

# Optimointi ja mallinnus taloustieteissä

Veli-Matti Hokkanen

11. helmikuuta 2002



# Sisältö

|                                                                    |           |
|--------------------------------------------------------------------|-----------|
| <b>0 Johdanto</b>                                                  | <b>v</b>  |
| <b>1 Yleistä tieteenfilosofiaa</b>                                 | <b>1</b>  |
| 1.1 Teoreettisen fysiikan menestystarina . . . . .                 | 1         |
| 1.2 Selittäminen . . . . .                                         | 2         |
| 1.3 Ihmiset järjestömiä luontokappaleita? . . . . .                | 3         |
| 1.3.1 Teleologia ja variaatioperiaatteet . . . . .                 | 3         |
| 1.4 Ymmärtäminen . . . . .                                         | 4         |
| 1.5 Selittäminen ja ymmärtäminen . . . . .                         | 5         |
| 1.5.1 Poliittisesta taloustieteestä . . . . .                      | 6         |
| 1.6 Harjoitustehtäviä . . . . .                                    | 7         |
| <b>2 Erityistä tieteenfilosofiaa ja -metodologiaa</b>              | <b>9</b>  |
| 2.1 Aksiomaattinen tieteenihanne . . . . .                         | 9         |
| 2.1.1 Fyysikot ja filosofia . . . . .                              | 9         |
| 2.1.2 Fysiikka ja filosofia . . . . .                              | 10        |
| 2.1.3 Fysiikan maailmankuva . . . . .                              | 10        |
| 2.2 Tieteellisen teorian rakenne . . . . .                         | 11        |
| 2.2.1 Miksi fysiikan teorian on aksiomatisoitava? . . . . .        | 11        |
| 2.2.2 Esimerkki aksiomatisoidusta teoriasta: virtapiirit . . . . . | 11        |
| 2.3 Mallit ja mallinnus . . . . .                                  | 13        |
| 2.3.1 Esimerkki: laskuvarjohyppy . . . . .                         | 13        |
| 2.3.2 Esimerkki: loisiminen . . . . .                              | 13        |
| 2.4 Dimensioanalyysiä . . . . .                                    | 15        |
| 2.4.1 Esimerkki: yksinkertainen heiluri . . . . .                  | 15        |
| 2.4.2 Suureet ja mittayksiköt . . . . .                            | 15        |
| 2.4.3 Dimensiokuvaus . . . . .                                     | 16        |
| 2.4.4 Dimensiohomogeenisuus . . . . .                              | 17        |
| 2.4.5 Buckinghamin $\pi$ -teoreema . . . . .                       | 18        |
| 2.4.6 Esimerkkejä . . . . .                                        | 19        |
| 2.4.7 Huomioita dimensioanalyysistä . . . . .                      | 20        |
| 2.5 Aika, avaruus, järjestelmä, tila, kenttä . . . . .             | 21        |
| 2.5.1 Aika ja avaruus . . . . .                                    | 21        |
| 2.5.2 Järjestelmä ja tila . . . . .                                | 22        |
| 2.5.3 Kenttä . . . . .                                             | 22        |
| 2.6 Harjoitustehtäviä . . . . .                                    | 22        |
| <b>3 Taloustieteiden tärkeimpiä suureita</b>                       | <b>25</b> |
| 3.1 Määräsuureet . . . . .                                         | 25        |
| 3.2 Arvosuureet . . . . .                                          | 25        |
| 3.3 Nopeussuureet . . . . .                                        | 26        |
| 3.4 Kertymäsuureet . . . . .                                       | 26        |
| 3.4.1 Diracin $\delta$ -funktio . . . . .                          | 27        |
| 3.5 Rajasuureet . . . . .                                          | 28        |
| 3.6 Rajoitteet . . . . .                                           | 29        |

|          |                                                      |           |
|----------|------------------------------------------------------|-----------|
| 3.7      | Informaatio . . . . .                                | 29        |
| 3.7.1    | Esimerkki: osakesijoittaja . . . . .                 | 30        |
| 3.8      | Harjoitustehtäviä . . . . .                          | 31        |
| <b>4</b> | <b>Kaksi numeerisen matematiikan menetelmää</b>      | <b>33</b> |
| 4.1      | Iterointi . . . . .                                  | 33        |
| 4.1.1    | Keplerin yhtälön ratkaiseminen . . . . .             | 33        |
| 4.1.2    | Banachin kiintopistelause . . . . .                  | 34        |
| 4.1.3    | Diskreetit dynaamiset järjestelmät . . . . .         | 36        |
| 4.2      | Gradienttimenetelmä . . . . .                        | 36        |
| 4.2.1    | Vapaa markkinatalous vs. suunnitelmatalous . . . . . | 38        |
| 4.3      | Harjoitustehtäviä . . . . .                          | 39        |
| <b>5</b> | <b>Variaatiolaskentaa</b>                            | <b>41</b> |
| 5.1      | Suunnattu derivaatta ja ääriarvot . . . . .          | 41        |
| 5.1.1    | Ääriarvon olemassaolon riittäviä ehtoja . . . . .    | 43        |
| 5.1.2    | Variaatiolemma . . . . .                             | 46        |
| 5.2      | Eulerin-Lagrangen yhtälö . . . . .                   | 47        |
| 5.2.1    | Eulerin-Lagrangen kenttäyhtälöt . . . . .            | 49        |
| 5.3      | Lagrangen kertojat . . . . .                         | 50        |
| 5.4      | Harjoitustehtäviä . . . . .                          | 51        |
| <b>6</b> | <b><math>\epsilon</math>-optimaalisuus</b>           | <b>53</b> |
| 6.1      | Harjoitustehtäviä . . . . .                          | 56        |
| <b>7</b> | <b>Diskreettiä optimointia</b>                       | <b>59</b> |
| 7.1      | Bellmanin optimaalisuusperiaate . . . . .            | 59        |
| 7.2      | Optimaalisuusyhtälö ja säätöteoriaa . . . . .        | 60        |
| 7.2.1    | Diskonttaus . . . . .                                | 62        |
| 7.3      | Optimaalinen kuluttaminen ja sijoittaminen . . . . . | 63        |
| 7.4      | Harjoitustehtäviä . . . . .                          | 64        |

# Luku 0

## Johdanto

Tämä nidos sisältää Jyväskylän yliopiston matematiikan ja tilastotieteen laitoksessa lukuvuonna 2001-2002 luennoitun matematiikan valinnaisen cum laude -kurssin luennot ja harjoitustehtävät. Kurssin nimi oli Optimointi ja mallinnus taloustieteissä ja sen tarkoituksena on johdattaa matemaattiseen mallintamiseen ja optimointiin taloustieteissä ja edelleen tasoittaa matemaattisen koulutuksen saaneiden tai saavien tietä niin matemaattisen rahoitusteorian kuin yleensäkin matemaattisen taloustieteen maailmaan. Yleisinä matemaattisina pohjatietoina vaaditaan vain perusosaaminen moniulotteisesta differentiaali- integraalilaskennasta (Differentiaali- ja integraalilaskenta 1), joka sisältyy kansantaloustieteen opiskelijoille pakollisiin menetelmäkursseihin (Matematiikan peruskurssi ja Matemaattinen taloustiede 1), joten tämä kurssi soveltuu myös taloustieteiden perus- ja jatko-opiskelijoille.

Koska kurssi on luonteeltaan peruskurssi ja osalla opiskelijoista ei ole ennestään takana taloustieteiden tai yleensäkin yhteiskuntatieteiden opintoja, on ensimmäisessä luvussa käsitelty luonnontieteiden ja yhteiskuntatieteiden välisiä eroja. G.H. von Wrightin teoksen *Explanation and Understanding* hengessä kyseenalaistetaan, onko ylipäätensä hyväksyttävää ja menestyksestä soveltaa taloudellisten järjestelmien tarkastelussa samoja käsitteellisiä keinoja kuin fysiikassa ja matemaattisessa biologiassa (esim. populaatiomallit). Esitetään ratkaisu, joka sallii samanaikaisen taloudellisen järjestelmän *ymmärtämisen* ja *selittämisen*: v. Wrightin praktinen syllogismi tulkitaan rationaalisuuden määritelmäksi ja taloustieteiden aksiomaksi asetetaan, että talouden toimijat ovat rationaalisia.

Toisessa luvussa käsitellään erityisempiä tieteenfilosofisia kysymyksiä. Aluksi tutustutaan aksiomaattiseen tieteenihanteeseen ja sen toteutumiseen fysiikassa Mario Bungen johdattamana (*Philosophy of Physics*). Teorian aksiomatisoimiseen tutustutaan esimerkin avulla. Paljastuu, että mitään yleistä yhtenäistä fysiikan maailmankuvaa ei ole olemassa, vaan fysiikka on tilkkutäkki eri teorioita, jotka ovat osin päällekkäisiä ja jopa ristiriidassa keskenään. Määritellään malli mini- tai esiteoriaksi ja tutustutaan mallintamiseen esimerkkien avulla. Toisessa luvussa käsitellään lyhyesti myös dimensioanalyysiä ja opitaan vähentämään Buckinhamin  $\pi$ -teoreeman avulla toisistaan riippuvien muuttujien määrää. Se on menetelmä, jolla on runsaasti käyttöä empiirisessä tutkimuksessa. Lopuksi toisessa luvussa käsitellään lyhyesti ajan, avaruuden ja kentän käsitteitä.

Kolmannessa luvussa käsitellään taloustieteiden tavallisimpia suureita: kertymä-, arvo-, nopeus- ja rajasuureita sekä rajoitteita ja informaatiota. Alenevaan rajahyödyn lakiin tutustutaan.

Neljännessä luvussa tutustutaan kahteen numeerisen matematiikan menetelmään, iterointiin yhtälön ratkaisemiseksi ja gradienttimenetelmään funktion ääriarvon löytämiseksi. Molemmat osoittautuvat taloustieteen kannalta perustaviksi: iterointikaavat vastaavat dynaamisia järjestelmiä diskreetin ajan kera; gradienttimenetelmä antaa mallin, miten voittoa maksimoivan yrittäjän on

järkevä toimia, kun hän tietää voittofunktiostaan vain sen arvon nykyhetkellä toteutuneiden muuttujien arvoilla ja joinakin menneinä hetkinä toteutuneiden muuttujien arvoilla. Gradienttimenetelmää voidaan pitää markkinatalouden “alkusymbolina” Goethen tai Spenglerin mielessä.

Viidennessä luvussa käsitellään ääretönulotteisissa joukoissa määriteltyjen funktioiden ääriarvo-oppia. Sovellutuksena johdetaan klassisen variaatiolaskennan Eulerin-Lagrangen yhtälöt. Mainitaan myös niiden muoto sidotuille ääriarvotähtäville sekä kentille. Jälkimmäisistä esitetään esimerkki, jossa optimaalinen hinnoittelu ajan ja paikan funktiona saadaan osittaisdifferentiaaliyhtälön ratkaisuna.

Kuudennessa luvussa tarkastellaan ääriarvotähtäviä, joilla ei klassisessa mielessä ole ratkaisua. Yksi tällainen on junalle kiirehtijän ongelma: hän pyrkii minimoimaan junan odotusajan, kun täsmälleen junan lähtöhetkellä sen ovet lukkiutuvat. Minimien sijasta haetaan  $\epsilon$ -minimejä eli niitä muuttujan arvoja, joilla funktio on korkeintaan  $\epsilon$ :n päässä infimumistaan. Reaalilukujen täydellisyysaksiooma takaa, että jokaisella alhaalta rajoitetulla funktiolla on infimum. Esitetään ja todistetaan Ekelandin lause, joka sanoo karkeasti, että  $\epsilon$ -minimissä funktion derivaatta on normiltaan pienempi kuin  $\epsilon$ .

Seitsemännessä luvussa käsitellään diskreetin optimoinnin alkeita. Aluksi tarkastellaan tieverkossa liikkumisen kustannusten minimointia ja tutustutaan Bellmanin optimaalisuusperiaatteeseen. Siitä muokataan yleinen optimaalisuusperiaate, jota sovelletaan säätö- eli kontrolliteoriaan, kun aika on diskreetti.

# Luku 1

## Yleistä tieteenfilosofiaa

Tässä luvussa pohdimme, miksi ylipäätään ihmisten ja heidän muodostamiensa yhteisöjen taloudellista toimintaa voidaan menestyksekkäästi selittää, ennustaa ja mallintaa samalla tavalla kuin elotonta luontoa. Tutustumme lyhyesti käsitteisiin *selittää* ja *ymmärtää* sekä niihin liittyvään kiistaan ihmistieteiden (kuten historia) oikeasta metodologiasta. Lopuksi omaksumme kannan, että selittäminen ja ymmärtäminen ovatkin mahdollisia samanaikaisesti ja että ei ole mitään periaatteellista estettä fysiikan käyttämien matemaattisten rakenteiden soveltamiselle taloustieteisiin tai muihinkaan ihmistieteisiin.

### 1.1 Teoreettisen fysiikan menestystarina

Modernin tieteen ja fysiikan alkuna pidämme *Galileo Galilein* kokeita, *Johannes Keplerin* systematisointia planeettojen liikkeistä (Keplerin lait) ja lopullisena läpimurtona Isaac Newtonin yleistä teoriaa mekaniikasta ja gravitaatiosta. Newton ja hänen aikalaisensa *Gotlieb Leibniz* kehittivät differentiaali- ja integraalilaskentaa fysikaalista teoriaansa varten ja niin tuo matematiikan ala syntyi vähintäänkin fysiikan kieleksi. Myöhempi fysiikan kehitys vain vahvisti tätä fysiikan ja matematiikan sisäistä sukulaisuutta: analyyttinen mekaniikka Hamiltonin yhtälöineen ja yleistettyine koordinaatteineen johdatti matematiikkaa kohti koordinaatistosta riippumatonta esitystä, klassisen sähkömagnetismin yhtälöt (Maxwellin yhtälöt) ovat osittaisdifferentiaaliyhtälöitä, samoin kuin kvanttimekaniikan tilayhtälöt.

Moderni fysiikka on tehnyt ihmisestä luonnon valtiaan<sup>1</sup>: jo pelkät mekaniikan, termodynamiikan ja sähkömagnetismin klassiset teoriat edistivät noiden alojen insinööritaitoja niin, että vähistä materiaaleista rakennetaan lujia ja tehokkaita taloja, siltoja, kulku- ja viestintävälineitä<sup>2</sup>. . . Tietokoneiden kehitys mahdollistaa yhä monimutkaisempien järjestelmien ja ilmiöiden laskennallisen tutkimisen ja suunnittelemisen. Esimerkiksi sopii mainita uusien nopeiden paperikoneiden suunnittelemisen (Metso Oyj ja Jyväskylän yliopisto): paperinvalmistusprosessia kuvaaviin virtausdynamiikan osittaisdifferentiaaliyhtälöihin on täytynyt ja voitu sisällyttää epälineaarisia viskositeettitermejä.

Moderni fysiikka on mullistanut myös ihmisen maailmankuvan. Ympäristön maailman ilmiöitä ja prosesseja on voitu yhä tyhjentävämmin ja tyhjentävämmin selittää ja ennustaa. Maailma ei näyttäytynyt enää arvaamattomana ja järjettömänä paikkana, vaan sen lainalaisuudet voidaan paljastaa ja ymmärtää. Kun maailmankuva on muuttunut, ovat muuttuneet usein myös käsitykset ihmisen arvosta, oikeuksista ja velvollisuuksista.

<sup>1</sup>Tätä käsitystä ei esimerkiksi T.S. Kuhn kiistä kuuluisassa teoksessaan *The Structure of Scientific Revolutions*. Ihmisen kyky hallita luontoa on kasvanut alati tieteen vallankumouksista huolimatta.

<sup>2</sup>Kvantti- ja ydinfysiikan sovellutukset tulivat merkittäviksi vasta 1940-luvulla puolijohteiden, ydinaseiden ja -energian myötä.

Fysiikan ja vielä tarkemmin teoreettisen fysiikan menestys on tehnyt siitä esimerkin tai jopa esikuvan muille tieteille. Kuvaavaa on, että merkittävimmät tieteenfilosofit (P. Duhem, K. Popper, T.S. Kuhn, I. Lakatos, P. Feyerabend, L. Laudan ...) ovat itse asiassa kehittäneet tieteenteorioitaan ensisijaisesti fysiikasta ja silti niitä pidetään yleisenä tieteenfilosofiana. Varsin usein taloustieteilijät ja toisinaan jopa jotkut sosiologit kaipaavat tieteenalansa Newtonia, joka klassisen mekaniikan kaltaisella uudella teorialla loisi pohjan tieteenalansa kehittymiselle teoreettisen fysiikan kaltaiseksi “oikeaksi tieteeksi”.

## 1.2 Selittäminen

Karl Popper määritteli teoksessaan [10, §12], että tapahtuman (*kausaalinen selittäminen*) tarkoittaa sitä, että se päätellään kahdesta lausejoukosta, joista ensimmäinen koostuu *yleisistä laeista* ja toinen *erityisistä lauseista*<sup>3</sup>, joita sanotaan myös *alkuehdoiksi*. Esimerkiksi vuorikiipeilijän putoaminen turvaköyden katkettua selitetään seuraavasti. Yleinen laki sanoo, että jokainen tyypin A köysi katkeaa, jos sitä kuormittaa suurempi voima kuin 1,0 kN. Alkuehtoina on, että vuorikiipeilijä N.N. joutui A-tyyppisen turvaköyden varaan, hänen massansa oli 150 kg ja paikkakunnalla vapaan putoamisliikkeen kiihtyvyys oli 9,8 m/s<sup>2</sup>. Alkuehdoista seuraa, että N.N. kuormitti köyttä voimalla 1,5 kN, jolloin se yleisen lain mukaan katkesi. Tämä tietyn köyden tietyssä tilanteessa katkeavaksi väittävä lause on *ennustus*. Lausumme tämän esimerkin yleisemmin termein. Selittäminen on loogiselta rakenteeltaan seuraava:

- |                       |       |
|-----------------------|-------|
| 1. Yleiset lait $L$ . |       |
| 2. Alkuehdot $A$ .    | (PLM) |
| 3. Ennustus $P$ .     |       |

Tätä päättelykaaviota, syllogismia, vastaa logiikan lause “Jos  $L$  ja  $A$ , niin  $P$ ”. Lyhennys (PLM) tulee sanoista peittävän lain malli; Carl Hempel käytti tästä syllogismista sitä nimitystä. Ennustus selittämisen lopputuloksena tarkoittaa niin tulevaisuuteen sijoittuvaa tapahtumaa (prediktio), menneisyyden tapahtumaa (retrodiktio) tai sitten ajallisesti epämääräistä tapahtumaa. Retrodiktio tuli tähtitieteeseen ja fysiikkaan jo Keplerin lakien myötä: planeetan sijainti kullakin ajanhetkellä — menneenä tai tulevana — voidaan ratkaista sen rataelementeistä (alkuehdot) ja Keplerin laeista (yleiset lait)<sup>4</sup>. Klassisen mekaniikan avulla planeetan paikan ennustaminen käy siten, että alkuehtona on planeetan sijainti ja nopeus jollakin ajanhetkellä ja yleisenä lakina planeetan liikeyhtälö, joka on toisen kertaluvun tavallinen differentiaaliyhtälö. Ennustaminen on nyt seuraavan alkuarvotehtävän ratkaisemista<sup>5</sup>:

$$\mathbf{x}''(t) = -\frac{\gamma M_0 \mathbf{x}(t)}{\|\mathbf{x}(t)\|^3}, \quad t \in \mathbb{R}, \quad (1.1)$$

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{x}'(0) = \mathbf{v}_0, \quad (1.2)$$

missä  $\gamma$  on gravitaatiovakio,  $M_0$  esittää Auringon massaa,  $t$  ajanhetkeä ja  $\mathbf{x}(t)$  planeetan sijaintia hetkellä  $t$ . Voimme ennustaa siis planeetan paikkaa nyt, menneisyydessä tai tulevaisuudessa<sup>6</sup>.

<sup>3</sup>Engl. singular statements.

<sup>4</sup>Ks. esimerkiksi Hannu Karttunen, Johdatus taivaanmekaniikkaan. Limes ry, Helsinki, 1980

<sup>5</sup>Oletamme, että planeetan massa on hyvin pieni Auringon massa verrattuna; oletuksesta voidaan luopua siirtymällä massakeskipistekoordinaatistoon, jolloin saadaan kerrointa  $\gamma M_0$  vaille sama differentiaaliyhtälö

<sup>6</sup>Ahti Karjalainen lausuikin taannoin, että “ennustaminen on vaikeaa, varsinkin tulevaisuuden ennustaminen”.

## 1.3 Ihmiset järjettämiä luontokappaleita?

On ilmeistä että selittäminen syllogismin (PLM) mukaan on oikea ja sopiva menetelmä luonnontieteille aina biologiaa ja fysiologiaa myöten<sup>7</sup>. Mutta sopiiko se esimerkiksi historiaan. Jo Popper itse oli epäileväinen: (PLM) olisi kyllä sinänsä sovellettavissa, mutta historian lainalaisuudet ovat niin banaaleja, että niitä ei edes vaivauduta muotoilemaan. Ja jos ne muotoiltaisiin, saataisiin niistä vain harvoin täsmällisiä ennustuksia. Toisenlaisen kannan esitti William Dray teoksessaan *Laws and Explanation in History* (1957): historian yleiset lainalaisuudet ovat niin yksityiskohtaisia, että ne eivät ole enää yleisiä. Tällainen on esimerkiksi se ”yleinen” historian laki, joka selittää, miksi Ludwig XIV joutui epäsuosioon. Tuo laki lausuu itse asiassa, että jokainen hallitsija, joka harjoittaa sellaista politiikkaa kuin Ludwig XIV niissä oloissa kuin Ludwig XIV, joutuu epäsuosioon. Tämä ”yleinen” laki soveltuu vain Ludwig XIV:een. Historian tapahtumien selityksien looginen rakenne on siis jotakin muuta kuin luonnontieteissä pätevä (PLM),

Lisäongelmiin (PLM) joutuu siitä, että ennustuksen esittäminen voi saada ihmiset muuttamaan käyttäytymistään ja siten kumoamaan ennustuksen. Ihmiset voivat olla patologisesti niin itsepäisiä, että he tekevät päinvastoin kuin ennustetaan. Tällöin heidän käyttäytymisensä ennustaminen on loogisesti mahdotonta. Ennustus voi olla myös itsensä toteuttava tai kumoava: uskottavan ennustuksen julkistaminen talouslaman alkamisesta saa kuluttajat supistamaan kulutustaan, jolloin talous ajautuu lamaan, johon se ei muuten päätyisi; julkistettu ennustus jonkin osakkeen kurssinoususta lisää sen kysyntää niin, että sen kurssi nousee.

Aristoteles tunnisti neljänlaisia syitä. *Causa materialis* eli aineksellinen syy on se materiaali, josta olio on muodostunut. Tietystä materiaalista ei sovi odottaa saavansa mitä hyvänsä. Suomalaisissa sotilaissa ei tunnetusti ole ainesta virheettömästi marssivaan paraatijoukkoon. *Causa formalis* eli muodollinen syy on rakenne tai yleinen laki, jonka mukaan tapahtuminen kävisi, jos ulkoisia tekijöitä ei olisi. *Causa efficiens* eli vaikuttava syy on se ulkoinen tekijä, joka panee tapahtumisen liikkeelle. *Causa finalis* eli päämääräisyys on se, mitä kohti tapahtuminen kehittyy, mihin se päättyy ja jonka luonnollinen tarkoitus se on.

Nykyaikaisessa mielessä tapahtuman (kausaalinen) selittäminen on lähinnä Aristoteleen muodollisen ja vaikuttavan syyn hakemista. Kuitenkin tällaisissa rinnastuksissa on oltava varovainen, sillä Aristoteleen maailmankuva ja peruskäsitteet poikkeavat oleellisesti nykyisistä<sup>8</sup>

### 1.3.1 Teleologia ja variaatioperiaatteet

Aristoteleen selitysmalli lukuisine syineen näyttää perustuvan analogiaan ihmisen käyttäytymisestä: ihmisillä on tavoitteita, joita he pyrkivät saavuttamaan toiminnallaan. Aristoteles laajensi mallinsa myös fysiikkaan: kivi *pyrkii* kohti luonnollista tilaansa, siksi kivi putoaa. Aristoteleen maailmankuva oli fysiikkaa myöten *teleologinen*.

Fysiikassa tunnetaan joukko variaatioperiaatteita. Yksi niistä on *Fermat'n periaate*, joka sanoo valonsäteen kulkevan taittavassa väliaineessa pisteestä *a* pisteeseen *b* sitä reittiä, jota pitkin matkaan kuluu vähiten aikaa, eli valonsäteen reitti eli polku  $\gamma: [0, 1] \rightarrow \mathbb{R}^3$  on seuraavan ääriarvotehtävän ratkaisu:

$$\min \left\{ \frac{1}{c} \int_0^1 n((\gamma(\sigma))) \|\gamma'(\sigma)\| d\sigma \mid \gamma \in C^1([0, 1]; \mathbb{R}^3), \gamma(0) = a, \gamma(1) = b \right\},$$

<sup>7</sup>Georg Henrik von Wrightin teos *Vetenskapen och förnuftet*, Bonnier, Tukholma 1986, on yritys alustavasti hahmottaa toisenlaista menetelmää ja ajattelutapaa luonnontieteille. Holistista, sanoi v. Wright. Romanttista antiteknologiaa, sanoi Erik Allardt.

<sup>8</sup>Tätä valaisevat hyvin Jaakko Hintikan esseet kokoelmassa Tieto on valtaa. Esimerkiksi mahdollinen oli Aristoteleelle sellaista, joka toteutuu joskus.

missä  $n(\mathbf{x})$  on taitekerroin paikassa  $\mathbf{x}$  ja  $c$  valon nopeus tyhjiössä (valon nopeus taittavassa väliaineessahan on  $c/n$ ). Toinen kuuluisa fysiikan variaatioperiaate on *Hamiltonin pienimmän vaikutuksen periaate*. Sen mukaan järjestelmä kehittyy ajanhetkien  $t_0$  ja  $t_1$ ,  $t_0 < t_1$ , välillä niin, että vaikutusintegraali

$$\int_{t_0}^{t_1} L(t, \mathbf{x}(t), \mathbf{x}'(t)) dt$$

saavuttaa pienimmän arvonsa; tässä  $L(t, \mathbf{x}, \mathbf{x}')$  on järjestelmän *Lagrangen funktio*, esimerkiksi  $m$ -massaiselle planeetalle  $M$ -massaisen auringon gravitaatiokentässä se on

$$L(t, \mathbf{x}, \mathbf{x}') = T - V = \frac{1}{2}m\|\mathbf{x}'\|^2 + \frac{\gamma mM}{\|\mathbf{x}\|},$$

missä  $T$  on liike-energia ja  $V$  potentiaalienergia. Newtonin jälkeisestä fysiikasta teleologiset selitykset on järjestelmällisesti karkotettu pois. Kuitenkin variaatioperiaatteita on — sinänsä virheellisesti — tulkittu teleologisina: valonsäde *pyrkii* minimoimaan matkaansa kuluvan ajan. Oikea variaatioperiaatteiden tulkinta käy ilman tarkoituksia tai pyrkimyksiä: jos  $u$  on funktionaalin  $J(\cdot)$  minimikohta ja suunnattu derivaatta eli Gâteaux'n derivaatta

$$\lim_{\epsilon \rightarrow 0+} \frac{J(u + \epsilon h) - J(u)}{\epsilon}$$

on olemassa, on se nolla. Toisaalta Gâteaux'n derivaatta osataan laskea kaikkiin suuntiin<sup>9</sup>, jolloin Hamiltonin pienimmän vaikutuksen periaatteesta seuraa Lagrangen liikeyhtälö

$$\frac{d}{dt} \frac{\partial L}{\partial \mathbf{x}'} - \frac{\partial L}{\partial \mathbf{x}} = 0,$$

joka esimerkiksi yllä mainitun planeetan tapauksessa on tuttu Newtonin liikeyhtälö. Implikaation saamme pätemään toisinkin päin, kun muotoilemme variaatioperiaatteet ekstremaaliperiaatteina eli että ääriarvokohdan sijasta haetaan funktiota, joka on kyseisen funktionaalin ekstremaali eli se piste, jossa Gâteaux'n derivaatta on nolla kaikkiin suuntiin.

Biologiassa evoluutioajatuksia muotoillaan usein variaatioperiaatteina ja niihin vedoten huonossa popularisomisessa puhutaan luonnon *viisaudesta* tai *tarkoituksenmukaisuudesta*. Yksi kehitysbiologinen teoria väittää, että ne hyppyrötaurokset, jotka tuottavat mahdollisimman paljon jälkeläisiä, menestyvät geneettisesti. Tästä ei kuitenkaan seuraa, että luonnon, hyppyröttäpopulaation tai edes yksittäisten hyppyröttakoiraiden *tarkoituksena* olisi maksimoida jälkeläistensä määrää.

## 1.4 Ymmärtäminen

Olemme jo todenneet koko joukon ongelmia, joihin peittävän lain malli jouuu, jos sitä tarjotaan yleiseksi selityksen malliksi ihmistieteissä<sup>10</sup>. G.H. von Wright kirjoittaa merkittävässä teoksessaan [12] kahdesta tieteellisestä perinteestä, *aristotelisestä ja galileisesta*. Niin ansiokas kuin galileinen perinne on ollutkin luonnontieteissä, ei sitä ole pidettävä ihmistieteissä pätevänä. Tällöin *metodologisen monismin* ihanteesta joudutaan luopumaan.

Aristoteleen syiden oppi ja sen ilmeinen perustuminen ihmisten tavalliseen käyttäytymiseen tavoitteitaan toteuttavina oliona tekee ymmärrettäväksi, miksi koko aristotelinen traditio (Aristoteles-Tuomas Akvinolainen-Hegel-Marx-hermeneutikot...) tarjoaa kelpoisemman pohjan hakea ihmistieteiden selitysten yleistä rakennetta. Avainkäsite on aikomus, intentio. Tärkeä vaihe aristotelisessä traditiossa on 1800-luvun hermeneutikot, jotka pyrkivät tulkitsemaan ja

<sup>9</sup>Oletetaan, että  $h$  on jatkuvasti differentioituva ja  $h(t_0) = h(t_1) = 0$ .

<sup>10</sup>Ks. tämän jakson aiheesta laajemmin [12] ja [8]

ymmärtämään uskonnollisia tekstejä. Saksalainen filosofi ja historioitsija Droysen lieene ensimmäinen, joka käytti käsiteparia *selittää-ymmärtää* (Erklären-Verstehen) painottamaan metodologista dualismia: luontoa sopii kyllä selittää, mutta ihmistä on ymmärrettävä. Nämä ajatukset esitti systemaattisesti William Dilthey, joka kutsui tieteenaloja, joilla ymmärtämismenetelmää on käytettävä, termillä *Geisteswissenschaften*. Käännämme sen suomeksi sanalla ihmistieteet. Ymmärtämiseen liittyy sellaisia psykologisia osasia, joita selittämisessä ei ole. Simmel painotti, että ymmärtämiseen sisältyy *eläytyminen* (Einfühlen): tutkijan on pyrittävä luomaan uudestaan tutkimuskohteensa henkinen ilmapiiri, käsitykset, motivaatiot, tunteet.

Näiden psykologisten osasten lisäksi ymmärtämiseen kuuluu aikomuksen eli *intention* käsite. von Wright esitti seuraavan *praktisen syllogismin*, [12, IV.4]:

1. Henkilö *A* aikoo aiheuttaa tapahtuman *p*.
2. *A* otaksuu, että hän ei voi aiheuttaa *p*:tä, ellei hän tee tekoa *a*. (PS)

---

3. *A* alkaa tehdä tekoa *a*.

Tästä päättelykaaviosta von Wright esittää joukon muunnoksia ja täydennyksiä<sup>11</sup>, mutta perusidean näemme parhaiten tästä yksinkertaisimmasta versiosta. Hän tarjoaa praktista syllogismia yleiseksi selityksen malliksi ihmistieteisiin.

## 1.5 Selittäminen ja ymmärtäminen

Metodologinen monismi oli yksi *positivismin* tunnusmerkeistä. Kuitenkin niin klassinen positivismi kuin uuspositivismi ovat aikoja sitten olleet *passé*. Tuleeko ihmisen käyttäytymistä selittää vai ymmärtää? Onko tämä kysymys edes hyvin asetettu? Vai-sanahan sulkee pois sen, että molempia menetelmiä voidaan käyttää samasta tutkimuskohteesta. Muistellaanpa esimerkkiä, jonka Rudolf Carnap<sup>12</sup> esitti teonvapautta pohtiessaan<sup>13</sup>. Hänellä oli ystävä, Jussi, joka rakasti klassista musiikkia ja tämän kotikaupungissa oli tunnetun artistin konsertti, joka oli kuitenkin jo loppuunmyyty. Carnap tunsu ystävänsä hyvin ja tiesi, että jos tälle annetaan pääsylippu tuohon konserttiin, tämä menee sinne vuorenvaramasti<sup>14</sup>. Tässä tapauksessa on yleinen laki, alkuehdot, ennustus aivan kuten peittävän lain malli vaatii, joten Jussin meneminen konserttiin voidaan selittää. Mutta me voimme myös soveltaa praktista syllogismia tähän tilanteeseen: (1) Jussi aikoo kuulla ihailemaansa artistia, (2) hän tietää, että hän ei kuule tätä, jos ei käytä samaansa pääsylippua, (3) niinpä hän menee konserttiin. Huomaamme, että selittäminen ja ymmärtäminen ovat vastauksia *eri* kysymyksiin, jotka voidaan esittää *samasta* kohteesta<sup>15</sup>.

Kehitämme tätä ajatusta pidemmälle taloustieteitä silmällä pitäen. Muutamme von Wrightin praktisen syllogismin rationaalisuuden *määritelmäksi*<sup>16</sup>:

1. Henkilö *A* aikoo aiheuttaa tapahtuman *p*.
2. *A* otaksuu, että hän ei voi aiheuttaa *p*:tä, ellei hän tee tekoa *a*.

---

3. *A*:n on rationaalista alkaa tehdä tekoa *a*.

Jos asetamme taloustieteiden aksioomaksi seuraavan oletuksen

<sup>11</sup>Ks. myös [8].

<sup>12</sup>Sattumoisin hänen teoksensa *Logische Aufbau der Welt*, Berliini 1928, on uuspositivismin klassikko. Maailma yritetään rakentaa välittömiä aistimuksia kuvaavista lauseista.

