2.4.1. Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} 0, \text{ kun } x = 0 \text{ tai } y = 0, \\ 1, \text{ muulloin} \end{cases}$$

Selvästi $\partial_1 f(0,0) = 0 = \partial_2 f(0,0)$, sillä f saa arvon 0 koordinaattiakselien muodostamalla ristikolla. Funktio f ei kuitenkaan ole jatkuva origossa, koska sillä ei ole origossa raja-arvoa.

Edellinen esimerkki osoittaa, että osittaisderivaattojen olemassaolosta tapauksessa $n \geq 2$ ei seuraa funktion jatkuvuus kyseisessä pisteessä. Tilanne on kuitenkin korjattavissa.

Tarvitaan käsite, että funktiota f voidaan lähellä annettua pistettä approksimoida affiinilla kuvauksella. Usein tämä ilmaistaan siten, että funktion f graafia voidaan, lähellä annettua pistettä x_0, approksimoida *tangenttitasollaan*. Tämä vastaa tapauksen $n = 1$ approksimointikaavaa

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + |x - x_0|\varepsilon(x - x_0),$$

missä $\varepsilon(x - x_0) \to 0$ kun $x \to x_0$. Ylläoleva esitys kirjoitetaan usein muotoon

$$f(x_0 + h) = f(x_0) + f'(x_0)h + |h|\varepsilon(h), \; h \to 0.$$

Kuvauksen $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ graafi on suora, joka kulkee pisteen $(x_0, f(x_0))$ kautta ja jonka kulmakerroin on $f'(x_0)$. Tämä on f:n tangenttisuora pisteessä $(x_0, f(x_0))$. Kuvauksen f approksimointi tarkoittaa siis kuvauksen f approksimointia affiinilla kuvauksella

$$x \mapsto f(x_0) + f'(x_0)(x - x_0).$$

Tällä ilmiöllä on luonnollinen vastine $\mathbb{R}^n$:ssä. Suoraa $\mathbb{R}^2$:ssa vastaa nyt n-taso $\mathbb{R}^{n+1}$:ssä. Tällainen taso (joka ei ole kohtisuorassa tasoa

$x_{n+1} = 0$ vastaan) on aina kuvauksen $\mathbb{R}^n \to \mathbb{R}$,

(2.4.2) $$(x_1, \ldots, x_n) \mapsto a_0 + a_1x_1 + \ldots + a_nx_n,$$

graafi

$$G = \{(x_1, \ldots, x_n, a_0 + a_1x_1 + \ldots + a_nx_n) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in \mathbb{R}^n\}.$$

Kätevä tapa esittää taso on käyttää vektoreiden $a = (a_1, \ldots, a_n)$ ja $x = (x_1, \ldots, x_n)$ pistetuloa. Nyt kuvaus (2.4.2) saa muodon

$$x \mapsto a_0 + a \cdot x.$$

Tällaista kuvausta sanotaan *affiiniksi*. Sen graafi on n-taso avaruudessa $\mathbb{R}^{n+1}$. Affiini kuvaus saadaan yhdistämällä lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}$, $L(x) = a \cdot x$, ja $\mathbb{R}$:n siirto $t \mapsto a_0 + t$. Luonnollisesti myös kuvaus

$$x \mapsto a_0 + a \cdot (x - x_0)$$

on affiini, vaikkakaan ei yleensä sama kuvaus.

Olkoon $D \subset \mathbb{R}^n$ avoin, $f : D \to \mathbb{R}$ ja $x_0 \in D$. Asetetaan taso kulkemaan pisteen $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ kautta ja tulkitaan se kuvauksen

$$x \mapsto a_0 + a \cdot (x - x_0),$$

missä $a_0 \in \mathbb{R}$ ja $a \in \mathbb{R}^n$, graafina. Tämä kulkee pisteen $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ kautta täsmälleen silloin, kun $a_0 = f(x_0)$, eli kun

$$x \mapsto f(x_0) + a \cdot (x - x_0).$$

Tämä johtaa seuraavaan käsitteeseen: funktio $f : D \to \mathbb{R}$ on *derivoituva eli differentioituva pisteessä* $x_0 \in D$, jos on olemassa $a \in \mathbb{R}^n$, jolla pätee

(2.4.3) $$f(x) = f(x_0) + a \cdot (x - x_0) + |x - x_0|\varepsilon(x - x_0),$$

2.4.1. Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} 0, \text{ kun } x = 0 \text{ tai } y = 0, \\ 1, \text{ muulloin} \end{cases}$$

Selvästi $\partial_1 f(0,0) = 0 = \partial_2 f(0,0)$, sillä f saa arvon 0 koordinaattiakselien muodostamalla ristikolla. Funktio f ei kuitenkaan ole jatkuva origossa, koska sillä ei ole origossa raja-arvoa.

Edellinen esimerkki osoittaa, että osittaisderivaattojen olemassaolosta tapauksessa $n \geq 2$ ei seuraa funktion jatkuvuus kyseisessä pisteessä. Tilanne on kuitenkin korjattavissa.

Tarvitaan käsite, että funktiota f voidaan lähellä annettua pistettä approksimoida affiinilla kuvauksella. Usein tämä ilmaistaan siten, että funktion f graafia voidaan, lähellä annettua pistettä x_0, approksimoida *tangenttitasollaan*. Tämä vastaa tapauksen $n = 1$ approksimointikaavaa

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + |x - x_0|\varepsilon(x - x_0),$$

missä $\varepsilon(x - x_0) \to 0$ kun $x \to x_0$. Ylläoleva esitys kirjoitetaan usein muotoon

$$f(x_0 + h) = f(x_0) + f'(x_0)h + |h|\varepsilon(h), \ h \to 0.$$

Kuvauksen $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ graafi on suora, joka kulkee pisteen $(x_0, f(x_0))$ kautta ja jonka kulmakerroin on $f'(x_0)$. Tämä on f:n tangenttisuora pisteessä $(x_0, f(x_0))$. Kuvauksen f approksimointi tarkoittaa siis kuvauksen f approksimointia affiinilla kuvauksella

$$x \mapsto f(x_0) + f'(x_0)(x - x_0).$$

Tällä ilmiöllä on luonnollinen vastine $\mathbb{R}^n$:ssä. Suoraa $\mathbb{R}^2$:ssa vastaa nyt n-taso $\mathbb{R}^{n+1}$:ssä. Tällainen taso (joka ei ole kohtisuorassa tasoa

$x_{n+1} = 0$ vastaan) on aina kuvauksen $\mathbb{R}^n \to \mathbb{R}$,

(2.4.2) $$(x_1, \ldots, x_n) \mapsto a_0 + a_1x_1 + \ldots + a_nx_n,$$

graafi

$$G = \{(x_1, \ldots, x_n, a_0 + a_1x_1 + \ldots + a_nx_n) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in \mathbb{R}^n\}.$$

Kätevä tapa esittää taso on käyttää vektoreiden $a = (a_1, \ldots, a_n)$ ja $x = (x_1, \ldots, x_n)$ pistetuloa. Nyt kuvaus (2.4.2) saa muodon

$$x \mapsto a_0 + a \cdot x.$$

Tällaista kuvausta sanotaan *affiiniksi*. Sen graafi on n-taso avaruudessa $\mathbb{R}^{n+1}$. Affiini kuvaus saadaan yhdistämällä lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}$, $L(x) = a \cdot x$, ja $\mathbb{R}$:n siirto $t \mapsto a_0 + t$. Luonnollisesti myös kuvaus

$$x \mapsto a_0 + a \cdot (x - x_0)$$

on affiini, vaikkakaan ei yleensä sama kuvaus.

