Olkoon $A \subset \mathbb{R}^n$. Kuvaus $f : A \to \mathbb{R}$ on *n:n muuttujan reaalifunktio*. Se kuvaa A:n pisteet $x = (x_1, \ldots, x_n) \in A$ $(x_1, \ldots, x_n \in \mathbb{R})$ reaaliluvuiksi $f(x)$ ja koko A:n $\mathbb{R}$:n osajoukoksi $f(A) \subset \mathbb{R}$. Tapauksissa $n = 2$ tai $n = 3$ käytetään myös vaihtoehtoista merkintää

$\bar{r} = (x, y) = x\bar{i} + y\bar{j}$ $(\bar{i} = (1, 0),\ \bar{j} = (0, 1))$ tai

$\bar{r} = (x, y, z) = x\bar{i} + y\bar{j} + z\bar{k}$ $(\bar{i} = (1, 0, 0),\ \bar{j} = (0, 1, 0),\ \bar{k} = (0, 0, 1))$

merkintöjen $x = (x_1, x_2)$ ja $x = (x_1, x_2, x_3)$ sijaan.

Kertaan aluksi lineaarialgebran määritelmiä ja etäisyyteen liittyviä käsitteitä ja tuloksia:

Jos $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ ja $a \in \mathbb{R}$, niin

(3.1)
$$\begin{cases} x + y = (x_1 + y_1, \ldots, x_n + y_n), \text{ summa} \\ ax = (ax_1, \ldots, ax_n), \text{ luvun ja vektorin tulo} \\ x \cdot y = x_1 y_1 + \ldots + x_n y_n, \text{ pistetulo} \\ \|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + \ldots + x_n^2}, \text{ normi} \\ d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}, \text{ etäisyys} \\ B_r(x) = \{y \in \mathbb{R}^n \mid d(x, y) < r\}, \text{ } r\text{-säteinen, } x\text{-keskinen, } n\text{-ulotteinen avoin kuula} \\ \overset{\circ}{B}_r(x) = B'_r(x) = B_r(x) \setminus \{x\}, \text{ punkteerattu } n\text{-kuula} \end{cases}$$

3.2 Lause. (i) $|x \cdot y| \le \|x\|\|y\|$ $\forall x, y \in \mathbb{R}^n$ *(Schwarz)*

(ii) $\|ax\| = |a|\|x\|$ *(*$a \in \mathbb{R}$, $x \in \mathbb{R}^n$*)*

(iii) $\|x + y\| = \|x\| + \|y\|$, *Kolmioepäyhtälö* $\mathbb{R}^n$*:ssä.*

(iv) $d(x, z) \le d(x, y) + d(y, z)$, *Kolmioepäyhtälö* $\mathbb{R}^n$*:ssä.*

(v) *Tasossa* $\mathbb{R}^2$ $\|(x, y)\| = \sqrt{x^2 + y^2}$ *ja* $\mathbb{R}^3$*:ssa* $\|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$. $\square$

Joukko $A \subset \mathbb{R}^n$ on *avoin*, jos

$$\forall x \in A\ \exists r > 0 \text{ s.e. } B_r(x) \subset A, \tag{3.3}$$

ts. jos jokainen A:n piste on sisäpiste määritelmän 2.6 mielessä. Jos $\mathbb{R}^n \setminus A$ on avoin, A on *suljettu*. Pätee:

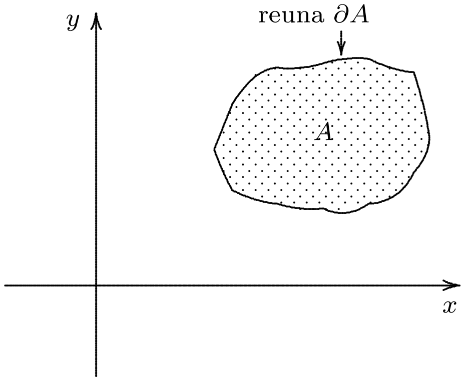
A suljettu $\Leftrightarrow$ $\partial A \subset A$ (eli A sisältää reunapisteidensä joukon ∂A, ks. 2.6).

3.4 Esimerkki. (i) Kolmioepäyhtälöstä seuraa, että avoin kuula $B_r(x)$ $(x \in \mathbb{R}^n$, $r > 0)$ on avoin ja suljettu kuula

$$\overline{B}_r(x) = \{y \in \mathbb{R}^n \mid d(x, y) \le r\}$$

on suljettu joukko.

(ii)



Jos kuvan joukon A reunakäyrä $\partial A \subset A$, A on suljettu. Jos $\partial A \subset \mathbb{R}^2 \setminus A$, A on avoin. Muuten A ei ole avoin eikä suljettu.

Useamman muuttujan funktion raja-arvo ja jatkuvuus

Perusidea: Lausekkeilla määritellyistä funktioista tulee jatkuvia muualla kuin mahdollisissa ongelmakohdissa (joita ovat esimerkiksi nimittäjän nollakohdat, negatiiviset argumentit neliöjuuren alla ja logaritmeissa jne.).

3.5 Määritelmä. Olkoon $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$ funktio ja $x_0 \in A$. Funktio f *on jatkuva pisteessä* x_0, jos

$$\forall \epsilon > 0\ \exists \delta > 0 \text{ s.e. } f(B_\delta(x_0) \cap A) \subset \,]f(x_0) - \epsilon, f(x_0) + \epsilon[.$$

Toinen muotoilu:

$$f \text{ jatkuva } x_0\text{:ssa } \Leftrightarrow\ \forall \epsilon > 0\ \exists \delta > 0 \text{ s.e. } \forall x \in A \text{ pätee ehto}$$

$$(*) \qquad d(x, x_0) = \|x - x_0\| < \delta \ \Rightarrow\ |f(x) - f(x_0)| < \epsilon.$$

Jos f on jatkuva x_0:ssa $\forall x_0 \in A$, sanomme, että f on *jatkuva* (määrittelyjoukossaan A).

3.6 Huomautus. (i) Jos $x_0 \in A$ on *eristetty piste* eli jos on olemassa $\delta > 0$ s.e. $B_\delta(x_0) \cap A = \{x_0\}$, jokainen $f : A \to \mathbb{R}$ on jatkuva x_0:ssa, sillä ehto $(*)$ pätee tällä δ:n arvolla (ja pienemmillä) triviaalisti.

(ii) Muissa kuin eristetyissä A:n pisteissä x_0 jatkuvuusehto voidaan pukea yhden muuttujan teoriasta tuttuun muotoon "raja-arvo on arvo":

$$f \text{ jatkuva } x_0\text{:ssa } \Leftrightarrow\ \lim_{x \to x_0} f(x) = f(x_0).$$

Tätä varten joudumme tutkimaan f:n raja-arvoa x_0:ssa joukossa A ja havaitsemme hyödylliseksi tutkia raja-arvoa x_0:ssa "pitkin joukkoa" $B \subset A$, kun pistettä x_0 "voidaan lähestyä" pitkin B:tä.

3.7 Esimerkki. Tutkitaan funktiota $f(x,y) = \dfrac{xy}{x^2+y^2}$ kun $(x,y) \to (0,0)$ pitkin suoraa $y = kx$ ($k \in \mathbb{R}$). Tällöin $f(x,y) = f(x,kx) = \dfrac{kx^2}{x^2+k^2x^2} = \dfrac{k}{1+k^2}$ on vakio, joten sen raja-arvona on sama k:sta riippuva vakio. Koska nämä raja-arvot eroavat toisistaan, on syytä olettaa, ettei f:llä ole minkään "järkevän" määritelmän mielessä raja-arvoa $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ ja arvoa $f(0,0)$ ei voi määritellä niin, että f:stä tulisi myös origossa jatkuva.

3.8 Määritelmä. (i) Piste $x_0 \in \mathbb{R}^n$ on joukon $A \subset \mathbb{R}^n$ *kasaantumispiste*, jos x_0:n jokaisessa ympäristössä on A:n x_0:sta eroavia pisteitä:

$$\forall \epsilon > 0 \ : \ B'_\epsilon(x_0) \cap A \neq \emptyset.$$

(ii) Olkoon $B \subset A \subset \mathbb{R}^n$ ja olkoon x_0 B:n kasaantumispiste sekä $f : A \to \mathbb{R}$ funktio. (Tällöin x_0 on myös A:n kasaantumispiste ja pistettä x_0 "voidaan lähestyä" pitkin joukkoja B ja A.) Sanomme, että *funktiolla f on pisteessä x_0 raja-arvo $a \in \mathbb{R}$ pitkin joukkoa B* ja merkitsemme

$$\lim_{x\to x_0,\ x\in B} f(x) = a, \quad \text{jos } \forall\epsilon > 0\ \exists\delta > 0 \text{ s.e.}$$

$$0 < \|x - x_0\| < \delta \text{ ja } x \in B \ \Rightarrow\ |f(x) - a| < \epsilon$$

Toinen muotoilu:

$$\lim_{x\to x_0,\ x\in B} f(x) = a \ \Leftrightarrow\ \forall\epsilon > 0\ \exists\delta > 0 \text{ s.e. } f(B'_\delta(x_0) \cap B) \subset\ ]a - \epsilon, a + \epsilon[\,.$$

Tapauksessa $B = A$ (f:n määrittelyjoukko) merkitään lyhyemmin

$$\lim_{x\to x_0,\ x\in A} f(x) = \lim_{x\to x_0} f(x)$$

3.9 Lause. *Jos $\lim\limits_{x\to x_0,\ x\in B} f(x) = a$ ja $C \subset B \subset A$ (x_0 C:n kasautumispiste, $f : A \to \mathbb{R}$), niin myös $\lim\limits_{x\to x_0,\ x\in C} f(x) = a$.* [Ts. pienempää joukkoa pitkin tulee sama raja-arvo kuin isompaa joukkoa pitkin.]

