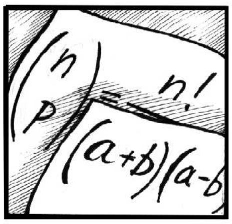

Sturmin lause

Antti Rasila
Matematiikan laitos, Teknillinen korkeakoulu

Sturmin ongelma

Aina joskus matematiikassa törmää ongelmaan, joka näyttää vaikealta, mutta kuitenkin ratkeaa suhteellisen yksinkertaisella päättelyllä. Tässä esitettävä Sturmin lause on tällainen tulos. Se liittyy läheisesti Algebran peruslauseeseen ja Descartesin merkinvaihtosääntöön.

Tulos kuuluu ranskalaiselle matematikolle Jacques Sturmille (1803-1855), joskin Joseph Fourier oli tiettävästi ratkaissut ongelman jo aiemmin. Fourierin paperi kuitenkin ilmeisesti katosi ja unohdettiin Ranskan vallankumouksen melskeissä. Tuloksen mielenkiintoisuudesta huolimatta se ei ole käsitykseni mukaan kovin laajasti tunnettu edes matemaatikkojen keskuudessa. Syynä tähän lienee se, että ratkaisun löytymisen jälkeen matemaatikot menettivät kiinnostuksensa ongelmaan, eikä ratkaisulla pitkään aikaan ollut erityistä teoreettista tai käytännöllistä merkitystä. Tietokoneiden myötä tilanne on muuttunut, itse törmäsin tulokseen ensimmäisen kerran symbolisen laskennan kurssilla.

Ongelma: Etsittävä reaalikertoimisen polynomiyhtälön

$$P(x) = 0$$

reaalisten ratkaisujen lukumäärä välillä (α, β), missä $\alpha < \beta$.

Sturmin ongelman ratkaisu

Seuraavaksi paneudutaan Sturmin ongelmaan vuonna 1829 esittämään yksinkertaiseen ja eleganttiin ratkaisuun. Tarkastellaan kahta erillistä tapausta:

1. Kaikki annetun yhtälön reaaliset juuret ovat yksinkertaisia välillä (α, β).

2. Annetulla yhtälöllä on kyseisellä välillä myös moninkertaisia juuria.

Tarkastellaan aluksi tapauksista jälkimmäistä. Oletetaan siis, että yhtälöllä $P(x) = 0$ on erilliset juuret $r_1, r_2, r_3, \ldots$, ja merkitään juuren r_j kertalukua k_j, kun $j = 1, 2, \ldots$. Voidaan kirjoittaa

$$P(x) = a(x - r_1)^{k_1}(x - r_2)^{k_2}(x - r_3)^{k_3} \ldots$$

Polynomin $P(x)$ derivaatalle $P'(x)$ saadaan esitys

$$\begin{aligned}\frac{P'(x)}{P(x)} &= \frac{k_1}{x - r_1} + \frac{k_2}{x - r_2} + \frac{k_3}{x - r_3} + \ldots \\ &= \frac{k_1(x - r_2)(x - r_3) \ldots + k_2(x - r_1)(x - r_3) \ldots + \ldots}{(x - r_1)(x - r_2)(x - r_3) \ldots} \\ &\equiv \frac{p(x)}{q(x)}.\end{aligned}$$

Merkitään lisäksi $G(x) \equiv P(x)/q(x)$. Saadaan yhtälöt $P(x) = G(x)q(x)$ ja $P'(x) = G(x)p(x)$.

Voidaan olettaa, että lausekkeilla $p(x)$ ja $q(x)$ ei ole yhteisiä tekijöitä, joten $G(x)$ on $P(x)$:än ja $P'(x)$:än suurin yhteinen tekijä. Siten yhtälö $P(x) = 0$ on hajotettu tapauksiin $q(x) = 0$ ja $G(x) = 0$. Näistä ensimmäisellä on vain yksinkertaisia juuria, ja jälkimmäiseen voidaan iteratiivisesti soveltaa samaa menettelyä.

Seuraus: Riittää ratkaista ongelma ensimmäisessä tapauksessa.