<sup>13</sup>Rudolf Carnap, *Philosophical Foundations of Physics*, luku 22, Basic Books, Lontoo 1966

<sup>14</sup>Carnap antaa Jussille pääsylipun ja saa hänet siten menemään konserttiin. Oliko Jussin konserttiin meneminen vapaa teko?

<sup>15</sup>Näin Matti Sintonen on huomauttanut.

<sup>16</sup>Toki *a*:n tekeminen oletetaan vain välttämättömäksi, ei riittäväksi toimeksi. Emme puutu tähän.

(RA) Kaikki taloudelliset toimijat ovat rationaalisia.

niin selittäminen ja ymmärtäminen ovat saman kolikon eri puolia.

Havainnollistetaan tätä konkreettisella esimerkillä. Olkoon  $A$  yrittäjä, joka tuottaa tiettyä tuotetta määrän  $q$  päivässä. Tuotantokustannukset päivässä olkoot  $C(q) = q^2 + \frac{1}{2}$  ja hän voi myydä tuotteet yksikkö hinnalla  $p = 2$ . Tällöin hänen voittonsa päivässä on

$$\Pi(q) = pq - C(q) = 2q - q^2 - \frac{1}{2} = -(q-1)^2 + \frac{1}{2}.$$

Yrittäjän voitto on suurin, kun hän tuottaa tuotetta päivässä määrän  $q = 1$ . Muodostamme seuraavat lauseet.

- (i) Kaikki yrittäjät tuottavat niin, että heidän otaksumansa voitto on suurin mahdollinen.
- (ii) Yrittäjä  $A$  otaksuu voittofunktionsa olevan  $\Pi(q)$ .
- (iii) Jokainen yrittäjä aikoo maksimoida voittonsa.
- (iv) Yrittäjä  $A$  otaksuu, että hän ei maksimoi voittoaan ellei tuota määrää  $q^*$ , missä  $\Pi(q^*) \geq \Pi(q)$  kaikilla  $q$ .

Yleinen laki (i) seuraa lauseesta (iii), rationaalisuuden määritelmästä ja aksiomasta (RA). Lauseesta (ii) seuraa lause (iv), jos yrittäjä osaa funktion ja sen suurimman arvon käsitteet. Nyt praktinen syllogismi kuuluu:

1. Yrittäjä  $A$  aikoo maksimoida voittonsa.
  2.  $A$  otaksuu, että hän ei maksimoi voittoaan, ellei hän tuota määrää  $q^*$ .
- 
3.  $A$  tuottaa määrän  $q^*$ .

Peittävän lain malli taasen sisältää lauseen (i) yleisenä lakina ja lauseen (ii) alkuuehtona ja johtopäätöksenä on se, että  $A$  tuottaa määrän  $q^*$ .

Fysiikan vahvimmat käsitteelliset menetelmät ovat differentiaali- ja osittaisdifferentiaaliyhtälöitä. Niitä voidaan johtaa vastaavista variaatioperiaatteista. Näin idea taloudellisista toimijoista rationaalisina voittonsa tai hyötynsä maksimoijina johtaa luonnollisella tavalla samanlaisiin differentiaali- ja osittaisdifferentiaaliyhtälöihin kuin fysiikassa ja silti ei oleteta, että ihmiset olisivat järjettömiä tai tavoitteettomia luontokappaleita<sup>17</sup>. Eksakti ja matemaattinen käsitteen ja teorianmuodostus on siis mahdollista niin taloustieteissä kuin muissakin yhteiskuntatieteissä. Tällä kurssilla tulemme soveltamaan suoraan muun muassa mekaniikan tai lämpöopin matemaattista kalustoa taloustieteisiin.

### 1.5.1 Poliittisesta taloustieteestä

Karl Marxin Pääoman esityön nimi oli Poliittisen taloustieteen kritiikkiä. Taavoitteena se ei ole vanhentunut. Yksi kurssimme tavoitteista on oppia muodostamaan taloustieteellisiä malleja, joita ja joiden antamia tuloksia voidaan käyttää argumentteina poliittisessa keskustelussa ja propagandassa. Opimme esimerkiksi todistelemaan, että sama verojärjestelmän muutos aiheuttaa massatyöttömyyden tai työvoiman kysynnän kasvun, kumpaa tulosta toimeksiantajamme sitten sattuu haluamaan. Tutustumme seuraavassa luvussa teorioiden aksiomatisoimiseen myös siksi, että aksiomatisoitu teoria tai malli paljastaa ideologiset olettamuksensa.

<sup>17</sup>Ja jos he ovat järjettömiä luontokappaleita, niin silloin fysiikan menetelmät ovat omiaan heidän käyttäytymisensä selittämiseen.

## 1.6 Harjoitustehtäviä

1. Olkoon  $x \mapsto (x, f(x))$ ,  $[a_1, b_1] \rightarrow \mathbb{R}^2$ ,  $a_1 < b_1$ , jatkuvasti differentioituva polku pisteestä  $(a_1, a_2)$  pisteeseen  $(b_1, b_2)$ . Sen pituus on

$$\int_{a_1}^{b_1} \sqrt{1 + f'(x)^2} dx.$$

Osoita Eulerin-Lagrangen yhtälön avulla, että lyhin polku yhdestä pisteestä  $(a_1, a_2)$  toiseen pisteeseen  $(b_1, b_2)$  on niitä yhdistävä jana.

2. Osoita Fermat'n periaatteen ja Eulerin-Lagrangen yhtälön avulla, että homogeenisessa väliaineessa (eli taitekerroin on vakio  $n_0 \geq 1$ ) valo kulkee yhdestä pisteestä toiseen pitkin niitä yhdistävää janaa. Tarkastele vain 2-ulotteinen tapaus ja käytä  $x$ -koordinaattia polun parametrina eli polku on  $x \mapsto (x, f(x))$ .
3. Määrittää valonsäteen reitti origosta pisteeseen  $(1, 1, 0)$ , kun väliaineen taitekerroin on  $n(x, y, z) = 1 + y^2$ . Ongelma pelkistyy kaksiulotteiseksi, joten riittää hakea polku  $x \mapsto (x, f(x))$ ,  $[0, 1] \rightarrow \mathbb{R}^2$ . Jos et osaa ratkaista samaasi differentiaaliyhtälöä, kokeile polun parametrisointia  $y \mapsto (g(y), y)$ . ( $\text{Darsinh } x = 1/\sqrt{1+x^2}$ .)
4. Yrittäjän voitto aikayksikössä hetkellä  $t$  olkoon

$$\Pi(t, q(t)) = p(t)q(t) - C(q(t)),$$

missä  $p(t)$  on tuotteen yksikköhinta hetkellä  $t$ ,  $q(t)$  on tuotteen tuotantonopeus hetkellä  $t$  ja differentioituva funktio  $C: \mathbb{R} \rightarrow \mathbb{R}$  esittää tuotantokustannuksia aikayksikössä. Tällöin hänen voittonsa ajanjaksolla  $[t_0, t_1]$ ,  $t_0 < t_1$ , on

$$\int_{t_0}^{t_1} \Pi(t, q(t)) dt.$$

Yrittäjä ei voi vaikuttaa yksikköhintaan  $p(t)$ , mutta tuotantonopeutensa  $q(t)$  hän voi valita vapaasti. Osoita Eulerin-Lagrangen yhtälön avulla, että hetkellisen voiton maksimointi antaa parhaan voiton myös ajalla  $[t_0, t_1]$ .

5. Muutetaan edellisen tehtävän tapausta niin, että tuotantonopeuden muuttamisesta tulee sopeutumiskustannuksia; olkoon esimerkiksi

$$\Pi(t, q, q') = pq - \frac{1}{2}\alpha q'^2 - C(q), \quad t \in [t_0, t_1],$$

missä vakio  $\alpha > 0$ . Johda voiton maksimoimisesta ja Eulerin-Lagrangen yhtälöstä liikeyhtälö

$$\alpha q''(t) + p(t) - C'(q(t)) = 0, \quad t \in [t_0, t_1].$$

Ratkaise se kun  $C(q) = \frac{1}{2}\beta q^2$ , vakio  $\beta > 0$ ,  $p(t) \equiv 1$ ,  $t_0 = 0$ ,  $t_1 = 1$   $q(0) = 0$  ja  $q(1) = 1$ . (Lineaarisen epähomogeenisen differentiaaliyhtälön yleinen ratkaisu on yksittäisratkaisun ja vastaavan homogeenisen differentiaaliyhtälön yleisen ratkaisun summa.)

6. Vertaile kahden edellisen tehtävän liikeyhtälöitä ja niiden ratkaisuja. Mitä tapahtuu, kun  $\alpha \rightarrow 0+$ ?



## Luku 2

# Erityistä tieteenfilosofiaa ja -metodologiaa

Luomme tässä luvussa yleissilmäyksen tieteellisen teorian rakenteeseen ja luonnehdimme mallit — mallinnuksen kohteet tai tulokset — mini- tai esiteorioina. Erityistä huomiota kiinnitämme siihen, että malleissa esiintyvät *suureet* eivät voi riippua toisistaan miten tahansa (ks. jakso dimensioanalyysistä). Lopuksi käsittelemme aikaa, avaruutta, tilaa ja kenttää.

## 2.1 Aksiomaattinen tieteenihanne

Käsittelemme tässä ja seuraavassa jaksossa nykyfysiikan rakennetta, ei fysiikan syntyä eikä fysiikan historiaa. Monelta osin nojaudumme argentiinalais-kanadalaisen filosofin ja fyysikon Mario Bungen teokseen *Philosophy of Physics* (D. Reidel, 1973). Emme kuitenkaan sitoudu Bungen ankaraan *tieteelliseen realismiin*<sup>1</sup>, vaan pyrimme esitykseen, joka on sopusuhteissa sekä tieteellisen realismin että *tieteellisen instrumentalismin*<sup>2</sup> kanssa. Tieteellisen instrumentalismin klassikkona pidetään ranskalaisen Pierre Duhemin teosta *La théorie physique* (1914), josta otamme joitakin tarkasteluja.

### 2.1.1 Fyysikot ja filosofia

Useimmat nykyfyysikot suhtautuvat filosofiaan vastenmielisesti: se on hyödytöntä käsitevääntelyä, josta oivallinen esimerkki on G.W. Hegelin esittämä sähkön määritelmä. Fyysikoiden vieroksunta ei kohdistu vain spekulatiiviseen Hegelin filosofiaan vaan jopa analyttiseen filosofiaan. Lisäksi moni fyysikko, teoreettinenkin, vieroksuu matematiikkaa. Voidaankin sanoa, että yhteistä ihmistieteille ja fyysikoille on täsmällisen käsitteen- ja teorianmuodostuksen väheksyminen ja vieroksuminen. Teorioita muodostellaan ja kokeita tehdään välittämättä tehdä selväksi, mitä oikeastaan on tehty.

Kysymykseen, miksi teoriaa ei ole aksiomatisoitu tai miksi teorian matemaattiset käsitteet eivät ole hyvin määriteltäviä, fyysikko tavallisesti vastaa, että se ei ole tarpeen tai että se on jopa haitallista tieteelliselle luovuudelle. Kuitenkin varsinkin nimekkäimmät fyysikot ovat vanhoilla päivillään alkaneet kiinnostua ja kirjoitella fysiikan filosofiasta ja filosofiasta yleensä (Einstein, Bohr, Heisenberg, Wigner, von Weizsäcker, Laurikainen...). Esimerkiksi kvanttimekaniikasta ja sen filosofisesta merkityksestä filosofoivat fyysikot kiistelevät yhä.

<sup>1</sup>Tieteenfilosofinen kanta, jonka mukaan teorian teoreettiset termit viittaavat todellisuuden olioihin tai esittävät niitä. Esimerkiksi sähkö- ja magneettikenttä ovat olemassa.

<sup>2</sup>Tieteenfilosofinen kanta, jonka mukaan teorian teoreettiset termit ovat vain lyhennysmerkintöjä. Siis sähkö- ja magneettikenttiä ei ole olemassa.

### 2.1.2 Fysiikka ja filosofia

Matematiikan harjoittamisessa on tuloksen löytäminen ja todistaminen eri asioita: tuloksen saa arvata tai nähdä unessa, mutta sen todistaminen on sidottua; tulos on pääteltävä loogisesti määritelmistä ja jo todistetuista lauseista. Samoin fysiikan teorioiden muodostamisessa saadaan olla huolettomia, mutta aikanaan teoriat olisi systematisoitava, lausuttava täsmällisesti, itse asiassa ne tulisi *aksiomatisoida*. Siinä vaiheessa filosofia ja fysiikka ovat erottamattomia, sillä silloin heräävät muun muassa seuraavat kysymykset:

1. Tulkitseeko matemaattinen formalismi itse itsensä vai tarvitaanko tulkin-  
taolettamuksia, ja jos tarvitaan, kuinka nämä olettamukset tulee ymmär-  
tää?
2. Mistä fysiikan teoriat kertovat: fysikaalisista järjestelmistä, laboratorio-  
järjestelyistä, tiedostamme maailmasta vai jostakin muusta?
3. Miten teorian peruskäsitteet on otettu käyttöön: viittaamalla mittaukseen,  
määrittelemällä suoraan vai aksiomaattisesti?
4. Mitä käyttöä aksiomatiikalla on fysiikassa?
5. Miten eri fysiikan teoriat suhtautuvat toisiinsa?
6. Miten teoria kohtaa kokeen: suoraan vai muiden teorioiden avulla?

Näitä joudutaan kysymään viimeistään silloin, kun halutaan ymmärtää, mitä kukin fysiikan teoria kertoo maailmasta. Tai mitä fysiikka yleensä kertoo maailmasta.

### 2.1.3 Fysiikan maailmankuva

Fysiikan yliopisto-opetuksessa Jyväskylässä opiskelija perehtyy aluksi välttämättömimpään matemaattiseen käsitteistöön ja menetelmiin. Rinnan sen kanssa tutustutaan aluksi klassiseen mekaniikkaan, erikoiseen suhteellisuusteorian, sitten sähkö- ja magnetismioppiin, termodynamiikkaan, kvanttimekaniikkaan, tilastolliseen fysiikkaan (sekä klassiseen että kvanttimekaaniseen), atomi-, ydin- ja alkeishiukkasfysiikkaan. Varsinkin alkeishiukkasfysiikan osalta opetus jää varsin pinnalliseksi ja vain tuloksia esitteleväksi: alan teoria (kvanttikenttä- ja yhtenäisteoriat) on niin monimutkaista ja vaikeaa, että sitä ei voida cum laude -tasolla esittää. Teorioita ei opeteta aksiomaattisessa muodossa, eksplisiittisesti ei tehdä selkoa, mistä teoriat kertovat. Ei korosteta, että fysiikka koostuu useista teorioista, joita ei voida toisistaan päätellä, jotka menevät toistensa kanssa päällekkäin ja jotka ovat usein jopa ristiriidassa keskenään.

Albert Einsteinin unelma oli muodostaa yleinen kenttäteoria, joka sisältäisi erikoistapauksina tunnetut ja testeistä selvinneet fysiikan teoriat. Einsteinin unelman toteuttamiseksi on ponnisteltu ahkerasti ja on edistyttykin, mutta neljästä perusvuorovaikutuksesta on vasta kolme saatu saman teorian alle (sähköinen, heikko ja vahva) eikä esimerkiksi hiukkasmekaniikkaa ja tilastollista mekaniikkaa ole saatu yhdistettyä. Yleistä fysikaalista teoriaa ei siis vielä ole. Toisaalta kokemus osoittaa, että vanhojen teorioiden käytännöllinen arvo säilyy: autoinsinööri haaskaa vain aikaansa, jos hän käyttää työssään klassisen mekaniikan sijasta yleistä suhteellisuusteoriaa, sillä jälkimmäinen vie vain hankaliin laskuihin. Samoin maailman ja sen lainalaisuuksien ymmärtämisen kannalta yksinkertaisempi, mutta pätevyysalueeltaan rajallisempi teoria, on usein parempi.

Yleistä yhtenäistä fysikaalista maailmankuvaa ei siis ole: fysiikka on eri teorioiden verkosto. Tämän luulisi olevan kuolinisku tieteelliselle realismille, mutta juuri Mario Bunge korostaa, että yhtenäisteoriaa ei ole ja silti hän puolustaa vakain äänenpainon käsitystä, että fysiikan teoriat kertovat todellisista fysikaalisista olioista, että elektronilla ja magneettikentällä on fysikaaliset referentit.

Mikä fysiikan maailmankuva kulloinkin on ollut, on vaihdellut fyysikoiden muotifilosofian mukaan. 1950- ja 60-luvuilla vallitsi uuspositivismi: silloin katsottiin, että esimerkiksi kvanttimekaniikan aaltofunktio ei olekaan fysikaalinen suure, vaan se kuvaa vain tietoa tutkimuskohteesta, jolloin mittauksen myötä aaltofunktion tulee muuttua. Samoin erikoista suhteellisuusteoriaa esitettiin havaitsijoiden avulla niin kiihkeästi, että päädyttiin jopa käsitykseen, että järjestelmä on olemassa vain, jos sitä joku havaitsee. Sitten realismi on voittanut alaa ja tullut ilmeisesti vallitsevaksi muotifilosofiaksi. Kun ennen fyysikot pitivät itsestään selvänä, että kvanttimekaaniset järjestelmät eivät ole kovinkaan todellisia, vaan vain mitatut energiatasot (spektriviivat) ovat todellisia, ovat nyttemmin realistiset dogmit jälleen saaneet itsestäänselvyyden asemaa. Millainen fysiikan maailmakuva on, riippuu siten oleellisesti niistä filosofisista ja metodologisista periaatteista, joihin kulloinkin sitoudutaan.

## 2.2 Tieteellisen teorian rakenne

### 2.2.1 Miksi fysiikan teorial on aksiomatisoitava?

Tieteellinen teoria voidaan esittää kolmella tavalla: historiallisesti, heuristisesti tai aksiomaattisesti. Historiallinen esitys hahmottaa alkuperäisen ongelman ja sen erilaiset ratkaisuyritykset, tavan, jolla omaksuttu ratkaisu perusteltiin ja ratkaisun todelliset tai todennäköiset vaikutukset myöhempään alan kehitykseen. Heuristinen lähestymistapa tarttuu teorian kaikkein hyödyllisimpiin kaavoihin (ei välttämättä kaikkein perustavimpiin) ja se ryntää johtamaan niistä seurauksia ja soveltamaan niitä. Heuristinen lähestymistapa on vallitseva fysiikan opetuksessa ja julkaisuissa. Sekä historiallinen että heuristinen lähestymistapa ovat valaisevia, mutta historiallinen on epäkäytännöllinen: jos halutaan nopeita tuloksia, heuristista lähestymistapaa tulee suosia.

Molemmat nämä lähestymistavat epäonnistuvat teorian esittämisessä kokonaisuudessaan: ne vaikenavat useimmista teorian esiolettamuksista, ne eivät esitä kaikkia sen perusoletuksia, ne jättävät sen loogisen rakenteen laajalti pimeyteen ja ne ovat monimielisiä tai epämääräisiä puhuessaan sen fysikaalisesta merkityksestä. Näihin pulmiin tarjoaa ratkaisun fysikaalisen teorian aksiomatisointi.

Heuristisesti tai historiallisesti esitetty teorial ovat yksinkertaisesti niin epämääräisiä, että niiden avulla ei voida vastata, millainen on fysiikan maailmankuva. Huolellinen aksiomatisointi sen mahdollistaa, mutta silloin paljastuu, että saatu maailmankuva riippuu niistä filosofisista opeista, joista on pidetty kiinni (esim. tieteellinen instrumentalismi tai realismi).

### 2.2.2 Esimerkki aksiomatisoidusta teoriasta: virtapiirit

Teoksessaan [2, ss. 152-153] Mario Bunge aksiomatisoi Kirchhoffin-Helmholtzin teorian virtapiireistä seuraavalla tavalla.

*Muodollinen tausta:* tavallinen logiikka (toisen kertaluvun predikaattilogiikka), graafiteoria, matemaattinen alkeisanalyysi.

*Filosofinen tausta:* semantiikka (teoria merkityksestä ja totuudesta) ja metafysiset tieteellisen tutkimuksen oletukset (oletukset ulkoisen maailman riippumattomuudesta ja käsitettävyydestä).

*Prototysikaalinen tausta:* alkeissysteemitheoria, alkeisteoria yleisestä ajasta ja dimensioanalyysi.

*Teorian primitiivit:*  $T$  (aika),  $G$  (graafi),  $V$  (potentiaali),  $e$  (sähkömotorinen voima),  $i$  (virta),  $R$  (ohminen vastus),  $C$  kapasitanssi,  $L$  (itseinduktanssi) ja  $M$  (keskinäinen induktanssi).

*Aksioomat ajasta*

- (1a)  $T$  on lukusuoran väli (MO).
- (1b) Jokainen  $t \in T$  esittää ajan hetkeä. Relatio  $\leq$ , joka järjestää (osittain) joukon  $T$ , esittää ajanhetkien välisen relaation “olla aiempi tai samanaikainen” (SO).

*Aksioomat piiristä*

- (2a)  $\mathcal{G}$  on epättyhjä perhe suunnistettuja graafeja (MO).
- (2b) Jokaista virtapiiriä kohti on olemassa perheen  $\mathcal{G}$  alkio  $G$  siten, että jokaista virtapiirin päätekohtaa tai liitosta vastaa  $G$ :n verteksi ja jokaista virtapiirin komponenttia vastaa  $G$ :n sivu (SO).

*Aksioomat potentiaalista ja virrasta*

- (3a)  $e$  ja  $i$  ovat kuvauksia parilta (sivu,  $t$ ) reaalityluville ja  $V$  on kuvaus parilta (verteksi,  $t$ ) reaalityluville siten, että ne ovat  $t$ -muuttujansa suhteen jatkuvia ja rajoitetusti heilahtelevia<sup>3</sup> (MO).
- (3b) Jos  $n$  on virtapiiriä esittävän graafin  $G \in \mathcal{G}$  sivu ja  $a$  on  $G$ :n verteksi, niin  $e_n(t)$  esittää jännitettä ja  $i_n(t)$  sähkövirran voimakkuutta tuon virtapiirin  $n$ :nnessä komponentissa sekä  $V_n(t)$  esittää sähköistä potentiaalia  $a$ :nnessä päätekohtassa tai liitoksessa (SO).

*Aksioomat parametreista*

- (4a)  $R, C$  ja  $L$  ovat kuvauksia joukolta  $G$  reaalityluville ja  $M$  on symmetrinen neliömatriisi, jonka alkiot ovat kuvauksia joukolta  $G \times G$  reaalityluville; merkitsemme  $n \mapsto R_n$ ,  $n \mapsto C_n$ ,  $n \mapsto L_n$  ja  $(m, n) \mapsto M_{mn}$ , missä  $m$  ja  $n$  ovat graafin  $G$  sivuja. (MO).
- (4b) Jos  $n$  ja  $m$  ovat kaksi virtapiiriä esittävän graafin  $G \in \mathcal{G}$  sivua, niin  $R_n$  esittää  $n$ :nnessä komponentin ohmista vastusta,  $C_n$  kapasitanssia,  $L_n$  induktanssia ja  $M_{nm}$  esittää  $n$ :nnessä ja  $m$ :nnessä komponentin keskinäistä induktanssia (SO).

*Aksioomat laeista*

Esittäköön  $G \in \mathcal{G}$  virtapiiriä.

- (5a) Jokaisessa  $G$ :n verteksissä siihen liittyvien sivujen esittämien komponenttien virranvoimakkuuksien summa on nolla.
- (5b) Jokaiselle  $G$ :n silmukalle sen sivujen esittämien komponenttien potentiaalierojen summa häviää.
- (5c) Jokaiselle  $t \in T$  ja  $G$ :n sivulle  $n$  verteksien  $a$  ja  $b$  välissä pätee

$$L_n i'_n(t) + R_n i_n(t) + \frac{1}{C_n} \int i_n(t) dt + \sum_{n,p} M_{np} i'_p(t) + e_n(t) = V_a(t) - V_b(t).$$

<sup>3</sup>Kuvaus  $f: T \rightarrow \mathbb{R}$  on rajoitetusti heilahteleva, jos jokaisella rajoitetulla välillä  $[s_1, s_2] \subset T$  pätee

$$\sup \left\{ \sum_{i=1}^N |f(t_{i+1}) - f(t_i)| \mid \{t_1, \dots, t_n\} \text{ on välin } [s_1, s_2] \text{ jako} \right\} < \infty.$$

Aksioomat (1)-(4) luonnehtivat yhdeksän primitiiviä kunkin yksinään, muttei niiden suhteita. Tämä luonnehdinta on sekä muodollinen (matemaattinen) että faktuaalinen (fysikaalinen). Edellisen tekevät aksioomat, joita merkittiin (MO) niin kuin muodollinen oletus. Jälkimmäisen tekevät aksioomat, joita merkittiin (SO) niin kuin semanttinen oletus. Niissä esiintyy avainsana *esittää*. Siksi aksiooma (2b) ei väitä, että virtapiiri *on* suunnistettu graafi, vaan että se on virtapiirin edustaja tai esittäjä. Graafi on abstraktio, virtapiiri todellisuuden olio. Tällöin sama graafi voi esittää kokonaista ekvivalenssiluokkaa eri virtapiirejä.

Semanttisten oletustensa takia Bungen aksiomatisointi on tieteellisen realismin mukainen, mutta muuttamalla vain niitä, saataisiin tästä aksiomatisoinnista tieteellisen instrumentalismin mukainen "empiiristen lainalaisuuksien luonnollinen luokittelu". Esimerkkiaksiomatisointi valaissee riittävästi, kuinka hedelmällistä aksiomatisointi on pohdittaessa kysymyksiä, mikä on eri fysiikan teorioiden suhde toisiinsa, voidaanko yksi teoria palauttaa toiseen (reduktio), onko yksi teoria toisen rajatapa...

## 2.3 Mallit ja mallinnus

Kahdessa edellisessä jaksossa tutustuimme Mario Bungen näkemykseen fysiikasta eri teorioiden tilkkutäkinä. Sellainen on myös mallintamisen tuloksena syntyneiden *mallien* joukko. Ne ovat *mini- tai esiteorioita* ja siksi raja teorian ja mallin välillä on epämääräinen.

### 2.3.1 Esimerkki: laskuvarjohyppy

Varsin usein mallin muodostamisessa käytetään hyväksi yleisempiä teorioita. Muodostamme malliksi mallin laskuvarjohypystä. Esittäköön  $y(t)$  hyppääjän etäisyyttä maanpinnasta hetkellä  $t$ ,  $A_0$  hänen poikkileikkauspinta-alansa pituusakselinsa suhteen,  $m$  hänen massaansa ja  $g$  vapaan putoamisliikkeen kiihtyvyyttä. Newtonin toisen lain mukaan

$$my''(t) = -mg + F_\mu, \quad t \in \mathbb{R}_+,$$

missä  $F_\mu$  esittää ilmanvastusta. Kokeellisesta virtausdynamikasta tiedämme, että ilmanvastus on suoraan verrannollinen sekä kappaleen poikkileikkauspinta-alaan  $A(t)$  että vauhdin neliöön eli

$$F_\mu = C_w A(t) y'(t)^2, \quad t \geq 0,$$

missä  $C_w > 0$  on jokin vakio. Hetkellä  $t = 0$  olkoon hyppääjän etäisyys maasta  $y_0$ . Hyppyä kuvaa siten seuraava alkuarvotehtävä

$$\begin{cases} my''(t) = -mg + C_w A(t) y'(t)^2, & t \in \mathbb{R}_+, \\ y(0) = y_0, \quad y'(0) = 0. \end{cases} \quad (2.1)$$

missä

$$A(t) = \begin{cases} A_0, & \text{jos } 0 \leq t \leq t_1, \\ 3A_0, & \text{jos } t_1 < t \leq t_2, \\ 200A_0, & \text{jos } t_2 < t, \end{cases} \quad (2.2)$$

kertoo, että ensin hyppääjä sukeltaa, sitten kääntyy poikittain ja lopulta avaa varjonsa.

### 2.3.2 Esimerkki: loisiminen

Tarkastellaan kahta populaatiota: isännät ja loiset. Esittäkööt  $H(t)$  ja  $P(t)$  niiden määriä hetkellä  $t$ . Arvaamme kokemuksemme perusteella, että isäntien määrän muutosnopeus on suoraan verrannollinen niiden määrään ja että verrannollisuuskerroin on positiivinen kun loisia on vähän ja negatiivinen kun loisia on

paljon. Yksinkertaisin tuollainen kerroin on  $a_{11} - a_{12}P(t)$ , missä  $a_{11}, a_{12} \in \mathbb{R}_+$  ovat vakioita. Vastaavan riippuvuuden arvaamme loisien määrälle, jolloin saamme seuraavan parin differentiaaliyhtälöitä:

$$\begin{cases} H'(t) = (a_{11} - a_{12}P(t))H(t), & t \in \mathbb{R}_+ \\ P'(t) = (-a_{21} + a_{22}H(t))P(t), & t \in \mathbb{R}_+, \end{cases} \quad (2.3)$$

missä  $a_{21}, a_{22} \in \mathbb{R}_+$ . Sanomme yhtälöitä (2.3) lyhyesti malliksi loisimisesta, mutta implisiittisesti liitämme niihin koko joukon oletuksia. Näillä yhtälöillä on nimi: Lotkan-Volterran yhtälöt, ja voisimme myös puhua loisimisen teorias-  
ta. Tarvittaessa osaisimme esittää sen aksiomaattisesti kaikkine aidon teorian osasineen<sup>4</sup>.

Lotkan-Volterran mallilla loisimisesta on kuitenkin yksi huono puoli: sen epälineaaristen differentiaaliyhtälöiden ryhmä on vaikea ratkaista. Muodostamme uuden ja yksinkertaisemman mallin *linearisoimalla* yhtälöt (2.3). Merkitsemme

$$H_\infty = \frac{a_{21}}{a_{22}}, \quad P_\infty = \frac{a_{11}}{a_{12}};$$

nämä vakiot ovat järjestelmän tasapainoarvot eli  $(H_\infty, P_\infty)$  on järjestelmän *stationaarinen tila*, sillä ne toteuttavat yhtälöt (2.3). Tutkimme, mitä tapahtuu, jos  $H(t)$  ja  $P(t)$  ovat lähellä noita vakioita. Kirjoitamme siis

$$H(t) = H_\infty + h(t) \text{ ja } P(t) = P_\infty + p(t),$$

missä  $|h(t)| \ll H_\infty$  ja  $|p(t)| \ll P_\infty$ . Tällöin

$$\begin{cases} h'(t) = a_{12}p(t)(H_\infty + h(t)) \approx a_{12}H_\infty p(t), \\ p'(t) = a_{22}h(t)(P_\infty + p(t)) \approx a_{22}P_\infty h(t). \end{cases}$$

Pudotamme tulot  $p(t)h(t)$  pois, koska ne ovat pieniä<sup>5</sup>. Silloin saamme seuraavan parin *lineaarisia* differentiaaliyhtälöitä:

$$\begin{cases} h'(t) = a_{12}H_\infty p(t), & t \in \mathbb{R}_+, \\ p'(t) = a_{22}P_\infty h(t), & t \in \mathbb{R}_+. \end{cases} \quad (2.4)$$

Helposti löydämme tämän yhtälöparin yleisen ratkaisun<sup>6</sup>

$$\begin{aligned} h(t) &= A \sin(\omega t + \delta), \quad t \geq 0, \\ p(t) &= -\frac{\omega A}{a_{12}H_\infty} \cos(\omega t + \delta) = \frac{\omega A}{a_{12}H_\infty} A \sin\left(\omega t + \delta - \frac{\pi}{2}\right), \quad t \geq 0, \end{aligned}$$

missä  $\omega = \sqrt{a_{11}a_{12}}$  ja vakiot  $A, \delta \in \mathbb{R}$  määräytyvät yksikäsitteisesti esimerkiksi alkuehdoista

$$H(0) = H_0, \quad P(0) = P_0.$$

Näin järjestelmässä esiintyy jaksollista heilahtelua: jos alkutila ei ole stationaarinen tila  $(H_\infty, P_\infty)$  mutta silti lähellä sitä, niin järjestelmä ei koskaan asetu tasapainoon, vaan värähtelee alati sinimuotoisesti stationaarisen tilan ympärillä. Taloustieteilijä puhuisi loisimisen *suhdannevaihtelusta*. Huomaamme, että loisien määrä seuraa isäntien määrää ajan  $\frac{\pi}{2\omega}$  jäljessä.

Voisimme alkaa tutkia, kuinka suuren virheen teimme, kun korvasimme epälineaariset differentiaaliyhtälöt (2.3) lineaarisilla yhtälöillä (2.4).

<sup>4</sup>Ks. Bunge, Mario: Sosiaalisen rakenteen käsite, "Yhteiskuntatieteiden eksakti metodologia", toim. R. Tuomela. Gaudeamus, Helsinki, 1975.

<sup>5</sup>Jäljelle jääneet termit ovat ensimmäisen asteen polynomeja pienistä luvuista  $p(t)$  ja  $q(t)$ , joten sanomme aproksimaatiomme olevan ensimmäistä kertalukua.

<sup>6</sup>Muista, että  $\sin(\omega t + \delta) = \cos \delta \sin \omega t + \sin \delta \cos \omega t$ .

## 2.4 Dimensioanalyysiä

Jos tunnemme johonkin fysikaaliseen järjestelmään liittyvät suureet, tiedämme implisiittisesti jo jotain siitä, miten nämä suureet riippuvat toisistaan. Menetelmää, jolla tämä tieto saadaan esiin, kutsutaan *dimensioanalyysiksi*. Se on hyödyllinen ja laajasti käytetty menetelmä erityisesti virtausdynamiikassa, mutta sitä voidaan soveltaa kaikilla tieteenoaloilla, joissa käytetään *suureita*. Toisaalta dimensioanalyysi —oikeammin käsitteet suure, mittaluku, mittayksikkö ja dimensio— ovat yksi tieteellisen teorian osa, jonka mukanaolo on tehtävä eksplisiittiseksi viimeistään silloin kun teoria aksiomatisoidaan. Tässä jaksossa tutustumme dimensioanalyysiin sekä näppäränä mallintamisen apukeinona että tieteellisen teorian osasena.