Todistus. Olkoon $\epsilon > 0$. Oletuksesta $\lim\limits_{x\to x_0,\ x\in B} f(x) = a$ seuraa, että on olemassa $\delta > 0$ s.e. $f(B'_\delta(x_0) \cap B) \subset\ ]a - \epsilon, a + \epsilon[$. Tällöin

$$f(B'_\delta(x_0) \cap C) \subset f(B'_\delta(x_0) \cap B) \subset\ ]a - \epsilon, a + \epsilon[$$

ja väite seuraa. $\square$

3.10 Lause. *Jos raja-arvo* $\lim\limits_{x\to x_0,\ x\in B} f(x)$ *on olemassa, se on yksikäsitteinen.*

Todistus. Kuten yhden muuttujan teoriassa. $\square$

3.11 Seuraus. *Jos* $B \subset A$ *ja* $C \subset A$ *ja* x_0 *on* B*:n ja* C*:n kasaantumispiste ja* $f : A \to \mathbb{R}$ *toteuttaa*

$$\lim_{x\to x_0,\ x\in B} f(x) \neq \lim_{x\to x_0,\ x\in C} f(x),$$

niin raja-arvoa $\lim\limits_{x\to x_0} f(x)$ *ei ole olemassa.*

Todistus. Lauseet 3.9 ja 3.10. $\square$

Jos siis "kahta pienempää joukkoa pitkin tulee eri raja-arvot", f:llä ei ole raja-arvoa.

3.12 Esimerkki. (i) Koska $\lim\limits_{(x,y)\to(0,0), y=kx} \dfrac{xy}{x^2+y^2} = \dfrac{k}{1+k^2}$ esimerkin 3.7 nojalla, seuraa 3.11:stä, että $\nexists \lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ (valitsemalla kaksi eri k:n arvoa, jotka antavat 3.11:n joukoiksi B ja C kaksi eri suoraa $y = kx$).

(ii) Jos $f(x,y) = \dfrac{xy^2}{x^2+y^4}$, niin joukossa $B_k = \{(x,y) \mid y = kx,\ x \neq 0\}$ f:llä on kaikilla $k \in \mathbb{R}$ raja-arvo 0 origossa, sillä

$$y = kx \ \Rightarrow\ \frac{xy^2}{x^2+y^4} = \frac{x\cdot k^2x^2}{x^2+k^4x^4} = \frac{k^2x}{1+k^4x^2} \xrightarrow[x\to 0]{} 0 \quad (\text{ja } B_k\text{:ssa } (x,y)\to 0 \ \Leftrightarrow\ x \to 0.)$$

Siitä huolimatta $\nexists \lim\limits_{(x,y)\to(0,0)} f(x,y)$!

Perustelu: Käyrää $x = y^2$ pitkin raja-arvoksi tuleekin $\frac{1}{2} \neq 0$, sillä

$$x = y^2 \ \Rightarrow\ \frac{xy^2}{x^2+y^4} = \frac{y^4}{y^4+y^4} = \frac{1}{2} \xrightarrow[y\to 0]{} \frac{1}{2}$$

Nytkin 3.11 $\Rightarrow$ $\nexists \lim\limits_{(x,y)\to(0,0)} f(x,y)$.

Positiivinen raja-arvotulos todistetaan arvioimalla f:n ja väitetyn raja-arvon erotuksen itseisarvoa sopivalla tavalla.

3.13 Esimerkki. (i) *Väite:* $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2y^2}{x^2+y^2} = 0.$

Todistus:

$$\left|\frac{x^2y^2}{x^2+y^2} - 0\right| \leq \frac{(x^2+y^2)y^2}{x^2+y^2} = y^2$$

[joten tässä $\left|\dfrac{x^2y^2}{x^2+y^2} - 0\right| < \epsilon$, kun $0 < \|(x,y)\| = \sqrt{x^2+y^2} < \sqrt{\epsilon} = \delta$, jolloin $y^2 \leq x^2+y^2 < \epsilon$]. Johtopäätökseen riittää siis se havainto, että saatu yläraja $y^2 \to 0$, kun $(x,y) \to (0,0)$.

(ii) Todistetaan kohdan (i) väite myös napakoordinaattien avulla

$$\begin{aligned} &\left|\frac{x^2y^2}{x^2+y^2} - 0\right| = \left|\frac{r^4\sin^2\varphi\cos^2\varphi}{r^2}\right| = |r^2\sin^2\varphi\cos^2\varphi| \\ &\leq r^2 = x^2+y^2 = \|(x,y)\|^2 \underset{(x,y)\to(0,0)}{\longrightarrow} 0, \end{aligned}$$

joten $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2} = 0$. Tässä on arvioissa päästävä eroon φ:stä ja se käy havaitsemalla, että $|\sin^2\varphi\cos^2\varphi| \leq 1$.

(iii) Projektiokuvauksilla $\mathrm{pr}_i : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{pr}_i(x) = x_i \quad \forall x = (x_1,\ldots,x_n) \in \mathbb{R}^n \quad (i \in \{1,\ldots,n\}),$$

raja-arvo on arvo: Jos $y = (y_1,\ldots,y_n) \in \mathbb{R}^n$, niin

Väite: $\displaystyle\lim_{x\to y} \mathrm{pr}_i(x) = y_i \ \big(= \mathrm{pr}_i(y)\big)$

Todistus: Jos $\epsilon > 0$ ja $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$, niin arvion

$$0 \leq |x_i - y_i| \leq \sqrt{(x_1-y_1)^2 + \ldots + (x_i-y_i)^2 + \ldots + (x_n-y_n)^2} = \|x-y\|$$

nojalla saadaan valitsemalla $\delta = \epsilon$, että

$$\|x-y\| < \delta = \epsilon \ \Rightarrow\ |x_i - y_i| = |\,\mathrm{pr}_i(x) - \mathrm{pr}_i(y)| < \epsilon. \quad \square$$

Äsken saadun tuloksen voi myös muotoilla seuraavasti: "projektiokuvaukset $\mathrm{pr}_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1,\ldots,n)$ ovat jatkuvia":

3.14 Lause. *Funktio $f : A \to \mathbb{R}$ $(A \subset \mathbb{R}^n)$ on jatkuva joukon A kasaantumispisteessä $x_0 \in A$, jos ja vain jos f:n raja-arvo x_0:ssa on f:n arvo. Siis*

$$f \text{ jatkuva } x_0\text{:ssa} \ \Leftrightarrow\ \lim_{x\to x_0} f(x) = f(x_0), \quad \text{(vrt. 3.6(ii))}.$$

Todistus. Suora seuraus määritelmistä 3.5 ja 3.8. $\square$

Myös useamman muuttujan teoriassa pätevät tutut raja-arvojen laskusäännöt:

3.15 Lause. *Olkoot $f : A \to \mathbb{R}$ ja $g : A \to \mathbb{R}$ $(A \subset \mathbb{R}^n)$ funktioita, $x_0 \in \mathbb{R}^n$ osajoukon $B \subset A$ kasaantumispiste ja*

$$\lim_{x\to x_0, x\in B} f(x) = a, \qquad \lim_{x\to x_0, x\in B} g(x) = b, \quad (a,b \in \mathbb{R}).$$

Tällöin

(i) $$\lim_{x\to x_0, x\in B}\big(f(x)\pm g(x)\big)=a\pm b,$$

(ii) $$\lim_{x\to x_0, x\in B} cf(x)=ca\ \forall c\in\mathbb{R},$$

(iii) $$\lim_{x\to x_0, x\in B}\big(f(x)g(x)\big)=ab, \quad ja$$

(iv) $$\lim_{x\to x_0, x\in B}\frac{f(x)}{g(x)}=\frac{a}{b}, \quad kun \quad b\neq 0$$

Todistus. Sivuutetaan (ei eroa paljonkaan vastaavan yhden muuttujan tuloksen todistuksesta). □

3.16 Seuraus. *Jatkuvien funktioiden summa, erotus, tulo ja osamäärä ovat jatkuvia (osamäärän tapauksessa ei kuitenkaan nimittäjän nollakohdissa).*

Todistus. Lauseet 3.15 ja 3.14. □

3.17 Huomautus. Lauseen 3.15 voi laajentaa koskemaan myös raja-arvoja $\pm\infty$ (määrittele nämä käsitteet!), kunhan lasketaan syksyn kurssista tutuilla ”sallituilla muodoilla”.

3.18 Esimerkki. Koska $\lim_{(x,y)\to(1,0)} x = 1$ (3.13(ii)) ja $\lim_{(x,y)\to(1,0)}\frac{1}{y^2}=\infty$ $\Big($sillä $\frac{1}{y^2}>M$ (>0), kun $|y|<\frac{1}{\sqrt{M}}$ eli esimerkiksi ystössä $B_{\frac{1}{\sqrt{M}}}((1,0))\Big)$, saamme, että

$$\lim_{(x,y)\to(1,0)}(x+\frac{1}{y^2})=1+\infty=\infty.$$

Vektoriarvoisen funktion $f: A\to\mathbb{R}^p$ $(A\subset\mathbb{R}^n)$ raja-arvo ja jatkuvuus voidaan määritellä suoraan tai palauttaa ne f:n *komponenttifunktioiden*

(3.19) $$f_i=\mathrm{pr}_i\circ f: A\to\mathbb{R}\quad (i=1,\dots,p)$$

vastaaviin käsitteisiin (pr_i ks. 3.13(iii)). Esimerkiksi jatkuvuuden määritelmät ovat

(i) $$f: A\to\mathbb{R}^p \text{ jatkuva pisteessä } x_0\in A \iff$$
$$\forall\epsilon>0\ \exists\delta>0 \text{ s.e. } \|x-x_0\|<\delta \Rightarrow \|f(x)-f(x_0)\|<\epsilon$$

(vasemman puolen normi $\mathbb{R}^n$:ssä oikeanpuoleinen $\mathbb{R}^p$:ssä) ja toisella tavalla

(ii) $$f: A\to\mathbb{R}^p \text{ jatkuva pisteessä } x_0\in A \iff$$
$$f_i=\mathrm{pr}_i\circ f \text{ jatkuva } x_0\text{:ssa } \forall i\in\{1,\dots,p\}.$$

Lopputulos on sama ja tämän todistus ei ole vaikeaa.

3.20 Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x,y) = (x, xy, x^2 + y)$. Tällöin $f_1(x,y) = x$, $f_2(x,y) = xy$ ja $f_3(x,y) = x^2 + y$ ovat jatkuvia (3.15). Siis lauseen 3.13(iii) nojalla $f = (f_1, f_2, f_3)$ on jatkuva (koska sen komponenttifunktiot $f_i = \mathrm{pr}_i \circ f : \mathbb{R}^2 \to \mathbb{R}$ ovat jatkuvia).

