Brutte cose di analisi funzionale

De Donato Paolo

18 marzo 2022

Prefazione

Inizialmente questi appunti erano semplicemente una trascrizione delle lezioni del corso di Analisi Funzionale tenute nell'anno accademico 2017/2018 dal professore Carbone Luciano presso l'Università degli Studi di Napoli "Federico II".

Col passare dei giorni questi appunti sono stati corretti, rielaborati ed ampliati. Con l'aumentare delle modifiche e delle correzioni l'obbiettivo principale di mera trascrizione di lezioni orali è stato mano mano messo in secondo piano rispetto alla volontà di redigere un manuale introduttivo dell'analisi funzionale liberamente disponibile ad ogni studente universitario che sia studiando per l'esame o semplicemente interessato all'analisi funzionale.

Questo manuale procede per gradi, partendo da oggetti più generali e aggiungendo via via nuovi assiomi. Il primo capitolo in realtà è solamente introduttivo e richiama i concetti base di topologia necessari alla trattazione degli argomenti presenti nei capitoli successivi. Nel secondo capitolo verranno presentati gli spazi vettoriali dotati di topologie per le quali le operazioni di somma e prodotto esterno risultano continue, chiamati spazi vettoriali topologici. Nel terzo capitolo invece ci concentreremo sugli spazi metrici e in particolare sui teoremi di compattezza e il teorema di Ascoli-Arzelà nella sua forma più generale. Le caratteristiche degli spazi vettoriali topologici e degli spazi metrici vengono ereditati ed assimilati all'interno degli spazi normati e di Banach, i quali rivestono un ruolo così importante all'interno dell'analisi funzionale da essere trattati in due capitoli anziché uno: nel capitolo 4 verranno analizzate le proprietà più elementari mentre nel quindo ci si concentrerà sugli operatori lineari tra spazi normati. Il sesto capitolo infine è incentrato sugli spazi di Hilbert e sulle sue proprietà più elementari senza addentrarsi eccessivamente su di esso.

I seguenti appunti sono da considerarsi tutt'altro che definitivi e potrebbero contenere errori o inesattezze di vario genere. Chiunque riscontrasse errori nelle dimostrazioni o sviste al livello grammaticale o più semplicemente volesse darmi qualche consiglio a riguardo è pregato di segnalarli all'indirizzo email `pdd.math@gmail.com`.

Indice

Notazioni

$\#(X)$	Cardinalità dell'insieme X.	1
$\mathcal{P}(X)$	Famiglia dei sottoinsiemi di A.	1
$\mathcal{F}(X)$	Famiglia dei sottoinsiemi finiti di A.	1
$\{X_\alpha\}_{\alpha\in I}$	Famiglia di insiemi su I.	1, 2
$F(A,B)$	Insieme delle funzioni definite su A a valori in B.	1
$f\big\|_C$	Restrizione di f su C.	1
$F(I,\mathscr{X})$	Sottoinsieme di $F\left(I,\bigcup\mathscr{X}\right)$ per cui $f(i)\in\mathscr{X}(i)$ per ogni $i\in I$.	1
$MV(I,\mathscr{X})$	Spazio delle funzioni definite su qualche sottoinsieme finito di I a valori in $\mathscr{X}$.	1
$\operatorname{dom} f$	Dominio della funzione f.	2
$\operatorname{im} f$	Immagine della funzione f.	2
$\overline{\mathbb{R}}$	$\mathbb{R}$ esteso.	2
$\operatorname{span} S$	Sottospazio vettoriale generato da S.	5
$\mathcal{L}(V,W)$	Insieme delle funzioni lineari da V in W.	5
V^+	Duale algebrico di V.	5
$\mathbb{K}$	Un campo che può essere o $\mathbb{R}$ o $\mathbb{C}$.	6
$z^\dagger$	Coniugato di z.	7
$\mathcal{UL}\left(V,W\right)$	Insieme delgli operatori lineari da V in W.	10
$\mathscr{G}(L)$	Grafico dell'operatore L.	10
$x\vee y$	Massimo tra x ed y.	12
$x\wedge y$	Minimo tra x ed y.	12
$\operatorname{Cl}(I)$	Chiusura di I.	17
$\operatorname{Int}(I)$	Interno di I	17
$C(X,Y)$	Insieme delle funzioni continue da X in Y.	22
id_X	Funzione identica da X in sé stesso.	23
$\operatorname{dist}(M,N;d)$	La distanza tra M ed N rispetto alla metrica d.	60
$B(x,r,d)$	Palla aperta di centro x e raggio r rispetto alla metrica d.	61
$D(x,r,d)$	Palla chiusa di centro x e raggio r rispetto alla metrica d.	61
$L(V,W)$	Insieme delle funzioni lineari e continue da V in W.	90
V^*	Duale continuo di V.	90
$\operatorname{epi}_c f$	Epigrafico chiuso di f	109
$\operatorname{epi}_o f$	Epigrafico aperto di f	109
δ_{ij}	Delta di Dirac.	130

Capitolo 1

Concetti preliminari

1.1 Insiemi e funzioni

Dato un insieme finito X indichiamo con il simbolo $\#(X)$ la cardinalità di X, che sarà quindi un numero intero non negativo dato che non abbiamo bisogno di definire la cardinalità per insiemi infiniti. Inoltre come da tradizione indicheremo con i simboli $\mathcal{P}(X)$ e $\mathcal{F}(X)$ rispettivamente l'insieme delle parti e l'insieme delle parti finite di X.

Non avendo intenzione di penetrare a fondo nella teoria assiomatica degli insiemi si dà per scontato che il lettore conosca già le definizioni di insiemi e di funzioni, e di conseguenza tutti i risultati di base su di essi. In particolare con il termine **famiglia di insiemi** con indici in I indicheremo semplicemente una funzione con dominio in I e codominio $\mathcal{P}(D)$ per un qualche insieme D, in altre parole una famiglia di insiemi su I associa ad ogni elemento i di I un insieme. Utilizzeremo perciò il simbolo $\{X_\alpha\}_{\alpha\in I}$ per indicare quella famiglia di insiemi su I che assume il valore X_α per ogni $\alpha \in I$.

Dato che abbiamo supposto l'esistenza di un insieme D per una famiglia di insiemi $\mathscr{F}\{U_i\}_{i\in I}$ grazie alla teoria assiomatica degli insiemi possiamo definirne l'unione come

$$\bigcup \mathscr{F} = \bigcup_{i\in I} U_i$$

definita come quell'unico insieme $D' \subseteq D$ tale che $x \in D'$ se e solo se esiste un $i \in I$ tale che $x \in U_i$. In maniera analoga se ne definisce l'intersezione.

Dati due insiemi A e B indichiamo con il simbolo $F(A,B)$ l'insieme delle funzioni da A in B. Preso $C \subseteq A$ ed $f \in F(A,B)$ definiamo la **restrizione** di f su C con il simbolo $f\big|_C$.

Ora dato un insieme di indici I e una famiglia di insiemi $\mathscr{X} = \{X_\alpha\}_{\alpha\in I}$ indicheremo con il simbolo $F(I,\mathscr{X})$ lo spazio di tutte le funzioni f definite su I tali che $f(\alpha) \in X_\alpha$ per ogni $\alpha \in I$. Analogamente indicheremo con il simbolo $MV(I,\mathscr{X})$ lo spazio di tutte le funzioni f definite su un qualche sottoinsieme finito I' di I tale che $f(\alpha) \in X_\alpha$ per ogni $\alpha \in I'$, ovvero

$$MV(I,\mathscr{X}) = \bigcup_{I'\in\mathcal{F}(I)} F\left(I',\mathscr{X}\right)$$

Consideriamo ora a titolo di esempio $I = \{1,2,3\}$ e $\mathscr{X} = \{X_1, X_2, X_3\}$ famiglia di insiemi generica. Allora per ogni $f \in F(I,\mathscr{X})$ avremo che $f(1) \in X_1, f(2) \in X_2, f(3) \in X_3$. Quindi ogni funzione in $F(I,\mathscr{X})$ può essere espressa univocamente come una terna ordinata $[f(1), f(2), f(3)]$ in $X_1 \times X_2 \times X_3$ e viceversa. In particolare per una generica famiglia di insiemi

$\{X_\alpha\}_{\alpha\in I}$ definiamo il suo **prodotto cartesiano** come

$$\prod_{\alpha\in I} X_\alpha = F(I,\mathscr{X})$$

i cui elementi sono di conseguenza non più semplicemente delle I-ple, ma delle vere e proprie funzioni definite su I.

Dati un insieme di indici I ed una famiglia di insiemi $\mathscr{X} = \{X_i\}_{i\in I}$ per ogni $i \in I$ possiamo definire la **proiezione** i-esima l'applicazione $\pi_i : F(I,\mathscr{X}) \to X_i$ in modo tale che

$$\pi_i(f) = f(i)$$

più in generale per ogni $J \subseteq I$ non vuoto possiamo restringere la famiglia $\mathscr{X}$ su J per ottenere la nuova famiglia $\mathscr{Y} = \{X_i\}_{i\in J}$ e definire stavolta la proiezione $\pi_J : F(I,\mathscr{X}) \to F(J,\mathscr{Y})$ in modo tale che ad ogni $f \in \mathscr{F}(I,\mathscr{X})$ associ la sua restrizione su J.

Nel seguito daremo per vero il cosiddetto assioma della scelta. Benché possa sembrare un fatto noto e comprensibile esso risulta non solo indipendente dagli altri assiomi della teoria degli insiemi ma anche molto potente, il che ha spinto molti matematici a farne a meno o comunque di indebolirlo sensibilmente. Nella sua forma più generale esso afferma che

Assioma della scelta. *Data una famiglia di insiemi $\mathscr{X} = \{X_i\}_{i\in I}$ con $I \neq \emptyset$. Se $X_i \neq \emptyset$ per ogni $i \in I$ allora $F(I,\mathscr{X}) = \prod_{i\in I} X_i \neq \emptyset$.*

Se I è finito esso può essere dimostrato molto facilmente, e non è difficile dimostrarlo a partire dagli altri assiomi quando $I = \mathbb{N}$. Quando invece I è un insieme generico come abbiamo detto prima non è possibile dimostrarlo a partire dagli altri assiomi e perciò bisogna darlo per vero. L'assioma della scelta è una componente fondamentale per la dimostrazione del lemma di Zorn che useremo moltissime volte in queste note.

Data una funzione generica f indicheremo con $\operatorname{dom} f$ e $\operatorname{im} f$ rispettivamente il dominio e l'immagine (l'insieme di tutti gli elementi nella forma $f(x)$) di f.

1.2 Calcolo in $\mathbb{R}$ esteso

Come da tradizione indicheremo con il simbolo $\mathbb{R}$ l'insieme dei numeri reali dotato delle usuali operazioni di addizione e sottrazione. In alcune momenti avremo bisogno di introdurre in $\mathbb{R}$ il concetto di elemento infinito, o più precisamente di infinito positivo e infinito negativo.

Indicheremo questi due elementi con i simboli $+\infty$ e $-\infty$ rispettivamente, e definiamo $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$. Induciamo anche su $\overline{\mathbb{R}}$ l'ordine naturale ottenuto ponendo in aggiunta che $-\infty < a < +\infty$ per ogni $a \in \mathbb{R}$. Poniamo anche $-(+\infty) = -\infty$ e $-(-\infty) = +\infty$. Discorso a parte va fatto per le operazioni su $\overline{\mathbb{R}}$, in quanto per poter definire la somma dobbiamo rinunciare a varie sue proprietà tra cui l'invertibilità.

Porremo da questo momento in poi che

- $a + \infty = +\infty$ per ogni $a \in \overline{\mathbb{R}}$;
- $a - \infty = -\infty$ per ogni $a \in \overline{\mathbb{R}}$ diverso da $+\infty$.

in particolare è bene notare che

$$+\infty - \infty = +\infty$$

e quindi gli elementi infiniti non possiedono l'elemento inverso. Precisiamo che alcune proprietà della somma continuano a valere mentre altre no, ad esempio si noti che

$$a - a = \begin{cases} 0 & \text{se } a \neq \pm\infty \\ +\infty & \text{se } a = +\infty \text{ o } a = -\infty \end{cases} \geq 0$$

Per comodità elenchiamo di seguito alcune implicazioni valide comunque per ogni $a, b, c \in \overline{\mathbb{R}}$

$$a \geq b \Rightarrow a + c \geq b + c$$
$$a \geq b + c \Rightarrow a - c \geq b$$
$$a + c \geq b \text{ e } b \in \mathbb{R} \Rightarrow a \geq b - c$$

Per quanto riguarda la moltiplicazione lasceremo indeterminata l'espressione $0 \cdot (\pm\infty)$ definendo le altre combinazioni normalmente.

1.3 Spazi vettoriali

Definizione 1.3.1. Siano $(V, +)$ un gruppo abeliano e $(\mathbb{K}, +, \cdot)$ campo con elemento neutro 0 ed unità 1. Allora fissata un'applicazione binaria nella seguente forma

$$h : \mathbb{K} \times V \to V$$

diremo che la struttura $(V, \mathbb{K}, h)$ è uno **spazio vettoriale** su $\mathbb{K}$ se e solo se valgono le seguenti proprietà:

- $h(1, v) = v$ per ogni $v \in V$;
- $h[\alpha, u + v] = h(\alpha, u) + h(\alpha, v)$ per ogni $u, v \in V$, $\alpha \in \mathbb{K}$;
- $h(\alpha + \beta, u) = h(\alpha, u) + h(\beta, u)$ per ogni $u \in V$, $\alpha, \beta \in \mathbb{K}$;
- $h(\alpha\beta, u) = h[\alpha, h(\beta, u)]$ per ogni $u \in V$, $\alpha, \beta \in \mathbb{K}$;

D'ora in avanti identificheremo un generico spazio vettoriale $(V, \mathbb{K}, h)$ con il suo supporto V e identificheremo il prodotto esterno con il simbolo $\cdot$ in ogni caso.

Proposizione 1.3.2. *Se $(V, \mathbb{K}, +, \cdot)$ è uno spazio vettoriale allora*

- $0v = 0$ *per ogni $v \in V$, si noti che lo zero al primo membro è il numero reale* 0*, mentre quello al secondo membro è il vettore nullo ovvero l'elemento neutro del gruppo abeliano $(V, +)$;*
- $(-1)v = -v$ *per ogni $v \in V$, si noti che il vettore al secondo membro è il vettore inverso di v nel gruppo abeliano $(V, +)$;*

Definizione 1.3.3. Se V è uno spazio vettoriale su $\mathbb{K}$ e $W \subseteq V$ un suo sottoinsieme non vuoto allora W è un **sottospazio** di V, e si indica con $W \leq V$, se e solo se la somma e la moltiplicazione per uno scalare di elementi di W rimane in W. In altre parole $(W, \mathbb{K}, +, \cdot)$ è esso stesso uno spazio vettoriale.

Dati due sottoinsiemi non vuoti A e B dello spazio vettoriale V ed $\alpha \in \mathbb{K}$ possiamo sempre costruire i seguenti insiemi

$$\begin{aligned} A+B &= \{a+b \mid a \in A, b \in B\} \\ A-B &= \{a-b \mid a \in A, b \in B\} \\ \alpha A &= \{\alpha a \mid a \in A\} \end{aligned}$$

che chiameremo rispettivamente somma algebrica, differenza algebrica e omotetia per α. Se invece almeno uno tra A e B è l'insieme vuoto poniamo comunque $A+B=A-B=\emptyset$ e se $A=\emptyset$ allora $\alpha A=\emptyset$ per ogni $\alpha \in \mathbb{K}$.

Non è difficile verificare che valgono le seguenti implicazioni, che valgono anche quando si prendono insiemi vuoti

$$\begin{gathered} 0 \in A-A \\ A \subseteq B \Rightarrow A+C \subseteq B+C \\ A+x \subseteq B+x \text{ se e solo se } A \subseteq B \\ (A \cap B)+C \subseteq (A+C) \cap (B+C) \\ (A+B) \cap C \subseteq [A \cap (C-B)]+B \\ A \subseteq B \text{ se e solo se } \alpha A \subseteq \alpha B \text{ per qualche } \alpha \neq 0 \\ (\alpha+\beta)A \subseteq \alpha A+\beta A \\ \alpha(A+B)=\alpha A+\alpha B \\ \alpha(\beta A)=(\alpha\beta)A \end{gathered}$$

Altre relazioni che sono verificate tra vettori in uno spazio vettoriale non valgono se estese ai suoi sottoinsiemi. In particolare si osservi che in generale si ha

$$\begin{gathered} A-A \neq \{0\} \\ (\alpha+\beta)A \neq \alpha A+\beta A \\ A+C \subseteq B+C \not\Rightarrow A \subseteq B \end{gathered}$$

Proposizione 1.3.4. *Un sottoinsieme A di uno spazio vettoriale V su $\mathbb{K}$ è un sottospazio vettoriale se e solo se è non vuoto e per ogni $\alpha, \beta \in \mathbb{K}$ si ha*

$$\alpha A+\beta A \subseteq A \tag{1.3.1}$$

e in tal caso l'inclusione in realtà è un'uguaglianza se almeno uno tra α e β è non nullo.

Dimostrazione. Poniamo innanzitutto $\beta=-\alpha$ allora $0 \in A-A \subseteq A$ e quindi contiene l'origine e dalla (1.3.1) segue che A è un sottospazio vettoriale. Per ogni $\alpha \in \mathbb{K}$ avremo $\alpha A = \alpha A+0A \subseteq A$ e quindi per ogni $\alpha \neq 0$ si ha inoltre

$$A=\alpha\left(\frac{1}{\alpha}A\right) \subseteq \alpha A$$

ovvero $A=\alpha A$.

Supponiamo ora che $\alpha \neq 0$ allora

$$A=\alpha A \subseteq \alpha A+\beta A \subseteq A$$

e quindi $\alpha A+\beta A=A$. ■

Dato uno spazio vettoriale V su $\mathbb{K}$ ed $S \subseteq V$ non vuoto poniamo

$$\operatorname{span} S = \left\{ \sum_{i=1}^{N} k_i x_i \;\middle|\; N \in \mathbb{N}, x_i \in S, k_i \in \mathbb{K} \right\}$$

chiaramente $\operatorname{span} S$ è un sottospazio di V ed è il più piccolo sottospazio contenente S.

Definizione 1.3.5. Siano V e W due spazi vettoriali sullo stesso campo $\mathbb{K}$, diremo che l'applicazione $f : V \to W$ è **lineare** se e solo se per ogni $x, y \in X$, $\alpha, \beta \in \mathbb{K}$ vale

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

Indicheremo inoltre con $\mathcal{L}(V, W)$ l'insieme delle funzioni lineari da V in W.

Possiamo dotare lo spazio $\mathcal{L}(V, W)$ di una struttura di $\mathbb{K}$-spazio vettoriale con somma $+$ e prodotto esterno $\cdot$ definiti nella seguente maniera.

$$(f + g)(x) = f(x) + g(x)$$
$$(\alpha f)(x) = \alpha f(x)$$

per ogni $f, g \in \mathcal{L}(V, W)$, $\alpha \in \mathbb{K}$ ed $x \in V$. Un qualunque campo $(\mathbb{K}, +, \cdot)$ può essere visto anche come $\mathbb{K}$-spazio vettoriale, possiamo allora porre

$$V^+ = \mathcal{L}(V, \mathbb{K})$$

e V^+ sarà il **duale algebrico** di V. Per ogni funzione lineare $f \in \mathcal{L}(V, V')$ possiamo definire il **kernel** e l'**immagine** di f nella seguente maniera

$$\ker f = \{x \in V \mid f(x) = 0\}$$
$$\operatorname{im} f = \{f(x) \in V' \mid x \in V\}$$

Essi sono chiaramente anche sottospazi di V e V' rispettivamente.

Definizione 1.3.6. Siano $V, V_1, V_2, \dots, V_n$ spazi vettoriali generici su un campo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, diciamo che l'applicazione $f : V_1 \times V_2 \times \cdots \times V_n \to V$ è **multilineare** se e solo se risulta essere lineare rispetto a ciascuna delle precedenti variabili.

Consideriamo adesso uno spazio vettoriale V, un suo sottospazio $W \leq V$ ed un'applicazione lineare $p : V \to V$. Diremo che p è una **proiezione** lineare su W se e solo se soddisfa le seguenti proprietà:

- $p(v) \in W$ per ogni $v \in V$;
- $p(v) = v$ per ogni $v \in W$.

In particolare segue che $W = \operatorname{im} p$ e quindi ha senso parlare solo di proiezione senza specificare necessariamente il sottospazio sul quale proiettare. Vale il seguente risultato di caratterizzazione:

Teorema 1.3.7. *Un'applicazione lineare $p : V \to V$ è una proiezione se e solo se $p \circ p = p$.*

Dimostrazione. Se p è una proiezione allora si ha chiaramente che $p \circ p = p$, viceversa poniamo

$$W = \{v \in V \mid p(v) = v\} = \ker(p - \mathrm{id}_V)$$

per la linearità di p segue immediatamente che W è un sottospazio vettoriale di V, inoltre $p(v) = v$ per ogni $v \in W$ per definizione, dobbiamo quindi verificare solamente che $\operatorname{im} p \leq W$.

Preso un qualunque $v \in V$ per ipotesi abbiamo $p[p(v)] = (p \circ p)(v) = p(v)$ ovvero $p(v) \in W$ per ogni $v \in V$, di conseguenza $\operatorname{im} p = W$ e p è una proiezione. ■

Lo spazio $\mathbb{K}^n$

Un esempio di spazio vettoriale finitamente generato su $\mathbb{K}$ è lo spazio $\mathbb{K}^n$ delle n-ple in $\mathbb{K}$, che può essere ricordiamo o $\mathbb{R}$ o $\mathbb{C}$. Per ogni $x \in \mathbb{K}^n$ e per ogni $p \in [1, +\infty)$ definiamo le quantità

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$
$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Esse soddisfano le seguenti relazioni per ogni $x, y \in \mathbb{K}^n$ e per ogni $p, q, r \in [1, +\infty]$

$$\|x\|_p = 0 \text{ se e solo se } x = 0$$
$$\|\lambda x\|_p \leq |\lambda| \, \|x\|_p \text{ per ogni } \lambda \in \mathbb{K}$$
$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$
$$\|x\|_p \leq \|x\|_q \text{ se } q \leq p$$
$$\|xy\|_r \leq \|x\|_p \, \|x\|_q \text{ dove } (xy)_i = x_i y_i \text{ e } 1/p + 1/q = 1/r$$

Nell'ultima relazione si pone $1/(+\infty) = 0$.

1.4 Somma diretta

Consideriamo una famiglia di spazi vettoriali $\mathscr{X} = \{X_a\}_{a \in A}$ rispetto allo stesso campo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, allora possiamo dotare l'insieme $F(A, \mathscr{X}) = \prod_{a \in A} X_a$ di una struttura di spazio vettoriale su $\mathbb{K}$ rispetto alle seguenti operazioni:

$$(f + g)(a) = f(a) + g(a) \qquad (\alpha f)(a) = \alpha f(a)$$

per ogni $f, g \in F(A, \mathscr{X})$, $\alpha \in \mathbb{K}$ ed $a \in A$. In questo modo $\prod_{a \in A} X_a$ è uno spazio vettoriale e le proiezioni π_a da $F(A, \mathscr{X})$ in X_a sono anche lineari. In particolare se restringiamo il dominio di π_a al sottospazio $\tilde{X}_a$, formato da tutte e sole le funzioni f in $F(A, \mathscr{X})$ tali che $f(b) = 0$ per ogni $b \neq a$, avremo che π_a è anche biettiva e quindi gli spazi vettoriali X_a ed $\tilde{X}_a$ sono isomorfi tra loro.

Questo spazio vettoriale risulta però essere troppo grande per i nostri scopi, per questo motivo introduciamo il seguente sottospazio

$$\bigoplus_{a \in A} X_a = \left\{ f \in \prod_{a \in A} X_a \;\middle|\; \exists A' \subseteq A \text{ finito tale che } f(b) = 0 \ \forall b \in A \setminus A' \right\}$$

che sarà detto **somma diretta** degli spazi vettoriali $\{X_a\}_{a\in A}$. Notiamo soprattutto che i $\tilde{X}_a$ sono ancora suoi sottospazi e che quando A è finito allora $\bigoplus_{a\in A} X_a = \prod_{a\in A} X_a$. In tal caso indicheremo la somma diretta anche nel seguente modo:

$$X_1 \oplus X_2 \oplus X_3 \oplus \cdots \oplus X_{n-1} \oplus X_n$$

Le proprietà più interessanti della somma diretta sono però le seguenti:

1. $\bigoplus_{a\in A} X_a = \operatorname{span}\left(\bigcup_{a\in A} \tilde{X}_a\right)$;

2. $\tilde{X}_a \cap \operatorname{span}\left(\bigcup_{\substack{b\in A\\ b\neq a}} \tilde{X}_a\right) = \{0\}$ per ogni $a \in A$.

da queste proprietà non è difficile verificare che qualunque $f \in \bigoplus_{a\in A} X_a$ può essere scritto in maniera univoca come una somma di un numero finito di elementi di $\tilde{X}_a$. In altre parole se $f \neq 0$ allora esistono sempre e sono univocamente determinati $n \in \mathbb{N}, a_1, a_2, \ldots, a_n \in A$, $f_i \in \tilde{X}_{a_i} \setminus \{0\}$ tali che

$$f = f_1 + f_2 + \cdots + f_n$$

Generalizziamo il seguente concetto a qualunque spazio vettoriale X e non solo alla somma diretta.

Definizione 1.4.1. Dati uno spazio vettoriale X ed una famiglia di suoi sottospazi $\{X_a\}_{a\in A}$ diremo che X è **somma diretta** degli X_a se e solo se

1. $X = \operatorname{span}\left(\bigcup_{a\in A} X_a\right)$;

2. $X_a \cap \operatorname{span}\left(\bigcup_{\substack{b\in A\\ b\neq a}} X_a\right) = \{0\}$ per ogni $a \in A$.

Abbiamo utilizzato lo stesso il nome di somma diretta per via del seguente risultato:

Teorema 1.4.2. *Se lo spazio vettoriale X è somma diretta dei suoi sottospazi $\{X_a\}_{a\in A}$ allora X è isomorfo ad $\bigoplus_{a\in A} X_a$ e tale isomorfismo identifica X_a con $\tilde{X}_a$ per ogni $a \in A$.*

In particolare per ogni $a \in A$ sono univocamente determinate le proiezioni $p_a : X \to X_a$ per ogni $a \in A$ e quindi possiamo scrivere per ogni $x \in X$

$$x = \sum_{a\in A} p_a(x)$$

dove la somma a destra contiene sempre un numero finito di termini non nulli.

1.5 Spazi vettoriali reali e complessi

Tutti gli spazi vettoriali che considereremo da qui in poi avranno come campo $\mathbb{R}$ o $\mathbb{C}$, quindi ogni volta che considereremo uno spazio vettoriale su $\mathbb{K}$ daremo per scontato che $\mathbb{K}$ sia $\mathbb{R}$ oppure $\mathbb{C}$. In questi appunti se $z = x + iy$ indicheremo il complesso coniugato di z con il simbolo $z^\dagger = x - iy$.

Chiaramente ogni spazio vettoriale V su $\mathbb{C}$ può essere visto anche come uno spazio vettoriale su $\mathbb{R}$ semplicemente restringendo il dominio del prodotto esterno. Indicheremo questo nuovo spazio vettoriale con $V_{\mathbb{R}}$ il quale, benché contenga gli stessi vettori di V, possiede proprietà diverse e quindi risulta opportuno differenziarli.

Difatti presi due spazi vettoriali V, W su $\mathbb{C}$ avremo che $\mathcal{L}(V,W) \subseteq \mathcal{L}\left(V_{\mathbb{R}}, W_{\mathbb{R}}\right)$ e l'inclusione in certi casi è stretta come mostra il seguente esempio.

Esempio 1.5.1. Poniamo $V = \mathbb{C}$ e dotiamolo della struttura canonica di $\mathbb{C}$-spazio vettoriale. Allora $V_{\mathbb{R}} \equiv \mathbb{R}^2$ è un $\mathbb{R}$-spazio vettoriale di dimensione 2 e l'applicazione

$$f : (x, y) \in \mathbb{R}^2 \to (y, x) \in \mathbb{R}^2$$

è chiaramente lineare da $\mathbb{R}^2$ in $\mathbb{R}^2$ rispetto al campo $\mathbb{R}$. D'altronde $f(x + iy) = y + ix$ non è $\mathbb{C}$-lineare in quanto

$$f(i) = 1 \neq -1 = if(1)$$

dunque $f \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ ma $f \notin \mathcal{L}(\mathbb{C}, \mathbb{C})$.

Inoltre il sottoinsieme $M = \left\{(x, 0) \in \mathbb{R}^2 \;\middle|\; x \in \mathbb{R}\right\}$ è un sottospazio vettoriale di $\mathbb{R}^2$ ma non di $\mathbb{C}$ in quanto $1 = (1, 0)$ appartiene a M ma $i \equiv i \cdot (1, 0) \equiv (0, 1)$ no.

Ciononostante se V è uno spazio vettoriale complesso allora vi è un isomorfismo tra $V^+ = \mathscr{L}(V, \mathbb{C})$ e $V_{\mathbb{R}}^+ = \mathscr{L}(V_{\mathbb{R}}, \mathbb{R})$, come si evince dal seguente risultato

Proposizione 1.5.2. *Se V è uno spazio vettoriale complesso allora l'applicazione*

$$T : f \in V^+ \to \operatorname{Re} f \in V_{\mathbb{R}}^+$$

è biettiva.

Dimostrazione. Chiaramente $(Tf)(x+y) = (Tf)(x)+(Tf)(y)$ e $(Tf)(rx) = r(Tf)(x)$ per ogni $r \in \mathbb{R}$ e quindi $Tf \in V_{\mathbb{R}}^+$ e quindi l'applicazione è ben definita. Sia ora $f \in V^+$ tale che $(Tf)(x) = \operatorname{Re} f(x) = 0$ per ogni x, allora

$$\operatorname{Im} f(x) = -\operatorname{Re} if(x) = -\operatorname{Re} f(ix) = 0$$

e dunque $f = 0$ e T è iniettiva.

Consideriamo ora una qualunque $g : V \to \mathbb{R}$ $\mathbb{R}$-lineare, possiamo allora definire $g' : V \to \mathbb{C}$ in modo tale che

$$g'(x) = g(x) - ig(ix)$$

Chiaramente g' è ancora $\mathbb{R}$-lineare mentre $g'(ix) = g(ix) - ig(-x) = (-i)ig(ix) + ig(x) = ig'(x)$ e quindi g' è $\mathbb{C}$-lineare. Infine $Tg' = g$ e quindi T è biettiva con

$$\left(T^{-1}f\right)(x) = f(x) - if(ix)$$

∎

Una proprietà utile della funzione T che sarà utilizzata in seguito è la seguente

Proposizione 1.5.3. *Dato uno spazio vettoriale complesso X e una funzione lineare $f : X \to \mathbb{C}$ allora per ogni $x \in X$ per cui $f(x) \neq 0$*

$$(Tf)\left(\frac{x}{f(x)}\right) = 1$$

$$(Tf)\left(\frac{ix}{f(x)}\right) = 0$$

Dimostrazione. Per ogni $f \in X^+$ si ha

$$f(x) = (Tf)(x) - i(Tf)(ix)$$

Dunque se $f(x) \neq 0$ allora basta porre $\lambda = f(x)/\left|f(x)\right| \in \mathbb{C}$ e per la linearità di f avremo così che

$$\left|f(x)\right| = \frac{f(x)}{\lambda} = f\left(\frac{x}{\lambda}\right) = (Tf)\left(\frac{x}{\lambda}\right) - i(Tf)\left(\frac{ix}{\lambda}\right)$$

ma $\left|f(x)\right| \in \mathbb{R}$ e Tf ha valori reali, dunque segue immediatamente che

$$(Tf)\left(\frac{x}{\lambda}\right) = \left|f(x)\right|(Tf)\left(\frac{x}{f(x)}\right) = \left|f(x)\right|$$
$$\left|f(x)\right|(Tf)\left(\frac{ix}{f(x)}\right) = 0$$

■

Dato uno spazio vettoriale reale V non è sempre possibile trovare uno spazio vettoriale complesso W tale che $W_{\mathbb{R}} = V$. Infatti se W è finitamente generato e ha dimensione n su $\mathbb{C}$ allora $W_{\mathbb{R}}$ ha dimensione $2n$ su $\mathbb{R}$, quindi se V ha dimensione dispari non è possibile trovare un siffatto spazio.

D'altronde W esiste se e solo se esiste un operatore $\hat{\imath} \in \mathcal{L}(V,V)$ tale che

$$(\hat{\imath})^2(x) = \hat{\imath}\left[\hat{\imath}(x)\right] = -x \ \forall x \in V \tag{1.5.1}$$

in tal caso possiamo definire su V un prodotto esterno $* : \mathbb{C} \times V \to V$ in modo tale che

$$(a+ib) * v = av + b\hat{\imath}(v)$$

che soddisfa tutti gli assiomi degli spazi vettoriali. Dimostreremo solo l'ultima in quanto le prime sono semplici da dimostrare. Siano $a,b,c,d \in \mathbb{R}$ e $v \in V$ allora per la linearità di $\hat{\imath}$

$$\begin{aligned}\left[(a+ib)*\left((c+id)*v\right)\right] &= (a+bi)*\left[cv + d\hat{\imath}(v)\right]\\ &= acv + ad\hat{\imath}(v) + bc\hat{\imath}(v) + bd\hat{\imath}[\hat{\imath}(x)]\\ &= (ac-bd)v + (ad+bc)\hat{\imath}(v)\\ &= [(a+ib)(c+id)]*v\end{aligned}$$

Se V è uno spazio vettoriale reale finitamente generato di dimensione $2k$ pari allora, presa una sua qualunque base $e_1, e_2, \ldots, e_{2k}$ possiamo sempre definire $\hat{\imath}$ ponendo per ogni $1 \leq j \leq k$

$$\hat{\imath}(e_{2j}) = e_{2j-1}$$
$$\hat{\imath}(e_{2j-1}) = -e_{2j}$$

cambiando la base cambia anche $\hat{\imath}$ e quindi la struttura di spazio vettoriale complesso.

1.6 Operatori lineari

Dati due spazi vettoriali V e W su uno stesso campo $\mathbb{K}$ definiamo un **operatore lineare** su V a valori in W come una coppia (D, L) dove D è un sottospazio vettoriale di V ed L è un'applicazione lineare definita su D e a valori in W.

Per comodità di notazione rappresenteremo un operatore lineare (D, L) solamente con il simbolo dell'applicazione lineare L e ponendo $D = \operatorname{dom} L$ quando vogliamo specificarne il dominio, tenendo sempre ben in mente che un operatore è identificato anche dal suo dominio. In altre parole dati due operatori L, R su V a valori in W abbiamo

$$L = R \Leftrightarrow \operatorname{dom} L = \operatorname{dom} R \text{ e } L(x) = R(x)\ \forall x \in \operatorname{dom} L = \operatorname{dom} R$$

Se indichiamo con $\mathcal{UL}(V, W)$ l'insieme di tutti gli operatori lineari da V in W possiamo definire la somma ed il riscalamento di operatori lineari nella seguente maniera:

$$\operatorname{dom}(L + R) = \operatorname{dom} L \cap \operatorname{dom} R \qquad (L + R)(x) = L(x) + R(x)$$
$$\operatorname{dom}(\alpha L) = \operatorname{dom} L \qquad (\alpha L)(x) = \alpha L(x)$$

per ogni $L, R \in \mathcal{UL}(V, W)$, $\alpha \in \mathbb{K}$. Si tenga però presente che la somma in $\mathcal{UL}(V, W)$ non è invertibile e dunque lo spazio degli operatori lineari **non è uno spazio vettoriale**. Infatti l'unico elemento neutro è la funzione identicamente nulla che indicheremo con 0, chiaramente $\operatorname{dom} 0 = V$ ma preso un qualunque $L \in \mathcal{UL}(V, W)$ con $\operatorname{dom} L < V$ si ha $\operatorname{dom}(L - L) = \operatorname{dom} L \neq V = \operatorname{dom} 0$ e così $L - L \neq 0$.

Dati anche $L \in \mathcal{UL}(V, W)$, $S \in \mathcal{UL}(W, U)$ definiamo la composizione $S \circ L \in \mathcal{UL}(V, U)$ in modo tale che

$$\operatorname{dom}(S \circ L) = \{x \in \operatorname{dom} L \mid L(x) \in \operatorname{dom} S\} \qquad (S \circ L)(x) = S[L(x)]$$

Per ogni operatore lineare $L \in \mathcal{UL}(V, W)$ definiamo il **grafico** di L il seguente insieme:

$$\mathcal{G}(L) = \{(x, y) \in V \times W \mid x \in \operatorname{dom} L \text{ e } L(x) = y\}$$

Chiaramente $\mathcal{G}(L)$ è un sottospazio vettoriale di $V \times W$ con la struttura standard di spazio vettoriale, inoltre vale il seguente risultato di caratterizzazione di immediata verifica:

Proposizione 1.6.1. *Dato un sottospazio vettoriale $G \leq V \times W$ allora esiste $L \in \mathcal{UL}(V, W)$ tale che $\mathcal{G}(L) = G$ se e solo se per ogni $y \in W$ tale che $(0, y) \in G$ si ha $y = 0$. Inoltre $L = R$ se e solo se $\mathcal{G}(L) = \mathcal{G}(R)$.*

Un operatore $E \in \mathcal{UL}(V, W)$ è un'**estensione** di $L \in \mathcal{UL}(V, W)$ se e solo se $\operatorname{dom} L \subseteq \operatorname{dom} E$ ed $L(x) = E(x)$ per ogni $x \in \operatorname{dom} L$, ed in tal caso scriveremo $L \subseteq E$.

Proposizione 1.6.2. *Dati due operatori $L, R \in \mathcal{UL}(V, W)$ abbiamo $L \subseteq R$ se e solo se $\mathcal{G}(L) \subseteq \mathcal{G}(R)$.*

1.7 Insieme quoziente

Definizione 1.7.1. Dato un insieme A non vuoto una relazione binaria $\Re$ definita sugli elementi di A è una **relazione di equivalenza** se e solo se soddisfa le seguenti proprietà:

- **riflessività**: $a\Re a$ per ogni $a \in A$;
- **simmetria**: $a\Re b$ se e solo se $b\Re a$;
- **transitività**: se $a\Re b$ e $b\Re c$ allora $a\Re c$.

Inoltre se $\Re$ è una relazione di equivalenza su A definiamo

$$[a]_{\Re} = \{b \in A \mid b\mathfrak{a}\} \qquad \forall a \in A$$
$$A/\Re = \left\{[a]_{\Re} \;\middle|\; a \in A\right\}$$

l'insieme $A/\Re$ sarà detto **insieme quoziente** di A rispetto alla relazione $\Re$.

A partire da una relazione di equivalenza $\Re$ su A è sempre possibile definire la seguente funzione biettiva:

$$\pi_{\Re} : a \in A \to [a]_{\Re} \in A/\Re \tag{1.7.1}$$

Ora che abbiamo introdotto i concetti di spazio vettoriale e di insieme quoziente mettiamoli insieme attraverso il seguente risultato:

Teorema 1.7.2. *Dati uno spazio vettoriale V su $\mathbb{K}$ ed una relazione di equivalenza $\sim$ su V allora le seguenti affermazioni sono equivalenti:*

1. *$V/\sim$ è uno spazio vettoriale su $\mathbb{K}$ e $\pi_{\sim}$ è lineare;*
2. *per ogni $x, y, z \in V$, $\lambda \in \mathbb{K}$ per cui $x \sim y$ si ha $x + z \sim y + z$ e $\lambda x \sim \lambda y$;*
3. *per ogni $x, y \in V$, $\lambda \in \mathbb{K}$ si ha $x \sim y$ se e solo se $\lambda(x - y) \in [0]_{\sim}$ per ogni $\lambda \in \mathbb{K}$.*

Dimostrazione. Osserviamo che se $V/\sim$ è uno spazio vettoriale allora $\pi_{\sim}$ è lineare se e solo se per ogni $x, y \in V$, $\lambda \in \mathbb{K}$ si ha

$$[x]_{\sim} + [y]_{\sim} = [x + y]_{\sim} \qquad\qquad \lambda[x]_{\sim} = [\lambda x]_{\sim}$$

e quindi dal punto 1 segue immediatamente il punto 2. Viceversa dal punto 2 segue immediatamente il punto 1 in quanto le definizioni di somma e prodotto esterno in $V/\sim$ come sopra sono ben definite grazie proprio alle ipotesi fatte nel punto 2.

Supponiamo ora che valga il punto 2 e prendiamo $a, b \in V$ generici, allora

$$a \sim b \Leftrightarrow \lambda a \sim \lambda b\ \forall \lambda \in \mathbb{K} \Leftrightarrow \lambda(a - b) \sim 0\ \forall \lambda \in \mathbb{K} \Leftrightarrow \lambda(a - b) \in [0]_{\sim}\ \forall \lambda \in \mathbb{K}$$

e il punto 3 è verificato. Viceversa se supponiamo vero il punto 3 allora si avrà

$$x + z \sim y + z \Leftrightarrow \lambda(x + z) - \lambda(y + z) = \lambda(x - y) \in [0]_{\sim}\ \forall \lambda \in \mathbb{K} \Leftrightarrow x \sim y$$
$$x \sim y \Leftrightarrow \lambda(x - y) \in [0]_{\sim}\ \forall \lambda \in \mathbb{K} \Rightarrow \lambda\mu(x - y) \in [0]_{\sim}\ \forall \lambda \in \mathbb{K} \Leftrightarrow \mu x \sim \mu y$$

∎

In particolare si osserva che se $V/\sim$ è uno spazio vettoriale e $\pi_{\sim}$ è lineare allora $[0]_{\sim}$ è un sottospazio vettoriale di V ed $[x]_{\sim} = x + [0]_{\sim}$. Quindi da una tale relazione di equivalenza possiamo sempre ricondurci ad un sottospazio vettoriale che la genera. In tal caso diremo che $V/\sim$ è uno **spazio vettoriale quoziente**.

Possiamo al contempo lavorare anche a ritroso: preso un qualunque sottospazio vettoriale $N \leq V$ possiamo sempre trovare un'unica relazione di equivalenza $\sim_N$ su V in modo tale che $N = [0]_{\sim_N}$ e che $pi_{\sim_N}$ sia lineare. Più nel dettaglio avremo che

$$x \sim_N y \Leftrightarrow x - y \in N$$
$$[x]_{\sim_N} = x + N$$

e non è difficile verificare che soddisfano tutte le ipotesi richieste. In tal caso scriveremo V/N e π_N al posto di $V/\sim_N$ e $\pi_{\sim_N}$.

1.8 Relazioni d'ordine

Dato un qualunque insieme non vuoto A una relazione binaria $\leq$ su A è un **preordine** se e solo se soddisfa le seguenti proprietà:

$$x \leq x \qquad \text{(Riflessività)}$$
$$x \leq y \text{ e } y \leq z \Rightarrow x \leq z \qquad \text{(Transitività)}$$

e se $\leq$ soddisfa anche la seguente proprietà

$$x \leq y \text{ e } y \leq x \Rightarrow x = y \qquad \text{(Asimmetria)}$$

allora $\leq$ è un **ordine** sull'insieme A ed in tal caso la coppia $(A, \leq)$ sarà detta **insieme ordinato**. Se $\leq$ è una relazione d'ordine (non un preordine) possiamo definire la relazione di ordine stretto associato $<$ nella seguente maniera:

$$x < y \Leftrightarrow x \leq y \text{ e } x \neq y$$

Un preordine su A è detto **totale** se e solo se per ogni coppia di elementi $x, y \in A$ si ha $x \leq y$ oppure $y \leq x$. L'insieme $\mathbb{R}$ è un esempio di insieme totalmente ordinato, mentre l'insieme delle parti $\mathcal{P}(A)$ con la relazione di inclusione quasi mai è totale. In un insieme totalmente ordinato possiamo quindi sempre definire il minimo e il massimo di due o più elementi, per i quali utilizzeremo la seguente notazione:

$$x \vee y = \max\{x, y\}$$
$$x \wedge y = \min\{x, y\}$$

Se la relazione d'ordine non è totale diremo semplicemente che è parziale, e nel seguito incontreremo spesso insiemi solo parzialmente ordinati e non totalmente. Per poter operare su di essi abbiamo bisogno di indebolire il concetto di ordine totale con una versione più debole.

Definizione 1.8.1. Un insieme non vuoto A con una relazione di preordine $\leq$ è detto **diretto** se e solo se per ogni $x, y \in A$ esiste $z \in A$ tale che $x \leq z$ e $y \leq z$.

Per esempio se prendiamo due insiemi totalmente ordinati A, B allora il loro prodotto cartesiano $A \times B$ con la seguente relazione d'ordine

$$(a, x) \leq (b, y) \Leftrightarrow a \leq b \text{ e } x \leq y$$

è diretto pur non essendo necessariamente totalmente ordinato.

Un sottoinsieme non vuoto B di un insieme diretto A sarà detto **cofinale** se e solo se per ogni $a \in A$ esiste $b \in B$ tale che $a \leq b$.

Prendiamo adesso un insieme generico U non vuoto dotato di una relazione d'ordine parziale $\preceq$, un suo elemento u è detto **massimale** se e solo se per ogni $v \in U$ vale l'implicazione

$$u \preceq v \Rightarrow u = v$$

ovvero non esistono elementi in U strettamente più grandi di u. Diremo inoltre che u è un **massimo** se e solo se $v \preceq u$ per ogni $v \in U$. Il massimo è anche un elemento massimale e se un insieme possiede un elemento massimo u allora non esistono altri elementi massimali diversi da u. Ovviamente un insieme potrebbe contenere diversi elementi massimali pur non possedendo alcun massimo.

Un qualunque sottoinsieme non vuoto W di U è una **catena** se e solo se due suoi elementi a e b sono sempre confrontabili cioè $a \preceq b$ oppure $b \preceq a$, un **maggiorante** di W è un elemento $u \in U$ tale che $a \preceq u$ per ogni $a \in W$.

Un insieme U è detto **induttivo** se e solo se ogni sua catena possiede almeno un maggiorante. Un noto e fondamentale risultato della teoria degli insiemi, che useremo in varie dimostrazioni, fornisce un criterio per stabilire l'esistenza di elementi massimali di insiemi ordinati

Lemma 1.8.2 (Zorn)**.** *Ogni insieme induttivo possiede elementi massimali.*

Il lemma di Zorn risulta fondamentale per la maggior parte dell'analisi, difatti una buona parte dei risultati presenti in queste dispense lo utilizza. Un esempio di applicazione del lemma di Zorn si può osservare nel seguente risultato sugli spazi vettoriali

Proposizione 1.8.3. *Sia X spazio vettoriale e $M \leq X$, allora esiste un unico $N \leq X$ tale che $X = M \oplus N$. Inoltre esiste un isomorfismo (applicazione lineare invertibile) tra N e X/M.*

Dimostrazione. Se $M = X$ allora $N = \{0\}$, analogo se $M = \{0\}$.

Sia allora $x \in X \setminus M$ dunque $\operatorname{span}\{x\} \cap M = \{0\}$ e quindi $\operatorname{span}\{x\}$ appartiene alla classe

$$\mathcal{N} = \{A \leq X \mid A \cap M = \{0\}\}$$

che risulta quindi non vuota, dimostriamo che $\mathcal{N}$ è induttiva.

Consideriamo una qualunque sottoclasse $\mathcal{F} \subseteq \mathcal{N}$ non vuota totalmente ordinata rispetto all'inclusione $\subseteq$, poniamo

$$U = \bigcup_{A \in \mathcal{F}} A$$

vogliamo dimostrare che $U \in \mathcal{N}$, in tal caso infatti U risulta essere un maggiorante per $\mathcal{F}$. Presi $u, v \in U$ esisteranno $A, B \in \mathcal{F}$ per cui $u \in A$ e $v \in B$, ma $\mathcal{F}$ è totalmente ordinato dunque $A \subseteq B$ oppure $B \subseteq A$. Poiché sia A che B sono sottospazi avremo che $u + v$ apparterrà ad uno dei due e quindi $a + b \in U$. Con lo stesso ragionamento si verifica che se $u \in U$ e $\alpha \in \mathbb{K}$ allora $\alpha u \in U$ e quindi U è un sottospazio vettoriale di X. Infine osserviamo che

$$U \cap \operatorname{span}\{x\} = \left(\bigcup_{A \in \mathcal{F}} A \right) \cap \operatorname{span}\{x\} = \bigcup_{A \in \mathcal{F}} (A \cap \operatorname{span}\{x\}) = \{0\}$$

e quindi $U \in \mathcal{N}$ ed $\mathcal{N}$ è induttivo.

Per il lemma di Zorn $\mathcal{N}$ ammette un elemento massimale N che soddisfa l'uguaglianza $M \cap N = \{0\}$. Se per assurdo non fosse $X = M \oplus N$ allora esisterebbe un elemento $x \in X \setminus \{0\}$ che non può essere scritto come somma di elementi di M e N da cui segue immediatamente che $M \cap \left(N + \operatorname{span}\{x\}\right) = \{0\}$ assurdo per la massimalità di N.

Definiamo adesso l'applicazione

$$f : n \in N \to [n]_\sim \in X/M$$

che è chiaramente lineare e ben definita. Siano ora $n, n' \in N$ tali che $[n]_\sim [n']_\sim$ allora $n - n' \in N \cap M = \{0\}$ quindi $n = n'$ ed f è iniettiva. Per far vedere la suriettività osserviamo che per ogni $x \in X$ esistono $m \in M, n \in N$ tali che $x = m + n$ dunque $x \sim n$ e $f(n) = [x]_\sim$ e la suriettività è così dimostrata. ■

Il lemma di Zorn viene usato anche per garantire l'esistenza di una **base di Hamel** per ogni spazio vettoriale, ovvero di un sottoinsieme P di uno spazio vettoriale V su $\mathbb{K}$ tale che

1. $V = \operatorname{span} P$;

2. Per ogni $x_1, \dots, x_k \in P$ e per ogni $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tali che $\lambda_1 x_1 + \cdots + \lambda_k x_k = 0$ si ha $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

o equivalentemente V è somma diretta dei sottospazi $\operatorname{span}\{x\} = \left\{\lambda x \mid \lambda \in \mathbb{K}\right\}$ al variare di $x \in P$.

Teorema 1.8.4. *Ogni spazio vettoriale possiede una base di Hamel.*

Dimostrazione. Diremo che un qualunque sottoinsieme non vuoto di V è indipendente se e solo se soddisfa il punto 2 della definizione di base di Hamel. Dimostriamo adesso che la famiglia di tutti i sottoinsiemi indipendenti di V è induttiva.

Presa allora una qualunque famiglia $\left\{Q_i\right\}_{i \in I}$ di sottoinsiemi indipendenti totalmente ordinata rispetto all'inclusione, prendiamo adesso $x_1, x_2, \dots, x_k \in Q = \bigcup_{i \in I} Q_i$ e $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tali che $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k = 0$. Ma allora esisteranno $i_1, \dots, i_k \in I$ tali che $x_j \in Q_{i_j}$ per ogni j e poiché la famiglia è totalmente ordinata esisterà $i \in I$ tale che Q_i contiene tutti gli x_j. Poiché Q_i è indipendente avremo che $\lambda_j = 0$ per ogni j, abbiamo così mostrato che anche Q è indipendente e contiene tutti gli Q_i al variare di $i \in I$.

Per il lemma di Zorn esisterà un sottoinsieme P indipendente massimale in V, prendiamo adesso un generico $x \in V \setminus P$. Per massimalità l'insieme $P \cup \{x\}$ non può essere indipendente, ovvero esistono $x_1, \dots, x_k \in P$ e $\lambda, \lambda_1, \dots, \lambda_k \in \mathbb{K}$ tali che

$$\lambda x + \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k = 0$$

Se fosse $\lambda = 0$ allora P non sarebbe indipendente in quanto almeno un λ_i non è nullo, quindi $\lambda \neq 0$ e

$$x = \frac{\lambda_1}{\lambda} x_1 + \cdots + \frac{\lambda_k}{\lambda} x_k$$

ovvero $x \in \operatorname{span} P$ e P è una base di Hamel. ■

1.9 Somme infinite

Definizione 1.9.1. Dato un insieme non vuoto A e una funzione $f : A \to [0,+\infty)$ si definisce

$$\sum_{i\in A} f(i) = \sup\left\{ \sum_{i\in A'} f(i) \,\middle|\, A' \in \mathcal{F}(A)\setminus\{\emptyset\}\right\}$$

Per comodità di notazione poniamo $A_+^\infty = F(A,[0,+\infty))$. Dalla definizione segue immediatamente che se $f \in A_+^\infty$ e $\alpha \geq 0$ allora $\sum_{i\in A}[\alpha f(i)] = \alpha\sum_{i\in A} f(i)$, dimostriamo invece che

$$\sum_{i\in A}[f(i)+g(i)] = \sum_{i\in A} f(i) + \sum_{i\in A} g(i) \tag{1.9.1}$$

Se A fosse finito allora l'uguaglianza è banale, dunque poiché l'estremo superiore della somma è minore o uguale della somma degli estremi superiori segue la disuguaglianza $\leq$. Viceversa per ogni $\varepsilon > 0$ esistono $A', A'' \in \mathcal{F}(A)\setminus\{\emptyset\}$ tali che $\sum_{i\in A'} f(i) > \sum_{i\in A} f(i) - \varepsilon$ e $\sum_{i\in A''} g(i) > \sum_{i\in A} g(i) - \varepsilon$, allora

$$\sum_{i\in A} f(i) + \sum_{i\in A} g(i) - 2\varepsilon < \sum_{i\in A'} f(i) + \sum_{i\in A''} g(i) \leq \sum_{i\in A'\cup A''}[f(i)+g(i)] \leq \sum_{i\in A}[f(i)+g(i)]$$

e per l'arbitrarietà di ε si ha la (1.9.1).

Sempre dalla definizione osserviamo che le sommatorie generalizzate sono monotone, ovvero se $f(i) \leq g(i)$ per ogni $i \in A$ allora $\sum_{i\in A} f(i) \leq \sum_{i\in A} g(i)$.

Proposizione 1.9.2. *Per ogni $f \in A_+^\infty$ per cui $L = \sum_{a\in A} f(a) < +\infty$ esiste un sottoinsieme finito o numerabile A' di A tale che per ogni $i \in A\setminus A'$ si ha $f(i) = 0$.*

Dimostrazione. Definiamo per ogni $n \in \mathbb{N}$

$$A_n = \left\{ i \in A \,\middle|\, f(i) \geq \frac{1}{n}\right\}$$

allora A_n è una successione crescente di sottoinsiemi finiti di A con $\#(A_n) \leq nL$. Dunque $f(i) = 0$ per ogni $i \notin \bigcup_{n\in\mathbb{N}} A_n = A'$ e la tesi segue dall'osservazione che A' è chiaramente finito o numerabile. ■

Teorema 1.9.3 (Beppo-Levi)**.** *Data una successione crescente di funzioni $f_n \in A_+^\infty$ convergente puntualmente ad $f \in A_+^\infty$ si ha*

$$\lim_{n\to+\infty}\sum_{i\in A} f_n(i) = \sup_{n\in\mathbb{N}}\sum_{i\in A} f_n(i) = \sum_{i\in A} f(i)$$

Dimostrazione. Poiché $f_n(i) \leq f_{n+1}(i)$ allora $f_n(i) \leq f(i)$ e per monotonia si ottiene la disuguaglianza $\leq$. Se A fosse finito allora la somma è finita e quindi per la continuità della somma si ha la tesi, preso allora un qualunque sottoinsieme A' di A finito si avrebbe

$$\sum_{i\in A'} f(i) \leq \lim_{n\to+\infty}\sum_{i\in A'} f(i) \leq \lim_{n\to+\infty}\sum_{i\in A} f(i)$$

e poiché l'ultimo membro non dipende da A' passando all'estremo superiore si ha la tesi. ■

Corollario 1.9.4. *Se $A = \mathbb{N}$ per ogni $f \in A_+^\infty$ si ha*

$$\lim_{n\to+\infty} \sum_{i=1}^{n} f(i) = \sum_{i=1}^{+\infty} f(i) = \sum_{a\in A} f(a)$$

Dimostrazione. La successione $f_n \in A_+^\infty$ definita come

$$f_n(i) = \begin{cases} f(i) & \text{se } i \leq n \\ 0 & \text{se } i > n \end{cases}$$

soddisfa tutte le ipotesi del teorema di Beppo-Levi. ■

Teorema 1.9.5 (Fatou). *Se $f_n \in A_+^\infty$ è una successione generica allora*

$$\sum_{i\in A} \liminf_{n\to+\infty} f_n(i) \leq \liminf_{n\to+\infty} \sum_{i\in A} f_n(i)$$

se inoltre esiste $g \in A_\infty^+$ per cui $\sum_{i\in A} g(i) < +\infty$ e $f_n \leq g$ per ogni $n \in \mathbb{N}$ allora

$$\sum_{i\in A} \limsup_{n\to+\infty} f_n(i) \geq \limsup_{n\to+\infty} \sum_{i\in A} f_n(i)$$

Dimostrazione. Basta applicare Beppo-Levi alla successione $g_n(i) = \inf_{m\geq n} f_m(i)$. Per la seconda disuguaglianza basta applicare la prima alla successione $g_n = g - f_n$ e quindi sottrarre ■

Capitolo 2

Spazi topologici

2.1 Spazi topologici

In questa sezione richiamiamo alcuni concetti di topologia generale necessari alla trattazione degli argomenti successivi.

I vari risultati saranno presentati senza dimostrazione, per maggiori informazioni si veda [10].

Definizione 2.1.1. Sia X insieme non vuoto e τ contenente alcuni dei sottoinsiemi di X. Allora la struttura (X,τ) è uno **spazio topologico** se e solo se valgono le seguenti affermazioni:

- $\emptyset, X \in \tau$;
- Se $A, B \in \tau$ allora $A \cap B \in \tau$;
- Se $\mathscr{A} \subseteq \tau$ sottoinsieme non vuoto di elementi di τ allora $\bigcup_{A \in \mathscr{A}} A \in \tau$.

Gli elementi di τ sono detti **aperti** di X, mentre $C \subseteq X$ è **chiuso** se e solo se $X \setminus C \in \tau$.

Se (X,τ) è uno spazio topologico e Y è un sottoinsieme non vuoto di X allora Y sarà automaticamente della **topologia indotta**, in altre parole diremo che $V \subseteq Y$ è aperto in Y se e solo se esiste un aperto $U \in \tau$ tale che

$$V = U \cap Y$$

Date due topologie τ ed η sullo stesso insieme X diremo che la topologia τ è **più fine** della topologia η, e di conseguenza η è **meno fine** di τ, se e solo se $\eta \subseteq \tau$ ovvero se gli aperti nella topologia η sono aperti anche nella topologia τ.

Per ogni sottoinsieme I di X la **chiusura** di I, che viene indicata con $\mathrm{Cl}(I)$, è il più piccolo insieme chiuso contenente I, mentre l'**interno** di I, indicato con $\mathrm{Int}(I)$, è il più grande aperto contenuto in I. In formule

$$\mathrm{Cl}(I) = \bigcap \{C \supseteq I \mid C \text{ chiuso}\}$$
$$\mathrm{Int}(I) = \bigcup \{U \subseteq I \mid U \text{ aperto}\}$$

e quindi

$$X \setminus \mathrm{Cl}(I) = \mathrm{Int}(X \setminus I)$$
$$X \setminus \mathrm{Int}(I) = \mathrm{Cl}(X \setminus I)$$

Proposizione 2.1.2. *Dato uno spazio topologico* (X,τ) *e* $A \subseteq X$ *allora* $x \in \mathrm{Cl}(A)$ *se e solo se per ogni aperto* U *contenente* x *si ha* $A \cap U \neq \emptyset$.

Invece $x \in \mathrm{Int}(A)$ *se e solo se esiste un aperto* U *contenente* x *tale che* $U \subseteq A$.

Definizione 2.1.3. Uno spazio topologico X è di **Hausdorff** se e solo se per ogni $x, y \in X$ con $x \neq y$ esistono due aperti A, B di X tali che $x \in A$, $y \in B$ e $A \cap B = \emptyset$.

Uno spazio topologico X si dice **regolare** se e solo se per ogni chiuso $C \subseteq X$ e per ogni $x \in X$ non appartenente a C esistono due aperti A e B di X contenenti rispettivamente x e C tali che $A \cap B = \emptyset$.

Proposizione 2.1.4. *In uno spazio topologico di Hausdorff i punti sono chiusi. In altre parole se* (X,τ) *è uno spazio topologico di Hausdorff allora per ogni* $x \in X$ *l'insieme* $\{x\}$ *è chiuso in* X.

In generale non vale il viceversa, ovvero esistono spazi topologici non di Hausdorff dove tutti i punti sono chiusi, a meno che lo spazio non sia anche regolare. Inoltre uno spazio regolare può non essere uno spazio di Hausdorff in quanto i punti non sono necessariamente chiusi.

2.2 Limiti e continuità

Definizione 2.2.1. Dato uno spazio topologico (X,τ), $Z \subseteq X$ e $x \in X$ diremo che x è un **punto di accumulazione** per Z se e solo se per ogni intorno $I \subseteq X$ di x si ha $I \cap Z \neq \emptyset$.

Un punto $x \in Z$ che però non è un punto di accumulazione per Z è detto **punto isolato**.

Osservazione. Se $Z = X$ allora $x \in X$ è un punto di accumulazione (per X) se e solo se $\{x\}$ non è aperto in X.

Osserviamo anche che, presi uno spazio topologico X, un sottoinsieme non vuoto Z di X ed un elemento $x \in Z$ le seguenti affermazioni sono equivalenti:

- x è un punto di accumulazione per Z visto come sottoinsieme di X;
- x è un punto di accumulazione in Z visto come spazio topologico con la topologia indotta.

In analisi il limite di una funzione in un punto dà un quadro generale di come la funzione si avvicina al punto senza però considerare il valore che la funzione assume in quel punto. Per questo motivo possiamo calcolare il limite di una funzione tra spazi topologici solamente nei punti di accumulazione del suo dominio, in quanto dobbiamo essere in grado sia di considerare i punti vicini al punto limite sia escludere tale punto.

Definizione 2.2.2. Dati due spazi topologici X, Y, un sottoinsieme non vuoto $Z \subseteq X$, una funzione $f : Z \to Y$ e due punti $x \in X, y \in Y$ con x punto di accumulazione per Z, diremo che la funzione $f(z)$ **tende** a y al tendere di $z \in Z$ a x, e scriveremo $f(z) \to y$, se e solo se per ogni aperto $U \subseteq Y$ contenente y esiste un altro aperto $V \subseteq X$ contenente x tali che

$$V \cap Z \setminus \{x\} \subseteq f^{-1}(U) \tag{2.2.1}$$

il valore y sarà allora un **limite** di $f(z)$ al tendere di $z \in Z$ ad x.

In altre parole $f(z) \to y$ al tendere di $z \in Z$ ad x se e solo se per ogni aperto U di Y contenente y esiste un aperto U' di X contenente x tale che $f(z) \in U$ per ogni $z \in U' \cap Z$ **diverso** da x. Il fatto che z sia diverso da x è importante in quanto permette di slegare l'esistenza o meno del limite dal valore che f assume in x.

In generale una funzione può avere più di un limite in x o non averne nessuno. Se però lo spazio di arrivo Y è uno spazio di Hausdorff allora il limite, se esiste, è unico e in tal caso saremo autorizzati a scrivere

$$y = \lim_{\substack{\bar{x}\to x \\ \bar{x}\in Z}} f(\bar{x})$$

o nel caso in cui $Z = X$

$$y = \lim_{\bar{x}\to x} f(\bar{x})$$

Proposizione 2.2.3. *Dati due spazi topologici X, Y, un sottoinsieme non vuoto Z di X e $x \in X$ punto di accumulazione di Z. Allora presi un insieme chiuso $C \subseteq Y$, una funzione $f : Z \to Y$ e un valore $y \in Y$ tali che $f(z) \to y$ al tendere di $z \in Z$ ad x e $f(z) \in C$ per ogni $z \in Z \setminus \{x\}$ avremo che $y \in C$.*

Dimostrazione. Se per assurdo $y \notin C$ allora $y \in Y \setminus C$ il quale è aperto in Y, dunque esisterà $U \subseteq X$ aperto contenente x tale che $Z \cap U \setminus \{x\} \subseteq f^{-1}(Y \setminus C)$. Dato che x è un punto di accumulazione per Z esisterà sempre $\bar{z} \in Z$ diverso da x tale che $f(\bar{z}) \in Y \setminus C$, il che va in contraddizione con le nostre ipotesi iniziali. ■

Teorema 2.2.4 (Cambio di variabile nei limiti)**.** *Siano $x \in X$, $y \in Y$ ed $f : X \to Y$ tale che $f(\bar{x}) \to y$ al tendere di $\bar{x}$ a x, presa allora una funzione $p : E \to X$ e un punto di accumulazione e di E che soddisfano le seguenti condizioni*

- *$p(\bar{e})$ tende a x al tendere di $\bar{e}$ a e;*
- *Esiste un aperto $U \subseteq E$ contenente e tale che $p(\bar{e}) \neq x$ per ogni $\bar{e} \in U \setminus \{e\}$.*

avremo che la composizione $f[p(\bar{e})]$ tende a y al tendere di $\bar{e}$ ad e.

Dimostrazione. Per ogni aperto $I \subseteq Y$ contenente y esistono due aperti $x \in J \subseteq X$ e $e \in K \subseteq E$ tali che

$$J \setminus \{x\} \subseteq f^{-1}(I)$$
$$K \setminus \{e\} \subseteq p^{-1}(J)$$

Sia ora $U \subseteq E$ l'aperto come nella definizione, allora avremo che $K \cap U \setminus \{e\} \subseteq p^{-1}(J \setminus \{x\})$, difatti se non valesse questa inclusione esisterebbe $u \in U$ diverso da e per cui $p(u) = x$ il che è assurdo. Dato che $U \cap K$ è un aperto contenente e avremo la tesi. ■

Grazie a questo risultato è possibile effettuare un cambio di variabili all'interno dei limiti. In particolare se Y è uno spazio di Hausdorff e p soddisfa le ipotesi della proposizione precedente allora vale la seguente implicazione

$$\lim_{\bar{x}\to x} f(\bar{x}) = y \Rightarrow \lim_{\bar{e}\to e} f[p(\bar{e})] = y$$

Dato uno spazio topologico X definiamo una **net** x_a in X semplicemente come una qualunque applicazione definita su un qualche insieme diretto $(A, \preceq)$ a valori in X, in particolare se $A = \mathbb{N}$ con l'ordine usuale diremo che x_n è semplicemente una **successione** in X.

Vogliamo adesso definire su A una topologia in modo da poter definire il limite di una net semplicemente come un caso particolare di un limite di funzioni tra spazi topologici.

Preso un qualunque insieme diretto A con preordine $\preceq$ definiamo innanzitutto $\tilde{A} = A \cup \{+\infty\}$, dove con il simbolo $+\infty$ indichiamo semplicemente un qualunque elemento che non appartiene ad A, e poniamo $x \preceq +\infty$ per ogni $x \in A$. In questo modo l'insieme $\tilde{A}$ continua ad essere un insieme diretto rispetto al preordine $\preceq$. Per ogni $a \in A$ poniamo

$$S_a = \left\{x \in \tilde{A} \;\middle|\; a \preceq x\right\}$$

chiaramente avremo sempre che $a \in S_a$ ed $+\infty \in S_a$ per ogni $a \in A$, in particolare $A \cap S_a \neq \emptyset$. Diremo ora che un sottoinsieme U di $\tilde{A}$ è aperto se e solo se $U = \emptyset$ oppure esiste $a \in A$ tale che $S_a \subseteq U$, per dimostrare che questa è effettivamente una topologia dobbiamo solamente dimostrare che l'intersezione di aperti è ancora un aperto.

Per ogni coppia di aperti non vuoti $U, V \subseteq \tilde{A}$ esistono $a, b \in A$ tali che $S_a \subseteq U$ ed $S_b \subseteq V$. Per definizione di insieme diretto esiste $c \in AQ$ tale che $a \preceq c$ ed $b \preceq c$, per la proprietà transitiva avremo che $S_c \subseteq S_a$ ed $S_c \subseteq S_b$. Dunque

$$S_c \subseteq S_a \cap S_b \subseteq U \cap V$$

e quindi anche $U \cap V$ è un aperto di $\tilde{A}$ e quindi formano una topologia. Infine per quanto detto prima abbiamo che tutti gli aperti di $\tilde{A}$ non vuoti contengono $+\infty$ ed $A \cap U \neq \emptyset$ per ogni aperto U contenente $+\infty$, quindi $+\infty$ è un punto di accumulazione per l'insieme A.

Quindi le net sono a tutti gli effetti applicazioni tra spazi topologici e dunque ha perfettamente senso parlare di limite a $+\infty$. Dunque presa una net x_a se esiste almeno un valore $x \in X$ a cui la successione tende diremo che la net **converge** ad x e scriviamo semplicemente $x_a \to x$.

Proposizione 2.2.5. *Una net x_a in X definita su A converge ad un certo $x \in X$ se e solo se per ogni aperto $U \subseteq X$ contenente x esiste $a \in A$ tale che $x_b \in U$ per ogni $b \in A$ tale che $a \preceq b$.*

Anche per le net continuano a valere tutti i risultati ottenuti per i limiti di funzioni. In particolare la proposizione 2.2.3 assumerà la seguente forma per le successioni:

Proposizione 2.2.6. *Il sottoinsieme $C \subseteq X$ è chiuso se e solo se per ogni net $x_a \in C$ tale che $x_a \to x \in X$ si ha $x \in C$. Inoltre $x \in \mathrm{Cl}(C)$ se e solo se esiste una net x_a contenuta interamente in C convergente ad x.*

Dimostrazione. Se C è chiuso allora per la proposizione 2.2.3 per ogni net $x_a \in C$ convergente a qualche $x \in X$ si avrà $x \in C$. Viceversa prendiamo un qualunque punto $x \in X \setminus C$, mostriamo che esiste un aperto U contenente x per cui $C \cap U = \emptyset$. Se questo non fosse vero per ogni aperto U contenente x esisterebbe $x_U \in U \cap C$, la quale costituirebbe una net di elementi di C definita sugli aperti di X contenenti x rispetto al preordine $U \leq V \Leftrightarrow V \subseteq U$.

Per definizione di convergenza abbiamo $x_u \to x$ e per ipotesi $x \in C$ il che è assurdo. Dunque il complemento di C è unione di aperti e quindi è aperto anch'esso, perciò C è chiuso.

Per quanto riguarda l'enunciato sulla chiusura la dimostrazione è simile se si sfrutta la proposizione 2.1.2. ■

Presa una generica net x_a in X definita su qualche insieme diretto A se prendiamo un qualunque suo sottoinsieme cofinale $B \subseteq A$ la nuova net $\{x_a\}_{a \in B}$ ottenuta restringendo x_a

su B sarà detta **estratta** di x_a. Invece date adesso due net x_a, y_i definite entrambe nello stesso spazio topologico X ma definite su due insiemi diretti A, I che possono non coincidere. Diremo allora che y_i è una **subnet** di x_a se e solo se esiste un'applicazione $k : I \to A$ che soddisfa le seguenti proprietà:

1. per ogni $i, j \in I$ se $i \preceq j$ allora $k_i \preceq k_j$ (k è crescente);

2. per ogni $a \in A$ esiste $i \in I$ tale che $a \preceq k_i$ ($k(I)$ è cofinale in A),

tale che $y_i = x_{k_i}$ per ogni $i \in I$.

Osservazione. Ogni estratta è anche una subnet, ma non tutte le subnet sono estratte. Questo fatto va tenuto sempre in mente, soprattutto quando si parla di compattezza. Le estratte di successioni sono ancora successioni, mentre le subnet di successioni non sempre sono ancora successioni.

Proposizione 2.2.7. *Presi uno spazio topologico X, un punto $x \in X$ e una net $x_a \in X$ tale che ogni sua estratta possiede una subnet convergente ad x. Allora x_a converge ad x.*

Dimostrazione. Per assurdo esiste un aperto $U \subseteq X$ contenente x tale che per ogni $a \in A$ esiste $b_a \in A$ con $a \preceq b_a$ ed $x_{b_a} \notin U$. L'insieme $B = \{b_a \mid a \in A\}$ è un sottoinsieme cofinale di A e dunque $\{x_b\}_{b \in B}$ è un'estratta di x_a per cui $x_b \notin U$ per ogni $b \in B$. Per ipotesi però esiste $k : I \to B$ tale che la subnet x_{k_i} converge ad x e perciò esisterà $i \in I$ tale che $x_{k_i} \in U$. Ma essendo $k_i \in B$ si ottiene un assurdo. ■

Il seguente risultato fornisce delle condizioni per l'esistenza di subnet con particolari condizioni.

Lemma 2.2.8. *Dati un insieme X non vuoto, una net $x_a \in X$ ed una famiglia di sottoinsiemi non vuoti $\mathscr{F}$ di X tali che*

1. *per ogni $F \in \mathscr{F}$ e per ogni $a \in A$ esiste $b \succeq a$ tale che $x_b \in F$;*

2. *per ogni $F, G \in \mathscr{F}$ esiste $E \in \mathscr{F}$ tale che $E \subseteq F \cap G$.*

Allora esiste una subnet x_{f_i} di x_a definita su I tale che per ogni $F \in \mathscr{F}$ esiste $j \in I$ tale che $x_{f_i} \in F$ per ogni $i \succeq j$.

Dimostrazione. La famiglia $\mathscr{F}$ è un'insieme diretto con il preordine $F \preceq G \Leftrightarrow G \subseteq F$. Poniamo quindi

$$I = \{(F, a) \in \mathscr{F} \times A \mid x_a \in F\}$$

e lo dotiamo del preordine prodotto, verifichiamo che è un insieme diretto. Prendiamo allora $(F, a), (G, b) \in I$ per ipotesi esiste $E \in \mathscr{T}$ tale che $E \subseteq F \cap G$ e dunque $F, G \preceq E$. Ora essendo A diretto esiste $c' \in A$ tale che $a, b \preceq c'$, per la prima ipotesi esisterà $c \succeq c'$ tale che $x_c \in E$. Dunque $(F, a), (G, b) \preceq (E, c)$ ed $(E, c) \in I$ e perciò è un insieme diretto.

Posto adesso $f : I \to A$ in modo tale che $f(F, a) = a$ avremo chiaramente che f è crescente, inoltre per ogni $a \in A$ e per ogni $F \in \mathscr{F}$ esiste $b \succeq a$ tale che $x_b \in F$ e dunque $a \preceq f(F, b)$. Quindi $x_{f(F,a)}$ è una subnet di x_a, rimane da verificare l'ultima parte del lemma.

Sia $F \in \mathscr{F}$ generico, per quanto detto prima esiste $a \in A$ tale che $(F, a) \in I$, di conseguenza per ogni $(G, b) \succeq (F, a)$ si ha

$$x_{f(G,b)} = x_b \in G \subseteq F$$

come volevasi dimostrare. ■

Consideriamo adesso una funzione $f : Z \to Y$ definita su un sottoinsieme non vuoto Z di uno spazio topologico X, prendiamo inoltre un qualunque punto di accumulazione $x \in X$ di Z ed un generico $y \in Y$. È possibile dimostrare che $f(z) \to y$ al tendere di $z \in Z$ ad X se e solo se per ogni net $z_a \in Z$ definita su qualche insieme diretto A convergente ad x per cui esiste $b \in A$ tale che $z_a \neq x$ per ogni $a \geq b$ si abbia $f\left(z_a\right) \to y$. In particolare $f\left(z_n\right) \to y$ per ogni successione $z_n \in Z$ convergente ad x e che non assuma il valore x da un certo punto in poi.

Si noti che anche se $f\left(z_n\right) \to y$ per ogni successione z_n convergente ad x non è detto che il limite di f in x sia ancora y, anzi potrebbe non esistere. Per spazi topologici generici non è sufficiente considerare solo le successioni ma abbiamo bisogno necessariamente delle net.

Definizione 2.2.9. Dati due spazi topologici X, Y e una funzione $f : X \to Y$ diremo che f è **continua** in un punto $x_0 \in X$ se e solo se o x_0 è un punto isolato oppure x_0 è un punto di accumulazione per X ed $f(x) \to f(x_0)$ al tendere di x a x_0.

Diremo invece che f è continua su X se e solo se è continua su tutti i punti di X, in altre parole che vale una delle seguenti affermazioni:

- Tutti i punti di X sono punti isolati;
- Per ogni punto di accumulazione $x_0 \in X$ si ha $f(x) \to f\left(x_0\right)$ al tendere di x ad x_0.

Quando calcoliamo il limite di una funzione continua possiamo in realtà rilassare le condizioni imposte dalla definizione di limite che altrimenti saremo costretti a utilizzare per funzioni generiche. In altre parole

Proposizione 2.2.10. *Una funzione $f : X \to Y$ è continua in $x \in X$ se e solo se per ogni aperto $U \subseteq Y$ contenente $f(x)$ esiste un aperto $V \subseteq X$ contenente x tale che $V \subseteq f^{-1}(U)$. Inoltre f è continua su tutto X se e solo se la controimmagine degli aperti di Y è un aperto di X.*

Dimostrazione. Supponiamo che f sia continua in $x \in X$ e fissiamo un aperto U di Y contenente $f(x)$, allora x può essere un punto isolato o di accumulazione per X. Se x è isolato allora $\{x\}$ è un aperto di X con $\{x\} \subseteq f^{-1}(U)$ automaticamente. Se invece x è un punto di accumulazione esisterà un aperto $V \subseteq X$ contenente x tale che $V \setminus \{x\} \subseteq f^{-1}(U)$. Ora dato che $f(x) \in U$ si ha $x \in f^{-1}(U)$ e perciò $V \subseteq f^{-1}(U)$ e l'implicazione desiderata è dimostrata. L'implicazione opposta è banale.

Supponiamo ora f continua su tutto X e prendiamo un qualunque aperto U di Y, se $f^{-1}(U) = \emptyset$ l'asserto è automaticamente verificato in quanto l'insieme vuoto è sempre aperto, quindi supponiamo che esista un $x \in f^{-1}(U)$. Dato che f è aperta in x e U è un aperto contenente $f(x)$ per quanto detto prima esisterà un altro aperto $V_x \subseteq X$ contenente x tale che $V_x \subseteq f^{-1}(U)$. Ma allora

$$f^{-1}(U) = \bigcup_{x \in f^{-1}(U)} V_x$$

e dato che l'unione arbitraria di aperti è ancora aperta abbiamo che $f^{-1}(U)$ è aperto in X. L'implicazione opposta segue dal punto precedente semplicemente ponendo $V = f^{-1}(U)$. ■

Per indicare lo spazio di tutte le funzioni continue da X in Y utilizzeremo il simbolo $C(X, Y)$.

Da questo risultato possiamo indebolire le condizioni del teorema di cambio di variabile nei limiti quando f è continua in x. In tal caso possiamo richiedere sulla funzione p solamente che $p(\bar{e}) \to x$ al tendere di $\bar{e}$ ad e. Abbiamo così

Proposizione 2.2.11. *Data una funzione $f : X \to Y$ definita tra spazi topologici avremo che f è continua in $x \in X$ se e solo se per ogni net $x_a \in X$ convergente ad x si ha $f(x_a) \to f(x)$.*

Torniamo ora alle funzioni continue. Una funzione biettiva f tra due spazi topologici è un **omeomorfismo** se e solo se sia f che f^{-1} sono continue, mentre due spazi topologici sono **omeomorfi** se e solo se esiste un omeomorfismo tra di essi. Una proprietà $\mathscr{P}$ che può valere o no sugli spazi topologici è una **proprietà topologica** se e solo se preso un qualunque spazio topologico X che soddisfa la proprietà $\mathscr{P}$ possiamo dedurre che ogni altro spazio topologico omeomorfo a X deve soddisfare la proprietà $\mathscr{P}$.

Proposizione 2.2.12. *Data una funzione tra spazi topologici $f : X \to Y$ allora le seguenti affermazioni sono equivalenti tra loro:*

- *f è continua (su tutto X);*
- *per ogni aperto A di Y l'insieme $f^{-1}(A)$ è aperto in X;*
- *per ogni chiuso C di Y l'insieme $f^{-1}(C)$ è chiuso in X.*

Osservazione. Dato un insieme X allora la topologia τ su X è più fine della topologia η sempre su X se e solo se l'applicazione

$$\mathrm{id}_X : x \in (X, \tau) \to x \in (X, \eta)$$

è continua.

2.3 Basi e sistemi fondamentali di intorni

Definizione 2.3.1. Fissato $x \in X$, diremo che una famiglia di intorni $\mathscr{I}$ di un punto $x \in X$ è un **sistema fondamentale di intorni** di x se e solo se per ogni aperto A di X contenente x esiste $I \in \mathscr{I}$ tale che $I \subseteq A$.

Diremo invece che $\mathscr{I}$ è un **sistema essenziale di intorni** di x se e solo se per ogni aperto $A \subseteq X$ contenente x esistono $I_1, I_2, \ldots, I_n \in \mathscr{I}$ tali che $I_1 \cap I_2 \cap \cdots \cap I_n \subseteq A$.

Chiaramente ogni sistema fondamentale di intorni è anche un sistema essenziale di intorni, viceversa de $\mathscr{I}$ è un sistema essenziale di intorni per $x \in X$ allora la famiglia

$$\mathscr{I}' = \left\{ I_1 \cap I_2 \cap \cdots \cap I_n \,\middle|\, n \in \mathbb{N}, I_l \in \mathscr{I} \right\} \tag{2.3.1}$$

è un sistema fondamentale di intorni sempre per x.

Prendendo l'interno degli elementi di un sistema fondamentale di intorni è possibile generare un altro sistema fondamentale di intorni dello stesso punto formato però esclusivamente da insiemi aperti. Non è però sempre vero che un punto possiede sempre un sistema fondamentale di intorni chiusi, per averlo dobbiamo infatti richiedere almeno la regolarità dello spazio.

Proposizione 2.3.2. *Se X è uno spazio regolare e $\mathscr{I}$ è un sistema fondamentale di intorni per un qualche $x \in X$ allora per ogni $I \in \mathscr{I}$ esiste $J \in \mathscr{I}$ tale che*

$$x \in \mathrm{Cl}(J) \subseteq \mathrm{Int}(I)$$

e quindi ogni punto possiede sempre un sistema fondamentale di intorni chiusi.

Dimostrazione. L'insieme $X \setminus \mathrm{Int}(I)$ è chiuso e non contiene x per la definizione di intorno, quindi per regolarità esistono due aperti disgiunti A e B con $x \in A$ e $X \setminus \mathrm{Int}(I) \subseteq B$.

Poiché $\mathscr{I}$ è un sistema fondamentale di intorni per x esisterà $J \in \mathscr{I}$ tale che $J \subseteq A \subseteq (X \setminus B)$ e quindi per definizione di chiusura avremo che

$$\mathrm{Cl}(J) \subseteq (X \setminus B) \subseteq \mathrm{Int}(I)$$

■

Grazie ai sistemi fondamentali di intorni è possibile testare più facilmente i limiti di funzioni. Vale infatti il seguente risultato di semplice verifica:

Proposizione 2.3.3. *Dati due spazi topologici X, Y, un sottoinsieme non vuoto $Z \subseteq X$ ed una funzione $f : Z \to Y$. Presi $x \in X$ punto di accumulazione per Z ed $y \in Y$ e dati un sistema fondamentale di intorni $\mathscr{I}$ di x ed un sistema essenziale di intorni $\mathscr{J}$ di y avremo che*

$$f(z) \to y \Leftrightarrow \forall U \in \mathscr{J}\ \exists V \in \mathscr{I} \text{ tale che } f(V \cap Z \setminus \{x\}) \subseteq U$$

Nella proposizione precedente non possiamo sostituire $\mathscr{I}$ con un sistema essenziale di intorni (ovviamente), però nulla vieta di trasformare un sistema essenziale di intorni di x in uno fondamentale tramite la (2.3.1) per poi applicargli la proposizione precedente.

Corollario 2.3.4. *Sia X spazio topologico, $x \in X$ ed $\mathscr{I}$ sistema essenziale di intorni per x. Data una net x_a in X definita su A allora $x_a \to x$ se e solo se per ogni $I \in \mathscr{I}$ esiste $a \in A$ tale che $x_b \in I$ per ogni $b \geq a$.*

Dimostrazione. Chiaramente $x_a \to x$ se e solo se per ogni $I_1, I_2, \dots, I_n \in \mathscr{I}$ esiste $\tilde{b} \in A$ tale che $x_b \in I_1 \cap I_2 \cap \cdots I_n$ per ogni $b \geq \tilde{b}$. Per ipotesi sappiamo che esistono $b_i \in A$ tali che $x_a \in I_j$ per ogni $a \geq b_j$ e dunque, essendo A un insieme diretto, esisterà sicuramente un $\tilde{b} \in A$ tale che $\tilde{b} \geq b_i$ per ogni $1 \leq i \leq n$ che soddisfa la tesi. ■

I sistemi essenziali di intorni tra l'altro permettono di stabilire anche delle corrispondenze tra topologie, infatti abbiamo il seguente risultato:

Proposizione 2.3.5. *Dato un insieme non vuoto X e due topologie τ_1, τ_2 su di esso. Se per ogni $x \in X$ esiste un sistema essenziale di intorni $\{U_i\}_{i \in I}$ di x rispetto a τ_1 formato interamente da intorni di τ_2 allora $\tau_1 \subseteq \tau_2$.*

Dimostrazione. Prendiamo un aperto $U \in \tau_1$ ed un generico $x \in U$, esisteranno di conseguenza $U_1, U_2, \dots, U_n$ intorni di x rispetto a τ_1 per cui $U_1 \cap U_2 \cap \cdots \cap U_n \subseteq U$. Per ipotesi allora esistono degli aperti $V_i \in \tau_2$ tale che $x \in V_i \subseteq U_i$ per ogni i.

Ma dato che $V_1 \cap V_2 \cap \cdots \cap V_n$ è comunque un aperto di (X, τ_2) contenuto in U avremo che l'applicazione identica

$$\mathrm{id}_X : x \in (X, \tau_2) \to x \in (X, \tau_1)$$

è continua e perciò $\tau_1 \subseteq \tau_2$. ■

Definizione 2.3.6. Preso un qualunque spazio topologico X una famiglia di aperti $\mathscr{B}$ è una **base** per la topologia su X se e solo se ogni aperto di X è unione di elementi di $\mathscr{B}$.

Diremo invece che la famiglia di aperti $\mathscr{B}$ è una **semibase** per X se e solo se la famiglia

$$\mathscr{B}' = \left\{A_1 \cap A_2 \cap \cdots \cap A_n \;\middle|\; n \in \mathbb{N}, A_i \in \mathscr{B}\right\}$$

è una base di X.

La comodità nel lavorare con sistemi fondamentali di intorni e con le basi risiede principalmente nella possibilità che una data proprietà verificata solamente sui loro elementi risulti verificata per ogni intorno o aperto della topologia rispettivamente.

Per esempio per stabilire la convergenza di una successione in x ci si può limitare a considerare solamente gli elementi di un generico sistema fondamentale di intorni di x, mentre per dimostrare la continuità di una funzione $f : X \to Y$ su tutto X basterebbe considerare solamente le controimmagini di una qualunque base di Y.

Proposizione 2.3.7. *Preso A insieme non vuoto e $\mathscr{A}$ una collezione non vuota di sottoinsiemi di A che soddisfi le seguenti proprietà:*

- $\bigcup_{B\in\mathscr{A}} B = A$;
- *Per ogni $B_1, B_2 \in \mathscr{A}$ esiste $\mathscr{B} \subseteq \mathscr{A}$ non vuoto tale che $\bigcup_{B\in\mathscr{B}} B = B_1 \cap B_2$.*

Allora la famiglia

$$\tau = \left\{ \bigcup_{B\in\mathscr{B}} B \;\middle|\; \mathscr{B} \subseteq \mathscr{A} \right\}$$

è una topologia su A e $\mathscr{A}$ ne è una base.

Un risultato analogo vale anche per le semibasi

Proposizione 2.3.8. *Preso A insieme non vuoto e $\mathscr{A} \subseteq \mathcal{P}(A)$ non vuoto tale che $\bigcup_{B\in\mathscr{A}} B = A$. Allora la famiglia*

$$\mathscr{A}' = \{A_1 \cap A_2 \cap \cdots \cap A_n \mid n \in \mathbb{N}, A_i \in \mathscr{A}\}$$

è una base della seguente topologia su A:

$$\tau = \left\{ \bigcup_{B\in\mathscr{B}'} B \;\middle|\; \mathscr{B}' \subseteq \mathscr{A}' \right\}$$

In particolare $\mathscr{A}$ è una semibase di (X, τ).

Esempio 2.3.9 (Topologia di $\overline{\mathbb{R}}$)**.** Consideriamo in $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ i seguenti intervalli

$$\begin{aligned} (a,b) &= \left\{c \in \overline{\mathbb{R}} \mid a < c < b\right\} \\ [-\infty, b) &= \left\{c \in \overline{\mathbb{R}} \mid c < b\right\} \\ (a, +\infty] &= \left\{c \in \overline{\mathbb{R}} \mid a < c\right\} \end{aligned}$$

al variare di $a, b \in \overline{\mathbb{R}}$. Il lettore può verificare facilmente che la famiglia $\mathscr{A}$ formata da tutti e soli gli intervalli in tale forma soddisfa le ipotesi della proposizione 2.3.7 e quindi induce una topologia su $\overline{\mathbb{R}}$ che ha $\mathscr{A}$ come base.

Definizione 2.3.10. Sia X spazio topologico, allora un suo sottoinsieme $Y \subseteq X$ è **denso** in X se e solo se la chiusura di Y coincide con X, mentre è **raro** se e solo se la sua chiusura non ha punti interni.

Diremo inoltre che Y è di **prima categoria** in X se e solo se è unione numerabile di sottoinsiemi rari di X, altrimenti è detto di **seconda categoria**.

Diciamo invece che uno spazio topologico X è **separabile** se e solo se possiede un sottoinsieme denso e numerabile.

Proposizione 2.3.11. *Un sottoinsieme Y di X è denso se e solo se interseca ogni aperto non vuoto di X.*

Dimostrazione. Dalla definizione di chiusura l'insieme $A = X \setminus \mathrm{Cl}(Y)$ è l'unione di tutti gli aperti che hanno intersezione nulla con Y. Da ciò segue immediatamente il teorema. ∎

Proposizione 2.3.12. *Se Y è un sottoinsieme raro di uno spazio topologico X allora $X \setminus Y$ è denso in X.*

Il viceversa non è sempre vero: in $\mathbb{R}$ sia $\mathbb{Q}$ che $\mathbb{R} \setminus \mathbb{Q}$ sono densi in $\mathbb{R}$ e quindi non sono rari. Se però $Y \subseteq X$ è aperto e denso allora $X \setminus Y$ è chiuso e raro.

Esempio 2.3.13. Dotiamo $\mathbb{R}$ della topologia usuale, allora $\mathbb{N}$ è un sottoinsieme raro mentre $\mathbb{Q}$ è denso e numerabile, quindi $\mathbb{R}$ è uno spazio topologico separabile.

Con un procedimento analogo si dimostra facilmente che anche $\mathbb{R}^n$ è separabile per ogni $n \in \mathbb{N}$

Proposizione 2.3.14. *Se X è uno spazio topologico di seconda categoria e U_i è un sottoinsieme aperto denso per ogni $i \in \mathbb{N}$ allora $\bigcap_{i=1}^{+\infty} U_i$ è denso in X.*

Proposizione 2.3.15. *Se X è uno spazio topologico di seconda categoria allora per ogni famiglia numerabile $\{C_i\}_{i\in\mathbb{N}}$ di sottoinsiemi chiusi con $X = \bigcup_{i=1}^{+\infty} C_i$ almeno uno dei C_i possiede un punto interno.*

2.4 Compattezza

Definizione 2.4.1. Uno spazio topologico X si dice **compatto** se e solo se per ogni ricoprimento aperto $\{V_i\}_{i\in I}$ di X esiste un sottoinsieme I' di I finito tale che

$$X = \bigcup_{i\in I'} V_i$$

Dalla definizione di chiusura seguono immediatamente i seguenti risultati:

Proposizione 2.4.2. *Uno spazio topologico X è compatto se e solo se per ogni famiglia $\{C_i\}_{i\in I}$ di sottoinsiemi chiusi di X tale che $\bigcap_{i\in I'} C_i \neq \emptyset$ per ogni $I' \in \mathcal{F}(I)$ si ha $\bigcap_{i\in I} C_i \neq \emptyset$.*

Dato uno spazio topologico X diremo che un suo sottoinsieme Y è compatto se e solo se Y con la topologia indotta da X è compatto. Osserviamo quindi che, a differenza della proprietà di essere chiusi o aperti, la proprietà di compattezza di un sottoinsieme si mantiene anche se restringiamo lo spazio topologico ambiente. In altre parole dati uno spazio topologico X e due sottoinsiemi $Z \subseteq Y \subseteq X$ se dotiamo Y della topologia indotta allora Z è compatto in Y se e solo se è compatto in X.

Proposizione 2.4.3. *Se X è uno spazio topologico compatto allora ogni suo sottoinsieme chiuso è compatto.*

Proposizione 2.4.4. *Se X è uno spazio compatto ed $f : X \to Y$ è continua allora $f(X)$ è un sottospazio compatto di Y con la topologia indotta.*

Proposizione 2.4.5. *Se X è uno spazio topologico generico allora la chiusura di ogni sottoinsieme compatto è compatto (sempre con la topologia indotta). Se X è anche uno spazio di Hausdorff allora tutti i suoi sottoinsiemi compatti sono chiusi.*

Teorema 2.4.6. *Uno spazio topologico (X,τ) è compatto se ogni net x_a su X possiede una subnet convergente.*

Dimostrazione. Supponiamo che X sia compatto e prendiamo una qualunque net x_a definita su A. Per ogni $a \in A$ poniamo $C_a = \mathrm{Cl}(D_a)$ dove $D_a = \{x_b \mid b \geq a\}$. Dalla definizione di insieme diretto avremo che le intersezioni di un numero finito di C_a non è mai vuota, la proposizione 2.4.2 ci dirà dunque che esiste $x \in C_a$ per ogni $a \in A$. Definiamo ora la famiglia

$$\mathscr{F} = \{U[x] \mid U \in \mathscr{T}\}$$

poiché $(U \cap V)[x] = U[x] \cap V[x]$ abbiamo che l'intersezione finita degli elementi di $\mathscr{F}$ appartiene ancora ad F. Sia ora $a \in A$ generico, poiché x appartiene alla chiusura di D_a avremo che $U[x] \cap D_a \neq \emptyset$ ovvero esiste $b \geq a$ tale che $x_b \in U[x]$.

Possiamo quindi applicare il lemma 2.2.8 e trovare una subnet x_{a_i} su I di x_a tale che per ogni $U \in \mathscr{T}$ esiste $j \in I$ tale che $x_{a_i} \in U[x]$ per ogni $i \geq j$, ovvero $x_{a_i} \to x$.

Viceversa supponiamo che ogni net in (X,τ) possiede una subnet convergente e per assurdo esiste un ricoprimento aperto $\mathscr{C}$ di X che non possiede alcun sottoricoprimento finito. Posto allora

$$A = \mathcal{F}(\mathscr{C})$$

per ogni $a = \{U_1, U_2, \dots, U_n\} \in A$ esisterà qualche $x_a \in X \setminus \bigcup_{i=1}^n U_i$. Inoltre A è un insieme infinito diretto rispetto all'inclusione dunque x_a è una net su X. Per ipotesi esiste un'estratta x_{a_k} definita su qualche insieme diretto K convergente a qualche $x \in X$.

Ora esisterà necessariamente $U \in \mathscr{C}$ tale che $x \in U$ in quanto $\mathscr{C}$ è un ricoprimento di X, dunque esisterà $h \in K$ tale che $x_{a_k} \in U$ per ogni $k \geq h$. Per definizione di insieme diretto e di estratta di una net esisterà $k \geq h$ tale che $a_k \geq \{U\}$ e dunque $x_{a_k} \in U$. Ma dato che U appartiene ad a_k dovrebbe essere $x_{a_k} \notin U$ assurdo. Quindi $\mathscr{C}$ possiede un sottoricoprimento finito ed X è compatto. ■

Osservazione. Questo risultato non vale più se al posto delle subnet consideriamo le estratte.

Da questi risultati è possibile determinare un utilissimo criterio di omeomorfismo per funzioni continue biettive.

Lemma 2.4.7 (Criterio di omeomorfismo)**.** *Siano X spazio compatto, Y spazio di Hausdorff e $f : X \to Y$ applicazione continua biettiva. Allora anche l'inversa f^{-1} è continua e quindi gli spazi X ed Y sono omeomorfi.*

Dimostrazione. Basta dimostrare che per ogni chiuso $C \subseteq X$ anche $f(C)$ è chiuso. Per la compattezza di X avremo che C è compatto e quindi $f(C)$ è un sottospazio compatto di Y. Ma Y è uno spazio di Hausdorff quindi $f(C)$ è un chiuso di Y. ■

Proposizione 2.4.8. *Se X è uno spazio regolare allora per ogni chiuso $C \subseteq X$ e per ogni compatto $K \subseteq X$ tali che $C \cap K = \emptyset$ esistono due aperti A e B contenenti rispettivamente K e C tali che $A \cap B = \emptyset$.*

Dimostrazione. Dalla definizione di regolarità per ogni $x \in K$ esistono due aperti A_x e B_x tali che $x \in A_x$, $C \subseteq B_x$ e $A_x \cap B_x = \emptyset$. Dato che gli A_x ricoprono K allora possiamo scegliere un numero finito di punti $x_1, \dots, x_n$ tali che A_{x_i} ricopre ancora K.

Posto allora

$$A = \bigcup_{i=1}^{n} A_{x_i} \qquad\qquad B = \bigcap_{i=1}^{n} B_{x_i}$$

chiaramente A, B sono aperti che contengono rispettivamente K e C e inoltre

$$A \cap B \subseteq \bigcup_{i=1}^{n} (A_{x_i} \cap B_{x_i}) = \emptyset$$

■

2.5 Assiomi di numerabilità

Nei capitoli precedenti abbiamo ottenuto dei risultati di equivalenza tra definizioni date tramite aperti e tramite net. Dato che le successioni sono un caso particolare di net risulta ragionevole introdurre nuove definizioni dei concetti di continuità e compattezza tramite successioni.

Definizione 2.5.1. Una funzione $f : X \to Y$ definita tra due spazi topologici è **sequenzialmente continua** in $x \in X$ se e solo se per ogni successione $x_n \in X$ convergente ad x la successione $f(x_n)$ converge a $f(x)$.

Chiaramente tutte le funzioni continue sono sequenzialmente continue mentre esistono funzioni sequenzialmente continue ma non continue.

Definizione 2.5.2. Uno spazio topologico X si dice **sequenzialmente compatto** se e solo se ogni successione in X possiede un'estratta convergente.

Nonostante quanto detto prima per la continuità, compattezza e sequenziale compattezza sono due concetti del tutto indipendenti, ed esistono spazi topologici compatti ma non sequenzialmente compatti e spazi sequenzialmente compatti ma non compatti.

Il problema principale in questi due casi è che ci sono così tanti aperti nella topologia che i numeri naturali non sono abbastanza da poterli enumerare, e quindi le successioni non bastano ad enumerarli tutti. Per questa ragione introduciamo la seguente definizione:

Definizione 2.5.3. Uno spazio topologico X **soddisfa l'assioma** $\mathscr{N}_1$ se e solo se ogni suo punto possiede un sistema fondamentale di intorni al più numerabile.

In tali spazi quindi possiamo sempre trovare sistemi fondamentali di intorni al più numerabili superando in tal modo l'ostacolo della cardinalità. Innanzitutto osserviamo che se $\{U_i\}_{i\in\mathbb{N}}$ è un sistema fondamentale di intorni per x allora posto

$$V_1 = U_1$$
$$V_{i+1} = U_{i+1} \cap V_i$$

avremo che $\{V_i\}_{i\in\mathbb{N}}$ è ancora un sistema fondamentale di intorni per x e inoltre $V_{i+1} \subseteq V_i$ per ogni i.

Gli spazi $\mathscr{N}_1$ soddisfano molte proprietà che riguardano le successioni che precedentemente erano date dalle net, questo per via del seguente risultato:

Proposizione 2.5.4. *Sia X uno spazio topologico $\mathcal{N}_1$, presi $x \in X$ ed una qualunque net $x_a \in X$ su A che possieda una subnet convergente ad x esiste sempre un'applicazione $a : \mathbb{N} \to A$ crescente per cui $x_{a_n} \to x$.*

Se $A = \mathbb{N}$ possiamo supporre che a_i sia strettamente crescente e quindi x_{a_i} è un'estratta di x_a.

Dimostrazione. Prendiamo un sistema fondamentale di intorni $\{V_n\}_{n\in\mathbb{N}}$ per x come sopra ed una subnet x_{k_r} definita su R convergente ad x. Chiaramente esisterà $a_1 = k_{r_1} \in A$ tale che $x_{a_1} \in V_1$. Supponiamo di aver definito a_i con $i \geq 1$, definiamo allora $a_{i+1} \in A$ in modo tale che $a_i \leq a_{i+1}$ ed $x_{a_{i+1}} \in V_{i+1}$. Dato che la successione di intorni V_i è decrescente segue immediatamente che $x_{a_i} \to x$.

Quando $A = \mathbb{N}$ il preordine è una relazione d'ordine e dunque possiamo imporre $a_i < a_{i+1}$ nella definizione di prima. ■

Osservazione. La funzione trovata sopra non è una subnet di x_a in quanto l'immagine di a_i non è necessariamente cofinale.

Corollario 2.5.5. *Se X è uno spazio $\mathcal{N}_1$ e $Y \subseteq X$ allora*

$$x \in \mathrm{Cl}(I) \Leftrightarrow \exists x_n \in I \text{ tale che } x_n \to x$$

Proposizione 2.5.6. *Se $f : X \to Y$ è sequenzialmente continua in x ed X è $\mathcal{N}_1$ allora f è anche continua in x.*

Dimostrazione. Per assurdo f è sequenzialmente continua ma non continua, quindi esiste una net $x_a \to x$ per cui $f(x_a) \nrightarrow f(x)$. Ma allora esistono un aperto U contenente $f(x)$ ed un'estratta x_b su $B \subseteq A$ tale che $f(x_b) \notin U$. Poiché $x_b \to x$ per la proposizione 2.5.4 esiste $k : \mathbb{N} \to B$ tale che $x_{k_n} \to x$ e per ipotesi dunque $f\left(x_{k_n}\right) \to f(x)$ assurdo in quanto $k_n \in B$. ■

Proposizione 2.5.7 (Equivalenza)**.** *Dato un insieme non vuoto X con τ_1, τ_2 due topologie su X che soddisfano entrambe l'assioma $\mathcal{N}_1$. Allora $\tau_1 = \tau_2$ se e solo se per ogni successione x_n in X e per ogni $x \in X$ avremo che $x_n \to x$ in τ_1 se e solo se $x_n \to x$ in τ_2.*

Dimostrazione. Basta dimostrare che $\tau_1 \subseteq \tau_2$ se e solo se ogni successione convergente in τ_2 converge anche in τ_1 (in quanto quest'ultimo conterrà meno aperti di τ_2). Dato che l'implicazione diretta è banale dimostreremo solo l'implicazione inversa.

Sia $U \in \tau_1$ aperto non vuoto con $x \in U$ generico, preso un sistema di intorni numerabile $\{V_i\}_{i\in\mathbb{N}}$ rispetto alla topologia τ_2 con $V_{i+1} \subseteq V_i$ facciamo vedere che esiste $j \in \mathbb{N}$ tale che $V_j \subseteq U$. Se questo non fosse vero esisterebbero degli $x_i \in V_i \setminus U$ per ogni $i \in \mathbb{N}$ e dato che V_i è un fiftema fondamentale di intorni per x questo equivale ad affermare che $x_i \to x$ in τ_2.

Ma allora $x_i \to x$ anche in τ_1 e perciò esisterà $n \in \mathbb{N}$ tale che $x_n \in U$ assurdo, quindi $V_j \subseteq U$ per qualche $j \in \mathbb{N}$. Per l'arbitrarietà di x l'insieme U è unione di aperti in τ_2 e dunque $U \in \tau_2$. ■

Dalla proposizione 2.5.4 segue immediatamente che

Proposizione 2.5.8. *Se X è uno spazio compatto ed $\mathcal{N}_1$ allora è anche sequenzialmente compatto.*

Oltre all'assioma $\mathcal{N}_1$ vi è un altro assioma di separabilità molto più forte del precedente

Definizione 2.5.9. Uno spazio topologico X **soddisfa l'assioma** $\mathscr{N}_2$ se e solo se una base al più numerabile.

Uno spazio $\mathscr{N}_2$ è sia $\mathscr{N}_1$ che separabile, mentre il viceversa non è verificato per spazi topologici generici. Se infatti consideriamo la retta $\mathbb{R}$ dotata della topologia generata dagli intorni nella forma

$$[a, b) \tag{2.5.1}$$

con $-\infty < a < b < +\infty$, allora si può verificare che è uno spazio $\mathscr{N}_1$ con $\mathbb{Q}$ come sottoinsieme denso ma non è $\mathscr{N}_2$. Se però ci limitiamo a considerare solamente gli spazi metrici allora è verificata anche l'implicazione opposta come vedremo più avanti.

Un risultato degno di nota per gli spazi topologici $\mathscr{N}_2$ è il seguente lemma dovuto a Lindelöf:

Lemma 2.5.10 (Lindelöf)**.** *Se X è spazio topologico $\mathscr{N}_2$ allora per ogni famiglia di aperti Σ di S esiste un sottoinsieme $\Sigma' \subseteq \Sigma$ numerabile tale che*

$$\bigcup_{A \in \Sigma} A = \bigcup_{A \in \Sigma'} A$$

Dimostrazione. Se indichiamo con $\mathscr{B}$ una base numerabile di X allora anche l'insieme

$$\left\{ B \in \mathscr{B} \;\middle|\; B \subseteq A \text{ per qualche } A \in \Sigma \right\}$$

è al più numerabile, e quindi possiamo enumerare i suoi elementi, che perciò indicheremo come $B_1, B_2, B_3, \dots$. Per ogni $n \in \mathbb{N}$ esisterà $A_n \in \Sigma$ tale che $B_n \subseteq A_n$, dimostriamo ora che la famiglia

$$\Sigma' = \left\{ A_n \;\middle|\; n \in \mathbb{N} \right\}$$

soddisfa la tesi del lemma. Le inclusioni $\Sigma' \subseteq \Sigma$ e $\bigcup_{A \in \Sigma'} A \subseteq \bigcup_{A \in \Sigma} A$ sono banali, viceversa per ogni $x \in \bigcup_{A \in \Sigma} A$ esiste $A \in \Sigma$ per cui $x \in A$. Per definizione di base topologica esisterà $n \in \mathbb{N}$ tale che $x \in B_n \subseteq A_n$ dunque $x \in \bigcup_{A \in \Sigma'} A$ e il lemma è così dimostrato. ■

Il lemma di Lindelöf in particolare ci garantisce che da un ricoprimento aperto possiamo sempre estrarre un sottoricoprimento al più numerabile.

Corollario 2.5.11. *Gli spazi topologici $\mathscr{N}_2$ sono sequenzialmente compatti se e solo se sono compatti.*

Dimostrazione. Se Σ è una famiglia di aperti che ricopre X ed X è uno spazio $\mathscr{N}_2$ sequenzialmente compatto, per il lemma di Lindelöf possiamo supporre Σ numerabile e quindi possiamo enumerare i suoi elementi $A_1, A_2, A_3, \dots$.

Se X non fosse compatto allora nessun sottoinsieme finito di Σ ricoprirà X, in particolare per ogni $n \in \mathbb{N}$ si avrà

$$X \neq (A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n)$$

e quindi esisterà $x_n \in X$ che non appartiene ad alcun A_i per $1 \leq i \leq n$.

Per la sequenziale compattezza di X esisterà un'estratta x_{n_k} convergente ad un certo $x \in X$. Ora poiché gli elementi di Σ ricoprono X esisterà $A_h \in \Sigma$ tale che $x \in A_h$ e dunque esisterà $N \in \mathbb{N}$ tale che $x_{n_k} \in A_h$ per ogni $k > N$.

Ora se prendiamo $k > N \vee h$ poiché $n_k \geq k > h$ per ipotesi $x_{n_k} \notin A_h$ assurdo, quindi X deve essere compatto. ■

2.6 Semicontinuità inferiore

Consideriamo un qualunque punto l di $\overline{\mathbb{R}}$, diremo che un intorno $U \subseteq \overline{\mathbb{R}}$ di l è un **intorno superiore** di l se e solo se per ogni $l' \in U$ e per ogni $l'' \leq l'$ si ha $l'' \in U$. In particolare dato un intorno superiore U di $l \in \overline{\mathbb{R}}$ avremo che

- se $l \in \mathbb{R}$ allora $[-\infty, l+\varepsilon) \subseteq U$ per qualche $\varepsilon > 0$;
- se $l = -\infty$ allora $[-\infty, R) \subseteq U$ per qualche $R \in \mathbb{R}$;
- se $l = +\infty$ allora $U = \overline{\mathbb{R}}$.

Supponendo invece che $l'' \geq l'$ si definiscono gli **intorni inferiori** di l.

Dato un sottoinsieme non vuoto A di $\overline{\mathbb{R}}$ definiamo l'**estremo inferiore** di A quell'elemento l di $\overline{\mathbb{R}}$, che indicheremo con $\inf A$, tale che:

1. per ogni intorno inferiore I di l si ha $A \subseteq I$;
2. per ogni intorno superiore S di l si ha $A \cap S \neq \emptyset$.

In particolare se $l \in \mathbb{R}$ ciò è equivalente ad affermare che

1. per ogni $a \in A$ si ha $a \geq l$;
2. per ogni $\varepsilon > 0$ esiste $a' \in A$ per cui $a < l + \varepsilon$.

Scambiando i ruoli degli intorni inferiori e superiori otteniamo la definizione di **estremo superiore**. Per convenzione poniamo anche

$$\inf \emptyset = +\infty \qquad\qquad \sup \emptyset = -\infty$$

e inoltre data una funzione $f : X \to \overline{\mathbb{R}}$ poniamo per comodità di notazione

$$\inf_{x \in X} f(x) = \inf f(X)$$

e in maniera analoga facciamo con l'estremo superiore.

Definizione 2.6.1. Dati uno spazio topologico X, una funzione $f : X \to \overline{\mathbb{R}}$ e un punto x di accumulazione per X. Diremo che $l \in \overline{\mathbb{R}}$ è il **limite inferiore** di f in x, e scriviamo $l = \liminf_{z \to x} f(z)$, se e solo se l è il massimo di di tutti i numeri $t \in \overline{\mathbb{R}}$ che soddisfano la seguente proprietà: per ogni intorno inferiore $U \subseteq \overline{\mathbb{R}}$ di t esiste un aperto $V \subseteq X$ contenente x tale che $V \setminus \{x\} \subseteq f^{-1}(U)$. In maniera analoga si definisce il **limite superiore** di f in x.

A differenza del limite di una funzione, il limite inferiore e il limite superiore esistono sempre in quanto vale il seguente risultato:

Teorema 2.6.2. *Se x è un punto di accumulazione allora per ogni sistema fondamentale di intorni $\mathscr{I}$ di x*

$$\liminf_{z \to x} f(z) = \sup_{U \in \mathscr{I}} \inf_{z \in U \setminus \{x\}} f(z)$$
$$\limsup_{z \to x} f(z) = \inf_{U \in \mathscr{I}} \sup_{z \in U \setminus \{x\}} f(z)$$

e in particolare si ha $\liminf_{z \to x} f(z) \leq \limsup_{z \to x} f(z)$. Inoltre si ottiene l'uguaglianza l'uguaglianza $\liminf_{z \to x} f(z) = \limsup_{z \to x} f(z) = L$ se e solo se il limite $\lim_{z \to x} f(z)$ esiste in $\overline{\mathbb{R}}$, e in tal caso il limite sarà proprio L.

Dimostrazione. Dimostriamo solo l'uguaglianza per il limite inferiore in quanto l'altra si ottiene semplicemente sostituendo f con $-f$. Per ogni intorno U di x poniamo

$$L(U) = \inf_{z\in U\setminus\{x\}} f(z)$$
$$L = \sup_{U\in\mathscr{I}} L(U)$$

verifichiamo che $L = \liminf_{z\to x} f(z)$.

Sia ora $V \subseteq \overline{\mathbb{R}}$ un qualunque intorno inferiore aperto di L, per come abbiamo definito l'estremo superiore esisterà $U \in \mathscr{I}$ tale che $L(U) \in V$. Ma chiaramente V è un intorno aperto superiore anche di $L(U) = \inf_{z\in U\setminus\{x\}} f(z)$ e perciò $U \setminus \{x\} \subseteq f^{-1}(V)$.

Quindi L è uno dei possibili "limiti inferiori" per f, per dimostrare che è proprio L il limite inferiore ci basta verificare che è il più grande tra i possibili candidati. Sia allora $K \in \overline{\mathbb{R}}$ con $K > L$, dato che K non è l'estremo superiore degli $L(U)$ esisterà un intorno inferiore V di K tale che $L(U) \notin V$ per ogni intorno U di x (dato che gli intorni superiori di K sono intorni superiori anche di L). Dunque esisteranno $z_U \in U$ diversi da x tali che $f\left(z_U\right) \notin V$ e quindi $U \setminus \{x\} \nsubseteq f^{-1}(V)$ per alcun intorno di x. In altre parole K non è un buon candidato per essere l'estremo inferiore di f e perciò $\liminf_{z\to x} f(z) = L$.

Adesso per ogni coppia di intorni U e V di x

$$\inf_{z\in U\setminus\{x\}} f(z) \leq \inf_{z\in U\cap V\setminus\{x\}} f(z) \leq \sup_{z\in U\cap V\setminus\{x\}} f(z) \leq \sup_{z\in V\setminus\{x\}} f(z)$$

e facendo l'estremo superiore in U e l'estremo inferiore in V per quanto detto prima avremo che $\liminf_{z\to x} f(z) \leq \limsup_{z\to x} f(z)$. Supponiamo adesso che $\lim_{z\to x} f(z)$ esiste ed è pari ad L, dato che sia gli intorni superiori che quelli inferiori sono pur sempre intorni ne deduciamo che il limite inferiore ed il limite superiore devono anch'essi coincidere con L.

Viceversa supponiamo che $\liminf_{z\to x} f(z) = \limsup_{z\to x} f(z) = L$ e sia $U \subseteq \overline{\mathbb{R}}$ un intorno aperto di L, se $L \in \mathbb{R}$ allora esisterà $\varepsilon > 0$ per cui $\left(L-\varepsilon, L+\varepsilon\right) \subseteq U$. Posto così $S = \left[-\infty, L+\varepsilon\right)$ e $I = \left(L-\varepsilon, +\infty\right]$ avremo che $S \cap I \subseteq U$. Notiamo comunque che anche se $L = \pm\infty$ esisteranno sempre un intorno superiore aperto S ed un intorno inferiore aperto I di L tali che $S \cap I \subseteq U$.

Esisteranno così V_1, V_2 aperti di X contenenti x tali che $V_1 \setminus \{x\} \subseteq f^{-1}(I)$ ed $V_2 \setminus \{x\} \subseteq f^{-1}(S)$ e perciò

$$(V_1 \cap V_2) \setminus \{x\} \subseteq f^{-1}(S \cap I) \subseteq f^{-1}(U)$$

dunque poiché $V = U_1 \cap V_2$ è ancora aperto per l'arbitrarietà di U segue che $\lim_{z\to x} f(z) = L$. ■

In analisi funzionale in varie occasioni risulta comodo indebolire l'ipotesi di continuità di una funzione. Difatti alcune funzioni di particolare interesse non sono continue, ciononostante conservano ancora delle buone proprietà che discendono da una condizione più debole di continuità.

Definizione 2.6.3. Sia X spazio topologico e $f : X \to \overline{\mathbb{R}}$ funzione generica. Allora f è **semicontinua inferiormente** in $x \in X$ se e solo se

$$f(x) \leq \liminf_{y\to x} f(y) \Leftrightarrow f(x) \leq \sup_{U\in\mathscr{I}} \inf_{y\in U} f(y)$$

per un qualunque sistema fondamentale di intorni $\mathcal{I}$ di x. Se $f(x) \in \mathbb{R}$ possiamo riformularlo anche nella seguente versione, più agevole da maneggiare

$$\forall \epsilon > 0 \exists U \in \mathcal{I} \text{ tale che } \forall y \in U\, f(x) - \epsilon \leq f(y)$$

mentre se $f(x) = +\infty$ diventa

$$\forall N > 0 \exists U \in \mathcal{I} \text{ tale che } \forall y \in U\, N \leq f(y)$$

ovvero $\lim_{y\to x} f(y) = +\infty$.

Invece diciamo che f è **sequenzialmente semicontinua inferiormente** in $x \in X$ se e solo se per ogni x_n successione in X convergente ad x si ha

$$f(x) \leq \liminf_{n\to+\infty} f(x_n)$$

Con un procedimento analogo a quello utilizzato per le funzioni continue non è difficile dimostrare che funzioni semicontinue inferiormente sono anche sequenzialmente semicontinue inferiormente, e che se X soddisfa l'assioma $\mathcal{N}_1$ allora funzioni sequenzialmente semicontinue inferiormente sono anche semicontinue inferiormente.

Teorema 2.6.4. *Sia X spazio topologico sequenzialmente compatto e $f : X \to \mathbb{R} \cup \{+\infty\}$ sequenzialmente semicontinua inferiormente, allora f possiede un minimo in X.*

Dimostrazione. Posto $m = \inf_{x\in X} f(x)$ se $m = +\infty$ allora ogni suo punto è di minimo, altrimenti esisterà a meno di passare ad un'estratta una successione $x_n \in X$ e un elemento $x \in X$ tale che $x_n \to x$ e $f(x_n) \to m$. Segue immediatamente che

$$m \leq f(x) \leq \liminf_{n\to+\infty} f(x_n) = m$$

e quindi $f(x) = m \leq f(y)$ per ogni $y \in X$ e x è un punto di minimo per f. ■

Lemma 2.6.5. *Sia X spazio topologico e $f : X \to \overline{\mathbb{R}}$. Allora*

$$f \text{ semicontinua inferiormente} \Leftrightarrow \forall \alpha \in \mathbb{R}\ \left\{x \in X \,\middle|\, f(x) \leq \alpha\right\} \text{ è chiuso}$$

Dimostrazione. Supponiamo f semicontinua inferiormente e prendiamo $\alpha \in \mathbb{R}$ e $x \in X$ tale che $f(x) > \alpha$, quindi esisterà un $\epsilon > 0$ tale che $f(x) - \epsilon > \alpha$ per esempio

$$\epsilon = \frac{f(x) - \alpha}{2}$$

Esisterà allora un intorno di x contenuto interamente in $\left\{y \in X \,\middle|\, f(y) > \alpha\right\}$ che è perciò aperto.

Dimostriamo l'implicazione inversa, sia $x \in X$ tale che $f(x) \in \mathbb{R}$ e $\epsilon > 0$ allora l'insieme

$$\left\{y \in X \,\middle|\, f(y) > f(x) - \epsilon\right\}$$

è aperto quindi esiste un intorno $U \in \mathcal{I}$ di x contenuto in tale aperto, da cui segue che $f(x) \leq \liminf_{y\to x} f(y)$. Quest'ultima disuguaglianza è automaticamente verificata quando $f(x) = -\infty$, dobbiamo solamente verificare il caso $f(x) = +\infty$.

Innanzitutto osserviamo che

$$\{f = +\infty\} = \bigcap_{N\in\mathbb{N}} \{f > N\}$$

ovvero è intersezione di aperti, quindi per ogni $N \in \mathbb{N}$ esisterà un aperto U contenente x e contenuto in $\{f > N\}$ e quindi $\lim_{y\to x} f(y) = +\infty$. ■

Teorema 2.6.6. *Se X è uno spazio topologico compatto ed $f : X \to \mathbb{R} \cup \{+\infty\}$ è semicontinua inferiormente allora f ammette minimo in X.*

Dimostrazione. Supponiamo che esista $x \in X$ tale che $f(x) \in \mathbb{R}$ altrimenti il teorema è banalmente verificato. Sia $l = \inf\{f(x) \mid x \in X\} < +\infty$ e supponiamo che $l \neq -\infty$, allora per ogni $\varepsilon > 0$ l'insieme

$$C_\varepsilon = \{x \in X \mid f(x) \leq l + \varepsilon\}$$

per la proposizione precedente è chiuso in X ed è non vuoto per la definizione di estremo inferiore, quindi anche l'intersezione di un numero finito di essi è non vuota.

Per la definizione di compattezza tramite insiemi chiusi l'intersezione di tutti i C_ε al variare di ε è non vuota, quindi esisterà un $\bar{x} \in X$ tale che $f(\bar{x}) \leq l$ e quindi $f(\bar{x}) = \min f$ su X.

Se invece $l = -\infty$ allora possiamo ripetere l'intero procedimento ponendo però

$$C_\varepsilon = \left\{x \in X \,\middle|\, f(x) \leq -\frac{1}{\varepsilon}\right\}$$

ottenendo in tal modo un assurdo in quanto $\bigcap_{\varepsilon>0} C_\varepsilon = \emptyset$. ∎

I teoremi 2.6.4 e 2.6.6 sono conosciuti anche col nome di **teoremi di Weierstrass**.

2.7 Topologia iniziale

Una particolare classe di topologie è costituita dalle topologie generate da una classe di funzioni, tramite le quali possiamo trasportare la struttura topologica delle relative immagini sull'insieme di partenza in modo tale che queste funzioni risultino continue.

Fissiamo un generico insieme non vuoto X, un insieme degli indici A, una famiglia di spazi topologici $\{(Y_a, \tau_a)\}_{a\in A}$ ed una famiglia di funzioni $\mathscr{F} = \{f_a : X \to Y_a\}_{a\in A}$. Vogliamo in qualche modo trasportare la struttura topologica degli spazi Y_a su X mediante le funzioni f_a, e per farlo possiamo considerare la famiglia di insiemi nella forma

$$\mathscr{S} = \left\{f_a^{-1}(U) \subseteq X \,\middle|\, a \in A, U \in \tau_a\right\}$$

poiché $X \in \mathscr{S}$ per la proposizione 2.3.8 esisterà una topologia su X con semibase $\mathscr{S}$. Tale topologia verrà indicata con il simbolo $\mathscr{W}(\mathscr{F})$ e sarà detta **topologia iniziale** di X generata dalle funzioni f_a.

Se dotiamo X della topologia iniziale $\mathscr{W}(\mathscr{F})$ è immediato constatare che tutte le funzioni $f \in \mathscr{F}$ sono continue, viceversa $\mathscr{W}(\mathscr{F})$ è la topologia meno fine tra tutte quelle topologie su X che rendono le funzioni $f \in \mathscr{F}$ continue. Analizziamo adesso altre proprietà delle topologie iniziali.

Proposizione 2.7.1. *Data una famiglia di funzioni $\mathscr{F}$ come sopra, se tutti gli spazi Y_a sono di Hausdorff e se per ogni $x, y \in X$ con $x \neq y$ esiste $f \in \mathscr{F}$ tale che $f(x) \neq f(y)$ allora lo spazio topologico topologico $[X, \mathscr{W}(\mathscr{F})]$ è di Hausdorff.*

Dimostrazione. Prendiamo due punti di X distinti x ed y, esisterà allora una funzione $f_a : X \to Y_a$ in $\mathscr{F}$ tale che $f_a(x) \neq f_a(y)$. Ora essendo Y_a uno spazio di Hausdorff esistono due aperti disgiunti $U, V \subseteq Y_a$ tali che $f_a(x) \in U$ ed $f_a(y) \in V$ ovvero $x \in f_a^{-1}(U)$ e $y \in f_a^{-1}(V)$ ed $f_a^{-1}(U) \cap f_a^{-1}(V) = \emptyset$. ∎

Osserviamo inoltre che preso un qualunque punto $x \in X$ ed un sistema essenziale di intorni $\left\{V_{a,i}\right\}_{i\in I_a}$ di $f_a(x)$ in Y_a per ogni $a \in A$ la famiglia

$$I_{a,i} = f_a^{-1}\left(V_{a,i}\right) \tag{2.7.1}$$

è ancora un sistema essenziale di intorni di x rispetto allo spazio topologico $[X, \mathscr{W}(\mathscr{F})]$ con $\mathscr{F} = \{f_a : X \to Y_a\}_{a\in A}$.

Esempio 2.7.2. Sia X spazio topologico generico ed $Y \subseteq X$ non vuoto. Se consideriamo l'applicazione inclusione

$$i : y \in Y \to y \in X$$

chiaramente la famiglia $\mathscr{F} = \{i\}$ genera una topologia iniziale su Y che è la topologia meno fine tra quelle che rendono i continua. Ovviamente tale topologia è esattamente la topologia di sottospazio indotta da X su Y, possiamo così considerare la topologia di sottospazio come un caso particolare di topologia iniziale.

Le topologie iniziali di solito non sono $\mathscr{N}_1$ nemmeno quando gli spazi Y_a lo sono dato che la famiglia $\mathscr{F}$ è quasi sempre più che numerabile, quindi non sempre la continuità sequenziale implica la continuità. Ciononostante valgono altri risultati sulla continuità tra i quali ricordiamo i seguenti. Nei risultati che seguono assumeremo sempre che X è uno spazio topologico dotato della topologia iniziale $\mathscr{W}(\mathscr{F})$ indotta da una famiglia di funzioni $\mathscr{F} = \{f_\alpha\}_{\alpha\in A}$.

Proposizione 2.7.3. *Sia Z spazio topologico e $\psi : Z \to X$ un'applicazione generica. Allora*

$$\psi \textit{ è continua } \Leftrightarrow f_\alpha \circ \psi \textit{ è continua } \forall\, \alpha \in A$$

Dimostrazione. Chiaramente se ψ è continua allora tutte le applicazioni £$f_\alpha \circ \psi$ sono continue per ogni $\alpha \in A$, perciò supponiamo ora che $f_\alpha \circ \psi$ è continua per ogni $\alpha \in A$.

Per la proposizione 2.3.3 per dimostrare la continuità di ψ ci basta verificare che per ogni $x \in Z$ punto di accumulazione esiste un sistema essenziale di intorni $\mathscr{I}$ di $\psi(x)$ in X tale che $\psi^{-1}(I)$ è un intorno di x per ogni $I \in \mathscr{I}$. Ma gli insiemi nella forma

$$I_{\alpha,U^\alpha} = f_\alpha^{-1}\left(U^\alpha\right)$$

formano un sistema essenziale di intorni per $\psi(x)$ al variare di $U^\alpha \subseteq Y_\alpha$ aperto contenente $f_\alpha[\psi(x)]$ e così si avrà.

$$\psi^{-1}\left(I_{\alpha,U^\alpha}\right) = \psi^{-1}\left[f_\alpha^{-1}\left(U^\alpha\right)\right] = \left(f_\alpha \circ \psi\right)^{-1}\left(U^\alpha\right)$$

ma per ipotesi la funzione $f_\alpha \circ \psi$ è continua, così gli insiemi precedenti sono aperti in Z contenenti x. Abbiamo così dimostrato che il limite di p in x è esattamente $p(x)$ e così p è continua su tutto X. ■

Proposizione 2.7.4. *Fissati un punto $x \in X$ e una successione $x_n \in X$, allora*

$$x_n \to x \textit{ in } X \Leftrightarrow f_\alpha(x_n) \to f_\alpha(x) \textit{ in } Y_\alpha \; \forall\, \alpha \in A$$

Corollario 2.7.5. *Sia Z uno spazio topologico generico e $\psi : Z \to X$ applicazione generica. Allora ψ è sequenzialmente continua su X se e solo se $f_\alpha \circ \psi : Z \to Y_\alpha$ è sequenzialmente continua per ogni $\alpha \in A$.*

$$\begin{tikzcd}
Z \arrow[r, "f_\alpha"] \arrow[d, "\psi"'] & Y_\alpha \\
X \arrow[ur, dashed, "f_\alpha \circ \psi"'] &
\end{tikzcd}$$

Figura 2.1: Rappresentazione grafica delle funzioni nella proposizione 2.7.3

I due risultati precedenti possono essere riassunti nel seguente modo: se Y e Z sono spazi topologici e X è dotato della topologia iniziale $\mathscr{W}(\mathscr{F})$ i cui elementi hanno codominio Y. Allora per ogni applicazione $\psi : Z \to X$

- ψ è sequenzialmente continua $\Leftrightarrow f_\alpha \circ \psi$ è sequenzialmente continua per ogni $\alpha \in A$;
- ψ è continua $\Leftrightarrow f_\alpha \circ \psi$ è continua per ogni $\alpha \in A$.

In generale le topologie iniziali non sono $\mathscr{N}_1$ per cui bisogna sempre distinguere continuità e sequenziale continuità.

La proposizione 2.7.3 è molto importante in quanto fornisce un criterio di continuità molto utile in seguito. Una prima conseguenza è dato dal seguente risultato:

Proposizione 2.7.6. *Supponiamo che la topologia di X è la topologia iniziale generata dalla famiglia di funzioni F. Dati E insieme generico ed $p : E \to X$ applicazione qualunque allora la topologia iniziale su E generata da $\{p\}$ coincide con la topologia iniziale generata dalla famiglia*

$$F' = \{f_a \circ p : E \to Y_a \mid f_a \in F\}$$

Dimostrazione. Rispetto alla topologia $\mathscr{W}(\{p\})$ l'applicazione p è continua da E in X, dunque anche $f \circ p$ è continua per ogni $f \in F$ dato che X è dotato di una topologia iniziale, perciò abbiamo $\mathscr{W}(\mathscr{F}') \subseteq \mathscr{W}(\{p\})$.

viceversa se consideriamo la topologia $\mathscr{W}(\mathscr{F}')$ su E avremo che $f \circ p$ è continua per ogni $f \in F$, ma grazie alla proposizione 2.7.3 ciò implica immediatamente che l'applicazione $p : E \to X$ è continua anche quando E è dotato della topologia $\mathscr{W}(\mathscr{F}')$ e perciò $\mathscr{W}(\{p\}) \subseteq \mathscr{W}(\mathscr{F}')$. ■

Corollario 2.7.7. *Se X è uno spazio topologico con la topologia iniziale indotta dalla famiglia di funzioni F ed $Y \subseteq X$ allora la topologia di Y indotta da X coincide con la topologia iniziale di Y generate dalle funzioni in F dopo averne ristretto il dominio ad Y.*

Topologia prodotto e convergenza puntuale

Data una famiglia di spazi topologici $\mathscr{X} = \{X_i\}_{i \in I}$ possiamo dotare l'insieme $F(I.\mathscr{X})$ della topologia iniziale indotta dalle proiezioni $\pi_i : F(I, \mathscr{X}) \to X_i$ come detto prima. Questa topologia è chiamata **topologia prodotto** di $F(I, \mathscr{X}) = \prod_{i \in I} X_i$.

In particolare una successione di funzioni $f_n \in F(X, Y) = \prod_{x \in X} Y$ converge ad f con la topologia prodotto se e solo se $f_n(x)$ converge a $f(x)$ in Y per ogni $x \in X$. Per questo motivo la topologia prodotto sullo spazio $F(X, Y)$ è detta anche topologia della **convergenza puntuale**.

Proposizione 2.7.8. *Dati una famiglia di spazi topologici $\mathscr{X} = \{X_i\}_{i\in I}$, un insieme non vuoto J ed un'applicazione $f : J \to I$ allora il pull-back $f_\# : F(I,\mathscr{X}) \to F(J,\mathscr{X}\circ f)$ (con $\mathscr{X}\circ f(j) = \mathscr{X}[f(j)] = X_{f(j)}$) definito come*

$$\left[f_\#(x)\right]_j = x_{f(j)}$$

è un'applicazione continua rispetto alla topologia prodotto.

Dimostrazione. Siano π_i le proiezioni su $F(I,\mathscr{X})$ e η_j le proiezioni su $F(J,\mathscr{X}\circ f)$. Per ogni $x \in F(I,\mathscr{X})$ e per ogni $j \in J$ si ha

$$\eta_j\left[f_\#(x)\right] = \left[f_\#(x)\right]_j = x_{f(j)} = \pi_{f(j)}(x)$$

ovvero $\eta_j \circ f_\# = \pi_{f(j)}$ che è continua. Per l'arbitrarietà di j segue che $f_\#$ è continua. ■

Osservazione. Se $J \subseteq I$ e $f(j) = j$ allora la restrizione del dominio $|_J = f_\#$ è continua.

Un notevole risultato sulle topologie prodotto che utilizzeremo in seguito è il seguente

Teorema 2.7.9 (Tychonoff)**.** *Se gli spazi topologici X_i sono compatti per ogni $i \in I$ allora anche $\prod_{i\in I} X_i$ è compatto rispetto alla topologia prodotto.*

Per maggior chiarezza enunceremo i prossimi risultati limitandoci a considerare esclusivamente prodotti finiti nonostante valgano, con opportune modifiche, anche per prodotti qualunque. In questo modo il lettore ritroverà tutti i risultati noti tramite la definizione naïf di prodotto cartesiano visto come insieme di n-ple ordinate.

Siano $Y_1, Y_2, \ldots, Y_n$ spazi topologici e $X = \prod_{i=1}^n Y_i$ dotato della topologia prodotto. Allora

Proposizione 2.7.10. *Gli insiemi nella forma $U_1 \times U_2 \times \cdots \times U_n$, dove U_i è un aperto di Y_i, formano una base per la topologia di X.*

Inoltre

$$\mathrm{Cl}\left(A_1 \times \cdots \times A_n\right) = \mathrm{Cl}\left(A_1\right) \times \cdots \times \mathrm{Cl}\left(A_n\right)$$

per ogni $A_i \subseteq Y_i$.

Proposizione 2.7.11. *Un'applicazione tra spazi topologici $f : Z \to X$ è continua se e solo se tutte le proiezioni i-esime $f_i = \pi_i \circ f$ sono continue da Z in Y_i. Lo stesso vale per la continuità sequenziale.*

Proposizione 2.7.12. *Lo spazio X è di Hausdorff se e solo se tutti gli Y_i sono di Hausdorff, inoltre X è regolare se e solo se tutti gli Y_i sono regolari.*

2.8 Topologia finale

La topologia iniziale introdotta nella precedente sezione permette di trasportare la topologia dal codominio al dominio di una o più funzioni. È possibile in realtà effettuare anche il procedimento inverso, ovvero trasportare la topologia di una funzione dal dominio al codominio di una o più funzioni in modo da renderle continue. A differenza però della topologia iniziale questo procedimento verrà utilizzato esclusivamente per poter parlate di topologia quoziente e quindi in queste note verrà trattata solo brevemente.

In questo caso $\{(X_a,\tau_a)\}_{a\in A}$ è una famiglia di spazi topologici generici ed $\mathscr{F}$ è una famiglia di funzioni nella forma $f_a : X_a \to Y$ che partono da ciascun X_a ed hanno codominio tutte nello stesso insieme non vuoto Y. Vogliamo così portare la topologia delle X_a su Y.

Definiamo allora

$$\tau = \left\{ U \subseteq Y \middle| \; f_a^{-1}(U) \in \tau_a \; \forall a \in A \right\}$$

ovvero $U \in \tau$ se e solo se $f_a^{-1}(U)$ è un aperto di X_a per ogni $a \in A$. Chiaramente $\emptyset, Y \in \tau$ e τ è stabile rispetto all'intersezione finita e all'unione arbitraria di suoi elementi. Tale insieme è dunque una topologia su Y e sarà detta **topologia finale** generata dalla famiglia di funzioni $F = \{f_a\}_{a \in A}$.

Proposizione 2.8.1. *La topologia finale su Y generata da F è la più fine topologia su Y che rende tutte le funzioni f_a continue.*

Dimostrazione. Chiaramente le f_a sono continue rispetto alla topologia finale su Y. Viceversa considerata una qualunque topologia τ' su Y che renda le f_a continue avremo che $f_a^{-1}(U)$ è un aperto di X_a per ogni $u \in \tau'$ per continuità. Ma allora per definizione di topologia iniziale si ha $U \in \tau$ e così $\tau' \subseteq \tau$. ∎

Proposizione 2.8.2. *Se Y è dotato della topologia finale generata dalla famiglia di funzioni $F = \{f_a\}_{a \in A}$ ed Z è un qualunque spazio topologico allora $\psi : Y \to Z$ è continua se e solo se $\psi \circ f_a$ è continua da X_a in Z per ogni $a \in A$.*

Dimostrazione. Se ψ fosse continua per la proposizione precedente anche le $\psi \circ f_a$ devono essere continue. Viceversa supponiamo che ψ sia una funzione per cui $\psi \circ f_a$ è continua per ogni $a \in A$ e prendiamo un aperto generico U di Z. Per continuità di $\psi \circ f_a$ avremo allora che

$$f_a^{-1}\left[\psi^{-1}(U)\right] = \left(\psi \circ f_a\right)^{-1}(U) \in \tau_a$$

e quindi per definizione di topologia finale $\psi^{-1}(U) \in \tau$. La funzione ψ è così continua da X in Z. ∎

Corollario 2.8.3. *Se Y è dotato della topologia finale generata dalla famiglia di funzioni $F = \{f_a\}_{a \in A}$ ed Z è un qualunque spazio topologico allora $\psi : Y \to Z$ è sequenzialmente continua se e solo se $\psi \circ f_a$ è sequenzialmente continua da X_a in Z per ogni $a \in A$.*

Dalla proposizione 2.4.4 segue immediatamente il seguente criterio di compattezza

Proposizione 2.8.4. *Se esiste $a \in A$ tale che X_a è compatto/sequenzialmente compatto ed $f_a : X_a \to Y$ è suriettiva allora lo spazio Y con la topologia finale è compatto/sequenzialmente compatto.*

Topologia quoziente

Consideriamo adesso un qualunque spazio topologico X e una qualunque relazione di equivalenza $\sim$ su X, possiamo indurre su $X/\sim$ una nuova topologia, detta **topologia quoziente**, che non sarà altro che la topologia finale su $X/\sim$ generata dalla singola funzione $\pi_\sim$ definita su X.

Non è difficile verificare inoltre che

$$A' \text{ aperto in } X/\sim \Leftrightarrow \pi_\sim^{-1}(A') = \bigcup_{[a]_\sim \in A'} [a]_\sim \text{ aperto in } X$$

Si noti però che gli $[a]_\sim$ potrebbero non essere aperti nella topologia quoziente.

Chiaramente se X fosse compatto/sequenzialmente compatto allora anche $X/\sim$ lo è. Inoltre la topologia su $X/\sim$ è la topologia più fine che renda l'applicazione $\pi_\sim$ continua e vale il seguente risultato:

Proposizione 2.8.5. *Dati due spazi topologici X, Y e una relazione di equivalenza $\sim$ su X. Allora data una qualunque applicazione $f : X/\sim \to Y$ avremo che f è continua se e solo se $f \circ \pi_\sim$ è continua.*

Capitolo 3

Spazi uniformi

Questo capitolo si dedica allo studio dei cosiddetti spazi uniformi, ovvero spazi topologici che possiedono un'ulteriore struttura grazie alla quale possiamo parlare di funzioni uniformemente continue e di successione di Cauchy, concetti che ritroveremo più avanti negli spazi metrici e negli spazi vettoriali topologici.

3.1 Strutture uniformi

Consideriamo un generico insieme X che non sia vuoto, per ogni coppia di sottoinsiemi U, V del prodotto cartesiano $X \times X$ definiamo $U \diamond V \subseteq X \times X$ in modo tale che $(x, y) \in U \diamond V$ se e solo se esiste $z \in X$ tale che $(x, z) \in U$ ed $(z, y) \in V$. Poniamo anche $U^{-1} = \{(y, x) \mid (x, y) \in U\}$ e diremo che U è simmetrico se e solo se $U = U^{-1}$.

Chiaramente l'operazione $\diamond$ è associativa, ovvero

$$(U \diamond V) \diamond W = U \diamond (V \diamond Z) = U \diamond V \diamond Z$$

però di solito non è commutativa. In compenso se U e V sono simmetrici allora $U \diamond V = V \diamond U$.

Per ogni $U \subseteq X \times X$ e per ogni $A \subseteq X$ non vuoto definiamo inoltre

$$U[A] = \{x \in X \mid \exists a \in A \text{ tale che } (x, a) \in U\}$$

in particolare $U[x] = \{y \mid (y, x) \in U\}$ ed $U[A] = \bigcup_{a \in A} U[a]$.

Definizione 3.1.1. Presi un insieme non vuoto X ed una famiglia $\mathscr{T} \subseteq \mathcal{P}(X \times X)$ di sottoinsiemi del prodotto cartesiano $X \times X$. Diremo che $\mathscr{T}$ è una **struttura uniforme** su X se e solo se $X \times X \in \mathscr{T}$ e per ogni $U \in \mathscr{T}$ valgono:

1. $(x, x) \in U$ per ogni $x \in X$;
2. se $U \subseteq V \subseteq X \times X$ allora $V \in \mathscr{T}$;
3. per ogni $V \in \mathscr{T}$ si ha $U \cap V \in \mathscr{T}$;
4. si ha $U^{-1} \in \mathscr{T}$;
5. esiste $V \in \mathscr{T}$ tale che $V \diamond V \subseteq U$.

In tal caso gli elementi di $\mathcal{T}$ saranno detti **entourage**.

Osservazione. Se $U \in \mathcal{T}$ ed $\mathcal{T}$ è una struttura uniforme avremo chiaramente che $V = U \cap U^{-1} \in \mathcal{T}$ e V è simmetrico. Quindi ogni entourage conterrà a sua volta un entourage simmetrico al suo interno.

Dalle proprietà appena elencate segue immediatamente che $A \subseteq U[A]$ per ogni $A \subseteq X$, $U \in \mathcal{T}$ e che se $y \in U[x]$ per qualche $x, y \in X$ allora $x \in U^{-1}[y]$. Inoltre

$$(U \diamond V)[A] = \bigcup_{a\in A} \bigcup_{y \in V[a]} U[y] = U[V[A]]$$

che combinata con la proprietà 1 ci permette di affermare che $U \cup V \subseteq U \diamond V$. Per chi conoscesse già gli spazi metrici non può non notare le affinità degli entourage con le metriche.

Proposizione 3.1.2. *Siano $A, B \subseteq X$ non vuoti ed $U \in \mathcal{T}$ entourage, allora*

$$U[B] \cap A \neq \emptyset \Leftrightarrow B \cap U^{-1}[A] \neq \emptyset \tag{3.1.1}$$

Dimostrazione. Esiste $x \in U[B] \cap A$ se e solo se esiste un $y \in B$ tale che $(x, y) \in U$ e dunque $(y, x) \in U^{-1}$ ovvero $y \in B \cap U^{-1}[A]$. ■

Osservazione. Non stiamo dicendo che $U[B] \cap A = B \cap U^{-1}[A]$, anzi i due sottoinsieme di solito sono distinti.

Un sottoinsieme non vuoto $\mathcal{S}$ di $\mathcal{P}(X \times X)$ è una **base uniforme** per la struttura uniforme $\mathcal{T}$ se e solo se è contenuta in $\mathcal{T}$ ed ogni elemento di $\mathcal{T}$ contiene qualche elemento di $\mathcal{S}$, invece diremo che $\mathcal{S}'$ è una **semibase uniforme** se e solo se la famiglia

$$\mathcal{S} = \left\{ U_1 \cap U_2 \cap \cdots \cap U_n \;\middle|\; n \in \mathbb{N}, U_i \in \mathcal{S}' \right\}$$

è una base uniforme per $\mathcal{T}$. Per la proprietà 2 infatti in una struttura uniforme contano più gli insiemi piccoli che non quelli grandi.

Esempio 3.1.3. Poniamo $X = \mathbb{R}^n$ per qualche $n \in \mathbb{N}$ con la topologia usuale. Se prendiamo un qualunque intorno dell'origine I e definiamo $\tilde{I} \subseteq \mathbb{R}^n \times \mathbb{R}^n$ in modo tale che

$$(x, y) \in \tilde{I} \Leftrightarrow x - y \in I$$

allora gli $\tilde{I}$ formano una struttura uniforme per $\mathbb{R}^n$ che genera la topologia usuale di $\mathbb{R}^n$. Inoltre la famiglia di insiemi

$$U(\varepsilon) = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \;\middle|\; \left| x - y \right| < \varepsilon \right\}$$

al variare di $\varepsilon > 0$ è una base uniforme per $\mathbb{R}^n$. Da questo momento $\mathbb{R}^n$ sarà uno spazio uniforme dotato di tale struttura, se non diversamente specificato.

Come per la proposizione 2.3.7 per quanto riguarda gli spazi topologici vale un risultato del tutto analogo per quanto riguarda le basi uniformi:

Proposizione 3.1.4. *Presa una famiglia non vuota $\mathcal{S} \subseteq \mathcal{P}(X \times X)$ che soddisfa le seguenti proprietà:*

1. $(x, x) \in S$ *per ogni* $x \in X$ *e per ogni* $S \in \mathcal{S}$*;*

2. *per ogni* $S, T \in \mathscr{S}$ *esiste* $U \in \mathscr{S}$ *tale che* $U \subseteq S \cap T$;

3. *per ogni* $S \in \mathscr{S}$ *si ha* $(x, y) \in S$ *se e solo se* $(y, x) \in S$;

4. *per ogni* $S \in \mathscr{S}$ *esiste* $T \in \mathscr{S}$ *tale che* $T \diamond T \subseteq S$.

Allora la famiglia di insiemi

$$\mathscr{T} = \left\{T \subseteq X \times X \mid \exists S \in \mathscr{S} \text{ tale che } S \subseteq T\right\}$$

è l'unica struttura uniforme su X *che ha* $\mathscr{S}$ *come base.*

Se invece $\mathscr{S}'$ *soddisfa solamente i punti 1, 3, 4 allora esiste un'unica struttura uniforme* $\mathscr{T}$ *che ha* $\mathscr{S}'$ *come semibase uniforme.*

Ora che abbiamo introdotto il concetto di struttura uniforme possiamo considerare associata a tale struttura.

Definizione 3.1.5. Dati un insieme non vuoto X, una topologia τ ed una struttura uniforme $\mathscr{T}$ su X diremo che $(X, \tau, \mathscr{T})$ è uno **spazio uniforme** se e solo se per ogni $x \in X$ la famiglia $\{U[x] \mid U \in \mathscr{T}\}$ è un sistema fondamentale di intorni in x per la topologia τ.

Notare che gli elementi $V[x]$ potrebbero non essere aperti in tale topologia. Inoltre se $\mathscr{S}$ è una base uniforme di $\mathscr{T}$ allora la famiglia di insiemi $V[x]$ al variare di $V \in \mathscr{S}$ è comunque un sistema fondamentale di intorni per x.

Proposizione 3.1.6. *Gli spazi uniformi sono regolari, in particolare se* $(X, \mathscr{T})$ *è uno spazio uniforme* $C \subseteq X$ *chiuso ed* $K \subseteq X$ *compatto tali che* $C \cap K = \emptyset$ *allora esiste* $U \in \mathscr{T}$ *simmetrico tale che*

$$U[C] \cap U[K] = \emptyset$$

Inoltre uno spazio uniforme è di Hausdorff se e solo se l'intersezione di tutti gli entourage coincide con l'insieme $\Delta = \left\{(x, x) \;\middle|\; x \in X\right\}$.

Dimostrazione. Per definizione di spazio uniforme per ogni $x \in K$ esiste $A_x \in \mathscr{T}$ tale che $A_x[x] \cap C = \emptyset$, dato che gli $A_x[x]$ sono intorni di x per compattezza possiamo scegliere degli $B_1, B_2, \dots, B_n \in \mathscr{T}$ e dei punti $x_1, x_2, \dots, x_n \in K$ tali che

$$B_i \diamond B_i \subseteq A_{x_i}$$
$$K \subseteq \bigcup_{i=1}^{n} U_i[x_i]$$

con U_i simmetrico in modo tale che $U_i \diamond U_i \subseteq B_i$.

Per la proprietà di intersezione finita esiste A entourage simmetrico tale che $U \subseteq U_i$ per ogni $1 \leq i \leq n$, mostriamo che $U[K] \subseteq \bigcup_{i=1}^{n} B_i[x_i]$. Sia $y \in U[K]$ per definizione esiste $x \in K$ tale che $(y, x) \in U$ e quindi esiste i tale che $(x, x_i) \in U_i$, ma allora $(y, x_i) \in U_i \diamond U_i \subseteq B_i$ come volevasi dimostrare.

Adesso per la (3.1.1) si ha $x_i \notin A_{x_i}^{-1}[C]$ e dunque a maggior ragione $x_i \notin (B_i \diamond B_i)^{-1}[C] = B_i^{-1}[B_i^{-1}[C]]$ e quindi applicando di nuovo la (3.1.1) $B_i[x_i] \cap B_i^{-1}[C] = \emptyset$. Abbiamo quindi che

$$U[K] \cap U[C] \subseteq \bigcup_{i=1}^{n} B_i[x_i] \cap U[C] \subseteq \bigcup_{i=1}^{n} \left(B_i[x_i] \cap B_i^{-1}[C]\right) = \emptyset$$

Dato che $B[y]$ è un intorno di y per ogni y esistono due aperti U e V contenenti K e C rispettivamente tali che $U \cap V = \emptyset$ e quindi X è uno spazio regolare.

Supponiamo adesso che l'intersezione degli entourage coincide con Δ, se prendiamo $x, y \in X$ distinti esisterà $U \in \mathscr{T}$ tale che $(x, y) \notin U$. Preso $V \in \mathscr{T}$ per cui $V \diamond V \subseteq U$ avremo che $x \notin V[V[y]]$ e per la (3.1.1) $V^{-1}[x] \cap V[y] = \emptyset$ ed X è uno spazio di Hausdorff.

