Polynomifunktioiden juuret

Heli Mattila

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Elokuu 2018

Tiivistelmä: Heli Mattila, *Polynomifunktioiden juuret* (engl. *Roots of Polynomial Functions*), matematiikan pro gradu -tutkielma, 37. s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, elokuu 2018.

Tämän tutkielman tarkoituksena on koota polynomifunktioiden juurien olemassaoloon ja ratkaisemiseen liittyviä tuloksia. Polynomifunktiot ovat matematiikassa perustyökaluja, joiden juurien selvittämisellä on pitkä historia. Juurien arvojen selvittämisen lisäksi usein on mielenkiintoista ja riittävää tietää vain juurien lukumäärä. Kompleksilukujen joukossa jokaisella polynomilla on astelukunsa verran juuria. Reaalilukujen joukossa juurien määrä kuitenkin vaihtelee riippuen polynomien kertoimista.

Ensimmäisen asteen polynomifunktion juurien ratkaiseminen on triviaalia. Toisen asteen polynomiyhtälöiden juuret saadaan yksinkertaisella ratkaisukaavalla, joka on paljon käytetty matemaattinen työkalu. Kolmannen asteen yhtälön ratkaisukaavoja kutsutaan Cardanon kaavoiksi. Nämä kaavat ovat monimutkaisempia ja vaativat kompleksiluvuilla laskemista, minkä vuoksi niiden käyttäminen ei ole yhtä yleistä kuin toisen asteen yhtälön ratkaisukaavan. Toisen ja kolmannen asteen yhtälöille voidaan polynomin kertoimien avulla määritellä niin kutsuttu diskriminantti, joka kertoo kyseisen polynomin reaalijuurien lukumäärän ilman, että juurten arvoja täytyy määrittää. Myös neljännen asteen yhtälön juuret voidaan ratkaista algebrallisesti ratkaisukaavan avulla. Tätä ratkaisukaavaa käytetään kuitenkin vain harvoin käytännön tilanteissa, sillä sen käyttäminen työlästä vaatiessaan ensin kolmannen asteen yhtälön ratkaisemisen.

Korkeamman kuin neljännen asteen polynomiyhtälöille ei voida määrittää ratkaisukaavaa eikä niitä siis voida ratkaista algebrallisesti. Näiden polynomifunktioiden reaalijuurille voidaan kuitenkin etsiä likiarvoja numeerisilla menetelmillä, kuten merkin muuttumiseen perustuvalla menetelmällä tai Newtonin menetelmällä. Hyvä keino korkeamman asteisen yhtälön ratkaisemisessa on myös alentaa tutkittavan polynomin astelukua. Nimittäin jos polynomifunktion eräs juuri tiedetään, voidaan polynomi ilmaista juuren lineaarisen tekijän sekä alkuperäistä polynomia asteluvultaan yhden alemman polynomin tulona. Tulon nollasäännön nojalla alkuperäisen polynomin muut juuret saadaan ratkaisemalla tämän tulon tekijänä olevan alempiasteisen polynomin juuret. Polynomin erästä juurta voidaan etsiä esimerkiksi rationaalijuuritestillä, joka kertoo mitkä ovat polynomin mahdolliset rationaalijuuret.

Sisältö

LUKU 1

Johdanto

Tämän tutkielman aiheena ovat polynomifunktiot ja erityisesti niiden juurien määrittäminen. Tutkielmassa esitellään ja johdetaan algebralliset ratkaisukaavat toisen, kolmannen ja neljännen asteen polynomifunktioiden juurille. Tutkielman pääpainotus on erityisesti reaalijuurien etsimisessä, mutta ratkaisukaavojen yhteydessä selvitetään myös aidot kompleksijuuret. Lisäksi tutkielmassa esitellään muun muassa Sturmin lause, rationaalijuuritesti sekä reaalijuurien likiarvojen ratkaisemiseen käytettäviä numerisia menetelmiä.

Tulevan matematiikan aineenopettajan pro gradu -tutkielmalle aihe valikoitui, koska se sivuaa koulumatematiikkaa monilla luokka-asteilla. Tutkielman ajatuksena onkin koota oleellisimmat tulokset polynomifunktioiden juurien määrittämisestä ja olla myös soveltuvin osin erityisesti lukion oppilaiden ymmärrettävissä. Idea aiheelle syntyi Eero Saksmanin Matematiikkalehti Solmun artikkelista *Kolmannen asteen yhtälöä ratkaisemassa* [**9**], jossa Saksman pyrkii johtamaan kolmannen asteen yhtälön ratkaisukaavan siten, että se on ymmärrettävissä lukion pitkän matematiikan kurssien pohjalta. Tutkielman aihe tarjoaa kuitenkin myös yliopiston syventävien kurssien tasoisia haasteita kompleksilukujen muodossa.

Polynomiyhtälöt ovat askarruttaneet ihmisiä läpi historian aina muinaisesta Egyptistä lähtien. Tarve ensimmäisen ja toiseen asteen yhtälöiden ratkaisemiseen kumpusi pituuksiin ja pinta-aloihin liittyvistä käytännön ongelmista. Kolmannen asteen yhtälöt nousivat kiehtovaksi kysymykseksi viimeistään antiikin Kreikassa muiden matemaattisten ongelmien kuten kulman kolmijaon johtaessa näihin yhtälöihin. Alemman asteisten yhtälöiden pohtiminen herätti kiinnostusta myös korkeamman asteisiin yhtälöihin. Aluksi yhtälöille esitettiin sanallisia ratkaisualgoritmeja sekä geometrisia ratkaisuja matemaattisten merkintöjen puutteellisuuden vuoksi. Myös itse juuren käsite on muuttunut matemaattisen tiedon ja käsitteistön lisääntyessä. Aluksi juuriksi hyväksyttiin vain positiivisia rationaalilukuja. Myöhemmin alettiin ymmärtää mahdollisiksi ratkaisuiksi myös negatiiviset luvut ja edelleen irrationaaliluvut. Kolmannen asteen yhtälön ratkaisukaavan löytyminen sai matemaatikot huomaamaan kompleksilukujen merkityksen ja sitä kautta kehittämään kompleksilukujen teoriaa. [**9**] [**10**] [**12**]

Ensimmäisen asteen yhtälöiden ratkaiseminen on triviaalia ja opitaan jo yläkoulun ensimmäisillä luokilla. Myös toisen asteen yhtälöitä käsitellään jo yläkoulussa ja opitaan ratkaisemaan ainakin vaillinaiset toisen asteen yhtälöt [**7**]. Viimeistään lukiossa käsitellään yleinen toisen asteen yhtälön ratkaisukaava. Lisäksi jotkin lukiokirjat (esimerkiksi [**3**]) tarjoavat syventävänä tietona kolmannen asteen ratkaisukaavan. Pitkän matematiikan syventävissä opinnoissa tutustutaan myös numeeriisiin ratkaisumenetelmiin kuten iterointiin ja Newtonin menetelmään [**6**].

Tutkielman toisessa luvussa esitellään polynomifunktioihin liittyviä peruskäsitteitä ja -tuloksia sekä johdetaan Sturmin lause. Kyseisen luvun päälähteenä on [**12**].

Kolmannessa luvussa johdetaan toisen asteen yhtälöön liittyvät tulokset. Luvussa neljä kootaan myöhemmissä luvuissa tarvittavia tietoja kompleksiluvuista perustuen lähteisiin [**4**] ja [**12**]. Luvut viisi ja kuusi käsittelevät kolmannen ja neljännen asteen yhtälöitä ja niissä päälähteinä ovat [**9**] ja [**12**]. Seitsemännessä luvussa pohditaan korkeamman asteen yhtälöitä sekä tutustutaan mahdollisiin keinoihin ratkaista niitä rationaalijuuritestin ja numeeristen menetelmien avulla. Tämän luvun tulokset perustuvat pääosin lähteisiin [**1**], [**5**] ja [**12**].

LUKU 2

Polynomifunktiot ja juurien olemassaolo

Määritellään aluksi polynomifunktio ja siihen liittyviä käsitteitä:

MÄÄRITELMÄ 2.1. Yleinen n:nnen asteen polynomifunktio P on muotoa

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \tag{2.1}$$

missä $a_n \neq 0$ ja n on polynomin *asteluku*. Vakiot $a_k \in \mathbb{C}$, $k \in \{1, ..., n\}$ ovat polynomin *kertoimia* ja $a_0 \in \mathbb{C}$ on polynomin *vakiotermi*.

Polynomin kertoimet voivat olla mielivaltaisia kompleksilukuja. Tässä tutkielmassa perehdytään kuitenkin pääasiassa reaalikertoimisiin polynomeihin. Määritellään sitten polynomin juuren käsite:

MÄÄRITELMÄ 2.2. Polynomifunktion P juuret ovat sen nollakohtia eli yhtälön

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{2.2}$$

ratkaisuja $x \in \mathbb{C}$.

HUOMAUTUS 2.3. Jos yhtälön (2.2) molemmat puolet kerrotaan nollasta eroavalla vakiolla, saadulla yhtälöllä on edelleen samat juuret kuin alkuperäisellä yhtälöllä. Siispä polynomifunktion juuret eivät muutu nollasta eroavalla vakiolla kerrottaessa.

Seuraava lause kertoo juurien olemassaolosta ja se tunnetaan nimellä *algebran peruslause*:

LAUSE 2.4. *Olkoon $n \geq 1$. Tällöin jokaisella n:nnen asteen polynomifunktiolla on ainakin yksi juuri.*

On huomioitavaa, että algebran peruslause ei kerro mitään siitä, onko kyseinen juuri reaalinen vai imaginaarinen. Algebran peruslauseen todistus sivuutetaan tässä tutkielmassa, mutta siihen voi perehtyä esimerkiksi [**4**, §6.19].

Määritellään seuraavaksi polynomin jaollisuus ja todistetaan sitten polynomin jaollisuudesta kertova lause.

MÄÄRITELMÄ 2.5. Olkoon B polynomi, joka ei ole nollapolynomi. Polynomi A on jaollinen polynomilla B, jos on olemassa polynomi Q siten, että

$$A(x) = Q(x)B(x)$$

kaikilla $x \in \mathbb{C}$.

Kuten kokonaisluvuille, myös polynomeille pätee jakoyhtälö

$$A(x) = Q(x)B(x) + R(x), \tag{2.3}$$

missä A, B, Q ja R ovat polynomeja. Jakajan B asteluku on jaettavan polynomin A astelukua yhden pienempi. Polynomia Q kutsutaan osamääräksi ja polynomia R

jakojäännökseksi [**5**, §7]. Edellisen määritelmän mukaan polynomi A on jaollinen polynomilla B vain jos jakoyhtälössä jakojäännös R on nollapolynomi.

LAUSE 2.6. *Olkoon x_1 polynomin P juuri. Tällöin polynomi P on jaollinen lausekkeella $x - x_1$.*

TODISTUS. Olkoon polynomin P asteluku n. Jaetaan polynomi P lausekkeella $x - x_1$, jolloin P saadaan jakoyhtälön (2.3) nojalla muotoon

$$P(x) = (x - x_1)P_1(x) + R,$$

missä P_1 on jokin astelukua $n-1$ oleva polynomi ja jakojäännös R on vakiopolynomi [**5**, §7]. Koska vakio R on sama kaikilla muuttujan x arvoilla, voidaan vakion R arvo ratkaista sijoittamalla ylläolevaan yhtälöön $x = x_1$, jolloin saadaan

$$P(x_1) = R.$$

Toisaalta x_1 on polynomin P nollakohta, joten jakojäännökselle saadaan $R = 0$. Polynomi P on siis jaollinen lausekkeella $x - x_1$. □

HUOMAUTUS 2.7. Jos siis tiedetään polynomin P jokin nollakohta x_1, polynomin P juuret ovat ratkaisuja yhtälölle

$$P(x) = (x - x_1)P_1(x) = 0,$$

missä polynomin P_1 asteluku on yhden pienempi kuin polynomin P. Tulon nollasäännön nojalla muiden nollakohtien löytäminen palautuu yhtälön $P_1(x) = 0$ ratkaisemiseen, mikä taas on yleisesti helpompaa, koska polynomin P_1 asteluku on pienempi. Siispä yhden nollakohdan löytäminen mahdollistaa ratkaistavan yhtälön asteluvun alentamisen ja sitä kautta ratkaiseminen helpottuu.

Edellisestä lauseesta sekä algebran peruslauseesta seuraa:

LAUSE 2.8. *Jokainen polynomifunktio P voidaan esittää lineaaristen tekijöiden tulona eli muodossa*

$$P(x) = a(x - x_1)(x - x_2)\cdots(x - x_n), \tag{2.4}$$

missä a on nollasta eriävä vakio.

