

Solmu

Matematiikkalehti
1/2007

<http://solmu.math.helsinki.fi/>



Solmu 1/2007

ISSN 1458-8048 (Verkkolehti)

ISSN 1459-0395 (Painettu)

Matematiikan ja tilastotieteen laitos

PL 68 (Gustaf Hållströmin katu 2b)

00014 Helsingin yliopisto

<http://solmu.math.helsinki.fi/>

Päätoimittaja:

Matti Lehtinen, dosentti, Maanpuolustuskorkeakoulu

Toimitussihteerit:

Mika Koskenoja, tohtoriassistentti, Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

Antti Rasila, tutkija, Matematiikan laitos, Teknillinen korkeakoulu

Sähköposti *toimitus@solmu.math.helsinki.fi*

Toimituskunta:

Pekka Alestalo, dosentti, Matematiikan laitos, Teknillinen korkeakoulu

Heikki Apiola, dosentti, Matematiikan laitos, Teknillinen korkeakoulu

Aapo Halko, FT, Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

Ari Koistinen, FM, Helsingin ammattikorkeakoulu Stadia

Marjatta Näättänen, dosentti, Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

Tommi Sottinen, dosentti, Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

Graafinen avustaja *Marjaana Beddard*

Yliopistojen ja korkeakoulujen yhteyshenkilöt:

Virpi Kauko, yliopistonopettaja, virpi@maths.jyu.fi

Jyväskylän Avoin yliopisto

Jorma K. Mattila, professori, jorma.mattila@lut.fi

Sovelletun matematiikan laitos, Lappeenrannan teknillinen yliopisto

Jorma Merikoski, professori, jorma.merikoski@uta.fi

Matematiikan, tilastotieteen ja filosofian laitos, Tampereen yliopisto

Kalle Ranto, assistentti, kalle.ranto@utu.fi

Matematiikan laitos, Turun yliopisto

Matti Nuortio, opiskelija, mnuortio@paju.oulu.fi

Matematiikan tieteen laitos, Oulun yliopisto

Timo Tossavainen, lehtori, timo.tossavainen@joensuu.fi

Savonlinnan opettajankoulutuslaitos, Joensuun yliopisto

Numeroon 2/2007 tarkoitetut kirjoitukset pyydämme lähettämään 15.4.2007 mennessä.

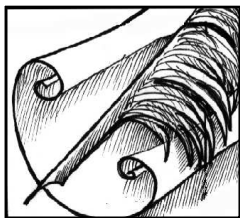
Kiitämme taloudellisesta tuesta Jenny ja Antti Wihurin rahastoa.

Huom! Solmun paperiversio postitetaan vain niihin kouluihin, jotka ovat sitä erikseen pyytäneet. Toivomme, että lehteä kopioidaan kouluissa kaikille halukkaille.

Kansi: Gerardus Mercator

Sisällys

Pääkirjoitus: Saako oppilas tietää liikaa? (Matti Lehtinen).....	4
Mercatorin kartta (Antti Rasila)	5
Kaksi syntymäpäiväsankaria (Matti Lehtinen).....	10
Vektorit, matriisit, Händel ja vaalit (Heikki Apiola)	15
Yleistetyt funktiot eli distribuutiot (Jukka Liukkonen)	25
Kirja-arvostelu: Alkulukujen lumoissa (Matti Lehtinen)	30
Ratkaisu aikaisempaan tehtävään (Pekka Alestalo)	32



Saako oppilas tietää liikaa?

Näinä aikoina testataan jälleen sitäkin, mitä 12 vuoden matematiikan opinnoista on jäänyt mieleen. Ylioppilaskirjoitukset ovat suomalaisen yhteiskunnan vuotuisen merkkitapahtuma, ja ne herättävät aina ajatuksia ja keskustelua. Matematiikka on yksi ylioppilaskirjoitusaine, jonka ympäriltä on käyty mielipiteiden vaihtoa vuosikymmenestä toiseen. En nyt ala taittaa peistä matematiikan kokeen pakollisuuden puolesta, vaikka sitä lämpimästi kannatankin. Otan esiin tehtävien alueen ja arvostelun. Perimätieto, joka ei liene perätön (itse olen tämän kuullut pitkäaikaiselta Ylioppilastutkintolautakunnan matematiikan jaoston vetäjältä Lauri Myrberg-vainajalta) kertoo, että eräänä vuonna tehtävänä oli määrittää muotoa

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$$

oleva raja-arvo, missä f ja g olivat polynomeja, joille $f(a) = g(a) = 0$. Tehtävä oli tietenkin tarkoitettu ratkaistavaksi jakamalla jaettava ja jakaja tekijöihin, $f(x) = (x-a)f_1(x)$ ja $g(x) = (x-a)g_1(x)$, supistamalla $x-a$ ja laskemalla triviaali raja-arvo

$$\lim_{x \rightarrow a} \frac{f_1(x)}{g_1(x)} = \frac{f_1(a)}{g_1(a)}.$$

Eräs kokelas ratkaisi kuitenkin tehtävän käyttämällä erittäin klassista, tähän tehtävään mainiosti sopivaa l'Hospitalin sääntöä, jonka mukaan, jos $g'(a) \neq 0$, raja-arvo on sama kuin $f'(a)/g'(a)$. Ratkaisua ei hyväksytty. Kokelas – tai hänen opettajansa – ei ollut tähän arvosteluun tyytyväinen. Valitusprosessi eteni Eduskunnan oikeusasiamiehelle asti, joka juridisessa

viisaudessaan ratkaisi, että l'Hospitalin sääntö ei ollut legitiimi keino esillä olleeseen tehtävään.

Ylioppilastutkinnon arvostelu on – tai ainakin sen uskotaan olevan – edelleen tällainen. Erinäiset kokemukset Suomen oloissa jopa ylivertaisesti kilpailuissa menestyneiden nuorten jäämisestä ylioppilaskirjoituksissa eximian tasolle tuntuvat vahvistavan käsitystä. Sallittuja keinoja ovat perusoppiaineeseen sisältyvät totuudet. Toki muitakin tuloksia voi käyttää, jos niiden käytön perustelee. Mutta ajatellaanpa nuorta, joka harrastaa matematiikkaa ja tuntee sitä hyvin, laajemmin kuin koulukurssi edellyttää. On aika turhaa odottaa, että tällainen henkilö tahtooi täsmällisesti opiskella juuri sen koulutiedon, jotta hän kykenisi perustelemaan mahdollisesti itselleen hyvinkin tutut asiat, jotka nyt eivät satu kurssiin kuulumaan. Asiaa mutkistaa se, että sallitut apuvälineet, taulukkokirjat, sisältävät ainesta, jota ei löydy vakiintuneiden oppikirjojen sisällön yhteisestä osasta. Tällaisia ovat esimerkiksi erinäiset kolmionmittaamiseen kuuluvat kaavat, kuten Heronin kaava. Sallittua tietoa vai ei?

Mielestäni matematiikan taidon kokeisiin perustuvan arvioinnin – tapahtuu se sitten millä tasolla tahansa – olisi kannustettava matematiikan oikeaa osaamista. Matematiikka on toki täsmällistä, ja täsmällisyyttä sen esityksen tulee olla. Mutta oikeaa matematiikkaa ei sinänsä voi jakaa jonkin kaanonin mukaan hyväksyttävään ja kelpaamattomaan.

Matti Lehtinen

Pääkirjoitus



Mercatorin kartta

Antti Rasila

Matematiikan laitos, Teknillinen korkeakoulu

Matematiikasta yleistajuisesti kirjoittamisen tekee usein haasteelliseksi konkreettisten esimerkkien ja sovellusten puute, käsitteiden abstraktius sekä se, että lukijalla vain harvoin on aikaisempaa kokemusta kirjoituksessa käsiteltävistä asioista. Arkipäivän matematiikkaa esittelevät kirjoitukset keskittyvät yleensä kauppalaskujen kaltaisiin yksinkertaisiin sovelluksiin, jotka eivät ole matemaattisina ongelmina mielenkiintoisia.

Tässä kirjoituksessa tarkoitukseni on kertoa Mercatorin kartasta, jonka uskoisin olevan entuudestaan tuttu jokaiselle lukijalle. Se on maailmankartta, joka löytyy jokaisesta koulun maantiedonkirjasta. Aikanaan Mercatorin kartta mullisti merenkulun. Se on auttanut laivoja, lentokoneita ja valloitusarmeijoita löytämään perille vuosisatojen ajan. Eräille tämä kartta on eurooppalaisen hegemonian ja kolonialismin ajan ajattelutavan symboli. Jotkin maat ovat kieltäneet sen käytön kouluissa. Lisäksi Mercatorin kartta on puhtaan matematiikan kannalta mielenkiintoinen konstruktio ja sisältää lukuisia geometrisen funktioteorian ja differentiaaligeometrian keskeisistä ajatuksista. Hämmästyttävintä on, että kartta on peräisin vuodelta 1569, ajalta ennen differentiaali- ja integraalilaskentaa ja Descartesin analyyttistä geometriaa.

Oma kiinnostukseni Mercatorin karttaan heräsi kuul-

tuani aiheesta tieteellisessä konferenssissa pari vuotta sitten. Aihe ei esiinny maantiedon koulukursseissa – ehkä siksi, että se on matemaattisesti liian vaativa – eikä myöskään matematiikan kursseissa koulussa tai yliopistossa. Maanmittauksen alan kirjat esittävät konstruktion lopputuloksena olevan kaavan mutta eivät itse konstruktiota. Tämä on mielestäni vahinko, koska kartta on eräs parhaista esimerkeistä, jossa deduktiivista matemaattista päättelyä voidaan käyttää konkreettisen ja maallikollekin helposti esitettävän ongelman ratkaisemiksi. Konstruktio toimii myös havainnollisen johdantona moniin matemaattisiin ilmiöihin, kuten derivaattaan, differentiaaliyhtälöihin, konformikuvauxiin, Riemannin pintoihin ja epäeuklidiseen geometriaan.

Kartan piirtämisestä

lhanteellisessa tapauksessa maailmankartta olisi isometrinen. Tämä tarkoittaa sitä, että kartalla mitatut etäisyydet vastaavat mittakaavan suhteessa suoraan etäisyyksiä maastossa. Maapallon pyöreyydestä johtuen tällaisen kartan laatiminen on todellisuudessa mahdollista, paitsi tietenkin pallokarttana.

Koska pallokartat ovat epäkäytännöllisiä, tässä tarkastellaan ainoastaan tasoon piirrettyjä karttoja. Isomet-

risen kartan mahdottomuus voidaan havaita ajattelemalla pallon pinnalle piirrettyä ympyrää ja sen sädettä. Tässä keskipisteen oletetaan olevan myös pallon pinnalla. Tasossa ympyrän säteen ja kehän pituuden suhde on 2π . Kartan oletetusta isometrisyydestä johtuen saman tulisi siis olla totta myös pallon pinnalla, kun kahden pisteen välinen etäisyys ymmärretään lyhimmän pallon pinnalla olevan polun pituutena. Pallon pinnalla kuitenkin suhde on aina eri, vääristymisen suuruus riippuu ympyrän säteestä. Siksi tasoon piirretty kartta kuvaa aina pallon pinnan osaa jossakin mittasuhteita vääristävässä karttaprojektiossa, vaikka karttoihin liittyvästä käsitteestä mittakaava voisikin päätellä toisin. Virheen suuruus riippuu kartalla esitettävän alueen laajuudesta. Monissa karttaprojektioissa, mm. Mercatorin projektiossa, navat sijaitsevat ”äärettömän kaukana” mistä tahansa muusta pisteestä.

Gerardus Mercator

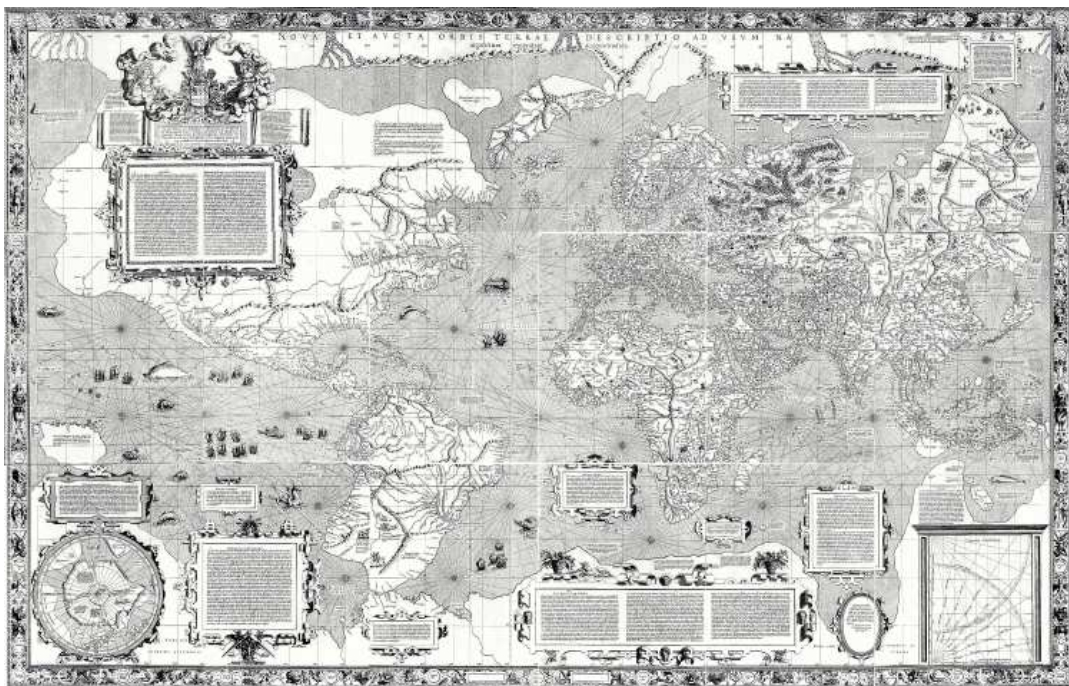
Gerard Kremer (5.3.1512 – 2.12.1594) syntyi Rupelmondessa, lähellä Antwerpeniä. Hänen vanhempansa olivat saksalaisia kauppiaita, jotka olivat paenneet Flanderiin uskonsotia ja protestantteihin kohdistuneita vainoja. Monien aikansa oppineiden tavoin hän latinisoi nimensä, ja se tunnetaan paremmin muodossa Gerardus Mercator. Mercator opiskeli Leuvenin yliopistossa kuuluisan humanistin ja näytelmäkirjailijan Macrope-diuksen oppilaana.

Hän kuitenkin pettyi teologian ja filosofian opintoihin, joiden tarjoamat selitykset eivät hänen mielestään

tydyttävällä tavalla kuvanneet maailmaa. Mercator matkusteli laajasti, mutta palasi kuitenkin takaisin Leuveniin, tällä kertaa aikakauden johtavan matemaatikon ja tähtitieteilijän Gemma Frisiuksen oppilaaksi. Frisiuksella oli muitakin kuuluisia oppilaita, kuten tähtitieteilijä Johannes Stadius ja myös okkultistina tunnettu Englannin kuningatar Elisabet ensimmäisen neuvonantaja John Dee. Frisiuksen mukaan on myös nimetty kraateri Kuussa.



KUVA: Gerardus Mercator (lähde: Wikipedia).



KUVA: Mercatorin maailmankartta vuodelta 1569 (lähde: Wikipedia).

Mercator perehtyi karttojen valmistamiseen työskennellessään Frisiuksen oppilaana, joka oli tehnyt monia merkittäviä alaan liittyviä keksintöjä, kuten kolmiomittauksen sekä menetelmän pituuspiirin määrittämiseksi, kun kellonaika tunnetaan tarkasti. Ensimmäinen Mercatorin itsenäisesti valmistama kartta on vuodelta 1537 ja esittää Palestiinaa. Tätä seurasi maailmankartta (1538) ja Flanderin kartta (1540). Vuonna 1544 Mercator sai tuomion kerettiläisyydestä, syynä tähän hänen runsas matkustelunsa sekä protestanttiset näkemyksensä. Seuraavat seitsemän kuukautta kuluivat vankilassa.

Vuonna 1554 Mercator muutti Duisburgiin, syitä tähän ei tiedetä. On arveltu, että hän saattoi lähteä Alankomaista uskonnollisista syistä, tai hän oli kuullut suunnitelmista uuden yliopiston perustamiseksi. Mercator opetti matematiikkaa Duisburgin yliopistossa ja samalla kehitti uutta karttaprojektiota. Vuonna 1569 hän julkaisi merkittävimmän työnsä, kuuluisan maailmankartan *Nova et Aucta Orbis Terrae Descriptio ad Usum Navigantium Emendate*, uusi ja tarkka maapallon esitys käytettäväksi merenkulussa. Mercator otti myös käyttöön sanan atlas merkitsemään useista lehdistä muodostettua karttaa, jonka lehdet yhdessä muodostavat kolmiulotteisen avaruuden pinnan esityksen. Vaikka sanalla nykyään tarkoitetaan usein mitä hyvänsä karttakirjaa, sitä käytetään edelleen (lähes) alkuperäisessä merkityksessä Riemannin pintojen teoriassa ja yleisemmin differentiaaligeometriassa.

Sylinteriprojektio

*Sylinteriprojektio*ksi kutsutaan sellaista karttaprojektiota, jossa leveyspiirit kuvautuvat kartalla vaakasuoriksi viivoiksi ja pituuspiirit pystysuoriksi viivoiksi. Esimerkki sylinteriprojektioista on keskisylinteriprojektio, joka saadaan kuvaamalla pallon pinnalla olevat pisteet pallon keskipisteen kautta piirrettyä suoraa pitkin palloa sivuavan lieriön pinnalle ja lopuksi levittämällä lieriö auki. Keskisylinteriprojektio ei kuitenkaan ole Mercatorin projektio.

