PRO GRADU -TUTKIELMA

Teemu Ojansivu

Polynomien resultanteista

TAMPEREEN YLIOPISTO
Informaatiotieteiden yksikkö
Matematiikka
Helmikuu 2015

Tampereen yliopisto

Matematiikan ja tilastotieteen laitos

Ojansivu, Teemu: Polynomien resultanteista

Pro gradu -tutkielma, 27 s.

Matematiikka

Helmikuu 2015

Tiivistelmä

Resultantti on tärkeä ja käyttökelpoinen tulos algebrassa. Tutkielmassa käydään läpi matriisimuotoinen Sylvesterin resultantti ja resultantti kahdesta polynomista muodostettujen lineaaristen polynomien tulona. Ensin kuitenkin tarkastellaan polynomeja yleisesti. Sekä symmetrisiä että homogeenisia polynomeja tarkastellaan erikseen erikoistapauksina. Lopuksi käydään tutkielman pääaihe, eli resultantti, ja sille yleisiä tuloksia lävitse. Viimeiseksi tarkastellaan diskriminanttia ja resultantin käyttöä eliminoinnissa.

Tutkielman päälähteinä on käytetty Serge Langin kirjaa Algebra ja Coxin, Littlen sekä O'Shean teosta Ideals, Varieties And Algorithms.

Sisältö

1 Johdanto

Resultantti on kahden polynomin kertoimista muodostuva polynomi, jonka arvolla on yhteys polynomien juuriin. Sillä jos polynomeilla on yhteinen juuri, on niiden resultantin arvo nolla. Myös diskriminantti on kerrointa vaille polynomin ja sen derivaatan resultantti. Resultantti on tärkeä tulos lukuteoriassa ja esimerkiksi eliminoinnissa.

Tässä tutkielmassa tarkastellaan yleisesti polynomien resultantteja, niiden ominaisuuksia ja soveltamista eliminointiin. Tutkielmassa on käytetty pääosin lähteinä Serge Langin kirjaa Algebra ja Coxin, Littlen sekä O'Shean teosta Ideals, Varieties And Algorithms. Haivainnollistamiseen on käytetty yksinkertaisia keksittyjä esimerkkejä.

Aluksi luvussa 2 tarkastellaan polynomeja algebrallisina konstruktioina. Määritellään polynomit äärettöminä jonoina, vaikkakin tutkielman tarkastelu keskittyy niihin muodollisina kirjoitelmina eli monomien summina. Määritellään polynomit yhdelle ja useammalle muuttujalle sekä polynomirengas. Lisäksi määritellään yksi olennainen konstruktio - nimittäin algebrallinen riippumattomuus.

Luvussa 3 tarkastellaan symmetrisiä ja homogeenisia polynomeja sekä niiden ominaisuuksia ja keskeisiä tuloksia. Alkeissymmetriset polynomit käsitellään tärkeänä erikoistapauksena yksityiskohtaisesti.

Luvussa 4 esitellään ja määritellään resultantti ensin matriisin, jota kutsutaan Sylvesterin matriisiksi, determinantin muodossa. Tällöin resultantti saadaan yleisesti missä tahansa kommutatiivisessa renkaassa. Seuraavaksi tutkitaan resultantin yhteyttä polynomien juuriin sekä esitetään että todistetaan yleisiä tuloksia resultanteille, mistä päästään tutkielman päätulokseen eli resultantin määritelmään lineaaristen polynomien tulona. Diskriminantilla on suora yhteys resultanttiin ja se käsitelläänkin erikseen tarkemmin. Lopuksi tarkastellaan resultanttien käyttöä eliminoinnissa ja yhtälöryhmien ratkaisemisessa.

Tutkielmassa edellytetään lukijalta algebran perusasioiden tuntemista. Esimerkiksi lauseen 4.4 todistuksessa tarvitaan Cramerin sääntöä ja joissain lauseissa resultantin ominaisuuksia algebrallisen sulkeuman käsitettä.

2 Valmistelevia tarkasteluja

2.1 Polynomit

Polynomeja käsitellään perinteisesti yleensä laskulausekkeina, joissa *muuttujan* x paikalle sijoitetaan jokin luku tai muuttujan arvo selvitetään yhtälöstä. Nyt tarkastelemme polynomeja kuitenkin sellaisinaan eli muodollisina summina termeistä. Kertoimet voidaan valita mielivaltaisesta renkaasta ja juuri kertoimista ollaan kiinnostuneita. Tällä tavalla saamme polynomeille yleisiä tuloksia riippumatta muuttujien arvoista. Siksi algebrassa usein muuttujaa kutsutaankin *tuntemattomaksi* (erityisesti englanninkielisessä kirjallisuudessa) ja usein sitä merkitään isolla kirjaimella.

Tarkastellaan ensin polynomeja funktioina. Olkoon R rengas. Funktioiden $f : R \to R$ joukko on selvästi rengas. Funktioiden f ja g summa $(f+g)(x) = f(x) + g(x)$ ja tulo $(fg)(x) = f(x)g(x)$. Polynomien alijoukko

$$\{a_0 + a_1x + \cdots + a_nx^n \mid a_i \in R\}$$

on tämän renkaan alirengas. On helppo osoittaa, että yllämainitut polynomien yhteen- ja kertolasku muodostuvat seuraavasti:

$$\begin{aligned}&(a_0 + a_1xa_2x^2 + \dots) + (b_0 + b_1x + b_2x^2 + \dots)\\ &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots \quad \text{ja}\end{aligned}$$

$$\begin{aligned}&(a_0 + a_1xa_2x^2 + \dots) + (b_0 + b_1x + b_2x^2 + \dots)\\ &= (a_0b_0) + (a_1b_0 + a_0b_0)x + (a_2b_0 + a_1b_1 + a_0b_2)x^2 + \dots\ .\end{aligned}$$

Näin määriteltyinä yhteen- ja kertolaskun tuloksena saadut polynomit ovat myös R-kertoimisia polynomeja. Nyt polynomeja halutaan tarkastella kuitenkin algebrallisina objekteina eikä niinkään osana funktiorengasta. Nyt muuttujia merkitään tästä eteenpäin isoilla kirjaimilla.

Muuttujalla X ei ole näiden tarkasteluiden kannalta merkitystä, joten polynomi voidaan määritellä täysin ainoastaan kertoimiensa perusteella. Muuttujaa X tarvitaan ainoastaan merkitsemään kertoimien a_i paikkoja, jolloin kerroin a_i vastaa muuttujan potenssia X^i. Voisimme täten määritellä polynomit äärettöminä jonoina renkaan R alkioita eli $(a_0, a_1, \dots, a_n, 0, 0, \dots)$, jossa alkiot a_i kuuluvat renkaaseen R ja muuttuja X voidaan mieltää ainoastaan määräämättömänä symbolina, tuntemattomana.

Määritelmä 2.1. Olkoon R kommutatiivinen rengas. Tällöin polynomi, jonka kertoimet kuuluvat renkaaseen R, on ääretön jono

$$(a_i) = (a_0, a_1, a_2, \dots),$$

missä $a_i \in R$ kaikilla $i \in \mathbb{N}$ ja vain äärellisen moni a_i poikkeaa nollasta.

Näin voidaan määritellä muuttuja jonona $(0, 1, 0, 0, 0, \dots)$ ja vakio vastaavasti jonona $(a, 0, 0, 0, \dots)$.

Esimerkki 2.1. Polynomia $-2X^5 + X^3 + 5X^2 - 3X + 13$ vastaa ääretön jono $(13, -3, 5, 1, 0, -2, 0, 0, \dots)$. Jonosta voidaan myös jättää pois termit, joiden kerroin on nolla. Tällöin edellinen polynomi olisi siis (äärellistä) muotoa $(13, -3, 5, 1, 0, -2)$.

Tässä gradussa käytetään pääasiassa yleisempää merkintätapaa eli polynomit merkitään monomien summina ja jonomuotoa käytetään polynomien määrittelyn apukeinona.

2.2 Polynomirengas

Käsitellään polynomeja nyt algebrallisina objekteina. Muuttujaa käsitellään muodollisena symbolina, jonka arvosta ei olla kiinnostuneita.