TODISTUS. Olkoon $n \geq 1$. Algebran peruslauseen nojalla tiedetään, että n:nnen asteen polynomilla P on ainakin yksi nollakohta x_1. Siten edellisen lauseen nojalla P voidaan esittää muodossa

$$P(x) = (x - x_1)P_1(x),$$

missä P_1 on astetta $n-1$ oleva polynomi. Jos nyt $n - 1 = 0$, ollaan löydetty haluttu muoto. Jos taas $n - 1 \geq 1$, niin algebran peruslauseen nojalla polynomilla P_1 on olemassa jokin nollakohta x_2. Siten polynomi P_1 voidaan edelleen esittää muodossa

$$P_1(x) = (x - x_2)P_2(x),$$

missä P_2 on astetta $n-2$ oleva polynomi. Siten alkuperäinen polynomi P saa muodon

$$P(x) = (x - x_1)(x - x_2)P_2(x).$$

Jos edelleen $n-2 \geq 1$, voidaan jatkaa samaan tapaan eli jakaa polynomi P_2 tekijöihin juuren avulla. Tätä voidaan jatkaa kunnes päädytään tilanteeseen, jossa polynomi P saa muodon

$$P(x) = (x - x_1)(x - x_2)\cdots(x - x_n)P_n(x),$$

missä P_n on astetta $n - n = 0$ oleva polynomi eli jokin vakio $a \neq 0$. Polynomi P on siis esitetty lineaaristen tekijöiden tulona. $\square$

Tulon nollasäännön nojalla

$$P(x) = a(x - x_1)(x - x_2) \cdots (x - x_n) = 0$$

jos ja vain jos jokin tulon tekijöistä $(x - x_k)$, $k \in \{1, ..., n\}$ on nolla. Tämä vastaa sitä, että jokainen luvuista x_k on polynomin P juuri. Lisäksi tulon nollasäännön nojalla muita juuria ei voi olla.

HUOMAUTUS 2.9. (1) Edelleen on huomattava, että polynomin P lausekkeessa (2.4) juuret x_k, $k \in \{1, ..., n\}$ voivat olla reaalisia tai imaginaarisia. Lausekkeen mukainen lineaarisiin tekijöihin jako toimii siis aina kompleksilukujen joukossa, mutta ei välttämättä reaalilukujen joukossa.

(2) Jotkin lausekkeen (2.4) juurista x_k, $k \in \{1, ..., n\}$ voivat olla yhtä suuria. Jos erisuuret juuret ovat $x_1, x_2, ..., x_j$, $j \leq n$, niin lauseke (2.4) saadaan muotoon

$$P(x) = a(x - x_1)^{m_1}(x - x_2)^{m_2} \cdots (x - x_j)^{m_j},$$

missä $m_1 + m_2 + \cdots + m_j = n$. Juuri x_t, $t \in \{1, ..., j\}$ on tällöin *m_t-kertainen juuri.*

Aiemmin todetun tulon nollasäädön nojalla voidaan todeta:

LAUSE 2.10. *Jokaisella n:nnen asteen polynomilla on aina tasan n juurta kompleksilukujen joukossa, kun jokainen juuri huomioidaan yhtä monta kertaa kuin sen kertaluku on.*

Tähän asti on puhuttu polynomin kaikista mahdollisista juurista, myös imaginaarisista. Tutkitaan seuraavaksi pelkästään reaalijuuria ja pyritään löytämään yleisiä lainalaisuuksia reaalijuurien lukumäärälle. Todistetaan ensin seuraava tulos paritonta astelukua oleville polynomeille:

LAUSE 2.11. *Jos polynomin asteluku on pariton, sillä on ainakin yksi reaalijuuri.*

TODISTUS. Olkoon polynomin $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ asteluku n pariton. Jos nyt $a_n > 0$, niin

$$\lim_{x \to +\infty} P(x) = +\infty \qquad \text{ja} \qquad \lim_{x \to -\infty} P(x) = -\infty.$$

Jos taas $a_n < 0$, niin

$$\lim_{x \to +\infty} P(x) = -\infty \qquad \text{ja} \qquad \lim_{x \to -\infty} P(x) = +\infty.$$

Molemmissa tapauksissa polynomi P saa sekä positiivisia että negatiivisia arvoja. Koska P on polynomifunktiona jatkuva, sillä täytyy siis Bolzanon lauseen nojalla olla ainakin yksi nollakohta reaaliakselilla. $\square$

Samaa menettelyä ei kuitenkaan voida soveltaa asteluvultaan parillisten polynomien tapauksessa, koska niiden raja-arvot molemmissa äärettömyyksissä ovat samanmerkkiset. Tarvitaan siis muitakin keinoja reaalijuurien lukumäärän selvittämiseen.

2.1. Sturmin lause

Johdetaan seuraavaksi niin kutsuttu *Sturmin lause*, joka kertoo polynomin erisuurien reaalijuurien lukumäärän välillä $]a, b[$. Kyseistä lausetta voidaan käyttää ainoastaan polynomeihin, joilla on vain yksinkertaisia juuria. Jotta lausetta voitaisiin soveltaa myös mielivaltaisen polynomin erisuurien reaalijuurien lukumäärän selvittämiseen, on yleiselle polynomille P ensiksi löydettävä polynomi P_0, jolla on samat juuret kuin alkuperäisellä polynomilla, mutta jokainen juuri on yksinkertainen. Ennen Sturmin lauseen johtamista näytetään aluksi, miten tämä tapahtuu. Tutkitaan mielivaltaista polynomia

$$P(x) = a(x - x_1)^{m_1}(x - x_2)^{m_2} \cdots (x - x_j)^{m_j} \tag{2.5}$$

ja todistetaan ensin seuraava aputulos:

Lemma 2.12. *Polynomifunktion P m-kertainen juuri on sen derivaattafunktion P' $(m - 1)$-kertainen juuri.*

Todistus. Olkoon x_j polynomin P m-kertainen juuri. Tällöin polynomi P voidaan esittää muodossa

$$P(x) = (x - x_j)^m Q(x),$$

missä Q on polynomi, jolle $Q(x_j) \neq 0$. Polynomin P derivaataksi saadaan tästä lausekkeesta

$$\begin{aligned} P'(x) &= m(x - x_j)^{m-1}Q(x) + (x - x_j)^m Q'(x) \\ &= (x - x_j)^{m-1}[mQ(x) + (x - x_j)Q'(x)] \\ &= (x - x_j)^{m-1}R(x), \end{aligned}$$

missä merkittiin $R(x) = mQ(x) + (x - x_j)Q'(x)$. Nyt $R(x_j) = mQ(x_j) \neq 0$, joten x_j on derivaattafunktion $(m - 1)$-kertainen juuri. □

Huomautus 2.13. Sovitaan, että 0-kertainen juuri ei ole varsinaisesti juuri vaan vakio, jolloin edellinen lemma pätee myös, kun $m - 1 = 0$.

Polynomifunktion derivaatta voidaan siten kirjoittaa muodossa

$$P'(x) = (x - x_1)^{m_1 - 1}(x - x_2)^{m_2 - 1} \cdots (x - x_j)^{m_j - 1}S(x), \tag{2.6}$$

missä S on polynomi, jolle $S(x_t) \neq 0$ kaikilla $t = 1, 2, ..., j$. Polynomi S ei siten ole tutkittavan polynomin P tekijä.

Määritellään seuraavaksi polynomien suurin yhteinen tekijä, joka vastaa kokonaislukujen suurinta yhteistä tekijää [**5**, §7].

Määritelmä 2.14. Polynomi D on polynomien A ja B suurin yhteinen tekijä, jos seuraavat ehdot täyttyvät:

(1) Polynomit A ja B ovat jaollisia polynomilla D.
(2) Jos polynomit A ja B ovat jaollisia jollakin polynomilla C, niin myös polynomi D on jaollinen polynomilla C.

Polynomien suurin yhteinen tekijä voidaan määrittää Eukleideen algoritmilla vastaavasti kuin kokonaisluvuille [**5**, §7].

Vertaamalla nyt polynomin P lauseketta (2.5) ja sen derivaatan P' lauseketta (2.6), huomataan, että niiden suurin yhteinen tekijä on polynomi

$$D(x) = (x - x_1)^{m_1-1}(x - x_2)^{m_2-1} \cdots (x - x_j)^{m_j-1}. \tag{2.7}$$

Saadaan tulos:

Lause 2.15. *Polynomien P ja P' suurin yhteinen tekijä on polynomi D, jonka juuret ovat samat kuin polynomin P moninkertaiset juuret mutta kertaluvultaan yhtä pienemmät.*

Edelleen yhtälöistä (2.5) ja (2.7) huomataan, että

$$\frac{P(x)}{D(x)} = a(x - x_1)(x - x_2) \cdots (x - x_j),$$

joten saadaan:

Lause 2.16. *Polynomifunktion $\frac{P(x)}{D(x)}$ juuret ovat samat kuin polynomin P juuret ja lisäksi jokainen juuri on yksinkertainen.*

Jokaisesta polynomista P voidaan siis suurimman yhteisen tekijän D avulla saada polynomi, jolla on samat juuret kuin alkuperäisellä polynomilla mutta yksinkertaisina. Jos siis osataan selvittää polynomin $\frac{P(x)}{D(x)}$ erisuurten reaalijuurien lukumäärä, niin alkuperäisen polynomin P reaalijuurien lukumäärä on sama.

Esimerkki 2.17. Tutkitaan polynomifunktiota

$$P(x) = x^6 - 3x^2,$$

jonka derivaattafunktio on

$$P'(x) = 6x^5 - 6x.$$

Etsitään näiden suurin yhteinen tekijä Eukleideen algoritmilla. Jakamalla polynomi P derivaatallaan P' saadaan jakoyhtälö

$$P(x) = \frac{1}{6}xP'(x) - 2x^2.$$

Jakamalla derivaatta P' jakojäännöksellä $-2x^2$ saadaan

$$P'(x) = -3x^3(-2x^2) - 6x.$$

Jaetaan edelleen lauseke $-2x^2$ uudella jakojäännöksellä $-6x$, jolloin saadaan

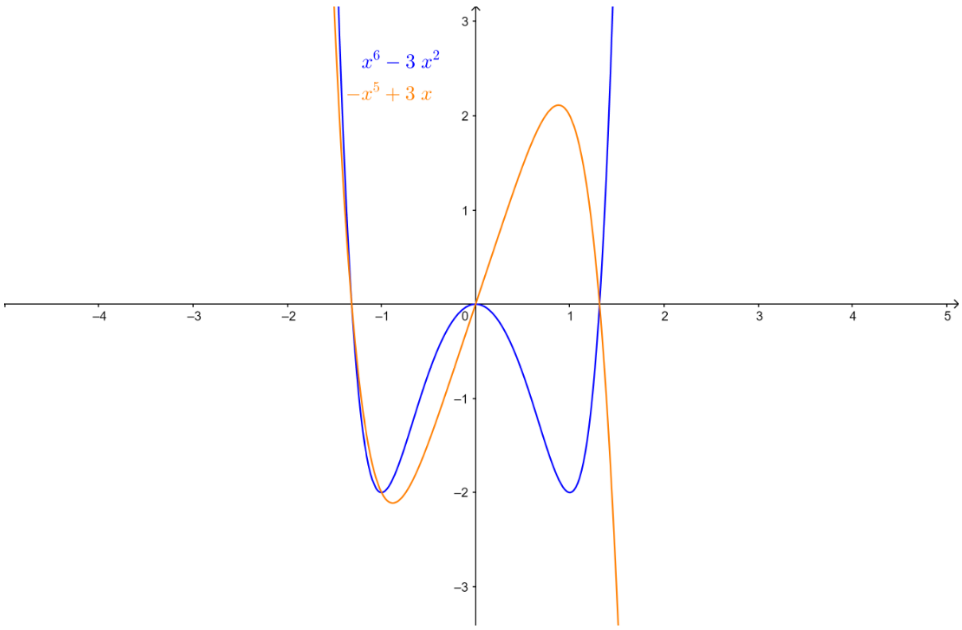
$$-2x^2 = \frac{1}{3}x(-6x) + 0.$$

Viimeisin nollasta poikkeava jakojäännös on siis $-6x$, joten se on myös polynomien P ja P' suurin yhteinen tekijä. Merkitään $D(x) = -6x$. Nyt siis polynomifunktiolla

$$\frac{P(x)}{D(x)} = -\frac{1}{6}x^5 + \frac{1}{2}x$$

on samat juuret kuin polynomilla P, mutta jokainen juuri on yksinkertainen. Koska polynomin juuret eivät muutu kerrottaessa polynomi vakiolla, voidaan saatu polynomi yksinkertaistaa myöhempää juurien etsintää varten muotoon

$$P_0(x) = -x^5 + 3x.$$



Kuva 2.1. Polynomifunktioiden $P(x) = x^6 - 3x^2$ (sininen käyrä) ja $P_0(x) = -x^5 + 3x$ (oranssi käyrä) kuvaajat piirrettynä GeoGebralla. Kuvaajien perusteella molemmilla funktioilla näyttäisi olevan samat nollakohdat.

Kuvassa 2.1 on esitetty polynomien P ja P_0 kuvaajat, joista voidaan myös huomata, että kyseisillä polynomeilla on samat nollakohdat.

Huomautus 2.18. (1) Esimerkissä polynomin asteluku pieneni yhdellä, joten sillä oli siis yksi kaksinkertainen juuri.
(2) Tällä keinolla päädytään asteluvultaan alempaan polynomifunktioon, jonka tarkastelu on yleisesti helpompaa. Siispä jos ei tarvitse selvittää tiettyjen juurien kertalukuja vaan pelkät juurien arvot tai reaalijuurien lukumäärä riittävät, voidaan ratkaisemista helpottaa tällä tavalla.

