

Epästandardi analyysi, mitä se on?

Jukka Tuomela

Itä-Suomen yliopisto, Joensuu
jukka.tuomela@uef.fi

Differentiaalilaskennan alkuvaiheet

Ensimmäinen differentiaalilaskentaa perusteellisesti käsittelevä kirja julkaistiin 1696 [2]. Täältä ovat peräisin monet käytännöt, joita vieläkin noudatetaan: esimerkiksi vakiot valitaan aakkosten alkupäästä, a , b , kun taas muuttujat otetaan aakkosten loppupäästä, x , y . Lisäksi kirjaimella d on oma erikoisasemansa, joka on säilynyt nykypäiviin saakka, vaikka sen merkitys onkin muuttunut matkan varrella.

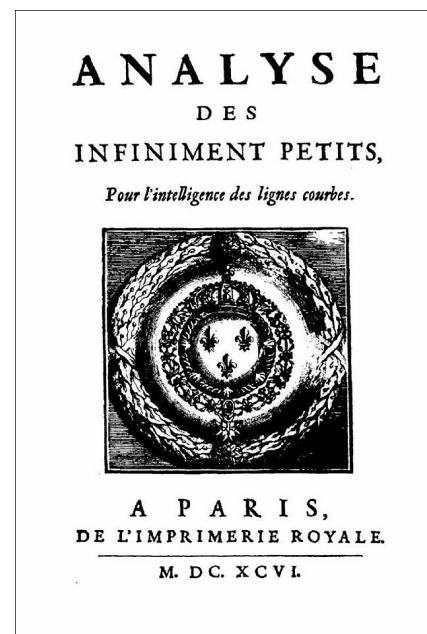
Erikoisesti kirjassa ei kerrota, kuka sen on kirjoittanut (kirjan kansi on kuvassa 1). Kirjoittajana on pidetty L'Hospitalia, mutta sittemmin on osoittautunut, että se perustuu suurelta osin Johann Bernoullin käsi-kirjoitukseen [15]. Kirjan esipuhe on kirjoitettu minämuodossa, ja se varmaankin on L'Hospitalin kirjoittama, koska siinä hän oleellisesti myöntää, että tulokset perustuvat Johannin ja tämän isoveljen Jacobin töihin.

Je reconnois devoir beaucoup aux lumières de M^{rs} Bernoulli, sur tout à celles du jeune presentement Professeur à Groningue. Je me suis servi sans façon de leurs découvertes & de celles de M. Leibnis. C'est-pourquoy je consens qu'ils en revendiquent tout ce qu'il leur plaira, me contentant de ce qu'ils voudront bien me laisser.

Olen suuresti velkaa Bernoullien [Jacob & Johann] loistaville ideoille, etenkin nuoremmalle [Johann], joka on nykyään professorina Groningenissa. Olen vapaasti käyttänyt heidän ja Leibnizin tuloksia. Sik-

si hyväksyn, jos he haluavat pitää esitettyjä asioita ominaan, ja tyydyn siihen mitä he haluavat jättää minulle.

L'Hospitalin sääntö löytyy sivulta 145. Kirjan [2] nykyisessä painoksessa on myös Varignonin kirjoittama teksti vuodelta 1725, joka on tarkoitettu täydentämään ja selventämään alkuperäistä teosta.



Kuva 1: Johann Bernoullin ja L'Hospitalin kirjan kansi.

Differentiaalilaskenta aloitti aivan uuden kauden matematiikassa, ja koko 1700-luvun ajan sen avulla saatiinkin valtavasti uusia tuloksia. Kuitenkin koko jutun taustalla oli käsitteellisiä vaikeuksia: erityisesti miten pitäisi ymmärtää äärettömän pieni luku eli infinitesimaali (infiniment petit, infinitely small)? Tuo on terminäkin outo, koska äärettömän avulla yritetään kuvaila jotain hyvin pientä.

1800-luvulla sitten differentiaalilaskenta muotoiltiin aivan uudella tavalla ε - δ -tekniikan avulla, missä infinitesimaaleja ei edes mainita. Kuitenkin matemaatikkoja jäi mietityttämään ajatus siitä, miten infinitesimaalit pitäisi ymmärtää. Tulokset, joita niitten avulla oli saatu, olivat tavallaan ”liian hyviä” siinä mielessä, että tuntui mahdottomalta, että kaikki olisikin ollut vain tyhjän päällä.

Sitten Robinson selvensi infinitesimaalien käsitettä 1960-luvulla matemaattisen logiikan avulla ja näin saadusta systeemistä ruvettiin käyttämään nimitystä epästandardi analyysi [12]. Hiukan myöhemmin Nelson esitteli oman versionsa epästandardista analyysistä [10]. Nelsonin systeemi on kätevämpi, jos nimenomaan halutaan tehdä differentiaalilaskentaa, eikä oleteta teorian käyttäjältä pitkälle meneviä logiikan esitietoja. Tässä kirjoituksessa noudatetaan siis Nelsonin käytäntöjä, joihin voi tarkemmin perehtyä teosten [5, 4, 9] avulla. Epästandardin analyysin historiaa ja merkitystä on myös mielenkiintoisesti pohdittu kirjassa [3].

Mutta siirrytään sitten itse asiaan: miten Nelsonin systeemillä lasketaan?

