

# MS-A0207 Differentiaali- ja integraalilaskenta 2 (Chem) Yhteenveto ja esimerkkejä ym., osa I

G. Gripenberg

Aalto-yliopisto

21. tammikuuta 2016

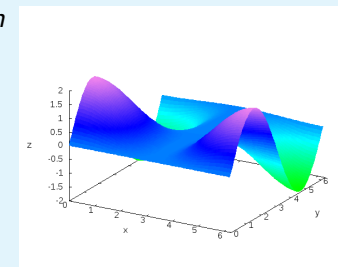
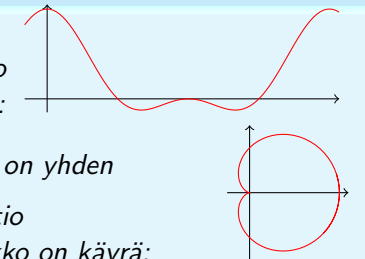
- 1 Usean muuttujan funktiot
- 2 Raja-arvot
- 3 Jatkuvat funktiot
- 4 Osittaisderivaatat
- 5 Derivaatta eli gradientti
- 6 Lineaarinen approksimointi
- 7 Newtonin menetelmä

## 💡 Vektoreista

- $\mathbb{R}$  on reaalilukujen joukko ja  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\}$  on  $xy$ -tason pisteitä.
- Jokaista pistettä tai paria  $(x, y)$  voimme käsitellä vektorina, jonka voimme myös esittää joko pystyvektorina ( $2 \times 1$ -matriisina)  $\begin{bmatrix} x \\ y \end{bmatrix}$ , vaakavektorina ( $1 \times 2$ -matriisina)  $[x \ y]$  tai muodossa  $x\mathbf{i} + y\mathbf{j}$  (missä  $\mathbf{i}$  on positiivisen  $x$ -akselin suuntainen yksikkövektori jne.)
- Jos  $\mathbf{u}$  on vaakavektori  $[u(1) \ u(2)]$  ja  $\mathbf{v}$  on pystyvektori  $\begin{bmatrix} v(1) \\ v(2) \end{bmatrix}$  niin niiden matriisitulo on  $\mathbf{u}\mathbf{v} = u(1)v(1) + u(2)v(2)$ . Jos emme tee eroa vaak- ja pystyvektoriden välillä voimme myös kirjoittaa tätä pistetulona (sisätulona, skalaaritulona)  $\mathbf{u} \cdot \mathbf{v}$ .
- Vektorin pituuden määritelmäksi otamme  $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$ .
- Vektorit  $\mathbf{u}$  ja  $\mathbf{v}$  ovat kohtisuorassa toisiaan vastaan jos  $\mathbf{u} \cdot \mathbf{v} = 0$ .
- $\mathbb{R}^d = \{(x_1, x_2, \dots, x_d) : x_j \in \mathbb{R}, j = 1, \dots, d\}$  ja kun  $\mathbb{R}^2$  korvataan  $\mathbb{R}^d$ :llä niin "ainoastaan" kuvien piirtäminen tulee vaikeammaksi.

## 😊 Funktioista

- $f(x) = (1 + \cos(x)) \cos(x)$  on yhden reaalimuuttujan reaaliarvoinen funktio (eli funktio:  $\mathbb{R} \rightarrow \mathbb{R}$ ) jolla on kuvaaja:
- Funktio  $\mathbf{g}(t) = \begin{bmatrix} (1 + \cos(t)) \cos(t) \\ (1 + \cos(t)) \sin(t) \end{bmatrix}$  on yhden reaalimuuttujan vektoriarvoinen funktio (eli funktio:  $\mathbb{R} \rightarrow \mathbb{R}^2$ ) jonka kuvajoukko on käyrä:
- Funktio  $h(x, y) = (1 + \cos(x)) \sin(y)$  on kahden reaalimuuttujan reaaliarvoinen funktio (eli funktio:  $\mathbb{R}^2 \rightarrow \mathbb{R}$ ), jota voi myös käsitellä yhden vektorimuuttujan funktiona  $h\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = (1 + \cos(x)) \sin(y)$  ja jolla on seuraava kuvaaja:



Huomaa, että  $f$ :n kuvaaja on käyrä, jolla on parametriesitys  $t \mapsto \begin{bmatrix} t \\ (1 + \cos(t)) \cos(t) \end{bmatrix}$ .

## 💡 Usean muuttujan vai vektorimuuttujan funktiot

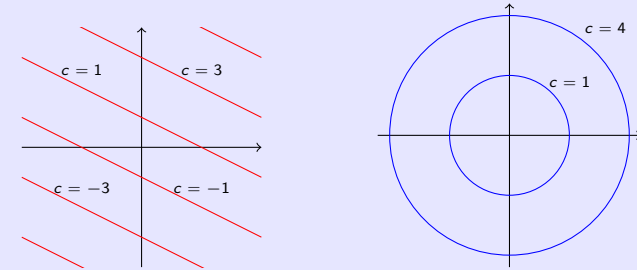
- Funktio  $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}$  on "sääntö" tai "yhteys" (usein mutta ei suinkaan välttämättä, kaava) joka **jokaiseen** määrittelyjoukon  $D$  alkioon  $(x_1, x_2, \dots, x_d)$  liittyy **yksikäsitteisen** "arvon"  $f(x_1, x_2, \dots, x_d) \in \mathbb{R}$ .
- Funktion  $\mathbf{f} : \mathbb{R}^d \rightarrow \mathbb{R}^m$  arvot ovat vektoreita ja vektorin  $\mathbf{f}$  jokainen komponentti on funktio:  $\mathbb{R}^d \rightarrow \mathbb{R}$  eli

$$\mathbf{f}(x_1, \dots, x_d) = \begin{bmatrix} f_1(x_1, \dots, x_d) \\ \vdots \\ f_m(x_1, \dots, x_d) \end{bmatrix}$$

- Usein voi olla hyödyllistä, että sen sijaan että käsittelemme funktiota  $f(x_1, x_2, \dots, x_d)$  usean muuttujan funktiona käsittelemme sitä yhden **vektori**muuttujan  $\mathbf{x}$  funktiona  $f(\mathbf{x})$  missä vektorilla  $\mathbf{x}$  on  $d$  komponenttia.
- Useimmiten otamme funktion argumentit pystyvektorina jolloin funktion derivaatta eli gradientti (kun funktio on reaaliarvoinen) on vaakavektori.

## 💡 Tasa-arvokäyrät

Jos  $f(x, y)$  on kahden muuttujan reaaliarvoinen funktio ja  $c \in \mathbb{R}$  niin joukko  $\{(x, y) : f(x, y) = c\}$  on usein käyrä (tai sitten käyrien unioni) eli funktion **tasa-arvokäyrä**. Kuvio, jossa on piirrettynä monta tällaista tasa-arvokäyrää antaa tietynlaista informaatiota funktiosta. Funktion  $f(x, y) = x + 2y$  tasa-arvokäyrät ovat suoria kun taas funktion  $g(x, y) = x^2 + y^2$  tasa-arvokäyrät ovat ympyröitä:



Kolmen muuttujan tapauksessa saadaan vastaavasti **tasa-arvopintoja**.

## 😊 Muuttujat, parametrit, vakiot

Jos sylinterin muotoisella putkella on pituus  $L$ , sisähalkaisija  $r$ , seinämän paksuus  $h$  ja materiaalin tiheys on  $\rho$  niin putken massa

$$m = \rho \cdot L \cdot \pi((r + h)^2 - r^2).$$

Riippuen tilanteesta voimme tämän kaavan avulla määritellä erilaisia funktioita:

- $m = f(L) = \rho \cdot L \cdot \pi((r + h)^2 - r^2)$  kun pidämme suureita  $\rho$ ,  $r$  ja  $h$  parametreina, eli käsittelemme eripituisia putken pätkiä.
- $m = f(L, r, h, \rho) = \rho \cdot L \cdot \pi((r + h)^2 - r^2)$  missä meillä on neljä muuttujaa.
- $m = f \left( \begin{bmatrix} L \\ r \\ h \end{bmatrix} \right) = \rho \cdot L \cdot \pi((r + h)^2 - r^2)$  missä meillä on yhden vektorimuuttujan funktio ja pidämme  $\rho$ :ta parametrina.

Kaikissa näissä tapauksissa pidämme  $\pi$ :tä vakiona mutta jos  $\pi$ :n paikalle sijoitetaan (epätarkka) likiarvo ja haluamme selvittää mikä on tämän approksimaation vaikutus voi syntyä tilanteita missä  $\pi$ :tä käsitellään muuttujana tai parametrina.

## 😊 Muuttujat, parametrit, vakiot, jatk.

Jos käytämme Matlab/Octavea voimme esimerkiksi kirjoittaa

`rho=7.87`

`f=@(L,r,h) rho*L*pi*((r+h)^2-r^2)`

tai jos käytämme vektoriarargumenttiä

`rho=7.87`

`f=@(X) rho*X(1)*pi*((X(2)+X(3))^2-X(2)^2)`

Huomaa, että jos muutamme  $\rho$ :n arvoa niin meidän pitää toistaa funktion  $f$  määritelmää!