Jatkuvien kuvausten yhdistetty kuvaus on aina jatkuva.

3.21 Lause. *Olkoot $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^p$ sekä $f : A \to \mathbb{R}^p$, $g : B \to \mathbb{R}^m$ jatkuvia. Jos $f(A) \subset B$, niin yhdistetty kuvaus $g \circ f : A \to \mathbb{R}^m$ on jatkuva.*

Kuva.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & B \arrow[r, "g"] & \mathbb{R}^m \end{tikzcd}$$

Todistus. Olkoon $\epsilon > 0$, $x_0 \in A$. Koska g on jatkuva $f(x_0)$:ssa on olemassa $\delta_1 > 0$ s.e.

$$(*) \qquad \|y - f(x_0)\| < \delta_1 \quad \text{ja} \quad y \in B \;\Rightarrow\; \|g(y) - g(f(x_0)\| < \epsilon$$

Koska f on jatkuva x_0:ssa on olemassa $\delta_2 > 0$ s.e. jos $x \in A$, niin

$$(**) \qquad \|x - x_0\| < \delta_2 \;\Rightarrow\; \|f(x) - f(x_0)\| < \delta_1$$

Valitsemalla $\delta = \delta_2$ saadaan nyt

$$x \in A \quad \text{ja} \quad \|x - x_0\| < \delta \overset{(**)}{\Rightarrow} \|\underbrace{f(x)}_{\in B} - f(x_0)\| < \delta_1$$

$$\overset{(*)}{\Rightarrow} \|g(f(x)) - g(f(x_0))\| < \epsilon \quad \text{eli} \quad \|(g \circ f)(x) - (g \circ f)(x_0)\| < \epsilon. \quad \square$$

Kuten yhden muuttujan funktioiden teoriassa, nämä lauseet takaavat sen, että ”lausekkeilla määritellyt funktiot ovat jatkuvia muualla kuin lausekkeiden muodostamiseen liittyvissä ongelmakohdissa”.

3.22 Esimerkki. i) *Rationaalifunktiot* ovat jatkuvia muualla kuin nimittäjän nollakohdissa. Esimerkiksi rationaalifunktio

$$f(x,y,z) = \frac{xy + x + z^2}{(x-1)^2 yz}$$

on x:n y:n ja z:n *polynomien*

$$P(x,y,z) = xy + x + z^2 \quad \text{(aste 2)} \quad \text{ja}$$
$$Q(x,y,z) = (x-1)^2 yz = x^2yz - 2xyz + yz \quad \text{(aste 4)}$$

osamäärä $f = \dfrac{P}{Q}$. Se on määritelty ja jatkuva $\mathbb{R}^3$:n osajoukossa

$$A = \{(x,y,z) \mid Q(x,y,z) \neq 0\} = \{(x,y,z) \mid x \neq 1,\ y \neq 0,\ z \neq 0\}.$$

ii) Funktio $f(x,y) = \sqrt{\sin(xy) + |x|}$ on jatkuva määrittelyjoukossaan

$$A = \{(x,y) \mid \sin(xy) + |x| \geq 0\} \subset \mathbb{R}^2$$

Perustelu: Esimerkin 3.13 (iii) nojalla $(x,y) \mapsto x$ ja $(x,y) \mapsto y$ ovat jatkuvia kaikkialla (joten niiden rajoittumat A:han ovat myös jatkuvia (lauseet 3.9 ja 3.14)), siten seurauksen 3.16 nojalla $(x,y) \mapsto xy$ on jatkuva A:ssa ja edelleen lauseesta 3.21 seuraa, että $(x,y) \mapsto \sin(xy)$ on jatkuva A:ssa (sillä $\sin : \mathbb{R} \to \mathbb{R}$ on jatkuva). Koska $(x,y) \mapsto |x|$ on jatkuva $\big[3.21 :\ (x,y) \overset{\mathrm{pr}_1}{\mapsto} x \underset{\text{arvo}}{\overset{\text{itseis}}{\mapsto}} |x|$ on jatkuvien kuvausten yhdiste$\big]$, niin lauseesta 3.15 seuraa, että $(x,y) \mapsto \sin(xy) + |x|$ on jatkuva A:ssa. Koska se on A:ssa ≥ 0 ja koska neliöjuuri $[0,\infty[\to [0,\infty[$ on jatkuva, niin lauseen 3.21 nojalla f on jatkuva A:ssa.

Myös useamman muuttujan funktioiden yhteydessä voidaan tietysti tarkastella erilaisia lukujonoja tai vektorijonoja ja niiden raja-arvoja.

3.23 Esimerkki. Määritä $\lim\limits_{n\to\infty} \left(2^{\frac{n-1}{n}}, \ln\left(\left(1 + \frac{\pi}{n}\right)^n\right)\right)$.

Ratkaisu. Koska $\dfrac{n-1}{n} = 1 - \dfrac{1}{n} \underset{n\to\infty}{\longrightarrow} 1$, koska $\left(1 + \dfrac{\pi}{n}\right)^n \underset{n\to\infty}{\longrightarrow} e^\pi$, ja koska funktiot $x \mapsto 2^x$ ja $x \mapsto \ln x$ ovat jatkuvia pisteissä 1 ja e^π on

$$\lim_{n\to\infty} \left(2^{\frac{n-1}{n}}, \ln\left(\left(1 + \frac{\pi}{n}\right)^n\right)\right) = \left(2^1, \ln(e^\pi)\right) = \underline{\underline{(2,\pi)}}.$$

Osittaisderivaatat ja gradientit, differentioituvuus

Olkoon $x_0 = (a_1, \ldots, a_n) \in A$ joukon $A \subset \mathbb{R}^n$ sisäpiste ja olkoon $f : A \to \mathbb{R}$ funktio. Yhden muuttujan funktiot

$$g_i(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n) \quad (i = 1, \ldots, n)$$

ovat tällöin määritellyt pienillä a_i-keskisillä väleillä (sillä x_0 oli A:n sisäpiste) ja riippuvat vain f:n arvoista x_0:n kautta kulkevalla x_i-akselin suuntaisella suoralla. Jos g_i on derivoituva pisteessä $t = a_i$, niin sen derivaatta

$$g_i'(a_i) = \lim_{h\to 0} \frac{g_i(a_i + h) - g_i(a_i)}{h} \tag{3.24}$$

on f:n *osittaisderivaatta muuttujan x_i suhteen pisteessä x_0*. Merkintöjä:

$$g_i'(a_i) = D_i f(x_0) = \frac{\partial f}{\partial x_i}(x_0) = D_{x_i} f(x_0) = f_{x_i}(x_0). \tag{3.25}$$

Jos A on avoin (ts. jokainen $x_0 \in A$ on sisäpiste) ja f:llä on jokaisessa $x_0 \in A$ kaikki osittaisderivaatat $D_i f(x_0)$ $(i = 1, \ldots, n)$, f on *derivoituva* A*:ssa*. Osittaisderivaatoista koottu n-vektori

$$\operatorname{grad} f(x_0) = \nabla f(x_0) = (D_1 f(x_0), \ldots, D_n f(x_0)) \tag{3.26}$$

on f:n *gradientti(vektori)* x_0:ssa. Se on määritelty vain kun f on *derivoituva* x_0*:ssa* (ts. $\exists\, D_i f(x_0)\ \forall i$).

3.27 Huomautus. (i) Gradientin määritelmässä esiintyvä symboli ∇ (”nabla”) on ”derivaattaoperaattori” $\nabla = (D_1, \ldots, D_n)$, jonka ajatellaan operoivan reaalifunktioon f seuraavasti:

$$\nabla f = (D_1, \ldots, D_n) f = (D_1 f, \ldots, D_n f).$$

Näemme myöhemmin, että gradientti on jonkinlainen yhden muuttujan funktioiden derivaatan vastine: Funktio f esimerkiksi kasvaa x_0:ssa nopeimmin gradienttivektorinsa $\nabla f(x_0)$ suuntaan.

(ii) Osittaiderivaattojen käytännön laskeminen sujuu peruskurssista tutuilla derivoimissäännöillä, sillä kyseessä on yhden muuttujan funktioiden $g_i(t)$ tavallinen derivointi; muut kuin i:s muuttuja x_i ajatellaan vakioiksi ja derivoidaan x_i:n suhteen.

(iii) Vektoriarvoisen funktion $f : A \to \mathbb{R}^p$ $(A \subset \mathbb{R}^n)$ derivointi palautetaan koordinaattifunktioiden $f_j = \mathrm{pr}_j \circ f$ $(j = 1, \ldots, p)$ derivointiin kaavalla

$$D_i f(x_0) = (D_i f_1(x_0), \ldots, D_i f_p(x_0)) \quad (i = 1, \ldots, n),$$

missä x_0 on A:n sisäpiste.

3.28 Esimerkki. i) $f(x_1, x_2, x_3, x_4) = x_1 e^{x_2 x_3} + x_3^2 x_4$. Tällöin

$$D_1 f(x) = e^{x_2 x_3}, \quad D_2 f(x) = x_1 x_3 e^{x_2 x_3}, \quad D_3 f(x) = x_1 x_2 e^{x_2 x_3} + 2 x_3 x_4 \quad \text{ja}$$
$$D_4 f(x) = x_3^2 \quad \forall x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4.$$

ii) $f(x, y) = \dfrac{x}{y^2} + \sin(x + \ln y)$. Tällöin f on määritelty joukossa $A = \{(x, y) \mid y > 0\}$ ja tämä joukko A on avoin $\mathbb{R}^2$:ssa. Nyt

$$D_1 f(x, y) = \frac{1}{y^2} + \cos(x + \ln y) \quad \text{ja}$$

$$D_2 f(x, y) = -\frac{2x}{y^3} + \frac{\cos(x + \ln y)}{y} \ \forall (x, y) \in A,$$

joten esimerkiksi pisteessä $(0, 1) \in A$ on

$$\begin{aligned}\operatorname{grad} f(0,1) &= \nabla f(0,1) = (D_1 f(0,1), D_2 f(0,1)) = \\ &= \left(\frac{1}{1^2} + \cos(0 + \ln 1), -\frac{2 \cdot 0}{1^3} + \frac{\cos(0 + \ln 1)}{1}\right) = (2, 1).\end{aligned}$$