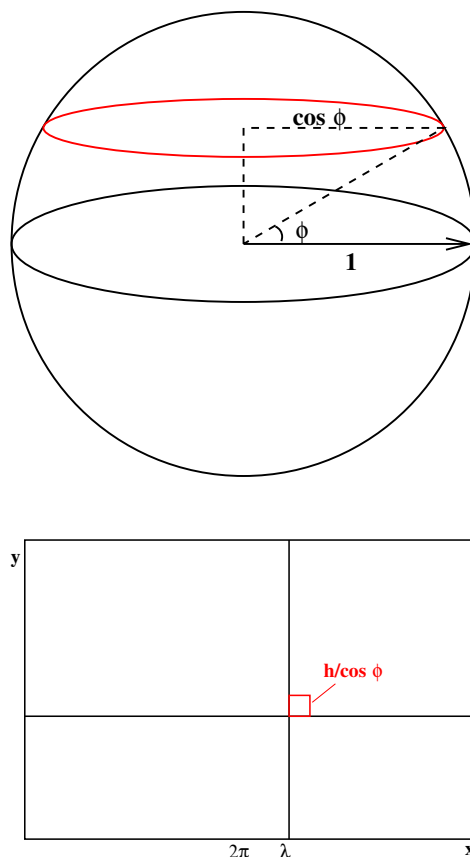


KUVA: Sylinteriprojektio (lähde: Carlos Furuti <http://www.progonos.com/furuti/>).

Mercatorin projektio on *konforminen*, eli se säilyttää kahden käyrän välisen kulman niiden leikkauspisteessä. Se on ainoa konforminen sylinteriprojektio. Tämä ominaisuus on hyödyllinen karttaprojektioilla, koska tällöin kartalta tehdyt mittaukset vastaavat maastosta mitattuja suuntia. Koska pituus- ja leveyspiiri on mahdollista selvittää mittaamalla taivaankappaleiden korkeuksia, ja suunta kompassia käyttämällä, se soveltuu erittäin hyvin navigointiin.

Mercatorin projektion konstruktio

Mercator ei esittänyt karttaprojektioilleen matemaattista selitystä. Vuonna 1599 englantilainen matemaatikko Edward Wright keksi miten projektio tehdään matemaattisesti. Tarkastellaan pientä tonttia, jonka eteläraja on leveyspiirillä ϕ . Tontin rajat ovat pituus- ja leveyspiirien suuntaiset ja sekä länsi- että etelärajan pituus on h . Jotta Mercatorin projektio voisi toimia, on tontin kuvan kartalla oltava neliö. Merkitään tontin länsirajan pituuspiiriä λ :lla.



KUVA: Mercatorin karttaprojektion konstruktio.

Merkitään kuvapistettä (x, y) ja valitaan $x = \lambda$. Jäljelle jää laskea miten saadaan y . Geometrisesti voidaan päätellä (ks. kuva), että leveyspiiriä ϕ vastaava venytys

karttaprojektiossa on $1/\cos \phi$. Siis tontin, jonka leveys on h , leveys kartalla on $h/\cos \phi$. Siksi myös korkeuden on oltava $h/\cos \phi$.

Selvästi kuvapisteen y -koodinaatti riippuu vain leveyspiiristä ϕ . Voidaan siis merkitä $y = F(\phi)$, missä F on aidosti kasvava funktio. On selvitettävä mikä F on. Tontin kuvasta kartalla saatiin yhtälö, joka voidaan kirjoittaa F :n avulla

$$F(\phi + h) = F(\phi) + h/\cos \phi,$$

eli

$$\frac{F(\phi + h) - F(\phi)}{h} = \frac{1}{\cos \phi}.$$

Kun $h \rightarrow 0$, saadaan $F'(\phi) = 1/\cos \phi$. Kiinnittämällä päiväntasaajan kuva kartalla tasolle $y = 0$, saadaan y -koordinaatille kaava

$$F(\phi) = \int_0^\phi \frac{dt}{\cos(t)}. \quad (1)$$

Valittavasti geometrinen konstruktio tehtiin ennen differentiaali- ja integraalilaskentaa, ja integraalia (1) ei osattu laskea, vaikka se nykyään onkin peruskursitasoa. Integraalin likiarvoja julkaistiin taulukkoina käytettäväksi merenkulussa. John Napier keksi vuonna 1614 logarifmifunktion, ja 1620 julkaistiin trigonometristen funktioiden logaritmeja sisältänyt taulukkokirja. Taulukkokirjoja tutkiessaan Henry Bond huomasi satumalta 1640, että

$$\int_0^\phi \frac{dt}{\cos(t)} = \ln(\tan(\phi/2 + \pi/4)).$$

Tämän tuloksen todistaminen säilyi kuitenkin avoimena ongelmana aina vuoteen 1668, jolloin James Gregory julkaisi sille (erittäin monimutkaisen) todistuksen. Mercatorin projektio voidaan siis määritellä kaavalla

$$(x, y) = (\lambda, \ln(\tan(\phi/2 + \pi/4))),$$

missä ϕ on pallon pinnalla olevan pisteen leveyspiiri ja λ sen pituuspiiri.

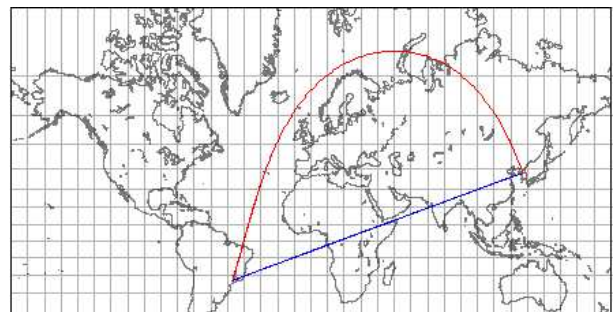
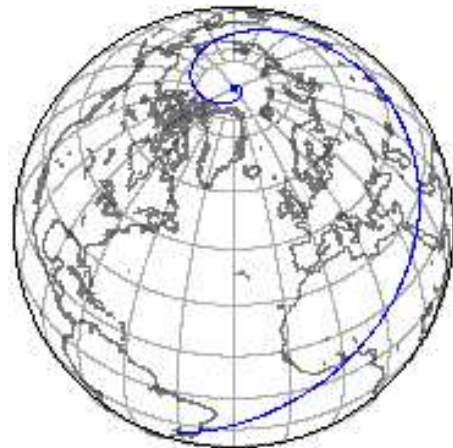
Loksodromit

Loksodromi on käyrä, joka syntyy edettäessä johonkin kiinnitettyyn kompassisuuntaan. Loksodromi kohtaa jokaisen pituuspiirin samassa kulmassa. Mercatorin projektiossa suorat kartalla vastaavat loksodromeja. Loksodromin idean keksijä on portugalilainen matemaatikko Pedro Nunes (1502–1578).

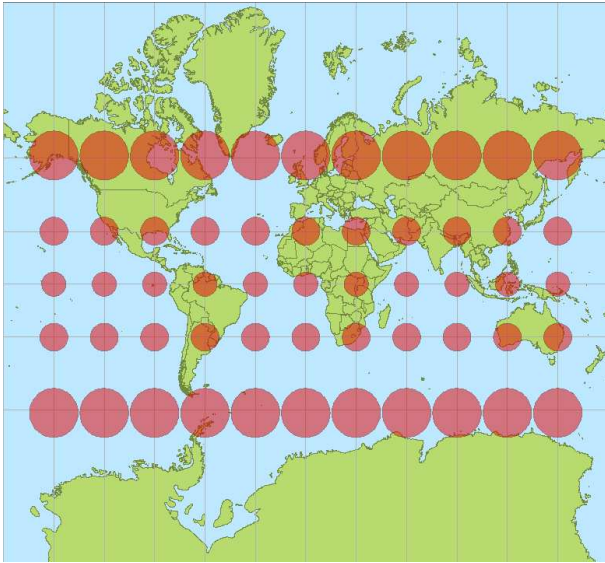


KUVA: Pedro Nunes kuvattuna portugalilaisessa postimerkissä (lähde: Wikipedia).

Loksodromit ovat hyödyllisiä navigoitaessa kompassin avulla, mutta eivät kuitenkaan (yleensä) anna lyhintä reittiä (ks. kuva) kahden pisteen välillä. Tästä huolimatta loksodromeja käytettiin mm. lentoliikenteen reittien suunnittelussa 1960-luvulle saakka, jolloin ne korvautuivat isoympyrän kaariin perustuvilla reiteillä.



KUVA: Loksodromi muistuttaa pallokartalla spiraalia, mutta sen kuva Mercatorin kartalla on suora. Loksodromi ei kuitenkaan anna lyhintä reittiä, joka on alemmassa kuvassa suoran päätepisteitä yhdistävä kaari (lähde: Carlos Furuti <http://www.progonos.com/furuti/>).



KUVA: Pinta-alojen vääristyminen Mercatorin projektiossa (lähde: Wikipedia).

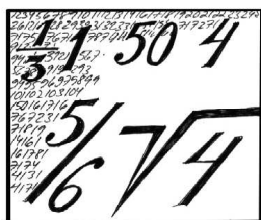
Poliittinen maailmankartta?

Maailmankarttaa laatiessaan Mercatorin tavoitteena oli luoda mahdollisimman hyvä apuväline merenkul-

kijoiden käytettäväksi. Mercatorin projektio ei säilytä pinta-aloja. Esimerkiksi Grönlanti näyttää projektiossa suunnilleen samankokoiselta kuin Afrikka. Todellisuudessa Afrikka on pinta-alaltaan noin 13-kertainen.

Kartan ominaisuuksista seuraa, että kaukana päiväntasaajalta olevat alueet ovat kartassa suurempia kuin todellisuudessa. Saksalainen yhteiskuntatieteilijä Arno Peters (1916 – 2002) esitti teoksessaan *Die Neue Kartographie/The New Cartography* (1983) voimakkaita syytöksiä kartanpiirtäjiä kohtaan. Hänen mukaansa karttojen pyrkimyksenä on esittää rikkaat pohjoiset teollisuusmaat todellista suurempina ja siten merkittävämpinä. Petersin esittämän ”tasa-arvoisen” maailmankartan voi tilata osoitteesta <http://www.petersmap.com/>.

Vaikka Petersin kritiikki vaikuttaakin perusteettomalta kartan takana olevaa matematiikkaa tuntevalle, kartat toki muokkaavat ihmisten maailmankuvaa monin tavoin. Mercatorin kartan hyödyllisyys kompassin avulla tapahtuvassa merenkulussa on tehnyt siitä suosittua monissa sellaisissa asioissa, joihin se ei sovellu. Esimerkiksi koulumaantiedon opetuksessa pinta-alat ja etäisyydet ovat varmastikin kartan konformisuutta tärkeämpiä ominaisuuksia.



Kaksi syntymäpäiväsankaria

Matti Lehtinen

Maanpuolustuskorkeakoulu

Sata ei ole matemaattisesti kovin kiinnostava luku, mutta ihmisellä sattuu olemaan kymmenen sormea ja kymmeneen perustuva tapa nimetä ja merkitä lukuja. Tapahtumat, joista on kulunut tasamäärä satoja vuosia, tulevat usein huomion kohteiksi.

Tämän vuoden huhtikuussa tulee kuluneeksi täysiä satoja vuosia kahden merkittävän matemaatikon syntymästä. Heillä on muitakin yhteisiä piirteitä: Kumpikin teki suuren osan merkittävimmistä työstään muualla kuin kotimaassaan ja kummankin elämä koskettaa Suomea, toisen tosin aika ohuesti. Kummankin uraa auttoi alkuun merkittävän avoimen ongelman ratkaisussa onnistuminen, ja molempien matematiikassa yksi merkittävä osa liittyi aikaisemman hämärästi itseään ilmaissien matemaatikon ajatusten ymmärrettäväksi tekemiseen. Henkilöt ovat *Leonhard Euler* ja *Lars Valerian Ahlfors*.

Eulerin elämä

Leonhard Euler syntyi Baselissa Sveitsissä 15. huhtikuuta 1707. Baselin porvarit olivat noina aikoina päättäneet, että matematiikka on koulussa turha aine, ja poistaneet sen kokonaan. Nuoren Eulerin koulutoveri sattui kuitenkin olemaan aikansa merkittävimpiin kuuluneen matemaatikon *Johann Bernoullin* poika,

ja isä-Bernoulli johdatti Eulerin matematiikan piiriin. Opetusmenetelmä oli tehokas: Bernoulli antoi Eulerille luku- ja perehtymistehtäviä viikoksi ja sallii oppilaan tulla sunnuntaina kyselemään, jos jotain oli jäänyt epäselväksi. Kunnianhimoinen Euler pyrki ymmärtämään itse ja kysymään niin vähän kuin mahdollista. Menetelmä toimi. Euler valmistui maisteriksi Baselin yliopistosta 16-vuotiaana. Hiukan vaikea yhdistelmä työnsaannin kannalta.

1700-luvun alussa Itämeren rannoilla tapahtui. Tsaari *Pietari Suuri* ryhtyi tosissaan johtamaan maataan Eurooppaan. Nevan suun soille perustettiin Pietarin kaupunki ja maata vallattiin Ruotsilta; Suomikin oli pitkään miehitettynä Ison Vihan aikaan. Pietarin pyrkimykset eivät olleet pelkästään sotaisia. Eräiden muiden aikansa hallitsijoiden tavoin hänkin halusi rakentaa ympärilleen oppineiden piirin, tiedeakatemia. Pietari ehti kuolla, mutta vuonna 1725 Akatemia aloitti toimintansa. Sen jäseniksi oli värvätty kaksi Johann Bernoullin poikaa, Daniel ja Nicolaus, ja myös tuolloinkin vielä varsin nuorelle Eulerille tarjottiin paikkaa Pietarissa. Euler oli saavuttanut ensimmäisen tieteellisen menestyksensä osallistumalla Pariisin tiedeakatemiaan kilpailuun, jossa aiheena oli purjelaivan mastojen konstruktio. Ei ihan sveitsiläisen aihe.

Eulerin ensimmäiset tehtävät Pietarissa olivat vaihtelee-

via. Hän opetti Tiedeakatemiassa ja Merisotakoulussa vaihtelevia aineita, fysiologiasta alkaen. Pian tehtävä vaihtui ensin fysiikan ja sitten matematiikan professoriksi. Missään vaiheessa Euler ei ollut fakki-idiootti. Hän antoi neuvoja hallitukselle eri asioissa, osallistui Venäjän kartoitukseen, auttoi laivanrakennuksessa ja suunnitteli paloruiskuja. Kruununperijän horoskoopin laadinnasta hän kuitenkin kieltäytyi.

Euler oli Pietarissa kahteen otteeseen, ensin vuoteen 1741 ja sitten vuodesta 1766 elämänsä loppuun. Väliajan Euler toimi Berliinissä. Preussin kuningas *Fredrik II* eli Fredrik Suuri halusi oman statussymbolinsa, Berliinin Akatemian, olevan numero yksi tiedeakatemioiden joukossa ja houkutteli Eulerin sen jäseneksi ja johtoon. Eulerin ja Fredrikin välit eivät olleet parhaat mahdolliset: ahkera, asiallinen ja mutta väritön matemaatikko ei viehättänyt monarkkia samalla tavoin kuin esimerkiksi kirjailija-filosofi *Voltaire*, Fredrikin suosikki. Lopulta Euler kyllästyi ja hakeutui takaisin Pietariin, jossa valtaa nyt piti saksalaissyntyinen keisarinna *Katariina*, myös II, myös Suuri.

Euler avioitui Pietarissa saksalaissyntyisen *Katharina Gsellin* kanssa. Eulereilla oli 13 lasta. Eulerin tasaisen ja työteliään elämän yksi tragedia oli näön menetys. Eulerin oikean silmän näkö meni jo vuonna 1738. Perimätieto sanoo syyn olleen auringon katsomisen kaukoputken läpi, mutta luultavampi syy on infektiosairaus. Melkein kaikki Eulerin muotokuvat on maalattu niin, että malli on kääntänyt maalariin päin kasvojensa vasemman puolen. Fredrik Suuri nimitteli Euleria ilkeyksissään kykloopikseen. Pian Eulerin palattua toisen kerran Pietariin hän menetti kaihin takia näkönsä toisestakin silmästä. Vuonna 1771 Eulerin kaihi leikattiin ja hän näki muutaman päivän ajan, mutta sitten silmiin iski infektio, joka sokeutti Eulerin täydellisesti ja lopullisesti.

Näön menetyksellä ei näytä olleen mitään vaikutusta Eulerin työtahtiin. Hän pystyi käsittelemään matemaatiikkaa päässään ja ajattelemaan valmiiksi tuloksensa, jotka hän sitten saneli assistenteilleen.

Tässä tulemme Eulerin ja Suomen yhteyteen. Yksi Eulerin tärkeimpiä assistentteja oli *Anders Lexell*, Turussa jouluaattona 1740 syntynyt raatimies Jonas Lexellin poika. Lexell oli 1763 tullut Turun Akatemian dosentiksi, mutta hänelle kutsu Pietariin vuonna 1768, siis kaksi vuotta Eulerin toisen Pietarin-kauden alun jälkeen, oli merkittävä askel tieteen suureen maailmaan. Kutsun takana oli Euler. Lexell nimitettiin jo 1775 professoriksi Turkuun, mutta hän otti jatkuvasti virkavapautta voidakseen toimia Eulerin luona Pietarissa. Kun Euler kuoli 1783, Lexell peri hänen virkansa. Hän kuoli kuitenkin jo seuraavana vuonna. Lexell on ensimmäinen kansainvälisesti merkittävä suomalainen matemaatikko. Hänen saavutuksiinsa kuuluu Uranus-planeetan tunnistaminen planeetaksi. (Uranuk-

sen löytäjä, englantilainen William Herschel oli tulkinut löytönsä komeetaksi.) Uranuksen ratalaskelmista on kunnian saanut kuuluisa ranskalaismatemaatikko *Pierre Simon Laplace*, mutta Lexell oli tehnyt omat laskunsa Laplacesta riippumatta ja tietämättä.

