Appunti di Analisi convessa

Paolo Acquistapace

15 marzo 2017

Indice

Capitolo 1

Spazi vettoriali topologici

1.1 Insiemi convessi

Questo corso è dedicato alla descrizione, necessariamente sommaria, di alcuni aspetti e metodi dell'analisi non lineare, con particolare riguardo a qualcuna delle sue innumerevoli applicazioni.
L'analisi convessa è il principale argomento di cui ci occuperemo. Esso per certi versi è il più vicino alle tematiche di analisi funzionale lineare svolte in corsi precedenti (e che supponiamo note): frequentemente, infatti, i risultati dell'analisi convessa sono le prime, naturali estensioni a casi non lineari di proprietà già dimostrate ed analizzate nel caso lineare.
Ricordiamo, per cominciare, la nozione di insieme convesso.

Definizione 1.1.1 *Sia X uno spazio vettoriale su $\mathbb{R}$ o su $\mathbb{C}$. Un sottoinsieme $E \subseteq X$ si dice* convesso *se si ha*

$$\lambda x + (1-\lambda)x' \in E \qquad \forall \lambda \in [0,1],\ \forall x, x' \in E.$$

Osservazione 1.1.2 Le intersezioni, finite od infinite, di insiemi convessi sono insiemi convessi. Inoltre se $E, F \subseteq X$ sono convessi, allora gli insiemi

$$\alpha E = \{x \in X : x = \alpha y,\ y \in E\} \quad (\alpha \in \mathbb{R} \ \text{ oppure } \ \alpha \in \mathbb{C}),$$

$$E + F = \{x \in X : x = u + v,\ u \in E,\ v \in F\}$$

sono convessi.

Inoltre, se X, Y sono spazi vettoriali e se $F : X \to Y$ è un'applicazione lineare, allora valgono le implicazioni seguenti:

$$A \subseteq X \ \text{ convesso } \Longrightarrow F(A) \subseteq Y \ \text{ convesso},$$

$$B \subseteq Y \ \text{ convesso } \Longleftrightarrow \ F^{-1}(B) \subseteq X \ \text{ convesso}.$$

Definizione 1.1.3 *Sia E un sottoinsieme dello spazio vettoriale X. L'* inviluppo convesso *di E è l'insieme*

$$\text{co}(E) = \left\{ \sum_{k=1}^{n} \lambda_k x_k : \ n \in \mathbb{N}^+, \ x_k \in E, \ \lambda_k \geq 0, \ \sum_{k=1}^{n} \lambda_k = 1 \right\}.$$

Dunque $\text{co}(E)$ è il più piccolo insieme convesso contenente E.

Esercizi 1.1

1. Sia X uno spazio vettoriale e sia $\{U_t\}_{t \in]a,b[}$ una famiglia di sottoinsiemi convessi di X tale che $U_t \subseteq U_s$ per $t \leq s$. Si mostri che $\bigcup_{t \in]a,b[} U_t$ è un insieme convesso.

2. Sia X uno spazio vettoriale. Si provi che:

   (i) se $U \subseteq X$ e $\alpha, \beta > 0$ si ha $\alpha U + \beta U \supseteq (\alpha + \beta) U$ con inclusione stretta in generale;

   (ii) un sottoinsieme $U \subseteq X$ è convesso se e solo se $\alpha U + \beta U = (\alpha + \beta) U$ per ogni $\alpha, \beta > 0$.

3. Sia X uno spazio vettoriale. Un sottoinsieme U di X si dice *radiale*, od *assorbente*, se $0 \in U$ e se per ogni $x \in X$ esiste $\lambda > 0$ tale che $\alpha x \in U$ per ogni $\alpha \in]0, \lambda[$. Si provi che le proprietà "U è convesso e contiene 0" e "U è radiale" sono fra loro indipendenti.

4. Un sottoinsieme E di uno spazio vettoriale X si dice *bilanciato*, o *cerchiato*, se risulta $\alpha E \subseteq E$ per $|\alpha| \leq 1$. Si provi che:

   (i) un convesso che contiene l'origine non è necessariamente bilanciato;

   (ii) un insieme bilanciato contiene l'origine ma non è necessariamente convesso;

(iii) un insieme radiale non è necessariamente bilanciato;

(iv) un insieme bilanciato non è necessariamente radiale.

5. Si determini l'inviluppo convesso dei seguenti insiemi:

(i) $E = \{(x, y) \in \mathbb{R}^2 : \ |x| \le 1, \ |y| \le 1, \ xy = 0\}$;

(ii) $E = \{(x, y) \in \mathbb{R}^2 : \ y \le e^x\}$;

(iii) $E = \{z \in \mathbb{C} : \ z = e^{i\vartheta}, \ \vartheta \in \mathbb{Z}\}$;

(iii) $E = \left\{ \left(k, \frac{1}{|k|}\right) \in \mathbb{R}^2 : \ k \in \mathbb{Z} \setminus \{0\} \right\}$.

6. Sia X uno spazio vettoriale. Si mostri che

$$t(U \cap V) = tU \cap tV, \quad t(U \cup V) = tU \cup tV \qquad \forall t \in \mathbb{R}, \ \forall U, V \subseteq X.$$

7. Sia E un sottoinsieme di uno spazio vettoriale X e sia τ una qualunque topologia su X (che lo renda uno spazio vettoriale topologico). L'*inviluppo convesso chiuso* di E è il più piccolo insieme convesso e chiuso che contiene E. Si verifichi che esso coincide con $\overline{\mathrm{co}(E)}$.

1.2 Funzionale di Minkowski

Ad ogni sottoinsieme E di uno spazio vettoriale X possiamo associare una funzione reale p_E definita su X, dotata di interessanti proprietà che sono particolarmente significative, come vedremo, allorché E è un convesso contenente l'origine.

Definizione 1.2.1 *Sia E un sottoinsieme dello spazio vettoriale X. Il* funzionale di Minkowski *di E è la funzione $p_E : X \to [0, \infty]$ definita da*

$$p_E(x) = \inf\{\rho > 0 : \ x \in tE \quad \forall t \ge \rho\}$$

(con la convenzione che $p_E(x) = +\infty$ quando l'insieme a secondo membro è vuoto).

Si osservi che il numero $p_E(x)$ è l'estremo inferiore di una semiretta, eventualmente vuota: può benissimo succedere, in effetti, che sia $p_E(x) = +\infty$ per qualche x. Ad esempio, se $E = \{y\}$ con $y \ne 0$, si ha $p_E(x) = +\infty$ per ogni $x \in X$; se invece E è il semidisco di $\mathbb{R}^2$ di centro $(0, 0)$ e raggio 1 contenuto nel semipiano $y \ge 0$, allora $p_E(x, y) = +\infty$ se e solo se $y < 0$. Ha interesse, perciò, sapere per quali insiemi E il funzionale p_E è finito su X.

Proposizione 1.2.2 *Sia X uno spazio vettoriale, sia $E \subseteq X$. Risulta $p_E(x) < \infty$ per ogni $x \in X$ se e solo se E è radiale (esercizio 1.1.3).*

Dimostrazione ($\Longrightarrow$) Sia $p_E(x) < \infty$ per ogni $x \in X$. Dalle proprietà dell'estremo inferiore segue che $x \in tE$ per ogni $t > p_E(x)$; pertanto $\alpha x \in E$ per ogni $\alpha < \frac{1}{p_E(x)}$ (intendendo $\alpha < +\infty$ se $p_E(x) = 0$). In particolare $0 \in tE$ per ogni $t > p_E(0)$ e pertanto $0 = 0/t \in E$. Ciò prova che E è radiale.
($\Longleftarrow$) Sia $x \in X$. Per ipotesi si ha $x \in \frac{1}{\alpha}E$ per ogni $\alpha \in]0, \lambda[$, ossia $x \in tE$ per ogni $t > \frac{1}{\lambda}$. Ciò prova che $p_E(x) \leq \frac{1}{\lambda} < \infty$. $\square$

Vediamo ora le principali proprietà del funzionale di Minkowski.

Proposizione 1.2.3 *Sia E un sottoinsieme dello spazio vettoriale X. Valgono i fatti seguenti:*

(i) *$0 \in E$ se e solo se $p_E(0) = 0$;*

(ii) *se $x \neq 0$, allora $p_E(x) = 0$ se e solo se $\{\alpha x : \ \alpha > 0\} \subseteq E$;*

(iii) *si ha $p_E(\lambda x) = \lambda p_E(x)$ per ogni $\lambda > 0$ e per ogni $x \in X$;*

(iv) *se E è bilanciato (esercizio 1.1.4), allora $p_E(\lambda x) = |\lambda| p_E(x)$ per ogni $\lambda \in \mathbb{C} \setminus \{0\}$ e per ogni $x \in X$;*

(v) *se E è convesso, allora $p_E(x + x') \leq p_E(x) + p_E(x')$ per ogni $x, x' \in X$.*

Dimostrazione (i) Basta osservare che se $t > 0$ si ha $0 \in tE$ se e solo se $0 \in E$.
(ii) Sia $x \neq 0$; per definizione si ha $p_E(x) = 0$ se e solo se $x \in tE$ per ogni $t > 0$, ossia se e solo se $\alpha x \in E$ per ogni $\alpha > 0$.
(iii) Se $p_E(x) = +\infty$, allora chiaramente $p_E(\lambda x) = +\infty$. Se invece $p_E(x) < +\infty$, si ha

$$\begin{aligned} p_E(\lambda x) &= \inf\{\rho > 0 : \ \lambda x \in tE \ \ \forall t \geq \rho\} = \\ &= \inf\left\{\rho > 0 : \ x \in sE \ \ \forall s \geq \frac{\rho}{\lambda}\right\} = \\ &= \lambda \inf\{\sigma > 0 : \ x \in sE \ \ \forall s \geq \sigma\} = \lambda p_E(x). \end{aligned}$$

(iv) Proviamo anzitutto che $p_E(e^{i\vartheta}x) = p_E(x)$ per ogni $x \in X$ e $\vartheta \in \mathbb{R}$. Dal fatto che E è bilanciato segue facilmente che $p_E(x) = +\infty$ se e solo se

$p_E(e^{i\vartheta}x) = +\infty$. Se invece queste quantità sono finite, se $t > p_E(x)$ si ha $x \in tE$; quindi, essendo E bilanciato,

$$e^{i\vartheta}x \in e^{i\vartheta}(tE) = t(e^{i\vartheta}E) = tE \qquad \forall t > p_E(x),$$

da cui $p_E(e^{i\vartheta}x) \le p_E(x)$. Scambiando i ruoli di x e $e^{i\vartheta}x$ si ottiene l'altra disuguaglianza.
Se ora $\lambda = |\lambda|e^{i\vartheta} \in \mathbb{C} \setminus \{0\}$, si ha, usando **(iii)**,

$$p_E(\lambda x) = p_E(|\lambda|e^{i\vartheta}x) = |\lambda|p_E(e^{i\vartheta}x) = |\lambda|p_E(x),$$

da cui la tesi.

(v) Possiamo supporre che $p_E(x)$ e $p_E(x')$ siano entrambi finiti. Poiché E è convesso, si ha $\alpha E + \beta E = (\alpha + \beta)E$ per ogni $\alpha, \beta \ge 0$ (esercizio 1.1.2). Poiché risulta $x \in tE$ per ogni $t > p_E(x)$ e $x' \in sE$ per ogni $s > p_E(x')$, si deduce $x + x' \in (t + s)E$ per ognuno di tali t, s; quindi $x + x' \in \tau E$ per ogni $\tau > p_E(x) + p_E(x')$. Ciò implica $p_E(x + x') \le p_E(x) + p_E(x')$. $\square$

Osservazione 1.2.4 Se lo spazio X è reale, la proprietà **(iv)** della proposizione precedente vale per ogni $\lambda \in \mathbb{R}\setminus\{0\}$ supponendo solamente che l'insieme E sia *simmetrico*, ossia $E = -E$.

Una nozione che useremo spesso nel seguito è quella di seminorma: le seminorme sono importanti perché, come vedremo, consentono di costruire topologie dotate di buone proprietà.

Definizione 1.2.5 *Una* seminorma *nello spazio vettoriale X è una funzione $p : X \to [0, \infty[$ tale che*

(i) *$p(\lambda x) = |\lambda|p(x)$ per ogni $\lambda \in \mathbb{R}$ (oppure per ogni $\lambda \in \mathbb{C}$) e per ogni $x \in X$;*

(ii) *$p(x + x') \le p(x) + p(x')$ per ogni $x, x' \in X$.*

Osservazione 1.2.6 Dalla definizione segue immediatamente che ogni seminorma p su X verifica $p(0) = 0$ e $|p(x) - p(x')| \le p(x - x')$ per ogni $x, x' \in X$. In particolare, ogni seminorma verifica $p(0) = 0$, ma non è detto che valga l'implicazione $p(x) = 0 \implies x = 0$; ciò vale se e solo se p è una norma su X. Inoltre si verifica subito che se p è una seminorma su X allora gli insiemi

$$U_c = \{x \in X : \ p(x) < c\}, \quad c > 0,$$

sono convessi, radiali e bilanciati.

Esempio 1.2.7 Se E è un sottoinsieme dello spazio vettoriale X che è radiale, bilanciato e convesso, allora il funzionale di Minkowski p_E è una seminorma su X in virtù della proposizione 1.2.3. È vero anche il viceversa: se p è una seminorma sullo spazio vettoriale X, allora p coincide col funzionale di Minkowski p_U, ove

$$U = \{x \in X : \; p(x) < 1\}$$

(esercizio 1.2.4).

Esercizi 1.2

1. Si forniscano esempi di insiemi contenenti l'origine che siano:

   (a) convessi ma non radiali;

   (b) radiali ma non convessi;

   (c) convessi ma non simmetrici;

   (d) simmetrici ma non convessi;

   (e) simmetrici ma non radiali;

   (f) radiali ma non simmetrici;

   (g) simmetrici ma non bilanciati.

2. Si provi che se E è un sottoinsieme convesso dello spazio vettoriale X che contiene 0, allora

   $$p_E(x) = \inf\{r > 0 : \; x \in rE\} \qquad \forall x \in X;$$

   si dia anche un esempio di sottoinsieme $E \subseteq X$ tale che $0 \in E$ ma per il quale risulti, per almeno un $x \in X$,

   $$\inf\{r > 0 : \; x \in rE\} < p_E(x).$$

3. Si provi che se X è uno spazio vettoriale, e se $U \subseteq V \subseteq X$, allora $p_U \geq p_V$. Se ne deduca che se $\{U_i\}_{i \in I}$ è un'arbitraria famiglia di sottoinsiemi di X, allora

   $$p_{\bigcap_{i \in I} U_i} = \sup_{i \in I} p_{U_i} \,, \qquad p_{\bigcup_{i \in I} U_i} = \inf_{i \in I} p_{U_i} \,.$$

4. Sia p una seminorma definita sullo spazio vettoriale X. Posto per $c > 0$

$$U_c = \{x \in X :\ p(x) < c\},$$

si provi che $p_{U_c}(x) = \frac{p(x)}{c}$ per ogni $x \in X$.

5. Sia X uno spazio vettoriale, sia $E \subseteq X$. Si provi che

$$\{x \in X :\ p_E(x) < 1\} \subseteq E,$$

e che inoltre se E è convesso e $0 \in E$, allora

$$E \subseteq \{x \in X : p_E(x) \le 1\}.$$

6. Sia X uno spazio normato. Se U è un intorno di 0, si provi che esiste $M \ge 0$ tale che $p_U(x) \le M\|x\|_X$ per ogni $x \in X$.

7. Si determini il funzionale di Minkowski p_E nei casi seguenti:

   (i) X spazio normato, $E = X$;

   (ii) X spazio normato, $E = B(0, r)$;

   (iii) $X = \mathbb{R}^2$, $E = (x, y) \in \mathbb{R}^2 :\ ax + by \le c\}$, con $a, b \in \mathbb{R}$ e $c > 0$ fissati;

   (iv) $X = \mathbb{R}^2$, $E = \{(x, y) \in \mathbb{R}^2 :\ x^2 + y^2 \le 1,\ y \ge 0\}$;

   (v) $X = \ell^2$, $E = \{x = \{x_n\}_{n\in\mathbb{N}} :\ |x_j| \le 1\}$, con $j \in \mathbb{N}$ fissato.

1.3 Topologie su uno spazio vettoriale

In molte applicazioni dell'analisi non lineare gli spazi di Banach non sono sufficientemente generali. Ad esempio, per ottenere l'esistenza di soluzioni per certi problemi occorrono risultati di compattezza che valgono in molti di questi spazi soltanto se si fa uso di topologie più deboli di quella indotta dalla norma; rispetto a tali topologie lo spazio perde la sua struttura metrica, mantenendo solo quella di "spazio vettoriale topologico", in cui l'unico requisito richiesto è la continuità delle operazioni di somma e prodotto per scalari rispetto alla topologia. È quindi utile fare un esame sommario delle caratteristiche di questi spazi.

Definizione 1.3.1 *Uno* spazio vettoriale topologico *è uno spazio vettoriale X dotato di una topologia τ, rispetto alla quale sono continue le due applicazioni*

$$(x,y) \mapsto x+y \qquad da\ X \times X\ in\ X,$$
$$(\alpha,x) \mapsto \alpha x \qquad da\ \mathbb{R} \times X\ in\ X,\ oppure\ da\ \mathbb{C} \times X\ in\ X.$$

Da questa definizione segue subito che le traslazioni $x \mapsto x+x_0$ e le omotetie $x \mapsto \alpha x$, con $x_0 \in X$ e $\alpha \in \mathbb{C}$ fissati, sono omeomorfismi di X in sé. È anche chiaro che la topologia di uno spazio vettoriale topologico è univocamente individuata se si fissano gli intorni U di 0: gli intorni di ogni altro punto $x \neq 0$ sono semplicemente gli insiemi traslati $x+U$.

Osservazione 1.3.2 In un qualunque spazio vettoriale topologico, ogni intorno U di 0 contiene un intorno W di 0 che è bilanciato. Infatti, per la continuità del prodotto esistono $\varepsilon > 0$ ed un intorno V di 0 tali che $\alpha V \subseteq U$ per $|\alpha| < \varepsilon$; quindi, posto

$$W = \bigcup_{|\alpha|<\varepsilon} \alpha V,$$

risulta $W \subseteq U$, e W è ovviamente bilanciato. Si noti che W è anche radiale (esercizio 1.3.1). Se l'intorno U è anche convesso, non è detto che anche W lo sia; tuttavia in tal caso è facile verificare che l'insieme $\mathrm{co}(W)$ è ancora un intorno di 0 che è contenuto in U, convesso, radiale, nonché bilanciato in virtù dell'esercizio 1.3.9.

La proposizione che segue caratterizza gli spazi vettoriali topologici che sono di Hausdorff.

Proposizione 1.3.3 *Sia X uno spazio vettoriale topologico, sia $\{U_i\}_{i\in I}$ un sistema fondamentale di intorni dell'origine. Lo spazio X è di Hausdorff se e solo se*

$$\bigcap_{i\in I} U_i = \{0\}.$$

Dimostrazione ($\Longrightarrow$) Sia X di Hausdorff: allora se $x_0 \neq 0$ esiste un intorno U_i tale che $x_0 \notin U_i$, e quindi $x_0 \notin \bigcap_{i\in I} U_i$.
($\Longleftarrow$) Sia $x_0 \neq 0$: per ipotesi esiste $i \in I$ tale che $x_0 \notin U_i$. Poiché l'applicazione $(x,y) \mapsto x-y$ è continua, esiste $j \in I$ tale che $x-y \in U_i$ per ogni $x,y \in U_j$; ciò significa che $U_j - U_j \subseteq U_i$. Se ne deduce che U_j è disgiunto da $x_0 + U_j$: infatti in caso contrario troveremmo $x_1, x_2 \in U_j$ tali che

$x_0+x_1=x_2$, da cui $x_0=x_2-x_1\in U_j-U_j\subseteq U_i$, cosa che non è. Quindi x_0 e 0 hanno intorni disgiunti. Per traslazione, due arbitrari punti distinti x, x' hanno intorni disgiunti. $\square$

Definizione 1.3.4 *Uno spazio vettoriale topologico X si dice* localmente convesso *se esiste un sistema fondamentale di intorni di* 0 *costituito da insiemi convessi.*

Gli spazi vettoriali topologici localmente convessi sono importanti perché la loro topologia può essere caratterizzata mediante una famiglia di seminorme, il che ne rende più comodo lo studio.
Sia $\{p_i\}_{i\in I}$ una famiglia arbitraria di seminorme definite su uno spazio vettoriale X: vediamo ora come esse "generino" una topologia τ che rende X uno spazio vettoriale topologico localmente convesso. Fissati $k\in\mathbb{N}^+$ e $i_1,\ldots,i_k\in I$, definiamo per $x\in X$ ed $\varepsilon>0$

$$V_{i_1,\ldots,i_k,\varepsilon}(x)=\{\xi\in X:\ p_{i_h}(\xi-x)<\varepsilon \ \text{ per } \ h=1,\ldots,k\}.$$

Gli insiemi $V_{i_1,\ldots,i_k,\varepsilon}(x)$ sono convessi (osservazione 1.2.6), ed è facile constatare che per ogni $x\in X$ la famiglia

$$\mathcal{V}(x)=\{V_{i_1,\ldots,i_k,\varepsilon}(x):\ k\in\mathbb{N}^+,\ i_1,\ldots,i_k\in I,\ \varepsilon>0\},$$

verifica gli assiomi relativi ad un sistema fondamentale di intorni di x; quindi esiste una ed una sola topologia τ su X per la quale, per ogni $x\in X$, $\mathcal{V}(x)$ è un sistema fondamentale di intorni di x. Rispetto a questa topologia le applicazioni $(\alpha,x)\mapsto\alpha x$ e $(x,y)\mapsto x+y$ sono continue, come facile conseguenza delle relazioni

$$p_i(\alpha x)=|\alpha|p_i(x), \qquad p_i(x+y)\le p_i(x)+p_i(y);$$

dunque la topologia τ rende X uno spazio vettoriale topologico localmente convesso.
Si noti che tutte le seminorme p_i sono automaticamente continue rispetto a τ, a causa della disuguaglianza $|p_i(x)-p_i(y)|\le p_i(x-y)$. In effetti τ è precisamente la meno fine topologia rispetto a cui tutte le p_i sono continue.
Vale anche il viceversa: se (X,τ) è uno spazio vettoriale topologico localmente convesso, esiste una famiglia di seminorme su X che genera la topologia τ. Infatti, sia $\{U_i\}_{i\in I}$ un sistema fondamentale di intorni convessi di 0 nella

topologia τ, che possiamo sempre supporre bilanciati e radiali (osservazione 1.3.2): i funzionali di Minkowski p_{U_i} ad essi associati sono seminorme (esempio 1.2.7) che verificano, in virtù dell'esercizio 1.3.2,

$$\{x \in X : \ p_{U_i}(x) < 1\} \subseteq U_i \subseteq \{x \in X : \ p_{U_i}(x) \leq 1\},$$

e quindi generano τ.
Vi è un facile criterio, basato sulle seminorme, per stabilire se un dato spazio vettoriale topologico è di Hausdorff o no. Esso è espresso dall'enunciato che segue.

Proposizione 1.3.5 *Sia X uno spazio vettoriale, sia $\{p_i\}_{i \in I}$ una famiglia di seminorme definite su X. La topologia generata dalle p_i è di Hausdorff se e solo se le p_i separano i punti, ossia per ogni $x \in X \setminus \{0\}$ esiste $i \in I$ tale che $p_i(x) > 0$.*

Dimostrazione Per la proposizione 1.3.3, X è di Hausdorff se e solo se l'intersezione di tutti gli intorni

$$V_{i_1,\dots,i_k,\varepsilon} = \{x \in X : \ p_{i_h}(x) < \varepsilon \ \text{ per } \ h = 1, \dots, k\},$$

al variare di $k \in \mathbb{N}^+$, $i_1, \dots, i_k \in I$ ed $\varepsilon > 0$, è uguale a $\{0\}$. Poiché tale intersezione coincide con $\bigcap_{i \in I} p_i^{-1}(0)$, X è di Hausdorff se e solo se non vi è alcun $x \neq 0$ tale che $p_i(x) = 0$ per ogni $i \in I$. Ciò prova la tesi. $\square$

Alcuni spazi vettoriali topologici di Hausdorff sono particolarmente importanti perché su di essi è possibile definire una metrica:

Proposizione 1.3.6 *Sia X uno spazio vettoriale topologico localmente convesso e di Hausdorff. Allora X è metrizzabile se e solo se esiste un sistema fondamentale di intorni di 0 che è numerabile.*

L'ipotesi di locale convessità sarebbe superflua, ma semplifica la dimostrazione, e d'altronde è il caso che più ci interessa.

Dimostrazione È chiaro che se X è metrizzabile allora le palle di centro 0 e raggio $\frac{1}{n}$ formano un sistema fondamentale di intorni di 0 che è numerabile. Viceversa, sia $\{U_n\}_{n \in \mathbb{N}}$ un sistema fondamentale di intorni di 0 numerabile e fatto di intorni convessi, bilanciati e radiali (osservazione 1.3.2). Allora, in virtù dell'esercizio 1.2.5, gli insiemi $\{x \in X : \ p_{U_n}(x) < 1\}$ costituiscono un sistema fondamentale di intorni di 0 equivalente a quello degli U_n. Perciò, definendo

$$d(x,y) = \sum_{n=0}^{\infty} 2^{-n} \frac{p_{U_n}(x-y)}{1+p_{U_n}(x-y)} \qquad \forall x,y \in X,$$

si ottiene una distanza su X che induce la topologia data. $\square$

Osservazione 1.3.7 Si noti che la distanza sopra definita è invariante per traslazioni, cioè $d(x+z, y+z) = d(x,y)$ per ogni $x,y,z \in X$. Questa proprietà non è automatica: ad esempio, nell'insieme $\mathbb{R}$, che è uno spazio vettoriale topologico metrizzabile con l'ovvia distanza $d(x,y) = |x-y|$, vi è la distanza equivalente $\delta(x,y) = |\arctan x - \arctan y|$ che induce la stessa topologia ma non è invariante per traslazioni.

Esercizi 1.3

1. Sia X uno spazio vettoriale topologico. Si provi che ogni intorno di 0 è radiale.

2. Sia X uno spazio vettoriale topologico e sia U un convesso contenente l'origine. Si provi che:

   (i) se U è aperto, allora $U = \{x \in X : \ p_U(x) < 1\}$;

   (ii) se U è chiuso, allora $U = \{x \in X : \ p_U(x) \leq 1\}$.

3. Sia X uno spazio vettoriale topologico. Si provi che se U è un intorno di 0, allora λU è un intorno di 0 per ogni $\lambda \neq 0$.

4. Sia X uno spazio vettoriale topologico. Si provi che per ogni intorno U di 0 esiste un altro intorno V di 0 tale che $V + V \subseteq U$.

5. Sia X uno spazio vettoriale topologico e sia U un intorno di 0. Si provi che $X = \bigcup_{s>0} sV$.

6. Si provi che in ogni spazio vettoriale topologico l'origine ha un sistema fondamentale di intorni *chiusi*.

7. Sia X uno spazio vettoriale topologico. Si provi che se $A \subseteq X$ è un aperto, allora il suo inviluppo convesso $\mathrm{co}(A)$ è aperto.

8. Sia X uno spazio vettoriale topologico, sia K un convesso non vuoto di X. Si provi che $\overset{\circ}{K} \neq \emptyset$ se e solo se le chiusure di $\overset{\circ}{K}$ e di K coincidono.

9. Sia W un insieme bilanciato in uno spazio vettoriale topologico X. Si provi che $\mathrm{co}(W)$ è bilanciato.

1.4 Completezza

Sia X uno spazio vettoriale topologico. Andiamo a definire la nozione di completezza dello spazio X. Le nozioni di successione di Cauchy e di successione convergente (a valori in X) sono ben note: $x_n \to x$ in X se per ogni intorno U di 0 esiste $\nu \in \mathbb{N}$ tale che $x_n - x \in U$ per ogni $n > \nu$; $\{x_n\}$ è di Cauchy in X se per ogni intorno U di 0 esiste $\nu \in \mathbb{N}$ tale che $x_n - x_m \in U$ per ogni $n, m > \nu$. Utilizzando la continuità della somma in X si ottiene anche che ogni successione convergente è di Cauchy. Se la topologia di X è generata da una famiglia di seminorme $\{p_j\}_{j \in J}$, si vede facilmente che $x_n \to x$ in X se e solo se si ha $p_j(x_n - x) \to 0$ in $\mathbb{R}$ per ogni seminorma p_j.
Cominciamo con il caso in cui 0 ha un sistema fondamentale di intorni numerabile.

Definizione 1.4.1 *Sia X uno spazio vettoriale topologico tale che 0 abbia un sistema fondamentale di intorni numerabile. Diciamo che X è* completo *se ogni successione di Cauchy in X converge ad un elemento di X.*

Nel caso generale, eliminando l'ipotesi che 0 abbia un sistema fondamentale di intorni numerabile, la completezza di X si fonda sul concetto di successione generalizzata, o net.

Definizione 1.4.2 *Un* insieme diretto*, o* filtrante*, è un insieme $(I, \leq)$ parzialmente ordinato, con la proprietà che per ogni $i, j \in I$ esiste $h \in I$ con $h \geq i$ e $h \geq j$. Una* successione generalizzata*, o* net*, a valori in un insieme X è una famiglia $\{x_i\}_{i \in I}$ di punti di X, ove I è un insieme filtrante.*

Esempi 1.4.3 **(1)** Poiché $\mathbb{N}$ è un insieme filtrante, ogni successione $\{x_n\}$ contenuta in un fissato insieme X è un net a valori in X.

(2) La famiglia U degli intorni di 0 in un arbitrario spazio topologico è un insieme filtrante rispetto all'ordinamento parziale dell'inclusione "rovesciata": se $U, V \in \mathcal{U}$, si definisce $U \leq V$ se $U \supseteq V$. Quindi ogni famiglia di punti

della forma $\{x_U\}_{U\in\mathcal{U}}$ è un net a valori in X.
(3) Un ordinamento parziale in $\mathbb{R}^2$ è dato da

$$(x,y) \leq (x',y') \quad \Longleftrightarrow \quad x < x' \text{ oppure } x = x' \text{ e } y \leq y'.$$

Con questo ordinamento, $\mathbb{R}^2$ è un insieme filtrante e ogni funzione $f : \mathbb{R}^2 \to X$ definisce il net $\{f(x,y)\}_{(x,y)\in\mathbb{R}^2}$ a valori in X.

Definizione 1.4.4 *Sia X uno spazio vettoriale topologico. Diciamo che un net $\{x_i\}_{i\in I}$ contenuto in X* converge *ad un punto $x \in X$ se per ogni intorno U di 0 esiste $i_0 \in I$ tale che $x_i - x \in U$ per ogni $i \in I$, $i \geq i_0$. Il net è* di Cauchy *in X se per ogni intorno U di 0 esiste $i_0 \in I$ tale che $x_i - x_j \in U$ per ogni $i, j \in I$, $i, j \geq i_0$.*

Nuovamente, ogni net convergente è di Cauchy. Se la topologia di X è generata da una famiglia di seminorme $\{p_j\}_{j\in J}$, un net $\{x_i\}_{i\in I} \subseteq X$ converge all'elemento x in X se e solo se per ogni $j \in J$ il net $\{p_j(x_i - x)\}_{i\in I} \subseteq \mathbb{R}$ è infinitesimo (tende a 0).

Definizione 1.4.5 *Uno spazio vettoriale topologico X si dice* completo *se ogni net di Cauchy in X converge ad un elemento x di X.*

Definizione 1.4.6 *Uno spazio vettoriale topologico localmente convesso, di Hausdorff, completo è detto* spazio di Fréchèt.

Vediamo ora alcuni esempi importanti di spazi vettoriali topologici, completi e non.

Esempio 1.4.7 Ogni spazio normato è uno spazio vettoriale topologico localmente convesso e di Hausdorff; le palle $B(0, \frac{1}{n})$ costituiscono un sistema fondamentale numerabile di intorni di 0 convessi. Se lo spazio normato è completo, esso è uno spazio di Fréchèt.

Esempio 1.4.8 Consideriamo per $p \in]0,1[$ lo spazio

$$\ell^p = \left\{ x = \{x_n\}_{n\in\mathbb{N}} \subset \mathbb{C} : \; \Delta(x) = \sum_{n=0}^{\infty} |x_n|^p < \infty \right\}.$$

La disuguaglianza $(a+b)^p \leq a^p + b^p$, valida per ogni $a, b \geq 0$, mostra che ℓ^p è uno spazio vettoriale, e che la quantità

$$d(x,y) = \Delta(x-y) = \sum_{n=0}^{\infty} |x_n - y_n|^p, \qquad x, y \in \ell^p,$$

è una distanza su ℓ^p; si noti che $\Delta(x)$ non è una norma perché $\Delta(\lambda x) = |\lambda|^p \Delta(x)$. È chiaro che somma ed il prodotto per scalari sono applicazioni continue rispetto a questa distanza; quindi ℓ^p è uno spazio vettoriale topologico di Hausdorff, ed è anche uno spazio metrico completo (esercizio 1.4.6). Tuttavia ℓ^p non è localmente convesso (e quindi non è uno spazio di Fréchèt). Infatti in caso contrario, posto $B(r) = \{x \in \ell^p : \ \Delta(x) \leq r\}$, esisterebbe un intorno convesso W di 0 tale che $W \subseteq B(1)$; scelto $m \in \mathbb{N}^+$ in modo che $B(m^{-p}) \subseteq W$, utilizzando l'omotetia $x \mapsto mx$ si dedurrebbe

$$B(1) = \frac{1}{m} B(m^{-p}) \subseteq \frac{1}{m} W \subseteq \frac{1}{m} B(1) = B(m^p) \subset B(m).$$

Consideriamo d'altronde la famiglia $\{e^{(n)}\}_{n\in\mathbb{N}}$, ove $e^{(n)}$ è la successione che ha tutte la componenti nulle tranne la n-sima che vale 1. Si ha $e^{(n)} \in B(1)$, quindi $e^{(n)}$ apparterrebbe al convesso mW per ogni $n \in \mathbb{N}$. Pertanto anche

$$g^{(n)} = \frac{1}{n} \sum_{k=0}^{n-1} e^{(k)}$$

apparterrebbe a mW per ogni $n \in \mathbb{N}^+$. D'altra parte

$$\Delta(g^{(n)}) = n^{1-p} \to \infty \qquad \text{per} \quad n \to \infty,$$

mentre, essendo $mW \subset B(m)$, dovrebbe aversi $\Delta(g^{(n)}) \leq m$ per ogni $n \in \mathbb{N}^+$. Ciò è assurdo.

Esempio 1.4.9 Sia Ω un aperto non vuoto di $\mathbb{R}^N$; fissato $m \in \mathbb{N}$, consideriamo lo spazio $C^m(\Omega)$ delle funzioni che sono continue su Ω insieme con le loro derivate fino all'ordine m incluso. Se $\{K_n\}_{n\in\mathbb{N}}$ è una successione strettamente crescente di compatti (ossia $K_n \subset \overset{\circ}{K}_{n+1}$), la cui unione è Ω, ponendo

$$p_n(f) = \sup_{K_n} \sum_{|\alpha| \leq m} |D^\alpha f|$$

si ottiene una famiglia numerabile di seminorme che separano le funzioni di $C^m(\Omega)$: infatti se $f \neq 0$ esiste certamente un intero $n \in \mathbb{N}$ per il quale $p_n(f) > 0$. Dunque, per ogni $m \in \mathbb{N}$, $C^m(\Omega)$ è uno spazio vettoriale topologico localmente convesso, di Hausdorff, metrizzabile; in particolare, la topologia è indipendente dalla scelta della successione di compatti $\{K_n\}$, e si ha $f_n \to f$ in $C^m(\Omega)$ se e solo se per ogni compatto $K \subset \Omega$ si ha $D^\alpha f_n \to D^\alpha f$ uniformemente in K per ogni multi-indice $\alpha \in \mathbb{N}^N$ con $|\alpha| \leq m$. Da noti teoremi sulla convergenza uniforme segue che $C^m(\Omega)$ è uno spazio di Fréchèt.

Esempio 1.4.10 Consideriamo lo spazio $C^\infty(\Omega)$ delle funzioni infinitamente derivabili definite sull'aperto non vuoto $\Omega \subseteq \mathbb{R}^N$. Ovviamente $C^\infty(\Omega) \subset C^m(\Omega)$ per ogni $m \in \mathbb{N}$; se vogliamo, come è naturale, che questa immersione sia continua, occorre mettere su $C^\infty(\Omega)$ una topologia τ tale che per ogni $m \in \mathbb{N}$ e per ogni compatto $K \subset \Omega$ gli insiemi

$$\left\{ f \in C^\infty(\Omega) : \sup_K \sum_{|\alpha| \leq m} |D^\alpha f| < \varepsilon \right\}, \quad \varepsilon > 0,$$

siano τ-intorni di 0. La meno fine topologia per cui questo accade si può costruire nel modo seguente: si sceglie una successione strettamente crescente di compatti $\{K_n\}_{n\in\mathbb{N}}$ che invada Ω, e si introduce la famiglia numerabile di seminorme $\{p_{mn}\}_{m,n\in\mathbb{N}}$ così definite:

$$p_{mn}(f) = \sup_{K_n} \sum_{|\alpha| \leq m} |D^\alpha f|.$$

Poiché, ovviamente, queste seminorme separano gli elementi di $C^\infty(\Omega)$, la topologia da esse generata è localmente convessa e di Hausdorff; in particolare $C^\infty(\Omega)$ è metrizzabile. Si ha $f_n \to f$ in $C^\infty(\Omega)$ se e solo se per ogni compatto $K \subset \Omega$ e per ogni $\alpha \in \mathbb{N}^N$ si ha $D^\alpha f_n \to D^\alpha f$ uniformemente in K. Si conclude che $C^\infty(\Omega)$ è uno spazio di Fréchèt immerso con continuità in tutti gli spazi $C^m(\Omega)$.

Esempio 1.4.11 Questo esempio è meno facile ma più importante. Per $m \in \mathbb{N}$ sia $C_0^m(\Omega)$ lo spazio delle funzioni $f \in C^m(\Omega)$ il cui supporto è compatto nell'aperto $\Omega \subseteq \mathbb{R}^N$. Osserviamo che se K è un fissato compatto contenuto in Ω, allora il sottospazio $C_K^m(\Omega)$, costituito dalle funzioni di $C_0^m(\Omega)$ che hanno supporto contenuto in K, è uno spazio di Banach con la norma

$$p_K^m(f) = \sup_K \sum_{|\alpha| \le m} |D^\alpha f|;$$

inoltre, se K e K' sono compatti tali che $K \subseteq K' \subset \Omega$, l'immersione naturale $C_K^m(\Omega) \subseteq C_{K'}^m(\Omega)$ è un'isometria di $C_K^m(\Omega)$ su un sottospazio chiuso di $C_{K'}^m(\Omega)$. Notiamo poi che, ovviamente,

$$C_0^m(\Omega) = \bigcup_{K \text{ compatto } \subset \Omega} C_K^m(\Omega).$$

Introduciamo su $C_0^m(\Omega)$ la topologia indotta dalla famiglia di seminorme costruite nel modo seguente: considerato l'insieme c_0^+ delle successioni infinitesime di numeri positivi, e fissata una successione crescente di compatti $\{K_n\}_{n\in\mathbb{N}}$ tale che $K_0 = \emptyset$ e $\bigcup_{n\in\mathbb{N}} K_n = \Omega$, per ogni successione $a = \{a_n\}_{n\in\mathbb{N}} \in c_0^+$ si definisce la seminorma

$$p_a(f) = \sup_{\nu\in\mathbb{N}} \sup_{x\in\Omega\setminus K_\nu} \frac{1}{a_\nu} \sum_{|\alpha|\le m} |D^\alpha f(x)|.$$

Si noti che in effetti, essendo $K_0 = \emptyset$, p_a è una norma. In particolare, la famiglia $\{p_a\}_{a\in c_0^+}$ definisce su $C_0^m(\Omega)$ una topologia localmente convessa e di Hausdorff. Il motivo per il quale si considera l'intera famiglia $\{p_a\}_{a\in c_0^+}$ anziché una singola norma è chiarito dal risultato seguente:

Proposizione 1.4.12 *La topologia sopra descritta è la più fine topologia localmente convessa nello spazio $C_0^m(\Omega)$ rispetto alla quale, per ogni compatto $K \subset \Omega$, l'immersione $C_K^m(\Omega) \subset C_0^m(\Omega)$ è continua.*

Dimostrazione Per provare la tesi dobbiamo mostrare che gli intorni dell'origine nella topologia sopra definita sono tutti e soli gli insiemi convessi dotati della proprietà che per ogni compatto $K \subset \Omega$ la loro intersezione con lo spazio $C_K^m(\Omega)$ è un intorno convesso dell'origine in $C_K^m(\Omega)$.
A questo scopo basta considerare il sistema fondamentale di intorni dell'origine della forma $V_a = \{f \in C_0^m(\Omega) : p_a(f) < 1\}$. Sia quindi $a = \{a_\nu\} \in c_0^+$: si ha

$$V_a = \left\{ f \in C_0^m(\Omega) : \sum_{|\alpha|\le m} |D^\alpha f(x)| < a_\nu \ \forall x \in \Omega \setminus K_\nu,\ \forall \nu \in \mathbb{N} \right\},$$

e quindi

$$V_a \cap C^m_{K_\nu}(\Omega) =$$
$$= \left\{ f \in C^m_{K_\nu}(\Omega) : \sum_{|\alpha|\le m} |D^\alpha f(x)| < a_j \ \forall x \in \Omega \setminus K_j, \ \forall j = 0, 1, \ldots, \nu - 1 \right\}.$$

In particolare

$$\{f \in C^m_{K_\nu}(\Omega) : p^m_{K_\nu}(f) < \min\{a_0, a_1, \ldots, a_{\nu-1}\}\} \subseteq V_a \cap C^m_{K_\nu}(\Omega),$$

cioè $V_a \cap C^m_{K_\nu}(\Omega)$ è un intorno convesso dell'origine in $C^m_{K_\nu}(\Omega)$ per ogni $\nu \in \mathbb{N}$.
Viceversa, sia $W \subseteq C^m_0(\Omega)$ un convesso tale che, per ogni $\nu \in \mathbb{N}$, $W \cap C^m_{K_\nu}(\Omega)$ sia un intorno dell'origine in $C^m_{K_\nu}(\Omega)$: proviamo che esiste una successione $a \in c_0^+$ tale che $W \supseteq V_a$.
Poiché $W \cap C^m_{K_{\nu+2}}(\Omega)$ è intorno di 0 in $C^m_{K_{\nu+2}}(\Omega)$, risulta

$$W \cap C^m_{K_{\nu+2}}(\Omega) \supseteq \{f \in C^m_{K_{\nu+2}}(\Omega) : \ p^m_{K_{\nu+2}}(f) < \eta_\nu\},$$

ove $\eta = \{\eta_\nu\}$ è un opportuno elemento di c_0^+. Sia $\{\varphi_\nu\}_{\nu\in\mathbb{N}}$ una partizione dell'unità di classe C^m associata al ricoprimento di Ω costituito dalle "corone" aperte $\overset{\circ}{K}_{\nu+2} \setminus K_\nu$. Scrivendo una generica $f \in C^m_0(\Omega)$ in forma di somma (necessariamente finita, in realtà)

$$f = \sum_{\nu=0}^{\infty} 2^{-\nu-1}(2^{\nu+1}\varphi_\nu f),$$

dalla convessità di W seguirà $f \in W$ purché si abbia $2^{\nu+1}\varphi_\nu f \in W$ per ogni $\nu \in \mathbb{N}$. Osserviamo che per ogni $\nu \in \mathbb{N}$ esiste $C_\nu > 0$ tale che, per qualsiasi $f \in C^m_0(\Omega)$,

$$\sup_{\Omega\setminus K_\nu} \sum_{|\alpha|\le m} |D^\alpha f| \le a_\nu \quad \Longrightarrow \quad \sup_{\Omega} 2^{\nu+1} \sum_{|\beta|\le m} |D^\beta(\varphi_\nu f)| \le C_\nu a_\nu \,,$$

in virtù del fatto che φ_ν è nulla su K_ν. Scegliamo la successione $a = \{a_\nu\}$ in modo che sia $C_\nu a_\nu < \eta_\nu$ per ogni $\nu \in \mathbb{N}$: allora se $f \in V_a$, per ogni $\nu \in \mathbb{N}$ si ha $2^{\nu+1}\varphi_\nu f \in C^m_{K_{\nu+2}}(\Omega)$ e $p^m_{K_{\nu+2}}(2^{\nu+1}\varphi_\nu f) < \eta_\nu$, ossia $2^{\nu+1}\varphi_\nu f \in W$. In definitiva W contiene l'intorno V_a e pertanto è un intorno convesso dell'origine in $C^m_0(\Omega)$. Ciò prova la tesi. $\Box$

Teorema 1.4.13 *Sia Ω un aperto di $\mathbb{R}^N$, e sia $m \in \mathbb{N}$. Valgono i seguenti fatti:*

(i) *$C_0^m(\Omega)$ non è metrizzabile;*

(ii) *$C_0^m(\Omega)$ è uno spazio di Fréchèt.*

Dimostrazione (i) Basta far vedere che per ogni sottofamiglia numerabile $\{V_{a^{(n)}}\}_{n\in\mathbb{N}}$ del sistema $\{V_a\}$ esiste un intorno V_a che non contiene alcun elemento della sottofamiglia. Possiamo supporre che per ogni $n \in \mathbb{N}$ la successione $a^{(n)} = \{a_\nu^{(n)}\}_{\nu\in\mathbb{N}}$ sia decrescente e contenuta in $]0,1[$. Per ogni $j \in \mathbb{N}$ sia

$$a_j = \prod_{n,\nu=0}^{j} a_\nu^{(n)}$$

e poniamo $a = \{a_j\}_{j\in\mathbb{N}}$: la successione a è un elemento di c_0^+ e inoltre $a_j < a_j^{(n)}$ per ogni $j \in \mathbb{N}$ e $n \leq j$. Sia, per assurdo, $V_{a^{(n)}} \subseteq V_a$: se $f \in C_0^m(\Omega)$ ha supporto contenuto in $\overset{\circ}{K}_{n+1} \setminus K_n$, e se la quantità

$$C = \sup_{\Omega} \sum_{|\alpha|\leq m} |D^\alpha f|$$

è sufficientemente grande, allora la funzione $g = \frac{1}{C+1} a_n^{(n)} f$ appartiene a $V_{a^{(n)}}$ ma non a V_a.

(ii) Sia $\{f_i\}_{i\in I}$ un net di Cauchy in $C_0^m(\Omega)$. Allora per ogni successione infinitesima a di numeri positivi esiste $i_a \in I$ tale che

$$\sum_{|\alpha|\leq m} |D^\alpha f_i(x) - D^\alpha f_j(x)| < a_\nu \quad \forall x \in \Omega \setminus K_\nu\,,\ \forall \nu \in \mathbb{N},\ \forall i,j \geq i_a\,.$$

Proviamo a questo punto il seguente

Lemma 1.4.14 *Se $\{f_i\}_{i\in I}$ è un net di Cauchy in $C_0^m(\Omega)$, allora esistono un compatto $K \subset \Omega$ e un indice $i_0 \in I$ tali che il supporto di tutte le f_i, per $i \geq i_0$, è contenuto in K.*

Dimostrazione Supponiamo, per assurdo, che la tesi del lemma non sia vera: allora per ogni $\nu \in \mathbb{N}^+$ e per ogni $i \in I$ esisteranno una funzione

$f_{j_{\nu,i}} \in \{f_j\}_{j\geq i}$ e un punto $x_{\nu,i} \in \Omega \setminus K_\nu$ tali che $f_{j_{\nu,i}}(x_{\nu,i}) \neq 0$. Per ogni $i \in I$ consideriamo l'insieme

$$W_i = \left\{ g \in C_0^m(\Omega) : |g(x_{\nu,i})| < \frac{1}{\nu}|f_{j_{\nu,i}}(x_{\nu,i})| \quad \forall \nu \in \mathbb{N}^+ \right\}$$

e scegliamo una successione $a^{(i)} = \{a_\nu^{(i)}\}_{\nu\in\mathbb{N}} \in c_0^+$ con $a_\nu^{(i)} < \frac{1}{\nu}|f_{j_{\nu,i}}(x_{\nu,i})|$ per ogni $\nu \in \mathbb{N}^+$; allora si ha per ogni $i \in I$

$$V_{a^{(i)}} = \left\{ g \in C_0^m(\Omega) : \sum_{|\alpha|\leq m} |D^\alpha g(x)| < a_\nu^{(i)} \quad \forall x \in \Omega \setminus K_\nu\,,\ \forall \nu \in \mathbb{N} \right\} \subseteq W_i\,.$$

Dunque, per ogni $i \in I$ l'insieme W_i è un intorno di 0 e inoltre $f_{j_{\nu,i}} \notin \nu W_i$ per ogni $\nu \in \mathbb{N}^+$. Pertanto

$$\{f_j\}_{j\geq i} \nsubseteq \beta W_i \qquad \forall \beta > 0, \quad \forall i \in I.$$

D'altra parte, in virtù dell'esercizio 1.4.2, per ogni $i \in I$ esistono $h_i \in I$ e $\lambda_i > 0$ tali che $f_j \in \beta W_i$ per ogni $\beta > \lambda_i$ e per ogni $j \geq h_i$: in altre parole,

$$\{f_j\}_{j\geq h_i} \subseteq \beta W_i \qquad \forall \beta > \lambda_i, \quad \forall i \in I.$$

Le ultime due proprietà si contraddicono l'una con l'altra e ciò prova il lemma. □

Dal lemma 1.4.14 segue subito che esistono un compatto $K \subset \Omega$ e un indice $i_o \in I$ tali che il net $\{f_i\}_{i\geq i_0}$ è di Cauchy rispetto alla norma di $C_K^m(\Omega)$, e quindi converge in tale norma a una funzione $f \in C_K^m(\Omega)$. Ma allora $f \in C_0^m(\Omega)$; inoltre, per ogni successione $\{a_n\} \in c_0^+$ si ha definitivamente rispetto a i

$$\sum_{|\alpha|\leq m} |D^\alpha f_i(x) - D^\alpha f(x)| < a_n \quad \forall x \in \Omega \setminus K_n\,,\ \forall n \in \mathbb{N};$$

infatti se $\nu \in \mathbb{N}$ è tale che $K \subseteq K_\nu$ avremo

$$\sum_{|\alpha|\leq m} |D^\alpha f_i(x) - D^\alpha f(x)| = 0 < a_n \quad \forall x \in \Omega \setminus K_n\,,\ \forall n \geq \nu,\ \forall i \geq i_0\,,$$

mentre per $n = 0, 1, \ldots \nu - 1$ le condizioni corrispondenti, in numero finito, saranno soddisfatte definitivamente rispetto a i. Quindi $f_i \to f$ in $C_0^m(\Omega)$. □

Esempio 1.4.15 In modo sostanzialmente analogo si prova che lo spazio $C_0^\infty(\Omega)$ delle funzioni di classe C^∞ aventi supporto compatto nell'aperto $\Omega \subseteq \mathbb{R}^N$ è uno spazio di Fréchèt rispetto alla topologia definita dalla seguente famiglia di seminorme: fissata una successione di compatti $\{K_n\}$ con $K_0 = \emptyset$ e $\bigcup_{n\in\mathbb{N}} K_n = \Omega$, per ogni successione $a \in c_0^+$ e per ogni successione divergente $m = \{m_\nu\}_{\nu\in\mathbb{N}}$ di interi positivi si pone

$$p_{m,a}(f) = \sup_{\nu\in\mathbb{N}} \sup_{x\in\Omega\setminus K_\nu} \frac{1}{a_\nu} \sum_{|\alpha|\le m_\nu} |D^\alpha f(x)|.$$

Questa topologia è la più fine tra quelle che rendono continue tutte le immersioni $C_K^\infty(\Omega) \subseteq C^\infty(\Omega)$, ove $C_K^\infty(\Omega)$ è il sottospazio di $C^\infty(\Omega)$ formato dalle funzioni il cui supporto è contenuto nel compatto $K \subset \Omega$. La verifica di queste proprietà è analoga alla dimostrazione della proposizione 1.4.12 e del teorema 1.4.13.
Gli spazi $C_0^\infty(\Omega)$ e $C_K^\infty(\Omega)$ sono anche comunemente denotati con $\mathcal{D}(\Omega)$ e $\mathcal{D}_K(\Omega)$.

Esempio 1.4.16 Sia $\mathcal{S}(\mathbb{R}^N)$ lo spazio di Schwartz delle funzioni di classe C^∞ "a decrescenza rapida": $f \in \mathcal{S}(\mathbb{R}^N)$ se e solo se

$$x \mapsto |x|^k D^\alpha f(x) \in L^\infty(\mathbb{R}^N) \qquad \forall \alpha \in \mathbb{N}^N, \quad \forall k \in \mathbb{N}.$$

Definiamo in $\mathcal{S}(\mathbb{R}^N)$ la famiglia di seminorme

$$p_{m,k}(f) = \sup_{|\alpha|\le m} \sup_{x\in\mathbb{R}^N} (1+|x|)^k |D^\alpha f(x)|, \qquad m,k \in \mathbb{N}.$$

È chiaro che queste seminorme separano i punti di $\mathcal{S}(\mathbb{R}^N)$, e che tale spazio è localmente convesso e metrizzabile. Si noti che esso è un sottospazio di $C^\infty(\mathbb{R}^N)$, ma che la sua topologia è strettamente più fine di quella indotta da $C^\infty(\mathbb{R}^N)$: infatti una successione $\{f_n\} \subset \mathcal{S}(\mathbb{R}^N)$ converge a 0 in $\mathcal{S}(\mathbb{R}^N)$ se e solo se le successioni $x \mapsto |x|^k D^\alpha f_n(x)$ convergono a 0 uniformemente in $\mathbb{R}^N$, qualunque sia $k \in \mathbb{N}$ e $\alpha \in \mathbb{N}^N$. In particolare tutte le derivate di f_n convergono uniformemente a 0 in $\mathbb{R}^N$ e ciò implica immediatamente che $f_n \to 0$ in $C^\infty(\mathbb{R}^N)$.
Proviamo che lo spazio $\mathcal{S}(\mathbb{R}^N)$ è completo. Sia $\{f_n\}$ una successione di Cauchy in $\mathcal{S}(\mathbb{R}^N)$: allora $\{f_n\}$ è anche una successione di Cauchy in $C^\infty(\mathbb{R}^N)$ e quindi converge in $C^\infty(\mathbb{R}^N)$ a una funzione $f \in C^\infty(\mathbb{R}^N)$. Dobbiamo provare

che $f \in \mathcal{S}(\mathbb{R}^N)$ e che $f_n \to f$ in $\mathcal{S}(\mathbb{R}^N)$.
Fissiamo $m, k \in \mathbb{N}$ e $\varepsilon > 0$: esiste un indice $\nu = \nu_{m,k} \in \mathbb{N}$ tale che

$$p_{m,k}(f_n - f_h) \leq \varepsilon \qquad \forall n, h \geq \nu;$$

scelto $h = \nu$, ne segue $p_{m,k}(f_n) \leq \varepsilon + p_{m,k}(f_\nu) = C_{m,k}$ per ogni $n \geq \nu$, e in particolare

$$\sup_{|\alpha| \leq m} |D^\alpha f_n(x)| \leq C_{m,k}(1 + |x|)^{-k} \qquad \forall x \in \mathbb{R}^N, \quad \forall n \geq \nu.$$

Poiché, in virtù della convergenza delle f_n in $C^\infty(\mathbb{R}^N)$, le derivate di f_n convergono uniformemente sui compatti alle corrispondenti derivate di f, si deduce che

$$\sup_{|\alpha| \leq m} |D^\alpha f(x)| \leq C_{m,k}(1 + |x|)^{-k} \qquad \forall x \in \mathbb{R}^N$$

e pertanto $f \in \mathcal{S}(\mathbb{R}^N)$. Ma dalla condizione di Cauchy sopra scritta deduciamo anche

$$\sup_{|\alpha| \leq m} |D^\alpha (f_n - f_h)(x)| \leq \varepsilon(1 + |x|)^{-k} \qquad \forall x \in \mathbb{R}^N, \quad \forall n, h \geq \nu,$$

e per $h \to \infty$ ricaviamo facilmente

$$p_{m,k}(f_n - f) < \varepsilon \qquad \forall n \geq \nu.$$

Per l'arbitrarietà di m e k si ha dunque $f_n \to f$ in $\mathcal{S}(\mathbb{R}^N)$. Ciò prova che $\mathcal{S}(\mathbb{R}^N)$ è uno spazio di Fréchèt.

Esercizi 1.4

1. Un sottoinsieme B di uno spazio vettoriale topologico X si dice *limitato* se per ogni intorno U di 0 esiste $\lambda > 0$ tale che $B \subseteq \alpha U$ per ogni $\alpha > \lambda$. Si provi che:

   (i) se A e B sono limitati, allora $A + B$ è limitato;

   (ii) se A è limitato e $t \in \mathbb{R}$, allora tA è limitato;

   (iii) se A è limitato allora $\overline{A}$ è limitato.

2. Sia X uno spazio vettoriale topologico e sia $\{x_i\}_{i\in I}$ un net di Cauchy contenuto in X.

   (i) Si provi che per ogni intorno W di 0 esistono $\lambda > 0$ e $i_0 \in I$ tali che $\{x_i\}_{i\geq i_0} \subseteq \beta W$ per ogni $\beta > \lambda$.

   (ii) Si mostri con un esempio che $\{x_i\}_{i\in I}$ non è necessariamente limitato in X.

   [**Traccia:** per (i) si supponga, come è sempre possibile, che W sia bilanciato e si determini un altro intorno V di 0 tale che $V + V \subseteq W$; si mostri poi che esiste $i_0 \in I$ tale che $x_i \in x_{i_0} + V$ per ogni $i \geq i_0$. Si scelga adesso $s > 1$ in modo che $x_{i_0} \in sV$ (esercizio 1.3.5), e si deduca che $x_i \in tW$ per ogni $t > s$ e $i \geq i_0$.]

3. Sia X l'insieme delle funzioni misurabili $f : [0, 1] \to \mathbb{R}$, con la topologia generata dalla famiglia di seminorme $p_x(f) = |f(x)|$ al variare di $x \in [0, 1]$. Si provi che $f_n \to f$ in X se e solo se $f_n \to f$ puntualmente in $[0, 1]$.

4. Sia (X, τ) lo spazio vettoriale topologico definito nell'esercizio precedente, e sia σ la topologia su X indotta dalla "pseudometrica" seguente:

$$d(f, g) = \int_0^1 \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|}\, dx$$

   (si noti che d non è una distanza, ma definisce una topologia compatibile con le operazioni dello spazio vettoriale X). Si provi che:

   (i) l'applicazione identica $i : (X, \tau) \to (X, \sigma)$ trasforma insiemi τ-limitati in insiemi σ-limitati;

   (ii) se $\{f_n\}_{n\in\mathbb{N}} \subset X$ e $f_n \to 0$ rispetto a τ, allora $f_n \to 0$ rispetto a σ;

   (iii) l'applicazione i, tuttavia, non è continua.

   Se ne deduca che 0 non ha un sistema fondamentale di intorni numerabile nella topologia τ e che quindi (X, τ) non è metrizzabile.

5. Nello spazio $C(\mathbb{R})$ si considerino le seminorme

$$p_n(f) = \sup_{[-n,n]} |f|, \qquad n \in \mathbb{N}^+,$$

e la distanza indotta

$$d(f,g) = \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(f-g)}{1+p_n(f-g)} \ .$$

Posto $f(x) = \max\{0, 1-|x|\}$, $g(x) = 100f(x-2)$ e $h(x) = \frac{1}{2}(f(x) + g(x))$, si calcolino $d(f,0)$, $d(g,0)$ e $d(h,0)$ e si concluda che la palla di centro 0 e raggio $\frac{1}{2}$ non è convessa, benché lo spazio metrico $(C(\mathbb{R}), d)$ sia localmente convesso.

6. Dimostrare che per $p \in]0,1[$ lo spazio metrico ℓ^p è completo.

7. Sia $\{x_i\}_{i\in I}$ un net in uno spazio topologico X; sia poi $\{\gamma_i\}_{i\in I}$ un net di numeri reali tali che $|\gamma_i| \le K$ per ogni $i \in I$. Si provi che se $x_i \to 0$ in X, allora $\gamma_i x_i \to 0$ in X.

8. Sia X uno spazio vettoriale topologico. Si provi che:

   (i) se A, B sono sottoinsiemi compatti di X, allora $A+B$ è compatto;

   (ii) se A è compatto e B è chiuso, allora $A+B$ è chiuso;

   (ii) se A e B sono chiusi, non è detto che $A+B$ sia chiuso.

9. Si provi che i sottoinsiemi compatti di uno spazio vettoriale topologico sono limitati.

1.5 Operatori lineari e continui

Siano X, Y spazi vettoriali topologici. Descriveremo brevemente alcune proprietà che sono equivalenti alla continuità di un operatore lineare $f : X \to Y$. Anzitutto, è chiaro che f è continuo se e solo se è continuo nel punto $0 \in X$. Un'altra facile caratterizzazione è fornita dalla seguente

Proposizione 1.5.1 *Siano X, Y spazi vettoriali topologici, sia $f : X \to Y$ lineare. Allora f è continuo se e solo se per ogni net $\{x_i\}_{i\in I}$ convergente a 0 in X il net $\{fx_i\}_{i\in I}$ converge a 0 in Y.*

Dimostrazione ($\Longrightarrow$) Sia U un arbitrario intorno di 0 in Y. Per continuità, esiste un intorno V di 0 in X tale che $f(V) \subseteq U$. Sia $\{x_i\}_{i\in I} \subseteq X$ un net convergente a 0 in X: per definizione, esiste $i_0 \in I$ tale che $x_i \in V$ per ogni

$i \in I$, $i \geq i_0$. Dunque, $fx_i \in U$ per ogni $i \in I$, $i \geq i_0$ e pertanto il net $\{fx_i\}_{i\in I}$ converge a 0 in Y.
($\Longleftarrow$) Sia $\mathcal{V}$ la famiglia degli intorni di $0 \in X$. Se f non fosse continuo, esisterebbe un intorno U di 0 in Y tale che per qualunque $V \in \mathcal{V}$ potremmo trovare $x_V \in V$ per cui $fx_V \notin U$. Ma allora il net $\{x_V\}_{V\in\mathcal{V}}$ convergerebbe a 0 in X, senza che il net $\{fx_V\}_{V\in\mathcal{V}}$ converga a 0 in Y, il che contraddirebbe l'ipotesi. $\square$

Nel caso speciale che f sia lineare da X a $\mathbb{R}$ (oppure $\mathbb{C}$), vale la seguente caratterizzazione:

Teorema 1.5.2 *Sia X uno spazio vettoriale topologico e sia $f : X \to \mathbb{C}$ un funzionale lineare. Sono fatti equivalenti:*

(i) *f è continuo;*

(ii) $\ker f$ *(il* nucleo *di f) è chiuso in X;*

(iii) *esiste un intorno U di 0 in X tale che $f|_U$ è limitato, ossia esiste $M > 0$ tale che $|fx| < M$ per ogni $x \in U$.*

Dimostrazione (i)$\Longrightarrow$ (ii) $\ker f$ è l'immagine inversa del chiuso $\{0\} \in \mathbb{C}$, e quindi è chiuso in X perché f è continuo.
(ii)$\Longrightarrow$ (iii) Se $f \equiv 0$, la tesi è evidente. Altrimenti, deve essere $\ker f \neq X$, quindi $A = (\ker f)^c$ è un aperto non vuoto. Scelto $x \in A$, sia U un intorno di 0 in X tale che $x + U \subseteq A$; possiamo supporre che U sia bilanciato (osservazione 1.3.2). Allora si ha

$$f(U) = \bigcup_{|\alpha|\leq 1} f(\alpha U) = \bigcup_{|\alpha|\leq 1} \alpha f(U),$$

quindi $f(U)$, essendo bilanciato in $\mathbb{C}$, è un disco di centro 0: dunque $f(U)$ è un disco limitato, oppure $f(U)$ è tutto $\mathbb{C}$. Se, per assurdo, fosse $f(U) = \mathbb{C}$, esisterebbe $y \in U$ tale che $fy = -fx$, ossia $x+y \in \ker f$; ma ciò è impossibile perché $x + y \in x + U \subseteq A$. Perciò $f(U) = \{z \in \mathbb{C} : |z| < M\}$ e si ha la tesi.
(iii)$\Longrightarrow$ (i) Sia $|fx| < M$ per ogni $x \in U$, con U intorno di 0 in X. Allora, fissato $\varepsilon > 0$, l'insieme $W = \frac{\varepsilon}{M}U$ è intorno di 0 (in quanto l'applicazione $x \mapsto \lambda x$ è un omeomorfismo), e si ha $|fx| < \varepsilon$ per ogni $x \in W$. Ciò prova che f è continuo in 0. $\square$

Nel caso in cui le topologie negli spazi X, Y siano generate da seminorme, la continuità di un operatore lineare si può descrivere in un modo particolarmente comodo, che generalizza la nozione di limitatezza per operatori lineari fra spazi normati.

Teorema 1.5.3 *Siano X, Y spazi vettoriali topologici, con topologie generate dalle famiglie di seminorme $\{p_i\}_{i\in I}$ e $\{q_j\}_{j\in J}$ rispettivamente. Sia $f : X \to Y$ un operatore lineare; allora f è continuo se e solo se per ogni $j \in J$ esistono $K_j > 0$, $n_j \in \mathbb{N}^+$ e $i_1, \ldots, i_{n_j} \in I$ tali che*

$$q_j(fx) \le K_j \max_{1\le h\le n_j} p_{i_h}(x) \qquad \forall x \in X.$$

Dimostrazione ($\Longleftarrow$) Per ipotesi, per ogni fissato $j \in J$ esistono $K_j > 0$, $n_j \in \mathbb{N}^+$ e $i_1, \ldots, i_{n_j} \in I$ per i quali

$$q_j(fx) \le K_j \max_{1\le h\le n_j} p_{i_h}(x) \qquad \forall x \in X.$$

Sia $\varepsilon > 0$. Sia U un intorno di $0 \in Y$, che sarà della forma

$$U = U_{j_1,\ldots,j_m,\varepsilon} = \bigcap_{s=1}^{m} q_{j_s}^{-1}([0, \varepsilon[),$$

e consideriamo l'insieme

$$V = \left\{ x \in X : p_{i_h}(x) < \frac{\varepsilon}{K_{j_s}},\ h = 1, \ldots, n_{j_s},\ s = 1, \ldots, m \right\},$$

il quale è un intorno di $0 \in X$. Allora la disuguaglianza precedente mostra che $fx \in U$ per ogni $x \in V$. Quindi f è continuo nel punto $0 \in X$.

($\Longrightarrow$) Siano $\varepsilon > 0$, $j \in J$ ed $U = q_j^{-1}([0, \varepsilon[)$. Poiché f è continuo, esiste un intorno W di 0 tale che $f(W) \subseteq U$, ossia esistono $\delta > 0$ e $i_1, \ldots, i_n \in I$ tali che

$$x \in X, \quad \max_{1\le h\le n} p_{i_h}(x) < \delta \qquad \Longrightarrow \qquad q_j(fx) < \varepsilon.$$

Sia ora $z \in X$ e poniamo

$$w = \frac{\delta}{\max_{1\le h\le n} p_{i_h}(z) + \eta}\, z,$$

dove η è un arbitrario numero positivo (questo numero η serve a dar senso alla frazione nel caso che si abbia $p_{i_h}(z) = 0$ per $h = 1, \ldots, n$). Allora

$$p_{i_h}(w) = \frac{\delta}{\max_{1\le h\le n} p_{i_h}(z) + \eta}\, p_{i_h}(z);$$

quindi

$$\max_{1 \le h \le n} p_{i_h}(w) < \delta,$$

da cui

$$q_j(fw) < \varepsilon,$$

ossia

$$q_j(fz) = \frac{\max_{1 \le h \le n} p_{i_h}(z) + \eta}{\delta}\, q_j(fw) < \frac{\varepsilon}{\delta} \left(\max_{1 \le h \le n} p_{i_h}(z) + \eta \right).$$

Questa relazione vale per ogni $\eta > 0$; ne segue la tesi prendendo $n_j = n$ e $K_j = \frac{\varepsilon}{\delta}$. □

Quando X e Y sono spazi normati, la continuità di un operatore $f : X \to Y$ equivale alla sua limitatezza; una condizione analoga (ma non identica!) vale anche nel caso di due spazi vettoriali topologici: ricordiamo che la nozione di sottoinsieme limitato in uno spazio vettoriale topologico è stata introdotta nell'esercizio 1.4.1.

Proposizione 1.5.4 *Siano X, Y spazi vettoriali topologici, sia $f : X \to Y$ un operatore lineare. Se f è continuo, allora f è* limitato, *ossia trasforma insiemi limitati di X in insiemi limitati di Y.*

Si noti che il viceversa di questa implicazione è falso (esercizio 1.4.4), ma è vero sotto opportune ipotesi di metrizzabilità (esercizio 1.5.1).

Dimostrazione Sia U un intorno di $0 \in Y$: per continuità, esiste un intorno V di $0 \in X$ tale che $f(V) \subseteq U$. Sia E limitato in X: per definizione esiste $\lambda > 0$ tale che $E \subseteq \alpha V$ per ogni $\alpha > \lambda$; ne segue

$$f(E) \subseteq f(\alpha V) = \alpha f(V) \subseteq \alpha U \qquad \forall \alpha > \lambda.$$

Per l'arbitrarietà di U, $f(E)$ è limitato. □

Consideriamo ora lo spazio vettoriale dei funzionali lineari e continui su uno spazio vettoriale topologico.

Definizione 1.5.5 *Il* duale *di uno spazio vettoriale topologico (X, τ) è lo spazio vettoriale X^* dei funzionali lineari e continui su X, ossia degli operatori lineari e continui da X in $\mathbb{R}$ (oppure da X in $\mathbb{C}$).*

Si noti che la continuità di un operatore lineare su X dipende dalla topologia τ, quindi X^* dipende a sua volta da τ e dovremmo denotarlo con X^*_τ; tuttavia se τ si pensa fissata, la notazione X^* non si presta ad equivoci.
Di per sé, X^* è uno spazio vettoriale privo di una propria topologia naturale (salvo che nel caso in cui X sia uno spazio normato). I sottoinsiemi di X^* determinano altre topologie su X, diverse da τ: infatti si osserva anzitutto che ogni applicazione lineare $\varphi : X \to \mathbb{C}$ individua su X la seminorma $p_\varphi(x) = |\varphi x|$, $x \in X$; di conseguenza, fissato un arbitrario sottoinsieme B di X^*, la *topologia indotta da B su X* è la topologia generata dalla famiglia di seminorme $\{p_\varphi\}_{\varphi \in B}$ e si indica con $\sigma(X, B)$. In particolare, quando $B = X^*$ si ha la *topologia debole* su X, cioè $\sigma(X, X^*)$,che viene spesso denotata con w (benché dipenda anche, attraverso X e X^*, da τ); essa verrà ampiamente descritta nel paragrafo che segue.
Si possono definire in modo analogo svariate topologie su X^* che lo rendono uno spazio vettoriale topologico localmente convesso: infatti per ogni $x \in X$ la funzione $q_x(\varphi) = |\varphi x|$, $\varphi \in X^*$, è una seminorma su X^*; quindi, fissato un sottoinsieme D di X, la famiglia di seminorme $\{q_x\}_{x \in D}$ genera una topologia su X^*. che si denota con $\sigma(X^*, D)$. In particolare è importante la *topologia debole** $\sigma(X^*, X)$ (si legge "debole-star"), che si ottiene scegliendo $D = X$, e che verrà analizzata nel prossimo paragrafo.
Un'altra importante topologia su X^* è la *topologia forte*, che è generata dalla più ricca famiglia di seminorme $\{p_B\}_{B \in \mathcal{B}}$ definite da

$$p_B(f) = \sup_{x \in B} |fx| \qquad \forall f \in X^*,$$

ove $\mathcal{B}$ è la famiglia dei sottoinsiemi limitati di X. Queste seminorme sono ben definite in virtù della proposizione 1.5.4.
Ovviamente, la topologia forte è più fine di quella debole*; in particolare, allorché X è uno spazio di Banach la topologia forte è quella indotta dalla norma di X^*.

Esempio 1.5.6 Consideriamo lo spazio $C_0^\infty(\Omega)$, altrimenti detto $\mathcal{D}(\Omega)$, introdotto nell'esempio 1.4.15. Il suo duale si denota con $\mathcal{D}'(\Omega)$ e gli elementi del duale si chiamano *distribuzioni* su Ω. Dalla definizione delle seminorme $p_{m,a}$ è facile dimostrare (esercizio 1.5.3) che un funzionale lineare T definito su $\mathcal{D}(\Omega)$ è una distribuzione se e solo se per ogni compatto $K \subset \Omega$ esistono due costanti $C > 0$ e $m \in \mathbb{N}$ tali che

$$|Tf| \le C \sum_{|\alpha| \le m} \sup_{x \in K} |D^\alpha f(x)| \qquad \forall f \in \mathcal{D}_K(\Omega).$$

Nel caso $\Omega = \mathbb{R}^N$, lo spazio $\mathcal{D}(\mathbb{R}^N)$ è immerso con continuità nello spazio di Schwartz $\mathcal{S}(\mathbb{R}^N)$ introdotto nell'esempio 1.4.16; dunque il duale dello spazio di Schwartz, che si denota con $\mathcal{S}'(\mathbb{R}^N)$, è immerso con continuità in $\mathcal{D}'(\mathbb{R}^N)$. Gli elementi di $\mathcal{S}'(\mathbb{R}^N)$ si dicono *distribuzioni temperate.*

Esercizi 1.5

1. Siano X, Y spazi vettoriali topologici e supponiamo che X sia metrizzabile, con distanza d tale che $d(kx, 0) \le k\ d(x, 0)$ per ogni $k \in \mathbb{N}$ e per ogni $x \in X$. Sia inoltre $f : X \to Y$ un'applicazione lineare che trasforma insiemi limitati di X in insiemi limitati di Y. Si provino i seguenti fatti:

   (i) se $\{x_n\}_{n\in\mathbb{N}} \subset X$ e $x_n \to 0$ in X, allora $\{fx_n\}_{n\in\mathbb{N}}$ è limitata in Y;

   (ii) se $\{x_n\}_{n\in\mathbb{N}} \subset X$ e $x_n \to 0$ in X, allora esiste $\{\gamma_n\}_{n\in\mathbb{N}} \subset]0, \infty[$ tale che $\gamma_n \to \infty$ e $\gamma_n x_n \to 0$ in X;

   (iii) se $\{x_n\}_{n\in\mathbb{N}} \subset X$ e $x_n \to 0$ in X, allora $fx_n \to 0$ in Y;

   (iv) f è continua.

2. Sullo spazio vettoriale $X = C[0, 1]$ definiamo il seguente sistema fondamentale di intorni di 0:

   $$V_{n,k} = \left\{ f \in X : \ m \left\{ x \in [0, 1] : |f(x)| < \frac{1}{n} \right\} \ge 1 - \frac{1}{k} \right\}, \ n, k \in \mathbb{N}^+,$$

   ove m è la misura di Lebesgue.

   (i) Si verifichi che X è uno spazio vettoriale topologico metrizzabile.

   (ii) Si mostri che $X^* = \{0\}$, ossia l'unico funzionale lineare e continuo su X è quello identicamente nullo.

   [**Traccia:** per (ii), fissato $F \in X^*$, si determini un intorno $V_{n,k}$ di 0 tale che $|F(f)| < 1$ per ogni $f \in V_{n,k}$. Quindi si utilizzi una partizione dell'unità con funzioni φ_j aventi supporto in intervalli di ampiezza minore di $\frac{1}{k}$, e si osservi che per ogni $f \in X$ e $\lambda \in \mathbb{R}$ si ha $\lambda\varphi_j f \in V_{n,k}$. Si mostri infine che $F(f) = 0$ sfruttando l'arbitrarietà di λ.]

3. Si provi che un funzionale lineare $T : \mathcal{D}(\Omega) \to \mathbb{C}$ è una distribuzione se e solo se per ogni compatto $K \subset \Omega$ esistono $C > 0$ e $m \in \mathbb{N}$ tali che
$$|Tf| \leq C \sum_{|\alpha| \leq m} \sup_{x \in K} |D^\alpha f(x)| \qquad \forall f \in \mathcal{D}_K(\Omega).$$

1.6 Topologie deboli

In uno spazio vettoriale topologico (X, τ), come abbiamo visto, è possibile definire, accanto alla sua topologia naturale, un'altra topologia meno fine: la topologia debole. Questo paragrafo è dedicato alla descrizione delle sue numerose proprietà, che hanno particolare rilevanza, sia teorica che applicativa, nel caso speciale in cui X è uno spazio di Banach.

Definizione 1.6.1 *Sia X uno spazio vettoriale topologico. Si chiama* topologia debole *su X la topologia generata dalla famiglia di seminorme $\{p_\varphi\}_{\varphi \in X^*}$ così definite:*
$$p_\varphi(x) = |\varphi x| \qquad \forall x \in X.$$
La topologia debole si denota con i simboli w ("weak") oppure $\sigma(X, X^*)$.

Dunque un sistema fondamentale di w-intorni di 0 è dato dagli insiemi della forma
$$V_{\varphi_1,\ldots,\varphi_n,\varepsilon} = \{x \in X : \ |\varphi_i x| \leq \varepsilon, \ i = 1, \ldots, n\},$$
al variare di $n \in \mathbb{N}^+$, $\varepsilon > 0$ e $\varphi_1, \ldots, \varphi_n \in X^*$. Ripetiamo ancora che la topologia debole dipende dalla topologia originaria τ di cui è dotato X attraverso la scelta dei funzionali di X^*: cambiando τ cambia lo spazio X^* dei funzionali lineari τ-continui.
Dalla definizione segue che la topologia debole w è la meno fine topologia su X che rende continui tutti i funzionali $\varphi \in X^*$; poiché tutti gli elementi di X^* sono continui per la topologia naturale di X, la topologia debole è meno fine di quest'ultima. Tuttavia, se X ha dimensione finita le due topologie coincidono (esercizio 1.6.7). In particolare, ogni insieme *debolmente chiuso* (cioè chiuso rispetto a w) è chiuso, mentre il viceversa non è vero in generale. Ad esempio, nello spazio di Banach ℓ^2 l'insieme $\Sigma = \{x \in \ell^2 : \|x\|_2 = 1\}$ è chiuso ma non debolmente chiuso: infatti, ricordando che la successione $\{e^{(n)}\}_{n \in \mathbb{N}}$ è contenuta in Σ e converge debolmente a 0 in ℓ^2, si deduce che 0 appartiene alla chiusura debole di Σ perché ogni suo w-intorno della forma $V_{\varphi_1,\ldots,\varphi_p,\varepsilon}$ contiene necessariamente qualcuno degli $e^{(n)}$.

Se X è uno spazio normato, da un noto corollario del teorema di Hahn-Banach segue che per ogni $x \in X \setminus \{0\}$ esiste $\varphi \in X^*$ tale che $\varphi(x) \neq 0$; quindi la topologia debole rende X uno spazio vettoriale topologico localmente convesso e di Hausdorff. Se in particolare X è uno spazio di Banach riflessivo, allora (X, w) è completo (esercizio 1.6.4). In generale, 0 non ha un sistema fondamentale di intorni numerabile e lo spazio (X, w) non è metrizzabile: si veda a questo proposito l'esercizio 1.6.15.
Il duale X^* di uno spazio vettoriale topologico X è esso stesso uno spazio vettoriale topologico (localmente convesso) con una delle due topologie (debole* e forte) introdotte alla fine del paragrafo precedente. La definizione di topologia debole si applica ovviamente anche al caso di X^*: è quella generata dalle seminorme $\{p_\alpha\}_{\alpha \in X^{**}}$ definite da

$$p_\alpha(\varphi) = |\alpha\varphi| \qquad \forall \varphi \in X^*.$$

Naturalmente, lo spazio X^{**} cambia a seconda che si scelga di mettere su X^* la topologia forte, oppure la topologia debole* (o altre ancora di cui non ci occuperemo). Richiamiamo la definizione di topologia debole*:

Definizione 1.6.2 *Sia X uno spazio vettoriale topologico, sia X^* il suo duale. Si chiama* topologia debole* *su X^* la topologia generata dalla famiglia di seminorme $\{p_x\}_{x \in X}$ così definite:*

$$p_x(\varphi) = |\varphi x| \qquad \forall \varphi \in X^*.$$

La topologia debole si denota con i simboli w^* oppure $\sigma(X^*, X)$.*

Un sistema fondamentale di w^*-intorni di 0 è fatto dagli insiemi della forma

$$V_{x_1,\ldots,x_n,\varepsilon} = \{\varphi \in X^* : \ |\varphi x_i| \leq \varepsilon, \ i = 1, \ldots, n\}.$$

Dato che, ovviamente, per ogni $\varphi \in X^* \setminus \{0\}$ esiste $x \in X$ tale che $\varphi x \neq 0$, le seminorme p_x separano i punti di X^* e dunque (X^*, w^*) è uno spazio vettoriale topologico localmente convesso e di Hausdorff. Se in particolare X è uno spazio di Banach, lo spazio (X^*, w^*) è anche completo (esercizio 1.6.4).
Quando X è uno spazio di Banach, ricordando come è definita l'immersione canonica $J : X \to X^{**}$, possiamo scrivere le seminorme che generano la topologia $\sigma(X^*, X)$ nel modo seguente:

$$p_x(\varphi) = |\varphi x| = |J_x\varphi| \qquad \forall \varphi \in X^*.$$

Quindi, la topologia debole* $\sigma(X^*, X)$ è la meno fine topologia su X^* che rende continui tutti i funzionali $J_x \in J(X)$, mentre, per definizione, *tutti* gli elementi di X^{**} sono continui su X^* per la topologia $\sigma(X^*, X^{**})$ (la topologia debole relativa alla topologia indotta dalla norma di X^*). Ne segue che su X^* la topologia debole* $\sigma(X^*, X)$ è meno fine della topologia debole $\sigma(X^*, X^{**})$, che a sua volta è meno fine di quella indotta dalla norma di X^*. Le due topologie deboli su X^* coincidono se e solo se $J(X) = X^{**}$, ossia se e solo se X è riflessivo.
Come si è già osservato, le topologie deboli in spazi normati sono importanti perché permettono di ottenere buoni teoremi di compattezza. Ne descriveremo adesso i principali, insieme ad alcuni risultati di metrizzabilità strettamente collegati.

Teorema 1.6.3 (di Banach-Alaoglu) *Sia X uno spazio normato e sia B^* la palla unitaria chiusa di X^*. Allora B^* è compatta nella topologia debole* $\sigma(X^*, X)$.*

Dimostrazione Per ogni $x \in X$ denotiamo con I_x l'intervallo chiuso reale $[-\|x\|_X, \|x\|_X]$. Consideriamo il prodotto cartesiano

$$P = \prod_{x \in X} I_x \,,$$

munito della topologia prodotto τ. Allora P, essendo prodotto di compatti, è τ-compatto; denotando i suoi elementi con $\{f(x)\}_{x \in X}$, possiamo descrivere tale insieme nel modo seguente:

$$P = \{f : X \to \mathbb{R} : \ |f(x)| \leq \|x\|_X \ \forall x \in X\}.$$

Poiché $B^* \subseteq P \cap X^*$, B^* eredita due topologie indotte, una da (X^*, w^*) e una da (P, τ). Proveremo che:

(a) B^* è chiuso in (P, τ),

(b) w^* coincide con τ su B^*.

Da (a) seguirà che B^* è τ-compatto, e da (b) dedurremo che B^* è w^*-compatto, ossia la tesi.
Proviamo dapprima (b). Siano $n \in \mathbb{N}^+$, $x_1, \ldots, x_n \in X$, $\delta_1, \ldots, \delta_n > 0$ e $g_0 \in B^*$: consideriamo i due insiemi

$$W_1 = \{g \in X^* : \ |gx_i - g_0 x_i| < \delta_i \,, \ i = 1, \ldots, n\},$$

$$W_2 = \{f \in P : \ |f(x_i) - g_0 x_i| < \delta_i \, , \ i = 1, \ldots, n\};$$

W_1 e W_2 sono generici intorni di g_0 nelle topologie w^* e τ rispettivamente. Dato che, banalmente, $W_1 \cap B^* = W_2 \cap B^*$, la proprietà (b) è provata. Dimostriamo (a). Sia f_0 un elemento della chiusura di B^* in (P, τ): proveremo che $f_0 : X \to \mathbb{C}$ è lineare. Dati $x, y \in X$, $\alpha, \beta \in \mathbb{C}$ ed $\varepsilon > 0$, l'insieme

$$\begin{aligned} U &= \{f \in P : |f(x) - f_0(x)| < \varepsilon \, , \ |f(y) - f_0(y)| < \varepsilon \, , \\ &\qquad\qquad |f(\alpha x + \beta y) - f_0(\alpha x + \beta y)| < \varepsilon\} \end{aligned}$$

è intorno di f_0 in (P, τ). Poiché f_0 è τ-aderente a B^*, esiste $f \in B^* \cap U$. Usando la linearità di f, ne segue

$$\begin{aligned} &|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)| = \\ &\quad = |(f_0 - f)(\alpha x + \beta y) - \alpha(f_0 - f)(x) - \beta(f_0 - f)(y)| < \\ &\quad < (1 + |\alpha| + |\beta|)\varepsilon, \end{aligned}$$

e per l'arbitrarietà di ε otteniamo che f_0 è lineare.
Resta da far vedere che $f_0 \in B^*$: fissati $\varepsilon > 0$ e $x \in X$ con $\|x\|_X \le 1$, sia

$$V = \{g \in P : \ |g(x) - f_0 x| < \varepsilon\}.$$

Dato che V è un τ-intorno di f_0, esiste $g \in B^* \cap V$; dunque si ha

$$|f_0 x| \le |f_0 x - gx| + |gx| < \varepsilon + \|x\|_X \le \varepsilon + 1,$$

e per l'arbitrarietà di ε, ricaviamo

$$|f_0 x| \le 1 \qquad \forall x \in X \text{ con } \|x\|_X \le 1.$$

Si conclude che $f_0 \in B^*$, e ciò prova (a). $\square$

Teorema 1.6.4 (di Eberlein-Smulyan) *Sia X uno spazio di Banach riflessivo, e sia B la palla unitaria chiusa di X. Allora B è compatta nella topologia debole $\sigma(X, X^*)$.*

Dimostrazione Sia $\{U_i\}_{i \in I}$ un ricoprimento w-aperto di B: per ogni $i \in I$ e per ogni $x \in U_i$ esiste un w-intorno V di x, della forma

$$V = \{y \in X : \ |\varphi_i(x - y)| < \varepsilon, \ i = 1, \ldots, k\},$$

che è contenuto in U_i. Applichiamo l'immersione canonica $J : X \to X^{**}$, che per ipotesi è un'isometria bigettiva: otteniamo che per ogni $i \in I$ e per ogni $J_x \in J(U_i)$ l'insieme

$$J(V) = \{J_y \in X^{**} : \ |J_x\varphi_i - J_y\varphi_i| < \varepsilon, \ i = 1, \dots, k\}$$

è contenuto in $J(U_i)$. Ma $J(V)$ è un intorno di J_x nella topologia debole* su X^{**}, e dunque $\{J(U_i)\}_{i\in I}$ è un ricoprimento w^*-aperto di $J(B)$, cioè della palla unitaria chiusa di X^{**}. Per il teorema di Banach-Alaoglu (teorema 1.6.3), $J(B)$ è w^*-compatta in X^{**}, quindi esiste un sottoricoprimento finito $\{J(U_{i_h})\}_{1\le h\le m}$ di $J(B)$. Di conseguenza, $\{U_{i_h}\}_{1\le h\le m}$ è un sottoricoprimento finito di B, estratto dal ricoprimento w-aperto originale $\{U_i\}_{i\in I}$. Pertanto B è w-compatta. $\square$

Corollario 1.6.5 *Sia X uno spazio normato. Allora:*

(i) *i limitati di X^* sono debolmente* relativamente compatti;*

(ii) *se X è uno spazio di Banach riflessivo, allora i limitati di X sono debolmente relativamente compatti.*

Dimostrazione Basta osservare che ogni limitato è incluso in una palla chiusa, ed applicare i due teoremi precedenti. $\square$

La topologia debole su uno spazio normato non è metrizzabile in generale, come abbiamo visto (esercizio 1.6.15); però vale il seguente risultato:

Teorema 1.6.6 *Sia X uno spazio normato e sia B^* la palla unitaria chiusa di X^*. Allora X è separabile se e solo se la topologia indotta su B^* dalla topologia debole* di X^* è metrizzabile.*

Dimostrazione ($\Longrightarrow$) Sia X separabile: fissata una successione $\{x_n\}_{n\in\mathbb{N}}$ densa nella palla unitaria B di X, poniamo

$$d(f, g) = \sum_{n=0}^{\infty} 2^{-n}|fx_n - gx_n|, \qquad f, g \in B^*.$$

È chiaro che d è una distanza su B^* e che è invariante per traslazioni. Quindi per provare la tesi sarà sufficiente mostrare che:

(a) ogni palla $Q_\varepsilon = \{f \in B^* : d(f, 0) < \varepsilon\}$ contiene $B^* \cap U$, ove U è un opportuno intorno di 0 in $\sigma(X^*, X)$;

(b) ogni intorno U di 0 in $\sigma(X^*, X)$ contiene una palla Q_ε , con ε opportuno.

Proviamo (a): fissato un arbitrario $\varepsilon > 0$, sia $N \in \mathbb{N}$ tale che $2^{-N} < \frac{\varepsilon}{4}$. Scegliamo

$$U = \left\{ f \in X^* : \ |fx_k| < \frac{\varepsilon}{4}, \ k = 0, \dots, N-1 \right\};$$

allora se $f \in B^* \cap U$ si ha

$$d(f, 0) = \sum_{n=0}^{N-1} 2^{-n} |fx_n| + \sum_{n=N}^{\infty} 2^{-n} |fx_n| < \frac{\varepsilon}{4} \sum_{n=0}^{N-1} 2^{-n} + \sum_{n=N}^{\infty} 2^{-n} \|f\|_{X^*} < \varepsilon,$$

cioè $f \in Q_\varepsilon$. Quindi $B^* \cap U \subseteq Q_\varepsilon$.
Proviamo (b): fissati $\delta > 0$ e $y_1, \dots, y_m \in X \setminus \{0\}$, poniamo

$$U = \{ f \in X^* : \ |fy_i| < \delta, \ i = 1, \dots, m \}$$

e cerchiamo Q_ε tale che $Q_\varepsilon \subseteq U$. Poiché $\{x_n\}_{n \in \mathbb{N}}$ è densa in B, esistono $n_1, \dots, n_m \in \mathbb{N}$ tali che

$$\left\| \frac{y_i}{\|y_i\|_X} - x_{n_i} \right\|_X < \frac{\delta}{2K} \quad \text{per} \ \ i = 1, \dots, m, \quad \text{ove} \ \ K = \max_{1 \le i \le m} \|y_i\|_X \ .$$

Siano $N = \max_{1 \le i \le m} \{n_i\}$ ed $\varepsilon = \frac{2^{-N}\delta}{2K}$; si ha per ogni $f \in Q_\varepsilon$ e per ogni $i = 1, \dots, m$

$$2^{-n_i} |fx_{n_i}| \le d(f, 0) < \varepsilon,$$

da cui

$$|fx_{n_i}| < 2^{n_i} \varepsilon \le 2^N \varepsilon = \frac{\delta}{2K} \, .$$

Pertanto per ogni $i = 1, \dots, m$ si ha, ricordando che $\|f\|_{X^*} = 1$,

$$\frac{|fy_i|}{\|y_i\|_X} \le \left| f \left(\frac{y_i}{\|y_i\|_X} - x_{n_i} \right) \right| + |fx_{n_i}| < \frac{\delta}{2K} + \frac{\delta}{2K} = \frac{\delta}{K}$$

ed infine

$$|fy_i| < \frac{\delta}{K} \|y_i\|_X \le \delta, \quad i = 1, \dots, m.$$

Ciò prova che $f \in U$, ossia $Q_\varepsilon \subseteq U$.

($\Longleftarrow$) Supponiamo che (B^*, w^*) sia metrizzabile; allora 0 ha un sistema fondamentale di intorni numerabile $\{U_n\}_{n\in\mathbb{N}}$, e si può supporre che tali U_n siano della forma

$$U_n = \{\varphi \in B^* : \ |\varphi x| < \varepsilon_n \ \forall x \in A_n\},$$

ove A_n è un sottoinsieme finito di X. Poniamo $A = \bigcup_{n\in\mathbb{N}} A_n$: dato che A è numerabile, il sottospazio chiuso $X_0 \subseteq X$ generato da A è separabile. Supponiamo ora che esista $x \in X \setminus X_0$: per un corollario del teorema di Hahn-Banach, vi sarebbe un funzionale $\varphi \in X^*$, di norma unitaria, tale che $\varphi x \neq 0$ e $\varphi z = 0$ per ogni $z \in X_0$. Ciò significherebbe che $\varphi \in U_n$ per ogni $n \in \mathbb{N}$, e questo implicherebbe che $\varphi \equiv 0$, contraddicendo il fatto che $\varphi x \neq 0$. Pertanto deve essere $X = X_0$ e dunque X è separabile. $\square$

Un'analoga proprietà di metrizzabilità vale negli spazi di Banach che hanno duale separabile.

