

Istituzioni di Analisi Superiore,
secondo modulo

Gianluca Gorni
Università di Udine
Anno Accademico 2021-22

13 giugno 2022

Indice

1	Spazi di Hilbert	5
1.1	Prodotto scalare hermitiano	5
1.2	Identità di polarizzazione e identità del parallelogrammo . . .	16
1.3	Il teorema della minima distanza	19
1.4	Riformulazioni della condizione di minima distanza	24
1.5	Ortogonalità	31
1.6	Sottospazi a dimensione finita	37
1.7	Somme infinite di vettori	44
1.8	Sistemi ortonormali infiniti	56
1.9	Basi hilbertiane	64
1.10	Basi hilbertiane esplicite	68
2	Insiemi compatti	71
2.1	Compatti in dimensione infinita	71
2.2	Infinite estrazioni	76
2.3	Il teorema di Ascoli-Arzelà	80
2.4	Il cubo di Hilbert	82
3	Operatori compatti e aggiunti	85
3.1	Operatori compatti	85
3.2	Duale ed operatori aggiunti	90
3.3	Aggiunti e compattezza	95
3.4	Operatori a rango finito	98
4	Compatti e aggiunti hilbertiani	103
4.1	Duale e biduale di uno spazio di Hilbert	103
4.2	Aggiunto hilbertiano	107
4.3	Convergenza debole e compattezza	110

5	Teoria spettrale	115
5.1	Autovalori, autovettori e autospazi	115
5.2	Autospazi di operatori compatti	116
5.3	Operatori autoaggiunti	117
5.4	Operatori autoaggiunti compatti	119
5.5	L'alternativa di Fredholm	127
6	Il problema di Sturm-Liouville	137
6.1	Operatore nucleo integrale	137
6.2	L'operatore differenziale di Sturm-Liouville	142
6.3	Variazione delle costanti	145
6.4	Problemi ai limiti omogenei	151
	6.4.1 Il caso non degenere	153
	6.4.2 Il caso degenere	160
7	La trasformata di Fourier	165
7.1	Convoluzione	165
7.2	La trasformata di Fourier in L^1	168
7.3	Le trasformate di Fourier in L^1 sono continue	175
7.4	Una famiglia di mollificatori	179
7.5	Convergenza delle mollificate	184
7.6	Il teorema di inversione	188
7.7	Il teorema di Plancherel	190

Capitolo 1

Spazi di Hilbert

Notazioni: dato uno spazio vettoriale V e un suo sottinsieme $E \subseteq V$, indicheremo con $\text{span } E$ il sottospazio vettoriale generato da E , cioè l'insieme delle combinazioni lineari finite di elementi di E .

Dato un numero complesso α , coi simboli $\bar{\alpha}$, $\Re\alpha$ e $\Im\alpha$ indicheremo rispettivamente il coniugio, la parte reale e la parte immaginaria di α .

La chiusura di un sottinsieme E di uno spazio topologico sarà indicata con $\text{cl } E$.

Dato uno spazio metrico (X, d) , un sottinsieme $E \subset X$ e un punto $x \in X$, la distanza fra x ed E è $\text{dist}(x, E) := \inf_{y \in E} d(x, y)$.

Se X è un insieme, l'operatore identità su X verrà scritto come id_X .

1.1 Prodotto scalare hermitiano

Diamo gli assiomi del prodotto scalare per spazi complessi. Per gli spazi su \mathbb{R} basterà ignorare i coniugi.

Definizione 1.1.1. Dato uno spazio vettoriale V su \mathbb{C} , un'applicazione $(x, y) \mapsto \langle x, y \rangle$ da $V \times V$ in \mathbb{C} si dice **prodotto scalare** (hermitiano) se valgono gli assiomi seguenti:

- a) $\langle y, x \rangle = \overline{\langle x, y \rangle} \quad \forall x, y \in V$ (hermitianità);
- b) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall x, y, z \in V$ (additività, o distributività, rispetto al primo fattore);
- c) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall x, y \in V \quad \forall \alpha \in \mathbb{C}$ (omogeneità rispetto al primo fattore);
- d) $\langle x, x \rangle \geq 0 \quad \forall x \in V$ (positività semidefinita sulla diagonale);

e) se $\langle x, x \rangle = 0$ allora $x = 0 \in V$ (non degenerazione sulla diagonale).

Non c'è standardizzazione sul nome e la notazione del prodotto scalare. Molti autori chiamano prodotto “interno” quello che qui chiamiamo prodotto scalare. Al posto di $\langle x, y \rangle$ si incontrano altre notazioni, come $\langle x | y \rangle$, (x, y) , $(x | y)$, $x \cdot y$. La notazione $\langle x, y \rangle$ viene talvolta usata col significato di valutazione di un funzionale lineare y sul vettore x .

Alcune semplici conseguenze degli assiomi:

- dall'omogeneità segue che $\langle 0, y \rangle = 0$ per ogni $y \in V$; in particolare $\langle x, x \rangle = 0 \iff x = 0$;
- l'omogeneità e l'additività insieme equivalgono a dire che l'applicazione $x \mapsto \langle x, y \rangle$ è un funzionale lineare $V \rightarrow \mathbb{C}$ per ogni $y \in V$ fissato;
- dall'hermitianità e dall'omogeneità segue che $\langle x, \alpha y \rangle = \bar{\alpha} \langle x, y \rangle$. In particolare, il prodotto scalare in uno spazio complesso *non* è lineare rispetto al secondo fattore; se $\alpha \in \mathbb{R}$ allora $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$; in particolare $\langle x, -y \rangle = -\langle x, \alpha y \rangle$;
- dall'hermitianità e dall'additività rispetto al primo fattore, segue l'additività rispetto al secondo fattore: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

Il prodotto scalare negli spazi complessi non è bilineare, ma è talvolta chiamato “sesquilineare”, cioè lineare una volta e mezza.

Definizione 1.1.2. Dato uno spazio vettoriale V su \mathbb{C} dotato di un prodotto scalare $\langle x, y \rangle$, diremo **norma** del vettore $x \in V$ (indotta dal prodotto scalare) la radice quadrata non negativa di $\langle x, x \rangle$, cioè, in simboli

$$\|x\| := \sqrt{\langle x, x \rangle}. \quad (1.1)$$

Dagli assiomi è chiaro che $\|x\| \geq 0$ e che $\|x\| = 0$ solo quando $x = 0$.

Non è ancora chiaro che questo oggetto $\|x\|$ meriti il nome di norma nel senso degli spazi normati, di cui rammentiamo gli assiomi:

Definizione 1.1.3. Dato uno spazio vettoriale V su \mathbb{C} , chiameremo *norma* su V una funzione $x \mapsto \|x\|$, $V \rightarrow \mathbb{R}$, tale che

- $\|x\| \geq 0 \forall x \in V$ (nonnegatività);

- $\|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall x \in V \quad \forall \alpha \in \mathbb{C}$ (omogeneità assoluta);
- $\|x\| = 0 \iff x = 0$ (non degenerazione);
- $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$ (subadditività, o disuguaglianza triangolare).

Se questi assiomi sono verificati, la coppia $(V, \|\cdot\|)$ è detta *spazio normato*.

Comunque si può dedurre subito dagli assiomi che per la funzione definita come $\|x\| := \sqrt{\langle x, x \rangle}$ valgono la nonnegatività, l'*omogeneità assoluta*:

$$\begin{aligned} \|\alpha x\| &= \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \langle x, \alpha x \rangle} = \sqrt{\alpha \bar{\alpha} \langle x, x \rangle} = \sqrt{|\alpha|^2 \langle x, x \rangle} = \\ &= |\alpha| \sqrt{\langle x, x \rangle} = |\alpha| \cdot \|x\|, \end{aligned}$$

la non degenerazione, nonché le formule per il “quadrato del binomio”:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \\ &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 = \\ &= \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2, \\ \|x - y\|^2 &= \langle x - y, x - y \rangle = \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle = \\ &= \|x\|^2 - \langle x, y \rangle - \overline{\langle x, y \rangle} + \|y\|^2 = \\ &= \|x\|^2 - 2\Re\langle x, y \rangle + \|y\|^2. \end{aligned}$$

Notare i “doppi prodotti” con le parti reali. Si usa il fatto che $\alpha \bar{\alpha} = |\alpha|^2$ e $\alpha + \bar{\alpha} = 2\Re\alpha$ per ogni $\alpha \in \mathbb{C}$.

Più in generale, il quadrato di una sommatoria su un $I = \{1, 2, \dots, n\}$ si può espandere così:

$$\begin{aligned} \left\| \sum_{k \in I} x_k \right\|^2 &= \left\langle \sum_{k \in I} x_k, \sum_{j \in I} x_j \right\rangle = \sum_{k, j \in I} \langle x_k, x_j \rangle = \\ &= \sum_{k \in I} \|x_k\|^2 + \sum_{\substack{k, j \in I \\ k \neq j}} \langle x_k, x_j \rangle = \\ &= \sum_{k \in I} \|x_k\|^2 + \sum_{\substack{k, j \in I \\ k < j}} (\langle x_k, x_j \rangle + \langle x_j, x_k \rangle) = \end{aligned}$$

$$= \sum_{k \in I} \|x_k\|^2 + 2 \sum_{\substack{k, j \in I \\ k < j}} \Re \langle x_k, x_j \rangle.$$

Nel caso speciale in cui i vettori x_k siano a due a due ortogonali, cioè $\langle x_k, x_j \rangle = 0$ quando $k \neq j$, la somma si semplifica:

$$\left\| \sum_{k \in I} x_k \right\|^2 = \sum_{k \in I} \|x_k\|^2. \quad (1.2)$$

Algebricamente, la norma quadra “si distribuisce” su una somma di vettori a due a due ortogonali, quasi come se fosse lineare. La formula (1.2) potrebbe anche essere chiamata *Teorema di Pitagora* per un parallelepipedo rettangolo in dimensione n : il quadrato costruito sulla diagonale principale è uguale alla somma dei quadrati costruiti sui lati che si incontrano in uno stesso vertice (Fig. 1.1 e Fig.1.2).