### 💡 Raja-arvo, määritelmä I

$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = L$  jos  $|f(x,y) - L|$  on "pieni" aina kun  $\|(x,y) - (x_0,y_0)\| = \sqrt{(x-x_0)^2 + (y-y_0)^2}$  on "riittävän pieni" ja  $(x,y) \neq (x_0,y_0)$ .

### 💡 Raja-arvo, määritelmä II

$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = L$  jos  $|f(\mathbf{x}) - L|$  on "pieni" kun  $\|\mathbf{x} - \mathbf{x}_0\|$  on "riittävän pieni" ja  $\mathbf{x} \neq \mathbf{x}_0$ .

### 💡 Vektorin pituus

Vektorin  $\mathbf{x}$  pituus, kun sen komponentit ovat  $x_1, x_2, \dots, x_d$ , on tässä

$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$  mutta raja-arvot eivät riipu siitä miten vektorin "pituus" on määritelty kunhan sillä on normaalit "pituuden" ominaisuudet.

### 😊 Raja-arvo, määritelmä III

$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = L$  jos jokaisella  $\epsilon > 0$  on olemassa  $\delta > 0$  siten, että jos  $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$  ja  $\mathbf{x} \in \Omega$  niin pätee  $|f(\mathbf{x}) - L| < \epsilon$ .

### 💡 Raja-arvo, ominaisuudet

Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = F$  ja  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = G$  niin pätee

- $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (\alpha f(\mathbf{x}) + \beta g(\mathbf{x})) = \alpha F + \beta G$ ,
- $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})g(\mathbf{x}) = FG$ ,
- $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{F}{G}$  jos  $G \neq 0$ .

Tästä seuraa, että raja-arvot, joiden määrittäminen on hankalaa ja näin ollen vaativat eniten työtä, ovat ne, joissa sijotus antaa tulokseksi  $\frac{0}{0}$ .

### 💡 Raja-arvo, kuristusperiaate

- Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = 0$  ja  $|f(\mathbf{x})| \leq g(\mathbf{x})$  (kun  $\mathbf{x} \neq \mathbf{x}_0$ ) niin  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = 0$ .
- Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} h(\mathbf{x}) = L$  ja  $g(\mathbf{x}) \leq f(\mathbf{x}) \leq h(\mathbf{x})$  (kun  $\mathbf{x} \neq \mathbf{x}_0$ ) niin  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = L$ .

### 💡 Kuristusperiaate

Jos meidän pitää määrittää raja-arvo

$$\lim_{(x,y) \rightarrow (0,0)} \frac{5x^2y^2}{x^4 + y^2},$$

niin voimme yrittää käyttää kuristusperiaatetta jos arvelemme, että raja-arvo on 0 ja silloin meidän pitää korvata  $\frac{5x^2y^2}{x^4+y^2}$  suuremmalla funktiolla (kun olemme ottaneet itseisarvon, mikä tässä ei ole tarpeen) jolle on helpompaa osoittaa, että raja-arvo on 0. Voimme ensin todeta, että  $x^4 \geq 0$  josta seuraa, että  $x^4 + y^2 \geq y^2$  jolloin

$$0 \leq \frac{5x^2y^2}{x^4 + y^2} \leq \frac{5x^2y^2}{y^2} = 5x^2.$$

Nyt on selvää (?), että

$$\lim_{(x,y) \rightarrow (0,0)} 5x^2 = \lim_{x \rightarrow 0} 5x^2 = 5 \cdot 0^2 = 0,$$

ja näin ollen saamme kuristusperiaatteen nojalla

$$\lim_{(x,y) \rightarrow (0,0)} \frac{5x^2y^2}{x^4 + y^2} = 0.$$

### 💡 Raja-arvo säteitä pitkin

Jos  $f$  on yhden reaaliuuttujan funktio niin sillä on raja-arvo pisteessä  $t_0$  jos ja vain jos oikean- ja vasemmanpuoliset raja-arvot ovat samat.

Useamman muuttujan tapauksessa tilanne on osittain toinen:

- Jos raja-arvo  $\lim_{t \rightarrow 0^+} f(x_0 + \alpha t, y_0 + \beta t)$  ei ole riippumaton parametrien  $\alpha$  ja  $\beta$  arvoista niin raja-arvo  $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y)$  ei ole olemassa.
- Jos  $\lim_{t \rightarrow 0^+} f(x_0 + \alpha t, y_0 + \beta t) = L$  kaikilla  $\alpha$  ja  $\beta$  joilla  $\alpha^2 + \beta^2 > 0$  niin tästä seuraa ainoastaan, että jos raja-arvo on olemassa niin se on  $L$ .

### 😊 Epäyhtälöitä ym.

- Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = F$ ,  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = G$  ja  $f(\mathbf{x}) \leq g(\mathbf{x})$  kun  $\mathbf{x} \neq \mathbf{x}_0$  niin pätee  $F \leq G$ .
- Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = F$  niin pätee  $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = F$ .
- Jos  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = G$  ja  $g(\mathbf{x}) \neq G$  kun  $\mathbf{x} \neq \mathbf{x}_0$  niin pätee  $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(g(\mathbf{x})) = \lim_{t \rightarrow G} f(t)$ .

### 💡 Esimerkki: Raja-arvo

Jos haluamme määrittää raja-arvon  $\lim_{(x,y) \rightarrow (0,0)} \frac{xy^2}{x^2+y^4}$ , niin voimme laskea raja-arvon sädetä  $(\alpha t, \beta t)$  pitkin ja kun määrittelemme  $f(x, y) = \frac{xy^2}{x^2+y^4}$  saamme

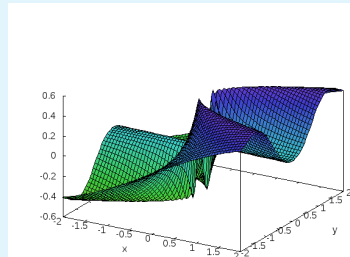
$$\lim_{t \rightarrow 0^+} f(\alpha t, \beta t) = \lim_{t \rightarrow 0^+} \frac{\alpha t \beta^2 t^2}{\alpha^2 t^2 + \beta^4 t^4} = \lim_{t \rightarrow 0} t \frac{\alpha \beta^2}{\alpha^2 + \beta^4 t^2} = 0$$

Tästä seuraa (ainoastaan!), että jos raja-arvo on olemassa niin se on 0.

Nyt osoittautuu, että raja-arvoa ei ole olemassa koska jos valitsemme  $x = t^2$  ja  $y = t$  ja laskemme raja-arvon kun  $t \rightarrow 0^+$

$$\lim_{t \rightarrow 0^+} f(t^2, t) = \lim_{t \rightarrow 0^+} \frac{t^2 t^2}{(t^2)^2 + t^4} = \frac{1}{2} \neq 0,$$

josta seuraa ettei  $f(x, y)$  ole "lähellä" 0 jos  $(x, y)$  on "riittävän lähellä"  $(0, 0)$ .



### 💡 Jatkuvat funktiot

- Funktio  $f : \Omega \rightarrow \mathbb{R}$  on jatkuva pisteessä  $\mathbf{x}_0$  jos  $\mathbf{x}_0 \in \Omega \subseteq \mathbb{R}^d$  ja

$$\lim_{\substack{\mathbf{x} \rightarrow \mathbf{x}_0 \\ \mathbf{x} \in \Omega}} f(\mathbf{x}) = f(\mathbf{x}_0)$$

eli  $\Omega$  on  $\mathbb{R}^d$ :n osajoukko,  $\mathbf{x}_0$  kuuluu joukkoon  $\Omega$ ,  $f(\mathbf{x}_0)$  on määritelty, raja-arvo on olemassa ja se on  $f(\mathbf{x}_0)$ .

- Funktion  $f : \Omega \rightarrow \mathbb{R}$  on jatkuva joukossa  $\Omega \subseteq \mathbb{R}^d$  jos

$$\lim_{\substack{\mathbf{x} \rightarrow \mathbf{x}_0 \\ \mathbf{x} \in \Omega}} f(\mathbf{x}) = f(\mathbf{x}_0) \quad \text{kaikilla } \mathbf{x}_0 \in \Omega.$$

eli jos se on jatkuva joukon  $\Omega$  jokaisessa pisteessä.

### 💡 Jatkuvat vektoriarvoiset funktiot

Funktio  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  on jatkuva pisteessä  $\mathbf{x}_0 \in \Omega$ , (joukossa  $\Omega$ ), jos jokainen komponentti on jatkuva pisteessä  $\mathbf{x}_0 \in \Omega$ , (joukossa  $\Omega$ ).