Olkoon $D \subset \mathbb{R}^n$ avoin, $f : D \to \mathbb{R}$ ja $x_0 \in D$. Asetetaan taso kulkemaan pisteen $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ kautta ja tulkitaan se kuvauksen

$$x \mapsto a_0 + a \cdot (x - x_0),$$

missä $a_0 \in \mathbb{R}$ ja $a \in \mathbb{R}^n$, graafina. Tämä kulkee pisteen $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ kautta täsmälleen silloin, kun $a_0 = f(x_0)$, eli kun

$$x \mapsto f(x_0) + a \cdot (x - x_0).$$

Tämä johtaa seuraavaan käsitteeseen: funktio $f : D \to \mathbb{R}$ on *derivoituva eli differentioituva pisteessä* $x_0 \in D$, jos on olemassa $a \in \mathbb{R}^n$, jolla pätee

(2.4.3) $$f(x) = f(x_0) + a \cdot (x - x_0) + |x - x_0|\varepsilon(x - x_0),$$

missä $\varepsilon(x - x_0) \to 0$ kun $x \to x_0$. Vektoria a sanotaan funktion f *gradientiksi pisteessä* x_0 ja merkitään $a = \nabla f(x_0)$. Toisaalta $h \mapsto \nabla f(x_0) \cdot h$ määrittelee lineaarikuvauksen $L : \mathbb{R}^n \to \mathbb{R}$ kaavalla $L(h) = \nabla f(x_0) \cdot h$. Tätä lineaarikuvausta merkitään yleensä $L = f'(x_0)$; siis $f'(x_0)(h) = \nabla f(x_0) \cdot h$.

2.4.4. Huomautus. Kaava (2.4.3) on aina voimassa millä hyvänsä $a \in \mathbb{R}^n$, sillä voidaan asettaa

$$\varepsilon(x - x_0) = \begin{cases} \dfrac{f(x) - f(x_0) - a \cdot (x - x_0)}{|x - x_0|} & \text{, kun } x \neq x_0,\ x \in A, \\ 0 & \text{, kun } x = x_0. \end{cases}$$

Sen sijaan ei yleensä ole totta, että $\varepsilon(x - x_0) \to 0$ kun $x \to x_0$.

2.4.5. Lause. *Olkoon f derivoituva pisteessä x_0. Tällöin*

$$\nabla f(x_0) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0)).$$

Toisin sanoen pisteessä x_0 derivoituvan funktion f osittaisderivaatat ovat kaikki olemassa pisteessä x_0 ja ne antavat gradientin $\nabla f(x_0)$.

Todistus. Lasketaan vektorin $\nabla f(x_0) = a$ ensimmäinen koordinaatti a_1. Muut koordinaatit lasketaan samoin. Olkoon $x = x_0 + he_1$. Nyt $x \in D$ kun $h \in \mathbb{R}$ on pieni. Kun $h \neq 0$, niin sijoitetaan x kaavaan (2.4.3). Tämä antaa

$$f(x_0 + he_1) - f(x_0) = a \cdot (he_1) + |h||e_1|\varepsilon(he_1).$$

Tästä seuraa, että

$$\frac{f(x_0 + he_1) - f(x_0)}{h} = a \cdot e_1 + \frac{|h|}{h}\,\varepsilon(he_1).$$

Rajankäynnillä, kun $h \to 0$, saamme $\partial_1 f(x_0) = a \cdot e_1 = a_1$. □

2.4.6. Lause. *Jos f on derivoituva pisteessä x_0, niin kaavan* (2.4.3) *vektori a on yksikäsitteisesti määrätty.*

Todistus. Seuraa lauseesta 2.4.5, koska osittaisderivaatat $\partial_i f(x_0)$ ovat yksikäsitteisesti määrättyjä. □

2.4.7. Huomautus. Kaava (2.4.3) kirjoitetaan usein muotoon

$$f(x_0 + h) = f(x_0) + \nabla f(x_0) \cdot h + |h|\varepsilon(h), \tag{2.4.8}$$

missä h on $\mathbb{R}^n$:n vektori (yleensä $|h|$ on pieni niin, että $f(x_0 + h)$ on määritelty) ja $\varepsilon(h) \to 0$ kun $h \to 0$.

2.4.9. Lause. *Olkoon $D \subset \mathbb{R}^n$ ja $x_0 \in D$. Jos funktio $f : D \to \mathbb{R}$ on derivoituva pisteessä x_0, niin f on jatkuva pisteessä x_0.*

Todistus. On näytettävä, että $\lim_{x \to x_0} f(x) = f(x_0)$, toisin sanoen on osoitettava, että ehdosta $x_i \to x_0$, $x_i \in D$ seuraa $f(x_i) \to f(x_0)$. Nyt

$$f(x_i) = f(x_0) + \nabla f(x_0) \cdot (x_i - x_0) + |x_i - x_0|\ \varepsilon(x_i - x_0)$$

ja siis

$$\begin{aligned} |f(x_i) - f(x_0)| &\leq |\nabla f(x_0) \cdot (x_i - x_0)| + |x_i - x_0||\varepsilon(x_i - x_0)| \\ &\leq |\nabla f(x_0)||(x_i - x_0)| + |x_i - x_0||\varepsilon(x_i - x_0)| \to 0. \end{aligned}$$

Tässä käytettiin Cauchy-Schwartzin epäyhtälöä, kaavaa (1.1.3). Tästä seuraa $f(x_i) \to f(x_0)$, kuten vaadittiinkin. □

2.4.10. Huomautus. Vaikka funktiolla f on osittaisderivaatat

$$\partial_1 f(x_0), \ldots, \partial_n f(x_0),$$

niin f ei välttämättä ole derivoituva x_0:ssa, sillä osittaisderivaatat voivat olla olemassa, vaikka f ei ole edes jatkuva pisteessä x_0.

Jos tiedetään, että f on derivoituva pisteessä x_0, ei ole mitään vaikeuksia määrätä gradienttia $\nabla f(x_0)$, sillä

$$\nabla f(x_0) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0)).$$

Pitää siis vain laskea osittaisderivaatat. Sen sijaan on vaikeampi päätellä, milloin f on derivoituva x_0:ssa. Huomaa, että yleensä kirjallisuudessa määritellään gradientti yllä olevalla kaavalla, kunhan vain osittaisderivaatat ovat olemassa pisteessä x_0. Gradientin olemassaolo ei siis edellytä derivoituvuutta.

