3.1. Konveksia analyysiä

Joukko $A \subseteq \mathbb{R}^d$ on affiini, jos $\forall x, y \in A$, $\lambda \in \mathbb{R}$,

$$(1-\lambda)x + \lambda y \in A.$$

Siis A sisältää kaikki pisteidensä kautta kulkevat suorat. Selvästi affiini joukko on aina konveksi. On suoraviivaista nähdä, että jos L on $\mathbb{R}^d$:n aliavaruus ja $x \in \mathbb{R}^d$, niin

$$(3.3) \qquad A = x + L = \{x + y \mid y \in L\}$$

on affiini. Kääntäen, jokainen affiini A on muotoa (3.3) ja L on esityksessä yksikäsitteinen. Affiinin joukon dimensio on $\dim L$, jos A on muotoa (3.3).

Selvästi affiinien joukkojen leikkaus on aina affiini. Joukon $S \subseteq \mathbb{R}^d$ affiini verho $\mathcal{A}(S)$ on

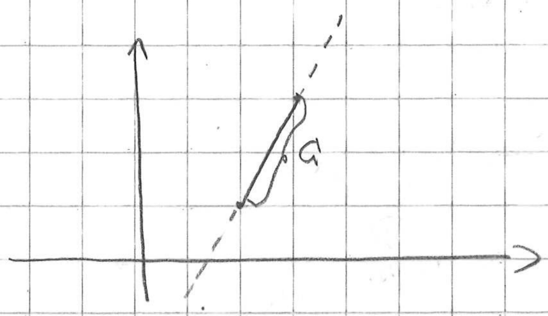
$$\mathcal{A}(S) = \bigcap \{A \subseteq \mathbb{R}^d \mid S \subseteq A,\ A \text{ affiini}\}.$$

Ilmeisesti $\mathcal{A}(S)$ on suppein $\mathbb{R}^d$:n affiini osajoukko, joka sisältää S:n.

Olkoon $C \subseteq \mathbb{R}^d$ konveksi. Merkitään symbolilla $\mathrm{ri}\, C$ (relative interior) C:n sisäosaa affiinissa verhossa $\mathcal{A}(C)$. Siis $x \in \mathrm{ri}\, C$ jos ja vain jos on olemassa $\delta > 0$:

$$y \in C \text{ aina, kun } y \in \mathcal{A}(C) \text{ ja } |x-y| < \delta$$

eli $B(x, \delta) \cap \mathcal{A}(C) \subseteq C$.



$\mathbb{R}^2$:ssa $\mathring{C} = \emptyset$ ja $\mathrm{ri}\, C$ = ko. jana pl. päätepisteet. $\mathcal{A}(C)$ saadaan jatkamalla janaa kulkuviivan mukaisesti.

C:n <u>dimensiolla</u> tarkoitetaan $\mathcal{A}(C)$:n dimensiota.

<u>Lemma 3.5.</u> Olkoon C ei-tyhjä konveksi joukko. Tällöin

(i) $\mathrm{ri}\, C \neq \emptyset$,

(ii) $x \in \mathrm{ri}\, C$ jos ja vain jos jokaiselle $z \in C$ on olemassa $\mu = \mu_z > 1$:

$$(1-\mu) z + \mu x \in C.$$

(iii) Jos $x \in \mathrm{ri}\, C$ ja $y \in \overline{C}$, niin

$$(1-\lambda) x + \lambda y \in \mathrm{ri}\, C, \ \forall \lambda \in [0,1).$$

(iv) Olkoot C_1 ja C_2 konvekseja joukkoja. Jos

$$\mathrm{ri}\, C_1 \cap \mathrm{ri}\, C_2 \neq \emptyset$$

niin

$$\overline{C_1 \cap C_2} = \overline{C}_1 \cap \overline{C}_2.$$

(v) Jos $x_0 \in \mathrm{ri}\, C$, niin $\exists\, x_1, \dots, x_r \in C$ ja $\delta > 0$:

$$B(x_0, \delta) \cap C \subseteq C_0, \quad \text{missä}$$

$$C_0 = \{\lambda_1 x_1 + \dots + \lambda_r x_r \mid \lambda_1, \dots, \lambda_r \geq 0, \ \lambda_1 + \dots + \lambda_r = 1\}.$$

Todistus. (ei käsitellä luennolla). (i). Olkoon ensin $\mathcal{A}(C)$ lineaariavaruus ja $0 \in C$. Olkoon m maksimaalinen määrä lineaarisesti riippumattomia vektoreita ja $\{x_1, \dots, x_m\} \subseteq C$ eräs tällainen vektorijoukko. Olkoon edelleen

$$M = \mathcal{A}(\{0, x_1, \dots, x_m\}) \subseteq \mathcal{A}(C).$$

Yleisesti: jos affiini joukko A sisältää nollavektorin, niin A on lineaariavaruus. Jos nimittäin $x \in A$, $\lambda \in \mathbb{R}$, niin

$$(1-\lambda)0 + \lambda x = \lambda x \in A$$

ja jos $x, y \in A$, niin $\frac{1}{2}(x+y) = (1-\frac{1}{2})x + \frac{1}{2}y \in A$, joten $x+y = 2(\frac{1}{2}(x+y)) \in A$.