Nelsonin systeemi

Epästandardin analyysin lähtökohta on laajentaa reaalilukujen joukkoa; tätä laajennusta merkitään edelleen symbolilla $\mathbb{R}$ ja tämän joukon alkioita sanotaan edelleen reaaliluvuiksi. ”Vanhoja” reaalilukuja sanotaan tavallisiksi reaaliluvuiksi, ja näitten lukujen kokoelmaa voidaan merkitä symbolilla $\mathbb{R}_{\text{tav}}$. Lukija kenties ihmettelee, että eikö tälle laajennetulle joukolle ja sen alueille pitäisi antaa uusi nimi, eikä nimetä perinteiset reaaliluvut uudelleen. Tietysti näinkin voisi menetellä, mutta tämä käytäntö on osoittautunut käteväksi. Syy tähän on seuraava tulos.

”Metalause”. *Se, mikä on ennenkin ollut totta, on edelleen totta.*

Toisin sanoen mikä tahansa lause, joka on todistettu perinteisessä matematiikassa, on automaattisesti totta tässä uudessa systeemissä. Esimerkiksi siis tässä uudessa systeemissä reaaliluvut muodostaa kunnan, algebran peruslause on voimassa jne. Huomaa, että käsitettä ”tavallinen reaaliluku” ja siitä johdettuja käsitteitä ei ole olemassa perinteisessä matematiikassa.

Koska reaalilukujen joukkoa laajennettiin, niin siellä on uudenlaisia alkioita, joille on syytä antaa nimi.

Määritelmä 1. (i) *Reaaliluku x on tosi-iso, jos $|x| > a$ kaikilla tavallisilla reaaliluvuilla a .*

(ii) *Reaaliluku x on tosipieni, jos $|x| < a$ kaikilla tavallisilla positiivisilla reaaliluvuilla a .*

(iii) *Reaaliluku x on rajoitettu, jos se ei ole tosi-iso.*

(iv) *Reaaliluku x on kohtuullinen, jos se ei ole tosi-iso eikä tosipieni.* ♦

Luonnollisesti tavalliset reaaliluvut eivät voi olla tosi-isoja tai tosipieniä. Vastaavasti positiivinen kokonaisluku on tosi-iso, jos se on suurempi kuin kaikki tavalliset kokonaisluvut. Huomaa erityisesti, että tosi-isoit luvut ovat äärellisiä, vaikka ne ovatkin hyvin isoja. Siis jos N on tosi-iso kokonaisluku, niin joukko

$$S = \{1, 2, 3, \dots, N-1, N\}$$

on äärellinen. Tämä onkin eräs epästandardin analyysin mielenkiintoinen piirre: riittää usein tarkastella äärellisiä joukkoja, koska ne ovat jo tarpeeksi isoja. Välttään siis äärettömiin joukkoihin liittyviltä vaikeuksilta. Ennen esimerkkiä tarvitaan kuitenkin vielä seuraava

Määritelmä 2. 1. *x ja y ovat ekvivalentteja, merkitään $x \simeq y$, jos $x - y$ on tosipieni.*

2. *x ja y ovat asympotoottisia, merkitään $x \sim y$, jos $x/y \simeq 1$.*

3. *Jos $x \simeq y$ ja y on tavallinen, niin sanotaan, että y on x :n standardiosa, merkitään $y = \text{st}(x)$.* ♦

Esimerkiksi jos x on tosipieni, niin $x \simeq 0$, joten $\text{st}(x) = 0$. Jos $f = \sum_{k=0}^n c_k x^k$ on polynomi, niin $f \sim c_0$, jos $x \simeq 0$. Nyt voidaan osoittaa [5, s. 36]:

Lause 1. *Jos x on rajoitettu, niin sen standardiosa on yksikäsitteinen.* ★

Lauseesta seuraa, että jos x ja y ovat rajoitettuja, niin

$$x \simeq y \implies \text{st}(x) = \text{st}(y). \quad (1)$$

Seuraavasta tuloksesta saa jonkinlaisen mielikuvan miten isoja äärelliset luvut voivat olla.

Lause 2. *Välillä $[0, 1]$ on vain äärellinen määrä tavallisia reaalilukuja.*

Todistus. Olkoon n jokin tosi-iso kokonaisluku, jolloin siis $h = 1/n$ on tosipieni. Jaetaan väli $[0, 1]$ osiin ja merkitään $I_k = [(k-1)h, kh]$, $k = 1, \dots, n$. Olkoon $x, y \in I_k$; tällöin $|x - y| \leq h$, joten $x \simeq y$ ja siis äskeisen lauseen perusteella $\text{st}(x) = \text{st}(y)$. Siis välillä I_k voi olla korkeintaan yksi tavallinen reaaliluku, joten yhteensä niitä on korkeintaan n kappaletta. ♦

Luonnollisesti klassisen matematiikan kannalta välillä $[0, 1]$ on ylinumeroituvasti ääretön määrä reaalitykijöitä. Tässä näyttäisi olevan ristiriita: miten äärellisen joukon osajoukko voisi olla (ylinumeroituvasti) ääretön? Ongelma ratkeaa sillä, että $\mathbb{R}_{\text{tav}}$ ei ole joukko.

Nelson itse sovelsi tätä ideaa todennäköisyyslaskentaan [11]. Kirjan nimessä sanat *radically elementary* viittaa juuri siihen, että koska äärelliset joukot voivat olla niin isoja, niin hyvin monet todennäköisyyslaskennan tulokset voidaan muotoilla äärellisten joukkojen avulla.