Olkoon R kommutatiivinen rengas ja $R[\mathbb{N}]$ sellainen funktioiden $f : \mathbb{N} \to R$ joukko, että $f(n) = 0$, kun $n \gg 0$. Funktioiden $f, g \in R[\mathbb{N}]$ summa määritellään $(f+g)(n) = f(n)+g(n)$. Kertolasku määritellään myös luonnollisesti

$$(fg)(n) = \sum_{i+j=n} f(i)g(j),$$

missä $i, j \in \mathbb{N}$. Esimerkiksi

$$(fg)(4) = f(4)g(0) + f(3)g(1) + f(2)g(2) + f(1)g(3) + f(0)g(4).$$

Tässä symbolilla $X^i \in R[\mathbb{N}]$ merkitään funktiota

$$X^i(n) = \begin{cases} 1, & \text{jos } n = i, \\ 0, & \text{jos } n \neq i. \end{cases}$$

Huomaa, että $X^i X^j = X^{i+j}$, missä $i, j \in \mathbb{N}$. Alkiota $a \in \mathbb{R}$ merkitään funktiona

$$a(n) = \begin{cases} a, & \text{jos } n = 0, \\ 0, & \text{jos } n > 0 \end{cases}$$

renkaassa $R[\mathbb{N}]$. Täten alkio $f \in R[\mathbb{N}]$ voidaan merkitä

$$f = a_0 + a_1 X + \cdots + a_n X^n,$$

missä $a_i = f(i)$ ja $f(i) = 0$, jos $i > n$. Huomaa, että $1 = X^0$ on neutraalialkio kertolaskun suhteen. Yhteenlaskun neutraalialkio on $0 \in R$. Huomataan, että selvästi $fg = gf$ ja $f(g + h) = fg + fh$, kun $f, g, h \in R[\mathbb{N}]$. On helppo osoittaa, esimerkiksi aluksi olettamalla $h = cX^m$, missä $c \in R$ ja $m \in \mathbb{N}$, että kertolasku on assosiatiivinen eli $f(gh) = (fg)h$ kaikille $f, g, h \in R[\mathbb{N}]$.

Määritelmä 2.2. Määritellään yhden muuttujan polynomien polynomirengas. Kaikkien tällaisten polynomien joukkoa merkitään

$$R[X] = \{f = a_0 + a_1X + \cdots + a_nX^n |\ n \geq 0, a_i \in R\ \ (i = 0, \ldots, n)\}.$$

Joukon alkioita sanotaan polynomeiksi yli renkaan R. Jos $a_n \neq 0$, sanotaan *termin* a_nX^n kerrointa a_n polynomin f johtavaksi kertoimeksi. Polynomin aste, jota merkitään $\deg(f)$, saadaan johtavasta ketoimesta $\deg(f) = n$. Jos johtava kerroin on 1, kutsutaan polynomia *pääpolynomiksi*. Polynomia, jossa on vain yksi termi, kutsutaan *monomiksi*.

Huomaa, että polynomin aste on kuvaus $\deg : R[X] \setminus 0 \to \mathbb{N}$. Nollapolynomin astetta ei ole määritelty, mutta joskus näkee käytettävän $-\infty$.

Polynomit $f = a_0 + \cdots + a_nX^n$ ja $g = b_0 + \cdots + b_mX^m$ ovat samat, jos ja vain jos $n = m$ ja $a_i = b_i$ kaikille $i \in \{1, \ldots, n\}$.

2.3 Usean muuttujan polynomit

Usean muuttujan polynomit voidaan määritellä hyvin samankaltaiseen tapaan kuin aiemmin käsitellyt yhden muuttujan polynomit. Määritellään n:n muuttujan $X_1, \ldots, X_n$ polynomirengas $R[X_1, \ldots, X_n]$ seuraavasti:

$$R[X_1, \ldots, X_n] = R[\mathbb{N}^n] = \{f : \mathbb{N}^n \to R \ \mid\ f(v) = 0, |v| \gg 0\},$$

missä $v = (v_1, \ldots, v_n) \in \mathbb{N}^n$ ja $|v| = v_1 + \cdots + v_n$. Polynomi $f \in R[X_1, \ldots, X_n]$ on siis kuvaus $f : \mathbb{N}^n \to R$ ja vain äärellisen moni $f(v) \in \mathbb{N}^n$ on nollasta eroava. Aivan kuten aiemmin, $X^v \in R[\mathbb{N}^n]$ vastaa siis kuvausta

$$X^v(w) = \begin{cases} 1, & \text{jos } v = w, \\ 0, & \text{jos } v \neq w. \end{cases}$$

Näin jokainen polynomi $f \in R[\mathbb{N}^n]$ voidaan kirjoittaa äärellisenä summana

$$f = \sum_{v \in \mathbb{N}^n} a_v X^v,$$

missä $a_v \in R$. Jos $f, g \in R[\mathbb{N}^n]$ määritellään polynomien summa $(f+g)(v) = f(v) + g(v)$ ja tulo äärellisenä summana

$$(fg)(v) = \sum_{v_1 + v_2 = v} f(v_1)g(v_2),$$

missä $v_1, v_2 \in \mathbb{N}^n$. On helppo todistaa, että $R[\mathbb{N}^n]$ on rengas. Yhteenlaskun nolla-alkio on $0 \in R$ ja kertolaskun neutraalialkio on kuvaus $X^{(0,0,\ldots,0)}$, joka kuvaa $\mathbb{N}^n$:n nollavektorin alkioksi $1 \in R$ ja kaiken muun nollaksi. Nyt huomataan myös, että näillä merkinnöillä muuttujaa X_1 vastaa $X^{(1,0,\ldots,0)}$ ja niin edelleen.

2.4 Algebrallinen riippumattomuus

Olkoon B rengas, R sen alirengas ja $a_{j_1,\ldots,j_m} \in R$. Jos $f \in R[x_1, \ldots, x_m]$ on polynomi, määritellään usean muuttujan polynomifunktio seuraavasti:

$$f_B : B^m \to B,$$

missä

$$f_B(x_1, \ldots, x_m) = \sum a_{j_1,\ldots,j_m} x_1^{j_1} \cdots x_m^{j_m}.$$

Jos nyt $x_1, \ldots, x_m \in B$ nähdään, että kuvaus

$$f \mapsto f(x_1, \ldots, x_m)$$

on rengashomomorfismi renkaalta $R[x_1, \ldots, x_m]$ renkaaseen B. Kuvausta voidaan kutsua myös sijoitushomomorfismiksi.

Erikoistapauksena yhden muuttujan funktio määritellään vastaavasti eli $f_B : B \to B$ ja $x \mapsto f(x)$.

Olkoon edelleen $x \in B$. Nyt nähdään, että renkaan B alirengas $R[x]$, jonka x virittää R:n yli, on kaikkien kuvauksen $f(x)$ arvojen rengas, kun $f \in R[X]$. Jos kuvaus $f \mapsto f(x)$ on isomorfismi $R[X]$:ltä $R[x]$:lle, sanotaan x:n olevan transsendenttinen yli R:n.

Määritelmä 2.3. Jos kuvaus $f \mapsto f(x)$ on injektiivinen, kutsutaan alkioita $x_1, \ldots, x_n \in B$ *algebrallisesti riippumattomiksi.*

Toisin sanoen, jos $f \in R[X]$ on polynomi ja $f(x) = 0$, seuraa että $f = 0$.

Lause 2.1. *Olkoon B rengas ja R sen alirengas. Alkiojono $x_1, ..., x_n \in B$ on algebrallisesti riippumaton, jos ja vain jos sijoitushomomorfismi $X_i \mapsto x_i$ on isomorfismi $f : R[X_1, ..., X_n] \to R[x_1, ..., x_n]$.*

Todistus. Oletetaan, että joukko $\{x_1, \ldots, x_n\}$ on algebrallisesti riippumaton ja $f \neq 0$, joten $f(x) \neq 0$. Tällöin $Ker(f) = \{0_R\}$, joten kuvaus on injektiivinen. (Rengashomomorfismi, jonka ydin koostuu ainoastaan nolla-alkiosta, on injektio.) Se on myös selvästi surjektiivinen, joten se on bijektio ja täten kuvaus on siis isomorfismi.

Nyt oletetaan, että sijoitushomomorfismi on isomorfismi (siis se on myös injektiivinen), jolloin todistus algebrallisen riippumattomuuden määritelmän 2.3 mukaan seuraa välittömästi. □

Huomaa, että sijoitushomomorfismin maalijoukko $R[x_1, \ldots, x_n]$ on renkaan B alirengas. Tarkemmin ottaen se on suppein sellainen rengas, joka sisältää alirenkaan R ja alkiot $x_1, \ldots, x_n$. Nyt $R[x_1, \ldots, x_n]$ siis koostuu alkioista $f(x_1, \ldots, x_n)$, missä $f \in R[X_1, \ldots, X_n]$. Alkiot $x_1, \ldots, x_n$ eivät kuitenkaan välttämättä generoi rengasta B.