### 2.4.1 Esimerkki: yksinkertainen heiluri

Tarkastelemme heiluria, jonka heilahdusaika on  $P$  sekuntia, langan pituus on  $l$  metriä ja punnuksen massa on  $m$  kilogrammaa. Heiluri on tasaisessa gravitaatiokentässä, jossa vapaan putoamisliikkeen kiihtyvyys on  $g$ . Kun heilahduskulma on pieni, oletamme, että mitkään muut tekijät eivät heilahdusaikaan vaikuta eli

$$P = f(m, l, g), \quad (2.5)$$

missä  $f$  on jokin tuntematon kolmen reaalimuuttujan reaaliarvoinen funktio. Käytämme massan, ajan ja pituuden *mittayksikköinä* SI-yksiköjä kg, s ja m. Mutta se on vain mielivaltainen valintamme: yhtä hyvin voisimme käyttää näiden *perussuureiden* yksikköinä leiviskää, lasia ja syltä tai mitä hyvänsä niiden monikertaa. Heilurin käyttäytyminen ei muutu, vaikka vaihtaisimme perussuureiden mittayksiköt toisiksi. Jos esimerkiksi massan mittayksikkö on kilogramman sijasta leiviskä, on massan *mittaluku* jaettava kymmenellä: Villen massa on 60 kg eli 6 leiviskää.

Vaihdetaan perussuureiden massa, pituus ja aika mittayksiköt monikerroiksi  $\mu$  kg,  $\lambda$  m ja  $\tau$  s, missä  $\lambda, \mu, \tau > 0$ . Silloin heilahdusajan mittaluku  $P$  on jaettava luvulla  $\tau$ , punnuksen massan mittaluku  $m$  luvulla  $\mu$  ja langan pituuden mittaluku  $l$  luvulla  $\lambda$ . Vakion  $g$  mittayksikkö on  $\text{m/s}^2$  SI-yksiköissä, joten sen mittaluku on jaettava luvulla  $\lambda/\tau^2$ . Koska mittayksikköjen muuttaminen ei saa muuttaa heilurin käyttäytymistä, ei sitä kuvaavan yhtälönkään tule muodoltaan muuttua. Siis

$$\tau^{-1}P = f(\mu^{-1}m, \lambda^{-1}l, \lambda^{-1}\tau^2g) \text{ kaikilla } \lambda, \mu, \tau > 0. \quad (2.6)$$

Valitaan nyt  $\mu = m$ ,  $\lambda = l$  ja  $\tau = \sqrt{l/g}$ , jolloin

$$P\sqrt{\frac{g}{l}} = f(1, 1, 1) = C \text{ eli } P = C\sqrt{\frac{l}{g}}, \quad (2.7)$$

missä  $C$  on jokin vakio. Sen arvo jäi nyt määrittämättä, mutta jos mittaamme yhden yksinkertaisen heilurin heilahdusajan ja langan pituuden, saamme kaavasta (2.7) laskettua vakion  $C$  ja niin meillä on yleinen kaava kaikille yksinkertaisille heilureille. Tässä tapauksessa tiedämme heilurin liikeyhtälön ja osaamme sen ratkaistakin (jolloin saamme, että  $C = 2\pi$ ), mutta merkittävää on, että (2.7) saatiin *ilman* tietoa tai oletuksia Newtonin laeista.

### 2.4.2 Suureet ja mittayksiköt

Jokainen fysiikan teoria  $T$  sisältää (äärettömän) joukon  $Q_T$  *suureita*, josta osa riippuu toisistaan sellaisten yhtälöiden kuin  $F = ma$  tai  $v = \frac{ds}{dt}$  kautta. Teorian  $T$  *perussuureiden* joukoksi  $Q_T^0$  sanotaan niiden teoriassa  $T$  määrittelemättömien suureiden joukkoa, joiden avulla muut teorian  $T$  suureet määritellään. Jokainen

fysikaalinen järjestelmä luonnehditaan reaaliarvoisten funktioiden avulla. Jos  $A_1, A_2, \dots, A_n$  ovat fysikaalisen järjestelmän piirteitä (hiiliatomeita, magneettikenttiä, nopeuksia, ...), niin on olemassa epätyhjä joukko  $U_P$  ja kuvaus

$$P: A_1 \times A_2 \times \dots \times A_n \times U_P \rightarrow \mathbb{R}_+ \quad (2.8)$$

siten että tämä kuvaus kuvaa jokaisen joukon  $A_1 \times \dots \times A_n$  alkion tiettyä ominaisuutta. Kuvausta  $P$  sanotaan *mittaluvuksi* ja joukkoa  $U_P$   *$P$ -mittayksiköiden* joukoksi. Funktion  $P$  arvo riippuu mielivaltaisesta mittayksiköstä  $u \in U_P$  siten, että

$$P(a, \lambda^{-1}u) = \lambda P(a, u) \text{ kaikilla } u \in U_P, \lambda > 0. \quad (2.9)$$

Esimerkiksi klassisessa mekaniikassa kappaleen massaa esittää *kahden* muuttujan kuvaus

$$m: B \times U_M \rightarrow \mathbb{R}_+, \quad (2.10)$$

missä  $B$  on massakappaleiden joukko ja  $U_M$  massayksiköiden joukko.

Suureen ei tarvitse olla mitattavissa: riittää, että sitä kuvaa positiiviarvoinen funktio, jolla on ominaisuudet (2.8)-(2.9) ja jota kutsutaan mittaluvuksi. Lukija pohtinee, mikä tavallisissa ilmauksissa  $F = ma$  ja  $m = 60$  kg esiintyvät  $m$  tai  $kg$  ovat. Lyhin vastaus on sanoa, että  $F = ma$  on mittalukuyhtälö, joka sitoo voiman, massan ja kiihtyvyyden mittaluvut. Yksiköt ovat käsitteellisiä konstruktioita: voimme sopia, että kilogramma *on*  $(1, 0, 0)$  ja leiviskä *on*  $(10, 0, 0)$ ; ne ovat tällöin siis vektoriavaruuden  $\mathbb{R}^3$  alkioita.

### 2.4.3 Dimensiokuvaus

Olkoot  $M_1, M_2, \dots, M_m$  teorian  $T$  perussuureet ja  $S$  jokin sen suure siten, että mittaluvuille pätee

$$S = f(M_1, M_2, \dots, M_m), \quad (2.11)$$

missä  $f: \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ . *Bridgmanin aksiooma* sanoo, että suureen  $S$  kahden eri arvon mittalukujen suhde ei muutu, kun niitä vastaavien perussuureiden mittayksiköt vaihdetaan  $\lambda_i^{-1}$ -kertaisiksi eli

$$\frac{S}{S'} = \frac{f(M_1, \dots, M_m)}{f(M'_1, \dots, M'_m)} = \frac{f(\lambda_1 M_1, \dots, \lambda_m M_m)}{f(\lambda_1 M'_1, \dots, \lambda_m M'_m)} \quad (2.12)$$

kaikilla  $\lambda_i > 0, i = 1, 2, \dots, m$ . Bridgmanin aksiooma siis sanoo esimerkiksi, että Villen ja Kallen massojen mittalukujen suhde on  $\frac{1}{2}$ , käytettiinpä sitten massan mittayksikkönä kilogrammaa, paunaa tai leiviskää.

Seuraavan lauseen avulla määrittelemme suureen *dimension*.

**Lause 2.4.1** *Olkoot  $M_1, \dots, M_m$  perussuureita ja  $S$  kaavalla (2.11) annettu suure sekä  $f: \mathbb{R}_+^m \rightarrow \mathbb{R}_+$  Lebesgue-integroituva jokaisessa määrittelyjoukkonsa kompaktissa osajoukossa eli  $f \in L^1_{loc}(\mathbb{R}_+^m)$ . Tällöin*

$$S = C \prod_{i=1}^m M_i^{a_i}, \text{ missä } a_i \in \mathbb{R}, C > 0, i = 1, \dots, m, \quad (2.13)$$

*jos ja vain jos Bridgmanin aksiooma pätee.*

Tämä lause voidaan todistaa seuraavan tuloksen avulla: Jos  $g \in L^1_{loc}(\mathbb{R}_+)$  ja

$$g(xy) = g(x)g(y) \text{ kaikilla } x, y \in \mathbb{R}_+,$$

niin silloin

$$g(x) = x^\alpha, \text{ kaikilla } x \in \mathbb{R}_+,$$

missä  $\alpha \in \mathbb{R}$ .

**Määritelmä 2.4.2** *Olkoon Bridgmanin aksiooma voimassa sekä suure  $S$  annettu yhtälöllä (2.11), missä  $f \in L^1_{loc}(\mathbb{R}^m_+)$ . Suureen  $S$  dimensio on*

$$[S] := \prod_{i=1}^m M_i^{a_i}$$

*missä esiintyvät reaalityluvut  $a_i$  antaa kaava (2.13).*

Mekaniikassa perussuureina pidetään tavallisimmin massaa  $M$ , pituutta  $L$  ja aikaa  $T$  jolloin massan dimensio on  $M = M^1 L^0 T^0$ , pituuden  $L = M^0 L^1 T^0$ , keskinopeuden  $L^1 T^{-1}$ , keskihihtyvyyden  $L^1 T^{-2}$ , keskivoiman  $M L T^{-2}$ , ... Kaikki mekaniikan suureet eivät kuitenkaan ole niin yksinkertaisia kuin  $S$  kaavassa (2.11). Esimerkiksi kappaleen nopeuden mittaluku  $v(t)$  hetkellä  $t$  on derivaatta

$$v(t) = x'(t) = \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{(t+h) - t},$$

missä  $x(t)$  on kappaleen sijainnin mittaluku hetkellä  $t$ . Toisaalta kertymäsuureet ovat vastaavasti integraaleja muista suureista. Integraalikin on raja-arvo, joten määrittelemme kaikkien integraalin tai derivaatan avulla määriteltyjen suureiden dimensiot samoiksi kuin sen suureen dimensio, josta vastaava rajankäynti tehdään. Siten esimerkiksi hetkellisen nopeuden dimensio on  $L T^{-1}$ . Fysiikan teorioissa ei esiinny muunlaisia suureita kuin perussuureita, niistä kaavan (2.11) mukaisesti johdettuja suureita tai näistä rajankäynnillä johdettuja suureita. Siten meillä on kullekin fysiikan teorialle  $T$  ja sen perussuureiden joukolle  $Q_T^0 = \{M_1, \dots, M_m\}$  yksikäsitteinen kuvaus  $[\cdot]: Q_T \rightarrow \langle M_1, \dots, M_m \rangle$ ,  $S \mapsto [S]$ , missä

$$\langle M_1, \dots, M_m \rangle := \left\{ \prod_{i=1}^m M_i^{b_i} \mid b_i \in \mathbb{R} \right\}.$$

Sanomme kuvausta  $[\cdot]$  *dimensiokuvaukseksi*. Jos suurelle  $S$  pätee  $[S] = 1$ , niin sanomme, että  $S$  on *dimensioton*.

**Lause 2.4.3** *Dimensiofunktio toteuttaa:*

1.  $[C] = 1$  kaikilla  $C \in \mathbb{R}_+$ .
2.  $[S^a R^b] = [S]^a [R]^b$  kaikilla  $a, b \in \mathbb{R}$ ,  $S, R \in Q_T$ .
3. Joukko  $\langle M_1, \dots, M_m \rangle$  on  $m$ -ulotteinen vektoriavaruus seuraavien laskutoimitusten suhteen:

$$\prod_{i=1}^m M_i^{a_i} + \prod_{i=1}^m M_i^{b_i} := \prod_{i=1}^m M_i^{a_i + b_i} \quad \text{ja} \quad \lambda \prod_{i=1}^m M_i^{a_i} := \prod_{i=1}^m M_i^{\lambda a_i}, \quad \lambda \in \mathbb{R}.$$

Ihmisen aistiminen noudattaa psykofyysistä lainalaisuutta, jonka mukaan aistimuksen voimakkuus on suoraan verrannollinen ärsykkeen logaritmiin. Joitakin siihen perustuvia mitta-asteikoita käytetään yhä (äänenvoimakkuuden ilmoittaminen desibeleissä ja tähtien kirkkauden magnitudeissa). Nämä mittaluvut eivät toteuta Bridgmanin aksioomaa kuten ei myöskään lämpötilan Celsius- tai Fahrenheit-asteikko. Kaikille näille suureille on toki olemassa asteikot, joilla mittaluku toteuttaa Bridgmanin aksiooman.

#### 2.4.4 Dimensiohomogeenisuus

Olko  $M_1, \dots, M_m$  teorian  $T$  perussuureet ja  $X_1, \dots, X_n$  sen suureita. Suureiden  $X_1, \dots, X_n$  *dimensiomatriisiksi* perussuurejärjestelmässä  $M_1, \dots, M_m$  sanotaan matriisia  $A = (a_{ij})$ , jolle

$$\prod_{j=1}^m M_j^{a_{ij}} = [X_i], \quad i = 1, \dots, n.$$

*Dimensiohomogeenisuuden periaate* sanoo, että mittalukuyhtälön muoto ei muutu, vaikka perusuureita  $M_1, \dots, M_m$  vastaavat yksiköt muutetaan  $\lambda_i^{-1}$ -kertaisiksi,  $\lambda_i > 0$ ,  $i = 1, \dots, m$ . Jos siis suureet  $X_1, \dots, X_n$  sitoo yhtälö

$$f(X_1, \dots, X_n) = 0,$$

niin

$$f\left(\prod_{j=1}^m \lambda_j^{a_{1j}} X_1, \dots, \prod_{j=1}^m \lambda_j^{a_{nj}} X_n\right) = 0 \text{ kaikilla } (\lambda_1, \dots, \lambda_m) \in \mathbb{R}_+^m.$$

Dimensiohomogeenisuuden periaate muistuttaa Einsteinin suhteellisuusteoriaa: luonnonlakien muotojen on oltava invariantteja yksikönmuutosten suhteen. Ilmiö ei saa riippua mielivaltaisista yksiköistä.

Dimensiohomogeenisuuden periaatteen mukaan yhteenlaskettavien suureiden dimensioiden on oltava samat. Samoin transsendenttisten funktioiden muutujien on oltava dimensiottomia, sillä niiden funktioiden potenssisarjat on päätymättömiä.

### 2.4.5 Buckinghamin $\pi$ -teoreema

Olkoot  $M_1, \dots, M_m$  teorian  $T$  perussuureet ja  $X_1, \dots, X_n$  sen suureita. Joukko  $\langle X_1, \dots, X_n \rangle$  on  $n$ -ulotteinen vektoriavaruus lauseen 2.4.3 kohdan 3 kaavoja vastaavien laskutoimitusten suhteen. Sen osajoukko  $\{P_1, \dots, P_p\}$  on lineaarisesti riippumaton, jos ja vain jos

$$\prod_{i=1}^p P_i^{\lambda_i} = 1 \implies \lambda_1 = \lambda_2 = \dots = \lambda_p = 0. \quad (2.14)$$

Dimensioanalyysissä dimensiottomat potenssitulot ovat tärkeitä. Lauseen 2.4.3 mukaan potenssitulo  $\prod_{i=1}^n X_i^{k_i}$ ,  $k = (k_1, \dots, k_n) \in \mathbb{R}^n$ , on dimensioton, jos ja vain jos

$$1 = \left[\prod_{i=1}^n X_i^{k_i}\right] = \prod_{i=1}^n [X_i]^{k_i} = \prod_{i=1}^n \prod_{j=1}^m M_j^{a_{ij}k_i} = \prod_{j=1}^m M_j^{\sum_{i=1}^n k_i a_{ij}}.$$

Tämä toteutuu täsmälleen silloin, kun matriisitulo

$$kA = \left(\sum_{l=1}^n k_l a_{li}\right) = 0. \quad (2.15)$$

Lineaarialgebrasta tiedämme, että lineaarikuvauksen  $\tilde{A}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $k \mapsto kA$  ydin on  $(n-r)$ -ulotteinen vektoriavaruus, missä  $r$  on matriisin  $A$  aste (eli sen suurimman kääntyvän neliömäisen alimatriisin rivien määrä). Täten suureista  $X_1, \dots, X_n$  voidaan muodostaa  $p = n-r$  kpl lineaarisesti riippumattomia dimensiottomia potenssituloja  $\pi_1, \dots, \pi_p$ . Ne eivät ole yksikäsitteisiä: mikä hyvänsä lineaarikuvauksen  $\tilde{A}$  ytimen kanta kelpaa antamaan niiden potenssitulojen eksponentit.

Nyt voimme esittää dimensioanalyysin tärkeimmän lauseen.

**Lause 2.4.4 (Buckinghamin  $\pi$ -teoreema)** *Olkoot  $M_1, \dots, M_m$  teorian perussuureita ja  $X_1, \dots, X_n$  suureita siten, että dimensiokuvaus  $[\cdot]$  on määriteltä joukossa  $\langle X_1, \dots, X_n \rangle$ . Merkitään  $r$ :llä dimensiomatriisin astetta ja olkoon  $f: \mathbb{R}_+^n \rightarrow \mathbb{R}$ . Tällöin on olemassa  $p = n-r$  kpl lineaarisesti riippumattomia dimensiottomia potenssituloja  $\pi_1, \dots, \pi_p$  ja kuvaus  $F: \mathbb{R}_+^p \rightarrow \mathbb{R}$  siten että suureet  $X_1, \dots, X_n$  sitoo dimensiohomogeeninen yhtälö*

$$f(X_1, \dots, X_n) = 0 \quad (2.16)$$

täsmälleen silloin, kun

$$F(\pi_1, \dots, \pi_p) = 0. \quad (2.17)$$

Jos  $r = n$  ja (2.16) on dimensiohomogeeninen, eivät suuret  $X_1, \dots, X_n$  riipu toisistaan. Jos  $r < n$ ,  $f(\cdot, X_2, \dots, X_n)$  on injektio ja  $X_1$  esiintyy potenssitulossa  $\pi_1$ , niin  $F(\cdot, \pi_2, \dots, \pi_p)$  on injektio.

Lineaarisesti riippumattomien potenssitulojen olemassaolo on jo todettu, joten riittää todistaa (2.16)  $\implies$  (2.17), sillä (2.17)  $\implies$  (2.16 on triviaali. Jos  $r = n$ , dimensiohomogeenisuus vaatii, että  $f$  on vakiofunktio eli suuret  $X_1, \dots, X_n$  eivät riipu toisistaan. Tapauksessa  $r < n$  sivuutamme todistuksen, joka ei ole kovin pitkä eikä vaikea (ks. matriisien ominaisuuksia käyttävä alkeellinen todistus Arkhimedes 1/1985, ss. 45-47, dualiteetin käsitteeseen perustuva todistus Arkhimedes 2/1985, ss. 117-119.)

### 2.4.6 Esimerkkejä

Sovelletaan Buckinghamin  $\pi$ -teoremaa yksinkertaiseen heiluriin. Perussuureet ovat siis  $M$ ,  $L$  ja  $T$  ja relevanteiksi otaksutut suuret  $P$ ,  $l$ ,  $m$  ja  $g$ . Muodostamme dimensiomatriisin  $A$ :

$$\begin{array}{c|ccc} A & M & L & T \\ \hline P & 0 & 0 & 1 \\ m & 1 & 0 & 0 \\ l & 0 & 1 & 0 \\ g & 0 & 1 & -2 \end{array} \quad \text{eli} \quad A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix}.$$

Dimensiomatriisin aste  $r = 3$  ja toisaalta  $n = 4$ , joten Buckinghamin  $\pi$ -teoreman nojalla on olemassa vain yksi dimensioton potenssitulo  $\pi_1$  siten, että järjestelmää kuvaa yhtälö  $F(\pi_1) = 0$ . Koska  $F$  ei ole vakio, niin  $\pi_1 = C^2$ , missä  $C > 0$  on jokin vakio. Potenssitulon

$$\pi_1 = \prod_{i=1}^4 X_i^{k_i} = P^{k_1} m^{k_2} l^{k_3} g^{k_4}$$

eksponentit  $k_i$  ratkaisemme yhtälöstä  $kA = 0$  eli  $A^t k^t = 0$  ( $t$  niin kuin transpoosi). Siis

$$\begin{cases} k_2 & & & = 0 \\ & k_3 & +k_4 & = 0 \\ k_1 & & -2k_4 & = 0. \end{cases}$$

Ratkaisut ovat  $k = (2\lambda, 0, -\lambda, \lambda)$ ,  $\lambda \in \mathbb{R}$ , joista valitsemme  $k = (-2, 0, -1, 1)$ . Siten saamme jälleen yksinkertaisen heilurin heilahdusajaksi

$$P = C \sqrt{\frac{l}{g}}.$$

Tarkastellaan seuraavaksi taloustieteellistä ongelmaa ([4, ss. 56-59]). Tehtailija pohtii, miten paljon hän tarvitsee työvoimaa aikayksikköä kohti. Merkitään tätä suuretta  $N$ . Perussuureet ovat  $M$ ,  $R_a$ ,  $R_k$ ,  $R_e$  ja  $T$ . Niistä  $M$  vastaa rahaa ja sen mittayksikkö on markka,  $R_a$  vastaa tehtyä työtä (miestyötunti),  $R_k$  vastaa tehtaaseen sijoitettua kiinteää pääomaa (laitteet, rakennukset),  $R_e$  vastaa tuotetta (pyykkipoika) ja  $T$  vastaa aikaa (tunti). Tehtailija arvioi, että työvoiman tarve riippuu vain palkkatasosta  $w$ , jonka dimensio  $[w] = MR_a^{-1}$ , tuotteen yksikkö hinnasta  $p$ ,  $[p] = MR_e^{-1}$ , kiinteästä pääomasta  $K$ ,  $[K] = R_k$  ja tuotantotekniikan indeksistä  $c_\alpha$ , joka esiintyy yhtälössä

$$u = c_\alpha N^\alpha K^{1-\alpha}, \quad \alpha \in ]0, 1[.$$

missä  $u$  on tuotantovauhti,  $[u] = R_e T^{-1}$ , ja  $N$  on työllistäminen,  $[N] = R_a T^{-1}$ . Itse asiassa tehtailija olettaa, että yhtälö

$$f(N, w, p, K, c_\alpha) = 0$$

on dimensiomogeeninen. Muodostetaan jälleen dimensiomatriisi  $A$ :

| $A$        | $M$ | $R_a$     | $R_k$        | $R_e$ | $T$            |
|------------|-----|-----------|--------------|-------|----------------|
| $N$        | 0   | 1         | 0            | 0     | -1             |
| $w$        | 1   | -1        | 0            | 0     | 0              |
| $p$        | 1   | 0         | 0            | -1    | 0              |
| $K$        | 0   | 0         | 1            | 0     | 0              |
| $c_\alpha$ | 0   | $-\alpha$ | $\alpha - 1$ | 1     | $\alpha - 1$ . |

Koska  $n = 5$  ja dimensiomatriisin aste on 4, on olemassa vain yksi dimensioton potenssitulo  $\pi_1$  ja pätee  $F(\pi_1) = 0$  eli  $\pi_1 = C^2$ , missä  $C > 0$  on vakio. Ratkaisemalla yhtälö  $kA = 0$  saadaan

$$\pi_1 = \frac{N}{K} \left( \frac{w}{pc_\alpha} \right)^{\frac{1}{1-\alpha}} = C \text{ eli } N = CK \left( \frac{pc_\alpha}{w} \right)^{\frac{1}{1-\alpha}}.$$

Työvoiman kysynnällä on siis ominaisuudet

$$N \propto K; N \rightarrow 0, \text{ kun } \frac{p}{w} \rightarrow 0; N \rightarrow \infty, \text{ kun } \frac{p}{w} \rightarrow \infty.$$

Koko kansantalouteen sovellettuna tämä tulos näyttäisi osoittavan, että investoinnit tuotantolaitoksiin ja palkkatason lasku vähentävät tehokkaasti työttömyyttä.

Olisimme voineet käyttää vain kahta perussuuretta:  $M$  ja  $T$ . Silloin olisimme samastaneet dimensioltaan neljä eri arvosuuretta  $M$ ,  $R_a$ ,  $R_e$  ja  $R_k$ . Siinä tapauksessa dimensiottomia potenssituloja olisi saatu ainakin kolme ja niin olisimme saaneet epämääräisemmän tuloksen

$$F(\pi_1, \pi_2, \pi_3) = 0.$$

Tämän vuoksi kannattaa käyttää mahdollisimman monta eri perussuuretta. Joissakin fysiikan ongelmissa on jopa mahdollista tunnistaa esimerkiksi kahta eri lajin pituutta: luotisuoran suuntainen ja sitä vastaan kohtisuora pituus.

### 2.4.7 Huomioita dimensioanalyysistä

Dimensioanalyysin teoriaa edellä luonnostellessamme emme vaatineet suureilta mitattavuutta: riitti että niillä on vain joitakin mitattavien suureiden ominaisuuksia. On ilmeistä, että dimensiomogeenisuus ei rajoitu vain fysiikkaan, vaan dimensiomogeenisia yhtälöitä voidaan rakentaa minkä tahansa inhimillisen tiedon piiriin kuuluvien suureiden välille, kunhan vain noiden suureiden dimensioiden tunnetaan. Tehtailijaesimerkissämme yhtenä suureena oli raha, joka on arvon mitta. Voisimme kuvitella dimensioanalyysiä sovellettavan muihinkin arvosuureisiin kuten kauneuteen, pyhyteen...

Millä edellytyksillä sitten dimensioanalyysin antamat tulokset ovat tosia tai käyttökelpoisia? Dimensioanalyttisen päättelyn yleinen kulku on seuraava:

(1) Järjestelmän oleelliset suureet ovat  $X_1, \dots, X_n$ , jotka sitoo dimensiomogeeninen yhtälö.

(2) Perussuurejärjestelmässä  $\{M_1, \dots, M_m\}$  suureen  $X_i$  dimensio on  $[X_i] = \prod_{j=1}^m M_j^{a_{ij}}$ ,  $i = 1, \dots, n$ .

(DA)

(3) Järjestelmää kuvaa yhtälö  $F(\pi_1, \dots, \pi_p) = 0$ .

Kyseessä on tavallinen deduktiivinen päättely, jonka premissit ovat (1) ja (2): kun on nähty ongelman kannalta oleelliset suureet, valittu perussuureet ja määritetty oleellisten suureiden dimensioiden valitussa perussuurejärjestelmässä, voidaan johtopäätöksen (3) laskeminen jättää yksinkertaisen tietokoneohjelman tehtäväksi. Todellinen vaikeus piilee premissien löytämisessä: jos lähtökohdat ovat päättömiä, ei johtopäätökseltä pidä odottaa sen enempää. Tarvitaan hyvä fyysikko, taloustieteilijä . . . , jonka intuitio ja kokemus opastavat hänet arvaamaan kohtuullisen hyvin perussuureet ja oleelliset suureet dimensioineen.

Tehtailijaesimerkissä tarkasteltiin mikrotalouden ongelmaa ja tarjottiin sen ratkaisua suoraan makrotalouteen: työttömyys voidaan poistaa sijoituksin tuotantolaitoksiin ja laskemalla palkkoja. Makrotaloudessa kuitenkin palkkatason lasku vähentää kuluttajien ostovoimaa, mikä puolestaan vähentää tuotteen kysyntää, mikä laskisi sen hintaa. Näin ollen tulos, joka pätee yhdelle tehtaalle, ei välttämättä yleistykään koko kansantalouteen kaikkine vuorovaikutuksineen ja takaisinkytkentöineen.

## 2.5 Aika, avaruus, järjestelmä, tila, kenttä

Käsitlemme tässä jaksossa lyhyesti taloudellisissa malleissa toistuvasti esiintyviä käsitteitä. Tavallisesti mallia muodostettaessa ei vaivauduta yhä uudestaan ja uudestaan toistamaan niiden luonnehdintoja.

### 2.5.1 Aika ja avaruus

Ajan sisällyttämme malleihimme silloin, kun haluamme kuvata tai ottaa huomioon muutoksia ajan suhteen. Silloin mallimme ovat *dynaamisia*. Asetamme seuraavat aksioomat ajasta.

- (A.1) Joukko  $T$  on epätyhjä lukusuoran osajoukko.
- (A.2) Jokainen  $t \in T$  esittää ajan hetkeä. Relaatio  $\leq$ , joka järjestää (osittain) joukon  $T$ , esittää ajanhetkien välisen relaation "olla aiempi tai samanlainen".

Näistä aksioomista ensimmäinen on muodollinen oletus eli se kertoo, millainen abstrakti olio  $T$  on. Toinen aksiooma on semanttinen oletus, joka sitoo joukon  $T$  alkioit todellisuuden ajanhetkiin. Voi olla, että  $T = \mathbb{R}$ , jolloin aika on tuttu klassisen mekaniikan jatkuva absoluuttinen aika. Voi myös olla  $T = \mathbb{R}_+ = [0, \infty[$ , jolloin aika alkaa siitä hetkestä, jota esittää reaalityluku 0. Jos taas  $T = ]-\infty, 10]$ , niin aika loppuu hetkeen, jota esittää luku 10. Voi olla myös  $T = \mathbb{Z}$  tai  $T = \mathbb{N}$ , jolloin aika on *diskreetti*. Tämä tapaus on tavallinen ja luonnollinen taloudellisissa malleissa etenkin silloin, kun yrittäjä tekee ratkaisunsa laskettuaan päivän kassan. Ääritapaus on se, että joukko  $T$  sisältää vain yhden alkion. Silloin mallissa ei tarkastella muutoksia ajan suhteen lainkaan.

Joskus tarvitsemme ajankäsitteen, jolla on enemmän rakennetta. Seuraava aksiooma lausuu ajan *puoliryhmäominaisuuden*:

- (A.3) Jos  $s, t \in T$ , niin  $s + t \in T$ .

Tästä aksioomasta seuraa ikuisen tulevaisuuden olemassaolo eli joukko  $T$  ei voi olla ylhäältä rajoitettu.

Newtonin mekaniikassa avaruutta esittää  $\mathbb{R}^3$ . Joskus meille riittävät sen osajoukot: mallimme voi olla kaksi- tai yksiulotteinen tai diskreetti avaruuden suhteen (katuosoitteetkin ovat diskreettejä).

Suhteellisuusteorioissa aika ja avaruus eivät ole toisistaan riippumattomia: aika-avaruutta esittää neliulotteinen differentioituva monisto. Tosin jo Newtonin gravitaatioteorian yhdessä muunnoksessa (Newtonin-Poissonin gravitaatioteoria) esitetään avaruutta kolmiulotteisella differentioituvalla monistolla. Voimme

toki muodostaa taloudellisia malleja, joissa käytetään vaikkapa yleisen suhteellisuusteorian kaltaisia rakenteita aika-avaruuden kuvaamiseen: tällainen malli sisältäisi äärellisen informaation siirtymisnopeuden.

### 2.5.2 Järjestelmä ja tila

Järjestelmän tilaa ajanhetkellä, jota esittää  $t \in T$ , esitetään oliolla  $y(t)$ . Tapauksesta riippuen  $y(t)$  on reaalityyppinen tai monimutkaisempi matemaattinen olio. Meillä on siis joukolla  $T$  määritelty kuvaus. Esimerkiksi sijoittajan tilaa ajanhetkellä  $t$  voisi esittää pari  $(m(t), s(t))$ , missä  $m(t)$  on hänen pankkitilinsä saldo markoissa ja  $s(t)$  hänen omistamiensa Olvi Oyj:n osakkeiden määrä. Järjestelmän kehittymistä kuvaavat lait muotoillaan sen tilaa esittävän funktion avulla. Tällöin pyrimme klassisen mekaniikan ja peittävän lain mallin esimerkin mukaisesti siten, että tila jollakin hetkellä määräisi yksikäsitteisesti järjestelmän myöhemmät tilat. Näin ollen klassisessa massipistemekaniikassa tila ei ole vain järjestelmän kappaleiden sijainnit vaan myös niiden nopeudet tai impulssit. Järjestelmän, joka koostuu  $n$ :stä kappaleesta, liikeyhtälö massipistemekaniikan mukaan on muotoa

$$y'(t) = Ay(t), \quad t \in \mathbb{R},$$

missä  $y: \mathbb{R} \rightarrow \mathbb{R}^{2n}$ . Kun siihen liitetään alkuehto  $y(t_0) = y_0$ , on ratkaisu – ainakin yleensä – olemassa ja yksikäsitteinen. Huomaa kuitenkin, että jo seuraavan yksinkertaisen alkuarvotehtävän ratkaisu ei ole yksikäsitteinen:

$$u'(t) = \sqrt{u(t)}, \quad t \in \mathbb{R}_+, \quad u(0) = 0.$$

### 2.5.3 Kenttä

Tavallisia fysiikan kenttiä ovat gravitaatiokenttä ja magneettikenttä. *Kenttää* esittää paikan ja ajan funktio eli sen muuttujina on ainakin  $\mathbf{x}$  ja  $t$ . Se voi olla arvoltaan reaalin, kompleksinen, vektori, matriisi...

Yksi taloustieteen kenttäsuure on maan yksikköhinta  $p$ , jonka mittayksikkö on  $\text{mk}/\text{m}^2$  ja joka riippuu sekä maa-alan sijainnista että ajasta. Voisimme suoraan tarjota kaksiulotteista lämpöyhtälöä

$$\frac{\partial p}{\partial t}(x, t) - \kappa \sum_{i=1}^2 \frac{\partial^2 p}{\partial x_i^2}(x, t) = 0$$

kuvaamaan maan yksikköhinnan kehittymistä kaupungin läheisyydessä ( $\kappa$  on jokin positiivinen vakio).

## 2.6 Harjoitustehtäviä

1. Tarkastellaan kauppiaan toimintaa. Hän myy yhtä tuotetta vapaasti valitsemallaan yksikköhinnalla  $p(t)$  ajanhetkellä  $t$ . Kuluttajien kysyntänopeus  $q_d(t)$  on jokin funktio tuotteen yksikkö hinnasta, siis  $q_d(t) = f(p(t))$ . Kauppias ostaa rajoittamattomasti tuotetta yksikkö hinnalla 1 juuri sillä nopeudella, jolla hän saa sitä myytyä, siis  $q_s(t) = q_d(t)$ . Kauppiaan voitto aikayksikössä on siten

$$\Pi = \Pi(p(t)) = p(t)q_d(t) - q_s(t) = p(t)f(p(t)) - f(p(t)). \quad (2.18)$$

Kauppias *ei tiedä*, millainen funktio  $f$  on, mutta hän osaa approksimoida  $f$ :n derivaattaa toteutuneesta kuluttajien käyttäytymisestä:

$$f'(p(t)) \approx \frac{q_d(t) - q_d(t-h)}{p(t) - p(t-h)}, \quad h \approx 0.$$

Näin kauppias osaa laskea voittonsa derivaatan yksikköhinnan suhteen ja niin hän laskee tai nostaa yksikköhintaa, jos se parantaa hänen voittoa. Yksinkertaisin tämän lausuva kaava on

$$p'(t) = \Pi'(p(t)), \quad t > 0. \quad (2.19)$$

Tämän differentiaaliyhtälön oikealla puolella voisi vielä olla uloimmaisena funktiona mikä hyvänsä aidosti kasvava funktio  $g: \mathbb{R} \rightarrow \mathbb{R}$ , jolle  $g(0) = 0$ .