Infinitesimaalin käsite 1700-luvulla

Infinitesimaalin eli tosipiienen luvun selittäminen 1700-luvun termein oli kuitenkin haastavaa. Kirja [2] perustui kahteen perusoletukseen, joista ensimmäinen on seuraava.

- (O1) On demande qu'on puisse prendre indifféremment l'une pour l'autre deux quantités qui ne diffèrent entr'elles que d'une quantité infiniment petite.

Kaksi suuretta voidaan samaistaa, jos niiden erotus on tosipieni.

Oletuksessa (O1) "quantité" on käännetty suureeksi, mutta tämä voidaan myös tulkita reaalityluvaksi, vaikka tietysti tuolloin käsitettä reaalityluku ei ollut olemassa. Tämän oletuksen voisi tulkita määritelmän 2 kohdaksi 1. Varignon yritti seuraavasti [2, Varignon, s. 13]:

L'on ne doit pas dire que $x + dx$ soit $= x$ [...] mais que $x + dx$, est une abscisse qui a dans ce lieu la même propriété que x [...] de sorte que l'on pourra faire de $x + dx$ la même chose que de x .

Ei pidä sanoa, että $x + dx = x$, vaan että arvolla $x + dx$ on samat ominaisuudet kuin arvolla x , joten sitä voi käyttää samaan tapaan.

Nykytermein tämän voisi tulkita niin, että $x + dx \simeq x$, jos x on kohtuullinen ja dx on tosipieni, siis jälleen määritelmän 2 kohta 1.

Bernoullin ja L'Hospitalin kirja, samoin kuin Varignonin täydennys, oli tarkoitettu matemaatikoille: tähän oli silloista huippututkimusta, jossa esiteltiin alan uusimpia tuloksia. Kun sitten differentiaalilaskennan käyttö koko ajan laajeni, niin sitä piti ruveta opettamaan, ja tuli tarve varsinaisille oppikirjoille. Euler kirjoitti tällaisen vuonna 1755 [6]. Eulerin kirjaa voi jo pitää tyyliään ja merkinnöitään varsin modernina: algebrallinen esitystapa on paikoin hyvinkin tutunnäköistä, kun taas Bernoulli ja L'Hospital esittivät asiat hyvin geometrisesti, jolloin tyyliä oltiin lähellä Eukleideen geometriaa. 50 vuodessa oli siis tapahtunut valtava muutos matemaattisen tekstin kirjoittamisessa.

Euler selitti tosipiienen lukujen olemusta näin [6, ss. 51-53]:

If we accept the notation used in the analysis of the infinite, then dx indicates a quantity that is infinitely small, so that both $dx = 0$ and $adx = 0$, where a is any finite quantity. Despite this, the geometric ratio adx/dx is finite, namely $a/1$. For this reason these two infinitely small quantities dx and adx , both being equal to 0, cannot be confused when we consider their ratio.

If 1 or any other finite quantity is divided by something infinitely small or 0, the quotient will be infinitely large, and thus an infinite quantity. Since the symbol ∞ stands for an infinitely large quantity, we have the equation $a/dx = \infty$. The truth of this is clear also when we invert: $a/\infty = dx = 0$.

Luonnollisesti näin kirjoitettuna tätä ei voi hyväksyä. Voisi kuvitella, että 1700-luvun opiskelijaparka on ollut lohduton kirjansa ääressä, tai ainakin vähintään yhtä hämmentynyt kuin moderni lukijakin. Mutta jos käyttää tämän kirjoituksen termejä, niin Eulerin teksti voidaan tulkita täysin ymmärrettävästi:

Jos dx on tosipieni ja a kohtuullinen, niin $dx \simeq adx \simeq 0$ ja $adx/dx = a/1 = a$.

Jos $y = a/dx$, niin y on tosi-iso.

Pitää siis huolellisesti erottaa a ja $\simeq$ sekä ääretön ja tosi-iso (mutta äärelliset) luvut. Lisäksi Euler myöhemmin [6, s. 116] kirjoittaa $\sin(dx) = dx$, vaikka pitäisi olla $\sin(dx) \sim dx$.

Mutta on helppoa uskoa, että tuo Eulerin selitys ei ollut kovin vakuuttava 1700-luvullakaan, joten tosipiienen lukujen syvin olemus jäi edelleen hämärän peittoon. Kuitenkin Eulerin kirjassa on paljon laskuja ja esimerkkejä, ja idea varmaankin oli, että kun näkee, miten homma toimii "käytännössä", niin sen perusteella ymmärtää jotenkin intuitiivisesti, mistä on kysymys. Esimerkiksi Eulerin kirjasta löytyy kaikki alkeisfunktioitten derivointikaavat, jotka tietysti ovat samat kuin nykyäänkin, vaikka perustelut ovatkin erilaiset.

HARJOITUSTEHTÄVÄ 1 Onko Eulerin kirjassa jokin kaava tai tulos, mikä on "oikeasti" väärin? Miksi? *

1700-luvun lopussa Lagrange kuitenkin päätti, että nyt tämä epämääräisyys saa riittää [8]. Kirjan alaotsikko ilmoittaa selkeästi, mistä on kysymys:

Les principes du calcul différentiel, dégagés de toute considération d'infiniment petits, d'évanouissans, de limites et de fluxions, et réduits à l'analyse algébrique des quantités finies.

Differentiaalilaskennan perusteet, josta on raivattu pois äärettömän pienet ja häviävän pienet suuret ja fluksiot, käyttäen vain algebrallista analyysiä ja äärellisiä suureita.