Algebrallinen riippumattomuus voidaan siis tulkita myös seuraavasti: joukon $S = \{x_1, x_2, \ldots, x_n\}$ alkiot ovat algebrallisesti riippumattomat, jos ja vain jos ei ole olemassa sellaista polynomia $0 \neq f \in R[X_1, \ldots, X_n]$, että

$$f(x_1, x_2, \ldots, x_n) = 0.$$

Seuraus 2.1. Polynomirengas $R[X_1, \ldots, X_n, Y]$ on isomorfinen renkaan $R[X_1, \ldots, X_n][Y]$ kanssa. $R[X_1, \ldots, X_n][Y]$ on siis muuttujan Y polynomirengas, jonka kertoimet ovat renkaasta $R[X_1, \ldots, X_n]$.

Esimerkki 2.2. Joukko $\{X_1, X_2\}$ on selvästi algebrallisesti riippumaton.

Joukko $\{X_1, X_2, X_1^3 + 2X_2^2\}$ ei ole, sillä jos sijoitetaan joukon alkiot polynomiin $f(Y_1, Y_2, Y_3) = Y_1^3 + 2Y_2^2 - Y_3$, saadaan

$$f(X_1, X_2, X_1^3 + 2X_2^2) = X_1^3 + 2X_2^2 - (X_1^3 + 2X_2^2) = 0.$$

3 Polynomien ominaisuuksia

3.1 Symmetriset polynomit

Tarkastellaan kolmannen asteen polynomia $f = X^3 + bX^2 + cX + d$, jolla olkoon juuret q_1, q_2 ja q_3. Tällöin

$$X^3 + bX^2 + cX + d = (X - q_1)(X - q_2)(X - q_3).$$

Kertomalla oikea puoli auki saadaan

$$X^3 + bX^2 + cX + d = X^3 - (q_1 + q_2 + q_3)X^2 + (q_1q_2 + q_1q_3 + q_2q_3)X - q_1q_2q_3,$$

joten on oltava

$$\begin{aligned} b &= -(q_1 + q_2 + q_3) \\ c &= q_1q_2 + q_1q_3 + q_2q_3 \\ d &= q_1q_2q_3. \end{aligned}$$

Nähdään, että polynomin f kertoimet ovat polynomeja juurien q_i suhteen. Juurien järjestyksen valinnalla ei ole vaikutusta polynomiin f, joten kertoimia a, b ja c kuvaavat polynomit pysyvät muuttumattomina kaikilla juurien q_1, q_2 ja q_3 permutaatioilla. Tällaisia polynomeja kutsutaan *symmetrisiksi*.

Määritelmä 3.1. Olkoon R rengas. Polynomia $f \in R[X_1, \ldots, X_n]$ kutsutaan *symmetriseksi* polynomiksi, jos

$$f(X_{i_1}, \ldots, X_{i_n}) = f(X_1, \ldots, X_n)$$

kaikilla muuttujien $X_1, \ldots, X_n$ permutaatioilla $X_{i_1}, \ldots, X_{i_n}$.

Esimerkiksi polynomit $X^2 + Y^2 + Z^2$ ja XYZ ovat symmetrisiä, mutta polynomit $X^4 + Y$ ja XY^2Z eivät ole.

Edellä tarkastellut polynomin juurista muodostuvat uudet polynomit ovat siis muotoa

$$\begin{aligned} s_1 &= q_1 + \cdots + q_n, \\ s_2 &= q_1q_2 + q_1q_3 \cdots + q_{n-1}q_n, \\ s_3 &= q_1q_2q_3 + q_1q_2q_4 + \cdots + q_{n-2}q_{n-1}q_n, \\ &\cdots \\ s_n &= q_1 \cdots q_n. \end{aligned}$$

Nyt siis polynomien $s_1, \ldots, s_n$ muuttujina ovat $q_1, \ldots, q_n$. Jokainen polynomi s_i on siis kaikkien niiden monomien, joissa on i eri muuttujaa, summa. Myös jokaisen tällaisen monomin aste on i.

Määritelmä 3.2. Olkoon $q_1, \ldots, q_n$ muuttujia. Polynomeja

$$s_k = s_k(q_1, \ldots, q_n) = \sum_{1 \leq j_1 < \cdots < j_k \leq n} q_{j_1} q_{j_2} \cdots q_{j_k}, \ k = 1, \ldots, n$$

kutsutaan *alkeissymmetrisiksi polynomeiksi.*

Lemma 3.1. *Alkeissymmetriset polynomit $s_1, \ldots, s_n$ ovat symmetrisiä polynomeja eli*

$$s_k(q_{i_1}, \ldots, q_{i_n}) = s_k(q_1, \ldots, q_n)$$

aina, kun $k = 1, \ldots, n$.

Todistus. Sijoitetaan alkiot $q_1, \ldots, q_n$ polynomin $f(X)$ juuriksi:

$$(*) \qquad f(X) = (X - q_1)(X - q_2) \cdots (X - q_n).$$

Nyt jos kerrotaan oikea puoli auki, saadaan

$$f(X) = X^n - s_1 X^{n-1} + s_2 X^{n-2} + \cdots + (1)^{n-1} s_{n-1} X + (-1)^n s_n.$$

Jos nyt järjestetään alkiot $q_1, \ldots, q_n$ uusiksi, muuttuu yhtälön (*) oikean puolen järjestys, mutta itse polynomi f ei muutu. Täten polynomin f kertoimet $(-1)^i s_i$ ovat symmetrisiä polynomeja. □

Seurauksena tästä on, että jokaisen polynomin, jonka korkeimman asteen termin kerroin on 1, muut kertoimet ovat itsessään alkeissymmetrisiä polynomeja, jotka muodostuvat polynomin juurista.

Huomaa, että alkeissymmetriset peruspolynomit ovat algebrallisesti riippumattomat, jos alkiot, joista ne on muodostettu, ovat algebrallisesti riippumattomat. Tässä siis polynomin kertoimet ovat algebrallisesti riippumattomat, jos polynomin juuret ovat.

Lause 3.1. *Alkeissymmetriset polynomit $s_1, \ldots, s_n$ ovat algebrallisesti riippumattomat yli renkaan R.*

Todistus. (vrt. [1, s. 134]) Oletetaan nyt, että $s_1, \ldots, s_n$ eivät ole algebrallisesti riippumattomat. Tarkastellaan polynomia $f(X_1, \ldots, X_n) \in R[X]$, jonka aste on mahdollisimman pieni mutta kuitenkin nollasta eroava, ja jolle pätee

$$f(s_1, \ldots, s_n) = 0.$$

Kirjoitetaan polynomi $f(X)$ kertoimilla renkaasta $R[X_1, \ldots, X_{n-1}]$, jolloin saadaan

$$f(X_1, \ldots, X_n) = f_0(X_1, \ldots, X_{n-1}) + \cdots + f_d(X_1, \ldots, X_{n-1}) X_n^d.$$

Nyt $f_0 \neq 0$, sillä muutoin saataisiin

$$f(X) = X_n g(X),$$

missä g on jokin polynomi ja täten olisi $s_n g(s_1, \ldots, s_n) = 0$. Tästä seuraisi, että $g(s_1, \ldots, s_n) = 0$ ja että polynomin g aste olisi pienempi kuin polynomin f.

Sijoitetaan muuttujien X_i tilalle s_i ja saadaan

$$0 = f_0(s_1, \ldots, s_{n-1}) + \cdots + f_d(s_1, \ldots, s_{n-1}) s_n^d.$$

Tämä yhtälö pätee siis renkaassa $R[q_1, \ldots, q_n]$ ja sijoitetaan q_n:n paikalle 0. Nyt kaikki muut termit ovat 0 paitsi ensimmäinen, josta saadaan

$$0 = f_0((s_1)_0, \ldots, (s_{n-1})_0).$$

Tästä saadaan, että nyt alkeissymmetriset polynomit $s_1, \ldots, s_{n-1}$ muuttujinaan $q_1, \ldots q_{n-1}$ eivät ole algebrallisesti riippumattomat. Induktiota käyttäen päästään ristiriitaan, joten todistuksen oletus on epätosi. □

Alkeissymmetrisistä polynomeista voidaan muodostaa uusia symmetrisiä polynomeja. Esimerkiksi

$$s_2^2 - s_1 s_3 = X^2Y^2 + X^2YZ + X^2Z^2 + XY^2Z + XYZ^2 + Y^2Z^2$$

on symmetrinen polynomi. Yllättävää on, että kaikki symmetriset polynomit kunnassa K voidaan esittää yksikäsitteisesti alkeissymmetrisistä polynomeista $s_1, \ldots, s_n$ muodostettuina polynomeina (*symmetristen polynomien peruslause*). Tässä tutkielmassa ei kuitenkaan käsitellä symmetrisiä polynomeja tämän tarkemmin.