Nyt voidaan johtaa Sturmin lause. Voidaan siis olettaa, että polynomilla

$$P_0(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

on vain yksinkertaisia juuria. Siten lauseen 2.15 nojalla polynomin P_0 ja sen derivaatan P_0' suurin yhteinen tekijä on polynomi, jolla ei ole juuria eli sen täytyy olla jokin nollasta poikkeava vakio. Merkitään $P_0' = P_1$. Sovelletaan polynomeihin P_0 ja P_1 Eukleideen algoritmia ja merkitään algoritmissa peräkkäisien jakojäännöksien *vastapolynomeja* P_j sekä jakolaskujen osamääriä Q_j, missä $j = 2, 3, ...$ Saadaan yhtälöketju

$$\begin{aligned}
P_0(x) &= Q_2(x)P_1(x) - P_2(x)\\
P_1(x) &= Q_3(x)(-P_2(x)) - P_3(x)\\
&\vdots\\
P_{m-3} &= Q_{m-1}(x)(-P_{m-2}(x)) - P_{m-1}(x)\\
P_{m-2} &= Q_m(x)(-P_{m-1}(x)) - P_m(x).
\end{aligned}$$

Lopetetaan algoritmi, kun jakojäännökseksi saadaan nollasta poikkeava vakio $P_m(x)$, joka saavutetaan, koska polynomien P_0 ja P_1 suurin yhteinen tekijä on vakio. Näin saadaan polynomien muodostama *Sturmin jono*

$$P_0, P_1, ..., P_{m-1}, P_m.$$

ESIMERKKI 2.19. Esimerkissä 2.17 löydettiin polynomi $P_0(x) = -x^5 + 3x$, jolla on vain yksinkertaisia juuria. Selvitetään tätä polynomia vastaava Sturmin jono. Polynomin P_0 derivaatta on $P_1(x) = -5x^4 + 3$. Jakamalla P_0 derivaatallaan P_1 jakojäännökseksi saadaan $\frac{12}{5}x$, joten sen vastapolynomi on $P_2(x) = -\frac{12}{5}x$. Jakamalla polynomi P_1 edellisellä jakojäännöksellä saadaan uudeksi jakojäännökseksi vakio 3. Tämän vastapolynomi on $P_3(x) = -3$. Sturmin jonon muodostavat siten polynomit

$$\begin{aligned}
P_0(x) &= -x^5 + 3x\\
P_1(x) &= -5x^4 + 3\\
P_2(x) &= -\frac{12}{5}x\\
P_3(x) &= -3.
\end{aligned}$$

Todistetaan seuraavaksi muutamia Sturmin jonoon liittyviä aputuloksia:

LEMMA 2.20. *Sturmin jonolla on seuraavat ominaisuudet:*

(1) *Jonon peräkkäisillä jäsenillä ei voi olla samaa nollakohtaa x.*
(2) *Jos x_0 on polynomin P_k nollakohta, niin arvot $P_{k-1}(x_0)$ ja $P_{k+1}(x_0)$ ovat erimerkkisiä.*
(3) *Polynomin P_k nollakohdan riittävän pienessä ympäristössä polynomit P_{k-1} ja P_{k+1} ovat aidosti positiivisia tai aidosti negatiivisia.*

TODISTUS. (1) Antiteesi: On olemassa $x_0 \in \,]a, b[$ siten, että peräkkäisille jonon jäsenille pätee $P_{k-1}(x_0) = 0$ ja $P_k(x_0) = 0$ jollain $k = 1, 2, ..., m-1$. Tällöin yhtälöketjun nojalla jonon seuraavalle jäsenellä pätee

$$P_{k+1}(x_0) = Q_k(x_0)P_k(x_0) - P_{k-1}(x_0) = 0.$$

Tästä seuraa edelleen, että jonon jokainen seuraava jäsen katoaa pisteessä x_0 eli myös $P_m(x_0) = 0$. Tämä on kuitenkin ristiriita, sillä aiemmin todettiin viimeisen jäsenen olevan nollasta eriävä vakio.

(2) Olkoon $P_k(x_0) = 0$. Yhtälöketjun nojalla pätee

$$P_{k-1}(x_0) = Q_{k+1}(x_0)P_k(x_0) - P_{k+1}(x_0),$$

mistä saadaan

$$P_{k-1}(x_0) = -P_{k+1}(x_0).$$

Lisäksi kohdan (1) nojalla $P_{k-1}(x_0) \neq 0$ ja $P_{k+1}(x_0) \neq 0$, joten kyseiset arvot ovat erimerkkiset.

(3) Olkoon polynomin P_k nollakohta x_0. Kohdan (1) nojalla x_0 ei voi olla polynomien P_{k-1} ja P_{k+1} nollakohta, joten kyseisessä pisteessä näiden polynomien arvot ovat joko aidosti positiivisia tai negatiivisia. Lisäksi P_{k-1} ja P_{k+1} ovat polynomifunktioina jatkuvia, joten kohdalla x_0 on olemassa pieni ympäristö, jossa kyseisten polynomien arvot ovat vastaavasti aidosti positiivisia tai negatiivisia.

□

Sturmin jonon polynomien arvojen $P_0(x), P_1(x), ..., P_{m-1}(x), P_m(x)$ etumerkit pisteessä x muodostavat *Sturmin merkkiketjun*. Määritellään funktio V siten, että funktion arvo $V(x)$ kertoo Sturmin merkkiketjussa pisteessä x tapahtuvien etumerkinvaihtojen ($+-$ tai $-+$) lukumäärän.

Esimerkki 2.21. Määritetään Esimerkin 2.19 Sturmin jonolle merkkiketjut sekä merkinvaihtojen lukumäärä $V(x)$ pisteissä $x = -2, -1, 0, 1, 2$. Laskemalla jokaisen Sturmin jonon polynomin arvo kyseisissä pisteissä saadaan selville arvojen etumerkit, jotka taulukoidaan:

x	$P_0(x)$	$P_1(x)$	$P_2(x)$	$P_3(x)$	$V(x)$
-2	+	-	+	-	3
-1	-	-	+	-	2
0	0	+	0	-	1
1	+	-	-	-	1
2	-	-	-	-	0

Taulukosta huomataan jo, että tutkittavan polynomin eräs nollakohta on $x = 0$.

Tutkitaan seuraavaksi funktion V arvon muuttumista välillä $]a, b[$. Funktion V arvo voi muuttua pisteessä $x \in]a, b[$ vain silloin, kun vähintään yksi Sturmin jonon polynomeista vaihtaa merkkiään kyseisessä pisteessä. Jatkuvina funktioina polynomit voivat vaihtaa merkkiään vain nollakohdissaan.

Olkoon siis x_0 Sturmin jonon polynomin P_k nollakohta. Valitaan $\delta > 0$ siten, että välillä $]x_0 - \delta, x_0 + \delta[$ polynomilla P_k ei ole muita nollakohtia ja lisäksi polynomit P_{k-1} ja P_{k+1} eivät vaihda merkkiään tällä välillä. Jälkimmäinen ehto onnistuu Lemman 2.20 kohdan (3) nojalla. Jaetaan tarkastelu kahteen tapaukseen:

Tapaus 1: $k > 0$. Tällöin Lemman 2.20 kohtien (2) ja (3) nojalla polynomien P_{k-1} ja P_{k+1} arvojen etumerkit eivät muutu välillä $]x_0 - \delta, x_0 + \delta[$ ja lisäksi näiden polynomien arvot ovat keskenään erimerkkiset. Arvojen $P_k(x)$ etumerkeistä riippumatta välillä $]x_0 - \delta, x_0 + \delta[$ on siis aina tasan yksi merkinvaihto eli polynomin P_k arvon muuttuminen nollakohdassaan ei aiheuta muutosta arvoon $V(x)$.

Tapaus 2: $k = 0$, jolloin tarkastellaan Sturmin jonon kahta ensimmäistä jäsentä P_0 ja P_1. Nyt Lemman 2.20 kohdan (3) nojalla välillä $]x_0 - \delta, x_0 + \delta[$ polynomin P_1 arvot ovat joko aidosti positiivisia tai negatiivisia. Koska polynomi P_1 on polynomin P_0 derivaattafunktio, täytyy siis polynomin P_0 olla kyseisellä välillä joko aidosti kasvava tai vähenevä. Jos siis P_1 on aidosti positiivista, niin P_0 on aidosti kasvava ja sen etumerkille pätee: $P_0(x) < 0$, kun $x \in]x_0 - \delta, x_0[$ ja $P_0(x) > 0$, kun $x \in]x_0, x_0 + \delta[$. Ennen nollakohtaa x_0 polynomien P_0 ja P_1 välillä on siis merkinvaihto, mutta nollakohdan jälkeen niiden merkit ovat samat. Vastaava päättely voidaan toistaa, kun P_1 on aidosti negatiivista. Tässä tapauksessa siis muuttujan x sivuuttaessa polynomin P_0 nollakohdan merkinvaihtojen lukumäärä V vähenee yhdellä.

Merkinvaihtojen lukumäärä muuttuu siis ainoastaan polynomin P_0 nollakohdissa ja erityisesti se vähenee yhden yksikön verran jokaisessa nollakohdassa. Toisaalta polynomin P_0 nollakohdat ovat juurikin tarkasteltavana olevan yhtälön juuret, jotka ovat tässä tapauksessa reaalisia muuttujan x ollessa reaaliakselilla. Saadaan siis seuraava Sturmin lause reaalijuurien lukumäärälle:

Lause 2.22. *Olkoon polynomi* $P_0(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ *siten, että sillä on vain yksinkertaisia juuria. Tällöin sen reaalijuurien lukumäärä välillä* $]a, b[$ *saadaan laskemalla*

$$V(a) - V(b).$$

Erityisesti reaalijuurien lukumäärä koko reaaliakselilla saadaan valitsemalla välin rajoiksi $-\infty$ ja $+\infty$.

Esimerkki 2.23. Määritetään polynomin $P_0(x) = -x^5 + 3x$ reaalijuurien lukumäärä. Esimerkissä 2.19 määriteltyjen Sturmin jonon polynomien avulla saadaan etumerkeille taulukko:

x	$P_0(x)$	$P_1(x)$	$P_2(x)$	$P_3(x)$	$V(x)$
$-\infty$	+	-	+	-	3
$+\infty$	-	-	-	-	0

Sturmin lauseen nojalla polynomilla P_0 on siis kolme reaalijuurta. Koska polynomin P_0 asteluku on viisi ja sen kaikki juuret ovat aiempien esimerkkien nojalla yksinkertaisia, sillä täytyy olla siis vielä kaksi imaginaarista juurta. Esimerkin 2.21 taulukosta voidaan lisäksi päätellä, missä reaalijuuret sijaitsevat sen perusteella, millä väleillä merkinvaihtojen luku pienenee yhdellä. Siispä yksi juurista on nolla, yksi on välillä $]-2, -1[$ ja kolmas on välillä $]1, 2[$.

Huomautus 2.24. Sturmin lause antaa siis myös mahdollisuuden eristää reaalijuuri tietylle välille. Juuren likiarvo voidaankin ratkaista numeerisesti tietyllä tarkkuudella pienentämällä juuren sisältävää väliä ja tarkkailemalla milloin merkinvaihtojen lukumäärä pienee. Edellisessä esimerkissä väliltä $]1, 2[$ voitaisiin valita välin keskikohta $x = 1,5$ ja laskea $V(1,5) = 0$, jolloin tiedettäisiin, että juuri on välillä $]1; 1,5[$. Väliä voitaisiin kaventaa edelleen ja saada siten juurelle tarkempi likiarvo. Tämä on eräs mahdollinen tapa reaalijuuren likiarvon numeeriseen ratkaisemiseen, jota käsitellään tarkemmin luvussa 7.

LUKU 3

Toisen asteen polynomifunktiot

Toisen asteen yhtälöt ovat kiehtoneet ihmisiä pitkään historian saatossa, sillä ne nousivat esiin pinta-aloja ja pituuksia käsittelevissä käytännön ongelmissa. Pitkään ratkaisuiksi hyväksyttiinkin vain positiivisia lukuja edustamaan pituuksia. Noin 1800 eaa. babylonialaiset tunsivatkin ratkaisukaavat tietyille täydellisille toisen asteen yhtälöille, joissa ei esiintynyt negatiivisia lukuja. Intialainen Brahmagupta esitti noin 600 jaa. lähes nykyisen kaltaisen ratkaisukaavan, joka hyväksyi myös negatiiviset luvut. Toisen asteen yhtälöitä ratkaistiinkin tehokkaasti eri puolilla maailmaa ja modernien merkintöjen kehittyessä ratkaisukaava sai ajan mittaan nykyisen muotonsa. [**10**, §1] [**2**, §13]

Johdetaan toisen asteen ratkaisukaava vielä tässäkin. Yleinen toisen asteen reaalikertoiminen yhtälö on muotoa

$$ax^2 + bx + c = 0 \tag{3.1}$$

missä $a \neq 0$ ja $a, b, c \in \mathbb{R}$. Ratkaisukaavan johtamiseksi siirretään aluksi vakiotermi yhtälön oikealle puolelle:

$$ax^2 + bx = -c.$$

Muokataan sitten yhtälön vasen puoli neliömuotoon kertomalla ensin yhtälön molemmat puolet luvulla $4a$ ja lisäämällä sitten molemmille puolille luku b^2. Päädytään yhtälöön

$$4a^2x^2 + 4abx + b^2 = b^2 - 4ac,$$

joka on yhden pitävä yhtälön

$$(2ax + b)^2 = b^2 - 4ac$$

kanssa. Ottamalla neliöjuuri saadaan

$$2ax + b = \pm\sqrt{b^2 - 4ac},$$

mistä voidaan ratkaista

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Kaavasta huomataan, että neliöjuuren sisällä olevan lausekkeen $b^2 - 4ac$ ollessa aidosti positiivista, molemmat juuret ovat reaalisia ja lisäksi erisuuria. Myös kyseisen lausekkeen ollessa nolla juuret ovat reaalisia, mutta tällöin juuret ovat yhtäsuuria. Jos taas lauseke on aidosti negatiivista, otetaan neliöjuuri negatiivisesta luvusta. Juuriksi saadaan siten aidot kompleksiluvut

$$x = -\frac{b}{2a} \pm \frac{\sqrt{-(b^2 - 4ac)}}{2a} i,$$

jotka ovat toistensa liittolukuja.