Teorema 1.6.7 *Sia X uno spazio di Banach e sia B la palla unitaria chiusa di X. Allora X^* è separabile se e solo se la topologia indotta su B dalla topologia debole di X è metrizzabile.*

Dimostrazione ($\Longrightarrow$) Sia X^* separabile: allora è noto che anche X è separabile; dunque, per il teorema precedente, la palla unitaria B^* di X^*, munita della topologia debole*, è metrizzabile. Inoltre B^* è w^*-compatta in virtù del teorema di Banach-Alaoglu. Perciò, dato che gli spazi metrici compatti sono separabili, (B^*, w^*) è separabile. Sia $\{f_n\}_{n\in\mathbb{N}^+}$ una successione densa in (B^*, w^*), e definiamo la distanza

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} |f_n(x - y)|, \qquad x, y \in B.$$

Proviamo che lo spazio metrico (B, d) coincide con lo spazio (B, w).
Sia $i : (B, w) \to (B, d)$ l'identità: essa è continua poiché, fissata arbitrariamente una palla $B(x_0, \delta)$, esiste $N_\delta \in \mathbb{N}^+$ tale che per il w-intorno $V = \{x \in B : |f_n(x - x_0)| < \delta/2, \ n \leq N_\delta\}$ risulta per ogni $x \in V$

$$d(x, x_0) \leq \sum_{n \leq N_\delta} 2^{-n} |f_n(x - x_0)| + \sum_{n > N_\delta} 2^{-n} < \sum_{n \leq N_\delta} 2^{-n} \frac{\delta}{2} + \frac{\delta}{2} < \delta.$$

Perciò l'applicazione i è biunivoca e continua dallo spazio compatto (B, w) nello spazio di Hausdorff (B, d): dunque è un omeomorfismo, e in particolare $(B, d) = (B, w)$.

($\Longleftarrow$) Proviamo anzitutto due lemmi.

Lemma 1.6.8 *Sia X uno spazio vettoriale e siano $g, f_1, \dots, f_m : X \to \mathbb{C}$ funzionali lineari. Vale l'inclusione $\bigcap_{i=1}^{n} \ker f_i \subseteq \ker g$ se e solo se esistono $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ tali che $g = \sum_{i=1}^{m} \lambda_i f_i$.*

Dimostrazione ($\Longleftarrow$) Evidente.

($\Longrightarrow$) Sia F l'operatore lineare da X in $\mathbb{C}^m$ dato da $Fx = (f_1 x, \dots, f_m x)$, e definiamo $h : R(F) \to \mathbb{C}$ nel modo seguente:

$$h(z) = gx \qquad \forall z = Fx \in R(F).$$

Questa definizione è ben posta perché se $z = Fx = Fy$ si ha $x - y \in \ker F = \bigcap_{i=1}^{n} \ker f_i$ e quindi, per ipotesi, $gx = gy$. Si noti che $R(F)$ è un sottospazio di $\mathbb{C}^m$ e quindi, avendo dimensione finita, è chiuso. Per il teorema di Riesz, applicato allo spazio di Hilbert $R(F) \subseteq \mathbb{C}^m$, esiste un unico elemento $v \in R(F)$ tale che $hz = \langle z, v \rangle_m$ per ogni $z \in R(F)$. Ne segue che il funzionale $H : \mathbb{C}^m \to \mathbb{C}$, definito da $Hz = \langle z, v \rangle_m$ per ogni $z \in \mathbb{C}^m$, estende h a tutto $\mathbb{C}^m$. In particolare, allora,

$$gx = h(Fx) = H(Fx) = \langle Fx, v \rangle_m = \sum_{i=1}^{m} f_i x \, \overline{v_i} \qquad \forall x \in X,$$

ovvero si ha la tesi con $\lambda_i = v_i$.

Lemma 1.6.9 *Sia X uno spazio di Banach e sia B la palla unitaria chiusa di X. Detta $J : X \to X^{**}$ l'immersione canonica, l'insieme $J(B)$ è $\sigma(X^{**}, X^*)$-denso nella palla unitaria chiusa B^{**} di X^{**} .*

Dimostrazione Sia, per assurdo, $\alpha \in B^{**} \setminus \overline{J(B)}^{w^*}$, ove $\overline{J(B)}^{w^*}$ è la chiusura di $J(B)$ rispetto a $\sigma(X^{**}, X^*)$. Esiste allora, per il teorema di Hahn-Banach, un funzionale $f : X^{**} \to \mathbb{C}$ lineare e $\sigma(X^{**}, X^*)$-continuo tale che

$$f\alpha \neq 0, \qquad f\beta = 0 \quad \forall \beta \in \overline{J(B)}^{w^*}.$$

Dimostriamo che esiste $\varphi \in X^*$ tale che

$$f\beta = \beta\varphi \qquad \forall \beta \in X^{**},$$

ovvero, in altri termini, proviamo che esiste $\varphi \in X^*$ tale che $f = J^*_\varphi$, essendo J^* l'immersione canonica di X^* nel proprio biduale. Ciò implicherà $0 = fJx = Jx\varphi = \varphi x$ per ogni $x \in B$, e quindi per linearità dedurremo $\varphi = 0$;

ma questo è assurdo essendo $\alpha\varphi = f\alpha \neq 0$.
Per provare l'asserzione osserviamo che f è $\sigma(X^{**}, X^*)$-continua, esiste un intorno di $0 \in X^{**}$ della forma $U = \{\alpha \in X^{**} : |\alpha\varphi_i| < \varepsilon,\ i = 1, \ldots, m\}$ tale che $|f\beta| < 1$ per ogni $\beta \in U$. Adesso, se $\beta \in \bigcap_{i=1}^m \ker J^*_{\varphi_i}$, ossia $\beta\varphi_i = 0$ per $i = 1, \ldots, m$, allora per ogni $k \in \mathbb{N}$ si ha $k|f\beta| = |f(k\beta)| < 1$, da cui $f\beta = 0$: ciò mostra che $\bigcap_{i=1}^m \ker J^*_{\varphi_i} \subseteq \ker f$. Ne segue, per il lemma 1.6.8, che esistono $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ tali che $f = \sum_{i=1}^m \lambda_i J^*_{\varphi_i}$. Dunque, posto $\varphi = \sum_{i=1}^m \lambda_i\varphi_i$, abbiamo provato che $f = J^*_\varphi$, ossia l'asserzione sopra scritta. Ciò conclude la dimostrazione del lemma. $\square$

Dimostriamo la seconda implicazione del teorema. Sia (B, w) metrizzabile: allora la topologia w ha come base d'intorni di 0 la famiglia numerabile $\{U_n\}_{n\in\mathbb{N}}$, con
$$U_n = \{x \in X : |\varphi x| < \varepsilon_n\ \forall\varphi \in A_n\},$$
ove $\varepsilon_n > 0$ e A_n è un sottoinsieme finito di X^*. Il sottospazio chiuso $X' \subseteq X$, generato da tutti i funzionali appartenenti all'unione degli A_n, è separabile: vogliamo provare che $X' = X^*$, e che pertanto X^* è separabile.
Supponiamo per assurdo che esista $\varphi \in X^+ \setminus X'$: allora
$$\delta = \inf_{\psi\in X'} \|\varphi - \psi\|_{X^*} > 0.$$
Per uno dei corollari del teorema di Hahn-Banach, esiste $\alpha \in X^{**}$ tale che
$$\|\alpha\|_{X^{**}} = \frac{1}{\delta}, \quad \alpha\varphi = 1, \quad \alpha|_{X'} = 0.$$
Consideriamo l'insieme $U = \{x \in B : |\varphi x| < \delta/2\}$, che è un w-intorno di $0 \in B$: esiste allora $n \in \mathbb{N}$ tale che $V \supseteq U_n$. Gli insiemi
$$V_n = \{\beta \in X^{**} : |\beta\psi| < \varepsilon_n\ \forall\psi \in A_n\}, \quad V = \{\beta \in X^{**} : |\beta\varphi| < \delta/2\}$$
sono $\sigma(X^{**}, X^*)$-intorni di $0 \in X^{**}$. Poiché $\delta\alpha \in B^{**}$, per il lemma 1.6.9 esiste $x_1 \in B$ tale che $J_{x_1} - \delta\alpha \in V_n \cap V$, ossia
$$|\psi x_1| = |J_{x_1}\psi - \delta\alpha\psi| < \varepsilon_n \quad \forall\psi \in A_n\,, \quad |\varphi x_1 - \delta| = |J_{x_1}\varphi - \delta\alpha\varphi| < \delta/2.$$
Ma allora $|\varphi x_1| > \delta/2$: dunque si ha $x_1 \in U_n \setminus V$, il che è assurdo essendo $U_n \subseteq V$. Ne segue $X^* = X$ e dunque la tesi. $\square$

Proviamo infine il seguente risultato, che va confrontato con quello dell'esercizio 1.6.14:

Teorema 1.6.10 *Sia X uno spazio normato con X^* separabile, e sia $M \subset X$ un insieme limitato. Allora la chiusura debole di M coincide con la chiusura debole sequenziale di M, cioè con l'insieme*

$$M^s = \{x \in X : \exists \{x_n\} \subseteq M : x_n \rightharpoonup x\}.$$

Dimostrazione È facile verificare che M^s è contenuto nella chiusura debole di M. Per provare l'inclusione opposta faremo uso del seguente

Lemma 1.6.11 *Sia X uno spazio normato, sia $M \subset X$ limitato e sia $u \in X$ un elemento della chiusura debole di M. Allora esiste un sottoinsieme al più numerabile $M_0 \subseteq M$ tale che u appartiene alla chiusura debole di M_0.*

Dimostrazione Sia B^* la palla unitaria chiusa di X^* e sia

$$B_n^* = B^* \times B^* \times \cdots \times B^* \quad (n \text{ volte}).$$

Per $m, n \in \mathbb{N}^+$ e $f = (f_1, \ldots, f_n) \in B_n^*$ consideriamo il w-intorno

$$U_{m,n,f}(u) = \left\{ v \in X : |f_j(v-u)| < \frac{1}{m} \,,\ j = 1, \ldots, n \right\}.$$

Per ipotesi, per ogni $m, n \in \mathbb{N}^+$ e $f \in B_n^*$ esiste $v \in M$ tale che $v \in U_{m,n,f}(u)$. Riscriviamo questa affermazione in modo diverso: per ogni $m, n \in \mathbb{N}^+$ e $v \in M$ sia

$$W_{m,n,v} = \{f \in B_n^* : v \in U_{m,n,f}(u)\};$$

allora possiamo dire che per ogni $m, n \in \mathbb{N}^+$ l'insieme B_n^* è ricoperto dalla famiglia $\{W_{m,n,v}\}_{v \in M}$.
Ora osserviamo che i $W_{m,n,v}$ sono aperti in B_n^* per la topologia indotta dalla topologia debole* di $(X^*)^n$; infatti se $f \in W_{m,n,v}$ l'insieme

$$\left\{ g \in B_n^* : |f_i(u-v) - g_i(u-v)| < \frac{1}{m} - \max_{1 \le j \le n} |f_j(u-v)|,\ i = 1, \ldots, n \right\}$$

è un intorno di f in tale topologia che è contenuto in $W_{m,n,v}$. Poiché, per il teorema di Banach-Alaoglu (teorema 1.6.3), B^* è debolmente* compatto in X^*, anche B_n^* è debolmente* compatto in $(X^*)^n$. Quindi esiste una sottofamiglia finita di $\{W_{m,n,v}\}_{v \in M}$ che è ancora un ricoprimento di B_n^*: in altre parole, esiste un sottoinsieme finito $S_{m,n} \subseteq M$ tale che

$$\forall f = (f_1, \ldots, f_n) \in B_n^* \quad \exists v \in S_{m,n} : \quad |f_j(v-u)| < \frac{1}{m} \text{ per } j = 1, \ldots, n.$$

Posto $M_0 = \bigcup_{m,n \in \mathbb{N}^+} S_{m,n}$, l'insieme M_0 è numerabile; essendo contenuto in M, è anche limitato in X. Per definizione di M_0, per ogni $f \in B_n^*$ esiste $v \in M_0$ tale che $v \in U_{m,n,f}(u)$: dunque u appartiene alla chiusura debole di M_0. $\square$

Riprendiamo la dimostrazione del teorema 1.6.10. Sia $u \in X$ un elemento della chiusura debole di M in X. Sia M_0 il sottoinsieme numerabile di M determinato dal lemma 1.6.11: dunque u appartiene alla chiusura debole di M_0 in X. Poiché M_0 è limitato, esiste una palla chiusa B_R di X, di raggio R opportuno, tale che $M_0 \subseteq B_R$. Essendo X^* separabile per ipotesi, per il teorema 1.6.7 la topologia debole su B_R è indotta da una distanza. Poiché negli spazi metrici la chiusura di un insieme coincide con la chiusura sequenziale, deduciamo che u appartiene alla chiusura debole sequenziale di M_0, da cui in particolare la tesi. $\square$

Esercizi 1.6

1. Sia X uno spazio vettoriale topologico, sia $\varphi : X \to \mathbb{C}$ lineare. Si provi che $\varphi \in X^*$ se e solo se φ è w-continua.

2. Sia X uno spazio normato, sia $\alpha : X^* \to \mathbb{C}$ lineare. Si provi che $\alpha \in X^{**}$ se e solo se α è w-continua (ove w è la topologia debole su X^*), e che $\alpha \in J(X)$ se e solo se α è w^*-continua.

3. Per ogni $t \in [a, b]$ sia $F_t : C[a, b] \to \mathbb{R}$ definito da $F_t(f) = f(t)$ per ogni $f \in C[a, b]$. Si provi che l'applicazione $t \mapsto F_t$ è continua da $[a, b]$ in $((C[a, b])^*, w^*)$.

4. Sia X uno spazio di Banach. Si provi che:

   (i) (X^*, w^*) è uno spazio vettoriale topologico completo;

   (ii) se lo spazio X è riflessivo, allora anche (X, w) è uno spazio vettoriale topologico completo.

5. Si provi che se X è lo spazio normato $(c_{00}, \|\cdot\|_{\ell^2})$, allora (X^*, w^*) è uno spazio vettoriale topologico non completo.
   [**Traccia:** si considerino i funzionali $T_k : c_{00} \to \mathbb{R}$ definiti da $T_k x = \sum_{j=0}^{k} j x_j \ldots$]

6. Si provi che se X è lo spazio normato $(c_0, \|\cdot\|_{\ell^\infty})$, allora (X, w) è uno spazio vettoriale topologico non completo.
[**Traccia:** fissato $x \in \ell^\infty \setminus c_0$, si consideri la successione $\{x^{(n)}\}$ ove $x^{(n)} = \{x_0, x_1, \ldots, x_{n-1}, x_n, 0, 0, \ldots\}\ldots$]

7. Sia X uno spazio normato. Si provi che la topologia debole di X coincide con quella indotta dalla norma se e solo se X ha dimensione finita.

8. Sia X uno spazio normato, sia E un sottoinsieme di X. Si provi che E è limitato in X se e solo se E è w-limitato.

9. Sia X uno spazio normato, sia F un sottoinsieme di X^*. Si provi che F è limitato in X^* se e solo se F è w^*-limitato.

10. Sia X uno spazio normato e sia B la palla unitaria chiusa di X. Si provi che $J(B)$ è w^*-denso nella palla unitaria chiusa B^{**} di X^{**} (qui w^* è la topologia debole* di X^{**}).
Traccia: Sia $\alpha \in B^{**}$ e sia $V = \{\beta \in X^{**} : |\beta\varphi - \alpha\varphi| < \varepsilon\}$: si osservi che $\alpha\varphi \in [-\|\varphi\|_{X^*}, \|\varphi\|_{X^*}]$ e che $\varphi(B) \supseteq] - \|\varphi\|_{X^*}, \|\varphi\|_{X^*}[$, e si deduca che esiste $x \in B$ tale che $J_x \in V$.]

11. Si provi che $C[a,b]$ è w^*-denso in $L^\infty(a,b)$.

12. Si provi che la palla unitaria chiusa di c_0 non è w-compatta.

13. Sia (X, τ) uno spazio vettoriale topologico tale che l'origine abbia un sistema fondamentale di intorni numerabile. Si provi che per ogni sottoinsieme $E \subseteq X$ la chiusura debole di E e la chiusura debole sequenziale di E coincidono.

14. Siano $X = \ell^2$ e $E = \{e^{(m)} + me^{(n)} : \ m, n \in \mathbb{N}, \ m \le n\}$. Si dimostri che:

    (i) la chiusura debole sequenziale di E è $E \cup \{e^{(m)}\}_{m \in \mathbb{N}}$;

    (ii) la chiusura debole di E contiene anche 0.

15. Si provi che lo spazio vettoriale topologico (ℓ^2, w) non è metrizzabile.

16. Sia (X, τ) uno spazio vettoriale topologico. Denotando con X_w lo spazio vettoriale X dotato della topologia debole relativa a τ, si provi che $(X_w)_w = X_w$, cioè che la topologia debole relativa alla topologia debole coincide con questa.

1.7 Separazione di insiemi convessi

Iniziamo il paragrafo con l'estensione agli spazi vettoriali topologici del teorema di Hahn-Banach. Di questo risultato fondamentale conosciamo una versione "inutilmente" restrittiva, in cui l'ambiente è uno spazio normato. In effetti, senza cambiare una virgola della dimostrazione relativa a quel caso, si prova il seguente enunciato più generale:

Teorema 1.7.1 (di Hahn-Banach, caso reale) *Sia X uno spazio vettoriale reale, sia $p : X \to \mathbb{R}$ una funzione subadditiva e positivamente omogenea (ossia, tale che $p(\lambda x) = \lambda p(x)$ e $p(x + y) \leq p(x) + p(y)$ per ogni $\lambda > 0$ e $x, y \in X$); siano poi M un sottospazio di X e $f : M \to \mathbb{R}$ un funzionale lineare tale che $fx \leq p(x)$ per ogni $x \in M$. Allora esiste un funzionale lineare $F : X \to \mathbb{R}$ tale che $F|_M = f$ e $Fx \leq p(x)$ per ogni $x \in X$.* $\square$

Se X è uno spazio vettoriale complesso, il teorema di Hahn-Banach vale nella forma seguente:

Teorema 1.7.2 (di Hahn-Banach, caso complesso) *Sia X uno spazio vettoriale complesso, sia $p : X \to \mathbb{R}$ una funzione subadditiva e positivamente omogenea; siano poi M un sottospazio di X e $f : M \to \mathbb{C}$ un funzionale lineare tale che $\mathrm{Re}\, fx \leq p(x)$ per ogni $x \in M$. Allora esiste un funzionale lineare $F : X \to \mathbb{C}$ tale che $F|_M = f$ e $\mathrm{Re}\, Fx \leq p(x)$ per ogni $x \in X$.*

Dimostrazione Ci ricondurremo alla dimostrazione del caso reale mediante le seguenti osservazioni di immediata verifica. Se $F : X \to \mathbb{C}$ è un funzionale lineare, allora $U = \mathrm{Re}\, F : X \to \mathbb{R}$ è un funzionale $\mathbb{R}$-lineare (ossia è additivo e $U(\lambda x) = \lambda Ux$ per ogni $\lambda \in \mathbb{R}$). Viceversa, se $U : X \to \mathbb{R}$ è $\mathbb{R}$-lineare, allora il funzionale $F : X \to \mathbb{C}$ definito da

$$Fx = Ux - iU(ix), \qquad x \in X,$$

è lineare su X e, naturalmente, $\mathrm{Re}\, F = U$.
Nelle ipotesi sopra scritte, allora, il funzionale $u = \mathrm{Re}\, f$ verifica le ipotesi

del teorema 1.7.1. Quindi esiste un funzionale $U : X \to \mathbb{R}$, $\mathbb{R}$-lineare, tale che $U|_M = u$ e $Ux \le p(x)$ per ogni $x \in X$. Ne segue che il funzionale $Fx = Ux - iU(ix)$ è lineare su X e soddisfa la tesi. □

Del teorema di Hahn-Banach esiste una versione geometrica altrettanto importante e utile, che si esprime in termini di separazione di insiemi convessi mediante gli insiemi di livello di opportuni funzionali lineari. Si hanno due enunciati abbastanza simili: entrambi si dimostrano sfruttando l'enunciato "algebrico", cioè il teorema 1.7.2 (si noti che il teorema 1.7.2, che riguarda il caso complesso, si riduce automaticamente al teorema 1.7.1, relativo al caso reale, quando lo spazio X è reale).

Teorema 1.7.3 (di Hahn-Banach, 1ª forma geometrica) *Sia X uno spazio vettoriale topologico, e siano K, H sottoinsiemi non vuoti, convessi e disgiunti di X, con K aperto. Allora esiste un funzionale lineare $F : X \to \mathbb{C}$, continuo e non nullo, tale che*

$$\sup_K \operatorname{Re} F \le \inf_H \operatorname{Re} F .$$

Dimostrazione Cominciamo con un caso particolare.

Lemma 1.7.4 *Sia X uno spazio vettoriale topologico reale, sia K un convesso aperto e non vuoto di X e sia $x_0 \in X \setminus K$. Allora esiste un funzionale lineare $F : X \to \mathbb{R}$, continuo e non nullo, tale che*

$$Fx < Fx_0 \qquad \forall x \in K.$$

Dimostrazione Supponiamo dapprima che $0 \in K$. Allora si ha $x_0 \neq 0$, quindi possiamo definire un funzionale lineare f sul sottospazio $M = [\{x_0\}]$ ponendo

$$f(tx_0) = t \qquad \forall t \in \mathbb{R}.$$

Essendo il convesso K aperto e contenente l'origine, si vede facilmente che, detto p_K il funzionale di Minkowski di K, si ha $p_K(x) < +\infty$ per ogni $x \in X$. Inoltre risulta

$$f(tx_0) = t \le p_K(tx_0) \qquad \forall t \in \mathbb{R} :$$

infatti ciò è ovvio quando $t \le 0$, e d'altra parte, essendo $x_0 \notin K$, se $t > 0$ si ha $tx_0 \notin tK$ e dunque, per definizione, $p_K(tx_0) \ge t$.
Per il teorema 1.7.1 esiste un funzionale lineare $F : X \to \mathbb{R}$ tale che $F|_M = f$

e $F \leq p_K$ in X. Ricordando che $p_K(x) < 1$ per ogni $x \in K$ (esercizio 1.3.2), si deduce

$$Fx \leq p_K(x) < 1 = fx_0 = Fx_0 \qquad \forall x \in K$$

e in particolare il funzionale F non è identicamente nullo. Resta da provare che F è continuo: si ha $F < 1$ su K, da cui anche $F > -1$ sul convesso $-K$. Di conseguenza,

$$|Fx| < 1 \qquad \forall x \in K \cap (-K).$$

Così F risulta limitato sull'intorno $K \cap (-K)$ di 0. Dal teorema 1.5.2 segue che F è continuo.

Supponiamo ora che $0 \notin K$. Sia $y_0 \in K$ e consideriamo il traslato $K - y_0$, che è un convesso aperto contenente 0. Per quanto già provato, esiste un funzionale lineare, continuo e non nullo $F : X \to \mathbb{R}$ tale che

$$Fy < F(x_0 - y_0) \qquad \forall y \in K - y_0\,,$$

ossia

$$Fx = F(x - y_0) + Fy_0 < F(x_0 - y_0) + Fy_0 = Fx_0 \qquad \forall x \in K.$$

Ciò prova la tesi del lemma. $\square$

Torniamo alla dimostrazione del teorema 1.7.3. Poniamo

$$C = K - H = \{x - y : \ x \in K, \ y \in H\};$$

C è un convesso aperto non vuoto, e $0 \notin C$ perché H e K sono disgiunti. Applicando il lemma 1.7.4 con $x_0 = 0$ e $K = C$, si ottiene l'esistenza di un funzionale $\mathbb{R}$-lineare, continuo e non nullo $U : X \to \mathbb{R}$ tale che $Uz < 0$ per ogni $z \in C$, ossia

$$Ux < Uy \qquad \forall x \in K, \ \forall y \in H.$$

Posto $Fx = Ux - iU(ix)$, il funzionale $F : X \to \mathbb{C}$ è lineare, continuo e non nullo, e verifica la tesi. $\square$

Vi è una versione "stretta" del teorema di separazione 1.7.3, che è la seguente:

Teorema 1.7.5 (di Hahn-Banach, 2ª forma geometrica) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff, e siano K, H convessi non vuoti e disgiunti di X, con K chiuso e H compatto. Allora esiste un funzionale $F : X \to \mathbb{C}$ lineare, continuo e non nullo tale che*

$$\sup_K \operatorname{Re} F < \inf_H \operatorname{Re} F.$$

Dimostrazione Premettiamo il seguente lemma.

Lemma 1.7.6 *Se B è un compatto in uno spazio topologico di Hausdorff Z, allora ogni net $\{z_i\}_{i\in I} \subseteq B$ ha almeno un punto di accumulazione $z \in B$.*

Dimostrazione Poniamo $I_i = \{j \in I : i < j\}$ e $A_i = \{z_j : j \in I_i\}$: dunque $A_i \subseteq B$ per ogni $i \in I$. Stiamo cercando un punto $z \in B$ tale che per ogni intorno U di z e per ogni $i \in I$ esista $z_j \in A_i \cap U$; cioè stiamo cercando un punto $z \in B$ che appartenga a tutte le chiusure $\overline{A_i}$, $i \in I$. Dato che, per definizione di net, si ha $\bigcap_{h=1}^{n} A_{i_h} \neq \emptyset$ per ogni sottoinsieme finito $\{i_1, \dots i_n\} \subseteq I$, a maggior ragione risulta

$$\bigcap_{h=1}^{n} \overline{A_{i_h}} \neq \emptyset \qquad \forall i_1, \dots i_n \in I,$$

e per la compattezza di B si ottiene

$$\bigcap_{i\in I} \overline{A_i} \neq \emptyset.$$

Dunque esiste un punto $z \in \bigcap_{i\in I} \overline{A_i} \subseteq \overline{B}$. Tale z appartiene anche a B poiché B, essendo compatto nello spazio di Hausdorff Z, è chiuso. Ciò prova il lemma. $\square$

Proviamo il teorema 1.7.5. Sia $\{U_i\}_{i\in I}$ un sistema fondamentale di intorni di 0, con gli U_i aperti e convessi. Per ogni $i \in I$ poniamo $H_i = H + U_i$; allora H_i è un convesso aperto. Dimostriamo che esiste $i_0 \in I$ tale che $K \cap H_i = \emptyset$ per ogni $i \in I$, $i \geq i_0$. In caso contrario, per ogni $k \in I$ esisterebbe $i_k \geq k$ per cui $K \cap H_{i_k} \neq \emptyset$, ossia esisterebbero $x_{i_k} \in K$ e $y_{i_k} \in H$ tali che $x_{i_k} - y_{i_k} \in U_{i_k}$. Quindi il net $\{x_{i_k} - y_{i_k}\}_{k\in I}$ convergerebbe a 0 in X. D'altronde, essendo H compatto, il lemma 1.7.6 ci dice che il net $\{y_{i_k}\}_{k\in I} \subseteq H$ avrebbe un punto di accumulazione $y \in H$. Poiché $x_{i_k} - y_{i_k} \to 0$, il punto y sarebbe di accumulazione anche per il net $\{x_{i_k}\}_{k\in I} \subseteq K$; ma essendo K chiuso, dovrebbe essere $y \in K$, il che è assurdo perché H e K sono disgiunti. L'asserzione precedente è provata.
Pertanto i convessi H_{i_0} e K verificano le ipotesi del teorema 1.7.3: quindi esiste un funzionale $G : X \to \mathbb{C}$ lineare, continuo e non nullo tale che

$$\sup_H \operatorname{Re} G + \sup_{U_{i_0}} \operatorname{Re} G = \sup_{H_{i_0}} \operatorname{Re} G \leq \inf_K \operatorname{Re} G.$$

Osserviamo che $\sup_{U_{i_0}} \mathrm{Re}\, G > 0$: infatti, essendo G non identicamente nullo, esiste $z \in X \setminus \{0\}$ tale che $\mathrm{Re}\, Gz \neq 0$; per la continuità delle omotetie, esisterà $\vartheta \in \mathbb{R} \setminus \{0\}$ per cui $\vartheta z \in U_{i_0}$ e $\mathrm{Re}\, G(\vartheta z) > 0$. Quindi $\sup_{U_{i_0}} \mathrm{Re}\, G > 0$ e pertanto

$$\sup_H \mathrm{Re}\, G < \sup_{H_{i_0}} \mathrm{Re}\, G \leq \inf_K \mathrm{Re}\, G.$$

Posto $F = -G$, si ha allora la tesi. $\square$

Vogliamo ora riformulare i teoremi 1.7.3 e 1.7.5 in termini più geometrici. A questo scopo ci occorre anzitutto una caratterizzazione dei sottospazi di uno spazio vettoriale topologico X che sono luogo di zeri di funzionali lineari su X.

Definizione 1.7.7 *Sia X uno spazio vettoriale e sia M un sottospazio di X. Diciamo che M è* massimale *se non vi è alcun sottospazio Y di X che verifichi le inclusioni proprie $M \subset Y \subset X$.*

Proposizione 1.7.8 *Sia X uno spazio vettoriale e sia M un sottospazio di X. Allora M è massimale se e solo se esiste un funzionale lineare f su X tale che $M = \ker f$. In tal caso il funzionale f è unico a meno di costanti moltiplicative.*

Dimostrazione ($\Longleftarrow$) Se $M = \ker f$ con $f : X \to \mathbb{C}$ lineare, allora ovviamente M è un sottospazio di X. Se Y è un sottospazio di X che contiene propriamente M, fissiamo $y_0 \in Y \setminus M$: allora per ogni $x \in X$ possiamo scrivere, essendo $fy_0 \neq 0$,

$$x = \left(x - \frac{fx}{fy_0}\, y_0 \right) + \frac{fx}{fy_0}\, y_0 \,;$$

ma a secondo membro il primo addendo è un elemento di M, quindi di Y, ed il secondo addendo sta in Y. Ciò mostra che $X \subseteq Y$, ossia $X = Y$.

($\Longrightarrow$) Sia M un sottospazio massimale. Se $M = X$, allora il funzionale identicamente nullo ha M come luogo di zeri. Se $M \subset X$, fissiamo $z_0 \in X \setminus M$: allora il sottospazio $[\{z_0\}, M]$ contiene propriamente M, e quindi deve coincidere con X. Quindi ogni elemento $x \in X$ si può decomporre (in modo unico) nella forma

$$x = u + \lambda z_0 \,, \qquad u \in M,\ \lambda \in \mathbb{C}.$$

Definiamo il funzionale f nel modo seguente:

$$f(u + \lambda z_0) = \lambda.$$

Chiaramente f è ben definito e lineare; inoltre $\ker f = M$.
Resta da verificare che se $M = \ker f = \ker g$, con f e g funzionali lineari su X, allora g è multiplo di f; e in effetti se $z_0 \in X \setminus M$, allora

$$x - \frac{fx}{fz_0}\, z_0 \in \ker f = \ker g \qquad \forall x \in X,$$

da cui

$$g\left(x - \frac{fx}{fz_0}\, z_0\right) = 0 \qquad \forall x \in X,$$

ossia

$$gx = \frac{gz_0}{fz_0}\, fx \qquad \forall x \in X.$$

Ciò prova la tesi. $\square$

Osservazione 1.7.9 Se M è un sottospazio massimale dello spazio vettoriale topologico X, allora M è chiuso in X oppure denso in X. Infatti, $\overline{M}$ è un sottospazio che contiene M: se $\overline{M} = M$, allora M è chiuso; se invece $\overline{M} \supset M$, allora per la massimalità di M si ha $\overline{M} = X$, ossia M è denso in X.

Corollario 1.7.10 *Sia X uno spazio vettoriale topologico e sia f un funzionale lineare su X. Allora* $\ker f$ *è chiuso in X oppure denso in X a seconda che f sia continuo oppure no.*

Dimostrazione Se f è continuo, allora $\ker f = f^{-1}(0)$ è l'immagine inversa di un chiuso di $\mathbb{C}$ e quindi è chiuso. Se f non è continuo, allora non è continuo nel punto $0 \in X$: quindi esiste $\varepsilon > 0$ ed esiste un net $\{x_i\}_{i\in I}$ tale che $x_i \to 0$ in X ma $|fx_i| > \varepsilon$ per ogni $i \in I$. Fissiamo $x \in X$: posto

$$y_i = x - \frac{fx}{fx_i}\, x_i\,, \qquad i \in I,$$

il net $\{y_i\}_{i\in I}$ converge a x, come è facile verificare (esercizio 1.4.7), ed è contenuto in $\ker f$. Quindi $\ker f$ è denso in X. $\square$

Adesso introduciamo la nozione di iperpiano.

Definizione 1.7.11 *Sia X uno spazio vettoriale topologico. Un* iperpiano affine *di X è un insieme della forma*

$$\Pi = \{x \in X : \ \mathrm{Re}\, fx = k\}, \qquad k \in \mathbb{R},$$

ove f è un funzionale lineare su X.

Si noti che se X è reale, ogni iperpiano affine è il traslato di un sottospazio massimale: infatti se $f : X \to \mathbb{R}$ è lineare e non identicamente nullo, e $\Pi = \{x \in X : fx = k\}$, $k \in \mathbb{R}$, allora si ha $\Pi = \{x \in X : fx = 0\} + x_0$, ove x_0 è scelto in modo che $fx_0 = k$. Osserviamo che ciò è falso se X è complesso (esercizio 1.7.7).

Osservazione 1.7.12 Dal corollario 1.7.10 segue che un iperpiano affine è chiuso o denso in X a seconda che il corrispondente funzionale sia continuo o no. Un iperpiano affine chiuso è anche detto *varietà affine*. Si noti che ogni iperpiano affine chiuso della forma $\Pi = \{x \in X : \mathrm{Re}\, fx = k\}$, con k reale, divide X in due semispazi chiusi

$$\Pi^+ = \{x \in X : \ Re\, fx \geq k\}, \qquad \Pi^- = \{x \in X : \ Re\, fx \leq k\}.$$

I teoremi di Hahn-Banach 1.7.3 e 1.7.5 si possono allora riformulare nel modo seguente.

Corollario 1.7.13 *Sia X uno spazio vettoriale topologico.*

(i) *Se K, H sono convessi non vuoti e disgiunti in uno spazio vettoriale topologico X, con K aperto, allora esiste un iperpiano affine chiuso Π che separa H e K (nel senso che H e K sono contenuti uno in Π^+ ed uno in Π^-).*

(ii) *Se lo spazio vettoriale topologico X è localmente convesso di Hausdorff, e se K, H sono convessi non vuoti disgiunti con K chiuso e H compatto, allora esiste un iperpiano affine chiuso Π che separa* strettamente*H e K (ossia H e K sono contenuti nei due semispazi* aperti *individuati da Π).* $\square$

Vediamo altre proprietà geometriche degli insiemi convessi.

Definizione 1.7.14 *Sia X uno spazio vettoriale topologico, sia $K \subseteq X$. Un* iperpiano di appoggio *per K è un iperpiano affine chiuso Π tale che:*

(i) $\Pi \cap K \neq \emptyset$,

(ii) *K è contenuto in uno dei due semispazi chiusi individuati da* Π.

Un punto $x \in \Pi \cap K$ si dice punto di appoggio *di* Π *su* K.

Corollario 1.7.15 *Sia X uno spazio vettoriale topologico, sia K un convesso di X con parte interna non vuota. Allora ogni $x_0 \in \partial K$ è punto di appoggio su $\overline{K}$ di qualche iperpiano affine chiuso.*

Dimostrazione Gli insiemi $\{x_0\}$ e $\overset{\circ}{K}$ sono convessi disgiunti di X, con $\overset{\circ}{K}$ aperto. Per il teorema 1.7.3, esiste un funzionale lineare e continuo $F : X \to \mathbb{C}$ tale che

$$\operatorname{Re} Fx_0 \leq \operatorname{Re} Fx \qquad \forall x \in \overset{\circ}{K};$$

quindi, per continuità,

$$\operatorname{Re} Fx_0 \leq \operatorname{Re} Fx \qquad \forall x \in \overline{\overset{\circ}{K}},$$

il che implica, per l'esercizio 1.3.8,

$$\operatorname{Re} Fx_0 \leq \operatorname{Re} Fx \qquad \forall x \in K.$$

Pertanto, posto $\alpha = \operatorname{Re} Fx_0$, l'iperpiano affine chiuso $\Pi = \{u \in X : \operatorname{Re} Fu = \alpha\}$ è di appoggio per $\overline{K}$ e x_0 è punto di appoggio di Π su $\overline{K}$. □

Corollario 1.7.16 *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff. Se $K \subseteq X$ è un convesso chiuso non vuoto, allora K è l'intersezione di tutti i semispazi chiusi che lo contengono.*

Dimostrazione Sia H l'intersezione di tutti i semispazi chiusi contenenti K: ovviamente $K \subseteq H$. Sia $x_0 \in H$: se fosse $x_0 \notin K$, i convessi $\{x_0\}$ e K sarebbero disgiunti, con $\{x_0\}$ compatto e K chiuso. Per il teorema 1.7.5, esisterebbe un iperpiano affine chiuso Π che separa strettamente i due convessi; quindi fra i due semispazi chiusi individuati da Π ce ne sarebbe uno contenente K ma non il punto x_0, contro l'ipotesi $x_0 \in H$. □

Dai risultati precedenti deduciamo facilmente questa importante proprietà:

Corollario 1.7.17 *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff. Se K è un convesso di X, allora K è chiuso se e solo se K è w-chiuso.*

Dimostrazione Poiché la topologia debole è meno fine di quella naturale di X, ogni insieme w-chiuso è necessariamente chiuso.
Viceversa, supponiamo che K sia chiuso. Per il corollario 1.7.16, K è l'intersezione di tutti i semispazi chiusi che lo contengono. Ma un semispazio chiuso ha la forma $\{x \in X : \operatorname{Re} Fx \leq k\}$, con $F \in X^*$ e $k \in \mathbb{R}$; dato che w è la meno fine topologia che rende continui gli elementi di X^*, ogni insieme della forma sopra scritta è w-chiuso, essendo la controimmagine del chiuso $]-\infty, k] \subset \mathbb{R}$ mediante l'applicazione w-continua $\operatorname{Re} F$. Dunque ogni semispazio chiuso è w-chiuso, e dunque H, essendo intersezione di w-chiusi, è un insieme w-chiuso. $\square$

Corollario 1.7.18 *Sia X uno spazio normato, sia $\{x_n\}_{n\in\mathbb{N}}$ una successione contenuta in X, e sia $x \in X$ tale che $x_n \rightharpoonup x$ in X. Allora esiste una successione $\{v_n\}_{n\in\mathbb{N}}$, costituita da combinazioni convesse degli x_n, tale che $\|v_n - x\|_X \to 0$.*

Dimostrazione Poiché $x_n \rightharpoonup x$, per ogni $n \in \mathbb{N}$ il punto x appartiene alla chiusura debole dell'insieme $\bigcup_{k=n}^{\infty}\{x_k\}$, ed a maggior ragione alla chiusura debole del suo inviluppo convesso. Ma, ovviamente, quest'ultimo insieme è un convesso w-chiuso, quindi anche chiuso nella topologia indotta dalla norma. In definitiva,

$$x \in \overline{\operatorname{co}\left(\bigcup_{k=n}^{\infty}\{x_k\}\right)} \qquad \forall n \in \mathbb{N}.$$

Perciò si può costruire una successione $\{v_n\}_{n\in\mathbb{N}}$ tale che $v_n \to x$ in X e $v_n \in \operatorname{co}\left(\bigcup_{k=n}^{\infty}\{x_k\}\right)$ per ogni $n \in \mathbb{N}$. Ciascun v_n, per definizione, è della forma

$$v_n = \sum_{k=p_n}^{q_n} \lambda_k^{(n)} x_k\,, \quad \lambda_k^{(n)} \in [0,1], \quad \sum_{k=p_n}^{q_n} \lambda_k^{(n)} = 1, \quad n \leq p_n \leq q_n\,.$$

Ciò prova la tesi. $\square$

Concludiamo il paragrafo con un altro risultato di compattezza negli spazi di Banach: questa proprietà è stabile nel passaggio da un insieme qualunque al suo inviluppo convesso.

Proposizione 1.7.19 (teorema di Mazur) *Sia K un sottoinsieme relativamente compatto di uno spazio di Banach X. Allora $\overline{\operatorname{co}(K)}$ è compatto.*

Dimostrazione Basta dimostrare che $\mathrm{co}(K)$ è totalmente limitato. Sia $\varepsilon > 0$: poiché K, essendo relativamente compatto, è totalmente limitato, esistono $z_1, \ldots, z_n \in K$ tali che

$$K \subseteq \bigcup_{i=1}^{n} B\left(z_i\,, \frac{\varepsilon}{2}\right).$$

Poniamo $H = \mathrm{co}\{z_1, \ldots, z_n\}$; allora H è compatto in X, perché è l'immagine del compatto $\{a \in [0,1]^n : \sum_{1=1}^{n} a_i = 1\}$ mediante l'applicazione continua $\psi(a) = \sum_{i=1}^{n} a_i z_i$. Dunque H è totalmente limitato e pertanto esistono $v_1, \ldots v_m \in H$ tali che

$$H \subseteq \bigcup_{j=1}^{m} B\left(v_j\,, \frac{\varepsilon}{2}\right).$$

Sia ora $y \in \mathrm{co}(K)$, ossia $y = \sum_{r=1}^{p} \lambda_r y_r$, con $y_r \in K$, $\sum_{r=1}^{p} \lambda_r = 1$ e $\lambda_r \geq 0$; sia $i_r \in \{1, \ldots m\}$ un indice tale che $y_r \in B(z_{i_r}\,, \frac{\varepsilon}{2})$. Allora, posto $u = \sum_{r=1}^{p} \lambda_r z_{i_r}$, si ha $u \in H$ e

$$\|y - u\|_X = \left\| \sum_{r=1}^{p} \lambda_r (y_r - z_{i_r}) \right\|_X \leq \sum_{r=1}^{p} \lambda_r \|y_r - z_{i_r}\|_X < \frac{\varepsilon}{2}\,,$$

e dunque

$$\mathrm{co}(K) \subseteq \bigcup_{u \in H} B\left(u, \frac{\varepsilon}{2}\right) \subseteq \bigcup_{j=1}^{m} B\left(v_j\,, \varepsilon\right),$$

da cui la tesi, essendo $v_1, \ldots v_m \in H \subseteq \mathrm{co(K)}$. $\square$

Esercizi 1.7

1. Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff. Se $x \in X \setminus \{0\}$, si provi che esiste un funzionale lineare e continuo $F : X \to \mathbb{C}$ tale che $Fx \neq 0$.

2. Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff e sia M un sottospazio proprio di X. Se $x \notin \overline{M}$, si provi che esiste un funzionale lineare e continuo $F : X \to \mathbb{C}$ tale che $\mathrm{Re}\, Fx = 1$ e $F|_M = 0$.

3. Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff, e sia B un sottoinsieme chiuso, convesso e bilanciato di X. Se $x_0 \notin B$, si provi che esiste un funzionale lineare e continuo $F : X \to \mathbb{C}$ tale che $|\mathrm{Re}\, F| \leq 1$ su B e $\mathrm{Re}\, Fx_0 > 1$.

4. Nello spazio $L^2(-1,1)$ si consideri il sottoinsieme $K_\alpha = \{f \in C[-1,1] : f(0) = \alpha\}$, $\alpha \in \mathbb{R}$. Si verifichi che ciascun K_α è convesso e denso in $L^2(-1,1)$, e si provi che per $\alpha \neq \beta$ gli insiemi K_α e K_β sono disgiunti ma non sono separati da alcun iperpiano affine.

5. Verificare che i due insiemi
$$K = \{(x,y) \in \mathbb{R}^2 : \ y \leq 0\}, \qquad H = \{(x,y) \in \mathbb{R}^2 : \ y \geq e^{-x}\}$$
sono convessi chiusi disgiunti di $\mathbb{R}^2$, che non possono essere separati strettamente da alcun funzionale $F \in (\mathbb{R}^2)^*$.

6. Sia X uno spazio vettoriale topologico reale. Si provi che ogni iperpiano affine chiuso in $X \times \mathbb{R}$ è della forma
$$M = \{(x,t) \in X \times \mathbb{R} : \ fx + \alpha t = \beta\},$$
ove f è un funzionale lineare e continuo su X e α, β sono numeri reali. Si mostri poi che se $\alpha \neq 0$, allora M è un opportuno traslato del *grafico* di un funzionale $\Gamma \in X^*$, ossia esiste $y \in X \times \mathbb{R}$ tale che M si può scrivere nella forma
$$M = y + \{(x, \Gamma x) : \ x \in X\}.$$

7. Sia X uno spazio vettoriale topologico complesso e sia $f \in X^* \setminus \{0\}$. Si mostri che se $k \in \mathbb{R}$ l'insieme $P = \{x \in X : \operatorname{Re} fx = k\}$ non è il traslato di alcun sottospazio di X.

8. Si estenda il teorema di Mazur al caso in cui X è uno spazio vettoriale topologico completo e metrizzabile.

9. Sia X uno spazio topologico di Hausdorff. Si provi che un sottoinsieme A di X è compatto se e solo se ogni net $\{x_i\}_{i \in I} \subseteq A$ ha almeno un punto di accumulazione $x \in A$.

1.8 Punti estremi di insiemi convessi

Ogni insieme convesso K in uno spazio vettoriale X è, ovviamente, l'inviluppo convesso di se stesso; ci chiediamo se sia possibile individuare un sottoinsieme $H \subseteq K$ "minimale" del quale K sia l'inviluppo convesso. Per rispondere a questa domanda è importante la nozione di punto estremo.

Definizione 1.8.1 *Sia X uno spazio vettoriale e sia K un sottoinsieme convesso e non vuoto di X. Un punto $x_0 \in K$ si dice* punto estremo *di K se l'insieme $K \setminus \{x_0\}$ è convesso.*

In altre parole, un punto $x_0 \in K$ è estremo per K se non vi è alcuna coppia di punti distinti $x, y \in K$ tali che si abbia $\lambda x + (1-\lambda)y = x_0$ per qualche $\lambda \in]0,1[$, cosicché togliere x_0 non distrugge la convessità di K.
Uno dei più profondi risultati relativi agli insiemi convessi e compatti è il seguente

Teorema 1.8.2 (di Krein-Milman) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff e sia $K \subseteq X$ un insieme compatto, convesso e non vuoto. Allora, detto K_e l'insieme dei punti estremi di K, si ha:*

(i) *K_e è non vuoto;*

(ii) *$K = \overline{\mathrm{co}(K_e)}$.*

Dimostrazione (i) Indichiamo con $\mathcal{K}$ la classe di tutti i sottoinsiemi chiusi e non vuoti $A \subseteq K$ tali che

$$x, y \in K, \quad \lambda \in]0,1[, \quad \lambda x + (1-\lambda)y \in A \qquad \Longrightarrow \qquad x, y \in A.$$

Poiché $K \in \mathcal{K}$, la classe $\mathcal{K}$ è non vuota (ed è costituita da insiemi compatti, essendo K compatto). Consideriamo su di essa l'ordinamento parziale dato dall'inclusione "rovesciata":

$$A_1 \leq A_2 \quad \Longleftrightarrow \quad A_1 \supseteq A_2\,, \qquad A_1\,, A_2 \in \mathcal{K}.$$

Osserviamo che se $\mathcal{A} \subseteq \mathcal{K}$ è un insieme totalmente ordinato, allora l'insieme $\bigcap_{A \in \mathcal{A}} A$ è non vuoto, essendo intersezione di una famiglia decrescente di compatti; inoltre applicando la definizione di $\mathcal{K}$ si vede subito che $\bigcap_{A \in \mathcal{A}} A$ è un elemento di $\mathcal{K}$, ed è evidentemente un maggiorante per $\mathcal{A}$. Per il lemma di Zorn, deduciamo che esiste almeno un elemento massimale $A^* \in \mathcal{K}$.
Affermiamo che ogni elemento massimale $A^* \in \mathcal{K}$ contiene esattamente un punto. Infatti, supposto per assurdo $x, y \in A^*$ e $x \neq y$, per il teorema 1.7.5 possiamo trovare un funzionale lineare e continuo $F \in X^*$ tale che, posto $G = \mathrm{Re}\, F$, sia $Gx < Gy$. Poniamo

$$A^\circ = \{u \in A^* : \; Gu = \inf_{A^*} G\};$$

allora $A^\circ \neq \emptyset$ perché il funzionale continuo G ha minimo sul compatto A^*. D'altra parte, $A^\circ \in \mathcal{K}$: infatti, se $u, v \in K$ e $\lambda u + (1-\lambda)v \in A^\circ$ per qualche $\lambda \in]0,1[$, si ha intanto $u, v \in A^*$ (in quanto $A^* \in \mathcal{K}$); di conseguenza, essendo

$$\inf_{A^*} G = \lambda \inf_{A^*} G + (1-\lambda) \inf_{A^*} G \leq \lambda Gu + (1-\lambda) Gv = G(\lambda u + (1-\lambda)v) = \inf_{A^*} G,$$

si deduce che $Gu = Gv = \inf_{A^*} G$, ossia $u, v \in A^\circ$. Ciò mostra appunto che $A^\circ \in \mathcal{K}$. Ma, dato che $y \in A^* \setminus A^\circ$, otteniamo che A° è contenuto propriamente in A^*, ossia $A^* \leq A^\circ$ e $A^* \neq A^\circ$, il che contraddice la massimalità di A^*. Si conclude che A^* ha al più un elemento; ma siccome $A^* \in \mathcal{K}$, deve essere $A^* \neq \emptyset$ e pertanto A^* ha esattamente un elemento. Tale elemento, per definizione di $\mathcal{K}$, è un punto estremo di K. Ciò prova (i).

(ii) Essendo K un convesso chiuso contenente K_e, è chiaro che $K \supseteq \overline{\mathrm{co}(K_e)}$; proviamo l'altra inclusione. Supponiamo per assurdo che esista $z \in K \setminus \overline{\mathrm{co}(K_e)}$. Per il teorema 1.7.5, esiste un funzionale lineare e continuo $F \in X^*$ tale che, posto $G = \mathrm{Re}\, F$, risulta

$$Gz < \inf\{Gu : \ u \in \overline{\mathrm{co}(K_e)}\}.$$

Posto

$$K^\circ = \left\{ x \in K : \ Gx = \inf_K G \right\},$$

chiaramente K° è un chiuso non vuoto disgiunto da K_e ed inoltre si ha, ragionando come in precedenza,

$$x, y \in K, \quad \lambda \in]0,1[, \quad \lambda x + (1-\lambda)y \in K^\circ \qquad \Longrightarrow \qquad x, y \in K^\circ,$$

ossia K° appartiene alla classe $\mathcal{K}$ introdotta all'inizio della dimostrazione. Inoltre K°, essendo compatto e convesso, ha punti estremi in virtù della parte **(i)** già provata.

Mostriamo adesso che ogni punto estremo di K° è punto estremo di K. Sia $x_0 \in (K^\circ)_e$: in particolare, $x_0 \in K^\circ$ e quindi $Gx_0 = \inf_K G$. Se fosse $x_0 \notin K_e$ esisterebbero $u, v \in K$ e $\lambda \in]0,1[$ tali che $\lambda u + (1-\lambda)v = x_0$. Ma allora, con l'ormai consueto ragionamento, deduciamo $\inf_K G = Gx_0 = Gu = Gv$, cioè $u, v \in K^\circ$; questo contraddice il fatto che x_0 è punto estremo di K°. Si conclude quindi che $x_0 \in K_e$.

Abbiamo così ottenuto $(K^\circ)_e \subseteq K^\circ \cap K_e = \emptyset$, e ciò è impossibile. Pertanto $K \subseteq \overline{\mathrm{co}(K_e)}$ e la tesi è dimostrata. $\square$

Esercizi 1.8

1. Sia K un convesso non vuoto in uno spazio vettoriale topologico localmente convesso di Hausdorff. Si provi che $K_e \subseteq \partial K$; si fornisca un esempio in cui vale l'inclusione stretta ed un esempio in cui vale l'uguaglianza.

2. Se K è un poliedro convesso in $\mathbb{R}^N$, si provi che K_e coincide con l'insieme dei vertici di K.

3. Si consideri l'insieme $K \subset \mathbb{R}^3$ così definito:

$$K = \mathrm{co}\{(1,0,1),\ (1,0,-1),\ (\cos\vartheta, \sin\vartheta, 0) :\ \vartheta \in [0, 2\pi]\}.$$

   Si determini K_e e si mostri che K è compatto mentre K_e non lo è.

4. Per ogni $p \in [1, \infty[$ sia

$$B_p = \{f \in L^p(0,1) :\ \|f\|_p \leq 1\}.$$

   Si provi che $(B_1)_e = \emptyset$, mentre

$$(B_p)_e = \{f \in L^p(0,1) :\ \|f\|_p = 1\} \qquad \forall p \in]1, \infty[.$$

5. Posto

$$B = \{f \in C([0,1], \mathbb{C}) :\ \|f\|_\infty \leq 1\}, \quad B' = B \cap C([0,1], \mathbb{R}),$$

   si provi che

$$B_e = \{f \in C([0,1], \mathbb{C}) :\ |f(t)| = 1\ \forall t \in [0,1]\}, \qquad B'_e = \{\pm 1\}.$$

6. Si determini B_e , ove $B = \{f \in L^\infty(0,1) :\ \|f\|_\infty \leq 1\}$.

7. Sia S un sottoinsieme non vuoto di $\mathbb{R}^N$. Si provi che ogni $x \in \mathrm{co}(S)$ è combinazione convessa di al più $N+1$ punti di S.
   [**Traccia:** Se $x \in \mathrm{co}(S)$, sarà $x = \sum_{i=1}^m \lambda_i x_i$, con $\lambda_i \in [0,1]$, $\sum_{i=1}^m \lambda_i = 1$, $x_i \in S$ e $m \in \mathbb{N}^+$. Se $m \leq N+1$ non c'è niente da dimostrare; altrimenti sarà $m > N+1$ e potremo assumere $\lambda_i > 0$ per ogni $i = 1, \ldots, m$. Si verifichi che esistono $\alpha_1, \ldots, \alpha_{m-1}$ non tutti nulli, tali che $\sum_{i=1}^{m-1} \alpha_i (x_i - x_m) = 0$, ossia $\sum_{i=1}^m \alpha_i x_i = 0$, ove $\alpha_m = -\sum_{i=1}^{m-1} \alpha_i$.

Si determini ora $\mu \in \mathbb{R}$ tale che: (a) $\lambda_i - \mu\alpha_i \geq 0$ per $i = 1, \dots, m$; (b) esiste $j \in \{1, \dots, m\}$ per cui $\lambda_j - \mu\alpha_j = 0$. Se ne deduca che x è combinazione convessa di $x_1, \dots, x_m$ con coefficienti $\lambda_i - \mu\alpha_i$, $i = 1, \dots m$, uno dei quali è nullo: dunque x è combinazione convessa di m elementi di S. Iterando questo argomento un numero finito di volte si ricavi la tesi.]

8. Provare che se K è un convesso compatto e non vuoto di $\mathbb{R}^N$, allora ogni $x \in K$ è combinazione convessa di al più $N + 1$ punti estremi di K.

Capitolo 2

Funzioni convesse

2.1 Generalità

Le funzioni convesse hanno un ruolo fondamentale nei problemi di minimo e nel calcolo delle variazioni. Questo capitolo è dedicato alla descrizione delle loro principali proprietà.

Definizione 2.1.1 *Sia X uno spazio vettoriale. Una funzione $f : X \to \overline{\mathbb{R}} = [-\infty, +\infty]$ si dice* convessa *se risulta*

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \qquad \forall t \in [0,1]$$

per ogni coppia di punti $x, y \in X$ per i quali il secondo membro sia ben definito (ossia non risulti $f(x) = -f(y) = \pm\infty$). Se, limitatamente agli x, y per i quali ha senso, la disuguaglianza è stretta per ogni $t \in]0,1[$ e $x \neq y$, la funzione si dice strettamente convessa. *La funzione si dice* concava, *o* strettamente concava, *se $-f$ è convessa o strettamente convessa.*

Abbiamo ammesso tra i valori di una funzione convessa anche $\pm\infty$. Ci sono almeno due buone ragioni per ammettere il valore $+\infty$. La prima è che se $U \subset X$ è un insieme convesso e se $f : U \to \mathbb{R}$ è una funzione convessa (cioè vale la disuguaglianza di convessità per ogni $t \in [0,1]$ e $x, y \in U$), allora l'estensione

$$\overline{f}(x) = \begin{cases} f(x) & \text{se } x \in U \\ +\infty & \text{se } x \notin U \end{cases}$$

risulta convessa: quindi possiamo limitarci a considerare funzioni convesse definite su tutto X. La seconda ragione è la seguente: introducendo la funzione *indicatrice* di un insieme $E \subseteq X$,

$$I_E(x) = \begin{cases} 0 & \text{se } x \in E \\ +\infty & \text{se } x \notin E, \end{cases}$$

si ha che E è convesso in X se e solo se I_E è una funzione convessa su X: dunque lo studio degli insiemi convessi può attuarsi tramite lo studio delle funzioni convesse.
Invece le funzioni convesse che assumono il valore $-\infty$ sono alquanto patologiche: se f è convessa e se $f(x_0) = -\infty$, allora su ogni semiretta $\{x_0+tv,\ t > 0\}$ il comportamento di f è il seguente: o essa vale costantemente $+\infty$, o vale costantemente $-\infty$, oppure per un opportuno $t_0 > 0$ si avrà $f(x_0+tv) = -\infty$ per $0 \leq t < t_0$ e $f(x_0 + tv) = +\infty$ per $t > t_0$, mentre $f(x_0 + t_0 v)$ sarà un qualunque valore in $\overline{\mathbb{R}}$. Questa affermazione segue facilmente utilizzando la disuguaglianza di convessità.
Per sgombrare il campo da certe patologie, sono utili le definizioni che seguono.

Definizione 2.1.2 *Sia X uno spazio vettoriale. Una funzione $f : X \to \overline{\mathbb{R}}$ si dice* propria *se $f(x) > -\infty$ per ogni $x \in X$ ed esiste $\overline{x} \in X$ per cui $f(\overline{x}) < +\infty$.*

Definizione 2.1.3 *Sia X uno spazio vettoriale e sia $f : X \to \overline{\mathbb{R}}$ una funzione. Il* dominio *di f è l'insieme*

$$D(f) = \{x \in X : \ f(x) < +\infty\}.$$

Dunque se f è convessa e propria si ha $D(f) = \{x \in X : f(x) \in \mathbb{R}\}$. Inoltre, se f è convessa, $D(f)$ è un sottoinsieme convesso di X.

Esempio 2.1.4 In uno spazio vettoriale X sono funzioni convesse e proprie: i funzionali lineari, i funzionali di Minkowski di insiemi convessi e radiali, le seminorme.

Osservazione 2.1.5 Se $f : X \to \overline{\mathbb{R}}$ è convessa, allora gli *insiemi di sottolivello*

$$E_\lambda = \{x \in X : \ f(x) \leq \lambda\}, \qquad \lambda \in \mathbb{R},$$

sono convessi. Il viceversa è falso, come mostra l'esempio della funzione x^3 in $\mathbb{R}$. Se $f : X \to \overline{\mathbb{R}}$ è una funzione i cui insiemi di sottolivello sono convessi, essa si dice *quasi convessa*; f si dirà *quasi concava* se $-f$ è quasi convessa.

Definizione 2.1.6 *Sia X uno spazio vettoriale, sia $f : X \to \overline{\mathbb{R}}$. L'*epigrafico *di f è il sottoinsieme di $X \times \mathbb{R}$ così definito:*

$$\mathrm{epi}(f) = \{(x,t) \in X \times \mathbb{R} : \ f(x) \leq t\} = \bigcup_{\lambda \in \mathbb{R}} (E_\lambda \times [\lambda, +\infty[\,).$$

La stretta parentela fra insiemi convessi e funzioni convesse è evidenziata dalla seguente

Proposizione 2.1.7 *Sia X uno spazio vettoriale. Una funzione $f : X \to \overline{\mathbb{R}}$ è convessa se e solo se il suo epigrafico è convesso in $X \times \mathbb{R}$.*

Dimostrazione ($\Longrightarrow$) Siano $(x,t), (y,s) \in \mathrm{epi}(f)$: allora in particolare né $f(x)$, né $f(y)$ valgono $+\infty$ e quindi si può sicuramente scrivere la disuguaglianza di convessità. Pertanto, per ogni $\lambda \in [0,1]$ si ha

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \leq (1-\lambda)t + \lambda s,$$

ossia

$$(1-\lambda)(x,t) + \lambda(y,s) \in \mathrm{epi}(f) \qquad \forall \lambda \in [0,1],$$

il che mostra che $\mathrm{epi}(f)$ è convesso.

($\Longleftarrow$) Se almeno uno fra $f(x)$ e $f(y)$ vale $+\infty$, la disuguaglianza di convessità, se ha senso, è ovvia. Se $f(x)$ e $f(y)$ sono entrambi reali, allora i punti $(x, f(x))$ e $(y, f(y))$ appartengono a $\mathrm{epi}(f)$, e quindi si ha, per ipotesi,

$$((1-\lambda)x + \lambda y, (1-\lambda)f(x) + \lambda f(y)) \in \mathrm{epi}(f) \qquad \forall \lambda \in [0,1],$$

cioè

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \qquad \forall \lambda \in [0,1].$$

Infine se almeno uno fra $f(x)$ e $f(y)$ vale $-\infty$, ad esempio $f(x) = -\infty$ (ma nessuno dei due vale $+\infty$), per ogni $u \in \mathbb{R}$ e per ogni $v \in [f(y), +\infty[$ si ha $(x,u), (y,v) \in \mathrm{epi}(f)$ e dunque come prima otteniamo

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)u + \lambda v \qquad \forall u \in \mathbb{R}, \quad \forall v \in [f(y), +\infty[,$$

da cui per $u \to -\infty$ ricaviamo $f((1-\lambda)x + \lambda y) = -\infty$, ossia f verifica la relazione di convessità anche in questo caso. $\square$

Esercizi 2.1

1. Sia X uno spazio vettoriale e sia $f : X \to \mathbb{R}$ convessa. Si provi che se $\varphi : \mathbb{R} \to \mathbb{R}$ è crescente e convessa, allora $\varphi \circ f$ è convessa su X.

2. Si provi che l'estremo superiore di una famiglia arbitraria di funzioni convesse è una funzione convessa.

3. Sia X uno spazio vettoriale e sia $f : X \to \mathbb{R}$ una funzione qualunque. Si provi che f è convessa su X se e solo se per ogni $x, v \in X$ il rapporto incrementale
$$t \mapsto \frac{f(x+tv) - f(x)}{t}, \qquad t \in \mathbb{R} \setminus \{0\},$$
è una funzione crescente.

4. Sia X uno spazio vettoriale. Si provi che se $f : X \to \overline{\mathbb{R}}$ è convessa, allora l'insieme $\{x \in X : f(x) = -\infty\}$ è convesso mentre in generale l'insieme $\{x \in X : f(x) = +\infty\}$ non lo è.

5. Sia X uno spazio di Hilbert. Si provi che $x \mapsto \|x\|^2$ è una funzione strettamente convessa, mentre $x \mapsto \|x\|$ è convessa ma non strettamente convessa.

6. Sia X uno spazio vettoriale topologico. Se $f : X \to \mathbb{R}$ è continua, si provi che f è convessa se e solo se
$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y) \qquad \forall x, y \in X.$$

7. Sia X uno spazio vettoriale. Si provi che l'insieme delle funzioni convesse è un convesso dello spazio vettoriale $Y = \{f : X \to \overline{\mathbb{R}}\}$, ed è chiuso rispetto alla convergenza puntuale.

2.2 Semicontinuità

Una fondamentale proprietà funzionale, di grande importanza nei problemi di minimo, è la semicontinuità.

Definizione 2.2.1 *Sia X uno spazio topologico e sia $f : X \to \overline{\mathbb{R}}$. Diciamo che f è* semicontinua inferiormente *nel punto $x_0 \in X$ se si ha*

$$f(x_0) \leq \liminf_{x \to x_0} f(x).$$

Diciamo che f è semicontinua superiormente *nel punto $x_0 \in X$ se si ha*

$$f(x_0) \geq \limsup_{x \to x_0} f(x).$$

La funzione f si dice semicontinua inferiormente, o superiormente, in X se lo è in ogni punto di X.

Ricordiamo che, detta $\mathcal{U}(x_0)$ la famiglia degli intorni di x_0, si ha per definizione

$$\liminf_{x \to x_0} f(x) = \sup_{V \in \mathcal{U}(x_0)} \inf_{y \in V \setminus \{x_0\}} f(y), \qquad \limsup_{x \to x_0} f(x) = \inf_{V \in \mathcal{U}(x_0)} \sup_{y \in V \setminus \{x_0\}} f(y).$$

È facile verificare che f è semicontinua inferiormente in x_0 se e solo se $-f$ è semicontinua superiormente in x_0; inoltre f è continua in x_0 se e solo se f è semicontinua sia inferiormente che superiormente in x_0. Vale poi la seguente caratterizzazione:

Proposizione 2.2.2 *Sia X uno spazio topologico e sia $f : X \to \overline{\mathbb{R}}$. I seguenti fatti sono equivalenti:*

(i) *f è semicontinua inferiormente in X;*

(ii) *gli insiemi di sottolivello $E_\lambda = \{x \in X : f(x) \leq \lambda\}$ sono chiusi per ogni $\lambda \in \mathbb{R}$;*

(iii) $\mathrm{epi}(f)$ *è chiuso in $X \times \mathbb{R}$.*

Dimostrazione (i) $\Longrightarrow$ (ii) Proviamo che E_λ^c è aperto. Sia $x_0 \notin E_\lambda$, ossia $f(x_0) > \lambda$; scelto $\varepsilon \in]0, f(x_0) - \lambda[$, per ipotesi esiste $V \in \mathcal{U}(x_0)$ tale che $f(x) \geq f(x_0) - \varepsilon > \lambda$ per ogni $x \in V \setminus \{x_0\}$. Ciò è ovviamente vero anche per $x = x_0$, il che mostra che $V \subseteq E_\lambda^c$; pertanto E_λ^c è aperto.
(ii) $\Longrightarrow$ (iii) Proviamo che $\mathrm{epi}(f)^c$ è aperto. Sia $(x_0, \alpha) \notin \mathrm{epi}(f)$, ossia $f(x_0) > \alpha$; scelto $\varepsilon \in]0, f(x_0) - \alpha[$, poiché x_0 appartiene all'aperto $E_{\alpha+\varepsilon}^c$ esiste $V \in \mathcal{U}(x_0)$ tale che $f(x) > \alpha + \varepsilon$ per ogni $x \in V$. Dunque, per

ogni $(x, \beta) \in V \times]\alpha - \varepsilon, \alpha + \varepsilon[$ si ha a maggior ragione $f(x) > \beta$, ossia $V \times]\alpha - \varepsilon, \alpha + \varepsilon[\subseteq \mathrm{epi}(f)^c$. Ciò mostra che $\mathrm{epi}(f)^c$ è aperto.

(iii) $\Longrightarrow$ (i) Proviamo che f è semicontinua inferiormente in x_0. Sia $\alpha < f(x_0)$: allora (x_0, α) appartiene all'aperto $\mathrm{epi}(f)^c$, cosicché esistono $V \in \mathcal{U}(x_0)$ ed $\varepsilon > 0$ tali che $V \times]\alpha - \varepsilon, \alpha + \varepsilon[\subseteq \mathrm{epi}(f)^c$; ciò significa $f(x) > \beta$ per ogni $x \in V$ e per ogni $\beta \in]\alpha - \varepsilon, \alpha + \varepsilon[$. Ne segue $\inf_V f \geq \alpha + \varepsilon > \alpha$, ed a maggior ragione

$$\sup_{V \in \mathcal{U}(x_0)} \inf_{V \setminus \{x_0\}} f > \alpha.$$

Per l'arbitrarietà di α, otteniamo $\liminf_{x \to x_0} f(x) \geq f(x_0)$. $\square$

Corollario 2.2.3 *L'estremo superiore di un'arbitraria famiglia di funzioni semicontinue inferiormente è una funzione semicontinua inferiormente.*

Dimostrazione Sia $\{f_i\}_{i \in I}$ una famiglia di funzioni semicontinue inferiormente nello spazio topologico X. Per ogni $x_0 \in X$ e per ogni $i \in I$ si ha

$$f_i(x_0) \leq \sup_{V \in \mathcal{U}(x_0)} \inf_{x \in V \setminus \{x_0\}} f_i(x);$$

quindi

$$f_i(x_0) \leq \sup_{V \in \mathcal{U}(x_0)} \inf_{x \in V \setminus \{x_0\}} \sup_{i \in I} f_i(x),$$

ed infine

$$\sup_{i \in I} f_i(x_0) \leq \sup_{V \in \mathcal{U}(x_0)} \inf_{x \in V \setminus \{x_0\}} \sup_{i \in I} f_i(x),$$

che è la tesi. $\square$

Per le funzioni semicontinue vale metà del classico teorema di Weierstrass.

Proposizione 2.2.4 *Sia X uno spazio topologico di Hausdorff e sia $f : X \to \overline{\mathbb{R}}$ una funzione semicontinua inferiormente in X. Se $K \subseteq X$ è un compatto, oppure sequenzialmente compatto, allora f ha minimo su K.*

Naturalmente, se vogliamo che il minimo di f su K sia finito, dovremo richiedere che in K si abbia $f \neq +\infty$ e $f > -\infty$.

Dimostrazione Se $f \equiv +\infty$ su K, allora $\inf_K f = \min_K f = +\infty$. Altrimenti, sarà $\lambda = \inf_K f < +\infty$; esiste allora una successione $\{x_n\} \subseteq K$ tale che $f(x_n) \to \lambda$ per $n \to \infty$. Supponendo K sequenzialmente compatto, esiste una sottosuccessione $\{x_{n_k}\} \subseteq \{x_n\}$ che converge ad un punto $x \in K$;

per semicontinuità, si conclude che $\lambda \leq f(x) \leq \liminf_{k\to\infty} f(x_{n_k}) = \lambda$, ossia $f(x) = \lambda = \min_K f$.
Se invece K è compatto, sia $\varepsilon > 0$: esiste $\nu \in \mathbb{N}$ tale che $\lambda \leq f(x_n) < \lambda+\varepsilon$ per ogni $n \geq \nu$. Utilizzando il lemma 1.7.6 otteniamo che la successione $\{x_n\}_{n\geq\nu}$ ha un punto d'accumulazione $x \in K$; ciò significa che per ogni $U \in \mathcal{U}(x)$ esiste un indice $n_U \geq \nu$ tale che $x_{n_U} \in U$. Ne segue che il net $\{x_{n_U}\}_{U\in\mathcal{U}(x)}$ converge a x. Scelto $U \in \mathcal{U}(x)$ in modo che $\inf_{y\in U} f(y) > \liminf_{y\to x} f(y) - \varepsilon$, per semicontinuità si conclude che

$$\lambda \leq f(x) \leq \inf_{y\in U} f(y) + \varepsilon \leq f(x_{n_U}) = \lambda + 2\varepsilon,$$

da cui $f(x) = \lambda = \min_K f$. $\square$

Esercizi 2.2

1. Sia X uno spazio topologico. Si provi che un sottoinsieme E di X è chiuso se e solo se la sua funzione indicatrice I_E è semicontinua inferiormente su X.

2. Esibire un esempio di funzione $f : \mathbb{R} \to \mathbb{R}$ che non è semicontinua inferiormente né superiormente in alcun punto.

2.3 Funzioni convesse semicontinue inferiormente

Sia X uno spazio vettoriale topologico: le funzioni convesse da X in $\overline{\mathbb{R}}$ che sono semicontinue inferiormente hanno un quadro più ricco di proprietà e rivestono grande importanza nelle applicazioni.
Tutti gli enunciati che daremo hanno una versione simmetrica relativa alle funzioni concave e semicontinue superiormente, di cui per brevità non parleremo; osserviamo tuttavia che sono proprio le funzioni concave e semicontinue superiormente quelle maggiormente usate nelle applicazioni all'economia.
Prima di tutto, osserviamo che, fra le funzioni convesse, quelle semicontinue inferiormente presentano meno patologie. Infatti:

Proposizione 2.3.1 *Sia X uno spazio vettoriale topologico e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa e semicontinua inferiormente. Se esiste $x_0 \in X$ tale che $f(x_0) = -\infty$, allora $|f(x)| = +\infty$ per ogni $x \in X$.*

Dimostrazione Se esistesse $y_0 \in X$ tale che $f(y_0) \in \mathbb{R}$, allora per convessità avremmo

$$f(\lambda x_0 + (1-\lambda)y_0) \leq \lambda f(x_0) + (1-\lambda)f(y_0) = -\infty,$$

e quando $\lambda \to 0$ otterremmo, per semicontinuità,

$$f(y_0) \leq \liminf_{\lambda \to 0} f(\lambda x_0 + (1-\lambda)y_0) = \liminf_{\lambda \to 0}(-\infty) = -\infty,$$

il che è assurdo. $\square$

In particolare, dunque, se una funzione convessa e semicontinua inferiormente assume un valore reale, allora non assume il valore $-\infty$.
Le funzioni convesse e semicontinue inferiormente possiedono la seguente proprietà, di fondamentale importanza per le applicazioni.

Proposizione 2.3.2 *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff. Se $f : X \to \overline{\mathbb{R}}$ è convessa e semicontinua inferiormente, allora f è semicontinua inferiormente per la topologia debole di X.*

Dimostrazione Per ipotesi, il convesso epi(f) è chiuso in $X \times \mathbb{R}$; quindi, per il corollario 1.7.17, esso è anche w-chiuso in $X \times \mathbb{R}$. I funzionali lineari e continui su $X \times \mathbb{R}$ sono tutti e soli quelli della forma

$$(x,t) \mapsto \varphi x + \lambda t, \qquad (x,t) \in X \times \mathbb{R},$$

con $\varphi \in X^*$ e $\lambda \in \mathbb{R}^* = \mathbb{R}$ (esercizio 2.3.8); quindi lo spazio topologico $(X \times \mathbb{R}, w)$ coincide con lo spazio prodotto $(X, w) \times \mathbb{R}$. Dunque epi(f) è chiuso in $(X, w) \times \mathbb{R}$ e pertanto f è semicontinua inferiormente in X relativamente a w. $\square$

Dalla proposizione precedente discendono alcuni criteri per l'esistenza, su opportuni insiemi, del minimo di una funzione convessa e semicontinua inferiormente.

Teorema 2.3.3 *Sia X uno spazio di Banach riflessivo e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa e semicontinua inferiormente. Se K è un sottoinsieme limitato, debolmente chiuso e non vuoto di X, allora f ha minimo su K. Se f è propria su K, tale minimo è finito. Se infine K è convesso e f è strettamente convessa, il punto di minimo è unico.*

Dimostrazione Se $f \equiv +\infty$ su K, allora $\min_K f = \inf_K f = +\infty$. Altrimenti, sarà $\lambda = \inf_K f < +\infty$; scegliamo una successione $\{x_n\} \subseteq K$ tale che $f(x_n) \to \lambda$ per $n \to \infty$. Poiché K è debolmente chiuso, esso è chiuso nonché limitato: essendo X riflessivo, per il teorema di Eberlein-Smulyan (più precisamente per il corollario 1.6.5) K è debolmente compatto; per il teorema 1.6.10, K è anche debolmente sequenzialmente compatto. Quindi esiste $\{x_{n_k}\} \subseteq \{x_n\}$ tale che $x_{n_k} \rightharpoonup x \in K$ per $k \to \infty$. Poiché f è w-semicontinua inferiormente (proposizione 2.3.2), otteniamo $\lambda \le f(x) \le \liminf_{k\to\infty} f(x_{n_k}) = \lambda$.
Se f è anche propria in K, la relazione $f(x) = \lambda$ mostra che $\lambda > -\infty$ e quindi il minimo di f è finito.
Infine se K è convesso e f è strettamente convessa, siano x_1 e x_2 due distinti punti di minimo per f in K: allora essendo $f(x_1) = f(x_2) = \lambda$, la stretta convessità implica

$$\lambda \le f\left(\frac{x_1 + x_2}{2}\right) < \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2) = \lambda,$$

il che è assurdo. $\square$

Corollario 2.3.4 *Sia X uno spazio di Banach riflessivo e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa e semicontinua inferiormente. Se risulta*

$$\lim_{\|x\|_X\to\infty} f(x) = +\infty,$$

allora f ha minimo su ogni sottoinsieme debolmente chiuso non vuoto $H \subseteq X$. Se f è anche propria in H, tale minimo è finito. Se inoltre H è convesso e f è strettamente convessa, il punto di minimo è unico.

Dimostrazione Se $f \equiv +\infty$, allora $\min_H f = \inf_H f = +\infty$. Altrimenti, fissato un numero reale $M > \inf_H f$, per ipotesi si avrà $f(x) > M$ per $\|x\|_X > r$, con r opportuno; dunque $\inf_H f = \inf_{H\cap\overline{B(0,r)}} f$, e quindi ci siamo ridotti al caso del teorema precedente. Ne segue la tesi. $\square$

Osservazione 2.3.5 Il risultato del teorema 2.3.3 e quindi del corollario 2.3.4 si estendono al caso delle funzioni quasi convesse, definite nell'osservazione 2.1.5: si veda in proposito l'esercizio 2.3.5.

Stabiliamo adesso un "lemma di separazione" fra l'epigrafico di una funzione convessa ed un convesso disgiunto; le due forme di questo enunciato, che seguono dalle due forme geometriche del teorema di Hahn-Banach, verranno usate ripetutamente nel seguito.

Lemma 2.3.6 **(i)** *Sia X uno spazio vettoriale topologico, sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa e sia $B \subset X \times \mathbb{R}$ un insieme convesso non vuoto. Se $\overset{\circ}{\text{epi}}(f) \neq \emptyset$, $B \cap \overset{\circ}{\text{epi}}(f) = \emptyset$ ed esistono $\xi \in \overset{\circ}{D}(f)$ e $s \in \mathbb{R}$ tali che $(\xi, s) \in B$, allora si possono trovare $G \in X^*$ e $\lambda > 0$ tali che, posto $\varphi = \operatorname{Re} G$, risulti*

$$\varphi y + \lambda s \leq \varphi x + \lambda t \qquad \forall (y, s) \in B, \quad \forall (x, t) \in \text{epi}(f).$$

(ii) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa, semicontinua inferiormente e tale che $f(x) > -\infty$ per ogni $x \in X$. Se $(y, s) \notin \text{epi}(f)$, allora si possono trovare $G \in X^*$ e $\lambda > 0$ tali che, posto $\varphi = \operatorname{Re} G$, risulti*

$$\varphi y + \lambda s < \inf\{\varphi x + \lambda t : (x, t) \in \text{epi}(f)\}.$$

Dimostrazione (i) Poiché $\overset{\circ}{\text{epi}}(f)$ e B sono convessi non vuoti disgiunti ed il primo è aperto, per il teorema 1.7.3 esiste un funzionale lineare e continuo su $X \times \mathbb{R}$, non nullo, che li separa: in altre parole, in virtù del corollario 1.7.13 e dell'esercizio 1.7.6, esistono $\lambda \in \mathbb{R}$ e $G : X \to \mathbb{C}$ lineare e continuo, non entrambi nulli e tali che, posto $\varphi = \operatorname{Re} G$,

$$\sup\{\varphi y + \lambda s : \ (y, s) \in B\} \leq \inf\{\varphi x + \lambda t : \ (x, t) \in \overset{\circ}{\text{epi}}(f)\}.$$

Questa disuguaglianza, per continuità, diventa

$$\sup\{\varphi y + \lambda s : \ (y, s) \in B\} \leq \inf\{\varphi x + \lambda t : \ (x, t) \in \overline{\overset{\circ}{\text{epi}}(f)}\},$$

e poiché, per l'esercizio 1.3.8, risulta

$$\overline{\overset{\circ}{\text{epi}}(f)} = \overline{\text{epi}(f)},$$

deduciamo in particolare

$$\varphi y + \lambda s \leq \varphi x + \lambda t \qquad \forall (y, s) \in B, \quad \forall (x, t) \in \text{epi}(f).$$

Dimostriamo che si ha necessariamente $\lambda > 0$. È chiaro che non può essere $\lambda < 0$ perché in tal caso per $t \to +\infty$ dedurremmo che il primo membro della disuguaglianza vale $-\infty$. D'altronde, se fosse $\lambda = 0$, la φ sarebbe non nulla;

prendendo $(\xi, s) \in B$ con $\xi \in \overset{\circ}{D(f)}$ e scegliendo $y = \xi$, la disuguaglianza si ridurrebbe a

$$\varphi\xi \leq \varphi x \qquad \forall x \in D(f).$$

Dato che ξ è un punto interno a $D(f)$, si potrebbe prendere un intorno bilanciato V di 0 tale che $\xi + V \subseteq D(f)$: quindi scegliendo $x = \xi \pm v$, $v \in V$, ricaveremmo

$$0 \leq \pm\varphi v \qquad \forall v \in V,$$

il che, per linearità, implicherebbe $\varphi \equiv 0$: ma ciò è assurdo, e pertanto deve essere $\lambda > 0$.

(ii) Se $f \equiv +\infty$, allora $\mathrm{epi}(f)$ è vuoto e la tesi è ovvia: supponiamo perciò che esista $z \in X$ tale che $f(z) \in \mathbb{R}$. Gli insiemi $\mathrm{epi}(f)$ e $\{(y, s)\}$ sono convessi non vuoti e disgiunti in $X \times \mathbb{R}$; inoltre il secondo è compatto ed il primo è chiuso in quanto f è semicontinua inferiormente. Per il teorema 1.7.5, esiste un funzionale lineare e continuo su $X \times \mathbb{R}$, non nullo, che li separa strettamente: dunque esistono $\lambda \in \mathbb{R}$ e $G : X \to \mathbb{C}$ lineare e continuo, non entrambi nulli e tali che, posto $\varphi = \mathrm{Re}\, G$,

$$\varphi y + \lambda s < \beta \leq \varphi x + \lambda t \qquad \forall (x, t) \in \mathrm{epi}(f).$$

Se fosse $\lambda < 0$, l'ultimo membro della disuguaglianza sarebbe $-\infty$ e ciò è assurdo; dunque si ha, intanto, $\lambda \geq 0$. Supponiamo dapprima che sia $f(y) \in \mathbb{R}$: allora $(y, f(y)) \in \mathrm{epi}(f)$ e quindi, scegliendo $x = y$ e $t = f(y)$ ricaviamo $\lambda s < \lambda f(y)$, da cui necessariamente $\lambda > 0$. Se invece $f(y) = +\infty$, può essere ancora $\lambda > 0$, e in tal caso abbiamo concluso, oppure $\lambda = 0$: mostreremo che in questo caso si possono comunque modificare λ e G prendendo λ positivo. Supponiamo dunque $\lambda = 0$, cosicché

$$\varphi y < \beta \leq \varphi x \qquad \forall x \in D(f).$$

Siccome $f(z) \in \mathbb{R}$, esiste $u \in \mathbb{R}$ tale che $(z, u) \notin \mathrm{epi}(f)$, e come abbiamo appena visto, in questa situazione esistono $\widetilde{G} \in X^*$ ed $\varepsilon > 0$ tali che, posto $\psi = \mathrm{Re}\, \widetilde{G}$,

$$\psi z + \varepsilon u < \gamma = \inf\{\psi x + \varepsilon t : (x, t) \in \mathrm{epi}(f)\}.$$

Scegliamo $c > 0$ abbastanza grande da far sì che

$$\varepsilon s < \varepsilon c(\beta - \varphi y) + \gamma - \psi y.$$

Allora risulta

$$(\varepsilon c\varphi + \psi)y + \varepsilon s < \gamma + \varepsilon c\beta \le (\varepsilon c\varphi + \psi)x + \varepsilon t \qquad \forall (x,t) \in \mathrm{epi}(f),$$

da cui, posto $\overline{G} = \varepsilon c G + \widetilde{G}$ e $\eta = \mathrm{Re}\,\overline{G} = \varepsilon c\varphi + \psi$, si ha

$$\eta y + \varepsilon s < \inf\{\eta x + \varepsilon t : (x,t) \in \mathrm{epi}(f)\}.$$

La tesi è provata anche in questo caso. $\square$

Le funzioni convesse e semicontinue inferiormente hanno la proprietà di essere approssimabili dal basso con funzioni affini.

Proposizione 2.3.7 *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa, semicontinua inferiormente e tale che $f(x) > -\infty$ per ogni $x \in X$. Allora:*

(i) *f ha una funzione affine minorante g, ossia esistono $b \in \mathbb{R}$ e $G : X \to \mathbb{C}$ lineare e continuo, tali che $f(x) \ge g(x) = \mathrm{Re}\,Gx + b$ per ogni $x \in X$;*

(ii) *f è l'inviluppo delle sue funzioni affini minoranti, ossia*

$$f(x) = \sup\{g(x) : \ g \text{ è una funzione affine, } g \le f \text{ in } X\} \quad \forall x \in X.$$

Dimostrazione Anzitutto, se $f \equiv +\infty$ si ha $f(x) = +\infty = \sup_{n\in\mathbb{N}} n$ e le proprietà (i), (ii) sono ovvie.
Se invece esiste $z \in X$ tale che $f(z) \in \mathbb{R}$, in particolare $\mathrm{epi}(f)$ è un convesso chiuso non vuoto; sia allora $y \in X$ e sia $a \in]-\infty, f(y)[$: mostreremo che esiste una funzione affine g, minorante f, tale che $a < g(y)$. Ciò proverà la tesi.
Dato che $(y,a) \notin \mathrm{epi}(f)$, applicando il lemma 2.3.6 (ii) troviamo $G : X \to \mathbb{C}$ lineare e continuo e $\lambda > 0$ tali che, posto $\varphi = \mathrm{Re}\,G$,

$$\varphi y + \lambda a < \beta = \inf\{\varphi x + \lambda t : \ (x,t) \in \mathrm{epi}(f)\}.$$

Perciò, posto $g(x) = -\frac{1}{\lambda}\,\varphi x + \frac{\beta}{\lambda}$, si ha

$$a < g(y), \qquad g(x) \le f(x) \qquad \forall x \in D(f),$$

e a maggior ragione

$$g(x) \le f(x) \qquad \forall x \in X,$$

che è quanto si voleva. $\square$

Se f è una funzione qualunque su uno spazio vettoriale topologico X, è possibile considerare la *più grande* funzione convessa semicontinua inferiormente e minore di f. Più precisamente:

Definizione 2.3.8 *Sia X uno spazio vettoriale topologico e sia $f : X \to \overline{\mathbb{R}}$ una funzione. La* regolarizzata semicontinua *di f è l'inviluppo delle funzioni convesse e semicontinue inferiormente minoranti di f, ossia è la funzione*

$$\overline{f}(x) = \sup\{g(x) : \begin{array}{l} g : X \to \overline{\mathbb{R}}, \ g \text{ è convessa} \\ \text{e semicontinua inferiormente}, \ \ g \leq f \text{ in } X\}, \quad x \in X. \end{array}$$

Notiamo che esistono sempre funzioni convesse e semicontinue minoranti di f: ad esempio, $g(x) \equiv -\infty$. Dunque $\overline{f}$ è sempre ben definita. È ovvio che $\overline{f}(x) \leq f(x)$ per ogni $x \in X$; si noti poi che per l'esercizio 2.1.2 $\overline{f}$ è convessa, e per il corollario 2.2.3 $\overline{f}$ è semicontinua inferiormente.
Il caso più significativo in cui si applica la definizione 2.3.8 è quello in cui la funzione f di partenza è già convessa. In tal caso si ha:

Proposizione 2.3.9 *Sia X uno spazio vettoriale topologico, sia $f : X \to \overline{\mathbb{R}}$ convessa, e sia $\overline{f}$ la sua regolarizzata semicontinua. Allora:*

(i) $\text{epi}(\overline{f}) = \overline{\text{epi}(f)}$*;*

(ii) *se $x_0 \in X$ e $\liminf_{x \to x_0} f(x) < +\infty$, allora $\overline{f}(x_0) = \liminf_{x \to x_0} f(x)$;*

(iii) *f è semicontinua inferiormente in un punto $x_0 \in X$ se e solo se risulta $f(x_0) = \overline{f}(x_0)$.*

Si noti che se la condizione $\liminf_{x \to x_0} f(x) < \infty$ è violata, la proprietà (ii) non vale in generale: ad esempio per la funzione $I_{\{0\}}$ in $\mathbb{R}$ si ha $\overline{I_{\{0\}}}(0) = I_{\{0\}}(0) = 0 < +\infty = \liminf_{x \to 0} I_{\{0\}}(x)$.

Dimostrazione (i) Poiché $\overline{f}$ è semicontinua inferiormente e $\overline{f} \leq f$, $\text{epi}(\overline{f})$ è chiuso e contiene $\text{epi}(f)$: quindi $\text{epi}(\overline{f}) \supseteq \overline{\text{epi}(f)}$.
Proviamo che, viceversa, $\text{epi}(\overline{f}) \subseteq \overline{\text{epi}(f)}$. Mostreremo che $\overline{\text{epi}(f)} = \text{epi}(g)$, con g funzione convessa, semicontinua inferiormente e tale che $g \leq f$ in X; da ciò, per definizione di $\overline{f}$, seguirà che $g \leq \overline{f}$ e dunque che $\overline{\text{epi}(f)} = \text{epi}(g) \supseteq \text{epi}(\overline{f})$. Osserviamo a questo scopo che per ogni $x \in X$ l'insieme $\overline{\text{epi}(f)} \cap (\{x\} \times \mathbb{R})$ è sempre una semiretta (eventualmente vuota, o coincidente

con tutto $\mathbb{R}$): infatti se $(x,t) \in \overline{\mathrm{epi}(f)}$ esiste un net $\{(x_i,t_i)\}_{i\in I} \subseteq \mathrm{epi}(f)$ tale che $x_i \to x$ in X e $t_i \to t$ in $\mathbb{R}$; quindi, fissato $\tau > t$ si troverà $i_0 \in I$ per cui risulta $t_i \leq \tau$ per $i \geq i_0$. Di conseguenza avremo $f(x_i) \leq t_i \leq \tau$, cioè $(x_i,\tau) \in \mathrm{epi}(f)$, per ogni $i \geq i_0$, e quindi passando al limite otteniamo $(x,\tau) \in \overline{\mathrm{epi}(f)}$. In altre parole, abbiamo mostrato che

$$(x,t) \in \overline{\mathrm{epi}(f)} \quad \Longrightarrow (x,\tau) \in \overline{\mathrm{epi}(f)} \quad \forall \tau > t.$$

Possiamo allora definire la funzione g nel modo seguente:

$$g(x) = \begin{cases} +\infty & \text{se } \overline{\mathrm{epi}(f)} \cap (\{x\} \times \mathbb{R}) = \emptyset, \\ \inf\{t \in \mathbb{R} : (x,t) \in \overline{\mathrm{epi}(f)}\} & \text{altrimenti.} \end{cases}$$

È facile verificare che $\mathrm{epi}(g) = \overline{\mathrm{epi}(f)}$, e dunque $\mathrm{epi}(g)$ è chiuso; essendo la chiusura del convesso $\mathrm{epi}(f)$, $\mathrm{epi}(g)$ è anche convesso e, ovviamente, contiene $\mathrm{epi}(f)$. Ne segue che g è semicontinua inferiormente e convessa, e che $g \leq f$ in X. Ciò prova (i).

(ii) Poiché $\overline{f}$ è semicontinua inferiormente e $\overline{f} \leq f$ in X, si ha

$$\overline{f}(x_0) \leq \liminf_{x\to x_0} \overline{f}(x) \leq \liminf_{x\to x_0} f(x) \qquad \forall x_0 \in X.$$

Proviamo la disuguaglianza opposta. Per la parte già provata, $\overline{f}(x_0) < +\infty$. Sia $t \in]\overline{f}(x_0), +\infty[$: allora $(x_0,t) \in \mathrm{epi}(\overline{f}) = \overline{\mathrm{epi}(f)}$. Quindi, se $\{U_i \times I_i\}_{i\in I}$ è un sistema fondamentale di intorni di (x_0,t) in $X \times \mathbb{R}$, per ogni $i \in I$ esiste un punto $(x_i,t_i) \in (U_i \times I_i) \cap \mathrm{epi}(f)$. Se risulta $x_i \neq x_0$ per infiniti indici $i \in I$, essendo $f(x_i) \leq t_i$ deduciamo che $\liminf_{x\to x_0} f(x) \leq t$.
Se invece risulta $x_i = x_0$ definitivamente, allora otteniamo $f(x_0) \leq t_i$ definitivamente, da cui $f(x_0) \leq t$; in questo caso, d'altronde, è facile provare che $\liminf_{x\to x_0} f(x) \leq f(x_0)$. Infatti, posto per comodità $\lambda = \liminf_{x\to x_0} f(x)$, fissato $\varepsilon > 0$ esiste un intorno V di 0, che possiamo supporre bilanciato, tale che

$$\lambda - \varepsilon < \inf_{y\in V\setminus\{0\}} f(x_0+y) \leq \lambda.$$

Per definizione di estremo inferiore, si può anche trovare un punto $\xi \in V \setminus \{0\}$ in modo che

$$\lambda - \varepsilon \leq \inf_{y\in V\setminus\{0\}} f(x_0+y) \leq f(x_0+\xi) \leq \lambda + \varepsilon.$$

Sia ora $\delta \in]0,1]$; poiché V è bilanciato, si ha $\delta\xi \in V$. Dunque, per convessità,

$$\begin{aligned} \lambda - \varepsilon &\leq f(x_0 + \delta\xi) = f((1-\delta)x_0 + \delta(x_0+\xi)) \leq \\ &\leq (1-\delta)f(x_0) + \delta f(x_0+\xi) \leq \\ &\leq (1-\delta)f(x_0) + \delta(\lambda+\varepsilon) < +\infty. \end{aligned}$$

Per $\delta \to 0^+$ si deduce $\lambda - \varepsilon \leq f(x_0)$, e per l'arbitrarietà di ε si ricava $\lambda \leq f(x_0)$.
Abbiamo così ottenuto, in tutti i casi, che $\liminf_{x\to x_0} f(x) \leq t$ per ogni $t > \overline{f}(x_0)$: dunque $\liminf_{x\to x_0} f(x) \leq \overline{f}(x_0)$ e (ii) è provata.