3.2 Homogeeniset polynomit

Määritelmä 3.3. Olkoon R kommutatiivinen rengas, $f \in R[X]$ ja $\deg(f) \geq 0$. Polynomia $f \in R[X_1, \ldots, X_n]$ kutsutaan *homogeeniseksi* polynomiksi, jos sen termit ovat monomeja, joiden on asteet ovat samat.

Esimerkki 3.1. Polynomi $X^3Y + 3X^4 + 2X^2Y^2$ on homogeeninen ja sen aste on 4.

Polynomi $Y^5X^2 - 4XYZ + Z^7$ ei ole homogeeninen, sillä termin $-4XYZ$ aste on 3, muiden termien aste on 7.

Esimerkki 3.2. Alkeissymmetriset polynomit ovat homogeenisia. Esimerkiksi s_4 on homogeeninen ja astetta 4. Symmetristen polynomien peruslauseesta seuraa, että kaikki symmetriset polynomit voidaan kuvata homogeenisten (ja symmetristen) polynomien summana.

Lause 3.2. *Olkoon R rengas ja $f \in R[X_1, \ldots, X_n]$ nollasta eroava polynomi. Tällöin f on homogeeninen ja astetta d, jos ja vain jos*

$$f(cX_1, \ldots, cX_n) = c^d f(X_1, \ldots, X_n),$$

kun $c, X_1, \ldots, X_n$ ovat algebrallisesti riippumattomia yli R:n.

Todistus. Oletetaan, että polynomi f on homogeeninen ja astetta d. Tällöin sen yleinen termi on muotoa $aX_1^{b_{X_1}} \cdots X_n^{b_{X_n}}$ ja $b_{X_1} + \cdots + b_{X_n} = d$. Kun sijoitetaan termiin $X_i = cX_i$, saadaan

$$a(cX)_1^{b_{X_1}} \cdots (cX)_n^{b_{X_n}} = ac^{(b_{X_1} + \cdots + b_{X_n})} X^{(b_{X_1} + \cdots + b_{X_n})}.$$

Koska f on homogeeninen ja sen aste on d, $c^{(b_{X_1} + \cdots + b_{X_n})} = c^d$ jokaisella termillä, joten se joidaan ottaa yhteiseksi tekijäksi, ja näin

$$f(cX_1, \ldots, cX_n) = c^d f(X_1, \ldots, X_n).$$

Oletetaan nyt, että

$$f(cX_1, \ldots, cX_n) = c^d f(X_1, \ldots, X_n).$$

Tästä seuraa, että yleinen termi on muotoa

$$a(cX)_1^{b_{X_1}} \cdots (cX)_n^{b_{X_n}} = c^d a X_1^{b_{X_1}} \cdots X_n^{b_{X_n}}.$$
$$c^{b_{X_1} + \cdots + b_{X_n}} a X_1^{b_{X_1}} \cdots X_n^{b_{X_n}} = c^d a X_1^{b_{X_1}} \cdots X_n^{b_{X_n}}.$$

Nyt $b_{X_1} + \cdots + b_{X_n} = d$, joten jokaisen termin aste on d, ja täten polynomi on homogeeninen. □

4 Resultantti

4.1 Sylvesterin resultantti

Lemma 4.1. *Olkoon K kunta ja $f, g \in K[X]$ polynomeja, joiden asteet ovat $m > 0$ ja $n > 0$. Silloin polynomeilla f ja g on positiivista astetta oleva yhteinen tekijä, jos ja vain jos on olemassa sellaiset polynomit $A, B \in K[X]$, että:*

(1) A ja B ovat nollasta eroavia polynomeja,

(2) polynomien A ja B asteet ovat korkeintaan $n-1$ ja $m-1$ ja

(3) $Af + Bg = 0$.

Todistus. (vrt. [2, s. 154]) Oletetaan, että polynomeilla f ja g on positiivista astetta oleva yhteinen tekijä $h \in K[X]$. Nyt $f = hf_1$ ja $g = hg_1$, missä $f_1, g_1 \in K[X]$. Huomaa, että $\deg(f_1) \leq m-1$ ja $\deg(g_1) \leq n-1$. Tällöin

$$g_1 \cdot f + (-f) \cdot g = g_1 \cdot hf_1 - f_1 \cdot hg_1 = 0.$$

Nähdään, että polynomeilla f_1 ja g_1 on halutut ominaisuudet, ja merkitään $A = g_1$ ja $B = f_1$.
Oletetaan nyt, että polynomeilla A ja B on yllä mainitut kolme ominaisuutta. Kohdan (1) mukaan $B \neq 0$. Jos polynomeilla f ja g ei ole yhtään positiivista astetta olevaa yhteistä tekijää, polynomien SYT $= 1$, ja voidaan löytää polynomit $A_1, B_1 \in K[X]$ joille $A_1 f + B_1 g = 1$. Kerrotaan yhtälö polynomilla B ja sijoitetaan $Bg = -Af$:

$$B = (A_1 f + B_1 g)B = A_1 Bf + B_1 Bg = A_1 Bf - B_1 Af = (A_1 B - B_1 A)f.$$

Koska $B \neq 0$, yhtälöstä nähdään, että polynomin B aste on oltava vähintään m, mikä on ristiriidassa kohdan (2) kanssa. Täten on oltava olemassa yhteinen tekijä, jonka aste on positiivinen. □

Lemman tulos ei kuitenkaan ole tyydyttävä, sillä siitä ei voi päätellä onko polynomeja A ja B ylipäätään olemassa. Lineaarialgebran keinoin voidaan muuttaa yhtälö $Af + Bg = 0$ lineaarisiksi yhtälöiksi. Kirjoitetaan

$$A = c_0 X^{n-1} + \ldots + c_{n-1},$$
$$B = d_0 X^{m-1} + \ldots + d_{m-1},$$

missä kertoimia $c_0, \ldots, c_{n-1}, d_0, \ldots, d_{m-1}$ käsitellään tuntemattomina. Tarkoitus on löytää sellaiset kertoimet $c_i, d_j \in K$, että $A, B \neq 0$, kertoimista ainakin osa nollasta eroavia ja että yhtälö

$$(*) \qquad Af + Bg = 0$$

on voimassa. Tästä saadaan Lemmassa 4.1 tarvittavat polynomit A ja B. Kirjoitetaan polynomit f ja g muotoon

$$\begin{aligned} f &= a_0X^m + \ldots + a_m, \quad (a_0 \neq 0), \\ g &= b_0X^n + \ldots + b_n, \quad (b_0 \neq 0), \end{aligned}$$

missä $a_i, b_j \in K$ $\forall i, j$. Sijoitetaan f, g, A ja B yhtälöön (*), jolloin saadaan

$$\begin{cases} a_0c_0 & + b_0d_0 & & = 0 & X^{m+n-1} : n \text{ kerroin} \\ a_1c_0 + a_0c_1 & + b_1d_0 + b_0d_1 & & = 0 & X^{m+n-2} : n \text{ kerroin} \\ \quad \ddots & \quad \ddots & \vdots & & \\ & a_mc_{n-1} \quad + & b_nd_{m-1} & = 0 & X^0 : n \text{ kerroin.} \end{cases}$$

Koska yhtälöitä on $m + n$ ja muuttujia $m + n$, tiedetään yhtälöryhmällä olevan nollasta eroava ratkaisu, jos ja vain jos yhtälöryhmän kerroinmatriisin determinantti on nolla. Toisin sanoen on olemassa ratkaisu, jos

$$\begin{vmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ a_2 & a_1 & \ddots & & b_2 & b_1 & \ddots & \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ & \vdots & & a_1 & & \vdots & & b_1 \\ a_m & & & & b_n & & & \\ & a_m & & \vdots & & b_n & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_m & & & & b_n \end{vmatrix} = 0,$$

missä tyhjät kohdat merkitään nollaksi. Tästä päästään erääseen resultantin määritelmään.

Määritelmä 4.1. Olkoon R kommutatiivinen rengas ja $u_0, \ldots, u_m, v_0, \ldots, v_n$ algebrallisesti riippumattomia alkioita renkaan R yli. Muodostetaan kaksi polynomia:

$$\begin{aligned} f_u(X) &= u_0X^m + \cdots + u_m \quad \text{ja} \\ g_v(X) &= v_0X^n + \cdots + v_n. \end{aligned}$$

Nyt resultantti voidaan määritellä joko alkioille (u, v) tai polynomeille f_u, g_v (riippuen kumpi on tarkoituksellisempaa) determinantilla

$$(**) \qquad \mathrm{m}+\mathrm{n}\left\{\underbrace{\begin{vmatrix} u_0 & & & & v_0 & & & \\ u_1 & u_0 & & & v_1 & v_0 & & \\ u_2 & u_1 & \ddots & & v_2 & v_1 & \ddots & \\ \vdots & & \ddots & u_0 & \vdots & & \ddots & v_0 \\ & \vdots & & u_1 & & \vdots & & v_1 \\ u_m & & & & v_n & & & \\ & u_m & & \vdots & & v_n & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & u_m & & & & v_n \end{vmatrix}}_{n \qquad\qquad m}\right. .$$

Tyhjät kohdat merkitään tässäkin nolliksi.