Dagli assiomi segue in modo meno ovvio una ben nota disuguaglianza:

Teorema 1.1.4 (Disuguaglianza di Schwarz). *In uno spazio con prodotto scalare vale sempre la disuguaglianza $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$.*

Dimostrazione. Di solito la disuguaglianza viene dimostrata con un conto puramente algebrico. La dimostrazione che propongo qui parte dal seguente vettore z , che ha il vantaggio di avere un'interpretazione geometrica abbastanza chiara (Fig. 1.3):

$$z := x - \left\langle x, \frac{y}{\|y\|} \right\rangle \frac{y}{\|y\|}. \quad (1.3)$$

Imponiamo la non negatività di $\langle z, z \rangle$ e applichiamo l'omogeneità assoluta e il quadrato del binomio:

$$\begin{aligned} 0 \leq \|z\|^2 &= \left\| x - \left\langle x, \frac{y}{\|y\|} \right\rangle \frac{y}{\|y\|} \right\|^2 = \\ &= \left\| x - \langle x, y \rangle \frac{y}{\|y\|^2} \right\|^2 = \\ &= \|x\|^2 - 2\Re \left\langle x, \langle x, y \rangle \frac{y}{\|y\|^2} \right\rangle + \left\| \langle x, y \rangle \frac{y}{\|y\|^2} \right\|^2 = \\ &= \|x\|^2 - 2\Re \left(\frac{\overline{\langle x, y \rangle}}{\|y\|^2} \langle x, y \rangle \right) + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \|y\|^2 = \end{aligned} \quad (1.4)$$

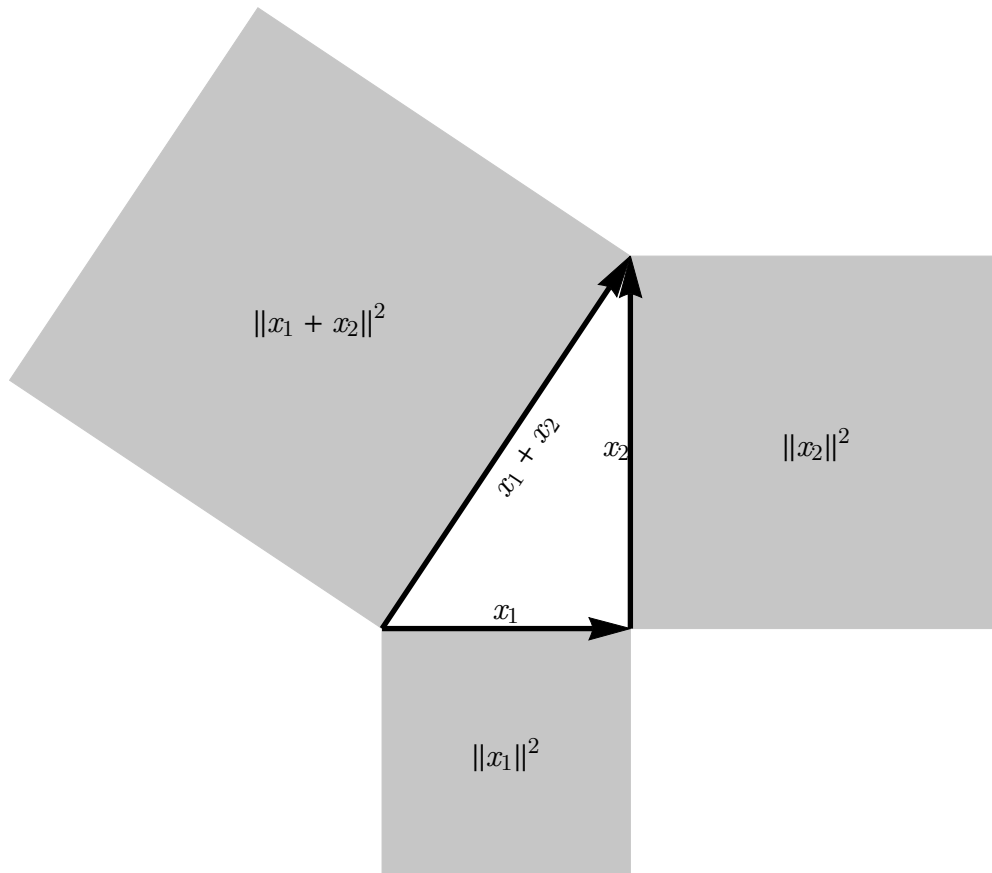


Figura 1.1: L'identità $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2$, valida quando x_1 e x_2 sono ortogonali in uno spazio con prodotto scalare, si può interpretare come il teorema di Pitagora della geometria euclidea piana

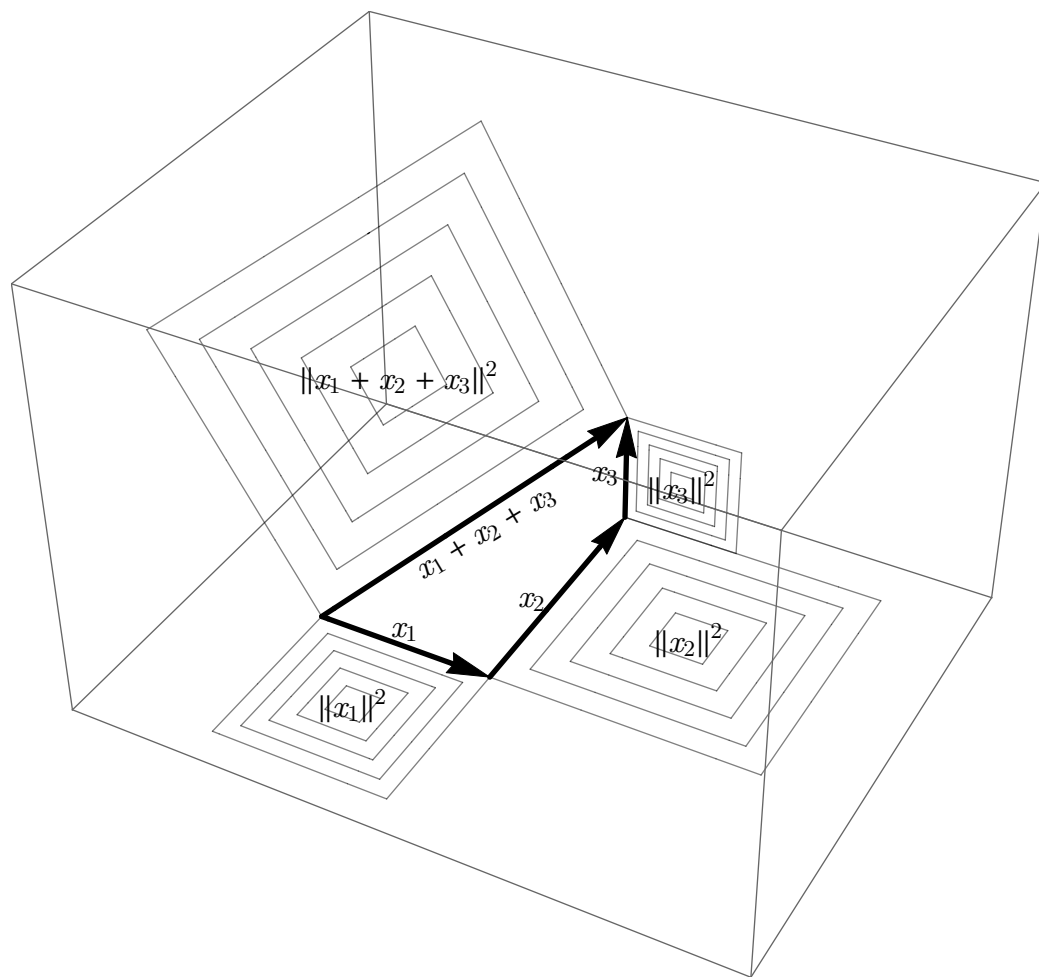


Figura 1.2: L'identità $\|x_1 + x_2 + x_3\|^2 = \|x_1\|^2 + \|x_2\|^2 + \|x_3\|^2$, valida quando x_1, x_2, x_3 sono a due a due ortogonali, si può interpretare come una generalizzazione tridimensionale del teorema di Pitagora classico. Volendo illustrare l'analoga formula (1.2) per una somma di quattro o più vettori, ci vorrebbe uno spazio di dimensione 4 o più

$$\begin{aligned}
&= \|x\|^2 - 2\Re\left(\frac{|\langle x, y \rangle|^2}{\|y\|^2}\right) + \frac{|\langle x, y \rangle|^2}{\|y\|^2} = \\
&= \|x\|^2 - 2\frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^2} = \\
&= \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}. \tag{1.5}
\end{aligned}$$

Moltiplicando il primo e l'ultimo membro per $\|y\|^2$ e riordinando:

$$|\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2, \tag{1.6}$$

Prendendo infine la radice quadrata si ottiene la disuguaglianza di Schwarz. Il conto precedente ha senso solo se $y \neq 0$, ma il caso con $y = 0$ è banale.

Se fossimo partiti con $\|y\|^2 z$ al posto di z non ci sarebbe stato bisogno di frazioni e di trattare il caso $y = 0$ a parte. Però il significato geometrico di $\|y\|^2 z$ è più oscuro. \square

Esercizio 1.1.5. Trovare una semplice condizione su x, y che equivalga all'*uguaglianza* $|\langle x, y \rangle| = \|x\| \cdot \|y\|$.

Con la disuguaglianza triangolare si completano gli assiomi della norma (Def. 1.1.3 p. 6):

Teorema 1.1.6 (Disuguaglianza triangolare). *In uno spazio V con prodotto scalare vale la disuguaglianza $\|x + y\| \leq \|x\| + \|y\|$ per ogni $x, y \in V$.*

Dimostrazione. Usando la formula del quadrato del binomio, dal fatto che $\Re \alpha \leq |\alpha|$ per ogni $\alpha \in \mathbb{C}$, e dalla disuguaglianza di Schwarz segue che

$$\begin{aligned}
\|x + y\|^2 &= \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \leq \\
&\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.
\end{aligned}$$

Estraendo la radice quadrata aritmetica del primo e dell'ultimo membro la disuguaglianza si conserva. \square

Il prodotto scalare induce su V la struttura di spazio normato, e quindi metrico e topologico. Qualora come spazio metrico risulti *completo*, cioè se ogni successione di Cauchy converge, allora si parla di *spazio di Hilbert*. A volte uno spazio con prodotto scalare non necessariamente completo viene

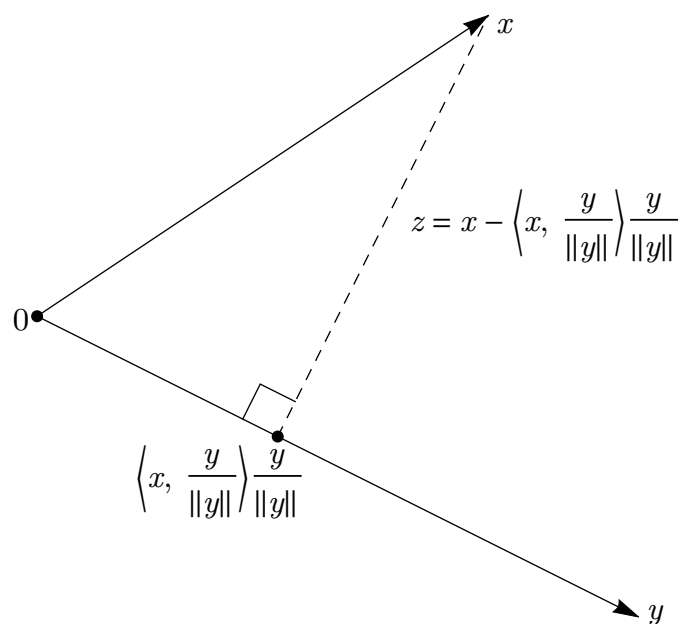


Figura 1.3: Interpretazione geometrica dei vettori che compaiono nella dimostrazione della disuguaglianza di Schwarz. $\hat{y} = y/\|y\|$ è il versore di y . Il prodotto scalare $\langle x, \hat{y} \rangle$ è la lunghezza (orientata) della componente di x in direzione di y . La proiezione di x sulla direzione di y è $\langle x, \hat{y} \rangle \hat{y}$. Infine si prende il vettore z che congiunge tale proiezione con x . Il valore $\|z\|^2$ che si ricava per via algebrica nelle formule (1.4) e (1.5) si può dedurre da questa figura tramite il teorema di Pitagora

detto pre-hilbertiano. È un fatto generale che uno spazio normato o metrico non completo si può sempre “completare”, cioè si può costruire uno spazio completo in modo tale che lo spazio iniziale sia isomorfo a un sottospazio denso del nuovo spazio completo.

Esempio 1.1.7. Dovrebbe essere familiare lo spazio \mathbb{C}^n delle ennuple di numeri complessi con il prodotto scalare usuale:

$$\langle (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \rangle := x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n = \sum_{k=1}^n x_k \bar{y}_k.$$

Nel caso dello spazio euclideo \mathbb{R}^n basta ignorare i coniugi. La norma associata al prodotto scalare è la norma euclidea:

$$\|(x_1, x_2, \dots, x_n)\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

(attenzione che in \mathbb{R}^n si possono ignorare i valori assoluti nella formula, ma in \mathbb{C}^n no). È ben noto che \mathbb{C}^n e \mathbb{R}^n sono completi come spazi metrici con tale norma.