Per l'implicazione opposta sappiamo già che l'intersezione di tutti gli entourage conterrà sicuramente Δ, quando X è uno spazio di Hausdorff se $(x, y) \in U$ per ogni $U \in \mathscr{T}$ allora avremo che $x \in U[y]$. Poiché esso è un sistema fondamentale di intorni per y segue che $x = y$ e dunque l'intersezione degli entourage deve coincidere con Δ. ■

Per ogni sottoinsieme $A \subseteq X$ non vuoto e per ogni entourage $U \subseteq \mathscr{T}$ possiamo definire i seguenti sottoinsiemi di X

$$\mathrm{Cl}_U(A) = \left\{x \in X \;\middle|\; x \in U[y] \text{ per qualche } y \in A\right\} = U[A] \tag{3.1.2}$$

$$\mathrm{Int}_U(A) = \{x \in X \mid U[x] \subseteq A\} \tag{3.1.3}$$

come suggeriscono i simboli utilizzati questi insiemi ci permettono di costruire la chiusura e l'interno di A nella seguente maniera:

Proposizione 3.1.7. *Per ogni $A \subseteq X$ non vuoto e per ogni base $\mathscr{S}$ della struttura uniforme $\mathscr{T}$ su X*

$$\mathrm{Int}(A) = \bigcup_{U \in \mathscr{S}} \mathrm{Int}_U(A)$$

$$\mathrm{Cl}(A) = \bigcap_{U \in \mathscr{S}} \mathrm{Cl}_U(A)$$

Dimostrazione. Poiché gli insiemi nella forma $U[x]$ formano un sistema fondamentale di intorni per x avremo che

$$\begin{aligned} x \in \mathrm{Int}(A) &\Leftrightarrow \exists U \in \mathscr{S} \text{ tale che } U[x] \subseteq A \\ &\Leftrightarrow \exists U \in \mathscr{S} \text{ tale che } x \in \mathrm{Int}_U(A) \\ &\Leftrightarrow x \in \bigcup_{U \in \mathscr{S}} \mathrm{Int}_U(A) \end{aligned}$$

ed

$$\begin{aligned} x \in \mathrm{Cl}(A) &\Leftrightarrow \forall U \in \mathscr{S} \;\; U[x] \cap A \neq \emptyset \\ &\Leftrightarrow \forall U \in \mathscr{S} \;\; x \in U[A] = \mathrm{Cl}_U(A) \\ &\Leftrightarrow x \in \bigcap_{U \in \mathscr{S}} \mathrm{Cl}_U(A) \end{aligned}$$

avendo usato la (3.1.1). ■

Dato una struttura uniforme $\mathscr{T}$ su X per ogni sottoinsieme Y di X non vuoto possiamo dotare su Y della seguente struttura uniforme

$$\mathscr{T} \cap Y = \{U \cap (Y \times Y) \mid U \in \mathscr{T}\}$$

chiaramente essa è ancora una struttura uniforme su Y, inoltre se dotiamo X della topologia generata dalla struttura uniforme $\mathscr{T}$ allora la struttura uniforme $\mathscr{T} \cap Y$ induce su Y esattamente la topologia indotta da X sul sottoinsieme Y.

3.2 Funzioni uniformemente continue

Definizione 3.2.1. Una funzione $f : X \to Y$ tra due spazi uniformi X ed Y con strutture uniformi $\mathscr{S}$ ed $\mathscr{T}$ rispettivamente è detta **uniformemente continua** se e solo se per ogni $V \in \mathscr{T}$ si ha

$$(f \times f)^{-1}(V) = \{(x, y) \in X \times X \mid (f(x), f(y)) \in V\} \in \mathscr{S}$$

Osserviamo che se f è uniformemente continua allora per ogni $x \in X$ e $V \in \mathscr{T}$ avremo che

$$(f \times f)^{-1}(V)[x] = f^{-1}(V[f(x)])$$

e in particolare $f^{-1}(V[f(x)])$ è un intorno di x. Dato che $V[f(x)]$ è un sistema fondamentale di intorni per $f(x)$ ne segue che f è continua in x e dunque è continua su tutto X. Le funzioni uniformemente continue sono allora anche continue, inoltre se $\mathscr{S}$ o $\mathscr{T}$ possiedono delle basi uniformi è possibile limitarsi ad utilizzare quelle al posto delle intere strutture uniformi.

Diremo che f è un **omeomorfismo uniforme** se e solo se è biettiva, uniformemente continua con inversa anch'essa uniformemente continua, mentre due spazi uniformi X ed Y sono **uniformemente equivalenti** se e solo se esiste un omeomorfismo uniforme da X in Y.

Dimostriamo adesso la seguente ben nota proprietà in analisi.

Proposizione 3.2.2. *Siano $(X, \tau, \mathscr{S})$, $(Y, \eta, \mathscr{T})$ spazi topologici uniformi con X compatto. Se $f : X \to Y$ è continua allora è anche uniformemente continua.*

Dimostrazione. Preso un qualunque entourage V in Y sappiamo che esiste $\tilde{V} \in \mathscr{T}$ con $\tilde{V} \diamond \tilde{V} \subseteq V$. Per continuità sappiamo che per ogni $x \in X$ e per ogni $V \in \mathscr{T}$ esiste $U_x \in \mathscr{S}$ tale che $(U_x \diamond U_x)[x] \subseteq f^{-1}(\tilde{V}[f(x)]) = (f \times f)^{-1}(\tilde{V})[x]$ per ogni x.

La compattezza di X implica che esistono $x_1, x_2, \dots, x_n \in X$ tali che $X = \bigcup_i U_{x_i}[x_i]$. Posto $U = \bigcap_i U_i \in \mathscr{S}$ per ogni $(x, y) \in U$ esiste $1 \le i \le n$ per cui $y \in U_{x_i}[x_i]$ ma allora $x \in U\left[U_{x_i}[x_i]\right]$ e così

$$\begin{aligned} y &\in U_{x_i}[x_i] \subseteq (f \times f)^{-1}(\tilde{V})[x_i] \\ x &\in \left(U_{x_i} \diamond U_{x_i}\right)[x_i] \subseteq (f \times f)^{-1}(\tilde{V})[x_i] \end{aligned}$$

e per simmetria si avrà dunque

$$(x, y) \in (f \times f)^{-1}(\tilde{V}) \diamond (f \times f)^{-1}(\tilde{V}) \subseteq (f \times f)^{-1}(\tilde{V} \diamond \tilde{V}) \subseteq (f \times f)^{-1}(V)$$

ovvero $U \subseteq (f \times f)^{-1}(V)$ e per definizione di struttura uniforme $(f \times f)^{-1}(V) \in \mathscr{S}$ ed f è uniformemente continua. ■

Osservazione. Il risultato precedente non vale se Y è compatto ma X non lo è, un controesempio è dato dalla funzione $f(t) = \sin(1/t)$ definita su $\mathbb{R} \setminus \{0\}$.

Consideriamo adesso una famiglia di spazi topologici uniformi $\{(Y_a, \tau_a, \mathscr{T}_a)\}_{a \in A}$ e di funzioni $\mathscr{F} = \{f_a : X \to Y_a\}_{a \in A}$, possiamo considerare al variare di $a \in A$ e di $U_a \in \mathscr{T}_a$ i seguenti insiemi $T_{a, U_a} = (f_a \times f_a)^{-1}(U_a)$. Vogliamo ora usare la proposizione nel capitolo precedente per dimostrare che $\tilde{T} = \left\{T_{a, U_a}\right\}$ è una struttura uniforme su X che genera la topologia iniziale indotta dalle funzioni f_a al variare di $a \in A$.

Innanzitutto poiché gli U_a sono entourage di Y_a avremo che $[f_a(x), f_a(x)] \in U_a$ per ogni $a \in A$, $U_a \in \mathscr{T}_a$ ovvero $(x,x) \in T_{a,U_a}$. Che gli insiemi T_{a,U_a} siano simmetrici è banale da verificare, dunque dobbiamo verificare solo la proprietà di composizione. Innanzitutto per ogni $U_a \in \mathscr{T}_a$ esiste $V_a \in \mathscr{T}_a$ tale che $V_a \diamond V_a \subseteq U_a$. Ma allora preso $(x,y) \in T_{a,V_a} \diamond T_{a,V_a}$ esisterà $z \in X$ tale che $(x,z),(z,y) \in T_{a,V_a}$ ovvero $[f_a(x), f_a(z)], [f_a(z), f_a(y)] \in V_a$, $[f_a(x), f_a(y)] \in V_a \diamond V_a \subseteq U_a$ e dunque

$$T_{a,V_a} \diamond T_{a,V_a} \subseteq T_{a,V_a \diamond V_a} \subseteq T_{a,U_a}$$

Osservazione. Si noti però che $T_{a,V_a} \diamond T_{a,V_a} \neq T_{a,V_a \diamond V_a}$ a meno che f_a non sia suriettiva.

Dunque la famiglia

$$\left\{ R \subseteq X \times X \;\middle|\; T_{a_1,U_{a_1}} \cap T_{a_1,U_{a_1}} \cap \cdots \cap T_{a_1,U_{a_1}} \subseteq R \text{ per qualche } n \in \mathbb{N}, a_i \in A, U_{a_i} \in \mathscr{T}_{a_i} \right\}$$

è una struttura uniforme su X con gli insiemi T_{a,U_a} come semibase. Chiaramente le funzioni f_a sono uniformemente continue rispetto a tale struttura, ed in particolare sono continue rispetto alla topologia τ generata da essi e dunque $\mathscr{W}(\mathscr{F}) \subseteq \tau$.

Viceversa presi $x \in X$, $a \in A$ ed $U_a \in \mathscr{T}_a$ sappiamo che esiste un aperto $V \subseteq Y_a$ tale che

$$f_a(x) \in V \subseteq U_a[f_a(x)]$$

e perciò $x \in f_a^{-1}(V) \subseteq f_a^{-1}(U_a[f_a(x)]) = (f_a \times f_a)^{-1}(U_a)[x]$ ovvero gli elementi nella forma $T_{a,U_a}[x]$ sono intorni di x anche rispetto alla topologia $\mathscr{W}(\mathscr{F})$. Dato che tali insiemi formano un sistema essenziale di intorni per τ la proposizione 2.3.5 implica dunque che $\mathscr{W}(\mathscr{F}) = \tau$.

Abbiamo quindi che

Teorema 3.2.3. *Le topologie iniziali sono generate da una struttura uniforme.*

teniamo comunque ben presente che uno stesso spazio topologico può essere generato da diverse strutture uniformi che possono tranquillamente non essere equivalenti. È perciò importante specificare se due spazi sono equivalenti solamente dal punto di vista topologico o anche per quanto riguarda lo spazio uniforme. Ciononostante quando si vedranno gli spazi vettoriali topologici introdurremo du si essi una struttura uniforme a partire dalla loro topologia, così da identificare struttura topologica ed uniforme per gli spazi vettoriali.

La funzione interpolante

Definizione 3.2.4. Uno spazio topologico X si dice **completamente regolare** se e solo se per ogni $x \in X$ e per ogni chiuso $C \subseteq X$ per cui $x \notin C$ esiste una funzione continua $f : X \to [0,1]$ tale che $f(x) = 1$ ed $f(y) = 0$ per ogni $y \in C$.

Osservazione. Equivalentemente per ogni aperto non vuoto $U \subseteq X$ e per ogni $x \in U$ esiste $f : X \to [0,1]$ tale che $f(x) = 1$ ed $f(y) = 0$ per ogni $y \notin U$. In particolare dati due aperti U_1, U_2 con $x \in U_1 \cap U_2$ se f_1, f_2 sono due funzioni continue come sopra definite a partire da U_1 e U_2 rispettivamente posto

$$f(z) = f_1(z) f_2(z)$$

avremo chiaramente che $f(x) = 1$ ed $f(y) = 0$ per ogni $y \in X \setminus (U_1 \cap U_2)$.

Dalla definizione segue immediatamente che

Proposizione 3.2.5. *Lo spazio topologico X è completamente regolare se e solo se per ogni $C \subseteq X$ chiuso esiste $F \subseteq C(X)$ tali che $C = \bigcap_{f \in F} \{f = 0\}$.*

Non è difficile verificare che gli spazi completamente regolari sono anche regolari. Bisogna tener presente che l'implicazione opposta non è in generale vera, ed esistono spazi regolari ma non completamente regolari. Se indichiamo con $C(X)$ l'insieme delle funzioni continue da X in $\mathbb{R}$ abbiamo il seguente risultato:

Teorema 3.2.6. *Se gli spazi topologici Y_a sono completamente regolari al variare di $a \in A$ allora presa una famiglia di funzioni $\mathscr{F} = \{f : X \to Y_a\}$ anche $[X, \mathscr{W}(\mathscr{F})]$ è completamente regolare.*

Dimostrazione. Per l'osservazione precedente possiamo ridurci a considerare $x \in X$ e gli elementi del sistema essenziale di intorni T_{a,U_a} per cui $x \in T_{a,U_a}$. Allora $f_a(x) \in U_a$ ed essendo Y_a completamente regolare esisterà $F \in C\left(Y_a, [0,1]\right)$ tale che $F\left[f_a(x)\right] = 1$ ed $F(y) = 0$ se $y \notin U_a$.

Posto così $G = F \circ f_a$ chiaramente è continua da X in $[0,1]$ e vale 1 in x. Infine $y \notin T_{a,U_a} \Leftrightarrow f_a(y) \notin U_a$ e dunque G è la funzione cercata. Dato che gli T_{a,U_a} formano un sistema essenziale di intorni aperti di x lo spazio X con la topologia $\mathscr{W}(\mathscr{F})$ è completamente regolare. ■

Corollario 3.2.7. *Sottospazi di spazi completamente regolari è ancora completamente regolare. Inoltre il prodotto cartesiano di spazi completamente regolari è ancora completamente regolare.*

Proposizione 3.2.8. *Uno spazio topologico (X, τ) è completamente regolare se e solo se la sua topologia coincide con la topologia iniziale generata da $C(X)$.*

Dimostrazione. Il teorema precedente dimostra una delle due implicazioni. Viceversa sappiamo che la topologia iniziale generata da tutte le funzioni continue da X in $\mathbb{R}$ è meno fine della topologia τ, dimostriamo che coincidono tramite la proposizione 2.3.5.

Prendiamo un qualunque $x \in X$ ed un aperto $U \in \tau$ contenente x, per la completa regolarità di X esiste $f : X \to [0,1]$ continua che vale 1 in x e 0 all'infuori di U. Ma $f \in C(X)$ ed $T_{f,\{t>0\}} = \left\{y \in X \,\middle|\, f(y) > 0\right\}$ è un aperto della topologia iniziale contenente x e contenuto in U, dunque ogni sistema fondamentale di intorni di x rispetto alla topologia τ è formato da intorni della topologia iniziale generata da $C(X)$ dunque le due topologie coincidono. ■

Combinando assieme il teorema 3.2.3 e la proposizione 3.2.8 avremo che ogni spazio topologico completamente regolare possiede almeno una struttura uniforme che ne genera la topologia. Si noti che possono esistere più di una struttura uniforme, tra loro non equivalenti, che genera la medesima topologia. In compenso la struttura uniforme ottenuta da questi risultati è la meno fine tra tutte quelle su X che rendono tutte le funzioni continue da X in $\mathbb{R}$ uniformemente continue.

Ora invece mostreremo che ogni spazio topologico uniforme è necessariamente completamente regolare, ed in virtù della proposizione 3.2.8 la sua topologia coincide con la topologia iniziale indotta da tutte le funzioni continue in $\mathbb{R}$. Si badi però che non stiamo dicendo affatto che tali funzioni siano uniformemente continue, in quanto la struttura uniforme data dal teorema 3.2.3 potrebbe non coincidere con quella data.

Prendiamo innanzitutto un entourage $U \in \mathscr{T}$ dello spazio uniforme $(X, \tau, \mathscr{T})$, posto $U_0 = U$ grazie agli assiomi 3, 4, 5 delle strutture uniformi per ogni $n \in \mathbb{N}$ esisterà $U_n \in \mathscr{T}$ simmetrico tale che

$$U_n \diamond U_n \subseteq U_{n-1}$$

in particolare $U_n \subseteq U_{n-1}$ per ogni $n \in \mathbb{N}$. Definiamo ora l'insieme

$$\mathcal{Q} = \{r \in [0,1) \mid 2^n r \in \mathbb{N}_0 \text{ per qualche } n \in \mathbb{N}\}$$

(si ricorda che $0 \in \mathbb{N}_0 \setminus \mathbb{N}$) osserviamo anche che se $2^n r \in \mathbb{N}_0$ allora $2^{n+k} r \in \mathbb{N}_0$ per ogni $k \in \mathbb{N}$, dunque per ogni $r \in \mathcal{Q}$ esisterà un unico $N_r \in \mathbb{N}_0$ tale che $2^{N_r} r \in \mathbb{N}_0$ e, se $N_r > 0$, $2^{N_r - 1} r \notin \mathbb{N}_0$.

Chiaramente $N_r = 0$ se e solo se $r = 0$, inoltre per ogni $r \in \mathcal{Q}$ diverso da 0 avremo $2^{N_r} r < 2^{N_r}$ e dunque esisteranno delle costanti $c_1(r), c_2(r), \dots, c_{N_r}(r) \in \{0,1\}$ tali che

$$2^{N_r} r = c_1(r) 2^{N_r - 1} + c_2(r) 2^{N_r - 2} + \cdots + c_{N_r - 1}(r) 2 + c_{N_r}(r)$$

con $c_{N_r}(r) = 1$ in quanto per la minimalità di N_r l'intero al primo membro deve essere dispari. Se poniamo $c_i(r) = 0$ per ogni $i > N_r$ ed $c_i(0) = 0$ per ogni $i \in \mathbb{N}$ dividendo la precedente relazione per 2^{N_r} avremo che

$$r = \sum_{i=1}^{+\infty} c_i(r) 2^{-i} \tag{3.2.1}$$

con la sommatoria a destra composta solamente da un numero finito di termini non nulli. Il ragionamento appena fatto ci garantisce che la scrittura così ottenuta nella (3.2.1) è unica (purché il numero di termini non nulli è finito).

Dimostriamo ora che

Proposizione 3.2.9. *L'insieme* $\mathcal{Q}$ *è denso in* $[0,1]$.

Dimostrazione. Chiaramente per ogni numero reale $l > 0$ esiste $n \in \mathbb{N}_0$ tale che $n \leq l < n+1$. Sia allora $l \in [0,1)$ chiaramente per ogni $n \in \mathbb{N}$ esisterà $k_n \in \mathbb{N}_0$ tale che $k_n \leq 2^n l < k_n + 1$ e dunque

$$0 \leq l - 2^{-n} k_n < 2^{-n}$$

In particolare abbiamo che $2^{-n} k_n \leq l < 1$ e dunque $2^{-n} k_n \in \mathscr{Q}$ per ogni $n \in \mathbb{N}$, poiché $2^{-n} \to 0$ al tendere di n all'infinito abbiamo che $2^{-n} k_n \to l$ e quindi l appartiene alla chiusura di $\mathcal{Q}$.

Se invece $l = 1$ allora la successione $1 - 2^{-n}$ apparterrà a $\mathcal{Q}$ e convergerà ad $1 = l$ dunque $\mathrm{Cl}(\mathcal{Q}) = [0,1]$. ■

Per ogni $l \in \mathcal{Q}$ possiamo definire l'insieme

$$\mathcal{U}(l) = c_1(l) U_1 \diamond c_2(l) U_2 \diamond c_3(l) U_3 \diamond \cdots$$

dove per ogni $i \in \{0,1\}$ e per ogni $V \in \mathscr{T}$ poniamo

$$iV = \begin{cases} V & \text{se } i = 1 \\ \Delta & \text{se } i = 0 \end{cases}$$

poiché $V \diamond \Delta = \Delta \diamond V = V$ per come abbiamo definito $\mathcal{Q}$ la scrittura precedente è consistente e la definizione di $\mathcal{U}(l)$ è ben posta. Inoltre per l'associatività dell'operazione $\diamond$ non è necessario esplicitare le parentesi, mentre avendo supposto gli U_i simmetrici possiamo tranquillamente scambiare l'ordine dei termini nella definizione di $\mathcal{U}(l)$, cosa che faremo tacitamente nei prossimi risultati.

Per ogni $x, y \in X$ definiamo la seguente funzione

$$D(x,y) = \begin{cases} 1 & \text{se } (x,y) \notin \mathcal{U}(l) \text{ per ogni } l \in \mathcal{Q} \\ \inf\{l \in \mathcal{Q} \mid (x,y) \in \mathcal{U}(l)\} & \text{altrimenti} \end{cases} \in [0,1] \tag{3.2.2}$$

che sarà detta **funzione interpolante** del sistema di entourages $\{U_n\}_{n\in\mathbb{N}}$. Abbiamo chiaramente che $D(x,x) = 0$ per ogni $x \in X$ e per la simmetria degli U_n segue anche che $D(x,y) = D(y,x)$. Per dimostrare le altre proprietà di D procediamo per gradi.

Parte 1 Dimostriamo che $\mathcal{U}(l) \diamond \mathcal{U}(r) \subseteq \mathcal{U}(l+r)$ per ogni $l, r \in \mathcal{Q}$ con $l + r < 1$.

Supponiamo inizialmente che $r = 2^{-k}$ per qualche $k \in \mathbb{N}$. Dato che $l + r < 1$ avremo di conseguenza che

$$l < 1 - 2^{-k} = \frac{1}{2}\frac{1-2^{-k}}{1-2^{-1}} = \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{k-1}} + \frac{1}{2^k}$$

e quindi $c_i(l) = 0$ per qualche $1 \le i \le k$, prendiamo allora il più grande tra tali indici che indicheremo con i semplicemente.

Se $i = k$ allora $c_j(l+r) = c_j(l) + c_j(r)$ per ogni $j \in \mathbb{N}$ e perciò $\mathcal{U}(l+r) = \mathcal{U}(l) \diamond \mathcal{U}(r)$. Invece se $i < k$ allora $c_i(r) = 0$ mentre $c_{i+1}(r) = c_{i+2}(r) = \cdots = c_k(r) = 1$ e quindi

$$\begin{aligned} &c_j(l + 2^{-k}) = c_j(l) \text{ se } j < i \text{ o } j > k \\ &c_i(l + 2^{-k}) = 1 \\ &c_j(l + 2^{-k}) = 0 \text{ se } i < j \le k \end{aligned}$$

abbiamo così per come abbiamo definito gli insiemi U_n e per l'associatività

$$\begin{aligned} &U_{i+1} \diamond U_{i+2} \diamond \cdots \diamond U_{k-2} \diamond U_{k-1} \diamond U_k \diamond U_k \\ &\qquad \subseteq U_{i+1} \diamond U_{i+2} \diamond \cdots \diamond U_{k-2} \diamond U_{k-1} \diamond U_{k-1} \\ &\qquad\quad \subseteq U_{i+1} \diamond U_{i+2} \diamond \cdots \diamond U_{k-2} \diamond U_{k-2} \subseteq \cdots \subseteq U_{i+1} \diamond U_{i+1} \subseteq U_i \end{aligned}$$

Aggiungendo gli altri termini ad entrambi i membri dell'inclusione otteniamo che $\mathcal{U}(l) \diamond \mathcal{U}\left(2^{-k}\right) = \mathcal{U}(l) \diamond U_k \subseteq \mathcal{U}\left(l + 2^{-k}\right)$.

Il caso in cui $r \in \mathcal{Q}$ è generico si ottiene scrivendo r nella forma (3.2.1) ed iterando il ragionamento di sopra rispetto a ciascun termine.

Parte 2 Se $l, r \in \mathcal{Q}$ e $r < l$ allora $\mathcal{U}(r) \subseteq \mathcal{U}(l)$.

Dato che $l - r \in \mathcal{Q}$ per quanto dimostrato nella parte precedente avremo che $\mathcal{U}(r) \subseteq \mathcal{U}(r) \diamond \mathcal{U}(l-r) \subseteq \mathcal{U}(l)$.

Parte 3 $D(x,z) \le D(x,y) + D(y,z)$ per ogni $x, y, z \in X$. Inoltre se $(x,y) \notin U$ allora $D(x,y) = 1$.

Dalla densità di $\mathcal{Q}$ per ogni $\varepsilon > 0$ esistono $l, r \in \mathcal{Q}$ tali che $D(x,y) < l < D(x,y) + \varepsilon$ e $D(y,z) < r < D(y,z) + \varepsilon$. Per la proprietà di monotonia dimostrata nel punto 2 avremo necessariamente che $(x,y) \in \mathcal{U}(l)$, $(y,z) \in \mathcal{U}(r)$ e dunque $(x,z) \in \mathcal{U}(l) \diamond \mathcal{U}(r) \subseteq \mathcal{U}(l+r)$ e quindi $D(x,z) \le l + r < D(x,y) + D(y,z) + 2\varepsilon$. Per l'arbitrarietà di ε avremo che $D(x,z) \le D(x,y) + D(y,z)$.

Infine essendo $U_i \diamond U_i \subseteq U_{i-1}$ per ogni $l \in \mathcal{Q}$ segue che $\mathcal{U}(l) \subseteq U_1 \diamond U_1 \subseteq U_0 = U$ e dunque $D(x,y) = 1$ se $(x,y) \notin U$.

Parte 4 L'applicazione $(x, y) \in X \times X \to D(x, y) \in \mathbb{R}$ è uniformemente continua se $X \times X$ è dotato della struttura uniforme prodotto (si veda il teorema 3.2.3). Inoltre per ogni $n \in \mathbb{N}$

$$U_{n+1} \subseteq \left\{(x, y) \;\middle|\; D(x, y) < 2^{-n}\right\} \subseteq U_n \tag{3.2.3}$$

Presi due entourage $U, V \in \mathscr{T}$ definiamo $P(U, V) \in \mathcal{P}(X \times X \times X \times X)$ in modo tale che

$$[(x, y), (x', y')] \in P(U, V) \Leftrightarrow (x, x') \in U \text{ e } (y, y') \in V$$

allora gli insiemi $P(U, V)$ sono una base uniforme di $X \times X$ e quindi possiamo lavorare su di essi.

Per ogni $a, b, c, d \in X$ per quanto detto nella parte precedente abbiamo che

$$D(a, b) \leq D(a, c) + D(c, d) + D(d, b)$$

e per simmetria

$$\left|D(a, b) - D(c, d)\right| \leq D(a, c) + D(b, d)$$

Adesso per ogni $r > 0$ osserviamo che

$$\left\{(x, y) \in X \times X \;\middle|\; D(x, y) < r\right\} = \bigcup_{s \in \mathbb{Q}, s < r} \mathcal{U}(s) \tag{3.2.4}$$

in particolare $U_{n+1} = \mathcal{U}\left(2^{-n-1}\right) \subseteq \left\{D(x, y) < 2^{-n}\right\}$. Sempre per monotonia abbiamo anche che $\mathcal{U}(r) \subseteq \mathcal{U}\left(2^{-n}\right) = U_n$ per ogni $r < 2^{-n}$ e perciò $\left\{D(x) < 2^{-n}\right\} \subseteq U_n$.

Per ogni $\varepsilon > 0$ esisterà dunque $i \in \mathbb{N}$ tale che per ogni $(x, y) \in U_i$ si abbia $D(x, y) < \varepsilon/2$. Posto così $V = P(U_i, U_i)$ avremo che V è un entourage di $X \times X$ e che per ogni $[(x, y), (x', y')] \in V$ si abbia

$$\left|D(x, y) - D(x', y')\right| \leq D(x, x') + D(y, y') < \varepsilon$$

e dunque D è uniformemente continua.

Osserviamo anche che $D(x, x) = 0$ ed $D(x, y) = D(y, x)$ per ogni $x, y \in X$. Siamo ora in gradi di dimostrare che ogni spazio uniforme è anche completamente regolare. Prendiamo $x \in X$ ed $A \subseteq X$ aperto contenente x, per definizione di spazio topologico uniforme esisterà un entourage U tale che $U[x] \subseteq A$.

Costruiamo la funzione interpolante D a partire da U come sopra, e definiamo $f(y) = 1 - D(y, x)$. Chiaramente $f \in C(X)$ con $f(x) = 1$ mentre per ogni $y \notin A$ avremo che $(y, x) \notin U$ e dunque $f(y) = 0$. Dunque X è uno spazio completamente regolare.

La funzione interpolante è molto importante nello studio degli spazi uniformi, in quanto possiedono delle proprietà molto utili nella dimostrazione dei prossimi risultati, in particolare la cosiddetta disuguaglianza triangolare dimostrata nella parte 3. La funzione interpolante è un caso particolare di **pseudometrica**, come afferma la seguente definizione

Definizione 3.2.10. Dato un insieme non vuoto S, l'applicazione $d : S \times S \to \mathbb{R}$ è una pseudometrica su S se e solo se soddisfa le seguenti proprietà:

(Riflessività) $d(x, x) = 0$;

(Simmetria) $d(x, y) = d(y, x)$;

(Disuguaglianza triangolare) $d(x,z) \leq d(x,y) + d(y,z)$.

Se inoltre d soddisfa la riflessività stretta, ovvero se $d(x,y) = 0$ allora $x = y$, diremo che d è una **metrica** su X.

Osservazione. Osserviamo innanzitutto che $d(x,y) \geq 0$ per ogni $x,y \in X$, si ha infatti $0 = d(x,x) \leq d(x,y) + d(y,x) = 2d(x,y)$.

Sottraendo $d(x,z)$ nella disuguaglianza triangolare avremo che $d(x,y) - d(x,z) \leq d(z,y)$, scambiando ora y con z e sfruttando l'ipotesi di simmetria $d(y,z) = d(z,y)$ avremo che

$$\left|d(x,y) - d(x,z)\right| \leq d(y,z) \tag{3.2.5}$$

disuguaglianza che useremo enormemente nel seguito.

Consideriamo adesso una famiglia di pseudometriche $\{d_i\}_{i\in I}$ definite su un insieme non vuoto S per ogni $i \in I$, $\varepsilon > 0$ gli insiemi nella forma

$$U(\varepsilon, d_i) = \left\{(x,y) \in S \times S \,\middle|\, d_i(x,y) < \varepsilon\right\}$$

generano una struttura uniforme su S di cui costituiscono una semibase uniforme. Infatti $\Delta \subseteq U(\varepsilon, d_i)$ per ogni $\varepsilon > 0$ e chiaramente $U(\varepsilon, d_i)^{-1} = U(\varepsilon, d_i)$ per la proprietà di simmetria. Infine abbiamo anche che

$$U(\varepsilon, d_i) \diamond U(\eta, d_i) \subseteq U(\varepsilon + \eta, d_i)$$

Tale struttura uniforme (e topologica) sarà detta struttura uniforme generata dalla famiglia di pseudometriche $\{d_i\}_{i\in I}$ e lo spazio $\left(S, \{d_i\}_{i\in I}\right)$ è detto spazio pseudometrico.

Proposizione 3.2.11. *La topologia di uno spazio pseudometrico S generato dalle pseudometriche d_i coincide con la topologia iniziale generata dalle funzioni $f_{i,x}(y) = d_i(x,y)$. Inoltre la struttura uniforme generata dalle $f_{i,x}$ è meno fine della struttura uniforme generata dalle pseudometriche di S.*

Dimostrazione. Sappiamo già che gli insiemi

$$F_{\varepsilon,i,u} = \left\{(x,y) \in S \times S \,\middle|\, \left|d_i(x,u) - d_i(y,u)\right| < \varepsilon\right\}$$

formano una semibase uniforme di S rispetto alla struttura uniforme generata dalle $f_{i,u}$ al variare di $\varepsilon > 0$, $i \in I$ ed $u \in S$. Grazie alla (3.2.5) abbiamo la seguente inclusione

$$U(\varepsilon, d_i) \subseteq F_{\varepsilon,i,u}$$

e così gli insiemi $F_{\varepsilon,i,u}$ appartengono anche alla struttura uniforme generata dalle pseudometriche. Abbiamo così dimostrato che la struttura uniforme generata dalle $f_{i,x}$ è meno fine della struttura uniforme generata dalle pseudometriche e così vale anche per le rispettive topologie.

Fissati ora $x \in S$, $\varepsilon > 0$ ed $i \in I$ abbiamo

$$U(\varepsilon, d_i)[x] = \left\{y \in S \,\middle|\, d_i(y,x) < \varepsilon\right\} = F_{\varepsilon,i,x}[x]$$

e quindi ngli intorni di X nella topologia indotta dalle pseudometriche sono intorni anche nella topologia iniziale. Per quanto detto prima le due topologie coincidono. ■

Ora che abbiamo introdotto le caratteristiche principali delle pseudometriche possiamo dimostrare il seguente teorema di equivalenza:

Teorema 3.2.12. *Ogni spazio uniforme* $(X, \tau, \mathcal{T})$ *è generato da una qualche famiglia di pseudometriche* $\{d_i\}_{i\in I}$*. Inoltre* X *è di Hausdorff se e solo se per ogni* $x, y \in X$ *con* $x \neq y$ *esiste* $i \in I$ *tale che* $d_i(x, y) \neq 0$.

Dimostrazione. Per ogni $U \in \mathcal{T}$ possiamo definire la funzione di interpolazione D_U rispetto ad U. Sia allora $\mathcal{S}$ la struttura uniforme su X generata dalla semibase $\{(x, y) \mid D_U(x, y) < \varepsilon\}$ al variare di $U \in \mathcal{T}$ e di $0 < \varepsilon < 1$. Per la (3.2.3) segue immediatamente che $\mathcal{S} \subseteq \mathcal{T}$, viceversa per ogni $U \in \mathcal{T}$ abbiamo che $\{(x, y) \mid D_U(x, y) < 1\} \subseteq U$ e dunque $\mathcal{T} \subseteq \mathcal{S}$. Le due strutture uniformi coincidono allora.

Lo spazio X è di Hausdorff se e solo se l'intersezione di tutti gli insiemi $U(\varepsilon, d_i)$ è esattamente Δ ovvero se e solo se per ogni x, y con $x \neq y$ esiste i tale che $d_i(x, y) > 0$. ■

3.3 Cauchy nets

Definizione 3.3.1. Dato uno spazio uniforme X una **Cauchy net** su X è una net x_a definita su un insieme diretto A tale che per ogni entourage U di X esiste $a \in A$ per cui $(x_b, x_c) \in U$ per ogni $b, c \geq a$.

Se $A = \mathbb{N}$ diremo che x_n è una **successione di Cauchy**.

Osservazione. Al posto di considerare ogni entourage U ci si può limitare a prendere solo quelli appartenenti a qualche semibase uniforme.

Notiamo che se la struttura uniforme di X è indotta da una famiglia di pseudometriche $\{d_i\}_{i\in I}$ allora la definizione di Cauchy net assume una forma molto simile a quella classica su $\mathbb{R}$. Una net x_a è di Cauchy se e solo se per ogni $\varepsilon > 0$ e per ogni $i \in I$ esiste $a \in A$ tale che $d_i(x_b, x_c) < \varepsilon$ per ogni $b, c \geq a$.

Osserviamo che tutte le net convergenti sono Cauchy nets, se per esempio prendiamo $x_a \to x$ ed un entourage $U \in \mathcal{T}$ esisterà chiaramente un altro entourage V simmetrico tale che $V \diamond V \subseteq U$ e quindi esiste $a \in A$ tale che $x_b \in V[x]$ per ogni $b \geq a$ per definizione di convergenza.

Ma allora presi $b, c \geq a$ avremo che $(x_b, x), (x, x_c) \in V$ e dunque $(x_b, x_c) \in V \diamond V \subseteq U$ e perciò x_a è una Cauchy net. In generale però non vale il viceversa, ovvero possono esistere Cauchy nets che non convergono ad alcun punto dello spazio. In compenso vale il seguente risultato:

Proposizione 3.3.2. *Se una Cauchy net* x_a *possiede una subnet convergente a qualche* $x \in X$ *allora* x_a *converge ad* x.

Dimostrazione. Sia x_{a_i} la subnet definita su I che converge a qualche $x \in X$ e prendiamo un qualunque entourage U di X, esisterà chiaramente $V \in \mathcal{T}$ simmetrico tale che $V \diamond V \subseteq U$ ed esiste $a \in A$ tale che $(x_b, x_c) \in V$ per ogni $b, c \geq a$.

D'altronde dalla definizione di insieme diretto e di subnet possiamo sempre trovare $i \in I$ tale che $a \leq k_i$ ed $\left(x_{k_i}, x\right) \in V$, dunque $\left(x_b, x_{k_i}\right), \left(x_{k_i}, x\right) \in V$ ed $(x_b, x) \in U$ per ogni $b \geq a$. Per l'arbitrarietà di U segue che $x_a \to x$ come volevasi dimostrare. ■

Dato che non in tutti gli spazi uniformi le Cauchy nets convergono risulta conveniente introdurre le seguenti definizioni:

Definizione 3.3.3. Uno spazio uniforme X è detto **uniformemente** $\mathcal{N}_1$ se e solo se possiede una base uniforme numerabile.

Diremo inoltre che X è **net-completo** se e solo se ogni Cauchy net converge in X, mentre diremo che è **(sequenzialmente) completo** se e solo se ogni successione di Cauchy converge.

Chiaramente ogni spazio net-completo è anche sequenzialmente completo mentre il viceversa in generale non vale, nemmeno se lo spazio soddisfa l'assioma $\mathcal{N}_1$. Per questa ragione è necessario rafforzarlo con l'uniforme $\mathcal{N}_1$ che è più forte rispetto al semplice assioma $\mathcal{N}_1$: data una base uniforme numerabile $\{U_n\}_{n\in\mathbb{N}}$ per ogni $x \in X$ la famiglia

$$\{U_n[x] \mid n \in \mathbb{N}\}$$

è un sistema fondamentale di intorni per x. Esistono però spazi uniformi $\mathcal{N}_1$ che non sono uniformemente $\mathcal{N}_1$.

Proposizione 3.3.4. *Dato uno spazio uniforme X uniformemente $\mathcal{N}_1$ allora lo spazio X è net-completo se e solo se è sequenzialmente completo.*

Dimostrazione. Supponiamo che X sia uniformemente $\mathcal{N}_1$ e sequenzialmente completo, e prendiamo una qualunque Cauchy net $x_a \in X$ definita su un insieme diretto A. Prendiamo una qualunque base uniforme numerabile $\{U_n\}_{n\in\mathbb{N}}$ per ogni $n \in \mathbb{N}$ esiste $a_n \in A$ tale che $a_{n+1} \geq a_n$ ed $(x_b, x_c) \in U_n$ per ogni $b, c \geq a_n$. Dunque x_{a_n} è una successione di Cauchy in X che per ipotesi convergerà a qualche $x \in X$.

Ora per definizione di base uniforme per ogni $n \in \mathbb{N}$ esiste $m \in \mathbb{N}$ tale che $U_m \diamond U_m \subseteq U_n$ e di conseguenza esisterà $N \geq m$ tale che $\left(x_{a_N}, x\right) \in U_m$, essendo $a_N \geq a_m$ avremo anche che $\left(x_{a_N}, x_b\right) \in U_m$ per ogni $b \geq a_m$ e dunque

$$(x_b, x) \in U_m \diamond U_m \subseteq U_n$$

ovvero $x_a \to x$ e X è net-completo. ∎

Proposizione 3.3.5. *Se $f : X \to Y$ è un'applicazione uniformemente continua tra spazi uniformi e $x_a \in X$ è una Cauchy net allora $f(x_a)$ è una Cauchy net in Y.*

Proposizione 3.3.6. *Se X è uno spazio uniforme net-completo e $y \subseteq X$ è chiuso e dotato della struttura uniforme indotta da X allora Y è ancora uno spazio uniforme net-completo.*

Dimostrazione. Prendiamo una generica Cauchy net $x_a \in Y$, per la net-completezza di X esisterà un $x \in X$ tale che $x_a \to x$. Poiché Y è chiuso allora $x \in Y$ e perciò Y è net-completo. ∎

Possiamo ridefinire la completezza di uno spazio uniforme anche in termini di sottoinsiemi chiusi, come mostra il seguente risultato:

Teorema 3.3.7. *Uno spazio uniforme $(X, \mathcal{T})$ è net-completo se e solo se presi un qualunque insieme diretto A ed una qualunque net $\{C_a\}_{a\in A}$ di sottoinsiemi di X tale che*

- *$\emptyset \neq C_b \subseteq C_a$ per ogni $b \geq a$ in A;*
- *per ogni $U \in \mathcal{T}$ esiste $a \in A$ tale che $C_a \times C_a \subseteq U$.*

Allora $C = \bigcap_{a\in A} \mathrm{Cl}(C_a) \neq \emptyset$.

Invece X *è sequenzialmente completo se e solo se vale l'asserto precedente con* $A = \mathbb{N}$.

Dimostrazione. Prendiamo $(X, \mathscr{T})$ net-completo e $\{C_a\}_{a\in A}$ che soddisfano le ipotesi sopra. Scegliamo $x_a \in C_a$ elementi generici, poiché C_a è decrescente si verifica immediatamente che x_a è una Cauchy net, dunque per la completezza di X esisterà $x \in X$ tale che $x_a \to x$. Dato che $x_b \in C_a$ per ogni $b \geq a$ l'estratta definita su $\{b \in A \mid b \geq a\} \neq \emptyset$ è definita sull'insieme chiuso $\mathrm{Cl}(C_a)$ e convergerà ancora ad x, dunque $x \in \mathrm{Cl}(C_a)$ per ogni $a \in A$ e così $x \in C \neq \emptyset$.

Dimostriamo adesso l'implicazione opposta. Considerata una qualunque Cauchy net $x_a \in X$ su A e poniamo $C_a = \{x_b \mid b \geq a\}$, dato che x_a è una Cauchy net per ogni $U \in \mathscr{T}$ esisterà $a \in A$ tale che $C_a \times C_a \subseteq U$. Per ipotesi esisterà perciò $x \in X$ tale che $x \in \mathrm{Cl}(C_a)$ per ogni $a \in A$, in altre parole $U[x] \cap C_a \neq \emptyset$ per ogni $a \in A$ e per ogni $U \in \mathscr{T}$.

Definiamo adesso

$$I = \{(U, a) \in \mathscr{T} \times A \mid x_a \in U[x]\}$$

con l'ordinamento $\leq$ definito in modo tale che $(U, a) \leq (V, b)$ se e solo se $V \subseteq U$ ed $a \leq b$, verifichiamo innanzitutto che l'insieme è diretto. Presi due suoi elementi $(U, a), (V, b)$ esisterà $c \in A$ tale che $a, b \leq c$, per quanto detto sopra $C_c \cap (U \cap V)[x] \neq \emptyset$ e dunque esisterà $d \geq c$ tale che $x_d \in (U \cap V)[x]$ e perciò $(U \cap V, d) \in I$ con $(U, a), (V, b) \leq (U \cap V, d)$. In particolare per ogni U esiste a tale che $(U, a) \in I$.

Posto dunque $f : (U, a) \in I \to a \in A$ avremo che $x_{f(U,a)} \in X$ è una subnet di x_a con $x_{f(U,a)} \in U[x]$. Per l'arbitrarietà di U non è difficile verificare che tale subnet converge esattamente ad x, e quindi per la proposizione 3.3.2 tutta x_a deve convergere ad x. ■

In generale una funzione continua definita solo su un sottospazio denso di uno spazio uniforme potrebbe non essere estendibile sull'intero spazio, e controesempi si possono trovare anche su $\mathbb{R}$. Per poter estendere una tale funzione abbiamo bisogno innanzitutto che la funzione sia uniformemente continua, poi abbiamo bisogno anche che il codominio sia completo in quanto uno spazio uniforme non completo potrebbe presentare dei "buchi" in cui non è possibile definire la nostra funzione.

Lemma 3.3.8 (Estensione per uniforme continuità)**.** *Siano* $(X, \mathscr{T}), (Y, \mathscr{S})$ *spazi uniformi con* Y *net-completo di Hausdorff e* $A \subseteq X$ *sottoinsieme denso. Presa una qualunque funzione* $f : A \to Y$ *uniformemente continua su* A *esisterà un'unica funzione* $g : X \to Y$ *uniformemente continua su* X *tale che* $f(x) = g(x)$ *per ogni* $x \in A$. *Se* X *è* $\mathscr{N}_1$ *allora* Y *può essere (sequenzialmente) completo al posto di net-completo.*

Dimostrazione. L'unicità è immediata da verificare, infatti presa una qualunque altra funzione continua g' da X in Y che coincide con f su A per ogni $x \in S$ esisterà una net $x_b \in A$ convergente a x. Dunque per il teorema di unicità del limite in uno spazio di Hausdorff

$$g(x) = \lim_{b\to+\infty} g(x_b) = \lim_{b\to+\infty} g'(x_b) = g'(x)$$

dobbiamo quindi dimostrare solo l'esistenza. Sia $\mathscr{T}$ la struttura uniforme di X, per ogni $U \in \mathscr{T}$ e per ogni $x \in X$ abbiamo $A \cap U[x] \neq \emptyset$ e perciò possiamo porre

$$B_U = f(A \cap U[x]) \neq \emptyset$$

Dotiamo $\mathscr{T}$ della relazione d'ordine $U \leq V \Leftrightarrow V \subseteq U$ in modo da renderlo un insieme diretto, per l'uniforme continuità di f per ogni $W \in \mathscr{S}$ esiste $U \in \mathscr{T}$ tale che $B_U \times B_U \subseteq V$, perciò possiamo applicare il teorema 3.3.7 e la completezza di Y per stabilire che $\bigcap_{U\in\mathscr{T}} \mathrm{Cl}(B_U) \neq \emptyset$. Poiché Y è uno spazio di Hausdorff, e quindi $\bigcap_{Y\in\mathscr{S}} Y = \Delta$, l'intersezione può contenere un solo punto che indicheremo con $g(x)$, quindi g è una funzione ben definita da X in Y.

Ora se $x \in A$ allora $f(x) \in B_U$ per ogni $U \in \mathscr{T}$, dunque per quanto detto prima deve essere necessariamente $f(x) = g(x)$. Dimostriamo adesso che anche g è uniformemente continua. Per ogni $V \in \mathscr{S}$ esiste $U \in \mathscr{T}$ simmetrico tale che

$$(U \diamond U \diamond U) \cap (A \times A) \subseteq (f \times f)^{-1}(V')$$

con $V' \in \mathscr{S}$ simmetrico in modo tale che $V' \diamond V' \diamond V' \subseteq V$. Siano ora $(x, y) \in U$ generici su X, per definizione abbiamo

$$g(x) \in V\big[f(A \cap U[x])\big]$$

e quindi esisterà $\tilde{x} \in A$ tale che $(x, \tilde{x}) \in U$ ed $(g(x), f(\tilde{x})) \in V'$, analogamente si può trovare $\tilde{y}$ per y.

Poiché $(x, y) \in U$ ed U è simmetrico avremo $(\tilde{x}, \tilde{y}) \in U \diamond U \diamond U$ e quindi $(f(\tilde{x}), f(\tilde{y})) \in V'$ da cui discende che $(g(x), g(y)) \in V' \diamond V' \diamond V' \subseteq V$. In altre parole g è uniformemente continua su tutto X.

Quando X è uniformemente $\mathscr{N}_1$ basta limitarsi a considerare gli elementi di una base uniforme numerabile ed applicare la forma sequenziale del teorema 3.3.7. ■

Definizione 3.3.9. Diremo che uno spazio uniforme $(X, \mathscr{T})$ è **totalmente limitato** se e solo se per ogni $U \in \mathscr{T}$ esiste un sottoinsieme finito $F \subseteq X$ tale che $U[F] = X$, in altre parole esistono $x_1, x_2, \dots, x_n$ tali che per ogni $x \in X$ esiste $1 \leq i \leq n$ tale che $(x, x_i) \in U$.

Osservazione. Per uno spazio uniforme non è possibile definire il concetto di limitatezza, il quale è proprio degli spazi metrici.

Abbiamo già definito all'inizio la compattezza e la compattezza sequenziale per spazi topologici generici, che quindi possono essere applicate anche ad uno spazio uniforme. Lo scopo di questa sezione è indagare la relazione che intercorre tra gli spazi totalmente limitati e le proprietà di compattezza. Introduciamo ora un nuovo concetto di compattezza il quale, a differenza dei precedenti, prende in considerazione anche la struttura uniforme dello spazio.

Definizione 3.3.10. Sia X spazio metrico, allora diremo che

- X è **relativamente compatto** se ogni net x_a di X possiede una subnet di Cauchy;
- X è **relativamente net compatto** se ogni net in X possiede una subnet di Cauchy;
- X è **relativamente sequenzialmente compatto** se ogni successione in X possiede un'estratta di Cauchy.

Diamo ora alcune proprietà degli spazi totalmente limitati

Proposizione 3.3.11. *Sottoinsiemi non vuoti di uno spazio totalmente limitato sono totalmente limitati.*

Dimostrazione. Sia X totalmente limitato e $Y \subseteq X$ non vuoto con la struttura uniforme indotta, allora per ogni U entourage di X esiste $F \subseteq X$ finito tale che $U[F] = X$. Fissiamo allora

un entourage generico V e scegliamo U simmetrico in modo tale che $U \diamond U \subseteq V$, posto $F = \{x_1, x_2, \dots, x_n\}$ ed
$$I = \{1 \leq i \leq n \mid U[x_i] \cap Y \neq \emptyset\}$$
per ogni $i \in I$ esisterà $y_i \in Y$ tale che $(y_i, x_i) \in U$.

Prendiamo ora un generico $y \in Y$, dato che $U[F] \supseteq Y$ esisterà $i \in I$ tale che $(y, x_i) \in U$ e dunque $(y, y_i) \in U \diamond U \subseteq V$. Quindi $Y \subseteq \bigcup_{i \in I} V[y_i]$ e la tesi segue immediatamente in quanto I è finito. ■

Teorema 3.3.12. *Sia $(X, \mathscr{T})$ spazio uniforme, allora X è relativamente compatto se e solo se è totalmente limitato. Inoltre ogni spazio relativamente net compatto è anche relativamente compatto.*

Dimostrazione. Supponiamo inizialmente che X non sia totalmente limitato, allora esiste $U \in \mathscr{T}$ simmetrico tale che per ogni sottoinsieme finito $F \subseteq X$ si abbia $X \setminus U[F] \neq \emptyset$. Per ricorrenza possiamo determinare una successione $x_n \in X$ tale che $x_n \notin U[x_i]$ per ogni $i < n$, dunque $(x_i, x_j) \notin U$ per ogni $i \neq j$. Questo implica che x_n non possiede né estratte né subnet di Cauchy, quindi se X è relativamente compatto o relativamente sequenzialmente compatto allora deve essere totalmente limitato.

Viceversa supponiamo che X sia totalmente limitato ed $x_a \in X$ una qualunque net, definiamo la famiglia $\mathscr{L}$ di tutti i sottoinsiemi H di X per cui esiste una subnet di x_a contenuta interamente in H. Chiaramente $X \in \mathscr{L}$ e per il lemma di Zorn esiste una sottofamiglia $\mathscr{F} \subseteq L$ massimale rispetto all'inclusione tale che per ogni $H, K \in \mathscr{F}$ si abbia $H \cap K \in \mathscr{F}$. Osserviamo che se $H \in \mathscr{F}$ e $K \supseteq H$ allora anche $K \in \mathscr{F}$ in quanto la famiglia
$$\{K \subseteq X \mid \exists H \in \mathscr{F} \text{ tale che } H \subseteq K\}$$
è contenuta in $\mathscr{L}$, contiene $\mathscr{F}$ ed è stabile rispetto all'intersezione finita, quindi per massimalità deve coincidere con $\mathscr{F}$. Il lemma 2.2.8 garantisce così l'esistenza di una subnet x_{a_i} definita su qualche insieme diretto I tale che per ogni $H \in \mathscr{F}$ esiste $j \in I$ tale che $x_{a_i} \in H$ per ogni $i \geq j$.

Dimostriamo ora le seguenti proprietà delle famiglie $\mathscr{L}$ ed $\mathscr{F}$.

- Se $H_1 \cup H_2 \cup \cdots \cup H_n \in \mathscr{L}$ allora esiste i tale che $H_i \in \mathscr{L}$. Per ogni i intero tra 1 ed n consideriamo gli insiemi
$$I_i = \{a \in A \mid x_a \in H_i\}$$
per ipotesi sappiamo che per ogni $a \in A$ esistono un indice i ed un $b \in I_i$ tali che $b \geq a$. Per assurdo supponiamo anche che nessuno di essi è un sottoinsieme cofinale di A, ovvero per ogni i esistono $a_i \in A$ per cui non esiste alcun $b_i \in I_i$ tale che $b_i \geq a_i$. Dato però che gli i sono in numero finito esiste $a \in A$ tale che $a \geq a_i$ per ogni i e dunque esistono $1 \leq j \leq n$ intero ed $b_j \in I_j$ tali che $b_j \geq a \geq a_j$ assurdo. Quindi esiste j tale che la net $\{x_b\}_{b \in I_j}$ è un'estratta di x_a e dunque $H_j \in \mathscr{L}$.

- Dati $H_1, H_2, \dots, H_n \in \mathscr{L}$ tali che $\bigcup_{i=1}^n H_i \in \mathscr{F}$ allora esiste i tale che $H_i \in \mathscr{F}$. Per ogni i possiamo considerare la famiglia
$$\mathscr{G}_i = \mathscr{F} \cup \{B \cap H_i \mid B \in \mathscr{F}\}$$

che contiene $\mathscr{F}$ ed è stabile rispetto all'intersezione finita. Se dimostriamo che è contenuta in $\mathscr{L}$ per qualche i per massimalità deve coincidere con $\mathscr{F}$ e perciò conterrà anche A_i. Per assurdo esistono $K_i \in \mathscr{F}$ tali che $H_i \cap K_i \notin \mathscr{L}$ ma allora

$$\bigcup_{i=1}^{n} H_i \cap \bigcap_{i=1}^{n} K_i \subseteq \bigcup_{i=1}^{n} (H_i \cap K_i)$$

per ipotesi l'insieme a sinistra apparterrà a $\mathscr{F}$ mentre l'insieme a destra non può appartenere ad $\mathscr{L}$ per quanto detto nel punto precedente, generando così un assurdo.

- Per ogni $H \subseteq X$ si ha $H \in \mathscr{F}$ oppure $X \setminus H \in \mathscr{F}$. Innanzitutto se $H, X \setminus H \in \mathscr{L}$ la tesi segue immediatamente dal punto precedente, quindi possiamo supporre che $H \notin \mathscr{L}$ e quindi esiste un $a \in A$ tale che se $x_b \in H$ allora $b \ngeq a$. In altre parole per ogni $b \geq a$ si ha $x_b \in X \setminus H$ e dunque $K \cap (X \setminus H) \in \mathscr{L}$ per ogni $K \in \mathscr{L}$. Per massimalità si ha allora $X \setminus H \in \mathscr{F}$ dato che possiamo costruire una famiglia contenente $\mathscr{F}$ e contenuta in $\mathscr{L}$ chiusa rispetto alle intersezioni finite.

Consideriamo adesso un qualunque entourage U di X, per la totale limitatezza esiste $F_U \subseteq X$ finito tale che $X = U[F_U]$. Per il primo punto l'insieme G_U dei punti $x \in F_U$ tali che $U[x] \in \mathscr{L}$ non è vuoto. Sempre per il primo punto abbiamo che $\bigcup_{x \in F_U \setminus G_U} U[x] \notin \mathscr{L}$ e dunque per il terzo punto si ha

$$X \setminus \bigcup_{x \in F_U \setminus G_U} U[x] \subseteq \bigcup_{y \in G_U} U[x] \in \mathscr{F}$$

che combinato con il secondo punto ci garantisce l'esistenza di $x_U \in X$ tale che $U[x_U] \in \mathscr{F}$.

Ricapitolando tutto quanto detto finora sappiamo che esiste una subnet x_{a_i} su I tale che per ogni U entourage esistono $x_U \in X$ ed $j \in I$ tali che $x_{a_i} \in U[x_U]$ per ogni $i \geq j$. D'altronde esisterà un entourage simmetrico V tale che $V \diamond V \subseteq U$ per cui $\left(x_{a_i}, x_V\right) \in V$ per ogni $i \geq k$, e così se anche $j \geq k$

$$\left(x_{a_i}, x_{a_j}\right) \in V \diamond V \subseteq U$$

ovvero x_{a_i} è di Cauchy. ■

Da questo risultato otteniamo immediatamente il seguente criterio di compattezza:

Teorema 3.3.13. *In uno spazio uniforme X sono equivalenti:*

- *X è compatto;*
- *ogni net in X possiede una subnet convergente;*
- *X è totalmente limitato e net-completo.*

3.4 Completamento

Precedentemente abbiamo introdotto la definizione di spazio uniforme net-completo, ovvero di spazi uniformi nei quali le Cauchy net convergono sempre. Dato che non tutti gli spazi uniformi sono completi vogliamo sapere se è sempre possibile "completare" uno spazio uniforme aggiungendo nuovi punti in modo tale da risultare completo.

Per esempio l'insieme dei numeri razionali $\mathbb{Q}$ con la struttura uniforme standard è uno spazio uniforme non completo in quanto la successione $x_n \in \mathbb{Q}$ definita per ricorrenza come $x_1 = 1$ ed

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)$$

è di Cauchy ma non converge in $\mathbb{Q}$. Invece lo spazio $\mathbb{R}$ è chiaramente completo, e in particolare la successione x_n convergerà a $\sqrt{2}$. Anzi, ogni elemento di $\mathbb{R}$ è il limite di qualche successione di Cauchy definita in $\mathbb{Q}$, per questo motivo possiamo vedere l'insieme dei numeri reali come il completamento di quello dei razionali.

Questo ragionamento può essere generalizzato ad ogni spazio uniforme: vale infatti il seguente risultato di completamento

Teorema 3.4.1. *Dato uno spazio uniforme di Hausdorff $(X, \mathscr{F})$ esistono uno spazio di Hausdorff uniforme net-completo $(Y', \mathscr{S})$ ed una funzione iniettiva uniformemente continua $f : X \to Y$ tale che $f(X)$ è denso in Y ed f^{-1} è uniformemente continua da $f(X)$ in X. Inoltre se Z è un altro spazio uniforme di Hausdorff net-completo che soddisfa le medesime condizioni di Y allora esiste un omeomorfismo uniforme tra Y e Z.*

Per dimostrare il teorema dobbiamo procedere per gradi, dimostrando prima l'unicità del completamento. Siano $(Y_1, \mathscr{S}_1), (Y_2, \mathscr{S}_2)$ spazi uniformi net-completi per cui esistono delle funzioni iniettive uniformemente continue $f : X \to Y_1$, $g : X \to Y_2$ come nell'asserto. Poiché f e g sono iniettive possiamo considerare la composizione $g \circ f^{-1}$ che va da $R_1 = f(X)$ in $R_2 = g(X)$ la quale è ancora un omeomorfismo uniforme tra spazi uniformi.

Per il lemma 3.3.8 possiamo allora estendere $g \circ f^{-1}$ su tutto Y_1 in maniera univoca ottenendo così una funzione $F : Y_1 \to Y_2$ sia uniformemente continua. Possiamo d'altronde estendere anche la funzione $f \circ g^{-1}$ ad una nuova funzione $G : Y_2 \to Y_1$ uniformemente continua, notiamo anche che la composizione $F \circ G$ è l'identità su R_2 mentre $G \circ F$ è l'identità su R_1. Dato che entrambe le composizioni sono uniformemente continue per l'unicità dell'estenzione avremo che $F \circ G = \mathrm{id}_{Y_2}$ e $G \circ F = \mathrm{id}_{Y_1}$ ovvero $G = F^{-1}$ e quindi F è un omeomorfismo uniforme.

Passiamo adesso a costruire effettivamente lo spazio uniforme completo Y a partire da $(X, \mathscr{T})$. Innanzitutto indichiamo con R l'insieme di tutte le Cauchy net in X e quindi definiamo la relazione $\sim$ su R tale che per ogni coppia di net $x_a, y_i \in R$

$$x_a \sim y_i \Leftrightarrow \forall U \in \mathscr{T}\ \exists a_U \in A, i_U \in I \text{ tali che } \left(x_a, y_i\right) \in U \text{ per ogni } a \geq a_U, i \geq i_U$$

Notiamo che $\sim$ è simmetrica e transitiva per la definizione di struttura uniforme, metre la riflessività discende dal fatto che gli elementi di R sono Cauchy net, quindi $\sim$ è una relazione di equivalenza e poniamo $Y = R/\sim$. Vogliamo ora definire una struttura uniforme su X. Innanzitutto per ogni $U \in \mathscr{T}$ definiamo l'insieme $\mathscr{H}(U) \subseteq R \times R$ in modo tale che

$$(x, y) \in \mathscr{H}(U) \Leftrightarrow \forall V \in \mathscr{T}\, \exists a_V \in A, i_V \in I \text{ tali che } \left(x_a, y_i\right) \in V \diamond U \diamond V^{-1} \text{ per ogni } a \geq a_V, i \geq i_V$$

Una definizione così controintuitiva è necessaria per poter essere compatibile con la relazione $\sim$: siano infatti x_a, x'_b, y_c, y'_d con $a \in A, b \in B, c \in C, d \in D$ diretti con $x \sim x'$, $y \sim y'$ ed $(x, y) \in \mathscr{H}(U)$, vogliamo dimostrare che $(x', y') \in \mathscr{H}(U)$. Sia allora $V \in \mathscr{T}$ che possiamo supporre simmetrico per comodità, esiste W entourage tale che $W \diamond W \subseteq V$ e quindi esisteranno a_W, b_W, c_W, d_W nei rispettivi spazi tali che $\left(x_a, x'_b\right), \left(y_c, y'_d\right) \in W$ quando $a \geq a_W$ eccetera.

Senza perdere in generalità possiamo anche supporre che $(x_a, y_c) \in W \diamond U \diamond W$ sugli stessi indici in quanto A e C sono insiemi diretti, e di conseguenza

$$(x'_b, y'_d) \in W \diamond W \diamond U \diamond W \diamond W \subseteq V \diamond U \diamond V$$

come volevasi dimostrare. Possiamo quindi porre

$$\mathcal{K}(U) = \frac{\mathcal{H}(U)}{\sim} \in Y$$

Per come abbiamo definito la relazione di equivalenza $\sim$ per ogni Cauchy net x_a su X abbiamo $(x, x) \in \mathcal{H}(U)$ e inoltre

$$U \subseteq V \Rightarrow \mathcal{H}(U) \subseteq \mathcal{H}(V)$$
$$\mathcal{H}(U \cap V) = \mathcal{H}(U) \cap \mathcal{H}(V)$$
$$\mathcal{H}(U)^{-1} = \mathcal{H}\left(U^{-1}\right)$$
$$\mathcal{H}(U \diamond V) = \mathcal{H}(U) \diamond \mathcal{H}(V)$$

relazioni che il lettore può dimostrare da solo. Quindi gli insiemi nella forma $\mathcal{K}(U)$ inducono una struttura uniforme su Y che indicheremo con $\mathcal{S}$. Non è nemmeno difficile da dimostrare che Y è automaticamente uno spazio di Hausdorff (anche nel caso in cui X non lo fosse).

Consideriamo adesso la funzione $f : X \to Y$ che ad ogni $x \in X$ associa la classe di equivalenza di una qualunque net che vale costantemente x. Chiaramente questa funzione è ben definita ed è iniettiva in quanto X è uno spazio di Hausdorff, è inoltre uniformemente continua in quanto

$$[f(x), f(y)] \in \mathcal{K}(U) \Leftrightarrow (x, y) \in U$$

Dobbiamo solo verificare che $f(X)$ è denso in Y. Sia allora $y_a \in R$ una Cauchy net in X definita su un insieme diretto A, per ogni $a \in A$ definiamo

$$[t^a]_\sim = f(y_a) \in Y$$

con $t^a \in R$ la successione di Cauchy definita su $\{1\}$ che vale y_a. Mostriamo che $[t^a]_\sim$ converge ad $[y]_\sim$ in Y. Per ogni $U \in \mathcal{T}$ per definizione di Cauchy net esiste $a \in A$ tale che $(y_b, y_c) \in U$ per ogni $b, c \geq a$. Ma dato che $(y_b, y_c) = \left[\left(t^b\right)_1, y_c\right]$ avremo che $\left(\left[t^b\right]_\sim, [y]_\sim\right) \in \mathcal{K}(U)$ per ogni $b \geq a$ e quindi vi è convergenza. Abbiamo così dimostrato che $\mathrm{Cl}\left(f(X)\right) = Y$.

La completezza di Y discende dal seguente risultato:

Proposizione 3.4.2. *Dato uno spazio uniforme $(X, \mathcal{T})$ e sia $Y \subseteq X$ denso in X, se ogni Cauchy net in Y ammette limite in X allora X è completo.*

Dimostrazione. Prendiamo una qualunque Cauchy net x_a su X, per ogni $U \in \mathcal{T}$ esiste $y_{a,U} \in Y$ tale che $y_{a,U} \in U[x_a]$ per ogni $a \in A$. L'insieme $A \times \mathcal{T}$ è diretto per quanto detto più volte precedentemente, inoltre per ogni $U \in \mathcal{T}$ esiste un entourage simmetrico V tale che $V \diamond V \diamond V \subseteq U$ ed esiste $a \in A$ tale che $(x_b, x_c) \in V$ per ogni $b, c \geq a$.

Se quindi $(b, V_1), (c, V_2) \geq (a, V)$ allora avremo che

$$\left(y_{b,V_1}, x_b\right) \in V_1 \subseteq V$$
$$\left(y_{c,V_2}, x_c\right) \in V_2 \subseteq V$$
$$\left(x_b, x_c\right) \in V$$

e perciò $(y_{b,V_1}, y_{c,V_2}) \in U$ ovvero $y_{a,U}$ è una Cauchy net. Per ipotesi $y_{a,U} \to x \in X$ facciamo vedere che anche $x_a \to x$.

Sia U entourage, esiste $V \in \mathcal{T}$ simmetrico tale che $V \diamond V \subseteq U$ e di conseguenza esistono $a \in A$, $W \in \mathcal{T}$ tali che $y_{b,W_1} \in V[x]$ per ogni $b \geq a$, $W_1 \subseteq W$. Ma essendo $(x_b, y_{b,W\cap V}) \in V$ per ogni $b \geq a$ avremo anche che $(x_a, x) \in V \diamond V \subseteq U$ e perciò $x_a \to x$. ■

3.5 Spazi metrici

Torniamo adesso allo studio delle proprietà delle pseudometriche e delle metriche definite su un particolare insieme S non vuoto.

Se $M, N \subseteq S$ allora poniamo

$$\operatorname{dist}(M, N; d) = \inf\left\{d(x,y) \;\middle|\; x \in M \wedge y \in N\right\}$$

che indicheremo anche con $\operatorname{dist}(M,N)$ se non vi sono ambiguità. Inoltre se $x \in S$ scriveremo direttamente $\operatorname{dist}(x, M) = \operatorname{dist}(\{x\}, M)$.

Ancora dato un sottoinsieme non vuoto T di S definiamo il **diametro** di T la quantità

$$\operatorname{diam} T = \sup\left\{d(x,y) \;\middle|\; x, y \in T\right\}$$

Proposizione 3.5.1. *Se $M, N \subseteq S$ allora*

$$\operatorname{dist}(M,N) = \inf\left\{\operatorname{dist}(x,N) \;\middle|\; x \in M\right\}$$

Dimostrazione. Chiaramente $\operatorname{dist}(M,N) \leq \operatorname{dist}(x,N)$ per ogni $x \in M$ quindi $\operatorname{dist}(M,N) \leq \inf\left\{\operatorname{dist}(x,N) \;\middle|\; x \in M\right\}$.

D'altronde per ogni $m \in M$, $n \in N$ vale $\operatorname{dist}(m,N) \leq d(m,n)$ e di conseguenza avremo che $\inf\left\{\operatorname{dist}(x,N) \;\middle|\; x \in M\right\} \leq d(m,n)$ e dunque $\inf\left\{\operatorname{dist}(x,N) \;\middle|\; x \in M\right\} \leq \operatorname{dist}(M,N)$ concludendo la dimostrazione. ■

Proposizione 3.5.2. *Dati tre sottoinsiemi non vuoti $M, N, P \subseteq S$ vale la disuguaglianza*

$$\operatorname{dist}(M,N) \leq \operatorname{dist}(M,P) + \operatorname{dist}(N,P) + \operatorname{diam} P$$
$$\operatorname{diam}(M \cup N) \leq \operatorname{diam} M + \operatorname{diam} N + \operatorname{dist}(M,N)$$

Dimostrazione. Per ogni $\varepsilon > 0$ esistono $m \in M, n \in N, p, p' \in P$ tali che

$$d(m,p) \leq \operatorname{dist}(M,P) + \varepsilon$$
$$d(n,p') \leq \operatorname{dist}(N,P) + \varepsilon$$

Per la disuguaglianza triangolare si avrà allora

$$\operatorname{dist}(M,N) \leq d(m,n) \leq d(m,p) + d(p,p') + d(p',n) \leq \operatorname{dist}(M,P) + \operatorname{diam} P + \operatorname{dist}(N,P) + 2\varepsilon$$

e la disuguaglianza finale si ottiene facendo tendere ε a 0.

Per quanto riguarda la seconda disuguaglianza prendiamo $x, y \in M \cup N$, se appartengono entrambi ad M allora $d(x,y) \leq \operatorname{diam} M$ e allo stesso modo se appartengono entrambi ad N. Consideriamo allora $x \in M, y \in N$ per ogni $x' \in M, y' \in N$ si ha

$$d(x,y) \leq d(x,x') + d(x',y') + d(y',y) \leq \operatorname{diam} M + d(x',y') + \operatorname{diam} N$$

dato che x, y, x', y' non dipendono l'uno dall'altro avremo che

$$\operatorname{diam}(M \cup N) \leq \operatorname{diam} M + \inf_{\substack{x' \in M \\ y' \in N}} d(x', y') + \operatorname{diam} N = \operatorname{diam} M + \operatorname{dist}(M, N) + \operatorname{diam} N$$

da cui la tesi. ■

Non è possibile eliminare il termine $\operatorname{diam} P$ dalla disuguaglianza precedente. Si consideri ad esempio $M = \{x\}, N = \{y\}, P = \{x, y\}$ con $x \neq y$.

Data una qualunque pseudometrica d su S definiamo anche i seguenti insiemi:

$$B(x, r, d) = \{y \in S \mid d(x, y) < r\}$$
$$D(x, r, d) = \{y \in S \mid d(x, y) \leq r\}$$

al variare di $x \in S$ ed $r > 0$. Essi saranno detti rispettivamente **palla aperta** e **palla chiusa** centrate in x, di raggio r rispetto alla pseudometrica d. Poiché $B(x, r, d) = U(\varepsilon, d)[x]$ se dotiamo S della struttura uniforme generata dalle pseudometriche $\{d_i\}_{i \in I}$ gli insiemi $B(x, r, d_i)$ ed $D(x, r, d_i)$ sono intorni di x che al variare di r ed i formano due sistemi essenziali di intorni per x, in particolare se $I = \{i_1\}$ allora essi formeranno anche un sistema fondamentale di intorni. Dimostriamo ora che

Proposizione 3.5.3. *Se dotiamo X della struttura uniforme generata dalle pseudometriche $\{d_i\}_{i \in I}$ allora le palle aperte sono aperte e le palle chiuse sono chiuse (guarda un po', chi lo avrebbe mai detto?).*

Dimostrazione. Per far vedere che $B(x, r, d_i)$ è aperto dobbiamo solamente far vedere che è un intorno di ogni suo punto. Prendiamo allora $y \in B(x, r, d_i)$ dunque $s = d_i(x, y) < r$ e per la disuguaglianza triangolare si avrà $B(y, r - s, d_i) \subseteq B(x, r, d_i)$, per l'arbitrarietà di y segue che la palla aperta è aperta.

Prendiamo ora $y \notin D(x, r, d_i)$ dunque $s = d_i(x, y) > r$ e perciò si avrà $B(y, s - r, d_i) \cap D(x, r, d_i) = \emptyset$, per l'arbitrarietà di y segue che la palla chiusa è chiusa. ■

Dalla proposizione di prima segue anche che $\operatorname{Cl}(B(x, r, d)) \subseteq D(x, r, d)$ per ogni $x \in S$, $r > 0$ e per ogni pseudometrica d. Ciononostante l'inclusione può essere anche stretta, ovvero la palla chiusa in genere non coincide con la chiusura della relativa palla aperta.

Dato che gli insiemi $U(\varepsilon, d_i)$ formano una semibase uniforme per X non è difficile dimostrare il seguente risultato:

Proposizione 3.5.4. *Data una net x_a avremo che $x_a \to x$ se e solo se $\lim_{a \to +\infty} d_i(x_a, x) = 0$ per ogni $i \in I$. Invece x_a è una Cauchy net se e solo se per ogni $\varepsilon > 0$ e per ogni $i \in I$ esiste $a \in A$ tale che $d_i(x_b, x_c) < \varepsilon$ per ogni $b, c \geq a$.*

Dato un insieme X ed una pseudometrica d chiameremo la coppia (X, d) uno **spazio pseudometrico**, mentre se d è una metrica lo chiameremo anche **spazio metrico**. Ogni spazio pseudometrico possiede un'unica struttura uniforme indotta dalla metrica d definita come sopra, viceversa uno spazio uniforme si dice **pseudometrizzabile** se e solo se possiede una pseudometrica che induce la sua struttura uniforme. Chiaramente tale metrica può non essere unica.

Teorema 3.5.5. *Uno spazio uniforme X è generato da una sola pseudometrica se e solo se è uniformemente $\mathcal{N}_1$, se X è anche di Hausdorff allora la pseudometrica in realtà è una metrica. Inoltre le palle aperte/chiuse in x con raggi in $\mathbb{Q}^+$ formano un sistema fondamentale di intorni per x.*

Dimostrazione. Chiaramente ogni spazio pseudometrico (X, d) possiede la base uniforme numerabile data da

$$\left\{(x, y) \;\middle|\; d(x, y) < \frac{1}{n}\right\}$$

con $n \in \mathbb{N}$. Sia allora V_n una base uniforme per X al variare di $n \in \mathbb{N}$, Posto $U_1 \subseteq V_1$ simmetrico e per ricorrenza definiamo $U_n \subseteq V_n$ simmetrico in modo tale che $U_{n+1} \diamond U_{n+1} \subseteq U_n$.

Ma allora possiamo costruire la pseudometrica D utilizzando proprio la successione $\{U_n\}_{n\in\mathbb{N}}$, inoltre essendo $U_n \subseteq V_n$ dalla (3.2.3) segue che la struttura uniforme generata da D è proprio la struttura uniforme di X.

Poiché X è di Hausdorff per ogni $x, y \in X$ distinti esiste V_n tale che $y \notin V_n[x]$ e per la (3.2.3) segue che $D(y, x) \geq 2^{-n}$, quindi D è una metrica su X. ■

Corollario 3.5.6. *Dati due spazi pseudometrici X ed Y allora anche $X \times Y$ con la struttura uniforme prodotto è uno spazio pseudometrico. In particolare de d_X e d_Y sono delle pseudometriche che inducono la topologia di X ed Y rispettivamente allora*

$$d[(x, y), (x', y')] = d_X(x, x') \vee d_Y(y, y')$$
$$e[(x, y), (x', y')] = d_X(x, x') + d_Y(y, y')$$

sono alcune delle metriche che inducono la struttura uniforme prodotto di $X \times Y$.

Osserviamo che in uno spazio pseudometrico i concetti di funzione continua e funzione uniformemente continua ricalcano fedelmente la definizione data nei corsi di Analisi 1 per quanto riguarda le funzioni da $\mathbb{R}$ in $\mathbb{R}$. Data una funzione $f : S \to T$, un punto di accumulazione $s \in S$ e un valore $t \in T$ generico. Allora $f(x) \to t$ al tendere di x ad s se e solo se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $x \in S$ tale che $d_S(x, s) < \delta$ si ha $d_T[f(x), t] < \varepsilon$.

Definizione 3.5.7. Un'applicazione $f : S \to T$ tra due spazi metrici è della **Lipschitziana** se e solo se esiste una costante $L \geq 0$ tale che

$$d_T[f(x), f(y)] \leq L d_S(x, y)$$

Se inoltre $L < 1$ f è detta **contrazione**.

Le funzioni lipschitziane sono uniformemente continue e quindi continue su tutto lo spazio. Ovviamente la condizione di lipschitzianità dipende enormemente dalle metriche scelte e non dalla struttura uniforme sottostante. Il fatto che gli spazi pseudometrici siano anche uniformemente $\mathcal{N}_1$ offre una definizione analoga di uniforme continuità tramite successioni di insiemi.

Proposizione 3.5.8. *Una funzione $f : S \to T$ tra spazi pseudometrici è uniformemente continua se e solo se per ogni successione di sottoinsiemi non vuoti B_n di S tali che* $\operatorname{diam} B_n \to 0$ *si ha* $\operatorname{diam} f(B_n) \to 0$. *Invece f è L-lipschitziana se e solo se* $\operatorname{diam} f(B) \leq L \operatorname{diam} B$ *per ogni $B \subseteq S$ non vuoto.*

Dimostrazione. Supponiamo innanzitutto che f sia uniformemente continua e per ogni $\varepsilon > 0$ definiamo δ in modo tale che $d[f(x), f(y)] < \varepsilon$ se $d(x, y) < \delta$. Presa una generica successione $B_n \subseteq S$ tale che $\operatorname{diam} B_n \to 0$ esisterà $N \in \mathbb{N}$ tale che $\operatorname{diam} B_n < \delta$ per ogni $n > N$, quindi per ogni $x, y \in B_n$ si ha $d(x, y) < \delta$ e perciò $d[f(x), f(y)] < \varepsilon$. Passando all'estremo superiore avremo che $\operatorname{diam} f(B_n) \leq \varepsilon$ per ogni $n > N$ e perciò $\operatorname{diam} f(B_n) \to 0$.

Viceversa supponiamo per assurdo che f non sia uniformemente continua su S. Esisterà dunque, a meno di passare ad un'estratta, una costante $\varepsilon_0 > 0$ tale che per ogni $n \in \mathbb{N}$ esistono $x_n, y_n \in S$ tali che $d(x_n, y_n) < 1/n$ ma $d[f(x_n), f(y_n)] \geq \varepsilon_0$. Posto $B_n = \{x_n, y_n\}$ avremo che $\operatorname{diam} B_n \to 0$ ma $\operatorname{diam} f(B_n) \geq \varepsilon_0$ e quindi non converge a 0.

Il caso in cui f sia lipschitziana si dimostra analogamente. ■

Proposizione 3.5.9. *Se (S, d) è uno spazio pseudometrico e $H \subseteq S$ non vuoto allora l'applicazione*

$$d_H : x \in S \to \operatorname{dist}(x, H) \in \mathbb{R}$$

è 1-Lipschitziana.

Lemma 3.5.10 (Caratterizzazione della chiusura)**.** *Sia (S, d) uno spazio pseudometrico e $A \subseteq S$ un suo sottoinsieme non vuoto. Allora le seguenti affermazioni sono equivalenti tra loro*

1. $x \in \operatorname{Cl}(A)$;
2. $\exists x_n \in A$ *tale che* $x_n \to x$;
3. $\operatorname{dist}(x, A) = 0$.

Dimostrazione. L'equivalenza 1⇔2 si ottiene verificando che ogni spazio metrico soddisfa l'assioma di numerabilità $\mathcal{N}_1$ mentre l'equivalenza 1⇔3 deriva dalla proposizione 3.1.7. ■

Corollario 3.5.11. *Dato un qualunque sottoinsieme non vuoto A di uno spazio pseudometrico (X, d) abbiamo che* $\operatorname{diam} A = \operatorname{diam} \operatorname{Cl}(A)$.

Dimostrazione. Chiaramente $\operatorname{diam} A \leq \operatorname{diam} \operatorname{Cl}(A)$, viceversa prendiamo $x, y \in \operatorname{Cl}(A)$ generici e per il lemma precedente avremo che $\operatorname{dist}(x, A) = \operatorname{dist}(y, A) = 0$. Ma allora per ogni $\varepsilon > 0$ esistono $x', y' \in A$ tali che $d(x, x'), d(y, y') < \varepsilon$ e perciò

$$d(x, y) \leq d(x, x') + d(x', y') + d(y', y) < 2\varepsilon + \operatorname{diam} A$$

Passando all'estremo superiore avremo $\operatorname{diam} \operatorname{Cl}(A) \leq 2\varepsilon + \operatorname{diam} A$ e la tesi si ottiene facendo tendere ε a 0. ■

Proposizione 3.5.12. *Se M è chiuso e N sequenzialmente compatto allora* $\operatorname{dist}(M, N) = 0 \Leftrightarrow M \cap N \neq \emptyset$.

Dimostrazione. L'implicazione ⇐ è banale, se $\operatorname{dist}(M, N) = 0$ allora esistono $\{x_n\}_{n\in\mathbb{N}} \subseteq M$, $\{y_n\}_{n\in\mathbb{N}} \subseteq N$ tali che $\lim_{n\to+\infty} d(x_n, y_n) = 0$. Dato che N è compatto per successioni avremo un'estratta y_{n_k} che convergerà ad un certo $y \in N$ da cui segue che

$$\lim_{k\to+\infty} d(x_{n_k}, y) \leq \lim_{k\to+\infty} d(x_{n_k}, y_{n_k}) + \lim_{k\to+\infty} d(y_{n_k}, y) = 0$$

e quindi x_{n_k} tende anch'esso a y. Dalla chiusura di M allora $y \in M$ e perciò $M \cap N \neq \emptyset$. ■

Per gli spazi pseudometrici vale una proprietà più forte della completa regolarità: per ogni coppia di insiemi chiusi disgiunti A,B esiste una funzione continua $f : S \to [0,1]$ tale che $x \in A$ se e solo se $f(x) = 0$ e $x \in B$ se e solo se $f(x) = 1$. Basta infatti prendere come funzione f la quantità

$$f(x) = \frac{\operatorname{dist}(x,A)}{\operatorname{dist}(x,A) + \operatorname{dist}(x,B)} \tag{3.5.1}$$

che per la chiusura di A e B è ben definita su tutto S e soddisfa le proprietà richieste.

A differenza di ciò che il buonsenso lascerebbe intendere, in uno spazio pseudometrico generico la chiusura della palla aperta non coincide quasi mai con la palla chiusa con lo stesso centro e lo stesso raggio. In generale vale il seguente risultato di equivalenza:

Proposizione 3.5.13. *Sia (S,d) spazio pseudometrico, allora sono equivalenti*

1. *per ogni $x \in S$, $r > 0$ si ha $\operatorname{Cl}(B(x,r)) = D(x,r)$;*
2. *per ogni $x,y \in S$ con $x \neq y$ e per ogni $\varepsilon > 0$ esiste $z \in S$ tale che $d(y,z) < \varepsilon$ e $d(x,z) < d(x,y)$.*

Dimostrazione. Supponiamo che vale il primo punto e siano $x,y \in S$ con $r = d(x,y) > 0$ allora $y \in \operatorname{Cl}(B(x,r))$. Dunque per ogni $\varepsilon > 0$ per la proposizione 2.1.2 l'intersezione $B(x,r) \cap B(y,\varepsilon)$ non è mai vuota, e ogni elemento z appartenente ad essa verificherà il secondo punto.

Viceversa supponiamo che valga il secondo punto e prendiamo un qualunque elemento $y \in D(x,r)$. Ma allora per ipotesi per ogni $\varepsilon > 0$ esiste $z_\varepsilon \in B(x,r)$ tale che $d(y,z_\varepsilon) < \varepsilon$ e perciò $B(x,r) \cap B(y,\varepsilon) \neq \emptyset$. Applicando sempre la proposizione 2.1.2 deduciamo che $y \in \operatorname{Cl}(B(x,r))$ e per l'arbitrarietà di y otteniamo l'uguaglianza richiesta. ∎

In uno spazio metrico net-completezza e sequenziale completezza coincidono, quindi si può parlare semplicemente di completezza. Un notevole risultato degli spazi metrici completi è il teorema delle contrazioni di Banach-Cacioppoli enunciato nella seguente maniera:

Teorema 3.5.14 (Banach-Cacioppoli)**.** *Se S è uno spazio metrico completo e $f : S \to S$ contrazione allora esiste un unico $x \in S$ tale che*

$$f(x) = x$$

Dimostrazione. Sia $L \in [0,1)$ tale che per ogni $x,y \in S$ vale $d[f(x),f(y)] \leq Ld(x,y)$, fissiamo $x_0 \in S$ allora definiamo la seguente successione per ricorrenza come

$$x_{n+1} = f(x_n)$$

vogliamo dimostrare che è una successione di Cauchy. Innanzitutto

$$d(x_{n+1},x_n) = d[f(x_n),f(x_{n-1})] \leq Ld(x_n,x_{n-1}) \leq L^2 d(x_{n-1},x_{n-2}) \leq L^n d(x_1,x_0) = CL^n$$

poi

$$d(x_{n+1},x_0) \leq \sum_{i=0}^{n} d(x_{i+1},x_i) \leq \sum_{i=0}^{n} CL^i = C\frac{1-L^{n+1}}{1-L} \leq \frac{C}{1-L}$$

infine

$$d(x_{n+k}, x_k) \leq L^k d(x_n, x_0) \leq \frac{CL^k}{1-L}$$

il termine al secondo membro non dipende da n e decresce al crescere di k, quindi è di Cauchy.

Poiché S è completo allora la successione tende a un certo x, dall'uniforme continuità di f si può passare al limite ottenendo $x = f(x)$. ■

Prima di passare al prossimo argomento diamo altre definizioni.

Definizione 3.5.15. Dati due spazi metrici (X, d_X), (Y, d_Y) e una applicazione $f : X \to Y$ diremo che f è

- **bilipschitziana** se e solo se esistono due costanti $M, N > 0$ tali che per ogni $a, b \in X$

$$N\, d_X(a, b) \leq d_Y[f(a), f(b)] \leq M\, d_X(a, b)$$

- una **isometria** se e solo se per ogni $a, b \in X$ si ha

$$d_Y[f(a), f(b)] = d_X(a, b)$$

Due spazi metrici X, Y si dicono **equivalenti** se e solo se esiste una isometria biettiva tra X ed Y.

Nel primo capitolo abbiamo definito per uno spazio topologico generico i sottoinsiemi rari. Per come abbiamo definito la topologia su uno spazio metrico gli insiemi rari sono tutti quei sottoinsiemi la cui chiusura non contiene palle aperte.

Il seguente risultato, che riguarda gli spazi metrici completi, avrà una grande importanza nei prossimi capitoli.

Teorema 3.5.16 (Baire)**.** *Ogni spazio metrico completo è di* ***seconda categoria****, ovvero non è unione numerabile di insiemi rari.*

Dimostrazione. Per assurdo $S = \bigcup_{n=1}^{+\infty} S_n$ con gli S_n chiusi e rari, allora esisterà un $x_1 \in S \setminus S_1$ e un $r_1 > 0$ tale che $D(x_1, r_1) \subseteq S \setminus S_1$ e $r_1 \leq \frac{1}{2}$. Essendo S_2 raro allora l'insieme $B(x_1, r_1) \setminus S_2$ è necessariamente non vuoto e aperto, possiamo perciò trovare un $x_2 \in B(x_1, r_1) \setminus S_2$ e $r_2 > 0$ tali che $D(x_2, r_2) \subseteq B(x_1, r_1) \setminus S_2$ e $r_2 \leq \frac{r_1}{2}$.

Continuando per ricorrenza otteniamo una successione di elementi x_n e reali r_n tali che

$$\begin{gathered} D(x_{n+1}, r_{n+1}) \subseteq B(x_n, r_n) \\ B(x_n, r_n) \cap S_n = \emptyset \\ r_{n+1} \leq \frac{r_n}{2} \leq \frac{1}{2^n} \end{gathered}$$

osserviamo che x_n è una successione di Cauchy: per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che $\frac{1}{2^{N+1}} < \epsilon$ e per ogni $m, n > N$ è immediato dimostrare che $d(x_m, x_n) < 2r_N < \epsilon$.

Ricordando che S è completo la successione x_n convergerà ad un certo x, inoltre per ogni $m \in \mathbb{N}$ chiaramente $x \in D(x_m, r_m)$ essendo tali insiemi chiusi, ciò implicherebbe che $x \notin \bigcup_{n=1}^{+\infty} S_n$ assurdo. ■

Corollario 3.5.17. *Se S è uno spazio metrico ed $\{X_i\}_{i\in\mathbb{N}}$ è una famiglia numerabile di sottoinsiemi chiusi e rari di S allora $X=\bigcup_{i=1}^{+\infty} X_i$ non contiene punti interni.*

Dimostrazione. Per assurdo esistono $x \in S$ ed $r > 0$ tali che $B(x,r) \subseteq X$ e quindi $D(x,r/2) \subseteq X$. Sottoinsiemi chiusi di spazi metrici completi sono ancora spazi metrici completi e $X_i \cap D(x,r/2)$ è raro in $D(x,r/2)$. Ma

$$\bigcup_{i=1}^{+\infty} [X_i \cap D(x,r/2)] = X \cap D(x,r/2) = D(x,r/2)$$

il che è assurdo, perciò X non possiede punti interni. ■

Corollario 3.5.18. *Se X è uno spazio metrico completo e $\{G_i\}_{i\in\mathbb{N}}$ è una famiglia numerabile di sottoinsiemi aperti e densi allora $\bigcap_{i=1}^{+\infty} G_i$ è denso in S.*

Dimostrazione. Se G è aperto e denso allora $S \setminus G$ è chiuso e raro. Dal corollario precedente l'insieme

$$\bigcup_{i=1}^{+\infty} (S \setminus G_i) = S \setminus \bigcap_{i=1}^{+\infty} G_i$$

non contiene aperti, ciò implica che $\bigcap_{i=1}^{+\infty} G_i$ interseca ogni aperto di S ed è perciò denso. ■

Per quanto riguarda i teoremi di compattezza il teorema 3.3.12 riadattato agli spazi metrici assumerà inoltre la seguente forma.

Teorema 3.5.19. *In uno spazio metrico (S,d) le seguenti affermazioni*

- *S relativamente compatto;*
- *S relativamente sequenzialmente compatto;*
- *S totalmente limitato*

sono equivalenti tra loro.

Dimostrazione. Supponiamo S totalmente limitato e x_n una successione generica un S, se essa è definitivamente costante allora ammetterà sicuramente un'estratta di Cauchy perciò supponiamola senza perdere in generalità non definitivamente costante e quindi l'insieme $A_0 = \{x_n \mid n \in \mathbb{N}\}$ conterrà un numero infinito di elementi distinti. Ora essendo S totalmente limitato esisterà un insieme finito B_0 i cui elementi sono sfere di raggio unitario che ricoprono S, perciò esisterà $A_1 \subseteq A_0$ infinito i cui elementi sono interamente contenuti in un unico elemento di B_0, esisterà inoltre $n_1 \in \mathbb{N}$ tale che $x_{n_1} \in A_1$. Esisterà ancora un insieme finito B_1 di sfere di raggio $\frac{1}{2}$ che ricopre S, allora esisterà un sottoinsieme $A_2 \subseteq A_1$ contenuto in un elemento fissato di B_1. Poniamo allora $n_2 > n_1$ tale che $x_{n_2} \in A_2$.

Procedendo per ricorrenza esisterà un ricoprimento B_k di palle aperte di raggio $\frac{1}{2^k}$ di S e un sottoinsieme infinito $A_{k+1} \subseteq A_k$ contenuto in un elemento fissato di B_k ed esisterà $n_{k+1} > n_k$ tale che $x_{n_{k+1}} \in A_{k+1}$. Ma allora per ogni $i > j$

$$d(x_{n_i}, x_{n_j}) \leq \frac{1}{2^{j-1}}$$

che è un'estratta di Cauchy della successione iniziale. ■

Teorema 3.5.20 (Hausdorff)**.** *Se S è uno spazio metrico allora le seguenti affermazioni sono equivalenti:*

1. *S compatto;*
2. *S sequenzialmente compatto;*
3. *S totalmente limitato e completo.*

Dimostrazione. Il lettore sfruttando il teorema e il lemma precedente può agevolmente dimostrare che i punti 2 e 3 sono equivalenti. Ci basta perciò da dimostrare che 1 implica 2 e che 3 implica la 1.

Supponiamo che S non sia sequenzialmente compatto ovvero esiste una successione x_n che non ha estratte convergenti a nessun punto di S, in altre parole per ogni $x \in S$ esisterà un $r_x > 0$ tale che l'insieme $\{n \in \mathbb{N} \mid x_n \in B(x, r_x)\}$ è finito. Allora la famiglia $\{B(x, r_x)\}_{x \in S}$ è un ricoprimento aperto e se esistesse un sottoricoprimento finito di esso ci sarebbero solo un numero finito di indici n tali che x_n sia contenuto in almeno uno di essi, quindi S non può essere compatto. Allora la compattezza implica la compattezza sequenziale.

Ci basta dimostrare allora che se S è completo e totalmente limitato allora è compatto. Per assurdo sia $\mathcal{V}$ ricoprimento di aperti di S che non ammette sottoricoprimenti finiti, per la totale limitatezza esisterà una palla $B(x_1, 1)$ tale che $\mathcal{V}$ non possiede sottoricoprimenti finiti di $B(x_1, 1)$.

Si dimostra facilmente che anche $B(x_1, 1)$ è totalmente limitato poiché contenuto in uno spazio totalmente limitato, quindi esiste una palla $B\left(x_2, \frac{1}{2}\right)$ con $d(x_1, x_2) < 1$ tale che $\mathcal{V}$ non ammetta sottoricoprimenti finiti. Continuando di questo passo troveremo una successione x_n tale che $d(x_n, x_{n+1}) < \frac{1}{2^{n-1}}$ tale che $\mathcal{V}$ non abbia sottoricoprimenti finiti di $B\left(x_n, \frac{1}{2^{n-1}}\right)$. Ma x_n è allora di Cauchy e per la completezza avrà limite $x \in S$.

Poiché esiste $V \in \mathcal{V}$ tale che $x \in V$ e sia $r > 0$ tale che $B(x, r) \subseteq V$ esisterà sicuramente $n \in \mathbb{N}$ tale che $\frac{1}{2^{n-1}} < \frac{r}{2}$ e $d(x_n, x) < \frac{r}{2}$ e quindi $B\left(x_n, \frac{1}{2^{n-1}}\right) \subseteq B(x, r) \subseteq V$ assurdo perché $\mathcal{V}$ non aveva sottoricoprimenti finiti su tale insieme.

Allora S deve essere per forza compatto. ■

Corollario 3.5.21 (Heine-Borel per spazi metrici)**.** *Se (S, d) è uno spazio metrico completo e $T \subseteq S$ allora T è compatto se e solo se è chiuso e totalmente limitato.*

Molto interessanti per gli spazi metrici sono anche le proprietà di separabilità. Valgono infatti i seguenti risultati:

Proposizione 3.5.22. *Gli spazi metrici totalmente limitati sono separabili. In particolare tutti gli spazi metrici compatti/sequenzialmente compatti sono separabili.*

Dimostrazione. Per ogni $k \in \mathbb{N}$ esiste un sottoinsieme finito $\Sigma_k \subseteq S$ tale che

$$S = \bigcup_{x \in \Sigma_k} B\left(x, \frac{1}{k}\right)$$

dimostriamo che l'insieme $\Sigma = \bigcup_{k=1}^{+\infty} \Sigma_k$ è denso in S.

Per ogni $\varepsilon > 0$ esisterà $k \in \mathbb{N}$ tale che $\varepsilon > 1/k$ e quindi

$$\mathrm{Cl}_\varepsilon(\Sigma) \supseteq \bigcup_{i=1}^{n_k} B\left(x_{i,k}, \varepsilon\right) \supseteq \bigcup_{i=1}^{n_k} B\left(x_{i,k}, \frac{1}{k}\right) = S$$

per la proposizione 3.1.7 segue infine che $\mathrm{Cl}(\Sigma) = \bigcap_{\varepsilon > 0} \mathrm{Cl}_\varepsilon(\Sigma) = S$. ■

Questo risultato mostra un collegamento tra compattezza e separabilità all'interno degli spazi metrici. Inoltre la separabilità limitata agli spazi metrici implica molte altre proprietà che non valgono in generale sugli spazi topologici separabili, in particolare vale il seguente risultato

Proposizione 3.5.23. *Gli spazi metrici separabili soddisfano l'assioma di numerabilità $\mathcal{N}_2$.*

Dimostrazione. Sia A sottoinsieme denso numerabile di S, ovvero per ogni $x \in S$, $r > 0$ si ha $S \cap B(x, r) \neq \emptyset$. Consideriamo ora la classe

$$\mathscr{B} = \{B(x, q) \mid x \in A \wedge q \in \mathbb{Q}^+\}$$

è chiaramente numerabile. Per dimostrare che è anche una base dobbiamo mostrare che per ogni aperto U e $x \in U$ esiste $B \in \mathscr{B}$ tale che $x \in B \subseteq U$.

Innanzitutto esiste un $r > 0$ e un elemento $y \in A$ tali che $y \in B(x, r) \subseteq U$ da cui, ponendo $r' = r - d(x, y) > 0$, abbiamo $B(y, r') \subseteq B(x, r)$. La tesi segue immediatamente ricordando che esiste sempre un razionale q tale che $0 < q < r'$. ■

Da questo risultato possiamo applicare il lemma di Lindelöf e quindi dimostrare l'equivalenza tra compattezza e sequenziale compattezza all'interno di spazi metrici separabili. Un altra notevole proprietà degli spazi metrici separabili è che la separabilità viene ereditata automaticamente da tutti i suoi sottospazi come mostra il seguente risultato.

Lemma 3.5.24. *Ogni sottoinsieme di uno spazio metrico separabile con la metrica indotta è ancora uno spazio metrico separabile.*

Dimostrazione. Sia $\{x_n\}_{n\in\mathbb{N}}$ denso in S e A un suo sottoinsieme, poniamo

$$W = \left\{(i, j) \in \mathbb{N}^2 \;\middle|\; A \cap B\left(x_i, \frac{1}{j}\right) \neq \emptyset\right\}$$

e $x_{i,j} \in A \cap B\left(x_i, \frac{1}{j}\right)$ con $(i, j) \in W$.

Dalla densità in S segue che per ogni $x \in A$ esiste necessariamente un n_k tale che $x \in B\left(x_{n_k}, \frac{1}{k}\right)$ e quindi $(n_k, k) \in W$ è numerabile. Infine osserviamo che

$$d(x, x_{n_k,k}) \leq d(x, x_{n_k}) + d(x_{n_k}, x_{n_k,k}) \leq \frac{2}{k}$$

Perciò il sottoinsieme $\left\{x_{n_k,k} \mid k \in \mathbb{N}\right\}$ è denso in A. ■

Osservazione. Si noti che questo risultato non vale per spazi topologici generici. Esistono infatti spazi topologici separabili che possiedono sottospazi non separabili.

3.6 Convergenza uniforme

Divideremo questa sezione in due parti: nella prima ci dedicheremo alle "metriche" d che, pur soddisfacendo i tre assiomi delle metriche, possono assumere anche il valore $+\infty$, mentre nella seconda applicheremo tale concetto allo studio di particolari spazi di funzioni dotati di una metrica di questo tipo.

Dato un insieme non vuoto X ed un'applicazione d definita su $X \times X$ ed a valori in $\overline{\mathbb{R}}$ tale che

- $d(x,y) \geq 0$ ed $d(x,x) = 0$ per ogni $x \in X$;
- $d(x,y) = d(y,x)$;
- $d(x,y) \leq d(x,z) + d(z,y)$.

L'unica proprietà che la distingue dalle pseudometriche è la possibilità che essa valga $+\infty$, per comodità diremo che d è una **pseudometrica foliata**. Possiamo però rendere X uno spazio pseudometrico nel senso usuale dotandolo della metrica d' definita come

$$d'(x,y) = d(x,y) \wedge 1$$

(al posto di 1 possiamo mettere qualunque numero positivo) che chiaramente è una pseudometrica su X. Osserviamo inoltre che la struttura uniforme su X generata da d' coincide con quella che si otterrebbe invece da d in quanto in una struttura uniforme "contano gli entourage piccoli" piuttosto che quelli grandi, ed infatti d e d' coincidono per valori piccoli. Quindi lo studio delle pseudometriche foliate si rifà allo studio delle proprietà uniformi degli spazi pseudometrici (anche se altre proprietà sono più delicate da affrontare, si veda la limitatezza).

Osserviamo però che i tre assiomi assicurano che la relazione su X definita come

$$x \sim y \Leftrightarrow d(x,y) < +\infty$$

è una relazione di equivalenza, quindi le classi di equivalenza

$$X_u = \{v \in X \mid v \sim u\}$$

formano una partizione dello spazio X sui quali la pseudometrica d è una pseudometrica nel senso usuale. Queste classi sono dette **foglie** di (X,d) che quindi può essere visto come l'unione disgiunta delle sue foglie. Viceversa infatti data una famiglia di spazi pseudometrici $\{(X_i, d_i)\}_{i\in I}$ a due a due disgiunti posto $X = \bigcup_{i\in I} X_i$ possiamo definire una pseudometrica foliata d su X in questa maniera:

$$d(x,y) = \begin{cases} d_i(x,y) & \text{se } x,y \in X_i \text{ per qualche } i \in I \\ +\infty & \text{altrimenti} \end{cases}$$

e le foglie di X coincidono esattamente con gli X_i.

Un notevole esempio di spazio metrico foliato è lo spazio $F(X,Y)$ delle funzioni definite su un insieme X a valori in uno spazio metrico (Y,d) con la metrica foliata

$$d_\infty(f,g) = \sup_{x\in X} d[f(x), g(x)]$$

e quindi $[F(X,Y), d_\infty]$ è uno spazio metrico foliato. Per comodità diremo che una successione di funzioni f_n **converge uniformemente** ad f se e solo se converge rispetto alla metrica d_∞, ovvero se per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che per ogni $x \in X$ e per ogni $n > N$ si ha $d[f_n(x), f(x)] < \varepsilon$.

Lemma 3.6.1 (Inversione dei limiti). *Dati uno spazio topologico X, uno spazio metrico Y con metrica d, un punto di accumulazione x di X e una successione $f_n \in F(X,Y)$ convergente* *uniformemente ad f tale che il limite $z_n = \lim_{y\to x} f_n(y)$ esiste per ogni $n \in \mathbb{N}$. Allora se Y è completo i due limiti*

$$\lim_{n\to+\infty}\left(\lim_{y\to x} f_n(y)\right) \tag{3.6.1}$$

$$\lim_{y\to x} f(y) \tag{3.6.2}$$

esistono e sono uguali. Se invece Y non è completo ed esiste almeno uno dei due limiti allora esisterà anche l'altro e i due limiti coincidono.

Dimostrazione. Supponiamo inizialmente che Y sia completo e dimostriamo che il limite (3.6.1) esiste sempre. Innanzitutto per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che per ogni $m, n > N$ e per ogni $y \in X$ si ha

$$d[f_m(y), f_n(y)] < \varepsilon$$

e quindi passando al limite in y si avrà $d(z_m, z_n) \leq \varepsilon$ per ogni $m, n > N$. Per la completezza di Y la successione z_n convergerà ad un certo $z \in Y$ e quindi il limite (3.6.1) esiste.

Supponiamo adesso che il limite (3.6.1) esiste ed è uguale a z, di conseguenza esisterà un indice $N \in \mathbb{N}$ tale che $d_\infty(f_N, f) < \varepsilon$ ed $d(z_N, z) < \varepsilon$. Allora

$$d[f(y), z] \leq d_\infty(f, f_N) + d[f_N(y), z_N] + d(z_N, z) < 2\varepsilon + d[f_N(y), z_N]$$

e dato che N non dipende da y passando al limite si ottiene che $\lim_{y\to x} d[f(y), z] \leq 2\varepsilon$ per ogni $\varepsilon > 0$. Ciò è equivalente ad affermare che $\lim_{y\to x} f(y) = z$ e quindi il limite (3.6.2) esiste e coincide con (3.6.1).

Viceversa supponiamo che il limite (3.6.2) esiste ed è uguale a z. Poiché la successione f_n converge uniformemente ad f esisterà $N \in \mathbb{N}$ tale che per ogni $y \in X$ e per ogni $n > N$ si abbia $d[f_n(y), f(y)] < \varepsilon$. Dato che n non dipende da y passando al limite avremo che $d(z_n, z) \leq \varepsilon$ per ogni $n > N$ e quindi il limite (3.6.1) esiste ed è uguale a (3.6.2). ∎

Corollario 3.6.2. *Se X è uno spazio topologico allora lo spazio $C(X, Y)$ è un sottospazio chiuso di $F(X, Y)$.*

Dimostrazione. Fissiamo un generico punto di accumulazione x di X e sia f_n una successione di funzioni continue convergenti uniformemente ad f. Allora

$$\lim_{n\to+\infty}\left(\lim_{y\to x} f_n(y)\right) = \lim_{n\to+\infty} f_n(x) = f(x)$$

e quindi il limite (3.6.1) esiste, per il lemma di inversione dei limiti avremo che $f(y) \to f(x)$ al tendere di y ad x e quindi per l'arbitrarietà di x si ha la tesi. ∎

Corollario 3.6.3. *Il limite uniforme di una successione di funzioni uniformemente continue, se esiste, è uniformemente continua.*

Dimostrazione. Data una successione di funzioni uniformemente continue $f_n \in C(X, Y)$ convergente ad $f \in C(X, Y)$, quindi per ogni $\varepsilon > 0$ esiste $n \in \mathbb{N}$ tale che $d_\infty(f_n, f) < \varepsilon/3$. Di conseguenza esisterà $\delta_n > 0$ tale che per ogni $x, y \in X$ con $d(x, y) < \delta_n$ si ha $d[f_n(x), f_n(y)] < \varepsilon/3$. Ma allora

$$d[f(x), f(y)] \leq d[f(x), f_n(x)] + d[f_n(x), f_n(y)] + d[f_n(y), f(y)] < \varepsilon$$

e dato che n dipende solo da ε segue che anche f è uniformemente continua. ∎

Teorema 3.6.4. *Se (Y, d) è uno spazio metrico completo allora anche $[F(X, Y), d_\infty]$ è completo.*

Dimostrazione. Sia f_n una successione di Cauchy in $F(X, Y)$ allora $f_n(x)$ è di Cauchy in Y per ogni $x \in X$. Per la completezza di Y esisterà $f \in F(X, Y)$ tale che $f(x) = \lim_{n\to+\infty} f_n(x)$. Ora per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che per ogni $m, n > N$ e per ogni $x \in X$ si ha

$$d[f_m(x), f_n(x)] < \varepsilon$$

Passando al limite su m avremo che $d[f(x), f_n(x)] \leq \varepsilon$ per ogni $x \in X$ e per ogni $n > N$ e dunque

$$d_\infty(f_n, f) = \sup_{x\in X} d[f_n(x), f(x)] \leq \varepsilon$$

e f_n converge uniformemente ad f. ∎

Corollario 3.6.5. *Se (Y, d) è completo e X è uno spazio topologico generico allora anche spazio $C(X, Y)$ è completo rispetto alla metrica d_∞.*

Il prossimo risultato che dimostreremo è il famosissimo teorema di Ascoli-Arzelà, un risultato molto utilizzato all'interno dell'analisi funzionale in quanto fornisce un fondamentale criterio di compattezza per gli spazi di funzioni (continue). In queste note daremo una versione più generale del teorema, che invece di limitarsi alle funzioni a valori in $\mathbb{R}$ o $\mathbb{C}$ considera qualunque spazio di funzioni a valori in un qualunque spazio metrico Y.

Consideriamo da questo momento due spazi metrici (X, d_X), (Y, d_Y), introduciamo le nozioni generali di equilimitatezza ed equicontinuità per una famiglia di funzioni $\mathscr{F}$ definite su X a valori in Y.

Definizione 3.6.6. Diremo che la famiglia $\mathscr{F}$ è **equilimitata** se e solo se $\mathscr{F}$ è d_∞-limitato e per ogni $x \in X$ l'orbita

$$\mathscr{O}(x) = \{f(x) \mid f \in \mathscr{F}\}$$

è totalmente limitato in Y.

Osservazione. Se $Y = \mathbb{R}^n$ allora $\mathscr{F}$ è equilimitata se e solo se esiste una costante $L > 0$ tale che $|f(x)| \leq L$ per ogni $x \in X$ e per ogni $f \in f$, come nella definizione usuale di equilimitatezza.

Definizione 3.6.7. Diremo che la famiglia $\mathscr{F}$ è **equicontinua** se e solo se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $x, y \in X$ con $d_X(x, y) < \delta$ e per ogni $f \in \mathscr{F}$ si ha $d_Y[f(x), f(y)] < \varepsilon$.

Osservazione. Chiaramente se $\mathscr{F}$ è equicontinua allora tutti i suoi elementi sono funzioni continue su X, perciò $\mathscr{F} \subseteq C(X, Y)$.

Teorema 3.6.8 (Ascoli-Arzelà)**.** *Se X è totalmente limitato allora ogni insieme equilimitato ed equicontinuo in $C(X, Y)$ è totalmente limitato.*

Viceversa se X è anche completo allora i sottoinsiemi totalmente limitati di $C(X, Y)$ sono tutti e soli quelli equilimitati ed equicontinui.

Dimostrazione. Definiamo gli $x_{i,k}$, Σ_k e Σ come nella dimostrazione della proposizione 3.5.22, in particolare Σ è denso in X e ne scegliamo una enumerazione $x_1, x_2, x_3, \ldots$.

Consideriamo adesso una successione $f_n \in \mathscr{F}$, allora la successione di reali $f_n(x_1)$ è relativamente sequenzialmente compatta in n e quindi esiste un'estratta f_n^1 tale che $f_n^1(x_1)$ è di Cauchy in Y. Anche $f_n^1(x_2)$ è totalmente limitata dunque esiste un'estratta f_n^2 di f_n^1 tale che $f_n^2(x_2)$ è di Cauchy in Y, chiaramente $f_n^2(x_1)$ sarà ancora una successione di Cauchy in quanto è un'estratta di $f_n^1(x_1)$.

Continuando in questo modo otteniamo per ogni $k \in \mathbb{N}$ un'estratta f_n^k di f_n^{k-1} tale che $f_n^k(x_i)$ è di Cauchy se $i \leq k$, definiamo perciò la seguente successione di funzioni:

$$g_k(x) = f_k^k(x)$$

chiaramente g_k sarà un'estratta di f_k e inoltre per ogni $x_i \in \Sigma$ e per ogni $k \geq i$ avremo che $g_k(x_i)$ al variare di $k \geq i$ è un'estratta della successione $f_k^i(x_i)$ e quindi $g_k(x_i)$ è una successione di Cauchy per ogni i.

Dimostriamo adesso che g_k è di Cauchy in $F(X,Y)$. Per la proprietà di equicontinuità avremo che per ogni $\varepsilon > 0$ esiste $K \in \mathbb{N}$ tale che per ogni $k \geq K$, per ogni $n \in \mathbb{N}$ e per ogni $x \in X$ esiste $x_{i,k} \in \Sigma_k$ tale che

$$d_Y\left[g_n(x), g_n\left(x_{i,k}\right)\right] < \epsilon$$

in quanto $d_X(x, x_{i,k}) < 1/k$. D'altronde essendo Σ_k finito esiste $N(k) \in \mathbb{N}$ tale che per ogni $x_{i,k} \in \Sigma_k$ e per ogni $m,n > N(k)$ vale

$$d_Y\left[g_n\left(x_{i,k}\right), g_m\left(x_{i,k}\right)\right] < \epsilon$$

Combinando tutto assieme per ogni $\varepsilon > 0$ esiste $K \in \mathbb{N}$ tale che per ogni $m,n > N(K)$ e per ogni $x \in X$ si ha

$$\begin{aligned} d_Y\left[g_m(x), g_n(x)\right] \leq d_Y\left[g_m(x), g_m\left(x_{i,K}\right)\right] &+ d_Y\left[g_m\left(x_{i,K}\right), g_n\left(x_{i,K}\right)\right] \\ &+ d_Y\left[g_n\left(x_{i,K}\right), g_n(x)\right] \leq 3\epsilon \end{aligned}$$

per un opportuno $x_{i,K} \in \Sigma_k$. Osserviamo che $N(K)$ non dipende in alcun modo dalla scelta di x, ma solamente da ε, dunque avremo che $d_\infty(g_m, g_n) \leq 3\varepsilon$ e quindi g_n è una successione di Cauchy.

Viceversa supponiamo che X sia uno spazio totalmente limitato e completo, ovvero compatto, e sia $\mathscr{F}$ un sottoinsieme totalmente limitato dello spazio metrico $C(X,Y)$. Allora per ogni $\varepsilon > 0$ esistono $f_1, f_2, \dots, f_k \in \mathscr{F}$ tali che per ogni $f \in \mathscr{F}$ si ha

$$f \in \bigcup_{i=1}^{k} B\left(f_i, \varepsilon, d_\infty\right)$$

e quindi per ogni $x \in X$ segue che $f(x) \in \bigcup_{i=1}^k B\left(f_i(x), \varepsilon, d_Y\right)$ e quindi $\mathscr{F}$ è equilimitato in $C(X,Y)$.

Per la proposizione 3.2.2 ogni $f \in \mathscr{F}$ è uniformemente continua, quindi esisterà $\delta > 0$ tale che $d_Y[f_i(x), f_i(y)] < \varepsilon$ per ogni $x, y \in X$ tali che $d_X(x,y) < \delta$ e per ogni $1 \leq i \leq k$, ma allora per ogni $f \in \mathscr{F}$ esisterà un i tale che per ogni $x, y \in X$ con $d_X(x,y) < \delta$

$$d_Y[f(x), f(y)] \leq d_Y[f(x), f_i(x)] + d_Y[f_i(x), f_i(y)] + d_Y[f_i(y), f(y)] < 3\varepsilon$$

e quindi $\mathscr{F}$ è anche equicontinua. ■

Prima di concludere l'argomento concentriamoci sul problema dell'estendibilità delle funzioni continue. Il lemma 3.3.8 ci permette di estendere le funzioni uniformemente continue definite su un sottoinsieme denso sull'intero spazio preservandone l'uniforme continuità e tale estensione d'altronde è anche unica.

Se però adesso consideriamo solamente le funzioni a valori in $\mathbb{R}$ possiamo indebolire enormemente le ipotesi del risultato richiedendo solamente una funzione continua definita su un sottoinsieme chiuso, perdendo però l'unicità dell'estensione. La dimostrazione di questo risultato si basa fondamentalmente sul seguente risultato:

Proposizione 3.6.9. *Se X è uno spazio metrico allora per ogni coppia di insiemi non vuoti chiusi A e B tali che $A \cap B = \emptyset$ esiste sempre una funzione continua $f : X \to [0,1]$ che vale 0 su A e 1 su B.*

Dimostrazione. Si veda (3.5.1). ■

Grazie ad esso possiamo dimostrare prima il seguente risultato preliminare:

Lemma 3.6.10. *Preso uno spazio metrico (X,d), un sottoinsieme chiuso C non vuoto ed una funzione continua e limitata $f : C \to \mathbb{R}$ con $R = \sup_{x\in C}\left|f(x)\right|$ esiste sempre una funzione continua $g : X \to \mathbb{R}$ tale che $\left|g(x)\right| \leq R/3$ per ogni $x \in X$ ed $\left|f(y) - g(y)\right| \leq 2R/3$ per ogni $y \in C$.*

Dimostrazione. Possiamo supporre che f non sia identicamente nulla e perciò $R > 0$. Definiamo innanzitutto

$$C_1 = \left\{x \in C \,\middle|\, f(x) \leq -\frac{R}{3}\right\}$$
$$C_2 = \left\{x \in C \,\middle|\, f(x) \geq \frac{R}{3}\right\}$$

poiché C è chiuso in X anche C_1 e C_2 lo saranno, quindi per la (3.5.1) esisterà una funzione continua $g' : X \to [0,1]$ che vale 0 su C_1 e 1 su C_2. Posto per ogni $x \in X$

$$g(x) = \frac{2R}{3} g'(x) - \frac{R}{3}$$

avremo che $\left|g(x)\right| \leq R/3$ e vale $-R/3$ su C_1 ed $R/3$ su C_2. Di conseguenza per ogni $y \in C$ si avrà

- Se $y \in C_1$ allora
$$\left|f(y) - g(y)\right| = -\frac{R}{3} - f(y) \leq \frac{2R}{3}$$
in quanto $-f(y) \leq R$;
- Se $y \in C_2$ allora
$$\left|f(y) - g(y)\right| = f(y) - \frac{R}{3} \leq \frac{2R}{3}$$
sempre perché $f(y) \leq R$;
- Infine se $y \in C \setminus (C_1 \cup C_2)$ allora $\left|f(y)\right| < R/3$ e dunque
$$\left|f(y) - g(y)\right| \leq \left|f(y)\right| + \left|g(y)\right| \leq \frac{2R}{3}$$

L'asserto è così dimostrato. ■

Dimostriamo ora il teorema di estensione vero e proprio.

Teorema 3.6.11 (estensione di Tietze)**.** *dato uno spazio metrico* (X,d)*, un suo sottoinsieme chiuso non vuoto* C *ed una funzione continua* $f : C \to \mathbb{R}$*. Allora esiste una funzione continua* $G : X \to \mathbb{R}$ *che coincide con* f *su* C*.*

Inoltre se $\left|f(x)\right| \leq R$ *su* C *allora* $|G(x)| \leq R$ *su tutto* X*, mentre se* $\left|f(x)\right| < R$ *su* C *allora* $|G(x)| < R$ *su tutto* X*.*

Dimostrazione. Procediamo per gradi, supponendo innanzitutto che f sia limitata e non identicamente nulla con $R = \sup_{x\in C}\left|f(x)\right| > 0$. Per il lemma 3.6.10 esisterà una funzione continua $g_1 : X \to \mathbb{R}$ tale che $\left|g_1(x)\right| \leq R/3$ su X ed $\left|f(y) - g_1(y)\right| \leq 2R/3$ su C.

D'altronde possiamo applicare il lemma 3.6.10 anche alla funzione $f - g_1$, e dunque esisterà un'altra funzione $g_2 : X \to \mathbb{R}$ tale che $\left|g_2(x)\right| \leq 2R/9$ su X ed $\left|f(y) - g_1(y) - g_2(y)\right| \leq 4R/9$ su C. Procedendo per induzione per ogni $n \in \mathbb{N}$ esisterà una funzione $g_n : X \to \mathbb{R}$ tale che

$$\left|g_n(x)\right| \leq \frac{R}{3}\left(\frac{2}{3}\right)^{n-1}$$
$$\left|f(y) - \sum_{k=1}^{n} g_k(y)\right| \leq R\left(\frac{2}{3}\right)^{n}$$

Definiamo $G_n(x) = \sum_{k=1}^{n} g_k(x)$ per ogni $x \in X$, osserviamo che G_n è una successione di Cauchy in $C(X,\mathbb{R})$ in quanto

$$\left|G_{n+m}(x) - G_n(x)\right| = \left|\sum_{k=n+1}^{n+m} g_k(x)\right| \leq \frac{R}{3}\left(\frac{2}{3}\right)^{n-1}\sum_{k=1}^{m}\left(\frac{2}{3}\right)^{k-1} \leq R\left(\frac{2}{3}\right)^{n-1}$$

per ogni $m,n \in \mathbb{N}$. Poiché $\mathbb{R}$ è completo anche $C(X,\mathbb{R})$ lo è dunque G_n converge uniformemente a qualche funzione continua $G : X \to \mathbb{R}$ con $|G(x)| \leq R$ su tutto X.

Infine poiché $d_\infty\left(f, G_n\right) \leq R(2/3)^n$ su C passando al limite in n avremo che $f = G$ su C come volevasi dimostrare.

Prima di passare al caso non limitato dimostriamo un passaggio intermedio, ovvero quando $\left|f(y)\right| < R$ per ogni $y \in C$. Definiamo $G : X \to \mathbb{R}$ come prima, allora $|G(x)| \leq R$ per ogni $x \in R$ e l'insieme

$$D = \{x \in X \mid \ |G(x)| = R\}$$

è chiuso e disgiunto da C in quanto su di esso abbiamo che $G(y) = f(y)$. Se allora $D \neq \emptyset$ possiamo applicare la (3.5.1) in modo da determinare $\phi : X \to \left[0,1\right]$ che vale 0 su D e vale 1 su C. In questo modo la nuova funzione $\phi(x)G(x)$ coincide ancora con f su C e inoltre $\left|\phi(x)G(x)\right|$ non assumerà mai il valore R su X, dunque $\left|\phi(x)G(x)\right| < R$ per ogni $x \in X$.

Consideriamo infine il caso generale, ovvero quando f è illimitata su C. Posto

$$g(y) = \arctan[f(y)]$$

avremo che $\left|g(y)\right| < \pi/2$ su C ed è ancora continua, per quanto detto poso fa possiamo estenderla su tutto X ad una funzione G' per cui $\left|G'(x)\right| < \pi/2$. Dato che G non assume mai i valori $\pi/2$ e $-\pi/2$ possiamo porre $G(x) = \tan\left[G'(x)\right]$, che di conseguenza sarà una funzione continua su X che coincide con f su C. ■

Dalla (3.5.1) discende un altro risultato che risulta sorprendente in tutta la sua semplicità:

Proposizione 3.6.12. *Se* (X, d) *è uno spazio metrico allora la sua topologia coincide con la topologia iniziale indotta da tutte le funzioni continue da* X *in* $\mathbb{R}$.

Dimostrazione. Chiaramente gli aperti di tale topologia iniziale, che indicheremo con τ, sono aperti anche rispetto alla topologia di X. dobbiamo quindi verificare solamente che gli aperti di X sono aperti anche rispetto alla topologia τ, ovvero per ogni aperto U di X rispetto alla metrica d e per ogni $x \in U$ esiste $V \in \tau$ tale che $x \in V \subseteq U$.

Poiché $X \setminus U$ è chiuso per la (3.5.1) esiste una funzione continua ϕ da X in $[0, 1]$ che vare 1 in x e 0 all'infuori di U. Per come abbiamo definito la topologia iniziale avremo che ϕ è continua anche rispetto a τ e dunque basta porre

$$V = \left\{ y \in X \,\middle|\, \phi(y) > \frac{1}{2} \right\}$$

il quale è chiaramente un aperto di τ, contiene x ed è contenuto in U. Per l'arbitrarietà di x segue che $U \in \tau$ e le due topologie coincidono. ∎

Capitolo 4

Spazi vettoriali topologici

4.1 Prime proprietà

Definizione 4.1.1. Sia $(V, \mathbb{K}, +, \cdot)$ uno spazio vettoriale con $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ e τ una topologia su V. Allora $(V, \mathbb{K}, \tau, +, \cdot)$ è uno **spazio vettoriale topologico** se e solo se le seguenti applicazioni

$$\pi_1 : (x, y) \in V \times V \to x + y \in V$$
$$\pi_2 : (x, \alpha) \in V \times \mathbb{K} \to \alpha x \in V$$

Si ricorda che su $\mathbb{R}$ e su $\mathbb{C}$ si considera sempre la topologia usuale.

Proposizione 4.1.2. *Se V è uno spazio vettoriale topologico allora per ogni $x \in V$ e $t \in \mathbb{K} \setminus \{0\}$ le applicazioni*

$$T_x : v \in V \to v + x \in V$$
$$P_t : v \in V \to tv \in V$$

sono omeomorfismi.

Dimostrazione. Basta dimostrare che sono continue in quanto

$$(T_x)^{-1} = T_{-x}$$
$$(P_t)^{-1} = P_{\frac{1}{t}}$$

Fissiamo un generico $x \in V$ e aperto A di V, per ogni $v \in (T_x)^{-1}(A)$ dalla continuità della esistono due aperti U, W di V tali che

$$(v, x) \in U \times W \subseteq \pi_1^{-1}(A)$$

in quanto $T_x v = \pi_1(v, x)$ e quindi $v \in U \subseteq T_x^{-1}(A)$ e perciò T_x è continua.
Vale un ragionamento analogo con P_t. ■

Dati dei sottoinsiemi $A, B, C, \ldots$ di uno spazio vettoriale V, dei vettori $u, v, \ldots \in V$ e delle costanti $\alpha, \beta, \ldots \in \mathbb{K}$ definiamo $A + B$, $A - B$ e αA come nel capitolo introduttivo.

Proposizione 4.1.3. *Per ogni $A \subseteq V$ e per ogni $x \in V$ avremo che A è aperto se e solo se $v + A$ è aperto.*

Dimostrazione. Segue immediatamente dalla proposizione 4.1.2, difatti $T_v(x) = v + x$ è un omeomorfismo di V in sé e quindi manda aperti in aperti. ■

Da questo risultato segue che i sistemi fondamentali di intorni di un qualunque punto di uno spazio vettoriale topologico è ottenuto traslando un opportuno sistema fondamentale di intorni dell'origine. In particolare uno spazio vettoriale topologico soddisfa l'assioma $\mathscr{N}_1$ se e solo se possiede un sistema fondamentale di intorni numerabile nell'origine.

Teorema 4.1.4. *Dati uno spazio vettoriale X, una famiglia di spazi vettoriali topologici $\mathscr{Y} = \{(Y_a, \tau_a)\}_{a\in A}$ ed una famiglia di applicazioni lineari $f_a : X \to Y_a$. Allora X munito della topologia iniziale generata dalle f_a è ancora uno spazio vettoriale topologico.*

Dimostrazione. Sappiamo che la somma in X è continua se e solo se per ogni $a \in A$ l'applicazione

$$(x, y) \in X \times X \to f_a(x + y) \in Y_a$$

è continua. Dalle proprietà della topologia iniziale le due applicazioni $x \to f_a(x)$ ed $y \to f_a(y)$ sono continue, mentre per linearità $f_a(x + y) = f_a(x) + f_a(y)$ ed essendo Y_a spazio vettoriale topologico la somma in esso è continua. Mettendo tutto assieme avremo così che anche in X la somma è continua.

Con un procedimento analogo si verifica la continuità del prodotto esterno. ■

Corollario 4.1.5. *Data una famiglia di spazi vettoriali topologici $\{Y_a\}_{a\in A}$ allora $\prod_{a\in A} Y_a$ e $\bigoplus_{a\in A} Y_a$ sono spazi vettoriali topologici rispetto alla topologia prodotto e le proiezioni standard π_a sono continue.*

Osservazione. Si tenga però presente che anche se uno spazio vettoriale X fosse somma diretta di una famiglia di suoi sottospazi Y_a non vuol dire assolutamente che le rispettive proiezioni siano continue né che la topologia di X coincida con la topologia prodotto dei suoi sottospazi.

In generale infatti l'isomorfismo tra X ed $\bigoplus_a Y_a$ potrebbe non essere continuo, per garantire la continuità è infatti necessario supporre ulteriori condizioni sulla topologia di X.

Lemma 4.1.6. *Il il quoziente di uno spazio vettoriale topologico (se ancora è uno spazio vettoriale) è ancora uno spazio vettoriale topologico e l'applicazione $\pi_\sim$ definita in* (1.7.1) *è aperta.*

Dimostrazione. Consideriamo un qualunque spazio vettoriale topologico V e una qualunque relazione di equivalenza ben formata $\sim$, verifichiamo che anche $V/\sim$ è uno spazio vettoriale topologico con la topologia quoziente. Innanzitutto verifichiamo che $\pi_\sim$ è aperta oltre che continua. Preso allora $U \subseteq V$ aperto generico abbiamo chiaramente che

$$\pi_\sim^{-1}\left[\pi_\sim(U)\right] = \bigcup_{x\in U}(x + [0]_\sim) = \bigcup_{y\in[0]_\sim}(y + U)$$

ma per la proposizione precedente sappiamo già che $y + U$ è aperto per ogni y e dunque $\pi_\sim(U)$ è aperto in $V/\sim$.

Prendiamo adesso $x, y \in V$ generici ed un aperto U di $V/\sim$ tale che $[x]_\sim + [y]_\sim \in U$, per linearità avremo che $x + y \in \pi_\sim^{-1}(U)$ e quindi esisteranno $A, B \subseteq V$ aperti tali che $x \in A$, $y \in B$ e $A + B \subseteq \pi_\sim^{-1}(U)$.

Per la suriettività di $\pi_\sim$ avremo che $\pi_\sim(A) + \pi_\sim(B) \subseteq U$ e gli insiemi $pi_\sim(A)$, $\pi_\sim(B)$ sono aperti in $V/\sim$, dunque la somma è continua su $V/\sim$, con un procedimento analogo si verifica la continuità anche del prodotto esterno. ■

In uno spazio vettoriale topologico i sottospazi potrebbero essere chiusi o potrebbero non esserlo, però come anche nel caso finito dimensionale non esisteranno sottospazi vettoriali aperti diversi dall'intero spazio. Consideriamo infatti uno spazio vettoriale topologico V su $\mathbb{K}$ con $W \leq V$ sottospazio aperto. Poiché il prodotto esterno è continuo per ogni $x \in V$ avremo che l'insieme

$$U = \{\alpha \in \mathbb{K} \mid \alpha x \in W\}$$

è un aperto di $\mathbb{K}$ contenente lo 0, dunque possiederà un elemento non nullo $\alpha \in U$ e $\alpha x \in W$. Ciò implica immediatamente che $x = (\alpha x)/\alpha \in W$ per ogni $x \in V$ dunque $W = V$.

Definizione 4.1.7. Dato uno spazio vettoriale V su $\mathbb{K}$ diremo che un suo sottoinsieme non vuoto G è **simmetrico** se e solo se $G = -G$, mentre diremo che è **bilanciato** se e solo se per ogni $v \in G$ si ha $\lambda v \in G$ per ogni $|\lambda| \leq 1$.

Se G è un sottoinsieme bilanciato di V allora per ogni $\lambda \in \mathbb{K}$ si avrà

$$\lambda G = |\lambda| G$$

in quanto per ogni $\lambda \neq 0$

$$\lambda G = |\lambda| \frac{\lambda}{|\lambda|} G \subseteq |\lambda| G = \lambda \frac{|\lambda|}{\lambda} G \subseteq \lambda G$$

Proposizione 4.1.8. *In uno spazio vettoriale topologico per ogni aperto G contenente l'origine esiste un aperto U bilanciato contenente l'origine tale che $U + U \subseteq G$.*

Dimostrazione. Dato che $0 + 0 = 0$ per la continuità della somma esistono U_1, U_2 aperti contenenti l'origine tali che $U_1 + U_2 \subseteq G$ e quindi poniamo $U' = U_1 \cap U_2$ che sarà ancora aperto.

Posto

$$U = \left\{u \in U' \,\middle|\, \exists W \text{ aperto contenente } u \text{ tale che } \lambda W \subseteq U' \ \forall |\lambda| \leq 1\right\}$$

dobbiamo solo dimostrare che U contiene l'origine in quanto le altre proprietà discendono dal fatto che P_λ è un omeomorfismo su V.

Dalla continuità del prodotto esterno esistono un aperto L contenente l'origine e una costante $c > 0$ tali che per ogni $u \in L$ e $\left|\mu\right| \leq c$ si ha $\mu u \in U'$ e quindi $\lambda(cL) \subseteq U'$ per ogni $|\lambda| \leq 1$ e quindi $0 \in U$. ∎

Nonostante la semplicità di tale proposizione le sue conseguenze sono molto importanti nello studio degli spazi vettoriali topologici in quanto permette di affermare che ogni spazio vettoriale topologico possiede una struttura uniforme che dipende esclusivamente dalla sua topologia. In generale infatti è la struttura uniforme a generare la topologia, e quindi diverse strutture uniformi possono indurre diverse topologie. Nel caso degli spazi vettoriali topologici tra tutte queste strutture uniformi ne possiamo sempre scegliere una che dipende solo dalla topologia, e quindi se due spazi hanno la medesima topologia avranno anche la stessa struttura uniforme.

Procediamo per gradi ora. Prendiamo un qualunque sottoinsieme U di uno spazio vettoriale topologico X contenente l'origine e poniamo

$$\mathscr{P}(U) = \left\{(x, y) \in X \times X \,\middle|\, x - y \in U\right\}$$

il lettore può velocemente dimostrare le seguenti relazioni

$$\begin{gathered}0 \in U \Leftrightarrow (x,x) \in \mathscr{P}(U)\ \forall x\\ \mathscr{P}(U \cap V) = \mathscr{P}(U) \cap \mathscr{P}(V)\\ \mathscr{P}(U + V) = \mathscr{P}(U) \diamond \mathscr{P}(V)\\ \mathscr{P}(-U) = \mathscr{P}(U)^{-1}\\ \mathscr{P}(U)[A] = U + A\ \forall A \subseteq X\end{gathered}$$

Adesso se prendiamo stavolta $U \subseteq X$ aperto bilanciato contenente l'origine in virtù della proposizione 4.1.8 e della proposizione 3.1.4 la famiglia

$$\mathscr{B} = \left\{\mathscr{P}(U) \;\middle|\; U \text{ intorno aperto bilanciato dell'origine}\right\}$$

genera su X una struttura uniforme che possiede $\mathscr{P}(U)$ come base uniforme.

Dato che le traslazioni sono omeomorfismi gli insiemi nella forma $\mathscr{P}(U)[x] = U+x$ sono aperti quando U è un intorno aperto bilanciato dell'origine, quindi la topologia indotta dalla struttura uniforme è meno fine (ha meno aperti) della topologia di X. Viceversa preso un qualunque aperto O di X contenente x avremo che $O - x$ è un aperto contenente l'origine, la proposizione 4.1.8 ci dice che esiste un aperto bilanciato U contenente l'origine tale che $U \subseteq O - x$ ovvero

$$O \supseteq U + x = \mathscr{P}(U)[x]$$

per l'arbitrarietà di x segue così che O è aperto anche nella topologia indotta dalla base uniforme, dunque le due topologie coincidono. In particolare gli spazi vettoriali topologici sono completamente regolari e sono di Hausdorff se e solo se i punti sono chiusi.

Questa struttura uniforme è generata dalla topologia del nostro spazio vettoriale topologico X, però potrebbero esistere altre strutture uniformi su X che inducono la stessa topologia su X (per esempio gli spazi con una topologia iniziale, che come abbiamo visto precedentemente possiedono una struttura uniforme indotta dalle applicazioni che la generano). Se però questa struttura uniforme possiede una base "invariante per traslazioni" allora essa dovrà necessariamente coincidere con la struttura uniforme generata dagli insiemi nella forma $\mathscr{P}(U)$. Più in generale si ha

Proposizione 4.1.9. *Sia V spazio vettoriale dotato di due strutture uniformi $\mathscr{T}_1, \mathscr{T}_2$ che possiedono rispettivamente le semibasi uniformi $\mathscr{B}_1, \mathscr{B}_2$ composte interamente da insiemi U che soddisfano la seguente proprietà:*

$$(x,y) \in U \Leftrightarrow (x+z, y+z) \in U$$

qualunque siano $x,y,z \in U$.

Se $\mathscr{T}_1$ e $\mathscr{T}_2$ inducono su X la medesima topologia allora $\mathscr{T}_1 = \mathscr{T}_2$.

Dimostrazione. Per ogni $B \in \mathscr{B}_1$ esistono $A_1, A_2, \dots, A_n \in \mathscr{B}_2$ tali che $\left(A_1 \cap A_2 \cap \cdots A_n\right)[0] \subseteq B[0]$ dato che generano la stessa topologia. Questo significa che se $(x,0) \in A_i$ per ogni $1 \leq i \leq n$ allora $(x,0) \in B$.

Siano ora $x, y \in V$ tali che $(x,y) \in A_i$ per ogni i, per l'invarianza abbiamo $(x-y,0) \in A_i$ per ogni i e così $(x-y,0) \in B$ ovvero $(x,y) \in B$. Abbiamo così dimostrato che per ogni $B \in \mathscr{B}_1$ esistono $A_1, A_2, \dots, A_n \in \mathscr{B}_2$ tali che $\bigcap_{i=1}^n A_i \subseteq B$ e così $\mathscr{T}_1 \subseteq \mathscr{T}_2$. Per lo stesso ragionamento si ha anche l'inclusione opposta e così le due strutture uniformi coincidono. ∎

Corollario 4.1.10. *Nelle stesse ipotesi del teorema 4.1.4 la struttura uniforme di X generata dalla sua topologia di spazio vettoriale topologico coincide con la struttura uniforme generata dalle funzioni f_a.*

Dimostrazione. Una semibase uniforme generata dalle f_a è data dagli insiemi nella forma

$$\left\{(x,y) \in X \,\middle|\, f_a(x) - f_a(y) \in U_a\right\}$$

al variare di $U_a \subseteq Y_a$ aperto bilanciato contenente l'origine. Per linearità si ha $f_a(x+z)-f_a(y+z) = f_a(x) - f_a(y)$ e dunque siamo esattamente nelle ipotesi della proposizione 4.1.9. ■

Teorema 4.1.11. *Se V è uno spazio vettoriale topologico allora valgono le seguenti affermazioni:*

1. *Se U è un aperto di V e A è un generico sottoinsieme non vuoto di V allora $U+A$ è aperto in V;*

2. *Se C è chiuso e K compatto con $C \cap K = \emptyset$ allora esiste un aperto U contenente l'origine tale che $(U+C) \cap (U+K) = \emptyset$;*

3. *Se C è chiuso e mentre K è compatto in V allora $C+K$ è chiuso in V.*

4. *Se K e H sono compatti in V allora anche $K+H$ è compatto.*

5. *Dati due sottoinsiemi A e B di V si ha* $\operatorname{Int}(A) + \operatorname{Int}(B) \subseteq \operatorname{Int}(A+B)$.

Dimostrazione. Innanzitutto il punto 1 segue dall'uguaglianza $U+A = \bigcup_{x\in A} T_x(U)$. Il punto 2 invece è una diretta conseguenza della proposizione 3.1.6 negli spazi uniformi.

Per dimostrare il punto 3 per ogni $x \notin C+K$ si ha $K \cap (x-C) = \emptyset$. Possiamo applicare il punto 2 e trovare U aperto contenente l'origine tale che $(K+U) \cap (x-C+U) = \emptyset$ e dunque $(x+U) \cap (K+C+U) = \emptyset$. Dato che $0 \in U$ l'ultima uguaglianza implica che esiste un aperto contenente x che non interseca $K+C$, per l'arbitrarietà di x segue perciò che $K+C$ è chiuso.

Supponiamo ora che K, H sono sottoinsiemi compatti di V, allora $K \times H$ è compatto in $V \times V$ (sia con la topologia prodotto che quella di sottospazio dato che coincidono) e quindi per la continuità della somma $K+H$ è compatto in V. L'ultimo punto discende dal punto 1 e dalla definizione di interno di un insieme. ■

Proposizione 4.1.12. *Se V è uno spazio vettoriale topologico e $A,B \subseteq V$ sottoinsiemi non vuoti, allora avremo che*

1. $\operatorname{Cl}(A) = \displaystyle\bigcap_{\substack{U \text{ aperto} \\ 0\in U}} (A+U)$

2. $\operatorname{Cl}(A) + \operatorname{Cl}(B) \subseteq \operatorname{Cl}(A+B)$

Dimostrazione. Il primo punto discende dalla proposizione 3.1.7 in quanto vale l'uguaglianza $A+U = \operatorname{Cl}_{\mathscr{P}(U)}(A)$.

Siano $a \in \operatorname{Cl}(A)$, $b \in \operatorname{Cl}(B)$ e U aperto contenente $a+b$, per la continuità della somma esistono due aperti U_1, U_2 contenenti rispettivamente a e b tali che $U_1 + U_2 \subseteq U$. Essendo $A \cap U_1 \neq \emptyset$ e $B \cap U_2 \neq \emptyset$ allora $\emptyset \neq (A+B) \cap (U_1+U_2) \subseteq (A+B) \cap U$ e quindi $a+b \in \operatorname{Cl}(A+B)$. ■

Corollario 4.1.13. *Siano $\alpha_1, \alpha_2, \dots, \alpha_n$ degli scalari in $\mathbb{K}$ e*

$$\mathscr{J} = \{\emptyset\} \cup \left\{A \subseteq V \,\middle|\, \alpha_1 A + \alpha_2 A + \cdots + \alpha_n A \subseteq A\right\}$$

Allora per ogni $A \in \mathscr{J}$ si ha $\mathrm{Cl}(A) \in \mathscr{J}$. Ancora dato $A \in \mathscr{J}$ se almeno uno degli α_i è diverso da 0 oppure se $0 \in \mathrm{Int}(A)$ allora avremo anche $\mathrm{Int}(A) \in \mathscr{J}$.

Inoltre se $\mathscr{F} \subseteq \mathscr{J}$ è non vuota allora $\bigcap_{F \in \mathscr{F}} F \in \mathscr{J}$ e se $\mathscr{F}$ è un insieme diretto rispetto all'inclusione allora $\bigcup_{A \in \mathscr{F}} A \in \mathscr{J}$.

Dimostrazione. Possiamo supporre tutti gli α_i diversi da 0, infatti se sono tutti nulli allora $A \in \mathscr{J}$ se e solo se $0 \in A$ altrimenti basta semplicemente eliminare gli scalari nulli.

Ricordiamo che P_{α_i} è un omeomorfismo per ogni i e quindi $\alpha_i \mathrm{Cl}(A) = \mathrm{Cl}(\alpha_i A)$, quindi se $A \in \mathscr{A}$ applicando il secondo punto della proposizione precedente si ha $\mathrm{Cl}(A) \in \mathscr{I}$.

Se invece gli α_i non sono tutti nulli allora per quanto detto sopra possiamo porre $\alpha_i \neq 0$ per ogni i e quindi

$$\alpha_1 \mathrm{Int}(A) + \cdots + \alpha_n \mathrm{Int}(A) \subseteq \alpha_1 A + \cdots + \alpha_n A \subseteq A$$

se $A \in \mathscr{J}$. Ma la somma di insiemi aperti è ancora aperta quindi $\mathrm{Int}(A) \in \mathscr{J}$.

Che l'intersezione di elementi di $\mathscr{J}$ appartenga al medesimo insieme è banale da verificare, dobbiamo solo provare che se $\mathscr{F}$ è un insieme diretto contenuto in $\mathscr{J}$ allora $F = \bigcup_{A \in \mathscr{F}} A \in \mathscr{J}$.

Presi $a_1, a_2, \dots, a_n \in F$ esisteranno $A_1, A_2, \dots, A_n \in \mathscr{F}$ tali che $a_i \in A_i$, ma esisterà $B \in \mathscr{F}$ tale che $A_1 \cup A_2 \subseteq B$ e in particolare $a_1, a_2 \in B$, ancora esisterà $C \in \mathscr{F}$ tale che $a_1, a_2, a_3 \in C$ sempre per lo stesso motivo e così via. Alla fine esisterà un $F' \subseteq \mathscr{F}$ tale che $a_i \in F'$ per ogni i, e poiché $F' \in \mathscr{J}$ avremo che $\sum_{i=1}^n \lambda_i \alpha_i \in F' \subseteq F$ e per l'arbitrarietà degli a_i segue così che $F \in \mathscr{J}$. ■

Da questo corollario discendono immediatamente i seguenti risultati per un insieme A con interno non vuoto:

- Se A è simmetrico allora anche $\mathrm{Cl}(A)$ e $\mathrm{Int}(A)$ sono simmetrici;
- Se A è bilanciato allora anche $\mathrm{Cl}(A)$ è bilanciato, inoltre se $0 \in \mathrm{Int}(A)$ allora anche $\mathrm{Int}(A)$ è bilanciato;
- Se A è un sottospazio vettoriale allora anche $\mathrm{Cl}(A)$ e $\mathrm{Int}(A)$ (quest'ultimo solo se contiene l'origine) lo sono.

Abbiamo visto che se uno spazio vettoriale topologico ha i punti chiusi allora è uno spazio di Hausdorff e in particolare abbiamo l'unicità del limite. Vedremo ora che anche se V non è uno spazio di Hausdorff è possibile quozientarlo con un sottospazio vettoriale per renderlo di Hausdorff.

In generale infatti il quoziente di uno spazio topologico di Hausdorff potrebbe non essere uno spazio di Hausdorff, per questo motivo introduciamo il risultato seguente:

Proposizione 4.1.14. *Dato uno spazio vettoriale topologico V e un sottospazio vettoriale chiuso W di V allora V/W sarà uno spazio vettoriale topologico di Hausdorff.*

Dimostrazione. Che V/W sia uno spazio vettoriale topologico è una conseguenza del lemma 4.1.6, dobbiamo solamente verificare che i punti in V/W sono chiusi, da cui per regolarità segue immediatamente che è uno spazio di Hausdorff.

La funzione π_W è aperta ma in genere non è chiusa, Poiché $V \setminus W$ è aperto in V avremo che

$$\pi_W(V \setminus W) = [(V \setminus W) + W]/W = V/W \setminus W/W$$

e quindi $W/W = \{W\} = \left\{[0]_W\right\}$ è chiuso in V/W. ■

Possiamo allora applicare questa proposizione all'insieme $W = \mathrm{Cl}(\{0\})$ il quale sarà ancora un sottospazio vettoriale di V, dunque $V/\mathrm{Cl}(\{0\})$ sarà sempre uno spazio vettoriale topologico di Hausdorff.

A differenza di uno spazio uniforme generico uno spazio vettoriale topologico $\mathscr{N}_1$ è anche uniformemente $\mathscr{N}_1$ in quanto preso un qualunque sistema fondamentale di intorni aperti bilanciati dell'origine U_n avremo che gli insiemi $\mathscr{P}\left(U_n\right)$ formano una base uniforme per X.

4.2 Spazi vettoriali metrici

In questa sezione daremo una definizione di limitatezza per i sottoinsiemi di uno spazio vettoriale topologico. Ricordiamo che gli spazi vettoriali topologici sono anche spazi uniformi ed in certi casi la loro struttura uniforme può essere generata da una metrica. Dato però che in base alla metrica scelta cambiano anche gli insiemi limitati è opportuno dare per gli spazi vettoriali topologici un'altra definizione di limitatezza che si rifà invece alla struttura topologica (e quindi a quella uniforme). Tale definizione di limitatezza è distinta da quella metrica e presenta delle proprietà interessanti.

Definizione 4.2.1. Un sottoinsieme A di uno spazio vettoriale topologico si dice **limitato** se e solo se per ogni intorno I dell'origine esiste una costante $t > 0$ tale che $A \subseteq \lambda I$ per ogni $\lambda \in \mathbb{K}$ con $|\lambda| > t$.

Osserviamo innanzitutto che i punti sono sempre insiemi limitati. Per ogni $x \in V$ infatti l'applicazione $\lambda \in \mathbb{K} \to \lambda x \in V$ è continua in zero, dunque per ogni aperto $U \subseteq V$ contenente l'origine esiste $t > 0$ tale che $\lambda x \in U$ per ogni $\lambda \in \mathbb{K}$ tale che $|\lambda| < t$. Dato che l'unione di un numero finito di insiemi numerabili è numerabile avremo che ogni sottoinsieme finito di uno spazio vettoriale topologico è limitato.

Come per gli spazi uniformi, anche per gli spazi vettoriali topologici possiamo parlare di sottoinsiemi totalmente limitati, e come suggerisce il nome la totale limitatezza è una condizione più forte della limitatezza:

Proposizione 4.2.2. *Gli insiemi totalmente limitati sono anche limitati.*

Dimostrazione. Sia $A \subseteq V$ sottoinsieme totalmente limitato, prendiamo un qualunque intorno aperto dell'origine U, esisterà un altro intorno aperto bilanciato dell'origine W tale che $W + W \subseteq U$.

Per la totale limitatezza di A esiste $B \subseteq A$ finito tale che $A \subseteq B + W$, per la finitezza di B e per la continuità del prodotto scalare esisterà $t > 0$ tale che $(1/t)B \subseteq W$. Poiché W è bilanciato avremo che

$$A \subseteq B + W \subseteq tW + W \subseteq (t \vee 1)W + (t \vee 1)W \subseteq (t \vee 1)U$$

e quindi A è anche limitato. ■

Corollario 4.2.3. *I sottoinsiemi compatti sono limitati.*

Proposizione 4.2.4. *Se A è un insieme limitato allora anche* $\mathrm{Cl}(A)$ *e λA per ogni $\varepsilon \in \mathbb{K}$ lo sono. Inoltre se A e B sono limitati anche $A+B$ lo sarà*

Dimostrazione. Per la proposizione 2.3.2 per ogni aperto U contenente l'origine esiste un altro aperto W, sempre contenente l'origine, tale che $\mathrm{Cl}(W) \subseteq U$. Dalla definizione di limitatezza esiste $t > 0$ tale che $A \subseteq \lambda W$ per ogni $|\lambda| > t$ e quindi $\mathrm{Cl}(A) \subseteq \lambda \mathrm{Cl}(W) \subseteq \lambda U$.

Preso ora $\lambda \in \mathbb{K}$ se $k = 0$ allora $\lambda A = \{0\}$ è chiaramente limitato, se $k \neq 0$ per ogni aperto U contenente l'origine esisterà $t > 0$ tale che $A \subseteq \mu(1/\lambda)U$ per ogni $|\mu| \geq t$ dato che $(1/\lambda)U$ è ancora un aperto, dunque $\lambda A \subseteq \mu U$ e per l'arbitrarietà di U segue la limitatezza di λA.

Dati due insiemi limitati A e B e un qualunque aperto contenente l'origine sappiamo che esiste un intorno aperto dell'origine W tale che $W + W \subseteq U$ e quindi esisteranno $r, t > 0$ tali che $A \subseteq \lambda W$ e $B \subseteq \mu W$ per ogni $|\lambda| \geq r$ e $|\mu| \geq t$. Consideriamo ora un qualunque $v \in \mathbb{K}$ tale che

$$|v| \geq t^* = r + t > 0$$

ma allora posto

$$v_1 = \frac{r}{r+t} v \qquad\qquad v_2 = \frac{t}{r+t} v$$

segue immediatamente che $v = v_1 + v_2$ e $|v_1| \geq r$, $|v_2| \geq t$ e dunque

$$A + B \subseteq v_1 W + v_2 W = vW \subseteq vU$$

ovvero $A + B$ è limitato. ■

Benché la definizione può sembrare semplice, la presenza di aperti limitati fornisce allo spazio molte proprietà molto interessanti, tra tutte la proprietà $\mathcal{N}_1$ come il seguente lemma mostra.

Lemma 4.2.5. *Sia V uno spazio vettoriale topologico allora*

1. *Per ogni successione $r_n > 0$ strettamente crescente tendente all'infinito e per ogni aperto U contenente l'origine si ha*
$$V = \bigcup_{n=1}^{+\infty} r_n U$$

2. *Se $s_n > 0$ è una successione strettamente decrescente tendente a* 0 *allora per ogni aperto limitato U contenente l'origine la famiglia $\{s_n U\}_{n \in \mathbb{N}}$ è un sistema fondamentale di intorni dell'origine.*

Dimostrazione. Il primo punto è facile da verificare: preso $x \in X$ dato che $\{x\}$ è totalmente limitato per quanto detto prima è anche limitato e dunque possiamo scegliere n sufficientemente grande in modo tale che $x \in r_n U$.