Nähdään, että M on lineaariavaruus ja $\dim M = m$. Ilmeisesti

$$M = \{\lambda_1 x_1 + \dots + \lambda_m x_m \mid \lambda_1, \dots, \lambda_m \in \mathbb{R}\}.$$

Nyt $\mathcal{A}(C) \subseteq M$: jos olisi $x \in C \setminus M$, niin x ei olisi vektoreiden $x_1, \dots, x_m$ lineaariyhdistelmä. Tämä on vastoin m:n maksimaalisuusoletusta.

Siis $G \subseteq M$, joten $A(G) \subseteq M$ ja itse asiassa $A(G) = M$. Erityisesti

$$\dim G = \dim M = m.$$

Jos $m = 0$ (eli $G = \{0\}$), on asia selvä: $0 \in \mathrm{ri}\, G$. Muutoin edellä esitetyn nojalla G sisältää pisteet $\{0, x_1, \dots, x_m\}$ ja konveksina joukon

$$G_0 = \{\lambda_0 \cdot 0 + \lambda_1 x_1 + \dots + \lambda_m x_m \mid \lambda_0, \dots, \lambda_m \geq 0,\ \lambda_0 + \dots + \lambda_m = 1\}$$

$$= \{\lambda_1 x_1 + \dots + \lambda_m x_m \mid \lambda_1, \dots, \lambda_m \geq 0,\ \lambda_1 + \dots + \lambda_m \leq 1\}.$$

Tarkastellaan lineaarikuvausta $T : M \to \mathbb{R}^m$,

$$T(\lambda_1 x_1 + \dots + \lambda_m x_m) = (\lambda_1, \dots, \lambda_m).$$

Tällöin T on jatkuva bijektio ja

$$T(G_0) = \{(\lambda_1, \dots, \lambda_m) \mid \lambda_1, \dots, \lambda_m \geq 0,\ \lambda_1 + \dots + \lambda_m \leq 1\}.$$

Tällä on ei-tyhjä sisäosa $\mathbb{R}^m$:ssä, nimittäin

$$\{(\lambda_1, \dots, \lambda_m) \mid \lambda_1, \dots, \lambda_m > 0,\ \lambda_1 + \dots + \lambda_m < 1\}.$$

Myös T^{-1} on lineaarinen ja siis jatkuva. Näin ollen joukolla

$$T^{-1}(T(G_0)) = G_0$$

on ei-tyhjä sisäosa M:ssä, joten $\mathrm{ri}\, G \neq \emptyset$.

Jos $A(G)$ ei ole lineaariavaruus, sovelletaan alkuosaa joukkoon $G - x_0$, missä $x_0 \in G$. Jos $x \in \mathrm{ri}(G - x_0)$, niin $x + x_0 \in \mathrm{ri}\, G$. Kohta (i) on todistettu.

Kohdan (ii) todistamiseksi oletetaan ensin, että $x \in \operatorname{ri} C$. Jos $z \in C$, niin

$$\{(1-\lambda)z + \lambda x \mid \lambda \in \mathbb{R}\} \subseteq \mathcal{A}(C).$$

Koska $x \in \operatorname{ri} C$, voidaan määrätä $\varepsilon > 0$ siten, että

$$B(x, \varepsilon) \cap \mathcal{A}(C) \subseteq C.$$

Siis $(1-\mu)z + \mu x \in C$ jollakin $\mu > 1$.

Olkoon nyt x sellainen, että kohdan (ii) μ on olemassa kaikilla $z \in C$. Valitaan erityisesti $z \in \operatorname{ri} C$ (mikä on mahdollista kohdan (i) nojalla). Olkoon $\varepsilon > 0$ sellainen, että

$$B(z, \varepsilon) \cap \mathcal{A}(C) \subseteq C.$$

Merkitään vielä $w = (1-\mu)z + \mu x$, $\mu > 1$, $w \in C$.
Jos $|x - y| < \varepsilon'$, $y \in \mathcal{A}(C)$, niin

$$\left| z - \frac{w - \mu y}{1-\mu} \right| = \left| \frac{\mu}{1-\mu} \right| |x - y| < \varepsilon,$$

kun valitaan $\varepsilon' < \frac{\mu - 1}{\mu}$. Koska $w \in C$, niin

$\frac{w - \mu y}{1-\mu} \in \mathcal{A}(C)$. Siispä $\frac{w - \mu y}{1-\mu} \in B(z, \varepsilon) \cap \mathcal{A}(C)$, joten

$$\frac{w - \mu y}{1-\mu} \doteq c \in C.$$

Edelleen $y = \frac{1}{\mu} w + \frac{\mu - 1}{\mu} c \in C$. Nähdään, että

$$B(x, \varepsilon') \cap \mathcal{A}(C) \subseteq C,$$

joten $x \in \operatorname{ri} C$.