Lagrange olisi voinut kirjoittaa neutraalisti *sans infiniment petits* (ilman infinitesimaaleja), mutta hän kirjoitti *dégagés de toute considération d'infiniment petits*, minkä lukija helposti tulkitsee niin, että infinitesimaalit on tarpeettomina ja ärsyttävinä heitetty roskikseen.

Fluksio oli Newtonin käyttämä termi. Algebrallisella analyysillä Lagrange tarkoitti potenssisarjoja; jos f on analyttinen funktio, niin se voidaan kirjoittaa potenssisarjana, joten

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k.$$

Derivaatat siis määrittyvät potenssisarjan kertoimien avulla. Tästä kirjasta on muuten peräisin sana derivaatta, ja merkintä f' , f'' jne derivaatoille. Lagrange kirjoitti, että f' on *première fonction dérivée*, siis ensimmäinen johdettu funktio [8, s. 19].¹ Funktio f itse oli *fonction primitive* eli alkuperäinen funktio.²

Lagrange'n kirja siis hylkäsi täysin perinteisen infinitesimaalianalyysin, ja hänen kirjaansa ainakin suurelta osin voidaan pitää pätevänä myös nykynäkökulmasta. Luonnollisesti ongelmana oli, että piti rajoittua analyttisiin funktioihin, joten tämän jälkeen kesti vielä pitkään ennen kuin päädyttiin moderniin derivaatan määritelmään.

Olemassaolo?

Skeptinen lukija saattaa ehkä ihmetellä, että onko tosi-pieniä ja tosi-isoja lukuja ”oikeasti” olemassa, vai onko jossain piilossa jokin ristiriita. Yleensäkin matemaattisten olioitten olemassaolo on hankala kysymys, eikä siihen ole yleisesti hyväksyttyä vastausta. Samoin voisi kysyä, onko tavallisia reaalilukuja olemassa, ja mitä tällä olemassaololla tarkoitetaan.

Normaalisti reaaliluvut luodaan joko Dedekindin leikkausten tai Cauchy-jonojen avulla rationaaliluvuista. Molemmissa tapauksissa tarvitaan äärettömiä joukkoja. Äärettömien joukkojen olemassaolo puolestaan on tavallisen joukko-opin aksiooma. Mutta jos asetettaisiin aksioomaksi saunatonttujen olemassaolo, niin luulisin, että aika harva pitäisi tätä todisteena saunatonttujen olemassaolosta. Tosin saunatontut ovat ideana olemassa: monet ihmiset ovat kuulleet niistä, kuten reaaliluvuistakin. Reaalilukujen ja saunatonttujen olemassaolo on siis varsin samankaltaista.

Joka tapauksessa matematiikassa on tapana sivuuttaa olemassaoloon liittyvät ongelmat, ja aksioomasysteemeiltä vaaditaan ”vain”, että ne ovat ristiriidattomia.

Tunnetusti Gödel sitten osoitti, että aksioomasysteemin sisällä ei voida todistaa ristiriidattomuutta. Voidaan kuitenkin vertailla eri systeemejä, ja onkin voinut osoittaa, että jos Nelsonin systeemissä on ristiriita, niin silloin on ristiriita myös tavallisessa joukko-opissa. Tosi-isojen ja tosi-pienten lukujen käyttö on siis yhtä ”turvallista” kuin tavallisten reaalilukujen käyttö.

Geometria

Katsotaan sitten pieni geometrinen esimerkki. Siinä tarvitaan seuraavaa aputulosta.

Lemma 1. *Olkkoon α tosi-pieni; tällöin*

$$\begin{aligned}\sin(\alpha) &\sim \tan(\alpha) \sim \alpha, \\ \cos(\alpha) &\simeq 1.\end{aligned}$$

Todistus. Metalauseen mukaan trigonometriset funktiot ovat analyttisiä, joten tulos saadaan potenssisarjojen avulla. $\diamond$

Määritellään luku π siten, että 2π radiaania on koko kierros.

Lause 3. *Piirretään r -säteisen ympyrän sisään säännöllinen n -kulmio. Olkkoon L monikulmion reunan pituus ja A sen pinta-ala. Jos n on tosi-iso, niin*

$$\text{st}(L) = 2\pi r \quad \text{ja} \quad \text{st}(A) = \pi r^2.$$

Todistus. Monikulmion reuna koostuu janoista, joitten pituus on h , joten reunan pituus on $L = nh$. Edelleen monikulmio koostuu tasasivuisista kolmioista K_h , joitten keskuskulma on $\alpha = 2\pi/n$.

Pienten laskujen jälkeen saadaan

$$L = nh = 2rn \sin(\alpha/2) = \frac{4\pi r \sin(\alpha/2)}{\alpha}.$$

Kun n on tosi-iso, niin α on tosi-pieni, joten yllä olevan Lemman perusteella saadaan

$$L \simeq 2\pi r \quad \Rightarrow \quad \text{st}(L) = 2\pi r.$$

Olkkoon $b = r \cos(\alpha/2)$ kolmion K_h korkeus, joten K_h :n pinta-ala on $A_h = \frac{1}{2}bh$. Tästä saadaan

$$\begin{aligned}A &= nA_h = \frac{nbh}{2} = nr^2 \sin(\alpha/2) \cos(\alpha/2) \\ &= \frac{nr^2 \sin(\alpha)}{2} = \frac{\pi r^2 \sin(\alpha)}{\alpha}.\end{aligned}$$

Jälleen äskeisen Lemman avulla saadaan

$$A \simeq \pi r^2 \quad \Rightarrow \quad \text{st}(A) = \pi r^2. \quad \diamond$$

¹Finanssimarkkinoilla kuitenkin puhutaan johdannaisista eikä derivaatoista, vaikka kyseessä on sama sana.