Olko  $f(p) = 3 - p$  ja  $p(0) = 0$ . Ratkaise  $p$  yhtälöstä (2.19) ja piirrä sen kuvaaja. Laske optimaalinen yksikköhinta  $p^*$  lausekkeesta (2.18). Mitä tapahtuu, kun  $t \rightarrow \infty$ ?

2. Tarkastellaan taloutta, jossa tuotetaan ja kulutetaan tasan yhtä nopeasti pilaantuvaa tuotetta. Tuottajat voivat päättää tuotantonopeutensa  $q_s(t)$  ja yksikköhinnan  $p(t)$  hetkellä  $t$ . Kysyntänopeus  $q_d(t)$  on jokin funktio tuotteen yksikköhinnasta  $p(t)$ , siis  $q_d(t) = f(p(t))$ . Tuottajien voitto aikayksikössä on siten kahden muuttujan funktio

$$(p, q_s) \mapsto \Pi(p, q_s) = p \min\{q_d, q_s\} - C(q_s), \quad (2.20)$$

missä  $C: \mathbb{R} \rightarrow \mathbb{R}$  esittää tuotantokustannuksia aikayksikössä. Tuottajat *eivät tiedä* funktioita  $f$  ja  $C$ , mutta he osaavat approksimoida voittonsa osittaisderivaattoja:

$$\begin{aligned} \frac{\partial \Pi}{\partial p} &= \frac{\partial}{\partial p} \min\{pq_s, pf(p)\}, \quad f'(p(t)) \approx \frac{q_d(t) - q_d(t-h)}{p(t) - p(t-h)}, \quad h \approx 0, \\ \frac{\partial \Pi}{\partial q_s} &= \frac{\partial}{\partial q_s} \min\{pq_s, pf(p)\}, \\ C'(q_s(t)) &\approx \frac{C(q_s(t)) - C(q_s(t-h))}{q_s(t) - q_s(t-h)}, \quad h \approx 0. \end{aligned}$$

Muodosta yksinkertaisimmat mahdolliset differentiaaliyhtälöt suureille  $p$  ja  $q_s$ , jotka kertovat, että tuottajat nostavat tai laskevat yksikköhintaa jos se näyttää kasvattavan heidän voittoa samoin kuin he lisäävät tai vähentävät tuotantonopeutta sen mukaan, onko  $\partial \Pi / \partial q_s$  positiivinen tai negatiivinen. Lausu tuo yhtälöpari vektorimuodossa.

3. Todista, että Pythagoraan lause  $c = \sqrt{a^2 + b^2}$  on dimensiohomogeeninen. Oleta, että  $[c] = [a] = [b] = L$  ja  $c = f(a, b)$  dimensiohomogeenisesti ja päättelee, millainen funktio  $f$  on.
4. Tarkastellaan ilmakuplan nousunopeutta  $v$  pystysuorassa putkessa, jonka halkaisija on  $d$  ja jossa on vettä, jonka tiheys on  $\rho_v$ . Vapaan putoamisliikkeen kiihtyvyys on  $g$  ja ilman tiheys  $\rho_i$ . Oleta, että nousunopeus riippuu vain näistä suureista, siis

$$v = f(d, g, \rho_i, \rho_v).$$

Määrrä ilman Buckinghamin  $\pi$ -teoreemaa ensin yleinen lauseke nousunopeudelle. Miten se yksinkertaistuu, kun ilman ja veden tiheys ovat vakioita?

5. Väitetään, että ns. isokenkäiset ajavat autoillaan vauhdikkaammin kuin pienikenkäiset. "Todista" dimensioanalyysin avulla, että se ei voi pitää paikkaansa. Oleta, että auton nopeus  $v$  (metri/sekunti) riippuu polttoaineen kulutuksesta  $\dot{V}$  (kuutiometri/sekunti) ja kuljettajan kengännumerosta, jota esittää kengän pituus  $l$  (metri). Mikä on johtopäätöksesi?

6. Jatketaan edellistä tehtävää. Oletetaan, että autoilijan nopeus riippuu hänen kengännumeron ja autonsa polttoaineenkulutuksen lisäksi vielä hänen tuloistaan  $S$  (mk/sekunti) ja polttoaineen yksikköhinnasta  $p$  (mk/kuutiometri). Siis  $v = f(\dot{V}, l, S, p)$ . Tee dimensioanalyysi. Miten ajonopeus jakautuu niiden autoilijoiden keskuudessa, joilla  $p\dot{V}/S$  on vakio?
7. Valmisbetonikauppias olettaa, että betonin menekki  $\dot{V}$  (kuutiometri/sekunti) paikkakunnalla riippuu vain betonin yksikköhinnasta (mk/kuutiometri), asuntolainojen korosta  $r$  (1/sekunti), paikkakunnan rakentamattomien tonttien pinta-alasta  $A$  (neliömetri) ja sen asukkaiden kokonaistuloista  $S$  (mk/sekunti). Siis  $\dot{V} = f(p, S, r, A)$ . Käytä perussuurejärjestelmää  $\{L, T, R\}$ , missä  $R$  vastaa rahaa. Määritä funktion  $f$  muoto dimensioanalyysin avulla.
8. Olkoon pystysuoran ohuen putken säde  $r$  ja olkoon siinä nestettä, jonka tiheys on  $\rho$  ja pintajännitys  $\gamma$ . Halutaan ratkaista nestepinnan nousukorkeus  $h$  painovoimakentässä, jossa vapaan putoamisliikkeen kiihtyvyys on  $g$ . Dimensiomatriisi on

| $A$      | $M$ | $L$ | $T$ |
|----------|-----|-----|-----|
| $h$      | 0   | 1   | 0   |
| $r$      | 0   | 1   | 0   |
| $\rho$   | 1   | -3  | 0   |
| $\gamma$ | 1   | 0   | -2  |
| $g$      | 0   | 1   | -2  |

Tee dimensioanalyysi Buckinghamin  $\pi$ -teoreeman avulla ja johda lauseke

$$h = r\phi\left(\frac{r^2\rho g}{\gamma}\right),$$

missä  $\phi: \mathbb{R}_+ \rightarrow \mathbb{R}$  on tuntematon funktio.

Tarkastellaan tilanteen geometriaa hieman huolellisemmin. Jaetaan pituus pystysuoraan ja vaakasuoraan komponenttiin  $Z$  ja  $R$ . Nyt  $h$  ja  $g$  vaikuttavat  $Z$ -suunnassa,  $\rho$  on taas symmetrinen joka suunnassa ja  $r$  vaikuttaa  $R$ -suunnassa. Mitä taas pintajännitykseen  $\gamma$  tulee, se määrittää voimana pinnan tangentin suuntaiseen tasoon kuuluvaa pituusyksikköä kohti, joten sen dimensio on  $MZT^{-2}R^{-1}$ . Dimensiomatriisi on nyt

| $A$      | $M$ | $R$ | $Z$ | $T$ |
|----------|-----|-----|-----|-----|
| $h$      | 0   | 0   | 1   | 0   |
| $r$      | 0   | 1   | 0   | 0   |
| $\rho$   | 1   | -2  | -1  | 0   |
| $\gamma$ | 1   | -1  | 1   | -2  |
| $g$      | 0   | 0   | 1   | -2  |

Tee dimensioanalyysi Buckinghamin  $\pi$ -teoreeman avulla ja päätele, että

$$h = C \frac{\gamma}{g\rho r},$$

missä  $C > 0$  on jokin vakio.

## Luku 3

# Taloustieteiden tärkeimpiä suureita

Mitään yleistä ja yksikäsitteistä perussuurejärjestelmää taloustieteissä ei ole, koska sellaista ei ole fysiikassakaan. Kulloistakin teoriaamme tai malliamme varten kiinnitämme jotkin suureet perussuureiksi ja muut esiintyvät suureet ovat niistä johdettuja. Riittää, että suureen mittaluku toteuttaa Bridgmanin aksiooman. Tuollaisen mittaluvun olemassaolo voidaan perustella olettamalla sopiva järjestysrelaatio: jos on olemassa ideaalinen orsivaaka, jolla voidaan järjestää massakappaleet, niin sitten on olemassa Bridgmanin aksiooman toteuttava massan mittalukukin<sup>1</sup>

Oletukset järjestysrelaatiosta ovat kuitenkin riittäviä, eivät välttämättömiä ehtoja. Teoriassamme tai mallissamme saa esiintyä suureita, joita emme periaatteessakaan voi mitata suoraan, koska sellaista esiintyy myös fysiikan teorioissa (vektoripotentiaali, aaltofunktio...) ja on myös muistettava Duhemin-Quinen teesi, jonka mukaan ainoatakaan fysiikan teoriaa ei voida suoraan testata empiirisesti, vaan aina testi kohdistuu koeteltavan teorian ja koejärjestelyä koskevien teorioiden ja alkuehtojen kompleksiin.

### 3.1 Määräsuureet

Määräsuureina käytetään kaikkia fysikaalisia määräsuureita: tilavuus, massa, pinta-ala, pituus... Rahan tai osakkeiden määrät ovat myös määräsuureita. Joskus on hyödyllistä panna esimerkiksi maidon ja bensiinin määrät dimensioltaan erisuuriksi: ne ovat molemmat nesteitä ja kelpaavat polttoaineeksi, mutta edellinen ihmisille, jälkimmäinen bensiinimoottoreille eikä niitä voi korvata toisillaan.

### 3.2 Arvosuureet

Taloustieteissä tavallisin arvosuure on vastaavan määräsuureen yksikköhinta. Maidon yksikköhinnan mittayksikkö on mk/l, hamppuköyden mk/m, metsätalousmaan mk/m<sup>2</sup>, Olvi Oyj:n osakkeen mk/kpl...

Ihmiset voivat tunnistaa useampaa lajia samaa rahaa (omaa, ansaittua, vierasta, perittyä) ja kohdella samoja markkoja eri lailla. Tällöin heidän käyttäytymisensä kuvaamisessa on paikallaan jakaa rahan määrä vastaavasti eridimensioisiksi suureiksi. Arvosuureina on myös pidettävä kauneuden tai pyhyiden

---

<sup>1</sup> Ks. Niiniluoto, Ilkka: *Approksimaatio ja idealisaatio teorianmuodostuksessa*. Arkhimedes 1A/1986, ss. A38-A48. Aiheesta myös: Krantz, David H.; Luce, R. Duncan; Suppes, Patrick; Tversky, Amos: "Foundations of measurement. Vol. 1: Additive and polynomial representations". Academic Press, New York & Lontoo, 1971. Vrt. janamitan yksikäsitteisyys geometrian kurssilla.

määriä. Molemmilla voisi olla käyttöä, jos kuvaisimme vaikkapa kynttilöiden ja ikoninkuvien menekkiä krasnodarilaisen kirkon eteisessä.

### 3.3 Nopeussuureet

Huomattava osa taloudellisista ilmiöistä on dynaamisia: maata kyllä omistetaan siitä siirtelemättä, mutta päivittäistavaroita ostetaan ja myydään joka päivä. Tärkeitä suureita ovat siksi kulutus-, tuotanto-, investointinopeudet. Niiden yksikkö on vastaavan määräsuureen yksikkö jaettuna aikayksiköllä. Jos aika on jatkuva, niin nopeussuureen  $q(t)$  yleinen määritelmä on vastaavan määräsuureen  $Q(t)$  derivaatta ajan suhteen:

$$q(t) = \lim_{h \rightarrow 0} \frac{Q(t+h) - Q(t)}{h}.$$

Jos taas aika on diskreetti, käytämme määräsuureita tai erotusosamäärällä

$$q(t_i) = \frac{Q(t_i) - Q(t_{i-1})}{t_i - t_{i-1}}$$

määriteltyä nopeussuuretta.

Tavallisia yksiköitä tuotanto- tai kulutusnopeudelle ovat m/kk, l/s, kpl/d. Investointinopeuden yksikkönä voi olla mk/d. Myös tuotantokustannukset aikayksikössä ja myyntitulot aikayksikössä ovat nopeussuureita. Niiden yksikkö voi olla myös mk/d. Huomaa, että korko on sikäli nopeussuure, että sen yksikkö on 1/d, dimensio  $T^{-1}$ , ja korko  $\times$  pääoma on selvästi nopeussuure.

### 3.4 Kertymäsuureet

Jos  $q$  on nopeussuure ja aika on diskreetti eli  $T' = \{\tau_i \in \mathbb{R} \mid i = 1, 2, \dots, b\}$ , missä  $b \in \mathbb{N} \cup \{\infty\}$ , niin summa

$$\sum_{i=1}^b q(\tau_i)(\tau_i - \tau_{i-1})$$

kertoo aikana  $T'$  kertyneen määrän. Jos aika on jatkuva ja  $[t_1, t_2] \subset T'$ , niin silloin integraali

$$\int_{t_1}^{t_2} q(t) dt$$

kertoo ajanhetkien  $t_1$  ja  $t_2$  välillä kertyneen määrän.

Jos siis tuotteen tuotantonopeus on  $q(t) = \alpha t^2$ ,  $\alpha > 0$  on vakio, niin ajanhetkien  $t = 0$  ja  $t = 3d$  välillä tuotettu tavaramäärä on

$$Q = \int_0^{3d} \alpha t^2 dt = \int_0^{3d} \frac{\alpha}{3} t^3 = 9\alpha d^3.$$

Huomaa, että integraalilaskennan peruslauseen mukaan kertymäsuureen derivaatta ajan suhteen on vastaava nopeussuure.

Tyypillinen taloudellinen kertymäsuure on päivittäin korkoa antavan pankkitilin korkotulot. Jos  $r(t)$  on korko ja  $y(t)$  on tilin saldo hetkellä  $t$ , niin ajanjaksolla  $[0, t]$  kertyneet korkotulot ovat

$$\int_0^t r(\tau)y(\tau) d\tau$$

Jos korko maksetaan päivittäin tilille, pätee kasvuyhtälö

$$y'(t) = r(t)y(t), \quad t > 0.$$

### 3.4.1 Diracin $\delta$ -funktio

Fysiikassa tarvitaan funktiota  $\delta$ , jolla on seuraavat ominaisuudet:

1.  $\delta: \mathbb{R} \rightarrow [0, \infty]$ .
2.  $\delta(x) = 0$  kaikilla  $x \neq 0$ .
3.  $\int_{-\infty}^{\infty} \delta(x)f(x) dx = f(0)$  kaikilla origossa jatkuvilla  $f: \mathbb{R} \rightarrow \mathbb{R}$ .

Tätä oliota sanotaan *Diracin  $\delta$ -funktioksi*.

Olkoon  $H: \mathbb{R} \rightarrow \mathbb{R}$  *Heavisiden askelfunktio eli*

$$H(x) = \begin{cases} 0, & \text{jos } x < 0, \\ 1, & \text{jos } x > 0. \end{cases}$$

Sen derivaatta on Diracin  $\delta$ -funktio, sillä jokaiselle funktiolle  $f \in C_0^1(\mathbb{R})$  pätee osittaisintegroitukaavan mukaan

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)H'(x) dx &= \int_{-\infty}^{\infty} f(x)H(x) - \int_{-\infty}^{\infty} f'(x)H(x) dx = \\ &= - \int_0^{\infty} f'(x) dx = f(0). \end{aligned} \quad (3.1)$$

Näin siis  $\delta(x) = H'(x)$  kaikilla  $x \in \mathbb{R}$ .

Esimerkiksi pisteessä  $a = (a_1, a_2, a_3)$  olevan elektronin varausjakaumaa esittää

$$\rho(y) = -e\delta_3(y - a) = -e\delta(y_1 - a_1)\delta(y_2 - a_2)\delta(y_3 - a_3),$$

missä  $\delta_3$  on kolmiulotteinen Diracin  $\delta$ -funktio ja  $e = 1,602 \cdot 10^{-16} Q$  on alkeisvaraus. Tällöin elektronin kokonaisvaraus on

$$\int_{\mathbb{R}^3} \rho(y) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} -e\delta(y_1 - a_1)\delta(y_2 - a_2)\delta(y_3 - a_3) dy_3 dy_2 dy_1 = -e.$$

Matemaattisesti Diracin  $\delta$ -funktio ei ole funktio ja yllä oleva lasku Heavisiden askelfunktion derivaatasta on hölynpölyä. Seikka synnytti kiihkeän selkkauksen teoreettisten fyysikkojen ja matemaatikkojen välille. Tästä dialektisesta ristiriidasta syntyi synteesi, *distribuutioteoria*, joka on matemaattisesti korrekti ja merkittävä ala modernia matematiikkaa (distributionaalinen derivaatta, Sobolevin avaruudet...). Oikeasti Diracin  $\delta$ -funktio ei ole funktio, vaan olio  $\delta(x) dx$  on *mitta*. Osittainintegroitukaava (3.1) *määrittelee* funktion (distributionaalisen) derivaatan.

Osakkeiden osinko  $R$  (mk/osake) maksetaan tavallisesti kerran vuodessa tietynä täsmäyshetkenä  $t_1$  omistettujen osakkeiden määrän  $y(t_1)$  mukaan, joten osinkojen kertymä hetkellä  $t$  on

$$O(t) = \begin{cases} Ry(t_1), & \text{jos } t \geq t_1, \\ 0, & \text{jos } t < t_1. \end{cases}$$

Diracin  $\delta$ -funktion avulla voidaan ilmaista osinkotulot samannäköisellä integraalilausekkeella kuin jatkuvat korkotulot:

$$\int_0^t R\delta(\tau - t_1)y(\tau) d\tau = \int_{-t_1}^{t-t_1} R\delta(\sigma)y(\sigma + t_1) d\sigma = \begin{cases} Ry(t_1), & \text{jos } t > t_1, \\ 0, & \text{jos } t < t_1. \end{cases}$$

Täten hetkellistä korkoa vastaava hetkellinen osinko on  $R\delta(t - t_1)$  ja vastaava differentiaaliyhtälö on

$$y'(t) = R\delta(t - t_1)y(t), \quad t > 0.$$

### 3.5 Rajasuureet

Suureita voidaan yrittää derivoida muidenkin muuttujiensa kuin ajan suhteen. Tuloverot Suomessa riippuu vuotuisista tuloista. Henkilön vuodessa maksettava tulovero on  $V(q)$ , missä  $q$  on hänen tulonsa vuodessa. Sekä  $V(q)$  että  $q$  ovat nopeussuureita. Funktio  $q \mapsto V(q)$ ,  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  on paloittain affiini ja jatkuva. *Rajaveroasteeksi* sanotaan derivaattaa  $V'(q)$ .

Kuluttajan mielihyvä  $P$  voi olla funktio hänen kulutusnopeudestaan  $q$  ja massastaan  $m$ . Tällöin siis  $P = P(m, q)$  eli mielihyvä on kahden muuttujan funktio. Nyt voidaan tutkia, onko sillä osittaisderivaatat noiden muuttujiensa suhteen. Jos on, niin niitä sanotaan *rajamielihyväksi* kulutusnopeuden ja massan suhteen, ja ne kertovat, miten kuluttajan mielihyvä muuttuu, jos täsmälleen toista noista muuttujista muutetaan. Yksi uskottava lauseke<sup>2</sup> kuluttajan mielihyvälle on

$$P(m, q) = \frac{\gamma m}{\alpha^2 + m^2} \ln \frac{q}{\beta},$$

missä  $\gamma > 0$  on vakio,  $\alpha = 100$  kg ja  $\beta = 100$  g/d. Huomaamme, että rajamielihyvät

$$\frac{\partial P}{\partial m} = \gamma \frac{\alpha^2 - m^2}{(\alpha^2 + m^2)^2} \ln \frac{q}{\beta}, \quad \frac{\partial P}{\partial q} = \frac{\gamma m}{\alpha^2 + m^2} \frac{1}{q},$$

käyttäytyvät eri lailla kun  $m$  ja  $q$  kasvavat. Kun  $m$  ylittää vakion  $\alpha$  ja  $q > \beta$ , tulee rajamielihyvä  $\partial P / \partial m$  negatiiviseksi ja siten massan lisäys vähentää kuluttajan mielihyvää. Kun taas kulutusnopeus kasvaa, lisääntyy mielihyvä alati. Mutta mitä suurempi  $q$  on, sitä pienemmäksi tulee rajamielihyvä  $\partial P / \partial q$  ja niin sitä vähemmän mielihyvä kasvaa. Tätä ilmiötä sanotaan *alenevan rajamielihyvän laiksi* (jos mielihyvän sijasta tarkastelisimme hyötyä, nimitys olisi alenevan rajahyödyn laki). Ensimmäiset leipäpalat ovat nälkäiselle mieluisimmat<sup>3</sup>.

Olkoon  $f: A \rightarrow \mathbb{R}$ , missä  $\emptyset \neq A \subset V$  ja  $V$  on vektoriavaruus. Sanotaan, että kuvaus  $f$  on *kupera*, jos kaikilla  $x, y \in A$  ja  $\lambda \in [0, 1]$  pätee

$$\lambda x + (1 - \lambda)y \in A \text{ ja } f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Jos  $-f$  on kupera, niin kuvausta  $f$  sanotaan *koveraksi*. Olkoon  $P: \mathbb{R}_+^n \rightarrow \mathbb{R}$  funktio joka ilmoittaa  $n$ :ää eri hyödykettä nopeuksilla  $q_1, \dots, q_n$  kuluttavan kuluttajan mielihyvän aikayksikössä. Alenevan rajahyödyn laki toteutuu, kun kuvaus  $P$  on kovera, sillä silloin kaikki osittaisderivaatat ovat väheneviä vastaavan muuttujan kasvaessa. Sanomme mielihyvän  $P$  gradienttia

$$h(q_1, \dots, q_n) = \left( \frac{\partial P}{\partial q_1}, \dots, \frac{\partial P}{\partial q_n} \right)$$

kuluttajan *maksuhalukkuudeksi*; sen  $i$ :s komponentti kertoo, paljonko kuluttaja on halukas maksamaan  $i$ :nnestä tuotteesta noilla kulutusnopeuksillaan. Alenevan rajamielihyvän lain mukaan maksuhalukkuus vähenee kulutuksen kasvaessa<sup>4</sup>. Klassisen potentiaalteorian termein maksuhalukkuus on skalaaripotentiaalilin  $P$  *vektorikenttä*<sup>5</sup>.

<sup>2</sup>Seuraavat raja-arvo-ominaisuudet ainakin toteutuvat:

$$\begin{aligned} \lim_{m \rightarrow 0+} P(m, q) &= 0, & \lim_{m \rightarrow \infty} P(m, q) &= 0, \\ \lim_{q \rightarrow 0+} P(m, q) &= -\infty, & \lim_{q \rightarrow \infty} P(m, q) &= \infty. \end{aligned}$$

<sup>3</sup>Aleneva rajamielihyvän laki on poliittista dynamiittia: tulonsiirrot pienituloisimmille liisäävät väestön kokonaismielihyvää.

<sup>4</sup>Jos  $f: \mathbb{R} \rightarrow \mathbb{R}$  ja kahdesti derivoituva, niin  $f$  on kupera täsmälleen silloin kun  $f''(x) \geq 0$  kaikilla  $x \in \mathbb{R}$ .

<sup>5</sup>Kuten klassisessa mekaniikassa gravitaatiokenttä tai sähköstatiikassa sähkökenttä. Huomaa kuitenkin, että mielihyvän muuttujat ovat nopeussuureita ja gradientti lasketaan niiden suhteen; gravitaatiokenttä ja staattinen sähkökenttä saadaan laskemalla vastaavan potentiaalilin gradientti paikkamuuttujien suhteen; paikkamuuttujat ovat määräsuureita.

Maksuhalukkuuden mittaaminen on yksinkertaista kyselytutkimuksin; toisaalta fysiikassa yleensäkin skalaaripotentialiaali on vain apusuure vektorikentän laskemiseksi. Siksi pidämme maksuhalukkuutta perustavampana suureena kuin mielihyvää<sup>6</sup>. Aivan kuten pyörteiset fysikaaliset vektorikentät voi maksuhalukkuus olla sellainen, että se ei ole skalaaripotentialin gradientti<sup>7</sup>.

Kuluttajan *hyöty*  $U(p, q)$  aikayksikössä kulutusnopeudella  $q$  on erotus<sup>8</sup>

$$U(p, q) = P(q) - p \cdot q = P(q_1, \dots, q_n) - \sum_{i=1}^n p_i q_i,$$

missä  $p_i$  on tuotteen  $i$  yksikköhinta ja  $q_i$  sen kulutusnopeus,  $i = 1, 2, \dots, n$ . Hyöty on siis mielihyvä pois kustannukset.

Tuottajalle mielihyvää ja hyötyä vastaavat suureet ovat tuotantokustannukset ja voitto. Tuottakoon tuottaja  $n$ :ää eri tuotetta nopeuksilla  $q_1, \dots, q_n$ . Merkitään *tuotantokustannusta* aikayksikössä  $C(q)$ . Varsin usein  $q \mapsto C(q)$  on kupera kuvaus, jolloin rajatuotantokustannukset ovat kasvavia. Tuotannon lisäämiseksi on laajennettava tuotantolaitosta, mistä tulee kuluja raaka-aine- ja työvoimakulujen lisäksi. Tuottajan *voitto* aikayksikössä on tulot aikayksikössä pois tuotantokustannus aikayksikössä. Jos tuottaja saa kaikki tuotteensa myytyä ja kunkin yksikköhinta on  $p_i$ , niin hänen voittonsa aikayksikössä on

$$\Pi(p, q) = p \cdot q - C(q) = \sum_{i=1}^n p_i q_i - C(q_1, \dots, q_n).$$

Tuotantokustannuksen kuperuudesta seuraa tällöin, että rajavoitot ovat vähenviä.

## 3.6 Rajoitteet

Taloudelliselle toiminnalle on ominaista sopeutuminen lukuisiin rajoitteisiin: käytettävissä olevien rahojen niukkuus rajoittaa kuluttamista budjettirajoittein, jotka ovat muotoa

$$p \cdot q = \sum_{i=1}^n p_i q_i \leq M,$$

missä  $p_i$  on  $i$ :nnen tuotteen yksikköhinta ja  $q_i$  sen kulutusnopeus. Negatiivinen tuotanto- tai kulutusnopeus ei ehkä olekaan mahdollinen, tuotannon kasvattaminen yli tietyn rajan voi olla kilpailunedistämislainsäädännön vastaista tai tuotannon laajentamiseen tarvittavia tehdaskiinteistöjä ei ole saatavissa.

## 3.7 Informaatio

Praktisessa syllogismissa oleellinen tekijä on toimijan otaksumat vaikutusmahdollisuuksistaan, siis hänen *informaationsa*<sup>9</sup> toimintaympäristöstään ja sen kehittämisestä. Ajan myötä tekijän informaatio muuttuu ja hän korjaa käyttäytymistään uusien otaksumiensa mukaiseksi.

<sup>6</sup>Mielihyvän käsitteen käyttöä vastaan on huomautettu, että se olisi muka huonosti määritelty, koska mikä hyvänsä funktio  $f(P(q))$ , missä  $f$  on aidosti kasvava, kertoisi samat ihmisen preferenssit, ja siten ihmisjoukon kokonaismielihyvä olisi sekin mielivaltainen käsite. Samalla perusteella ihmisen massa ja ihmisjoukon yhteismassa olisivat kelvottomia käsitteitä.

<sup>7</sup>Tällöin voidaan yhä puhua mielihyvästä, sillä klassisen potentiaaliteorian mukaan jokainen vektorikenttä on summa skalaaripotentialin gradientista ja vektoripotentialin roottorista.

<sup>8</sup>Tällöin oletetaan, että mielihyvän dimensio on sama kuin rahan määrän dimensio.

<sup>9</sup>Informaatio sinänsä ei ole tietoa, sillä tieto on totta tai ainakin "esiintyy totuuden vaatimuksin".

### 3.7.1 Esimerkki: osakesijoittaja

Tarkastellaan osakesijoittajaa, joka ostaa ja myy Olvi Oyj:n osakkeita. Hänen tilansa ajanhetkellä  $t$  on pari  $(y_1(t), y_2(t)) = y(t)$ , missä  $y_1(t)$  on hänen asiakastilinsä saldo (mk) ja  $y_2(t)$  on hänen arvo-osuustililleen kirjattujen Olvi Oyj:n osakkeiden määrä. Hetkellä  $t_0$  hänellä on tillillä rahaa määrä  $y_0$  ja osakkeita ei lainkaan. Hän haluaa toimia niin, että hänen varallisuutensa hetkellä  $t_1 > t_0$

$$W(y, t_1) = y_1(t_1) + p(t_1)y_2(t_1)$$

on mahdollisimman suuri; tässä  $p(t)$  on osakkeen yksikköhinta hetkellä  $t$ . Jos käytämme jatkuvaa aikaa ja jos sijoittajan asiakastilin korko maksetaan tillille heti, niin hänen asiakastilinsä saldo hetkellä  $t \in [t_0, t_1]$  on

$$y_1(t) = y_0 + \int_{t_0}^t r_1(s)y_1(s) ds + \sum_{i=1}^{N_t} r_2(\tau_i)y_2(\tau_i) - \int_{t_0}^t p(s)y_2'(s) ds, \quad (3.2)$$

missä  $r_1(s)$  on pankkitilin korko hetkellä  $s$  ja  $r_2(\tau_i)$  osakkeista hetkellä  $\tau_i$  omistettujen osakkeiden perusteella maksettava osinko;  $N_t \in \mathbb{N}$  siten, että  $\tau_1 < \tau_2 < \dots < \tau_{N_t} \leq t$  ja  $\tau_{N_t+1} > t$ . Toinen termi on korkotulot: lauseke  $r_1(s)y_1(s)ds$  on infinitesimaalisella<sup>10</sup> aikavälillä  $ds$  kertynyt korkotulo, jolloin integraali yli koko välin  $[t_0, t_1]$  antaa kaikki korkotulot. Tavallisesti korko maksetaan päivittäisen saldon mukaan ja siten toinen termi on tarkkojen korkotulojen likiarvo<sup>11</sup>. Summan korvaaminen integraalilla (ja erotusosamäärän derivaatalla) mahdollistaa vuosisatoja kehittyneen matemaattisen analyysin<sup>12</sup> ja sen vahvojen menetelmien soveltamisen. Kolmas termi antaa osinkotulot:  $\tau_1, \dots, \tau_{N_t} \in [t_0, t_1]$  ovat ajanhetket, joina omistettujen osakkeiden määrän perusteella jaetaan tuon tilivuoden osinko  $r_2(\tau_i)$ . Neljäs termi on myyntitulot osakkeista: jos  $-dy_2 = y_2'(s)ds$  osaketta myydään infinitesimalisessa ajassa  $ds$ , niin niistä saadaan rahaa määrä  $-p(s)y_2'(s)ds$ . Rajoitteet ovat

$$y_1(t) \geq 0, y_2(t) \geq 0 \text{ kaikilla } t \in [t_0, t_1], \quad (3.3)$$

$$y_1(t_0) = y_0, y_2(t_0) = 0. \quad (3.4)$$

Epäyhtälöt  $y_1(t) \geq 0$  ja  $y_2(t) \geq 0$  saadaan siitä, että tileillä ei ole luottoa. Viimeiset kaksi yhtälöä kertovat alkutilan.

Tulevaa korkoa, osinkoja eikä osakekurssia sijoittaja ei tiedä. Aina kun hänen informaationsa eli käsityksensä niiden tulevista arvoista muuttuu, on hänen haettava uudestaan ne funktiot  $y_1$  ja  $y_2$ , joilla hänen varallisuudestaan tulee suurin. Hänen on siis haettava ne derivoituvat *funktiot*  $y_1, y_2: [t_1, t_2] \rightarrow \mathbb{R}$ , jotka toteuttavat (3.2)–(3.4) ja joilla *funktionaali*<sup>13</sup>  $(y_1, y_2) \mapsto W(y, t_1)$  on suurin. Osakesijoittajamme ääriarvotehtävän osasia nimitetään toisinaan, etenkin *sääntöteoriassa*<sup>14</sup>, seuraavasti:  $y \mapsto W(y, t_1)$  on voittofunktio, (3.2) on tilayhtälö, (3.3) on rajoitteet ja (3.4) on alkuehdot.

Yllä rajoituttiin derivoituviin funktioihin  $y_1$  ja  $y_2$ . Maksimoitava funktionaali on lineaarinen ja se ei ole identtisesti nolla, joten jos ääriarvokohta on olemassa, niin ainakin toinen rajoite-epäyhtälöistä (3.3) toteutuu yhtälönä. Näinhän käy myös maksimoitaessa lineaarikuvausta  $\mathbb{R}^2 \rightarrow \mathbb{R}$  jossakin tason joukossa. Tällöin derivoituvuuden oletaminen funktioista  $y_1$  ja  $y_2$  on liian paljon. Paloittain jatkuvuuden on riitettävä, sillä selvästi näemme erikoistapauksia (sopivat  $p(t)$ ,  $r_1(t)$  ja  $r_2(t)$ ), joissa sijoittajan käyttäytyminen on ns. bang-bang-toimintaa<sup>15</sup>:

<sup>10</sup>Fyysikkojen kieltä, “äärettömän pieni”. Palaamme vielä infinitesimaalisiin tarkasteluihin perusteellisemmin.

<sup>11</sup>Nykyään pankkien olisi teknisesti helppoa laskea korko vaikkapa sekunneittain.

<sup>12</sup>Se matematiikka, joka perustuu raja-arvon käsitteeseen. Siis differentiaali- ja integraalilaskenta, differentiaaliyhtälöt yms.

<sup>13</sup>Ts. reaaliarvoinen funktio, jonka muuttuja on funktio.

<sup>14</sup>Engl. (optimal) control theory.

<sup>15</sup>Sääntöteorian termi, tarkoittaa menettelyä, jossa resursseja vuorotellen käytetään täydellä teholla ja ei lainkaan.

jos sijoittaja uskoo osakekurssien nousevan huomiseksi kaksinkertaiseksi ja laskevan takaisin ylihuomiseksi, niin hänen on ostettava tänään osakkeita kaikilla varoillaan ja myytävä ne kaikki huomenna. Integraaliyhtälöstä (3.2) saadaan integroimalla viimeinen termi osittain ja soveltamalla alkuehtoja (3.4), että

$$y_1(t) - \int_{t_0}^t r_1(s)y_1(s) ds = y_0 + \sum_{i=1}^{N_t} r_2(\tau_i)y_2(\tau_i) - y_2(t)p(t) + \int_{t_0}^t y_2(s)p'(s) ds \text{ kaikilla } t \in [t_0, t_1] \quad (3.5)$$

Tämä yhtälö ei vaadi funktioilta  $y_1$  ja  $y_2$  kuin integroituvuuden<sup>16</sup>, johon riittää paloittain jatkuvuus. Edelleen yhtälöstä (3.5) voidaan ratkaista vakion varioimiskeinolla  $y_1$   $y_2$ :n funktiona. Lauseke on affini<sup>17</sup>  $y_2$ :n suhteen. Korvaamme yhtälön (3.2) yhtälöllä (3.5), joilloin voimme eliminoida loppuvarallisuuden lausekkeesta funktion  $y_1$  pois ja saamme yhden tuntemattoman funktion ääriarvottehtävän rajoittein (3.3).