Euler kuoli Pietarissa 18. syyskuuta 1783. Hänet haudattiin Vasilinsaarelle, mutta vuonna 1956 hänen maalliset jäämistönsä siirrettiin Aleksanteri Nevskin luostarin liepeillä olevalle hautausmaalle. Eulerin kookas mutta mahtaimaton hautakivi löytyy vastapäätä usein turistikohteena olevaa ”Kuuluisuuksien hautausmaata”, sitä, johon mm. Tšaikovskin ja Dostojevskin maalliset jäännökset on sijoitettu. Pietarissa käydessä kannattaa poiketa.

Eulerin matematiikkaa

Eulerin matemaattisten pääsaavutusten pintapuolinenkaan luettelo ei ole näissä puitteissa mahdollista. Euler oli luultavasti kaikkien aikojen tuotteliain matemaatikko. On laskettu, että hän kirjoitti 866 tieteellistä artikkelia. Pietarin tiedeakatemian julkaisut täyttyivät Eulerin kirjoituksista kymmeniksi vuosiksi hänen kuolemansa jälkeen. Eulerin koottuja teoksia on Sveitsissä julkaistu yli 70 vankkaa osaa. Euler käsitteli tutkimuksissaan kaikkia tuohon aikaan tunnettuja matemaatiikan aloja ja pani alulle uusia. Google-haku sanalla Euler tuottaa 12 100 000 iskemää (helmikuun alussa 2007), kun Gauss, ”matemaatikkojen kuningas”, saa vain 11 600 000 (ja Ahlfors 75 000).

Paitsi tutkimuksia, Euler kirjoitti loisteliaita oppikirjoja, jotka pitkälle ovat määrittäneet matemaattisen oppikirjan olemuksen. Vähäinen ei ole hänen merkityksensä myöskään matematiikan merkintöjen ja nimitysten keksijänä ja vakiinnuttajana. Meille kaikille tuttu π on tullut matematiikkaan varsinaisesti Eulerin oppikirjassa *Introductio in Analysin Infinitorum*. Euler kirjoitti myös kansantajuisesti: kuuluisa on hänen 1768 julkaisemansa ranskankielinen teos *Kirjeitä eräälle saksalaiselle prinsessalle erinäisistä fysiikan ja filosofian aiheista*. Se on erittäin selkeää ja yksinkertaista, mutta asiallista tieteen popularisointia.

Matemaatikoilla on kaunis tapa ikuistaa kollegojaan nimeämällä käsitteitä tai matemaattisia tuloksia heidän mukaansa. Otetaan tähän muutama näyte asioista joihin lähes jokainen matematiikan kanssa toimiva väistämättä törmää:

Eulerin kaava

$$e^{ix} = \cos x + i \sin x$$

on keskeisin väline kompleksilukujen käytössä ja ymmärtämisessä.

Eulerin monitahokaskaava $V - E + S = 2$ sitoo yhdes- ti yhtenäisen monitahokkaan kärkien lukumäärän V ,

särmien lukumäärän E ja sivujen lukumäärän S yhteen. Laskepa suure kuutiolle, tetraedrille ja Kheopsin pyramidille!

Eulerin suora yhdistää missä tahansa kolmiossa keskijanojen leikkauspisteen M , sivujen keskinormaalien leikkauspisteen eli kolmion ympäri piirretyn ympyrän keskipisteen O ja kolmion korkeusjanojen leikkauspisteen H . Tämän asian todistus ei ole kovin hankala. Joskus ihmetellään, miksei Eukleides tiennyt tätä.

Eulerin funktio $\phi(n)$ kertoo niiden positiivisten kokonaislukujen $k < n$ lukumäärän, joilla ei ole luvun n kanssa muita tekijöitä kuin 1. Eulerin–Fermat’n lause kertoo, että jos a :lla ja n :llä ei ole yhteisiä tekijöitä, $a^{\phi(n)}$ antaa aina jakojäännöksen 1, kun jakaja on n . Kokeile, kun $n = 100$.

Numeerista matematiikkaan harrastava törmää Eulerin yksinkertaiseen likiarvomenetelmään differentiaaliyhtälön $y' = f(x, y)$ alkuarvotehtävän $y(x_0) = y_0$ ratkaisemiseksi on seuraava: valitaan askelpituus h ; määritellään $y_1 = y_0 + hf(x_0, y_0)$, $y_2 = y_1 + hf(x_1, y_1)$ jne. Silloin y_k on likimain yhtälön ratkaisun arvo pisteessä x_k .

Eulerin vakio γ on yksi matematiikan tärkeitä, monessa yhteydessä esiin tulevia erityisiä lukuja. Se on raja-arvo

$$\lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln n \right) \approx 0,5772156649 \dots$$

Hämmästyttävää on, että vieläköän emme tiedä, onko γ rationaali- vai irrationaaliluku.

Verkkoteorian peruskäsitteitä on *Eulerin ketju*. Sitä voisi havainnollistaa kävelynä kaikkia kaupungin katuja pitkin niin, mitään katuosuutta ei kävellä kahdesti. Euler ratkaisi vuonna 1735 ”ajanvietematematiikan” ongelman, joka koski mahdollisuutta tehdä Königsbergin kaupungissa Itä-Preussissa kävely, joka olisi ylittänyt kaupungissa olleet seitsemän siltaa kunkin vain kerran. Monesti lasketaan, että nykyään suuret ja tärkeät matematiikan alat topologia ja verkkoteoria ovat saaneet alkusysäyksensä juuri tästä Eulerin ratkaisusta. Euler käsitteli kirjoituksissaan myös muita ajanvietematematiikan aiheita, mm. latinalaisia neliöitä, sudokujen esimuotoa.

Yksi nuoren Eulerin suuria ja matematiikan kannalta kauaskantoisimpia saavutuksia ja hänen maineensa perusta oli niin sanotun Baselin ongelman ratkaisu. Kysymys oli sarjan

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots$$

summasta. Kysymystä oli pohtinut mm. Johann Bernoulli, tulokseen kuitenkin pääsemättä. Euler onnistui vuonna 1735 osoittamaan, että summa on

$$\frac{\pi^2}{6}.$$

(Tuolloin Euler ei vielä käyttänyt kirjainta π vaan kirjainta p osoittamaan ympyrän kehän ja halkaisijan suhdetta.) Pari vuotta myöhemmin Euler johti yleisemmän kaavan

$$\sum_{k=1}^{\infty} \frac{1}{k^s} = \prod \frac{1}{1 - p^{-s}},$$

missä oikean puolen tulo ottaa huomioon kaikki alkuluvut.

Eulerin havainto on lähtökohta tarkasteluihin, jotka johtavat matematiikan luultavasti kuuluisimpaan avoimeen ongelmaan, *Riemannin hypoteesiin*. Eulerin kaavan vasemman puolen summa on s :n funktio, ja se on mielekäs myös, kun s on kompleksiluku. On tullut tavaksi merkitä tätä funktiota symbolilla ζ . Monien lukuteorian ongelmien kannalta olisi tärkeää tietää, milloin $\zeta(s) = 0$. Saksalainen *Bernhard Riemann* esitti vuonna 1859, että jos $\zeta(s) = 0$, niin joko s on parillinen negatiivinen kokonaisluku tai s on kompleksiluku, jonka reaalisosa on $\frac{1}{2}$. Onko todella näin, on yhä ratkaisematta. Ratkaisijalle olisi tarjolla miljoonan dollarin palkkiokin. Myös se vielä kuuluisampi ongelma, kymmenisen vuotta sitten ratkaisunsa saanut Fermat’n suuren lauseen todistus, saa kiittää kuuluisuudestaan Euleria: Euler oli ensimmäinen, joka tarttui Fermat’n heittämään syöttiin ja todisti Fermat’n lauseen ensimmäisen vaikeamman tapauksen: sen, että yhtälöllä $x^3 + y^3 = z^3$ ei ole nollasta eroavia kokonaislukuratkaisuja.

Eulerin uskomatonta tuotteliaisuutta ja hänen tulosensa moninaisuutta esitellessä on tapana aina huomauttaa, että Eulerin matematiikka ei ollut kaikin puolin korrektia. Hän operoi paljon päättymättömillä sarjoilla. Nykymatemaatikko ja opiskelija tietää, että sarjoihin liittyy aina kysymys suppenemisesta ja siitä, voidaan sarjan yksittäisiin termeihin kohdistuvat toimet kuten derivointi tai integrointi periyttää sarjan summalle. Euler ei näistä kysymyksistä samalla tavalla huolehtinut. Tämä oli kuitenkin hänen ajalleen tyypillistä. Suppenemiseen ja yleisemmin raja-arvoihin liittyvät ongelmat tiedostettiin ja ratkaistiinkin vasta Euleria seuraavan vuosisadan aikana. Eulerin matemaattinen intuitio johti kuitenkin siihen, että hänen epäilyttävätkin päättelynsä yleensä johtivat oikeaan lopputulokseen.

Eulerin tuotantoa on paljon esillä netissäkin: Euler Archive osoitteessa

<http://www.math.dartmouth.edu/~euler/>

pyrkii saamaan Eulerin koko tuotannon kaikkien näkyviin.

Lars Valerian Ahlfors

Lars Ahlfors syntyi Helsingissä 18. huhtikuuta 1907, siis lähes päivälleen 200 vuotta myöhemmin kuin Euler. Kun Ahlfors alkoi matematiikan opintonsa Helsingin yliopistossa 1924, matematiikan ainoana professorina oli *Ernst Lindelöf*, mies, jota enemmän kuin ketään muuta voi kiittää matematiikan korkeatasoisen tutkimuksen alkuun panemisesta Suomessa. Lindelöf oli monipuolinen matemaatikko, mutta hänen pääkiinnostuksensa kohde oli 1800-luvun lopulla voimakkaasti kehittyä alkanut kompleksilukumuuttujan kompleksilukuarvoisten funktioiden, erityisesti niin sanottujen analyyttisten funktioiden tutkimus. Lindelöfin toinen lahjakas oppilas, *Rolf Nevanlinna*, joka oli juuri julkaissut sittemmin meromorffifunktioiden arvojenjakautumisteorian tai Nevanlinnan teorian nimellä tunnetut tuloksensa, tuli Ahlforsin opiskeluvuosina toiseksi matematiikan professoriksi. Lindelöfin ansiota lienee pitkälti se, että kun Nevanlinna vuonna 1928 matkusti joksikin aikaa Zürichiin vierailevaksi tutkijaksi, juuri maisteriksi valmistunut 21-vuotias Ahlfors sai seurata häntä.

Funktioteoria, ainakin siltä osin kuin se Suomessa tuli merkittäväksi tutkimuskohteeksi, oli lähtöisin Ranskasta. Ja Ranskassa oli vuonna 1909 väitellyt tohtoriksi *Arnaud Denjoy* tutkimuksella, joka koski erästä kompleksifunktioiden erityisluokkaa, ns. kokonaisfunktoita. Denjoy oli väitöskirjassaan esittänyt otaksuman tällaisten funktioiden asymptoottisten arvojen lukumäärää. Otaksuma oli jäänyt todistamatta, mutta ongelma oli kiinnostava, ja sen ratkaisua olivat miettineet monet alan arvostetut tutkijat, Nevanlinnakin. Väitöskirjaansa aiheutta kaipaavalle nuorelle Ahlforsille Nevanlinna ehdotti tutustumista Denjoyn ongelmaan. Nevanlinna oli varsin hämmästynyt, kun Ahlfors parin viikon kuluttua palasi ongelman ratkaisun kera.

Denjoyn ongelman ratkaisu sisältyy Ahlforsin vuonna 1930 (jolloin Ahlfors oli 23-vuotias) valmistuneeseen väitöskirjaan. Hänen tutkijanuransa eteni nyt nopeasti. Hänet nimitettiin Helsingin yliopiston apulaiseksi eli apulaisprofessoriksi vuonna 1932 ja professoriksi 1938. Lukuvuoden 1935–36 hän oli Harvardin yliopistossa. Tutkijana hän aluksi seurasi opettajaansa Nevanlinnaa, mutta varsin omintakeisesti. Merkittäväksi muodostui Ahlforsin 1935 julkaisema tutkimus *Zur Theorie der Überlagerungsflächen*.

Matemaatikot ovat jo yli sadan vuoden ajan kokoontuneet joka neljäs vuosi yleiseen kansainväliseen matemaattikkokongressiin, jossa esitellään laajasti tieteen uusimpia saavutuksia yli matematiikan monien erikolasalojen rajojen. Vuonna 1924 kokous pidettiin Torontossa Kanadassa. Kokouksen pääorganisoi oli *John Fields*. Hän oli niin taitava sponsoroinnin järjestäjä, että kokous tuotti pienen ylijäämän. Fields ehdotti,

että ylijäämä käytettäisiin tulevissa Kansainvälisissä matemaattikkokongresseissa erittäin ansioituneiden matemaatikkojen palkitsemiseen. Fields kuoli vuonna 1932, eikä ehtinyt nähdä ehdotuksensa toteutuvan. Se kuitenkin toteutui vuonna 1936 Oslossa pidetyssä kongressissa. Palkittuja matemaatikkoja oli kaksi: amerikkalainen minimipintojen tutkija *Jesse Douglas* ja suomalainen Lars Ahlfors. Ahlforsin erityiseksi ansioksi katsottiin juuri edellisessä kappaleessa mainittu tutkimus.

Kun matematiikka ei kuulu niihin tiedonaloihin, joita Alfred Nobel halusi palkittavan, niin Fieldsin mitali on muodostunut ”matematiikan Nobeliksi”, korkeimmaksi kunnianosoitukseksi, jonka matemaatikko voi saada. (Mitaliin liittyvä rahapalkinto, 15 000 Kanadan dollaria, on kuitenkin pieni murto-osa Nobelin palkinnosta.) Ahlforsin Fieldsin mitali on nykyään Suomessa. Sen voi nähdä Helsingin yliopiston matematiikan laitusrakennuksen Exactumin ala-aulan vitriinissä Kumpulankampusalueella.

Ahlfors jätti Suomen vuoden 1944 kesän sekavissa oloissa. Yksi syy pois siirtymiseen saattoi olla se, että Ahlforsin vaimo Erna oli itävaltalainen. Ahlfors siirtyi ensin Zürichiin ja sitten Harvardiin, jossa hän vuodesta 1946 oli matematiikan professorina vuoteen 1977, ja senkin jälkeen aktiivisena tutkijana. Ahlforsilla ohjasi noin 25 väitöskirjaa. (Pieni detalji, joka tavallaan osoittaa matematiikan paikkaa yleisessä arvostuskentässä: kun entinen pääministeri ja presidenttiehdokas Esko Aho vuonna 2000 lähti muutamaksi kuukaudeksi Harvardiin, Suomen Kuvalehti julkaisi laajan artikkelin Harvardissa opiskelleista tai vaikuttaneista suomalaisista. Lars Ahlforsia ei artikkelissa ollenkaan mainittu.)

Ahlfors oli – kuten matemaatikot keskimäärinkin ovat – ahkera ja kurinalainen tutkija. Hän oli kuitenkin myös, kuten akateemikko Olli Lehto Arkhimedes-lehdessä julkaistussa muistokirjoituksessaan toteaa ”värikä, vieraanvarainen, voimaa uhkuva *bon vivant*”. Ahlforsin alkoholinkäyttöön liittyvät anekdootit ovat matemaattikkokertomusten klassikkoja. Hänet tunteneet antavat ylistäviä lausuntoja hänen ja Ernan vieraanvaraisuudesta ja ystävällisyydestä. Ja mainitsevat, että juhlaillan jälkeen aamuna Ahlfors oli aina valmis jatkamaan työtään.

Lars Ahlfors eli korkeaan ikään ja jatkoi työtään lähes viimeisiin vuosiinsa asti. Ahlfors kuoli Pittsfieldissä Massachusettsissa 11. lokakuuta 1996.

Ahlforsin matematiikka ei tietenkään ole yhtä laajaa alaista kuin Eulerin. Kahdessasadassa vuodessa matemaattisen tiedon määrä oli valtavasti kasvanut ja tietämyksen rajat edenneet niin, että näitä rajoja ei enää kukaan yksilö voinut työntää edemmäs kovin monessa paikassa. Funktioteorian tai niin kuin sitä nyt tavallisemmin nimitetään kompleksianalyysin alalla Ahlfors

fors oli kuitenkin erittäin monipuolinen. Yksi hänen tutkimusaloistaan tuo uuden analogian Euleriin, joka lukuteoriassa paljon täydensi, selvensi ja korjasi Pierre de Fermat'n tuloksia.

Funktioteoreettikojen tutkimat analyyttiset funktiot ovat niin sanottuja konformikuvauksia. Se tarkoittaa, että vaikka funktion välittämä tason kuvaus vääntelisi kuvioita isossa mittakaavassa rajustikin, niin "mikroskooppisesti" kuvaus on yhdenmuotoisuuskuvauksia: kahden käyrän välinen kulma säilyy aina samana. 1920-luvulta alkaen jotkut matemaatikot, Ahlfors heidän joukossaan, alkoivat miettiä, mitä voisi tapahtua, jos kuvaukselle sallittaisiin hiukan enemmän vapautta. Ahlfors antoi nimen tälle konformikuvauksia laajemmalle luokalle: ne ovat kvasikonformikuvauksia. Mutta ilman muuta pisimmälle näiden funktioiden tutkimuksessa pääsi saksalainen Oswald Teichmüller. Teichmüllerin tapa esittää asiansa oli kui-

tenkin erittäin vaikeaselkoinen, eikä hänen töihinsä tutustumiseen myöskään varsinaisesti kannustanut tieto Teichmüllerin voimakkaista ja vastenmielisistä poliittisista mielipiteistä (jotka eivät kuitenkaan hänen matematiikassaan näy, vaikka 1930-luvun Saksassa sellais-takin ilmeni).

Ahlforsin kirjoitukset 1950-luvulla toivat Teichmüllerin ajatukset ymmärrettäviksi. Kvasikonformikuvauksista kasvoi merkittävä kompleksianalyysin osa-alue. Se on ollut yksi tärkeimpiä matematiikan tutkimusaloja Suomessa 1900-luvun loppupuoliskolla.