²Tämän takia ranskaksi määräämätön integraali on primitive, adjektiivi on muutettu substantiiviksi.

Epästandardin analyysin puitteissa voitaisiin siis sanoa, että ympyrä on n -kulmio, missä n on tosi-iso. Mielinkiintoisesti Johann Bernoulli ja L'Hospital ajattelivat täsmälleen näin. Heidän toinen perusoletuksensa oli [2, s. 2–3]:

(O2) On demande qu'une ligne courbe puisse être considérée comme assemblage d'une infinité de lignes droites, chacune infiniment petite.

Käyrä voidaan ajatella koostuvaksi janoista, joitten pituus on tosi-pieni.

Muistetaan, että pituuden ja pinta-alan laskeminen on oikeastaan integrointia, joten äskenen lause voidaan tulkita eräitten integraalien laskemiseksi äärellisten summien avulla. Oletetaan jatkossa useimmiten, että lausekkeissa esiintyvät funktiot ovat *tavallisia*.

Määritelmä 3. *Funktio $f: \mathbb{R} \rightarrow \mathbb{R}$ on tavallinen, jos sen määrittelyssä ei esiinny määritelmässä 1 ja 2 olevia käsitteitä tai näitten johdoksia.* ♦

Olkoon nyt $f: [0, 1] \rightarrow \mathbb{R}$ jokin tavallinen funktio ja asetetaan

$$\int_0^1 f(x) dx = \text{st} \left(\frac{1}{n} \sum_{k=1}^n f(k/n) \right), \quad (2)$$

missä n on tosi-iso. Yhtälön vasenta puolta on tapana sanoa määrätyksi integraaliksi, ja yhtälön oikea puoli kertoo, mitä tuolla hassulla merkinnällä tarkoitetaan. Otetaan esimerkiksi $f(x) = x^3$; tällöin

$$\frac{1}{n} \sum_{k=1}^n f(k/n) = \frac{1}{n^4} \sum_{k=1}^n k^3.$$

Tunnetusti

$$\sum_{k=1}^n k^3 = \frac{n^2(n+1)^2}{4},$$

mikä tulos löytyy Jacob Bernoullin kirjasta vuodelta 1713 [1]. Siispä

$$\int_0^1 x^3 dx = \text{st} \left(\frac{1}{4} + \frac{1}{2n} + \frac{1}{4n^2} \right) = \frac{1}{4}.$$

d , osa 0

Siirrytään sitten lopulta derivointiin, differensseihin ja differentiaaleihin. Jätin tämän viimeiseksi, koska se on niin hankala; ongelmana on, että kirjaimella d on ai- van liian monta merkitystä, kun lasketaan mukaan sekä vanhat että uudet merkitykset. Kuten yllä olevista esimerkeistä on käynyt ilmi, niin jos puhutaan tosi-pienistä luvuista, niin voidaan sanoa esimerkiksi, että jos a kohtuullinen ja h tosi-pieni, niin $a + h \simeq a$. Perinteisesti tosi-pienistä luvuista kuitenkin aina käytettiin merkintää dx ja kirjoitettiin $x + dx$. Miksi d ?

Tarkastellaan tasoa, jossa on koordinaatit (x, y) , ja olkoon $y = f(x)$ jokin käyrä. Muistetaan, että kirja [2] analysoi nimenomaan tasokäyrien ominaisuuksia. Määritellään nyt

$$dy = df = f(x + dx) - f(x).$$

Tässä d tulee sanasta differenssi; siis merkinnässä dx ajatellaan, että lukuun x lisätään jokin tosi-pieni luku dx , mutta merkinnässä df ajatellaan, että d operoi funktioon f .

Ajatuksena on siis, että jos dx on tosi-pieni, niin myös dy on tosi-pieni. Mutta mistä tiedetään, että dy on tosi-pieni? Itse asiassa tässä käytetään funktion f jatkuvuutta.

Lause 4. *Olkoon $f: \mathbb{R} \rightarrow \mathbb{R}$ tavallinen funktio. Seuraavat väitteet ovat ekvivalentteja.*

(i) f on jatkuva pisteessä a .

(ii) $f(a + h) \simeq f(a)$ kaikilla $h \simeq 0$. ★

1700-luvulla jatkuvuutta ei tietenkään oltu määritelty, vaan tavallaan automaattisesti ajateltiin, että ne funktiot, joita tarkastellaan, toteuttavat tuon ehdon. Ehkäpä siis hiukan yllättävästi tämä vastasi täysin nykyistä jatkuvuuden käsitettä.

Olettaen, että f on jatkuva, lasketaan käyrälle $y = f(x)$ tangentti oletuksen (O2) avulla [2, s. 11]. Tarkastellaan pisteitä (a, b) ja $(a + da, b + db)$. Suora, joka kulkee näitten pisteitten kautta, on

$$y - b = \frac{db}{da} (x - a). \quad (3)$$

Jos nyt $b = f(a)$, niin moderni lukijakin tunnistaa heti, että saatiin oikea vastaus, koska vieläkin saatetaan kirjoittaa $db/da = f'(a)$. Mutta nyt siis

$$\frac{db}{da} = \frac{f(a + da) - f(a)}{da}.$$

Koska oletettiin, että f on jatkuva, niin oikealla puolella on kahden tosi-pienen luvun suhde. Mutta tämän suhteen ei pitäisi oleellisesti riippua valitusta da :sta, jotta derivaatta olisi hyvin määritelty. Voidaan osoittaa:

Lause 5. *Olkoon $f: \mathbb{R} \rightarrow \mathbb{R}$ tavallinen funktio. Seuraavat väitteet ovat ekvivalentteja.*

(i) f on derivoituva pisteessä a .