(iii) La relazione $f(x_0) = \overline{f}(x_0)$ implica

$$f(x_0) \leq \liminf_{x\to x_0} \overline{f}(x) \leq \liminf_{x\to x_0} f(x).$$

Supponiamo viceversa che f sia semicontinua inferiormente in x_0. Se $\overline{f}(x_0) = +\infty$, allora essendo $f \geq \overline{f}$ si ha a maggior ragione $f(x_0) = +\infty = \overline{f}(x_0)$. Se invece $\overline{f}(x_0) < +\infty$, allora per ogni $t > \overline{f}(x_0)$ risulta $(x_0, t) \in \mathrm{epi}(\overline{f}) = \overline{\mathrm{epi}(f)}$: quindi esiste un net $\{(x_i, t_i)\}_{i\in I} \subseteq \mathrm{epi}(f)$ tale che $x_i \to x$ in X e $t_i \to t$ in $\mathbb{R}$. Se $x_i \neq x_0$ per infiniti indici avremo di conseguenza

$$f(x_0) \leq \liminf_i f(x_i) \leq \liminf_i t_i = t \qquad \forall t > \overline{f}(x_0),$$

mentre, nel caso in cui sia $x_i = x_0$ definitivamente, la stessa relazione segue in modo ancora più semplice. In definitiva, per l'arbitrarietà di t, otteniamo $f(x_0) \leq \overline{f}(x_0)$. Ciò implica l'uguaglianza. □

Esercizi 2.3

1. Esibire una funzione convessa, semicontinua inferiormente, non propria e non identicamente uguale né a $+\infty$, né a $-\infty$.

2. Esibire una funzione convessa, propria e non semicontinua inferiormente in almeno un punto.

3. Esibire una funzione convessa, propria, semicontinua inferiormente ma non continua in almeno un punto.

4. Sia X uno spazio di Banach. Si verifichi che una funzione convessa, semicontinua inferiormente e propria può non avere minimo su X.

5. Sia X uno spazio di Banach riflessivo, sia $K \subset X$ un convesso non vuoto, chiuso e limitato e sia $f : K \to \overline{\mathbb{R}}$ una funzione quasi convessa e semicontinua inferiormente. Si provi che f ha minimo su K.

6. Sia H uno spazio di Hilbert e poniamo
$$J(x) = \frac{1}{2}\|x\|_H^2 - \langle z, x\rangle_H\,, \qquad x \in H,$$
ove z è un fissato elemento di H. Si provi che:
   (i) J è una funzione strettamente convessa, continua e finita;
   (ii) $J(x) \to \infty$ per $\|x\|_H \to \infty$;
   (iii) J ha minimo su H ed il punto di minimo è z;
   (iv) se $K \subseteq H$ è un convesso chiuso, allora J ha minimo su K ed il punto di minimo u è l'unica soluzione in K della *disequazione variazionale*
   $$\mathrm{Re}\,\langle u, v-u\rangle_H \geq \mathrm{Re}\,\langle z, v-u\rangle_H \qquad \forall v \in K.$$

   [**Traccia:** per (iv) si usi la condizione $\left[\frac{d}{dt}J(tv+(1-t)u)\right]_{t=0} \geq 0$.]

7. Sia H uno spazio di Hilbert, sia $\{e_k\}_{k\in\mathbb{N}^+}$ un sistema ortonormale completo e sia $F = \overline{B(0,1)} \setminus B(0,\frac{1}{2})$. Posto
$$f(x) = I_F(x) + \sum_{k=1}^{\infty} \frac{|\langle x, e_k\rangle_H|}{k} \qquad \forall x \in H,$$
si provi che f è una funzione semicontinua inferiormente che non ha minimo nel chiuso limitato F.

8. Sia X uno spazio vettoriale topologico reale. Si provi che i funzionali lineari su $X \times \mathbb{R}$ sono tutti e soli quelli della forma $(x,t) \mapsto Gx + \lambda t$, con $\lambda \in \mathbb{R}$ e $G : X \to \mathbb{C}$ lineare. Si verifichi poi che i funzionali lineari e continui su $X \times \mathbb{R}$ sono tutti e soli quelli della forma precedente con $\lambda \in \mathbb{R}$ e $G : X \to \mathbb{C}$ lineare e continuo.

9. Sia f una funzione convessa su uno spazio vettoriale topologico X, e sia $\overline{f}$ la sua regolarizzata semicontinua. Se $\liminf_{x\to x_0} f(x) < +\infty$ per ogni $x_0 \in X$, si provi che
$$\{x \in X : \ \overline{f}(x) \leq \alpha\} = \bigcap_{\lambda>\alpha} \overline{\{x \in X : \ f(x) \leq \lambda\}} \quad \forall \alpha \in \mathbb{R}.$$

2.4 Funzioni convesse continue

Vediamo qualche condizione che assicuri la continuità di una funzione convessa. Cominciamo con un lemma che verrà usato sistematicamente nel seguito del paragrafo.

Lemma 2.4.1 *Sia X uno spazio vettoriale topologico, sia $f : X \to \overline{\mathbb{R}}$ convessa. Se $z_0 \in X$ e f è limitata superiormente in un intorno di z_0, allora f è continua in z_0.*

Dimostrazione Per ipotesi, esiste $K > 0$ ed esiste un intorno bilanciato V di 0 tale che $f(y) \leq K$ per ogni $y \in z_0 + V$. Sia $\varepsilon \in]0,1[$: se $x \in \varepsilon V$, si ha

$$\begin{aligned} f(z_0 + x) &= f\left((1-\varepsilon)z_0 + \varepsilon\left(z_0 + \frac{x}{\varepsilon}\right)\right) \leq \\ &\leq (1-\varepsilon)f(z_0) + \varepsilon f\left(z_0 + \frac{x}{\varepsilon}\right) \leq (1-\varepsilon)f(z_0) + \varepsilon K. \end{aligned}$$

Da qui segue intanto che se $f(z_0) = -\infty$, allora $f \equiv -\infty$ in $z_0 + \varepsilon V$, e dunque f è continua in z_0. Quindi possiamo supporre che $f(z_0) > -\infty$; in questo caso, per sottrazione otteniamo

$$f(z_0 + x) - f(z_0) \leq \varepsilon(|f(z_0)| + K) \quad \forall x \in \varepsilon V.$$

D'altra parte, se $x \in \varepsilon V$ si ha

$$f(z_0) = f\left(\frac{1}{2}(z_0 + x) + \frac{1}{2}(z_0 - x)\right) \leq \frac{1}{2}f(z_0 + x) + \frac{1}{2}f(z_0 - x),$$

da cui, osservando che se $x \in \varepsilon V$ anche $-x \in \varepsilon V$,

$$f(z_0) - f(z_0 + x) \leq f(z_0 - x) - f(z_0) \leq \varepsilon(|f(z_0)| + K) \quad \forall x \in \varepsilon V;$$

quindi si conclude che

$$|f(z_0 + x) - f(z_0)| \leq \varepsilon(|f(z_0)| + K) \quad \forall x \in \varepsilon V,$$

cosicché f è continua in z_0. $\square$

Teorema 2.4.2 *Sia X uno spazio vettoriale topologico e sia $f : X \to \overline{\mathbb{R}}$ convessa e propria. Sono fatti equivalenti:*

(i) *$\overset{\circ}{D(f)}$ è non vuoto e f è continua in ogni punto di $\overset{\circ}{D(f)}$;*

(ii) *esiste* $z_0 \in \overset{\circ}{D}(f)$ *tale che* f *è continua in* z_0*;*

(iii) *esiste* $z_0 \in D(f)$ *tale che* f *è limitata superiormente in un intorno di* z_0*.*

Dimostrazione (i) $\Longrightarrow$ **(ii)** Ovvio.

(ii) $\Longrightarrow$ **(iii)** Poiché f è continua nel punto $z_0 \in \overset{\circ}{D}(f)$, si ha $f(z_0) \in \mathbb{R}$ ed esiste un intorno bilanciato V di 0 tale che $f(y) < f(z_0) + 1$ per ogni $y \in z_0 + V$. Quindi f è limitata superiormente in un intorno di z_0.
(iii) $\Longrightarrow$ **(i)** Per ipotesi esiste un intorno bilanciato V di 0 tale che $f(y) \leq K < +\infty$ per ogni $y \in z_0 + V$; in particolare, $z_0 \in \overset{\circ}{D}(f)$. Sia z un arbitrario punto di $\overset{\circ}{D}(f)$: allora esiste $t > 0$ tale che $z' = z + t(z - z_0) \in D(f)$. Consideriamo l'intorno $W = z + \frac{t}{t+1}V$ di z: per ogni $u \in W$ si ha $u = z + \frac{t}{t+1}y$ con $y \in V$, da cui, essendo $z = \frac{1}{t+1}(z' + tz_0)$, otteniamo per ogni $y \in V$, ossia per ogni $u \in W$,

$$\begin{aligned} f(u) &= f\left(\frac{1}{t+1}(z' + tz_0) + \frac{t}{t+1}y\right) = f\left(\frac{1}{t+1}z' + \frac{t}{t+1}(z_0 + y)\right) \leq \\ &\leq \frac{1}{1+t}f(z') + \frac{t}{t+1}f(z_0 + y) \leq \frac{1}{t+1}|f(z')| + \frac{t}{t+1}K. \end{aligned}$$

Dunque f è limitata superiormente nell'intorno W di z. Per il lemma 2.4.1, f è continua nel punto z. $\square$

Corollario 2.4.3 *Sia* X *uno spazio vettoriale topologico e sia* $f : X \to \overline{\mathbb{R}}$ *convessa e propria. Se* f *è semicontinua superiormente in un punto* $x_0 \in D(f)$*, allora* f *è continua in* $\overset{\circ}{D}(f)$*.*

Dimostrazione Dall'ipotesi segue che $+\infty > f(x_0) \geq \limsup_{x \to x_0} f(x)$, quindi esiste un intorno V di x_0 tale che

$$\sup_V f \leq f(x_0) + 1 < +\infty.$$

Dal teorema 2.4.2 segue la tesi. $\square$

Dai risultati precedenti segue che in relazione alla continuità le funzioni convesse hanno un comportamento analogo a quello delle funzioni lineari: la continuità equivale alla continuità in un singolo punto ed alla locale limitatezza nell'intorno di un singolo punto. La differenza sta nel fatto che la continuità non è globale in X, ma solo in $\overset{\circ}{D}(f)$. Se però si considerano solo funzioni convesse *finite*, allora $\overset{\circ}{D}(f) = D(f) = X$ e l'analogia è perfetta. Vediamo ora qualche situazione particolare in cui sulla continuità delle funzioni convesse si può dire qualcosa di più.

Teorema 2.4.4 *Sia X uno spazio di Banach e sia $f : X \to \mathbb{R}$ una funzione convessa e semicontinua inferiormente. Allora f è continua su X.*

Dimostrazione Fissato $x_0 \in X$, basterà provare che f è limitata superiormente in un intorno di x_0. Sia

$$A = \{x \in X : \ f(x_0 \pm x) \leq |f(x_0)| + 1\};$$

poiché f è convessa e semicontinua inferiormente, A è convesso e chiuso; inoltre $0 \in A$ ed A è simmetrico. Se $x \in X$ si ha, scelto un intero $n > \max\{f(x_0 + x) - f(x_0), f(x_0 - x) - f(x_0)\}$ (il che è possibile essendo f a valori in $\mathbb{R}$),

$$\begin{aligned} f\left(x_0 \pm \frac{1}{n}\, x\right) &= f\left(\left(1 - \frac{1}{n}\right) x_0 + \frac{1}{n}(x_0 \pm x)\right) \leq \\ &\leq \left(1 - \frac{1}{n}\right) f(x_0) + \frac{1}{n}\, f(x_0 \pm x) = \\ &= f(x_0) + \frac{1}{n}\, [f(x_0 \pm x) - f(x_0)] \leq |f(x_0)| + 1. \end{aligned}$$

Dunque, con questa scelta di n si ha $x \in nA$. Ciò prova che $X = \bigcup_{n \in \mathbb{N}^+} nA$. Dato che X è uno spazio di Banach, il teorema di Baire ci dice che esiste $\nu \in \mathbb{N}^+$ tale che νA ha parte interna non vuota: dunque esiste una palla $B(y_0, r)$ contenuta in νA. Quindi, se $w \in B(0, r)$ si ha $y_0 \pm w \in \nu A$, ed essendo A simmetrico abbiamo anche $-y_0 \pm w \in \nu A$. Poiché A è convesso, deduciamo poi che $w = \frac{1}{2}(w + y_0) + \frac{1}{2}(w - y_0) \in \nu A$; si conclude pertanto che $B(0, r) \subseteq \nu A$, ossia $B(0, \frac{r}{\nu}) \subseteq A$. Per definizione di A, ciò significa che f è limitata superiormente in $B(x_0, \frac{r}{\nu})$. Ne segue la tesi. $\square$

Proposizione 2.4.5 *Sia X uno spazio normato e sia $f : X \to \overline{\mathbb{R}}$ convessa. Se f è limitata in una palla $B(x_0, r)$, allora f è localmente lipschitziana in $B(x_0, r)$.*

Dimostrazione Sia $x \in B(x_0, r)$ e sia $\delta > 0$ tale che $B(x, 2\delta) \subseteq B(x_0, r)$; proveremo che f è lipschitziana in $B(x, \delta)$. Per ipotesi, $N_x = \sup_{B(x,2\delta)} |f| <$

∞. Se $x_1, x_2 \in B(x, \delta)$, poniamo $d = \|x_1 - x_2\|_X$ e $x_3 = x_2 + \frac{\delta}{d}(x_2 - x_1)$. Allora $x_3 \in B(x, 2\delta)$ e $x_2 = \frac{\delta}{d+\delta}\, x_1 + \frac{d}{d+\delta}\, x_3$; quindi per convessità deduciamo

$$f(x_2) \le \frac{\delta}{d+\delta} f(x_1) + \frac{d}{d+\delta} f(x_3),$$

da cui

$$f(x_2) - f(x_1) \le \frac{d}{d+\delta}[f(x_3) - f(x_1)] \le \frac{d}{\delta}|f(x_3) - f(x_1)| \le \frac{2N_x}{\delta}\|x_1 - x_2\|_X.$$

Scambiando i ruoli di x_1 e x_2, si conclude che

$$|f(x_2) - f(x_1)| \le \frac{2N_x}{\delta}\|x_1 - x_2\|_X. \quad \square$$

Infine enunciamo un risultato di differenziabilità, dovuto a Mazur, valido in spazi di Banach separabili, la dimostrazione del quale sarà fornita nel capitolo 5.

Teorema 2.4.6 (di Mazur) *Sia X uno spazio di Banach separabile, sia $D \subseteq X$ un aperto convesso non vuoto e sia $f : D \to \mathbb{R}$ una funzione convessa e continua. Allora f è G-differenziabile in un insieme $D_0 \subseteq D$, denso in D, che è unione numerabile di aperti densi.*

Esercizi 2.4

1. Sia X uno spazio vettoriale topologico di dimensione finita. Se $f : X \to \overline{\mathbb{R}}$ è convessa e propria, si provi che f è continua su $\overset{\circ}{D(f)}$.

   [**Traccia:** Se $x_0 \in \overset{\circ}{D(f)}$, si fissi un cubo C di centro x_0 contenuto in $D(f)$, e si osservi che ogni $x \in C$ è combinazione convessa dei vertici x_i del cubo; se ne deduca che $f(x) \le \max\{f(x_i)\}$ per ogni $x \in C$ e di qui si ricavi la tesi.]

2. Sia X uno spazio vettoriale topologico e sia $K \subseteq X$ un convesso contenente 0. Si provi che il funzionale di Minkowski p_K è una funzione continua se e solo se $0 \in \overset{\circ}{K}$.

3. Sia X uno spazio normato. Esibire una funzione $f : X \to \mathbb{R}$ continua, ma non w-continua.

2.5 Funzioni coniugate

La nozione di funzione coniugata (relativa ad una funzione assegnata) è assai utile nelle applicazioni ai problemi di ottimizzazione, specialmente nel caso in cui la funzione assegnata sia convessa. Si può comunque definire la coniugata di una funzione qualunque.

Definizione 2.5.1 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. La* coniugata*, o* polare*, di f è la funzione $f^* : X^* \to \overline{\mathbb{R}}$ così definita:*

$$f^*(\varphi) = \sup_{x\in X}\{\varphi x - f(x)\} \qquad \forall \varphi \in X^*.$$

Osservazioni 2.5.2 **(1)** La definizione ha sempre senso, in quanto φx è un numero reale per ogni $x \in X$; risulta $f^*(\varphi) \le b$ se e solo se la funzione affine $x \mapsto \varphi x - b$ è minorante di f.

(2) La funzione f^* è sempre convessa e w^*-semicontinua inferiormente, perché è l'estremo superiore di funzioni che, rispetto alla variabile φ, sono convesse e w^*-semicontinue inferiormente.

(3) Può capitare che sia $f^* \equiv +\infty$: ad esempio, se $X = \mathbb{R}$ e $f(x) = -e^x$, si ha $f^*(y) = \sup_{x\in\mathbb{R}}\{yx + e^x\} = +\infty$ per ogni $y \in \mathbb{R}^* = \mathbb{R}$. Però quando f è convessa ciò non può accadere (sotto certe ipotesi), come mostra la proposizione che segue.

Proposizione 2.5.3 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$ una funzione convessa, semicontinua inferiormente e propria. Allora la coniugata f^* è propria.*

Dimostrazione Per ipotesi, $D(f)$ è non vuoto, quindi scelto $x \in D(f)$ si ha $f^*(\varphi) \ge \varphi x - f(x) > -\infty$. Inoltre, per la proposizione 2.3.7, esistono $\varphi \in X^*$ e $b \in \mathbb{R}$ tali che $f(x) \ge \varphi x + b$ per ogni $x \in X$: dunque, per l'osservazione precedente, $f^*(\varphi) \le -b$, cosicché f^* è propria. $\square$

Osservazione 2.5.4 È utile precisare la seguente proprietà: si ha $f^*(\varphi) > -\infty$ per ogni $\varphi \in X^*$ se e solo se f non è identicamente uguale a $+\infty$. La dimostrazione si lascia al lettore (esercizio 2.5.1).

Il termine "coniugata" di f si giustifica sulla base della seguente proprietà:

Proposizione 2.5.5 (disuguaglianza di Young) *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Se f^* è la coniugata di f, si ha*

$$f^*(\varphi) \geq \varphi x - f(x) \qquad \forall \varphi \in X^*, \quad \forall x \in X.$$

In particolare, se f è propria la disuguaglianza acquista la forma simmetrica

$$\varphi x \leq f(x) + f^*(\varphi) \qquad \forall \varphi \in X^*, \quad \forall x \in X.$$

Dimostrazione La prima forma della disuguaglianza è banale conseguenza della definizione di f^*. Per la seconda basta osservare che se f è propria, allora $f^* > -\infty$ per l'osservazione 2.5.4, e quindi la somma $f(x) + f^*(\varphi)$ è sempre ben definita. $\square$

Esempi 2.5.6 **(1)** Sia $X = \mathbb{R}$. Fissato $p \in]1, \infty[$, poniamo

$$f(x) = \frac{|x|^p}{p}\,, \qquad x \in \mathbb{R}.$$

Allora

$$f^*(y) = \sup_{x \in \mathbb{R}} \left\{ yx - \frac{|x|^p}{p} \right\} = \frac{|y|^q}{q} \qquad \forall y \in \mathbb{R},$$

ove $\frac{1}{p} + \frac{1}{q} = 1$. Infatti, $f^*(y) \leq |y|^q/q$ per la disuguaglianza di Young classica, e vale l'uguaglianza, come si verifica facilmente, quando $x = \mathrm{sgn}(y) \cdot |y|^{q/p}$.

(2) Sia $X = \mathbb{R}$ e sia $f(x) = |x|$. Allora

$$f^*(y) = \sup_{x \in \mathbb{R}} \{ yx - |x| \} = \begin{cases} 0 & \text{se } |y| \leq 1 \\ +\infty & \text{se } |y| > 1, \end{cases}$$

cosicché $f^* = I_{[-1,1]}$.

(3) Sia $X = \mathbb{R}$ e sia $f(x) = e^x$. Allora si trova facilmente che

$$f^*(y) = \sup_{x \in \mathbb{R}} \{ yx - e^x \} = \begin{cases} +\infty & \text{se } y < 0, \\ 0 & \text{se } y = 0, \\ y(\log y - 1) & \text{se } y > 0. \end{cases}$$

(4) Sia X uno spazio normato reale e sia U un sottoinsieme non vuoto di X. Per la funzione I_U si ha

$$I_U^*(\varphi) = \sup_{x \in X} \{ \varphi x - I_U(x) \} = \sup_{x \in U} \{ \varphi x \};$$

la funzione I_U^* si chiama *funzione di supporto* di U. Essa è positivamente omogenea e subadditiva; inoltre vale l'equivalenza

$$U \subseteq \{x \in X : \ \varphi x \le \alpha\} \qquad \Longleftrightarrow \qquad I_U^*(\varphi) \le \alpha.$$

Le proprietà della funzione coniugata sono raccolte nella seguente

Proposizione 2.5.7 *Sia X uno spazio normato reale e siano $f, g, f_i : X \to \overline{\mathbb{R}}$. Allora:*

(i) $f^*(0) = -\inf_{x\in X} f(x)$*;*

(ii) *se $f \le g$, allora $f^* \ge g^*$;*

(iii) $(\inf_{i\in I} f_i)^* = \sup_{i\in I} f_i^*$*;*

(iv) $(\sup_{i\in I} f_i)^* \le \inf_{i\in I} f_i^*$*;*

(v) $(\lambda f)^*(\varphi) = \lambda f^*(\frac{\varphi}{\lambda})$ *per ogni $\varphi \in X^*$ e per ogni $\lambda > 0$;*

(vi) *posto $f_u(x) = f(x-u)$, si ha $f_u^*(\varphi) = f^*(\varphi) + \varphi u$ per ogni $\varphi \in X^*$ e per ogni $u \in X$.*

Dimostrazione Sono tutte facili conseguenze della definizione 2.5.1. □

Possiamo anche considerare la *bipolare* di f, ossia la coniugata di f^*: sarà la funzione $f^{**} : X^{**} \to \overline{\mathbb{R}}$ definita da

$$f^{**}(\alpha) = \sup_{\varphi\in X^*} \{\alpha\varphi - f^*(\varphi)\} \qquad \forall \alpha \in X^{**}.$$

La funzione f^{**} è convessa e semicontinua inferiormente rispetto alla topologia debole* di X^{**}. Se calcoliamo f^{**} nel punto $\alpha = J_x$, ove $x \in X$ e J è l'immersione canonica di X in X^{**}, si ha per la disuguaglianza di Young

$$f^{**}(J_x) = \sup_{\varphi\in X^*} \{\varphi x - f^*(\varphi)\} \le f(x) \qquad \forall x \in X.$$

Tuttavia, sotto opportune ipotesi, $f^{**}(J_x)$ coincide con $f(x)$.

Teorema 2.5.8 (di Fenchel-Moreau) *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$ propria. Allora si ha $f^{**}(J_x) = f(x)$ per ogni $x \in X$ se e solo se f è convessa e semicontinua inferiormente.*

Dimostrazione ($\Longrightarrow$) Sappiamo che f^{**} è convessa e semicontinua inferiormente per la topologia debole* di X^{**}; quindi $f^{**} \circ J$ è convessa e semicontinua inferiormente per la topologia debole di X. Essendo $f(x) = f^{**}(J_x)$ per ogni $x \in X$, tale è anche f e a maggior ragione si ha la tesi.

($\Longleftarrow$) Poiché $f^{**}(J_x) \leq f(x)$, la tesi è banale se $f^{**}(J_x) = +\infty$. Inoltre, siccome f è propria, f^* non è identicamente $+\infty$ e di conseguenza $f^{**} > -\infty$ (osservazione 2.5.4). Perciò dobbiamo provare che se $f^{**}(J_x) \in \mathbb{R}$, allora $f^{**}(J_x) = f(x)$.
Supponiamo, per assurdo, che esista $x_0 \in X$ tale che $f^{**}(J_{x_0}) < f(x_0)$. Allora $(x_0, f^{**}(J_{x_0}))$ non appartiene al convesso chiuso $\mathrm{epi}(f)$; dunque, per il lemma 2.3.6 (ii), esistono $\varphi \in X^*$, $\alpha > 0$ e $\beta \in \mathbb{R}$ tali che

$$\varphi x_0 + \alpha f^{**}(J_{x_0}) < \beta = \inf_{(x,t)\in \mathrm{epi}(f)} \{\varphi x + \alpha t\}.$$

Osservato che

$$\begin{aligned} \beta &= \inf_{(x,t)\in \mathrm{epi}(f)} \{\varphi x + \alpha t\} = \inf_{x\in D(f)} \{\varphi x + \alpha f(x)\} = \\ &= -\sup_{x\in D(f)} \{-\varphi x - \alpha f(x)\} = -(\alpha f)^*(-\varphi) = -\alpha f^* \left(-\frac{\varphi}{\alpha}\right), \end{aligned}$$

dividendo la disuguaglianza precedente per α si ricava

$$\frac{1}{\alpha}\varphi x_0 + f^{**}(J_{x_0}) < -f^* \left(-\frac{\varphi}{\alpha}\right).$$

Questa relazione contraddice la disuguaglianza di Young: essendo $f^{**} > -\infty$, deve infatti essere

$$-\frac{1}{\alpha}\varphi x_0 = J_{x_0}\left(-\frac{\varphi}{\alpha}\right) \leq f^{**}(J_{x_0}) + f^* \left(-\frac{\varphi}{\alpha}\right). \quad \square$$

Osservazione 2.5.9 Se X è uno spazio normato reale e $f : X \to \overline{\mathbb{R}}$ è una funzione convessa che non è semicontinua inferiormente, allora $f^{**}(J_x)$ coincide con $\overline{f}(x)$, ove $\overline{f}$ è la regolarizzata semicontinua di f introdotta nella definizione 2.3.8. Infatti, $f^{**} \circ J$ è convessa, semicontinua inferiormente e minorante di f: quindi, per definizione di regolarizzata semicontinua, si ha $f^{**} \circ J \leq \overline{f}$. D'altra parte, se g è una funzione convessa, semicontinua inferiormente e minorante di f, dalla proposizione 2.5.7 segue $g^* \geq f^*$ e $g^{**} \leq f^{**}$. Ma dal teorema di Fenchel-Moreau otteniamo

$$g(x) = g^{**}(J_x) \leq f^{**}(J_x) \qquad \forall x \in X,$$

e quindi, per l'arbitrarietà di g, deduciamo $\overline{f} \leq f^{**} \circ J$.

Esempio 2.5.10 Sia X uno spazio normato reale. Se U è un sottoinsieme non vuoto di X, consideriamo la funzione $f = I_U$ (si veda l'esempio 2.5.6 (4)). L'osservazione precedente ci dice che $I_U^{**} \circ J = \overline{I_U}$; d'altra parte, poiché il minimo convesso chiuso contenente U è $\overline{\mathrm{co}(U)}$, si riconosce facilmente che $\overline{I_U} = I_{\overline{\mathrm{co}(U)}}$. Abbiamo così provato che

$$I_U^{**}(J_{x_0}) = I_{\overline{\mathrm{co}(U)}}(x_0) \qquad \forall x_0 \in X,$$

e naturalmente risulta in generale $I_{\overline{\mathrm{co}(U)}}(x) < I_U(x)$, salvo che U non sia un convesso chiuso di X; in questo caso, come è giusto, I_U è convessa e semicontinua inferiormente.

Esercizi 2.5

1. Sia X uno spazio normato e sia $f : X \to \overline{\mathbb{R}}$. Si provi che la coniugata f^* non assume il valore $-\infty$ se e solo se f non è identicamente uguale a $+\infty$.

2. Sia X uno spazio normato. Si calcoli f^* quando $f \in X^*$.

3. Sia $X = \mathbb{R}$. Determinare le coniugate delle funzioni:

$$f_1(x) = \cosh x, \qquad\qquad f_2(x) = \max\{e^x, e^{-x}\},$$

$$f_3(x) = \begin{cases} -\log x & \text{se } x > 0 \\ +\infty & \text{se } x \le 0, \end{cases} \qquad f_4(x) = \begin{cases} -\cos x & \text{se } |x| < 1 \\ +\infty & \text{se } |x| \ge 1. \end{cases}$$

4. Verificare con un esempio che in generale si ha

$$\min\{f_1^*, f_2^*\} \neq (\max\{f_1, f_2\})^*.$$

5. Sia $X = \mathbb{R}$. Per ogni $n \in \mathbb{N}^+$ si calcoli la coniugata della funzione $f_n(x) = \sqrt[n]{1 + |x|^n}$.

6. Sia $X = \mathbb{R}^N$. Si calcoli la coniugata della funzione $f(x) = \frac{1}{2}(Qx, x)_{\mathbb{R}^N}$, ove Q è una matrice $N \times N$ definita positiva.

7. Sia $f : \mathbb{R}^N \to \mathbb{R}$ una funzione convessa. Si provi che

$$\lim_{|y|_N \to \infty} \frac{f^*(y)}{|y|_N} = +\infty.$$

8. Sia X uno spazio normato e sia $H : \mathbb{R} \to \mathbb{R}$ una funzione convessa e semicontinua inferiormente. Posto $F(x) = H(\|x\|_X)$ per ogni $x \in X$ e $G(\varphi) = H^*(\|\varphi\|_{X^*})$ per ogni $\varphi \in X^*$, si provi che $F^* = G$.

9. Sia X uno spazio normato e sia $A \subseteq X$ un convesso contenente 0. Se p_A è il funzionale di Minkowski di A, si provi che
$$p_A^* = I_{\{\varphi \in X^* : \, I_A^*(\varphi) \leq 1\}}.$$

10. Sia X uno spazio di Banach riflessivo e sia U un sottoinsieme non vuoto, convesso e limitato di X. Si provi che si ha $I_U^*(\varphi) = \beta$ se e solo se il semipiano $\{x \in X : \varphi x \leq \beta\}$ contiene U e l'iperpiano affine $\{x \in X : \varphi x = \beta\}$ è di appoggio per $\overline{U}$.

11. Sia $U = \{(x, y) \in \mathbb{R}^2 : \, x > 0, \, xy < -1\}$. Si verifichi che $I_U^*(0, 1) = 0$, che U è contenuto nel semipiano $y \leq 0$ ma che la retta $y = 0$ non è di appoggio per $\overline{U}$.

12. Si consideri in $L^1(\mathbb{R})$ il convesso limitato $U = \mathrm{co}(\{\chi_{[n,n+1]}\}_{n \in \mathbb{Z}})$. Posto $F\psi = -\int_{\mathbb{R}} e^{-x^2} \psi(x) dx$ per ogni $\psi \in L^1(\mathbb{R})$, si provi che risulta $I_U^*(F) = 0$, che $U \subset \{\psi \in L^1(\mathbb{R}) : F\psi \leq 0\}$ ma che il semispazio $\{\psi \in L^1(\mathbb{R}) : F\psi \leq 0\}$ non è di appoggio per $\overline{U}$.

13. Sia X uno spazio normato e siano U, V convessi non vuoti contenuti in X. Si provi che $U \subseteq V$ se e solo se $I_U^* \leq I_V^*$, dove I_U^* e I_V^* sono le rispettive funzioni di supporto (esempio 2.5.6 (4)).

2.6 Problemi di ottimizzazione

Le funzioni coniugate mostrano tutta la loro utilità nei problemi di ottimizzazione che coinvolgono funzioni convesse. In questo paragrafo illustriamo alcuni esempi relativi a questa vasta tematica.
Ci occuperemo del problema
$$\exists \min_{u \in X} F(u)$$
ove $F : X \to \overline{\mathbb{R}}$ è una funzione convessa, semicontinua inferiormente e propria e X è uno spazio normato reale; cercheremo di caratterizzare i punti di

minimo attraverso delle opportune “relazioni di estremalità”. A questo scopo sarà conveniente “immergere” il problema sopra descritto in una famiglia di problemi del tipo

$$\exists \min_{u \in X} \phi(u, p)$$

ove p è un parametro che varia in un altro spazio normato reale Y, e ϕ è un'altra funzione convessa, semicontinua inferiormente e propria in $X \times Y$, tale che $\phi(u, 0) = F(u)$ per ogni $u \in X$.
Cominciamo con alcuni fatti preliminari.

Lemma 2.6.1 *Siano X, Y spazi normati reali e sia $\phi : X \times Y \to \overline{\mathbb{R}}$ una funzione convessa. Allora la funzione*

$$h(y) = \inf_{u \in X} \phi(u, y), \qquad y \in Y,$$

è convessa.

Dimostrazione Siano $p, q \in Y$ tali che non risulti $h(p) = -h(q) = \pm\infty$. Se $h(p)$ oppure $h(q)$ vale $+\infty$, la disuguaglianza di convessità è ovvia; supponiamo dunque $h(p) < +\infty$ e $h(q) < +\infty$. Fissiamo $a > h(p)$ e $b > h(q)$: allora esistono $u, v \in X$ tali che

$$h(p) \le \phi(u, p) < a, \qquad h(q) \le \phi(v, q) < b.$$

Dunque per ogni $\lambda \in]0, 1[$ si ha

$$\begin{aligned} h(\lambda p + (1 - \lambda) q) &= \inf_{w \in X} \phi(w, \lambda p + (1 - \lambda) q) \le \\ &\le \phi(\lambda u + (1 - \lambda) v, \lambda p + (1 - \lambda) q) \le \\ &\le \lambda \phi(u, p) + (1 - \lambda) \phi(v, q) \le \lambda a + (1 - \lambda) b, \end{aligned}$$

e per l'arbitrarietà di $a > h(p)$ e $b > h(q)$ si ottiene la disuguaglianza di convessità. $\square$

Lemma 2.6.2 *Nelle ipotesi del lemma 2.6.1, la coniugata della funzione $h(y) = \inf_{u \in X} \phi(u, y)$ è data da*

$$h^*(\varphi) = \phi^*(0, \varphi) \qquad \forall \varphi \in Y^*,$$

ove ϕ^* è la coniugata di ϕ, ossia

$$\phi^*(\psi, \varphi) = \sup_{u \in X,\, y \in Y} \{\psi u + \varphi y - \phi(u, y)\} \qquad \forall \psi \in X^*, \quad \forall \varphi \in Y^*.$$

Dimostrazione Si ha

$$
\begin{aligned}
h^*(\varphi) &= \sup_{y\in Y}\{\varphi y - h(y)\} = \sup_{y\in Y}\{\varphi y - \inf_{u\in X}\phi(u,y)\} = \\
&= \sup_{y\in Y}\left\{\varphi y + \sup_{u\in X}\{-\phi(u,y)\}\right\} = \\
&= \sup_{y\in Y}\sup_{u\in X}\{\varphi y - \phi(u,y)\} = \phi^*(0,\varphi). \quad \square
\end{aligned}
$$

Ciò premesso, fissiamo le nostre ipotesi, che varranno per tutto il resto del paragrafo. Sono dati due spazi normati reali X e Y, e una funzione $\phi : X\times Y \to \overline{\mathbb{R}}$ convessa, semicontinua inferiormente e propria, tale che $\phi(u,0) = F(u)$ per ogni $u \in X$. Chiameremo "problema primale" quello che consiste nel trovare il minimo di F su X; diremo invece "problema duale" il problema di massimizzazione seguente:

$$\exists \max_{\varphi\in Y^*}\{-\phi^*(0,\varphi)\}.$$

Osservazione 2.6.3 Per il lemma 2.6.2 e per definizione di bipolare, risulta

$$\sup_{\varphi\in Y^*}\{-\phi^*(0,\varphi)\} = \sup_{\varphi\in Y^*}\{-h^*(\varphi)\} = h^{**}(J_0) \le h(0) = \inf_{u\in X}\phi(u,0),$$

e tale quantità non vale $+\infty$, essendo $F = \phi(\cdot,0)$ propria; potrebbe però valere $-\infty$, come mostra l'esempio di $\phi(u,y) = u - y + e^{u-y}$ in $X \times Y = \mathbb{R}^2$: in questo caso si ha infatti $h(y) \equiv -\infty$ e $h^*(\varphi) \equiv +\infty$.

Andiamo ad analizzare le relazioni fra il problema primale e il problema duale, cercando condizioni che assicurino la risolubilità di entrambi: scopriremo che questo comporterà anche l'uguaglianza $\inf_{u\in X}\phi(u,0) = \sup_{\varphi\in Y^*}\{-\phi^*(0,\varphi)\}$.

Definizione 2.6.4 *Nelle ipotesi precedenti, diciamo che il problema primale è* normale *se* $h(0) = \inf_{u\in X}\phi(u,0) \in \mathbb{R}$ *e se la funzione* h *è semicontinua inferiormente nel punto* $0 \in Y$.

Definizione 2.6.5 *Nelle ipotesi precedenti, diciamo che il problema primale è* stabile *se* $h(0) \in \mathbb{R}$ *e se esiste* $\varphi \in Y^*$ *tale che* $h(y) \ge h(0) + \varphi y$ *per ogni* $y \in Y$.

Come vedremo più avanti nel corso, la condizione di stabilità equivale alla "sottodifferenziabilità" di h nel punto 0.

Proposizione 2.6.6 *Nelle ipotesi precedenti, il problema primale è normale se e solo se*

$$\sup_{\varphi \in Y^*} \{-\phi^*(0,\varphi)\} = \inf_{u \in X} \phi(u,0) \in \mathbb{R}.$$

Dimostrazione Supponiamo che il problema primale sia normale. Sia $\overline{h}$ la regolarizzata semicontinua di h: si ha, per il lemma 2.6.2, per l'osservazione 2.5.9, per la proposizione 2.3.9 (iii) e per ipotesi,

$$\sup_{\varphi \in Y^*} \{-\phi^*(0,\varphi)\} = h^{**}(J_0) = \overline{h}(0) = h(0) = \inf_{u \in X} \phi(u,0) \in \mathbb{R}.$$

Viceversa, se queste quantità coincidono e sono finite, allora $h^{**}(J_0) = \overline{h}(0) = h(0) \in \mathbb{R}$ e quindi il problema primale è normale in virtù della proposizione 2.3.9 (iii). $\square$

Proposizione 2.6.7 *Nelle ipotesi precedenti, il problema primale è stabile se e solo se esso è normale e il problema duale ha soluzione.*

Dimostrazione Se il problema primale è stabile, allora esiste $\varphi_0 \in Y^*$ tale che $h(y) \geq h(0) + \varphi_0 y$ per ogni $y \in Y$; ne segue

$$\liminf_{\|y\|_Y \to 0} h(y) \geq h(0),$$

quindi h è semicontinua inferiormente in 0. Inoltre per ogni $\varphi \in Y^*$ si ha, grazie al lemma 2.6.2,

$$-\phi^*(0,\varphi) = -h^*(\varphi) = J_0\varphi - h^*(\varphi) \leq h^{**}(J_0) = h(0) \leq h(y) - \varphi_0 y \quad \forall y \in Y,$$

da cui

$$-\phi^*(0,\varphi) \leq \inf_{y \in Y} \{h(y) - \varphi_0 y\} = -h^*(\varphi_0) = -\phi^*(0,\varphi_0) \qquad \forall \varphi \in Y^*.$$

Ciò prova che

$$\exists \max_{\varphi \in Y^*} \{-\phi^*(0,\varphi)\} = -\phi^*(0,\varphi_0).$$

Viceversa, se il problema primale è normale e il problema duale ha soluzione φ_0, allora

$$-h^*(\varphi_0) \geq -h^*(\varphi) \qquad \forall \varphi \in Y^*,$$

da cui

$$-h^*(\varphi_0) = \sup_{\varphi \in Y^*} \{-h^*(\varphi)\} = h^{**}(J_0) = h(0),$$

e dunque

$$h(0) = -h^*(\varphi_0) \le -\varphi_0 y + h(y) \qquad \forall y \in Y:$$

ciò mostra che il problema primale è stabile. $\square$

Osservazione 2.6.8 Di fatto, l'ipotesi che il problema primale sia normale è irrinunciabile se vogliamo che esso abbia minimo finito; con l'ipotesi ulteriore che il problema primale sia stabile, otteniamo la risolubilità del problema duale, e l'unica cosa che resta da fare è provare l'esistenza del minimo per il problema primale. L'ipotesi di stabilità è dunque naturale.

Un semplice criterio di stabilità per il problema primale è il seguente:

Lemma 2.6.9 *Oltre alle ipotesi precedenti, supponiamo che $h(0) > -\infty$ e che esista $u_0 \in X$ tale che la funzione $y \mapsto \phi(u_0, y)$ sia continua in 0, con $\phi(u_0, 0) \in \mathbb{R}$. Allora il problema primale è stabile.*

Dimostrazione Poiché $\phi(u_0, \cdot)$ è continua e finita in 0, essa è limitata in un opportuno intorno aperto V di 0 da una costante K. Dunque

$$h(y) = \inf_{u \in X} \phi(u, y) \le \phi(u_0, y) \le K \qquad \forall y \in V,$$

per cui h, essendo convessa e limitata superiormente in un intorno di 0, è continua e finita in tale punto (lemma 2.4.1).
Adesso notiamo che $(0, h(0))$ appartiene alla frontiera di epi(h) in $Y \times \mathbb{R}$, e che epi(h) ha parte interna non vuota poiché contiene l'aperto $V \times]K, +\infty[$.
Per il lemma 2.3.6 (i), esistono $\psi \in Y^*$ e $\lambda > 0$ tali che

$$\psi\, 0 + \lambda h(0) \le \psi y + \lambda t \qquad \forall (y, t) \in \mathrm{epi}(h),$$

e in particolare, scelti $y \in D(h)$ e $t > h(y)$, e posto $\varphi_0 = -\frac{\psi}{\lambda}$,

$$h(0) \le -\varphi_0 y + t \qquad \forall y \in D(h), \quad \forall t > h(y).$$

Per l'arbitrarietà di $t > h(y)$ e per il fatto che $h(0) \in \mathbb{R}$, deduciamo che h è propria e che

$$h(0) + \varphi_0 y \le h(y) \qquad \forall y \in Y.$$

Ciò mostra che il problema primale è stabile. $\square$

Veniamo ora alla questione dell'esistenza del minimo per il problema primale, e soprattutto alla sua caratterizzazione mediante opportune condizioni di estremalità.

Teorema 2.6.10 *Nelle ipotesi del lemma 2.6.9, se $\hat{u} \in X$ è punto di minimo per il problema primale e $\hat{\varphi} \in Y^*$ è punto di massimo per il problema duale, allora $\hat{u}$ e $\hat{\varphi}$ verificano la* relazione di estremalità

$$\phi(\hat{u}, 0) + \phi^*(0, \hat{\varphi}) = 0.$$

Viceversa, se $\hat{u} \in X$ e $\hat{\varphi} \in Y^*$ soddisfano tale relazione, allora

$$\phi(\hat{u}, 0) = \min_{u \in X} \phi(u, 0) = \max_{\varphi \in Y^*}\{-\phi^*(0, \varphi)\} = -\phi^*(0, \hat{\varphi}) \in \mathbb{R}.$$

Dimostrazione Per il lemma 2.6.9 il problema primale è stabile: quindi, se $\hat{u}$ minimizza il problema primale e $\hat{\varphi}$ massimizza il problema duale, allora

$$\phi(\hat{u}, 0) = \min_{u \in X} \phi(u, 0) = \max_{\varphi \in Y^*}\{-\phi^*(0, \varphi)\} = -\phi^*(0, \hat{\varphi}) \in \mathbb{R},$$

cosicché $\phi(\hat{u}, 0) + \phi^*(0, \hat{\varphi}) = 0$.
Viceversa, dalla stabilità del problema primale segue

$$\inf_{u \in X} \phi(u, 0) = \max_{\varphi \in Y^*}\{-\phi^*(0, \varphi)\} \in \mathbb{R}.$$

Dunque, essendo

$$\phi(\hat{u}, 0) \geq \inf_{u \in X} \phi(u, 0) = \max_{\varphi \in Y^*}\{-\phi^*(0, \varphi)\} \geq -\phi^*(0, \hat{\varphi}) = \phi(\hat{u}, 0),$$

si ha immediatamente la tesi. □

Corollario 2.6.11 *Oltre alle ipotesi del lemma 2.6.9, supponiamo che:*

(i) *X sia uno spazio di Banach riflessivo;*

(ii) $\lim_{\|u\|_X \to \infty} \phi(u, 0) = +\infty$.

Allora il problema primale ha minimo, il problema duale ha massimo e vale la relazione di estremalità.

Dimostrazione Basta osservare che, per il corollario 2.3.4 ed il lemma 2.6.9 il problema primale ha soluzione ed è stabile; per la proposizione 2.6.7 anche il problema duale ha soluzione. La tesi segue allora dal teorema precedente. □

Vi è un altro modo di caratterizzare i punti di minimo del problema primale e i punti di massimo del problema duale: essi si ottengono come "punti di sella" di un'opportuna funzione: la "lagrangiana".

Definizione 2.6.12 *La* funzione lagrangiana *del problema primale è la funzione $L : X \times Y^* \to \overline{\mathbb{R}}$ così definita:*

$$L(u, \varphi) = \inf_{y \in Y} \{\phi(u, y) - \varphi y\} \qquad \forall u \in X, \quad \forall \varphi \in Y^*.$$

In altre parole, per ogni fissato $u \in X$, $L(u, \varphi)$ è data da

$$L(u, \varphi) = \inf_{y \in Y} \{\phi(u, y) - \varphi y\} = -\sup_{y \in Y} \{\varphi y - \phi(u, y)\} = -[\phi(u, \cdot)]^*(\varphi),$$

cioè è l'opposta della coniugata di $\phi(u, \cdot)$.

Lemma 2.6.13 *Nelle ipotesi del lemma 2.6.1, si ha:*

(i) *$\varphi \mapsto L(u, \varphi)$ è concava e semicontinua superiormente in Y^* per ogni $u \in X$;*

(ii) *$u \mapsto L(u, \varphi)$ è convessa in X per ogni $\varphi \in Y^*$.*

Dimostrazione (i) Per $u \in X$ fissato, $[\phi(u, \cdot)]^*$ è convessa e semicontinua inferiormente, quindi $L(u, \cdot) = -[\phi(u, \cdot)]^*$ è concava e semicontinua superiormente.

(ii) Fissato $\varphi \in Y^*$, $L(\cdot, \varphi)$ è l'estremo inferiore delle funzioni convesse $(u, y) \mapsto \phi(u, y) - \varphi y$; quindi la convessità di $L(\cdot, \varphi)$ segue con lo stesso argomento usato nella dimostrazione del lemma 2.6.1. $\square$

Tramite la funzione lagrangiana, i problemi primale e duale si possono riformulare come problemi di "minimax": si tratta di un punto di vista che ritroveremo più avanti nel corso, parlando della teoria dei giochi.
Scrivendo

$$\begin{aligned}
\phi^*(\psi, \varphi) &= \sup_{u \in X,\ y \in Y} \{\psi u + \varphi y - \phi(u, y)\} = \\
&= \sup \left\{ \psi u + \sup_{y \in Y} \{\varphi y - \phi(u, y)\} \right\} = \sup_{u \in X} \{\psi u - L(u, \varphi)\},
\end{aligned}$$

ricaviamo

$$-\phi^*(0, \varphi) = -\sup_{u \in X} \{-L(u, \varphi)\} = \inf_{u \in X} L(u, \varphi),$$

e quindi il problema duale si scrive nella forma "maximin"

$$\sup_{\varphi \in Y^*} \inf_{u \in X} L(u, \varphi).$$

Similmente, poiché ϕ è convessa e semicontinua inferiormente,

$$\begin{aligned}\phi(u,y) &= \left([\phi(u,\cdot)]^{**} \circ J\right)(y) = \\ &= \sup_{\varphi \in Y^*} \{\varphi y - [\phi(u,\cdot)]^*(\varphi)\} = \sup_{\varphi \in Y^*} \{\varphi y + L(u,\varphi)\},\end{aligned}$$

da cui

$$\phi(u,0) = \sup_{\varphi \in Y^*} L(u,\varphi).$$

Pertanto il problema primale si scrive nella forma "minimax"

$$\inf_{u \in X} \sup_{\varphi \in Y^*} L(u,\varphi).$$

Osservazione 2.6.14 È immediato verificare che

$$\sup_{\varphi \in Y^*} \inf_{u \in X} L(u,\varphi) \leq \inf_{u \in X} \sup_{\varphi \in Y^*} L(u,\varphi) \qquad \forall u \in X, \quad \forall \varphi \in Y^* :$$

infatti

$$L(u,\psi) \leq \sup_{\varphi \in Y^*} L(u,\varphi) \qquad \forall u \in X, \quad \forall \psi \in Y^*,$$

da cui

$$\inf_{u \in X} L(u,\psi) \leq \inf_{u \in X} \sup_{\varphi \in Y^*} L(u,\varphi) \qquad \forall \psi \in Y^*$$

e quindi la tesi. Ritroviamo così il risultato dell'osservazione 2.6.3.

Definizione 2.6.15 *Diciamo che* $(\hat{u},\hat{\varphi}) \in X \times Y^*$ *è* punto di sella *per la funzione lagrangiana* L *se risulta*

$$L(\hat{u},\varphi) \leq L(\hat{u},\hat{\varphi}) \leq L(u,\hat{\varphi}) \qquad \forall u \in X, \quad \forall \varphi \in Y^*.$$

Teorema 2.6.16 *Nelle ipotesi del lemma 2.6.9, sia* $(\hat{u},\hat{\varphi}) \in X \times Y^*$*. Sono fatti equivalenti:*

(i) $(\hat{u},\hat{\varphi})$ *è punto di sella per la lagrangiana* L*,*

(ii) $\hat{u}$ *minimizza il problema primale,* $\hat{\varphi}$ *massimizza il problema duale e*

$$\min_{u \in X} \phi(u,0) = \max_{\varphi \in Y^*} \{-\phi^*(0,\varphi)\} \in \mathbb{R}.$$

Dimostrazione Ricordiamo anzitutto che, per i calcoli fatti in precedenza,

$$\inf_{u\in X} L(u,\hat\varphi) = -\phi^*(0,\hat\varphi), \qquad \sup_{\varphi\in Y^*} L(\hat u,\varphi) = \phi(\hat u,0).$$

Supponiamo che valga (i): allora si ha

$$-\phi^*(0,\hat\varphi) = \min_{u\in X} L(u,\hat\varphi) = L(\hat u,\hat\varphi) = \max_{\varphi\in Y^*} L(\hat u,\varphi) = \phi(\hat u,0) \in \mathbb{R},$$

da cui segue subito la relazione di estremalità $\phi(\hat u,0)+\phi^*(0,\hat\varphi) = 0$; dal teorema 2.6.10 segue (ii).

Supponiamo viceversa che valga (ii): essendo

$$-\phi^*(0,\hat\varphi) = \inf_{u\in X} L(u,\hat\varphi) \le L(\hat u,\hat\varphi), \quad \phi(\hat u,0) = \sup_{\varphi\in Y^*} L(\hat u,\varphi) \ge L(\hat u,\hat\varphi),$$

la relazione di estremalità, che vale in virtù del teorema 2.6.10, implica

$$L(\hat u,\hat\varphi) = \inf_{u\in X} L(u,\hat\varphi) = \sup_{\varphi\in Y^*} L(\hat u,\varphi),$$

cosicché $(\hat u,\hat\varphi)$ è punto di sella per L. $\square$

Corollario 2.6.17 *Nelle ipotesi del lemma 2.6.9, sia $\hat u \in X$. Allora $\hat u$ minimizza il problema primale se e solo se esiste $\hat\varphi \in Y^*$ tale che $(\hat u,\hat\varphi)$ sia punto di sella per la lagrangiana L.*

Dimostrazione Se tale $\hat\varphi$ esiste, la tesi segue dal teorema 2.6.16. Viceversa, se $\hat u$ minimizza il problema primale, poiché esso è stabile il problema duale ha una soluzione $\hat\varphi \in Y^*$ e si ha

$$\inf_{u\in X} \phi(u,0) = \sup_{\varphi\in Y^*} \{-\phi^*(0,\varphi)\} \in \mathbb{R}.$$

Dunque $(\hat u,\hat\varphi)$ è punto di sella per L in virtù del teorema 2.6.16. $\square$

Esempio 2.6.18 Consideriamo un caso particolare della situazione considerata fin qui: prendiamo la funzione F della forma

$$F(u) = f(u) + g(\Lambda u), \qquad u \in X,$$

ove Λ è un fissato operatore lineare e continuo da X in Y, mentre $f : X \to \overline{\mathbb{R}}$ e $g : Y \to \overline{\mathbb{R}}$ sono funzioni convesse, semicontinue inferiormente e proprie. Faremo le ipotesi seguenti:

(a) X è uno spazio di Banach riflessivo e Y è uno spazio normato (entrambi reali);

(b) $\lim_{\|u\|_X \to \infty}[f(u) + g(\Lambda u)] = +\infty$,

(c) esiste un punto $u_0 \in X$ tale che i numeri $f(u_0)$ e $g(\Lambda u_0)$ siano entrambi finiti e che g sia continua nel punto Λu_0.

Allora, scegliendo

$$\phi(u, y) = f(u) + g(\Lambda u - y) \qquad \forall (u, y) \in X \times Y,$$

per ogni $\varphi \in Y^*$ si ha, denotando con $\Lambda^* \in \mathcal{L}(Y^*, X^*)$ l'operatore aggiunto di Λ,

$$\begin{aligned}
\phi^*(0, \varphi) &= \sup_{u\in X} \sup_{y\in Y} \{\varphi y - f(u) - g(\Lambda u - y)\} = \qquad [\text{posto } \Lambda u - y = p] \\
&= \sup_{u\in X} \sup_{p\in Y} \{\varphi(\Lambda u - p) - f(u) - g(p)\} = \\
&= \sup_{u\in X} \sup_{p\in Y} \{(\Lambda^*\varphi)u - \varphi p - f(u) - g(p)\} = \\
&= \sup_{u\in X} \{(\Lambda^*\varphi)u - f(u)\} + \sup_{p\in Y} \{-\varphi p - g(p)\} = \\
&= f^*(\Lambda^*\varphi) + g^*(-\varphi).
\end{aligned}$$

Le ipotesi fatte su f e g garantiscono che sono soddisfatte le ipotesi del corollario 2.6.11. La condizione di estremalità diventa

$$f(\hat{u}) + g(\Lambda\hat{u}) + f^*(\Lambda^*\hat{\varphi}) + g^*(-\hat{\varphi}) = 0;$$

ma poiché, per la disuguaglianza di Young,

$$f(\hat{u}) + f^*(\Lambda^*\hat{\varphi}) \geq (\Lambda^*\hat{\varphi})\hat{u}, \qquad g(\Lambda\hat{u}) + g^*(-\hat{\varphi}) \geq -\hat{\varphi}(\Lambda\hat{u}) = -(\Lambda^*\hat{\varphi})\hat{u},$$

la relazione di estremalità si scinde nel sistema seguente:

$$f(\hat{u}) + f^*(\Lambda^*\hat{\varphi}) = (\Lambda^*\hat{\varphi})\hat{u} = -g(\Lambda\hat{u}) - g^*(-\hat{\varphi}).$$

Le soluzioni $\hat{u}, \hat{\varphi}$ di questo sistema risolvono i problemi primale e duale:

$$\begin{aligned}
f(\hat{u}) + g(\Lambda\hat{u}) &= \min_{u\in X} \{f(u) + g(\Lambda u)\} = \\
&= \max_{\varphi\in Y^*} \{-f^*(\Lambda^*\varphi) - g^*(-\varphi)\} = -f^*(\Lambda^*\hat{\varphi}) - g^*(-\hat{\varphi}).
\end{aligned}$$

La funzione lagrangiana è data da

$$\begin{aligned}
L(u,\varphi) &= \inf_{y\in Y}\{f(u)+g(\Lambda u-y)-\varphi y\} = \\
&= \inf_{p\in Y}\{f(u)+g(p)-\varphi(\Lambda u)+\varphi p\} = \\
&= f(u)-(\Lambda^*\varphi)u+\inf_{p\in Y}\{\varphi p+g(p)\} = \\
&= f(u)-(\Lambda^*\varphi)u-\sup_{p\in Y}\{-\varphi p-g(p)\} = f(u)-(\Lambda^*\varphi)u-g^*(-\varphi),
\end{aligned}$$

e il punto $(\hat{u},\hat{\varphi})$ è di sella per L. Questo ce lo dice il teorema 2.6.16, ma si può vedere direttamente: dal sistema che esprime le relazioni di estremalità, utilizzando la disuguaglianza di Young, segue infatti

$$\begin{aligned}
L(\hat{u},\hat{\varphi}) &= f(\hat{u})-(\Lambda^*\hat{\varphi})\hat{u}-g^*(-\hat{\varphi}) = -f^*(\Lambda^*\hat{\varphi})-g^*(-\hat{\varphi}) \le \\
&\le f(u)-(\Lambda^*\hat{\varphi})u-g^*(-\hat{\varphi}) = L(u,\hat{\varphi}) \qquad \forall u\in X,
\end{aligned}$$

e similmente

$$\begin{aligned}
L(\hat{u},\hat{\varphi}) &= f(\hat{u})-(\Lambda^*\hat{\varphi})\hat{u}-g^*(-\hat{\varphi}) = f(\hat{u})+g(\Lambda\hat{u}) \ge \\
&\ge f(\hat{u})-\varphi(\Lambda\hat{u})-g^*(-\varphi) = L(\hat{u},\varphi) \qquad \forall\varphi\in Y^*.
\end{aligned}$$

Esempio 2.6.19 Siano X,Y spazi di Banach reali con X riflessivo. Sia C un *cono* di Y (ciò significa che $C+C\subseteq C$ e $tC\subseteq C$ per ogni $t\ge 0$) che supponiamo non vuoto, convesso, chiuso e tale che $C\cap(-C)=\{0\}$. Esso induce una relazione d'ordine parziale $\le_Y$ su Y, definita per $y,z\in Y$ da

$$y\le_Y z \qquad \Longleftrightarrow \qquad z-y\in C$$

(le verifiche sono ovvie). Per costruzione risulta $C=\{y\in Y:0\le_Y y\}$, mentre $-C=\{y\in Y:y\le_Y 0\}$. Il *cono duale* è il cono di Y^* definito da

$$C^*=\{\varphi\in Y^*:\varphi y\ge 0 \ \ \forall y\in C\},$$

e scriveremo $\varphi\ge_{Y^*}0$ se e solo se $\varphi\in C^*$. Si verifica facilmente (esercizio 2.6.2) che

$$C=\{y\in Y:\varphi y\ge 0 \ \ \forall\varphi\in C^*\}.$$

Sia K un convesso chiuso non vuoto di X e sia $J:X\to\overline{\mathbb{R}}$ una funzione convessa, semicontinua inferiormente e propria, tale che $D(J)=K$. Sia infine

$B : X \to Y$ un operatore non necessariamente lineare, convesso rispetto alla relazione $\leq_Y$:

$$B(\lambda y + (1-\lambda)z) \leq_Y \lambda B(y) + (1-\lambda)B(z) \qquad \forall y, z \in X, \quad \forall \lambda \in [0,1].$$

Supponiamo che valgano le seguenti condizioni:

(a) per ogni $\varphi \in C^*$ l'applicazione $u \mapsto \varphi(B(u))$ è semicontinua inferiormente in K;

(b) l'insieme $\left\{x \in K : -B(x) \in \overset{\circ}{C}\right\}$ è non vuoto;

(c) risulta $\inf\{J(u) : u \in K,\ B(u) \leq_Y 0\} \in \mathbb{R}$;

(d) si ha $\displaystyle\lim_{\|u\|_X \to \infty,\ B(u) \leq_Y 0} J(u) = +\infty$.

Il nostro problema primale sarà

$$\exists \min_{u \in K,\ B(u) \leq_Y 0} J(u),$$

e sceglieremo la funzione ϕ così definita:

$$\phi(u, y) = J(u) + I_C(y - B(u)) = \begin{cases} J(u) & \text{se } u \in K \text{ e } B(u) \leq_Y y \\ +\infty & \text{altrimenti.} \end{cases}$$

Occorre verificare che le ipotesi del corollario 2.6.11 sono soddisfatte. È immediato verificare, grazie a (c), che ϕ e $\phi(\cdot, 0)$ sono proprie. Proviamo che ϕ è convessa e semicontinua inferiormente: posto

$$E = \{(u, y) \in X \times Y :\ u \in K,\ B(u) \leq_Y y\},$$

si può scrivere

$$\phi(u, y) = J(u) + I_E(u, y),$$

e quindi, ricordando che J è convessa e semicontinua inferiormente, basta far vedere che E è un convesso chiuso in $X \times Y$. Dato che

$$E = \bigcap_{\varphi \in C^*} \{(u, y) \in X \times Y :\ u \in K,\ \varphi(B(u) - y) \leq 0\},$$

in virtù di (a) l'insieme E è intersezione di convessi chiusi e quindi è un convesso chiuso.

Infine, la condizione di stabilità del problema primale segue da (b): infatti, scelto $u_0 \in K$ tale che $-Bu_0 \in \overset{\circ}{C}$, esiste un intorno V di $0 \in Y$ tale che $B(u_0) - y \leq_Y 0$ per ogni $y \in V$: ne segue $\phi(u_0, y) = J(u_0)$ per ogni $y \in V$, cosicché $y \mapsto \phi(u_0, y)$ è finita e continua nel punto 0. Notiamo anche che (d) implica l'esistenza del minimo per il problema primale.
Si ha allora, procedendo come nell'esempio 2.6.18,

$$\begin{aligned} \phi^*(0,\varphi) &= \sup_{u\in X}\sup_{y\in Y}\{\varphi y - J(u) - I_C(y - B(u))\} = \qquad [y - B(u) = p] \\ &= \sup_{u\in K}\sup_{p\in Y}\{\varphi(B(u)) + \varphi p - J(u) - I_C(p)\} = \\ &= \sup_{u\in K}\{\varphi(B(u)) - J(u)\} + I_C^*(\varphi), \end{aligned}$$

e ricordando l'esempio 2.5.6 (4),

$$\phi^*(0,\varphi) = \sup_{u\in K}\{\varphi(B(u)) - J(u)\} + \sup_{0\leq y}\{\varphi y\}.$$

Notando poi che, come si verifica facilmente,

$$I_C^*(\varphi) = \sup_{0\leq_Y y}\{\varphi y\} = I_{C^*}(-\varphi) \qquad \forall \varphi \in Y^*,$$

si conclude che

$$\phi^*(0,\varphi) = \sup_{u\in K}\{\varphi(B(u)) - J(u)\} + I_{C^*}(-\varphi).$$

Quindi il problema duale è

$$\exists \max_{\varphi\in Y^*}\left\{\inf_{u\in K}\{J(u) - \varphi(B(u))\} - I_{C^*}(-\varphi)\right\} = \max_{\varphi\in C^*}\{\inf_{u\in K}\{J(u) - \varphi(B(u))\}.$$

La condizione di estremalità è

$$J(\hat{u}) + I_{-C}(B(\hat{u})) + \sup_{u\in K}\{\hat{\varphi}(B(u)) - J(u)\} + I_{C^*}(-\hat{\varphi}) = 0;$$

essa implica $\hat{u} \in K$, $B(\hat{u}) \leq_Y 0$, $-\hat{\varphi} \in C^*$ (da cui $\hat{\varphi}(B(\hat{u})) \geq 0$) e

$$J(\hat{u}) + \sup_{u\in K}\{\hat{\varphi}(B(u)) - J(u)\} = 0.$$

Poiché, per definizione,

$$J(\hat{u}) + \sup_{u \in K}\{\hat{\varphi}(B(u)) - J(u)\} \geq \hat{\varphi}(B(\hat{u})) \geq 0,$$

la relazione di estremalità implica le semplici condizioni

$$\hat{u} \in K, \quad B(\hat{u}) \leq_Y 0, \quad \hat{\varphi} \leq_{Y^*} 0; \qquad \hat{\varphi}(B(\hat{u})) = 0,$$

le quali tuttavia non sono equivalenti alla relazione di estremalità: ad esempio esse sono soddisfatte per $\hat{\varphi} = 0$, ma non è detto che il funzionale $0 \in Y^*$ massimizzi $\varphi \mapsto -\phi^*(0, \varphi)$.
La lagrangiana è data da

$$\begin{aligned} L(u, \varphi) &= \inf_{y \in Y}\{J(u) + I_C(y - B(u)) - \varphi y\} = \\ &= \inf_{p \in Y}\{J(u) + I_C(p) - \varphi(B(u)) - \varphi p\} = \\ &= J(u) - \varphi(B(u)) + \inf_{p \in Y}\{I_C(-p) - \varphi p\} = \\ &= J(u) - \varphi(B(u)) - \sup_{p \in Y}\{\varphi p - I_C(p)\} = \\ &= J(u) - \varphi(B(u)) - I_C^*(\varphi). \end{aligned}$$

La relazione di estremalità, come sappiamo, equivale al fatto che $(\hat{u}, \hat{\varphi})$ è punto di sella per L, cosa che, volendo, si può anche verificare direttamente senza troppa fatica.

Esercizi 2.6

1. Sia X uno spazio normato, e sia $K \subseteq X$ un convesso chiuso non vuoto. Si provi che

$$d(x_0, K) = \max_{\|\varphi\|_{X^*}=1}\{\varphi x_0 - I_K^*(\varphi)\} \qquad \forall x_0 \in K.$$

2. Sia Y uno spazio di Banach, sia C un cono di Y che supponiamo non vuoto, convesso, chiuso e tale che $C \cap (-C) = \{0\}$. Detto C^* il cono duale di Y^* rispetto alla relazione d'ordine $\leq_Y$ indotta da C, si verifichi che C coincide con $\{y \in Y : \varphi y \geq 0 \;\; \forall \varphi \in C^*\}$.
[**Traccia:** per la seconda inclusione si ragioni per assurdo utilizzando il teorema di Hahn-Banach.]

3. *(Teorema di Karush-Kuhn-Tucker)* Per ogni $k \in \mathbb{N}^+$ si introduca in $\mathbb{R}^k$ l'ordinamento $\leq_k$ indotto dal cono $\Gamma_k = \{x \in \mathbb{R}^k : \ x_i \geq 0, i = 1, \ldots, k\}$. Sia poi $A = \{a_{ij}\}$ una matrice reale $m \times n$. Fissati $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, si consideri il problema primale

$$\exists \min_{u \in V} \langle c, u \rangle_n \,, \quad V = \{u \in \mathbb{R}^n : \ u \geq_n 0, \ Au \leq_m b\}.$$

(i) Posto $\phi(u, y) = \langle c, u \rangle_n + I_{\Gamma_n}(u) + I_{\Gamma_m}(b + y - Au)$, si provi che il problema duale è

$$\exists \max_{y \in W} \langle b, y \rangle_m \,, \quad W = \{y \in \mathbb{R}^m : \ y \leq_m 0, \ {}^t A y \leq_n c\}.$$

(ii) Si scriva la lagrangiana e si determini la relazione di estremalità.

(iii) Supponendo inoltre che $c_i > 0$ per $i = 1, \ldots, n$ e che l'insieme $\{u \in \Gamma_n : \ (Au)_i < b_i, \, i = 1, \ldots, m\}$ sia non vuoto, si verifichi che esiste un punto di minimo $\hat{u}$ per il problema primale e che esiste un punto di massimo $\hat{y}$ per il problema duale; se ne deduca che per ogni $i = 1, \ldots, m$ vale la seguente alternativa:

$$(A\hat{u})_i < b_i \ \text{ e } \ \hat{y}_i = 0 \qquad \text{oppure} \qquad (A\hat{u})_i = b_i \ \text{ e } \ \hat{y}_i \leq 0.$$

Capitolo 3

Calcolo in spazi di Banach

3.1 Integrale di Bochner

Nel capitolo precedente abbiamo visto, sotto opportune ipotesi, alcuni risultati di esistenza del minimo di funzioni convesse in spazi di Banach. Ma per trovare effettivamente i punti di minimo, occorre saper "fare le derivate" di questo tipo di funzioni. Questo capitolo è dedicato all'estensione delle usuali proprietà del calcolo differenziale al caso di funzioni tra spazi normati.
Preliminare a questo studio, però, è una breve descrizione della teoria dell'integrazione per funzioni a valori in uno spazio di Banach.
Siano dunque X uno spazio di Banach (reale o complesso) ed $(E, \mathcal{E}, \mu)$ uno spazio misurato: denoteremo con $\|\cdot\|$ la norma di X e considereremo funzioni $f : E \to X$.

Definizione 3.1.1 *Indichiamo con $\mathcal{S}(E, X)$ lo spazio vettoriale delle* funzioni semplici*, ossia delle funzioni $\varphi : E \to X$ tali che:*

(i) *φ assume un numero finito di valori $x_1, \ldots, x_k \in X$;*

(ii) *per ogni $i = 1, \ldots, k$ gli insiemi $A_i = \{u \in E : \varphi(u) = x_i\}$ appartengono alla σ-algebra $\mathcal{E}$ ed inoltre $\mu(A_i) < \infty$.*

La forma canonica *della funzione φ è dunque la seguente:*

$$\varphi(u) = \sum_{i=1}^{k} x_i \, \chi_{A_i}(u), \qquad u \in E.$$

Definizione 3.1.2 *Se $\varphi \in \mathcal{S}(E, X)$, con forma canonica $\varphi = \sum_{i=1}^k x_i \chi_{A_i}$, l'* integrale *di φ su E è l'elemento di X definito da*

$$\int_E \varphi \, d\mu = \sum_{i=1}^k x_i \, \mu(A_i).$$

Osservazione 3.1.3 L'integrale su $\mathcal{S}(E, X)$ gode delle usuali proprietà: in particolare è lineare, e in luogo della monotonia (che non ha senso nel generico spazio di Banach X) si ha la disuguaglianza

$$\left\| \int_E \varphi \, d\mu \right\| \le \int_E \|\varphi(\cdot)\| \, d\mu \qquad \forall \varphi \in \mathcal{S}(E, X).$$

Si noti che $\|\varphi(\cdot)\|$ è una funzione semplice su E, a valori reali.

Passiamo ad introdurre le funzioni misurabili definite su E, a valori in X.

Definizione 3.1.4 *Una funzione $f : E \to X$ è detta* fortemente misurabile *se esiste una successione $\{\varphi_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}(E, X)$ tale che*

$$\varphi_n(u) \to f(u) \quad \forall u \in E.$$

La funzione f è detta debolmente misurabile *se per ogni $T \in X^*$ la funzione reale o complessa $u \to Tf(u)$ è misurabile su E.*

È facile verificare che ogni funzione fortemente misurabile è debolmente misurabile; il viceversa è falso, come mostra l'esercizio 3.1.1. Inoltre si vede immediatamente che se f è fortemente misurabile allora la funzione reale $u \to \|f(u)\|$ è misurabile su E.
È importante la seguente proprietà delle funzioni fortemente misurabili:

Proposizione 3.1.5 *Sia $f : E \to X$ una funzione fortemente misurabile; allora per ogni aperto $A \subseteq X$ la controimmagine $f^{-1}(A)$ appartiene ad $\mathcal{E}$.*

Dimostrazione Sia $\{\varphi_n\}$ una successione di funzioni semplici che converge a f puntualmente in E. Allora, se A è un aperto di X, posto $A_k = \{x \in A : d(x, \partial A) > 1/k\}$ possiamo scrivere

$$f^{-1}(A) = \bigcup_{k \in \mathbb{N}^+} \bigcup_{n \in \mathbb{N}} \bigcap_{m=n}^{\infty} \varphi_m^{-1}(A_k);$$

ora, detta $\varphi_m(t) = \sum_{j=1}^{h_m} x_{j,m}\, \chi_{B_{j,m}}$ la forma canonica di φ_m, gli insiemi $B_{j,m}$ sono elementi disgiunti di $\mathcal{E}$, e dunque avremo

$$\varphi_m^{-1}(A_k) = \bigcup_{j \in S_k} B_{j,m}$$

ove S_k è l'insieme degli indici $j \in \{1, ..., h_m\}$ per i quali si ha $x_j \in A_k$. Quindi $\varphi_m^{-1}(A_k)$ è un elemento di $\mathcal{E}$, e dunque la relazione precedente mostra che $f^{-1}(A)$ appartiene ad $\mathcal{E}$. $\square$

Osservazione 3.1.6 Anche il viceversa della proposizione precedente è falso: esistono funzioni che non sono fortemente misurabili, per le quali tuttavia si ha $f^{-1}(A) \in \mathcal{E}$ per ogni aperto $A \subseteq X$: si veda l'esercizio 3.1.2.

Vediamo ora come e quando è possibile definire l'integrale di una funzione fortemente misurabile, il quale sarà un elemento di X.

Definizione 3.1.7 *Sia $f : E \to X$ una funzione fortemente misurabile. Diciamo che f è* sommabile *su E se si ha*

$$\int_E \|f(\cdot)\|\, d\mu < +\infty.$$

Come sappiamo, questo integrale ha senso perché $\|f(\cdot)\|$ è una funzione misurabile non negativa. Osserviamo che, ovviamente, ogni funzione semplice è sommabile; naturalmente, bisogna verificare che per le funzioni semplici la nuova definizione di integrale coincide con la vecchia, ma questo è faclie, usando la forma canonica di una funzione semplice. Stabiliamo adesso la seguente

Proposizione 3.1.8 *Sia $f : E \to X$ fortemente misurabile. I seguenti fatti sono equivalenti:*

(i) *f è sommabile;*

(ii) *esiste $\{\psi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{S}(E, X)$ tale che $\int_E \|f(\cdot) - \psi_n(\cdot)\| d\mu \to 0$ per $n \to \infty$.*

Dimostrazione (ii)$\Longrightarrow$ (i) Per ipotesi esiste $\nu \in \mathbb{N}$ tale che

$$\int_E \|f(\cdot) - \psi_n(\cdot)\| d\mu < 1 \qquad \forall n \geq \nu;$$

quindi

$$\int_E \|f(\cdot)\|d\mu \leq \int_E \|f(\cdot)-\psi_\nu(\cdot)\|d\mu + \int_E \|\psi_\nu(\cdot)\|d\mu \leq 1 + \int_E \|\psi_\nu(\cdot)\|d\mu < \infty.$$

Pertanto f è sommabile.

(i) $\Longrightarrow$ **(ii)** Poiché f è fortemente misurabile, esiste $\{\varphi_n\}_{n\in\mathbb{N}^+} \subseteq \mathcal{S}(E,X)$ tale che $\varphi_n(u) \to f(u)$ in X per ogni $u \in E$. Definiamo

$$\psi_n(u) = \begin{cases} 0 & \text{se } \varphi_n(u) = 0, \\ \frac{\varphi_n(u)}{\|\varphi_n(u)\|}\ h_n(u) & \text{se } \varphi_n(u) \neq 0, \end{cases}$$

ove la funzione $h_n : E \to \mathbb{R}$ è data da

$$h_n(u) = \begin{cases} n & \text{se } \|f(u)\| \geq n \\ \frac{k-1}{2^n} & \text{se } \frac{k-1}{2^n} \leq \|f(u)\| < \frac{k}{2^n} \quad (k = 1, 2, \ldots, n2^n). \end{cases}$$

Si verifica facilmente che $\{\psi_n\}_{n\in\mathbb{N}^+} \subseteq \mathcal{S}(E,X)$ e che $\|\psi_n(u)\| \leq \|f(u)\|$ per ogni $u \in E$; proviamo che si ha

$$\psi_n(u) \to f(u) \quad \text{in } X \qquad \forall u \in E.$$

In effetti, se $f(u) = 0$ si ha $\psi_n(u) = 0$ per definizione. Se invece $f(u) \neq 0$, esisterà $\nu \in \mathbb{N}^+$ tale che $\|f(u)\| \geq 2^{-\nu}$: quindi per $n \geq \nu$ risulta $\varphi_n(u) \neq 0$ e dalla definizione di $\psi_n(u)$ segue

$$\lim_{n\to\infty} \psi_n(u) = \lim_{n\to\infty} \frac{\varphi_n(u)}{\|\varphi_n(u)\|}\ h_n(u) = \frac{f(u)}{\|f(u)\|}\|f(u)\| = f(u) \qquad \forall u \in E.$$

Poiché inoltre $\|\psi_n(u) - f(u)\| \leq 2\|f(u)\|$, dato che f è sommabile la convergenza è dominata. Ne segue la tesi. □

Siamo ora in grado di definire l'integrale di una funzione sommabile $f : E \to X$. Infatti, per la proposizione 3.1.8 esiste $\{\psi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{S}(E,X)$ tale che $\lim_{n\to\infty} \int_E \|f(\cdot) - \psi_n(\cdot)\|\, d\mu = 0$. Ne segue che la successione $\left\{\int_E \psi_n\, d\mu\right\}_{n\in N}$ è di Cauchy in X, in quanto

$$\left\|\int_E \psi_n\, d\mu - \int_E \psi_m\, d\mu\right\| \leq \int_E \|\psi_n(\cdot) - \psi_m(\cdot)\|\, d\mu \leq$$
$$\leq \int_E \|\psi_n(\cdot) - f(\cdot)\|\, d\mu + \int_E \|f(\cdot) - \psi_m(\cdot)\|\, d\mu \to 0 \quad \text{per } n, m \to \infty.$$

Essendo X completo, tale successione converge in X ed il suo limite, per definizione, sarà l'integrale su E della funzione f. In altre parole, definiamo l'integrale di una funzione sommabile nel modo seguente:

Definizione 3.1.9 *Sia $(E, \mathcal{E}, \mu)$ uno spazio misurato, sia X uno spazio di Banach e sia $f : E \to X$ sommabile. L'* integrale (di Bochner) *di f in E è l'elemento di X definito da*

$$\int_E f \, d\mu = \lim_{n\to\infty} \int_E \psi_n \, d\mu,$$

ove il limite è fatto nella norma di X e $\{\psi_n\}_{n\in\mathbb{N}}$ è una qualunque successione di funzioni semplici tale che $\lim_{n\to\infty} \int_E \|f(\cdot) - \psi_n(\cdot)\| d\mu = 0$.

La definizione non dipende dalla scelta delle funzioni approssimanti ψ_n: infatti, se $\{\eta_n\}_{n\in\mathbb{N}}$ è un'altra successione in $\mathcal{S}(E, X)$ tale che

$$\lim_{n\to\infty} \int_E \|f(\cdot) - \eta_n(\cdot)\| \, d\mu = 0,$$

allora evidentemente

$$\lim_{n\to\infty} \left\| \int_E \psi_n \, d\mu - \int_E \eta_n \, d\mu \right\| \leq \lim_{n\to\infty} \int_E \|\psi_n(\cdot) - \eta_n(\cdot)\| \, d\mu = 0.$$

L'integrale di Bochner gode di tutte le proprietà usuali: ad esempio, se per $A \in \mathcal{E}$ si definisce

$$\int_A f \, d\mu = \int_E f \chi_A \, d\mu,$$

è facile dedurre (esercizio 3.1.3) che risulta

$$\int_{A\cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu \qquad \forall A, B \in \mathcal{E} \text{ con } A \cap B = \emptyset.$$

Si ha anche, come facile conseguenza della definizione,

$$\left\| \int_E f \, d\mu \right\| \leq \int_E \|f(\cdot)\| \, d\mu.$$

Introduciamo adesso gli spazi L^p per funzioni definite su E a valori in X.

Definizione 3.1.10 *Se $1 \le p < \infty$ denotiamo con $L^p(E;X)$ lo spazio delle funzioni $f : E \to X$ fortemente misurabili tali che $\int_E \|f(\cdot)\|^p \, d\mu < \infty$. Se $p = \infty$ denotiamo con $L^\infty(E;X)$ lo spazio delle funzioni $f : E \to X$ fortemente misurabili e tali che* $\mathrm{supess}_{u\in E}\|f(u)\| < \infty$.

Si verifica facilmente, riconducendosi al caso di funzioni da E in $\mathbb{R}$, che gli spazi $L^p(E;X)$, $1 \le p \le \infty$, sono spazi normati con le usuali rispettive norme:

$$\|f\|_{L^p(E;X)} = \left[\int_E \|f(\cdot)\|^p \, d\mu\right]^{\frac{1}{p}}, \qquad \|f\|_{L^\infty(E;X)} = \mathrm{supess}_{u\in E}\|f(u)\|$$

(si veda l'esercizio 3.1.7).
Inoltre, con la stessa dimostrazione che si fa nel caso $X = \mathbb{R}$, si ottiene che tali spazi sono di Banach. Quando X è riflessivo, si sa anche caratterizzare, in analogia col caso scalare, il duale di $L^p(E;X)$:

Teorema 3.1.11 (di Riesz-Fischer) *Siano $(E, \mathcal{E}, \mu)$ uno spazio misurato e X uno spazio di Banach riflessivo; sia inoltre $p \in [1,\infty[$. Allora $(L^p(E;X))^*$ è isomorfo ed isometrico a $L^q(E;X^*)$, ove $\frac{1}{p} + \frac{1}{q} = 1$; in altre parole, per ogni $T \in (L^p(E;X))^*$ esiste un'unica funzione $g \in L^q(E;X^*)$ tale che*

(i) $Tf = \int_E \langle g(\cdot), f(\cdot)\rangle_X \, d\mu$ *per ogni $f \in L^p(E;X)$, ove $\langle g(u), f(u)\rangle_X$ indica l'azione del funzionale $g(u) \in X^*$ sull'elemento $f(u) \in X$;*

(ii) $\|T\|_{(L^p(E;X))^*} = \|g\|_{L^q(E;X^*)}$.

Consideriamo in particolare il caso in cui $(E, \mathcal{E}, \mu) = ([a,b], \mathcal{M}, m)$, ove m è la misura di Lebesgue. Si osservi che lo spazio $C([a,b],X)$ delle funzioni continue da $[a,b]$ in X è contenuto (propriamente) in tutti gli spazi $L^p(a,b;X)$ (esercizio 3.1.4). In particolare, vale il seguente enunciato:

Proposizione 3.1.12 *Sia X uno spazio di Banach. Se $f \in C([a,b],X)$, allora la funzione $F(t) = \int_a^t f(s)\, ds$ appartiene a $C^1([a,b],X)$ e $F'(t) = f(t)$ per ogni $t \in [a,b]$.*

Dimostrazione Se $t, t+h \in [a,b]$, si ha

$$\begin{aligned} \left\| \frac{F(t+h) - F(t)}{h} - f(t) \right\| &= \left\| \frac{1}{h} \int_t^{t+h} [f(s) - f(t)]\, dt \right\| \le \\ &\le \frac{1}{|h|} \left| \int_t^{t+h} \|f(s) - f(t)\|\, ds \right|, \end{aligned}$$

ed in virtù della continuità di f, l'ultimo membro è infinitesimo per $h \to 0$. $\square$

Corollario 3.1.13 *Sia X uno spazio di Banach. Se $f \in C^1([a,b],X)$ risulta*

$$f(t) = f(s) + \int_s^t f'(r)\,dr \qquad \forall t, s \in [a,b].$$

Dimostrazione Fissato $s \in [a,b]$, la funzione

$$\gamma(t) = f(t) - \int_s^t f'(r)\,dr, \qquad t \in [a,b],$$

appartiene a $C^1([a,b],X)$ ed ha derivata nulla per la proposizione precedente. Perciò, fissato $T \in X^*$, la funzione reale o complessa $g(t) = T\gamma(t)$ ha ancora derivata nulla in quanto, in virtù della linearità e continuità di T,

$$\lim_{h\to 0} \frac{g(t+h) - g(t)}{h} = T\left(\lim_{h\to 0} \frac{\gamma(t+h) - \gamma(t)}{h}\right) = 0.$$

Dunque, g è costante su $[a,b]$; ne segue $T(\gamma(t) - \gamma(s)) = 0$ per ogni $t \in [a,b]$ e per ogni $T \in X^*$. Per il teorema di Hahn-Banach, fissato $t \in [a,b]$ si può scegliere T tale che $T(\gamma(t)-\gamma(s)) = \|\gamma(t)-\gamma(s)\|$; si conclude che $\gamma(t) = \gamma(s)$ per ogni $t \in [a,b]$. Essendo $\gamma(s) = f(s)$, si ha la tesi. $\square$

Esercizi 3.1

1. Sia $X = L^\infty(a,b)$ e sia $f : [a,b] \to X$ definita da

$$f(t) = \chi_{[a,t]} \qquad \forall t \in [a,b].$$

   Si provino i seguenti fatti:

   (i) Per ogni funzionale $T \in X^*$ *positivo* (ossia $Tg \geq 0$ se $g \geq 0$ q.o.) la funzione reale $t \mapsto Tf(t)$ è crescente.

   (ii) La funzione f è debolmente misurabile.

   (iii) Se $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{S}(X)$, allora $\{\varphi_n(t) : \ t \in [a,b], \ n \in \mathbb{N}\}$ è un sottoinsieme al più numerabile di X.

(iv) Se $g : [a, b] \to X$ è fortemente misurabile, allora $\{g(t) : \ t \in [a, b]\}$ è un sottoinsieme "quasi separabile" di X, ovvero esiste un insieme $B_0 \subset [a, b]$ di misura nulla, tale che $\{g(t) : \ t \in [a, b] \setminus B_0\}$ è separabile.

(v) La funzione f non è fortemente misurabile.

2. Si consideri lo spazio misurato $(E, \mathcal{P}(E), \nu)$, ove ν è la misura "cardinalità". Posto $X = L^\infty(0, 1)$, si verifichi che per la funzione f dell'esercizio precedente il viceversa della proposizione 3.1.5 è falso.

3. Sia $(E, \mathcal{E}, \mu)$ uno spazio misurato e sia X uno spazio di Banach. Si provi che se $f : E \to X$ è sommabile, allora per ogni $A, B \in \mathcal{E}$ disgiunti si ha
$$\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu.$$

4. Sia X uno spazio di Banach. Si provi che ogni funzione continua $f : [a, b] \to X$ è fortemente misurabile.

5. Sia $(E, \mathcal{E}, \mu)$ uno spazio misurato e sia X uno spazio di Banach. Si provi che una funzione $f : E \to X$ è fortemente misurabile se e solo se esiste una successione $\{f_n\}$ di funzioni fortemente misurabili che converge a f puntualmente in E.

6. Sia $(E, \mathcal{E}, \mu)$ uno spazio misurato e sia X uno spazio di Banach; sia $p \in [1, \infty]$ e sia q l'esponente coniugato di p. Se $f \in L^p(E; X)$ e $g \in L^q(E; X^*)$, si provi la *disuguaglianza di Hölder*:
$$\int_E |\langle g(\cdot), f(\cdot) \rangle| \, d\mu \le \|f\|_{L^p(E,X)} \cdot \|g\|_{L^q(E;X^*)}.$$

7. Sia $(E, \mathcal{E}, \mu)$ uno spazio misurato e sia X uno spazio di Banach. Fissato $p \in [1, \infty[$, se $f, g \in L^p(E; X)$ si dimostri la *disuguaglianza di Minkowski*:
$$\left[\int_E \|f(\cdot) + g(\cdot)\|^p \, d\mu \right]^{\frac{1}{p}} \le \left[\int_E \|f(\cdot)\|^p \, dt \right]^{\frac{1}{p}} + \left[\int_E \|g(\cdot)\|^p \, dt \right]^{\frac{1}{p}}.$$

3.2 Differenziale di Fréchèt

Veniamo ora alla generalizzazione delle usuali proprietà del calcolo differenziale.

Definizione 3.2.1 *Siano X, Y spazi normati e sia Ω un aperto di X. Una funzione $f : \Omega \to Y$ si dice* differenziabile secondo Fréchèt *nel punto $x_0 \in \Omega$ (più brevemente, F-*differenziabile *in x_0) se esiste un operatore $A \in \mathcal{L}(X, Y)$ tale che risulti*

$$\lim_{\|h\|_X \to 0} \frac{\|f(x_0 + h) - f(x_0) - Ah\|_Y}{\|h\|_X} = 0.$$

L'operatore A si dice differenziale di Fréchèt *di f in x_0 e si scrive $A = f'(x_0)$.*

Osservazione 3.2.2 Il differenziale di Fréchèt di f, se esiste, è unico. Infatti se $B \in \mathcal{L}(X, Y)$ verifica la stessa definizione, risulterà

$$\lim_{\|h\|_X \to 0} \frac{\|f(x_0 + h) - f(x_0) - Bh\|_Y}{\|h\|_X} = 0;$$

per differenza, deduciamo

$$\lim_{\|h\|_X \to 0} \frac{\|Ah - Bh\|_Y}{\|h\|_X} = 0.$$

Dunque, fissato $\varepsilon > 0$, si avrà per $\|h\|_X < \delta$, con δ opportuno,

$$\frac{\|Ah - Bh\|_Y}{\|h\|_X} < \varepsilon,$$

e per linearità

$$\|(A - B)u\|_Y \leq \varepsilon \qquad \forall u \in X \text{ con } \|u\|_X \leq 1.$$

Ciò prova che $\|A - B\|_{\mathcal{L}(X,Y)} \leq \varepsilon$, e dunque $A = B$ per l'arbitrarietà di ε. □

Si noti che la definizione di differenziabilità secondo Fréchèt dipende solo dalla topologia di X e di Y, ed è invariante rispetto al passaggio a norme equivalenti.
La regola di calcolo per il differenziale di funzioni composte è valida anche nel contesto degli spazi normati. Si ha infatti:

Proposizione 3.2.3 *Siano X, Y, Z spazi normati; siano Ω un aperto di X, Λ un aperto di Y, e siano $f : \Omega \to \Lambda$, $g : \Lambda \to Z$. Se f è F-differenziabile in un punto $x_0 \in \Omega$ e g è F-differenziabile nel punto $y_0 = f(x_0)$, allora $g \circ f$ è F-differenziabile in x_0 e si ha*

$$(g \circ f)'(x_0) = g'(y_0) \circ f'(x_0).$$

Dimostrazione Per ipotesi si ha

$$\lim_{\|h\|_X \to 0} \frac{\|\omega(h)\|_Y}{\|h\|_X} = 0, \qquad \text{ove} \quad \omega(h) = f(x_0 + h) - f(x_0) - f'(x_0)h,$$

$$\lim_{\|k\|_Y \to 0} \frac{\|\eta(k)\|_Z}{\|k\|_Y} = 0, \qquad \text{ove} \quad \eta(k) = g(y_0 + k) - g(y_0) - g'(y_0)k.$$

Scegliamo $k = f(x_0 + h) - f(x_0)$ ed osserviamo che se $\|h\|_X \to 0$ allora $\|k\|_Y \to 0$. Si trova allora

$$\begin{aligned} g(f(x_0 + h)) - g(f(x_0)) &= g(y_0 + k) - g(y_0) = g'(y_0)k + \eta(k) = \\ &= g'(y_0)[f'(x_0)h + \omega(h)] + \eta(f'(x_0)h + \omega(h)); \end{aligned}$$

posto $\lambda(h) = g'(y_0)[\omega(h)] + \eta(f'(x_0)h + \omega(h))$, si ricava

$$\lim_{\|h\|_X \to 0} \frac{\|\lambda(h)\|_Z}{\|h\|_X} \le \lim_{\|h\|_X \to 0} \|g'(y_0)\|_{\mathcal{L}(Y,Z)} \frac{\|\omega(h)\|_Y}{\|h\|_X} + \lim_{\|h\|_X \to 0} \frac{\|\eta(k)\|_Z}{\|h\|_X}.$$

D'altra parte se $\|h\|_X \to 0$ si ha

$$\frac{\|\eta(k)\|_Z}{\|h\|_X} = \mathrm{o}(1) \frac{\|k\|_Y}{\|h\|_X} \le \mathrm{o}(1) \left(\|f'(x_0)\|_{\mathcal{L}(X,Y)} + \frac{\|\omega(h)\|_Y}{\|h\|_X} \right) \to 0.$$

Pertanto, $\|\lambda(h)\|_Z = \mathrm{o}(\|h\|_X)$ e la relazione

$$g \circ f(x_0 + h) - g \circ f(x_0) - g'(y_0) \circ f'(x_0)h = \lambda(h)$$

ci dà la tesi. $\square$

Esempi 3.2.4 **(1)** Se $f(x)$ è costante in Ω, allora $f'(x_0) = 0$ per ogni $x_0 \in \Omega$.
(2) Se $f \in \mathcal{L}(X, Y)$, allora $f'(x_0) = f$ per ogni $x_0 \in X$.
(3) Se $f : X \times Y \to Z$ è un'applicazione bilineare e continua, allora f è F-differenziabile in ogni $(x_0, y_0) \in X \times Y$, e $f'(x_0, y_0)$ è l'applicazione lineare $(h, k) \mapsto f(h, y_0) + f(x_0, k)$.

Concludiamo il paragrafo con la seguente

Definizione 3.2.5 *Siano X, Y spazi normati e sia Ω un aperto di X. Se $f : \Omega \to Y$ è una funzione F-differenziabile in ogni punto $x_0 \in \Omega$, l'applicazione $f' : \Omega \to \mathcal{L}(X, Y)$ è detta* derivata prima *di f. Se f' è continua su Ω, diremo che f è* di classe C^1 *su Ω.*

Esercizi 3.2

1. Si verifichi che ogni funzione F-differenziabile in un punto x_0 è continua in x_0, e che il viceversa non è vero.

2. Sia $\phi : \mathbb{R} \to \mathbb{R}$ una funzione derivabile. Se X è uno spazio normato e $f : X \to \mathbb{R}$ è una funzione F-differenziabile, si provi che $\phi \circ f$ è F-differenziabile e
$$(\phi \circ f)'(x) = \phi'(f(x))f'(x) \qquad \forall x \in X.$$

3. Sia X uno spazio con prodotto scalare. Si provi che la funzione $f(x) = \|x\|^2$ è F-differenziabile in X, mentre la funzione $g(x) = \|x\|$ è F-differenziabile soltanto in $X \setminus \{0\}$.

4. Siano $X = \mathbb{R}^N$ e $Y = \mathbb{R}$; si verifichi che la definizione 3.2.5 relativa a funzioni $f : X \to Y$ si riduce a quella usuale.

5. Verificare che i seguenti funzionali a valori in $\mathbb{R}$ sono F-differenziabili negli spazi a fianco indicati, e scriverne la derivata di Fréchèt:

   (i) $J(y) = \int_a^b [y(x)^2 + y'(x)^2 - 2y(x)\sin x]\, dx, \quad X = C^1[a, b]$;

   (ii) $J(y) = \int_a^b \frac{y'(x)^2}{x^3}\, dx, \quad X = C^1[a, b]\ (a > 0)$;

   (iii) $J(y) = \int_a^b [16y(x)^2 - y''(x)^2 + x^3]\, dx, \quad X = C^2[a, b]$;

   (iv) $J(y) = \int_a^b [2xy(x) - y'''(x)^2]\, dx, \quad X = C^3[a, b]$.