Myös Ahlfors kirjoitti varsinaisten tutkimusten lisäksi oppikirjoja. Erityisesti vuonna 1953 ensi kerran ilmestynyt funktioteorian oppikirja *Complex Analysis* on säilyttänyt asemansa klassikkona. Ahlfors omistaa sen Ernst Lindelöfin muistolle.



Vektorit, matriisit, Händel ja vaalit

Heikki Apiola

Matematiikan laitos, Teknillinen korkeakoulu

Lineaarialgebran maistiaiset

Lineaarialgebra on matematiikan ala, joka on sekä teorian että sovellusten kannalta hyvin keskeinen. Se yleistää havainnollisen vektorikäsitteen mitä moninai-
simpiin ja yllättävimpiin tilanteisiin.

Matriisit ovat lukutaulukoita, joille määritellään lasku-
toimitukset. Taulukkoon koottua tietoa käsitellään yh-
tenä objektina samankaltaisin laskusäännöin kuin lu-
kuja. Kun matriisia sovelletaan vektoriin, syntyy lii-
kettä, dynamiikkaa.

Lineaarialgebran merkitys matematiikan sovelluksille
on kasvanut käsi kädessä tietotekniikan kehityksen
kanssa. Niinpä alan ohjelmisto on varsin kypsässä ja
kattavassa tilassa, toki kaiken aikaa kehittyen.

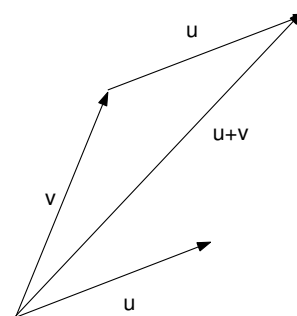
”Lineaarialgebran kieltä” puhuvia laadukkaita ohjel-
mistoja on saatavilla. Kirjoituksessa esittelemme esi-
merkien valossa vapaasti saatavissa olevaa ohjelmaa
SCILAB, joka noudattaa varsin uskollisesti tunnetun ja
maailmalla hyvin suosittun MATLAB-ohjelman syntak-
sia.

Kirjoituksessa pyrimme valottamaan mm. vektor-
ikäsitteen moninaisia ja kenties yllättäviä ilmenty-
miä havainnollisista tason suuntanuolista aina digitaal-
isiin musiikkivektoreihin. Matriisien alalla korostam-
me tässä dynamiikkaa ja prosessia aina eduskuntavaa-
leihin saakka.

Kyseessä ei ole yritys opettaa varsinaista line-
aarialgebran kieltä yhden kirjoituksen puitteissa,
vaan pikemminkin antaa maistiaisia ja ruokahalua
sekä näyttää kenties yllättäviäkin näiden käsitteiden
käyttöalueita ja -tapoja.

Geometriset ja fysiikan vektorit

Geometriset vektorit tasossa ja avaruudessa lienevät
koulumatematiikasta ja fysiikasta tuttuja olioita. Tyy-
pillinen fysiikan ensituttavuus on voima, jolla on suu-
ruus ja suunta. Tällaista suuretta sanotaan *vektorisuu-
reeksi*. Sitä voidaan geometrisesti kuvata voiman suun-
taa osoittavalla nuolella, jonka pituus kuvaa voiman
suuruutta.



KUVA 1. Vektorisumma $u + v$

Vektorien $\mathbf{u}$ ja $\mathbf{v}$ summa $\mathbf{u} + \mathbf{v}$ määritellään kuvan osoittamalla tavalla. On syytä panna merkille, että vektorin määrää suunta ja pituus, ts. kaikki **samansuuntaiset ja yhtä pitkät** suuntajanat alkupisteestä riippumatta **edustavat samaa vektoria**. Kuvassa se ilmenee kahtena $\mathbf{u}$:lla merkittynä suuntajanana.

Tason vektoreita voidaan toisaalta käsitellä puhtaasti algebrallisesti lukupareina $\mathbf{p} = (x, y)$. Geometrisena taustana on tällöin tason pisteen koordinaattiesitys. Yhteys edelliseen nuoliajatuksen saadaan piirtämällä vektorinuoli, jonka alkupiste asetetaan koordinaatiston origoon O . Pisteen $\mathbf{p}_1 = (x_1, y_1)$ ja $\mathbf{p}_2 = (x_2, y_2)$ summa määritellään luontevasti: $\mathbf{p}_1 + \mathbf{p}_2 = (x_1 + x_2, y_1 + y_2)$, ts. lasketaan vastinkoordinaatit yhteen. Geometrisesti tulos saadaan laskemalla vektorien $\mathbf{v} = \overline{O\mathbf{p}_1}$ ja $\mathbf{u} = \overline{\mathbf{p}_1\mathbf{p}_2}$ summavektori, jonka kärjen koordinaatit antavat juuri edellä määritellyn summapisteen $\mathbf{p}_1 + \mathbf{p}_2$.

Vektoreille määritellään myös toinen laskutoimitus, luvulla eli *skalaarilla*¹ kertominen.

Vektorin $\mathbf{v}$ kertominen skalaarilla c tarkoittaa geometrisessa mallissa vektorin pituuden kertomista $|c|$:lla ja lisäksi suunnan vaihtoa, jos $c < 0$. Algebrallisessa mallissa vektorin komponentit kerrotaan c :llä, ts. jos $\mathbf{v} = (x, y)$, niin

$$c\mathbf{v} = (cx, cy).$$

Edellä oleva yleistyy suoraan 3-ulotteiseen avaruuteen, algebrallisessa vektorimallissa otetaan mukaan kolmas koordinaatti: $\mathbf{v} = (x, y, z)$, geometrisessa suuntajanat leijailevat avaruudessa.

Järjestettyjen lukuparien joukkoa kutsutaan (vektori)tasoksi ja merkitään symbolilla $\mathbb{R}^2$. Vastaavasti lukukolmikoiden joukkoa kutsutaan (3-ulotteiseksi) vektoriavaruudeksi ja käytetään merkintää $\mathbb{R}^3$.

Geometriaan viittaavassa puheessamme nimitämme samaa oliota vuoroin ”pisteeksi” ja vuoroin ”vektoriksi”. Jälkimmäisessä tapauksessa meillä on takaraivossamme ajatus origosta pisteeseen piirretystä paikkavektorista. Esimerkiksi laskiessamme ”pisteitä” yhteen, laskemme geometrisesti ajatellen yhteen vastaavia paikkavektoreita.

Lisää ulottuvuuksia

Algebrallisen vektorimallin etu on, että se voidaan suoraan yleistää mielivaltaisen korkeaan dimensioon n yksinkertaisesti kutsumalla **vektoreiksi** lukujonoja $\mathbf{u} = (x_1, x_2, \dots, x_n)$ ja määrittelemällä laskutoimitukset **vektoreille** $\mathbf{u} = (x_1, x_2, \dots, x_n)$ ja $\mathbf{v} = (y_1, y_2, \dots, y_n)$ sekä skalaarille c asettamalla:

$$\begin{cases} \mathbf{u} + \mathbf{v} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ c\mathbf{u} = (cx_1, cx_2, \dots, cx_n) \end{cases}$$

Vektorien *vähennyslasku* suoritetaan vähentämällä toisistaan vastinkoordinaatit, *nollavektori* on vektori

$$\mathbf{0} = (0, 0, \dots, 0).$$

Vektorien yleiset laskusäännöt ovat aivan samat kuin vastaavat reaalityökalujen laskusäännöt, jos ajatteleme skaarilla kertomista tulona. Näitä sääntöjä ovat liitöntä- vaihdanta- ja osittelulait. Kaikki palautuvat suoraan koordinaateilla suoritettaviin lukujen vastaaviin sääntöihin.

Toinen yleistysmahdollisuus on sallia kompleksiset komponentit ja skalaarit. Pitäydymme tässä kirjoituksessa ”turvallisuussyistä” reaalissa vektorikäsitteissä, vaikka lopussa mainitussa ominaisarvoteoriassa kompleksisilta ei voida välttyä.

Sisätulo, pituus, kulma

Geometrisessa vektorimallissa vektorin $\mathbf{u}$ pituus on tunnettu peruskäsite, jota merkittäköön $\|\mathbf{u}\|$. Jos $\mathbf{u} = (x, y)$, niin *Pythagoraan lauseen* mukaan $\|\mathbf{u}\| = \sqrt{x^2 + y^2}$. Jos kyseessä on 3-ulotteisen avaruuden vektori $\mathbf{u} = (x, y, z)$, niin $\|\mathbf{u}\| = \sqrt{x^2 + y^2 + z^2}$.

Tässä on selvä malli yleiseen tilanteeseen, määritellään siis vektorin $\mathbf{u} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ pituus eli *normi* kaavalla

$$\|\mathbf{u}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Määritellään nyt suoraan yleisessä tilanteessa vektorien $\mathbf{u}$ ja $\mathbf{v}$ *sisätulo* (skalaaritulo, pistetulo) kaavalla

$$\mathbf{u} \cdot \mathbf{v} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Geometrisessa vektorimallissa sisätulo voidaan määritellä kordinaattivapaasti kaavalla $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \gamma$, missä γ on vektorien $\mathbf{u}$ ja $\mathbf{v}$ välinen kulma. Määritelmien yhtäpitävyys voidaan osoittaa esimerkiksi *kosinilauseella*.

Yleisessä n -ulotteisessa vektorimallissa ei luonnostaan ole käsitettä ”vektorien välinen kulma”. Nyt voimme halutessamme määritellä tuon kulman yllä olevalla sisätulokaavalla. Yleisen vektorikulman tärkein merkitys liittyy *kohtisuoruuden* eli *ortogonaalisuuden* käsitteen yleistykseen, ehtona on $\mathbf{u} \cdot \mathbf{v} = 0$

¹Skalaari, ”scalar” – sellainen, jolla skaalataan

Olemme irrottaneet vektoriopin geometrisesta taustastaan ja antaneet täsmällisen sisällön n -ulotteisen avaruuden vektoreille. Tämä on pohjana alkujohdannossa kuvatulle tärkeälle matematiikan alalle, *lineaarialgebralle*. Erityisen viehättävää lineaarialgebrassa on, että pohjana on koko ajan geometrinen intuitio 2- ja 3-ulotteisten geometrinen vektorien parissa, joka auttaa koko perusteorian ymmärtämisessä ja todistusten ideoinnissa olennaisella tavalla. Se toimii yleisen teorian ”mallikuvana” kaiken aikaa.

Vektorilaskentaa tietokoneella

Sovelluksissa suoritamme vektorilaskentaa usein hyvin moniulotteisilla vektoreilla. Laskut muodostuvat pitkiksi ja tuloksen jälkikäsitteily saattaa sisältää vaikkapa vektorin muuntamisen kuvaksi tai ääneksi.

Ohjelmat, joiden perustietorakenteisiin kuuluvat vektorit, ja laskuoperaatiot kattavat vektorialgebran, ovat korvaamattomia työvälineitä.

Olen aiemmin esitellyt symbolilaskentaohjelmistoa MAPLE

<http://solmu.math.helsinki.fi/1999/5/apiola/>
ja numeerista vektori/matriisiohjelmistoa MATLAB.
solmu.math.helsinki.fi/2004/3/apiola.pdf

Ongelmana lukijoille on ohjelmien maksullisuus. Esitelen tällä kertaa vapaasti saatavaa ohjelmaa

SCILAB, joka on peruseräpäätteiltään aivan samanlainen kuin MATLAB. Ohjelmassa on avoimesti pyritty varsin pitkälle menevään yhteensopivuuteen MATLAB’n kanssa. (Toinen vastaava vapaasti saatava ohjelmisto on OCTAVE.)

Ohjelman voit ladata itsellesi osoitteesta

<http://www.scilab.org/>. Käyttöohjeita on ranskan, saksan ja englanninkin kielellä saatavissa:
<http://www.scilab.org/publications/>

Lyhyt Scilab/Matlab-käyttöohje

SCILAB-ohjelman opiskelun voit aloittaa josta-kin MATLAB-oppaasta, suomenkielinen MATLAB-käyttöohje on viitteissä.

Kun SCILAB käynnistetään, avautuu komentoikkuna. Tähän voidaan kirjoittaa ohjelman tuntemia komentoja, jotka ohjelma tulkitsee ja toteuttaa samantien (tai antaa virheilmoituksen). Tarkoituksena on antaa pieniä maistiaisiasia ohjelman käytöstä sen verran, että esittämämme esimerkit voidaan suorittaa ja ymmärtää. Tarkoitus on innostaa lukijaa omiin kokeiluihin ohjelmien avustus- ja dokumentointijärjestelmien ja viitteissä mainittujen lähteiden avulla. Voit aloittaa kirjoittamalla jotain tämällylistä SCILAB- (tai MATLAB-) komentoikkunaan:

```
// Ensimmäinen Scilab/Matlab-istunto
// Scilab-kommentti: // Matlab-kommentti: %
u=[1 2 3 4] // Luodaan vektori u
v=[-2 2 -1 5] // samoin toinen vektori v
w=-u+2*v // Lineaarikombinaatio sijoitetaan
           // muuttuukaan w
u.*v // Kerrotaan vektorin vastinkomponentit
u*v' // u:n ja v:n sisätulo.
```

Selityksiä:

1. Lineaarikombinaatio tarkoittaa muotoa
 $c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k$ olevaa summaa.

2. Viimeistä edellinen rivi: $\mathbf{u}.*\mathbf{v}$ on pisteittäin eli vastinalkioittain tapahtuva kertolasku. Se ei siis ole varsinaisen vektorilaskutoimitus, mutta monessa yhteydessä, erityisesti vektorilausekkeissa tuiki tarpeellinen.

3. Viimeinen rivi on erikoistapaus matriisikertolaskusta, siitä enemmän tuonnempana.

Ohjelman käytön käytännön vihjeitä

Yllä olevia ja muutamia vastaavia kirjoitettua ei kannata jatkaa työskentelyä suoraan komentoikkunassa, koska työsi häviää taivaan tuuliin. Ainoa, mitä siinä voit tehdä, on selata komentopuskuria nuoli-ylösnäppäimellä ja editoida komentoriviä.

Sensijaan kannattaa avata otsikkopalkista Editor ja kirjoittaa komentoja kommentteineen editorikkunaan. SCILAB tarjoaa varsin kätevän toimintatavan: Maalataan halutut komentorivit ja painetaan CTR-Y tai valitaan otsikkoriviltä Execute. Kyseiset rivit siirtyvät välittömästi tulkille suoritettaviksi.

Toinen mahdollisuus on käyttää omaa mielieditoria, kuten EMACS, NOTEPAD, WORDPAD, ym. ja leikata/liimata editorista komentoja SCILAB-komentoikkunaan. Tämä vaatii yhden näppäinpainalluksen enemmän, mutta on ehkä tutumpana turvallisempaa.

Kuvia ja musiikkia vektoreilla

Siirrytään nyt katsomaan vektorikäsitteen erilaisia ilmenemismuotoja. Nähdään, että yleistys uusiin ulottuvuuksiin ei ole pelkkää matemaatikkojen harjoittamaa esteettistä, maailmalle vierasta abstrahointihöpinää, vaan se avaa niin silmiä kuin korviakin näkemään ja kuulemaan uusilla tavoilla. Viimemainitusta olkoon esimerkkinä digitaalinen äänen käsittely, josta itsekin voi nauttia CD-levyjä ja äänitiedostoja kuunnellesaan. Ensinmainitusta voitaisiin vastaavasti ottaa digitaalinen kuvankäsittely, mutta asiaa harkittuani valitsin tällä kertaa äänen. Toki kuviakin sentään katsellaan funktioiden kuvajien muodossa.

Funktion f kuvaajan piirto käsiopelilla tapahtui yksinkertaisimmillaan niin, että laskettaisiin tasavälisessä pisteistössä $x_1, x_2, \dots, x_n$ funktion arvot $y_i = f(x_i), i = 1, \dots, n$, pyöräyteltäisiin kynällä pisteet koordinaatistoon ja yhdistettäisiin peräkkäiset pisteet janapätkillä.

Juuri näin tekee `plot`-funktio MATLAB/SCILAB:ssa. Peruskutsu on `plot(x,y)`, missä x ja y ovat samanpituisia vektoreita. Ohjelmissamme on funktio `linspace`, jolla saadaan aikaan tasavälinen x -pisteistö. Seuraavassa esimerkissä muodostetaan x -vektori jakamalla väli $[-\pi, \pi]$ 15:een osaan, ja sitten y -vektori laskemalla sinifunktion arvot x -pisteissä. Kaikki matemaattiset funktiot näissä vektoriohjelmissa toimivat niin, että sovellettaessa funktiota (x -)vektoriin, tuloksena saadaan (y -)vektori, joka koostuu vastaavista funktion arvoista. Tämä tekee käytön hyvin vaivattomaksi.

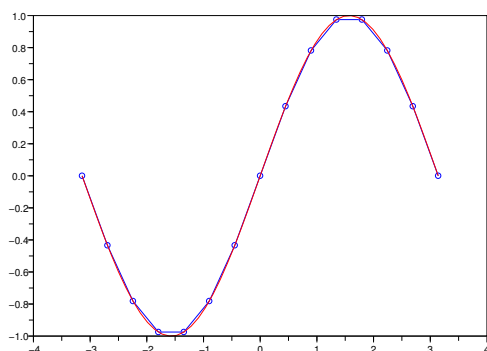
Pari pientä, hyödyllistä yksityiskohtaa:

1. Puolipiste (`:`) komennon perässä estää tulostuksen ruudulle.
2. Kummassakin ohjelmassa on erikoissymboleja joillekin vakioille, π on MATLAB:ssa `pi` ja SCILAB:ssa `%pi`. Toimimme tähän tapaan:

```
pi=%pi // Matlab-yhteensopivuuden vuoksi
x=linspace(-pi,pi,15); y=sin(x);
[x' y'] // xy-arvojen taulukko
clf() // Grafiikkaruudun pyyhkiminen
plot(x,y); // Pisteet ja murtoviiva
plot(x,y,'o') // Pelkät pisteet
```

Piirretään tarkemmin jakamalla väli 50:een osaan. Piirretään samaan kuvaan punaisella värillä. Nähdään, että tässä on riittävä määrä "näytteitä" sini-funktiosta $\sin(x)$ välillä, jotta silmä näkee kuvaajan kauniisti kaartuvana sinikäyränä.

```
x=linspace(-pi,pi,50); y=sin(x);
plot(x,y,'r') // 'r' viittaa väriin "red"
```



SININ KUVAAJA, näytteitä 15 ja 50

Musiikkia vektoreilla

Edellä näimme, miten vektoreilla voidaan saada aikaan silmänruokaa vaikkapa sini-funktion kuvaajan ihastelemiseksi. Saataisiinko myös korvin kuultavaa? Jatkaamme sini-funktion parissa.