(ii) On olemassa jokin tavallinen reaaliluku q siten, että

$$\text{st} \left(\frac{f(a + h) - f(a)}{h} \right) = q$$

kaikilla $h \simeq 0$.

Jos väite (ii) pätee, niin kirjoitetaan $f'(a) = q$. ★

Olkoon esimerkiksi $y = f(x) = x^3$, jolloin

$$\frac{f(a+h) - f(a)}{h} = \frac{(a+h)^3 - a^3}{h} = 3a^2 + 3ah + h^2.$$

Siispä

$$f'(a) = \text{st}\left(3a^2 + 3ah + h^2\right) = 3a^2,$$

jos $h \simeq 0$. Jos siis oletetaan, että f on derivoituva, niin yleisesti ottaen $\text{st}(db/da) = f'(a)$, joten yhtälöstä (3) saadaan

$$y - b \simeq f'(a)(x - a)$$

ja tangentiksi saatiin (ainakin melkein) se mitä pitääkin. Lagrange kommentoi tätä laskua näin [8, s. 3]:

En regardant une courbe comme un polygone d'un nombre infini de côtés chacun infiniment petit, et dont le prolongement est la tangente de la courbe, il est clair qu'on fait une supposition erronée; mais l'erreur se trouve corrigée dans le calcul par l'omission qu'on y fait des quantités infiniment petites.

Jos ajattelee, että käyrä koostuu tosipienistä janoista, joitten jatkeena saadaan tangentit, niin on selvää, että tämä oletus on virheellinen; mutta virhe korjaantuu, kun laskun kuluessa jätetään huomiotta tosipieniä suureita.

Lagrange on tässä aivan oikeassa: äsken ”tarpeettomista” tosipienistä suureista päästiin eroon ottamalla standardiosa. Tietenkin ennen vanhaan käsitettä standardiosa ei ollut, ja huolettomasti jätettiin sopivia termejä huomiotta. Mysteeriksi jäi, että jos lähdetään liikkeelle ”vääristä” oletuksista, niin miten ja miksi aina kuitenkin virheet korjaantuvat? Vai korjaantuvatko ne aina? Kuten Lagrange sanoo:

Il serait peut-être difficile de donner une démonstration générale.

Olisi ehkä vaikeaa todistaa tämä yleisesti.

Ei siis ihme, että Lagrange kirjansa johdannossa otti esille, että differentiaalilaskentaa on hyvin vaikeaa opettaa aloittelijoille (commençans).

Lasketaan vielä sinin ja arkussinin derivaatta Eulerin tapaan. Olkoon $y = \sin(x)$, jolloin

$$\begin{aligned} \frac{dy}{dx} &= \frac{\sin(x+dx) - \sin(x)}{dx} \\ &= \frac{\sin(x)\cos(dx) + \cos(x)\sin(dx) - \sin(x)}{dx} \\ &\simeq \frac{\sin(x) + \cos(x)\sin(dx) - \sin(x)}{dx} \simeq \cos(x), \end{aligned}$$

missä $\simeq$ seuraa Lemmasta 1, kun dx on tosipieni. Euler laski juuri näin, vaikka tietysti hän käytti yhtäsuuruusmerkkiä, eikä symbolia $\simeq$. Samoin hän vetosi potenssisarjoihin, eli käytännössä Lemmaan 1, kun hän perusteli siirtymistä toiselta riviltä kolmannelle yllä olevassa kaavassa.

Jos $y = \sin(x)$, niin $x = \arcsin(y)$, joten käyttämällä sitä, että $dy/dx \simeq \cos(x)$, saadaan

$$\begin{aligned} \frac{dx}{dy} &\simeq \frac{1}{\cos(x)} \\ &= \frac{1}{\sqrt{1 - \sin^2(x)}} = \frac{1}{\sqrt{1 - y^2}}. \end{aligned}$$

Yleisesti ottaen Eulerin laskut antavat tuttuja tuloksia, vaikka ne nykyään perustellaan toisin. Mutta epästandardin analyysin puitteissa Eulerin laskut (ainakin useimmiten) ovat täysin oikein, jos yhtäsuuruusmerkin tilalle tietyissä kohdissa laitetaan joko $\simeq$ tai $\sim$, ja lisäksi sopivissa paikoissa otetaan lausekkeesta standardiosa.