Pisteessä $(0,1)$ tämä funktio f kasvaa siis nopeimmin vektorin

$$\nabla f(0,1) = (2,1) = 2\bar{i} + \bar{j}$$

suuntaan.

iii) Vektorifunktion $f : \mathbb{R}^3 \to \mathbb{R}^2$,

$$f(x,y,z) = \big(f_1(x,y,z), f_2(x,y,z)\big) = (x^3y^4, x^2+y^2+z^2)$$

osittaisderivaattafunktiot ovat

$$\begin{aligned}
D_1 f(x,y,z) &= \big(D_x(x^3y^4), D_x(x^2+y^2+z^2)\big) = (3x^2y^4, 2x),\\
D_2 f(x,y,z) &= \big(D_y(x^3y^4), D_y(x^2+y^2+z^2)\big) = (4x^3y^3, 2y),\\
D_3 f(x,y,z) &= \big(D_z(x^3y^4), D_z(x^2+y^2+z^2)\big) = (0, 2z).
\end{aligned}$$

Jos funktion $f : A \to \mathbb{R}$ osittaisderivaatta $D_i f$ on olemassa eräissä A:n sisäpisteissä x_0, saadaan vastaava *osittaisderivaattafunktio* $D_i f : B \to \mathbb{R}$, missä

$$B = \{x_0 \in A \mid x_0 \text{ on } A\text{:n sisäpiste ja } \exists\, D_i f(x_0)\}.$$

Näillä voi edelleen olla *korkeampia osittaisderivaattoja* $D_j(D_i f) = D_{ij} f$. Jos $A \subset \mathbb{R}^n$ on avoin merkitään

$$\begin{aligned}
C^0(A) &= \{f \mid f : A \to \mathbb{R} \text{ jatkuva}\}\\
C^1(A) &= \{f \mid f : A \to \mathbb{R} \textit{ jatkuvasti derivoituva},\\
&\qquad\quad \text{ts. } D_i f : A \to \mathbb{R} \text{ jatkuva } \forall i \in \{1,\dots,n\}\}\\
C^2(A) &= \{f \mid D_j f \in C^1(A)\ \forall j \in \{1,\dots,n\}\} =\\
&= \{f \mid f \text{ kahdesti jatkuvasti derivoituva}\}
\end{aligned} \tag{3.29}$$

jne. Saadaan jono sisäkkäisiä A:ssa määriteltyjen reaalifunktioiden avaruuksia

$$C^0(A) \supset C^1(A) \supset C^2(A) \supset \dots$$

ja näiden leikkauksena

$$C^\infty(A) = \{f \mid f\text{:llä on kaikkien kertalukujen jatkuvat osittaisderivaatat}\}. \tag{3.30}$$

3.31 Huomautus. Koska osittaisderivaatat $D_i f(x_0)$ riippuvat vain f:n arvoista x_0:n kautta kulkevilla koordinaattiakseleiden suuntaisilla suorilla, on selvää, ettei pelkkä osittaisderivaattojen olemassaolo kerro mitään funktiosta f näiden suorien ulkopuolella.
Esimerkki

$$f(x,y) = \begin{cases} 0, x = 0 \quad \text{tai } \ y = 0 \\ 1, x \neq 0 \quad \text{ja } \ y \neq 0 \end{cases}$$

näyttää, ettei f:n edes tarvitse olla jatkuva pisteessä, jossa se on derivoituva. Tässä näet $D_1 f(0,0) = D_2 f(0,0) = 0$, mutta f ei ole jatkuva origossa.

Kohdassa 3.29 määritelty *jatkuva derivoituvuus x_0:ssa*, ts. osittaisderivaattafunktioiden olemassaolo ja jatkuvuus pisteessä $x_0 \in A$, sen sijaan kertoo jo paljon f:stä. Jatkuvien ja jatkuvasti derivoituvien funktioiden "väliin" mahtuvat vielä ns. *differentioituvat* funktiot, joita kohta käsitellään.

3.32 Esimerkki. Jos $f : \mathbb{R}^n \to \mathbb{R}$ on polynomi, niin $f \in C^\infty(\mathbb{R}^n)$ ja riittävän korkeat f:n osittaisderivaatat ovat nollia. Jos esimerkiksi $f : \mathbb{R}^3 \to \mathbb{R}$ on polynomi

$$f(x,y,z) = xy^2 + xyz + z^2 x,$$

niin osittaisderivaatat

$$D_1 f = y^2 + yz + z^2, \ D_2 f = 2xy + xz, \quad \text{ja } \ D_3 f = xy + 2xz$$

ovat jatkuvia, joten $f \in C^1(\mathbb{R}^3)$. Edelleen toisen kertaluvun osittaisderivaatat

$$\begin{aligned} &D_1(D_1 f) = D_{11} f = 0, \ D_2(D_1 f) = D_{12} f = 2y + z, \ D_3(D_1 f) = D_{13} f = y + 2z, \\ &D_1(D_2 f) = D_{21} f = 2y + z, \ D_{22} f = 2x, \ D_{23} f = x, \ D_{31} f = y + 2z, \ D_{32} f = x, \\ &D_{33} f = 2x \end{aligned}$$

ovat jatkuvia, joten $f \in C^2(\mathbb{R}^3)$. Toisen kertaluvun osittaisderivaatoista $D_{ij} f$ ja $D_{ji} f$ havaitaan tässä esimerkissä, että ne aina yhtyvät. Jatkuvasti derivoituville funktioille tämä seuraa ns. *derivoimisjärjestyksen vaihtosäännöstä*:

3.33 Lause. *Jos $A \subset \mathbb{R}^n$ on avoin ja $f \in C^k(A)$ on A:ssa k-kertaa jatkuvasti derivoituva ($k \in \mathbb{N}$, $k \geq 2$), niin f:n osittaisderivaatat kertalukuun k asti riippuvat ainoastaan siitä, montako kertaa kunkin muuttujan suhteen on derivoitu, eivät siis derivoimisjärjestyksestä.*

Todistus. Sivuutetaan liian työläänä. $\square$

Toisen kertaluvun osittaiderivaatoista voidaan muodostaa ns. *Hessen matriisi*, jolla on käyttöä mm. ääriarvojen teoriassa:

3.34 Määritelmä. Olkoon $A \subset \mathbb{R}^n$, x_0 A:n sisäpiste ja $f : A \to \mathbb{R}$ kahdesti derivoituva x_0:ssa. Matriisi

$$H_f(x_0) = H(x_0) = \begin{bmatrix} D_{11}f(x_0) & D_{12}f(x_0) & \dots & D_{1n}f(x_0) \\ \vdots & \vdots & & \vdots \\ D_{n1}f(x_0) & D_{n2}f(x_0) & \dots & D_{nn}f(x_0) \end{bmatrix}$$

$$\big(\text{ts. } [H(x_0)]_{ij} = D_{ij}f(x_0)\ \forall i,j \in \{1,\dots,n\}\big)$$

on f:n *Hessen matriisi pisteessä* x_0 ja sen determinantti $\det H(x_0) \in \mathbb{R}$ on *f:n Hessen determinantti x_0:ssa.*

Jos A on avoin ja $f \in C^2(A)$, f:n Hessen matriisi $H_f(x_0)$ on symmetrinen $\forall x_0 \in A$ lauseen 3.33 nojalla.

3.35 Esimerkki. (i) Jos $f(x,y)$ on kahden muuttujan x ja y *neliömuoto*

$$(*) \qquad \begin{cases} f(x,y) = ax^2 + 2bxy + cy^2 \\ \Big(= [\,x \quad y\,] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \text{ ”matriisikielellä”}\Big), \end{cases}$$

niin

$$D_1 f(x,y) = 2ax + 2by,\ D_2 f(x,y) = 2bx + 2cy,$$
$$D_{11} f = 2a,\ D_{12} f = 2b = D_{21} f,\ D_{22} f = 2c,$$

joten f:n gradientti on ”matriisikielellä” ilmaistuna

$$\nabla f(x,y) = 2 \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2A \begin{bmatrix} x \\ y \end{bmatrix}$$

ja f:n Hessen matriisi on vakiomatriisi

$$H_f = \begin{bmatrix} 2a & 2b \\ 2b & 2c \end{bmatrix} = 2A,$$

missä $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ on neliömuodon f määrittelevä symmetrinen matriisi (ks. $(*)$).

(ii) Vastaava pätee samanlaisella laskulla n:n muuttujan neliömuodolle

$$f(x_1,\dots,x_n) = [\,x_1 \quad \dots \quad x_n\,]\, A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

(A on se symmetrinen $n \times n$-matriisi, jonka (i,i)-alkiot ovat neliömuodon f termien x_i^2 kertoimet ja (i,j) ja (j,i)-alkiot ovat puolet termin $x_i x_j$ kertoimesta, kun $i \neq j$ ja samanmuotoiset termit $x_i x_j$ ja $x_j x_i$ on yhdistetty):

$$\nabla f(x) = 2A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{ja} \qquad H_f = 2A.$$

Yhden muuttujan funktioiden teoriasta tuttu differentiaalikehitelmä on tärkeä myös useamman muuttujan funktioille.

3.36 Määritelmä. Olkoon $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$ funktio ja x_0 A:n sisäpiste. Funktio f on *differentioituva* x_0*:ssa*, jos sillä on x_0:n ympäristössä ns. *differentiaalikehitelmä* eli jos f:n muutos $f(x_0+h) - f(x_0)$ voidaan esittää muodossa

$$(3.37)\quad f(x_0+h)-f(x_0) = a\cdot h+\|h\|\epsilon(h) = a_1h_1+\ldots+a_nh_n+\sqrt{h_1^2+\ldots+h_n^2}\,\epsilon(h),$$

missä $a \in \mathbb{R}^n$ on vakiovektori ja $\epsilon(h) \to 0$, kun $h \to \bar{0}$. Jos A on avoin ja f on differentioituva x_0:ssa kaikilla $x_0 \in A$, niin f on *differentioituva* (joukossa A).