Perussävelenä tarkastelkaamme yksiviivaista a -säveltä. Kyseessä on siniaalto, jonka värähtelytaajuus on 440 Hz, ts. 440 värähdystä sekunnissa. Värähtelyn kulma-
taajuus on tällöin $2\pi 440$ radiaania sekunnissa ja sitä esittää siniaalto $y(t) = \sin(2\pi 440 t)$.

Kun tällaista siniaaltoa digitoidaan, täytyy näytteenottotaajuuden olla mielellään selvästi enemmän kuin $2 \cdot 440$. Otetaan varman päälle: 10 000. Seuraavassa istunnossa jaetaan ensin 1 sekunnin aikaväli $(0, 1)$ 10 000:een osaan, piirretään 100 ensimmäistä kuvaajan pistettä vastaten 1/100 sekunnin aikaa ja sitten kuunnellaan. Funktio `sound` on samaa korvalle kuin `plot` on silmälle.

```
a1=2*pi*440; // Scilab
t=linspace(0,1,10000); y2=sin(a1*2*pi*t);
clf(); plot(t(1:100), y2(1:100))
sound(y2, 10000);
```

Koska lukija ei lehden sivulta saa ääntä kuuluviin, niin jätettäkään kuvakin tulostamatta.

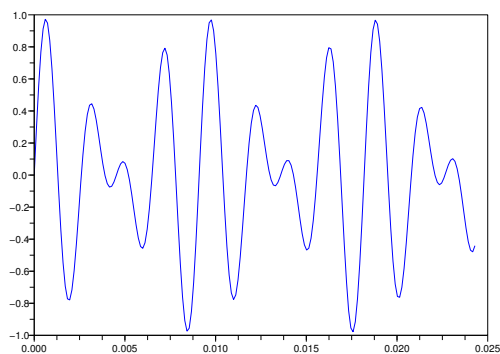
Sointuja

Esittelemme muutaman perusidean ja työkalun digitaalisen äänenkäsittelyn alalla. Muodostetaan yksiviivaisen oktaavin nuottien taajuuudet ja sijoitetaan muut-
tajiin

```
c=261.6; d=293.7; e=329.6; f=349.2;
g=392; a=440; h=493.9;
```

Jospa soittaisimme sekunnin verran intervallia *kvintti*, joka voidaan toteuttaa vaikkapa viulun d - ja a -kieliä yhtäaikaan soittamalla. Käytetään näytteenottotaajutta `nt=8192`. Intervalli saadaan laskemalla värähtelyjen summa, ts. laskemme näytevektorit yhteen. Skaalataan jakamalla 2:lla, koska `sound`-funktio haluaa värähtelyn arvoalueen (y -arvot) välille $[-1, 1]$.

```
nt=8192; pi=%pi; t=linspace(0,1,nt);
kvintti=0.5*sin(2*pi*d*t)+0.5*sin(2*pi*a*t);
sound(kvintti, nt);
clf(); plot(t(1:200), kvintti(1:200))
```



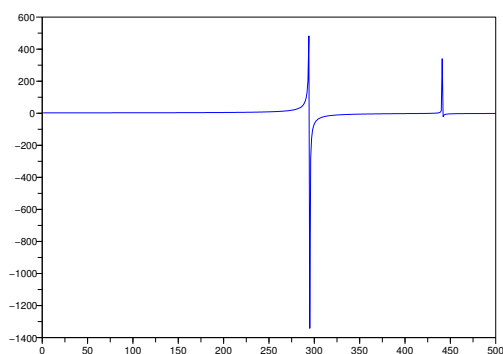
KVINTTI: Viulun d- ja a-kielet soivat.

Raottaaksemme ovea hiukan signaalinkäsittelyn matemaattisten perustyökalujen suhteen, otamme esimerkkinä nopeasta Fourier-muunnoksesta, FFT ("Fast Fourier Transform"). Kyseessä on väline, joka on mainittu joissain arvioissa 1900-luvun merkittävimpiin matemaattisiin algoritmeihin kuuluvaksi. Paino on sanalla "nopea", joka on mahdollistanut reaaliaikaiset digitaalisen signaalinkäsittelyn sovellukset.

Tässä yhteydessä emme selitä, mitä Fourier-muunnos tarkoittaa. Katsomme edellistä esimerkkiä jatkaen, minkälaiseen maailmaan tuo muunnos meidät vie.

MATLAB/SCILAB:ssa on funktio `fft`, jota käytämme "mustana laatikkona". Tämä muunnos toimii kaikkein tehokkaimmin vektoreille, joiden pituus on jokin 2^n potenssi. Näytteenottotaajuutemme valittiin siitä syystä luvuksi 8192, kun tämä sattuu olemaan 2^{13} . Mainittakoon, että CD-levyn musiikin näytteenottotaajuus on 44 100 näytettä sekunnissa. (Tämä ei ole 2^n potenssi, syy on se, että $2^{15} = 32768$ on liian pieni, jotta korkeimmat ylä-äänit eivät vääristyisi ja $2^{16} = 65536$ on aivan turhan suuri.)

Suoritetaanpa nyt kvintti-vektorillemme Fourier-muunnos SCILAB:n `fft`-funktia käyttäen.



KVINTIN FOURIER-MUUNNOS Viulun d- ja a-taajuudet erottuvat.

```
Fkvintti=fft(kvintti);
clf();plot(Fkvintti(1:500))
```

Kahden sinivärähtelyn kuva on muunnettu "taajuustasoon". Vaaka-akselilla on taajuudet ja pystyakseli edustaa kunkin taajuuden tehoon viittaavaa arvoa. Nähdään, että kvintin taajuuskuvassa kaikki muut taajuudet ovat nolla-tehoisia, paitsi ~ 293 Hz ja ~ 440 Hz.

Käänteismuunnoksella päästään takaisin. Se on nimikään suoraan saatavilla ohjemissamme nimellä `ifft`. Jos haluaisimme viulun kieliä vaikka vähän virittää, se tapahtuisi helposti taajuustasossa, sitten palaisimme käänteismuunnoksella aikatasoon soittamaan puhdasta kvinttiä.

Tuiki tuiki tähtönen – Halleluja

Tehdään pieni viihteellinen/taiteellinen sivuhyppäys, kun meillä nyt on koko nuotisto näpeissämme. Kirjoitetaan vektori `tuiki`, jossa on laulun nuotit. Oletetaan, että edellä olevat nuottien taajuusarvot on sijoitettu muuttujiin `c, d, ...`

```
tuiki=[c,c,g,g,a,a g f f e e d d];
tuiki=[tuiki c g g f f e e d];//loistat vaan..
```

Yksinkertaisella ohjelmasilmukalla voimme nyt soittaa suositun lastenlaulun. Komento `halt()` pysäyttää ohjelman siihen saakka, kunnes painetaan jotain näppäintä.

```
t=linspace(0,1,8192);
for k=1:length(tuiki)
    halt()
    y=sin(tuiki(k)*t);sound(y,8192)
end;
```

Jos laulat kuorossa, etkä soita sujuvasti mitään instrumenttia, voit tällä menetelmällä harjoitella helposti omaa stemmaasi, otin sen itse käyttöön näiden kehitysten seurauksena.

Jotta pääsisimme käsiksi hiukan vaativampaan taidenautintoon, mainittakoon, että MATLAB:ssa on digitoituna Ote *Händelin* Messias-oratorioon kuuluvasta Halleluja-kuorosta. Kas näin:

```
>> load handel
>> sound(y) % Halleluja halleluja ...
           % Ei Hard Rock!
```

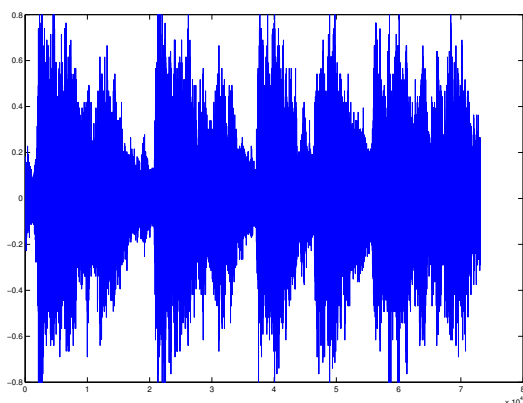
Katsotaan hiukan tuota Halleluja-vektoria:

```
>> length(y) % Vektorin pituus
ans =
    73113

>> y(1:5) % 5 ensimmäistä komponenttia

ans =
     0
-0.0062
-0.0750
-0.0312
    0.0062 % Ihan totta, pelkistä tylsistä
           % numeroista koostuva vektori!

>> plot(y)
```



HÄNDELIN MESSIAS: HALLELUJA-KUORO

Lineaarisia yhtälöryhmiä

Palatkaamme korkealentoisista taidenautintojen sfääreistä takaisin maanpinnalle arkisempaan aherukseen.

Matriisilaskennan ensimmäinen ja tärkeä käyttöalue on lineaaristen yhtälöryhmien teoria ja ratkaisumenetelmät.

En tällä kertaa paneudu tähän aiheeseen muuten kuin selittämällä aivan lyhyesti ratkaisutekniikan pääperiaatteet samalla johdatellen matriiseihin.

Katsotaanpa esimerkkiä:

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

Tällaisia yhtälöryhmiä ratkaistaan niin, että yhtälöitä muokataan operaatioilla, joissa yhtälöryhmän ratkaisujoukko pysyy samana. Näitä operaatioita

ovat yhtälöiden yhteenlasku, vakiolla kertominen ja yhtälöiden järjestyksen vaihtaminen.

Koulussahan eliminointitekniikkaa harjoitellaan, mutta tietääkseni ei esitetä systemaattista tapaa, *Gaussin eliminaatiomenetelmää*. Siinä päädytään aina ”yläkolmiomuotoon”, josta yhtälöryhmän ratkaisujen olemassaolo- ja lukumääräkysymykset voidaan selvittää ja samalla saada ratkaisut lasketuksi silloin, kun niitä on.

Huomataan, että yhtälöryhmä on annettu, kun kaikki tuntemattomien kertoimet ja tunnetut oikean puolen luvut on annettu. Niinpä kaikki yhtälöryhmää koskeva informaatio sisältyy vasemmanpuoleiseen ”matriisiin”, eli lukutaulukkoon. Merkitään symbolilla ($\sim$) yhtälöryhmiä edustavien matriisien ekvivalenssia, eli ratkaisujoukkojen samuutta.

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Kirjoitusvaikeuksia voidaan hiukan säästää kohdistamalla sallitut muokausoperaatiot (”Gaussin rivioperaatiot”) suoraan tähän matriisiin ja mikä tärkeämpää, tällöin menettelyn muuntaminen tietokoneohjelmaksi käy vattomasti.

Tavoitteena on saada nollat päälävistäjän alapuolelle. Ensimmäisessä sarakkeessa on jo yksi 0, toinen saadaan näin: **Kerrotaan 1. rivi 4:llä ja lisätään kolmanteen**, vain kolmas rivi muuttuu. Näin päädytään oikeanpuoleiseen matriisiin.

Nyt voidaan 2. rivi (yhtälö) jakaa 2:lla ja sen jälkeen kertoa 3:lla ja lisätä kolmanteen. Näin päädytään matriisiin

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \leftrightarrow \begin{cases} x_1 - 2x_2 + x_3 = 0 \\ x_2 - 4x_3 = 4 \\ x_3 = 3 \end{cases}$$

Eliminaatiovaihe on valmis, päädyimme yläkolmiomuotoon. Koska x_3 :n kerroin $\neq 0$, saamme yksikäsitteisen ratkaisun ratkaisemalla alhaalta ylöspäin edeten kolme erillistä ensimmäisen asteen yhtälöä. Näemme, että saadaan yksikäsitteinen ratkaisu, olipa oikean puolen ”pystyvektori” valittu miten hyvänsä. Tässä tapauksessa saadaan alimmasta: $x_3 = 3$, sitten toisesta: $x_2 = 16$ ja vihdoin ensimmäisestä: $x_1 = 29$.

Gaussin eliminaatiomenettely voidaan nähdä kaksivaiheisena prosessina: 1) Eteenpäin eliminointi (”forward elimination”) ja takaisinsijoitus (”backsubstitution”), jälkimmäinen siinä tapauksessa, että 1. vaihe lupaa ratkaisuja olevan.

Laskentaa matriiseilla

Matriisilla tarkoitetaan yksinkertaisesti suorakulmion muotoista lukutaulukkoa. Jos rivejä on m ja sarakkeita n kappaletta, puhutaan $m \times n$ -matriisista.

Merkitään edellä olevassa esimerkissä

A :lla yhtälöryhmän kertoimien muodostamaa 3×3 -matriisia, joka siis koostuu edellä olevan matriisin 3:sta ensimmäisestä sarakkeesta.

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

Otetaan myös käyttöön vektorit $\mathbf{x} = [x_1, x_2, x_3]^T$ ja $\mathbf{b} = [0, 4, 3]^T$. Yläindeksi T viittaa ns. *transpoosiin*, joka tarkoittaa tässä, että vektorit ajatellaan ”pystyvektoreiksi”, eli kyseessä on oikeastaan 3×1 -matriisi.²

Määrittelemme **matriisi kertaa vektori**-tulon niin, että yhtälöryhmämme voidaan kirjoittaa muotoon $A\mathbf{x} = \mathbf{b}$. Siten $A\mathbf{x} =$

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - 2x_2 + x_3 \\ 2x_2 - 8x_3 \\ -4x_1 + 5x_2 + 9x_3 \end{bmatrix}$$

Toisin sanoen: tulo $A\mathbf{x}$ on (pysty)vektori, jonka komponentit saadaan sisätuloina $\mathbf{a}_1 \cdot \mathbf{x}, \mathbf{a}_2 \cdot \mathbf{x}, \mathbf{a}_3 \cdot \mathbf{x}$, kun matriisin A vaakavektoreita merkitään: $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Matriisin rivin on oltava yhtä pitkä kuin kerrottava sarakevektori, ts. matriisin sarakkeiden lukumäärän ja kerrottavan vektorin pituuden on oltava samoja. Tulosvektorin pituus on sama kuin matriisin sarakkeen pituus (= rivien lukumäärä).

Tähän saakka esittämämme määritelmät näyttävät antavan meille ainakin esteettistä nautintoa, yhtälöryhmämme saa muodollisesti hyvin yksinkertaisen asun: $A\mathbf{x} = \mathbf{b}$.

Mikään ei estäisi käyttämästä analogiaa tavallisen ensimmäisen asteen yhtälön ratkaisulle ja merkitsemästä yhtälöryhmän ratkaisua jakolaskuun viittaavilla tavoilla:

$$\mathbf{x} = \frac{\mathbf{b}}{A} = A^{-1}\mathbf{b}.$$

Jakolaskumuoto ei ole yleisessä käytössä matemaattisena merkintätapana, sensijaan jälkimmäinen muoto, jossa ratkaisu esitetään ”käänteismatriisilla” kertomisenä, on matriisilaskennan arkipäivää. Mikään ei

toki estä käyttämästä jakolaskuasymboliikkaa ainaakaan tietokoneohjelmassa. Näin onkin tehty mm. MATLAB/SCILAB:ssa:

Esimerkkiyhtälöryhmämme ratkaistaisiin näin:

```
-->A=[1 -2 1; 0 2 -8;-4 5 9]
```

```
A =
    1.    -2.     1.
    0.     2.    -8.
   -4.     5.     9.
```

```
-->b=[0;8;-9]
```

```
b =
    0.
    8.
   -9.
```

```
-->x=A\b // "matriisilla A jako"
```

```
x =
   29.
   16.
    3.
```

```
// Tarkista kertomalla: A*x, antaako b:n ?
```

Jos kyseessä olisi lukuja (1×1 -matriiseja) koskeva yhtälö, olisi yhtälön ratkaisun eli jakolaskun onnistumisen ehtona $A \neq 0$. Yhtälöryhmän tapauksessa matriisia A koskeva ehto on juuri se, johon Gaussin eliminaatiomenettely johtaa. Ehto voidaan lausua monessa muodossa, joiden esittelyyn emme tässä ryhdy.

Matriisitulo yleisesti

Osaamme muodostaa tulon $A\mathbf{b}$, kun $\mathbf{b}$ on vektori, jonka pituus³ on sama kuin matriisin rivin pituus (ts. sarakkeiden lukumäärä). Jos vaikka A on 2×3 -matriisi ja $\mathbf{b}$ on $\mathbb{R}^3$:n vektori, niin

$$A\mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b} \\ \mathbf{a}_2 \cdot \mathbf{b} \\ \mathbf{a}_3 \cdot \mathbf{b} \end{bmatrix},$$

missä $\mathbf{a}_i$ tarkoittaa A -matriisin rivivektoria numero i .

Matriisitulo AB voidaan nyt määritellä soveltamalla tuloa $A\mathbf{b}$ kuhunkin B -matriisin sarakkeeseen ja latomalla näin saadut sarakevektorit vierekkäin. Matriisin A rivin on oltava samanpituinen kuin B :n sarake. Jos A on $m \times n$ ja B on $n \times p$, niin $C = AB$ on $m \times p$. Eriytyisesti neliömatriiseja voidaan aina kertoa keskenään ja tuloksena on samankokoinen neliömatriisi.