### 💡 Jatkuvien funktioiden summa, tulo ja osamäärä

Jos  $f$  ja  $g : \Omega \rightarrow \mathbb{R}$  ovat jatkuvia joukossa  $\Omega \subseteq \mathbb{R}^d$  niin

- $\alpha f(\mathbf{x}) + \beta g(\mathbf{x})$  ja  $f(\mathbf{x})g(\mathbf{x})$  ovat jatkuvia joukossa  $\Omega$ ,
- $\frac{f(\mathbf{x})}{g(\mathbf{x})}$  on jatkuva joukossa  $\{\mathbf{x} \in \Omega : g(\mathbf{x}) \neq 0\}$ .

### 💡 Jatkuvien funktioiden yhdistetty funktio on jatkuva

Jos  $g : \Omega_g \rightarrow \Omega_f$  ja  $f : \Omega_f \rightarrow \mathbb{R}^m$ , missä  $\Omega_g \subseteq \mathbb{R}^d$  ja  $\Omega_f \subseteq \mathbb{R}^p$ , ovat jatkuvia määrittelyjoukoissaan niin funktio  $(f \circ g)(\mathbf{x}) = f(g(\mathbf{x}))$  on jatkuva joukossa  $\Omega_g$ .

### 💡 Huom!

Jos funktio  $f(x, y)$  on jatkuva (esim.  $\mathbb{R}^2$ :ssa) niin funktio  $x \mapsto f(x, y)$  on jatkuva kaikilla parametrin  $y$  arvoilla ja funktio  $y \mapsto f(x, y)$  on jatkuva kaikilla parametrin  $x$  arvoilla.

Mutta jos ainoastaan oletamme, että funktio  $x \mapsto f(x, y)$  on jatkuva kaikilla  $y$  ja funktio  $y \mapsto f(x, y)$  on jatkuva kaikilla  $x$  niin tästä ei välttämättä seuraa, että funktio  $(x, y) \mapsto f(x, y)$  olisi jatkuva.

### 💡 Osittaisderivaatat

$$f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}$$

$$f_y(x, y) = \lim_{k \rightarrow 0} \frac{f(x, y+k) - f(x, y)}{k}$$

### 💡 Erilaisia merkintätapoja

$$f_x = \frac{\partial f}{\partial x} = D_x f = f_1 = D_1 f = D^{(1,0)} f \dots$$

$$f_{xy} = (f_x)_y = \frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x} = D^{(1,1)} f.$$

## 💡 Derivoimisjärjestyksen vaihto

Oleta, että funktion  $f(x, y)$  osittaisderivaatat  $f_x(x, y)$ ,  $f_y(x, y)$ ,  $f_{xy}(x, y)$  ja  $f_{yx}(x, y)$  ovat olemassa (ainakin) kaikissa pisteissä  $(x, y)$  joilla  $(x - x_0)^2 + (y - y_0)^2 < \delta^2$  ja, että funktiot  $f_{xy}(x, y)$  ja  $f_{yx}(x, y)$  ovat jatkuvia pisteessä  $(x_0, y_0)$ . Silloin pätee  $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$ .

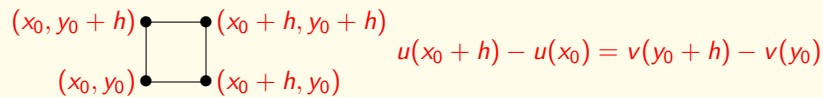
## 😊 Todistus

Olkoon  $0 < h < \frac{\delta}{2}$  ja määrittele funktiot  $u$  ja  $v$  siten, että

$$u(x) = f(x, y_0 + h) - f(x, y_0) \quad \text{ja} \quad v(y) = f(x_0 + h, y) - f(x_0, y).$$

Funktio  $u$  on derivoituva välillä  $(x_0, x_0 + h)$  ja jatkuva välillä  $[x_0, x_0 + h]$  ja näin ollen väliarvolauseesta seuraa, että on olemassa luku  $\theta_1 \in (0, 1)$  siten, että

$$u(x_0 + h) - u(x_0) = hu'(x_0 + \theta_1 h).$$



## 😊 Todistus, jatk.

Koska  $hu'(x_0 + \theta_1 h) = h(f_x(x_0 + \theta_1 h, y_0 + h) - f_x(x_0 + \theta_1 h, y_0))$  voimme soveltaa väliarvolauseetta vielä kerran funktioon  $y \mapsto f_x(x_0 + \theta_1 h, y)$  ja saamme

$$u(x_0 + h) - u(x_0) = h^2 f_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 h),$$

missä  $\theta_2 \in (0, 1)$ .

"Samalla tavalla" toteamme myös, että

$$v(y_0 + h) - v(y_0) = h^2 f_{yx}(x_0 + \theta_4 h, y_0 + \theta_3 h),$$

missä  $\theta_3$  ja  $\theta_4 \in (0, 1)$ . Nyt  $u(x_0 + h) - u(x_0) = v(y_0 + h) - v(y_0)$  koska molemmat lausekkeet ovat

$f(x_0 + h, y_0 + h) - f(x_0 + h, y_0) - f(x_0, y_0 + h) + f(x_0, y_0)$  josta seuraa, että

$$h^2 f_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 h) = h^2 f_{yx}(x_0 + \theta_4 h, y_0 + \theta_3 h).$$

Jos nyt jaamme  $h^2$ :lla, otamme raja-arvon kun  $h \rightarrow 0$  ja käytämme hyväksi oletusta, että  $f_{xy}$  ja  $f_{yx}$  ovat jatkuvia, niin saamme väitteen  $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$ .

## 💡 Derivaatta

Funktio  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  on derivoituva pisteessä  $\mathbf{x}$  jos on olemassa (rivi)vektori, joka merkitään  $f'(\mathbf{x})$ :llä siten, että

$$\lim_{\mathbf{h} \rightarrow 0} \frac{|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - f'(\mathbf{x})\mathbf{h}|}{\|\mathbf{h}\|} = 0.$$

Tässä  $f'(\mathbf{x})\mathbf{h}$  on  $1 \times d$ -rivivektorin  $f'(\mathbf{x})$  ja  $d \times 1$ -sarakevektorin  $\mathbf{h}$  matriisitulo, ja se voidaan myös esittää pistetulon muodossa  $f'(\mathbf{x}) \cdot \mathbf{h}$  jos ei haluta tehdä eroa rivi- ja pystyvektorien välillä.

Muita usein käytettyjä derivaatan merkintöjä ovat  $\nabla f(\mathbf{x})$  ja  $Df(\mathbf{x})$ .

## 💡 Derivaatta ja osittaisderivaatat

Jos funktiolla  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  on jatkuvat osittaisderivaatat niin  $f$  on derivoituva ja

$$f'(\mathbf{x}) = \nabla f(\mathbf{x}) = Df(\mathbf{x}) = [f_{x_1}(\mathbf{x}) \quad f_{x_2}(\mathbf{x}) \quad \dots \quad f_{x_d}(\mathbf{x})].$$

💡 "Jatkuvasti derivoituva"  $\Leftrightarrow$  "Derivoituva ja derivaatta on jatkuva"

## 💡 Esimerkki: Osittaisderivaattoja

Jos  $f(x, y, z) = \ln(1 + e^{x+yz^2})$  niin  $f$ :n osittaisderivaatat ovat seuraavat:

$$f_x(x, y, z) = \frac{\partial}{\partial x} \ln(1 + e^{x+yz^2}) = \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2},$$

$$f_y(x, y, z) = \frac{\partial}{\partial y} \ln(1 + e^{x+yz^2}) = \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} z^2,$$

$$f_z(x, y, z) = \frac{\partial}{\partial z} \ln(1 + e^{x+yz^2}) = \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} 2yz.$$

Näin ollen funktion  $f$  derivaatta eli gradientti on

$$\begin{aligned} f'(\mathbf{x}, y, z) &= Df(x, y, z) = \nabla f(x, y, z) \\ &= \left[ \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2}, \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} z^2, \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} 2yz \right] \\ &= \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} \mathbf{i} + \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} z^2 \mathbf{j} + \frac{1}{1 + e^{x+yz^2}} e^{x+yz^2} 2yz \mathbf{k}. \end{aligned}$$

## 💡 Vektoriarvoisen funktion derivaatta

Funktio  $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$  on derivoituva pisteessä  $\mathbf{x}$  jos jokainen komponentti on derivoituva pisteessä  $\mathbf{x}$ , eli on olemassa  $m \times d$  matriisi  $\mathbf{f}'(\mathbf{x})$  siten, että

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Matriisin  $\mathbf{f}'(\mathbf{x})$  rivivektorit ovat vektorin  $\mathbf{f}$  komponenttien derivaatat ja  $\mathbf{f}'(\mathbf{x})(i, j) = \frac{\partial f_i(\mathbf{x})}{\partial x_j}$ . (Tätä matriisiä kutsutaan usein Jacobin matriisiksi.)

