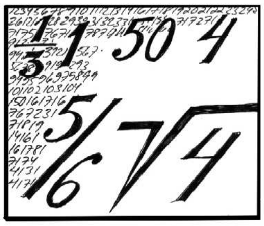

Tekijäfunktio ja muita lukuteoreettisia otuksia

Anne-Maria Ernvall-Hytönen

Ensi silmäyksellä tekijäfunktio

$$d(n) = \sum_{d|n} 1$$

ei välttämättä näytä maailman mielenkiintoisimmalta matemaattiselta objektilta. Se on kuitenkin yhteydessä esimerkiksi Riemannin kuuluisaan ζ-funktioon, sillä kun $\Re s > 1$, niin

$$\zeta(s)^2 = \left(\sum_{n=1}^{\infty} \frac{1}{n^s}\right)^2 = \sum_{n=1}^{\infty} \frac{d(n)}{n^s}.$$

Tekijäfunktion arvo on varsin helppo laskea, jos luvun alkutekijähajotelma on tiedossa, sillä

$$d\left(p_1^{\alpha_1} \cdots p_k^{\alpha_k}\right) = (\alpha_1 + 1) \cdots (\alpha_k + 1).$$

Valitettavasti luvun alkutekijähajotelmaa ei yleensä voi olettaa tunnetuksi (harvoinpa edes tiedämme, onko luku alkuluku vai ei). Tällöin on varsin hankala sanoa yhtään mitään tekijäfunktion suuruudesta. Onnellista kyllä, tekijäfunktion keskimääräisestä suuruudesta pystymmekin sanomaan aika paljon. Aloitetaan tarkastelemalla summaa

$$\sum_{n=1}^{M} d(n).$$

Huomataan aluksi, että sen sijaan, että laskettaisiin miten monta tekijää jollakin luvulla on, voidaan laskea miten moni luku on kullakin luvulla jaollinen:

$$\sum_{n=1}^{M} d(n) = \sum_{n=1}^{M} \left\lfloor \frac{M}{n} \right\rfloor.$$

Alaspäinpyöristykset hieman häiritsevät elämää yhtälön oikealla puolella, mutta kirjoitetaan nyt

$$\left\lfloor \frac{M}{n} \right\rfloor = \frac{M}{n} - c_n,$$

missä $0 \leq c_n < 1$. Oikea puoli voidaankin nyt kirjoittaa muodossa

$$\sum_{n=1}^{M} \left\lfloor \frac{M}{n} \right\rfloor = \sum_{n=1}^{M} \frac{M}{n} - \sum_{n=1}^{M} c_n = \sum_{n=1}^{M} \frac{M}{n} + O(M),$$

missä $O(M)$ tarkoittaa virhetermiä, joka on korkeintaan jokin vakio kertaa M, kun M on tarpeeksi suuri. Tässä tapauksessa on toki helppo nähdä, että virhetermi on itse asiassa korkeintaan M. Summa on puolestaan varsin helppo arvioida:

$$\sum_{n=1}^{M} \frac{M}{n} = M \sum_{n=1}^{M} \frac{1}{n} \leq M + M \int_1^M \frac{1}{x} dx = M + M \log M$$

ja toisaalta

$$\begin{aligned}\sum_{n=1}^{M} \frac{M}{n} = M \sum_{n=1}^{M} \frac{1}{n} &\geq \int_M^{M+1} \frac{1}{x} dx \\ &= M \log(M+1) \geq M \log M.\end{aligned}$$

Anne-Maria Ernvall-Hytönen toimii Vetenskapsrådetin rahoittamana postdoc-tutkijana Kungliga Tekniska Högskolanissa.

(Tässä artikkelissa log on luonnollinen logaritmi.) Kaiken kaikkiaan siis

$$\sum_{n=1}^{M} d(n) = M \log M + O(M),$$

missä merkinnän O implikoima vakio ei välttämättä ole sama kuin aiemmin. $O(M)$ on sinänsä varsin mukava virhetermi, että se on pienempi kuin $M \log M$, mutta toki voi miettiä, saisikohan tätä virhetermiä runnottua hieman pienemmäksi. Onnellista kyllä, virhetermi suostuu pienenemään varsin alkeellisilla menetelmillä suuruusluokkaan $O(\sqrt{M})$, mutta tällöin pitää ottaa toinenkin päätermi mukaan:

$$\sum_{n \le M} d(n) = M \log M + (2\gamma - 1)M + O(\sqrt{M}),$$

missä $\gamma = \lim_{x\to\infty} \left(\sum_{n\le x} \frac{1}{n} - \log x\right)$. Todistus jätetään innokkaalle lukijalle harjoitustehtäväksi.

