F} = 0\,\vec{e}_r + 0\,\vec{e}_\theta + \frac{1}{r}[\partial_r(rF_\theta) - \partial_\theta F_r]\vec{e}_\varphi = \vec{0} \quad (r \neq 0).$$

Koska origon poissulkeminen ei aiheuta yhtenäisyysongelmia kolmessa dimensiossa, niin $\vec{F}$ on gradienttikenttä. Potentiaali voidaan laskea ottamalla lähtöpisteeksi esimerkisi $\vec{r}_0 = (r_0, 0, 0)$, $r_0 > 0$, ja etenemällä koordinaattiviivojen suuntaisia polkuja:

$$\begin{aligned}
u(r,\theta,\varphi) &= -\int_{r_0}^{r} F_r(t,0,0)\,dt - \int_0^\theta F_\theta(r,t,0)r\,dt - \int_0^\varphi F_\varphi(r,\theta,t)r\sin\theta\,dt \\
&= -\int_{r_0}^{r} 2t^{-3}\,dt - \int_0^\theta r^{-2}\sin t\,dt \\
&= \Big|_{r_0}^{r} t^{-2} + r^{-2}\Big|_0^\theta \cos t \\
&= r^{-2}\cos\theta + \text{vakio } (= -r_0^{-2}).
\end{aligned}$$

Potentiaaliksi kelpaa siis $u(r,\theta,\varphi) = \underline{\underline{r^{-2}\cos\theta}}$. $\square$

HARJOITUSTEHTÄVIÄ

1. Totea vektorikenttä $\vec{F}(x,y) = (y+2x)\vec{i} + x\vec{j}$ pyörteettömäksi. Laske tätä havaintoa hyödyksi käyttäen polkuintegraali $\int_p \vec{F}\cdot d\vec{r}$, missä p kulkee pisteestä $(1,0)$ origon kautta pisteeseen $(1,2)$ pitkin ympyrän kaarta.

2. Määritä funktio f siten, että $f(0)=1$ ja tason vektorikenttä
$$\vec{F}(x,y) = f(y)[2xy\vec{i} + x^2(y+1)\vec{j}\,]$$
on pyörteetön. Mikä tällöin on kentän potentiaali?

3. Osoita seuraavat vektorikentät koko avaruudessa pyörteettömiksi ja määritä kenttien potentiaalit.
a) $z\vec{i} + z\cos y\vec{j} + (x+\sin y)\vec{k}$ b) $e^x \sin y\vec{j} + e^x\cos y\vec{j} + z\vec{k}$
c) $2xyz\vec{i} + x^2z\vec{j} + x^2y\vec{k}$ d) $-2xe^{-y}\vec{i} + (x^2e^{-y} + \sin z)\vec{j} + y\cos z\vec{k}$
e) $(x^2+y^2+z^2)(x\vec{i} + y\vec{j} + z\vec{k}) = r^2\vec{r}$

4. Vektorikentästä $\vec{F}$ tiedetään, että se on koko avaruudessa pyörteetön ja muotoa $\vec{F}(x,y,z) = xy^2\vec{i} + (\cos z + x^2y)\vec{j} + yf(z)\vec{k}$. Määrää kentän skalaaripotentiaali ja tämän avulla polkuintegraali $\int_p \vec{F}\cdot d\vec{r}$, missä p:n alkupiste on $(0,0,0)$ ja loppupiste $(2,-1,3)$.

5. (*) (Häiriö kentässä) Tason vektorikenttään $\vec{U}_0 = u_0\vec{i}$ (u_0 = vakio) tuodaan este
$$A = \{(x,y)\in\mathbb{R}^2 \mid x^2+y^2 \le R^2\},$$
jolloin kenttä häiriintyy muotoon $\vec{U} = u(x,y)\vec{i} + v(x,y)\vec{j}$. Häiriintyneestä kentästä $\vec{U}$ tiedetään, että se on sekä lähteetön että pyörteetön esteen ulkopuolella, ts. $\operatorname{div}\vec{U} = \operatorname{rot}\vec{U} = 0$ muualla kuin A:ssa. Lisäksi tiedetään, että $\vec{U}\cdot\vec{n} = 0$ reunalla ∂A (eli kun $x^2+y^2=R^2$, $\vec{n}$ = ∂A:n normaali) ja että $\vec{U}(x,y) \to \vec{U}_0$, kun $x^2+y^2\to\infty$. Ongelmana on löytää nämä ehdot täyttävä kenttä $\vec{U}$. Koska A:n ulkopuolinen alue ei ole yhdesti yhtenäinen, ei $\vec{U}$ välttämättä ole gradienttikenttä (vaikka onkin pyörteetön).

a) Näytä, että problemalla on ratkaisu, jonka napakoordinaattimuoto on
$$\vec{U} = -\nabla f(r)\cos\varphi + g(r)\vec{e}_\varphi$$
eli määritä funktiot $f(r), g(r)$ mahdollisimman yleisessä muodossa niin, että annetut ehdot toteutuvat. (Muita ratkaisuja ei ongelmalla ole).
b) Totea, että em. ratkaisusta tulee yksikäsitteinen, kun asetetaan lisäehto muotoa $v(R,0) = v_0$ (tässä karteesiset koordinaatit).