Tarkastellaan seuraavaksi tapauksista ensimmäistä. Oletetaan siis, että yhtälön $P(x) = 0$ kaikki juuret ovat yksinkertaisia. Tällöin derivaatta $P'(x)$ ei saa arvoa nolla yhdessäkään näistä juurista. Lisäksi $P(x)$:n ja $P'(x)$:än suurin yhteinen tekijä on nollasta poikkeava vakio C.

Merkitään $f_0(x) \equiv P(x)$, $f_1(x) \equiv P'(x)$. Määritellään rekursiivisesti $f_j(x)$ ja $q_j(x)$ seuraavan skeeman mukaisesti:

$$\begin{aligned}
&(0) \qquad f_0 = q_0 f_1 - f_2 \\
&(1) \qquad f_1 = q_1 f_2 - f_3 \\
&(2) \qquad f_2 = q_2 f_3 - f_4 \\
&\qquad\qquad \ldots
\end{aligned}$$

Toisin sanoen, kukin q_j on peräkkäisissä jakolaskuissa osamäärä ja $-f_{j+2}$ vastaavasti jakojäännös.

Lopulta skeemassa esiintyy nollasta poikkeava vakio C, jolla erityisesti on sama merkki koko välillä (α, β). Merkitään $f_s \equiv C$. Tällöin sanotaan, että konstruoidut *Sturmin funktiot*

$$\bar{f} = f_0, f_1, f_2, \ldots, f_s$$

muodostavat *Sturmin ketjun.*

Aputulos. Sturmin funktioilla on seuraavat ominaisuudet:

1. Kaksi peräkkäistä Sturmin funktiota eivät katoa samanaikaisesti yhdessäkään välin (α, β) pisteessä.
2. Kunkin Sturmin funktion f_j nollakohdassa viereiset funktiot f_{j-1}, f_{j+1} ovat erimerkkisiä.
3. Funktion f_0 nollakohdan riittävän pienen ympäristön sisällä funktio f_1 on joko aidosti positiivinen tai aidosti negatiivinen.

Todistus. (1) Jos esimerkiksi $f_2(\sigma) = 0$ ja $f_3(\sigma) = 0$ jossakin välin (α, β) pisteessä σ, niin tällöin

$$f_4(\sigma) = q_2(\sigma) f_3(\sigma) - f_2(\sigma) = 0,$$

eli σ on myös funktion f_4 nollakohta. Siten myös $f_5(\sigma) = 0$, jne. ja edelleen $f_s(\sigma) = 0$. Tämä on ristiriita, koska f_s on vakio $C \neq 0$.

(2) Jos funktiolla f_3 on nollakohta σ, niin määritelmän nojalla

$$f_2(\sigma) = -f_4(\sigma),$$

jne.

(3) Seuraa siitä, että f_1 on jatkuva funktio, eikä f_0:n nollakohta ole f_1:n nollakohta. □

Olkoon nyt $x \in (\alpha, \beta)$. Sturmin funktioiden f_j pisteessä x saavuttamien arvojen $f_0(x), f_1(x), \ldots, f_s(x)$ etumerkit muodostavat *Sturmin merkkiketjun.*

Merkitään $\mathrm{Var}_{\bar{f}}(x)$ Sturmin ketjussa $\bar{f}$ tapahtuvien etumerkin vaihtojen ($+-$ tai $-+$) lukumäärä pisteessä x. Määritellään edelleen polynomijonon etumerkin vaihtojen erotus

$$\mathrm{Var}_{\bar{f}}[\alpha, \beta] \equiv \mathrm{Var}_{\bar{f}}(\alpha) - \mathrm{Var}_{\bar{f}}(\beta).$$

Tutkitaan funktion $\mathrm{Var}\bar{f}(x)$ arvon muuttumista, kun x on välillä (α, β). Havaitaan, että funktion arvo voi muuttua pisteessä x vain, jos yksi tai useampi Sturmin funktio vaihtaa etumerkkiään kyseisessä pisteessä.

Olkoon $\sigma \in (\alpha, \beta)$ Sturmin funktion f_v nollakohta. Valitaan $\delta > 0$ siten, että välillä $\Delta = (\sigma - \delta, \sigma + \delta)$ seuraavat ehdot ovat voimassa:

1. Funktio $f_v(x) \neq 0$, kun $x \in \Delta$ ja $x \neq \sigma$.
2. Funktion f_v:n viereiset funktiot (f_{v-1}, f_{v+1}) eivät vaihda merkkiään välillä Δ.