Näiden laskujen tuloksena pääsemme johtopäätökseen, että tekijäfunktion arvo joukossa $\{1, 2, \ldots, M\}$ on keskimäärin $\log M$. Seuraava luonnollinen kysymys on se, kuinka paljon arvot tästä heittelehtivät. Selvästi kaikille alkuluvuille p pätee $d(p) = 2$ ja alkulukujen potensseille $d(p^\ell) = \ell + 1$. Toisin sanoen, todella pieniä arvoja kyllä löytyy, mutta ei välttämättä kovin tiheässä. Toisaalta voidaan arvioida:

$$d(n) = \sum_{d|n} 1 = 2 \sum_{d|n,\ d\le\sqrt{n}} \le 2\sqrt{n}.$$

Pienille luvun n arvoille tämä yläraja ei ole välttämättä kovinkaan huono: Esimerkiksi $d(6) = 4$, ja $2\sqrt{6} < 5$. Suurille luvun n arvoille tämä sen sijaan on jo pahasti pielessä. Tämän todistaminen on jo selvästi hankalampaa ja työläämpää, mutta koska se on kuitenkin tehtävissä alkeellisilla menetelmillä, siirrymme nyt siihen. Tästä lähtien voimme siis olettaa, että luku n on suuri. Kirjoitetaan $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, ja olkoon δ reaaliluku, jolle annetaan arvo myöhemmin. Tarkastellaan aluksi osamäärää

$$\frac{d(n)}{n^\delta} = \prod_{i=1}^{k} \left(\frac{\alpha_i + 1}{p_i^{\alpha_i\delta}}\right).$$

Oikean puolen tuloa voidaan arvioida ylöspäin, jos siitä poistetaan kaikki sellaiset termit, joilla $p_i^\delta > 2$, sillä tällöin $p_i^{\delta\alpha_i} > 2^{\alpha_i} \ge \alpha_i + 1$ (viimeinen vaihe on helppo todistaa vaikka induktiolla), eli $\frac{\alpha_i+1}{p_i^{\alpha_i\delta}} < 1$. Jäljelle jää tulo

$$\prod_{1\le i\le r,\ p_i^\delta\le 2} \left(\frac{\alpha_i + 1}{p_i^{\alpha_i\delta}}\right).$$

Tämän summan arviointia voidaan taas jatkaa manipuloimalla nimittäjää. Osoitetaan ensin, että

$$p^{\delta\alpha_1} = e^{\alpha_1\delta\log p_i} > 1 + \delta\alpha_i \log p_i.$$

Yhtäsuuruus on triviaali. Epäyhtälö puolestaan saadaan, kun tarkastellaan funktiota $f(y) = e^y - 1 - y$, ja huomataan, että $f'(y) = e^y - 1 = 0$, kun $y = 0$ ja $f'(1) > 0$, jolloin funktio on kasvava positiivisilla muuttujan y arvoilla. Alkuperäinen epäyhtälö on todistettu, kun vielä huomataan, että $f(0) = 0$, eli positiivisilla muuttujan y arvoilla myös funktion arvo on positiivinen. Kun otetaan luonnollinen logaritmi ehdosta $p_i^\delta \le 2$ molemmilta puolilta, saadaan $\delta p_i \le \log 2 < 1$, ja niinpä

$$1 + \delta\alpha_i \log p_i > \delta \log p_i(\alpha_i + 1).$$

Tämä muotoilu vaikuttaa varsin lupaavalta, sillä nyt voimme hankkiutua eroon osoittajasta. Lauseke siis muokkaantuu seuraavasti:

$$\prod_{1\le i\le r,\ p_i^\delta\le 2} \left(\frac{\alpha_i + 1}{p_i^{\alpha_i\delta}}\right) \le \prod_{1\le i\le r,\ p_i^\delta\le 2} \left(\frac{\alpha_i + 1}{(\alpha_i + 1)\delta\log p_i}\right) = \prod_{1\le i\le r,\ p_i^\delta\le 2} \left(\frac{1}{\delta\log p_i}\right).$$