Esempio 1.1.8. Dato uno spazio di misura positiva (X, \mathcal{M}, μ) , poniamo $V := L^2(\mu)$ e definiamo

$$\langle f, g \rangle := \int_X f \cdot \bar{g} d\mu \quad \text{per } f, g \in L^2(\mu).$$

Nel caso che lo spazio V sia fatto di sole funzioni a valori reali basta ignorare i coniugi. L'integrale è ben definito ed è un numero complesso finito perché $f \cdot \bar{g} \in L^1(\mu)$ dalla disuguaglianza di Hölder:

$$\int_X |f \cdot \bar{g}| d\mu = \int_X |f| \cdot |g| d\mu \leq \|f\|_{L^2} \|g\|_{L^2} < +\infty.$$

La norma associata è la nota norma di L^2 :

$$\|f\|_{L^2} := \left(\int_X |f|^2 d\mu \right)^{1/2}.$$

La disuguaglianza di Schwarz equivale al caso particolare $p = q = 2$ della *disuguaglianza di Hölder*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q} \quad \text{se } 1 < p, q < +\infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

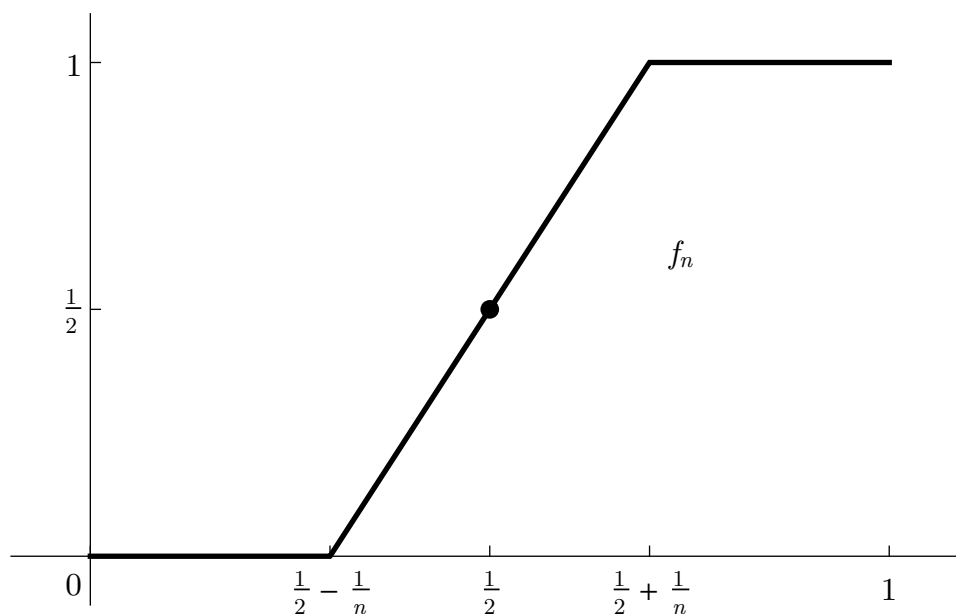


Figura 1.4: Fissato n consideriamo la funzione f_n lineare a tratti (in particolare, continua) fatta come in figura. Si può dimostrare che la successione f_n è di Cauchy per la norma di L^2 ma non esiste nessuna f continua tale che $\|f_n - f\|_2 \rightarrow 0$

È noto che $L^2(\mu)$ è completo come spazio normato, come tutti gli $L^p(\mu)$. Quindi $L^2(\mu)$ ha anche la struttura di spazio di Hilbert, oltre che di spazio normato.

Esempio 1.1.9. Sia V lo spazio vettoriale delle funzioni continue $[0, 1] \rightarrow \mathbb{C}$, dotato del prodotto scalare

$$\langle f, g \rangle := \int_0^1 f(t) \cdot \overline{g(t)} dt$$

che è ben definito come integrale elementare di Riemann. Questo spazio è pre-hilbertiano ma non è difficile (Fig. 1.4) trovare una successione di Cauchy in V che non converge (in V).

L'esempio \mathbb{C}^n può essere visto come un caso particolare di uno spazio L^2 . Infatti un'ennupla (x_1, \dots, x_n) si può interpretare come una funzione x con dominio l'insieme finito $X = \{1, 2, \dots, n\}$ e codominio \mathbb{C} . L'insieme X si

può dotare della sigma algebra dell'insieme delle parti, e della *misura # del conteggio* (“counting measure”), dove

$$\#(A) := \begin{cases} \text{cardinalità di } A & \text{se } A \text{ è finito} \\ +\infty & \text{se } A \text{ è infinito.} \end{cases} \quad (1.7)$$

Con queste convenzioni, dati $x, y: X \rightarrow \mathbb{C}$ si può scrivere

$$\begin{aligned} \langle x, y \rangle &= \langle (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \rangle = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n = \\ &= \int_X x \cdot \bar{y} d\#. \end{aligned}$$

Vedremo più avanti che, se accettiamo l'assioma di scelta, *ogni* spazio di Hilbert, anche a dimensione infinita, è isomorfo a un qualche spazio $L^2(X, \#)$ per un qualche insieme X , anche infinito. Questo particolare tipo di spazio L^2 viene talvolta indicato col simbolo $\ell^2(X)$. In un certo senso, la teoria degli spazi di Hilbert è equivalente alla teoria degli spazi ℓ^2 .

Proposizione 1.1.10. *In uno spazio con prodotto scalare le funzioni a valori scalari*

$$x \mapsto \langle x, y \rangle, \quad x \mapsto \langle y, x \rangle, \quad x \mapsto \|x\|$$

sono continue (per la topologia indotta dalla norma), per $y \in V$ fissato.

Dimostrazione. Dalla disuguaglianza di Schwarz segue che

$$\left| \langle x_1, y \rangle - \langle x_2, y \rangle \right| = \left| \langle x_1 - x_2, y \rangle \right| \leq \|x_1 - x_2\| \cdot \|y\|,$$

da cui segue facilmente che l'applicazione $x \mapsto \langle x, y \rangle$ è (uniformemente) continua su V (anzi, lipschitziana). Analogamente, oppure usando il coniugio, si vede che anche $x \mapsto \langle y, x \rangle$ è continua. Per dimostrare che la norma è continua, si può partire dalla disuguaglianza triangolare $\|x_1\| \leq \|x_1 - x_2\| + \|x_2\|$, da cui segue che

$$\|x_1\| - \|x_2\| \leq \|x_1 - x_2\|.$$

Invertendo i ruoli fra x_1 e x_2 si ottiene che

$$\|x_2\| - \|x_1\| \leq \|x_2 - x_1\|.$$

Combinando le due disuguaglianze e notando che i primi membri sono opposti mentre i secondi membri sono uguali otteniamo l'utile disuguaglianza

$$\left| \|x_1\| - \|x_2\| \right| \leq \|x_1 - x_2\|, \quad (1.8)$$

da cui segue in particolare che la norma è (uniformemente, anzi, lipschitzianamente) continua su V . \square

Esercizio 1.1.11. Dimostrare che l'applicazione (forma quadratica) $(x, y) \mapsto \langle x, y \rangle$ è uniformemente continua sui limitati di $V \times V$.

1.2 Identità di polarizzazione e identità del parallelogrammo

Abbiamo visto che da un prodotto scalare si può indurre una norma. Si può fare anche il viceversa: conoscendo la norma indotta, si può ricostruire il prodotto scalare, grazie alla cosiddetta *identità di polarizzazione*:

$$\langle x, y \rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4} + i \frac{\|ix + y\|^2 - \|ix - y\|^2}{4}, \quad (1.9)$$

dove i è l'unità immaginaria. Nel caso di uno spazio reale, la formula vale ignorando la parte immaginaria. La verifica è un semplice conticino che usa le proprietà algebriche del prodotto scalare (non servono quelle di positività). Svolgiamolo nel solo caso di uno spazio reale:

$$\begin{aligned} \frac{\|x + y\|^2 - \|x - y\|^2}{4} &= \frac{\|x\|^2 + 2\langle x, y \rangle + \|y\|^2 - (\|x\|^2 - 2\langle x, y \rangle + \|y\|^2)}{4} = \\ &= \frac{2\langle x, y \rangle + 2\langle x, y \rangle}{4} = \\ &= \langle x, y \rangle. \end{aligned}$$

Ci si può chiedere come mai l'identità è chiamata "di polarizzazione". Forse questo ha a che fare col fatto che tramite l'identità si può calcolare l'angolo fra due vettori, e quindi si possano introdurre le coordinate "polari". Si può trovare una discussione sulla storia del nome al link [4].

Esercizio 1.2.1. In \mathbb{R}^n col prodotto scalare canonico, l'angolo compreso fra due vettori x, y è dato da $\arccos \langle x, y \rangle$.

Potrebbe venire quest'idea: se partiamo da una qualsiasi norma, se la inseriamo nel secondo membro dell'identità di polarizzazione ne ricaviamo un prodotto scalare che induce quella norma. Attenzione: l'oggetto $\langle x, y \rangle$ che si ottiene è di sicuro una funzione da $V \times V$ in \mathbb{C} , ma chi ci dice che sia un prodotto scalare?

1.2 Identità di polarizzazione e identità del parallelogrammo 17

C'è un'identità che coinvolge soltanto la norma, ma nella cui dimostrazione si usa il prodotto scalare: si chiama *identità del parallelogrammo*:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \quad (1.10)$$

Il motivo del nome diventa ovvio guardando la Figura 1.5. Usando la formula del quadrato del binomio i “doppi prodotti” semplicemente si cancellano:

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \|x\|^2 + \|y\|^2 + 2\Re\langle x, y \rangle + \|x\|^2 - 2\Re\langle x, y \rangle + \|y\|^2 = \\ &= 2\|x\|^2 + 2\|y\|^2. \end{aligned}$$

L'identità del parallelogrammo vale quindi in tutti gli spazi con prodotto scalare. Però si vede rapidamente che *non* vale in altri semplici spazi normati. Prendete per esempio $V = \mathbb{R}^2$ con la norma $\|(x_1, x_2)\|_1 := |x_1| + |x_2|$, oppure $\|(x_1, x_2)\|_\infty := \max\{|x_1|, |x_2|\}$, e x, y la base canonica di \mathbb{R}^2 .