Kehitelmässä (3.37) merkitään usein $h_i = dx_i$ (x:n muutos) ja lukuja $a_i h_i = a_i dx_i$ kutsutaan *f:n osittaisdifferentiaaleiksi x_0:ssa* sekä lukua $a \cdot h = a_1 dx_1 + \ldots + a_n dx_n$ *f:n kokonaisdifferentiaaliksi pisteessä x_0*. On oltava $a = \nabla f(x_0)$:

3.38 Lause. *Olkoon f differentioituva x_0:ssa kuten (3.37):ssä. Tällöin f on jatkuva x_0:ssa ja derivoituva x_0:ssa. Lisäksi kehitelmä (3.37) on muotoa*

$$f(x_0+h) - f(x_0) = \nabla f(x_0) \cdot h + \|h\|\epsilon(h),$$

ts. $a = \nabla f(x_0)$ ja $a_i = D_i f(x_0)$ $\forall i \in \{1, \ldots, n\}$.

Todistus. Koska

$$\begin{aligned} |f(x_0+h) - f(x_0)| &= |a\cdot h + \|h\|\epsilon(h)| \le |a \cdot h| + \|h\||\epsilon(h)| \\ &\overset{3.2(\mathrm{i})}{\le} \|a\|\|h\| + \|h\||\epsilon(h)| = \|h\|(\|a\| + \epsilon(h)) \underset{h\to\bar{0}}{\longrightarrow} 0, \end{aligned}$$

niin f on jatkuva x_0:ssa.

Sen osoittaminen, että $D_i f(x_0) = a_i$ jätetään harjoitustehtäväksi lukijalle. (Vihje: Valitse $h = (0, \ldots, 0, h_i, 0, \ldots, 0)$ ja käytä erotusosamäärää.) □

Pelkkä osittaisderivaattojen olemassaolo ei takaa differentioituvuutta (ks. huomautus 3.31 ja lause 3.38), mutta jatkuva derivoituvuus takaa sen:

3.39 Lause. *Jos $f : A \to \mathbb{R}^n$ ($A \subset \mathbb{R}^n$) on jatkuvasti derivoituva A:n sisäpisteessä x_0, f on differentioituva x_0:ssa.*

Todistus. Sivuutetaan. $\square$

Käsitteiden jatkuva, derivoituva, jatkuvasti derivoituva ja differentioituva välillä on siis seuraavat implikaatiot:

$$\begin{array}{ccc}
\text{JATKUVASTI DERIVOITUVA} & & \\
\Downarrow & & \\
\text{DIFFERENTIOITUVA} & \Longrightarrow & \text{JATKUVA} \\
\Downarrow & & \\
\text{DERIVOITUVA} & &
\end{array}$$

Huomautus. Mitään yllä olevista implikaatioista ei voi yleisesti kääntää.

Huomautus. Differentiaalikehitelmässä (3.37) $a = \nabla f(x_0)$ esiintyy samassa roolissa kuin $f'(x_0)$ yhden muuttujan funktioille. Tämän mukaisesti käytetäänkin joskus gradientille myös merkintää

$$\nabla f(x_0) = \operatorname{grad} f(x_0) = f'(x_0).$$

Differentioituvalla funktiolla $f : A \to \mathbb{R}$ ($A \subset \mathbb{R}^n$) on *suunnatut derivaatat*

$$D_{\bar{a}} f(x_0) = \lim_{t \to 0} \frac{f(x_0 + t\bar{a}^\circ) - f(x_0)}{t}, \quad \bar{a}^\circ = \frac{\bar{a}}{\|\bar{a}\|} \tag{3.40}$$

myös muihin suuntiin $\bar{a} \neq \bar{0}$ kuin koordinaattiakselien suuntiin
$\bar{a} = e_i = (0, \ldots, 0, \overset{i:\text{s}}{1}, 0, \ldots, 0)$ (joihin ne ovat osittaisderivaatat $D_i f(x_0)$):

3.41 Lause. *Jos $A \subset \mathbb{R}^n$ ja $f : A \to \mathbb{R}$ on differentioituva A:n sisäpisteessä x_0, niin*

$$D_{\bar{a}} f(x_0) = \nabla f(x_0) \cdot \bar{a}^\circ = \nabla f(x_0) \cdot \frac{\bar{a}}{\|\bar{a}\|}$$

Todistus. Valitaan differentiaalikehitelmässä (3.37) muutosvektoriksi h $\bar{a}$:n suuntainen vektori $h = t\bar{a}^\circ$:

$$\begin{aligned}
&f(x_0 + t\bar{a}^\circ) - f(x_0) = \nabla f(x_0) \cdot (t\bar{a}^\circ) + |t| \epsilon(t\bar{a}^\circ) \;\Rightarrow \\
&\frac{f(x_0 + t\bar{a}^\circ) - f(x_0)}{t} = \nabla f(x_0) \cdot \bar{a}^\circ + \underbrace{\frac{|t|}{t}}_{\pm 1} \epsilon(t\bar{a}^\circ) \xrightarrow[t \to 0]{} \nabla f(x_0) \cdot \bar{a}^\circ. \quad \square
\end{aligned}$$

Suunnattu derivaatta $D_{\bar{a}} f(x_0)$ kuvaa f:n muutosnopeutta siirryttäessä x_0:sta suuntaan $\bar{a}$. Schwarzin epäyhtälön 3.2(i) avulla maksimaalinen kasvunopeus tulee gradientin suuntaan ja maksimaalinen vähenemisnopeus vastakkaiseen suuntaan $-\nabla f(x_0)$:

3.42 Seuraus. *Olkoon f differentioituva pisteessä x_0. Tällöin suunnattu derivaatta $D_{\bar{a}}f(x_0)$ on*

$$\begin{aligned} \textit{suurin} &= \|\nabla f(x_0)\| \;\Leftrightarrow\; \bar{a} \upuparrows \nabla f(x_0) \\ \textit{pienin} &= -\|\nabla f(x_0)\| \;\Leftrightarrow\; \bar{a} \updownarrows \nabla f(x_0) \end{aligned}$$

Lisäksi suunnattu derivaatta on 0 *gradienttia vastaan kohtisuoriin suuntiin.* □

Seurauksen 3.42 viimeisen toteamuksen sisältö voidaan ilmaista sanonnalla "Funktion muutosnopeus gradienttia vastaan kohtisuoriin suuntiin on nolla". Geometrisesti $\nabla f(x_0)$ on pisteessä $x_0 \in A$ kohtisuorassa $(n-1)$-ulotteista $\mathbb{R}^n$:n *tasa-arvopintaa* $f(x) = f(x_0)$ vastaan, kun f on differentioituva ja $\nabla f(x_0) \neq 0$. Tapauksessa $n = 2$ kyseinen tasa-arvopinta on käyrä $f(x,y) = f(x_0,y_0)$ pienessä (x_0,y_0):n ympäristössä, tapauksessa $n = 3$ se on pinta $f(x,y,z) = f(x_0,y_0,z_0)$ pienessä (x_0,y_0,z_0):n ympäristössä. Tämän pinnan pisteeseen $(x_0,y_0,z_0) = \bar{r}_0$ piirretyn *tangenttitason* yhtälö on siis

$$(3.43) \qquad \begin{cases} \nabla f(x_0,y_0,z_0)\cdot(x-x_0,y-y_0,z-z_0) = 0 \quad \text{eli} \\ D_1f(\bar{r}_0)(x-x_0) + D_2f(\bar{r}_0)(y-y_0) + D_3f(\bar{r}_0)(z-z_0) = 0. \end{cases}$$

3.44 Esimerkki. Harjoitellaan esitettyjä käsitteitä tarkastelemalla vaikkapa funktiota $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = 2x^2 + 3y^2 + 4z^2$ (aina ≥ 0), jonka tasa-arvojoukko $\{\bar{r} \mid f(\bar{r}) = c\}$ on

i) $\emptyset$, jos $c < 0$,

ii) $\{\bar{0}\}$, jos $c = 0$ ($f(\bar{r}_0) = 0 \;\Leftrightarrow\; \bar{r}_0 = \bar{0}$).

HUOM! tapauksissa i) ja ii) kyseessä ei ole pinta.

iii) ellipsoidi, jos $c > 0$ ($c = f(\bar{r}_0) > 0 \;\Rightarrow\; \nabla f(\bar{r}_0) \neq 0$ ja tulee pinta).

Selvästi $D_1f = 4x$, $D_2f = 6y$ ja $D_3f = 8z$ ovat jatkuvia, joten f on jatkuvasti derivoituva ($f \in C^1(\mathbb{R}^3)$, jopa $f \in C^\infty(\mathbb{R}^3)$) ja siten myös differentioituva.

Esim. pisteeseen $\bar{r}_0 = (1,1,-1)$ kuuluva f:n gradientti on $\nabla f(1,1,-1) = (4,6,-8)$ ja differentiaalikehitelmä

$$(*) \qquad \begin{cases} f(1+h_1, 1+h_2, -1+h_3) - f(1,1,-1) = \\ = \nabla f(1,1,-1)\cdot(h_1,h_2,h_3) + \sqrt{h_1^2+h_2^2+h_3^2}\,\epsilon(h) = \\ = 4h_1 + 6h_2 - 8h_3 + \|h\|\epsilon(h) \end{cases}$$

Tässä funktion $\epsilon(h)$ voisi ratkaistakin yhtälöstä $(*)$, mutta tiedämme myös muuten, että $\lim\limits_{h\to\bar{0}} \epsilon(h) = 0$. Toisella (perinteisellä) tavalla kirjoitettuna f:n kokonaisdifferentiaali pisteessä $\bar{r}_0 = (1,1,-1)$ on $4dx + 6dy - 8dz$, osittaisdifferentiaaliensa summa. Kokonaisdifferentiaali *df* on "pienillä" muutosvektoreilla $(dx,dy,dz) = (h_1,h_2,h_3) = \bar{h}$ hyvä approksimaatio f:n muutokselle $\Delta f = f(\bar{r}_0 + \bar{h}) - f(\bar{r}_0)$

$$\Delta f \approx df = 4dx + 6dy - 8dz.$$

Pisteeseen $\bar{r}_0$ liittyy f:n tasa-arvopinta

$$f(\bar{r}) = f(\bar{r}_0) \;\Leftrightarrow\; 2x^2 + 3y^2 + 4z^2 = 9 \;\Leftrightarrow\; \frac{x^2}{\left(\frac{3}{\sqrt{2}}\right)^2} + \frac{y^2}{\left(\sqrt{3}\right)^2} + \frac{z^2}{\left(\frac{3}{2}\right)^2} = 1$$