²Vektorikäsitteen kannalta on yhdentekevää, kirjoitetaanko koordinaatit pysty- tai vaakasuoraan (tai vaikka S:n muotoon tms.), kunhan komponenttien järjestys on selvä. Matriisilaskennan kannalta on huomattavaa laskentateknistä hyötyä apukäsitteistä pysty- ja vaakavektori.

³Matriisilaskennassa puhutaan usein vektorin ”pituudesta”, kun oikeasti tarkoitetaan dimensiota.

Matriisialgebraa

Samankokoisille matriiseille voidaan määritellä yhteenlasku vastinalkioittain aivan samoin kuin vektoreilla. Samoin määritellään luvun c ja matriisin tulo kertomalla kukin matriisin alkio c :llä. Samat laskusäännöt ovat voimassa. (Itse asiassa $(m \times n)$ -matriiseja voidaan näitä laskutoimituksia ajatellen pitää $m \times n$ -ulotteisina vektoreina.)

Kun otamme kertolaskun mukaan, puhumme jatkossa yksinkertaisuuden vuoksi vain neliömatriiseista, jolloin matriisit ovat aina kertomiskelpoisia.

Kaikki vektorilaskennasta tutut laskusäännöt pätevät, mutta miten toimivat laskusäännöt, kun mukaan otetaan matriisikertolasku. Osoittautuu, että yhtä poikkeusta lukuunotamatta kaikki tutut algebran säännöt ovat voimassa. Niinpä liitännäisyys ja osittelulait pätevät. Siis $A(BC) = (AB)C$, $A(B+C) = AB+AC$, jne. Todistukset seuraavat matriisitulon määritelmästä suoraviivaisesti, mutta eivät ole samalla tavoin itsestäänselvyyksiä kuin vektorien laskuominaisuudet.

Entä päteekö vaihdannaisuus, eli onko $AB = BA$?

Tämä voidaan heti osoittaa vastaesimerkillä vääräksi. Voit ottaa melkein mitkä tahansa, vaikkapa 2×2 -matriisit ja kertoa ne keskenään molemminpäin. Huomaat, että saat eri tulokset. Kokeile vaikka:

$$A = \begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}.$$

$$AB = \begin{bmatrix} 4 & -4 \\ 3 & -3 \end{bmatrix}, \quad BA = \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix}.$$

Suorita laskut käsin! Kokeile myös SCILAB:lla, siinä kertomerkki (*) tarkoittaa juuri matriisikertolaskua. kuten edellä jo näimme.

Kenties elämäsi ensimmäisen kerran olet tekemisissä sellaisen laskentajärjestelmän kanssa, jossa kertolasku ei ole vaihdannainen.

Neliömatriisille A voidaan määritellä **potenssi** aivan kuten luvuille:

$$A^p = \underbrace{AA \cdots A}_p \text{ kertaa}$$

Tehtävä: Päteekö neliömatriiseille binomin neliön kaava? Ellei, niin mihin muotoon on kirjoitettava:

$$(A+B)^2 = A^2 + B^2 + \text{---} ?$$

Dynaamisia systeemejä, Markovin prosesseja

Monilla sovellusalueilla, kuten ekologiassa, kansantaloudessa, erilaisissa insinööritieteissä mallinnetaan ajan

mukana muuttuvaa systeemiä. Systeemin tilaa ajanhetkellä t_k kuvatkoon vektori $\mathbf{x}_k$. Puhutaan diskreetistä dynaamisesta systeemistä. Yksinkertaisinta tyyppiä on lineaarinen systeemi, jossa lähdetään alkutilasta $\mathbf{x}_0$ ja seuraavaan tilaan päästään kertomalla neliömatriisiilla A . Tällöin $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$, ..., yleisesti:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, k = 0, 1, 2, \dots$$

Kuten heti nähdään, tehtävälle voidaan kirjoittaa ratkaisukaava: $\mathbf{x}_k = A^k\mathbf{x}_0$. Tässä yksi syy tarpeeseen oppia keinot, joilla matriisipotensseja voidaan tehokkaasti laskea.

Tarkastelemme erityistä tyyppiä olevia matriiseja, ns. *stokastisia matriiseja*, joita myös *Markovin matriiseiksi* kutsutaan. Näissä kaikki alkiot ovat ei-negatiivisia ($a_{ij} \geq 0$) ja sarakesummat ovat $= 1$. Tällaisen matriisin määräämää dynaamista systeemiä kutsutaan *Markovin prosessiksi*.

Eduskuntavaalit

Aiheen ajankohtaisuuden vuoksi valitsen esimerkin, jossa joudun sotkeutumaan puoluepolitiikkaan. Aineistona käytän HS:n sunnuntaina 25.2.07 s. A4 julkaisemia, Suomen Gallupin keräämiä puolueiden kannatuslukemia. Lehdessä on julkaistu 9×9 -matriisi, jossa riveillä on puolueiden kannatusprosentit. Matriisin sarakkeet edustavat kannatusmittauksia keskimäärin noin $2 : n$ kuukauden välein alkaen helmi-maaliskuusta 2006. (Sarake 1, Eduskuntavaalien 2003 tulos, jätetään pois.) Tehtävän yksinkertaistamiseksi tarkastelen vain kolmen suurimman puolueen keskinäisiä kannatussuoksia. Annetusta matriisista valitaan siten 3 ensimmäistä riviä ja kukin sarake jaetaan sarakesummalla. (Voidaan tehdä hyvin elegantilla matriisilausekkeella MATLAB/SCILAB:ssa.)

Näin meillä on Gallup-matriisi G :

$$\begin{bmatrix} 0.320 & 0.325 & 0.325 & 0.324 & 0.317 & 0.314 & 0.311 \\ 0.313 & 0.316 & 0.327 & 0.331 & 0.337 & 0.341 & 0.343 \\ 0.367 & 0.359 & 0.348 & 0.344 & 0.346 & 0.344 & 0.346 \end{bmatrix}.$$

Rivit ovat järjestyksessä (Kok, Ke, Sd) ja ajat 2/2006 – 2/2007.

Kysymykset 1. Voidaanko prosessia kuvata Markovin matriisilla A ?

2. Jos voidaan, niin mistä tuo matriisi saadaan?

Ohitetaan toistaiseksi nämä kysymykset ja tempaistaan ns. "hatusta" tällainen Markovin matriisi:

$$A = \begin{bmatrix} 0.93 & 0.02 & 0.04 \\ 0.03 & 0.93 & 0.04 \\ 0.04 & 0.05 & 0.92 \end{bmatrix}$$

Olkoon $\mathbf{x}_0 = [k_0, c_0, d_0]^T$, k =Kokoomus, c =Keskusta ("center"), d =Demarit. Nyt

$$A\mathbf{x}_0 = \begin{bmatrix} 0.93k_0 + 0.02c_0 + 0.04d_0 \\ 0.03k_0 + 0.93c_0 + 0.04d_0 \\ 0.04k_0 + 0.05c_0 + 0.92d_0 \end{bmatrix} = \mathbf{x}_1 = \begin{bmatrix} k_1 \\ c_1 \\ d_1 \end{bmatrix}$$

Rivi 1 sanoo: 93% Kok-kannattajista pysyy, 2% Ke-kannattajista ja 4% Sd-kannattajista siirtyy Kok-kannattajaksi. Aivan vastaavasti muut. Siten A-matriisin sarakkeet ilmaisevat sijamuotoa elätiivi (sta/stä) ja rivit illatiivia (een,uun,hin). Huomaa, että sarakesummat ovat 1 (Jokaisesta siirrytään jonnekin, mahdollisesti samaan.), mutta rivisummien ei tarvitse olla 1, voisihan joku puolue ääritapauksessa vaikka romahtaa totaalisesti ja päätyä 0-kannatukseen. Tämä näkyisi matriisissa ao. puolueen kohdalla nollarivinä. Matriisimme on varsin "lävistäjävaltainen", mikä merkitsee voimakasta puolueuskollisuutta.

Näin jatkaen johdumme Markovin prosessiin

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, k = 0, 1, 2, \dots$$

Matriisin A "tempaamisesta" todettakoon, että siinä on otettu huomioon lävistäjävaltaisuus ja lisäksi tehty kokeiluja ja säätöjä aineistoon nojautuen. Matriisia ei voida ratkaista aineiston perusteella saatavissa olevasta 9×9 -yhtälöryhmästä, tämä ei yleensä tuottaisi Markovin matriisia.

Vektoria $\mathbf{x}_0$ voimme kutsua todennäköisyysvektoriksi: alkioit ovat ei-negatiivisia ja niiden summa = 1.

Todetaanpa tärkeä perusominaisuus yleiselle ($n \times n$) Markovin matriisille: Jos $\mathbf{x}_0$ on todennäköisyysvektori, niin samoin ovat $\mathbf{x}_1, \mathbf{x}_2, \dots$

Perustelu:

1. Koska kukin $a_{ij} \geq 0$, ja kukin $x_i \geq 0$, lasketaan matriisi kertaa vektori-tulossa ei-negatiivisia termejä yhteen ja tulos ≥ 0 .

2. Vektorin $\mathbf{v}$ alkioiden summa on sama kuin sisätulo (eli matriisitulo) $[1, 1, \dots, 1] \mathbf{v}$ (Ajatellaan $\mathbf{v}$ pystyvektorina, kuten yleensäkin.) Markovin matriisin sarakesummat = 1, joten $[1, 1, \dots, 1] A = [1, 1, \dots, 1]$, ja niinpä $\mathbf{x}_1$:n alkioiden summa = $[1, 1, \dots, 1] (A\mathbf{x}_0) = ([1, 1, \dots, 1] A) \mathbf{x}_0 = [1, 1, \dots, 1] \mathbf{x}_0 = 1$.

Tässä käytimme matriisitulon liitännäisyyttä.

Mitäpä tapahtuu kelpo puolueillemme, saavatko selvää toisistaan?

Lähdetään iteroimaan helmikuusta 2006, ts. valitaan $\mathbf{x}_0 = G$ -matriisin ensimmäinen sarake: Iteroidaan ihan vaan käsin kirjoitellen ohjelmallemme:

```
--> A=[.93 .02 .04;.03 .93 .04;.04 .05 .92];
--> x0=G(:,1) // G:n 1. sarake
--> x1=A*x0; x2=A*x1; x3=A*x2; x4=A*x3;
--> x5=A*x4; x6=A*x5; x7=A*x6;
--> [x0 x2 x4 x7]
```

'2/06'	'6/06	'10/06'	'2/07'	
0.320	0.317	0.315	0.312	'Kok'
0.313	0.317	0.321	0.325	'Kepu'
0.367	0.365	0.364	0.363	'Sdp'

Kun verrataan viimeistä saraketta yllä annetun G -matriisin viimeiseen sarakkeeseen, nähdään, että Kokoomus on varsin tarkkaan kohdallaan, Kepu:n ja Sdp:n välillä mallinnuksemme on hiukan liikaa kallellaan jälkimmäiseen päin, tosin ajan kuluessa tuo näyttäisi hitaasti korjaantuvan. No, koska matriisimme näinkin hyvin ennustaa "oikeita" ennusteita, jatkamme eteenpäin mallia (A -matriisia) muuttamatta.

Miten sitten käy vaaleissa? Tällä mallilla ja aineistolla on turha toivoa saatavan mitään parempaa kuin G -matriisin viimeinen sarake. Joudun tässä suhteessa tuottamaan lukijalle pettymyksen, nuo mainiot matriisitkaan eivät auta meitä ennustamaan tulosta. Sensijaan voidaan kysyä, mitä tapahtuu pitkällä aikavälillä. Jatketaanpa iterointia, itse asiassa se on hyvin helppoa, koska $\mathbf{x}_n = A^n \mathbf{x}_0$ ja matriisipotenssi ohjelmissamme on A^n

```
--> [x0 (A^50)*x0 (A^70)*x0 (A^80)*x0]
ans =
    0.3201    0.3027    0.3025    0.3025
    0.3129    0.3360    0.3361    0.3361
    0.367     0.3613    0.3613    0.3613
```

Huomataan, että viimeistään 70:n aikajakson jälkeen, eli n. 10 vuoden kuluttua prosessi on päätynyt kiintopisteeseen, josta se ei enää etene. Entä mitä vaikuttaa alkujakau-
ma. Kokeillaan vaikka $\mathbf{x}_0 = [.6 \ .3 \ .1]^T$;

```
--> [x0 (A^10)*x0 (A^30)*x0 (A^50)*x0 (A^80)*x0]
ans =
    0.6    0.3989    0.3133    0.3038    0.3026
    0.3    0.3143    0.3318    0.3355    0.3361
    0.1    0.2868    0.3549    0.3607    0.3613
```

```
--> [x0 (A^10)*x0 (A^30)*x0 (A^50)*x0 (A^100)*x0]
ans =
    0.6    0.3989    0.3133    0.3038    0.3025
    0.3    0.3143    0.3318    0.3355    0.3361
    0.1    0.2868    0.3549    0.3607    0.3613
```

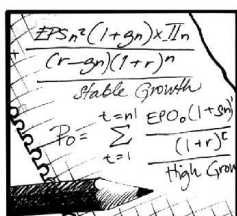
Huomataan, että aivan "vinossa olevasta" alkujakau-
masta päädytään rajalla **samaan tasapainojakau-
maan**. Toden totta, Markovin prosessi johtaa (tie-
tyin yleisin lisäehdoin) **samaan kiintopisteeseen**,

olipa **alkujakauma mikä tahansa**. Mutta miksi? Perustelu voidaan esittää hyvin tyylipuhtaasti ominaisarvoteorian sovelluksena. Sen esittely on pakko jättää johonkin myöhempään yhteyteen. Aihetta voit opiskella alla mainituista lähteistä sekä Googlella hakusanoilla *ominaisarvot*, *eigenvalues*, *eigenvectors*. Kerromme kuitenkin tuloksen: Rajajakauma on suurinta ominaisarvoa (1) vastaava ominaisvektori todennäköisyysvektoriksi normeerattuna. Annan lopuksi vaihteeksi aidon MATLAB-ajon, jossa on tätä matriisiamme koskevat ominaisarvo/-vektoriloitsut, jotka paljastavat taikuuden:

```
>> A=[.93 .02 .04;.03 .93 .04;.04 .05 .92];
>> [V,D]=eig(A);
>> lambda1=D(1,1);v1=V(:,1);
>> lambda1
lambda1 =
    1.0000
>> v1/sum(v1) % ominaisvektorin normeeraus.
    0.3025
    0.3361
    0.3613
```

Viitteet

1. David C. Lay *Linear Algebra and its applications*, 3rd ed., Addison Wesley, 2003.
2. Gilbert Strang *Introduction to Linear Algebra*, 3rd ed., Wellesley-Cambridge Press, 2003.
Luentovideot MIT:n sivulla:
<http://ocw.mit.edu/OcwWeb/Mathematics/18-06Spring-2005/VideoLectures/index.htm>
3. Wikipedia: Historiaa ja yleiskatsaus runsaine viitteineen
http://en.wikipedia.org/wiki/Linear_algebra
4. Heikki Apiola
Luentoja TKK:lla syksyllä 2006
<http://math.tkk.fi/opetus/kp3-ii/06/L/>
5. Matlab/Scilab-kotisivut
<http://www.mathworks.com/>
<http://www.scilab.org/>
5. Heikki Apiola–Marko Laine
Lyhyt Matlab-opas
<http://math.tkk.fi/~apiola/matlab/opas/lyhyt/>
6. Google-hakusanoja:
matriisi, *matrix*, *vector*, *lineaarialgebra*, *linear algebra*.



Yleistetyt funktiot eli distribuutiot

Jukka Liukkonen

Helsingin ammattikorkeakoulu Stadia

Johdanto

Matematiikan välineillä reaali maailman ilmiöt pyritään pelkistämään niin, että jäljelle jää oleellinen informaatio ja vain se. Esimerkiksi taivaankappaleiden liikkeitä tarkasteltaessa on useimmiten yhdentekevää, mikä on kappaleen tarkka muoto tai tiheysjakauma; sen sijaan tähti ajatellaan sen painopisteeseen sijoitettuna pistemassana vailla ulottuvuuksia. Tiheysjakauman ja muun turhan informaation ottaminen mukaan malliin tekisi sen vaikeammin käsiteltäväksi ilman, että monimutkaisesta mallista suurella vaivalla laskettu liikerata olisi oleellisesti tarkempi kuin yksinkertaisesta mallista helposti ratkaistava rata. Samalla tavoin seinän läpi kulkevaa ääniaaltoja laskettaessa seinän ajatellaan koostuvan yhdestä tai useammasta tasapaksusta kerroksesta, joilla ei ole hienorakennetta. Matemaattisen sileät pystysuorat tasot erottavat kerrokset terävästi toisistaan ja ympäröivästä homogeenisestä ilmasta.

Tässä kirjoituksessa suhtaudutaan kriittisesti funktion käsitteeseen ja yleistetään se distribuution käsitteeksi. Jälkimmäinen auttaa karsimaan mallista epäolennaisuuksia. Sen avulla differentiaalilaskennan valtava voima ulotetaan sinnekin, missä derivointia ei klassisessa mielessä voida tehdä. Distribuutiot tekevät mahdolliseksi kaikenulotteisten tiheysjakaumien (pistemassat, pistemäiset dipolit, pintavaraukset jne.) täsmällisen käsittelyn. Distribuutioderivaattaan on sisäänrakennettuna mekanismi, joka gene-

roi automaattisesti eräät fysiikan sovelluksissa esiintyvät alkuehdot ja materiaalirajapintaehdot. Distribuutioiden ansiosta taajuusanalyysissä ei tarvitse rajoittaa äärellisen energian omaaviin signaaleihin. Esityksen yksinkertaisuuden ja selkeyden vuoksi tässä käsitellään ainoastaan yksiulotteisia malleja. Tarkastelutavat käsitteet ja tulokset yleistyvät kuitenkin useampaan ulottuvuuteen.