Preso ora un qualunque aperto limitato U contenente l'origine e un altro aperto W sempre contenente l'origine, allora la nuova successione $r_n = 1/s_n$ tenderà all'infinito e quindi per definizione di limitatezza esisterà $\bar{n} \in \mathbb{N}$ tale che $U \subseteq 1/s_{\bar{n}} W$ ovvero $s_{\bar{n}} U \subseteq W$. ■

Per quanto riguarda invece i sottospazi limitato dobbiamo ridimensionare le nostre pretese.

Proposizione 4.2.6. *In uno spazio vettoriale topologico V gli unici sottospazi limitati sono quelli contenuti in $T = \mathrm{Cl}(\{0\})$ (T incluso).*

Dimostrazione. La proposizione 4.1.12 afferma che $T = \bigcap U$ al variare di U tra gli aperti contenenti l'origine, in particolare T è limitato. Prendiamo ora un qualunque $W \leq V$ sottospazio limitato, allora $W \leq \lambda U$ per ogni aperto U contenente l'origine e per qualche $\lambda \in \mathbb{K} \setminus \{0\}$, ma allora $W = W/\lambda \subseteq U$ e quindi $W \leq T$. ■

Come per gli spazi uniformi possiamo definire le funzioni uniformemente continue anche per le funzioni tra spazi vettoriali topologici. Introduciamo anche la nozione di uniforme continuità sequenziale che, sebbene può essere introdotta anche per gli spazi uniformi, la definiamo solo ora.

Definizione 4.2.7. Dati due spazi vettoriali topologici V, W, un sottoinsieme non vuoto U di V e una funzione $f : U \to W$ diremo che f è **uniformemente sequenzialmente continua** su U se e solo se per ogni coppia di successioni $x_n, y_n \in U$ tali che $x_n - y_n \to 0$ si ha $f(x_n) - f(y_n) \to 0$ in W.

Proposizione 4.2.8. *Dati due spazi vettoriali topologici V, W, un sottoinsieme U di V e una funzione $f : U \to W$. Se f è sequenzialmente continua e U è sequenzialmente compatto allora f è sequenzialmente uniformemente continua.*

Dimostrazione. Supponiamo U sequenzialmente compatto ed f sequenzialmente continua e prendiamo due successioni x_n, y_n tali che $x_n - y_n \to 0$. Consideriamo adesso la successione $f(x_n) - f(y_n)$ e ne prendiamo una qualunque estratta k_n, per la compattezza sequenziale esiste un'estratta k'_n di k_n tale che $x_{k'_n} \to x$ e $y_{k'_n} \to y$. Ma allora per la continuità della differenza deve essere $x = y$ e per la continuità sequenziale di f segue che $f(x_{k'_n}) - f(y_{k'_n}) \to f(x) - f(y) = 0$. Per la proposizione 2.2.7 tutta la successione $f(x_n) - f(y_n)$ converge a 0 e f è uniformemente sequenzialmente continua. ■

Concentriamoci ora sullo studio delle pseudometriche compatibili con la struttura uniforme di uno spazio vettoriale topologico X. La terna (X, τ, d) composta da uno spazio vettoriale X su $\mathbb{K}$, una topologia τ che rende (X, τ) uno spazio vettoriale topologico ed una pseudometrica d compatibile con la topologia τ è detta **spazio vettoriale pseudometrico**, e se d è una metrica sarà invece detto **spazio vettoriale metrico**. Dato che è possibile trovare diverse metriche che inducono la medesima struttura uniforme ci concentreremo a trovarne alcune che possiedano proprietà particolari.

Definizione 4.2.9. Una pseudometrica d su uno spazio vettoriale V su $\mathbb{K}$ si dice

- **invariante** per traslazioni se e solo se per ogni $x, y, z \in V$ si ha
$$d(x + z, y + z) = d(x, y)$$
- **bilanciata** se e solo se per ogni $r > 0$ e per ogni $\lambda \in \mathbb{K}$ con $|\lambda| \leq 1$ si ha $d(y + \lambda x, y) \leq d(y + x, y)$ per ogni $x, y \in X$.

- **regolare sugli scalari** se e solo se presi $x \in X$, $\lambda \in \mathbb{K}$ e delle successioni $x_n \in X, \lambda_n \in \mathbb{K}$ tali che $d(x_n, x) \to 0$ ed $\lambda_n \to \lambda$ allora $d(\lambda_n x_n, \lambda x) \to 0$.

Osservazione. Se d è invariante avremo che

$$B\big(x, r, d\big) = x + B\big(0, r, d\big) \qquad\qquad D\big(x, r, d\big) = x + D\big(0, r, d\big)$$

e quindi una metrica invariante è bilanciata se e solo se le palle aperte centrate nell'origine sono bilanciate.

Osservazione. Nell'ultimo punto non stiamo dicendo affatto che x_n converge ad x, anche perché V potrebbe non essere dotato di una topologia o potrebbe possederne una che non soddisfi l'assioma $\mathcal{N}_1$. Questa proprietà è infatti propria di della pseudometrica d e non deve dipendere dalla struttura topologica di X.

Se d è una metrica invariante su X è utile introdurre la seguente funzione

$$D : x \in X \to d(x, 0) \in \mathbb{R}$$

che permette di determinare univocamente d in quanto per l'invarianza si ha

$$d(x, y) = d(x - y, 0) = D(x - y)$$

la funzione D, che chiameremo sempre pseudometrica invariante, soddisfa le seguenti proprietà:

$$\begin{gathered} D(0) = 0 \\ D(x) = D(-x) \\ D(x + y) \le D(x) + D(y) \\ D(x) \ge 0 \end{gathered}$$

avremo inoltre che

- D è una metrica se e solo se $D(x) = 0 \Rightarrow x = 0$;
- D è bilanciata se e solo se $D(\lambda x) \le D(x)$ per ogni $\lambda \in \mathbb{K}$ con $|\lambda| \le 1$;
- D è regolare sugli scalari se e solo se per ogni $x_n, x \in X$ tali che $D(x_n - x) \to 0$ e per ogni $\lambda_n, \lambda \in \mathbb{K}$ con $\lambda_n \to \lambda$ si ha $D(\lambda_n x_n - \lambda x) \to 0$.

Consideriamo adesso uno spazio vettoriale V su $\mathbb{K}$ su cui introduciamo una famiglia di pseudometriche invarianti e regolari sugli scalari $\{D_i\}_{i \in I}$, poiché esse sono comunque delle pseudometriche possiamo applicare il teorema 3.2.12 per dimostrare che esse inducono su X una struttura uniforme, inoltre per l'invarianza X è di Hausdorff se e solo se per ogni $x \ne 0$ esiste $i \in I$ tale che $D_i(x) > 0$. Vogliamo dimostrare che X è uno spazio vettoriale topologico la cui struttura uniforme coincide con quella indotta dalle pseudometriche D_i.

Dimostriamo innanzitutto che la somma è continua. Poiché gli insiemi $B\big(x, r, D_i\big) = \big\{y \in X \,\big|\, D_i(y - x) < r\big\}$ formano un sistema essenziale di intorni per x ci possiamo limitare a lavorare con questi piuttosto che con gli aperti generici. Fissiamo $x, y \in X$, $\varepsilon > 0$ ed $i \in I$, per la proprietà di invarianza per ogni $x', y' \in X$ si ha

$$D_i[(x' + y') - (x + y)] = D_i(x' - x) + D_i(y' - y)$$

e quindi per ogni $x' \in B\left(x, \varepsilon/2, D_i\right)$, $y' \in B\left(y, \varepsilon/2, D_i\right)$ avremo che $x' + y' \in B\left(x + y, \varepsilon, D_i\right)$ e quindi la somma è continua.

Per la continuità del prodotto esterno prendiamo delle net $x_a \in X, \lambda_a \in \mathbb{K}$ convergenti rispettivamente ad $x \in X$ e $\lambda \in \mathbb{K}$. Per assurdo $\lambda_a x_a$ non converge ad λx e quindi esistono un $i \in I$, una subnet a_j su J ed $\varepsilon > 0$ tali che $D_i\left(\lambda_{a_j} x_{a_j} - \lambda x\right) \geq \varepsilon$. Dato che $D_i(x_a - x) \to 0$ esistono $j(n) \in J$ al variare di $n \in \mathbb{N}$ tali che $\lambda_{a_{j(n)}} \to \lambda$ ed $D_i\left(x_{a_{j(n)}} - x\right) \to 0$. Per la regolarità sugli scalari abbiamo allora $D_i\left(\lambda_{a_{j(n)}} x_{a_{j(n)}} - \lambda x\right) \to 0$ assurdo, quindi il prodotto esterno è continuo ed X è uno spazio vettoriale topologico.

Ricordiamo che per la struttura uniforme indotta da una famiglia di pseudometriche gli insiemi nella forma

$$P_{\varepsilon,i} = \left\{(x, y) \in X \times X \;\middle|\; D_i(x - y) < \varepsilon\right\}$$

al variare di $\varepsilon > 0$ ed $i \in I$ determinano una semibase uniforme della struttura uniforme indotta dalle pseudometriche D_i. Per la proposizione 4.1.9 la struttura unidorme generata dalle pseudometriche D_i coinciderà con la struttura uniforme generata dalla topologia dello spazio vettoriale topologico X.

Abbiamo così dimostrato il teorema

Teorema 4.2.10. *Sia V uno spazio vettoriale su $\mathbb{K}$ munito della struttura uniforme generata da una qualunque famiglia di pseudometriche invarianti regolari sugli scalari $\{D_\alpha\}_{\alpha \in A}$. Allora V è uno spazio vettoriale topologico la cui struttura uniforme coincide con quella indotta dalle pseudometriche D_α. Inoltre V è di Hausdorff se e solo se per ogni $x \neq 0$ esiste $\alpha \in A$ tale che $D_\alpha(x) > 0$.*

Si noti che la regolarità sugli scalari è strettamente necessaria, altrimenti il prodotto esterno potrebbe non essere continuo.

Il prossimo step è dimostrare che in ogni spazio vettoriale topologico la sua topologia e la sua struttura uniforme sono generati da una famiglia di pseudometriche invarianti, bilanciate e regolari sugli scalari. Per il teorema 3.2.12 la struttura uniforme di uno spazio vettoriale topologico, e quindi la sua topologia, possono essere generate da una famiglia di pseudometriche. Ci interessa però sapere se questa famiglia di pseudometriche può essere scelta in modo tale da soddisfare alcune nostre ipotesi, in particolare vogliamo che essa sia invariante, bilanciata e regolare sugli scalari.

Scegliamo così un aperto $U = U_0$ di uno spazio vettoriale topologico V su $\mathbb{K}$ contenente l'origine, per la proposizione 4.1.8 per ogni $n \in \mathbb{N}$ esiste un aperto bilanciato U_n contenente l'origine tale che $U_n + U_n \subseteq U_{n-1}$. Dato che $\mathscr{P}\left(U_m\right) \diamond \mathscr{P}\left(U_n\right) = \mathscr{P}\left(U_m + U_n\right)$ se poniamo

$$\mathcal{V}(l) = c_1(l)U_1 + c_2(l)U_2 + c_3(l)U_3 + \cdots \qquad \forall\, l \in \mathbb{Q}$$

$$E(x) = \begin{cases} 1 & \text{se } x \notin \mathcal{V}(l) \text{ per ogni } l \in \mathbb{Q} \\ \inf\{l \in \mathbb{Q} \mid x \in \mathcal{V}(l)\} & \text{altrimenti} \end{cases}$$

per ogni $x \in V$, non è difficile verificare che se $D(x, y)$ è la funzione interpolante generata dalla famiglia di entourage $\mathscr{P}\left(U_n\right)$ allora

$$D(x, y) = E(x - y)$$

per ogni $x, y \in V$. Dunque le proprietà della funzione interpolante valgono anche per la nuova funzione E che risulterà dunque essere una pseudometrica invariante. In particolare avremo che

- $E(0) = 0$;
- $E(x) = E(-x)$;
- $E(x + y) \leq E(x) + E(y)$;
- $U_{n+1} \subseteq \{x \in V \mid E(x) < 2^{-n}\} \subseteq U_n$.

e per il teorema 3.2.12 ogni spazio vettoriale topologico può essere generato da una famiglia di tali pseudometriche invarianti. Ci rimane così da dimostrare che E è anche bilanciata e regolare sugli scalari.

Dato che la somma di aperti bilanciati è ancora un aperto bilanciato avremo che gli insiemi $\mathcal{V}(l)$ sono tutti bilanciati, e quindi $E(\lambda x) = E(|\lambda|\, x)$ per ogni $\lambda \in \mathbb{K}$ e la pseudometrica è bilanciata. Per quanto riguarda la regolarità sugli scalari bisogna essere più cauti.

Prendiamo dei valori $x_i, x \in V$, $\lambda_i, \lambda \in \mathbb{K}$ al variare di $i \in \mathbb{N}$ tali che $\lambda_i \to \lambda$ ed $E\left(x_i - x\right) \to 0$. Questo significa in particolare che per ogni $n \in \mathbb{N}$ esiste $i_n \in \mathbb{N}$ tale che $i_{n+1} \geq i_n$ ed $E\left(x_i - x\right) < 2^{-n}$ per ogni $i \geq i_n$, in particolare $x_i - x \in U_n$. Possiamo tra l'altro supporre senza perdere in generalità che $\left|\lambda_i - \lambda\right| < 1$ e dunque essendo gli insiemi U_n bilanciati avremo anche che

$$(\lambda_i - \lambda)(x_i - x) = (\lambda_i x_i - \lambda x) - (\lambda_i x - \lambda x) - (\lambda x_i - \lambda x) \in U_n$$

mentre per la continuità del prodotto esterno possiamo supporre sempre senza perdere in generalità che $\lambda_i x - \lambda x \in U_n$ per ogni $i \geq i_n$.

Adesso essendo $2U_{n+1} \subseteq U_{n+1} + U_{n+1} \subseteq U_n$ per induzione avremo che

$$2^m U_{n+m} \subseteq U_n$$

per ogni $m, n \in \mathbb{N}$, possiamo quindi scegliere m sufficientemente grande in modo tale che $|\lambda| \leq 2^m$. Prendendo così $i \geq i_{m+n+1}$ avremo che

$$\begin{aligned}
\lambda_i x_i - \lambda x &= (\lambda_i - \lambda)(x_i - x) + (\lambda_i x - \lambda x) + (\lambda x_i - \lambda x) \\
&\in U_{n+2} + U_{n+2} + \lambda U_{m+n+1} \\
&= U_{n+2} + U_{n+2} + |\lambda|\, U_{m+n+1} \subseteq U_{n+1} + 2^m U_{n+1+m} \subseteq U_n
\end{aligned}$$

per l'arbitrarietà di n avremo così che $E\left(\lambda_i x_i - \lambda x\right) \to 0$.

Mettendo tutto assieme abbiamo dimostrato il seguente risultato:

Teorema 4.2.11. *Ogni spazio vettoriale topologico V è generato da una famiglia di pseudometriche invarianti, bilanciate e regolari sugli scalari. Inoltre se V soddisfa l'assioma $\mathcal{N}_1$ allora la topologia e la struttura uniforme standard di V possono essere generate da una sola pseudometrica invariante, bilanciata e regolare sugli scalari, che diventa una metrica quando V è uno spazio di Hausdorff (oppure se i punti sono chiusi).*

Corollario 4.2.12. *In uno spazio vettoriale topologico $\mathcal{N}_1$ compattezza e compattezza sequenziale sono equivalenti.*

Dato che la famiglia di pseudometriche genera anche la struttura uniforme del nostro spazio topologico i concetti di Cauchy nets e di totale limitatezza rimangono invariati se consideriamo le pseudometriche al posto degli intorni aperti dell'origine. Invece il concetto di limitatezza negli spazi vettoriali topologici $\mathcal{N}_1$ è completamente diverso da quello indotto da una qualunque pseudometrica invariante d che induce la topologia di V, anzi se $d(x, y) = E(x - y)$ allora tutto V sarebbe d-limitato mentre, a meno di casi degeneri, l'intero spazio non è mai limitato rispetto alla definizione data in questo capitolo.

4.3 Funzioni lineari

In questo capitolo ci concentreremo sulle funzioni lineari e continue tra spazi vettoriali topologici. Nel capitolo introduttivo abbiamo già definito i concetti di linearità e di continuità, prima di proseguire la nostra trattazione dobbiamo introdurre un nuovo concetto utile alla nostra trattazione

Definizione 4.3.1. Siano V, W spazi vettoriali topologici, allora un'applicazione $f : V \to W$ è **limitata** se e solo se per ogni $A \subseteq V$ limitato l'insieme $f(A)$ è limitato in W.

Diremo invece che f è **sequenzialmente limitata** se e solo se per ogni successione x_n convergente in V la successione $f(x_n)$ è limitata in W.

Osservazione. Non si confondano le funzioni limitate con le funzioni la cui immagine è limitata. Per esempio la funzione $f(x) = x$ definita da $\mathbb{R}$ in $\mathbb{R}$ con la topologia usuale è chiaramente una funzione limitata ma l'immagine di f, che coincide con tutto $\mathbb{R}$, è chiaramente illimitata.

Ogni funzione limitata è sempre sequenzialmente limitata, mentre il viceversa potrebbe non valere.

I prossimi due risultati valgono per le funzioni lineari definite su uno spazio vettoriale topologico su $\mathbb{K}$ a valori nello stesso campo $\mathbb{K}$.

Proposizione 4.3.2. *Sia V uno spazio vettoriale topologico su $\mathbb{K}$ e $f : V \to \mathbb{K}$ una funzione lineare. Se f non è identicamente nulla allora f è una funzione aperta.*

Dimostrazione. Dato un aperto A di V dimostriamo che se f non è identicamente nulla allora $f(A)$ è aperta in $\mathbb{K}$. Innanzitutto prendiamo un qualunque elemento $y \in f(A)$ e sia $x \in A$ tale che $y = f(x)$. Esisterà allora un aperto bilanciato B tale che $x + B \subseteq A$, ora la funzione f non può essere identicamente nulla su B in quanto per la limitatezza dei punti per ogni $v \in V$ esisterà $t > 0$ tale che $v/t \in B$.

Scegliamo perciò $z \in B$ tale che $f(z) \neq 0$ allora per ogni $\lambda \in \mathbb{K}$ con $|\lambda| < 1$ avremo che $x + \lambda z \in A$ e dunque

$$f(A) \supseteq \left\{ y + f(z)\lambda \;\middle|\; |\lambda| < 1 \right\}$$

ora poiché $f(z) \neq 0$ l'insieme a destra è un aperto in $\mathbb{K}$ (che può essere $\mathbb{R}$ o $\mathbb{C}$) contenente y, dunque per l'arbitrarietà di y l'insieme $f(A)$ è aperto in $\mathbb{K}$. ■

Proposizione 4.3.3. *Sia V uno spazio vettoriale topologico su $\mathbb{K}$ e $f : V \to \mathbb{K}$ una funzione lineare, allora sono equivalenti*

1. *f è continua su V;*
2. *$\ker f$ è chiuso in V;*
3. *Se f non è identicamente nulla allora $\ker f$ non è denso in V;*
4. *Esiste un intorno $U \subseteq V$ dell'origine tale che $f(U)$ è limitato in $\mathbb{K}$.*

Dimostrazione. Dal punto 1 segue immediatamente il punto 2, inoltre se assumiamo valido il punto 2 e $\ker f$ è denso in V allora avremo che

$$\ker f = \mathrm{Cl}(\ker f) = V$$

e quindi f è identicamente nulla. Supponiamo ora che vale il punto 3 e che f non è identicamente nulla, allora esiste un $x \in V$ e un aperto bilanciato A contenente l'origine tali che

$$(x + A) \cap \ker f = \emptyset$$

e quindi $-f(x) \neq f(y)$ per ogni $y \in A$. Ma essendo A bilanciato avremo che anche $f(A)$ è bilanciato in $\mathbb{K}$ per linearità e quindi, non potendo coincidere con tutto $\mathbb{K}$, esisterà $0 \leq r \leq \left|f(x)\right|$ tale che

$$f(A) = \{\lambda \in \mathbb{K} \mid |\lambda| \leq r\}$$

e quindi si ha il punto 4.

Ora supponiamo che vale la 4 e dimostriamo che f è continua in 0. Sia $U \subseteq V$ intorno aperto dell'origine per cui $\left|f(U)\right| \leq L$ e quindi U è contenuto nella controimmagine di una palla centrata nell'origine e di raggio L in $\mathbb{K}$. Per riscalamento le controimmagini di tutte le palle centrate in 0 saranno intorni in V dell'origine e quindi f sarà continua in 0 e di conseguenza su tutto V. ■

Osservazione. Si osservi che per una funzione lineare e continua tra spazi vettoriali topologici in generale **il kernel non è chiuso** in quanto il singleton $\{0\}$ può non essere un sottoinsieme chiuso. Se tutti i punti del codominio sono chiusi (o equivalentemente se il codominio è uno spazio vettoriale topologico di Hausdorff) allora si può dire che il kernel è un sottospazio chiuso.

Proposizione 4.3.4. *Se $f : V \to W$ è una funzione tra spazi vettoriali topologici lineare e continua in* 0 *allora è uniformemente continua su tutto V. Analogamente se f è sequenzialmente continua in* 0 *allora sarà sequenzialmente uniformemente continua su tutto V.*

Dimostrazione. Se f è continua in 0 allora per ogni aperto $A \subseteq W$ contenente l'origine esiste un aperto $B \subseteq V$ contenente l'origine tale che $f(B) \subseteq A$. Per la linearità di f avremo allora che

$$f(x + B) \subseteq f(x) + A$$

e quindi f è uniformemente continua su tutto V.

Per quanto riguarda la continuità sequenziale osserviamo che per ogni coppia di successioni $x_n, y_n \in V$ tali che $x_n - y_n \to 0$ avremo che $f(x_n) - f(y_n) = f(x_n - y_n) \to 0$ e quindi f è sequenzialmente uniformemente continua. ■

Indichiamo con $L(V, W)$ lo spazio vettoriale delle funzioni lineari continue tra gli spazi vettoriali topologici V e W su $\mathbb{K}$, inoltre poniamo $V^* = L(V, \mathbb{K})$ che sarà detto **spazio duale** o **duale continuo** di V.

Proposizione 4.3.5. *Le funzioni lineari continue sono limitate e le funzioni lineari sequenzialmente continue sono sequenzialmente limitate.*

Dimostrazione. Consideriamo una funzione $f : V \to W$ tra spazi vettoriali topologici. Supponiamo inizialmente che f è continua e sia A un insieme limitato di V, vogliamo dimostrare che $f(A)$ è limitato in W. Per ogni intorno aperto U dell'origine in W per continuità esiste un intorno aperto dell'origine U' in V tale che $f(U') \subseteq U$ e quindi dalla definizione di limitatezza esiste $t > 0$ tale che $E \subseteq \lambda U'$ per ogni $|\lambda| > t$ e quindi

$$f(E) \subseteq f(\lambda U') = \lambda f(U') \subseteq \lambda U$$

e quindi E è limitato.

Dalla definizione di continuità sequenziale e per la limitatezza delle successioni convergenti discende immediatamente che funzioni sequenzialmente continue sono sequenzialmente limitate. ■

In generale non tutte le funzioni limitate sono continue e non tutte le funzioni sequenzialmente limitate sono limitate. Se però supponiamo che lo spazio di partenza sia una spazio metrico (o equivalentemente se è $\mathscr{N}_1$ e di Hausdorff) allora le affermazioni precedentementi sono del tutto equivalenti:

Teorema 4.3.6. *Siano V, W spazi vettoriali topologici con V metrico, e sia $f : V \to W$ un'applicazione lineare. Allora le affermazioni seguenti*

1. *f è continua;*
2. *f è limitata;*
3. *f è sequenzialmente continua;*
4. *f è sequenzialmente limitata.*

sono tutte equivalenti.

Dimostrazione. Grazie alla proposizione 2.5.6 le seguenti implicazioni sono tutte verificate

$$\begin{array}{ccc} \text{continuità} & \Longleftrightarrow & \text{continuità sequenziale} \\ \Downarrow & & \Downarrow \\ \text{limitatezza} & \Longrightarrow & \text{limitatezza sequenziale} \end{array}$$

quindi se dimostriamo che le funzioni sequenzialmente limitate sono sequenzialmente continue il teorema è chiaramente dimostrato.

Presa una metrica d invariante su V la cui topologia indotta coincide con la topologia di V allora per induzione si verifica facilmente che $d(nx, 0) \leq nd(x, 0)$ per ogni $x \in V$ e per ogni $n \in \mathbb{N}$. Sia ora x_n una successione in V convergente a 0, allora $\lim_{n\to+\infty} d(x_n, 0) = 0$ e quindi possiamo supporre senza perdere in generalità che $d(x_n, 0) < 1$ per ogni n.

Inoltre per ogni $n \in \mathbb{N}$ esisterà $k_n \in \mathbb{N}$ tale che $k_n > k_{n-1}$ e per ogni $m > k_n$ si avrà

$$d(x_m, 0) < \frac{1}{(n+1)^2}$$

e poniamo anche $k_0 = 0$. Ora per ogni $m \in \mathbb{N}$ poniamo $\alpha_m = n$ se e solo se $k_{n-1} < m \leq k_n$, in tal caso si avrà

$$d(\alpha_m x_m, 0) \leq nd(x_m, 0) \leq \frac{1}{n}$$

e quindi avremo che $\alpha_m x_m \to 0$ e $\alpha_m \to +\infty$. Posto $y_n = f(x_n)$ per ipotesi e per la linearità di f avremo che $\alpha_n y_n$ è limitata in W, dimostriamo che y_n converge a 0.

Per ogni intorno aperto dell'origine U in W, che possiamo supporre bilanciato senza perdere in generalità, esisterà $t > 0$ tale che $\alpha_n y_n \in tU$ per ogni $n \in \mathbb{N}$, d'altronde esisterà $N \in \mathbb{N}$ tale che $\alpha_n \geq t$ per ogni $n > N$ e in tal caso avremo che

$$y_n \in \frac{t}{\alpha_n} U \subseteq U$$

e quindi $y_n \to 0$ ed f è sequenzialmente continua in 0. Per linearità d'altronde f sarà sequenzialmente continua su tutto V e il teorema è così dimostrato. ■

Consideriamo adesso uno spazio vettoriale topologico complesso X, nell'introduzione abbiamo definito lo spazio vettoriale reale $X_{\mathbb{R}}$ come lo spazio X con gli stessi elementi e la stessa operazione di somma ma con il prodotto esterno ristretto ai soli scalari reali. Abbiamo già visto all'inizio che non tutti gli operatori lineari su $X_{\mathbb{R}}$. Doteremo allora $X_{\mathbb{R}}$ della medesima topologia di X, chiaramente $X_{\mathbb{R}}$ sarà anch'esso uno spazio vettoriale topologico.

In questo modo l'applicazione

$$x \in X_{\mathbb{R}} \to ix \in X_{\mathbb{R}}$$

è un omeomorfismo tra spazi topologici anche se il prodotto per i non sarebbe definito in $X_{\mathbb{R}}$ (ma lo è in X quindi la scrittura di sopra è ben posta).

Lemma 4.3.7. *Sia X uno spazio vettoriale topologico complesso, allora l'applicazione T definita nella proposizione 1.5.2 è biiettiva da $L(X,\mathbb{C})$ in $L(X_{\mathbb{R}},\mathbb{R})$.*

Dimostrazione. Ricordiamo che l'applicazione $T : X^+ \to X^+_{\mathbb{R}}$ è definita come

$$(Tf)(x) = \mathrm{Re} f(x)$$
$$(T^{-1}g)(x) = g(x) - ig(ix)$$

e quindi $T[L(X,\mathbb{C})] \subseteq L(X_{\mathbb{R}},\mathbb{R})$ per la continuità della parte reale di un numero complesso.

Viceversa e $g \in L(X_{\mathbb{R}},\mathbb{R})$ allora $T^{-1}g$ è composizione di funzioni continue e di conseguenza $T^{-1}(g) \in L(X,\mathbb{C})$. ■

4.4 Il teorema della funzione aperta

Uno spazio vettoriale topologico V è un **F-spazio** se e solo se è $\mathcal{N}_1$, di Hausdorff e completo, dunque se la sua topologia è indotta da una metrica invariante, completa, bilanciata ed invariante sugli scalari. Nel capitolo sugli spazi uniformi abbiamo dimostrato il teorema di Baire senza però mostrare altri risultati correlati. L'importanza di tale teorema risiede infatti nella sua applicazione negli F-spazi e negli spazi di Banach che vedremo successivamente.

Lemma 4.4.1. *Dato un F-spazio V e un sottospazio vettoriale chiuso $N \leq V$ allora V/N è anch'esso un F-spazio rispetto alla topologia quoziente.*

Dimostrazione. Prendiamo una qualunque metrica invariante d su V e definiamo

$$d_N(x+N, y+N) = \inf_{z\in N} d(x+z, y) = \inf_{z\in N} d(x-y, z) = \mathrm{dist}(x-y, N)$$

Verifichiamo innanzitutto che d_N sia una metrica su V/N. Per ogni $x, y \in V$ si ha

$$0 = d_N(x+N, y+N) = \mathrm{dist}(x-y, N) \Leftrightarrow x-y \in \mathrm{Cl}(N) = N \Leftrightarrow x+N = y+N$$
$$d(x+N, y+N) = \inf_{z\in N} d(x+z, y) = \inf_{z\in N} d(x, y-z) = \inf_{z\in N} d(y+z, x) = d_N(y+N, x+N)$$

la disuguaglianza triangolare invece è una conseguenza della disuguaglianza

$$d(a+z+w, c) \leq d(a+z, b) + d(b+w, c)$$

ottenuta grazie all'invarianza di d. Dall'invarianza di d segue immediatamente anche l'invarianza di d_N sfruttando il fatto che N è pur sempre un sottospazio vettoriale.

Infine abbiamo che

$$x+N \in B(N,r,d_N) \Leftrightarrow d(x-z,0)<r \text{ per qualche } z\in N \Leftrightarrow x+N \in \pi_N\left[B(0,r,d)\right]$$

per il lemma 4.1.6 π_N è continua ed aperta se dotiamo V/N della topologia quoziente, quindi le palle aperte rispetto a d_N sono aperte in V/N. Viceversa se prendiamo un qualunque aperto $U\subseteq V/N$ ed $x+N\in U$ per definizione di topologia quoziente esisterà $r>0$ tale che $B(x,r,d)\subseteq \pi_N^{-1}(U)$ e dunque $B(x+N,r,d_N)\subseteq U$. Abbiamo così dimostrato che la topologia indotta dalla metrica invariante d_N coincide con la topologia quoziente, e quindi per quanto detto precedentemente anche le rispettive strutture topologiche indotte dovranno coincidere e perciò V/N è completo se e solo se d_N lo è.

Prendiamo adesso una qualunque successione di Cauchy $x_k+N\in V/N$, possiamo sempre scegliere un'estratta x_{n_k} in modo tale che $d_N\left(x_{n_k}+N, x_{n_{k+1}}+N\right)<2^{-k}$, sfruttando l'invarianza di d esistono $z_k\in N$ tali che $d\left(x_{n_k}+z_k, x_{n_{k+1}}+z_{k+1}\right)<2^{-k}$ e quindi $x_{n_k}+z_k$ è di Cauchy anche in V e perciò converge. Per la continuità di π_N anche $x_{n_k}+N$ converge in V/N e per la proposizione 3.3.2 anche x_k+N converge. ■

Definizione 4.4.2. Dati due spazi vettoriali topologici V, W e Γ una famiglia di funzioni lineari e continue da V in W. Diremo che Γ è **equicontinua** se e solo se per ogni intorno dell'origine $B\subseteq W$ esiste un altro intorno dell'origine $A\subseteq V$ tale che $f(A)\subseteq B$ per ogni $f\in\Gamma$.

Osservazione. Dato che le funzioni continue sono limitate allora se Γ è equicontinua allora per ogni insieme limitato $A\subseteq V$ esiste un altro insieme limitato $B\subseteq W$ tale che $f(A)\subseteq B$. Per questa ragione gli è stato dato questo nome.

Teorema 4.4.3 (Principio di uniforme limitatezza)**.** *Dati V e W spazi vettoriali topologici e Γ famiglia di funzioni lineari e continue da V in W. Posto dunque*

$$\mathcal{O}(x)=\left\{f(x)\in W \,\middle|\, f\in\Gamma\right\}$$
$$L=\left\{x\in V \mid \mathcal{O}(x) \text{ limitato in } W\right\}$$

se L è un insieme di seconda categoria allora Γ è equicontinuo e inoltre $L=V$.

Dimostrazione. Sia $U\subseteq W$ un aperto bilanciato contenente l'origine, allora per regolarità esiste un altro aperto bilanciato B contenente l'origine tale che $\mathrm{Cl}(B)+\mathrm{Cl}(B)\subseteq W$ e poniamo

$$0\in A=\bigcap_{f\in\Gamma} f^{-1}(\mathrm{Cl}(B))$$

Per ogni $x\in L$ esisterà chiaramente $n\in\mathbb{N}$ per cui $\mathcal{O}(x)\subseteq nB$ dunque per linearità $x\in nA$ dunque

$$L\subseteq\bigcup_{n=1}^{+\infty} nA$$

dato che A è un chiuso di V per definizione di seconda categoria almeno uno degli nA deve contenere un punto interno in L e quindi anche A contiene un punto interno che chiameremo x. Esisterà allora un aperto $C\subseteq A-x\subseteq V$ contenente l'origine e per ogni $f\in\Gamma$ si avrà allora che

$$f(C)\subseteq f(A)-f(x)\subseteq \mathrm{Cl}(B)-\mathrm{Cl}(B)\subseteq U$$

dunque Γ è equicontinuo. In particolare per l'osservazione di sopra le orbite $\mathcal{O}(x)$ saranno limitate per ogni $x\in V$ e perciò $L=V$. ■

Proposizione 4.4.4. *Siano V F-spazio, W spazio vettoriale topologico e $\Lambda_n \in L(V,W)$ tale che $\lim_{n\to+\infty} \Lambda_n x$ converge per ogni $x \in V$. Allora posto*

$$\Lambda x = \lim_{n\to+\infty} \Lambda_n x$$

avremo che $\Lambda \in L(V,W)$.

Dimostrazione. Se poniamo $\Gamma = \{\Lambda_n\}_{n\in\mathbb{N}}$ avremo che le orbite $\mathscr{O}(x)$ al variare di x in V sono tutte limitate. Dunque per il principio di uniforme limitatezza per ogni intorno aperto dell'origine $E \subseteq W$ esistono altri due intorni aperti dell'origine $F \subseteq W$, $A \subseteq V$ tali che $\Lambda_n(A) \subseteq F \subseteq \mathrm{Cl}(F) \subseteq E$.

Passando al limite in n avremo così che $\Lambda(A) \subseteq E$ e dunque Λ è continua nell'origine. Per linearità Λ sarà continua su tutto V e perciò $\Lambda \in L(V,W)$. ■

Dal principio di uniforme limitatezza seguono dei risultati molto importanti nello studio degli F-spazi e in particolare negli spazi di Banach: il teorema della funzione aperta e il teorema del grafico chiuso. In questa sezione li dimostreremo entrambi, a partire dal

Teorema 4.4.5 (della funzione aperta)**.** *Siano V e W F-spazi e $\Lambda \in L(V,W)$ suriettiva. Allora Λ è una funzione aperta.*

chiaramente la suriettività di Λ è strettamente necessaria, difatti se Λ fosse identicamente nulla allora la sua immagine sarebbe uguale a $\{0\}$ che non è chiaramente aperta. Da questo teorema discende il seguente corollario di notevole importanza nel seguito

Corollario 4.4.6. *Se Λ è lineare, continua e biettiva tra due F-spazi allora anche Λ^{-1} è continua.*

In questa sede dimostreremo un risultato più forte del teorema della funzione aperta:

Teorema 4.4.7. *Siano V F-spazio e W spazio vettoriale topologico di Hausdorff. Se $\Lambda : V \to W$ è una funzione lineare e continua per cui $\Lambda(V)$ è un sottoinsieme di seconda categoria in W allora*

- *$\Lambda(V) = W$;*
- *Λ è aperta;*
- *W è un F-spazio.*

Per dimostrare il seguente risultato procederemo per gradi, dimostrando mano mano vari risultati parziali.

Proposizione 4.4.8. *Siano V, W spazi vettoriali topologici e $\Lambda : V \to W$ lineare generica, allora sono equivalenti*

1. *Λ è aperta;*
2. *Per ogni intorno dell'origine $A \subseteq V$ esiste un aperto $B \subseteq W$ contenente l'origine tali che $\Lambda(A) \supseteq B$.*

Dimostrazione. Supponiamo innanzitutto che Λ sia aperta, poiché $\Lambda(0) = 0$ allora $\Lambda(A)$ è un aperto contenente $0 \in Y$, da cui segue immediatamente il punto 2.

Viceversa supponiamo ora che valga il secondo punto e prendiamo un generico aperto U in V, dimostriamo che $\Lambda(U)$ è ancora aperto in W. Preso un qualunque $y \in \Lambda(U)$ esisterà $x \in U$ tale che $\Lambda x = y$, esisterà inoltre un aperto bilanciato A di V contenente l'origine per cui $x + A \subseteq U$.

Per la linearità di Λ avremo dunque che

$$\Lambda(U) \supseteq \Lambda(x + A) = y + \Lambda(A) \supseteq y + B$$

per ogni $y \in \Lambda(U)$, dunque $\Lambda(U)$ è un aperto di W. ■

Lemma 4.4.9. *Dati V spazio vettoriale metrico, W spazio vettoriale topologico e $\Lambda \in L(V, W)$ tale che $\Lambda(V)$ sia di seconda categoria in W. Allora per ogni intorno aperto dell'origine $A \subseteq V$ esiste un altro intorno aperto dell'origine $B \subseteq W$ tale che*

$$\mathrm{Cl}(\Lambda(A)) \supset B$$

Dimostrazione. Innanzitutto esiste un aperto $A' \subseteq V$ bilanciato contenente l'origine tale che $A' + A' \subseteq A$. Per il lemma 4.2.5 avremo che

$$\Lambda(V) = \bigcup_{n=1}^{+\infty} n\Lambda(A')$$

poiché per ipotesi $\Lambda(V)$ è un insieme di seconda categoria esisterà un intero $n \in \mathbb{N}$ tale che $\mathrm{Cl}(n\Lambda(A')) = n\mathrm{Cl}(\Lambda(A'))$ possiede un punto interno e di conseguenza anche $\mathrm{Cl}(\Lambda(A'))$.

Quindi esisteranno $y \in \mathrm{Cl}(\Lambda(A'))$ e $B' \subseteq W$ intorno aperto bilanciato dell'origine tali che $\mathrm{Cl}(\Lambda(A')) \supseteq y + B'$ e dato che la chiusura di un insieme bilanciato è ancora bilanciata avremo anche che $\mathrm{Cl}(\Lambda(A)) \supseteq -y - B' = B' - y$. Applicando ora la proposizione 4.1.12 avremo infine che

$$\mathrm{Cl}(\Lambda(A)) \supseteq \mathrm{Cl}(\Lambda(A' + A')) \supseteq \mathrm{Cl}(\Lambda(A')) + \mathrm{Cl}(\Lambda(A')) \supseteq B' + B' = B$$

con B aperto contenente l'origine. ■

Lemma 4.4.10. *Siano V F-spazio, $A \subseteq V$ intorno aperto dell'origine, W spazio vettoriale topologico di Hausdorff e $\Lambda \in L(V, W)$ tale che $\mathrm{Cl}(\Lambda(A))$ è un intorno dell'origine per ogni aperto $A \subseteq V$ contenente 0. Allora anche $\Lambda(A)$ è un intorno dell'origine.*

Dimostrazione. Consideriamo una qualunque metrica invariante bilanciata d che induce la topologia su V e sia $A \subseteq V$ aperto generico contenente l'origine. Innanzitutto esiste $r > 0$ tale che $D(0, 2r, d) \subseteq A$, dunque per ogni $n \in \mathbb{N}_0$ poniamo di conseguenza

$$A_n = B(0, 2^{-n} r, d)$$

per la proprietà di invarianza abbiamo $A_n + A_n \subseteq A_{n-1}$ per ogni $n \in \mathbb{N}$.

Prendiamo innanzitutto un generico $y_0 \in \mathrm{Cl}(\Lambda(A_0))$, ora poiché A_1 è un aperto contenente l'origine avremo che $\mathrm{Cl}(\Lambda(A_1))$ è un intorno dell'origine in W e perciò $y_0 - \mathrm{Cl}(\Lambda(A_1))$ è un intorno di y_0. Per definizione di chiusura avremo che

$$[y_0 - \mathrm{Cl}(\Lambda(A_1))] \cap \Lambda(A_0) \neq \emptyset$$

e possiamo così scegliere $x_0 \in A_0$ in modo tale che $\Lambda x_0 \in y_0 - \mathrm{Cl}(\Lambda(A_1))$. Sia adesso $y_1 = y_0 - \Lambda x_0 \in \mathrm{Cl}(\Lambda(A_1))$ allora $y_1 - \mathrm{Cl}(\Lambda(A_2))$ è un intorno di y_1 e dunque

$$[y_1 - \mathrm{Cl}(\Lambda(A_2))] \cap \Lambda(A_1) \neq \emptyset$$

e quindi possiamo prendere $x_1 \in A_1$ in modo tale che $\lambda x_1 \in y_1 - \mathrm{Cl}(\Lambda(A_2))$. Iterando il procedimento per ogni $n \in \mathbb{N}$ si avrà

$$\begin{gathered} y_n = y_{n-1} - \Lambda x_{n-1} \in \mathrm{Cl}(\Lambda(A_n)) \\ x_n \in A_n \\ y_n = y_0 - \Lambda x_0 - \Lambda x_1 - \cdots - \Lambda x_{n-1} = y_0 - \Lambda \bar{x}_n \end{gathered}$$

A questo punto osserviamo che $x_{n-1} + x_n \in A_{n-1} + A_{n-1} \subseteq A_{n-1}$, iterando il procedimento per ogni $0 < m < n$ avremo che

$$x_m + x_{m+1} + \cdots + x_{n-1} + x_n \in A_{m-1} = B(0, 2^{-m} r, d) \tag{4.4.1}$$

e perciò $\bar{x}_m \in A_0 + A_0 \subseteq D(0, 2r, d)$. Ma la (4.4.1) ci dice anche che $\bar{x}_n$ è una successione di Cauchy in V, poiché V è un F-spazio esiste $x \in D(0, 2r, d)$ per cui $\bar{x}_n \to x \in A$. Ma allora il limite di y_n esiste ed è uguale a $y = y_0 - \Lambda x$, vogliamo ora dimostrare che $y = 0$ da cui segue che $y_0 \in \Lambda(A)$ e quindi $\mathrm{Cl}(\Lambda(A_0)) \subseteq \Lambda(A)$.

Per la continuità di Λ e per la regolarità degli spazi vettoriali topologici per ogni $U \subseteq W$ aperto contenente l'origine esiste $n \in \mathbb{N}$ tale che $y_n \in \mathrm{Cl}(\Lambda(A_n)) \subseteq U$ e perciò passando al limite $y \in \mathrm{Cl}(U)$. Per l'arbitrarietà di U ed essendo W uno spazio di Hausdorff avremo che $y = 0$. ■

In questa maniera abbiamo dimostrato che Λ è una applicazione aperta. Ora poiché $\Lambda(V)$ è un sottospazio vettoriale aperto di W segue immediatamente che $\Lambda(V) = W$ e dunque la nostra funzione è suriettiva. L'ultimo risultato da dimostrare è il seguente

Lemma 4.4.11. *Siano V F-spazio, W spazio vettoriale topologico di Hausdorff, $\Lambda \in L(V, W)$ funzione aperta. Allora anche W è un F-spazio.*

Dimostrazione. L'insieme $\ker \Lambda$ è chiuso in V e diverso da V per la suriettività di Λ, quindi per la proposizione 4.4.1 $V / \ker \Lambda$ è ancora un F-spazio. Possiamo definire l'applicazione $\Phi : V / \ker \Lambda \to W$ affinché

$$\Phi(x + \ker \Lambda) = \Lambda x$$

Chiaramente Φ è un'applicazione ben definita, lineare e biettiva.

Fissati due intorni aperti dell'origine $A \subseteq V/N$, $B \subseteq W$ osserviamo che

$$x + \ker \Lambda \in \Phi^{-1}(B) \Leftrightarrow \Lambda x \in B \Leftrightarrow x + \ker \Lambda \in \pi_\Lambda \left[\Lambda^{-1}(B) \right]$$

$$y \in \Phi(A) \Leftrightarrow y = \Lambda x \text{ per qualche } x + N = \pi_\Lambda(x) \in A \Leftrightarrow y \in \Lambda \left[\pi_\Lambda^{-1}(A) \right]$$

per il lemma 4.1.6 sappiamo che la funzione π_Λ è continua e aperta, perciò Φ è un omeomorfismo e W è un F-spazio. ■

La dimostrazione del teorema 4.4.7 si ottiene mettendo assieme questi risultati parziali.

Dopo aver dimostrato il teorema della funzione aperta dimostriamo il teorema successivo, quello del grafico chiuso.

Lemma 4.4.12. *Il prodotto di due F-spazi sul campo $\mathbb{K}$ è ancora un F-spazio.*

Dimostrazione. Siano V, W due F-spazi con metriche invarianti d_V, d_W rispettivamente, dimostriamo che la metrica definita su $V \times W$

$$d\left[(x,v),(y,w)\right] = d_V(x,y) \vee d_W(v,w)$$

è ancora una metrica invariante compatibile completa con la topologia prodotto di $V \times W$. La riflessività, l'invarianza e la simmetria sono immediate da verificare, dimostriamo che vane la disuguaglianza triangolare. Presi $x_1, x_2, x_3 \in V$, $v_1, v_2, v_3 \in W$ non e difficile verificare che

$$\begin{aligned} d[(x_1,v_1),(x_3,v_3)] &\leq \left[d_V(x_1,x_2) + d_V(x_2,x_3)\right] \vee \left[d_W(v_1,v_2) + d_W(v_2,v_3)\right] \\ &\leq \left[d_V(x_1,x_2) \vee d_W(v_1,v_2)\right] + \left[d_V(x_2,x_3) \vee d_W(v_2,v_3)\right] \end{aligned}$$

e quindi la disuguaglianza triangolare è verificata. Infine osserviamo che

$$B\big((0,0),r,d\big) = B\big(0,r,d_V\big) \times B\big(0,r,d_W\big)$$

dunque la metrica d induce la topologia prodotto di $V \times W$.

Poiché (x_n, v_n) è di Cauchy se e solo se x_n e v_n sono di Cauchy e (x_n, v_n) converge se e solo se x_n e v_n convergono avremo che anche $V \times W$ è un F-spazio. ■

4.5 Operatori chiusi

Consideriamo ora gli operatori lineari tra spazi vettoriali topologici X ed Y. Diremo che l'operatore $L \in \mathcal{UL}\left(X,Y\right)$ è continuo se e solo se $\operatorname{dom} L = X$ ed L è continuo su tutto X, in altre parole se e solo se $L \in L(X,Y)$. Dato che la continuità in certi casi è una condizione troppo stringente introduciamo una proprietà un po' più debole della continuità ma comunque utile.

Definizione 4.5.1. Un operatore $L \in \mathcal{UL}\left(X,Y\right)$ è detto **chiuso** se e solo se il suo grafico $\mathcal{G}(L)$ è un sottospazio chiuso in $X \times Y$ dotato della topologia prodotto. Diremo invece che L è **chiudibile** se e solo se possiede un'estensione chiusa.

Diremo invece che L è **densamente definito** se e solo se $\operatorname{dom} L$ è denso in X.

Proposizione 4.5.2. *Siano X, Y spazi vettoriali topologici, per ogni $L \in \mathcal{UL}\left(X,Y\right)$ le seguenti affermazioni sono equivalenti:*

1. *L è chiuso;*
2. *per ogni net $x_a \in \operatorname{dom} L$ tale che $x_a \to x \in X$ ed $L(x_a) \to y \in Y$ si ha $x \in \operatorname{dom} L$ e $L(x) = y$.*

Dimostrazione. Per quanto detto sugli spazi topologici l'operatore L è chiuso se e solo se per ogni net $x_a \in \operatorname{dom} L$ tale che $\left[x_a, L(x_a)\right] \to (x,y)$ in $X \times Y$ abbiamo $(x,y) \in \mathcal{G}(L)$ ovvero $L(x) = y$. La tesi si ottiene ricordando che nella topologia prodotto si ha convergenza se e solo se la convergenza è puntuale. ■

Osservazione. Se L è chiuso non è detto che il suo dominio sia anch'esso chiuso, anzi in molti casi esso risulta essere denso in X e allo stesso tempo strettamente contenuto in X. Questo perché la net x_a che confuterebbe la chiusura del dominio di L non può entrare in gioco poiché $L(x_a)$ non converge in Y, mentre noi abbiamo imposto che debba convergere.

Osservazione. Se L è un operatore chiuso allora l'immagine di L è un sottospazio vettoriale topologico di Hausdorff di Y.

Corollario 4.5.3. *Se $L \in \mathcal{UL}(X,Y)$ è chiuso allora il sottospazio $\ker L = \{x \in \operatorname{dom} L \mid L(x) = 0\}$ è chiuso in X.*

Proposizione 4.5.4. *Se X ed Y sono spazi vettoriali topologici allora per ogni $L \in \mathcal{UL}(X,Y)$ le seguenti affermazioni sono equivalenti:*

1. *L è chiudibile;*
2. *per ogni net $x_a \in \operatorname{dom} L$ tale che $x_a \to 0$ e $L(x_a) \to y \in Y$ si ha $y = 0$ (si noti che non abbiamo imposto che Y sia di Hausdorff);*
3. *$\mathrm{Cl}(\mathcal{G}(L))$ è un grafico.*

Dimostrazione. Supponiamo che L sia chiusibile con R una sua estensione chiusa. Allora per ogni net $x_a \in \operatorname{dom} L \subseteq \operatorname{dom} R$ tale che $x_a \to 0$ e $L(x_a) = R(x_a) \to y \in Y$ per la proposizione precedente abbiamo $y = R(0) = 0$.

Supponiamo ora che valga il punto 2. Dato che $\mathrm{Cl}(\mathcal{G}(L))$ è un sottospazio vettoriale di $X \times Y$ esso sarà il grafico di un operatore se e solo se per ogni $(0,y) \in \mathrm{Cl}(\mathcal{G}(L))$ si ha $y = 0$. Per le proprietà della chiusura esisterà una net $x_a \in \operatorname{dom} L$ tale che $x_a \to 0$ e $L(x_a) \to y$ e per ipotesi si ha così $y = 0$, dunque $\mathrm{Cl}(\mathcal{G}(L))$ è un grafico.

Supponiamo infine che $\mathrm{Cl}(\mathcal{G}(L))$ sia un grafico, possiamo allora definire il seguente operatore

$$\operatorname{dom} \mathrm{Cl}(L) = \left\{x \in X \,\middle|\, \exists y \in Y \text{ tale che } (x,y) \in \mathrm{Cl}(\mathcal{G}(L))\right\}$$
$$\mathrm{Cl}(L)(x) = y \Leftrightarrow (x,y) \in \mathrm{Cl}(\mathcal{G}(L)) \ \forall x \in \operatorname{dom} \mathrm{Cl}(L)$$

chiaramente $\mathcal{G}(\mathrm{Cl}(L)) = \mathrm{Cl}(\mathcal{G}(L))$ e quindi $\mathrm{Cl}(L)$ è un operatore chiuso. Dato che $\mathcal{G}(L) \subseteq \mathrm{Cl}(\mathcal{G}(L))$ avremo che $\mathrm{Cl}(L)$ è un'estensione di L e dunque L è chiudibile. ■

Dimostriamo ora il teorema del grafico chiuso, un importante teorema che fornisce un collegamento tra operatori chiusi e quelli continui quando sono definiti su F-spazi.

Teorema 4.5.5 (del grafico chiuso)**.** *Siano V, W F-spazi e $\Lambda \in \mathcal{UL}(V,W)$, allora*

$$\Lambda \textit{ è continua } \Leftrightarrow \operatorname{dom} \Lambda = V \textit{ e } \Lambda \textit{ è chiuso}$$

Dimostrazione. Supponiamo che il sottospazio vettoriale $\mathcal{G}(\Lambda)$ sia chiuso, allora il grafico è anch'esso un F-spazio. Ricordiamo che le due proiezioni

$$P_1 : (x,y) \in \mathcal{G}(\Lambda) \to x \in \operatorname{dom} \Lambda = V$$
$$P_2 : (x,y) \in \mathcal{G}(\Lambda) \to y = \Lambda x \in W$$

sono entrambe continue e P_1 è anche invertibile con inversa continua per il corollario 4.4.6, dunque la funzione $P_2 \circ P_1^{-1}$ che va da X in Y è continua e coincide con Λ.

Viceversa se Λ è continua allora per ogni successione $x_n \in V$ (siamo in uno spazio metrico) tale che $x_n \to x$ e $\Lambda(x_n) \to y$ abbiamo per continuità $\Lambda(x_n) \to \Lambda(x)$, dato che Y è uno spazio di Hausdorff questo implica che $\Lambda(x) = y$ e perciò Λ è chiuso. ■

Il teorema del grafico chiuso ci fornisce un utile criterio per determinare se una funzione lineare tra F-spazi è continua. Più nello specifico T è continua se e solo se per ogni successione $x_n \in V$ tale che $x_n \to x$ e $Tx_n \to y$ si ha $y = Tx$, in particolare possiamo dare per scontato che la successione Tx_n converge in W.

4.6 Spazi finitamente generati

Ricordiamo che uno spazio vettoriale V su $\mathbb{K}$ ha dimensione finita se e solo se esiste un'applicazione $I : V \to \mathbb{K}^n$ lineare biettiva per un certo $n \in \mathbb{N}$. In tal modo possiamo portare su V la topologia standard di $\mathbb{K}^n$ nella maniera più immediata:

$$U \text{ aperto di } V \Leftrightarrow I(U) \text{ aperto di } \mathbb{K}^n$$

Con questa topologia V è chiaramente uno spazio vettoriale topologico, e più nello specifico un F-spazio, e I è un omeomorfismo tra spazi topologici. Poiché tutte le applicazioni lineari biettive di $\mathbb{K}^n$ in sé sono omeomorfismi avremo che questa topologia su V è indipendente dalla scelta di I.

Possiamo però scegliere altre topologie su V che lo rendono uno spazio vettoriale topologico, ma ce ne sta solamente una che è anche di Hausdorff come vedremo nel prossimo risultato:

Teorema 4.6.1. *Se V è uno spazio vettoriale topologico di Hausdorff su $\mathbb{K}$ di dimensione $n \in \mathbb{N}$ allora ogni applicazione lineare biettiva $I : V \to \mathbb{K}^n$ è un omeomorfismo.*

In particolare V possiede un'unica struttura di spazio vettoriale topologico di Hausdorff.

Dimostrazione. Sia $f_1, f_2, \dots, f_k$ una base di V e consideriamo l'isomorfismo lineare

$$F : (c_1, c_2, \dots, c_n) \in \mathbb{K}^n \to c_1 f_1 + c_2 f_2 + \cdots + c_n f_n \in V$$

chiaramente per ogni $1 \le i \le n$ l'applicazione $e_i^* : c \in \mathbb{K}^n \to c_i \in \mathbb{K}$ è continua dunque avremo che

$$F = f_1 e_1^* + f_2 e_2^* + \cdots + f_n e_n^*$$

quindi F è somma e prodotto di funzioni continue. Poiché V è uno spazio vettoriale topologico, e quindi somma e prodotto sono continue, anche F sarà continua.

Poniamo

$$\begin{aligned} S &= \left\{ c \in \mathbb{K}^n \;\middle|\; \|c\|_\infty = 1 \right\} \\ B &= \left\{ c \in \mathbb{K}^n \;\middle|\; \|c\|_\infty < 1 \right\} \\ D &= \left\{ c \in \mathbb{K}^n \;\middle|\; \|c\|_\infty \le 1 \right\} \end{aligned}$$

sappiamo che S è compatto in $\mathbb{K}^n$ e dunque $F(S)$ è compatto in V con $0 \notin F(S)$. Poiché V è uno spazio di Hausdorff i punti sono chiusi e quindi per regolarità esisterà un aperto bilanciato $U \subseteq V$ contenente l'origine che non interseca $F(S)$, di conseguenza $U' = F^{-1}(U)$ sarà un aperto di $\mathbb{K}^n$ che contiene l'origine ma non interseca S.

Ma F è lineare e quindi anche U' sarà bilanciato, e dato che non può intersecare S dovrà essere interamente contenuto in B. Allora

$$F(B) \supseteq F(U') = U$$

poiché sia V che $\mathbb{K}^n$ sono spazi vettoriali topologici ed F è lineare seguirà che F è una funzione aperta e perciò F è un omeomorfismo. ■

Corollario 4.6.2. *Se V è uno spazio vettoriale topologico di Hausdorff finitamente generato allora $V^* = V^+$ e $\dim V = \dim V^* = \dim V^+$.*

Dimostrazione. Prendiamo un qualunque isomorfismo $I : V \to \mathbb{K}^n$, per ogni $f \in V^+$ abbiamo che $f' = f \circ I^{-1}$ è lineare da $\mathbb{K}^n$ in $\mathbb{K}$ e perciò è anche continua, ma allora $f = f' \circ I$ è continua da V in $\mathbb{K}$ dunque $f \in V^*$. ∎

Corollario 4.6.3. *Sia V spazio vettoriale topologico di Hausdorff e W sottospazio di V finitamente generato, allora W è chiuso in V.*

Dimostrazione. Utilizziamo la stessa notazione della dimostrazione precedente sostituendo chiaramente V con W e ponendo U aperto di V tale che $U' = F^{-1}(U \cap W)$. Sia $x \in \mathrm{Cl}(W)$ innanzitutto esisterà $t > 0$ tale che $x \in tU$ e quindi

$$x \in \mathrm{Cl}(W) \cap tU \subseteq \mathrm{Cl}(W \cap tU) \subseteq \mathrm{Cl}(F(tD))$$

Ma tD è compatto in $\mathbb{K}^n$ e per continuità allora $F(tD)$ è chiuso in V in quanto spazio di Hausdorff, avremo così che $x \in F(tD) \subseteq W$ e $\mathrm{Cl}(W) = W$. ∎

Osservazione. Nella dimostrazione abbiamo utilizzato questa inclusione: in uno spazio topologico X se A è sottoinsieme generico e B è un aperto di X allora

$$\mathrm{Cl}(A) \cap B \subseteq \mathrm{Cl}(A \cap B)$$

il lettore può dimostrarlo senza alcuna difficoltà.

In $\mathbb{K}^n$ sappiamo che un insieme è compatto se e solo se è chiuso e limitato, in particolare la chiusura di aperti limitati è compatta. Anche in alcuni spazi metrici la chiusura delle palle aperte è compatta oppure solo per alcune di esse, invece negli spazi vettoriali topologici questo non è mai possibile, a meno che lo spazio non abbia dimensione finita.

Teorema 4.6.4. *Se V è uno spazio vettoriale topologico di Hausdorff ed esiste un aperto U contenente l'origine tale che* $\mathrm{Cl}(U)$ *è compatto allora V è finitamente generato.*

Dimostrazione. Per il teorema 3.3.13 e per il lemma 4.2.5 segue che $2^{-n}U$ forma un sistema fondamentale di intorni dell'origine. Per compattezza inoltre esistono $x_1, \dots, x_k \in \mathrm{Cl}(U)$ tali che

$$\mathrm{Cl}(U) \subseteq \left(x_1 + \frac{1}{2}U\right) \cup \cdots \cup \left(x_k + \frac{1}{2}U\right)$$

e definiamo $W = \mathrm{span}\left\{x_n \;\middle|\; 1 \le i \le k\right\}$. Per la proposizione precedente W è chiuso in V, inoltre essendo $U \subseteq W + (1/2)U$ avremo che

$$U \subseteq W + \frac{1}{2}U \subseteq W + \frac{1}{2}W + \frac{1}{4}U = W + \frac{1}{4}U$$

essendo W un sottospazio vettoriale e quindi $U \subseteq W + 2^{-n}U$, ma allora

$$U \subseteq \bigcap_{n \in \mathbb{N}} [W + 2^{-n}U] = \mathrm{Cl}(W) = W$$

in quanto $2^{-n}U$ è un sistema fondamentale di intorni dell'origine, per il lemma 4.2.5 segue che $V = W$. ∎

4.7 Somma continua

Nel capitolo introduttivo abbiamo esteso il concetto di somma diretta anche a spazi vettoriali generici per evidenziarne la struttura interna. In altre parole uno spazio vettoriale X si dice somma diretta dei suoi sottospazi $\{Y_a\}_{a\in A}$ se e solo se X è generato dagli Y_a ed $Y_a \cap \operatorname{span}\left(\bigcup_{b\neq a} Y_b\right) = \{0\}$, in tal caso possiamo sempre trovare un'applicazione lineare biettiva

$$F : X \to \bigoplus_{a\in A} Y_a$$

che manda Y_a in $\tilde{Y}_a$ e viceversa. Questa funzione però non è quasi mai continua né tanto meno possiede inversa continua, in questa sezione ci dedicheremo quindi a studiare delle condizioni su X e sugli Y_a in modo che F sia un omeomorfismo.

Considereremo solo il caso A finito (per comodità porremo $A = \{1,2,\dots,n\}$), in questo modo l'inversa F^{-1} avrà la forma

$$F^{-1} : (x_1,\dots,x_n) \in X_1 \times \cdots \times X_n \to x_1 + x_2 + \cdots + x_n \in X$$

ragionando per induzione su n ed applicando la continuità della somma non è difficile dimostrare la continuità di F^{-1}. In generale diremo che uno spazio vettoriale topologico X è **somma continua** dei sottospazi $Y_1, Y_2,\dots, Y_n$ se e solo se X è somma diretta di tali sottospazi e l'applicazione F è continua (e quindi è un omeomorfismo).

Osservazione. Presi degli spazi vettoriali topologici generici Y_i su $\mathbb{K}$ al variare di $1 \leq i \leq n$ il nuovo spazio vettoriale $\bigoplus_{i=1}^n Y_i$ è chiaramente somma continua dei suoi sottogruppi $\tilde{Y}_i$ per come abbiamo definito la topologia prodotto. Per questo motivo la somma diretta di spazi vettoriali topologici la possiamo indicare anche con il simbolo $\tilde{\bigoplus}_i Y_i$.

Per la proposizione 2.7.3 X è somma continua se e solo se le proiezioni canoniche $p_a = \pi_a \circ F$ di X su Y_a sono tutte quanti continue. In questo modo otteniamo un'altra caratterizzazione della somma continua:

Proposizione 4.7.1. *Siano X spazio vettoriale topologico ed $Y_1, Y_2,\dots,Y_n \leq X$ tali che $Y_1 + Y_2 + \cdots + Y_n = X$. Allora X è somma continua dei sottospazi Y_i se e solo se per ogni net $x_{i,a} \in Y_i$ su A vale la seguente doppia implicazione*

$$x_{1,a} + x_{2,a} + \cdots + x_{n,a} \to 0 \Leftrightarrow x_{i,a} \to 0 \text{ per ogni } 1 \leq i \leq n$$

in altre parole la somma converge a 0 *se e solo se tutte le sue componenti vi convergono singolarmente.*

Dimostrazione. Poniamo $y_a = \sum_{i=1}^n x_{i,a} \in X$, se X è somma continua degli Y_i allora $p_i(y_a) = x_{i,a}$ e la tesi segue dalla continuità delle proiezioni.

Viceversa osserviamo innanzitutto che X deve essere somma diretta degli Y_i, altrimenti potremmo trovare delle net costanti ma non tutte nulle la cui somma fa 0, in netto contrasto con l'implicazione da sinistra a destra. Di conseguenza esistono le proiezioni p_i in Y_i e sono univocamente determinate.

L'ipotesi di convergenza per net che abbiamo imposto implica immediatamente che tali proiezioni p_i sono continue nell'origine, per linearità saranno allora continue su tutto lo spazio e così X è somma continua. ■

Consideriamo adesso un qualunque sottospazio H di X diverso da $\{0\}$ e da X, per la proposizione 1.8.3 esiste automaticamente $K < X$ (che chiameremo **supplemento algebrico** di H) tale che X è somma diretta di H e K, di conseguenza esisteranno le proiezioni lineari p_H ed p_K con $p_K = \mathrm{id}_X - p_H$. Perciò se X è uno spazio vettoriale topologico ed H un suo qualunque sottospazio avremo che X è somma continua di H e del sottospazio K definito precedentemente se e solo se la proiezione lineare di X su H è anche continua.

Questa osservazione è molto utile, in quanto permette di concentrarsi sulle proprietà topologiche dei sottospazi vettoriali presi singolarmente invece che cercare tra le interazioni tra essi. In altre parole introdotta la seguente proprietà dei sottospazi di uno spazio vettoriale topologico:

Definizione 4.7.2. Dati uno spazio vettoriale topologico X ed un suo sottospazio H, diremo che il sottospazio $K \leq X$ definito precedentemente è il **supplemento topologico** di H se e solo se X è somma continua di H e K.

per quanto detto precedentemente abbiamo la seguente caratterizzazione della somma continua:

Proposizione 4.7.3. *Se X è somma diretta di un numero finito di suoi sottospazi $Y_1, \ldots, Y_n$ allora ne è anche somma continua se e solo se ogni Y_i possiede un supplemento topologico in X.*

Supponiamo ora che il sottospazio non banale H di X possiede supplemento topologico e dunque la sua proiezione p è continua. dato che $p(x) = x$ se e solo se $x \in H$ possiamo scrivere

$$\begin{aligned} H &= \ker\left(\mathrm{id}_X - p\right) \\ K &= \ker p \end{aligned}$$

dove K è il supplemento algebrico di H. Dunque se X fosse uno spazio di Hausdorff il sottospazio H ed il suo supplemento algebrico devono essere necessariamente chiusi. Per un generico spazio vettoriale topologico di Hausdorff questa condizione è necessaria ma non sufficiente, per avere la sufficienza sono richieste condizioni molto più stringenti:

Lemma 4.7.4. *Se X è un F-spazio allora tutti i suoi sottospazi vettoriali chiusi con supplementi algebrici anch'essi chiusi possiedono supplementi topologici.*

Dimostrazione. Supponiamo che H sia somma diretta di H e K sottospazi chiusi, dimostriamo che la proiezione p_H è continua sfruttando il teorema del grafico chiuso in quanto X è anche uno spazio metrico completo, più precisamente ci basterà dimostrare che presi $x_n, x, y \in X$ tali che $x_n \to x$, $p_H(x_n) \to y$ allora $p_H(x) = y$. Dato che sia H che K sono chiusi abbiamo

$$\begin{aligned} y &= \lim_{n\to+\infty} p_H(x_n) \in H \\ x - y &= \lim_{n\to+\infty} \left[x_n - p_H(x_n)\right] = \lim_{n\to+\infty} p_K(x_n) \in K \end{aligned}$$

e così $y \in H$ e $x - y \in K$.

Ora poiché $x = y + (x - y)$ ed essendo X somma diretta di H e K la scrittura di x come somma di un elemento di H e di un elemento di K è unica, perciò $p_H(x) = y$ e $p_K(x) = x - y$, il teorema del grafico chiuso ci garantisce così che p_H è continua e perciò X è somma continua di H e K. ∎

Osservazione. Il lemma precedente non dice che se X è somma diretta di 3 o più sottospazi chiusi allora è anche somma continua degli stessi, in quanto abbiamo visto che la somma di sottospazi chiusi in generale non è chiusa. Per ognuno di tali sottospazi bisognerà comunque verificare se il proprio supplemento algebrico è chiuso.

Ricapitolando se X è uno spazio vettoriale topologico e $Y_1, Y_2, \dots, Y_n$ sono suoi sottospazi allora le seguenti affermazioni sono tutte equivalenti tra loro:

- X è somma continua degli Y_i;
- esistono, e sono univocamente determinate, delle proiezioni continue $p_i : X \to Y_i$ tali che $\sum_{i=1}^n p_i(x) = x$;
- X coincide con la somma degli Y_i e per ogni net $x_{i,a} \in Y_i$ sull'insieme diretto A vale la doppia implicazione
$$x_{1,a} + x_{2,a} + \cdots + x_{n,a} \to 0 \Leftrightarrow x_{i,a} \to 0 \text{ per ogni } 1 \leq i \leq n$$
- X è somma diretta degli Y_i ed ogni sottospazio Y_i possiede un supplemento topologico in X.

Inoltre se X è un F-spazio si aggiunge anche la seguente affermazione equivalente:

- X è somma diretta degli Y_i e per ogni i sia Y_i che il suo supplemento algebrico sono chiusi in X.

Capitolo 5

Convessità

In questo capitolo studieremo le proprietà di convessità e di conicità di sottoinsiemi di spazi vettoriali topologici e di come interagiscono con la topologia.

5.1 Insiemi convessi

Definizione 5.1.1. Un sottoinsieme K di un $\mathbb{K}$-spazio vettoriale V, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, è **convesso** se e solo se per ogni $x, y \in K$ e per ogni $\lambda \in [0,1]$ si ha $\lambda x + (1-\lambda)y \in K$ o equivalentemente

$$tK + (1-t)K \subseteq K$$

per ogni $t \in [0,1]$.

Diremo invece che K è un **cono** se e solo se per ogni $x \in K$ e per ogni $\lambda > 0$ si ha $\lambda x \in K$ ovvero

$$\lambda K = K$$

per ogni $\lambda > 0$.

In questo capitolo ci dedicheremo principalmente allo studio degli insiemi convessi e delle loro proprietà all'interno degli spazi vettoriali topologici.

Proposizione 5.1.2. *Un sottoinsieme K di V è convesso se e solo se per ogni α, β reali positivi si ha*

$$\alpha K + \beta K = (\alpha+\beta)K \tag{5.1.1}$$

Dimostrazione. Se K soddisfa la (5.1.1) allora sarà chiaramente convesso. Viceversa se K è convesso allora

$$\alpha K + \beta K = (\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta}K + \frac{\beta}{\alpha+\beta}K\right) \subseteq (\alpha+\beta)K$$

l'inclusione opposta vale in generale. ■

In uno spazio vettoriale topologico il corollario 4.1.13 nel capitolo precedente permette di affermare tra l'altro che la chiusura e l'interno di un insieme convesso/cono, se non vuoti, è ancora un insieme convesso/cono. Una conseguenza importante di questa osservazione è data dalla seguente proposizione

Proposizione 5.1.3. *Ogni intorno convesso dell'origine contiene un intorno aperto convesso bilanciato dell'origine.*

Dimostrazione. Dato che l'interno di un convesso è ancora convesso possiamo dimostrare l'asserto per un aperto U convesso contenente l'origine. Posto

$$G = \left\{u \in U \,\middle|\, \exists W \text{ aperto contenente } u \text{ tale che } \lambda W \subseteq U\ \forall\, |\lambda| \leq 1\right\}$$

è immediato verificare che G è un aperto bilanciato contenente l'origine (si veda la proposizione 4.1.8) dimostriamo che è anche convesso.

Presi $x, y \in G$ esisteranno due aperti U_1, U_2 contenenti x e y rispettivamente tali che $\lambda U_1, \lambda U_2 \subseteq U$ per ogni $|\lambda| \leq 1$, dalla definizione di convessità si avrà per ogni $t \in [0,1]$ e $|\lambda| \leq 1$

$$\lambda[tU_1 + (1-t)U_2] = t(\lambda U_1) + (1-t)(\lambda U_2) \subseteq U$$

ed essendo $tU_1 + (1-t)U_2$ un aperto contenente $tx + (1-t)y$ segue che anche G è convesso. ■

Definizione 5.1.4. Dato un qualunque sottoinsieme $P \subseteq V$ di uno spazio vettoriale reale V definiamo l'**inviluppo convesso** di P e il **cono generato da** P gli insiemi

$$\operatorname{con}(P) = \bigcap \{C \subseteq V \mid P \subseteq C \text{ e } C \text{ è convesso}\}$$
$$\operatorname{rad}(P) = \bigcap \{C \subseteq V \mid P \subseteq C \text{ e } C \text{ è un cono}\}$$

Se V è anche uno spazio vettoriale topologico definiamo anche i seguenti insiemi

$$\overline{\operatorname{con}}(P) = \bigcap \left\{C \subseteq V \mid P \subseteq C \text{ e } C \text{ è convesso e chiuso}\right\}$$
$$\overline{\operatorname{rad}}(P) = \bigcap \left\{C \subseteq V \mid P \subseteq C \text{ e } C \text{ è un cono chiuso}\right\}$$

Grazie al corollario 4.1.13 avremo che con, $\overline{\text{con}}$ sono insiemi convessi e rad, $\overline{\text{rad}}$ sono coni. Abbiamo inoltre che

Proposizione 5.1.5. *Per ogni $P \subseteq V$ non vuoto si ha*

$$\operatorname{con}(P) = \left\{\sum_{i=1}^{n} \lambda_i x_i \,\middle|\, n \in \mathbb{N}, x_i \in P, \lambda_i \in [0,1] \text{ e } \sum_{i=1}^{n} \lambda_i = 1\right\} \tag{5.1.2}$$

$$\operatorname{rad}(P) = \left\{\lambda x \,\middle|\, x \in P, \lambda > 0\right\} = \bigcup_{\lambda > 0} \lambda P \tag{5.1.3}$$

Dimostrazione. Posto

$$P_1 = \left\{\sum_{i=1}^{n} \lambda_i x_i \,\middle|\, n \in \mathbb{N}, x_i \in P, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1\right\}$$
$$P_2 = \left\{\lambda x \,\middle|\, x \in P, \lambda > 0\right\}$$

dimostriamo innanzitutto che P_1 è convesso e che P_2 è un cono. Presi $x = \sum_i \lambda_i x_i \in P_1$, $y = \sum_j \mu_j y_j \in P_1$ con $x_i, y_j \in P$ come nella definizione osserviamo che per ogni $t \in [0,1]$ abbiamo che $\sum_i t\lambda_i + \sum_j (1-t)\mu_j = 1$ e perciò

$$tx + (1-t)y = \sum_i t\lambda_i x_i + \sum_j (1-t)\mu_j y_j \in P_1$$

per definizione di P_1, dunque P_1 è convesso e contiene P. Da ciò segue immediatamente che $\operatorname{con}(P) \subseteq P_1$, viceversa preso un qualunque $x \in P_1$ esistono $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ tali che $\lambda_1 + \cdots + \lambda_n = 1$ e

$$x \in \lambda_1 P + \cdots + \lambda_n P \subseteq \lambda_1 \operatorname{con}(P) + \cdots + \lambda_n \operatorname{con}(P) = \left(\lambda_1 + \cdots + \lambda_n\right) \operatorname{con}(P) = \operatorname{con}(P)$$

e perciò $P_1 = \operatorname{con}(P)$.

Per quanto riguarda P_2 chiaramente è un cono e $\operatorname{rad}(P) \subseteq P_2$ e con un procedimento analogo al precedente si verifica facilmente che $\operatorname{rad}(P) = P_2$. ■

Corollario 5.1.6. *Sia V vettoriale topologico e $P \subseteq V$ non vuoto, allora*

$$\overline{\operatorname{con}}(P) = \operatorname{Cl}(\operatorname{con}(P))$$
$$\overline{\operatorname{rad}}(P) = \operatorname{Cl}\left(\operatorname{rad}(P)\right)$$

Dimostrazione. Dimostriamo solo l'uguaglianza dell'inviluppo convesso, in quanto quella del cono generato è uguale.

Innanzitutto $\operatorname{con}(P) \subseteq \overline{\operatorname{con}}(P)$ poiché è intersezione di un maggior numero di insiemi, per la definizione di chiusura allora $\operatorname{Cl}(\operatorname{con}(P)) \subseteq \overline{\operatorname{con}}(P)$. La disuguaglianza opposta è una conseguenza del fatto che la chiusura di insiemi convessi è ancora convessa. ■

Proposizione 5.1.7. *Sia V spazio vettoriale topologico di Hausdorff e un suo sottoinsieme finito $P = \left\{x_1, x_2, \dots, x_n\right\} \subseteq V$. Allora*

$$\operatorname{con}(P) = \overline{\operatorname{con}}(P)$$

e sono entrambi sia compatti che sequenzialmente compatti.

Dimostrazione. Dalla proposizione 5.1.5 e dalla finitezza di P abbiamo

$$\operatorname{con}(P) = \left\{ \sum_{i=1}^{n} \lambda_i x_i \;\middle|\; \lambda_i \in [0,1] \wedge \sum_{i=1}^{n} \lambda_i = 1 \right\}$$

dunque $\operatorname{con}(P)$ è contenuto in un sottospazio finitamente generato $R = \operatorname{span} P$ di V, e poiché V è uno spazio di Hausdorff per il teorema 4.6.1 R con la topologia indotta sarà omeomorfo a $\mathbb{R}^n$, in particolare i sottoinsiemi sequenzialmente compatti sono chiusi e compatti.

Dimostriamo che $\operatorname{con}(P)$ è sequenzialmente compatto in R. Per ogni successione $y_k \in \operatorname{con}(P)$ di suoi elementi possiamo determinare $\lambda_{i,k} \in [0,1]$ tali che $y_k = \sum_{i=1}^{n} \lambda_{i,k} x_i$ e quindi posso sceglierne un'estratta affinché i coefficienti convergano, ottenendo così un'estratta convergente. Infine la relazione $\sum_{i=1}^{n} \lambda_i = 1$ che è soddisfatta per ogni termine della successione per continuità è soddisfatta anche dal limite e quindi il limite dell'estratta giace ancora nello stesso insieme.

Dalla definizione di topologia indotta esisterà un chiuso C di V tale che $\operatorname{con}(P) = C \cap R$ ma per il corollario 4.6.3 R è chiuso in V dunque $\operatorname{con}(P)$ è chiuso e perciò $\overline{\operatorname{con}}(P) \subseteq \operatorname{con}(P)$. La disuguaglianza opposta è banale. ■

Proposizione 5.1.8. *Se V è uno spazio vettoriale reale e $K \subseteq V$ è un cono convesso allora*

$$\operatorname{span} K = K - K$$

Dimostrazione. Innanzitutto $0 \in K - K$ e $K \subseteq K - K$ in quanto $x = 2x - x \in K - K$ per ogni $x \in K$. Ancora $\alpha(K-K) = (-\alpha)(K-K) \subseteq K - K$ per ogni $\alpha \in \mathbb{R}$, dunque per ogni $\alpha, \beta \in \mathbb{R}$

$$\alpha(K-K) + \beta(K-K) = (|\alpha| + |\beta|)(K-K) = K - K$$

da cui segue che $K - K$ è un sottospazio vettoriale contenente K.

Prendiamo adesso un qualunque sottospazio L contenente K allora $-K \subseteq L$ e perciò $K - K \subseteq L$ e per l'arbitrarietà di L segue che $\operatorname{span} K = K - K$. ■

Proposizione 5.1.9. *Per ogni convesso C di V l'insieme* $\operatorname{rad}(C)$ *è un cono convesso. Invece se C fosse un cono allora* $\operatorname{con}(C)$ *è ancora un cono convesso.*

Dimostrazione. Prendiamo $x, y \in \operatorname{rad}(C) = \bigcup_{\lambda>0} \lambda C$ allora esistono $\alpha, \beta > 0$ tali che $x \in \alpha C, y \in \beta C$. Ma allora per ogni $\lambda \in [0,1]$ per la convessità di C segue che

$$\lambda x + (1-\lambda) y \in \lambda\alpha C + (1-\lambda)\beta C \subseteq [\lambda\alpha + (1-\lambda)\beta] C \subseteq \bigcup_{\lambda>0} \lambda C$$

e quindi è ancora convesso.

Poiché riscalamenti di insiemi convessi è ancora convesso si ha

$$\operatorname{con}(C) = \bigcap\{K \text{ convesso} \mid C \subseteq K\} = \bigcap\{\lambda K \text{ convesso} \mid C \subseteq K\} = \lambda \operatorname{con}(C)$$

per ogni $\lambda > 0$ e dunque $\operatorname{con}(C)$ è ancora un cono. ■

Definizione 5.1.10. Dato un insieme P di V definiamo

$$\operatorname{dir} P = \left\{ v \in V \,\middle|\, x + \lambda v \in P \; \forall x \in P, \lambda \geq 0 \right\}$$

Chiaramente $\operatorname{dir} P$ è un cono convesso tale che per ogni $C \subseteq V$ vale la seguente equivalenza

$$C \subseteq \operatorname{dir} p \Leftrightarrow P + \mu C \subseteq P \; \forall \mu > 0$$

infatti per ogni $\lambda > 0$ valgono le relazioni seguenti:

$$P + \mu(\lambda \operatorname{dir} P) = P + (\mu\lambda) \operatorname{dir} P \subseteq P$$
$$P + \mu \operatorname{dir} P + \mu \operatorname{dir} P \subseteq P + \mu \operatorname{dir} P \subseteq P$$

da cui discendono rispettivamente la positiva omogeneità e la convessità.

Osserviamo inoltre che

$$v \in \operatorname{dir} P \Leftrightarrow v \in \frac{P - x}{\lambda} \forall x \in P, \lambda > 0$$

e quindi

$$\operatorname{dir} P = \bigcap_{\lambda>0} \bigcap_{x \in P} \frac{P-x}{\lambda} \tag{5.1.4}$$

Possiamo però migliorare ancora questa stima quando P è convesso limitandoci a prendere $\lambda \leq 1$, infatti vale

Proposizione 5.1.11. *Se P è un qualunque sottoinsieme di V non vuoto allora*

$$\operatorname{dir} P = \bigcap_{\lambda \in [0,1]} \bigcap_{x \in P} \frac{P-x}{\lambda}$$

Dimostrazione. Posto $P' = \bigcap_{\lambda\in(0,1]}\bigcap_{x\in P}\frac{P-x}{\lambda}$ per ogni $y \in P'$ avremo che $y + x \in P$. Facciamo vedere per induzione che
$$ny + x \in P$$
per ogni $n \in \mathbb{N}$ e per ogni $x \in P$. Quando $n = 1$ è banale altrimenti supponiamolo vero per $n-1 \geq 1$, allora $(n-1)y + x \in P$ per ogni $x \in P$ e
$$ny + x = y + [(n-1)y + x] \in P$$
Prendiamo ora $\lambda > 0$ generico, esisterà $n \in \mathbb{N}$ tale che $\lambda' = \lambda - n \in (0,1]$ e quindi per ipotesi $\lambda' y + x \in P$ per ogni $x \in P$ e
$$\lambda y + x = ny + (\lambda' y + x) \in P$$
dunque $P' = \operatorname{dir} P$. ■

Proposizione 5.1.12. *Se C è un insieme convesso contenente l'origine allora $x \in \operatorname{dir} C$ se e solo se $\lambda x \in C$ per ogni $\lambda > 0$.*

Dimostrazione. Se $x \in \operatorname{dir} C$ segue direttamente dalla definizione che $\lambda x \in C$ per ogni $\lambda > 0$ in quanto $0 \in C$. Viceversa avremo per ogni $y \in C$
$$2\lambda x + y \in C + C \subseteq 2C$$
e quindi per l'arbitrarietà di y e λ avremo che
$$x \in \bigcap_{y\in V}\bigcap_{\lambda>0}\frac{C - y/2}{\lambda} = \operatorname{dir} C$$
e il teorema è così verificato. ■

5.2 Epigrafici

Adesso che abbiamo introdotto il concetto di convessità per gli insiemi definiamo un concetto analogo per le funzioni. Innanzitutto per qualunque funzione $f : V \to \overline{\mathbb{R}}$ definiamo il **dominio** di f come
$$\operatorname{dom} f = \left\{x \in V \,\middle|\, \left|f(i)\right| < +\infty\right\}$$
utilizziamo lo stesso simbolo del dominio di un multivettore in quanto le due nozioni coincidono quando il codominio è $\mathbb{R}$. Diremo inoltre che f è **propria** su V se e solo se

- $f(x) > -\infty$ per ogni $x \in K$;
- $f(x) < +\infty$ per qualche $x \in K$.

Per ogni funzione $f : V \to \overline{\mathbb{R}}$ definiamo l'**epigrafico aperto** e l'**epigrafico chiuso** di f come
$$\operatorname{epi}_c f = \left\{(x,r) \in V \times \mathbb{R} \,\middle|\, f(x) \leq r\right\}$$
$$\operatorname{epi}_o f = \left\{(x,r) \in V \times \mathbb{R} \,\middle|\, f(x) < r\right\}$$
Utilizzeremo entrambe le definizioni di epigrafico, così da poter utilizzare la più adatta a seconda della situazione. Osserviamo innanzitutto che se $(x,r) \in \operatorname{epi}_c f$ per qualche $x \in$

$V, r \in \mathbb{R}$ allora $(x, r+s) \in \mathrm{epi}_c f$ per ogni $s > 0$ e allo stesso modo si ha per $\mathrm{epi}_o f$. Dunque il nostro epigrafico (aperto e chiuso) è composto interamente da semirette verticali in $V \times \mathbb{R}$ con la coda che punta a $+\infty$, questa caratterizzazione ci permette di dare una definizione più generale di epigrafico. Diremo allora che un sottoinsieme E di $V \times \mathbb{R}$ è un **epigrafico** se e solo se

$$E + \vec{\varepsilon} \subseteq E$$

per ogni $\varepsilon > 0$ dove si è posto $\vec{\varepsilon} = (0, \varepsilon) \in V \times \mathbb{R}$.

Osserviamo inoltre che per ogni $\varepsilon > 0$

$$\mathrm{epi}_c f + \varepsilon \subseteq \mathrm{epi}_o f \subseteq \mathrm{epi}_c f$$
$$\mathrm{epi}_o f = \bigcup_{\varepsilon > 0} \left(\mathrm{epi}_c f + \vec{\varepsilon}\right)$$

Dati adesso due epigrafici $D, E \subseteq V \times \mathbb{R}$ definiamo la seguente relazione di equivalenza

$$D \sim E \Leftrightarrow D + \vec{\varepsilon} \subseteq E \text{ e } E + \vec{\varepsilon} \subseteq D \ \forall \varepsilon > 0$$

grazie alla relazione precedente avremo che $\mathrm{epi}_c f \sim \mathrm{epi}_o f$ per ogni funzione $f : V \to \overline{\mathbb{R}}$.

Proposizione 5.2.1 (Proprietà degli epigrafici)**.** *Dati due epigrafici $D, E \subseteq V \times \mathbb{R}$ con $D \sim E$ allora*

1. *Se $D \subseteq L \subseteq E$ con L generico allora L è un epigrafico e $L \sim D \sim E$;*
2. *$D + C \sim E + C$ per ogni epigrafico C e $\lambda D \sim \lambda E$ per ogni $\lambda > 0$;*
3. *Se $D_i \sim E_i$ per ogni $i \in I$ allora $\bigcap_{i \in I} D_i \sim \bigcap_{i \in I} E_i$ e $\bigcup_{i \in I} D_i \sim \bigcup_{i \in I} E_i$;*
4. $\mathrm{con}(D) \sim \mathrm{con}(E)$, $\mathrm{rad}(D) \sim \mathrm{rad}(E)$ *e* $\mathrm{dir}(D) \sim \mathrm{dir}(E)$.

Dimostrazione. Dimostriamo innanzitutto il punto 1. Preso un qualunque $(x, r) \in L \subseteq E$ poiché $E \sim D$ per ogni $\varepsilon > 0$ avremo $(x, r+\varepsilon) \in D \subseteq L$ e L è un epigrafico. Da ciò segue anche che $D + \vec{\varepsilon} \subseteq L + \vec{\varepsilon} \subseteq L$, viceversa avremo che $L + \vec{\varepsilon} \subseteq E + \vec{\varepsilon} \subseteq D$ e dunque $L \sim D \sim E$.

Per quanto riguarda il punto 2 abbiamo per ogni $\varepsilon > 0$

$$D + C + \vec{\varepsilon} = D + \vec{\varepsilon} + C \subseteq E + C$$
$$\lambda D + \vec{\varepsilon} = \lambda \left(D + \frac{\vec{\varepsilon}}{\lambda}\right) \subseteq \lambda E$$

e per simmetria segue che $D + C \sim E + C$ e $\lambda D \sim \lambda E$.

Analogamente si dimostra il punto 3, il punto 4 deriva direttamente dai punti precedenti, dalla proposizione 5.1.5, dalla (5.1.3) e dalla (5.1.4). ■

Dato un qualunque epigrafico T definiamo

$$T_o = \bigcup_{\varepsilon > 0} T + \vec{\varepsilon}$$
$$T_c = \bigcap_{\varepsilon > 0} T - \vec{\varepsilon}$$

chiaramente se $R \sim T$ allora $T_o \subseteq R \subseteq T_c$, in qualche modo T_c, T_o fungono da chiusura e interno di T. Sempre per le relazioni precedenti abbiamo $(\mathrm{epi}_c f)_o = \mathrm{epi}_o f$ e $(\mathrm{epi}_o f)_c = \mathrm{epi}_c f$.

Proposizione 5.2.2. *Dato un epigrafico convesso/conico T di $V \times \mathbb{R}$ allora anche T_c e T_o sono epigrafici convessi/conici.*

Dimostrazione. Direttamente dal corollario 4.1.13 in quanto la famiglia di insiemi $T + \vec{r}$ è monotona al variare di $r \in \mathbb{R}$ dato che T è un epigrafico. ■

Il seguente risultato permette di identificare le funzioni con gli epigrafici

Teorema 5.2.3 (Identificazione)**.** *Dati due epigrafici T, Q su $V \times \mathbb{R}$, allora sono equivalenti:*

1. $T \sim Q$;
2. $T_c = Q_c$ *e* $T_o = Q_o$;
3. $T_o \subseteq Q_c$ *e* $Q_o \subseteq T_c$.

Dimostrazione. Se $T \sim Q$ allora abbiamo

$$\begin{aligned} T_c &= \bigcap_{\varepsilon>0} T - \vec{\varepsilon} \subseteq \bigcap_{\varepsilon>0} Q - 2\vec{\varepsilon} = Q_c \\ T_o &= \bigcup_{\varepsilon>0} T + 2\vec{\varepsilon} \subseteq \bigcup_{\varepsilon>0} Q + \vec{\varepsilon} = Q_o \end{aligned}$$

e per simmetria abbiamo così che $T_c = Q_c$ e $T_o = Q_o$.

Adesso supponiamo che $T_o \subseteq Q_c$ e $Q_o \subseteq T_c$, preso un qualunque $\varepsilon > 0$ abbiamo che

$$T + \vec{\varepsilon} \subseteq T_o \subseteq Q_c \subseteq Q - \vec{\varepsilon}$$

e perciò $T + 2\vec{\varepsilon} \subseteq Q$ per ogni $\varepsilon > 0$, sempre per simmetria possiamo scambiare T con Q e perciò $T \sim Q$. ■

Corollario 5.2.4. *Dato un qualunque epigrafico T abbiamo che $T \sim T_c \sim T_o$ e*

$$\begin{aligned} T_{cc} &= T_{oc} = T_c \\ T_{oo} &= T_{co} = T_o \end{aligned}$$

inoltre $T \sim R$ se e solo se $T_o \subseteq R \subseteq T_c$.

Daremo ora una definizione un po' più generale di funzione convessa e di funzione positivamente omogenea, definizione utilizzata anche in [12].

Definizione 5.2.5. Diremo che la funzione $f : V \to \overline{\mathbb{R}}$ è **convessa** se e solo se esiste un insieme vuoto o convesso $K \subseteq V \times \mathbb{R}$ tale che

$$\mathrm{epi}_o f \subseteq K \subseteq \mathrm{epi}_c f$$

Diremo invece che f è **quasiconvessa** se e solo se per ogni $\alpha \in \mathbb{R}$ esiste un insieme convesso $C_\alpha \subseteq V$ tale che

$$\{f < \alpha\} \subseteq C_\alpha \subseteq \{f \leq \alpha\}$$

La funzione f è invece detta **positivamente omogenea** se e solo se esiste un cono $C \subseteq K \times \mathbb{R}$ tale che $\mathrm{epi}_o f \subseteq C \subseteq \mathrm{epi}_c f$.

Il vantaggio principale di questa definizione di convessità rispetto a quella usuale è che permette di considerare anche funzioni che assumono i valori $\pm\infty$ senza andare in contro a forme indeterminate. Inoltre si vede facilmente che tutte le funzioni convesse sono anche quasiconvesse in quanto

$$\begin{aligned}\{f \le \alpha\} &\cong \text{epi}_c f \cap H_\alpha \\ \{f < \alpha\} &\cong \text{epi}_o f \cap H_\alpha \\ C_\alpha &\cong K \cap H_\alpha\end{aligned}$$

dove $H_\alpha = \{(x,\alpha) \mid x \in V\}$.

Questa definizione di convessità per le funzioni si adatta perfettamente anche alle funzioni definite solamente su un sottoinsieme convesso di V, in tal caso estendendola a $+\infty$ su tutti gli altri punti otterremo una nuova funzione definita su V ma con lo stesso epigrafico della funzione di partenza.

Teorema 5.2.6. *Data una qualunque funzione $f : V \to \overline{\mathbb{R}}$ sono equivalenti*

1. *f è convessa;*
2. *$\text{epi}_o f$ è vuoto o convesso;*
3. *$\text{epi}_c f$ è vuoto o convesso;*
4. *Per ogni $x, y \in V$ e per ogni $\lambda \in [0,1]$ si ha*

$$f[\lambda x + (1-\lambda)y] \le \lambda f(x) + (1-\lambda) f(y) \tag{5.2.1}$$

ponendo però $+\infty - \infty = +\infty$.

Dimostrazione. L'equivalenza tra le prime tre affermazioni deriva dal teorema di identificazione degli epigrafici. Consideriamo adesso una funzione convessa f a valori in $\overline{\mathbb{R}}$, innanzitutto $\text{epi}_c f$ è vuoto se e solo se $f(x) = +\infty$ per ogni $x \in V$, quindi supponiamo che contenga almeno un elemento. Dati $x, y \in V$ se $f(x) = +\infty$ o $f(y) = +\infty$ allora la disuguaglianza (5.2.1) sarà sempre verificata, altrimenti per ogni $a > f(x), b > f(y)$ finiti avremo che $(x,a),(y,b) \in \text{epi}_o f$ e così

$$[\lambda x + (1-\lambda)y, \lambda a + (1-\lambda)b] \in \text{epi}_o f$$

per ogni $\lambda \in [0,1]$, da cui segue immediatamente la (5.2.1).

Viceversa se f soddisfa la (5.2.1) prendiamo due qualunque elementi $(x,a),(y,b)$ di $\text{epi}_c f$ allora $f(x), f(y) < +\infty$ e

$$f[\lambda x + (1-\lambda)y] \le \lambda f(x) + (1-\lambda) f(y) \le \lambda a + (1-\lambda) b$$

perciò $\lambda(x,a) + (1-\lambda)b \in \text{epi}_c f$ ed f è così convessa. ■

Questo risultato ci permette di introdurre una nozione più stringente di convessità. Diremo che $f : K \to \mathbb{R}$ è **strettamente convessa** se e solo se presi $x, y \in K$ distinti e $\lambda \in (0,1)$ si ha

$$f[\lambda x + (1-\lambda)y] < \lambda f(x) + (1-\lambda) f(y)$$

Corollario 5.2.7 (Disuguaglianza di Jensen)**.** *Data una funzione convessa $f : V \to \overline{\mathbb{R}}$ per ogni $x_1, x_2, \dots, x_k \in K$ e $\lambda_1, \dots, \lambda_k \geq 0$ con $\sum_{i=1}^{k} \lambda_i = 1$ si avrà*

$$f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i)$$

Esempio 5.2.8. Prendiamo una costante $p > 1$ e consideriamo la funzione

$$F(t) = t^p$$

definita per ogni $t \geq 0$, dimostriamo che F è strettamente convessa, ovvero

$$[\lambda s + (1-\lambda)t]^p < \lambda s^p + (1-\lambda)t^p$$

per ogni $0 < s < t$ e per ogni $0 < \lambda < 1$. Innanzitutto abbiamo $F''(t) = p(p-1)t^{p-2} > 0$ per ogni $t > 0$ e quindi la derivata prima è strettamente crescente, consideriamo adesso due casi

1. Se $t > s$ allora
$$t^p - s^p = \int_s^t p u^{p-1}\, du > \int_s^t p s^{p-1}\, du = p s^{p-1}(t-s)$$
2. Se $t < s$ allora
$$t^p - s^p = -\int_t^s p u^{p-1}\, du > -\int_t^s p s^{p-1}\, du = p s^{p-1}(t-s)$$

Posto $r = \lambda s + (1-\lambda)t$ con $0 < \lambda < 1$ allora r sarà diverso sia da t che da s e dunque

$$\begin{aligned} t^p &> r^p + p r^{p-1}(t-r) \\ s^p &> r^p + p r^{p-1}(s-r) \end{aligned}$$

da cui segue che
$$\lambda s^p + (1-\lambda)t^p > [\lambda s + (1-\lambda)t]^p$$
e quindi è strettamente convessa.

In maniera analoga si dimostra un risultato di equivalenza analogo per la quasiconvessità e per la positiva omogeneità:

Proposizione 5.2.9. *Data una funzione $f : V \to \overline{\mathbb{R}}$ sono equivalenti*

1. *f è quasiconvessa;*
2. *$\{f \leq \alpha\}$ è convesso per ogni $\alpha \in \mathbb{R}$;*
3. *$\{f < \alpha\}$ è convesso per ogni $\alpha \in \mathbb{R}$;*
4. *Per ogni $x, y \in K$ e per ogni $t \in [0,1]$ si ha*
$$f[tx + (1-t)y] \leq f(x) \vee f(y) \tag{5.2.2}$$

Proposizione 5.2.10. *Data una funzione $f : V \to \overline{\mathbb{R}}$ sono equivalenti*

1. *f è positivamente omogenea;*
2. $\mathrm{epi}_o f$ *è un cono;*
3. $\mathrm{epi}_c f$ *è un cono;*
4. *per ogni $x \in V$ e per ogni $\lambda > 0$ si ha $f(\lambda x) = \lambda f(x)$.*

Esempio 5.2.11. Dato uno spazio vettoriale reale V e una metrica d su V invariante e convessa (secondo la definizione data nel capitolo degli spazi vettoriali topologici) allora l'applicazione

$$x \in V \to d(x,0) \in [0,+\infty)$$

è una funzione quasiconvessa ma non necessariamente convessa, basta prendere ad esempio su $\mathbb{R}$

$$d(x,y) = \sqrt{|x-y|}$$

Una funzione $f : K \to \overline{\mathbb{R}}$ definita su un convesso K di V è **strettamente quasiconvessa** se e solo se la disuguaglianza (5.2.2) è stretta quando $x \neq y$ e $t \neq 0,1$. Chiaramente tutte le funzioni strettamente convesse sono anche strettamente quasiconvesse.

Per ogni sottoinsieme non vuoto K di V definiamo le seguenti funzioni:

- la funzione **indicatrice** $i_K : V \to \mathbb{R}^l$ definita come

$$i_K(x) = \begin{cases} 0 & \text{se } x \in K \\ +\infty & \text{se } x \notin K \end{cases}$$

- la funzione **supporto** $s_{K,F} : W \to \mathbb{R}^l$ definita a partire da un'applicazione bilineare $F : V \times W \to \mathbb{R}$ tra spazi vettoriali reali come

$$s_{K,F}(a) = \sup \{F(x,a) \mid x \in K\}$$

- il **funzionale di Minkowski** $p_K : V \to \overline{\mathbb{R}}$ definito in modo tale che

$$p_K(x) = \inf\{\lambda > 0 \mid x \in \lambda K\}$$

Queste funzioni ci saranno utili in varie occasioni più avanti. Innanzitutto osserviamo che il funzionale di Minkowski è una funzione positivamente omogenea mentre la funzione supporto è convessa per qualunque scelta di $K \subseteq V$ grazie al seguente risultato:

Proposizione 5.2.12. *Data una famiglia di funzioni convesse $\{f_i\}_{i \in I}$ da V in $\overline{\mathbb{R}}$ allora la funzione*

$$F(x) = \sup_{i \in I} f_i(x) \in \overline{\mathbb{R}}$$

è convessa su V.

Dimostrazione. Deriva immediatamente dall'uguaglianza

$$\mathrm{epi}_c F = \bigcap_{i \in I} \mathrm{epi}_c f_i$$

■

Finora abbiamo costruito epigrafici a partire da funzioni ben definite per studiarne le proprietà, da questo momento faremo l'operazione inversa ovvero costruiremo funzioni a partire da epigrafici generici. Per ogni sottoinsieme T di $V\times\mathbb{R}$ definiamo la funzione $INF(T)$ definita su V a valori in $\overline{\mathbb{R}}$ in modo tale che

$$INF(T)(x) = \begin{cases} +\infty & \text{se } (x,r)\notin T \text{ per ogni } r\in\mathbb{R} \\ \inf\left\{r\in\mathbb{R} \;\middle|\; (x,r)\in T\right\} & \text{altrimenti} \end{cases} \tag{5.2.3}$$

per ogni $x\in V$. Chiaramente se poniamo $\bar{T} = \bigcup_{R\geq 0} T + \vec{R}$ avremo che $INF(T) = INF(\bar{T})$ e dunque possiamo supporre senza perdere in generalità che T sia un epigrafico.

Proposizione 5.2.13. *Dato un qualunque epigrafico $T\subseteq V\times\mathbb{R}$ allora la funzione $INF(T): V\to\overline{\mathbb{R}}$ definita in* (5.2.3) *soddisfa le seguenti uguaglianze*

$$\begin{aligned} \operatorname{epi}_o INF(T) &= T_o \\ \operatorname{epi}_c INF(T) &= T_c \end{aligned} \tag{5.2.4}$$

In particolare $INF(T) = INF(Q)$ se e solo se $T\sim Q$.

Dimostrazione. Poniamo per comodità $f = INF(T)$, allora per ogni $(x,r)\in V\times\mathbb{R}$ tale che $f(x)\leq r$ esisterà una successione $r_n\in\mathbb{R}$ tale che $(x,r_n)\in T$, $r_n\geq r$ e $r_n\to r$. Ora per ogni $\varepsilon>0$ esiste $n\in\mathbb{N}$ tale che $r_n - r<\varepsilon$ e quindi

$$(x,r) = (x,r_n) + (0, r-r_n)\in T + (0,r-r_n)\subseteq T-\vec{\varepsilon}$$

e perciò $(x,r)\in T_c$.

Viceversa se $(x,r)\in T_c$ allora $(x,r+\varepsilon)\in T$ per ogni $\varepsilon>0$ e perciò $f(x)\leq r+\varepsilon$. Per l'arbitrarietà di ε avremo però che $f(x)\leq r$ e $(x,r)\in\operatorname{epi}_c f$, così $\operatorname{epi}_c f = T_c$. Con un procedimento analogo si verifica che $\operatorname{epi}_o f = T_o$. ■

Lemma 5.2.14. *Date inoltre due funzioni $\mathfrak{h}: V^k\to V$ e $\mathfrak{n}:\overline{\mathbb{R}}^k\to\overline{\mathbb{R}}$ che soddisfano le seguenti proprietà*

- *$\mathfrak{n}(a_1,\dots,a_k) = +\infty$ se e solo se $a_i = +\infty$ per almeno un $1\leq i\leq k$, mentre se abbiamo $\mathfrak{n}(a_1,\dots,a_k) = -\infty$ allora $a_i = -\infty$ per almeno un i;*
- *se $b_i\geq a_i$ per ogni i allora $\mathfrak{n}(b_1,\dots,b_k)\geq\mathfrak{n}(a_1,\dots,a_k)$ e inoltre per ogni $a_i\in[-\infty,+\infty)$*

$$\inf\left\{\mathfrak{n}(b_1,\dots,b_k) \;\middle|\; b_i>a_i\ \forall i\right\} = \mathfrak{n}(a_1,\dots,a_k)$$

- *$\mathfrak{h}$ e $\mathfrak{n}$ sono suriettive.*

Allora posto $\mathfrak{g}((x_1,r_1),\dots,(x_k,r_k)) = [\mathfrak{h}(x_1,\dots,x_k),\mathfrak{n}(r_1,\dots,r_k)]$ si ha

$$INF[\mathfrak{g}(C_1,\dots,C_k)](x) = \inf\left\{\mathfrak{n}[INF(C_1)(x_1),\dots,INF(C_k)(x_k)] \;\middle|\; \mathfrak{h}(x_1,\dots,x_k) = x\right\} \tag{5.2.5}$$

per ogni $C_1,\dots,C_k\subseteq V\times\mathbb{R}$.

Dimostrazione. Il primo membro della (5.2.5) è $+\infty$ se e solo se per ogni $r \in R$ si ha $(x, r) \notin \mathfrak{g}(C_1, \dots, C_k)$ ovvero per ogni $r_1, \dots, r_k \in \mathbb{R}$ deve essere $[x, \mathfrak{n}(r_1, \dots, r_k)] \notin \mathfrak{g}(C_1, \dots, C_k)$ e questo grazie alla suriettività di $\mathfrak{n}$. Ma questo è equivalente ad affermare che per ogni scelta di $(x_1, \dots, x_k) \in \mathfrak{h}^{-1}(x)$ esista un i tale che $(x_i, l) \notin C_i$ per ogni $l \in \mathbb{R}$ ovvero $INF(C_i)(x_i) = +\infty$ per almeno un i, e questo è equivalente ad affermare che il secondo membro della (5.2.5) sia $+\infty$. Da questo punto supporremo allora che tutte le quantità in gioco non siano $+\infty$.

Posto $t_i = INF(C_i)(x_i)$ allora per ogni $\varepsilon > 0$ esisteranno $r_i \in \mathbb{R}$ sufficientemente vicini a t_i, ovvero

$$t_i \leq r_i < t_i + \varepsilon \text{ se } t_i \in \mathbb{R}$$
$$r_i < -\frac{1}{\varepsilon} \text{ se } t_i = -\infty$$

tali che $(x_i, r_i) \in C_i$ per ogni i e di conseguenza $[x, \mathfrak{n}(r_1, \dots r_k)] \in \mathfrak{g}(C_1, \dots, C_k)$. Abbiamo così

$$INF\left[\mathfrak{g}(C_1, \dots, C_k)\right]\left[\mathfrak{h}(x_1, \dots, x_k)\right] \leq \mathfrak{n}(r_1, \dots, r_k)$$

e quindi facendo tendere ε a zero per la continuità destra di $\mathfrak{n}$ avremo dimostrato la disuguaglianza $\leq$ nella (5.2.5).

Viceversa preso un qualunque $x \in V$ posto $t = INF\left[\mathfrak{g}(C_1, \dots, C_k)\right](x)$ per ogni $\varepsilon > 0$ esisteranno $x_i \in V$, $r \in \mathbb{R}$ sufficientemente vicino a t e $r_i \in \mathbb{R}$ tali che

$$(x_i, r_i) \in C_i$$
$$x = \mathfrak{h}(x_1, \dots, x_k)$$
$$r = \mathfrak{n}(r_1, \dots, r_k)$$

e quindi avremo che

$$\inf\left\{\mathfrak{n}[INF(C_1)(y_1), \dots, INF(C_k)(y_k)] \;\middle|\; x = \mathfrak{h}(y_1, \dots, y_k)\right\}$$
$$\leq \mathfrak{n}[INF(C_1)(x_1), \dots, INF(C_k)(x_k)] \leq \mathfrak{n}(r_1, \dots, r_k) = r$$

Poiché il membro più a sinistra non dipende da ε facendolo tendere a 0 sempre per la continuità destra otteniamo la disuguaglianza opposta e così l'uguaglianza. ■

Proposizione 5.2.15. *Presa una famiglia di sottoinsiemi $\{C_i\}_{i\in I}$ di $V \times \mathbb{R}$ si avranno*

$$INF\left(\bigcup_{i\in I} C_i\right) = \inf_{i\in I} INF(C_i)(x) \tag{5.2.6}$$

$$INF\left(\bigcap_{i\in I} C_i\right) = \sup_{i\in I} INF(C_i)(x) \tag{5.2.7}$$

Dimostrazione. Osserviamo che la disuguaglianza $\leq$ è banale, prendiamo allora un qualunque $x \in V$ se $INF(C_i)(x) = +\infty$ per ogni $i \in I$ allora (x, r) non apparterrà a nessun c_i qualunque sia $r \in R$ e di conseguenza $INF\left(\bigcup_{i\in I} C_i\right)(x) = +\infty$, per questo supponiamo che il secondo membro sia finito.

Per ogni $\varepsilon > 0$ esisterà $i \in I$ tale che $(x, t + \bar{\varepsilon}) \in C_i$ dove $t = INF\left(\bigcup_{i\in I} C_i\right)(x)$ e $0 \leq \bar{\varepsilon} < \varepsilon$ e perciò

$$\inf_{j\in I} INF(C_j)(x) \leq INF(C_i)(x) \leq t + \bar{\varepsilon}$$

e facendo tendere ε a 0 si ha la tesi. L'ultima uguaglianza si dimostra analogamente. ■

Corollario 5.2.16. *Date due funzioni f e g definite su V a valori in $\mathbb{R}^l$ e $\varepsilon > 0$ si ha*

- $INF(\{(a,r)\}) = i_a + r$ *per ogni* $r \in \mathbb{R}$;
- $INF(\mathrm{epi}_o f) = INF(\mathrm{epi}_c f) = f$ *per il teorema di identificazione;*
- $INF(\mathrm{epi}_c f + \mathrm{epi}_c g)(x) = \inf\{f(y) + g(x-y) \mid y \in V\}$*:*
- $INF(\lambda\,\mathrm{epi}_c f)(x) = \lambda f(\lambda^{-1}x)$*;*
- $INF[\mathrm{con}(\mathrm{epi}_c f)](x) = \inf\left\{\sum_{i=1}^k \lambda_i f(x_i) \mid k \in \mathbb{N}, x = \sum_{i=1}^k \lambda_i x_i\right\}$*;*
- $INF\left(\mathrm{rad}(\mathrm{epi}_c f)\right)(x) = \inf_{\lambda>0} \lambda f(\lambda^{-1}x)$.

e inoltre abbiamo

$$INF\left[\mathrm{con}\left(\bigcup_{i\in I} \mathrm{epi}_c f_i\right)\right](x) = \inf\left\{\sum_{j=1}^k \lambda_j f_{i_j}(x_j) \;\middle|\; i_j \in I, x = \sum_{j=1}^k \lambda_j x_j, \lambda_i \geq 0, \sum_i \lambda_i = 1\right\}$$

Grazie alla proposizione 5.1.11 se f è anche propria allora avremo che

$$INF(\mathrm{dir}\,\mathrm{epi}_c f)(x) = INF\left(\bigcap_{y\in\mathrm{dom} f, \lambda\in(0,1]} \frac{\mathrm{epi}_c f - [y, f(y)]}{\lambda}\right) = \sup_{\substack{y\in\mathrm{dom} f\\ \lambda\in(0,1]}} \frac{f(y+\lambda x) - f(y)}{\lambda}$$

e se f fosse convessa e $\mathrm{dom} f \neq \emptyset$ allora possiamo porre $\lambda = 1$ in quanto per convessità avremo che

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

oppure possiamo applicare la proposizione 5.1.12 all'insieme convesso $\mathrm{epi}_c f - [x_0, f(x_0)]$ se $x_0 \in \mathrm{dom} f$, in tal caso avremo

$$\begin{aligned} INF(\mathrm{dir}\,\mathrm{epi}_c f)(x) &= \sup_{\lambda>0} \lambda^{-1}[f(x_0+\lambda x) - f(x_0)] \\ &= \lim_{\lambda\to+\infty} \lambda^{-1}[f(x_0+\lambda x) - f(x_0)] \\ &= \lim_{\lambda\to+\infty} \lambda^{-1} f(x_0+\lambda x) \end{aligned} \tag{5.2.8}$$

dove la seconda uguaglianza discende dal seguente risultato:

Proposizione 5.2.17. *Data una funzione convessa $f : V \to \overline{\mathbb{R}}$ allora per ogni $x \in \mathrm{dom} f, v \in V$ l'applicazione*

$$\lambda \in \mathbb{R}^+ \to \frac{f(x+\lambda v) - f(x)}{\lambda} \in \mathbb{R}^l$$

è crescente. Inoltre se f fosse strettamente convessa e $v \neq 0$ allora anche la funzione precedente sarà strettamente crescente.

Dimostrazione. Presi $0 < \lambda < \mu$ abbiamo

$$\begin{aligned} \frac{f(x+\lambda v) - f(x)}{\lambda} &= \frac{f\left[\frac{\lambda}{\mu}(x+\mu v) + \left(1-\frac{\lambda}{\mu}\right)x\right] - f(x)}{\lambda} \\ &\leq \frac{\frac{\lambda}{\mu} f(x+\mu v) + \left(1-\frac{\lambda}{\mu}\right) f(x) - f(x)}{\lambda} = \frac{f(x+\mu v) - f(x)}{\mu} \end{aligned}$$

la disuguaglianza sarà stretta quando f è strettamente convessa e $v \neq 0$ chiaramente. ■

Definizione 5.2.18. Date due funzioni f, g definite su tutto V a valori in $\overline{\mathbb{R}}$ definiamo la **convoluzione additiva** la funzione $f \diamond g$ definita su V che vale

$$f \diamond g = INF(\operatorname{epi}_c f + \operatorname{epi}_c g)$$

Inoltre per ogni $\lambda \geq 0$ definiamo l'**omotetia** di f mediante λ la funzione

$$f_\lambda = INF(\lambda \operatorname{epi}_c f)$$

in particolare si ha $f_0 = i_{\{0\}}$, mentre definiamo l'**inviluppo omogeneo** e l'**inviluppo convesso** di f rispettivamente come

$$\operatorname{rad} f = INF[\operatorname{rad}(\operatorname{epi}_c f)]$$
$$\operatorname{con} f = INF[\operatorname{con}(\operatorname{epi}_c f)]$$

Poniamo anche per ogni funzione f propria

$$\operatorname{dir} f = INF[\operatorname{dir}(\operatorname{epi}_c f)]$$

Combinando il lemma 5.2.14 e il teorema di identificazione avremo in particolare che la convoluzione additiva è associativa e commutativa, e tutte le funzioni precedenti saranno convesse/positivamente omogenee se le funzioni in gioco lo sono.

Proposizione 5.2.19. *Se f è una funzione convessa propria allora per ogni $x \in V$ e per ogni $x_0 \in \operatorname{dom} f$*

$$\begin{aligned}\operatorname{dir} f(x) &= \lim_{\lambda \to +\infty} \lambda^{-1} f(x_0 + \lambda x) \\ &= \sup_{y \in \operatorname{dom} f} [f(x+y) - f(y)]\end{aligned}$$

Dimostrazione. Direttamente dalla (5.2.8) e da quanto detto precedentemente. ∎

Proposizione 5.2.20. *Date due funzioni generiche f e g da V in $\mathbb{R}$ allora*

$$\operatorname{dom}(f \diamond g) \subseteq \operatorname{dom} f + \operatorname{dom} g$$

e

$$\operatorname{dom} f_\lambda \subseteq \lambda \operatorname{dom} f$$

Dimostrazione. Innanzitutto $x \in \operatorname{dom}(f \diamond g)$ se e solo se esistono $r, \in \mathbb{R}$ tali che $(x, r) \in E_f + E_g$ e $(x, s) \notin E_f + E_g$. Esisteranno dunque $(u, a) \in E_f$, $(v, b) \in E_g$ tali che $x = u + v$, inoltre $u \in \operatorname{dom} f$ altrimenti $(u, l) \in E_f$ per ogni $l \in \mathbb{R}$ e perciò

$$(x, s) = (u, s - b) + (v, b) \in E_f + E_g$$

assurdo. In modo analogo si verifica che $v \in \operatorname{dom} g$ e perciò $x \in \operatorname{dom} f + \operatorname{dom} g$.

Il procedimento per l'omotetia è analogo. ∎

Osservazione. Per ogni $f : V \to \mathbb{R}^l$ e per ogni $a \in V$ si ha

$$(f \diamond i_a)(x) = f(x - a)$$

come per la convoluzione usuale.

5.3 Spazi localmente convessi

In generale non è detto che l'origine in uno spazio vettoriale topologico generico possieda sempre un sistema fondamentale di intorni convesso (benché l'intero spazio sia esso stesso un aperto convesso), ed infatti possiamo trovare facilmente spazi vettoriali topologici che non possiedono alcun sistema fondamentale di intorni convessi. Per questa ragione introduciamo la nozione di spazio localmente convesso:

Definizione 5.3.1. Uno spazio vettoriale topologico si dice **localmente convesso** se e solo se esiste una base per la topologia composta interamente da insiemi aperti convessi.

Si osserva che grazie alla proposizione 5.1.3 ogni spazio vettoriale topologico localmente convesso possiede una base di aperti convessi bilanciati. Un concetto estremamente importante negli spazi vettoriali topologici localmente convessi è il concetto di **seminorma**, ovvero una qualunque funzione convessa $p : V \to \mathbb{R}$ definita su uno spazio vettoriale V sul campo $\mathbb{K}$ tale che $p(\lambda x) = |\lambda|\, p(x)$ per ogni $\lambda \in \mathbb{K}$. In particolare le seminorme sono anche positivamente omogenee.

Data una qualunque seminorma p su V osserviamo che essa è un'applicazione sublineare: infatti per ogni coppia di vettori $v, w \in V$ si ha

$$p(x+y) = p\left(\frac{2x+2y}{2}\right) \leq \frac{1}{2}p(2x) + \frac{1}{2}p(2y) = p(x) + p(y)$$

e inoltre $p(0) = p(0v) = 0p(v) = 0$. Da quest'ultima relazione segue che per ogni $v \in V$

$$0 = p(v-v) \leq p(v) + p(-v) = 2p(v)$$

e quindi $p(v) \geq 0$ ovvero le seminorme sono funzioni non negative. In generale però potrebbero esistere dei vettori v non nulli tali che $p(v) = 0$.

Proposizione 5.3.2. *Le seminorme sono pseudometriche invarianti bilanciate e regolari sugli scalari.*

Dimostrazione. Dobbiamo dimostrare solamente la proprietà di continuità del prodotto esterno per una qualunque seminorma p su V. Se prendiamo $\lambda_i, \lambda \in \mathbb{K}$ tali che $\lambda_i \to \lambda$ ed $x_i, x \in V$ con $p\left(x_i - x\right) \to 0$ allora avremo che

$$p\left(\lambda_i x_i - \lambda x\right) \leq \left|\lambda_i\right| p\left(x_i - x\right) + \left|\lambda_i - \lambda\right| p(x) \leq M p\left(x_i - x\right) + \left|\lambda_i - \lambda\right| p(x) \to 0$$

per qualche $M > 0$. ∎

La differenza più evidente tra le seminorme e le pseudometriche generiche è che per ogni seminorma $p : V \to \mathbb{R}$ gli insiemi nella forma

$$C_{x,r} = \left\{y \in V \,\middle|\, p(y-x) < r\right\}$$

sono a tutti gli effetti insiemi convessi: se prendiamo $a, b \in C_{x,r}$ allora

$$\begin{aligned} p[\lambda a + (1-\lambda)b - x] = p[\lambda(a-x) + (1-\lambda)(b-x)] \leq p[\lambda(a-x)] + p[(1-\lambda)(b-x)] \\ = \lambda p(a-x) + (1-\lambda)p(b-x) < r \end{aligned}$$

e dunque $\lambda a + (1-\lambda)b \in C_{x,r}$.

Il teorema 4.2.10 diventerà quindi

Teorema 5.3.3. *Dato uno spazio vettoriale V su $\mathbb{K}$ munito della topologia generata da una qualunque famiglia di seminorme $\mathscr{F} = \{p_\alpha\}_{\alpha\in A}$. Allora V è uno spazio vettoriale topologico localmente convesso e se $\mathscr{F}$ è totale allora V è uno spazio di Hausdorff.*

Consideriamo adesso un generico insieme aperto convesso bilanciato U contenente l'origine, per ogni $n \in \mathbb{N}$ poniamo

$$U_n = 2^{-n} U$$

per la (5.1.1) segue che $U_{n+1} + U_{n+1} = U_n$ e dunque possiamo applicare il teorema 4.2.11 per trovare una pseudometrica invariante $D : V \to [0,1]$ su V. Però per come abbiamo definito gli insiemi U_n abbiamo per ogni $r \in \mathbb{Q}$

$$\mathcal{U}(r) = rU$$

e quindi, sfruttando anche la densità di $\mathbb{Q}$ in $[0,1]$, possiamo definire D anche nella seguente maniera:

$$D(x) = \begin{cases} 1 & \text{se } x \notin U \\ \inf\{r > 0 \mid x \in rU\} & \text{altrimenti} \end{cases}$$

ovvero $D(x) = p_U(x) \wedge 1$. Concentriamoci da questo momento sulle funzioni nella forma p_U quando U non è necessariamente un aperto ma soddisfa una condizione significativamente più debole.

Definizione 5.3.4. Sia V spazio vettoriale generico e $K \subseteq V$, un punto $v \in V$ è **internale** per K se e solo se per ogni $u \in V$ esiste $\delta_0 > 0$ tale che per ogni $\delta_0 > \delta \geq 0$ si ha $v + \delta u \in K$. I punti di V che non sono internali né a K né a $V \setminus K$ sono detti **frontalieri** per K.

Proposizione 5.3.5. *In uno spazio vettoriale topologico V su $\mathbb{K}$ tutti i punti interni di un sottoinsieme $K \subseteq V$ sono internali per K mentre tutti i punti sulla frontiera per K sono frontalieri.*

Dimostrazione. L'applicazione

$$h : \delta \in \mathbb{K} \to x + \lambda v \in V$$

è continua per ogni $x, v \in V$, quindi essendo K un intorno di $x = h(0)$ per continuità esiste un $\delta > 0$ tale che per ogni $\lambda \in \mathbb{K}$ con $|\lambda| \leq \delta$ si ha $x + \lambda v = h(\lambda) \in K$ e in particolare per λ reale positivo. Quindi x è un punto internale.

Ripetendo lo stesso ragionamento per $V \setminus K$ si dimostra l'asserto per i punti di frontiera. ■

La definizione di punto internale è indipendente dalla topologia scelta su V sullo spazio vettoriale. In generale però esistono punti internali che non sono però interni in quanto la definizione di interno di un insieme dipende dalla topologia.

Abbiamo definito nella sezione degli epigrafici il funzionale di Minkowski p_K per un certo sottoinsieme non vuoto K di V. È immediato constatare che $p_K(x) \geq 0$ e che $p_K(0) = 0$ se e solo se $0 \in K$ altrimenti $p_K(0) = +\infty$. Il prossimo risultato mostrerà altre proprietà del funzionali di Minkowski un po' meno elementari di queste.

Proposizione 5.3.6. *Sia p_K il funzionale di Minkowski di un convesso K che possiede* 0 *come punto internale. Allora*

1. *$p_K(v) < +\infty$ e $v \in tK$ per ogni $t > p_K(v)$;*
2. *Per ogni $\alpha > 0$ si ha $p_{\alpha K}(v) = \alpha^{-1} p_K(v)$;*
3. *Se K è bilanciato allora per ogni $\alpha \in \mathbb{K}$ vale $p_K(\alpha v) = |\alpha|\, p_K(v)$;*
4. *Se $v \in K$ allora $p_K(v) \leq 1$, se $v \in V \setminus K$ allora $p_K(v) \geq 1$;*
5. *$p_K(u+v) \leq p_K(u) + p_K(v)$, in particolare p_K è convessa;*
6. *Il punto v è internale a K se e solo se $p_K(v) < 1$;*
7. *Il punto v è frontaliero rispetto a K se e solo se $p_K(v) = 1$.*

In particolare se K è bilanciato allora p_K è una seminorma.

Dimostrazione. Analizziamo i vari casi:

1. Dalla definizione di internale esiste $\delta_0 > 0$ tale che per ogni $\delta_0 > \delta > 0$ si ha $\delta v \in K$ ovvero $v \in \delta^{-1} K$ e quindi $p_K(v)$ è finito.

   Per ogni $t > p_K(v)$ per definizione di estremo inferiore esiste $t_0 \in \left[p_K(v), t\right)$ tale che $v \in t_0 K$ e quindi dato che K è convesso e contiene l'origine avremo che

   $$t_0 K = t\left[\frac{t_0}{t} K + \left(1 - \frac{t_0}{t}\right) K\right] \subseteq tK$$

   e quindi $v \in tK$.

2. Ricordiamo che $v \in \alpha K$ se e solo se $\alpha^{-1} v \in K$. Abbiamo così che

   $$\begin{aligned} p_{\alpha K}(v) &= \inf\{t > 0 \mid v \in t(\alpha K)\} \\ &= \inf\{t > 0 \mid v \in (\alpha t) K\} \\ &= \inf\left\{\alpha^{-1} t' > 0 \,\middle|\, v \in t' K\right\} \\ &= \alpha^{-1} p_K(v) \end{aligned}$$

3. Dimostriamo che per ogni $\lambda \in \mathbb{K}$ con $|\lambda| = 1$ si ha $p_K(v) = p_K(\lambda v)$. Se $\alpha_n > 0$ è una successione minimizzante tendente ad $p_K(v)$ con $\alpha_n^{-1} v \in K$ allora per ipotesi $\alpha_n^{-1}(\lambda v) \in K$ per ogni $\lambda \in \mathbb{K}$ con $|\lambda| = 1$ e perciò $p_K(v) \geq p_K(\lambda v)$. Sostituendo v con λv e λ con λ^{-1} otteniamo l'uguaglianza.

4. La prima disuguaglianza è banale, se invece $v \in V \setminus K$ per il punto 1 esisterà un $\delta > 0$ tale che $v \in \delta K$. Per il punto 1 dato che $v \notin K$ deve essere $\delta > 1$ dunque $p_K(v) \geq 1$.

5. Prendiamo due successioni $\alpha_n, \beta_n > 0$ tendenti rispettivamente a $p_K(u)$ e $p_K(v)$ tali che $u \in \alpha_n K$, $v \in \beta_n K$. Dalla (5.1.1) segue che $u + v \in (\alpha_n + \beta_n) K$ e quindi $p_K(u+v) \leq \alpha_n + \beta_n$. Passando al limite avremo che $p_K(u+v) \leq p_K(u) + p_K(v)$.

6. Supponiamo che v sia un punto internale per K, esisterà un certo $\delta > 0$ tale che $v + \delta v \in K$ ovvero $v \in [1/(1+\delta)] K$ e perciò $p_K(v) < 1$.

   Viceversa sia $\epsilon \in \left(0, 1\right)$ tale che $p_K(v) < 1 - \epsilon$, per ogni $u \in V$ e $\delta > 0$ avremo allora che

   $$p_K(v + \delta u) \leq p_K(v) + \delta p_K(u) < 1 - \epsilon + \delta p_K(u)$$

Se $p_K(u) > 0$ poniamo

$$\delta_0 = \frac{\epsilon}{p_K(u)}$$

altrimenti scegliamo un $\delta_0 > 0$ arbitrario, in entrambi i casi per ogni $0 < \delta < \delta_0$ avremo che $v + \delta u \in \bar{\delta} K \subseteq K$ dove $\bar{\delta}$ è una costante minore di 1 dipendente da δ.

7. Dobbiamo solo dimostrare che v è internale a $V \setminus K$ se e solo se $p_K(V) > 1$. Se v è internale a $V \setminus K$ allora esiste $\delta \in (0,1)$ tale che $v + \delta(-v) \notin K$. Quindi per il punto 1 avremo che $p_K(v) \geq 1/(1-\delta) > 1$.

   Viceversa esiste $\epsilon > 0$ tale che $p_K(v) > 1 + \epsilon$, in tal caso per ogni $u \in V$ e $\delta > 0$ si ha, modificando opportunamente la disuguaglianza triangolare

   $$p_K(v + \delta u) = p_K(v - \delta(-u)) \geq p_K(v) - \delta p_K(-u) > 1 + \varepsilon - \delta p_K(-u)$$

   Se $p_K(-u) > 0$ poniamo

   $$\delta_0 < \frac{\epsilon}{p_K(-u)}$$

   mentre se $p_K(-u) = 0$ scegliamo un δ_0 arbitrario. Ragionando come nel punto precedente si dimostra che $v + \delta u \in V \setminus K$ per ogni $0 < \delta < \delta_0$ e quindi v è internale in $V \setminus K$.

■

Di conseguenza se V è uno spazio vettoriale topologico localmente convesso ed U è un intorno aperto convesso bilanciato dell'origine allora p_U è una seminorma su V. Quindi la funzione D ottenuta precedentemente è a tutti gli effetti la restrizione di una seminorma su U. Possiamo quindi modificare il teorema 4.2.11 sugli spazi vettoriali topologici localmente convessi nella seguente maniera:

Teorema 5.3.7 (Von Neumann)**.** *Consideriamo un generico spazio vettoriale topologico localmente convesso V con topologia τ, allora esiste una famiglia di seminorme $\{p_\alpha\}_{\alpha\in A}$ che la genera. Se V è uno spazio di Hausdorff allora tale famiglia di seminorme è anche totale.*

Prima di passare alla prossima sezione mostreremo altre proprietà degli spazi localmente convessi.

Proposizione 5.3.8. *Se V è uno spazio vettoriale topologico localmente convesso e $\{p_\alpha\}_{\alpha\in A}$ una famiglia di seminorme che genera la topologia di V. Allora un sottoinsieme K di V è limitato se e solo se per ogni $\alpha \in A$ esiste $T_\alpha > 0$ tale che $p_\alpha(x) < T_\alpha$ per ogni $x \in K$.*

Dimostrazione. Se K è uno spazio limitato allora per ogni $\alpha \in A$ esiste $T_\alpha > 0$ tale che $K \subseteq T_\alpha \cdot I(0, 1_\alpha)$ (qui 1_α è l'1-mutliindice definito sull'insieme $\{\alpha\}$ che vale 1) ovvero per ogni $x \in K$ si ha

$$p_\alpha\left(\frac{x}{T_\alpha}\right) < 1$$

e quindi si ha la tesi.

Viceversa supponiamo che $p_\alpha(x) < T_\alpha$ per ogni $\alpha \in A$ e per ogni $x \in K$, dimostriamo che per ogni multivettore $\boldsymbol{\varepsilon}$ esiste $T_{\boldsymbol{\varepsilon}} > 0$ tale che per ogni $|\lambda| > T_{\boldsymbol{\varepsilon}}$ si ha $K \subseteq \lambda I(0, \boldsymbol{\varepsilon})$. Essendo $\operatorname{dom} \boldsymbol{\varepsilon}$ finito definiamo

$$T_{\boldsymbol{\varepsilon}} = \max_{\alpha \in \operatorname{dom} \boldsymbol{\varepsilon}} \frac{T_\alpha}{\varepsilon_\alpha}$$

che sarà chiaramente finito, per ogni $x \in K$, $\alpha \in \operatorname{dom}\boldsymbol{\varepsilon}$ e per ogni $|\lambda| > T_{\varepsilon}$ si avrà allora

$$p_\alpha\left(\frac{x}{\lambda}\right) = \frac{p_\alpha(x)}{|\lambda|} < T_\alpha \cdot \frac{\varepsilon_\alpha}{T_\alpha} = \varepsilon_\alpha$$

ovvero $K \subseteq \lambda I(0, \boldsymbol{\varepsilon})$. ■

Proposizione 5.3.9. *Se V è uno spazio vettoriale topologico localmente convesso generato dalla famiglia di seminorme $\{p_\alpha\}_{\alpha\in A}$ allora x_n è di Cauchy se e solo se per ogni $\varepsilon > 0$ e per ogni $\alpha \in A$ esiste $N \in \mathbb{N}$ tale che per ogni $m, n > N$*

$$p_\alpha(x_m - x_n) < \varepsilon$$

All'inizio del capitolo abbiamo visto che la chiusura convessa di un sottoinsieme finito di uno spazio vettoriale topologico è sia compatto che sequenzialmente compatto in quanto è contenuto in un sottospazio vettoriale finitamente generato con una topologia meno fine di quella euclidea, il quale è a sua volta uno spazio localmente convesso. Il prossimo risultato generalizza questo risultato:

Lemma 5.3.10. *Se V è uno spazio vettoriale topologico localmente convesso allora per ogni $K \subseteq V$ totalmente limitato l'insieme* con$K$ *è ancora totalmente limitato in V.*

Dimostrazione. Sia $K \subseteq V$ totalmente limitato e $U \subseteq V$ intorno aperto dell'origine, essendo V localmente convesso esiste U' intorno aperto convesso dell'origine tale che $U' + U' \subseteq U$. Per definizione di totale limitatezza esisterà $F_1 \subseteq K$ finito tale che $K \subseteq F_1 + U'$, dato che la somma di convessi è convessa abbiamo

$$\operatorname{con}(K) \subseteq \operatorname{con}\left(F_1 + U'\right) \subseteq \operatorname{con}\left(F_1\right) + \operatorname{con}\left(U'\right) = \operatorname{con}\left(F_1\right) + U'$$

D'altronde per la 5.1.7 $\operatorname{con}\left(F_1\right)$ è compatto e dunque totalmente limitato, esisterà $F_2 \subseteq \operatorname{con}\left(F_1\right) \subseteq \operatorname{con}(K)$ finito tale che

$$\operatorname{con}(K) \subseteq F_2 + U' + U' \subseteq F_2 + U$$

e quindi $\operatorname{con}(K)$ è totalmente limitato. ■

Corollario 5.3.11. *Se V è localmente convesso e completo e K è totalmente limitato allora $\overline{\operatorname{con}}(K)$ è compatto e net compatto in V.*

Per quanto riguarda gli spazi vettoriali metrici diremo che una metrica d su uno spazio vettoriale V è convessa se e solo se tutte le palle aperte $B(x, r)$ sono insiemi convessi. Inoltre dato uno spazio vettoriale topologico $\mathscr{N}_1$ localmente convesso di Hausdorff possiamo supporre nella dimostrazione del teorema 4.2.11 che gli insiemi U_n siano tutti convessi. In tal modo la metrica d che si otterrà oltre ad essere bilanciata e invariante per traslazioni è anche convessa. In letteratura gli F-spazi localmente convessi sono detti **spazi di Fréchet** e comprendono buona parte degli spazi vettoriali topologici studiati in matematica, tra cui gli spazi normati.

Prima di passare al prossimo capitolo ci teniamo a far notare che V è uno spazio vettoriale metrico localmente convesso se e solo se la sua topologia può essere generata da una famiglia totale finita o numerabile di seminorme. Se indichiamo con $\{p_i\}_{i\in\mathbb{N}}$ una di queste famiglie allora la quantità

$$d(v, w) = \sum_{i=1}^{+\infty} \frac{1}{2^i} \frac{p_i(v - w)}{1 + p_i(v - w)}$$

è effettivamente una metrica su V compatibile con la sua topologia.

5.4 Il teorema di Hahn-Banach

Dato uno spazio vettoriale V consideriamo l'insieme $\mathscr{V}$ di tutte le seminorme su V, possiamo definire su $\mathscr{V}$ la seguente relazione d'ordine:

$$q \preceq p \Leftrightarrow q(x) \leq p(x) \ \forall x \in V$$

in modo da poter parlare di famiglie dirette di seminorme su V, sottointendendo che lo siano rispetto alla relazione $\preceq$.

Ogni spazio vettoriale topologico localmente convesso ammette una famiglia diretta di seminorme che lo genera. Infatti se $\mathscr{F} = \{p_\alpha\}_{\alpha \in A}$ è una famiglia di seminorme che genera la topologia di V allora per ogni $I \subseteq A$ finito non vuoto definiamo

$$p_I(x) = \max_{i \in I} p_i(x)$$

e per comodità di notazione poniamo $p_\emptyset \equiv 0$, allora $\mathscr{G} = \{p_I\}_{i \in \mathcal{F}(A)}$ è una famiglia diretta di seminorme su V.

Se indichiamo con τ e σ le topologie su V indotte rispettivamente da $\mathscr{F}$ e da $\mathscr{G}$ dimostriamo che $\tau = \sigma$. Innanzitutto per ogni multivettore $\boldsymbol{\varepsilon}$ su $\mathbb{R}^+$ con dominio in A (ovvero $\operatorname{dom} \boldsymbol{\varepsilon} \subseteq A$) possiamo definire il multivettore $\boldsymbol{\mu}$ su $\mathbb{R}^+$ con dominio in $\mathcal{F}(A)$ in modo tale che

$$i \in \operatorname{dom} \boldsymbol{\varepsilon} \Leftrightarrow \{i\} \in \operatorname{dom} \boldsymbol{\mu}$$
$$\varepsilon_i = \mu_{\{i\}}$$

e dato che $p_{\{i\}} = p_i$ allora avremo che $I(\mathscr{F}, x, \boldsymbol{\varepsilon}) = I(\mathscr{G}, x, \boldsymbol{\mu})$ e quindi $\tau \subseteq \sigma$.

Ora per ogni sottoinsieme finito I di A avremo che $p_I(x) < k$ se e solo se $p_i(x) < k$ per ogni $i \in I$. Sia ora $\boldsymbol{\mu}$ un multivettore su $\mathbb{R}^+$ con dominio in $\mathcal{F}(A)$ allora $\operatorname{dom} \boldsymbol{\mu}$ è un insieme finito i cui elementi sono a loro volta sottoinsiemi finiti di A. Definiamo allora un multivettore $\boldsymbol{\varepsilon}$ su $\mathbb{R}^+$ con dominio su A nella seguente maniera

$$\operatorname{dom} \boldsymbol{\varepsilon} = \bigcup_{I \in \operatorname{dom} \boldsymbol{\mu}} I$$
$$\varepsilon_i = \min \{\mu_I \mid I \in \operatorname{dom} \boldsymbol{\mu}, i \in I\}$$

allora per ogni $x \in V$ non è difficile dimostrare che

$$p_i(x) < \varepsilon_i \ \forall i \in \operatorname{dom} \boldsymbol{\varepsilon} \Leftrightarrow p_I(x) < \mu_I \ \forall I \in \operatorname{dom} \boldsymbol{\mu}$$

dunque $I(\mathscr{F}, x, \boldsymbol{\varepsilon}) = I(\mathscr{G}, x, \boldsymbol{\mu})$ e per l'arbitrarietà di $\boldsymbol{\mu}$ avremo che $\sigma \subseteq \tau$.

Corollario 5.4.1. *Una famiglia di seminorme $\mathscr{F} = \{p_\alpha\}_{\alpha \in A}$ su V è diretta se e solo se per ogni multivettore $\boldsymbol{\varepsilon}$ su $\mathbb{R}^+$ esiste $\gamma \in A$ che dipende solamente dall'insieme* $\operatorname{dom} \boldsymbol{\varepsilon}$ *tale che per ogni $x \in V$*

$$B\left(x, \min \boldsymbol{\varepsilon}, p_\gamma\right) \subseteq I(\mathscr{F}, x, \boldsymbol{\varepsilon})$$

in particolare gli insiemi $B(0, \varepsilon, p_\alpha)$ formano un sistema fondamentale di intorni aperti dell' origine al variare di $\alpha \in A$ e $\varepsilon > 0$.

Proposizione 5.4.2. *Siano V, W spazi vettoriali topologici localmente convessi le cui topologie sono generate rispettivamente dalle famiglie di seminorme $\mathscr{F} = \{p_\alpha\}_{\alpha \in A}$ e $\mathscr{G} = \left\{q_\beta\right\}_{\beta \in B}$ con $\mathscr{F}$ diretta.*

Allora ogni applicazione lineare $f : V \to W$ è continua se e solo se per ogni $\beta \in B$ esistono $\alpha \in A$ ed $M_{\alpha\beta} > 0$ tali che

$$q_\beta[f(x)] \leq M_{\alpha\beta} p_\alpha(x) \tag{5.4.1}$$

Osservazione. Grazie al ragionamento precedente la proposizione continua a valere anche se la famiglia $\mathscr{F}$ non è diretta. In questo caso però la tesi deve essere modificata così: per ogni $\beta \in B$ esistono $\alpha_1, \alpha_2, \ldots, \alpha_k \in A$ ed esiste $M > 0$ tali che

$$q_\beta[f(x)] \leq M\left[p_{\alpha_1}(x) \vee p_{\alpha_2}(x) \vee \cdots \vee p_{\alpha_k}(x)\right]$$

Dimostrazione della proposizione 5.4.2. Per ogni $\beta \in B$ poniamo

$$U = \left\{y \in W \middle|\ q_\beta(y) < 1\right\}$$

esso sarà un intorno aperto dell'origine, se f è continua allora dato che la famiglia p_α è diretta esisterà $\alpha \in A$ e $s_{\alpha\beta} > 0$ tali che

$$B\left(0, s_{\alpha\beta}, p_\alpha\right) \subseteq f^{-1}(U)$$

ovvero che per ogni $x \in V$ che soddisfa la disuguaglianza $p_\alpha(x) < s_{\alpha\beta}$ si avrà $q_\beta(f(x)) < 1$. Dunque per ogni $x \in V$ e per ogni $\varepsilon > 0$ si avrà

$$1 > q_\beta\left[f\left(s_{\alpha\beta}\frac{x}{p_\alpha(x)+\varepsilon}\right)\right] = \frac{s_{\alpha\beta}}{p_\alpha(x)+\varepsilon} q_\beta(f(x)) \Rightarrow q_\beta(f(x)) < M_{\alpha\beta}[p_\alpha(x)+\varepsilon]$$

dato che $M_{\alpha\beta}$ non dipende da ε facendolo tendere a 0 si ha la tesi.

Viceversa supponiamo che la f verifichi la (5.4.1) e scegliamo un qualunque intorno $I(\mathscr{G}, 0, \boldsymbol{\mu})$ in W, per ogni $\beta \in \operatorname{dom} \boldsymbol{\mu}$ esisterà $\alpha \in A$ e $M_{\alpha\beta}$ come in (5.4.1), definiamo allora un nuovo multivettore $\boldsymbol{\varepsilon}$ con dominio in A tale che

$$\varepsilon_{\alpha'} = \min_{\beta'} \frac{\mu_{\beta'}}{M_{\alpha'\beta'}}$$

dove β' varia tra tutti gli elementi in $\operatorname{dom} \boldsymbol{\mu}$ che soddisfano la (5.4.1) con $\alpha = \alpha'$ e $\beta = \beta'$. Dunque per ogni $x \in I(\mathscr{F}, 0, \boldsymbol{\varepsilon})$ avremo che

$$q_\beta(f(x)) \leq M_{\alpha\beta} p_\alpha(x) < M_{\alpha\beta}\varepsilon_\alpha \leq \mu_\beta$$

e quindi $f(x) \in I(\mathscr{G}, 0, \boldsymbol{\mu})$ ed f è così continua in 0, dunque per linearità è continua su tutto V. ■

Osservazione. Se la famiglia $\mathscr{F}$ non fosse diretta allora il risultato precedente continuerebbe a valere se riformuliamo la (5.4.1) nella seguente maniera: per ogni $\beta \in B$ esistono una costante $M_\beta > 0$ e un sottoinsieme finito I di A tali che $q_\beta(f(v)) \leq M_\beta p_\alpha(v)$ per ogni $v \in V$ e per ogni $\alpha \in I$.

Se V è uno spazio vettoriale topologico localmente convesso su $\mathbb{K}$ e $W = \mathbb{K}$ allora la famiglia $\mathscr{G}$ contiene solamente una seminorma (il valore assoluto per i numeri reali o il modulo per i complessi) e quindi la (5.4.1) diventa

$$\left|f(v)\right| \leq M p_\alpha(v)$$

per una qualche $\alpha \in A$ dipendente solo da f.

Enunciamo adesso il teorema di Hahn-Banach tramite il quale possiamo estendere una funzione lineare continua definita su un sottospazio di uno spazio vettoriale topologico localmente convesso sull'intero spazio.

Teorema 5.4.3 (Hahn-Banach analitico)**.** *Sia V uno spazio vettoriale su $\mathbb{R}$ e $W \leq V$ sottospazio vettoriale dotato della topologia indotta. Data una funzione lineare $f : W \to \mathbb{R}$ e una applicazione $p : V \to \mathbb{R}$ tali che*

$$\begin{gathered} p(x+y) \leq p(x) + p(y) \text{ per ogni } x, y \in V \\ p(tx) = t p(x) \text{ per ogni } t \geq 0 \\ f(x) \leq p(x) \text{ per ogni } x \in W \end{gathered}$$

allora esisterà un'altra funzione lineare $g : V \to \mathbb{K}$ tale che $f(x) = g(x)$ per ogni $x \in W$ e per ogni $x \in V$

$$-p(-x) \leq g(x) \leq p(x)$$

Dimostrazione. Dati W, f e p come nelle ipotesi del teorema definiamo la classe $\mathscr{L}$ di tutte le funzioni lineari e continue $g : W_g \to \mathbb{R}$ definite su un qualche sottospazio $W \leq W_g \leq V$ tali che $f(x) = g(x)$ per ogni $x \in W$ e $g(x) \leq p(x)$ per ogni $x \in W_g$.

Chiaramente $f \in \mathscr{L}$ e definiamo la seguente relazione d'ordine su $\mathscr{L}$:

$$g \preceq h \Leftrightarrow W_g \leq W_h \wedge g(x) = h(x) \text{ per ogni } x \in W_g$$

e verifichiamo che la classe $\mathscr{L}$ è induttiva rispetto a tale relazione d'ordine. Preso un qualunque sottoinsieme $\mathscr{I} \subseteq \mathscr{L}$ totalmente ordinato poniamo

$$\begin{gathered} \hat{W} = \bigcup_{g \in \mathscr{I}} W_g \\ \hat{g} : x \in W \to g(x) \text{ se } x \in W_g \text{ per qualche } g \in \mathscr{I} \end{gathered}$$

dimostriamo innanzitutto che $\hat{W}$ è un sottospazio vettoriale di V. Se $x \in \hat{W}$ allora $x \in W_g$ per un certo $g \in \mathscr{I}$ e quindi $\lambda x \in W_g \subseteq \hat{W}$ per ogni $\lambda \in \mathbb{R}$, se prendiamo anche $y \in \hat{W}$ allora esisterà $h \in \mathscr{I}$ tale che $y \in W_h$. Ora l'insieme $\mathscr{I}$ è totalmente ordinato e dunque possiamo supporre senza perdere in generalità che $g \preceq h$ e dunque $W_g \leq W_h$, questo significa che $x + y \in W_h \leq \hat{W}$ e abbiamo così dimostrato che $\hat{W}$ è un sottospazio vettoriale di V.

Dimostriamo ora che $\hat{g}$ è ben definita e lineare. Se $x \in W_g \cap W_h$ per due funzioni distinte $g, h \in \mathscr{I}$ allora o $g \preceq h$ oppure $h \preceq g$, in entrambi i casi avremo che $g(x) = h(x)$ e quindi $\hat{g}$ è ben definita. Prendiamo ora $x, y \in \hat{W}$, $\alpha, \beta \in \mathbb{R}$ per quanto detto nel paragrafo precedente esisterà $g \in \mathscr{I}$ tale che $x, y \in W_g$ e quindi

$$\hat{g}(\alpha x + \beta y) = g(\alpha x + \beta y) = \alpha g(x) + \beta g(y) = \alpha \hat{g}(x) + \beta \hat{g}(y)$$

e quindi $\hat{g}$ è lineare. Con lo stesso approccio si verifica che $\hat{g}(x) \leq p(x)$ u $\hat{W}$ e dunque $\hat{g} \in \mathscr{L}$ e $g \preceq \hat{g}$ per ogni $g \in \mathscr{I}$ dunque $\mathscr{L}$ è induttivo.

Dal lemma di Zorn ammette un elemento massimale $g : U \to \mathbb{R}$. Per assurdo U è diverso da V e quindi possiamo scegliere $x \in V \setminus U$, definiamo perciò la seguente funzione

$$h : u + tx \in U \oplus \operatorname{span}\{x\} \to g(u) + \alpha t \in \mathbb{R}$$

con $\alpha \in \mathbb{R}$ scelto opportunamente. L'applicazione h è chiaramente lineare definita su un sottospazio strettamente più grande di U ed estende la funzione di partenza f, dobbiamo solo dimostrare che

$$h(u + tx) = g(x) + \alpha t \leq p(u + tx)$$

Supponiamo di aver già scelto α in modo tale da soddisfare il teorema, allora dividendo per $|t|$ otteniamo le seguenti disuguaglianze

$$\begin{aligned} g(u) - \alpha &\leq p(u - x) \\ g(v) + \alpha &\leq p(v + x) \end{aligned}$$

ottenute rispettivamente per t negativo e positivo. La costante α deve perciò soddisfare

$$g(u) - p(u - x) \leq \alpha \leq p(v + x) - g(v) \text{ per ogni } u, v \in U$$

L'esistenza di un tale α equivale perciò a verificare se per ogni $u, v \in U$ vale

$$g(u) + g(v) \leq p(v + x) + p(u - x)$$

che è banalmente vera come il lettore può dimostrare senza alcuna difficoltà. Quindi possiamo prendere α in modo tale che

$$\sup_{u \in U} \left[g(u) - p(u - x) \right] \leq \alpha \leq \inf_{v \in U} \left[p(v + x) - g(v) \right]$$

e quindi $h \in \mathscr{L}$, ottenendo un assurdo. Poiché $-g(x) = g(-x) \leq p(-x)$ avremo che $-p(-x) \leq g(x) \leq p(x)$ su tutto V. ■

Corollario 5.4.4. *Dato uno spazio vettoriale V su $\mathbb{K}$, un sottospazio vettoriale $W \leq V$ e una seminorma $p : V \to \mathbb{R}$ allora per ogni funzione lineare $f : W \to \mathbb{K}$ tale che per ogni $v \in W$ si ha $\left|f(v)\right| \leq p(v)$ esiste un'altra funzione lineare $g : V \to \mathbb{K}$ che coincide con f su W tale che $\left|g(v)\right| \leq p(v)$ per ogni $v \in V$.*

Dimostrazione. Se $\mathbb{K} = \mathbb{R}$ il risultato è una semplice conseguenza del teorema di Hahn-Banach, supponiamo allora che $\mathbb{K} = \mathbb{C}$. Prendiamo allora l'applicazione T ottenuta nella proposizione 1.5.2 dunque $T(f)$ è $\mathbb{R}$-lineare su W e

$$\left|(Tf)(x)\right| = \left|\operatorname{Re} f(x)\right| \leq \left|f(x)\right| \leq p(x)$$

Esisterà allora $g' : V \to \mathbb{R}$ lineare tale che $g'(x) = \operatorname{Re} f(x)$ su W mentre $\left|g'(x)\right| \leq p(x)$ su V. Poniamo $g(x) = (T^{-1}g')(x)$ allora per la proposizione 1.5.3 per ogni $x \in V$ tale che $g(x) \neq 0$

$$\left|g(x)\right| = g'\left(\frac{\left|g(x)\right|}{g(x)} x\right) \leq p\left(\frac{\left|g(x)\right|}{g(x)} x\right) = p(x)$$

in quanto p è una seminorma su V. ■

Il teorema di Hahn-Banach permette di dedurre alcune proprietà geometriche degli spazi vettoriali, in particolare quando è possibile separare due sottoinsiemi convessi tramite un iperpiano.