Un fatto interessante è stato determinato da Jordan e von Neumann nel 1935: una norma è indotta da un prodotto scalare se e solo se per essa vale l'identità del parallelogrammo. Consiglio la lettura dell'articolo originale [3]. Qui riporto brevemente l'avvio della dimostrazione. Supponiamo di avere una norma $\|x\|$ su uno spazio reale (per semplicità). Poniamo, come nell'identità di polarizzazione,

$$\langle x, y \rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4}. \quad (1.11)$$

Definiamo anche la seguente funzione di due variabili:

$$\varphi(x, y) = \|x + y\|^2 + \|x - y\|^2 - 2\|x\|^2 - 2\|y\|^2. \quad (1.12)$$

Allora valgono le identità seguenti:

$$4\left(\langle 2x, y \rangle - 2\langle x, y \rangle\right) = \varphi(x + y, x) - \varphi(x - y, x) \quad (1.13)$$

e

$$\begin{aligned} 4\left(\langle x, z \rangle + \langle y, z \rangle - 2\left\langle \frac{x + y}{2}, z \right\rangle\right) &= \\ &= \varphi\left(\frac{x + y}{2} + z, \frac{x - y}{2}\right) - \varphi\left(\frac{x + y}{2} - z, \frac{x - y}{2}\right). \end{aligned} \quad (1.14)$$

Queste identità si verificano con un po' di pazienza usando semplicemente le formule (1.11) e (1.12) qui sopra e l'omogeneità assoluta della norma

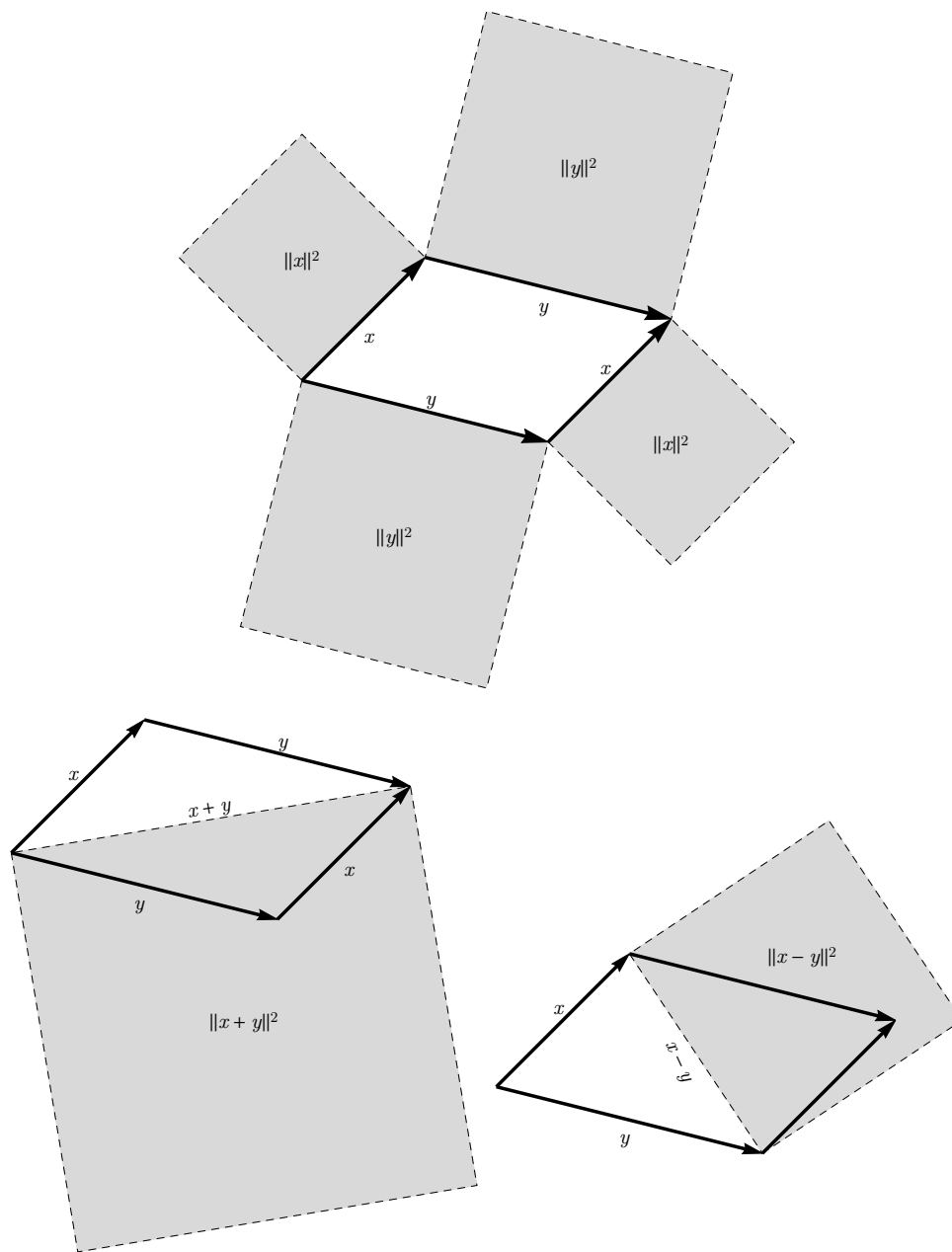


Figura 1.5: Identità del parallelogrammo: la somma dei quadrati costruiti sui quattro lati è uguale alla somma dei quadrati costruiti sulle due diagonali

$\|\alpha x\| = |\alpha| \cdot \|x\|$. Attenzione a *non* usare gli assiomi del prodotto scalare! Quelli sono nella tesi, non nell'ipotesi. Supponiamo ora che valga l'identità del parallelogrammo. La funzione φ diventa identicamente nulla, e quindi il secondo membro della formula (1.13) è zero. Ne deduciamo un caso particolare dell'omogeneità del prodotto scalare, limitata al fattore 2:

$$\langle 2x, y \rangle = 2\langle x, y \rangle. \quad (1.15)$$

D'altra parte, se vale l'identità del parallelogrammo, si annulla il secondo membro anche della formula (1.14), per cui

$$\langle x, z \rangle + \langle y, z \rangle = 2\left\langle \frac{x+y}{2}, z \right\rangle. \quad (1.16)$$

Grazie alla formula (1.15) possiamo far entrare il fattore 2 dentro al prodotto scalare al secondo membro della (1.16) e ottenere l'additività:

$$\langle x, z \rangle + \langle y, z \rangle = \langle x+y, z \rangle.$$

Gli interessati possono proseguire, anche aiutati dall'articolo originale.

1.3 Il teorema della minima distanza

Un *convesso* in uno spazio vettoriale è un sottinsieme C tale che per ogni $x, y \in C$ e per ogni $t \in [0, 1]$ si ha che anche $tx + (1-t)y \in C$. Notare che anche se lo spazio è complesso, il parametro t viene sempre fatto variare nell'intervallo $[0, 1] \subset \mathbb{R}$.

Il seguente è il teorema che sta alla base della teoria degli spazi di Hilbert. Il significato geometrico dell'enunciato è semplice, come si vede in Figura 1.7.

Teorema 1.3.1 (della minima distanza). *Dati uno spazio di Hilbert H , un sottinsieme C convesso, chiuso e non vuoto, e un punto $x \in H$, esiste un unico $\bar{y} \in C$ che ha la minima distanza da x fra tutti i punti di C , cioè tale che*

$$\|x - \bar{y}\| \leq \|x - y\| \quad \forall y \in C,$$

o, equivalentemente,

$$\|x - \bar{y}\| = \inf_{y \in C} \|x - y\|.$$

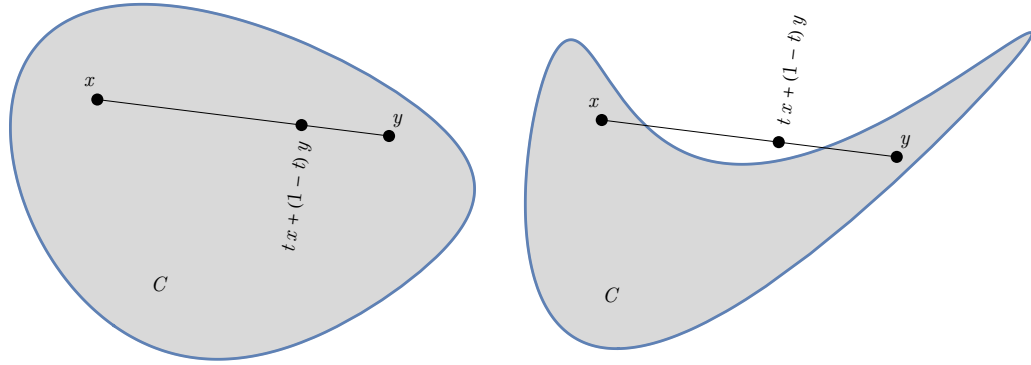


Figura 1.6: Un insieme convesso e uno non convesso. Nell'insieme convesso ogni segmento che abbia gli estremi nell'insieme è tutto contenuto nell'insieme. Un insieme non è convesso se esiste almeno un segmento che ha gli estremi nell'insieme ma non è interamente contenuto nell'insieme. Il segmento di estremi x, y è parametrizzato da $tx + (1-t)y$ al variare del parametro $t \in [0, 1]$. Per $t = 0$ siamo in y , per $t = 1$ siamo in x , per $t = 1/2$ siamo nel punto medio del segmento, ecc. L'espressione $tx + (1-t)y$ può essere vista come una media pesata dei due vettori x, y , oppure come il baricentro di una distribuzione di massa in cui x ha massa t e y ha massa $1-t$

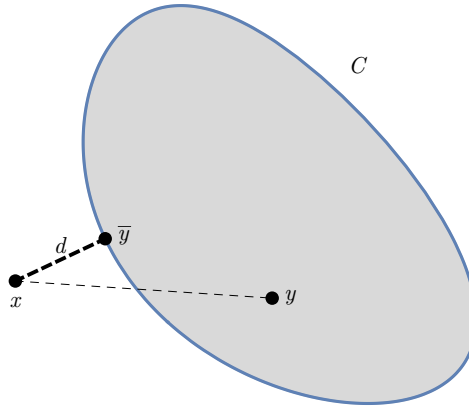


Figura 1.7: L'enunciato del teorema della minima distanza concerne un insieme chiuso e convesso C , un punto x dello spazio (magari non in C , se no tutto si banalizza), il punto $\bar{y} \in C$ che ha minima distanza da x , e un punto generico $y \in C$, la cui distanza d da x viene confrontata con quella di \bar{y}

Dimostrazione. Poniamo per cominciare

$$d := \inf_{y \in C} \|x - y\| = \text{dist}(x, C).$$

Certamente $0 \leq d < +\infty$ perché C è non vuoto e la norma è sempre ≥ 0 .

Stabiliamo alcune formule. Prendiamo due punti generici $y_1, y_2 \in C$ e consideriamo il parallelogrammo di vertici x, y_1, y_2 e $x + (y_1 - x) + (y_2 - x) = y_1 + y_2 - x$, come in Figura 1.8. Applichiamo l'identità del parallelogrammo:

$$\|(y_1 + y_2 - x) - x\|^2 + \|y_1 - y_2\|^2 = 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2. \quad (1.17)$$

A noi interessa la distanza fra y_1 e y_2 , e quindi isoliamo quel termine:

$$\|y_1 - y_2\|^2 = 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - \|y_1 + y_2 - 2x\|^2.$$

Raccogliamo un fattore 2 dentro l'ultimo addendo e poi portiamolo fuori dalla norma quadrata:

$$\begin{aligned} \|y_1 - y_2\|^2 &= 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - \left\| 2\left(\frac{y_1 + y_2}{2} - x\right) \right\|^2 = \\ &= 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4\left\| \frac{y_1 + y_2}{2} - x \right\|^2. \end{aligned} \quad (1.18)$$

Lo scopo della manipolazione precedente è di evidenziare il termine $(y_1 + y_2)/2$, che è il punto medio del segmento che congiunge y_1 con y_2 . Dato che C è convesso, anche il punto medio appartiene a C . La quantità

$$\left\| \frac{y_1 + y_2}{2} - x \right\|$$

è la distanza del punto medio da x . Quindi per definizione di d , tale distanza è $\geq d$. Possiamo dedurre da (1.18) che

$$\|y_1 - y_2\|^2 \leq 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4d^2 \quad \forall y_1, y_2 \in C. \quad (1.19)$$

Già ora possiamo stabilire l'unicità del punto di minima distanza (qualora ci sia): se y_1, y_2 sono entrambi di minima distanza, allora $\|y_1 - x\| = \|y_2 - x\| = d$, e quindi, sostituendo in (1.19),

$$\|y_1 - y_2\|^2 \leq 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4d^2 = 2d^2 + 2d^2 - 4d^2 = 0,$$