2.4.11. Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(xy)$. Tällöin $\partial_1 f(x, y) = \partial_x f(x, y) = y\cos(xy)$ ja $\partial_2 f(x, y) = \partial_y f(x, y) = x\cos(xy)$. Siten

$$\nabla f(x, y) = (y\cos(xy), x\cos(xy)).$$

Annamme pian ehdon funktion f derivoituvuudelle pisteessä x_0. Ehto osoittaa muun muassa, että esimerkin 2.4.11 funktio f on derivoituva $\mathbb{R}^2$:n jokaisessa pisteessä.

Jos f on derivoituva pisteessä x_0, niin tasoa

$$x_{n+1} = f(x_0) + \partial_1 f(x_0)(x_1 - x_{0,1}) + \ldots + \partial_n f(x_0)(x_n - x_{0,n})$$

sanotaan f:n *tangenttitasoksi pisteessä* $(x_0, f(x_0))$. Tämä on n-taso $\mathbb{R}^{n+1}$:ssä. Tätä tasoa kannattaa yleensä ajatella funktion $\mathbb{R}^n \to \mathbb{R}$

$$x \mapsto f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

graafina $\mathbb{R}^{n+1}$:ssä. Graafi on taso, joka kulkee pisteen $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ kautta ja on kohtisuorassa $\mathbb{R}^{n+1}$:n vektoria

$$(\nabla f(x_0), -1) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0), -1)$$

vastaan.

2.4.12. Huomautus. Tapaus $n = 1$:

$$y = f(x_0) + f'(x_0)(x - x_0),$$

missä $y = x_{n+1}$. Tämän suoran normaali on $\mathbb{R}^2$:n vektori $(f'(x_0), -1)$.

Jos f derivoituva x_0:ssa, niin f:ää voidaan approksimoida pisteen x_0 pienessä ympäristössä affiinilla kuvauksella

$$x \mapsto f(x_0) + \nabla f(x_0) \cdot (x - x_0), \tag{2.4.13}$$

katso kaava (2.4.2). Tämä on pisteessä x_0 derivoituvan funktion f tärkein ominaisuus. Tähän ei riitä osittaisderivaattojen eli gradientin olemassaolo pisteessä x_0, vaan approksimointi vaatii funktion f derivoituvuuden pisteessä x_0. Huomaa, että kuvaus (2.4.13) on määritelty koko $\mathbb{R}^n$:ssä.

2.4.14. Huomautus. Kuvaus $h \mapsto \nabla f(x_0) \cdot h$ on lineaarinen kuvaus $\mathbb{R}^n \to \mathbb{R}$; tähän palataan kappaleessa 2.6. Itse asiassa lineaarialgebran tietojen nojalla jokainen lineaarinen kuvaus $L : \mathbb{R}^n \to \mathbb{R}$ on muotoa $L(h) = a \cdot h$ jollakin $a \in \mathbb{R}^n$.

HARJOITUSTEHTÄVIÄ

2.4:1 Määritä funktion $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$, $f(x) = |x|^{|x|}$, gradientti.

2.4:2 Osoita, että funktio $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x) = \begin{cases} x_1^3/|x|, \text{ kun } x \neq 0, \\ 0, \text{ kun } x = 0, \end{cases}$$

on derivoituva origossa.

2.4:3 Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$ ja $\partial_1 f = 0 = \partial_2 f$. Osoita, että f on vakiofunktio.

2.4:4 Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 - 4xy - 2y^2 + 12x - 12y - 1$. Missä pisteissä $(x_0, y_0) \in \mathbb{R}^2$ funktion f tangenttitaso on horisontaalinen eli missä pisteissä affiinin kuvauksen

$$(x,y) \mapsto f(x_0,y_0) + \nabla f(x_0,y_0) \cdot (x,y)$$

graafi on (x,y)-tason suuntainen?

2.4:5 Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = (x, xe^y)$. Määritä $\nabla f(1,1)$. Onko olemassa pisteitä $(x,y) \in \mathbb{R}^2$, missä $\nabla f(1,1) = (0,0)$?

2.5. Derivoimissääntöjä

Osittaisderivaattojen laskeminen tapahtuu samoilla säännöillä kuin yksiulotteisessa tapauksessa. Kuitenkin nämä säännöt yhdistettyjen funktioiden tapauksessa edellyttävät usein derivoituvuutta (eivät pelkästään osittaisderivaattojen olemassaoloa), joten tutkimme ensin seuraavaa kysymystä. Milloin voidaan päätellä, että funktio $f : D \to \mathbb{R}$ on derivoituva pisteessä $x_0 \in D$?

2.5.1. Määritelmä. Olkoon $D \subset \mathbb{R}^n$ avoin ja $f : D \to \mathbb{R}$. Funktio f *kuuluu luokkaan* $C^1(D)$ (kerran jatkuvasti derivoituvat funktiot D:ssä), jos

a) osittaisderivaatat $\partial_1 f(x_0), \ldots, \partial_n f(x_0)$ ovat olemassa D:n jokaisessa pisteessä ja
b) nämä ovat jatkuvia D:ssä.

Usein yllä olevassa määritelmässä oletetaan, että myös f on jatkuva joukossa D. Tämä on kuitenkin tarpeetonta. Seuraavan lauseen nojalla ehdoista a) ja b) seuraa funktion f jatkuvuus, sillä derivoituva funktio on aina jatkuva, kuten lauseessa 2.4.9 osoitettiin.

2.5.2. Lause. *Jos $f \in C^1(D)$, niin f on derivoituva D:n jokaisessa pisteessä.*

Todistus tapauksessa $n = 2$. Olkoon $f : D \to \mathbb{R}$ ja $x_0 = (a,b) \in D$. On näytettävä, että f on derivoituva pisteessä x_0, toisin sanoen

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + |x - x_0|\, \varepsilon(x - x_0), \tag{2.5.3}$$

missä $\varepsilon(x - x_0) \to 0$ kun $x \to x_0$. Huomaa, että $\nabla f(x_0)$ on olemassa, koska olettamuksen mukaan osittaisderivaatat $\partial_1 f(x_0)$ ja $\partial_2 f(x_0)$ ovat olemassa. Kuten aiemmin totesimme, kaava (2.5.3) pitää aina paikkansa; ainoa ongelma on funktion ε raja-arvo.