Kohdan (iii) todistus on samantapainen. On osoitettava, että voidaan määrätä sellainen $\varepsilon > 0$, että

$$(3.4)\quad w \in \mathcal{A}(G),\ |(1-\lambda)x + \lambda y - w| < \varepsilon \Rightarrow w \in G,$$

Olkoon $\varepsilon > 0$. Määritään sellainen $z \in G$, että $|z - y| < \varepsilon$. Jos myös (3.4) toteutuu (paitsi että $w \in G$), niin

$$|(1-\lambda)x + \lambda z - w| \le (1+\lambda)\varepsilon,$$

joten

$$\left| x - \frac{w - \lambda z}{1-\lambda} \right| \le \frac{(1+\lambda)\varepsilon}{1-\lambda}.$$

Koska $w, z \in \mathcal{A}(G)$ ja $x \in \operatorname{ri} G$, niin riittävän pienellä $\varepsilon > 0$,

$$\frac{w - \lambda z}{1-\lambda} \in G.$$

Siis $w = \lambda z + (1-\lambda)c$, $c \in G$, joten $w \in G$.

Todistetaan kohta (iv). Olkoon $x \in \operatorname{ri} G_1 \cap \operatorname{ri} G_2$ kiinteä. Jos $y \in \bar{G}_1 \cap \bar{G}_2$, niin kohdan (iii) nojalla, $\forall \lambda \in [0,1)$

$$(1-\lambda)x + \lambda y \in \operatorname{ri} G_1 \cap \operatorname{ri} G_2 \subseteq G_1 \cap G_2.$$

Koska $y = \lim_{\lambda \nearrow 1}((1-\lambda)x + \lambda y)$, niin $y \in \overline{G_1 \cap G_2}$. Siis

$$\bar{G}_1 \cap \bar{G}_2 \subseteq \overline{G_1 \cap G_2}.$$

Olkoon nyt $y \in \overline{G_1 \cap G_2}$. Tällöin kin

$$(1-\lambda)x + \lambda y \in G_1,\ G_2,$$

joten $y \in \bar{G}_1$, $y \in \bar{G}_2$. Siis $\overline{G_1 \cap G_2} \subseteq \bar{G}_1 \cap \bar{G}_2$.

(v) Voidaan olettaa, että G ei ole singleton. Olkoon $\dim G = m \geq 1$.

Oletetaan ensin, että $A(G)$ on lineaariavaruus. Tällöin

$$A(G) = \{\lambda_1 y_1 + \dots + \lambda_m y_m \mid \lambda_1, \dots, \lambda_m \in \mathbb{R}\}$$

eräille $y_1, \dots, y_m \in \mathbb{R}^d$. Määritellään lineaarikuvaus $T : A(G) \to \mathbb{R}^m$ ehdosta

$$T(\lambda_1 y_1 + \dots + \lambda_m y_m) = (\lambda_1, \dots, \lambda_m).$$

Tällöin T ja T^{-1} ovat jatkuvia bijektioita. Merkitään

$$G' = T(G), \quad x_0' = T(x_0).$$

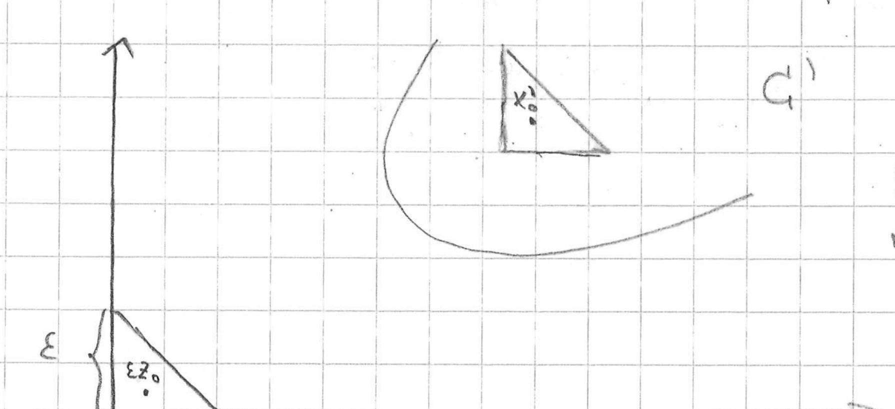
Tällöin x_0' on G':in sisäpiste $\mathbb{R}^m$:ssä. Olkoon e_i $\mathbb{R}^m$:in yksikkövektori, jonka i. komponentti on $=1$ (ja muut nollia). Olkoon $z_i = e_i$, $i=1,\dots,m$, $z_{m+1} = 0$ ja $n = m+1$. Tällöin

$$K \doteq \{\lambda_1 z_1 + \dots + \lambda_n z_n \mid \lambda_1 \geq 0, \dots, \lambda_n \geq 0, \lambda_1 + \dots + \lambda_n = 1\}$$

$$= \{(t_1, \dots, t_m) \in \mathbb{R}^m \mid t_1 \geq 0, \dots, t_m \geq 0, t_1 + \dots + t_m \leq 1\}.$$