Consideriamo innanzitutto lo spazio $\mathbb{R}^n$, sappiamo che gli iperpiani affini di $\mathbb{R}^n$ sono insiemi di livello di qualche funzione lineare da $\mathbb{R}^n$ in $\mathbb{R}$ non identicamente nulla, e dunque due sottoinsiemi A e B di $\mathbb{R}^n$ possono essere separati da un iperpiano se e solo se esiste un'applicazione lineare $f : \mathbb{R}^n \to \mathbb{R}$ e un valore $\alpha \in \mathbb{R}$ tali che $f(a) \leq \alpha \leq f(b)$ per ogni $a \in A$, $b \in B$. Estendiamo questo concetto ad un qualunque spazio vettoriale.

Definizione 5.4.5. Sia X uno spazio vettoriale sul campo $\mathbb{K}$, presi due qualunque suoi sottoinsiemi non vuoti A e B diremo che un'applicazione lineare $f : X \to \mathbb{K}$ **separa** A e B se e solo se esiste $\alpha \in \mathbb{R}$ tale che per ogni $x \in A$ e per ogni $y \in B$ si ha

$$\operatorname{Re} f(x) \leq \alpha \leq \operatorname{Re} f(y)$$

Diremo invece che f **separa strettamente** A e B se e solo se esistono $\alpha, \beta \in \mathbb{R}$ tali che

$$\operatorname{Re} f(x) \leq \alpha < \beta \leq \operatorname{Re} f(y)$$

per ogni $x \in A$, $y \in B$.

Lemma 5.4.6. *Sia V spazio vettoriale su $\mathbb{K}$ e $K \subseteq V$ convesso con punti internali. Allora per ogni $v_0 \in V \setminus K$ esiste $f \in V^+ \setminus \{0\}$ tale che per ogni $v \in K$*

$$\operatorname{Re} f(v) \leq \operatorname{Re} f(v_0)$$

Dimostrazione. Supponiamo per il momento che 0 sia un punto internale di K, definiamo allora l'applicazione $\mathbb{R}$-lineare

$$f : t v_0 \in \operatorname{span}\{v_0\} \subseteq V_{\mathbb{R}} \to t \in \mathbb{K}$$

Poiché v_0 non si trova in K convesso si ha $\left|f(v_0)\right| \leq p_K(v_0)$ dove p_K è il funzionale di Minkowski di K (non è in generale una seminorma poiché K non è bilanciato), per ogni $t \in \mathbb{R}$ abbiamo inoltre

$$\begin{gathered} t \geq 0 \Rightarrow f(t v_0) = t \leq t p_K(v_0) = p_K(t v_0) \\ t < 0 \Rightarrow f(t v_0) = t < 0 \leq p_K(t v_0) \end{gathered}$$

possiamo perciò applicare Hahn-Banach ed estendere f su tutto V sempre con $f(v) \leq p_K(v)$. Ora $f(v_0) = 1$ mentre per ogni $v \in K$ $f(v) \leq p_K(v) \leq 1$.

Consideriamo adesso un generico punto u internale in K allora 0 è internale in $K - u$ e quindi $f(v-u) \leq f(v_0 - u)$ e la tesi segue dalla linearità di f per $\mathbb{K} = \mathbb{R}$. Se invece $\mathbb{K} = \mathbb{C}$ allora possiamo applicare la proposizione 1.5.2 e quindi trovare un'applicazione lineare $g : X \to \mathbb{C}$ tale che $f = \operatorname{Re} g$. ■

Teorema 5.4.7 (Hahn-Banach geometrico o primo teorema di Mazur)**.** *Dati due sottoinsiemi convessi non vuoti A e B di uno spazio vettoriale V tali che $A \cap B = \emptyset$ se A possiede punti internali allora A e B possono essere separati da una funzione lineare f.*

Inoltre se tutti i punti di A sono internali allora esiste $\alpha \in \mathbb{R}$ tale che per ogni $a \in A$, $b \in B$

$$\operatorname{Re} f(a) < \alpha \leq \operatorname{Re} f(b)$$

mentre se V è uno spazio vettoriale topologico e A è aperto allora f è anche continua.

Dimostrazione. L'insieme $K = A - B$ è convesso, non vuoto e non contiene l'elemento 0 dimostriamo che contiene punti internali. Sia allora x_0 un punto internale in A e x un generico elemento di B allora per ogni $v \in V$ esisterà $\delta_0 > 0$ tale che per ogni $0 < \delta \leq \delta_0$

$$(x_0 - x) + \delta v = (x_0 + \delta v) - x \in A - B$$

e dunque $x_0 - x$ è internale in $A - B$.

Se applichiamo il lemma precedente con $v_0 = 0$ allora esisterà un funzionale $f \in V^+$ tale che per ogni $v \in K$ si ha $\operatorname{Re} f(v) \leq 0$ e quindi per ogni $v \in A$, $w \in B$ segue che $\operatorname{Re} f(v) \leq \operatorname{Re} f(w)$ ed esisterà allora $\alpha \in \mathbb{R}$ tale che

$$\sup_{v \in A} \operatorname{Re} f(v) \leq \alpha \leq \inf_{w \in B} \operatorname{Re} f(w)$$

e quindi f separa in senso largo i due insiemi. Supponiamo ora che tutti i punti di A sono internali e per assurdo esiste $a \in A$ tale che $\operatorname{Re} f(a) = \alpha$, poiché f non è identicamente nulla esisterà $x \in V$ tale che $\operatorname{Re} f(x) > 0$ ed esisterà $\delta > 0$ tale che $a + \delta x \in A$ e quindi

$$\alpha \geq \operatorname{Re} f(a + \delta x) = \operatorname{Re} f(a) + \delta \operatorname{Re} f(x) = \alpha + \delta \operatorname{Re} f(x) > \alpha$$

assurdo.

Supponiamo ora che V sia uno spazio vettoriale topologico e A sia un suo aperto, allora $\left|\operatorname{Re} f\right|$ è limitata su $A \cap (-A)$ dunque per la proposizione 4.3.3 e per il lemma 4.3.7 segue che f è continua su V. ■

Osservazione. Nell'enunciato del teorema di Hahn-Banach geometrico se B è un sottospazio vettoriale allora possiamo supporre $\alpha = 0$ e $\operatorname{Re} f(b) = 0$ identicamente su B.

Se infatti esistesse $b \in B$ per cui $\operatorname{Re} f(b) \neq 0$ allora esisterebbe $k \in \{-1, 1\}$ tale che

$$\lim_{t \to +\infty} \operatorname{Re} f(tkb) = -\infty$$

e dato che $tkb \in B$ ciò porta ad un assurdo.

Il prossimo risultato fornisce un criterio di separazione più stretto del precedente, che però vale solamente in spazi localmente convessi

Teorema 5.4.8 (Secondo teorema di Mazur)**.** *Sia V uno spazio vettoriale topologico localmente convesso, allora per ogni coppia di sottoinsiemi convessi non vuoti K e C di V con K compatto, C chiuso e $K \cap C = \emptyset$ esiste una funzione lineare continua $f : V \to \mathbb{K}$ che separa strettamente K da C.*

Se inoltre C è anche un sottospazio vettoriale di V allora f è identicamente nulla su C.

Dimostrazione. Per il teorema 4.1.11 e per la locale convessità esiste un aperto convesso contenente l'origine $U \subseteq V$ tale che $(K + U) \cap C = \emptyset$. Per il primo teorema di Mazur esiste $f : V \to \mathbb{K}$ lineare continua ed $\beta \in \mathbb{R}$ tali che per ogni $u \in K + U$, $c \in C$

$$\operatorname{Re} f(u) < \beta \leq \operatorname{Re} f(c)$$

Ora K è compatto e f è continua, dal teorema di Weierstrass $\operatorname{Re} f$ ha massimo in K che indicheremo con $\alpha \in \mathbb{R}$, questo significa che per ogni $k \in K$, $c \in C$

$$\operatorname{Re} f(k) \leq \alpha < \beta \leq \operatorname{Re} f(c)$$

Supponiamo ora che C sia anche un sottospazio vettoriale, allora $0 \in C$ e dunque $0 \leq \alpha$. Fissato un qualunque $x \in C$ per ogni $t \in \mathbb{R}$ dovremmo avere $t \operatorname{Re} f(x) = \operatorname{Re} f(tx) \leq \alpha$ il che è possibile se e solo se $\operatorname{Re} f(x) = 0$ per ogni $x \in C$ il che è vero se e solo se f è identicamente nulla su C. ■

Utilizzeremo ora il secondo teorema di Mazur per dimostrare il seguente risultato sulle basi duali che utilizzeremo in seguito. Prima però dimostriamo la proposizione seguente:

Proposizione 5.4.9. *Dati uno spazio vettoriale X su $\mathbb{K}$ e delle funzioni $f, f_1, f_2, \ldots, f_k$ lineari da X in $\mathbb{K}$ tali che*

$$\bigcap_{i=1}^{k} \ker f_i \leq \ker f$$

allora esisteranno $\lambda_1, \lambda_2, \ldots, \lambda_k$ tali che per ogni $x \in X$

$$f(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \cdots + \lambda_k f_k(x)$$

Dimostrazione. Consideriamo il sottospazio di $\mathbb{K}^{k+1}$ dato da

$$F = \left\{ [f(x), f_1(x), \ldots, f_k(x)] \in \mathbb{K}^{k+1} \;\middle|\; x \in X \right\}$$

per ipotesi il vettore $(1, 0, 0, \ldots, 0)$ non può appartenere a F, dunque poiché stiamo lavorando su spazi finitamente generati esisterà $\vec{\lambda} \in \mathbb{K}^{k+1}$ tale che $\vec{\lambda} \cdot [f(x), f_1(x), \ldots, f_k(x)] = 0$ e $\vec{\lambda} \cdot (1, 0, \ldots, 0) \neq 0$.

Quindi posto $\vec{\lambda} = (\lambda, \lambda_1, \ldots, \lambda_k)$ segue che $\lambda \neq 0$ e

$$\lambda f(x) + \sum_{i=1}^{k} \lambda_i f_i(x) = 0$$

per ogni $x \in X$. ■

Lemma 5.4.10 (della base duale)**.** *Prendiamo uno spazio vettoriale topologico X su $\mathbb{K}$. Allora*

1. *Se X è localmente convesso allora per ogni scelta di $x_1, x_2, \ldots x_n \in X$ linearmente indipendenti esistono delle funzioni lineari continue $x_1^*, x_2^*, \ldots, x_n^* \in X^*$ tali che*

$$x_i^*(x_j) = \delta_{ij}$$

2. *Per ogni scelta di $x_1^*, x_2^*, \ldots x_n^* \in V^*$ linearmente indipendenti esistono $x_1, x_2, \ldots, x_n \in X$ tali che*

$$x_i^*(x_j) = \delta_{ij}$$

Dimostrazione. Dimostriamo i vari punti.

1. Per ogni i vale

$$x_i \notin \operatorname{span}\left\{ x_j \;\middle|\; j \neq i \right\} = F_i$$

poiché la dimensione di F_i è finita allora per il corollario 4.6.3 F_i è chiuso e dal secondo teorema di Mazur esistono $x_i^* \in L(X, \mathbb{K})$ tali che $x_i^*(x_i) = 1$ e $x_i^*(F_i) = \{0\}$.

2. Usiamo l'induzione su n. Per $n = 1$ allora $x_1^* \neq 0$ e quindi la tesi è immediata, supponiamolo valido per $n-1$ quindi esistono $x_1', x_2', \dots, x_{n-1}' \in X$ tali che per ogni $i, j < n$ si ha $x_i^*(x_j') = \delta_{ij}$.

   Quindi per la proposizione 5.4.9 esisterà un certo $x_n \in X$ che annullerà ogni x_i^* per $i < n$ ma $x_n^*(x_n) = 1$, ciò implicherà anche che x_n non può dipendere linearmente dagli x_i'. Dunque posto

$$\begin{aligned} x_1 &= x_1' - x_n^*(x_1')x_n \\ x_2 &= x_2' - x_n^*(x_2')x_n \\ &\dots \\ x_{n-1} &= x_{n-1}' - x_n^*(x_{n-1}')x_n \end{aligned}$$

   avremo che $x_1, \dots, x_{n-1}, x_n$ sono linearmente indipendenti e soddisfano tutte le ipotesi richieste. Per induzione il lemma è così dimostrato.

■

Capitolo 6

Dualità per funzioni convesse

6.1 La trasformata di Legendre

In questa sezione consideriamo due spazi vettoriali reali V e W e una applicazione bilineare $F : V \times W \to \mathbb{R}$. Poniamo anche

$$\begin{aligned} \sup \emptyset &= -\infty \\ \inf \emptyset &= +\infty \end{aligned}$$

Vogliamo chiederci per quali funzioni $f : V \to \overline{\mathbb{R}}$ esistano $a \in W$, $r \in \mathbb{R}$ tali che

$$f(x) \geq F(x, a) - r$$

per ogni $x \in V$, ovvero quali funzioni possono essere minorate da una qualche funzione affine, e in tal caso se esse stesse possono essere ricostruite a partire dal loro **sottografico**, ovvero dal sottoinsieme di $W \times \mathbb{R}$

$$G_f = \left\{(a, \alpha) \in W \times \mathbb{R} \;\middle|\; f(x) \geq F(x, a) - \alpha \;\forall x \in V\right\}$$

infatti mentre l'epigrafico di f è contenuto in $V \times \mathbb{R}$ il suo sottografico sarà invece contenuto in $W \times \mathbb{R}$. Fissato $a \in W$ avremo che $(a, \alpha) \in G_f$ se e solo se

$$\alpha \geq \sup_{x \in V} [F(x, a) - f(x)]$$

Da questa relazione definiamo la **trasformata di Legendre associata alla terna** (V, W, F) come quella applicazione che ad ogni funzione $f : V \to \overline{\mathbb{R}}$ associa la nuova funzione $f^* : W \to \overline{\mathbb{R}}$ definita come

$$f^*(a) = \sup_{x \in V} [F(x, a) - f(x)]$$

Data una funzione $f : V \to \overline{\mathbb{R}}$ non è difficile verificare che valgono le seguenti proprietà:

- se $g(x) = f(x) + l$ per qualche $l \in \mathbb{R}$ allora $g^*(a) = f^*(a) - l$;
- se $g(x) = f(\alpha x)$ per qualche $\alpha \neq 0$ allora $g^*(a) = f^*(a/\alpha)$;
- se $g(x) = f(x - y)$ per qualche $y \in V$ allora $g^*(a) = f^*(a) + F(y, a)$, viceversa se $g(x) = f(x) + F(x, b)$ per qualche $b \in W$ allora $g^*(a) = f^*(a - b)$;

- dato un sottoinsieme K di V si ha $i_K^*(a) = \sup_{x\in K} F(x,a) = s_{K,F}(a)$ e quindi le funzioni supporto sono la trasformata di Legendre di qualche funzione indicatrice;
- data $g : V \to \overline{\mathbb{R}}$ in modo tale che $f(x) \leq g(x)$ per ogni $x \in V$ allora $g^*(a) \leq f^*(a)$ per ogni $a \in W$.

Inoltre se consideriamo la terna $(W, V, \tilde{F})$ con

$$\tilde{F}(a,x) = F(x,a)$$

allora possiamo considerare la trasformata di Legendre associata a quest'altra terna che associa ad f^* un'altra funzione f^{**} su V tale che

$$f^{**}(x) = \sup_{a\in W}[F(x,a) - f^*(a)] = \sup_{a\in W}\inf_{y\in V}[F(x-y,a) + f(y)]$$

In generale vale la disuguaglianza $f^{**} \leq f$ ma la disuguaglianza di solito è stretta, anche se f fosse una funzione lineare di V in quanto dipenderà, oltre che da V, anche da W e da F.

Una notevole disuguaglianza che utilizzeremo spesso nel seguito discende immediatamente dalla definizione di trasformata di Legendre:

$$F(x,a) \leq f(x) + f^*(a) \tag{6.1.1}$$

per ogni $x \in V, a \in W$, sempre ponendo $+\infty - \infty = +\infty$. Fissato f questa disuguaglianza risulta essere ottimale, nel senso che non esiste un'altra funzione $g : W \to \overline{\mathbb{R}}$ che offra una stima migliore di f^*. Lo studio dei casi in cui valga l'uguaglianza nella (6.1.1) richiede degli strumenti aggiuntivi, si veda più avanti la proposizione 6.4.2.

Una funzione $f : V \to \overline{\mathbb{R}}$ è **sempropria** se e solo se $f \equiv -\infty$ oppure se $f(x) > -\infty$ per ogni $x \in V$. Dimostriamo adesso alcune proprietà della trasformata di Legendre.

Proposizione 6.1.1. *Date due funzioni f, g da V in $\overline{\mathbb{R}}$ e $\lambda \geq 0$ si ha*

$$f^* + g^* = (f \diamond g)^*$$
$$\lambda f^* = f_\lambda^*$$

ponendo eventualmente $+\infty \times 0 = 0$.

Dimostrazione. Si ha per ogni $a \in W$

$$\begin{aligned} f^*(a) + g^*(a) &= \sup_{x,y\in V}[F(x+y,a) - f(x) - g(y)] \\ &= \sup_{x,y\in V}\sup_{b\in W}[F(x,b) + F(y,a-b) - f(x) - g(y)] = (f \diamond g)^*(a) \end{aligned}$$

Se $\lambda = 0$ allora $f_0 = i_0$ e perciò $i_0^* = s_0^* = 0 = 0f^*$.

$$\lambda f^*(a) = \sup_{x\in V}[\lambda F(x,a) - \lambda f(x)] = \sup_{x\in V}[F(x,a) - \lambda f(\lambda^{-1}x)]$$

■

Proposizione 6.1.2. *Data una qualunque funzione $f : X \to \overline{\mathbb{R}}$ la trasformata di Legendre f^* è convessa e semipropria e $f \equiv +\infty \Leftrightarrow f^* \equiv -\infty$. Inoltre se f^* è propria anche f lo sarà.*

Dimostrazione. Per ogni $x \in V$ la funzione $k_x(a) = F(x,a) - f(x)$ è convessa su W, poiché vale la

$$\mathrm{epi}_c f^* = \bigcap_{x\in V} \mathrm{epi}_c\, k_x$$

avremo che $\mathrm{epi}_c f^*$ è convesso in $W \times \overline{\mathbb{R}}$ e perciò f^* è convessa. Se esistesse $a \in W$ tale che $f^*(a) = -\infty$ allora $F(x,a) - f(x) = -\infty$ per ogni $x \in V$ e questo è possibile se e solo se $f \equiv +\infty$ identicamente, dunque $f^* \equiv -\infty$ e perciò f^* è semipropria.

Chiaramente se $f \equiv +\infty$ allora $f^* \equiv -\infty$, se invece $f^* \equiv -\infty$ allora avremo che $-\infty = f^*(0) = \sup_{x\in V}[-f(x)] = -\inf_{x\in V} f(x)$ ovvero $f \equiv +\infty$.

Prendiamo ora una funzione f tale che f^* sia propria, poiché $f^*(0) = -\inf_{x\in V} f(x) > -\infty$ avremo che f non può essere identicamente $+\infty$, inoltre esisterà $b \in W$ tale che $f^*(b) \in \mathbb{R}$ e questo implica che

$$f(x) \geq F(x,b) - f^*(b) > -\infty$$

per ogni $x \in V$, dunque f è propria. ■

Proposizione 6.1.3. *Data una qualunque funzione $f : V \to \overline{\mathbb{R}}$ tale che f^* sia propria avremo*

$$\left(i_{\mathrm{dom} f}\right)^* = \mathrm{dir} f^*$$

Dimostrazione. Poiché f^* è propria per la proposizione 6.1.2 anche f è propria, dunque

$$\left(i_{\mathrm{dom} f}\right)^*(a) = \sup_{x\in \mathrm{dom} f} F(x,a)$$

Prendiamo ora $b \in \mathrm{dom} f^*$ generico, allora $F(x,b) - f(x) \leq f^*(b) < +\infty$ per ogni $x \in V$. Ma da quanto abbiamo detto prima avremo

$$\begin{aligned}
\mathrm{dir} f^*(a) = \sup_{\lambda>0} \frac{f^*(b+\lambda a) - f^*(b)}{\lambda} &= \sup_{\lambda\geq 0}\sup_{x\in\mathrm{dom} f} \frac{F(x,b+\lambda a) - f(x) - f^*(b)}{\lambda} \\
&= \sup_{x\in\mathrm{dom} f}\sup_{\lambda>0}\left[F(x,a) + \frac{F(x,b) - f(x) - f^*(b)}{\lambda}\right] \\
&= \sup_{x\in\mathrm{dom} f} F(x,a) = \left(i_{\mathrm{dom} f}\right)^*(a)
\end{aligned}$$

in quanto $F(x,b) - f(x) - f^*(b) \leq 0$. Abbiamo inoltre ristretto il dominio dell'estremo superiore da V a $\mathrm{dom} f$ in quanto gli elementi in $V \setminus \mathrm{dom} f$ non contribuiscono al calcolo di $f^*(b+\lambda a)$ in quanto

$$x \in V \setminus \mathrm{dom} f \Rightarrow f(x) = +\infty \Rightarrow F(x, b+\lambda a) - f(x) = -\infty < f^*(b+\lambda a)$$

■

Proposizione 6.1.4. *Data una qualunque funzione $f : V \to \overline{\mathbb{R}}$ propria positivamente omogenea allora f^* è la funzione indicatrice di qualche sottoinsieme convesso di W.*

Dimostrazione. Basta solamente verificare che f^* assume solamente i valori 0 e $+\infty$, in tal caso infatti segue che $f^* = i_{\{f^*=0\}}$. Per ogni $\lambda > 0$ dato che $f(\lambda x) = \lambda f(ax)$ abbiamo

$$f^*(a) = \sup_{x\in V}[F(x,a) - f(x)] = \sup_{x\in V}[F(\lambda x, a) - f(\lambda x)] = \lambda f^*(a)$$

per ogni $a \in W$. Infatti poiché f possiede almeno un valore diverso da $+\infty$ abbiamo visto che f^* non può assumere il valore $-\infty$ dunque gli unici valori ammissibili per f^* sono 0 e $+\infty$. ■

Proposizione 6.1.5. *Data una qualunque funzione propria $f : V \to \overline{\mathbb{R}}$ si ha*

$$\left(\mathrm{rad} f\right)^* = i_{\{f^* \leq 0\}}$$

Dimostrazione. Abbiamo

$$\left(\inf_{\lambda>0} f_\lambda\right)^*(a) = \sup_{x\in V,\lambda>0}\left[F(x,a) - \lambda f(\lambda^{-1}x)\right] = \sup_{\lambda>0} \lambda f^*(a)$$

che fa 0 se e solo se $f^* \leq 0$ altrimenti è $+\infty$. ■

Per ottenere un criterio di dualità della trasformata di Legendre abbiamo bisogno di estendere il secondo teorema di Mazur, con il quale siamo riusciti a separare i punti dagli insiemi convessi chiusi tramite iperpiani, alla terna (V, W, F). Per farlo dobbiamo definire su V una topologia che soddisfi i seguenti assiomi:

- lo spazio V è uno spazio vettoriale topologico localmente convesso di Hausdorff (svc);
- per ogni $a \in W$ l'applicazione lineare

  $$\hat{a} : y \in V \to F(y, a) \in \mathbb{R}$$

  è anche continua;
- dato un qualunque insieme chiuso convesso $C \subseteq V$ per ogni $x \in V \setminus C$ esistono $a \in W$, $r \in \mathbb{R}$ e $\varepsilon > 0$ tali che

  $$F(y, a) \leq r - \varepsilon < r + \varepsilon \leq F(x, a)$$

  per ogni $y \in C$.

Definizione 6.1.6. Dati due spazi vettoriali reali V, W e data un'applicazione bilineare $F : V \times W \to \mathbb{R}$ diremo che la terna (V, W, F) è **duale in un verso** rispetto alla topologia τ su V se e solo se V soddisfa i tre assiomi precedenti per τ.

Diremo invece che (V, W, F) è **duale nel verso opposto** rispetto ad una topologia η di W se e solo se $(W, V, \bar{F})$, con $\bar{F}(w, v) = F(v, w)$, è duale in un verso rispetto a τ. Infine diremo che la terna è **duale in entrambi i versi** rispetto a τ ed ad η se e solo se è duale in un verso rispetto a τ ed è duale nel verso opposto rispetto ad η.

Lemma 6.1.7. *Dato uno spazio vettoriale topologico localmente convesso V allora (V, W, F) è duale in un verso se e solo se:*

- *per ogni $x \in V \setminus \{0\}$ esiste $a \in W$ tale che $F(x, a) > 0$;*
- *si ha $f \in V^*$ se e solo se $f = \hat{a}$ per qualche $a \in W$.*

Inoltre se (V, W, F) è duale in un verso allora anche $\left(V \times \mathbb{R}, W \times \mathbb{R}, \ddot{F}\right)$ lo è con la topologia prodotto dove

$$\ddot{F}[(x, r), (a, s)] = F(x, a) + rs$$

Dimostrazione. Supponiamo che la topologia di V sia duale rispetto a W e F, dato che V è uno spazio di Hausdorff i punti sono chiusi, dunque per ogni $x \neq 0$ esisterà $a \in W$ tale che $0 = F(0, a) < F(x, a)$. Preso ora un qualunque $f \in V^*$ se $f \equiv 0$ chiaramente avremo che $f = \hat{0}$, altrimenti esisterà $x \in V$ tale che $f(x) < 0$. L'insieme $\left\{f \geq 0\right\}$ è chiuso e convesso in V

e non contiene x dunque esisteranno $a \in W$ e $r \in \mathbb{R}$ tali che per ogni $y \in V$ per cui $f(y) \geq 0$ si ha $F(y,a) + r = \hat{a}(y) + r \geq 0$.

Ora per ogni $y \in \ker f$ dovremmo avere $\hat{a}(ty) + r \geq 0$ per ogni $t \in \mathbb{R}$ e ciò è possibile se e solo se $\hat{a}(y) = 0$ e così $\ker f \subseteq \ker \hat{a}$. Ma allora per la proposizione 5.4.9 esiste $\lambda \in \mathbb{K}$ tale che $f(y) = \lambda \hat{a}(y) = F(y, \lambda a)$ per ogni $y \in V$.

Viceversa se V soddisfa le due ipotesi nell'enunciato dobbiamo solamente verificare che in V i punti sono chiusi, in quanto per regolarità avremo che V è uno spazio di Hausdorff. Per ogni $x \neq 0$ abbiamo detto che esiste $a_x \in W$ tale che $F\left(x, a_x\right) > 0$ e dunque

$$\{0\} = \left\{z \in V \,\middle|\, F\left(z, a_x\right) = 0 \ \forall x \in V \setminus \{0\}\right\} = \bigcap_{x \neq 0} \ker \hat{a}_x$$

e quindi $\{0\}$ è chiuso, poiché siamo in uno spazio vettoriale topologico tutti i punti saranno chiusi.

Per quanto riguarda la topologia prodotto su $V \times \mathbb{R}$ se la topologia su V è duale allora per ogni $g \in L(V \times \mathbb{R}, \mathbb{R})$ esisterà $a \in W$ tale che

$$g(x,r) = g(x,0) + rg(0,1) = F(x,a) + rg(0,1) = \ddot{F}[(x,r),[a,g(0,1)]]$$

in quanto $x \to g(x,0)$ è lineare e continua da V in $\mathbb{R}$. Per concludere per la continuità della somma e del prodotto l'applicazione $(x,r) \in V \times \mathbb{R} \to F(x,a) + rs$ è continua per ogni $(a,s) \in W \times \mathbb{R}$. ■

Grazie al lemma 6.1.7 se V è finitamente generato e (V,W,F) è duale in un verso allora possiamo supporre che anche W sia finitamente generato. Infatti presa una qualunque base $x_1, x_2, \ldots, x_n$ di V esisteranno $x_1^*, x_2^*, \ldots, x_n^* \in V^* = V^+$ tali che $x_i^*\left(x_j\right) = \delta_{ij}$ e quindi esisteranno $a_1, \ldots, a_n \in W$ tali che $F(x_i, a_j) = \delta_{ij}$. In tal modo possiamo scrivere

$$\begin{gathered} W = W_1 \oplus W_2 \\ W_1 = \operatorname{span} a_1, a_2, \ldots, a_n \\ W_2 = \left\{b \in W \,\middle|\, F(x,b) = 0 \ \forall x \in V\right\} \end{gathered}$$

e quindi possiamo sostituire W con W_1, il quale è banalmente finitamente generato.

Consideriamo adesso due qualunque spazi vettoriali reali V e W e una funzione bilineare $F : V \times W \to \mathbb{R}$ non degenere, ovvero per ogni $x \in V \setminus \{0\}, a \in W \setminus \{0\}$ esistono $y \in V, b \in W$ tali che

$$F(x,b) \neq 0 \qquad\qquad F(y,a) \neq 0$$

Dotiamo allora V della topologia iniziale generata dai funzionali $\hat{a}$ al variare di $a \in W$ e W della topologia iniziale generata dai funzionali nella forma $\hat{x}$ al variare di $x \in V$. Diremo che queste sono le **topologie duali** di V e W.

Teorema 6.1.8. *Gli spazi V e W dotati di queste topologie iniziali sono svc duali tra loro rispetto ad F. Inoltre ogni altra topologia su V che lo renda uno svc duale rispetto ad W e F sarà sempre più fine di questa topologia iniziale.*

Dimostrazione. Verifichiamo innanzitutto che V è uno spazio vettoriale topologico localmente convesso di Hausdorff, in maniera analoga si verificherà anche per W.

Prima di tutto dimostriamo che la somma è continua, per la proposizione 2.7.3 la somma è continua se e solo se per ogni $a \in W$ l'applicazione $(x, y) \in V \times V \to \hat{a}(x + y) \in \mathbb{R}$ è continua. Ma $\hat{a}(x + y) = \hat{a}(x) + \hat{a}(y)$ per linearità dunque le applicazioni seguenti

$$(x, y) \in V \times V \to [\hat{a}(x), \hat{a}(y)] \in \mathbb{R} \times \mathbb{R} \to \hat{a}(x) + \hat{a}(y) \in \mathbb{R}$$

sono tutte continue in quanto le $\hat{a}$ sono sempre continue nella loro topologia iniziale, quindi anche la somma è continua. In maniera analoga si verifica la continuità del prodotto e quindi V è uno spazio vettoriale topologico.

Presi ora $x, y \in V$ distinti allora esisterà $a \in W$ tale che $\hat{a}(x - y) = F(x - y, a) \neq 0$ e $\hat{a}(x) \neq \hat{a}(y)$, poiché la topologia su $\mathbb{R}$ è di Hausdorff anche V sarà uno spazio di Hausdorff.

Per verificare la locale convessità osserviamo che gli insiemi $B_\varepsilon = \left(-\varepsilon, \varepsilon\right)$ formano un sistema fondamentale di intorni convesso in $\mathbb{R}$ dell'origine, per linearità anche $\hat{a}^{-1}(B_\varepsilon)$ sarà aperto convesso di V contenente l'origine e gli insiemi

$$\mathscr{W}(B_{\boldsymbol{\varepsilon}}) = \bigcap_{a \in \operatorname{dom} \boldsymbol{\varepsilon}} \hat{a}^{-1}\left(B_{\varepsilon_a}\right)$$

dove $\mathscr{W} = \{\hat{a} \mid a \in W\}$ sarà un sistema fondamentale di intorni convessi dell'origine al variare del multivettore $\boldsymbol{\varepsilon}$ in $MV(W, \mathbb{R}^+)$, quindi V è uno svc.

Dobbiamo solamente provare che $V^* = \mathscr{W}$ ora, in quanto dal lemma 6.1.7 seguirà immediatamente che questa topologia di V è duale rispetto a W e F. Prendiamo allora un qualunque funzionale lineare continuo $f : V \to \mathbb{R}$ rispetto a questa topologia iniziale, allora $f^{-1}(B_1)$ è un aperto di V contenente l'origine e perciò esisterà un multivettore $\boldsymbol{\varepsilon} \in MV(W, \mathbb{R}^+)$ tale che

$$\mathscr{W}(B_{\boldsymbol{\varepsilon}}) = \left\{x \in V \middle| \; \left|F(x, a)\right| < \varepsilon_a \; \forall a \in \operatorname{dom} \boldsymbol{\varepsilon}\right\} \subseteq f^{-1}(B_1)$$

In altre parole se $\operatorname{dom} \boldsymbol{\varepsilon} = \left\{a_1, a_2, \ldots, a_n\right\}$ e $\varepsilon = \min \boldsymbol{\varepsilon} > 0$ allora $\left|\hat{a}_i(x)\right| < \varepsilon$ per ogni $1 \leq i \leq n$ implica che $f(x) < 1$.

Ma allora per la linearità di f e delle $\hat{a}_i$ abbiamo

$$\bigcap_{i=1}^{n} \ker \hat{a}_i \leq \ker f$$

e possiamo allora applicare la proposizione 5.4.9 per cui esisteranno $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tali che

$$f = \lambda_1 \hat{a}_1 + \cdots + \lambda_n \hat{a}_n = \hat{a}$$

e $a \in W$, quindi possiamo applicare il lemma 6.1.7 e quindi la topologia iniziale di V è duale rispetto a W e F. Possiamo ripetere gli stessi ragionamenti anche per W.

Poiché la topologia iniziale è la topologia meno fine tra quelle che rendono tutte le $\hat{a}$ continue ogni altra topologia su V che lo renda una svc duale rispetto a W e F dovrà essere più fine della rispettiva topologia iniziale, così l'ultima parte del teorema è dimostrata. ■

Corollario 6.1.9. *Dato un qualunque svc V e un suo qualunque sottoinsieme C convesso allora C è chiuso in V se e solo se è chiuso rispetto alla topologia ottenuta nel teorema precedente ponendo $W = V^*$ e $F(x, f) = f(x)$.*

Dimostrazione. In V i punti sono chiusi, dunque per il secondo teorema di Mazur per ogni $x \in V \setminus \{0\}$ esiste $f \in V^*$ tale che $f(x) \neq 0$, dunque possiamo applicare il teorema precedente. ■

Quindi per quanto riguarda gli insiemi e le funzioni convesse non ha importanza quale topologia si induca su V e W purché esse siano duali tra loro rispetto ad F. Da questo momento supporremo quindi che V e W siano svc duali tra loro, infatti anche se non fossero spazi topologici per il teorema precedente li potremo sempre dotare di una tale topologia.

In generale $\mathrm{epi}_c f$ non è chiuso e $\mathrm{epi}_o f$ non è aperto in $V \times \mathbb{R}$ nemmeno se f fosse convessa, inoltre le funzioni convesse possono non essere continue (dato che possono esistere anche funzioni lineari non continue) anche se $\mathrm{epi}_c f$ fosse chiuso.

In generale si ha infatti

Lemma 6.1.10. *Dati uno svc V e una funzione $f : V \to \overline{\mathbb{R}}$ sono equivalenti*

1. $\mathrm{epi}_c f$ *chiuso in* $V \times \mathbb{R}$*;*
2. $\{f \leq \alpha\}$ *chiuso in V per ogni* $\alpha \in \mathbb{R}$*;*
3. *la funzione f è semicontinua inferiormente.*

Dimostrazione. L'equivalenza tra gli ultimi due punti è esattamente quanto detto nel lemma 2.6.5, inoltre gli iperpiani $P_\alpha = \{(x,\alpha) \mid x \in V\}$ sono chiaramente chiusi in $V \times \mathbb{R}$ e perciò il primo punto implica il secondo.

Invece se $\{f \leq \alpha\}$ è chiuso per ogni $\alpha \in \mathbb{R}$ prendiamo $(x,r) \notin \mathrm{epi}_c f$ allora $r < f(x)$ e quindi esisterà $\varepsilon > 0$ tale che $r + \varepsilon < f(x)$. Da questo possiamo dedurre che l'insieme

$$U = \{f > r + \varepsilon\} \times (-\infty, r+\varepsilon)$$

contiene il punto (x,r), è aperto in quanto $\{f > \alpha\}$ è aperto e non interseca $\mathrm{epi}_c f$ quindi $\mathrm{epi}_c f$ è chiuso in $V \times \mathbb{R}$. ■

Proposizione 6.1.11. *Dati due epigrafici E, D su $V \times \mathbb{R}$ e V è un qualunque svc se $E \sim D$ allora* $\mathrm{Cl}(E) = \mathrm{Cl}(D)$ *e* $\mathrm{Int}(E) = \mathrm{Int}(D)$.

Dimostrazione. Per ogni $\varepsilon > 0$ abbiamo

$$D \subseteq \bigcap_{\varepsilon>0} E - \vec{\varepsilon} \subseteq \bigcap_{\varepsilon>0} \mathrm{Cl}(E) - \vec{\varepsilon} \subseteq \mathrm{Cl}(E)$$

dove l'ultima inclusione segue dal fatto che $V \times \mathbb{R}$ è dotato della topologia prodotto, dunque se $(x, r+\varepsilon) \in \mathrm{Cl}(E)$ per ogni $\varepsilon > 0$ allora anche $(x,r) \in \mathrm{Cl}(E)$,dunque $\mathrm{Cl}(D) \subseteq \mathrm{Cl}(E)$ e l'inclusione opposta si dimostra in maniera analoga. Lo stesso vale per l'interno di E e D. ■

Data una qualunque funzione f da V in $\overline{\mathbb{R}}$ la funzione $INF[\mathrm{Cl}(\mathrm{epi}_c f)]$ sarà sempre semicontinua inferiormente per il teorema di identificazione, inoltre è la più grande funzione semicontinua inferiormente ad essere minore o uguale a f. Data una funzione convessa f poniamo anche

$$\mathrm{Cl}(f) = \begin{cases} -\infty & \text{se } f(x) = -\infty \text{ per qualche } x \in V \\ INF[\mathrm{Cl}(\mathrm{epi}_c f)] & \text{se } f(x) > -\infty \text{ per ogni } x \in V \end{cases}$$

mentre per una funzione generica f poniamo $\overline{\mathrm{con}}\, f = \mathrm{Cl}(\mathrm{con} f)$. Osserviamo che per un qualunque sottoinsieme K di V si ha $\overline{\mathrm{con}}\ i_K = i_{\overline{\mathrm{con}(K)}}$ mentre per funzioni convesse semiproprie e semicontinue inferiormente si ha $\mathrm{Cl}(f) = f$.

Ancora una conseguenza del corollario precedente è che $\overline{\text{con}}f$ non dipende solamente da W e da F e non dalla topologia scelta su V (purché sia duale a W ed F) o dalla topologia di W, quindi possiamo considerarla come un'operazione prettamente vettoriale e non topologica.

Proposizione 6.1.12. *Data una funzione convessa propria $f : V \to \mathbb{R}^l$ allora*

$$\operatorname{Cl}(f)(x) = \liminf_{y \to x} f(y)$$

Dimostrazione. Prendiamo un qualunque intorno aperto bilanciato dell'origine $U \subseteq V$ e per ogni $\varepsilon > 0$ avremo che $INF(U \times (-\varepsilon, \varepsilon))(x) = -\varepsilon i_U$. Dunque

$$INF(\operatorname{epi}_c f + U \times (-\varepsilon, \varepsilon))(x) = \inf_{y \in x+U} f(y) - \varepsilon$$

$$\begin{aligned} INF(\operatorname{Cl}(\operatorname{epi}_c f))(x) &= INF\left[\bigcap_{U,\varepsilon}(\operatorname{epi}_c f + U \times (-\varepsilon, \varepsilon))\right](x) \\ &= \inf_U \sup_{y \in x+U} f(y) = \liminf_{y \to x} f(y) \end{aligned}$$

■

Il risultato fondamentale di dualità che ci permette di passare dalla trasformata di Legendre alla funzione di partenza è il seguente:

Teorema 6.1.13. *Per ogni funzione $f : V \to \overline{\mathbb{R}}$ convessa semipropria e semicontinua inferiormente avremo per ogni $x \in V$*

$$f(x) = \sup\left\{F(x,a) - \alpha \;\middle|\; a \in W, \alpha \in \mathbb{R} \; con \, f(y) \geq F(y,a) - \alpha \; \forall y \in V\right\} \tag{6.1.2}$$

Dimostrazione. Possiamo tranquillamente supporre $f(x) > -\infty$ per ogni $x \in V$ esiste $r \in \mathbb{R}$ tale che $(x, r) \notin \operatorname{epi}_c f$. Poiché f è semicontinua inferiormente e V è localmente convesso per il lemma 6.1.7 esisteranno $a \in W$, $v, \beta \in \mathbb{R}$ e $\varepsilon > 0$ tali che $F(y, a) + vt \geq \beta + \varepsilon$ per ogni $(y, t) \in \operatorname{epi}_c f$ e $F(x, a) + vr \leq \beta - \varepsilon$. Per il momento supponiamo che $f(x) < +\infty$ e quindi $f(x) > r$ e $[x, f(x)] \in \operatorname{epi}_c f$, in tal caso avremo per linearità

$$F(x,a) + rv < F(x,a) + vf(x) \Rightarrow [f(x) - r]v > 0$$

e perciò $v > 0$ e il nostro iperpiano non sarà verticale.

Per ogni $z \in \operatorname{dom} f$ avremo allora

$$vf(z) + F(z,a) - \varepsilon \geq vr + F(x,a) + \varepsilon \Rightarrow f(z) > F(z, -a/v) - \alpha$$

dove $\alpha = F(x, a/v) - r$ e quindi il secondo insieme non è vuoto (anche quando $f(x) = +\infty$, in tal caso basta scegliere un qualunque altro $x' \in \operatorname{dom} f$) e facendo tendere r a $f(x)$ abbiamo che l'estremo superiore converge a $f(x)$.

Invece ora supponiamo che $f(x) = +\infty$, sappiamo però che esiste $x' \in V$ tale che $f(x') < +\infty$ e quindi esisteranno $a \in W, \alpha \in \mathbb{R}$ tali che $f(z) \geq F(z, a) - \alpha$. Per ogni $r \in \mathbb{R}$ avremo che $(x, r) \notin \operatorname{epi}_c f$, possiamo allora supporre che per qualche r l'iperpiano che li separa è verticale, ovvero che esisterà $b \in W$ tale che $F(y, b) \geq \beta + \varepsilon$ per ogni $(y, t) \in \operatorname{epi}_c f$ e $F(x, b) \leq \beta - \varepsilon$.

In particolare per ogni $y \in \operatorname{dom} f$ avremo $F(y,b) \geq \beta + \varepsilon$ e cosi per ogni $\lambda \geq 0$ e per ogni $y \in V$ abbiamo

$$F(y,a) - \alpha - \lambda[F(y,b) - \beta - \varepsilon] = F(y, a - \lambda b) - \gamma_\lambda \leq f(z)$$

Ora posto $z = x$ il termine $F(x,b) - \beta - \varepsilon$ è strettamente positivo e quindi

$$\lim_{\lambda \to +\infty} [F(x, a - \lambda b) - \gamma_\lambda] = +\infty = f(x)$$

e quindi il risultato è dimostrato anche quando $x \notin \operatorname{dom} f$. ■

Corollario 6.1.14. *Data una generica funzione $f : V \to \overline{\mathbb{R}}$ avremo che*

$$\overline{\operatorname{con}} f = \sup \left\{ F(x,a) - \alpha \;\middle|\; a \in W, \alpha \in \mathbb{R} \text{ con } f(z) > F(z,a) - \alpha \; \forall z \in V \right\}$$

Dimostrazione. Per linearità e continuità abbiamo che $\operatorname{Cl}(\operatorname{con} \hat{a}) = \hat{a}$ e dunque $F(x,a) + \alpha \leq f(x)$ se e solo se $F(x,a) + \alpha \leq \operatorname{Cl}\big(\operatorname{con} f\big)(x) \leq f(x)$. ■

Proposizione 6.1.15. *Data una qualunque funzione $f : V \to \overline{\mathbb{R}}$ la trasformata di Legendre f^* sarà convessa semipropria e semicontinua inferiormente su W.*

Dimostrazione. Per ogni $x \in V$ la funzione $k_x(a) = F(x,a) - f(x) = \hat{x}(a) - f(x)$ è continua su W, poiché vale la

$$\operatorname{epi}_c f^* = \bigcap_{x \in V} \operatorname{epi}_c k_x$$

avremo che $\operatorname{epi}_c f^*$ è chiuso e perciò f^* è semicontinua inferiormente. Il resto deriva dalla proposizione 6.1.2. ■

Ora abbiamo tutto il necessario per dimostrare il nostro teorema di dualità e poter stabilire quando una funzione è la trasformata di Legendre di qualche altra funzione:

Teorema 6.1.16 (Fenchel-Morenau)**.** *Data una funzione $f : V \to \overline{\mathbb{R}}$ convessa semipropria e semicontinua inferiormente allora $f^{**} = f$ ed f è propria se e solo se f^* è propria.*

Inoltre se invece f è una funzione generica su V allora $f^* = (\overline{\operatorname{con}} f)^*$ e $f^{**} = \overline{\operatorname{con}} f$.

Dimostrazione. Sappiamo già che $f^{**} \leq f$ su V, dobbiamo verificare la disuguaglianza opposta. Presi $a \in W, \alpha \in \mathbb{R}$ tali che $f(z) \geq F(z,a) - \alpha$ per ogni $z \in V$ per definizione di trasformata di Legendre abbiamo che $\alpha \geq f^*(a)$ e di conseguenza $-\alpha \leq -f^*(a)$.

Il teorema 6.1.13 garantisce allora che

$$f^{**}(x) = \sup_{a \in V} [F(x,a) - f^*(a)] \geq \sup \left\{ F(x,a) - \alpha \;\middle|\; f(z) \geq F(z,a) - \alpha \; \forall z \in V \right\} = f(x)$$

e l'uguaglianza è così verificata, inoltre se $f = f^{**}$ fosse propria per la proposizione 6.1.2 anche f^* è propria.

Invece prendiamo adesso f applicazione generica. Poiché $\overline{\operatorname{con}} f \leq f$ avremo che $f^* \leq (\overline{\operatorname{con}} f)^*$, prendiamo $a \in W$ e $\mu \in \mathbb{R}$ tali che $f^*(a) \leq \mu$ allora $f(x) \geq F(x,a) - \mu$ per ogni $x \in V$.

Grazie al corollario 6.1.14 avremo che $\overline{\operatorname{con}} f(x) \geq F(x, \underline{a) - \mu}$ per ogni $x \in \mu$, di conseguenza $(\overline{\operatorname{con}} f)^*(a) \leq \mu$ e per l'arbitrarietà di μ allora $f^* = (\overline{\operatorname{con} f})^*$. Poiché $\overline{\operatorname{con}} f$ è convessa semipropria semicontinua inferiormente avremo anche $f^{**} = \overline{\operatorname{con}} f$. ■

Corollario 6.1.17 (Disuguaglianza di Young)**.** *Dati uno svc V e una funzione $f : V \to \overline{\mathbb{R}}$ convessa propria semicontinua inferiormente allora per ogni $x \in V$, $L \in V^*$ si ha*

$$L(x) \leq f(x) + f^*(L) \tag{6.1.3}$$

Inoltre per ogni altra coppia di funzioni $g : V \to \overline{\mathbb{R}}$, $h : V^* \to \overline{\mathbb{R}}$ tali che $g \leq f$ e $h \leq f^*$ che soddisfano la medesima disuguaglianza al posto di f e f^* si deve avere $g = f$ e $h = f^*$.

La dimostrazione si basa esclusivamente sulla definizione di f^* e sull'uguaglianza $f^{**} = f$ ed è lasciata al lettore.

Per quanto riguarda la funzione indicatrice abbiamo che

$$i_K^{**}(x) = i_{\overline{\text{con}(K)}}(x) = \sup_{a \in W} F(x,a) - i_K^*(a)$$

Da ciò segue immediatamente il seguente risultato

Proposizione 6.1.18. *Sia V spazio vettoriale topologico localmente convesso di Hausdorff e $K \subseteq V$ convesso non vuoto. Allora $x \in \mathrm{Cl}(K)$ se e solo se per ogni $f \in V^*$ si ha*

$$\sup_{y \in K} f(y - x) \geq 0$$

Rivedremo una disuguaglianza simile molto più avanti nel capitolo sugli spazi di Hilbert.

Da questa osservazione deduciamo che da ogni insieme convesso chiuso di V è associata un unica funzione supporto e che quindi insiemi chiusi convessi distinti avranno necessariamente funzioni supporto distinte. La domanda che sorge spontanea è allora quando una funzione $f : V^* \to \mathbb{R}$ è la funzione supporto di qualche sottoinsieme di V? Per rispondere a questa domanda dobbiamo necessariamente metterci in dualità in quanto non tutte le funzioni convesse semiproprie da W in $\overline{\mathbb{R}}$ sono la trasformata di Legendre di qualche altra funzione.

Concludiamo invece questa sezione calcolando la trasformata di Legendre in $\mathbb{R}$ della funzione $f(t) = |t|^p/p$ al variare di $p \in [1, +\infty)$. Chiaramente poniamo $V = W = \mathbb{R}$ e $F(r,t) = rt$ con la topologia usuale, in questo modo la terna sarà duale in entrambi i sensi e quindi potremmo applicare il teorema di dualità.

Per ogni $r \in \mathbb{R}$ vogliamo studiare i massimi della funzione $f_r(t) = rt - |t|^p/p$ al variare di t in $\mathbb{R}$, innanzitutto osserviamo che:

- quando $p > 1$ allora f_r tende a $-\infty$ quando t tende a $\pm\infty$ per ogni scelta di r, quindi f_r possiede almeno un punto di massimo assoluto;
- quando $p = 1$ e $|r| > 1$ avremo che f_r non ha massimo in $\mathbb{R}$ in quanto $\sup_{t \in \mathbb{R}} f_r(t) = +\infty$, mentre se $|r| \leq 1$ avremo che $f_r(t) \leq 0$ per ogni $t \in \mathbb{R}$.

Derivando rispetto a t quando $p > 1$ e $t \neq 0$ avremo che

$$0 = r - t|t|^{p-2} \Rightarrow t = \frac{r}{|r|}|r|^{\frac{1}{p-1}}$$

sostituendo si avrà

$$f^*(r) = |r|^{\frac{p}{p-1}} - \frac{1}{p}|r|^{\frac{p}{p-1}} = \frac{|r|^q}{q} \tag{6.1.4}$$

dove
$$\frac{1}{p}+\frac{1}{q}=1$$

Invece se $p=1$ per l'osservazione di prima avremo che

$$f^*(r)=\max_{t\in[-1,1]}(rt-|t|)=i_{[-1,1]}(r)$$

Avremmo potuto in realtà usare direttamente la proposizione 6.1.5 alla funzione $g(x)=i_{[-1,1]}(x)+1$ in quanto, come vedremo meglio più avanti, $|x|=\inf_{\lambda>0}g_\lambda(x)$.

6.2 Polari

La proposizione 6.1.4 afferma che la trasformata di Legendre di una funzione positivamente omogenea è una funzione indicatrice su W. Quindi per ogni sottoinsieme C di V esisterà un unico sottoinsieme $C^\circ\subseteq W$ tale che

$$\left[\mathrm{rad}\left(i_C+1\right)\right]^*=i_{C^\circ}$$

e l'insieme C° sarà la **polare** di C rispetto a W e F.

La proposizione 6.1.5 ci fornisce un'altra formulazione per C°

$$C^\circ=\left\{(i_C+1)^*\leq 0\right\}=\left\{i_C^*-1\leq 0\right\}=\left\{a\in W\,\middle|\,F(x,a)\leq 1\ \forall x\in C\right\}$$

chiaramente C° sarà sempre convesso e chiuso in W qualunque sia C.

Proposizione 6.2.1. *Dato un cono C in V allora C° è un cono convesso chiuso in W con $i_C^*=i_{C^\circ}$ e*
$$C^\circ=\left\{a\in W\,\middle|\,F(x,a)\leq 0\ \forall x\in C\right\}$$

Dimostrazione. Innanzitutto i_C è positivamente omogenea in quanto il suo epigrafico è un cono. Inoltre per ogni $\lambda>0$ e per ogni $x\in V$

$$\lambda\left[i_C(\lambda^{-1}x)+1\right]=i_C(x)+\lambda\to i_C(x)$$

e perciò $\mathrm{rad}(i_C+1)=i_C$ e la tesi segue dalla proposizione 6.1.5. ■

Vogliamo ora ottenere un risultato di dualità per le polari come nella sezione precedente abbiamo fatto per la trasformata di Legendre, e per farlo abbiamo bisogno di alcuni risultati preliminari. Innanzitutto per ogni sottoinsieme non vuoto C di V abbiamo che

$$\mathrm{rad}\left(i_C+1\right)(x)=\inf_{\lambda>0}\lambda\, i_C(\lambda^{-1}x)+\lambda=\inf\left\{\lambda>0\mid x\in\lambda C\right\}=p_C(x)$$

e dunque avremo che
$$p_C^*=i_{C^\circ}\tag{6.2.1}$$

Dimentichiamo per il momento le polari e concentriamoci sulle proprietà del funzionale di Minkowski e delle funzioni positivamente omogenee non negative. Per ogni $r\geq 0$ e per ogni insieme non vuoto $C\subseteq V$ definiamo innanzitutto

$$\mathrm{rad}_r(C)=\bigcap_{\varepsilon>0}\bigcup_{\lambda\in(0,r+\varepsilon)}\lambda C$$

osserviamo in particolare che per ogni $C \subseteq D$ ed $r \leq s$ abbiamo

$$\operatorname{rad}_r(C) \subseteq \operatorname{rad}_r(D) \subseteq \operatorname{rad}_s(D)$$

e per ogni $\lambda > 0$

$$\operatorname{rad}_{\lambda r}(C) = \operatorname{rad}_r(\lambda C) = \lambda \operatorname{rad}_r(C)$$

Proposizione 6.2.2. *Data una qualunque funzione $q : V \to [0, +\infty]$ positivamente omogenea avremo che*

$$q = p_{\{q \leq 1\}}$$

Inoltre dato un qualunque insieme C si ha

$$\{p_C \leq r\} = \operatorname{rad}_r(C)$$

per ogni $r \geq 0$.

Dimostrazione. Posto $C = \{q \leq 1\}$ per ogni $x \in V$ tale che $q(x) \neq +\infty$ si ha per omogeneità $q(x/\lambda) \leq 1$ per ogni $\lambda \geq q(x)$ e dunque $p_C(x) \leq q(x)$. Ora se $p_C(x) = l \neq +\infty$ allora per ogni $\varepsilon > 0$ esiste $l' < l + \varepsilon$ tale che $x \in l'C$, dunque $q(x) \leq l' < l + \varepsilon$ e per l'arbitrarietà di ε avremo che $q(x) \leq l = p_C(x)$.

L'ultimo punto è banale in quanto

$$p_C(x) \leq l \Leftrightarrow \exists l_n > 0 \text{ tale che } \lim_{n \to +\infty} l_n \leq l \text{ e } x \in l_n C \Leftrightarrow x \in \operatorname{rad}_l(C)$$

■

Lemma 6.2.3. *Dato uno svc V e un sottoinsieme D chiuso convesso di V contenente l'origine avremo che*

$$\operatorname{rad}_r(D) = rD \text{ se } r > 0$$
$$\overline{\operatorname{con}}(p_D) = p_D$$

Più in generale per ogni sottoinsieme non vuoto C di V si ha

$$\overline{\operatorname{con}}\left(p_C\right) = p_{\overline{\operatorname{con}}(C \cup \{0\})}$$

Dimostrazione. Osserviamo che $p_D = INF(E_D)$ dove

$$E_D = \left\{(x, r) \in V \times \mathbb{R} \,\middle|\, x \in \operatorname{rad}_r(D)\right\}$$

poiché D è convesso e contiene l'origine per il lemma 5.3.6 anche p_D è convessa. Per dimostrare che $\operatorname{rad}_r(D) = rD$ osserviamo innanzitutto che $rD \subseteq \operatorname{rad}_r(D)$ per ogni $r > 0$, viceversa per ogni $x \in \operatorname{rad}_r(D)$ esiste una successione $r_n > 0$ tale che $r' = \lim_{n\to+\infty} r_n \leq r$ e $x \in r_n D$ per ogni $n \in \mathbb{N}$. Ora poiché D è convesso e contiene l'origine avremo che $aD \subseteq bD$ per ogni $a \leq b$ e dunque possiamo supporre che $r_n > r$ e $r' = r$. Ma allora $x/r_n \in D$ e poiché D è chiuso in uno spazio vettoriale topologico avremo che $x/r \in D$ e $x \in rD$. Dal lemma 6.1.10 segue immediatamente che p_D è semicontinua inferiormente in quanto $\{p = 0\} = \bigcap_{r>0} rD$ è comunque chiuso.

Prendiamo adesso un qualunque sottoinsieme C di V non vuoto, allora banalmente avremo $\overline{\operatorname{con}}(E_C) \subseteq E_{\overline{\operatorname{con}}(C \cup \{0\})}$, la disuguaglianza opposta è equivalente a dimostrare che

$$E_{\overline{\operatorname{con}}(C \cup \{0\})} \subseteq \overline{\operatorname{con}}\left(E_C\right)$$

Preso allora (x, r) tale che $x \in \mathrm{rad}_r[\overline{\mathrm{con}}(C \cup \{0\})]$ e un qualunque aperto U contenente (x, r) esisterà una successione $r_n > r$ per cui $\lim_{n\to+\infty} r_n = r$ e $x/r_n \in \overline{\mathrm{con}}(C \cup \{0\})$.

Poniamo allora

$$U' = \left\{y \in V \middle| \ (r_n y, r_n) \in U\right\}$$

il quale è un aperto di V contenente x/r_n che a sua volta appartiene alla chiusura di $\mathrm{con}(C \cup \{0\})$, allora per il corollario 5.1.7 e per la proposizione 5.1.5 esistono dei punti $x_1, \ldots, x_k \in C$ e degli scalari $\lambda_1, \ldots, \lambda_k \in [0, 1]$ tali che

$$\bar{\lambda} = \sum_{i=1}^{k} \lambda_i \leq 1$$
$$\sum_{i=1}^{k} \lambda_i x_i \in U'$$

e dunque $(r_n x_i, r_n) \in E_C$ e $\sum_{i=1}^{k} \lambda_i (r_n x_i, r_n) + (1 - \bar{\lambda})(0, r_n) \in U$.

Prendiamo adesso un qualunque insieme $D \subseteq V \times \mathbb{R}$ convesso chiuso contenente E_C. Poiché C è non vuoto scegliamo un qualunque suo elemento $\bar{x}$ e perciò $(\bar{x}, 1) \in D$. Ma allora per ogni $t > 0$ avremo che $(\bar{x}/t, 1/t + r_n) \in E_C \subseteq D$ e dunque poiché D è chiuso passando al limite $(0, r_n) \in D$.

Dunque

$$\sum_{i=1}^{k} \lambda_i (r_n x_i, r_n) + (1 - \bar{\lambda})(0, r_n) \in U \cap D$$

e perciò $U \cap D \neq \emptyset$ per ogni aperto U contenente (x, r) ovvero $(x, r) \in D$ e perciò $E_{\overline{\mathrm{con}}(C\cup\{0\})} \subseteq D$, per l'arbitrarietà di D invece segue che $E_{\overline{\mathrm{con}}(C\cup\{0\})} \subseteq \overline{\mathrm{con}}(E_C)$ e il lemma è così dimostrato. ■

Osservazione. A differenza del caso finito dimensionale il funzionale p_C potrebbe non essere semicontinuo inferiormente nemmeno se $W = V^*$. Basta prendere ad esempio una funzione lineare $f : V \to \mathbb{R}$ non continua e porre $q(x) = \left|f(x)\right|$ in quanto $\{q = 0\} = \ker f$ non è chiuso in V per la proposizione 4.3.3.

Possiamo ora dimostrare il teorema di dualità per le polari:

Teorema 6.2.4. *Se (V, W, F) è duale in un verso allora per ogni sottoinsieme C di V si ha*

$$C^{\circ\circ} = \overline{\mathrm{con}}(C \cup \{0\})$$
$$C^{\circ} = [\overline{\mathrm{con}}(C \cup \{0\})]^{\circ}$$

mentre se $(V, W mF)$ è duale nel verso opposto allora C° è chiuso, convesso e contiene l'origine

Dimostrazione. Applicando i risultati precedenti e il teorema di dualità si ha

$$i_{C^{\circ}} = p_C^* = p^*_{\overline{\mathrm{con}}(C\cup\{0\})} = i_{[\overline{\mathrm{con}}(C\cup\{0\})]^{\circ}}$$
$$p_{\overline{\mathrm{con}}(C\cup\{0\})} = i^*_{C^{\circ}}$$

e quindi $C^{\circ} = [\overline{\mathrm{con}}(C \cup \{0\})]^{\circ}$. Inoltre per la proposizione 6.1.5 applicata alla definizione di polare

$$i_{C^{\circ\circ}} = i_{\{p_{\overline{\mathrm{con}}(C\cup\{0\})} \leq 1\}} = i_{\overline{\mathrm{con}}(C\cup\{0\})}$$

in quanto i^*_C è una funzione positivamente omogenea, dunque $C^{\circ\circ} = \overline{\mathrm{con}}(C \cup \{0\})$. ■

Una volta definito l'insieme polare possiamo introdurre il concetto di funzione polare. A differenza della trasformata di Legendre non definiremo la polare per tutte le possibili funzioni, ma solo per quelle appartenenti alla classe di funzioni

$$\mathfrak{P}(V) = \left\{ f : V \to \overline{\mathbb{R}} \;\middle|\; f \geq 0 \text{ e } f(0) = 0 \right\}$$

Osserviamo anche che tutti gli elementi di $\mathfrak{P}(V)$ sono funzioni proprie e per ogni $f \in \mathfrak{P}(V)$ abbiamo chiaramente che $f^* \in \mathfrak{P}(W)$ in quanto

$$f^*(a) \geq F(0,a) - f(0) = 0$$
$$f^*(0) = -\inf_{x\in V} f(x) = 0$$

Definizione 6.2.5. Data una qualunque applicazione $f \in \mathfrak{P}(V)$ definiamo la **polare** di f la funzione

$$f^\circ(a) = INF\left[L(\text{epi}_c f)^\circ\right]$$

dove $L : (a,r) \in W \times \mathbb{R} \to (a,-r) \in W \times \mathbb{R}$.

L'applicazione L è necessaria per rendere la polare di un epigrafico ancora un epigrafico. Osserviamo infatti che

$$\begin{aligned}(a,s) \in \left(\text{epi}_c f\right)^\circ &\Leftrightarrow F(x,a) + rs \leq 1 \ \forall x \in \text{dom} f, 0 \leq f(x) \leq r \\ &\Leftrightarrow s \leq \frac{1 - F(x,a)}{r} \ \forall x \in \text{dom} f, 0 \leq f(x) \leq r\end{aligned}$$

e quindi $L\left(\text{epi}_c f\right)^\circ$ è un epigrafico contenuto nel semispazio $W \times [0, +\infty)$. Più nel dettaglio abbiamo che

$$\begin{aligned}(a,s) \in L(\text{epi}_c f)^\circ &\Leftrightarrow (a,-s) \in (\text{epi}_c f)^\circ \\ &\Leftrightarrow F(x,a) - sr \leq 1 \ \forall x \in \text{dom} f, f(x) \leq r \\ &\Leftrightarrow s \geq 0 \text{ e } F(x,a) - sf(x) \leq 1 \ \forall x \in \text{dom} f\end{aligned}$$

e dunque $f^\circ \in \mathfrak{P}(W)$. Abbiamo così dimostrato che

$$f^\circ(a) = \inf\left\{\lambda > 0 \;\middle|\; F(x,a) \leq 1 + \lambda f(x) \ \forall x \in \text{dom} f\right\}$$

Un caso degno di nota è il calcolo della polare di una funzione positivamente omogenea. Se infatti $p \in \mathfrak{P}(V)$ è anche positivamente omogenea allora avremmo che

$$p^\circ(a) = \inf\left\{\lambda > 0 \;\middle|\; F(x,a) \leq \lambda p(x) \ \forall x \in \text{dom}\, p\right\}$$

in particolare anche p° è positivamente omogenea ma soprattutto

$$p^\circ(a) = \sup_{p(x)\leq 1} F(x,a)$$

Proposizione 6.2.6. *Dato un qualunque insieme $C \subseteq V$cvx:kza si ha*

$$i_C^\circ = i_{C^\circ} \tag{6.2.2}$$
$$p_C^\circ = p_{C^\circ} \tag{6.2.3}$$
$$p_C^* = i_C^\circ \tag{6.2.4}$$
$$i_C^* = p_C^\circ \tag{6.2.5}$$

Dimostrazione. La (6.2.2) discende immediatamente dalle definizioni di polare di un insieme e di polare di una funzione, e di conseguenza si ottiene anche la (6.2.4).

Se prendiamo $\lambda > 0$ in modo tale che $p_C^\circ(a) < \lambda$ per qualche $a \in W$ allora $F(x,a) < \lambda p_C(x) = p_C(\lambda x)$ per ogni $x \in V$ e in particolare $F(x, a/\lambda) \leq 1$ per ogni $x \in C$ e dunque $p_{C^\circ}(a) \leq \lambda$. Viceversa se prendiamo $p_{C^\circ}(a) < \lambda$ essendo C° convesso contenente l'origine avremo che $x/\lambda \in C^\circ$ e perciò $F(x,a) \leq \lambda$ per ogni $x \in C$. Se prendiamo $x \in V$ per cui $p(x) < +\infty$ esisterà $l > p(x)$ sufficientemente vicino ad esso per cui $x/l \in C$ e dunque $F(x,a) \leq \lambda l$ e perciò $F(x,a) \leq \lambda p(x)$. Dato che questa disuguaglianza vale anche per $p(x) = +\infty$ si ha $p_C^\circ(a) \leq \lambda$ dimostrando in tal modo la (6.2.3).

Per quanto riguarda la (6.2.5) osserviamo semplicemente che $i_C^*(a) = \sup_{x \in C} F(x,a)$ e quindi ci basta applicare la proposizione 6.2.2 e la (6.2.3). ■

Corollario 6.2.7. *Se $p \in \mathfrak{P}(V)$ è positivamente omogenea allora $\{p^\circ \leq 1\} = \{p \leq 1\}^\circ$.*

Dimostrazione. Posto $C = \{p \leq 1\}$ sappiamo già che $p = p_C$. Ora poiché C° è chiuso, convesso e contiene l'origine per la (6.2.3) e per il lemma 6.2.3 abbiamo che $\{p^\circ \leq 1\} = \{p_{C^\circ} \leq 1\} = C^\circ$. ■

Questo risultato vale solo se p è positivamente omogenea ma non in generale.

Come per la trasformata di Legendre abbiamo anche per le polari un risultato di omogeneità:

Teorema 6.2.8. *Data una qualunque funzione $f \in \mathfrak{P}(V)$ con (V, W, F) duale in un verso si ha*

$$f^\circ = (\overline{\mathrm{con}} f)^\circ$$
$$f^{\circ\circ} = \overline{\mathrm{con}} f$$

Dimostrazione. Dobbiamo solo verificare che $\left(\mathrm{epi}_c f^\circ\right)^\circ = \left[L\left(\mathrm{epi}_c f\right)\right]^{\circ\circ}$. Dati due epigrafici $A, B \subseteq V \times \mathbb{R}$ contenenti l'origine tali che $A \sim B$ abbiamo già dimostrato che $\overline{\mathrm{con}}(A) = \overline{\mathrm{con}}(B)$ e perciò

$$A^\circ = \left[\overline{\mathrm{con}}(A)\right]^\circ = \left[\overline{\mathrm{con}}(B)\right]^\circ = B^\circ$$

La tesi segue quindi osservando che $\mathrm{epi}_c f^\circ \sim L\left(\mathrm{epi}_c f\right)^\circ$ e applicando il teorema di dualità. ■

Inoltre come per la trasformata di Legendre se $f \in \mathfrak{P}(V)$ è convessa e semicontinua inferiormente allora per ogni $x \in V, a \in W$ avremo che

$$F(x,a) \leq 1 + f(x) f^\circ(a)$$

Consideriamo adesso una funzione $r : [0, +\infty] \to [0, +\infty]$ strettamente crescente con $r(0) = 0$ e $r(+\infty) = +\infty$, per il prossimo risultato risulta utile conoscere la trasformata di Legendre della funzione $r'(t) = r(|t|)$, per cui poniamo per ogni $s \geq 0$

$$r^+(s) = \left(r'\right)^*(s) = \sup_{t \in \mathbb{R}} \left[st - r'(t)\right] = \sup_{t \geq 0} [st - r(t)]$$

Proposizione 6.2.9. *Date una funzione $p \in \mathfrak{P}(V)$ positivamente omogenea ed una funzione $r : [0, +\infty] \to [0, +\infty]$ strettamente crescente con $r(0) = 0$ e $r(+\infty) = +\infty$ allora*

$$(r \circ p)^* = r^+ \circ p^\circ$$

Dimostrazione. per ogni $a \in W$ si ha

$$(r \circ p)^*(a) = \sup_{x \in V}[F(x,a) - r[p(x)]] = \sup_{\lambda>0, p(x)\leq\lambda}[F(x,a) - r(\lambda)]$$
$$= \sup_{\lambda>0}[\lambda p^\circ(a) - r(\lambda)] = r^+\left[p^\circ(a)\right]$$

■

Corollario 6.2.10. *Data una qualunque funzione $f \in \mathfrak{P}(V)$ positivamente omogenea per ogni $p, q > 1$ tali che $1/p + 1/q = 1$ si ha*

$$\left(\frac{f^p}{p}\right)^*(a) = \frac{[f^\circ(a)]^q}{q}$$

per ogni $a \in W$.

Esempio 6.2.11. Presi $p, q > 1$ tali che $1/p + 1/q = 1$ per ogni $x \in \mathbb{R}^n$ definiamo

$$p_1(x) = \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}$$
$$p_2(x) = \sqrt[q]{|x_1|^q + \cdots + |x_n|^q}$$

Entrambe sono funzioni positivamente omogenee che si annullano nell'origine, vogliamo dimostrare che sono l'una la polare dell'altra. Innanzitutto abbiamo che

$$\left(\frac{p_1^p}{p}\right)^*(y) = \sup_x\left[x_1y_1 + \cdots + x_ny_n - \frac{|x_1|^p + \cdots + |x_n|^p}{p}\right]$$
$$= \sup_{x_1}\left(x_1y_1 - \frac{|x_1|^p}{p}\right) + \cdots + \sup_{x_n}\left(x_ny_n - \frac{|x_n|^p}{p}\right) = \frac{|y_1|^q + \cdots + |y_n|^q}{q}$$

e perciò per la proposizione precedente

$$\frac{p_2^q}{q} = \left[\frac{p_1^p}{p}\right]^* = \frac{(p_1^\circ)^q}{q}$$

ovvero $p_1^\circ = p_2$.

Concludiamo la sezione con un un risultato che ci permetta di passare dalla trasformata di Legendre di funzioni non negative alla polare e viceversa. Definiamo prima di tutto l'applicazione O_V che ad ogni $f \in \mathfrak{P}(V)$ associ la funzione $O_V(f) \in \mathfrak{P}(V)$ in modo tale che

$$O_V(f)(x) = \inf\{\lambda > 0 \mid f_\lambda(x) \leq 1\}$$

osserviamo che se f fosse convessa allora $\mathrm{epi}_c f$ sarebbe un insieme convesso contenente l'origine e dunque $l\,\mathrm{epi}_c f \subseteq r\,\mathrm{epi}_c f$ per ogni $l \leq r$ e perciò l'applicazione

$$\lambda > 0 \to f_\lambda(x)$$

è decrescente per ogni $x \in V$.

Teorema 6.2.12. *Per ogni $f \in \mathfrak{P}(V)$ abbiamo $O_W(f^*) = f^\circ$ e $O_W(f^\circ) = f^*$.*

Dimostrazione. Per ogni $\lambda > 0$ abbiamo

$$\begin{aligned} f^\circ(a) < \lambda &\Leftrightarrow F(x,a) \le 1 + \lambda f(x) \ \forall x \in V \\ &\Leftrightarrow \lambda F(x, \lambda^{-1}a) - \lambda f(x) \le 1 \ \forall x \in V \\ &\Leftrightarrow \lambda f^*(\lambda^{-1}a) \le 1 \\ &\Leftrightarrow O_W(f^*)(a) < \lambda \end{aligned}$$

e quindi la prima disuguaglianza è verificata. Per quanto riguarda la seconda uguaglianza abbiamo

$$\begin{aligned} f^*(a) < \lambda &\Leftrightarrow F(x,a) < \lambda + f(x) \ \forall x \in V \\ &\Leftrightarrow F(x, \lambda^{-1}a) < 1 + \lambda^{-1} f(x) \ \forall x \in V \\ &\Leftrightarrow f^\circ(\lambda^{-1}a) < \lambda^{-1} \\ &\Leftrightarrow O_W(f^\circ)(a) < \lambda \end{aligned}$$

e quindi anche l'uguaglianza opposta è verificata. ■

Corollario 6.2.13. *Data una qualunque funzione $f \in \mathfrak{P}(V)$ si ha $f^{*\circ} = f^{\circ *}$.*

Dimostrazione. Dal risultato precedente abbiamo che

$$f^{*\circ} = O_V(f^{**}) = O_V(\overline{\mathrm{con}} f) = O_V(f^{\circ\circ}) = f^{\circ *}$$

■

6.3 Il teorema di Fenchel-Rockafellar

Il lemma 6.1.7 ci permette anche di usare il primo teorema di Mazur, ovvero di separare due insiemi convessi di cui uno dei due è aperto. In questo caso però le topologie scelte su V in modo tale che la terna (V, W, F) sia duale in un verso non sono equivalenti poiché un insieme convesso potrebbe essere aperto rispetto ad una topologia ma non rispetto ad un'altra, a differenza di come avviene per gli insiemi convessi chiusi.

Ricordiamo che in uno spazio vettoriale topologico chiusura e interno di un insieme convesso sono ancora insiemi convessi. Dimostriamo ora che preso un qualunque insieme convesso $C \subseteq V$ se $\mathrm{Int}(C) \neq \emptyset$ allora

$$\mathrm{Cl}(C) = \mathrm{Cl}\big(\mathrm{Int}(C)\big)$$

Innanzitutto $\mathrm{Cl}\big(\mathrm{Int}(C)\big) \subseteq \mathrm{Cl}(C)$, supponiamo adesso che $\mathrm{Int}(C) \neq \emptyset$ e prendiamo $x \in C$ generico, allora avremo che

$$\lambda x + (1-\lambda)\mathrm{Int}(C) \subseteq C$$

e dato che il primo membro è aperto segue che $\lambda x + (1-\lambda)x_0 \in \mathrm{Int}(C) \subseteq \mathrm{Cl}\big(\mathrm{Int}(C)\big)$ per ogni $x_0 \in \mathrm{Int}(C)$. Fissato x_0 e facendo tendere λ a 1 avremo così che $x \in \mathrm{Cl}\big(\mathrm{Int}(C)\big)$ e perciò $C \subseteq \mathrm{Cl}\big(\mathrm{Int}(C)\big)$ da cui discende l'uguaglianza cercata.

Prendiamo ora una qualunque funzione $f : V \to \overline{\mathbb{R}}$ non identicamente $+\infty$ e quindi con $\mathrm{epi}_o f \neq \emptyset$. Osserviamo che

$$\mathrm{Int}\big(\mathrm{epi}_o f\big) \neq \emptyset \Leftrightarrow \exists U \subseteq V \text{ aperto}, r \in \mathbb{R} \text{ tale che } f(x) \le r \ \forall x \in U$$

in particolare se f è continua in un punto di $\operatorname{dom} f$ allora $\operatorname{Int}(\operatorname{epi}_o f) \neq \emptyset$. Dimostreremo ora che se l'interno dell'epigrafico è non vuoto e se f è convessa e propria allora sarà continua in $\operatorname{Int}(\operatorname{dom} f) \neq \emptyset$, prima di tutto abbiamo bisogno di un lemma tecnico

Lemma 6.3.1. *Dato uno spazio vettoriale reale V e una funzione $f : V \to \overline{\mathbb{R}}$ convessa, siano $x, y \in V$ tali che $x, x+y, x-y \in \operatorname{dom} f$. Allora per ogni $\delta \in [0,1]$*

$$\left|f(x \pm \delta y) - f(x)\right| \leq \delta \left[(f(x+y) - f(x)) \vee (f(x-y) - f(x))\right]$$

Dimostrazione. Per convessità abbiamo

$$f(x+\delta y) = f[\delta(x+y) + (1-\delta)x] \leq \delta f(x+y) + (1-\delta) f(x)$$

e quindi abbiamo

$$\begin{aligned} f(x+\delta y) - f(x) &\leq \delta[f(x+y) - f(x)] \\ f(x-\delta y) - f(x) &\leq \delta[f(x-y) - f(x)] \end{aligned}$$

Sempre sfruttando la convessità di f abbiamo che $2f(x) \leq f(x-\delta y) + f(x+\delta y)$ e perciò

$$\begin{aligned} f(x) - f(x+\delta y) &\leq f(x-\delta y) - f(x) \\ f(x) - f(x-\delta y) &\leq f(x+\delta y) - f(x) \end{aligned}$$

che combinate con le precedenti disuguaglianze ci permettono di applicare i valori assoluti ottenendo così la tesi. ■

Teorema 6.3.2 (Continuità locale)**.** *Dato uno spazio vettoriale topologico V e una funzione $f : V \to \overline{\mathbb{R}}$ convessa propria. Se esistono $x \in \operatorname{dom} f$, $r \geq 0$ tali che $f(y) \leq f(x) + r$ per ogni $y \in U$ aperto contenente x allora f è continua in x.*

Dimostrazione. Possiamo tranquillamente supporre che $U = x + U'$ con U' aperto bilanciato contenente l'origine. Preso $\varepsilon > 0$ poniamo $\delta = \varepsilon/(r+1) > 0$, allora per ogni $y \in x + \delta U'$ esisterà $y' \in U'$ tale che $y = x + \delta y'$. Poiché $x + y', x - y' \in U$ allora per il lemma precedente si ha

$$\left|f(y) - f(x)\right| \leq \delta \left[(f(x+y') - f(x)) \vee (f(x-y') - f(x))\right] \leq \varepsilon$$

e per l'arbitrarietà di $y \in x + \delta U'$ segue la tesi. ■

Osservazione. Se $f : V \to \overline{\mathbb{R}}$ è convessa e continua in un punto $x_0 \in \operatorname{dom} f$ allora dalla definizione di continuità avremo che $x_0 \in \operatorname{Int}(\operatorname{dom} f)$ e inoltre f è propria. Infatti se esistesse $x \in V$ per cui $f(x) = -\infty$ per continuità esisterebbe $\lambda \in [0,1]$ tale che $\lambda x + (1-\lambda)x_0 \in \operatorname{dom} f$ e perciò

$$-\infty < f[\lambda x + (1-\lambda)x_0] \leq \lambda f(x) + (1-\lambda) f(x_0) = -\infty$$

assurdo.

Data una generica funzione $f : V \to \overline{\mathbb{R}}$ definita su uno spazio vettoriale topologico V diremo che f è **continua sul dominio** se e solo se $D = \operatorname{Int}(\operatorname{dom} f) \neq \emptyset$ ed f risulta continua in tutti i punti di D. Si noti che una funzione continua sul dominio può tranquillamente non essere semicontinua inferiormente su tutto V, nemmeno se f fosse convessa.

Teorema 6.3.3 (Continuità globale)**.** *Data una funzione convessa propria $f : V \to \overline{\mathbb{R}}$ definita su uno spazio vettoriale topologico e sia $D = \operatorname{Int}(\operatorname{dom} f) \neq \emptyset$, allora sono equivalenti:*

1. *f è continua sul dominio;*
2. *$-f$ è semicontinua inferiormente su D;*
3. *per ogni $x \in D$ la funzione f è superiormente limitata in un intorno aperto di x;*
4. *esiste $x \in D$ tale che f è superiormente limitata in un intorno aperto di x;*
5. $\operatorname{Int}(\operatorname{epi}_o f) \neq \emptyset$*;*
6. *f è continua in un punto di* $\operatorname{dom} f$.

Dimostrazione. Le implicazioni $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4$ sono banali da verificare in quanto $-f$ è semicontinua inferiormente se e solo se $f(x) \geq \limsup_{y \to x} f(y)$. Chiaramente 4 e 5 sono equivalenti, il teorema di continuità locale permette di affermare che $4 \Leftrightarrow 6$.

Dimostriamo ora che $5 \Rightarrow 1$ per concludere la dimostrazione. Prendiamo $x \in D$, $r \in \mathbb{R}$ tali che $(x,r) \in \operatorname{Int}(\operatorname{epi}_o f)$ e sia $y \in D$ generico con $x \neq y$. Poiché D è aperto in V esisterà $\delta > 0$ tale che $\bar{y} = y + \delta(y-x) \in D$ quindi possiamo prendere $s \in \mathbb{R}$ tale che $(\bar{y}, s) \in \operatorname{epi}_o f$.

Adesso osserviamo che

$$y = \frac{1}{1+\delta} y + \frac{\delta}{1+\delta} y + \frac{\delta}{1+\delta} x - \frac{\delta}{1+\delta} x = \frac{1}{1+\delta}\left[y + \delta(y-x)\right] + \frac{\delta}{1+\delta} x$$

e questo implica che esiste $t \in \mathbb{R}$ tale che

$$(y,t) \in \frac{1}{1+\delta}(\bar{y}, s) + \frac{\delta}{1+\delta} \operatorname{Int}(\operatorname{epi}_o f) \subseteq \operatorname{Int}(\operatorname{epi}_o f)$$

possiamo allora applicare l'implicazione $5 \Rightarrow 6$ con y al posto di x. Ma per l'arbitrarietà di $y \in D$ avremo che f è continua sull'intero dominio, ottenendo in tal modo la tesi. ■

Ora dimostriamo il

Teorema 6.3.4 (Fenchel-Rockafellar)**.** *date due funzioni convesse f, g su uno svc V in modo tale che (V, W, F) sia duale in un verso. Se esiste $x_0 \in$* $\operatorname{dom} f \cap \operatorname{dom} g$ *tale che f sia continua in x_0 allora*

$$\inf_{x \in V} \left[f(x) + g(x)\right] = \max_{e \in W} \left[-f^*(-e) - g^*(e)\right]$$

Dimostrazione. Poniamo per comodità

$$\begin{aligned} \mathfrak{a} &= \inf_{x \in V} \left[f(x) + g(x)\right] \\ \mathfrak{b} &= \sup_{e \in W} \left[-f^*(-e) - g^*(e)\right] \end{aligned}$$

se $\mathfrak{a} = -\infty$ allora almeno uno tra $f^*(0)$ e $g^*(0)$ vale $+\infty$, dunque possiamo supporre che $\mathfrak{a}$ sia finito in quanto non può assumere per ipotesi il valore $+\infty$ e in particolare che le funzioni f, g siano proprie. Ora per ogni $x \in V$, $e \in W$ utilizzando la (6.1.1) abbiamo $g(x) + g^*(e) \geq F(x,e)$ e $f(x) + f^*(-e) \geq -F(x,e)$ dunque $g(x) + g^*(e) \geq -f(x) - f^*(-e)$ da cui segue immediatamente che $\mathfrak{a} \geq \mathfrak{b}$.

Adesso essendo f continua in un punto del dominio avremo che $\operatorname{Int}(\operatorname{dom} f) \neq \emptyset$ e per il teorema di continuità globale $O = \operatorname{Int}(\operatorname{epi}_o f) \neq \emptyset$. Se poniamo

$$P = \left\{(x,r) \in V \times \mathbb{R} \;\middle|\; g(x) + r \leq \mathfrak{a}\right\}$$

possiamo scegliere r sufficientemente piccolo in modo tale che $(x_0, r) \in P$, inoltre se $(x, r) \in O$ allora $f(x) < r$ e dunque $\mathfrak{a} \leq f(x) + g(x) < g(x) + r$ quindi $O \cap P = \emptyset$. Possiamo così applicare il primo teorema di Mazur alla terna $(V \times \mathbb{R}, W \times \mathbb{R}, \tilde{F})$ così da trovare $a \in W$, $t, u \in \mathbb{R}$ tali che

$$(x, r) \in O \Rightarrow F(x, a) + rt > u$$
$$(x, r) \in P \Rightarrow F(x, a) + rt \leq u$$

Essendo $\mathrm{Cl}(O) = \mathrm{Cl}(\mathrm{epi}_c f)$ possiamo separare anche $\mathrm{epi}_c f$ da P, preso allora $\lambda \geq f(x_0)$ osserviamo che $F(x_0, a) + \lambda t \geq u$ facendo tendere λ a $+\infty$ avremo che $t \geq 0$. Se $t = 0$ allora l'applicazione $x \to F(x, a)$ non può essere identicamente nulla, d'altronde esiste un aperto $U \subseteq V$ contenente l'origine tale che $x_0 + U \subseteq \mathrm{dom} f$ e quindi $F(x_0 + y, a) \geq u$ per ogni $y \in U$. Ma allora

$$u \geq F(x_0, a) \geq u - F(y, a)$$

il che è assurdo in quanto possiamo sempre trovare $y \in U$ in modo tale che $F(y, a) < 0$, dunque $t > 0$.

Poiché $-F(x, a/t) - f(x) = F(x, -a/t) - f(x) \leq -u/t$ si ha $-f^*(-a/t) \geq u/t$, inoltre essendo $[x, \mathfrak{a} - g(x)] \in P$ per ogni $x \in \mathrm{dom} f$ segue che $F(x, a/t) + \mathfrak{a} - g(x) \leq u/t$ e dunque $g^*(a/t) \leq u/t - \mathfrak{a}$. Combinando le due disuguaglianze avremo che

$$\mathfrak{b} \geq -f^*\left(-\frac{a}{t}\right) - g^*\left(\frac{a}{t}\right) \geq \mathfrak{a}$$

dunque $\mathfrak{a} = \mathfrak{b}$ e il punto massimo è raggiunto da $e = a/t$. ∎

Corollario 6.3.5. *Se f e g soddisfano le ipotesi del teorema di Frenchel-Rockafellar allora*

$$(f + g)^*(a) = \left(f^* \diamond g^*\right)(a) = \min_{b \in W} [f^*(a - b) + g^*(b)]$$

Dimostrazione. Innanzitutto ponendo $h(x) = g(x) - F(x, a)$ non è difficile dimostrare che $h^*(b) = \sup_{x \in V} [F(x, b) + F(x, a) - g(x)] = g^*(a + b)$, dunque

$$(f + g)^*(a) = -\inf_{x \in V} [f(x) + g(x) - F(x, a)] = -\max_{b \in W} [-f^*(b) - g^*(a - b)] = (f^* \diamond g^*)(a)$$

∎

Diamo ora uno sguardo all'uguaglianza (6.1.2) che ci è stata fondamentale per dimostrare il teorema di Fenchel-Morenau. In realtà l'estremo superiore potrebbe non essere mai raggiunto, nemmeno se $f(x) \in \mathbb{R}$. Per l'esistenza del massimo abbiamo bisogno della continuità della funzione in un punto, in quanto

Corollario 6.3.6. *Se $f : V \to \overline{\mathbb{R}}$ è convessa e continua in $x_0 \in \mathrm{dom} f$ allora esiste $a \in W$ tale che*

$$f(x_0) + f^*(a) = F(x, a)$$

Dimostrazione. Per convessità abbiamo $f(x_0) \leq f(x_0 + y)/2 + f(x_0 - y)/2$ mentre per continuità segue immediatamente che

$$f(x_0) = \inf_{y \in V} \left[\frac{f(x_0 + y)}{2} + \frac{f(x_0 - y)}{2}\right]$$

Posto $f_1(y) = f(x_0 + y)$ e $f_2(y) = f(x_0 - y)$ per le proprietà della trasformata di Legendre abbiamo che $f_1^*(a) = f^*(a) - F(x_0, a)$, $f_2^*(a) = f^*(-a) + F(x_0, a)$ quindi per il teorema di Fenchel-Rockafellar avremo che

$$2f(x_0) = \max_{b\in W}\left[-f^*(b) + F(x_0, b) - f^*(b) + F(x_0, b)\right] = 2[F(x_0, a) - f^*(a)]$$

per un qualche $a \in W$, da cui la tesi. ■

6.4 Sottodifferenziale

Data una qualunque funzione convessa $\varphi : V \to \overline{\mathbb{R}}$ definita su uno spazio vettoriale reale V per la proposizione 5.2.17 per ogni $x \in \operatorname{dom}\varphi$ e per ogni $v \in V \setminus \{0\}$ il rapporto incrementale

$$\frac{\varphi(x + \lambda v) - \varphi(x)}{\lambda}$$

è crescente in $\lambda > 0$. Dunque il limiti destro e sinistro quando λ va a 0 esistono sempre anche se potrebbero non coincidere. Questa proprietà è stata utilizzata per definire il sottodifferenziale delle funzioni convesse su $\mathbb{R}$, in questa sezione estenderemo il concetto di sottodifferenziale di funzioni convesse su spazi vettoriali generici.

Innanzitutto consideriamo come sempre due spazi vettoriali reali V, W e una funzione bilineare $F : V \times W \to \mathbb{R}$, data una qualunque funzione convessa $f : V \to \overline{\mathbb{R}}$ definiamo il **sottodifferenziale** di f l'insieme

$$\partial^* f = \begin{cases} \left\{(x, a) \in V \times W \,\middle|\, f(y) - f(x) \geq F(y - x, a)\ \forall y \in V\right\} & \text{se } x \in \operatorname{dom} f \\ \emptyset & \text{se } x \notin \operatorname{dom} f \end{cases}$$

e definiamo anche i seguenti insiemi

$$\partial^* f(x) = \left\{a \in W \,\middle|\, (x, a) \in \partial^* f\right\}$$
$$\operatorname{dom}\partial^* f = \left\{x \in V \,\middle|\, \partial^* f(x) \neq \emptyset\right\}$$

A differenza del caso unidimensionale il sottodifferenziale di una funzione convessa può essere vuoto in ogni punto, anche se la funzione assume solamente valori reali, per esempio se (V, W, F) è duale in un verso basta prendere una qualunque $f \in V^+ \setminus V^*$ ovvero una funzione lineare non continua.

Proposizione 6.4.1. *Il sottodifferenziale $\partial^* f(x)$ è un sottoinsieme convesso di W, inoltre se (V, W, F) è duale nel verso opposto allora è anche chiuso.*

Dimostrazione. Presi $a, b \in \partial^* f(x)$ per ogni $\lambda \in [0, 1]$ e per ogni $y \in V$ si avrà

$$F[y - x, \lambda a + (1 - \lambda)b] \leq \lambda[f(y) - f(x)] + (1 - \lambda)[f(y) - f(x)] = f(y) - f(x)$$

e quindi il sottodifferenziale è convesso. Supponiamo ora che la terna (V, W, F) sia duale nel verso opposto, in particolare per ogni $a_n, a \in W$ tali che $a_n \to a$ e per ogni $x \in V$ si avrà per continuità $F(x, a_n) \to F(x, a)$.

Sia allora $a_n \in \partial^* f(x)$ e $a \in W$ tali che $a_n \to a$, per ogni $y \in V$ avremo che

$$F(y - x, a) = \lim_{n\to+\infty} F(y - x, a_n) \leq f(y) - f(x)$$

e perciò $a \in \partial^* f(x)$ e il sottodifferenziale è chiuso. ■

Proposizione 6.4.2. *Per ogni* $(x,a) \in V \times W$ *si ha*

$$(x,a) \in \partial^* f \Leftrightarrow f(x) + f^*(a) = F(x,a)$$

Dimostrazione. Supponiamo inizialmente che $(x,a) \in \partial^* f$, per la definizione di trasformata di Legendre per ogni $y \in V$ si avrà

$$f^*(a) \geq F(x,a) - f(x) \geq F(y,a) - f(y)$$

dove la seconda disuguaglianza discende dalla definizione di sottodifferenziale, passando all'estremo superiore la prima disuguaglianza diventa un uguaglianza.

Viceversa se $f(x) + f^*(a) = F(x,a)$ allora $x \in \operatorname{dom} f$ e $F(x,a) - f(x) = f^*(a) \geq F(y,a) - f(y)$ e dunque $a \in \partial^* f(x)$. ■

Corollario 6.4.3. *Data una funzione convessa* $f : V \to \overline{\mathbb{R}}$ *e due elementi* $x \in V, a \in W$ *se* $(x,a) \in \partial^* f$ *allora* $(a,x) \in \partial^* f^*$. *In particolare* $\partial^* f \subseteq \partial^* f^{**}$.

Dimostrazione. Discende immediatamente dalla catena di disuguaglianze $f(x) + f^*(a) \geq f^{**}(x) + f^*(a) \geq F(x,a)$. ■

Corollario 6.4.4. *Data una funzione* f *convessa propria semicontinua inferiormente definita su* V *avremo che* $\partial^* f^*(0)$ *è l'insieme dei minimi di* f.

Dimostrazione. Osserviamo che $x \in V$ è un punto di minimo di f se e solo se $x \in \operatorname{dom} f$ e $f(y) \geq f(x)$ per ogni $y \in V$ ovvero $0 \in \partial^* f(x)$. Essendo semicontinua inferiormente avremo che $\partial^* f = \partial^* f^{**}$ dimostrando così l'asserto. ■

Proposizione 6.4.5. *Date due funzioni convesse* f, g *su* V *e un elemento* $x \in \operatorname{dom} \partial^* f \cap \operatorname{dom} \partial^* g$. *Allora*

$$\partial^* f(x) + \partial^* g(x) \subseteq \partial^*(f+g)(x)$$

Dimostrazione. Chiaramente si ha $x \in \operatorname{dom} f \cap \operatorname{dom} g = \operatorname{dom}(f+g)$, siano $a \in \partial^* f(x)$, $b \in \partial^* g(x)$ allora si ha

$$F(y-x, a+b) \leq f(y) - f(x) + g(y) - g(x) = (f+g)(y) - (f+g)(x)$$

e perciò $a + b \in \partial^*(f+g)(x)$. ■

In generale però non vale l'uguaglianza e l'inclusione può essere anche stretta. Per avere l'uguaglianza abbiamo bisogno che almeno una delle funzioni in gioco sia anche continua, in quanto

Teorema 6.4.6. *Date due funzioni convesse proprie* f, g *su* V, *se la terna* (V, W, F) *è duale in un verso e se* f *è continua in un punto* $x_0 \in \operatorname{dom} f \neq \emptyset$, *allora*

1. $\partial^* f(x_0) \neq \emptyset$;
2. $\partial^* f(x_0) + \partial^* g(x_0) = \partial^*(f+g)(x_0)$ *se* $x_0 \in \operatorname{dom} \partial^* g$.

Dimostrazione. Il primo punto è una diretta conseguenza del corollario 6.3.6, sia ora $a \in \partial^*(f+g)(x_0)$, per il teorema di Fenchel-Rockafellar applicato alle funzioni

$$\begin{aligned} f_1(y) &= f(y) - f(x_0) - F(y - x_0, a) \\ f_2(y) &= g(y) - g(x_0) \end{aligned}$$

in quanto f_1 è continua in x_0, dunque esisterà $b \in W$ tale che

$$0 \leq F(x_0, a) - f(x_0) - f^*(a-b) - g^*(b) - g(x_0) \leq 0$$

dove l'ultima disuguaglianza deriva dalla (6.1.1). Sempre dalla (6.1.1) abbiamo che

$$F(x_0, b) = g(x_0) + g^*(b) \text{ e } F(x_0, a-b) = f(x_0) + f^*(a-b)$$

e perciò $b \in \partial^* g(x_0)$ e $a - b \in \partial^* f(x_0)$. ■

Se ci mettiamo sul piano cartesiano il sottodifferenziale di una funzione convessa corrisponde all'insieme di tutti i coefficienti angolari delle rette che toccano il grafico in un punto e che rimangono sotto al grafico. Ciò è chiaramente correlato al concetto di retta tangente al grafico e di conseguenza alla derivata classica. Infatti abbiamo

Teorema 6.4.7. *Se f è convessa su V e continua in $x_0 \in \operatorname{dom} f$ allora per ogni $v \in V$*

$$\lim_{\lambda \to 0^+} \frac{f(x_0 + \lambda v) - f(x_0)}{\lambda} = \max_{a \in \partial^* f(x_0)} F(v, a)$$

Dimostrazione. Per la proposizione 5.2.17 il limite a sinistra coincide con l'estremo inferiore per $\lambda > 0$ che per comodità indicheremo con d. Applicando il teorema di Fenchel-Rockafellar alle funzioni

$$\begin{aligned} f_1(z) &= f(x_0 + z) - f(x_0) \\ f_2(z) &= \begin{cases} -\lambda d & \text{se } \lambda \geq 0, z = \lambda v \\ +\infty & \text{altrimenti} \end{cases} \\ f_1^*(a) &= f^*(a) - F(x_0, a) + f(x_0) \\ f_2^*(a) &= i_{F(v,a) \leq -d}(a) \end{aligned}$$

si ottiene

$$0 = \inf_{\lambda \geq 0} f(x_0 + \lambda v) - f(x_0) - \lambda d = \max_{F(v,a) \geq d} \left[F(x_0, a) - f^*(a) - f(x_0) \right]$$

e quindi esiste $a \in \partial^* f(x_0)$ tale che $F(v, a) \geq d$. D'altronde dalla definizione di sottodifferenziale abbiamo che $F(v, b) \leq d$ per ogni $b \in \partial^* f(x_0)$ da cui discende la tesi. ■

Capitolo 7

Spazi normati e di Banach

7.1 Norme

Definizione 7.1.1. Sia X spazio vettoriale sul campo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, una **norma** su X è un'applicazione $\|\cdot\| : X \to \mathbb{R}$ tale che:

- $\|x\| = 0$ se e solo se $x = 0$;
- $\|tx\| = |t| \|x\|$ per ogni $x \in X$ e $t \in \mathbb{K}$;
- $\|x + y\| \leq \|x\| + \|y\|$ per $x, y \in X$.

Una norma è semplicemente una seminorma che si annulla solamente nell'origine, ma allora $\|x\| \geq 0$ per ogni $x \in X$ e $\left| \|x\| - \|y\| \right| \leq \|x - y\|$.

Una norma genera su X una topologia τ che lo rende uno spazio vettoriale topologico localmente convesso $\mathcal{N}_1$ di Hausdorff per il teorema 5.3.3 e quindi X è anche uno spazio metrico per il teorema 4.2.11. Una metrica che induce la topologia su X è la seguente:

$$d(x, y) = \|x - y\| \tag{7.1.1}$$

essa è chiaramente invariante, bilanciata e convessa e soddisfa la seguente proprietà di omogeneità

$$d(\lambda x, \lambda y) = |\lambda| \, d(x, y)$$

per ogni $x, y \in X$ e $\lambda \in \mathbb{K}$. Da questa proprietà segue infatti che per ogni $x \in X$, $r, s > 0$

$$B(x, r, d) = x + rB_1 = \frac{r}{s} B\left(\frac{s}{r} x, s, d\right)$$

dove $B_1 = \{x \in X \mid \|x\| < 1\}$, e quindi le palle aperte rispetto alla metrica d coincidono con gli intorni convessi della topologia indotta dalla norma.

Segue anche che le palle aperte, e quindi anche quelle chiuse, della metrica d sono limitate e dunque gli insiemi d-limitati di X sono tutti e i soli insiemi limitati (rispetto alla topologia di spazio vettoriale topologico) in X. Da questo momento in poi per ogni spazio normato X l'unica metrica che verrà considerata sarà la metrica d definita in (7.1.1) e non ne verranno considerate altre se non diversamente specificato, perciò indicheremo le palle aperte/chiuse con i simboli $B(x, r)$, $D(x, r)$ senza specificare la metrica scelta.

In uno spazio metrico generico abbiamo detto che la palla chiusa non è in generale la chiusura della palla aperta, ma può essere strettamente più grande di esso. Quando stiamo in uno spazio normato le due nozioni coincidono invece, infatti

Proposizione 7.1.2. *In uno spazio normato le palle chiuse sono le chiusure delle rispettive palle aperte.*

Dimostrazione. Vogliamo applicare la proposizione 3.5.13, scelti $x, y \in X$ spazio normato con $x \neq y$ e $\varepsilon > 0$ basta allora porre

$$\lambda = \frac{\varepsilon}{\varepsilon + \left\| x - y \right\|}$$
$$z = \lambda x + (1 - \lambda) y$$

Infatti così avremo che

$$\left\| z - y \right\| = \lambda \left\| y - x \right\| < \varepsilon$$
$$\| z - x \| = (1 - \lambda) \left\| y - x \right\| < \left\| y - x \right\|$$

e le ipotesi della proposizione 3.5.13 sono soddisfatte, quindi in uno spazio normato le palle chiuse coincidono con la chiusura delle rispettive palle aperte. ■

Proposizione 7.1.3. *Uno spazio vettoriale topologico X possiede una norma che induce la sua topologia se e solo se è di Hausdorff e possiede un intorno convesso limitato dell'origine.*

Dimostrazione. Esisterà allora un aperto U convesso bilanciato limitato contenente l'origine. Per il lemma 4.2.5 la famiglia $\{(1/n)U\}_{n \in \mathbb{N}}$ è un sistema fondamentale di intorni dell'origine.

Definiamo
$$\|x\| = p_U(x)$$
con p_U il funzionale di Minkowski del convesso U, allora p_U è una seminorma in quanto U è bilanciato e l'origine è un punto internale. Dato che X è di Hausdorff per ogni $x \neq 0$ esiste $n \in \mathbb{N}$ tale che $nx \notin U$ e quindi $p_U(x) \geq n > 0$, perciò p_U è una norma su X.

Ora U è un aperto di V e per ogni $r > 0$ si ha

$$rU = \left\{ rx \;\middle|\; p_U(x) < 1 \right\} = \left\{ x \;\middle|\; p_U(x) < r \right\} = B\left(0, r, p_U\right)$$

e quindi le palle aperte di p_U sono aperte in X, viceversa la famiglia $\{rU\}_{r>0}$ è un sistema fondamentale di intorni quindi gli aperti di V ono aperti anche rispetto alla topologia generata dalla norma p_U. ■

Proposizione 7.1.4. *Lo spazio quoziente di uno spazio normato X rispetto ad un suo sottospazio chiuso N è anch'esso uno spazio normato con norma*

$$\|x + N\|_N = \inf\{\|x + z\| \mid z \in N\} = \operatorname{dist}(x, N)$$

Dimostrazione. Si dimostra alla stessa maniera del lemma 4.4.1. ■

Per il momento non tratteremo il prodotto di spazi normati, dato che verranno trattati più nel dettaglio nel prossimo capitolo. Per il momento ci basta sapere che se $(X, \|\cdot\|_1)$ e $(Y, \|\cdot\|_2)$ sono spazi normati su $\mathbb{K}$ allora $X \times Y$ con la topologia prodotto è ancora uno spazio normato con norma
$$\left\|(x, y)\right\| = \|x\|_1 \vee \left\|y\right\|_2$$
Notiamo che questa non è l'unica norma possibile su $X \times Y$ per cui $\left\|(x, 0)\right\| = \|x\|_1$ e $\left\|(0, y)\right\| = \left\|y\right\|_2$

Definizione 7.1.5. Uno spazio normato X è di **Banach** se e solo se X è anche uno spazio vettoriale topologico completo, o equivalentemente la metrica in (7.1.1) è completa in X.

Poiché gli spazi di Banach sono spazi di Fréchet per il lemma 4.4.1 avremo che

Proposizione 7.1.6. *Se X è uno spazio di Banach e $N \leq X$ è un sottospazio chiuso allora anche N è uno spazio di Banach rispetto alla medesima norma.*

Se X e X' sono spazi normati e $f : X \to X'$ lineare allora f è continua e e solo se è limitata, inoltre la definizione di limitatezza di una funzione lineare tra due spazi normati è equivalente ad affermare che esiste $L > 0$ tale che per ogni $x \in X$

$$\left\| f(x) \right\| \leq L \left\| x \right\|$$

Per questo motivo dati due spazi normati X ed Y su $\mathbb{K}$ possiamo dotare anche lo spazio $L(X, Y)$ delle funzioni lineari e continue da X in Y della norma definita come

$$\left\| f \right\| = \sup_{x \neq 0} \frac{\left\| f(x) \right\|}{\left\| x \right\|} = \sup_{\left\| x \right\| \leq 1} \left\| f(x) \right\|$$

Il lettore può verificare con estrema facilità che questa applicazione è effettivamente una norma su $L(X, Y)$.

Proposizione 7.1.7. *Siano X, Y, Z spazi normati e $f \in L(X, Y)$, $g \in L(Y, Z)$. Allora $g \circ f \in L(X, Z)$ e*

$$\left\| g \circ f \right\| \leq \left\| g \right\| \left\| f \right\|$$

Dimostrazione. per ogni $x \in X$

$$\left\| (g \circ f)(x) \right\| = \left\| g\left[f(x) \right] \right\| \leq \left\| g \right\| \left\| f(x) \right\| \leq \left\| g \right\| \left\| f \right\| \left\| x \right\|$$

e per l'arbitrarietà di x segue la tesi. ■

Proposizione 7.1.8. *Se X è uno spazio normato e Y spazio di Banach allora anche $L(X, Y)$ è uno spazio di Banach.*

Dimostrazione. Sia f_n successione di Cauchy su $L(X, Y)$, ovvero per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che per ogni $m, n > N$ e per ogni $x \in X$

$$\left\| f_n(x) - f_m(x) \right\| \leq \epsilon \left\| x \right\|$$

Quindi $f_n(x)$ è una successione di Cauchy per ogni $x \in X$ perciò ammette limite. Poniamo

$$f(x) = \lim_{n \to +\infty} f_n(x)$$

che è ancora lineare, passando ancora al limite in m si ha

$$\left\| f_n(x) - f(x) \right\| \leq \epsilon \left\| x \right\| \Rightarrow \left\| f(x) \right\| \leq \left(\left\| f_n \right\| + \epsilon \right) \left\| x \right\|$$

per ogni $n > N$ ed è quindi limitata. L'unica cosa che bisogna dimostrare che f_n converge a f in norma, difatti per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che per ogni $n > N$ e per ogni $x \in X$

$$\frac{\left\| f_n(x) - f(x) \right\|}{\left\| x \right\|} \leq \epsilon \Leftrightarrow \left\| f_n - f \right\| \leq \epsilon$$

dimostrando così la tesi. ■

Osservazione. Se X è uno spazio normato generico allora X^* è sempre uno spazio di Banach.

Ricordiamo che dato uno spazio normato complesso X il lemma 4.3.7 garantisce che l'applicazione T è biettiva tra X^* e $X^*_{\mathbb{R}}$. Inoltre la norma $\|\cdot\|$ definita su X è in realtà una norma anche su $X_{\mathbb{R}}$ in quanto hanno la medesima topologia, dimostriamo ora che per ogni $f \in X^*$ allora $\|Tf\| = \|f\|$.

Abbiamo innanzitutto che

$$\left|f(x)\right| = \sqrt{[\mathrm{Re} f(x)]^2 + [\mathrm{Im} f(x)]^2} \geq \left|\mathrm{Re} f(x)\right|$$

dunque Tf è limitata e $\|Tf\| \leq \|f\|$. Se f non è identicamente nulla per la proposizione 1.5.3 per ogni $x \neq 0$ abbiamo

$$\|Tf\| \geq \frac{\left|f(x)\right|}{\|x\|}$$

e dunque $\|f\| \leq \|Tf\|$. Abbiamo appena dimostrato che

Proposizione 7.1.9. *L'applicazione $T : X^* \to X^*_{\mathbb{R}}$ definita precedentemente è biettiva e soddisfa le seguenti proprietà:*

- $T(a+b) = T(a) + T(b)$ *per ogni* $a, b \in X^*$*;*
- $T(\lambda a) = \lambda T(a)$ *per ogni* $\lambda \in \mathbb{R}$.

Non possiamo dire che T sia lineare in quanto X^* è di solito uno spazio vettoriale su $\mathbb{C}$ mentre $X^*_{\mathbb{R}}$ è uno spazio vettoriale su $\mathbb{R}$.

Quando X è uno spazio reale osserviamo che la terna (X, X^*, F), dove $F : X \times X^* \to \mathbb{R}$ è il funzionale bilineare definito come

$$F(x, f) = f(x)$$

è una terna duale in un senso secondo la definizione 6.1.6, e inoltre la norma standard di X^* è la polare della norma standard di X. Perciò possiamo applicare il teorema 6.2.8 in modo tale che

$$\|x\| = \sup_{\|f\| \leq 1} \left|f(x)\right|$$

In realtà questo risultato continua a valere anche se X è uno spazio normato complesso grazie alla proposizione 7.1.9. Inoltre utilizzando il teorema di Hahn-Banach o il teorema di Fenchel-Rockafellar l'estremo superiore sopra viene raggiunto da un funzionale di norma unitaria. Per il momento ricordiamo che se dotiamo X^* della norma duale la terna (X, X^*, F) non sarà duale nel verso opposto, in quanto potrebbero esistere dei funzionali lineari continui su X^* che non si possono rappresentare tramite gli elementi di X.

Teorema 7.1.10. *Sia X spazio normato e $N \leq X$ chiuso e diverso da X. Allora*

$$X \text{ è uno spazio di Banach} \Leftrightarrow N \text{ e } X/N \text{ sono spazi di Banach}$$

Dimostrazione. Sia X uno spazio di Banach, allora anche N è uno spazio di Banach e dobbiamo perciò dimostrare la completezza solo di X/N. Sia $x_n + N$ successione di Cauchy quindi per ogni $k \in \mathbb{N}$ esiste $N_k \in \mathbb{N}$ **crescente** tale che per ogni $m, n > N_k$

$$\|x_n - x_m + N\| < \frac{1}{2^k}$$

esisterà allora un $z_k \in N$ tale che

$$\left\| x_{N_{k+1}} - x_{N_k} + z_k \right\| < \frac{1}{2^k}$$

Definiamo adesso $y_k = x_{N_{k+1}} - x_{N_1} + \sum_{i=1}^k z_i$ essa rimane ancora di Cauchy e pertanto ha un'estratta convergente a y. Infine si ha

$$\left\| x_{N_k} - x_{N_1} - y + N \right\| \leq \left\| x_{N_k} - x_{N_1} - y + \sum_{i=1}^k z_i \right\| = \| y_k - y \|$$

e quindi $x_{N_k} + N$ converge ad $y + x_{N_1} + N$.

Supponiamo ora che N e X/N sono spazi di Banach e consideriamo una successione di Cauchy $x_n \in E$, allora anche $x_n + N$ è di Cauchy in X/N e per ipotesi esiste $x \in X$ tale che $x_n + N \to x + N$. Esisterà inoltre per ogni $k \in \mathbb{N}$ una successione crescente $N_k \in \mathbb{N}$ tale che

$$\left\| x_{N_k} - x + N \right\| < \frac{1}{2^k}$$

e quindi esisteranno $y_k \in N$ tali che $\left\| x_{N_k} - x - y_k \right\| < 1/2^k$.

Poiché x_{N_k} è una successione di Cauchy anche y_k lo è quindi esisterà $y \in N$ tale che $y_k \to y$. Mettendo insieme tutti i pezzi avremo che

$$\left\| x_{N_k} - x - y \right\| \leq \left\| x_{N_k} - x - y_k \right\| + \| y_k - y \| \to 0$$

e il teorema è così dimostrato. ■

Prima di andare avanti generalizziamo un risultato ben noto di algebra lineare, ovvero che ogni isometria su $\mathbb{R}^n$ che fissa l'origine è anche lineare. Quando passiamo da $\mathbb{R}^n$ agli spazi normati dobbiamo però supporre anche che l'isometria sia suriettiva, in quanto potrebbero esistere isometrie non affini. Per rimuovere l'ipotesi di suriettività dobbiamo supporre che le norme sugli spazi normati in gioco siano strettamente convesse, come vedremo alla fine del capitolo.

Innanzitutto abbiamo bisogno del seguente lemma preparatorio:

Proposizione 7.1.11. *Dati due spazi vettoriali topologici V, W reali e un'applicazione $f : V \to W$ continua tale che*

$$f\left(\frac{a+b}{2}\right) = \frac{f(a) + f(b)}{2}$$

allora f è affine, ovvero esistono $y \in W$ e $L \in L(V, W)$ tali che $f(v) = y + L(v)$.

Dimostrazione. Poniamo $L(v) = f(v) - f(0)$ e $y = f(0)$, chiaramente $L \in C(V, W)$ dato che siamo tra spazi vettoriali topologici. Una volta che abbiamo dimostrato che $L(\alpha v) = \alpha L(v)$ per ogni $\alpha \in \mathbb{K}, v \in V$ avremo che

$$L(a+b) = 2L\left(\frac{a+b}{2}\right) = f(a) + f(b) - 2y = L(a) + L(b)$$

e quindi L è lineare.

Posto $b = 0$ e $a = 2v$ abbiamo che $L(2v) = 2L(v)$, supponiamo ora che $L(nv) = nL(v)$ per un certo intero $n \geq 2$ e per ogni $v \in V$ allora

$$L[(n+1)v] = L\left[\frac{n(2v) + 2v}{2}\right] = \frac{L[n(2v)] + L(2v)}{2} = \frac{2nL(v) + 2L(v)}{2} = (n+1)L(v)$$

e perciò $L(rv) = rL(v)$ per ogni $v \in V$ e per ogni $r \in \mathbb{Q}$. Poiché f è continua e poiché V e W sono spazi vettoriali topologici passando al limite avremo che l'uguaglianza vale anche per $r \in \mathbb{R}$. ■

Teorema 7.1.12 (Mazur-Ulam). *Data una isometria biettiva $f : X \to Y$ tra spazi normati reali, allora f è affine.*

Dimostrazione. Qui seguiremo la dimostrazione in [15]. Per ogni $z \in V$ poniamo $R_z x = 2z - x$, chiaramente R_z è una isometria di X in sé con $R_z^{-1} = R_z$. Adesso scegliamo $x, y \in X$ qualunque e definiamo I l'insieme di tutte le isometrie biettive $g : X \to X$ tali che $g(a) = a$ e $g(b) = b$.

Poniamo $z = (x + y)/2$ e

$$\zeta = \sup_{g \in I} \left\| g(z) - z \right\|$$

osserviamo che ζ è finito in quanto g è una isometria e dunque

$$\left\| g(z) - z \right\| \leq \left\| g(z) - x \right\| + \| x - z \| = \left\| g(z) - g(x) \right\| + \| a - z \| = 2 \| x - z \|$$

Costruiamo ora una applicazione $T : I \to I$ in modo tale che

$$T(g) = R_z \circ g^{-1} \circ R_z \circ g$$

che è ben formata in quanto R_z scambia x con y, e quindi

$$\zeta \geq \left\| (Tg)(z) - z \right\| = \left\| g^{-1} \left[R_z(g(z)) \right] - z \right\| = \left\| R_z[g(z)] - g(z) \right\| = 2 \left\| z - g(z) \right\|$$

Per come abbiamo definito ζ questo è possibile se e solo se $\zeta = 0$ in quanto non può essere $+\infty$, ma allora $g(z) = z$ per ogni $g \in I$.

Torniamo alla nostra f e poniamo $w = [f(x) + f(y)]/2 \in Y$, ma allora $f' = R_z \circ f^{-1} \circ R_w \circ f \in I$ e perciò

$$f(z) = R_w[f(z)]$$

e questo è possibile se e solo se $f(z) = w$. Applicando così la proposizione precedente si ha la tesi. ■

Osservazione. Questo risultato non vale per spazi normati complessi, basta prendere $X = \mathbb{C}$ e

$$f(z) = z^{\dagger}$$

che risulta essere una isometria $\mathbb{R}$-lineare ma non $\mathbb{C}$-lineare.

7.2 Il teorema di Hahn-Banach in spazi normati

Abbiamo già enunciato il teorema di Hahn-Banach nel capitolo degli spazi vettoriali topologici. In questa sezione lo applicheremo più in particolare agli spazi normati.

Teorema 7.2.1. *Sia X spazio normato su $\mathbb{K}$ e $Y \leq X$ dotato della stessa norma di X, allora per ogni $f \in Y^*$ esiste un funzionale $F \in X^*$ tale che $\|F\| = \left\|f\right\|$ e per ogni $y \in Y$ si ha $f(y) = F(y)$*

Dimostrazione. Invece di mostrare la dimostrazione classica che usa il teorema di Hahn-Banach negli spazi vettoriali topologici in questa sede utilizzeremo invece il teorema di Fenchel-Rockafellar applicato alle funzioni

$$F_1(x) = i_{\|x\|\leq 1}$$
$$F_2(x) = \begin{cases} f(x) & \text{se } x \in Y \\ +\infty & \text{se } x \notin Y \end{cases}$$

per la continuità della norma infatti f_1 è continua nel suo dominio e $f_1(0) = f_2(0) = 0$. Dato che

$$F_1^*(g) = \|g\|$$
$$F_2^*(g) = \begin{cases} 0 & \text{se } g = f \text{ su } Y \\ +\infty & \text{altrimenti} \end{cases}$$

applicando il teorema avremo così che

$$-\|f\| = -\min_{g=f \text{ su } Y} \|g\| = -\|F\|$$

come volevasi dimostrare. Nonostante questa dimostrazione così come scritta vale solamente per gli spazi vettoriali reali la proposizione 7.1.9 la rende valida anche se X fosse uno spazio normato complesso. ■

Corollario 7.2.2. *Sia X uno spazio normato su $\mathbb{K}$ allora per ogni $x \in X \setminus \{0\}$ esiste $f \in X^*$ con $\|f\| = 1$ tale che*

$$\|x\| = |f(x)| = \max_{g\neq 0} \frac{|g(x)|}{\|g\|}$$

Dimostrazione. Consideriamo il funzionale

$$g : tx \in \operatorname{span}\{x\} \to t\|x\| \in \mathbb{K}$$

chiaramente g è lineare e $\|g\| = 1$, quindi da Hahn-Banach esiste $f \in X^*$ che estende g su tutto X e $\|f\| = 1$. In particolare avremo che $f(x) = g(x) = \|x\|$ poiché $x \in \operatorname{span}\{x\}$. ■

Proposizione 7.2.3. *Se X è uno spazio normato e $Y < X$ sottospazio chiuso diverso da X e fissiamo $x \in X \setminus Y$. Allora esiste $f \in X^*$ di norma unitaria che vale 0 su Y e $f(x) = \operatorname{dist}(x, Y)$.*

Dimostrazione. Applichiamo il corollario precedente allo spazio X/Y rispetto al punto $x + Y \neq Y$ avremo così un funzionale $F : X/Y \to \mathbb{K}$ lineare continuo di norma unitaria tale che $F(x+Y) = \|x+Y\| = \operatorname{dist}(x, Y)$.

Definiamo allora

$$f : z \in X \to F(z+Y) \in \mathbb{K}$$

allora f è ancora lineare in X e si annulla su Y. Ora per ogni $z \notin Y$ abbiamo chiaramente $|f(z)|/\|z\| \leq |f(z)|/d(z,Y)$, inoltre per ogni $z \notin Y$ e $\epsilon > 0$ esiste $y \in Y$ tale che $1/d(z,Y) \leq 1/\|z-y\| + \epsilon$ e quindi

$$\frac{|f(z)|}{\operatorname{dist}(z,Y)} \leq \frac{|f(z-y)|}{\|z-y\|} + \epsilon \leq \sup_{u\neq 0} \frac{|f(u)|}{\|u\|} + \epsilon$$

quindi $\|f\| = \|F\| = 1$, in particolare $f \in X^*$. Per concludere basta osservare che $f(x) = F(x+Y) = \operatorname{dist}(x, Y)$ e la dimostrazione è così conclusa. ■

Osservazione. Se X è uno spazio normato generico non possiamo assolutamente dire se, fissato $f \in X^*$ la quantità

$$\frac{|f(x)|}{\|x\|}$$

ammette massimo su $X \setminus \{0\}$. Anzi per molti spazi normati l'estremo superiore di tale quantità non viene mai raggiunto da alcun elemento di X.

Grazie al teorema di Hahn-Banach possiamo invertire il risultato della proposizione 7.1.8 sfruttando il fatto che se $X \neq \{0\}$ allora $X^* \neq \{0\}$ per il corollario 7.2.2.

Proposizione 7.2.4. *Siano X ed Y spazi normati con $X \neq \{0\}$. Allora $L(X,Y)$ è uno spazio di Banach se e solo se Y è uno spazio di Banach.*

Dimostrazione. Sia $y_n \in Y$ una successione di Cauchy e supponiamo che $L(X,Y)$ sia uno spazio di Banach, esisteranno dunque $f \in X^* \setminus \{0\}$ ed $z \in X$ tali che $f(z) = 1$. Definiamo allora la seguente applicazione

$$F_n(x) = y_n f(x) \in L(X,Y)$$

chiaramente $\|F_m - F_n\| \leq \|y_m - y_n\| \|f\|$ e quindi F_n è una successione di Cauchy nello spazio $L(X,Y)$ che dovrà necessariamente convergere ad un certo $F \in L(X,Y)$. Dunque

$$y_n = F_n(z) \to F(z) = y \in Y$$

e quindi anche Y è uno spazio di Banach. ■

Lemma 7.2.5. *Sia X spazio normato, se X^* è separabile allora anche X è separabile.*

Dimostrazione. Possiamo supporre che X sia uno spazio reale, la dimostrazione sui complessi è la stessa. Dal lemma 3.5.24 esisterà una successione $f_n \in X^*$ che risulti densa sull'insieme $\left\{f \in X^* \mid \|f\| = 1\right\}$. Esisteranno allora $x_n \in X$ di norma unitaria con $f_n(x_n) \geq \frac{1}{2}$, l'insieme $U = \{\sum_{i=1}^m q_i x_i \mid m \in \mathbb{N} \wedge q_i \in \mathbb{Q}\}$ è numerabile e inoltre

$$\mathrm{Cl}(U) = \mathrm{Cl}\left(\mathrm{span}\{x_n \mid n \in \mathbb{N}\}\right)$$

Difatti per ogni $x \in \mathrm{Cl}\left(\mathrm{span}\{x_n \mid n \in \mathbb{N}\}\right)$ e per ogni $\epsilon > 0$ esistono delle costanti $m \in \mathbb{N}$ ed $r_1, \dots, r_m \in \mathbb{R}$ tali che $\|x - \sum_{i=1}^m r_i x_i\| < \epsilon$. Per la densità dei numeri razionali possiamo trovare $q_1, \dots, q_m \in \mathbb{Q}$ tali che $|r_i - q_i| < \epsilon / m$ e quindi

$$\left\|x - \sum_{i=1}^m q_i x_i\right\| \leq \left\|x - \sum_{i=1}^m r_i x_i\right\| + \sum_{i=1}^m |r_i - q_i| < 2\epsilon$$

e quindi $x \in \mathrm{Cl}(U)$.

Se per assurdo non fosse denso in X allora per la proposizione 7.2.3 esisterà un $y \in X \setminus \mathrm{Cl}\left(\mathrm{span}\{x_n \mid n \in \mathbb{N}\}\right)$ e un $f \in X^*$ con $\|f\| = 1$ che vale 0 sugli x_n e $f(y) \neq 0$. Ma allora

$$\frac{1}{2} \leq |f(x_n)| + |f_n(x_n) - f(x_n)| \leq \|f_n - f\|$$

raggiungendo un assurdo poiché $\{f_n \mid n \in \mathbb{N}\}$ è denso. ■

Fissiamo adesso un elemento $x \in X$ e definiamo a partire da esso una nuova applicazione definita come

$$\hat{x} : f \in X^* \to f(x) \in \mathbb{K}$$

È banale constatare che $\hat{x}$ è lineare e continua, mentre dal corollario 7.2.2 segue immediatamente che $\|\hat{x}\| = \|x\|$. L'applicazione

$$J_X : x \in X \to \hat{x} \in X^{**} \tag{7.2.1}$$

è un'**isometria** ($\left\|J_X(x)\right\| = \|x\|$) lineare e quindi iniettiva.

Teorema 7.2.6. *Sia X spazio normato, allora esistono uno spazio di Banach B e un'isometria lineare f da X in B tale che $f(X)$ è denso in B.*

Dimostrazione. Innanzitutto $J_X(X)$ è un sottospazio di X^{**} isomorfo a X, poniamo $B = \mathrm{Cl}\left(J_X(X)\right)$ (la chiusura viene effettuata su tutto X^{**}) e $f = J_X$. Dato che X^{**} è uno spazio di Banach e B è un suo sottospazio chiuso allora per la proposizione 7.1.6 anche B è uno spazio di Banach. ■

Il risultato continua a valere anche se X è solamente uno spazio metrico, ovviamente la funzione f non sarà più lineare.

Teorema 7.2.7. *Se X è uno spazio metrico allora per ogni $x \in X$ esistono uno spazio di Banach B_x e una isometria $f : X \to B_x$ tali che $f(x) = 0$ e $f(X)$ è denso in B_x.*

Dimostrazione [6]. Consideriamo lo spazio $\mathrm{Lip}(X, Y)$ delle funzioni $f : X \to Y$ lipschitziane tra lo spazio metrico X e lo spazio normato Y, allora $\mathrm{Lip}(X, Y)$ è uno spazio vettoriale con somma $(f + g)(x) = f(x) + g(x)$, prodotto esterno $(\alpha f)(x) = \alpha f(x)$ ed elemento neutro la funzione identicamente nulla.

Per ogni $f \in \mathrm{Lip}(X, Y)$ definiamo

$$\|f\|_{\mathrm{Lip}} = \sup_{a \neq b} \frac{\|f(a) - f(b)\|}{d(a, b)}$$

essa non sarà una norma su $\mathrm{Lip}(X, Y)$ in quanto si annulla sulle funzioni costanti. Se però fissiamo $x \in X$ allora sarà una norma sullo spazio

$$C_x = \left\{f \in \mathrm{Lip}(X, Y) \mid f(x) = 0\right\}$$

Poniamo adesso $Y = \mathbb{R}$ e consideriamo lo spazio duale C_x^*, chiaramente per ogni $y \in X$ possiamo definire $[y] \in C_x^*$ in modo tale che $[y](f) = f(y) \in \mathbb{R}$ per ogni $f \in C_x$. Esso sarà sicuramente un funzionale lineare continuo in quanto $\left|[y](f)\right| = \left|f(y) - f(x)\right| \leq d(y, x)\left\|f\right\|_{\mathrm{Lip}}$ e quindi $\left\|[y]\right\|_{C_x^*} \leq d(y, x)$, in realtà abbiamo l'uguaglianza prendendo $f(y) = d(y, x)$. Più in generale per ogni $y, z \in X$ si avrà

$$\left\|[y] - [z]\right\|_{C_x^*} = d(y, z)$$

e quindi l'applicazione $F : y \in X \to [y] \in C_x^*$ è una isometria.

Basterà allora prendere

$$B_x = \mathrm{Cl}\left(\mathrm{span}\, f(X)\right) \leq C_x^*$$

e B_x sarà allora un sottospazio chiuso di uno spazio di Banach, e perciò è uno spazio di Banach. ■

7.3 Norme su spazi vettoriali finitamente generati

Due norme $\|\cdot\|_1$, $\|\cdot\|_2$ su uno stesso spazio vettoriale X si dicono **equivalenti** se e solo se inducono la medesima topologia su X, ovvero se e solo se l'applicazione seguente

$$x \in (X, \|\cdot\|_1) \to x \in (X, \|\cdot\|_2) \tag{7.3.1}$$

è un omeomorfismo tra spazi topologici.

Per quanto detto precedentemente ciò equivale ad affermare che esistano due costanti $c, C > 0$ tali che

$$c\,\|x\|_1 \leq \|x\|_2 \leq C\,\|x\|_1$$

per ogni $x \in X$, in altre parole se e solo se l'applicazione (7.3.1) è bilipschitziana. Più in generale ogni omeomorfimo lineare tra spazi normati è bilipschitziana, cosa ovviamente non più vera se togliamo la condizione di linearità.

Una conseguenza del teorema 4.6.1 permette di stabilire che tutte le norme sono equivalenti:

Teorema 7.3.1. *Se V è uno spazio vettoriale di dimensione finita allora tutte le norme su V sono equivalenti.*

Dimostrazione. Le norme inducono su V una struttura di spazio vettoriale topologico di Hausdorff, dato che su V vi è un unica struttura di spazio vettoriale topologico di Hausdorff l'applicazione (7.3.1) è sempre un omeomorfismo dunque due qualsiasi norme sono equivalenti. ■

Quindi per gli spazi vettoriali di dimensione finita esiste un'unica topologia indotta da una norma, quindi ogniqualvolta non viene specificata la topologia di uno spazio vettoriale di dimensione finita si intende sempre quella generata da una norma.

Corollario 7.3.2. *Gli spazi vettoriali di dimensione finita sono completi.*

Abbiamo già dimostrato che solamente negli spazi vettoriali topologici finitamente generati la chiusura della palla unitaria è compatta. Dimostreremo di nuovo questo risultato negli spazi normati per la compattezza sequenziale, la quale è equivalente alla compattezza in quanto gli spazi normati sono anche metrici.

Lemma 7.3.3 (dei quasi ortogonali di Riesz)**.** *Sia X spazio normato e $M < X$ sottospazio chiuso. Allora per ogni $\epsilon > 0$ esiste $x \in X$ con $\|x\| = 1$ e*

$$d(x, M) \geq 1 - \epsilon$$

Dimostrazione. Sia $x_1 \in X \setminus M$ generico e poniamo $R = d(x_1, M)$, dalla proposizione 3.5.12 segue che $R > 0$, quindi per ogni $0 < \epsilon < 1$ esiste $y_1 \in M$ tale che

$$\|x_1 - y_1\| \leq \frac{R}{1-\epsilon}$$

Poniamo ora

$$x = \frac{x_1 - y_1}{\|x_1 - y_1\|} \in X \setminus M$$

ha chiaramente norma unitaria e per ogni $y \in M$

$$\|x-y\| = \frac{\|(x_1 - y_1 - \|x_1 - y_1\| y)\|}{\|x_1 - y_1\|} \geq 1-\epsilon$$

ottenendo la tesi passando all'estremo inferiore. ■

Teorema 7.3.4 (Riesz)**.** *Sia X spazio normato, se $D(0,1)$ è relativamente sequenzialmente compatto allora X ha dimensione finita.*

Dimostrazione. Per assurdo X ha dimensione infinita, sia $x_1 \in X$ con $\|x_1\| = 1$, costruiamo per ricorrenza grazie al lemma precedente una successione $x_n \in X$ con $\|x_n\| = 1$ e

$$d(x_n, \operatorname{span}\{x_1, x_2, \dots, x_{n-1}\}) \geq \frac{1}{2}$$

da cui segue che $d(x_n, x_m) \geq \frac{1}{2}$ per ogni m, n e quindi la successione non può avere estratte di Cauchy. Raggiungiamo così un assurdo poiché $\{x_n\} \subseteq D(0,1)$ è relativamente sequenzialmente compatto. ■

Questo teorema mostra che le norme che rendono le palle chiuse compatte sono relativamente rare in quanto definibili esclusivamente sugli spazi vettoriali di dimensione finita. Questa osservazione giustifica il fatto che, nel momento in cui definiamo la norma di un operatore tramite la formulazione equivalente

$$\|f\| = \sup_{\|x\|=1} \|f(x)\|$$

non possiamo in alcun modo sostituire l'estremo superiore con un massimo senza condizioni aggiuntive sullo spazio X.

7.4 Topologie deboli

Prima di passare oltre mostriamo alcune conseguenze del principio di uniforme limitatezza negli spazi di Banach.

Teorema 7.4.1 (Banach-Steinhaus)**.** *Siano X spazio di Banach, Y spazio normato e $\{\Lambda_i\}_{i\in\mathbb{N}} \subseteq L(X,Y)$ tali che per ogni $x \in X$*

$$\sup_{i\in\mathbb{N}} \|\Lambda_i x\| < +\infty$$

Allora

$$\sup_{i\in\mathbb{N}} \|\Lambda_i\| < +\infty$$

Dimostrazione. Per il teorema 4.4.3, essendo X uno spazio di Banach e quindi un F-spazio, la famiglia $\{\Lambda_i\}_{i\in\mathbb{N}}$ è equicontinua e quindi esisterà $M > 0$ tale che $\Lambda_i(B(0,1)) \subseteq D(0,M)$ o equivalentemente

$$\|\Lambda_i\| = \sup_{x<1} \|\Lambda_i x\| \leq T$$

passando all'estremo superiore sugli i si ha la tesi. ■

Corollario 7.4.2. *Siano X spazio di Banach, Y spazio normato e $\{\Lambda_i\}_{i\in\mathbb{N}} \subseteq L(X,Y)$ tali che per ogni $x \in X$ $\Lambda_n x$ converge in Y ad una certa quantità Λx. Allora Λ è un operatore lineare continuo e vale*

$$\|\Lambda\| \leq \liminf_{n\to+\infty} \left\|\Lambda_n\right\|$$

Dimostrazione. L'operatore Λ è chiaramente lineare in x, per quanto riguarda la continuità osserviamo che per ogni $x \in X$

$$\|\Lambda x\| = \lim_{n\to+\infty} \left\|\Lambda_n x\right\| \leq \left(\liminf_{n\to+\infty} \left\|\Lambda_n\right\|\right) \|x\|$$

Il teorema di Banach-Steinhaus garantisce che la quantità $\liminf_{n\to+\infty} \left\|\Lambda_n\right\|$ sia finita e dunque Λ è continua e vale la disuguaglianza della tesi. ∎

Dato uno spazio normato X su $\mathbb{K}$ con norma $\|\cdot\|$ abbiamo indicato con X^* lo spazio delle funzioni lineari continue da X in $\mathbb{K}$. Se prendiamo ora un sottospazio non vuoto F di X^* possiamo dotare X della topologia iniziale generata da tutti i funzionali di F, come d'altronde abbiamo già fatto a pagina 137.

Adattando opportunamente il teorema 6.1.8 e la proposizione 7.1.9 notiamo che la topologia duale rende X uno spazio vettoriale (reale o complesso) topologico localmente convesso che sarà meno fine della topologia usuale di X, detta anche **topologia forte**. Da questo momento chiameremo tale topologia la **topologia debole di X generata da F** e sarà indicata con il simbolo $\sigma(X,F)$. Gli aperti e i chiusi della topologia $\sigma(X,F)$ saranno detti rispettivamente insiemi F-**debolmente aperti** e insiemi F-**debolmente chiusi**.

Poiché le topologie deboli sono meno fini della topologia forte avremo che:

1. Gli insiemi F-debolmente aperti sono aperti;
2. Gli insiemi F-debolmente chiusi sono chiusi;
3. La chiusura forte di un insieme è contenuta nella sua chiusura F-debole;
4. L'interno forte di un insieme contiene l'interno F-debole;
5. Se una successione x_n converge ad x nella topologia forte allora convergerà ad x anche nella topologia debole generata da F;
6. Ogni successione di Cauchy nella topologia forte è di Cauchy nella topologia debole generata da F;
7. Gli insiemi compatti sono F-debolmente compatti;
8. Gli insiemi sequenzialmente compatti sono F-debolmente sequenzialmente compatti;
9. Gli insiemi limitati sono F-debolmente limitati.

Inoltre grazie alla proposizione 2.7.4 avremo che $x_n \to x$ F-debolmente se e solo se $f(x_n) \to f(x)$ per ogni $f \in F$. In altre parole le topologie deboli sono topologie "meno fini e più grossolane" rispetto alla topologia forte, per questo motivo sono chiamate topologie deboli.

Sappiamo già che F è totale se e solo se per ogni $x \in X \setminus \{0\}$ esiste $f \in F$ tale che $f(x) \neq 0$, in questo caso X con la topologia F-debole è anche uno spazio di Hausdorff e dunque possiamo applicare i risultati ottenuti nel precedente capitolo sull'analisi convessa. Notiamo però che la norma $\|\cdot\|$ potrebbe non essere continua una volta che dotiamo X della topologia F-debole, quindi il teorema 6.2.8 dice solamente che la bipolare di $\|\cdot\|$ rispetto ad F

$$\|x \llcorner F\| = \sup_{f\in F\setminus\{0\}} \frac{|f(x)|}{\|f\|}$$

è una norma F-debolmente semicontinua inferiormente e più piccola della norma usuale. Diremo perciò che lo spazio F è **totalizzante** se e solo se è totale e $\|x \llcorner F\| = \|x\|$ per ogni $x \in X$.

Proposizione 7.4.3. *Dato uno spazio normato X e un sottospazio $F \leq X^*$ allora*

1. *Se F è chiuso in X^* gli insiemi F-debolmente limitati sono limitati;*
2. *Se x_n converge F-debolmente ad x allora*
$$\|x \llcorner F\| \leq \liminf_{n\to+\infty} \|x_n \llcorner F\|$$
3. *Se x_n converge a x F-debolmente con $\|x_n\|$ limitato e $f_n \to f$ fortemente in F allora $f_n(x_n) \to f(x)$.*

Dimostrazione. Dimostriamo i vari punti:

1. Un sottoinsieme A di X è limitato se e solo se tutte le successioni contenute in esso sono limitate in norma. Consideriamo una generica successione $x_n \in A$ allora essa è F-debolmente limitata e quindi per ogni $f \in F$ esiste $M_f > 0$ tale che per ogni $n \in \mathbb{N}$
$$|\hat{x}_n(f)| = |f(x_n)| \leq M_f$$
Ora essendo F chiuso e X^* uno spazio di Banach segue che anche F è uno spazio di Banach e dunque possiamo applicare il teorema di Banach-Steinhaus alla famiglia di funzionali $\hat{x}_n \in X^{**}$ da cui discende che x_n è limitato in norma.
2. Banale.
3. È immediata in quanto $\|x_n\|$ è limitato e
$$|f_n(x_n) - f(x)| \leq \|f_n - f\| \, \|x_n\| + |f(x_n) - f(x)|$$

∎

Lemma 7.4.4. *Siano X spazio normato e $F \leq X^*$ sottospazio su X. Allora:*

- *se F è totalizzante e X è separabile allora esiste una sottofamiglia numerabile e totalizzante $\mathscr{F} \subseteq F$ di funzioni tale che $\|f\| = 1$ per ogni $f \in \mathscr{F}$;*
- *se F è separabile e totale su X allora esiste una sottofamiglia numerabile e totale $\mathscr{F} \subseteq F$ di funzioni tale che $\|f\| = 1$ per ogni $f \in \mathscr{F}$.*

In entrambi i casi la topologia $\sigma(X, F)$ indotta su un qualunque sottoinsieme F-debolmente compatto di X è metrizzabile.

Dimostrazione. Supponiamo inizialmente che X sia separabile ed F totalizzante. Possiamo allora prendere una successione $x_n \in X \setminus \{0\}$ densa in X ed un'altra successione $0 < \varepsilon_n < 1/2$ convergente a 0, per la proprietà totalizzante esisteranno $f_n \in F$ tali che $\|f_n\| = 1$ e $|f_n(x_n)| > (1-\varepsilon_n)\|x_n\|$, dimostriamo che la famiglia di seminorme $\mathscr{F} = \left\{ |f_n| \;\middle|\; n \in \mathbb{N} \right\}$ è totalizzante su X. Presi $x \in X \setminus \{0\}$ ed $0 < \varepsilon < 1/3$ generici esisterà x_n tale che $\|x - x_n\| \leq \varepsilon \|x\|$ e di conseguenza $\|x_n\| \geq (1-\varepsilon)\|x\|$, allora

$$|f_n(x)| \geq |f_n(x_n)| - \|f_n\| \|x - x_n\| > \left[(1-\varepsilon_n)(1-\varepsilon) - \varepsilon \right] \|x\|$$

per l'arbitrarietà di ε e di n segue così che $\mathscr{F}$ è totalizzante.

Supponiamo ora che F sia separabile e totale e sia $f_n \in F$ un sottoinsieme denso. Per ogni $x \in X \setminus \{0\}$ esistono $f \in F$ e $\varepsilon > 0$ tale che $|f(x)| > \varepsilon \|x\|$, per densità esiste f_n tale che $\|f - f_n\| < \varepsilon/2$. Dunque avremo che

$$|f_n(x)| \geq |f(x)| - \|f_n - f\| \|x\| > \frac{\varepsilon}{2} \|x\|$$

e quindi anche $\mathscr{F} = \left\{ |f_n| \;\middle|\; n \in \mathbb{N} \right\}$ è una famiglia totale di seminorme.

La famiglia $\mathscr{F}$ induce su X una topologia τ che lo rende uno spazio vettoriale topologico localmente convesso $\mathscr{N}_1$ e di Hausdorff, per il teorema 4.2.11 la topologia τ è indotta da una metrica invariante. Prendiamo W sottoinsieme F-debolmente compatto di X e lo dotiamo della topologia indotta da $\sigma(X,F)$, che indicheremo per comodità con τ', quindi (W,τ') è uno spazio topologico compatto e (W,τ) è uno spazio metrico, in particolare è uno spazio di Haudorff.

Poiché $\mathscr{F}$ è ottenuto da un sottoinsieme di F avremo che $\tau \subseteq \tau'$ e quindi l'applicazione

$$x \in (W,\tau') \to x \in (W,\tau)$$

è continua. Grazie al lemma 2.4.7 questa applicazione è un omeomorfismo e quindi anche la topologia τ', ovvero la topologia su W indotta da $\sigma(X,F)$, è generata da una metrica. ■

Osservazione. Il lemma di sopra non implica che la topologia F-debole di tutto X sia metrizzabile, anzi in generale le topologie deboli non sono quasi mai metrizzabili.

Due notevoli topologie deboli di uno spazio normato che risultano molto importanti in varie applicazioni si ottengono per particolari scelte della famiglia $\mathscr{F}$:

- Se $F = X^*$ allora la topologia $\sigma(X,X^*)$ sarà detta semplicemente **topologia debole** di X;

- Se esiste uno spazio normato Y tale che $X = Y^*$ allora la topologia generata dalla famiglia
$$F = \left\{ \hat{y} \in Y^{**} = X^* \;\middle|\; y \in Y \right\} = J_Y(Y)$$
è detta **topologia debole*** generata da Y su Y^* e utilizzeremo il simbolo $\sigma(Y^*,Y)$ per indicarla invece del più corretto $\sigma(Y^*, J_Y(Y))$.

Inoltre dati due spazi normati X ed Y diremo che $f : X \to Y$ è **debolmente continua** se e solo se è continua con X ed Y dotati delle rispettive topologie deboli, analogamente definiamo le funzioni debolmente* continue. Si noti però che funzioni continue possono non essere debolmente continue e funzioni debolmente continue possono non essere continue. Utilizzeremo la notazione $x_n \rightharpoonup x$ per dire che x_n converge a x nella topologia debole, e $x_n \overset{*}{\rightharpoonup} x$ per dire che vi converge con la topologia debole*.

In generale il sottospazio F che genera la topologia debole è sempre totalizzante e chiuso per il corollario 7.2.2, mentre quello che genera la topologia debole* è totalizzante ma è chiusa se e solo se X è uno spazio di Banach, quindi la proposizione 7.4.3 si applica alla topologia debole* di X^* solo quando X è uno spazio di Banach.

Proposizione 7.4.5. *Se* $\dim X < +\infty$ *allora tutte le topologie deboli generate da sottospazi totali coincidono con la topologia forte.*

Dimostrazione. Fissiamo una base $e_1, \dots, e_n$ di X, poiché tutte le norme su X sono equivalenti tra loro possiamo considerare la seguente norma su X

$$\left\| x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \right\|_\infty = \left| x_1 \right| \vee \left| x_2 \right| \vee \cdots \vee \left| x_n \right|$$

dove $x_i \in \mathbb{K}$. Se per ogni $1 \leq i \leq n$ consideriamo il funzionale lineare $e_i^* \in X^*$ definito in modo tale che $e_i^*(e_j) = \delta_{ij}$ (si ricordi che $\delta_{ij} = 1$ se e solo se $i = j$ altrimenti vale 0) allora gli e_i^* formano una base per X^* che avrà dunque dimensione finita. Preso un qualunque sottospazio totale $Z \leq X^*$ se per assurdo fosse $Z \neq X^*$ allora esisterebbe $f = \sum_{i=1}^n f_i e_i^* \in X^*$ tale che per ogni $g = \sum_{i=1}^n g_i e_i^* \in Z$ si avrebbe

$$f \cdot g = \sum_{i=1}^n f_i g_i = 0$$

Prendiamo allora $x = \sum_{i=1}^n f_i e_i \in X$ per ipotesi deve esistere $g \in Z$ tale che $g(x) \neq 0$ ma

$$g(x) = \sum_{i=1}^n \sum_{j=1}^n g_i f_j e_i^*(e_j) = \sum_{i=1}^n g_i f_i = g \cdot f = 0$$

assurdo, quindi $Z = X^*$ e in particolare l'unica topologia F debole di Hausdorff su X è quella debole.

Inoltre si avrà anche $e_i^* \in Z$, questo implica che

$$B(x,r) = \left\{ y \in X \;\middle|\; \left\| y - x \right\|_\infty < 1 \right\} = \left\{ y \in X \;\middle|\; e_i^*(y - x) < 1 \text{ per ogni } 1 \leq i \leq n \right\}$$

e quindi gli insiemi aperti sono anche debolmente aperti. ■

Proposizione 7.4.6. *Siano* X, Y *spazi normati e* $T : X \to Y$ *un'applicazione lineare. Se* T *è anche continua allora è debolmente continua.*

Dimostrazione. Per la proposizione 2.7.3 abbiamo che $T : X \to Y$ è debolmente continua se e solo se per ogni $f \in Y^*$ l'applicazione $f \circ T : X \to \mathbb{K}$ è continua rispetto alla topologia debole di X. Se T è continua allora $f \circ T \in X^*$, da cui segue immediatamente la tesi in quanto la topologia debole su X è la più piccola topologia rispetto alla quale tutti i funzionali in X^* risultano continui. ■

Dalla definizione di successione di Cauchy in uno spazio vettoriale topologico, dalla proposizione 5.3.9 e dalla linearità degli elementi dell'insieme $Z \leq X^*$ sappiamo che una successione x_n è F-debolmente di Cauchy se e solo se per ogni $f \in Z$ la successione $f(x_n) \in \mathbb{K}$

è di Cauchy (e quindi converge), e lo spazio X è F-debolmente completo se e solo se per ogni successione di Cauchy x_n esiste $x \in X$ tale che $f(x_n) \to f(x)$ per ogni $f \in Z$.

Nonostante l'intuito porti a pensare che tutti gli spazi siano debolmente completi in realtà questa proprietà non è sempre verificata, nemmeno quando X è uno spazio di Banach.

Dal corollario 7.4.2 segue un utile criterio di completezza per la topologia debole*

Proposizione 7.4.7. *Se X è di Banach allora X^* è debolmente* completo.*

Non è possibile utilizzare lo stesso approccio per dimostrare un analogo risultato per la debole completezza, in quanto il limite Λ del corollario 7.4.2 potrebbe appartenere a X^{**} e non a X anche se X fosse uno spazio di Banach. Per ottenere la completezza debole dobbiamo supporre per X condizioni più stringenti dell'essere uno spazio di Banach come vedremo più avanti.

Il seguente risultato invece discende direttamente dal lemma 6.1.7

Proposizione 7.4.8. *Se K è un sottoinsieme convesso di uno spazio normato X allora*

$$K \text{ debolmente chiuso} \Leftrightarrow K \text{ fortemente chiuso}$$

Teorema 7.4.9 (Terzo teorema di Mazur)**.** *Sia ora X spazio di Banach con $x_n \rightharpoonup x$ in X, allora esiste una applicazione strettamente crescente $N : \mathbb{N} \to \mathbb{N}$ tale che per ogni $n \in \mathbb{N}$ esistono $c_1^n, c_2^n, \dots, c_{N(n)}^n \in [0,1]$ tali che*

$$\sum_{i=1}^{N(n)} c_i^n = 1$$
$$\lim_{n\to+\infty} \left\| \sum_{i=1}^{N(n)} c_i^n x_i - x \right\| = 0$$

Dimostrazione. Consideriamo l'insieme $A = \operatorname{con}\left(\{x_n \mid n \in \mathbb{N}\}\right)$, chiaramente A è convesso ed, essendo X uno spazio vettoriale topologico, anche la chiusura di A sarà convessa e dunque $\mathrm{Cl}(A)$ è debolmente chiuso. Poiché $x_n \in A$ allora $x_n \rightharpoonup x \Rightarrow x \in \mathrm{Cl}(A) \Rightarrow$ esiste una successione in A che tende fortemente a x. ■

Osservazione. Questi due risultati valgono solamente per la topologia debole e non per la topologia debole*, potrebbero infatti esistere sottoinsiemi chiusi convessi che però non sono debolmente* chiusi.

Il teorema 6.1.8 afferma che la terna $(X, X^*, \mathscr{F})$ è duale in entrambi i sensi quando X è dotato della topologia debole e X^* della topologia debole*, dal lemma 6.1.7 segue così che

Proposizione 7.4.10. *Dato uno spazio normato X per ogni applicazione lineare debolmente* continua $F : X^* \to \mathbb{K}$ esiste un $z \in X$ tale che $F(f) = f(z)$.*

Lemma 7.4.11. *Sia X spazio normato, allora l'applicazione $J_X : X \to J_X(X)$ è un omeomorfismo se X è dotato della topologia debole e $J_X(X)$ è dotato della topologia indotta dalla topologia debole* su X^{**}.*

Dimostrazione. Dimostriamo prima di tutto che $J_X : X \to X^{**}$ è continua, per la proposizione 2.7.3 avremo che J_X è continua se e solo se $J_{X^*}(f) \circ J_X$ è una funzione debolmente continua da X in $\mathbb{K}$ per ogni $f \in X^*$. Ma per ogni $x \in X$ abbiamo che

$$J_{X^*}(f)[J_X(x)] = J_X(x)(f) = f(x)$$

ovvero $J_{X^*}(f) \circ J_X = f \in X^*$ che è chiaramente una funzione debolmente continua, dunque J_X è continua.