Tutkitaan erikseen tapauksia $v > 0$ ja $v = 0$. Ensimmäisessä tapauksessa tarkastellaan kolmikkoa f_{v-1}, f_v, f_{v+1}, toisessa paria f_0, f_1.

Kolmikon tapauksessa havaitaan, että f_{v-1},f_{v+1} ovat erimerkkiset (Aputulos/kohta 2.), joten riippumatta f_v:n etumerkistä välillä Δ on aina täsmälleen yksi merkinvaihto. Täten f_v:n nollakohdan ohittaminen ei muuta $\mathrm{Var}_{\bar{f}}(x)$:n arvoa.

Parin tapauksessa havaitaan, että f_1:llä on tarkasteluvälillä vakio etumerkki $+$ tai $-$. Ensimmäisessä tapauksessa f_0 on kasvava ja $f_0(\sigma - \delta) < 0$, $f_0(\sigma + \delta) > 0$. Toisessa tapauksessa f_0 vähenevä ja etumerkit muuttuvat päinvastaisesti. Molemmissa tapauksissa nollakohdan σ jälkeen funtioilla f_0, f_1 on sama etumerkki.

Havaitaan, että funktion $\mathrm{Var}_{\bar{f}}(x)$ arvossa tapahtuu muutos täsmälleen silloin, kun ohitetaan funktion f_0 nollakohta. Edelleen, funktion f_0 nollakohdassa $\mathrm{Var}_{\bar{f}}(x)$:n arvo vähenee yhdellä.

Olemme siis osoittaneet seuraavan tuloksen:

Lause. (Sturmin lause) Olkoot $P(x)$ reaalikertoiminen polynomi ja $\alpha < \beta$ pisteitä, jotka eivät ole $P(x)$:an nollakohtia. Tällöin funktio $\mathrm{Var}_{\bar{f}}[\alpha, \beta]$, missä $\bar{f}$ on $P(x)$:än Sturmin ketju, antaa $P(x)$:än nollakohtien määrän välillä (α, β).

Sovelluksia

Esimerkki 1. Määritä yhtälön $x^5 + 3x - 1 = 0$ ratkaisujen lukumäärä ja sijainti.

Ratkaisu. Sturmin ketjuksi saadaan

$$\begin{aligned} f_0 &= x^5 - 3x - 1, \\ f_1 &= 5x^4 - 3, \\ f_2 &= 12x + 5, \\ f_3 &= 1. \end{aligned}$$

Polynomin $\bar{f}$ etumerkit vastaten muuttujan x arvoja $x = -2, -1, 0, +1, +2$ ovat:

x	f_0	f_1	f_2	f_3
-2	$-$	$+$	$-$	$+$
-1	$+$	$+$	$-$	$+$
0	$-$	$-$	$+$	$+$
$+1$	$-$	$+$	$+$	$+$
$+2$	$+$	$+$	$+$	$+$

Täten yhtälöllä on kolme reaaliratkaisua, yksi välillä $(-2, -1)$, toinen välillä $(-1, 0)$ ja kolmas välillä $(+1, +2)$. Loput ratkaisut ovat kompleksisia. □

Esimerkki 2. Määritä yhtälön $x^5 - ax - b = 0$ reaalisten juurten lukumäärä, kun a ja b ovat positiivisia ja $4^4a^5 > 5^5b^4$.

Ratkaisu. Sturmin ketjuksi saadaan

$$\begin{aligned} f_0 &= x^5 - ax - b, \\ f_1 &= 5x^4 - a, \\ f_2 &= 4ax + 5b, \\ f_3 &= 4^4a^5 - 5^5b^4. \end{aligned}$$

Kun $x \to -\infty$, saadaan raja-arvona ketjun etumerkeiksi $-+-+$. Vastaavasti, kun $x \to +\infty$, saadaan etumerkit $++++$. Siten yhtälöllä on kolme reaalista ja kaksi kompleksista ratkaisua. □