(origokeskinen ellipsoidi, puoliakselit $\frac{3}{\sqrt{2}}\bar{\imath}$, $\sqrt{3}\,\bar{\jmath}$, $\frac{3}{2}\,\bar{k}$). Tämän pinnan pisteeseen $\bar{r}_0 = (1,1,-1)$ piirretty tangenttitaso on siis kohtisuorassa vektoria $\nabla f(\bar{r}_0) = (4,6,-8)$ vastaan ja sen yhtälö on

$$\nabla f(\bar{r}_0) \cdot (x-1, y-1, z+1) = 0 \;\Leftrightarrow\; 4x + 6y - 8z - 4 - 6 - 8 = 0 \;\Leftrightarrow$$
$$\Leftrightarrow\; \underline{\underline{2x + 3y - 4z = 9}} \quad \text{(ellipsoidin tangenttitaso pisteessä } (1,1,-1)\text{)}$$

Pisteessä $(1,1,-1)$ funktio f siis pysyy likimain muuttumattomana tämän tangenttitason vektorien suuntaan, kasvaa nopeimmin suuntaan $\nabla f(1,1,-1) = (4,6,-8)$ ja vähenee nopeimmin suuntaan $(-4,-6,8)$. Esimerkiksi suuntaan $3\bar{\imath} + 4\bar{\jmath} = (3,4,0) = \bar{a}$ f:n suunnattu derivaatta $(1,1,-1)$:ssä on

$$D_{(3,4,0)} f(1,1,-1) = \nabla f(1,1,-1) \cdot \frac{(3,4,0)}{\sqrt{3^2 + 4^2 + 0^2}} =$$
$$= (4,6,-8) \cdot \left(\frac{3}{5}, \frac{4}{5}, 0\right) = \frac{12}{5} + \frac{24}{5} = \frac{36}{5} = 7\frac{1}{5}$$

ja tämä on pienempi kuin maksimaalinen suunnattu derivaatta

$$\|\nabla f(1,1,-1)\| = \sqrt{4^2 + 6^2 + (-8)^2} = \sqrt{116}.$$

Vektorifunktion $f : A \to \mathbb{R}^p$ ($A \subset \mathbb{R}^n$) differentioituvuus ja differentiaalit määritellään koordinaateittain:

$$f \text{ on differentioituva } \;\Leftrightarrow\; f_i = \mathrm{pr}_i \circ f \text{ on differentioituva } \forall i \in \{1, \ldots, p\}$$

3.45 Esimerkki. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x^2y + x, xy + x^2)$. Tällöin
$f_1(x,y) = x^2y + x$ ja $f_2(x,y) = xy + x^2$, joten
$D_1 f_1 = 2xy + 1$, $D_2 f_1 = x^2$, $D_1 f_2 = y + 2x$ ja $D_2 f_2 = x$ ja edelleen

$$\nabla f_1(2,3) = (2 \cdot 2 \cdot 3 + 1, 2^2) = (13,4) \quad \text{ja} \quad \nabla f_2(2,3) = (3 + 2 \cdot 2, 2) = (7,2)$$

ja f:n kokonaisdifferentiaali $(2,3)$:ssa on

$$df = \big(\nabla f_1(2,3) \cdot (dx, dy), \nabla f_2(2,3) \cdot (dx, dy)\big) = (13dx + 4dy, 7dx + 2dy).$$

Differentioituvien kuvausten yhdistetty kuvaus on differentioituva

3.46 Lause (Ketjusääntö). *Olkoot $A \subset \mathbb{R}^m$ avoin, $g = (g_1, \ldots, g_n) : A \to \mathbb{R}^n$, $g(A) \subset B \subset \mathbb{R}^n$, B avoin ja $f : B \to \mathbb{R}$.*

$$\begin{tikzcd}
\mathbb{R}^m & \mathbb{R}^n & \\
\cup & \cup & \\
A \arrow[r, "g = (g_1, \ldots, g_n)"] & B \arrow[r, "f"] & \mathbb{R} \\
x \arrow[r, mapsto] & g(x) \arrow[r, mapsto] & (f \circ g)(x)
\end{tikzcd}$$

Jos g ja f ovat differentioituvia, niin yhdistetty kuvaus $f \circ g$ on differentioituva ja $\forall x \in A$ on

$$D_i(f \circ g)(x) = \sum_{j=1}^{n} D_j f(g(x)) D_i g_j(x) \quad (i = 1, \ldots, m)$$

(eli vanhanaikaisesti "ketjun" muotoon kirjoitettuna

$$\frac{\partial (f \circ g)}{\partial x_i} = \frac{\partial f}{\partial g_1} \frac{\partial g_1}{\partial x_i} + \frac{\partial f}{\partial g_2} \frac{\partial g_2}{\partial x_i} + \ldots + \frac{\partial f}{\partial g_n} \frac{\partial g_n}{\partial x_i} \Big).$$

Todistus. Sivuutetaan. $\square$

Ketjusäännöllä on käyttöä lähinnä teoreettisissa yhteyksissä, koska käytännössä yhdistetyn funktion derivoiminen yleensä sujuu nopeiten muodostamalla ensin yhdistetyn kuvauksen lauseke ja derivoimalla suoraan sitä.

3.47 Esimerkki. i) Derivoi $w(x, y) = (x^2 + 2y - 3)^3$ suoraan ja ketjusäännön avulla.

Ratkaisu. Suora derivointi antaa

$$\begin{aligned}
D_x w &= 3 \cdot (x^2 + 2y - 3)^2 \cdot 2x = 6x(x^2 + 2y - 3)^2 \\
D_y w &= 3 \cdot (x^2 + 2y - 3)^2 \cdot 2 = 6(x^2 + 2y - 3)^2
\end{aligned}$$

Ketjusäännön soveltamiseksi tulkitaan w yhdistetyksi kuvaukseksi $w = f \circ g$, $g(x, y) = x^2 + 2y - 3$, $f(t) = t^3$, jolloin $g : \mathbb{R}^2 \to \mathbb{R}$ ja $f : \mathbb{R} \to \mathbb{R}$ ja $m = 2$, $n = 1$ ketjusäännössä. Nyt $g = g_1$ ja saamme

$$\begin{aligned}
D_1 w(x, y) &= f'(g(x, y)) D_1 g(x, y) = 3 \cdot (x^2 + 2y - 3)^2 \cdot 2x \quad \text{ja} \\
D_2 w(x, y) &= f'(g(x, y)) D_2 g(x, y) = 3 \cdot (x^2 + 2y - 3)^2 \cdot 2,
\end{aligned}$$

siis sama tulos kuin suoraan derivoitaessa.

ii) Laske ketjusäännöllä $D_3(f \circ g)$, kun $g : \mathbb{R}^3 \to \mathbb{R}^2$ on kuvaus $g = (g_1, g_2)$, $g_1(x, y, z) = x^2 e^{-z} + yz^3$, $g_2(x, y, z) = yz^2 + \dfrac{z}{x^2 + 1}$ ja $f : \mathbb{R}^2 \to \mathbb{R}$ on kuvaus $f(x, y) = xy$.

Ratkaisu. Nyt $m = 3$, $n = 2$, $D_1 f = y$, $D_2 f = x$ ja saamme

$$
\begin{aligned}
D_3(f \circ g)(x,y,z) &= \sum_{j=1}^{2} D_j f(g(x,y,z)) D_3 g_j(x,y,z) = \\
&= D_1 f\left(x^2e^{-z} + yz^3, yz^2 + \frac{z}{x^2+1}\right) D_3(x^2e^{-z} + yz^3) \\
&\quad + D_2 f\left(x^2e^{-z} + yz^3, yz^2 + \frac{z}{x^2+1}\right) D_3\left(yz^2 + \frac{z}{x^2+1}\right) = \\
&= \left(yz^2 + \frac{z}{x^2+1}\right) \cdot (-x^2e^{-z} + 3yz^2) + (x^2e^{-z} + yz^3) \cdot \left(2yz + \frac{1}{x^2+1}\right)
\end{aligned}
$$

Sama tulos saadaan suoralla derivoinnilla yhdistetyn kuvauksen $f \circ g$ lausekkeesta

$$
(f \circ g)(x,y,z) = (x^2e^{-z} + yz^3) \cdot \left(yz^2 + \frac{z}{x^2+1}\right).
$$

Ketjusäännöstä seuraa heti, että jatkuvasti derivoituvien funktioiden yhdistetty kuvaus on jatkuvasti derivoituva.

3.48 Lause. *Olkoot $A \subset \mathbb{R}^m$ avoin, $g : A \to \mathbb{R}^n$ jatkuvasti derivoituva, $B \subset \mathbb{R}^n$ avoin, $g(A) \subset B$ ja $f : B \to \mathbb{R}$ jatkuvasti derivoituva. Tällöin yhdistetty kuvaus $f \circ g : A \to \mathbb{R}$ on jatkuvasti derivoituva.*

Todistus. Ketjusäännön nojalla

$$
D_i(f \circ g) = \sum_{j=1}^{n} \big((D_j f) \circ g\big) \cdot (D_i g_j)
$$

on jatkuva kaikilla $i \in \{1, \ldots, m\}$ koska $D_j f$ ja $D_i g_j$ ovat jatkuvia ja g on jatkuva. $\square$

Implisiittifunktiot

Tarkastellaan kysymystä, millä ehdoilla esim. yhtälö $F(x,y) = 0$ määrittelee ainakin paikallisesti (x_0, y_0):n $\big(F(x_0,y_0) = 0\big)$ ympäristössä y:n x:n funktiona $y = y(x)$ ja tämän funktion derivaatan $y'(x)$ laskemista ilman tietoa funktion lausekkeesta. Yleisemmin x voi tässä olla n-vektori $x = (x_1, \ldots, x_n)$, jolloin F on $(n+1)$:n muuttujan funktio ja y on n:n muuttujan funktio ja y:n derivaatat osittaisderivaattoja muuttujien $x_1, \ldots, x_n$ suhteen. On selvää, että jotain rajoittavia ehtoja F:stä tarvitaan, sillä jos esim. F on vakiofunktio $F(x,y) \equiv 0$, ei y ratkea x:n funktiona.

3.49 Esimerkki. Yhtälö $F(x,y) = x^2 + y^2 - 1 = 0$ määrittelee yksikköympyrän C: $x^2 + y^2 = 1$ pisteen (x_0, y_0) ympäristössä y:n x:n funktiona, jos $y_0 \neq 0$:

Jos $y_0 > 0$, $y = +\sqrt{1-x^2}$ C:ssä lähellä (x_0, y_0):aa.