Mostriamo ora che $J_X^{-1} : J_X(X) \to X$ è continua, sempre per la proposizione 2.7.3 equivale a dimostrare che $f \circ J_X^{-1} : J_X(X) \to \mathbb{K}$ è continua con la topologia debole* indotta su $J_X(X)$ per ogni $f \in X^*$. Ma per ogni $x \in X$ avremo che

$$f[J_X^{-1}(\hat{x})] = f(x) = \hat{x}(f) = J_{X^*}(f)(\hat{x})$$

quindi $f \circ J_X^{-1}$ è la restrizione della funzione $J_{X^*}(f) \in J_{X^*}(X^*)$ che come sappiamo è una funzione debolmente* continua da X^{**} su $\mathbb{K}$ quindi anche le sue restrizioni saranno continue e J_X^{-1} sarà anch'esso continuo. ∎

Dimostriamo ora il teorema di Banach-Alaoglu il quale afferma che le palle chiuse sono debolmente* compatte. Per snellire la notazione utilizzeremo il simbolo $B(x, r, X)$ al posto di $B(x, r, \|\cdot\|)$ per indicare la palla aperta di centro x e raggio r nello spazio normato X con norma $\|\cdot\|$, analogamente faremo con la palla chiusa.

Teorema 7.4.12 (Banach-Alaoglu)**.** *Sia X spazio normato su* $\mathbb{K}$*, allora l'insieme*

$$D(0, 1, X^*) = \left\{f \in X^* \;\middle|\; \|f\| \leq 1\right\}$$

è debolmente compatto.*

Dimostrazione. Nel primo capitolo abbiamo dotato l'insieme $F(X, \mathbb{K})$ della topologia prodotto definita come la topologia iniziale indotta dalle proiezioni

$$\pi_x : f \in F(X, \mathbb{K}) \to f(x) \in \mathbb{K}$$

al variare di $x \in X$. Chiaramente avremo che $X^* \subseteq F(X, \mathbb{K})$ e la topologia debole* di X^* non è nient'altro che la topologia prodotto di $F(X, \mathbb{K})$ indotta su X^* in quanto $\pi_x(f) = \hat{x}(f)$ per ogni $f \in X^*$. Quindi se dimostriamo che $D(0, 1, X^*)$ è un sottoinsieme compatto di $F(X, \mathbb{K})$ allora sarà chiaramente compatto anche in X^* rispetto alla topologia debole*.

Ricordiamo che X^+ denota l'insieme di tutti i funzionali lineari su X in $\mathbb{K}$, dimostriamo che è chiuso in $F(X, \mathbb{K})$. Infatti per ogni $x, y \in X$ e $\lambda \in \mathbb{K}$ le applicazioni

$$A_{x,y} : f \in F(X, \mathbb{K}) \to \pi_{x+y}(f) - \pi_x(f) - \pi_y(f) \in \mathbb{K}$$
$$B_{x,\lambda} : f \in F(X, \mathbb{K}) \to \pi_{\lambda x}(f) - \lambda \pi_x(f) \in \mathbb{K}$$

sono continue in quanto composizione di funzioni continue e dunque

$$\bigcap_{x,y \in X} A_{x,y}^{-1}(\{0\}) \cap \bigcap_{\substack{x \in X \\ \lambda \in \mathbb{R}}} B_{x,\lambda}^{-1}(\{0\}) = X^+$$

è un sottoinsieme chiuso chiuso. Ora

$$f \in D(0, 1, X^*) \Leftrightarrow f \text{ lineare e } |f(x)| \leq \|x\| \;\; \forall x \in X \Leftrightarrow f \in X^+ \cap \prod_{x \in X} D(0, \|x\|, \mathbb{K})$$

Ma $D(0, r, \mathbb{K})$ sono le palle chiuse in $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ e quindi sono compatte, dal teorema di Tychonoff anche il loro prodotto sarà compatto nella topologia di $F(X, Y)$. Poiché X^+ è chiuso e $D(0, 1, X^*)$ è intersezione di un chiuso e di un compatto segue che sarà anch'esso compatto. ∎

Osservazione. Dal teorema di Banach-Alaoglu segue che tutte le palle sono debolmente* compatte in X^* e non solo quella unitaria centrata nell'origine. Infatti lo spazio X^* con la topologia debole* è pur sempre uno spazio vettoriale topologico e quindi le traslazioni e i riscalamenti per un fattore non nullo sono omeomorfismi e quindi mandano compatti in compatti.

In generale però $D(0,1,X^*)$ non è debolmente* sequenzialmente compatto ed esistono spazi normati che possiedono successioni limitate in norma che non hanno estratte convergenti debolmente*.

7.5 Spazi riflessivi

L'applicazione (7.2.1) è una isometria lineare da X in X^{**}. In generale non è biiettiva, ovvero esistono funzionali in X^{**} che non possono essere rappresentati da elementi di X. Per questo motivo introduciamo una nuova classe di spazi per cui l'applicazione J_X definita precedentemente risulti biiettiva.

Definizione 7.5.1. Uno spazio normato è **riflessivo** se e solo se l'applicazione $J_X : X \to X^{**}$ è biettiva.

Questi spazi possiedono molte proprietà interessanti, per esempio abbiamo

Proposizione 7.5.2. *Se X è riflessivo allora le topologie debole e debole* su X^* coincidono.*

Proposizione 7.5.3. *Poniamo X spazio riflessivo, allora per ogni $f \in X^*$ esiste $x \in X$ tale che $\|x\| = 1$ e*

$$\|f\| = |f(x)| = \max_{x \neq 0} \frac{|f(x)|}{\|x\|}$$

Dimostrazione. Dal corollario 7.2.2 ed essendo X riflessivo esiste $x \in X$ tale che

$$\|f\| = |\hat{x}(f)| = |f(x)|$$

in quanto tutti gli elementi di X^{**} possono essere determinati da X. ■

Proposizione 7.5.4. *Gli spazi riflessivi sono spazi di Banach.*

Dimostrazione. Se x_n è una successione di Cauchy in X allora lo è anche $\hat{x}_n$ in X^{**}. Ora X^{**} è il duale di X^* e quindi è uno spazio di Banach, esisterà allora $F \in X^{**}$ tale che $x_n \to F$ in X^{**}, ma per la riflessività di X esisterà $x \in X$ tale che $F = \hat{x}$ e $x_n \to x$ in X. ■

Negli spazi riflessivi oltre alla completezza forte vale anche quella debole, difatti

Proposizione 7.5.5. *Gli spazi riflessivi sono debolmente completi.*

Dimostrazione. Prendiamo x_n successione debolmente di Cauchy nello spazio riflessivo X, ovvero per ogni $f \in X^*$ la successione $f(x_n)$ ammette limite. possiamo applicare il corollario 7.4.2 alla sequenza di funzionali $\hat{x}_n \in X^{**}$ dunque esiste $F \in X^{**}$ tale che

$$F(f) = \lim_{n\to+\infty} \hat{x}_n(f) = \lim_{n\to+\infty} f(x_n)$$

La tesi segue immediatamente dalla riflessività di X. ■

Teorema 7.5.6. *Sia X spazio di Banach, allora X è riflessivo se e solo se X^* è riflessivo.*

Dimostrazione. Supponiamo innanzitutto che X sia riflessivo e prendiamo $F \in X^{***}$ che sarà un'applicazione lineare continua da X^{**} in $\mathbb{K}$. Posto

$$l : x \in X \to F\left[J_X(x)\right] \in \mathbb{K}$$

allora $l \in X^*$ e per ogni $f \in X^{**}$ avremo

$$F(f) = l\left[J_X^{-1}(f)\right] = f(l)$$

ovvero $F = J_{X^*}(l)$.

Supponiamo adesso che X^* sia uno spazio riflessivo, poniamo $M = J_X(X)$ che sarà un sottospazio di X^{**} isomorfo ad X e dimostriamo che è chiuso. Sia $y_n \in M$ ed $y \in X^{**}$ tale che $y_n \to y$ e quindi esisteranno $x_n \in X$ tali che $y_n = \hat{x}_n$. Poiché y_n è una successione di Cauchy, essendo J_X isometria anche x_n è una successione di Cauchy e avendo supposto che X sia uno spazio di Banach avremo dunque che $x_n \to x \in X$. Perciò sempre per continuità avremo che $y = J_X(x)$ e quindi M è chiuso.

Se per assurdo $M \neq X^{**}$ dalla proposizione 7.2.3 esiste $F \in X^{***}$ non nullo tale che $F(M) = 0$, per ipotesi avremo un $f \in X^*$ tale che $F = J_{X^*}(f)$ e quindi per ogni $x \in X$

$$f(x) = J_X(x)(f) = F\left[J_X(x)\right] = 0$$

ma allora $f = 0$ il che implicherebbe che $F = 0$ assurdo, perciò $M = X^{**}$ e X è riflessivo. ■

Teorema 7.5.7. *Sia X spazio di Banach, allora X è riflessivo e separabile se e solo se X^* è riflessivo e separabile.*

Dimostrazione. L'implicazione da destra a sinistra è una semplice applicazione del teorema precedente e del lemma 7.2.5.

Viceversa supponiamo che X sia separabile e riflessivo, allora anche X^{**} è separabile e per il lemma 7.2.5 X^* è separabile. Sempre dal teorema 7.5.6 segue che X^* è riflessivo. ■

Abbiamo già visto che i sottoinsiemi chiusi degli spazi di Banach sono anch'essi spazi di Banach, vale un risultato analogo anche per gli spazi riflessivi come afferma il seguente teorema.

Teorema 7.5.8. *Se X è riflessivo allora ogni suo sottospazio chiuso è riflessivo.*

Dimostrazione. Sia $Y \leq X$ un sottospazio chiuso e $F \in Y^{**}$ un'applicazione lineare da Y^* in $\mathbb{K}$. Possiamo considerare allora l'applicazione $f \in X^* \to f\big|_Y \in Y^*$ che ad ogni f associa la sua restrizione su Y. Questa applicazione è ben definita, lineare e continua e quindi esiste un unico $G \in X^{**}$ tale che per ogni $f \in X^*$

$$G(f) = F\left(f\big|_Y\right) = f(x) \tag{7.5.1}$$

con $x \in X$ per riflessività. Se $x \in Y$ allora per ogni $g \in Y^*$ possiamo applicare la (7.5.1) ad una sua qualunque estensione $\tilde{g}$ su tutto X mediante il teorema di Hahn-Banach per ottenere $F(g) = \tilde{g}(x) = g(x)$.

Supponiamo allora che $x \notin Y$ dunque per la proposizione 7.2.3 esiste $f \in X^*$ che si annulla su M ma non in x, il che è assurdo in quanto $f(x) = F\left(f\big|_Y\right) = 0$. Quindi per ogni $F \in Y^{**}$ esiste $y \in Y$ tale che $F(f) = f(y)$ e Y è riflessivo. ■

Abbiamo dimostrato precedentemente che uno spazio normato ha dimensione finita se e solo se la palla chiusa unitaria è compatta rispetto alla topologia forte. Con il teorema di Banach-Alaoglu abbiamo mostrato che le palle chiuse sono debolmente* compatte, ora dimostreremo un risultato analogo per la topologia debole in spazi riflessivi.

Il teorema di Kakutani afferma inoltre non solo che la palla chiusa è debolmente compatta in uno spazio riflessivo, ma anche che se la palla chiusa è debolmente compatta in uno spazio di Banach allora lo spazio deve essere riflessivo.

Dimostreremo le due implicazioni del teorema separatamente. Mentre la prima implicazione è una semplice applicazione del teorema di Banach-Alaoglu l'implicazione opposta è più difficile e richiede prima dei lemmi accessori.

Teorema 7.5.9 (Kakutani - parte 1)**.** *Se X è uno spazio riflessivo allora la palla chiusa unitaria è debolmente compatta.*

Dimostrazione. Per il lemma 7.4.11 J_X^{-1} manda insiemi debolmente* compatti in X^{**} in insiemi debolmente compatti in X. Possiamo quindi applicare il teorema di Banach-Alaoglu e dedurre così che l'insieme $D(0,1,X) = J_X^{-1}(D(0,1,X^{**}))$ è debolmente compatto. ∎

Lemma 7.5.10 (Helly)**.** *Sia X spazio di Banach, $f_1, f_2, \dots, f_k \in X^*$ ed $\alpha_1, \dots, \alpha_k \in \mathbb{K}$. Allora sono equivalenti*

1. *Per ogni $\epsilon > 0$ esiste $x \in D(0,1,X)$ tale che*
$$\left|f_i(x) - \alpha_i\right| < \epsilon$$
per ogni i;
2. *Per ogni $\beta_1, \dots, \beta_k \in \mathbb{K}$*
$$\left|\sum_{i=1}^k \alpha_i\beta_i\right| \leq \left\|\sum_{i=1}^k \beta_i f_i\right\|$$

Dimostrazione. Supponiamo che X soddisfi la 1 e prendiamo $\beta_1, \dots, \beta_k \in \mathbb{K}$, allora per ogni ϵ esiste $x \in D(0,1,X)$ tale che
$$\left|\sum_{i=1}^k \beta_i f_i(x) - \sum_{i=1}^k \beta_i\alpha_i\right| \leq \sum_{i=1}^k |\beta_i|\left|f_i(x) - \alpha_i\right| \leq \epsilon\sum_{i=1}^k |\beta_i|$$
e quindi
$$\left|\sum_{i=1}^k \beta_i\alpha_i\right| \leq \left|\sum_{i=1}^k \beta_i f_i(x)\right| + \epsilon\sum_{i=1}^k |\beta_i| \leq \left\|\sum_{i=1}^k \beta_i f_i\right\| + \epsilon\sum_{i=1}^k |\beta_i|$$
e il membro a destra non dipende da x.

Se invece vale la 2 per comodità poniamo $\vec{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{K}^k$ e
$$\|\vec{\alpha}\| = \max_i |\alpha_i|$$
$$\varphi : x \in X \to [f_1(x), \dots, f_k(x)] \in \mathbb{K}^k$$
per assurdo $\operatorname{dist}[\vec{\alpha}, \varphi[D(0,1,X)]] > 0$ e quindi $\vec{\alpha} \notin \operatorname{Cl}(\varphi[D(0,1,X)])$, dal secondo teorema di Mazur esiste $\vec{\beta} \in \mathbb{K}^k$ ed $a \in \mathbb{R}$ tali che per ogni $x \in D(0,1,X)$
$$\operatorname{Re}\left[\vec{\beta}\cdot\varphi(x)\right] < a < \operatorname{Re}(\vec{\beta}\cdot\vec{\alpha})$$

ovvero per ogni $x \in D(0,1,X)$

$$\sum_{i=1}^{k} \mathrm{Re}[\beta_i f_i(x)] < a < \sum_{i=1}^{k} \mathrm{Re}(\beta_i \alpha_i)$$

Il lemma 7.1.9 afferma che per ogni $g \in X^*$ si ha $\|\mathrm{Re}\, g\| = \|g\|$, quindi

$$\left\| \sum_{i=1}^{k} \beta_i f_i \right\| < a < \mathrm{Re}\left(\sum_{i=1}^{k} \beta_i \alpha_i \right)$$

assurdo. ■

Lemma 7.5.11 (Goldstine)**.** *Se X è uno spazio di Banach allora $J_X[D(0,1,X)]$ è denso in $D(0,1,X^{**})$ rispetto alla topologia debole* di X^{**}.*

Dimostrazione. Preso un qualunque aperto non vuoto V di X^{**} che interseca la palla chiusa $D(0,1,X^{**})$ vogliamo dimostrare che interseca anche $J_X[D(0,1,X)]$. Sia $F \in V \cap D(0,1,X^{**})$ senza perdere in generalità possiamo supporre che V sia uno degli intorni fondamentali centrati in F che genera la topologia debole* in X^{**}, ovvero

$$\begin{aligned} V &= \left\{ G \in X^{**} \;\middle|\; \left| J_{X^*}(f_i)(G - F) \right| < \epsilon, i = 1,2,\dots,k \right\} \\ &= \left\{ G \in X^{**} \;\middle|\; \left| G(f_i) - F(f_i) \right| < \epsilon \right\} \end{aligned}$$

per opportuni $f_1, \dots, f_k \in X^*$.

Posto $\alpha_i = F(f_i)$ avremo per ogni $\beta_i \in \mathbb{K}$

$$\left| \sum_{i=1}^{k} \beta_i \alpha_i \right| = \left| F\left(\sum_{i=1}^{k} \beta_i f_i \right) \right| \le \left\| \sum_{i=1}^{k} \beta_i f_i \right\|$$

e per il lemma di Helly per ogni $\epsilon > 0$ esiste $x_\epsilon \in D(0,1,X)$ tale che per ogni i

$$\left| f_i(x_\epsilon) - F(f_i) \right| < \epsilon$$

e quindi $J_X(x_\epsilon) \in V$. ■

Teorema 7.5.12 (Kakutani - parte 2)**.** *Se X è uno spazio di Banach e se $D(0,1,X)$ è debolmente compatta allora X è riflessivo.*

Dimostrazione. Per il lemma 7.4.11 $J_X : X \to X^{**}$ è continua quando X è dotato della topologia debole e X^{**} della topologia debole*. Ma allora $J_X[D(0,1,X)]$ è debolmente* compatto in $D(0,1,X^{**})$ ed essendo la topologia debole* di Hausdorff sarà anche chiuso in $D(0,1,X^{**})$.

Dal lemma di Goldstine tale insieme è anche denso dunque

$$J_X[D(0,1,X)] = D(0,1,X^{**})$$

per linearità di J_X lo spazio X è riflessivo. ■

Grazie alla compattezza debole della palla unitaria possiamo enunciare la seguente versione del teorema di Weierstrass in spazi riflessivi, molto utilizzata in varie branche dell'analisi.

Teorema 7.5.13. *Sia X spazio riflessivo e $K \subseteq X$ chiuso convesso e limitato. Posto $F : K \to \mathbb{R}$ quasiconvessa e semicontinua inferiormente allora F ha un punto di minimo in K. Se F è strettamente quasiconvessa allora il punto di minimo è unico.*

Dimostrazione. Per dimostrare l'esistenza del minimo bisogna solo dimostrare che F rimane semicontinua inferiormente quando sostituiamo la topologia forte di X con la topologia debole. In tal caso infatti per la proposizione 7.4.8 l'insieme K è debolmente chiuso e contenuto in un insieme debolmente compatto per la proposizione 7.5.9 e quindi K sarà debolmente compatto. L'esistenza del minimo deriverà allora dal teorema 2.6.6.

Sappiamo già che F è debolmente semicontinua inferiormente se e solo se i suoi sottolivelli $P_\alpha = \{x \in X \mid F(x) \leq \alpha\}$ sono debolmente chiusi, per ipotesi sappiamo solamente che sono fortemente chiusi. Per la proposizione 5.2.9 i sottolivelli P_α sono convessi e perciò sono anche debolmente chiusi e dunque F è debolmente semicontinua inferiormente.

Ricordiamo che F è strettamente quasiconvessa se e solo se per ogni $x \neq y$ e $\lambda \in (0,1)$ vale

$$F[\lambda x + (1-\lambda)y] < F(x) \vee F(y)$$

Per assurdo esistono due punti distinti $x, y \in K$ con $m = F(x) = F(y)$ il minimo di F, allora per ogni $\lambda \in (0,1)$

$$m \leq F[\lambda x + (1-\lambda)y] < m$$

assurdo. ∎

Il teorema di Banach-Alaoglu e il teorema di Kakutani forniscono dei criteri sufficienti di compattezza debole e debole* dei sottoinsiemi limitati, il primo vale per X normato mentre il secondo richiede che X sia riflessivo. Questi risultati, quando X oltre ad essere normato è anche uno spazio di Banach, forniscono anche una condizione necessaria alla compattezza in quanto limitatezza in norma e limitatezza debole coincidono per la proposizione 7.4.3. Più nello specifico abbiamo

Teorema 7.5.14 (Criterio di compattezza debole e debole*)**.** *Se X è uno spazio di Banach allora un sottoinsieme $K \subseteq X^*$ è debolmente* compatto se e solo se è debolmente* chiuso e limitato in norma.*

Se invece X è riflessivo allora $K \subseteq X$ è debolmente compatto se e solo se è debolmente chiuso e limitato.

Per quanto riguarda la compattezza sequenziale dimostreremo più avanti che nel caso della topologia debole essa è equivalente alla compattezza, mentre abbiamo già accennato che nella topologia debole* i compatti potrebbero non essere sequenzialmente compatti. Se però X è uno spazio normato separabile allora il lemma 7.4.4 insieme al teorema di Hausdorff e al teorema di Banach-Alaoglu forniscono una condizione sufficiente per la compattezza sequenziale nella topologia debole*

Teorema 7.5.15 (Criterio di compattezza debole* sequenziale)**.** *Sia X spazio normato separabile, allora ogni successione $f_n \in X^*$ limitata in norma possiede un'estratta debolmente* convergente.*

7.6 Il teorema di Eberlein–Šmulian

Il teorema di Kakutani fornisce una condizione necessaria e sufficiente per la compattezza debole. In questa sezione dimostreremo che è anche una condizione necessaria e sufficiente per la compattezza debole sequenziale, in altre parole in uno spazio riflessivo ogni successione limitata possiede un'estratta debolmente* convergente.

Per comodità di notazione dato un sottoinsieme A di X indicheremo con $\mathrm{Cl}(A,F)$ la chiusura di A nella topologia $\sigma(X,F)$.

Teorema 7.6.1 (Eberlein–Šmulian)**.** *Se X è uno spazio normato e $K \subseteq X$ non vuoto, allora sono equivalenti*

1. $\mathrm{Cl}(K,X^*)$ *è debolmente compatto in X;*
2. *ogni successione in K ammette un'estratta debolmente convergente in X.*

Inoltre se K soddisfa almeno una delle seguenti affermazioni allora per ogni $x \in \mathrm{Cl}(K,X^*)$ esiste una successione $x_k \in K$ tale che $x_k \rightharpoonup x$.

L'ultimo punto non è banale in quanto la topologia debole non è quasi mai $\mathcal{N}_1$ e quindi non è possibile caratterizzare la chiusura tramite successioni.

Dimostriamo innanzitutto che il primo punto implica il secondo. Sia $K \subseteq X$ con chiusura debolmente compatta e sia $x_n \in K$ successione generica, definiamo

$$K_0 = \mathrm{Cl}(\mathrm{span}\{x_n\}, X^*)$$

allora K_0 è un sottoinsieme debolmente chiuso separabile e la topologia debole $\sigma(K_0, K_0^*)$ per Hahn-Banach coincide con la topologia debole di X quindi $W_0 = K_0 \cap \mathrm{Cl}(K,X^*)$ è debolmente compatto in K_0. Per il lemma 7.4.4 e per il teorema di Hausdorff W_0 è sequenzialmente debolmente compatto e quindi $x_n \in W_0$ possiede un'estratta debolmente convergente in $W_0 \subseteq X$.

Supponiamo ora che tutte le successioni in K hanno un'estratta debolmente convergente, dimostriamo che ha chiusura compatta. Dimostriamo innanzitutto il seguente risultato

Proposizione 7.6.2. *Se X è uno spazio normato e $F \leq X^*$ ha dimensione finita allora esistono $x_1, x_2, \dots, x_k \in B(0,1,X)$ tali che*

$$\max_{1\leq i\leq k} \left|f(x_i)\right| \geq \frac{\|f\|}{2}$$

per ogni $f \in F$.

Dimostrazione. Dato che F ha dimensione finita l'insieme $\left\{f \in F \mid \|f\| = 1\right\}$ è compatto e quindi per il teorema di Hausdorff è totalmente limitato, dunque esisteranno $f_1, f_2, \dots, f_k \in F$ di norma unitaria tali che per ogni $f \in F$ di norma unitaria esiste f_i tale che

$$\|f - f_i\| < \frac{1}{4}$$

Esistono allora $x_i \in B(0,1,X)$ tali che $\left|f_i(x_i)\right| \geq 3/4$ dunque per ogni $f \in F$ esisterà $1 \leq i \leq k$ tale che

$$\left|f(x_i)\right| \geq \|f\| \left|f_i(x_i)\right| - \|f\| \left\| \frac{f}{\|f\|} - f_i \right\| \geq \left(\frac{3}{4} - \frac{1}{4}\right) \|f\|$$

e quindi si ha la tesi. ■

Lemma 7.6.3. *Sia $K \subseteq X$ tale che ogni successione in esso possiede un'estratta debolmente convergente in X, allora per ogni $\zeta \in \mathrm{Cl}(J_X(K), X^*) \subseteq X^{**}$ esiste $x \in \mathrm{Cl}(K, X^*)$ tale che $\zeta = J_X(x)$ ed esiste una successione $x_n \in K$ convergente debolmente a x.*

Dimostrazione. Sappiamo che ζ appartiene alla chiusura debole* di $J_X(K)$ se e solo se ogni insieme debolmente* aperto in X^{**} contenente ζ deve intersecare $J_X(K)$. Quindi se $f_1 \in B(0,1,X^*)$ allora l'insieme $\left\{F \in X^{**} \mid \left|F(f_1) - \zeta(f_1)\right| < \varepsilon\right\}$ è per definizione un insieme debolmente* aperto contenente ζ, dunque esisterà $x_1 \in K$ tale che

$$\left|f_1(x_1) - \zeta(f_1)\right| < 1$$

Poniamo allora $W_1 = \mathrm{span}\{\zeta, \hat{x}_1\} \leq X^{**}$ per la proposizione precedente esisteranno delle funzioni $f_2, f_3, \dots, f_{l_2} \in B(0,1,X^*)$ tale che per ogni $\xi \in W_1$

$$\max_{l_1 < i \leq l_2} \left|\xi(f_i)\right| \geq \frac{\|\xi\|}{2}$$

dove $l_1 = 1$.

Ancora ripetendo il ragionamento di sopra esisterà $x_2 \in K$ tale che per ogni $1 \leq i \leq l_2$

$$\left|f_i(x_2) - \zeta(f_i)\right| < \frac{1}{2}$$

e quindi posto $W_2 = \mathrm{span}\{\zeta, \hat{x}_1, \hat{x}_2\}$ sempre per la proposizione precedente esisteranno altre funzioni $f_{l_2+1}, \dots, f_{l_3} \in B(0,1,X^*)$ tale che per ogni $\xi \in W_2$

$$\max_{l_2 < i \leq l_3} \left|\xi(f_i)\right| \geq \frac{\|\xi\|}{2}$$

Per induzione allora per ogni $k \in \mathbb{N}$ esisteranno $x_k \in K, f_k \in B(0,1,X^*)$ tali che

$$\left|f_i(x_k) - \zeta(f_i)\right| < \frac{1}{k} \text{ per ogni } 1 \leq i \leq l_k$$
$$\max_{l_k < i \leq l_{k+1}} \left|\xi(f_i)\right| \geq \frac{\|\xi\|}{2} \text{ per ogni } \xi \in W_k$$

con $W_k = \mathrm{span}\{\zeta, \hat{x}_1, \hat{x}_2, \dots, \hat{x}_k\}$.

Ora $W = \bigcup_{k=1}^{+\infty} W_k$ è un sottospazio vettoriale di X^{**}, chiaramente per ogni $\xi \in W$ si avrà

$$\sup_{i \in \mathbb{N}\setminus\{1\}} \left|\xi(f_i)\right| \geq \frac{\|\xi\|}{2} \tag{7.6.1}$$

dimostriamo che vale la (7.6.1) anche se $\xi \in \mathrm{Cl}(W)$ (chiusura nella topologia forte). Se per assurdo la (7.6.1) non fosse vera allora esisterebbero $\xi_0 \in \mathrm{Cl}(W)$ e $t > 0$ tali che per ogni $i > 1$

$$\left|\xi_0(f_i)\right| \leq t < \frac{\left\|\xi_0\right\|}{2}$$

Poiché stiamo lavorando con la chiusura rispetto ad una norma esisterà $\xi \in W$ tale che $3\left\|\xi - \xi_0\right\| < \left\|\xi_0\right\| - 2t$, per ogni $i > 1$ avremo allora

$$\left|\xi(f_i)\right| \leq \left|\xi_0(f_i)\right| + \left\|\xi - \xi_0\right\| \leq t + \left\|\xi - \xi_0\right\| < \frac{\left\|\xi_0\right\|}{3} + \frac{t}{3}$$

Ora stimiamo meglio la norma di ξ_0:

$$\begin{aligned}\left\|\xi_0\right\| - \|\xi\| < \frac{\left\|\xi_0\right\|}{3} - \frac{2}{3}t \Rightarrow \left\|\xi_0\right\| &< \|\xi\| + \frac{\left\|\xi_0\right\|}{3} - \frac{2}{3}t \\ \Rightarrow \left\|\xi_0\right\| &< \frac{3}{2}\|\xi\| - t\end{aligned}$$

da cui segue che $\left|\xi(f_i)\right| < \|\xi\|/2$ per ogni $i > 1$. Questo però è impossibile in quanto $\xi \in W_k$ per qualche k dunque la (7.6.1) è verificata anche su $\mathrm{Cl}(W)$.

Per ipotesi su K esisterà un'estratta x_{n_k} convergente debolmente a un qualche x appartenente a $\mathrm{Cl}\big(K, X^*\big)$, per il terzo teorema di Mazur x sarà il limite di una composizione convessa delle x_{x_k} dunque $\hat{x} \in \mathrm{Cl}(W)$. Ancora essendo $x_{n_k} \rightharpoonup x$ allora per ogni $k \in \mathbb{N}$ esisterà $k' \in \mathbb{N}$ tale che $n_{k'} > k$ e

$$\left|f_i(x) - f_i(x_{n_{k'}})\right| < \frac{1}{k} \text{ per ogni } 1 \le i \le l_k$$

e dunque

$$\left|f_i(x) - \zeta(f_i)\right| \le \left|f_i(x) - f_i(x_{n_{k'}})\right| + \left|f_i(x_{n_{k'}}) - \zeta(f_i)\right| < \frac{2}{k}$$

per ogni $1 \le i \le l_k$ in quanto $l_k < l_{n_{k'}}$. Ma il termine a sinistra non dipende da k' quindi facendo tendere k a infinito avremo che

$$f_i(x) = \xi(f_i)$$

per ogni $i \in \mathbb{N}$. Ma $\hat{x} - \xi \in \mathrm{Cl}(W)$ in quanto la chiusura di sottospazi vettoriali è ancora un sottospazio vettoriale e dunque

$$0 = \sup_{i>1}(\hat{x} - \xi)(f_i) \ge \frac{\|\hat{x} - \xi\|}{2}$$

e dunque $\xi = \hat{x} = J_X(x)$ e inoltre $x_{n_k} \rightharpoonup x$. ■

Dimostriamo allora il primo punto a partire dal secondo. Per il lemma appena dimostrato avremo che $\mathrm{Cl}\big(J_X(K), X^*\big) \subseteq J_X(X)$ e per il lemma 7.4.11 avremo che

$$J_X\left[\mathrm{Cl}\big(K, X^*\big)\right] = \mathrm{Cl}\big(J_X(K), X^*\big)$$

Sappiamo anche che K è limitato quindi anche $J_X(K)$ è limitato, per il teorema di Banach-Alaoglu ed essendo la chiusura di compatti ancora compatta $\mathrm{Cl}\big(J_X(K), X^*\big)$ è un insieme debolmente* compatto, poiché J_X è un omeomorfismo allora $\mathrm{Cl}\big(K, X^*\big)$ è debolmente compatto.

Sia adesso $x \in \mathrm{Cl}\big(K, X^*\big)$ allora $\xi = J_X(x) \in \mathrm{Cl}\big(J_X(K), X^*\big)$ e dunque il lemma precedente garantisce l'esistenza di una successione $x_n \in K$ convergente debolmente a x.

7.7 Spazi strettamente e uniformemente convessi

In uno spazio normato la palla unitaria chiusa $D\big(0, 1\big)$ è chiaramente un insieme convesso per le proprietà della norma. Introduciamo adesso delle proprietà aggiuntive per spazi normati che interessano le proprietà di convessità della pala unitaria, e quindi di tutte le sue palle.

Definizione 7.7.1. Uno spazio normato X si dice **strettamente convesso** se e solo se per ogni $x,y\in X$ con $\|x\|,\|y\|\leq 1$ e $x\neq y$ e per ogni $\lambda\in(0,1)$ si ha $\|\lambda x+(1-\lambda)y\|<1$.

Per ogni $x\in X$ spazio normato per il teorema di Hahn-Banach, e più nello specifico il corollario 7.2.2, l'insieme

$$\mathscr{D}(x)=\left\{f\in X^{*}\;\middle|\;\|f\|=\|x\|\text{ e }f(x)=\|x\|^{2}\right\}\tag{7.7.1}$$

non è mai vuoto. Dimostriamo che l'insieme $\mathscr{D}(x)$ è debolmente* chiuso e convesso.

Presa una qualunque successione $f_n\in\mathscr{D}(x)$ convergente debolmente* ad un qualche $f\in X^{*}$ per la proposizione 7.4.3 avremo che $\|f\|\leq\liminf_{n\to+\infty}\|f_n\|=\|x\|$ e

$$f(x)=\lim_{n\to+\infty}f_n(x)=\|x\|^{2}$$

perciò $\|f\|=\|x\|$ e $f\in\mathscr{D}(x)$. Se ora prendiamo $f,g\in\mathscr{D}(x)$ e $\lambda\in[0,1]$ allora

$$\begin{gathered}\|\lambda f+(1-\lambda)g\|\leq\lambda\|f\|+(1-\lambda)\|g\|=\|x\|\\ [\lambda f+(1-\lambda)g](x)=\|x\|^{2}\end{gathered}$$

e perciò $\mathscr{D}(x)$ è anche convesso.

Proposizione 7.7.2. *Se X^{*} è uno spazio normato strettamente convesso allora $\mathscr{D}(x)$ conterrà un solo elemento.*

Dimostrazione. Tutti gli elementi di $\mathscr{D}(x)$ hanno la medesima norma $\|x\|=r$ e in particolare $\|f/2+g/2\|=r$ per ogni $f,g\in\mathscr{D}(x)$, ma la palla chiusa $D(0,r)$ che contiene $\mathscr{D}(x)$ è strettamente convessa, dunque deve essere necessariamente $f=g$ e perciò $\mathscr{D}(x)$ possiede un unico elemento. ■

La proprietà di stretta convessità dipende dalla norma utilizzata oltre che alla topologia, difatti su $X=\mathbb{R}^{2}$ la norma $\|(x,y)\|_2=\sqrt{x^2+y^2}$ è strettamente convessa mentre $\|(x,y)\|_1=|x|+|y|$ no.

Proposizione 7.7.3. *Se X è uno spazio strettamente convesso allora l'applicazione $x\to\|x\|^{p}$ è strettamente convessa per ogni $p>1$.*

Viceversa se $x\to\|x\|^{p}$ è strettamente convessa per qualche $p>1$ allora la norma di X è strettamente convesso.

Dimostrazione. Per la stretta crescenza di t^{p} se $\|x\|\neq\|y\|$ e $0<\lambda<1$ allora per quanto detto sopra abbiamo

$$\|\lambda x+(1-\lambda)y\|^{p}\leq\left[\lambda\|x\|+(1-\lambda)\|y\|\right]^{p}<\lambda\|x\|^{p}+(1-\lambda)\|y\|^{p}$$

se invece $\|x\|=\|y\|$ ma $x\neq y$ allora per la stretta convessità di X si ha

$$\|\lambda x+(1-\lambda)y\|^{p}<\|x\|^{p}=\lambda\|x\|^{p}+(1-\lambda)\|y\|^{p}$$

Viceversa se $\|x\|^{p}$ è strettamente convessa per qualche $p>1$ allora per ogni $x,y\in D(0,1)$ con $x\neq y$ e $0<\lambda<1$ si ha $\|\lambda x+(1-\lambda)y\|^{p}<\lambda\|x\|^{p}+(1-\lambda)\|y\|^{p}=1$ e perciò X è strettamente convesso. ■

Non tutte le norme su uno spazio normato sono strettamente convesse. Ciononostante in alcuni casi è possibile trovare una norma equivalente che non solo sia strettamente convessa, ma che anche la sua polare sia strettamente convessa.

Teorema 7.7.4. *Sia X uno spazio di Banach separabile, allora esiste una norma su X equivalente alla norma iniziale che rende sia X che X^* spazi normati strettamente convessi.*

Dimostrazione. Poiché sottoinsiemi di spazi metrici separabili sono ancora separabili esiste $\{x_n\}_{n\in\mathbb{N}} \in D(0,1,X)$ denso in $D(0,1,X)$. Definiamo per ogni $f \in X^*$

$$\|f\|_1 = \sqrt{\|f\|^2 + \sum_{i=1}^{+\infty} \frac{f(x_i)^2}{2^i}}$$

non è difficile dimostrare che è effettivamente una norma su X^* con $\sqrt{2}\|f\| \geq \|f\|_1 \geq \|f\|$, vogliamo ora applicare la proposizione 7.7.3 per dimostrare che questa norma è anche strettamente convessa.

Innanzitutto per ogni $i \in \mathbb{N}$ non è difficile verificare che

$$\left|\lambda f(x_i) + (1-\lambda)g(x_i)\right|^2 + \lambda(1-\lambda)\left|f(x_i) - g(x_i)\right|^2 = \lambda\left|f(x_i)\right|^2 + (1-\lambda)\left|g(x_i)\right|^2 \tag{7.7.2}$$

e poiché se $f \neq g$ esisterà $i \in \mathbb{N}$ tale che $f(x_i) \neq g(x_i)$ sommando in i avremo che la funzione $f \to \|f\|_1^2$ è strettamente convessa e dunque la norma è convessa.

Indichiamo sempre con $\|x\|_1$ la polare su X della norma ottenuta precedentemente. Siano ora $f_n \in D(0,1,X^*)$ come nella prima parte del lemma 7.4.4 e definiamo

$$\|x\|_2 = \sqrt{\|x\|_1^2 + \sum_{i=1}^{+\infty} \frac{f_i(x)^2}{2^i}}$$

come prima è possibile dimostrare che $\|\cdot\|_2$ è equivalente alla norma di X ed è anche strettamente convessa (dato che la famiglia delle f_n è totale), ci rimane da dimostrare che anche la sua polare è strettamente convessa.

Posto $[x] = \sqrt{\sum_{i=1}^{+\infty} f_i(x)^2/2^i}$ e $[f]$ la sua polare, pur non essendo norme sono comunque convesse e positivamente omogenee e perciò possiamo applicare il corollario 6.2.10 in modo tale che

$$\frac{[f]^2}{2} = \left(\frac{[x]^2}{2}\right)^*(f) = \sup_{x\in X} f(x) - \frac{[x]^2}{2}$$

Possiamo ottenere un'uguaglianza simile alla (7.7.2) per la polare semplicemente ponendo

$$\begin{aligned}
&\frac{[\lambda f + (1-\lambda)g]^2}{2} + \frac{\lambda(1-\lambda)[f-g]^2}{2} \\
&\quad = \sup_{x,y\in X}\left[\lambda f(x) + (1-\lambda)g(x) + \lambda(1-\lambda)[f(y) - g(y)] - \frac{[x]^2}{2} - \lambda(1-\lambda)\frac{[y]^2}{2}\right] \\
&= \sup_{z,w\in X}\Bigg[\lambda f[\lambda z + (1-\lambda)w] + (1-\lambda)g[\lambda z + (1-\lambda)w] + \lambda(1-\lambda)[f(z-w) - g(z-w)] \\
&\qquad\qquad -\frac{[\lambda z + (1-\lambda)w]^2}{2} - \lambda(1-\lambda)\frac{[z-w]^2}{2}\Bigg] = \frac{\lambda[f]^2}{2} + \frac{(1-\lambda)[g]^2}{2}
\end{aligned}$$

avendo posto

$$\begin{aligned} z &= x + (1-\lambda)y \\ w &= x - \lambda y \end{aligned}$$

dunque anche la polare è strettamente convessa. Per il corollario 6.3.5 abbiamo anche

$$\frac{\|f\|_2^2}{2} = \left(\frac{\|x\|_1^2}{2} + \frac{[x]^2}{2}\right)^*(f) = \min_{g\in X^*}\left[\frac{\|f-g\|_1^2}{2} + \frac{[g]^2}{2}\right]$$

in quanto le norme considerate sono equivalenti alla norma standard e perciò continue, dunque per ogni f, g con $f \neq g$ esistono $h_f, h_g \in X^*$ tali che

$$\begin{aligned} \|f\|_2^2 &= \|f-h_f\|_1^2 + [h_f]^2 \\ \|g\|_2^2 &= \|g-h_g\|_1^2 + [h_g]^2 \end{aligned}$$

Infine osserviamo che per ogni $\lambda \in (0,1)$

$$\|\lambda f + (1-\lambda)g\|_2^2 \leq \|\lambda(f-h_f) + (1-\lambda)(g-h_g)\|_1^2 + [\lambda h_f + (1-\lambda)h_g]^2$$

e quindi sfruttando la stretta convessità delle due seminorme otteniamo che $\|\cdot\|_2^2$ è strettamente convessa anche in X^* e dunque la tesi. ■

Come abbiamo già anticipato all'inizio negli spazi strettamente convessi possiamo rimuovere dall'asserto del teorema di Mazur-Ulam l'ipotesi di suriettività. Vale infatti

Lemma 7.7.5. *Dati due spazi normati reali X, Y strettamente convessi e $f : X \to Y$ una qualunque isometria allora f è affine.*

Dimostrazione. Prendiamo $x, y \in X$ con $r = \|x-y\| > 0$, poiché X ed Y sono strettamente convessi avremo che $(x+y)/2$ sarà l'unico elemento di $D(x, r/2) \cap D(y, r/2)$ mentre $[f(x) + f(y)]/2$ sarà l'unico elemento di $D(f(x), r/2) \cap D(f(y), r/2)$.

Poiché f è una isometria abbiamo anche che

$$f\left(D(x,r/2)\cap D(y,r/2)\right) = f\left[D(x,r/2)\right] \cap f\left[D(y,r/2)\right] = D(f(x),r/2)\cap D(f(y),r/2)$$

e perciò $f[(x+y)/2] = [f(x)+f(y)]/2$. La tesi segue perciò dalla proposizione 7.1.11. ■

Definizione 7.7.6. Uno spazio di Banach X si dice **uniformemente convesso** se e solo se per ogni $\epsilon > 0$ esiste un $\delta > 0$ tale che per ogni $x, y \in X$ con $\|x\|, \|y\| \leq 1$ e $\|x-y\| > \epsilon$ si abbia

$$\left\|\frac{x+y}{2}\right\| < 1 - \delta$$

In altre parole per ogni coppia di elementi distinti della palla chiusa unitaria il loro punto medio è sufficientemente distante dalla sua frontiera.

Teorema 7.7.7. *Dato uno spazio X uniformemente convesso allora per ogni costante $p \in (1, +\infty)$, $\alpha \in (0, 1/2)$, $R > 0$, $\varepsilon > 0$ esiste $\delta > 0$ tale che*

$$\|\lambda x + (1-\lambda)y\|^p \leq \lambda\|x\|^p + (1-\lambda)\|y\|^p - \delta$$

per ogni $\lambda \in [\alpha, 1-\alpha]$ e per ogni $x, y \in X$ con $\|x\|, \|y\| \leq R$ e $\|x-y\| > \varepsilon$.

In particolare gli spazi uniformemente convessi sono strettamente convessi.

Dimostrazione. La funzione t^p con $p > 1$ è continua e strettamente convessa su $\mathbb{R}^+$, fissati inoltre $R > 0, \varepsilon > 0, \alpha \in (0, 1/2)$ l'insieme

$$\{(s,t,\lambda) \in [0,R] \times [0,R] \times [\alpha, 1-\alpha] \mid |s-t| \geq \varepsilon\}$$

è compatto in $\mathbb{R}^3$, quindi dal teorema dell'esistenza dei minimi di Weierstrass esisterà $\Lambda > 0$ tale che

$$[\lambda_n s + (1-\lambda_n)t]^p < \lambda s^p + (1-\lambda)t^p - \Lambda \tag{7.7.3}$$

per ogni s, t, λ come sopra.

Supponiamo ora per assurdo che l'asserto sia falso, esisteranno allora p, α, R, ε tali che per ogni $n \in \mathbb{N}$ esisteranno $\lambda_n \in [\alpha, 1-\alpha]$, $x_n, y_n \in D(0,R)$ con $\|x_n - y_n\| > \varepsilon$ tali che

$$\|\lambda x_n + (1-\lambda)y_n\|^p > \lambda_n \|x_n\|^p + (1-\lambda_n)\|y_n\|^p - \frac{1}{n}$$

Se $\|x_n\| - \|y_n\|$ non convergesse a 0 allora potremmo, a meno di passare ad un'estratta, supporre che $\big|\|x_n\| - \|y_n\|\big| > \varepsilon_0 > 0$ e quindi applicando la (7.7.3) seguirebbe che

$$\lambda_n \|x_n\|^p + (1-\lambda_n)\|y_n\|^p - \frac{1}{n} < \|\lambda_n x_n + (1-\lambda_n)y_n\|^p < \lambda_n \|x_n\|^p + (1-\lambda_n)\|y_n\|^p - \Lambda$$

con Λ indipendente da n, il che è assurdo.

Supponiamo allora che $\|x_n\| - \|y_n\| \to 0$, poniamo senza perdere in generalità $2\lambda_n \leq 1$ (altrimenti consideriamo $1 - \lambda_n$ al suo posto) allora definiamo

$$z_n = 2\lambda_n x_n + (1-2\lambda_n)y_n$$

segue immediatamente che $(z_n + y_n)/2 = \lambda_n x_n + (1-\lambda_n)y_n$ e di conseguenza $\|z_n - y_n\| = 2\lambda_n \|x_n - y_n\| > 2\alpha\varepsilon$ dunque esisterà $\delta_0 > 0$ tale che

$$\|\lambda x_n + (1-\lambda)y_n\| = \left\|\frac{z_n + y_n}{2}\right\| < (1-\delta_0)(\|x_n\| \vee \|y_n\|)$$

in quanto $\|z_n\| \leq \|x_n\| \vee \|y_n\|$.

Poiché $\|x_n - y_n\| > \varepsilon$ segue che $\|x_n\| \vee \|y_n\| > \varepsilon/2$ e quindi sostituendo nella disuguaglianza iniziale e dividendo avremo che

$$(1-\delta_0)^p > \lambda_n \left(\frac{\|x_n\|}{\|x_n\| \vee \|y_n\|}\right)^p + (1-\lambda_n)\left(\frac{\|y_n\|}{\|x_n\| \vee \|y_n\|}\right)^p - \frac{2}{\varepsilon n}$$

ciascuno dei due rapporti al secondo membro tende a 1 in quanto

$$\left|\frac{\|x_n\|}{\|x_n\| \vee \|y_n\|} - 1\right| \leq \frac{2\big|\|x_n\| - \|y_n\|\big|}{\varepsilon} \to 0$$

quindi il secondo membro tende a 1 al tendere di n a $+\infty$, il che è assurdo. ■

Teorema 7.7.8 (Milman-Pettis)**.** *Gli spazi di Banach uniformemente convessi sono riflessivi.*

Dimostrazione. Sia X spazio di Banach uniformemente convesso e $J_X : X \to X^{**}$ l'operatore definito in (7.2.1). Dobbiamo solo dimostrare che ogni elemento $F \in X^{**}$ di norma unitaria appartiene alla chiusura di $J_X\left[D\left(0,1,X\right)\right]$. Infatti poiché X è uno spazio di Banach isomorfo a $J_X(X)$ allora esso sarà un sottospazio chiuso di X^{**} contenente tutti gli elementi di norma unitaria, dunque dovrà coincidere con X^{**}.

Scegliamo un qualunque $\epsilon > 0$ allora esisterà un $\delta > 0$ che soddisfa la condizione di uniforme continuità, d'altronde esisterà $f \in X^*$ con $\left\|f\right\| = 1$ tale che

$$F(f) \geq 1 - \frac{\delta}{2}$$

Inoltre l'insieme

$$Y = \left\{ G \in X^{**} \;\middle|\; \left|F(f) - G(f)\right| < \frac{\delta}{2} \right\}$$

è un intorno di F nella topologia debole*. Dal lemma di Goldstine avremo che

$$J_X[D\left(0,1,X\right)] \cap Y \neq \emptyset$$

e perciò esiste $x \in D\left(0,1,X\right)$ tale che $\left|f(x) - F(f)\right| < \delta/2$.

Per assurdo esiste $\epsilon > 0$ tale che $\left\|F - J_X(x)\right\| > \epsilon$ e, poiché $D\left(0,1,X^{**}\right)$ è chiuso rispetto anche alla topologia debole* per Banach-Alaoglu, avremo che

$$F \in Z = X^{**} \setminus \left[J_X(x) + \epsilon D\left(0,1,X^{**}\right)\right]$$

e Z è debolmente* aperto. Sempre dal lemma di Goldstine $J_X(D\left(0,1,X\right)) \cap Y \cap Z \neq \emptyset$ ovvero esiste $y \in D\left(0,1,X\right)$ tale che $J_X(y) \in Y \cap Z$ ovvero

$$\left|f(y) - F(f)\right| < \frac{\delta}{2}$$
$$\left\|y - x\right\| > \epsilon$$

ma allora

$$1 - \frac{\delta}{2} \leq F(f) < \delta + f(x) + f(y) \leq \left\|x + y\right\| + \delta$$

raggiungendo in tal modo un assurdo. ∎

Proposizione 7.7.9. *Sia X spazio uniformemente convesso e siano $x_n, x \in X$. Se $x_n \rightharpoonup x$ e $\left\|x_n\right\| \to \|x\|$ allora $x_n \to x$.*

Dimostrazione. Per assurdo supponiamo che $x \neq 0$ ed esiste $\epsilon_0 > 0$ tale che $\left\|x_n - x\right\| > \epsilon_0$ per ogni n. Per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che $\left\|x_n\right\| \leq \|x\| + \varepsilon$ per ogni $n > N$ e per l'uniforme convessità esisterà $\delta > 0$ dipendente solo da ε_0 per cui $\left\|(x_n + x)/2\right\| < (\|x\| + \varepsilon)(1 - \delta)$ e così

$$\limsup_{n\to+\infty} \left\|\frac{x_n + x}{2}\right\| \leq \|x\|\,(1 - \delta)$$

D'altronde poiché x_n converge debolmente a x allora anche $(x_n + x)/2$ converge debolmente a x dunque

$$\|x\| \leq \liminf_{n\to+\infty} \left\|\frac{x_n + x}{2}\right\| \leq \limsup_{n\to+\infty} \left\|\frac{x_n + x}{2}\right\| \leq \|x\|\,(1 - \delta)$$

assurdo. ∎

Se X è uniformemente convesso non è detto che X^* sia strettamente convesso, quindi l'insieme $\mathscr{D}(x)$ potrebbe contenere più di un elemento. Ciononostante vale il seguente risultato

Proposizione 7.7.10. *Se X è uniformemente convesso allora gli insiemi $\mathscr{D}(x)$ formano una partizione di X^*, inoltre $\mathscr{D}(x) = \mathscr{D}(y)$ se e solo se $x = y$.*

Dimostrazione. Dal teorema di Milman-Pettis X è riflessivo, quindi per ogni $f \in X^* \setminus \{0\}$ per il corollario 7.4.3 esisterà $z \in X$ tale che $\|z\| = 1$ e $f(z) = \|f\|$, quindi basta porre $x = \|f\| z$ da cui segue che $f \in \mathscr{D}(x)$.

Supponiamo adesso che $f \in \mathscr{D}(x) \cap \mathscr{D}(y)$ allora $\|x\| = \|f\| = \|y\| = r$ e $f(x) = f(y) = r^2$. Se per assurdo $x \neq y$ allora $\|(x+y)/2\| < r$ per l'uniforme convessità da cui segue che

$$r^2 > \left\| \frac{x+y}{2} \right\| \|x\| \geq f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2} = r^2$$

assurdo, quindi $x = y$ e in particolare $\mathscr{D}(x)$ formano una partizione di X^*. ■

Capitolo 8

Algebre di Banach

8.1 Applicazioni multilineari

In questa sezione ci occupiamo principalmente delle applicazioni multilineari tra spazi normati. Abbiamo già dato la definizione di applicazione multilineare tra spazi vettoriali generici nel primo capitolo, per questo non la ripetiamo anche qui.

Esempi di applicazioni multilineari sono il prodotto scalare di uno spazio di Hilbert o l'applicazione definita in (9.1.1). Se $V, V_1, V_2, \dots, V_n$ sono spazi vettoriali topologici allora indicheremo con il simbolo $L^n(V_1, \dots, V_n, W)$ lo spazio di tutte le applicazioni multilineari continue con $V_1 \times V_2 \times \cdots \times V_n$ dotato della topologia prodotto, inoltre se $V_1 = V_2 = \cdots = V_n = W$ lo indicheremo semplicemente con $L^n(W, V)$.

Da adesso considereremo solamente spazi normati.

Proposizione 8.1.1. *Siano $Y, X_1, X_2, \dots, X_n$ spazi normati su $\mathbb{K}$ e l'applicazione multilineare $v : X_1 \times \cdots \times X_n \to X$, allora le affermazioni seguenti sono equivalenti:*

1. *$v \in L^n(X_1, \dots, X_n, Y)$;*
2. *v è continua in $(0, 0, \dots, 0)$;*
3. *Esiste una costante $C \geq 0$ tale che per ogni $x_i \in X_i$*

$$\left\| v(x_1, x_2, \dots, x_n) \right\|_Y \leq C \left\| x_1 \right\|_{X_1} \left\| x_2 \right\|_{X_2} \cdots \left\| x_n \right\|_{X_n}$$

Dimostrazione. L'implicazione $1 \Rightarrow 2$ è banale, dimostriamo ora che $2 \Rightarrow 3$. Supponiamo allora che per ogni $m \in \mathbb{N}$ esistono $x_m^i \in X_i$ per ogni $1 \leq i \leq n$ tali che $\left\| x_m^i \right\|_{X_i} = 1$ e

$$\left\| v(x_m^1, x_m^2, \dots, x_m^n) \right\|_Y \geq m^n$$

e quindi

$$\left\| v\left(\frac{x_m^1}{m}, \frac{x_m^2}{m}, \dots, \frac{x_m^n}{m} \right) \right\|_Y \geq 1$$

Ma $x_m^i / m \to 0$ e per la 2 v dovrebbe convergere a 0 giungendo così ad una contraddizione.

Per mostrare l'implicazione $3 \Rightarrow 1$ per comodità supponiamo $n = 2$. Siano $x_m \in X_1$, $y_m \in X_2$ convergenti ad $x \in X_1, y \in X_2$ rispettivamente, allora

$$\left\| v(x_m, y_m) - v(x, y) \right\|_Y \leq C \left\| x_m - x \right\|_{X_1} \left\| y_m \right\|_{X_2} + C \left\| x \right\|_{X_1} \left\| y_m - y \right\|_{X_2}$$

Il secondo membro converge a 0 n quanto $\left\| y_m \right\|_{X_2} \to \left\| y \right\|_{X_2}$ e quindi $\left\| y_m \right\|_{X_2}$ è limitato. ■

Quindi possiamo dotare lo spazio $L^n(X_1,\dots,X_n,Y)$ della seguente norma

$$\begin{aligned}\|v\| = \sup_{x_i\in X_i\setminus\{0\}} \frac{\left\|v(x_1,\dots,x_n)\right\|_Y}{\left\|x_1\right\|_{X_1}\cdots\left\|x_n\right\|_{X_n}} &= \sup_{\|x_i\|_{X_i}=1}\left\|v(x_1,\dots,x_n)\right\|_Y \\ &= \sup_{\|x_i\|_{X_i}\leq 1}\left\|v(x_1,\dots,x_n)\right\|_Y\end{aligned}$$

che lo rende così uno spazio normato.

Proposizione 8.1.2. *Dati degli spazi normati $X_1,X_2,\dots,X_n,Y$ su $\mathbb{K}$, allora per ogni $0\leq i\leq n$ esiste una isometria lineare biettiva dallo spazio normato $L^n(X_1,\dots,X_n;Y)$ nello spazio normato $L^i\left[X_1,\dots,X_i;L^{n-i}(X_{i+1},\dots,X_n;Y)\right]$ (ponendo $L^0(Y)=Y$).*

Dimostrazione. Possiamo tranquillamente supporre che $1\leq i<n$, prendiamo innanzitutto $f\in L^n\left(X_1,\dots,X_n;Y\right)$ allora per ogni $v_1\in X_1,\dots,v_i\in X_i$ definiamo

$$f_{v_1,v_2,\dots,v_i}:\left(v_{i+1},\dots,v_n\right)\in X_{i+1}\times\cdots\times X_n\to f\left[v_1,\dots v_n\right]\in Y$$

non è difficile verificare che effettivamente $f_{v_1,\dots,v_n}\in L^{n-i}\left(X_{i+1},\dots,X_n;Y\right)$ con $\left\|f_{v_1,\dots,v_n}\right\|\leq\left\|f\right\|\left\|v_1\right\|\cdots\left\|v_i\right\|$. Analogamente si verifica che l'applicazione

$$T_f:\left(v_1,\dots,v_i\right)\in X_1\times\cdots\times X_i\to f_{v_1,\dots,v_i}\in L^{n-i}\left(X_{i+1},\dots,X_n;Y\right)$$

appartiene a $L^i\left[X_1,\dots,X_i;L^{n-i}\left(X_{i+1},\dots,X_n;Y\right)\right]$ ed inoltre

$$\begin{aligned}\left\|T_f\left[v_1,\dots,v_i\right]\right\| &= \sup_{v_{i+1},\dots,v_n}\frac{\left\|f\left[v_1,\dots,v_n\right]\right\|}{\left\|v_{i+1}\right\|\cdots\left\|v_n\right\|} \\ &\leq\left\|f\right\|\left\|v_1\right\|\cdots\left\|v_i\right\|\end{aligned}$$

ed in particolare $\left\|T_f\right\|\leq\left\|f\right\|$. Ma allora l'applicazione l'applicazione $T:f\to T_f$ è lineare e continua con norma minore o uguale ad 1. Dato che è semplice verificare che sia biettiva dimostreremo solamente che è effettivamente una isometria tra spazi normati, ovvero $\left\|T_f\right\|=\left\|f\right\|$.

Per ogni $\varepsilon>0$ e per ogni f esistono $x_1\in X_1,\dots,v_n\in X_n$ di norma unitaria tali che

$$\begin{aligned}\left\|f\right\| &< \left\|f\left[x_1,x_2,\dots,x_n\right]\right\|+\varepsilon \\ &\leq\left\|f_{x_1,x_2,\dots,x_i}\right\|+\varepsilon \\ &\leq\left\|T_f\right\|+\varepsilon\end{aligned}$$

ma dato che f non dipende da ε quest'ultimo lo possiamo far tendere a 0 da cui segue così che $\left\|T_f\right\|=\left\|f\right\|$ e per linearità avremo che T è a tutti gli effetti una isometria biettiva. ■

Questo risultato ci permette tra l'altro di identificare gli spazi di funzioni $L^n(X;Y)$ con $L(X;L^{n-1}(X,Y))$, quindi se $T\in L^3(X,Y,Z;W)$ allora per ogni $x\in X$, $y\in Y$ e $z\in Z$ avremo $T(x)\in L^2(Y,Z;W)$, $T(x,y)\in L(Z,W)$ con

$$T(x)(y,z)=T(x,y)(z)=T(x,y,z)=T(x)(y)(z)$$

e così via.

Un'applicazione multilineare $F \in L^k(X,Y)$ è detta **simmetrica** se e solo se il suo valore non cambia se permutiamo le variabili, in altre parola per ogni $x_1, x_2, \dots, x_k \in X$ e per ogni permutazione σ di $\{1,2,\dots,k\}$ si ha

$$F(x_1, x_2, \dots, x_k) = F\left(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k)}\right)$$

e indicheremo con $Sym_k(X,Y)$ l'insieme delle applicazioni k-multilineari simmetriche da X^k in Y.

8.2 Algebre di Banach

Definizione 8.2.1. Un'**algebra** sul campo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ è uno spazio vettoriale V su $\mathbb{K}$ dotato di un'applicazione bilineare $\cdot : V \times V \to V$ associativa, ovvero per ogni $x, y, z \in V$

$$(x \cdot y) \cdot z = x \cdot (y \cdot z)$$

tale funzione sarà detta **prodotto interno**.

Se V è uno spazio di Banach allora diremo che $(V, \cdot)$ è un'**algebra di Banach** se e solo se è un'algebra e per ogni $x, y \in V$

$$\left\| x \cdot y \right\| \leq \|x\| \left\| y \right\|$$

ovvero $\cdot \in L^2(X;X)$ con norma minore o uguale ad 1.

Un'algebra V è detta **unitaria** se e solo se esiste un elemento $\mathbb{e} \in V$ tale che $\mathbb{e} \cdot x = x \cdot \mathbb{e} = x$ per ogni $x \in V$. L'elemento $\mathbb{e}$ (che se esiste è banalmente unico) è detto **unità** di V. Se però $(V, \cdot)$ è un'algebra di Banach unitaria supporremo anche che $\|\mathbb{e}\| = 1$.

Un esempio di algebra di Banach è dato dallo spazio delle funzioni lineari e continue $L(X,X)$, con X spazio di Banach, che ha come prodotto interno l'operazione di composizione di funzioni e la norma usuale di $L(X,X)$.

Definizione 8.2.2. Date due algebre X ed Y diremo che l'applicazione $f : X \to Y$ è **moltiplicativa** se e solo se per ogni $x, y \in X$

$$f(xy) = f(x) f(y)$$

Diremo che f è un **omomorfismo di algebre** se e solo se è lineare, continua e moltiplicativa. Se X ed Y sono algebre di Banach unitarie diremo che l'omomorfismo $f : X \to Y$ è **unitaria** se e solo se $f(\mathbb{e}) = \mathbb{e}$.

Se invece f è una isometria lineare, moltiplicativa allora sarà una **immersione di algebre**, mentre se è anche suriettiva allora è un **isomorfismo di algebre**.

Proposizione 8.2.3. *Data un'algebra di Banach X su $\mathbb{K}$ esistono sempre un'altra algebra di Banach unitaria Y e una immersione di algebre F da X in Y.*

Dimostrazione. Poniamo $Y = X \times \mathbb{K}$ con le operazioni usuali e con prodotto interno e norma pari a

$$(x, \lambda) \cdot (y, \mu) = (x \cdot y + \lambda y + \mu x, \lambda \mu)$$
$$\left\| (x, \lambda) \right\| = \|x\| \vee |\lambda|$$

e definiamo $F(x) = (x, 0)$. Osserviamo che

$$[(x,\lambda)+(z,\xi)]\cdot(y,\mu) = (x+z,\lambda+\xi)\cdot(y,\mu) = (x\cdot y + z\cdot y + \lambda y + \xi y + \mu x + \mu z, \lambda\mu + \xi\mu)$$
$$= (x,\lambda)\cdot(y,\mu) + (z,\xi)\cdot(y,\mu)$$
$$(x,\lambda)\cdot(0,1) = (0,1)\cdot(x,\lambda) = (x,\lambda)$$

Da questo punto è immediato che l'algebra di Banach Y soddisfa tutte le ipotesi dell'enunciato. ■

Definizione 8.2.4. Data un'algebra di Banach unitaria X diremo che un elemento x di X è

- **invertibile a destra** se e solo se esiste $y \in X$ tale che $xy = \mathbb{e}$;
- **invertibile a sinistra** se e solo se esiste $y \in X$ tale che $yx = \mathbb{e}$;
- **invertibile** se e solo se esiste $y \in X$ tale che $xy = yx = \mathbb{e}$.

Indicheremo inoltre con gli insiemi $G_d(X), G_s(X), G(X)$ i sottoinsiemi di X formati rispettivamente dagli elementi invertibili a destra, invertibili a sinistra e semplicemente invertibili.

Osservazione. Gli elementi invertibili a destra sono chiamati anche **suriettivi** mentre gli elementi invertibili a sinistra sono detti anche **iniettivi** su alcuni testi. Questo perché preso un qualunque insieme non vuoto V ed $X = F(V,V)$ con il prodotto interno $fg = f \circ g$ avremo che l'elemento neutro è l'applicazione identica id_V mentre f è invertibile a sinistra se e solo se è iniettiva, mentre è invertibile a destra se e solo se è suriettiva.

Chiaramente $\mathbb{e} \in G(X)$ mentre $0 \in G(X)$ se e solo se $X = \{0\}$, quindi possiamo tranquillamente supporre che $0 \notin G(X)$, inoltre se X è un'algebra commutativa allora $G_d(X) = G_s(X) = G(X)$.

Proposizione 8.2.5. *Per ogni algebra di Banach unitaria X si ha*

$$G(X) = G_d(X) \cap G_s(X)$$

in altre parole un elemento $x \in X$ è invertibile se e solo se possiede un'inversa destra e un'inversa sinistra. In tal caso esiste un unico elemento di X, che indicheremo sempre con x^{-1}, tale che $xx^{-1} = x^{-1}x = \mathbb{e}$.

Dimostrazione. L'inclusione $\subseteq$ è banale. Viceversa se $x \in G_d(X) \cap G_s(X)$ esisteranno $y, z \in X$ tali che $xy = zx = \mathbb{e}$. Per l'associatività del prodotto interno si ha

$$z = z\cdot\mathbb{e} = z\cdot x\cdot y = \mathbb{e}\cdot y = y$$

e quindi $x \in G(X)$. Di conseguenza l'inversa di x, se esiste, è unica. ■

Osservazione. Può capitare che un elementi di un'algebra non commutativa possa avere più di un'inversa sinistra (anche infinite) ma nessuna inversa destra. In tal caso infatti non è possibile applicare la proposizione 8.2.5 e dunque tale elemento non è invertibile.

Corollario 8.2.6. *Siano $x, y \in X$ tali che $x\cdot y = y\cdot x$, allora le seguenti affermazioni sono equivalenti:*

- *$x, y \in G(X)$;*

- $xy \in G(X)$.

Inoltre se $x \in G(X)$ allora $\lambda x \in G(X)$ per ogni $\lambda \in \mathbb{K} \setminus \{0\}$ con $(\lambda x =^{-1}= \lambda^{-1}x^{-1}$, e inoltre se $x, y \in G(X)$ allora $x^{-1}, xy \in G(X)$ con $\left(x^{-1}\right)^{-1} = x$ e $(xy)^{-1} = y^{-1}x^{-1}$. Si noti che quando l'algebra non è commutativa bisogna stare attenti all'ordine delle operazioni.

Proposizione 8.2.7. *Se $f : X \to Y$ è un omomorfismo di algebre unitario allora*

$$
\begin{aligned}
f[G_s(X)] &\subseteq G_s(Y) \\
f[G_d(X)] &\subseteq G_d(Y) \\
f[G(X)] &\subseteq G(Y)
\end{aligned}
$$

e se y è un inverso (destro, sinistro) di x allora un inverso (destro o sinistro) di $f(x)$ è $f(y)$.

Dimostrazione. Se y è l'inverso sinistro di x allora

$$f(y)f(x) = f(yx) = f(\mathbb{e}) = \mathbb{e}$$

■

Data un'algebra di Banach X per ogni $n \in \mathbb{N}$ definiamo $x^n \in X$ per ricorrenza come

$$
\begin{aligned}
x^1 &= x \\
x^{n+1} &= xx^n = x^n x
\end{aligned}
$$

se X è anche unitaria poniamo $x^0 = \mathbb{e}$ mentre se anche $x \in G(X)$ allora poniamo $x^{-n} = \left(x^{-1}\right)^n$. Non è difficile verificare che per ogni $m, n \in \mathbb{Z}$ si ha

$$x^{m+n} = x^m x^n \qquad\qquad x^{mn} = \left(x^m\right)^n$$

sempre se le operazioni siano permesse. Se però prendiamo due elementi distinti $x, y \in X$ può capitare che $(xy)^n \neq x^n y^n$, ma se $xy = yx$ allora l'uguaglianza vale per ogni $n \in \mathbb{N}$. Difatti se valesse per qualche $n \geq 1$ allora

$$(xy)^{n+1} = xy(xy)^n = xyx^n y^n = xxyx^{n-1}y^n = \cdots = xx^n yy^n = x^{n+1}y^{n+1}$$

e se X è unitaria vale anche quando $n \in \mathbb{Z}$.

Definizione 8.2.8. Data un'algebra di Banach X ed un suo sottospazio vettoriale $Y \leq X$ diremo che Y è una **sottoalgebra** di X se e solo se per ogni $x, y \in Y$ si ha $xy, yx \in Y$. Inoltre se X possiede una unità $\mathbb{e}$ diremo che Y è unitaria se e solo se $\mathbb{e} \in Y$.

Diremo invece che Y è un **ideale destro** di X se e solo se $xy \leq Y$ per ogni $x \in X$ e per ogni $y \in Y$. Se invece $yx \in Y$ per ogni $x \in X$ e per ogni $y \in Y$ allora diremo che Y è un **ideale sinistro**. Infine diremo che Y è semplicemente un **ideale bilatero** (o semplicemente **ideale**) se e solo se è un ideale destro e sinistro.

Il concetto di ideale si rivela essere molto utile nel caso di algebre di Banach unitarie, soprattutto per via del seguente risultato:

Proposizione 8.2.9. *Se X è un'algebra di Banach unitaria ed $Y \leq X$ è un ideale sinistro/destro allora sono equivalenti:*

1. *$Y = X$;*
2. *$\mathbb{e} \in Y$;*
3. *Y possiede un elemento invertibile a sinistra/destra.*

Dimostrazione. Chiaramente il punto 1 implica il punto 2 che a sua volta implica il punto 3 nel caso in cui X sia un ideale sinistro. Supponiamo che valga il punto 3 e sia $z \in X$ generico, allora per definizione di ideale sinistro avremo che $zxy \in Y$, ma essendo $xy = \mathbb{e}$ avremo che $z \in Y$ e perciò $X = Y$. ■

Proposizione 8.2.10. *La chiusura di un ideale in un'algebra di Banach è ancora un ideale. Inoltre se X è un'algebra di Banach unitaria allora la chiusura di ideali propri è ancora propria.*

Dimostrazione. Sia Y ideale di X, chiaramente $\mathrm{Cl}(Y)$ è ancora un sottospazio vettoriale di X e perciò è uno spazio di Banach. Sia $y \in \mathrm{Cl}(Y)$, poiché siamo in uno spazio $\mathcal{N}_1$ esiste una successione $y_n \in Y$ convergente ad y, inoltre per ogni $a \in X$ si ha $ay_n \in Y$. Per continuità

$$ay = \lim_{n\to+\infty} ay_n \in \mathrm{Cl}(Y)$$

e perciò $\mathrm{Cl}(Y)$ è un ideale.

Supponiamo ora che X possieda un'unità ed $Y \cap G(X) = \emptyset$ per ogni ideale proprio Y di X. Ma allora $\mathrm{Cl}(Y) \cap G(X) = \emptyset$ e dunque anche la chiusura è propria. ■

Osservazione. Con lo stesso procedimento si verifica che la chiusura di sottoalgebre di un'algebra di Banach è ancora una sottoalgebra, anzi sarà essa stessa un'algebra di Banach. Con lo stesso ragionamento si verifica che la chiusura di sottoalgebre commutative è ancora una sottoalgebra commutativa.

Lemma 8.2.11 (Krull)**.** *Sia X algebra di Banach unitaria, allora esiste in X almeno un ideale proprio massimale, ovvero un ideale proprio che non è contenuto in nessun altro ideale proprio di X.*

Dimostrazione. Sappiamo che l'ideale I di X è proprio se e solo se non contiene l'unità. Indichiamo con $\mathcal{M}$ la classe di tutti gli ideali propri di X, dato che $\{0\} \in \mathcal{M}$ esso non è vuoto. Vogliamo utilizzare il lemma di Zorn per garantire l'esistenza di elementi massimali.

Sia $\mathcal{I} \subseteq \mathcal{M}$ totalmente ordinata rispetto all'inclusione, poniamo $I = \bigcup_{J\in\mathcal{I}} J$. Poiché due qualunque elementi di $\mathcal{I}$ sono confrontabili rispetto all'inclusione anche I è un ideale di X, e poiché nessuno degli elementi di $\mathcal{I}$ contiene l'unità nemmeno I la conterrà, dunque $I \in \mathcal{M}$.

Per il lemma di Zorn esiste in X almeno un ideale massimale. ■

Osservazione. Il risultato continua a valere se ci limitiamo agli ideali destri o sinistri. Non vale però se X non possiede l'unità, in tal caso possiamo trovare spazi che non contengono alcun ideale massimale proprio.

Osservazione. Se I è un ideale massimale di un'algebra di Banach unitaria la proposizione 8.2.10 ci garantisce che anche la sua chiusura $\mathrm{Cl}(I)$ è un ideale proprio e contiene I. Per massimalità si deve così avere $I = \mathrm{Cl}(I)$ e dunque tutti gli ideali massimali in un'algebra unitaria sono chiusi. Se però l'algebra non è unitaria potrebbero esistere ideali propri densi.

I risultati ottenuti precedentemente continuano a valere anche se, invece degli ideali, consideriamo solamente gli ideali destri o sinistri. Il prossimo risultato invece vale solo per ideali destri e sinistri, e può essere esteso agli ideali solamente se l'algebra è commutativa.

Proposizione 8.2.12. *Se X è un'algebra di Banach unitaria allora per ogni $x \in X \setminus G_d(X)$ esiste un ideale destro proprio di X contenente x.*

Dimostrazione. Definiamo l'insieme

$$I = \{xa \mid a \in X\}$$

chiaramente I è un ideale destro di X per la bilinearità del prodotto interno. Ancora, $x = x\mathbb{e}$ e così $x \in I$. Per far vedere che è proprio basta verificare che non contiene l'unità, infatti se la contenesse ciò significherebbe che $xy = \mathbb{e}$ per qualche $y \in X$ e di conseguenza x possiede un'inversa destra, assurdo. ■

Prendiamo adesso un sottoinsieme non vuoto P di un'algebra di Banach X su $\mathbb{K}$, possiamo definire i seguenti sottoinsiemi di X:

$$\mathbb{K}(P) = \bigcap \{K \text{ sottoalgebra di } X \mid P \subseteq K\}$$
$$\overline{\mathbb{K}}(P) = \bigcap \{K \text{ sottoalgebra chiusa di } X \mid P \subseteq K\}$$

Non è difficile verificare che sono entrambe sottoalgebre di X contenenti P, e sono rispettivamente la più piccola sottoalgebra e la più piccola sottoalgebra chiusa di X che contengono P. Per la proposizione 8.2.10 riadattata alle sottoalgebre abbiamo anche che

$$\overline{\mathbb{K}}(P) = \mathrm{Cl}(\mathbb{K}(P))$$

Infine non è difficile verificare che la sottoalgebra $\mathbb{K}(P)$ corrisponde esattamente all'insieme

$$\mathbb{K}(P) = \left\{ \sum_{i=0}^{N} \lambda_i p_{i,1} p_{i,2} \cdots p_{i,k_i} \,\middle|\, N, k_i \in \mathbb{N}_0, \lambda_i \in \mathbb{K}, p_{i,j} \in P \right\}$$

8.3 L'algebra dei polinomi

Fissato un campo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ ed un intero $n \in \mathbb{N}$, definiamo lo **spazio dei polinomi** con n incognite come la somma diretta di $\mathbb{N}_0^n$ copie del campo $\mathbb{K}$ ovvero

$$\mathbb{K}[n] = \bigoplus_{\boldsymbol{\gamma} \in \mathbb{N}_0^n} \mathbb{K}$$

dove ricordiamo che $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ e quindi gli elementi di $\mathbb{N}_0^n$ sono semplicemente n-ple di numeri interi non negativi. Un polinomio in n incognite non sarà quindi altro che un'applicazione $p : \mathbb{N}_0^+ \to \mathbb{K}$ che ad ogni n-pla $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ associa lo scalare $p_{\gamma_1,\ldots,\gamma_n} \in \mathbb{K}$ il quale sarà diverso da 0 solamente per un numero finito di n-ple per definizione di somma diretta. Possiamo così definire per un polinomio $p \in \mathbb{K}[n]$ il suo **grado** come

$$\deg p = \max\left\{ |\boldsymbol{\gamma}| \,\middle|\, p_{\gamma_1,\ldots,\gamma_n} \neq 0 \right\}$$

dove $|\boldsymbol{\gamma}| = \gamma_1 + \gamma_2 + \cdots + \gamma_n$. Naturalmente $\deg p = -\infty$ se e solo se $p \equiv 0$.

Dato che lo spazio $\mathbb{K}[n]$ possiede già la struttura di spazio vettoriale su $\mathbb{K}$ lo dotiamo anche di un'operazione di prodotto interno definito nella seguente maniera: per ogni coppia di polinomi $p,q \in \mathbb{K}[n]$ definiamo il polinomio prodotto $pq \in \mathbb{K}[n]$ in modo tale che per ogni n-pla $\mathbf{a} \in \mathbb{N}_0^n$

$$(pq)_{\mathbf{a}} = \sum \left\{ p_{\mathbf{b}} q_{\mathbf{c}} \;\middle|\; \mathbf{a} = \mathbf{b} + \mathbf{c} \right\}$$

dove la somma degli elementi di $\mathbb{N}_0^n$ è intesa termine a termine, ovvero $a_i = b_i + c_i$ per ogni $1 \leq i \leq n$. Si noti che, per come abbiamo definito i polinomi, la sommatoria a destra è sempre finita e perciò il prodotto di due polinomi è ben definito. Inoltre per la commutatività del prodotto in $\mathbb{K}$ avremo che anche il prodotto tra polinomi commuta ovvero $pq = qp$.

Lemma 8.3.1. *Lo spazio dei polinomi in n incognite $\mathbb{K}[n]$ è un'algebra commutativa unitaria su $\mathbb{K}$ con prodotto interno $p \cdot q = pq$ ed unità il polinomio $\mathbb{1}$ definito in modo tale che $\mathbb{1}_{0,0,\cdots,0} = 1$ ed $\mathbb{1}_{\mathbf{i}} = 0$ per ogni altra scelta di $\mathbf{i} \in \mathbb{N}_0^n$ con almeno una componente non nulla. Inoltre*

$$\deg(pq) = \deg p + \deg q$$

Dimostrazione. Esercizio. ∎

Prendiamo adesso un'algebra di Banach unitaria X sul campo $\mathbb{K}$ ed n suoi elementi $x_1, x_2, \dots, x_n$ di X che commutano tra loro, ovvero $x_i x_j = x_j x_i$ per ogni coppia di indici i,j. Posto $\mathbf{x} = \left(x_1, x_2, \dots, x_n\right) \in X^n$ per comodità di notazione scriviamo

$$\mathbf{x}^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \in X$$

per ogni $\mathbf{i} \in \mathbb{N}_0^n$, ponendo sempre $0^0 = \mathbb{e}$. L'ipotesi di commutatività degli x_i implica inoltre che $\mathbf{x}^{\mathbf{i}+\mathbf{j}} = \mathbf{x}^{\mathbf{i}} \mathbf{x}^{\mathbf{j}}$ per ogni $\mathbf{i}, \mathbf{j} \in \mathbb{N}_0^n$.

Posto allora $C_n(X)$ l'insieme di tutte le n-ple di elementi di X formate da elementi che commutano tra loro ogni polinomio $p \in \mathbb{K}[n]$ induce un'applicazione definita su $C_n(X)$ a valori in X definita nella seguente maniera:

$$p(\mathbf{x}) = \sum_{\mathbf{a} \in \mathbb{N}_0^n} p_{\mathbf{a}} \mathbf{x}^{\mathbf{a}} = \sum_{a_1, a_2, \dots, a_n} p_{a_1, a_2, \dots, a_n} x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \in X$$

che come ben sappiamo è sempre ben definito. Osserviamo in particolare che $(p+q)(x) = p(x) + q(x)$ per ogni $x \in X$ ed $\mathbb{1}(x) = \mathbb{e}$, ma soprattutto che il polinomio p da solo non ha nulla a che fare con l'algebra di Banach X scelta oltre al campo in comune $\mathbb{K}$, e quindi tutte le applicazioni che il polinomio p può indurre conservano molte caratteristiche in comune tra loro. Osserviamo inoltre che $C_1(X) = X$ mentre quando $n \geq 2$ allora $C_n(X)$ è un sottoinsieme chiuso non vuoto di X^n con la topologia prodotto, infatti posto

$$R_{ij}(\mathbf{x}) = x_i x_j - x_j x_i$$

avremo che R_{ij} è continuo ed $C_n(X) = \bigcap_{i,j} \left\{ R_{ij} = 0 \right\}$.

Il prossimo passo è dimostrare che se X è un'algebra di Banach unitaria allora il polinomio p visto come applicazione da $C_n(X)$ in X è continua su tutto il dominio e lipschitziana sugli insiemi limitati rispetto alla norma prodotto. Osserviamo innanzitutto che per ogni $n \in \mathbb{N}$ e per ogni $x, y \in B\left(0, r\right)$ coppia di elementi di X

$$\left\| x^n - y^n \right\| \leq n r^{n-1} \left\| x - y \right\|$$

La disuguaglianza è chiaramente vera quando $n = 1$, se invece è vera per qualche $n \in \mathbb{N}$ allora

$$\begin{aligned}\left\|x^{n+1} - y^{n+1}\right\| &= \left\|x^{n+1} - y \cdot x^n + y \cdot x^n - y^{n+1}\right\| \\ &\leq \left\|x - y\right\| r^n + r\left\|x^n - y^n\right\| \leq (n+1) r^n \left\|x - y\right\|\end{aligned}$$

Successivamente se prendiamo $x, y, z, w \in B(0, r)$ ed $m, n \in \mathbb{N}_0$ avremo che

$$\begin{aligned}\left\|x^m y^n - z^m w^n\right\| &= \left\|x^m y^n - x^m w^n + x^m w^n - z^m w^n\right\| \\ &\leq \|x\|^m \left\|y^n - w^n\right\| + \left\|x^m - z^m\right\| \|w\|^n \\ &\leq n r^{m+n} \left\|y - w\right\| + m r^{m+n} \|x - z\| \\ &\leq (m+n) r^{m+n} \left(\|x - z\| \vee \left\|y - w\right\|\right)\end{aligned}$$

lavorando alla stessa maniera sui polinomi in n variabili si deduce che p è un'applicazione lipschitziana sull'aperto $B(0, r)$ per ogni $r > 0$ dello spazio di Banach X^n dotato della metrica prodotto e dunque è continua.

Proposizione 8.3.2. *Se $x_1, x_2, \ldots, x_n \in X$ commutano a due a due allora per ogni coppia di polinomi $p, q \in \mathbb{K}[n]$ si ha*

$$p(x_1, \ldots, x_n) q(x_1, \ldots, x_n) = q(x_1, \ldots, x_n) p(x_1, \ldots, x_n) = (pq)(x_1, \ldots, x_n)$$

Dimostrazione. Per comodità poniamo sempre $\mathbf{x} = \left(x_1, x_2, \ldots, x_n\right) \in C_n(X)$, e supponiamo anche che i gradi di p e q siano pari rispettivamente ad M ed $N \in \mathbb{N}_0$. Allora

$$\begin{aligned}\sum_{|\mathbf{a}| \leq M+N} (pq)_{\mathbf{a}} \mathbf{x}^{\mathbf{a}} &= \sum_{|\mathbf{a}| \leq M+N} \sum_{\mathbf{a} = \mathbf{b}+\mathbf{c}} p_{\mathbf{b}} q_{\mathbf{c}} \mathbf{x}^{\mathbf{b}+\mathbf{c}} \\ &= \sum_{|\mathbf{b}+\mathbf{c}| \leq M+N} \left(p_{\mathbf{b}} \mathbf{x}^{\mathbf{b}}\right)\left(q_{\mathbf{c}} \mathbf{x}^{\mathbf{c}}\right) \\ &= \sum_{j=0}^{M+N} \sum_{i=0}^{M+N-j} P_i Q_j\end{aligned}$$

dove $P_i = \sum_{|\mathbf{b}|=i} p_{\mathbf{b}} \mathbf{x}^{\mathbf{b}}$, $Q_j = \sum_{|\mathbf{c}|=j} q_{\mathbf{c}} \mathbf{x}^{\mathbf{c}}$, in quanto $|\mathbf{b} + \mathbf{c}| = |\mathbf{b}| + |\mathbf{c}|$.

Ora osserviamo che se $j > N$ allora $Q_j = 0$ e quindi anche la somma più interna è nulla, mentre se $j \leq N$ allora $M + N - j \geq M$ e quindi per lo stesso motivo possiamo limitarci a considerare i che va da 0 ad M. Abbiamo così che

$$\begin{aligned}(pq)(\mathbf{x}) &= \sum_{j=0}^{N} \sum_{i=0}^{M} P_i Q_j \\ &= \left(\sum_{i=0}^{M} P_i\right)\left(\sum_{j=0}^{N} Q_j\right) \\ &= p(\mathbf{x}) q(\mathbf{x})\end{aligned}$$

Per concludere essendo $\mathbf{x}^{\mathbf{b}} \mathbf{x}^{\mathbf{c}} = \mathbf{x}^{\mathbf{c}} \mathbf{x}^{\mathbf{b}}$ ripetendo i passaggi di prima avremo che $(pq)(\mathbf{x}) = q(\mathbf{x}) p(\mathbf{x})$ concludendo in tal modo la dimostrazione. ∎

Corollario 8.3.3. *Siano X, Y algebre unitarie su $\mathbb{K}$ ed $f : X \to Y$ applicazione lineare unitaria e moltiplicativa. Allora posto $\tilde{f}\left(x_1, x_2, \ldots, x_n\right) = \left[f(x_1), f(x_2), \ldots, f(x_n)\right]$ abbiamo che $\tilde{f}[C_n(X)] \subseteq C_n(Y)$ e per ogni $p \in \mathbb{K}[n]$ si ha*

$$f[p(\mathbf{x})] = p\left[\tilde{f}(\mathbf{x})\right]$$

Dimostrazione. Presi $x, y \in X$ tali che $xy = yx$ per la moltiplicatività di f si ha $f(x)f(y) = f(xy) = f(yx) = f(y)f(x)$ e così $\tilde{f}[C_n(X)] \subseteq C_n(Y)$.

Prendiamo ora $\mathbf{x} \in C_n(X)$ e $\boldsymbol{\gamma} \in \mathbb{N}_0^n$, essendo f moltiplicativa e unitaria è immediato verificare che

$$f\left(\mathbf{x}^{\gamma}\right) = \left[\tilde{f}(\mathbf{x})\right]^{\gamma}$$

mentre per linearità si avrà anche

$$f[p(\mathbf{x})] = f\left(\sum_{\gamma \in \mathbb{N}_0^n} p_{\gamma} \mathbf{x}^{\gamma}\right) = \sum_{\gamma \in \mathbb{N}_0^n} p_{\gamma} \left[\tilde{f}(\mathbf{x})\right]^{\gamma} = p\left[\tilde{f}(\mathbf{x})\right]$$

■

Il seguente risultato discende velocemente da quanto detto in questa e nella precedente sezione:

Proposizione 8.3.4. *Data un'algebra unitaria X su $\mathbb{K}$ ed n elementi $x_1, x_2, \dots, x_n \in X$ che commutano a due a due allora*

$$\mathbb{K}\left(\mathbb{e}, x_1, x_2, \dots, x_n\right) = \left\{p\left(x_1, \dots, x_n\right) \mid p \in \mathbb{K}[n]\right\}$$

Osservazione. Chiaramente a partire un polinomio $p \in \mathbb{K}[1]$ possiamo definire $p(x)$ per qualunque elemento x di una qualunque algebra di Banach unitaria definita sul capo $\mathbb{K}$, in particolare anche su $\mathbb{K}$ e perciò per ogni $\lambda \in \mathbb{K}$ si ha

$$p(\lambda \mathbb{e}) = p(\lambda)\mathbb{e}$$

Ora se scegliamo un qualunque insieme non vuoto O lo spazio di funzioni $F(O, X)$ con le operazioni standard di somma, prodotto esterno ed interno ricavate a partire dall'algebra unitaria X è essa stessa un'algebra unitaria (e di conseguenza anche $\mathbb{K}[n]$ è un'algebra unitaria con $O = C_n(X)$ grazie alla proposizione 8.3.2). Di conseguenza un qualunque polinomio $p \in \mathbb{K}[m]$ induce un'applicazione da $C_m[F(O, X)]$ in $F(O, X)$, e in particolare manda polinomi in sé.

Se ora scegliamo $\left(f_1, f_2, \dots, f_m\right) \in C_m[F(O, X)]$ per come abbiamo definito il prodotto interno in $F(O, X)$ si ha

$$\left(f_i^k f_j^h\right)(o) = \left[f_i(o)\right]^k \left[f_j(o)\right]^h$$

per ogni $o \in O$ e dunque

$$p\left(f_1, f_2, \dots, f_m\right)(o) = p\left[f_1(o), f_2(o), \dots, f_m(o)\right] = \left(p \circ \vec{f}\right)(o)$$

sempre per ogni $o \in O$.

Da questa osservazione discende immediatamente che

Teorema 8.3.5. *Per ogni $p \in \mathbb{K}[1]$ e per ogni $\lambda \in \mathbb{K}$ esiste un unico polinomio $q \in \mathbb{K}[1]$ tale che per ogni algebra unitaria X e per ogni $x \in X$*

$$q(x) = p(x - \lambda \mathbb{e})$$

Osservazione. Il polinomio q non dipende né da x né dallo spazio X.

Dimostrazione. Definiamo il polinomio $l \in \mathbb{K}[1]$ in modo tale che $l_0 = -\lambda$, $l_1 = 1$ ed $l_i = 0$ altrimenti, osserviamo che $l(x) = x - \lambda e$ per ogni $x \in X$. Basta così porre $q = p(l) = p \circ l$ per dimostrare il teorema. ■

Corollario 8.3.6. *Sia X algebra di Banach su $\mathbb{K}$, $p \in \mathbb{K}[1]$ e $\lambda \in \mathbb{K}$ tale che $p(\lambda) = 0$. Allora esiste $q \in \mathbb{K}[1]$ tale che per ogni $x \in X$*

$$p(x) = (x - \lambda e)q(x)$$

Dimostrazione. Supponiamo innanzitutto che $\lambda = 0$, allora avremo che

$$p_0 = p(\lambda) = 0$$

quindi se definiamo il polinomio $q \in \mathbb{K}[1]$ in modo tale che $q_i = p_{i+1}$ per la proposizione 8.3.2 avremo che $p(x) = q(x)x = xq(x)$ per ogni x appartenente a qualche algebra unitaria.

Se invece $\lambda \neq 0$ per il teorema 8.3.5 esisterà un qualche polinomio q in una variabile tale che $q(r) = p(r + \lambda)$ per ogni $r \in \mathbb{K}$ e quindi $q(0) = 0$. Per quanto detto prima abbiamo così che

$$p(x) = (x - \lambda e)q(x - \lambda e) = (x - \lambda e)q'(x)$$

come volevasi dimostrare. ■

Teorema 8.3.7 (Teorema fondamentale dell'algebra)**.** *Se $\mathbb{K} = \mathbb{C}$ allora per ogni $p \in \mathbb{C}[1]$ con grado ≥ 1 esiste $\lambda \in \mathbb{C}$ tale che $p(\lambda) = 0$.*

Adesso che abbiamo dotato lo spazio $\mathbb{K}[n] = \bigoplus_{\boldsymbol{\gamma} \in \mathbb{N}_0^n} \mathbb{K}$ di una struttura di algebra commutativa unitaria su $\mathbb{K}$ possiamo fare lo stesso anche con lo spazio $\prod_{\boldsymbol{\gamma} \in \mathbb{N}_0^n} \mathbb{K} = F(\mathbb{N}_0^n, \mathbb{K})$ usando il medesimo prodotto interno. A differenza dei polinomi non è sempre possibile definire un'analoga funzione da $C_n(X)$ in X in quanto potremmo ottenere una somma infinita di elementi di X, ed è in questo momento che dobbiamo utilizzare la struttura di spazio di Banach di X.

Innanzitutto presa una qualunque funzione $f \in F(\mathbb{N}_0^n, \mathbb{K})$ ed una n-pla $\mathbf{x} \in C_n(/X)$ diremo che la **serie di potenze** di termine generale $f_{\boldsymbol{\gamma}}$ **converge in x** se e solo se la nuova applicazione definita su $\mathbb{N}_0^n$

$$g(\boldsymbol{\gamma}) = f_{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\gamma}} = f_{\gamma_1, \ldots, \gamma_n} x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$$

appartiene allo spazio $\ell^1(\mathbb{N}_0^n, X)$. Grazie quindi al teorema 12.2.11 ed alla completezza di X possiamo definire $f(\mathbf{x})$ nella seguente maniera:

$$f(\mathbf{x}) = \int_{\mathbb{N}_0^n} f_{\mathbf{j}} \mathbf{x}^{\mathbf{j}} \, d\#(\mathbf{j}) = \sum_{\mathbf{j} \in \mathbb{N}_0^n} f_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}$$

Osserviamo immediatamente che preso $\mathbf{0} = (0, 0, \ldots, 0)$ allora la serie di potende di termine generale f convergerà sempre in $\mathbf{0}$ con $f(\mathbf{0}) = f_{0,0,\ldots,0} e$. Inoltre

Proposizione 8.3.8. *Se la serie di potenze di termine generale f converge in $\mathbf{x}$ allora $f(\mathbf{x}) = y \in X$ se e solo se per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}_0$ tale che per ogni $j_i \geq N$ in $\mathbb{N}_0$ si ha*

$$\left\| \sum_{l_i \leq j_i} f_{l_1, \ldots, l_n} x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n} - y \right\| < \varepsilon$$

Dimostrazione. Introduciamo su $\mathbb{N}_0^n$ la relazione di equivalenza $\boldsymbol{\gamma} \leq \boldsymbol{\eta} \Leftrightarrow \gamma_i \leq \eta_i$ per ogni $1 \leq i \leq n$ e quindi $\left(\mathbb{N}_0^n, \leq\right)$ è un insieme diretto. Possiamo quindi definire una topologia su $\tilde{\mathbb{N}}_0^n = \mathbb{N}_0^n \cup \{+\infty\}$ come per qualunque altro insieme diretto e considerare così l'applicazione tra spazi topologici

$$F : \boldsymbol{\gamma} \in \mathbb{N}_0^n \to \sum_{\boldsymbol{\eta} \leq \boldsymbol{\gamma}} f_{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\eta}} \in X$$

Il teorema della convergenza dominata ci dice che $\lim_{m\to+\infty} F\left(\boldsymbol{\gamma}_m\right) = f(\mathbf{x})$ per ogni $\boldsymbol{\gamma}_n \to +\infty$ ma essendo X uno spazio $\mathscr{N}_1$ allora $f(\mathbf{x}) = \lim_{\boldsymbol{\gamma}\to+\infty} F\left(\boldsymbol{\gamma}\right) = y$ come volevasi dimostrare. ■

In questo modo ad esempio possiamo definire su qualunque algebra di Banach l'applicazione esponenziale

$$e^x = \sum_{i=0}^{+\infty} \frac{x^i}{i!} \in \overline{\mathbb{K}}(\mathbb{e}, x)$$

che converge per ogni $x \in X$ in quanto $\sum_i \left\| x^i / i! \right\| \leq \sum_i r^i / i! < +\infty$. Osserviamo inoltre che se $xy = yx$ allora $e^{x+y} = e^x e^y$ ma se $xy \neq yx$ allora l'uguaglianza precedente non è più valida in genere. In particolare $e^x \in G(X)$ per ogni $x \in X$ con $\left(e^x\right)^{-1} = e^{-x}$.

Sempre in un'algebra di Banach unitaria su $\mathbb{K}$ è possibile calcolare la potenza di un vettore che risulti conforme con l'operazione potenza in $\mathbb{R}$. Innanzitutto per ogni $\alpha \in \mathbb{K}$ e per ogni $k \in \mathbb{N}_0$ definiamo per ricorrenza

$$\Delta_0(\alpha) = 1$$
$$\Delta_{k+1}(\alpha) = \frac{\alpha - k}{k+1} \Delta_k(\alpha)$$

Per ogni $x \in X$ tale che $\|x\| < 1$ definiamo l'elemento $P_\alpha(x) \in X$ in modo tale che

$$P_\alpha(x) = \sum_{i=0}^{+\infty} \Delta_i(\alpha) x^i$$

dimostriamo ora che la quantità è ben definita.

Innanzitutto si ha

$$\left| \frac{\Delta_{k+1}(\alpha)}{\Delta_k(\alpha)} \right| = \left| \frac{\alpha - k}{k+1} \right|$$

che tende ad 1 al tendere di k a $+\infty$. Questo significa che per ogni $\varepsilon > 0$ esiste un $N \in \mathbb{N}$ tale che per ogni $n \geq N$

$$\left| \Delta_n(\alpha) \right| \leq (1+\varepsilon) \left| \Delta_{n-1}(\alpha) \right| \leq (1+\varepsilon)^{n-N} \left| \Delta_N(\alpha) \right|$$

quindi scegliendo ε in modo tale che $\|x\|(1+\varepsilon) < 1$ poiché N è fisso avremo che $\Delta_i(\alpha) x^i \in \ell^1(\mathbb{N}_0, X)$ e quindi la quantità $P_\alpha(x)^\alpha$ esiste ed è ben definita.

Osserviamo che quando $\alpha \in \mathbb{N}_0$ allora la serie è definitivamente nulla e perciò può essere definita su tutto X, in particolare si ha

$$\sum_{i=0}^{+\infty} \Delta_i(2) x^i = \mathbb{e} + 2x + x^2 = (\mathbb{e} + x)^2$$
$$\sum_{i=0}^{+\infty} \Delta_i(3) x^i = \mathbb{e} + 3x + 3x^2 + x^3 = (\mathbb{e} + x)^3$$
$$\dots$$

quindi sembra ragionevole porre $P_\alpha(x) = (\mathbb{e}+x)^\alpha$ per ogni $\alpha \in \mathbb{C}$. Infatti si ha per ogni $\alpha, \beta \in \mathbb{K}$ si ha

$$\Delta_k(\alpha+\beta) = \sum_{i=0}^{k} \Delta_i(\alpha)\Delta_{k-i}(\beta)$$

esistono diverse dimostrazioni di tale identità, ma la più semplice consiste nell'utilizzare la serie di Taylor della funzione $(1+t)^\alpha$ quando $t \in \mathbb{R}$ ed $\alpha \in \mathbb{C}$ in quanto i coefficienti $\Delta_k(\alpha)$ non dipendono in alcun modo dallo spazio X.

Quindi se X è un'algebra di Banach su $\mathbb{K}$ allora per ogni $\alpha, \beta \in \mathbb{K}$ e per ogni $x \in X$ con $\|x\| < 1$ si ha

$$(\mathbb{e}+x)^\alpha(\mathbb{e}+x)^\beta = (\mathbb{e}+x)^{\alpha+\beta} \tag{8.3.1}$$

come nel caso scalare. In particolare per ogni $x \in B(\mathbb{e}, 1)$ si ha $(\sqrt{x})^2 = x$, però si noti che in generale $\sqrt{x^2} \neq x$.

Generalizziamo quanto detto nei due esempi precedenti con un criterio di convergenza e continuità delle serie di potenze.

Teorema 8.3.9 (Cauchy-Hadamard)**.** *Data una serie di potenze di termine generale f per cui*

$$m = \limsup_{n\to+\infty} \sqrt[n]{\sum_{|\gamma|=n} |f_\gamma|} < +\infty$$

allora se $m > 0$ l'applicazione $\mathbf{x} \to f(\mathbf{x})$ è ben definita e continua su $C_n(X) \cap B(0, 1/m)$ rispetto alla norma prodotto, mentre se $m = 0$ allora è ben definita e continua su tutto $C_n(X)$. Inoltre per ogni $0 < s < 1/m$ finito avremo che l'applicazione è uniformemente continua su $C_n(X) \cap D(0, s)$.

Dimostrazione. Per comodità di notazione scriviamo $g_r = \sum_{|\gamma|=r} |f_\gamma| \in \mathbb{K}$ e la somma al secondo membro è finita. Dalla definizione di estremo superiore per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che $g_r \leq (m+\varepsilon)^r$ per ogni $r \geq N$. Prendiamo adesso un qualunque $\mathbf{x} \in C_n(X) \cap B(x, 1/m)$, possiamo scegliere ε sufficientemente piccolo in modo tale che $\|\mathbf{x}\| = \max_i \|x_i\| < 1/(m+\varepsilon)$ e così per ogni $r \geq N$

$$\sum_{|\gamma|=r} \|f_\gamma \mathbf{x}^\gamma\| \leq \sum_{|\gamma|=r} |f_\gamma| \, \|x_1\|^{\gamma_1} \|x_2\|^{\gamma_2} \cdots \|x_n\|^{\gamma_n} \leq g_r \|\mathbf{x}\|^r$$

Applicando ora il teorema di Fubini-Tonelli si avrà

$$\sum_{\gamma \in \mathbb{N}_0^n} \|f_\gamma \mathbf{x}^\gamma\| = \sum_{r=0}^{+\infty} \sum_{|\gamma|=r} \|f_\gamma \mathbf{x}^\gamma\| \leq D + \sum_{r=N}^{+\infty} [(m+\varepsilon)\|\mathbf{x}\|]^r < +\infty$$

in quanto $(m+\varepsilon)\|x\| < 1$ e quindi la serie di potenze converge in $\mathbf{x}$.

Per quanto riguarda la continuità ricordiamo che l'applicazione

$$h_\gamma : \mathbf{x} \in C_n(X) \cap D(0, s) \to f_\gamma \mathbf{x}^\gamma \in X$$

è lipschitziana per ogni $0 \leq s < 1/m$ e quindi è uniformemente continua e limitata. Posto allora

$$CC = \left\{ f : C_n(X) \cap D(0, s) \to X \text{ uniformemente continua} \;\middle|\; \sup_{\|\mathbf{x}\| \leq s} \|f\|_\infty = \|f(\mathbf{x})\| < +\infty \right\}$$

con la metrica $\|\cdot\|_\infty$ allora per il corollario 3.6.3 CC è uno spazio di Banach su $\mathbb{K}$ ($C_n(X)$ è chiuso) e l'applicazione $h : \boldsymbol{\gamma} \in \mathbb{N}_0^n \to h_{\boldsymbol{\gamma}} \in CC$ appartiene allo spazio $\ell^1\left(\mathbb{N}_0^n, CC\right)$ in quanto per $r = \left|\boldsymbol{\gamma}\right|$ sufficientementemente grande e per ogni $\mathbf{x} \in C_n(X) \cap D\left(0, s\right)$

$$\left\| h_{\boldsymbol{\gamma}}(\mathbf{x}) \right\| \leq \left| g_r \right| \|x\|^r < [(m+\varepsilon)s]^r$$

con $\varepsilon > 0$ scelto in modo tale che $s < 1/(m+\varepsilon)$. Quindi le somme parziali finite convergono uniformemente ad una qualche funzione $F \in CC$ e dunque per ogni $x \in X$

$$F(x) = \sum_{i=0}^{+\infty} \sum_{|\gamma|=i} f_{\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\gamma}} = f(\mathbf{x})$$

Ciò implica in particolare che il valore di F non dipende dalla scelta di s, che è continua su tutto $C_n(X) \cap B\left(0, 1/m\right)$ ed è uniformemente continua su ogni $C_n(X) \cap D\left(0, s\right)$ al variare di $s < 1/m$. ■

Osservazione. Poiché $\lim_{n\to+\infty} 1/\sqrt[n]{n!} = 0$ l'esponenziale e^x è ben definito e continuo su tutto X.

Utilizziamo i risultati precedenti sulle serie di potenze siamo in grado di dimostrare il seguente risultato:

Teorema 8.3.10. *Data un'algebra di Banach unitaria X allora l'insieme $G(X)$ è aperto in X e l'applicazione $x \in G(X) \to x^{-1} \in G(X)$ è continua.*

Dimostrazione. Innanzitutto sappiamo che $\mathbb{e} \in G(X)$ e dunque non è vuoto. Inoltre per ogni $x \in B\left(0, 1\right)$ grazie alla (8.3.1) abbiamo che

$$(\mathbb{e} + x)^{-1} = \sum_{i=0}^{+\infty} \Delta_i(-1) x^i = \sum_{i=0}^{+\infty} (-x)^i$$

è effettivamente l'inversa di x ed è un'applicazione continua per Cauchy-Hadamard, dunque $B\left(\mathbb{e}, 1\right) \subseteq G(X)$ e l'applicazione $x \in B\left(\mathbb{e}, 1\right) \to x^{-1}$ è continua.

Prendiamo adesso $z \in G(X)$, $x \in X$ generici generico tali che $\|x - z\| < 1/\left\|z^{-1}\right\|$, ma allora

$$\left\|\mathbb{e} - z^{-1}x\right\| \leq \left\|z^{-1}\right\| \|z - x\| < 1$$

e quindi $z^{-1}x \in G(X)$ e di conseguenza anche $x = zz^{-1}x \in G(X)$. Infine le applicazioni

$$x \in B\left(z, \frac{1}{\left\|z^{-1}\right\|}\right) \to z^{-1}x \in B\left(\mathbb{e}, 1\right) \to \left(z^{-1}x\right)^{-1} \to x^{-1}$$

sono tutte continue nei rispettivi domini, dunque l'applicazione inversa è continua in z e di conseguenza è continua su tutto $G(X)$. ■

Per quanto riguarda l'applicazione $x \in B\left(\mathbb{e}, 1\right) \to x^\alpha \in X$ con $\alpha \in \mathbb{K}$ possiamo modificare un po' la dimostrazione del teorema di Cauchy-Hadamard per dimostrare che essa è ben definita e continua sul suo intervallo. Quando $\alpha = 0$ oppure $\operatorname{Re}\alpha > 0$ possiamo definire x^α anche quando $\|x - \mathbb{e}\| = 1$, per farlo dobbiamo prima dimostrare alcuni risultati preliminari.

Proposizione 8.3.11. *Sia a_n una successione di numeri reali non negativi tale che $\sum_{i=1}^{+\infty} a_i < +\infty$, allora $\prod_{i=1}^{+\infty}\left(1 + a_n\right) < +\infty$.*

Dimostrazione. Sia $\sum_{i=1}^{+\infty} a_n = M \in \mathbb{R}$, dato che per ogni $t \geq 0$ abbiamo che $\ln(1+t) \leq t$ avremo per la continuità del logaritmo

$$0 \leq \ln \prod_{i=1}^{+\infty} \left(1+a_n\right) = \sum_{i=1}^{+\infty} \ln\left(1+a_n\right) \leq \sum_{i=1}^{+\infty} a_n = M$$

e quindi $\prod_{i=1}^{+\infty} \left(1+a_n\right) \leq e^M$, in particolare il limite esiste ed è finito. ∎

Lemma 8.3.12. *Dato un numero complesso $\alpha \in \mathbb{C}$ con $\operatorname{Re}\alpha > 0$ esiste una costante $M > 0$ tale che per ogni $k \in \mathbb{N}_0$*

$$\left|\Delta_k(\alpha)\right| \leq \frac{\sqrt{M}}{(k+1)^{1+\operatorname{Re}\alpha}} \tag{8.3.2}$$

Dimostrazione. Poniamo innanzitutto $\alpha = a + ib$ con $a > 0$, allora $|\alpha - k|^2 = |k-a|^2 + b^2$ e dunque

$$\begin{aligned}\frac{\frac{|\alpha-k|^2}{(k+1)^2}}{\left(\frac{k+1}{k+2}\right)^{2(1+a)}} &= \frac{|(k+1)-(a+1)|^2 + b^2}{(k+1)^2}\left(\frac{k+2}{k+1}\right)^{2(1+a)} \\ &= \left[\left|1 - \frac{a+1}{k+1}\right|^2 + \left(\frac{b}{k+1}\right)^2\right]\left(1 + \frac{1}{k+1}\right)^{2(a+1)} = 1 + A_k\end{aligned}$$

Per il teorema di Taylor con il resto di Peano avremo che

$$\left(1 + \frac{1}{k+1}\right)^{2(1+a)} = 1 + \frac{2(a+1)}{k+1} + o\left(\frac{1}{(k+1)^2}\right)$$

dove $o(t)$ è una opportuna funzione per cui sappiamo che $\limsup_{t\to 0} |o(t)|/|t| < +\infty$. In particolare per $|t|$ sufficientemente piccolo avremo che $|o(t)| \leq L|t|$ per una opportuna costante $L > 0$. Abbiamo quindi che

$$\begin{aligned}\left|A_k\right| &= \left|\left[\left|1 - \frac{a+1}{k+1}\right|^2 + \left(\frac{b}{k+1}\right)^2\right]\left(1 + \frac{1}{k+1}\right)^{2(a+1)} - 1\right| \\ &= \left|\left(1 - \frac{2(a+1)}{k+1} + \frac{(a+1)^2}{(k+1)^2} + \frac{b^2}{(k+1)^2}\right)\left(1 + \frac{1}{k+1}\right)^{2(a+1)} - 1\right| \\ &= \left|1 - \frac{2(a+1)}{k+1} + \frac{2(a+1)}{k+1} + o\left(\frac{1}{(k+1)^2}\right) - 1\right| \\ &\leq \frac{L}{(k+1)^2}\end{aligned}$$

e per la proposizione precedente abbiamo così che $\prod_{i=0}^{+\infty}\left(1 + A_i^+\right) = M \in \mathbb{R}$ dove $A_i^+ = A_i \vee 0$.

Dimostriamo ora che, posto $M_0 = 1$ ed $M_{k+1} = \prod_{i=0}^{k}\left(1 + A_i^+\right)$,

$$\left|\Delta_k(\alpha)\right| \leq \frac{\sqrt{M_k}}{(k+1)^{1+a}}$$

essa vale chiaramente quando $k = 0$, se la supponiamo valida per qualche $k \geq 0$ allora

$$\left|\Delta_{k+1}(\alpha)\right| = \frac{|k-\alpha|}{k+1}\left|\Delta_k(\alpha)\right| \leq \left(\frac{k+1}{k+2}\right)^{1+a}\frac{\sqrt{\left(1+A_k\right)M_k}}{(k+1)^{1+a}} = \frac{\sqrt{M_{k+1}}}{(k+2)^{1+a}}$$

per induzione quindi la (8.3.2) vale per ogni $k \in \mathbb{N}_0$ in quanto $M_k \leq M$ per ogni $k \in \mathbb{N}_0$. ∎

Di conseguenza abbiamo che

$$\sum_{k=1}^{+\infty} \left|\Delta_k(\alpha)\right| \leq M \sum_{k=1}^{+\infty} \frac{1}{k^{1+\operatorname{Re}\alpha}}$$

e dato che $\operatorname{Re}\alpha > 0$ la serie al secondo membro converge, di conseguenza possiamo definire x^α per ogni x tale che $\|x - \mathbb{e}\| \leq 1$.

8.4 Spettro

Definizione 8.4.1. Siano dati un'algebra di Banach X ed un elemento $x \in X$, definiamo il **risolvente** di x è l'insieme

$$\rho(x) = \left\{\lambda \in \mathbb{K} \mid x - \lambda\mathbb{e} \in G(X)\right\}$$

mentre definiamo lo **spettro** di c come $\sigma(x) = \mathbb{K} \setminus \rho(x)$.

Proposizione 8.4.2. *Lo spettro di un elemento $x \in X$, se non vuoto, è compatto e*

$$\sigma(x) \subseteq \left\{\lambda \in \mathbb{K} \mid |\lambda| \leq \|x\|\right\}$$

Dimostrazione. Prendiamo un qualunque $|\lambda| > \|x\|$, allora $\left\|\lambda^{-1}x\right\| < 1$ e per quanto detto nella sezione precedente $\mathbb{e} - \lambda^{-1}x \in G(X)$ e perciò $\lambda \in \rho(x)$.

Dimostriamo ora che $\rho(x)$ è aperto in $\mathbb{K}$, da cui discende immediatamente la compattezza di $\sigma(x)$. Sia $\lambda \in \rho(x)$ allora $z = x - \lambda\mathbb{e} \in G(X)$ e dato che $G(X)$ è aperto esisterà $\varepsilon > 0$ tale che $B(z, \varepsilon) \subseteq G(X)$. Scelto $\mu \in \mathbb{K}$ con $\left|\lambda - \mu\right| < \varepsilon$ abbiamo che

$$\left\|(x - \lambda\mathbb{e}) - (x - \mu\mathbb{e})\right\| = \left|\lambda - \mu\right| \|\mathbb{e}\| < \varepsilon$$

e quindi $\mu \in \rho(x)$ e $\rho(x)$ è aperto. ∎

Se $\mathbb{K} = \mathbb{R}$ lo spettro di un elemento può essere vuoto, mentre per $\mathbb{K} = \mathbb{C}$ lo spettro conterrà sicuramente almeno un elemento. Per dimostrarlo però si richiede la conoscenza di concetti base dell'analisi complessa, in particolare il teorema di Liouville.

Teorema 8.4.3. *Se X è un'algebra di Banach unitaria su $\mathbb{C}$ allora $\sigma(x) \neq \emptyset$ per ogni $x \in X$.*

Dimostrazione. Per assurdo sia $x \in X$ tale che $\sigma(x) = \emptyset$ allora $x - \lambda\mathbb{e} \in G(X)$ per ogni $\lambda \in \mathbb{C}$. Consideriamo l'applicazione

$$T : \lambda \in \mathbb{C} \to (x - \lambda\mathbb{e})^{-1} \in X$$

la quale è chiaramente ben definita, dimostriamo che è una funzione differenziabile (olomorfa) su tutto $\mathbb{C}$ con derivata $(x - \lambda\mathbb{e})^{-2}$.

Abbiamo per ogni $\lambda, \mu \in \mathbb{C}$

$$[x - (\lambda + \mu)\mathbb{e}]^{-1} - (x - \lambda\mathbb{e})^{-1} = [x - (\lambda + \mu)\mathbb{e}]^{-1} \cdot \left(-\mu\mathbb{e}\right) \cdot (x - \lambda\mu\mathbb{e})^{-1}$$

e quindi

$$\lim_{\mu \to 0} \frac{[x - (\lambda + \mu)\mathbb{e}]^{-1} - (x - \lambda\mathbb{e})^{-1}}{\mu} = \lim_{\mu \to 0} [x - (\lambda + \mu)\mathbb{e}]^{-1} \cdot (x - \lambda\mu\mathbb{e})^{-1} = (x - \lambda\mathbb{e})^{-2}$$

in quanto per il teorema 8.3.10 l'inversa è continua. Ma allora per ogni $\varphi \in X^*$ anche la composizione $\varphi \circ T$ è olomorfa e quindi è una funzione intera. Ma preso $|\lambda| > \|x\|$ avremo che $T(x) = (-\lambda)^{-1} \sum_{i=0}^{+\infty} \left(\lambda^{-1} x\right)^i$ che tende a 0 al tendere di λ all'infinito, quindi $\varphi \circ T$ si annulla all'infinito. Per il teorema di Liouville $\varphi \circ T$ è identicamente nulla per ogni $\varphi \in X^*$ ovvero $(x - \lambda \mathbb{e})^{-1} = 0$ assurdo. ■

Corollario 8.4.4. *Se X è un'algebra di Banach unitaria su $\mathbb{C}$ tale che $G(X) = X \setminus \{0\}$ allora $X = \mathbb{C}\mathbb{e} = \left\{\lambda \mathbb{e} \mid \lambda \in \mathbb{C}\right\}$.*

Dimostrazione. Per ogni $x \in X$ per il teorema precedente esiste sempre $\lambda \in \sigma(x) \neq \emptyset$, ma l'unico elemento non invertibile in X è lo 0 e perciò $x = \lambda \mathbb{e}$. ■

Consideriamo adesso un'algebra di Banach unitaria X sul campo $\mathbb{K}$ e consideriamo il polinomio $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \mathbb{e}$ su X a coefficienti in $\mathbb{K}$.

Teorema 8.4.5. *Scelto x elemento di un'algebra di Banach unitaria X su $\mathbb{K}$ ed un polinomio $p(x)$ a coefficienti in $\mathbb{K}$ si ha*

$$p\,[\sigma(x)] \subseteq \sigma\left[p(x)\right]$$

e se $\mathbb{K} = \mathbb{C}$ allora l'inclusione di sopra è un'uguaglianza.

Dimostrazione. Prendiamo innanzitutto $\lambda \in p\,[\sigma(x)]$ allora esiste $\mu \in \mathbb{K}$ tale che $x - \mu \mathbb{e} \notin G(X)$ ed $p(\mu) = \lambda$. Ma allora esiste un polinomio q a coefficienti in $\mathbb{K}$ tale che $p(x) - \lambda \mathbb{e} = (x - \mu \mathbb{e}) q(x)$ e inoltre $x - \mu \mathbb{e}$ ed $q(x)$ commutano tra loro. Quindi se per assurdo fosse $\lambda \in \rho[p(x)]$ avremmo che $x - \mu \mathbb{e}$ è invertibile grazie alla proposizione 8.2.5 il che è assurdo, perciò $\lambda \in \sigma[p(x)]$.

Viceversa supponiamo $\mathbb{K} = \mathbb{C}$ e $\lambda \in \sigma[p(x)]$, per il teorema fondamentale dell'aritmetica esistono $\eta, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$ con $\eta \neq 0$ tali che

$$p(x) - \lambda \mathbb{e} = \eta (x - \lambda_1 \mathbb{e}) \cdots (x - \lambda_n \mathbb{e})$$

e quindi deve esistere almeno un i tale che $\lambda_i \in \sigma(x)$, in tal caso $p(\lambda_i) = \lambda$ e perciò $\lambda \in [\sigma(x)]$. ■

Fissato $x \in X$ poniamo per ogni $n \in \mathbb{N}$

$$a_n = \ln \left\|x^n\right\|$$

vogliamo utilizzare questa successione per stimare lo spettro di x con maggior precisione.

Proposizione 8.4.6. *Per ogni $m, n \in \mathbb{N}$*

$$a_{m+n} \leq a_m + a_n$$

Inoltre

$$\lim_{n \to +\infty} \frac{a_n}{n} = \inf_{n \in \mathbb{N}} \frac{a_n}{n}$$

Dimostrazione. La prima affermazione segue dalla disuguaglianza $\left\|x^{m+n}\right\| \leq \left\|x^m\right\| \left\|x^n\right\|$. Scegliamo ora un intero $m \in \mathbb{N}$ allora per ogni $n \in \mathbb{N}$ esistono e sono unici $q, r \in \mathbb{N}_0$ con $r < m$ tali che $n = mq + r$ e quindi

$$a_n = a_{mq+r} \leq q a_m + a_r \leq \frac{n}{m} a_m + (a_0 \vee a_1 \vee \cdots \vee a_{m-1})$$

Ma allora $\limsup_{n \to +\infty} (a_n / n) \leq a_m / m$ per ogni $m \in \mathbb{N}$ e quindi anche il secondo punto è verificato. ■

Una conseguenza immediata di questo risultato è che il limite $\lim_{n\to+\infty} \left\|x^n\right\|^{1/n}$ esiste finito ed è più piccolo di $\|x\|$. Tale quantità è il **raggio spettrale** del vettore x e lo indicheremo con $r(x)$.

Il raggio spettrale di un operatore permette di ottenere una stima migliore di quella ottenuta nella proposizione 8.4.2. Vale infatti

Proposizione 8.4.7. *Sia $x \in X$ allora*

$$\sigma(x) \subseteq \left\{\lambda \in \mathbb{K} \mid |\lambda| \le r(x)\right\}$$

Dimostrazione. Per ogni $n \in \mathbb{N}$ e per ogni $\lambda \in \sigma(x)$ sappiamo già che $\lambda^n \in \sigma\left(x^n\right)$ e quindi per la proposizione 8.4.2 si ha $|\lambda| \le \sqrt[n]{\|x^n\|}$. ■

Anche in questo caso se $\mathbb{K} = \mathbb{C}$ abbiamo un risultato più forte, sempre utilizzando la teoria delle funzioni olomorfe. Innanzitutto per ogni algebra di Banach unitaria X definiamo per ogni $x \in X$

$$t(x) = \begin{cases} \max\left\{|\lambda| \mid \lambda \in \sigma(x)\right\} & \text{se } \sigma(x) \neq \emptyset \\ -\infty & \text{se } \sigma(x) = \emptyset \end{cases}$$

La proposizione precedente ci permette di affermare che $t(x) \le r(x) \le \|x\|$ per ogni $x \in X$. Quando X è uno spazio complesso abbiamo una relazione molto più forte:

Teorema 8.4.8. *Se X è un'algebra di Banach unitaria su $\mathbb{C}$ allora $t(x) = r(x)$ per ogni $x \in X$.*

Dimostrazione. Dal teorema 8.4.3 si ha $t(x) \ge 0$ per ogni $x \in X$. Per ogni $\varphi \in X^*$ l'applicazione $g : \lambda \in \rho(x) \to \varphi\left[(x - \lambda \mathbb{e})^{-1}\right] \in \mathbb{C}$ è analitica sull'aperto $\mathbb{C}\backslash\sigma(x)$ il quale è anche un intorno dell'infinito, e dato che g si annulla all'infinito possiamo esprimerla in serie di Laurent su $\left\{|\lambda| > t(x)\right\}$. Ma per ogni $|\lambda| > \|x\|$ si ha

$$g(\lambda) = \sum_{i=0}^{+\infty} \frac{\varphi(x^n)}{\lambda^{n+1}} \tag{8.4.1}$$

e quindi, dato che l'espansione in serie di Laurent esiste, deve coincidere con la (8.4.1). Dunque $\varphi(x^n/\lambda^{n+1})$ converge a 0 al tendere di n all'infinito. Ma allora per il principio di uniforme limitatezza la successione x^n/λ^{n+1} è limitata in X e dunque $\left\|x^n\right\| \le C|\lambda|\,|\lambda|^n$ per qualche $C > 0$ e dunque

$$r(x) = \lim_{n\to+\infty} \sqrt[n]{\|x^n\|} \le |\lambda|$$

e perciò $r(x) \le t(x) \le r(x)$. ■

Proposizione 8.4.9. *Siano X, Y due algebre di Banach unitarie su $\mathbb{K}$ ed $f : X \to Y$ un'applicazione lineare ed unitaria, allora:*

- *se f è continua e moltiplicativa allora $r[f(x)] \le r(x)$;*
- *se f è unitaria allora $\sigma[f(x)] \subseteq \sigma(x)$, in particolare $t[f(x)] \le t(x)$ e se $\mathbb{K} = \mathbb{C}$ avremo anche che $r[f(x)] \le r(x)$.*

Dimostrazione. Sia $f : X \to Y$ lineare, continua e moltiplicativa, allora per ogni $n \in \mathbb{N}$ abbiamo che

$$\left\|f(x)^n\right\| = \left\|f\left(x^n\right)\right\| \le \|f\|\,\left\|x^n\right\|$$

calcolando la radice n-esima ad entrambi i membri e passando al limite avremo così che $r[f(x)] \leq r(x)$.

Supponiamo ora che f sia lineare ed unitaria, preso $\lambda \in \rho(x)$ avremo di conseguenza che $x - \lambda \mathbb{e} \in G(X)$. Di conseguenza anche $f(x - \lambda \mathbb{e}) = f(x) - \lambda \mathbb{e}$ è invertibile in Y ovvero $\lambda \in \rho[f(x)]$. Nel caso in cui $\mathbb{K} = \mathbb{C}$ ed f è solamente lineare ed unitaria la tesi segue dal teorema 8.4.8. ■

8.5 C^*-algebre

Definizione 8.5.1. Data un'algebra di Banach X su $\mathbb{K}$ diremo che un'applicazione $x \in X \to x^* \in X$ è una **involuzione** su X se e solo se valgono le seguenti relazioni per ogni $x, y \in X$, $\lambda \in \mathbb{K}$

- $(x + y)^* = x^* + y^*$;
- $(\lambda x)^* = \lambda^\dagger x^*$;
- $(xy)^* = y^* x^*$;
- $x^{**} = x$.

In tal caso la coppia $(X, *)$ è detta $*$**-algebra** su $\mathbb{K}$. Se $\mathbb{K} = \mathbb{C}$ e se l'involuzione soddisfa anche la seguente relazione

$$\left\| x^* x \right\| = \|x\|^2 \tag{8.5.1}$$

allora diremo che la coppia $(X, *)$ è una C^***-algebra**.

Il fatto di lavorare esclusivamente sugli spazi vettoriali complessi per quanto riguarda le C^*-algebre ci permette di utilizzare alcuni importanti risultati dell'analisi complessa, i quali risultano fondamentali per la gran parte dei risultati in questa sezione. Questi risultati infatti possono non valere all'interno di spazi vettoriali reali, i quali necessitano di una trattazione completamente diversa.

Se $\mathbb{e}$ è l'unità di X allora avremo che

$$\mathbb{e}^* = \mathbb{e}^* \mathbb{e} = \left(\mathbb{e}^* \mathbb{e} \right)^* = \mathbb{e}^{**} = \mathbb{e}$$

in generale se $x^* = x$ diremo che x è **autoaggiunto**, mentre se $x^* x = x x^*$ diremo che x è **normale**. Se invece $x^* = -x$ diremo che x è **antisimmetrico**. Chiaramente se $x \in X$ è autoaggiunto/antisimmetrico allora per ogni $\lambda \in \mathbb{R}$ anche λx lo è.

In una C^*-algebra dalla (8.5.1) segue che $\|x\|^2 \leq \left\| x^* \right\| \|x\|$ e quindi $\|x\| \leq \left\| x^* \right\|$, poiché $x^{**} = x$ avremo che

$$\|x\| = \left\| x^* \right\| \tag{8.5.2}$$

Se X ed Y sono due $*$-algebre allora l'applicazione $f : X \to Y$ si dice $*$**-compatibile** se e solo se $f\left(x^*\right) = f(x)^*$ per ogni $x \in X$. Un esempio di C^*-algebra è lo spazio $\mathbb{C}$ con l'operazione di involuzione $z^* = z^\dagger$, quindi se $f : X \to \mathbb{C}$ è $*$-compatibile allora $f\left(x^*\right) = f(x)^\dagger$. Definiamo di conseguenza gli $*$-omomorfismi tra $*$-algebre semplicemente come omomorfismi che sono anche $*$-compatibili.

Proposizione 8.5.2. *Data una $*$-algebra X sul campo $\mathbb{K}$ per ogni $x \in X$ esistono a autoaggiunto ed b antisimmetrico tali che $x = a + b$.*

Inoltre se $\mathbb{K} = \mathbb{C}$ allora $x \in X$ è antisimmetrico se e solo se ix è autoaggiunto.

Dimostrazione. Sia $x \in X$ generico e poniamo $a = \left(x + x^*\right)/2$, $b = \left(x - x^*\right)/2$. Chiaramente $x = a + b$ e $x^* = a - b$ mentre a è autoaggiunto e b è antisimmetrico. Supponiamo ora che $\mathbb{K} = \mathbb{C}$, per ogni $x \in X$ si ha

$$(ix)^* = -ix^*$$

e quindi x è antisimmetrico se e solo se ix è autoaggiunto. ■

Dalla (8.5.2) segue inoltre che $\left\|x^* x\right\| = \|x\|^2 = \left\|x^*\right\| \|x\|$. Osserviamo che per ogni $x \in X$ l'elemento $x^* x$ è chiaramente autoaggiunto in quanto $\left(x^* x\right)^* = x^* x^{**} = x^* x$, quindi la (8.5.1) ci dice anche che la norma di qualunque elemento di X coincide con la radice quadrata della norma di qualche elemento autoaggiunto.

Proposizione 8.5.3. *Se X è una C^*-algebra allora per ogni $x \in X$ normale si ha $r(x) = \|x\|$.*

Dimostrazione. Se x è autoaggiunto allora $\left\|x^2\right\| = \left\|x^* x\right\| = \|x\|^2$, dato che anche x^2 è autoaggiunto ragionando per induzione si ha che $\left\|x^{2^n}\right\| = \|x\|^{2^n}$. Dato che precedentemente abbiamo dimostrato che la successione $\sqrt[n]{\|x^n\|}$ è decrescente segue che quando x è autoaggiunto essa sarà costante e perciò $r(x) = \|x\|$.

Se invece x è normale allora $\left(x^* x\right)^n = \left(x^*\right)^n x^n$, di conseguenza si ha

$$r(x)^2 \leq \|x\|^2 = \left\|x^* x\right\| = \lim_{n\to+\infty} \sqrt[n]{\|(x^* x)^n\|} \leq \lim_{n\to+\infty} \sqrt[n]{\|(x^*)^n\| \|x^n\|} = r\left(x^*\right) r(x)$$

e sempre sfruttando la $x^{**} = x$ si ha $r(x) = r\left(x^*\right)$ e quindi $r(x) = \|x\|$. ■

Corollario 8.5.4. *Se X ed Y sono due C^*-algebre ed $f : X \to Y$ è uno $*$-omomorfismo allora $\left\|f\right\| \leq 1$. In particolare se X, Y ed f fossero anche unitari allora $\left\|f\right\| = 1$*

Dimostrazione. Se $x \in X$ è autoaggiunto allora anche $f(x)$ è autoaggiunto in quanto $f(x)^* = f\left(x^*\right) = f(x)$, per la proposizione 8.4.9 e per la proposizione precedente si ha

$$\left\|f(x)\right\| = r[f(x)] \leq r(x) = \|x\|$$

invece se x è un elemento generico di X per la (8.5.1)

$$\left\|f(x)\right\|^2 = \left\|f(x)^* f(x)\right\| = \left\|f\left(x^* x\right)\right\| \leq \left\|x^* x\right\| = \|x\|^2$$

e perciò $\left\|f\right\| \leq 1$. Per il caso unitario l'uguaglianza segue in quanto $f(\mathbb{e}) = \mathbb{e}$. ■

Si noti che non stiamo affatto dicendo che f sia una isometria, anzi possono tranquillamente esistere $*$-omomorfismi unitari non iniettivi.

Sempre applicando le proposizioni 8.4.9 ed 8.5.3 nella stessa maniera del corollario precedente abbiamo il seguente criterio di continuità:

Corollario 8.5.5. *Date due C^*-algebre X, Y unitarie allora ogni applicazione lineare, moltiplicativa, unitaria e $*$-compatibile è continua con norma uguale ad 1, in particolare è uno $*$-omomorfismo unitario.*

Combinando questo risultato con il teorema (8.4.8) avremo che

Teorema 8.5.6. *Se X è una C^*-algebra allora per ogni $x \in X$ si ha*

$$\|x\| = \sqrt{\|x^* x\|} = \sqrt{\sup\left\{|\lambda| \mid \lambda \in \sigma(x^* x)\right\}}$$

quindi la norma di una C^*-algebra può essere dedotta univocamente dal suo prodotto interno.

8.6 Algebre commutative

Abbiamo detto precedentemente che un'algebra di Banach è commutativa se e solo se $xy = yx$ per ogni coppia di elementi x e y. Dato uno spazio topologico compatto T possiamo rendere lo spazio $C(T, \mathbb{K})$ delle funzioni continue da X in $\mathbb{K}$ uno spazio vettoriale su $\mathbb{K}$ con le definizioni standard di somma e prodotto esterno. Grazie al corollario 3.6.5 lo spazio $C(T, \mathbb{K})$ è uno spazio di Banach con la norma

$$\|f\|_\infty = \max_{x \in T} |f(x)|$$

in quanto il massimo esiste per via del teorema 2.6.6 opportunamente adattato. D'altronde possiamo renderlo una C^*-algebra unitaria commutativa semplicemente ponendo

$$\begin{aligned} (f \cdot g)(x) &= f(x)g(x) \\ \mathbb{e}(x) &= 1 \qquad \forall x \in X \\ f^*(x) &= [f(x)]^\dagger \end{aligned}$$

in quanto $f^*(x)f(x) = |f(x)|^2$ per ogni $x \in T$ e quindi vale la (8.5.1).

Questa algebra è fondamentale per gran parte dei risultati sulle algebre di Banach, perché come vedremo dopo ogni C^*-algebra unitaria è isomorfa a $C(T, \mathbb{C})$ per qualche spazio topologico compatto di Hausdorff T. Un altro risultato famoso per le algebre $C(T, \mathbb{K})$ è il teorema di Stone-Weierstrass, utilizzato principalmente per dimostrare la separabilità dello spazio delle funzioni continue.

Innanzitutto se X è una $*$-algebra diremo che una sua sottoalgebra Y di X è anche **$*$-compatibile** se e solo se per ogni $y \in Y$ si ha anche $y^* \in Y$ e viceversa. Non richiediamo che sia chiusa necessariamente.

Teorema 8.6.1 (Stone-Weierstrass)**.** *Dato uno spazio topologico compatto T ed una sottoalgebra $*$-compatibile V di $C(T, \mathbb{K})$ (contenente l'unità). Se per ogni $x, y \in T$ distinti esiste $f \in V$ tale che $f(x) \neq f(y)$ allora V è denso in $C(T, \mathbb{K})$.*

Dimostrazione. Supponiamo inizialmente che $\mathbb{K} = \mathbb{R}$. Per ogni $x, y \in T$ con $x \neq y$ e per ogni $f \in C(T, \mathbb{R})$ possiamo sempre trovare $h_{x,y} \in V$ tale che $h_{x,y}(x) = f(x)$ ed $h_{x,y}(y) = f(y)$. Per ogni $\varepsilon > 0$ possiamo così definire

$$W_{x,y} = \left\{ z \in T \,\middle|\, h_{x,y}(z) < f(z) + \varepsilon \right\}$$

essi saranno chiaramente aperti e conterranno sia x che y. Quindi la famiglia di insiemi

$$\left\{ W_{x,y} \,\middle|\, y \in T \setminus \{x\} \right\}$$

è un ricoprimento di T per ogni $x \in T$, per compattezza esisteranno $y_1, y_2, \dots, y_k \in X$ diversi da x tali che $\bigcup_{i=1}^k W_{x,y_i} = T$. Dimostriamo ora che la nuova funzione

$$g_x(z) = h_{x,y_1}(z) \wedge \cdots \wedge h_{x,y_k}(z)$$

appartiene a $\mathrm{Cl}(V)$.

Per la continuità del prodotto interno abbiamo che $\mathrm{Cl}(V)$ è un'algebra di Banach, quindi per ogni $f \in \mathrm{Cl}(V)$ abbiamo chiaramente che $f^2 \in \mathrm{Cl}(V)$ ed $f^2(x) \geq 0$ per ogni $x \in T$. Questo significa quindi che

$$\left\| \frac{f^2}{\left\|f^2\right\|_\infty} - 1 \right\| \leq 1$$

e per quanto detto precedentemente sulle serie di potenze possiamo definire la radice quadrata $\sqrt{f^2 / \left\|f^2\right\|_\infty}$ che apparterrà a $\mathrm{Cl}(V)$ in quanto l'abbiamo definita tramite una serie di potenze. Ma f è una funzione a valori in $\mathbb{R}$ e quindi, utilizzando stavolta la serie di Taylor della radice quadrata in $\mathbb{R}$

$$\sqrt{\frac{f^2}{\left\|f^2\right\|_\infty}}(x) = \sqrt{\frac{f(x)^2}{\left\|f^2\right\|_\infty}} = \frac{\left|f(x)\right|}{\left\|f\right\|_\infty}$$

e perciò abbiamo che $\left|f\right| \in \mathrm{Cl}(V)$. Dato che $f \wedge g = f - (f - g + \left|f - g\right|)/2$ abbiamo anche che $f \wedge g \in \mathrm{Cl}(V)$ e analogamente si vede per $f \vee g$.

Tornando alla funzione g_x osserviamo innanzitutto che $g_x(x) = f(x)$, poi per ogni $z \in T$ esiste un i tale che $z \in W_{x,y_i}$ e dunque

$$g_x(z) \leq h_{x,y_i}(z) < f(z) + \varepsilon$$

Definiamo per ogni $x \in T$ il seguente insieme aperto non vuoto

$$W_x = \left\{ z \in T \,\middle|\, g_x(z) > f(z) - \varepsilon \right\}$$

che chiaramente contiene x, sempre per compattezza esisteranno $x_1, \ldots, x_l \in T$ tali che gli W_{x_j} ricoprano T. Definiamo infine

$$g(z) = g_{x_1}(z) \vee \cdots \vee g_{x_l}(z)$$

per lo stesso motivo di prima $g \in \mathrm{Cl}(V)$ e inoltre si ha $\left|g(z) - f(z)\right| < \varepsilon$ per ogni $z \in T$. Ma dato che $\mathrm{Cl}(V)$ è chiuso per l'arbitrarietà di ε segue che $f \in \mathrm{Cl}(V)$ e quindi $\mathrm{Cl}(V) = C(T, \mathbb{R})$.

Supponiamo adesso che $\mathbb{K} = \mathbb{C}$. Posto

$$V_1 = \left\{ \mathrm{Re} f \,\middle|\, f \in V \right\}$$

abbiamo che $V_1 \leq C(T, \mathbb{R})$ e inoltre poiché per ogni $v, w \in \mathbb{C}$

$$\mathrm{Im}\, v = -\mathrm{Re}(i v)$$
$$\mathrm{Re}\, v\, \mathrm{Re}\, w - \mathrm{Im}\, v\, \mathrm{Im}\, w = \frac{1}{2} \mathrm{Re}\left(v w + v^\dagger w^\dagger \right)$$

segue che V_1 è un'algebra di Banach reale contenuta in $C(T, \mathbb{R})$. Inoltre per ogni $x, y \in T$ distinti esiste $f \in V$ tale che $f(x) \neq f(y)$, in altre parole o $\mathrm{Re} f(x) \neq \mathrm{Re} f(y)$ oppure $\mathrm{Im} f(x) \neq \mathrm{Im} f(y)$, nel secondo caso al posto di prendere f si prende $if \in V$.

Quindi V_1 soddisfa tutti i punti dell'enunciato, per quanto detto prima allora V_1 è denso in $C(T, \mathbb{R})$. Quindi per ogni $f \in C(T, \mathbb{C})$ e per ogni $\varepsilon > 0$ esistono $g_1, g_2 \in V_1$ tali che

$$\left\| \mathrm{Re} f - g_1 \right\|_\infty \leq \varepsilon$$
$$\left\| \mathrm{Im} f - g_2 \right\|_\infty \leq \varepsilon$$

Ma essendo $\mathrm{Re} f = (f + f^*)/2$ segue che $V_1 \subseteq V$ e perciò posto $g = g_1 + i g_2 \in V$ si ha la tesi. ■

Consideriamo ora una generica algebra di Banach X sul campo $\mathbb{K}$, diremo che l'applicazione $f : X \to \mathbb{K}$ è un **carattere** di X se e solo se è lineare, moltiplicativa e non identicamente nulla. Indicheremo con il simbolo $\operatorname{chr}(X)$ l'insieme dei caratteri di X.

Proposizione 8.6.2. *Per ogni $f \in \operatorname{chr}(X)$ si ha:*

1. $\ker f$ *è un ideale proprio di X;*
2. $f(\mathbb{e}) = 1$*;*
3. $f(x) \in \sigma(x)$ *per ogni $x \in X$;*
4. $f \in X^*$ *con* $\|f\| = 1$.

In particolare i caratteri sono omomorfismi di algebre.

Dimostrazione. Dimostriamo i vari punti:

1. Per ogni $y \in \ker f$ e per ogni $x \in X$ abbiamo che $f(xy) = f(yx) = f(x)f(y) = 0$ e perciò $xy, yx \in \ker f$ ed il kernel di f è dunque un ideale di X, poiché f non è identicamente nulla avremo che $\ker f \neq X$ e quindi è proprio.
2. Dato che f non è identicamente nullo esisterà $x \in X$ tale che $f(x) \neq 0$, allora avremo che
$$f(x) = f(x\mathbb{e}) = f(x)f(\mathbb{e})$$
e poiché possiamo dividere per $f(x)$ segue che $f(\mathbb{e}) = 1$.
3. Essendo $\ker f$ un ideale proprio di X abbiamo visto che non può contenere elementi invertibili. Ma $x - f(x)\mathbb{e}$ è chiaramente un elemento di $\ker f$ per quanto visto poco fa, dunque non può essere invertibile ed $f(x) \in \sigma(x)$.
4. Per quanto riguarda la continuità di f per la proposizione 8.4.2 che $\left|f(x)\right| \leq \|x\|$ e perciò $\|f\| \leq 1$. L'uguaglianza si ottiene valutando f in $\mathbb{e}$.

■

Da questo risultato segue che se $\mathbb{K} = \mathbb{R}$ lo spazio $\operatorname{chr}(X)$ può essere vuoto. Infatti in un'algebra di Banach reale quasi sempre esistono elementi il cui spettro è vuoto, e ciò implica per la proposizione precedente la non esistenza di alcun carattere per X. Come vedremo più avanti per le algebre di Banach complesse esisterà sempre almeno un carattere.

Proposizione 8.6.3. *Siano X, Y due algebre di Banach su $\mathbb{K}$ ed $f : X \to Y$ omomorfismo unitario di algebre. Allora per ogni $x \in X$*

$$f(e^x) = e^{f(x)}$$

Inoltre se X è una $*$-algebra ed $x \in X$ è autoaggiunto allora $\left\|e^{ix}\right\| = 1$.

Dimostrazione. Per la definizione di omomorfismo di algebre e di esponenziale si ha

$$f\left(e^x\right) = \sum_{i=0}^{+\infty} \frac{f\left(x^n\right)}{n!} = \sum_{i=0}^{+\infty} \frac{[f(x)]^n}{n!} = e^{f(x)}$$

Per la continuità e l'additività dell'involuzione per ogni $x \in X$ autoaggiunto si ha

$$\left(e^{ix}\right)^* = \sum_{n=0}^{+\infty} \left[\frac{(ix)^n}{n!}\right]^* = \sum_{n=0}^{+\infty} \frac{[-ix^*]^n}{n!} = e^{-ix}$$
$$\left\|e^{ix}\right\|^2 = \left\|e^{ix}e^{-ix}\right\| = \|\mathbb{e}\| = 1$$

e quindi ha norma unitaria. ■

Lemma 8.6.4. *Se X è una C^*-algebra allora per ogni carattere $f \in \operatorname{chr}(X)$ e per ogni $x \in X$ si ha*

$$f\left(x^*\right) = [f(x)]^\dagger$$

e quindi i caratteri sono anche $*$-omomorfismi.

Dimostrazione. Dimostriamo innanzitutto che $f(x) \in \mathbb{R}$ per ogni $x \in X$ autoaggiunto, per ogni $t \in \mathbb{R}$ consideriamo la quantità e^{itx}, per la proposizione 8.6.3 ci dice che $\left\|e^{itx}\right\| = 1$. La proposizione 8.6.2 ci dice che f è anche continua su X, dunque $\left\|e^{itf(x)}\right\| = \left\|f(e^{itx})\right\| \leq 1$. Ma $e^{itf(x)}$ è l'esponenziale nel senso classico, che ha norma limitata per ogni $t \in \mathbb{R}$ se e solo se $f(x) \in \mathbb{R}$.

Sia ora $x \in X$ generico, allora esistono $a, b \in X$ autoaggiunti tali che $x = a + ib$ e perciò

$$f(x) = f(a) + if(b)$$
$$f(x^*) = f(a) - if(b)$$

ed essendo $f(a), f(b) \in \mathbb{R}$ si ha $f(x^*) = [f(x)]^\dagger$. ■

Lemma 8.6.5. *Se X è un'algebra di Banach unitaria complessa commutativa allora* $\operatorname{chr}(X) \neq \emptyset$.

Dimostrazione. Per il teorema 8.4.3 abbiamo $\sigma(x) \neq \emptyset$ per ogni $x \in X$. Dimostreremo più nello specifico che per ogni $x \in X$ e per ogni $\lambda \in \sigma(x)$ esiste un'applicazione $f \in \operatorname{chr}(X)$ tale che

$$f(x) = \lambda \in \mathbb{C}$$

Dato che $x - \lambda\mathbb{e}$ non è invertibile ed X è un'algebra commutativa per la proposizione 8.2.12 e per il lemma di Krull esso è contenuto in qualche ideale chiuso massimale I. Essendo I chiuso lo spazio quoziente X/I è uno spazio di Banach con la norma quoziente, mentre il prodotto interno di X può essere trasportato su X/I in quanto

$$(a + I) \cdot (b + I) = ab + aI + bI + I \cdot I = ab + I$$

e quindi X/I è un'algebra di Banach commutativa. Per la massimalità di I gli unici ideali di X/I sono $\{I\}$ ed X/I. Quindi per la proposizione 8.2.12 tutti i suoi elementi non nulli sono invertibili, e il corollario 8.4.4 ci dice che per ogni $a \in X$ esiste un unico $\lambda_a \in \mathbb{C}$ tale che

$$a + I = \lambda_a\mathbb{e} + I$$

Possiamo così definire l'applicazione $f : a \in X \to \lambda_a \in \mathbb{C}$ che risulta essere chiaramente un carattere in quanto $f(\mathbb{e}) = 1$. Infine poiché $x - \lambda\mathbb{e} \in I$ avremo che $x + I = \lambda\mathbb{e} + I$ e perciò $\lambda_x = \lambda$. ■

Corollario 8.6.6. *Se X è un'algebra di Banach commutativa unitaria su $\mathbb{C}$ allora per ogni $x \in X$ l'applicazione*
$$\ddot{x} : f \in \operatorname{chr}(X) \to f(x) \in \sigma(x)$$
è suriettiva e continua se dotiamo $\operatorname{chr}(X) \leq X^*$ *della topologia debole*. Inoltre* $\operatorname{chr}(X)$ *è uno spazio topologico compatto di Hausdorff.*

Dimostrazione. Che l'applicazione sia continua discende dalla definizione di topologia debole*, mentre il lemma precedente ne garantisce la suriettività. Se dimostriamo che $\operatorname{chr}(X)$ è debolmente* chiuso in X^* per il teorema di Banach-Alaoglu e per la proposizione 8.6.2 segue che $\operatorname{chr}(X)$ è anche debolmente* compatto.

Prendiamo una successione di funzioni $f_n \in \operatorname{chr}(X)$ convergente debolmente* a qualche $f \in X^*$, ma allora per ogni $a, b \in X$

$$f(ab) = \lim_{n\to+\infty} f_n(ab) = \lim_{n\to+\infty} f_n(a) f_n(b) = f(a) f(b)$$

e quindi f è moltiplicativa. Per far vedere che f non è identicamente nulla (cosa in generale non sempre vera per la convergenza debole) ci basta osservare che per la proposizione 8.6.2 $f_n(\mathbb{e}) = 1$ e quindi $f(\mathbb{e}) = 1$. Dunque $f \in \operatorname{chr}(X)$ e $\operatorname{chr}(X)$ è debolmente* chiuso. ■

In altre parole abbiamo $\ddot{x} \in C(\operatorname{chr}(X), \mathbb{C})$ per ogni $x \in X$. Per come abbiamo definito $\operatorname{chr}(X)$ avremo che l'applicazione

$$T : x \in X \to \ddot{x} \in C(\operatorname{chr}(X), \mathbb{C})$$

è lineare, moltiplicativa ed $\ddot{\mathbb{e}}$ coincide con la funzione identicamente 1 su $\operatorname{chr}(X)$ la quale è anche l'unità di $C(\operatorname{chr}(X), \mathbb{C})$. Inoltre se X è anche una C^*-algebra allora $\ddot{x^*} = \ddot{x}^*$ per ogni x. Enunciamo allora il risultato fondamentale di questa sezione:

Teorema 8.6.7. *Sia X un'algebra di Banach commutativa, unitaria e complessa, allora esiste un omomorfismo unitario T da X in $C(\operatorname{chr}(X), \mathbb{C})$.*

Inoltre se X è anche una C^*-algebra allora tale applicazione T è anche uno $*$-isomorfismo.

Dimostrazione. Dimostriamo prima la continuità di T. Per ogni $f \in \operatorname{chr}(X)$ sappiamo che $\|f\| = 1$ e quindi per ogni $x \in X$

$$\|\ddot{x}\|_\infty = \max_{f \in \operatorname{chr}(X)} |\ddot{x}(f)| = \max_{f \in \operatorname{chr}(X)} |f(x)| \leq \|x\|$$

Sia ora X una C^*-algebra, mostriamo prima che T è una isometria. Se prendiamo $x \in X$ autoaggiunto per la suriettività di $\ddot{x}$ su $\sigma(x)$ avremo che

$$\|x\| = r(x) = \sup\{|\lambda| \mid \lambda \in \sigma(x)\} = \sup\left\{|\ddot{x}(f)| \;\middle|\; f \in \operatorname{chr}(X)\right\} = \|\ddot{x}\|_\infty$$

se $x \in X$ invece è un elemento generico allora per definizione di C^*-algebra

$$\|x\|^2 = \|x^* x\| = \|T(x^* x)\| = \|T(x)^* T(x)\| = \|T(x)\|^2$$

e quindi per linearità T è una isometria.

Definiamo adesso $Y = T(X) \leq C(\operatorname{chr}(X), \mathbb{C})$, per quanto detto prima per ogni $x, y \in Y$ avremo che $xy = yx \in Y$ ed $x^* \in Y$, quindi per applicare il teorema di Stone-Weierstrass dobbiamo solamente verificare che Y separa i punti di $\mathrm{chr}(X)$. Presi $f, g \in \mathrm{chr}(X)$ distinti esiste chiaramente $x \in X$ tale che $f(x) \neq g(x)$ ovvero $\ddot{x}(f) \neq \ddot{x}(g)$ e chiaramente $\ddot{x} \in Y$.

Il teorema di Stone-Weierstrass ci garantisce così che Y è denso in $C(\mathrm{chr}(X), \mathbb{C})$, ma T è una isometria lineare biettiva da X in Y, ed essendo X completo anche Y deve essere uno spazio di Banach, quindi Y è chiuso in $C(\mathrm{chr}(X), \mathbb{C})$ e perciò T è suriettiva. ∎

Ora che abbiamo un risultato di struttura per C^*-algebre commutative unitarie proveremo ad utilizzare alcuni risultati in algebre non commutative. Ricordiamo che $\overline{\mathbb{C}}(x)$ è la più piccola algebra di Banach unitaria contenuta in X che a sua volta contiene x, la quale è chiaramente commutativa. Se però x fosse un elemento normale di una C^*-algebra unitaria X, ovvero $x^*x = xx^*$, allora

$$\mathbb{C}(\mathbb{e}, x, x^*) = \left\{ p\left(x, x^*\right) \;\middle|\; p \in \mathbb{C}[2] \right\} \tag{8.6.1}$$

è una sottoalgebra complessa unitaria commutativa in X contenente x ed x^*, per la linearità e continuità dell'involuzione anche $\overline{\mathbb{C}}(\mathbb{e}, x, x^*)$ è una C^*-algebra contenuta in X.