LAUSE 3.1. *Toisen asteen yhtälön* $ax^2 + bx + c = 0$ *juuret ovat*

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \in \mathbb{R}, \tag{3.2}$$

kun $b^2 - 4ac \geq 0$ *ja*

$$x = -\frac{b}{2a} \pm \frac{\sqrt{-(b^2 - 4ac)}}{2a} i, \tag{3.3}$$

kun $b^2 - 4ac < 0$.

Ensimmäistä kaavaa kutsutaan yleisesti toisen asteen yhtälön ratkaisukaavaksi ja se on hyvin merkittävä matemaattinen työkalu.

Kuten yllä todettiin, lauseke $b^2 - 4ac$ kertoo siis toisen asteen yhtälön juurien reaalisuuden ja onkin siksi tärkeä työkalu. Määritellään lauseke:

MÄÄRITELMÄ 3.2. Toisen asteen yhtälön $ax^2 + bx + c = 0$ diskriminantti on

$$D = b^2 - 4ac.$$

Toisen asteen yhtälön reaalijuurten lukumäärän määrittämiseksi ei siis tarvitse ratkaista itse yhtälöä, vaan riittää tarkastella ainoastaan yhtälön diskriminantin merkkiä. Kootaan aiemmat huomiot seuraavaan lauseeseen:

LAUSE 3.3. *Toisen asteen yhtälöllä* $ax^2 + bx + c = 0$ *on*

(1) *kaksi erisuurta reaalijuurta, jos* $D > 0$,
(2) *yksi kaksinkertainen reaalijuuri, jos* $D = 0$,
(3) *kaksi kompleksijuurta, jotka ovat toistensa liittolukuja, jos* $D < 0$.

LUKU 4

Kompleksiluvut ja niiden juurtaminen

Seuraavissa luvuissa tutustutaan kolmannen ja neljännen asteen yhtälöiden ratkaisukaavoihin. Niiden johtamisessa ja hyödyntämisessä tarvitaan kompleksilukuja ja erityisesti niiden juurtamista. Kootaankin aluksi tässä luvussa tarvittavat tiedot kompleksilukuihin liittyen.

Kompleksiluku $z = a + ib$ voidaan ajatella tason pisteenä kuvan 4.1 mukaisesti. Sille saadaan *napakoordinaattiesitys*

$$z = a + ib = r(cos\varphi + i\sin\varphi),$$

missä *moduli* $r \in \mathbb{R}$ sekä *vaihekulma* $\varphi \in [0, 2\pi[$ lasketaan kaavoilla:

$$r = \sqrt{a^2 + b^2} \tag{4.1}$$

$$\tan\varphi = \frac{b}{a}. \tag{4.2}$$

Jälkimmäisessä kaavassa samaa suhdetta b/a vastaa useampia kulmia. Vaihekulman φ valinnassa täytyykin huomioida, missä tason neljänneksessä kompleksilukua vastaava piste sijaitsee.

Esimerkki 4.1. Kompleksiluvun $z = -1 - i\sqrt{3}$ vaihekulmalle saadaan laskukaava $\tan\varphi = \sqrt{3}$. Kaavan toteuttavia kulman arvoja ovat $\varphi = \frac{\pi}{3} + 2 \cdot \pi n$ ja $\varphi = \frac{4\pi}{3} + 2 \cdot \pi n$, missä $n \in \mathbb{Z}$. Koska kompleksiluvun z reaali- ja imaginaariosien merkit ovat negatiivisia, sijaitsee luku z tason kolmannessa neljänneksessä. Siksi vain jälkimmäistä muotoa olevat vaihekulmat ovat mahdollisia. Vaihekulmaksi väliltä $[0, 2\pi[$ on siis valittava $\varphi = \frac{4\pi}{3}$. Moduliksi saadaan $r = 2$, joten napakoordinaattiesitys on

$$z = 2\left(\cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}\right).$$

Kompleksilukujen $z_j = r_j(\cos\varphi_j + i\sin\varphi_j)$, $j = 1, 2, ..., n$ tulo saadaan kaavalla

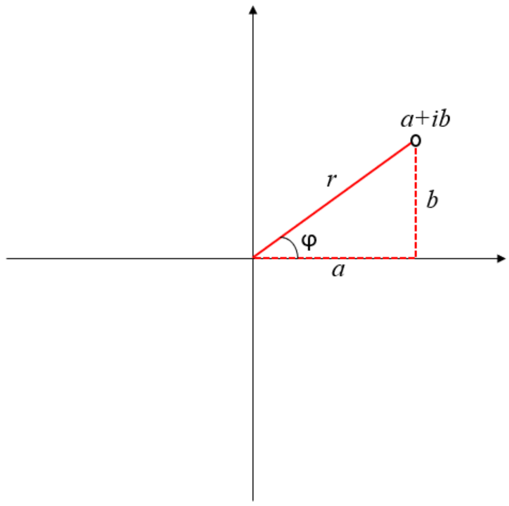
$$z_1 z_2 \cdots z_n = r_1 r_2 \cdots r_n \left[\cos\left(\sum_{j=1}^{n} \varphi_j\right) + i\sin\left(\sum_{j=1}^{n} \varphi_j\right)\right]. \tag{4.3}$$

Tästä seuraa suoraan *De Moivren kaava* kompleksiluvun $z = r(cos\varphi + i\sin\varphi)$ potensseille:

$$z^n = r^n\left(\cos(n\varphi) + i\sin(n\varphi)\right). \tag{4.4}$$

Kompleksiluvun z n-juuret ovat yhtälön $x^n = z = r(cos\varphi + i\sin\varphi)$ ratkaisut x. Jos $z \neq 0$, niin näitä juuria on n kappaletta ja ne saadaan kaavalla

$$x_\nu = \sqrt[n]{r}\left[\cos\left(\frac{\varphi}{n} + \nu\frac{2\pi}{n}\right) + i\sin\left(\frac{\varphi}{n} + \nu\frac{2\pi}{n}\right)\right], \tag{4.5}$$



KUVA 4.1. Kompleksiluku $a + ib$ esitettynä tason pisteenä sekä kompleksiluvun moduli r ja vaihekulma φ.

missä $\nu = 0, 1, ..., n-1$. Kaavan toimivuus voidaan tarkistaa korottamalla saatu x_ν potenssiin n, jolloin De Moivren kaavalla saadaan:

$$x_\nu^n = r\left[\cos(\varphi + 2\pi\nu) + i\sin(\varphi + 2\pi\nu)\right] = r(cos\varphi + i\sin\varphi) = z$$

Toisaalta voidaan merkitä

$$\varrho = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}. \tag{4.6}$$

Jos nyt jokin kompleksiluvun z n-juurista on $x_0 = \sqrt[n]{r}\left(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}\right)$, niin tällöin muut kaavan (4.5) juuret saadaan kompleksilukujen tulon kaavan (4.3) nojalla tuloina

$$x_\nu = \varrho^\nu x_0, \tag{4.7}$$

missä $\nu = 0, 1, ..., n-1$.

ESIMERKKI 4.2. Myös jokaisella nollasta eroavalla reaaliluvulla on tasan n kappaletta n-juuria kompleksilukujen joukossa. Selvitetään reaaliluvun $a = 8$ kuutiojuuret. Kyseiselle luvulle $r = 8$ ja $\varphi = 0$. Kaavan (4.5) nojalla juuriksi saadaan

$$\begin{cases} x_0 &= 2 \\ x_1 &= 2\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right) = 2\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = -1 + i\sqrt{3} \\ x_2 &= 2\left(\cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}\right) = 2\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = -1 - i\sqrt{3}. \end{cases}$$

Toisaalta juuret voidaan laskea myös kaavalla (4.7). Nyt $\varrho = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}$ ja De Moivren kaavan nojalla $\varrho^2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}$. Kun on edelleen $x_0 = 2$, saadaan samat juuret kuin aiemminkin.

LUKU 5

Kolmannen asteen polynomifunktio ja Cardanon kaavat

Tässä kappaleessa johdetaan ratkaisukaava kolmannen asteen yhtälöille sekä pohditaan juurien reaalisuuden selvittämistä. Yleinen kolmannen asteen reaalikertoiminen polynomifunktio on muotoa

$$f(z) = a_3 z^3 + a_2 z^2 + a_1 z + a_0,$$

missä $a_3 \neq 0$ ja $a_k \in \mathbb{R}$ kaikilla $k = 0, 1, 2, 3$. Sen juuret ovat ratkaisuja kolmannen asteen yhtälölle

$$a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0. \tag{5.1}$$

Voidaan olettaa, että $a_3 = 1$, sillä tähän tilanteeseen päästään aina jakamalla yhtälö puolittain luvulla a_3. Yhtälöä voidaan yksinkertaistaa poistamalla toisen asteen termi valitsemalla uudeksi muuttujaksi x siten, että

$$z = x - \frac{a_2}{3}. \tag{5.2}$$

Sijoittamalla tämä yleiseen kolmannen asteen yhtälöön (5.1) ja olettamalla $a_3 = 1$ päädytään muotoon

$$\left(x - \frac{a_2}{3}\right)^3 + a_2 \left(x - \frac{a_2}{3}\right)^2 + a_1 \left(x - \frac{a_2}{3}\right) + a_0 = 0.$$

Tämä yhtälö sieventyy muotoon

$$x^3 + px + q = 0, \tag{5.3}$$

missä vakioilla p ja q on lausekkeet

$$p = a_1 - \frac{1}{3}{a_2}^2 \qquad \text{ja} \qquad q = a_0 - \frac{1}{3}a_1 a_2 + \frac{2}{27}{a_2}^3. \tag{5.4}$$

Huomautus 5.1. Koska kertoimet $a_k \in \mathbb{R}$ kaikilla $k = 0, 1, 2, 3$, niin myös $p \in \mathbb{R}$ ja $q \in \mathbb{R}$.

Jokainen yleinen kolmannen asteen yhtälö voidaan johtaa muuttujan vaihdolla yhtälön (5.3) mukaiseen yksinkertaisempaan muotoon. Siispä jos jokainen muotoa (5.3) oleva yhtälö osataan ratkaista, osataan ratkaista myös kaikki kolmannen asteen yhtälöt. Yleisen kolmannen asteen yhtälön ratkaisut saadaan yksinkertaisemman yhtälön ratkaisuista muuttujanvaihdon yhtälön (5.2) avulla.

Esimerkki 5.2. Yhtälö $x^3 + 6x^2 + 2x - 4 = 0$ saadaan muokattua yksinkertaisempaan muotoon $x^3 - 10x + 8 = 0$ laskemalla vakioiden p ja q arvot yhtälöiden (5.4) avulla. Muuttujanvaihdon yhtälön (5.2) nojalla alkuperäisen yhtälön ratkaisut saadaan vähentämällä luku 2 yksinkertaisemman yhtälön ratkaisuista.

Siten riittää jatkossa tarkastella ainoastaan muotoa $x^3+px+q=0$ olevia yhtälöitä. Tehdään tähän yhtälöön aluksi sijoitus

$$x = u + v, \tag{5.5}$$

jolloin yhtälö saa muodon

$$(u+v)^3 + p(u+v) + q = 0.$$

Tämä saadaan muokattua muotoon

$$(u^3+v^3+q) + (3uv+p)(u+v) = 0,$$

mistä nähdään, että yhtälö toteutuu, kun u ja v toteuttavat yhtälöparin

$$\begin{cases} u^3+v^3 & = -q \\ uv & = -\frac{p}{3}. \end{cases} \tag{5.6}$$

Korottamalla jälkimmäinen yhtälö kolmanteen potenssiin, saadaan seuraava yhtälöpari:

$$\begin{cases} u^3+v^3 & = -q \\ u^3v^3 & = -\left(\frac{p}{3}\right)^3. \end{cases} \tag{5.7}$$

Tulee siis löytää luvuille u ja v lausekkeet siten, että ne toteuttavat kyseiset yhtälöparit, jolloin niiden avulla saadaan juuret x alkuperäiselle yhtälölle.

Tutkitaan toisen asteen yhtälöä

$$(y-u^3)(y-v^3) = 0,$$

jonka juuret ovat tulon nollasäännön nojalla u^3 ja v^3. Yhtälö sievenee muotoon

$$y^2 - y(u^3+v^3) + u^3v^3 = 0.$$

Sijoittamalla tähän yhtälöparin (5.7) mukaiset lausekkeet, saadaan

$$y^2 + qy - \left(\frac{p}{3}\right)^3 = 0.$$

Kyseinen yhtälö on tutkittavaa kolmannen asteen yhtälöä (5.3) vastaava *resolventtiyhtälö*. Resolventtiyhtälön juuret saadaan toisen asteen yhtälön ratkaisukaavalla:

$$y = -\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$$

Toisaalta yhtälön asettelusta tiedetään, että juuret ovat u^3 ja v^3, joten saadaan:

$$u^3 = -\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} \quad \text{ja} \quad v^3 = -\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} \tag{5.8}$$

Muuttujat u ja v ovat yhtälöissä symmetrisiä, joten merkit voitaisiin valita myös toisinpäin. Koska resolventtiyhtälö saatiin sijoittamalla yhtälöparin (5.7) mukaiset lausekkeet, vastaavasti luvuille u^3 ja v^3 saadut lausekkeet toteuttavat kyseisen yhtälöparin.