Lauseke ei ainakaan kasva, jos alkuperäisen luvun n alkutekijöiden $p_1, p_2, \ldots, p_k$ joukkoon lisätään ne joukon ulkopuoliset alkuluvut, jotka toteuttavat ehdon $p^\delta \le 2$, eli $p \le 2^{1/\delta}$. Kaikki alkuluvut p ovat suuruudeltaan vähintään ≥ 2, ja korkeintaan luvun $2^{1/\delta}$. Siispä

$$\prod_{1\le i\le r,\ p_i^\delta\le 2} \left(\frac{1}{\delta\log p_i}\right) \le \prod_{p\le 2^{1/\delta}} \left(\frac{1}{\delta\log p}\right) \le \prod_{p\le 2^{1/\delta}} \left(\frac{1}{\delta\log 2}\right) \le \left(\frac{1}{\delta\log 2}\right)^{2^{1/\delta}}.$$

Olemme nyt päätyneet arvioon

$$\frac{d(n)}{n^\delta} \le \left(\frac{1}{\delta\log 2}\right)^{2^{1/\delta}}.$$

Otetaan tästä logaritmit puolittain, jolloin päädytään lausekkeeseen (jota on ehkä hieman helpompi manipuloida kuin ylläolevaa, vaikka periaatteessa ideat menevät läpi kummassa tahansa tilanteessa)

$$\log d(n) \le \delta \log n + 2^{1/\delta} \log\left(\frac{1}{\delta\log 2}\right).$$

Valitaan nyt

$$\delta = \frac{(1 + \frac{\varepsilon}{2})\log 2}{\log\log n},$$

missä ε on positiivinen kiinteä reaaliluku, joka voidaan valita niin pieneksi kuin halutaan. Tämän sijoituksen jälkeen voimmekin olla toistaiseksi ihan tyytyväisiä oikean puolen ensimmäiseen termiin. Toinen termi sen sijaan vaatii vielä työtä. Jaetaan toinen termi osiin (tulontekijöihin) ja käsitellään nämä osat erikseen. Ensimmäiseksi työvuorossa on $2^{1/\delta}$:

$$2^{1/\delta} = 2^{\frac{\log\log n}{\left(1+\frac{\varepsilon}{2}\right)\log 2}} = e^{\frac{\log\log n \log 2}{\left(1+\frac{\varepsilon}{2}\right)\log 2}} = e^{\frac{\log\log n}{\left(1+\frac{\varepsilon}{2}\right)}} = (\log n)^{\frac{1}{1+\varepsilon}}.$$

Tähän arvioon voimme hetken olla aivan tyytyväisiä, ja siirtyä toiseen palikkaan. Kun n on riittävän suuri, pätee varmasti $\log 2 > \delta$, joten $\frac{1}{\delta \log 2} < \frac{1}{\delta^2}$ ja tästä saadaankin

$$\log\left(\frac{1}{\delta\log 2}\right) < 2\log\frac{1}{\delta}.$$

Tätä lauseketta onkin jo varsin helppo muokata:

$$2\log\frac{1}{\delta} = 2\log\left(\frac{\log\log n}{\left(1+\frac{\varepsilon}{2}\right)\log 2}\right) < 2\log\left(\frac{\log\log n}{\log 2}\right) = 2\log\log\log n - 2\log\log 2.$$