Kiinnitetään neliö $Q \subset D$, jonka sivujen pituus on $2r$ ja jonka keskipiste on $x_0 = (a,b)$. Olkoon $x = (a+h, b+k) \in Q$, missä $(h,k) \neq 0$. Tällöin $|h|, |k| < r$. Nyt pätee

$$\begin{aligned} f(x) - f(x_0) &= f(a+h, b+k) - f(a,b) \\ &= f(a+h, b+k) - f(a, b+k) + f(a, b+k) - f(a,b). \end{aligned}$$

Kiinteällä k funktio $\varphi(t) = f(a+t, b+k) - f(a, b+k)$ on määritelty välillä $(-r, r)$ ja derivoituva kyseisellä välillä, sillä $\varphi'(t) = \partial_1 f(a + t, b+k)$. Soveltamalla tavallista väliarvolausetta välillä $[0,h]$ (tai välillä $[h,0]$) saadaan

$$\begin{aligned} f(a+h, b+k) - f(a, b+k) &= \varphi(h) - \varphi(0) \\ &= \varphi'(\xi)(h - 0) = \partial_1 f(a + \xi, b+k)h. \end{aligned}$$

Samoin soveltamalla väliarvolausetta funktioon

$$\psi(s) = f(a, b+s) - f(a,b)$$

saadaan

$$f(a,b+k)-f(a,b)=\psi(k)-\psi(0)$$
$$=\psi'(\theta)(k-0)=\partial_2 f(a,b+\theta)k.$$

Yhdistämällä nämä aikaisempaan yhtälöön saadaan

$$f(x)-f(x_0)=\partial_1 f(a+\xi,b+k)h+\partial_2 f(a,b+\theta)k$$
$$=(\partial_1 f(a+\xi,b+k),\partial_2 f(a,b+\theta))\cdot(h,k).$$

Luku ξ on 0:n ja h:n välissä ja luku θ on vastaavasti 0:n ja k:n välissä; huomaa, että h ja k voivat olla myös negatiivisia. Soveltamalla tätä kaavaan (2.5.3) saadaan funktiolle ε lauseke

$$\varepsilon(x-x_0)=\frac{1}{|x-x_0|}\,[(\partial_1 f(a+\xi,b+k),\ \partial_2 f(a,b+\theta))$$
$$-\ (\partial_1 f(x_0),\ \partial_2 f(x_0))]\cdot(h,k)$$
$$=\frac{1}{|x-x_0|}\,[(\partial_1 f(a+\xi,b+k)-\partial_1 f(x_0),\ \partial_2 f(a,b+\theta)-\partial_2 f(x_0))\cdot(h,k)].$$

Koska $x-x_0=(h,k)$, tästä seuraa

$$|\varepsilon(x-x_0)|\le\frac{1}{|(h,k)|}\,|(\partial_1 f(a+\xi,b+k)-\partial_1 f(x_0),\ \partial_2 f(a,b+\theta)$$
$$-\ \partial_2 f(x_0)||(h,k)|$$
$$=|(\partial_1 f(a+\xi,b+k)-\partial_1 f(x_0),\partial_2 f(a,b+\theta)-\partial_2 f(x_0)|\to 0,$$

kun $(h,k)\to 0$. Tämä pätee, koska $\partial_1 f$ ja $\partial_2 f$ ovat jatkuvia pisteessä $x_0=(a,b)$; tällöin $\partial_1 f(a+\xi,b+k)\to\partial_1 f(x_0)$ ja $\partial_2 f(a,b+\theta)\to\partial_2 f(x_0)$, kun $(h,k)\to 0$. Tästä seuraa väite. □

2.5.4. Esimerkki. Määrätään funktion $f:\mathbb{R}^n\to\mathbb{R}$,

$$f(x)=|x|=\sqrt{x_1^2+\ldots+x_n^2},$$

gradientti $\nabla f(x)$ ja tutkitaan funktion f derivoituvuutta. Osittaisderivaattoja ei ole olemassa pisteessä $x=0$. Jätämme tämän todistamisen harjoitustehtäväksi; tämä vastaa yksiulotteista tapausta, jossa funktiolla $x\mapsto|x|$ ei ole derivaattaa origossa. Nyt funktio f ei ole derivoituva pisteessä 0. Lasketaan $\nabla f(x)$, kun $x\neq 0$. Kiinnitetään $i=1,2,\ldots,n$ ja tutkitaan yhden muuttujan, x_i, funktiota

$$x_i\mapsto\sqrt{x_1^2+\ldots+x_i^2+\ldots+x_n^2}.$$

Laskemalla tämän derivaatta saadaan

$$\partial_i f(x)=1/2\,(x_1^2+\ldots+x_n^2)^{-1/2}\,2x_i=\frac{x_i}{|x|}.$$

Funktio f on derivoituva joukossa $\mathbb{R}^n\backslash\{0\}$. Syy tähän on seuraava. Funktio $x\mapsto|x|$ on jatkuva $\mathbb{R}^n$:ssä, mistä seuraa tapausta $n=1$ vastaavasti, että $x\mapsto 1/|x|$ on jatkuva joukossa $\mathbb{R}^n\backslash\{0\}$. Funktio $x\mapsto x_i$ on selvästi jatkuva $\mathbb{R}^n$:ssä ja kahden jatkuvan funktion tulo on jatkuva, joten funktio $x\mapsto x_i/|x|$ on jatkuva joukossa $\mathbb{R}^n\backslash\{0\}$. Tämä pätee kaikilla $i=1,2,\ldots,n$. Lauseen 2.5.2 nojalla f on siten derivoituva joukossa $\mathbb{R}^n\backslash\{0\}$. Funktion f gradientille $\nabla f(x)$ saadaan pisteessä $x\neq 0$ lauseke

$$\nabla f(x)=(\partial_1 f(x),\ldots,\partial_n f(x))=\left(\frac{x_1}{|x|},\ldots,\frac{x_n}{|x|}\right)$$
$$=\frac{1}{|x|}(x_1,\ldots,x_n)=\frac{x}{|x|}.$$

2.5.5. Huomautus. Tapaus $n=1$: Olkoon $f:\mathbb{R}\to\mathbb{R}$, $f(x)=|x|$. Tällöin

$$f'(x)=\nabla f(x)=\frac{x}{|x|}=\begin{cases}=1, \text{ kun } x>0,\\ =-1, \text{ kun } x<0.\end{cases}$$

Tarkastellaan seuraavassa f:n osittaisderivaattojen määräämistä tapauksessa, kun f on yhdistetty funktio, eli johdetaan niin sanotut yhdistetyn funktion derivoimisen *ketjusäännöt*. Tarkastelemme tilannetta ensin erikoistapausten kautta. Yleinen tilanne käsitellään kappaleessa 2.6.

A) Olkoon $D \subset \mathbb{R}^n$ avoin, $\Delta \subset \mathbb{R}$ väli, $h : D \to \Delta$ ja $g : \Delta \to \mathbb{R}$. Tarkastelemme funktion $f = g \circ h$ derivoimista.