## 💡 Ketjusääntö

Jos  $\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{g}(\mathbf{x}))$  missä  $\mathbf{f}$  ja  $\mathbf{g}$  ovat derivoituvia niin pätee

$$\mathbf{h}'(\mathbf{x}) = \mathbf{f}'(\mathbf{g}(\mathbf{x}))\mathbf{g}'(\mathbf{x}).$$

Huomaa, että tässä oletetaan, että  $\mathbf{x}$ ,  $\mathbf{g}$  ja  $\mathbf{f}$  ovat sarakevektoreita (mahdollisesti vain yhdellä komponentilla) ja on tärkeää että ketjusääntö kirjoitetaan järjestyksessä  $\mathbf{f}'(\mathbf{g})\mathbf{g}'$  koska jos  $\mathbf{g}: \mathbb{R}^d \rightarrow \mathbb{R}^p$  ja  $\mathbf{f}: \mathbb{R}^p \rightarrow \mathbb{R}^m$  niin  $\mathbf{f}'$  on  $m \times p$ -matriisi ja  $\mathbf{g}'$  on  $p \times d$ -matriisi jolloin matriisitulo  $\mathbf{f}'(\mathbf{g})\mathbf{g}'(\mathbf{x})$  on hyvin määritelty.

## 💡 Suunnattu derivaatta

Funktion  $f$  suunnattu derivaatta pisteessä  $\mathbf{x}$  suuntaan  $\mathbf{u}$  on funktion  $t \mapsto f(\mathbf{x} + t \frac{1}{\|\mathbf{u}\|}\mathbf{u})$  derivaatta pisteessä 0 ja on näin ollen

$$D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\mathbf{u} \frac{1}{\|\mathbf{u}\|} = \nabla f(\mathbf{x}) \cdot \left( \frac{1}{\|\mathbf{u}\|}\mathbf{u} \right),$$

ja se kertoo miten nopeasti funktio  $f$  kasvaa tai vähenee kun kuljetaan pisteestä  $\mathbf{x}$  suuntaan  $\mathbf{u}$ .

Huomaa, että  $D_i f(x, y, z) = f_x(x, y, z)$ ,  $D_j f(x, y, z) = f_y(x, y, z)$  ja  $D_k f(x, y, z) = f_z(x, y, z)$ , eli osittaisderivaatat ovat suunnatut derivaatat koordinaattiakselien (positiivisiin) suuntiin.

## 💡 Mihin suuntaan osoittaa gradientti?

Suunnattu derivaatta  $D_{\mathbf{u}}f(\mathbf{x})$  pisteessä  $\mathbf{x}$  on suurimmillaan kun  $\mathbf{u} \uparrow Df(\mathbf{x})$  (ja pienin vastakkaiseen suuntaan) joten **funktio kasvaa nopeimmin gradientin suuntaan**.

Suunnattu derivaatta on 0 kun  $\mathbf{u} \perp Df(\mathbf{x})$  joten **gradientti on kohtisuorassa tasa-arvokäyrä (tai -pintoja) kohti**.

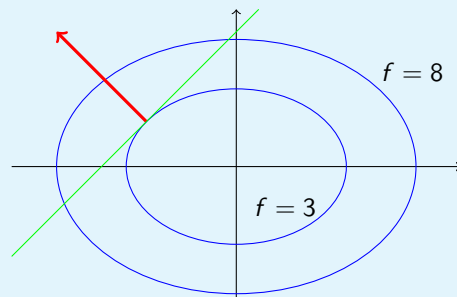
## 😊 Esimerkki: Gradientti ja tasa-arvokäyrä

Jos  $f(x, y) = \frac{1}{2}x^2 + y^2$  niin sen gradientti eli eli derivaatta on

$$Df(x, y) = [x \quad 2y] = x\mathbf{i} + 2y\mathbf{j}$$

ja erityisesti pisteessä  $(-2, 1)$  gradientti on  $-2\mathbf{i} + 2\mathbf{j}$ . Funktion pisteen  $(-2, 1)$  kautta kulkeva tasa-arvokäyrä on (ellipsi)  $\{(x, y) : \frac{1}{2}x^2 + y^2 = 3\}$  ja alla olevasta kuvioista nähdään, että gradientti on kohtisuorassa tasa-arvokäyrää ja sen tangenttia pisteessä  $(-2, 1)$  vastaan.

Koska gradientti pisteessä  $(2, 1)$  on  $-2\mathbf{i} + 2\mathbf{j}$  niin tätä vastaan kohtisuorassa olevan, pisteen  $(-2, 1)$  kautta kulkevan suoran suuntavektori on  $2\mathbf{i} + 2\mathbf{j}$  jolloin tämän suoran yhtälö on  $y - 1 = x + 2$  eli  $y = x + 3$ .



## 😊 Derivaatta ja koordinaattisysteemi

Edellä annetun määritelmän mukaan esimerkiksi funktion  $f(x, y) = xy$  derivaatta eli gradientti on  $\mathbf{f}'(x, y) = [y \quad x] = y\mathbf{i} + x\mathbf{j}$ , mutta tässä määritelmässä meillä on oletuksena, että koordinaattisysteemi on kiinnitetty. Jos valitsemme toisen, st-koordinaattisysteemin, esimerkiksi (kiertämällä xy-systeemin koordinaattiakseleita  $45^\circ$ ) siten, että

$$\begin{bmatrix} s \\ t \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x + y \\ y - x \end{bmatrix},$$

niin  $f(s, t) = \frac{1}{2}(s^2 - t^2)$  ja  $\mathbf{f}'(s, t) = [s \quad -t] = s\mathbf{u} - t\mathbf{v}$  missä  $\mathbf{u}$  ja  $\mathbf{v}$  ovat s- ja t-akselien suuntaisia kantavektoreita, eli  $\mathbf{u} = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j})$  ja  $\mathbf{v} = \frac{1}{\sqrt{2}}(\mathbf{j} - \mathbf{i})$ .

Näin ollen voimme todeta, että derivaatta pisteessä  $\mathbf{x}$  "oikeasti" on lineaarikuvauksena  $\mathbf{h} \mapsto \mathbf{f}'(\mathbf{x})\mathbf{h}$  ja tämän lineaarikuvauksen matriisiesitys (jota sitten käytetään laskuissa) riippuu käytetystä koordinaattisysteemistä. Samalla tavalla voimme sanoa, että funktion  $g(x) = x^2$  derivaatta pisteessä  $x = 2$  ei ole luku  $g'(2) = 4$  vaan lineaarikuvauksena "neljällä kertominen":  $t \mapsto 4t$ .

## 😊 Derivaatan tulosääntö ketjusäännön avulla

Derivaatan tulosääntö sanoo tunnetusti, että  $(fg)' = f'g + fg'$ .

Vaihtoehtoinen lähestymistapa on seuraava: Kirjoita  $p(t) = f(t)g(t)$  yhdistettynä funktiona  $p(t) = (h \circ k)(t) = h(k(t))$  missä  $h: \mathbb{R}^2 \rightarrow \mathbb{R}$  on

$h\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = xy$  ja  $k(t) = \begin{bmatrix} f(t) \\ g(t) \end{bmatrix}$ . Derivaatan määritelmän nojalla

$$h'\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = [y \quad x] \quad \text{ja} \quad k'(t) = \begin{bmatrix} f'(t) \\ g'(t) \end{bmatrix}.$$

Ketusäännön mukaan pätee silloin

$$p'(t) = h'(k(t))k'(t) = [g(t) \quad f(t)] \begin{bmatrix} f'(t) \\ g'(t) \end{bmatrix} = g(t)f'(t) + f(t)g'(t).$$

## 💡 Laplacen yhtälö $u_{xx} + u_{yy} = 0$ napakoordinaateissa

Oletamme, että  $u$  toteuttaa ns. Laplacen yhtälön  $u_{xx} + u_{yy} = 0$  ja tehtävänä on selvittää minkä yhtälön funktio  $v(r, \theta) = u(r \cos(\theta), r \sin(\theta))$  toteuttaa. (Tässä siis  $v$  on funktio  $u$  esitettynä napakoordinaattien funktiona).