Saadut lausekkeet eivät kuitenkaan anna vielä ratkaisuja tutkittavalle kolmannen asteen yhtälölle, sillä yhtälön (5.5) nojalla juurien x ratkaisemiseksi tarvitaan luvut u ja v niiden kuutioiden sijaan. Tarvittavat luvut saadaan ottamalla kuutiojuuret saaduista lausekkeista. Tässä on kuitenkin otettava huomioon, että algebran peruslauseen nojalla jokaisella luvulla on tasan kolme kuutiojuurta kompleksilukujen joukossa. Kuutiojuurten arvot on siis valittava siten, että ne toteuttavat edelleen ensimmäisen yhtälöparin (5.6). Ylemmän yhtälön $u^3 + v^3 = -q$ toteutuminen on selvää millä tahansa kuutiojuurten arvoilla

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \qquad \text{ja} \qquad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}},$$

sillä kun nämä korotetaan kolmanteen potenssiin ja summataan saadaan tulokseksi $-q$ kuten halutaankin.

Alemman yhtälön toteutuminen vaatii kuitenkin oikeiden kuutiojuuriparien valitsemista. Valitaan aluksi luvuksi u_0 mikä tahansa seuraavista kuutiojuuren arvoista:

$$u_0 = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

Kuutiojuuren

$$v_0 = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

arvo on nyt valittava siten, että se toteuttaa alemman yhtälön eli

$$u_0 v_0 = -p/3. \tag{5.9}$$

Tällöin luvut u_0 ja v_0 toteuttavat yhtälöparin (5.6) ja luku $x = u_0 + v_0$ on siis alkuperäisen kolmannen asteen yhtälön ratkaisu.

Tämä on kuitenkin vasta yksi ratkaisu. Kahden muun ratkaisun löytämiseksi käytetään kaavaa (4.7). Nyt

$$\varrho = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$

ja

$$\varrho^2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Lisäksi huomataan, että

$$\varrho^3 = \cos\left(3\frac{2\pi}{3}\right) + i\sin\left(3\frac{2\pi}{3}\right) = 1.$$

Luvut ϱu_0 ja $\varrho^2 u_0$ ovat myös kuutiojuuren $\sqrt[3]{u^3}$ arvoja ja vastaavasti luvut ϱv_0 ja $\varrho^2 v_0$ ovat kuutiojuuren $\sqrt[3]{v^3}$ arvoja. Nyt lukuparit $(\varrho u_0, \varrho^2 v_0)$ ja $(\varrho^2 u_0, \varrho v_0)$ toteuttavat myös yhtälöparin (5.6), sillä

$$\varrho u_0 \varrho^2 v_0 = \varrho^3 u_0 v_0 = u_0 v_0 = -\frac{p}{3}$$

ja

$$\varrho^2 u_0 \varrho v_0 = \varrho^3 u_0 v_0 = u_0 v_0 = -\frac{p}{3}.$$

Yhtälön (5.3) kaikki kolme ratkaisua on siis löydetty. Tulokset voidaan koota seuraaviin *Cardanon kaavoina* tunnettuihin ratkaisukaavoihin:

Lause 5.3. *Olkoon* $\varrho = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$. *Kolmannen asteen yhtälön* $x^3 + px + q = 0$ *juuret ovat*

$$\begin{cases} x_1 = u_0 + v_0 \\ x_2 = \varrho u_0 + \varrho^2 v_0 \\ x_3 = \varrho^2 u_0 + \varrho v_0, \end{cases} \tag{5.10}$$

missä kuutiojuuren arvot

$$u_0 = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \quad \text{ja} \quad v_0 = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{5.11}$$

on valittu siten, että $u_0 v_0 = -p/3$.

Yllä näytettiin siis, että kaavan (5.10) mukaiset luvut x_1, x_2 ja x_3 toteuttavat tutkitun kolmannen asteen yhtälön. Tämä ei kuitenkaan vielä takaa, että kaavat antavat kaikki kolme *eri* juurta. Tämä voitaisiin näyttää toteamalla, ettei lauseke $(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$ häviä identtisesti kaikilla yhtälön (5.3) kertoimilla p ja q. Tätä todistusta ei kuitenkaan johdeta tässä tutkielmassa, vaan jätetään tarkistettavaksi lukijalle. Tarkastellaan sen sijaan esimerkkiä kaavojen käytöstä ja todetaan kaavojen antavan kolme eri suurta juurta:

Esimerkki 5.4. Yksinkertaisen yhtälön $x^3 + x = 0$ juuret voidaan päätellä ilman Cardanon kaavojakin, sillä

$$x^3 + x = x(x^2 + 1) = x(x + i)(x - i),$$

joten juuret ovat tulon nollasäännön nojalla 0 ja $\pm i$. Näytetään vielä varmistukseksi, että Cardanon kaavoilla saadaan sama tulos. Nyt $p = 1$ ja $q = 0$, joten

$$\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 = \frac{1}{27}.$$

Valitaan juureksi u_0 reaalinen arvo

$$u_0 = \sqrt[3]{\sqrt{\frac{1}{27}}} = \frac{1}{\sqrt{3}},$$

jolloin arvo v_0 voidaan laskea ehdosta (5.9) ja saadaan

$$v_0 = -\frac{1}{3u_0} = -\frac{1}{\sqrt{3}}.$$

Tällöin Cardanon kaavoilla ensimmäiseksi juureksi saadaan $x_1 = u_0 + v_0 = 0$. Lisäksi saadaan

$$x_2 = \varrho u_0 + \varrho^2 v_0 = \frac{1}{\sqrt{3}} \left(\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2} \right) - \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2} \right) \right) = i$$

ja

$$x_3 = \varrho^2 u_0 + \varrho v_0 = \frac{1}{\sqrt{3}} \left(\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2} \right) - \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2} \right) \right) = -i,$$

eli Cardanon kaavat antavat halutut juuret kuten pitääkin.

Kappaleessa 4 todettiin, että toisen asteen yhtälön reaalijuurien lukumäärän selvittämiseksi ei tarvitse ratkaista juuria vaan pelkästään yhtälön diskriminantin tarkastelu riittää. Määritellään vastaavasti diskriminantti kolmannen asteen yhtälölle (5.3) Cardanon kaavoissa esiintyvien neliöjuurien (5.11) sisällä olevien lausekkeiden perusteella:

MÄÄRITELMÄ 5.5. Kolmannen asteen yhtälön $x^3 + px + q = 0$ diskriminantti D on

$$D = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3.$$

Seuraava lause kertoo, kuinka diskriminantin merkin avulla voidaan määrittää yhtälön (5.3) reaalijuurien lukumäärä:

LAUSE 5.6. *Yhtälön $x^3 + px + q = 0$ reaalijuurten lukumäärä riippuu diskriminantista D seuraavasti:*

(1) *Jos $D > 0$, yhtälöllä on yksi reaalijuuri ja kaksi kompleksijuurta, jotka ovat toistensa liittolukuja.*
(2) *Jos $D = 0$, yhtälöllä on kolme reaalijuurta, joista jokin on moninkertainen.*
(3) *Jos $D < 0$, yhtälöllä on kolme erisuurta reaalijuurta.*

TODISTUS. (1) Olkoon $D > 0$, jolloin $\sqrt{D} \in \mathbb{R}$. Siten kuutiojuurilla

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{D}} \qquad \text{ja} \qquad v = \sqrt[3]{-\frac{q}{2} - \sqrt{D}}$$

on olemassa reaaliset arvot. Olkoon siis u_0 yllä olevan kuutiojuuren u reaalinen arvo. Nyt ehdosta (5.9) saadaan, että myös $v_0 \in \mathbb{R}$, koska

$$v_0 = -\frac{p}{3u_0}.$$

Siten myös Cardanon kaavojen ensimmäiselle juurelle $x_1 = u_0 + v_0 \in \mathbb{R}$. Kahdelle muulle juurelle saadaan:

$$x_2 = \varrho u_0 + \varrho^2 v_0 = -\frac{1}{2}(u_0 + v_0) + i\frac{\sqrt{3}}{2}(u_0 - v_0)$$

ja

$$x_3 = \varrho^2 u_0 + \varrho v_0 = -\frac{1}{2}(u_0 + v_0) - i\frac{\sqrt{3}}{2}(u_0 - v_0).$$

Koska $D > 0$, niin $u_0 \neq v_0$, joten juuret x_2 ja x_3 ovat kompleksisia ja lisäksi toistensa liittolukuja.

(2) Olkoon $D = 0$. Tällöin diskriminantin lausekkeen perusteella joko $p < 0$ tai $p = q = 0$. Nyt

$$u = v = -\sqrt[3]{\frac{q}{2}},$$

joten luvuiksi u_0 ja v_0 voidaan valita kuutiojuuren reaalinen arvo kuten edellisessä tapauksessa eli

$$u_0 = v_0 = -\sqrt[3]{\frac{q}{2}} \in \mathbb{R}.$$

Siispä ensimmäiselle juurelle

$$x_1 = u_0 + v_0 = -2\sqrt[3]{\frac{q}{2}} \in \mathbb{R}.$$

Muut kaksi juurta ovat symmetrisiä, koska $u_0 = v_0$, ja juuriksi saadaan

$$x_2 = x_3 = \varrho u_0 + \varrho^2 u_0 = (\varrho + \varrho^2)u_0 = -u_0 = \sqrt[3]{\frac{q}{2}} \in \mathbb{R}.$$

Kaikki juuret ovat siis reaalisia ja lisäksi yksi on kaksoisjuuri. Lisäksi jos $q = 0$, niin $x_1 = x_2 = x_3 = 0$ eli yhtälöllä on vain yksi kolminkertainen juuri.

(3) Olkoon $D < 0$, jolloin diskriminantin lausekkeen nojalla on oltava $p < 0$. Koska nyt $\sqrt{D} \in \mathbb{C}$, voidaan merkitä $\sqrt{D} = i\sqrt{-D}$. Kuutioille (5.8) saadaan siten lausekkeet

$$u^3 = -\frac{q}{2} + i\sqrt{-D} \qquad \text{ja} \qquad v^3 = -\frac{q}{2} - i\sqrt{-D}. \tag{5.12}$$

Määritetään näiden kompleksilausekkeiden napakoordinaattiesitykset. Yhtälöllä (4.1) ja diskriminantin lausekkeella saadaan

$$r = \sqrt{\left(-\frac{q}{2}\right)^2 + \sqrt{-D}^2} = \sqrt{\left(\frac{q}{2}\right)^2 - D} = \left(\sqrt{-\frac{p}{3}}\right)^3. \tag{5.13}$$

Nyt koska $p < 0$, niin neliöjuuren sisällä on positiivinen luku, joten voidaan valita kuutiojuuren arvo niin, että $r \in \mathbb{R}$ ja lisäksi tällöin $r > 0$. Vaihekulmalle φ saadaan lauseke

$$\cos\varphi = -\frac{q}{2r}. \tag{5.14}$$

Näillä parametreilla kuutioiden lausekkeiksi saadaan

$$u^3 = r(\cos\varphi + i\sin\varphi) \qquad \text{ja} \qquad v^3 = r(\cos\varphi - i\sin\varphi). \tag{5.15}$$

Yhtälöitä (5.12) ja (5.15) vertaamalla huomataan, että koska $\sqrt{-D} > 0$, täytyy olla myös $\sin\varphi > 0$, jotta imaginääriosien merkit ovat samat molemmissa muodoissa. Siten φ on valittava ensimmäisestä tai toisesta neljänneksestä riippuen vakion q merkistä yhtälössä (5.14).

Nyt kuutiojuuret voidaan ottaa lausekkeista (5.15) ja saadaan

$$\begin{aligned} u_\nu &= \sqrt{-\frac{p}{3}}\left[\cos\left(\frac{\varphi}{3} + \nu\frac{2\pi}{3}\right) + i\sin\left(\frac{\varphi}{3} + \nu\frac{2\pi}{3}\right)\right] \\ v_\nu &= \sqrt{-\frac{p}{3}}\left[\cos\left(\frac{\varphi}{3} + \nu\frac{2\pi}{3}\right) - i\sin\left(\frac{\varphi}{3} + \nu\frac{2\pi}{3}\right)\right], \end{aligned}$$

missä $\nu = 0, 1, 2$. Merkitään $\alpha = \frac{\varphi}{3} + \nu\frac{2\pi}{3}$ ja tutkitaan samaa vakion ν arvoa vastaavien juurien u_ν ja v_ν tuloa:

$$\begin{aligned} u_\nu v_\nu &= -\frac{p}{3}\left(\cos\alpha + i\sin\alpha\right)\left(\cos\alpha - i\sin\alpha\right) \\ &= -\frac{p}{3}\left(\cos^2\alpha + \sin^2\alpha\right) \\ &= -\frac{p}{3} \end{aligned}$$

Samaa vakion ν arvoa vastaavat juuret u_ν ja v_ν toteuttavat siis ehdon (5.9), joten niiden summa on alkuperäisen yhtälön juuri. Juuriksi saadaan siten

$$\begin{cases} x_1 = 2\sqrt{-\dfrac{p}{3}}\cos\dfrac{\varphi}{3} \\ x_2 = 2\sqrt{-\dfrac{p}{3}}\cos\left(\dfrac{\varphi}{3}+\dfrac{2\pi}{3}\right) \\ x_3 = 2\sqrt{-\dfrac{p}{3}}\cos\left(\dfrac{\varphi}{3}+\dfrac{4\pi}{3}\right). \end{cases} \tag{5.16}$$

Kaavoista on selvää, että kaikki juuret ovat reaalisia ja lisäksi erisuuria.