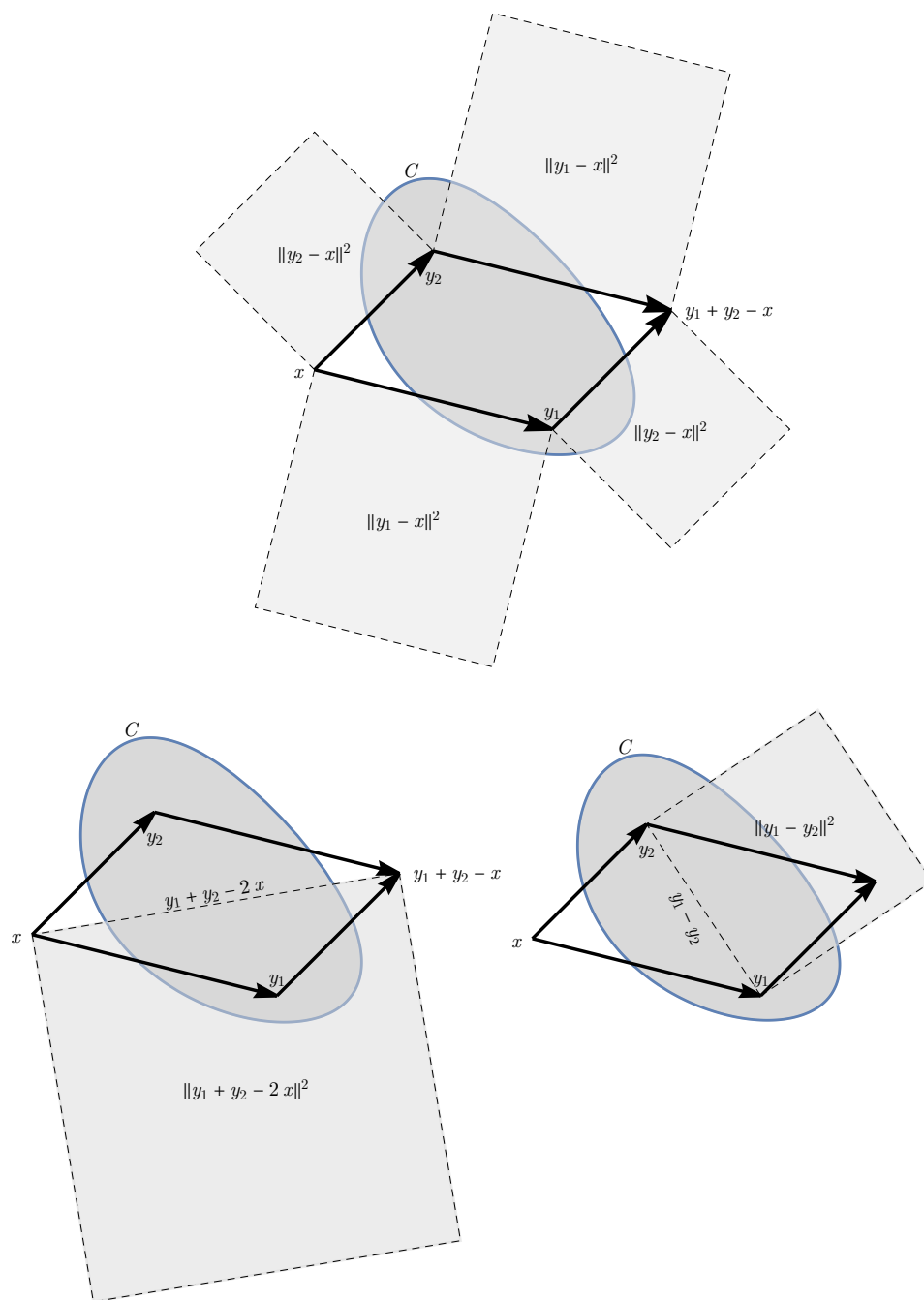


Figura 1.8: Interpretazione geometrica delle quantità che compaiono nella formula (1.17)

da cui $y_1 = y_2$.

Passiamo all'esistenza. Per definizione di estremo inferiore, esiste una successione $y_n \in C$ che è minimizzante, cioè tale che $\|y_n - x\| \rightarrow d^+$ per $n \rightarrow +\infty$. Intendiamo mostrare che y_n è di Cauchy. Fissiamo $\varepsilon > 0$. Poiché $\|y_n - x\|^2 \rightarrow (d^2)^+$, esiste un indice $N_\varepsilon \in \mathbb{N}$ tale che

$$\|y_n - x\|^2 < d^2 + \frac{\varepsilon^2}{4} \quad \forall n \geq N_\varepsilon.$$

Applichiamo la disuguaglianza (1.19) con y_n, y_m al posto di y_1, y_2 , per $n, m \geq N_\varepsilon$:

$$\begin{aligned} \|y_n - y_m\|^2 &\leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2 < \\ &< 2\left(d^2 + \frac{\varepsilon^2}{4}\right) + 2\left(d^2 + \frac{\varepsilon^2}{4}\right) - 4d^2 = \\ &= \varepsilon^2, \end{aligned}$$

da cui, estraendo le radici quadrate,

$$\|y_n - y_m\| < \varepsilon \quad \forall n, m \geq N_\varepsilon.$$

Effettivamente la successione y_n è di Cauchy. Essendo H completo, esiste $\bar{y} \in H$ tale che $y_n \rightarrow \bar{y}$. Dato che $y_n \in C$ per ogni n e che C è chiuso, anche il limite \bar{y} è in C . Infine, la successione $\|y_n - x\|$ converge da una parte a d , dall'altra a $\|\bar{y} - x\|$, perché la norma è continua. Per l'unicità del limite concludiamo che

$$\|\bar{y} - x\| = d,$$

cioè che \bar{y} è di minima distanza. \square

Osservazione 1.3.2. Strettamente parlando, per l'esistenza e unicità del punto di minima distanza, non è necessario che tutto H sia completo, ma *basta che il convesso C sia completo.*

Una successione minimizzante della distanza è automaticamente limitata. Se lo spazio ha dimensione finita la successione ha una sottosuccessione convergente, da cui si deduce l'esistenza del punto di minima distanza, senza bisogno di usare né la convessità né l'identità del parallelogrammo, essendo sufficiente che C sia chiuso. L'interesse del teorema è tutto *in dimensione infinita*, dove le successioni limitate possono benissimo non avere alcuna sottosuccessione convergente.

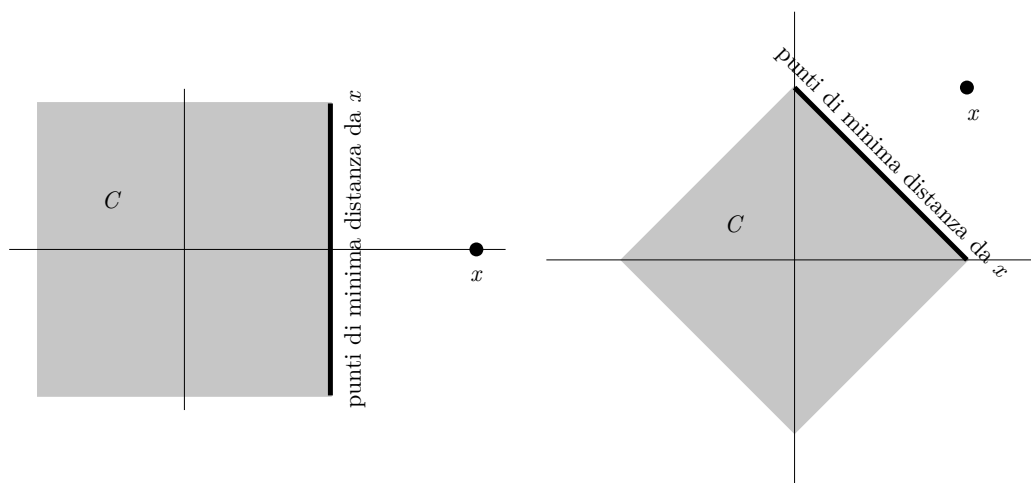


Figura 1.9: Se dotiamo \mathbb{R}^2 della norma $\|\cdot\|_\infty$ o della norma $\|\cdot\|_1$ (definite a pag. 17), la palla unitaria chiusa C di centro l'origine e raggio 1 è un quadrato, coi lati paralleli agli assi (figura a sinistra) o le diagonali parallele agli assi (figura a destra). Per certe posizioni del punto $x \in \mathbb{R}^2$, un intero lato del quadrato è formato da punti di C di minima distanza da x . Viene quindi a mancare l'unicità del punto di minima distanza.

Quando la norma non deriva da un prodotto scalare, può venire a mancare l'unicità del punto di minima distanza, anche in dimensione finita. Per due esempi semplicissimi si veda la Fig. 1.9. L'esistenza può mancare se lo spazio non è completo (esercizio 1.5.9 p. 35).

Esercizio 1.3.3. Dimostrare che la funzione che associa al vettore x la sua distanza d da C è continua. A questo scopo è importante che C sia convesso?

1.4 Riformulazioni della condizione di minima distanza

La condizione che $\bar{y} \in C$ sia di minima distanza

$$\|x - \bar{y}\| \leq \|x - y\| \quad \forall y \in C, \quad (1.20)$$

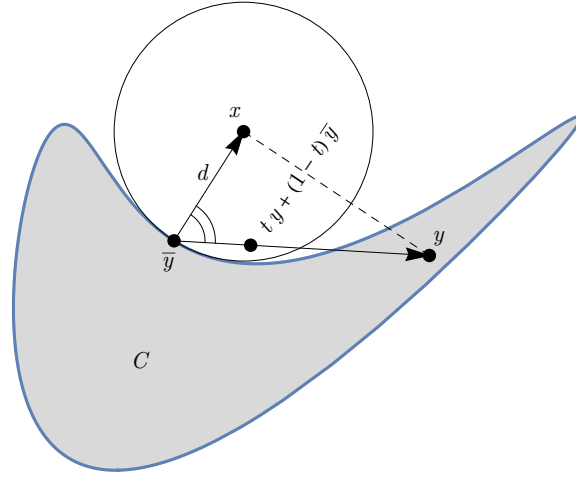


Figura 1.10: La condizione che \bar{y} sia di minima distanza da x in un insieme C non convesso ha perfettamente senso e si interpreta geometricamente dicendo che C giace all'esterno del disco (aperto) di centro x e raggio $d = \|x - \bar{y}\|$. Se però prendiamo un generico punto $y \in C$, il punto $ty + (1-t)\bar{y}$ del segmento che congiunge y con \bar{y} può non appartenere a C , e non si può ragionare come nella proposizione 1.4.1

coinvolge solo la norma e non traspare alcun ruolo per la convessità di C , come si vede dalla figura 1.10. La condizione si può riformulare tirando in ballo in modo essenziale il prodotto scalare e la convessità nel modo seguente.