HARJOITUSTEHTÄVÄ 2 Valitse muutama Eulerin lasku ja laita yhtäsuuruusmerkin tilalle oikeissa kohdissa joko $\simeq$ tai $\sim$. *

d , osa 1

Äsken jo nähtiin, että d voidaan tulkita operaattoriksi: jos f on funktio, niin

$$df = f(x+h) - f(x),$$

missä $h \simeq 0$. Selvästi tämä on lineaarinen; jos g on jokin toinen funktio, niin

$$d(af + bg) = a df + b dg,$$

missä a ja b ovat vakioita. Entä tulo? Nyt voidaan laskea

$$\begin{aligned} d(fg) &= f(x+h)g(x+h) - f(x)g(x) \\ &= f(x+h)g(x+h) - f(x+h)g(x) + \\ &\quad f(x+h)g(x) - f(x)g(x) \\ &= f(x+h)(g(x+h) - g(x)) + \\ &\quad (f(x+h) - f(x))g(x) \\ &\sim f dg + g df. \end{aligned}$$

Kun nyt merkitään $u = fg$, niin Lauseen 5 perusteella saadaan, että

$$u' = f g' + g f',$$

jos f ja g ovat derivoituvia, joten päädyttiin tuttuun tulon derivointisääntöön, jota myös Leibnizin säännökksi sanotaan. Vastaavasti voidaan sitten määritellä osittaisderivaattoja, jos f on monen muuttujan funktio.

HARJOITUSTEHTÄVÄ 3 Olkoon f ja g jatkuvia funktioita. Osoita, että

$$d\left(\frac{f}{g}\right) \sim \frac{g df - f dg}{g^2}.$$

*

d , osa 2

Jotta sekaannus olisi mahdollisimman suuri, niin modernissa differentiaaligeometriassa käytetään samoja symboleja (d , dx , df jne) kuin 1700-luvulla, niille on vain annettu uusi merkitys. Differentiaaligeometriassa käytettiin vanhaa terminologiaa (tosipieni jne) paljon kauemmin kuin analyysin puolella. Syy oli se, että ongelmana oli tiettyjen käsitteitten määrittely, eikä tämä ratkennut sillä, että derivaatat määriteltiin ε - δ -menetelmällä. Spivak selittää asian näin [14, s. 111]:

Once this realization came, it was only a matter of making new definitions, which preserved the *old* notation, and waiting for everybody to catch up. In short, all classical notions involving infinitely small quantities became functions on tangent vectors, like df , except for quotients of infinitely small quantities, which became tangent vectors, like dc/dt .

”This realization” tarkoittaa sitä, että ymmärrettiin, miten tietyt käsitteet on syytä määritellä. Spivak ei kerro, kuinka pitkään piti odottaa, mutta ehkäpä yhden matemaatikkosukupolven piti vaihtua ennen kuin kaikki olivat tilanteen tasalla (catch up).

Jos siis $g: [0, 1] \rightarrow \mathbb{R}$ ja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ on (x, y) tasossa määritelty funktio, niin nykyään voidaan kirjoittaa esimerkiksi

$$\int_0^1 g(x) dx, \\ df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Nyt symbolit dx , dy ja df ovat kovektoreita, siis lineaarikuvauksia, jotka operoivat (tangentti)vektoreihin. Kovektorit ovat toiselta nimeltä differentiaali ykkös-muotoja (differential one form), ja ylipäättään differentiaalimuodot ovat objekteja, joita voi integroida. Lisäksi kuvaus $d: f \rightarrow df$ on ulkoderivaatta (exterior derivative). Joka tapauksessa tämä on johtanut siihen, että opetuksessa käytetään merkintöjä, joitten merkitystä ei oikeasti selitetä; differentiaaligeometriaahan opetetaan vasta yliopiston jatkokursseilla. Minulle on epäselvää miten lukion opettajat nykyään selittävät symbolia dx integroinnin yhteydessä, tai onko ylipäättään olemassa jotain ”virallista” selitystä tähän käytäntöön. Voidaan siis todeta Lagrangen tapaan, että differentiaalilaskennan opettaminen on edelleen vaikeaa.

Differentiaaligeometriaan voi perehtyä esimerkiksi kirjan [7] avulla. Kirjan otsikko on ehkä hämäävästi *Vector analysis*, mutta tämä viittaa siihen, että lähdetään

liikkeelle siitä, että huolellisesti määritellään, mitä ovat (tangentti)vektorit, ko(tangentti)vektorit, ja mitä eroa näillä on.

Mekaniikassa saatetaan edelleen käyttää vanhoja termejä; esimerkiksi voidaan edelleen puhua infinitesimaalisesta siirtymästä (infinitesimal displacement), vaikka kaiketi nykyään enemmän kuitenkin käytetään termiä virtuaalinen siirtymä (virtual displacement). Joka tapauksessa tällä ei ole mitään tekemistä epästandardin analyysin tai infinitesimaalien kanssa, vaan kyseessä on konfiguraatioavaruuden tangenttivektori. Näille vektoreille saatetaan myös käyttää merkintöjä dx tai δx , mikä tietysti sekoittaa asioita lisää.

Askelfunktio

Edellä on ollut puhetta tavallisista funktioista, joten annetaan vielä esimerkki funktiosta, joka ei ole tavallinen. Olkoon

$$H(x) = \frac{1}{\pi} \arctan(x/h) + \frac{1}{2},$$

missä h on tosipieni ja positiivinen. Jos x on kohtuullinen ja positiivinen, niin $H(x) \simeq 1$, ja jos x on kohtuullinen ja negatiivinen, niin $H(x) \simeq 0$. Metalauseen mukaan H on analyyttinen; voitaisiin siis sanoa, että H on ”analyyttinen askelfunktio”. Koska H on analyyttinen, niin kaikki derivaatat ovat olemassa, ja on tapana sanoa, että $H' = \delta$ on Diracin delta. Tämä avaa uudenlaisia mahdollisuuksia tutkia funktioita, jotka eivät tavallisessa mielessä ole derivoituvia. Perinteinen tapa käsitellä tätä on Schwartzin distribuutioteoria [13]. Huomaa esimerkiksi, että δ^2 on hyvin määritelty, koska δ on analyyttinen funktio, kun taas δ^2 ei ole määritelty distribuutioteoriassa.