□

HUOMAUTUS 5.7. Tapauksissa $D > 0$ ja $D = 0$ yhtälön (5.3) juuret voitiin määrittää helposti suoraan Cardanon kaavoilla, koska kuutiojuurille (5.11) saatiin yhdet reaaliset arvot. Tapauksessa $D < 0$ näin ei kuitenkaan ollut, joten yhtälön juurien määrittämisessä täytyi juurtaa aitoja kompleksilukuja. Tämä tuotti ongelmia Cardanon kaavojen keksijöille sekä niiden ensimmäisille käyttäjille, koska kompleksiluvuilla ei osattu vielä laskea sujuvasti. Heitä hämmästytti myös se, että etukäteen helpoimmaksi luultu kolmen reaalijuuren tapaus osoittautuikin vaikeimmaksi ratkaistavaksi. Huomattiinkin, ettei Cardanon kaavoja tässä tapauksessa voitu ratkaista pelkästään algebrallisin keinoin, vaan käyttöön oli otettava trigonometriset funktiot. Lisäksi alettiin ymmärtää kompleksilukujen merkityksellisyys, koska niitä voitiin tarvita myös täysin reaalisten tulosten saamiseksi.

Kyseinen tapaus tunnetaan nimellä *casus irreducibilis* ("jakautumaton tapaus") ja sen vaikutus matematiikan kehitykseen on ollut merkittävä, sillä se innosti ottamaan kompleksiluvut käyttöön ja kehittämään uusia teorioita. Kompleksiluvuilla laskemisen yleistyttyä onnistuttiinkin johtamaan kaavat (5.16), joilla pystytään siis kyseisessä tapauksessa laskemaan juurien arvot suoraan ilman kompleksilukujen juurtamista.

ESIMERKKI 5.8. Tutkitaan polynomia $P(x) = x^3 - 12x - 8$, jonka diskriminantiksi saadaan $D = -48 < 0$. Polynomilla P on siis kolme reaalijuurta. Juuret voidaan siis ratkaista kaavojen (5.16) avulla. Selvitetään ensin kulma φ. Nyt yhtälöstä (5.13) saadaan

$$r = \left(\sqrt{-\frac{-12}{3}}\right)^3 = 8,$$

jolloin yhtälön (5.14) nojalla $\cos\varphi = \frac{1}{2}$. Koska kulman φ tulee olla ensimmäisessä tai toisessa neljänneksessä, voidaan yhtälön toteuttavaksi kulmaksi valita $\varphi = \frac{\pi}{3}$. Alkuperäisen yhtälön juuriksi saadaan siten

$$\begin{cases} x_1 = 4\cos\dfrac{\pi}{9} \\ x_2 = 4\cos\dfrac{7\pi}{9} \\ x_3 = 4\cos\dfrac{13\pi}{9}. \end{cases}$$

LUKU 6

Neljännen asteen yhtälö

Kuten toisen ja kolmannen asteen yhtälöille, myös neljännen asteen yhtälölle on olemassa ratkaisukaava, jolla juurien tarkat arvot voidaan ratkaista algebrallisesti. Tässä kappaleessa esitetään kyseinen ratkaisukaava, mutta kaavan johtamista ei perustella.

Yleinen neljännen asteen reaalikertoiminen polynomifunktio on muotoa

$$f(z) = a_4z^4 + a_3z^3 + a_2z^2 + a_1z + a_0,$$

missä $a_4 \neq 0$ ja $a_k \in \mathbb{R}$ kaikilla $k = 0, 1, 2, 3, 4$. Sen juuret ovat ratkaisuja neljännen asteen yhtälölle

$$a_4z^4 + a_3z^3 + a_2z^2 + a_1z + a_0 = 0. \tag{6.1}$$

Kuten kolmannen asteen yhtälön tapauksessa, voidaan olettaa, että $a_4 = 1$. Vastaavasti yksinkertaistetaan yhtälöä poistamalla kolmannen asteen jäsen tekemällä sijoitus

$$z = x - \frac{1}{4}a_3.$$

Yhtälö (6.1) sieventyy tällöin muotoon

$$x^4 + px^2 + qx + r = 0, \tag{6.2}$$

missä

$$\begin{cases} p = a_2 - \frac{3}{8}a_3^2 \\ q = \frac{5}{32}a_3^3 - \frac{1}{32}a_3^2 - \frac{1}{2}a_3a_2 + a_1 \\ r = \frac{1}{4}\left(-\frac{3}{4}a_3^4 + \frac{1}{4}a_3^2a_2 - a_3a_1\right). \end{cases}$$

Koska $a_k \in \mathbb{R}$ kaikilla $k = 0, 1, 2, 3$, myös kertoimet p, q ja r ovat reaalisia.

Jos tätä muotoa oleva neljännen asteen yhtälö osataan ratkaista, osataan ratkaista siis kaikki neljännen asteen yhtälöt. Yhtälön ratkaisukaava voidaan johtaa samankaltaisin periaattein kuin kolmannen asteen yhtälön tapauksessakin tekemällä sijoitus $x = u + v + w$ ja muodostamalla resolventtiyhtälö. Kaava ei kuitenkaan tässä johdeta, vaan siirrytään suoraan tuloksiin. Johtamiseen voi perehtyä esimerkiksi Väisälän teoksessa [**12**, §82]. Määritellään ensiksi yhtälön (6.2) resolventtiyhtälö:

Määritelmä 6.1. Yhtälön $x^4 + px^2 + qx + r = 0$ resolventtiyhtälö on kolmannen asteen yhtälö

$$y^3 + \frac{p}{2}y^2 + \frac{p^2 - 4r}{16}y - \frac{q^2}{64} = 0.$$

Esimerkki 6.2. Yhtälössä $x^4 + 6x^2 + 48x + 33 = 0$ on $p = 6$, $q = 48$ ja $r = 33$. Siten resolventtiyhtälöksi saadaan

$$y^3 + 3y^2 - 6y - 36 = 0.$$

Esitetään sitten ratkaisukaava:

LAUSE 6.3. *Yhtälön* $x^4 + px^2 + qx + r = 0$ *juuret ovat*

$$\begin{cases} x_1 = & u_0 + v_0 + w_0 \\ x_2 = & u_0 - v_0 - w_0 \\ x_3 = - & u_0 + v_0 - w_0 \\ x_4 = - & u_0 - v_0 + w_0, \end{cases} \tag{6.3}$$

missä u_0, v_0 *ja* w_0 *ovat resolventtiyhtälön juurien* y_1, y_2 *ja* y_3 *neliöjuuria, joiden merkit on valittu siten, että*

$$u_0 v_0 w_0 = -\frac{q}{8}. \tag{6.4}$$

Näytetään vielä, että kaavat (6.3) todella antavat kaikki tutkittavan yhtälön juuret eli että juuret voivat olla keskenään erisuuria. Tutkitaan tuloa

$$(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4),$$

joka on nolla jos ja vain jos jotkin juurista x_k, $k \in \{1, 2, 3, 4\}$ ovat yhtäsuuria. Tämä tulo saadaan kaavojen (6.3) perusteella muotoon

$$(2v_0 + 2w_0)(2u_0 + 2w_0)(2u_0 + 2v_0)(2u_0 - 2w_0)(2u_0 - 2v_0)(2v_0 - 2w_0),$$

josta saadaan edelleen muokattua muoto

$$2^6 \left(u_0^2 - v_0^2\right) \left(u_0^2 - w_0^2\right) \left(v_0^2 - w_0^2\right).$$

Luvut u_0, v_0 ja w_0 on valittu siten, että ne vastaavat resolventtiyhtälön juurien y_1, y_2 ja y_3 neliöjuuria. Olkoon valittu niin, että $u_0^2 = y_1$, $v_0^2 = y_2$ ja $w_0^2 = y_3$. Tällöin edellinen tulo saa muodon

$$2^6 (y_1 - y_2)(y_1 - y_3)(y_2 - y_3)$$

Tämä on nolla, jos ja vain jos jotkin resolventtiyhtälön juurista ovat yhtäsuuria eli resolventtiyhtälöllä on moninkertainen juuri. Voidaan kuitenkin näyttää, ettei resolventtiyhtälölle ole aina moninkertaista juurta. Esimerkiksi kun yhtälössä (6.2) on $p = r = 0$ ja $q \neq 0$, niin resolventtiyhtälö saa muodon

$$y^3 - \frac{q^2}{64} = 0,$$

jolla on kolme erisuurta juurta. Siispä kaavat (6.3) voivat esittää neljää erisuurta juurta ja antavat kaikki ratkaisut yhtälölle (6.2).

Lisäksi seuraa tulos:

LAUSE 6.4. *Yhtälöllä* $x^4 + px^2 + qx + r = 0$ *on moninkertainen juuri jos ja vain jos sen resolventtiyhtälöllä on moninkertainen juuri eli resolventtiyhtälön diskriminantti on nolla.*

Havainnollistetaan sitten neljännen asteen yhtälön ratkaisemista todetuilla kaavoilla:

Esimerkki 6.5. Tutkitaan neljännen asteen yhtälöä $x^4 + 6x^2 + 48x + 33 = 0$, jonka resolventtiyhtälöksi saatiin edellisessä esimerkissä $y^3 + 3y^2 - 6y - 36 = 0$. Jotta resolventtiyhtälön juuret voidaan ratkaista Cardanon kaavoilla, on se ensin muutettava muotoon $z^3 + pz + q = 0$. Kertoimien arvoiksi saadaan yhtälöillä (5.4) $p = -9$ ja $q = -28$.

Ratkaistaan saatu yhtälö $z^3 - 9z - 28 = 0$ Cardanon kaavoilla. Saadaan $u_0 = \sqrt[3]{27}$ ja $v_0 = \sqrt[3]{1}$. Kun valitaan kuutiojuurten reaaliset arvot $u_0 = 3$ ja $v_0 = 1$, niin ehto $u_0 v_0 = -\frac{p}{3}$ toteutuu. Yhtälön ratkaisut ovat siten

$$\begin{cases} z_1 = 4 \\ z_2 = -2 + i\sqrt{3} \\ z_3 = -2 - i\sqrt{3}. \end{cases}$$

Näistä saadaan resolventtiyhtälön $y^3 + 3y^2 - 6y - 36 = 0$ ratkaisut muuttujanvaihdon yhtälön (5.2) nojalla vähentämällä luku 1, joten saadaan

$$\begin{cases} y_1 = 3 \\ y_2 = -3 + i\sqrt{3} \\ y_3 = -3 - i\sqrt{3}. \end{cases}$$

Lauseen 6.3 perusteella näistä ratkaisuista on otettava neliöjuuret alkuperäisen yhtälön ratkaisemiseksi. Neliöjuuren ottamiseksi määritetään ensin juurten y_2 ja y_3 napakoordinaattiesitykset. Molempien juurien moduliksi saadaan

$$r = \sqrt{(-3)^2 + (-\sqrt{3})^2} = \sqrt{12}.$$

Juuren y_2 vaihekulma

$$\varphi_2 = \arctan\left(-\frac{\sqrt{3}}{3}\right) = \arctan\left(-\frac{1}{\sqrt{3}}\right)$$

on valittava toisesta neljänneksestä, joten saadaan $\varphi_2 = \frac{5\pi}{6}$. Juuren y_3 vaihekulma

$$\varphi_3 = \arctan\left(\frac{-\sqrt{3}}{-3}\right) = \arctan\left(\frac{1}{\sqrt{3}}\right)$$

taas on valittava kolmannesta neljänneksestä, joten saadaan $\varphi_3 = \frac{7\pi}{6}$. Napakoordinaattiesityksiksi saadaan siis

$$y_2 = \sqrt{12}\left(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}\right) \qquad \text{ja} \qquad y_3 = \sqrt{12}\left(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\right).$$

Resolventtiyhtälön juurien neliöjuuriksi saadaan siten

$$\begin{cases} u = \pm\sqrt{3} \\ v = \sqrt[4]{12}\left[\cos\left(\frac{5\pi}{12} + \nu\pi\right) + i\sin\left(\frac{5\pi}{12} + \nu\pi\right)\right] \\ w = \sqrt[4]{12}\left[\cos\left(\frac{7\pi}{12} + \nu\pi\right) + i\sin\left(\frac{7\pi}{12} + \nu\pi\right)\right], \end{cases}$$

missä $\nu \in \{0,1\}$. Kun valitaan

$$\begin{cases} u_0 = \sqrt{3} \\ v_0 = \sqrt[4]{12}\left(\cos\dfrac{5\pi}{12} + i\sin\dfrac{5\pi}{12}\right) \\ w_0 = \sqrt[4]{12}\left(\cos\dfrac{7\pi}{12} + i\sin\dfrac{7\pi}{12}\right), \end{cases}$$

niin saadaan

$$u_0v_0w_0 = \sqrt{3}\sqrt{12}\left[\cos\left(\frac{5\pi}{12}+\frac{7\pi}{12}\right) + i\sin\left(\frac{5\pi}{12}+\frac{7\pi}{12}\right)\right] = -6.$$

Koska myös $-\frac{48}{8} = -6$, ehto (6.4) toteutuu.