Tämä on helppoa, sillä osittaisderivaatat lasketaan rajoittamalla funktio $\mathbb{R}^n$:n koordinaattiakselien suuntaisille suorille. Määrätään tässä tapauksessa $\partial_1 f$. Määritelmän mukaan

$$\partial_1 f(x_0) = \lim_{t \to 0} \frac{g \circ h(x_0 + te_1) - g \circ h(x_0)}{t}.$$

Koska $x_0 \in D$ ja D on avoin, funktio

$$\varphi(t) = g \circ h(x_0 + te_1)$$

on määritelty jollakin välillä $(-\delta, \delta)$, $\delta > 0$. Laskemalla tämän funktion derivaatta 0:ssa tavallisella yksiulotteisella ketjusäännöllä saadaan

$$\begin{aligned} \varphi'(0) &= g'(h(x_0 + 0e_1)) \cdot \partial_1 h(x_0 + 0e_1) \\ &= g'(h(x_0))\partial_1 h(x_0). \end{aligned}$$

Toisaalta

$$\begin{aligned} \varphi'(0) &= \lim_{t \to 0} \frac{g \circ h(x_0 + te_1) - g \circ h(x_0)}{t} \\ &= \partial_1 f(x_0). \end{aligned}$$

Havaitaan siis, että $\partial_1 f(x_0) = g'(h(x_0))\partial_1 h(x_0)$, jos derivaatat $\partial_1 h(x_0)$ ja $g'(h(x_0))$ ovat olemassa. Yleisesti saadaan ketjusääntö

$$\partial_i f(x_0) = g'(h(x_0))\, \partial_i h(x_0), \; i = 1, \ldots, n.$$

2.5.6. Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \sin(xy)$. Nyt $f = g \circ h$, missä $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x,y) = xy$ ja $g : \mathbb{R} \to \mathbb{R}$, $g(s) = \sin s$. Määrätään $\partial_2 f(x,y)$. Ylläolevasta kaavasta saadaan

$$\begin{aligned} \partial_2 f(x,y) &= g'(h(x,y)) \cdot \partial_2 h(x,y) \\ &= \cos(h(x,y))x = \cos(xy)x. \end{aligned}$$

Sama tulos saadaan tietysti suoraan kiinnittämällä x ja derivoimalla y:n suhteen.

B) Olkoot $D, D' \subset \mathbb{R}^n$ avoimia, $w : D \to D'$ ja $h : D' \to \mathbb{R}$. Tarkastelemme funktion $f = h \circ w$ derivoimista.

Kuvauksia eli funktioita w, joiden arvot ovat $\mathbb{R}^n$:ssä, ei ole aikaisemmin käsitelty. Nämä palautetaan reaaliarvoiseen tapaukseen seuraavasti. Kuvauksen $w : D \to D'$ määritelmän mukaan jokaiseen $x \in D$ on liitetty yksikäsitteinen $w(x) \in D'$, toisin sanoen $w(x) = (y_1, \ldots, y_n)$, missä y_i on $w(x)$:n koordinaatti. Siten w määrää n kappaletta funktioita $y_i = w_i(x)$, $x \in D$. Nämä ovat kuvauksia $w_i : D \to \mathbb{R}$, toisin sanoen $w(x) = (w_1(x), \ldots, w_n(x))$, missä w_i on kuvauksen w *i. koordinaattifunktio*. Siten

$$f(x) = h(w(x)) = h(w_1(x), \ldots, w_n(x)).$$

2.5.7. Lause. *Olkoot funktiot w_i derivoituvia pisteessä $x_0 \in D$ ja h derivoituva pisteessä $w(x_0)$. Tällöin f on derivoituva pisteessä x_0 ja*

$$\partial_i f(x_0) = \sum_{j=1}^{n} \partial_j h(w(x_0))\partial_i w_j(x_0), \; i = 1, 2, \ldots, n.$$

Todistus. Todistetaan kaava tapauksessa $i = 1$. Muut indeksit i menevät samoin. On näytettävä, että

$$\partial_1 f(x_0) = \sum_{j=1}^{n} \partial_j h(w(x_0))\partial_1 w_j(x_0).$$

Määritelmän nojalla

$$\partial_1 f(x_0) = \lim_{t\to 0} \frac{f(x_0 + te_1) - f(x_0)}{t}.$$

Olkoon $t \neq 0$ niin pieni, että $x_0 + te_1 \in D$. Nyt

$$\frac{f(x_0 + te_1) - f(x_0)}{t} = \frac{h(w(x_0 + te_1)) - h(w(x_0))}{t}.$$

Koska h on derivoituva pisteessä $w(x_0)$, pätee jokaisella $y \in D'$

$$h(y) - h(w(x_0)) = \nabla h(w(x_0)) \cdot (y - w(x_0)) + |y - w(x_0)|\, \varepsilon(y - w(x_0)),$$

vertaa kaava (2.4.8). Asettamalla $y = w(x_0 + te_1) \in D'$ ja jakamalla puolittain t:llä saadaan

$$\begin{aligned}
&\frac{h(w(x_0 + te_1)) - h(w(x_0))}{t} \\
&= \nabla h(w(x_0)) \cdot \frac{w(x_0 + te_1) - w(x_0)}{t} \\
&\qquad + \frac{|w(x_0 + te_1) - w(x_0)|}{t}\, \varepsilon(w(x_0 + te_1) - w(x_0)).
\end{aligned}$$

Koordinaattifunktioiden avulla esittämällä saadaan

$$\begin{aligned}
&\frac{w(x_0 + te_1) - w(x_0)}{t} \\
&= \left(\frac{w_1(x_0 + te_1) - w_1(x_0)}{t}, \dots, \frac{w_n(x_0 + te_1) - w_n(x_0)}{t} \right) \\
&\to (\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0)),
\end{aligned}$$

kun $t \to 0$. Yhdistämällä tämä aikaisempaan kaavaan ja antamalla $t \to 0$ havaitaan, että

$$\begin{aligned}
&\partial_1 f(x_0) \\
&= \nabla h(w(x_0)) \cdot (\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0)) + |(\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0))| \cdot 0 \\
&= \sum_{j=1}^{n} \partial_j h(w(x_0))\partial_1 w_j(x_0).
\end{aligned}$$

Tämä on vaadittu summakaava. Funktion f derivoituvuus seuraa suhteellisen helposti funktioiden h ja w_i, $i = 1, \dots, n$, derivoituvuudesta. Jätämme todistamisen harjoitustehtäväksi. □

2.5.8. Esimerkki. Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = \sin(\sin(x)e^{xy}).$$

Osittaisderivaatta $\partial_1 f(x, y)$ on mahdollista määrittää tavalliseen tapaan kiinnittämällä y ja derivoimalla x:n suhteen. Käytetään kuitenkin ketjusääntöä asettamalla

$$\begin{aligned}
h(x, y) &= \sin(xy), \\
w(x, y) &= (w_1(x, y), w_2(x, y)) = (\sin x, e^{xy}).
\end{aligned}$$

Tällöin

$$\begin{aligned}
(h \circ w)(x, y) &= h(w(x, y)) = h(w_1(x, y), w_2(x, y)) \\
&= \sin(\sin(x)e^{xy}) = f(x, y).
\end{aligned}$$