Proposizione 1.4.1. *Siano V uno spazio con prodotto scalare, C un convesso in V ed $x \in V$, $\bar{y} \in C$. Allora la condizione che \bar{y} sia il punto di C più vicino a x , cioè*

$$\|x - \bar{y}\| \leq \|x - y\| \quad \forall y \in C, \quad (1.21)$$

equivale a dire che

$$\Re \langle x - \bar{y}, y - \bar{y} \rangle \leq 0 \quad \forall y \in C. \quad (1.22)$$

Dimostrazione. Per ogni $y \in C$, tutto il segmento che congiunge \bar{y} con y è contenuto in C . Quindi (1.21) equivale a

$$\|x - \bar{y}\| \leq \left\| x - (ty + (1-t)\bar{y}) \right\| \quad \forall y \in C, \forall t \in [0, 1] \quad (1.23)$$

(Fig. 1.11), cioè, raccogliendo t ,

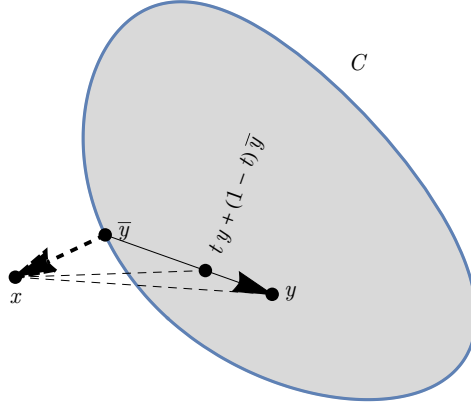


Figura 1.11: La condizione che \bar{y} sia di minima distanza da x in un convesso si può scrivere coinvolgendo un generico punto $y \in C$ e un punto $ty + (1-t)\bar{y}$ del segmento che congiunge y con \bar{y}

$$\|x - \bar{y}\| \leq \|x - \bar{y} - t(y - \bar{y})\| \quad \forall y \in C, \forall t \in [0, 1]. \quad (1.24)$$

Eleviamo al quadrato ambo i membri, e applichiamo il quadrato del binomio al secondo membro:

$$\|x - \bar{y}\|^2 \leq \|x - \bar{y}\|^2 + \|t(y - \bar{y})\|^2 - 2\Re\langle x - \bar{y}, t(y - \bar{y}) \rangle \quad \forall y \in C, \forall t \in [0, 1]. \quad (1.25)$$

Semplificando e riordinando:

$$2\Re\langle x - \bar{y}, t(y - \bar{y}) \rangle \leq \|t(y - \bar{y})\|^2 \quad \forall y \in C, \forall t \in [0, 1]. \quad (1.26)$$

Portando fuori i coefficienti reali t ,

$$2t\Re\langle x - \bar{y}, y - \bar{y} \rangle \leq t^2\|y - \bar{y}\|^2 \quad \forall y \in C, \forall t \in [0, 1]. \quad (1.27)$$

Notare che la (1.27) è *equivalente* alla 1.21. Quando $t > 0$ possiamo dividere ambo i membri per t , ottenendo

$$2\Re\langle x - \bar{y}, y - \bar{y} \rangle \leq t\|y - \bar{y}\|^2 \quad \forall y \in C, \forall t \in]0, 1]. \quad (1.28)$$

Dalla (1.28), passando al limite per $t \rightarrow 0^+$, segue che

$$2\Re\langle x - \bar{y}, y - \bar{y} \rangle \leq 0 \quad \forall y \in C. \quad (1.29)$$

Questa disuguaglianza, a parte il fattore 2, è la (1.22). Viceversa, dalla (1.29) segue chiaramente la (1.27). \square

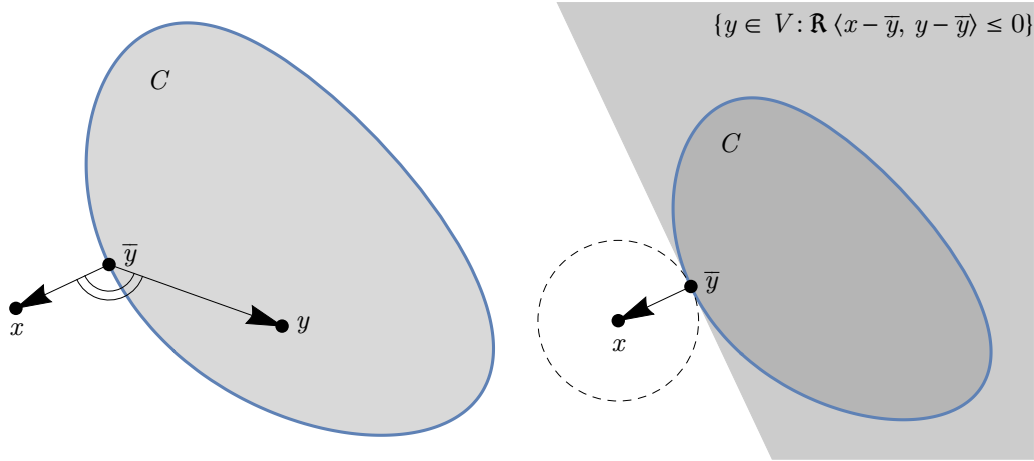


Figura 1.12: Interpretazioni geometriche della condizione che \bar{y} sia di minima distanza da x fra tutti gli y di un convesso C . A sinistra la formulazione secondo cui l'angolo $\widehat{x\bar{y}y}$ è ottuso, in quanto il prodotto scalare di due vettori è uguale al prodotto delle lunghezze per il coseno dell'angolo compreso. A destra quella secondo cui C è contenuto nel semispazio affine il cui bordo passa per \bar{y} ed è perpendicolare a $x - \bar{y}$, ed è tangente all'ipersfera di centro x e raggio $\|x - \bar{y}\|$

La condizione di minima distanza su un convesso si può riscrivere in termini più geometrici (Fig. 1.12), qualora x non appartenga al convesso e si abbia un'interpretazione del prodotto scalare. Naturalmente il tutto diventa più tangibile quando lo spazio è su \mathbb{R} .

Corollario 1.4.2. *Siano V uno spazio con prodotto scalare, C un convesso in V ed $x \in V \setminus C$. Allora la condizione che \bar{y} sia il punto di C più vicino a x equivale a dire che*

$$\text{per ogni } y \in C \text{ l'angolo } \widehat{x\bar{y}y} \text{ è ottuso } (\geq 90^\circ), \quad (1.30)$$

o anche che

$$C \text{ è contenuto nel semispazio affine } \{y \in V : \Re\langle x - \bar{y}, y - \bar{y} \rangle \leq 0\}. \quad (1.31)$$

Il bordo del semispazio passa per \bar{y} ed è ortogonale alla congiungente di x con \bar{y} , e tangente all'ipersfera di centro x e raggio $\|\bar{y} - x\|$.

Esercizio 1.4.3. Dimostrare che l'applicazione che associa al vettore x il punto \bar{y} di C più vicino a x è continua. A tal fine è cruciale che C sia convesso?

A noi nel séguito interesserà esclusivamente il caso particolare in cui il convesso C è in realtà un *sottospazio vettoriale* di V , come in Figura 1.13 e 1.14.

Proposizione 1.4.4. *Siano V uno spazio con prodotto scalare, M un sottospazio vettoriale di V ed $x \in H$, $\bar{y} \in M$. Allora la condizione che \bar{y} sia il punto di M più vicino a x equivale a dire che*

$$\langle x - \bar{y}, y \rangle = 0 \quad \forall y \in M. \quad (1.32)$$

Dimostrazione. Un sottospazio vettoriale è automaticamente anche un convesso. Quindi possiamo partire dalla formulazione (1.22):

$$\Re \langle x - \bar{y}, y - \bar{y} \rangle \leq 0 \quad \forall y \in M. \quad (1.33)$$

È ovvio che la condizione (1.32) implica la (1.33), perché $y - \bar{y} \in M$ per ogni $y \in M$. Resta da dimostrare il viceversa.

Se $y \in M$, allora anche $\bar{y} + ty \in M$ per ogni $t \in \mathbb{R}$. Quindi da (1.33) segue

$$\Re \langle x - \bar{y}, (\bar{y} + ty) - \bar{y} \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.34)$$

Semplificando

$$\Re \langle x - \bar{y}, ty \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.35)$$

Il coefficiente reale t si può portare fuori dal prodotto scalare e anche dalla parte reale:

$$t \cdot \Re \langle x - \bar{y}, y \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.36)$$

Ponendo successivamente $t = 1$ e $t = -1$ otteniamo

$$\Re \langle x - \bar{y}, y \rangle \leq 0 \text{ e } -\Re \langle x - \bar{y}, y \rangle \leq 0 \quad \forall y \in M, \quad (1.37)$$

da cui segue che

$$\Re \langle x - \bar{y}, y \rangle = 0 \quad \forall y \in M. \quad (1.38)$$

Nel caso di spazio reale questo già basterebbe. Se V è complesso, osserviamo che anche $\bar{y} + ity \in M$ per ogni $t \in \mathbb{R}$, e quindi da (1.33) segue

$$\Re \langle x - \bar{y}, (\bar{y} + ity) - \bar{y} \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.39)$$

Semplificando

$$\Re \langle x - \bar{y}, ity \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.40)$$

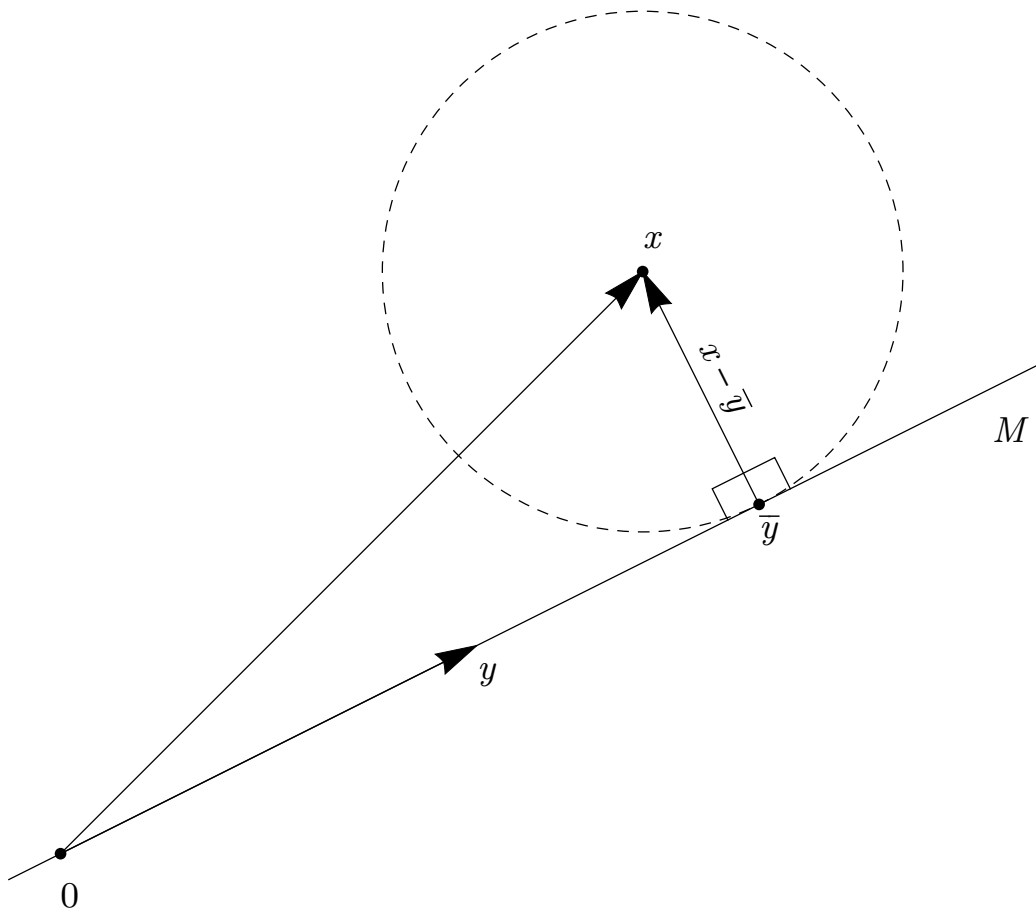


Figura 1.13: Interpretazione geometrica della condizione che \bar{y} sia di minima distanza da x fra tutti gli y di un sottospazio vettoriale M di dimensione 1. Qui V è il piano e M è una retta. La retta è ortogonale alla congiungente fra x e il punto di minima distanza. La retta è anche tangente al cerchio di centro x e avente per raggio la minima distanza