Olkoon $\varepsilon > 0$ ja $\varepsilon K = \{\varepsilon z \mid z \in K\}$. Selvästi $\mathring{K} \neq \emptyset$ ja jos $z_0 \in \mathring{K}$, niin $\varepsilon z_0 \in \varepsilon \mathring{K}$. Valitaan ε pieneksi ja $\delta > 0$ siten, että

$$B(\varepsilon z_0, \delta) \subseteq \varepsilon \mathring{K}, \quad x_0' + \varepsilon z - \varepsilon z_0 \in G', \quad \forall z \in K.$$



Siirto tehdään kuvauksella

$\varepsilon z \mapsto x_0' + \varepsilon z - \varepsilon z_0$

$(\varepsilon K \mapsto G')$

Valitaan $x_i' = x_0' + \varepsilon z_i - \varepsilon z_0$, $i=1,\dots,r$. Jos $|x' - x_0'| < \delta$, niin

$$x' = x_0' - \varepsilon z_0 + \underbrace{(x' - x_0' + \varepsilon z_0)}_{\in B(\varepsilon z_0, \delta) \subseteq \varepsilon \mathring{K}}$$

$$= x_0' - \varepsilon z_0 + \varepsilon(\lambda_1 z_1 + \dots + \lambda_r z_r) \quad (\lambda_1, \dots, \lambda_r \geq 0,\ \lambda_1 + \dots + \lambda_r = 1)$$

$$= \lambda_1 x_1' + \dots + \lambda_r x_r'.$$

Siis

$$B(x_0', \delta) \subseteq \{\lambda_1 x_1' + \dots + \lambda_r x_r' \mid \lambda_1, \dots, \lambda_r \geq 0,\ \lambda_1 + \dots + \lambda_r = 1\},$$

missä $x_1', \dots, x_r' \in G'$. Jos nyt $x_i = T^{-1}(x_i')$, $\forall i$, ja $x \in T^{-1}(B(x_0', \delta))$, niin x on muotoa

$$x = \lambda_1 T^{-1}(x_1') + \dots + \lambda_r T^{-1}(x_r') = \lambda_1 x_1 + \dots + \lambda_r x_r.$$

Koska $T^{-1}(B(x_0', \delta))$ on $\mathcal{A}(G)$:n avoin joukko ja $x_0 \in T^{-1}(B(x_0, \delta))$, niin

$$B(x_0, \delta_0) \cap \mathcal{A}(G) \subseteq T^{-1}(B(x_0', \delta))$$

eräälle $\delta > \delta_0 > 0$. Lisäksi $x_i \in G$, $\forall i$, joten on saatu vaadittu tulos.

Jos $\mathcal{A}(G)$ ei ole lineaariavaruus, pätee tulos joukolle $G - x_0$ (sillä $0 \in \mathcal{A}(G - x_0)$, joten $\mathcal{A}(G - x_0)$ on lineaariavaruus). Jos $x_1, \dots, x_r$ ovat lemman mukaiset vektorit koskien pistettä $0 \in \mathrm{ri}(G - x_0)$ niin $x_0 + x_1, \dots, x_0 + x_r$ ovat vaaditut vektorit pisteelle $x_0 \in \mathrm{ri}\, G$. $\square$

Lemma 3.6. Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi. Jos $x_0 \in \mathrm{ri}\, D(f)$, niin $\exists\, \delta > 0$ ja $\alpha \in (0,\infty)$:

(3.4.1) $$f(x) \le \alpha \,, \quad \forall\, x \in B(x_0,\delta) \cap D(f).$$

Lisäksi $\underline{f}(x) > -\infty$, $\forall\, x \in \mathbb{R}^d$.

Todistus. Lemman 3.5 kohdan (iv) nojalla $\exists\, x_1,\dots,x_n \in D(f)$ ja $\delta > 0$:

$$B(x_0,\delta) \cap D(f) \subseteq \{\lambda_1 x_1 + \dots + \lambda_n x_n \mid \lambda_1,\dots,\lambda_n \ge 0,\ \lambda_1 + \dots + \lambda_n = 1\}$$

Jos siis $x \in B(x_0,\delta) \cap D(f)$, niin

$$f(x) \le \lambda_1 f(x_1) + \dots + \lambda_n f(x_n)$$
$$\le \max(0, f(x_1), \dots, f(x_n)),$$

josta lemman ensimmäinen väite seuraa.