6. In uno spazio di Hilbert, le proiezioni su convessi chiusi sono applicazioni F-differenziabili?

3.3 Differenziale di Gâteaux

La nozione di differenziabilità secondo Gâteaux è l'estensione del concetto di derivata direzionale. Si tratta quindi di una proprietà meno forte della F-differenziabilità.

Definizione 3.3.1 *Siano X, Y spazi normati, sia Ω un aperto di X. Una funzione $f : \Omega \to Y$ si dice* differenziabile secondo Gâteaux *nel punto $x_0 \in \Omega$ (più brevemente, G-*differenziabile *in x_0) se esiste un operatore $A \in \mathcal{L}(X, Y)$ tale che*

$$\lim_{t \to 0} \left\| \frac{1}{t}[f(x_0 + tv) - f(x_0)] - Av \right\|_Y = 0 \qquad \forall v \in X.$$

L'operatore A si dice differenziale di Gâteaux *di f in x_0 e si scrive $A = f'_G(x_0)$.*

Si noti che in questa definizione si richiede non solo che esista la derivata di f in x_0 secondo qualunque direzione $v \in X$, ma anche che la dipendenza di tale derivata da v sia *lineare*: ciò non è sempre vero (esercizio 3.3.3).
Per l'unicità del limite, il G-differenziale di f in x_0 è unico (se esiste); l'applicazione $x_0 \mapsto f'_G(x_0)$ si chiama *derivata di Gâteaux* di f.
È noto da esempi elementari che già nel caso $X = \mathbb{R}^2$, $Y = \mathbb{R}$ un'applicazione G-differenziabile può non essere continua e quindi, a maggior ragione, non essere F-differenziabile (esercizio 3.2.1). Però dalle definizioni si vede subito che se f è una funzione F-differenziabile in x_0, allora f è anche G-differenziabile in x_0 ed in tal caso si ha $f'_G(x_0) = f'(x_0)$.
La seguente proposizione generalizza il classico teorema di Lagrange.

Proposizione 3.3.2 *Siano X, Y spazi normati, sia Ω un aperto di X, sia $f : \Omega \to Y$ una funzione G-differenziabile in Ω. Se x_1 e x_2 sono punti di Ω tali che l'intero segmento di estremi x_1, x_2 sia contenuto in Ω, allora*

$$\|f(x_2) - f(x_1)\|_Y \le \sup_{t \in [0,1]} \|f'_G((1-t)x_1 + tx_2)\|_{\mathcal{L}(X,Y)} \|x_2 - x_1\|_X \,.$$

Si noti che il secondo membro della disuguaglianza potrebbe valere $+\infty$ (esercizio 3.3.2).

Dimostrazione Consideriamo la funzione $F : [0, 1] \to \mathbb{R}$ definita da

$$F(t) = \varphi[f((1-t)x_1 + tx_2)], \qquad t \in [0, 1],$$

ove φ è un fissato elemento di Y^*. Essa è derivabile perché, posto $x(t) = (1-t)x_1 + tx_2$, risulta

$$\lim_{h\to 0} \frac{F(t+h)-F(t)}{h} = \lim_{h\to 0} \frac{\varphi[f(x(t)+h(x_2-x_1))-f(x(t))]}{h} =$$
$$= \varphi\left[\lim_{h\to 0} \frac{f(x(t)+h(x_2-x_1))-f(x(t))}{h}\right] = \varphi[f'_G(x(t))(x_2-x_1)].$$

Per il teorema di Lagrange si ha

$$\varphi[f(x_2)] - \varphi[f(x_1)] = F(1) - F(0) = F'(\vartheta) = \varphi[f'_G(x(\vartheta))(x_2-x_1)]$$

con $\vartheta \in]0,1[$ opportuno. Di conseguenza

$$|\varphi[f(x_2)] - \varphi[f(x_1)]| \le \|\varphi\|_{Y^*} \|f'_G(x(\vartheta))\|_{\mathcal{L}(X,Y)} \|x_2 - x_1\|_X\,,$$

e scegliendo φ tale che $\|\varphi\|_{Y^*} = 1$ e $\varphi[f(x_2) - f(x_1)] = \|f(x_2) - f(x_1)\|_Y$ si ottiene la tesi. $\Box$

Il risultato che segue fornisce una condizione sufficiente per la F-differenziabilità ed è la generalizzazione del classico teorema del differenziale totale.

Corollario 3.3.3 *Siano X, Y spazi normati, sia Ω un aperto di X. Se $f : \Omega \to Y$ è una funzione G-differenziabile, e se inoltre $f'_G : \Omega \to \mathcal{L}(X,Y)$ è continua nel punto $x_0 \in \Omega$, allora f è F-differenziabile in x_0 e $f'(x_0) = f'_G(x_0)$.*

Dimostrazione Esiste un intorno U di $0 \in X$ nel quale è definita la funzione

$$\omega(h) = f(x_0 + h) - f(x_0) - f'_G(x_0)h. \qquad h \in U.$$

Questa funzione è G-differenziabile, perché tali sono gli addendi che la compongono (esercizio 3.3.1), e si ha

$$\omega'_G(h) = f'_G(x_0 + h) - f'_G(x_0) \qquad \forall h \in U.$$

Poiché $\omega(0) = 0$, dalla proposizione 3.3.2 otteniamo

$$\|\omega(h)\|_Y \le \sup_{t\in[0,1]} \|f'_G(x_0 + th) - f'_G(x_0)\|_{\mathcal{L}(X,Y)} \|h\|_X\,,$$

e dunque, per la continuità di f'_G in x_0, si ha

$$\lim_{\|h\|_X \to 0} \frac{\|\omega(h)\|_Y}{\|h\|_X} = 0. \quad \Box$$

Esercizi 3.3

1. Siano X, Y spazi normati. Si verifichi che se $f : X \to Y$ è una funzione G-differenziabile, allora per ogni $x \in X$ la funzione $g(h) = f(x+h)$ è a sua volta G-differenziabile con $g'_G(h) = f'_G(x+h)$.

2. Sia $f : \mathbb{R} \to \mathbb{R}$ definita da
$$f(x) = \begin{cases} 0 & \text{se } x = 0, \\ |x|^{3/2} \sin \frac{1}{|x|} & \text{se } x \neq 0. \end{cases}$$
Si provi che f è derivabile in ogni punto di $\mathbb{R}$, ma che
$$\sup_{x \in [-1,1]} |f'(x)| = +\infty.$$

3. Si provi che la funzione $f : \mathbb{R}^2 \to \mathbb{R}$, definita da
$$f(x,y) = \begin{cases} 0 & \text{se } |y| \leq |x|, \\ \operatorname{sgn}(y) \cdot \sqrt{y^2 - x^2} & \text{se } |y| > |x|, \end{cases} \qquad (x,y) \in \mathbb{R}^2,$$
non è G-differenziabile in $(0,0)$, benché abbia tutte le derivate direzionali in tale punto.

4. Posto $X = L^1(a,b)$, si provi che l'applicazione $f \mapsto \|f\|_1$ non è in generale G-differenziabile.
[**Traccia:** si consideri la derivata nel punto χ_I secondo la direzione χ_J, ove I e J sono intervalli disgiunti di $]a,b[$.]

5. Sia X uno spazio normato e sia $f : X \to \mathbb{R}$ una funzione convessa e G-differenziabile. Si provi che
$$f(x_0 + v) \geq f(x_0) + f'_G(x_0)v \qquad \forall x_0, v \in X.$$
[**Traccia:** si utilizzi la crescenza del rapporto incrementale nella direzione v (esercizio 2.1.3).]

3.4 L'operatore di superposizione

Esaminiamo in dettaglio un esempio interessante e non banale di funzione F-differenziabile.
Sia $\Omega \subset \mathbb{R}^N$ un aperto di misura di Lebesgue finita (limitato o no) e sia $f : \Omega \times \mathbb{R} \to \mathbb{R}$ una funzione tale che:

(i) $x \mapsto f(x,t)$ è misurabile in Ω per ogni $t \in \mathbb{R}$;

(ii) esiste un insieme misurabile $\Omega_0 \subseteq \Omega$ tale che $m_N(\Omega \backslash \Omega_0) = 0$ e $t \mapsto f(x,t)$ è continua su $\mathbb{R}$ per ogni $x \in \Omega_0$.

Poniamo

$$[\Phi(u)](x) = f(x, u(x)), \qquad x \in \Omega;$$

per comprensibili ragioni, l'applicazione $u \mapsto \Phi(u)$ si chiama *operatore di superposizione.*

Si tratta ora di scegliere lo spazio in cui far variare le funzioni u. Osserviamo anzitutto che se $u : \Omega \to \mathbb{R}$ è misurabile, allora $\Phi(u)$ è a sua volta misurabile. Infatti, questo è vero quando u è una funzione semplice (se $u = \sum_{i=1}^k \alpha_i \chi_{E_i}$, con gli E_i disgiunti e $\bigcup_{i=1}^k E_i = \Omega$, si ha $\Phi(u) = \sum_{i=1}^k f(\cdot, \alpha_i)\chi_{E_i}$ e dunque $\Phi(u)$ è misurabile); poi, se u è misurabile e $\{u_n\}$ è una successione di funzioni semplici che converge ad u puntualmente in Ω, allora $\Phi(u)$ è il limite puntuale in Ω_0, dunque q.o. in Ω, delle funzioni misurabili $\Phi(u_n)$, e dunque è misurabile in virtù della completezza della misura di Lebesgue.

Vale inoltre la seguente proprietà:

Lemma 3.4.1 *Nelle ipotesi precedenti, se $\{u_n\}$ è una successione di funzioni misurabili su Ω tale che $u_n \to u$ in misura, allora $\Phi(u_n) \to \Phi(u)$ in misura.*

Dimostrazione Sia $\varepsilon > 0$; dobbiamo provare che

$$\lim_{n\to\infty} m_N(\{x \in \Omega_0 : \ |\Phi(u_n)(x) - \Phi(u)(x)| > \varepsilon\}) = 0.$$

Per ogni $k \in \mathbb{N}^+$ consideriamo l'insieme

$$\Omega_k = \left\{ x \in \Omega_0 : \ |f(x, u(x)) - f(x,t)| \le \varepsilon \quad \forall t \in \left[u(x) - \frac{1}{k}, u(x) + \frac{1}{k} \right] \right\}.$$

Non è evidente che l'insieme Ω_k sia misurabile: la dimostrazione di questo fatto è tratteggiata nell'esercizio 3.4.1. Risulta inoltre $\Omega_k \supseteq \Omega_{k+1}$ e, per la continuità di f rispetto alla variabile t, si ha $\bigcup_{k=1}^\infty \Omega_k = \Omega_0$; essendo $m_N(\Omega_0) = m_N(\Omega) < \infty$, se ne deduce, in virtù della misurabilità di Ω_k, che $m_N(\Omega_0 \setminus \Omega_k) \to 0$ per $k \to \infty$. Dunque, fissato $\eta > 0$, esiste $k \in \mathbb{N}^+$ (che pensiamo d'ora in poi fissato) tale che $m_N(\Omega_0 \setminus \Omega_k) < \eta/2$.

Definiamo ora per ogni $n \in \mathbb{N}$ l'insieme

$$A_n = \left\{ x \in \Omega_0 : \ |u_n(x) - u(x)| \le \frac{1}{k} \right\};$$

poiché $u_n \to u$ in misura, si ha $\lim_{n\to\infty} m_N\Omega_0 \setminus A_n) = 0$ e dunque esiste $\overline{n} \in \mathbb{N}$ tale che

$$m_N(\Omega_0 \setminus A_n) < \frac{\eta}{2} \qquad \forall n \geq \overline{n}.$$

D'altra parte, se $x \in A_n \cap \Omega_k$ si ha $u_n(x) \in [u(x) - 1/k, u(x) + 1/k]$ e quindi $|\Phi(u_n)(x) - \Phi(u)(x)| \leq \varepsilon$; ne segue, per ogni $n \geq \overline{n}$,

$$\begin{aligned} &m_N(\{x \in \Omega_0 : \ |\Phi(u_n)(x) - \Phi(u)(x)| > \varepsilon\}) \leq \\ &\quad \leq m_N(\Omega_0 \setminus (A_n \cap \Omega_k)) \leq m_N(\Omega_0 \setminus A_n) + m_N(\Omega_0 \setminus \Omega_k) \leq \frac{\eta}{2} + \frac{\eta}{2} = \eta\,, \end{aligned}$$

cioè la tesi. $\square$

Una conseguenza del lemma precedente è la continuità dell'operatore Φ su opportuni spazi L^p.

Proposizione 3.4.2 *Supponiamo che, in aggiunta alle ipotesi precedenti, esistano $p, q \geq 1$ tali che*

$$|f(x,t)| \leq a(x) + b|t|^{p/q} \qquad \textit{per q.o. } x \in \Omega, \ \forall t \in \mathbb{R},$$

ove $a \in L^q(\Omega)$ con $a(x) \geq 0$ q.o., e $b \geq 0$. Allora l'operatore di superposizione Φ manda con continuità $L^p(\Omega)$ in $L^q(\Omega)$.

Dimostrazione Anzitutto, se $u \in L^p(\Omega)$ si ha $\Phi(u) \in L^q(\Omega)$ perché

$$\int_\Omega |f(x,u(x))|^q\,dx \leq 2^{q-1}\left[\int_\Omega a(x)^q\,dx + b^q \int_\Omega |u(x)|^p\,dx\right];$$

dobbiamo provare la continuità di Φ, ossia mostrare che, se $\{u_n\} \subseteq L^p(\Omega)$ è una successione tale che $u_n \to u$ in $L^p(\Omega)$, allora $\Phi(u_n) \to \Phi(u)$ in $L^q(\Omega)$. Dato che, in particolare, $u_n \to u$ in misura, il lemma 3.4.1 ci dice che $\Phi(u_n) \to \Phi(u)$ in misura. Inoltre vale la seguente stima:

$$\begin{aligned} |\Phi(u_n)(x) - \Phi(u)(x)|^q &\leq 2^{q-1}\Big[|\Phi(u_n)(x)|^q + |\Phi(u)(x)|^q\Big] \leq \\ &\leq 4^{q-1}\Big[a(x)^q + b^q|u_n(x)|^p + a(x)^q + b^q|u(x)|^p\Big] \leq \\ &\leq 4^{q-1}\Big[2a(x)^q + 2^{p-1}b^q|u_n(x) - u(x)|^p + (2^{p-1}+1)b^q|u(x)|^p\Big] \leq \\ &\leq c_{p,q}\Big[a(x)^q + |u_n(x) - u(x)|^p + |u(x)|^p\Big]. \end{aligned}$$

Sia allora $\varepsilon > 0$: esiste $\overline{n} \in \mathbb{N}$ tale che $\|u_n - u\|^p_{L^p(\Omega)} < \varepsilon$ per ogni $n \geq \overline{n}$; quindi, per l'assoluta continuità dell'integrale e per la stima precedente, esiste $\eta > 0$ per cui si ha

$$m_N(E) < \eta, \quad n \geq \overline{n} \qquad \Longrightarrow \qquad \int_E |\Phi(u_n) - \Phi(u)|^q \, dx \leq 3c_{p,q}\, \varepsilon.$$

Scegliamo ora $\delta > 0$ tale che $\delta^q < \varepsilon$; posto

$$E_n = \{x \in \Omega : \ |\Phi(u_n)(x) - \Phi(u)(x)| \geq \delta\},$$

sarà $m_N(E_n) < \eta$ per ogni $n \geq \overline{\overline{n}} \geq \overline{n}$. Ne segue, finalmente,

$$\begin{aligned} &\int_\Omega |\Phi(u_n) - \Phi(u)|^q \, dx = \\ &\quad = \int_{E_n} |\Phi(u_n) - \Phi(u)|^q \, dx + \int_{\Omega \setminus E_n} |\Phi(u_n) - \Phi(u)|^q \, dx \leq \\ &\quad \leq 3c_{p,q}\, \varepsilon + \delta^q \, m_N(\Omega) < C\varepsilon \qquad \forall n \geq \overline{\overline{n}}, \end{aligned}$$

e ciò prova la tesi. □

Vediamo ora quali ulteriori condizioni garantiscono la F-differenziabilità dell'applicazione Φ.

Teorema 3.4.3 *Supponiamo che f soddisfi tutte le ipotesi della proposizione 3.4.2 con $p > 2$ e $q = \frac{p}{p-1}$. Supponiamo inoltre che la funzione $\frac{\partial f}{\partial t}$ esista e verifichi le seguenti condizioni:*

(i) *$x \mapsto \frac{\partial f}{\partial t}(x,t)$ è misurabile in Ω per ogni $t \in \mathbb{R}$;*

(ii) *$t \mapsto \frac{\partial f}{\partial t}(x,t)$ è continua su $\mathbb{R}$ per q.o. $x \in \Omega$;*

(iii) *vale la disuguaglianza*

$$\left| \frac{\partial f}{\partial t}(x,t) \right| \leq \alpha(x) + \beta |t|^{p-2} \qquad \text{per q.o. } x \in \Omega, \ \forall t \in \mathbb{R},$$

ove $\alpha \in L^{\frac{p}{p-2}}(\Omega)$ con $\alpha(x) \geq 0$ q.o., e $\beta \geq 0$.

Allora l'operatore di superposizione $\Phi : L^p(\Omega) \to L^q(\Omega)$ è F-differenziabile in ogni punto $u \in L^p(\Omega)$ e

$$[\Phi'(u)v](x) = \frac{\partial f}{\partial t}(x, u(x))v(x), \quad x \in \Omega, \qquad \forall v \in L^p(\Omega).$$

Dimostrazione Anzitutto, osserviamo che l'applicazione Ψ definita da

$$\Psi(u)(x) = \frac{\partial f}{\partial t}(x, u(x)), \quad x \in \Omega, \qquad \forall u \in L^p(\Omega)$$

manda con continuità $L^p(\Omega)$ in $L^{\frac{p}{p-2}}(\Omega)$ in virtù della proposizione 3.4.2; quindi, per la disuguaglianza di Hölder con esponenti $p-1$ e $\frac{p-1}{p-2}$, se $v \in L^p(\Omega)$ si ha $\Psi(u)v \in L^q(\Omega)$ e

$$\|\Psi(u)v\|_{L^q(\Omega)} \leq \|\Psi(u)\|_{L^{\frac{p}{p-2}}(\Omega)} \cdot \|v\|_{L^p(\Omega)}.$$

Poniamo adesso per $u, v \in L^p(\Omega)$

$$\omega(v) = \Phi(u+v) - \Phi(u) - \Psi(u)v;$$

allora

$$\begin{aligned}
&\|\omega(v)\|^q_{L^q(\Omega)} = \\
&= \int_\Omega \left| f(x, u(x) + v(x)) - f(x, u(x)) - \frac{\partial f}{\partial t}(x, u(x))v(x) \right|^q dx = \\
&= \int_\Omega \left| \int_0^1 \left[\frac{\partial f}{\partial t}(x, u(x) + \tau v(x)) - \frac{\partial f}{\partial t}(x, u(x)) \right] v(x)\ d\tau \right|^q dx\,;
\end{aligned}$$

dunque, posto

$$z(x) = \int_0^1 \left[\frac{\partial f}{\partial t}(x, u(x) + \tau v(x)) - \frac{\partial f}{\partial t}(x, u(x)) \right] d\tau,$$

si ha

$$\|\omega(v)\|^q_{L^q(\Omega)} = \int_\Omega |z(x)v(x)|^q\ dx \leq \|z\|^q_{L^{\frac{p}{p-2}}(\Omega)} \cdot \|v\|^q_{L^p(\Omega)}\ .$$

Valutiamo adesso la norma $\|z\|_{L^{\frac{p}{p-2}}(\Omega)}$. Usando ancora una volta la disuguaglianza di Hölder ed il teorema di Tonelli, si trova

$$\begin{aligned}
\int_\Omega |z|^{\frac{p}{p-2}}\ dx \ &\leq \ \int_\Omega \int_0^1 \left| \frac{\partial f}{\partial t}(x, u(x) + \tau v(x)) - \frac{\partial f}{\partial t}(x, u(x)) \right|^{\frac{p}{p-2}} d\tau dx = \\
&= \ \int_0^1 \|\Psi(u + \tau v) - \Psi(u)\|^{\frac{p}{p-2}}_{L^{\frac{p}{p-2}}(\Omega)}\ d\tau.
\end{aligned}$$

Per la continuità dell'applicazione Ψ, l'integrando all'ultimo membro tende a 0 per $\|v\|_{L^p(\Omega)} \to 0$. D'altronde, in virtù dell'ipotesi (iii), per $\|v\|_{L^p(\Omega)} \le 1$ esso è dominato dalla costante

$$c_p\Big[\|\alpha\|^{\frac{p}{p-2}}_{L^{\frac{p}{p-2}}(\Omega)} + \beta^{\frac{p}{p-2}}(\|u\|^p_{L^p(\Omega)} + 1)\Big]$$

(che ovviamente è una funzione sommabile in $[0,1]$ rispetto a τ); ne segue, per il teorema di Lebesgue, che $\|z\|_{L^{\frac{p}{p-2}}(\Omega)} \to 0$ per $\|v\|_{L^p(\Omega)} \to 0$. Ciò prova che

$$\lim_{\|v\|_{L^p(\Omega)}\to 0} \frac{\|\omega(v)\|_{L^q(\Omega)}}{\|v\|_{L^p(\Omega)}} = 0. \quad \square$$

Abbiamo così provato la F-differenziabilità di Φ sugli spazi L^p con $p > 2$. Se $p = 2$ vale un risultato più debole:

Teorema 3.4.4 *Supponiamo che f verifichi tutte le ipotesi della proposizione 3.4.2 con $p = 2$ e $q = 2$. Supponiamo inoltre che la funzione $\frac{\partial f}{\partial t}$ esista e verifichi le seguenti condizioni:*

(i) *$x \mapsto \frac{\partial f}{\partial t}(x,t)$ è misurabile in Ω per ogni $t \in \mathbb{R}$;*

(ii) *$t \mapsto \frac{\partial f}{\partial t}(x,t)$ è continua su $\mathbb{R}$ per q.o. $x \in \Omega$;*

(iii) *vale la disuguaglianza*

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \le M \qquad \text{per q.o. } x \in \Omega,\ \forall t \in \mathbb{R},$$

ove $M \ge 0$.

Allora l'operatore di superposizione $\Phi : L^2(\Omega) \to L^2(\Omega)$ è G-differenziabile in ogni $u \in L^2(\Omega)$ e

$$[\Phi'_G(u)v](x) = [\Psi(u)v](x) = \frac{\partial f}{\partial t}(x,u(x))v(x), \quad x \in \Omega, \qquad \forall v \in L^2(\Omega).$$

Dimostrazione Sia $v \in L^2(\Omega)$. Per $\lambda \ne 0$ e per q.o. $x \in \Omega$ si ha

$$\frac{1}{\lambda}\Big[[\Phi(u+\lambda v)](x) - [\Phi(u)](x)\Big] - [\Psi(u)](x)v(x) =$$
$$= v(x)\int_0^1 \left[\frac{\partial f}{\partial t}(x,u(x)+t\lambda v(x)) - \frac{\partial f}{\partial t}(x,u(x))\right] dt.$$

Se $\lambda \to 0$, si ha $t\lambda v(x) \to 0$ per ogni $t \in [0,1]$ e per q.o. $x \in \Omega$; quindi, utilizzando la continuità di $\frac{\partial f}{\partial t}(x,\cdot)$, per q.o. $x \in \Omega$ l'integrando a secondo membro converge puntualmente a 0 per $\lambda \to 0$. Esso è anche dominato dalla costante $2M$, e quindi il teorema di Lebesgue implica che

$$\lim_{\lambda\to 0}\left[\frac{1}{\lambda}\Big[[\Phi(u+\lambda v)](x) - [\Phi(u)](x)\Big] - [\Psi(u)](x)v(x)\right] = 0 \quad \text{per q.o. } x \in \Omega.$$

Dato che si ha anche

$$\left|\frac{1}{\lambda}\Big[[\Phi(u+\lambda v)](x) - [\Phi(u)](x)\Big] - [\Psi(u)](x)v(x)\right|^2 \le (2M)^2|v(x)|^2 \text{ q.o. in } \Omega,$$

un'altra applicazione del teorema di Lebesgue porta a concludere che

$$\lim_{\lambda\to 0}\left\|\frac{1}{\lambda}\Big[[\Phi(u+\lambda v)](x) - [\Phi(u)](x)\Big] - [\Psi(u)](x)v(x)\right\|^2_{L^2(\Omega)} = 0,$$

che è la tesi. □

Il risultato del teorema precedente si può ulteriormente precisare: se $p = q = 2$, l'applicazione Φ è F-differenziabile in un punto fissato di $L^p(\Omega)$ *soltanto* nel caso in cui la funzione $f(x,t)$ è affine rispetto alla variabile t (e naturalmente in questo caso la Φ sarà F-differenziabile in ogni punto).

Teorema 3.4.5 *Siano $p = q = 2$. Nelle stesse ipotesi del teorema 3.4.4, se l'operatore di superposizione Φ è F-differenziabile in un punto $u_0 \in L^2(\Omega)$, allora la funzione f è necessariamente della forma*

$$f(x,t) = h(x) + tk(x) \qquad \textit{per q.o. } x \in \Omega,\ \forall t \in \mathbb{R},$$

ove $h \in L^2(\Omega)$ e $k \in L^\infty(\Omega)$ sono funzioni opportune.

Dimostrazione Anzitutto notiamo che $\Phi'(u_0)$, che esiste per ipotesi, deve coincidere con $\Phi'_G(u_0)$, che esiste per il teorema 3.4.4 e coincide con $\Psi(u_0)$. Per ogni $x \in \Omega$ sia $B(x,\delta_x)$ una palla contenuta in Ω. Poniamo per $x \in \Omega$ e $\delta \le \delta_x$

$$v_{\delta,x}(\xi) = \lambda\, \chi_{B(x,\delta)}(\xi), \quad \xi \in \Omega,$$

con $\lambda \in \mathbb{R}$ fissato: chiaramente, qualunque sia $x \in \Omega$, $v_{\delta,x} \to 0$ in $L^2(\Omega)$ per $\delta \to 0$; quindi dall'ipotesi segue che, posto

$$\omega(v_{\delta,x}) = \Phi(u_0 + v_{\delta,x}) - \Phi(u_0) - \Psi(u_0)v_{\delta,x}\,,$$

risulta per ogni $x \in \Omega$

$$\lim_{\delta \to 0} \frac{\|\omega(v_{\delta,x})\|_{L^2(\Omega)}}{\|v_{\delta,x}\|_{L^2(\Omega)}} = 0.$$

D'altra parte, osservato che

$$\frac{\|\omega(v_{\delta,x})\|^2_{L^2(\Omega)}}{\|v_{\delta,x}\|^2_{L^2(\Omega)}} = \frac{1}{|\lambda|^2} \cdot \frac{1}{m_N(B(x,\delta))} \cdot$$
$$\cdot \int_{B(x,\delta)} \left[f(\xi, u_0(\xi) + \lambda) - f(\xi, u_0(\xi)) - \frac{\partial f}{\partial t}(\xi, u_0(\xi))\lambda \right]^2 d\xi,$$

si ha anche

$$\lim_{\delta \to 0} \frac{\|\omega(v_{\delta,x})\|^2_{L^2(\Omega)}}{\|v_{\delta,x}\|^2_{L^2(\Omega)}} = \frac{1}{|\lambda|^2} \left| f(x, u_0(x) + \lambda) - f(x, u_0(x)) - \frac{\partial f}{\partial t}(x, u_0(x))\lambda \right|^2$$

in ogni punto $x \in \Omega$ che sia punto di Lebesgue per la funzione integranda; quindi tale uguaglianza vale per q.o. $x \in \Omega$. Più precisamente, mettendo insieme i due risultati otteniamo che per ogni $\lambda \in \mathbb{R}$ esiste un insieme misurabile $A_\lambda \subseteq \Omega$ con $m_N(\Omega \setminus A_\lambda) = 0$, tale che

$$\left| f(x, u_0(x) + \lambda) - f(x, u_0(x)) - \frac{\partial f}{\partial t}(x, u_0(x))\lambda \right| = 0 \qquad \forall x \in A_\lambda.$$

Facendo variare λ in $\mathbb{Q}$ otteniamo, per intersezione numerabile, un insieme misurabile $A \subseteq \Omega$, con $m_N(\Omega \setminus A) = 0$, tale che

$$\left| f(x, u_0(x) + \lambda) - f(x, u_0(x)) - \frac{\partial f}{\partial t}(x, u_0(x))\lambda \right| = 0 \qquad \forall x \in A, \ \forall \lambda \in \mathbb{Q};$$

dato che f è continua (anzi di classe C^1) nella variabile t, la relazione precedente implica che

$$\left| f(x, u_0(x) + \lambda) - f(x, u_0(x)) - \frac{\partial f}{\partial t}(x, u_0(x))\lambda \right| = 0 \qquad \forall x \in A, \ \forall \lambda \in \mathbb{R}.$$

Posto $t = u_0(x) + \lambda$, con $x \in A$, t è un arbitrario numero reale: otteniamo così

$$f(x,t) = f(x, u_0(x)) + \frac{\partial f}{\partial t}(x, u_0(x))(t - u_0(x)) \qquad \forall x \in A, \ \forall t \in \mathbb{R}.$$

Ne segue la tesi, prendendo le funzioni

$$h(x) = f(x, u_0(x)) - \frac{\partial f}{\partial t}(x, u_0(x))u_0(x), \qquad k(x) = \frac{\partial f}{\partial t}(x, u_0(x)).$$

Si noti che $h \in L^2(\Omega)$ e $k \in L^\infty(\Omega)$, come richiesto. $\square$

Esercizi 3.4

1. Sia f una funzione verificante le ipotesi elencate all'inizio del paragrafo 3.4 e sia $u : \Omega \to \mathbb{R}$ una funzione misurabile. Per $\varepsilon, r > 0$ poniamo
$$\Omega_{\varepsilon,r} = \{x \in \Omega_0 : |f(x, u(x)) - f(x,t)| \le \varepsilon \text{ per } |t - u(x)| \le r\}.$$
Si provi che:
   (i) l'insieme $\Omega_{\varepsilon,r}$ coincide con
$$\{x \in \Omega_0 : \ |f(x, u(x)) - f(x,q)| \le \varepsilon \ \forall q \in \mathbb{Q} \text{ con } |q - u(x)| \le r\};$$
   (ii) se u è una funzione semplice, allora $\Omega_{\varepsilon,r}$ è misurabile;
   (iii) se u è arbitraria, se u_n è una successione di funzioni semplici che converge puntualmente a u in Ω, e se poniamo
$$\Omega^n_{\varepsilon,r} = \{x \in \Omega_0 : |f(x, u_n(x)) - f(x,t)| \le \varepsilon \text{ per } |t - u_n(x)| \le r\},$$
allora risulta
$$\Omega_{\varepsilon,r} = \bigcap_{h\in\mathbb{N}^+} \limsup_{n\to\infty} \Omega^n_{\varepsilon+\frac{1}{h},r} = \bigcap_{h\in\mathbb{N}^+} \bigcap_{m\in\mathbb{N}} \bigcup_{n=m}^{\infty} \Omega^n_{\varepsilon+\frac{1}{h},r},$$
cosicché $\Omega_{\varepsilon,r}$ è misurabile per ogni $\varepsilon, r > 0$.

2. Sia f una funzione verificante le ipotesi della proposizione 3.4.2 con $\frac{1}{p} + \frac{1}{q} = 1$. Si provi che il funzionale
$$J(u) = \int_\Omega \int_0^{u(x)} f(x,t)\ dtdx, \qquad u \in L^p(\Omega),$$
è F-differenziabile in tutti i punti di $L^p(\Omega)$, con
$$J'(u)v = \int_\Omega f(x, u(x))v(x)\, dx \qquad \forall v \in L^p(\Omega).$$

3.5 Derivate successive

Vediamo brevemente come si possono introdurre le derivate successive di una applicazione f fra spazi normati.

Definizione 3.5.1 *Siano X, Y spazi normati, sia Ω un aperto di X e sia $f : \Omega \to Y$ una funzione F-differenziabile. Se la derivata di Fréchèt $f' : \Omega \to \mathcal{L}(X, Y)$ è F-differenziabile in un punto $x_0 \in \Omega$, l'applicazione $[f']'(x_0) \in \mathcal{L}(X, \mathcal{L}(X, Y))$ è detta* differenziale secondo *di f in x_0, e si denota con $f''(x_0)$. Se f' è F-differenziabile in Ω, l'applicazione $f'' : \Omega \to \mathcal{L}(X, \mathcal{L}(X, Y))$ è detta* derivata seconda *di f; se essa è continua, la funzione f si dirà* di classe C^2 *in Ω.*

Esempi 3.5.2 **(1)** Se $f \in \mathcal{L}(X, Y)$, allora $f'(x_0) = f$ per ogni $x_0 \in X$, e $f''(x_0) = 0$ per ogni $x_0 \in X$.
(2) Sia $X = C[a, b]$, e poniamo $f(g) = g^2$ per ogni $g \in X$, ossia $[f(g)](t) = g(t)^2$ per ogni $t \in [a, b]$. La funzione f applica $C[a, b]$ in sé, e si verifica facilmente che essa è F-differenziabile in ogni punto $g_0 \in C[a, b]$, con

$$[f'(g_0)h](t) = 2g_0(t)h(t), \quad t \in [a, b], \qquad \forall h \in C[a, b].$$

Si ha poi, per ogni $k \in C[a, b]$,

$$f'(g_0 + h)k - f'(g_0)k = 2(g_0 + h)k - 2g_0k = 2hk,$$

e dunque

$$[f''(g_0)h]k = 2hk \qquad \forall h, k \in C[a, b].$$

L'applicazione f'' è dunque bilineare su $C[a, b] \times C[a, b]$. Questo non è casuale, come mostra la proposizione che segue.

Proposizione 3.5.3 *Siano X, Y spazi normati. Lo spazio $\mathcal{L}(X, \mathcal{L}(X, Y))$ è isomorfo ed isometrico allo spazio $\mathcal{L}_2(X, Y)$ delle applicazioni $\varphi : X \times X \to Y$ bilineari e continue, munito della norma*

$$\|\varphi\|_{\mathcal{L}_2(X,Y)} = \sup_{\|h\|_X,\, \|k\|_X \le 1} \|\varphi(h, k)\|_Y \ .$$

Dimostrazione Sia $\varphi \in \mathcal{L}(X, \mathcal{L}(X, Y))$ e siano $h, k \in X$. Allora $\varphi h \in \mathcal{L}(X, Y)$ e $[\varphi h]k \in Y$. La dipendenza di $[\varphi h]k$ da h e k è ovviamente bilineare, nonché continua: quindi esiste $\psi \in \mathcal{L}_2(X, Y)$ tale che

$$\psi(h, k) = [\varphi h]k \qquad \forall h, k \in X,$$

e risulta

$$\begin{aligned}\|\psi\|_{\mathcal{L}_2(X,Y)} &= \sup_{\|h\|_X,\,\|k\|_X\le 1}\|\psi(h,k)\|_Y = \sup_{\|h\|_X,\,\|k\|_X\le 1}\|[\varphi h]k)\|_Y = \\ &= \sup_{\|h\|_X\le 1}\|\varphi h\|_{\mathcal{L}(X,Y)} = \|\varphi\|_{\mathcal{L}(X,\mathcal{L}(X,Y))}\,.\end{aligned}$$

Dunque, esiste una corrispondenza $i : \mathcal{L}(X,\mathcal{L}(X,Y)) \to \mathcal{L}_2(X,Y)$, definita da $i(\varphi) = \psi$, che è un'isometria lineare. Resta da far vedere che i è anche surgettiva. Sia dunque $\psi \in \mathcal{L}_2(X,Y)$ e consideriamo per ogni fissato $h \in X$ l'applicazione $\psi_h : X \to Y$ definita da $\psi_h k = \psi(h,k)$ per ogni $k \in X$; ovviamente, $\psi_h \in \mathcal{L}(X,Y)$ per ogni $h \in X$. La corrispondenza $h \mapsto \psi_h$ è un'applicazione lineare $\varphi : X \to \mathcal{L}(X,Y)$, che è anche continua perché per ogni $h \in X$ si ha

$$\begin{aligned}\|\varphi h\|_{\mathcal{L}(X,Y)} &= \|\psi_h\|_{\mathcal{L}(X,Y)} = \sup_{k\in X\setminus\{0\}}\frac{\|\psi_h k\|_Y}{\|k\|_X} = \\ &= \sup_{k\in X\setminus\{0\}}\frac{\|\psi(h,k)\|_Y}{\|k\|_X} \le \|\psi\|_{\mathcal{L}_2(X,Y)}\|h\|_X\,.\end{aligned}$$

Dunque $\varphi \in \mathcal{L}(X,\mathcal{L}(X,Y))$, e naturalmente

$$i(\varphi)(h,k) = [\varphi h]k = \psi_h k = \psi(h,k) \qquad \forall h,k \in X. \quad \square$$

In definitiva, si può sempre vedere la derivata seconda come una forma bilineare continua su $X \times X$. In modo analogo, definendo induttivamente la derivata n-sima come la derivata prima di $f^{(n-1)} : \Omega \to \mathcal{L}_{n-1}(X,Y)$, si trova che tale derivata n-sima è un'applicazione n-lineare $f^{(n)} : \Omega \to \mathcal{L}_n(X,Y)$, dove per ogni $n \in \mathbb{N}^+$ lo spazio $\mathcal{L}_n(X,Y)$ è la classe delle applicazioni di X^n in Y n-lineari e continue, con la norma

$$\|\varphi\|_{\mathcal{L}_n(X,Y)} = \sup\{\|\varphi(h_1,\ldots,h_n)\|_Y \;:\; \|h_i\|_X \le 1 \text{ per } i = 1,\ldots n\}.$$

Vediamo infine come si scrive la formula di Taylor (arrestata al secondo ordine, per semplicità) nel caso di applicazioni fra spazi normati.

Proposizione 3.5.4 *Siano X, Y spazi normati, sia Ω un aperto di X e sia $f : \Omega \to Y$ di classe C^2. Se $x_0 \in \Omega$ e se $B(x_0,R) \subseteq \Omega$, allora per ogni $h \in B(0,R)$ si ha*

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{1}{2}[f''(x_0)h]h + \eta(h),$$

ove $\eta : B(0,R) \to Y$ è una funzione tale che

$$\lim_{\|h\|_X \to 0} \frac{\|\eta(h)\|_Y}{\|h\|_X^2} = 0.$$

Dimostrazione Sia $h \in B(0,R)$. Fissato $\psi \in Y^*$, la funzione reale

$$g_h(t) = \psi(f(x_0 + th)), \qquad t \in [-1,1],$$

è di classe C^2: infatti per la proposizione 3.2.3 si ha

$$g_h'(t) = \psi(f'(x_0 + th)h) \qquad \forall t \in [-1,1],$$
$$g_h''(t) = \psi([f''(x_0 + th)h]h) \qquad \forall t \in [-1,1],$$

ed inoltre g_h' e g_h'' sono funzioni continue, essendo f di classe C^2.
Dalla formula di Taylor per la funzione g_h segue

$$\psi(f(x_0 + h) - f(x_0)) = g_h(1) - g_h(0) = g_h'(0) + \frac{1}{2} g_h''(\xi_h)$$

ove ξ_h è un punto opportuno in $]0,1[$. Ne segue

$$\begin{aligned} \psi(f(x_0 + h) - f(x_0)) &= \psi(f'(x_0)h) + \frac{1}{2}\psi([f''(x_0 + \xi_h h)h]h) = \\ &= \psi(f'(x_0)h) + \frac{1}{2}\psi([f''(x_0)h]h) + \psi(\eta(h)), \end{aligned}$$

ove

$$\eta(h) = \frac{1}{2}[(f''(x_0 + \xi_h h) - f''(x_0))h]h.$$

Dunque

$$\begin{aligned} &\left| \psi \left[f(x_0 + h) - f(x_0) - f'(x_0)h - \frac{1}{2}[f''(x_0)h]h \right] \right| \le \\ &\qquad \le |\psi(\eta(h))| \le \|\psi\|_{Y^*} \|\eta(h)\|_Y . \end{aligned}$$

Per l'arbitrarietà di $\psi \in Y^*$, si ricava

$$\left\| f(x_0 + h) - f(x_0) - f'(x_0)h - \frac{1}{2}[f''(x_0)h]h \right\|_Y \le \|\eta(h)\|_Y .$$

Poiché, per la continuità di f'',

$$\lim_{\|h\|_X \to 0} \frac{\|\eta(h)\|_Y}{\|h\|_X^2} = 0,$$

si ha la tesi. $\square$

Esempio 3.5.5 Sia H uno spazio di Hilbert (reale, per semplicità), e si consideri la funzione

$$F(x) = \|x\|_H^4 - \|x\|_H^2, \qquad x \in H.$$

Poiché

$$\begin{aligned} F(x+h) - F(x) &= (\|x+h\|_H^2 - \|x\|_H^2)(\|x+h\|_H^2 + \|x\|_H^2 - 1) = \\ &= (2\langle x, h\rangle_H + \|h\|_H^2)(\|x+h\|_H^2 + \|x\|_H^2 - 1), \end{aligned}$$

si vede subito che la funzione F è differenziabile secondo Fréchèt e $F'(x) = (4\|x\|_H^2 - 2)x$, ossia

$$\langle F'(x), h\rangle_H = (4\|x\|^2 - 2)\langle x, h\rangle_H \qquad \forall h \in H.$$

Analogamente,

$$F'(x+k) - F'(x) = 4(\|x+k\|_H^2 - \|x\|_H^2)x + (4\|x+k\|_H^2 - 2)k,$$

da cui è facile dedurre che anche F' è differenziabile secondo Fréchèt in H e

$$\langle F''(x)k, h\rangle_H = (4\|x\|_H^2 - 2)\langle k, h\rangle_H + 8\langle x, k\rangle_H \,\langle x, h\rangle_H \qquad \forall k, h \in H.$$

Esercizi 3.5

1. Siano X, Y spazi normati, sia Ω un aperto di X e sia $f : \Omega \to Y$ una funzione di classe C^2.

   (i) Fissato $x_0 \in \Omega$, si provi che, posto

   $$\psi(h,k) = f(x_0+h+k) - f(x_0+k) - f(x_0+h) + f(x_0), \qquad h, k \in X,$$

   se $\|h\|_X$ e $\|k\|_X$ sono sufficientemente piccole si ha

   $$\|\psi(h,k) - [f''(x_0)h]k\|_Y \le \varepsilon(\|h\|_X + 2\|k\|_X)\|k\|_X,$$

   $$\|\psi(h,k) - [f''(x_0)k]h\|_Y \le \varepsilon(2\|h\|_X + \|k\|_X)\|h\|_X.$$

   (ii) Si provi che per $\|h\|_X$ e $\|k\|_X$ abbastanza piccole si ha anche

   $$\|[f''(x_0)h]k - [f''(x_0)k]h\|_Y \le 2\varepsilon(\|h\|_X^2 + \|h\|_X\|k\|_X + \|k\|_X^2).$$

(iii) Si concluda che l'applicazione $f''(x_0) : \Omega \to \mathcal{L}_2(X, Y)$ è *simmetrica*, ossia
$$[f''(x_0)h]k = [f''(x_0)k]h \qquad \forall h, k \in X.$$

2. Siano X, Y, Z spazi normati. sia Ω un aperto di $X \times Y$ e sia $f : \Omega \to Z$ una funzione. Le *derivate parziali* $f_x(x_0, y_0)$ e $f_y(x_0, y_0)$ sono così definite:
$$f_x(x_0, y_0) = [f(\cdot, y_0)]'(x_0), \qquad f_y(x_0, y_0) = [f(x_0, \cdot)]'(y_0).$$
(i) Si verifichi che se f è F-differenziabile in (x_0, y_0) allora esistono $f_x(x_0, y_0)$ e $f_y(x_0, y_0)$, e si ha
$$f_x(x_0, y_0)h = [f'(x_0, y_0)](h, 0) \qquad \forall h \in X,$$
$$f_y(x_0, y_0)k = [f'(x_0, y_0)](0, k) \qquad \forall k \in Y.$$
(ii) Si provi che se f è di classe C^2, allora
$$[f_{xy}(x_0, y_0)h]k = [f''(x_0, y_0)(h, 0)](0, k) \qquad \forall (h, k) \in X \times Y,$$
$$[f_{yx}(x_0, y_0)k]h = [f''(x_0, y_0)(0, k)](h, 0) \qquad \forall (h, k) \in X \times Y.$$
(iii) Si deduca che se f è di classe C^2 le applicazioni bilineari $f_{xy}(x_0, y_0)$ e $f_{yx}(x_0, y_0)$ coincidono, cioè che
$$[f_{xy}(x_0, y_0)](h, k) = [f_{yx}(x_0, y_0)](k, h) \qquad \forall (h, k) \in X \times Y.$$

3. Sia $J : X \to \mathbb{R}$ un funzionale di classe C^2 definito sullo spazio normato X. Si scriva la derivata seconda del funzionale $E(x) = e^{J(x)}$.

3.6 Massimi e minimi relativi

Analizziamo brevemente alcune condizioni necessarie o sufficienti per l'esistenza di punti di massimo o di minimo relativo nel caso di funzionali definiti su spazi normati. Come vedremo, si tratta di facili generalizzazioni delle analoghe condizioni che valgono in $\mathbb{R}^N$.

Proposizione 3.6.1 *Sia X uno spazio normato, sia Ω un aperto di X e sia $f : \Omega \to \mathbb{R}$ una funzione G-differenziabile. Se $x_0 \in \Omega$ è un punto di massimo o di minimo locale per f, allora $f'_G(x_0) = 0$.*

Dimostrazione Supponiamo ad esempio che x_0 sia punto di massimo locale; allora esiste un intorno U di x_0 tale che

$$f(x) \leq f(x_0) \qquad \forall x \in U.$$

Se $v \in X$ allora si ha $x_0 + tv \in U$ per $|t| < \delta$, con δ sufficientemente piccolo; quindi

$$\frac{f(x_0 + tv) - f(x_0)}{t} \begin{cases} \geq 0 & \text{se} \quad -\delta < t < 0, \\ \leq 0 & \text{se} \quad 0 < t < \delta. \end{cases}$$

Per $t \to 0$ ne segue $f'_G(x_0)v = 0$ per ogni $v \in X$, cioè la tesi. $\square$

Questa condizione, benché soltanto necessaria per l'esistenza del minimo, è la più utile nelle applicazioni al calcolo delle variazioni che vedremo nel prossimo paragrafo.
Vediamo ora un'altra condizione legata al comportamento della derivata seconda.

Proposizione 3.6.2 *Sia X uno spazio normato, sia Ω un aperto di X e sia $f : \Omega \to \mathbb{R}$ una funzione di classe C^2. Se $x_0 \in \Omega$ è punto di massimo relativo, allora*

$$[f''(x_0)v]v \leq 0 \qquad \forall v \in X.$$

Se x_0 è punto di minimo relativo, si ha invece

$$[f''(x_0)v]v \geq 0 \qquad \forall v \in X.$$

Dimostrazione Supponiamo ad esempio che x_0 sia punto di massimo locale. Sia $\delta > 0$ tale che $B(x_0, \delta) \subseteq \Omega$. Fissato $v \in X \setminus \{0\}$, dalla formula di Taylor (proposizione 3.5.4), osservato che si ha $f'(x_0) = 0$ per la proposizione precedente, otteniamo

$$f(x_0 + tv) - f(x_0) - \frac{t^2}{2}[f''(x_0)v]v = \eta(tv) \qquad \forall t \in \left] -\frac{\delta}{\|v\|_X}, \frac{\delta}{\|v\|_X} \right[,$$

dove la funzione $\eta(x)/\|x\|_X^2$ è infinitesima per $\|x\|_X \to 0$. Dunque, se $|t|$ è sufficientemente piccolo si deduce

$$0 \geq f(x_0 + tv) - f(x_0) = t^2 \left[\frac{1}{2}[f''(x_0)v]v + \|v\|_X^2 \frac{\eta(tv)}{\|tv\|_X^2} \right].$$

Dividendo per t^2 e facendo tendere t a 0 si ottiene la tesi. $\square$

Questa condizione necessaria, lievemente rafforzata, diventa sufficiente.

Proposizione 3.6.3 *Sia X uno spazio normato, sia Ω un aperto di X e sia $f : \Omega \to \mathbb{R}$ una funzione di classe C^2. Se in un punto x_0 si ha $f'(x_0) = 0$ e $[f''(x_0)v]v \le -c\|v\|_X^2$ per ogni $v \in X$, con $c > 0$, allora x_0 è punto di massimo relativo per f. Se invece si ha $f'(x_0) = 0$ e $[f''(x_0)v]v \ge c\|v\|_X^2$ per ogni $v \in X$, allora x_0 è punto di minimo relativo per f.*

Dimostrazione Supponiamo che $f'(x_0) = 0$ e $[f''(x_0)v]v \le -c\|v\|_X^2$ per ogni $v \in X$, con $c > 0$. Dalla formula di Taylor (proposizione 3.5.4) segue che

$$\begin{aligned} f(x_0+v) - f(x_0) &= \frac{1}{2}[f''(x_0)v]v + \eta(v) \le \\ &\le -\frac{c}{2}\|v\|_X^2 \left[1 - \frac{2}{c}\,\frac{\eta(v)}{\|v\|_X^2}\right] \qquad \forall v \in X \setminus \{0\}. \end{aligned}$$

Poiché $\eta(v)/\|v\|_X^2$ è infinitesima per $\|v\|_X \to 0$, se δ è sufficientemente piccolo si deduce che $f(x_0+v) \le f(x_0)$ per $\|v\|_X < \delta$. Quindi x_0 è punto di massimo relativo. $\square$

Esempio 3.6.4 Riconsideriamo la funzione F dell'esempio 3.5.5: risulta

$$F'(0) = 0, \qquad F''(0) = -2I,$$

dove $I : H \to H$ è l'applicazione identità. Per la proposizione 3.6.3, l'origine è punto di massimo relativo per F (come era facile prevedere, ponendo $t = \|x\|$ e studiando la funzione $g(t) = t^4 - t^2$).

Incontreremo esempi più significativi più avanti nel corso, nell'ambito del calcolo della variazioni. Va comunque osservato che la condizione sufficiente fornita dalla proposizione 3.6.3 è troppo restrittiva per le applicazioni.

Esercizi 3.6

1. Sia $F : \ell^2 \to \mathbb{R}$ definito da

   $$F(x) = \sum_{n=1}^{\infty} \left(\frac{|x_n|^2}{n} - |x_n|^4 \right), \qquad x \in \ell^2.$$

   Si provi che si ha $F'(0) = 0$ e $\langle F''(0)u, u\rangle_{\ell^2} > 0$ per ogni $u \in \ell^2 \setminus \{0\}$, e che tuttavia 0 non è un punto di minimo relativo per F.

3.7 Funzioni implicite

Siano X, Y, Z spazi di Banach e sia $F : X \times Y \to Z$ un'applicazione. Ci poniamo il problema di sapere sotto quali ipotesi l'equazione

$$F(x, y) = 0$$

definisce implicitamente il grafico di una funzione $y : X \to Y$ tale che $F(x, y(x)) = 0$. La soluzione di questo problema si otterrà, come vedremo, sotto ipotesi del tutto analoghe a quelle classiche del teorema del Dini in $\mathbb{R}^2$ od in $\mathbb{R}^{n+m}$.

Teorema 3.7.1 (delle funzioni implicite) *Siano X, Y, Z spazi di Banach, sia $(x_0, y_0) \in X \times Y$ e sia $V \times W$ un intorno di tale punto. Sia poi $F : V \times W \to Z$ una funzione dotata delle seguenti proprietà:*

(i) *F è continua in (x_0, y_0) e $F(x_0, y_0) = 0$;*

(ii) *per ogni $x \in V$ esiste la derivata di Fréchèt $F_y(x, y)$ della funzione $y \to F(x, y)$ in ogni punto $y \in W$, ed inoltre $(x, y) \to F_y(x, y)$ è continua in (x_0, y_0);*

(iii) *esiste l'inversa $F_y(x_0, y_0)^{-1} \in \mathcal{L}(Z, Y)$.*

Allora esistono $\delta, \varepsilon > 0$ tali che per ogni $x \in B(x_0, \delta)$ l'equazione $F(x, y) = 0$ ha un'unica soluzione $y \in B(y_0, \varepsilon)$: in particolare, esiste $\psi : B(x_0, \delta) \to B(y_0, \varepsilon)$ per cui risulta

$$x \in B(x_0, \delta),\ y \in B(y_0, \varepsilon),\ F(x, y) = 0 \qquad \Longleftrightarrow \qquad y = \psi(x).$$

Inoltre ψ è continua in x_0; se poi F è anche continua in $V \times W$, allora ψ è continua in $B(x_0, \delta)$. Infine, se $F \in C^1(V \times W, Z)$ allora $\psi \in C^1(B(x_0, \delta), Y)$ e si ha

$$\psi'(x) = -F_y(x, \psi(x))^{-1} F_x(x, \psi(x)) \qquad \forall x \in B(x_0, \delta).$$

Osserviamo che, quando F è differenziabile secondo Fréchèt, gli operatori $F_y(x, y)$ e $F_x(x, y)$ sono le derivate parziali della funzione F, definite nell'esercizio 3.5.2, ed appartengono rispettivamente a $\mathcal{L}(Y, Z)$ e $\mathcal{L}(X, Z)$. Notiamo anche che se F è di classe C^1 l'ipotesi (iii) implica che gli operatori $F_y(x, y)$ hanno inversa continua per ogni (x, y) in un opportuno intorno di (x_0, y_0)

(esercizi 3.7.1 e 3.7.2).

Dimostrazione Anzitutto, non è restrittivo supporre che sia $V = B(x_0, \delta_0)$ e $W = B(y_0, \varepsilon_0)$. Per ogni $x \in B(x_0, \delta_0)$ consideriamo l'applicazione $T_{(x)} : B(y_0, \varepsilon_0) \to Y$ definita da

$$T_{(x)}(y) = y - F_y(x_0, y_0)^{-1} F(x, y).$$

L'equazione $F(x, y) = 0$ equivale a $T_{(x)}(y) = y$; dunque occorre cercare un punto fisso di $T_{(x)}$ per ogni x sufficientemente vicino a x_0. Osserviamo che $T_{(x)}$ è F-differenziabile in ogni punto $y \in B(y_0, \varepsilon_0)$ e

$$T'_{(x)}(y) = I - F_y(x_0, y_0)^{-1} F_y(x, y) = F_y(x_0, y_0)^{-1}[F_y(x_0, y_0) - F_y(x, y)];$$

quindi, per la continuità di F_y in (x_0, y_0), fissato $q \in]0, 1[$ esistono $\rho \in]0, \delta_0]$ ed $\varepsilon \in]0, \varepsilon_0[$ tali che

$$\|x - x_0\|_X < \rho, \quad \|y - y_0\|_Y \le \varepsilon \quad \Longrightarrow \quad \|T'_{(x)}(y)\|_{\mathcal{L}(Y)} \le q.$$

Di conseguenza, per la proposizione 3.3.2 si ha per ogni $y, y' \in \overline{B(y_0, \varepsilon)}$

$$\|T_{(x)}(y) - T_{(x)}(y')\|_Y \le \sup_{s \in [0,1]} \|T'_{(x)}((1-s)y + sy')\|_{\mathcal{L}(Y)} \|y - y'\|_Y \le q\|y - y'\|_Y \, .$$

Dunque $T_{(x)}$ è una contrazione in $\overline{B(y_0, \varepsilon)}$. Inoltre, per la continuità di F in (x_0, y_0), in corrispondenza di ε e q esiste $\sigma \in]0, \delta_0]$ tale che

$$\|x - x_0\|_X < \sigma \implies \|T_{(x)}(y_0) - y_0\|_Y = \|F_y(x_0, y_0)^{-1} F(x, y_0)\|_Y < \frac{1}{2}\,\varepsilon(1-q).$$

Sia allora $\delta = \min\{\rho, \sigma\}$: per ogni $x \in B(x_0, \delta)$, l'applicazione $T_{(x)}$ manda $\overline{B(y_0, \varepsilon)}$ in sé, anzi in $B(y_0, \varepsilon)$, poiché

$$\begin{aligned} \|T_{(x)}(y) - y_0\|_Y &\le \|T_{(x)}(y) - T_{(x)}(y_0)\|_Y + \|T_{(x)}(y_0) - y_0\|_Y \le \\ &\le q\|y - y_0\|_Y + \frac{1}{2}\,\varepsilon(1 - q) < \\ &< q\varepsilon + \varepsilon(1 - q) = \varepsilon \qquad \forall y \in \overline{B(y_0, \varepsilon)}. \end{aligned}$$

Pertanto, per il teorema delle contrazioni, per ogni $x \in B(x_0, \delta)$ l'applicazione $T_{(x)}$ ha un unico punto fisso $\psi(x) \in B(y_0, \varepsilon)$. In particolare, $T_{(x_0)}(y_0) = y_0$ e quindi $\psi(x_0) = y_0$. Inoltre, la funzione $\psi : B(x_0, \delta) \to B(y_0, \varepsilon)$ è continua in x_0 perché

$$\begin{aligned}
\|\psi(x) - \psi(x_0)\|_Y &= \|T_{(x)}(\psi(x)) - y_0\|_Y \leq \\
&\leq \|T_{(x)}(\psi(x)) - T_{(x)}(y_0)\|_Y + \|T_{(x)}(y_0) - y_0\|_Y \leq \\
&\leq q\|\psi(x) - y_0\|_Y + \|F_y(x_0, y_0)^{-1}F(x, y_0)\|_Y\,,
\end{aligned}$$

da cui

$$(1-q)\|\psi(x) - \psi(x_0)\|_Y \leq \|F_y(x_0, y_0)^{-1}F(x, y_0)\|_Y \to 0 \quad \text{per } \|x - x_0\|_X \to 0.$$

Supponiamo ora che F sia continua in $B(x_0, \delta) \times B(y_0, \varepsilon_0)$. Allora, con un calcolo analogo al precedente, per ogni $x, x' \in B(x_0, \delta)$ si ha

$$\begin{aligned}
&\|\psi(x) - \psi(x')\|_Y = \|T_{(x)}(\psi(x)) - T_{(x')}(\psi(x'))\|_Y \leq \\
&\quad \leq \|T_{(x)}(\psi(x)) - T_{(x)}(\psi(x'))\|_Y + \|T_{(x)}(\psi(x')) - T_{(x')}(\psi(x'))\|_Y \leq \\
&\quad \leq q\|\psi(x) - \psi(x')\|_Y + \|F_y(x_0, y_0)^{-1}\|_{\mathcal{L}(Z,Y)}\|F(x, \psi(x')) - F(x', \psi(x'))\|_Z
\end{aligned}$$

(si noti che, peraltro, $F(x', \psi(x')) = 0$); da qui, per $\|x - x'\|_X \to 0$ si ottiene

$$\begin{aligned}
&(1-q)\|\psi(x) - \psi(x')\|_Y \leq \\
&\quad \leq \|F_y(x_0, y_0)^{-1}\|_{\mathcal{L}(Z,Y)}\ \|F(x, \psi(x')) - F(x', \psi(x'))\|_Z \to 0.
\end{aligned}$$

Supponiamo infine che F sia di classe C^1. Allora $F_x(x, y) \in \mathcal{L}(X, Z)$ e $F_y(x, y) \in \mathcal{L}(Y, Z)$ per ogni $(x, y) \in B(x_0, \delta) \times B(y_0, \varepsilon_0)$, e possiamo supporre, come osservato in precedenza, che in ogni punto (x, y) di tale intorno risulti $F_y(x, y)^{-1} \in \mathcal{L}(Z, Y)$. Fissiamo $\overline{x} \in B(x_0, \delta)$ e poniamo per comodità $\overline{y} = \psi(\overline{x})$. Osserviamo che $F_x(\cdot, \overline{y})$ è limitata in un'opportuna palla $B(\overline{x}, \eta) \subseteq B(x_0, \delta)$, in virtù della relazione di continuità

$$\|F_x(x, \overline{y}) - F_x(\overline{x}, \overline{y}\|_{\mathcal{L}(X,Z)} \leq 1 \qquad \text{per } \|x - \overline{x}\|_X < \eta;$$

dalla stima precedente e dalla proposizione 3.3.2 segue allora

$$\begin{aligned}
&\|\psi(x) - \psi(\overline{x})\|_Y \leq \\
&\quad \leq \frac{C}{1-q} \sup_{s \in [0,1]} \|F_x((1-s)x + s\overline{x}, \psi(\overline{x}))\|_{\mathcal{L}(X,Z)}\|x - \overline{x}\|_X \leq \\
&\quad \leq K\|x - \overline{x}\|_X \qquad \forall x \in B(\overline{x}, \eta).
\end{aligned}$$

Per ipotesi, esiste $F_y(\overline{x},\overline{y})^{-1} \in \mathcal{L}(Z,Y)$; quindi, usando la F-differenziabilità di F in $(\overline{x},\overline{y})$, si ha

$$\begin{aligned}
&\left\|\psi(x) - \psi(\overline{x}) + \left[F_y(\overline{x},\overline{y})^{-1}F_x(\overline{x},\overline{y})\right](x-\overline{x})\right\|_Y = \\
&\quad = \left\|F_y(\overline{x},\overline{y})^{-1}\left[F_y(\overline{x},\overline{y})(\psi(x)-\psi(\overline{x})) + F_x(\overline{x},\overline{y})(x-\overline{x})\right]\right\|_Y \leq \\
&\quad \leq \|F_y(\overline{x},\overline{y})^{-1}\|_{\mathcal{L}(Z,Y)}\Big[\|F(x,\psi(x)) - F(\overline{x},\psi(\overline{x}))\|_Z + \\
&\qquad + \|\omega(x-\overline{x},\psi(x)-\psi(\overline{x}))\|_Z\Big].
\end{aligned}$$

Essendo

$$\begin{aligned}
\|\omega(x-\overline{x},\psi(x)-\psi(\overline{x}))\|_Z &= \mathrm{o}(\|x-\overline{x}\|_X + \|\psi(x)-\psi(\overline{x})\|_X) \leq \\
&\leq \mathrm{o}((1+K)\|x-\overline{x}\|_X),
\end{aligned}$$

tale relazione ci dice che ψ è F-differenziabile nel punto $\overline{x}$, e che la sua derivata è $F_y(\overline{x},\overline{y})^{-1}F_x(\overline{x},\overline{y})$; ma, per ipotesi e per l'esercizio 3.7.2, questa derivata è continua rispetto alla variabile $\overline{x}$, e ciò prova la tesi. $\square$

Osservazione 3.7.2 La formula che fornisce la derivata della funzione implicita ψ dipende solo da F_x, F_y e ψ stessa. Quindi, se si suppone che F sia di classe C^m in $V \times W$, $m \geq 1$, si può derivare tale formula, ottenendo così che anche ψ è di classe C^m, e ricavando una formula per la derivata $\psi^{(m)}$ in termini delle derivate di F fino all'ordine m e di ψ fino all'ordine $m-1$.

Esempio 3.7.3 Tra le numerose applicazioni del teorema delle funzioni implicite, ne illustriamo una di fondamentale importanza nello studio delle equazioni differenziali in spazi di Banach: la regolarità delle soluzioni rispetto ai dati iniziali. Consideriamo il problema di Cauchy

$$\begin{cases} u'(t) &= f(t,u(t)), \quad t \in [t_0,T], \\ u(t_0) &= u_0 \end{cases}$$

ove $u_0 \in X$ e $f : [t_0,T] \times X \to X$, essendo X uno spazio di Banach. Supponiamo che f sia continua, insieme con la sua derivata di Fréchèt rispetto alla seconda variabile, in $[t_0,T] \times X$; facciamo anche l'ipotesi che tale derivata sia uniformemente continua nella seconda variabile. Per questo problema di Cauchy si ha esistenza ed unicità locale della soluzione, ossia esiste $t_1 = t_1(u_0) \in]t_0,T]$, dipendente dalla scelta di u_0, ed esiste un'unica funzione $u \in C^1([t_0,t_1],X)$ per cui $u(t_0) = u_0$ e l'equazione è verificata in ogni punto di $[t_0, t_1]$: questo fatto si dimostra nello stesso modo con cui si prova il caso scalare, vale a dire trasformando il problema nell'equazione integrale equivalente

$$u(t) = u_0 + \int_{t_0}^{t} f(s, u(s))\, ds, \quad t \in [t_0, T],$$

e risolvendo quest'ultima in $[t_0, t_1]$ con il metodo delle approssimazioni successive.
Definiamo la seguente funzione F:

$$[F(u_0, u)](t) = u(t) - u_0 - \int_{t_0}^{t} f(s, u(s))\, ds, \quad t \in [t_0, \tau],$$

ove τ è un fissato punto di $]t_0, T]$. Questa funzione è definita su $X \times C^1([t_0, \tau], X)$, a valori in $C^1([t_0, \tau], X)$, ed è continua, come è facile verificare; essa è anche di classe C^1, in virtù dell'uniforme continuità di f_u, ed in particolare si ha

$$F_{u_0}(u_0, u)x = -x \qquad \forall x \in X,$$

$$[F_u(u_0, u)g](t) = g(t) - \int_{t_0}^{t} [f_u(s, u(s))]g(s)\, ds,\ t \in [t_0, \tau],\ \forall g \in C^1([t_0, \tau], X).$$

Verifichiamo che $F_u(u_0, u) : C^1([t_0, \tau], X) \to C^1([t_0, \tau], X)$ è invertibile. Fissata $g \in C^1([t_0, \tau], X)$, l'equazione $F_u(u_0, u)h = g$ (nell'incognita h) equivale all'equazione integrale

$$h(t) - \int_{t_0}^{t} [f_u(s, u(s))]h(s)\, ds = g(t), \quad t \in [t_0, \tau],$$

che a sua volta equivale al problema di Cauchy

$$\begin{cases} h'(t) &= f_u(t, u(t))h(t) + g'(t), \quad t \in [t_0, \tau], \\ h(t_0) &= g(t_0). \end{cases}$$

Questo problema è lineare, a coefficienti continui: con lo stesso metodo che si usa nel caso scalare, si dimostra che esso ha soluzione unica *globale*, ossia definita nello stesso intervallo in cui sono definiti i coefficienti (nel nostro caso, $[t_0, \tau]$), e dipendente con continuità dai dati. Esiste dunque un'unica $h \in C^1([t_0, \tau], X)$ tale che $F_u(u_0, u)h = g$, e la sua norma in $C^1([t_0, \tau], X)$ è controllata dalla norma del secondo membro g nello stesso spazio. Questo fatto ci dice che l'operatore $F_u(u_0, u)$ ha inverso continuo in $C^1([t_0, \tau], X)$.

Sono dunque verificate per la funzione F tutte le ipotesi del teorema delle funzioni implicite nell'intorno di un arbitrario punto $(\overline{u}_0, \overline{u})$ dello spazio $X \times C^1([t_0, \tau], X)$, nel quale risulti $F(\overline{u}_0, \overline{u}) = 0$: questa relazione equivale a dire che $\overline{u}$ verifica, nell' intervallo $[t_0, \tau]$,

$$\begin{cases} \overline{u}'(t) &= f(t, \overline{u}(t)), \quad t \in [t_0, \tau] \\ \overline{u}(t_0) &= \overline{u}_0 \,. \end{cases}$$

Pertanto, scelto $\tau = t_1(\overline{u_0})$, per il teorema 3.7.1 esistono un intorno V di $\overline{u}_0$ in X ed un intorno W di $\overline{u}$ in $C^1([t_0, t_1], X)$ tali che l'equazione $F(u_0, u) = 0$ ha soluzione unica in W per ogni $x \in V$, che dipende in modo C^1 da u_0. In altre parole, per ogni u_0 in V la soluzione u del problema di Cauchy

$$\begin{cases} u'(t) &= f(t, u(t)), \quad t \in [t_0, t_1], \\ u(t_0) &= u_0 \,, \end{cases}$$

che è appunto la funzione $u = \psi(u_0)$ definita implicitamente dall'equazione $F(u_0, u) = 0$, dipende in modo C^1 del dato iniziale u_0.

Osservazione 3.7.4 In modo analogo si può dimostrare che la soluzione del problema di Cauchy dipendente da un parametro

$$\begin{cases} u'(t) &= f(t, u(t), \alpha), \quad t \in [t_0, T], \quad \alpha \in \mathbb{R} \text{ fissato}, \\ u(t_0) &= u_0 \,, \end{cases}$$

sotto le ipotesi precedenti e supponendo inoltre che la derivata f_α sia continua, è differenziabile con continuità rispetto al parametro α.

Esercizi 3.7

1. Siano X, Y spazi di Banach. Provare che l'insieme

$$\mathcal{I} = \{A \in \mathcal{L}(X, Y) : \ \exists A^{-1} \in \mathcal{L}(Y, X)\}$$

   è aperto in $\mathcal{L}(X, Y)$.
   [**Traccia:** fissati $A, B \in \mathcal{I}$, si provi che $A^{-1} = B^{-1}(I - (A - B)A^{-1})$ e si deduca che se la norma $\|A - B\|_{\mathcal{L}(X,Y)}$ è sufficientemente piccola, allora $A^{-1}(I - (A - B)A^{-1})^{-1} = \sum_{n=0}^{\infty} A^{-1}[(A - B)A^{-1}]^n$ è effettivamente l'inverso di B.]

2. Siano X, Y, Z spazi di Banach e sia $A(z) \subseteq \mathcal{L}(X,Y)$ una famiglia di operatori dipendenti da una variabile $z \in Z$. Se $z \mapsto A(z)$ è di classe C^1 in una palla $B(z_0, R)$ e se esiste $A(z_0)^{-1} \in \mathcal{L}(Y,X)$, si provi che:

   (i) esiste $A(z)^{-1}$ in tutti i punti z di un'opportuna palla $B(z_0, r) \subseteq B(z_0, R)$;

   (ii) che l'applicazione $z \mapsto A(z)^{-1}$ è a sua volta di classe C^1 in $B(z_0, r)$, e che la derivata di tale applicazione in un fissato punto $z \in B(z_0, r)$ è l'elemento di $\mathcal{L}(Z, \mathcal{L}(Y,X))$ così definito:

$$(A(\cdot)^{-1})'(z)h = A(z_0)^{-1}[A'(z)h]A(z_0)^{-1} \qquad \forall h \in Z.$$

3. Con riferimento all'enunciato del teorema 3.7.1, sia $\psi : B(x_0, \delta) \to B(y_0, \varepsilon)$ la funzione definita implicitamente dall'equazione $F(x,y) = 0$. Supponendo F di classe C^2, si calcoli $\psi''(x)$ in termini delle derivate prime e seconde di F, nonché di ψ e ψ'.

4. Con riferimento alla dimostrazione del teorema 3.7.1, si definisca per ogni $n \in \mathbb{N}$

$$\begin{cases} y_0(x) = y_0 \\ y_{n+1}(x) = T_{(x)}(y_n(x)), \quad n \in \mathbb{N}, \end{cases} \qquad x \in B(x_0, \delta);$$

   si provi che per ogni $x \in B(x_0, \delta)$ la successione $\{y_n(x)\}_{n\in\mathbb{N}}$ converge in Y al punto fisso dell'applicazione $T_{(x)}$.

5. Si dimostri il risultato di esistenza ed unicità locale per il problema di Cauchy

$$\begin{cases} u'(t) &= f(t, u(t)), \quad t \in [t_0, T], \\ u(t_0) &= u_0 \end{cases}$$

   in uno spazio di Banach X. Si provi anche che se il secondo membro è lineare o affine rispetto a u, la soluzione è globale.

3.8 Varietà tangente ad una curva di livello

Dall'Analisi II sappiamo che se F è una funzione reale di classe C^1 definita su $\mathbb{R}^2$, e Z è l'insieme dei suoi zeri oppure una sua qualunque altra curva di livello non vuota, allora in ogni punto $P_0 = (x_0, y_0) \in Z$ nel quale il gradiente di F sia non nullo esiste la retta r tangente a Z in P_0. L'equazione di r è

$$F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0) = 0,$$

ovvero $\langle \nabla F(P_0), P - P_0 \rangle_2 = 0$; in altre parole, si ha $P \in r$ se e solo se il vettore $P - P_0$ appartiene al nucleo dell'applicazione lineare $v \mapsto \nabla F(P_0) \cdot v$. La retta r è l' *unica* fra quelle passanti per P_0 per la quale risulti

$$\lim_{P \to P_0,\, P \in Z} \frac{d(P, r)}{|P - P_0|_2} = 0.$$

Una situazione analoga si ha, sotto opportune ipotesi, nel caso di una funzione $F : X \to Y$ di classe C^1, allorché X e Y sono spazi di Banach.

Teorema 3.8.1 (di Lyusternik) *Siano X ed Y spazi di Banach, sia $F : X \to Y$ una applicazione di classe C^1 e poniamo $Z_0 = \{x \in X : F(x) = 0\}$. Fissato $x_0 \in Z_0$, supponiamo che il sottospazio $T_0 = \ker F'(x_0)$ sia non banale e definiamo la varietà affine $T_{x_0} = x_0 + T_0$. Se l'operatore lineare $F'(x_0) : X \to Y$ è surgettivo, allora valgono le seguenti proprietà:*

(i) $\displaystyle \lim_{x \to x_0,\, x \in Z_0} \frac{d(x, T_{x_0})}{\|x - x_0\|_X} = 0,$

(ii) $\displaystyle \lim_{v \to x_0,\, v \in T_{x_0}} \frac{d(v, Z_0)}{\|v - x_0\|_X} = 0.$

Se, in particolare, il sottospazio chiuso T_0 è complementabile *in X, ossia esiste un sottospazio chiuso $S_0 \subseteq X$ tale che $X = T_0 \oplus S_0$, allora esiste un intorno U di x_0 ed un omeomorfismo $\gamma : T_{x_0} \cap U \to Z_0 \cap U$, di classe C^1, tali che*

$$\lim_{v \to x_0,\, v \in T_{x_0}} \frac{\|v - \gamma(v)\|_X}{\|v - x_0\|_X} = 0, \qquad \lim_{x \to x_0,\, x \in Z_0} \frac{\|x - \gamma^{-1}(x)\|_X}{\|x - x_0\|_X} = 0.$$

La varietà affine T_{x_0} è dunque "tangente" alla curva di livello Z_0 nel punto x_0. Si osservi che se X è uno spazio di Hilbert allora T_0 è sempre complementabile: basta prendere $S_0 = T_0^{\perp}$.
Osserviamo che il teorema di Lyusternik è falso quando $\ker F'(x_0) = \{0\}$. Ad esempio, se $Y = X$, $x_0 = 0$ e $F(x) = x$ per ogni $x \in X$, si ha $T_0 = \ker F'(x_0) = \ker I = \{0\}$ e $Z_0 = \{0\}$: in questo caso né (i), né (ii)

hanno senso.

Dimostrazione Ci limitiamo per semplicità a provare la tesi nel caso complementabile (per il caso generale si vedano gli esercizi 3.8.1 e 3.8.2).
Sia dunque $T_0 \neq \{0\}$ e $X = T_0 \oplus S_0$. Osserviamo che se $x = h + g$ con $h \in T_0$ e $g \in S_0$, la norma $\|x\|_X$ è equivalente alla norma $\|h\|_X + \|g\|_X$: infatti l'applicazione $(h,g) \mapsto h+g$ è lineare, continua e bigettiva da $T_0 \times S_0$ su X, e dunque per il teorema dell'applicazione aperta essa è un isomorfismo.
Per ogni $x \in X$ scriviamo $x - x_0 = h + g$, con $h \in T_0$ e $g \in S_0$. Allora l'equazione $F(x) = 0$ equivale a

$$\Phi(h,g) \stackrel{\text{def}}{=} F(x_0 + h + g) = 0,$$

ed è soddisfatta per $(h,g) = (0,0)$. Vogliamo applicare a questa equazione il teorema delle funzioni implicite (teorema 3.7.1). Osserviamo che Φ è di classe C^1, che $\Phi'_g(0,0) = F'(x_0)|_{S_0}$ e che tale applicazione è bigettiva: infatti se $g \in S_0$ e $F'(x_0)g = 0$, allora $g \in S_0 \cap T_0 = \{0\}$. D'altra parte, fissato $y \in Y$, possiamo scegliere, in virtù dell'ipotesi di surgettività, un elemento $v \in X$ tale che $F'(x_0)v = y$: scrivendo $v = h + g$, con $h \in T_0$ e $g \in S_0$, otteniamo perciò $y = F'(x_0)(h+g) = F'(x_0)g = F'(x_0)|_{S_0}g$. In definitiva, $F'(x_0)|_{S_0}$ è un isomorfismo fra S_0 e Y.
Per il teorema delle funzioni implicite, esistono un intorno V di 0 in T_0, un intorno W di 0 in S_0 ed un'applicazione $\psi : V \to W$ di classe C^1, tali che $\psi(0) = 0$ e

$$\Phi(h,g) = 0, \quad (h,g) \in V \times W \qquad \Longleftrightarrow \qquad g = \psi(h).$$

Inoltre si ha $\psi'(h) = -\Phi_g(h,\psi(h))^{-1}\Phi_h(h,\psi(h))$ per ogni $h \in V$; in particolare

$$\psi'(0) = -\Phi_g(0,0)^{-1}F'(x_0)|_{T_0} = 0.$$

Siano $U = x_0 + V + W$ e $\Lambda = x_0 + V$: U è un intorno di x_0 in X, mentre Λ è intorno di x_0 in T_{x_0}. Consideriamo l'applicazione $\gamma : \Lambda \to X$ definita da

$$\gamma(v) = v + \psi(v - x_0), \qquad v \in \Lambda.$$

Posto $x = \gamma(v)$, risulta $x \in U \cap Z_0$: infatti, detto $h = v - x_0$, si ha $h \in V$, da cui $\psi(h) \in W$, $x = v + \psi(v - x_0) = x_0 + h + \psi(h) \in U$ e $F(x) = \Phi(h,\psi(h)) = 0$. In definitiva γ manda Λ in $U \cap Z_0$ ed è continua, come composizione di funzioni continue. Si ha anche, in particolare, $\gamma(x_0) = x_0 + \psi(0) = x_0$.

Verifichiamo che γ è bigettiva. Se $\gamma(v) = \gamma(v')$, allora posto $h = v - x_0$ e $h' = v' - x_0$ abbiamo $x_0 + h + \psi(h) = x_0 + h' + \psi(h')$, da cui $h - h' = \psi(h') - \psi(h) \in T_0 \cap S_0 = \{0\}$; ne segue $h = h'$ e dunque $v = v'$. Se poi $x \in U \cap Z_0$, si ha $x = x_0 + h + g$, con $h \in V$ e $g \in W$, e $F(x) = \Phi(h, g) = 0$: quindi $g = \psi(h)$, e detto $v = x_0 + h$ otteniamo $x = v + \psi(v - x_0) = \gamma(v)$.
Quindi, esiste γ^{-1}; proviamo che γ^{-1} è continua. Se $x = x_0 + h + g$ e $x' = x_0+h'+g'$ sono punti di $U\cap Z_0$, allora $\gamma^{-1}(x) = h-x_0$ e $\gamma^{-1}(x') = h'-x_0$, da cui

$$\|\gamma^{-1}(x) - \gamma^{-1}(x')\|_X = \|h - h'\|_X \leq \|h - h'\|_X + \|g - g'\|_X \leq c\|x - x'\|_X.$$

Infine, per $v \in \Lambda$ e $v \to x_0$ si ha

$$\begin{aligned} \|\gamma(v) - v\|_X &= \|\psi(v - x_0)\|_X = \\ &= \|\psi(v - x_0) - \psi(0) - \psi'(0)(v - x_0)\|_X = \mathrm{o}(\|v - x_0\|_X), \end{aligned}$$

mentre per $x \in U \cap Z_0$ e $x \to x_0$, posto $v = \gamma^{-1}(x)$, si ha anche $v \to x_0$ ed inoltre

$$\begin{aligned} &\frac{\|\gamma^{-1}(x) - x\|_X}{\|x - x_0\|_X} = \frac{\|v - \gamma(v)\|_X}{\|x - x_0\|_X} = \\ &\quad = \frac{\|v - \gamma(v)\|_X}{\|v - x_0\|_X} \cdot \frac{\|\gamma^{-1}(x) - \gamma^{-1}(x_0)\|_X}{\|x - x_0\|_X} \leq c\,\frac{\mathrm{o}(\|v - x_0\|_X)}{\|v - x_0\|_X}. \end{aligned}$$

Ciò prova l'ultima parte dell'enunciato. □

Esempio 3.8.2 Siano $X = L^2(a, b)$, $Y = \mathbb{R}$ e $F(g) = \|g\|_2^2$ per ogni $g \in X$. Fissata $g \in X$ con $\|g\|_2^2 = \lambda > 0$, determiniamo la varietà tangente alla curva di livello $Z = \{f \in X : F(f) = \lambda\}$. La derivata di Fréchèt $F'(g)$ è

$$F'(g)h = 2\langle g, h\rangle_2 \qquad \forall h \in X,$$

e tale applicazione è ovviamente surgettiva da X su $\mathbb{R}$. Si ha

$$\ker F'(g) = [\{g\}]^\perp$$

e dunque, per il teorema di Lyusternik,

$$T_g = g + \ker F'(g) = g + [\{g\}]^\perp,$$

il che era prevedibile data la simmetria circolare della curva di livello Z.

Esercizi 3.8

1. Si dimostri l'enunciato (i) del teorema 3.8.1 nel caso generale.
[**Traccia:** si introduca in X la relazione di equivalenza $h_1 \simeq h_2 \iff h_1 - h_2 \in T_0$, e si osservi che $F'(x_0)$ induce un operatore $A : X/T_0 \to Y$, definito da $A[h] = F'(x_0)k$ per ogni $k \in [h]$, il quale è bigettivo e continuo: dunque $A^{-1} \in \mathcal{L}(Y, X/T_0)$. Ciò premesso, sia $\varepsilon > 0$ e sia $\delta > 0$ tale che per $\|x - x_0\|_X < \delta$ risulti
$$\|F(x) - F(x_0) - F'(x_0)(x - x_0)\|_X \le \frac{\varepsilon}{2\|A^{-1}\|_{\mathcal{L}(Y,X/T_0)}}\|x - x_0\|_X \,;$$
se ne deduca, per $x = x_0 + d \in Z_0$,
$$\|d\|_X \le \delta \quad \Longrightarrow \quad \|[d]\|_{X/T_0} \le \frac{\varepsilon}{2}\|d\|_X \,,$$
da cui
$$\|d\|_X \le \delta, \quad g \in [d] \cap B(0, 2\|[d]\|_{X/T_0}) \quad \Longrightarrow \quad \|g\|_X \le \varepsilon \|d\|_X \,.$$
Da qui si ricavi la tesi.]

2. Si dimostri l'enunciato (ii) del teorema 3.8.1 nel caso generale.
[**Traccia:** posto $c = \|A^{-1}\|_{\mathcal{L}(Y,X/T_0)}$, si fissi $\varepsilon \in \left]0, \frac{1}{4c}\right]$, e si determini $\eta > 0$ tale che si abbia: (a) $\|F'(x) - F'(x_0)\|_{\mathcal{L}(X,Y)} \le \varepsilon$ per $\|x - x_0\|_X \le \eta$, e (b) per $\|k\|_X, \|h\|_X \le \eta$ risulti
$$\|F(x_0 + k) - F(x_0 + h) - F'(x_0)(k - h)\|_X \le \varepsilon \|k - h\|_X \,.$$
Ciò premesso, fissato $v = x_0 + h \in T_{x_0}$, si ponga $g_0 = 0$, $[h_1] = -A^{-1}F(x_0 + h)$, e si scelga $g_1 \in [h_1]$ con $\|g_1\|_X \le 2\|[h_1]\|_{X/T_0}$; poi, noti $[h_i]$ e g_i per $i = 1, \ldots, n$, si ponga $[h_{n+1}] = [h_n] - A^{-1}F(x_0 + h + g_n)$ e si scelga $g_{n+1} \in [h_{n+1}]$ con $\|g_{n+1} - g_n\|_X \le 2\|[h_{n+1}] - [h_n]\|_{X/T_0}$. Si mostri che esiste $\delta \in \left]0, \frac{\eta}{2}\right]$ tale che $\|g_1\|_X \le \min\{\varepsilon, \frac{1}{2}\}\|h\|_X$; poi si dimostrino per induzione le seguenti disuguaglianze:
$$\|g_n\|_X \le \sum_{k=0}^{n-1} 2^{-k}\|g_1\|_X \,,$$
$$\|[h_{n+1}] - [h_n]\|_{X/T_0} \le 2^{-n-1}\|g_1\|_X \,, \quad \|g_{n+1} - g_n\|_X \le 2^{-n}\|g_1\|_X \,.$$
Si concluda che esistono $g \in X$ e $[d] \in X/T_0$ tali che $\|g_n - g\|_X \to 0$ e $\|[h_n] - [d]\|_{X/T_0} \to 0$. Si deduca infine che $x = x_0 + h + g \in Z_0$ e che $\|v - x\|_X = \|g\|_X \le 2\varepsilon\|h\|_X = 2\varepsilon\|v - x_0\|_X$.]

3.9 Massimi e minimi vincolati

Siano X ed Y spazi di Banach e siano $F : X \to \mathbb{R}$, $\Phi : X \to Y$ funzioni assegnate. Ci poniamo il problema di trovare il minimo o il massimo di F sul vincolo costituito dall'insieme $Z = \{x \in X : \Phi(x) = 0\}$. Otterremo risultati analoghi al caso dei vincoli in $\mathbb{R}^N$. Il risultato che segue è una versione generalizzata del metodo dei moltiplicatori di Lagrange.

Teorema 3.9.1 *Siano X ed Y spazi di Banach reali, siano $F : X \to \mathbb{R}$ e $\Phi : X \to Y$ funzioni di classe C^1 e sia $Z = \{x \in X : \Phi(x) = 0\}$. Fissiamo poi $x_0 \in Z$ e supponiamo che l'applicazione lineare $\Phi'(x_0)$ abbia immagine chiusa in Y. Se x_0 è punto di massimo relativo, o di minimo relativo, per $F|_Z$, allora esistono $\lambda_0 \in \mathbb{R}$ e $\varphi \in Y^*$, non entrambi nulli, tali che*

$$\lambda_0 F'(x_0) + \Phi'(x_0)^* \varphi = 0.$$

Si osservi che l'uguaglianza sopra scritta è fra elementi di X^*; in particolare, $\Phi'(x_0)^*$ è l'operatore aggiunto di $\Phi'(x_0)$, definito da

$$\Phi'(x_0)^* \psi = \psi \circ \Phi'(x_0) \qquad \forall \psi \in Y^*.$$

Nelle applicazioni ha interesse soprattutto il caso in cui $\lambda_0 \neq 0$, che si realizza, come vedremo dalla dimostrazione, quando $\Phi'(x_0)$ è un'applicazione surgettiva.
Nel caso particolare in cui $X = \mathbb{R}^N$ e $Y = \mathbb{R}^k$ $(N > k)$, il funzionale $\varphi \in (\mathbb{R}^k)^*$ è della forma $x \mapsto \langle \lambda, x \rangle_k$, con $\lambda \in \mathbb{R}^k$; la derivata $F'(x_0) \in (\mathbb{R}^N)^*$ si esprime come $x \mapsto \langle \nabla F(x_0), x \rangle_N$, e la derivata $\Phi'(x_0) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ è la matrice Jacobiana $(k \times N)$ di Φ nel punto x_0. Quindi la tesi del teorema si riduce al fatto che esistono $\lambda_0, \lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tali che

$$\lambda_0 \frac{\partial F}{\partial x_j}(x_0) + \sum_{i=1}^{k} \lambda_i \frac{\partial \Phi^i}{\partial x_j}(x_0) = 0, \qquad j = 1, \ldots, N.$$

Se la matrice Jacobiana di Φ in x_0 ha rango massimo, allora Z è una varietà, e in particolare $\Phi'(x_0)$ è surgettiva. In questo caso, come si è detto, possiamo scegliere $\lambda_0 \neq 0$: si ritrova così l'usuale regola dei moltiplicatori di Lagrange.

Dimostrazione Sia K l'immagine di $\Phi'(x_0)$: per ipotesi, K è un sottospazio chiuso di Y. Distinguiamo i due casi $K \neq Y$ e $K = Y$.

Se $K \neq Y$, per il teorema di Hahn-Banach esiste $\varphi \in Y^* \setminus \{0\}$ tale che $\varphi|_K = 0$. Ne segue

$$[\Phi'(x_0)^*\varphi](x) = \varphi(\Phi'(x_0)x) = 0 \qquad \forall x \in X,$$

e quindi si ha la tesi del teorema scegliendo $\lambda_0 = 0$.
Supponiamo invece $K = Y$: osserviamo che se $\ker \Phi'(x_0) = \{0\}$ allora $\Phi'(x_0)$ è un isomorfismo fra X e Y, quindi $\Phi'(x_0)^*$ è un isomorfismo fra Y^* e X^*: pertanto scelto $\varphi = [\Phi'(x_0)^*]^{-1}F'(x_0)$ si ha subito $-F'(x_0) + \Phi'(x_0)^*\varphi = 0$ e la tesi è provata con $\lambda_0 = -1$. Altrimenti sarà $\ker \Phi'(x_0) \neq \{0\}$: in tal caso applichiamo a Φ il teorema di Lyusternik (teorema 3.8.1). Posto $T_{x_0} = x_0 + \ker \Phi'(x_0)$, e fissato $\varepsilon > 0$, per la condizione (ii) si ha

$$\lim_{v \to x_0,\, v \in T_{x_0}} \frac{d(v, Z)}{\|v - x_0\|_X} = 0.$$

Scegliamo $v = x_0 + th$, ove $t \in \mathbb{R} \setminus \{0\}$ e h è un fissato elemento di $\ker \Phi'(x_0)$ con norma unitaria; allora

$$\lim_{t \to 0} \frac{d(x_0 + th, Z)}{|t|} = 0.$$

Per ogni fissato $t \neq 0$ esiste $x_t \in Z$ tale che, posto $r(t) = x_0 + th - x_t$, risulta $\|r(t)\|_X = \|x_0 + th - x_t\|_X < d(x_0 + th, Z) + t^2$; quindi si ha anche

$$\lim_{t \to 0} \frac{\|r(t)\|_X}{|t|} = 0.$$

Di conseguenza $r(0) = 0$ e $r'(0) = 0$. Sia ora

$$f(t) = F(x_t) = F(x_0 + th - r(t)), t \in \mathbb{R} :$$

per l'ipotesi fatta su F la funzione $f : \mathbb{R} \to \mathbb{R}$ ha un punto di massimo o di minimo relativo in $t = 0$, e pertanto $0 = f'(0) = F'(x_0)(h - r'(0)) = F'(x_0)h$. Se ne deduce, per linearità,

$$F'(x_0)h = 0 \qquad \forall h \in \ker \Phi'(x_0).$$

A questo punto ci serve un lemma di analisi funzionale.

Lemma 3.9.2 *Siano X, Y spazi di Banach e sia $A \in \mathcal{L}(X, Y)$ surgettivo. Allora*

$$R(A^*) = \{\eta \in X^* : \ \eta x = 0 \ \ \forall x \in \ker A\}.$$

Prima di dimostrare il lemma, concludiamo la dimostrazione del teorema 3.9.1: poiché $F'(x_0)$ ha valore nullo in ogni punto di $\ker \Phi'(x_0)$, il lemma ci garantisce che $F'(x_0) \in R(\Phi'(x_0)^*)$, e quindi esiste $\varphi \in Y^*$ tale che $F'(x_0) = \Phi'(x_0)^*\varphi$. Ne segue la tesi del teorema scegliendo $\lambda_0 = -1$. $\Box$

Dimostrazione del lemma 3.9.2 Sia $\eta \in R(A^*)$: allora $\eta = A^*\psi$, con $\psi \in Y^*$, da cui $\eta x = \psi(Ax) = \psi(0) = 0$ per ogni $x \in \ker A$.
Viceversa, sia $\eta \in X^*$ tale che $\eta x = 0$ per ogni $x \in \ker A$; cerchiamo $\psi \in Y^*$ tale che $\eta = A^*\psi$. Per ipotesi, $\ker \eta \supseteq \ker A$; fissato $y \in Y$, e scelti $x, x' \in X$ tali che $Ax = Ax' = y$ (tali punti esistono perché A è surgettivo), avremo $x - x' \in \ker A$ e dunque $\eta x = \eta x'$. In altri termini, η è costante sull'insieme $A^{-1}(y)$. È lecito allora definire $\psi : Y \to \mathbb{R}$ nel modo seguente:

$$\psi(y) = \eta x \qquad \forall x \in A^{-1}(y),$$

ed ovviamente ψ è un funzionale lineare. Proviamo che ψ è continuo: se G è un aperto di $\mathbb{R}$, si ha

$$\begin{aligned} \psi^{-1}(G) &= \{y \in Y : \psi y \in G\} = \{Ax : x \in X,\ \eta x \in G\} = \\ &= A\{x \in X : \eta x \in G\} = A\eta^{-1}(G). \end{aligned}$$

Poiché η è continua, $\eta^{-1}(G)$ è un aperto di X, mentre $A\eta^{-1}(G)$ è aperto in Y in quanto A è un'applicazione aperta (essendo lineare, continua e surgettiva). Ne segue che $\psi^{-1}(G)$ è aperto in Y e pertanto ψ è continuo, ossia $\psi \in Y^*$. Per definizione, si ha $\eta x = \psi(Ax) = A^*\psi(x)$ per ogni $x \in X$, cioè $\eta = A^*\psi$.
$\Box$

Osservazione 3.9.3 La condizione necessaria per l'esistenza di un massimo o di un minimo vincolato espressa dal teorema 3.9.1 (nel caso $\lambda_0 = -1$) dice che è nullo il differenziale di Fréchèt in $X \times Y^*$ della funzione

$$H(x, \varphi) = F(x) + \varphi(\Phi(x)).$$

Infatti si verifica senza difficoltà che

$$\begin{aligned} H'(x,\varphi)(h,\psi) &= F'(x)h + \varphi(\Phi'(x)h) + \psi(\Phi(x)) = \\ &= [F'(x) + \Phi'(x)^*\varphi]h + \psi(\Phi(x)), \end{aligned}$$

quindi tale differenziale è nullo se e solo se

$$\begin{cases} [F'(x) + \Phi'(x)^*\varphi]h = 0 & \forall h \in X \\ \psi(\Phi(x)) = 0 \qquad \forall \psi \in Y^*, \end{cases}$$

ossia se e solo se

$$\begin{cases} F'(x) + \Phi'(x)^*\varphi = 0 \\ \Phi(x) = 0. \end{cases}$$

Esempio 3.9.4 Sia H uno spazio di Hilbert. Vogliamo determinare i punti di massimo e di minimo della funzione $F(x) = \langle x, y\rangle_H$ sull'insieme

$$K = \{x \in H : \ \|x\|_H = 1, \ \langle x, z\rangle_H = 0\},$$

ove y, z sono fissati elementi di H.
Posto $X = H$, $Y = \mathbb{R}^2$ e

$$\Phi : H \to \mathbb{R}^2, \qquad \Phi(x) = \big(\|x\|_H^2 - 1, \langle x, z\rangle_H\big) \quad \forall x \in H,$$

il nostro problema è quello di trovare il massimo ed il minimo di F sul vincolo $K = \{x \in H : \Phi(x) = 0\}$. Naturalmente si può supporre che z ed y siano linearmente indipendenti (cosicché in particolare $z \neq 0$), poiché in caso contrario avremmo $F|_K = 0$ e quindi ci sarebbe ben poco da dire.
Questo problema si risolve facilmente con considerazioni geometriche, ma vogliamo mostrare come la sua risoluzione sia possibile tramite il teorema 3.9.1. Per applicare tale teorema occorre verificare che l'immagine di $\Phi'(x)$ è chiusa in $\mathbb{R}^2$; in effetti non è difficile verificare che ciò è vero, ed anzi, poiché $z \neq 0$, l'applicazione $\Phi'(x)$ è surgettiva per ogni $x \in K$ in virtù dell'esercizio 3.9.3.
Sia allora x un punto di massimo relativo o di minimo relativo per $F|_K$: per il teorema 3.9.1 esistono $\alpha \in \mathbb{R}$ e $\varphi \in (\mathbb{R}^2)^* = \mathbb{R}^2$ tali che $\alpha F'(x) + \Phi'(x)^*\varphi = 0$, e dalla dimostrazione del teorema 3.9.1 segue che si può scegliere $\alpha = -1$; in altre parole, posto $\varphi = (\lambda, \mu)$, si ha

$$-\langle y, v\rangle_H + 2\lambda\langle x, v\rangle_H + \mu\langle z, v\rangle_H = 0 \qquad \forall v \in H,$$

il che implica (scrivendo λ in luogo di 2λ)

$$y = \lambda x + \mu z.$$

Dai due vincoli $\|x\|_H^2 = 1$ e $\langle x, z\rangle_H = 0$ segue subito, essendo $z \neq 0$, che

$$\lambda = \langle y, x\rangle_H, \qquad \mu = \frac{\langle y, z\rangle_H}{\|z\|_H^2},$$

e di conseguenza, con facili calcoli,

$$x = \pm \frac{y - \frac{\langle y,z\rangle_H}{\|z\|_H^2}\, z}{\left\| y - \frac{\langle y,z\rangle_H}{\|z\|_H^2}\, z \right\|_H}.$$

È anche facile verificare che, come era prevedibile, i punti trovati sono addirittura punti di minimo e di massimo assoluti per F sull'insieme K.