### 3.8 Harjoitustehtäviä

1. Tarkastellaan palkansaaajan käyttötilin saldoa  $y(t)$  hetkellä  $t$  vuoden aikana. Hän saa jokaisen kuukauden viimeinen päivä tililleen palkkaa 8000 mk, 30.6 lisäksi lomarahaa 5000 mk, jokaisen kuukauden 15. päivä hän maksaa vuokraa asunnostaan 2000 mk. Käyttötililtään hän nostaa pankkikorttiostoksina joka päivä 100 mk. Käyttötilin korko on 2 %/vuosi, se lasketaan päivittäisestä saldosta ja maksetaan käyttötilille 30.11. Käytä jatkuvaa aikaa ja oletta, että vuodessa on tasan 12 kpl 30-päiväistä kuukautta. Laske  $y(t)$  jokaiselle  $0 \leq t \leq 360$  vrk, kun  $y(0) = 8000$  mk.
2. Jatketaan edellistä tehtävää. Laske  $y'(t)$  ja lausu se Diracin  $\delta$ -funktion avulla. Tarkasta tuloksesi laskemalla  $\int_0^{361 \text{ vrk}} y'(t) dt$ .
3. Poliittisessa propagandassa esitetään toisinaan yleispätevänä tosiasiana väite, että yrityksen optimaalinen työvoiman kysyntä vähenee, jos palkkoja korotetaan ja korkotaso laskee. Siten palkkojen korottaminen heikentäisi työllisyyttä. Tutki väitteen todenperäisyyttä seuraavan yrityksen avulla. Sen voiton aikayksikössä hetkellä  $t$  lausuu funktio

$$\Pi: [0, \infty]^2 \rightarrow \mathbb{R}, \Pi(K, L) = \alpha K^{\frac{1}{2}} L^{\frac{1}{4}} - rK - wL,$$

missä  $K$  on tuotantolaitteisiin sijoitetun pääoman määrä (mk) ja  $L$  on käytetyn työvoiman määrä aikayksikössä (miestyötunti/tunti). Ensimmäinen termi antaa tuotteen myyntitulot; vakio  $\alpha > 0$  sisältää tuotteen yksikköhinnan, joka on vakio, ja edelleen tuotantonopeus on kertovaa vakioita vaille  $K^{\frac{1}{2}} L^{\frac{1}{4}}$ . Toinen termi esittää työvoimakustannuksia aikayksikössä ja kolmas pääomakustannuksia aikayksikössä ( $r$  on korko, joka joudutaan maksamaan laitteiden hankkimislainasta). Osoita, että  $\Pi$  saavuttaa suurimman arvonsa yksikäsitteisessä pisteessä  $(K_*, L_*) \in (0, \infty)^2$  ja ilmoita se parametrien  $\alpha$ ,  $r$  ja  $w$  avulla. Valitse edelleen  $w = w_0(1 + a)$  ja  $r = r_0(1 - a)$ ,  $-1 < a < 1$ ,  $w_0, r_0 > 0$ , jolloin  $L_* = L_*(a)$ . Mitä tapahtuu kun  $a$  kasvaa? (Muista, että  $\alpha\beta \leq \epsilon\alpha^2 + \frac{1}{4\epsilon}\beta^2$  kaikilla  $\alpha, \beta \geq 0$  ja  $\epsilon > 0$  ja että kompaktissa joukossa jatkuva kuvaus saavuttaa suurimman ja pienimmän arvonsa.)

4. Poliitikko Piia-Noora Kaupin mielestä<sup>18</sup> olisi oikeudenmukaista jakaa miljoonan markan perintö rutiköyhän ja upporikkaan veljeksien välillä siten,

<sup>16</sup>Riemann-integroituvuuden tai Lebesgue-integroituvuuden.

<sup>17</sup>Affini kuvaus on lineaarikuvaus ynnä vakio.

<sup>18</sup>NYT, 7/2000, s. 10.

että köyhimys saa 100 000 mk ja pohatta 900 000 mk, *koska* sellainen jako olisi molemmille suhteellisesti sama.

Oletetaan, että kummakin veljeksien mielihyvä  $U$  riippuu hänen varallisuudestaan  $r$  (1000 mk) ja elävien sukulaistensa määrästä  $s$  (1 kpl) lausekkeen  $U(r, s) = \sqrt{r + 1000s}$  mukaisesti. Olkoon alkutilanne  $(r_0, s_0) = (0, 2)$  ja  $(\tilde{r}_0, \tilde{s}_0) = (16000, 2)$ . Sukulaisen kuoleman ja perinnönjaon jälkeen tilanne on  $(r_1, s_1) = (x, 1)$  ja  $(\tilde{r}_1, \tilde{s}_1) = (17000 - x, 1)$ , missä  $x$  on siis köyhimyksen osa perinnönjaossa. Totea ensiksi, että rajamielihyvä on positiivinen ja vähenevä niin varallisuuden kuin sukulaisten määrän kasvaessa eli kyseessä on aivan tavallinen mielihyväfunktio. Määrää  $x$  Kaupin ehdosta

$$\frac{U(r_1, s_1) - U(r_0, s_0)}{U(r_0, s_0)} = \frac{U(\tilde{r}_1, \tilde{s}_1) - U(\tilde{r}_0, \tilde{s}_0)}{U(\tilde{r}_0, \tilde{s}_0)}. \quad (3.6)$$

Laske vielä, mitä ehtoa (3.6) noudattavan tuomarin tulisi päättää, jos perintö onkin vain 1000 mk.

## Luku 4

# Kaksi numeerisen matematiikan menetelmää

Tarkastelemme tässä luvussa kahta numeerisen matematiikan menetelmää, jotka ovat perustavia taloudellisten mallien ja taloustieteiden kannalta. Nämä menetelmät ovat *iterointi* yhtälön ratkaisemiseksi ja *gradienttimenetelmä* funktion ääriarvon löytämiseksi.

### 4.1 Iterointi

#### 4.1.1 Keplerin yhtälön ratkaiseminen

Keplerin ensimmäinen laki sanoo, että planeetat kiertävät ellipsinmuotoisia ratoja, joiden toisessa polttopisteessä on Aurinko on. Keplerin toisen lain mukaan Aurinko-planeetta-yhdysjanan ja ellipsin isoakselin rajaama pinta-ala kasvaa vakionopeudella (eli pintaanopeus on vakio<sup>1</sup>). Keplerin kolmannen lain mukaan planeetan radan isoakselin kuutio on suoraan verrannollinen planeetan kiertoaajan neliöön. Nämä lait mahdollistavat planeetan paikan laskemisen, kunhan kuusi kullekin planeetalle ominaista vakiota, ns. rataelementtiä, tiedetään<sup>2</sup>. Tämän laskun vaikein askel on seuraavan *Keplerin yhtälön*

$$E - e \sin E = M \quad (4.1)$$

ratkaiseminen;  $M$  on planeetan *keskianomalia*,

$$M = 2\pi \frac{t - T}{P},$$

missä  $T$  on aika, jolloin planeetta on perihelissä<sup>3</sup>, ja  $P$  on planeetan kiertoaika;

$$0 \leq e = \sqrt{1 - \frac{b^2}{a^2}} < 1 \quad (4.2)$$

on planeetan radan *eksentrisyys* ( $a$  ja  $b$  ovat radan iso- ja pikkuakselin puolikaat). Tuntematon suure  $E$  on planeetan *eksentrisen anomalia* eli kulma, jonka sivuilla planeetta ja periheli ovat ja jonka kärkenä on radan keskipiste.

---

<sup>1</sup>Pintaanopeus  $A(t) = \frac{1}{2} \mathbf{x}(t) \times \mathbf{x}'(t)$ , missä  $\times$  on vektoritulo, jolloin keskeisvoimakentässä  $2A'(t) = \mathbf{x}'(t) \times \mathbf{x}'(t) + \mathbf{x}(t) \times \mathbf{x}''(t) = 0$ .

<sup>2</sup>Ks. Heikki Oja, Taivaanmekaniikka, Limes ry, Helsinki 1977, ss. 10-17, tai Hannu Karttunen, Johdatus taivaanmekaniikkaan, Limes ry, Helsinki 1980, ss. 26-35.

<sup>3</sup>Kiertolaisen radan piste, joka on lähinnä keskuskappaletta. Jupiterin kuilla vastaavan pisteen nimi on perijovium, Kuulla perigeum.

Yhtälön (4.1) tarkka ratkaiseminen onnistuu<sup>4</sup> helposti vain joissakin erikoistapauksissa (esim. jos  $M = 0$ ), joten yritetään seuraavaa menettelyä. Yhtälöstä (4.1) saamme

$$E = M + e \sin E, \quad (4.3)$$

joten määrittelemme induktiivisesti lukujonon  $(E_n)$ :

$$E_0 = 0, \quad E_{n+1} = M + e \sin E_n, \quad n = 0, 1, 2, \dots \quad (4.4)$$

Osoitetaan, että Keplerin yhtälön ratkaisu on yksikäsitteinen. Olkoot  $E$  ja  $\tilde{E}$  kaksi sen ratkaisua siten, että  $E \neq \tilde{E}$ . Silloin väliarvolauseen nojalla on niiden välissä luku  $\xi$  siten, että

$$\begin{aligned} |E - \tilde{E}| &= e |\sin E - \sin \tilde{E}| = e |\cos \xi| \cdot |E - \tilde{E}| \\ &\leq e |E - \tilde{E}| < |E - \tilde{E}|, \end{aligned}$$

mikä on mahdotonta. Siten  $E = \tilde{E}$ .

Jono  $(E_n)$  on rajoitettu, sillä  $|E_n| \leq |M| + e$ . Siten sillä on osajono, joka suppenee kohti jotakin lukua  $E$ . Sinin jatkuvuuden perusteella tuo  $E$  on Keplerin yhtälön ratkaisu. Jos jono  $(E_n)$  ei suppenisi kohti lukua  $E$ , niin jonolla  $(E_n)$  olisi osajono, jonka kaikki alkiot olisivat kauempana kuin jokin kiinteä  $\epsilon > 0$  luvusta  $E$ . Kuitenkin tuolla osajonolla on suppeneva osajono, joka suppenisi kohti Keplerin yhtälön ratkaisua. Koska ratkaisu on yksikäsitteinen, saadaan ristiriitaa. Siten koko jono  $(E_n)$  suppenee. Tarvitsemme kuitenkin myös virhearvion termille  $E_n$ . Haemme siksi toisenlaista todistusta ratkaisun olemassaololle.

Olkoot  $n, p \in \mathbb{N}$ . Huomaamme ensin — kuten ratkaisun yksikäsitteisyyden todistuksessa —, että

$$|E_{n+1} - E_n| = e |\sin E_n - \sin E_{n-1}| \leq e |E_n - E_{n-1}| \leq \dots \leq e^n |E_1 - E_0| = e^n |M|.$$

Kolmioepäyhtälön nojalla edelleen

$$\begin{aligned} |E_{n+p} - E_n| &\leq |E_{n+p} - E_{n+p-1}| + |E_{n+p-1} - E_{n+p-2}| + \dots + \\ &\quad + |E_{n+1} - E_n| \leq e^{n+p-1} |M| + e^{n+p-2} |M| + \dots + e^n |M| = \\ &= e^n |M| (1 + e + \dots + e^{p-1}) \leq e^n \frac{|M|}{1 - e}. \end{aligned} \quad (4.5)$$

Täten  $|E_n - E_m| \rightarrow 0$ , kun  $n, m \rightarrow \infty$ , eli  $(E_n)$  on Cauchyn jono. Koska jokainen lukusuoran Cauchyn jono suppenee, on olemassa  $E = \lim_{n \rightarrow \infty} E_n$ . Tällöin (4.5) antaa virhearvion

$$|E - E_n| \leq e^n \frac{|M|}{1 - e}, \quad n \in \mathbb{N}. \quad (4.6)$$

Suppeneminen on nopeaa, sillä esimerkiksi Marsille  $e = 0,093$ .

### 4.1.2 Banachin kiintopistelause

Yleistämme edellisessä jaksossa löytämämme ongelman ja sen ratkaisumenetelmän. Olkoon  $V$  epätyhjä joukko ja  $d: V \times V \rightarrow [0, \infty)$  kuvaus, joka toteuttaa seuraavat kolme ehtoa.

1.  $d(x, y) = d(y, x)$  kaikilla  $x, y \in V$ . (Symmetrisyys).

<sup>4</sup>Jos kuvauksella  $f: [a, b] \rightarrow [c, d]$ ,  $a < b$ ,  $c < d$ , on käänteisfunktio  $f^{-1} = g$ , jonka pisteessä  $c$  kehitetty Taylorin sarja suppenee välillä  $[c, d]$ , niin yhtälön  $f(x) = y$ ,  $y \in [c, d]$ , ratkaisu on

$$x = g(y) = \sum_{j=0}^{\infty} \frac{g^{(j)}(c)}{j!} (y - c)^j.$$

Derivaatat  $g^{(j)}(c)$  voidaan laskea ketjusäännön avulla ja Taylorin kaava antaa arvion sarjan katkaisemisesta syntyvälle virheelle.

2.  $d(x, z) \leq d(x, y) + d(y, z)$  kaikilla  $x, y, z \in V$ . (Kolmioepäyhtälö).
3.  $d(x, y) = 0$ , jos ja vain jos  $x = y$ . (Definiittisyys).

Sanomme paria  $(V, d)$  *metriseksi avaruudeksi* ja kuvausta  $d$  sen *metriikaksi*. Jono  $(x_n)$  metrisen avaruuden  $(V, d)$  alkioita on *Cauchyn jono*, jos

$$d(x_m, x_n) \rightarrow 0, \text{ kun } m, n \rightarrow \infty.$$

Metrisen avaruus  $(V, d)$  on *täydellinen*, jos sen jokainen Cauchyn jono  $(x_n)$  *suppenee* eli on olemassa  $x \in V$  siten, että  $d(x_n, x) \rightarrow 0$ , kun  $n \rightarrow \infty$ .

Olkoon  $(V, d)$  metrisen avaruus ja  $F: V \rightarrow V$ . Pistettä  $x \in V$ , joka toteuttaa yhtälön  $F(x) = x$ , sanomme kuvauksen  $F$  *kiintopisteeksi*. Kuvausta  $F$  sanomme *vahvaksi kutistuskuvaukseksi*, jos on olemassa  $\alpha \in [0, 1)$  siten että

$$d(F(x), F(y)) \leq \alpha d(x, y) \text{ kaikilla } x, y \in V. \quad (4.7)$$

Jos (4.7) toteutuu luvulla  $\alpha = 1$ , sanomme kuvausta  $F$  *kutistuskuvaukseksi*. Lukua  $\alpha$  sanomme kuvauksen  $F$  *suurennussuhteeksi*.

**Lause 4.1.1** (*Banachin kiintopistelause*). *Olkoon  $(V, d)$  täydellinen metrisen avaruus,  $x_0 \in V$ ,  $F: V \rightarrow V$  vahva kutistuskuvauksena ja  $\alpha$  sen suurennussuhde. Silloin kuvauksella  $F$  on yksikäsitteinen kiintopiste  $x \in V$ . Lisäksi jono  $(x_n)$ ,  $x_{n+1} = F(x_n)$ ,  $n = 0, 1, 2, \dots$ , toteuttaa epäyhtälön*

$$d(x, x_n) \leq \alpha^n \frac{d(x_0, F(x_0))}{1 - \alpha}, \quad n = 1, 2, \dots \quad (4.8)$$

*Todistus.* Olkoot  $n, p \in \mathbb{N}$ . Koska  $F$  on vahva kutistuskuvauksena,

$$d(x_{n+1}, x_n) = d(F(x_n), F(x_{n-1})) \leq \alpha d(x_n, x_{n-1}) \leq \dots \leq \alpha^n d(x_1, x_0).$$

Kolmioepäyhtälön nojalla edelleen

$$\begin{aligned} d(x_{n+p}, x_n) &\leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots \\ &\quad \dots + d(x_{n+1}, x_n) \leq (\alpha^{n+p-1} + \alpha^{n+p-2} + \dots + \alpha^n) d(x_1, x_0) \leq \\ &\leq \alpha^n (1 + \alpha + \alpha^2 + \dots) d(x_1, x_0) \leq \alpha^n \frac{d(x_1, x_0)}{1 - \alpha}. \end{aligned}$$

Täten  $d(x_n, x_m) \rightarrow 0$ , kun  $n, m \rightarrow \infty$ , eli  $(x_n)$  on Cauchyn jono. Koska  $(V, d)$  on täydellinen,  $(x_n)$  suppenee kohti jotakin  $x \in V$ . Kolmioepäyhtälön nojalla nyt

$$d(x, x_n) \leq d(x, x_{n+p}) + d(x_{n+p}, x_n) \leq d(x, x_{n+p}) + \alpha^n \frac{d(x_1, x_0)}{1 - \alpha},$$

mistä (4.8) seuraa, kun  $p \rightarrow \infty$ . Edelleen kolmioepäyhtälön nojalla

$$\begin{aligned} d(x, F(x)) &\leq d(x, x_{n+1}) + d(x_{n+1}, F(x)) \\ &= d(x, x_{n+1}) + d(F(x_n), F(x)) \leq \\ &\leq d(x, x_{n+1}) + \alpha d(x_n, x) \rightarrow 0, \text{ kun } n \rightarrow \infty. \end{aligned}$$

Metriikan definiittisyyden nojalla  $x = F(x)$  eli  $x$  on kuvauksen  $F$  kiintopiste.

Osoitetaan vielä kiintopisteen yksikäsitteisyys. Olkoot  $x$  ja  $\tilde{x}$  kiintopisteitä. Silloin

$$d(x, \tilde{x}) = d(F(x), F(\tilde{x})) \leq \alpha d(x, \tilde{x}) < d(x, \tilde{x}),$$

josta  $d(x, \tilde{x}) = 0$  eli  $x = \tilde{x}$ .

**m.o.t.**

Banachin kiintopistelause sisältää myös numeerisen menetelmän virhearvioineen kiintopisteen laskemiseksi. Tätä menetelmää sanotaan *iteroimiseksi* yhtälön  $f(x) = 0$  ratkaisemiseksi. Ensin muokkaamme alkuperäisestä yhtälöstä  $f(x) = 0$  lausekkeen  $x = F(x)$ . Kuvaus  $F$  ei luonnollisestikaan määräydy yksikäsitteisesti, joten pyrimme valitsemaan sellaisen  $F$ , joka on vahva kutistuskuvauksena ja vielä kohtuullisen pienellä suurennussuhteella  $\alpha$ , sillä silloin iterointi suppenee nopeasti.

### 4.1.3 Diskreetit dynaamiset järjestelmät

Kiintopistelauseet ovat sovellutustensa kannalta tärkeä osa matematiikkaa: differentiaali- ja integraaliyhtälöiden ratkeavuus voidaan usein palauttaa kysymykseksi sopivan apufunktion kiintopisteen olemassaolosta. Esimerkiksi Banachin kiintopistelauseella näytetään, että alkuarvotehtävällä

$$y'(t) + Ly(t) = 0, \quad t > 0, \quad y(0) = y_0,$$

on yksikäsitteinen jatkuvasti differentioituva ratkaisu  $y: [0, \infty) \rightarrow \mathbb{R}^m$ ,  $m = 1, 2, \dots$ , kunhan  $L: \mathbb{R}^m \rightarrow \mathbb{R}^m$  on Lipschitz-jatkuva.

Tarkastellaan seuraavaa differentiaaliyhtälöä alkuehdon kera

$$y'(t) = f(y(t)), \quad t > 0, \quad y(0) = y_0, \quad (4.9)$$

missä  $f: \mathbb{R}^m \rightarrow \mathbb{R}^m$ ,  $m = 1, 2, \dots$ . Yksinkertainen numeerinen menetelmä tämän ratkaisemiseksi perustuu derivaatan korvaamiseen erotusosamäärällä. Silloin saamme *differenssiyhtälön*

$$\frac{y(t_{n+1}) - y(t_n)}{t_{n+1} - t_n} = f(y(t_n)), \quad n = 0, 1, 2, \dots, \quad (4.10)$$

alkuehdolla  $y(t_0) = y_0$ ; tässä  $0 = t_0 < t_1 < t_2 < \dots$ . Merkitsemme

$$x_n = y(t_n), \quad F_n(x) = x + (t_{n+1} - t_n)f(x), \quad n = 0, 1, 2, \dots \quad (4.11)$$

Silloin differenssiyhtälöstä (4.7) tulee iterointikaava

$$x_{n+1} = F_n(x_n), \quad n = 0, 1, 2, \dots \quad (4.12)$$

Tähän samaan kaavaan päädytään, kun taloudellisen järjestelmän aika on *diskreetti*. Näin on esimerkiksi silloin, kun yrittäjä päättää illalla suljettuaan liikkeensä ja laskettuaan päivän ostot ja myynnit, miten hän seuraavana päivänä käyttäytyy. Ajan askelpituus on tällöin yksi päivä. Kun aika on diskreetti, ei iterointikaava (4.12) ole approksimaatio vaan tarkka.

Olkoon aika diskreetti ja sen askelpituus vakio. Silloin  $F_n = F$  kaavassa (4.11) ja Banachin kiintopistelause antaa oitis riittävän ehdon sille, että dynaaminen järjestelmä (4.12) kehittyy kohti stationaarista tilaa, kun  $n \rightarrow \infty$ . Riittää, että kuvaus  $F$  on vahva kutistuskuvaus, mikä toteutuu, jos  $F(x) = x - \beta x$  ja  $\beta \in (0, 2)$ . Itse asiassa

$$x_n = (1 - \beta)^n x_0 = \begin{cases} \rightarrow 0, & \text{jos } 0 < \beta < 2, \\ = x_0, & \text{jos } \beta = 0 \text{ tai } x_0 = 0, \\ \rightarrow \infty, & \text{jos } \beta < 0 \text{ ja } x_0 > 0, \\ \rightarrow -\infty, & \text{jos } \beta < 0 \text{ ja } x_0 < 0, \\ \nrightarrow & \text{muulloin.} \end{cases}$$

## 4.2 Gradienttimenetelmä

Olkoon<sup>5</sup>  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $m = 1, 2, \dots$ , differentioituva. Sen differentiaali  $df = (\nabla f(x) | \cdot)$  pisteessä  $x \in \mathbb{R}^m$  approksimoi funktion  $f$  muutosta tuossa pisteessä<sup>6</sup>:

$$f(x+h) - f(x) = (\nabla f(x) | h) + \epsilon_x(h) \|h\|,$$

<sup>5</sup>Tämän jakson yksityiskohdista tarkemmin esim. Kaisa Miettinen, Optimointi, Jyväskylän yliopisto, matematiikan laitos, luentomoniste 33, 1995, ss. 33-39.

<sup>6</sup>Merkitään

$$\|x\| = \left( \sum_{i=1}^m x_i^2 \right)^{1/2} \text{ ja } (x|y) = \sum_{i=1}^m x_i y_i \text{ kaikilla } x, y \in \mathbb{R}^m.$$

missä  $h \in \mathbb{R}^m$  ja  $\epsilon_x(v) \rightarrow 0$  kun  $\|v\| \rightarrow 0$ . Tästä saadaan edelleen, että funktion  $f$  suunnattu derivaatta  $D_h f(x)$  pisteessä  $x \in \mathbb{R}^m$  suuntaan  $h \in \mathbb{R}^m$ ,  $\|h\| = 1$ , on funktion muutosnopeus tuohon suuntaan ja

$$D_h f(x) = \lim_{\delta \rightarrow 0^+} \frac{f(x + \delta h) - f(x)}{\delta} = (\nabla f(x)|h).$$

Cauchyn-Bunjakovskin-Schwarzin epäyhtälö sanoo, että

$$- \|a\| \|b\| \leq (a|b) \leq \|a\| \|b\|,$$

missä jälkimmäisessä pätee yhtäsuuruus vain jos vektorit  $a \neq 0$  ja  $b \neq 0$  ovat samansuuntaisia eli  $a = \lambda b$ ,  $\lambda > 0$ . Näin funktio  $f$  kasvaa pisteessä  $x \in \mathbb{R}^m$  nopeiten suuntaan<sup>7</sup>  $\nabla f(x)/\|\nabla f(x)\|$ . Tähän huomioon perustuu *gradienttimenetelmä* eli jyrkimmän nousun menetelmä funktion  $f: \mathbb{R}^m \rightarrow \mathbb{R}$  (lokaalin) maksimin<sup>8</sup> löytämiseksi. Sen algoritmi on seuraava:

1. Valitse lopetusparametri  $\epsilon > 0$  ja alkupiste  $x_0 \in \mathbb{R}^m$ . Aseta  $n = 0$ .
2. Laske  $d_n := \nabla f(x_n)$ . Jos  $\|d_n\| < \epsilon$ , lopeta.
3. Ratkaise yksiulotteinen maksimointitehtävä

$$\max\{f(x_n + \lambda d_n) \mid \lambda \geq 0\}$$

ja merkitse sen ratkaisua  $\lambda_n$ .

4. Aseta  $x_{n+1} := x_n + \lambda_n d_n$  ja  $n := n + 1$ . Jatka kohdasta 2.

Lopetusehto  $\|d_n\| < \epsilon$  korvataan toisinaan jollakin seuraavista ehdoista:

$$\max_{i=1,\dots,m} \left\{ \left| \frac{\partial f}{\partial x_i} \right| \right\} < \epsilon, \quad \|\nabla f\| < \epsilon, \quad |f(x_{n+1}) - f(x_n)| < \epsilon, \quad \|x_{n+1} - x_n\| < \epsilon.$$

Funktion  $f$  gradientti lasketaan tavallisesti numeerisesti kaavasta

$$\frac{\partial f}{\partial x_i} \approx \frac{f(x + \delta e_i) - f(x)}{\delta},$$

missä  $\delta \approx 0$  ja  $e_i$  on  $i$ :nneen koordinaattiakselin suuntainen yksikkövektori.

Yksinkertaistetussa gradienttimenetelmässä algoritmin toisen askeleen maksimointitehtävä jätetään pois: valitaan joko  $\lambda_n$  kiinteäksi positiiviseksi parametrikseksi tai sitten sitä muutellaan algoritmin edetessä. Tällainen yksinkertaistettu gradienttimenetelmä on taloustieteen ja taloudellisen mallintamisen kannalta muutakin kuin näppärä numeerisen matematiikan keino. Voittoaan tai hyötyään maksimoiva puutteellisesti informoitu yrittäjä tai kuluttaja käyttäytyy kuin yksinkertaistettu gradienttimenetelmä. He laskevat voittonsa tai hyötynsä gradientin kullakin ajan hetkellä ja muuttavat käyttäytymistään siihen suuntaan, mihin voitto tai hyöty näyttää kasvavan nopeiten.

Gradienttimenetelmää ilman yksinkertaistuksia käyttävä yrittäjä voisi tehdä tai teettää markkinatutkimuksia, mutta ne olisivat niin kalliita tai vaikeita, että voittofunktion lauseke on kohtuullisin vaivoin ja kustannuksin saatavissa vain yhteen suuntaan.

Numeerisia menetelmiä maksimin löytämiseksi on muitakin ja ne soveltuvat myös mallintamiseen samoin suoraan: gradienttimenetelmää parempi menetelmä on *Newtonin menetelmä*<sup>9</sup>, jonka algoritmi on seuraava:

<sup>7</sup>Suunnattu derivaatta on kasvunopeus, jos suunnan normi on 1. Muuten kelpaisi pelkkä  $\nabla f(x)$ .

<sup>8</sup>Jos täytyisi hakea funktion  $f$  minimiä, käytettäisiin hyväksi sitä, että suuntaan  $-\nabla f(x)/\|\nabla f(x)\|$  funktio  $f$  pienenee nopeiten. Silloin algoritmin toista askelta muutetaan seuraavasti:  $d_n := -\nabla f(x_n)$ . Menetelmän nimi on tällöin jyrkimmän laskun menetelmä.

<sup>9</sup>Newtonin menetelmä vaatii enemmän säännöllisyyttä maksimoitavalta funktioltä ja siinä mielessä se on huonompi kuin gradienttimenetelmä. Toisaalta Newtonin menetelmä suppenee nopeammin etenkin ääriarvokohdan lähellä.

1. Valitse lopetusparametri  $\epsilon > 0$  ja alkupiste  $x_0 \in \mathbb{R}^m$ . Aseta  $n = 0$ .
2. Laske  $d_n := -H(x_n)^{-1} \nabla f(x_n)$  missä  $H(x_n)$  on funktion  $f$  Hessen matriisi<sup>10</sup> pisteessä  $x_n \in \mathbb{R}^m$ . Jos  $\|d_n\| < \epsilon$ , lopeta.
3. Aseta  $x_{n+1} := x_n + d_n$  ja  $n := n + 1$ . Jatka kohdasta 2.

Newtonin menetelmä perustuu siihen, että maksimoitavaa kuvausta approksimoidaan pisteessä  $x_n$  sen toisen asteen Taylorin polynomilla ja pisteeksi  $x_{n+1}$  valitaan approksimaatin maksimikohta<sup>11</sup>. Täten Newtonin menetelmä ratkaisee samalla yhtälöä  $\nabla f(x) = 0$ . Newtoninkaan menetelmä ei vaadi sitä soveltavaa yrittäjää tietämään tulevaisuutta, vaan laskemaan toteutuneista myynneistään ja kuluistaan ensimmäisen ja toisen kertaluvun osittaisderivaatat.

#### 4.2.1 Vapaa markkinatalous vs. suunnitelmatalous

Niin gradienttimenetelmä kuin Newtonin menetelmä soveltuvat lokaalien ääriarvokohtien hakemiseen. On helppo muodostaa kuvauksia  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , joiden globaaleja ääriarvoja kumpikaan niistä ei löydy (tämän havainnollistaminen käy parhaiten piirtämällä korkeuskäyriä sopivasta maastosta tai muovaamalla lumesta pinta, jolla lumipallo saa hakea alinta kohtaa). Edelleen tästä huomiosta seuraa, että voittoa tai hyötyään maksimoivien taloudellisten toimijoiden yhteisö ei välttämättä osaa hakeutua tilaan, jossa sen enempää kokonaisvoiton ja -hyödyn summa tai edes yksittäisten toimijoiden voitto tai hyöty olisi mahdollisimman suuri. Voi jopa käydä niin, että yksilöiden voitontavoittelu tuhoaa koko yhteisön. Tavallinen historiallinen esimerkki talouden ohjaamisesta väkisin kohti suurempia voittoja ja hyötyjä on sähköistämisen tai rautateiden rakentamisen kaltaiset suurinvestoinnit: Euroopassa ne on toteuttanut valtio.

Tarkastellaan taloutta, jossa tuotetaan ja kulutetaan vain yhtä tuotetta, joka on nopeasti pilaantuva ja jonka yksikköhinta hetkellä  $t$  on  $p(t)$ . Sen tuotantonopeus hetkellä  $t$  on  $q_s(t)$  ja kysyntänopeus on  $q_d(t)$ . Tuottajat voivat päättää vapaasti tuotantonopeutensa ja kuluttajat kysyntänopeutensa. Tuottajien voitto aikayksikössä hetkellä  $t$  on

$$\Pi(t) = \min\{q_d(t), q_s(t)\} - C(q_d(t)),$$

missä  $C: \mathbb{R} \rightarrow \mathbb{R}$  esittää tuotantokustannuksia aikayksikössä hetkellä  $t$ . Nyt

$$\frac{\partial \Pi}{\partial q_s} = \begin{cases} p - C'(q_s), & \text{jos } q_s \leq q_d, \\ -C'(q_s), & \text{jos } q_s \geq q_d, \end{cases} = H(q_d - q_s)p - C'(q_s),$$

missä  $H$  on Heavisiden askelfunktio. Kuluttajien hyöty aikayksikössä hetkellä  $t$  on

$$U(t) = P(q_d(t)) - p(t)q_d(t),$$

missä  $P(q_d(t))$  esittää heidän mielihyväänsä aikayksikössä. Nyt

$$\frac{\partial U}{\partial q_d} = h(q_d) - p,$$

missä  $h: \mathbb{R} \rightarrow \mathbb{R}$  esittää kuluttajien maksuhalukkuutta. Yksinkertaistetun gradienttimenetelmän mukaisesti

$$q_d'(t) = \alpha \left( h(q_d(t)) - p(t) \right), \quad t \in \mathbb{R}, \quad (4.13)$$

$$q_s'(t) = \beta \left( p(t)H(q_d(t) - q_s(t)) - C'(q_s(t)) \right), \quad t \in \mathbb{R}, \quad (4.14)$$

<sup>10</sup>Kuvauksen  $f: \mathbb{R}^m \rightarrow \mathbb{R}$  toisen kertaluvun differentiaali  $h \mapsto d^2 f(x; h)$  on neliömuoto  $\mathbb{R}^m \rightarrow \mathbb{R}$ ,  $h \mapsto (h_1 \ h_2 \ \dots \ h_m) H(x) (h_1 \ h_2 \ \dots \ h_m)^T$ .

<sup>11</sup>Jotta Newtonin menetelmän algoritmi olisi kaikille  $x_0 \in \mathbb{R}^m$  hyvin määritelty, on toisen kertaluvun differentiaalin  $h \mapsto d^2 f(x; h)$  oltava kaikilla  $x \in \mathbb{R}^m$  olemassa ja definiitti. Jos se on positiivisesti definiitti kaikkialla, ei funktiolla ole maksimeita.

missä  $\alpha$  ja  $\beta$  ovat positiivisia vakioita.