Osserviamo però che se restringiamo l'algebra di Banach di riferimento X lo spettro di un suo elemento può variare, poiché la nuova sottoalgebra potrebbe non contenere gli inversi di alcuni suoi elementi che invece stanno in X. Se però la nuova sottoalgebra è commutativa e chiusa possiamo dormire tranquilli:

Proposizione 8.6.8. *Sia X una C^*-algebra unitaria ed $Y \leq X$ una sua sottoalgebra chiusa, unitaria, $*$-compatibile e commutativa, allora $G(Y) = G(X) \cap Y$ ovvero l'inversa degli elementi in Y invertibili su X appartiene ancora ad Y. In particolare lo spettro degli elementi di Y rimane lo stesso se calcolato in X oppure in Y.*

Dimostrazione. Chiaramente $G(Y) \subseteq G(X) \cap Y$, sia allora $y \in Y$ invertibile con inversa $y^{-1} \in X$. Allora anche $y^* \in Y$ è invertibile e possiede inversa in X e perciò $y^*y \in G(X) \cap Y$. Il corollario 8.6.6 e il lemma 8.6.4 garantiscono che

$$\begin{aligned}
\sigma_Y(y^*y) &= \left\{ f(y^*y) \;\middle|\; f \in \mathrm{chr}(Y) \right\} \\
&= \left\{ f(y^*)f(y) \;\middle|\; f \in \mathrm{chr}(Y) \right\} \\
&= \left\{ [f(y)]^\dagger f(y) \;\middle|\; f \in \mathrm{chr}(Y) \right\} \\
&= \left\{ |f(y)|^2 \;\middle|\; f \in \mathrm{chr}(Y) \right\}
\end{aligned}$$

dove con $\sigma_Y(z)$ indichiamo lo spettro di z calcolato in Y, in particolare si ha

$$\sigma(y^*y) \subseteq \sigma_Y(y^*y) \subseteq \left[0, \|y^*y\|\right]$$

Dato che $y^*y \in G(X)$ per la compattezza dello spettro (che sappiamo non essere vuoto) esiste $\varepsilon > 0$ tale che $\lambda \geq \varepsilon$ per ogni $\lambda \in \sigma(y^*y)$.

Adesso per ogni $\lambda \geq 0$ abbiamo

$$(\|y^*y\| \mathbb{e} - y^*y) - \lambda \mathbb{e} = -\left[y^*y - \left(\|y^*y\| - \lambda\right) \mathbb{e} \right]$$

e quindi se $\lambda \in \sigma\left(\|y^*y\| \mathbb{e} - y^*y\right)$ allora $\|y^*y\| - \lambda \in \sigma(y^*y)$ e quindi $0 \leq \lambda \leq \|y^*y\| - \varepsilon$. Ma $\|y^*y\| \mathbb{e} - y^*y$ è autoaggiunto, dunque

$$\left\| \|y^*y\| \mathbb{e} - y^*y \right\| = r\left(\|y^*y\| \mathbb{e} - y^*y\right) \leq \|y^*y\| - \varepsilon$$

e perciò $\|y^*y\| - |f(y)|^2 = f\left(\|y^*y\| \mathbb{e} - y^*y\right) < \|y^*y\|$ per ogni $f \in \mathrm{chr}(Y)$. Abbiamo così che $f(y) \neq 0$ per ogni $f \in \mathrm{chr}(Y)$ e perciò $0 \notin \sigma_Y(y)$ ovvero $y \in G(Y)$. ∎

Con queste premesse possiamo rafforzare il corollario 8.6.6 restringendo opportunamente il nostro insieme chr (X):

Teorema 8.6.9. *Data una C^*-algebra unitaria e commutativa X per ogni $x \in X$ poniamo per comodità di notazione*

$$C_x = \overline{\mathbb{C}}\left(\mathbb{e}, x, x^*\right)$$

allora l'applicazione $\ddot{x} : f \in \operatorname{chr}\left(C_x\right) \to f(x) \in \sigma(x)$ è un omeomorfismo tra spazi topologici.

Dimostrazione. Sappiamo già che $\ddot{x}$ è continua e suriettiva grazie al corollario 8.6.6 sostituendo X con C_x. Prendiamo ora $f, g \in \operatorname{chr}\left(C_x\right)$ tali che $f(x) = g(x)$, allora $f(x^*) = g(x^*)$ ed $f(\mathbb{e}) = g(\mathbb{e}) = 1$. Per la (8.6.1) f e g coincidono su $\mathbb{C}\left(\mathbb{e}, x, x^*\right)$ e per continuità dovranno coincidere anche su $C_x = \operatorname{Cl}\left(\mathbb{C}\left(\mathbb{e}, x, x^*\right)\right)$.

Consideriamo adesso una successione di funzioni $f_n \in \operatorname{chr}\left(C_x\right)$ tali che $f_n(x) \to f(x)$ per qualche $f \in \operatorname{chr}\left(C_x\right)$, essendo $f(x^*) = [f(x)]^\dagger$ avremo che $f_n(x^*) \to f(x^*)$ e quindi $f_n(y) \to f(y)$ per ogni $y \in \mathbb{C}\left(\mathbb{e}, x, x^*\right)$.

Per ogni $y \in C_x$ e per ogni $\varepsilon > 0$ esiste $y' \in \mathbb{C}\left(\mathbb{e}, x, x^*\right)$, tale che $\left\| y - y' \right\| < \varepsilon/2$. Poiché $f_n(y') \to f(y')$ per la proposizione 8.6.2 si avrà

$$\begin{aligned}
\lim_{n\to+\infty} \left\| f_n(y) - f(y) \right\| &\leq \lim_{n\to+\infty} \left[\left\| f_n(y) - f_n(y') \right\| + \left\| f_n(y') - f(y') \right\| + \left\| f(y') - f(y) \right\| \right] \\
&\leq 2 \left\| y - y' \right\| + \lim_{n\to+\infty} \left\| f_n(y') - f(y') \right\| \\
&\leq \varepsilon
\end{aligned}$$

e perciò $f_n(y) \to f(y)$ e quindi anche l'inversa di $\ddot{x}$ è continua. ■

Ora se X è una C^*-algebra unitaria non necessariamente commutativa se $x \in X$ è normale allora $C_x = \overline{\mathbb{C}}(\mathbb{e}, x, x^*)$ ed è una sottoalgebra commutativa. Quindi per il teorema precedente l'applicazione

$$\mathscr{L}_x : f \in C\left(\operatorname{chr}\left(C_x\right), \mathbb{C}\right) \to f \circ \ddot{x}^{-1} \in C(\sigma(x), \mathbb{C})$$

è uno $*$-isomorfismo tra C^*-algebre. Il teorema 8.6.7 diventa così

Teorema 8.6.10. *Se X è una C^*-algebra unitaria ed $x \in X$ è un elemento normale di X allora l'applicazione $T_x = \mathscr{L}_x \circ T$ è uno $*$-isomorfismo da $\overline{\mathbb{C}}(\mathbb{e}, x, x^*)$ in $C[\sigma(x), \mathbb{C}]$ tale che*

$$\begin{aligned}
T_x(\mathbb{e})(\lambda) &= 1 \\
T_x(x)(\lambda) &= \lambda \\
T_x(x^*)(\lambda) &= \lambda^\dagger
\end{aligned}$$

Vogliamo studiare più nel dettaglio la struttura di T_x e della sua inversa. Per ogni $y \in \overline{\mathbb{C}}(\mathbb{e}, x, x^*)$ e per ogni $f \in \operatorname{chr}\left(\overline{\mathbb{C}}(\mathbb{e}, x, x^*)\right)$ abbiamo quindi

$$T_x(y)[f(x)] = T(y)(f) = f(y)$$

se poniamo nell'uguaglianza precedente $y = T_x^{-1}(\phi)$ per qualche $\phi \in C[\sigma(x), \mathbb{C}]$ avremo di conseguenza

$$\phi[f(x)] = f\left[T_x^{-1}(\phi)\right] \tag{8.6.2}$$

Osserviamo inoltre che se $p \in \mathbb{C}[2]$ è un polinomio in 2 incognite allora per ogni $x \in X$ normale si ha

$$T_x\left[p\left(x,x^*\right)\right](\lambda) = p\left[T_x(x),T_x(x^*)\right](\lambda) = p\left(\lambda,\lambda^\dagger\right)$$

quindi la funzione T_x prende un elemento di $\mathbb{C}\left(x^*,x\right)$ e gli associa il polinomio che lo ha generato. Di conseguenza se $\phi(\lambda) = p\left(\lambda,\lambda^\dagger\right)$ per qualche $p \in \mathbb{C}[2]$ allora

$$T_x^{-1}(\phi) = p\left(x,x^*\right)$$

Per questa ragione è lecito porre, per ogni $x \in X$ normale e per ogni $\phi \in C\left(\sigma(x),\mathbb{C}\right)$

$$\phi\{x\} = T_x^{-1}(\phi) \in \overline{\mathbb{C}}\left(\mathbb{e},x,x^*\right)$$

e la (8.6.2) assumerà la forma

$$\phi[f(x)] = f\left[\phi\{x\}\right] \tag{8.6.3}$$

per ogni $x \in X$, $\phi \in C[\sigma(x),\mathbb{C}]$ e $f \in \operatorname{chr}\left(\overline{\mathbb{C}}(\mathbb{e},x,x^*)\right)$.

Quindi ogni applicazione continua ϕ da $\mathbb{C}$ in sé induce una nuova applicazione continua $x \to \phi\{x\}$ definita sugli elementi normali x di X il cui valore che assume in x dipende solamente dai valori assunti da ϕ su $\sigma(x)$. Inoltre essendo T_x^{-1} uno $*$-isomorfismo per ogni $p \in \mathbb{C}[m]$ avremo anche che

$$p(\phi_1,\phi_2,\dots,\phi_m)\{x\} = p[\phi_1\{x\},\phi_2\{x\},\dots,\phi_m\{x\}]$$

ed in particolare $\phi^2\{x\} = \phi\{x\}^2$ e $\left|\phi\right|^2\{x\} = \phi\{x\}\phi\{x\}^*$.

Da questo teorema discende immediatamente un'importante caratterizzazione dello spettro degli elementi autoaggiunti:

Corollario 8.6.11. *Se X è una C^*-algebra unitaria ed $x \in X$ è normale allora x è autoaggiunto se e solo se $\sigma(x) \subseteq \mathbb{R}$.*

Dimostrazione. Se x è autoaggiunto allora abbiamo $\lambda = T_x(x)(\lambda) = T_x(x^*)(\lambda) = \lambda^\dagger$ per ogni $\lambda \in \sigma(x)$ e quindi $\lambda \in \mathbb{R}$. Viceversa se x è normale e $\sigma(x) \subseteq \mathbb{R}$ allora $T_x(x) = T_x(x^*)$ e per l'iniettività di T_x segue che $x = x^*$. ■

Lemma 8.6.12. *Date due C^*-algebre unitarie X ed Y ed uno $*$-omomorfismo unitario $f : X \to Y$. Allora per ogni $x \in X$ normale e per ogni $\xi \in C[\sigma(x),\mathbb{C}]$ anche $f(x)$ è normale e*

$$f(\xi\{x\}) = \xi\left\{f(x)\right\}$$

Inoltre se f è uno $*$-omomorfismo iniettivo unitario allora $\sigma[f(x)] = \sigma(x)$.

Dimostrazione. Poniamo

$$X' = \overline{\mathbb{C}}(\mathbb{e},x,x^*) \leq X$$
$$Y' = \overline{\mathbb{C}}[\mathbb{e},f(x),f(x^*)] \leq Y$$

chiaramente $f(X') \subseteq Y'$ ed T_x^{-1} manda $C[\sigma(x),\mathbb{C}]$ in X' mentre $T_{f(x)}^{-1}$ manda $C[\sigma[f(x)],\mathbb{C}]$ in Y'. Se prendiamo $\psi \in \operatorname{chr}\left(Y'\right)$ allora $\psi \circ f \in \operatorname{chr}\left(X'\right)$ e la (8.6.3) ci dice che

$$\psi\left[T_{f(x)}^{-1}(\xi)\right] = \xi[\psi[f(x)]] = (\psi\circ f)\left[T_x^{-1}(\xi)\right]$$

per ogni ψ. Ma il teorema 8.6.7 applicato alla C^*-algebra Y' ci dice che l'applicazione $T : x \to \ddot{x}$ è iniettiva, perciò $f[T_x^{-1}(\xi)] = T_{f(x)}^{-1}(\xi)$ come volevasi dimostrare.

L'inclusione $\subseteq$ discende dalla proposizione 8.4.9. Supponiamo adesso che f sia iniettiva e per assurdo esiste $\lambda \in \sigma(x) \setminus \sigma[f(x)] \subseteq \mathbb{C}$, poiché lo spettro è compatto possiamo sempre trovare $\phi \in C[\sigma(x), \mathbb{C}]$ che vale 0 su $\sigma[f(x)]$ ed 1 in λ. Dato che ϕ non è identicamente nullo su $\sigma(x)$ avremo che $z = \psi\{x\} \neq 0$ e $\psi\{f(x)\} = 0$, però allora $f(z) = 0$ il che va in contraddizione con l'iniettività di f. ■

Corollario 8.6.13. *Data una C^*-algebra unitaria X ed $Y \leq X$ una sottoalgebra chiusa $*$-compatibile unitaria. Per ogni $x \in Y$ si ha*

$$\sigma_Y(x) = \sigma(x)$$

Dimostrazione. Se x fosse normale allora la tesi seguirebbe dal teorema di sopra applicato all'omomorfismo $x \in Y \to x \in X$. Sia allora $z \in G(X) \cap Y$, avremo che $z^*z, zz^* \in G(Y)$ essendo entrambi autoaggiunti e quindi normali. Ma allora

$$z\left[z^*\left(zz^*\right)^{-1}\right] = \mathbb{e}$$
$$\left[\left(z^*z\right)^{-1}z^*\right]z = \mathbb{e}$$

perciò z è invertibile anche in Y e dunque $G(Y) = G(X) \cap Y$. ■

Corollario 8.6.14. *Date due C^*-algebre unitarie X ed Y allora ogni $*$-omomorfismo unitario iniettivo da X in Y è anche una immersione.*

Dimostrazione. Dobbiamo solamente verificare che f è una isometria. Sappiamo già che $\|f\| = 1$ e il lemma precedente ci dice che quando x è autoaggiunto si ha

$$\|f(x)\| = r[f(x)] = r(x) = \|x\|$$

Di conseguenza se x è generico allora

$$\|f(x)\|^2 = \left\|f\left(x^*x\right)\right\| = \left\|x^*x\right\| = \|x\|^2$$

e quindi f è anche una isometria. ■

Corollario 8.6.15. *Per ogni $x \in X$ elemento normale di una C^*-algebra unitaria X e per ogni $f \in C[\sigma(x), \mathbb{C}]$ vale l'uguaglianza*

$$\sigma[f\{x\}] = f[\sigma(x)]$$

Dimostrazione. Sappiamo che una funzione $F \in C(K, \mathbb{C})$ con K compatto di Hausdorff è invertibile se e solo se $0 \notin F(K)$ e quindi $\sigma(F) = F(K)$.

Ora T_x^{-1} è uno $*$-isomorfismo unitario da $C[\sigma(x), \mathbb{C}]$ in $Y = \overline{\mathbb{C}}(\mathbb{e}, x, x^*)$ e quindi per il lemma precedente

$$\sigma_Y(f\{x\}) = \sigma_Y\left(T_x^{-1}(f)\right) = \sigma(f) = f[\sigma(x)]$$

e la tesi segue dalla proposizione 8.6.8. ■

8.7 Positività

Abbiamo concluso il capitolo definendo a partire da una qualche applicazione continua $\phi \in C(\mathbb{C}, \mathbb{C})$ un'altra applicazione $x \in X \to \phi\{x\} \in \overline{\mathbb{C}}(e, x, x^*)$ con X una C^*-algebra. Vogliamo ora dare un'applicazione di questo fatto approfondendo lo studio delle radici quadrate di elementi in una C^*-algebra. Nel capitolo sulle serie di potenze abbiamo visto che preso $x \in X$ se esiste $t > 0$ per cui $\|te - x\| \leq t$ allora x ammette una radice quadrata in $\overline{\mathbb{C}}(e, x)$ ovvero un certo y tale che $y^2 = x$.

Ovviamente la radice quadrata anche se esiste potrebbe non essere unica, cerchiamo allora di fornire alcune condizioni per ottenere una sorta di "risultato di unicità" per le radici quadrate.

Definizione 8.7.1. Un elemento x di un'algebra di Banach complessa X è detto **semidefinito positivo** se e solo se esiste $y \in X$ tale che $x = y^* y$.

Chiaramente gli elementi semidefiniti positivi sono autoaggiunti e quindi normali, in particolare il loro spettro è contenuto interamente in $\mathbb{R}$. Vogliamo dimostrare quindi che un elemento x di X è positivo se e solo se il suo spettro si trova in $[0, +\infty)$ e per farlo procederemo per gradi.

Consideriamo le funzioni $R(t) = t \vee 0$ e $S(t) = (-t) \vee 0$ definite su $\mathbb{R}$ e continue, di conseguenza le quantità $R\{x\}$ ed $S\{x\}$ sono ben definite quando x è autoaggiunto, inoltre la funzione $Q(t) = \sqrt{R(t)}$ è chiaramente ben definita, continua e a valori in $\mathbb{R}$. Poiché R ed S hanno valori in $\mathbb{R}$ avremo che $R\{x\}^* = R^*\{x\} = R\{x\}$ e lo stesso vale per S, e quindi $R\{x\}$, $S\{x\}$, $Q\{x\}$ sono automaticamente autoaggiunti. Per le proprietà di linearità e moltiplicatività si avranno le seguenti relazioni

$$x = R\{x\} - S\{x\} \qquad\qquad Q\{x\}^2 = R\{x\}$$

ma $R\{x\}$ è uguale anche a $Q\{x\}^* Q\{x\}$ e quindi $R\{x\}$ (come anche $S\{x\}$ e $Q\{x\}$) è semidefinito positivo e possiede una radice anch'essa semidefinita positiva.

Se ora prendiamo $x \in X$ tale che $\sigma(x) \geq 0$ allora S coincide con l'applicazione identicamente nulla su $\sigma(x)$, per la biettività di T_x segue che anche $S\{x\}$ deve essere nullo e perciò $x = R\{x\}$. Abbiamo appena dimostrato la prima parte del

Teorema 8.7.2. *Se $x \in X$ normale, con X C^*-algebra unitaria, ha lo spettro contenuto interamente in $[0, +\infty)$ allora x è semidefinito positivo. In tal caso inoltre $\sqrt{x} = \sqrt{\lambda}\{x\}$ è l'unica radice quadrata di x che sia semidefinita positiva.*

Dimostrazione. Essendo lo spettro di x contenuto in $[0, +\infty)$ l'elemento $\sqrt{x}$ è ben definito ed appartiene a $\overline{\mathbb{C}}(e, x)$. Sia ora $y \in X$ semidefinito positivo per cui $y^2 = x$ allora $\overline{\mathbb{C}}(e, x) \subseteq \overline{\mathbb{C}}(e, y)$ il quale è una C^*-algebra commutativa unitaria. Per il teorema 8.6.7 possiamo così ricondurci al caso $X = C(K, \mathbb{C})$ con K spazio topologico compatto di Hausdorff.

Un elemento $f \in C(K, \mathbb{C})$ è semidefinito positivo se e solo se esiste $g : K \to \mathbb{C}$ continua tale che per ogni $a \in K$

$$f(a) = (g^* g)(a) = g(a)^\dagger g(a) = |g(a)|^2$$

e quindi f è semidefinito positivo se e solo se $f(a) \geq 0$ per ogni $a \in K$, viceversa infatti ponendo $h(a) = \sqrt{f(a)}$ avremo chiaramente $h \in C(K, \mathbb{C})$ ed $f = h^2 = h^* h$. Ora scegliamo un qualunque $g \in C(K,\mathbb{C})$ per cui $f = g^2$ se e solo se $f(a) = g(a)^2$ ma essendo sia $f(a)$ che $g(a)$ numeri reali non negativi allora $g(a)$ è univocamente determinato da $f(a)$ e quindi anche g sarà univocamente determinato da f, perciò in $C(K,\mathbb{C})$ la radice semidefinita positiva di una funzione semidefinita positiva esiste ed è unica.

Ora essendo $y, \sqrt{x} \in \overline{\mathbb{C}}(\mathbb{e}, y)$ ed $y^2 = \left(\sqrt{x}\right)^2 = x$ per quanto detto prima deve essere $y = \sqrt{x}$ e l'unicità è dimostrata. ■

Il prossimo passo è dimostrare che ogni elemento semidefinito positivo ha spettro non negativo. Prima però facciamo un resoconto sui diversi modi per calcolare la radice di un elemento x:

- se esiste $t > 0$ per cui $\|\mathbb{e} - x/t\| \leq 1$ allora si può calcolare tramite la serie di potenze della radice calcolata in x/t per determinare una sua radice;

- se x ha spettro non negativo allora ha un'unica radice semidefinita positiva definita come $\sqrt{x} = \sqrt{\lambda}\{x\}$.

e quindi si potrebbe utilizzare l'uno o l'altro in base ai casi in cui ci si trova. In realtà se x è autoaggiunto allora o entrambi i metodi sono validi o nessuno dei due è utilizzabile per via di questo lemma:

Lemma 8.7.3. *Se X è una C^*-algebra ed $x \in X$ è autoaggiunto allora sono equivalenti:*

1. $\sigma(x) \subseteq [0, +\infty)$;
2. *per ogni $t \geq \|x\|$ si ha $\|t\mathbb{e} - x\| \leq t$;*
3. *esiste $t > 0$ tale che $\|t\mathbb{e} - x\| \leq t$.*

Dimostrazione. Per ogni $t \geq \|x\|$ definiamo la funzione $f_t(\lambda) = t - \lambda$ su $\sigma(x)$. Essendo x autoaggiunto avremo che f_t ha valori reali, anzi abbiamo proprio $f_t(\lambda) \geq 0$ per ogni $\lambda \in \sigma(x) \subseteq [-\|x\|, \|x\|]$. Dato che lo spettro è compatto indichiamo con $\bar{\lambda}$ il minimo di $\sigma(x)$, quindi x ha spettro non negativo se e solo se $\bar{\lambda} \geq 0$.

Ora $f_t\{x\} = t\mathbb{e} - x$ ma essendo T_x una isometria tra C^*-algebre avremo anche che

$$\|t\mathbb{e} - x\| = \|f_t\|_\infty = \max_{\lambda \in \sigma(x)} (t - \lambda) = t - \bar{\lambda}$$

e quindi la 1 implica la 2 e la 3 implica la 1 quando $t \geq \|x\|$. Supponiamo ora che la 3 è soddisfatta per un qualche $0 < t < \|x\|$, allora $x/t \in D(\mathbb{e}, 1)$ ed essendo anche $0 \in D(\mathbb{e}, 1)$ ed $t/\|x\| < 1$ per convessità avremo che

$$\frac{x}{\|x\|} = \frac{t}{\|x\|}\frac{x}{t} + \frac{\|x\| - t}{\|x\|} 0 \in D(\mathbb{e}, 1)$$

e quindi la 3 è soddisfatta anche quando $t = \|x\|$ e quindi implica la 1. ■

Osservazione. Se x è autoaggiunto allora è possibile trovare una sua radice quadrata con la serie di potenze o con l'isomorfismo T_x se e solo se lo spettro di x è non negativo. Nel caso la radice trovata con la serie di potenze è semidefinita positiva (cosa che verificheremo tra poco) allora coinciderà con l'unica radice semidefinita positiva ottenuta in questa sezione. Quando invece lo spettro di x contiene elementi negativi pur rimanendo autoaggiunta questi due metodi non funzionano più, mentre togliendo l'ipotesi che x sia autoaggiunto la serie di Taylor potrebbe ancora convergere ma ci troverebbe una sola radice che non sappiamo come distinguere dalle altre.

Corollario 8.7.4. *La somma di elementi con spettro non negativo ha ancora spettro non negativo.*

Dimostrazione. Siano $x, y \in X$ con spettro non negativo, allora $\| \|x\| \mathbb{e} - x\| \leq \|x\|$ e lo stesso vale per y. Posto $t = \|x\| + \|y\|$ avremo che

$$\left\|(\|x\| + \|y\|)\mathbb{e} - x - u\right\| \leq \| \|x\| \mathbb{e} - x\| + \left\| \|y\| \mathbb{e} - y\right\| \leq \|x\| + \|y\|$$

e per il lemma precedente $\sigma(x + y) \geq 0$. ■

Un altro risultato preliminare di cui abbiamo bisogno è il seguente:

Proposizione 8.7.5. *Siano x, y elementi di un'algebra unitaria X allora*

$$\sigma(xy) \setminus \{0\} = \sigma(yx) \setminus \{0\}$$

Dimostrazione. Sia $\lambda \neq 0$ appartenente ad $\rho(xy)$ ovvero $xy - \lambda\mathbb{e} \in G(X)$. Osserviamo che

$$y(xy - \lambda\mathbb{e}) = (yx - \lambda\mathbb{e})y$$

il che implica che $y = (yx - \lambda\mathbb{e})y(xy - \lambda\mathbb{e})^{-1}$. Di conseguenza

$$\mathbb{e} = \lambda^{-1}(\lambda\mathbb{e} - yx) + yx = (yx - \lambda\mathbb{e})\left(-\lambda^{-1} + y(xy - \lambda\mathbb{e})^{-1}x\right)$$

e quindi $yx - \lambda\mathbb{e}$ è invertibile a destra. Con lo stesso procedimento si vede che è invertibile anche a sinistra e dunque $\lambda \in \rho(yx) \setminus \{0\}$. Scambiando x con y si ha la tesi. ■

Siamo ora in grado di dimostrare il seguente

Teorema 8.7.6. *Data una C^*-algebra X ed un elemento normale $x \in X$ allora X è semidefinito positivo se e solo se il suo spettro è non negativo.*

Dimostrazione. Sia x semidefinito positivo, ovvero esiste $y \in X$ tale che $x = y^* y$ e definiamo $z = yS\{x\}$ (che è ben definita in quanto x è autoaggiunto). Di conseguenza

$$z^* z = S\{x\}y^* yS\{x\} = S\{x\}xS\{x\} = -S\{x\}^3$$

in quanto $tS(t) = -S(t)^2$.

La proposizione 8.5.2 ci dice che esistono a, b autoaggiunti tali che $z = a + ib$ e $z^* = a - ib$ e così

$$zz^* = a^2 - iab + iba + b^2 = 2a^2 + 2b^2 - z^* z = 2a^2 + 2b^2 + S\{x\}^3$$

Ora x, a, b sono tutti autoaggiunti, posto $f(\lambda) = \lambda^2$ il corollario 8.6.15 garantisce che gli spettri di $a^2 = f\{a\}$, b^2 e $S\{x\}$ sono non negativi e quindi anche la loro somma zz^* ha lo spettro contenuto in $[0, +\infty)$ mentre lo spettro di $z^* z$ sarà formato interamente da elementi non positivi.

La proposizione 8.7.5 d'altronde ci dice che gli spettri di $z^* z$ e di zz^* privati eventualmente dello zero devono coincidere, ma d'altronde essendo entrambi non vuoti (siamo in $\mathbb{C}$) avremo così $\sigma(z^* z) = \sigma(zz^*) = \{0\}$ e perciò

$$\{0\} = \sigma\left(-S^3\{x\}\right) = -S^3\left[\sigma(x)\right]$$

e questo è possibile se e solo se lo spettro di x è non negativo. Il teorema è così dimostrato. ■

Osservazione. Si tenga presente che y non è detto che sia normale, quindi anche se poniamo $f(\lambda) = |\lambda|^2$ non è possibile definire $f\{y\}$ né tanto meno sarà uguale a $y^* y = x$.

Possiamo allora identificare gli elementi semidefiniti positivi con gli elementi che hanno lo spettro non negativo. In questo modo abbiamo i seguenti risultati:

Corollario 8.7.7. *Se f è una funzione continua da $\mathbb{C}$ in $\mathbb{R}$ per ogni x normale avremo che $f\{x\}$ è autoaggiunto. Se poi f ha valori positivi allora $f\{x\}$ è anche semidefinito positivo.*

Corollario 8.7.8. *L'insieme X^+ degli elementi semidefiniti positivi della C^*-algebra unitaria X è un cono convesso di X. Inoltre l'insieme $X^+ - X^+$ è un sottospazio vettoriale di $X_{\mathbb{R}}$ che coincide con l'insieme di tutti gli elementi autoaggiunti. Quindi se x ed y sono autoaggiunti allora anche $x + y$ e tx lo sono per ogni $t \in \mathbb{R}$.*

Dimostrazione. L'ultima parte discende dalla proposizione 5.1.8. ∎

Capitolo 9

Operatori in spazi normati

9.1 Aggiunto di operatori

Consideriamo un generico spazio normato X con spazio duale X^*. Possiamo definire la seguente applicazione bilineare $\langle \cdot : \cdot \rangle : X^* \times X \to \mathbb{K}$ tale che

$$\langle f : x \rangle = f(x) \tag{9.1.1}$$

Proposizione 9.1.1. *L'applicazione $\langle \cdot : \cdot \rangle$ è continua, ovvero per ogni $x_n \to x \in X$ e $f_n \to f \in X^*$ si ha*

$$\langle f_n : x_n \rangle \to \langle f : x \rangle$$

Dimostrazione. Innanzitutto per ogni $f \in X^*$ e $x \in X$ abbiamo

$$|\langle f : x \rangle| \leq \|f\| \, \|x\|$$

nei rispettivi spazi. Quindi se $f_n \to f$ in X^* allora esiste $M > 0$ tale che $\|f_n\| \leq M$ per ogni $n \in \mathbb{N}$. Ma allora

$$|\langle f_n : x_n \rangle - \langle f : x \rangle| = |\langle f_n : x_n - x \rangle + \langle f_n - f : x \rangle| \leq M \|x_n - x\| + \|f_n - f\| \, \|x\| \to 0$$

e quindi $\langle \cdot : \cdot \rangle$ è continuo. ∎

In realtà grazie alla proposizione 7.4.3 abbiamo anche i seguenti risultati di continuità per successioni

- Se $x_n \rightharpoonup x$ e $f_n \to f$ allora $\langle f_n : x_n \rangle \to \langle f : x \rangle$;
- Se $x_n \to x$ e $f_n \overset{*}{\rightharpoonup} f$ allora $\langle f_n : x_n \rangle \to \langle f : x \rangle$.

Non possiamo però fare a meno della convergenza forte come si può vedere nell'esempio seguente

Esempio 9.1.2. Sia $X = \ell^2$ e consideriamo le successioni $e_n \in \ell^2$, $e_n^* \in (\ell^2)^*$ definite nella seguente maniera

$$e_n(i) = \begin{cases} 1 & \text{se } i = n \\ 0 & \text{se } i \neq n \end{cases}$$

$$e_n^*(d) = d(n) \quad \forall d \in \ell^2$$

Allora $e_n \rightharpoonup 0$ in ℓ^2 e $e_n^* \overset{*}{\rightharpoonup} 0$ in $(\ell^2)^*$. Ma $\langle e_n^* : e_n \rangle = 1$ per ogni $n \in \mathbb{N}$ e dunque avremo che $\langle e_n^* : e_n \rangle \nrightarrow \langle 0 : 0 \rangle = 0$.

Ricordiamo che $\mathcal{UL}(X, Y)$ è lo spazio di tutti gli operatori lineari da X in Y, e che l'operatore L è densamente definito se e solo se $\operatorname{dom} L$ è denso in X mentre è chiuso se e solo se $\mathcal{G}(L)$ è chiuso in $X \times Y$ rispetto alla topologia prodotto. Nonostante rappresentino una classe molto generica di funzioni è possibile definirvi l'aggiunto in modo del tutto naturale

Definizione 9.1.3. Dati due spazi normati X, Y sul campo $\mathbb{K}$ e $T \in \mathcal{UL}(X, Y)$ operatore lineare densamente definito definiamo l'operatore **aggiunto** di T l'operatore $T^* \in \mathcal{UL}(Y^*, X^*)$ definito nella seguente maniera:

$$f \in \operatorname{dom} T^* \leq Y^* \Leftrightarrow \text{ l'applicazione } x \in \operatorname{dom} T \to \langle f : T(x) \rangle \text{ è continua}$$
$$T^*(f)(x) = \langle f : S(x) \rangle \ \forall f \in \operatorname{dom} T^*, x \in X, S \in L(X, Y) \text{ tale che } T \subseteq S$$

L'operatore aggiunto è sempre ben definito ed univocamente determinato, infatti l'operatore $x \to f[T(x)]$ per ipotesi è continuo su $\operatorname{dom} T$ per ogni $f \in \operatorname{dom} T^*$. D'altronde essendo $\operatorname{dom} T$ denso in X esiste un'unica estensione continua S di T e quindi $T^*(f)$ è univocamente determinata su tutto X ed è perciò un elemento di X^*.

Proposizione 9.1.4. *Se $T \in L(X, Y)$ allora $T^* \in L(Y^*, X^*)$ e inoltre*

$$\|T\| = \|T^*\|$$

Dimostrazione. Per ogni $f, g \in Y^*$, $\alpha, \beta \in \mathbb{R}$ e $x \in X$

$$[T^*(\alpha f + \beta g)](x) = \alpha f(Tx) + \beta g(Tx) = \alpha [T^* f](x) + \beta [T^* g](x)$$

Il lettore può dimostrare facilmente che per ogni $f \in Y^*$

$$\|f \circ T\| \leq \|f\| \|T\|$$

e quindi T^* è un operatore lineare continuo con $\|T^*\| \leq \|T\|$. Viceversa per ogni $x \in X$ esiste $g_x \in Y^*$ di norma unitaria tale che $g_x(Tx) = \|Tx\|$ e quindi

$$\|Tx\| = \|(T^* g_x)(x)\| \leq \|T^* g_x\| \|x\| \leq \|T^*\| \|x\|$$

che ci permette di ottenere la disuguaglianza opposta. ∎

Proposizione 9.1.5. *Dati $A, B \in \mathcal{UL}(X, Y)$ densamente definiti e $\alpha \in \mathbb{K}$ si ha $A^* + B^* \subseteq (A + B)^*$ e $(\alpha A)^* = \alpha A^*$.*

Dimostrazione. Dimostriamo solo il risultato per la somma. Per ogni $f \in \operatorname{dom} A^* \cap \operatorname{dom} B^*$ abbiamo che $\langle f : Ax \rangle$ e $\langle f : Bx \rangle$ sono continue in x, ma allora anche la loro somma

$$\langle f : Ax \rangle + \langle f : Bx \rangle = \langle f : (A + B)(x) \rangle$$

è continua su $\operatorname{dom} A \cap \operatorname{dom} B = \operatorname{dom}(A + B)$ e dunque $f \in \operatorname{dom}(A^* + B^*)$. Per l'unicità dell'operatore aggiunto abbiamo anche

$$(A + B)^*(f)(x) = f[A(x) + B(x)] = f[A(x)] + f[B(x)] = [A^*(f) + B^*(f)](x)$$

sul dominio di $A^* + B^*$. ∎

Proposizione 9.1.6. *Per ogni $L \in \mathcal{UL}(X, Y)$ densamente definito l'operatore L^* è chiuso.*

Dimostrazione. Consideriamo una qualunque successione $f_n \in \operatorname{dom} L^*$ tale che $f_n \to f \in Y^*$ e $L^*(f_n) \to g \in X^*$. Mostriamo inizialmente che $f \in \operatorname{dom} L^*$, prendiamo un qualunque $x \in \operatorname{dom} L$ per avere

$$\langle f : L(x) \rangle = \lim_{n \to +\infty} \langle f_n : L(x) \rangle = \lim_{n \to +\infty} \langle L^*(f_n) : x \rangle = \langle g : x \rangle$$

che è continua in quanto $g \in X^*$ dunque $f \in \operatorname{dom} L^*$. L'unicità dell'aggiunto implica che $g = L^*(f)$ e quindi L^* è sempre chiuso. ■

Osservazione. Non è detto che L^* sia anche densamente definito.

Definizione 9.1.7. Prendiamo $H \subseteq X$, definiamo l'**annichilatore destro** di H l'insieme

$$H^{\perp} = \left\{ x^* \in X^* \;\middle|\; \langle x^* : x \rangle = 0 \text{ per ogni } x \in H \right\}$$

Considerando invece $H \subseteq X^*$ definiamo l'**annichilatore sinistro** di H l'insieme

$${}^{\perp}H = \left\{ x \in X \;\middle|\; \langle x^* : x \rangle = 0 \text{ per ogni } x^* \in H \right\}$$

Un sottospazio H di X ha **codimensione finita** se e solo se esiste un sottospazio $N \leq X^*$ finitamente generato tale che $H = {}^{\perp}N$, in tal caso la **codimensione** di H è la dimensione di N.

Proposizione 9.1.8. *Gli annichilatori destro e sinistro sono sempre sottospazi vettoriali chiusi.*

Dimostrazione. Per ogni $H \subseteq X^*$, $K \subseteq X$ abbiamo

$$\begin{aligned} {}^{\perp}H &= \bigcap_{f \in H} \ker f \\ K^{\perp} &= \bigcap_{x \in K} \ker \hat{x} \end{aligned}$$

poiché l'intersezione arbitraria di sottospazi chiusi è chiusa anche $K^{\perp}$ e ${}^{\perp}H$ sono chiusi. ■

Proposizione 9.1.9. *Se $M \leq X$ è chiuso allora*

$${}^{\perp}\left(M^{\perp}\right) = M$$

Dimostrazione. Per la proposizione precedente si ha

$${}^{\perp}\left(M^{\perp}\right) = \left\{ x \in X \;\middle|\; \langle f : x \rangle = 0 \; \forall f \in X^* \text{ tale che } \langle f : y \rangle = 0 \; \forall y \in M \right\}$$

e dunque si ha $M \subseteq {}^{\perp}\left(M^{\perp}\right)$. D'altronde essendo M un sottospazio vettoriale chiuso per il secondo teorema di Mazur se $x \notin M$ esiste $f \in X^*$ che si annulla su M ma non in x e dunque $x \notin {}^{\perp}\left(M^{\perp}\right)$. ■

Osservazione. Dato che i sottospazi finitamente generati sono sempre chiusi allora un sottospazio H ha codimensione finita se e solo se è chiuso e $H^{\perp}$ ha dimensione finita.

Proposizione 9.1.10. *Se $M \subseteq X$ allora*

$$^{\perp}\left(M^{\perp}\right) = \mathrm{Cl}\left(\mathrm{span}\{M\}\right)$$

Dimostrazione. Definiamo $Y = \mathrm{Cl}\left(\mathrm{span}\{M\}\right)$ sottospazio chiuso, dimostriamo innanzitutto che $Y^{\perp} = M^{\perp}$. Chiaramente $Y^{\perp} \subseteq M^{\perp}$, prendiamo ora un generico $f \in M^{\perp}$ per cui segue immediatamente che $M \subseteq \ker f$ e per definizione di sottospazio vettoriale generato si ha $\mathrm{span}\{M\} \leq \ker f$.

Ma $\ker f$ è anche chiuso e perciò $\mathrm{Cl}\left(\mathrm{span}\{M\}\right) = Y \leq \ker f$ ovvero $f \in Y^{\perp}$. Infine per la proposizione precedente si ha $Y = {}^{\perp}Y^{\perp} = {}^{\perp}M^{\perp}$. ■

Teorema 9.1.11 (Operatori a codominio chiuso)**.** *Prendiamo due spazi normati X, Y d un operatore $A \in \mathcal{UL}\left(X, Y\right)$ densamente definito, allora*

$$\mathrm{im}\, A \textit{ chiuso } \Leftrightarrow \mathrm{im}\, A = {}^{\perp}\ker A^*$$

Dimostrazione. L'implicazione $\Leftarrow$ deriva direttamente dalla proposizione 9.1.8, basta allora dimostrare che $(\mathrm{im}\, A)^{\perp} = \ker A^*$ per qualunque operatore chiuso e densamente definito A. Per ogni $y^* \in Y^*$ osserviamo che

$$\begin{aligned} y^* \in (\mathrm{im}\, A)^{\perp} \Leftrightarrow \ \forall x \in \mathrm{dom}\, A \left\langle y^* : Ax\right\rangle = 0 \Leftrightarrow \ \forall x \in X \left\langle A^* y^* : x\right\rangle = 0 \\ \Leftrightarrow A^* y^* = 0 \Leftrightarrow y^* \in \ker A^* \end{aligned}$$

E applicando la proposizione 9.1.9 otteniamo la tesi. ■

Corollario 9.1.12. *Sia $A \in L(X, Y)$ operatore con* $\mathrm{im}\, A$ *chiuso, allora*

$$A \textit{ suriettivo } \Leftrightarrow A^* \textit{ iniettivo}$$

Teorema 9.1.13. *Siano stavolta X, Y spazi di Banach e $A \in L(X, Y)$, allora*

$$A \textit{ iniettivo e a codominio chiuso } \Leftrightarrow \ \exists c > 0 \textit{ tale che } \|x\| \leq c \|Ax\|$$

Dimostrazione. Sia A iniettivo e a codominio chiuso, allora $\mathrm{im}\, A$ è uno spazio di Banach e dal teorema della funzione aperta esiste $A^{-1} \in L(\mathrm{im}\, A, X)$ inversa di A e quindi

$$\|x\| \leq \left\|A^{-1}\right\| \|Ax\|$$

Supponiamo ora che esiste $c > 0$ tale che $\|x\| \leq c \|Ax\|$ allora A è immediatamente iniettiva, dobbiamo solo dimostrare che il suo codominio è chiuso. Per ogni $x_n \in X$ tale che $y_n = Ax_n$ tende ad un certo $y \in Y$ vale per ipotesi

$$\left\|x_n - x_m\right\| \leq c \left\|y_n - y_m\right\|$$

e quindi x_n è una successione di Cauchy che tende ad un certo $x \in X$ (spazio di Banach). Per continuità di A si ha $y = Ax$. ■

Ricordando la definizione di spazio quoziente l'operatore

$$\mathscr{A} : x + \ker A \in X/\ker A \to Ax \in Y$$

è chiaramente ben definito, lineare e iniettivo e

$$\|\mathscr{A}(x + \ker A)\| = \|A(x+z)\| \le \|A\| \, \|x+z\|$$

per ogni $z \in \ker A$ e quindi $\mathscr{A}$ rimane continuo. Infine $\operatorname{im} A = \operatorname{im} \mathscr{A}$.

Adesso abbiamo tutte le carte in regola per generalizzare il teorema precedente:

Teorema 9.1.14. *Siano stavolta X, Y spazi di Banach e $A \in L(X, Y)$, allora*

$$A \text{ a codominio chiuso } \Leftrightarrow \exists c > 0 \text{ tale che } d(x, \ker A) \le c\,\|Ax\|$$

Teorema 9.1.15. *Siano X, Y spazi di Banach e $A \in L(X, Y)$ a codominio chiuso, allora A^* è a codominio chiuso e inoltre*

$$\operatorname{im} A^* = \ker A^\perp$$

Dimostrazione. Consideriamo $x^* \in Y^*$ e $x \in \ker A$ allora

$$\langle A^* x^* : x \rangle = \langle x^* : Ax \rangle = 0$$

e quindi $\operatorname{im} A^* \subseteq \ker A^\perp$.

Prendiamo ora $x^* \in \ker A^\perp$ ovvero per ogni $x \in \ker A$ $\langle x^* : x \rangle = 0$ e quindi l'applicazione lineare

$$\mathscr{M} : y \in \operatorname{im} A \to \langle x^* : x \rangle \in \mathbb{R} \text{ con } y = Ax$$

è ben definita sullo spazio di Banach $\operatorname{im} A$ (chiuso) dal teorema 9.1.14 esiste $c > 0$ tale che

$$d(x, \ker A) \le c\,\|Ax\| = c\,\|y\|$$

allora per ogni $z \in \ker A$ si ha

$$\begin{aligned} |\mathscr{M}(y)| = |\langle x^* : x \rangle| = |\langle x^* : x + z \rangle| \le \|x^*\| \, \|x+z\| \\ \Rightarrow |\mathscr{M}(y)| \le \|x^*\| \, d(x, \ker A) \le c\,\|x^*\| \, \|y\| \end{aligned}$$

e quindi è anche una funzione continua, per Hahn-Banach possiamo estendere $\mathscr{M}$ su tutto Y e quindi $\mathscr{M} \in Y^*$.

Quindi per ogni $x \in X$ $\langle A^* \mathscr{M} : x \rangle = \langle \mathscr{M} : Ax \rangle = \langle x^* : x \rangle$ ovvero $x^* \in \operatorname{im} A^*$ e la dimostrazione è conclusa. ■

Ricapitolando finora abbiamo ottenuto i seguenti risultati:

- Se X, Y sono spazi normati e $A \in L(X, Y)$ allora $\operatorname{im} A^\perp = \ker A^*$;
- Se X, Y sono spazi normati e $A \in L(X, Y)$ allora $\ker A = {}^\perp \operatorname{im} A^*$;
- Se X, Y sono spazi normati, $A \in L(X, Y)$ con codominio chiuso allora $\operatorname{im} A = {}^\perp \ker A^*$;
- Se X, Y sono spazi di Banach, $A \in L(X, Y)$ con codominio chiuso allora $\ker A^\perp = \operatorname{im} A^*$;
- Se X, Y sono spazi normati e $A \in L(X, Y)$ allora A iniettivo se e solo se A^* suriettivo;
- Se X, Y sono spazi normati, $A \in L(X, Y)$ con codominio chiuso allora A suriettivo se e solo se A^* iniettivo.

9.2 Operatori di Fredholm

Nella sezione 4.7 degli spazi vettoriali topologici abbiamo parlato della somma continua di sottospazi di uno spazio vettoriale topologico. Riprendiamo quei concetti ed applichiamoli al caso in cui X è uno spazio normato.

Proposizione 9.2.1. *Ogni sottospazio di dimensione finita di X ha un supplemento topologico.*

Dimostrazione. I sottospazi $N \leq X$ di dimensione finita sono chiaramente chiusi, consideriamo $e_1, \dots, e_n$ una sua base, per il lemma 5.4.10 esistono $e_1^*, \dots, e_n^* \in X^*$ tali che $e_i^*(e_j) = \delta_{ij}$.

Poniamo

$$p : x \in X \to e_1^*(x)e_1 + e_2^*(x)e_2 + \cdots + e_n^*(x)e_n \in N$$

chiaramente p è una proiezione lineare e continua su N dunque per la proposizione precedente ammette supplemento topologico. ∎

Corollario 9.2.2. *Sia $N \leq X$ sottospazio chiuso tale che* $\dim N^\perp = n$ *finito, allora esiste un sottospazio B di X di dimensione n tale che X è somma continua di N e B.*

Dimostrazione. Se $e_1^*, \dots, e_n^*$ è una base di $N^\perp$ allora per il lemma 5.4.10 esistono dei vettori linearmente indipendenti $e_1, e_2, \dots, e_n \in X$ tali che $e_i^*(e_j) = \delta_{ij}$. Se definiamo p come nella proposizione precedente allora

$$x \in \ker p \Leftrightarrow e_i^*(x) = 0\ \forall i \Leftrightarrow f(x) = 0\ \forall f \in N^\perp \Leftrightarrow x \in {}^\perp\left(N^\perp\right)$$

e per la proposizione 9.1.9 segue che $\ker p = N$. ∎

Proposizione 9.2.3. *Consideriamo uno spazio topologico X e $R \leq X^*$ tale che esiste $N \leq X$ di dimensione n e $R = N^\perp$, allora esiste $S \leq X^*$ di dimensione n tale che $X^* = R \oplus S$.*

Dimostrazione. Analoga alle precedenti. ∎

Definizione 9.2.4. Siano X e Y spazi di Banach, un operatore $A \in L(X, Y)$ è **semifredholmiano**, e si indica con $A \in \Phi_+(X.Y)$ se e solo se $\dim\ker A < +\infty$ e $\operatorname{im} A$ è chiuso.

$A \in L(X, Y)$ è invece **fredholmiano**, e si indica con $A \in \Phi(X, Y)$, se e solo se $A \in \Phi_+(X, Y)$ e $A^* \in \Phi_+(Y^*, X^*)$ e il numero $\operatorname{ind}(A) = \dim\ker A - \dim\ker A^*$ è detto **indice** di A.

Ricordiamo che se X e Y sono spazi di Banach allora $\operatorname{im} A^\perp = \ker A^*$.

Proposizione 9.2.5. *Siano X e Y spazi di Banach e $A \in L(X, Y)$. Allora $A \in \Phi(X, Y)$ se e solo se $A \in \Phi_+(X, Y)$ e il supplemento topologico di* $\operatorname{im} A$ *ha dimensione finita.*

Dimostrazione. Supponiamo A fredholmiano, allora è anche semifredholmiano e il supplementare topologico di $\operatorname{im} A$ ha la stessa dimensione di $\ker A^*$ per la proposizione precedente.

Viceversa essendo spazi di Banach $\operatorname{im} A^*$ è chiuso mentre la dimensione di $\operatorname{im} A^\perp$ coincide con quella del supplementare topologico di $\operatorname{im} A$. ∎

Proposizione 9.2.6. *Se X è uno spazio di Banach e M, N due sottospazi chiusi di X tali che $X = M \oplus N$ allora le applicazioni*

$$P_M : m + n \in X \to m \in M$$
$$P_N : m + n \in X \to n \in N$$

sono continue.

Dimostrazione. Ci basta far vedere che il grafico di P_M è chiuso, scegliamo allora $x_n, x \in X$, $y \in M$ tali che $x_n \to x$ e $P_M x_n \to y$ dobbiamo dimostrare che $y = P_M x$. Innanzitutto la successione $x_n - P_M x_n \in N$ converge a $x - y \in N$ per la chiusura di N, ancora $x_n = P_M x_n + (x_n - P_M x_n)$ e passando al limite si ha $P_M x + P_N x = y + (x - y) \Rightarrow P_M x = y$ e dal teorema del grafico chiuso P_M è continua. ■

Teorema 9.2.7. *Siano X, Y spazi di Banach e $A \in \Phi_+(X, Y)$. Allora esiste un sottospazio chiuso $X_0 \leq X$ e posto $A_0 : x \in X_0 \to Ax \in \operatorname{im} A$ si ha A_0 invertibile e $A_0^{-1} \in L(\operatorname{im} A, X_0)$.*

Se inoltre A è fredholmiano allora esiste $\hat{A} \in L(Y, X)$ tale che per ogni $y \in \operatorname{im} A$

$$\hat{A} y = A_0^{-1} y$$

Dimostrazione. Prendiamo X_0 il supplemento topologico di $\ker A$ che è chiuso, quindi si ha $A_0 \in L(X_0, \operatorname{im} A)$. Sia ora $x \in X_0$ tale che $A_0 x = 0$ allora $x \in \ker A \cap X_0 = \{0\}$ e quindi è iniettivo, se ora $y \in \operatorname{im} A$ esiste $x' \in X$ tale che $y = Ax' = A_0(P_{X_0} x')$ e quindi è anche suriettivo. Dal corollario 4.4.6 l'inversa è continua.

Se A è fredholmiano allora $\ker A^* = \operatorname{im} A^{\perp}$ ha dimensione finita e quindi esiste il supplementare topologico B di $\operatorname{im} A$. La tesi segue osservando che

$$\hat{A} : A_0^{-1} \circ P_{\operatorname{im} A}$$

è continua. ■

Definizione 9.2.8. Poniamo ora X spazio vettoriale e $\|\cdot\|_s$, $\|\cdot\|_w$ due norme di X. Allora diremo che la norma $\|\cdot\|_w$ è **più debole** della norma $\|\cdot\|_s$ se e solo se ogni successione limitata in $(X, \|\cdot\|_s)$ ammette un'estratta di Cauchy in $(X, \|\cdot\|_w)$

Teorema 9.2.9 (Peetre). *Siano X, Y spazi di Banach e $A \in L(X, Y)$. Allora $A \in \Phi_+(X, Y)$ se e solo se esistono $c > 0$ e una norma $\|\cdot\|_w$ più debole di $\|\cdot\|$ tali che $\|x\| \leq c\,\|Ax\| + \|x\|_w$.*

Dimostrazione. Sia A semifredholmiano, allora esiste il supplementare topologico F di $\ker A$ e possiamo restringere A ad F ottenendo l'operatore $A_{|F}$ lineare continuo invertibile con inversa $B \in L(\operatorname{im} A, F) \Rightarrow \exists c > 0$ tale che $\|f\| \leq c\,\|Af\|$ per ogni $f \in F$.

Per ogni $x \in X$ poniamo $\|x\|_w = \|P_{\ker A} x\|$ che è ancora una norma su X, allora

$$\|x\| = \|P_F x + P_{\ker A} x\| \leq c\,\|A(P_F x)\| + \|x\|_w = c\,\|Ax\| + \|x\|_w$$

dobbiamo dimostrare che la norma appena definita è più debole di $\|\cdot\|$. Se $\|x_n\|$ è limitata allora per continuità anche $\|x_n\|_w$ è limitata, però $\ker A$ ha dimensione finita e quindi ammette un'estratta convergente in $\ker A$.

Dimostriamo ora l'implicazione inversa. Per ogni $x \in \ker A$ allora $\|x\| \leq \|x\|_w$ con $\|\cdot\|_w$ una norma più debole, quindi ogni successione limitata in $\|\cdot\|$ ha un'estratta di Cauchy in

$\|\cdot\|_w$ e quindi anche in $\|\cdot\|$ e quindi le palle chiuse in $\ker A$ sono relativamente sequenzialmente compatte, per il teorema di Riesz allora $\dim\ker A < +\infty$.

Sia ora $x_n \in X$ e $y \in Y$ tale che $Ax_n \to y$ e F il supplementare topologico di $\ker A$ e quindi x_n può essere scomposto univocamente come somma di $z_n \in \ker A$ e $w_n \in F$. Per assurdo $\|w_n\|$ è illimitato allora potremmo trovarne una sottosuccessione che la farebbe tendere a $+\infty$, allora poniamo

$$q_n = \frac{w_n}{\|w_n\|}$$

e quindi ammette un'estratta q_{n_k} di Cauchy rispetto alla norma debole ma allora

$$\left\|q_{n_i} - q_{n_j}\right\| \leq \frac{\left\|Aw_{n_i}\right\|}{\left\|w_{n_i}\right\|} + \frac{\left\|Aw_{n_j}\right\|}{\left\|w_{n_j}\right\|} + \left\|q_{n_i} - q_{n_j}\right\|$$

il lettore può constatare allora che per ipotesi segue che q_{n_k} è di Cauchy anche rispetto alla norma iniziale e quindi ammette limite q di norma unitaria. Per continuità però $Aq = 0$ e $q \in \ker A \cap F$ assurdo.

Abbiamo dimostrato così che $\|w_n\|$ è limitata e perciò ammette un'estratta di Cauchy rispetto alla norma $\|\cdot\|_1$ ma è di Cauchy anche rispetto alla norma iniziale e perciò ammette limite $x \in X$. La tesi allora segue per continuità in quanto $y = Ax = \lim_{n\to+\infty} Ax_{k_n} = \lim_{n\to+\infty} Aw_{k_n}$. ■

9.3 Operatori compatti

Definizione 9.3.1. Siano X, Y spazi normati e $\Lambda : X \to Y$ applicazione lineare, allora Λ è **compatto** se e solo se per ogni $K \subseteq X$ limitato $\mathrm{Cl}(\Lambda(K))$ è compatto.

Proposizione 9.3.2. *Un operatore $\Lambda : X \to Y$ è compatto se e solo se per ogni successione $x_n \in X$ limitata esiste un'estratta n_k tale che Λx_{n_k} converge in Y*

Dimostrazione. Consideriamo un qualunque sottoinsieme K limitato in X e sia y_n una successione generica di elementi di $\mathrm{Cl}(\Lambda(K))$, allora esiste $x_n \in K$ tale che $\|y_n - \Lambda x_n\| < \frac{1}{n}$, possiamo quindi trovare un'estratta tale che Λx_{n_k} converge ad un certo $y \in \mathrm{Cl}(\Lambda(K))$ e quindi $y_{n_k} \to y$ e $\mathrm{Cl}(\Lambda(K))$ è sequenzialmente compatto, quindi compatto. ■

Proposizione 9.3.3. *Gli operatori compatti sono continui.*

Dimostrazione. Se per assurdo non lo fosse allora esisterebbe una successione $x_n \in X$ con $\|x_n\| = 1$ tale che $\|\Lambda x_n\| \to +\infty$ e quindi non possiede estratte convergenti. ■

Definiamo l'insieme degli operatori compatti come

$$K(X,Y) = \left\{\Lambda \in L(X,Y) \;\middle|\; \Lambda \text{ compatto}\right\}$$

Proposizione 9.3.4. *Se* $\dim Y < +\infty$ *allora* $K(X,Y) = L(X,Y)$.

Dimostrazione. Dalla continuità segue che $\Lambda(K)$ è limitato per ogni K limitato, quindi l'insieme $\mathrm{Cl}(\Lambda(K))$ è compatto in quanto Y ha dimensione finita. ■

Proposizione 9.3.5. *Se X è uno spazio normato e Y spazio di Banach allora $K(X,Y)$ è un sottospazio chiuso di $L(X,Y)$, in particolare è uno spazio di Banach.*

Dimostrazione. È immediato dimostrare che è un sottospazio, ovvero somma e prodotto per uno scalare di operatori compatti rimane compatto. Sia Λ_n successione di operatori compatti convergente in norma a $\Lambda \in L(X,Y)$, ovvero

$$\forall \epsilon > 0\, \exists N_\epsilon \in \mathbb{N} \text{ tale che } \forall n \geq N_\epsilon\, \forall x \in X \ \left\|\Lambda_n x - \Lambda x\right\| \leq \epsilon \|x\|$$

Se prendiamo adesso una successione x_m limitata da $L > 0$ allora dalla compattezza di Λ_n con un procedimento analogo a quello usato nella dimostrazione del teorema di Ascoli-Arzelà esiste una successione x_k^n estratta da x_k^{n-1} (con $x_k^0 = x_k$) tale che

$$\forall \epsilon > 0\, \exists N(n,\epsilon) \in \mathbb{N} \text{ tale che } \forall m, m' > N(n,\epsilon) \ \left\|\Lambda_n x_m^n - \Lambda_n x_{m'}^n\right\| < \epsilon$$

Allora per ogni $m, n > N(N_\epsilon, \epsilon) \vee N_\epsilon$ si ha

$$\left\|\Lambda x_m^m - \Lambda x_n^n\right\| \leq \left\|\Lambda x_m^m - \Lambda_{N_\epsilon} x_m^m\right\| + \left\|\Lambda_{N_\epsilon} x_m^m - \Lambda_{N_\epsilon} x_n^n\right\| + \left\|\Lambda_{N_\epsilon} x_n - \Lambda x_n\right\| < \epsilon(1+2L)$$

poiché vale

$$\left\|\Lambda_{N_\epsilon} x_m^{N_\epsilon} - \Lambda_{N_\epsilon} x_n^{N_\epsilon}\right\| < \epsilon$$

e da questa possiamo ricondurci alla precedente in quanto per ogni $n \geq N_\epsilon$ esiste $n' \geq n$ intero tale che

$$x_n^n = x_{n'}^{N_\epsilon}$$

dato che stiamo lavorando per estratte. Dato che Y è uno spazio di Banach Λ è compatto. ■

Teorema 9.3.6 (Schauder)**.** *Siano X, Y spazi normati, allora*

$$T \in K(X,Y) \Rightarrow T^* \in K(Y^*, X^*)$$

Dimostrazione. Consideriamo $y_n^* \in Y^*$ successione limitata da $L > 0$, vogliamo dimostrare che $T^* y_n^*$ è di Cauchy.

Poniamo $B = \{x \in X \mid \|x\| \leq 1\}$ allora $T(B)$ è relativamente sequenzialmente compatto (la sua chiusura è compatta). Per comodità indichiamo con $f\big|_A$ la restrizione di f su A, poniamo allora

$$\mathscr{F} = \left\{ z_n^* = y_n^*\big|_{T(B)} \;\middle|\; n \in \mathbb{N} \right\}$$

classe di funzioni "lineari" continue.

- $\mathscr{F}$ equilimitato. Banalmente $\left\| y_n^*\big|_{T(B)} \right\| \leq \left\|y_n^*\right\| \leq L$ e quindi per ogni $x \in B$

$$\left\|z_n^*(Tx)\right\| \leq L \|T\|$$

- $\mathscr{F}$ equicontinua. Basta dimostrare che

$$\exists C > 0 \text{ tale che } \forall n \in \mathbb{N}\, \forall x, y \in T(B) \ \left\|z_n^*(x) - z_n^*(y)\right\| \leq C \left\|x - y\right\|$$

  ma è banale in quanto

$$\left\|z_n^*(x) - z_n^*(y)\right\| = \left\|y_n^*(x) - y_n^*(y)\right\| \leq \left\|y_n^*\right\| \left\|x - y\right\| \leq L \left\|x - y\right\|$$

Possiamo applicare perciò il teorema di Ascoli-Arzelà ($T(B)$ relativamente sequenzialmente compatto) trovando in tal modo un'estratta che renda $z^*_{n_k}$ di Cauchy ovvero

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \text{ tale che } \forall i,j > N \forall y \in T(B) \ \left\| z^*_{n_i}(y) - z^*_{n_j}(y) \right\| \leq \epsilon$$

ovvero per la linearità di T

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \text{ tale che } \forall i,j > N \forall x \in X \ \left\| y^*_{n_i}(Tx) - y^*_{n_j}(Tx) \right\| \leq \epsilon \left\| x \right\|$$

e quindi $y^*_{n_i} \circ T = T^*(y^*_{n_i})$ è di Cauchy in X^*. La tesi è immediata poiché la chiusura di sottospazi relativamente sequenzialmente compatti in spazi completi (X^* è di Banach) è compatta. ■

9.4 Teorema del punto fisso di Shauder

Innanzitutto diamo alcuni risultati senza dimostrazione che utilizzeremo più avanti.

Teorema 9.4.1 (del punto fisso di Brouwer)**.** *Sia $D(0,1) \subseteq \mathbb{R}^n$ la palla chiusa unitaria rispetto alla norma standard e sia $f : D(0,1) \to D(0,1)$ continua, allora esiste un $x \in D(0,1)$ tale che $f(x) = x$.*

Lemma 9.4.2. *Sia X spazio normato di dimensione finita e $K \subseteq H$ convesso, chiuso e limitato che non sia un singleton. Allora esiste una costante $m \in \mathbb{N}$ ed un omeomorfismo tra K e $D(0,1) \subseteq \mathbb{R}^m$.*

Proposizione 9.4.3. *Sia X spazio normato di dimensione finita e $K \subseteq X$ convesso, chiuso e limitato. Allora ogni applicazione continua $f : K \to K$ ha un punto fisso.*

Dimostrazione. Se $\varphi : K \to D(0,1)$ è l'omeomorfismo del lemma precedente allora $\varphi \circ f \circ \varphi^{-1}$ soddisfa il teorema di Brouwer. ■

Dal teorema di Hausdorff e dalla proposizione 3.3.6 se X è uno spazio di Banach e $K \subseteq X$ è chiuso e totalmente limitato allora è anche compatto. Dimostriamo il seguente risultato

Lemma 9.4.4. *Sia X spazio normato e $K \subseteq X$ totalmente limitato, allora anche* $\mathrm{con}(K)$ *è totalmente limitato.*

Dimostrazione. Poiché K è totalmente limitato per ogni $\epsilon > 0$ esiste un sottoinsieme finito $F = \{x_1, x_2, \dots, x_n\} \subseteq K$ tale che

$$K \subseteq \bigcup_{i=1}^{n} B(x_i, \epsilon/2) \tag{9.4.1}$$

L'insieme $\mathrm{con}(F)$ è un sottoinsieme chiuso e limitato del sottospazio $\mathrm{span}\{x_1, x_2, \dots, x_n\}$ di dimensione finita, quindi $\mathrm{con}(F)$ è sequenzialmente compatto e quindi totalmente limitato, esisteranno allora $y_1, y_2, \dots, y_m \in \mathrm{con}(F)$ tali che

$$\mathrm{con}(F) \subseteq \bigcup_{i=1}^{m} B(y_i, \epsilon/2) \tag{9.4.2}$$

Dimostriamo adesso l'asserto iniziale. Sia $z \in \operatorname{con}(K)$ allora esistono $z_1, z_2, \dots z_k \in K$ e $\lambda_1, \lambda_2, \dots, \lambda_k \in [0,1]$ tali che $\sum_i \lambda_i = 1$ e $\sum_i \lambda_i z_i = z$. Dalla (9.4.1) esistono $u_1, \dots, u_k \in F$ non necessariamente distinti tali che $\|z_i - u_i\| < \epsilon/2$ per ogni indice i, allora

$$\left\| z - \sum_{i=1}^{k} \lambda_i u_i \right\| < \frac{\epsilon}{2}$$

Sfruttando ora la (9.4.2) esiste y_j tale che

$$\|z - y_j\| \leq \left\| z - \sum_{i=1}^{k} \lambda_i u_i \right\| + \left\| \sum_{i=1}^{k} \lambda_i u_i - y_j \right\| < \epsilon$$

e quindi

$$\operatorname{con}(K) \subseteq \bigcup_{i=1}^{m} B(y_i, \epsilon)$$

ovvero $\operatorname{con}(K)$ è totalmente limitati. ■

Corollario 9.4.5. *Sia X spazio di Banach e $K \subseteq X$ totalmente limitato, allora $\overline{\operatorname{con}}(K)$ è compatto*

Dimostrazione. Segue dal lemma precedente e dalla proposizione 3.3.6. ■

Siamo ora in grado di dimostrare il teorema più importante di questa sottosezione:

Teorema 9.4.6 (del punto fisso di Shauder)**.** *Sia X spazio di Banach e K sottoinsieme non vuoto compatto convesso, allora ogni applicazione continua $f : K \to K$ ha un punto fisso.*

Dimostrazione. Grazie al teorema di Hausdorff per ogni intero $m \in \mathbb{N}$ esisteranno dei punti $x_1, x_2, \dots, x_n \in X$ tali che $K \subseteq \bigcup_{i=1}^{n} B\left(x_i, \frac{1}{m}\right)$, poniamo $K_m = \operatorname{con}(\{x_i \mid i \leq n\}) \subseteq K$. Definiamo la funzione

$$\varphi_m : x \in K \to \frac{\sum_{i=1}^{n} \operatorname{dist}\left[x, K \setminus B\left(x_i, \frac{1}{m}\right)\right] x_i}{\sum_{i=1}^{n} \operatorname{dist}\left[x, K \setminus B\left(x_i, \frac{1}{m}\right)\right]} \in K_m$$

Per come sono stati definiti gli x_i il denominatore non si annulla mai e per la proposizione 3.5.9 è continuo. Quindi

$$\|x - \varphi_m(x)\| \leq \frac{\sum_{i=1}^{n} \operatorname{dist}\left[x, K \setminus B\left(x_i, \frac{1}{m}\right)\right] \|x - x_i\|}{\sum_{i=1}^{n} \operatorname{dist}\left[x, K \setminus B\left(x_i, \frac{1}{m}\right)\right]}$$

ricordiamo che $\operatorname{dist}\left[x, K \setminus B\left(x_i, \frac{1}{m}\right)\right] > 0 \Leftrightarrow x \in B\left(x_i, \frac{1}{m}\right) \Leftrightarrow \|x - x_i\| \leq \frac{1}{m}$ e quindi

$$\|x - \varphi_m(x)\| \leq \frac{1}{m}$$

Presa una qualunque funzione continua f da K in sé alla restrizione di $\varphi_m \circ f$ sull'insieme K_m che sappiamo essere convesso, chiuso, limitato e contenuto in un sottospazio di dimensione finita può essere applicato il teorema di Bouwer, esisterà dunque una successione $y_m \in K_m \subseteq K$ tale che $y_m = \varphi[f(y_m)]$ per ogni $m \in \mathbb{N}$. Esisterà allora un'estratta di y_m che convergerà ad un certo $y \in K$ e quindi dalla relazione

$$\|y_m - f(y_m)\| = \|\varphi[f(y_m)] - f(y_m)\| \leq \frac{1}{m}$$

segue che $f(y) = y$ completando così la tesi. ■

Teorema 9.4.7 (Leray-Schauder)**.** *Sia sempre X spazio di Banach ed $f : X \to X$ una generica funzione continua che manda successioni limitate in successioni con un'estratta convergente in X. Supponiamo anche che l'insieme*

$$\left\{x \in X \,\middle|\, \exists \lambda \in [0,1] \text{ tale che } x = \lambda f(x)\right\}$$

è limitato, allora f possiede un punto fisso.

Dimostrazione. Sia $M > 0$ la costante che delimita tale insieme ovvero

$$\mathscr{A} = \left\{x \in X \,\middle|\, \exists \lambda \in [0,1] \text{ tale che } x = \lambda f(x)\right\} \subseteq B\big(0, M\big)$$

Definiamo l'applicazione

$$g(x) = \begin{cases} f(x) & \text{se } \left\|f(x)\right\| \leq M \\ \frac{M}{\|f(x)\|} f(x) & \text{se } \left\|f(x)\right\| > M \end{cases}$$

la funzione g potremmo vederla anche come applicazione da $D\big(0,M\big)$ in sé stesso e continua a mandare successioni chiuse in successioni con un'estratta convergente. Posto $K = \mathrm{Cl}(\mathrm{con})\left(g(D\big(0,M\big))\right) \subseteq D\big(0,M\big)$ allora dal corollario 9.4.5 K è convesso e compatto poiché $g(D\big(0,M\big))$ è precompatto, difatti

$$y_n \in g(D\big(0,M\big)) \to y_n = g(x_n) \text{ con } \left\|x_n\right\| < M \Rightarrow y_n \text{ ha estratta di Cauchy}$$

e dal teorema 3.5.19 è precompatto. Perciò $g(K) \subseteq K$ e possiamo porre $g : K \to K$, applichiamo il teorema del punto fisso di Schauder e quindi esisterà un punto fisso x di g.

Se $\left\|f(x)\right\| \leq M$ allora è un punto fisso anche per f altrimenti si avrebbe $x = \frac{M}{\|f(x)\|} f(x) = \lambda f(x)$ e quindi $x \in \mathscr{A}$, ma allora $M > \|x\| = \left\|g(x)\right\| = M$ il che è assurdo. ■

9.5 Operatori di Riesz

Definizione 9.5.1. Consideriamo ora uno spazio di Banach X. Un operatore $A \in L(X,X)$ è di **Riesz** se e solo se esiste $K \in K(X,X)$ tale che

$$A = \mathrm{id}_X - K$$

con id_X l'applicazione identica su X.

Proposizione 9.5.2. *Sia $L \in L(X,X)$ allora*

$$\left(\mathrm{id}_X - L\right)^* = \mathrm{id}_{X^*} - L^*$$

Dimostrazione. per ogni $x \in X$ e $f \in X^*$

$$\left\langle \left(\mathrm{id}_X - L\right)^* f : x \right\rangle = \left\langle f : x - Lx \right\rangle = \left\langle f : x \right\rangle - \left\langle L^* f : x \right\rangle = \left\langle \left(\mathrm{id}_{X^*} - L^*\right) f : x \right\rangle$$

■

Corollario 9.5.3. *Se A è un operatore di Riesz allora anche A^* è di tipo Riesz.*

Proposizione 9.5.4. *Gli operatori di Riesz sono fredholmiani.*

Dimostrazione. Basta dimostrare che ogni operatore A di Riesz è semifredholmiano dato che anche l'aggiunto è un operatore di Riesz.

Sia $A+K=\mathrm{id}_X$ con K compatto, allora la norma $\|x\|_1=\|Kx\|$ è chiaramente più debole di $\|\cdot\|$ quindi $\|x\|\leq\|Ax\|+\|x\|_1$ e la tesi segue da Peetre. ∎

Teorema 9.5.5 (Primo teorema di Fredholm)**.** *Se X è uno spazio di Banach allora l'operatore di Riesz A è iniettivo se e solo se è suriettivo.*

Dimostrazione. Per assurdo $A=\mathrm{id}_X-K$ è iniettivo ma non suriettivo, essendo semifredholmiano $X_1=A(X)\leq X$ è chiuso, dimostriamo che $A(X_1)<X_1$. Infatti se fossero uguali e $x\in X\setminus X_1$ allora $Ax\in X_1=A(X_1)$ e quindi esiste $y\in X_1$ con $Ax=Ay$ assurdo poiché A è iniettiva.

Possiamo quindi costruire una successione strettamente decrescente di sottospazi chiusi $X_n=A(X_{n-1})<X_{n-1}$ in quanto la restrizione di operatori di Riesz su sottospazi chiusi rimane di Riesz. Dal lemma 7.3.3 esiste una successione $x_n\in E_{n-1}$ di norma unitaria con $d(x_n,E_n)\geq\frac{1}{2}$, prendiamo ora due interi $i<j$ allora

$$X_j<X_{j-1}\leq X_i<X_{i-1}$$

e

$$\left\|Kx_i-Kx_j\right\|=\left\|Ax_j-Ax_i+x_i-x_j\right\|\geq d(x_i,X_i)\geq\frac{1}{2}$$

ovvero Kx_n non ammette estratte di Cauchy raggiungendo così un assurdo che può essere risolto solo supponendo A suriettivo.

Vivecersa se A è suriettivo dalle relazioni di ortogonalità A^* è iniettivo e per il punto precedente A^* è suriettiva e perciò $\{0\}={}^{\perp}\mathrm{im}\,A^*=\ker A$. ∎

Teorema 9.5.6 (Secondo teorema di Fredholm)**.** *Sia $A=\mathrm{id}_X-K$ di Riesz sullo spazio di Banach X, allora*

$$A\text{ iniettivo}\Leftrightarrow A^*\text{ suriettivo}\Leftrightarrow A^*\text{ iniettivo}\Leftrightarrow A\text{ suriettivo}$$

Dimostrazione. Segue immediatamente dalla dimostrazione del teorema precedente. ∎

Il seguente teorema è conosciuto anche come **teorema indice zero**.

Teorema 9.5.7 (Terzo teorema di Fredholm)**.** *Per ogni operatore $A=\mathrm{id}_X-K$ di Riesz vale* $\mathrm{ind}\,(A)=0$.

Dimostrazione. Prendiamo $m=\dim\ker A$ e $n=\dim\ker A^*$, se $n=0$ dal secondo teorema di Fredholm anche $m=0$ e viceversa, quindi possiamo supporli entrambi positivi. Due casi

Supponiamo $n<m$, sia B il supplementare topologico di $\ker A$ e S il supplementare topologico di $\mathrm{im}\,A$ con $\dim S=m$ poiché $\mathrm{im}\,A={}^{\perp}\ker A^*$.

Prendiamo $x_1,\dots,x_n$ base di $\ker A$ e $y_1,\dots,y_m$ base di S e ricordiamo che le proiezioni $P_{\ker A}$ e P_B sono continue. Ora definiamo l'applicazione

$$K_1:\sum_{i=1}^{n}\alpha_ix_i+x_B\in\ker A\oplus B\to\sum_{i=1}^{n}\alpha_iy_i\in X$$

che è lineare e continua, è anche compatto poiché $\mathrm{im}\,K_1$ ha dimensione finita. Posto $A_1=A-K_1$ è di Riesz.

Dimostriamo che

- A_1 è iniettiva. Per ogni $x \in \ker A_i$

$$x - Kx - K_1 x = 0 \Rightarrow x - Kx \in S \cap \operatorname{im} A \Rightarrow x \in \ker A \cap B \Rightarrow x = 0$$

- A_1 non è suriettivo. Per assurdo esiste $x \in X$ tale che $A_1 x = y_m$ ovvero $Ax = y_m + sum_{i=1}^{n} \alpha_i y_i \in S \cap \operatorname{im} A$ ma è assurdo poiché gli y_i sono indipendenti e $\alpha_m = 1$.

Ciò però porta ad un assurdo per il primo teorema di Fredholm, e questo assurdo è nato nell'aver posto $n < m$.

Supponiamo infine $m < n$, costruiamo l'operatore

$$K_2 : \sum_{i=1}^{n} \alpha_i x_i + x_B \in \ker A \oplus B \to \sum_{i=1}^{m} \alpha_i y_i \in X$$

e $A_2 = A - K_2$ è ancora di Riesz. Dimostriamo che

- A_2 è suriettiva. Infatti per ogni $y = \sum_{i=1}^{m} \alpha_i y_i + y_B \in S \oplus \operatorname{im} A = X$ esiste $x_B \in B$ tale che $y_B = Ax_B$ poiché B è il supplemento topologico di $\ker A$. Quindi l'immagine di

$$-\sum_{i=1}^{m} \alpha_i x_i + x_B$$

coincide con y.

- A_2 non è iniettiva. Difatti $A_2(x_n) = A(x_n) = 0$ poiché x_n è un elemento della base del kernel di A.

Abbiamo quindi raggiunto di nuovo un assurdo, perciò deve per forza valere $m = n$. ■

Rimettendo insieme tutti i risultati precedenti otteniamo il seguente

Teorema 9.5.8 (dell'alternativa di Fredholm)**.** *Sia X spazio di Banach, $K \in K(X,X)$ e $A = \mathrm{id}_X - K$ operatore di Riesz. Allora una ed una sola delle seguenti affermazioni sarà vera:*

1. $\ker A = \{0\}, \operatorname{im} A = X, \ker A^* = \{0\}$ *e* $\operatorname{im} A^* = X^*$*;*
2. $\dim \ker A = \dim \ker A^* \in \mathbb{N}$ *e* $\operatorname{im} A = {}^{\perp}\ker A^*, \operatorname{im} A^* = \ker A^{\perp}$.

Oppure in maniera analoga

1. *Il seguente sistema nelle variabili $x \in X$ e $x^* \in X^*$*

$$\begin{cases} Ax = y \\ A^* x^* = y^* \end{cases}$$

anno un'unica soluzione per ogni $y \in X$ e per ogni $y^* \in X^*$.

2. *Le soluzioni del sistema in x e x^**

$$\begin{cases} Ax = 0 \\ A^* x^* = 0 \end{cases}$$

sono infinite e generano due sottospazi di dimensione finita di X e X^* rispettivamente con la stessa dimensione.

9.6 Teoria spettrale

Nel capitolo delle algebre di Banach abbiamo già definito il risolvente e lo spettro di un vettore. Dato che lo spazio $L(X,X)$ è un'algebra di Banach con prodotto interno $ST = S \circ T$ quando X è uno spazio di Banach queste definizioni e i risultati ottenuti continuano a valere anche in $L(X,X)$.

Per quanto riguarda l'algebra degli operatori possiamo introdurre delle definizioni aggiuntive per quanto riguarda lo spettro:

Definizione 9.6.1. Sia $T \in L(X,X)$ definito si X spazio di Banach, oltre al risolvente e allo spettro di T definiti come nel capitolo sulle algebre di Banach definiamo lo **spettro puntuale** di T è l'insieme

$$\sigma_p(T) = \left\{\lambda \in \mathbb{K} \;\middle|\; T - \lambda \mathrm{id}_X \text{ non è iniettiva su } X\right\}$$

Gli elementi di $\sigma_p(T)$ sono detti **autovalori** di T, mentre per ogni $\lambda \in \sigma_p(T)$ un elemento $x \in X \setminus \{0\}$ è un **autovettore** di T associato a λ se e solo se

$$Tx = \lambda x$$

ovvero se e solo se $x \in \ker\left(\lambda \mathrm{id}_X - T\right) \setminus \{0\}$.

Lo **spettro continuo** di T invece è l'insieme

$$\sigma_c(T) = \left\{\lambda \in \sigma(T) \setminus \sigma_p(T) \;\middle|\; (T - \lambda \mathrm{id}_X)(X) \text{ è denso in } X\right\}$$

mentre lo **spettro residuo** di T è l'insieme

$$\sigma_r(T) = \sigma(T) \setminus \left[\sigma_p(T) \cup \sigma_c(T)\right]$$

Nonostante gran parte dei risultati sullo spettro si riconducono al capitolo precedente la struttura dell'algebra $L(X,X)$ permette di ottenere altri risultati di notevole interesse che non possono essere ottenuti direttamente dalla definizione di algebra di Banach o di C^*-algebra.

Quando X è uno spazio di Banach complesso il teorema 8.4.3 garantisce che lo spettro di ogni operatore in $L(X,X)$ non è mai vuoto, mentre se X fosse uno spazio di Banach reale ciò potrebbe non essere vero. Invece lo spettro puntuale $\sigma_p(T)$ può essere vuoto anche se lo spazio è complesso, basta porre ad esempio $X = \ell^2(\mathbb{N},\mathbb{C})$ e l'operatore $T \in L(X,X)$ definito come

$$T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$$

Il prossimo risultato invece si concentra sugli autovettori:

Lemma 9.6.2. *Autovettori associati a diversi autovalori sono linearmente indipendenti.*

Dimostrazione. Dimostriamolo per induzione sul numero di autovalori. Per $n = 1$ è banalmente verificato, quindi supponiamolo vero per $n-1$ e scegliamo n autovettori $x_1, \dots, x_n$ associati rispettivamente agli autovalori $\lambda_1, \dots, \lambda_n$ distinti tra loro, siano allora $\alpha_i \in \mathbb{K}$ tali che

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i T x_i = \sum_{i=1}^{n} \alpha_i \lambda_i x_i = 0$$

Inoltre

$$\sum_{i=1}^{n} \alpha_i \lambda_n x_i = 0 \Rightarrow \sum_{i=1}^{n-1} \alpha_i (\lambda_i - \lambda) x_i = 0 \Rightarrow \alpha_i = 0 \text{ per ogni } i < n$$

ovvero $\alpha_n x_n = 0$ ed essendo x_n non nullo anche $\alpha_n = 0$ e sono linearmente indipendenti. ∎

Corollario 9.6.3. *Data una qualunque funzione $L \in L(X)$, per ogni $\lambda \in \mathbb{K}$ poniamo*

$$X_\lambda = \{x \in X \mid Lx = \lambda x\}$$

se $\lambda \neq \mu$ allora $X_\lambda \cap X_\mu = \{0\}$, più in generale per ogni $\lambda \in \mathbb{K}$ vale

$$X_\lambda \cap \operatorname{span}\{x \in X \mid Lx = \mu x, \mu \neq \lambda\} = \{0\}$$

Abbiamo detto sopra che lo spettro di un operatore lineare continuo può essere vuoto se siamo in uno spazio normato reale. Però se il nostro operatore è anche compatto allora il suo spettro non sarà mai vuoto dato che vale il seguente risultato:

Proposizione 9.6.4. *Se $T \in K(X,X)$ allora*

$$\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$$

Se poi X ha dimensione infinita allora $0 \in \sigma(T)$.

Dimostrazione. Se $0 \notin \sigma(T)$ allora T^{-1} è continuo e $\mathrm{id}_X = T \circ T^{-1}$ rimane compatto e quindi le palle chiuse sono sequenzialmente compatte, ovvero $\dim X < +\infty$.

Se ora $\lambda \in \sigma(T) \setminus \{0\}$ allora $\mathrm{id}_X - \frac{1}{\lambda} T$ è un operatore di Riesz non invertibile, perciò non è né iniettiva né suriettiva e quindi $\lambda \in \sigma_p(T) \setminus \{0\}$. ∎

Osservazione. Lo spazio degli operatori compatti $K(X,X)$ è un ideale in $L(X,X)$. Infatti per ogni operatore continuo $f : X \to X$ la disuguaglianza $\|f(x)\| \leq \|f\| \|x\|$ implica che:

- f manda insiemi limitati in insiemi limitati;
- f manda insiemi relativamente compatti in insiemi relativamente compatti.

Quindi la composizione di un operatore lineare e continuo con un operatore compatto è ancora un operatore compatto.

Proposizione 9.6.5. *Sia $T \in K(X,X)$ con X di Banach allora l'unico punto di accumulazione di $\sigma(T) \setminus \{0\}$ è proprio 0.*

Dimostrazione. Sia $\lambda_n \in \sigma_n(T) \setminus \{0\}$ convergente ad un certo $\lambda \in \mathbb{K}$ e sia x_n autovettore di λ_n. Definiamo

$$X_n = \operatorname{span}\{x_1, \dots, x_n\}$$

ed essendo linearmente indipendenti avremo che $X_{n-1} < X_n$ e di conseguenza per il lemma dei quasi ortogonali esiste $y_n \in X_n$ di norma 1 tale che $\operatorname{dist}(y_n, X_{n-1}) \geq \frac{1}{2}$.

Dimostriamo che $(T - \lambda_n \mathrm{id}_X)(X_n) \leq X_{n-1}$, infatti

$$(T - \lambda_n \mathrm{id}_X) \left(\sum_{i=1}^{n} \alpha_i x_i \right) = \sum_{i=1}^{n-1} \alpha_i (\lambda_i - \lambda_n) x_i \in X_{n-1}$$

quindi per $m < n$ si ha $X_{m-1} < X_m \le X_{n-1} < X_n$ e

$$\left\| \frac{Ty_m}{\lambda_m} - \frac{Ty_n}{\lambda_n} \right\| = \left\| \frac{Ty_m - \lambda_m y_m}{\lambda_m} - \frac{Ty_n - \lambda_n y_n}{\lambda_n} + y_m - y_n \right\| \ge \frac{1}{2}$$

Se ora λ_n non tendesse a 0 allora esiste $L > 0$ tale che $\left|\lambda_n\right| > L$ per n abbastanza grande e quindi, a meno di passare ad un'estratta, $\frac{1}{\lambda_n}$ è di Cauchy e

$$\left\| \frac{Ty_m}{\lambda_m} - \frac{Ty_n}{\lambda_n} \right\| = \left\| \frac{Ty_m}{\lambda_m} - \frac{Ty_n}{\lambda_m} + \frac{Ty_n}{\lambda_m} - \frac{Ty_n}{\lambda_n} \right\| \le \frac{\left\|Ty_m - Ty_n\right\|}{L} + \left| \frac{1}{\lambda_m} - \frac{1}{\lambda_n} \right| \|T\|$$

raggiungendo un assurdo poiché T è un operatore compatto. Quindi l'unico valore a cui una successione di autovalori può convergere è 0. ■

Corollario 9.6.6. *Lo spettro di un operatore compatto T è al più numerabile.*

Dimostrazione. Innanzitutto $\sigma_p(T) \subseteq [-\|T\|, \|T\|]$ inoltre

$$\sigma_p(T) \setminus \{0\} = \sigma_p(T) \cap \bigcup_{i=2}^{n} \left(\left[-\|T\|, -\frac{\|T\|}{n} \right] \cup \left[\frac{\|T\|}{n}, \|T\| \right] \right)$$

e ogni insieme nella forma $\sigma_p(T) \cap \left(\left[-\|T\|, -\frac{\|T\|}{n} \right] \cup \left[\frac{\|T\|}{n}, \|T\| \right] \right)$ deve essere necessariamente finito. Quindi lo spettro è al più numerabile. ■

Lavorando come nel teorema 8.4.5 per ogni polinomio p abbiamo che

$$p\left[\sigma_p(T)\right] \subseteq \sigma_p[p(T)] \tag{9.6.1}$$

Dimostriamo ora che se $\mathbb{K} = \mathbb{C}$ allora le inclusioni sono uguaglianze. Sia $\lambda \in \sigma_p[p(T)]$ e per assurdo $\lambda \notin p[\sigma_p(T)]$. Dal teorema fondamentale dell'algebra esistono $l, l_1, l_2, \ldots, l_k \in \mathbb{C}$ non necessariamente distinti tali che

$$p(x) - \lambda = l(x - l_1)(x - l_2) \cdots (x - l_k) \tag{9.6.2}$$

con $l_i \notin \sigma_p(T)$ altrimenti $\lambda = p(l_i)$. Inoltre esiste $x \ne 0$ tale che $p(T)(x) = \lambda x$. Ma $T - l_i \mathrm{id}_X$ è iniettivo per ogni i e quindi si avrebbe $x = 0$ assurdo.

Capitolo 10

Spazi di Hilbert

10.1 Spazi prehilbertiani e di Hilbert

Definizione 10.1.1. Sia H spazio vettoriale su $\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$, un'applicazione $(\cdot : \cdot) : H \times H \to \mathbb{K}$ è un **prodotto scalare** se e solo se valgono

1. $(x : x) \geq 0$ e vale l'uguaglianza se e solo se $x = 0$;
2. Per ogni $u, v, w \in H$, $\alpha, \beta \in \mathbb{K}$ vale $\left(w : \alpha u + \beta v\right) = \alpha\,(w : u) + \beta\,(w : v)$ in altre parole è lineare nella seconda variabile;
3. $(u : v) = (v : u)^{\dagger}$.

Osserviamo che se $\mathbb{K} = \mathbb{R}$ allora il punto 3 diventa

$$(u : v) = (v : u)$$

se però $\mathbb{K} = \mathbb{C}$ il prodotto scalare non risulta lineare nella prima variabile in quanto si ha

$$\left(\alpha u + \beta v : w\right) = \left(w : \alpha u + \beta v\right)^{\dagger} = \alpha^{\dagger}(w : u)^{\dagger} + \beta^{\dagger}(w : v)^{\dagger} = \alpha^{\dagger}\,(u : w) + \beta^{\dagger}\,(v : w)$$

Nella prossima sezione proveremo almeno in parte a come ovviare a questo inconveniente quando $\mathbb{K} = \mathbb{C}$.

Se H è uno spazio prehilbertiano allora possiamo definire la quantità

$$\|x\| = \sqrt{(x : x)}$$

e dimostreremo poco più avanti che essa è a tutti gli effetti una norma su H. Per definizione di prodotto scalare avremo chiaramente che $\|x\| = 0$ se e solo se $x = 0$ e inoltre $\|\alpha x\| = \sqrt{\alpha\alpha^{\dagger}\,(x : x)} = |\alpha|\,\|x\|$ e inoltre vale

$$\begin{aligned}\left\|x + y\right\|^2 = \left(x + y : x + y\right) &= (x : x) + \left(x : y\right) + \left(y : x\right) + \left(y : y\right)\\ &= \|x\|^2 + \left\|y\right\|^2 + 2\,\mathrm{Re}\left(x : y\right)\end{aligned}$$

Per dimostrare la disuguaglianza triangolare abbiamo bisogno del seguente risultato.

Proposizione 10.1.2 (Disuguaglianza di Cauchy-Schwarz)**.** *Per ogni $x, y \in H$ si ha*

$$\left|\left(x : y\right)\right| \leq \|x\|\,\left\|y\right\|$$

Dimostrazione. Prendiamo $\lambda \in \mathbb{R}$ e $k \in \mathbb{K}$ con $|k| = 1$ allora vale

$$0 \le \left\|\lambda k x - y\right\|^2 = \lambda^2 \|x\|^2 - 2\lambda \operatorname{Re}\left(k\left(x : y\right)\right) + \left\|y\right\|^2$$

quindi il discriminante $[\operatorname{Re}(k\left(x : y\right))]^2 - \|x\|^2 \left\|y\right\|^2$ deve essere necessariamente minore o uguale a 0 per ogni k. Se $\left(x : y\right) \neq 0$ poniamo

$$k = \frac{\left(y : x\right)}{\left|\left(x : y\right)\right|}$$

e quindi $\operatorname{Re}(k\left(x : y\right)) = \left|\left(x : y\right)\right|$ da cui segue immediatamente la disuguaglianza di Cauchy-Schwartz. ■

Proposizione 10.1.3. *Per ogni $x, y \in H$ si ha*

$$\left\|x + y\right\| \le \|x\| + \left\|y\right\|$$

Dimostrazione.

$$\left\|x + y\right\|^2 \le \|x\|^2 + \left\|y\right\|^2 + 2\|x\|\left\|y\right\| = (\|x\| + \left\|y\right\|)^2$$

■

Quindi ogni spazio prehilbertiano è anche uno spazio normato, dunque possiamo dotare H della topologia generata dalla norma $\|x\| = \sqrt{(x : x)}$.

Corollario 10.1.4. *L'applicazione*

$$(u, v) \in H \times H \to (u : v) \in \mathbb{K}$$

è continua con $H \times H$ dotato della topologia prodotto.

Dimostrazione. Presi $a, b, c, d \in H$ dalla disuguaglianza di Cauchy-Schwartz avremo che

$$\left|\left(a : b\right) - \left(c : d\right)\right| = \left|\left(a - c : b\right) + \left(c : b - d\right)\right| \le \|a - c\|\,\|b\| + \|c\|\,\|b - d\|$$

se adesso facciamo tendere la coppia (a, b) a (c, d) dalla definizione di topologia prodotto seguirà che $b \to d$ in H e dunque $\|b\| \to \|d\|$. Applicando la disuguaglianza di sopra avremo così che $\left(a : b\right) \to \left(c : d\right)$ e dunque il prodotto scalare è continuo. ■

Definizione 10.1.5. Uno spazio prehilbertiano è **di Hilbert** se e solo se il relativo spazio normato è di Banach.

Definizione 10.1.6. Dati due spazi H, K prehilbertiani su $\mathbb{K}$ diremo che l'applicazione lineare $F : H \to K$ è **ortogonale** se e solo se per ogni $x, y \in H$

$$\left(x : y\right) = \left(F(x) : F(y)\right)$$

Osservazione. L'applicazione F è di conseguenza anche una isometria tra H e K visti come spazi normati, e in particolare è un omeomorfismo. Più avanti vedremo che ogni isometria lineare tra due spazi prehilbertiani è anche ortogonale.

Sia H uno spazio prehilbertiano complesso, allora possiamo dotare $H_{\mathbb{R}}$ (che ricordiamo essere H dotato di una struttura di spazio vettoriale reale) del seguente prodotto scalare reale

$$(x:y)_{\mathbb{R}} = \operatorname{Re}(x:y)$$

in particolare se H è uno spazio di Hilbert anche $H_{\mathbb{R}}$ è uno spazio di Hilbert.

Proposizione 10.1.7. *Sia H uno spazio vettoriale complesso e $[\cdot : \cdot]$ un prodotto scalare reale su $H_{\mathbb{R}}$ tale che*

$$[x : iy] = -[ix : y] \tag{10.1.1}$$

per ogni $x, y \in H$. Allora esiste un unico prodotto scalare $(\cdot:\cdot)$ su H tale che $[x : y] = \operatorname{Re}(x:y)$ per ogni $x, y \in H$.

Dimostrazione. L'unicità discende immediatamente dall'uguaglianza

$$\operatorname{Im}(x:y) = i\operatorname{Re}(x:iy) = i[x : iy]$$

Prendiamo ora un prodotto scalare reale $[x : y]$ su $H_{\mathbb{R}}$ che soddisfi la (10.1.1), definiamo

$$(x:y) = [x : y] + i[x : iy]$$

se $x = y$ allora avremo dalla che $[ix : x] = [x : ix] = -[ix : x]$ e quindi $[x : ix] = 0$ e $(x:x) \geq 0$ e si annulla se e solo se $x = 0$.

Abbiamo inoltre

$$(x:y) = [x : y] + i[x : iy] = [y : x] - i[y : ix] = (y:x)^{\dagger}$$

mentre l'applicazione $x \to (x:y)$ è $\mathbb{R}$-lineare per ogni $y \in H$. Inoltre

$$(ix:y) = [ix : y] + i[ix : iy] = -[x : iy] - i[x : -y] = i(x:y)$$

e quindi è anche $\mathbb{C}$-lineare. ■

Se H è uno spazio prehilbertiano (reale o complesso) allora non è difficile verificare che vale l'uguaglianza

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{10.1.2}$$

per ogni $x, y \in H$, detta **uguaglianza del parallelogramma**. In realtà si può dimostrare che vale anche il viceversa, ovvero che se X è uno spazio normato (reale o complesso) che soddisfa la (10.1.2) allora possiede un prodotto scalare la cui norma indotta coincide con la norma propria di X.

La dimostrazione di tale risultato è piuttosto lunga, e considereremo separatamente il caso reale e quello complesso.

Teorema 10.1.8 (Fréchet–von Neumann–Jordan, Caso reale)**.** *Sia $(X, \|\cdot\|)$ uno spazio normato reale, allora esiste un prodotto scalare $(\cdot:\cdot)$ su X tale che $(x:x) = \|x\|^2$ se e solo se vale la* (10.1.2). *In tal caso vale l'uguaglianza*

$$(x:y) = \frac{\|x+y\|^2 - \|x-y\|^2}{4}$$

Se X è uno spazio prehilbertiano allora

$$\left\|x+y\right\|^2-\left\|x-y\right\|^2=\|x\|^2+\left\|y\right\|^2+2\left(x:y\right)-\|x\|^2-\left\|y\right\|^2+2\left(x:y\right)=4\left(x:y\right)$$

Sia X uno spazio normato la cui norma soddisfi la (10.1.2), per ogni $x,y\in X$ poniamo

$$\left(x:y\right)=\frac{\left\|x+y\right\|^2-\|x\|^2-\left\|y\right\|^2}{2}=\frac{\left\|x+y\right\|^2-\left\|x-y\right\|^2}{4}$$

Chiaramente $\left(x:y\right)=\left(y:x\right)$ e $(x:x)=\|x\|^2$, dobbiamo solo dimostrare le proprietà di linearità e di omogeneità, procediamo per gradi.

- Dimostriamo che $\left(x:2y\right)=2\left(x:y\right)$ e $\left(x:-y\right)=-\left(x:y\right)$. Infatti abbiamo

$$\begin{aligned}\left(x:2y\right)&=\frac{\left\|y+\left(x+y\right)\right\|^2-\|x\|^2-4\left\|y\right\|^2}{2}\\&=\frac{2\left\|x+y\right\|^2+2\left\|y\right\|^2-\left\|\left(x+y\right)-y\right\|^2-\|x\|^2-4\left\|y\right\|^2}{2}\\&=\left\|x+y\right\|^2-\|x\|^2-\left\|y\right\|^2\\&=2\left(x:y\right)\end{aligned}$$

  mentre

$$\begin{aligned}\left(x:-y\right)&=\frac{\left\|x-y\right\|^2-\|x\|^2-\left\|y\right\|^2}{2}\\&=\frac{\|x\|^2+\left\|y\right\|^2-\left\|x+y\right\|^2}{2}\\&=-\left(x:y\right)\end{aligned}$$

- Dimostriamo l'additività nella seconda coordinata. Per ogni $x,y,z\in X$ applicando più volte l'uguaglianza del parallelogramma

$$\begin{aligned}2\|x\|^2+2\left\|y\right\|^2&=\left\|x+y\right\|^2+\left\|x-y\right\|^2\\&=\left\|x+y\right\|^2+2\|x+z\|^2+2\left\|y+z\right\|^2-\left\|x+y+2z\right\|^2\\&=\left\|x+y\right\|^2+2\|x+z\|^2+2\left\|y+z\right\|^2-2\left\|x+y+z\right\|^2\\&-2\|z\|^2+\left\|x+y\right\|^2\end{aligned}$$

  ovvero

$$\left\|x+y+z\right\|^2+\|x\|^2+\left\|y\right\|^2+\|z\|^2=\left\|x+y\right\|^2+\|x+z\|^2+\left\|y+z\right\|^2$$

  e così abbiamo che

$$\begin{aligned}2\left(x:y+z\right)&=\left\|x+y+z\right\|^2-\|x\|^2-\left\|y+z\right\|^2\\&=\left(\left\|x+y\right\|^2-\|x\|^2-\left\|y\right\|^2\right)+\left(\|x+z\|^2-\|x\|^2-\|z\|^2\right)\\&=2\left(x:y\right)+2\left(x:z\right)\end{aligned}$$

- Per induzione $(x : ny) = n(x : y)$ per ogni $n \in \mathbb{N}$. Ma allora

$$\left(x : \frac{m}{n} y\right) = \frac{m}{n}(x : y)$$

per ogni $m \in \mathbb{Z}, n \in \mathbb{N}$. Sia ora $\lambda \in \mathbb{R}$ allora esisterà una successione $\lambda_n \in \mathbb{Q}$ tale che $\lambda_n \to \lambda$, dunque

$$\begin{aligned}\lambda(x : y) = \lim_{n\to+\infty}(x : \lambda_n y) = \lim_{n\to+\infty} \frac{\|x + \lambda_n y\|^2 - \|x\|^2 - \|\lambda_n y\|^2}{2} \\ = \frac{\|x + \lambda y\|^2 - \|x\|^2 - \|\lambda y\|^2}{2}\end{aligned}$$

e quindi $\lambda(x : y) = (x : \lambda y)$.

La dimostrazione del caso reale è conclusa.

Teorema 10.1.9 (Fréchet–von Neumann–Jordan, Caso complesso)**.** *Sia* $(X, \|\cdot\|)$ *uno spazio normato complesso, allora esiste un prodotto scalare* $(\cdot : \cdot)$ *su* X *tale che* $(x : x) = \|x\|^2$ *se e solo se vale la* (10.1.2)*. In tal caso vale l'uguaglianza*

$$(x : y) = \frac{\|x + y\|^2 - \|x - y\|^2 + i\|ix + y\|^2 - i\|ix - y\|^2}{4}$$

Dimostrazione. Se X è uno spazio prehilbertiano allora

$$\begin{aligned}\|x + y\|^2 - \|x - y\|^2 &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(x : y) - \|x\|^2 - \|y\|^2 + 2\operatorname{Re}(x : y) \\ &= 4\operatorname{Re}(x : y) \\ \|ix + y\|^2 - \|ix - y\|^2 &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(ix : y) - \|x\|^2 - \|y\|^2 \\ &\quad + 2\operatorname{Re}(ix : y) \\ &= 4\operatorname{Re}[-i(x : y)] = 4\operatorname{Im}(x : y)\end{aligned}$$

Se X è uno spazio normato complesso che soddisfa l'uguaglianza del parallelogramma allora anche $X_{\mathbb{R}}$ lo soddisfa. Quindi per il risultato precedente $X_{\mathbb{R}}$ è dotato di un prodotto scalare reale il quale induce un unico prodotto scalare complesso su X, e questo prodotto scalare assume la medesima forma dell'enunciato. ■

Questo teorema ci permette così di vedere gli spazi prehilbertiani non più solamente come spazi vettoriali dotati di un prodotto scalare, ma come spazi normati la cui norma soddisfa l'identità del parallelogramma.

Corollario 10.1.10. *Se* $f : H \to K$ *è una isometria lineare tra due spazi prehilbertiano* H *e* K *allora* f *è anche ortogonale. In particolare se due spazi prehilbertiani sono isomorfi se visti come spazi normati allora possono essere identificati anche come spazi prehilbertiani.*

Proposizione 10.1.11. *Gli spazi di Hilbert sono uniformemente convessi, in particolare sono riflessivi.*

Dimostrazione. Siano $x, y \in H$ con $\|x\|, \|y\| \leq 1$ e sia $\epsilon > 0$ tale che $\|x-y\| > \epsilon$. Per l'uguaglianza del parallelogramma avremo allora che

$$\left\|\frac{x+y}{2}\right\|^2 = \frac{\|x\|^2 + \|y\|^2}{2} - \frac{\|x-y\|^2}{4} < 1 - \frac{\epsilon^2}{4}$$

e quindi prendendo $\delta = 1 - \sqrt{1-\epsilon^2/4}$ si ha la tesi. ■

Sia $x \in H$ elemento fissato, allora l'applicazione

$$L_y : x \in H \to (y : x) \in \mathbb{K}$$

è lineare per definizione di prodotto scalare e continua dalla disuguaglianza di Cauchy-Schwartz con

$$\|L_y\| = \|y\|$$

in quanto $L_x(x) = \|x\|^2$.

10.2 L'operazione di coniugio

Consideriamo innanzitutto la funzione f che fa da $\mathbb{C}$ in $\mathbb{C}$ definita nella seguente maniera:

$$f(z) = (1+i)z$$

Chiaramente f è una funzione lineare che a prima vista sembrerebbe rappresentata dalla costante $1+i$. In realtà poiché il prodotto scalare standard definito su $\mathbb{C}$ è $(x:y) = x^\dagger y$ la costante che meglio rappresenta f è in realtà $1-i$ poiché $f = L_{1-i}$, invece la costante che rappresenta if non è $i(1-i) = 1+i$ ma $(-i)(1-i) = -1-i$ e perciò l'applicazione che ad ogni funzione lineare da $\mathbb{C}$ in sé associa la relativa costante non è lineare. Questo esempio ci mostra che se H è uno spazio prehilbertiano complesso allora non è del tutto ragionevole identificare gli elementi dello spazio duale H^* con i vettori di H.

Per questo motivo definiamo a partire da uno spazio normato H su $\mathbb{C}$ lo spazio normato $H^\dagger$ sempre su $\mathbb{C}$ che avrà gli stessi elementi e la stessa operazione di somma di H, ma il cui prodotto esterno coincida con il prodotto esterno di partenza con il coniugato dello scalare. Più formalmente diremo che uno spazio normato $H^\dagger$ su $\mathbb{C}$ è un **coniugato** di H se e solo se esiste una isometria biettiva

$$\dagger : x \in H \leftrightarrow x^\dagger \in H^\dagger$$

tale che per ogni $x, y \in H$ e per ogni $\alpha \in \mathbb{K}$

$$(x+y)^\dagger = x^\dagger + y^\dagger \qquad (\alpha x)^\dagger = \alpha^\dagger x^\dagger \qquad \|x^\dagger\| = \|x\|$$

dove si ricorda che $\alpha^\dagger$ è il complesso coniugato di α. Dato che due coniugati di uno stesso spazio H sono isomorfi tra loro possiamo fissarne uno $H^\dagger$ che chiameremo semplicemente **coniugato** di H. Se $\mathbb{K} = \mathbb{R}$ poniamo $H^\dagger = H$ e $z^\dagger = z$ per uniformare la notazione successiva in modo che sia valida sia per spazi reali che complessi.

Chiaramente se H fosse uno spazio prehilbertiano allora possiamo rendere anche $H^\dagger$ prehilbertiano dotandolo del prodotto scalare seguente

$$\left(u^\dagger : v^\dagger\right) = (v : u) = (u : v)^\dagger$$

Esso soddisfa chiaramente il primo e il terzo assioma dei prodotti scalari, per quanto riguarda la seconda ci basta osservare solo che

$$\left(u^\dagger : \alpha v^\dagger\right) = \left(u^\dagger : \left(\alpha^\dagger v\right)^\dagger\right) = \left(u : \alpha^\dagger v\right)^\dagger = \alpha (u : v)^\dagger = \alpha \left(u^\dagger : v^\dagger\right)$$

Inoltre $\left\|x^\dagger\right\| = \|x\|$ e $\dagger$ è perciò anche un omeomorfismo, in particolare H è uno spazio di Hilbert se e solo se $H^\dagger$ lo è.

Esempio 10.2.1. Preso $H = \mathbb{C}$ visto come $\mathbb{C}$-spazio vettoriale con il prodotto scalare usuale $\left(x + iy : a + ib\right) = (x + iy)(a - ib) = (ax + by) + i(ay - bx)$ allora possiamo prendere $H^\dagger = \mathbb{C}$ ponendo

$$(x + iy)^\dagger = x - iy$$

ovvero l'operazione di coniugio usuale.

Esempio 10.2.2. Prendiamo adesso uno spazio di Hilbert su $\mathbb{C}$ generico $(H, +, \cdot)$, dove abbiamo evidenziato le operazioni di somma e prodotto esterno, e consideriamo l'applicazione $* : \mathbb{C} \times H \to H$ definita in modo tale che

$$\alpha * z = \alpha^\dagger z$$

Non è difficile verificare che lo spazio $\bar{H} = (H, +, *)$ è ancora uno spazio vettoriale su $\mathbb{C}$, inoltre per ogni $z \in H$ se poniamo $z^\dagger = z \in \bar{H}$ avremo che

$$(\alpha z)^\dagger = \alpha z = \alpha^\dagger * z = \alpha^\dagger * z^\dagger$$

e quindi $\bar{H} = H^\dagger$ in quanto l'operazione di prodotto esterno di $\bar{H}$ è $*$ e non $\cdot$ che invece è proprio di H. Poiché H e $\bar{H}$ hanno gli stessi elementi il lettore potrebbe confondersi con i passaggi utilizzati, la cosa principale da tenere in mente è che cambia la definizione di prodotto esterno, e dunque la struttura di spazio vettoriale è distinta.

Definiamo anche una forma bilineare tra $H^\dagger$ e H tale che per ogni $u^\dagger \in H^\dagger$, $v \in H$

$$\left\langle u^\dagger : v \right\rangle = (u : v)$$

la linearità sulla prima componente segue infatti da

$$\left\langle \alpha u^\dagger : v \right\rangle = \left\langle \left(\alpha^\dagger u\right)^\dagger : v \right\rangle = \left(\alpha^\dagger u : v\right) = \alpha \left\langle u^\dagger : v \right\rangle$$

Rispetto al prodotto scalare questa forma risulta essere lineare su entrambe le variabili, al costo però di considerare due spazi differenti e non uno solo. Si noti che abbiamo usato lo stesso simbolo per la forma bilineare tra X^* e X con X spazio normato (quindi è ben definita anche su spazi prehilbertiani), questa coincidenza non è casuale come vedremo più avanti.

Poiché anche $H^\dagger$ è uno spazio prehilbertiano possiamo definire nella stessa maniera lo spazio $H^{\dagger\dagger} = \left(H^\dagger\right)^\dagger$ i cui elementi sono nella forma $x^{\dagger\dagger} = \left(x^\dagger\right)^\dagger$ per ogni $x \in H$. Allora per ogni $\alpha \in \mathbb{K}$

$$\alpha x^{\dagger\dagger} = \left(\alpha^\dagger x^\dagger\right)^\dagger = (\alpha x)^{\dagger\dagger}$$

e per ogni $u, v \in H$

$$\left(u^{\dagger\dagger} : v^{\dagger\dagger}\right) = \left(v^\dagger : u^\dagger\right) = (u : v)$$

quindi H e $H^{\dagger\dagger}$ sono spazi prehilbertiani equivalenti e possono essere identificati, in particolare porremo $u^{\dagger\dagger} = u$ per ogni $u \in H$.

Questo implica che la forma bilineare canonica su $H^{\dagger\dagger} \times H^{\dagger}$ in realtà è su $H \times H^{\dagger}$, e sarà definita in questa maniera

$$\left\langle u : v^{\dagger} \right\rangle = \left(u^{\dagger} : v^{\dagger} \right) = (v : u) = \left\langle v^{\dagger} : u \right\rangle = \left\langle u^{\dagger} : v \right\rangle^{\dagger}$$

In generale per ogni $u \in H$ e $\mathbf{x} \in H^{\dagger}$ avremo le seguenti uguaglianze

$$\langle \mathbf{x} : u \rangle = \langle u : \mathbf{x} \rangle \qquad \left\langle \mathbf{x}^{\dagger} : u^{\dagger} \right\rangle = \langle \mathbf{x} : u \rangle^{\dagger}$$
$$\langle \mathbf{x} : u \rangle = \left(\mathbf{x}^{\dagger} : u \right) \qquad \left(\mathbf{x}^{\dagger} : u \right) = \left(u^{\dagger} : \mathbf{x} \right)$$

Dati due spazi normati H, K su $\mathbb{C}$ anche $L(H,K)$ è uno spazio normato su $\mathbb{C}$, e quindi ha perfettamente senso definirne il coniugato $L(H,K)^{\dagger}$. In realtà possiamo definire esplicitamente il coniugato di tale spazio in una forma che ci sarà utile in seguito grazie al seguente risultato:

Proposizione 10.2.3. *Lo spazio $L\left(H^{\dagger}, K^{\dagger}\right)$ è un coniugato di $L(H,K)$.*

Dimostrazione. Definiamo l'applicazione $\dagger : L(H,K) \to L\left(H^{\dagger}, K^{\dagger}\right)$ in modo tale che per ogni $f \in L(H,K)$ e per ogni $v \in H$

$$f^{\dagger}\left(v^{\dagger}\right) = f(v)^{\dagger} \tag{10.2.1}$$

notiamo innanzitutto che effettivamente $f^{\dagger} \in L\left(H^{\dagger}, K^{\dagger}\right)$ ovvero che è lineare e continua. Difatti si ha

$$f^{\dagger}\left(\alpha v^{\dagger}\right) = f^{\dagger}\left[\left(\alpha^{\dagger} v\right)^{\dagger}\right] = \alpha f(v)^{\dagger} = \alpha f^{\dagger}\left(v^{\dagger}\right)$$
$$\left\| f^{\dagger}\left(v^{\dagger}\right) \right\| = \| f(v) \| \leq \| f \| \, \| v^{\dagger} \|$$

L'applicazione $\dagger$ appena definita è chiaramente iniettiva ed additiva, inoltre per ogni $\alpha \in \mathbb{C}$ si ha

$$(\alpha f)^{\dagger}\left(v^{\dagger}\right) = [\alpha f(v)]^{\dagger} = \alpha^{\dagger} f(v)^{\dagger} = \left(\alpha^{\dagger} f^{\dagger}\right)\left(v^{\dagger}\right)$$

ovvero $(\alpha f)^{\dagger} = \alpha^{\dagger} f^{\dagger}$. Per quanto riguarda la suriettività essendo $H^{\dagger\dagger} = H$ e $K^{\dagger\dagger} = K$ avremo che $f^{\dagger\dagger} = f$ e quindi la nostra applicazione è anche biettiva. Per concludere osserviamo che

$$\| f^{\dagger} \| = \sup_{\| x^{\dagger} \| = 1} \left\| f^{\dagger}\left(x^{\dagger}\right) \right\| = \sup_{\| x \| = 1} \| f(x)^{\dagger} \| = \| f \|$$

e quindi $\dagger$ soddisfa tutte le nostre ipotesi. La proposizione è così dimostrata. ∎

Possiamo perciò porre $L(H,K)^{\dagger} = L\left(H^{\dagger}, K^{\dagger}\right)$ grazie all'uguaglianza (10.2.1), in particolare se $K = \mathbb{C} = K^{\dagger}$ allora avremo che $H^{\dagger *} = H^{*\dagger}$.

Proposizione 10.2.4. *Sia H uno spazio prehilbertiano su $\mathbb{K}$, allora l'applicazione $L_H : H^{\dagger} \to H^{*}$ definita come*

$$\left(L_H w^{\dagger}\right)(x) = (w : x) = \left\langle w^{\dagger} : x \right\rangle \tag{10.2.2}$$

è una isometria lineare tra spazi normati.

Dimostrazione. Applicando la disuguaglianza di Cauchy-Schwartz segue immediatamente che $\left\| L_H w^\dagger \right\| \leq \left\| w^\dagger \right\|$ mentre se $w \neq 0$ ponendo $x = w / \|w\|$ avremo che $\left\| L_H w^\dagger \right\| = \left\| w^\dagger \right\|$. La (10.2.2) è chiaramente additiva e inoltre

$$L_H\left(\alpha w^\dagger\right)(x) = \left(\alpha^\dagger w : x\right) = \alpha\,(w : x) = \alpha L_H\left(w^\dagger\right)(x)$$

e quindi L_H è lineare. ■

Chiaramente $L_w = L_H w^\dagger$ per ogni $w \in H$, però l'applicazione $w \to L_w$ non è lineare se $\mathbb{K} = \mathbb{C}$ quindi nel seguito utilizzeremo principalmente L_H. Chiaramente L_H essendo una isometria lineare sarà anche iniettiva mentre non è detto che sia suriettiva, in quanto potrebbero esistere funzionali in H^* che non possono essere espressi tramite prodotti scalari.

O almeno se H è solamente uno spazio prehilbertiano.

10.3 Proiezioni ortogonali

Definizione 10.3.1. Due elementi $x, y \in H$ con H spazio prehilbertiano sono **ortogonali** se e solo se $\left(x : y\right) = 0$ e lo indichiamo con la notazione $x \perp y$.

Per ogni $U \subseteq H$ definiamo

$$U^p = \left\{x \in H \,\middle|\, x \perp y \;\forall y \in U\right\}$$

il quale, come è facile da verificare, è un sottospazio prehilbertiano chiuso di H.

Ricordando la definizione 9.1.7 allora per ogni $x \in H$

$$x \in U^p \Leftrightarrow L_x \in U^\perp$$

Proposizione 10.3.2. *Sia $U \subseteq H$ prehilbertiano allora*

$$U^p = \mathrm{Cl}\left(\mathrm{span}\,\{U\}\right)^p$$

Dimostrazione. Chiaramente $U^p \supseteq \mathrm{Cl}\left(\mathrm{span}\,\{U\}\right)^p$, viceversa se $x \in U^p$ allora $U \subseteq \ker L_x \Leftrightarrow \mathrm{Cl}\left(\mathrm{span}\,\{U\}\right) \leq \ker L_x \Leftrightarrow x \in \mathrm{Cl}\left(\mathrm{span}\,\{U\}\right)^p$ e quindi la tesi. ■

quindi ci possiamo limitare a considerare esclusivamente i sottospazi chiusi.

Proposizione 10.3.3. *Se $x \perp y$ allora $\left\|x + y\right\|^2 = \|x\|^2 + \left\|y\right\|^2$.*

Teorema 10.3.4. *Sia H spazio di Hilbert e M un suo sottoinsieme chiuso non vuoto e convesso, allora esiste un unico elemento x di M di norma minima.*

Dimostrazione. Prendiamo $\delta = \inf\{\|x\| \mid x \in M\}$ quindi per ogni $n \in \mathbb{N}$ esiste $y_n \in M$ tale che $\delta \leq \left\|y_n\right\| < \delta + \frac{1}{n}$. Ma allora per ogni $m, n \in \mathbb{N}$ avremo che $(y_m + y_n)/2 \in M$ e

$$\begin{aligned}\left\|y_m - y_n\right\|^2 = 2\left\|y_m\right\|^2 + 2\left\|y_n\right\|^2 - \left\|y_m + y_n\right\|^2 = 4\left(\frac{\left\|y_m\right\|^2 + \left\|y_n\right\|^2}{2} - \left\|\frac{y_m + y_n}{2}\right\|^2\right)\\ \leq 4\delta\left(\frac{1}{m} + \frac{1}{n}\right) + 2\left(\frac{1}{m^2} + \frac{1}{n^2}\right)\end{aligned}$$

quindi y_n è una successione di Cauchy.

Poiché H è uno spazio di Hilbert allora la successione convergerà ad un certo $x \in M$ tale che $\|x\| = \delta$.

Se per assurdo esistessero due elementi $x, y \in M$ di norma minima allora sempre per convessità vale

$$\|x-y\|^2 = 4\left(\frac{\|x\|^2+\|y\|^2}{2} - \left\|\frac{x+y}{2}\right\|^2\right) \leq 0$$

e quindi $x = y$. ∎

Corollario 10.3.5. *Sia H spazio di Hilbert e M un suo sottoinsieme chiuso non vuoto e convesso. Prendiamo inoltre $x \in H$ allora esiste un unico elemento y di M che minimizza la quantità $\|z-x\|$ al variare di z in M.*

*L'elemento y è detto **proiezione** di x su M e lo indicheremo con il simbolo $P_M x$.*

Dimostrazione. Per ogni $x \in H$ l'insieme $M - x$ è ancora convesso, d'altronde gli spazi di Hilbert sono chiaramente spazi vettoriali topologici quindi $M - x$ è chiuso. Possiamo quindi applicare il teorema precedente all'insieme $M - x$ e quindi esisterà $y - x \in M - x$ tale che $\|y-x\|$ sia minima. ∎

Proposizione 10.3.6. *Nelle stesse ipotesi del teorema precedente allora $\tilde{x}$ è la proiezione di x su M se e solo se per ogni $y \in M$*

$$\operatorname{Re}\left(x-\tilde{x} : y-\tilde{x}\right) \leq 0 \tag{10.3.1}$$

Dimostrazione. Se $M = \{\tilde{x}\}$ allora è banale, quindi possiamo supporre che esistano almeno due elementi distinti in M.

Se $\tilde{x}$ è la proiezione di x su M sia $y \in M \setminus \{\tilde{x}\}$ e definiamo l'applicazione

$$\varphi : t \in [0,1] \to \left\|x - [(1-t)\tilde{x} + ty]\right\|^2$$

che ha minimo in $t = 0$. Infine

$$\varphi(t) = \left\|x-\tilde{x}-t(y-\tilde{x})\right\|^2 = \|x-\tilde{x}\|^2 - 2t\operatorname{Re}\left(x-\tilde{x} : y-\tilde{y}\right) + t^2\left\|y-\tilde{x}\right\|$$

che ha derivata

$$\varphi'(t) = -2\operatorname{Re}\left(x-\tilde{x} : y-\tilde{y}\right) + 2t\left\|y-\tilde{x}\right\|$$

e perciò deve essere $\varphi'(0) \geq 0$.

Supponendo ora che vale la (10.3.1), in tal caso $\varphi'(0) \geq 0$ e $\varphi''(t) > 0$ per ogni t. Dunque φ' è strettamente crescente perciò $\varphi(0) < \varphi(1)$ ovvero $\|x-\tilde{x}\| < \|x-y\|$. ∎

Corollario 10.3.7. *Sia $M \leq H$ sottospazio chiuso di uno spazio di Hilbert e $x \in H$. Allora l'elemento $\tilde{x} \in M$ è la proiezione di x su M se e solo se*

$$x - \tilde{x} \in M^p$$

Dimostrazione. È immediato constatare che se vale $x - \tilde{x} \in M^p$ allora $\left(x-\tilde{x} : y-\tilde{x}\right) = 0$ per ogni $y \in M$ e quindi è una proiezione. Viceversa supponiamo che $\operatorname{Re}\left(x-\tilde{x} : y-\tilde{x}\right) \leq 0$ per ogni $y \in M$, allora per ogni $z \in M$ e $\lambda \in \mathbb{K}$ con $|\lambda| = 1$ avremo $\lambda z \in M$ e

$$\operatorname{Re}\left(x-\tilde{x} : \lambda z\right) = \operatorname{Re}\left(x-\tilde{x} : \left(\lambda z + \tilde{x}\right) - \tilde{x}\right) \leq 0$$

Ponendo $\lambda = \pm 1$ avremo che $\operatorname{Re}(x-\tilde{x} : z) = 0$, inoltre se $\mathbb{K} = \mathbb{C}$ ponendo $\lambda = \pm i$ avremo anche $\operatorname{Im}(x-\tilde{x} : z) = 0$ e quindi $x - \tilde{x} \in M^p$. ∎

Proposizione 10.3.8. *Sia H spazio di Hilbert e $M \leq H$ sottospazio chiuso, allora esistono e sono uniche le applicazioni lineari continue $P_M : X \to M$ e $Q_M : X \to M^p$ tali che*

$$x = P_M x + Q_M x \in M + M^p$$

e tale scrittura è unica per x.

Inoltre se M è diverso da $\{0\}$ e da H avremo che $\|P_M\| = \|Q_M\| = 1$.

Dimostrazione. Per ogni $x \in H$ indichiamo con $P_M x$ la sua proiezione su x, allora per il corollario precedente $x - P_M x \in M^p$.

Se prendiamo $x, y \in H$ e $\alpha, \beta \in \mathbb{R}$ osserviamo che per ogni $z \in M$

$$\left(\alpha x + \beta y - \alpha P_M x - \beta P_M y : z\right) = \alpha\left(x - P_M x : z\right) + \beta\left(y - P_M y : z\right) = 0$$

e quindi $\alpha P_M x + \beta P_M y = P_M(\alpha x + \beta y)$ per l'unicità della proiezione. Posto invece $Q_M x = x - P_M x$ allora Q_M è ancora lineare ed è contenuto interamente in M^p sempre per il corollario 10.3.7.

Siano ora $y \in M$, $z \in M^p$ tali che $y + z = P_M x + Q_M x$ allora $y - P_M x = Q_M x - z \in M \cap M^p = \{0\}$ ovvero $y = P_M x$ e $z = Q_M x$ garantendo l'unicità.

Per concludere osserviamo che per ogni $x \in H$ vale

$$\|x\|^2 = \left\|Px + Qx\right\|^2 = \|Px\|^2 + \left\|Qx\right\|^2 + 2\left(Px : Qx\right) = \|Px\|^2 + \left\|Qx\right\|^2$$

da cui segue la continuità delle proiezioni con norma $\|P\| \leq 1$ e $\left\|Q\right\| \leq 1$. Poiché M è diverso dal sottospazio banale e P fissa gli elementi di M allora $\|P\| = 1$. Poiché M è chiuso e diverso da H allora M^p è diverso da $\{0\}$, infatti preso $z \in H \setminus M$ allora $y = Qz = z - Pz$ deve essere necessariamente diverso dal vettore nullo ed essendo $y = Qy$ la norma di Q vale anch'essa 1. ∎

Corollario 10.3.9. *Se H è uno spazio di Hilbert e $M < H$ è un suo sottospazio chiuso proprio allora $M^p \neq \{0\}$.*

Corollario 10.3.10 (Teorema di Pitagora)**.** *Preso H spazio di Hilbert e $x \in H$ un suo generico elemento, allora*

$$\|x\|^2 = \left\|P_M x\right\|^2 + \left\|Q_M x\right\|^2$$

Corollario 10.3.11. *Se H spazio di Hilbert e M sottospazio chiuso, allora $\left(M^p\right)^p = M$.*

Dimostrazione. Banalmente $M \subseteq \left(M^p\right)^p$, sia $x \in \left(M^p\right)^p$ e P_M, Q_M le proiezioni riferite al sottospazio chiuso M. Perciò

$$0 = \left(x : Q_M x\right) = \left(P_M x : Q_M x\right) + \left(Q_M x : Q_M x\right) = \left\|Q_M x\right\|^2$$

ovvero $x = P_M x \in M$. ∎

Teorema 10.3.12 (Rappresentazione di Riesz)**.** *Se H è uno spazio di Hilbert allora l'applicazione L_H definita in* (10.2.2) *è una isometria lineare biettiva.*

$$H \xleftrightarrow{\dagger} H^{\dagger} \xleftrightarrow{L_H} H^{*} \xleftrightarrow{\dagger} H^{*\dagger} \xleftrightarrow{K_{H^*}} H^{**}$$

Figura 10.1: Isometrie biettive tra spazi di Hilbert

Dimostrazione. Dobbiamo dimostrare solo la suriettività di L_H. Se $f \in H^* \setminus \{0\}$ allora il suo kernel sarà chiuso e non banale, per cui esisterà $z \in \left(\ker f\right)^p$ non nullo tale che $f(z) = 1$. Poniamo allora
$$w = \frac{z}{\|z\|^2} \in (\ker f)^p$$

Per ogni $x \in H$ si ha $x - f(x)z \in \ker f$ dunque $\left(w : x - f(x)z\right) = 0$ e

$$\left\langle L_H w^{\dagger} : x \right\rangle = (w : x) = \left(\frac{z}{\|z\|^2} : f(x)z\right) = \langle f : x\rangle$$

quindi per l'arbitrarietà di x segue che $f = L_H w^{\dagger}$. ■

Da questo risultato possiamo dimostrare la riflessività degli spazi di Hilbert anche senza sfruttare l'ipotesi di uniforme convessità. Innanzitutto indichiamo con $K_H : H^* \to H^{\dagger}$ l'inversa di L_H, per il corollario 10.1.10 H^* è uno spazio di Hilbert con prodotto scalare

$$\left(f : g\right)_* = \left(K_H f : K_H g\right)$$

Ricapitolando se H è uno spazio di Hilbert possiamo sistemare tutte le isometrie ottenute precedentemente come nella figura 10.1. Osserviamo che l'applicazione $L_{H^{*\dagger}} \circ \dagger \circ L_H \circ \dagger$ è una isometria lineare biettiva da H in H^{**}, ma soprattutto coincide con la mappa J_H definita nel capitolo sugli spazi normati in quanto per ogni $f \in H^*$

$$\langle f : x\rangle = \langle K_H f : x\rangle = \left(x^{\dagger} : K_H f\right) = \left(L_H x^{\dagger} : f\right)_* = \left\langle \left(L_H x^{\dagger}\right)^{\dagger} : f\right\rangle = \left\langle L_{H^{*\dagger}}\left[\left(L_H x^{\dagger}\right)^{\dagger}\right] : f\right\rangle$$

e quindi H è uno spazio riflessivo.

Corollario 10.3.13. *Se H è uno spazio di Hilbert allora per ogni $M \leq H$ M^p è isomorfo a $M^{\perp}$ e per ogni $U \leq H^*$ sottospazio chiuso esiste un altro sottospazio chiuso $M \leq H$ isomorfo a U tale che ${}^{\perp}U = M^p$.*

Per gli spazi di Hilbert quindi i tre sottospazi M^p, $M^{\perp}$ e ${}^{\perp}M$ coincidono a meno di isometrie invertibili.

Prima di concludere facciamo altre considerazioni sulle mappe L_H e K_H ottenute precedentemente. Innanzitutto osserviamo che la composizione $T_1 = \dagger \circ L_H$ manda $H^{\dagger}$ in $H^{*\dagger}$ mentre la composizione $T_2 = L_{H^{\dagger}} \circ \dagger$ manda $H^{\dagger}$ in $H^{\dagger *}$. Dato che abbiamo posto $H^{*\dagger} = H^{\dagger *}$ ci chiediamo se è vero che $T_1 = T_2$ e che quindi il prodotto scalare su $H^{*\dagger}$ non risulta ambiguo. La risposta è affermativa in quanto per ogni $v \in H$, $x \in H^{\dagger}$

$$\left\langle \left[L_H x\right]^{\dagger} : v^{\dagger}\right\rangle = \langle L_H x : v\rangle^{\dagger} = \langle x : v\rangle^{\dagger} = \left\langle x^{\dagger} : v^{\dagger}\right\rangle = \left\langle L_{H^{\dagger}} x^{\dagger} : v^{\dagger}\right\rangle$$

e quindi T_1 e T_2 inducono su $H^{*\dagger}$ il medesimo prodotto scalare.

Proposizione 10.3.14. *Sia $T : H \to K$ un qualunque funzionale lineare con H e K spazi di Hilbert, allora*

$$\|T\| = \sup\left\{(Tx : y) \mid x \in H, y \in K, \|x\|, \|y\| \leq 1\right\}$$

Dimostrazione. Innanzitutto abbiamo che $\|y\| = \sup_{\|z\|\leq 1}(y : z)$, quindi avremo che

$$\|T\| = \sup_{\|x\|\leq 1} \|Tx\| = \sup_{\|x\|,\|y\|\leq 1} (Tx : y)$$

e quindi la tesi. ∎

10.4 Serie di Fourier in spazi di Hilbert

Definizione 10.4.1. Sia H spazio di Hilbert su $\mathbb{K}$ e S sottoinsieme non vuoto. S è un **sistema ortonormale** se e solo se per ogni $a, b \in S$ si ha $\|a\| = \|b\| = 1$ e vale la relazione $(a : b) = 0 \Leftrightarrow a \neq b$.

Innanzitutto dato un insieme non vuoto A del tutto arbitrario definiamo lo spazio di Hilbert

$$\ell^2(A, \mathbb{K}) = \left\{ f : A \to \mathbb{K} \;\middle|\; \sum_{\alpha\in A} |f(\alpha)|^2 < +\infty \right\}$$

con prodotto scalare

$$(f, g)_{\ell^2} = \sum_{\alpha\in A} f(\alpha)^\dagger g(\alpha)$$

che non è nient'altro che l'integrale di Bochner rispetto alla misura cardinalità. Consideriamo adesso il sistema ortonormale $S = \{u_\alpha \in H \mid \alpha \in A\}$ con A l'insieme degli indici, definiamo per ogni $x \in H$ l'applicazione

$$\hat{x} : \alpha \in A \to (x : u_\alpha) \in \mathbb{K}$$

In questa sezione dimostreremo che $\hat{x} \in \ell^2(A, \mathbb{K})$ per ogni $x \in H$ e che H e $\ell^2(A, \mathbb{K})$ sono isomorfi.

Teorema 10.4.2 (Disuguaglianza di Bessel)**.** *Per ogni $x \in H$ vale la disuguaglianza*

$$\|x\|^2 \geq \sum_{\alpha\in A} |\hat{x}(\alpha)|^2 = \|\hat{x}\|^2_{\ell^2}$$

Dimostrazione. Sia $A' \subseteq A$ sottoinsieme finito, $M = \operatorname{span}\{u_\alpha \mid \alpha \in A'\}$ allora l'elemento $\tilde{x} = \sum_{\alpha\in A'} \hat{x}(\alpha) u_\alpha$ apparterrà ad M. Per il corollario 4.6.3 M è un sottospazio chiuso di H e per ogni $\alpha \in A'$

$$(x - \tilde{x} : u_\alpha) = (x : u_\alpha) - \hat{x}(\alpha) = 0$$

ovvero $\tilde{x} = P_M x$. Di conseguenza avremo che

$$\|x\|^2 \geq \|\tilde{x}\|^2 = \left\| \sum_{\alpha\in A'} \hat{x}(\alpha) u_\alpha \right\|^2 = \sum_{\alpha\in A'} |\hat{x}(\alpha)|^2$$

Passando all'estremo superiore otteniamo così la tesi. ∎

Corollario 10.4.3. *Fissato un sistema ortonormale* $S = \{u_\alpha \mid \alpha \in A\}$ *per ogni* $x \in H$ *si ha* $\hat{x} \in \ell^2(A,\mathbb{K})$ *e l'applicazione*

$$\mathscr{F} : x \in H \to \hat{x} \in \ell^2(A,\mathbb{K}) \tag{10.4.1}$$

è lineare e continua su H con $\|\mathscr{F}\| \leq 1$*. Inoltre risulta iniettiva sul sottospazio* $\operatorname{span} S$ *(ma non sull'intero spazio H).*

Teorema 10.4.4 (Riesz-Fischer)**.** *L'applicazione* (10.4.1) *è suriettiva.*

Dimostrazione. Prendiamo $\varphi \in \ell^2(A,\mathbb{K})$ generica e per ogni $n \in \mathbb{N}$ definiamo gli insiemi

$$A_n = \left\{ |\varphi| > \frac{1}{n} \right\}$$

nella dimostrazione della proposizione 1.9.2 si vede che A_n è un insieme finito per ogni $n \in \mathbb{N}$, posto

$$x_n = \sum_{\alpha \in A_n} \varphi(\alpha) u_\alpha \in \operatorname{span}\{S\}$$

dimostriamo che $\hat{x}_n = \mathscr{F}(x_n)$ converge in norma $\ell^2(A,\mathbb{K})$ a φ. Per ogni $\alpha \in A$ si ha allora

$$\hat{x}_n(\alpha) = \begin{cases} \varphi(\alpha) & \text{se } \alpha \in A_n \\ 0 & \text{se } \alpha \notin A_n \end{cases}$$

e in particolare se $\varphi(\alpha) \neq 0$ allora per n sufficientemente grande avremo che $\hat{x}_n(\alpha) = \varphi(\alpha)$. Quindi $\hat{x}_n$ converge puntualmente a φ e inoltre $|\hat{x}_n| leq |\varphi|$ e possiamo così applicare il teorema della convergenza per dimostrare che $\hat{x}_n$ converge in norma ℓ^2 a φ.

Poiché $\mathscr{F}$ è un'isometria su $\operatorname{span} S$ allora

$$\|x_m - x_n\| = \|\hat{x}_m - \hat{x}_m\|_{\ell^2}$$

e quindi per completezza la successione x_n convergerà ad un certo $x \in H$. Dalla continuità di $\mathscr{F}$ segue immediatamente che $\mathscr{F}(x) = \varphi$ e perciò $\mathscr{F}$ è suriettiva. ■

Osservazione. Sia $x \in H$, posto $\varphi = \hat{x}$ allora la successione x_n ottenuta nel teorema 10.4.4 di solito **non** converge a x ma alla **proiezione** di x su $M = \operatorname{Cl}(\operatorname{span}\{S\})$.

Infatti se x_n converge a $y \in H$ allora $y \in M$ e $(x - y : u_\alpha) = x \hat{-} y(\alpha) = \hat{x}(\alpha) - \hat{y}(\alpha) = 0$.

Corollario 10.4.5. *Nelle stesse ipotesi del teorema precedente si ha*

$$\ker \mathscr{F} = S^\perp$$

Dimostrazione.

$$z \in \ker \mathscr{F} \Leftrightarrow \hat{z}(\alpha) = 0 \text{ per ogni } \alpha \in A \Leftrightarrow (z : u_\alpha) = 0 \Leftrightarrow z \in S^\perp$$

■

Basi ortonormali

Definizione 10.4.6. Un sistema ortonormale S è una **base ortonormale** se e solo se risulta massimale rispetto all'inclusione.

Dal lemma di Zorn segue immediatamente che ogni spazio di Hilbert ammette almeno una base ortonormale.

Proposizione 10.4.7. *Sia S un sistema ortonormale, allora le seguenti affermazioni sono tutte equivalenti:*

1. *S è una base ortonormale;*
2. *$S^{\perp} = \{0\}$;*
3. *$\mathrm{Cl}\big(\mathrm{span}\{S\}\big) = H$;*
4. *Per ogni $x \in H$ vale l'**identità di Parseval***
$$\|x\|^2 = \sum_{\alpha \in A} |\hat{x}(\alpha)|^2 = \|\hat{x}\|^2_{\ell^2}$$
5. *$\big(x : y\big) = \sum_{\alpha \in A} \hat{x}(\alpha)^{\dagger} \hat{y}(\alpha)$;*
6. *$\mathscr{F}$ è invertibile su H.*

Dimostrazione. L'equivalenza tra i primi due punti è immediata da verificare in quanto gli eventuali elementi di $S^{\perp}$ potrebbero essere aggiunti al sistema nel caso non fosse massimale. Anche l'equivalenza tra il secondo e il terzo è dimostrabile con facilità.

Dimostriamo ora il punto 5 a partire dal 4: sia H che $\ell^2(A, \mathbb{K})$ sono spazi di Hilbert e quindi verificano l'uguaglianza del parallelogramma, per l'identità di Parseval avremo così che
$$\big(\hat{x} : \hat{y}\big) = \frac{\|\hat{x} + \hat{y}\|^2 - \|\hat{x} - \hat{y}\|^2}{4} = \frac{\|x + y\|^2 - \|x - y\|^2}{4} = \big(x : y\big)$$

Dalla proposizione 10.4.5 e dal teorema 10.4.4 si ha l'equivalenza tra i punti 2 e 6, mentre supponendo vera la 4 dimostrare la 6 non richiede un grande sforzo.

Bisogna a questo punto dimostrare che il punto 3 implichi il 4 per concludere la proposizione. L'applicazione $\mathscr{F}$ è un'isometria su $\mathrm{span}\{S\}$, per ipotesi fissato $x \in H$ esiste una successione $x_n \in \mathrm{span}\{S\}$ convergente a y in H, per continuità la successione $\hat{x_n}$ convergerà a $\hat{y}$ in norma l^2 e quindi per continuità della norma
$$\|\hat{y}\|_{l^2} = \lim_{n \to +\infty} \|\hat{x}_n\|_{l^2} = \lim_{n \to +\infty} \|x_n\| = \|y\|$$
e la dimostrazione è conclusa. ■

Teorema 10.4.8. *Consideriamo uno spazio di Hilbert H non banale, allora H è separabile se e solo se ammette una base ortonormale finita o numerabile.*

Dimostrazione. Supponiamo H separabile con $\{x_n \mid n \in \mathbb{N}\}$ denso in H, dal lemma di Zorn H ammette una base ortonormale $S = \{u_\alpha \mid \alpha \in A\}$. Il lettore può verificare che le palle aperte $B\left(u_\alpha, \frac{1}{\sqrt{2}}\right)$ sono tutte disgiunte tra loro.

Per definizione di densità l'applicazione

$$\alpha \in A \to \min\left\{ n \in \mathbb{N} \,\middle|\, x_n \in B\left(u_\alpha, \frac{1}{\sqrt{2}}\right)\right\} \in \mathbb{N}$$

è ben definita ed iniettiva, quindi A è al più numerabile.

Viceversa se S è una base ortonormale finita o numerabile allora l'insieme composto dalle combinazioni finite di elementi di S con scalari in $\mathbb{Q}$ (o in $\mathbb{Q} + i\mathbb{Q}$) è denso in H oltre che numerabile. ■

Il concetto di somma diretta introdotto all'inizio di queste note nello studio degli spazi di Hilbert risulta essere sovrabbondante, per questa ragione introduciamo un concetto più debole della somma diretta:

Definizione 10.4.9. Sia H spazio di Hilbert e $\{H_n\}_{n\in\mathbb{N}}$ sottospazi chiusi di H, allora H è **somma hilbertiana** di $\{H_n\}_{n\in\mathbb{N}}$ e si indica col simbolo

$$H = \bigoplus_{i=1}^{+\infty}{}^{\mathbb{H}} H_i$$

se e solo se valgono le seguenti proprietà:

- $H = \mathrm{Cl}\left(\mathrm{span}\left\{H_n \,\middle|\, n \in \mathbb{N}\right\}\right)$;
- per ogni $n \neq m$ si ha $H_n \perp H_m$ ovvero $(x : y) = 0$ per ogni $x \in H_n$ e $y \in H_m$ (in particolare $H_n \cap \mathrm{span}\left\{H_m \,\middle|\, m \neq n\right\} = \{0\}$).

10.5 Il teorema di Lax-Milgram

Definizione 10.5.1. Sia H spazio di Hilbert su $\mathbb{K}$ e $T \in L(H) = L(H,H)$. Diremo che T è **coerciva** se e solo se esiste $\beta > 0$ tale che per ogni $x \in H$

$$|(Tx : x)| \geq \beta \|x\|^2$$

Il concetto di operatore coercivo si ricollega a quello di semidefinito positivo che abbiamo introdotto per le algebre di Banach, come vedremo più avanti.

Teorema 10.5.2. *Sia $T \in L(H)$ tale che $|(Tx : x)| \geq \beta \|x\|^2$ per ogni $x \in H$, allora T è biiettiva e $\left\|T^{-1}\right\| \leq 1/\beta$.*

Dimostrazione. Dalla disuguaglianza di Cauchy-Schwartz abbiamo immediatamente che

$$\beta \|u\|^2 \leq \|Tu\| \|u\| \Rightarrow \|u\| \leq \frac{1}{\beta} \|Tu\| \tag{10.5.1}$$

e quindi per il teorema 9.1.13 l'operatore T è iniettivo e $\operatorname{im} T$ è chiuso in H. Se per assurdo $\operatorname{im} T \neq H$ allora per la proposizione 10.3.8 esisterà $y \in \operatorname{im} T^\perp$ non nullo, ma allora

$$\beta \left\|y\right\|^2 \leq \left|(Ty : y)\right| = 0$$

il che è assurdo, quindi T deve essere suriettivo. Dalla (10.5.1) avremo che T^{-1} è anch'esso continuo con $\left\|T^{-1}\right\| \leq 1/\beta$. ■

Teorema 10.5.3 (Lax-Milgram)**.** *Sia $F : H \times H \to \mathbb{K}$ definita sullo spazio di Hilbert H sul campo $\mathbb{K}$ tale che*

1. *le applicazioni $x \to F(x,a)$, $x \to [F(a,x)]^\dagger$ sono lineari per ogni $a \in H$;*
2. *esiste $\alpha > 0$ tale che $\left|F(x,y)\right| \leq \alpha \left\|x\right\| \left\|y\right\|$ per ogni $x,y \in H$;*
3. *esiste $\beta > 0$ tale che $\left|F(x,x)\right| \geq \beta \left\|x\right\|^2$ per ogni $x \in H$.*

Allora per ogni $f \in H^*$ esiste un unico $w \in H$ tale che per ogni $x \in H$

$$f(x) = F(x,w)$$

Dimostrazione. Per ogni $u \in H$ definiamo l'applicazione $g_u : v \in H \to F(v,u) \in \mathbb{K}$, chiaramente g_u è lineare e continua quindi per il teorema di rappresentazione di Riesz esisterà $Tu \in H$ tale che

$$F(v,u) = g_u(v) = (v : Tu)$$

per ogni $u, v \in H$. Per l'unicità nell'asserto del teorema di rappresentazione di Riesz l'applicazione $T : H \to H$ è ben definita, inoltre $T(u+v) = Tu + Tv$ e per ogni $\alpha \in \mathbb{K}$

$$(v : T(\alpha u)) = F(v, \alpha u) = \alpha^\dagger F(v,u) = (v : \alpha Tu)$$

e quindi T è lineare. La continuità di T segue dalla coercività di F in quanto

$$\beta \left\|Tu\right\|^2 \leq \left|F(Tu,u)\right| \leq \alpha \left\|Tu\right\| \left\|u\right\|$$

per ogni $u \in H$.

Abbiamo così che $T \in L(H)$ e $|(Tu : u)| = \left|F(u,u)\right| \geq \beta \left\|u\right\|^2$ dunque per il teorema 10.5.2 T è invertibile. Allora per ogni $f \in H^*$ posto

$$w = T^{-1}\left[\left(R_H f\right)^\dagger\right]$$

avremo per ogni $x \in H$

$$\langle f : x \rangle = \langle R_H f : x \rangle = \left(x : \left(R_H f\right)^\dagger\right) = (x : Tw) = F(x,w)$$

Ora se esistesse un altro $w' \in H$ tale che $f(x) = F(x,w')$ per ogni $x \in H$ allora avremo per la linearità di F

$$0 = \left|F(w-w', w-w')\right| \geq \beta \left\|w-w'\right\|^2$$

e quindi $w = w'$. ■

Proposizione 10.5.4. *Nelle stesse ipotesi del teorema di Lax-Milgram, se $\mathbb{K} = \mathbb{R}$ e $F(x,y) = F(y,x)$ per ogni $x,y \in H$ allora per ogni $f \in H^*$ avremo che $F(w,x) = f(x)$ per ogni x se e solo se w è il minimo del funzionale*

$$J(u) = \frac{1}{2}F(u,u) - f(u)$$

Dimostrazione. Sia $w \in H$ tale che $F(w,x) = f(x)$ per ogni $x \in H$, allora preso un generico $u \in H$ si ha

$$J(w+u) = J(w) + \frac{1}{2}F(u,u) + F(w,u) - f(u) \geq J(w) + \frac{\beta}{2}\|u\|^2$$

quindi $J(w)$ è un minimo di J su H.

Viceversa se w è un punto di minimo di J allora $J(w) \leq J(w+\lambda u)$ per ogni $u \in H$ e per ogni $\lambda \in \mathbb{R}$, ovvero

$$G(\lambda) = \frac{\lambda^2}{2}F(u,u) + \lambda[F(w,u) - f(u)] \geq 0$$

Dunque $G(\lambda)$ è una parabola non negativa passante per l'origine, dunque $F(w,u) - f(u) = 0$ per ogni $u \in H$. ■

Osservazione. Il teorema di Lax-Milgram implica allora che il funzionale $j(w)$ possiede un unico minimo su H. Questa conseguenza ha molte applicazioni in vari ambiti dell'analisi, in particolare nel calcolo delle variazioni.

10.6 Algebra degli operatori

Dati due spazi di Hilbert H e K, presi $T \in L(H,K)$ e $T^* \in L(K^*, H^*)$ il suo aggiunto possiamo definire l'applicazione

$$T^* \in L(K^*, H^*) \leftrightarrow L_H^{-1} \circ T^* \circ L_K \in L\left(K^\dagger, H^\dagger\right)$$

la quale risulta essere una isometria biettiva tra due spazi di Banach. Possiamo dunque porre $T^* \in L\left(K^\dagger, H^\dagger\right)$ e quindi per ogni $x \in H$, $y \in K$ avremo che

$$\left\langle y^\dagger : Tx \right\rangle = \left\langle T^* y^\dagger : x \right\rangle$$

di conseguenza avremo che $T^{*\dagger} \in L(K,H)$ e $(Tx : y) = \left(x : T^{*\dagger} y\right)$.

Proposizione 10.6.1. *Si ha* $T^{*\dagger} = T^{\dagger *}$.

Dimostrazione. Per ogni $x \in H$, $y \in K$ si ha

$$\left(T^{*\dagger} y : x\right) = \left\langle T^* y^\dagger : x \right\rangle = \left\langle y^\dagger : Tx \right\rangle = \left\langle y : T^\dagger x^\dagger \right\rangle^\dagger = \left\langle T^{\dagger *} y : x^\dagger \right\rangle^\dagger = \left(T^{\dagger *} y : x\right)$$

come volevasi dimostrare. ■

Per snellire la notazione scriveremo da questo momento $L(H)$ al posto di $L(H,H)$. Abbiamo visto nel capitolo sugli operatori tra spazi normati che lo spazio $L(H)$ può essere considerato un'algebra di Banach con l'operazione di prodotto interno $ST = S \circ T$. Quando H è uno spazio di Hilbert complesso si verifica facilmente che l'applicazione $T \in L(H) \to T^{*\dagger} \in L(H)$ è una involuzione che rende $L(H)$ una C^*-algebra unitaria. Infatti per ogni $T \in L(H)$ si ha

$$\begin{aligned}
\left\|T^{*\dagger} T\right\| &= \sup\left\{\left(T^{*\dagger} Tx : y\right) \;\middle|\; x, y \in H, \|x\|, \|y\| \leq 1\right\} \\
&= \sup\left\{(Tx : Ty) \;\middle|\; \|x\|, \|y\| \leq 1\right\} \\
&= \sup\left\{\|Tx\|^2 \;\middle|\; \|x\| \leq 1\right\} \\
&= \|T\|^2
\end{aligned}$$

e quindi possiamo applicare tutti i risultati delle C^*-algebre unitarie all'algebra degli operatori lineari e continui $L(H)$ quando H è uno spazio di Hilbert complesso. Se invece H è uno spazio di Hilbert reale non possiamo utilizzare la maggior parte dei risultati del capitolo sulle C^*-algebre, ma dobbiamo prima immergerlo in uno spazio di Hilbert complesso in modo da conservare alcune caratteristiche dello spazio originario.