Suoraan laskemalla saadaan funktion h gradientiksi

$$\begin{aligned}
\nabla h(x, y) &= (\partial_1 h(x, y), \partial_2 h(x, y)) \\
&= (y \cos(xy), x \cos(xy))
\end{aligned}$$

ja osittaisderivaatoiksi

$$\partial_1 w_1(x,y) = \cos x, \quad \partial_1 w_2(x,y) = ye^{xy}.$$

Kaikki funktiot ovat jatkuvia, joten voidaan käyttää ketjusääntöä. Se antaa

$$\partial_1 f(x,y) = e^{xy}\cos(\sin xe^{xy}) \cdot \cos x + \sin x \cos(\sin xe^{xy}) \cdot ye^{xy}.$$

C) Seuraava tilanne tulee usein vastaan. Olkoon $\Delta \subset \mathbb{R}$ väli, $D \subset \mathbb{R}^n$ avoin joukko ja $g : D \to \mathbb{R}$ sekä $\gamma : \Delta \to D$ kuvauksia. Funktio γ on siis kuvaus väliltä Δ avoimeen joukkoon $D \subset \mathbb{R}^n$. Jokaisella $t \in D$ kuvauksen γ arvo pisteessä t on $\mathbb{R}^n$:n vektori

$$(\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t)),$$

toisin sanoen γ määrittelee koordinaattifunktiot $\gamma_i : \Delta \to \mathbb{R}$, $i = 1, \ldots, n$. Kuvausta γ sanotaan jatkuvaksi, jos jokainen γ_i on jatkuva, ja jatkuvaa kuvausta $\gamma : \Delta \to D$ sanotaan *poluksi* joukossa D. Kuvauksen γ derivaatta $\gamma'(t)$ pisteessä $t \in \Delta$ määritellään vektoriksi

$$\gamma'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t)),$$

missä $\gamma_i'(t)$ on tavallinen derivaatta. Tämä tietenkin edellyttää, että jokainen derivaatoista $\gamma_i'(t)$ on olemassa pisteessä t. Nyt on määritettävä tavallisen reaaliarvoisen funktion $u : \Delta \to \mathbb{R}$ derivaatta pisteessä $x_0 \in \Delta$, kun u on muotoa $u = g \circ \gamma$.

2.5.9. Lause. *Olkoon $t_0 \in \Delta$, $\gamma'(t_0)$ olemassa ja $g : D \to \mathbb{R}$ derivoituva pisteessä $\gamma(t_0)$. Tällöin*

$$h'(t_0) = \nabla g(\gamma(t_0)) \cdot \gamma'(t_0).$$

Todistus. Harjoitustehtävä. □

Polkuja ja muita vektoriarvoisia funktioita käsitellään tarkemmin luvussa 3.

HARJOITUSTEHTÄVIÄ

2.5:1 Todista väliarvolause avaruudessa $\mathbb{R}^n$: Olkoon $D \subset \mathbb{R}^n$, $f \in C^1(D)$, $x, y \in D$ ja jana

$$I = \{ty + (1-t)x : t \in [0,1]\} \subset D.$$

Tällöin on olemassa sellainen $\xi \in I$, että

$$f(y) - f(x) = \nabla f(\xi) \cdot (y - x).$$

Vihje: Tarkastele funktiota $\varphi(t) = f(ty + (1-t)x)$, $t \in [0,1]$.

2.5:2 Muotoile ongelma virheen arvioinnista käyttäen hyväksi edellistä tehtävää tapauksessa $D \subset \mathbb{R}^3$ ja ratkaise se.

2.5:3 Kuinka monta desimaalia luvuista π ja e on tunnettava, jotta luku e^π saadaan laskettua tarkkuudella $0,001$? Pyri karkeaan, mutta varmasti riittävään arvioon.

2.5:4 Olkoon $\gamma : \mathbb{R} \to \mathbb{R}^2$, $\gamma(t) = (e^t, \cos t)$, ja $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = xy$. Määritä $f \circ \gamma$ ja $(f \circ \gamma)'(t)$.

2.5:5 Määritä edellisen tehtävän derivaatta $(f \circ \gamma)'(t)$ ketjusäännön avulla.

2.5:6 Todista lause 2.5.9.

2.5:7 Olkoot $D, D' \subset \mathbb{R}^n$ avoimia, $w : D \to D'$ ja $h : D' \to \mathbb{R}$. Oletetaan, että kuvauksen w koordinaattifunktiot $w_i : D \to \mathbb{R}$, $i = 1, \ldots, n$, ovat derivoituvia pisteessä $x_0 \in D$ ja että funktio h on derivoituva pisteessä $w(x_0) \in D'$. Osoita, että kuvaus $f : D \to \mathbb{R}$, $f = h \circ w$ on derivoituva pisteessä $x_0 \in D$.

2.6. Derivaatta ja ketjusääntö yleisessä tilanteessa

Olkoon $D \subset \mathbb{R}^n$ ja $f : D \to \mathbb{R}^p$, $p \geq 1$, kuvaus. Tällöin jokaisella $x \in D$

$$f(x) = (f_1(x), \ldots, f_p(x)),$$

missä $f_j : D \to \mathbb{R}$, $j = 1, \ldots, p$, ovat f:n *koordinaattifunktiot*. Kuvausta f sanotaan *jatkuvaksi pisteessä* $x_0 \in D$, jos jokaisella joukon D jonolla (x_i), jolla $x_i \to x_0$, pätee $f(x_i) \to f(x_0)$. Kuten aikaisemmin todettiin, tämä on yhtäpitävää sen kanssa, että jokainen koordinaattifunktio f_j, $j = 1, \ldots, p$, on jatkuva pisteessä x_0. Kuvaus on *jatkuva joukossa* D, jos se on jatkuva D:n jokaisessa pisteessä.

Oletetaan, että D on $\mathbb{R}^n$:n avoin osajoukko ja $x_0 \in D$. Kuvaus $f : D \to \mathbb{R}^p$ on *derivoituva* eli *differentioituva* pisteessä x_0, jos on olemassa sellainen lineaarikuvaus

$$L : \mathbb{R}^n \to \mathbb{R}^p,$$

että

$$f(x_0 + h) - f(x_0) = L(h) + |h|\varepsilon(h), \tag{2.6.1}$$

missä $\varepsilon(h) \to 0$, kun $h \to 0$. Siis jos (h_i) on sellainen jono vektoreita, että $h_i \to 0$ ja $h_i \neq 0$, niin $\varepsilon(h_i) \to 0$. Tämä on yhtäpitävää sen kanssa, että jokaisella $\varepsilon > 0$ on olemassa sellainen $\delta > 0$, että

$$|\varepsilon(h)| < \varepsilon, \text{ kun } |h| < \delta. \tag{2.6.2}$$

Joukon D avoimuudesta seuraa, että ε on määritelty ainakin jossain pallossa $B(0, r)$, $r > 0$. Asettamalla $\varepsilon(0) = 0$ havaitaan, että ehto (2.6.2) on yhtäpitävä sen kanssa, että ε on jatkuva origossa. Kuten

aikaisemminkin, on syytä huomata, että kaava (2.6.1) pätee jokaiselle lineaarikuvaukselle $L : \mathbb{R}^n \to \mathbb{R}^p$; derivoituvuus sen sijaan riippuu ehdosta $\varepsilon(h) \to 0$ kun $h \to 0$.