Kerroinmatriisia, josta determinantti (**) on muodostettu, kutsutaan *Sylvesterin matriisiksi* ja siitä muodostettua determinanttia (**) kutsutaan *Sylvesterin resultantiksi*. Kirjallisuudessa esiintyy usein myös matriisin transpoosi. Transponointihan ei muuta determinantin eli tässä tapauksessa resultantin arvoa.

Huomataan, että resultantti on kokonaislukukertoiminen polynomi, joten voitaisiin olettaa, että $R = \mathbb{Z}$. Lisäksi resultantissa $\mathrm{Res}(f_u, g_v)$ esiintyy monomi $u_0^n v_0^m$ kertoimella 1.

Sijoittamalla kertoimien $(u) = (u_0, \ldots, u_m)$ ja $(v) = (v_0, \ldots, v_n)$ paikalle renkaan R alkiot $(a) = (a_0, \ldots, a_m)$ ja $(b) = (b_0, \ldots, b_n)$ saadaan R-kertoimiset polynomit f_a ja g_b, joiden *resultantti* määritellään sijoittamalla determinanttiin polynomien kertoimien (u) ja (v) paikalle kertoimiksi (a) ja (b). Resultantti saadaan nyt siis sijoitushomomorfismilla mille tahansa renkaan R alkioille a ja b. Polynomien f_a, g_b resultanttia merkitään $\mathrm{Res}(f_a, g_b)$ tai $\mathrm{Res}(a, b)$ (tai lyhyemmin $\mathrm{R}(f_a, g_b)$).

Lause 4.1. *Olkoon c algebrallisesti riippumaton yli R:n. Nyt*

$$\mathit{Res}(cf_u, g_v) = c^n \mathit{Res}(f_u, g_v) \qquad \mathit{ja} \qquad \mathit{Res}(f_u, cg_v) = c^m \mathit{Res}(f_u, g_v).$$

Todistus. Seuraa suoraan determinanttien ominaisuuksista. Ensimmäisessä tapauksessa otetaan determinantista c yhteiseksi tekijäksi n:ltä ensimmäiseltä pystyriviltä ja jälkimmäisessä m:ltä pystyriviltä. □

Jos oletetaan polynomien f ja g asteiden olevan m ja n ja määritetään asteiden summalla matriisin pysty- ja vaakarivien lukumäärä kuten määritelmässä 4.1, polynomit määräävät resultantin yksikäsitteisesti (mukaan lukien triviaalit tapaukset $f = 0$ tai $g = 0$). Näin ei kuitenkaan aina välttämättä tarvitse olla, sillä joskus voi olla hyödyllistä määritellä yleisempi muoto valitsemalla $m \geq \deg(f)$ ja $n \geq \deg(g)$, jolloin $\mathrm{Res}(f, g)$ määräytyy polynomeista

ja annetuista arvoista m ja n. Tällöin voidaan sekaannuksen välttämiseksi käyttää myös merkintätapaa $\mathrm{Res}_{m,n}(f,g)$.

Lause 4.2. $\mathrm{Res}_{m,n}(f,g)$ *on kokonaislukukertoiminen homogeeninen polynomi muuttujinaan polynomien kertoimet* a_i, b_j.

(1) $\mathrm{Res}_{m,n}(f,g)$ *on homogeeninen ja astetta* n *muuttujien* $a_0,\ldots,a_m$ *suhteen ja astetta* m $b_0,\ldots,b_n$ *suhteen.*

(2) Jos muuttujien a_i *ja* b_i *asteet ovat* i*, on silloin* $\mathrm{Res}_{m,n}(f,g)$ *homogeeninen astetta* mn.

Todistus. (vrt. [4, s. 7]) Selvästikin $\mathrm{Res}(f,g)$ on homogeeninen kokonaislukukertoiminen polynomi, jonka aste on $m+n$. Jos nyt sijoitamme kertoimen a_i paikalle $c\,a_i, c \in \mathbb{N}$ ja kertoimen b_j paikalle $d\,b_j, d \in \mathbb{N}$ Sylvesterin resultanttiin, on ensimmäiset n pystyriviä kerrottu c:llä ja jälkimmäiset m pystyriviä d:llä, jolloin resultantin $\mathrm{Res}(f,g)$ kertoimeksi muodostuu $c^n d^m$, joten kohta (1) seuraa lauseesta 3.2.

Kohdan (2) todistus etenee samalla tavalla, jos sijoitamme a_i:n paikalle $c^i a_i$ ja b_j:n paikalle $c^j b_j$. Täten Sylvesterin resultantin kertoimeksi saadaan nyt c^{mn}. □

Esimerkki 4.1. Olkoon

$$f_a(X) = a_0X^m + \cdots + a_m \quad \text{ja}$$
$$g_b(X) = b_0X^n + \cdots + b_n.$$

Jos nyt $l = 2$ ja $m = 3$, niin tällöin

$$\mathrm{Res}(f_a, g_b) = \begin{vmatrix} a_0 & 0 & 0 & b_0 & 0 \\ a_1 & a_0 & 0 & b_1 & b_0 \\ a_2 & a_1 & a_0 & b_2 & b_1 \\ 0 & a_2 & a_1 & b_3 & b_2 \\ 0 & 0 & a_2 & 0 & b_3 \end{vmatrix} = a_0^3b_3^2 + a_2^3b_0^2.$$

Aiemmin määriteltiin resultantti polynomeille, joiden asteet ovat positiivisia kokonaislukuja. Miten resultantti määritellään vakiopolynomeille tai muille erikoistapauksille? Jos molemmat polynomit ovat vakioita ($f = a_0$ ja $g = b_0$), määritellään resultantti yleensä seuraavasti:

$$\mathrm{Res}(a_0, b_0) = \begin{cases} 1, & \text{jos } a_0 = 0 \text{ tai } b_0 = 0, \\ 0, & \text{jos } a_0 \neq 0 \text{ ja } b_0 \neq 0. \end{cases}$$

Jos toinen polynomi on vakio ja toisen aste on suurempi kuin nolla, eli $\deg(f) = m > 0$ ja $g = b_0 \in \mathbb{Z}$, on Sylvesterin resultantti diagonaalimatriisi, jonka diagonaalilla on b_0. Koska $\deg(g) = 0\ (= n)$, ei Sylvesterin resultantissa esiinny polynomin f kertoimia. Täten $\mathrm{Res}(f,g) = b_0^m$. Samoin nähdään, jos f on vakio ja g:n aste suurempi kuin nolla, on resultantti a_0^n.

Sylvesterin resultantista nähdään suoraan, että $\mathrm{Res}(g,f)$ saadaan resultantista $\mathrm{Res}(f,g)$ rivejä vaihtamalla.

Lause 4.3. *Jos polynomin f aste on m ja polynomin g aste on n, on resultantilla seuraava symmetrinen ominaisuus:*

$$\mathrm{Res}(f, g) = (-1)^{mn}\mathrm{Res}(g, f).$$

Todistus. (vrt. [4, s. 4]) Siirtääkseemme pystyrivin $(n+1)$ ensimmäiseksi vaatii se n siirtoa. Samoin pystyrivin $(n+2)$ siirtäminen toiseksi vaatii n siirtoa. Kaiken kaikkiaan siirtoja on m, ja jokainen siirto aiheuttaa determinantille kertoimen (-1), joten kokonaiskertoimeksi muodostuu $(-1)^{mn}$. □

4.2 Resultantti ja polynomien juuret

Käytetään samoja merkintöjä kuin resultantin määritelmässä 4.1 ja tarkastellaan resultanttia nyt renkaan $\mathbb{Z}$ yli.

Lause 4.4. *Jos f ja g ovat polynomeja, joiden asteet ovat positiivisia, on olemassa sellaiset polynomit $A_{u,v}$ ja $B_{u,v} \in \mathbb{Z}[u,v][X]$, että*

$$A_{u,v}\, f_u + B_{u,v}\, g_v = \mathrm{Res}(u,v) = \mathrm{Res}(f_u, g_v).$$

Nyt siis $A_{u,v}$ ja $B_{u,v}$ ovat polynomeja, joiden kertoimet ovat polynomien f ja g kertoimista muodostuvat kokonaislukukertoimiset polynomit.