Alkuperäisen neljännen asteen yhtälön ratkaisuiksi saadaan siten kaavoilla (6.3):

$$\begin{cases} x_1 = \sqrt{3} + \sqrt[4]{12}\dfrac{\sqrt{6}+\sqrt{2}}{2}i \\ x_2 = \sqrt{3} - \sqrt[4]{12}\dfrac{\sqrt{6}+\sqrt{2}}{2}i \\ x_3 = -\sqrt{3} + \sqrt[4]{12}\dfrac{\sqrt{6}-\sqrt{2}}{2} \\ x_4 = -\sqrt{3} - \sqrt[4]{12}\dfrac{\sqrt{6}-\sqrt{2}}{2} \end{cases}$$

Yhtälöllä on siis kaksi reaalijuurta ja kaksi aitoa kompleksijuurta.

Neljännen asteen yhtälön ratkaisemiseksi on siis ensin ratkaistava kolmannen asteen resolventtiyhtälö Cardanon kaavoilla. Kuten esimerkistä huomataan, tämäkin on jo monivaiheinen ja melko työläs tehtävä. Siksi neljännen asteen yhtälön ratkaisukaavaa käytetään harvoin käytännön ongelmissa, joissa pelkkä likiarvo tietyllä tarkkuudella on riittävä tulos. Ratkaisukaavan olemassaolon löytäminen olikin merkittävämpää puhtaan matematiikan kannalta kuin käytännön hyötynsä puolesta.

Mainitaan vielä, että muotoa

$$a_4x^4 + a_2x^2 + a_0 = 0$$

olevien neljännen asteen yhtälöiden algebrallinen ratkaiseminen onnistuu helposti, sillä tehtävä voidaan palauttaa toisen asteen yhtälön ratkaisemiseen muuttujanvaihdolla $x^2 = u$:

Esimerkki 6.6. Tutkitaan neljännen asteen yhtälöä $x^4 + 2x^2 - 8 = 0$. Tehdään sijoitus $x^2 = u$, jolloin saadaan toisen asteen yhtälö $u^2+2u-8 = 0$. Tämän ratkaisuiksi saadaan toiseen asteen ratkaisukaavalla $u = 2$ ja $u = -4$. Alkuperäisen neljännen asteen yhtälön ratkaisuiksi saadaan siten muuttujanvaihdon yhtälön $x^2 = u$ nojalla luvut $x = \pm\sqrt{2}$ ja $x = \pm 2i$.

LUKU 7

Korkeamman asteen yhtälöt ja juurien likiarvojen määrittäminen

Toisen, kolmannen ja neljännen asteen yhtälöille löydettiin aiemmissa kappaleissa algebralliset ratkaisukaavat, joilla juurille saadaan määritettyä tarkat arvot. Seuraavaksi olisi loogista pyrkiä löytämään vastaavanlainen kaava yleiselle viidennen asteen polynomiyhtälölle ja edelleen korkeamman asteen yhtälöille. Tätä matemaatikot yrittivätkin satojen vuosien ajan, mutta yritykset päätyivät lähes aina siihen, että ratkaistavaksi tuli alkuperäistä yhtälöä asteluvultaan korkeampi yhtälö, jota ei myöskään osattu ratkaista. Tämä johti pohdintoihin, onko korkeamman asteisille yhtälöille edes mahdollista löytää algebrallista ratkaisukaavaa. Vuosien saatossa pystyttiinkin todistamaan, että tietynlaiset korkeamman asteen yhtälöt olivat algebrallisesti ratkeamattomia [**10**, §5.6]. Lopulta vuonna 1826 norjalainen Niels Henrik Abel todisti seuraavan lauseen:

Lause 7.1. *Viidennen ja sitä korkeamman asteen yleiset polynomiyhtälöt ovat algebrallisesti ratkeamattomia.*

Todistusta ei esitetä tässä tutkielmassa. Tämä lause lopetti siis yleisten korkeamman asteisten yhtälöiden ratkaisukaavojen etsimisen. On kuitenkin huomattava, että lause käsittelee *yleisiä* muotoa

$$a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_0 = 0,$$

olevia n:nnen asteen yhtälöitä, missä kertoimet $a_0, a_1, ..., a_n$ ovat riippumattomia ja mielivaltaisia. Toki siis jotkin korkeamman asteen yhtälöt voidaan ratkaista algebrallisesti, esimerkiksi yksinkertaiset yhtälöt $x^n + a_0 = 0$. Niin kutsuttu *Galois´n teoria* määrittelee, millaista muotoa olevat korkeamman asteen yhtälöt ovat algebrallisesti ratkaistavissa. Tätä teoriaa ei kuitenkaan käsitellä tässä tarkemmin, vaan tyydytään vain toteamaan yleisten yhtälöiden ratkeamattomuus.

Kuten Huomautuksessa 2.7 todettiin, korkeamman asteen yhtälöiden ratkaisemista voidaan usein helpottaa löytämällä yhtälölle aluksi yksi juuri, jolloin riittää enää löytää juuret alemman asteiselle yhtälölle. Tätä yhtä juurta voi lähteä etsimään esimerkiksi seuraavaksi esiteltävällä *rationaalijuuritestillä*, joka kertoo millaista muotoa yhtälön mahdolliset rationaalijuuret voivat olla.

Lause 7.2. *Olkoon* $P(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ *kokonaislukukertoiminen polynomi, jolle* $a_n \neq 0$ *ja* $a_0 \neq 0$*. Olkoon* $\frac{p}{q}$ *polynomin* P *rationaalijuuri, jolle* $syt(p, q) = 1$*. Tällöin vakiotermi* a_0 *on jaollinen osoittajalla* p *ja kerroin* a_n *on jaollinen nimittäjällä* q*.*

TODISTUS. Koska $\frac{p}{q}$ on polynomin P juuri, se toteuttaa yhtälön

$$P\left(\frac{p}{q}\right) = a_n\left(\frac{p}{q}\right)^n + a_{n-1}\left(\frac{p}{q}\right)^{n-1} + \cdots + a_1\left(\frac{p}{q}\right) + a_0 = 0. \tag{7.1}$$

Siirtämällä vakiotermi a_0 yhtälön oikealle puolella ja kertomalla molemmat puolet luvulla q^n saadaan yhtälö muotoon

$$p\left(a_n p^{n-1} + a_{n-1}p^{n-2}q + \cdots + a_1 q^{n-1}\right) = -a_0 q^n,$$

missä otettiin lisäksi vasemmalla puolella yhteiseksi tekijäksi luku p. Koska luvut p ja q sekä kertoimet a_i, $i = 1, ..., n$ ovat kokonaislukuja, vasemmalla puolella sulkujen sisällä oleva lauseke on myös kokonaisluku. Siispä yhtälön oikea puoli $-a_0 q^n$ on jaollinen luvulla p. Koska $\mathrm{syt}(p, q) = 1$, luku p ei jaa lukua q^n. Siispä vakiotermin a_0 täytyy olla jaollinen luvulla p.

Toisaalta siirtämällä yhtälössä (7.1) korkeimman asteen termi oikealle puolelle ja kertomalla luvulla q^n saadaan yhtälö

$$q\left(a_{n-1}p^{n-1} + a_{n-2}p^{n-2}q + \cdots + a_1 p q^{n-2} + a_0 q^{n-1}\right) = -a_n p^n.$$

Tästä voidaan vastaavasti päätellä, että luvun q täytyy jakaa kerroin a_n. □

HUOMAUTUS 7.3. Rationaalijuuritestissä oletettiin tutkittavan polynomin olevan kokonaislukukertoiminen. Testi toimii kuitenkin myös rationaalilukukertoimiselle polynomille, sillä siitä saadaan kokonaislukukertoiminen esimerkiksi kertomalla kaikkien kertoimien nimittäjillä. Koska polynomin kertominen vakiolla ei muuta sen juuria, saadaan testillä alkuperäisen polynomin rationaalijuuret.

Lisäksi oletettiin vakiotermin a_0 olevan nollasta eriävä. Jos näin ei kuitenkaan ole, voidaan muuttuja x ottaa polynomin lausekkeessa yhteiseksi tekijäksi:

$$P(x) = x(a_n x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_1).$$

Tällöin voidaan käyttää rationaalijuuritestiä sulkujen sisällä olevaan $n-1$ -asteiseen polynomiin.

Rationaalijuuritesti kertoo siis millaista muotoa kaikki mahdolliset polynomin P rationaalijuuret ovat. On huomattava, että kaikki tätä muotoa olevat rationaaliluvut eivät suinkaan ole polynomin juuria, vaan nämä ovat ainoat mahdolliset juuret. Jokainen kyseistä muotoa oleva luku onkin testattava erikseen, jotta saadaan selville kaikki rationaalijuuret.

ESIMERKKI 7.4. Tutkitaan polynomia $P(x) = 3x^3 - 5x^2 + 5x - 2$. Vakiotermi -2 on jaollinen luvuilla $p \in \{\pm 1, \pm 2\}$ ja kerroin 3 on jaollinen luvuilla $q \in \{\pm 1, \pm 3\}$. Rationaalijuuritestin nojalla mahdolliset rationaalijuuret ovat siten

$$\frac{p}{q} \in \left\{\pm\frac{1}{3}, \pm\frac{2}{3}, \pm 1, \pm 2\right\}.$$

Näistä vain $P(\frac{2}{3}) = 0$, joten $\frac{2}{3}$ on polynomin P ainoa rationaalijuuri. Muut juuret ovat siten irrationaalisia tai aitoja kompleksilukuja.

Kaikille yhtälöille ei kuitenkaan löydy rationaalijuuria, joilla yhtälöiden ratkaiseminen voitaisiin palauttaa alemman asteisen yhtälön ratkaisemiseen. Koska yleisiin korkeamman asteisiin yhtälöihin voidaan törmätä käytännön ongelmissa, niille onkin täytynyt kehittää muunlaisia ratkaisutapoja, joilla saadaan selville reaalijuurten likiarvot riittävällä tarkkuudella. Näitä menetelmiä käytetään usein myös kolmannen ja neljännen asteen yhtälöiden juurien etsinnässä, sillä kuten aiemmin todettiin niille löydetyt ratkaisukaavat ovat melko työläitä käyttää. Seuraavaksi esitetäänkin muutamia menetelmiä reaalijuurien likiarvojen määrittämiseksi.

7.1. Merkin muuttumiseen perustuva menetelmä

Tutkitaan yleistä polynomifunktiota

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

Jos löydetään $a, b \in \mathbb{R}$, $a < b$ siten, että $P(a)$ ja $P(b)$ ovat erimerkkiset, tiedetään Bolzanon lauseen nojalla, että jatkuvalla polynomifunktiolla P on ainakin yksi nollakohta välillä $]a, b[$. Nollakohtia voi kyseisellä välillä olla myös useampia, mutta oletetaan tässä, että niitä on vain yksi, mikä voidaan tarkistaa esimerkiksi tutkimalla funktion kuvaajaa.

Valitaan sitten $c \in]a, b[$. Nyt jos $P(a)$ ja $P(c)$ ovat erimerkkiset, funktiolla on nollakohta välillä $]a, c[$. Vastaavasti nollakohta on välillä $]c, b[$, jos $P(b)$ ja $P(c)$ ovat erimerkkiset. Nollakohdan sisältävä väli siis pienenee, jolloin juurelle saadaan tarkempi arvio. Tätä voidaan jatkaa ottamalla edelleen uusi luku saadulta väliltä ja tutkimalla funktion P merkkiä kyseisessä pisteessä ja välin päätepisteissä. Näin väliä saadaan pienennettyä aina niin kauan, että saadaan tarvittavan tarkka likiarvo juurelle x.

Usein merkin muuttumiseen perustuvassa menetelmässä valitaan aina tutkittavan välin keskikohta. Tällöin puhutaan *puolitusmenetelmästä*.

Tämä menetelmä on melko työläs, erityisesti jos halutaan likiarvo hyvin suurella tarkkuudella. Ensimmäinen väli kannattaakin valita mahdollisimman pieneksi esimerkiksi polynomifunktion kuvaajaa tarkastelemalla, jotta juuren likiarvolle saadaan haluttu tarkkuus mahdollisimman nopeasti. Tällä menetelmällä ei myöskään löydetä mahdollisia moninkertaisia juuria, joissa polynomifunktio ei vaihda merkkiään. Toisaalta menetelmän hyviä puolia ovat sen yksinkertaisuus ja se, että saatujen arvojen jono konvergoi aina välillä olevaan juureen. Siksi sitä käytetäänkin usein aluksi juuren karkean likiarvon määrittämiseen, jonka jälkeen arvoa tarkennetaan muilla menetelmillä. [**1**, §2.1]

7.2. Newtonin menetelmä

Newtonin tai Newton-Raphsonin menetelmä on yksi eniten käytetyistä ja tehokkaimista numeerisista keinoista funktioiden juurien likiarvojen ratkaisemiseen. Johdetaan kyseinen menetelmä tässä hyödyntämällä Taylorin sarjaa polynomille P [**1**, §2.3].