Ositaisderivaattoja laskemalla (ja ketjusääntöä hyväksi käyttäen) toteamme, että

$$v_r(r, \theta) = u_x(r \cos(\theta), r \sin(\theta)) \cos(\theta) + u_y(r \cos(\theta), r \sin(\theta)) \sin(\theta),$$
$$v_\theta(r, \theta) = u_x(r \cos(\theta), r \sin(\theta))(-r \sin(\theta)) + u_y(r \cos(\theta), r \sin(\theta))r \cos(\theta).$$

Osoittautuu, että meidän ei tarvitse laskea  $v_{r\theta} = v_{\theta r}$  (mutta tämä ei ole lainkaan etukäteen selvää) mutta sen sijaan meidän pitää laskea

$$v_{rr}(r, \theta) = u_{xx}(r \cos(\theta), r \sin(\theta)) \cos^2(\theta) + 2u_{xy}(r \cos(\theta), r \sin(\theta)) \cos(\theta) \sin(\theta) + u_{yy}(r \cos(\theta), r \sin(\theta)) \sin^2(\theta),$$

## 💡 Laplacen yhtälö napakoordinaateissa, jatkuu

ja

$$v_{\theta\theta}(r, \theta) = u_{xx}(r \cos(\theta), r \sin(\theta))r^2 \sin^2(\theta) + 2u_{xy}(r \cos(\theta), r \sin(\theta))(-r^2 \cos(\theta) \sin(\theta)) + u_{yy}(r \cos(\theta), r \sin(\theta))r^2 \cos^2(\theta) - u_x(r \cos(\theta), r \sin(\theta))r \cos(\theta) - u_y(r \cos(\theta), r \sin(\theta))r \sin(\theta).$$

Oletuksen mukaan  $u_{xx}(x, y) + u_{yy}(x, y) = 0$ , erityisesti

$$u_{xx}(r \cos(\theta), r \sin(\theta)) + u_{yy}(r \cos(\theta), r \sin(\theta)) = 0,$$

ja tämän sekä kaavan  $\cos^2(\theta) + \sin^2(\theta) = 1$  avulla näemme, että

$$v_{rr}(r, \theta) + \frac{1}{r}v_r(r, \theta) + \frac{1}{r^2}v_{\theta\theta}(r, \theta) = 0$$

## 💡 Esimerkki: Suunnatut derivaatat

Jos funktio  $f$  on kolmen muuttujan derivoituva funktio ja en suunnatut derivaatat pisteessä  $\mathbf{x}_0$  suuntiin  $\mathbf{u}_1 = \mathbf{i} + \mathbf{j}$ ,  $\mathbf{u}_2 = \mathbf{j} - \mathbf{k}$  ja  $\mathbf{u}_3 = -\mathbf{i} + \mathbf{k}$  ovat 1, 2 ja 3 niin voimme määrittää derivaatan pisteessä  $\mathbf{x}_0$  koska vektorit  $\mathbf{u}_1$ ,  $\mathbf{u}_2$  ja  $\mathbf{u}_3$  ovat lineaarisesti riippumattomia (eli erisuuntaisia): Oletamme, että  $f'(\mathbf{x}_0) = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$  jolloin suunnatun derivaatan määritelmän nojalla saamme yhtälösystemin

$$1 = D_{\mathbf{u}_1}f(\mathbf{x}_0) = (A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot (\mathbf{i} + \mathbf{j}) \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}A + \frac{1}{\sqrt{2}}B,$$

$$2 = D_{\mathbf{u}_2}f(\mathbf{x}_0) = (A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot (\mathbf{j} - \mathbf{k}) \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}B - \frac{1}{\sqrt{2}}C,$$

$$3 = D_{\mathbf{u}_3}f(\mathbf{x}_0) = (A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot (-\mathbf{i} + \mathbf{k}) \frac{1}{\sqrt{2}} = -\frac{1}{\sqrt{2}}A + \frac{1}{\sqrt{2}}C.$$

Matriisimuodossa tämä systeemi ja sen ratkaisu ovat

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \Rightarrow \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} -2\sqrt{2} \\ 3\sqrt{2} \\ \sqrt{2} \end{bmatrix}.$$

### 💡 Esimerkki: Ketjusääntö ja virtausmekaniikka

Oletamme, että neste virtaa tasossa siten, että nopeus pisteessä  $(x, y)$  hetkellä  $t$  on  $u(x, y, t)\mathbf{i} + v(x, y, t)\mathbf{j}$ . Jos nyt haluamme määrittää hetkellä  $t$  pisteessä  $(x, y)$  sijaitsevan "nestepartikkelin" kiihtyvyys niin ensin oletamme, että partikkelin sijainti hetkellä  $t$  on  $(X(t), Y(t))$  (jolloin siis  $X(t) = x$  ja  $Y(t) = y$  jos se on pisteessä  $(x, y)$  hetkellä  $t$ ). Silloin nopeus on  $X'(t)\mathbf{i} + Y'(t)\mathbf{j}$ , josta seuraa, että  $X'(t) = u(X(t), Y(t), t)$  ja  $Y'(t) = v(X(t), Y(t), t)$ . Kiihtyvyys taas on  $X''(t)\mathbf{i} + Y''(t)\mathbf{j}$  ja ketjusäännön nojalla

$$\begin{aligned} X''(t) &= u_x(X(t), Y(t), t)X'(t) + u_y(X(t), Y(t), t)Y'(t) \\ &\quad + u_t(X(t), Y(t), t) \\ &= u_x(X(t), Y(t), t)u(X(t), Y(t), t) + u_y(X(t), Y(t), t)v(X(t), Y(t), t) \\ &\quad + u_t(X(t), Y(t), t) \\ &= u_x(x, y, t)u(x, y, t) + u_y(x, y, t)v(x, y, t) + u_t(x, y, t), \end{aligned}$$

missä viimeisellä rivillä käytimme oletusta  $x(t) = x$  ja  $Y(t) = y$ .

### 💡 Esimerkki: Ketjusääntö ja virtausmekaniikka, jatk.

Samalla tavalla saamme

$$Y''(t) = v_x(x, y, t)u(x, y, t) + v_y(x, y, t)v(x, y, t) + v_t(x, y, t).$$

Näin ollen kiihtyvyys on

$$\begin{aligned} &(u_x(x, y, t)u(x, y, t) + u_y(x, y, t)v(x, y, t) + u_t(x, y, t))\mathbf{i} \\ &\quad + (v_x(x, y, t)u(x, y, t) + v_y(x, y, t)v(x, y, t) + v_t(x, y, t))\mathbf{j}. \end{aligned}$$

Huomaa, että tämä on epälineaarinen lauseke eli jos  $u$  korvataan  $u_1 + u_2$ :lla ja samoin  $v$  korvataan  $v_1 + v_2$ :lla niin tulos ei ole kahden lausekkeen summa missä toisessa esiintyy  $u_1$  ja  $v_1$  ja toisessa  $u_2$  ja  $v_2$ . Suuri osa virtausmekaniikan hankaluuksista on seuraus tästä epälinearisuudesta.

### 😊 Milloin pätee $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq C\|\mathbf{x} - \mathbf{y}\|$ kun $\|\mathbf{x}\|, \|\mathbf{y}\| \leq \rho$ ?

Jos  $\mathbf{f}: \mathbb{R}^d \rightarrow \mathbb{R}^m$  on jatkuvasti derivoituva, niin jokainen komponentti  $f_j$ ,  $1 \leq j \leq m$ , on jatkuvasti derivoituva ja löytyy luvut  $c_j$  siten, että  $\|f'_j(\mathbf{v})\| \leq C$  kun  $\|\mathbf{v}\| \leq \rho$ . Jos nyt  $\mathbf{x}$  ja  $\mathbf{y}$  ovat sellaisia, että  $\|\mathbf{x}\| \leq \rho$  ja  $\|\mathbf{y}\| \leq \rho$  niin määritellään funktio  $h$  kaavalla  $h_j(t) = f_j((1-t)\mathbf{y} + t\mathbf{x})$ . Silloin  $h_j$  on derivoituva reaaliarvoinen funktio ja väliarvolauseen nojalla

$$f_j(\mathbf{x}) - f_j(\mathbf{y}) = h_j(1) - h_j(0) = h'_j(t_j)(1 - 0),$$

missä  $t_j \in (0, 1)$ . Ketjusäännön nojalla

$$h'_j(t_j) = f'_j((1-t_j)\mathbf{y} + t_j\mathbf{x})(\mathbf{x} - \mathbf{y}),$$

joten

$$|f_j(\mathbf{x}) - f_j(\mathbf{y})| = |h_j(1) - h_j(0)| \leq \|f'_j((1-t_j)\mathbf{y} + t_j\mathbf{x})\| \|\mathbf{x} - \mathbf{y}\| \leq c_j \|\mathbf{x} - \mathbf{y}\|,$$

koska  $\|(1-t_j)\mathbf{y} + t_j\mathbf{x}\| \leq (1-t_j)\|\mathbf{y}\| + t_j\|\mathbf{x}\| \leq \rho$ . Näin ollen  $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq C\|\mathbf{x} - \mathbf{y}\|$

missä  $C = \sqrt{c_1^2 + \dots + c_m^2}$ .

### 😊 Milloin pätee $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \geq c\|\mathbf{x} - \mathbf{y}\|$ kun $\|\mathbf{x}\|, \|\mathbf{y}\| \leq \rho$ ?