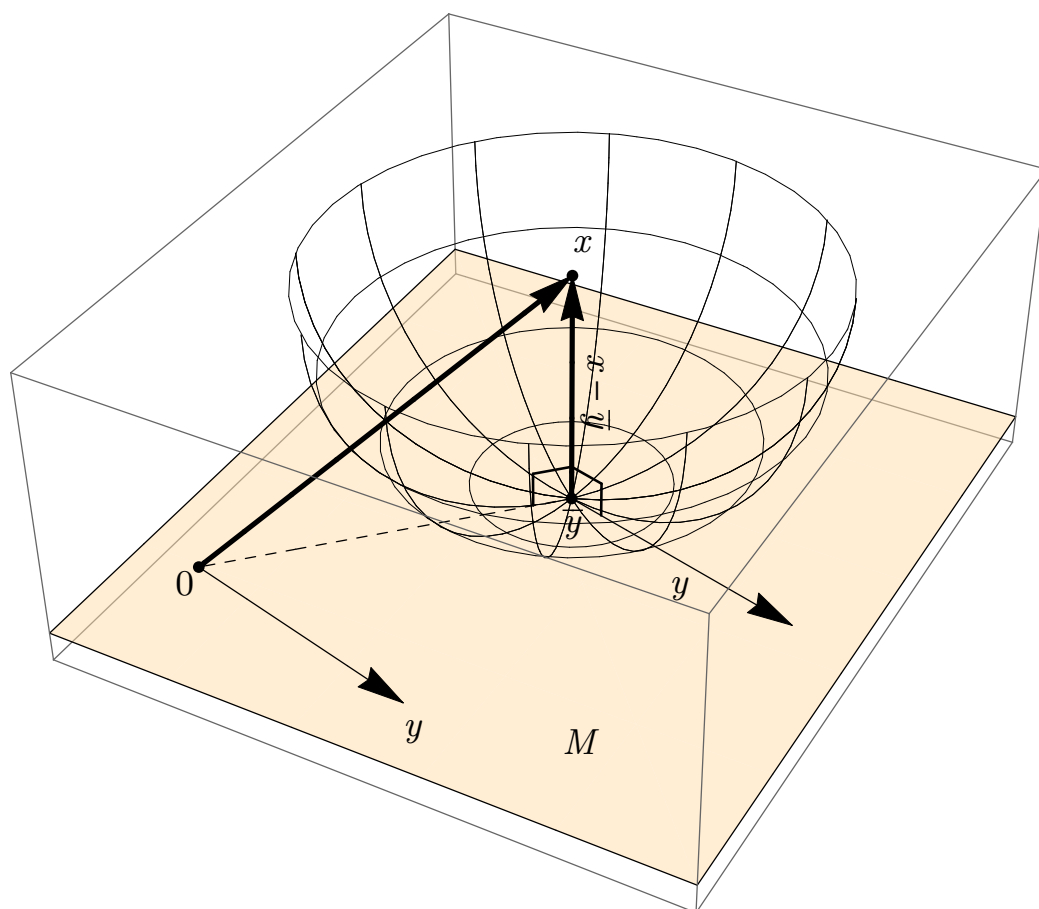


Figura 1.14: Interpretazione geometrica della condizione che \bar{y} sia di minima distanza da x fra tutti gli y di un sottospazio vettoriale M di dimensione 2. Qui V è lo spazio tridimensionale e M è un piano. Il piano è ortogonale alla congiungente fra x e il punto di minima distanza. Il piano è anche tangente alla sfera di centro x e avente per raggio la minima distanza

Il coefficiente immaginario puro it esce dal prodotto scalare coniugato:

$$\Re(-it\langle x - \bar{y}, y \rangle) \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.41)$$

Il coefficiente reale $-t$ esce anche dalla parte reale:

$$-t \cdot \Re(i\langle x - \bar{y}, y \rangle) \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.42)$$

Ora usiamo il fatto che nei numeri complessi $i(a + ib) = ia - b$, e quindi $\Re(i\alpha) = -\Im\alpha$:

$$t \cdot \Im\langle x - \bar{y}, y \rangle \leq 0 \quad \forall y \in M \quad \forall t \in \mathbb{R}. \quad (1.43)$$

Ponendo successivamente $t = 1$ e $t = -1$ deduciamo che

$$\Im\langle x - \bar{y}, y \rangle \leq 0 \text{ e } -\Im\langle x - \bar{y}, y \rangle \leq 0 \quad \forall y \in M, \quad (1.44)$$

da cui deduciamo che

$$\Im\langle x - \bar{y}, y \rangle = 0 \quad \forall y \in M, \quad (1.45)$$

che, insieme alla (1.38) dà

$$\langle x - \bar{y}, y \rangle = 0 \quad \forall y \in M, \quad (1.46)$$

come desiderato. \square

1.5 Ortogonalità

La proprietà di minima distanza da un sottospazio vettoriale è strettamente legata al concetto di ortogonalità.

Definizione 1.5.1. In uno spazio V con prodotto scalare, due vettori x, y si dicono (fra loro) **ortogonali** se $\langle x, y \rangle = 0$, e si scrive in tal caso $x \perp y$. L'insieme dei vettori ortogonali a x si indica con

$$x^\perp := \{y \in V : x \perp y\} = \{y \in V : \langle x, y \rangle = 0\},$$

e, più in generale, l'insieme dei vettori ortogonali a ogni elemento del sottinsieme $E \subseteq V$ è detto *l'ortogonale di E* e si indica con

$$\begin{aligned} E^\perp &:= \{y \in V : x \perp y \ \forall x \in E\} = \{y \in V : \langle x, y \rangle = 0 \ \forall x \in E\} = \\ &= \bigcap_{x \in E} x^\perp. \end{aligned}$$

Esercizio 1.5.2. Sia V uno spazio con prodotto scalare sul campo \mathbb{R} . Dimostrare che $x \perp y$ equivale a $\|x + y\| = \|x - y\|$, una condizione che coinvolge solo la norma. E se lo spazio è complesso?

Proposizione 1.5.3. *Sia V uno spazio con prodotto scalare, $x, y \in V$, $E, F \subseteq V$. Allora x^\perp ed E^\perp sono sempre sottospazi vettoriali chiusi di V . Inoltre:*

$$x \perp y \iff y \perp x, \quad E \subseteq F \Rightarrow E^\perp \supseteq F^\perp.$$

Dimostrazione. Dall'assioma di hermitianità segue che la relazione $x \perp y$ è simmetrica, cioè $x \perp y \iff y \perp x$. Inoltre l'insieme x^\perp è sempre un sottospazio vettoriale chiuso. Infatti, se definiamo $L_x: V \rightarrow \mathbb{C}$ come $L_x(y) := \langle y, x \rangle$, sappiamo che L_x è un'applicazione lineare e continua, e inoltre $x^\perp = L_x^{-1}(\{0\})$. Quindi x^\perp è un sottospazio vettoriale perché è il nucleo di un'applicazione lineare, ed è chiuso perché è la contrimmagine del chiuso $\{0\} \subset \mathbb{C}$ tramite la funzione continua L_x . Se $E \subseteq V$, anche E^\perp è sempre un sottospazio vettoriale chiuso, perché intersezione di una famiglia di sottospazi vettoriali chiusi. Infine

$$E \subseteq F \Rightarrow E^\perp = \bigcap_{x \in E} x^\perp \supseteq \bigcap_{x \in F} x^\perp = F^\perp,$$

perché più insiemi intersechiamo, più si stringe l'intersezione. □

Esercizio 1.5.4. Sia V uno spazio normato sul campo \mathbb{R} . Diremo che i vettori x, y sono ortogonali se $\|x + y\| = \|x - y\|$, una condizione che coinvolge solo la norma. La relazione è simmetrica? L'insieme x^\perp è necessariamente un sottospazio vettoriale chiuso di V ? Se no, mostrare un contreesempio.

Proposizione 1.5.5. *Sia V uno spazio con prodotto scalare ed $E \subseteq V$. Allora*

$$\begin{aligned} E^\perp &= (\text{cl } E)^\perp = (\text{span } E)^\perp = (\text{cl span } E)^\perp \\ E \cap E^\perp &\subseteq \{0\}, \quad E \subseteq \text{cl span } E \subseteq (E^\perp)^\perp, \\ V &\supseteq (\text{cl span } E) \oplus E^\perp. \end{aligned}$$

Dimostrazione. Chiaramente $E \subseteq \text{cl } E$ e quindi passando agli ortogonali l'inclusione si inverte: $E^\perp \supseteq (\text{cl } E)^\perp$. Per dimostrare che vale anche l'inclusione

opposta $E^\perp \subseteq (\text{cl } E)^\perp$, prendiamo $y \in E^\perp$ e $\bar{x} \in \text{cl } E$. Esiste una successione $x_n \in E$ tale che $x_n \rightarrow \bar{x}$. Allora, passando al limite per $n \rightarrow +\infty$,

$$y \in E^\perp \Rightarrow (\forall n \in \mathbb{N} \quad \langle y, x_n \rangle = 0) \Rightarrow \langle y, \bar{x} \rangle = 0 \iff y \perp \bar{x}.$$

Quindi ogni $y \in E^\perp$ è ortogonale a ogni $\bar{x} \in \text{cl } E$, e quindi appartiene a $(\text{cl } E)^\perp$.

Analogamente è chiaro che $E^\perp \supseteq (\text{span } E)^\perp$. Per dimostrare l'inclusione opposta, siano $y \in E^\perp$ e $x \in \text{span } E$. Esistono $x_1, \dots, x_n \in E$ ed $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ tali che $x = \alpha_1 x_1 + \dots + \alpha_n x_n$ e quindi

$$\begin{aligned} y \in E^\perp &\implies (\forall n \in \mathbb{N} \quad \langle y, x_n \rangle = 0) \implies \\ &\implies \left\langle y, \sum_{k=1}^n \alpha_k x_k \right\rangle = 0 \iff \langle y, x \rangle = 0 \iff \\ &\iff y \perp \bar{x}. \end{aligned}$$

Quindi ogni $y \in E^\perp$ è ortogonale a ogni $x \in \text{span } E$, e quindi appartiene a $(\text{span } E)^\perp$.

Si ha sempre $E \cap E^\perp \subseteq \{0\}$ perché $y \in E \cap E^\perp \implies y \perp y \iff \langle y, y \rangle = 0 \iff \|y\| = 0 \iff y = 0$. Inoltre $E \subseteq (E^\perp)^\perp$ perché

$$x \in E \implies (\forall y \in E^\perp \quad x \perp y) \iff x \in (E^\perp)^\perp.$$

Poiché $(E^\perp)^\perp$ è un sottospazio vettoriale chiuso e $E \subseteq (E^\perp)^\perp$, si ha che

$$E \subseteq \text{cl span } E \subseteq (E^\perp)^\perp.$$

Infine, poiché $(\text{cl span } E) \cap E^\perp = (\text{cl span } E) \cap (\text{cl span } E)^\perp = \{0\}$ (l'intersezione di due sottospazi non è mai vuota), ricaviamo che la somma diretta $(\text{cl span } E) \oplus E^\perp$ ha senso ed è un sottospazio di V . \square

La Proposizione 1.4.4 si può riformulare così in termini di ortogonalità (Fig. 1.15):

Proposizione 1.5.6. *Siano V uno spazio con prodotto scalare, M un sottospazio vettoriale di V ed $x \in H$, $\bar{y} \in M$. Allora la condizione che \bar{y} sia il punto di M più vicino a x equivale a dire che*

$$x - \bar{y} \in M^\perp.$$

Il teorema della minima distanza verrà usato soprattutto attraverso questa sua conseguenza:

Teorema 1.5.7 (della decomposizione ortogonale). *Sia H uno spazio di Hilbert e M un sottospazio vettoriale chiuso. Allora lo spazio si può decomporre nella somma diretta di M con M^\perp , ed M coincide con l'ortogonale dell'ortogonale:*

$$H = M \oplus M^\perp, \quad M = (M^\perp)^\perp.$$

Dimostrazione. Già sappiamo che M e M^\perp sono sottospazi vettoriali e che si intersecano solo nell'origine. Resta da vedere che ogni vettore $x \in H$ si può scrivere come somma di un elemento di M e di uno di M^\perp . Osserviamo che M è un sottinsieme convesso, chiuso e non vuoto di H . Per il teorema di minima distanza, esiste $\bar{y} \in M$ che ha minima distanza da x . Per la Proposizione 1.5.6, abbiamo che $x - \bar{y} \in M^\perp$. Ma allora il vettore x si può scrivere come

$$x = \underbrace{\bar{y}}_{\in M} + \underbrace{(x - \bar{y})}_{\in M^\perp}$$

cioè effettivamente come somma di un elemento di M e di uno di M^\perp , e la decomposizione è unica perché $M \cap M^\perp = \{0\}$.