Esercizi 3.9

1. Sia H uno spazio di Hilbert reale e sia $A \in \mathcal{L}(H)$. Posto $f(x) = \langle Ax, x\rangle_H$, si provi che se x_0 è punto di massimo o di minimo relativo per la restrizione di f all'insieme $B = \{x \in H : \|x\|_H = 1\}$, allora esiste $\lambda \in \mathbb{R}$ tale che $Ax_0 + A^* x_0 = 2\lambda x_0$.

2. Nelle stesse ipotesi dell'esercizio precedente, supponiamo anche che A sia autoaggiunto. Si provi che se x_0 è punto di massimo assoluto per f su B, allora $Ax_0 = \|A\|_{\mathcal{L}(H)} \cdot x_0$, cosicché $\max_B f = \|A\|_{\mathcal{L}(H)}$.

3. Sia H uno spazio di Hilbert e siano $y, z \in H$. Si provi che l'applicazione $\psi : H \to \mathbb{R}^2$, definita da $\psi(v) = (\langle v, y\rangle_H, \langle v, z\rangle_H)$, ha immagine chiusa; si provi che essa coincide con $\mathbb{R}^2$ se e solo se i vettori y, z sono linearmente indipendenti.

Capitolo 4

Calcolo delle variazioni

4.1 Funzionali del calcolo delle variazioni

In questo capitolo ci limiteremo ad alcuni cenni su una problematica che è vastissima ed alla quale sono dedicati interi scaffali di biblioteche ed innumerevoli articoli di ricerca. Il problema principale del calcolo delle variazioni è la minimizzazione di funzionali dipendenti da variabili che si muovono in spazi di funzioni; la scelta degli spazi in cui cercare la soluzione di un problema di minimo è in genere parte integrante del problema stesso.
Limitandoci al caso di funzioni di una sola variabile, considereremo soprattutto il seguente problema tipico: si vuole minimizzare un funzionale della forma

$$J(y) = \int_a^b F(x, y(x), y'(x))\, dx$$

nella classe delle $y \in C^1[a, b]$ soggette (eventualmente) a certe condizioni negli estremi a, b. Dunque la classe X delle funzioni ammissibili è un sottospazio (lineare od affine) di $C^1[a, b]$. Supporremo, come ipotesi minima, che l'integrando F sia di classe C^1 nei suoi argomenti $(x, y, p) \in [a, b] \times \mathbb{R}^2$. Senza grosse complicazioni, incontreremo anche funzionali che dipendono da funzioni a valori in $\mathbb{R}^m$: in tal caso la F sarà di classe C^1 in $[a, b] \times \mathbb{R}^{2m}$.
Verifichiamo anzitutto che nelle ipotesi fatte, il funzionale J è differenziabile.

Proposizione 4.1.1 *Sia* $J(y) = \int_a^b F(x, y(x), y'(x))\, dx$ *per* $y \in C^1[a, b]$. *Se* $F \in C^1([a, b] \times \mathbb{R}^2)$, *allora il funzionale* J *è* F*-differenziabile in* $C^1[a, b]$, *e* *per ogni* $v \in C^1[a,b]$ *si ha*

$$J'(y)v = \int_a^b \Big[F_y(x, y(x), y'(x)) \cdot v(x) + F_p(x, y(x), y'(x)) \cdot v'(x)\Big] dx.$$

Dimostrazione Sia $y \in C^1[a,b]$. Per ogni $v \in C^1[a,b]$ possiamo scrivere

$$\begin{aligned}
J(y+v) - J(y) &= \\
&= \int_a^b \Big[F(x, y(x)+v(x), y'(x)+v'(x)) - F(x, y(x), y'(x))\Big] dx = \\
&= \int_a^b \int_0^1 \frac{d}{dt} F(x, y(x)+tv(x), y'(x)+tv'(x))\, dt dx = \\
&= \int_a^b \int_0^1 \Big[F_y(x, y(x)+tv(x), y'(x)+tv'(x)) \cdot v(x) + \\
&\quad +F_p(x, y(x)+tv(x), y'(x)+tv'(x)) \cdot v'(x)\Big] dt dx;
\end{aligned}$$

applicando il teorema della media, per ogni $x \in [a,b]$ esiste un opportuno $\tau_x \in]0,1[$ tale che l'ultimo membro si può riscrivere come

$$\begin{aligned}
\int_a^b &\Big[F_y(x, y(x)+\tau_x v(x), y'(x)+\tau_x v'(x)) \cdot v(x) + \\
&+F_p(x, y(x)+\tau_x v(x), y'(x)+\tau_x v'(x)) \cdot v'(x)\Big] dx.
\end{aligned}$$

Pertanto possiamo decomporre l'incremento $J(y+v) - J(y)$ nella forma

$$\begin{aligned}
J(y&+v) - J(y) = \\
&= \int_a^b \Big[F_y(x, y(x), y'(x)) \cdot v(x) + F_p(x, y(x), y'(x)) \cdot v'(x)\Big] dx + \\
&+ \int_a^b \Big[[F_y(x, y(x)+\tau_x v(x), y'(x)+\tau_x v'(x)) - F_y(x, y(x), y'(x))] \cdot v(x) + \\
&+[F_p(x, y(x)+\tau_x v(x), y'(x)+\tau_x v'(x)) - F_p(x, y(x), y'(x))] \cdot v'(x)\Big] dx.
\end{aligned}$$

Verifichiamo che il secondo integrale è infinitesimo per $\|v\|_{C^1[a,b]} \to 0$. Le derivate di F sono per ipotesi continue; fissato $M > 2\|y\|_{C^1[a,b]}$, esse sono uniformemente continue nel compatto $[a,b] \times [-M,M] \times [-M,M]$. Pertanto, dato $\varepsilon > 0$, esiste $\delta \in]0, M/2[$ per cui, quando $\|v\|_{C^1[a,b]} < \delta$, si ha

$$\left| \int_a^b \Big[[F_y(x, y(x) + \tau_x v(x), y'(x) + \tau_x v'(x)) - F_y(x, y(x), y'(x))] \cdot v(x) + \right.$$

$$\left. + [F_p(x, y(x) + \tau_x v(x), y'(x) + \tau_x v'(x)) - F_p(x, y(x), y'(x))] \cdot v'(x) \Big] dx \right| \le$$

$$\le \varepsilon \int_a^b [|v(x)| + |v'(x)|] dx \le \varepsilon (b-a) \|v\|_{C^1[a,b]} .$$

Ciò prova la tesi. □

Osservazioni 4.1.2 (1) Nel caso in cui $F \in C^1([a,b] \times \mathbb{R}^{2m})$, la derivata di Fréchèt di J assume esattamente la stessa forma, con l'unica differenza che nell'integrale i prodotti $F_y \cdot v + F_p \cdot v'$ diventano prodotti scalari in $\mathbb{R}^m$.
(2) Se $F \in C^2([a,b] \times \mathbb{R}^2)$, con procedimento analogo si può dimostrare che il funzionale J è differenziabile due volte e che

$$\begin{aligned} [J''(y)v]w &= \int_a^b \Big[F_{yy}(x, y(x), y'(x)) v(x) w(x) + \\ &+ F_{yp}(x, y(x), y'(x)) [v(x) w'(x) + w(x) v'(x)] + \\ &+ F_{pp}(x, y(x), y'(x)) v'(x) w'(x) \Big] dx. \end{aligned}$$

Un analogo risultato vale nel caso in cui $F \in C^2([a,b] \times \mathbb{R}^{2m})$ (esercizio 4.1.1).

Esercizi 4.1

1. Dimostrare che se $F \in C^2([a,b] \times \mathbb{R}^{2m})$ allora il funzionale

$$J(y) = \int_a^b F(x, y(x), y'(x))\, dx, \qquad y \in C^1([a,b], \mathbb{R}^m)$$

è differenziabile due volte e che

$$\begin{aligned} [J''(y)v]w &= \int_a^b \Big[\langle F_{yy}(x, y(x), y'(x)) v(x), w(x) \rangle_m + \\ &+ \langle F_{yp}(x, y(x), y'(x)) v(x), w'(x) \rangle_m + \\ &+ \langle F_{yp}(x, y(x), y'(x)) w(x), v'(x) \rangle_m + \\ &+ \langle F_{pp}(x, y(x), y'(x)) v'(x), w'(x) \rangle_m \Big] dx. \end{aligned}$$

2. Fissata una curva regolare $\gamma_0 \in C^1[a,b],\mathbb{R}^N)$, si provi che il funzionale "lunghezza d'arco" $J(\gamma) = \int_a^b |\gamma'(t)|_N\,dt$ è F-differenziabile due volte nello spazio $C^1([a,b],\mathbb{R}^N)$ per $\gamma = \gamma_0$, mentre è semicontinuo inferiormente, ma non continuo, in $C^1[a,b],\mathbb{R}^N)$ rispetto alla topologia indotta dalla norma uniforme.

4.2 Lemmi fondamentali del calcolo delle variazioni

Per trovare i punti di minimo del funzionale J introdotto nel paragrafo precedente, cercheremo anzitutto di utilizzare la condizione necessaria espressa dalla proposizione 3.6.1. A questo scopo, è chiaro che le migliori informazioni per la determinazione dei punti di minimo si hanno quando si è in grado di dedurre condizioni necessarie legate all'integrando F. Per procurarci queste informazioni sono utilissimi i seguenti lemmi, globalmente denominati "lemmi fondamentali del calcolo delle variazioni".

Lemma 4.2.1 *Sia $\alpha \in C([a,b],\mathbb{R}^N)$ tale che*

$$\int_a^b \langle \alpha(x), \varphi(x)\rangle_N\,dx = 0 \qquad \forall \varphi \in C_0^1([a,b],\mathbb{R}^N);$$

allora $\alpha = 0$ in $[a,b]$.

Ricordiamo che $C_0^1([a,b],\mathbb{R}^N) = \{\varphi \in C^1([a,b],\mathbb{R}^N) : \varphi(a) = \varphi(b) = 0\}$.

Dimostrazione Basta provare, in virtù della continuità di α, che $\alpha(x) = 0$ per ogni $x \in]a,b[$. Se in un punto $x_0 \in]a,b[$ fosse $\alpha(x_0) \neq 0$, allora esiste una componente $i \in \{1,\ldots,N\}$ tale che $\alpha_i(x_0) \neq 0$ (e possiamo supporre che sia $\alpha_i(x_0) > 0$). Quindi, sempre per continuità, esiste $[h,k] \subseteq [a,b]$ tale che $x_0 \in]h,k[$ e $\alpha_i(x) > 0$ per ogni $x \in [h,k]$.
Scegliamo una funzione $\varphi \in C_0^1([a,b],\mathbb{R}^N)$ nel modo seguente: la componente i-esima di φ è

$$\varphi_i(x) = \begin{cases} (x-h)^2(k-x)^2 & \text{se } x \in [h,k], \\ 0 & \text{se } x \in [a,h[\cup]k,b], \end{cases}$$

mentre se $j \neq i$, poniamo $\varphi_j(x) = 0$ per ogni $x \in [a,b]$. È chiaro che $\varphi \in C_0^1[a,b]$. Allora dall'ipotesi segue

$$0 = \int_a^b \langle \alpha(x), \varphi(x)\rangle_N\,dx = \int_h^k \alpha_i(x)(x-h)^2(k-x)^2 dx > 0,$$

il che è assurdo. Dunque α è nulla su $[a,b]$. $\square$

Lemma 4.2.2 *Sia $\alpha \in C([a,b],\mathbb{R}^N)$ tale che*

$$\int_a^b \langle \alpha(x), \varphi'(x)\rangle_N \, dx = 0 \qquad \forall \varphi \in C_0^1([a,b],\mathbb{R}^N);$$

allora α è costante in $[a,b]$.

Dimostrazione Definiamo

$$c = \frac{1}{b-a}\int_a^b \alpha(\xi)\, d\xi, \qquad \varphi(x) = \int_a^x [\alpha(\xi) - c]\, d\xi, \quad x \in [a,b];$$

allora $\varphi \in C_0^1([a,b],\mathbb{R}^N)$ per la scelta di c. Poiché $\varphi'(x) = \alpha(x) - c$, si ha dall'ipotesi

$$\int_a^b |\alpha(x) - c|_N^2 \, dx = \int_a^b \langle \alpha(x) - c, \varphi'(x)\rangle_N \, dx = 0 - \langle c, \varphi(b) - \varphi(a)\rangle_N = 0,$$

da cui la tesi. $\square$

Lemma 4.2.3 *Siano $\alpha, \beta \in C([a,b],\mathbb{R}^N)$ tali che*

$$\int_a^b [\langle \alpha(x), \varphi(x)\rangle_N + \langle \beta(x), \varphi'(x)\rangle_N]\, dx = 0 \qquad \forall \varphi \in C_0^1([a,b],\mathbb{R}^N);$$

allora $\beta \in C^1([a,b],\mathbb{R}^N)$ e $\beta' = \alpha$ in $[a,b]$.

Dimostrazione Poniamo $A(x) = \int_a^x \alpha(\xi)\, d\xi$, $x \in [a,b]$; allora per ogni $\varphi \in C_0^1([a,b],\mathbb{R}^N)$ si ha, usando l'ipotesi ed integrando per parti,

$$0 = \int_a^b [\langle \alpha(x), \varphi(x)\rangle_N + \langle \beta(x), \varphi'(x)\rangle_N]\, dx = \int_a^b \langle -A(x) + \beta(x), \varphi'(x)\rangle_N \, dx;$$

dal lemma 4.2.2 segue allora che $-A + \beta$ è costante in $[a,b]$. Dunque, $\beta = A + c \in C^1([a,b],\mathbb{R}^N)$ e $\beta' = A' = \alpha$ in $[a,b]$. $\square$

Definiamo lo spazio

$$C_0^2([a,b],\mathbb{R}^N) = \{\varphi \in C^2([a,b],\mathbb{R}^N) : \ \varphi(a) = \varphi(b) = \varphi'(a) = \varphi'(b) = 0\}.$$

I due lemmi che seguono costituiscono la generalizzazione dei due precedenti.

Lemma 4.2.4 *Sia $\alpha \in C([a,b],\mathbb{R}^N)$ tale che*

$$\int_a^b \langle \alpha(x), \varphi''(x)\rangle_N \, dx = 0 \qquad \forall \varphi \in C_0^2([a,b],\mathbb{R}^N);$$

allora α è un polinomio di grado non superiore a 1 in $[a,b]$.

Dimostrazione Poniamo

$$\varphi(x) = \int_a^x (x-\xi)[\alpha(\xi) - c - d(\xi - a)]\, d\xi, \quad x \in [a,b],$$

e cerchiamo $c, d \in \mathbb{R}^N$ tali che $\varphi \in C_0^2([a,b],\mathbb{R}^N)$. Con facili calcoli si trova che occorre scegliere

$$c = -\frac{2}{(b-a)^2}\left[\int_a^b \alpha(\xi)\, d\xi - \frac{3}{b-a}\int_a^b (b-\xi)\alpha(\xi)\, d\xi\right],$$

$$d = \frac{6}{(b-a)^2}\left[\int_a^b \alpha(\xi)\, d\xi - \frac{2}{b-a}\int_a^b (b-\xi)\alpha(\xi)\, d\xi\right].$$

Con questa scelta si ottiene

$$\begin{aligned}
&\int_a^b |\alpha(x) - c - d(x-a)|_N^2 \, dx = \\
&\quad = \int_a^b \langle \alpha(x) - c - d(x-a), \varphi''(x)\rangle_N \, dx = \\
&\quad = 0 - \langle c, \varphi'(b) - \varphi'(a)\rangle_N - (b-a)\langle d, \varphi'(b)\rangle_N + \langle d, \varphi(b) - \varphi(a)\rangle_N = 0,
\end{aligned}$$

cioè la tesi. $\square$

Lemma 4.2.5 *Siano $\alpha, \beta, \gamma \in C([a,b],\mathbb{R}^N)$ tali che risulti, per ogni $\varphi \in C_0^2([a,b],\mathbb{R}^N)$,*

$$\int_a^b [\langle \alpha(x), \varphi(x)\rangle_N + \langle \beta(x), \varphi'(x)\rangle_N + \langle \gamma(x), \varphi''(x)\rangle_N]\, dx = 0;$$

allora $\gamma \in C^1([a,b],\mathbb{R}^N)$, $\beta - \gamma' \in C^1([a,b],\mathbb{R}^N)$ e $(\beta - \gamma')' = \alpha$ in $[a,b]$.

Dimostrazione Definiamo

$$A(x) = \int_a^x (x-\xi)\alpha(\xi)\,d\xi, \quad B(x) = \int_a^x \beta(\xi)\,d\xi, \qquad x \in [a,b].$$

Dall'ipotesi, integrando per parti, segue

$$\begin{aligned} 0 &= \int_a^b [\langle \alpha(x), \varphi(x)\rangle_N + \langle \beta(x), \varphi'(x)\rangle_N + \langle \gamma(x), \varphi''(x)\rangle_N]\,dx = \\ &= \int_a^b \langle A(x) - B(x) + \gamma(x), \varphi''(x)\rangle_N\,dx \quad \forall \varphi \in C_0^2([a,b], \mathbb{R}^N). \end{aligned}$$

Per il lemma 4.2.4, si ha $\gamma(x)+A(x)-B(x) = c+d(x-a)$ per ogni $x \in [a,b]$; quindi, per differenza, $\gamma \in C^1([a,b],\mathbb{R}^N)$ e $\gamma' = d - A' + \beta$ in $[a,b]$. Da qui ricaviamo $\gamma' - \beta = d - A' \in C^1([a,b],\mathbb{R}^N)$ e $(\gamma'-\beta)' = -\alpha$ in $[a,b]$. Ne segue la tesi. □

Esercizi 4.2

1. Si enunci e si dimostri la generalizzazione del lemma 4.2.4 al caso dell'equazione

$$\int_a^b \langle \alpha(x), \varphi^{(n)}(x)\rangle_N\,dx = 0 \qquad \forall \varphi \in C_0^n([a,b],\mathbb{R}^N),$$

ove

$$\begin{aligned} C_0^n([a,b],\mathbb{R}^N) &= \{\varphi \in C^n([a,b],\mathbb{R}^N) : \\ &\qquad \varphi^{(k)}(a) = \varphi^{(k)}(b) = 0 \text{ per } k = 0,1,\ldots,n-1\}. \end{aligned}$$

2. Si enunci e si dimostri la generalizzazione del lemma 4.2.5 al caso dell'equazione

$$\int_a^b \sum_{k=0}^n \langle \alpha_k(x), \varphi^{(k)}(x)\rangle_N\,dx = 0 \qquad \forall \varphi \in C_0^n([a,b],\mathbb{R}^N).$$

3. Si dimostri il lemma 4.2.1 supponendo $\alpha \in L^1(a,b)$, anziché $\alpha \in C([a,b],\mathbb{R}^N)$.
[**Traccia:** fissato $\varepsilon > 0$, si fissi $g \in C_0[a,b]$ tale che $\|\alpha - g\|_1 < \varepsilon$ e se ne

deduca che $\left|\int_a^b g\varphi\,dt\right| < \varepsilon\|\varphi\|_\infty$ per ogni $\varphi \in C_0[a,b]$). Per stimare $\|g\|_1$, si definisca $K_1 = \{t \in [a,b] : g(t) \geq 1\}$, $K_2 = \{t \in [a,b] : g(t) \leq -1\}$ e $K = K_1 \cup K_2$. Si scelga una funzione $\psi \in C_0[a,b]$ tale che $|\psi| \leq 1$, $\psi = 1$ su K_1 e $\psi = -1$ su K_2: osservato che in $[a,b] \setminus K$ si ha $|g\psi| \leq |g| < \varepsilon$, si provi che $\|g\|_1 \leq \varepsilon + 2(b-a)\varepsilon$, e si concluda che $\|\alpha\|_1 < 2(1+b-a)\varepsilon$ per ogni $\varepsilon > 0$.]

4. Si dimostri il lemma 4.2.1 supponendo $\alpha \in L^1(a,b;\mathbb{R}^N)$, anziché $\alpha \in C([a,b],\mathbb{R}^N)$.
   [**Traccia:** si utilizzi l'esercizio precedente.]

5. Si dimostrino i lemmi 4.2.2, 4.2.3, 4.2.4 e 4.2.5 supponendo le funzioni α, β, γ in $L^1(a,b;\mathbb{R}^N)$ anziché in $C([a,b],\mathbb{R}^N)$.

6. Sia X uno spazio di Banach e siano $u, v \in L^p(a,b;X)$, $1 < p < \infty$. Se risulta
$$\int_a^b [u(t)\varphi(t) + v(t)\varphi'(t)]\,dt = 0 \qquad \forall \varphi \in C_0^\infty(]a,b[,\mathbb{R}),$$
   si provi che v è derivabile q.o. e che $v'(t) = -u(t)$ q.o. in $[a,b]$.
   [**Traccia:** si fissi $\psi \in X^*$ e si osservi che $\int_a^b [\psi u(t)\varphi(t) + \psi v(t)\varphi'(t)]\,dt = 0$ per ogni $\varphi \in C_0^\infty(]a,b[,\mathbb{R})$; se ne deduca che $\psi(\int_s^b u(t)dt + v(s) - v(b)) = 0$ per ogni $s \in [a,b]$. Fissato s e scelto opportunamente ψ, si mostri che $v(s) + \int_s^b u(t)dt = v(b)$ e si ricavi la tesi.]

4.3 Equazione di Eulero

Veniamo al risultato principale del calcolo delle variazioni: i punti di minimo o di massimo relativo del funzionale J sono soluzioni di una equazione differenziale che è legata soltanto all'integrando F.

Teorema 4.3.1 *Sia $F \in C^1([a,b] \times \mathbb{R}^2)$ e sia J il funzionale definito da*
$$J(y) = \int_a^b F(x, y(x), y'(x))\,dx$$
nella classe $X = \{y \in C^1[a,b] : y(a) = A,\ y(b) = B\}$. Se y_0 è un punto di minimo relativo, o di massimo relativo, per J in X, allora y_0 è soluzione

dell' equazione differenziale

$$\frac{d}{dx}F_p(x,y(x),y'(x)) = F_y(x,y(x),y'(x)) \quad \text{in } [a,b],$$

che è detta equazione di Eulero *del funzionale J.*

Dimostrazione Supponiamo che y_0 sia punto di minimo locale. Possiamo scrivere, in virtù della proposizione 4.1.1,

$$J'(y_0)v = \int_a^b \Big[F_y(x,y_0(x),y_0'(x))\cdot v(x) + F_p(x,y_0(x),y_0'(x))\cdot v'(x)\Big]$$

per ogni $v \in C^1[a,b]$. Poiché in particolare si ha $y_0 + v \in X$ per ogni $v \in C_0^1[a,b]$, ripetendo l'argomentazione usata per provare la proposizione 3.6.1 si ricava

$$J'(y_0)v = \int_a^b \Big[F_y(x,y_0(x),y_0'(x))\cdot v(x) + F_p(x,y_0(x),y_0'(x))\cdot v'(x)\Big]dx = 0$$

per ogni $v \in C_0^1[a,b]$. Dal lemma 4.2.3 segue allora la tesi. $\square$

Osservazioni 4.3.2 (1) Le soluzioni dell'equazione di Eulero di un funzionale J si dicono *estremali* del funzionale J nella classe X. Un punto $y_0 \in X$ è un'estremale per J se e solo se $J'(y_0)v = 0$ per ogni $v \in C_0^1[a,b]$. Non è detto, in generale, che un'estremale sia punto di massimo o di minimo relativo per J.

(2) Se però il funzionale J, oltre che F-differenziabile, è anche *convesso*, allora ogni estremale di J è punto di minimo *assoluto* per J: infatti in virtù dell'esercizio 3.3.5 si ha

$$J(y_0+v) \geq J(y_0) + J'(y_0)v = J(y_0) \qquad \forall v \in C_0^1[a,b].$$

Se inoltre J è strettamente convesso, il punto di minimo è unico.
Una condizione che assicura la convessità di J è la convessità dell'integrando F nelle variabili $(y,p) \in \mathbb{R}^2$ (esercizio 4.3.5).

(3) Quando $F \in C^1([a,b],\mathbb{R}^{2m})$, l'equazione di Eulero è un sistema di m equazioni della forma

$$\frac{d}{dx}F_{p_i}(x,y(x),y'(x)) = F_{y_i}(x,y(x),y'(x)) \quad \text{in } [a,b], \quad i = 1,\dots,m.$$

Esempio 4.3.3 Dati due punti $A, B \in \mathbb{R}^N$, e fissati $a, b \in \mathbb{R}$ con $a < b$, vogliamo determinare la curva regolare $\gamma : [a, b] \to \mathbb{R}^N$, con sostegno avente per estremi i punti A e B, che ha lunghezza minima; chiaramente, ci aspettiamo di trovare il segmento di estremi A e B. Dobbiamo determinare il punto di minimo assoluto del funzionale

$$J(\gamma) = \int_a^b |\gamma'|_N \, dt$$

nella classe $X = \{\gamma \in C^1([a, b], \mathbb{R}^N) : |\gamma'|_N \neq 0, \ \gamma(a) = A, \ \gamma(b) = B\}$. Poiché si suppone $|\gamma'|_N \neq 0$, la funzione integranda $F(t, y, p) = |p|_N$ è di classe C^1, anzi C^∞, ed in particolare $F_{p_i}(t, y, p) = p_i/|p|_N$ e $F_{y_i} = 0$. Dunque, se una curva γ realizza il minimo di J, per il teorema 4.3.1 essa deve soddisfare l'equazione di Eulero, cioè il sistema seguente:

$$\frac{d}{dt} \frac{\gamma_i'(t)}{|\gamma'(t)|_N} = 0 \quad \forall t \in [a, b], \qquad i = 1, \dots, N,$$

da cui

$$\frac{\gamma_i'(t)}{|\gamma'(t)|_N} = c_i \quad \forall t \in [a, b], \qquad i = 1, \dots, N.$$

Questa relazione esprime il fatto che il versore tangente alla curva è lo stesso in ogni punto e vale $c = (c_1, \dots, c_N)$; ovviamente, $|c|_N = 1$. Del resto, integrando fra a e b troviamo

$$B - A = \int_a^b \gamma'(t) \, dt = c \int_a^b |\gamma'(t)|_N \, dt = c \, J(\gamma),$$

da cui, come era da aspettarsi, $J(\gamma) = |B - A|_N$, e di conseguenza

$$c = \frac{B - A}{|B - A|_N} \, .$$

Dunque il versore tangente alla curva ha in ogni punto la direzione di $B - A$; quindi il sostegno di γ è il segmento AB. Le equazioni parametriche di γ sono

$$\gamma(t) = A + \frac{B - A}{|B - A|_N} \int_a^t |\gamma'(\tau)|_N \, d\tau \qquad \forall t \in [a, b].$$

Si noti che la nostra impostazione del problema fornisce infinite curve soluzioni, tutte parametrizzate su $[a, b]$ ed aventi come sostegno il segmento di

estremi A e B; il passaggio dall'una all'altra avviene tramite cambiamenti di parametro $\phi : [a,b] \to [a,b]$ di classe C^1, i quali non alterano il funzionale lunghezza ma solo la velocità istantanea di percorrenza del segmento AB.
In questo caso, la condizione necessaria ci ha permesso di trovare i punti di minimo. Si noti che non abbiamo *dimostrato* che le curve γ del tipo sopra scritto siano punti di minimo del funzionale J; tuttavia questo fatto (molto verosimile, del resto!) segue osservando che J è convesso e ricordando l'osservazione 4.3.2(2). Più semplicemente ancora, esso può essere verificato nel modo seguente: per ogni curva $\psi : [a,b] \to \mathbb{R}^N$ di estremi A, B, si ha

$$J(\psi) = \int_a^b |\psi'(t)|_N \, dt \geq \left| \int_a^b \psi'(t) \, dt \right|_N = |\psi(b) - \psi(a)|_N = |B - A|_N,$$

mentre per γ valeva l'uguaglianza.

Esercizi 4.3

1. Trovare le estremali dei seguenti funzionali in $C^1[a,b]$:

   (i) $J(y) = \int_a^b [y(x)^2 + y'(x)^2 + 2y(x)e^x] \, dx$,

   (ii) $J(y) = \int_a^b [y(x)^2 - y'(x)^2 - 2y(x) \cosh x] \, dx$,

   (iii) $J(y) = \int_a^b y'(x)[1 + x^2 y'(x)] \, dx$,

   (iv) $J(y) = \int_a^b [x^2 y'(x)^2 + 2y(x)^2 + 2xy(x)] \, dx$.

2. Risolvere l'equazione di Eulero del funzionale

   $$J(y) = \int_a^b f(x)\sqrt{1 + y'(x)^2} \, dx, \qquad y \in C^1[a,b],$$

   ove $f(x) > 0$ è una funzione assegnata, considerando in particolare i casi $f(x) = \sqrt{x}$ e $f(x) = x$.

3. Determinare, se esiste, il minimo del funzionale

   $$\int_0^1 \left[\frac{1}{2} \, y'(x)^2 + y(x)y'(x) + y'(x) + y(x) \right] dx$$

   in $C^1[0,1]$ ed in $\{y \in C^1[0,1] : y(0) = y(1) = 0\}$.

4. *(Problemi ad estremi variabili)* Si provi che se y_0 è un'estremale per il funzionale
$$J(y) = \int_a^b F(x, y(x), y'(x))\, dx$$
nella classe $X = \{y \in C^1[a,b] : y(a) \in [A_1, A_2], y(b) \in [B_1, B_2]\}$, allora y_0 verifica, oltre all'equazione di Eulero, le condizioni agli estremi
$$F_p(a, y_0(a), y_0'(a)) = 0, \qquad F_p(b, y_0(b), y_0'(b)) = 0.$$

5. Si provi che se $F(x, y, p)$ è di classe C^1 ed è convessa rispetto a (y, p), allora il funzionale $J(y) = \int_a^b F(x, y(x), y'(x))dx$ è convesso. Si provi anche che la stretta convessità di F in (y, p) implica la stretta convessità di J.

4.4 Integrali primi

Spesso lo studio e la risoluzione dell'equazione di Eulero di un funzionale sono tutt'altro che agevoli. In certi casi particolari, tuttavia, la struttura dell'equazione è tale da semplificare la situazione: ciò accade ad esempio allorché l'equazione possiede un "integrale primo", cioè si riesce a ricavare un'equazione differenziale più semplice della quale ogni estremale deve essere soluzione. Vale a questo proposito la seguente

Proposizione 4.4.1 *Sia $J(y) = \int_a^b F(x, y(x), y'(x))\, dx$, con $F \in C^1([a,b] \times \mathbb{R}^2)$. Se l'integrando F è di classe C^2 e non dipende dalla variabile x, ossia $F = F(y, p) \in C^2(\mathbb{R}^2)$, allora ogni estremale y di J verifica anche l'equazione*
$$F(y(x), y'(x)) - y'(x)F_p(y(x), y'(x)) = c \qquad \forall x \in [a, b],$$
ove c è una costante. Si dice che l'equazione qui scritta è un integrale primo *dell'equazione di Eulero.*

Si noti che la presenza di un integrale primo permette di abbassare il grado dell'equazione differenziale da risolvere.

Dimostrazione Sia y un'estremale: dunque vale l'equazione di Eulero
$$\frac{d}{dx} F_p(y(x), y'(x)) = F_y(y(x), y'(x)).$$

Osserviamo anzitutto che se in un punto $x \in [a, b]$ si ha $F_{pp}(y(x), y'(x)) \neq 0$, allora in quel punto esiste la derivata seconda $y''(x)$. Infatti, utilizzando il teorema di Lagrange,

$$\begin{aligned}
F_y(y(x), y'(x)) &= \lim_{h \to 0} \frac{1}{h}[F_p(y(x+h), y'(x+h)) - F_p(y(x), y'(x))] = \\
&= \lim_{h \to 0} \frac{1}{h}[F_{py}(\xi_h, \eta_h)(y(x+h) - y(x)) + F_{pp}(\xi'_h, \eta'_h)(y'(x+h) - y'(x))]
\end{aligned}$$

ove (ξ_h, η_h), (ξ'_h, η'_h) sono punti opportuni nel segmento di estremi $(y(x), y'(x))$ e $(y(x+h), y'(x+h))$. Nell'uguaglianza precedente, quando $h \to 0$ il primo termine all'ultimo membro converge a $F_{py}(y(x), y'(x))y'(x)$; quindi anche il secondo converge: indichiamo il suo limite con $G(x)$. Siccome $F_{pp}(y(x), y'(x)) \neq 0$, si ha che

$$\exists y''(x) = \lim_{h \to 0} \frac{y'(x+h) - y'(x)}{h} = \frac{F_y(y(x), y'(x)) - F_{py}(y(x), y'(x))y'(x)}{F_{pp}(y(x), y'(x))},$$

cosicché $G(x) = y''(x)F_{pp}(y(x), y'(x))$ nei punti in cui $F_{pp}(y(x), y'(x)) \neq 0$. Ciò prova l'asserzione precedente. In definitiva, vale l'equazione

$$\begin{aligned}
0 &= F_y(y(x), y'(x)) - \frac{d}{dx} F_p(y(x), y'(x)) = \\
&= F_y(y(x), y'(x)) - F_{py}(y(x), y'(x))y'(x) - G(x)
\end{aligned}$$

in *ogni* punto $x \in [a, b]$.
D'altra parte, calcolando il rapporto incrementale della quantità

$$F(y(x), y'(x)) - y'(x)F_p(y(x), y'(x)),$$

si ottiene, decomponendo opportunamente,

$$\begin{aligned}
&\frac{1}{h}\Big[F(y(x+h),y'(x+h)) - y'(x+h)F_p(y(x+h),y'(x+h)) - \\
&\qquad -F(y(x),y'(x)) + y'(x)F_p(y(x),y'(x))\Big] = \\
&\quad = \frac{1}{h}\left[F(y(x+h),y'(x+h)) - F(y(x),y'(x+h))\right] + \\
&\qquad + \frac{1}{h}\left[F(y(x),y'(x+h)) - F(y(x),y'(x))\right] - \\
&\qquad - \frac{1}{h}\left[y'(x+h) - y'(x)\right]F_p(y(x+h),y'(x+h)) - \\
&\qquad - \frac{1}{h}\,y'(x)[F_p(y(x+h),y'(x+h)) - F_p(y(x),y'(x+h))] - \\
&\qquad - \frac{1}{h}\,y'(x)[F_p(y(x),y'(x+h)) - F_p(y(x),y'(x))] = \\
&\quad = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}.
\end{aligned}$$

Quando $h \to 0$ si ha:

$$\mathrm{I} = \mathrm{o}(1) + F_y(y(x),y'(x))y'(x),$$
$$\mathrm{IV} = \mathrm{o}(1) - F_{py}(y(x),y'(x))y'(x)^2;$$

inoltre, in virtù del teorema di Lagrange,

$$\mathrm{II} = \mathrm{o}(1) + F_p(y(x),y'(x))\ \frac{y'(x+h)-y'(x)}{h},$$
$$\mathrm{III} = \mathrm{o}(1) - F_p(y(x),y'(x))\ \frac{y'(x+h)-y'(x)}{h},$$

ed infine

$$\mathrm{V} = \mathrm{o}(1) - G(x)y'(x),$$

quantità che ha sempre senso per quanto osservato in precedenza. Si conclude allora, sommando i cinque termini e ricordando l'equazione precedentemente ottenuta, che

$$\begin{aligned}
&\exists \frac{d}{dx}[F(y(x),y'(x)) - y'(x)F_p(y(x),y'(x))] = \\
&\quad = F_y(y(x),y'(x))y'(x) - F_{py}(y(x),y'(x))y'(x)^2 - G(x)y'(x) = \\
&\quad = 0 \qquad \forall x \in [a,b],
\end{aligned}$$

e ciò prova la tesi. $\square$

Esempio 4.4.2 Vogliamo trovare una curva cartesiana $y \in C^1[a,b]$, strettamente positiva, con valori fissati negli estremi, la cui rotazione attorno all'asse x generi una superficie di area minima. Fissati dunque $A, B > 0$, cerchiamo una funzione $y(x)$ che renda minimo il funzionale

$$J(y) = 2\pi \int_a^b y(x)\sqrt{1 + y'(x)^2}\, dx$$

nella classe

$$X = \{y \in C^1[a,b] : \ y(a) = A, \ y(b) = B, \ y(x) > 0 \ \ \forall x \in [a,b]\}.$$

Se $y \in X$, allora quando $v \in C_0^1[a,b]$ si ha $y + v \in X$, purché la norma $\|v\|_{C^1[a,b]}$ sia sufficientemente piccola. Inoltre, se y_0 è punto di minimo per J, varrà l'equazione di Eulero: quindi, essendo

$$F(x,y,p) = y\sqrt{1+p^2}, \qquad F_p(x,y,p) = \frac{yp}{\sqrt{1+p^2}}, \qquad F_y = \sqrt{1+p^2},$$

avremo

$$\frac{d}{dx}\frac{y(x)y'(x)}{\sqrt{1+y'(x)^2}} = \sqrt{1+y'(x)^2} \qquad \forall x \in [a,b].$$

Questa equazione è piuttosto complicata: fortunatamente però l'integrando F non dipende dalla variabile x e quindi, per la proposizione 4.4.1, si ha l'integrale primo

$$y(x)\sqrt{1+y'(x)^2} - \frac{y(x)y'(x)^2}{\sqrt{1+y'(x)^2}} = c \qquad \forall x \in [a,b],$$

ossia, semplificando,

$$\frac{y(x)}{\sqrt{1+y'(x)^2}} = c \qquad \forall x \in [a,b].$$

Se $c = 0$ otteniamo $y(x) = 0$, non ammissibile. Se invece $c \neq 0$, per risolvere l'equazione differenziale introduciamo una variabile ausiliaria $t = t(x)$, ponendo

$$y'(x) = \sinh t(x), \qquad \text{ossia} \qquad t(x) = \log\left(y'(x) + \sqrt{1+y'(x)^2}\right).$$

Si ha allora

$$y(x) = c\sqrt{1 + y'(x)^2} = c\sqrt{1 + \sinh^2 t(x)} = c\cosh t(x) \quad \forall x \in [a, b];$$

inoltre

$$\sinh t(x) = y'(x) = \frac{d}{dx}\, c\cosh t(x) = c\sinh t(x)\, t'(x),$$

da cui $t(x) = 0$ oppure $t'(x) = \frac{1}{c}$. Il primo caso corrisponde a $y(x) = c$, non ammissibile se $A \neq B$; nel secondo caso segue che $t(x)$ è invertibile e l'inversa $x(t)$ verifica $x'(t) = c$, per cui $x(t) = ct + d$. Quindi $t = \frac{x(t)-d}{c}$, ed in definitiva

$$y(x) = c\cosh t(x) = c\cosh\frac{x-d}{c}\,, \qquad x \in [a, b].$$

È facile verificare che queste funzioni risolvono effettivamente l'equazione differenziale.
Abbiamo così ottenuto come estremali delle *catenarie*; le corrispondenti superfici di rotazione si chiamano *catenoidi*. Occorre ora imporre le condizioni agli estremi:

$$y(a) = c\cosh\frac{a-d}{c} = A, \qquad y(b) = c\cosh\frac{b-d}{c} = B;$$

da queste si dovrebbero ricavare le costanti c, d. Si trovano due, una o nessuna soluzione a seconda di come sono scelti i valori A, B (vedere l'esercizio 4.4.2).
Anche in questo caso non abbiamo dimostrato che le estremali sono in effetti punti di minimo del funzionale: questo comunque è vero poiché il funzionale è convesso nella regione $\{(y, p) :\ y > 0,\ p \in \mathbb{R}\}$.

Esempio 4.4.3 Uno dei problemi che storicamente hanno motivato la nascita del calcolo delle variazioni è il problema della *brachistocrona*: dati due punti $(a, A), (b, B)$ in $\mathbb{R}^2$ con $A > B$, si cerca una curva cartesiana γ tale che un punto materiale vincolato a muoversi lungo γ vada dal primo al secondo estremo impiegando un tempo minimo.
Possiamo fissare l'origine delle coordinate in modo che $a = 0$ e $B = 0$, e supporre che il punto abbia massa unitaria. Per il principio di conservazione dell'energia, fissato $x \in [0, b]$ e detti $y(x)$ l'ordinata del punto e $v(x)$ il modulo della sua velocità, si avrà

$$\frac{1}{2}v(x)^2 + gy(x) = gA,$$

ove g è la costante di gravitazione; quindi

$$v(x) = \sqrt{2g(A - y(x))}.$$

Poiché il vettore velocità è tangente al grafico di γ in ogni punto, lo spazio s percorso lungo γ è pari a vt, dove t è il tempo impiegato. Dunque, il funzionale da minimizzare, cioè il tempo totale per andare dal primo estremo al secondo lungo γ è dato dall'integrale curvilineo

$$J(y) = T = \int_\gamma \frac{1}{v}\, ds = \int_0^b \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(A - y(x))}}\, dx.$$

La classe delle funzioni ammissibili è

$$X = \{y \in C[0,b] \cap C^1]0,b] : \ y(0) = A, \ y(b) = 0, \ 0 < y(x) < A \ \ \forall x \in]0,b[\,\}.$$

Si noti che a priori non possiamo prendere $y \in C^1[0,b]$, perché la soluzione potrebbe avere pendenza $-\infty$ nel punto iniziale. L'integrando è

$$F(y,p) = \frac{\sqrt{1 + p^2}}{\sqrt{2g(A - y)}}\ ,$$

quindi non dipende da x; si ha poi

$$F_y(y,p) = -\frac{\sqrt{1 + p^2}}{2\sqrt{2g}(A - y)^{3/2}}\ , \quad F_p(y,p) = \frac{p}{\sqrt{2g(A - y)(1 + p^2)}}\ .$$

Si noti che se $y \in X$ allora $y + v \in X$ per ogni $v \in C_0^1[0,b[$ sufficientemente piccola (in particolare, v deve annullarsi in un intorno sinistro di b). Dunque, possiamo dire che vale l'equazione di Eulero (dapprima soltanto in $]0, b - \varepsilon]$, con ε arbitrario, e poi di conseguenza in $]0,b]$); anzi, per la proposizione 4.4.1, esiste l'integrale primo

$$\frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(A - y(x))}} - \frac{y'(x)^2}{\sqrt{2g(A - y(x))(1 + y'(x)^2)}} = c \qquad \forall x \in]0,b],$$

che, dopo semplificazioni, diventa

$$\sqrt{2g(A - y(x))(1 + y'(x)^2)} = \frac{1}{c} \qquad \forall x \in]0,b].$$

Quadrando e ponendo $u(x) = A - y(x)$, si ottiene

$$u(x)(1 + u'(x)^2) = K \qquad \forall x \in]0, b].$$

Come nell'esempio 4.4.2, introduciamo la variabile ausiliaria $t(x)$ mediante la relazione

$$u'(x) = \frac{1}{\tan \frac{t(x)}{2}}, \qquad x \in]0, b]$$

(si noti che $t(x) \neq 0$ per ogni $x \in]0, b]$). Con procedura analoga, troviamo per ogni $x \in]0, b]$

$$u(x) = \frac{K}{1 + u'(x)^2} = \frac{K}{1 + \frac{1}{\tan^2 \frac{t(x)}{2}}} = \frac{K}{2}(1 - \cos t(x)),$$

da cui

$$\frac{1}{\tan \frac{t(x)}{2}} = u'(x) = \frac{K}{2} \sin t(x)\, t'(x)$$

e dunque

$$t'(x) = \frac{2}{K} \frac{1}{\sin t(x)\ \tan \frac{t(x)}{2}} = \frac{1}{K \sin^2 \frac{t(x)}{2}} = \frac{2}{K(1 - \cos t(x))} \qquad \forall x \in]0, b].$$

Perciò $t(x)$ è invertibile e l'inversa $x(t)$ verifica

$$x'(t) = \frac{K}{2}(1 - \cos t),$$

da cui $x(t) = \frac{K}{2}(t - \sin t)$; ne segue che la curva $(x, u(x))$ è data, in forma parametrica, da

$$x(t) = \frac{K}{2}(t - \sin t), \quad u(x(t)) = \frac{K}{2}(1 - \cos t).$$

Si tratta quindi di un arco di *cicloide*. Tornando alla funzione originaria y, essa è data da

$$x(t) = \frac{K}{2}(t - \sin t), \quad y(x(t)) = A - \frac{K}{2}(1 - \cos t).$$

Questa curva, che è ancora una cicloide, passa per il punto $(0, A)$ in corrispondenza del valore $t = 0$; passerà per l'altro estremo $(b, 0)$ in corrispondenza di un certo valore $t_0 > 0$ tale che

$$x(t_0) = b = \frac{K}{2}(t_0 - \sin t_0), \qquad y(x(t_0)) = 0 = A - \frac{K}{2}(1 - \cos t_0).$$

Il valore t_0 e la costante K si determinano risolvendo questo sistema. Si noti che la curva trovata è ammissibile, ossia è cartesiana ed appartiene a X, anche se non è possibile ricavare un'espressione esplicita per $y(x)$.
Si può verificare che il funzionale J è convesso nella regione $\{y \in X,\ p \in \mathbb{R}\}$, cosicché le estremali trovate sono in effetti punti di minimo del funzionale.

Esercizi 4.4

1. Si determini una curva $\gamma \in C[0,b] \cap C^1]0,b]$, con $\gamma(0) = A$, tale che un punto materiale vincolato a muoversi lungo γ, soggetto soltanto alla forza di gravità, raggiunga la retta $x = b$ in tempo minimo. Si provi che nel punto di contatto con tale retta la curva ottimale ha tangente orizzontale.
   [**Traccia:** con riferimento alle notazioni dell'esempio 4.4.3, si verifichi che imponendo la condizione $F_p(y(b), y'(b)) = 0$ si trova $t_0 = \pi$ e $K = 2b/\pi$.]

2. Si consideri il problema dell'esempio 4.4.2 con $[a,b] = [-1,1]$ e $A = B > 0$.

   (i) Si provi che se vi è un'estremale della forma
   $$y(x) = c \cosh\left(\frac{x}{c} + \mu\right), \qquad (\mu \in \mathbb{R},\ c > 0),$$
   allora deve essere $\mu = 0$ e $B = c \cosh \frac{1}{c}$.

   (ii) Se ne deduca che, posto $\beta = \min_{t>0} t \cosh \frac{1}{t}$, si hanno due estremali di tale forma se $B > \beta$, un'estremale di tale forma se $B = \beta$ e nessuna estremale di tale forma se $0 < B < \beta$.

4.5 Condizione di Legendre

Descriviamo brevemente una ulteriore condizione necessaria per l'esistenza di un'estremale, legata stavolta alla derivata seconda del funzionale J.

Proposizione 4.5.1 *Sia $J(y) = \int_a^b F(x, y(x), y'(x))\, dx$, con $F \in C^2([a,b] \times \mathbb{R}^2)$. Se y è punto di minimo locale per J nella classe $X = \{y \in C^1[a,b] : y(a) = A,\ y(b) = B\}$, allora*
$$F_{pp}(x, y(x), y'(x)) \geq 0 \qquad \forall x \in [a,b].$$

Se y è punto di massimo locale, allora

$$F_{pp}(x, y(x), y'(x)) \leq 0 \qquad \forall x \in [a, b].$$

Questa condizione è detta condizione necessaria di Legendre.

Dimostrazione Supponiamo che y sia punto di minimo locale. Se in un punto $x_0 \in]a, b[$ fosse $F_{pp}(x_0, y(x_0), y'(x_0)) < 0$, allora avremmo

$$F_{pp}(x, y(x), y'(x)) \leq -\beta < 0 \qquad \forall x \in [x_0 - \alpha, x_0 + \alpha],$$

con α e β numeri positivi opportuni. Consideriamo la funzione

$$u(x) = \begin{cases} \sin^2 \frac{\pi(x-x_0)}{\alpha} & \text{se } |x - x_0| \leq \alpha, \\ 0 & \text{altrimenti.} \end{cases}$$

Se α è sufficientemente piccolo, si ha $u \in C_0^1[a, b]$ e con calcoli elementari si vede che

$$\|u\|^2_{L^2(a,b)} = \int_{x_0-\alpha}^{x_0+\alpha} \sin^4 \frac{\pi(x - x_0)}{\alpha}\, dx = \frac{\alpha}{\pi} \int_{-\pi}^{\pi} \sin^4 t\, dt = \frac{3}{4}\,\alpha,$$

$$\|u'\|^2_{L^2(a,b)} = \int_{x_0-\alpha}^{x_0+\alpha} \frac{\pi^2}{\alpha^2} \sin^2 \frac{2\pi(x - x_0)}{\alpha}\, dx = \frac{\pi}{2\alpha} \int_{-2\pi}^{2\pi} \sin^2 t\, dt = \frac{\pi^2}{\alpha}\,.$$

Adesso calcoliamo la derivata seconda del funzionale J: per l'osservazione 4.1.2(2) si ha

$$\begin{aligned} [J''(y)v]w &= \int_a^b \Big[F_{yy}(x, y(x), y'(x))v(x)w(x) + \\ &+ F_{yp}(x, y(x), y'(x))[v(x)w'(x) + w(x)v'(x)] + \\ &+ F_{pp}(x, y(x), y'(x))v'(x)w'(x)\Big]dx. \end{aligned}$$

Quindi, sostituendo u al posto di v e w ed utilizzando le stime precedenti, si ottiene

$$\begin{aligned} [J''(y)u]u &= \int_{x_0-\alpha}^{x_0+\alpha} \Big[F_{yy}(x, y(x), y'(x))u(x)^2 + \\ &+ 2F_{yp}(x, y(x), y'(x))u(x)u'(x) + F_{pp}(x, y(x), y'(x))u'(x)^2\Big]dx \leq \\ &\leq \frac{3\alpha}{4}\|F_{yy}(\cdot, y(\cdot), y'(\cdot))\|_{L^\infty(a,b)} + \\ &+ \pi\sqrt{3}\|F_{py}(\cdot, y(\cdot), y'(\cdot))\|_{L^\infty(a,b)} - \frac{\beta\pi^2}{\alpha} = c_1\alpha + c_2 - \frac{c_3}{\alpha}\,, \end{aligned}$$

ove c_1, c_2 e c_3 sono costanti positive. Se α è sufficientemente piccolo, si ottiene allora

$$[J''(y)u]u < 0,$$

il che è assurdo, in quanto per la proposizione 3.6.2 dovrebbe essere

$$[J''(y)v]v \geq 0 \qquad \forall v \in C^1[a,b]. \quad \square$$

Osservazione 4.5.2 Nel caso di un funzionale definito su $C^1([a,b],\mathbb{R}^m)$, se $F \in C^2([a,b] \times \mathbb{R}^{2m})$ la condizione necessaria di Legendre assume la forma seguente: per ogni $x \in [a,b]$, la matrice $\{F_{p_i p_j}(x, y(x), y'(x))\}_{i,j=1,\ldots,m}$ è semi-definita positiva quando y è punto di minimo locale, ed è semidefinita negativa quando y è punto di massimo locale.

Esercizi 4.5

1. Fissata $c \in \mathbb{R}$, si consideri il funzionale

$$J(u) = \int_0^1 [u'(x)^2 + cu(x)^2]\, dx, \qquad u \in C_0^1[0,1].$$

   (i) Si trovino le estremali di J.

   (ii) Si verifichi che J'' è costante e verifica la condizione di Legendre in ogni punto di $C_0^1[0,1]$.

   (iii) Si provi che 0 è punto di minimo assoluto per J se e solo se risulta $c \geq -\pi^2$.

   (iv) Si mostri che se $c < -\pi^2$ si ha $\inf J = -\infty$.

4.6 Il problema isoperimetrico

Dedichiamo questo paragrafo all'analisi di un problema classico, vale a dire il *problema isoperimetrico*, che è il seguente: fra tutte le curve chiuse di classe C^1 in $\mathbb{R}^2$ aventi lunghezza assegnata L, si vuole determinare quella che delimita la regione di area massima.
Ricordando che per ogni dominio $D \subset \mathbb{R}^2$ con frontiera di classe C^1 vale la formula

$$m_2(D) = \frac{1}{2}\int_{+\partial D} (x\, dy - y\, dx),$$

ove l'orientazione positiva su ∂D è quella antioraria, occorrerà massimizzare la funzione

$$F(\gamma) = \frac{1}{2}\int_{+\gamma} (x\,dy - y\,dx)$$

fra tutte le curve chiuse γ in $\mathbb{R}^2$, di classe C^1, aventi lunghezza L. Fissiamo un intervallo $[a,b]$ e scegliamo $X = C^1([a,b],\mathbb{R}^2)$; cercheremo il massimo del funzionale $F(\gamma) = \frac{1}{2}\int_a^b (xy' - yx')\,dt$ sull'insieme

$$Z = \left\{\gamma \in X:\ \gamma(a) = \gamma(b),\ \int_a^b |\gamma'|_2\,dt = L\right\}.$$

Si tratta di un problema di massimo con vincolo: quindi andiamo ad applicare il teorema 3.9.1. Poniamo $Y = \mathbb{R}$ e $\Phi(\gamma) = \int_a^b |\gamma'|_2\ dt - L$. Denotiamo con $\gamma(\cdot) = (x(\cdot), y(\cdot))$ una generica curva appartenente a X; se $h(\cdot) = (u(\cdot), v(\cdot))$ è un altro elemento di X si ha

$$\begin{aligned} F'(\gamma)h &= [F'(x(\cdot), y(\cdot))](u(\cdot), v(\cdot)) = \\ &= \frac{1}{2}\int_a^b [(y'u - yu') + (xv' - x'v)]dt = \\ &= \frac{1}{2}\int_a^b \left[\left\langle \begin{pmatrix} y' \\ -x' \end{pmatrix}, h \right\rangle_2 + \left\langle \begin{pmatrix} -y \\ x \end{pmatrix}, h' \right\rangle_2\right] dt \quad \forall h \in X, \end{aligned}$$

e anche

$$\Phi'(\gamma)h = \int_a^b \frac{\langle \gamma', h' \rangle_2}{|\gamma'|_2}\ dt \quad \forall h \in X.$$

Si noti che $\Phi'(\gamma)$ non esiste in generale se la curva γ non è una curva regolare: quindi dovremo escludere eventuali soluzioni non regolari. Osserviamo che, al contrario, se γ è una curva regolare allora $\Phi'(\gamma) : X \to \mathbb{R}$ è un funzionale surgettivo: se si sceglie $h = r\gamma$, con $\gamma \in X$, al variare di $r \in \mathbb{R}$ la quantità $\Phi'(\gamma)r\gamma = r\int_a^b |\gamma'|_2\,dt = r\ell(\gamma)$ descrive tutto $\mathbb{R}$. Dunque, per il teorema 3.9.1,se γ è punto di massimo di F su Z, esiste $\lambda \in \mathbb{R}^* = \mathbb{R}$ tale che

$$F'(\gamma)h + [\Phi'(\gamma)^*\lambda]h = 0 \qquad \forall h \in X,$$

ossia

$$\frac{1}{2}\int_a^b \left[(y'u - yu') + (-x'v + xv') + \lambda\frac{x'u' + y'v'}{\sqrt{x'^2 + y'^2}}\right] dt = 0 \quad \forall (u(\cdot), v(\cdot)) \in X.$$

Dato che u e v variano indipendentemente nella classe delle funzioni $w \in C^1[a,b]$ con $w(a) = w(b)$, scegliendo prima l'una e poi l'altra uguale a 0 l'equazione si scinde nel sistema

$$\begin{cases} \int_a^b \Big(y'u - yu' + \lambda \frac{x'u'}{\sqrt{x'^2+y'^2}} \Big) dt = 0 \\ \int_a^b \Big(- x'v + xv' + \lambda \frac{y'v'}{\sqrt{x'^2+y'^2}} \Big) dt = 0 \end{cases} \qquad \forall (u(\cdot), v(\cdot)) \in X.$$

Dal lemma 4.2.3 segue

$$\begin{cases} \frac{d}{dt} \Big[- y + \lambda \frac{x'}{\sqrt{x'^2+y'^2}} \Big] - y' = 0 \\ \frac{d}{dt} \Big[x + \lambda \frac{y'}{\sqrt{x'^2+y'^2}} \Big] + x' = 0 \end{cases} \qquad \text{in } [a,b],$$

da cui, integrando,

$$\lambda \frac{x'}{\sqrt{x'^2 + y'^2}} = 2(c_1 + y), \quad \lambda \frac{y'}{\sqrt{x'^2 + y'^2}} = 2(c_2 - x) \qquad \text{in } [a,b].$$

Se ne deduce, quadrando e sommando,

$$\frac{\lambda^2}{4} = (c_1 + y(t))^2 + (c_2 - x(t))^2 \qquad \forall t \in [a,b].$$

Ciò prova che i punti di massimo di F in Z sono circonferenze di raggio $\frac{|\lambda|}{2}$ e, naturalmente, centro $(c_2, -c_1)$ arbitrario. Dato che la circonferenza deve avere lunghezza L, si ricava $L = \pi|\lambda|$; dunque le curve di lunghezza L che delimitano una regione di area massima, *se esistono*, sono le circonferenze di raggio $\frac{L}{2\pi}$. I metodi qui illustrati non ci permettono di dedurre che le circonferenze effettivamente risolvono il nostro problema; di questo fatto peraltro esiste una dimostrazione di tipo geometrico, non banale ma del tutto elementare.

Esercizi 4.6

1. Dati due punti distinti $A, B \in \mathbb{R}^2$, sia γ_0 una fissata curva regolare di estremi A e B, il cui sostegno giaccia in uno dei due semipiani individuati dalla retta passante per A e per B. Fra tutte le curve di fissata lunghezza ℓ, aventi per estremi A e B ed il cui sostegno non interseca quello di γ_0, si determini quella che insieme a γ_0 delimita una regione di area massima.

4.7 Geodetiche su una superficie

In questo paragrafo ci occupiamo della determinazione delle curve *geodetiche*, ossia di minima lunghezza, giacenti su una superficie regolare assegnata. Si tratta di un problema classico e piuttosto intricato, di fondamentale importanza in geometria differenziale; cercheremo di affrontarlo con i metodi dei paragrafi precedenti.

Sia dunque $\Sigma \subseteq \mathbb{R}^3$ una superficie regolare connessa, di equazione $g(x, y, z) = 0$, e supponiamo che g sia di classe C^3 (questo fatto ci servirà in seguito); dati due punti distinti $A, B \in \Sigma$, si cerca la curva (regolare) γ_0 giacente su Σ, di estremi A e B, che abbia lunghezza minima. Non è detto a priori che tale curva esista, né che sia unica.

Fissiamo $[a, b] \subseteq \mathbb{R}$ e consideriamo curve $\gamma : [a, b] \to \Sigma$ di classe C^1. Dobbiamo minimizzare il funzionale

$$J(\gamma) = \int_a^b |\gamma'|_3 \, dt$$

sull'insieme $K = \{\gamma \in C^1([a, b], \mathbb{R}^3) : \ \gamma(a) = A, \ \gamma(b) = B\}$, sotto il vincolo che γ sia contenuta in Σ, ossia sotto la condizione $g(\gamma(t)) = 0$ per ogni $t \in [a, b]$. Abbiamo a disposizione il teorema 3.9.1, che fornisce una condizione necessaria affinché una fissata curva γ_0 sia una geodetica di Σ; ma occorre che l'insieme dove è definita la funzione da minimizzare sia uno spazio di Banach, mentre nel nostro caso K è soltanto una varietà affine di $C^1([a, b], \mathbb{R}^3)$. Conviene allora porre $\gamma = \gamma_0 + \psi$, in modo che la curva ψ appartenga allo spazio di Banach $X = C_0^1([a, b], \mathbb{R}^3)$. Poniamo poi $Y = C^1[a, b]$, e definiamo le funzioni $F : X \to \mathbb{R}$ e $\Phi : X \to Y$ nel modo seguente:

$$F(\psi) = \int_a^b |\gamma_0' + \psi'|_3 \, dt, \qquad \Phi(\psi) = g \circ (\gamma_0 + \psi),$$

cosicché il fatto che γ_0 sia una geodetica di Σ di estremi A, B significa che 0 è punto di minimo per la restrizione di F al vincolo $Z = \{\psi \in X : \ \Phi(\psi) = 0\}$. Verifichiamo le ipotesi del teorema 3.9.1. Anzitutto, avendo supposto γ_0 regolare, sia F che Φ sono F-differenziabili in un intorno di $\psi = 0$, con

$$F'(\psi)\varphi = \int_a^b \frac{\langle \gamma_0' + \psi', \varphi' \rangle_3}{|\gamma_0' + \psi'|_3} \, dt, \quad \Phi'(\psi)\varphi = \langle \nabla g(\gamma_0 + \psi), \varphi \rangle_3 \qquad \forall \varphi \in X.$$

Verifichiamo che $\Phi'(0) : X \to Y$ è un operatore surgettivo: poiché la superficie Σ è regolare, si ha $\nabla g(\gamma_0(t)) \neq 0$ per ogni $t \in [a, b]$. Essendo $[a, b]$

compatto, esiste $\delta > 0$ tale che

$$|\nabla g(\gamma_0)(t)|_3 \geq \delta \qquad \forall t \in [a,b].$$

Quindi, posto

$$A_i = \left\{ t \in [a,b] : \; |D_i g(\gamma_0(t))| > \frac{\delta}{2\sqrt{3}} \right\}, \quad i = 1,2,3,$$

la terna $\{A_1, A_2, A_3\}$ è un ricoprimento aperto di $[a,b]$. Sia $\{\varphi_1, \varphi_2, \varphi_3\}$ una partizione dell'unità di classe C^∞ associata a tale ricoprimento (ossia $\varphi_i \in C_0^\infty(A_i)$, $\varphi_i \geq 0$, e $\sum_{i=1}^3 \varphi_i \equiv 1$). Allora, se v è un arbitrario elemento di Y, scelta $h = (h_1, h_2, h_3)$ ove $h_i = \frac{v\varphi_i}{D_i g(\gamma_0)}$, si ha $h \in X$ e

$$\Phi'(0)h = \langle \nabla g(\gamma_0), h \rangle_3 = \sum_{i=1}^{3} D_i g(\gamma_0) h_i = \sum_{i=1}^{3} \varphi_i v = v,$$

il che mostra la surgettività di $\Phi'(0)$.
Per il teorema 3.9.1, esiste $T \in Y^* = (C^1[a,b])^*$ tale che

$$F'(0)h + T(\Phi'(0)h) = 0 \qquad \forall h \in X.$$

Per interpretare questa uguaglianza occorre caratterizzare lo spazio duale $(C^1[a,b])^*$.

Lemma 4.7.1 *Lo spazio $(C^1[a,b])^*$ è costituito dai funzionali T della forma*

$$Tf = \alpha f(a) + \mu f' \qquad f \in C^1[a,b],$$

dove $\alpha \in \mathbb{R}$ e $\mu \in (C[a,b])^*$, ossia $\mu = \mu^+ - \mu^-$, con μ^+ e μ^- misure di Lebesgue-Stieltjes associate a funzioni $p^+, p^- : [a,b] \to \mathbb{R}$ crescenti e continue a sinistra.

Dimostrazione È chiaro che ogni funzionale T della forma sopra scritta appartiene a $(C^1[a,b])^*$.
Viceversa, sia $T \in (C^1[a,b])^*$. Consideriamo il funzionale $T|_M$, ove $M = \{f \in C^1[a,b] : \; f(a) = 0\}$. Osserviamo che l'operatore $A = \frac{d}{dt} : M \to C[a,b]$ è bigettivo e continuo, perché se $\eta \in C[a,b]$ vi è una ed una sola funzione $f \in M$ tale che $f' = \eta$, e precisamente $f(t) = \int_0^t \eta(s)ds$. Per il lemma 3.9.2, l'operatore aggiunto $A^* : (C[a,b])^* \to M^*$ è a sua volta surgettivo:

quindi esiste $\mu \in (C[a,b])^*$ tale che $A^*\mu = \psi|_M$. Il funzionale μ è della forma $\mu = \mu^+ - \mu^-$ con μ^+ e μ^- misure di Lebesgue-Stieltjes su $[a,b]$ associate a funzioni p^+, p^- crescenti e continue a sinistra in $[a,b]$.
Se ora $f \in C^1[a,b]$, possiamo scrivere $f = f(a) + g$, con $g \in M$; ne segue, indicando con $I(t)$ la funzione costantemente uguale a 1,

$$Tf = Tf(a) + (A^*\mu)g = f(a)TI + \mu g'.$$

Scelto $\alpha = TI \in \mathbb{R}$, si ha la tesi del lemma. $\square$

Possiamo ora interpretare l'identità fornita dal teorema 3.9.1. Supporremo però che la curva geodetica γ_0 sia di classe C^2, come del resto è pratica comune in geometria differenziale, dove non è mai utile fare economia nella regolarità. L'equazione $F'(0)h + T(\Phi'(0)h) = 0$, valida per ogni $h \in X$, diventa, tenuto conto dell'espressione di T dedotta dal lemma 4.7.1,

$$\int_a^b \frac{\langle \gamma_0', h' \rangle_3}{|\gamma_0'|_3}\, dt + \alpha \langle \nabla g(\gamma_0(a)), h(a) \rangle_3 + \int_a^b \frac{d}{dt} \langle \nabla g(\gamma_0(t)), h(t) \rangle_3 \, d\mu = 0$$

per ogni $h \in X$, ove il simbolo $d\mu$ è una forma compatta per indicare $d\mu^+ - d\mu^-$. E poiché $h(a) = 0$, si ottiene più semplicemente

$$\int_a^b \frac{\langle \gamma_0', h' \rangle_3}{|\gamma_0'|_3}\, dt + \int_a^b \frac{d}{dt} \langle \nabla g(\gamma_0(t)), h(t) \rangle_3 \, d\mu = 0 \quad \forall h \in X.$$

Adesso vorremmo integrare per parti il secondo termine del primo membro, sfruttando il fatto che g è di classe C^3 e γ_0 di classe C^2. A questo scopo è fondamentale la seguente generalizzazione della formula di integrazione per parti:

Lemma 4.7.2 *Sia $p : [a,b] \to \mathbb{R}$ una funzione crescente e continua a sinistra, e sia μ la corrispondente misura di Lebesgue-Stieltjes. Allora per ogni $f \in AC[a,b]$ si ha*

$$\int_a^b f \, d\mu = f(b)p(b) - f(a)p(a) - \int_a^b f' p \, dt.$$

Dimostrazione La formula è vera quando $p \in AC[a,b]$: basta infatti osservare che in tal caso $fp \in AC[a,b]$, scrivere la formula che lega fp all'integrale della sua derivata ed osservare che $d\mu = p'(t)dt$.
Nel caso generale, esiste una successione $\{p_n\} \subset AC[a,b]$ tale che ogni p_n

è crescente in $[a, b]$ e $p_n(t) \to p(t)$ puntualmente in $[a, b]$ (basta ad esempio prendere le convoluzioni $p \star \varphi_n$, ove $\varphi_n(t) = n\varphi(nt)$, con $\varphi \in C^\infty(\mathbb{R})$, $\varphi \geq 0$, $\varphi(t) = 0$ per $|t| \geq 1$ e $\int_{-1}^{1} \varphi(t)\, dt = 1$). Allora si ha

$$\mu([x, y[\,) = p(y) - p(x) = \lim_{n\to\infty} [p_n(y) - p_n(x)] = \lim_{n\to\infty} \int_x^y p_n'\, dt.$$

Di conseguenza, per ogni funzione η costante a tratti, $\eta = \sum_{k=1}^{N} \alpha_k \chi_{[c_k, d_k[}$, si ha

$$\int_a^b \eta\, d\mu = \sum_{k=1}^{N} \alpha_k\, \mu([c_k, d_k[\,) = \lim_{n\to\infty} \sum_{k=1}^{N} \alpha_k \int_{c_k}^{d_k} p_n'\, dt = \lim_{n\to\infty} \int_a^b \eta p_n'\, dt.$$

Vogliamo ottenere la stessa relazione quando η è una funzione crescente ed assolutamente continua. Scegliamo una successione $\{\eta_k\}$ di funzioni costanti a tratti tale che $\eta_k \leq \eta$ e $\eta_k(t) \to \eta(t)$ uniformemente in $[a, b]$; si ha allora

$$\int_a^b \eta_k\, d\mu = \lim_{n\to\infty} \int_a^b \eta_k p_n'\, dt \leq \lim_{n\to\infty} \int_a^b \eta p_n'\, dt \quad \forall k \in \mathbb{N},$$

da cui

$$\int_a^b \eta\, d\mu = \lim_{k\to\infty} \int_a^b \eta_k\, d\mu \leq \lim_{n\to\infty} \int_a^b \eta p_n'\, dt.$$

D'altra parte, fissato $\varepsilon > 0$ e scelto k sufficientemente grande,

$$\begin{aligned}
\lim_{n\to\infty} \int_a^b \eta p_n'\, dt &\leq \limsup_{n\to\infty} \int_a^b (\eta - \eta_k) p_n'\, dt + \lim_{n\to\infty} \int_a^b \eta_k p_n'\, dt \leq \\
&\leq \lim_{n\to\infty} \varepsilon[p_n(b) - p_n(a)] + \int_a^b \eta_k\, d\mu \leq \\
&\leq \varepsilon[p(b) - p(a)] + \int_a^b \eta\, d\mu.
\end{aligned}$$

Per l'arbitrarietà di ε, si deduce che

$$\int_a^b \eta\, d\mu = \lim_{n\to\infty} \int_a^b \eta p_n'\, dt$$

per ogni funzione η assolutamente continua e crescente; per differenza, questa relazione vale per ogni $\eta \in AC[a, b]$.

Sia finalmente $f \in AC[a,b]$: possiamo scrivere, utilizzando il teorema di Lebesgue,

$$\begin{aligned}
\int_a^b f\,d\mu &= \lim_{n\to\infty}\int_a^b f p_n'\,dt = \\
&= \lim_{n\to\infty}\left[f(b)p_n(b) - f(a)p_n(a) - \int_a^b f' p_n\,dt\right] = \\
&= f(b)p(b) - f(a)p(a) - \int_a^b f' p\,dt,
\end{aligned}$$

e ciò prova la tesi. $\square$

Consideriamo ora le funzioni $h \in C_0^2(\,]a,b[\,,\mathbb{R}^3)$, che costituiscono un sottospazio denso in X: allora per tali h, tenendo conto del lemma 4.7.2 (applicato separatamente alle misure μ^+ e μ^-), la condizione necessaria per il minimo vincolato si può scrivere nella forma seguente:

$$\int_a^b \frac{\langle \gamma_0', h'\rangle_3}{|\gamma_0'|_3}\,dt - \int_a^b \frac{d^2}{dt^2}\langle \nabla g(\gamma_0(t)), h(t)\rangle_3\, p(t)\,dt,$$

da cui, sviluppando le derivate,

$$\int_a^b \left[-p\langle[\nabla g(\gamma_0)]'', h\rangle_3 + \left\langle \frac{\gamma_0'}{|\gamma_0'|_3} - 2p[\nabla g(\gamma_0)]', h'\right\rangle_3 - p\langle \nabla g(\gamma_0), h''\rangle_3\right] dt = 0$$

per ogni $h \in C_0^2(\,]a,b[\,,\mathbb{R}^3)$. Utilizzando il lemma 4.2.5, otteniamo le seguenti informazioni: anzitutto, $-p\nabla g(\gamma_0) \in C^1([a,b],\mathbb{R}^3)$, il che implica $p \in C^1[a,b]$ essendo $\nabla g(\gamma_0)$ di classe C^2 e diverso da 0; in secondo luogo, la funzione

$$\frac{\gamma_0'}{|\gamma_0'|_3} - 2p[\nabla g(\gamma_0)]' + (p[\nabla g(\gamma_0)])' = \frac{\gamma_0'}{|\gamma_0'|_3} - p[\nabla g(\gamma_0)]' + p'[\nabla g(\gamma_0)]$$

deve essere in $C^1([a,b],\mathbb{R}^3)$: ciò implica, per differenza, che $p'[\nabla g(\gamma_0)]$ è di classe C^1 e dunque $p \in C^2[a,b]$. La terza informazione è la relazione

$$\frac{d}{dt}\left[\frac{\gamma_0'}{|\gamma_0'|_3} - p[\nabla g(\gamma_0)]' + p'[\nabla g(\gamma_0)]\right] = -p[\nabla g(\gamma_0)]'' \quad \text{in } [a,b],$$

dalla quale, svolgendo le derivate (lecito, visto che sappiamo che $p \in C^2[a,b]$), ricaviamo dopo alcune semplificazioni

$$\frac{d}{dt}\frac{\gamma_0'}{|\gamma_0'|_3} + p''\nabla g(\gamma_0) = 0 \qquad \text{in } [a,b].$$

Questo è il sistema differenziale che caratterizza le geodetiche: esso ci dice che, lungo la geodetica γ_0, la normale principale alla curva (ossia la derivata del versore tangente $\frac{\gamma_0'}{|\gamma_0'|_3}$) è parallela alla normale alla superficie Σ, che è il vettore ∇g. Se, in particolare, rappresentiamo la curva γ_0 mediante il parametro lunghezza d'arco s, e poniamo $\lambda = -p''$, il sistema acquista la forma più semplice

$$\gamma_0''(s) - \lambda(s)\nabla g(\gamma_0(s)) = 0.$$

Questo sistema, che pure costituisce soltanto una condizione necessaria per l'esistenza di geodetiche, permette in molti casi concreti di calcolarle esplicitamente (esempi 4.7.3 e 4.7.4, nonché esercizio 4.7.1).
Naturalmente, resta il problema di provare l'esistenza di almeno una geodetica. Dimostreremo che esiste almeno una geodetica nella classe delle curve lipschitziane su $[a, b]$ (ossia assolutamente continue in $[a, b]$ con derivata essenzialmente limitata).
Osserviamo prima di tutto che, essendo Σ una superficie regolare e connessa, vi è almeno una curva regolare $\varphi : [a, b] \to \Sigma$ di estremi A e B; posto $M = \|\varphi'\|_\infty$, l'insieme

$$N = \{\gamma \in AC([a, b], \Sigma) : \gamma(a) = A, \gamma(b) = B, \gamma' \in L^\infty(a, b; \mathbb{R}^3), \|\gamma'\|_\infty \le M\}$$

è dunque un sottoinsieme non vuoto dello spazio

$$Z = \{\gamma \in AC([a, b], \mathbb{R}^3) : \gamma' \in L^\infty(a, b; \mathbb{R}^3)\}$$

che è di Banach con la norma

$$\|\gamma\|_Z = \|\gamma\|_\infty + \|\gamma'\|_\infty .$$

Proviamo che N è compatto in Z: per ogni $\gamma \in N$ si ha

$$|\gamma(t)|_3 \le |\gamma(a)|_3 + \int_a^t |\gamma'(\tau)|_3 \, d\tau \le |A|_3 + M(b - a) < \infty \qquad \forall t \in [a, b],$$

$$|\gamma(t) - \gamma(s)|_3 = \left| \int_s^t \gamma'(r) \, dr \right|_3 \le M|t - s| \qquad \forall s, t \in [a, b],$$

e dunque le funzioni di N costituiscono una famiglia equicontinua ed equilimitata. Per il teorema di Ascoli-Arzelà, N è relativamente compatto rispetto alla convergenza uniforme; la compattezza in Z si ottiene osservando che se $\{\gamma_n\}_{n\in\mathbb{N}} \subset N$ e $\gamma_n \to \gamma$ uniformemente, allora otteniamo anche

$|\gamma(t) - \gamma(s)|_3 \le M|t-s|$ e quindi γ è assolutamente continua: inoltre la sua derivata, che esiste quasi ovunque, è limitata da M, il che mostra che $\gamma \in N$. Osserviamo adesso che il funzionale lunghezza $F(\gamma) = \int_a^b |\gamma'(t)|_3 \, dt$ è ovviamente continuo su Z (e addirittura semicontinuo inferiormente rispetto alla convergenza uniforme, in virtù dell'esercizio 4.1.2); quindi F, che evidentemente è limitato inferiormente, ha minimo sul compatto N, ossia esiste una geodetica in N.
Resta aperto il delicato problema della regolarità: sappiamo che esiste una geodetica lipschitziana, ma la condizione necessaria da noi dimostrata in precedenza vale solo per geodetiche di classe C^2. Questa questione è però al di là della portata degli strumenti che abbiamo.

Esempio 4.7.3 Consideriamo il caso della superficie sferica di raggio r, centrata nell'origine. Si ha $g(x,y,z) = x^2+y^2+z^2-r^2$, e dunque il sistema che caratterizza le geodetiche assume la forma

$$\gamma''(s) = 2\lambda(s)\gamma(s),$$

dove abbiamo scritto per comodità γ in luogo di γ_0. Moltiplicando scalarmente per $\gamma(s)$ otteniamo

$$\frac{d}{ds}\langle \gamma(s), \gamma'(s)\rangle_3 - |\gamma'(s)|_3^2 = \langle \gamma(s), \gamma''(s)\rangle_3 = 2\lambda(s)|\gamma(s)|_3^2;$$

d'altra parte si ha $|\gamma(s)|_3 = r$, $|\gamma'(s)|_3 = 1$ ed inoltre $\langle \gamma(s), \gamma'(s)\rangle_3 = \frac{1}{2}\frac{d}{ds}|\gamma(s)|_3^2 = 0$. Quindi l'equazione si riduce a $-1 = 2\lambda(s)r^2$, il che significa $\lambda(s) \equiv -\frac{1}{2r^2}$; in particolare, λ è indipendente da s. Arriviamo così al sistema

$$\gamma''(s) = 2\lambda\gamma(s) = -\frac{1}{r^2}\,\gamma(s).$$

Scegliamo i punti A, B: non è restrittivo per uno dei due scegliere il più comodo possibile dal punto di vista dei calcoli, mentre per il secondo la scelta deve essere generica. Poniamo quindi

$$A = (r,0,0), \qquad B = (x_0, y_0, z_0).$$

Risolvendo il sistema si trova

$$\begin{cases} x(s) = P\cos\frac{s}{r} + Q\sin\frac{s}{r} \\ y(s) = R\cos\frac{s}{r} + S\sin\frac{s}{r} \\ z(s) = T\cos\frac{s}{r} + U\sin\frac{s}{r}\ , \end{cases}$$

con P, Q, R, S, T, U costanti da determinare. Imponendo che $\gamma(0) = A$ si trova subito $P = r$, $R = 0$ e $T = 0$; dall'identità $|\gamma'(s)|_3 = 1$, ponendo dapprima $s = 0$, si trova con facili calcoli $Q^2+S^2+U^2 = r^2$, e di conseguenza, con $s \neq 0$, si deduce $Q = 0$ e $S^2 + U^2 = r^2$. In definitiva

$$\begin{cases} x(s) = r \cos \frac{s}{r} \\ y(s) = S \sin \frac{s}{r} \\ z(s) = U \sin \frac{s}{r} \end{cases},$$

da cui ricaviamo

$$\frac{y(s)}{S} = \frac{z(s)}{U},$$

ossia la curva γ giace sul piano di equazione $Uy - Sz = 0$. Questo piano passa per l'origine (contiene anzi l'asse x), e dunque taglia la superficie sferica Σ in un cerchio massimo. Ne segue che la geodetica γ è necessariamente un arco di cerchio massimo.

Esempio 4.7.4 Consideriamo il caso di un cilindro, con asse che possiamo supporre coincidente con l'asse z. Avremo perciò $g(x, y, z) = x^2 + y^2 - r^2$. Il sistema delle geodetiche diventa

$$\begin{cases} x''(s) = 2\lambda(s)x(s) \\ y''(s) = 2\lambda(s)y(s) \\ z''(s) = 0. \end{cases}$$

Scegliamo i punti $A = (r, 0, 0)$ e $B = (x_0, y_0, z_0)$. Moltiplicando la prima equazione per $x(s)$, la seconda per $y(s)$, e sommando, otteniamo

$$x(s)x''(s) + y(s)y''(s) = 2\lambda(s)(x(s)^2 + y(s)^2),$$

ed osservato che $xx'' + yy'' = (xx' + yy')' - (x'^2 + y'^2) = -(x'^2 + y'^2)$ (perché dall'equazione $x^2 + y^2 = r^2$ segue $xx' + yy' = 0$), deduciamo $2\lambda(s) = -r^{-2}(x'(s)^2+y'(s)^2)$. D'altronde, risolvendo la terza equazione si ha $z'(s) = c$ e $z(s) = cs + d$, con $d = 0$ (per come è stato scelto A); quindi

$$\lambda(s) = -\frac{x'(s)^2 + y'(s)^2}{2r^2} = -\frac{1 - z'(s)^2}{2r^2} = -\frac{1 - c^2}{2r^2},$$

quindi nuovamente λ non dipende da s. Si noti che, necessariamente, $|c| \leq 1$. Le prime due equazioni del sistema diventano allora

$$x''(s) = -\frac{1 - c^2}{r^2}\, x(s), \qquad y''(s) = -\frac{1 - c^2}{r^2}\, y(s),$$

da cui

$$x(s) = P\cos\frac{\sqrt{1-c^2}}{r}\,s + Q\sin\frac{\sqrt{1-c^2}}{r}\,s,$$
$$y(s) = R\cos\frac{\sqrt{1-c^2}}{r}\,s + S\sin\frac{\sqrt{1-c^2}}{r}\,s.$$

Dalla condizione iniziale ricaviamo $P = r$ e $R = 0$, mentre dall'identità $x(s)x'(s) + y(s)y'(s) = 0$, scelto $s = 0$, segue che $Q = 0$; di conseguenza $S^2 - r^2 = 0$, ossia $S = \pm r$. In definitiva

$$\begin{cases} x(s) = r\ \cos\frac{\sqrt{1-c^2}}{r}\ s, \\ y(s) = \pm r\ \sin\frac{\sqrt{1-c^2}}{r}\ s, \\ z(s) = cs, \end{cases}$$

ove la costante $c \in [-1, 1]$ dipende dalla scelta del punto B. Queste sono le equazioni di un'elica cilindrica; nel caso particolare in cui $c = 0$ si ha un arco di cerchio orizzontale, mentre quando $c = \pm 1$ si ottiene un segmento verticale.

Esercizi 4.7

1. *(Geodetiche su un cono)* Dato il cono di equazione $z = a\sqrt{x^2+y^2}$, ove $a > 0$, sia $\gamma(s) = (x(s), y(s), z(s))$ una geodetica di estremi $A = (x_0, y_0, a\sqrt{x_0^2+y_0^2})$ e $B = (x_1, y_1, a\sqrt{x_1^2+y_1^2})$, espressa in funzione del parametro lunghezza d'arco.

   (i) Si scrivano le equazioni differenziali soddisfatte da $x(s)$, $y(s)$, $z(s)$.

   (ii) Si provi che
   $$\frac{d^2}{ds^2}|\gamma(s)|_3^2 = 2, \qquad \frac{d^2}{ds^2}\,z(s)^2 = \frac{2a^2}{1+a^2}\,.$$

   (iii) Si deduca che
   $$z(s)^2 = \frac{a^2}{1+a^2}\,s^2 + bs + c,$$
   con b, c costanti univocamente individuate dalle condizioni
   $$z(0) = a\sqrt{x_0^2+y_0^2}, \qquad z(\ell) = a\sqrt{x_1^2+y_1^2}.$$

(iv) Scritte $x(s)$ e $y(s)$ nella forma

$$x(s) = \frac{1}{a}\, z(s) \cos\vartheta(s), \qquad y(s) = \frac{1}{a}\, z(s) \sin\vartheta(s),$$

si verifichi che

$$\begin{cases} a[\cos\vartheta(s)\, x'(s) + \sin\vartheta(s)\, y'(s)] = z'(s) \\ a[\sin\vartheta(s)\, x'(s) - \cos\vartheta(s)\, y'(s)] = -z(s)\vartheta'(s), \end{cases}$$

e se ne deduca che

$$\vartheta''(s) = -2\frac{z'(s)\vartheta'(s)}{z(s)}\,.$$

(v) Si concluda che

$$\vartheta(s) = \vartheta_0 + (\vartheta_1 - \vartheta_0)\frac{\int_0^s \frac{dr}{z(r)^2}}{\int_0^\ell \frac{dr}{z(r)^2}}\,,$$

dove ϑ_0 e ϑ_1 sono individuati dalle relazioni

$$\cos\vartheta_0 = \frac{x_0}{\sqrt{x_0^2 + y_0^2}}\,, \qquad \sin\vartheta_0 = \frac{y_0}{\sqrt{x_0^2 + y_0^2}}\,,$$

$$\cos\vartheta_1 = \frac{x_1}{\sqrt{x_1^2 + y_1^2}}\,, \qquad \sin\vartheta_1 = \frac{y_1}{\sqrt{x_1^2 + y_1^2}}\,.$$

Capitolo 5

Sottodifferenziale

5.1 Definizione e proprietà

Se occorre minimizzare un funzionale che non è differenziabile, i metodi ed i risultati visti nei capitoli precedenti non sono utilizzabili. Vi sono però altre tecniche, legate a generalizzazioni del concetto di derivata, che funzionano in questi contesti più generali. La più importante di tali generalizzazioni, nell'ottica dei problemi di minimo ed in stretta relazione con la convessità, è la nozione di sottodifferenziabilità.

Definizione 5.1.1 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Diciamo che f è* sottodifferenziabile *nel punto $x_0 \in X$ se esiste un'applicazione affine $\psi : X \to \mathbb{R}$ tale che*

$$\psi(x) \leq f(x) \quad \forall x \in X, \qquad \psi(x_0) = f(x_0).$$

Se f è definita in un aperto $U \subset X$, diremo che f è sottodifferenziabile nel punto $x_0 \in U$ se è sottodifferenziabile in x_0 l'estensione di f che vale $+\infty$ in $X \setminus U$.

Osservazione 5.1.2 Dalla sottodifferenziabilità di una funzione f in un punto x_0 seguono subito due fatti. Il primo, banale ma importante, è che $f(x_0) \in \mathbb{R}$; il secondo è che esiste $\varphi \in X^*$ tale che

$$f(x) \geq f(x_0) + \varphi(x - x_0) \qquad \forall x \in X.$$

Infatti, se $\psi(x) = \varphi x + b$ è l'applicazione affine che verifica la definizione di sottodifferenziabilità, si ha per ogni $x \in X$

$$f(x) \geq \psi(x) = \varphi(x-x_0) + \varphi x_0 + b = \varphi(x-x_0) + \psi(x_0) = \varphi(x-x_0) + f(x_0);$$

si noti che, di conseguenza, f è semicontinua inferiormente nel punto x_0. D'altra parte, se f è semicontinua inferiormente nel punto x_0, allora per la proposizione 2.3.7 f possiede funzioni affini minoranti, e quindi f è sottodifferenziablie in x_0.

Definizione 5.1.3 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Se f è sottodifferenziabile in x_0, ogni $\varphi \in X^*$ tale che*

$$f(x) \geq f(x_0) + \varphi(x - x_0) \qquad \forall x \in X$$

si chiama sottogradiente *di f in x_0; l'insieme dei sottogradienti di f in x_0 si chiama* sottodifferenziale *di f in x_0 e si indica col simbolo $\partial f(x_0)$.*

Osservazione 5.1.4 Se non esistono sottogradienti di f in x_0, evidentemente $\partial f(x_0) = \emptyset$; questo accade certamente se $f(x_0) = \pm\infty$, ma può anche capitare in un punto x_0 in cui $f(x_0) \in \mathbb{R}$, ad esempio se f è una funzione concava su $\mathbb{R}$. Se $\varphi \in X^*$, si ha $\varphi \in \partial f(x_0)$ se e solo se $f(x_0) \in \mathbb{R}$ e $f(x) \geq f(x_0) + \varphi(x - x_0)$ per ogni $x \in X$. In particolare, $\partial f(x_0)$ è un sottoinsieme convesso (eventualmente vuoto) di X^*.

Esempi 5.1.5 **(1)** Sia X uno spazio normato reale e sia $f : X \to \mathbb{R}$ convessa e G-differenziabile; allora si ha $\partial f(x_0) = \{f'_G(x_0)\}$ per ogni $x_0 \in X$. Infatti, se $\varphi \in \partial f(x_0)$ si ha, per l'osservazione precedente,

$$f(x_0 + tv) - f(x_0) \geq t\, \varphi v \qquad \forall t \in \mathbb{R} \setminus \{0\},\ \forall v \in X;$$

dividendo per t si ottiene per $t \to 0$

$$\begin{cases} f'_G(x_0)v \geq \varphi v & \text{se } t \to 0^+, \\ f'_G(x_0)v \leq \varphi v & \text{se } t \to 0^-, \end{cases}$$

da cui $f'_G(x_0) = \varphi$. Viceversa, poiché f è convessa, il suo rapporto incrementale in una qualunque direzione $v \in X$ è crescente (esercizio 2.1.3): quindi, scegliendo $v = x - x_0$, si ottiene

$$f(x) - f(x_0) = f(x_0 + v) - f(x_0) \geq f'_G(x_0)v = f'_G(x_0)(x - x_0) \qquad \forall x \in X,$$

da cui $f'_G(x_0) \in \partial f(x_0)$.

(2) Sia X uno spazio di Banach reale. Calcoliamo il sottodifferenziale $\partial f(0)$, dove $f(x) = \|x\|_X$. Per definizione si ha $\varphi \in \partial f(0)$ se e solo se $\|x\|_X \geq \varphi x$ per ogni $x \in X$; scrivendo $-x$ al posto di x otteniamo anche $\|x\|_X \geq -\varphi x$, quindi $\varphi \in \partial f(0)$ se e solo se $\|\varphi\|_{X^*} \leq 1$.

Vediamo adesso alcune proprietà del sottodifferenziale.

Proposizione 5.1.6 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Per ogni $x_0 \in X$ il sottodifferenziale $\partial f(x_0)$ è un sottoinsieme convesso e debolmente* chiuso di X^*; in particolare, $\partial f(x_0)$ è chiuso in X^*. Se, inoltre, f è continua in x_0, allora $\partial f(x_0)$ è debolmente* compatto.*

Dimostrazione Se $f(x_0) = \pm\infty$, allora $\partial f(x_0) = \emptyset$ e quindi la tesi è provata. Sia dunque $f(x_0) \in \mathbb{R}$: allora la convessità di $\partial f(x_0)$ segue dall'osservazione 5.1.4. Proviamo che $\partial f(x_0)$ è debolmente* chiuso: se $\varphi \notin \partial f(x_0)$, esiste $x \in X$ tale che $f(x_0) + \varphi(x - x_0) - f(x) = \varepsilon_0 > 0$; fissato l'intorno

$$U = \left\{ \psi \in X^* : \ |(\varphi - \psi)(x - x_0)| < \frac{\varepsilon_0}{2} \right\},$$

si ha $\psi \notin \partial f(x_0)$ per ogni $\psi \in U$, in quanto

$$f(x_0) + \psi(x - x_0) - f(x) \geq f(x_0) + \varphi(x - x_0) - f(x) - \frac{\varepsilon_0}{2} = \frac{\varepsilon_0}{2} > 0.$$

Dunque il complementare di $\partial f(x_0)$ è aperto nella topologia debole*.
Supponiamo adesso che f sia continua in x_0: scelto $\varepsilon = 1$, esiste $\delta > 0$ tale che $|f(x) - f(x_0)| \leq 1$ per $\|x - x_0\|_X \leq \delta$. Sia allora $\varphi \in \partial f(x_0)$: dalla relazione

$$\varphi(x - x_0) \leq f(x) - f(x_0) \qquad \forall x \in X$$

segue subito

$$|\varphi(x - x_0)| \leq |f(x) - f(x_0)| \leq 1 \quad \text{per } \|x - x_0\|_X \leq \delta,$$

e quindi, scelto $x = x_0 + \delta u$, con $\|u\|_X \leq 1$,

$$\delta|\varphi u| \leq 1 \qquad \text{per } \|u\|_X \leq 1.$$

Ciò significa che

$$\|\varphi\|_{X^*} \leq \frac{1}{\delta} \qquad \forall \varphi \in \partial f(x_0).$$

Dunque $\partial f(x_0)$ è un sottoinsieme w^*-chiuso e limitato in X^*; pertanto, in virtù del teorema di Banach-Alaoglu (teorema 1.6.3), $\partial f(x_0)$ è w^*-compatto.
□

Osservazione 5.1.7 Dalla dimostrazione precedente segue facilmente che se f è lipschitziana in una palla $B(x_0, \delta)$ di uno spazio normato reale X, con costante di Lipschitz M, allora per $x \in B(x_0, \delta/2)$ i sottodifferenziali $\partial f(x)$ sono insiemi equilimitati dalla costante M.