Ei välttämättä jos  $d > 1$ ,  $\mathbf{g}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  on jatkuvasti derivoituva,  $\rho > 0$  ja  $\|\mathbf{g}'(\mathbf{v})\mathbf{u}\| > 0$  kun  $\|\mathbf{v}\| \leq \rho$ ,  $\mathbf{u} \in \mathbb{R}^d$  ja  $\|\mathbf{u}\| = 1$  kuten seuraava esimerkki näyttää:

$$\text{Olkoon } \mathbf{g}(\mathbf{v}) = \begin{bmatrix} e^{\frac{2\pi s}{\rho}} \cos(\frac{2\pi t}{\rho}) \\ e^{\frac{2\pi s}{\rho}} \sin(\frac{2\pi t}{\rho}) \end{bmatrix} \text{ kun } \mathbf{v} = \begin{bmatrix} s \\ t \end{bmatrix}. \text{ Silloin}$$

$$\mathbf{g}'(\mathbf{v}) = \frac{2\pi}{\rho} \begin{bmatrix} e^{\frac{2\pi s}{\rho}} \cos(\frac{2\pi t}{\rho}) & -e^{\frac{2\pi s}{\rho}} \sin(\frac{2\pi t}{\rho}) \\ e^{\frac{2\pi s}{\rho}} \sin(\frac{2\pi t}{\rho}) & e^{\frac{2\pi s}{\rho}} \cos(\frac{2\pi t}{\rho}) \end{bmatrix},$$

Jos  $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$  niin

$$\mathbf{g}'(\mathbf{v})\mathbf{u} = \frac{2\pi}{\rho} e^{\frac{2\pi s}{\rho}} \begin{bmatrix} \cos(\frac{2\pi t}{\rho})u_1 - \sin(\frac{2\pi t}{\rho})u_2 \\ \sin(\frac{2\pi t}{\rho})u_1 + \cos(\frac{2\pi t}{\rho})u_2 \end{bmatrix},$$



😊 Milloin pätee  $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \geq c\|\mathbf{x} - \mathbf{y}\|$  kun  $\|\mathbf{x}\|, \|\mathbf{y}\| \leq \rho$ ?  
jatk.

joten

$$\begin{aligned} \|\mathbf{g}'(\mathbf{v})\mathbf{u}\|^2 &= \frac{4\pi^2}{\rho^2} e^{\frac{4\pi x}{\rho}} \left( \cos\left(\frac{2\pi t}{\rho}\right)^2 u_1^2 - 2 \cos\left(\frac{2\pi t}{\rho}\right) \sin\left(\frac{2\pi t}{\rho}\right) u_1 u_2 \right. \\ &\quad \left. + \sin\left(\frac{2\pi t}{\rho}\right)^2 u_2^2 + \sin\left(\frac{2\pi t}{\rho}\right)^2 u_1^2 \right. \\ &\quad \left. + 2 \cos\left(\frac{2\pi t}{\rho}\right) \sin\left(\frac{2\pi t}{\rho}\right) u_1 u_2 + \cos\left(\frac{2\pi t}{\rho}\right)^2 u_2^2 \right) = e^{2x} \|\mathbf{u}\|^2. \end{aligned}$$

Tästä päättelemme, että jos  $\|\mathbf{v}\| \leq \rho$ , niin  $s \geq -\rho$  ja

$$\|\mathbf{g}'(\mathbf{v})\mathbf{u}\| \geq \frac{2\pi}{\rho} e^{-2\pi} \|\mathbf{u}\|,$$

kaikilla vektoreilla  $\mathbf{u}$ . Jos nyt  $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  ja  $\mathbf{y} = \begin{bmatrix} 0 \\ \rho \end{bmatrix}$  niin  $\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y}) = 0$ .

💡💡 Lineaarinen approksimointi

$$f(\mathbf{x} + \mathbf{h}) \approx f(\mathbf{x}) + f'(\mathbf{x})\mathbf{h} = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{h},$$

tai toisella tavalla esitettynä

$$\begin{aligned} f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) \\ \approx f_x(x, y, z)\Delta x + f_y(x, y, z)\Delta y + f_z(x, y, z)\Delta z. \end{aligned}$$

😊 Tangenttitaso

Koska gradientti on kohtisuorassa tasa-arvopintoja kohti niin pinnan  $f(x, y, z) = c$  normaali pisteessä  $(x_0, y_0, z_0)$  on

$$f'(x_0, y_0, z_0) = f_x(x_0, y_0, z_0)\mathbf{i} + f_y(x_0, y_0, z_0)\mathbf{j} + f_z(x_0, y_0, z_0)\mathbf{k},$$

ja pinnan tangenttitasolla pisteessä  $(x_0, y_0, z_0)$  on yhtälö

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0$$

😊 Esimerkki: Tangenttitaso

Jos haluamme määrittää funktion  $f(x, y) = x^2 - y^2$  kuvaajan tangenttitason kun  $x = 2$  ja  $y = 1$  niin kirjoitamme ensin yhtälön muodossa  $z = x^2 - y^2$  eli  $g(x, y, z) = 0$  missä  $g(x, y, z) = x^2 - y^2 - z$

(vaikka se voisi yhtä hyvin olla  $g(x, y, z) = -x^2 + y^2 + z$ ). Funktion  $g$  derivaatta on  $Dg(x, y, z) = 2x\mathbf{i} - 2y\mathbf{j} - \mathbf{k}$  ja pisteessä  $(2, 1, f(2, 1)) = (2, 1, 3)$  tämä derivaatta on  $\mathbf{n} = 4\mathbf{i} - 2\mathbf{j} - \mathbf{k}$ . Tämä vektori on pinnan  $g(x, y, z) = 0$  normaali pisteessä  $(2, 1, 3)$  ja samalla pinnan tässä pisteessä otetun tangenttitason normaali. Koska  $(2, 1, 3)$  on tangenttitason piste niin  $\mathbf{v} = (x - 2)\mathbf{i} + (y - 1)\mathbf{j} + (z - 3)\mathbf{k}$  on tangenttitason suuntainen vektori jos  $(x, y, z)$  on myös tangenttitason piste. Näin ollen  $\mathbf{v}$  on kohtisuorassa normaalia  $\mathbf{n}$  vastaan eli  $\mathbf{v} \cdot \mathbf{n} = 0$  ja tästä ehdosta saamme tangenttitason yhtälön

$$4(x - 2) - 2(y - 1) - (z - 3) = 0 \quad \text{eli} \quad 4x - 2y - z = 3.$$

💡💡 Lineaarinen approksimointi ja suhteelliset virheet, erikoistapaus

Jos

$$f(x_1, x_2, \dots, x_n) = c x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n},$$

niin lineaarisella approksimoinnilla saadaan

$$\begin{aligned} \frac{\Delta f}{f} &= \frac{f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) - f(x_1, x_2, \dots, x_n)}{f(x_1, x_2, \dots, x_n)} \\ &\approx \alpha_1 \frac{\Delta x_1}{x_1} + \alpha_2 \frac{\Delta x_2}{x_2} + \dots + \alpha_n \frac{\Delta x_n}{x_n}, \end{aligned}$$

ja erityisesti

$$\left| \frac{\Delta f}{f} \right| \lesssim |\alpha_1| \left| \frac{\Delta x_1}{x_1} \right| + |\alpha_2| \left| \frac{\Delta x_2}{x_2} \right| + \dots + |\alpha_n| \left| \frac{\Delta x_n}{x_n} \right|.$$

### 💡 Lineaarinen approksimointi, suhteellinen virhe

Jos meidän pitää arvioida virhettä, joka syntyy kun ympyräkartion tilavuus lasketaan kaavalla  $V = \frac{1}{3}\pi r^2 h$  ja tiedämme ainoastaan, että säteen  $r$  mitatun arvon virhe on korkeintaan 3% ja korkeuden  $h$  mitatun arvon virhe on korkeintaan 2% niin emme saa absoluuttista ylärajaa tilavuuden virheelle mutta saamme helposti approksimatiivisen ylärajan suhteelliselle virheelle seuraavalla tavalla: Lineaarisen approksimoinnin perusteella:

$$\Delta V = V(r + \Delta r, h + \Delta h) - V(r, h) \approx V_r(r, h)\Delta r + V_h(r, h)\Delta h,$$

josta seuraa, että

$$\frac{\Delta V}{V} \approx \frac{\frac{1}{3}2\pi r h \Delta r}{\frac{1}{3}\pi r^2 h} + \frac{\frac{1}{3}\pi r^2 \Delta h}{\frac{1}{3}\pi r^2 h} = 2 \cdot \frac{\Delta r}{r} + 1 \cdot \frac{\Delta h}{h}.$$

Koska oletamme, että  $|\frac{\Delta r}{r}| \leq 0.03$  ja  $|\frac{\Delta h}{h}| \leq 0.02$  niin

$$\left| \frac{\Delta V}{V} \right| \lesssim 2 \cdot 0.03 + 1 \cdot 0.02 = 0.08,$$

eli suhteellinen virhe on korkeintaan 8%.