Käytämme seuraavassa lineaarialgebran kurssin tietoja, kun käsittelemme lineaarikuvauksia $L : \mathbb{R}^n \to \mathbb{R}^p$. Muista, että kuvaus $L : \mathbb{R}^n \to \mathbb{R}^p$ on *lineaarinen*, jos $L(\alpha h + \beta h') = \alpha L(h) + \beta L(h')$ kaikilla $h, h' \in \mathbb{R}^n$ ja $\alpha, \beta \in \mathbb{R}$.

Jos f on derivoituva pisteessä x_0, niin lineaarikuvausta L kaavassa (2.6.1) merkitään $f'(x_0)$:lla. Seuraava lause ei ole vaikea todistaa; sen todistus on sama kuin tapauksessa $p = 1$, vertaa lause 2.4.9.

2.6.3. Lause. *Olkoon $f : D \to \mathbb{R}^p$ derivoituva pisteessä $x_0 \in D$. Tällöin*

(i) *f on jatkuva pisteessä x_0,*
(ii) *lineaarikuvaus $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ on yksikäsitteisesti määrätty,*
(iii) *koordinaattifunktiot f_j, $j = 1, \ldots, p$, ovat derivoituvia pisteessä x_0.* □

Ehdosta (iii) seuraa myös kuvauksen f derivoituvuus pisteessä x_0.

Oletetaan, että f on derivoituva pisteessä x_0. Tällöin $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ on lineaarikuvaus, ja jokaisella lineaarikuvauksella on esitysmatriisi, joka tässä tapauksessa on $(p \times n)$-matriisi

$$f'(x_0) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^p$ standardikannassa. Tämä tarkoittaa, että kun $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, niin $f'(x_0)\,h$ on $\mathbb{R}^p$:n vektori

$$\begin{aligned} f'(x_0)\,h &= \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \\ a_{p1} & \dots & a_{pn} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \\ &= (a_{11}\,h_1 + \dots + a_{1n}\,h_n, \ \dots, \ a_{p1}\,h_1 + \dots + a_{pn}\,h_n). \end{aligned}$$

Edellä on merkitty, kuten lineaarikuvausten yhteydessä on tavallista, $f'(x_0)(h) = f'(x_0)\,h$, toisin sanoen kuvauksen argumentin ympäriltä jätetään sulut pois.

Jos f on derivoituva pisteessä x_0 ja lineaarikuvauksella $f'(x_0)$ on edellä mainittu esitysmatriisi, niin määräämme seuraavaksi luvut a_{ji}, $1 \le j \le p$, $1 \le i \le n$, funktion f koordinaattifunktioiden f_j osittaisderivaattojen avulla lausuttuna. Valitaan kaavassa (2.6.1) $h = te_1$, $t \neq 0$ ja $L = f'(x_0)$. Nyt kuvauksen $f'(x_0)$ lineaarisuuden takia

$$f'(x_0)(te_1) = t\,f'(x_0)\,e_1$$

ja toisaalta

$$f(x_0 + te_1) - f(x_0) = (f_1(x_0 + te_1) - f_1(x_0), \dots, f_p(x_0 + te_1) - f_p(x_0)).$$

Kaava (2.6.1) saadaan, jakamalla puolittain t:llä, muotoon

$$\begin{aligned} &\left(\frac{1}{t}\,(f_1(x_0 + te_1) - f_1(x_0)), \dots, \frac{1}{t}\,(f_p(x_0 + te_1) - f_p(x_0))\right) \\ &= f'(x_0)\,e_1 + \frac{|t|}{t}\,\varepsilon(te_1) \end{aligned}$$

$$\begin{aligned} &= \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \\ a_{p1} & \dots & a_{pn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} + \frac{|t|}{t}\,\varepsilon(te_1) \\ &= (a_{11}, a_{21}, \dots, a_{p1}) + \frac{|t|}{t}\,\varepsilon(te_1). \end{aligned}$$

Koska $\varepsilon(te_1) \to 0$, kun $t \to 0$, ja koska koordinaattifunktioilla f_j on osittaisderivaatat

$$\frac{1}{t}\,(f_j(x_0 + te_1) - f_j(x_0)) \to \partial_1 f_j(x_0),$$

niin antamalla $t \to 0$ nähdään, että vektorit

$$(\partial_1 f_1(x_0), \partial_1 f_2(x_0), \dots, \partial_1 f_p(x_0))$$

ja $(a_{11}, a_{21}, \dots, a_{p1})$ ovat samat ja siis $a_{j1} = \partial_1 f_j(x_0)$ jokaisella $j = 1, \dots, p$. Vastaavasti jokaisella $i = 1, \dots, n$ saadaan

$$a_{ji} = \partial_i f_j(x_0),$$

missä $1 \le i \le n$ ja $1 \le j \le p$. Siten lineaarikuvauksen $f'(x_0)$ esitysmatriisi on muotoa

$$f'(x_0) = \begin{bmatrix} \partial_1 f_1(x_0) & \partial_2 f_1(x_0) & \dots & \partial_n f_1(x_0) \\ \partial_1 f_2(x_0) & \partial_2 f_2(x_0) & \dots & \partial_n f_2(x_0) \\ \vdots & & \ddots & \\ \partial_1 f_p(x_0) & \partial_2 f_p(x_0) & \dots & \partial_n f_p(x_0) \end{bmatrix}. \tag{2.6.4}$$

Yllä on itse asiassa todistettu aikaisempi tulos, lause 2.4.5, yleisemmässä muodossa. Lauseen 2.4.5 kaavassa vaakarivejä on vain yksi.

Samoin kuin aiemmin yksinkertaisemmassa tapauksessa pätee, että kaikkien koordinaattifunktioiden derivoituvuudesta seuraa f:n derivoituvuus. Lisäksi pätee, että jos kaikki funktiot $f_j \in C^1(D)$, $1 \le j \le p$, niin f on derivoituva D:n jokaisessa pisteessä. Näiden todistaminen ei ole oleellisesti vaikeampaa kuin aikaisempien tulosten.

Jos $p = 1$ eli $f : D \to \mathbb{R}$, olemme määritelleet derivoituvuuden pisteessä $x_0 \in D$ kaavalla (2.4.3), missä vektori $a = \nabla f(x_0)$. Derivoituvuus on tässä tapauksessa täsmälleen sama kuin yllä määritelty. Syy on se, että lineaarikuvaus $f'(x_0) : \mathbb{R}^n \to \mathbb{R}$ tulee muotoon

$$f'(x_0)\, h = \nabla f(x_0) \cdot h, \tag{2.6.5}$$

mikä nähdään suoraan esityksestä (2.6.4). Kuten aiemmin todettiin, on jokainen lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}$ muotoa $Lh = a \cdot h$, $h \in \mathbb{R}^n$, jollakin kiinteällä $a \in \mathbb{R}^n$.