L'importanza della nozione di sottodifferenziale in relazione ai problemi di minimo è ben evidenziata dalla seguente

Proposizione 5.1.8 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Se $x_0 \in X$, risulta*
$$f(x_0) = \min_{x \in X} f(x) \in \mathbb{R}$$
se e solo se $0 \in \partial f(x_0)$.

Dimostrazione È una diretta conseguenza della definizione di sottodifferenziale. $\square$

Il sottodifferenziale di una funzione ha una chiara parentela con la bipolare (della stessa funzione), introdotta nel paragrafo 2.5. Il legame fra questi due oggetti viene messo in luce dalla proposizione che segue (si veda anche il teorema 2.5.8).

Proposizione 5.1.9 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Si ha:*

(i) *se $\partial f(x_0) \neq \emptyset$, allora $f(x_0) = f^{**}(J_{x_0})$;*

(ii) *se $f(x_0) = f^{**}(J_{x_0})$, allora $\partial f(x_0) = \partial(f^{**} \circ J)(x_0)$.*

Dimostrazione (i) Sia ψ una funzione affine tale che $\psi(x) \leq f(x)$ per ogni $x \in X$ e $\psi(x_0) = f(x_0)$; allora, essendo in particolare ψ convessa e semicontinua inferiormente, per l'osservazione 2.5.9 si ha $\psi(x) \leq f^{**}(J_x) \leq f(x)$ per ogni $x \in X$; in particolare, per $x = x_0$ si ottiene $f^{**}(J_{x_0}) = \psi(x_0) = f(x_0)$.
(ii) Supponiamo che $f^{**}(J_{x_0}) = f(x_0)$. Se $f(x_0) = \pm\infty$, allora $f^{**}(J_{x_0}) = \pm\infty$ e quindi $\partial f(x_0) = \partial(f^{**} \circ J)(x_0) = \emptyset$. Se invece $f(x_0) \in \mathbb{R}$ (il che non esclude che $\partial f(x_0)$ sia ugualmente vuoto!), sia $\varphi \in \partial f(x_0)$: allora per definizione
$$f(x) \geq f(x_0) + \varphi(x - x_0) = f^{**}(J_{x_0}) + \varphi(x - x_0) \qquad \forall x \in X;$$
poiché $x \mapsto f^{**}(J_{x_0}) + \varphi(x - x_0)$ è una funzione affine, dunque convessa e semicontinua inferiormente, si deduce
$$f^{**}(J_x) \geq f^{**}(J_{x_0}) + \varphi(x - x_0) \qquad \forall x \in X,$$

da cui $\varphi \in \partial(f^{**} \circ J)(x_0)$.
Se, viceversa, $\varphi \in \partial(f^{**} \circ J)(x_0)$, allora

$$f(x) \geq f^{**}(J_x) \geq f^{**}(J_{x_0}) + \varphi(x - x_0) = f(x_0) + \varphi(x - x_0) \qquad \forall x \in X,$$

cioè $\varphi \in \partial f(x_0)$. $\square$

Un'altra caratterizzazione del sottodifferenziale è fornita dal seguente risultato:

Proposizione 5.1.10 *Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Fissati $\varphi \in X^*$ e $x_0 \in X$, risulta $\varphi \in \partial f(x_0)$ se e solo se $f(x_0) \in \mathbb{R}$ e $f(x_0) + f^*(\varphi) = \varphi x_0$; in particolare, in tal caso si ha $f^*(\varphi) \in \mathbb{R}$.*

Dimostrazione Per l'osservazione 5.1.4, si ha $\varphi \in \partial f(x_0)$ se e solo se $f(x_0) \in \mathbb{R}$ e

$$f(x) \geq f(x_0) + \varphi(x - x_0) \qquad \forall x \in X,$$

ovvero, moltiplicando per -1, se e solo se $f(x_0) \in \mathbb{R}$ e

$$f^*(\varphi) = \sup_{x \in X}\{\varphi x - f(x)\} = \varphi x_0 - f(x_0).$$

In particolare, se $\varphi \in \partial f(x_0)$ questa relazione implica $f^*(\varphi) \in \mathbb{R}$. $\square$

Corollario 5.1.11 *Sia X uno spazio normato reale, sia $f : X \to \overline{\mathbb{R}}$ e siano $x_0 \in X$ e $\varphi \in X^*$. Si ha $\varphi \in \partial f(x_0)$ se e solo se $J_{x_0} \in \partial f^*(\varphi)$ e $f(x_0) = f^{**}(J_{x_0})$.*

Dimostrazione ($\Longrightarrow$) Sia $\varphi \in \partial f(x_0)$; allora $f(x_0) \in \mathbb{R}$ e, per la proposizione 5.1.9, si ha $f(x_0) = f^{**}(J_{x_0})$. Dalla proposizione 5.1.10 otteniamo allora $f^*(\varphi) \in \mathbb{R}$ e

$$\varphi x_0 = f(x_0) + f^*(\varphi) = f^{**}(J_{x_0}) + f^*(\varphi),$$

da cui, sempre per la proposizione 5.1.10 (applicata stavolta a f^* nel punto φ) segue $J_{x_0} \in \partial f^*(\varphi)$.
($\Longleftarrow$) Sia $J_{x_0} \in \partial f^*(\varphi)$: per la proposizione 5.1.10 si ha $f^*(\varphi) \in \mathbb{R}$ e $f^{**}(J_{x_0}) + f^*(\varphi) = J_{x_0}\varphi$, da cui, essendo $f(x_0) = f^{**}(J_{x_0})$,

$$f(x_0) + f^*(\varphi) = f^{**}(J_{x_0}) + f^*(\varphi) = J_{x_0}\varphi = \varphi x_0\,.$$

Quindi $f(x_0) \in \mathbb{R}$ ed infine, applicando ancora una volta la proposizione 5.1.10, si conclude che $\varphi \in \partial f(x_0)$. $\square$

Esercizi 5.1

1. Sia X uno spazio normato reale e sia $f : X \to \overline{\mathbb{R}}$. Provare che
$$\partial(\lambda f)(x_0) = \lambda \partial f(x_0) \qquad \forall \lambda \geq 0, \quad \forall x_0 \in X.$$

2. Sia X uno spazio normato reale e siano $f, g : X \to \overline{\mathbb{R}}$. Provare che
$$\partial(f + g)(x_0) \supseteq \partial f(x_0) + \partial g(x_0);$$
si mostri che l'inclusione è in generale stretta.

3. Fissato $\alpha > 0$, si determini $\partial f(0)$, ove $f(x) = |x|^\alpha$, $x \in \mathbb{R}$.

4. Sia X uno spazio di Banach reale e sia K un sottoinsieme convesso e chiuso di X. Si provi che, posto
$$N_K(x_0) = \{\varphi \in X^* : \ \varphi(x - x_0) \leq 0 \ \ \forall x \in K\},$$
risulta
$$\partial I_K(x_0) = \begin{cases} N_K(x_0) & \text{se } x_0 \in K \\ \emptyset & \text{se } x_0 \notin K. \end{cases}$$
Nel caso $X = \mathbb{R}^N$, si interpreti geometricamente questo risultato. (L'insieme $N_K(x_0)$ si chiama *cono normale* a K in x_0.)

5. Sia $g \in L^\infty(a, b)$ e poniamo $f(x) = \int_a^x g(t)dt$. Si provi che $\partial f(x)$ coincide con l'intervallo (eventualmente vuoto) $[g^-(x), g^+(x)]$, ove
$$g^-(x) = \sup_{a \leq \xi < x} \frac{1}{\xi - x} \int_x^\xi g(t)\, dt, \quad g^+(x) = \inf_{x < \xi \leq b} \frac{1}{\xi - x} \int_x^\xi g(t)\, dt.$$

5.2 Funzioni convesse sottodifferenziabili

La nozione di sottodifferenziale, come si è già osservato, è particolarmente importante ed utile quando si considerano funzioni convesse. Per semplicità, considereremo solamente funzioni a valori in $]-\infty, +\infty]$. Innanzi tutto, il sottodifferenziale di una funzione convessa è "quasi sempre" non vuoto, come mostra la proposizione che segue.

Proposizione 5.2.1 *Sia $f : X \to]-\infty, +\infty]$ una funzione convessa su uno spazio normato reale X. Se esiste $x_0 \in D(f)$ tale che f è continua in x_0, allora si ha $\partial f(x) \neq \emptyset$ per ogni $x \in \overset{\circ}{D}(f)$ e, in particolare, $\partial f(x_0) \neq \emptyset$.*

Dimostrazione Poiché f non assume il valore $-\infty$, si ha $f(x_0) \in \mathbb{R}$; la continuità di f implica allora che f è limitata in un intorno U di x_0, da cui $U \subseteq D(f)$ e pertanto $x_0 \in \overset{\circ}{D}(f)$. Per il teorema 2.4.2, si deduce che f è continua in ogni punto di $\overset{\circ}{D}(f)$ ed è limitata in un opportuno intorno di ciascun punto di $\overset{\circ}{D}(f)$.

Ciò premesso, sia $x \in \overset{\circ}{D}(f)$. Adoperiamo un ragionamento già impiegato nella dimostrazione del lemma 2.6.9: l'insieme $\mathrm{epi}(f)$ è convesso in $X \times \mathbb{R}$; proviamo che esso ha parte interna non vuota. Sia V un intorno aperto di x tale che $|f(\xi)| \le K$ per ogni $\xi \in V$: allora $V \times]K, +\infty[$ è un aperto contenuto in $\mathrm{epi}(f)$, cosicché per ogni $t > K$ il punto (x,t) è interno ad $\mathrm{epi}(f)$. Si osservi inoltre che $(x, f(x)) \in \partial\, \mathrm{epi}(f)$, in quanto $(x, f(x) - \varepsilon) \notin \mathrm{epi}(f)$ per ogni $\varepsilon > 0$: in particolare, $(x, f(x)) \notin \overset{\circ}{\mathrm{epi}}(f)$.
Dunque, per il lemma 2.3.6 (i), esistono $\varphi \in X^*$ e $t > 0$ tali che

$$\varphi x + t f(x) = \beta \le \varphi\xi + ta \qquad \forall (\xi, a) \in \mathrm{epi}(f).$$

Dividendo per t e scegliendo $a = f(\xi)$, troviamo

$$\frac{\beta}{t} - \frac{\varphi}{t}\,\xi \le f(\xi) \quad \forall \xi \in X, \qquad \frac{\beta}{t} - \frac{\varphi}{t}\,x = f(x).$$

Per sottrazione si ottiene

$$-\frac{\varphi}{t}(\xi - x) + f(x) \le f(\xi) \qquad \forall \xi \in X,$$

e dunque $-\frac{\varphi}{t} \in \partial f(x)$. Abbiamo così provato che $\partial f(x) \neq \emptyset$ per ogni $x \in \overset{\circ}{D}(f)$. Poiché $x_0 \in \overset{\circ}{D}(f)$, in particolare $\partial f(x_0) \neq \emptyset$. $\square$

Osserviamo che se f assume anche il valore $-\infty$, il risultato della proposizione precedente non è vero (esercizio 5.2.1).
Vediamo ora come è strutturato il sottodifferenziale di una funzione convessa.

Teorema 5.2.2 *Sia $f : X \to]-\infty, +\infty]$ una funzione convessa su uno spazio normato reale X. Allora, posto per ogni $x \in D(f)$*

$$D^+ f(x,v) = \lim_{h \to 0^+} \frac{f(x+hv) - f(x)}{h}\ , \ D^- f(x,v) = \lim_{h \to 0^-} \frac{f(x+hv) - f(x)}{h}\ ,$$

si ha per $x \in D(f)$ *e* $v \in X$

$$\begin{aligned}\partial f(x) &= \{\varphi \in X^* : \ \varphi v \le D^+ f(x,v) \ \ \forall v \in X\} = \\ &= \{\varphi \in X^* : \ \varphi v \ge D^- f(x,v) \ \ \forall v \in X\}.\end{aligned}$$

Inoltre, se esiste $x_0 \in D(f)$ *tale che* f *è continua in* x_0*, allora per ogni* $x \in \overset{\circ}{D}(f)$ *e per ogni* $v \in X$ *risulta*

$$D^+ f(x,v) = \sup\{\varphi v : \ \varphi \in \partial f(x)\}, \quad D^- f(x,v) = \inf\{\varphi v : \ \varphi \in \partial f(x)\}.$$

Dimostrazione Per l'esercizio 2.1.3, per ogni $x \in D(f)$ i rapporti incrementali di f lungo un'arbitraria direzione $v \in X$ sono crescenti e quindi i limiti $D^\pm f(x,v)$ esistono, finiti od infiniti, con $D^- f(x,v) \le D^+ f(x,v)$.
Dimostriamo le inclusioni ($\subseteq$). Sia $x \in D(f)$ e sia $\varphi \in \partial f(x)$: si ha

$$f(\xi) - f(x) \ge \varphi(\xi - x) \qquad \forall \xi \in X,$$

quindi

$$f(x + tv) - f(x) \ge t\ \varphi v \qquad \forall t \in \mathbb{R} \setminus \{0\}, \ \forall v \in X.$$

Dividendo per t otteniamo per $t \to 0^+$ e $t \to 0^-$

$$D^+ f(x,v) \ge \varphi v \ge D^- f(x,v) \qquad \forall v \in X.$$

Viceversa, se $\varphi \in X^*$ verifica la prima di queste due disuguaglianze, allora per la crescenza dei rapporti incrementali di f si ha a maggior ragione

$$f(x + v) - f(x) \ge \varphi v \qquad \forall v \in X,$$

da cui, posto $v = \xi - x$, si deduce

$$f(\xi) \ge f(x) + \varphi(\xi - x) \qquad \forall \xi \in X,$$

ossia $\varphi \in \partial f(x)$. Se invece φ verifica la seconda delle due disuguaglianze precedenti, allora analogamente

$$\varphi v \ge f(x) - f(x - v) \qquad \forall v \in X,$$

e posto $v = x - \xi$ si arriva nuovamente alla conclusione che $\varphi \in \partial f(x)$. Ciò prova le inclusioni ($\supseteq$).

Proviamo la seconda parte della tesi. Per la proposizione 5.2.1, $\partial f(x)$ è non vuoto per ogni $x \in \overset{\circ}{D}(f)$. È chiaro che per ogni $v \in X$ si ha

$$D^+ f(x,v) \geq \sup\{\varphi v : \ \varphi \in \partial f(x)\}, \ D^- f(x,v) \leq \inf\{\varphi v : \ \varphi \in \partial f(x)\};$$

dimostriamo le disuguaglianze opposte. Fissiamo $x \in \overset{\circ}{D}(f)$. Essendo f continua in x_0, per il teorema 2.4.2 f è continua in ogni punto di $\overset{\circ}{D}(f)$, quindi f è limitata in un'opportuna palla $B(x, 2\delta)$; per la proposizione 2.4.5, f è localmente lipschitziana in $B(x, 2\delta)$ e pertanto esiste $K \geq 0$ tale che

$$\begin{aligned} -K\|v\|_X &\leq f(x-v) - f(x) \leq D^- f(x,v) \leq \\ &\leq D^+ f(x,v) \leq f(x+v) - f(x) \leq K\|v\|_X \qquad \forall v \in B(0,\delta). \end{aligned}$$

Dato che le funzioni $v \mapsto D^{\pm} f(x,v)$ sono positivamente omogenee (oltre che convesse), deduciamo che la stessa relazione vale per ogni $v \in X$, ossia

$$-K\|v\|_X \leq D^- f(x,v) \leq D^+ f(x,v) \leq K\|v\|_X \qquad \forall v \in X.$$

Fissato $v \in X \setminus \{0\}$, andiamo ora a costruire un sottogradiente $\overline{\psi} \in \partial f(x)$ che verifichi l'uguaglianza $\overline{\psi} v = D^+ f(x,v)$: ciò mostrerà che

$$D^+ f(x,v) \leq \sup\{\varphi v : \ \varphi \in \partial f(x)\}.$$

Consideriamo il funzionale $y \mapsto D^+ f(x,y)$, che è positivamente omogeneo e convesso, e definiamo un funzionale lineare $\psi : [\{v\}] \to \mathbb{R}$ ponendo

$$\psi(tv) = t\ D^+ f(x,v) \qquad \forall t \in \mathbb{R}.$$

Si ha $\psi(tv) \leq D^+ f(x, tv)$ per ogni $t \in \mathbb{R}$: infatti ciò è banale se $t \geq 0$, mentre se $t < 0$ possiamo scrivere, in virtù della relazione $D^+ f(x,-v) = -D^- f(x,v)$,

$$\begin{aligned} \psi(tv) &= -\psi(|t|v) = \\ &= -D^+ f(x,|t|v) = -D^+ f(x,-tv) = D^- f(x,tv) \leq D^+ f(x,tv). \end{aligned}$$

Per il teorema di Hahn-Banach esiste $\overline{\psi} : X \to \mathbb{R}$ tale che

$$\overline{\psi}(tv) = \psi(tv) \quad \forall t \in \mathbb{R}, \qquad \overline{\psi} y \leq D^+ f(x,y) \quad \forall y \in X.$$

Si ha $\overline{\psi} \in X^*$, perché

$$|\overline{\psi}y| = \max\{\overline{\psi}y, \overline{\psi}(-y)\} \leq \max\{D^+f(x,y), D^-f(x,y)\} \leq K\|y\|_X \quad \forall y \in X.$$

Di conseguenza, $\overline{\psi} \in \partial f(x)$ per la caratterizzazione già dimostrata. Inoltre, per costruzione, $\overline{\psi}v = D^+f(x,v)$.
L'analoga disuguaglianza relativa a $D^-f(x,v)$ segue utilizzando nuovamente la relazione $D^-f(x,v) = -D^+f(x,-v)$. $\square$

Vediamo ora la relazione fra sottodifferenziabilità e G-differenziabilità di funzioni convesse, andando a completare il risultato dell'esempio 5.1.5(1).

Proposizione 5.2.3 *Sia X uno spazio normato reale, e sia $f : X \to]-\infty, +\infty]$ una funzione convessa e propria; sia inoltre $x_0 \in D(f)$.*

(i) *Se f è G-differenziabile in x_0, allora f è sottodifferenziabile in x_0 e $\partial f(x_0) = \{f'_G(x_0)\}$.*

(ii) *Se f è continua in x_0 e $\partial f(x_0) = \{\varphi_0\}$, allora f è G-differenziabile in x_0 e $f'_G(x_0) = \varphi_0$.*

Dimostrazione (i) Basta ripetere il ragionamento adoperato nell'esempio 5.1.5 (1).
(ii) Dal teorema 5.2.2 segue, per ogni $v \in X$,

$$D^+f(x_0,v) = \sup\{\varphi v : \ \varphi \in \partial f(x_0)\} = \varphi_0 v,$$

$$D^-f(x_0,v) = \inf\{\varphi v : \ \varphi \in \partial f(x_0)\} = \varphi_0 v,$$

e dunque esiste la derivata di Gâteaux $f'_G(x_0)v = \varphi_0 v$ per ogni $v \in X$. $\square$

Dimostriamo adesso il teorema di Mazur (teorema 2.4.6), che qui richiamiamo:

Teorema 2.4.6 (di Mazur) Sia X uno spazio di Banach separabile, sia $D \subseteq X$ un aperto convesso non vuoto e sia $f : D \to \mathbb{R}$ una funzione convessa e continua. Allora f è G-differenziabile in un insieme $D_0 \subseteq D$, denso in D, che è unione numerabile di aperti densi.

Dimostrazione Sia

$$D_0 = \{x \in D : \ \exists f'_G(x)\},$$

e dimostriamo che $D \setminus D_0$ è unione numerabile di insiemi chiusi in D.
Sia $\{x_n\}_{n\in\mathbb{N}^+}$ una successione densa nella palla unitaria B di X. Per ogni $n, m \in \mathbb{N}^+$ poniamo

$$A_{nm} = \left\{ x \in D : \ \exists \varphi, \psi \in \partial f(x) : \ (\varphi - \psi)(x_n) \geq \frac{1}{m} \right\}.$$

Osserviamo che, per la proposizione 5.2.3, $f'_G(x)$ non esiste se e solo se $\partial f(x)$ contiene almeno due elementi distinti: dato che due elementi distinti di X^* devono assumere valori distinti in almeno un punto della palla unitaria B, dalla densità di $\{x_n\}_{n\in\mathbb{N}^+}$ si ricava facilmente che

$$D \setminus D_) = \bigcup_{n,m\in\mathbb{N}^+} A_{nm} \, .$$

Basta allora provare che ciascun A_{nm} è chiuso in D. Sia dunque $\{z_k\}_{k\in\mathbb{N}} \subseteq A_{nm}$ una successione convergente ad un elemento $z \in D$, e proviamo che $z \in A_{nm}$. Per ogni $k \in \mathbb{N}$ esistono $\varphi_k, \psi_k \in \partial f(z_k)$ tali che $(\varphi_k - \psi_k)(x_n) \geq \frac{1}{m}$. Notiamo adesso che, per la proposizione 2.4.5, f è lipschitziana in una palla $B(z, \delta)$; l'osservazione 5.1.7 ci dice allora che i sottodifferenziali $\partial f(z_k)$ sono equilimitati per ogni $k \geq k_0$, ove k_0 è la soglia a partire dalla quale si ha $z_k \in B(z, \delta/2)$. Quindi esiste $M \geq 0$ tale che $\|\varphi_k\|_{X^*}, \|\psi_k\|_{X^*} \leq M$ per ogni $k \in \mathbb{N}$. Per debole* compattezza possiamo supporre, a meno di sottosuccessioni, che $\varphi_k \overset{*}{\rightharpoonup} \varphi$ e $\psi_k \overset{*}{\rightharpoonup} \psi$ in X^*. Ne segue, facilmente, $\varphi_k z_k \to \varphi z$ in $\mathbb{R}$ e pertanto

$$\varphi(y - x) = \lim_{k\to\infty} \varphi_k(y - z_k) \leq \lim_{k\to\infty} [f(y) - f(z_k)] = f(y) - f(z) \quad \forall y \in X,$$

cioè $\varphi \in \partial f(z)$; analogamente si vede che $\psi \in \partial f(z)$. Inoltre

$$(\varphi - \psi)(x_n) = \lim_{k\to\infty} (\varphi_k - \psi_k)(x_n) \geq \frac{1}{m} \, ,$$

il che prova che $z \in A_{nm}$ e dunque A_{nm} è chiuso in D.
Dimostriamo adesso che $D \setminus A_{nm}$ è denso in D: da ciò seguirà che anche $D \setminus D_0 = \bigcap_{n,m\in\mathbb{N}^+} [D \setminus A_{nm}]$ è denso in D, concludendo così la dimostrazione.
Sia $x_0 \in D$. Per ogni $n \in \mathbb{N}^+$ sia $I_n = \{s \in \mathbb{R} : \ x_0 + s x_n \in D\}$: I_n è un intervallo aperto non vuoto. La funzione $f_1(r) = f(x_0 + r x_n)$, $r \in I_N$, è derivabile salvo al più un insieme numerabile di punti (esercizio 5.2.5): dunque esiste una successione di punti della forma $x'_h = x_0 + r_h x_n$, convergente a x_0, tale

che $f_1'(r_h)$ esiste per ogni $h \in \mathbb{N}$. Ne segue che se $\varphi, \psi \in \partial f(x_h')$ le restrizioni di φ e ψ allo spazio $[\{x_n\}]$ appartengono a $\partial f_1(r_h)$ per ogni h, ed essendo $\partial f_1(r_h) = \{f_1'(r_h)\}$ si deduce $\varphi x_n = \psi x_n$. Ciò mostra che $x_h' \notin A_{nm}$ per ogni $m \in \mathbb{N}^+$, e che quindi il punto x_0 è approssimabile da punti $x_h \in D \setminus A_{nm}$. La tesi è così provata. $\square$

Proviamo infine due risultati che estendono al caso di spazi normati due ben note caratterizzazioni della convessità per funzioni derivabili su $\mathbb{R}$.

Proposizione 5.2.4 *Sia X uno spazio normato reale, sia $K \subseteq X$ un convesso non vuoto e sia $f : K \to \mathbb{R}$ una funzione G-differenziabile. Allora f è convessa se e solo se*

$$f(\xi) \geq f(x) + f_G'(x)(\xi - x) \qquad \forall \xi, x \in K.$$

Dimostrazione Sia f convessa: dopo averla prolungata a tutto X ponendola uguale a $+\infty$ nei punti di $X \setminus K$, siamo nelle ipotesi della proposizione 5.2.3; ne segue $\partial f(x) = \{f_G'(x)\}$ per ogni $x \in K$, e da qui segue subito la disuguaglianza cercata.
Viceversa, applichiamo la disuguaglianza alle coppie di punti $(u, (1-\lambda)u+\lambda v)$ e $(v, (1-\lambda)u + \lambda v)$, con $u, v \in K$ e $\lambda \in [0, 1]$: si ottiene

$$f(u) \geq f((1-\lambda)u + \lambda v) + \lambda f_G'((1-\lambda)u + \lambda v)(u - v),$$

$$f(v) \geq f((1-\lambda)u + \lambda v) + (1-\lambda) f_G'((1-\lambda)u + \lambda v)(v - u).$$

Moltiplichiamo la prima disuguaglianza per $(1-\lambda)$ e la seconda per λ, e poi sommiamo: il risultato è

$$(1-\lambda)f(u) + \lambda f(v) \geq f((1-\lambda)u + \lambda v),$$

cosicché f è convessa. $\square$

Proposizione 5.2.5 *Sia X uno spazio normato reale, sia $K \subseteq X$ un convesso non vuoto e sia $f : K \to \mathbb{R}$ una funzione G-differenziabile. Allora f è convessa se e solo se il suo differenziale di Gâteaux $f_G' : K \to X^*$ è un operatore* monotono*, ossia*

$$[f_G'(u) - f_G'(v)](u - v) \geq 0 \qquad \forall u, v \in K.$$

Dimostrazione Sia f convessa: fissati $u, v \in K$, per la proposizione 5.2.4 si ha
$$f(v) \geq f(u) + f'_G(u)(v-u), \quad f(u) \geq f(v) + f'_G(v)(u-v);$$
sommando queste due disuguaglianze si ha
$$0 \geq f'_G(u)(v-u) + f'_G(v)(u-v),$$
da cui la monotonia di f'_G.
Viceversa, sia f'_G monotono. Fissati $u, v \in K$, poniamo
$$\phi(\lambda) = f((1-\lambda)u + \lambda v), \qquad \lambda \in [0,1].$$
Chiaramente ϕ è derivabile e
$$\phi'(\lambda) = f'_G((1-\lambda)u + \lambda v)(v-u) \qquad \forall \lambda \in [0,1];$$
dalla monotonia di f'_G segue che se $\lambda < \mu$ si ha
$$\begin{aligned} \phi'(\mu) - \phi'(\lambda) &= \Big[f'_G((1-\mu)u + \mu v) - f'_G((1-\lambda)u + \lambda v)\Big](v-u) = \\ &= \frac{1}{\mu - \lambda}\Big[f'_G((1-\mu)u + \mu v) - f'_G((1-\lambda)u + \lambda v)\Big] \cdot \\ &\quad \cdot \Big[[(1-\mu)u + \mu v] - [(1-\lambda)u + \lambda v]\Big] \geq 0. \end{aligned}$$
In altre parole, $\phi' : \mathbb{R} \to \mathbb{R}$ è crescente: ne segue che ϕ è convessa in $[0,1]$. Pertanto
$$f((1-\lambda)u + \lambda v) = \phi(\lambda) \leq (1-\lambda)\phi(0) + \lambda\phi(1) = (1-\lambda)f(u) + \lambda f(v),$$
cioè f è convessa su K. $\square$

Osservazione 5.2.6 Con lo stesso ragionamento usato nella prima parte della dimostrazione precedente, è facile dedurre la seguente proprietà: se X è uno spazio normato reale e se $f : X \to \overline{\mathbb{R}}$ è una funzione convessa, allora si ha
$$(\psi - \varphi)(u - v) \geq 0 \qquad \forall \psi \in \partial f(u), \ \forall \varphi \in \partial f(v)$$
(talvolta naturalmente potrà succedere che i sottodifferenziali in questione siano vuoti). Si esprime questo fatto dicendo che il sottodifferenziale è monotono, quantunque esso non sia un operatore da X in X^*; ciò troverà piena giustificazione più avanti, nel più ampio contesto delle multifunzioni.

Esercizi 5.2

1. Si costruisca una funzione $f : \mathbb{R} \to \overline{\mathbb{R}}$ convessa, tale che $0 \in \overset{\circ}{D}(f)$ e f sia continua in 0, ma risulti $\partial f(x) = \emptyset$ per ogni $x \in \overset{\circ}{D}(f)$.

2. Sia K un convesso non vuoto dello spazio normato reale X, e sia $f : K \to \mathbb{R}$ una funzione G-differenziabile. Si provi che f è strettamente convessa se e solo se
$$f(v) > f(u) + f'_G(u)(v-u) \qquad \forall u, v \in K \text{ con } u \neq v.$$

3. Sia X uno spazio normato reale e siano $f, g : X \to]-\infty, +\infty]$ convesse e proprie. Se esiste $x_0 \in D(f) \cap D(g)$ tale che f è continua in x_0, si provi che risulta
$$\partial(f+g)(x) = \partial f(x) + \partial g(x) \qquad \forall x \in D(f) \cap D(g).$$
[**Traccia:** si osservi che, per l'esercizio 5.1.2, basta provare l'inclusione $(\subseteq)$. Fissato $x \in D(f) \cap D(g)$, sia $\varphi \in X^*$ tale che $f(\xi) + g(\xi) \geq \varphi(\xi - x) + f(x) + g(x)$ per ogni $\xi \in X$. Si definiscano
$$C = \{(\xi, a) \in X \times \mathbb{R} : \ f(\xi) - f(x) - \varphi(\xi - x) \leq a\},$$
$$D = \{(\eta, b) \in X \times \mathbb{R} : \ b \leq g(x) - g(\eta)\};$$
si mostri che C e D sono convessi e che $\overset{\circ}{C} \cap \overset{\circ}{D} = \emptyset$, con $\overset{\circ}{C} \neq \emptyset$. Si applichi il lemma 2.3.6 ai convessi D e $\overset{\circ}{C}$, e si deduca che esistono $\psi \in X^*$, $t > 0$ e $\beta \in \mathbb{R}$ tali che
$$\psi\eta + tb \leq \beta \leq \psi\xi + ta \qquad \forall (\eta, b) \in D, \ \forall (\xi, a) \in C.$$
Scegliendo $\eta = \xi$ e $a = f(\xi) - f(x) - \varphi(\xi - x)$ e $b = g(x) - g(\xi)$, si ricavi la relazione
$$g(x) - g(\xi) \leq \frac{\beta}{t} - \frac{\psi}{t}\,\xi \leq f(\xi) - f(x) - \varphi(\xi - x) \quad \forall \xi \in X.$$
Si concluda che $\beta = \psi x$ e che $\frac{\psi}{t} \in \partial g(x)$, $\varphi - \frac{\psi}{t} \in \partial f(x)$.]

4. Siano X, Y spazi normati reali, sia $\Lambda \in \mathcal{L}(X,Y)$ e sia $f : Y \to]-\infty, +\infty]$ una funzione convessa; supponiamo che esista $y_0 \in D(f)$ tale che f è continua in y_0. Si provi che

$$\partial f \circ \Lambda(x) = \Lambda^*(\partial f(\Lambda x)) \qquad \forall x \in X.$$

[**Traccia:** l'inclusione ($\supseteq$) è facile. Per provare l'altra, si fissi $\varphi \in \partial f \circ \Lambda(x)$, e si definisca

$$L = \{(\Lambda\xi, \varphi(\xi - x) + f(\Lambda x)) \in Y \times \mathbb{R} : \ \xi \in X\}.$$

Si provi che L è convesso ed è disgiunto dal convesso (non vuoto) $\overset{\circ}{\mathrm{epi}(f)}$; si determini un funzionale non nullo $(\psi, t) \in (Y \times \mathbb{R})^*$ che separi i due convessi, si mostri che $t > 0$ e si deduca che $-\frac{\psi}{t} \in \partial f(\Lambda x)$ e che $\varphi = \Lambda^*(-\frac{\psi}{t})$.]

5. Sia $f : [a,b] \to \mathbb{R}$ una funzione convessa. Si provi che f è derivabile in $[a,b] \setminus N$, ove N è un insieme al più numerabile.
[**Traccia:** si utilizzi la ben nota relazione

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(u)}{w - u} \leq \frac{f(w) - f(w)}{w - v} \quad \text{per } a \leq u < v < w \leq b$$

per dimostrare che

$$D^+f(u) \leq D^-f(v) \leq D^+f(v) \leq D^-f(w) \quad \text{per } a \leq u < v < w \leq b;$$

se ne deduca che se f non è derivabile in x_0 allora x_0 è un punto di discontinuità della funzione crescente $x \mapsto D^+f(x)$...]

5.3 Regolarizzata di una funzione convessa

Vi è un modo "canonico" di approssimare le funzioni convesse, valido in ogni spazio di Hilbert reale ed assai utile in molte applicazioni, che andiamo ad illustrare.
Sia H uno spazio di Hilbert reale e sia $f : H \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria. Per ogni $\varepsilon > 0$ ed ogni $x \in H$, consideriamo la funzione

$$g_{\varepsilon,x}(y) = \frac{1}{2\varepsilon}\|x - y\|_H^2 + f(y), \qquad y \in H.$$

Si vede immediatamente che $g_{\varepsilon,x}$ è strettamente convessa, semicontinua inferiormente e propria.

Lemma 5.3.1 *Nelle ipotesi sopra scritte, comunque si fissino $\varepsilon > 0$ ed $x \in H$, la funzione $g_{\varepsilon,x}$ ha un unico punto di minimo in H. Inoltre, un punto $x_0 \in H$ è il punto di minimo di $g_{\varepsilon,x}$ se e solo se risulta $\frac{1}{\varepsilon}(x - x_0) \in \partial f(x_0)$.*

Dimostrazione In virtù della proposizione 2.3.7, esiste una funzione affine ψ tale che $f \geq \psi$ in H, ossia esistono $z \in H$ e $b \in \mathbb{R}$ tali che

$$g_{\varepsilon,x}(y) = \frac{1}{2\varepsilon}\|x - y\|_H^2 + f(y) \geq \frac{1}{2\varepsilon}\|x - y\|_H^2 + \langle z, y\rangle_H + b \qquad \forall y \in H.$$

Ne segue che $g_{\varepsilon,x}(y) \to +\infty$ per $\|y\|_H \to \infty$. Dunque, per il corollario 2.3.4, la funzione $g_{\varepsilon,x}$ ha almeno un punto di minimo in H. Poiché inoltre $g_{\varepsilon,x}$ è strettamente convessa, il punto di minimo è unico.
Supponiamo che $x_0 \in H$ sia tale che $\frac{1}{\varepsilon}(x - x_0) \in \partial f(x_0)$. Allora, per definizione di sottodifferenziale, $f(x_0) \in \mathbb{R}$ e

$$\begin{aligned}
f(y) - f(x_0) &\geq \frac{1}{\varepsilon}\langle x - x_0, y - x_0\rangle_H \geq \\
&\geq \frac{1}{\varepsilon}\langle x - x_0, y - x_0\rangle_H - \frac{1}{2\varepsilon}\|y - x_0\|_H^2 = \\
&= \frac{1}{2\varepsilon}\langle 2x - 2x_0 - y + x_0, y - x_0\rangle_H = \\
&= \frac{1}{2\varepsilon}\langle x - x_0 + x - y, y - x + x - x_0\rangle_H = \\
&= \frac{1}{2\varepsilon}\left(\|x - x_0\|_H^2 - \|x - y\|_H^2\right);
\end{aligned}$$

ne segue che $g_{\varepsilon,x}$ ha minimo nel punto x_0.
Viceversa, sia $x_0 \in H$ l'unico punto di minimo di $g_{\varepsilon,x}$: allora, fissato $y \in H$ e scelto $z = (1 - t)x_0 + ty$, con $t \in]0, 1[$, utilizzando la convessità di f ed osservando che $x - z = x - x_0 - t(y - x_0)$, si può scrivere

$$\begin{aligned}
t[f(y) - f(x_0)] &\geq f(z) - f(x_0) \geq \\
&\geq \frac{1}{2\varepsilon}\left(\|x - x_0\|_H^2 - \|x - z\|_H^2\right) = \\
&= \frac{t}{\varepsilon}\langle x - x_0, y - x_0\rangle_H - \frac{t^2}{2\varepsilon}\|x_0 - y\|_H^2 \qquad \forall t \in]0, 1[,
\end{aligned}$$

da cui, dividendo per t e passando al limite per $t \to 0^+$,

$$f(y) - f(x_0) \geq \frac{1}{\varepsilon}\langle x - x_0, y - x_0\rangle_H \qquad \forall y \in H.$$

Ciò mostra che $\frac{1}{\varepsilon}(x-x_0)$, o meglio il funzionale $z \mapsto \frac{1}{\varepsilon}\langle x-x_0, z\rangle_H$, appartiene a $\partial f(x_0)$. Il lemma è provato (si noti che l'intero enunciato poteva anche dedursi dall'esercizio 5.2.3). $\square$

Indicheremo con $y_{\varepsilon,x}$ l'unico punto di minimo della funzione $g_{\varepsilon,x}$.
Ciò premesso, andiamo a definire la regolarizzata di una funzione convessa.

Definizione 5.3.2 *Sia $f : H \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria su uno spazio di Hilbert reale H. Fissato $\varepsilon > 0$, la* regolarizzata *di f è la funzione così definita:*

$$f_\varepsilon(x) = \min_{y\in H} g_{\varepsilon,x}(y) = g_{\varepsilon,x}(y_{\varepsilon,x}), \qquad x \in H.$$

Naturalmente, la regolarizzata f_ε di una funzione f convessa, semicontinua inferiormente e propria è cosa diversa dalla regolarizzata semicontinua $\overline{f}$ di una funzione convessa f, introdotta con la definizione 2.3.8.
Vediamo ora le principali proprietà della regolarizzata f_ε .

Proposizione 5.3.3 *Sia $f : H \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria su uno spazio di Hilbert reale H. Per ogni $\varepsilon > 0$ sia f_ε la regolarizzata di f. Allora:*

(i) *f_ε è convessa, assume valori reali e verifica*

$$f(y_{\varepsilon,x}) \leq f_\varepsilon(x) \leq f(x) \qquad \forall x \in H;$$

(ii) *f_ε è F-differenziabile in H e*

$$f'_\varepsilon(x) = \frac{1}{\varepsilon}(x - y_{\varepsilon,x}) \qquad \forall x \in H;$$

(iii) *posto $D = \{x \in H : \ \partial f(x) \neq \emptyset\}$, si ha*

$$\left\|\frac{1}{\varepsilon}(x - y_{\varepsilon,x})\right\|_H \leq \min\{\|y\|_H : \ y \in \partial f(x)\} \qquad \forall x \in D;$$

(iv) *l'insieme $\overline{D}$ è convesso e*

$$\lim_{\varepsilon\to 0^+} \|y_{\varepsilon,x} - P_{\overline{D}}(x)\|_H = 0 \qquad \forall x \in H;$$

(v) *risulta*

$$\lim_{\varepsilon\to 0^+} f_\varepsilon(x) = f(x) \qquad \forall x \in H.$$

Dimostrazione (i) Anzitutto $f_\varepsilon(x)$, essendo un minimo, è un numero reale. Per ogni $x \in H$ si ha poi, per definizione,

$$f(y_{\varepsilon,x}) \le \frac{1}{2\varepsilon}\|x - y_{\varepsilon,x}\|_H^2 + f(y_{\varepsilon,x}) = f_\varepsilon(x) = \min_{y\in H} g_{\varepsilon,x}(y) \le g_{\varepsilon,x}(x) = f(x).$$

Proviamo la convessità di f_ε: per ogni $x, x' \in H$ e $\lambda \in [0,1]$ si ha

$$\begin{aligned}
&f_\varepsilon(\lambda x + (1-\lambda)x') = g_{\varepsilon,\lambda x+(1-\lambda)x'}(y_{\varepsilon,\lambda x+(1-\lambda)x'}) \le \\
&\quad \le g_{\varepsilon,\lambda x+(1-\lambda)x'}(\lambda y_{\varepsilon,x} + (1-\lambda)y_{\varepsilon,x'}) = \\
&\quad = \frac{1}{2\varepsilon}\|\lambda(x - y_{\varepsilon,x}) + (1-\lambda)(x' - y_{\varepsilon,x'})\|_H^2 + f(\lambda y_{\varepsilon,x} + (1-\lambda)y_{\varepsilon,x'}) \le \\
&\quad \le \frac{1}{2\varepsilon}\left[\lambda\|x - y_{\varepsilon,x}\|_H^2 + (1-\lambda)\|x' - y_{\varepsilon,x'}\|_H^2\right] + \\
&\qquad\quad + \lambda f(y_{\varepsilon,x}) + (1-\lambda) f(y_{\varepsilon,x'}) = \\
&\quad = \lambda g_{\varepsilon,x}(y_{\varepsilon,x}) + (1-\lambda) g_{\varepsilon,x'}(y_{\varepsilon,x'}) = \lambda f_\varepsilon(x) + (1-\lambda) f_\varepsilon(x').
\end{aligned}$$

(ii) Proviamo anzitutto la disuguaglianza

$$\left\|\frac{1}{\varepsilon}(x - y_{\varepsilon,x}) - \frac{1}{\varepsilon}(x' - y_{\varepsilon,x'})\right\|_H \le \frac{1}{\varepsilon}\|x - x'\|_H \qquad \forall x, x' \in H.$$

In virtù del lemma 5.3.1, si ha $\xi = \frac{1}{\varepsilon}(x - y_{\varepsilon,x}) \in \partial f(y_{\varepsilon,x})$ e $\xi' = \frac{1}{\varepsilon}(x' - y_{\varepsilon,x'}) \in \partial f(y_{\varepsilon,x'})$; ne segue, essendo $x = y_{\varepsilon,x} + \varepsilon\xi$ e $x' = y_{\varepsilon,x'} + \varepsilon\xi'$, e tenendo conto dell'osservazione 5.2.6,

$$\langle \xi - \xi', x - x' \rangle_H = \langle \xi - \xi', \varepsilon(\xi - \xi')\rangle_H + \langle \xi - \xi', y_{\varepsilon,x} - y_{\varepsilon,x'}\rangle_H \ge \varepsilon\|\xi - \xi'\|_H^2\,,$$

da cui la tesi.
Ciò premesso, osservato che risulta, per definizione di sottodifferenziale,

$$f(y_{\varepsilon,x'}) - f(y_{\varepsilon,x}) \ge \langle \xi, y_{\varepsilon,x'} - y_{\varepsilon,x}\rangle_H\,,$$

possiamo scrivere

$$\begin{aligned} f_\varepsilon(x') - f_\varepsilon(x) &= f(y_{\varepsilon,x'}) + \frac{1}{2\varepsilon}\|x' - y_{\varepsilon,x'}\|_H^2 - f(y_{\varepsilon,x}) - \frac{1}{2\varepsilon}\|x - y_{\varepsilon,x}\|_H^2 \geq \\ &\geq \frac{\varepsilon}{2}(\|\xi'\|_H^2 - \|\xi\|_H^2) + \langle \xi, y_{\varepsilon,x'} - y_{\varepsilon,x}\rangle_H = \\ &= \frac{\varepsilon}{2}(\|\xi'\|_H^2 - \|\xi\|_H^2) + \langle \xi, x' - x\rangle_H - \varepsilon\langle \xi, \xi' - \xi\rangle_H\,, \end{aligned}$$

da cui

$$\begin{aligned} f_\varepsilon(x') - f_\varepsilon(x) - \langle \xi, x' - x\rangle_H &\geq \frac{\varepsilon}{2}\left(\|\xi'\|_H^2 - \|\xi\|_H^2 - 2\langle \xi, \xi' - \xi\rangle_H\right) = \\ &= \frac{\varepsilon}{2}\left(\|\xi'\|_H^2 + \|\xi\|_H^2 - 2\langle \xi, \xi'\rangle_H\right) = \\ &= \frac{\varepsilon}{2}\|\xi' - \xi\|_H^2 \geq 0. \end{aligned}$$

D'altra parte, scambiando i ruoli di x e x' abbiamo anche

$$f_\varepsilon(x) - f_\varepsilon(x') - \langle \xi', x - x'\rangle_H \geq 0,$$

e quindi, in virtù della disuguaglianza dimostrata preliminarmente,

$$\begin{aligned} &f_\varepsilon(x') - f_\varepsilon(x) - \langle \xi, x' - x\rangle_H = \\ &\quad = f_\varepsilon(x') - f_\varepsilon(x) - \langle \xi', x' - x\rangle_H + \langle \xi' - \xi, x' - x\rangle_H \leq \\ &\quad \leq \langle \xi' - \xi, x' - x\rangle_H \leq \frac{1}{\varepsilon}\|x' - x\|_H^2\,. \end{aligned}$$

Ne segue

$$0 \leq f_\varepsilon(x') - f_\varepsilon(x) - \langle \xi, x' - x\rangle_H \leq \frac{1}{\varepsilon}\|x' - x\|_H^2\,,$$

cosicché f_ε è F-differenziabile in ogni $x \in H$, con $f_\varepsilon'(x) = \xi = \frac{1}{\varepsilon}(x - y_{\varepsilon,x})$. Ciò prova (ii).

(iii) Dal lemma 5.3.1 segue che $\frac{1}{\varepsilon}(x - y_{\varepsilon,x}) \in \partial f(y_{\varepsilon,x})$. Quindi

$$f(z) - f(y_{\varepsilon,x}) \geq \frac{1}{\varepsilon}\langle x - y_{\varepsilon,x}, z - y_{\varepsilon,x}\rangle_H \qquad \forall z \in H,$$

da cui, scelto $z = x$, ricaviamo

$$\frac{1}{\varepsilon}\,\|x - y_{\varepsilon,x}\|_H^2 \leq f(x) - f(y_{\varepsilon,x}) \leq \langle y, x - y_{\varepsilon,x}\rangle_H \qquad \forall y \in \partial f(x).$$

Ne segue

$$\left\|\frac{1}{\varepsilon}(x-y_{\varepsilon,x})\right\|_H^2 \le \left\langle y, \frac{1}{\varepsilon}(x-y_{\varepsilon,x})\right\rangle_H \qquad \forall y \in \partial f(x),$$

e ciò implica la tesi, essendo l'insieme $\{\|y\|_H : \ y \in \partial f(x)\}$ non vuoto, chiuso e limitato inferiormente in $\mathbb{R}$.

(iv) Sia $x \in H$. Mostriamo anzitutto che l'insieme $\{y_{\varepsilon,x}\}_{\varepsilon\in\,]0,1]}$ è limitato in H. Fissiamo $u \in D$ e $v \in \partial f(u)$: ad esempio potremmo prendere $u = y_{1,x}$ e $v = x - y_{1,x}$. Allora per il lemma 5.3.1 e l'osservazione 5.2.6 si ha

$$\langle x - y_{\varepsilon,x} - \varepsilon v, y_{\varepsilon,x} - u\rangle_H \ge 0 \qquad \forall u \in D, \ \forall v \in \partial f(u).$$

Ne segue la seguente fondamentale disuguaglianza:

$$\|y_{\varepsilon,x}\|_H^2 \le \langle x, y_{\varepsilon,x} - u\rangle_H + \langle y_{\varepsilon,x}, u\rangle_H - \varepsilon\langle v, y_{\varepsilon,x} - u\rangle_H \,.$$

Da questa relazione è facile dedurre una disequazione di secondo grado del tipo

$$\|y_{\varepsilon,x}\|_H^2 \le A\|y_{\varepsilon,x}\|_H + B,$$

la quale implica la limitatezza di $\|y_{\varepsilon,x}\|_H$.

Pertanto l'insieme $\{y_{\varepsilon,x}\}_{\varepsilon\in\,]0,1]}$ è debolmente relativamente compatto nello spazio di Hilbert H ed è contenuto nel convesso chiuso $K = \overline{\mathrm{co}(\overline{D})}$. Sia $\{\varepsilon_n\}$ una successione contenuta in $]0,1]$ e tendente a 0, tale che $y_{\varepsilon_n,x} \rightharpoonup x_0$ in H; allora $x_0 \in K$. Passando al minimo limite nella disuguaglianza fondamentale, e ricordando che $\liminf_{\varepsilon_n\to 0}\|y_{\varepsilon_n,x}\|_H \ge \|x_0\|_H$, si ricava

$$\langle x - x_0, u - x_0\rangle_H \le 0 \qquad \forall u \in D;$$

per chiusura, per convessità e per densità, si deduce

$$\langle x - x_0, u - x_0\rangle_H \le 0 \qquad \forall u \in K.$$

Ma questa disequazione è precisamente quella che caratterizza la proiezione sul convesso chiuso K: quindi x_0 coincide con $P_K(x)$. Dato che lo stesso ragionamento si può fare per *qualunque* successione $\{\varepsilon_n\}$ contenuta in $]0,1]$ ed infinitesima, concludiamo che $y_{\varepsilon,x} \rightharpoonup P_K(x)$ per $\varepsilon \to 0$.

Adesso osserviamo che dalla disuguaglianza fondamentale segue, al limite per $\varepsilon \to 0$,

$$\limsup_{\varepsilon\to 0}\|y_{\varepsilon,x}\|_H^2 \le \langle x, P_K(x) - u\rangle_H + \langle P_K(x), u\rangle_H \qquad \forall u \in D,$$

e ancora una volta questa relazione si estende ad ogni $u \in K$. Scegliendo $u = P_K(x)$, si deduce

$$\limsup_{\varepsilon\to 0} \|y_{\varepsilon,x}\|_H^2 \le \|P_K(x)\|_H^2 \le \liminf_{\varepsilon\to 0} \|y_{\varepsilon,x}\|_H^2 \,,$$

e quindi $\|y_{\varepsilon,x}\|_H \to \|P_K(x)\|_H$ per $\varepsilon \to 0$. Essendo anche $y_{\varepsilon,x} \rightharpoonup P_K(x)$, per una nota proprietà della convergenza debole si conclude che $y_{\varepsilon,x} \to P_K(x)$ in H per $\varepsilon \to 0$.
Infine, essendo $\{y_{\varepsilon,x}\} \subseteq D$, si ha $P_K(x) \in \overline{D}$ per ogni $x \in H$, e quindi $K = R(P_K) \subseteq \overline{D}$. Ma per definizione di K si ha ovviamente $\overline{D} \subseteq K$; dunque $K = \overline{D}$ e pertanto $\overline{D}$, coincidendo col suo inviluppo convesso, è esso stesso convesso.
(v) Osserviamo prima di tutto che se $0 < \varepsilon' < \varepsilon$ risulta

$$f_{\varepsilon'}(x) = \inf_{y\in H} \left[\frac{1}{2\varepsilon'}\|x-y\|_H^2 + f(y)\right] \ge \inf_{y\in H} \left[\frac{1}{2\varepsilon}\|x-y\|_H^2 + f(y)\right] = f_\varepsilon(x);$$

dunque esiste il limite di $f_\varepsilon(x)$ per $\varepsilon \to 0$ e tenuto conto di (i) e (iii) si ha, essendo f semicontinua inferiormente,

$$f(P_{\overline{D}}(x)) \le \liminf_{\varepsilon\to 0} f(y_{\varepsilon,x}) \le \lim_{\varepsilon\to 0} f_\varepsilon(x) \le f(x) \qquad \forall x \in H.$$

Se $x \in \overline{D}$, questa relazione fornisce subito la tesi. Se invece $x \notin \overline{D}$, allora $d = \mathrm{dist}(x, \overline{D}) > 0$: quindi

$$\frac{1}{2\varepsilon}\|x - y_{\varepsilon,x}\|_H^2 \ge \frac{d^2}{2\varepsilon}\,,$$

e pertanto

$$f(x) \ge \lim_{\varepsilon\to 0} f_\varepsilon(x) \ge \liminf_{\varepsilon\to 0} \frac{d^2}{2\varepsilon} + f(P_{\overline{D}}(x)) = +\infty,$$

da cui, ugualmente, la tesi. □

Osservazione 5.3.4 Nelle ipotesi della proposizione 5.3.3, si vede facilmente che $\overline{D(f)}$ coincide con $\overline{D}$, ove $D = \{x \in H : \partial f(x) \ne \emptyset\}$. Infatti se $x \notin \overline{D}$ si è visto nella dimostrazione di (v) che $f(x) = +\infty$ e quindi $x \notin D(f)$; pertanto $D(f) \subseteq \overline{D}$ e dunque $\overline{D(f)} \subseteq \overline{D}$. D'altra parte, l'osservazione 5.1.2 ci dice che $D \subseteq D(f)$, il che implica $\overline{D} \subseteq \overline{D(f)}$.

Esercizi 5.3

1. Sia H uno spazio di Hilbert reale e sia $f : H \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria. Si provi che esiste sempre almeno un punto x_0 tale che $\partial f(x_0) \neq \emptyset$.

2. Sia H uno spazio di Hilbert reale e sia $f : H \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria. Detto $y_{\varepsilon,x}$ il punto di minimo della funzione $y \mapsto \frac{1}{2\varepsilon}\|x-y\|_H^2 + f(y)$, si provi la disuguaglianza
$$\|y_{\varepsilon,x} - y_{\varepsilon,x'}\|_H \leq 2\|x - x'\|_H \qquad \forall x, x' \in H.$$

3. Si calcoli esplicitamente la regolarizzata f_ε nei casi seguenti:
   (a) $H = \mathbb{R}^N$, $f(x) = |x|_N^2$;
   (b) $H = \mathbb{R}$, $f(x) = \text{dist}(x, [a, b])$;
   (c) H spazio di Hilbert, f funzione affine;
   (d) H spazio di Hilbert, $f = I_K$, K convesso chiuso di H;
   (e) H spazio di Hilbert, $f(x) = \text{dist}(x, K)$, K convesso chiuso di H.

Capitolo 6

Disequazioni variazionali

6.1 Problemi variazionali con vincoli convessi

Negli spazi di Banach riflessivi le funzioni convesse e semicontinue inferiormente hanno minimo sui chiusi limitati ed anche, sotto ipotesi convenienti, su quelli non limitati (teorema 2.3.3 e corollario 2.3.4). Una caratteristica tipica di questo genere di problemi (ossia dei problemi di minimizzazione di funzionali convessi su vincoli a loro volta convessi) è la possibilità di caratterizzare i punti di minimo come soluzioni di opportune *disequazioni variazionali*. Il prototipo di tali disequazioni è quello che contraddistingue la proiezione su un convesso chiuso K in uno spazio di Hilbert reale X: $u = P_K(x)$ se e solo se u è il punto di minimo in K della funzione $v \mapsto \|x - v\|_X^2$, e ciò accade se e solo se

$$u \in K, \qquad \langle u - x, v - u \rangle_X \geq 0 \quad \forall v \in K$$

(esercizio 2.3.6). Più in generale si ha il seguente risultato:

Teorema 6.1.1 *Sia X uno spazio normato reale, sia $K \subseteq X$ un convesso non vuoto, e sia $f : K \to \mathbb{R}$ convessa e G-differenziabile, con $f'_G : K \to X^*$ continua. Allora sono fatti equivalenti:*

(i) $u \in K$ *e* $f(u) = \min_{v \in K} f(v)$*;*

(ii) $u \in K$ *e* $f'(u)(v - u) \geq 0$ *per ogni* $v \in K$*;*

(iii) $u \in K$ *e* $f'(v)(v - u) \geq 0$ *per ogni* $v \in K$*.*

Si noti che f è di classe C^1 per il corollario 3.3.3, quindi la notazione f' anziché f'_G è giustificata.

Dimostrazione (i) $\Longrightarrow$ (ii) Per ogni $v \in K$ e $\lambda \in]0,1[$ si ha $f(u) \leq f(u + \lambda(v-u))$; quindi

$$\frac{1}{\lambda}[f(u+\lambda(v-u)) - f(u)] \geq 0 \qquad \forall \lambda \in]0,1[.$$

Per $\lambda \to 0^+$ si ottiene $f'(u)(v-u) \geq 0$ per ogni $v \in K$. Si osservi che in questa implicazione non si è usata l'ipotesi di convessità.

(ii) $\Longrightarrow$ (i) Poiché f è convessa, i rapporti incrementali di f sono crescenti (esercizio 2.1.3); quindi, se $u \in K$ verifica (ii) si ha

$$f(v) - f(u) \geq \frac{1}{\lambda}[f(u+\lambda(v-u)) - f(u)] \geq f'(u)(v-u) \geq 0 \qquad \forall v \in K.$$

(ii) $\Longrightarrow$ (iii) Per la proposizione 5.2.5 f' è un operatore monotono, cioè

$$[f'(v) - f'(u)](v-u) \geq 0 \qquad \forall u, v \in K;$$

quindi se $u \in K$ verifica (ii) si ottiene immediatamente che vale anche (iii).

(iii) $\Longrightarrow$ (ii) Se $w \in K$ e $\lambda \in]0,1[$, scegliamo in (iii) $v = (1-\lambda)u + \lambda w$: si trova

$$f'((1-\lambda)u + \lambda w)(\lambda(w-u)) \geq 0,$$

e dopo aver diviso per λ, per $\lambda \to 0$ si ottiene (ii) in virtù della continuità di f' (si noti che tale ipotesi viene sfruttata esclusivamente qui). □

Nei prossimi paragrafi vedremo risultati più generali ed incontreremo altri esempi di disequazioni variazionali.

Esercizi 6.1

1. Sia $K \subseteq \mathbb{R}^N$ un convesso chiuso non vuoto e sia A un aperto di $\mathbb{R}^N$ contenente K. Si provi che se $f : A \to \mathbb{R}$ è una funzione di classe C^1 che ha minimo in K nel punto x_0, allora

$$\langle \mathrm{grad} f(x_0), x - x_0 \rangle_N \geq 0 \qquad \forall x \in K;$$

si deduca che se $x_0 \in \overset{\circ}{K}$ allora (ovviamente) $\mathrm{grad} f(x_0) = 0$.

2. Siano $f_1, \dots, f_k$ funzioni convesse di classe C^1 definite su $\mathbb{R}^N$; posto

$$K = \{x \in \mathbb{R}^N : \ f_i(x) \le 0, \ i = 1, \dots, k\},$$

si supponga che $\overset{\circ}{K} \neq \emptyset$ e che $\mathrm{grad} f_i(x) \neq 0$ per ogni $x \in K$ e per ogni $i = 1, \dots, k$. Sia poi $F : K \to \mathbb{R}^N$ una funzione continua. Se in un punto $x_0 \in K$ risulta

$$\langle F(x_0), x - x_0 \rangle_N \ge 0 \qquad \forall x \in K,$$

si provi che esistono $\lambda_1, \dots, \lambda_k \ge 0$ tali che

$$F(x_0) + \sum_{i=1}^{k} \lambda_i \, \mathrm{grad} f_i(x_0) = 0; \qquad \lambda_i f_i(x_0) = 0, \quad i = 1, \dots, k$$

(i numeri λ_i si dicono *moltiplicatori di Lagrange*).
[**Traccia:** l'unico caso non banale è quando $x_0 \in \partial K$ e $F(x_0) \neq 0$. In questo caso si provi preliminarmente che il cono normale $N_K(x_0)$ (definito nell'esercizio 5.1.4) è dato da

$$N_K(x_0) = \left\{ \sum_{i \in I} \lambda_i \, \mathrm{grad} f_i(x_0), \ \lambda_i \ge 0 \right\},$$

ove $I = \{i \in \{1, \dots, k\} : \ f_i(x_0) = 0\}$ (a tale scopo si ragioni per assurdo utilizzando il teorema di Hahn-Banach). Ciò premesso, si osservi che l'iperpiano $\langle F(x_0), x - x_0 \rangle_N = 0$ è di appoggio per K in x_0 e se ne deduca il risultato.]

6.2 Forme bilineari

Consideriamo un caso speciale, ma importante, del teorema 6.1.1. Sia X uno spazio di Banach reale riflessivo e sia $a : X \times X \to \mathbb{R}$ una *forma bilineare*, ossia un'applicazione lineare separatamente nei suoi due argomenti. Supponiamo che la forma a sia *simmetrica* (ovvero verifichi $a(u,v) = a(v,u)$ per ogni $u, v \in X$), *continua* (cioè esista $M \ge 0$ tale che

$$|a(u,v)| \le M \|u\|_X \|v\|_X \qquad \forall u, v \in X)$$

e *coerciva* (ossia esista $\nu > 0$ tale che

$$a(u,u) \geq \nu \|u\|_X^2 \qquad \forall u \in X).$$

Fissato $\varphi \in X^*$, il funzionale

$$f(u) = a(u,u) - 2\varphi u, \qquad u \in X,$$

verifica le condizioni del teorema 6.1.1. Infatti f è F-differenziabile perché per ogni $u, v \in K$ si ha

$$\begin{aligned} f(v) - f(u) &= a(v-u,u) + a(u,v-u) + a(v-u,v-u) - 2\varphi(v-u) = \\ &= 2a(u,v-u) - 2\varphi(v-u) + a(v-u,v-u), \end{aligned}$$

cosicché

$$f'(u)w = 2a(u,w) - 2\varphi w, \quad w \in X, \qquad \forall u \in X.$$

Si noti che $f' : X \to X^*$ è continua, anzi lipschitziana:

$$|[f'(v) - f'(u)]w| = |2a(v-u,w)| \leq 2M\|v-u\|_X \|w\|_X \qquad \forall w \in X,$$

da cui $\|f'(v) - f'(u)\|_{X^*} \leq 2M\|u-v\|_X$. Poi, f è strettamente convessa: infatti si ha

$$a(v-u,v-u) \geq 0, \quad a(v-u,v-u) = 0 \iff v = u,$$

da cui $2a(u,v) \leq a(u,u) + a(v,v)$ con l'uguaglianza se e solo se $u = v$; ne segue, per ogni $\lambda \in]0,1[$,

$$\begin{aligned} &a((1-\lambda)u + \lambda v, (1-\lambda)u + \lambda v) = \\ &\quad = (1-\lambda)^2 a(u,u) + 2\lambda(1-\lambda)a(u,v) + \lambda^2 a(v,v) \leq \\ &\quad \leq (1-\lambda)^2 a(u,u) + \lambda(1-\lambda)[a(u,u) + a(v,v)] + \lambda^2 a(v,v) = \\ &\quad = (1-\lambda)a(u,u) + \lambda a(v,v), \end{aligned}$$

con l'uguaglianza se e solo se $u = v$. Pertanto $u \mapsto a(u,u)$ è strettamente convessa e dunque tale è anche f. Infine, $f(u)$ tende a $+\infty$ per $\|u\|_X \to \infty$, in quanto

$$f(u) \geq \nu\|u\|_X^2 - 2\|\varphi\|_{X^*}\|u\|_X = \|u\|_X \cdot [\nu\|u\|_X - 2\|\varphi\|_{X^*}].$$

Possiamo allora concludere che f ha minimo unico in ogni chiuso convesso non vuoto contenuto in X. Per il teorema 6.1.1, per ogni convesso chiuso non vuoto $K \subseteq X$ esiste un unico $u \in K$ che risolve la disequazione variazionale

$$a(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K,$$

od equivalentemente

$$a(v, v-u) \geq \varphi(v-u) \qquad \forall v \in K.$$

Si noti che se K è limitato, per avere la sola esistenza di punti di minimo basta che la forma $a(u, v)$ sia continua, simmetrica e non negativa.

Osservazione 6.2.1 L'equivalenza fra il problema di minimo e la risolubilità della disequazione variazionale nella situazione precedente è dovuta alla simmetria della forma bilineare: quando a non è simmetrica, un punto u è di minimo per $u \mapsto a(u, u) - 2\varphi u$ in K se e solo se verifica

$$a(u, v-u) + a(v-u, u) \geq 2\varphi(v-u) \qquad \forall v \in K,$$

e questa disequazione variazionale *non* è equivalente alla precedente.

È interessante osservare che, anche quando la forma bilineare non è simmetrica, la disequazione variazionale

$$a(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K,$$

è risolubile, come mostra il risultato che segue: naturalmente la soluzione di essa *non* sarà, in generale, punto di minimo per $u \mapsto a(u, u) - 2\varphi u$.

Proposizione 6.2.2 *Sia X uno spazio di Banach reale riflessivo e sia $a : X \times X \to \mathbb{R}$ una forma bilineare, continua e coerciva. Se $K \subseteq X$ è un convesso chiuso non vuoto, e se $\varphi \in X^*$, allora esiste un unico $u \in K$ tale che*

$$a(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K,$$

o equivalentemente

$$a(v, v-u) \geq \varphi(v-u) \qquad \forall v \in K.$$

Dimostrazione Per l'equivalenza fra le due formulazioni della disequazione variazionale, rimandiamo all'esercizio 6.2.1.
Poniamo

$$\alpha(u,v) = \frac{a(u,v) + a(v,u)}{2}\,, \qquad \beta(u,v) = \frac{a(u,v) - a(v,u)}{2}\;;$$

le forme bilineari α e β sono la *parte simmetrica* e la *parte antisimmetrica* di a. È chiaro che α è simmetrica, continua e coerciva con le stesse costanti di a, mentre β è continua (e $\beta(u,v) = -\beta(v,u)$ per ogni $u, v \in X$). Sia ν la costante di coercività di a ed α; sia poi

$$N = \sup_{u,v \in X \setminus \{0\}} \frac{|\beta(u,v)|}{\|u\|_X \|v\|_X}\,.$$

e definiamo per ogni $t \in [0,1]$

$$a_t(u,v) = \alpha(u,v) + t\,\beta(u,v);$$

anche la forma a_t è coerciva con la stessa costante ν. Dimostreremo fra poco il seguente

Lemma 6.2.3 *Nelle ipotesi precedenti, sia $\tau \in [0,1[$ tale che per ogni $\varphi \in X^*$ esista un unico $u \in K$ per cui*

$$a_\tau(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K.$$

Allora la stessa proprietà vale per ogni $t \in \left[\tau, \tau + \frac{\nu}{N}\right[$.

Da questo lemma segue subito la tesi della proposizione 6.2.2: infatti la disequazione variazionale

$$a_\tau(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K$$

è risolubile per $\tau = 0$ (perché la forma $a_0 = \alpha$ è simmetrica); quindi ripetute applicazioni del lemma ci danno la risolubilità per ogni $\tau \in \left[0, \frac{\nu}{N}\right[$, per ogni $\tau \in \left[0, \frac{2\nu}{N}\right[$, per ogni $\tau \in \left[0, \frac{3\nu}{N}\right[$, eccetera, finché dopo un numero finito di passi avremo la risolubilità per ogni $\tau \in [0,1]$. Dato che $a_1 = a$, avremo la tesi. $\square$

Dimostrazione del lemma 6.2.3 Fissato $t_0 \in \left[0, \frac{\nu}{N}\right[$, sia $t \in [\tau, \tau + t_0]$. Riscriviamo la disequazione variazionale associata a t nel modo seguente:

$$a_\tau(u, v-u) \geq -(t-\tau)\beta(u, v-u) + \varphi(v-u) \qquad \forall v \in K.$$

Per ogni $w \in X$ sia ϕ_w il funzionale definito da

$$\phi_w(v) = -(t-\tau)\beta(w,v) + \varphi v, \qquad v \in X;$$

allora $\phi_w \in X^*$ e

$$\|\phi_w\|_{X^*} \le (t-\tau)N\|w\|_X + \|\varphi\|_{X^*}\,.$$

Quindi, per ipotesi, esiste un unico $u \in K$ tale che

$$a_\tau(u, v-u) \ge \phi_w(v-u) \qquad \forall v \in K;$$

poiché u è funzione di w, scriviamo $u = \psi(w)$. Proviamo che $\psi : X \to K$ è una contrazione: se $w_1, w_2 \in X$ e $u_1 = \psi(w_1)$, $u_2 = \psi(w_2)$, si ha

$$\begin{aligned}
\nu\|u_2 - u_1\|_X^2 &\le a_\tau(u_2 - u_1, u_2 - u_1) = \\
&= a_\tau(u_2, u_2 - u_1) - a_\tau(u_1, u_2 - u_1) \le \\
&\le \phi_{w_2}(u_2 - u_1) - \phi_{w_1}(u_2 - u_1) = (\phi_{w_2} - \phi_{w_1})(u_2 - u_1) \le \\
&\le \|\phi_{w_2} - \phi_{w_1}\|_{X^*}\|u_2 - u_1\|_X\,.
\end{aligned}$$

Poiché $\phi_{w_2} - \phi_{w_1} = -(t-\tau)\beta(w_2 - w_1, \cdot)$, dalla relazione precedente segue

$$\begin{aligned}
\|u_2 - u_1\|_X &\le \frac{t-\tau}{\nu}\|\beta(w_2 - w_1, \cdot)\|_{X^*} \le \\
&\le \frac{(t-\tau)N}{\nu}\|w_2 - w_1\|_X \le \frac{t_0 N}{\nu}\|w_2 - w_1\|_X;
\end{aligned}$$

dato che $t_0 < \frac{\nu}{N}$, ψ è una contrazione.
Essendo X uno spazio di Banach, l'applicazione $\psi : X \to K$ ha un unico punto fisso $w \in K$: l'equazione $w = \psi(w)$ significa

$$a_\tau(w, v-w) \ge -(t-\tau)\beta(w, v-w) + \varphi(v-w) \qquad \forall v \in K,$$

cioè w risolve la disequazione variazionale relativa alla forma a_t. $\square$

Osservazione 6.2.4 Se X è uno spazio di Hilbert reale, ogni forma bilineare continua $a : X \times X \to \mathbb{R}$ individua univocamente, in virtù del teorema di Riesz-Fréchèt, un operatore $A \in \mathcal{L}(X)$ tale che

$$a(u,v) = \langle Au, v\rangle_X \qquad \forall u, v \in X.$$

Le proprietà della forma a si traducono in corrispondenti proprietà dell'operatore A: in effetti, a è simmetrica se e solo se A è autoaggiunto, a è non negativa se e solo se A è monotono, ed infine a è coerciva se e solo se A è anch'esso coercivo, nel senso che $\langle Au, u\rangle_X \geq \nu\|u\|_X^2$ per ogni $u \in X$. Si noti che ciò implica $\|Au\|_X \geq \nu\|u\|_X$ per ogni $u \in X$, ma che quest'ultima proprietà non implica la coercività dell'operatore A e della forma a (esercizio 6.2.2).

Corollario 6.2.5 (teorema di Lax-Milgram) *Sia X uno spazio di Hilbert reale e sia $A \in \mathcal{L}(X)$ un operatore coercivo, ossia tale che*

$$\langle Au, u\rangle_X \geq \nu\|u\|_X^2 \qquad \forall u \in X,$$

con $\nu > 0$. Allora per ogni $\varphi \in X$ esiste un unico $u \in X$ tale che $Au = \varphi$.

Dimostrazione La forma bilineare $a(u,v) = \langle Au, v\rangle_X$ verifica tutte le ipotesi della proposizione 6.2.2 sul convesso $K = X$. Quindi esiste un unico $u \in X$ tale che

$$\langle Au, v-u\rangle_X \geq \langle \varphi, v-u\rangle_X \qquad \forall v \in X,$$

e scegliendo $v = u \pm w$, con $w \in X$, si trova

$$\pm\langle Au, w\rangle_X \geq \langle \varphi, w\rangle_X \qquad \forall w \in X,$$

ossia

$$\langle Au, w\rangle_X = \langle \varphi, w\rangle_X \qquad \forall w \in X.$$

Ciò prova che $Au = \varphi$. $\square$

Esercizi 6.2

1. Sia X uno spazio di Banach reale riflessivo e sia $a : X \times X \to \mathbb{R}$ una forma bilineare, continua e coerciva. Se $K \subseteq X$ è un convesso chiuso non vuoto, e se $\varphi \in X^*$, si provi che un elemento $u \in K$ risolve la disequazione variazionale

$$a(u, v-u) \geq \varphi(v-u) \qquad \forall v \in K,$$

se e solo se risolve quest'altra:

$$a(v, v-u) \geq \varphi(v-u) \qquad \forall v \in K.$$

2. Sia H uno spazio di Hilbert reale e sia $a : H \times H \to \mathbb{R}$ una forma bilineare continua e simmetrica. Si dia un esempio in cui l'operatore A ad essa associato verifica $\|Au\|_X \geq \nu \|u\|_X$ per ogni $u \in X$ senza che a sia coerciva.

3. Stabilire per quali $p, q, r, s \in \mathbb{R}$ la forma quadratica

$$a(u, v) = \int_0^1 [pu'v' + qu'v + ruv' + suv] \, dx$$

è coerciva sullo spazio di Hilbert $H = \{u \in AC[0, 1] : \ u' \in L^2(0, 1)\}$, rispetto alla norma naturale $\|u\|_H = \sqrt{\|u\|^2_{L^2(0,1)} + \|u'\|^2_{L^2(0,1)}}$.

6.3 Operatori monotoni

L'osservazione 6.2.4 lascia intuire che per una forma bilineare la proprietà di monotonia dell'operatore associato è un buon surrogato della coercività della forma, con l'ulteriore vantaggio, come già sappiamo, che la nozione di operatore monotono ha senso anche nel caso di operatori non lineari. In effetti vi è una ricca teoria che riguarda le disequazioni variazionali in cui compaiono operatori monotoni, e che trova applicazione in problemi non lineari con vincoli convessi. Uno dei più semplici risultati in questo ambito è la seguente facile generalizzazione dell'enunciato del teorema 6.1.1:

Teorema 6.3.1 *Sia X uno spazio di Banach reale riflessivo, sia $K \subseteq X$ un convesso chiuso non vuoto e sia $f : K \to]-\infty, +\infty]$ una funzione della forma $f = f_1 + f_2$, con f_1 e f_2 funzioni convesse e semicontinue inferiormente; supponiamo inoltre che f_1 sia G-differenziabile, con $f_1' : K \to X^*$ continua, e che f_2 sia propria. Allora sono fatti equivalenti:*

(i) $u \in K$ *e* $f(u) = \min_{v \in K} f(v)$*;*

(ii) $u \in K$ *e* $f_1'(u)(v - u) \geq f_2(u) - f_2(v)$ *per ogni* $v \in K$*;*

(iii) $u \in K$ *e* $f_1'(v)(v - u) \geq f_2(u) - f_2(v)$ *per ogni* $v \in K$*.*

Dimostrazione Si tratta di una sostanziale ripetizione della dimostrazione del teorema 6.1.1.

(i) $\Longrightarrow$ **(ii)** Poiché f_1 è a valori in $\mathbb{R}$ (essendo G-differenziabile) e f_2 è propria, si ha $f(u) \in \mathbb{R}$. Poiché $f(u) \leq f((1-\lambda)u + \lambda v)$ per ogni $\lambda \in]0,1]$ e per ogni $v \in K$, dalla convessità di f_2 segue

$$\begin{aligned} f(u) = f_1(u) + f_2(u) &\leq f_1((1-\lambda)u + \lambda v) + f_2((1-\lambda)u + \lambda v) \leq \\ &\leq f_1((1-\lambda)u + \lambda v) + (1-\lambda) f_2(u) + \lambda f_2(v). \end{aligned}$$

Mettendo a sinistra i termini contenenti f_1 e a destra quelli contenenti f_2, e dividendo per λ, si trova

$$\frac{1}{\lambda}[f_1((1-\lambda)u + \lambda v) - f_1(u)] \geq f_2(u) - f_2(v);$$

per $\lambda \to 0^+$ si ottiene (ii).
(ii) $\Longrightarrow$ (i) Per la crescenza dei rapporti incrementali di f_1 e per ipotesi,

$$f_1(v) - f_1(u) \geq f_1'(u)(v-u) \geq f_2(u) - f_2(v) \qquad \forall v \in K,$$

cioè $f(v) \geq f(u)$ per ogni $v \in K$.
(ii) $\Longrightarrow$ (iii) Facile conseguenza della monotonia di f_1'.
(iii) $\Longrightarrow$ (ii) Se $v = (1-\lambda)u + \lambda w$, con $\lambda \in]0,1]$ e $w \in K$, si ha da (iii)

$$\lambda f_1'((1-\lambda)u + \lambda w)(w-u) \geq f_2(u) - f_2((1-\lambda)u + \lambda w),$$

e per la convessità di f_2

$$\lambda f_1'((1-\lambda)u+\lambda w)(w-u) \geq f_2(u)-(1-\lambda)f_2(u)-\lambda f_2(w) = \lambda[f_2(u)-f_2(w)];$$

dividendo per λ ed utilizzando la continuità di f_1', per $\lambda \to 0^+$ si ottiene (ii).
$\square$

Esempio 6.3.2 Sia X uno spazio di Hilbert reale. Poniamo, per un fissato $x \in X$,

$$f_1(u) = \frac{1}{2}\|u - x\|_X^2\,, \quad f_2(u) = \varphi(u) \qquad \forall u \in X,$$

ove $\varphi : X \to]-\infty, +\infty]$ è una funzione convessa, semicontinua inferiormente e propria. Osserviamo che $f = f_1 + f_2$ tende a $+\infty$ per $\|u\|_X \to \infty$: infatti, φ ha una minorante affine (proposizione 2.3.7), cioè esistono $y \in X$ e $\alpha \in \mathbb{R}$ per cui $\varphi(u) \geq (y, u)_X + \alpha$ per ogni $u \in X$; ne segue

$$f(u) \geq \frac{1}{2}\|u-x\|_X^2 - \|y\|_X\|u\|_X - |\alpha| \to +\infty \quad \text{per } \|u\|_X \to \infty.$$

Per il corollario 2.3.4, f ha un unico punto di minimo su X, essendo strettamente convessa. In particolare, per il teorema 6.3.1, esiste un unico $u \in K$ tale che

$$\langle u-x, v-u\rangle_X \geq \varphi(u) - \varphi(v) \qquad \forall v \in X.$$

Tale u, evidentemente, dipende dal prefissato x; scriviamo dunque $u = p_\varphi(x)$. L'applicazione $p_\varphi : X \to D(\varphi)$ si chiama *applicazione di prossimità* (relativa alla φ prescelta).
Se, ad esempio, K è un convesso chiuso non vuoto di X, e scegliamo

$$\varphi(u) = I_K(u) = \begin{cases} 0 & \text{se } u \in K \\ +\infty & \text{se } u \notin K, \end{cases}$$

allora non è difficile vedere che $p_\varphi = P_K$ (la proiezione sul convesso K).
Osserviamo che p_φ è continua, anzi lipschitziana: infatti se $x_1, x_2 \in X$ e $u_1 = p_\varphi(x_1)$, $u_2 = p_\varphi(x_2)$, sostituendo u_1 e u_2 nella disequazione variazionale risolta rispettivamente da u_2 e u_1, si ottiene

$$\langle u_1 - x_1, u_2 - u_1\rangle_X \geq \varphi(u_1) - \varphi(u_2),$$

$$\langle u_2 - x_2, u_1 - u_2\rangle_X \geq \varphi(u_2) - \varphi(u_1);$$

sommando le due relazioni deduciamo

$$\langle x_1 - x_2, u_1 - u_2\rangle_X - \|u_1 - u_2\|_X^2 \geq 0,$$

da cui

$$\|u_1 - u_2\|_X^2 \leq \langle x_1 - x_2, u_1 - u_2\rangle_X \leq \|x_1 - x_2\|_X \|u_1 - u_2\|_X\,,$$

e quindi

$$\|u_1 - u_2\|_X = \|p_\varphi(x_1) - p_\varphi(x_2)\|_X \leq \|x_1 - x_2\|_X\,.$$

Il più importante risultato nella teoria delle disequazioni variazionali è il teorema che segue.

Teorema 6.3.3 (di Lions-Stampacchia) *Sia X uno spazio di Banach reale riflessivo, sia $g : X \to\,]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria, e sia $A : X \to X^*$ un operatore (non necessariamente lineare) tale che:*

(i) A *è* demicontinuo *(cioè continuo rispetto alla topologia debole* di* X^* *sui sottospazi di* X *finito-dimensionali);*

(ii) *A è monotono;*

(iii) *esiste* $v_0 \in D(g)$ *per cui*

$$\lim_{\|v\|_X \to \infty} \frac{A(v)(v - v_0) + g(v)}{\|v\|_X} = +\infty.$$

Allora per ogni $\varphi \in X^*$ *esiste* $u \in D(g)$ *tale che*

$$A(u)(v - u) + g(v) - g(u) \geq \varphi(v - u) \qquad \forall v \in X.$$

Inoltre se A è strettamente monotono, ossia $[A(v) - A(u)](v - u) > 0$ *per ogni* $u, v \in X$ *con* $u \neq v$*, allora la soluzione della disequazione variazionale è unica.*

L'ipotesi (ii) significa che se $Y \subseteq X$ è un sottospazio finito-dimensionale, allora per ogni fissato $u \in X$ la funzione reale $y \mapsto A(y)u$ è continua.

Dimostrazione Proveremo la tesi in tre passi.
1º passo: $\dim(X) < \infty$ e $D(g)$ limitato in X.
Poiché X ha dimensione finita, $X \simeq X^* \simeq \mathbb{R}^N$ e la dualità fra X e X^* è espressa dal prodotto scalare di $\mathbb{R}^N$. Si noti che $A : \mathbb{R}^N \to \mathbb{R}^N$ è continua. Risulta

$$u \in D(g), \qquad A(u)(v - u) + g(v) - g(u) \geq \varphi(v - u) \quad \forall v \in X$$

se e solo se

$$u \in D(g), \qquad \langle u - (u + \varphi - A(u)), v - u\rangle_N + g(v) - g(u) \geq 0 \quad \forall v \in X,$$

ossia se e solo se

$$u = p_g(u + \varphi - A(u)),$$

ove p_g è l'applicazione di prossimità relativa a g, definita nell'esempio 6.3.2. Poiché $D(g)$ è convesso e limitato, $\overline{D(g)}$ è convesso e compatto. La restrizione a $\overline{D(g)}$ della funzione

$$r(u) = p_g(u + \varphi - A(u)), \qquad u \in X,$$

manda $\overline{D(g)}$ in sé, anzi addirittura in $D(g)$, ed è continua; per il teorema di Brouwer, esiste almeno un punto fisso $u = r(u) \in D(g)$. Tale punto fisso risolve, per costruzione, la disequazione variazionale.

2$^{\mathbf{o}}$ passo: $\dim(X) < \infty$.

Notiamo che, in particolare, $X^* \simeq X$. Per $r > 0$ definiamo $g_r : X \to]-\infty, +\infty]$ nel modo seguente:

$$g_r(u) = g(u) + I_{\{u \in X:\ \|u\|_X \le r\}}(u) = \begin{cases} g(u) & \text{se } \|u\|_X \le r \\ +\infty & \text{se } \|u\|_X > r. \end{cases}$$

La funzione g_r è convessa e semicontinua inferiormente; se $r > \|v_0\|_X$, essa è anche propria con dominio $D(g_r) \subseteq \{u \in X : \|u\|_X \le r\}$, dunque limitato. Quindi, per il 1$^{\text{o}}$ passo, esiste $u_r \in D(g_r)$ tale che

$$\langle A(u_r), v - u_r \rangle_X + g_r(v) - g_r(u_r) \ge \langle \varphi, v - u_r \rangle_X \qquad \forall v \in X.$$

Se $r > \|v_0\|_X$, scegliendo $v = v_0$ otteniamo $g_r(v_0) = g(v_0) < \infty$ e

$$\langle A(u_r), v_0 - u_r \rangle_X + g(v_0) - g(u_r) \ge \langle \varphi, v_0 - u_r \rangle_X ,$$

da cui

$$\frac{\langle A(u_r), u_r - v_0 \rangle_X + g(u_r)}{\|u_r\|_X} \le \frac{1}{\|u_r\|_X} [\langle \varphi, u_r - v_0 \rangle_X + g(v_0)].$$

Quindi se per una successione $r_k \to \infty$ si avesse $\|u_{r_k}\|_X \to \infty$, dall'ipotesi (iii) dedurremmo che il primo membro della disuguaglianza tenderebbe a $+\infty$ per $k \to \infty$, mentre il secondo sarebbe limitato. Ciò è assurdo e pertanto esiste $K > 0$ tale che

$$\sup_{r>0} \|u_r\|_X = K < \infty.$$

Poiché X ha dimensione finita, deduciamo che esiste una successione $\{u_{r_n}\}$ estratta dalla famiglia $\{u_r\}_{r>0}$ tale che $u_{r_n} \to u$ per $n \to \infty$. Dall'ipotesi (i) segue che $A(u_{r_n}) \to A(u)$ in $X^* \simeq X$. Passando al limite per $n \to \infty$ nella relazione

$$\langle A(u_{r_n}), v - u_{r_n} \rangle_X + g_{r_n}(v) - g_{r_n}(u_{r_n}) \ge \langle \varphi, v - u_{r_n} \rangle_X \qquad \forall v \in X,$$

e tenendo conto che definitivamente risulta $g_{r_n}(u_{r_n}) = g(u_{r_n})$ e $g_{r_n}(v) = g(v)$ per ogni $v \in X$, si ottiene

$$\langle A(u), v - u \rangle_X + g(v) + \limsup_{n\to\infty}(-g(u_{r_n})) \ge \langle \varphi, v - u \rangle_X \qquad \forall v \in X,$$

da cui, per la semicontinuità inferiore di g,

$$\langle A(u), v - u\rangle_X + g(v) - g(u) \geq \langle \varphi, v - u\rangle_X \qquad \forall v \in X,$$

da cui la tesi del 2^{O} passo.

3^{O} passo: caso generale.

Sia $\mathcal{V}$ la famiglia dei sottospazi $M \subseteq X$ finito-dimensionali e contenenti v_0. Per il 2^{O} passo, per ogni $M \in \mathcal{V}$ esiste $u_M \in D(g) \cap M$ tale che

$$A(u_M)(v - u_M) + g(v) - g(u_M) \geq \varphi(v - u_M) \qquad \forall v \in M.$$

Scelto $v = v_0 \in M$, ripetendo l'argomentazione svolta nel 2^{O} passo si deduce

$$\sup_{M \in \mathcal{V}} \|u_M\|_X = K < \infty.$$

A questo punto ci serve un lemma per mettere la disequazione variazionale in forma più comoda.

Lemma 6.3.4 *Nelle ipotesi precedenti, se M è un sottospazio di X, per un elemento $u \in M \cap D(g)$ i seguenti fatti sono equivalenti:*

(i) $A(u)(v - u) + g(v) - g(u) \geq \varphi(v - u)$ *per ogni* $v \in M$*;*

(ii) $A(v)(v - u) + g(v) - g(u) \geq \varphi(v - u)$ *per ogni* $v \in M$*.*

Dimostrazione Si fa esattamente come nella dimostrazione (ii)$\Longleftrightarrow$ (iii) dei teoremi 6.1.1 e 6.3.1; nella parte ($\Longrightarrow$) si usa l'ipotesi di monotonia di A, mentre nella parte ($\Longleftarrow$) si utilizza la demicontinuità di A. $\square$

Continuiamo la dimostrazione del teorema di Lions-Stampacchia. Sia $v \in X$ e poniamo

$$S(v) = \{u \in D(g) : \ \|u\|_X \leq K, \ A(v)(v - u) + g(v) - g(u) \geq \varphi(v - u)\}.$$

L'insieme $S(v)$ è un convesso chiuso (eventualmente vuoto) di X, in quanto g è convessa e semicontinua inferiormente. Quindi $S(v)$ è debolmente chiuso. Essendo anche limitato, dalla riflessività di X segue che $S(v)$ è debolmente compatto. Se dimostriamo che

$$\bigcap_{v \in X} S(v) \neq \emptyset,$$

allora esiste $u \in X$ tale che $\|u\|_X \le K$ e

$$A(v)(v-u) + g(v) - g(u) \ge \varphi(v-u) \qquad \forall v \in X;$$

dal lemma 6.3.4 seguirà che u risolve la disequazione variazionale e ciò proverà l'enunciato principale del teorema.
Per mostrare che gli $S(v)$ hanno intersezione non vuota, trattandosi di insiemi w-compatti è sufficiente far vedere che vale la proprietà dell'intersezione finita, ossia che per ogni $N \in \mathbb{N}^+$ e per ogni $v_1, \ldots, v_N \in X$ si ha $\bigcap_{i=1}^N S(v_i) \ne \emptyset$. Ma, fissati $v_1, \ldots, v_N \in X$ e posto $M_N = [\{v_0, v_1, \ldots, v_N\}]$, si ha $M_N \in \mathcal{V}$ e quindi, per il 2^{O} passo, esiste $u_N \in M_N \cap D(g)$ tale che

$$A(u_N)(v-u_N) + g(v) - g(u_N) \ge \varphi(v-u_N) \qquad \forall v \in M_N,$$

ossia, per il lemma 6.3.4,

$$A(v)(v-u_N) + g(v) - g(u_N) \ge \varphi(v-u_N) \qquad \forall v \in M_N.$$

Se in questa relazione sostituiamo $v = v_i$, $i = 1, \ldots N$, otteniamo che $u_N \in \bigcap_{i=1}^N S(v_i)$. Ciò, come si è osservato, prova la tesi del 3^{O} passo.
Resta da provare l'unicità della soluzione quando A è strettamente monotono. Se u_1, u_2 sono due soluzioni della disequazione variazionale, possiamo sostituire ciascuno dei due nella disequazione verificata dall'altro: sommando le due relazioni si ottiene

$$[A(u_1) - A(u_2)](u_2 - u_1) \ge 0.$$

D'altra parte, se $u_1 \ne u_2$ deve aversi $[A(u_1) - A(u_2)](u_2 - u_1) < 0$: ne segue $u_1 = u_2$. $\square$

Vediamo ora alcuni casi particolari del teorema di Lions-Stampacchia.

Corollario 6.3.5 *Sia X uno spazio di Banach reale riflessivo e sia $K \subseteq X$ un convesso chiuso contenente l'origine; sia inoltre $A : X \to X^*$ un operatore demicontinuo, monotono e tale che*

$$\lim_{\|v\|_X \to \infty} \frac{A(v)v}{\|v\|_X} = +\infty.$$

Allora per ogni $\varphi \in X^*$ esiste $u \in K$ tale che

$$A(u)(v-u) \ge \varphi(v-u) \qquad \forall v \in K.$$

Dimostrazione Basta porre $g = I_K$ (la funzione indicatrice di K) e $v_0 = 0$, ed applicare il teorema 6.3.3. $\square$

Corollario 6.3.6 *Sia X uno spazio di Banach reale riflessivo e sia $A : X \to X^*$ un operatore demicontinuo, monotono e tale che*

$$\lim_{\|v\|_X \to \infty} \frac{A(v)v}{\|v\|_X} = +\infty.$$

Allora A è surgettivo.

Dimostrazione Fissato $\varphi \in X^*$, si applichi il corollario precedente con $K = X$: si ottiene l'esistenza di un $u \in X$ tale che

$$A(u)(v-u) \geq \varphi(v-u) \qquad \forall v \in X.$$

Scelto $v = u \pm w$ con $w \in X$, si deduce

$$(A(u) - \varphi)w = 0 \qquad \forall w \in X,$$

cioè $A(u) = \varphi$. $\square$

Osservazione 6.3.7 Se g è convessa, semicontinua inferiormente e propria, un elemento $u \in X$ risolve la disequazione variazionale

$$A(u)(v-u) + g(v) - g(u) \geq \varphi(v-u) \qquad \forall v \in X$$

se e solo se, per definizione di sottodifferenziale, $\varphi - A(u) \in \partial g(u)$. In particolare, u risolve

$$u \in K, \qquad A(u)(v-u) \geq \varphi(v-u) \quad \forall v \in K$$

se e solo se $\varphi - A(u) \in \partial I_K(u)$.