Jos $y_0 < 0$, $y = -\sqrt{1-x^2}$ C:ssä lähellä (x_0, y_0):aa.

Pisteiden $(1,0)$ ja $(-1,0)$ ympäristössä y ei ole x:n funktio, mutta x on kyllä y:n funktio:

$x = \sqrt{1-y^2}$ $(1,0)$:n ympäristössä C:ssä.

$x = -\sqrt{1-y^2}$ $(-1,0)$:n ympäristössä C:ssä.

Annamme heti yleisen tuloksen *implisiittifunktion* olemassaolosta ja osittaiderivaatoista ilman todistusta:

3.50 Lause. *Olkoon $A \subset \mathbb{R}^{n+1}$ avoin, $F : A \to \mathbb{R}$ jatkuvasti derivoituva, $(x_0, y_0) = (x_{01}, \ldots, x_{0n}, y_0) \in A$, $F(x_0, y_0) = 0$ ja*

$$D_{n+1}F(x_0, y_0) = \frac{\partial F(x_0, y_0)}{\partial y} \neq 0.$$

Tällöin on olemassa (x_0, y_0)-keskinen $(n+1)$-ulotteinen suorakulmainen särmiö $S = T \times [y_0 - \delta, y_0 + \delta]$, missä $\delta > 0$ ja T on x_0-keskinen n-ulotteinen särmiö, $T = [x_{01} - \delta_1, x_{01} + \delta_1] \times \ldots \times [x_{0n} - \delta_n, x_{0n} + \delta_n]$, niin että $S \subset A$ ja joukossa S yhtälö $F(x,y) = F(x_1, \ldots, x_n, y) = 0$ määrittelee y:n muuttujien $x_1, \ldots, x_n$ funktiona $y = f(x_1, \ldots, x_n) : T \to \mathbb{R}$. Siis

$$(x,y) \in S \quad \text{ja} \quad F(x,y) = F(x_1 \ldots, x_n, y) = 0 \;\Leftrightarrow\; x \in T \quad \text{ja} \quad y = f(x_1, \ldots, x_n).$$

Saatu implisiittifunktio $y = f(x) : T \to \mathbb{R}$ on jatkuvasti derivoituva ja

$$D_i f(x) = -\frac{D_i F(x, y(x))}{D_{n+1}F(x, y(x))} \quad \forall x \in T, \; (i = 1, \ldots, n). \quad \Box$$

3.51 Esimerkki. (i) Olkoon $F(x,y) = x^2 + y^2 - 1$, jolloin F:n nollajoukko on yksikköympyrä

$$C = F^{-1}(0) = \{(x,y) \mid x^2 + y^2 = 1\}.$$

Nyt $F : \mathbb{R}^2 \to \mathbb{R}$ on jatkuvasti derivoituva ja

$$D_2F(x,y) = \frac{\partial F(x,y)}{\partial y} = 2y \neq 0, \quad \text{kun} \quad y \neq 0.$$

Jos siis $(x_0, y_0) \in C$ ja $y_0 \neq 0$, lause 3.50 (jossa nyt $n = 1$) takaa implisiittifunktion $y = f(x) = y(x)$ olemassaolon pienellä x_0-keskisellä välillä $T = [x_0 - \delta_1, x_0 + \delta_1]$ (nyt $x_0 = x_{01}$). Tässä tapauksessa implisiittifunktio f saadaan myös eksplisiittisesti ratkaistua: $f(x) = \pm\sqrt{1-x^2}$ y_0:n merkin mukaan valittavalla etumerkillä (ks. esimerkki 3.49). Sen derivaataksi saadaan esim. tapauksessa $y_0 > 0$ suoralla derivoinnilla

$$f'(x) = D\left(+\sqrt{1-x^2}\right) = \frac{-x}{\sqrt{1-x^2}} = \frac{-x}{y(x)}$$

tai lauseen 3.50 kaavalla (jossa $n = 1$, $D_1 = D$)

$$f'(x) = D_1 f(x) = -\frac{D_1 F(x, y(x))}{D_2 F(x, y(x))} = -\frac{2x}{2y(x)}$$

eli sama tulos. (Lisähuomio: Kaava $f'(x) = -\frac{x}{y}$ pätee myös, kun $y_0 < 0$.)

Käytännössä implisiittifunktion derivointi suoritetaan yleensä ns. *implisiittisenä derivointina* ajatellen tiedetyksi, että $y = y(x)$ ja derivoimalla tämän mukaisesti. Esim. äskeinen lasku sujuisi implisiittisellä derivoinnilla seuraavasti:

Oletetaan, että yhtälö $x^2 + y^2 = 1$ määrittelee (x_0, y_0):n ympäristössä y:n x:n funktiona $y = y(x)$. Tällöin

$$x^2 + \big(y(x)\big)^2 = 1 \Rightarrow D\left(x^2 + \big(y(x)\big)^2\right) = 0 \Rightarrow 2x + 2y(x)y'(x) = 0 \Rightarrow y'(x) = \frac{-x}{y(x)}$$

(ja viimeksi saadusta muodosta nähdään, että tapauksessa $y_0 = 0$ tulisi luultavasti ongelmia).

(ii) Osoita, että yhtälö $xyz - \ln(zy) + 2 = 0$ määrittelee pisteen $(-2, 1, 1)$ ympäristössä z:n x:n ja y:n funktiona $z = z(x, y)$. Laske $D_x z(-2, 1)$ ja $D_y z(-2, 1)$.

Ratkaisu. Merkitään $A = \{(x, y, z) \mid z > 0,\ y > 0\}$ ja $F(x, y, z) = xyz - \ln(zy) + 2$, jolloin $F(-2, 1, 1) = -2 - \ln(1 \cdot 1) + 2 = 0$ ja $F : A \to \mathbb{R}$ on jatkuvasti derivoituva, sillä $D_1 F = yz$, $D_2 F = xz - \frac{1}{y}$ ja $D_3 F = xy - \frac{1}{z}$ ovat A:ssa jatkuvia funktioita. Koska

$$D_3 F(-2, 1, 1) = -2 \cdot 1 - \frac{1}{1} = -3 \neq 0,$$

lauseesta 3.50 seuraa, että yhtälö $F(x, y, z) = 0$ määrittelee pisteen $(-2, 1, 1)$ ympäristössä z:n funktiona $z = z(x, y)$ ja

$$D_x z = D_1 z = -\frac{D_1 F\big(x, y, z(x, y)\big)}{D_3 F\big(x, y, z(x, y)\big)} = -\frac{yz}{xy - \frac{1}{z}}$$

$$D_y z = D_2 z = -\frac{D_2 F\big(x, y, z(x, y)\big)}{D_3 F\big(x, y, z(x, y)\big)} = -\frac{xz - \frac{1}{y}}{xy - \frac{1}{z}}.$$

Kun $(x, y) = (-2, 1)$ on $z = z(x, y) = 1$ ja siten

$$\underline{\underline{D_x z(-2, 1)}} = -\frac{1 \cdot 1}{-2 \cdot 1 - \frac{1}{1}} = \underline{\underline{\frac{1}{3}}} \quad \text{ja}$$

$$\underline{\underline{D_y z(-2, 1)}} = -\frac{-2 \cdot 1 - \frac{1}{1}}{-2 \cdot 1 - \frac{1}{1}} = \underline{\underline{-1}}.$$

Johdetaan samat tulokset implisiittisellä derivoinnilla:

$$
\begin{aligned}
& z = z(x,y) \quad \text{ja} \quad F(x,y,z) = 0 \\
& \Rightarrow \; xyz(x,y) - \ln\bigl(z(x,y)y\bigr) + 2 = 0 \\
& \Rightarrow \; D_x(xyz - \ln(zy) + 2) = 0 \quad \text{ja} \quad D_y(xyz - \ln(zy) + 2) = 0 \\
& \Rightarrow \; yz + xyD_xz - \frac{D_xz}{z} = 0 \quad \text{ja} \quad xz + xyD_yz - \frac{(z + yD_yz)}{zy} = 0 \\
& \Rightarrow \; D_xz = -\frac{yz}{xy - \frac{1}{z}} \quad \text{ja} \quad xz - \frac{1}{y} + \left(xy - \frac{1}{z}\right) D_yz = 0 \\
& \Rightarrow \; D_xz = -\frac{yz}{xy - \frac{1}{z}} \quad \text{ja} \quad D_yz = -\frac{xz - \frac{1}{y}}{xy - \frac{1}{z}},
\end{aligned}
$$

missä siis $z = z(x,y)$.

3.52 Huomautus. (i) Samanlaista teoriaa voi tietysti soveltaa myös esim. muuttujan y ratkaisemiseen yhtälöstä $F(x,y,z) = 0$: Jos F on jatkuvasti derivoituva ja $D_2F(x_0,y_0,z_0) \neq 0$ ja $F(x_0,y_0,z_0) = 0$, niin y ratkeaa lokaalisti x.n ja z:n funktiona ja

$$D_xy(x,z) = -\frac{D_xF(x,y(x,z),z)}{D_yF(x,y(x,z),z)} \quad \text{ja} \quad \text{vast. } D_zy = -\frac{D_zF}{D_yF}$$

(ii) Yleisemmin implisiittifunktioiden teoriaa voi soveltaa yhtälöryhmiin ratkaisemalla muuttujia $y_1, \ldots, y_m$ muuttujien $x_1, \ldots, x_n$ avulla $(m+n)$:n tuntemattoman m:n yhtälön yhtälöryhmästä. Lyhyt esitys löytyy esim. Timo Patovaaran monisteesta.

(iii) Jos $f : A \to \mathbb{R}$ ($A \subset \mathbb{R}^n$ avoin) on jatkuvasti derivoituva ($f \in C^1(A)$) ja jos $\nabla f(x_0) \neq \bar{0}$, niin yllä olevan nojalla ne muuttujista x_i ($i = 1, \ldots, n$), joilla $D_if(x_0) \neq 0$, ratkeavat muiden funktioina vastaavan tasa-arvopinnan yhtälöstä $f(x) = f(x_0)$. Tämä takaa mm. sen, että tämä pinta "leikkaa itseään" lähellä pistettä x_0 ja tuloksena on x_0:n ympäristössä "siisti" $(n-1)$-ulotteinen pinta, jolla on $\nabla f(x_0)$:aa vastaan kohtisuora $(n-1)$-ulotteinen tangenttitaso. Sivuutamme tähän johtavat tarkastelut liian mutkikkaina peruskurssille.