Jos $\underline{f}(x) = f(x)$ tai $f(x) = +\infty$, niin $\underline{f}(x) > -\infty$. Muuten

$$\underline{f}(x) = \lim f(x_n)$$

eräälle jonolle $(x_n) \in D(f)$, $x_n \to x$. Olkoon $x_0 \in \mathrm{ri}\, D(f)$ ja $\delta > 0$: (3.4.1) pätee. Määrätään $\mu_n > 1$:

$$(1-\mu_n) x_n + \mu_n x_0 =: y_n \in B(x_0,\delta) \cap D(f),$$

lemma 3.5, (ii). (μ_n voidaan valita mielivaltaisen läheltä ykköstä). Tällöin

$$x_0 = (1-\lambda_n) x_n + \lambda_n y_n,$$

$\lambda_n = \frac{1}{\mu_n} \in (0,1)$. Siispä

$$f(x_0) \leq (1-\lambda_n) f(x_n) + \lambda_n f(y_n)$$

$$\leq (1-\lambda_n) f(x_n) + \lambda_n \alpha \leq (1-\lambda) f(x_n) + \alpha.$$

Nähdään, että $f(x_n)$ on alhaalta rajoitettu, sillä voidaan valita esim. $\lambda_n = \frac{1}{2}$, $\forall n$. $\square$

<u>Lemma 3.7.</u> Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi ja $\mathring{D}(f) \neq \emptyset$. Silloin

$$\mathring{E}(f) = \{(x, M) \mid x \in \mathring{D}(f),\ f(x) < M < \infty\}.$$

<u>Todistus.</u> Olkoon $x_0 \in \mathring{D}(f)$ ja $M_0 > f(x_0)$. Osoitetaan, että

$$(x_0, M_0) \in \mathring{E}(f).$$

Olkoot $\delta > 0$ ja $\alpha > 0$ kuten lemmassa 3.6. Tällöin

$$(3.5) \quad \{(x, R) \mid x \in B(x_0, \delta),\ R > \alpha\} \subseteq \mathring{E}(f).$$

Jos $R_0 > \alpha$, niin

$$(x_0, M_0) = (1 - \lambda)(x_0, R_0) + \lambda (x_0, f(x_0))$$

eräälle $\lambda \in [0, 1)$. Lemman 3.5 kohdan (iii) nojalla

$$(x_0, M_0) \in \mathring{E}(f).$$

Selvää on, että jos $(x_0, M_0) \in \mathring{E}(f)$, niin $x_0 \in \mathring{D}(f)$ ja $M_0 > f(x_0)$. $\square$

<u>Huomautus.</u> Edellä on käytetty tietoa: (A konveksi)

$$\mathring{A} \neq \emptyset \Rightarrow \operatorname{ri} A = \mathring{A}.$$

Jos nimittäin $0 \in \mathring{A}$, niin $\varepsilon e_1, \ldots, \varepsilon e_d \in \mathring{A}$, ε pieni $\Rightarrow \mathcal{A}(A) = \mathbb{R}^d$ (koska $\mathcal{A}(A)$ on lineaariavaruus). Jos $x_0 \in \mathring{A}$, on siis $\mathcal{A}(A - x_0) = \mathbb{R}^d \Rightarrow \mathcal{A}(A) = \mathbb{R}^d$.

<u>Lemma 3.8.</u> Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi ja $\underline{f}$ f:n alaverho. Oletetaan, että $\mathring{D}(f) \neq \emptyset$. Silloin $\underline{f}$ on apj ja konveksi ja

$$\underline{f}(x) = f(x), \quad \forall x \in \mathring{D}(f) \cup (\bar{D}(f))^c.$$

<u>Todistus.</u> Olkoon $x_0 \in \mathring{D}(f)$ ja

$$B = \{(x_0, M) \mid M \in \mathbb{R}\}.$$

Lemman 3.7 nojalla $B \cap \mathring{E}(f) \neq \emptyset$. Lemman 3.5 kohdan (iv) nojalla

$$B \cap \bar{E}(f) = \bar{B} \cap \bar{E}(f) = \overline{B \cap E(f)} = B \cap E(f).$$

Ilmeisesti $\bar{E}(f)$ on $\underline{f}$:n epigraafi, joten $\underline{f}(x) = f(x)$.

Olkoon $x_0 \in (\bar{D}(f))^c$. Tällöin $f(x) = +\infty$ eräässä pisteen x_0 ympäristössä, joten

$$\underline{f}(x_0) = \lim_{\delta \to 0+} \{\inf\{f(x) \mid |x - x_0| < \delta\}\} = \infty = f(x_0).$$

Kuten edellä todettiin,

$$\bar{E}(f) = E(\underline{f}).$$

On helppo nähdä, että konveksin joukon sulkeuma on konveksi. Siis $\underline{f}$ on konveksi (Lemman 3.6 nojalla $\underline{f}(x) > -\infty$, $\forall x$). Triviaalisti $\underline{f}$ on apj. $\square$

Lemman 3.4 todistus. On siis näytettävä, että konveksi funktio f on jatkuva $\mathring{D}(f)$:ssä. Olkoon $x_0 \in \mathring{D}(f)$. Koska f on apj, niin lemman 3.8 nojalla

$$\liminf_{x \to x_0} f(x) = \liminf_{x \to x_0} \underline{f}(x) \geq \underline{f}(x_0) = f(x_0).$$