Todistus. (vrt. [1, s. 136]) Oletetaan, että $\mathrm{Res}(f_u, g_v) \neq 0$. Muodostetaan lineaariset yhtälöt

$$\begin{aligned}
X^{n-1} f_u(X) &= u_0 X^{m+n-1} + u_1 X^{m+n-2} + \ldots + u_m X^{n-1}, \\
X^{n-2} f_u(X) &= u_0 X^{m+n-2} + \ldots + u_m X^{n-2}, \\
&\ldots \\
f_u(X) &= u_0 X^m + u_m, \\
X^{m-1} g_v(X) &= v_0 X^{m+n-1} + v_1 X^{m+n-2} + \ldots + v_n X^{m-1}, \\
X^{m-2} g_v(X) &= v_0 X^{m+n-2} + \ldots + v_n X^{m-2}, \\
&\ldots \\
g_v(X) &= v_0 X^n + v_n.
\end{aligned}$$

Olkoon C vasemman puolen pystyvektori ja olkoon

$$C_0, \ldots, C_{n+m-1}$$

yhtälöiden oikean puolen polynomien kertoimista muodostetut pystyvektorit. Täten voidaan kirjoittaa yhtälöt muodossa

$$C = X^{m+n-1} C_0 + \cdots + 1 \cdot C_{n+m-1}.$$

Koska viimeisen pystyvektorin C_{m+n-1} kerroin on 1, on Cramerin säännön mukaan ([1, s. 330])

$$1 \cdot \det(C_0, \ldots, C_{n+m-1}) = \det(C_0, \ldots, C_{n+m-2}, C).$$

Huomataan myös, että pystyvektoreista $C_0, \ldots, C_{m+n-1}$ muodostettu determinantti on sama kuin määritelmän 4.1 resultantti, joten

$$\mathrm{Res}(u,v) = \det(C_0, \ldots, C_{n+m-1}) = \det(C_0, \ldots, C_{n+m-2}, C).$$

Tästä nähdään, että on olemassa polynomit $A_{u,v}$ ja $B_{u,v} \in \mathbb{Z}[u,v][X]$ siten, että

$$A_{u,v} f_u + B_{u,v} g_v = \mathrm{Res}(u,v) = \mathrm{Res}(f_u, g_v).$$

Jos $\mathrm{Res}(f_u, g_v) = 0$, niin voidaan valita $A_{u,v} = 0$ ja $B_{u,v} = 0$. □

Huomaa yllä, että $\mathrm{Res}(u, v) \in \mathbb{Z}[u, v]$, mutta vasemman puolen polynomeissa esiintyy muuttuja X.

Jos $h : \mathbb{Z}[u, v] \mapsto R$ on homomorfismi kommutatiiviselle renkaalle R ja olkoon $h(u_i) = a_i$, $h(v_j) = b_j$, niin tällöin

$$A_{a,b}\, f_a + B_{a,b}\, g_b = \mathrm{Res}(a, b) = \mathrm{Res}(f_a, g_b).$$

Näin lauseen resultanttia koskeva yhtälö yli renkaan $\mathbb{Z}$ saadaan yleistettyä yhtälöksi mille tahansa parille polynomeja missä tahansa kommutatiivisessa renkaassa R.

Lause 4.5. *Olkoon K kunnan L alikunta ja $f_a, g_b \in K[X]$ polynomeja, joilla on yhteinen juuri r kunnassa L. Tällöin* $\mathrm{Res}(a, b) = 0$.

Todistus. (vrt. [1, s. 136]) Jos $f_a(r) = g_b(r) = 0$, sijoitetaan r muuttujan X paikalle saatuun resultantin $\mathrm{Res}(a, b)$ muotoon

$$A_{a,b}\, f_a(r) + B_{a,b}\, g_b(r) = R(a, b),$$

josta nähdään, että $\mathrm{Res}(a, b) = 0$. □

Tutkitaan nyt tarkemmin resultantin yhteyttä polynomien juuriin. Sitä varten ensiksi tarvitaan kaksi lemmaa.

Lemma 4.2. *Olkoon $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ n:n muuttujan polynomi. Jos f on yhtä suuri kuin* 0*, kun sijoitamme muttujan X_2 paikalle X_1 pitäen muut muuttujat $X_i (i \neq 2)$ kiinteinä, niin silloin $X_1 - X_2$ jakaa polynomin $f(X_1, \ldots, X_n)$ renkaassa $\mathbb{Z}[X_1, \ldots, X_n]$.*

Todistus. Kirjoitetaan X_1^k muodossa

$$X_1^k = (X_1 - X_2)X_1^{k-1} + X_2 X_1^{k-1}.$$

Nyt saadaan

$$f(X_1, \ldots, X_n) = (X_1 - X_2)P(X_1, \ldots, X_n) + Q(X_2, \ldots, X_n),$$

mikä on yhtä suuri kuin 0, kun $X_1 = X_2$, joten polynomin Q on oltava nollapolynomi. Tästä nähdään, että f on jaollinen polynomilla $X_1 - X_2$. □

Olkoon $u_0, q_1, \ldots, q_m, v_0, r_1, \ldots, r_n$ algebrallisesti riippumattomia alkioita yli renkaan $\mathbb{Z}$. Tällöin voidaan muodostaa polynomit

$$f_u = u_0(X - q_1) \cdots (X - q_m) = u_0 X^m + \cdots + u_m,$$
$$g_v = v_0(X - r_1) \cdots (X - r_n) = v_0 X^n + \cdots + v_n.$$

Nyt nähdään, että polynomien kertoimet muodostuvat juurista q_i ja r_j, joten

$$(*) \qquad u_i = (-1)^i u_0 s_i(q) \quad \text{ja} \quad v_j = (-1)^j v_0 s_j(r).$$

Lemma 4.3. *Alkiot $u_0, u_1, \ldots, u_m, v_0, v_1, \ldots, v_n$ ovat algebrallisesti riippumattomat yli renkaan $\mathbb{Z}$.*

Todistus. Jotta lemman alkiot olisivat algebrallisesti riippumattomat, ei siis voi olla olemassa sellaista nollasta eroavaa polynomia f, että kun sijoitetaan sen muuttujien paikalle lemman alkiot, olisi $f = 0$. Oletetaan nyt, että

$$f(u_0, u_1, \ldots, u_m, v_0, v_1, \ldots, v_n) = 0,$$

missä $f \in \mathbb{Z}[X_0, X_1, \ldots, X_m, Y_0, \ldots, Y_n]$. Kirjoitetaan f muotoon

$$f = \sum_{i,j} f_{i,j} \ ,$$

missä $f_{i,j}$ on homogeeninen sekä astetta i muuttujien $X_0, \ldots, X_m$ suhteen ja astetta j muuttujien $Y_0, \ldots, Y_n$ suhteen. Täten

$$\begin{aligned} &f(u_0, \ldots, u_m, v_0, \ldots, v_n) \\ &= \sum_{i,j} f_{i,j}(s_1(q), \ldots, s_m(q), s_1(r), \ldots, s_n(r))(-1)^{\frac{i(i+1)}{2}+\frac{j(j+1)}{2}} u_0^i v_0^j \\ &= 0. \end{aligned}$$

Tällöin siitä, että u_0 ja v_0 ovat algebrallisesti riippumattomia yli renkaan $\mathbb{Z}[q_1 \ldots, q_m, r_1, \ldots, r_n]$, seuraa, että kertoimet $f_{i,j}(s_1(q), \ldots, s_n(r))$ ovat nollia. Hajoitetetaan jokainen kerroin $f_{i,j}$ vielä muotoon

$$\sum g_{m_1,\ldots,m_n}(X_1, \ldots, X_m)) Y_1^{m_1} \cdots Y_n^{m_n},$$

siis

$$\sum g_{m_1,\ldots,m_n}(s_1(q), \ldots, s_m(q)) s_1(r)^{m_1} \cdots s_n(r)^{m_n} = 0.$$

Lauseen 3.1 nojalla $s_1(r), \ldots, s_n(r)$ ovat algebrallisesti riippumattomia yli renkaan $\mathbb{Z}[q_1, \ldots, q_m]$. Tämän vuoksi kertoimet $g_{m_1,\ldots,m_n}(s_1(q), \ldots, s_m(q))$ ovat nollia. Mutta lauseesta 3.1 seuraa jälleen, että polynomit $g_{m_1,\ldots,m_n}$ ovat nollia. Täten polynomit $f_{i,j}$ ovat nollia ja siis lopulta myös $f = 0$.