### 😊 Esimerkki: Lineaarinen approksimointi

Sylinterimäisen säiliön sisältämä nestemäärä, kun säiliön akseli on vaakasuorassa on  $V = L(r^2 \arccos(1 - \frac{h}{r}) - (r - h)\sqrt{2hr - h^2})$  missä  $L$  on säiliön pituus, joka on noin 2 m,  $r$  on säde, joka on noin 50 cm ja  $h$  on nestepinnan korkeus, joka myös on noin 50 cm. Pituuden  $L$  ja säteen  $r$  pystymme mittaamaan 1 cm tarkkuudella. Miten tarkasti meidän on mitattava  $h$  jotta saisimme nestemäärän lasketuksi 50 litran tarkkuudella? Määrittelemme

$$f(L, r, h) = L\left(r^2 \arccos\left(1 - \frac{h}{r}\right) - (r - h)\sqrt{2hr - h^2}\right).$$

Silloin (funktion  $\arccos(t)$  derivaatta on  $-\frac{1}{\sqrt{1-t^2}}$ )

$$f_L(L, r, h) = (r^2 \arccos(1 - \frac{h}{r}) - (r - h)\sqrt{2hr - h^2})$$

$$f_r(L, r, h) = 2Lr \arccos(1 - \frac{h}{r}) - Lr^2 \frac{1}{\sqrt{1 - (1 - \frac{h}{r})^2}} \frac{h}{r^2} - L\sqrt{2hr - h^2} - L(r - h) \frac{h}{\sqrt{2hr - h^2}},$$

### 😊 Esimerkki: Lineaarinen approksimointi, jatk.

$$f_h(L, r, h) = Lr^2 \frac{1}{\sqrt{1 - (1 - \frac{h}{r})^2}} \cdot \frac{1}{r} + L\sqrt{2hr - h^2} - L(r - h) \frac{r - h}{\sqrt{2hr - h^2}},$$

ja erityisesti

$$f_L(200, 50, 50) = \frac{\pi}{2} \cdot 2500 \approx 3927$$

$$f_r(200, 50, 50) = \frac{\pi}{2} \cdot 20000 - 10000 - 10000 \approx 11416,$$

$$f_h(200, 50, 50) = 10000 + 10000 = 20000.$$

Lineaarisella approksimoinnilla saamme

$$\Delta f \approx f_L \Delta L + f_r \Delta r + f_h \Delta h.$$

### 😊 Esimerkki: Lineaarinen approksimointi, jatk.

josta seuraa, että

$$|\Delta f| \lesssim |f_L| |\Delta L| + |f_r| |\Delta r| + |f_h| |\Delta h| \leq 3927 \cdot 1 + 11416 \cdot 1 + 20000 \cdot |\Delta h|$$

koska  $|\Delta L| \leq 1$  ja  $|\Delta r| \leq 1$ . Jos nyt haluamme, että  $|\Delta f| \leq 50 \cdot 10^3$  niin meidän täytyy vaatia, että

$$3927 \cdot 1 + 11416 \cdot 1 + 20000 \cdot |\Delta h| \leq 50000,$$

ja tästä seuraa, että pitää olla  $|\Delta h| \leq 2.4$ .

### 💡 Huom

Tässä kuten muissa vastaavissa laskuissa  $\Delta L$  on "virhe" tai "poikkeama" muuttujan  $L$  arvossa, eli  $\Delta L = L - L_0$  mutta  $\Delta$ :lla merkitään myös Laplacen differentiaalioperaattoria:  $\Delta u = u_{xx} + u_{yy} + u_{zz}$ .

### 💡 Esimerkki: Lineaarinen approksimointi

Funktiosta  $f$  tiedämme, että  $f(3, 2) = 2.7$ ,  $f(3.1, 2.2) = 2.75$  ja  $f(2.9, 2.1) = 2.68$ . Nyt voimme seuraavalla tavalla arvioida derivaattaa käyttäen mikä on  $f(3.2, 1.9)$ :

Linearisella approksimoinnilla saamme

$$f(3 + \Delta x, 2 + \Delta y) \approx f(3, 2) + f_x(3, 2)\Delta x + f_y(3, 2)\Delta y.$$

Jos ensin valitsemme  $\Delta x = 0.1$  ja  $\Delta y = 0.2$  ja sitten  $\Delta x = -0.1$  ja  $\Delta y = 0.1$  niin saamme "yhtälösystemin"

$$2.75 = f((3 + 0.1, 2 + 0.2) \approx 2.7 + f_x(3, 2) \cdot 0.1 + f_y(3, 2) \cdot 0.2,$$

$$2.68 = f(3 - 0.1, 2 + 0.1) \approx 2.7 + f_x(3, 2) \cdot (-0.1) + f_y(3, 2) \cdot 0.1,$$

eli

$$0.5 \approx f_x(3, 2) + 2f_y(3, 2),$$

$$-0.2 \approx -f_x(3, 2) + f_y(3, 2).$$

### 💡 Esimerkki: Lineaarinen approksimointi, jatk.

Laskemalla yhteen saamme ensin  $f_y(3, 2) \approx 0.1$  ja sitten sijoittamalla tämän tuloksen ensimmäiseen yhtälöön saamme  $f_x(3, 2) \approx 0.5 - 0.2 = 0.3$ . Jos nyt valitsemme  $\Delta x = 0.2$  ja  $\Delta y = -0.1$  niin linearisella approksimoinnilla saamme

$$\begin{aligned} f(3.2, 1.9) &= f(3 + 0.2, 2 - 0.1) \\ &\approx f(3, 2) + f_x(3, 2) \cdot 0.2 + f_y(3, 2) \cdot (-0.1) \\ &\approx 2.7 + 0.3 \cdot 0.2 + 0.1 \cdot (-0.1) \\ &= 2.7 + 0.06 - 0.01 = 2.75. \end{aligned}$$

Huomaa, että tässä virhelähteenä on myös epätarkkuudet osittaisderivaattojen arvoissa.

### 😊🚧 Differentiaali??

Sanomme, että

$$\psi = a(x, y) dx + b(x, y) dy,$$

on 1. kertaluvun differentiaalimuoto (tai differentiaalimuotokenttä koska kertoimet  $a$  ja  $b$  ovat muuttujien  $x$  ja  $y$  funktioita) tai lyhyemmin vain differentiaali. Tässä 1. kertaluvun tapauksessa voimme pitää  $dx$  symbolina, jota vastaa  $x$ -akselin suuntaista yksikkövektoria ja samoin  $dy$  symbolina jota vastaa  $y$ -akselin suuntaista yksikkövektoria jolloin differentiaalia  $\psi$  vastaa vektorifunktio  $[a(x, y) \quad b(x, y)]$ .

Tällaista differentiaalia voimme (kunhan kertoimet ovat esim. jatkuvia) integroida käyrää pitkin: Jos  $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ ,  $t \in [a, b]$  on suunnatun käyrän  $C$  parametrisitys (ja funktiot  $x(t)$  ja  $y(t)$  ovat esim. paloittain jatkuvasti derivoituvia) niin

$$\int_C \psi = \int_a^b (a(x(t), y(t))x'(t) + b(x(t), y(t))y'(t)) dt.$$

### 😊🚧 Differentiaali??, jatk.

Jos nyt  $a(x, y) = f_x(x, y)$  ja  $b(x, y) = f_y(x, y)$  niin ketjusäännön nojalla integroitavana on derivaatta jolloin saamme integraalin suoraan sijoittamalla ja

$$\int_C \psi = f(\mathbf{r}(b)) - f(\mathbf{r}(a)) = f(C:n \text{ loppupiste}) - f(C:n \text{ alkupiste}).$$

Jos siis  $a(x, y) = f_x(x, y)$  ja  $b(x, y) = f_y(x, y)$  niin kirjoitamme  $\psi = df$ , sanomme, että  $\psi$  on funktion  $f$  kokonaisdifferentiaali ja

$$df = f_x dx + f_y dy,$$

ja tämän vastine on approksimointikaavaa

$$\Delta f \approx f_x \Delta x + f_y \Delta y.$$

Mutta jos ei päde  $a(x, y) = f_x(x, y)$  ja  $b(x, y) = f_y(x, y)$  jollain funktiolla  $f$  niin  $\int_C \psi$  on edelleen laskettavissa mutta tulos ei enää riipu pelkästään differentiaalista  $\psi$  ja käyrän päätepisteistä vaan myös siitä miten käyrä kulkee alkupisteestä loppupisteeseen.

## 💡💡 Newtonin menetelmä: $\mathbf{f}(\mathbf{x}) = \mathbf{0} \Rightarrow \mathbf{x} = ?$

Jos  $\mathbf{f}(\mathbf{x}_n + \Delta\mathbf{x}) = \mathbf{0}$  niin pätee

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_n + \Delta\mathbf{x}) \approx \mathbf{f}(\mathbf{x}_n) + \mathbf{f}'(\mathbf{x}_n)\Delta\mathbf{x} \Rightarrow \Delta\mathbf{x} \approx -\mathbf{f}'(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n)$$

ja jotta  $\mathbf{x}_{n+1} \approx \mathbf{x}_n + \Delta\mathbf{x}$  valitsemme

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{f}'(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n).$$

Käänteismatriisin  $\mathbf{f}'(\mathbf{x}_n)^{-1}$  laskeminen ei ole välttämätöntä mutta meidän pitää ratkaista yhtälösystemi

$$\mathbf{f}'(\mathbf{x}_n)(\mathbf{x}_{n+1} - \mathbf{x}_n) = -\mathbf{f}(\mathbf{x}_n).$$

## 💡 Milloin Newtonin menetelmä suppenee?

Jos  $\mathbf{f}(\mathbf{x})$  on jatkuvasti derivoituvia,  $\mathbf{f}(\mathbf{x}_*) = \mathbf{0}$ ,  $\mathbf{f}'(\mathbf{x}_*)$  on kääntyvä matriisi ja jos  $\|\mathbf{x}_0 - \mathbf{x}_*\|$  on riittävän pieni (mille voi antaa riittävä ehto) niin pätee  $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}_*$ , mutta muuten ei ole takeita siitä, että menetelmä konvergoi ja vaikeus on löytää sopiva alkuarvo (ja tietää milloin pitää luovuttaa jos näyttää siltä ettei menetelmä konvergoikaan).