Esercizi 6.3

1. Si verifichi che ogni applicazione di prossimità è un operatore monotono.

2. Sia X uno spazio di Banach reale riflessivo, sia $g : X \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria, e sia K un convesso chiuso di X tale che $K \cap D(g) \neq \emptyset$. Sia inoltre $A : X \to X^*$ un operatore demicontinuo, monotono e tale che esista $v_0 \in D(g) \cap K$ per cui
$$\lim_{\|v\|_X \to \infty} \frac{A(v)(v - v_0) + g(v)}{\|v\|_X} = +\infty.$$
Si provi che per ogni $\varphi \in X^*$ esiste almeno un $u \in K$ tale che
$$A(u)(v - u) + g(v) - g(u) \geq \varphi(v - u) \qquad \forall v \in K.$$

6.4 Filo elastico teso sopra un ostacolo

Come applicazione della teoria delle disequazioni variazionali, consideriamo un filo elastico teso fra due estremi di ascisse $-a$ e a, con $a > 0$ fissato, e vincolato a stare al di sopra di un ostacolo $\psi(x)$. La configurazione $u(x)$ assunta dal filo sarà tale da minimizzare l'energia in gioco, che, in assenza di forze esterne, è descritta dal funzionale
$$J(u) = \frac{1}{2} \int_{-a}^{a} [p(x)u'(x)^2 + q(x)u(x)^2]dx,$$
dove $p(x)$ è la densità del filo nel punto di ascissa x e $q(x)$ è il coefficiente di elasticità del filo nello stesso punto; il primo termine nell'integrale è l'energia potenziale dovuta alla sua massa, mentre il secondo è l'energia di deformazione elastica immagazzinata dal filo.
Supporremo $p, q, \psi \in C[-a, a]$ con $p > 0$ e $q \geq 0$ in $[-a, a]$; affinché l'ostacolo non interferisca col filo nei punti di fissaggio, supporremo anche $\psi(-a) < 0$ e $\psi(a) < 0$. Il problema è quello di trovare una funzione u che minimizzi il funzionale J nell'insieme delle funzioni per le quali esso ha senso, e che sono in ogni punto non inferiori a $\psi(x)$.
Come spazio ambiente è naturale scegliere
$$X = \{u \in AC[-a, a] : \ u' \in L^2(-a, a), \ u(-a) = u(a) = 0\},$$
che è uno spazio di Hilbert con il prodotto scalare
$$\langle u, v \rangle_X = \int_{-a}^{a} [u(x)v(x) + u'(x)v'(x)]dx \qquad \forall u, v \in X.$$

La classe delle funzioni ammissibili è descritta dall'insieme

$$K = \{u \in X : \ u(x) \geq \psi(x) \ \forall x \in [-a, a]\}.$$

Per riuscire a minimizzare J bisognerà verificare che K è un convesso chiuso e che J verifica le ipotesi del teorema 2.3.3 o del corollario 2.3.4.
Proviamo anzitutto il seguente

Lemma 6.4.1 *Nelle ipotesi precedenti, risulta*

$$\|u\|_{L^2(-a,a)} \leq \sqrt{2a}\, \|u\|_{C[-a,a]} \leq 2a\, \|u'\|_{L^2(-a,a)} \leq 2a\, \|u\|_X \quad \forall u \in X.$$

Dimostrazione Se $u \in X$ si ha

$$\sqrt{\int_{-a}^{a} |u|^2\, dx} \leq \sqrt{2a}\|u\|_{C[-a,a]} = \sqrt{2a} \sup_{x\in[-a,a]} |u(x) - u(-a)| =$$
$$= \sqrt{2a} \sup_{x\in[-a,a]} \left| \int_{-a}^{x} u'(t)\, dt \right| \leq 2a\|u'\|_{L^2(-a,a)} \leq 2a\|u\|_X . \quad \square$$

È facile vedere che K è un convesso. Esso è ovviamente chiuso in $C[-a, a]$ (con la norma uniforme) e quindi è anche chiuso in X in virtù della stima fornita dal lemma precedente. Grazie alla limitatezza delle funzioni p e q, è chiaro che J è continuo rispetto alla norma di X; inoltre risulta

$$\lim_{\|u\|_X \to \infty} J(u) = +\infty :$$

per provare ciò, notiamo che

$$J(u) \geq \min_{[-a,a]} p \cdot \|u'\|^2_{L^2(-a,a)} ;$$

d'altronde se $\|u\|_X \to \infty$ allora si ha anche $\|u'\|_{L^2(-a,a)} \to \infty$, poiché altrimenti otterremmo, dal lemma 6.4.1, che nemmeno $\|u\|_{L^2(-a,a)}$ potrebbe tendere all'infinito, contro l'ipotesi $\|u\|_X \to \infty$. Si verifica anche, facilmente, che il funzionale J è strettamente convesso.
Dal corollario 2.3.4 segue che J ha un unico punto di minimo u sul convesso K. Poiché J è differenziabile secondo Fréchèt, con

$$J'(u)h = \int_{-a}^{a} [p(x)u'(x)h'(x) + q(x)u(x)h(x)]dx \qquad \forall h \in X,$$

per il teorema 6.1.1 il punto di minimo u è l'unica soluzione della disequazione variazionale

$$u \in K, \qquad J'(u)(v-u) \geq 0 \quad \forall v \in K.$$

Vogliamo ora trovare le eventuali ulteriori proprietà di regolarità del punto di minimo vincolato u. A questo scopo conviene anzitutto analizzare il comportamento dell'operatore $J' : X \to X^*$, che in questo caso è ovviamente lineare e che denoteremo con A.

Lemma 6.4.2 *Sia u il punto di minimo di J su K e sia $Au = J'(u)$, cosicché*

$$Au(v) = \int_{-a}^{a} (pu'v' + quv)\, dx \qquad \forall v \in X.$$

Allora il funzionale $Au \in X^*$ verifica le seguenti proprietà:

(i) *Au è positivo, ossia*

$$Au(g) \geq 0 \qquad \forall g \in X \text{ con } g \geq 0 \text{ in } [-a,a];$$

(ii) *Au è nullo sul complementare $E = [-a,a] \setminus D$ dell'insieme di contatto $D = \{x \in]-a,a[:\ u(x) = \psi(x)\}$, nel senso che*

$$Au(g) = 0 \qquad \forall g \in X \text{ con } g = 0 \text{ in } D.$$

Dimostrazione (i) Sia $g \in X$ una funzione non negativa. Allora, posto $v = u + g$, si ha $v \in K$; quindi, utilizzando la disequazione variazionale soddisfatta da u, si ricava

$$Au(g) = J'(u)g = J'(u)(v-u) \geq 0.$$

(ii) Sia $g \in X$ nulla in D: se $g \equiv 0$, ovviamente $Au(g) = 0$. Altrimenti, il supporto di g sarà un compatto contenuto in E, sul quale la funzione continua $u - \psi$ deve avere un minimo δ necessariamente positivo. Di conseguenza, si ha $v = u + tg \in K$ per $|t| < \frac{\delta}{\|g\|_\infty}$ e dunque, ancora dalla disequazione variazionale,

$$t(Au)(g) = Au(tg) = J'(u)(v-u) \geq 0 \quad \text{per } |t| < \frac{\delta}{\|g\|_\infty} :$$

dato che il segno di t è arbitrario, ciò significa $Au(g) = 0$. $\square$

La prossima tappa, che è il punto chiave di tutto il discorso, consiste nell'estendere il funzionale Au ad un elemento di $(C[-a,a])^*$. Preliminare a ciò è il seguente risultato:

Lemma 6.4.3 *Sia u il punto di minimo di J su K e sia Au il funzionale definito nel lemma precedente. Allora esiste $M \geq 0$ tale che*

$$|Au(g)| \leq M\|g\|_{L^\infty(D)} \qquad \forall g \in X,$$

ove $D = \{x \in [-a, a] : \ u(x) = \psi(x)\}$.

Dimostrazione Supponiamo per assurdo che la tesi non sia vera: allora per ogni $j \in \mathbb{N}^+$ esiste $g_j \in X$ tale che $\|g_j\|_{L^\infty(D)} \leq \frac{1}{j}$ e $|Au(g_j)| = 1$. Notando che D è un compatto contenuto in $]-a, a[$, possiamo fissare una funzione $\varphi \in C_0^1(\,]-a, a[\,)$ tale che $0 \leq \varphi \leq 1$ e $\varphi = 1$ su D. Per ogni $j \in \mathbb{N}^+$ sia poi $\delta_j > 0$ tale che

$$|g_j(x)| \leq \frac{2}{j} \qquad \text{per } d(x, D) = \inf_{y \in D} |x - y| < \delta_j \,,$$

e fissiamo per ogni $j \in \mathbb{N}^+$ una funzione $\varphi_j \in C_0^1(\,]-a, a[\,)$ tale che

$$0 \leq \varphi_j \leq \varphi, \quad \varphi_j = 1 \text{ su } D, \quad \varphi_j = 0 \text{ per } d(x, D) \geq \delta_j \,.$$

Allora $|g_j\varphi_j| \leq \frac{2}{j}\varphi$ in $[-a, a]$, cosicché, utilizzando le proprietà (i) e (ii) del lemma 6.4.2, risulta

$$\begin{aligned} 1 &= |Au(g_j)| = |Au(g_j\varphi_j) + Au(g_j(1 - \varphi_j))| = \\ &= |Au(g_j\varphi_j)| \leq Au(|g_j|\varphi_j) \leq \frac{2}{j}\, Au(\varphi). \end{aligned}$$

Per $j \to \infty$ si ottiene l'assurdo. $\square$

Corollario 6.4.4 *Sia u il punto di minimo di J su K e sia Au il funzionale definito nel lemma 6.4.2. Allora esiste un'unica misura finita μ, definita su una σ-algebra contenente i boreliani di $[-a, a]$, tale che*

$$Au(g) = \int_{-a}^{a} g \, d\mu \qquad \forall g \in X;$$

inoltre μ è concentrata sull'insieme di contatto D.

Dimostrazione Dal lemma 6.4.3 segue, in particolare,

$$|Au(g)| \leq M\|g\|_{C[-a,a]} \qquad \forall g \in X;$$

quindi utilizzando il teorema di Hahn-Banach il funzionale Au si estende a un funzionale lineare e continuo su $C[-a,a]$. Esiste dunque una misura di Lebesgue-Stieltjes μ, definita su una σ-algebra contenente i boreliani di $[-a,a]$ e associata a una funzione $m(x)$ crescente e continua a sinistra in $[-a,a]$, tale che
$$Au(g) = \int_{-a}^{a} g\, d\mu \qquad \forall g \in X.$$
Tale misura verifica però, sempre per il lemma 6.4.3,
$$\left| \int_{-a}^{a} g\, d\mu \right| \le M\|g\|_{L^\infty(D)} \qquad \forall g \in X,$$
e quindi, per densità, tale relazione vale per ogni $g \in C[-a,a]$ (si ricordi che $D \subset]-a,a[$). Ciò implica che μ è concentrata sull'insieme D.
Inoltre la misura μ è univocamente determinata: se μ_1 è un'altra misura con le stesse proprietà, si ha infatti
$$\int_D g\, d\mu = \int_{-a}^{a} g\, d\mu = Au(g) = \int_{-a}^{a} g\, d\mu_1 = \int_D g\, d\mu_1 \qquad \forall g \in X,$$
da cui, per densità,
$$\int_D g\, d\mu = \int_D g\, d\mu_1 \qquad \forall g \in C(D),$$
ossia $\mu = \mu_1$. $\square$

Analizziamo finalmente le proprietà di regolarità del punto di minimo u, sotto opportune ipotesi sull'ostacolo ψ.

Proposizione 6.4.5 *Sia u il punto di minimo di J su K e sia $E = \{x \in [-a,a] : u(x) > \psi(x)\}$; allora $pu' \in C^1(E)$ e $(pu')' = qu$ in E. Se inoltre l'ostacolo ψ ha le derivate destra e sinistra $\psi'(x^+)$ e $\psi'(x^-)$, e se risulta $\psi'(x^+) \ge \psi'(x^-)$ in ogni punto $x \in [-a,a]$, allora la funzione u è derivabile in ogni punto di $[-a,a]$.*

Questo enunciato ci dice che l'ipotesi chiave per avere la derivabilità di u in $[-a,a]$ è che i punti angolosi dell'ostacolo siano "rivolti verso il basso", ossia non siano punti di massimo locale: se ad esempio uno di quei punti fosse di massimo assoluto per ψ, esso, fatalmente, sarebbe un punto angoloso di u.

Dimostrazione Partiamo dalla relazione dimostrata nel corollario 6.4.4:
$$Au(g) = \int_{-a}^{a} g\,d\mu \qquad \forall g \in X,$$
e ricordiamo che la misura μ è associata ad una funzione $m : [-a, a] \to \mathbb{R}$ crescente, continua a sinistra e nulla per $x = -a$. Dal lemma 4.7.2 segue, integrando per parti,
$$Au(g) = \int_{-a}^{a} g\,d\mu = -\int_{-a}^{a} g' m\,dx \qquad \forall g \in X.$$
Dunque, ricordando la definizione del funzionale Au,
$$\int_{-a}^{a} [(pu' + m)g' + qug]\,d\mu = 0 \qquad \forall g \in X.$$
Applichiamo ora il lemma 4.2.2 alla funzione $pu' + m$, nella versione più generale dell'esercizio 6.4.2: si noti infatti che a priori $pu' + m$ è soltanto sommabile in $[-a, a]$, e non continua. Si ottiene che
$$[p(x)u'(x) + m(x)]' = q(x)u(x) \qquad \text{q.o. in } [-a, a],$$
e in particolare $pu' + m \in C^1[-a, a]$. Ma poiché μ è concentrata su $D = [-a, a] \setminus E$, in ogni componente connessa di E la funzione m è costante, cosicché la derivata m' esiste ed è nulla. Ne segue $pu' \in C^1(E)$ e $(pu')' = qu$ in E.
Osserviamo poi che si ha, per un'opportuna costante c,
$$u'(x) = -\frac{m(x)}{p(x)} + \frac{c}{p(x)} + \frac{1}{p(x)} \int_{-a}^{x} q(t)u(t)\,dt \qquad \forall x \in [-a, a].$$
e quindi u' ha limiti destro e sinistro in ogni punto, con
$$u'(x^+) - u'(x^-) = -\frac{m(x^+) - m(x^-)}{p(x)} \le 0 \qquad \forall x \in]-a, a[\,;$$
in particolare u' è continua nei punti di E.
D'altra parte, nei punti di D la funzione $u - \psi$ ha minimo uguale a 0, da cui
$$(u - \psi)'(x^+) \ge 0 \ge (u - \psi)'(x^-) \qquad \forall x \in D.$$
Dunque, per l'ipotesi fatta su ψ,
$$u'(x^+) \ge \psi'(x^+) \ge \psi'(x^-) \ge u'(x^-) \qquad \forall x \in D,$$
e pertanto
$$u'(x^+) = u'(x^-) \qquad \forall x \in D.$$
Ciò prova che u è derivabile anche nei punti di D. $\square$

Esercizi 6.4

1. Si provi che nello spazio $X = \{u \in AC[a,b] : \ u(a) = u(b) = 0\}$ la quantità $\|u'\|_{L^2(a,b)}$ è una norma equivalente alla norma naturale
$$\|u\|_X = \sqrt{\|u\|^2_{L^2(a,b)} + \|u'\|^2_{L^2(a,b)}}\,.$$

2. Si provino queste generalizzazioni dei lemmi 4.2.1, 4.2.2 e 4.2.3:

   (i) Sia $\alpha \in L^1(a,b)$ tale che
   $$\int_a^b \alpha(x)\varphi(x)\,dx = 0 \qquad \forall \varphi \in C^1_0[a,b];$$
   allora $\alpha = 0$ q.o. in $[a,b]$.

   (ii) Sia $\alpha \in L^1(a,b)$ tale che
   $$\int_a^b \alpha(x)\varphi'(x)\,dx = 0 \qquad \forall \varphi \in C^1_0[a,b];$$
   allora esiste $c \in \mathbb{R}$ tale che $\alpha = c$ q.o. in $[a,b]$.

   (iii) Siano $\alpha, \beta \in L^1(a,b)$ tali che
   $$\int_a^b [\alpha(x)\varphi(x) + \beta(x)\varphi'(x)]dx = 0 \qquad \forall \varphi \in C^1_0[a,b];$$
   allora $\beta \in AC[a,b]$, e $\beta' = \alpha$ q.o. in $[a,b]$.

3. Si supponga che l'ostacolo ψ sia una funzione assolutamente continua in $[-a,a]$; si provi che se il funzionale
$$A\psi(v) = \int_{-a}^a (p\psi'v' + q\psi v)\,dx, \qquad v \in X,$$
è positivo, allora l'insieme di contatto D è un intervallo.
[**Traccia:** Si ragioni per assurdo, supponendo che esistano $s,t \in D$ e $x \in [s,t]$ tali che $u(x) > \psi(x)$: posto $v(x) = \psi(x)$ per $x \in [s,t]$ e $v(x) = u(x)$ per $x \in [-a,a] \setminus [s,t]$, si verifichi che $v \in K$ e che $v \le u$; si deduca che $[A\psi - Au](v-u) \le 0$ e che di conseguenza $(\psi - u)' = 0$ in $[s,t]$. Di qui si arrivi all'assurdo.]

Capitolo 7

Multifunzioni

7.1 Nomenclatura

Le multifunzioni, o funzioni multivoche, sono applicazioni che ad ogni punto associano un insieme. Esse rivestono grande importanza in svariati problemi della matematica applicata, e comprendono entro il proprio ambito, generalizzandone la portata, la teoria delle disequazioni variazionali e dei problemi al contorno per equazioni differenziali.

Definizione 7.1.1 *Siano X, Y insiemi. Una* multifunzione *è un'applicazione $F : X \to \mathcal{P}(Y)$. L'insieme $F(x)$, per un fissato $x \in X$, è il* valore*, o* immagine*, di x. Gli insiemi*

$$D(F) = \{x \in X : \; F(x) \neq \emptyset\}, \qquad R(F) = F(X) = \bigcup_{x \in X} F(x)$$

sono rispettivamente il dominio *e l'* immagine *di F. Il* grafico *di F è l'insieme*

$$G(F) = \{(x, y) \in X \times Y : \; y \in F(x)\}.$$

Se A è un sottoinsieme di X, l' immagine *di A mediante F è l'insieme*

$$F(A) = \bigcup_{x \in A} F(x);$$

se B è un sottoinsieme di Y, la controimmagine*di B mediante F è l'insieme*

$$F^{-1}(B) = \{x \in X : F(x) \cap B \neq \emptyset\}.$$

Osservazione 7.1.2 Se K è un sottoinsieme di X e $F : K \to \mathcal{P}(Y)$ è una multifunzione con dominio $D(F) \subseteq K$, possiamo sempre estendere F ad una multifunzione $\overline{F} : X \to \mathcal{P}(Y)$, ponendo

$$\overline{F}(x) = \begin{cases} \emptyset & \text{se } \ x \notin K, \\ F(x) & \text{se } \ x \in K. \end{cases}$$

Si ottiene in questo modo una multifunzione definita su tutto X, tale che $D(\overline{F}) = D(F) \subseteq K$. Quindi nel seguito, quando ci farà comodo, potremo supporre di aver a che fare con multifunzioni definite su tutto X.

Osservazione 7.1.3 Ogni multifunzione $F : X \to \mathcal{P}(Y)$ è univocamente determinata dal suo grafico $G(F)$, nel senso che l'applicazione $F \mapsto G(F)$ è iniettiva ed anche surgettiva su $\mathcal{P}(X \times Y)$. Infatti l'iniettività è pressocché ovvia, mentre per la surgettività basta osservare che ogni $\Gamma \subseteq X \times Y$ è il grafico della multifunzione $J : X \to \mathcal{P}(Y)$ così definita:

$$J(x) = \{y \in Y : \ (x, y) \in \Gamma\}, \qquad x \in X.$$

Definizione 7.1.4 *Sia $F : X \to \mathcal{P}(Y)$ un'arbitraria multifunzione. La sua* inversa *è la multifunzione $F^{-1} : Y \to \mathcal{P}(X)$ definita da:*

$$F^{-1}(y) = \{x \in X : \ y \in F(x)\}, \qquad y \in Y.$$

È facile verificare che $D(F^{-1}) = R(F)$ e $R(F^{-1}) = D(F)$; inoltre per ogni $x \in X$ e per ogni $y \in Y$ si ha

$$(x, y) \in G(F) \qquad \Longleftrightarrow \qquad (y, x) \in G(F^{-1}).$$

Si noti che, coerentemente con la definizione di controimmagine mediante F, risulta

$$F^{-1}(B) = \{x \in X : F(x) \cap B \neq \emptyset\} = \bigcup_{y \in B} F^{-1}(y) \qquad \forall B \subseteq Y.$$

Osservazione 7.1.5 Ad ogni funzione $F : X \to Y$ corrisponde la multifunzione $\overline{F} : X \to \mathcal{P}(Y)$ definita da $\overline{F}(x) = \{F(x)\}$. In tal caso la multifunzione $\overline{F}^{-1}$ è semplicemente l'immagine inversa (in senso insiemistico) della funzione F, ossia

$$\overline{F}^{-1}(y) = F^{-1}(\{y\}) = \{x \in X : \ F(x) = y\} \qquad \forall y \in Y.$$

Esempi 7.1.6 **(1)** Se X è uno spazio normato e $f : X \to \overline{\mathbb{R}}$ è una funzione fissata, il sottodifferenziale di f è una multifunzione $\partial f : X \to \mathcal{P}(X^*)$, il cui dominio è l'insieme dei punti in cui f è sottodifferenziabile.
Si osservi che se X è riflessivo e f è convessa, semicontinua inferiormente e propria, allora dal corollario 5.1.11 segue che

$$\varphi \in \partial f(x) \quad \iff \quad x \in J^{-1} \circ \partial f^*(\varphi) \qquad \forall x \in X, \quad \forall \varphi \in X^*,$$

ove $J : X \to X^{**}$ è l'immersione canonica; ne segue che le due multifunzioni $\partial f : X \to \mathcal{P}(X^*)$ e $J^{-1} \circ \partial f^* : X^* \to \mathcal{P}(X)$ hanno dominio non vuoto e sono una l'inversa dell'altra.
(2) Sia X uno spazio di Banach; per ogni $x \in X$ poniamo

$$F(x) = \{\varphi \in X^* : \ \varphi x = \|x\|_X^2 = \|\varphi\|_{X^*}^2\}.$$

Il teorema di Hahn-Banach ci garantisce che $F(x) \neq \emptyset$ per ogni $x \in X$; quindi la multifunzione F ha per dominio tutto X. Essa si chiama *applicazione di dualità*.

Riprenderemo in esame questi esempi più avanti.

Andiamo ora a definire le nozioni di semicontinuità e di continuità per multifunzioni, che sono sostanzialmente differenti da quelle analoghe relative alle funzioni. La continuità per una multifunzione non ha una naturale definizione: se si prova ad estendere alle multifunzioni qualcuna delle svariate caratterizzazioni di questa proprietà valide per le funzioni, si ottengono condizioni non più equivalenti fra loro. Bisogna dunque fare una scelta fra le varie possibilità, e la bontà di tale scelta sarà giustificata dalla sua utilità nelle applicazioni.

Definizione 7.1.7 *Siano X, Y spazi topologici e sia $F : X \to \mathcal{P}(Y)$ una multifunzione. Diciamo che F è* semicontinua superiormente *nel punto $x_0 \in X$ se per ogni aperto $A \supseteq F(x_0)$ esiste un intorno U di x_0 in X tale che*

$$F(U) \subseteq A.$$

Diciamo che F è semicontinua superiormente in X se F è semicontinua superiormente in ogni punto di X.

Si noti che se F è una funzione questa definizione si riduce alla usuale nozione di continuità.

Definizione 7.1.8 *Siano X, Y spazi topologici e sia $F : X \to \mathcal{P}(Y)$ una multifunzione. Diciamo che F è* semicontinua inferiormente *nel punto $x_0 \in X$ se per ogni $y_0 \in F(x_0)$ e per ogni intorno N di y_0 in Y esiste un intorno U di x_0 in X tale che*

$$F(x) \cap N \neq \emptyset \qquad \forall x \in U.$$

Diciamo che F è semicontinua inferiormente in X se F è semicontinua inferiormente in ogni punto di X.

Anche questa definizione, allorché F è una funzione, si riduce a quella di continuità, in quanto la condizione $F(x) \cap N \neq \emptyset$ diventa $F(x) \in N$.

Definizione 7.1.9 *Siano X, Y spazi topologici e sia $F : X \to \mathcal{P}(Y)$ una multifunzione. Diciamo che F è* continua *nel punto $x_0 \in X$ se F è sia semicontinua superiormente che semicontinua inferiormente nel punto x_0. Diciamo che F è continua in X se F è continua in ogni punto di X.*

Esempio 7.1.10 Siano $F, G : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ definite da:

$$F(x) = \begin{cases} [-1,1] & \text{se } x = 0 \\ \{0\} & \text{se } x \neq 0, \end{cases} \qquad G(x) = \begin{cases} \{0\} & \text{se } x = 0 \\ [-1,1] & \text{se } x \neq 0. \end{cases}$$

Allora si verifica facilmente che F è semicontinua superiormente in $\mathbb{R}$, ma non è semicontinua inferiormente nel punto 0, mentre G è semicontinua inferiormente in $\mathbb{R}$ ma non è semicontinua superiormente nel punto 0.

Come vedremo, la semicontinuità superiore è una proprietà importante nello studio dei *punti fissi* delle multifunzioni, mentre la semicontinuità inferiore è utile per costruire *selezioni* di multifunzioni. Negli esercizi 7.1.1 e 7.1.2 sono esposte caratterizzazioni di questi due tipi di semicontinuità che, in qualche modo, motivano la scelta degli attributi "superiore" ed "inferiore" loro assegnati.

Esercizi 7.1

1. Siano X, Y spazi topologici. Si provi che una multifunzione $F : X \to \mathcal{P}(Y)$ è semicontinua superiormente se e solo se $F^{-1}(B)$ è chiuso in X per ogni chiuso $B \subseteq Y$.

2. Siano X, Y spazi topologici. Si provi che una multifunzione $F : X \to \mathcal{P}(Y)$ è semicontinua inferiormente se e solo se $F^{-1}(A)$ è aperto in X per ogni aperto $A \subseteq Y$.

3. Siano X, Y, Z spazi topologici e siano $F : X \to \mathcal{P}(Y)$ e $G : Y \to \mathcal{P}(Z)$ due multifunzioni. Definiamo la multifunzione composta $G \circ F$ così:

$$G \circ F(x) = \bigcup_{y \in F(x)} G(y), \qquad x \in X.$$

   Si provi che se F e G sono semicontinue superiormente allora $G \circ F$ è semicontinua superiormente, e che se F e G sono semicontinue inferiormente allora $G \circ F$ è semicontinua inferiormente.

4. Siano X, Y spazi metrici e sia $F : X \to \mathcal{P}(Y)$ una multifunzione. Si provi che:

   (i) se F è semicontinua superiormente e $F(x)$ è chiuso in Y per ogni $x \in X$, allora il grafico $G(F)$ è chiuso in $X \times Y$;

   (ii) se il grafico $G(F)$ è chiuso in $X \times Y$ ed Y è compatto, allora F è semicontinua superiormente;

   (iii) la multifunzione $F : \mathbb{R} \to \mathcal{P}(\mathbb{R}^2)$, definita da

$$F(\xi) = \{(x, y) \in \mathbb{R}^2 : \; y = \xi x\}, \qquad \xi \in \mathbb{R},$$

   ha grafico chiuso in $\mathbb{R}^3$, $F(\xi)$ è chiuso in $\mathbb{R}^2$ per ogni $\xi \in \mathbb{R}$, ma F non è semicontinua superiormente nel punto 0.

5. Si provi che la multifunzione $F : \mathbb{R} \to \mathcal{P}(\mathbb{R}^2)$ definita da

$$F(\xi) = \{(x, y) \in \mathbb{R}^2 : \; y = \xi\}, \qquad \xi \in \mathbb{R}$$

   è semicontinua inferiormente ma non semicontinua superiormente.

7.2 Punti fissi

L'esistenza di punti fissi per le multifunzioni è una proprietà di grande importanza nelle applicazioni, come cercheremo di mostrare nel seguito.

Definizione 7.2.1 *Sia X un insieme e sia $F : X \to \mathcal{P}(X)$ una multifunzione. Diciamo che un elemento $x \in X$ è un* punto fisso *per F se si ha $x \in F(x)$.*

Il primo risultato sui punti fissi di multifunzioni è una generalizzazione del classico teorema delle contrazioni di Banach. Introduciamo anzitutto una nozione di distanza fra insiemi.

Definizione 7.2.2 *Sia (X, d) uno spazio metrico. La quantità*

$$D(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \ \sup_{b \in B} d(b, A) \right\}, \qquad A, B \subseteq X,$$

si chiama metrica di Hausdorff.

Ricordiamo che $d(a, B) = \inf_{b \in B} d(a, b)$. È abbastanza facile verificare che la funzione D è una distanza sull'insieme $\mathcal{C}$ dei sottoinsiemi chiusi e limitati di X (esercizio 7.2.1).
Si ha il seguente risultato:

Teorema 7.2.3 *Sia (X, d) uno spazio metrico completo, sia M un sottoinsieme non vuoto e chiuso di X, e sia $F : X \to \mathcal{P}(M)$ una multifunzione con $D(F) \supseteq M$, tale che $F(x)$ sia chiuso per ogni $x \in M$. Se esiste $k \in]0, 1[$ tale che*

$$D(F(x), F(y)) \leq k \, d(x, y) \qquad \forall x, y \in M,$$

allora F ha almeno un punto fisso.

Si osservi che non possiamo aspettarci l'unicità del punto fisso: se ad esempio $F(x) = M \subseteq X$ per ogni $x \in X$, e M ha più di un elemento, le ipotesi del teorema sono banalmente verificate e ogni punto di M è punto fisso per F.

Dimostrazione Fissiamo $h \in]k, 1[$ e sia $x_0 \in D(F)$. Scegliamo $x_1 \in F(x_0)$ tale che $d(x_0, x_1) > 0$ (se tale x_1 non esiste, vuol dire che $F(x_0) = \{x_0\}$, cosicché in particolare x_0 è un punto fisso di F ed abbiamo finito). Essendo $x_1 \in F(x_0) \subseteq M \subseteq D(F)$, si ha $F(x_1) \neq \emptyset$. Allora

$$\begin{aligned} d(x_1, F(x_1)) &\leq \sup_{a \in F(x_0)} d(a, F(x_1)) \leq \\ &\leq D(F(x_0), F(x_1)) \leq k \ d(x_0, x_1) < h \ d(x_0, x_1); \end{aligned}$$

quindi, per definizione di $d(x_1, F(x_1))$, esiste $x_2 \in F(x_1)$ tale che

$$d(x_1, x_2) < h \ d(x_0, x_1).$$

Se $d(x_2, x_1) = 0$, allora $x_2 = x_1 \in F(x_1)$, e x_1 è un punto fisso di F; altrimenti, sarà $d(x_2, x_1) > 0$ e possiamo iterare questo procedimento. Se non si trova un punto fisso dopo un numero finito di passi, si costruirà una successione $\{x_n\}$ tale che

$$x_{n+1} \in F(x_n) \subseteq M, \quad d(x_n, x_{n+1}) < h^n\, d(x_0, x_1) \quad \forall n \in \mathbb{N}.$$

È facile riconoscere che $\{x_n\}$ è una successione di Cauchy contenuta nel chiuso M; per la completezza di X essa converge ad un elemento $x \in M$. Ne segue

$$d(x_{n+1}, F(x)) \leq D(F(x_n), F(x)) \leq k\ d(x_n, x) \to 0 \quad \text{per } n \to \infty,$$

da cui $x \in \overline{F(x)} = F(x)$, che è la tesi. $\square$

Osservazione 7.2.4 Non è difficile verificare che nelle ipotesi del teorema precedente la multifunzione F è semicontinua inferiormente, ma non è in generale semicontinua superiormente: lo è, ad esempio, nei punti x tali che $F(x)$ è compatto (si vedano gli esercizi 7.2.2 e 7.2.3).

Uno dei più importanti teoremi di punto fisso per le multifunzioni è quello che segue; da esso discendono come corollari molti dei classici risultati di punto fisso per funzioni.

Teorema 7.2.5 (di Kakutani) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff, sia K un sottoinsieme non vuoto, convesso e compatto di X, e sia $F : K \to \mathcal{P}(K)$ una multifunzione semicontinua superiormente e tale che $F(x)$ sia non vuoto, convesso e chiuso per ogni $x \in K$. Allora F ha almeno un punto fisso.*

Prima di dimostrare questo teorema vediamo alcune delle sue conseguenze.

Corollario 7.2.6 (teorema di Tikhonov) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff, sia K un sottoinsieme non vuoto, convesso e compatto di X, e sia $F : K \to K$ un'applicazione continua. Allora F ha almeno un punto fisso.*

Dimostrazione Poiché $F : K \to K$ è continua, la multifunzione $\overline{F}(x) = \{F(x)\}$ è continua e verifica le ipotesi del teorema 7.2.5. Quindi esiste $x \in K$ tale che $x \in \overline{F}(x)$, ossia $x = F(x)$. $\square$

Corollario 7.2.7 (teorema di Schauder) *Sia X uno spazio di Banach e sia M un sottoinsieme non vuoto, convesso, chiuso e limitato di X. Se $F : M \to M$ è un'applicazione continua e compatta (ossia trasforma insiemi limitati in insiemi relativamente compatti), allora F ha almeno un punto fisso.*

Dimostrazione Per ipotesi, l'insieme $F(M)$ è relativamente compatto: dunque, per il teorema di Mazur (teorema 1.7.19), l'insieme $K = \overline{\mathrm{co}(F(M))}$ è convesso e compatto; inoltre $K \subseteq M$ perché M è un convesso chiuso che contiene $F(M)$. Consideriamo la restrizione $F|_K$: questa applicazione verifica le ipotesi del teorema di Tikhonov (corollario 7.2.6). Quindi esiste $x \in K \subseteq M$ tale che $F(x) = x$. $\square$

Il teorema di Schauder può essere generalizzato alle multifunzioni come segue:

Corollario 7.2.8 *Sia X uno spazio di Banach e sia M un sottoinsieme non vuoto, convesso e chiuso di X. Sia $F : M \to \mathcal{P}(M)$ una multifunzione semicontinua superiormente, tale che $F(M)$ sia relativamente compatto e che $F(x)$ sia convesso, chiuso e non vuoto per ogni $x \in M$; allora F ha almeno un punto fisso.*

Dimostrazione Sia $K = \mathrm{co}(\overline{F(M)})$; K è un convesso compatto non vuoto contenuto in M. La restrizione $F|_K$ è a valori in $\mathcal{P}(K)$: infatti se $x \in K$ allora $F(x)$ è un convesso chiuso contenuto in $F(M)$, e quindi in K. Dunque, per il teorema 7.2.5, $F|_K$ ha un punto fisso; ne segue che anche F ha un punto fisso. $\square$

Dimostrazione del teorema 7.2.5 Proviamo anzitutto la seguente proposizione, che fornisce la tesi del teorema sotto differenti ipotesi e che ha interesse di per sé.

Proposizione 7.2.9 *Sia X uno spazio vettoriale topologico di Hausdorff, sia $K \subseteq X$ non vuoto, convesso e compatto, e sia $F : K \to \mathcal{P}(K)$ una multifunzione (non necessariamente semicontinua superiormente) tale che:*

(i) *$F(x)$ è non vuoto e convesso per ogni $x \in K$;*

(ii) *$F^{-1}(y)$ è aperto in K per ogni $y \in K$.*

Allora F ha almeno un punto fisso.

Si verifica facilmente che la condizione (ii) implica la semicontinuità inferiore di F: dunque siamo in ipotesi diverse da quelle del teorema di Kakutani.

Dimostrazione La famiglia di aperti $\{F^{-1}(y)\}_{y\in K}$ è un ricoprimento di K: infatti, se $x \in K$, scelto $y \in F(x)$, si ha $y \in K$ e, per definizione, $x \in F^{-1}(y)$. Dato che K è compatto, K è ricoperto da una unione finita $\bigcup_{i=1}^{n} F^{-1}(y_i)$, con $y_1, \ldots, y_n \in K$. Sia $\{f_i\}$ una partizione dell'unità associata a $\{F^{-1}(y_i)\}$, $i = 1, \ldots, n$ (essa esiste perché il compatto K, essendo di Hausdorff, è paracompatto e in particolare è normale, ossia tale che due chiusi disgiunti possiedono intorni disgiunti). Poniamo

$$p(x) = \sum_{i=1}^{n} f_i(x)y_i \qquad \forall x \in K_0 = \mathrm{co}(\{y_1, \ldots y_n\}) \subseteq K.$$

Il convesso K_0 è contenuto in uno spazio al più n-dimensionale, vale a dire quello generato dai vettori $y_1, \ldots, y_n$, ed è compatto. L'applicazione p è continua e $p(K_0) \subseteq K_0$ in quanto $\sum_{i=1}^{n} f_i(x) = 1$. Per il teorema di Brouwer p ha un punto fisso $\overline{x} \in K_0$. Basta allora dimostrare che $p(x) \in F(x)$ per ogni $x \in K_0$, per dedurre che $\overline{x}$ è un punto fisso di F in $K_0 \subseteq K$.
Sia $x \in K_0$: poiché $\sum_{i=1}^{n} f_i(x) = 1$ su K, esiste almeno un $j \in \{1, \ldots, n\}$ per il quale $f_j(x) \neq 0$. Per ciascun indice j tale che $f_j(x) \neq 0$ si ha $x \in F^{-1}(y_j)$ e quindi $y_j \in F(x)$. Ne segue, essendo $F(x)$ convesso,

$$p(x) = \sum_{i=1}^{n} f_i(x)y_i = \sum_{j:\ f_j(x)\neq 0} f_j(x)y_j \in F(x),$$

da cui la tesi. $\square$

La dimostrazione del teorema 7.2.5 prosegue introducendo una disequazione variazionale.

Proposizione 7.2.10 *Sia X uno spazio vettoriale topologico di Hausdorff, sia $K \subseteq X$ non vuoto, convesso e compatto, e sia $T : K \to X^*$ continua rispetto alla topologia forte di X^*. Allora la disequazione variazionale*

$$[T(x_0)](x_0 - x) \geq 0 \qquad \forall x \in K$$

ha soluzione $x_0 \in K$.

Dimostrazione Ragioniamo per assurdo: se così non fosse, per ogni $x_0 \in K$ troveremmo un elemento $x \in K$ tale che $[T(x_0)](x_0 - x) < 0$. Consideriamo la multifunzione $F : K \to \mathcal{P}(K)$ definita nel modo seguente:

$$F(x_0) = \{x \in K : \ [T(x_0)](x_0 - x) < 0\}, \qquad x_0 \in K.$$

Chiaramente $F(x_0)$ è non vuoto e convesso. Proviamo che $F^{-1}(x)$ è aperto in K per ogni $x \in K$: per definizione di controimmagine di F si ha

$$F^{-1}(x) = \{x_0 \in X : \ [T(x_0)](x_0 - x) < 0\}.$$

Sia $\{x_i\}_{i\in I}$ un net contenuto in $K \setminus F^{-1}(x)$ che converge ad un elemento $\overline{x} \in K$ in X; allora vale la relazione

$$[T(x_i)](x_i - x) \geq 0 \qquad \forall i \in I.$$

Quindi possiamo scrivere

$$\begin{aligned} &[T(\overline{x})](\overline{x} - x) = \\ &\quad = [T(\overline{x})](\overline{x} - x_i) + [T(\overline{x}) - T(x_i)](x_i - x) + [T(x_i)](x_i - x) \geq \\ &\quad \geq [T(\overline{x})](\overline{x} - x_i) + [T(\overline{x}) - T(x_i)](x_i - x) \qquad \forall i \in I. \end{aligned}$$

Passando al limite rispetto ad i, il primo termine all'ultimo membro tende a 0 perché $T(\overline{x}) \in X^*$. Notiamo poi che K, essendo compatto, è limitato in X (esercizio 1.4.9). Poiché (vedere il paragrafo 1.5) la topologia forte di X^* è generata dalle seminorme $p_A(\varphi) = \sup_{y\in A} |\varphi y|$, con A sottoinsieme limitato di X, dalla continuità di T rispetto a tale topologia e dalla limitatezza di $K - K$ si deduce, passando al limite rispetto ad i,

$$|[T(x_i) - T(\overline{x})](x_i - x)| \leq p_{K-K}(T(x_i) - T(\overline{x})) \to 0;$$

quindi si conclude che

$$[T(\overline{x})](\overline{x} - x) \geq 0,$$

ossia $\overline{x} \in K \setminus F^{-1}(x)$. Ciò prova che $F^{-1}(x)$ è aperto.
Sono allora verificate le ipotesi della proposizione 7.2.9, e pertanto esiste un punto fisso $x_0 \in K$ per F; quindi, per definizione di $F(x_0)$, otteniamo

$$0 = [T(x_0)](x_0 - x_0) < 0,$$

il che è assurdo. $\square$

Il teorema di Kakutani si deduce immediatamente dal risultato che segue:

Proposizione 7.2.11 (teorema di Browder) *Sia X uno spazio vettoriale topologico localmente convesso di Hausdorff, sia $K \subseteq X$ non vuoto, convesso e compatto; sia inoltre $F : K \to \mathcal{P}(X)$ una multifunzione semicontinua superiormente, tale che $F(x)$ sia chiuso, convesso e non vuoto per ogni $x \in K$. Supponiamo inoltre che si abbia*

(a) *per ogni $x \in \partial K$ esistono $y \in F(x)$, $u \in K$, $\lambda > 0$ tali che $y = x + \lambda(u - x)$,*

oppure

(b) *per ogni $x \in \partial K$ esistono $y \in F(x)$, $u \in K$, $\lambda < 0$ tali che $y = x + \lambda(u - x)$.*

Allora F ha almeno un punto fisso.

Il teorema 7.2.5 corrisponde all'ipotesi (a): infatti nel caso del teorema di Kakutani si ha $F(x) \subseteq K$ e quindi, fissato $y \in F(x)$, si può scegliere $u = y$ e $\lambda = 1$. Pertanto avremo provato il teorema 7.2.5 se proveremo il teorema di Browder.

Dimostrazione Supponiamo che valga la condizione (a) e ragioniamo per assurdo. Assumiamo che F sia priva di punti fissi: ciò significa che $x \notin F(x)$ per ogni $x \in K$.
L'insieme $\{x\}$ è convesso e compatto, mentre $F(x)$ è convesso e chiuso. Per il teorema di Hahn-Banach (teorema 1.7.5) esiste un funzionale $w(x) \in X^*$ tale che

$$w(x)x < \inf_{y \in F(x)} w(x)y \ ;$$

quindi, per continuità, esiste un intorno W_x di x tale che

$$\sup_{z \in W_x} w(x)z < \inf_{y \in F(x)} w(x)y \,.$$

Posto

$$U_x = \{y \in X : \ \sup_{z \in W_x} w(x)z < w(x)y\},$$

U_x è un aperto contenente $F(x)$. Adesso per ogni $x \in K$ definiamo

$$N_x = \{v \in K : \ w(x)(v - y) < 0 \ \forall y \in F(v)\},$$

e proviamo che N_x è un intorno di x. Dato che F è semicontinua superiormente, in corrispondenza dell'aperto $U_x \supseteq F(x)$ esiste un intorno V_x di

x, aperto in K e che non è restrittivo supporre contenuto in W_x, tale che $F(v) \subseteq U_x$ per ogni $v \in V_x$; ciò significa che

$$\sup_{z \in W_x} w(x)z < w(x)y \qquad \forall y \in F(v), \qquad \forall v \in V_x\,.$$

A maggior ragione, quindi,

$$w(x)(v-y) < 0 \qquad \forall y \in F(v), \qquad \forall v \in V_x\,,$$

e dunque $V_x \subseteq N_x$.
Poiché K è compatto, esistono $x_1, \ldots, x_n \in K$ tali che $K = \bigcup_{i=1}^n V_{x_i}$; scegliamo una partizione dell'unità $\{f_i\}$ associata a $\{V_{x_i}\}$ (nuovamente, ciò è possibile perché K è uno spazio topologico compatto e di Hausdorff, dunque paracompatto), e poniamo

$$r(x) = \sum_{i=1}^{n} f_i(x) w(x_i) \qquad \forall x \in K.$$

È facile verificare che la funzione $r : K \to X^*$ è continua rispetto alla topologia forte di X^*; proviamo che

$$r(x)(x-y) < 0 \qquad \forall y \in F(x).$$

Per ogni indice i tale che $f_i(x) > 0$ (e di tali indici ce n'è almeno uno) si ha $x \in V_{x_i} \subseteq N_{x_i}$; quindi, per definizione di N_{x_i},

$$w(x_i)(x-y) < 0 \qquad \forall y \in F(x),$$

ed a maggior ragione, sommando su i, si ottiene $r(x)(x-y) < 0$.
Adesso osserviamo che r e K verificano le ipotesi della proposizione 7.2.10: quindi esiste $x \in K$ tale che

$$r(x)(x-v) \geq 0 \qquad \forall v \in K.$$

Da qui dedurremo l'assurdo. Infatti, il punto x sarà interno a K oppure sulla frontiera di K. Se x è interno a K, allora per ogni fissato $z \in X$ esiste $\varepsilon > 0$ tale che $x \pm \varepsilon z \in K$; se nella disequazione variazionale prendiamo $v = x \pm \varepsilon z$, otteniamo $r(x)z = 0$, e per l'arbitrarietà di z si ricava $r(x) \equiv 0$, contro il fatto che $r(x)(x-y) < 0$ per $y \in F(x)$. Se invece $x \in \partial K$, entra in gioco la condizione (a): esistono $y \in F(x)$, $u \in K$ e $\lambda > 0$ tali che $y = x + \lambda(u-x)$.

Scelto $v = u$, la disequazione variazionale ci dice che $r(x)(x-u) \geq 0$, da cui $r(x)(x-y) \geq 0$; ma dato che $y \in F(x)$, deve essere invece $r(x)(x-y) < 0$. Ciò prova l'assurdo.
Se si assume la condizione (b) in luogo della (a), si sceglie con il teorema 1.7.5 un funzionale $w(x) \in X^*$ tale che

$$w(x)x > \sup_{y \in F(x)} w(x)y \,,$$

e si ripete tutta l'argomentazione con la disuguaglianza $>$ in luogo di $<$. $\square$

Esercizi7.2

1. Dimostrare che la metrica di Hausdorff definisce una distanza sulla classe dei sottoinsiemi chiusi e limitati di uno spazio metrico (X, d).

2. Sia (X, d) uno spazio metrico e sia $F : X \to \mathcal{P}(X)$ una multifunzione con le seguenti proprietà:

   (i) $F(x)$ è chiuso e non vuoto per ogni $x \in X$;

   (ii) esiste $k \in]0,1[$ per cui

   $$D(F(x), F(x')) \leq k\ d(x, x') \qquad \forall x, x' \in X,$$

   ove D è la metrica di Hausdorff. Si provi che F è semicontinua inferiormente in ogni punto, e che è semicontinua superiormente nei punti $x \in X$ tali che $F(x)$ è compatto.

3. Sia $F : \mathbb{R}^2 \to \mathcal{P}(\mathbb{R}^2)$ definita da

   $$F(x, y) = \left]-\infty, \frac{x}{2}\right] \times \left]-\infty, \frac{y}{2}\right] \qquad \forall (x, y) \in \mathbb{R}^2.$$

   Si provi che F non è semicontinua superiormente in alcun punto, benché risulti

   $$D(F(x, y), F(x', y')) \leq \frac{1}{2}\sqrt{(x-x')^2 + (y-y')^2} \;\; \forall (x, y), (x', y') \in \mathbb{R}^2.$$

7.3 Punti di sella

Vogliamo applicare la teoria dei punti fissi di multifunzioni all'esistenza di punti di sella per funzioni definite sul prodotto di spazi di Banach; il teorema che dimostreremo troverà a sua volta applicazione nella teoria dei giochi. In accordo con la definizione 2.6.15, poniamo:

Definizione 7.3.1 *Siano* A, B *insiemi non vuoti, e sia* $f : A \times B \to \mathbb{R}$ *una funzione. Diciamo che* $(x_0, y_0) \in A \times B$ *è* punto di sella *per la funzione* f *se*

$$f(x_0, y) \leq f(x_0, y_0) \leq f(x, y_0) \qquad \forall (x, y) \in A \times B,$$

oppure se

$$f(x_0, y) \geq f(x_0, y_0) \geq f(x, y_0) \qquad \forall (x, y) \in A \times B.$$

Dunque un punto $(x_0, y_0) \in A \times B$ è di sella per f se e solo se

$$\max_{y \in B} f(x_0, y) = f(x_0, y_0) = \min_{x \in A} f(x, y_0)$$

oppure

$$\min_{y \in B} f(x_0, y) = f(x_0, y_0) = \max_{x \in A} f(x, y_0).$$

Dato che i punti di sella per f sono di sella anche per $-f$, a meno di cambiare segno alla f si può sempre supporre, e così faremo nel seguito, che in un punto di sella (x_0, y_0) si abbia

$$f(x_0, y) \leq f(x_0, y_0) \leq f(x, y_0) \qquad \forall (x, y) \in A \times B.$$

Ad esempio, se $A = B = \mathbb{R}$ il punto $(0, 0)$ sarà di sella per la funzione $f(x, y) = x^2 - y^2$ (e non, in base alla nostra convenzione, per la sua opposta). Si noti che la nozione di punto di sella che ci interessa qui è molto più restrittiva di quella utilizzata nel calcolo differenziale: ad esempio, la funzione $g : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$g(x, y) = (y - x^2/2)(y - 2x^2)$$

ha nell'origine un punto stazionario che non è né di massimo, né di minimo relativo (essendo $g(x, x^2) = -x^4/2 < 0 = g(0, 0) < y^2 = g(0, y)$ per ogni $x, y \neq 0$), ma che non è di sella secondo la definizione 7.3.1 poiché

$$g(0, y) > g(0, 0) \quad \forall y \neq 0, \qquad g(x, 0) > g(0, 0) \quad \forall x \neq 0.$$

Inoltre considereremo solamente punti di sella *globali*: per quelli locali basterà restringere la funzione ad un opportuno intorno del punto (x_0, y_0).

Proposizione 7.3.2 *Siano A, B insiemi non vuoti, sia $f : A \times B \to \mathbb{R}$ una funzione. La f possiede un punto di sella in $A \times B$ se e solo se la funzione $x \mapsto \sup_{y\in B} f(x, y)$ ha minimo in A, la funzione $y \mapsto \inf_{x\in A} f(x, y)$ ha massimo in B, e*

$$\min_{x\in A} \sup_{y\in B} f(x, y) = \max_{y\in B} \inf_{x\in A} f(x, y) = \lambda \in \mathbb{R}.$$

In tal caso, se (x_0, y_0) è punto di sella per f si ha $f(x_0, y_0) = \lambda$.

Dimostrazione Anzitutto definiamo per comodità

$$\varphi(x) = \sup_{y\in B} f(x, y), \qquad \psi(y) = \inf_{x\in A} f(x, y),$$

ed osserviamo che, come nell'osservazione 2.6.14, è facile verificare che

$$\sup_{y\in B} \inf_{x\in A} f(x, y) = \sup_{y\in B} \psi(y) \le \inf_{x\in A} \varphi(x) = \inf_{x\in A} \sup_{y\in B} f(x, y).$$

Notiamo che un punto $(x_0, y_0) \in A\times B$ è punto di sella per f, per definizione, se e solo se

$$\varphi(x_0) = \max_{y\in B} f(x_0, y) = f(x_0, y_0) = \min_{x\in A} f(x, y_0) = \psi(y_0);$$

quindi se f ha un punto di sella (x_0, y_0) dalla relazione precedente segue

$$\inf_{x\in A} \varphi(x) \le \varphi(x_0) = \psi(y_0) \le \sup_{y\in B} \psi(y)$$

e pertanto vale l'uguaglianza

$$\inf_{x\in A} \varphi(x) = \sup_{y\in B} \psi(y).$$

Detto λ tale valore, si ha necessariamente $f(x_0, y_0) = \lambda$.
Viceversa, supponiamo che φ abbia minimo nel punto $x_0 \in A$ e che ψ abbia massimo nel punto $y_0 \in B$, con

$$\varphi(x_0) = \min_{x\in A} \varphi(x) = \max_{y\in B} \psi(y) = \psi(y_0) = \lambda.$$

Allora il punto (x_0, y_0) è di sella per f poiché

$$f(x_0, y) \le \varphi(x_0) = \lambda = f(x_0, y_0) = \psi(y_0) \le f(x, y_0) \qquad \forall x \in A, \quad \forall y \in B.$$

La tesi è provata. $\square$

Dimostriamo ora il seguente teorema, la cui importanza applicativa sarà evidenziata nel prossimo paragrafo:

Teorema 7.3.3 (del minimax, o di von Neumann) *Siano X, Y spazi di Banach riflessivi, e siano $A \subset X$, $B \subset Y$ non vuoti, convessi, chiusi e limitati. Sia inoltre $f : A \times B \to \mathbb{R}$ una funzione tale che:*

(i) *$x \mapsto f(x, y)$ è quasi convessa e semicontinua inferiormente per ogni $y \in B$;*

(ii) *$y \mapsto f(x, y)$ è quasi concava e semicontinua superiormente per ogni $x \in A$.*

Allora f ha un punto di sella in $A \times B$.

Ricordiamo (osservazione 2.1.5) che f è quasi convessa se i suoi insiemi di sottolivello sono convessi, e che f è quasi concava se $-f$ è quasi convessa, ossia gli insiemi di sopralivello di f, vale a dire $F_\lambda = \{x \in A : f(x) \geq \lambda\}$, sono convessi.

Dimostrazione Poniamo

$$a = \min_{x\in A} \max_{y\in B} f(x, y), \qquad b = \max_{y\in B} \min_{x\in A} f(x, y);$$

grazie alla proposizione 7.3.2, basterà provare che i numeri a e b esistono e coincidono.
Fissato $y \in B$, la funzione $x \mapsto f(x, y)$ è quasi convessa e semicontinua inferiormente in A, e per l'osservazione 2.3.5 ha minimo in A in un punto $z(y) \in A$:

$$f(z(y), y) = \min_{x\in A} f(x, y) \qquad \forall y \in B.$$

Poniamo ora

$$h(y) = -f(z(y), y), \qquad y \in B,$$

e mostriamo che h è quasi convessa e semicontinua inferiormente in B. Fissato $t \in \mathbb{R}$, consideriamo gli insiemi

$$H = \{y \in B : \; h(y) \leq t\}, \quad G(w) = \{y \in B : \; -f(z(w), y) \leq t\}, \; w \in B.$$

Osserviamo che $H \subseteq G(w)$ per ogni $w \in B$, essendo $f(z(y), y) \leq f(z(w), y)$ per ogni $w, y \in B$; inoltre $G(w)$ è convesso e chiuso, in quanto la funzione $y \mapsto -f(z(w), y)$ è quasi convessa e semicontinua inferiormente. Proviamo che H è chiuso: se $\{y_n\} \subseteq H$ e $y_n \to y$ in B, allora si ha $y \in G(w)$ per ogni $w \in B$; in particolare $y \in G(y)$, ossia $y \in H$. Proviamo che H è convesso:

se $y, v \in H$ e $\lambda \in]0,1[$, allora posto $u = (1-\lambda)y + \lambda v$ si ha $u \in G(w)$ per ogni $w \in B$; in particolare $u \in G(u)$, ossia $u \in H$. Essendo H convesso e chiuso, si deduce che h è quasi convessa e semicontinua inferiormente per la proposizione 2.2.2. Pertanto, dall'osservazione 2.3.5 segue che h ha minimo m sul convesso chiuso e limitato B; posto $b = -m$, si ha

$$b = -\min_{y \in B} \left[-\min_{x \in A} f(x,y) \right] = \max_{y \in B} \min_{x \in A} f(x,y).$$

In modo analogo si prova che esiste il numero a. Inoltre risulta

$$\min_{x \in A} f(x,y) \le \min_{x \in A} \max_{y \in B} f(x,y) = a \qquad \forall y \in B,$$

da cui la disuguaglianza $b \le a$.
Proviamo ora che $a \le b$. Consideriamo le topologie deboli su X e su Y, fissiamo $\varepsilon > 0$ e introduciamo la multifunzione $T : A \times B \to \mathcal{P}(A \times B)$ definita da

$$T(x,y) = \{(u,v) \in A \times B : \; f(u,y) < b + \varepsilon, \; f(x,v) > a - \varepsilon\}.$$

Si ha $T(x,y) \neq \emptyset$ perché $\min_{u \in A} f(u,y) \le b$ e $\max_{v \in B} f(x,v) \ge a$; inoltre $T(x,y)$ è convesso in virtù delle ipotesi fatte su f. Proviamo che per ogni $(u,v) \in A \times B$ l'insieme $T^{-1}(u,v)$ è aperto nello spazio $X \times Y$, munito della topologia prodotto delle due topologie deboli su X e su Y (si noti che $X \times Y$ è uno spazio vettoriale topologico localmente convesso di Hausdorff, e che $A \times B$ è compatto e convesso in tale spazio). Si ha infatti

$$\begin{aligned} T^{-1}(u,v) &= \{(x,y) \in A \times B : \; (u,v) \in T(x,y)\} = \\ &= \{(x,y) \in A \times B : \; f(u,y) < b + \varepsilon, \; f(x,v) > a - \varepsilon\} = \\ &= \{x \in A : \; f(x,v) > a - \varepsilon\} \times \{y \in B : \; f(u,y) < b + \varepsilon\}; \end{aligned}$$

dal momento che $\{x \in A : \; f(x,v) \le a - \varepsilon\}$ e $\{y \in B : \; f(u,y) \ge b + \varepsilon\}$ sono convessi chiusi, dunque debolmente chiusi, in A ed in B rispettivamente, essi sono anche debolmente chiusi in X ed in Y, cxosicché i loro complementari sono debolmente aperti in X ed in Y. Pertanto il loro prodotto cartesiano è aperto nel prodotto $X \times Y$.
In definitiva, la multifunzione T verifica le ipotesi della proposizione 7.2.9 nello spazio $X \times Y$ munito del prodotto delle topologie deboli: pertanto esiste

un punto fisso per T, ossia esiste $(x_0, y_0) \in A \times B$ tale che $(x_0, y_0) \in T(x_0, y_0)$, ossia

$$a - \varepsilon < f(x_0, y_0) < b + \varepsilon.$$

Quindi $a < b + 2\varepsilon$, ed essendo ε arbitrario si conclude che $a \leq b$.
Per la proposizione 7.3.2, esiste un punto di sella per f in $A \times B$. $\square$

Esercizi 7.3

1. Siano $A = B = \{1, 2, 3\}$ e siano $F_1, F_2 : A \times B \to \mathbb{R}$ le funzioni i cui valori sono definiti dalle matrici

$$F_1 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 2 & 3 & 1 \end{pmatrix}, \qquad F_2 = \begin{pmatrix} -1 & 3 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   (per $i, j \in \{1, 2, 3\}$ i valori $F_1(i, j), F_2(i, j)$ sono dati dall'elemento di posto (i, j) della matrice corrispondente). Verificare che F_1 ha punti di sella (quanti?) mentre F_2 non ne ha.

2. Siano A, B insiemi non vuoti, e sia $f : A \times B \to \mathbb{R}$ una funzione. È vero che che se $(x_0, y_0) \in A \times B$ verifica

$$f(x_0, y_0) = \min_{x \in A} \max_{y \in B} f(x, y) = \max_{y \in B} \min_{x \in A} f(x, y),$$

   allora (x_0, y_0) è necessariamente punto di sella per f oppure per $-f$?

7.4 Teoria dei giochi

Consideriamo un gioco fra due giocatori P e Q: per vincere a questo gioco, P e Q hanno a disposizione due insiemi di strategie A e B. Ogni $x \in A$ ed ogni $y \in B$ costituiscono una possibile "mossa" del gioco, o sequenza di mosse del gioco, per P e per Q; se P sceglie la strategia x e Q sceglie la strategia y, vi sono due "funzioni di utilità" $f_P(x, y)$ e $f_Q(x, y)$ che rappresentano la vincita (o la perdita, se negative) per P e per Q conseguenti alla strategia scelta. Nel caso più semplice, che è quello dei cosiddetti *giochi a somma zero*, si ha $f_P(x, y) = -f_Q(x, y)$, il che significa che la funzione di utilità $g(x, y) = f_Q(x, y)$ rappresenta simultaneamente il guadagno di Q e la perdita di P (se positiva), o viceversa la perdita di Q e il guadagno di P (se

negativa).
Nei giochi a somma zero esiste un criterio soddisfacente per scegliere la strategia “migliore” relativa a ciascun giocatore.

Definizione 7.4.1 *La coppia* $(x_0, y_0) \in A \times B$ *si dice* coppia strategica ottimale *se* (x_0, y_0) *è punto di sella per la funzione di utilità* g *in* $A \times B$*.*

Dalla proposizione 7.3.2 segue:

Proposizione 7.4.2 *Siano* A, B *insiemi non vuoti e sia* $g : A \times B \to \mathbb{R}$ *la funzione di utilità di un gioco a somma zero con insiemi di strategie* A *e* B*. Esiste una coppia strategica ottimale relativa alla funzione* g *se e solo se*

$$\min_{x \in A} \max_{y \in B} g(x, y) = \max_{y \in B} \min_{x \in A} g(x, y) = \lambda;$$

in tal caso, se (x_0, y_0) *è una coppia strategica ottimale, allora* $g(x_0, y_0) = \lambda$*. Il numero* λ *si dice* valore *del gioco.* $\square$

Interpretiamo il risultato della proposizione precedente: ciascun giocatore gioca con una strategia volta ad ottimizzare il proprio interesse (così si suppone). Poiché $g(x, y)$ rappresenta ciò che guadagna Q ed insieme ciò che perde P, chiaramente P punta a minimizzare g mentre Q punta a massimizzare g. Supponiamo, per fissare le idee, che il valore del gioco sia positivo. Se P sceglie la mossa x_0, il guadagno di Q sarà, nella migliore delle ipotesi, pari al più a $\max_{y \in B} g(x_0, y)$; la scelta della strategia y_0 garantisce a Q esattamente tale guadagno, poiché $g(x_0, y_0) = \max_{y \in B} g(x_0, y)$. Simmetricamente, se Q opta per la mossa y_0, la perdita di P sarà, nella migliore delle ipotesi, pari a $\min_{x \in A} g(x, y_0)$; scegliendo la mossa x_0, P avrà la certezza di non perdere di più, in quanto $g(x_0, y_0) = \min_{x \in A} g(x, y_0)$.
Più in generale, se P gioca una mossa qualunque x, Q sa che, giocando bene, può arrivare a una vincita al più uguale a $\max_{y \in B} g(x, y)$; d'altronde Q non può impedire a P di rendere minima tale vincita riducendola a $\min_{x \in A} \max_{y \in B} g(x, y)$ (ossia a $g(x_0, y_0)$). Tuttavia, se Q sceglie y_0, otterrà una vincita $g(x, y_0)$ che supererà comunque tale stima pessimistica. Similmente, ad una mossa qualsiasi y da parte di Q, P potrà rispondere in modo ottimale perdendo soltanto $\min_{x \in A} g(x, y)$, ma non potrà farci niente se la mossa di Q era tale da massimizzare tale perdita elevandola a $\max_{y \in B} \min_{x \in A} g(x, y)$; però se P sceglie x_0 la sua perdita sarà $g(x_0, y)$, cioè comunque inferiore al valore precedente.

Tutto questo mostra che la coppia strategica ottimale (x_0, y_0) garantisce un bilanciamento ottimale degli interessi contrapposti dei due giocatori.
Il valore assunto da g nel punto (x_0, y_0), ossia il valore del gioco (che è univocamente determinato anche se la coppia strategica ottimale non è unica, in virtù della proposizione 7.4.2), ci dice se il gioco è *equo* oppure no: se $g(x_0, y_0) = 0$, e se i giocatori giocano entrambi in modo ottimale, nessuno dei due vincerà né perderà; se $g(x_0, y_0) > 0$, il gioco è squilibrato a favore di Q, che scegliendo una strategia ottimale guadagnerà comunque, mentre se $g(x_0, y_0) < 0$ il favorito dal gioco sarà P.

Esempio 7.4.3 (pari o dispari) I giocatori P e Q dichiarano ciascuno un numero: se la somma dei numeri è pari, P vince 1 euro, se è dispari P perde 1 euro; per Q avviene l'opposto. Le strategie di P e Q sono $A = B = \{p, d\}$, e la funzione di utilità $g = f_Q = -f_P$ è

$$g(p, d) = g(d, p) = 1, \qquad g(p, p) = g(d, d) = -1.$$

Si verifica subito che

$$\min_{x \in A} \max_{y \in B} g(x, y) = 1, \qquad \max_{y \in B} \min_{x \in A} g(x, y) = -1,$$

cosicché non esistono strategie ottimali. La cosa non sorprende, perché A e B non sono convessi!

Per applicare la teoria precedente al caso di insiemi di strategie finiti, bisogna introdurre una versione probabilistica del gioco. Si assume che il gioco (a somma zero) sia stato giocato molte volte, e che i giocatori P e Q scelgano le loro strategie sulla base di fissate probabilità, o meglio di frequenze di successi: se le strategie di P e Q sono $E = \{E_1, \ldots, E_N\}$ e $F = \{F_1, \ldots, F_M\}$, e se $g(E_i, F_j)$ è la funzione di utilità risultante dalle scelte E_i e F_j, supporremo che P scelga la mossa E_i con probabilità p_i, e che Q scelga la mossa F_j con probabilità q_j. Poniamo allora

$$A = \left\{ p \in \mathbb{R}^N : \ 0 \le p_i \le 1, \ \sum_{i=1}^{N} p_i = 1 \right\},$$

$$B = \left\{ q \in \mathbb{R}^M : \ 0 \le q_j \le 1, \ \sum_{j=1}^{M} q_j = 1 \right\};$$

un elemento $p \in A$ rappresenta dunque la strategia di P consistente nello scegliere la mossa E_1 con probabilità p_1, la mossa E_2 con probabilità p_2, eccetera.

Definizione 7.4.4 *Ogni coppia* $(p, q) \in A \times B$ *si dice* coppia strategica mista. *La quantità*

$$L(p, q) = \sum_{i=1}^{N} \sum_{j=1}^{M} g(E_i, F_j) p_i q_j$$

è la speranza di vincita *per il giocatore* Q.

La funzione $L(p, q)$ esprime il valore statisticamente probabile di vincita per Q, misurato dopo molte "mani" del gioco.

Definizione 7.4.5 *Una coppia* $(p_0, q_0) \in A \times B$ *è una* coppia strategica mista ottimale *se* (p_0, q_0) *è punto di sella per la funzione* $L(p, q)$.

Dunque una coppia (p_0, q_0) è ottimale se

$$L(p_0, q) \leq L(p_0, q_0) \leq L(p, q_0) \qquad \forall (p, q) \in A \times B.$$

Si ha allora questo risultato:

Teorema 7.4.6 *Nelle ipotesi precedenti esiste almeno una coppia strategica mista ottimale* (p_0, q_0), *e si ha*

$$L(p_0, q_0) = \min_{p \in A} \max_{q \in B} L(p, q) = \max_{q \in B} \min_{p \in A} L(p, q).$$

Dimostrazione Scelti $X = \mathbb{R}^N$ e $Y = \mathbb{R}^M$, basta applicare il teorema del minimax (teorema 7.3.3) alla funzione $L(p, q)$ che, essendo lineare in p e q, è continua, nonché simultaneamente concava e convessa in ciascuna variabile sui convessi A e B, i quali sono limitati e chiusi. □

Esempio 7.4.7 Nel caso del gioco "pari o dispari" (esempio 7.4.3), le mosse sono $E_1 = F_1 = p$, $E_2 = F_2 = d$, le strategie miste sono

$$A = B = \{p = (p_1, p_2) : \ p_1, p_2 \in [0, 1], \ p_1 + p_2 = 1\}$$

e la funzione di utilità è

$$g(E_i, F_j) = \begin{cases} -1 & \text{se } E_i = F_j \\ \ \ 1 & \text{se } E_i \neq F_j \ ; \end{cases}$$

la funzione speranza di vincita è

$$L(p,q) = \sum_{i,j=1}^{2} g(E_i, F_j)p_i q_j = -p_1q_1 - p_2q_2 + p_1q_2 + p_2q_1 = -(p_1 - p_2)(q_1 - q_2).$$

Si vede subito che un punto di sella è $(\overline{p}, \overline{q})$ con $\overline{p} = \overline{q} = (\frac{1}{2}, \frac{1}{2})$, e il valore del gioco è, come potevamo aspettarci,

$$L(\overline{p}, \overline{q}) = -\left(\frac{1}{2} - \frac{1}{2}\right)\left(\frac{1}{2} - \frac{1}{2}\right) = 0.$$

Si noti che $(\frac{1}{2}, \frac{1}{2})$ è l'unico punto di sella per L: infatti se (p^*, q^*) è un altro punto di sella, deve essere $L(p^*, q^*) = 0$, e quindi si ha $p_1^* = p_2^*$ oppure $q_1^* = q_2^*$, da cui $p^* = \overline{p}$ oppure $q^* = \overline{p}$. Ma se $p^* = \overline{p} = (\frac{1}{2}, \frac{1}{2})$, dalle relazioni

$$0 = L(\overline{p}, q^*) = L(p^*, q^*) \leq L(p, q^*) \qquad \forall p \in A,$$

scegliendo $p = (\frac{1}{4}, \frac{3}{4})$ oppure $p = (\frac{3}{4}, \frac{1}{4})$ si deduce $0 \leq \pm\frac{1}{2}(q_1^* - q_2^*)$, e quindi $q^* = \overline{p} = (\frac{1}{2}, \frac{1}{2})$. In modo analogo, se $q^* = (\frac{1}{2}, \frac{1}{2})$ si deduce $p^* = \overline{q} = (\frac{1}{2}, \frac{1}{2})$.
In definitiva, il gioco è equo e la strategia ottimale per entrambi i giocatori è quella di dichiarare un numero pari con probabilità $\frac{1}{2}$ e un numero dispari con probabilità $\frac{1}{2}$. (In effetti, trattandosi di un gioco essenzialmente di fortuna, i due giocatori, mediamente, vinceranno ciascuno con probabilità $\frac{1}{2}$, e il guadagno di ciascuno sarà mediamente nullo.)

Per i giochi non a somma zero la teoria è meno semplice. Per mostrare quello che può accadere, è utile considerare il seguente

Esempio 7.4.8 (dilemma dei prigionieri) Sono stati arrestati dalla polizia due malfattori che devono scontare un anno di prigione ciascuno per un reato di poco conto. Si sospetta però che i due siano stati complici in un delitto più grave, punibile con ulteriori cinque anni di prigione; tuttavia non si è in grado di provarlo. Nel tentativo di incriminarli per il reato maggiore, il procuratore fa a ciascuno di loro, separatamente, la seguente promessa: se egli accuserà il socio del crimine maggiore, gli verrà abbuonato l'anno di galera per il reato minore, e se poi il socio non lo accuserà del delitto maggiore (nel qual caso dovrà invece scontare i relativi 5 anni), sarà liberato.
La situazione può essere descritta come un gioco fra i due malfattori P e Q. Entrambi possono scegliere fra le due strategie di accusare l'altro (a) o di non accusarlo (n); le corrispondenti funzioni di utilità si definiscono come l'opposto degli anni da scontare (in modo da dover essere massimizzate). Quindi risulterà:

$$f_P(a,a) = -5, \quad f_P(a,n) = 0, \quad f_P(n,a) = -6, \quad f_P(n,n) = -1;$$

$$f_Q(a,a) = -5, \quad f_Q(a,n) = -6, \quad f_Q(n,a) = 0, \quad f_Q(n.n) = -1.$$

Si riconosce immediatamente che la coppia strategica consistente nell'accusarsi a vicenda fornisce un esito peggiore per tutti e due rispetto alla scelta di omertà per entrambi. Eppure, ciascun malfattore nota che la scelta di accusare l'altro gli arrecherà il beneficio di scontare, in qualunque caso, un anno in meno: infatti, se l'altro non lo accusa uscirà subito invece di star dentro un anno, mentre se l'altro lo accusa a sua volta, sconterà 5 anni anziché 6. Il risultato finale è che entrambi i rei si accuseranno, e dunque staranno in galera per 5 anni: con la strategia omertosa, invece, se la sarebbero cavata con un anno per ciascuno.
Come si vede, la strategia omertosa, che da un punto di vista cooperativo è la soluzione migliore, non lo è dal punto di vista individuale: anche in presenza di un accordo clandestino fra i malfattori per tacere entrambi, nessuno dei due può avere la garanzia che, non accusando l'altro, costui mantenga la parola di non accusarlo a sua volta; è troppo grande il rischio di beccarsi 6 anni vedendo l'altro uscire di galera.

L'idea che la razionalità individuale preceda quella collettiva, fenomeno tipico dei giochi non cooperativi (in cui non c'è possibilità di accordi o alleanze fra i giocatori), è alla base della nozione di equilibrio di Nash.

Definizione 7.4.9 *La coppia* $(x_0, y_0) \in A \times B$ *si dice* punto di equilibrio di Nash *se per ogni altra strategia* $(x, y) \in A \times B$ *valgono le disuguaglianze*

$$f_P(x_0, y_0) \geq f_P(x, y_0), \qquad f_Q(x_0, y_0) \geq f_Q(x_0, y).$$

In altre parole, un punto di equilibrio di Nash è una coppia di strategie che massimizza l'utilità di un singolo giocatore quando l'avversario si attiene alla propria. Il senso è questo: a P conviene scegliere la strategia x_0, perché ciò invoglierà Q a scegliere y_0 onde massimizzare f_Q, e in questo modo P massimizzerà f_P. Il discorso è simmetrico dal punto di vista di Q. Il guaio è che l'esistenza di un punto di equilibrio di Nash non esclude affatto che possa esistere una coppia strategica (x, y) per la quale risulti simultaneamente $f_P(x, y) > f_P(x_0, y_0)$ e $f_Q(x, y) > f_Q(x_0, y_0)$: è precisamente ciò che accade nel dilemma dei prigionieri con la strategia della non accusa reciproca.

Osservazioni 7.4.10 **(1)** Se il gioco è a somma zero (ossia $f_Q = -f_P = g$), allora un punto (x_0, y_0) fornisce un equilibrio di Nash se e solo se esso è punto di sella per la funzione g. Infatti dalla definizione 7.4.9 segue subito, per ogni $(x, y) \in A \times B$,

$$g(x_0, y) = f_Q(x_0, y) \leq f_Q(x_0, y_0) = -f_P(x_0, y_0) \leq -f_P(x, y_0) = g(x, y_0)$$

cosicché (x_0, y_0) è punto di sella per g; il viceversa è del tutto analogo.

(2) Nell'esempio del dilemma dei prigionieri è facile verificare che l'unico punto di equilibrio di Nash è la strategia dell'accusa reciproca.

Teorema 7.4.11 *Siano X, Y spazi di Banach e siano $A \subset X$, $B \subset Y$ non vuoti, convessi e compatti. Siano $f, g : A \times B \to \mathbb{R}$ funzioni continue tali che $x \mapsto f(x, y)$ sia concava per ogni $y \in B$ e $y \mapsto g(x, y)$ sia concava per ogni $x \in X$. Allora esiste un punto di equilibrio di Nash $(x_0, y_0) \in A \times B$ per le funzioni f e g.*

Dimostrazione Introduciamo la funzione $F : (A \times B) \times (A \times B) \to \mathbb{R}$, definita da

$$F(p, q) = f(p_1, q_2) + g(q_1, p_2) \qquad \forall p, q \in A \times B,$$

e osserviamo che F è continua, nonché concava in p per ogni fissato q. Per questa funzione vale il seguente

Lemma 7.4.12 *Nelle ipotesi precedenti, esiste $q_0 \in A \times B$ tale che*

$$\max_{p \in A \times B} F(p, q_0) = F(q_0, q_0).$$

Dimostrazione La disuguaglianza $\geq$ è banale per ogni $q_0 \in A \times B$. Ragioniamo per assurdo, ammettendo che per ogni $q \in A \times B$ esista $p \in A \times B$ per cui risulti $F(p, q) > F(q, q)$. Poniamo

$$G_p = \{q \in A \times B : F(p, q) > F(q, q)\}, \qquad p \in A \times B :$$

l'insieme G_p è aperto in $A \times B$, e per l'ipotesi fatta si ha $A \times B \subseteq \bigcup_{p \in A \times B} G_p$. L'insieme $A \times B$ è compatto nello spazio $X \times Y$, quindi esistono $p_1, \ldots, p_k \in A \times B$ tali che $A \times B \subseteq \bigcup_{i=1}^{k} G_{p_i}$. Definiamo adesso le funzioni

$$\varphi_i(q) = \max\{F(p_i, q) - F(q, q), 0\}, \quad q \in A \times B, \quad i = 1, \ldots, k,$$

le quali sono continue e non negative; osserviamo che per ogni $q \in A \times B$ esiste $i \in \{1, \dots k\}$ tale che $q \in G_{p_i}$ e dunque $\varphi_i(q) > 0$. Normalizziamo le φ_i ponendo

$$\psi_i(q) = \frac{\varphi_i(q)}{\sum_{j=1}^k \varphi_j(q)},$$

cosicché $\sum_{i=1}^k \psi_i(q) = 1$ per ogni $q \in A \times B$. Posto $\psi(q) = \sum_{i=1}^k \psi_i(q) p_i$, l'applicazione ψ manda il convesso $A \times B$ in sé ed è continua. Poiché $A \times B$ è compatto, per il teorema di Tikhonov (teorema 7.2.6) esiste un punto fisso $q' \in A \times B$ per ψ. Si ha dunque $q' = \sum_{i=1}^k \psi_i(q') p_i$, da cui per concavità

$$F(q', q') = F\left(\sum_{i=1}^k \psi_i(q') p_i, q'\right) \geq \sum_{i=1}^k \psi_i(q') F(p_i, q').$$

D'altra parte, se $\psi_i(q') > 0$ si ha $F(p_i, q') > F(q', q')$, e ciò implica

$$F(q', q') \geq \sum_{i=1}^k \psi_i(q') F(p_i, q') > \sum_{i=1}^k \psi_i(q') F(q', q') = F(q', q'),$$

il che è assurdo. Il lemma è provato. $\square$

Torniamo alla dimostrazione del teorema 7.4.11. Per il lemma, esiste $q_0 \in A \times B$ tale che

$$\max_{p \in A \times B} F(p, q_0) = F(q_0, q_0).$$

Posto $(x_0, y_0) = q_0$, si ha allora

$$f(x, y_0) + g(x_0, y) \leq f(x_0, y_0) + g(x_0, y_0) \qquad \forall (x_0, y_0) \in A \times B;$$

scegliendo $x = x_0$ si deduce $g(x_0, y) \leq g(x_0, y_0)$ per ogni $y \in B$, mentre scegliendo $y = y_0$ si trova $f(x, y_0) \leq f(x_0, y_0)$ per ogni $x \in A$. Pertanto (x_0, y_0) è punto di equilibrio di Nash. $\square$

Esercizi 7.4

1. Si analizzi, come si è fatto negli esempi 7.4.3 e 7.4.7, il gioco "sasso-carta-forbici": due giocatori P e Q scelgono simultaneamente S, C o F. Tenuto conto che il sasso spunta le forbici, le forbici tagliano la

carta e la carta nasconde il sasso, e che il gioco vuole essere a somma zero, si consideri la seguente funzione di utilità per il giocatore Q:

	S	C	F
S	0	-1	1
C	1	0	-1
F	-1	1	0

Si provi che il gioco, nella sua versione probabilistica, è equo e che ad entrambi i giocatori conviene alternare le tre mosse con probabilità $\frac{1}{3}$ ciascuna.

2. Siano $A = \{E_1, \ldots, E_N\}$ e $B = \{F_1, \ldots, F_M\}$ le strategie di un gioco a somma zero per due giocatori P e Q, e si supponga che la funzione di utilità g del giocatore Q non abbia punti di sella. Si provi che Q, giocando intelligentemente, può garantirsi un guadagno non inferiore a $\underline{v} = \max_j \min_i g(E_i, F_j)$, e che similmente P, giocando con astuzia, non permetterà a tale guadagno di superare $\overline{v} = \min_i \max_j g(E_i, F_j)$.

3. *(Battaglia fra i sessi)* Due innamorati U e D devono decidere come passare la serata: seguendo i rispettivi stereotipi culturali, U vorrebbe andare alla partita di calcio mentre D preferirebbe andare a ballare, ma soprattutto entrambi desiderano trascorrere la serata assieme. Indicate con c e b le corrispondenti scelte, definiamo quindi le funzioni di utilità f_U e f_D nel modo seguente (in verticale le scelte di U, in orizzontale quelle di D):

$$
(f_U, f_D) : \quad \begin{array}{c|cc} b & (-2,-2) & (1,2) \\ c & (2,1) & (-1,-1) \\ \hline & c & b \end{array}
$$

Si determinino i punti di equilibrio di Nash per questo gioco.

7.5 Selezioni

L'uso di una multifunzione si presta a rappresentare un'incertezza nella determinazione del valore di una data funzione f: dato x, può capitare che il valore $f(x)$ sia noto solo pprossimativamente, e quindi si conosca solo un insieme $F(x)$ al quale $f(x)$ certamente appartiene. Una selezione f di una multifunzione F è un modo di scegliere per ogni x, in assenza di più precise informazioni, un possibile valore $f(x)$ tra tutti quelli ammissibili, cioè fra gli elementi di $F(x)$.

Definizione 7.5.1 *Siano W, M insiemi non vuoti e sia $F : W \to \mathcal{P}(M)$ una multifunzione. Una* selezione *di F è un'applicazione $m : W \to M$ tale che $m(w) \in F(w)$ per ogni $w \in W$.*

Un'ovvia condizione necessaria affinché almeno una selezione esista è che si abbia $F(w) \neq \emptyset$ per ogni $w \in W$; in questo caso, l'esistenza di una selezione è sempre e comunque garantita dall'assioma della scelta. Il problema che ci poniamo, e che è importante nelle applicazioni, è quello di costruire selezioni dotate di particolari proprietà: quando W e M sono insiemi finiti, ci interesseranno selezioni iniettive; quando W è uno spazio misurato e M è, ad esempio, $\mathbb{R}$, vorremo trovare selezioni misurabili; quando W e M sono spazi topologici, andremo in cerca di selezioni continue.
Cominciamo dal caso di insiemi finiti.

Teorema 7.5.2 (del matrimonio) *Siano W, M insiemi non vuoti, con W finito; sia $F : W \to \mathcal{P}(M)$ una multifunzione. Allora esiste una selezione iniettiva $m : W \to M$ se e solo se risulta*

$$\#S \leq \#F(S) \qquad \forall S \subseteq W,$$

ove $\#S$ è la cardinalità di S.

Prima di dimostrare il teorema, vediamo come esso si applica al *problema del matrimonio*. Vi è un certo insieme $W \neq \emptyset$ di donne ed un certo insieme $M \neq \emptyset$ di uomini; l'obiettivo è far sposare più donne possibile. Però il matrimonio deve aver luogo fra individui "compatibili", o "amici" (mutuamente attraenti). La multifunzione F è l' "amicizia": ad ogni donna $w \in W$ viene associato l'insieme $F(w) \subseteq M$ dei suoi amici di sesso maschile, tra i quali ella dovrà scegliersi lo sposo. La condizione $\#S \leq \#F(S)$ dice che un arbitrario insieme di donne non è più grosso dell'insieme di tutti gli amici di esse; in particolare è escluso il caso in cui due donne debbono per forza litigarsi lo stesso uomo. La selezione $m : W \to M$ è il matrimonio: $m(w)$ è il marito scelto da w nella cerchia dei suoi amici (per garantire la monogamia femminile, m è una funzione e non una multifunzione); m deve essere iniettiva per garantire, simmetricamente, la monogamia dei maschi. Il teorema 7.5.2 dice allora che tutte le donne possono sposarsi.

Dimostrazione È chiaro che la condizione $\#S \le \#F(S)$ per ogni $S \subseteq W$ è necessaria per l'esistenza di una selezione iniettiva: infatti se questa esiste si ha

$$\#S = \#m(S) \le \#F(S) \qquad \forall S \subseteq W.$$

Proviamo che la condizione è anche sufficiente. Ragioneremo per induzione sulla cardinalità di W.
Se $\#W = 1$, c'è un unico elemento $w \in W$, al quale possiamo associare un qualsiasi elemento $\mu \in F(w)$ (per ipotesi, $\#F(w) \ge 1$); posto $m(w) = \mu$, abbiamo finito.
Fissato $n > 1$, supponiamo di sapere che se $\#W < n$ ogni multifunzione $G : W \to \mathcal{P}(M)$ abbia una selezione iniettiva $m : W \to M$; consideriamo il caso $\#W = n$ e la multifunzione F. Vi sono due sottocasi:

(I) $\#S < \#F(S)$ per ogni $S \subset W$ non vuoto;

(II) esiste $S' \subset W$, non vuoto, tale che $\#S' = \#F(S')$.

Nel caso (I), scegliamo $\overline{w} \in W$ e $\overline{\mu} \in F(\overline{w})$, e definiamo $m(\overline{w}) = \overline{\mu}$. Resta da definire m sull'insieme $W \setminus \{\overline{w}\}$, che è non vuoto e per il quale quindi si ha, per ipotesi,

$$n - 1 = \#(W \setminus \{\overline{w}\}) < \#F(W \setminus \{\overline{w}\}).$$

Consideriamo la multifunzione $F_1 : W \setminus \{\overline{w}\} \to \mathcal{P}(M \setminus \{\overline{\mu}\})$, definita da

$$F_1(w) = F(w) \setminus \{\overline{\mu}\} \qquad \forall w \in W \setminus \{\overline{w}\}.$$

Dalla definizione di F_1 segue che

$$F_1(S) = F(S) \setminus \{\overline{\mu}\} \qquad \forall S \subseteq W \setminus \{\overline{w}\},$$

ed in particolare

$$\#F_1(S) \ge \#F(S) - 1 \ge \#S \qquad \forall S \subseteq W \setminus \{\overline{w}\}.$$

Dunque, per ipotesi induttiva, la multifunzione F_1 ha una selezione iniettiva $m_1 : W \setminus \{\overline{w}\} \to M \setminus \{\overline{\mu}\}$, e si ha $m_1(w) \in F_1(w) \subseteq F(w)$ per ogni $w \in W \setminus \{\overline{w}\}$. La funzione

$$m(w) = \begin{cases} \overline{\mu} & \text{se } w = \overline{w} \\ m_1(w) & \text{se } w \ne \overline{w} \end{cases}$$

è la selezione di F cercata.
Nel caso (II), essendo $\#S' < \#W$, per ipotesi induttiva esiste una selezione iniettiva $m' : S' \to F(S')$. Resta da definire la selezione su $W \setminus S'$. A questo scopo consideriamo la multifunzione $F_1 : W \setminus S' \to \mathcal{P}(M \setminus F(S'))$ definita da

$$F_1(w) = F(w) \setminus F(S') \qquad \forall w \in W \setminus S'.$$

A priori $F_1(w)$ potrebbe essere vuoto per qualche w; dimostreremo che si ha, al contrario,

$$\#S \leq \#F_1(S) \qquad \forall S \subseteq W \setminus S'$$

(cosicché, in particolare, $1 \leq \#F_1(w)$ per ogni $w \in W \setminus S'$). Infatti, se per assurdo esistesse $T \subseteq W \setminus S'$ tale che $\#T > \#F_1(T)$, allora dall'inclusione

$$F(S' \cup T) \subseteq F(S') \cup (F(T) \setminus F(S'))$$

dedurremmo

$$\begin{aligned} \#F(S' \cup T) &\leq \#F(S') + \#(F(T) \setminus F(S')) = \\ &= \#S' + \#F_1(T) < \#S' + \#T = \#(S' \cup T), \end{aligned}$$

il che contraddirebbe l'ipotesi $\#S \leq \#F(S)$ per ogni $S \subseteq W$.
Pertanto possiamo applicare l'ipotesi induttiva alla multifunzione F_1: si ottiene che F_1 ha una selezione iniettiva $m_1 : W \setminus S' \to M \setminus F(S')$, e si ha $m_1(w) \in F_1(w) \subseteq F(w)$ per ogni $w \in W \setminus S'$. Posto

$$m(w) = \begin{cases} m'(w) & \text{se } w \in S' \\ m_1(w) & \text{se } w \in W \setminus S', \end{cases}$$

m è la selezione di F cercata. Dunque vale la tesi del teorema quando $\#W = n$, e ciò conclude la dimostrazione. □

Passiamo ora al caso degli spazi misurati (in una situazione particolarmente semplice).

Teorema 7.5.3 (di Kuratowski e Ryll-Nardzewski) *Sia $[a, b] \subset \mathbb{R}$, sia X uno spazio di Banach separabile e sia $F : [a, b] \to \mathcal{P}(X)$ una multifunzione tale che:*

(i) *$F(t)$ è non vuoto e chiuso in X per ogni $t \in [a, b]$;*

(ii) *$F^{-1}(A)$ è misurabile secondo Lebesgue per ogni aperto $A \subseteq X$.*

Allora F ha una selezione $f : [a, b] \to X$ fortemente misurabile.

Dimostrazione Sia $D = \{c_j\}_{j\in\mathbb{N}} \subset X$ un sottoinsieme numerabile e denso: allora a maggior ragione, D è denso in X anche rispetto alla distanza

$$d_X(x, y) = \frac{\|x - y\|}{1 + \|x - y\|} \qquad \forall x, y \in X.$$

In questa dimostrazione lavoreremo con la distanza d_X in luogo della norma di X: notiamo che $d_X(x, y) < 1$ per ogni $x, y \in X$. Inoltre, definiamo

$$B(x, r) = \{y \in X : d_X(x, y) < r\} \qquad \forall x \in X,\ \forall r > 0,$$

$$d_X(x, A) = \inf_{y\in A} d_X(x, y) \qquad \forall x \in X,\ \forall A \subseteq X.$$

Utilizzando la successione $\{c_j\}$, dimostriamo il seguente

Lemma 7.5.4 *Nelle ipotesi del teorema, esiste una successione $\{\varphi_n\}_{n\in\mathbb{N}}$ di funzioni da $[a, b]$ in X con le seguenti proprietà:*

(a) *φ_n è fortemente misurabile;*

(b) *$d_X(\varphi_n(t), F(t)) < 2^{-n}$ per ogni $t \in [a, b]$;*

(c) *$d_X(\varphi_n(t), \varphi_{n-1}(t)) < 2^{1-n}$ per ogni $t \in [a, b]$.*

Dimostrazione Fissiamo $c_{j_0} \in D$: ovviamente si ha $d_X(c_{j_0}, F(t)) < 1$ per ogni $t \in [a, b]$, cosicché, posto $\varphi_0(t) = c_{j_0}$, è evidente che φ_0 verifica (a) e (b). Fissato $p \in \mathbb{N}^+$, supponiamo che siano già state costruite $\varphi_0, \varphi_1, \ldots, \varphi_{p-1}$ verificanti (a) e (b); andiamo a definire una funzione φ_p con le stesse proprietà. Poniamo

$$A_{j,p} = F^{-1}(B(c_j, 2^{-p})) \cap \varphi_{p-1}^{-1}(B(c_j, 2^{1-p})) \qquad \forall j \in \mathbb{N},$$

$$E_{0,p} = A_{0,p}, \quad E_{j,p} = A_{j,p} \setminus \bigcup_{k=0}^{j-1} E_{k,p} \qquad \forall j \in \mathbb{N}^+,$$

ed osserviamo che, ovviamente, gli $E_{j,p}$ sono tutti disgiunti; proviamo, preliminarmente, che risulta $\bigcup_{j=0}^{\infty} E_{j,p} = [a, b]$.
Sia $t \in [a, b]$; poiché $d_X(\varphi_{p-1}(t), F(t)) < 2^{1-p}$, esisterà un punto $\xi_{p,t} \in F(t)$ tale che $d_X(\varphi_{p-1}(t), \xi_{p,t}) < 2^{1-p}$. Adesso scegliamo un numero τ tale che

$$0 < \tau < \frac{1}{2 + \|\varphi_{p-1}(t) - \xi_{p,t}\|_X},$$

il che equivale, come è facile verificare, alla condizione

$$\frac{\tau\|\varphi_{p-1}(t)-\xi_{p,t}\|_X}{1+\tau\|\varphi_{p-1}(t)-\xi_{p,t}\|_X} < \frac{1}{2}\,\frac{\|\varphi_{p-1}(t)-\xi_{p,t}\|_X}{1+\|\varphi_{p-1}(t)-\xi_{p,t}\|_X}\,.$$

Posto $\eta_{p,t} = \tau\varphi_{p-1}(t) + (1-\tau)\xi_{p,t}$, si ha allora

$$d_X(\xi_{p,t}, \eta_{p,t}) < \frac{1}{2} d_X(\varphi_{p-1}(t), \xi_{p,t}) < 2^{-p}.$$

Per densità, esisterà $c_{j_p} \in D$ tale che

$$d_X(c_{j_p}, \eta_{p,t}) < 2^{-p} - \frac{1}{2} d_X(\varphi_{p-1}(t), \xi_{p,t}) < 2^{-p} - d_X(\eta_{p,t}, \xi_{p,t}).$$

Da questa relazione segue intanto

$$d_X(c_{j_p}, F(t)) \le d_X(c_{j_p}, \xi_{p,t}) \le d_X(c_{j_p}, \eta_{p,t}) + d_X(\eta_{p,t}, \xi_{p,t}) < 2^{-p};$$

inoltre si ricava

$$\begin{aligned} d_X(\varphi_{p-1}(t), c_{j_p}) &\le d_X(\varphi_{p-1}(t), \eta_{p,t}) + d_X(\eta_{p,t}, c_{j_p}) < \\ &< d_X(\varphi_{p-1}(t), \eta_{p,t}) + 2^{-p} - \frac{1}{2} d_X(\varphi_{p-1}(t), \xi_{p,t}) < \\ &< d_X(\varphi_{p-1}(t), \xi_{p,t}) + 2^{-p} - \frac{1}{2} d_X(\varphi_{p-1}(t), \xi_{p,t}) = \\ &= \frac{1}{2} d_X(\varphi_{p-1}(t), \xi_{p,t}) + 2^{-p} < 2^{1-p}. \end{aligned}$$

Pertanto, dalla definizione degli $A_{j,p}$ segue che $t \in A_{j_p,p}$. Di conseguenza, o $t \in E_{j_p,p}$ oppure $t \in E_{i,p}$ per qualche $i < j_p$; in ogni caso sarà $t \in \bigcup_{j=0}^{\infty} E_{j,p}$, come si voleva.
Definiamo allora la funzione $\varphi_p : [a,b] \to X$ nel modo seguente:

$$\varphi_p(t) = c_j \qquad \forall t \in E_{j,p}\,, \quad \forall j \in \mathbb{N}.$$

Ora osserviamo che ciascun $E_{j,p}$ è misurabile secondo Lebesgue: infatti sono misurabili gli $A_{j,p}$, a causa dell'ipotesi (ii) del teorema e della forte misurabilità di φ_{p-1} . Dunque, essendo

$$\varphi_p(t) = \lim_{n\to\infty} \sum_{j=0}^{n} c_j \chi_{E_{j,p}}(t) \qquad \forall t \in [a,b],$$

la funzione φ_p è limite puntuale di funzioni semplici e pertanto è fortemente misurabile, ossia verifica (a). Proviamo che φ_p verifica anche (b). Sia $t \in [a, b]$; esiste j tale che $t \in E_{j,p} \subseteq A_{j,p}$. Ciò significa tre cose:

$$\varphi_p(t) = c_j\,, \qquad d_X(\varphi_{p-1}(t), c_j) < 2^{1-p}, \qquad \exists x \in F(t) : \; d_X(x, c_j) < 2^{-p}.$$

Da questi fatti segue

$$d_X(\varphi_p(t), F(t)) \le d_X(c_j, x) < 2^{-p}, \qquad d_X(\varphi_p(t), \varphi_{p-1}(t)) < 2^{1-p},$$

e dunque, qualsiasi sia $t \in [a, b]$, è soddisfatta non solo la proprietà (b) (il che conclude il passo induttivo) ma anche la (c). Dunque la successione $\{\varphi_n\}$ verifica (a) e (b), per induzione, e anche (c), per costruzione. Il lemma 7.5.4 è dimostrato. $\square$

Il teorema di Kuratowski e Ryll-Nardzewski si dimostra ora facilmente. La proprietà (c) del lemma 7.5.4 mostra che $\{\varphi_n\}$ converge puntualmente in X a una funzione f che, dunque, è fortemente misurabile (esercizio 3.1.5). Inoltre, dalla (b) del lemma segue $d_X(f(t), F(t)) = 0$, ed essendo $F(t)$ chiuso si conclude che $f(t) \in F(t)$. Ciò prova il teorema 7.5.3. $\square$

Infine esaminiamo il caso degli spazi topologici.