Distribuution erikoistapaus on mitta. Sitä voidaan havainnollistaa jakaumana (engl. ja ransk. *distribution*), joka voi kuvata esimerkiksi massan, sähköjohtavuuden tai todennäköisyyden jakautumista. Fyysikko *Paul Diracin* (1902–1984) mukaan nimetty Diracin mitta tai Diracin delta (ks. kappale Distribuutiot ja Diracin delta) kuvaa yhteen ainoaan pisteeseen keskittynyttä tiheysjakaumaa. Deltaa voitaisiin luonnehtia sellaiseksi normaalijakaumaksi, jossa keskihajonta on degeneroitunut nolaksi. Dirac esitteli deltan 1920-luvulla, mutta tietävästi sitä käytti jo 1800-luvulla *Oliver Heaviside* (1850–1925) — kauan ennen varsinaisen distribuutioteorian syntymä. Heaviside ja Dirac käsittelivät deltaa kuin se olisi vain omituinen funktio. Eräät lähteet pitävät distribuutioiden isänä venäläistä matemaatikkoa *Sergei Sobolevia* (1908–1989). Hän käytti nykydistribuutioita läheisesti muistuttavia funktion yleistymiä tutkimuksissaan 1930-luvun loppupuolella. Varsinaisen distribuutioteorian kehitti kuitenkin ranskalainen *Laurent Schwartz* (1915–2002) vuosina 1944–1948, Sobo-

levista riippumatta. Vaikka nimitys distribuutio viittaa mittaamiseen, kaikki distribuutiot eivät ole mittoja. Esi-merkkinä mainittakoon deltan derivaatta.

Mitä vikaa funktioissa?

Reaaliarvoisen funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ on sääntö, jolla reaaliin x liitetään säännön yksikäsitteisesti määräämä reaali $y = f(x)$, joka siis yleensä riippuu luvusta x . Tällainen funktiokäsitys sallii esimerkiksi funktion g määrittämisen asettamalla

$$g(x) = \begin{cases} x^2 + 1, & x \in \mathbb{Z} \\ x^2, & x \notin \mathbb{Z}. \end{cases} \quad (2)$$

Useimmissa reaalimaailman malleissa funktion arvot yksittäisissä pisteissä ovat kuitenkin merkityksettömiä; sen sijaan keskiarvon tyyppiset tunnusluvut ovat oleellisia. Esim. tyypillisessä fysiikan mittauksessa havainnoidaan jotakin ilmiötä hetken aikaa, ja mittaustulos on jonkinlainen keskiarvo — sen muodostumiseen tarvitaan pieni aikaväli. Integroituvan funktion f välillä $[a, b]$ saamien arvojen keskiarvo on

$$Af(a, b) = \frac{1}{b-a} \int_a^b f(x) dx = \int_{-\infty}^{\infty} f(x) \varphi_{a,b}(x) dx, \quad (3)$$

missä

$$\varphi_{a,b}(x) = \begin{cases} 1/(b-a), & x \in [a, b] \\ 0, & x \notin [a, b]. \end{cases} \quad (4)$$

Funktion f arvo kussakin pisteessä x voitaisiin korvata raja-arvolla

$$\bar{f}(x) = \lim_{\varepsilon \rightarrow 0^+} Af(x-\varepsilon, x+\varepsilon) = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} f(t) dt. \quad (5)$$

Valitettavasti raja-arvoa (5) ei ole aina olemassa.⁴ Niiden pisteiden joukko, joissa raja-arvoa ei ole olemassa, on kuitenkin hyvin pieni; matemaatikot puhuvat *nollamittaisesta* joukosta. Hieman yksinkertaistaen sanottuna sellaisen joukon osasten yhteenlaskettu pituus on nolla. Ei merkitse oleellista vääryyttä sopia funktion $\bar{f}(x)$ arvoksi nolla näissä poikkeuspisteissä. Tällä tavoin keskiarvoistetut funktiot sisältäisivät useimmiten kaiken oleellisen informaation. Lisäksi niissä on se hyvä piirre, että kaksi keskiarvoistettua funktiota eroavat toisistaan vain, jos ne eroavat toisistaan oleellisesti — niiden olisi mahdotonta saada eri arvot pelkästään esimerkiksi erillisistä pisteistä koostuvassa, kokonaislukujen joukon $\mathbb{Z}$ kaltaisessa joukossa. Esimerkin (2) tapauksessa $\bar{g}(x) = x^2$.

⁴ Jos oletetaan, että funktio f on rajoitetusti heilahteleva rajoitetuilla väleillä, raja-arvo (5) on olemassa kaikissa pisteissä x . Funktio heilahtelee rajoitetusti välillä $[a, b]$, jos summilla

$$\sum_{k=1}^n |f(t_k) - f(t_{k-1})|, \quad a \leq t_0 < t_1 < \dots < t_n \leq b,$$

on olemassa yhteinen äärellinen yläraja, jota suuremmiksi ne eivät voi tulla.

Testifunktiot

Funktion tärkeiden piirteiden tuntemiseksi riittää siis tietää integraalien

$$\int_{-\infty}^{\infty} f(x) \varphi_{a,b}(x) dx$$

arvot ”testifunktioilla” $\varphi_{a,b}$, kun $a, b \in \mathbb{R}$, $a < b$. Distribuutioteoriassa on tämänkaltaisen lähtökohta. Siellä otetaan **testifunktioksi** epäjatkuvien funktioiden $\varphi_{a,b}$ asemesta kaikki sellaiset funktiot φ , jotka häviävät (so. saavat arvon nolla) jonkin funktiosta φ riippuvan rajoitetun välin ulkopuolella, ja joilla on kaikkien kertalukujen derivaatat jokaisessa lukusuoran $\mathbb{R}$ pisteessä. Silloin sekä φ että kyseiset derivaatat ovat derivoituvina funktioina myös jatkuvia. Tällaisten funktioiden joukkoa merkitään $\mathcal{D}(\mathbb{R})$. Suljetun välin $[a, b]$ ulkopuolella häviävien testifunktioiden joukolle käytetään tässä merkintää $\mathcal{D}(a, b)$. Integraalia

$$\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x) \varphi(x) dx. \quad (6)$$

sanotaan **funktion f arvoksi testifunktiolla φ** . Jos φ ei saa negatiivisia arvoja, ja

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1, \quad (7)$$

funktio φ on ns. **painofunktio**, ja integraali (6) on funktion f painotettu keskiarvo. Kun f on jatkuva, $x \in \mathbb{R}$ ja ε on riittävän pieni positiivinen luku, painofunktiolla $\varphi \in \mathcal{D}(x-\varepsilon, x+\varepsilon)$ painotettu keskiarvo on lähellä funktion tarkkaa arvoa $f(x)$, sitä lähempänä mitä pienempi ε on.

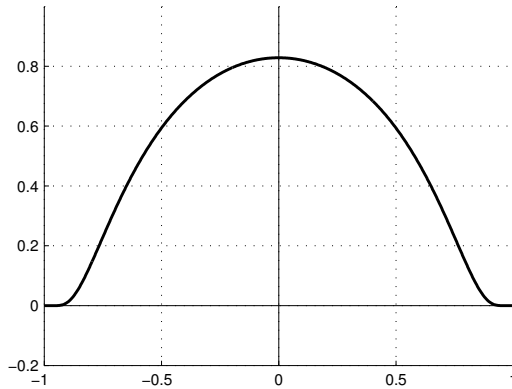
Millaisia testifunktioita sitten on olemassa? Ainakin vakiofunktio 0 kelpaa testifunktioksi. Toinen kelvollinen testifunktio on

$$\phi(x) = \begin{cases} M^{-1} e^{1/(x^2-1)}, & -1 < x < 1 \\ 0, & \text{muulloin,} \end{cases} \quad (8)$$

missä

$$M = \int_{-1}^1 e^{1/(x^2-1)} dx.$$

Sen derivoituvuusominaisuuksien tarkistaminen välin $[-1, 1]$ päätepisteissä jää lukijalle.⁵ Funktion ϕ kuvaaja on piirretty kuvaan 1.



Kuva 1. Testifunktion ϕ kuvaaja välillä $[-1, 1]$.

Siirtämällä ja skaalaamalla funktiosta ϕ saadaan kokonainen parvi

$$\phi_{x_0, \rho}(x) = \frac{1}{\rho} \phi\left(\frac{x - x_0}{\rho}\right), \quad x_0 \in \mathbb{R}, \quad \rho \in \mathbb{R}, \quad \rho > 0, \quad (9)$$

testifunktioita, jotka toteuttavat ehdot

- (a) $\phi_{x_0, \rho}(x) \geq 0$ kaikilla $x \in \mathbb{R}$,
- (b) $\phi_{x_0, \rho}(x) > 0$ vain, kun $x_0 - \rho < x < x_0 + \rho$, ja

$$(c) \int_{-\infty}^{\infty} \phi_{x_0, \rho}(x) dx = 1.$$

Distributiot ja Diracin delta

Koska funktion f arvoista yksittäisissä pisteissä ei olla kiinnostuneita, riittää tuntea kuvaus

$$\mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}, \quad \varphi \mapsto \langle f, \varphi \rangle, \quad (10)$$

millä tarkoitetaan siis funktion f arvojen tuntemista eri testifunktioilla. Integraalin ominaisuuksista seuraa testifunktioille φ ja ψ sekä vakioille α ja β yhtälö

$$\langle f, \alpha\varphi + \beta\psi \rangle = \alpha\langle f, \varphi \rangle + \beta\langle f, \psi \rangle. \quad (11)$$

Tällöin sanotaan, että kuvaus (10) on **lineaarinen**. Se on myös jatkuva, mutta ennen jatkuvuuden täsmällistä

määrittelemistä pitää sopia siitä, mitä tarkoitetaan testifunktiojonon suppenemisella. Alla ”nollas derivaatta” $\varphi^{(0)}$ tarkoittaa itse funktiota $\varphi \in \mathcal{D}(a, b)$ ja

$$\|\varphi\| = \max_{a \leq x \leq b} |\varphi(x)| = \max_{x \in \mathbb{R}} |\varphi(x)| \quad (12)$$

on suurin jatkuvan funktion $x \mapsto |\varphi(x)|$ suljetulla välillä $[a, b]$ saamista arvoista.

Määritelmä 1. Testifunktioiden $\varphi_n \in \mathcal{D}(\mathbb{R})$ jono suppenee kohti testifunktiota φ eli $\lim_{n \rightarrow \infty} \varphi_n = \varphi$, jos seuraavat ehdot ovat voimassa:

- (a) On olemassa sellaiset luvut a ja b että $\varphi_n \in \mathcal{D}(a, b)$ olipa n mikä positiivinen kokonaisluku tahansa.
- (b) Jos r on positiivinen kokonaisluku tai nolla,

$$\lim_{n \rightarrow \infty} \|\varphi_n^{(r)} - \varphi^{(r)}\| = 0. \quad (13)$$

Lineaarisen kuvauksen $u : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$ arvoja on tapana merkitä $u(\varphi) = \langle u, \varphi \rangle$, vaikka kyseessä ei olisikaan nimenomainen kuvaus (10).

Määritelmä 2. Lineaarinen kuvaus

$$u : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}, \quad \varphi \mapsto \langle u, \varphi \rangle$$

on **jatkuva**, jos

$$\lim_{n \rightarrow \infty} |\langle u, \varphi_n \rangle| = 0 \quad \text{aina, kun} \quad \lim_{n \rightarrow \infty} \varphi_n = 0. \quad (14)$$

On helpohko harjoitustehtävä osoittaa, että kuvaus (10) on jatkuva edellisen määritelmän mielessä.⁶ Tässä vaiheessa monella lukijalla on varmaankin jo aavistus siitä, miten funktion käsite yleistetään distribuution käsitteeksi:

Määritelmä 3. Jatkuvia lineaarisia kuvauksia $u : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$ sanotaan **yleistetyiksi funktioiksi** eli **distributioksi**. Niiden joukkoa merkitään $\mathcal{D}'(\mathbb{R})$.

Kun funktio $f : \mathbb{R} \rightarrow \mathbb{R}$ liittyy reaalityyppiseen x reaalityyppiseen $f(x)$, distributio $u : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$ liittyy testifunktioon φ reaalityyppiseen $\langle u, \varphi \rangle$. Jos funktio f tulkitaan distributioksi yhtälön (6) kautta, sen arvo $\langle f, \varphi \rangle$ on eräänlainen yleistetty, paikallinen, testifunktiolla φ painotettu keskiarvo funktion f tavallisista arvoista $f(x)$. Kaikki distributiot eivät suinkaan ole funktioita. Kuuluisin tällainen distributio on **deltadistributio** eli **Diracin delta** δ_{x_0} , joka määritellään yhtälöllä

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0). \quad (15)$$

Lause 1. Diracin delta on distributio.

⁵Lukion pitkän matematiikan tiedoilla tehtävä lienee vaativa. Ratkaisussa voidaan käyttää differentiaalilaskennan väliarvolausetta ja eksponenttifunktion kasvuominaisuuksia suhteessa rationaalifunktion kasvuominaisuuksiin.

⁶Integraali $\langle f, \varphi_n \rangle$ lähestyy nollaa jo silloin, kun pelkästään $\|\varphi_n\|$ lähestyy nollaa. Derivaattoja ei siis tarvitse käsitellä.

Todistus: Jos α ja β ovat vakioita ja φ sekä ψ testifunktioita,

$$\begin{aligned}\langle \delta_{x_0}, \alpha\varphi + \beta\psi \rangle &= (\alpha\varphi + \beta\psi)(x_0) \\ &= \alpha\varphi(x_0) + \beta\psi(x_0) \\ &= \alpha\langle \delta_{x_0}, \varphi \rangle + \beta\langle \delta_{x_0}, \psi \rangle.\end{aligned}$$

Siis δ_{x_0} on lineaarinen. Jos testifunktioiden φ_n jono suppenee kohti nollatestifunktiota, erityisesti

$$\lim_{n \rightarrow \infty} \|\varphi_n\| = 0.$$

Silloin

$$|\langle \delta_{x_0}, \varphi_n \rangle| = |\varphi_n(x_0)| \leq \max_{x \in \mathbb{R}} |\varphi_n(x)| = \|\varphi_n\| \rightarrow 0,$$

kun $n \rightarrow \infty$. Siis δ_{x_0} on jatkuva. $\square$

Vaikka δ_{x_0} ei olekaan funktio, sen voidaan ajatella synnyttävän yhtälön (9) määrittelemien testifunktioiden $\phi_{x_0, \rho}$ raja-arvona, sillä

$$\begin{aligned}\lim_{\rho \rightarrow 0^+} \langle \phi_{x_0, \rho}, \varphi \rangle &= \lim_{\rho \rightarrow 0^+} \int_{-\infty}^{\infty} \phi_{x_0, \rho}(x) \varphi(x) dx \\ &= \varphi(x_0) = \langle \delta_{x_0}, \varphi \rangle.\end{aligned}$$

Lukijaa kehoitetaan näyttämään toteen, että $\alpha u + \beta v$ on distribuutio aina, kun α ja β ovat vakioita ja u sekä v distribuutioita. Tämä havainto antaa menetelmän konstruoida tunnetuista distribuutioista uusia. Testifunktioita rakennettaessa voidaan käyttää hyväksi myös seuraavaa aputulosta, jonka todistus sivuutetaan:

Lemma 2. *Olkkoon funktiolla f kaikkien kertalukujen derivaatat avoimella välillä $]a, b[$. Jos $a < c < d < b$, on olemassa testifunktio $\varphi \in \mathcal{D}(a, b)$, joka yhtyy funktion f suljetulla välillä $[c, d]$; ts. $\varphi(x) = f(x)$ kaikilla $x \in [c, d]$.*

Esimerkki 1. *Ajatellaan kahta toisiaan vakioetäisyydellä kiertävää tähteä, joiden massat ovat m_1 ja m_2 . Sijoitetaan x -akseli kulkemaan tähtien painopisteiden kautta. Silloin x -akseli pyörii tähtien mukana. Valitaan tähtiparin painopiste origoksi, ja merkitään yksittäisten tähtien paikkakoordinaatteja x_1 ja x_2 . Jos tähtipari halutaan nähdä kokonaisuutena, sen määrittämiseksi malliksi voidaan valita distribuutio $u_* = (m_1 + m_2)\delta_0$. Mikäli tähtien pitää erottua toisistaan, malliksi otetaan $v_* = m_1\delta_{x_1} + m_2\delta_{x_2}$. Tähtien massat ja sijainnit saadaan tarvittaessa selville ”pommitamalla” malleja testifunktioilla samaan tapaan kuin hiukkasfyysikko pommittaa atomiytimiä:*

$$\langle u_*, \varphi \rangle = (m_1 + m_2)\varphi(0), \quad (16)$$

$$\langle v_*, \varphi \rangle = m_1\varphi(x_1) + m_2\varphi(x_2). \quad (17)$$

Testauksessa on ikäänkuin käytettävissä tietokoneohjelma, joka tulostaa tapauksessa (16) luvun $\langle u_*, \varphi \rangle$ ja tapauksessa (17) vastaavasti luvun $\langle v_*, \varphi \rangle$ millä tahansa

testifunktiosyötteellä φ . Sopivia syötteitä käyttämällä ja ohjelman tulosteita tutkimalla pitäisi selvittää, mitkä ovat ohjelman sisäiset parametrit m_1 , m_2 , x_1 ja x_2 . Jos esim. tähtien yhteenlaskettua massaa ei tiedetä, se saadaan selville mallista u_* sellaisella testifunktiolla φ_0 , jolle $\varphi_0(0) \neq 0$; nimittäin $\langle u_*, \varphi_0 \rangle = (m_1 + m_2)\varphi_0(0)$, jolloin $m_1 + m_2 = \langle u_*, \varphi_0 \rangle / \varphi_0(0)$.