(Tätä lauseketta katsoessa on varmaan helppo arvata miksi vitsaillaan, että hukkuvasta analyyttisestä lukuteoreetikosta lähtevä ääni on logloglog. Metodologisesti tämä vääntö ei tosin ihan analyyttistä lukuteoriaa ole, mutta samankaltaisista ongelmista ja tuloksista on kuitenkin kyse.) Koska $\log 2 < 1$, niin $\log\log 2 < 0$. Arvioidaan nyt ylöspäin: $-\log\log 2 < \log\log\log n$ riittävän suurille n. Päädytään siis tulokseen

$$\log\left(\frac{1}{\delta\log 2}\right) < 3\log\log\log n.$$

Seuraavaksi väitetään, että

$$3\,(\log n)^{\frac{1}{1+\varepsilon}}\log\log\log n \le \frac{\varepsilon\log 2\log n}{2\log\log n}$$

millä tahansa kiinteällä positiivisella ε. Tämän todistus (sopivan funktion käytöksen, eli käytännössä derivaatan tarkastelu) jätetään epäluuloisille lukijoille harjoitustehtäväksi. Olemme näin vihdoinkin päässeet tulokseen

$$\log d(n) < \frac{(1+\varepsilon)\log 2\log n}{\log\log n}.$$

Tämän tuloksen voi kirjoittaa useassakin muodossa, mutta ehkä tätä on helpoin verrata aikaisempaan neliöjuurirajaan, jos yksinkertaisesti kirjoitetaan

$$d(n) < n^{\frac{(1+\varepsilon)\log 2}{\log\log n}}.$$

Tämä arvio on itse asiassa todella hyvä, sillä jos eksponentin osoittajan summattavan ε eteen laittaisi miinusmerkin plusmerkin tilalle, niin tulos ei enää pätisi vaan löytyisi äärettömän paljon sellaisia lukuja n, joilla

$$d(n) > n^{\frac{(1-\varepsilon)\log 2}{\log\log n}},$$

vaikka ε olisi kuinka pieni kiinteä positiivinen luku. Tämän väitteen todistaminen on varsin helppoa, mutta tarvitaan hieman taustatietoja. Määritellään aluksi funktio $\pi(x)$:

$$\pi(x) = \sum_{p\le x} 1,$$

missä p on alkuluku. Funktio $\pi(x)$ siis laskee korkeintaan luvun x suuruiset alkuluvut. Esimerkiksi $\pi(6) = 3$. Alkulukulauseena tunnettu tulos kertoo, että

$$\pi(x) \sim \frac{x}{\log x},$$

missä merkintä $\sim$ tarkoittaa, että tulos ei ole välttämättä tarkka, mutta virhetermit ovat selvästi pienempiä kuin päätermi (joka on siis kirjoitettu oikealle puolelle). Toisaalta voimme myös määritellä toisenlaisen alkulukujen laskemiseen tarkoitetun funktion:

$$\vartheta(x) = \sum_{p\le x}\log p.$$

Tälle puolestaan pätee arvio $\vartheta(x) \ge cx$ jollakin vakiolla $0 < c \le 1$. Nyt pääsemmekin asiaan. Tarkastellaan lukua $n = 2\cdot 3\cdot 5\cdots q$, joka on siis tulo alkuluvuista johonkin alkulukuun q saakka. Nyt

$$\log n = \log 2 + \log 3 + \log 5 + \cdots + \log q \le \pi(q)\log q.$$

Nyt

$$d(n) = 2^{\pi(q)} \ge 2^{\log n/\log q}.$$

Huomataan toisaalta, että $\vartheta(q) = \log n$, joten voidaan arvioida $\log n \ge cq$, josta saadaan $\log q \le \log\log n - \log c$. Saammekin nyt

$$d(n) \ge 2^{\frac{\log n}{\log\log n - \log c}} = n^{\frac{\log 2}{\log\log n - \log c}}.$$

Seuraavaksi pitääkin vain osoittaa, että

$$n^{\frac{(1-\varepsilon)\log 2}{\log\log n}} < n^{\frac{\log 2}{\log\log n - \log c}}.$$

Tätä lauseketta voi jonkin verran yksinkertaistaa, eli itse asiassa riittää osoittaa

$$\frac{(1-\varepsilon)}{\log\log n} < \frac{1}{\log\log n - \log c},$$

mutta ristiin kertomalla voidaan tämä todeta paikkansa pitäväksi kunhan luku n on riittävän suuri. Siispä, luvuilla $n = 2\cdot 3\cdot 5\cdots q$ on liikaa alkutekijöitä, jotta luvun ε eteen voisi laittaa miinusmerkin.

Tätä tekstiä kirjoitettaessa on käytetty lähteenä K. Ramachandran teosta Theory of Numbers.