□

Lause 4.6. *Olkoon f ja g polynomeja, $u_0, q_1, \ldots, q_m, v_0, r_1, \ldots, r_n$ algebrallisesti riippumattomia yli $\mathbb{Z}$:n ja*

$$\begin{aligned} f_u &= u_0(X - q_1) \cdots (X - q_m) = u_0 X^m + \cdots + u_m, \\ g_v &= v_0(X - r_1) \cdots (X - r_n) = v_0 X^n + \cdots + v_n. \end{aligned}$$

Nyt saadaan

$$\mathrm{Res}(f_u, g_v) = u_0^n v_0^m \prod_{i=1}^{m} \prod_{j=1}^{n} (q_i - r_j).$$

Todistus. (vrt. [1, s. 137-138]) Merkitään S:llä yllä olevan yhtälön oikeaa puolta.

Yhtälöistä (*) ja siitä, että resultantti $\mathrm{Res}(f_u, g_v)$ on homogeeninen ja sen aste on n ensimmäisten muuttujiensa suhteen ja m toisten muuttujiensa suhteen, nähdään, että

$$\mathrm{Res}(f_u, g_v) = u_0^n v_0^m h(t, u),$$

missä $h(t, u) \in \mathbb{Z}[t, u]$.

Lauseesta 4.5 nähdään suoraan, että resultantti on yhtä suuri kuin nolla, jos sijoitamme q_i:n paikalle r_j, eli jos polynomeilla on vähintään yksi yhteinen juuri. Jos ajatellaan resultantin kuuluvan renkaaseen $\mathbb{Z}[q, r]$, lemmasta 4.2 seuraa, että se on jaollinen polynomilla $q_i - r_j$ jokaisella parilla (i, j). Täten S jakaa resultantin renkaassa $\mathbb{Z}[q, r]$, koska jokainen $q_i - r_j$ on jaoton renkaassa $\mathbb{Z}[q, r]$.

Muodosta

$$S = u_0^n v_0^m \prod_{i=1}^{m} \prod_{i=1}^{n} (q_i - r_j)$$

saadaan

$$\prod_{i=1}^{m} g_v(q_i) = v_0^m \prod_{i}^{m} \prod_{j=1}^{n} g_v(q_i - r_j),$$

missä

$$S = u_0^n \prod_{i=1}^{m} g_v(q_i). \tag{1}$$

Vastaavasti saadaan

$$S = (-1)^{mn} v_0^m \prod_{j=1}^{n} f_u(r_j). \tag{2}$$

Yhtälöstä (1) nähdään, että S on homogeeninen ja astetta m (v:n suhteen) ja yhtälöstä (2) nähdään, että S on homogeeninen ja astetta n (u:n suhteen). Koska resultantilla $\mathrm{Res}(f_u, g_v)$ on tismalleen samat homogeeniset ominaisuudet ja se on jaollinen S:llä, voidaan päätellä, että $\mathrm{Res}(f_u, g_v) = cS$ jollain luvulla c. Koska molempiin muotoihin $\mathrm{Res}(f_u, g_v)$ ja S kuuluu monomi $u_0^n v_0^m$ kertoimella 1, on oltava $c = 1$. □

Sijoitushomomorfismilla saadaan taas vastaavasti $\mathrm{Res}(f_a, g_b)$ mille tahansa renkaan R alkioille.

Seuraus 4.1. Olkoon f_a ja $g_b \in K[X]$ polynomeja siten, että $a_0 b_0 \neq 0$ ja että polynomit hajoavat astetta 1 oleviin tekijöihin renkaassa $K[X]$. Tällöin $\mathrm{Res}(f_a, g_b) = 0$ silloin ja vain silloin, kun polynomeilla f_a ja g_b on yhteinen juuri.

Todistus. Oletaan ensin, että resultantti on 0. Jos polynomit hajoavat tekijöihin

$$f_a = a_0(X - \alpha_1) \cdots (X - \alpha_m)$$
$$g_b = b_0(X - \beta_1) \cdots (X - \beta_m),$$

on olemassa sellainen homomorfismi

$$\mathbb{Z}[u_0, q, v_0, r] \mapsto K,$$

että $u_0 \mapsto a_0, v_0 \mapsto b_0, q_i \mapsto \alpha_i$ ja $r_j \mapsto \beta_j \quad \forall\, i, j$. Nyt on

$$0 = \operatorname{Res}(f_a, g_b) = a_0^n b_0^m \prod_i \prod_j (\alpha_i - \beta_j),$$

joten polynomeilla f_a ja g_b on yhteinen juuri.

Lause 4.5 todistaa seurauksen silloin kun oletetaan, että polynomeilla on yhteinen juuri. □

Esimerkki 4.2. Onko polynomeilla $f = 3X^2 + 2X + 5 \in \mathbb{Q}[X]$ ja $g = X^2 - 5X + 3 \in \mathbb{Q}[X]$ yhteinen juuri?

Lasketaan nyt polynomien resultantti käyttäen Sylvesterin resultanttia:

$$\operatorname{Res}(f, g) = \begin{vmatrix} 3 & 0 & 1 & 0 \\ 2 & 3 & -5 & 1 \\ 5 & 2 & 3 & -5 \\ 0 & 5 & 0 & 3 \end{vmatrix} = 543 \neq 0,$$

joten seurauksen 4.1 mukaan polynomeilla f ja g ei ole yhteistä juurta.

Polynomeilla $f = X^2 + X - 6$ ja $g = X^2 - 5X + 6$ on yhteinen juuri, sillä

$$\operatorname{Res}(f, g) = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & -5 & 1 \\ -6 & 1 & 6 & -5 \\ 0 & -6 & 0 & 6 \end{vmatrix} = 0.$$

Juuresta ei kuitenkaan näillä keinoilla tiedetä muuta kuin että polynomeilla on yhteinen juuri joko kunnassa $\mathbb{Q}$ tai jossain sen laajennuksessa.

Seuraavissa lauseissa tarvitaan polynomien juuria, mutta ne eivät välttämättä sisälly käsiteltävään renkaaseen tai kuntaan. Aiemmassa käsittelyssä polynomit muodostettiin juuriensa avulla, mutta aina ei näin ole. Tällöin tarvitaan algebrallisen sulkeuman käsitettä. Kuntaa K kutsutaan *algebrallisesti suljetuksi*, jos kaikkien polynomirenkaan $K[X]$ polynomien juuret kuuluvat kuntaan K. Voidaan osoittaa, että jokaisella kunnalla on laajennus, joka on algebrallisesti suljettu. Suppeinta tällaista laajennusta kutsutaan kunnan *algebralliseksi sulkeumaksi*.

Lause 4.7. *Olkoon f, h ja $g \in K[X]$ polynomeja, joille* $\deg(f) \leq m_1$, $\deg(h) \leq m_2$ *ja* $\deg(g) \leq n$ *ja*

$$\begin{aligned} f(X) &= u_0 X^{m_1} + \cdots + u_{m_1}, \\ h(X) &= w_0 X^{m_2} + \cdots + w_{m_2}, \\ g(X) &= v_0 X^n + \cdots + v_n. \end{aligned}$$

Tällöin

$$R_{m_1+m_2,n}(fh, g) = R_{m_1,n}(f, g) R_{m_2,n}(h, g).$$

Samoin jos f, g ja $h \in K[X]$ ovat polynomeja, joille $\deg(f) \leq m$, $\deg(g) \leq n_1$ *ja* $\deg(h) \leq n_2$, *on tällöin*

$$R_{m,n_1+n_2}(f, gh) = R_{m,n_1}(f, g) R_{m,n_2}(f, h).$$

Todistus. (vrt. [4, s. 9]) Voidaan olettaa, että nyt $\deg(g) = n$, joten polynomilla g on n juurta jossain kunnan K laajennuksessa, jolloin ensimmäinen kohta seuraa lauseen 4.6 resultantin muodosta (2):

$$\begin{aligned} S = R_{m_1+m_2,n}(fh, g) &= (-1)^{(m_1+m_2)n} v_0^{m_1+m_2} \prod_{j=1}^{n} (fh)(r_j) \\ &= (-1)^{m_1 n} v_0^{m_1} \prod_{j=1}^{n} f(r_j) \; (-1)^{m_2 n} v_0^{m_2} \prod_{j=1}^{n} h(r_j) \\ &= R_{m1,n}(f, g) R_{m_2,n}(h, g), \end{aligned}$$

missä r_j ovat polynomin g juuria.

Samoin jälkimmäinen seuraa vastaavasti kohdasta (1). □

Olkoon f ja g kuten aiemmin, ja jos määrittelemme "käänteiset" polynomit f^* ja $g^* \in K[X]$ seuraavasti:

$$\begin{aligned} f^*(X) &= X^m f(1/X) = a_0 + a_1 X + \cdots + a_m X^m, \\ g^*(X) &= X^n g(1/X) = b_0 + b_1 X + \cdots + b_m X^n. \end{aligned}$$

Nyt resultantilla on myös seuraava symmetrinen ominaisuus.