Tutkimme lopuksi ketjusääntöä yleisessä tapauksessa. Olkoon $D \subset \mathbb{R}^n$ ja $D' \subset \mathbb{R}^p$ avoimia joukkoja sekä $f : D \to D'$ ja $g : D' \to \mathbb{R}^q$ kuvauksia. Seuraava lause ei ole vaikeampi todistaa kuin aikaisemmat ketjusäännöt.

2.6.6. Lause. *Jos f on derivoituva pisteessä x_0 ja g on derivoituva pisteessä $f(x_0)$, niin yhdistetty kuvaus $g \circ f : D \to \mathbb{R}^q$ on derivoituva pisteessä x_0 ja*

$$(g \circ f)'(x_0) = g'(f(x_0))\, f'(x_0). \tag{2.6.7}$$

□

Kaava (2.6.7) tarkoittaa, että lineaarikuvaus $(g \circ f)'(x_0) : \mathbb{R}^n \to \mathbb{R}^q$ saadaan yhdistämällä lineaarikuvaukset $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ ja $g'(f(x_0)) : \mathbb{R}^p \to \mathbb{R}^q$; kuten tavallista, ei yhdistettyjen lineaarikuvausten välissä käytetä merkintää $\circ$. Lineaarikuvausten $(g \circ f)'(x_0)$ esitysmatriisi saadaan kertomalla $g'(f(x_0))$:n ja $f'(x_0)$:n esitysmatriisit. Kaikki aikaisemmat ketjusäännöt ovat seurauksia lauseesta 2.6.6.

2.6.8. Esimerkki. Olkoot $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (\sin x, xy)$ ja $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x,y) = xy$. Tällöin $g \circ f : \mathbb{R}^2 \to \mathbb{R}$,

$$(g \circ f)(x,y) = g(f(x,y)) = g(\sin x, xy) = xy \sin x.$$

Nyt g, f, $g \circ f \in C^1(\mathbb{R}^2)$, joten kaikki kuvaukset ovat derivoituvia. Laskemme gradientin $\nabla(g \circ f)(x,y)$ lauseen 2.6.6 kaavasta

$$(g \circ f)'(x,y) = \underbrace{\underbrace{g'(f(x,y))}_{\text{lineaarikuvaus } \mathbb{R}^2 \to \mathbb{R}} \circ \underbrace{f'(x,y)}_{\text{lineaarikuvaus } \mathbb{R}^2 \to \mathbb{R}^2}}_{\text{lineaarikuvausten yhdistetty lineaarikuvaus } \mathbb{R}^2 \to \mathbb{R}}$$

käyttämällä derivaatan ja gradientin välistä yhteyttä (2.6.5),

$$(g \circ f)'(x,y)h = \nabla(g \circ f)(x,y) \cdot h,$$

missä $h \in \mathbb{R}^2$. Kaavasta 2.6.4 saadaan lineaarikuvauksen $f'(x,y)$ esitysmatriisiksi

$$f'(x,y) = \begin{bmatrix} \partial_1 f_1(x,y) & \partial_2 f_1(x,y) \\ \partial_1 f_2(x,y) & \partial_2 f_2(x,y) \end{bmatrix} = \begin{bmatrix} \cos x & 0 \\ y & x \end{bmatrix}.$$

Käyttämällä kaavaa (2.6.5) kuvaukseen g saadaan

$$\begin{aligned} \nabla g(x,y) \cdot h = g'(x,y)h &= (\partial_1 g(x,y), \partial_2 g(x,y)) \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \\ &= (y,x) \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = yh_1 + xh_2 = (y,x) \cdot h. \end{aligned}$$

Lasketaan seuraavaksi yhdistetyn lineaarikuvauksen $g'(f(x,y))\circ f'(x,y)$ arvo vektorilla $h \in \mathbb{R}^2$; tämä tulee suoraan edellisistä kaavoista. Saadaan

$$\begin{aligned} g'(f(x,y))\, f'(x,y)\, h &= g'(f(x,y))(f'(x,y)\, h) \\ &= (xy, \sin x) \left(\begin{bmatrix} \cos x & 0 \\ y & x \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right) \\ &= (xy, \sin x) \begin{bmatrix} (\cos x)\, h_1 \\ y\, h_1 + x\, h_2 \end{bmatrix} \\ &= (xy\,(\cos x)\, h_1 + (\sin x)(y\, h_1 + x\, h_2) \\ &= (xy\ \cos x + y\ \sin x) h_1 + x\,(\sin x)\, h_2 \\ &= (xy\ \cos x + y\ \sin x, x\ \sin x) \cdot (h_1, h_2). \end{aligned}$$

Saadaan siis $\nabla(g \circ f)(x,y) = (xy\ \cos x + y\ \sin x, x\ \sin x)$. Saman tuloksen saa helpommin laskemalla suoraan osittaisderivaatat.

HARJOITUSTEHTÄVIÄ

2.6:1 Olkoon $f : \mathbb{R}^2 \to \mathbb{R}^2$ ja $f = (f_1, f_2)$. Osoita, että jos f_1 ja f_2 ovat derivoituvia pisteessä $x_0 \in \mathbb{R}^2$, niin f on derivoituva pisteessä x_0.

2.6:2 Määritä kuvauksen $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (e^y, xy)$, derivaatan $f'(x,y)$ esitysmatriisi pisteessä $(x,y) = (1,1)$.

2.7. SUUNNATTU DERIVAATTA JA GRADIENTIN GEOMETRINEN MERKITYS

Käsittelemme asiaa vain, kun $n = 2$. Yleistys tapaukseen $n \geq 3$ on helppo. Reaaliarvoisen funktion osittaisderivaatat pisteessä x_0 on otettu koordinaattiakselien suuntaan. Yhtä hyvin osittaisderivaatta voidaan määritellä mihin suuntaan hyvänsä.

Olkoon $f : D \to \mathbb{R}$, D avoin $\mathbb{R}^2$:ssa ja $x_0 \in D$ sekä $v \in \mathbb{R}^2$, $|v| = 1$ eli v on yksikkövektori. Funktion f:n *suunnattu derivaatta pisteessä* x_0 *suuntaan* v on

$$f'_v(x_0) = \partial_v f(x_0) = D_v f(x_0) = \lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

mikäli kyseinen raja-arvo on olemassa. Havaitaan, että

$$\partial_1 f(x_0) = \partial_{e_1} f(x_0), \ \partial_2 f(x_0) = \partial_{e_2} f(x_0).$$

2.7.1. Lause. *Olkoon f derivoituva pisteessä x_0. Tällöin*

$$\partial_v f(x_0) = \nabla f(x_0) \cdot v$$

ja

$$|\partial_v f(x_0)| \leq |\nabla f(x_0)|.$$

2.7.2. Huomautus. Lauseen 2.7.1 epäyhtälössä vasen puoli on reaaliluvun itseisarvo ja oikea on vektorin pituus.

Lauseen 2.7.1 todistus. Koska f on derivoituva pisteessä x_0, pätee

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