## 💡💡 Newtonin menetelmä

Jos haluamme käyttää Newtonin menetelmää yhtälösystemiin  $x^2 + 4y^2 = 4$ ,  $x^2 - 2x = y - 1$  ratkaisemiseksi niin kirjoitamme ensin systeemin muodossa  $\mathbf{f}(\mathbf{X}) = \mathbf{0}$ , missä siis

$$\mathbf{f}(\mathbf{X}) = \begin{bmatrix} x^2 + 4y^2 - 4 \\ x^2 - 2x - y + 1 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Silloin

$$\mathbf{f}'(\mathbf{X}) = \begin{bmatrix} 2x & 8y \\ 2x - 2 & -1 \end{bmatrix},$$

ja meidän pitää laskea

$$\mathbf{X}_{n+1} = \mathbf{X}_n - \mathbf{f}'(\mathbf{X}_n)^{-1}\mathbf{f}(\mathbf{X}_n), \quad n \geq 0.$$

Jos valitsemme  $X_0 = \frac{3}{2}$  ja  $y_0 = \frac{1}{2}$  niin saamme

$$\mathbf{X}_1 = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \end{bmatrix} - \begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} -\frac{3}{4} \\ -\frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{7} & \frac{4}{7} \\ \frac{1}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} -\frac{3}{4} \\ -\frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{7}{4} \\ \frac{1}{2} \end{bmatrix}.$$

## 💡💡 Newtonin menetelmä, jatk.

Jos käytämme laskuihin Matlab/Octavea niin määrittelemme ensin funktion  $\mathbf{f}$  komennolla

$$\mathbf{f} = @(\mathbf{X}) [\mathbf{X}(1)^2 + 4*\mathbf{X}(2)^2 - 4; \mathbf{X}(1)^2 - 2*\mathbf{X}(1) - \mathbf{X}(2) + 1]$$

ja sitten sen derivaatan  $\mathbf{f}'$  komennolla

$$\mathbf{df} = @(\mathbf{X}) [2*\mathbf{X}(1), 8*\mathbf{X}(2); 2*\mathbf{X}(1) - 2, -1]$$

Sitten valitsemme  $\mathbf{X} = [1.5; 0.5]$  ja laskemme montokertaa  $\mathbf{X} = \mathbf{X} - \mathbf{df}(\mathbf{X})^{-1}*\mathbf{f}(\mathbf{X})$  (tai  $\mathbf{X} = \mathbf{X} - \mathbf{df}(\mathbf{X}) \setminus \mathbf{f}(\mathbf{X})$ ) ja tuloksena saamme

$$\mathbf{X}_1 = \begin{bmatrix} 1.7500000000000000 \\ 0.5000000000000000 \end{bmatrix} \quad \mathbf{X}_2 = \begin{bmatrix} 1.717105263157895 \\ 0.513157894736842 \end{bmatrix}$$

$$\mathbf{X}_3 = \begin{bmatrix} 1.716438375580136 \\ 0.513283501264863 \end{bmatrix} \quad \mathbf{X}_4 = \begin{bmatrix} 1.716438125054029 \\ 0.513283587030869 \end{bmatrix}$$

$$\mathbf{X}_5 = \begin{bmatrix} 1.716438125053991 \\ 0.513283587030878 \end{bmatrix} \quad \mathbf{X}_6 = \begin{bmatrix} 1.716438125053991 \\ 0.513283587030878 \end{bmatrix}$$

## 😊 Miksi Newtonin menetelmä konvergoi niin nopeasti kun se konvergoi?

Oletamme, että funktio  $\mathbf{f} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  on derivoituva ja sellainen, että on olemassa vakioita  $L$  ja  $M$  siten, että

$$\|\mathbf{f}'(\mathbf{x} + \mathbf{v}) - \mathbf{f}'(\mathbf{x})\| \leq L\|\mathbf{v}\| \quad \text{ja} \quad \|\mathbf{f}'(\mathbf{x})^{-1}\| \leq M,$$

kaikilla  $\mathbf{x}$  ja  $\mathbf{v} \in \mathbb{R}^d$  (tai ainakin kun  $\mathbf{x} = \mathbf{x}_n$  ja  $\|\mathbf{v}\| \leq \|\mathbf{x}_{n+1} - \mathbf{x}_n\|$ ). Lisäksi oletamme, että  $\mathbf{f}(\mathbf{x}_*) = \mathbf{0}$ . Jos käytämme Newtonin menetelmää yhtälösystemiin  $\mathbf{f}(\mathbf{x}) = \mathbf{0}$  ratkaisemiseksi ja laskemme  $(\mathbf{x}_n)_{n=0}^{\infty}$  niin pätee

$$\|\mathbf{x}_{n+1} - \mathbf{x}_*\| \leq \frac{1}{2}ML\|\mathbf{x}_n - \mathbf{x}_*\|^2, \quad n \geq 1.$$

Perustelu tähän on seuraava: Jos  $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^d$  on jatkuvasti derivoituva niin pätee  $\mathbf{g}(1) - \mathbf{g}(0) - \mathbf{g}'(0) = \int_0^1 (\mathbf{g}'(t) - \mathbf{g}'(0)) dt$ . Jos  $\mathbf{g}(t) = \mathbf{f}(\mathbf{x} + t\mathbf{h})$  niin  $\mathbf{g}'(t) = \mathbf{f}'(\mathbf{x} + t\mathbf{h})\mathbf{h}$  jolloin

$$\begin{aligned} \mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h} &= \mathbf{g}(1) - \mathbf{g}(0) - \mathbf{g}'(0) \\ &= \int_0^1 (\mathbf{f}'(\mathbf{x} + t\mathbf{h})\mathbf{h} - \mathbf{f}'(\mathbf{x})\mathbf{h}) dt. \end{aligned}$$

😊 Miksi Newtonin menetelmä konvergoi niin nopeasti kun se konvergoi? jatk.

Tästä epäyhtälöstä ja oletuksista seuraa, että

$$\begin{aligned}\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\| &\leq \int_0^1 \|\mathbf{f}'(\mathbf{x} + t\mathbf{h})\mathbf{h} - \mathbf{f}'(\mathbf{x})\mathbf{h}\| dt \\ &\leq \int_0^1 Lt\|\mathbf{h}\|^2 dt = \frac{1}{2}L\|\mathbf{h}\|^2.\end{aligned}$$

Jos nyt valitsemme  $\mathbf{x} = \mathbf{x}_n$  ja  $\mathbf{h} = \mathbf{x}_* - \mathbf{x}_n$  niin  $\mathbf{f}(\mathbf{x}_n + \mathbf{h}) = \mathbf{0}$  ja

$$\mathbf{x}_{n+1} - \mathbf{x}_* = \mathbf{x}_n - \mathbf{x}_* - \mathbf{f}'(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n) = \mathbf{f}'(\mathbf{x}_n)^{-1}(\mathbf{f}(\mathbf{x}_n + \mathbf{h}) - \mathbf{f}(\mathbf{x}_n) - \mathbf{f}'(\mathbf{x}_n)\mathbf{h}),$$

ja oletuksesta edellä johdetusta epäyhtälöstä seuraa, että

$$\|\mathbf{x}_{n+1} - \mathbf{x}_*\| \leq M\|\mathbf{f}(\mathbf{x}_n + \mathbf{h}) - \mathbf{f}(\mathbf{x}_n) - \mathbf{f}'(\mathbf{x}_n)\mathbf{h}\| \leq \frac{1}{2}ML\|\mathbf{x}_n - \mathbf{x}_*\|^2.$$

Tästä tuloksesta seuraa, että **ainakin** jos  $\frac{1}{2}ML\|\mathbf{x}_m - \mathbf{x}_*\| < 1$  jollain  $m \geq 0$  niin vektorit  $\mathbf{x}_n$  suppenevat kohti ratkaisua  $\mathbf{x}_*$  ja kun  $\|\mathbf{x}_n - \mathbf{x}_*\|$  on riittävän pieni, niin etäisyys ratkaisuun pienenee hyvin nopeasti.