Markkinataloudessa tuotteen yksikköhinta määräytyy kysynnän ja tarjonnan lain mukaan. Paul Samuelson<sup>12</sup> ilmaisi sen differentiaaliyhtälöllä

$$p'(t) = \gamma(q_d(t) - q_s(t)), \quad t \in \mathbb{R}, \quad (4.15)$$

missä  $\gamma > 0$  on vakio. Differentiaaliyhtälöt (4.13)-(4.15) kuvaavat yksinkertaista vapaata markkinataloutta: tuottajat tavoittelevat suurinta mahdollista voittoa, kuluttajat suurinta mahdollista hyötyään ja tuotteen hinta määräytyy markkinoilla. Yhtälöistä (4.13)-(4.15) voidaan muodostaa myös diskreetit versiot, joko itse malliksi tai sitten jatkuvan mallin numeeriseksi ratkaisemiseksi. Joka tapauksessa voimme nyt konstruoida esimerkkejä, joissa vapaa markkinatalous hakeutuu tai ei hakeudu ilman ulkopuolista ohjausta kohti tilaa, joka antaisi tuottajille suurimman voiton ja kuluttajille suurimman hyödyn (erikseen tai minä hyvänsä painotettuna summana).

### 4.3 Harjoitustehtäviä

1. Tarkastellaan yhtälöä  $f(x) = 2x + e^x = 0$ ,  $x \in \mathbb{R}$ . Osoita jatkuvien funktioiden teorian avulla, että sillä on täsmälleen yksi ratkaisu. Hae ratkaisun likiarvo *haarukoimalla* eli seuraavalla algoritmilla:

- (a) Valitse lopetusparametri  $\epsilon > 0$  sekä ja väli  $[x_0, x_1]$  siten, että  $f(x_1)f(x_2) < 0$ . Aseta  $n := 0$ .
- (b) Jos  $|x_{n+1} - x_n| < \epsilon$ , niin lopeta. Muuten aseta  $x_{n+2} := \frac{1}{2}(x_n + x_{n+1})$ . Jos  $f(x_n)f(x_{n+2}) < 0$ , aseta  $x_{n+3} := x_n$ . Jos  $f(x_n)f(x_{n+2}) > 0$ , aseta  $x_{n+3} := x_{n+1}$ . Jos  $f(x_n)f(x_{n+2}) = 0$ , aseta  $x_{n+3} := x_{n+2}$ .
- (c) Aseta  $n := n + 2$  ja jatka kohdasta (b).

Hae ratkaisun likiarvo taskulaskimen laskentatarkkuudella.

2. Ratkaise edellisen tehtävän yhtälö iteroiden taskulaskimen laskentatarkkuudella. Yritä ratkaista iteroiden yhtälö  $-2x + e^x = 0$ ,  $x \in \mathbb{R}$ . Kokeile alkuarvoja  $x_0 = 1$  ja  $x_0 = 2$ .
3. Hae funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto -x^4 + 2x^3 + x$  maksimit taskulaskimen laskentatarkkuudella yksinkertaistetulla gradienttimenetelmällä ( $\lambda_n = 1$ ).
4. Ratkaise edellinen tehtävä taskulaskimen laskentatarkkuudella Newtonin menetelmällä. Hessen matriisin käänteismatriisi on nyt  $1/f''(x)$ .
5. Muodosta sellainen  $C^\infty$ -funktio (kaikki derivaatat ovat olemassa ja jatkuvia)  $f: \mathbb{R} \rightarrow \mathbb{R}$ , jonka kaikki lokaalit ja globaalit maksimit jäävät niin gradientti- kuin Newtonin menetelmällä löytymättä, kun alkuarvo  $x_0 \in \mathbb{R}$  pidetään kiinnitettynä.
6. Kuvatkoot vapaata markkinataloutta differentiaaliyhtälöt (4.13)-(4.15). Olkoon edelleen  $\alpha = \beta = \gamma = 1$ ,

$$\begin{aligned} C(q) &= \begin{cases} \frac{1}{2}q^2, & \text{jos } q \leq 1, \\ \frac{1}{3}(q-2)^3 + \frac{5}{6}, & \text{jos } q > 1, \end{cases} \\ P(q) &= \begin{cases} q - \frac{1}{2}q^2 + \frac{5}{6}, & \text{jos } q \leq 1, \\ -\frac{1}{3}(q-2)^3 + q, & \text{jos } q > 1, \end{cases} \end{aligned} \quad (4.16)$$

<sup>12</sup>Ks. Samuelson, P. A: *The stability of equilibrium*, Econometrica 9, (1941), ss. 97-120. Uusintapainos teoksessa "The Collected Scientific Papers of Paul A. Samuelson I", toim. Joseph E. Stiglitz, ss. 539-562, 565-589. M.I.T. Press, Cambridge Mass., 1966.

Todistele laskemalla tai hahmottelemalla funktioiden  $p(t)$ ,  $q_d(t)$  ja  $q_s(t)$  kuvaajia, että jos alkuehtona on  $(p(0), q_d(0), q_s(0)) = (0, 0, 0)$ , niin

$$p(t) \rightarrow \frac{1}{2}, \quad q_d(t) \rightarrow \frac{1}{2}, \quad q_s(t) \rightarrow \frac{1}{2}, \quad U(t) \rightarrow \frac{3}{4}, \quad \Pi(t) \rightarrow \frac{1}{8},$$

kun  $t \rightarrow \infty$ .

7. Edellisessä tehtävässä vapaa markkinatalous sai kehittyä omien lakiansa mukaan. Oleta, että hallinnollisella päätöksellä sen kulkuun puututaan ja ajanhetkellä  $t = 1$  pakotetaan sen tilaksi  $(p(1), q_d(1), q_s(1)) = (\frac{1}{2}, 2, 2)$ . Todistele, että sen jälkeen sekä kuluttajien hyöty että tuottajien voitto hakeutuvat kohti suurempia arvoja kuin edellisessä tehtävässä.
8. Muutetaan tehtävää 6 seuraavasti:

$$C(q) = \frac{1}{2}q^2, \quad P(q) = \frac{1}{3}q^3, \quad p(0) = 1, \quad q_d(0) = 2, \quad q_s(0) = 1.$$

Todistele tai todista, että on olemassa  $t_1 > 0$  siten, että  $q_d(t) \rightarrow \infty$ , kun  $t \rightarrow t_1 -$ . Mitä psykologista ilmiötä kasvava maksuhalukkuus voisi esittää? Mikä mahtaisi olla tuote?

# Luku 5

## Variaatiolaskentaa

Variaatiolaskennassa haetaan kuvauksen  $J: A \rightarrow \mathbb{R}$  ääriarvoja, kun  $A$  on jokin funktioiden joukko. Kuvauksen eli nyt funktionaalin  $J$  derivaattakuvausta on perinteisesti sanottu ”variaatioksi” ja se syntyy, kun funktionaalin arvon antavaa pistettä (funktiota) ”varioidaan”. Ääriarvokohdassa variaation tulee hävitä. Tästä tulee nimitys variaatiolaskenta. Todistamme tässä luvussa välttämättömän ehdon (Eulerin-Lagrangen yhtälö) sille, että funktio  $\phi$  on integraalilausekkeen avulla määritellyn funktionaalin  $\phi \mapsto J(\phi)$  ääriarvokohta. Opimme myös johtamaan osittaisdifferentiaaliyhtälöitä variaatioperiaatteista. Sidottuja ääriarvotehtäviä tarkastelemme omassa jaksossaan: yleistämme Lagrangen kertojen menetelmän variaatiolaskentaan.

### 5.1 Suunnattu derivaatta ja ääriarvot

Yleistämme aluksi kuvauksen  $F: \mathbb{R}^m \rightarrow \mathbb{R}^n$  suunnatun derivaatan ja reaalfunktion lokaalin ääriarvon käsitteet. Vektorit ja matriisit -kurssilta muistamme, että kuvausten  $h: B \rightarrow \mathbb{R}^m$ ,  $B \neq \emptyset$ , joukko on vektoriavaruus laskutoimitusten  $(g, h) \mapsto g + h$  ja  $(\lambda, h) \mapsto \lambda h$  suhteen, missä

$$(g + h)(x) = g(x) + h(x), (\lambda h)(x) = \lambda h(x) \text{ kaikilla } x \in B, \lambda \in \mathbb{R}.$$

Näin meidän riittää olettaa, että kuvauksen  $F$  määrittelyjoukko on yleinen vektoriavaruus, joka voi olla ääretönulotteinen. Toisaalta erotusosamäärän muodostamiseksi riittää, että kuvauksen maalijoukko on sekin vektoriavaruus. Mutta erotusosamäärästä olisi otettava vielä raja-arvo. Siksi maaliavaruuden on oltava myös topologinen avaruus. Määrittelemme sen alaviitteessä<sup>1</sup>; tässä riittää huomata, että jokainen metrinen avaruus on myös topologinen avaruus ja siten

---

<sup>1</sup>Olkoon  $X$  epätyhjä joukko ja  $\mathcal{T}$  jokin sen osajoukkojen joukko. Sanotaan, että  $\mathcal{T}$  on *topologia* ja  $(X, \mathcal{T})$  on *topologinen avaruus*, jos

1.  $\emptyset, X \in \mathcal{T}$ .
2. Jos  $A, B \in \mathcal{T}$ , niin  $A \cap B \in \mathcal{T}$ .
3. Jos  $I$  on epätyhjä ja  $A_\alpha \in \mathcal{T}$  kaikilla  $\alpha \in I$ , niin  $\cup_{\alpha \in I} A_\alpha \in \mathcal{T}$ .

Käsitteet funktion jatkuvuus ja raja-arvo voidaan määritellä avoimien joukkojen avulla: funktio on jatkuva, jos jokaisen avoimen joukon alkukuva on avoin;  $x_n \rightarrow x$ , kun  $n \rightarrow \infty$ , jos jokaiselle avoimelle joukolle  $A$  on olemassa  $N \in \mathbb{N}$  siten että  $x_n, x \in A$ , kun  $n > N$ . Joukossa  $\mathbb{R}$  nämä määriteltiin itseisarvon avulla. Joukossa  $\mathbb{R}^m$  käytettiin samoja määritelmiä siten korjattuina, että itseisarvon paikalla on normi  $\|\cdot\|$ . Samat määritelmät kelpasivat yleiseen normiavaruuteen (vektoriavaruus, jossa on määritelty normi). Seuraava askel oli siirtyä metriseen avaruuteen: ilmaisut  $\|x - y\|$  korvattiin ilmaisulla  $d(x, y)$ . Topologinen avaruus on seuraava askel tässä yleistämisen ketjussa. Otetaan määritelmiksi sellaiset lukusuoralla  $\mathbb{R}$  pätevät jatkuvuuden ja raja-arvon yhtäpitävät ehdot, jotka käyttävät avoimen joukon käsitettä. Topologia on joukon  $X$  kaikkien avoimien joukkojen joukko; kun sanotaan, mikä topologia  $\mathcal{T}$  kullekin joukolle  $X$  on, silloin itse asiassa luetellaan kokoelma joukon  $X$  osajoukkoja, joita päätetään kutsua avoimiksi.

jokainen vektoriavaruus, jossa on metriikka, on topologinen avaruus. Metrisessä avaruudessa määritelmän mukaan

$$\lim_{\epsilon \rightarrow 0} x_\epsilon = x \text{ jos ja vain jos } \lim_{\epsilon \rightarrow 0} d(x_\epsilon, x) = 0.$$

Olkoon  $F: A \rightarrow W$ ,  $\emptyset \neq A \subset V$ , missä  $V$  on vektoriavaruus ja  $W$  topologinen vektoriavaruus<sup>2</sup>. Jos raja-arvo

$$D_h F(x) := \lim_{\epsilon \rightarrow 0} \frac{F(x + \epsilon h) - F(x)}{\epsilon}, \quad h \in V, \quad (5.1)$$

on olemassa, niin sanotaan, että  $D_h F(x)$  on kuvauksen  $F$  *suunnattu derivaatta pisteessä*  $x \in A$  *suuntaan*  $h \in V$ . Kuvausta  $h \mapsto D_h F(x)$  sanotaan kuvauksen  $F$  *Gâteaux-derivaataksi pisteessä*  $x$ . Huomaamme oitis<sup>3</sup>, että Gâteaux-derivaatta on homogeeninen funktio eli jos  $D_h F(x)$  on olemassa, niin myös  $D_{\alpha h} F(x)$  on olemassa ja

$$D_{\alpha h} F(x) = \alpha D_h F(x) \text{ kaikilla } \alpha \in \mathbb{R}. \quad (5.2)$$

Kuvauksen  $F: A \rightarrow \mathbb{R}$ ,  $A \neq \emptyset$ , suurin ja pienin arvo on jo reaalifunktion tapauksessa määritelty riittävän yleisesti: esimerkiksi  $F(x_0)$  on suurin arvo, jos

$$x_0 \in A \text{ ja } F(x_0) \geq F(x) \text{ kaikilla } x \in A.$$

Mutta lokaalien ääriarvojen käsite vaatii täydennystä. Tehdään se metrisen avaruuden tapauksessa<sup>4</sup>.

Olkoon  $(V, d)$  metrinen avaruus. Olkoot  $x_0 \in V$  ja  $r > 0$ . Joukkoa

$$B(x_0, r) := \{x \in V \mid d(x_0, x) < r\}$$

sanotaan  $x_0$ -*keskiseksi* ja  $r$ -*säteiseksi avoimeksi palloksi*. Piste  $x_0 \in A$  on kuvauksen  $F: A \rightarrow \mathbb{R}$  *lokaali minimikohta* ja  $F(x_0)$  on sen *lokaali minimi*, jos on olemassa  $r > 0$  siten että

$$F(x_0) \leq F(x) \text{ kaikilla } x \in A \cap B(x_0, r).$$

Samoin piste  $x_0 \in A$  on kuvauksen  $F: A \rightarrow \mathbb{R}$  *lokaali maksimikohta* ja  $F(x_0)$  on sen *lokaali maksimi*, jos on olemassa  $r > 0$  siten että

$$F(x_0) \geq F(x) \text{ kaikilla } x \in A \cap B(x_0, r).$$

Palautamme vielä mieleen normiavaruuden käsitteen. Pari  $(V, \|\cdot\|)$  on *normiavaruus* ja  $\|\cdot\|: V \rightarrow [0, \infty[$  on *normi*, jos seuraavat ehdot pätevät.

1. Joukko  $V$  on vektoriavaruus.

2.  $\|x - y\| \leq \|x\| + \|y\|$  kaikilla  $x, y \in V$ . (Kolmioepäyhtälö).

---

<sup>2</sup>Siis sekä vektoriavaruus että topologinen avaruus. Lisäksi kuvaukset  $(u, v) \mapsto u + v$ ,  $V \times V \rightarrow V$  ja  $(\lambda, u) \mapsto \lambda u$ ,  $\mathbb{R} \times V \rightarrow V$  ovat jatkuvia.

<sup>3</sup>Kun  $\epsilon \rightarrow 0$ ,

$$\frac{F(x + \epsilon \alpha h) - F(x)}{\epsilon} = \alpha \frac{F(x + \epsilon \alpha h) - F(x)}{\alpha \epsilon} \rightarrow \alpha D_h F(x).$$

<sup>4</sup>Yleisessä topologisessa avaruudessa  $(X, \mathcal{T})$  määrittely käy seuraavasti. Olkoon  $F: A \rightarrow \mathbb{R}$ ,  $\emptyset \neq A \subset X$ . Piste  $x_0 \in A$  on kuvauksen  $F: A \rightarrow \mathbb{R}$  *lokaali minimikohta* ja  $F(x_0)$  on sen *lokaali minimi*, jos on olemassa avoin joukko  $B \subset X$  eli  $B \in \mathcal{T}$  siten että

$$x_0 \in A \cap B \text{ ja } F(x_0) \leq F(x) \text{ kaikilla } x \in A \cap B.$$

Samoin piste  $x_0 \in A$  on kuvauksen  $F: A \rightarrow \mathbb{R}$  *lokaali maksimikohta* ja  $F(x_0)$  on sen *lokaali maksimi*, jos on olemassa avoin joukko  $B \subset X$  eli  $B \in \mathcal{T}$  siten että

$$x_0 \in A \cap B \text{ ja } F(x_0) \geq F(x) \text{ kaikilla } x \in A \cap B.$$

3.  $\|\lambda x\| = |\lambda| \|x\|$  kaikilla  $\lambda \in \mathbb{R}$  ja  $x \in V$ . (Homogeenisuus).

4. Jos  $\|x\| = 0$ , niin  $x = 0$ . (Definiittisyys).

Yksi tavallinen normiavaruus on  $(\mathbb{R}^m, \|\cdot\|_{\mathbb{R}^m})$ ,  $m = 1, 2, \dots$ , kun

$$\|x\|_{\mathbb{R}^m} = \left( \sum_{i=1}^m x_i^2 \right)^{1/2}.$$

Yksi ääretönulotteinen normiavaruus on välillä  $[0, 1]$  jatkuvien reaaliarvoisten funktioiden joukko  $C([0, 1])$ , kun normina on

$$\|f\| = \sup_{x \in [0, 1]} |f(x)|.$$

**Lause 5.1.1** *Olkoon  $V$  normiavaruus,  $\emptyset \neq A \subset V$ ,  $F: A \rightarrow \mathbb{R}$  ja  $x_0 \in A$ . Jos  $F(x_0)$  on kuvauksen  $F$  lokaali ääriarvo,  $h \in V$  ja suunnattu derivaatta  $D_h F(x_0)$  on olemassa, niin  $D_h F(x_0) = 0$ .*

*Todistus.* Todistus käy samoin kuin vastaavan reaalfunktioiden tuloksen todistaminen. Tehdään vastaoletus:  $D_h F(x_0) = \alpha > 0$  jollakin  $\alpha \in \mathbb{R}$ . Raja-arvon määritelmän mukaan on olemassa  $\delta > 0$  siten, että

$$\left| \frac{F(x_0 + \epsilon h) - F(x_0)}{\epsilon} - \alpha \right| < \frac{\alpha}{2}, \text{ kun } 0 < |\epsilon| < \delta.$$

Itseisarvon ominaisuuksien nojalla

$$\frac{F(x_0 + \epsilon h) - F(x_0)}{\epsilon} > \frac{\alpha}{2}, \text{ kun } 0 < |\epsilon| < \delta.$$

Tästä seuraa, että

$$\begin{aligned} F(x_0 + \epsilon h) &> \frac{\alpha}{2} \epsilon + F(x_0) > F(x_0), \text{ jos } \delta > \epsilon > 0, \\ F(x_0 + \epsilon h) &< \frac{\alpha}{2} \epsilon + F(x_0) < F(x_0), \text{ jos } -\delta < \epsilon < 0, \end{aligned}$$

joten  $F(x_0)$  ei olekaan kuvauksen  $F$  ääriarvo, sillä  $\|x_0 + \epsilon h - x_0\| = |\epsilon| \|h\|$  ja niin löytyy  $\epsilon$  siten, että  $x_0 + \epsilon h \in B(x_0, r)$ , oli  $r > 0$  sitten kuinka pieni tahansa. Samoin saadaan ristiriita, jos oletetaan, että  $D_h F(x_0) = \beta < 0$ . Siten  $D_h F(x_0) = 0$ . **m.o.t.**

Aivan samoin todistetaan seuraava tulos. Huomaa, että kuvauksen määrittelyjoukon riittää olla pelkkä vektoriavaruus.

**Lause 5.1.2** *Olkoon  $V$  vektoriavaruus,  $\emptyset \neq A \subset V$ ,  $F: A \rightarrow \mathbb{R}$  ja  $x_0 \in A$ . Jos  $F(x_0)$  on kuvauksen  $F$  globaali ääriarvo,  $h \in V$  ja suunnattu derivaatta  $D_h F(x_0)$  on olemassa, niin  $D_h F(x_0) = 0$ .*

### 5.1.1 Ääriarvon olemassaolon riittäviä ehtoja

Reaalfunktioiden teoriassa seuraavat tulokset takaavat ääriarvon olemassaolon.

**Lause 5.1.3** (*Weierstraßin lause*). *Olkoon  $f: [a, b] \rightarrow \mathbb{R}$ ,  $a < b$ , jatkuva. Silloin kuvauksella  $f$  on sekä suurin että pienin arvo.*

**Lause 5.1.4** *Olkoon  $f: ]a, b[ \rightarrow \mathbb{R}$ ,  $a < b$ , kahdesti derivoituva,  $x_0 \in ]a, b[$  ja  $f'(x_0) = 0$ . Jos  $f''(x) \leq 0$  kaikilla  $x \in ]a, b[$ , niin  $f(x_0)$  on kuvauksen suurin arvo. Jos  $f''(x) \geq 0$  kaikilla  $x \in ]a, b[$ , niin  $f(x_0)$  on kuvauksen  $f$  pienin arvo.*

*Todistus.* Todistetaan vain lauseen maksimiväittämä, koska minimiväittämä saadaan soveltamalla sitä kuvaukseen  $-f$ . Olkoon siis  $f''(x) \leq 0$  kaikilla  $x \in ]a, b[$ . Oletetaan antiteesiksi, että on olemassa  $x_1 \in ]a, b[$  siten, että  $x_1 \neq x_2$  ja  $f(x_1) > f(x_0)$ . Väliarvolauseetta kaksi kertaa soveltamalla saadaan, että

$$\begin{aligned} 0 < f(x_1) - f(x_0) &= f'(\xi)(x_1 - x_0) = (f'(\xi) - f'(x_0))(x_1 - x_0) = \\ &= f''(\eta)(\xi - x_0)(x_1 - x_0), \end{aligned}$$

missä  $\xi \in ]x_0, x_1[$  ja  $\eta \in ]x_0, \xi[$  tai  $\xi \in ]x_1, x_0[$  ja  $\eta \in ]\xi, x_0[$ . Koska kummassakin tapauksessa  $(\xi - x_0)(x_1 - x_0) > 0$ , onkin  $f''(\eta) > 0$ . Ristiriita, joten  $f(x_0)$  on suurin arvo. **m.o.t.**

Lause 5.1.4 antaa oitis riittävän ehdon kuvauksen  $]c, d[ \rightarrow R$ ,  $c < d$ , lokaalin ääriarvon olemassaololle: riittää, että sen rajoittuma johonkin väliin täyttää lauseen 5.1.4 ehdot. Weierstraßin lauseessa on oleellista se, että kuvauksen määrittelyjoukko on suljettu ja rajoitettu eli *kompakti*<sup>5</sup>. Lauseessa 5.1.4 avainehto on se, että kuvaus  $f$  on *kovera* tai *kupea*.

Usein haemme joko vain minimiä tai vain maksimia, joten Weierstraßin lause ja sen oletama funktion jatkuvuus on pilkottava osiin. Olkoon  $(X, d)$  metrinen avaruus ja  $F: X \rightarrow \mathbb{R}$ . Sanomme, että  $F$  on *alhaalta puolijatkuva pisteessä*  $x_0 \in X$ , jos

$$F(x_0) \leq \liminf_{x \rightarrow x_0} F(x)$$

eli jokaisella jonolla  $(x_n)$ , joka suppenee kohti pistettä  $x_0$ , pätee

$$F(x_0) \leq \liminf_{n \rightarrow \infty} F(x_n).$$

Tässä esiintyvä reaalilukujonon  $(\alpha_j)$  *alaraja-arvo* määritellään seuraavasti:

$$\liminf_{n \rightarrow \infty} \alpha_n = \sup_{k=1,2,\dots} \inf_{j=k,k+1,\dots} \alpha_j.$$

Jos  $\liminf_{n \rightarrow \infty} \alpha_n = \beta$ , niin on olemassa osajono  $(\alpha_{n_k})$  siten, että  $\lim \alpha_{n_k} = \beta$  ja mikään muu osajono ei suppene kohti pienempää lukua kuin  $\beta$ .

Sanomme, että  $F: X \rightarrow \mathbb{R}$  on *ylhäältä puolijatkuva pisteessä*  $x_0 \in X$ , jos

$$F(x_0) \geq \limsup_{x \rightarrow x_0} F(x)$$

eli jokaisella jonolla  $(x_n)$ , joka suppenee kohti pistettä  $x_0$ , pätee

$$F(x_0) \geq \limsup_{n \rightarrow \infty} F(x_n).$$

Tässä esiintyvä reaalilukujonon  $(\alpha_j)$  *yläraja-arvo* määritellään seuraavasti:

$$\limsup_{n \rightarrow \infty} \alpha_n = \inf_{k=1,2,\dots} \sup_{j=k,k+1,\dots} \alpha_j.$$

Jos  $\limsup_{n \rightarrow \infty} \alpha_j = \beta$ , niin on olemassa osajono  $(\alpha_{n_k})$  siten, että  $\lim \alpha_{n_k} = \beta$  ja mikään muu osajono ei suppene kohti suurempaa lukua kuin  $\beta$ . Edelleen aina pätee  $\liminf \alpha_n \leq \limsup \alpha_n$  ja —mikä parasta— ylä- ja alaraja-arvot ovat aina olemassa (mahdollisesti arvoinaan  $\infty$  tai  $-\infty$ ). Jos ylä- ja alaraja-arvo ovat samat, on lukujonolla tuo sama raja-arvo.

Näin funktion jatkuvuuden käsite on jaettu kahteen osaan: alhaalta ja ylhäältä puolijatkuvuuteen. Jos kuvaus on sekä alhaalta että ylhäältä puolijatkuva pisteessä  $x_0 \in X$ , niin

$$F(x_0) \leq \liminf_{x \rightarrow x_0} F(x) \leq \limsup_{x \rightarrow x_0} F(x) \leq F(x_0)$$

eli  $F$  on jatkuva pisteessä  $x_0 \in X$ :

$$\lim_{x \rightarrow x_0} F(x) = F(x_0).$$

<sup>5</sup>Ääretönulotteisissa avaruuksissa jokainen suljettu ja rajoitettu joukko ei ole kompakti.

Nyt voimme esittää Weierstraßin lauseen yleisen muodon:

**Lause 5.1.5** *Olkoon  $X \neq \emptyset$  sellainen joukko, että sen jokaisella jonolla on suppeneva osajono<sup>6</sup> ja  $F: X \rightarrow \mathbb{R}$ . Jos  $F$  on alhaalta puolijatkuva, niin sillä on pienin arvo. Jos  $F$  on ylhäältä puolijatkuva, niin sillä on suurin arvo.*

*Todistus.* Todistetaan vain maksimiväittämä, sillä minimiväittämä saadaan siitä tarkastelemalla kuvausta  $-F$ , sillä reaali lukujonojen  $(\alpha_n)$  ylä- ja alaraja-arvoille pätee

$$\liminf_{n \rightarrow \infty} \alpha_n = - \limsup_{n \rightarrow \infty} -\alpha_n.$$

Olkoon siis  $F$  ylhäältä puolijatkuva ja merkitään

$$\beta = \sup\{F(y) \mid y \in X\}.$$

Luku  $\beta \in \mathbb{R}$  tai  $\beta = \infty$ . Joka tapauksessa on olemassa jono  $(x_n)$  joukon  $X$  alkioita siten, että  $\lim_{n \rightarrow \infty} F(x_n) = \beta$ . Oletuksen mukaan tällä jonolla on osajono  $(x_{n_k})$ , joka suppenee kohti jotakin joukon  $X$  alkioita  $x$ . Ylhäältä puolijatkuvuuden nojalla

$$F(x) \geq \limsup_{k \rightarrow \infty} F(x_{n_k}) = \lim_{n \rightarrow \infty} F(x_n) = \beta \geq F(x),$$

joten  $F(x) = \beta \in \mathbb{R}$ . Siten  $F(x)$  on kuvauksen  $F$  suurin arvo. **m.o.t.**

**Lause 5.1.6** *Olkoon  $(V, \|\cdot\|)$  normiavaruus, jonka jokaisella rajoitetulla jonolla on suppeneva osajono, ja  $F: V \rightarrow \mathbb{R}$ . Jos  $F$  on alhaalta puolijatkuva ja*

$$\lim_{\|y\| \rightarrow \infty} F(y) = \infty, \quad (5.3)$$

*niin kuvauksella  $F$  on pienin arvo. Jos  $F$  on ylhäältä puolijatkuva ja*

$$\lim_{\|y\| \rightarrow \infty} F(y) = -\infty, \quad (5.4)$$

*niin kuvauksella  $F$  on suurin arvo.*

*Todistus.* Todistetaan vain minimiväittämä. Merkitään

$$\beta = \inf\{F(y) \mid y \in V\}.$$

Luku  $\beta \in \mathbb{R}$  tai  $\beta = -\infty$ . On olemassa jono  $(x_n)$  joukon  $V$  alkioita siten, että  $\lim_{n \rightarrow \infty} F(x_n) = \beta$ . Raja-arvoehdon (5.3) mukaan jono  $(x_n)$  on rajoitettu, jolloin sillä on osajono  $(x_{n_k})$ , joka suppenee kohti jotakin pistettä  $x \in V$ . Alhaalta puolijatkuvuuden nojalla

$$F(x) \leq \liminf_{k \rightarrow \infty} F(x_{n_k}) = \lim_{n \rightarrow \infty} F(x_n) = \beta \leq F(x).$$

Siten  $F(x) = \beta \in \mathbb{R}$  eli  $F(x)$  on kuvauksen  $F$  pienin arvo. **m.o.t.**

Esitetään seuraavaksi lauseen 5.1.4 yleistys. Huomaa, että kuvausta ei oleteta edes jatkuvaksi.

<sup>6</sup>Tällainen joukko on esimerkiksi kompakti joukko. Joukko on määritelmän mukaan *kompakti*, jos ja vain jos sen jokaisella avoimella peitteellä on äärellinen osapeite eli

$$X \subset \cup_{i \in I} A_i, \quad A_i \text{ on avoin} \quad \text{kaikilla } i \in I \implies X \subset A_{i_1} \cup \dots \cup A_{i_n}.$$

**Lause 5.1.7** Olkoon  $V$  vektoriavaruus  $F: A \rightarrow \mathbb{R}$ ,  $\emptyset \neq A \subset V$  ja  $x_0 \in A$  siten, että suunnattu derivaatta  $D_h F(x_0)$  on olemassa ja

$$D_h F(x_0) = 0 \text{ kaikilla } h \in \{w \in V \mid x_0 + w \in A\}.$$

Jos  $F$  on kupera, niin  $F(x_0)$  on sen pienin arvo. Jos  $F$  on kovera, niin  $F(x_0)$  on sen suurin arvo.

*Todistus.* Todistetaan vain minimiväittämä. Olkoon siis  $F$  kupera. Silloin

$$\begin{aligned} 0 = D_h F(x_0) &= \lim_{\epsilon \rightarrow 0} \frac{F(x_0 + \epsilon h) - F(x_0)}{\epsilon} = \\ &= \lim_{\epsilon \rightarrow 0+} \frac{F((1-\epsilon)x_0 + \epsilon(h+x_0)) - F(x_0)}{\epsilon} \leq \\ &\leq \limsup_{\epsilon \rightarrow 0+} \frac{(1-\epsilon)F(x_0) + \epsilon F(h+x_0) - F(x_0)}{\epsilon} = F(h+x_0) - F(x_0) \end{aligned}$$

kaikilla  $h \in V$  siten, että  $h+x_0 \in A$ . Siten  $F(x_0)$  on kuvauksen  $F$  pienin arvo. **m.o.t.**

Lopetetaan ääriarvojen olemassaolon riittävien ehtojen käsittely mainintaan, että refleksiivisessä Banachin avaruudessa<sup>7</sup> jokainen rajoitettu jono suppenee heikosti<sup>8</sup>. Edelleen osoittautuu, että jokainen alhaalta puolijatkuva kupera kuvaus on alhaalta puolijatkuva myös heikon suppenemisen suhteen. Tällöin lauseen 5.1.6 mukaan kuvauksella on pienin arvo. Tulos on epätriviaali ja hyvin käyttökelpoinen: se takaa esimerkiksi, että funktionaalilla

$$L^2(0,1) \rightarrow \mathbb{R}, y \mapsto \int_0^1 (y(x)^2 + g(y(x))) dx$$

on pienin arvo, jos  $g: \mathbb{R} \rightarrow \mathbb{R}$  on jokin alhaalta puolijatkuva kupera kuvaus.

### 5.1.2 Variaatiolemma

Esitetään tärkeä variaatiolaskennan aputulos. Olkoon  $\Omega \subset \mathbb{R}^m$ ,  $m = 1, 2, \dots$ , avoin joukko. Merkitään

$$\begin{aligned} C^\infty(\Omega) &= \{\phi: \Omega \rightarrow \mathbb{R} \mid \text{kaikki } \phi\text{:n osittaisderivaatat ovat} \\ &\quad \text{olemassa ja jatkuvia}\}, \\ C_0^\infty(\Omega) &= \{\phi \in C^\infty(\Omega) \mid \text{on olemassa kompakti}^9 K_\phi \subset \Omega \text{ siten,} \\ &\quad \text{että } \phi(x) = 0 \text{ kaikilla } x \in \Omega \setminus K_\phi\}. \end{aligned}$$

**Apulause 5.1.1** (*Variaatiolemma*). Olkoon  $\Omega \subset \mathbb{R}^m$ ,  $m = 1, 2, \dots$ , avoin ja  $f: \Omega \rightarrow \mathbb{R}$  integroituva<sup>10</sup> Jos

$$\int_\Omega f(x)\phi(x) dx = 0 \text{ kaikilla } \phi \in C_0^\infty(\Omega),$$

niin  $f = 0$ .

Variaatiolemman todistus perustuu siihen, että  $C_0^\infty$  funktiot ovat tiheässä integroituvien funktioiden joukossa  $L^1(\Omega)$  eli jokaiselle  $\epsilon > 0$  on olemassa  $\phi_\epsilon \in C_0^\infty(\Omega)$  siten, että

$$\|f - \phi_\epsilon\|_{L^1(\Omega)} := \int_\Omega |f(x) - \phi_\epsilon(x)| dx < \epsilon.$$

<sup>7</sup>Banachin avaruus on normiavaruus, jonka jokainen Cauchyn jono suppenee. Esimerkiksi  $L^p(\Omega)$  on Banachin avaruus, kunhan  $p \in [1, \infty]$ ,  $\Omega \subset \mathbb{R}^m$  avoin,  $m = 1, 2, \dots$ . Se on refleksiivinen, jos ja vain jos  $1 < p < \infty$ .

<sup>8</sup>Banachin avaruuden  $V$  jono  $(x_n)$  suppenee heikosti kohti alkia  $x$ , jos  $f(x_n) \rightarrow f(x)$  jokaisella jatkuvalla lineaarikuvauksella  $f: V \rightarrow \mathbb{R}$ .

<sup>10</sup>Riemann-integroituva tai Lebesgue-integroituva. Variaatiolemma pätee myös jos  $f \in L^p(\Omega)$ , missä  $p \in [1, \infty]$ .

## 5.2 Eulerin-Lagrangen yhtälö

Olkoon  $L: [t_0, t_1] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $t_0 < t_1$ , ja kirjoitetaan

$$J(y) = \int_{t_0}^{t_1} L(t, y(t), y'(t)) dt,$$

missä  $y: [t_0, t_1] \rightarrow \mathbb{R}$  on jatkuvasti derivoituva.