Lopuksi

Kirjan [2] esipuheessa L’Hospital tarkastelee matemaatiikan historiaa; ajanjaksoa antiikin ajasta Descartesiin hän arvioi näin:

On ne sçauroit assés s’étonner [...] que par une admiration presque superstitieuse pour leur ouvrages, ils se soient contentés de les lire & de les commenter, [...] sans oser commettre le crime de penser quelques fois par eux-mêmes. [...] Les livres se multiplioient, & cependant rien n’avançoit.

Ei voi kuin ihmetellä [...] miten lähes taikauskoinen ihailu antiikin matemaatikoita kohtaan aiheutti sen, että vanhoja töitä luettiin ja kommentoitiin, [...] mutta kukaan ei uskaltanut tehdä sellaista rikosta, että olisi edes joskus ajatellut ihan itsenäisesti. [...] Kirjoja kirjoitettiin, mutta mikään ei edistynyt.

Kuten tästä lainauksesta voi päätellä, niin L'Hospital oli täysin vakuuttunut siitä, että lopulta yli 1500 vuoden pysähtyneisyyden ajan jälkeen differentiaalilaskenta aiheuttaa suuren mullistuksen matematiikassa. Täsähän hän oli aivan oikeassa. Esipuhe päättyy seuraavasti:

Les deux suppositions que j'ai faites au commencement de ce Traité [...] me paroissent si évidentes, que je ne croy pas qu'elles puissent laisser aucun doute dans l'esprit de Lecteurs attentifs. Je les aurois même pû démontrer facilement à la manière des Anciens.

Kaksi oletusta [oletukset (O1) ja (O2)], joihin tämä teos perustuu, ovat mielestäni niin selkeitä, etten usko, että tarkkaavaiselle lukijalle jää mitään epäilyksiä niitten suhteen. Olisin voinut jopa helposti todistaa ne vanhaan tapaan.

Tässä taas L'Hospital on niin väärässä, että koko kapale kuulostaa vitsiltä. Lisäksi on täysin epäselvää, mitä L'Hospital tarkoitti *todistamisella vanhaan tapaan* tässä yhteydessä.

Epästandardia analyysiä ei käytetä kovin laajasti, vaan se on jäänyt varsin pienen piirin harrastukseksi. Kuitenkin voisi sanoa, että esimerkiksi lauseen 4 jatkuvuuden karakterisointi on intuitiivisesti selkeämpi kuin perinteinen ε - δ -määritelmä. Periaatteessa epästandardia analyysiä voisi siis käyttää myös differentiaalilaskennan perusopetuksessa; näin ei ole käynyt, mutta ehkäpä joskus tulevaisuudessa tilanne voi muuttua.

Viitteet

- [1] Jacob Bernoulli, *The art of conjecturing*, Johns Hopkins University Press, 2006, Translation of *Ars conjectandi*, 1713, with an introduction and notes by Edith Dudley Sylla.
- [2] Johann Bernoulli and G. L'Hospital, *Analyse des infiniment petits pour l'intelligence des lignes courbes*, ACL-éditions, Paris, 1988, réimpression de l'édition originale de 1696, suivi des *Eclaircissements sur l'analyse des infiniment petits* par P. Varignon, 1725.
- [3] P. Davis, R. Hersh, and E. Marchisotto, *The mathematical experience, study edition*, Modern Birkhäuser Classics, Birkhäuser/Springer, New York, 2012.
- [4] F. Diener and M. Diener, *Nonstandard analysis in practice*, Springer Science & Business Media, 2012.
- [5] F. Diener and G. Reeb, *Analyse non standard*, Collection Enseignement des Sciences, vol. 40, Hermann, Paris, 1989.
- [6] L. Euler, *Foundations of differential calculus*, Springer-Verlag, New York, 2000, kirjan Institutiones calculi differentialis, 1755, 1. osan käännös.
- [7] K. Jänich, *Vector analysis*, Undergraduate Texts in Mathematics, Springer, 2001.
- [8] J.-L. Lagrange, *Théorie des fonctions analytiques: contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits, d'évanouissans, de limites et de fluxions, et réduits à l'analyse algébrique des quantités finies*, Imprimerie de la République, Paris, 1797.
- [9] C. Lobry, *Et pourtant... ils ne remplissent pas $\mathbb{N}$* , Aléas, 1989.
- [10] E. Nelson, *Internal set theory: a new approach to nonstandard analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 6, 1165–1198.
- [11] ———, *Radically elementary probability theory*, Annals of Mathematics Studies, vol. 117, Princeton University Press, 1987.
- [12] A. Robinson, *Non-standard analysis*, Princeton Landmarks in Mathematics, Princeton University Press, 1996, Reprint of the second (1974) edition, With a foreword by W. A. J. Luxemburg.
- [13] L. Schwartz, *Théorie des distributions*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, vol. IX-X, Hermann, Paris, 1966, Nouvelle édition, entièrement corrigée, refondue et augmentée.
- [14] M. Spivak, *A comprehensive introduction to differential geometry. Vol. I*, 3rd ed., Publish or Perish, 1999.
- [15] C. Truesdell, *The new Bernoulli edition*, Isis **49** (1958), 54–62.