Polynomifunktiona P on jatkuva ja kaksi kertaa jatkuvasti differentioituva. Olkoon nyt $p \in [a, b]$ polynomin P nollakohta ja olkoon $p_0 \in [a, b]$ nollakohdan p approksimaatio siten, että $|p - p_0|$ on "pieni". Pohditaan myöhemmin lisää, mikä on riittävän pieni ero approksimaation ja oikean nollakohdan välillä, mutta tyydytään tässä kohtaa siihen, että riittävän tarkka approksimaatio voidaan löytää esimerkiksi merkin muuttumiseen perustuvalla menetelmällä.

Polynomin P toiseksi Taylorin polynomiksi pisteen p_0 ympäristössä saadaan [**11**, §7.12]

$$P(x) = P(p_0) + P'(p_0)(x - p_0) + \frac{P''(p_0)}{2}(x - p_0)^2.$$

Koska p on polynomin P nollakohta, saadaan Taylorin polynomille pisteessä p:

$$0 = P(p_0) + P'(p_0)(p - p_0) + \frac{P''(p_0)}{2}(p - p_0)^2$$

Toisaalta oletettiin, että $|p - p_0|$ on pieni, joten lausekkeen $(p - p_0)^2$ sisältävän termin voidaan ajatella olevan merkityksettömän pieni. Siispä saadaan:

$$0 \approx P(p_0) + P'(p_0)(p - p_0)$$

Tästä yhtälöstä voidaan ratkaista

$$p \approx p_0 - \frac{P(p_0)}{P'(p_0)},$$

mikä antaa uuden approksimaation polynomin juurelle. Kun tiedetään siis tarpeeksi tarkka juuren alkuapproksimaatio p_0, voidaan seuraava approksimaatio laskea edellisestä kaavalla

$$p_n = p_{n-1} - \frac{P(p_{n-1})}{P'(p_{n-1})}. \tag{7.2}$$

Kyseinen kaava voidaan tulkita siten, että edellistä approksimaatiota p_{n-1} vastaavaan polynomin käyrän pisteeseen $(p_{n-1}, P(p_{n-1}))$ asetetaan polynomin tangentti. Uusi approksimaatio p_n on tällöin tangentin ja x-akselin leikkauspiste, sillä tangentin yhtälö on

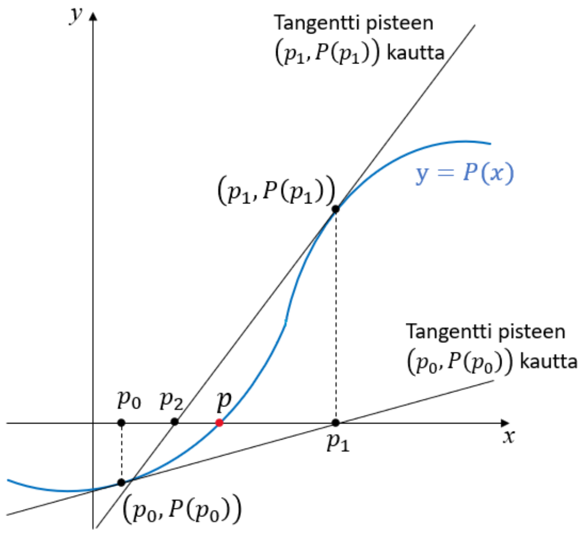
$$y - P(p_{n-1}) = P'(p_{n-1})(x - p_{n-1}),$$

mistä saadaan sijoittamalla $y = 0$ x-akselin leikkauspisteelle kaava (7.2) [**12**, §74]. Tätä graafista tulkintaa on havainnollistettu kuvassa 7.1.

Esimerkki 7.5. Tutkitaan polynomia $P(x) = -3x^7 + 3x^4 + 7x - 2$, jonka derivaatta on $P'(x) = -21x^6 + 12x^3 + 7$. Nyt $P(0) = -2$ ja $P(1) = 5$, joten merkin muuttumiseen perustuvan menetelmän nojalla polynomilla P on nollakohta välillä $]0, 1[$. Määritetään nollakohdan likiarvo seitsemän desimaalin tarkkuudella. Valitaan alkuapproksimaatioksi $p_0 = 0$. Seuraaviksi approksimaatioksi kaavalla (7.2) saadaan:

$$\begin{aligned} p_0 &= 0 \\ p_1 &= 0{,}285714285 \\ p_2 &= 0{,}283027967 \\ p_3 &= 0{,}283026634 \\ p_4 &= 0{,}283026634. \end{aligned}$$

Tästä saadaan nollakohdan likiarvoksi seitsemän desimaalin tarkkuudella $p = 0{,}2830266$. Valittu alkuapproksimaatio oli tarpeeksi lähellä nollakohtaa, jolloin haluttu tarkkuus saavutettiin jo neljännellä approksimaatiolla.



Kuva 7.1. Newtonin menetelmän graafinen tulkinta. Todellista nollakohtaa p on aluksi approksimoitu luvulla p_0, jonka kautta piirretyn tangentin ja x-akselin leikkauspiste antaa seuraavan approksimaation p_1. Tämän approksimaation avulla on saatu edelleen seuraava approksimaatio p_2, joka antaa paremman arvion nollakohdalle kuin aiemmat approksimaatiot.

Jos taas alkuapproksimaatioksi valitaan välin $]0,1[$ toinen päätepiste $p_0 = 1$, saadaan seuraava approksimaatioiden jono:

$$\begin{aligned} p_0 &= 1 \\ p_1 &= 3,5 \\ p_2 &= 3,005565282 \\ p_3 &= 2,584183768 \\ &\vdots \\ p_{12} &= 1,245278347 \\ p_{13} &= 1,245278346. \end{aligned}$$

Huomataan, että nyt jono ei konvergoikaan välillä $]0,1[$ olevaan juureen. Saatu likiarvo vastaa kuitenkin polynomin toista juurta. Tämän likiarvon saavuttamiseksi jouduttiin kuitenkin laskemaan useampia approksimaatioita.

Esimerkin jälkimmäisestä approksimaatiojonosta huomataan myös, että alkuapproksimaatio voi olla lähempänä todellista juuren arvoa kuin sitä välittömästi seuraavat approksimaatiot eli likiarvo ei välttämättä heti tarkennu Newtonin menetelmällä. Esimerkki korosti myös, kuinka tärkeää on valita alkuapproksimaatio tarpeeksi läheltä todellista juuren arvoa. Kuinka läheltä todellista juuren arvoa alkuarvo on valittava riippuu tutkittavasta polynomifunktiosta. Riittävän tarkka alkuapproksimaatio on kuitenkin mahdollista löytää kaikille polynomeille ja kaikille juurille [**1**, §2.3].

Kaavasta (7.2) nähdään myös, ettei menetelmä toimi, jos $P'(p_n) = 0$ jollekin approksimaatiolle p_n. Toisaalta esimerkin ensimmäisestä approksimaatiojonosta huomataan, että parhaimmillaan Newtonin menetelmällä saadaan tarkkojakin likiarvoja nopeasti.

7.3. Sekanttimenetelmä

Tässä menetelmässä tarkennetaan juuren likiarvoa määrittämällä polynomille sekantti aiempien approksimaatioiden avulla. Olkoon $P(a)$ ja $P(b)$ erimerkkisiä ja $a < b$, jolloin polynomilla P on juuri välillä $]a, b[$. Valitaan a ja b approksimaatioiksi juuren likiarvolle. Polynomin sekantin yhtälö pisteiden $(a, P(a))$ ja $(b, P(b))$ kautta on muotoa

$$y - P(a) = \frac{P(b) - P(a)}{b - a}(x - a).$$

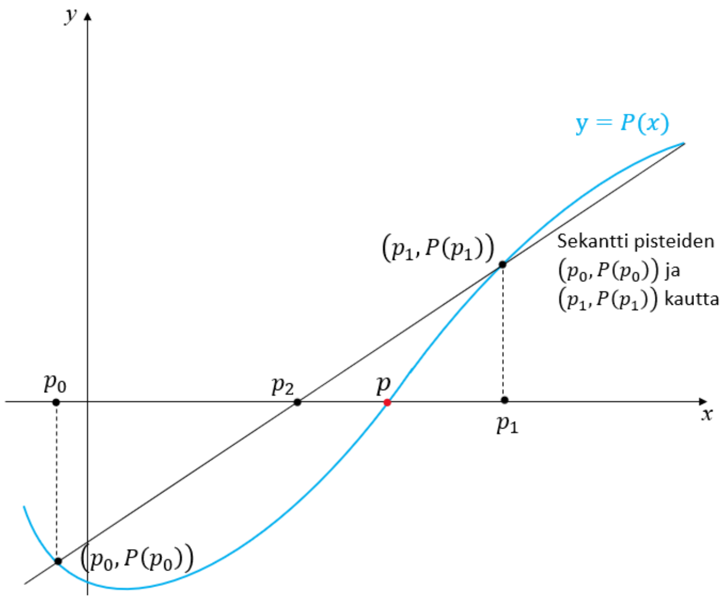
Sekantin ja x-akselin leikkauspiste saadaan asettamalla $y = 0$:

$$x = a - P(a)\frac{b - a}{P(b) - P(a)}.$$

Tämä leikkauspiste antaa uuden approksimaation juuren likiarvolle. Uuden approksimaation avulla saadaan edelleen seuraava approksimaatio kaavalla

$$p_n = p_{n-1} - P(p_{n-1})\frac{p_{n-2} - p_{n-1}}{P(p_{n-2}) - P(p_{n-1})}. \tag{7.3}$$

Näin saadaan jono (p_n), joka konvergoi kohti polynomin P juurta, jos alkuapproksimaatiot p_0 ja p_1 on valittu riittävän tarkasti [**1**, §2.3]. Kuva 7.2 havainnollistaa sekanttimenetelmän toimintaa.



KUVA 7.2. Sekanttimenetelmän graafinen esitys. Todellista nollakohtaa p approksimoidaan aluksi luvuilla p_0 ja p_1. Niiden avulla määritetään polynomin käyrälle sekantti, jonka leikkauspiste x-akselin kanssa antaa seuraavan approksimaation p_2.

Esimerkki 7.6. Tarkastellaan polynomia $P(x) = -3x^7 + 3x^4 + 7x - 2$, jolle määritettiin nollakohta $p = 0,2830266$ seitsemän desimaalin tarkkuudella Esimerkissä 7.5. Määritetään kyseinen nollakohta samalla tarkkuudella sekanttimenetelmällä. Valitaan alkuapproksimaatioiksi $p_0 = 0$ ja $p_1 = 1$. Kaavalla (7.3) saadaan approksimaatiojono:

$$
\begin{aligned}
p_0 &= 0 \\
p_1 &= 1 \\
p_2 &= 0,285714285 \\
p_3 &= 0,282914008 \\
p_4 &= 0,283026578 \\
p_5 &= 0,283026634 \\
p_6 &= 0,283026634
\end{aligned}
$$

Sekanttimenetelmällä päädytään siis samaan likiarvoon $p = 0,2830266$. Tällä menetelmällä jouduttiin kuitenkin laskemaan useampia approksimaatioita, jotta saavutettiin sama tarkkuus kuin Newtonin menetelmällä. Newtonin menetelmä konvergoikin yleensä nopeammin kuin sekanttimenetelmä, mutta alkuapproksimaatioiden valinta molemmissa menetelmissä vaikuttaa paljon konvergointinopeuteen. [**1**, §2.4]

Tässä tutkielmassa tutustuttiin erilaisiin menetelmiin niin polynomien reaalijuurien lukumäärän kuin juurien arvojenkin selvittämiseksi. Mikä menetelmistä on milloinkin hyödyllisin riippuu tutkittavasta polynomista sekä siitä, mitä halutaan selvittää ja millä tarkkuudella.

Kirjallisuutta

[1] R.L. Burden, J.D. Faires, *Numerical Analysis: Fourth Edition.* PWS Publishing Co., Boston, 1988.

[2] B. Datta, A.N. Singh, *History of Hindu Mathematics: A Source Book.* Asia Publishing House, Bombay, 1962.

[3] J. Kangasaho, J. Mäkinen, J. Oikkonen, J. Paasonen, M. Salmela, J. Tahvanainen, *Pitkä matematiikka: 2 Polynomifunktio.* WSOY, Helsinki, 2004.

[4] J.H. Matthews, R.W. Howell, *Complex Analysis for Mathematics and Engineering: Fourth Edition.* Jones and Bartlett Publishers, 2001.

[5] F. Nevanlinna, *Johdatus lukuteoriaan ja algebraan.* Otava, Helsinki, 1944.

[6] Opetushallitus, *Lukion opetussuunnitelman perusteet 2015.*

[7] Opetushallitus, *Perusopetuksen opetussuunnitelman perusteet 2014.*

[8] A. Rasila, *Sturmin lause.* Matematiikkalehti Solmu 2/2009, 2009.

[9] E. Saksman, *Kolmannen asteen yhtälöä ratkaisemassa.* Matematiikkalehti Solmu 1/2000-2001, 2000.

[10] J. Sesiano, *An Introduction to the History of Algebra: Solving Equations from Mesopotamian Times to the Renaissance.* American Mathematical Society, 2009

[11] B.S. Thomson, J.B. Bruckner, A.M. Bruckner, *Elementary Real Analysis: Second Edition.* Prentice Hall, 2008.

[12] K. Väisälä, *Lukuteorian ja korkeamman algebran alkeet: Toinen painos.* Otava, Helsinki, 1961.