Riittää siis näyttää, että

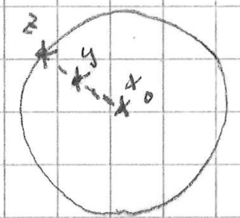
$$(3.6) \qquad \limsup_{x \to x_0} f(x) \leq f(x_0).$$

Lemman 3.6 nojalla voidaan valita $\delta > 0$ ja $\alpha > 0$:

$$f(x) \leq \alpha, \quad \forall x \in \bar{B}(x_0, \delta).$$

Olkoon $\varepsilon \in (0, \delta)$ ja $|y - x_0| = \varepsilon$. Merkitään

$$z = x_0 + \frac{\delta}{\varepsilon}(y - x_0).$$



Tällöin $|z - x_0| = \delta$ ja

$$y = \left(1 - \frac{\varepsilon}{\delta}\right) x_0 + \frac{\varepsilon}{\delta} z.$$

Siispä

$$f(y) \leq \left(1 - \frac{\varepsilon}{\delta}\right) f(x_0) + \frac{\varepsilon}{\delta} \alpha,$$

joten

$$\sup_{0 < |x_0 - y| \leq \varepsilon} f(y) \leq f(x_0) + \frac{\varepsilon}{\delta} |f(x_0)| + \frac{\varepsilon}{\delta} \alpha.$$

Tästä seuraa (3.6). $\square$

3.20.1.

Kuvaus $h : \mathbb{R}^d \to \mathbb{R}$ on <u>affiini</u>, jos

$$h(x) = \langle b, x \rangle + a \quad , \forall x \in \mathbb{R}^d,$$

joillain $a \in \mathbb{R}$, $b \in \mathbb{R}^d$.

Lemma 3.9. Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ apj ja konveksi. Silloin f on majoroimiensa affiinisten kuvausten pisteittäinen supremum ts.

$$(3.7) \quad f(x) = \sup\{h(x) \mid h \le f,\ h : \mathbb{R}^d \to \mathbb{R} \text{ affiini}\}.$$

(erityisesti f majoroi ainakin yhtä affiinista funktiota).

Lemman todistuksessa hyödynnetään seuraavaa tulosta.

Lemma 3.10. Olkoot $G_1, G_2 \subseteq \mathbb{R}^d$ ei-tyhjiä erillisiä konvekseja joukkoja. Oletetaan, että G_1 on suljettu ja G_2 kompakti. Silloin on olemassa lineaarikuvaus $T : \mathbb{R}^d \to \mathbb{R}$ ja vakiot $\gamma_1, \gamma_2 \in \mathbb{R}$ siten, että

$$T(x) > \gamma_1 > \gamma_2 > T(y), \quad \forall x \in G_1,\ y \in G_2.$$

Todistus on esitetty esim. lähteessä

Rudin (1985) Functional Analysis (kyse on ns. Hahn-Banach teoreemoista).

Kuvauksen T avulla voidaan siis erotella joukot G_1 ja G_2.

Lemman 3.9 todistus. Voidaan olettaa, että $\mathcal{E}(f) \neq \emptyset$. Sovelletaan lemmaa 3.10 epigraafiin $\mathcal{E}(f)$, joka on oletuksen nojalla suljettu. Olkoon $(x_0, M_0) \in \mathcal{E}(f)^c$ mielivaltainen, $x_0 \in \mathbb{R}^d$, $M_0 \in \mathbb{R}$. Valitaan

$$C_1 = \mathcal{E}(f) \quad \text{ja} \quad C_2 = \{(x_0, M_0)\}.$$

Olkoon T lemman mukainen lineaarikuvaus. Tällöin

(3.8) $\quad T(x, M) = \langle b, x \rangle + \alpha M$

eräälle $b \in \mathbb{R}^d$, $\alpha \in \mathbb{R}$. Olkoon $x_0 \in D(f)$. Koska $(x_0, M) \in \mathcal{E}(f)$, $\forall M \geq f(x_0)$ ja

$$T(x_0, M) > \text{vakio aina, kun } M \geq f(x_0),$$

on välttämättä $\alpha \geq 0$. Koska $x_0 \in D(f)$, on

$$T(x_0, f(x_0)) = \langle b, x_0 \rangle + \alpha f(x_0)$$

$$> \langle b, x_0 \rangle + \alpha M_0 .$$

Nähdään, että välttämättä $\alpha > 0$. Olkoon

(3.9) $\quad h(x) = -\alpha^{-1} \langle b, x \rangle + \alpha^{-1} \langle b, x_0 \rangle + M_0 .$

Jos $x \in D(f)$, niin $T(x, f(x)) > T(x_0, M_0)$, joten

$$\alpha f(x) > -\langle b, x \rangle + \langle b, x_0 \rangle + \alpha M_0 .$$

Siis $h(x) \leq f(x)$, $\forall x \in D(f)$, joten $h(x) \leq f(x)$, $\forall x \in \mathbb{R}^d$.

On siis ainakin yksi affiini funktio, jota f majoroi.