**Apulause 5.2.1** *Olkoon  $L(t, \cdot, \cdot): \mathbb{R}^2 \rightarrow \mathbb{R}$  differentioituva tasaisesti paramet-  
rin  $t \in [t_1, t_2]$  suhteen<sup>11</sup> ja olkoot  $L, \partial L / \partial y, \partial L / \partial y': [t_0, t_1] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  jat-  
kuvia. Tällöin<sup>12</sup> kuvauksen  $J$  suunnattu derivaatta pisteessä  $y \in C^1([t_1, t_2])$   
suuntaan  $h \in C^1([t_0, t_1])$  on olemassa ja se on*

$$D_h J(y) = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial y}(t, y(t), y'(t)) h(t) + \frac{\partial L}{\partial y'}(t, y(t), y'(t)) h'(t) \right) dt.$$

*Todistus.* Ensinnäkin  $J(y)$  ja  $D_h J(y)$  ovat hyvin määriteltyjä, sillä ne saadaan integroimalla jatkuvia funktioita. Olkoon  $\epsilon > 0$ . Silloin differentioituvuuden määritelmän mukaan

$$\begin{aligned} \frac{J(y + \epsilon h) - J(y)}{\epsilon} &= \\ &= \frac{1}{\epsilon} \int_{t_0}^{t_1} \left( L(t, y(t) + \epsilon h(t), y'(t) + \epsilon h'(t)) - L(t, y(t), y'(t)) \right) dt = \\ &= \frac{1}{\epsilon} \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial y}(t, y(t), y'(t)) \epsilon h(t) + \frac{\partial L}{\partial y'}(t, y(t), y'(t)) \epsilon h'(t) + \right. \\ &\quad \left. + \rho(t, \epsilon h(t), \epsilon h'(t)) \|(\epsilon h(t), \epsilon h'(t))\| \right) dt = \\ &= D_h J(y) + \int_{t_0}^{t_1} \rho(t, \epsilon h(t), \epsilon h'(t)) \|(h(t), h'(t))\| dt. \end{aligned}$$

Viimeinen termi lähestyy nollaa kun  $\epsilon \rightarrow 0$ , sillä sekä  $h$  että  $h'$  ovat jatkuvia, jolloin  $(h(t), h'(t))$  on rajoitettu tasaisesti kaikilla  $t \in [t_0, t_1]$ . **m.o.t.**

**Apulause 5.2.2** *Oletetaan apulauseen 5.2.1 ehtojen lisäksi, että kuvaus  $h \in C_0^\infty([t_0, t_1])$  ja että*

$$t \mapsto \frac{\partial L}{\partial y'}(t, y(t), y'(t))$$

*on paloittain jatkuvasti differentioituva. Silloin*

$$D_h J(y) = \int_{t_0}^{t_1} h(t) \left( \frac{\partial L}{\partial y}(t, y(t), y'(t)) - \frac{d}{dt} \frac{\partial L}{\partial y'}(t, y(t), y'(t)) \right) dt.$$

*Todistus.* Sovelletaan suunnatun derivaatan apulauseen 5.2.1 mukaisen lausek-  
keen toiseen termiin osittainintegroitikaavaa. Sijoitustermit häviävät, koska  
 $h(t_0) = h(t_1) = 0$ . **m.o.t.**

**Lause 5.2.1** *(Eulerin-Lagrangen yhtälö). Olkoon  $L: [t_0, t_1] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$   $C^2$ -  
funktio<sup>13</sup> ja olkoon*

$$V = \{w \in C^2([t_0, t_1]) \mid y(t_0) = a, y(t_1) = b\},$$

<sup>11</sup>Tällä tarkoitetaan sitä, että differentioituvuuden määritelmän jäännöstermissä funktio  $u \mapsto \rho(t, u)$ , jolle  $\rho(t, u) \rightarrow 0$  kun  $u \rightarrow 0$ , toteuttaa  $\|\rho(t, u)\| \leq \|\rho(t_0, u)\|$  kaikilla  $t \in [t_0, t_1]$ . Näin on Taylorin kaavan nojalla esimerkiksi silloin, kun  $L$  on  $C^2$ -funktio.

<sup>12</sup>Todistus perustuu rajankäynnin ja integroinnin järjestyksen vaihtamiseen. Lebesguen dom-  
minoidun konvergenssin lauseen avulla oletuksia voidaan lieventää.

<sup>13</sup>Hieman vähemmän säännöllisyyttä riittäisi, etenkin ensimmäisen muuttujan suhteen.

missä  $a, b \in \mathbb{R}$ . Jos  $y \in V$  on funktionaalin  $J: V \rightarrow \mathbb{R}$  ääriarvo, niin

$$\frac{d}{dt} \frac{\partial L}{\partial y'}(t, y(t), y'(t)) - \frac{\partial L}{\partial y}(t, y(t), y'(t)) = 0 \text{ kaikilla } t \in [t_0, t_1]. \quad (5.5)$$

*Todistus.* Kuvaus  $L$  toteuttaa  $C^2$ -funktiona apulauseen 5.2.1 oletukset. Koska myös  $y$  on  $C^2$ -funktio, niin apulauseen 5.2.2 oletukset toteutuvat, ja niin

$$\int_{t_0}^{t_1} h(t) \left( \frac{\partial L}{\partial y}(t, y(t), y'(t)) - \frac{d}{dt} \frac{\partial L}{\partial y'}(t, y(t), y'(t)) \right) dt = 0$$

kaikilla  $h \in C_0^\infty([t_0, t_1])$ , mistä variaatiolemman nojalla seuraa (5.5). **m.o.t.**

Eulerin-Lagrangen yhtälöiden yleistäminen vektoriarvoisten kuvausten tapaukseen käy suoraviivaisesti:

**Lause 5.2.2** (*Eulerin-Lagrangen yhtälöt*). Olkoon  $L: [t_0, t_1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $m = 1, 2, \dots$ ,  $C^2$ -funktio ja olkoon

$$V = \{y \in C^2([t_0, t_1])^m \mid y(t_0) = a, y(t_1) = b\},$$

missä  $a, b \in \mathbb{R}^m$ . Jos  $y \in V$  on funktionaalin  $J: V \rightarrow \mathbb{R}$  ääriarvo, niin

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial y'_i}(t, y(t), y'(t)) - \frac{\partial L}{\partial y_i}(t, y(t), y'(t)) &= 0 \\ \text{kaikilla } i = 1, 2, \dots, m \text{ ja } t \in [t_0, t_1]. \end{aligned} \quad (5.6)$$

Tämä muoto sopii juuri Hamiltonin pienimmän vaikutuksen periaatteelle. Kun mekaanisen järjestelmän Eulerin-Lagrangen yhtälöt kirjoitetaan, ne paljastuvat Newtonin liikeyhtälöiksi. Klassisen mekaniikan esitystä, jossa liikeyhtälöt johdetaan Lagrangen funktiosta  $L(t, \mathbf{q}, \mathbf{q}')$  sanotaan *Lagrangen mekaniikaksi*. Lagrangen funktio muodostetaan lausekkeesta  $L = T - V$ , missä  $T$  on liike-energia ja  $L$  on potentiaalienergia. On paljon helpompaa löytää ja laskea yhteen skalaariarvoisia energioita kuin vektoriarvoisia voimia. Tässä on Lagrangen mekaniikan vahvuus. Heikkoutena on se, että Eulerin-Lagrangen yhtälöt ovat toisen kertaluvun differentiaaliyhtälöitä, joiden ratkaiseminen on jo siksi vaikeaa ja lisäksi järjestelmän omat symmetriat kätkeytyvät niihin: kullakin mekaanisella järjestelmällä on oma luonnollinen koordinaatistonsa. Liikeyhtälöt tulisi voida kirjoittaa siinä, jolloin ne ovat rakenteeltaan hyvin yksinkertaisia. Tähän antaa keinon *Hamiltonin mekaniikka*. Järjestelmän *Hamiltonin funktio* on  $H(\mathbf{q}, \mathbf{p}) = T + V$  ja sen käyttäytymistä kuvaavat *Hamiltonin yhtälöt*

$$\mathbf{p}' = -\frac{\partial H}{\partial \mathbf{q}}, \quad \mathbf{q}' = \frac{\partial H}{\partial \mathbf{p}}, \quad (5.7)$$

missä  $\mathbf{p}$  on kanoninen impulssi

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{q}'}$$

Hamiltonin liikeyhtälöiden ratkaisusta saadaan palaamalla muuttujiin  $(\mathbf{q}, \mathbf{q}')$  funktionaalin

$$\mathbf{q} \mapsto S(\mathbf{q}) = \int_{t_0}^{t_1} L(t, \mathbf{q}(t), \mathbf{q}'(t)) dt$$

ääriarvokohta.

Hamiltonin mekaniikkaa tutkitaan ja sovelletaan ahkerasti modernissa säätöteoriassa<sup>14</sup>.

<sup>14</sup>Hamiltonin-Jacobin yhtälö yms. Tärkeä innoituksen lähde on L. Landau & E. Lifschitz, *Mekanika*, Mir, Moskova 1988.

### 5.2.1 Eulerin-Lagrangen kenttäyhtälöt

Hamiltonin pienimmän vaikutuksen periaate ei rajoitu vain klassiseen mekaniikkaan, vaan esimerkiksi sähködynamiikan liikeyhtälöt eli Maxwellin yhtälöt saadaan tietyn funktionaalien ääriarvosta. Maxwellin yhtälöt ovat ensimmäisen kertaluvun osittaisdifferentiaaliyhtälöitä, ja yhdessä reuna- ja alkuehtojen kanssa ne määräävät sähkö- ja magneettikentän. Niinpä sanomme kuvausta

$$\mathcal{L}: [0, 1] \times [t_0, t_1] \times \mathbb{R}^3 \rightarrow \mathbb{R}, \quad t_0 < t_1,$$

Lagrangen tiheysfunktiksi ja haemme funktionaalien

$$\phi \mapsto J(\phi) = \int_{t_0}^{t_1} \int_0^1 \mathcal{L}(x, t, \phi(x, t), \phi_x(x, t), \phi_t(x, t)) \, dx \, dt, \quad (5.8)$$

ääriarvoja joukossa  $V$ , kun

$$V = \{\phi \in C^2([0, 1] \times [t_0, t_1]) \mid \phi(0, \cdot) = a, \phi(1, \cdot) = b, \\ \phi(\cdot, t_0) = c, \phi(\cdot, t_1) = d\}, \quad (5.9)$$

missä  $a, b \in C^1([t_0, t_1])$  ja  $c, d \in C^1([0, 1])$ . Yllä merkittiin osittaisderivaattoja  $\phi_t = \partial\phi/\partial t$  ja  $\phi_x = \partial\phi/\partial x$ . Huomaa, että funktion  $\phi$  arvot joukon  $[0, 1] \times [t_0, t_1]$  koko reunalla  $\{0, 1\} \times [t_0, t_1] \cup [0, 1] \times \{t_0, t_1\}$  on kiinnitetty eli ääriarvoa haetaan ”varioimalla” funktiota  $\phi$  funktioilla  $h \in C_0^\infty([0, 1] \times [t_0, t_1])$ . Samoin kuin lause 5.2.1 todistetaan seuraava tulos:

**Lause 5.2.3** (Eulerin-Lagrangen kenttäyhtälö). *Olko  $\mathcal{L}: [0, 1] \times [t_0, t_1] \times \mathbb{R}^3 \rightarrow \mathbb{R}$   $C_2$ -funktio ja olko  $V$  annettu kaavalla (5.9), missä  $a, b \in C^2([t_0, t_1])$  ja  $c, d \in C^2([0, 1])$ . Jos  $\phi \in V$  on kaavoilla (5.8)-(5.9) määritellyn funktionaalien ääriarvokohta, niin*

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \phi_t} - \frac{\partial \mathcal{L}}{\partial \phi} + \frac{\partial}{\partial x} \frac{\partial \mathcal{L}}{\partial \phi_x} = 0 \text{ kaikilla } (t, x) \in [t_0, t_1] \times [0, 1]. \quad (5.10)$$

Tarkastellaan esimerkkinä suurta kiinteistönomistajaa joka vuokraa samankokoisia liikehuoneistoja Kauppakadun varrelta. Liikehuoneistojen vuokra  $p$  riippuu ajasta ja sijainnista kadun varrella eli yksiulotteisesta paikkakoordinaatista  $x$ . Siis  $p = p(x, t)$ . Kiinteistönomistajan voitto aikana  $[t_0, t_1]$  on

$$\int_{t_0}^{t_1} \int_0^1 \Pi(p(x, t), p_x(x, t), p_t(x, t)) \, dx \, dt, \quad (5.11)$$

missä  $\Pi$  on hetkellisen paikallinen voitto eli Lagrangen tiheysfunktio,

$$\Pi(p, p_x, p_t) = -\frac{1}{2}p_t^2 + pq_d(p, p_x) - \frac{1}{2}p_x^2 - \frac{1}{2}. \quad (5.12)$$

Ensimmäinen ja kolmas termi esittävät vuokran muuttumisesta tulevia sopeutumiskustannuksia (vuokrankorotus voi johtaa vuokralaisen lähtemiseen, vuokranalennus puolestaan voi saada uusia vuokralaisia tulemaan). Toinen termi esittää vuokratuloja;  $q_d$  esittää kussakin osoitteessa olevien liikehuoneistojen kysyntää,

$$q_d(p, p_x) = 4 - p + p_x. \quad (5.13)$$

Tällöin vuokran kohoaminen vähentää kysyntää ja vuokran kasvaminen kadun suuntaan lisää kysyntää.

Jos  $p$  on kaavoilla (5.11)-(5.13) annetun funktionaalien ääriarvokohta, niin Eulerin-Lagrangen kenttäyhtälöiden avulla saadaan osittaisdifferentiaaliyhtälö

$$\frac{\partial^2 p}{\partial t^2}(x, t) + \frac{\partial^2 p}{\partial x^2}(x, t) - 2p(x, t) = 4 \text{ kaikilla } (x, t) \in [0, 1] \times [t_0, t_1].$$

Tämän reunaehtoina ovat kiinnitetyt vuokrat eli esimerkiksi

$$\begin{aligned} p(0, t) = 0, \quad p(1, t) = 0 \text{ kaikilla } t \in [t_0, t_1] \\ p(x, t_0) = 0, \quad p(x, t_1) = x(1 - x) \text{ kaikilla } x \in [0, 1]. \end{aligned}$$

Yksiulotteisen avaruusmuuttujan ja välin  $[0, 1]$  sijasta voimme tarkastella muuttujaa  $x \in \Omega$ , missä  $\Omega \subset \mathbb{R}^n$ ,  $n = 1, 2, \dots$ , on avoin joukko. Lagrangen tiheysfunktio voi riippua kaikista funktion  $\phi : \Omega \times [t_0, t_1]$  osittaisderivaatoista. Tässäkin tapauksessa voisimme muotoilla ja todistaa lauseen 5.2.3 vastineen.

Lopetamme tämän jakson tulokseen, joka takaa, että tutkittavalla funktio-naalilla on suurin arvo. Tuloksen todistuksen ydin on jo hahmoteltu ääriarvojen olemassaoloa käsittelevän jakson lopussa. Merkitään symbolilla  $L^2(G)$  niiden funktioiden  $G \rightarrow \mathbb{R}$  joukkoa, jotka ovat mitallisia ja joiden neliö on integroituva.

**Lause 5.2.4** *Olko  $\Omega \subset \mathbb{R}^n$  avoin,  $n = 1, 2, \dots$ ,  $t_0 < t_1$ ,  $M > 0$  ja  $g : \Omega \times [t_0, t_1] \times \mathbb{R}^3 \rightarrow \mathbb{R}$  jatkuva siten, että*

$$g(x, t, u, v, w) \leq M(1 + |u| + |v| + |w|)$$

*kaikilla  $(x, t, u, v, w) \in \Omega \times [t_0, t_1] \times \mathbb{R}^3$ . Määritellään  $J : L^2(\Omega \times ]t_0, t_1[) \rightarrow \mathbb{R} \cup \{-\infty\}$  lausekkeella*

$$\begin{aligned} J(\phi) = \int_{\Omega \times [t_0, t_1]} \Big( -\phi(x, t)^2 - \phi_x(x, t)^2 - \phi_t(x, t)^2 + \\ + g(x, t, \phi(x, t), \phi_x(x, t), \phi_t(x, t)) \Big) d(x, t), \end{aligned}$$

*jos tämä integraali on olemassa. Jos ei ole, niin asetetaan  $J(\phi) = -\infty$ . Nyt funktionaalilla  $J$  on suurin arvo.*

### 5.3 Lagrangen kertojat

Haetaan funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $n = 1, 2, \dots$ , ääriarvoja joukossa

$$\{x \in \mathbb{R}^n \mid s_1(x) = 0, s_2(x) = 0, \dots, s_m(x) = 0\},$$

missä kuvaukset  $s_1, \dots, s_m : \mathbb{R}^n \rightarrow \mathbb{R}$  lausuvat side-ehdot. Tämän *sidotun ääriarvotehtävän* ratkaisemiseen on kaksi menetelmää. Ensimmäinen niistä on *si-joitusmenetelmä*, jossa ratkaistaan side-ehdoista osa muuttujista muiden muuttujien funktiona ja ongelma palautetaan alempiulotteiseksi tavalliseksi ääriarvotehtäväksi. Toinen menetelmä on *Lagrangen kertojien menetelmä*. Se perustuu siihen, että toisaalta kuvaus  $f$  kasvaa nopeiten gradienttinsa suuntaan ja toisaalta side-ehto  $s_i(x) = 0$  määrää tasa-arvopinnan, jonka normaali pisteessä  $x$  on  $\nabla s_i(x)$ . Ääriarvokohdassa nämä kaksi vektoria ovat yhdensuuntaisia, joten on olemassa luvut  $\lambda_i \in \mathbb{R}$  siten, että

$$\nabla f(x) + \lambda_i \nabla s_i(x) = 0 \text{ kaikilla } i = 1, 2, \dots, m. \quad (5.14)$$

Lukuja  $\lambda_1, \lambda_2, \dots, \lambda_m$  sanotaan *Lagrangen kertojiksi*.

Tämä Lagrangen kertojien menetelmä yleistyy myös ääretönulotteiseen tapaukseen. Esitämme seuraavan tuloksen ilman todistusta; sen geometrinen perustelu on sama kuin äärellisulotteisessa tapauksessa.

**Lause 5.3.1** *Olko  $L, S_1, \dots, S_m : [t_0, t_1] \times \mathbb{R}^2$ ,  $t_0 < t_1$ ,  $C^2$ -funktioita. Jos  $y$  on funktionaalin*

$$z \mapsto \int_{t_0}^{t_1} L(t, z(t), z'(t)) dt$$

ääriarvokohta joukossa

$$V = \{w \in C^2([t_0, t_1]) \mid w(t_0) = a, w(t_1) = b, \int_{t_0}^{t_1} S_i(t, y(t), y'(t)) dt = 0 \text{ kaikilla } i = 1, \dots, m\}, \quad (5.15)$$

missä  $a, b \in \mathbb{R}$ , niin on olemassa luvut  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  siten, että

$$\frac{d}{dt} \frac{\partial(L + \lambda_i S_i)}{\partial y'}(t, y(t), y'(t)) - \frac{\partial(L + \lambda_i S_i)}{\partial y}(t, y(t), y'(t)) = 0 \quad (5.16)$$

kaikilla  $t \in [t_0, t_1]$ ,  $i = 1, \dots, m$ .

## 5.4 Harjoitustehtäviä

1. Osoita Weierstraßin lauseen yleisten versioiden avulla, että kuvaus  $f: \mathbb{R} \rightarrow \mathbb{R}$  saavuttaa pienimmän arvonsa ja kuvaus  $g: \mathbb{R} \rightarrow \mathbb{R}$  saavuttaa suurimman arvonsa, kun

$$f(x) = \begin{cases} \frac{\sin x + x^3}{x}, & \text{jos } x \neq 0, \\ 0, & \text{jos } x = 0, \end{cases} \quad g(x) = \begin{cases} -2, & \text{jos } |x| \geq 1, \\ 1 - x^2, & \text{jos } |x| < 1. \end{cases}$$

2. Laske määritelmän avulla funktionaalin

$$y \mapsto F(y) := \int_0^1 \left( y(t)^2 + y'(t)^2 (1 + \sin y(t)) \right) dt, \quad C^2([0, 1]) \rightarrow \mathbb{R},$$

suunnattu derivaatta  $D_h F(y)$  jokaiseen suuntaan  $h \in C^2([0, 1])$ .

3. Hae funktionaalin

$$y \mapsto J(y) := -\frac{1}{2} \int_0^1 (y(t)^2 + y'(t)^2) dt$$

mahdollinen suurin arvo joukossa

$$V = \{z \in C^2([0, 1]) \mid z(0) = 0, z(1) = 0\}$$

- (a) Eulerin-Lagrangen yhtälöiden avulla
- (b) ilman niitä.

4. Onko seuraava päättely oikein? Haetaan funktionaalin

$$y \mapsto J(y) := -\frac{1}{2} \int_0^1 (y(t)^2 + y'(t)^2) dt$$

suurinta arvoa joukossa

$$V = \{z \in C^2([0, 1]) \mid z(0) = 0, z'(0) = 1\}.$$

Olkoon  $J(y)$ ,  $y \in V$ , funktionaalin  $J$  suurin arvo joukossa  $V$ . Suunnattu derivaatta  $D_h J(y)$  on olemassa kaikkiin suuntiin  $h \in C_0^\infty([0, 1])$  ja  $y + \epsilon h \in V$  kaikilla  $\epsilon \in \mathbb{R}$ , joten

$$D_h J(y) = 0 \text{ kaikilla } h \in C_0^\infty([0, 1]).$$

Koska kuvaus

$$t \mapsto \frac{d}{dt} \frac{\partial}{\partial z'} \frac{1}{2} (z^2 + z'^2) = z''(t)$$

on jatkuva kaikilla  $z \in V$ , niin integroimalla osittain saadaan

$$\int_0^1 h(t) \left( \frac{d}{dt} \frac{\partial}{\partial y'} \frac{1}{2} (y^2 + y'^2) - \frac{\partial}{\partial y} \frac{1}{2} (y^2 + y'^2) \right) dt = 0$$

kaikilla  $h \in C_0^\infty(]0, 1[)$ , mistä variaatiolemman nojalla seuraa, että

$$y''(t) - y(t) = 0 \text{ kaikilla } t \in [0, 1].$$

Tämän differentiaaliyhtälön yleinen ratkaisu on  $y(t) = A \cosh t + B \sinh t$ , missä vakiot  $A$  ja  $B$  määräytyvät alkuehdoista  $y(0) = 0$ ,  $y'(0) = 1$ . Siis  $y(t) = \sinh t$ .

5. Jatketaan edellisen tehtävän tarkastelua. Laske  $J(\sinh t)$  ja  $J(t)$ . Voiko edellisen tehtävän päättely olla oikein? Missä vika? Osoita, että

$$J(z) < 0 \text{ kaikilla } z \in V.$$

Muodosta jono  $(z_n)$ ,  $z_n \in V$  siten, että  $J(z_n) \rightarrow 0-$ , kun  $n \rightarrow \infty$ .

6. Johdetaan *ketjuviivan* yhtälö eli tarkastellaan narua, joka on päistään kiinnitetty pisteisiin  $(-1, a)$  ja  $(1, a)$ ,  $a > 0$ . Narun pituus olkoon  $l > 2$ . Narua esitetään  $C^2$ -kuvauksella  $f: [-1, 1] \rightarrow \mathbb{R}$ ;  $(x, f(x))$  on narun sijainti. Naru asettuu siten, että sen potentiaalienergia on pienin mahdollinen eli funktionaali

$$f \mapsto \int_0^l f ds = \int_{-1}^1 f(x) \sqrt{1 + f'(x)^2} dx$$

saavuttaa pienimmän arvonsa joukossa

$$V = \{w \in C^2([-1, 1]) \mid w(-1) = w(1) = a, \int_{-1}^1 \sqrt{1 + w'(x)^2} dx = l\}.$$

Kirjoita ja ratkaise yhtälö (5.16). Voit käyttää hyväksi symmetriaa  $y$ -akselin suhteen. Muista, että  $D \operatorname{arcosh} x = 1/\sqrt{x^2 - 1}$ .

# Luku 6

## $\epsilon$ -optimaalisuus

Edellisessä luvussa<sup>1</sup> löysimme funktionaalin

$$J: \{y \in C^2([0, 1]) \mid y(0) = 0, y'(0) = 1\} \rightarrow \mathbb{R}, \quad J(y) = - \int_0^1 (y^2(t) + y'(t)^2) dt,$$

jolla ei ole suurinta arvoa. Vielä yksinkertaisempi on seuraava kuvaus:

$$C: \mathbb{R} \rightarrow \mathbb{R}, \quad C(t) = \begin{cases} -t, & \text{jos } t < 0, \\ 1 & \text{muulloin.} \end{cases}$$

Sillä ei pienintä arvoa ja silti siihen liittyy taloudellisesti relevantti ääriarvotekniikka:  $C(t)$  esittää kustannuksia, jotka aiheutuvat, kun hetkellä  $t = 0$  lähtevälle junalle saavutaan hetkellä  $t \in \mathbb{R}$ .

**Määritelmä 6.0.1** Merkitään  $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ . Olkoon  $X$  epätyhjä joukko,  $F: X \rightarrow \overline{\mathbb{R}}$  ja  $\epsilon > 0$ . Piste  $x_0 \in X$  on kuvauksen  $F$   $\epsilon$ -maksimikohta ja  $F(x_0)$   $\epsilon$ -maksimi, jos

$$F(x_0) \geq \sup\{F(x) \mid x \in X\} - \epsilon.$$

Piste  $y_0 \in X$  on kuvauksen  $F$   $\epsilon$ -minimikohta ja  $F(y_0)$   $\epsilon$ -minimi, jos

$$F(y_0) \leq \inf\{F(y) \mid y \in X\} + \epsilon.$$

Heti saamme seuraavan huomion.

**Lause 6.0.2** Olkoon  $X \neq \emptyset$  ja  $F: X \rightarrow \overline{\mathbb{R}}$ . Jos  $x_0 \in X$  on kuvauksen  $F$  minimikohta, niin se on myös  $\epsilon$ -minimikohta kaikilla  $\epsilon > 0$ . Jos  $x_0 \in X$  on kuvauksen  $F$  maksimikohta, niin se on myös  $\epsilon$ -maksimikohta kaikilla  $\epsilon > 0$ .

**Lause 6.0.3** (Ekeland). Olkoon  $(X, d)$  täydellinen metrinen avaruus ja  $f: X \rightarrow [0, \infty]$  alhaalta puolijatkuva. Olkoon  $\epsilon > 0$  ja  $x_0 \in X$  siten, että  $F(x_0) < \infty$ . Tällöin on olemassa ainakin yksi piste  $\bar{x} \in X$  siten, että

$$f(\bar{x}) + \epsilon d(x_0, \bar{x}) \leq f(x_0), \quad (6.1)$$

$$f(\bar{x}) < f(x) + \epsilon d(x, \bar{x}) \text{ kaikilla } x \in X \setminus \{x_0\}. \quad (6.2)$$

*Todistus.* Voidaan olettaa, että  $\epsilon = 1$ . Asetetaan kuvausta  $f$  vastaamaan moniarvoinen funktio  $F$  kaavalla

$$F(x) := \{y \in X \mid f(y) + d(x, y) \leq f(x)\}, \quad x \in X. \quad (6.3)$$

Tällöin siis  $F$  kuvaa pisteen  $x \in X$  joksikin joukon  $X$  osajoukoksi ja sillä seuraavat ominaisuudet:

<sup>1</sup>Tämän luvun aiheista ks. tarkemmin Jean-Pierre Aubin, "L'analyse non linéaire et ses motivations économiques". Masson, Pariisi, 1984, venäjäksi Mir, Moskova, 1988, luku 1, ja I. Ekeland & R. Temam, "Analyse convexe et problèmes variationnels", Dunod, Gauthier-Villars, Pariisi, 1974, luku 1.

- (i)  $F(x)$  on suljettu jokaisella  $x \in X$ .
- (ii)  $y \in F(y)$  jokaisella  $y \in X$  (refleksiivisyys).
- (iii) Jos  $y \in F(x)$ , niin  $F(y) \subset F(x)$  (transitiivisuus).

Ensimmäinen näistä seuraa kuvauksen  $f$  alhaalta puolijatkuvuudesta. Kolmas toteutuu selvästi jos  $f(x) = \infty$ . Olkoon siis  $x \in X$  siten, että  $f(x) < \infty$ . Olkoot  $y \in F(x)$  ja  $z \in F(y)$ . Silloin

$$f(z) + d(y, z) \leq f(y) \text{ ja } f(y) + d(x, y) \leq f(x),$$

joista kolmioepäyhtälön nojalla seuraa, että  $f(z) + d(x, z) \leq f(x)$  eli  $z \in F(x)$ . Siten  $F$  on transitiivinen.

Määritellään kuvaus  $v: \{x \in X \mid f(x) < \infty\} \rightarrow \mathbb{R}$  kaavalla

$$v(y) := \inf\{f(z) \mid z \in F(y)\}.$$

Selvästi

$$d(x, y) \leq f(x) - v(x) \text{ kaikilla } y \in F(x),$$

mistä seuraa arvio joukon  $F(x)$  halkaisijalle:

$$\text{diam}(F(x)) \leq 2(f(x) - v(x)).$$

Määritellään induktiivisesti jono  $(x_n)$ : valitaan  $x_{n+1} \in F(x_n)$  siten, että

$$f(x_{n+1}) \leq v(x_n) + 2^{-n};$$

tämä on mahdollista infimumin määritelmän perusteella. Koska  $F$  on transitiivinen, niin  $F(x_{n+1}) \subset F(x_n)$  ja siten

$$v(x_n) \leq v(x_{n+1}). \text{ kaikilla } n \in \mathbb{N}.$$

Toisaalta aina pätee  $v(y) \leq f(y)$ , joten

$$v(x_{n+1}) \leq f(x_{n+1}) \leq v(x_n) + 2^{-n} \leq v(x_{n+1}) + 2^{-n}.$$

Tästä seuraa, että

$$0 \leq f(x_{n+1}) - v(x_{n+1}) \leq 2^{-n}.$$

Joukkojono  $(F(x_n))$  on vähenevä, joten koska  $X$  on täydellinen, niin on olemassa  $\bar{x} \in X$  siten, että

$$\bigcap_{n \in \mathbb{N}} F(x_n) = \{\bar{x}\}.$$

Koska  $\bar{x} \in F(x_0)$ , niin (6.1) pätee. Toisaalta  $\bar{x} \in F(x_n)$  kaikilla  $n \in \mathbb{N}$ , joten  $F(\bar{x}) \subset F(x_n)$  kaikilla  $n \in \mathbb{N}$ . Siten

$$F(\bar{x}) = \{\bar{x}\}.$$

Jos  $x \neq \bar{x}$ , niin  $x \notin F(\bar{x})$ . Siten  $f(x) + d(x, \bar{x}) > f(\bar{x})$ . Tämä on (6.2). **m.o.t.**

**Seuraus 6.0.1** *Olkoon  $(X, d)$  täydellinen metrinen avaruus ja  $f: X \rightarrow [0, \infty]$  alhaalta puolijatkua. Olkoot  $\epsilon, \lambda > 0$  ja  $x_0 \in X$  siten, että  $f(x_0)$  on äärellinen ja  $\epsilon\lambda$ -minimi. Tällöin on olemassa piste  $\bar{x} \in X$  siten, että*

$$f(\bar{x}) \leq f(x_0), \tag{6.4}$$

$$d(x_0, \bar{x}) \leq \lambda, \tag{6.5}$$

$$f(\bar{x}) \leq f(x) + \epsilon d(x, \bar{x}) \text{ kaikilla } x \in X. \tag{6.6}$$

*Todistus.*<sup>2</sup> Ekelandin lauseen nojalla on olemassa  $\bar{x} \in X$  siten, että

$$f(\bar{x}) + \epsilon d(x_0, \bar{x}) \leq f(x_0).$$

Siten (6.4) pätee. Toisaalta

$$f(x_0) \leq \inf\{f(x) \mid x \in X\} + \epsilon\lambda \leq f(\bar{x}) + \epsilon\lambda,$$

joten  $d(x_0, \bar{x}) \leq \lambda$ . Epäyhtälöstä (6.2) seuraa (6.6), jos  $x \neq x_0$ . Jos taas  $x = x_0$ , niin (6.6) seuraa epäyhtälöstä (6.4). **m.o.t.**

**Seuraus 6.0.2** *Olkoon  $(V, \|\cdot\|)$  täydellinen normiavaruus,  $A \subset V$  suljettu ja  $f: A \rightarrow [0, \infty]$  alhaalta puolijatkuva. Olkoot  $\epsilon > 0$  ja  $x_0 \in A$  siten, että  $f(x_0)$  on äärellinen ja  $\epsilon$ -minimi sekä suunnattu derivaatta  $D_h f(x_0)$  on olemassa suuntaan  $h \in V$ . Tällöin on olemassa piste  $\bar{x} \in A$  siten, että*

$$f(\bar{x}) \leq f(x_0), \quad (6.7)$$

$$\|x_0 - \bar{x}\| \leq \sqrt{\epsilon}, \quad (6.8)$$

$$|D_h f(\bar{x})| \leq \sqrt{\epsilon}\|h\|. \quad (6.9)$$

*Todistus.* Valitaan seurauksessa 6.0.1  $\lambda = \sqrt{\epsilon}$  ja  $\epsilon := \sqrt{\epsilon}$ , jolloin saadaan (6.7) ja (6.8). Todistetaan (6.9). Olkoon  $\eta < 0$  ja  $h \in V$ . Epäyhtälön (6.6) nojalla tällöin

$$f(\bar{x} + \eta h) - f(\bar{x}) \geq -\sqrt{\epsilon}\|\eta h\|,$$

mistä seuraa, että

$$D_h f(\bar{x}) \leq \sqrt{\epsilon}\|h\|.$$

Jos taas yllä valitaan  $\eta > 0$ , niin saadaan samoin

$$D_h f(\bar{x}) \geq -\sqrt{\epsilon}\|h\|.$$

**m.o.t.**

**Seuraus 6.0.3** *Olkoon  $(V, \|\cdot\|)$  täydellinen normiavaruus,  $A \subset V$  suljettu ja  $f: A \rightarrow [0, \infty]$  alhaalta puolijatkuva. Olkoot  $\epsilon > 0$  ja  $x_0 \in A$  siten, että  $f(x_0)$  on äärellinen ja  $\epsilon$ -minimi sekä suunnattu derivaatta  $D_h f(x_0)$  on olemassa suuntaan  $h \in V$ . Tällöin on olemassa piste  $\bar{x} \in A$  siten, että*

$$f(\bar{x}) \leq f(x_0), \quad (6.10)$$

$$\|x_0 - \bar{x}\| \leq 1, \quad (6.11)$$

$$|D_h f(\bar{x})| \leq \epsilon\|h\|. \quad (6.12)$$

*Todistus.* Valitaan seurauksessa 6.0.1  $\lambda = 1$ , jolloin saadaan (6.10) ja (6.11). Todistetaan (6.12). Olkoon  $\eta < 0$  ja  $h \in V$ . Epäyhtälön (6.6) nojalla tällöin

$$f(\bar{x} + \eta h) - f(\bar{x}) \geq -\epsilon\|\eta h\|,$$

mistä seuraa, että

$$D_h f(\bar{x}) \leq \epsilon\|h\|.$$

Jos taas yllä valitaan  $\eta > 0$ , niin saadaan samoin

$$D_h f(\bar{x}) \geq -\epsilon\|h\|.$$

**m.o.t.**

---

<sup>2</sup>Äärellisulotteisissa tapauksissa tämä tulos voidaan todistaa suoraan soveltamalla apufunktion  $f_\delta(x) = f(x) + \delta\|x - x_0\|$ ,  $\delta > 0$ , ehtoa  $\nabla f_\delta(x) = 0$ .