Neliömuodot ja niiden tyyppi

Esimerkissä 3.35 on jo käsitelty n:n muuttujan $x_1, \ldots, x_n$ *neliömuotoa*

$$Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = x \cdot f(x) = x^TAx, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \tag{3.53}$$

missä $\cdot$ on pistetulo ja $f : \mathbb{R}^n \to \mathbb{R}^n$ on se lineaarikuvaus, jonka matriisi on $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Siis $f(x) = \big(f_1(x), \dots, f_n(x)\big)$, missä koordinaattifunktiolla $f_i(x)$ on lauseke

$$f_i(x) = \sum_{j=1}^{n} a_{ij}x_j \quad (i = 1, \dots, n).$$

Tässä A vaaditaan (yleensä) *symmetriseksi*, $a_{ij} = a_{ji}\ \forall i, j$, jolloin f on ns. *itseadjungoitu lineaarikuvaus* ja neliömuodot Q, itseadjungoidut lineaarikuvaukset f ja symmetriset $n \times n$-matriisit A vastaavat toisiaan kääntäen yksikäsitteisesti kaavan (3.53) kautta. Neliömuodossa $Q(x)$ yhdistetään lisäksi yleensä termit $a_{ij}x_ix_j$ ja $a_{ji}x_jx_i$, jolloin

$$Q(x) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{i<j} a_{ij}x_ix_j. \tag{3.54}$$

3.55 Esimerkki.

$$Q(x,y) = x^2 + xy + y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (x,y)\cdot f(x,y),$$

missä $f(x,y) = \big(f_1(x,y), f_2(x,y)\big) = (x + \frac{1}{2}y, \frac{1}{2}x + y)$. Tällöin todella

$$(x,y)\cdot f(x,y) = x(x + \frac{1}{2}y) + y(\frac{1}{2}x + y) = x^2 + xy + y^2 = Q(x,y)$$

ja matriisikertolaskulla

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} x + \frac{1}{2}y \\ \frac{1}{2}x + y \end{bmatrix} = \begin{bmatrix} x^2 + xy + y^2 \end{bmatrix} =$$

$$= \underbrace{\begin{bmatrix} Q(x,y) \end{bmatrix}}_{1\times 1\text{-matriisi}} \text{"} = \text{"} \underbrace{Q(x,y)}_{\text{luku}}$$

ii) Neliömuotoon

$$x_1^2 + x_3^2 + 2x_1x_2 + 3x_4^2 + x_3x_4 = Q(x_1,x_2,x_3,x_4) \quad (Q : \mathbb{R}^4 \to \mathbb{R})$$

liittyvä symmetrinen (4×4)-matriisi on

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 3 \end{bmatrix}$$

a_{ii} on x_i^2:n kerroin, $a_{ij} = a_{ji}$ on x_ix_j:n kertoimen puolikas (esim. $a_{34} = \frac{1}{2}$, koska x_3x_4:n kerroin on $= 1$ ja $a_{42} = 0$ koska termiä x_2x_4 ei ole.).

3.56 Määritelmä. Olkoon $Q : \mathbb{R}^n \to \mathbb{R}$ symmetrisen $A \in \mathbb{R}^n \times \mathbb{R}^n$ indusoima neliömuoto: $Q(x) = x^T Ax$.

(i) Jos $Q(x) > 0$ $\forall x \in \mathbb{R}^n \setminus \{\bar{0}\}$, Q ja A ovat *positiivisesti definiittejä*; merk. $A > 0$.

(ii) Jos $Q(x) < 0$ $\forall x \in \mathbb{R}^n \setminus \{\bar{0}\}$, Q ja A ovat *negatiivisesti definiittejä*; merk. $A < 0$.

(iii) Jos $Q(x) \geq 0$ $\forall x \in \mathbb{R}^n$, Q ja A ovat *positiivisesti semidefiniittejä*; merk. $A \geq 0$.

(iv) Jos $Q(x) \leq 0$ $\forall x \in \mathbb{R}^n$, Q ja A ovat *negatiivisesti semidefiniittejä*; merk. $A \leq 0$.

(v) Jos $\exists\, x, y \in \mathbb{R}^n$ s.e. $Q(x) < 0$ ja $Q(y) > 0$, niin Q ja A ovat *indefiniittejä*.

3.57 Lause. *Olkoon $Q : \mathbb{R}^n \to \mathbb{R}$ kuten määritelmässä 3.56*

(i) A on positiivisesti definiitti $\Leftrightarrow$ A:n johtavat alideterminantit

$$d_1 = a_{11}, \quad d_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad d_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ \ldots, \quad d_n = \det A$$

ovat kaikki positiivisia.

(ii) A on negatiivisesti definiitti $\Leftrightarrow$ $-A$ on positiivisesti definiitti

(iii) A on positiivisesti semidefiniitti $\Leftrightarrow$ $\det B \geq 0$ aina kun B on sellainen matriisi, joka saadaan jättämällä A:sta pois r riviä ja vastaavat r saraketta ($r = 0, \ldots, n-1$).

(iv) A on negatiivisesti semidefiniitti $\Leftrightarrow$ $-A$ on positiivisesti semidefiniitti.

(v) A on positiivisesti (negatiivisesti) definiitti $\Leftrightarrow$ A:n ominaisarvot ovat kaikki positiivisia (negatiivisia).

(vi) A on positiivisesti (negatiivisesti) semidefiniitti $\Leftrightarrow$ A:n ominaisarvot ovat kaikki ≥ 0 (≤ 0).

Todistus. Väitteet (v) ja (vi) on tarkoitettu *lisätiedoiksi* kurssin Lineaarialgebra ja matriisilaskenta II tiedot hallitsevilla. Ne seuraavat heti diagonalisoimalla Q. A:n ominaisvektoreiden ortonormaalissa kannassa $(v_1, \ldots, v_n)$ Q:lla on näet esitys

$$Q(y_1 v_1 + \ldots + y_n v_n) = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2.$$

kun v_i on A:n ominaisarvoon λ_i liittyvä kannan ominaisvektori.

Väitteet (ii) ja (iv) seuraavat heti väitteistä (i) ja (iii). Väite (i) voidaan todistaa induktiolla alideterminantin kertaluvun suhteen. Joudumme sivuuttamaan väitteiden (i) ja (iii) todistukset liian työläinä. Joskus voi olla nopeampaa osoittaa neliömuoto semidefiniitiksi väitteen (vi) avulla kuin väitteen (iii) determinanteilla. (Väite (iii) kuitenkin toimii aina, kun sen sijaan ominaisarvoja ei aina saada ratkaistua.) $\square$

3.58 Esimerkki. (i) $Q(x,y) = x^2 + 2xy + y^2 = (x+y)^2$ on selvästi positiivisesti semidefiniitti, sillä $Q(x,y) \geq 0 \ \forall (x,y) \in \mathbb{R}^2$. Se ei ole definiitti muoto, sillä $Q(1,-1) = 0$ ja $(1,-1) \neq (0,0)$.

(ii) $Q(x,y) = -6xy + y^2$ on indefiniitti, sillä $Q(0,1) = 1 > 0$ ja $Q(1,1) = -6+1 = -5 < 0$.

(iii) $Q(x,y,z) = 3x^2 - 4xy + z^2$. Tällöin vastaava symmetrinen matriisi A on

$$A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Nyt

$$d_1 = 3 > 0, \quad d_2 = \begin{vmatrix} 3 & -2 \\ -2 & 0 \end{vmatrix} = -4 < 0,$$

joten A ei ole positiivisesti definiitti ja koska $-A$:ssa $d_1 = -3 < 0$, myöskään $-A$ ei ole positiivisesti definiitti. Siten A ei ole definiitti muoto. Koska $d_1 = 3 > 0$ ja $d_2 = -4 < 0$, niin A ei ole positiivisesti semidefiniitti (nämä determinantit ovat muotoa B 3.57 (iii):ssä) ja koska $-A$:ssa $d_1 = -3 < 0$, niin A ei ole negatiivisesti semidefiniitti. Johtopäätös: A on indefiniitti. Tämän näkisi suoraankin, sillä esim. $Q(0,0,1) = 1 > 0$ ja $Q(1,1,0) = -1 < 0$.

(iv) $Q(x,y) = x^2 + xy + y^2 = (x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 \geq 0$ ja $Q(x,y) = 0 \ \Leftrightarrow \ x = y = 0$, joten Q on positiivisesti definiitti.

Sama determinanttiehdolla 3.57 (i):

Q:n matriisi $A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$ toteuttaa $d_1 = 1 > 0$ ja $d_2 = \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{vmatrix} = 1 - \dfrac{1}{4} = \dfrac{3}{4} > 0,$

joten 3.57 (i):stä seuraa, että Q on positiivisesti definiitti.

(v) $Q(x,y,z) = x^2 + 2y^2 + 3z^2 + xy + 2yz$. Nyt

$$A = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \text{ toteuttaa } d_1 = 1 > 0, \ d_2 = 1 \cdot 2 - \frac{1}{2} \cdot \frac{1}{2} > 0 \quad \text{ja}$$

$$d_3 = \det A = 1 \cdot \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} - \frac{1}{2} \begin{vmatrix} \frac{1}{2} & 1 \\ 0 & 3 \end{vmatrix} + 0 = 5 - \frac{3}{4} > 0,$$

joten Q ja A ovat positiivisesti definiittejä.

(vi) $-x^2 + 4xy - 4y^2 = -(x-2y)^2 \leq 0$ on negatiivisesti semidefiniitti. Sama determinanttiehdoilla 3.57 (iv ja iii): Nyt

$$-A = -\begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \quad \text{ja}$$

$$\underset{(r=1)}{d_1} = 1 > 0, \quad \underset{(r=0)}{d_2} = 4 - 4 = 0 \geq 0, \quad \underset{(r=1)}{\overbrace{\det[4]}^{=4}} > 0,$$

joten 3.57 (iii):stä seuraa, että $-A$ on positiivisesti semidefiniitti. Siis 3.57 (iv):n nojalla A ja Q ovat negatiivisesti semidefiniittejä.