Osoitetaan, että (3.7) on tosi, jos $x = x_0 \in D(f)$. Jos $M_0 < f(x_0)$, niin edellä todetun nojalla kaavan (3.9) mukainen h on f:n majorantti $((x_0, M_0) \in \mathcal{E}(f)^c$, koska $M_0 < f(x_0))$. Koska

$$(3.10) \qquad h(x_0) = M_0, \quad h \leq f,$$

on kaavan (3.7) supremum vähintään M_0. Koska $M_0 < f(x_0)$ on mielivaltainen, on ko. supremum $= f(x_0)$ kuten väitettiin.

Olkoon nyt $x = x_0 \in D(f)^c$. Tällöin $(x_0, M_0) \in \mathcal{E}(f)^c$ kaikilla $M_0 > 0$. Lemma 3.10 soveltuu edelleen, joten erottava lineaarikuvaus (3.8) on nytkin olemassa ja $\alpha \geq 0$. Jos $\alpha > 0$, niin kaavan (3.9) mukainen h toteuttaa ehdot (3.10).

Olkoon nyt $\alpha = 0$ ts. $T(x, M) = \langle b, x \rangle$. Lemman 3.10 nojalla eräälle $\varepsilon > 0$,

$$T(x, M) > \langle b, x_0 \rangle + \varepsilon, \quad \forall x \in D(f), \ M \geq f(x).$$

Olkoon

$$h_1(x) = -\langle b, x \rangle + \langle b, x_0 \rangle + \varepsilon.$$

Tällöin

$$h_1(x) < 0, \ \forall x \in D(f), \quad h_1(x_0) = \varepsilon > 0.$$

Olkoon $h_2 \leq f$, h_2 affiini (tällainen on olemassa alkuosan nojalla: lähdetään liikkeelle jostain $D(f)$:n pisteestä). Olkoon $\lambda > 0$ mielivaltainen ja

$$h(x) = \lambda h_1(x) + h_2(x), \quad \forall x \in \mathbb{R}^d.$$

Tällöin h on alfiini ja

$$\begin{cases} h(x) \le h_2(x) \le f(x), & \forall x \in D(f), \\ h(x_0) = \lambda\varepsilon + h_2(x_0). \end{cases}$$

Antamalla $\lambda \to \infty$ nähdään, että $h(x_0) > M_0$.

Tapauksessa $\alpha > 0$ siis (3.10) toteutuu eräälle alfiinille $h \le f$ ja tapauksessa $\alpha = 0$

$$h(x_0) \ge M_0, \quad h \le f$$

eräälle alfiinille kuvaukselle h. Koska M_0 on mielivaltainen, on

$$\sup\{h(x_0) \mid h \le f,\ h \text{ alfiini}\} = \infty = f(x_0). \quad \square$$

Lemma 3.10. Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi. Silloin

$$(3.11) \qquad f^* = \underline{f}^*$$

ja

$$(3.12) \qquad \underline{f} = (f^*)^*.$$

Erityisesti $f = (f^*)^*$, jos f on lisäksi apj.

Todistus. Koska $\underline{f}(t) \le f(t)$, $\forall t \in \mathbb{R}^d$, niin

$$\underline{f}^*(x) = \sup_{t \in \mathbb{R}^d} \{ \langle t, x \rangle - \underline{f}(t) \}$$

$$\ge \sup_{t \in \mathbb{R}^d} \{ \langle t, x \rangle - f(t) \} = f^*(x), \ \forall x \in \mathbb{R}^d.$$

Olkoon nyt $x \in \mathbb{R}^d$ kiinteä. Tuloksen (3.11) todistamiseksi on osoitettava, että $\underline{f}^*(x) \le f^*(x)$. Oletetaan ensin, että $\underline{f}^*(x) < \infty$. Olkoon $\varepsilon > 0$ ja $t_\varepsilon \in \mathbb{R}^d$:

$$\underline{f}^*(x) \le \langle t_\varepsilon, x \rangle - \underline{f}(t_\varepsilon) + \varepsilon.$$

Tällöin

$$f^*(x) \ge \lim_{\delta \to 0+} \Big\{ \sup_{|t - t_\varepsilon| \le \delta} \{ \langle t, x \rangle - f(t) \} \Big\}$$

$$= \langle t_\varepsilon, x \rangle - \lim_{\delta \to 0+} \Big\{ \inf_{|t - t_\varepsilon| \le \delta} f(t) \Big\}$$

$$= \langle t_\varepsilon, x \rangle - \underline{f}(t_\varepsilon) \ge \underline{f}^*(x) - \varepsilon.$$

Siis $f^*(x) \ge \underline{f}^*(x)$. Tapaus $\underline{f}^*(x) = \infty$ jätetään harjoitustehtäväksi.