*Esimerkki.* Olkoon

$$V = \{y \in C^1([0, 1]) \mid y(0) = 0\}, \quad \|y\|_V = \sup_{0 \leq t \leq 1} |y'(t)|.$$

Nyt  $(V, \|\cdot\|_V)$  on täydellinen normiavaruus. Olkoon edelleen

$$A = \{y \in V \mid y'(0) = 1\}, \quad J: V \rightarrow \mathbb{R}, \quad J(y) = \frac{1}{2} \int_0^1 (y(t)^2 + y'(t)^2) dt.$$

Tällöin  $A$  on suljettu ja  $J: A \rightarrow \mathbb{R}$  on jatkuva. Suunnattu derivaatta on olemassa kaikkiin suuntiin  $h \in V$  ja

$$D_h J(y) = \int_0^1 (y(t)h(t) + y'(t)h'(t)) dt,$$

joten  $J'(y): V \rightarrow \mathbb{R}, h \mapsto J'(y)h := D_h J(y)$  on lineaarikuvaus eli  $J'(y)$  on kuvauksen  $J$  *Gâteaux-derivaatta* pisteessä  $y \in A$ . Jos  $y \in A$  on  $\epsilon$ -minimi, niin seurauksen 6.0.3 nojalla on olemassa toinen  $\epsilon$ -minimi  $\bar{y} \in A$  siten, että  $\|y - \bar{y}\|_V \leq 1$  ja

$$\|J'(\bar{y})\| := \sup \left\{ \frac{D_h J(\bar{y})}{\|h\|_V} \mid h \in V, h \neq 0 \right\} \leq \epsilon.$$

Toisaalta osittainintegroinnilla saadaan, että

$$D_h J(\bar{y}) = \int_0^1 (-z(t) + z''(t))h'(t) dt \quad \text{kaikilla } h \in C_0^\infty(]0, 1[),$$

missä  $z(t) = 1 + \int_0^t \bar{y}(s) ds, \quad t \in [0, 1]$ . Koska  $C_0^\infty(]0, 1[) \subset V$  tiheästi,

$$\int_0^1 |z''(t) - z(t)| dt = \sup \left\{ \frac{\int_0^1 (z''(t) - z(t))h'(t) dt}{\|h\|_V} \mid h \in C_0^\infty(]0, 1[), h \neq 0 \right\}.$$

Täten

$$\int_0^1 |z''(t) - z(t)| dt \leq \epsilon. \quad (6.13)$$

Huomaa, että tämä on *vain* välttämätön ehto sille, että  $\bar{y}$  on  $\epsilon$ -minimi. Tämän hän toteuttaa myös se  $z$ , jolle  $z'' - z = 0$  ja joka itse asiassa on  $z(t) = \cosh t$ , ja  $J(z')$  ei ole funktionaalin  $J$  minimi eikä  $\epsilon$ -minimi, jos  $\epsilon$  on tarpeeksi pieni (ks. edellisen luvun harjoitustehtävät 4-5).

## 6.1 Harjoitustehtäviä

1. Hae kuvauksen

$$C: \mathbb{R} \rightarrow \mathbb{R}, \quad C(t) = \begin{cases} -t, & \text{jos } t < 0, \\ 1 & \text{muulloin,} \end{cases}$$

kaikki  $0, 1$ -minimit.

2. Tee sama kuvaukselle

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{1 + x^2},$$

(a) suoraan määritelmän avulla,

(b) seurauksen 6.0.3 avulla (vihje:  $f'(x) = D_h f(x)$ , kun  $h = 1$ ).

3. Hae kaikki kuvauksen

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = \frac{1 + x^2}{1 + x^2 + y^2} + x^2 + 1,$$

0, 1-minimit. Vihje:  $|D_h f(x, y)| \leq \epsilon \|h\|$  kaikilla  $h \in \mathbb{R}^2$ , jos ja vain jos  $\|\nabla f(x, y)\| \leq \epsilon$ .

4. Osoita, että edellisessä esimerkissä joukko  $A$  on suljettu eli jos on olemassa jono  $(y_n)$  joukon  $A$  alkioita ja  $y \in V$  siten, että  $\|y_n - y\|_V \rightarrow 0$ , kun  $n \rightarrow \infty$ , niin  $y \in A$ .

5. Olkoon  $\epsilon > 0$ . Ratkaise seuraava differentiaaliepäyhtälö

$$|u'(t) - u(t)| \leq \epsilon \text{ kaikilla } t \in [0, 1]$$

alkuehdolla  $u(0) = 0$ . Vihje: kokeile sijoitusta  $u(t) = e^t v(t)$ . Miten löytäisit yhden perheen epäyhtälön (6.13) ratkaisuja?

6. Miksi epäyhtälöllä (6.13) yhdessä alkuehtojen  $z(0) = z''(0) = 1$ ,  $z'(0) = 0$  kanssa on muitakin ratkaisuja? Hae jokin sellainen.



# Luku 7

## Diskreettiä optimointia

### 7.1 Bellmanin optimaalisuusperiaate

Tarkastellaan epätyhjää joukkoa  $X$  paikkoja, joita yhdistää verkko teitä, rautateitä, laiva- ja lentoreittejä. Merkitään järjestettyä monikkoa  $(x_0, x_1, \dots, x_n) \in X^{n+1}$ ,  $n \in \mathbb{N}$ , lyhyesti  $x_0, x_1, \dots, x_n$ , ja sanotaan sitä *poluksi*, joka alkaa pisteestä  $x_0$  ja joka kulkee pisteiden  $x_1, \dots, x_{n-1}$  kautta pisteeseen  $x_n$ . Olkoot kuvaukset

$$c, K: X \times X \rightarrow ]-\infty, \infty].$$

Luku  $c(x, y)$  esittää matkan pisteestä  $x \in X$  pisteeseen  $y \in X$  aiheuttamia kustannuksia. Jos  $c(x, y) < 0$ , niin matka tuottaa enemmän kuin maksaa (maisemien katselu antaa mielihyvää, matkalla kerätään marjoja tai otetaan talteen yliajetut rusakot). Jos pisteestä  $x$  ei pääse pisteeseen  $y$ , asetetaan  $c(x, y) = \infty$ . Aika on tässä mallissa diskreetti ja parametri  $n$  viittaa ajanhetkeen. Tällöin  $c(x, y) = \infty$  pitää sisällään sen, että yhdellä ajan askeleella ei ehditä matkustaa paikasta  $x$  paikkaan  $y$ . Paikallaan olemisesta ei tule kustannuksia, jos

$$c(x, x) = 0, \quad K(x, x) = 0 \text{ kaikilla } x \in X. \quad (7.1)$$

Matkan kokonaiskustannuksia pitkin polkua  $x_0, x_1, \dots, x_n$ ,  $n = 1, 2, \dots$ , esittää

$$C(x_0, x_1, \dots, x_n; y) = \sum_{i=1}^n c(x_{i-1}, x_i) + K(x_n, y), \quad (7.2)$$

missä  $K(x, y)$  esittää sakkoa, joka tulee siitä, että paikkaan  $y \in X$  ei saavuttu. Asetetaan

$$C(x_0, y) = K(x_0, y). \quad (7.3)$$

Jos  $K(x, y) \equiv 0$ , ei ole väliä, mihin pisteeseen matka päättyy. Jos

$$K(x, y) = \begin{cases} \infty, & \text{jos } x \neq y, \\ 0, & \text{jos } x = y, \end{cases} \quad (7.4)$$

niin kokonaiskustannus (7.2) on äärellinen vain jos  $x_n = y$  eli sakko (7.4) pakottaa saapumaan pisteeseen  $y \in X$ . Toisaalta (7.1) sallii saapua pisteeseen  $y$  vähemmällä kuin  $n$ :llä askeleella ilman että kokonaiskustannukset siitä kasvavat. Luonteva optimointitehtävä on siis seuraava:

( $\mathbf{P}_\infty$ ) Olkoot  $x, y \in X$  annettuja. Hae sellaiset  $n \in \mathbb{N}$  ja  $(x_0, x_1, \dots, x_n) \in X^{n+1}$ , joille

$$C(x_0, x_1, \dots, x_n; y) = \inf \{ C(z_0, z_1, \dots, z_m; y) \mid m \in \mathbb{N}, (z_0, z_1, \dots, z_m) \in X^{m+1}, z_0 = x \}. \quad (7.5)$$

Tässä optimointitehtävässä ajan askelten määrää ei ole kiinnitetty. Jos se kiinnitetään, niin saadaan seuraava tehtävä:

( $\mathbf{P}_n$ ) Olkoot  $n \in \mathbb{N}$  ja  $x, y \in X$  annettuja. Hae sellainen  $(x_0, x_1, \dots, x_n) \in X^{n+1}$ , jolle

$$C(x_0, x_1, \dots, x_n; y) = \inf \{ C(z_0, z_1, \dots, z_n; y) \mid (z_0, z_1, \dots, z_n) \in X^{n+1}, z_0 = x \}. \quad (7.6)$$

Huomaamme, että jos joukko  $X$  on äärellinen, niin infimumia haetaan äärellisen monen reaaliluvun ja luvun  $\infty$  joukosta ja siten se saavutetaan. Tehtävällä ( $\mathbf{P}_n$ ) on tällöin ratkaisu.

Määritellään kuvaukset  $F_n, F_\infty: X \times X \rightarrow [-\infty, \infty]$ ,  $n \in \mathbb{N}$ ,

$$\begin{aligned} F_n(x, y) &= \inf \{ C(z_0, z_1, \dots, z_n; y) \mid z_0, z_1, \dots, z_n \in X^{n+1}, z_0 = x \}, \\ F_\infty(x, y) &= \inf \{ C(z_0, z_1, \dots, z_m; y) \mid m \in \mathbb{N}, z_0, z_1, \dots, z_m \in X^{m+1}, z_0 = x \}, \end{aligned}$$

**Lause 7.1.1** *Olkoot  $x, y \in X$  ja  $c(x, x) \leq 0$ . Tällöin lukujono  $(F_n(x, y))$  on vähenevä ja*

$$F_\infty(x, y) = \inf \{ F_n(x, y) \mid n \in \mathbb{N} \}.$$

*Todistus.* Harjoitustehtävä.

**m.o.t.**

Seuraava lause on *Bellmanin optimaalisuusperiaate*.

**Lause 7.1.2** *Olkoot  $x, y \in X$  ja  $n \in \mathbb{N}$ . Tällöin pätee*

$$F_0(x, y) = K(x, y), \quad (7.7)$$

$$F_m(x, y) = \inf \{ c(x, x') + F_{m-1}(x', y) \mid x' \in X \}, \quad m = 1, 2, \dots, n. \quad (7.8)$$

*Todistus.* Ensimmäinen yhtälö seuraa suoraan  $C(x_0)$ :n määritelmästä. Kaksi infimumia saadaan ottaa missä järjestyksessä hyvänsä tai yhtäaikaan. Siksi

$$\begin{aligned} F_m(x, y) &= \inf \left\{ \sum_{j=1}^m c(z_{j-1}, z_j) + K(z_m, y) \mid z_1, \dots, z_m \in X^m, z_0 = x \right\} \\ &= \inf_{z_1 \in X} \inf \left\{ c(x, z_1) + \sum_{j=2}^m c(z_{j-1}, z_j) + K(z_m, y) \mid z_2, \dots, z_m \in X^{m-1} \right\} \\ &= \inf_{z_1 \in X} \{ c(x, z_1) + F_{m-1}(z_1, y) \}. \end{aligned}$$

**m.o.t.**

Huomaa, että Bellmanin optimaalisuusperiaate mahdollistaa optimointitehtävän *rekursiivisen* ratkaisemisen.

## 7.2 Optimaalisuusyhtälö ja säätöteoriaa

Tarkastelemme tässä jaksossa diskreettejä *säätöteoreettisia* optimointiongelmia, jollainen on esimerkiksi talon lämmityksen säätäminen niin, että lämpötilajakauma talon sisällä on mahdollisimman lähellä haluttua ja lämmityskustannukset jäävät samalla mahdollisimman pieniksi. Tällöin tunnistetaan kaksi eri muuttujaa: säätö (lämmittimen teho) ja tila (lämpötila). Edellisen arvoa ajanhetkellä  $t$  merkitään tavallisesti  $u(t)$  ja jälkimmäisen  $x(t)$ . Aika on yhä diskreetti, joten merkitsemme lyhyesti  $u_n$  ja  $x_n$  pidempien ilmaisujen  $u(t_n)$  ja  $x(t_n)$  sijasta.

Yleistetään hieman Bellmanin optimaalisuusperiaatetta. Olkoot epätyhjät joukot  $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_{n-1}$ ,  $n \in \mathbb{N}$ , sekä

$$G: \mathcal{U}_0 \times \mathcal{U}_1 \times \dots \times \mathcal{U}_{n-1} \rightarrow [-\infty, \infty]. \quad (7.9)$$

Määritellään kuvaukset

$$\begin{aligned} G_j: \mathcal{U}_0 \times \mathcal{U}_1 \times \dots \times \mathcal{U}_{j-1} &\rightarrow [-\infty, \infty], \quad G_j(u_0, u_1, \dots, u_{j-1}) \\ &= \inf \{ G(u_0, u_1, \dots, u_{n-1}) \mid u_k \in \mathcal{U}_k, k = j, j+1, \dots, n-1 \}. \end{aligned}$$

**Lause 7.2.1** (*Optimaalisuusyhtälö*). Jokaiselle  $j = 0, 1, \dots, n-1$  on voimassa

$$G_j(u_0, u_1, \dots, u_{j-1}) = \inf_{u_j \in \mathcal{U}_j} G_{j+1}(u_0, u_1, \dots, u_j), \quad (7.10)$$

$$G_n(u_0, u_1, \dots, u_{n-1}) = G(u_0, u_1, \dots, u_{n-1}). \quad (7.11)$$

*Todistus.* Tämä todistetaan samoin kuin lause 7.1.2.

**m.o.t.**

Olkoon  $n \in \mathbb{N}$  ja epätyhjät joukot  $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_{n-1}$ . Sanomme kuvausta  $j \mapsto u_j \in \mathcal{U}_j$ ,  $j \in \{0, 1, \dots, n-1\}$  *säätömuuttujaksi*. Olkoon  $\mathcal{X}$  epätyhjä joukko; sanomme sitä *tila-avaruudeksi*. Kuvausta  $j \mapsto \omega_j$ ,  $j \in \{0, 1, \dots, n\}$ , sanotaan säätömuuttujaan  $u$  liittyväksi *tilamuuttujaksi*, jos

- (i) optimaalinen  $u_j$  riippuu vain olioista  $\omega_j$  ja  $j$  kaikilla  $j = 0, 1, \dots, n$ ,
- (ii)  $\omega_{j+1}$  riippuu vain olioista  $\omega_j$ ,  $j$  ja  $u_j$  kaikilla  $j = 0, 1, \dots, n-1$ ,
- (iii) sekä  $\omega_j \in \mathcal{X}$  kaikilla  $j = 0, 1, \dots, n$ .

Olkoot kuvaukset  $a_j: \mathcal{X} \times \mathcal{U}_j \rightarrow \mathcal{X}$ ,  $g_j: \mathcal{X} \times \mathcal{U}_j \rightarrow [-\infty, \infty]$  ja  $K: \mathcal{X} \rightarrow [-\infty, \infty]$ ,  $j = 0, 1, \dots, n$ . Oletamme, että kuvaus  $j \mapsto x_j$  toteuttaa

$$x_{j+1} = a_j(x_j, u_j), \quad j = 0, 1, \dots, n-1. \quad (7.12)$$

Tämä vastaa talon lämmittämisen tapauksessa lämpöyhtälöä, joka määrää, miten lämpötila  $x$  muuttuu kullakin lämmitysteholla  $u_j$ . Mekaniikan piiriin kuuluvissa sovellutuksissa yhtälöä (7.12) sanotaan liikeyhtälöksi; sanomme sitä yleisesti *tilayhtälöksi*. Minimioimme lauseketta

$$C = \sum_{j=0}^{n-1} g_j(x_j, u_j) + K(x_n) \quad (7.13)$$

säätömuuttujan  $j \mapsto u_j$  suhteen. Merkitään

$$\begin{aligned} F_j(x_j) &= F_j(x_0, x_1, \dots, x_j, u_0, u_1, \dots, u_{j-1}) = \\ &= \inf \left\{ \sum_{k=j}^{n-1} g_k(x_k, u_k) + K(x_n) \mid u_k \in \mathcal{U}_k, k = j, j+1, \dots, n-1 \right\}. \end{aligned} \quad (7.14)$$

Tämä on on pienin mahdollinen tuleva kustannus, kun takana on säätö- ja tilamuuttujien historia hetkestä  $k = 0$  hetkeen  $k = j$ .

**Lause 7.2.2** Olkoot  $n \in \mathbb{N}$ , epätyhjät joukot  $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_{n-1}$  ja  $\mathcal{X}$  sekä  $x_0 \in \mathcal{X}$ . Olkoon tilayhtälö (7.12) voimassa ja säätömuuttujan  $u$  optimaalisuuden määrittämään kaavan (7.13) kustannusfunktio. Tällöin pätee:

- (i) Optimaalinen  $u_j$  riippuu vain olioista  $j$  ja  $x_j$ .
- (ii)  $j \mapsto x_j$ ,  $\{0, 1, \dots, n\} \rightarrow \mathcal{X}$ , on säätömuuttujan  $j \mapsto u_j$  tilamuuttuja.

(iii) Lausekkeet  $F_j(x_j)$  riippuvat vain olioista  $j$  ja  $x_j$  eli

$$F_j: \mathcal{X} \rightarrow [-\infty, \infty], \quad j = 0, 1, \dots, n.$$

(iv) Kaikilla  $j = 0, 1, \dots, n-1$

$$F_j(x_j) = \inf_{v \in \mathcal{U}_j} \left( g_j(x_j, v) + F_{j+1}(a_j(x_j, v)) \right), \quad (7.15)$$

$$F_n(x_n) = K(x_n). \quad (7.16)$$

*Todistus.* Todistus on induktiivinen. Määritelmän (7.14) mukaan  $F_n(x_n) = K(x_n)$  eli (7.16) pätee. Samalla pätee, että  $F_n(x_0, \dots, x_n, u_0, \dots, u_{n-1})$  pelkistyy vain olioista  $x_n$  riippuvaksi. Olkoon tämä pelkistyminen voimassa oliolle  $F_{k+1}(x_0, \dots, x_{k+1}, u_0, \dots, u_k)$ , kun  $k \in \{0, 1, \dots, j-1\}$ , missä  $j \leq n$ . Optimaalisuusyhtälö (7.10) sanoo, että

$$F_k(x_k) = \inf_{u_k \in \mathcal{U}_k} \left( g_k(x_k, u_k) + F_{k+1}(a_k(x_k, u_k)) \right). \quad (7.17)$$

Lauseke, josta infimum otetaan, riippuu vain olioista  $x_k$ ,  $u_k$  ja  $k$ . Siten infimum riippuu vain olioista  $x_k$  ja  $k$ . Näin tämä pelkistyminen toteutuu myös oliolle  $F_k(x_0, \dots, x_k, u_0, \dots, u_{k-1})$ . Induktioperiaatteen nojalla se pätee kaikille  $k = n, n-1, \dots, 0$ . Täten (iii) ja (iv) pätevät. Yhtälöstä (7.17) nähdään, että optimaalinen  $u_k$  riippuu vain olioista  $x_k$  ja  $k$  eli (i) pätee. Tilayhtälön nojalla  $x_{k+1} \in \mathcal{X}$  riippuu vain olioista  $x_k$ ,  $u_k$  ja  $k$ , joten  $k \mapsto x_k$  on tilamuuttuja eli (ii) pätee. **m.o.t.**

### 7.2.1 Diskonttaus

Olkoon  $\beta > 0$ . Tarkastellaan *diskontattua*<sup>1</sup> kustannusfunktiota

$$C = \sum_{j=0}^{n-1} \beta^j g_j(x_j, u_j) + \beta^n K(x_n). \quad (7.18)$$

Määritellään oliot  $F_j(x_j)$ ,  $j = 0, 1, \dots, n$ , uudestaan:

$$\begin{aligned} F_j(x_j) &= F_j(x_0, x_1, \dots, x_j, u_0, u_1, \dots, u_{j-1}) = \\ &= \inf \left\{ \sum_{k=j}^{n-1} \beta^{k-j} g_k(x_k, u_k) + \beta^{n-j} K(x_n) \mid u_k \in \mathcal{U}_k, \quad k = j, j+1, \dots, n-1 \right\}. \end{aligned} \quad (7.19)$$

Tämä on on pienin mahdollinen tuleva diskontattu kustannus, kun takana on säätö- ja tilamuuttujien historia hetkestä  $k = 0$  hetkeen  $k = j$ . Huomaa, että jos  $\beta = 1$ , niin tilanne on sama kuin lauseen 7.2.2 tapauksessa.

**Lause 7.2.3** *Olkoot  $n \in \mathbb{N}$ ,  $\beta > 0$ , epättyhjät joukot  $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_n$  ja  $\mathcal{X}$  sekä  $x_0 \in \mathcal{X}$ . Olkoon tilayhtälö (7.12) voimassa ja säätömuuttujan  $u$  optimaalisuuden määrittämään kaavan (7.18) kustannusfunktio. Tällöin pätee:*

(i) Optimaalinen  $u_j$  riippuu vain olioista  $j$  ja  $x_j$ .

(ii)  $j \mapsto x_j$ ,  $\{0, 1, \dots, n\} \rightarrow \mathcal{X}$ , on säätömuuttujan  $j \mapsto u_j$  tilamuuttuja.

<sup>1</sup>Diskontto on lainan ennakokorko, joka maksetaan ennen lainan erääntymistä ns. diskonttauspäivänä ja jonka laskemisessa korko lasketaan diskonttaus- ja erääntymispäivän väliseltä ajalta. Esimerkki: 10 000 markan laina, jonka vuosikorko on 10 %, otetaan vuodeksi 1.1., koronmaksu- ja diskonttauspäivä ovat 30.6. Tällöin 30.6 maksetaan korko 1000 mk, josta 500 mk on taannehtivasti maksettava alkuvuoden korko ja loput 500 mk on ennakoiden (diskontaten) maksettava loppuvuoden korko.

(iii) Lausekkeet  $F_j(x_j)$  riippuvat vain olioista  $j$  ja  $x_j$  eli

$$F_j: \mathcal{X} \rightarrow [-\infty, \infty], \quad j = 0, 1, \dots, n.$$

(iv) Kaikilla  $j = 0, 1, \dots, n-1$

$$F_j(x_j) = \inf_{v \in \mathcal{U}_j} \left( g_j(x_j, v) + \beta F_{j+1}(a_j(x_j, v)) \right), \quad (7.20)$$

$$F_n(x_n) = K(x_n). \quad (7.21)$$

*Todistus.* Harjoitustehtävä.

**m.o.t.**

### 7.3 Optimaalinen kuluttaminen ja sijoittaminen

Tarkastellaan tapausta, jossa yksi henkilö voi kuluttaa tai sijoittaa rahansa. Kuluttaminen tuottaa mielihyvää, sijoitus korkotuloja, joilla voidaan kuluttaa. Henkilön tavoitteena on maksimoida diskontattu mielihyvänsä. Olkoot  $n \in \mathbb{N}$ ,  $\nu \in ]0, 1[$  ja  $\alpha, \beta > 0$ . Tila-avaruuksia olkoon  $\mathbb{R}$  ja tilamuuttuja  $j \mapsto x_j$ ; sen arvo  $x_j$  esittää henkilön rahavarojen määrää hetkellä  $t_j$ . Aika on siis jälleen diskreetti. Säättömuuttuja on  $j \mapsto u_j \in \mathcal{U}_j$  ja sen arvo  $u_j$  esittää hetkellä  $t_j$  kulutetun rahan määrää. Kuluttaa ei voi velaksi, joten  $u_j \leq x_j$ . Toisaalta tuottaminenkaan ei käy päinsä, joten  $u_j \geq 0$ . Täten  $\mathcal{U}_j = [0, x_j]$ . Raha on sijoitettu pankkitilille, jonka korko<sup>2</sup> on  $\alpha - 1$ , joten tilayhtälö on

$$x_{j+1} = \alpha(x_j - u_j), \quad j = 0, 1, \dots, n. \quad (7.22)$$

Kuluttajan mielihyvää hetkellä  $t_j$  esittää  $g(u_j)$ . Oletamme, että

$$g: [0, \infty[ \rightarrow \mathbb{R}, \quad g(r) = r^{1-\nu}. \quad (7.23)$$

Tällöin mielihyväfunktio on alenevan rajamielihyvän lain mukaisesti kovera. Hetkellisten mielihyvien summan sijasta maksimoidaan *diskontanttujen* mielihyvien summaa

$$\sum_{j=0}^{n-1} \beta^j g(u_j) + \beta^n K(x_n), \quad (7.24)$$

missä  $K(x_n)$  esittää mielihyvää loppuhetkellä olevasta rahamäärästä. Alenevan rajamielihyvän lain vuoksi — ja laskujen helpottamiseksi — oletamme, että

$$K: [0, \infty[ \rightarrow \mathbb{R}, \quad K(r) = A_n r^{1-\nu}, \quad (7.25)$$

missä  $A_n > 0$  on vakio. Meillä on *maksimointitehtävä*. Mutta koska kustannusfunktiossa kaikki yhteenlaskettavat ovat äärellisiä, niin yhtäpitävä minimointitehtävä on hakea olion

$$-\left( \sum_{j=0}^{n-1} \beta^j g(u_j) + \beta^n K(x_n) \right) = \sum_{j=0}^{n-1} \beta^j (-g(u_j)) + \beta^n (-K(x_n)) \quad (7.26)$$

minimiä. Koska jokaiselle reaalilukujoukolle  $B$  pätee  $\sup B = -\inf -B$ , niin lauseen 7.2.3 nojalla optimaalisuusyhtälöt ovat

$$-F_j(x_j) = \sup_{0 \leq v \leq x_j} \left( v^{1-\nu} - \beta F_{j+1}(\alpha(x_j - v)) \right), \quad j = 0, \dots, n-1, \quad (7.27)$$

$$-F_n(x_n) = K(x_n), \quad (7.28)$$

<sup>2</sup>Kyseessä on reaalikorko, jolloin inflaatio voi tehdä siitä negatiivisen.

missä jokaiselle  $j = 0, 1, \dots, n$

$$\begin{aligned} -F_j(x_j) &= -F_j(x_0, x_1, \dots, x_j, u_0, u_1, \dots, u_{j-1}) = \\ &= -\inf \left\{ \sum_{k=j}^{n-1} -\beta^{k-j} g(u_k) - \beta^{n-j} K(x_n) \mid u_k \in \mathcal{U}_k, k = j, j+1, \dots, n-1 \right\} \\ &= \sup \left\{ \sum_{k=j}^{n-1} \beta^{k-j} g(u_k) + \beta^{n-j} K(x_n) \mid 0 \leq u_k \leq x_k, k = j, \dots, n-1 \right\}. \end{aligned} \quad (7.29)$$

Optimaalisuusyhtälöiden ratkaisuna saadaan (harjoitustehtävä), että

$$-F_j(x_j) = A_j x_j^{1-\nu}, \quad u_j = \frac{x_j}{A_j^{1/\nu}}, \quad j = n, n-1, \dots, 0, \quad (7.30)$$

$$A_j^{1/\nu} = 1 + \gamma A_{j+1}^{1/\nu}, \quad j = 0, 1, \dots, n-1, \quad (7.31)$$

$$\gamma = (\beta \alpha^{1-\nu})^{1/\nu}. \quad (7.32)$$

Tällöin pätee

$$A_{n-j}^{1/\nu} = A_n^{1/\nu} \gamma^{n-j} + \frac{1 - \gamma^{n-j}}{1 - \gamma}, \quad j = 0, 1, \dots, n-1. \quad (7.33)$$

Tarkastellaan lopuksi likimääräisesti tapausta, jossa edessä on runsaasti aikaa eli  $n - j$  on suuri. Jos  $\gamma > 1$ , niin

$$A_{n-j}^{1/\nu} \approx (A^{1/\nu} - \frac{1}{1-\gamma}) \gamma^{n-j},$$

jolloin

$$-F_{n-j}(x_{n-j}) = A_{n-j} x_{n-j}^{1-\nu} \approx M_1 \gamma^{\nu(n-j)} x^{1-\nu}, \quad u_{n-j} \approx M_2 \frac{x_j}{\gamma^{n-j}}, \quad (7.34)$$

missä  $M_1$  ja  $M_2$  ovat positiivisia vakioita. Jälkimmäinen näistä yhtälöistä kertoo, että kulutusta kannattaa myöhentää. Jos  $\gamma < 1$  ja  $n - j$  on suuri, niin  $A_{n-j}^{1/\nu} \approx 1/(1 - \gamma)$ , jolloin

$$-F_{n-j}(x_{n-j}) \approx \frac{x_{n-j}^{1-\nu}}{(1-\gamma)^\nu}, \quad u_{n-j} \approx (1-\gamma)x_{n-j}. \quad (7.35)$$

Tällöin rahoista kulutetaan vakio-osa kullakin ajan askeleella.

## 7.4 Harjoitustehtäviä

1. Olkoon  $X = \{a, b, c\}$ . Määrittele sellaiset  $c, K: X \times X \rightarrow ]-\infty, \infty]$ , että

$$(a) \quad F_\infty(a, c) = -\infty,$$

$$(b) \quad F_\infty(a, c) = \infty,$$

$$(c) \quad F_\infty(a, c) = 10 \text{ ja } F_\infty(c, a) = 1.$$

2. Todista lause 7.1.1.

3. Olkoon  $X = \{\alpha, \beta, \delta, \gamma\}$  ja  $c, K: X \times X \rightarrow ]-\infty, \infty]$ ,

$$(c(\eta, \xi)) = \begin{pmatrix} 0 & 1 & 20 & 13 \\ 2 & 0 & 5 & 7 \\ 1 & 5 & 0 & 2 \\ 2 & 7 & 1 & 0 \end{pmatrix}, \quad K(\eta, \xi) = \begin{cases} \infty, & \text{jos } \eta \neq \xi, \\ 0, & \text{jos } \eta = \xi. \end{cases}$$

Selitä, miksi matriisi  $(c(\eta, \xi))$  voi olla epäsymmetrinen. Hae Bellmanin optimaalisuusperiaatteen avulla optimaalinen polku pisteestä  $\alpha$  pisteeseen  $\delta$ .

4. Olkoon  $n \in \mathbb{N}$  ja tilayhtälö muotoa

$$x_{k+1} = a_k(x_k, x_{k-1}, u_k), \quad k = 1, 2, \dots, n,$$

missä  $a_k: \mathbb{R}^3 \rightarrow \mathbb{R}$ . Osoita, että  $k \mapsto (x_k, x_{k-1})$  on tilamuuttuja, kun optimaalisuuden määrää kaavan (7.13) kustannusfunktio.

5. Johda diskontatun kustannusfunktion optimaalisuusehdot (7.20)-(7.21).
6. Johda kaavat (7.30)-(7.33).
7. Osoita, että yhtälöiden (7.35) mukaan toimittaessa pääoma kasvaa jokaisella ajan askeleella  $\theta = \alpha\gamma = (\alpha\beta)^{1/\nu}$  -kertaiseksi. Millainen tämä kerroin voi olla, kun  $\gamma < 1$ ?



# Kirjallisuutta

- [1] Jean-Pierre Aubin, “L’analyse non linéaire et ses motivations économiques”. Masson, Pariisi, 1984, venäjäksi Mir, Moskova, 1988.
- [2] Mario Bunge: “Philosophy of Physics”, D. Reidel, Dordrecht, 1973.
- [3] Pierre Duhem, “La théorie physique : son objet - sa structure”, 2. éd. rev. et augm., VRIN, Pariisi, 1981.
- [4] Frits J. De Jong: “Dimensional Analysis for Aconomists: with a Mathematical Appendix on the Algebraic Structure of Dimensional Analysis .” North Holland, Amsterdam, 1967.
- [5] I. Ekeland & R. Temam, “Analyse convexe et problèmes variationnels”, Dunod, Gauthier-Villars, Pariisi, 1974.
- [6] Veli-Matti Hokkanen & Eero Sarkkinen: *Dimensioanalyysi — salatiedettäkö?*, Arkhimedes 1/1985.
- [7] Hannu Karttunen: “Johdatus taivaanmekaniikkaan”, Limes ry, Helsinki, 1980.
- [8] Martin Kusch: “Ymmärtämisen haaste”, Pohjoinen, Oulu, 1986.
- [9] Kaisa Miettinen, “Optimointi”, Jyväskylän yliopisto, matematiikan laitos, luentomoniste 33, 1995.
- [10] Karl R. Popper: “The Logic of Scientific Discovery”, Hutchinson, Lontoo, 1987.
- [11] Peter Whittle: “Optimization over Time : Dynamic Programming and Stochastic Control”, Vol. 1, Wiley, Chichester, 1982.
- [12] Georg Henrik von Wright: “Explanation and Understanding”, Cornell University Press, Ithaca, N.Y., 1990.