Lause 4.8. *Olkoon f ja $g \in K[X]$ polynomeja, joille* $\deg(f) \leq m$ *ja* $\deg(g) \leq n$. *Tällöin on voimassa symmetrinen ominaisuus*

$$R_{m,n}(f^*, g^*) = R_{n,m}(g, f) = (-1)^{mn} R_{m,n}(f, g).$$

Todistus. (vrt. [4, s. 9]) Lauseen 4.3 ja Sylvesterin resultantin mukaan tulos seuraa välittömästi, sillä $\mathrm{Syl}_{m,n}(f^*, g^*)$ saadaan matriisista $\mathrm{Syl}_{n,m}(g, f)$ pysty- ja vaakarivit vaihtamalla, joten niiden determinantit ovat samat. □

4.3 Diskriminantti

Resultantilla on myös yhteys polynomin diskriminanttiin. Diskriminanttia varten määritellään polynomin muodollinen derivaatta lyhyesti.

Määritelmä 4.2. Olkoon R rengas, $a_1, \ldots, a_m \in R$ ja polynomi

$$f = a_0X^m + a_{m-1}X^{m-1} + \cdots + a_m.$$

Tällöin polynomin f *muodollinen derivaatta*

$$D(f) = f' = a_0mX^{m-1} + a_1(m-1)X^{m-2} + \cdots + a_{m-1}.$$

Muodostetaan polynomi f kuten aiemmin, eli

$$f(X) = u_0(X - q_1)\cdots(X - q_m) = u_0X^m + \cdots + u_{m-1}X + u_m.$$

Lauseen 4.6 muodosta (1) ja polynomin f muodollisesta derivaatasta f' saadaan

$$\operatorname{Res}(f, f') = u_0^{m-1}\prod_i f'(q_i). \tag{3}$$

Tulon derivoimissäännöstä saadaan edelleen

$$f'(X) = \sum_i u_0(X - q_1)\cdots(X - q_{i-1})(X - q_{i+1})\cdots(X - q_m),$$

$$f'(q_i) = u_0(q_i - q_1)\cdots(q_i - q_{i-1})(q_i - q_{i+1})\cdots(q_i - q_m). \tag{4}$$

Määritellään nyt polynomin diskriminantti.

Määritelmä 4.3. Olkoon

$$f(X) = u_0(X - q_1)\cdots(X - q_m) = u_0X^m + \cdots + u_{m-1}X + u_m.$$

Tällöin sen diskriminantti $D(f)$ on

$$D(f) = (-1)^{m(m-1)/2}u_0^{2m-2}\prod_{i\neq j}(q_i - q_j).$$

Diskriminantti on tällöin määritelty polynomeille, jonka juuret ovat tunnetut. Jos halutaan päästä yleisempään tulokseen, voidaan käyttää seuraavia lauseita.

Lause 4.9. *Olkoon edelleen f kertoiminaan algebrallisesti riippumattomat alkiot yli renkaan $\mathbb{Z}$. Tällöin on*

$$\operatorname{Res}(f, f') = u_0^{2m-1}\prod_{i\neq j}(q_i - q_j) = (-1)^{m(m-1)/2}u_0D(f).$$

Todistus. Kun sijoitetaan aiemmin saatu derivaatan $f'(q_i)$ muoto (4) yhtälöön (3), tulos seuraa välittömästi. □

Määritelmä 4.4. Olkoon $f \in R[X]$ ja $f = a_0X^m + \cdots + a_{m-1}X + a_m$. Diskriminantti voidaan nyt määritellä myös sijoitushomomorfismin ja lauseen 4.9 avulla resultanttia käyttäen, jolloin

$$D(f) = \frac{(-1)^{m(m-1)/2}}{a_0}\text{Res}(f, f').$$

Nyt siis f on mikä tahansa polynomi, jonka kertoimet ovat mitkä tahansa alkiot renkaasta R.

Algebrassa diskriminantti on siis polynomin kertoimista muodostuva polynomi. Diskriminantin termien lukumäärä kasvaa eksponentiaalisesti polynomin asteen kasvaessa. Jos polynomilla on useampikertainen juuri, diskriminantti häviää eli on yhtä suuri kuin nolla. Tämä nähdään suoraan diskriminantin määritelmästä 4.3.

Lause 4.10. *Olkoon $f \in K[X]$ polynomi,* $\deg(f) = m \geq 1$ *ja f:n kertoimet kuuluvat kuntaan K. Tällöin polynomin juuret ovat yksinkertaiset jossain kunnan K laajennuksessa, jos ja vain jos diskriminantti on eri suuri kuin nolla.*

Toisin sanoen:

$D(f) = 0 \Leftrightarrow$ *polynomilla on useampikertainen juuri jossain kunnan K laajennuksessa.*

Todistus. Tulos seuraa suoraan diskriminantin määritelmästä 4.3. □

Esimerkki 4.3. Olkoon $f(X) = aX^2 + bX + c$, joten $f'(X) = 2aX + b$. Tällöin $\deg(f) = 2$ ja (Sylvesterin) resultantti on

$$\text{Res}(f, f') = \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = -a(b^2 - 4ac).$$

Lauseen 4.9 mukainen kerroin on siis $(-1)^{2(2-1)/2}a = -a$. Polynomin f diskriminantti on siis $D(f) = b^2 - 4ac$.

Determinantin monimutkaisuuden kasvun polynomin asteen kasvaessa huomaa jo kolmannen asteen polynomista. Jos nyt $f(X) = aX^3 + bX^2 + cX + d$, niin

$$\text{Res}(f, f') = a(27a^2d^2 + 4b^3d + 4ac^3 - b^2c^2 - 18abcd)$$

ja

$$D(f) = \frac{(-1)^{3(3-1)/2}}{a}\text{Res}(f, f') = b^2c^2 + 18abcd - 27a^2d^2 - 4b^3d - 4ac^3.$$

4.4 Resultantit ja eliminointi

Resultantit ovat käyttökelpoisia eliminoinnissa. Lasketaan esimerkiksi polynomien $f = XY + 1$ ja $g = X^2 + Y^2 + 1$ resultantti muuttujan X suhteen:

$$\operatorname{Res}(f_X, g_X) = \begin{vmatrix} Y & 0 & 1 \\ 1 & Y & 0 \\ 0 & 1 & Y^2 + 1 \end{vmatrix} = Y^4 + Y^2 + 1.$$

Nyt $Y^4 + Y^2 + 1 \neq 0$, joten yhtälöillä ei ole yhteistä juurta.

Yleisemmin nähdään, että jos f ja g ovat mitä tahansa polynomeja (joiden asteet ovat positiivisia kokonaislukuja eliminoitavan muuttujan suhteen), voidaan resultantti laskea tietyn muuttujan suhteen samoin kuin yllä. Järjestetään polynomit niin, että eliminoitavan muuttujan kertoimet ovat polynomien kertoimista ja muista muuttujista muodostettuja polynomeja. Nyt myös eliminoitavan muuttujan suhteen laskettu resultantti on polynomi, joka muodostuu f:n ja g:n kertoimista ja muista kuin eliminoitavasta muuttujasta. Yllä siis $\operatorname{Res}(f_X, g_X)$ on polynomi, jonka muuttujana on Y, joten näin voidaan resultanttia käyttämällä eliminoida muuttuja X. Vastaavasti voidaan myös muut muuttujat eliminoida.

Esimerkki 4.4. Onko yhtälöparilla

$$\begin{cases} X^2 + Y^2 = 2 \\ Y = 2X - 1 \end{cases}$$

epätriviaalia ratkaisua?

Lasketaan polynomien resultantti muuttujan X suhteen:

$$\operatorname{Res}(f_X, g_X) = \begin{vmatrix} 1 & 0 & -2 \\ 0 & 1 & Y + 1 \\ Y^2 - 2 & 0 & 0 \\ 0 & Y^2 - 2 & 0 \end{vmatrix} = 0,$$

joten yhtälöparilla on yhteinen juuri muuttujalla X. Samoin

$$\operatorname{Res}(f_Y, g_Y) = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 - 2X \\ X^2 - 2 & 0 & 0 \\ 0 & X^2 - 2 & 0 \end{vmatrix} = 0.$$

Viitteet

[1] Lang, S.: *Algebra*. Addison-Wesley Publishing Company Inc., 1965.

[2] Cox, D., Little, J., O'Shea, D.: *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. New York: Springer, 2000.

[3] Lauritzen, N.: *Concrete Abstract Algebra: From Numbers to Gröbner Bases*. Cambridge, Cambridge University Press, 2003.

[4] Janson, S.: *Resultant And Discriminant Of Polynomials*. http://www2.math.uu.se/ svante/papers/sjN5.pdf, 2007.