On vielä todistettava (3.12). Riittää todistaa väite tapauksessa, jossa f on apj. Sovelletaan lemmaa 3.9. Olkoon $b \in \mathbb{R}^d$ kiinteä. Selvästi

$$\langle b, t \rangle + a \leq f(t), \quad \forall t \in \mathbb{R}^d,$$

jollekin $a \in \mathbb{R}^d$ jos ja vain jos

$$\sup_{t \in \mathbb{R}^d} \{ \langle b, t \rangle - f(t) \} < \infty$$

eli jos ja vain jos $f^*(b) < \infty$. Siis lemmassa 3.9

$$f(t) = \sup \{ \langle b, t \rangle + a \mid b \in D(f^*), \langle b, t \rangle + a \leq f(t), \forall t \}.$$

Kun $b \in D(f^*)$ on kiinteä, voidaan määritä $a = a(b)$, joka antaa maksimaalisen affiinin funktion edellisessä supremumissa, valitaan nimittäin

$$a(b) = \sup \{ a \mid \langle b, t \rangle + a \leq f(t), \forall t \}$$

$$= \sup \{ a \mid a \leq \inf_{t \in \mathbb{R}^d} \{ f(t) - \langle b, t \rangle \} \}$$

$$= -f^*(b).$$

Siis

$$f(t) = \sup_{b \in D(f^*)} \{ \langle b, t \rangle - f^*(b) \}$$

$$= (f^*)^*(t). \quad \square$$

Lemma 3.11. Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi. Oletetaan, että $\nabla f(x_0) = 0$ eräälle $x_0 \in \mathbb{R}^d$. Silloin

$$\min_{x \in \mathbb{R}^d} f(x) = f(x_0).$$

Todistus. Olkoon aluksi $x_0 = 0$ ja $f(x_0) = f(0) = 0$. Olkoon $y \in \mathbb{R}^d$ kiinteä. Määritellään $g : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ ehdosta

$$g(\alpha) = f(\alpha y).$$

Tällöin g on konveksi ja $g(1) = f(y)$. Koska

$$f(h) = f(0) + \langle \nabla f(0), h \rangle + o(|h|) = o(|h|), \quad h \to 0,$$

niin

$$\frac{g(\alpha) - g(0)}{\alpha} = \frac{f(\alpha y)}{\alpha} = o(1), \quad \alpha \to 0.$$

Siis $g'(0) = 0$. Mielivaltaiselle $\alpha \in (0,1)$,

$$g(\alpha) = g((1-\alpha) \cdot 0 + \alpha \cdot 1) \leq \alpha g(1),$$

joten

$$f(y) = g(1) \geq \lim_{\alpha \to 0+} \frac{g(\alpha)}{\alpha} = 0.$$

Tämä todistaa lemman väitteen, kun $x_0 = 0$ ja $f(0) = 0$. Määrittelemällä

$$h(x) = f(x_0 + x) - f(x_0), \quad x \in \mathbb{R}^d,$$

nähdään, että $h(0) = 0$, $\nabla h(0) = \nabla f(x_0) = 0$, ja selvästi h on konveksi. Alkuosan nojalla

$$\min_{x \in \mathbb{R}^d} f(x) = f(x_0) + \min_{x \in \mathbb{R}^d} h(x) = f(x_0) + h(0) = f(x_0). \quad \square$$

<u>Seuraus 3.11.</u> Jos $\nabla f(t_x) = x \in \mathbb{R}^d$, niin

$$f^*(x) = \langle t_x, x\rangle - f(t_x).$$

<u>Lemma 3.12.</u> Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi. Jos $f(0) = 0$ ja $\nabla f(0) = 0$, niin $f^*(0) = 0$ ja

$$f^*(x) > f^*(0), \quad \forall x \neq 0.$$

<u>Todistus.</u> Seurauksen 3.11 nojalla $f^*(0) = 0$. Olkoon $x \in \mathbb{R}^d$, $x \neq 0$. Kirjoitetaan $x = (x_1, \dots, x_d)$. Olkoon $\varepsilon > 0$ ja $x_i \neq 0$. Olkoon edelleen $t = t_\varepsilon = (t_1, \dots, t_d)$ vektori, jonka i:s komponentti on ε / x_i ja muut nollia. Tällöin

$$f^*(x) \geq \langle t_\varepsilon, x\rangle - f(t_\varepsilon)$$

$$= \varepsilon - \left[f(0) + \langle \nabla f(0), t_\varepsilon\rangle + o(|t_\varepsilon|)\right] = \varepsilon + o(\varepsilon), \ \varepsilon \to 0+,$$

Nähdään, että $f^*(x) > 0 = f^*(0)$. $\square$

<u>Seuraus 3.13.</u> Jos vain $\nabla f(0)$ $\exists$, niin $f^*(\nabla f(0)) = 0$ ja

$$f^*(x) > 0, \quad \forall x \neq \nabla f(0).$$

<u>Tod.</u> Sovelletaan lemmaa funktioon g

$$g(t) = f(t) - \langle t, \nabla f(0)\rangle,$$

$$g^*(x) = \sup \{\langle t, x\rangle - f(t) + \langle t, \nabla f(0)\rangle\}$$

$$= f^*(x + \nabla f(0)) > 0, \quad \forall x \neq 0. \ \square$$