Sia allora H uno spazio di Hilbert reale (in particolare $(x : y) = (y : x)$ per ogni $x, y \in H$) e poniamo adesso $H_{\mathbb{C}} = H \times H$ e lo rendiamo uno spazio vettoriale con le operazioni standard di somma e prodotto esterno per uno scalare reale. Per renderlo uno spazio prehilbertiano complesso poniamo semplicemente

$$\begin{aligned} i(x,y) &= (-y,x) \\ \left((x,a):(y,b)\right) &= (x:y) + (a:b) + i\left[(x:b)-(y:a)\right] \end{aligned}$$

per ogni $x, y, a, b \in H$. Non è difficile verificare che questo nuovo prodotto scalare è $\mathbb{R}$-lineare nella seconda variabile. Inoltre si ha

$$\begin{aligned} \left((x,a):i(y,b)\right) &= \left((x,a):(-b,y)\right) \\ &= -(x:b) + (a:y) + i\left[(x:y)+(b:a)\right] \\ &= i\left[i\left[(x:b)-(y:a)\right] + (x:y) + (a:b)\right] \\ &= i\left((x,a):(y,b)\right) \\ \left((y,b):(x,a)\right) &= (y:x) + (b:a) + i\left((y:a)-(x:b)\right) \\ &= (x:y) + (a:b) - i\left[(x:b)-(y:a)\right] \\ &= \left((x,a):(y,b)\right)^{\dagger} \\ \left\|(x,a)\right\|^2 &= \|x\|^2 + \|a\|^2 \end{aligned}$$

ed essendo $\|x\| \vee \|a\| \leq \sqrt{\|x\|^2 + \|a\|^2} \leq \sqrt{2}\,(\|x\| \vee \|a\|)$ avremo che $H_{\mathbb{C}}$ è uno spazio di Hilbert la cui topologia indotta dalla norma coincide con la topologia prodotto. Infine l'applicazione $x \in H \to (x,0) \in H_{\mathbb{C}}$ è una isometria $\mathbb{R}$-lineare tra spazi di Hilbert, quindi possiamo considerare $H_{\mathbb{C}}$ come la "complessificazione" di H. Se identifichiamo H con il relativo sottoinsieme di $H_{\mathbb{C}}$ allora per ogni $z \in H_{\mathbb{C}}$ esistono e sono unici $x, y \in H$ tali che $x + iy = z$, in altre parole

$$H_{\mathbb{C}} = H \tilde{\oplus} iH$$

se vediamo $H_{\mathbb{C}}$ come uno spazio vettoriale reale in quanto le relative proiezioni sono chiaramente continue.

Consideriamo adesso una qualunque funzione $T \in L\left(H_{\mathbb{C}}\right)$, per come abbiamo definito la somma continua esisteranno $T_a, T_b \in L\left(H \times H, H\right)$ $\mathbb{R}$-lineari tali che per ogni $x, y \in H$

$$\begin{aligned} T(x+iy) &= T_a(x+iy) + iT_b(x+iy) \\ &= T_a(x) - T_b(y) + iT_a(y) + iT_b(x) \end{aligned} \tag{10.6.1}$$

e quindi possiamo scomporre T a partire da due applicazioni T_a, T_b appartenenti ad $L(H)$. Viceversa se $T_a, T_b \in L(H)$ allora è banale constatare che l'applicazione T definita in (10.6.1) è lineare anche sugli scalari complessi, e per la continuità della somma e del prodotto esterno segue che T è anche continua. In particolare per ogni $T \in L(H)$ possiamo definire il suo complessificato $T_{\mathbb{C}} \in L\left(H_{\mathbb{C}}\right)$ in modo tale che $T_{\mathbb{C}}(x+iy) = T(x) + iT(y)$ (ovvero ponendo

$T_a = T$ e $T_b = 0$). Poiché in $H_{\mathbb{C}}$ si ha $\left\|x+iy\right\| = \sqrt{\|x\|^2 + \left\|y\right\|^2}$ per ogni $x, y \in H$ è chiaro che $\|T\| \leq \left\|T_{\mathbb{C}}\right\|$ mentre per la disuguaglianza opposta ci basta osservare che

$$\left\|T_{\mathbb{C}}(x+iy)\right\|^2 = \|T(x)\|^2 + \left\|T(y)\right\|^2 \leq \|T\|^2 \left(\|x\|^2 + \left\|y\right\|^2\right) = \|T\|^2 \left\|x+iy\right\|^2$$

Infine osserviamo che $S_{\mathbb{C}} = T_{\mathbb{C}}$ se e solo se $S = T$, $(S \circ T)_{\mathbb{C}} = S_{\mathbb{C}} \circ T_{\mathbb{C}}$ e $\left(\mathrm{id}_H\right)_{\mathbb{C}} = \mathrm{id}_{H_{\mathbb{C}}}$, in particolare $T_{\mathbb{C}}$ è invertibile se e solo se lo è anche T con $T_{\mathbb{C}}^{-1} = \left(T_{\mathbb{C}}\right)^{-1}$.

Tutto questo pippone ci è servito per dimostrare che possiamo sempre vedere gli spazi di operatori su uno spazio di Hilbert reale come sottoinsieme dello spazio di operatori del relativo spazio di Hilbert complessificato. Questa nuova "sottoalgebra" è anche $*$-compatibile in quanto per ogni $T \in L(X)$ e per ogni $x, y, a, b \in H$

$$\begin{aligned}\left(Tx + iTa : y + ib\right) &= \left(Tx : y\right) + \left(Ta : b\right) + i\left[\left(Tx : b\right) - \left(Ta : y\right)\right] \\ &= \left(x : T^{*\dagger}y\right) + \left(a : T^{*\dagger}b\right) + i\left[\left(x : T^{*\dagger}b\right) - \left(a : T^{*\dagger}y\right)\right] \\ &= \left(x + ia : T^{*\dagger}y + iT^{*\dagger}b\right)\end{aligned}$$

ovvero $T_{\mathbb{C}}^{*\dagger} = \left(T_{\mathbb{C}}\right)^{*\dagger}$.

Ricapitolando abbiamo il seguente risultato:

Teorema 10.6.2. *Se H è uno spazio di Hilbert complesso allora $L(H)$ è una C^*-algebra unitaria con involuzione $T \to T^{\dagger *}$. Invece se H è uno spazio di Hilbert reale esistono una C^*-algebra unitaria X ed una isometria $\mathbb{R}$-lineare $i : L(H) \to X$ che risulta essere anche moltiplicativa, unitaria e $*$-compatibile. Infine posto $H' = i[l(H)] \subseteq X$ allora avremo che*

$$X_{\mathbb{R}} = H' \dot{\oplus} iH'$$

Possiamo così adattare tutti i risultati ottenute per le C^*-algebre complesse anche agli spazi $L(H)$ quando H è reale. In particolare il corollario 8.6.11 ci dice che:

Teorema 10.6.3. *Se H è uno spazio di Hilbert reale ed $T \in L(H)$ è autoaggiunto allora $\sigma(T) \neq \emptyset$.*

Corollario 10.6.4. *Se H è uno spazio di Hilbert reale ed $T \in L(H)$ è autoaggiunto allora per ogni polinomio $p \in \mathbb{R}[1]$ si ha $p[\sigma(T)] = \sigma[p(T)]$.*

Osservazione. Se T è un'applicazione lineare da H in sé con H prehilbertiano su $\mathbb{C}$ tale che $\left(x : Ty\right) = \left(y : Tx\right)$ per ogni $x, y \in H$ allora T deve essere identicamente nulla.

Infatti avremo per la linearità di T su $\mathbb{C}$

$$i\left(x : Ty\right) = \left(x : T[iy]\right) = \left(iy : Tx\right) = -i\left(x : Ty\right)$$

e quindi $\left(x : Ty\right) = 0$ per ogni $x, y \in H$ e perciò $T \equiv 0$.

Proposizione 10.6.5. *Preso un qualunque operatore $T \in L(H)$ su uno spazio di Hilbert H, definiamo*

$$\mathcal{N}(T) = \left\{ \frac{[x,x]_T}{\|x\|^2} \;\middle|\; x \in H \setminus \{0\} \right\}$$

dove $[x, y]_T = \left(x : Ty\right)$. Allora

$$\sigma(T) \subseteq \mathrm{Cl}(\mathcal{N}(T))$$

Dimostrazione. Sia $\lambda \notin \mathrm{Cl}(\mathcal{N}(T))$ allora $\alpha = \mathrm{dist}(\lambda, \mathcal{N}(T)) > 0$ e

$$\left|(Tx - \lambda x : x)\right| \geq \alpha \left\|x\right\|^2$$

per ogni $x \in H$. Dunque dal teorema di Lax-Milgram avremo che $\lambda \in \rho(T)$. ■

Corollario 10.6.6. *Se H è uno spazio di Hilbert complesso e $T \in L(H)$ è normale allora si hanno le seguenti relazioni:*

1. *T è autoaggiunto se e solo se $(Tx : x) \in \mathbb{R}$ per ogni $x \in H$;*

2. *T è semidefinito positivo se e solo se $(Tx : x) \geq 0$ sempre per ogni $x \in H$.*

Dimostrazione. Se T è autoaggiunto allora $(Tx : x) = (x : Tx) = (Tx : x)^\dagger$ e perciò appartiene ad $\mathbb{R}$, se invece $T = S^{*\dagger} \circ S$ per qualche $S \in L(H)$ allora $(Tx : x) = \|Sx\|^2 \geq 0$. Le implicazioni inverse invece derivano direttamente dalla proposizione 10.6.5 e da quanto detto nel capitolo sulle algebre di Banach. ■

Proposizione 10.6.7. *Prendiamo T autoaggiunto (e quindi il suo spettro è non vuoto e appartiene ad $\mathbb{R}$). Posto*

$$\begin{aligned} m &= \inf_{\|u\|=1} [u,u]_T \\ M &= \sup_{\|u\|=1} [u,u]_T \end{aligned}$$

Allora $\sigma(T) \subset [m, M]$ e $m, M \in \sigma(T)$. Infine se $\sigma(T) = \{0\}$ allora T è nullo.

Dimostrazione. Innanzitutto dalla proposizione 10.6.5 $\sigma(T) \subseteq \left[m, M\right]$ mentre dalla proposizione 8.4.2 sia m che M sono finiti. Dimostriamo ora che M si trova nello spettro. È banale constatare che l'applicazione

$$[\![u,v]\!] = M(u:v) - [u,v]_T = (Mu - Tu : v)$$

è lineare nella seconda variabile, $[\![u,u]\!] \geq 0$ e $[\![u,v]\!] = [\![v,u]\!]^\dagger$. Ragionando come nella dimostrazione della disuguaglianza di Cauchy-Schwartz avremo che $[\![u,v]\!]^2 \leq [\![u,u]\!][\![v,v]\!]$ e quindi

$$\|Mu - Tu\| = \sup_{\|v\|=1} (Mu - Tu : v) \leq C\sqrt{[\![u,u]\!]}$$

con $C = \sqrt{M + \|T\|}$. Per come abbiamo definito M esiste una successione di vettori $x_n \in H$ di norma unitaria tale che $[x_n, x_n]_T \to M$, per la disuguaglianza appena dimostrata avremo così che $Mx_n - Tx_n \to 0$, però se per assurdo fosse $M \in \rho(T)$ allora $M\mathrm{id}_H - T$ dovrebbe avere inversa continua e quindi per continuità anche x_n tenderebbe a 0, assurdo.

Valgono analoghe proprietà anche per m sostituendo $-T$ a T. Infine se lo spettro contiene solo lo zero allora $[u,u]_T = 0$ per ogni $u \in H$ e quindi

$$0 = [u + Tu, u + Tu]_T = 0 + 2[u, Tu]_T + 0 = \|Tu\|^2$$

e quindi T è identicamente nullo. ■

Corollario 10.6.8. *Ogni operatore lineare continuo non identicamente nullo e autoaggiunto possiede un autovalore reale non nullo.*

Ricapitolando i risultati ottenuti precedentemente sullo spettro abbiamo che, dati uno spazio normato X su $\mathbb{K}$ e una funzione $T \in L(X)$ se almeno una delle seguenti condizioni è verificata:

- $\mathbb{K} = \mathbb{C}$,
- X ha dimensione infinita e T è compatto,
- X è uno spazio di Hilbert e T è autoaggiunto,

allora $\sigma(T) \neq \emptyset$. Mettendo però assieme gli ultimi due risultati abbiamo il seguente risultato

Teorema 10.6.9. *Se H è uno spazio di Hilbert (reale o complesso) ed $T \in L(H)$ è non identicamente nullo, compatto e autoaggiunto allora l'insieme $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$ è non vuoto, in particolare possiede sempre almeno un autovalore non nullo.*

Dimostrazione. Poiché T è compatto la proposizione 9.6.4 garantisce l'uguaglianza di sopra tra spettro e spettro puntuale, mentre essendo T autoaggiunto avremo che il suo spettro è non vuoto. Infine la proposizione 10.6.7 ci dice che se lo spettro $\sigma(T)$ deve contenere necessariamente almeno un elemento non nullo, che di conseguenza apparterrà anche allo spettro puntuale. ■

Decomposizione spettrale per operatori compatti

Sia ora $T \in L(H, H)$ operatore autoaggiunto compatto, sia μ_n la successione dei suoi autovalori ponendo $\mu_0 = 0$ e poniamo

$$H_n = \ker(T - \mu_n \mathrm{id}_H)$$

se $n \neq 0$ allora H_n ha dimensione finita per la proposizione 9.5.4. Allora

Teorema 10.6.10 (Decomposizione spettrale)**.** *Sia H spazio di Hilbert e $T \in L(H, H)$ compatto e autoaggiunto allora*

$$H = \bigoplus_{n=0}^{+\infty}{}^{\mathbb{H}} H_n \tag{10.6.2}$$

Dimostrazione. Per il corollario 9.6.3 abbiamo $H_n \cap \operatorname{span}\{H_m \mid m \neq n\} = \{0\}$, posto ora $W = \operatorname{span}\{H_n \mid n \in \mathbb{N}_0\}$ vogliamo dimostrare che è denso. Dato che $T(H_n) = \mu_n H_n$ si ha $T(W) \leq W$, poiché T è autoaggiunto per ogni $x \in W^\perp$ e per ogni $w \in W$ avremo che

$$(Tx : w) = (x : Tw) = 0$$

e dunque $T\left(W^\perp\right) \leq W^\perp$.

Poiché $W^\perp$ è un sottospazio chiuso di H se per assurdo $W^\perp \neq \{0\}$ allora la restrizione di T su $W^\perp$, che indicheremo con $R \in L\left(W^\perp\right)$, è ancora lineare, continua, autoaggiunta e compatta. Osserviamo che ogni autovalore di R è necessariamente un autovalore anche di T e perciò coinciderà con qualche μ_n, il che è impossibile, dunque $\sigma_p(R) = \emptyset$. Poiché R è autoaggiunto il suo spettro non può essere vuoto, ma allora $\sigma(R) = \{0\}$ e quindi R è identicamente nullo.

Abbiamo allora che $W^\perp \subseteq \ker T = H_0 \leq W$ e questo è possibile se e solo se $W^\perp = \{0\}$ assurdo. Dunque $\mathrm{Cl}(W) = H$. ■

Corollario 10.6.11. *Lo spettro residuo di un operatore compatto e autoaggiunto è vuoto.*

Dimostrazione. L'unico elemento dello spettro che potrebbe non appartenere allo spettro puntuale è lo 0. Supponiamo allora che non si trovi nello spettro puntuale, quindi T è iniettivo ovvero $H_0 = \{0\}$. Per la (10.6.2) H è separabile ed ha un sistema denso numerabile composto interamente da autovettori associati ad autovalori non nulli. Questo implica che tutti gli autovettori appartengono a $\operatorname{im} T = (T - 0\mathrm{id}_H)(H)$ e quindi lo 0 appartiene allo spettro continuo. ∎

10.7 Il teorema di rappresentazione di Gelfand-Naimark

Questa sezione sarà interamente dedicata alla dimostrazione del seguente teorema

Teorema 10.7.1 (Gelfand-Naimark)**.** *Data una qualunque C^*-algebra unitaria X esistono uno spazio di Hilbert complesso H ed una isometria lineare $f : X \to L(H)$ che risulta essere anche moltiplicativa e $*$-compatibile rispetto alla struttura standard di $L(H)$ come C^*-algebra.*

Vogliamo quindi costruire uno spazio di Hilbert H complesso a partire dalla struttura di X. Iniziamo introducendo la seguente definizione:

Definizione 10.7.2. Sia X una C^*-algebra unitaria, un'applicazione lineare e continua $\phi \in X^*$ è detta **positiva** se e solo se per ogni $x \in X$ semidefinito positivo si ha $\phi(x) \geq 0$. Se ϕ è positivo e $\phi(\mathbb{e}) = 1$ allora diremo che ϕ è uno **stato** su X.

I funzionali positivi soddisfano una disuguaglianza molto simile alla disuguaglianza di Cauchy-Schwartz: se $\phi \in X^*$ è positivo allora per ogni $x, y \in X$ si ha

$$\left|\phi(x^* y)\right|^2 \leq \phi(x^* x)\phi(y^* y) \tag{10.7.1}$$

Dalla definizione di positività segue infatti che per ogni $\lambda \in \mathbb{C}$

$$0 \leq \phi\left[(\lambda x + y)^*(\lambda x + y)\right] = |\lambda|^2 \phi(x^* x) + \lambda^\dagger \phi(x^* y) + \lambda\phi(y^* x) + \phi(y^* y)$$

ponendo prima $\lambda = 1$ e poi $\lambda = i$ otteniamo le seguenti relazioni:

$$\begin{aligned} \operatorname{Im}\phi(x^* y) &= -\operatorname{Im}\phi(y^* x) \\ \operatorname{Re}\phi(x^* y) &= \operatorname{Re}\phi(y^* x) \end{aligned}$$

e quindi $\phi(x^* y) = \phi(y^* x)^\dagger$, in particolare

$$\phi(x^*) = \phi(x)^\dagger$$

Sostituendo λ con $t\lambda$ dove $t \in \mathbb{R}$ otteniamo l'equazione di una parabola positiva in t e dunque

$$\operatorname{Re}\left[\lambda\phi(x^* y)\right]^2 \leq \phi(x^* x)\phi(y^* y)$$

Ma allora lavorando come della dimostrazione della disuguaglianza di Cauchy-Schwartz otteniamo proprio la (10.7.1). Un altro risultato utile sui funzionali positivi è il seguente:

Proposizione 10.7.3. *Se $\phi \in X^*$ è positivo allora $\|\phi\| = \phi(\mathbb{e})$ e per ogni $x, y \in X$*

$$\phi(x^* y^* y x) \leq \|y\|^2 \phi(x^* x)$$

Dimostrazione. Innanzitutto $\|\phi\| \geq \phi(\mathbb{e})$ mentre per ogni $x \in X$ grazie alla disuguaglianza (10.7.1)

$$|\phi(x)|^2 \leq \phi(x^* x)\phi(\mathbb{e}) \leq \|\phi\| \|x^* x\| \phi(\mathbb{e}) = \|\phi\| \|x\|^2 \phi(\mathbb{e})$$

e quindi $\|\phi\| \leq \phi(\mathbb{e})$.

Siano ora $x, y \in X$ generici, possiamo considerare l'applicazione $\varphi(z) = \phi(x^* z x)$ che risulterà ancora lineare, continua e positiva per definizione di elemento semidefinito positivo. Dunque $\|\varphi\| = \varphi(\mathbb{e}) = \phi(x^* x)$ e così

$$\phi(x^* y^* y x) = \varphi(y^* y) \leq \|\varphi\| \|y^* y\| = \phi(x^* x) \|y\|^2$$

■

Lemma 10.7.4. *Per ogni $\phi \in X^*$ avremo che ϕ è uno stato se e solo se $\|\phi\| = 1 = \phi(\mathbb{e})$.*

Dimostrazione. Se ϕ è uno stato allora la tesi segue immediatamente dalla proposizione precedente. Viceversa supponiamo ora che $\phi \in X^*$ e $\phi(\mathbb{e}) = \|\phi\| = 1$, dimostriamo prima che se $x = x^*$ allora $\phi \in \mathbb{R}$. Per assurdo supponiamo che $l = \operatorname{Im} \phi(x) > 0$ per ipotesi abbiamo $l \leq \|x\|$ e possiamo scegliere un qualunque $0 < r < l$, ora definiamo la seguente funzione su $s \geq 0$

$$h(s) = \sqrt{s^2 + \|x\|^2} - s$$

chiaramente $r \leq h(0)$ mentre $\lim_{s \to +\infty} h(s) = 0$, ora essendo r positivo per la continuità di h esisterà $s \geq 0$ tale che $r = h(s)$. Sia ora $\lambda \in \sigma(x) \subseteq \mathbb{R}$ (x è autoaggiunto) ed in particolare $|\lambda| \leq \|x\|$ abbiamo così

$$|\lambda + is| = \sqrt{\lambda^2 + s^2} \leq \sqrt{\|x\|^2 + s^2} = s + r$$

Ora dato che $\sigma(x + \lambda\mathbb{e}) = \sigma(x) + \lambda$ per ogni $\lambda \in \mathbb{C}$ la proposizione 8.5.3 ed il teorema 8.4.8 implicano che $\|x + is\mathbb{e}\| \leq r + s$. Ma

$$|\phi(x) + is| \geq \operatorname{Im} \phi(x) + s > r + s$$

per linearità e per il fatto che $\phi(\mathbb{e}) = \|\phi\| = 1$ abbiamo $\phi(x) - is = \phi(x - is\mathbb{e})$ raggiungendo così un assurdo.

Sia ora $x \in X$ semidefinito positivo, allora per il lemma 8.7.3 $\| \|x\| \mathbb{e} - x\| \leq \|x\|$ e quindi $\left| \|x\| - \phi(x) \right| \leq \|\phi\| \|x\| = \|x\|$. Dato che abbiamo dimostrato che $\phi(x) \in \mathbb{R}$ la disuguaglianza di sopra implica che $\phi(x) \geq 0$ come volevasi dimostrare. ■

Da questo risultato discende immediatamente il seguente corollario

Corollario 10.7.5. *Se X è una C^*-algebra unitaria allora possiede sempre almeno uno stato.*

Dimostrazione. Basta applicare il corollario 7.2.2 del teorema di Hahn-Banach negli spazi normati al vettore $\mathbb{e}$. ■

Abbiamo ora tutti gli strumenti per costruire H a partire da un qualunque stato ϕ che sappiamo esistere sempre. Preso un qualunque stato ϕ di X definiamo innanzitutto

$$N_\phi = \left\{x \in X \,\middle|\, \phi(x^* x) = 0\right\}$$

chiaramente $\lambda N_\phi \subseteq N_\phi$ per ogni $\lambda \in \mathbb{C}$, se almeno uno tra $x, y \in X$ appartiene ad N_ϕ la (10.7.1) implica che $\phi(x^* y) = \phi(y^* x) = 0$ e perciò $x + y \in N_\phi$ per ogni $x, y \in N_\phi$ e quindi N_ϕ è un sottospazio vettoriale di X che sarà chiuso per la continuità di ϕ.

Ora siano $x \in N_\phi$ ed $y \in X$ generico, per la proposizione 10.7.3 avremo che $\phi(x^* y^* y x) \leq \phi(x^* x) \left\|y\right\|^2 = 0$ e quindi $yx \in N_\phi$ ed N_ϕ è anche un ideale sinistro.

Il fatto che N_ϕ non è in generale un ideale bilatero non permette di rendere X/N_ϕ un'algebra pur possedendo una struttura di spazio di Banach. Possiamo però definirvi comunque il seguente prodotto scalare:

$$\left(x + N_\phi : y + N_\phi\right) = \phi(x^* y)$$

il quale è ben definito in quanto per ogni $a, b \in N_\phi$ si avrebbe

$$\phi[(x + a)^* (y + b)] = \phi(x^* y) + \phi(x^* b) + \phi(a^* y) + \phi(a^* b) = \phi(x^* y)$$

in quanto $\phi(a^* y) = \phi\left[(y^* a)^*\right] = \phi(y^* a)^\dagger = 0$.

Osservazione. Il prodotto scalare su X/N_ϕ induce una topologia che di solito non coincide con la topologia quoziente di X/N_ϕ ciononostante essendo $\left\|x + N_\phi\right\|^2 = \phi(x^* x) \leq \left\|x^* x\right\| = \|x\|^2$ avremo che la topologia indotta da questo prodotto scalare è meno fine della topologia quoziente.

Quindi X/N_ϕ è uno spazio prehilbertiano e quindi il suo completamento H_ϕ è uno spazio di Hilbert. Ancora possiamo definire una nuova applicazione lineare $L_\phi : X \to L(H_\phi)$ in modo tale che per ogni $y \in X$

$$L_\phi(x)[y + N_\phi] = xy + N_\phi$$

Osserviamo anche che

$$\begin{aligned}
L_\phi(xy)[z + N_\phi] &= xyz + N = L_\phi(x)[L_\phi(y)[z + N_\phi]] \\
\left(L_\phi(x^*)[y + N_\phi] : z + N_\phi\right) = \left(x^* y + N_\phi : z + N_\phi\right) = \phi(y^* x z) &= \left(y + N_\phi : xz + N_\phi\right) \\
&= \left(y + N_\phi : L_\phi(x)[z + N_\phi]\right) \\
&= \left(L_\phi(x)^{*\dagger}[y + N_\phi] : z + N_\phi\right)
\end{aligned}$$

e per densità avremo in generale $L_\phi(xy) = L_\phi(x) \circ L_\phi(y)$, $L_\phi(\mathbb{e}) = \mathrm{id}_H$ e $L_\phi(x^*) = L_\phi(x)^{*\dagger}$. Infine per ogni $b \in X$ per cui $\left\|b + N_\phi\right\|^2 = \phi(b^* b) = 1$ abbiamo

$$\left\|L_\phi(x)[b + N_\phi]\right\|^2 = \|xb + N\|^2 = \phi(b^* x^* x b) \leq \phi(b^* b) \|x\|^2$$

e dunque $\left\|L_\phi(x)\right\| \leq \|x\|$ ed L_ϕ è continua.

Ciononostante non è H_ϕ il nostro spazio di Hilbert desiderato in quanto l'applicazione L_ϕ non è in generale iniettiva. In compenso siamo molto vicini in quanto dobbiamo dimostrare solamente che

Lemma 10.7.6. *Se X è una C^*-algebra unitaria allora per ogni $x \in X$ autoaggiunto non nullo esiste uno stato $\phi : X \to \mathbb{C}$ tale che $\left\| L_\phi(x) \right\| = \|x\|$.*

Dimostrazione. Sappiamo già che gli spazi topologici $\operatorname{chr}\left(\overline{\mathbb{C}}(\mathbb{e}, x)\right)$ e $\sigma(x)$ sono omeomorfi e $\|x\| = r(x) = t(x)$. Essendo $\sigma(x)$ compatto esiste $\phi \in \operatorname{chr}\left(\overline{\mathbb{C}}(\mathbb{e}, x)\right)$ tale che $\left|\phi(x)\right| = \|x\|$, e per quanto detto sui caratteri sappiamo già che $\phi(\mathbb{e}) = \left\|\phi\right\| = 1$.

Dal teorema di Hahn-Banach possiamo estendere ϕ con continuità su tutto X mantenendo invariata la norma e renderlo così uno stato. Ora sappiamo che $\left\| L_\phi(x)[y + N_\phi] \right\|^2 \leq \|x\|^2 \phi(y^* y)$ per ogni y, posto allora $y = x / \|x\|$ avremo anche che $\phi(x) \in \mathbb{R}$ e

$$\left\| L_\phi(x)[y + N_\phi] \right\|^2 = \left\| \frac{x^2}{\|x\|} + N_\phi \right\|^2 = \frac{\phi\left(x^4\right)}{\|x\|^2} = \frac{\left|\phi(x)\right|^4}{\|x\|^2} = \|x\|^2$$

in quanto ϕ è moltiplicativa su $\overline{\mathbb{C}}(\mathbb{e}, x)$. ■

Possiamo ora trovare H. Sia $\mathbb{S}^X$ l'insieme di tutti i vettori di X con norma unitaria, per ogni $x \in \mathbb{S}^X$ possiamo sempre scegliere uno stato ϕ_x in modo tale che $\left\| L_{\phi_x}(x^* x) \right\| = \left\| x^* x \right\|$ ed in particolare $\left\| L_x(x) \right\| = \|x\|$ con $L_x = L_{\phi_x}$. Posto allora $\mathscr{H} = \left\{ H_x \right\}_{x \in \mathbb{S}^X}$ poniamo

$$H = \ell^2 \left(\mathbb{S}^X, \mathscr{H} \right)$$

definito come nella sezione **??** a pagina **??**. Chiaramente H è uno spazio di Hilbert con il prodotto scalare standard, dobbiamo quindi definire solamente $L : X \to L(H)$. Per ogni $x \in X$ e per ogni $f \in \ell^2 \left(\mathbb{S}^X, \mathscr{H} \right)$, $z \in \mathbb{S}^X$ poniamo quindi

$$L(x)[f](z) = L_z(x)[f(z)] \in H_z$$

Essendo $\left\| L_z(x) \right\| \leq 1$ per ogni $z \in \mathbb{S}^X$ avremo che $L(x)(f) \in \ell^2 \left(\mathbb{S}^X, \mathscr{H} \right)$ e dunque $L(x) \in L(H)$ per ogni $x \in X$, perciò L è ben definita. Chiaramente poi L è lineare, moltiplicativa e $*$-compatibile dato che tutte le L_z le soddisfano, dobbiamo quindi far vedere solo che è una isometria.

Prendiamo $x \in X, f \in H$ generici, allora

$$\left\| L(x)(f) \right\|_{\ell^2}^2 = \sum_{z \in \mathbb{S}^X} \left\| L_z(x)[f(z)] \right\|^2 \leq \sum_{z \in \mathbb{S}^X} \left\| L_z(x) \right\|^2 \left\| f(z) \right\|^2 \leq \sum_{x \in \mathbb{S}^X} \|x\|^2 \left\| f(z) \right\|^2 = \|x\|^2 \left\| f \right\|_{\ell^2}^2$$

e quindi $\|L(x)\| \leq \|x\|$ e $\|L\| \leq 1$. Viceversa essendo H_z il completamento di X / N_z per ogni $x \in X \setminus \{0\}$ posto $z = x / \|x\|$ esisterà una successione $z_n \in H_z$ tale che $\left\| z_n \right\| = 1$ e $L_z(x)(z_n) \to \|x\|$ per il lemma appena dimostrato. Infine ci basta porre

$$f_n(w) = \begin{cases} z_n & \text{se } w = z \\ 0 & \text{altrimenti} \end{cases}$$

per ogni $w \in \mathbb{S}^X$, in questo modo avremo sia che $\left\| f_n \right\|_{\ell^2} = \left\| z_n \right\|_{H_z} = 1$ e sia che $\left\| L(x)(f_n) \right\|_{\ell^2} = \left\| L_z(x)(z_n) \right\| \to \|x\|$, dunque $\|L(x)\| = \|x\|$ ed L è una isometria.

Abbiamo quindi dimostrato il teorema di Gelfand-Naimark, o meglio abbiamo dimostrato una sua versione leggermente più debole in quanto non si è voluto entrare più nel dettaglio nello studio delle rappresentazioni di C^*-algebre. Ci basta aver dimostrato che ogni C^*-algebra unitaria è $*$-isomorfa ad una sottoalgebra di un'algebra di operatori in spazi di Hilbert, e quindi limitare lo studio delle C^*-algebre allo studio degli spazi di operatori tra spazi di Hilbert non è riduttivo.

Capitolo 11

Differenziazione

11.1 Differenziale di Fréchet

In questa sezione lavoreremo con funzioni che ad ogni valore del dominio associano un'altra funzione lineare o multilineare. Per agevolare la lettura utilizzeremo le parentesi quadre quando valutiamo una funzione lineare o multilineare su uno o più valori, mentre le parentesi tonde saranno riservate a funzioni che non sono necessariamente lineari oppure che devono essere trattate alla stregua di funzioni generiche.

Definizione 11.1.1. Siano X, Y spazi normati su uno stesso campo $\mathbb{K}$, $x \in X$ e $Z \leq X$ sottospazio chiuso diverso da $\{0\}$. Una funzione $f \in F(U, Y)$ si dice **differenziabile secondo Fréchet** in x **rispetto a** Z se e solo se esiste $\partial_Z f(x) \in L(Z, Y)$ tale che

$$\lim_{\substack{h \in Z \setminus \{0\} \\ h \to 0}} \frac{f(x+h) - f(x) - \partial_Z f(x)[h]}{\|h\|} = 0 \tag{11.1.1}$$

e $\partial_Z f(x)$ sarà il **differenziale** di f in x su Z.

Inoltre se U è un aperto non vuoto di X diremo che f è differenziabile secondo Fréchet su U se e solo se è differenziabile secondo Fréchet su tutti i punti di U, in tal caso $\partial_Z f$ è una funzione che ad ogni punto di U associa il suo differenziale. Se inoltre $Z = X$ allora poniamo per comodità $\partial_X f = df$.

Proposizione 11.1.2. *Il differenziale di Fréchet, se esiste, è unico.*

Dimostrazione. Sia $f : U \to Y$ differenziabile in $x \in U$ con U aperto di X e prendiamo $S, T \in L(X, Y)$ due diversi differenziali di Fréchet di f in x, se per assurdo $S \neq T$ allora esisterà $v \in X$ di norma unitaria tale che $S[v] \neq T[v]$

Dunque avremo che

$$0 = \lim_{h \to 0} \frac{S[h] - T[h]}{\|h\|} = \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{S[tv] - T[tv]}{t} = S[v] - T[v] \neq 0$$

il che è assurdo, quindi S deve coincidere con T. ∎

Dall'unicità del differenziale segue che, se f è differenziabile in x rispetto a Z, allora sarà differenziabile anche rispetto a $W \leq Z$ e $\partial_Z f(x)[v] = \partial_W f(x)[v]$ per ogni $v \in W$. Dunque se f

è differenziabile rispetto all'intero spazio allora sarà differenziabile su tutti i sottospazi non banali.

Dato che il differenziale di una funzione a valori su uno spazio normato è ancora una funzione a valori su un altro spazio normato, quindi ci si può chiedere se il differenziale è ancora differenziabile e, nel caso lo fosse, differenziarlo una seconda volta. Ricorsivamente possiamo allora definire i differenziali di ordine due, tre, quattro... semplicemente applicando più volte l'operatore di differenziazione.

Se $T \in L(X, Y)$ allora T è automaticamente differenziabile secondo Fréchet su tutto X con $dT(x) = T$ per ogni $x \in X$, in particolare $d^2 T = 0$. Inoltre l'operatore di differenziazione è lineare su $\mathbb{K}$ ovvero

$$\partial_Z(f + g) = \partial_Z f + \partial_Z g$$
$$\partial_Z(\alpha f) = \alpha \partial_Z f$$

Proposizione 11.1.3. *Una funzione differenziabile secondo Fréchet su tutto lo spazio è continua.*

Dimostrazione. Dalla definizione di derivata secondo Fréchet segue immediatamente che $\lim_{h\to 0} f(x+h) = f(x)$ e quindi f è continua in x. ■

Lemma 11.1.4 (Regola della catena)**.** *Dati tre spazi normati X, Y, Z, un sottospazio non banale $X' \le X$ ed un punto $x \in X$, siano f una funzione differenziabile in x rispetto ad X' a valori in Y ed g funzione differenziabile in $f(x) \in Y$ a valori in Z. Allora $g \circ f$ è differenziabile secondo Fréchet in x rispetto a X' e*

$$\partial_{X'}(g \circ f)(x)[w] = dg(f(x))\left[\partial_{X'} f(x)[w]\right]$$

Dimostrazione. Basta dimostrare che l'applicazione lineare

$$dg(f(x)) \circ \partial_{X'} f(x) : X' \to Z$$

soddisfa la (11.1.1). Per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $h \in X'$ con $\|h\| \le \delta$

$$\frac{\left\|f(x+h) - f(x)\right\|}{\|h\|} \le \left\|\partial_{X'} f(x)\left[\frac{h}{\|h\|}\right]\right\| + \epsilon \le \left\|\partial_{X'} f(x)\right\| + \epsilon$$

Poniamo allora $l = f(x+h) - f(x) \in X$, per continuità $l \to 0$ e

$$\begin{aligned}\frac{g(f(x+h)) - g(f(x)) - dg(f(x))[\partial_{X'} f(x)[h]]}{\|h\|} \\ = \frac{g(f(x)+l) - g(f(x)) - dg(f(x))[l]}{\|l\|}\frac{\left\|f(x+h) - f(x)\right\|}{\|h\|} \\ + dg(f(x))\left[\frac{f(x+h) - f(x) - \partial_{X'} f[h]}{\|h\|}\right]\end{aligned}$$

e l'ultimo addendo tende a 0 in quanto $dg(f(x)) \in L(Y, Z)$.

Abbiamo allora che

$$\begin{aligned}\lim_{h\to 0}\frac{g(f(x+h)) - g(f(x)) - dg(f(x))[\partial_{X'} f(x)[h]]}{\|h\|} \\ \le \left\|\partial_{X'} f(x)\right\| \lim_{l\to 0}\frac{g(f(x)+l) - g(f(x)) - dg(f(x))[l]}{\|l\|} = 0\end{aligned}$$

da cui discende la tesi. ■

Dati due sottospazi X_1, X_2 di X se F è differenziabile secondo Fréchet in $x \in U$ rispetto ad X allora per ogni $v \in X_1, w \in X_2$ avremo che

$$dF(x)[v+w] = dF(x)[v] + dF(x)[w] = \partial_{X_1} F(x)[v] + \partial_{X_2} F(x)[w] \tag{11.1.2}$$

In generale se F è differenziabile rispetto a X_1 e X_2 potrebbe non essere differenziabile su tutto X anche se quest'ultimo fosse somma continua dei due sottospazi, e controesempi se ne possono trovare anche in $\mathbb{R}^2$. Se però queste derivate sono continue su un intorno aperto di x allora la nostra funzione sarà differenziabile rispetto a tutto lo spazio, ma prima di poterlo dimostrare abbiamo bisogno di una versione debole del teorema di Lagrange

Teorema 11.1.5 (Lagrange). *Se $f : U \to Y$ è differenziabile su tutto $U \subseteq X$ allora per ogni $x \in U$ e $h \in X$ per cui $x + th \in U$ per ogni $0 \le t \le 1$ si ha*

$$\left\| f(x+h) - f(x) \right\| \le \sup_{0 \le t \le 1} \left\| df(x+th)[h] \right\|$$

Dimostrazione. Possiamo supporre che $\mathbb{K} = \mathbb{R}$, se infatti X ed Y fossero spazi complessi li possiamo sostituire rispettivamente con gli spazi reali $X_{\mathbb{R}}$ ed $Y_{\mathbb{R}}$. In tal modo F sarà ancora differenziabile da $X_{\mathbb{R}}$ in $Y_{\mathbb{R}}$ con lo stesso differenziale, che sarà visto come elemento di $L(X_{\mathbb{R}}, Y_{\mathbb{R}})$.

Fissiamo $h \in X$ per il corollario 7.2.2 esiste $F \in Y^*$ di norma unitaria tale che $F[f(x+h) - f(x)] = \left\| f(x+h) - f(x) \right\|$. Poiché F è lineare e continua dalla regola della catena anche $F \circ f$ sarà differenziabile su U, inoltre la funzione

$$\Lambda : t \in \mathbb{R} \to F[f(x+th)] \in \mathbb{R}$$

è derivabile e per la regola della catena $\Lambda'(t) = F[df(x+th)[h]]$. Per il teorema di Lagrange usuale avremo allora che

$$\left\| f(x+h) - f(x) \right\| = F[f(x+h)] - F[f(x)] = F[df(x+\bar{t}h)[h]] \le \sup_{0 \le t \le 1} \left\| df(x+th)[h] \right\|$$

per qualche $\bar{t} \in (0,1)$ dipendente da h e da F. ■

Teorema 11.1.6. *Se X è somma continua dei sottospazi X_1, X_2, se $f : U \to Y$ è differenziabile su tutto $U \subseteq X$ rispetto ai sottospazi X_1, X_2 e se i differenziali $\partial_{X_1} f, \partial_{X_2} f$ sono continui in $x \in U$ allora f è differenziabile in x rispetto a X e vale la* (11.1.2).

Dimostrazione. Applichiamo il teorema di Lagrange alla funzione

$$u : v \in X_1 \to f(x+w+v) - \partial_{X_1} f(x)[v] \in \mathbb{K}$$

avremo allora per ogni $v \in X_1, w \in X_2$

$$\left\| f(x+v+w) - f(x+w) - \partial_{X_1} f(x)[v] \right\| \le \sup_{0 \le t \le 1} \left\| \partial_{X_1} f(x+w+tv) - \partial_{X_1} f(x) \right\| \|v\|$$

e quindi

$$\begin{aligned}
\lim_{v+w\to 0} & \frac{\left\| f(x+v+w) - f(x) - \partial_{X_1} f(x)[v] - \partial_{X_2} f(x)[w] \right\|}{\|v+w\|} \\
&\leq \lim_{v+w\to 0} \frac{\left\| f(x+v+w) - f(x+w) - \partial_{X_1} f(x)[v] \right\|}{\|v+w\|} \\
&\quad + \lim_{v+w\to 0} \frac{\left\| f(x+w) - f(x) - \partial_{X_2} f(x)[w] \right\|}{\|v+w\|} \\
&\leq \lim_{v+w\to 0} \sup_{0\leq t\leq 1} \frac{\|v\|}{\|v+w\|} \left\| \partial_{X_1} f(x+w+tv) - \partial_{X_1} f(x) \right\| \\
&\quad + \lim_{v+w\to 0} \frac{\|w\|}{\|v+w\|} \frac{\left\| f(x+w) - f(x) - \partial_{X_2} f(x)[w] \right\|}{\|w\|}
\end{aligned}$$

Ora essendo X somma continua di X_1 e X_2 le proiezioni sono continue, in particolare $\|v\| \leq C\|v+w\|$ e $\|w\| \leq C\|v+w\|$ e quindi i rispettivi rapporti sono limitati. Ancora se $v+w \to 0$ allora sempre per la continuità delle proiezioni $v \to 0$ e $w \to 0$ e quindi per ogni $\varepsilon > 0$ possiamo sempre scegliere v e w sufficientemente vicini all'origine in modo tale che

$$\begin{gathered}
\sup_{0\leq t\leq 1} \left\| \partial_{X_1} f(x+w+tv) - \partial_{X_1} f(x) \right\| \leq \frac{\varepsilon}{2C} \\
\frac{\left\| f(x+w) - f(x) - \partial_{X_2} f(x)[w] \right\|}{\|w\|} \leq \frac{\varepsilon}{2C}
\end{gathered}$$

e quindi il limite fa 0 e f è differenziabile in x rispetto a tutto X. ■

Prendiamo ora due spazi normati X e Y sul campo $\mathbb{K}$, sappiamo che $X \times Y$ è uno spazio normato su $\mathbb{K}$ ed è somma continua di X ed Y. Prendiamo ora un aperto non vuoto U di uno spazio normato Z, delle funzioni $f \in F(U,X)$, $g \in F(U,Y)$ e $F \in F(X\times Y, V)$ differenziabili sui loro domini. Allora se poniamo $G(x) = [f(x), g(x)]$ per ogni $x \in U$ avremo che

$$df(x)[u] = p_X\left[dG(x)[u]\right]$$

e quindi

$$\begin{aligned}
d(F\circ G)(x)[u] &= dF(G(x))[dG(x)[u]] \\
&= \partial_X F(G(x))[p_X\left[dG(x)[u]\right]] + \partial_Y F(G(x))[p_Y\left[dG(x)[u]\right]] \qquad (11.1.3)\\
&= \partial_X F(f(x), g(x))[df(x)[u]] + \partial_Y F(f(x), g(x))[dg(x)[u]]
\end{aligned}$$

Proposizione 11.1.7 (Derivazione del prodotto)**.** *Dati un aperto U di X contenente x, due funzioni $f \in F(U,Y)$, $g \in F(U,Z)$ differenziabili in x e una forma bilineare $F \in L^2(Y,Z,V)$ si avrà*

$$d(F[f,g])(x)[u] = F[df(x)[u], g(x)] + F[f(x), dg(x)[u]]$$

Dimostrazione. Definiamo $\tilde{F} : Y \times Z \to V$ in modo tale che $\tilde{F}(y+z) = F[y,z]$, chiaramente $\tilde{F}$ non è una funzione lineare su $Y \times Z$ ma per la bilinearità di F avremo che

$$\begin{gathered}
\partial_A \tilde{F}(y+z)[h] = F[h,z] \\
\partial_B \tilde{F}(y+z)[h] = F[y,h]
\end{gathered}$$

e per continuità F sarà differenziabile rispetto a tutto $Y \times Z$ e soddisfa la (11.1.2).

Se quindi applichiamo la (11.1.3) avremo che per ogni $l \in D$

$$\begin{aligned} d(F[f,g])(x)[l] &= \partial_A \tilde{F}(f(x),g(x))[df(x)[l]] + \partial_B \tilde{F}(f(x),g(x))[dg(x)[l]] \\ &= F[df(x)[l],g(x)] + F[f(x),dg(x)[l]] \end{aligned}$$

■

11.2 Derivate successive e spazi C^k

Se una funzione f è differenziabile su un aperto non vuoto U di X ed ha valori in Y allora il suo differenziale df può essere visto anche come una funzione da $U \subseteq X$ in $L(X,Y)$.

Poiché $L(X,Y)$ è comunque uno spazio normato è perfettamente lecito chiedersi non solo se df è continua su U, ma anche se è differenziabile esso stesso in qualche punto se non su tutto U. Possiamo chiaramente iterare il procedimento ed introdurre così il concetto di differenziale di ordine superiore.

Definizione 11.2.1. Sia f una funzione definita su un sottoinsieme non vuoto di uno spazio normato X a valori in Y, e sia $U \subseteq X$ un aperto non vuoto contenuto nel dominio di f. Diremo allora che

- $f \in C^0(U,Y)$ se e solo se f è continua su U;
- $f \in C^1(U,Y)$ se e solo se f è differenziabile su tutto U e il suo differenziale è un'applicazione continua su U;
- $f \in C^k(U,Y)$ se e solo se f è differenziabile su U ed $df \in C^{k-1}[U,L(X,Y)]$;
- $f \in C^\infty(U,Y)$ se e solo se $f \in C^k(U,Y)$ per ogni $k \in \mathbb{N}$.

Inoltre se $f \in C^k(U,Y)$ per ogni $0 \le i \le k$ definiamo l'applicazione $d^i f$ su U in modo tale che per ogni $x \in U$

$$\begin{aligned} d^0 f(x) &= f(x) \in Y = L^0(Y) \\ d^1 f(x) &= df(x) \in L(X,Y) \\ d^2 f(x) &= d(df)(x) \in L[X,L(X,Y)] = L^2(X;Y) \\ &\cdots \\ d^{i+1} f(x) &= d\left(d^i f\right)(x) \in L\left[X,L^i(X;Y)\right] = L^{i+1}(X;Y) \end{aligned}$$

Se f è differenziabile su U allora il differenziale $\partial_Z f$ è una funzione definita su U a valori nello spazio normato $L(Z,Y)$. Quindi possiamo definire il differenziale della funzione $\partial_Z f$ in $x \in U$ rispetto ad un sottospazio non banale $W \le X$ nella stessa maniera, questo nuovo differenziale $\partial_W \partial_Z f(x)$ sarà il **differenziale al second'ordine** in x rispetto a W e a Z e apparterrà allo spazio $L^2(W,Z,Y)$.

Teorema 11.2.2 (Schwartz). *Sia f differenziabile su un intorno U di $x \in X$ rispetto a due sottospazi X_1 e X_2. Supponiamo anche che $\partial_{X_1} f$ e $\partial_{X_2} f$ siano differenziabili in x rispetto a X, allora per ogni $h \in X_1$, $k \in X_2$ si ha*

$$\partial_{X_2}\partial_{X_1} f(x)[k,h] = \partial_{X_1}\partial_{X_2} f(x)[h,k]$$

Dimostrazione. Fissato $x \in U$ per ogni $h \in X_1, k \in X_2$ poniamo

$$\Delta(k,h) = f(x+h+k) - f(x+k) - f(x+h) + f(x)$$
$$G_k(h) = f(x+k+h) - f(x+h)$$

allora $\Delta(k,h) = G_k(h) - G_k(0)$ e $\Delta(h,k) = \Delta(k,h)$. Dal teorema di Lagrange avremo che

$$\begin{aligned}&\left\| G_k(h) - G_k(0) - \partial_{X_1}\partial_{X_2} f(x)[k,h] \right\| \\ &\qquad \le \sup_{0\le t\le 1} \left\| \partial_{X_1} f(x+k+th) - \partial_{X_1} f(x+th) - \partial_{X_1}\partial_{X_2} f(x)[k] \right\| \|h\|\end{aligned}$$

e per la differenziabilità della derivata prima esisterà una funzione $r : U \to L(X,Y)$ tale che $r(l)/\|l\| \to 0$ al tendere di $l \in X$ a 0 e

$$\partial_{X_1} f(x+l) = \partial_{X_1} f(x) + d\partial_{X_1} f(x)[l] + r(l)$$

e quindi $\partial_{X_1} f(x+k+th) - \partial_{X_1} f(x+th) = d\partial_{X_1} f(x)[k] + r(k+th) - r(th)$ per ogni $0 \le t \le 1$. Se fissiamo $h,k \in X$ allora per T reale avremo che

$$\lim_{T\to 0^+} \left\| \frac{\Delta(Tk,Th)}{T^2} - \partial_{X_2}\partial_{X_1} f(x)[k,h] \right\| \le \lim_{T\to 0^+} \frac{\|Th\|\, R(Tk,Th)}{T^2} \tag{11.2.1}$$

dove

$$\begin{aligned}R(Tk,Th) &= \sup_{0\le t\le 1} \|r(Tk+tTh) - r(tTh)\| \\ &\le T\big(\|k\|+\|h\|\big) \sup_{0\le t\le 1}\left[\frac{\|r(Tk+tTh)\|}{\|Tk+tTh\|} + \frac{\|r(tTh)\|}{\|tTh\|} \right]\end{aligned}$$

e quindi l'ultimo membro della (11.2.1) è nullo. Per la simmetria di $\Delta(k,h)$ segue immediatamente che $\partial_{X_2}\partial_{X_1} f(x)[k,h] = \partial_{X_1}\partial_{X_2} f(x)[h,k]$. ■

Osservazione. Le ipotesi del teorema possono essere indebolite supponendo che le derivate parziali siano differenziabili in x rispettivamente a $X_1 + X_2$ invece che su tutto lo spazio. In particolare se $X_1 + X_2$ è somma continua di X_1 e X_2 e le derivate seconde rispetto a X_1 e X_2 sono continue in x allora il teorema di Schwartz è ancora valido grazie al teorema 11.1.6.

Corollario 11.2.3. *Se f è differenziabile k-volte in un intorno aperto di x, allora $d^k f(x) \in Sym_k(X;Y)$.*

Dimostrazione. Dimostriamo innanzitutto che per ogni $1 \le i < k$ e per ogni $h_1, h_2, \dots, h_k \in X$

$$d^k f(x)\left[h_k, h_{k-1}, \dots, h_{i+1}, h_i, \dots, h_2, h_1\right] = d^k f(x)\left[h_k, h_{k-1}, \dots, h_i, h_{i+1}, \dots, h_2, h_1\right]$$

osserviamo che possiamo supporre senza perdere in generalità che $i = 1$, altrimenti basta applicare lo stesso risultato a $d^{i-1} f$ invece che ad f. Definiamo $F_1, F_2 \in L\left[L^2(X;Y), Y\right]$ in modo tale che

$$F_1[A] = A[h_2, h_1]$$
$$F_2[A] = A[h_1, h_2]$$

allora $F_1[d^2f(y)] = F_2[d^2f(y)] = d^2f(y)[h_1,h_2]$. Per la regola della catena applicata $k-2$ volte avremo così che

$$F_1\left[d^kf(y)\left[h_k,h_{k-1},\dots,h_4,h_3\right]\right] = F_2\left[d^kf(y)\left[h_k,h_{k-1},\dots,h_4,h_3\right]\right]$$

ottenendo così la tesi.

In questo modo abbiamo dimostrato che il differenziale è invariante per permutazioni di termini consecutivi, vogliamo dimostrare che presa una qualunque permutazione σ di $\{1,2,\dots,k\}$ si ha

$$d^kf(y)\left[h_k,h_{k-1},\dots,h_1\right] = d^kf(y)\left[h_{\sigma(k)},h_{\sigma(k-1)},\dots,h_{\sigma(1)}\right]$$

Se $k=1,2$ ciò segue immediatamente da quanto detto prima, supponiamolo allora valido per qualche $k-1\geq 2$ e sia σ una permutazione dei primi k numeri interi. Se $\sigma(k)=k$ allora l'asserto segue direttamente dal passo induttivo, altrimenti $\sigma(j)=k$ con $j<k$.

Per induzione si avrà allora

$$d^kf(y)\left[h_k,h_{k-1},\dots,h_1\right] = d^kf(y)\left[h_k,h_{\sigma(k)},h_{\sigma(k-1)},\dots,h_{\sigma(j+1)},h_{\sigma(j-1)},\dots,h_{\sigma(1)}\right]$$

a questo punto basterà semplicemente spostare $h_k = h_{\sigma(j)}$ al suo posto tramite permutazioni di termini consegutivi, dimostrando così l'asserto. ■

Esempio 11.2.4. Sia X spazio normato generico, $x,v\in X$ e consideriamo l'applicazione

$$\sigma : t\in\mathbb{R}\to x+tv\in X$$

per induzione si può verificare facilmente che $\sigma\in C^k(\mathbb{R},X)$ per ogni $k\in\mathbb{N}_0$ con

$$d\sigma(t)[h] = hv \qquad\qquad d^k\sigma(t)=0\quad\forall k\geq 2$$

Prendiamo adesso una funzione $f\in C^k(X,Y)$, per la regola della catena avremo che per ogni $h_1,h_2,h_3\in\mathbb{R}$

$$d(f\circ\sigma)(t)[h_1] = h_1\,df(x+tv)[v] \qquad d^2(f\circ\sigma)(t)[h_1,h_2] = h_1h_2d^2f(x+tv)[v,v]$$

e così via. Per snellire la notazione se $F\in L^k(X,Y)$ e $x\in X$ scriveremo $F[x^k]$ al posto di $F[x,x,\dots,x]$ k-volte, in questo modo avremo che

$$d^k(f\circ\sigma)(t)[h_1,h_2,\dots,h_k] = h_1h_2\cdots h_k\,d^kf(x+tv)[v^k]$$

Prendiamo ora una funzione multilineare $F\in L^k(X,Y)$ e definiamo la seguente funzione

$$f : x\in X\to F[x^k]\in Y$$

facciamo vedere che f è differenziabile infinite volte su X. Per comodità consideriamo solo il caso $k=3$ ma l'idea si può estendere tranquillamente per k generico. Per la regola di derivazione del prodotto si ha

$$df(x)[h] = F[h,x,x]+F[x,h,x]+F[x,x,h] = F_1[x^2][h]+F_2[x^2][h]+F_3[x^2][h]$$

e di conseguenza avremo anche

$$d^2f(x)[k,h] = F[h,k,x] + F[h,x,k] + F[k,h,x] + F[x,h,k] + F[k,x,h] + F[x,k,h]$$
$$d^3f(x)[r,k,h] = F[h,k,r] + F[h,r,k] + F[k,h,r] + F[r,h,k] + F[k,r,h] + F[r,k,h]$$
$$d^nf(x) = 0 \text{ se } n \geq 4$$

e lo stesso ragionamento varrà per k generico. Osserviamo che $df(x)$ possiede 3 termini distinti in base a dove il fattore lineare viene posizionato all'interno di F dunque se F fosse k-multilineare allora $df(x)$ sarebbe formato da $\frac{k!}{(k-1)!} = k$ termini distinti.

Il termine d^2f è formato invece da 6 termini tutti ottenuti permutando i due fattori lineari e la variabile x, dunque se F fosse k-multilineare allora $d^2f(x)$ possiederebbe $\frac{k!}{(k-2)!} = k(k-1)$ termini distinti. In generale allora $d^if(x)$ possiederà $\frac{k!}{(k-i)!}$ termini distinti se $i \leq k$, e se $F \in Sym_k(X,Y)$ allora

$$d^if(x)[h_1,h_2,\dots,h_i] = \begin{cases} \dfrac{k!}{(k-i)!}F[h_1,\dots,h_i][x^{k-i}] & \text{se } i \leq k \\ 0 & \text{altrimenti} \end{cases} \tag{11.2.2}$$

Come per $\mathbb{R}^n$ anche per funzioni tra spazi normati è possibile considerare i polinomi di Taylor

$$p_k(v) = f(0) + df(0)[v] + \frac{1}{2}d^2f(0)[v^2] + \dots + \frac{1}{k!}d^kf(0)[v^k]$$

a partire da una funzione $f : X \to Y$ differenziabile k-volte in x e vogliamo trovare delle stime sul resto $R_k(v) = f(v) - p_k(v)$. Purtroppo se $Y \neq \mathbb{R}$ la formula per il resto di Lagrange non è una uguaglianza ma solamente una stima sulle norme.

Teorema 11.2.5. *Sia U aperto convesso contenente l'origine e $f : U \to Y$ differenziabile k volte nell'origine, allora*

1. *(Resto di Peano) Si ha*
$$\lim_{h\to 0}\frac{\|R_k(h)\|}{\|h\|^k} = 0$$

2. *(Resto di Lagrange)Se d^kf è differenziabile su U allora*
$$\|R_k(h)\| \leq \frac{1}{(k+1)!}\sup_{0\leq t\leq 1}\left\|d^{k+1}f(th)\right\|\|h\|^{k+1}$$

Dimostrazione. Se $k=1$ allora il primo punto deriva direttamente dalla definizione di differenziale secondo Fréchet, per induzione supponiamo allora che vale per un certo $k \geq 1$. Consideriamo il polinomio di Taylor $p_{k+1}(v)$ di una funzione f differenziabile $k+1$ volte nell'origine, grazie al teorema di Schwartz i differenziali $d^if(0)$ sono simmetrici e quindi possiamo utilizzare la (11.2.2) per dimostrare che dp_{k+1} è il polinomio di Taylor di df di grado k e quindi per induzione avremo che $\lim_{h\to 0}\frac{\|dp_{k+1}(h)-df(h)\|}{\|h\|^k} = 0$.

Dal teorema di Lagrange segue quindi che

$$\lim_{h\to 0}\frac{\|p_{k+1}(h) - f(h)\|}{\|h\|^{k+1}} \leq \lim_{h\to 0}\sup_{0\leq t\leq 1}\frac{\|dp_{k+1}(th) - df(th)\|}{\|h\|^k} = 0$$

e quindi il teorema di Taylor con il resto di Peano è verificato.

Fissiamo adesso $h \in U$ e f differenziabile $(k+1)$-volte nell'origine esisterà $F \in Y^*_{\mathbb{R}}$ di norma unitaria tale che $\left\|p_k(h) - f(h)\right\| = F[p_k(h) - f(h)]$, inoltre $F \circ p_k$ è il polinomio di Taylor di $F \circ f$ quindi passando al caso unidimensionale esisterà $t \in \left(0,1\right)$ tale che

$$\begin{aligned}\left\|R_k(h)\right\| = F[p_k(h) - f(h)] = \frac{1}{(k+1)!}F[d^{k+1}f(th)[h^{k+1}]]\\ \leq \frac{1}{(k+1)!}\sup_{0\leq t\leq 1}\left\|d^{k+1}f(th)\right\|\|h\|^{k+1}\end{aligned}$$

■

11.3 Il teorema della funzione implicita

Dati due spazi vettoriali topologici X ed Y su uno stesso campo $\mathbb{K}$ indichiamo con il simbolo $GL(X,Y)$ lo spazio di tutte le funzioni lineari e continue da X in Y invertibili che hanno inversa ancora continua. Chiaramente gli elementi di $GL(X,Y)$ sono sia omeomorfismi che isomorfismi tra spazi vettoriali quindi se $GL(X,Y)$ non è vuoto allora gli spazi X ed Y possiederanno le stesse proprietà vettoriali e topologiche e possono essere identificati.

In questa sezione inoltre se $S \in L(X,Y)$ e $T \in L(Y,Z)$ scriveremo TS al posto di $T \circ S$, in quanto la composizione di funzioni lineari e continue è essa stessa una forma bilineare continua tra spazi normati.

Innanzitutto dal teorema 8.3.10 non è difficile dimostrare che

Lemma 11.3.1. *Dati due spazi di Banach X ed Y su $\mathbb{K}$ allora per ogni $T \in GL(X,Y)$ e $S \in L(X,Y)$ se*

$$\left\|T^{-1}S\right\| < 1$$

allora $S+T \in GL(X,Y)$ e

$$\left\|(S+T)^{-1}\right\| \leq \frac{\left\|T^{-1}\right\|}{1-\left\|T^{-1}S\right\|}$$

In particolare $GL(X,Y)$ è un sottoinsieme aperto di $L(X,Y)$.

Possiamo chiederci allora se la funzione

$$I : S \in GL(X,Y) \to S^{-1} \in GL(Y,X)$$

sia differenziabile sull'aperto $GL(X,Y)$. Fissiamo adesso due funzioni $T \in GL(X,Y)$ e $S \in B\left(0, \left\|T^{-1}\right\|^{-1}, L(X,Y)\right)$ allora $S+T \in GL(X,Y)$ e

$$\begin{aligned}(S+T)^{-1} - T^{-1} &= (S+T)^{-1}TT^{-1} - (S+T)^{-1}(S+T)T^{-1}\\ &= -(S+T)^{-1}ST^{-1}\\ &= (S+T)^{-1}T\left(-T^{-1}ST^{-1}S - T^{-1}S + T^{-1}ST^{-1}S\right)T^{-1}\\ &= (S+T)^{-1}T\left[-T^{-1}(S+T)T^{-1}S + T^{-1}ST^{-1}S\right]T^{-1}\\ &= -T^{-1}ST^{-1} + (S+T)^{-1}\left(ST^{-1}\right)^2\end{aligned}$$

l'applicazione $S \to T^{-1}ST^{-1}$ è lineare e continua dunque

$$\left\|(S+T)^{-1} - T^{-1} + T^{-1}ST^{-1}\right\| \le \left\|(S+T)^{-1}\right\| \|S\|^2 \left\|T^{-1}\right\|^2 \le \|S\|^2 \frac{\left\|T^{-1}\right\|^3}{1 - \|S\| \left\|T^{-1}\right\|}$$

e quindi I è differenziabile su tutto $GL(X,Y)$ con

$$dI(T)[S] = -T^{-1}ST^{-1} = F_2\left[T^{-1}, T^{-1}\right][S]$$

in particolare I è di classe C^1. Supponiamo adesso che I è di classe C^k con $k \ge 1$ allora anche $F_2[T^{-1}, T^{-1}]$ deve essere di classe C^k in quanto composizione di funzioni di classe C^k, dunque $I \in C^{k+1}$ e quindi I è differenziabile infinite volte su $GL(X,Y)$.

Adesso abbiamo tutti gli strumenti per dimostrare il

Teorema 11.3.2 (della funzione implicita)**.** *Dati tre spazi di Banach X, Y, Z, un aperto U di $X \times Y$ contenente (x_0, y_0) e una funzione $F : U \to Z$ differenziabile rispetto a Y su tutto U con $\partial_Y F \in C[U, L(Y,Z)]$. Se $F(x_0, y_0) = 0$ e $\partial_Y F(x_0, y_0) \in GL(Y,Z)$ allora esistono un aperto $U_0 \subseteq X$ contenente x_0, un aperto $V_0 \subseteq Y$ contenente y_0 e una funzione continua $f : U_0 \to V_0$ tale che*

1. *$f(x_0) = y_0$;*
2. *$F[x, f(x)] = 0$ per ogni $x \in U_0$;*
3. *Se $x \in U_0, y \in V_0$ e $F(x,y) = 0$ allora $y = f(x)$;*
4. *Se $F \in C^k$ su U allora $f \in C^k$ e per ogni $h \in X$*

$$df(x) = -\left[\partial_Y F(x, f(x))\right]^{-1} \circ \partial_X F(x, f(x))$$

Innanzitutto poniamo per comodità $T = \partial_Y F(x_0, y_0) \in GL(Y,Z)$ e $R(h,k) = F(x_0 + h, y_0 + k) - Tk$, vogliamo dimostrare che esistono $r, s > 0$ tali che l'applicazione

$$\varphi_h(k) = T^{-1}R(h,k)$$

è una contrazione su $D(0, r, Y)$ per ogni $h \in B(0, s, X)$. Poiché $\partial_Y F$ è continua su U dato che $B((x,y), r, X \times Y) = B(x, r, X) \times B(y, r, Y)$ per come abbiamo definito la norma sul prodotto di spazi normati, esisterà $r > 0$ tale che per ogni $h \in X$, $k \in Y$ con $\|h\|, \|k\| \le r$

$$\sup_{0 \le t \le 1} \left\|\partial_Y F(x_0 + h, y_0 + k) - T\right\| \le \frac{1}{2\left\|T^{-1}\right\|}$$

ed esisterà $s \in (0, r)$ tale che per ogni $h \in B(0, s, X)$

$$\left\|F(x_0 + h, y_0)\right\| \le \frac{r}{2\left\|T^{-1}\right\|}$$

In questo modo per ogni $h \in B(0, s, X)$ e per ogni $k, k' \in D(0, r, Y)$ abbiamo per il teorema di Lagrange

$$\begin{aligned}
&\left\|T^{-1}R(h,k) - T^{-1}R(h,k')\right\| \\
&\qquad \le \left\|T^{-1}\right\| \left\|F(x_0 + h, y_0 + k) - F(x_0 + h, y_0 + k') - T[k - k']\right\| \\
&\qquad \le \sup_{0 \le t \le 1} \left\|T^{-1}\right\| \left\|\partial_Y F(x_0 + h, y_0 + k' + t(k - k')) - T\right\| \left\|k - k'\right\| \le \frac{\left\|k - k'\right\|}{2}
\end{aligned}$$

e inoltre

$$\left\|T^{-1}R(h,k)\right\| \le \left\|T^{-1}F(h,0)\right\| + \frac{\|k\|}{2} \le r$$

quindi φ_h manda $D(0,r,Y)$ in $D(0,r,Y)$ ed è una contrazione, dunque per il teorema di Banach-Cacioppoli per ogni $h \in U_0 - x_0 = B(0,s,X)$ esiste $g(h) \in V_0 - y_0 = D(0,r,Y)$ tale che $R(h,g(h)) = Tg(h)$ e quindi $F(h+x_0, g(h)+y_0) = 0$. Per l'unicità del punto fisso la funzione $f(x) = y_0 + g(x-x_0)$ soddisfa tutte le ipotesi di univocità del teorema.

Dimostriamo ora la continuità di f. Per ogni $h, h' \in B(0,s,X)$

$$\begin{aligned}
\left\|f(x_0+h) - f(x_0+h')\right\| &= \left\|\varphi_h[g(h)] - \varphi_{h'}[g(h')]\right\| \\
&\le \frac{\left\|g(h)-g(h')\right\|}{2} + \left\|\varphi_h(g(h')) - \varphi_{h'}(g(h'))\right\| \\
&\le \frac{\left\|g(h)-g(h')\right\|}{2} + \left\|T^{-1}\right\| \left\|R(h,g(h')) - R(h',g(h'))\right\| \\
&= \frac{\left\|g(h)-g(h')\right\|}{2} + \left\|T^{-1}\right\| \left\|F(h,g(h')) - F(h',g(h'))\right\|
\end{aligned}$$

per la continuità di F mantenendo fisso h' segue la continuità di f.

Supponiamo ora che F sia di classe C^1 su U, per la (11.1.2) abbiamo per ogni $x \in U_0$, $h \in X$ e $k \in Y$

$$dF(x,f(x))[(h,k)] = \partial_X F(x,f(x))[h] + \partial_Y F(x,f(x))[k]$$

posto $y = f(x)$ per il lemma 11.3.1 e dalla continuità di f e di $\partial_Y F$ per x sufficientemente vicino anche $\partial_Y F(x,y)$ sarà invertibile con inversa continua. Restringiamo dunque U_0 in modo che $\partial_Y F$ sia invertibile su tutto U_0 e $\left\|[\partial_Y F(x,f(x))]^{-1}\right\| \le 2\left\|[\partial_Y F(x_0,y_0)]^{-1}\right\| = 2\left\|T^{-1}\right\|$.

Allora posto anche $k = f(x+h) - f(x) \in Y$ sufficientemente piccolo in norma e $Sh = -[\partial_Y F(x,y)]^{-1} \circ \partial_X F(x,y)$

$$\begin{aligned}
\frac{\left\|f(x+h)-f(x)-Sh\right\|}{\|h\|} &= \frac{\left\|[\partial_Y F(x,y)]^{-1}[\partial_Y F(x,y)[k] + \partial_X F(x,y)[h]]\right\|}{\|h\|} \\
&\le 2\left\|T^{-1}\right\| \frac{\left\|\mathcal{E}(h,k)\right\|}{\|h\|} \\
&\le 2\left\|T^{-1}\right\| \left[\frac{\left\|f(x+h)-f(x)-Sh\right\|}{\|h\|} + \|S\| + 1\right] \frac{\left\|\mathcal{E}(h,k)\right\|}{\|h\| \vee \|k\|}
\end{aligned} \tag{11.3.1}$$

dove $\mathcal{E}(h,k) = F(x+h,y+k) - F(x,y) - \partial_Y F(x,y)[k] - \partial_X F(x,y)[h]$ in quanto $F(x,y) = F(x+h,y+k) = 0$. Se la quantità tra parentesi quadre all'ultimo membro della (11.3.1) fosse illimitata si otterrebbe un assurdo e dunque deve necessariamente essere limitato in h e di conseguenza il primo membro tenderebbe a 0.

Per capire meglio questo passaggio muoviamoci in $\mathbb{R}^+$. Per ogni $t > 0$ consideriamo una qualunque funzione $f(t)$ positiva che tende a 0 all'infinito, se esistessero $l, k > 0$ tali che

$$t \le k(t+l)f(t)$$

allora avremo che $kf(t) \ge t/(t+l)$ il che è assurdo in quanto f pur andando a 0 all'infinito dovrebbe essere più grande di una quantità che mano mano si allontana dall'origine.

Sia ora $F \in C^k$ e sappiamo che $f \in C^0$ ovvero che è continua. Supponiamo ora che $f \in C^i$ con $i < k$, poiché $\partial_X F$, $\partial_Y F$ sono di classe C^{k-1} e $\partial_Y F$ è sempre invertibile allora la loro composizione con $(x, f(x))$ sarà ancora di classe C^i in quanto l'applicazione $S \to S^{-1}$ è differenziabile infinite volte. Dunque $df \in C^i$ e $f \in C^{i+1}$, per induzione avremo $f \in C^k$.

Corollario 11.3.3 (Teorema di invertibilità locale)**.** *Se f è definita su un aperto di X a valori in Y ed è di classe C^1 ed esiste x_0 nel dominio tale che $df(x_0) \in GL(X,Y)$ allora esistono due intorni aperti di x_0 ed $f(x_0)$ rispetto ai quali la funzione f è invertibile con inversa f^{-1} di classe C^1 e*

$$df^{-1}(f(x)) = [df(x)]^{-1}$$

Inoltre se $f \in C^k$ allora anche l'inversa sarà di classe C^k.

Dimostrazione. Basta applicare il teorema della funzione implicita alla funzione

$$F(x,y) = f(x) - y$$

rispetto però alla variabile x, dato che $\partial_X F(x_0, f(x_0)) = df(x_0) \in GL(X,Y)$. ■

11.4 Differenziazione di funzioni convesse

Abbiamo già introdotto il concetto di sottodifferenziale per quanto riguarda le funzioni convesse. In questa sezione studieremo invece la differenziabilità delle funzioni convesse definite su uno spazio normato reale X.

Il primo risultato ci mostra che per una funzione differenziabile i concetti di differenziale e sottodifferenziale coincidono:

Proposizione 11.4.1. *Se $F : X \to \overline{\mathbb{R}}$ è differenziabile in $x \in \operatorname{dom} F$ allora $\partial^* F(x) = \{dF(x)\}$.*

Dimostrazione. Dato che f è differenziabile in x sarà anche continua in tale punto, dunque per il teorema 6.4.6 il sottodifferenziale non è vuoto. Sia allora $f \in \partial^* F(x)$, per ogni $y \in X, \lambda > 0$ si avrà dunque che

$$F(x+\lambda y) - F(x) - dF(x)[\lambda y] \geq [f - dF(x)][\lambda y]$$

e perciò facendo tendere λ a 0 si ha $dF(x)[y] \geq f[y]$ per ogni $y \in Y$ ovvero $f = dF(x)$. ■

Dato uno spazio normato reale X con norma $\|\cdot\|$ ci chiediamo in quali casi l'applicazione

$$n : x \in X \to \|x\|^2 \in \mathbb{R}$$

risulta essere differenziabile su X. Nel caso in cui X fosse uno spazio di Hilbert si verifica velocemente che la norma è differenziabile con differenziale $dn(x)[v] = 2(x : v)$. Invece se X fosse uno spazio vettoriale complesso allora in quasi nessun caso la norma risulta essere differenziabile e un controesempio si trova anche in $\mathbb{C}$.

Ricordiamo la definizione di $\mathscr{D}(x)$ data in (7.7.1), dimostriamo innanzitutto che

Proposizione 11.4.2. *L'insieme $\mathscr{D}(x)$ è il sottodifferenziale di $\|x\|^2/2$ in x.*

Dimostrazione. Se $f \in \mathscr{D}(x)$ allora per ogni $y \in V$

$$\frac{\|y\|^2 - \|x\|^2}{2} = \frac{\|y\|^2 + \|f\|^2}{2} - \|x\|^2 \geq f(y-x)$$

viceversa applicando la proposizione 6.4.2 e la definizione di polare avremo che $\|x\|^2/2 + \|f\|^2/2 = f(x) = \|f\| \, \|x\|$ e dunque $(\|f\| - \|x\|)^2 = 0$ da cui discende che $f \in \mathscr{D}(x)$. ■

Supponiamo ora che la polare della norma di X sia strettamente convessa. Per la proposizione 7.7.2 questo insieme possiede un solo elemento, dunque $\mathscr{D}$ diventa un'applicazione da X in X^* e per il teorema 6.4.7 abbiamo

$$\lim_{\lambda \to 0} \frac{\|x + \lambda y\|^2 - \|x\|^2}{\lambda} = 2\mathscr{D}(x)[y]$$

per ogni $y \in V$. In generale però la norma non è differenziabile, bisogna imporre condizioni più restrittive.

Da questo momento supponiamo che la norma di X^* sia uniformemente convessa, vogliamo dimostrare che in questo caso la norma $\|x\|^2$ non solo è differenziabile su tutto X ma che è di classe C^1 su di esso.

Innanzitutto per ogni $x, y \in X$ non è difficile dimostrare la seguente uguaglianza simile all'identità del parallelogramma

$$[\mathscr{D}(x) + \mathscr{D}(y)][x + y] + [\mathscr{D}(x) - \mathscr{D}(y)][x - y] = 2\|x\|^2 + 2\|y\|^2$$

quindi se $\|x\| = \|y\| = 1$ allora $\|\mathscr{D}(x)\| = \|\mathscr{D}(y)\| = 1$ e per la disuguaglianza di sopra abbiamo che $\|\mathscr{D}(x) + \mathscr{D}(y)\| + \|x - y\| \geq 2$. Usiamolo per dimostrare che

Proposizione 11.4.3. *Se X^* è uniformemente convesso allora $\mathscr{D}$ è uniformemente continua sui sottoinsiemi limitati di X.*

Dimostrazione. Per assurdo il risultato non è vero e dunque esiste $R > 0$ tale che $\mathscr{D}$ non è uniformemente continua su $B(0, R)$ e dunque esisterà $\varepsilon > 0$ tale che per ogni $\delta > 0$ esistono $x_\delta, y_\delta \in X$ tali che $\|x_\delta\|, \|y_\delta\| < R$, $\|x_\delta - y_\delta\| < \delta$ e $\|\mathscr{D}(x_\delta) - \mathscr{D}(y_\delta)\| \geq \varepsilon$.

Dato che $\|\mathscr{D}(x_\delta)\|, \|\mathscr{D}(y_\delta)\| < R$ per l'uniforme continuità esiste $\bar{\delta} > 0$ che dipende solamente da ε tale che

$$\|\mathscr{D}(x_\delta) + \mathscr{D}(y_\delta)\| < 2R(1 - \bar{\delta})$$

per ogni $\delta > 0$, e quindi

$$2R < \|\mathscr{D}(x_\delta) + \mathscr{D}(y_\delta)\| + \|x_\delta - y_\delta\| < 2R(1 - \bar{\delta}) + \delta$$

e facendo tendere δ a 0 si ottiene una contraddizione. ■

Possiamo adesso dimostrare che

Teorema 11.4.4. *Sia X spazio riflessivo reale in modo tale che X^* sia uniformemente convesso, allora l'applicazione $n(x) = \|x\|^2$ è di classe C^1 su X e $dn(x)[v] = 2\mathscr{D}(x)[v]$.*

Dimostrazione. Innanzitutto osserviamo che per definizione di sottodifferenziale

$$2\mathscr{D}(x)[y] \leq \|x + y\|^2 - \|x\|^2$$
$$2\mathscr{D}(x + y)[-y] \leq \|x\|^2 - \|x + y\|^2$$

da cui discende che

$$0 \leq \|x + y\|^2 - \|x\|^2 - 2\mathscr{D}(x)[y] \leq 2[\mathscr{D}(x + y) - \mathscr{D}(x)][y]$$

e la tesi segue dalla continuità di $\mathscr{D}$. ■

In realtà il risultato appena ottenuto ci fornisce un criterio di differenziabilità per le funzioni convesse: presa una qualunque funzione $f : X \to \mathbb{R}$ convessa definita su uno spazio normato X ed un punto $x_0 \in X$ tali che

1. esiste un aperto U di X contenente x_0 tale che per ogni $x \in U$ il sottodifferenziale $\partial^* f(x)$ è formato da un solo elemento;

2. $\partial^* f(x) \in X^*$ è continuo in x_0.

Allora f è differenziabile in x_0 con $df(x_0) = \partial^* f(x_0)$.

Capitolo 12

Integrazione

12.1 Misurabilità

In questa sezione diamo per scontato che il lettore conosca i fondamenti della teoria della misura e dell'integrazione astratta, si veda ad esempio [8]. In questa sezione indicheremo con M un qualunque insieme non vuoto e con X uno spazio di Banach reale o complesso.

Definizione 12.1.1. Data una famiglia non vuota $\mathscr{M}$ di sottoinsiemi di M, consideriamo le seguenti possibili proprietà per $\mathscr{M}$:

(A) se $A \in \mathscr{M}$ allora $M \setminus A \in \mathscr{M}$;

(B) se $A, B \in \mathscr{M}$ allora $A \cup B \in \mathscr{M}$;

(C) se $A, B \in \mathscr{M}$ allora $A \cap B \in \mathscr{M}$;

(D) se $A_n \in \mathscr{M}$ per ogni $n \in \mathbb{N}$ allora $\bigcup_{n=1}^{+\infty} A_n \in \mathscr{M}$;

(E) se $A_n \in \mathscr{M}$ per ogni $n \in \mathbb{N}$ ed $A_m \cap A_n = \emptyset$ quando $m \neq n$ allora $\bigcup_{n=1}^{+\infty} A_n \in \mathscr{M}$;

(F) se $A \in \mathscr{M}$ e $B \subseteq A$ allora $B \in \mathscr{M}$.

Allora:

- se $\mathscr{M}$ soddisfa (A) e (B) diremo che è un'**algebra**;
- se soddisfa (A) e (D) diremo che è una **σ-algebra**;
- se soddisfa (A) ed (E) diremo che è un **λ-sistema**;
- se soddisfa (C) diremo che è un **π-sistema**;
- se soddisfa (D) ed (F) diremo che è un **σ-ideale**.

Osservazione. Dato che la famiglia $\mathscr{M}$ non può essere vuota essa dovrà necessariamente contenere M se $\mathscr{M}$ è un'algebra, una σ-algebra o un λ-sistema. Conterrà anche $\emptyset$ se è un'algebra, una σ-algebra, un λ-sistema o un σ-ideale. Le algebre tra l'altro sono chiuse rispetto alle intersezioni finite mentre le σ-algebre lo sono anche rispetto alle intersezioni numerabili.

Un λ-sistema che è anche un π-sistema è automaticamente una σ-algebra come è facile verificare.

Data una famiglia non vuota di insiemi $\mathscr{I}$ esistono sempre e sono univoci la σ-algebra generata, che indicheremo con $\sigma(\mathscr{I})$, e il λ-sistema generato, che indicheremo con $\lambda(\mathscr{I})$, ovvero rispettivamente la più piccola σ-algebra e il più piccolo λ-sistema che contengono $\mathscr{I}$. Chiaramente $\lambda(\mathscr{I}) \subseteq \sigma(\mathscr{I})$.

Lemma 12.1.2 (Dynkin)**.** *Se $\mathscr{I}, \mathscr{D}$ sono rispettivamente un π-system ed un λ-system su D per cui $\mathscr{I} \subseteq \mathscr{D}$ allora la σ-algebra generata da $\mathscr{I}$ è ancora contenuta in $\mathscr{D}$. In altre parole $\sigma(\mathscr{I}) = \lambda(\mathscr{I}) \subseteq \mathscr{D}$.*

Dimostrazione. Osserviamo che se $A, B \in \mathscr{D}$ con $A \subseteq B$ allora $B \setminus A \in \mathscr{D}$. Infatti A e $M \setminus B$ sono disgiunti e così $B \setminus A = M \setminus [(M \setminus B) \cup A] \in D$.

Consideriamo adesso la famiglia

$$I_1 = \left\{U \in \lambda(\mathscr{I}) \mid U \cap V \in \lambda(\mathscr{I})\ \forall V \in \mathscr{I}\right\}$$

chiaramente $M \in I_1$ e se $U \in I_1$ allora $M \setminus U \in \lambda(\mathscr{I})$ e per ogni $I \in \mathscr{I}$

$$(M \setminus U) \cap I = I \setminus (U \cap I)$$

e poiché $U \cap I \in \lambda(\mathscr{I})$ per quanto detto prima avremo $M \setminus U \in I_1$.

Ora prendiamo $A_n \in I_1$ a due a due disgiunti, chiaramente $A_n \cap I$ appartengono ad $\lambda(\mathscr{I})$ per ogni $I \in \mathscr{I}$ e continuano ad essere a due a due disgiunti, così

$$\bigcup_{n=1}^{+\infty} A_n \cap I = \bigcup_{n=1}^{+\infty} (A_n \cap I) \in \lambda(\mathscr{I})$$

e quindi l'unione degli A_n appartiene ad I_1. Ora essendo $\mathscr{I}$ un π-sistema si avrà $\mathscr{I} \subseteq I_1$ e perciò I_1 è un λ-sistema contenente $\mathscr{I}$ e contenuto in $\lambda(\mathscr{I})$, dunque $I_1 = \lambda(\mathscr{I})$.

Definiamo adesso

$$I_2 = \left\{U \in \lambda(\mathscr{I}) \mid U \cap V \in \lambda(\mathscr{I})\ \forall V \in \lambda(\mathscr{I})\right\}$$

dato che $\lambda(\mathscr{I}) = I_1$ avremo $\mathscr{I} \subseteq I_2$. Con un ragionamento analogo al precedente si verifica che I_2 è ancora un λ-sistema e perciò $I_2 = \lambda(\mathscr{I})$. Abbiamo così dimostrato che $\lambda(\mathscr{I})$ è anche un π-sistema e dunque è una σ-algebra. ■

Date una σ-algebra $\mathscr{M}$ ed un σ-ideale $\mathscr{Z}$ sull'insieme non vuoto M chiameremo la coppia $(\mathscr{M}, \mathscr{Z})$ una **σ-algebra completabile** sull'insieme M. Gli elementi di $\mathscr{Z}$ saranno detti **insiemi nulli** rispetto alla σ-algebra completabile $(\mathscr{M}, \mathscr{Z})$, questi insiemi rappresentano in qualche modo il grado di errore ammissibile per le proprietà che dovrebbero soddisfare le funzioni $(\mathscr{M}, \mathscr{Z})$-misurabili (il concetto classico di quasi ovunque).

Ogni σ-algebra $\mathscr{M}$ da sola induce naturalmente la σ-algebra completabile $(\mathscr{M}, \{\emptyset\})$ e quindi anch'esse possono essere considerate come un caso particolare di σ-algebre completabili.

Proposizione 12.1.3. *Se $(\mathscr{M}, \mathscr{Z})$ è una σ-algebra completabile allora*

$$\sigma(\mathscr{M} \cup \mathscr{Z}) = \left\{P \setminus Z_1 \cup Z_2 \;\middle|\; P \in \mathscr{M}, Z_1, Z_2 \in \mathscr{Z}\right\}$$

*Chiameremo tale σ-algebra il **completamento** di $(\mathscr{M}, \mathscr{Z})$ e la indicheremo con il simbolo $\mathscr{CC}(\mathscr{M}, \mathscr{Z})$.*

Dimostrazione. Posto

$$\mathscr{I} = \mathscr{M} \cup \mathscr{Z}$$
$$\mathscr{D} = \left\{ P \setminus Z_1 \cup Z_2 \;\middle|\; P \in \mathscr{M}, Z_1, Z_2 \in \mathscr{Z} \right\}$$

chiaramente $\mathscr{I} \subseteq \mathscr{D} \subseteq \sigma(\mathscr{I})$. Se verifichiamo che $\mathscr{D}$ è una σ-algebra la tesi è immediata.

Osserviamo innanzitutto che per ogni $P \in \mathscr{M}$, $Z_1, Z_2 \in \mathscr{Z}$

$$M \setminus (P \setminus Z_1 \cup Z_2) = (M \setminus P) \setminus Z_2 \cup Z_1 \setminus Z_2 \in \mathscr{D}$$

mentre scelti $P^k \setminus Z_1^k \cup Z_2^k \in \mathscr{D}$ si ha

$$\bigcup_{k=1}^{+\infty} \left(P^k \setminus Z_1^k \cup Z_2^k \right) = \left(\bigcup_{k=1}^{+\infty} P^k \right) \setminus K \cup \bigcup_{k=1}^{+\infty} Z_2^k$$

con

$$K = \bigcup_{k=1}^{+\infty} Z_1^k \setminus \left(\bigcup_{i \neq k} P^i \setminus Z_i^i \right) \in \mathscr{Z}$$

e quindi $\mathscr{D}$ è una σ-algebra. ■

Un sottoinsieme $E \subseteq M$ diremo che è $(\mathscr{M}, \mathscr{Z})$-misurabile se e solo se $E \in \mathscr{C}\mathscr{C}\left(\mathscr{M}, \mathscr{Z}\right)$, in tal caso utilizzeremo anche la notazione $E \in (\mathscr{M}, \mathscr{Z})$. Diciamo che $(\mathscr{M}, \mathscr{Z}) \subseteq (\mathscr{N}, \mathscr{W})$ se e solo se $\mathscr{C}\mathscr{C}\left(\mathscr{M}, \mathscr{Z}\right) \subseteq \mathscr{C}\mathscr{C}\left(\mathscr{N}, \mathscr{W}\right)$.

Diremo inoltre che $(\mathscr{M}, \mathscr{Z}) \subset\subset (\mathscr{N}, \mathscr{W})$ se e solo se $(\mathscr{M}, \mathscr{Z}) \subseteq (\mathscr{N}, \mathscr{W})$ e $Z \subseteq W$, aggiungendo così una limitazione agli insiemi nulli. Definiamo anche la seguente operazione tra σ-algebre completabili:

$$(\mathscr{M}, \mathscr{Z}) \cap (\mathscr{N}, \mathscr{W}) = (\mathscr{M} \cap \mathscr{N}, \mathscr{Z} \cap \mathscr{W})$$

Dato un σ-ideale $\mathscr{Z}$ su M, diremo che una successione di funzioni $f_n : M \to X$ converge puntualmente **$\mathscr{Z}$-quasi ovunque** ad $f : M \to X$ se e solo se esiste $Z \in \mathscr{Z}$ tale che $\lim_{n \to +\infty} f_n(x) = f(x)$ per ogni $x \in M \setminus Z$. Per coerenza con la notazione successiva diciamo anche che converge puntualmente $(\mathscr{M}, \mathscr{Z})$-quasi ovunque se e solo se vi converge puntualmente $\mathscr{Z}$-quasi ovunque.

Diremo che una qualunque funzione $s : M \to X$ è **semplice** se e solo se l'immagine di s è un sottoinsieme finito di X. Se $\mathscr{M}$ è una σ-algebra su M diremo che s è $\mathscr{M}$-semplice se e solo se è semplice e $\{s = x\} \in \mathscr{M}$ per ogni $x \in X$. Prima di introdurre il concetto di misurabilità è opportuno fare alcune precisazioni sulla struttura di X. Se X fosse uno spazio di Banach separabile allora sarebbe anche $\mathscr{N}_2$ e quindi lo studio della misurabilità funzionerebbe in maniera analoga al caso in dimensione finita.

Se invece X non fosse separabile la sua topologia possiede troppi aperti e quindi le funzioni misurabili in senso classico non sarebbero approssimabili tramite funzioni semplici. D'altronde considerare come funzioni misurabili solo quelle approssimabili esattamente tramite funzioni semplici risulta essere troppo riduttivo in quando, nel momento in cui vogliamo integrare una funzione, possiamo permetterle di variare una sua parte in modo del tutto arbitrario finché questa variazione sia limitata ad un insieme "trascurabile". È proprio la necessità degli insiemi trascurabili a motivare l'accorpamento del σ-ideale $\mathscr{Z}$ alla σ-algebra $\mathscr{M}$.

Date una generica funzione $f : M \to X$ ed una σ-algebra completabile $(\mathscr{M}, \mathscr{Z})$ su M introduciamo tre diverse definizioni di misurabilità per f:

1. f è $(\mathcal{M},\mathcal{Z})$-misurabile se e solo se $f^{-1}(U)$ è $(\mathcal{M},\mathcal{Z})$-misurabile per ogni aperto U di X;

2. f è fortemente $(\mathcal{M},\mathcal{Z})$-misurabile se e solo se esistono una successione di funzioni $\mathcal{M}$-semplici convergenti puntualmente $\mathcal{Z}$-quasi ovunque ad f;

3. f è debolmente $(\mathcal{M},\mathcal{Z})$-misurabile se e solo se $L\circ f$ è $(\mathcal{M},\mathcal{Z})$-misurabile per ogni $L\in X^*$.

Osserviamo che solo nella definizione di misurabilità forte gli insiemi nulli giocano un ruolo fondamentale per via dell'osservazione precedente, ed infatti definiremo l'integrale solamente per le funzioni fortemente misurabili. Diamo qualche risultato elementare per questi concetti di misurabilità.

Proposizione 12.1.4. *Se $(\mathcal{M},\mathcal{Z})\subseteq(\mathcal{N},\mathcal{W})$ ed $f:M\to X$ è (debolmente) $(\mathcal{M},\mathcal{Z})$-misurabile allora è anche (debolmente) $(\mathcal{N},\mathcal{W})$-misurabile.*

Se invece $(\mathcal{M},\mathcal{Z})\subset\subset(\mathcal{N},\mathcal{W})$ ed $f:M\to X$ è fortemente $(\mathcal{M},\mathcal{Z})$-misurabile allora è anche fortemente $(\mathcal{N},\mathcal{W})$-misurabile.

Dimostrazione. Se f è fortemente misurabile esistono s_n $\mathcal{M}$-semplici ed $Z\in\mathcal{Z}$ tali che $s_n(x)\to f(x)$ per ogni $x\in M\setminus Z$.

Grazie alla proposizione 12.1.3 per ogni n esistono $t_n:M\to X$ $\mathcal{N}$-semplice ed $W_n\in\mathcal{W}$ tali che $s_n(x)=t_n(x)$ per ogni $x\in M\setminus W_n$. Poiché $\mathcal{Z}\subseteq W$ abbiamo

$$W=Z\cup\bigcup_{n=1}^{+\infty}W_n\in\mathcal{W}$$

e $t_n(x)\to f$ per ogni $x\in M\setminus W$. ■

Chiaramente ogni funzione misurabile è anche debolmente misurabile, per vedere che le funzioni fortemente misurabili sono anche misurabili abbiamo bisogno di un piccolo risultato intermedio prima:

Proposizione 12.1.5. *Date f_n $(\mathcal{M},\mathcal{Z})$-misurabili convergenti puntualmente $\mathcal{Z}$-quasi ovunque ad una qualche funzione f allora f è automaticamente $(\mathcal{M},\mathcal{Z})$-misurabile indipendentemente da come è definita su Z.*

Dimostrazione. Possiamo supporre che $\mathcal{Z}\subseteq\mathcal{M}$ altrimenti basta sostituire $\mathcal{M}$ con $\mathscr{C}\mathscr{C}(\mathcal{M},\mathcal{Z})$. Preso un qualunque aperto U di X per ogni $r>0$ definiamo $\operatorname{Int}_r(U)=\operatorname{Int}_{D_r}(U)$ come nella (3.1.3) dove

$$D_r=\left\{(x,y)\;\middle|\;\|x-y\|<r\right\}$$

Preso $a\in P=M\setminus Z$ per cui $f(a)\in U$ allora $f(a)\in\operatorname{Int}_{1/m}(U)$ per qualche $m\in\mathbb{N}$ e quindi, essendo quest'ultimo insieme ancora aperto, si ha $f_n(a)\in\operatorname{Int}_{1/m}(U)$ da un certo n in poi. Viceversa se $f_n(a)\in\operatorname{Int}_{1/m}(U)$ per qualche m e da un certo n in poi passando al limite in n avremo che $f(a)\in\operatorname{Int}_{1/(m+1)}(U)$ in quanto per le proprietà degli spazi metrici

$$\operatorname{Cl}\left(\operatorname{Int}_{1/m}(U)\right)\subseteq\operatorname{Int}_{1/(m+1)}(U)$$

e quindi $f(a)\in U$. In formule

$$f^{-1}(U)\setminus Z=\bigcap_{m\in\mathbb{N}}\bigcup_{n=1}^{+\infty}\bigcap_{i\geq n}^{+\infty}\left\{a\in M\setminus Z\;\middle|\;f_j(a)\in\operatorname{Int}_{1/m}(U)\right\}\setminus Z$$

e poiché $\mathcal{Z}\subseteq\mathcal{M}$ avremo che $f^{-1}(U)\setminus Z\in\mathcal{M}$. Ma essendo $f^{-1}(U)\cap Z\in\mathcal{Z}$ abbiamo infine $f^{-1}(U)\in\mathcal{M}$. ■

Osservazione. Per ipotesi stiamo dando per scontato che la successione converga su un insieme prefissato. Se invece abbiamo una successione generica f_n di funzioni misurabili l'insieme massimale su cui f_n converge puntualmente potrebbe non essere misurabile quando X non è separabile. Infatti esisterebbero aperti di $X \times X$ che non possono essere rappresentati tramite unione numerabile di aperti rettangolari. Per questo motivo serve la forte misurabilità.

Diciamo inoltre che f è a **valori $\mathscr{Z}$-separabili** se e solo se esistono $Z \in \mathscr{Z}$ ed un sottospazio $Y \leq X$ chiuso e separabile tali che ed $f(M \setminus Z) \subseteq Y$. Dimostriamo ola il teorema di caratterizzazione della misurabilità:

Teorema 12.1.6 (Pettis)**.** *Una funzione $f : M \to X$ è $(\mathscr{M}, \mathscr{Z})$-fortemente misurabile se e solo se è debolmente $(\mathscr{M}, \mathscr{Z})$-misurabile e a valori $\mathscr{Z}$-separabili.*

Dimostrazione. Supponiamo che f sia il limite puntuale su $M \setminus Z$ di una successione di funzioni semplici $\mathscr{M}$-misurabili s_i. Per quanto detto prima è anche debolmente misurabile mentre posto

$$Y = \mathrm{Cl}\left(\mathrm{span}\left(\bigcup_{i=1}^{+\infty} s_i(M)\right)\right)$$

essendo $s_i(M)$ finiti avremo che Y è chiuso e separabile. Se prendiamo un qualunque $a \in M \setminus Z$ poiché $s_i(a) \in Y$ passando al limite in i avremo che $f(a) \in Y$, dunque $f(M \setminus Z) \subseteq Y$.

Viceversa supponiamo che f sia debolmente misurabile ed $f(a) \in Y$ con Y sottospazio chiuso e separabile per ogni $a \in M \setminus Z$ con $Z \in \mathscr{Z}$. Per il lemma 7.4.4 esisterà una successione totalizzante di funzionali continui $f_n \in Y^*$ tali che $\|f_n\| = 1$, che per il teorema di Hahn-Banach possiamo estendere su tutto X mantenendo inalterata la norma. Dato che l'estremo superiore di una quantità numerabile di funzioni misurabili a valori in $\mathbb{R}$ è ancora misurabile avremo che la funzione

$$g_y : a \in M \setminus Z \to \sup_{n \in \mathbb{N}} f_n\left[f(a) - y\right] = \|f(a) - y\| \in \mathbb{R}$$

è misurabile per ogni $y \in Y$.

Dato che Y è separabile possiamo sempre scegliere una successione densa y_n composta da termini distinti, da cui possiamo definire una successione di funzioni $t_n : Y \to Y$ tali che $t_n(y) = y_{i_n(y)}$ dove

$$i_n(y) = \min\left\{ i \in \mathbb{N} \;\middle|\; \|y - y_i\| \leq \|y - y_j\| \quad \forall 1 \leq j \leq n \right\}$$

Poiché y_i è denso in Y si può facilmente verificare che $t_n(y) \to y$ e soprattutto che $t_n(Y) \subseteq \{y_1, y_2, \ldots, y_n\}$ e quindi le t_n sono funzioni semplici.

Posto allora $r_n = t_n \circ f$ da quanto detto sopra r_n è una funzione semplice convergente puntualmente ad f su $M \setminus Z$. Precedentemente abbiamo osservato che le funzioni $g_{y_i}(a) = \|f(a) - y_i\|$ sono effettivamente delle funzioni misurabili e quindi

$$\begin{aligned} J_{n,i} &= \left\{ x \in M \setminus Z \;\middle|\; r_n(x) = y_i \right\} = \left\{ a \in M \;\middle|\; t_n[f(a)] = y_i \right\} \\ &= \left[\bigcap_{j=1}^{i-1} \left\{ a \in M \setminus Z \;\middle|\; g_{y_i}(a) < g_{y_j}(a) \right\} \cap \bigcap_{k=i+1}^{n} \left\{ a \in P \;\middle|\; g_{y_i}(a) \leq g_{y_k}(a) \right\} \right] \in \sigma(\mathscr{M} \cup \mathscr{Z}) \end{aligned}$$

e per quanto detto prima esisteranno $M_{n,i} \in \mathscr{M}$ tali che $Z_{n,i} = M_{n,i} \triangle J_{n,i} \in \mathscr{Z}$.

Definiamo le funzioni semplici $s_n : M \to X$ in modo tale che

$$s_n(x) = \sum_{i=1}^{n} y_i \mathcal{X}_{M_{n,i}}$$

e l'insieme

$$Z' = Z \cup \bigcup_{n,i} Z_{n,i}$$

Chiaramente s_n è $\mathscr{M}$-misurabile mentre $Z' \in \mathscr{Z}$, inoltre per ogni $x \in M \setminus Z'$ si ha $x \in M \setminus Z$ e $s_n(x) = r_n(x) \to f_n(x)$ e dunque f è $(\mathscr{M}, \mathscr{Z})$-misurabile. ■

Corollario 12.1.7. *Se X è separabile allora le tre definizioni di misurabilità coincidono.*

Corollario 12.1.8. *Data una successione di funzioni f_n $(\mathscr{M}, \mathscr{Z})$-fortemente misurabili e X spazio di Banach, la funzione $f : M \to X$ definita come*

$$f(x) = \begin{cases} \lim_n f_n(x) & \textit{se il limite esiste} \\ 0 & \textit{altrimenti} \end{cases}$$

è ancora $(\mathscr{M}, \mathscr{Z})$-fortemente misurabile.

Dimostrazione. Data che le f_n sono in quantità al più numerabile avremo che le f_n sono debolmente misurabili ed inoltre esistono $Z \in \mathscr{Z}$ ed $Y \leq X$ chiuso e separabile tali che $f_n(M \setminus Z) \subseteq Y$ per ogni $n \in \mathbb{N}$.

Anche lo spazio $Y \times Y$ è separabile, il che implica che la funzione

$$x \in M \to [f_a(x), f_b(x)] \in X \times X$$

è ancora debolmente misurabile e a valori separabili, quindi è fortemente misurabile e l'insieme

$$C = \{x \in X \mid f_n(x) \text{ converge}\} = \bigcap_{n \in \mathbb{N}} \bigcup_{l \in \mathbb{N}} \bigcap_{a,b \geq l} \left\{ x \in M \;\middle|\; \left\| f_a(x) - f_b(x) \right\| < \frac{1}{n} \right\}$$

appartiene ad $\mathscr{C}\mathscr{C}(\mathscr{M}, \mathscr{Z})$.

La funzione f definita nell'enunciato è dunque $(\mathscr{M}, \mathscr{Z})$-misurabile in quanto limite puntuale delle funzioni misurabili $\hat{f}_n = f_n \mathcal{X}_C$, per la chiusura di Y si ha anche $f(M \setminus Z) \subseteq Y$ e dunque f è fortemente misurabile. ■

Date una σ-algebra completabile $(\mathscr{M}, \mathscr{Z})$ ed una σ-algebra $\mathscr{N}$ diremo che $\mathscr{N}$ è **densa** in $(\mathscr{M}, \mathscr{Z})$ se e solo se $\mathscr{N} \subseteq (\mathscr{M}, \mathscr{Z})$ e $\mathscr{C}\mathscr{C}(\mathscr{N}, \mathscr{Z}) = \mathscr{C}\mathscr{C}(\mathscr{M}, \mathscr{Z})$.

Corollario 12.1.9 (troncamento dell'errore)**.** *Se X è uno spazio di Banach allora per ogni $f : M \to X$ $(\mathscr{M}, \mathscr{Z})$-fortemente misurabile e per ogni σ-algebra $\mathscr{N}$ densa in $(\mathscr{M}, \mathscr{Z})$ esiste $g : M \to X$ $\mathscr{N}$-fortemente misurabile per cui $\{f \neq g\} \in \mathscr{Z}$.*

Dimostrazione. Per densità abbiamo che f è $(\mathscr{N}, \mathscr{Z})$-fortemente misurabile e dunque esistono s_n funzioni $\mathscr{N}$-semplici che convergono puntualmente ad f su $M \setminus Z$ per qualche $Z \in \mathscr{Z}$. Chiaramente le s_n sono $\mathscr{N}$-fortemente misurabili e per il corollario precedente esiste una funzione g $\mathscr{N}$-fortemente misurabile che è uguale al limite delle s_n quando esiste.

Ma allora $g = f$ su $M \setminus Z$ e così $\{f \neq g\} \subseteq Z$ appartiene a $\mathscr{Z}$. ■

12.2 L'integrale di Bochner

Una **misura esterna** sull'insieme non vuoto M è una qualunque applicazione $\mu : \mathcal{P}(M) \to [0,+\infty]$ tale che $\mu(\emptyset) = 0$ e $\mu(A) \leq \sum_{i=1}^{+\infty} \mu(A_i)$ se $A \subseteq \bigcup_{i\in\mathbb{N}} A_i$. Le proprietà base delle misure esterne dimostrate in [8] saranno date per scontato.

Definiamo i seguenti insiemi:

$$\mathcal{M}(\mu) = \left\{A \subseteq M \,\middle|\, \mu(B) = \mu(B\cap A) + \mu(B\setminus A)\ \forall B \subseteq M\right\}$$
$$\mathcal{Z}(\mu) = \left\{A \subseteq M \,\middle|\, \mu(A) = 0\right\}$$

ovvero $\mathcal{M}(\mu)$ è la famiglia dei sottoinsiemi μ-misurabili di M e $\mathcal{Z}(\mu)$ la famiglia dei sottoinsiemi di misura nulla. Chiaramente $\mathcal{Z}(\mu) \subseteq \mathcal{M}(\mu)$, $\mathcal{M}(\mu)$ è una σ-algebra mentre $\mathcal{Z}(\mu)$ è un σ-ideale, dunque ogni misura esterna μ determina univocamente la σ-algebra completabile $(\mathcal{M}(\mu), \mathcal{Z}(\mu))$. Utilizzeremo il simbolo μ anche per indicare la σ-algebra completabile $(\mathcal{M}(\mu), \mathcal{Z}(\mu))$ mentre scriveremo $(\mathcal{M}, \mu)$ per indicare $(\mathcal{M}, \mathcal{Z}(\mu))$, quindi parlare di μ-misurabilità, di intersezioni ed inclusioni di misure esterne con σ-algebre completabili assume un senso.

Una utilissima conseguenza del lemma di Dynkin è il seguente corollario che fornisce un criterio di uguaglianza tra misure esterne:

Corollario 12.2.1. *Consideriamo la σ-algebra $\mathcal{M}$ su M generata dal π-sistema $\mathcal{I}$ contenente l'insieme M. Se $\mathcal{M} \subseteq \mu \cap \nu$, con μ e ν misure esterne finite che coincidono sugli elementi di $\mathcal{I}$, allora $\mu(E) = \nu(E)$ per ogni $E \in \mathcal{M}$.*

Dimostrazione. Se $\mathcal{D}$ è la famiglia di elementi di $\mathcal{M}$ su cui le due misure coincidono allora essendo $M \in \mathcal{I}$ e le misure finite la famiglia $\mathcal{D}$ è un λ-sistema, la tesi segue perciò dal lemma di Dynkin. ■

Teorema 12.2.2 (Egoroff)**.** *Data una successione di funzioni μ-misurabili $f_n : M \to X$ convergenti puntualmente μ-quasi ovunque ad $f : M \to X$ se $\mu(M) < +\infty$ allora per ogni $\varepsilon > 0$ esiste un sottoinsieme N μ-misurabile di M tale che $\mu(M \setminus N) < \varepsilon$ ed f_n converge uniformemente ad f su N.*

Dimostrazione. Definiamo

$$C_{ij} = \bigcup_{k=j}^{+\infty} \left\{x \in M \,\middle|\, \left\|f_k(x) - f(x)\right\| > 2^{-i}\right\}$$

allora C_{ij} è una successione di insiemi misurabili decrescente in j, posto $B_i = \bigcap_j C_{ij}$ per ipotesi avremo che $\mu(B_i) = 0$ per ogni i. Essendo $\mu(M) < +\infty$ avremo

$$\lim_{j\to+\infty} \mu\left(C_{ij}\right) = \mu(B_i) = 0$$

ed esisterà così $m_i \in \mathbb{N}$ tale che $\mu\left(C_{in(i)}\right) \leq 2^{-i}\varepsilon$.

Posto allora $N = M \setminus \bigcup_{i=1}^{+\infty} C_{in_i}$ avremo così

$$\mu(M\setminus N) \leq \sum_{i=1}^{+\infty} \mu\left(C_{in_i}\right) \leq \varepsilon$$

ed $\left\|f_k(x) - f(x)\right\| \leq 2^{-i}$ per ogni $x \in N$, $k > n_i$ in altre parole f_n converge uniformemente ad f su N. ■

Corollario 12.2.3. *Se $f : M \to X$ è fortemente misurabile con $\mu(M) < +\infty$ allora per ogni $\varepsilon > 0$ esiste $N \subseteq M$ misurabile tale che $\mu(M \setminus N) < \varepsilon$ ed $f(N)$ è totalmente limitato in X.*

Dimostrazione. Dato che f è fortemente misurabile esisterà una successione di funzioni semplici s_n convergente ad f puntualmente quasi ovunque. Per il teorema di Egoroff per ogni $\varepsilon > 0$ esisterà N tale che s_n converge ad f uniformemente su di esso.

Fissato così un qualche altro $\varepsilon' > 0$ esisterà $k \in \mathbb{N}$ per cui $\left\| f(x) - s_k(x) \right\| < \varepsilon'$ per ogni $x \in N$, e dato che s_k è semplice ciò equivale ad affermare che $f(N)$ è contenuto nell'unione di un numero finito di palle aperte di raggio al più ε'. Dato che ε' non dipende da ε avremo così che $f(N)$ è totalmente limitato. ■

Possiamo adesso parlare finalmente di integrabilità. Innanzitutto diremo che una funzione semplice μ-misurabile s da M in X è μ-**integrabile** se e solo se per ogni $x \in X$ diverso da 0 si ha $\mu(\{s = x\}) < +\infty$, e in tal caso si pone

$$\int_M s \, d\mu = \sum_{x \in s(M) \setminus \{0\}} x \mu(\{s = x\}) \in X$$

e poniamo $\int_M 0 \, d\mu = 0$ banalmente. Non è difficile verificare che se s è una funzione semplice integrabile ed $L : X \to Y$ è una funzione lineare non necessariamente continua allora $L \circ s$ è ancora semplice ed integrabile con

$$\int_M L \circ s \, d\mu = L\left(\int_M s \, d\mu \right) \tag{12.2.1}$$

Proposizione 12.2.4. *Per ogni funzione semplice integrabile $s : M \to X$ si ha*

$$\left\| \int_M s \, d\mu \right\| \leq \int_M \|s\| \, d\mu$$

Dimostrazione. Per ogni $L \in X^*$ con norma unitaria si ha

$$L\left(\int_M s \, d\mu \right) = \int_M L \circ s \, d\mu \leq \int_M \|s\| \, d\mu$$

e scegliendo opportunamente la funzione L per la proposizione 7.2.2 abbiamo la disuguaglianza cercata. ■

Diremo che una funzione $f : M \to X$ (fortemente misurabile) è **Bochner-integrabile** se e solo se esiste una successione di funzioni semplici μ-integrabili $s_n : M \to X$ convergenti puntualmente μ-quasi ovunque ad f tale che

$$\lim_{n \to +\infty} \int_M \left\| f - s_n \right\| d\mu = 0$$

Dimostriamo adesso che

Proposizione 12.2.5. *Se f è una funzione Bochner-integrabile ed s_n è una qualunque successione di funzioni semplici μ-integrabili convergenti ad f quasi ovunque e $\int_M \left\| s_n - f \right\| d\mu \to 0$. Allora il limite*

$$\lim_{n \to +\infty} \int_M s_n \, d\mu$$

*esiste sempre in X e non dipende dalla scelta della successione s_n. Tale limite sarà detto **integrale di Bochner** di f e si indicherà con il simbolo $\int_M f \, d\mu$.*

Dimostrazione. Osserviamo innanzitutto che per ogni $m, n \in \mathbb{N}$

$$\left\| \int_M s_m \, d\mu - \int_M s_n \, d\mu \right\| \leq \int_M \| s_m - s_n \| \, d\mu \leq \int_M \| s_m - f \| \, d\mu + \int_M \| s_n - f \| \, d\mu$$

dunque $\int_M s_n \, d\mu$ è una successione di Cauchy nello spazio di Banach X e quindi convergerà in X.

Se ora prendiamo un'altra successione t_n che soddisfa le medesime ipotesi allora

$$\left\| \int_M s_n \, d\mu - \int_M t_n \, d\mu \right\| \leq \int_M \| s_n - t_n \| \, d\mu \leq \int_M \| s_n - f \| \, d\mu + \int_M \| t_n - f \| \, d\mu$$

e dato che entrambe possiedono limite essi dovranno coincidere. ∎

Chiaramente le funzioni semplici integrabili sono anch'esse Bochner-integrabili ed il loro integrale coincide con quello già noto. Se ora $L \in L(X, Y)$ è una funzione lineare e continua tra spazi di Banach ed s_n è una successione di funzioni semplici integrabili che soddisfano le ipotesi di integrabilità di Bocher per una funzione f allora

- $L \circ s_n \to L \circ f$ puntualmente quasi ovunque su M;
- $\int_M \| L \circ s_n - L \circ f \| \, d\mu \leq \|L\| \int_M \| s_n - f \| \, d\mu \to 0.$

Applicando così la (12.2.1) e la proposizione precedente

$$\int_M L \circ f \, d\mu = \lim_{n \to +\infty} \int_M L \circ s_n \, d\mu = \lim_{n \to +\infty} L\left(\int_M s_n \, d\mu \right) = L\left(\int_M f \, d\mu \right)$$

e di conseguenza

$$\left\| \int_M f \, d\mu \right\| \leq \int_M \|f\| \, d\mu$$

Da ciò segue che

Proposizione 12.2.6. *Se $f : M \to X$ è Bochner-integrabile ed $\mu(M) > 0$ allora posto $I = f(M)$ si ha*

$$\frac{1}{\mu(M)} \int_M f \, d\mu \in \overline{\operatorname{con}}(I)$$

Dimostrazione. Supponiamo inizialmente che $\mu(M) < +\infty$. Per assurdo la quantità a sinistra non appartiene ad $\overline{\operatorname{con}}(I)$, per il secondo teorema di Mazur esistono $L \in X^*_{\mathbb{R}}$ (che come sappiamo corrisponde univocamente a qualche elemento di X^*) ed $\alpha \in \mathbb{R}$ tali che

$$\frac{1}{\mu(M)} \int_M L \circ f \, d\mu < \alpha \leq L[f(a)]$$

per ogni $a \in M$. Integrando di nuovo entrambi i membri si ha

$$\int_M L \circ f \, d\mu < \alpha \mu(M) \leq \int_M L \circ f \, d\mu$$

assurdo.

Se invece $\mu(M) = +\infty$ l'asserto equivale ad affermare che $0 \in \operatorname{Cl}(I) \subseteq \overline{\operatorname{con}}(I)$. Se infatti non fosse vero allora esisterebbe $\delta > 0$ tale che $\|f(x)\| > \delta$ per ogni $x \in M$. Poiché f è Bochner-integrabile deve esistere una funzione semplice μ-integrabile s tale che $\int_M \|f - s\| \, d\mu <$

$+\infty$, poiché $\mu(M) = +\infty$ esisterà $N \subseteq M$ tale che $\mu(N) = +\infty$ e $\left\|f(x) - s(x)\right\| \leq \delta/2$ per ogni $x \in N$.

Ma allora

$$\|s(x)\| \geq \left\|f(x)\right\| - \left\|f(x) - s(x)\right\| \geq \frac{\delta}{2}$$

per ogni $x \in N$, poiché s è semplice esisterà $y \in X$ diverso da 0 per cui $\mu\left(\{s = y\}\right) = +\infty$ il che è assurdo poiché s è μ-integrabile. Per questa ragione $0 \in \mathrm{Cl}(I)$. ■

Corollario 12.2.7. *Sia $f : M \to X$ una funzione Bochner-integrabile ed $A \subseteq X$ che soddisfa la seguente proprietà: esiste una successione di sottoinsiemi chiusi e convessi $C_1, C_2, \dots$ di X tali che*

$$X \setminus A = \bigcup_{i=1}^{+\infty} C_i$$

Se $\frac{1}{\mu(N)} \int_N f\, d\mu \in A$ per ogni $N \subseteq M$ misurabile tale che $\mu(N) > 0$ allora $\mu\Big(\{x \in M \mid f(x) \notin A\}\Big) = 0$.

Dimostrazione. Per ogni $i \in \mathbb{N}$ definiamo $N_i = \{f \in C_i\}$ che è misurabile in quanto f è μ-misurabile, se fosse $\mu(N_i) > 0$ allora il teorema precedente implicherebbe che $\frac{1}{\mu(N_i)} \int_{N_i} f\, d\mu \in C_i$ che è impossibile per ipotesi, perciò $\mu(N_i) = 0$.

Allora essendo $\{f \notin A\} = \bigcup_{i\in\mathbb{N}} N_i$ anche questo insieme deve avere misura nulla. ■

Corollario 12.2.8. *Sia $f : M \to X$ μ-misurabile tale che $\int_M \left\|f\right\| d\mu < +\infty$ (in particolare $\left\|f\right\|$ è Bochner-integrabile), se esiste $L > 0$ tale che $\frac{1}{\mu(N)} \int_N \left\|f\right\| d\mu \leq L$ per ogni $N \subseteq M$ misurabile con $\mu(N) > 0$ allora $\left\|f\right\| \leq L$ quasi ovunque su M.*

Data una misura esterna μ su M ed $F : M \to P$ applicazione generica possiamo definire su P una nuova misura esterna $F_\# \mu$ in modo tale che $F_\# \mu(I) = \mu[F^{-1}(I)]$ chiaramente se $F^{-1}(I)$ è μ-misurabile allora I è $F_\# \mu$-misurabile. Abbiamo il seguente risultato

Proposizione 12.2.9 (Cambio di variabili)**.** *Siano μ misura esterna su M, $F : M \to P$ ed $g : P \to X$ tale che $g \circ F$ è μ-fortemente misurabile. Allora g è $F_\#\mu$-fortemente misurabile, inoltre è $F_\#\mu$-Bochner integrabile se e solo se $g \circ F$ è μ-Bochner integrabile.*

In tal caso

$$\int_M g \circ F\, d\mu = \int_P g\, dF_\#\mu \tag{12.2.2}$$

Dimostrazione. Innanzitutto osserviamo che g e g' differiscono su un insieme di misura $F_\#\mu$ nulla se e solo se $g \circ F$ e $g' \circ F$ differiscono di un insieme di misura μ nulla in quanto

$$\{g \circ F \neq g' \circ F\} = F^{-1}\left(\{g' \neq g\}\right)$$

e dunque $F_\#\mu\left(\{g \neq g'\}\right) = \mu\left(g \circ F \neq g' \circ F\right)$. Possiamo perciò sempre sostituire g con un'altra funzione che vi coincide $F_\#\mu$-quasi ovunque senza temere ripercussioni sul problema. Inoltre $F_\#\mu[P \setminus F(M)] = 0$ quindi g può variare liberamente sui punti che non appartengono all'immagine di F. Possiamo perciò assumere F suriettiva.

Dato che $g \circ F$ è fortemente misurabile è a valori quasi ovunque separabili ed è anche μ-misurabile, ovvero gli insiemi $F^{-1}[g^{-1}(U)] \in \mathscr{M}(\mu)$ per ogni $U \subseteq X$ aperto. Perciò $g^{-1}(U)$ è anche $F_\#\mu$-misurabile ed avrà valori $F_\#\mu$-quasi ovunque separabili e così g è $F_\#\mu$-fortemente misurabile per il teorema di Pettis.

Supponiamo ora che $g(x) = \sum_{i=1}^{n} x_i \mathcal{X}_{A_i}(x)$ sia semplice $F_{\#}\mu$-integrabile con f una funzione arbitraria, allora si avrà

$$g \circ F = \sum_{i=1}^{n} x_i \mathcal{X}_{F^{-1}(A_i)}$$

e dunque

$$\int_M g \circ F \, d\mu = \sum_{i=1}^{n} x_i \mu\left[F^{-1}(A_i)\right] = \sum_{i=1}^{n} x_i F_{\#}\mu\left[A_i\right] = \int_P g \, dF_{\#}\mu$$

Lavorando su successioni di funzioni semplici ed applicando la forma classica della proposizione per funzioni a valori positivi si dimostra la (12.2.2) per g fortemente misurabile. ■

Finora non abbiamo fatto altre considerazioni sul secondo membro della disuguaglianza precedente, che quindi potrebbe essere anche infinito senza però creare problemi. In realtà se f è Bochner-integrabile allora tale integrale è necessariamente finito: per definizione possiamo sempre trovare una funzione semplice ed integrabile s per cui $\int_M \|f - s\| \, d\mu < \varepsilon$ e dunque

$$\int_M \|f\| \, d\mu \leq \int_M \|s - f\| \, d\mu + \int_M \|s\| \, d\mu < +\infty$$

Il prossimo teorema afferma che una funzione fortemente misurabile $f : M \to X$ è Bochner-integrabile se e solo se

$$\int_M \|f\| \, d\mu < +\infty$$

prima di dimostrare tale risultato soffermiamoci sulle funzioni misurabili per cui l'integrale della norma sia finito. Dato che f può essere approssimato tramite una successione di funzioni semplici integrabili, e che quindi sono diverse da 0 solo su un insieme di misura finita, è lecito chiedersi se tale proprietà possa estendersi anche ad f. La risposta è affermativa e dimostreremo che l'insieme $\{f \neq 0\}$ è a tutti gli effetti unione di una quantità numerabile di insiemi di misura finita.

Nello specifico diremo che un sottoinsieme F di M è σ-finito rispetto a μ se e solo se esistono una successione F_n di sottoinsiemi μ-misurabili di M tali che $\mu(F_n) < +\infty$ ed $F = \bigcup_{n=1}^{+\infty} F_n$.

Proposizione 12.2.10. *Data una funzione $f : M \to X$ μ-misurabile tale che $\int_M \|f\| \, d\mu < +\infty$ allora esiste un sottoinsieme N di M σ-finito tale che $f(a) = 0$ per ogni $a \in M \setminus N$.*

Inoltre se f è anche $(\mathcal{M}, \mu)$-misurabile con $\mathcal{M} \subseteq \mathcal{M}(\mu)$ allora è possibile prendere $N \in \mathcal{M}$.

Dimostrazione. Il caso $X = \mathbb{R}$ viene trattato abbondantemente nei libri di teoria della misura, da cui ponendo $g(a) = \|f(a)\|$ discende anche il caso generale. ■

Su ogni insieme non vuoto M possiamo definire la seguente misura esterna $\#(A)$ che associa la cardinalità di A se è finito, altrimenti vale $+\infty$. Tutti i sottoinsiemi di M sono $\#$-misurabili e quindi ogni applicazione $f : M \to X$ è $\#$-misurabile. Ma allora se $\int_M \|f\| \, d\# < +\infty$ la proposizione 12.2.10 ci dice essenzialmente che f è diversa da 0 solamente su un sottoinsieme finito o numerabile di M che indicheremo semplicemente con la successione $\{a_1, a_2, \ldots\}$ composta esclusivamente da termini distinti.

Se questo insieme è finito allora f è automaticamente semplice e integrabile con

$$\int_M f \, d\# = \sum_{i=1}^{n} f(a_i)$$

se invece è numerabile allora possiamo costruire facilmente una successione di funzioni semplici f_n integrabili che convergono puntualmente ad f per cui

$$\int_M \|f_n - f\| \, d\# = \sum_{i=n+1}^{+\infty} \|f(a_i)\| \to 0$$

per cui avremo che

$$\int_M f \, d\# = \sum_{n=1}^{+\infty} f(a_i) \in X$$

dato che X è pur sempre uno spazio di Banach. Dato che il valore del limite a destra non dipende dall'enumerazione a_i quando $\mu = \#$ scriveremo brevemente che

$$\sum_{a\in M} f(a) = \int_M f \, d\#$$

quando l'integrale a destra è finito.

Teorema 12.2.11. *Se $f : M \to X$ è fortemente misurabile allora f è Bochner-integrabile se e solo se $\int_M \|f\| \, d\mu < +\infty$.*

Dimostrazione. Sia f fortemente misurabile per cui $\int_M \|f\| \, d\mu < +\infty$, sia s_n una successione di funzioni semplici convergenti puntualmente μ-quasi ovunque ad f. Per la proposizione 12.2.10 possiamo supporre che esistono dei sottoinsiemi misurabili M_n di M tali che

- $M_n \subseteq M_{n+1}$;
- $\mu(M_n) < +\infty$;
- $M = \bigcup_{i=1}^{+\infty} M_n$

Di conseguenza possiamo definire le funzioni $t_n : M \to X$ in modo tale che

$$t_n(a) = \begin{cases} s_n(a) & \text{se } a \in M_n \text{ e } \|s_n(a)\| \leq 2\|f(a)\| \\ 0 & \text{altrimenti} \end{cases}$$

non è difficile dimostrare che anche t_n converge puntualmente quasi ovunque ad f e inoltre $\{t_n = x\} = \{s_n = x\} \cap M_n \cap \left\{2\|f\| \geq \|x\|\right\}$ per ogni $x \in X \setminus \{0\}$ e quindi t_n è integrabile.

Per verificare che f è Bochner-integrabile possiamo considerare le funzioni

$$\begin{aligned} \phi_n(a) &= \|t_n(a) - f(a)\| \\ \psi(a) &= 3\|f(a)\| \end{aligned}$$

che assumono valori in $\mathbb{R}$. Poiché ϕ_n converge puntualmente quasi ovunque a 0, $\phi_n \leq \psi$ puntualmente ed $\int_M \psi \, d\mu < +\infty$ per il teorema della convergenza dominata usuale si ha

$$\lim_{n\to+\infty} \int_M \|t_n - f\| \, d\mu = 0$$

e quindi f è Bochner-integrabile. ■

Corollario 12.2.12. *Data una funzione $f : M \to X$ che è $(\mathscr{M}, \mu)$-fortemente misurabile con $\mathscr{M} \subseteq \mu$ abbiamo che f è Bochner-integrabile se e solo se esiste una successione di funzioni semplici s_n $\mathscr{M}$-misurabili e μ-integrabili convergenti μ-quasi ovunque ad f per cui si ha $\lim_{n\to+\infty} \int_M \|f - s_n\| \, d\mu = 0$.*

Grazie a questo risultato possiamo integrare su M anche funzioni a valori in X che però non sono definite su tutto M ma su un sottoinsieme μ-misurabile P con $\mu(M \setminus P) = 0$. Quindi quando sceglieremo una funzione $f : M \to X$ che sia fortemente misurabile considereremo anche il caso in cui f non sia definita su tutto M ma solo su un suo sottoinsieme il cui complementare ha misura nulla.

Corollario 12.2.13. *Data una successione $x_n \in X$ spazio di Banach, se $\sum_{n=1}^{+\infty} \|x_n\| < +\infty$ allora esiste il limite $\lim_{n\to+\infty} \sum_{i=1}^{n} x_i$ in X.*

Corollario 12.2.14 (convergenza dominata)**.** *Sia f_n una successione di funzioni fortemente misurabili convergenti puntualmente ad una certa funzione f. Se esiste $\psi : M \to \mathbb{R}$ tale che $\|f_n(a)\| \le \psi(a)$ per quasi ogni $a \in M$ ed $\int_M \psi \, d\mu < +\infty$ allora*

$$\int_M \|f_n - f\| \, d\mu \to 0$$

e in particolare le funzioni f_n, f sono Bochner-integrabili ed

$$\lim_{n\to+\infty} \int_M f_n \, d\mu = \int_M f \, d\mu$$

Dimostrazione. Il teorema precedente ci garantisce che le funzioni in gioco sono Bochner-integrabili, quindi basta applicare il teorema della convergenza dominara usuale alla successione $\phi_n = \|f_n - f\|$ in quanto $\|f_n - f\| \le \|f_n\| + \|f\| \le 2\psi$. ■

Sfruttando invece il teorema di Fubini-Tonelli (si veda [8]) per le funzioni a valori reali è possibile estenderlo tranquillamente anche alle funzioni a valori vettoriali.

Teorema 12.2.15 (Fubini-Tonelli)**.** *Siano μ, η due misure esterne su M e P rispettivamente e sia $\mu \times \eta$ la misura prodotto su $N \times P$. Se prendiamo una funzione $f : M \times P \to X$ $(\mu \times \eta)$-fortemente misurabile allora valgono le seguenti implicazioni:*

- *se la quantità $\int_M \left[\int_P \|f(x,y)\| \, d\eta(y) \right] d\mu(x)$ è finita allora f è Bochner-integrabile su $M \times P$;*
- *se f è Bochner-integrabile su $M \times P$ allora anche le funzioni*

$$x \in M \to \int_P f(x,y) \, d\eta(y)$$
$$y \in P \to \int_M f(x,y) \, d\mu(x)$$

sono ben definite per quasi ogni $x \in M$, $y \in P$, sono Bochner-integrabili ed

$$\begin{aligned}\int_{M\times P} f \, d(\mu \times \eta) &= \int_M \left[\int_P f(x,y) \, d\eta(y) \right] d\mu(x) \\ &= \int_P \left[\int_M f(x,y) \, d\mu(x) \right] d\eta(y)\end{aligned}$$

Dimostrazione. Si applichino i teoremi in [8, pp. 22-26] ed in [14, pp. 164-166] al teorema 12.2.11. ■

12.3 Spazi L_p

Definizione 12.3.1. Dati un insieme non vuoto M, una misura esterna μ su M ed una σ-algebra $\mathscr{M}$ tale che $\mathscr{M} \subseteq \mu$ indichiamo con $\mathscr{S}(\mathscr{M},\mu,X)$ lo spazio delle funzioni semplici $\mathscr{M}$-misurabili e μ-integrabili.

Inoltre per ogni $0 < p < +\infty$ indichiamo con $L_p(\mathscr{M},\mu,X)$ lo spazio delle funzioni $f : M \to X$ $(\mathscr{M},\mu)$-fortemente misurabili tali che $\int_M \|f\|^p \, d\mu < +\infty$. Definiamo anche

$$D_{L_p(O)}(f) = \int_O \|f\|^p \, d\mu$$
$$\|f\|_{L_p(O)} = \sqrt[p]{\int_O \|f\|^p \, d\mu}$$

per ogni $O \subseteq M$ μ-misurabile. Quando $\mathscr{M} = \mathscr{M}(\mu)$ ometteremo il simbolo $\mathscr{M}$.

Se $\mu = \#$ è la misura cardinalità scriveremo $\ell^p(M,X)$ al posto di $L_p(\#,X)$. Osserviamo che per ogni $f,g \in L_p$

$$\int_M \|f-g\|^p \, d\mu = 0 \Leftrightarrow \mu\Big(\{a \in M \,|\, f(a) \neq g(a)\}\Big) = 0$$

e dunque pare logico identificare le funzioni in L_p che coincidono a meno di un insieme di misura nulla.

Vogliamo dimostrare ora che se $p \geq 1$ allora L_p è uno spazio di Banach con norma $\|\cdot\|_{L_p}$ mentre se $0 < p < 1$ lo spazio L_p è semplicemente un F-spazio con parametrica D_{L_p}. Per dimostrare la disuguaglianza triangolare abbiamo bisogno prima di dimostrare le seguenti disuguaglianze:

Proposizione 12.3.2. *Date due numeri reali $s,t \geq 0$ per ogni $p > 0$ si ha:*

- *se $p \geq 1$ allora*
$$s^p + t^p \leq (s+t)^p \leq 2^{p-1}\left(s^p + t^p\right) \tag{12.3.1}$$
- *se $0 < p \leq 1$ allora*
$$2^{p-1}\left(s^p + t^p\right) \leq (t+s)^p \leq s^p + t^p \tag{12.3.2}$$

Dimostrazione. Se $p \geq 1$ la funzione $t \to t^p$ è convessa, mentre se $0 < p \leq 1$ la funzione $t \to -t^p$ è convessa. Dimostreremo solo la (12.3.1) in quanto la (12.3.2) si dimostra in modo del tutto analogo.

Per convessità si ha innanzitutto

$$\left(\frac{s+t}{2}\right)^p \leq \frac{s^p + t^p}{2}$$

e quindi $(s+t)^p \leq 2^{p-1}\left(s^p + t^p\right)$, per dimostrare l'altra disuguaglianza osserviamo che

$$s = \frac{s}{s+t}(s+t) + \frac{t}{s+t}0$$
$$t = \frac{s}{s+t}0 + \frac{t}{s+t}(s+t)$$

quando s e t non sono entrambe nulle. Applichiamo la convessità della funzione potenza alle uguaglianze di sopra e sommando avremo così

$$s^p + t^p \leq \frac{s}{s+t}(s+t)^p + \frac{t}{s+t}0 + \frac{s}{s+t}0 + \frac{t}{s+t}(s+t)^p = (s+t)^p$$

come volevasi dimostrare. ■

Dalla (12.3.2) segue immediatamente che D_{L_p} soddisfa la disuguaglianza triangolare quando $p \leq 1$, inoltre $D_{L_p}(f) = 0$ se e solo se $f \cong 0$ e per dimostrare che D_{L_p} è una parametrica dobbiamo solamente dimostrare la continuità del prodotto. Poiché $D_{L_p}(\lambda f) = |\lambda|^p D_{L_p}(f)$ per ogni $\lambda \in \mathbb{K}$ avremo che

$$D_{L_p}(\lambda_n f_n - \lambda f) \leq |\lambda_n|^p D_{L_p}(f_n - f) + |\lambda_n - \lambda|^p D_{L_p}(f)$$

e quindi D_{L_p} è una parametrica ed L_p è uno spazio vettoriale metrico quando $p \leq 1$.

La dimostrazione che useremo per il caso $p \geq 1$ ricalca fedelmente la procedura classica usata nella teoria classica dell'integrazione astratta. Abbiamo infatti che

Teorema 12.3.3 (Hölder). *Sia (M, μ) spazio di misura, X, Y spazi normati su $\mathbb{K}$ ed $f : M \to X$, $g : M \to Y$ funzioni misurabili. Allora per ogni spazio normato Z e per ogni $F \in L^2(X, Y; Z)$ (bilineare e continua) vale la seguente disuguaglianza*

$$\int_M \|F[f,g]\| \, d\mu \leq \|F\| \sqrt[p]{\int_M \|f\|^p \, d\mu} \sqrt[q]{\int_M \|g\|^q \, d\mu} \tag{12.3.3}$$

per ogni $p, q > 1$ con $1/p + 1/q = 1$.

Dimostrazione. Si ha per continuità $\|F[f,g]\| \leq \|F\| \|f\| \|g\|$ e la tesi segue dalla disuguaglianza di Hölder classica. ■

Come per il caso $X = \mathbb{R}$ avremo che $\|f + g\|_{L_p} \leq \|f\|_{L_p} + \|g\|_{L_p}$ e quindi L_p è uno spazio normato quando $p \geq 1$.

Per dimostrare la completezza di $L_p(\mathcal{M}, \mu, X)$ abbiamo bisogno prima di alcuni risultati preparatori.

Proposizione 12.3.4. *Date una successione di funzioni $f_n \in L_p(\mathcal{M}, \mu, X)$ con $p > 0$ ed f_n convergente puntualmente quasi ovunque ad una qualche funzione $f : M \to X$, se f_n è una successione di Cauchy rispetto alla norma L_p allora $f \in L_p$ ed f_n converge ad f anche in L_p.*

Dimostrazione. Chiaramente f è automaticamente $(\mathcal{M}, \mu)$-fortemente misurabile, dobbiamo quindi verificare solamente la convergenza in L_p. Per ogni $\varepsilon > 0$ esiste $N \in \mathbb{N}$ tale che $D_{L_p}(f_n - f_m) < \varepsilon$ per ogni $m, n > N$. D'altronde $\lim_{m \to +\infty} \|f_m(a) - f_n(a)\| = \|f(a) - f_n(a)\|$ per μ-quasi ogni $a \in M$, $n \geq N$, dunque per il lemma di Fatou [8, pagina 19] avremo che per ogni $n \geq N$

$$\int_M \|f - f_n\|^p \, d\mu \leq \liminf_{m \to +\infty} \int_M \|f_m - f_n\|^p \, d\mu \leq \varepsilon$$

Per concludere che $f_n \to f$ in L_p ci basta solamente mostrare che $f \in L_p$, ma ciò deriva immediatamente dalla seguente disuguaglianza

$$D_{L_p}(f) \leq c_p D_{L_p}(f - f_N) + c_p D_{L_p}(f_N) < +\infty$$

dove $c_p = 1 \vee 2^{-p}$. ■

Lemma 12.3.5. *Data una successione di Cauchy $f_n \in L_p\left(\mathscr{M},\mu,X\right)$ esiste un'estratta f_{n_j} che converge ad una qualche funzione fortemente misurabile $f \in L_p$ puntualmente quasi ovunque.*

Dimostrazione. Possiamo sempre trovare un'estratta n_i tale che

$$D_{L_p}\left(f_{n_{i+1}} - f_{n_i}\right) \leq 2^{-i}$$

e poniamo

$$\begin{aligned} h_i(a) &= \left\|f_{n_1}(a)\right\| + \sum_{j=1}^{i-1} \left\|f_{n_{j+1}}(a) - f_{n_j}(a)\right\| \\ h(a) &= \left\|f_{n_1}(a)\right\| + \sum_{j=1}^{+\infty} \left\|f_{n_{j+1}}(a) - f_{n_j}(a)\right\| \end{aligned}$$

allora $\left\|f_{n_i}(a)\right\| \leq h_i(a) \to h(a)$ puntualmente. Per il lemma di Fatou avremo che

$$\int_M h^p \, d\mu \leq \liminf_{i\to+\infty} \int_M h_i^p \, d\mu \leq \sum_{j=1}^{+\infty} 2^{-j} = 1$$

Da ciò segue chiaramente che esiste $P \subseteq M$ misurabile tale che $\mu(M \setminus P) = 0$ ed $h(a) < +\infty$ per ogni $a \in P$, il corollario 12.2.13 ci dice allora che $\lim_{i\to+\infty} f_{n_i}(a) = f(a) \in X$ per ogni $a \in P$. Se poniamo $f(a) = 0$ per ogni $a \in M \setminus P$ avremo che f_{n_i} converge puntualmente quasi ovunque alla funzione f che sarà di conseguenza fortemente misurabile in quanto limite puntuale di funzioni fortemente misurabili. ■

Abbiamo di conseguenza il seguente risultato:

Teorema 12.3.6. *Se X è uno spazio di Banach allora $L_p(\mathscr{M},\mu,X)$ è uno spazio di Banach per ogni $p \geq 1$. Se inoltre X è uno spazio di Hilbert allora $L_2(\mathscr{M},\mu,X)$ è uno spazio di Hilbert con prodotto scalare*

$$(f : g) = \int_M \left(f(x) : g(x)\right) d\mu(x)$$

Dimostrazione. Se X è uno spazio di Hilbert ci basta verificare solamente che il prodotto scalare è ben definito. Dato che il prodotto scalare è continuo l'applicazione $x \in M \to \left(f(x) : g(x)\right)$ è fortemente misurabile, per le disuguaglianza di Cauchy-Schwartz e di Hölder abbiamo

$$\begin{aligned} \left|\int_M \left(f(x) : g(x)\right) d\mu(x)\right| \leq \int_M \left|\left(f(x) : g(x)\right)\right| d\mu(x) \leq \int_M \left\|f(x)\right\| \left\|g(x)\right\| d\mu(x) \\ \leq \left\|f\right\|_{L_2} \left\|g\right\|_{L_2} < +\infty \end{aligned}$$

■

Combinando il lemma 12.3.5 con la proposizione 12.3.4 avremo che f_{n_j} converge ad f anche rispetto alla metrica standard L_p, ma poiché f_i è una successione di Cauchy tutta la successione f_i convergerà ad f in norma L_p. Ciononostante è bene tenere in mente che f_i potrebbe non convergere puntualmente quasi ovunque ad f pur convergendovi in norma L_p.

Per ogni sottoinsieme A di M indichiamo con $\mathcal{X}_A$ la funzione su M a valori in $\mathbb{R}$ che vale 1 su A e 0 altrimenti. Se indichiamo con $\mathscr{F}(\mu)$ lo spazio dei sottoinsiemi μ-misurabili di M con misura finita per ogni $A, B \in \mathscr{F}(\mu)$ possiamo porre

$$\rho_\mu(A,B) = \int_M \left|\mathcal{X}_A - \mathcal{X}_B\right| d\mu = \mu\left[A \triangle B\right]$$

se identifichiamo gli insiemi A e B per cui $\rho_\mu(A,B) = 0$ avremo che ρ_μ è una metrica su $\mathscr{F}(\mu)$. Osserviamo che se $A_i \in \mathscr{F}(\mu)$ è una successione di Cauchy rispetto a ρ_μ allora $\mathcal{X}_{A_i} \to f$ in norma L_1 ed esiste un'estratta che vi converge puntualmente quasi ovunque. Ma poiché le funzioni caratteristiche possono solamente assumere i valori 0 ed 1 anche f assumerà quasi ovunque i valori 0 ed 1, e poiché $f \in L_1$ esisterà $A \in \mathscr{F}(\mu)$ tale che $F = \mathcal{X}_A$ ovvero $A_i \to A$.

Generalizziamo ora il teorema 12.2.11 al caso di funzioni in L_p invece che in L_1.

Teorema 12.3.7 (Densità in L_p)**.** *Dati $p > 0$ ed una funzione $(\mathscr{M},\mu)$-fortemente misurabile $f : M \to X$ avremo che $f \in L_p(\mathscr{M},\mu,X)$ se e solo se esiste una successione di funzioni semplici $\mathscr{M}$-misurabili μ-integrabili s_n convergenti puntualmente quasi ovunque ad f ed $\int_M \|s_n - f\|^p \, d\mu \to 0$. In altre parole $\mathscr{S}(\mathscr{M},\mu,X)$ è denso in $L_p(\mathscr{M},\mu,X)$.*

La dimostrazione è la stessa del teorema 12.2.11 considerando però stavolta le funzioni $\phi_n(a) = \|t_n(a) - f(a)\|^p$ ed $\psi(a) = c_p 3^p \|f(a)\|^p$ e al contempo sfruttare la (12.3.1) o la (12.3.2) a seconda dei casi.

Teorema 12.3.8 (Separabilità in L_p)**.** *Data una σ-algebra $\mathscr{M} \subseteq \mu$ se X e $\left(\mathscr{M},\rho_\mu\right)$ sono spazi topologici separabili allora anche $L_p(\mathscr{M},\mu,X)$ è separabile per ogni $1 \le p < +\infty$.*

Dimostrazione. Prendiamo C denso e numerabile in X ed $\mathscr{F}$ denso e numerabile in $\mathscr{M}$, definiamo

$$\mathscr{S} = \left\{ s = \sum_{i=1}^{N} x_i \mathcal{X}_{A_i} \;\middle|\; N \in \mathbb{N}, x_i \in C, A_i \in \mathscr{F} \right\}$$

abbiamo che $\mathscr{S}$ è numerabile, per dimostrare che è denso in L_p grazie al teorema precedente ci basta solamente verificare che tutte le funzioni semplici $\mathscr{M}$-misurabili e μ-integrabili appartengono alla chiusura di $\mathscr{S}$.

Per ogni $\varepsilon > 0$ e per ogni $s : M \to X$ semplice $\mathscr{M}$-misurabile e μ-integrabile esistono $y_j \in X \setminus \{0\}$, $B_j \in \mathscr{M}$ tali che $s(a) = \sum_{j=1}^N y_j \mathcal{X}_{B_j}$. Prendiamo allora $x_j \in C$, $A_j \in \mathscr{F}$ in modo tale che $\|x_j - y_j\| \le \varepsilon$ ed $\rho_\mu\left(A_j, B_j\right) \le \varepsilon$ allora

$$\left\| s - \sum_{j=1}^{N} x_j \mathcal{X}_{A_j} \right\|_{L_p} \le \sum_{j=1}^{N} \left\| y_j \mathcal{X}_{B_j} - x_j \mathcal{X}_{A_j} \right\|_{L_p}$$

con

$$\begin{aligned}
\int_M \left\| y_j \mathcal{X}_{B_j} - x_j \mathcal{X}_{A_j} \right\|^p d\mu &= \|y_j\|^p \mu\left(B_j \setminus A_j\right) + \|x_j - y_j\|^p \mu\left(A_j \cap B_j\right) + \|x_j\|^p \mu\left(A_j \setminus B_j\right) \\
&\le \|y_j\|^p \mu\left(B_j \setminus A_j\right) + \|x_j - y_j\|^p \mu\left(B_j\right) + \left(\|y_j\| + \varepsilon\right)^p \mu\left(A_j \setminus B_j\right) \\
&\le M\varepsilon
\end{aligned}$$

dove M dipende solo da s e non da x_j o A_j. Il teorema è così dimostrato. ■

Una conseguenza molto utile di tale criterio è data dal seguente risultato:

Proposizione 12.3.9. *Se X è separabile, μ è una misura finita su M e la σ-algebra $\mathscr{M}$ è generata da una successione di insiemi μ-misurabili E_n allora $L_p(\mathscr{M},\mu,X)$ è separabile per ogni* $1 \leq p < +\infty$.

Dimostrazione. Sia $S \subseteq \mathscr{M}$ l'algebra generata dagli E_n. Chiaramente S è numerabile, vogliamo mostrare che è denso in $\mathscr{M}$. Per farlo basta dimostrare che la famiglia

$$\mathscr{S} = \left\{A \in \mathscr{M} \;\middle|\; \forall \varepsilon > 0 \exists A' \in S, \phi_\nu(A,A') < \varepsilon\right\}$$

è una σ-algebra e di conseguenza coincide con $\mathscr{M}$.

Chiaramente $\emptyset, M, E_n \in \mathscr{S}$ e $\rho_\mu(A,A') = \rho_\mu(M\setminus A, M\setminus A')$, inoltre presi $A,B \in \mathscr{S}$, $A',B' \in S$ avremo

$$\rho_\mu\left(A\cup B, A'\cup B'\right) \leq \rho_\mu\left(A,A'\right) + \rho_\mu\left(B,B'\right)$$

e quindi anche le unioni finite vi appartengono.

Consideriamo ora una successione di insiemi disgiunti $A_i \in \mathscr{S}$, dato che μ è finita esisterà $N \in \mathbb{N}$ tale che $\mu\left(\bigcup_{i=N}^{+\infty} A_i\right) < \varepsilon$, così possiamo comunque ricondurci al caso di unioni finite. Abbiamo così $\mathscr{S} = \mathscr{M}$.

Lo spazio $(\mathscr{M},\rho_\mu)$ è perciò separabile, di conseguenza $L_p(\mathscr{M},\mu,X)$ è separabile. ■

Per esempio possiamo considerare un qualunque spazio topologico $\mathscr{N}_2$ (M,τ) e definire su M la σ-algebra generata da tutti gli aperti di M. Tale σ-algebra $\mathscr{B}(M,\tau)$ è detta σ-algebra di Borel di M. Dato che esiste una base di aperti numerabile la proposizione precedente implica che gli spazi $L_p\left[\mathscr{B}(M,\tau),\mu\right]$ sono separabili per ogni $1 \leq p < +\infty$ e per ogni misura esterna μ tale che $\mathscr{B}(M,\tau) \subseteq \mathscr{M}(\mu)$.

Quando $p \in (0,1)$ la disuguaglianza di Hölder perde di validità in quanto la funzione t^p non è più convessa. In generale lo spazio L_p quando $p < 1$ non è localmente convesso e quindi la sua topologia non è indotta da alcuna norma. Più nel dettaglio

Teorema 12.3.10. *Se esistono un numero infinito di insiemi $A_i \subseteq M$ misurabili per cui*

$$0 < \mu\left(A_i\right) < +\infty$$
$$\mu\left(A_i \cap A_j\right) = 0 \text{ se } i \neq j$$

allora $L_p(\mu,X)$ non è localmente convesso quando $0 < p < 1$.

Dimostrazione. Per assurdo L_p possiede un sistema fondamentale di intorni convessi dell'origine, ovvero esistono $0 < r < s$ ed un insieme convesso C tale che

$$D\left(0,r,D_{L_p}\right) \subseteq C \subseteq B\left(0,s,D_{L_p}\right)$$

fissiamo anche $x \in X$ di norma unitaria, e per ogni A misurabile con $0 < \mu(A) < +\infty$ definiamo

$$\delta_A = \frac{r^{1/p}x}{\mu(A)}\chi_A$$

chiaramente $\delta_A \in D(0,r)$. Di conseguenza scelta una famiglia numerabile di insiemi A_n come sopra per cui $\mu\left(A_m \cap A_n\right) = 0$ se $m \neq n$ allora possiamo porre

$$\sigma_n = \frac{\delta_{A_1} + \delta_{A_2} + \cdots + \delta_{A_n}}{n} \in C \subseteq B(0,s)$$

ma

$$D_{L_p}(\delta_n) = \sum_{i=1}^{n} D_{L_p}\left(\frac{r^{1/p}x}{n\mu(A)}\mathcal{X}_A\right) = n^{1-p}r$$

e poiché $1 - p > 0$ la quantità di sopra tende a $+\infty$ al crescere di n, in contraddizione con l'ipotesi di limitatezza di C. ■

Se M è σ-finito ed $\mu(M) = +\infty$ allora soddisfa automaticamente le ipotesi del teorema. Ma anche lo spazio $M = [0, T]$ con la misura di Lebesgue soddisfa le ipotesi del teorema e dunque $L_p([0, T], X)$ non è localmente convesso. Facciamo vedere però che in questo spazio l'uniso sottoinsieme aperto convesso non vuoto è l'intero spazio.

Preso infatti un qualunque insieme aperto convesso C di $L_p([0, T], X)$ contenente l'origine esisterà $r > 0$ tale che $B\left(0, r, D_{L_p}\right) \subseteq C$, dunque per ogni $f \in L_p$ esisterà $n \in \mathbb{N}$ tale che $n^{(p-1)/p}f \in B(0, r)$ in quanto $p - 1 < 0$. Per continuità esistono $0 = x_0 < x_1 < \cdots < x_n = 1$ tali che

$$\int_{x_{i-1}}^{x_i} \|f\|^p \, dx = \frac{1}{n}\int_0^T \|f\|^p \, dx$$

Definiamo allora $g_i = n\mathcal{X}_{[x_{i-1}, x_i)}f$ per ogni $1 \leq i \leq n$, avremo innanzitutto che

$$D_{L_p}(g_i) = n^p \int_{x_i}^{x_{i-1}} \|f\|^p \, dx = n^{p-1}D_{L_p}(f) < r$$

e dunque $g_i \in C$, per convessità allora $f = (g_1 + \cdots + g_n)/n \in C$ e perciò $C = L_p$.

Corollario 12.3.11. *L'unica applicazione lineare continua da $L_p([0, T], X)$ in $\mathbb{K}$ quando $0 < p < 1$ è l'applicazione identicamente nulla.*

Lo spazio L_∞

Come per il caso $X = \mathbb{R}$ anche quando X è un generico spazio di Banach è possibile definire lo spazio L_∞. Data una funzione misurabile $f : M \to \mathbb{R}$ ricordiamo innanzitutto

$$\text{ess sup} f = \|f\|_{L_\infty(\mu)} = \inf\{\lambda \in \mathbb{R} \mid \mu(\{f > \lambda\}) = 0\}$$

e quindi definiamo lo spazio

$$L_\infty(\mu, X) = \left\{f : M \to X \text{ fortemente misurabile} \mid \text{ess sup}\|f\| < +\infty\right\}$$

è uno spazio normato con norma $\|f\|_{L_\infty}$. È possibile dimostrare che L_∞ è uno spazio di Banach come nel caso classico.

Come nel caso reale non è difficile dimostrare che

Proposizione 12.3.12. *Per ogni $f \in L_\infty(\mu, X)$, $g \in L_p(\mu, Y)$, $F \in L^2(X, Y; Z)$ si ha $F[f, g] \in L_p(\mu, Z)$ con $\|F[f, g]\|_{L_p} \leq \|F\| \|f\|_{L_\infty} \|g\|_{L_p}$.*

In generale $\mathscr{S}(\mu, X)$ non è denso in $L_\infty(\mu, X)$, per avere un risultato di densità dobbiamo considerare invece le funzioni a valori numerabili ovvero che possono essere scritte nella forma

$$s(x) = \sum_{i=1}^{+\infty} x_i\mathcal{X}_{A_i} \tag{12.3.4}$$

con A_i a due a due disgiunti. Infatti essendo ogni $g \in L_\infty$ fortemente misurabile ha valori separabili, quindi sarà sempre possibile trovare una successioni di funzioni nella forma (12.3.4) che converge in norma L_∞ a g.