Tehtävä 1. *Lukijaa pyydetään miettimään testifunktiostrategia, jolla saadaan mahdollisimman vaivattomasti selville kummankin tähden massa ja sijainti mallista u_* .*

Distribuution derivaatta

Jos funktiot f ja g derivaattoineen ovat jatkuvia välillä $[a, b]$, pätee osittaisintegroitikaava

$$\int_a^b f'(x)g(x) dx = \left[f(x)g(x) \right]_a^b - \int_a^b f(x)g'(x) dx.$$

Korvaamalla g testifunktiolla $\varphi \in \mathcal{D}(a, b)$ keskimäinen termi häviää ja integroinnit voidaan ulottaa yli koko lukusuoran:

$$\int_{-\infty}^{\infty} f'(x)\varphi(x) dx = - \int_{-\infty}^{\infty} f(x)\varphi'(x) dx. \quad (18)$$

Yhtälön (18) oikea puoli on määritelty silloinkin, kun f ei ole derivoituva eikä edes jatkuva. Riittää, kun f on integroituva välillä $[a, b]$. Jos integroituvalla funktiolla h pätee

$$\int_{-\infty}^{\infty} h(x)\varphi(x) dx = - \int_{-\infty}^{\infty} f(x)\varphi'(x) dx \quad (19)$$

kaikilla $\varphi \in \mathcal{D}(\mathbb{R})$, funktiota h sanotaan integroituvan funktion f **heikoksi derivaataksi**. Yhtälö (18) osoittaa, että heikko derivaatta on derivaatan käsitteen yleistys. Kirjoittamalla kaava (18) muotoon

$$\langle f', \varphi \rangle = -\langle f, \varphi' \rangle \quad (20)$$

saadaan vihdoin ilmiselvä vihje siitä, miten distributioderivaatta tulee määritellä:

Määritelmä 4. *Distribuution u distribuutioderivaatta eli lyhyesti derivaatta on kuvaus*

$$u' : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}, \quad \langle u', \varphi \rangle = -\langle u, \varphi' \rangle. \quad (21)$$

Yhtälö (18) osoittaa, että derivoituvan funktion klasinen derivaatta distribuutioksi tulkittuna on samalla distribuutioksi tulkitun funktion distribuutioderivaatta.

Lemma 3. *Distribuution derivaatta on distribuutio.*

Todistus: Jos α ja β ovat vakioita ja φ sekä ψ testifunktioita, distribuution u lineaarisuuden nojalla

$$\begin{aligned}\langle u', \alpha\varphi + \beta\psi \rangle &= -\langle u, (\alpha\varphi + \beta\psi)' \rangle \\ &= -\langle u, \alpha\varphi' + \beta\psi' \rangle \\ &= -\alpha\langle u, \varphi' \rangle - \beta\langle u, \psi' \rangle \\ &= \alpha\langle u', \varphi \rangle + \beta\langle u', \psi \rangle.\end{aligned}$$

Siis u' on lineaarinen. Jos testifunktiojono φ_n suppenee kohti nollatestifunktiota, erityisesti

$$\lim_{n \rightarrow \infty} \|\varphi'_n\| = 0.$$

Koska u on jatkuva,

$$|\langle u', \varphi_n \rangle| = |\langle u, \varphi'_n \rangle| \rightarrow 0,$$

kun $n \rightarrow \infty$. Siis u' on jatkuva. $\square$

Funktiota

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \leq x \end{cases}$$

sanotaan **Heavisiden funktioksi** omaperäisen englantilaisen fyysikon ja sähköinsinöörin Oliver Heavisiden mukaan. Funktio H ei selvästikään ole derivoituva eikä edes jatkuva kohdassa $x = 0$, mutta mikä on sen distribuutioderivaatta H' ? Koska H on integroitava rajoitetuilla väleillä, distribuution H arvot voidaan laskea integraaleina:

$$\begin{aligned}\langle H', \varphi \rangle &= -\langle H, \varphi' \rangle = -\int_{-\infty}^{\infty} H(x)\varphi'(x) \, dx \\ &= -\int_0^{\infty} \varphi'(x) \, dx \\ &= -\left(\lim_{x \rightarrow \infty} \varphi(x) - \varphi(0)\right) = \varphi(0)\end{aligned}$$

kaikilla testifunktioilla φ . Heavisiden funktion distribuutioderivaatta on näin ollen deltadistribuutio:

$$H' = \delta_0. \quad (22)$$

Määritelmän mukaan deltan δ_{x_0} distribuutioderivaatta on kuvaus

$$\langle \delta'_{x_0}, \varphi \rangle = -\langle \delta_{x_0}, \varphi' \rangle = -\varphi'(x_0). \quad (23)$$

Lopuksi helppo harjoitustehtävä lukijalle:

Tehtävä 2. *Määritä distribuutioderivaatat funktioille*

$$|x|, \quad \frac{x}{|x|}, \quad H(x) - H(x-1) \quad \text{ja} \quad \frac{\cos x}{|\cos x|}.$$

Jos funktion määrittelevä lauseke on osamäärä, sen arvolla nimittäjän nollakohdassa, kuten muissakaan yksittäisissä pisteissä, ei ole merkitystä; se voidaan sopia vaikkapa nolllaksi.

Viitteet

- [1] Dautray, R. & Lions, J.-L.: *Mathematical Analysis and Numerical Methods for Science and Technology*. Vol. 2 *Functional and Variational Methods*. Springer-Verlag, Berlin 1988.
- [2] Hörmander, L.: *The Analysis of Linear Partial Differential Operators I* (2nd ed.). Springer-Verlag, Berlin 1990.
- [3] Rudin, W.: *Functional Analysis*. McGraw-Hill, New York 1973.
- [4] Schwartz, L.: *Théorie des distributions*. Hermann & Cie, Paris 1966.



Riemannin hypoteesi ymmärrettäväksi

Matti Lehtinen

Maanpuolustuskorkeakoulu

John Derbyshire: Alkulukujen lumoissa. Suom. Juha Pietiläinen. Terra Cognita 2006. 404 s. Ovh. 40 euroa.

Matematiikan tutkimus ylittää uutiskynnyksen yleensä vain silloin, kun jokin kuuluisa ongelma saa ratkaisunsa. *Andrew Wiles* ja Fermat'n suuren lauseen todistus oli esillä kymmenisen vuotta sitten, Poincarén konjektuuri ja sen originelli ratkaisija *Grigori Perelman* viime vuonna. Matemaatikkojen laaja konsensus on, että eräs tärkeimmistä ellei tärkein suuri avoin ongelma on *Riemannin hypoteesi*. Se on myös yksi seitsemästä amerikkalaisen Clay-instituutin miljoonan dollarin Millennium-ongelmasta. Lukuteoreettinen kirjallisuus vilisee tuloksia, jotka alkavat ”jos oletamme, että Riemannin hypoteesi on tosi, niin...”

John Derbyshire, englantilaissyntyinen mutta Yhdysvalloissa asuva kirjailija, joka tunnetaan mm. oikeaa laitaa kulkevan *National Review* -lehden kommentaattorina ja mitä erilaisimpiin asioihin kantaa ottavana ja kiistoja herättävänä nettikirjoittajana, on ottanut tehtäväkseen kertoa maallikkolukijalle, mistä Riemannin hypoteesissa oikeastaan on kysymys ja *Kimmo Pietiläisen* kulttuurikustantamo *Terra Cognita* on tuoreeltaan saattanut kirjan suomenkielisten lukijoiden ulottuville.

Kun tiedämme, että Riemannin hypoteesi on lyhyesti seuraava: kompleksimuuttujan s funktion eli *Rieman-*

nin zetafunktion

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$$

epätriviaalien nollakohtien reaaliosa on $\frac{1}{2}$, ymmärrämme, että Derbyshiren tehtävä on perin haastava. Derbyshire kohdistaa kirjansa ”älykkäälle ja uteliaalle, mutta matemaattisesti oppimattomalle lukijalle” ja arvelee kirjan vaikeustason sopivaksi lukion matematiikan edes välttävästi suorittaneelle.

Derbyshire pyrkii ratkaisemaan tehtävänsä, ei siis itse Riemannin hypoteesia vaan sen tavalliselle lukijalla kuvailemisen ongelman, kuljettamalla kirjassaan kahta juonta lomittain. Luvut, joiden järjestysnumero on pariton, kehittelevät Riemannin hypoteesin ymmärtämiseksi tarvittavaa matematiikkaa. (Derbyshire kertoo aikoneensa ensin käyttää tähän tarkoitukseen lukuja, joiden järjestysnumero on alkuluku, mutta arvelleensa sitten, että menetelmä olisi vaikuttanut liialliselta kikkailulta. – Hiukan huvittavaa tässä yhteydessä on suomen kielen sanan *luku* kaksoismerkitys.) Niissä kohdataan päättymätön sarja, alkuluvun käsite ja alkulukujen lukumäärälauseet, potenssin ja logaritmin käsitteet, ζ -funktio reaaliaralueella ja sen yhteys alkulukuihin, derivaatta ja integraali, kompleksiluvut, kompleksimuuttujan funktiot, kunnat ja viimein ζ -funktion ja alkulukujen jakautumisen kytkentä. On ymmärrettävää, että syvemmälle analyttisen lukuteorian viidakkoon kuljettaessa kirjoittaja joutuu

enemmän ja enemmän luopumaan johdonmukaisesta esityksestä ja turvautumaan ”uskokaa minua, näin se on” -tyyppisiin perusteluihin.

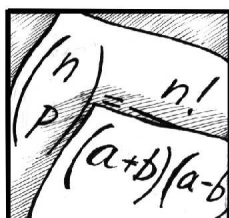
Kirjan parillisnumeroisissa luvuissa puhutaan enemmän ihmisistä ja historiasta: *Bernhard Riemannista* itsestään ja hänen edeltäjistään *Eulerista*, *Gaussista* ja *Dirichlet’stä*, alkulukulauseen todistuksen vaiheista, *David Hilbertin* kuuluisasta vuoden 1900 matemaatikkokongressin esitelmästä, joka viimeistään kanonisoi Riemannin hypoteesin matematiikan tavoittelulimpien hedelmien joukkoon, ja 1900-luvun lukuteoriasta aina melko uusiin tietokoneisiin tukeutuviin Riemannin hypoteesin kukistamiseen tähtääviin yrityksiin asti. Runsaat alaviitteet sisältävät vielä monia mielenkiintoisia lisätietoja.

Kirjan liitteeksi on painettu tunnetun matemaatikon *Tom Apostolin* kirjoittama monisäkeistöinen Riemannin hypoteesin ympärille kiertyvä laulunsanoitus. Kun laulussa mainitaan useamman kerran (*Ernst*) *Lindelöf*, Suomen matematiikan tutkimuksen isäksi luonnehdittavissa oleva matemaatikko, on Derbyshire kirjoittanut vielä toisen liitteen, jossa on laajahko esitys Riemannin hypoteesille läheistä sukua olevasta *Lindelöfin hypoteesista*. Lindelöfin ihmeellisen monipuolisen matemaati-

kontyön kietoutuminen myös Riemannin hypoteesiin ei ole kovin laajalti tunnettua, Suomessakaan.

Derbyshiren yritys on hyvä, mutta en kuitenkaan tullut aivan vakuutetuksi siitä, että hänen tarkoittamansa lukija lopulta ymmärtäisi hypoteesin merkittävyyden ja sen ratkaisemisen vaikeuden. Matematiikan opettajalle samoin kuin matematiikasta kiinnostuneelle lukiolaiselle kirjan lukemista voin kuitenkin lämpimästi suositella jo siksikin, että vaikka suomenkielistä matematiikan populaarikirjallisuutta alkaa aika paljon olla, kiitos mm. puheena olevan kirjan kustantajan, ei niistä juuri löydy sellaista tietoa, joka jotenkin kytkisi koulumatematiikan matematiikan tutkimuksen keskeisiin osiin. Ja kuitenkin, kun oppiaineemme nimi on matematiikka eikä laskento, olisi kohtuullista odottaa, että meillä olisi edes aavistuksia siitä, mitä matematiikka tieteenä on.

Kirjan suomennoksen on tarkastanut paras suomalainen analyyttisen lukuteorian asiantuntija, professori *Matti Jutila*. Terminologisia epäselvyyksiä tai ilmeisiä väärinymmärryksiä ei juuri tullut vastaan. Sujuvaa käännöstä tuntuu paikoin rasittavan hiukan liian tarkka pitäytyminen alkukielen ilmauksiin.



Ratkaisu aikaisempaan tehtävään

Pekka Alestalo

Matematiikan laitos, Teknillinen korkeakoulu

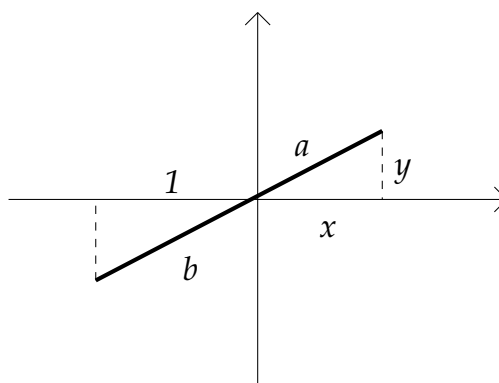
Esitin Solmun numerossa 2/2006 kaksi tehtävää, joista ensimmäisen ratkaisu ilmeistyi edellisessä numerossa; nyt on toisen vuoro.

Tilanne 2: Lieriön muotoiseen tyhjään juomalasiin asetetaan mehupilli. Pilli on kallellaan niin, että sen alaosa vastaa pohjan reunaan ja yläosa yltää juuri ja juuri lasin reunaan (vastakkaisella puolella). Lasiin kaadetaan hitaasti limonadia, jolloin pilliin kiinnittyy kuplia ja se alkaa nousta lasista. Oletetaan, että pillin alapää nousee suoraan ylöspäin lasin sivua pitkin ja että pillin tukipiste lasin yläreunassa pysyy samana (eli tilanne on tiettyssä mielessä kaksiulotteinen). Juoman kaatamista jatketaan niin kauan, että lasi täyttyy ja pilli on lopuksi vaakasuorassa.

Ongelma 2: Oletetaan, että lasin poikkileikkauksen halkaisija on 1 ja pillin pituus 2. Kuinka korkealla (lasin yläreunasta mitattuna) pillin yläpää enimmillään on, ja mikä on tällöin pillin kaltevuuskulma vaakatasoon nähden? Anna vastauksena korkeuden tarkka ja likiarvo sekä kulman likiarvo.

Ratkaisu 2: Kuten tehtävässä huomautetaan, tilannetta voidaan käsitellä kaksiulotteisena. Asetetaan koordinaatiston origo lasin reunaan niin, että pillin

päät ovat aluksi pisteissä $(-1, -\sqrt{3})$ ja $(0, 0)$, lopuksi pisteissä $(-1, 0)$ ja $(1, 0)$. Pillin yläpään korkeutta kuvaa silloin funktio $f: [0, 1] \rightarrow \mathbf{R}$, jolle $f(x) \geq 0$ kaikilla x ja lisäksi $f(0) = f(1) = 0$.



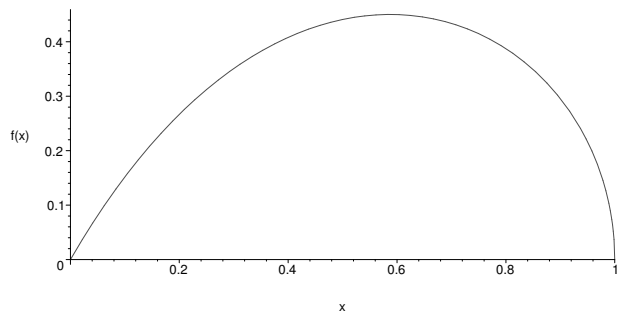
Aloitamme funktion f lausekkeen määrittämisestä. Jos pillin yläpää on pisteessä (x, y) ja origo jakaa pillin kuvion mukaisesti kahteen osaan, joiden pituudet ovat a ja b , niin $a + b = 2$. Toisaalta yhdenmuotoisista kolmioista saadaan

$$\frac{a}{x} = \frac{b}{1} = b,$$

joten $a = 2x/(1+x)$. Näin ollen

$$y = \sqrt{a^2 - x^2} = \frac{x}{1+x} \sqrt{3-2x-x^2} = f(x)$$

on etsimämme funktion lauseke. Tarkistuksena voidaan laskea $f(0) = f(1) = 0$ ja piirtää funktion f kuvaaja.



Kyseessä on derivoituva funktio, jolla sekä kuvion että intuitiivisen päättelyn mukaan on yksikäsitteinen maksimikohta. Se saadaan selville ratkaisemalla yhtälö $f'(x) = 0$, jossa derivaatan laskeminen jätetään lukijan

harteille. Kun saatu yhtälö kerrotaan puolittain termillä $\sqrt{3-2x-x^2}$, se sievenee lopulta muotoon

$$\frac{x^3 + 3x^2 + 3x - 3}{(1+x)^2} = 0.$$

Tehtäväksi jää silloin kolmannen asteen yhtälön $x^3 + 3x^2 + 3x - 3 = 0$ ratkaiseminen. Ennen suoraviivaista numeerista ratkaisua kannattaa yrittää yhtälön sieventämistä siirtämällä kolmannen asteen termin käännepiste origoon. Koska

$$\frac{d^2}{dx^2}(x^3 + 3x^2 + 3x - 3) = 6x + 6 = 0$$

pisteessä $x = -1$, niin muuttujanvaihdoilla $z = x - (-1) = x + 1$, eli $x = z - 1$, alkuperäinen yhtälö yksinkertaistuu sievennysten jälkeen muotoon $z^3 - 4 = 0$, ja ratkaisuksi saadaan $x_0 = z_0 - 1 = \sqrt[3]{4} - 1 \approx 0,587$. Huomautettakoon, että yllä käytetyn siirtomenetelmän avulla jokaisesta 3. asteen yhtälöstä voidaan poistaa 2. asteen termi, mutta tässä tapauksessa myös 1. asteen termi sattumalta (?) hävisi.

Pillin yläpää nousee siis lasin yläreunasta enimmillään korkeudelle

$$f(x_0) \approx 0,45.$$

Kyseisessä kohdassa pillin kaltevuuskulmalle α_0 pätee $\tan \alpha_0 = f(x_0)/x_0$, josta saadaan likiarvoksi $\alpha_0 \approx 0,654$ eli noin 37,5 astetta.