Teorema 7.5.5 (di Michael) *Sia X uno spazio topologico paracompatto, sia Y uno spazio di Banach e sia $F : X \to \mathcal{P}(Y)$ una multifunzione semicontinua inferiormente e tale che $F(x)$ sia non vuoto, chiuso e convesso per ogni $x \in X$. Allora F ha una selezione continua $f : X \to Y$.*

Si noti che se F è semicontinua superiormente anziché inferiormente il teorema non è vero: ad esempio la multifunzione $F : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ definita da

$$F(x) = \begin{cases} \{0\} & \text{se } x > 0 \\ [0, 1] & \text{se } x = 0 \\ \{1\} & \text{se } x < 0 \end{cases}$$

è semicontinua superiormente, non è semicontinua inferiormente in $x = 0$ e, pur verificando gli insiemi $F(x)$ le ipotesi richieste, F è priva di selezioni continue.

Dimostrazione Tutto si basa sul seguente lemma di approssimazione:

Lemma 7.5.6 *Sia W uno spazio topologico paracompatto, sia M uno spazio di Banach e sia $G : W \to \mathcal{P}(M)$ una multifunzione semicontinua inferiormente, tale che $G(w)$ sia non vuoto e convesso per ogni $w \in W$. Allora per ogni $\varepsilon > 0$ esiste una funzione continua $g : W \to M$ che verifica*

$$d(g(w), G(w)) < \varepsilon \qquad \forall w \in W.$$

Dimostrazione Fissiamo un'arbitraria selezione $m : W \to M$ della multifunzione G (costruibile, ad esempio, con l'assioma della scelta). Poiché G è semicontinua inferiormente, fissati $w \in W$ ed $\varepsilon > 0$ esiste un intorno aperto $U(w)$ di w in W tale che

$$G(x) \cap B(m(w), \varepsilon) \neq \emptyset \qquad \forall x \in U(w).$$

Poiché W è paracompatto, il ricoprimento $\{U(w)\}_{w \in W}$ ha un raffinamento $\{V_i\}_{i \in I}$ localmente finito: ciò significa che per ogni $i \in I$ esiste $w_i \in W$ tale che $V_i \subseteq U(w_i)$, ed inoltre ogni $w \in W$ ha un intorno che interseca un numero finito di intorni V_i. Sia $\{f_i\}_{i \in I}$ una partizione dell'unità associata al ricoprimento $\{V_i\}$: poniamo allora

$$g(x) = \sum_i f_i(x) m(w_i) \qquad \forall x \in W;$$

la somma è finita per ogni fissato $x \in W$, quindi g è ben definita ed è continua perché tali sono le f_i. Inoltre se $f_i(x) > 0$ si ha $x \in V_i \subseteq U(w_i)$ e dunque, per definizione di $U(w_i)$,

$$d(m(w_i), G(x)) < \varepsilon.$$

Ne segue che $g(x)$ è una combinazione convessa di elementi di $G(x) + \varepsilon B(0,1)$; poiché per ipotesi $G(x)$ è convesso, anche tale insieme è convesso e pertanto $g(x) \in G(x) + \varepsilon B(0,1)$. Ciò prova il lemma. $\square$

Proviamo il teorema di Michael. Utilizzando il lemma 7.5.6, per ogni $n \in \mathbb{N}$ andiamo a costruire una funzione continua $g_n : X \to Y$ con le seguenti proprietà:

$$d(g_n(x), F(x)) < 2^{-n} \qquad \forall x \in X,$$
$$\|g_n(x) - g_{n-1}(x)\|_X < 2^{2-n} \qquad \forall x \in X.$$

La costruzione ricalca quella usata nella dimostrazione del lemma 7.5.4. Se $n = 0$, basta applicare il lemma 7.5.6 con $M = X$, $W = Y$, $\varepsilon = 1$ e $G = F$, e poi prendere $g_0 = g$. Se sono state costruite $g_0, g_1, \ldots, g_{n-1}$, consideriamo la multifunzione $G_n : X \to \mathcal{P}(Y)$ definita da

$$G_n(x) = F(x) \cap B(g_{n-1}(x), 2^{1-n});$$

per ipotesi induttiva, $G_n(x)$ è non vuoto e convesso per ogni $x \in X$.
Proviamo che G_n è semicontinua inferiormente: fissato $x \in X$, sia $y \in G_n(x)$ e sia N un intorno di y in Y. Osservato che $y \in B(g_{n-1}(x), 2^{1-n})$, sia $\delta > 0$ tale che $B(y, \delta) \subseteq N \cap B(g_{n-1}(x), 2^{1-n})$, cosicché $\|y - g_{n-1}(x)\|_Y < 2^{1-n} - \delta$; allora, essendo F semicontinua inferiormente, esiste un intorno U_1 di x in X tale che $F(v) \cap B(y, \frac{\delta}{2}) \neq \emptyset$ per ogni $v \in U_1$. D'altra parte, essendo g_{n-1} continua, esiste un altro intorno U_2 di x in X tale che $g_{n-1}(U_2) \subseteq B(g_{n-1}(x), \frac{\delta}{2})$. Di conseguenza, posto $U = U_1 \cap U_2$, se $v \in U$ l'insieme $F(v) \cap B(y, \frac{\delta}{2})$ contiene almeno un elemento z, e si ha

$$\begin{aligned} &\|z - g_{n-1}(v)\|_Y \leq \\ &\quad \leq \|z - y\|_Y + \|y - g_{n-1}(x)\|_Y + \|g_{n-1}(x) - g_{n-1}(v)\|_Y < \\ &\quad < \frac{\delta}{2} + 2^{1-n} - \delta + \frac{\delta}{2} = 2^{1-n}, \end{aligned}$$

cosicché $z \in G_n(v) \cap B(y, \frac{\delta}{2}) \subseteq G_n(v) \cap N$. Ciò prova che $G_n(v) \cap N \neq \emptyset$ per ogni $v \in U$, e dunque G_n è semicontinua inferiormente.
Possiamo perciò applicare il lemma 7.5.6 con $W = X$, $M = Y$, $G = G_n$ e $\varepsilon = 2^{-n}$, trovando una funzione continua $g_n : X \to Y$ tale che

$$d(g_n(x), G_n(x)) < 2^{-n} \qquad \forall x \in X;$$

ne segue che, per costruzione, g_n soddisfa alle due condizioni richieste, le quali per induzione sono vere per ogni n.
La successione $\{g_n\}$ così costruita è di Cauchy rispetto alla convergenza uniforme: poiché Y è uno spazio di Banach, essa converge ad una funzione continua $f : X \to Y$; questa funzione è una selezione di F in quanto per ogni $x \in X$ si ha $d(f(x), F(x)) = 0$ e $F(x)$ è chiuso. □

Nel caso speciale in cui Y è uno spazio di Hilbert e la multifunzione F è continua, il teorema 7.5.5 si può migliorare e precisare.

Teorema 7.5.7 *Sia X uno spazio topologico, sia Y uno spazio di Hilbert e sia $F : X \to \mathcal{P}(Y)$ una multifunzione continua e tale che $F(x)$ sia non vuoto,* *convesso e chiuso per ogni* $x \in X$*. Allora la proiezione di* 0 *sul convesso* $F(x)$*, ossia la funzione*

$$f : X \to Y, \qquad f(x) = P_{F(x)}(0) \quad \forall x \in X,$$

è una selezione continua di F*. Essa viene chiamata* selezione minimale *di* F*.*

Dimostrazione Poiché F è continua, F è sia semicontinua superiormente che inferiormente; in particolare, fissati $\varepsilon > 0$ e $x_0 \in X$, esiste un intorno U di x_0 in X con le seguenti proprietà:

$$F(x) \subseteq F(x_0) + \varepsilon B(0,1) \qquad \forall x \in U,$$

$$F(x) \cap B(f(x_0), \varepsilon) \neq \emptyset \qquad \forall x \in U,$$

ove $f(x) = P_{F(x)}(0)$.
Sia dunque $x \in U$: per ogni $y \in F(x)$, scelto $y_0 \in F(x_0)$ tale che $\|y - y_0\|_Y < \varepsilon$, dalla minimalità di $\|f(x_0)\|_Y$ si deduce

$$\|y\|_Y \geq \|y_0\|_Y - \varepsilon \geq \|f(x_0)\|_Y - \varepsilon,$$

e dunque, ancora per minimalità,

$$\|f(x)\|_Y \geq \|f(x_0)\|_Y - \varepsilon \qquad \forall x \in U;$$

dalla seconda relazione si ricava invece, scelto $y \in F(x) \cap B(f(x_0), \varepsilon)$,

$$\|f(x)\|_Y \leq \|y\|_Y \leq \|f(x_0)\|_Y + \varepsilon \qquad \forall x \in U.$$

Ciò prova che per ogni $x \in U$ il vettore $f(x)$ appartiene alla corona S_ε di centro 0 e raggi $\|f(x_0)\|_Y \pm \varepsilon$; ne segue

$$f(x) \in S_\varepsilon \cap [F(x_0) + \varepsilon B(0,1)] \qquad \forall x \in U.$$

D'altra parte, per ogni $y \in S_\varepsilon \cap [F(x_0) + \varepsilon B(0,1)] \cap F(x)$, con $x \in U$, risulta, posto $y = y_0 + \varepsilon z$ ove $y_0 \in F(x_0)$ e $z \in B(0,1)$:

$$\begin{aligned}
\|f(x_0) - y\|_Y^2 &= \|f(x_0)\|_Y^2 + \|y\|_Y^2 - 2\langle f(x_0), y\rangle_Y \leq \\
&\leq \|f(x_0)\|_Y^2 + (\|f(x_0)\|_Y + \varepsilon)^2 - 2\langle f(x_0), y_0 + \varepsilon z\rangle_Y = \\
&= 2\|f(x_0)\|_Y^2 + \varepsilon[2\|f(x_0)\|_Y - 2\langle f(x_0), z\rangle_Y + \varepsilon] - 2\langle f(x_0), y_0\rangle_Y \leq \\
&\leq 2\langle f(x_0), f(x_0) - y_0\rangle_Y + 4\varepsilon\|f(x_0)\|_Y + \varepsilon^2.
\end{aligned}$$

Dalla disequazione variazionale di cui, per definizione, $f(x_0)$ è soluzione, concludiamo che si ha

$$\|f(x_0) - y\|_Y^2 \le 4\varepsilon\|f(x_0)\|_Y + \varepsilon^2,$$

ed in particolare, scelto $y = f(x)$ con $x \in U$, si ottiene

$$\|f(x_0) - f(x)\|_Y < c\varepsilon \qquad \forall x \in U.$$

Ciò prova che $x \mapsto f(x)$ è continua; ovviamente, poi, $f(x) = P_{F(x)}(0) \in F(x)$, e quindi f è una selezione di F. $\square$

Esercizi 7.5

1. Sia X uno spazio topologico paracompatto, sia Y uno spazio di Banach e sia $F : X \to \mathcal{P}(Y)$ una multifunzione semicontinua inferiormente, tale che $F(x)$ sia non vuoto, chiuso e convesso per ogni $x \in X$. Si provi che per ogni $x_0 \in X$ e per ogni $y_0 \in F(x_0)$ esiste una selezione continua $f_0 : X \to Y$ di F, tale che $f_0(x_0) = y_0$.

2. Nelle ipotesi dell'esercizio precedente, sia $\Phi : X \to \mathcal{P}(Y)$ un'altra multifunzione tale che $F(x) \cap \Phi(x) \neq \emptyset$ per ogni $x \in X$; supponiamo inoltre che il suo grafico $G(\Phi)$ sia aperto in $X \times Y$. Si provi che per ogni $x_0 \in X$ e per ogni $y_0 \in F(x_0) \cap \Phi(x_0)$ esistono un intorno U di x_0 in X ed una funzione continua $f : U \to Y$, tali che $f(x_0) = y_0$ e $f(x) \in F(x) \cap \Phi(x)$ per ogni $x \in U$.

3. Nelle ipotesi dell'esercizio precedente, si supponga inoltre che $F(x) \cap \Phi(x)$ sia convesso per ogni $x \in X$. Si costruisca una selezione continua di $F \cap \Phi$.
[**Traccia:** utilizzando l'esercizio 7.5.2, per ogni $x \in X$ e $z \in F(x)$ si trovino un intorno U_x di x in X ed una funzione continua $f_x : U_x \to Y$ tali che $f_x(x) = z$ e $f_x(v) \in F(v) \cap \Phi(v)$ per ogni $v \in U_x$; si costruisca un raffinamento localmente finito $\{V_i\} \subseteq \{U_{x_i}\}$ di $\{U_x\}_{x \in X}$, si fissi una partizione dell'unità $\{g_i\}$ ad esso associata e si provi che $f(v) = \sum_i g_i(v) f_{x_i}(v)$ è la selezione cercata.]

4. Sia B la palla unitaria chiusa di $\mathbb{R}^2$. Per ogni $x \in B \setminus \{0\}$ siano ϑ_x, ρ_x definiti dalla relazione $x = (\rho_x \cos\vartheta_x, \rho_x \sin\vartheta_x)$ e si definisca la

multifunzione $F : B \to \mathcal{P}(\partial B)$ nel modo seguente:

$$F(x) = \begin{cases} \partial B & \text{se } x = 0, \\ \{(\cos\omega, \sin\omega) : |\omega - \vartheta_x - \pi| \le \pi(1 - \rho_x)]\} & \text{se } x \neq 0. \end{cases}$$

Provare che:

(i) $F : B \to \mathcal{P}(\partial B)$ è continua;

(ii) $F(x)$ è compatto per ogni $x \in B$;

(iii) F non ha alcuna selezione continua.

[**Traccia:** per (iii), ragionare per assurdo, utilizzando il teorema di Brouwer ed il fatto che $F(x) = \{-x\}$ per ogni $x \in \partial B$.]

7.6 Inclusioni differenziali

Consideriamo una multifunzione F definita su un aperto $U \subseteq \mathbb{R}^{N+1}$, a valori in $\mathcal{P}(\mathbb{R}^N)$. Vogliamo determinare una funzione $x : \mathbb{R} \to \mathbb{R}^N$, con derivata continua, che sia soluzione dell'*inclusione differenziale*

$$\begin{cases} x'(t) \in F(t, x(t)), & t \in \mathbb{R}, \\ x(t_0) = x_0, \end{cases}$$

ove (t_0, x_0) è un fissato punto di U. Il senso della questione è il seguente: la multifunzione F associa agli stati $x(t)$ di un sistema l'insieme delle sue possibili velocità. Si modellano in questo modo, attraverso le inclusioni differenziali, quei sistemi dinamici nei quali la complessità o la grandezza è tale da non consentire una completa descrizione dei fenomeni e delle interazioni presenti, generando dunque un certo grado di indeterminazione. Svariati esempi di sistemi di questo tipo compaiono in biologia, in economia e nelle scienze sociali.
Una semplice applicazione del teorema di Michael (teorema 7.5.5) porta al seguente risultato:

Teorema 7.6.1 *Sia $U \subseteq \mathbb{R}^{N+1}$ un aperto, sia $F : U \to \mathcal{P}(\mathbb{R}^N)$ una multifunzione semicontinua inferiormente, tale che $F(t, x)$ sia non vuoto, chiuso e convesso per ogni $(t, x) \in U$. Fissato $(t_0, x_0) \in U$, valgono i seguenti fatti:*

(i) *esiste una funzione x di classe C^1, definita in un intorno J di t_0, a valori in $\mathbb{R}^N$, tale che*

$$\begin{cases} x'(t) \in F(t, x(t)), & t \in J, \\ x(t_0) = x_0; \end{cases}$$

(ii) *esistono un intervallo aperto $I =]a, b[$ contenente t_0 ed una funzione $x \in C^1(I)$ che risolve l'inclusione differenziale sopra scritta nell'intervallo I, la quale è* massimale*: ciò significa che si ha $a = -\infty$, oppure $a \in \mathbb{R}$ con $(t, x(t)) \to \partial U$ per $t \to a^+$; un enunciato analogo vale per l'estremo b.*

Dimostrazione Per il teorema 7.5.5, esiste una selezione continua f di F. Allora si può considerare il problema di Cauchy

$$\begin{cases} x'(t) = f(t, x(t)), & t \in J, \\ x(t_0) = x_0; \end{cases}$$

questo problema di Cauchy ha soluzione in virtù del teorema di esistenza di Peano, valido quando f è continua nella coppia (t, x). Ogni soluzione di questo problema è una soluzione dell'inclusione differenziale, e verifica la (ii) se opportunamente prolungata. $\square$

Esercizi 7.6

1. Sia $F : \mathbb{R}^2 \to \mathcal{P}(\mathbb{R})$ definita da $F(t, x) = \{-1, 1\}$. Determinare le soluzioni dell'inclusione differenziale
$$x'(t) \in F(t, x(t)), \qquad x(0) = 0.$$

2. Sia $F : \mathbb{R}^2 \to \mathcal{P}(\mathbb{R})$ definita da $F(t, x) = [-|t|, |t|]$. Determinare le soluzioni dell'inclusione differenziale
$$x'(t) \in F(t, x(t)), \qquad x(0) = 0.$$

7.7 Multifunzioni monotone

La nozione di monotonia, per le multifunzioni così come per gli operatori, è molto utile nelle applicazioni: è una proprietà che in un certo senso sostituisce

la continuità o la semicontinuità.
Cominciamo col definire la monotonia per insiemi. In questo paragrafo e nei successivi faremo uso della notazione $\langle \varphi, u \rangle_X$, già usata nell'enunciato del teorema 3.1.11, in luogo di quella usuale $\varphi(u)$, per indicare l'azione del funzionale $\varphi \in X^*$ sull'elemento $u \in X$.

Definizione 7.7.1 *Sia X uno spazio di Banach. Se X è reale, un insieme $S \subseteq X \times X^*$ si dice* monotono *se si ha*

$$\langle \varphi - \psi, u - v \rangle_X \geq 0 \qquad \forall (u, \varphi), (v, \psi) \in S.$$

Se X è complesso, l'insieme S è monotono se

$$\mathrm{Re}\, \langle \varphi - \psi, u - v \rangle_X \geq 0 \qquad \forall (u, \varphi), (v, \psi) \in S.$$

Un insieme $S \subseteq X \times X^*$ si dice massimale monotono *se è monotono e se non esiste alcun insieme monotono $S' \subseteq X \times X^*$ tale che $S \subset S'$.*

Ad esempio, se $X = \mathbb{R}$ (cosicché $X^* = \mathbb{R}$) un insieme $S \subseteq \mathbb{R}^2$ è monotono se $(y - y')(x - x') \geq 0$ per ogni $(x, y), (x', y') \in S$, ovvero se vale l'implicazione

$$(x, y), (x', y') \in S, \quad x \leq x' \qquad \Longrightarrow \qquad y \leq y'.$$

In particolare, in $\mathbb{R}^2$ il segmento di estremi $(0, 0)$ e $(1, 1)$ è un insieme monotono ma non massimale monotono; lo stesso segmento, unito alle semirette verticali $\{0\} \times]-\infty, 0]$ e $\{1\} \times [1, \infty[$, è un insieme massimale monotono, mentre il segmento di estremi $(0, 1)$ e $(1, 0)$ non è un insieme monotono.
Passiamo ora a definire la monotonia per multifunzioni.

Definizione 7.7.2 *Sia X uno spazio di Banach e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione. Diciamo che A è* monotona *se il suo grafico $G(A)$ è un sottoinsieme monotono di $X \times X^*$. Diciamo che A è* massimale monotona *se $G(A)$ è massimale monotono in $X \times X^*$.*

Osservazione 7.7.3 È facile verificare che se X è uno spazio di Banach una multifunzione $A : X \to \mathcal{P}(X^*)$ è massimale monotona se e solo se è monotona e

$$(u, \varphi) \in X \times X^*, \ \mathrm{Re}\, \langle \varphi - \psi, u - v \rangle_X \geq 0 \ \forall (v, \psi) \in G(A) \ \Longrightarrow \ (u, \varphi) \in G(A).$$

È chiaro che la definizione di monotonia per multifunzioni si riduce a quella usuale (vedere la proposizione 5.2.5) nel caso in cui la multifunzione sia ad un sol valore, cioè sia una funzione. In particolare, un operatore monotono $A : D(A) \subseteq X \to X^*$ è massimale monotono se e solo se

$$\begin{cases} (u, \varphi) \in X \times X^*, \\ \operatorname{Re} \langle \varphi - A(v), u - v \rangle_X \geq 0 \quad \forall v \in D(A) \end{cases} \implies \quad u \in D(A) \text{ e } A(u) = \varphi.$$

Ad esempio, se $X = X^* = \mathbb{R}$ ogni funzione crescente $f : \mathbb{R} \to \mathbb{R}$ è monotona, e se è anche continua allora è massimale monotona (esercizio 7.7.2). Se invece la funzione crescente f ha un salto nel (solo) punto $x_0 \in \mathbb{R}$, risulta massimale monotona la multifunzione $F : \mathbb{R} \to \mathcal{P}(\mathbb{R}^2)$ definita da

$$F(x) = \begin{cases} \{f(x)\} & \text{se } x \neq x_0, \\ [f(x_0^-), f(x_0^+)] & \text{se } x = x_0. \end{cases}$$

Vediamo alcune proprietà delle multifunzioni massimali monotone, la cui fondamentale importanza applicativa sarà chiara nel seguito.

Proposizione 7.7.4 *Sia X uno spazio di Banach e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione. Detta $J : X \to X^{**}$ l'immersione canonica, la multifunzione A è monotona, oppure massimale monotona, se e solo se $J \circ A^{-1} : X^* \to \mathcal{P}(X^{**})$ è monotona, oppure massimale monotona.*

Dimostrazione Per definizione di grafico e di multifunzione inversa si ha

$$\begin{aligned} G(J \circ A^{-1}) &= \{(\varphi, \alpha) : \alpha \in (J \circ A^{-1})(\varphi)\} = \{(\varphi, J_x) : x \in A^{-1}(\varphi)\} = \\ &= \{(\varphi, J_x) : (\varphi, x) \in G(A^{-1})\} = \{(\varphi, J_x) : (x, \varphi) \in G(A)\}; \end{aligned}$$

è facile allora riconoscere che $G(A)$ è monotono in $X \times X^*$ se e solo se lo è $G(J \circ A^{-1})$ in $X^* \times X^{**}$.
Sia poi $G(J \circ A^{-1})$ massimale monotono e sia, per assurdo, S un insieme monotono tale che $G(A) \subset S \subseteq X \times X^*$; allora è immediato verificare che l'insieme

$$S' = \{(\varphi, J_x) : (x, \varphi) \in S\}$$

è un insieme monotono tale che $G(J \circ A^{-1}) \subset S' \subseteq X^* \times X^{**}$, contro l'ipotesi. Il viceversa è analogo. $\square$

Proposizione 7.7.5 *Sia X uno spazio di Banach e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione massimale monotona. Allora per ogni $u \in X$ l'insieme $A(u)$ è convesso e chiuso per la topologia debole* di X^*.*

Dimostrazione Supporremo X reale, ma quando X è complesso il calcolo è del tutto analogo. Sia $u \in X$ e siano $\varphi, \psi \in A(u)$. Se $t \in]0,1[$, poniamo $\eta_t = (1-t)\varphi + t\psi$ e mostriamo che $\eta_t \in A(u)$. Per ogni $(v, \zeta) \in G(A)$ si ha

$$\langle \eta_t - \zeta, u - v\rangle_X = (1-t)\langle \varphi - \zeta, u - v\rangle_X + t\langle \psi - \zeta, u - v\rangle_X \geq 0,$$

dunque $(u, \eta_t) \in G(A)$ per l'osservazione 7.7.3; pertanto $\eta_t \in A(u)$ e $A(u)$ è convesso.
Proviamo che $A(u)$ è debolmente* chiuso in X^*. Sia $\varphi \notin A(u)$: per definizione si ha $(u, \varphi) \notin G(A)$, e per l'osservazione 7.7.3 deve esistere $(v, \psi) \in G(A)$ tale che

$$-\varepsilon = \langle \varphi - \psi, u - v\rangle_X < 0.$$

Poniamo

$$U = \{\eta \in X^* : \ |\langle \eta - \varphi, u - v\rangle_X| < \varepsilon\}.$$

L'insieme U è aperto nella topologia debole* di X^*, e se $\eta \in U$ si ha

$$\langle \eta - \psi, u - v\rangle_X = \langle \eta - \varphi, u - v\rangle + \langle \varphi - \psi, u - v\rangle_X < \varepsilon - \varepsilon = 0.$$

Dunque $(u, \eta) \notin G(A)$ per ogni $\eta \in U$, ossia risulta $U \cap A(u) = \emptyset$. Ciò prova che $A(u)^c$ è w^*-aperto, il che significa che $A(u)$ è debolmente* chiuso. □

Consideriamo ora multifunzioni ad un sol valore: questo corrisponde ad operatori $A : D(A) \subseteq X \to X^*$, ove il dominio di A è l'insieme degli $u \in X$ per i quali la multifunzione corrispondente ad A assume "valore" $A(u) \neq \emptyset$.

Proposizione 7.7.6 *Sia X uno spazio di Banach riflessivo e sia $A : D(A) \subseteq X \to X^*$ un operatore monotono ed* emicontinuo*, ossia continuo sui sottospazi mono-dimensionali rispetto alla topologia debole* di X^*. Se $D(A) = X$, allora A è massimale monotono.*

L'ipotesi di emicontinuità significa che per ogni $u, v \in D(A)$ e $w \in X$ la funzione reale $t \mapsto \langle A(u + tv), w\rangle_X$ è continua.

Dimostrazione Sia $(u, \varphi) \in X \times X^*$ tale che

$$\mathrm{Re}\,\langle \varphi - A(v), u - v\rangle_X \geq 0 \qquad \forall v \in X.$$

In virtù dell'osservazione 7.7.3, basterà provare che $A(u) = \varphi$. Scegliamo $v = u - tw$, con $t > 0$ e $w \in X$: si ha

$$\mathrm{Re}\,\langle \varphi - A(u - tw), tw\rangle_X \geq 0 \qquad \forall w \in X,\ \forall t > 0,$$

e dividendo per t

$$\operatorname{Re}\langle\varphi - A(u - tw), w\rangle_X \geq 0 \qquad \forall w \in X;$$

per $t \to 0^+$ si deduce, grazie all'emicontinuità di A, che

$$\operatorname{Re}\langle\varphi - A(u), w\rangle_X \geq 0 \qquad \forall w \in X,$$

e scrivendo $-w$ in luogo di w si conclude che

$$\operatorname{Re}\langle\varphi - A(u), w\rangle_X = 0 \qquad \forall w \in X.$$

In modo analogo, scegliendo $v = u - itw$, si dimostra che

$$\operatorname{Im}\langle\varphi - A(u), w\rangle_X = 0 \qquad \forall w \in X.$$

Ne segue che $\varphi = A(u)$. $\square$

Teorema 7.7.7 (di Minty) *Sia X uno spazio di Hilbert e sia $A : D(A) \subseteq X \to X$ un operatore monotono. Allora A è massimale monotono se e solo se l'operatore $I + A$ è surgettivo, ove $I : X \to X$ è l'identità.*

Questo risultato spiega l'importanza della proprietà di massimale monotonia: essa garantisce la surgettività dell'operatore $I + A$ e quindi la risolubilità di equazioni della forma $u + A(u) = g$, con A massimale monotono e g assegnata. Vedremo nel seguito che molti operatori che compaiono nelle applicazioni sono in effetti massimali monotoni.

Dimostrazione ($\Longrightarrow$) Proveremo questa implicazione più avanti, come corollario di un più generale teorema di Browder (corollario 7.9.7).

($\Longleftarrow$) Sia $R(I + A) = X$. Vale il seguente

Lemma 7.7.8 *Nelle ipotesi del teorema 7.7.7, sia $I + A$ surgettivo; allora per ogni $\mu > 0$ l'operatore $I + \mu A$ è bigettivo e l'inverso $(I + \mu A)^{-1}$ è un operatore* non espansivo*, ossia lipschitziano con costante uguale a* 1.

Dimostrazione Se $u, v \in D(A)$ si ha, grazie alla monotonia di A,

$$\begin{aligned}
&\|(I + \mu A)(u) - (I + \mu A)(v)\|_X^2 = \\
&\qquad = \|u - v\|_X^2 + \mu^2\|A(u) - A(v)\|_X^2 + 2\operatorname{Re}\mu\langle A(u) - A(v), u - v\rangle_X \geq \\
&\qquad \geq \|u - v\|_X^2\,;
\end{aligned}$$

quindi $I + \mu A$ è iniettivo per ogni $\mu > 0$. In particolare, $I + A$ è bigettivo e $(I + A)^{-1}$ è non espansivo. Sia ora $\mu > \frac{1}{2}$: fissato $w \in X$, risolvere rispetto a u l'equazione $u + \mu A(u) = w$ equivale a trovare $u \in D(A)$ tale che $\mu A(u) = w - u$, ossia $u + A(u) = u + \frac{1}{\mu}(w - u)$, ed infine $(I + A)(u) = u + \frac{1}{\mu}(w - u)$. Quindi si cerca $u \in D(A)$ tale che

$$u = (I + A)^{-1}\left[u + \frac{1}{\mu}(w - u)\right] = L_w(u).$$

Dobbiamo dunque trovare un punto fisso $u \in D(A)$ per l'operatore L_w, ove $w \in X$ è fissato. Dato che $(I + A)^{-1}$ è non espansivo, risulta per $u, v \in X$

$$\begin{aligned}
&\|L_w(u) - L_w(v)\|_X = \\
&\quad = \left\|(I + A)^{-1}\left[u + \frac{1}{\mu}(w - u)\right] - (I + A)^{-1}\left[v + \frac{1}{\mu}(w - v)\right]\right\|_X \le \\
&\quad \le \left\|\left[u + \frac{1}{\mu}(w - u)\right] - \left[v + \frac{1}{\mu}(w - v)\right]\right\|_X = \\
&\quad = \left|1 - \frac{1}{\mu}\right| \|u - v\|_X\,;
\end{aligned}$$

poiché $|1 - \frac{1}{\mu}| < 1$, l'operatore L_w è una contrazione nello spazio di Hilbert X, con valori in $D(A)$. Per il teorema delle contrazioni esiste un unico punto fisso $u \in D(A)$. Pertanto, per ogni $w \in X$ vi è un unico $u \in D(A)$ tale che $u + \mu A(u) = w$, ossia l'operatore $I + \mu A$ è bigettivo per ogni $\mu > \frac{1}{2}$.
Se ripetiamo questa argomentazione con $I + \lambda A$ al posto di $I + A$, ove $\lambda > \frac{1}{2}$, otteniamo la bigettività di $I + \mu A$ per ogni $\mu > \frac{\lambda}{2}$, e quindi per ogni $\mu > \frac{1}{4}$. Iterando questo procedimento, si ha la tesi del lemma per ogni $\mu > 2^{-n}$, con n arbitrario, e dunque il lemma è provato. □

Torniamo alla dimostrazione del teorema di Minty. Sia $(u, w) \in X \times X$ tale che

$$\operatorname{Re}\langle w - A(v), u - v\rangle_X \ge 0 \qquad \forall v \in D(A);$$

poniamo $v_t = (I + A)^{-1}(u + w + tz)$, ove $t > 0$ e $z \in X$. Poiché

$$\operatorname{Re}\langle w - A(v_t), u - v_t\rangle_X \ge 0, \qquad \operatorname{Re}\langle u - v_t, u - v_t\rangle_X = \|u - v_t\|_X^2 \ge 0,$$

sommando si ha

$$\operatorname{Re}\langle -tz, u - v_t\rangle_X = \operatorname{Re}\langle w + u - A(v_t) - v_t, u - v_t\rangle_X \ge 0,$$

da cui, dividendo per t,

$$\mathrm{Re}\,\langle z, u - (I + A)^{-1}(u + w + tz)\rangle_X = \langle z, u - v_t\rangle_X \le 0,$$

e per $t \to 0^+$

$$\mathrm{Re}\,\langle z, u - (I + A)^{-1}(u + w)\rangle_X \le 0 \qquad \forall z \in X.$$

Similmente, ponendo stavolta $v_t = (I + A)^{-1}(u + w + itz)$, si ottiene

$$\mathrm{Im}\,\langle z, u - (I + A)^{-1}(u + w)\rangle_X \le 0 \qquad \forall z \in X.$$

Per l'arbitrarietà di z, ciò implica $u = (I + A)^{-1}(u + w)$, ovvero $u + A(u) = u + w$, e dunque $A(u) = w$. Per l'osservazione 7.7.3 si conclude che l'operatore A è massimale monotono. $\square$

Esercizi 7.7

1. Sia X uno spazio di Banach e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione monotona. Provare che A^{-1} e λA (con $\lambda > 0$) sono monotone; provare anche che sono monotone la *chiusura* di A, definita da

$$\overline{A}(x) = \text{chiusura debole di } A(x) \text{ in } X^*$$

   e l'*inviluppo convesso chiuso* di A, cioè

$$\overline{\mathrm{co}(A)}(x) = \overline{\mathrm{co}(A(x))}.$$

2. Provare che ogni funzione $f : \mathbb{R} \to \mathbb{R}$ continua e crescente è un'applicazione massimale monotona.

3. Sia $f : \mathbb{R} \to \mathbb{R}$ crescente. Si provi che la multifunzione $\overline{f}$, definita da $\overline{f}(x) = [f(x^-), f(x^+)]$ per ogni $x \in \mathbb{R}$, è massimale monotona. Si provi anche che *ogni* multifunzione massimale monotona $g : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ è di questo tipo.

4. Sia X uno spazio di Hilbert. Se $J : X \to X$ è una contrazione, si verifichi che $I + J$ è monotona, ove I è l'identità su X.

5. Sia X uno spazio di Hilbert e sia $A : X \to X$ un operatore monotono. Se $\mathcal{A} : L^2(a, b; X) \to L^2(a, b; X)$ è definito da

$$[\mathcal{A}(u)](\cdot) = A(u(\cdot)) \qquad \forall u \in L^2(a, b; X),$$

si provi che $\mathcal{A}$ è monotono.

6. Sia X uno spazio di Banach riflessivo, e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione monotona ed *emicontinua*, ossia tale che per ogni $u, v \in X$ la multifunzione $t \mapsto A(tu + (1-t)v)$, da $[0, 1]$ in $\mathcal{P}(X^*)$, sia semicontinua superiormente rispetto alla topologia debole di X. Supponiamo anche che $A(u)$ sia convesso e debolmente chiuso in X^* per ogni $u \in X$. Si provi che se $(u, \varphi) \in D(A) \times X^*$ e se

$$\mathrm{Re}\,\langle \varphi - \psi, u - v\rangle_X \geq 0 \qquad \forall (v, \psi) \in D(A),$$

allora $(u, \varphi) \in G(A)$.
[**Traccia:** supponendo $\varphi \notin A(u)$, si faccia uso del teorema di Hahn-Banach (teorema 1.7.1).]

7. Sia X uno spazio di Banach riflessivo, e sia $A : X \to \mathcal{P}(X^*)$ una multifunzione monotona, emicontinua e tale che $A(u)$ sia non vuoto, convesso e chiuso in X^* per ogni $u \in X$. Si provi che A è massimale monotono.
[**Traccia:** utilizzare l'esercizio precedente.]

7.8 Alcune multifunzioni massimali monotone

Raggruppiamo in questo paragrafo alcuni esempi significativi di multifunzioni o di operatori massimali monotoni.

Esempio 7.8.1 (applicazioni di dualità) Sia X uno spazio di Banach: consideriamo l'applicazione di dualità introdotta nell'esempio 7.1.6 (2), che è definita per ogni $x \in X$ da

$$F(x) = \{\varphi \in X^* : \ \langle \varphi, x\rangle_X = \|x\|_X^2 = \|\varphi\|_{X^*}^2\}.$$

Le proprietà elementari di questa multifunzione sono raggruppate nel seguente

Lemma 7.8.2 *Sia X uno spazio di Banach (reale o complesso). L'applicazione di dualità F sopra definita verifica:*

(i) *$F(x)$ è un convesso chiuso non vuoto di X^* per ogni $x \in X$;*

(ii) *si ha $F(\alpha x) = \overline{\alpha} F(x)$ per ogni $\alpha \in \mathbb{C}$ e per ogni $x \in X$;*

(iii) *F è monotona;*

(iv) *se X è riflessivo, allora $\bigcup_{x \in X} F(x) = X^*$;*

(v) *se X è uno spazio di Hilbert, allora $F(x) = \{x\}$ per ogni $x \in X$.*

Dimostrazione (i) $F(x)$ è ovviamente chiuso, ed è non vuoto in virtù del teorema di Hahn-Banach. Proviamo la convessità: se $\varphi, \psi \in F(x)$ e $\lambda \in]0,1[$, allora

$$\langle (1-\lambda)\varphi + \lambda\psi, x\rangle_X = (1-\lambda)\langle \varphi, x\rangle_X + \lambda\langle \psi, x\rangle_X = (1-\lambda)\|x\|_X^2 + \lambda\|x\|_X^2 = \|x\|_X^2$$

da cui

$$\|(1-\lambda)\varphi + \lambda\psi\|_{X^*} \geq \|x\|_X \,;$$

d'altronde

$$\|(1-\lambda)\varphi + \lambda\psi\|_{X^*} \leq (1-\lambda)\|\varphi\|_{X^*} + \lambda\|\psi\|_{X^*} = \|x\|_X \,,$$

e dunque $(1-\lambda)\varphi + \lambda\psi \in F(x)$.
(ii) La tesi è conseguenza del fatto, di immediata verifica, che $\varphi \in F(\alpha x)$ se e solo se $(1/\overline{\alpha})\varphi \in F(x)$.
(iii) Dobbiamo provare che

$$\mathrm{Re}\,\langle \varphi - \psi, x - y\rangle_X \geq 0 \quad \forall x, y \in X, \ \forall \varphi \in F(x), \ \forall \psi \in F(y).$$

In effetti, poiché $\langle \varphi, x\rangle_X = \|x\|_X^2 = \|\varphi\|_{X^*}^2$ e $\langle \psi, y\rangle_X = \|y\|_X^2 = \|\psi\|_{X^*}^2$ si ha

$$\begin{aligned} \mathrm{Re}\,\langle \varphi - \psi, x - y\rangle_X &= \mathrm{Re}\,(\langle \varphi, x\rangle_X - \langle \varphi, y\rangle_X - \langle \psi, x\rangle_X + \langle \psi, y\rangle_X) \geq \\ &\geq \|x\|_X^2 - 2\|x\|_X\|y\|_X + \|y\|_X^2 = (\|x\|_X - \|y\|_X)^2 \geq 0. \end{aligned}$$

(iv) Sia X riflessivo: fissato $\varphi \in X^*$, per il teorema di Hahn-Banach esiste un elemento $\alpha \in X^{**}$, che sarà della forma $\alpha = J_x$ con $x \in X$, tale che

$$\langle \varphi, x\rangle_X = \langle J_x, \varphi\rangle_{X^*} = \|\varphi\|_X \,, \qquad \|x\|_X = \|J_x\|_{X^{**}} = 1.$$

Ne segue che l'elemento $y = tx$, ove $t = \|\varphi\|_{X^*}$, verifica

$$\langle \varphi, y\rangle_X = \|\varphi\|^2_{X^{**}} = \|y\|^2_X \,,$$

cioè $\varphi \in F(y)$.

(v) Sia X uno spazio di Hilbert. Se $x = 0$, allora per definizione $F(0) = \{0\}$. Se invece $x \neq 0$ e $z, w \in F(x) \subseteq X^* = X$, si ha

$$\langle z, x\rangle_X = \|z\|^2_X = \|x\|^2_X = \|w\|^2_X = \langle w, x\rangle_X \neq 0.$$

Quindi, essendo $2\|x\|^2_X = \langle z{+}w, x\rangle_X \leq \|z{+}w\|_X \|x\|_X$, si deduce $\|z{+}w\|_X \geq 2\|x\|_X = \|z\|_X + \|w\|_X$, ovvero $\|z + w\|_X = \|z\|_X + \|w\|_X$. Ciò implica che esiste $\alpha > 0$ tale che $\alpha w = z$; dalla relazione $\alpha\langle w, x\rangle_X = \langle z, x\rangle_H = \langle w, x\rangle_H \neq 0$ segue $\alpha = 1$ e quindi $w = z$. Pertanto $F(x)$ ha un solo elemento. Dato che,ovviamente, $x \in F(x)$, si conclude che $F(x) = \{x\}$. $\square$

Proposizione 7.8.3 *Sia X uno spazio di Banach reale o complesso. L'applicazione di dualità F sopra definita è massimale monotona.*

Dimostrazione Sia $(x, \varphi) \in X \times X^*$ tale che

$$\mathrm{Re}\,\langle \varphi - \psi, x - y\rangle_X \geq 0 \qquad \forall y \in X,\ \forall \psi \in F(y),$$

e proviamo che $\varphi \in F(x)$; dall'osservazione 7.7.3 seguirà che F è massimale monotona.

Cominciamo col provare che $\|\varphi\|_{X^*} \leq \|x\|_X$. Sia $\{v_n\}_{n\in\mathbb{N}}$ una successione di elementi di X tale che $\|v_n\|_X = 1$ e $\langle \varphi, v_n\rangle_X \to \|\varphi\|_{X^*}$ per $n \to \infty$ (ed in particolare $\mathrm{Im}\langle \varphi, v_n\rangle_X \to 0$). Scegliamo nella disuguaglianza precedente $y = x + tv_n$, con $t > 0$: poiché $\|\psi\|_{X^*} = \|x + tv_n\|_X$ e $\|v_n\|_X = 1$ otteniamo

$$\begin{aligned} -t\,\mathrm{Re}\,\langle \varphi, v_n\rangle_X &= \mathrm{Re}\,\langle \varphi, x - y\rangle_X \geq \mathrm{Re}\,\langle \psi, x - y\rangle_X = \\ &= \mathrm{Re}\,\langle \psi, -tv_n\rangle_X \geq -t\,\|x + tv_n\|_X \,, \end{aligned}$$

da cui, dividendo per t,

$$\mathrm{Re}\,\langle \varphi, v_n\rangle_X \leq \|x + tv_n\|_X \qquad \forall t > 0.$$

Se $t \to 0^+$ ricaviamo $\mathrm{Re}\,\langle \varphi, v_n\rangle_X \leq \|x\|_X$ per ogni $n \in \mathbb{N}$; infine per $n \to \infty$ si conclude che $\|\varphi\|_{X^*} \leq \|x\|_X$.

Adesso mostriamo che $\mathrm{Re}\,\langle\varphi,x\rangle_X = \|x\|_X^2$ e che quindi $\|x\|_X \le \|\varphi\|_{X^*}$. Ciò implicherà $\varphi \in F(x)$ e dunque la tesi. Sappiamo che risulta

$$\mathrm{Re}\,\langle\varphi,x\rangle_X \ge \mathrm{Re}\,\langle\psi,x\rangle_X + \mathrm{Re}\,\langle\varphi,y\rangle_X - \|y\|_X^2 \quad \forall y \in X,\ \forall \psi \in F(y).$$

Scegliamo $y = \lambda x$, con $\lambda \in \mathbb{R}$: allora essendo $\psi \in F(\lambda x) = \lambda F(x)$, sarà $\psi = \lambda\eta$ con $\eta \in F(x)$. Quindi $\langle\psi,x\rangle_X = \lambda\langle\eta,x\rangle_X = \lambda\|x\|_X^2$ e pertanto

$$\mathrm{Re}\,\langle\varphi,x\rangle_X \ge \lambda\|x\|_X^2 + \lambda\mathrm{Re}\,\langle\varphi,x\rangle_X - \lambda^2\|x\|_X^2 \quad \forall \lambda \in \mathbb{R}.$$

Se $\lambda \ne 1$, dividendo per $1-\lambda$ otteniamo

$$\mathrm{Re}\,\langle\varphi,x\rangle_X \ge \lambda\|x\|_X^2 \quad \forall \lambda < 1, \qquad \mathrm{Re}\,\langle\varphi,x\rangle_X \le \lambda\|x\|_X^2 \quad \forall \lambda > 1.$$

Al limite rispettivamente per $\lambda \to 1^-$ e per $\lambda \to 1^+$ deduciamo $\mathrm{Re}\,\langle\varphi,x\rangle_X = \|x\|_X^2$, il che è quanto si voleva. $\square$

Esempio 7.8.4 (derivata seconda) Sia A l'operatore $u \mapsto -u''$, definito sullo spazio di Hilbert $X = L^2(a,b)$, con dominio

$$D(A) = \{u \in C^1[a,b] : \ u' \in AC[a,b],\ u'' \in L^2(a,b),\ u(a) = u(b) = 0\}$$

(ma funzionano altrettanto bene le condizioni $u'(a) = u'(b) = 0$, oppure $u(a) = u'(b) = 0$, oppure $u'(a) = u(b) = 0$).
Verifichiamo che A è monotono: se $u \in D(A)$, integrando per parti (il che è lecito, dato che u e u' sono assolutamente continue) si ha

$$\langle Au,u\rangle_X = \int_a^b (-u'')u\,dx = -[u'u]_a^b + \int_a^b |u'|^2\,dx = \int_a^b |u'|^2\,dx \ge 0,$$

ed essendo A lineare, la monotonia è provata.
Verifichiamo che A è massimale monotono: per il teorema di Minty, basta provare che $R(I+A) = L^2(a,b)$. Ed infatti, se $f \in L^2(a,b)$ il problema

$$\begin{cases} -u'' + u = f & \text{in } [a,b] \\ u(a) = u(b) = 0 \end{cases}$$

ha l'unica soluzione

$$u(x) = \frac{\int_a^b \sinh(b-s)\,f(s)\,ds}{\sinh(b-a)}\,\sinh(x-a) - \int_a^x \sinh(x-s)\,f(s)\,ds.$$

Esempio 7.8.5 (derivata prima) Sia H uno spazio di Hilbert, che identificheremo, al solito, con il suo duale H^*, e poniamo, per un fissato $T > 0$,

$$X = L^2(0, T; H)$$

(si veda il paragrafo 3.1). Esso è uno spazio di Hilbert con il prodotto scalare

$$\langle f, g \rangle_X = \int_a^b \langle f(t), g(t) \rangle_H \, dt \qquad \forall f, g \in X.$$

Nel lemma che segue vengono dimostrate alcune proprietà importanti per il seguito.

Lemma 7.8.6 *Sia H uno spazio di Hilbert. Fissato $T > 0$, sia $X = L^2(0, T; H)$ e sia Y lo spazio costituito da tutte le funzioni $u : [0, T] \to H$ appartenenti a X, tali che u è derivabile q.o. in $[0, T]$, $u' \in X$, e vale la formula*

$$u(t) - u(s) = \int_s^t u'(\tau) \, d\tau \qquad \forall s, t \in [0, T].$$

Allora:

(i) *Y è uno spazio di Hilbert con il prodotto scalare*

$$\langle u, v \rangle_Y = \langle u, v \rangle_X + \langle u', v' \rangle_X;$$

(ii) *$C^1([0, T], H)$ è denso in Y;*

(iii) *Y è incluso con continuità in $C([0, T], H)$;*

(iv) *vale la formula di integrazione per parti*

$$\begin{aligned} &\langle u(t), v(t) \rangle_H - \langle u(s), v(s) \rangle_H = \\ &\quad = \int_s^t [\langle u'(r), v(r) \rangle_H + \langle v'(r), u(r) \rangle)_H] \, dr \qquad \forall u, v \in Y. \end{aligned}$$

Dimostrazione (i) Anzitutto, lo spazio Y non si riduce a $\{0\}$, dato che contiene le funzioni costanti $f(t) \equiv v$, con $v \in H$. Sia $\{u_n\}$ una successione di Cauchy in Y: poiché X è completo, esistono $u, v \in X$ tali che

$$u_n \to u \text{ in } L^2(0, T; H), \qquad u'_n \to v \text{ in } L^2(0, T : H).$$

Per la continuità del prodotto scalare, per ogni $\varphi \in C_0^\infty(]0,T[,H)$ si ha

$$\frac{d}{dt}\langle u_n(t),\varphi(t)\rangle_H = \langle u_n'(t),\varphi(t)\rangle_H + \langle u_n(t),\varphi'(t)\rangle_H \qquad \text{q.o. in } [0,T],$$

da cui, integrando,

$$\int_0^T \langle u_n'(t),\varphi(t)\rangle_H\,dt = -\int_0^T \langle u_n(t),\varphi'(t)\rangle_H\,dt$$

e per $n\to\infty$

$$\int_0^T \langle v(t),\varphi(t)\rangle_H\,dt + \int_0^T \langle u(t),\varphi'(t)\rangle_H\,dt = 0 \qquad \forall\varphi\in C(]0,T[,H).$$

Utilizzando un'opportuna generalizzazione del lemma 4.2.3 (esercizio 7.8.1), si deduce che

$$\exists u'(t) = v(t) \qquad \text{q.o. in } [0,T],$$

e ciò prova che $u\in Y$ e che $u_n\to u$ in Y; quindi Y è completo.

(ii) Sia $u\in Y$; estendiamo la u fuori di $[0,T]$ in modo qualunque, ad esempio ponendola uguale a 0 fuori di $[0,T]$. Sia $\varphi\in C_0^\infty(\mathbb{R})$ una funzione non negativa, nulla fuori di $[-1,1]$ e tale che $\int_{-1}^1\varphi(t)\,dt=1$; poniamo per $\varepsilon>0$:

$$u_\varepsilon(t) = \int_{\mathbb{R}} u(t-\varepsilon\sigma)\varphi(\sigma)\,d\sigma = \frac{1}{\varepsilon}\int_{\mathbb{R}} u(s)\varphi\left(\frac{t-s}{\varepsilon}\right)ds, \qquad t\in[0,T].$$

Le funzioni u_ε sono in $C^1([0,T],H)$ con

$$\begin{aligned} u_\varepsilon'(t) &= \frac{1}{\varepsilon^2}\int_{\mathbb{R}} u(s)\varphi'\left(\frac{t-s}{\varepsilon}\right)ds = \\ &= \frac{1}{\varepsilon}\int_{\mathbb{R}} u'(s)\varphi\left(\frac{t-s}{\varepsilon}\right)ds = \int_{\mathbb{R}} u'(t-\varepsilon\sigma)\varphi(\sigma)\,d\sigma \qquad \forall t\in[0,T]. \end{aligned}$$

Si ha allora

$$\begin{aligned}
\|u_\varepsilon - u\|_X^2 &= \int_0^T \left\| \frac{1}{\varepsilon} \int_{\mathbb{R}} [u(s) - u(t)] \, \varphi\left(\frac{t-s}{\varepsilon}\right) ds \right\|_H^2 dt \leq \\
&\leq \int_0^T \left\{ \int_{\mathbb{R}} \|u(s) - u(t)\|_H \left[\frac{1}{\varepsilon} \varphi\left(\frac{t-s}{\varepsilon}\right)\right]^{\frac{1}{2}+\frac{1}{2}} ds \right\}^2 dt \leq \\
&\leq \int_0^T \left[\int_{\mathbb{R}} \|u(s) - u(t)\|_H^2 \frac{1}{\varepsilon} \varphi\left(\frac{t-s}{\varepsilon}\right) ds \right] \left[\int_{\mathbb{R}} \frac{1}{\varepsilon} \varphi\left(\frac{t-s}{\varepsilon}\right) ds \right] dt = \\
&= \int_0^T \int_{\mathbb{R}} \|u(t - \varepsilon\sigma) - u(t)\|_H^2 \, \varphi(\sigma) \, d\sigma dt = \\
&= \int_{-1}^1 \int_0^T \|u(t - \varepsilon\sigma) - u(t)\|_H^2 \, dt \, \varphi(\sigma) \, d\sigma,
\end{aligned}$$

e quando $\varepsilon \to 0^+$ risulta $u_\varepsilon \to u$ in virtù della continuità in L^2 delle traslazioni e del teorema di Lebesgue.
In modo del tutto analogo,

$$\|u'_\varepsilon - u'\|_H^2 \leq \int_{-1}^1 \int_0^T \|u'(t - \varepsilon\sigma) - u'(t)\|_H^2 \, dt \, \varphi(\sigma) \, d\sigma \;\to\; 0 \qquad \text{per } \varepsilon \to 0^+.$$

Ciò prova che $C^1([0, T], H)$ è denso in Y.
(iii) Sia $u \in C^1([0, T], H)$ e sia $t_0 \in [0, T]$ un punto di minimo per $\|u(\cdot)\|_H$, cosicché

$$\|u(t_0)\|_H \leq \frac{1}{T} \int_0^T \|u(s)\|_H \, ds.$$

Allora si ha per ogni $t \in [0, T]$

$$\begin{aligned}
\|u(t)\|_H^2 &= \|u(t_0)\|_H^2 + \int_{t_0}^t \frac{d}{ds} \|u(s)\|_H^2 \, ds = \\
&= \|u(t_0)\|_H^2 + 2 \int_{t_0}^t \langle u'(s), u(s) \rangle_H \, ds \leq \\
&\leq \left[\frac{1}{T} \int_0^T \|u(s)\|_H \, ds \right]^2 + 2\|u'\|_X \|u\|_X \leq \\
&\leq \frac{1}{T^2} \|u\|_X^2 + 2\|u'\|_X \|u\|_X \leq c_T \|u\|_Y^2,
\end{aligned}$$

da cui

$$\|u\|_{C([0,T],H)} \le c_T \|u\|_Y \qquad \forall u \in C^1([0,T],H).$$

Questa stima, in virtù della densità (già provata) di $C^1([0,T],H)$ in Y, si estende a tutte le funzioni $u \in Y$, mostrando così che Y è incluso in $C([0,T],H)$ e che tale inclusione è continua.

(iv) La formula di integrazione per parti vale quando $u, v \in C^1([0,T],H)$: infatti

$$\begin{aligned}&\langle u(t), v(t)\rangle_H - \langle u(s), v(s)\rangle_H = \\ &\quad = \int_s^t \frac{d}{dr}\langle u(r), v(r)\rangle_H \, dr = \int_s^t [\langle u'(r), v(r)\rangle_H + \langle u(r), v'(r)\langle_H] \, dr;\end{aligned}$$

dato che $C^1([0,T],H)$ è denso in Y, e che Y è immerso con continuità in $C([0,T],H)$, si ottiene che $C^1([0,T],H)$ è anche denso nello spazio Y munito della norma di $C([0,T],H)$. Quindi la formula precedente si estende per densità a tutte le funzioni $u, v \in Y$. Il lemma 7.8.6 è completamente dimostrato.
□

Dopo tutte queste premesse, definiamo finalmente il dominio dell'operatore derivata prima $\frac{d}{dx}$. Abbiamo due opzioni:

$$D(L_1) = \{u \in Y : u(0) = 0\}, \qquad L_1 u = u' \quad \forall u \in D(L_1);$$

$$D(L_2) = \{u \in Y : u(0) = u(T)\}, \qquad L_2 u = u' \quad \forall u \in D(L_2).$$

Proposizione 7.8.7 *Sotto le ipotesi del lemma 7.8.6, per $i = 1, 2$ gli operatori $L_i : D(L_i) \subset X \to X^*$ sopra definiti sono massimali monotoni.*

Dimostrazione Ovviamente, L_i è lineare; inoltre se $u \in D(L_i)$ si ha

$$\begin{aligned}\langle L_i u, u\rangle_X &= \int_0^T \langle u'(t), u(t)\rangle_H \, dt = \frac{1}{2}\int_0^T \frac{d}{dt}\|u(t)\|_H^2 \, dt = \\ &= \frac{1}{2}\left[\|u(T)\|_H^2 - \|u(0)\|_H^2\right] \ge 0,\end{aligned}$$

e ciò prova la monotonia di L_i, $i = 1, 2$.
Proviamo che L_i è massimale monotono: in virtù del teorema di Minty (teorema 7.7.7) basterà provare che l'equazione $u + L_i u = f$ ha soluzione per ogni $f \in L^2(0,T;H)$. Tale equazione equivale a uno dei due problemi di Cauchy

$$\begin{cases} u + u' = f & \text{in } [0,T] \\ u(0) = 0, \end{cases} \qquad \text{oppure} \qquad \begin{cases} u + u' = f & \text{in } [0,T] \\ u(0) = u(T). \end{cases}$$

Tali problemi hanno rispettivamente le soluzioni

$$u_1(t) = \int_0^t e^{-(t-s)} f(s)\, ds,$$

$$u_2(t) = \frac{e^{-t}}{1-e^{-T}} \int_0^T e^{-(T-s)} f(s)\, ds + \int_0^t e^{-(t-s)} f(s)\, ds,$$

il che prova la massimale monotonia di L_1 e L_2. $\square$

Esempio 7.8.8 (sottodifferenziale) Il sottodifferenziale di una funzione $f : X \to \overline{\mathbb{R}}$, ove X è uno spazio di Banach, è una multifunzione $\partial f : X \to \mathcal{P}(X^*)$. Se f è convessa, a valori in $]-\infty, +\infty]$, semicontinua inferiormente e propria, tale multifunzione è massimale monotona, come vedremo. Questo fatto è vero per qualunque spazio di Banach, ma per semplicità ci limiteremo a considerare il caso più facile in cui X è uno spazio di Hilbert.

Teorema 7.8.9 (di Rockafellar) *Sia X uno spazio di Hilbert e sia $f : X \to]-\infty, +\infty]$ una funzione convessa, semicontinua inferiormente e propria. Allora $\partial f : X \to \mathcal{P}(X^*)$ è massimale monotono ed inoltre il suo dominio verifica $\overline{D(\partial f)} = \overline{D(f)}$.*

Dimostrazione Come anticipato nell'osservazione 5.2.6, ∂f è monotono. Proviamo adesso che $R(I + \alpha\partial f) = X$ per ogni $\alpha > 0$. Questo non implica ancora la tesi: infatti il teorema di Minty (teorema 7.7.7) è applicabile solo alle funzioni; tuttavia questa proprietà ci servirà nel corso della dimostrazione.
Sia y un arbitrario punto di X: fissato $\alpha > 0$, consideriamo la funzione

$$g_{\alpha,y}(x) = f(x) + \frac{1}{2\alpha}\, \|x-y\|_X^2, \qquad x \in X,$$

già introdotta nel paragrafo 5.3. Per il lemma 5.3.1, essa ha un unico punto di minimo $x_0 \in X$, e si ha $\frac{1}{\alpha}(y - x_0) \in \partial f(x_0)$; ciò significa

$$y \in (I + \alpha\partial f)(x_0),$$

che è quanto si voleva.
Proviamo che ∂f è massimale monotono. Siano $u, w \in X$ tali che

$$\langle w - x, u - v\rangle_X \geq 0 \qquad \forall v \in D(\partial f),\ \forall x \in \partial f(v);$$

se proveremo che $u \in D(\partial f)$ e che $w \in \partial f(u)$, avremo dimostrato la tesi in virtù dell'osservazione 7.7.3.
Consideriamo, per $t > 0$ e $z \in X$ fissati, l'elemento $u + w + tz \in X$: poiché $R(I + \partial f) = X$, esiste $v_t \in D(\partial f)$ tale che

$$u + w + tz \in (I + \partial f)(v_t);$$

in altri termini possiamo scrivere

$$u + w + tz = v_t + \xi_t, \qquad \xi_t \in \partial f(v_t).$$

Ora osserviamo che, per come sono stati scelti u, w,

$$\begin{aligned} -\langle tz, u - v_t\rangle_X &= \langle u - v_t + w - \xi_t, u - v_t\rangle_X = \\ &= \|u - v_t\|_X^2 + \langle w - \xi_t, u - v_t\rangle_X \geq \|u - v_t\|_X^2 \geq 0, \end{aligned}$$

il che implica, intanto, $\|u - v_t\|_X \leq t\|z\|_X$ e quindi $v_t \to u$ per $t \to 0^+$.
Inoltre, applicando nuovamente la proprietà $R(I+\partial f) = X$, esiste $v \in D(\partial f)$ tale che $u + w \in (I + \partial f)(v)$, cioè

$$u + w - v \in \partial f(v).$$

Dimostriamo adesso il seguente facile

Lemma 7.8.10 *Nelle ipotesi del teorema 7.8.9, siano $w, v \in D(\partial f)$: allora*

$$\|w - v\|_X^2 \leq \|w - v + \mu(x - y)\|_X^2 \qquad \forall \mu > 0,\ \forall x \in \partial f(w),\ \forall y \in \partial f(v).$$

Dimostrazione Basta osservare che

$$\|w-v+\mu(x-y)\|_X^2 = \|w-v\|_X^2+\mu^2\|x-y\|_X^2+2\mu\,\langle w-v, x-y\rangle_X \geq \|w-v\|_X^2. \quad \square$$

Scegliamo nel lemma

$$\mu = 1, \quad w = v_t, \quad x = u + w + tz - v_t, \quad y = u + w - v;$$

si ottiene

$$\|v_t - v\|_X^2 \leq \|v_t - v + (u + w + tz - v_t) - (u + w - v)\|_X^2 = t^2\|z\|_X^2$$

da cui $v_t \to v$ per $t \to 0^+$. Ma allora $u = v$, cioè

$$w \in \partial f(u).$$

Ciò dimostra che ∂f è massimale monotono.
Infine, la relazione $\overline{D(\partial f)} = \overline{D(f)}$ è stata dimostrata nell'osservazione 5.3.4.
$\square$

Esercizi 7.8

1. Sia H uno spazio di Hilbert e siano $u \in L^p(0,T;H)$ e $v \in L^q(0,T;H)$ $(\frac{1}{p}+\frac{1}{q}=1)$. Se risulta

$$\int_0^T [\langle \varphi'(t), u(t)\rangle_H + \langle v(t), \varphi(t)\rangle_H]\, dt = 0 \qquad \forall \varphi \in C_0^\infty(]0,T[,H),$$

si provi che u è derivabile q.o. a valori in H e che $u' = v$ q.o. in $[0,T]$.

2. Sia H uno spazio di Hilbert e sia K un convesso chiuso non vuoto contenuto in H. Detta I_K la funzione indicatrice di K, introdotta nel capitolo 2, si provi che

$$(I + \partial I_K)^{-1}(v) = \{P_K(v)\} \qquad \forall v \in H,$$

ove P_K è la proiezione sul convesso K.

7.9 Inclusioni funzionali

Sia X uno spazio di Banach, sia $K \subseteq X$ un convesso chiuso non vuoto e siano $A, B : K \to \mathcal{P}(X^*)$ due multifunzioni, con A massimale monotona. La multifunzione B rappresenta una "perturbazione" di A. Fissato un elemento $b \in X^*$, vogliamo studiare l' *inclusione funzionale*

$$\begin{cases} u \in K \\ b \in A(u) + B(u). \end{cases}$$

ove u è l'incognita. Il senso di questa scrittura è il seguente: si cerca un elemento $u \in K$ tale che il funzionale b si possa scrivere come $b = v + w$, con $v \in A(u)$ e $w \in B(u)$. Gli esempi che seguono forniscono alcune delle numerose motivazioni che inducono a questo studio.

Esempio 7.9.1 (equazioni integrali) Un'equazione integrale della forma

$$u(x) + \int_\Omega k(x,y) f(y, u(y))\, dy = g(x), \qquad x \in \Omega,$$

ove Ω è un aperto di $\mathbb{R}^n$ e $g \in L^p(\Omega)$ è una funzione assegnata, si può mettere nella forma astratta seguente:

$$u + KF(u) = g,$$

dove $F(u) = f(\cdot, u(\cdot))$, $Kv = \int_\Omega k(\cdot, y)v(y)\,dy$. Utilizzando la multifunzione K^{-1}, si può scrivere anche

$$F(u) \in K^{-1}(g - u),$$

ovvero

$$0 \in K^{-1}(g - u) - F(u).$$

Questa è un'inclusione funzionale che rientra nel nostro schema.

Esempio 7.9.2 (equazioni di evoluzione del 1^{o} ordine) Si considerino i problemi

$$\begin{cases} u' + B(u) = f \\ u(0) = 0, \end{cases} \qquad\qquad \begin{cases} u' + B(u) = f \\ u(0) = u(T), \end{cases}$$

dove B è un operatore, lineare o no, di $L^p(0, T; X)$ in sé, ed f è un fissato elemento di $L^p(0, T; X)$ (X è uno spazio di Banach). Utilizzando i risultati dell'esempio 7.8.5, i due problemi si riscrivono così:

$$f = L_i u + B(u), \qquad i = 1, 2,$$

e questo è un caso particolare della nostra inclusione funzionale.

Esempio 7.9.3 (problemi di minimo) Il problema di minimizzare una funzione $f : X \to \mathbb{R}$, ove X è uno spazio di Banach, equivale per la proposizione 5.1.8 alla condizione

$$0 \in \partial f(u),$$

che è ancora un'inclusione funzionale.

Esempio 7.9.4 (disequazioni variazionali) Se X è uno spazio di Banach, $B : X \to X^*$ è un operatore (lineare o no) e $\varphi \in X^*$, la disequazione variazionale sul convesso chiuso $K \subseteq X$

$$\begin{cases} u \in K, \\ \langle B(u) - \varphi, v - u \rangle_X \geq 0 \qquad \forall v \in K \end{cases}$$

si traduce, in virtù dell'osservazione 6.3.7, nell'inclusione funzionale

$$\begin{array}{ll} \varphi \in \partial I_K(u) + B(u) \qquad & \text{se } K \subset X, \\ \varphi = B(u) & \text{se } K = X. \end{array}$$

Più in generale, se $\psi : X \to]-\infty, +\infty]$ è una funzione convessa, semicontinua inferiormente e propria, la disequazione variazionale

$$\langle B(u) - \varphi, v - u\rangle_X + \psi(v) - \psi(u) \geq 0 \qquad \forall v \in X$$

equivale all'inclusione funzionale

$$\varphi \in \partial\psi(u) + B(u).$$

Nella teoria delle inclusioni funzionali è di fondamentale importanza il seguente risultato:

Teorema 7.9.5 (di Browder) *Sia X uno spazio di Banach riflessivo, e sia $b \in X^*$. Supponiamo che:*

(i) *$K \subseteq X$ sia un convesso chiuso non vuoto;*

(ii) *$A : K \to \mathcal{P}(X^*)$ sia una multifunzione massimale monotona;*

(iii) *$B : K \to X^*$ sia un'applicazione limitata (ossia trasforma limitati in limitati),* demicontinua *(ossia debolmente continua sui sottospazi finito-dimensionali) e monotona;*

(iv) *nel caso che K sia illimitato, esistano $u_0 \in D(A)$, $\varphi_0 \in A(u_0)$, $r > 0$ tali che*

$$\langle B(u) + \varphi_0, u - u_0\rangle_X > \langle b, u - u_0\rangle_X \qquad \forall u \in K \text{ con } \|u\|_X > r.$$

Allora esiste $u \in D(A)$ che verifica l'inclusione funzionale $b \in A(u) + B(u)$.

Prima di dimostrare il teorema, segnaliamo i seguenti corollari di dimostrazione immediata.

Corollario 7.9.6 *Sia X uno spazio di Banach riflessivo. Supponiamo vere le ipotesi (i), (ii) e (iii) del teorema 7.9.5; supponiamo poi che, nel caso che K sia illimitato, esista $u_0 \in D(A)$ tale che*

$$\frac{\langle B(u), u - u_0\rangle_X}{\|u\|_X} \to +\infty \qquad \text{per } u \in K,\ \|u\|_X \to +\infty.$$

Allora $R(A + B) = X^*$. $\square$

Corollario 7.9.7 *Sia X uno spazio di Hilbert e sia $A : X \to \mathcal{P}(X)$ una multifunzione massimale monotona. Allora $R(I + A) = X$.* $\square$

Si noti che in particolare il corollario 7.9.7 dimostra la seconda implicazione del teorema 7.7.7.

Dimostrazione del teorema 7.9.5 Cominciamo con alcune osservazioni preliminari. Anzitutto, come è naturale aspettarsi, si ha $D(A) \neq \emptyset$: altrimenti avremmo $G(A) = \emptyset$, cosicché definendo ad esempio

$$\overline{A}(x) = \begin{cases} \{0\} & \text{se } x = 0 \\ \emptyset & \text{se } x \neq 0, \end{cases}$$

la multifunzione $\overline{A}$ sarebbe un'estensione propria e monotona di A, contraddicendo la massimale monotonia di A.
Inoltre possiamo supporre che risulti $u_0 = 0$ e $\varphi_0 = 0$, ossia:

(a) che $0 \in D(A)$ e $0 \in A(0)$;

(b) che se K è illimitato la proprietà (iv) sia vera per $u_0 = 0$ e $\varphi = 0$.

Infatti nessuna di queste due richieste è restrittiva: l'inclusione funzionale

$$u \in K, \qquad b \in A(u) + B(u)$$

equivale, posto $v = u - u_0$, all'inclusione funzionale

$$v \in K - u_0, \qquad b \in A(v + u_0) - \varphi_0 + B(v + u_0) + \varphi_0.$$

Ne segue che, definendo $\overline{A}(v) = A(v + u_0) - \varphi_0$ e $\overline{B}(v) = B(v + u_0) + \varphi_0$, le multifunzioni $\overline{A}$ e $\overline{B}$ verificano le ipotesi del teorema con K, u_0 e φ_0 rimpiazzati rispettivamente da $K - u_0$, $0 \in X$ e $0 \in X^*$; inoltre $0 \in \overline{A}(0)$. Quindi valgono (a) e (b). Supporremo pertanto nel seguito che $u_0 = 0 \in D(A)$ con $\varphi_0 = 0 \in A(0) \cap K$.
La dimostrazione vera e propria è alquanto articolata e si divide in vari passi.

1° passo: conversione dell'inclusione funzionale in una disequazione variazionale equivalente.

Proviamo che per un elemento $u \in K$ si ha

$$\begin{cases} u \in D(A) \\ b \in A(u) + B(u) \end{cases} \iff \begin{cases} u \in K \\ \langle b - B(u) - \varphi, u - v \rangle_X \geq 0 \ \forall (v, \varphi) \in G(A). \end{cases}$$

Infatti, se vale l'inclusione funzionale allora $(u, b-B(u)) \in G(A)$ e quindi, per monotonia, si ottiene la disequazione variazionale; viceversa, dalla validità di quest'ultima segue che $(v, b-B(u))$ è un elemento di $X \times X^*$ che verifica la condizione dell'osservazione 7.7.3: dato che A è massimale monotona, si deduce $(u, b-B(u)) \in G(A)$, cosicché concludiamo che u verifica l'inclusione funzionale.

2° passo: stima a priori per le soluzioni della disequazione variazionale.

Se K è limitato non c'è niente da dimostrare: K sarà incluso in una palla di centro 0 e raggio r e quindi se u risolve la disequazione variazionale risulterà $\|u\|_X \leq r$.
Se K è illimitato e $u \in K$ è soluzione, allora poiché $(0,0) \in G(A)$ si ha

$$\langle b - B(u), u \rangle_X \geq 0.$$

D'altra parte l'ipotesi (iv) dice che esiste $r > 0$ per cui

$$\langle b - B(u), u \rangle_X < 0 \qquad \text{per } u \in K,\ \|u\|_X > r;$$

dunque deve essere $\|u\|_X \leq r$.

3° passo: risoluzione della disequazione variazionale in dimensione finita.

Sia $\mathcal{V}$ l'insieme di tutti i sottospazi finito-dimensionali di X. Sia $Y \in \mathcal{V}$, e consideriamo la disequazione variazionale

$$u_Y \in K \cap Y, \quad \langle b - B(u_Y) - \varphi, u_Y - v \rangle_X \geq 0 \quad \forall (v, \varphi) \in G(A) \text{ con } v \in Y.$$

Per $R > 0$ e $Y \in \mathcal{V}$, poniamo

$$K_R = \{v \in K \cap Y : \ \|v\|_X \leq R\}, \qquad G_R = \{(v, \varphi) \in G(A) : \ v \in K_R\},$$

e studiamo la disequazione variazionale in K_R:

$$u_R \in K_R\,, \quad \langle b - B(u_R) - \varphi, u_R - v \rangle_X \geq 0 \quad \forall (v, \varphi) \in G_R\,.$$

Poiché K_R è un convesso chiuso di Y e Y ha dimensione finita, questa disequazione variazionale ha soluzione $u_R \in K_R$ in virtù del lemma che segue.

Lemma 7.9.8 *Sia X uno spazio di Banach, sia K un convesso compatto non vuoto di X; siano poi $M \subseteq K \times X^*$ un insieme monotono e $T : K \to X^*$ un'applicazione w-continua. Allora esiste $u \in K$ che risolve la disequazione variazionale*

$$\langle T(u) - f, u - v \rangle_X \geq 0 \qquad \forall (v, f) \in M.$$

Dimostrazione Supponiamo per assurdo che non vi sia alcuna soluzione in K: allora per $(v, f) \in M$ poniamo

$$U(v, f) = \{u \in K : \langle T(u) - f, u - v\rangle_X < 0\}.$$

L'insieme $U(v, f)$ è aperto in K: infatti $K \setminus U(v, f)$ è chiuso, come segue facilmente dalla w-continuità di T. Inoltre, poiché nessun $u \in K$ risolve la disequazione variazionale, si ha

$$K = \bigcup_{(v,f)\in M} U(v, f);$$

per compattezza, esistono $(v_i\,, f_i) \in M$, $i = 1, \ldots, m$, tali che

$$K = \bigcup_{i=1}^{m} U(v_i\,, f_i).$$

Sia $\{\beta_i\}_{1\le i\le m}$ una partizione dell'unità associata al ricoprimento $\{U(v_i\,, f_i)\}$. Poniamo $K_1 = \mathrm{co}\{v_1, \ldots, v_m\}$ e definiamo

$$p : K_1 \to K_1\,, \qquad p(u) = \sum_{j=1}^{m} \beta_j(u)v_j \quad \forall u \in K_1,$$

$$q : K_1 \to X^*, \qquad q(u) = \sum_{i=1}^{m} \beta_i(u)f_i \quad \forall u \in K_1.$$

Dato che p è continua, per il teorema di Brouwer esiste $\overline{u} \in K_1 \subseteq K$ tale che $p(\overline{u}) = \overline{u}$. Definiamo le quantità

$$\Delta_{ij} = \langle T(\overline{u}) - f_i, \overline{u} - v_j\rangle_X\ ;$$

dalla monotonia di M si ha

$$\begin{aligned}
\Delta_{ij} + \Delta_{ji} &= \langle T(\overline{u}) - f_i, \overline{u} - v_j\rangle_X + \langle T(\overline{u}) - f_j, \overline{u} - v_i\rangle_X = \\
&= \Delta_{jj} + \langle f_j - f_i, \overline{u} - v_j\rangle_X + \Delta_{ii} + \langle f_i - f_j, \overline{u} - v_i\rangle_X = \\
&= \Delta_{jj} + \Delta_{ii} + \langle f_j - f_i, v_i - v_j\rangle_X \le \Delta_{jj} + \Delta_{ii}\ .
\end{aligned}$$

Ne segue

$$\begin{aligned}
0 &= \langle T(\overline{u}) - q(\overline{u}), \overline{u} - p(\overline{u})\rangle_X = \\
&= \sum_{i,j=1}^{m} \langle \beta_i(\overline{u})(T(\overline{u}) - f_i), \beta_j(\overline{u})(\overline{u} - v_j)\rangle_X = \\
&= \sum_{i,j=1}^{m} \beta_i(\overline{u})\beta_j(\overline{u})\Delta_{ij} = \sum_{i,j=1}^{m} \beta_i(\overline{u})\beta_j(\overline{u})\frac{\Delta_{ij} + \Delta_{ji}}{2} \leq \\
&\leq \sum_{i,j=1}^{m} \beta_i(\overline{u})\beta_j(\overline{u})\frac{\Delta_{jj} + \Delta_{ii}}{2}.
\end{aligned}$$

D'altra parte, se $\beta_i(\overline{u})\beta_j(\overline{u}) > 0$ allora $\overline{u} \in U(v_i\,, f_i) \cap U(v_j\,, f_j)$ e quindi $\Delta_{jj} < 0$, $\Delta_{ii} < 0$: ne segue che gli addendi non nulli di questa somma sono negativi e, visto che $\overline{u} \in K$, essi non sono tutti nulli. Ciò è assurdo. □

Applicando il lemma precedente, come già osservato si ha che per ogni $R > 0$ la disequazione variazionale

$$u_R \in K_R\,, \quad \langle b - B(u_R) - \varphi, u_R - v\rangle_X \geq 0 \quad \forall (v, \varphi) \in G_R$$

ha soluzione. Sia S_R l'insieme delle soluzioni:

$$S_R = \{u \in K_R : \ \langle b - B(u) - \varphi, u - v\rangle_X \geq 0 \ \ \forall (v, \varphi) \in G_R\}.$$

Oltre che non vuoto, S_R è limitato: infatti per ogni $u \in S_R$ si ha necessariamente $\|u\|_X \leq r$ (altrimenti, per l'ipotesi (iv), scegliendo $(v, \varphi) = (0, 0)$ avremmo $\langle b - B(u), u\rangle_X < 0$, in contraddizione con la disequazione variazionale). Poiché B è demicontinuo, S_R è chiuso; dato che $\dim Y < \infty$, S_R è compatto. Inoltre

$$r \leq R < R' \qquad \Longrightarrow \qquad G_R \subseteq G_{R'} \qquad \Longrightarrow \qquad S_R \supseteq S_{R'}\,,$$

cosicché $\{S_R\}_{R \geq r}$ è una famiglia di compatti che ha la proprietà dell'intersezione finita. Pertanto

$$\exists u_Y \in \bigcap_{R \geq r} S_R\,;$$

di conseguenza $u_Y \in K$ e

$$\langle b - B(u_Y) - \varphi, u_Y - v\rangle_X \geq 0 \qquad \forall (v, \varphi) \in G_R\,,\ \forall R \geq r.$$

Quindi u_Y è soluzione della disequazione variazionale in Y. Si noti che $\|u_Y\|_X \leq r$ per ogni $Y \in \mathcal{V}$.

4$^{\text{o}}$ passo: passaggio al limite quando Y "invade" X.

Per $Z \in \mathcal{V}$ definiamo

$$\begin{aligned} M_Z \;=\; & \{(u, B(u)) \in K \times X^* : \ \exists Y \in \mathcal{V},\ Y \supseteq Z \text{ tale che} \\ & \langle b - B(u) - \varphi, u - v\rangle_X \geq 0 \ \ \forall (v, \varphi) \in G(A) \text{ con } v \in Y\}. \end{aligned}$$

Osserviamo che se $Z, W \in \mathcal{V}$ e $Z \subseteq W$, allora per definizione $M_Z \supseteq M_W$. Proveremo che

$$\exists(\overline{u}, \overline{\varphi}) \in \bigcap_{Z \in \mathcal{V}} \overline{M_Z}^w$$

(ove $\overline{M_Z}^w$ è la chiusura debole di M_Z), e che il corrispondente $\overline{u}$ risolve la disequazione variazionale originaria.

Poiché l'operatore B è limitato, se $(u, B(u)) \in M_Z$ si ha $\|u\|_X \leq r$ e dunque $\|B(u)\|_{X^*} \leq K_r$, con K_r costante opportuna; quindi esiste una palla chiusa $S \subset X \times X^*$ tale che

$$\bigcup_{Z \in \mathcal{V}} M_Z \subseteq S.$$

Poiché X è riflessivo, tali sono X^* e $X \times X^*$; quindi S è debolmente compatta. Essendo $\overline{M_Z}^w$ debolmente chiuso e contenuto in S, $\overline{M_Z}^w$ è debolmente compatto e

$$\bigcup_{Z \in \mathcal{V}} \overline{M_Z}^w \subseteq S.$$

Siano ora $Z_1, \ldots, Z_m \in \mathcal{V}$: posto $W = [\{Z_1, \ldots, Z_m\}]$ (il minimo sottospazio contenente gli Z_i), avremo $M_W \subseteq \bigcap_{i=1}^m M_{Z_i}$, da cui

$$\overline{M_W}^w \subseteq \overline{\bigcap_{i=1}^m M_{Z_i}}^w \subseteq \bigcap_{i=1}^m \overline{M_{Z_i}}^w.$$

Quindi, per la proprietà dell'intersezione finita, esiste $(\overline{u}, \overline{\varphi}) \in \bigcap_{Z \in \mathcal{V}} \overline{M_Z}^w$.

5$^{\text{o}}$ passo: alcuni lemmi ausiliari.

Lemma 7.9.9 *Nelle ipotesi del teorema 7.9.5, se $(\overline{u}, \overline{\varphi}) \in \bigcap_{Z \in \mathcal{V}} \overline{M_Z}^w$ allora per ogni $Z \in \mathcal{V}$ esiste una successione $\{(u_n, B(u_n))\} \subseteq M_Z$ tale che $u_n \rightharpoonup \overline{u}$ in X e $B(u_n) \rightharpoonup \overline{\varphi}$ in X^*. In particolare, $\overline{u} \in K$.*

Dimostrazione Poiché M_Z è limitato in $K \times X^*$, la sua chiusura debole coincide con la sua chiusura debole sequenziale in virtù del lemma 1.6.10. Ne segue la tesi. $\square$

Lemma 7.9.10 *Nelle ipotesi del teorema 7.9.5, sia* $(\overline{u}, \overline{\varphi}) \in \bigcap_{Z \in \mathcal{V}} \overline{M_Z}^w$*; allora esiste* $(\overline{v}, \overline{\psi}) \in G(A)$ *tale che*

$$\langle b - \overline{\varphi} - \overline{\psi}, \overline{u} - \overline{v} \rangle_X \leq 0.$$

Dimostrazione Altrimenti avremmo $\langle b - \overline{\varphi} - \psi, \overline{u} - v \rangle_X > 0$ per ogni $(v, \psi) \in G(A)$; quindi otterremmo $\overline{u} \in D(A)$ e $b - \overline{\varphi} \in A(\overline{u})$ per la massimale monotonia di A. Ma allora, scelti $v = \overline{u}$ e $\psi = b - \overline{\varphi}$, dedurremmo $0 > 0$. $\square$

Lemma 7.9.11 *Sia X uno spazio di Banach, sia $K \subseteq X$ un convesso chiuso non vuoto e sia $B : K \to X^*$ un operatore monotono ed* emicontinuo *(ossia debolmente continuo sulle rette). Allora B è* pseudomonotono*, cioè verifica la proprietà seguente:*

$$\begin{aligned} &\{u_n\} \subseteq K, \quad u_n \rightharpoonup u, \quad \limsup_{n\to\infty} \langle B(u_n), u_n - u \rangle_X \leq 0 \quad \Longrightarrow \\ &\qquad \Longrightarrow \quad \langle B(u), u - v \rangle_X \leq \liminf_{n\to\infty} \langle B(u_n), u_n - v \rangle_X \quad \forall v \in K. \end{aligned}$$

Dimostrazione Sia $\{u_n\} \subseteq K$ tale che

$$u_n \rightharpoonup u \text{ in } X, \qquad \limsup_{n\to\infty} \langle B(u_n), u_n - u \rangle_X \leq 0.$$

Per monotonia

$$\langle B(u), u_n - u \rangle_X \leq \langle B(u_n), u_n - u \rangle_X \qquad \forall n \in \mathbb{N},$$

e poiché $u_n \rightharpoonup u$ si deduce

$$\lim_{n\to\infty} \langle B(u_n), u_n - u \rangle_X = 0.$$

Ancora per monotonia,

$$\langle B(v), u_n - v \rangle_X \leq \langle B(u_n), u_n - v \rangle_X \qquad \forall n \in \mathbb{N}, \quad \forall v \in K,$$

e per $n \to \infty$

$$\langle B(v), u - v \rangle_X \leq \liminf_{n\to\infty} \langle B(u_n), u_n - v \rangle_X \qquad \forall v \in K.$$

Adesso, fissato $w \in K$, scegliamo $v = u + \lambda(w - u)$, con $\lambda \in]0, 1]$: si trova

$$\begin{aligned}
&\langle B(u + \lambda(w - u)), \lambda(u - w)\rangle_X \leq \\
&\quad \leq \liminf_{n\to\infty} [\langle B(u_n), u_n - u\rangle_X + \lambda\langle B(u_n), u - w\rangle_X] = \\
&\quad = \lambda \liminf_{n\to\infty} \langle B(u_n), u - w\rangle_X = \\
&\quad = \lambda \lim_{n\to\infty} \langle B(u_n), u - u_n\rangle_X + \lambda \liminf_{n\to\infty} \langle B(u_n), u_n - w\rangle_X = \\
&\quad = \lambda \liminf_{n\to\infty} \langle B(u_n), u_n - w\rangle_X \,.
\end{aligned}$$

Dividendo per λ e usando l'emicontinuità di B, si deduce per $\lambda \to 0^+$

$$\langle B(u), u - w\rangle_X \leq \liminf_{n\to\infty} \langle B(u_n), u_n - w\rangle_X \qquad \forall w \in K,$$

e ciò prova la tesi. $\square$

6º passo: calcolo finale.
Sia $(\overline{u}, \overline{\varphi}) \in \bigcap_{Z\in\mathcal{V}} \overline{M_Z}^w$. Per il lemma 7.9.10, esiste $(\overline{v}, \overline{\psi}) \in G(A)$ tale che

$$\langle b - \overline{\varphi} - \overline{\psi}, \overline{u} - \overline{v}\rangle_X \leq 0.$$

Sia $Y \in \mathcal{V}$ tale che $\overline{v} \in Y$. Dimostreremo che

$$\langle b - B(\overline{u}) - \psi, \overline{u} - v\rangle_X \geq 0 \qquad \forall (v, \psi) \in G(A) \text{ con } v \in Y;$$

per l'arbitrarietà di Y, ciò proverà che $\overline{u}$ è soluzione della disequazione variazionale originaria.
Per il lemma 7.9.9, esiste $\{(u_n, B(u_n))\} \subseteq M_Y$ tale che $u_n \rightharpoonup \overline{u}$ in X e $B(u_n) \rightharpoonup \overline{\varphi}$ in X^*; si ha, per definizione di M_Y,

$$\langle b - B(u_n) - \psi, u_n - v\rangle_X \geq 0 \qquad \forall (v, \psi) \in G(A) \text{ con } v \in Y.$$

Scriviamo questa relazione nella forma

$$\begin{aligned}
\langle B(u_n), u_n - \overline{u}\rangle_X \leq \langle \psi - b, v - u_n\rangle_X + \langle B(u_n), v - \overline{u}\rangle_X& \\
\forall (v, \psi) \in G(A) \text{ con } v \in Y.&
\end{aligned}$$

Scelto $(v, \psi) = (\overline{v}, \overline{\psi})$, si trova per $n \to \infty$

$$\begin{aligned}
&\limsup_{n\to\infty} \langle B(u_n), u_n - \overline{u}\rangle_X \leq \\
&\quad \leq \langle \overline{\psi} - b, \overline{v} - \overline{u}\rangle_X + \langle \overline{\varphi}, \overline{v} - \overline{u}\rangle_X = \langle b - \overline{\varphi} - \overline{\psi}, \overline{u} - \overline{v}\rangle_X \leq 0.
\end{aligned}$$

Dal lemma 7.9.10 deduciamo che B è pseudomonotono; quindi

$$\langle B(\overline{u}), \overline{u} - v\rangle_X \leq \liminf_{n\to\infty} \langle B(u_n), u_n - v\rangle_X \qquad \forall v \in K \cap Y,$$

e dunque, usando la disequazione variazionale verificata da u_n, otteniamo per ogni $(v, \psi) \in G(A)$ con $v \in Y$

$$\langle B(\overline{u}), \overline{u} - v\rangle_X \leq \liminf_{n\to\infty} \langle b - \psi, u_n - v\rangle_X = \langle b - \psi, \overline{u} - v\rangle_X \,.$$

Pertanto

$$\langle b - B(\overline{u}) - \psi, \overline{u} - v\rangle_X \geq 0 \qquad \forall (v, \psi) \in G(A) \text{ con } v \in Y, \ \forall Y \in \mathcal{V},$$

cioè la tesi. Il teorema 7.9.5 è così provato. $\square$

Esercizi 7.9

1. Sia X uno spazio di Banach. Se $B : X \to X^*$ è monotono ed emicontinuo, si provi che B è pseudomonotono.

2. Sia X uno spazio di Banach. Se $B : X \to X^*$ è tale che

$$x_n \rightharpoonup x \text{ in } X \qquad \Longrightarrow \qquad B(x_n) \to B(x) \text{ in } X^*,$$

   si provi che B è pseudomonotono.

3. Sia X uno spazio finito-dimensionale. Se $B \in \mathcal{L}(X, X^*)$, si provi che B è pseudomonotono.

4. Sia X uno spazio di Banach riflessivo. Se $A : X \to X^*$ e $B : X \to X^*$ sono pseudomonotoni, si provi che $A + B$ è pseudomonotono.

5. Sia X uno spazio di Banach riflessivo. Se $A : X \to X^*$ è monotono ed emicontinuo, e $B : X \to X^*$ è tale che

$$x_n \rightharpoonup x \text{ in } X \qquad \Longrightarrow \qquad B(x_n) \to B(x) \text{ in } X^*,$$

   si provi che $A + B$ è pseudomonotono.

6. Sia X uno spazio di Banach. Si provi che ogni operatore $A : X \to X^*$ lineare e monotono è continuo.

7. Sia X uno spazio di Banach riflessivo. Se $A : X \to X^*$ è monotono ed emicontinuo, si provi che A è demicontinuo.

8. Sia X uno spazio di Banach. Si provi che se $A : X \to X^*$ è tale che

$$x_n \rightharpoonup x \text{ in } X \qquad \Longrightarrow \qquad A(x_n) \to A(x) \text{ in } X^*,$$

allora A è un operatore *compatto*, ossia trasforma insiemi limitati in insiemi relativamente compatti; si mostri che se A è lineare, allora vale anche il viceversa.

9. Sia X uno spazio di Banach. Se $A : X \to X^*$ è pseudomonotono e $B : X \to X^*$ è monotono ed emicontinuo, si provi che $A + B$ è pseudomonotono.

10. Sia X uno spazio di Banach. Se $A : X \to X^*$ è pseudomonotono e $B : X \to X^*$ è tale che

$$x_n \rightharpoonup x \text{ in } X \qquad \Longrightarrow \qquad B(x_n) \to B(x) \text{ in } X^*,$$

si provi che $A + B$ è pseudomonotono.

Bibliografia

[1] A. Andreotti, Lezioni sugli spazi vettoriali topologici, Ist. Mat. “L.Tonelli”, Pisa 1965.

[2] J. P. Aubin, A. Cellina, Differential inclusions. Set-valued maps and viability theory, Springer Verlag, Berlin 1984.

[3] J. P. Aubin, H. Frankowska, Set-valued analysis, Birkhauser, Boston 1990.

[4] C. Baiocchi, A. Capelo, Disequazioni variazionali e quasi-variazionali. Applicazioni a problemi di frontiera libera, Pitagora ed., Bologna 1976.

[5] V. Barbu, Th. Precupanu, Convexity and optimization in Banach spaces, D. Reidel Publ. Comp., Dordrecht 1986.

[6] H. Brézis, Analyse fonctionnelle. Théorie et applications, Masson, Paris 1983.

[7] H. Brézis, Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, North-Holland Publ. Comp., Amsterdam 1973.

[8] E. Burger, Introduction to the theory of games, Prentice Hall, Englewood Cliffs 1963.

[9] K. Deimling, Nonlinear functional analysis, Springer Verlag, Berlin 1985.

[10] N. Dunford, J. T. Schwartz, Linear operators, Interscience Publ., New York 1958.

[11] H. G. Eggleston, Convexity, Cambridge Univ. Press, Cambridge 1977.

[12] I. Ekeland, R.Temam, Analyse convexe et problèmes variationnels, Dunod, Paris 1974.

[13] I. M. Gelfand, S. V. Fomin, Calculus of variations, Prentice-Hall, Englewood Cliffs 1963.

[14] E. Hille, R. S. Phillips, Functional analysis and semigroups, Amer. Math. Soc., New York 1957.

[15] D. Kinderlehrer, G. Stampacchia, An introduction to variational inequalities and their applications, Academic Press, London 1980.

[16] A. N. Kolmogorov, S. V. Fomin, Elementi di teoria delle funzioni e di analisi funzionale, Mir, Mosca 1980.

[17] G. Prodi, A. Ambrosetti, Analisi non lineare, Scuola Norm. Sup., Pisa 1973.

[18] T. Parthasarathy, T. E. S. Raghavan, Some topics in two-person games, Elsevier, New York 1971.

[19] R. T. Rockafellar, R. J.-B. Wets, Variational analysis, Springer Verlag, Berlin 1988.

[20] W. Rudin, Functional analysis, McGraw-Hill, New York 1973.

[21] H. Sagan, Introduction to the calculus of variations, McGraw-Hill, New York 1969.

[22] F. Treves, Topological vector spaces, distributions and kernels, Academic Press, New York 1967.

[23] J. L. Troutman, Variational calculus with elementary convexity, Springer Verlag, Berlin 1983.

[24] E. Zeidler, Nonlinear functional analysis and its applications, Vol. I - IIA - IIB, Springer Verlag, Berlin 1984.

Indice analitico