Per questo motivo L_∞ non è quasi mai separabile, nemmeno se lo fossero anche X ed $\left(\mathscr{M}, \rho_\mu\right)$.

12.4 Risultati topologici

Gli integrali di Bochner possono essere utilizzati anche per dimostrare risultati topologici degli spazi di Banach. Alcuni di questi risultati sono

Teorema 12.4.1 (Kreĭn–Šmulian)**.** *Dato uno spazio di Banach X se K è debolmente compatto in X allora anche $\overline{\mathrm{con}}(K)$ è debolmente compatto.*

Dimostrazione. Per il teorema di Eberlein–Šmulian K è sequenzialmente debolmente compatto, sempre per lo stesso motivo ci basta dimostrare che $\overline{\mathrm{con}}(K)$ è debolmente sequenzialmente compatto.

Prendiamo una qualunque successione $x_n \in \overline{\mathrm{con}}(K)$, per ogni $n \in \mathbb{N}$ esistono $x_{ni} \in K$, $\lambda_{ni} \in [0,1]$ tali che $\lambda_{ni} = 0$ per ogni $i > N_n \in \mathbb{N}$, ${}_{i=1}^{+\infty}\lambda_{ni} = 1$ e $x_n = \sum_{i=1}^{+\infty} \lambda_{ni} x_{ni}$. Poniamo perciò

$$\begin{gathered} Y = \mathrm{Cl}\Big(\mathrm{span}\left\{x_{ni} \;\middle|\; n, i \in \mathbb{N}\right\}\Big) \\ K' = K \cap Y \end{gathered}$$

allora Y è uno spazio di Banach separabile, K' è debolmente sequenzialmente compatto rispetto alla topologia debole di Y e $x_n \in \overline{\mathrm{con}}(K')$, in particolare K' è separabile grazie al lemma 3.5.24. Questo ci permette di affermare senza perdere in generalità che K è separabile nella topologia forte, mentre per la proposizione 7.4.3 K è automaticamente limitato in norma.

Posto $g : x \in K \to x \in K$ chiaramente G ha valori separabili ed inoltre per ogni $f \in X^*$ abbiamo che fg è continua (K ha la topologia debole indotta), così il teorema di Pettis implica che g è μ-fortemente misurabile per ogni misura esterna μ borelliana rispetto alla topologia di K, e quindi $g \in L_\infty(\mu, X)$. Sia $C(K)$ lo spazio delle funzioni continue da K in $\mathbb{R}$, poiché K è uno spazio topologico compatto e di Hausdorff il teorema di rappresentazione di Riesz implica che per ogni $\Lambda \in C(K)^*$ esiste un'unica coppia di misure di Radon esterne $\mu = (\mu^+, \mu^-)$ su K tali che

- esiste $U \subseteq K$ borelliano tale che $\mu^+(K \setminus U) = 0 = \mu^-(U)$;
- per ogni $f : K \to \mathbb{R}$ continua si ha
$$\Lambda(f) = \int_K f\, d\mu^+ - \int_K f\, d\mu^- = \int_K f\, d\mu$$
- $\|\Lambda\| = \mu^+(K) + \mu^-(K)$.

Possiamo allora definire l'applicazione lineare e continua

$$T : \mu \in C(K)^* \to \int_K g\, d\mu \in X$$

con $\|T(f)\| \leq \|g\|_{L_\infty} \|f\|$, verifichiamo che è continua anche quando X ha la topologia debole e $C(K)^*$ quella debole*, in altre parole $f \circ T$ è debolmente* continua per ogni $f \in X^*$. Per le proprietà degli integrali abbiamo per ogni $\mu \in C(K)^*$

$$f[T(\mu)] = f\left(\int_K g\, d\mu\right) = \int_K f \circ g\, d\mu = \mu(f \circ g) = J_{C(K)}(f \circ g)(\mu)$$

in quanto $f \circ g \in C(K)$. Dato che $f \circ T = J_{C(K)}(f \circ g)$ è ovviamente debolmente* continua per l'arbitrarietà di $f \in X^*$ segue la tesi.

Il teorema di Banach-Alaoglu afferma che la palla chiusa $U = D\big(0, 1, C(K)^*\big)$ è debolmente* compatto, per continuità avremo $T(U)$ relativamente debolmente compatto e convesso in X, dunque la sua chiusura forte coincide con la chiusura debole. Per ogni $x \in K$ la misura esterna δ_x definita su K come

$$\delta_x(A) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

appartiene ad U e $T(\delta_x) = x$, abbiamo così $K \subseteq T(U)$ e perciò $\overline{\mathrm{con}}(K) \subseteq \mathrm{Cl}(T(U))$. Dato che $\overline{\mathrm{con}}(K)$ è convesso e chiuso è anche debolmente chiuso e contenuto in un insieme debolmente compatto, di conseguenza è debolmente compatto e la successione x_n ammette un'estratta debolmente convergente. ■

Teorema 12.4.2 (Mazur)**.** *Dato uno spazio di Banach X se K è compatto in X allora anche $\overline{\mathrm{con}}(K)$ è compatto.*

Dimostrazione. La dimostrazione è analoga al risultato precedente ma con alcune piccole modifiche. Chiaramente ora K è automaticamente separabile in X e lo dotiamo della topologia forte indotta. Inoltre K è anche totalmente limitato e perciò per ogni $i \in \mathbb{N}$ esistono $x_{i1}, x_{i2}, \ldots, x_{in_i}$ tali che

$$K \subseteq \bigcup_{j=1}^{n_i} B\left(x_{ij}, \frac{1}{i}\right)$$

Posto

$$A_{ij} = K \cap B\left(x_{ij}, \frac{1}{i}\right) \setminus \bigcup_{k=1}^{j-1} B\left(x_{ik}, \frac{1}{i}\right)$$

$$g_i = \sum_{j=1}^{n_i} x_{ij} \chi_{A_{ij}}$$

abbiamo che A_{ij} è un insieme di Borel ed g_i converge a g uniformemente su K. Dato che g_i è semplice l'applicazione

$$T_i : \mu \in C(K)^* \to \int_K g_i\, d\mu = \sum_{j=1}^{n_i} x_{ij} \mu\big(A_{ij}\big) \in X$$

ha valori in un sottospazio finitamente generato e perciò è un operatore compatto, dato che X è uno spazio di Banach il limite di operatori compatti è ancora un operatore compatto ed essendo

$$\|T_i - T\| \leq \|g_i - g\|_{L_\infty} \to 0$$

abbiamo che T è compatto. Lavorando ora come il precedente teorema si ha la tesi. ■

12.5 La proprietà di Radon-Nikodym

Dati uno spazio di Banach X, un insieme M non vuoto, una σ-algebra $\mathcal{M}$ su di esso ed un'applicazione $\nu : \mathcal{M} \to X$ diremo che la coppia $(\mathcal{M}, \nu)$ è una **misura vettoriale** se e solo se soddisfa le seguenti proprietà:

1. $\nu(\emptyset) = 0$;

2. presi $A_n \in \mathcal{M}$ a due a due disgiunti allora $\nu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \nu\left(A_n\right)$.

Quando $X = \mathbb{R}$ e $\mu(A) \geq 0$ per ogni $A \in \mathcal{M}$ diremo più semplicemente che μ è una **misura positiva**. Da notare che le misure positive non sono misure esterne in quanto sono sempre finite e sono determinate anche dalla σ-algebra e non solo dai valori che assume su di essa.

Date una misura vettoriale $(\mathcal{M}, \nu)$ ed una misura esterna μ entrambe su M diremo che ν è **μ-assolutamente continua** se e solo se $\mathcal{M} \subseteq \mathcal{M}(\mu)$ e per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $A \in \mathcal{M}$ tale che $\mu(A) < \delta$ si ha $\|\nu(A)\| < \varepsilon$.

Come per le misure esterne vale il seguente risultato di continuità:

Proposizione 12.5.1. *Data una successione crescente/decrescente di insiemi $E_i \in \mathcal{M}$ e sia $E \in \mathcal{M}$ la loro unione/intersezione. Per ogni misura vettoriale ν su $\mathcal{M}$ si ha*

$$\nu(E) = \lim_{n \to +\infty} \nu(E_i)$$

Osserviamo ora che la condizione di μ-assoluta continuità per una misura vettoriale $(\mathcal{M}, \nu)$ è equivalente ad affermare che $\nu(A) = 0$ per ogni $A \in \mathcal{M}$ tale che $\mu(A) = 0$. Se infatti esistesse $\varepsilon_0 > 0$ tale che per ogni $n \in \mathbb{N}$ esiste $A_n \in \mathcal{M}$ per cui $\mu(A_n) < 1/2^n$ e $\left\|\nu(A_n)\right\| \geq \varepsilon_0$ allora posto

$$B_n = \bigcup_{m=n}^{+\infty} A_m$$
$$B = \bigcap_{n \in \mathbb{N}} B_n$$

avremo che B_n è una successione decrescente di insiemi μ-misurabili con

$$\mu(B_n) \leq \sum_{m=n}^{+\infty} \frac{1}{2^m} = 2^{1-n}$$

e quindi $\mu(B) = 0$.

D'altronde abbiamo

$$\|\nu(B)\| = \lim_{n \to +\infty} \left\|\nu(B_n)\right\| \geq \varepsilon_0$$

ovvero $\nu(B) \neq 0$.

Data una misura vettoriale $(\mathcal{M}, \nu)$ definiamo per ogni $E \in \mathcal{M}$ la quantità

$$\|\nu\|_{\mathcal{M}}(E) = \sup\left\{ \sum_{i=1}^{n} \left\|\nu(E_i)\right\| \;\middle|\; n \in \mathbb{N}, E_i \in \mathcal{M}, E_i \cap E_j = \emptyset \text{ se } i \neq j, E = \bigcup_{i=1}^{n} E_i \right\}$$

chiaramente abbiamo $\|\nu(E)\| \leq \|\nu\|_{\mathcal{M}}(E)$, inoltre scelti $E \subseteq F$ in $\mathcal{M}$ ed $E_1, E_2, \ldots, E_n$ una partizione finita di E in $\mathcal{M}$ essa sarà anche una partizione di F se vi aggiungiamo $E_0 = F \setminus E$ e dunque $\|\nu\|_{\mathcal{M}}(E) \leq \|\nu\|_{\mathcal{M}}(F)$.

Perciò possiamo porre stavolta per ogni $A \subseteq M$

$$\|\nu\|_{\mathscr{M}}(A) = \inf\left\{\|\nu\|_{\mathscr{M}}(E) \;\middle|\; E \in \mathscr{M}, A \subseteq E\right\}$$

vogliamo ora dimostrare che $\|\nu\|_{\mathscr{M}}$ è una misura esterna. Presi $A, A_i \subseteq M$ tali che $A = \bigcup_{i=1}^{+\infty} A_i$ supponiamo inizialmente che $A, A_i \in \mathscr{M}$ e che gli A_i siano a due a due disgiunti. Per ogni $\varepsilon > 0$ esiste una partizione finita $E_1, \dots, E_n$ di A tale che $\sum_{j=1}^{n} \left\|\nu(E_j)\right\| > \|\nu\|_{\mathscr{M}}(A) - \varepsilon$.

Posto $E_{ij} = E_j \cap A_i$ avremo che E_{ij} sono sia una partizione di A_i che una di E_j a seconda dell'indice che lasciamo invariato. Per la continuità delle misure vettoriali abbiamo così

$$\|\nu\|_{\mathscr{M}}(A) < \varepsilon + \sum_{j=1}^{n} \left\|\nu(E_j)\right\| = \varepsilon + \sum_{j=1}^{n} \left\|\sum_{i=1}^{+\infty} \nu(E_{ij})\right\| \leq \varepsilon + \sum_{i=1}^{+\infty}\sum_{j=1}^{n} \left\|\nu(E_{ij})\right\| \leq \varepsilon + \sum_{i=1}^{+\infty} \|\nu\|_{\mathscr{M}}(A_i)$$

e la tesi segue facendo tendere ε a 0.

Se adesso gli A_i non fossero a due a due disgiunti definiamo

$$C_1 = A_1 \qquad\qquad C_{i+1} = A_{i+1} \setminus \bigcup_{j=1}^{i} A_j$$

così i C_i sono disgiunti, $C_i \subseteq A_i$ e $A = \bigcup_{i=1}^{+\infty} C_i$, per monotonia si ha

$$\|\nu\|_{\mathscr{M}}(A) \leq \sum_{i=1}^{+\infty} \|\nu\|_{\mathscr{M}}(C_i) \leq \sum_{i=1}^{+\infty} \|\nu\|_{\mathscr{M}}(A_i)$$

Infine se abbiamo solamente $A, A_i \subseteq M$ e $A \subseteq \bigcup_{i=1}^{+\infty} A_i$, per ogni $\varepsilon > 0$ esisterà $E_i \in \mathscr{M}$ per cui $A_i \subseteq E_i$ e $\|\nu\|_{\mathscr{M}}(A_i) > \|\nu\|_{\mathscr{M}}(E_i) - \varepsilon 2^{-i}$. Posto anche $E = \bigcup_{i=1}^{+\infty} E_i \in \mathscr{M}$ abbiamo $A \subseteq E$ e dunque

$$\|\nu\|(A) \leq \|\nu\|_{\mathscr{M}}(E) \leq \sum_{i=1}^{+\infty} \|\nu\|_{\mathscr{M}}(E_i) < \varepsilon + \sum_{i=1}^{+\infty} \|\nu\|_{\mathscr{M}}(A_i)$$

da cui discende la tesi. La misura esterna $\|\nu\|_{\mathscr{M}}$ sarà detta **variazione totale** della misura vettoriale $(\mathscr{M}, \nu)$. Si noti che la variazione totale di ν dipende anche dalla σ-algebra scelta, in particolare presa un'altra σ-algebra $\mathscr{M} \subseteq \mathscr{N}$ possiamo avere $\|\nu\|_{\mathscr{N}} \neq \|\nu\|_{\mathscr{M}}$ anche su $\mathscr{N}$! Quando non vi è ambiguità ometteremo il simbolo $\mathscr{M}$ dall'espressione della variazione totale.

Proposizione 12.5.2. *Per ogni misura vettoriale $(\mathscr{M}, \nu)$ si ha $\mathscr{M} \subseteq \mathscr{M}(\|\nu\|)$, in particolare ν sarà sempre $\|\nu\|$-assolutamente continua.*

Dimostrazione. Presi ora $A \in \mathscr{M}$ ed $E \subseteq M$ dimostreremo che quando $\|\nu\|(E) < +\infty$

$$\|\nu\|(E) \geq \|\nu\|(E \cap A) + \|\nu\|(E \setminus A)$$

Per ogni $\varepsilon > 0$ esiste $F \in \mathscr{M}$ tale che $E \subseteq F$ e $\|\nu\|(F) \leq \|\nu\|(E) + \varepsilon$, d'altronde poiché $F \cap A, F \setminus A \in \mathscr{M}$ esistono $A_1, A_2, \dots, A_n, B_1, \dots, B_m \in \mathscr{M}$ a due a due disgiunti tali che $F \cap A = \bigcup_i A_i$, $F \setminus A = \bigcup_j B_j$ ed

$$\|\nu\|(F \cap A) \leq \sum_{i=1}^{n} \left\|\nu(A_i)\right\| + \frac{\varepsilon}{2} \qquad\qquad \|\nu\|(F \setminus A) \leq \sum_{j=1}^{m} \left\|\nu(B_j)\right\| + \frac{\varepsilon}{2}$$

ma allora essi insieme formano una partizione di F e dunque

$$\begin{aligned}\|\nu\|(E) + \varepsilon \geq \|\nu\|(F) \geq \sum_{i=1}^{n} \left\|\nu(A_i)\right\| + \sum_{j=1}^{m} \left\|\nu(B_j)\right\| \geq \|\nu\|(F \cap A) + \|\nu\|(F \setminus A) - \varepsilon \\ \geq \|\nu\|(E \cap A) + \|\nu\|(E \setminus A) - \varepsilon\end{aligned}$$

per l'arbitrarietà di ε segue così la tesi. ■

Definizione 12.5.3. Diremo che la misura vettoriale ν ha **variazione limitata** se e solo se $\|\nu\|(M) < +\infty$.

Data adesso una misura esterna generica μ, $f \in L_1(\mu, X)$ ed $\mathcal{M} \subseteq \mu$ possiamo porre per ogni $A \in \mathcal{M}$

$$\nu(E) = \int_E f\, d\mu$$

si noti che f può non essere $\mathcal{M}$-misurabile. Osserviamo che $(\mathcal{M}, \nu)$ è una misura vettoriale, infatti se $E_n \in \mathcal{M}$ è una successione di insiemi misurabili disgiunti con $E = \bigcup_{i=1}^{+\infty} E_i$ allora è possibile applicare il teorema della convergenza dominata per dimostrare che

$$\int_E f\, d\mu = \sum_{i=1}^{+\infty} \int_{E_i} f\, d\mu$$

Osserviamo anche che ν ha variazione limitata in quanto per ogni $E_1, E_2, \dots, E_n \in \mathcal{M}$ disgiunti si ha

$$\sum_{i=1}^{n} \|\nu(E_i)\| = \sum_{i=1}^{n} \left\| \int_{E_i} f\, d\mu \right\| \le \sum_{i=1}^{n} \int_{E_i} \|f\|\, d\mu = \int_{\bigcup_{i=1}^{n}} \|f\|\, d\mu$$

da cui segue che

$$\|\nu\|(E) \le \int_E \|f\|\, d\mu < +\infty$$

e dunque ha variazione limitata.

Se però f fosse anche $(\mathcal{M}, \mu)$-misurabile, e dunque $f \in L_1(\mathcal{M}, \mu, X)$ allora la disuguaglianza precedente è in realtà un uguaglianza. Infatti essendo f Bocher-integrabile per ogni $\varepsilon > 0$ esiste una funzione semplice s $\mathcal{M}$-misurabile μ-integrabile nella forma

$$s = \sum_{i=0}^{I} x_i \mathcal{X}_{E_i}$$

con gli E_i disgiunti tale che $\int_M \|f - s\|\, d\mu < \varepsilon$. Esisterà però anche un'altra partizione F_j di M composta da elementi di $\mathcal{M}$ per cui

$$\|\nu\|(M) < \sum_{j=1}^{J} \|\nu(F_j)\| + \varepsilon$$

d'altronde $G_{ij} = E_i \cap F_j$ è ancora una partizione di M, perciò

$$0 \le \|\nu\|(M) - \sum_{ij} \|\nu(G_{ij})\| < \varepsilon$$

Mettendo tutto assieme

$$\begin{aligned}\left| \|\nu\|(M) - \int_M \|s\|\, d\mu \right| &= \left| \|\nu\|(M) - \sum_{i,j} \|x_i\|\, \mu\left(E_i \cap F_j\right) \right| \\ &\le \left| \sum_{i,j} \left\| \int_{G_{ij}} f\, d\mu \right\| - \sum_{i,j} \left\| \int_{G_{ij}} s\, d\mu \right\| \right| + \varepsilon \le 2\varepsilon\end{aligned}$$

e quindi $\left| \|\nu\|(M) - \int_M \|f\|\, d\mu \right| \le 3\varepsilon$ e perciò coincidono.

Diremo che uno spazio di Banach X ha la **proprietà di Radon-Nikodym** se e solo se per ogni spazio di misura (M,μ) con μ finita e per ogni misura vettoriale ν a valori in X μ-continua e a variazione limitata esiste $f \in L_1(M,X)$ tale che

$$\nu(E) = \int_M f\, d\mu$$

Bisogna precisare che non tutti gli spazi di Banach hanno questa proprietà, anzi esistono spazi di Banach non troppo complessi che non la soddisfano come per esempio lo spazio $L_1([0,1])$ rispetto alla misura di Lebesgue.

Prima di andare avanti dimostriamo la versione classica del teorema di Radon-Nikodym:

Teorema 12.5.4 (Radon-Nikodym, forma classica)**.** *Date due misure esterne finite μ,η su M ed una σ-algebra $\mathscr{F}$ tale che*

$$\mathscr{F} \subseteq \mathscr{M}(\mu) \cap \mathscr{M}(\eta)$$

esistono una funzione $f \in L_1(\mathscr{F},\mu)$ per cui $f \geq 0$ ed un insieme $Z \in \mathscr{F}$ con $\mu(B) = 0$ tali che per ogni $A \in \mathscr{F}$

$$\eta(A) = \int_A f\, d\mu + \eta(A \cap Z)$$

Dimostrazione (von Neumann). Poniamo $\xi = \mu + \eta$ che è ancora una misura esterna finita, consideriamo l'applicazione lineare

$$T : f \in L_2(\mathscr{F},\xi) \to \int_M f\, d\eta \in \mathbb{R}$$

che è ben definito, poiché ogni funzione a valori in $\mathbb{R}$ può essere vista come la differenza di due funzioni non negative per ogni $f \geq 0$ si avrà

$$\int_M f\, d\eta \leq \int_M f\, d\xi \leq \sqrt{\xi(M)}\, \|f\|_{L_2(\xi)}$$

quindi T è anche continua. Dato che $L_2(\mathscr{F},\xi)$ è uno spazio di Hilbert esisterà $h \in L_2(\mathscr{F},\xi)$ tale che

$$\int_M f\, d\eta = \int_M f\, h\, d\xi = \int_M f\, h\, d\mu + \int_M f\, h\, d\eta \Rightarrow \int_M f\, h\, d\mu = \int_M f(1-h)\, d\eta$$

Poiché μ,η sono finite le funzioni limitate sono chiaramente L_2, quindi possiamo sostituire f con funzioni caratteristiche. Abbiamo allora

$$0 \geq \int_{\{h<0\}} h\, d\mu = \int_{\{h<0\}} (1-h)\, d\eta \geq 0$$

che implica $\mu(\{h<0\}) = \eta(\{h<0\}) = 0$. Invece posto $Z = \{h \geq 1\}$

$$\eta(Z) = \int_Z h\, d\xi \geq \mu(Z) + \eta(Z)$$

e perciò $\mu(Z) = 0$.

Infine per ogni $n \in \mathbb{N}$ poniamo

$$B_n = \{0 \leq h \leq 1 - 1/n\} \in \mathscr{F}$$

allora

$$\eta(B_n) = \int_{B_n} \frac{1-h}{1-h}\, d\eta = \int_{B_n} \frac{h}{1-h}\, d\mu$$

posto $f(x) = \lim_{n\to+\infty} \mathcal{X}_{B_n} h/(1-h)$ per il teorema di Beppo-Levi della convergenza monotona si avrà

$$\begin{aligned}\eta(A) = \eta(A\cap\{0\leq h<1\}) + \eta(A\cap Z) &= \lim_{n\to+\infty} \eta(A\cap\{0\leq h\leq 1-1/n\}) + \eta(A\cap Z)\\ &= \int_A f\, d\mu + \eta(A\cap Z)\end{aligned}$$

dato che η è una misura finita avremo che $f\in L_1(\mathscr{F},\mu)$. ■

Corollario 12.5.5. *Gli spazi vettoriali finitamente generati soddisfano la proprietà di Radon-Nikodym.*

Dimostrazione. Supponiamo innanzitutto $X=\mathbb{R}$. Sia ν una misura reale μ-continua e a variazione limitata, possiamo allora definire le seguenti misure finite positive

$$\nu^+ = \frac{|\nu|+\nu}{2} \qquad\qquad \nu^- = \frac{|\nu|-\nu}{2}$$

chiaramente $\nu=\nu^+-\nu^-$ e $|\nu|=\nu^++\nu^-$, inoltre dato che ν è μ-continua anche ν^+ e ν^- saranno μ-continue. Possiamo così applicare la forma classica del teorema di Radon-Nikodym a ν^+ e ν^-. Il teorema è così dimostrato in quanto la μ continuità implica $\nu^+(Z_1)=\nu^-(Z_2)=0$.

Se ora X è finitamente generato possiamo sempre trovare una base di $X_{\mathbb{R}}$ e scomporre ν nelle sue componenti reali. ■

Un'altra conseguenza della forma classica del teorema di Radon-Nikodym che ci servirà in seguito è il seguente teorema:

Teorema 12.5.6 (Hellinger)**.** *Date due misure esterne μ,ν su M con $\mu(M)=\nu(N)=1$ per cui esistono un'altra misura esterna finita λ ed $f,g\in L_1(\lambda)$ tali che $\mu=f\lambda$ ed $\nu=g\lambda$. Allora la quantità*

$$\mathscr{H}(\mu,\nu) = \int_M \sqrt{fg}\, d\lambda$$

non dipende da λ,f,g e inoltre

$$2\left[1-\mathscr{H}(\mu,\nu)\right] \leq \left\|\mu-\nu\right\|(M) \leq 2\sqrt{1-\mathscr{H}(\mu,\nu)^2}$$

Dimostrazione. Possiamo supporre inizialmente che $\lambda=\mu+\nu$ così da ottenere $\left\|\mu-\nu\right\|(X) = \left\|f-g\right\|_{L_1(\lambda)}$. Dunque le disuguaglianze diventano semplicemente

$$\begin{aligned}2\left(1-\int_M\sqrt{fg}\,d\lambda\right) &= \int_M\left(\sqrt{f}-\sqrt{g}\right)^2 d\lambda \leq \int_M |f-g|\, d\lambda\\ &= \int_M\left|\sqrt{f}-\sqrt{g}\right|\left|\sqrt{f}+\sqrt{g}\right| d\lambda \leq \sqrt{\int_M\left(\sqrt{f}-\sqrt{g}\right)^2 d\lambda}\sqrt{\int_M\left(\sqrt{f}+\sqrt{g}\right)^2 d\lambda}\\ &= 2\sqrt{1-\left(\int_M\sqrt{fg}\,d\lambda\right)^2}\end{aligned}$$

Sia ora η una misura esterna finita generica rispetto alla quale sia μ che ν sono assolutamente continue. Ma allora anche λ è assolutamente continua rispetto a η e $\lambda = \psi\eta$. Dunque $\mu = f\psi\eta$ e $\nu = g\psi\eta$ e così

$$\int_M \sqrt{fg}\,d\lambda = \int_M \sqrt{fg}\psi\,d\eta = \int_M \sqrt{(f\psi)(g\psi)}\,d\eta$$

e quindi $\mathscr{H}(\mu,\nu)$ è ben definito. ■

Corollario 12.5.7. *Per ogni coppia di misure positive μ,ν si ha $\mathscr{H}(\mu,\nu) = 0$ se e solo se esiste $K \in \mathscr{M}(\mu)\cap\mathscr{M}(\nu)$ tale che $\mu(B) = \mu(B\cap K)$ e $\nu(B) = \nu(B\setminus K)$ per ogni B.*

Dimostrazione. Chiaramente $\mathscr{H}(\mu,\nu) = 0$ se e solo se $\left\|\mu-\nu\right\|(M) = 2$. Possiamo supporre che le funzioni di densità f e g ottenute nel teorema di Hellinger siano misurabili rispetto ad una σ-algebra comune sia a μ che a ν, cosicché gli insiemi

$$P = \{fg \neq 0\} \qquad Z_1 = \{f = 0, g > 0\}$$
$$Z_2 = \{f > 0, g = 0\} \qquad Z = \{f = g = 0\}$$

siano sia μ che ν misurabili. Abbiamo così per ogni $E \subseteq M$

$$\mu(E) = \mu(E\cap P) + \mu(E\cap Z_2)$$
$$\nu(E) = \nu(E\cap P) + \mu(E\cap Z_1)$$

chiaramente $Z_1 \cap Z_2 = \emptyset$. Ora $\mathscr{H}(\mu,\nu) = 0$ se e solo se $\mu(P) = \nu(P) = 0$ e quindi si ha la tesi. ■

Possiamo ora dimostrare il teorema più importante della sezione:

Teorema 12.5.8 (Dunford-Pettis)**.** *Dato uno spazio normato X, se X^* è separabile allora X^* soddisfa la proprietà di Radon-Nikodym.*

Dimostrazione. Supponiamo senza perdere in generalità che $\mathbb{K} = \mathbb{R}$. Sia ν una misura vettoriale a valori in X^* μ-continua e con variazione limitata, per il teorema di Radon-Nikodym classico si ha

$$\|\nu\|(A) = \int_A v\,d\mu$$

con $v \in L_1(\mu)$ e $v \geq 0$.

Ancora per ogni $x \in X$ poniamo

$$\nu_x(A) = \langle \nu(A) : x\rangle \in \mathbb{K}$$

chiaramente ν_x è una misura a valori scalari ed è $\|\nu\|$-continua, per il teorema di Radon-Nikodym esisterà $g_x \in L_1(\|\nu\|,\mathbb{K})$ tale che

$$\nu_x(A) = \int_A g_x\,d\|\nu\|$$

Ora osserviamo che per ogni $A \subseteq M$ misurabile

$$\left\|\nu_x(A)\right\| \leq \|\nu(A)\|\,\|x\| \leq \|\nu\|(A)\,\|x\|$$

e dunque $\left\|\nu_x\right\|(A) \leq \|x\| \|\nu\|(A)$, per il corollario 12.2.8 esiste $N_x \subseteq M$ tale che $\|\nu\|(N_x) = 0$ ed $\left|g_x(z)\right| \leq \|x\|$ per ogni $z \in M \setminus N_x$ ed $g_x \in L_\infty$.

Ora essendo X^* separabile per il lemma 7.2.5 anche X è separabile e possiamo trovare un sottoinsieme denso e numerabile $\{x_1, x_2, \dots\}$. Posto

$$D = \left\{ \sum_{i=1}^{N} q_i x_i \;\middle|\; N \in \mathbb{N}, q_i \in \mathbb{Q} \right\}$$

inoltre per ogni A misurabile e per ogni $x \in D$

$$\int_A g_x \, d\|\nu\| = \nu_x(A) = \langle \nu(A) : x \rangle = \sum_{i=1}^{N} q_i \langle \nu(A) : x_i \rangle = \sum_{i=1}^{N} \int_A q_i g_{x_i} \, d\|\nu\|$$

e così esiste N_x' tale che $\|\nu\|(N_x') = 0$ e per ogni $z \in M \setminus N_x'$

$$g_{\sum_i q_i x_i}(z) = \sum_i q_i g_{x_i}(z)$$

Posto $N = \bigcup_{x \in D} N_x \cup N_x'$ avremo che $\|\nu\|(N) = 0$ e per ogni $z \in M \setminus N$, $x \in D$

$$\left| \sum_{i=1}^{N} q_i g_{x_i}(z) \right| \leq \left\| \sum_{i=1}^{N} q_i x_i \right\|$$

per continuità possiamo assumere i q_i reali e così la disuguaglianza è valida per ogni $x \in \operatorname{span} D$. Da Hahn-Banach esiste un'applicazione $g : M \setminus N \to X^*$ tale che $g(z)(y) = g_y(z)$ per ogni $y \in D$ e $\left\|g(z)\right\| \leq 1$. Possiamo estendere g su N semplicemente ponendo $g(z) = 0$ e ricordiamo che per ogni $y \in D$

$$\nu(A)(y) = \int_A g(z)(y) \, d\|\nu\|(z) \tag{12.5.1}$$

Fissiamo adesso $x \in X$ ed $y_i \in D$ convergente ad x, per continuità abbiamo $g(z)(y_n) \to g(z)(x)$ per ogni $z \in M$ ed

$$\left|g(z)(y_i)\right| \leq \sup_i \left\|y_i\right\| \left\|g(z)\right\| \leq \sup_i \left\|y_i\right\| \;\; \forall z$$

per il teorema della convergenza dominata l'uguaglianza (12.5.1) vale anche per ogni $y \in X$ ed inoltre l'applicazione $z \to g(z)(x)$ è $\|\nu\|$-misurabile per ogni $x \in X$.

Poiché X^* è separabile ed X è totalizzante su X^* modificando un po' la dimostrazione del teorema di Pettis avremo che g è fortemente misurabile, di conseguenza $g \in L_\infty(M, X^*) \subseteq L_1(M, X^*)$ e quindi è anche Bochner-integrabile. Per l'unicità dell'integrale di Bochner abbiamo così

$$\nu(A) = \int_A g \, d\|\nu\| = \int_A g v \, d\mu$$

e il teorema è così dimostrato. ■

Vogliamo però ottenere delle condizioni più deboli del teorema, in modo da poter ammettere una classe più ambia di spazi ammissibili. Per ottenere un tale risultato dobbiamo innanzitutto lavorare sui funzionali lineari e continui $T : L_1(\mu) \to X$ nella forma

$$T(f) = \int_M f g \, d\mu \tag{12.5.2}$$

con μ misura esterna finita ed $g \in L_\infty(\mu, X)$.

Dalla disuguaglianza di Hölder si verifica immediatamente che $\|T(f)\| \le \|g\|_{L_\infty} \|f\|_{L_1}$ e così $\|T\| \le \|g\|_{L_\infty}$. Però essendo μ finita abbiamo anche $g \in L_1(\mu, X)$ e dunque $\nu(E) = \int_E g\, d\mu = T(\mathcal{X}_E)$ è una misura vettoriale a variazione limitata con

$$\|\nu\|(E) = \int_E \|g\|\, d\mu$$

Per ogni partizione finita E_i di E misurabile avremo anche

$$\sum_{i=1}^n \|\nu(E_i)\| = \sum_{i=1}^n \|T(\mathcal{X}_{E_i})\| \le \sum_{i=1}^n \|T\|\,\mu(E_i) = \|T\|\,\mu(E)$$

e così $\frac{1}{\mu(E)} \int_E \|g\|\, d\mu \le \|T\|$ per ogni E misurabile con $\mu(E) > 0$, il corollario 12.2.8 implica così che $\|g\|_{L_\infty} \le \|T\|$ e otteniamo così l'uguaglianza.

Teorema 12.5.9. *Uno spazio di Banach X soddisfa la proprietà di Radon-Nikodym se e solo se per ogni applicazione lineare e continua T da $L_1(\mu)$ ad X esiste $g \in L_\infty(M, X)$ tale che T è nella forma* (12.5.2).

Dimostrazione. Se X ha la proprietà di Radon-Nikodym allora l'applicazione $\nu(E) = T(\mathcal{X}_E)$ è una misura vettoriale con $\|\nu\|(M) \le \|T\|\,\mu(M)$, esisterà dunque $g \in L_1(M, X)$ tale che

$$T(\mathcal{X}_A) = \int_A g\, d\mu = \int_M \mathcal{X}_A g\, d\mu$$

Lavorando come nel ragionamento precedente abbiamo che $g \in L_\infty(\mu, X)$ con $\|g\|_{L_\infty} = \|T\|$. Poiché le funzioni semplici sono dense in $L_1(\mu)$ per la linearità e continuità di T avremo la (12.5.2) per ogni $f \in L_1(\mu)$.

Se invece ogni T è nella forma (12.5.2) prendiamo una qualunque misura vettoriale ν μ-continua e a variazione limitata, sappiamo che esiste $h \in L_1(\mu)$ tale che $\|\nu\|(A) = \int_A h\, d\mu$. Posto

$$E_n = \{n - 1 \le h \le n\}$$

per ogni funzione semplice $s = \sum_i x_i \mathcal{X}_{A_i}$ a valori in $\mathbb{K}$ possiamo definire per ogni $n \in \mathbb{N}$

$$T_n(s) = \sum_i x_i \nu(A_i \cap E_n) \in X$$

chiaramente T_n è lineare e inoltre

$$\|T_n(s)\| \le \sum_i |x_i|\, \|\nu\|(A_i \cap E_n) = \sum_i \int_{E_n} |x_i|\, \mathcal{X}_{A_i} h\, d\mu \le n \int_{A_i} |x_i|\, d\mu = n\,\|s\|_{L_1}$$

Per Hahn-Banach possiamo estendere T_n su tutto $L_1(\mu)$, così esiste $g_n \in L_\infty(\mu, X)$ tale che

$$T_n(f) = \int_M f g_n\, d\mu$$

in particolare $\nu(A \cap E_n) = \int_A g_n\, d\mu$.

Essendo gli E_n disgiunti possiamo porre $g = \sum_{n=1}^{+\infty} g_n \mathcal{X}_{E_n}$, inoltre

$$\lim_{n \to +\infty} \int_{A \cap \bigcup_{i=1}^n} g\, d\mu = \nu(A)$$

per ogni n abbiamo così

$$\int_{\bigcup_{i=1}^n E_i} \|g\| \, d\mu = \|\nu\| \left(\bigcup_{i=1}^n E_i \right) \leq \|\nu\| (M) < +\infty$$

e perciò $g \in L_1$, in particolare X soddisfa la proprietà di Radon-Nikodym. ■

Il prossimo passo è dimostrare che tutti gli spazi riflessivi possiedono la proprietà di Radon-Nikodym. Innanzitutto un insieme $F \subseteq L_1(\mu)$ è detto uniformemente integrabile se e solo se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $E \in \mathcal{M}(\mu)$ per cui $\mu(E) < \delta$ e per ogni $f \in F$ si ha $\int_E |f| \, d\mu < \varepsilon$.

Questa definizione ci serve per introdurre il seguente lemma tecnico:

Lemma 12.5.10. *Dati $g \in L_\infty(\mu, X)$ ed $F \subseteq L_1(\mu)$ limitato ed uniformemente integrabile posto*

$$T_g : f \in L_1(\mu) \to \int_M f g \, d\mu \in X$$

avremo che $T_g(F)$ è totalmente limitato in X.

Dimostrazione. Prendiamo una qualunque successione di funzioni semplici g_i convergenti puntualmente quasi ovunque a g. Per il teorema di Egoroff e per l'uniforme integrabilità di F per ogni $\varepsilon > 0$ esiste E misurabile tale che g_i converge uniformemente a g su E ed inoltre per ogni $f \in K$

$$\int_{M \setminus E} |f| \, d\mu < \frac{\varepsilon}{\|T\| + 1}$$

Inoltre la convergenza uniforme di funzioni semplici implica che $g\mathcal{X}_E$ ha valori in un insieme totalmente limitato e perciò $T_{g\mathcal{X}_E}$ è un operatore compatto e dunque $T_{g\mathcal{X}_E}(K)$ è totalmente limitato.

Ora essendo $\|T_g\| = \|g\|_{L_\infty}$ abbiamo anche

$$\left\| \int_{M \setminus E} f g \, d\mu \right\| \leq \varepsilon$$

per ogni $f \in K$ il che implica che anche $T_g(K)$ è totalmente limitato. ■

Teorema 12.5.11. *Dato uno spazio di Banach X, se ogni sottospazio chiuso separabile di X soddisfa la proprietà di Radon-Nikodym allora anche X la soddisfa.*

Dimostrazione. Supponiamo che ogni sottospazio chiuso e separabile di X soddisfa la proprietà di Radon-Nikodym, sia $T : L_1(\mu) \to X$ lineare e continua ed $E_i \in \mathcal{M}(\mu)$ una successione di insiemi misurabili. Sia $\mathcal{S}$ la σ-algebra generata da questi E_i, allora $L_1(\mathcal{S}, \mu)$ è un sottospazio chiuso e separabile di $L_1(\mu)$. Posto

$$S : f \in L_1(\mathcal{S}, \mu) \to T(f) \in \mathrm{Cl}\big(T[L_1(\mathcal{S}, \mu)]\big)$$

per ipotesi esisterà $g \in L_\infty(\mathcal{S}, \mu, X)$ per cui $S(f) = \int_M f g \, d\mu$.

Ora posto $F = \left\{ \mathcal{X}_{E_i} \;\middle|\; i \in \mathbb{N} \right\}$ esso sarà ovviamente uniformemente integrabile e limitato (μ è finita) ed il lemma appena dimostrato implica che $S(F) = T(F)$ è totalmente limitato in X ed ammette così estratte convergenti. Ma per l'arbitrarietà della successione E_i anche l'insieme

$$\left\{ T(\mathcal{X}_E) \;\middle|\; E \in \mathcal{M}(\mu) \right\}$$

è totalmente limitato in X, poiché le funzioni semplici sono dense in $L_1(\mu)$ avremo che T ha valori in un sottospazio separabile che per ipotesi soddisfa ancora la proprietà di Radon-Nikodym.

Per l'arbitrarietà di T segue allora che tutto X possiede la proprietà di Radon-Nikodym. ■

Corollario 12.5.12. *Gli spazi riflessivi soddisfano la proprietà di Radon-Nikodym.*

Dimostrazione. Sappiamo che i sottospazi chiusi di spazi riflessivi sono anch'essi riflessivi, inoltre se uno di questi sottospazi è anche separabile allora il rispettivo spazio duale è ancora riflessivo e separabile. Perciò ogni sottospazio chiuso e separabile soddisfa la proprietà di Radon-Nikodym e perciò l'intero spazio la soddisfa. ■

In certe occasioni risulta necessario utilizzare il teorema di Radon-Nikodym per rappresentare misure vettoriali con una particolare struttura. Per alcuni di questi casi però il teorema di Radon-Nikodym è troppo potente in quanto è possibile ottenere un risultato analogo anche per spazi che non soddisfano la proprietà di Radon-Nikodym:

Teorema 12.5.13. *Siano μ una misura esterna finita su M ed $\mathscr{M} \subseteq \mu$. Allora esiste un'applicazione lineare e continua $\mathscr{E} : L_1(\mu, X) \to L_1(\mathscr{M}, \mu, X)$ tale che per ogni $1 \leq p < +\infty$*

- *$\mathscr{E}f = f$ se $f \in L_1(\mathscr{M}, \mu, X)$;*
- *$\int_U f\, d\mu = \int_U \mathscr{E}f\, d\mu$ per ogni $U \in \mathscr{M}$;*
- *se $f \in L_p(\mu, X)$ allora $\mathscr{E}f \in L_p(\mathscr{M}, \mu, X)$ ed $\|\mathscr{E}f\|_{L_p} \leq \|f\|_{L_p}$.*

Dimostrazione. Supponiamo inizialmente che $X = \mathbb{R}$, per ogni $f \in L_1(\mu)$ definiamo la misura vettoriale

$$\nu : U \in \mathscr{M} \to \int_U f\, d\mu \in \mathbb{R}$$

ed il teorema di Radon-Nikodym implica che esiste un unica $\mathscr{E}f \in L_1(\mathscr{M}, \mu)$ per cui

$$\int_U f\, d\mu = \int_U \mathscr{E}f\, d\mu$$

ed in particolare lavorando sulle variazioni totali $\|\mathscr{E}f\|_{L_1} \leq \|f\|_{L_1}$. Si tenga presente che poiché f di solito non è $\mathscr{M}$-misurabile la disuguaglianza può essere stretta.

L'unicità implica che $\mathscr{E}f = f$ quando f è anche $(\mathscr{M}, \mu)$-fortemente misurabile e perciò $\mathscr{E}$ è continua su L_1.

Se $f \in L_p(\mu) \subseteq L_1(\mu)$ con $p > 1$ essendo μ finita dimostriamo che

$$\|\mathscr{E}f\|_{L_p} \leq \|f\|_{L_p}$$

Notiamo innanzitutto che possiamo assumere $f \geq 0$ senza perdere in generalità, inoltre se $f \leq g$ quasi ovunque avremo anche $\mathscr{E}f \leq \mathscr{E}g$ quasi ovunque. Per convessità abbiamo per ogni $t, t_0 \geq 0$

$$t^p \geq p t_0^{p-1}(t - t_0) + t_0^p = p t_0^{p-1} t + (1-p) t_0^p$$

ora per ogni $x \in \mathbb{Q}$ con $x \geq 0$ definiamo

$$a_x = p x^{p-1} \qquad\qquad b_x = (1-p) x^p$$

chiaramente $t^p \geq a_x t + b_t$ per ogni $t \geq 0$.

Ora prendiamo un qualunque $t_0 \geq 0$, per ogni $\varepsilon > 0$ esisterà $x \in \mathbb{Q}$ tale che $x \geq 0$ e $t_0 - \varepsilon \leq x \leq t_0$. Abbiamo allora

$$t_0^p - a_x t_0 - b_x = t_0^p - x^p - a_x \left(t_0 - x\right) \leq t_0^p - x^p + p t_0^{p-1}\left(t_0 - x\right)$$

che tende a 0 quando $\varepsilon \to 0^+$. Abbiamo così dimostrato che esistono $a_n, b_n \in \mathbb{R}$ tali che

$$t^p = \sup_{n\in\mathbb{N}} a_n t + b_n$$

per ogni $t \geq 0$.

Per quanto detto prima avremo $\mathscr{E}\left(a_n f + b_n\right) \leq \mathscr{E}\left(f^p\right)$ quasi ovunque su M per ogni $n \in \mathbb{N}$. D'altronde

$$(\mathscr{E}f)^p = \sup_{n\in\mathbb{N}} a_n \mathscr{E}f + b_n = \sup_n \mathscr{E}\left(a_n f + b_n\right) \leq \mathscr{E}\left(f^p\right)$$

quasi ovunque, e perciò

$$\int_M (\mathscr{E}f)^p \, d\mu \leq \int_M \mathscr{E}\left(f^p\right) d\mu = \int_M f^p \, d\mu < +\infty$$

ovvero $\mathscr{E}f \in L_p(\mathscr{M}, \mu)$ con $\left\|\mathscr{E}f\right\|_{L_p} \leq \left\|f\right\|_{L_p}$.

Passiamo ora al caso X generico, definiamo inizialmente $\mathscr{E}$ sulle funzioni semplici $s \in L_1(\mu, X)$ nella seguente maniera:

$$\mathscr{E}\left(\sum_{i=1}^n x_i \mathcal{X}_{A_i}\right) = \sum_{i=1}^n x_i \mathscr{E}\left(\mathcal{X}_{A_i}\right)$$

tale definizione rimane coerente con la precedente grazie alla linearità di $\mathscr{E}$.

Dato che la linearità è semplice da verificare passiamo direttamente a dimostrarne la continuità. Si ha per ogni $1 \leq p < +\infty$

$$\begin{aligned}\int_M \left\|\mathscr{E}\left(\sum_{i=1}^n x_i \mathcal{X}_{A_i}\right)\right\|^p d\mu &= \int_M \left\|\sum_{i=1}^n x_i \mathscr{E}\left(\mathcal{X}_{A_i}\right)\right\|^p d\mu \leq \int_M \left(\sum_{i=1}^n \left\|x_i\right\| \mathscr{E}\left(\mathcal{X}_{A_i}\right)\right)^p d\mu \\ &= \int_M \left(\mathscr{E}\left(\sum_{i=1}^n \left\|x_i\right\| \mathcal{X}_{A_i}\right)\right)^p d\mu = \left\|\mathscr{E}\left(\sum_{i=1}^n \left\|x_i\right\| \mathcal{X}_{A_i}\right)\right\|_{L_p}^p \leq \left\|\sum_{i=1}^n \left\|x_i\right\| \mathcal{X}_{A_i}\right\|_{L_p}^p = \left\|\sum_{i=1}^n x_i \mathcal{X}_{A_i}\right\|_{L_p}^p\end{aligned}$$

e per il teorema di Hahn-Banach possiamo così estendere $\mathscr{E}$ su tutto $L_1(\mu, X)$ in modo tale che $\left\|\mathscr{E}f\right\|_{L_1} \leq \left\|f\right\|_{L_1}$.

Poiché le funzioni semplici sono dense in ogni spazio L_p con $p < +\infty$ la disuguaglianza di sopra implica anche che $\left\|\mathscr{E}f\right\|_{L_p} \leq \left\|f\right\|_{L_p}$, ancora essendo le funzioni fortemente $(\mathscr{M}, \mu)$-misurabili approssimabili tramite funzioni semplici $\mathscr{M}$-misurabili otteniamo $\mathscr{E}f = f$ per ogni $f \in L_1(\mathscr{M}, \mu, X)$. ■

12.6 Il duale degli spazi L_p

Il prossimo passo è lo studio dello spazio duale di $L_p(\mu, X)$ quando $1 \leq p < +\infty$, in particolare mostreremo che per ogni $\Lambda \in \left[L_p(\mu, X)\right]^*$ esiste un unica $g \in L_q(\mu, X^*)$ tale che per ogni $f \in L_p(\mu, X)$

$$\Lambda(f) = \int_B g(a)[f(a)] \, d\mu(a)$$

e inoltre $\|\Lambda\|_{(L_p)^*} = \|g\|_{L_q}$.

Per comodità di notazione poniamo sempre $\langle f : x\rangle = f(x)$ per ogni $x \in X$, $f \in X^*$.

Proposizione 12.6.1. *Presi μ misura finita su M, $1 \leq p \leq +\infty$, $q = p/(p-1)$, $g \in L_q(\mu, X^*)$ e posto*

$$\Lambda_g(f) = \int_M \langle g : f\rangle \, d\mu$$

per ogni $f \in L_p(\mu, X)$ allora Λ_g è ben definita, appartiene a $\left[L_p(\mu, X)\right]^*$ e $\|\Lambda_g\| = \|g\|_{L_q}$.

Dimostrazione. Dalla disuguaglianza di Hölder abbiamo che Λ_g è ben definita e inoltre $\|\Lambda_g\| \leq \|g\|_{L_q}$ e quindi Λ_g appartiene al duale di $L_p(\mu, X)$.

Sia ora $1 < p < +\infty$, il teorema di densità ci permette di affermare che esiste una successione di funzioni $s_n \in L_q$ a valori in un insieme numerabile che approssimano g sia puntualmente quasi ovunque sia in norma L_q. Prendiamo così un qualunque $\varepsilon > 0$, esisterà $s = \sum_{i=1}^{+\infty} x_i^* \mathcal{X}_{A_i}$ con gli A_i a due a due disgiunti per cui $\|g - s\|_{L_q} \leq \varepsilon$ e possiamo supporre che g ed s non siano nulle quasi ovunque. Ma allora avremo anche degli $x_i \in X$ di norma unitaria tali che $\left\langle x_j^* : x_j\right\rangle \geq (1-\varepsilon)\left\|x_j^*\right\|$. Definiamo così

$$t(a) = \sum_{i=i}^{+\infty} x_i \left(\frac{\|x_i^*\|}{\|s\|_{L_q}}\right)^{q-1} \mathcal{X}_{A_i}(a)$$

Ricordiamo che $p(q-1) = q$ e dunque

$$\|t\|_{L_p}^p = \frac{1}{\|s\|_{L_q}^q} \sum_{i=1}^{+\infty} \|x_i^*\|^q \mu(A_i) = 1$$

$$\Lambda_s(t) \geq (1-\varepsilon) \sum_{i=1}^{+\infty} \frac{\|x_i^*\|^q}{\|s\|_{L_q}^{q-1}} \mathcal{X}_{A_i} = (1-\varepsilon)\|s\|_{L_q}$$

Dalla disuguaglianza di Hölder abbiamo anche che $\|\Lambda_g - \Lambda_s\| \leq \varepsilon$ e quindi mettendo tutto assieme

$$(1-\varepsilon)\left(\|g\|_{L_q} - \varepsilon\right) \leq (1-\varepsilon)\|s\|_{L_q} \leq \|\Lambda_s\| \leq \|\Lambda_g\| + \varepsilon$$

e facendo tendere ε a 0 si ha la disuguaglianza opposta.

Sia ora $q = 1$ e $p = +\infty$, in questo modo ci basta porre direttamente

$$t(a) = \sum_{i=i}^{N} x_i \mathcal{X}_{A_i}(a)$$

in quanto $\|t\|_{L_\infty} = 1$ e la dimostrazione procede alla stessa maniera.

Se invece $q = +\infty$ e $p = 1$ sia $i \in \mathbb{N}$ tale che $\mu(A_i) > 0$ ed $\|x_i^*\| > \|s\|_{L_\infty} - \varepsilon$, poniamo stavolta

$$t(a) = \frac{x_i}{\mu(A_i)} \mathcal{X}_{A_i}(a)$$

in questo modo

$$\|t\|_{L_1} = 1$$

$$\Lambda_s(t) \geq (1-\varepsilon)\|x_i^*\| \geq (1-\varepsilon)\|s\|_{L_\infty} - (1-\varepsilon)\varepsilon$$

e procedendo come a prima si ha la tesi. ■

Abbiamo così verificato che quando X è uno spazio di Banach l'applicazione $g \in L_q\left(M, X^*\right) \to \Lambda_g \in \left[L_p\left(M, X\right)\right]^*$ è una isometria lineare tra spazi di Banach. Sotto certe condizioni è possibile dimostrare che questa isometria è anche suriettiva e quindi è possibile rappresentare i funzionali di $\left(L_p\right)^*$ tramite funzioni L_q.

Teorema 12.6.2. *Sia μ una misura finita su M, $1 \leq p < +\infty$ ed X uno spazio di Banach tale che X^* soddisfa la proprietà di Radon-Nikodym, allora posto $q = p(p-1)$ per ogni $\Lambda \in \left(L_p(\mu, X)\right)^*$ esiste $g \in L_q(\mu, X^*)$ tale che $\Lambda = \Lambda_g$.*

Dimostrazione. Consideriamo $\Lambda \in \left(L_p(\mu, X)\right)^*$ e $\nu : \mathscr{M}(\mu) \to X^*$ definita in modo tale che per ogni $x \in X$

$$\nu(E)(x) = \Lambda\left(x\mathcal{X}_E\right) \leq \|\Lambda\| \, \|x\| \sqrt[p]{\mu(E)}$$

È immediato constatare che ν è una misura vettoriale μ-continua, mostriamo che ha variazione limitata.

Scelta una partizione finita E_i di M per ogni $\varepsilon > 0$ esistono $x_i \in X$ tali che $\left\|x_i\right\| = 1$ ed $\nu(E_i)(x_i) \geq \left\|\nu(E_i)\right\| - \varepsilon / n$ e perciò

$$\sum_{i=1}^n \left\|\nu(E_i)\right\| - \varepsilon \leq \sum_{i=1}^n \nu(E_i)(x_i) = \Lambda\left(\sum_{i=1}^n x_i \mathcal{X}_{E_i}\right) \leq \|\Lambda\| \sqrt[p]{\sum_{i=1}^n \left\|x_i\right\|^p \mu(E_i)} = \|\Lambda\| \sqrt[p]{\mu(M)}$$

e dunque $\|\nu\|(M) \leq \|\Lambda\| \sqrt[p]{\mu(M)}$. Il teorema di Radon-Nikodym implica che esiste $g \in L_1(\mu, X^*)$ tale che

$$\nu(E) = \int_E g \, d\mu$$

in particolare avremo

$$\Lambda\left(\sum_{i=1}^n x_i \mathcal{X}_{A_i}\right) = \sum_{i=1}^n \nu(A_i)(x_i) = \sum_{i=1}^n \left(\int_{A_i} g \, d\mu\right)(x_i) = \int_M \left\langle g : \sum_{i=1}^n x_i \mathcal{X}_{A_i}\right\rangle d\mu$$

e quindi $\Lambda = \Lambda_g$ sulle funzioni semplici.

Per ogni $n \in \mathbb{N}$ definiamo

$$E_n = \left\{x \in M \;\middle|\; \left\|g(x)\right\| \leq n\right\} \in \mathscr{M}(\mu)$$

allora $g\mathcal{X}_{E_n} \in L_\infty(\mu, X^*)$ ed $\left|\Lambda(s\mathcal{X}_{E_n})\right| \leq \left\|g\mathcal{X}_{E_n}\right\|_{L_q} \|s\|_{L_p}$ per ogni funzione semplice s. Quindi

$$\Lambda_n(f) = \Lambda(f\mathcal{X}_{E_n} g) = \int_M \left\langle g\mathcal{X}_{E_n} : f\right\rangle d\mu$$

per ogni $f \in L_p(\mu, X)$ e dunque per quanto detto precedentemente $\left\|g\mathcal{X}_{E_n}\right\|_{L_q} = \left\|\Lambda_n\right\| \leq \|\Lambda\|$.

La successione di funzioni $h_n(x) = \left\|g(x)\mathcal{X}_{E_n}(x)\right\|^q$ è crescente e converge puntualmente ad $\left\|g(x)\right\|^q$, applicando così il teorema di Beppo-Levi abbiamo $\left\|g\right\|_{L_q} \leq \|\Lambda\|$ ed $g \in L_q(\mu, X^*)$. Applicando Hahn-Banach e la densità delle funzioni semplici avremo infine che $\Lambda = \Lambda_g$ ed il teorema è così dimostrato. ■

Corollario 12.6.3. *Se X è riflessivo, μ finita ed $1 < p < +\infty$ allora $L_p(\mu, X)$ è riflessivo con $\left[L_(\mu, X)\right]^* = L_q(\mu, X^*)$.*

12.7 Il teorema fondamentale del calcolo integrale

Torniamo ora al nostro discorso sulla differenziabilità, e poniamo $M = [0,T]$ e sia μ la misura di Lebesgue usuale. Data quindi una funzione $f : [0,T] \to X$ differenziabile in $t \in (0,T)$ osserviamo che esiste $f'(t) \in X$ tale che

$$df(t)[r] = f'(t)r$$

per ogni $r \in \mathbb{R}$ e inoltre

$$\lim_{r\to 0} \frac{f(t+r)-f(t)}{r} = f'(t)$$

come per la derivata usuale in una variabile. Anche in questo caso vale il teorema fondamentale del calcolo integrale:

Teorema 12.7.1 (fondamentale del calcolo integrale)**.** *Sia $f : [0,T] \to X$ continua su tutto il dominio e differenziabile su $(0,T)$ in modo tale che la derivata f' sia estendibile per continuità su tutto $[0,T]$. Allora f' è Bochner-integrabile su $[0,T]$ ed*

$$f(T) - f(0) = \int_0^T f'(t)\, dt$$

Dimostrazione. La funzione f' è uniformemente continua in quanto definita su un compatto e dunque l'insieme $f'([0,T])$ è totalmente limitato e dunque separabile. Poiché f' è continua allora è anche debolmente misurabile e per il teorema di Pettis è anche fortemente misurabile. Essendo f' limitata su $[0,T]$ avremo che

$$\int_0^T \|f'(x)\|\, dx \le \|f'\|_\infty T$$

e quindi f' è Bochner-integrabile.

Consideriamo adesso la funzione $F(t) = \int_0^t f'(s)\, ds$ definita su $[0,T]$ con $F(0) = 0$, dimostriamo che F è derivabile e la sua derivata coincide con f'. Si ha per $r > 0$

$$\left\| \frac{F(t+r)-F(t)}{r} - f'(t) \right\| = \left\| \frac{1}{r}\int_t^{t+r} f'(s)\, ds - f'(t) \right\|$$
$$\le \frac{1}{r}\int_t^{t+r} \|f'(s) - f'(t)\|\, ds \le \varepsilon$$

sfruttando la continuità di f' in t. Un ragionamento analogo vale quando $r < 0$ e dunque $F'(t) = f'(t)$.

Prendiamo adesso una qualunque funzione $L \in X^*$, allora $L(F-f)$ è ancora una funzione differenziabile da $(0,T)$ in $\mathbb{R}$ con derivata

$$[L(F-f)]' = L\left(F' - f'\right) = 0$$

quindi $L(F-f)$ è una funzione costante. Se per assurdo $F - f$ non fosse costante allora esisterebbero $t_1, t_2 \in (0,T)$ tale che $F(t_1) - f(t_1) \neq F(t_2) - f(t_2)$. Ma per il secondo teorema di Mazur esisterebbe $L \in X^*$ tale che $L[F(t_1) - f(t_1)] \neq L[F(t_2) - f(t_2)]$ generando cos' un assurdo.

Esiste così una costante $c \in X$ tale che $f(t) = F(t) + c$ per ogni $t \in [0,T]$ e dunque $f(0) = F(0) + c = c$ ovvero $F(T) = f(T) - f(0)$. ■

Sia f una funzione di classe C^1 su un insieme aperto U di uno spazio di Banach X a valori in un altro spazio di Banach Y. Allora per ogni $x, y \in U$ se poniamo $F(t) = f(x + t(y - x))$ ed applichiamo il teorema fondamentale del calcolo integrale abbiamo che

$$f(y) - f(x) = \int_0^1 df(x + t(y - x))[y - x]\ dt$$

Consideriamo stavolta una funzione f differenziabile con continuità $k + 1$ volte in un intorno dell'origine, quindi possiamo porre $F_k(x) = \sum_{i=0}^{k} \frac{1}{i!} d^i f(0)[x^i]$ e porre $R_k = f - F_k$. Dimostriamo che

$$R_k(x) = \frac{1}{k!} \int_0^1 d^{k+1} f(tx)[x] \left[(1 - t)^k x^k\right]\ dt$$

innanzitutto supponiamo che $k = 0$, allora la tesi segue direttamente dal teorema fondamentale del calcolo integrale

$$f(x) = f(0) + \int_0^1 df(tx)[x]\ dt$$

Supponiamo ora che $k = 1$ e poniamo $L(t) = df(tx)[(1-t)x]$, per la regola di derivazione del prodotto si ha

$$L'(t)r = d^2 f(tx)[rx][(1 - t)x] - df(tx)[rx]$$

ovvero $L'(t) = d^2 f(tx)[x, (1 - t)x] - df(tx)[x]$. Essendo $L(0) = df(0)[x]$ dal teorema fondamentale del calcolo integrale si ha

$$-df(0)[x] = \int_0^1 d^2 f(tx)[x][(1 - t)x]\ dt - \int_0^1 df(tx)[x]\ dt$$

e sostituendo all'equazione precedente

$$f(x) = f(0) + df(0)[x] + \int_0^1 d^2 f(tx)[x, (1 - t)x]\ dt$$

Poniamo adesso $k = 2$ e $L(t) = d^2 f(tx)[(1 - t)x, (1 - t)x]$, derivando sempre per parti avremo che

$$L'(t) = d^3 f(tx)[x, (1 - t)x, (1 - t)x] - d^2 f(tx)[x, (1 - t)x] - d^2 f(tx)[(1 - t)x, x]$$

e per il teorema del calcolo integrale e il teorema di Schwartz

$$-d^2 f(0)[x^2] = \int_0^1 d^3 f(tx)[x] \left[(1 - t)^2 x^2\right]\ dt - 2\int_0^1 d^2 f(tx)[x][(1 - t)x]\ dt$$

sostituendo

$$f(x) = f(0) + df(0)[x] + \frac{1}{2} d^2 f(0)[x^2] + \frac{1}{2} \int_0^1 d^3 f(tx)[x, (1 - t)^2 x^2]\ dt$$

iternando il procedimento si ottiene la tesi.

Capitolo 13

Misure gaussiane

13.1 Insiemi cilindrici

Consideriamo adesso uno spazio vettoriale topologico reale X, ed indichiamo con $\mathscr{B}(X)$ la σ-algebra di Borel indotta dalla topologia di X. Quando parliamo di misure esterne Borel regolari non possiamo applicare i risultati di [8] direttamente ad ogni spazio vettoriale topologico in quanto non sono spazi localmente compatti in generale. Un modo per ovviare a questo inconveniente è costruire una σ-algebra più piccola con migliori proprietà.

Dato $F \subseteq X^*$ definiamo $\mathscr{E}(X, F)$ come la più piccola σ-algebra su X che renda misurabili tutte le funzioni di F, equivalentemente è la σ-algebra generata dai sottoinsiemi nella forma $f^{-1}(I)$ al variare di $f \in F$ ed I aperto di $\mathbb{R}$. Quando $F = X^*$ scriveremo semplicemente $\mathscr{E}(X)$. Chiaramente $\mathscr{E}(X, F) \subseteq \mathscr{B}(X)$ ma in generale l'inclusione è stretta, anzi di solito $\mathscr{E}(X)$ non coincide nemmeno con la σ-algebra generata dagli aperti deboli di X in quanto non sempre è presente una base numerabile.

Un sottoinsieme C di X è detto **cilindrico** se e solo se esistono $n \in \mathbb{N}$, $F \in L(X, \mathbb{R}^n)$ e $B \in \mathscr{B}\left(\mathbb{R}^n\right)$ tali che $C = F^{-1}(B)$. Analogamente una funzione $f : X \to \mathbb{R}^m$ è detta cilindrica se e solo se esistono $F \in L(X, \mathbb{R}^n)$ ed $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ Borel misurabile tali che $f = \varphi \circ F$.

Proposizione 13.1.1. *Gli insiemi cilindrici appartengono generano $\mathscr{E}(X)$, in particolare le funzioni cilindriche sono $\mathscr{E}(X)$-misurabili.*

Dimostrazione. Sia $\mathscr{C}$ la famiglia degli insiemi cilindrici. Poiché $X^* = L(X, \mathbb{R})$ abbiamo $\mathscr{E}(X) \subseteq \sigma(\mathscr{C})$. Mostriamo ora che $\mathscr{C}$ è un π-sistema.

Presi $C_1 = F_1^{-1}(B_1)$, $C_2 = F_2^{-1}(B_2)$ insiemi cilindrici con $B_1 \subseteq \mathbb{R}^n$, $B_2 \subseteq \mathbb{R}^m$ definiamo $G \in L(X, \mathbb{R}^{m+n})$ in modo tale che

$$G^i(x) = F_1^i(x) \text{ se } 1 \leq i \leq n$$
$$G^i(x) = F_2^{i-n}(x) \text{ se } n+1 \leq i \leq n+m$$

chiaramente G è lineare e continua. Posto $B = B_1 \times B_2$ è ben noto che $B \in \mathscr{B}(\mathbb{R}^{m+n})$ e quindi l'insieme

$$G^{-1}(B) = F_1^{-1}(B_1) \cap F_2^{-1}(B_2) = C_1 \cap C_2$$

è ancora un insieme cilindrico e quindi $\mathscr{C}$ è un π-sistema. D'altronde la σ-algebra $\mathscr{B}(\mathbb{R}^n)$ è generata dagli n-rettangoli mentre le funzioni in $L(X, \mathbb{R}^n)$ si ottengono tramite una combinazione di elementi di X^*, possiamo così ripetere il ragionamento precedente per dimostrare che $\mathscr{C} \subseteq \mathscr{E}(X)$ e la tesi segue dal lemma di Dynkin. ∎

Un'altro motivo per cui conviene lavorare con gli insiemi cilindrici è che rimangono stabili rispetto all'operazione di prodotto di σ-algebre, definita nella seguente maniera:

Definizione 13.1.2. Dati due insiemi M, N non vuoti che possiedono rispettivamente le σ-algebre $\mathscr{M}$, $\mathscr{N}$ definiamo sul prodotto cartesiano $M \times N$ la **σ-algebra prodotto** $\mathscr{M} \times \mathscr{N}$ come la σ-algebra generata dagli insiemi nella forma $A \times B$ con $A \in \mathscr{M}$ e $B \in \mathscr{N}$.

Se X ed Y sono spazi vettoriali topologici generici l'inclusione $\mathscr{B}(X) \times \mathscr{B}(Y) \subseteq \mathscr{B}(X \times Y)$ è quasi sempre stretta, a meno che non siano entrambi separabili. Con gli insiemi cilindrici invece le cose vanno molto meglio:

Proposizione 13.1.3. *Dati due spazi vettoriali localmente convessi X, Y si ha*

$$\mathscr{E}(X) \times \mathscr{E}(Y) = \mathscr{E}(X \times Y)$$

Dimostrazione. Innanzitutto vale l'inclusione $\mathscr{E}(X) \times \mathscr{E}(Y) \subseteq \mathscr{E}(X \times Y)$, infatti presi $L \in L(X, \mathbb{R}^n)$, $R \in L(Y, \mathbb{R}^m)$ possiamo definire $L \oplus R \in L(X \times Y, \mathbb{R}^{m+n})$ in modo tale che

$$(L \oplus R)(x, y) = \left[L(x), R(y)\right]$$

e non è difficile verificare che vale l'uguaglianza $(L \oplus R)^{-1}(A \times B) = L^{-1}(A) \times L^{-1}(B)$.

Viceversa prendiamo un qualunque $L \in L(X \times Y, \mathbb{R}^n)$, esistono e sono univocamente determinate $L_1 \in L(X, \mathbb{R}^n)$, $L_2 \in L(Y, \mathbb{R}^n)$ tali che $L(x, y) = L_1(x) + L_2(y) = S[(L_1 \oplus L_2)(x, y)]$ dove $S(a, b) = a + b$. Per quanto detto prima e dalla definizione di σ-algebra prodotto abbiamo $(L_1 \oplus L_2)^{-1}(U) \in \mathscr{E}(X) \times \mathscr{E}(Y)$ per ogni $U \in \mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(\mathbb{R}^n)$.

D'altronde è ben noto che $\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(\mathbb{R}^n) = \mathscr{B}(\mathbb{R}^n \times \mathbb{R}^n)$ e perciò S è $\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(\mathbb{R}^n)$-misurabile. Questo implica così che $L^{-1}(U) \in \mathscr{E}(X) \times \mathscr{E}(Y)$ per ogni $U \in \mathscr{B}(\mathbb{R}^n)$ e per ogni $L \in L(X \times Y, \mathbb{R}^n)$ dimostrando in tal modo la tesi. ■

Proposizione 13.1.4. *Se X è uno spazio di Frechet separabile allora esiste $F \subseteq X$ al più numerabile tale che $\mathscr{B}(X) = \mathscr{E}(X) = \mathscr{E}(X, F)$.*

Dimostrazione. Fissiamo un qualunque $r > 0$ e sia d la metrica (invariante, regolare sugli scalari, bilanciata e convessa) di X, dato che $D(0, r)$ è chiuso e convesso per il secondo teorema di Mazur esiste $G_r \subseteq X^* \times \mathbb{R}$ tale che

$$D(0, r) = \bigcap_{(f,c) \in G} \left\{x \in X \;\middle|\; f(x) \leq c\right\}$$

per il teorema di Lindelöf possiamo trovare $F_r' \subseteq G_r$ al più numerabile che soddisfi comunque la medesima uguaglianza.

Se poniamo

$$F = \left\{f \in X^* \;\middle|\; \exists r \in \mathbb{Q}^+, c \in \mathbb{R} \text{ tali che } (f, c) \in F_r\right\}$$

abbiamo ancora che F è numerabile, ed inoltre tutti gli aperti di X sono unione numerabile di palle chiuse con raggio razionale, dunque $\mathscr{B}(X) = \mathscr{E}(X) = \mathscr{E}(X, F)$. ■

Consideriamo adesso lo spazio vettoriale reale $\mathbb{R}^\infty = F(\mathbb{N}, \mathbb{R})$ con la topologia prodotto, chiaramente è uno spazio di Frechet separabile con le applicazioni $r_n : \mathbb{R}^\infty \to \mathbb{R}^n$ definite come

$$r_n(x_1, x_2, \dots) = (x_2, x_2, \dots, x_n)$$

continue, vogliamo ottenere un risultato di struttura per $\mathscr{E}(\mathbb{R}^\infty)$ che coincide con $\mathscr{B}(\mathbb{R}^\infty)$ dato che $\mathbb{R}^\infty$ è uno spazio di Frechet separabile.

Proposizione 13.1.5. *L'insieme $E \subseteq \mathbb{R}^\infty$ è cilindrico se e solo se esistono $n \in \mathbb{N}$ ed $B \in \mathscr{B}(\mathbb{R}^n)$ tali che $E = r_n^{-1}(B)$.*

Dimostrazione. La continuità di r_n implica che $r_n^{-1}(B)$ è cilindrico. Presa una qualunque funzione $f \in L\left(\mathbb{R}^\infty, \mathbb{R}^n\right)$ dimostriamo che esiste $N \in \mathbb{N}$ tale che $L(x) = 0$ per ogni $x \in \mathbb{R}^\infty$ per cui $x_i = 0$ per ogni $1 \leq i \leq N$.

Per assurdo esiste L tale che per ogni $N \in \mathbb{N}$ esiste $x^N \in \mathbb{R}^\infty$ tale che $x_i^N = 0$ se $i \leq N$ ma $L\left(x^N\right) = 1$. Chiaramente $\lim_{n\to+\infty} x_i^N = 0$ per ogni i e per la convergenza puntuale della topologia prodotto si ha $x^N \to 0$. Ma allora per continuità di L

$$1 = \lim_{n\to+\infty} L\left(x^N\right) = L(0) = 0$$

generando così un assurdo.

Per ogni $L \in L\left(\mathbb{R}^\infty, \mathbb{R}^n\right)$ esistono così $m \in \mathbb{N}$ ed $L^m : \mathbb{R}^m \to \mathbb{R}^n$ lineare per cui $L = L^m \circ r_m$ e dunque gli insiemi cilindrici di $\mathbb{R}^\infty$ sono tutti e soli nella forma $r_n^{-1}(B)$. ■

Lemma 13.1.6. *Se X è uno spazio vettoriale localmente convesso allora*

$$\mathscr{E}(X) = \left\{L^{-1}(B) \;\middle|\; L \in L\left(X, \mathbb{R}^\infty\right), B \in \mathscr{E}\left(\mathbb{R}^\infty\right) = \mathscr{B}\left(\mathbb{R}^\infty\right)\right\}$$

Dimostrazione. Sia T la famiglia al secondo membro, mostriamo che è una σ-algebra, in particolare che è chiusa rispetto alle intersezioni numerabili. Scegliamo una qualunque applicazione biettiva $p : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ e per ogni $i \in \mathbb{N}$ definiamo l'applicazione

$$q^i : j \in \mathbb{N} \to p^{-1}(i,j) \in \mathbb{N}$$

e definiamone il pull-back $q_\#^i : \mathbb{R}^\infty \to \mathbb{R}^\infty$ alla solita maniera.

Supponiamo che per ogni $i \in \mathbb{N}$ abbiamo definito $L_i \in F(X, \mathbb{R}^\infty)$ ed $A_i \in \mathscr{E}(\mathbb{R}^\infty)$, definiamo $L : X \to \mathbb{R}^\infty$ in modo tale che per ogni $x \in X$

$$L(x)_j = \left[L_{p_1(j)}(x)\right]_{p_2(j)}$$

che è banalmente continua. Definiamo inoltre

$$C = \bigcap_{i\in\mathbb{N}} \left(q_\#^i\right)^{-1}(A_i)$$

con $C \in \mathscr{E}(\mathbb{R}^\infty)$ in quanto i pull-back dei rimescolamenti di indici sono applicazioni continue rispetto alla topologia prodotto.

Osserviamo adesso che per ogni $i, j \in \mathbb{N}$ abbiamo

$$\left[q_\#^i\left[L(x)\right]\right]_j = \left[L(x)\right]_{q^i(j)} = \left[L_{p_1(q^i(j))}(x)\right]_{p_2(q^i(j))} = \left[L_i(x)\right]_j$$

ovvero $q_\#^i \circ L = L_i$ e quindi $L(x) \in \left(q_\#^i\right)^{-1}(A_i)$ se e solo se $L_i(x) \in A_i$. Abbiamo così ottenuto

$$x \in L^{-1}(C) \Leftrightarrow x \in \bigcap_{i\in\mathbb{N}} L_i^{-1}(A_i)$$

e quindi T è stabile sotto l'intersezione numerabile e dunque è una σ-algebra.

Dato che T contiene gli insiemi cilindrici abbiamo $\mathscr{E}(X) \subseteq T$, il prossimo passo è dimostrare l'inclusione opposta. Per ogni $L \in L\left(X, \mathbb{R}^\infty\right)$ definiamo la famiglia di insiemi

$$R_L = \left\{B \in \mathscr{E}(\mathbb{R}^\infty) \;\middle|\; L^{-1}(B) \in \mathscr{E}(X)\right\}$$

chiaramente R_L è una σ-algebra e se $R_L = \mathscr{E}(\mathbb{R}^\infty)$ per ogni L avremo $\mathscr{E}(X) = T$.

Sia $B \subseteq \mathbb{R}^\infty$ cilindrico, per la proposizione precedente esistono $n \in \mathbb{N}$ ed $C \in \mathscr{B}(\mathbb{R}^n)$ tali che $B = r_n^{-1}(C)$ e dunque $L^{-1}(B) = \left(r_n \circ L\right)^{-1}$ è cilindrico in X per definizione, dunque $B \in R_L$ e perciò $\mathscr{E}(\mathbb{R}^\infty) = R_L$. ■

13.2 Trasformata di Fourier

Definizione 13.2.1. Sia μ una misura esterna finita su un F-spazio separabile X compatibile con $\mathscr{E}(X)$, definiamo la **trasformata di Fourier** di μ l'applicazione $\tilde{\mu} : X^* \to \mathbb{C}$ definita in modo tale che

$$\tilde{\mu}(f) = \int_X e^{if(x)}\, d\mu(x)$$

Questa definizione è ben definita in quanto f è $\mathscr{E}(X)$ fortemente misurabile per ogni $f \in X^*$ dato che ha valori in uno spazio separabile.

Per quanto detto poco prima ed essendo $\left|e^{it}\right| = 1$ la trasformata di Fourier è sempre ben definita con $\left|\tilde{\mu}(f)\right| \leq \mu(X)$ e quindi $\tilde{\mu}$ è limitata. Per il teorema della convergenza dominata $\tilde{\mu}$ è anche continua su X^*, alla luce di queste osservazioni possiamo definire la **trasformazione di Fourier** l'applicazione $\mathscr{F}$ che ad ogni misura esterna finita μ associ la quantità $\mathscr{F}(\mu) = \tilde{\mu} \in C(X^*, \mathbb{C})$ con

$$\left\|\mathscr{F}(\mu)\right\|_\infty \leq \mu(X)$$

Osserviamo anche che, posto $T_{a,k}(x) = kx + a$ per ogni $a, x \in X, k \in \mathbb{K}$, per la (12.2.2) abbiamo

$$\mathscr{F}\left[\left(T_{a,1}\right)_\# \mu\right](f) = \int_X e^{if(x)}\, d\left(T_{a,1}\right)_\# \mu(x) = \int_X e^{if(x)+if(a)}\, d\mu(x) = e^{if(a)}\tilde{\mu}(f)$$

Per il teorema di convergenza dominata si ha chiaramente $\tilde{\mu} \in C(X^*, \mathbb{C})$ e perciò possiamo porre $\mathscr{F}$ come un'applicazione che associa ad ogni misura finita la sua trasformata di Fourier in $C(X^*, \mathbb{C})$.

Il prossimo passo è dimostrare che la trasformata di Fourier è iniettiva, per farlo però dobbiamo momentaneamente lavorare su $\mathbb{R}$. Consideriamo lo spazio vettoriale

$$S = \left\{\phi \in C^\infty(\mathbb{R}, \mathbb{C}) \;\middle|\; \lim_{t \to \pm\infty} t^m \phi^{(n)}(t) = 0\ \forall m, n \in \mathbb{N}_0\right\}$$

osserviamo che $S \subseteq L_1(\mathbb{R})$ in quanto

$$\left\|\phi\right\|_{L_1} = \int \left(1 + x^2\right)\left|\phi(x)\right| \frac{1}{1 + x^2}\, dx \leq \left\|(1 + x^2)\phi\right\|_{L_\infty} \left\|\frac{1}{1 + x^2}\right\|_{L_1}$$

Possiamo così definire la quantità

$$\mathscr{F}(\phi)(x) = \int_{-\infty}^{+\infty} e^{ixy}\phi(y)\, dy$$

definita per ogni $\phi \in S$, dimostriamo ora che $\mathscr{F}(\phi) \in S$ per ogni $\phi \in S$. Dalla formula di integrazione per parti si dimostra facilmente che $\mathscr{F}(\phi') = -ix\mathscr{F}(\phi)$, verifichiamo ora che $\mathscr{F}(\phi)$ è di classe C^∞.

Per ogni $x \in \mathbb{R}$ e per ogni successione $t_n \neq 0$ convergente a 0 si ha

$$\frac{\mathscr{F}(\phi)(x+t_n) - \mathscr{F}(\phi)(x)}{t_n} = \int_{-\infty}^{+\infty} \frac{e^{it_n y} - 1}{t_n} e^{ixy}\phi(y)\, dy$$

possiamo così applicare la ben nota formula

$$\left|e^{it} - 1\right| \leq |\sin t| + |\cos t - \cos 0| \leq |t| + \frac{1}{2}|t|^2 \tag{13.2.1}$$

in modo da ottenere$\left|\frac{e^{it_n y-1}}{t_n}\right| leq \left|y\right| + \frac{1}{2}\left|y\right|^2$ (supponendo $\left|t_n\right| \leq 1$). Dato che sia $y\phi(y)$ che $y^2\phi(y)$ appartengono ad $L_1(\mathbb{R})$ possiamo applicare il teorema della convergenza dominata ed ottenere così

$$\left[\mathscr{F}(\phi)\right]'(x) = \int_{-\infty}^{+\infty} iye^{ixy}\phi(y)\, dy = i\mathscr{F}(y\phi)(x)$$

Posto $\psi = \mathscr{F}(\phi)$ osserviamo che $\lim_{x\to\pm\infty}\psi(x) = 0$ in quanto

$$\left|x\psi(x)\right| \leq \left\|i\mathscr{F}(\phi')\right\|_\infty \leq \left\|\phi'\right\|_1 < +\infty$$

e questo è possibile se e solo se ψ converge a 0 all'infinito. Per concludere per ogni $m, n \in \mathbb{N}_0$ si ha

$$\lim_{x\to\pm\infty} x^n \frac{d^m\psi}{dx^m}(x) = \lim_{x\to\pm\infty} x^n \mathscr{F}\left(i^m y^m \phi\right)(x) = \lim_{x\to\pm\infty} \mathscr{F}\left[i^{m+n}\frac{d^n}{dy^n}\left(y^m\phi\right)\right](x) = 0$$

Abbiamo così dimostrato il seguente risultato:

Teorema 13.2.2. *Per ogni $\phi \in S$ si ha $\mathscr{F}(\phi) \in S$.*

Ora è perfettamente sensato definire $\mathscr{F}^2(\phi) = \mathscr{F}\left[\mathscr{F}(\phi)\right]$ per ogni $\phi \in S$. Prima di determinare esattamente $\mathscr{F}^2$ è necessario fare qualche piccolo calcolo preliminare.

Scelto un qualunque $a > 0$ e posto $\phi(t) = e^{-at^2/2}$ sappiamo che $\phi \in S$ ed $\phi' = -at\phi$ dunque posto $\psi = \mathscr{F}\phi$ si avrà $-it\psi = ai\psi'$ e quindi $\psi(t) = k_a e^{-t^2/(2a)}$ con

$$k_a = \psi(0) = \int_{\mathbb{R}} e^{-at^2/2}\, dt = a^{-1/2}\int_{\mathbb{R}} e^{-t^2/2}\, dt = \frac{k_1}{\sqrt{a}}$$

Quando però andiamo a calcolare $\mathscr{F}^2(\phi)$ per $\phi \in S$ non possiamo invertire l'ordine degli integrali in quanto e^{ixy} non è una funzione L_1. Ciò può essere ovviato aggiungendo un termine mollificante dipendente da un parametro $\varepsilon >$ che si andrà poi a far tendere a 0, per più nel dettaglio

$$\begin{aligned}\mathscr{F}\left[e^{-\varepsilon y^2/2}\mathscr{F}(\phi)\right] &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{ixy}e^{iyz}e^{-\varepsilon y^2/2}\phi(z)\, dzdy = \int_{\mathbb{R}}\phi(z)\int_{\mathbb{R}} e^{i(x+z)y}e^{-\varepsilon y^2/2}\, dydz \\ &= \frac{k_1}{\sqrt{\varepsilon}}\int_{\mathbb{R}}\phi(z)e^{-(x+z)^2/(2\varepsilon)}\, dz = k_1\int_{\mathbb{R}}\phi\left(\sqrt{\varepsilon}y - x\right)e^{-y^2/2}\, dy\end{aligned}$$

e per il teorema della convergenza dominata si avrà

$$\mathscr{F}^2\phi(x) = k_1^2\phi(-x)$$

e quindi $\mathscr{F}$ è invertibile. Lo stesso ragionamento vale anche su $\mathbb{R}^n$ per ogni $n \in \mathbb{N}$.

Ora che abbiamo ottenuto tutti i requisiti necessari possiamo dimostrare l'iniettività della trasformata di Fourier sulle misure esterne finite e Borel regolari:

Teorema 13.2.3. *Date μ, ν misure esterne finite su uno spazio vettoriale localmente convesso X compatibili con la σ-algebra $\mathscr{E}(X)$ se $\tilde{\mu} = \tilde{\nu}$ allora μ e ν coincidono su $\mathscr{E}(X)$.*

Dimostrazione. Supponiamo inizialmente che $X = \mathbb{R}^n$. Poniamo $f_y(x) = y \cdot x$ con $y \in \mathbb{R}^n$, per ogni aperto non vuoto $U \subseteq \mathbb{R}^n$ esiste una successione di funzioni $\phi_n \in S$ tale che

$$\mathcal{X}_{U_{1/n}} \leq \phi_n \leq \mathcal{X}_U$$

dove $U_r = \mathrm{Int}_r(U) = \left\{x \in U \,\middle|\, \mathrm{dist}(x, \partial U) > r\right\}$ e quindi ϕ_n converge puntualmente ad $\mathcal{X}_U$. Dal teorema di Fubini abbiamo anche

$$\int_{\mathbb{R}^n} \mathscr{F}^{-1}\phi_n(x)\hat{\mu}(x)\,dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathscr{F}^{-1}\phi_n(x)e^{ix\cdot y}\,d\mu(y)dx = \int_{\mathbb{R}^n}\phi_n(y)\,d\mu(y)$$

ripetendo lo stesso procedimento per ν si avrà

$$\int_{\mathbb{R}^n}\phi_n(y)\,d\mu(y) = \int_{\mathbb{R}^n}\phi_n(y)\,d\nu(y)$$

e passando al limite per il teorema della convergenza dominata si ha $\mu(U) = \nu(U)$ e coincidono sugli aperti. Per il corollario 12.2.1 coincideranno su tutti i borelliani.

Sia ora X localmente convesso, per ogni $f \in L(X, \mathbb{R}^n)$ e per ogni $y \in \mathbb{R}^n$ abbiamo

$$\mathscr{F}(f_\#\mu)(y) = \int_{\mathbb{R}^n} e^{iy\cdot z}\,df_\#\mu(z) = \int_X e^{iy\cdot f(x)}\,d\mu(x) = \mathscr{F}(\mu)(f_y)$$

con $f_y \in X^*$. Dunque se $\tilde{\mu} = \tilde{\nu}$ allora $f_\#\mu, f_\#\nu$ coincidono su tutti i borelliani di $\mathbb{R}^n$ e quindi μ e ν coincidono sugli insiemi cilindrici di X. Per il corollario 12.2.1 queste due misure coincideranno su $\mathscr{E}(X)$. ■

13.3 Misure gaussiane

In generale non è possibile definire misure di Legesgue in spazi di dimensione infinita. Vale infatti il seguente risultato:

Lemma 13.3.1. *Se X è uno spazio di Banach separabile di dimensione infinita e μ è una misura di Borel su X tale che*

- *$\mu(U + x) = \mu(U)$ per ogni U aperto di X;*
- *esiste U aperto contenente l'origine tale che $\mu(U) < +\infty$*

allora μ è identicamente nulla.

Dimostrazione. Per ipotesi esiste $r > 0$ tale che $\mu(B_r) < +\infty$, grazie al lemma 7.3.3 possiamo trovare una successione numerabile di palle disgiunte $B(x_i, \delta)$ con $\delta > 0$ fissato tutte contenute in B_r. Per l'invarianza per traslazioni abbiamo

$$n\mu(B_\delta) = \sum_{i=1}^{n} \mu\left(B(x_i, \delta)\right) \leq \mu(B_r) < +\infty$$

per ogni n, il che è possibile se e solo se $\mu(B_\delta) = 0$.

Dato che X è separabile, e quindi $\mathcal{N}_2$, possiamo ricoprire X con una quantità numerabile di palle di raggio δ e quindi $\mu(X) = 0$. ■

Una misura esterna Borel regolare μ su $\mathbb{R}$ è detta **gaussiana** se e solo se esistono $a \in \mathbb{R}$ (detta media di μ) ed $\sigma^2 > 0$ (detta varianza di μ) tali che per ogni borelliano $E \in \mathscr{B}(\mathbb{R})$

$$\mu(E) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_E \exp\left(-\frac{|x-a|^2}{2\sigma^2}\right) dx$$

si ricordi che $\mu(\mathbb{R}) = 1$ grazie alla scelta della costante di fronte all'integrale, inoltre se $a = 0$ diremo che μ è anche **simmetrica**. Scopo di questo capitolo è estendere il concetto di misura gaussiana agli spazi di Banach. Inoltre per ogni $a \in \mathbb{R}$, $\lambda \neq 0$ abbiamo che μ è gaussiana se e solo se $\left(T_{a,\lambda}\right)_\# \mu$ è gaussiana.

Definizione 13.3.2. Una misura esterna μ definita sulla spazio vettoriale topologico localmente convesso X e compatibile con $\mathscr{E}(X)$ è detta **gaussiana** se e solo se per ogni $f \in X^*$ la misura reale $f_\#\mu$ è gaussiana. Se inoltre tutte le $f_\#\mu$ sono simmetriche anche μ sarà detta simmetrica.

Dato che per ogni $f \in X^*$ si ha $f \circ T_{a,\lambda} = T_{f(a),\lambda} \circ f$ per la formula di cambiamento di variabile (12.2.2) avremo che $\left(T_{a,\lambda}\right)_\# \mu$ è gaussiana per ogni μ gaussiana. Inoltre per ipotest tutte le funzioni $f \in X^*$ sono μ-fortemente misurabili rispetto ad una qualunque misura gaussiana μ, più in generale $G \circ f$ è misurabile per ogni G borelliana definita su $\mathbb{R}$ a valori in qualche spazio di Banach separabile in quanto

$$(G \circ f)^{-1}(B) = f^{-1}[G^{-1}(B)]$$

che è misurabile in quanto $G^{-1}(B)$ è borelliano. Notiamo anche che $\mu(X) = 1$, ovvero le misure gaussiane sono finite.

È banale verificare anche che per ogni $F \in L(X, Y)$, con X ed Y spazi vettoriali topologici localmente convessi, se μ è una misura gaussiana su X allora $F_\#\mu$ è gaussiana su Y. Più in generale basterebbe prendere $F \in \mathscr{L}(X, Y)$ (non necessariamente continua) per cui però $f \circ F \in X^*$ per ogni $f \in Y^*$, ovvero F possiede l'auggunto.

Proposizione 13.3.3. *Se μ è una misura gaussiana allora $f \in L_p(\mu)$ per ogni $f \in X^*$ e per ogni $1 \leq p < +\infty$.*

Dimostrazione. Per la formula del cambio variabili si ha

$$\int_X |f|^p \, d\mu = \int_{-\infty}^{+\infty} |x|^p \; d\left(f_\#\mu\right)(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} |x|^p \exp\left(-\frac{|x-a|^2}{2\sigma^2}\right) dx$$

che è chiaramente finito per ogni p finito. ■

Un criterio utilissimo per determinare se una misura esterna è gaussiana è dato dal seguente risultato:

Teorema 13.3.4. *Una misura esterna finita μ sullo spazio vettoriale localmente convesso X è gaussiana se e solo se*

$$\mathscr{F}(\mu)(f) = e^{iL(f)-\frac{1}{2}B(f,f)}$$

per qualche $L \in (X^*)^+$ e $B : X^* \times X^* \to \mathbb{R}$ bilineare simmetrica tale che $B(f,f) \geq 0$.

Inoltre se μ è gaussiana allora

$$L(f) = a_\mu(f) = \int_X f\, d\mu$$
$$B(f,g) = V_\mu(f,g) = \int_X [f(x) - L(f)]\,[g(x) - L(g)]\, d\mu(x)$$

Dimostrazione. Se μ è gaussiana allora

$$\mathscr{F}(\mu)(f) = \int_X e^{if(x)}\, d\mu(x) = \int_{-\infty}^{+\infty} e^{it}\, d(f_\# \mu)(t) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \int_{-\infty}^{+\infty} e^{it} e^{-\left(t-a_f\right)^2/\left(2\sigma_f^2\right)}\, dt$$

dato che la funzione

$$g(t) = \frac{1}{\sqrt{2\pi\sigma_f^2}} e^{-\left(t-a_f\right)^2/\left(2\sigma_f^2\right)}$$

appartiene ad S possiamo applicare i risultati ottenuti precedentemente. Derivando

$$g' = -\frac{t-a_f}{\sigma_f^2} g = -\frac{t}{\sigma_f^2} g + \frac{a_f}{\sigma_f^2} g$$

e applicando la trasformazione di Fourier ad entrambi i membri

$$-it\tilde{g} = -\frac{i}{\sigma_f^2}\tilde{g}' + \frac{a_f}{\sigma_f^2}\tilde{g}$$

ed essendo $\tilde{g}(0) = 1$ abbiamo

$$\tilde{g}(t) = \exp\left(-ia_f t + \frac{1}{2}\sigma_f^2 t^2\right)$$

e la prima implicazione si ottiene osservando che $\mathscr{F}(\mu)(f) = \tilde{g}(1)$, $a_f = L(f)$ e $\sigma_f^2 = B(f,f)$.

Supponiamo ora che μ sia una misura esterna finita che soddisfi l'uguaglianza di sopra per ogni $f \in X^*$, possiamo definire su $\mathbb{R}$ la seguente misura

$$\nu_f(E) = \frac{1}{\sqrt{2\pi B(f,f)}} \int_{-\infty}^{+\infty} \exp\left[-\frac{\left(t - L(f)\right)^2}{2B(f,f)}\right]\, dt$$

per ipotesi abbiamo $\mathscr{F}\left(f_\# \mu\right) = \mathscr{F}\left(\nu_f\right)$ e dunque per il teorema 13.2.3 $f_\# \mu = \nu_f$ su $\mathscr{B}(\mathbb{R})$. Per l'arbitrarietà di f segue così che μ è gaussiana. ∎

Questo criterio ci permette di dimostrare molte proprietà delle misure gaussiane, ad esempio

Corollario 13.3.5. *Il prodotto di misure gaussiane è gaussiana.*

Dimostrazione. Siano μ, η misure gaussiane su X ed Y rispettivamente. Per ogni $F \in (X \times Y)^*$ abbiamo per linearità

$$F(x,y) = F(x,0) + F(0,y) = f(x) + g(y)$$

con $f \in X^*$, $g \in Y^*$. Posto $m = \mu \times \eta$ essendo $\mathscr{E}(X) \times \mathscr{E}(Y) = \mathscr{E}(X \times Y)$ abbiamo che gli insiemi cilindrici di $X \times Y$ sono m-misurabili. Per il teorema di Fubini si ha

$$\begin{aligned}\mathscr{F}(m)(F) = \int_{X\times Y} e^{iF(x,y)}\,dm(x,y) &= \int_X e^{if(x)}\,d\mu(x) \int_Y e^{ig(y)}\,d\eta(y) \\ &= \exp\left[ia_\mu(f) + ia_\eta(g) - \frac{1}{2}\left[V_\mu(f,f) + V_\eta(g,g)\right]\right]\end{aligned}$$

Adesso essendo $\mu(X) = \eta(Y) = 1$ abbiamo per Fubini

$$\begin{aligned}a_\mu(f) + a_\eta(g) &= \int_X f(x)\,d\mu(x) + \int_Y g(y)\,d\eta(y) \\ &= \int_{X\times Y} f(x)\,dm(x,y) + \int_{X\times YY} g(y)\,dm(x,y) = a_m(F)\end{aligned}$$

allo stesso modo si verifica l'uguaglianza $V_\mu(f,f) + V_\eta(g,g) = V_m(F,F)$ in quanto sempre per Fubini

$$\int_{X\times Y}\left[f(x) - a_\mu(f)\right]\left[g(y) - a_\eta(g)\right] dm(x,y) = 0$$

e così m è una misura gaussiana. ■

Corollario 13.3.6. *Una misura gaussiana μ è simmetrica se e solo se $\mu(B) = \mu(-B)$ per ogni $B \in \mathscr{E}(X)$.*

Dimostrazione. Posto $\eta(B) = \mu(-B)$ abbiamo

$$\mathscr{F}(\eta)(f) = \mathscr{F}\left(R_\#\mu\right)(f) = \mathscr{F}(\mu)(f \circ R) = \mathscr{F}\mu(-f) = \exp\left[-ia_\mu(f) - \frac{1}{2}V_\mu(f,f)\right]$$

dove $R(x) = -x$. Perciò $\mu = \eta$ su $\mathscr{E}(X)$ se e solo se a_μ è identicamente nulla, in altre parole se e solo se μ è simmetrica. ■

Posto $S : (x,y) \in X \times X \to x + y \in X$ definiamo la **convoluzione** delle misure esterne μ, η su X con $\mathscr{E}(X) \subseteq \mu \cap \eta$ la misura $\mu * \eta = S_\#(\mu \times \eta)$. Dato che S è lineare e continua si ha $\mathscr{E}(X) \subseteq \mu * \eta$ e per ogni $B \in \mathscr{E}(X)$

$$(\mu * \eta)(B) = \int_{X\times X} \mathcal{X}_B(x+y)\,d(\mu\times\eta)(x,y) = \int_X \mu(B-y)\,d\eta(y)$$

mostriamo che la convoluzione di misure gaussiane è ancora gaussiana.

Per ogni $f \in X^*$ abbiamo infatti

$$\mathscr{F}(\mu * \eta)(f) = \int_{X\times X} e^{if(x+y)}\,d(\mu\times\eta)(x,y) = \mathscr{F}(\mu)(f)\mathscr{F}(\eta)(f)$$

e così

$$\begin{aligned}a_{\mu*\eta} &= a_\mu + a_\eta \\ V_{\mu*\eta} &= V_\mu + V_\eta\end{aligned}$$

in particolare se $\eta(B) = \mu(-B)$ per ogni $B \in \mathscr{E}(X)$ allora $\mu * \eta$ è sempre una misura gaussiana simmetrica.

13.4 Lo spazio di Cameron-Martin

Data una misura gaussiana μ su uno spazio vettoriale localmente convesso X consideriamo l'applicazione

$$A_\mu : f \in X^* \to f - a_\mu(f) \in L_2(\mu)$$

chiaramente A_μ è lineare e iniettiva (non possiamo parlare di continuità dato che X non è normato) con $\|A_\mu f\|_{L_2}^2 = V_\mu(f,f)$. Indichiamo con $\tilde{X}_\mu^*$ l'immagine di A_μ in $L_2(\mu)$ e con X_μ^* la sua chiusura, chiaramente X_μ^* è uno spazio di Hilbert con il prodotto scalare di L_2. Definiamo anche

$$|x|_{H(\mu)} = \sup\left\{f(x) \;\middle|\; f \in X^*, V_\mu(f,f) \le 1\right\} \;\forall x \in X$$
$$H(\mu) = \left\{x \in X \;\middle|\; |x|_{H(\mu)} < +\infty\right\}$$

chiaramente $H(\mu)$ è uno spazio normato, detto **spazio di Cameron-Martin**.

Possiamo definire anche l'applicazione $R_\mu : X_\mu^* \to (X^*)^+$ (non definiamo su X^* alcuna topologia al momento, per questo introduciamo il duale algebrico) tale che per ogni $f \in X_\mu^*$, $g \in X^*$

$$R_\mu(f)(g) = \int_X f(x)\left[g(x) - a_\mu(g)\right] d\mu(x)$$

chiaramente R_μ è iniettiva. Mostriamo ora che possiamo trasportare la struttura di spazio di Hilbert da X_μ^* ad $H(\mu)$ grazie al seguente risultato:

Teorema 13.4.1. *Dato $x \in X$ abbiamo $x \in H(\mu)$ se e solo se esiste $f \in X_\mu^*$ tale che $g(x) = R_\mu(f)(g)$ per ogni $g \in X^*$. In tal caso*

$$|x|_{H(\mu)} = \|f\|_{L_2}$$

Dimostrazione. Supponiamo inizialmente che $x \in H(\mu)$, definiamo l'applicazione lineare $P_x : \tilde{X}_\mu^* \to \mathbb{R}$ tale che per ogni $f \in X^*$

$$P_x(A_\mu(f)) = f(x)$$

che è ben definita per l'iniettività di A_μ. Questa applicazione è anche continua in X_μ^* in quanto

$$\left|P_x(A_\mu(f))\right| \le \sqrt{V_\mu(f,f)}\,|x|_{H(\mu)} = \left\|A_\mu(f)\right\|_{L_2} |x|_{H(\mu)}$$

e dunque può essere estesa univocamente su tutto X_μ^*. Dato che siamo in uno spazio di Hilbert esiste un unico $f \in X_\mu^*$ tale che per ogni $g \in X^*$

$$g(x) = \int_X f(y)\left[g(y) - a_\mu(g)\right] d\mu(y) = R_\mu(f)(g)$$

e dalla definizione di $|\cdot|_{H(\mu)}$ segue immediatamente che $\|f\|_{L_2} = |x|_{H(\mu)}$.

Viceversa se vale l'uguaglianza abbiamo per ogni $g \in X^*$

$$g(x) = R_\mu(f)(g) = \int_X f(y)\left[g(y) - a_\mu(g)\right] d\mu(y) \le \|f\|_{L_2} \sqrt{V_\mu(g,g)}$$

per la disuguaglianza di Hölder. ■

L'iniettività di R_μ ci permette così di costruire un'applicazione iniettiva $x \in H(\mu) \to \hat{x} \in X^*_\mu$ definita in modo tale che $x = R_\mu(\hat{x})$. Possiamo così dotare $H(\mu)$ del seguente prodotto interno

$$(x:y) = \int_X \hat{x}\hat{y}\, d\mu$$

rendendo $H(\mu)$ uno spazio prehilbertiano. Se fosse $R_\mu(X^*_\mu) \subseteq X$ allora $H(\mu)$ sarebbe automaticamente uno spazio di Hilbert. In particolare per quanto detto precedentemente sugli integrali di Bochner

Teorema 13.4.2. *Se X fosse uno spazio di Banach separabile allora per ogni misura gaussiana μ su X abbiamo*

$$a_\mu(f) = f\left(\int_X x\, d\mu(x)\right) = f(m_\mu)\ \forall f \in X^*$$
$$R_\mu(f) = \int_X \left(x - m_\mu\right) f(x)\, d\mu(x)$$

in particolare H è uno spazio di Hilbert separabile.

Dimostrazione. Usa il teorema di Fernique per stabilire che l'applicazione identica $f(x) = x$ appartiene ad $L_p(\mu)$ per ogni misura gaussiana μ e per ogni p finito. ∎

Torniamo ad X spazio vettoriale localmente convesso. Innanzitutto abbiamo bisogno di un piccolo risultato di continuità per passare da X^* ad X^*_μ:

Proposizione 13.4.3. *Per ogni $g \in X^*_\mu$ la misura $\nu = g_\# \mu$ è gaussiana e simmetrica.*

Dimostrazione. Siano $f_n \in X^*$ tali che $h_n = f_n - a_\mu(f_n) \to g$ in norma L_2. Ma allora convergono puntualmente μ-quasi ovunque e perciò $(h_n)_\# \mu(E) \to g_\#\mu(E)$ per ogni $E \in \mathscr{B}(\mathbb{R})$. D'altronde

$$(h_n)_\# \mu(E) = \frac{1}{\sqrt{2\pi V_\mu(f_n, f_n)}} \int_E \exp - \frac{t^2}{2V_\mu(f_n, f_n)}\, dt$$

passando al limite in n si avrà che $g_\#\mu$ è gaussiana con media nulla e varianza $\|g\|^2_{L_2}$. ∎

Proposizione 13.4.4. *Per ogni $g \in X^*_\mu$ la misura ν su X definita come*

$$\nu(E) = \int_X \exp\left[g(x) - \frac{1}{2} V_\mu \|g\|^2\right] d\mu(x)$$

per ogni $E \in \mathscr{E}(X)$ è ancora gaussiana con

$$\mathscr{F}(\nu)(f) = \exp\left[iR_\mu(g)(f) + ia_\mu(f) - \frac{1}{2} V_\mu(f, f)\right]$$

Dimostrazione. Per ogni $g \in X^*_\mu$ è possibile verificare che

$$\int_X e^{|g(x)|}\, d\mu(x) = \int_{-\infty}^{+\infty} e^{|t|}\, d(g_\#\mu)(t) < +\infty$$

e quindi l'esponendiale è sommabile. D'altronde con un po' di calcoli si vede anche che ν è una misura di probabilità.

Per ogni $z \in \mathbb{C}$ e per ogni $f \in X^*$ definiamo

$$F_1(z) = \exp\left[-\frac{1}{2}\|g\|^2\right]\int_X \exp\left[i[f(x) - zg(x)]\right] d\mu(x)$$
$$F_2(z) = \exp\left[zR_\mu(g)(f) - \frac{1+z^2}{2}\|g\|^2 + ia_\mu(f) - \frac{1}{2}V_\mu(f,f)\right]$$

vogliamo dimostrare che coincidono. Prendiamo allora $z = t \in \mathbb{R}$ e per densità possiamo valutare la trasformata di Fourier anche su X_ν^* oltre che su X^*, quindi

$$\begin{aligned}F_1(t) &= \exp\left[-\frac{1}{2}\|g\|^2\right]\int_X \exp\left[i[f(x) - tg(x)]\right] d\mu(x)\\ &= \exp\left[-\frac{1}{2}\|g\|^2\right]\exp\left[ia_\mu(f - tg) - \frac{1}{2}\|f - tg - a_\mu(f)\|^2\right]\\ &= \exp\left[-\frac{1}{2}\|g\|^2\right]\exp\left[ia_\mu(f) - \frac{t^2}{2}\|g\|^2 - \frac{1}{2}V_\mu(f,f) + tR_\mu(g)(f)\right]\\ &= \exp\left[ia_\mu(f) - \frac{1+t^2}{2}\|g\|^2 - \frac{1}{2}V_\mu(f,f) + tR_\mu(g)(f)\right] = F_2(t)\end{aligned}$$

Trattandosi di funzioni olomorfe coincideranno su tutto $\mathbb{C}$, in particolare su $z = i$:

$$\mathscr{F}(\nu)(f) = \int_X \exp\left[if(x) + g(x) - \frac{1}{2}\|g\|^2\right] d\mu(x) = \exp\left[iR_\mu(g)(f) + ia_\mu(f) - \frac{1}{2}V_\mu(f,f)\right]$$

■

Teorema 13.4.5 (Cameron-Martin)**.** *Per ogni $x \in X$ sia $\mu_x(E) = \mu(E - x)$. Se $x \in H(\mu)$ allora posto*

$$\phi_x(y) = \exp\left[\hat{x}(y) - \frac{1}{2}|x|^2_{H(\mu)}\right]$$

si ha $\mu_x = \phi_x\mu$ su $\mathscr{E}(X)$. Se invece $x \in X \setminus H$ allora μ e μ_h sono concentrati su sottoinsiemi disgiunti.

Dimostrazione. Per ogni $f \in X^*$ abbiamo

$$\mathscr{F}(\mu_x)(f) = \int_X e^{if(y)}\, d\mu_x(y) = \int_X e^{if(x+y)}\, d\mu(y) = \exp\left[iR_\mu(\hat{x})(f) + ia_\mu(f) - \frac{1}{2}V_\mu(f,f)\right]$$

e per l'iniettività della trasformata di Fourier abbiamo $\mu_x = \phi_x\mu$.

Consideriamo ora il caso $X = \mathbb{R}$ e μ misura gaussiana con media a e varianza σ^2. Per ogni $x \in \mathbb{R}$ si ha

$$\mu_x(E) = K\int_E e^{-\frac{(t+x-a)^2}{2\sigma^2}}\, dt = K\int_E e^{-\frac{(t-a)^2+x^2-2x(t-a)}{2\sigma^2}}\, dt = \int_E e^{-\frac{x^2}{2\sigma^2}+\frac{x(t-a)}{\sigma^2}}\, d\mu(t)$$

e quindi μ_x è assolutamente continua rispetto a μ.

Ricordando il teorema di Hellinger

$$\mathscr{H}(\mu_x,\mu) = K\int_{-\infty}^{+\infty} e^{-\frac{x^2}{4\sigma^2}+\frac{x(t-a)}{2\sigma^2}-\frac{(t-a)^2}{2\sigma^2}}\, dt = K\int_{-\infty}^{+\infty} e^{-\frac{x^2}{8\sigma^2}-\frac{(t-a-x/2)^2}{2\sigma^2}}\, dt = e^{-\frac{x^2}{8\sigma^2}}$$

dunque

$$\left\|\mu_x - \mu\right\|(\mathbb{R}) \geq 2\left(1 - e^{-\frac{x^2}{8\sigma^2}}\right)$$

Supponiamo ora che X generico e $|x|_{H(\mu)} = +\infty$, notiamo anche che

$$f_{\#}\mu_x(E) = \mu[f^{-1}(E) - x] = \mu[f^{-1}(E - f(x))] = \left(f_{\#}\mu\right)_{f(x)}(E)$$

per ogni $f \in X^*$. Dato che il push-forward risuce il numero di partizioni possibili per le misure vettoriali si ha la disuguaglianza

$$\left\|f_{\#}\mu - \left(f_{\#}\mu\right)_{f(x)}\right\|(\mathbb{R}) \leq \left\|\mu - \mu_x\right\|(X)$$

Per ipotesi esistono $f_n \in X^*$ tali che $V_\mu(f_n, f_n) = 1$ ma $f_n(x) \geq n$, mettendo tutto assieme

$$\left\|\mu - \mu_x\right\|(X) \geq \left\|(f_n)_{\#}\mu - \left((f_n)_{\#}\mu\right)_{f_n(x)}\right\|(\mathbb{R}) \geq 2\left(1 - e^{-\frac{f_n(x)^2}{8V_\mu(f,f)}}\right)$$

il secondo membro tende a 2 al crescere di n, ottenendo così $\left\|\mu - \mu_x\right\|(X) = 2$ ovvero $H(\mu, \mu_x) = 0$. La tesi segue dal teorema di Hellinger. ■

Il teorema di Cameron-Martin ci dice che il nostro spazio prehilbertiano $H(\mu)$ è esattamente lo spazio per cui vale la regola di spostamento della media per una misura gaussiana. Altre direzioni portano infatti a misure indipendenti da quella originale.

13.5 Il teorema di Fernique

Da questo punto supporremo che X sia uno spazio di Banach separabile in modo da poter utilizzare i risultati del capitolo precedente sugli integrali di Bochner e verificare che la media a_μ di una misura gaussiana μ è proprio un elemento di X. Essendo X separabile sappiamo già che la funzione identità $\mathrm{id}_X(x) = x$ è fortemente misurabile su $\mathscr{B}(X) = \mathscr{E}(X)$, abbiamo però bisogno di un risultato di sommabilità prima di poterne calcolare l'integrale.

A questo scopo ci viene in aiuto il

Teorema 13.5.1 (Fernique). *Se X è uno spazio di Banach separabile e μ è una misura gaussiana simmetrica su X allora esistono $\alpha, C > 0$ tali che per ogni $t \geq 0$*

$$\mu(\{x \in X \mid \|x\| \geq t\}) \leq Ce^{-\alpha t^t}$$

Dimostrazione. Consideriamo la funzione lineare e continua

$$R : (x, y) \in X \times X \to \left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right)$$

dimostriamo innanzitutto che $R_{\#}\left(\mu \times \mu\right) = \mu \times \mu$ su $\mathscr{E}(X)$. Per ogni $F \in (X \times X)^*$ con $F(x, y) = f(x) + g(y)$ abbiamo

$$\begin{aligned}\mathscr{F}\left[R_{\#}\left(\mu \times \mu\right)\right](F) &= \int_{X\times X} e^{\frac{i}{\sqrt{2}}(f(x)+f(y)+g(x)-g(y))}\, d\left(\mu \times \mu\right)(x, y)\\ &= \int_X e^{\frac{i}{\sqrt{2}}(f+g)}\, d\mu \int_X e^{\frac{i}{\sqrt{2}}(f-g)}\, d\mu = \exp\left[-\frac{1}{4}\left[V_\mu(f+g, f+g) + V_\mu(f-g, f-g)\right]\right]\\ &= \exp\left[-\frac{1}{2}\left[V_\mu(f, f) + V_\mu(g, g)\right]\right] = \mathscr{F}(\mu \times \mu)\end{aligned}$$

per simmetria.

Per comodità poniamo $\mu^2 = \mu \times \mu$, abbiamo allora

$$\begin{aligned}
\mu(\|x\| \geq t)\,\mu(\|x\| \leq s) &= \mu^2\left(\|x\| \geq t, \|y\| \leq s\right) \\
&= \mu^2\left(\frac{1}{\sqrt{2}}\|x+y\| \geq t, \frac{1}{\sqrt{2}}\|x-y\| \leq s\right) \\
&\leq \mu^2\left(\frac{1}{\sqrt{2}}\|x\| + \|y\| \geq t, \frac{1}{\sqrt{2}}\big|\|x\| - \|y\|\big| \leq s\right) \\
&\leq \mu^2\left(\|x\|, \|y\| \geq \frac{t-s}{\sqrt{2}}\right) \\
&= \left[\mu\left(\|x\| \geq \frac{t-s}{\sqrt{2}}\right)\right]^2
\end{aligned}$$

abbiamo così mostrato che

$$\frac{\mu(\|x\| \geq t)}{\mu(\|x\| \leq s)} \leq \left[\frac{\mu\left(\|x\| \geq \frac{t-s}{\sqrt{2}}\right)}{\mu(\|x\| \leq s)}\right]^2$$

■

Appendice A

Miscellanea

A.1 Equivalenza delle forme di Hahn-Banach

È possibile dimostrare il teorema di Hahn-Banach anche partendo dal teorema di Hahn-Banach geometrico (senza usare l'assioma della scelta). La dimostrazione qui proposta è stata ricavata dall'autore e potrebbe risultare sbagliata o inutilmente complicata, perciò chi sta studiando questi appunti per la preparazione di un esame può tranquillamente tralasciare questa parte.

Dato uno spazio vettoriale V su $\mathbb{R}$ una **quasinorma** su V è un'applicazione $p : V \to \mathbb{R}$ tale che

$$\begin{gathered} p(0) = 0 \\ p(x+y) \leq p(x) + p(y) \text{ per ogni } x, y \in V \\ p(tx) = tp(x) \text{ per ogni } t \geq 0 \end{gathered}$$

In questa sezione daremo per scontato il teorema di Hahn-Banach geometrico

Teorema A.1.1. *Dato uno spazio vettoriale V su $\mathbb{R}$ e due sottoinsiemi convessi non vuoti A e B tali che $A \cap B = \emptyset$ allora se tutti i punti di A sono internali esisterà una funzione lineare $f : V \to \mathbb{R}$ non identicamente nulla e una costante $\alpha \in \mathbb{R}$ tale che*

$$f(b) \leq \alpha < f(a)$$

per ogni $a \in A$ e $b \in B$.

e lo useremo per dimostrare il teorema di Hahn-Banach analitico:

Teorema A.1.2. *Dato uno spazio vettoriale V su $\mathbb{R}$, un sottospazio vettoriale $W \leq V$, una funzione lineare $f : W \to \mathbb{R}$ e una quasinorma $p : V \to \mathbb{R}$ tali che $f(x) \leq p(x)$ per ogni $x \in W$ allora esiste $g : V \to \mathbb{R}$ lineare che coincide con f su W tale che $g(x) \leq p(x)$ per ogni $x \in V$.*

Il seguente lemma può essere verificato velocemente nel caso di seminorme su spazi vettoriali reali e complessi grazie ai risultati ottenuti precedentemente, però risulta opportuno ripetere la dimostrazione nel caso delle quasinorme

Lemma A.1.3. *Sia V spazio vettoriale reale, $f : V \to \mathbb{R}$ applicazione lineare e $p : V \to \mathbb{R}$ quasinorma. Se esistono $y \in V$, $r \in \mathbb{R}$, $s > 0$ tali che*

$$p(x-y) < s \Rightarrow f(x) \leq r$$

allora esisterà $M > 0$ tale che per ogni $x \in V$

$$f(x) \leq M p(x)$$

Dimostrazione. Per comodità di notazione poniamo $C = \{x \in V \mid p(x) < s\}$ allora per ipotesi sappiamo che $f(y + C) \subseteq (-\infty, r]$.

Possiamo supporre che f non sia identicamente nulla, allora per ogni $x \in V$ e per ogni $\varepsilon > 0$ avremo che

$$f\left(y + s\frac{x}{p(x)+\varepsilon}\right) \leq r \Rightarrow f(x) \leq \frac{r - f(y)}{s}[p(x) + \varepsilon]$$

ma ε non dipende da x e quindi avremo che $sM = r - f(y) > 0$ e $f(x) \leq M p(x)$. ■

Prima di proseguire con la trattazione introduciamo gli strumenti che ci serviranno. Fissiamo una funzione lineare $f : V \to \mathbb{R}$ non identicamente nulla e una quasinorma $p : V \to \mathbb{R}$ tali che $f(x) \leq p(x)$ su V, definiamo $T_f = \{x \in V \mid f(x) > 0\} \neq \emptyset$ e poniamo

$$\lambda^+(f) = \sup_{x \in T_f} \frac{f(x)}{p(x)} \in (0, 1]$$

questa quantità è ben definita in quanto se $f(x) > 0$ allora $p(x) > 0$ e per tali x si avrà $f(x) \leq \lambda^+(f)p(x)$. Se invece $f(x) < 0$ allora $-x \in T_f$ e dunque $f(x) \geq -\lambda^+(f)p(-x)$ ottenendo alla fine che

$$-\lambda^+(f)p(-x) \leq f(x) \leq \lambda^+(f)p(x)$$

stavolta per tutti gli $x \in V$.

Questa disuguaglianza può essere ulteriormente migliorata sostituendo la quasinorma p con la seguente funzione

$$p_f : x \in V \to \inf_{y \in \ker f} p(x - y)$$

che ha il vantaggio di fornire una stima migliore rispetto a p in quanto $f(x) \leq p_f(x) \leq p(x)$. Per verificare che è una seminorma osserviamo che $p_f(0) = 0$ e che $p_f(tx) = t p_f(x)$ in quanto $\ker f$ è comunque un sottospazio di W, per quanto riguarda la subadditività per ogni $x, y \in V$ e per ogni $\varepsilon > 0$ esistono $x_f, y_f \in \ker f$ tali che

$$p(x - x_k) \leq p_f(x) + \varepsilon \qquad\qquad p(y - y_k) \leq p_f(y) + \varepsilon$$

ma $x_f + y_f \in \ker f$ e dunque

$$p_f(x + y) \leq p(x - x_f + y - y_f) \leq p(x - x_f) + p(y - y_f) \leq p_f(x) + p_f(y) + 2\varepsilon$$

e per l'arbitrarietà di ε segue la tesi. Infine non è difficile verificare che

$$\lambda^+(f) = \sup_{x \in T_f} \frac{f(x)}{p_f(x)}$$
$$-\lambda^+(f)p_f(-x) \leq f(x) \leq \lambda^+(f)p_f(x)$$

Dimostriamo ora il seguente lemma

Lemma A.1.4. *Siano V spazio vettoriale reale, $f : V \to \mathbb{R}$ lineare e p una quasinorma tale che $f(x) \leq p(x)$ su V. Allora per ogni $x \in T_f$*

$$f(x) = \lambda^+(f)p_f(x)$$

Dimostrazione. Possiamo porre senza perdere in generalità che f non sia identicamente nulla. Per assurdo supponiamo che esiste $\varepsilon_0 > 0$ e un $x_0 \in T_f$ tali che $f(x_0) \leq \lambda^+(f)p_f(x_0) - \varepsilon_0$, tramite un riscalamento possiamo supporre che $p_f(x_0) = 1$. Dalla definizione di $\lambda^+(f)$ esisterà $y_0 \in T_f$ tale che $p_f(y_0) = 1$ e $\lambda^+(f) - \varepsilon_0 < f(y_0) \leq \lambda^+(f)$. Dimostreremo che $f(x_0) = f(y_0)$ per giungere così ad un assurdo.

Innanzitutto esisterà $t > 0$ tale che $f(x_0) = tf(y_0)$ e per linearità allora $x_0 - ty_0 \in \ker f$ e quindi $1 = p_f(x_0) = p_f(ty_0) = tp_f(y_0) = t$ ottenendo così la tesi. ∎

Una conseguenza di questo risultato è che per ogni $x, y \in V$ tali che $f(x), f(y) > 0$ allora

$$p_f(x + y) = p_f(x) + p_f(y)$$

Procediamo ora alla dimostrazione del teorema di Hahn-Banach analitico a partire dalla versione geometrica. Sia allora W sottospazio di V, $f : W \to \mathbb{R}$ lineare tale che $f(x) \leq p(x)$ su W allora consideriamo lo spazio vettoriale $X = V \times \mathbb{R}$ e il sottospazio

$$G = \left\{[x, f(x)] \in X \,\middle|\, x \in W\right\}$$

introduciamo su X anche la seminorma

$$p_*(x, r) = p(-x) + \alpha\,|r|$$

con $\alpha > 0$ una costante che determineremo dopo, per comodità poniamo $\lambda = \lambda^+(f) > 0$. Chiaramente $(0, 1) \notin G$ e vogliamo stimarne la distanza da G rispetto alla quasinorma p_* definita come

$$d_G = \inf_{x \in W} p_*(-x, 1 - f(x))$$

Per ogni $x \in W$ e per ogni $t \geq 0$ consideriamo la funzione

$$l(t) = p_*[(0, 1) - (tx, f(tx))] = tp(x) + \alpha\left|1 - tf(x)\right|$$

per x fissato questa funzione possiede due punti critici: uno in $t = 0$ con $l(0) = \alpha$ e un altro, che esiste solo se $x \in T_f$, in $tf(x) = 1$ dove $l(t) = p(x)/f(x)$. Quindi scegliamo $\alpha > 0$ sufficientemente grande in modo tale che esista $x \in T_f$ tale che $p(x)/f(x) < \alpha$. In questo modo avremo che

$$d_G = \inf_{x \in T_f} \frac{p(x)}{f(x)} = \frac{1}{\lambda}$$

Quindi tutti i punti dell'insieme convesso $C = \left\{(x, r) \in X \,\middle|\, p_*(-x, 1 - r) < 1/\lambda\right\}$ sono internali e C non interseca G, per il teorema di Hahn-Banach geometrico esisterà $F : X \to \mathbb{R}$ lineare tale che $F(x, f(x)) = 0$ per ogni $x \in W$ e $F(x, r) > 0$ se $p_*(-x, 1 - r) < 1/\lambda$.

Riadattando opportunamente il lemma A.1.3 esisterà $M > 0$ tale che per ogni $(x, r) \in X$

$$F(x, r) \geq -Mp_*(-x, -r) \Rightarrow F(x, r) \leq Mp_*(x, r)$$

riscalando opportunamente F avremo anche che $M = 1$ e $\lambda^+(F) = 1$.

Ora poiché $\ker F \supseteq G$ abbiamo che per ogni $(x,r) \in X$

$$\left(p_*\right)_F(0,1) = \inf_{(y,s)\in\ker f} p_*(-y,1-s) \le \inf_{(y,s)\in G} p_*(-y,1-s) = \frac{1}{\lambda}$$

ma poiché il convesso C non interseca $\ker F$ la disuguaglianza è in realtà un'uguaglianza e quindi per il lemma A.1.4 avremo che

$$F(0,1) = \left(p_*\right)_F(0,1) = \frac{1}{\lambda}$$

Definiamo infine

$$g(x) = -\lambda F(x,0)$$

innanzitutto avremo per ogni $x \in W$

$$0 = F(x, f(x)) = F(x,0) + f(x)F(0,1) = -\frac{g(x)}{\lambda} + \frac{f(x)}{\lambda} \Rightarrow g(x) = f(x)$$

mentre per ogni $x \in V$

$$g(x) = \lambda F(-x,0) \le \lambda p_*(-x,0) = \lambda p(x) \le p(x)$$

e il teorema di Hahn-Banach analitico è dimostrato.

A.2 Metriche non archimedee

Consideriamo uno spazio metrico (S,d) allora la metrica d è detta **non archimedea** se e solo se per ogni $x,y,z \in S$ vale la disuguaglianza

$$d(x,y) \le \max\{d(x,z), d(z,y)\} \tag{A.2.1}$$

questa disuguaglianza è una versione più forte della disuguaglianza triangolare tipica degli spazi metrici usuali, ciononostante le metriche non archimedee hanno proprietà davvero insolite anche per uno spazio metrico generico.

Proposizione A.2.1. *Se d è una metrica non archimedea allora per ogni $x,y,z \in S$ vale almeno una delle seguenti formule*

$$\begin{aligned} d(x,y) &\le d(x,z) = d(z,y) \\ d(x,z) &\le d(x,y) = d(y,z) \\ d(z,y) &\le d(z,x) = d(x,y) \end{aligned}$$

Dimostrazione. Se poniamo

$$\begin{aligned} a &= d(x,y) \\ b &= d(x,z) \\ c &= d(z,y) \end{aligned}$$

allora a meno di un riordinamento possiamo supporre $a \le b \le c$. Dalle proprietà della metrica si ha $c \le \max\{a,b\} = b$ e quindi $c = b$. ∎

In maniera euristica potremmo affermare che in una metrica non archimedea **tutti i triangoli sono isosceli o equilateri**. Le proprietà insolite di queste metriche non si esauriscono qui, difatti vale anche il seguente risultato

Proposizione A.2.2. *Fissato $r > 0$ la relazione su S*

$$x \sim y \Leftrightarrow d(x,y) < r$$

è una relazione di equivalenza in S. Quindi la famiglia

$$\{B(x,r)\}_{x\in S}$$

determina una partizione di S.

Dimostrazione. Le proprietà di riflessività e simmetria sono sempre verificate, dimostriamo ora la transitività: se $x \sim y$ e $y \sim z$ allora

$$d(x,z) \leq \max\{d(x,y), d(y,z)\} < r$$

e la transitività è così verificata. ∎

Corollario A.2.3. *Ogni punto di una palla aperta è anche il suo centro ovvero*

$$y \in B(x,r) \Leftrightarrow B(y,r) = B(x,r)$$

Corollario A.2.4. *Gli insiemi $B(x,r)$ sono sia aperti che chiusi, quindi se S ha almeno due elementi distinti allora è sconnesso.*

Corollario A.2.5. *Siano $x, y \in S$ ed $0 < r \leq R$ tali che $B(x,r) \cap B(y,R) \neq \emptyset$ allora $B(x,r) \subseteq B(y,R)$.*

Dimostrazione. Sia $z \in B(x,r) \cap B(y,R)$ quindi $B(x,r) = B(z,r) \subseteq B(z,R) = B(y,R)$. ∎

Corollario A.2.6. *Per ogni $x \in S$ ed $r > 0$ l'insieme $D(x,r)$ è aperto.*

La proposizione A.2.2 e il corollario A.2.5 ci danno molte informazioni su come definire nuove metriche non archimedee su un insieme non vuoto S. Le famiglie di aperti $\mathscr{F}_r = \{B(x,r)\}_{x\in S}$ con $r \in L = \mathbb{R}^+$ soddisfano le seguenti proprietà:

1. $\mathscr{F}_r$ è una partizione di S per ogni $r \in L$;
2. Siano $r, R \in L$ tali che $r < R$ e presi $A \in \mathscr{F}_r$, $B \in \mathscr{F}_R$ allora vale l'implicazione

$$A \cap B \neq \emptyset \Rightarrow A \subseteq B$$

3. Per ogni $x, y \in S$ **distinti** l'insieme

$$\{r \in I \mid \exists A \in \mathscr{F}_r \text{ tale che } \{x,y\} \subseteq A\} \tag{A.2.2}$$

è non banale, ovvero è diverso sia da $\emptyset$ che da L.

Definiamo su $\mathbb{Q}$ una metrica non archimedea molto importante in vari aspetti della teoria dei numeri. Innanzitutto fissiamo un numero primo p e definiamo la funzione

$$v_p : x \in \mathbb{Z} \to \inf\{n \in \mathbb{Z} \mid p^n x \text{ è un numero intero}\} \in \mathbb{Z} \cup \{-\infty\}$$

Chiaramente $v_p(x) \leq 0$, $v_p(x) = -\infty \Leftrightarrow x = 0$, $v_p(-x) = v_p(x)$ e sfruttando risultati basiliari dell'aritmetica $v_p(ab) = v_p(a) + v_p(b)$. Quest'ultima proprietà in particolare ci garantisce che l'applicazione

$$u_p : \frac{a}{b} \in \mathbb{Q} \to v_p(a) - v_p(b) \in \mathbb{Z} \cup \{-\infty\}$$

è ben definita. Dimostriamo ora che per ogni $a, b \in \mathbb{Z}$ si ha

$$v_p(a+b) \leq \max\left\{v_p(a), v_p(b)\right\}$$

Possiamo supporre sia a che b diversi da 0 e $m = v_p(a)$, $n = v_p(b)$. Ora $m \leq n$ implica che $p^n(a+b) = p^{n-m}p^m a + p^n b$ è un numero intero e quindi $v_p(a+b) \leq n$. Se ora $x = \frac{a}{b}$, $y \in \frac{c}{d}$ con $a, b, c, d \in \mathbb{Z} \setminus \{0\}$ si ha

$$\begin{aligned} u_p\left(x+y\right) = v_p(ad+bc) - v_p(b) - v_p(d) &\leq \max\left\{v_p(ad), v_p(bc)\right\} - v_p(b) - v_p(d) \\ &= \max\left\{v_p(a) - v_p(b), v_p(c) - v_p(d)\right\} = \max\left\{u_p(x), u_p(y)\right\} \end{aligned}$$

e la disuguaglianza continua a valere anche per i razionali.

Definiamo allora la "norma" p-adica $|\cdot|_p$ in modo tale che per ogni $x \in \mathbb{Q}$

$$|x|_p = p^{u_p(x)} \in [0, +\infty[$$

che soddisfa le seguenti proprietà:

$$\begin{gathered} |x|_p = 0 \Leftrightarrow x = 0 \\ \left|xy\right|_p = |x|_p \left|y\right|_p \\ |-x| = |x|_p \\ \left|x+y\right|_p \leq \max\left\{|x|_p, \left|y\right|_p\right\} \end{gathered}$$

Abbiamo usato le virgolette quando abbiamo nominato la norma p-adica poiché essa non è affatto una norma su $\mathbb{Q}$ in quanto quest'ultimo non è uno spazio vettoriale reale e perciò non potrebbe essere definito il prodotto per uno scalare. Se presi $x, y \in \mathbb{Q}$ poniamo

$$d(x,y) = \left|x-y\right|_p$$

allora lo spazio $(\mathbb{Q}, d)$ è uno spazio metrico e d è una metrica non archimedea.

Come per i numeri reali possiamo completare $\mathbb{Q}$ rispetto alla metrica p-adica ottenendo un nuovo campo non isomorfo ad $\mathbb{R}$ e con proprietà del tutto peculiari.

Ricaviamoci ora la metrica p-adica tramite le relazioni di equivalenza. Fissato $p \in \mathbb{N}$ numero primo definiamo l'insieme

$$\mathbb{H}_p = \{p^n \mid n \in \mathbb{Z}\} \subseteq \mathbb{R}^+$$

e la famiglia di relazioni di equivalenza $\left\{\equiv_{p^n}\right\}_{p^n \in \mathbb{H}_p}$ tali che per ogni $x, y \in \mathbb{Q}$

$$x \equiv_{p^n} y \Leftrightarrow \text{ esiste } k \in \mathbb{N} \text{ coprimo con } p \text{ tale che } kp^n(x-y) \in \mathbb{Z}$$

Non è difficile dimostrare le proprietà di trasferibilità, separabilità e confrontabilità di questa famiglia di relazioni di equivalenza. Inoltre ogni sezione finale di $\mathbb{H}_p$ non banale possiede minimo quindi $\equiv_i$ è una famiglia non archimedea inferiormente chiusa su ogni punto di $\mathbb{Q}$.

Preso $p^n \in \mathbb{H}_p$ per l'osservazione precedente abbiamo che

$$D\left(0, p^n, d^+\right) = \left\{ \frac{a}{p^k b} \,\middle|\, a, b \in \mathbb{Z} \text{ non divisibili per } p \text{ e } k \leq n \right\}$$

e quindi mantenendo la notazione

$$d^+\left(0, \frac{a}{p^n b}\right) = p^n$$

inoltre poiché le relazioni di equivalenza sono invarianti per traslazioni anche le relative metriche lo sono, quindi per ogni $x \in \mathbb{Q}$

$$d^+\left(x, x + \frac{a}{p^n b}\right) = p^n$$

Per maggiori informazioni sulle norme p-adiche si consulti [9].

A.3 Tetraedri

Dato uno spazio metrico (S, d) e un suo sottoinsieme non vuoto A con $\operatorname{diam} A = l > 0$ osserviamo che $A \subseteq D\left(x, l\right)$ per ogni $x \in A$, in quanto per ogni $x, y, z \in A$ abbiamo chiaramente che $d(x, y), d(x, z) \leq l$ mentre ricordiamo che $\operatorname{diam} D\left(x, l\right) \leq 2l$.

Non è sempre possibile restringere ulteriormente il raggio della palla mantenendo l'inclusione precedente, né in generale si ha $A \subseteq D\left(x, l/2\right)$ e controesempi se ne possono trovare anche in $\mathbb{R}^2$ con la metrica usuale. La minima costante $0 < \gamma \leq 1$ per cui $A \subseteq D\left(x, \gamma \operatorname{diam} A\right)$ per un certo $x \in S$ scelto ad hoc infatti dipende dalle proprietà della metrica scelta ed esula dagli argomenti di queste note, per cui ci limiteremo a calcolarla solamente in $\mathbb{R}^n$.

Innanzitutto mettiamoci in $\mathbb{R}^n$ con la metrica usuale $d(x, y) = \left|x - y\right| = \sqrt{\sum_{i=1}^n \left|x_i - y_i\right|^2}$, vogliamo determinare γ_n al variare di $n \in \mathbb{N}$. In particolare vogliamo dimostrare che gli insiemi che minimizzano γ_n sono proprio i tetraedri definiti nella seguente maniera:

Definizione A.3.1. Per ogni $k \in \mathbb{N}$ e per ogni $x_0, x_1, x_2, \ldots, x_k \in \mathbb{R}^n$ tali che

$$d\left(x_i, x_j\right) = 1 \ \forall i \neq j$$

diremo che l'insieme $T_k = \operatorname{con}\left(x_0, x_1, \ldots, x_k\right)$ è un k-**tetraedro** e i punti x_i saranno i suoi vertici.

Osserviamo innanzitutto che $\operatorname{diam} T_k = 1$, difatti per la proposizione 5.1.5 per ogni scelta di $\lambda_i, \mu_j \in [0, 1]$ con $\sum_{i=0}^k \lambda_i = \sum_{j=0}^k \mu_j = 1$ si ha

$$\left| \sum_{i=0}^k \lambda_i x_i - \sum_{j=0}^k \mu_j x_j \right| = \left| \sum_{i,j} \lambda_i \mu_j \left(x_i - x_j\right) \right| \leq \sum_{i,j} \lambda_i \mu_j \left|x_i - x_j\right| \leq 1$$

senza perdere in generalità infatti possiamo limitarci a considerare solamente gli insiemi con diametro unitario. Se prendiamo ora $k \geq 2$ e tre vertici distinti x, y, z di un tetraedro osserviamo che

$$1 = \left|x - y\right|^2 = |x - z|^2 + \left|z - y\right|^2 + 2(x - z) \cdot (z - y) = 2 + 2(x - z) \cdot (z - y)$$

e dunque

$$(x - z) \cdot (y - z) = 1/2 \tag{A.3.1}$$

In altre parola i vettori di lunghezza unitaria $x - z$ ed $y - z$ formano un angolo di 60° poiché $\cos(\pi/3) = 1/2$.

Definiamo ora il centro di T_k come

$$\tilde{T}_k = \frac{x_0 + x_1 + \cdots + x_k}{k + 1}$$

che sappiamo appartenere a T_k, la (A.3.1) ci permette di affermare così che quando $k \geq 2$

$$\begin{aligned}
\left|x_i - \tilde{T}_k\right|^2 = \frac{1}{(k+1)^2}\left|\sum_{j=0}^{k}\left(x_i - x_j\right)\right|^2 &= \frac{k}{(k+1)^2} + \frac{2}{(k+1)^2}\sum_{0 \leq a < b \leq k}(x_i - x_a) \cdot (x_i - x_b) \\
&= \frac{k}{(k+1)^2} + \frac{1}{(k+1)^2}\frac{k(k-1)}{2} \\
&= \frac{k}{2(k+1)}
\end{aligned}$$

e il risultato continua a valere anche quando $k = 1$. Per come abbiamo definito T_k abbiamo poi che $T_k \subseteq D\left(\tilde{T}_k, \gamma_k\right)$ con

$$\gamma_k = \sqrt{\frac{k}{2(k+1)}}$$

Il prossimo passo è quello di dimostrare che ogni sottoinsieme C di $\mathbb{R}^n$ con diametro unitario è contenuto in $D\left(\tilde{x}, \gamma_n\right)$ per un qualche $\tilde{x} \in \mathbb{R}^n$. Innanzitutto osserviamo che $\operatorname{diam} C = \operatorname{diam} \overline{\operatorname{con}}(C)$ e perciò possiamo supporre che C sia chiuso e convesso. Prima di proseguire abbiamo bisogno del teorema di Radon per la geometria discreta:

Teorema A.3.2 (Radon)**.** *Dato un sottoinsieme finito I di $\mathbb{R}^n$ se $\#(I) > n + 1$ allora esistono $J, K \subseteq I$ non vuoti tali che*

$$J \cap K = \emptyset \qquad J \cup K = I \qquad \operatorname{con}(J) \cap \operatorname{con}(K) \neq \emptyset$$

Dimostrazione. Prendiamo $a_1, a_2, \dots, a_{n+2} \in I \subseteq \mathbb{R}^n$ distinti, per comodità di notazione indichiamo con $e_1, e_2, \dots, e_{n+1}$ la base standard di $\mathbb{R}^{n+1}$ e poniamo $\mathbb{R}^n \leq \mathbb{R}^{n+1} \leq \mathbb{R}^{n+2}$ nella maniera classica, ovvero vedendo i vettori di $\mathbb{R}^n$ come vettori di $\mathbb{R}^{n+1}$ con l'ultima coordinata nulla. Possiamo così definire un'applicazione lineare $L : \mathbb{R}^{n+2} \to \mathbb{R}^{n+1}$ in modo tale che

$$L(e_i) = a_i + e_{n+1}$$

questa applicazione chiaramente non può essere iniettiva, dunque esisterà $b \in \mathbb{R}^{n+2}$ non nullo tale che $0 = L(b) = \sum_i (b \cdot e_i) a_i + e_{n+1} \sum_i b \cdot e_i$ e quindi

$$\sum_{i=1}^{n+2} (b \cdot e_i) a_i = 0$$

$$\sum_{i=1}^{n+2} b \cdot e_i = 0$$

Definiamo così

$$J = \{a_i \mid b \cdot e_i > 0\}$$
$$K = \{a_i \mid b \cdot e_i \leq 0\}$$

chiaramente $J \cap K = \emptyset$ e nessuno di essi è vuoto. Infatti $K \neq \emptyset$ sempre mentre se fosse $J = \emptyset$ dovremmo avere $b \cdot e_i = 0$ per ogni i ovvero $b = 0$ assurdo. Ma allora $l = \sum_{a_i \in J} b \cdot e_i = -\sum_{a_i \in K} b \cdot e_i > 0$ e perciò sempre per la caratterizzazione dell'inviluppo convesso

$$\sum_{a_i \in J} \frac{b \cdot e_i}{l} = \sum_{a_i \in K} \frac{-b \cdot e_i}{l} \in \operatorname{con}(J) \cap \operatorname{con}(K) \neq \emptyset$$

Aggiungendo i punti rimanenti di I a J o a K avremo infine che $I = J \cup K$ e il teorema è così verificato. ■

Questo teorema innanzitutto ci fornisce un criterio di esistenza per k-tetraedri:

Proposizione A.3.3. *Esiste un k-tetraedro in $\mathbb{R}^n$ se e solo se $k \leq n$.*

Dimostrazione. Per assurdo esiste un $(n+1)$-tetraedro in $\mathbb{R}^n$ formato da $n+2$ vertici, il teorema di Radon ci permette di dividere i vertici in due gruppi disgiunti $x_1, x_2, \dots, x_l$ ed $y_1, y_2, \dots, y_{n+2-l}$ con

$$\lambda_i, \mu_j \geq 0 \qquad \sum_{i=1}^{l} \lambda_i = \sum_{j=1}^{n+2-l} \mu_j = 1$$
$$\sum_{i=1}^{l} \lambda_i x_i = \sum_{j=1}^{n+2-l} \mu_j y_j$$

e quindi esisteranno due indici α, β per cui $\lambda_\alpha \mu_\beta > 0$. Sottraendo x_α ad entrambi i membri e moltiplicando scalarmente per $(y_\beta - x_\alpha)$ poiché $x_i \neq y_j$ avremo che

$$\frac{1}{2}(1-\lambda_\alpha) = \sum_{i=1}^{l} (x_i - x_\alpha) \cdot (y_\beta - x_\alpha) = \sum_{j=1}^{n+2-l} (y_j - x_\alpha) \cdot (y_\beta - x_\alpha) = \frac{1}{2}(1-\mu_\beta) + \mu_\beta$$

e quindi $-\lambda_\alpha/2 = \mu_\beta/2$. Ciò è possibile solo se $\lambda_\alpha = \mu_\beta = 0$ il che va contro le nostre ipotesi, perciò un tale tetraedro non può esistere.

Viceversa se dimostriamo che in $\mathbb{R}^n$ esiste sempre un n-tetraedro allora ogni sottoinsieme di k vertici distinti sono a loro volta i vertici di un k-tetraedro, e quindi l'implicazione opposta è dimostrata. Se $n = 1$ il segmento $T_1 = [0,1]$ è anche un 1-tetraedro, quindi procediamo per induzione supponendo che per qualche $n \geq 1$ esiste un n-tetraedro in $\mathbb{R}^n$ di vertici $x_0, x_1, \dots, x_n$ con centro $\tilde{x} \in \mathbb{R}^n$.

Poniamo sempre $\mathbb{R}^n \leq \mathbb{R}^{n+1}$ e per ogni $t > 0$ definiamo

$$y_t = \tilde{x} + t e_{n+1}$$

allora per ogni vertice x_i si ha

$$|x_i - y_t|^2 = |x_i - \tilde{x}|^2 + t^2 = \gamma_k^2 + t^2$$

che non dipende da i, quindi possiamo scegliete t in modo tale che $x_0, x_1, \dots, x_k, y$ siano i vertici di un $(n+1)$-tetraedro in $\mathbb{R}^{n+1}$. ■

Il prossimo risultato di geometria discreta riguarda invece l'intersezione di insiemi convessi.

Teorema A.3.4 (Helly)**.** *Siano $C_1, C_2, \dots, C_k$ sottoinsiemi convessi di $\mathbb{R}^n$ con $k > n+1$ per cui l'intersezione di al più $n+1$ di essi è non vuota. Allora l'intersezione di tutti questi sottoinsiemi è anch'essa non vuota.*

Dimostrazione. Lavoriamo per induzione supponendo inizialmente che $k = n+2$. Per ogni i intero tra 1 ed $n+2$ sia x_i un elemento di $\bigcap_{j\neq i} C_j$ che per ipotesi sappiamo non essere vuoto. Per il teorema di Radon dunque possiamo partizionare i punti in due classi disgiunte A e B i cui inviluppi convessi si intersecano.

Per comodità siano $x_1, x_2, \dots, x_l$ gli elementi di A e $x_{l+1}, \dots, x_{n+2}$ gli elementi di B, ma allora avremo che $A \subseteq \bigcap_{i=1}^{l} C_i$ e $B \subseteq \bigcap_{j=l+1}^{n+2} C_j$ e così per convessità

$$\bigcap_{i=1}^{n+2} \supseteq \operatorname{con}(A) \cap \operatorname{con}(B) \neq \emptyset$$

e il caso base è dimostrato.

Poniamo ora $k > n+2$ e supponiamo per ipotesi induttiva che il teorema sia valido per $k-1$, il passo induttivo implica così che l'intersezione di al più $k-1$ dei k insiemi di partenza ha intersezione non vuota, perciò applicando ancora il teorema di Radon e ragionando come per il passo base abbiamo ancora che l'intersezione di tutti e k gli insiemi è non vuota. ∎

Possiamo allora dimostrare il risultato principale. Procediamo però per gradi:

Lemma A.3.5. *Dato un sottoinsieme finito I di $\mathbb{R}^n$ con diametro unitario e $\#(I) \leq n+1$ allora esiste $\tilde{x} \in \mathbb{R}^n$ tale che $I \subseteq D(\tilde{x}, \gamma_n)$.*

Dimostrazione. Poniamo innanzitutto

$$r = \inf\left\{t > 0 \,\middle|\, I \subseteq D(y,t) \text{ per qualche } y \in \mathbb{R}^n\right\}$$

poiché I è limitato per compattezza l'estremo inferiore è raggiunto da qualche $\tilde{r} > 0$ e quindi $I \subseteq D(\tilde{x}, \tilde{r})$. Poiché I è finito esisteranno $x_0, x_1, \dots, x_l \in I$ distinti con $l \geq 0$ tali che $|x_i - \tilde{x}| = \tilde{r}$, vogliamo far vedere che $\tilde{r} \leq \gamma_n$.

Se $\tilde{x} \notin \operatorname{con}(I)$ per il secondo teorema di Mazur possono essere separati da un iperpiano, ovvero esistono $z, y_0 \in \mathbb{R}^n$ con z di norma unitaria tali che

$$\tilde{x} \cdot z > y_0 \cdot z > x_i \cdot z$$

per ogni i. Sostituendo $\tilde{x}$ con $\tilde{x} - [(\tilde{x} - y_0) \cdot z]\, z$ avremo per ogni i

$$\begin{aligned}
\left|x_i - \tilde{x} + [(\tilde{x} - y_0)\cdot z]\, z\right|^2 &= r^2 + [(\tilde{x} - y_0)\cdot z]^2 + 2\,[(\tilde{x} - y_0)\cdot z]\,(x_i - \tilde{x})\cdot z \\
&= r^2 + [(\tilde{x} - y_0)\cdot z]\,[(2x_i - \tilde{x} - y_0)\cdot z] \\
&< r^2
\end{aligned}$$

che va contro la nostra ipotesi di minimalità, perciò possiamo scrivere $\tilde{x} = \sum_i \lambda_i x_i$ e inoltre $1 \geq |x_i - x_j|^2 = 2\tilde{r}^2 - 2(x_i - \tilde{x})\cdot(x_j - \tilde{x})$. Per ogni indice i avremo così che

$$1 - \lambda_i \geq \sum_j \lambda_j |x_i - x_j|^2 = 2r^2 - (x_j - \tilde{x}) \cdot \left(\sum_j \lambda_j x_j - \tilde{x}\right) = 2r^2$$

e sommando stavolta in i avremo che $n \geq 2(n+1)r^2$ ovvero $r \leq \gamma_n$ e quindi $I \subseteq D(\tilde{x}, \gamma_n)$. ∎

Lemma A.3.6. *Dato un sottoinsieme finito I di $\mathbb{R}^n$ con diametro unitario esiste $\tilde{x} \in \mathbb{R}^n$ tale che $I \subseteq D(\tilde{x}, \gamma_n)$.*

Dimostrazione. Possiamo supporre tranquillamente che $\#(I) > n+1$ e siano $x_0, x_1, \dots, x_l$ gli elementi di I. Posto $C_i = D(x_i, \gamma_n)$ per il lemma precedente avremo che l'intersezione di al più $n+1$ dei C_i è non vuota, possiamo quindi applicare il teorema di Helly in modo tale da avere $\tilde{x} \in \bigcap_i D(x_i, \gamma_n)$ ovvero $I \subseteq D(\tilde{x}, \gamma_n)$. ∎

Teorema A.3.7. *Dato un sottoinsieme generico C di $\mathbb{R}^n$ con diametro unitario esiste $\tilde{x} \in \mathbb{R}^n$ tale che $C \subseteq D(\tilde{x}, \gamma_n)$.*

Dimostrazione. Poiché il diametro di C è limitato possiamo supporre che C sia contenuto in una palla chiusa di raggio finito, consideriamo dunque una famiglia di insiemi chiusi

$$\mathscr{C} = \{D(x, \gamma_n) \mid x \in C\}$$

che formano un ricoprimento di C. Il lemma precedente implica poi che ogni sottofamiglia finita di $\mathscr{C}$ ha intersezione non vuota in $\mathbb{R}^n$, mentre poiché in $\mathbb{R}^n$ le palle chiuse sono compatte per la proposizione 2.4.2 l'intersezione di tutti gli elementi della famiglia $\mathscr{C}$ è non vuota, in particolare per ogni $\tilde{x}$ nell'intersezione avremo che $C \subseteq D(\tilde{x}, \gamma_n)$. ∎

Bibliografia

[1] Charalambos D. Aliprantis e Kim C. Border. *Infinite dimensional analysis*. Springer, 2006.

[2] Tom M. Apostol. *Mathematical Analysis: A Modern Approach to Advanced Calculus*. Pearson, 1974.

[3] Vladimir I. Bogachev. *Gaussian measures*. Mathematical Surveys. American Mathematical Society, 1991.

[4] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer.

[5] L Danzer, B Grünbaum e V Klee. «Helly's Theorem and its Relatives». In: *Proceedings of Symposia in Pure Mathematics* 7 (1963), pp. 101–180.

[6] Camillo De Lellis e Emanuele Nunzio Spadaro. «Q-Valued Functions Revisited». In: *Memoirs of the American Mathematical Society* 211.991 (2011).

[7] J Diestel e JJ Uhl Jr. *Vector measures*. Mathematical Surveys. American Mathematical Society, 1977.

[8] Lawrence C. Evans e Ronald F. Gariepy. *Measure Theory and Fine properties of Functions*. Studies in Advanced Mathematics. CRC Press, 1992.

[9] Neal Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*. Vol. 58. New York: Springer Verlag, 1977.

[10] Marco Manetti. *Topologia Generale*. Springer, 2014.

[11] K. R. Parthasarathy. *Probability measures on metric spaces*. New York: Academic Press Inc., 1967.

[12] Tyrrell R. Rockafellar. *Convex Analysis*. Princeton: Princeton University Press, 1972.

[13] Walter Rudin. *Functional Analysis*. New York: McGraw-Hill, 1991.

[14] Walter Rudin. *Real and Complex analysis*. Mathematics series. New York: McGraw-Hill, 1987.

[15] Jussi Väisälä. «A proof of the Mazur-Ulam theorem». In: *The American mathematical monthly* 110.7 (2003), pp. 633–635.

Indice analitico