Sappiamo che $M \subseteq (M^\perp)^\perp$. Supponiamo per assurdo che non siano uguali, cioè $M \subsetneq (M^\perp)^\perp$. Possiamo applicare la decomposizione ortogonale in $(M^\perp)^\perp$ invece che in H : lo spazio $(M^\perp)^\perp$ è di Hilbert perché chiuso in H , ed M è sottospazio chiuso anche di $(M^\perp)^\perp$, e quindi $(M^\perp)^\perp$ si decompone come somma diretta di M più il suo complemento ortogonale (complemento nel senso di $(M^\perp)^\perp$, non di H):

$$(M^\perp)^\perp = M \oplus M', \quad \text{dove } M' = \{y \in (M^\perp)^\perp \mid y \perp M\}, \quad (1.47)$$

con $M' \neq \{0\}$. Notare che uso qui la notazione M' , perché riservo la notazione M^\perp al complemento in H . Per ipotesi esiste un $x' \in M' \setminus \{0\}$. Allora $x' \perp M$, cioè $x' \in M^\perp$, e contemporaneamente $x' \in (M^\perp)^\perp$. Insomma, x' è ortogonale a se stesso, cioè $x' = 0$, contraddizione. \square

Nell'ambito degli spazi di Hilbert, M^\perp merita di essere chiamato *complemento ortogonale* del sottospazio chiuso M . Occasionalmente bisogna specificare rispetto a che spazio ambiente si fa il complemento.

Esercizio 1.5.8. Se M, N sono sottospazi vettoriali chiusi di uno spazio di Hilbert H , allora

$$M = N^\perp \iff M^\perp = N, \quad M = N \iff M^\perp = N^\perp.$$

Se M, N sono sottospazi vettoriali di H non necessariamente chiusi, allora

$$M = N^\perp \Rightarrow M^\perp = \text{cl } N, \quad \text{cl } M = N^\perp \iff M^\perp = \text{cl } N. \quad (1.48)$$

Esercizio 1.5.9. Sia V lo spazio delle funzioni continue $[0, 1] \rightarrow \mathbb{R}$ dotato del prodotto scalare di $L^2([0, 1])$. Sia M l'insieme delle $f \in V$ tali che $\int_0^{1/2} f(t) dt - \int_{1/2}^1 f(t) dt = 0$. Mostrare che M è un sottospazio vettoriale chiuso e proprio di V e inoltre $M^\perp = \{0\}$, $M \subsetneq (M^\perp)^\perp = V$ e $V \not\supseteq M \oplus M^\perp$, e se $f \in V \setminus M$ non esiste alcun punto di M con minima distanza da f . (Immergere V in L^2).

Quando uno spazio vettoriale si decompone in somma diretta di due sottospazi, alla decomposizione si associano le due proiezioni sui due sottospazi, che sono operatori lineari dallo spazio nei rispettivi sottospazi. Nel caso della decomposizione ortogonale $H = M \oplus M^\perp$, le proiezioni sono dette *proiezioni ortogonali*, e le indicheremo con P_M e P_{M^\perp} . Quindi per ogni $x \in H$, se M è un sottospazio vettoriale chiuso, abbiamo

$$x = \underbrace{P_M x}_{\in M} + \underbrace{P_{M^\perp} x}_{\in M^\perp}.$$

$P_M x$ è il punto di M più vicino a x , $P_{M^\perp} x$ è il punto di M^\perp più vicino a x . Le due proiezioni sono ortogonali fra loro, per cui abbiamo

$$\begin{aligned} \|x\|^2 &= \|P_M x + P_{M^\perp} x\|^2 = \|P_M x\|^2 + \|P_{M^\perp} x\|^2 + 2\Re \underbrace{\langle P_M x, P_{M^\perp} x \rangle}_{\substack{\in M \\ \in M^\perp}} = \\ &= \|P_M x\|^2 + \|P_{M^\perp} x\|^2 + 0 = \|P_M x\|^2 + \|P_{M^\perp} x\|^2. \end{aligned}$$

L'uguaglianza

$$\|x\|^2 = \|P_M x\|^2 + \|P_{M^\perp} x\|^2 \quad (1.49)$$

può essere chiamata *teorema di Pitagora* (Figura 1.16 in alto). Una conseguenza immediata è che

$$\|P_M x\| \leq \|x\|, \quad \|P_{M^\perp} x\| \leq \|x\| \quad \forall x \in H, \quad (1.50)$$

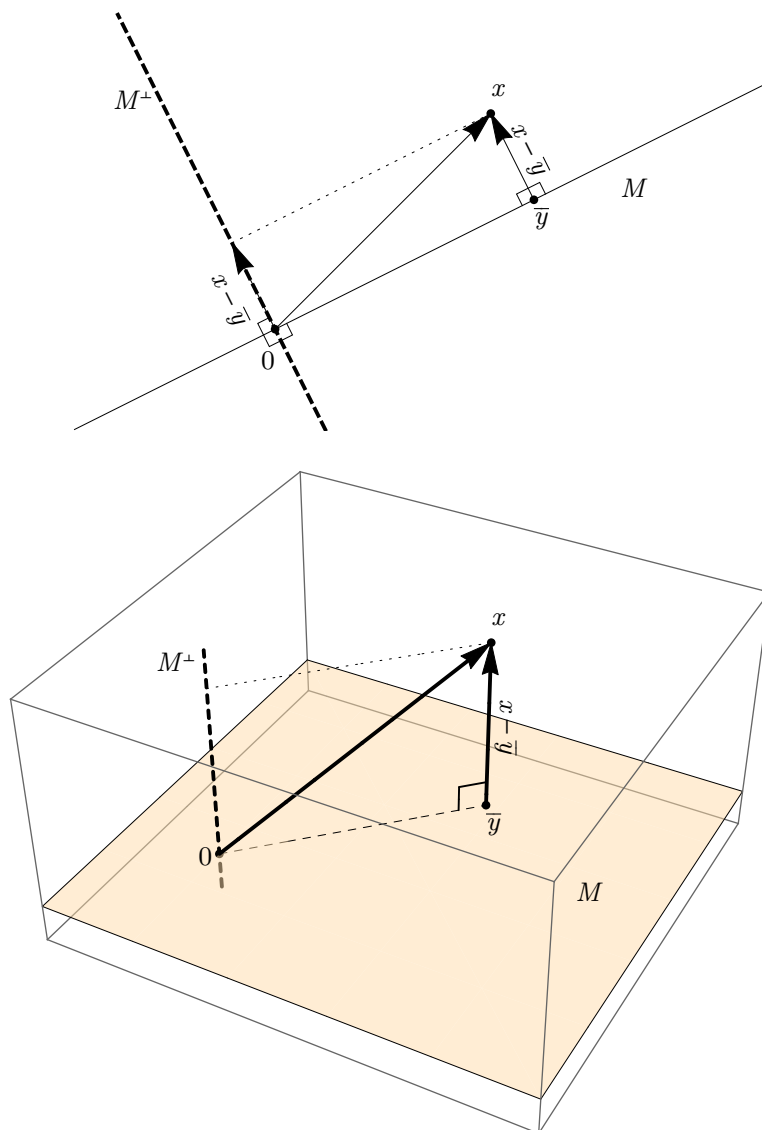


Figura 1.15: Interpretazione geometrica della decomposizione ortogonale di uno spazio di Hilbert $H = M \oplus M^\perp$ (Teorema 1.5.7 p. 34). Sopra, il caso in cui $H = \mathbb{R}^2$ e M ed M^\perp sono rette (Fig. 1.13 p. 29). Sotto $H = \mathbb{R}^3$, M è un piano e M^\perp è una retta (Fig. 1.14 p. 30). Il punto \bar{y} è il punto di M con minima distanza da x

(l'ipotenusa è maggiore o uguale a ogni cateto) e

$$\|P_M x\| = \|x\| \iff x \in M \iff P_{M^\perp} x = 0.$$

In particolare, *le proiezioni ortogonali sono operatori continui*. La loro norma operatoriale è 1, eccetto nel caso banale in cui $M = \{0\}$.

1.6 Sottospazi a dimensione finita

Lemma 1.6.1. *Sia V uno spazio normato, M un suo sottospazio vettoriale chiuso, e sia $\bar{y} \in V \setminus M$. Allora $\text{span}(M \cup \{\bar{y}\})$ è anch'esso chiuso.*

Dimostrazione. Sia $z_n \in \text{span } M \cup \{\bar{y}\}$ tale che $z_n \rightarrow \bar{z} \in V$ per $n \rightarrow +\infty$. Vogliamo dimostrare che anche $\bar{z} \in \text{span } M \cup \{\bar{y}\}$.

Ognuno degli z_n è una combinazione lineare $z_n = x_n + \lambda_n \bar{y}$ con $x_n \in M$ e λ_n scalare. Poiché le successioni convergenti sono in particolare limitate, esiste $r > 0$ tale che $\|x_n\| \leq r$ per ogni n . Voglio dimostrare che λ_n è limitata. Se non lo fosse, esisterebbe una sottosuccessione λ_{n_k} tale che $|\lambda_{n_k}| \rightarrow +\infty$. Dividiamo per $|\lambda_{n_k}|$ ambo i membri della disuguaglianza $\|z_{n_k}\| \leq r$:

$$\frac{\|x_{n_k} + \lambda_{n_k} \bar{y}\|}{|\lambda_{n_k}|} \leq \frac{r}{|\lambda_{n_k}|},$$

cioè, portando il λ_{n_k} dentro alla norma,

$$\left\| \frac{x_{n_k}}{\lambda_{n_k}} + \bar{y} \right\| \leq \frac{r}{|\lambda_{n_k}|},$$

cioè

$$\left\| \bar{y} - \underbrace{\left(-\frac{x_{n_k}}{\lambda_{n_k}} \right)}_{\in M} \right\| \leq \frac{r}{|\lambda_{n_k}|} \rightarrow 0,$$

in cui il primo membro è la distanza fra \bar{y} e $-x_{n_k}/\lambda_{n_k}$, per cui

$$\underbrace{-\frac{x_{n_k}}{\lambda_{n_k}}}_{\in M} \rightarrow \bar{y}.$$

Quindi \bar{y} risulterebbe essere limite di una successione di elementi di M , che è chiuso. Anche \bar{y} apparterrebbe a M , contro l'ipotesi. Effettivamente λ_n è limitata.

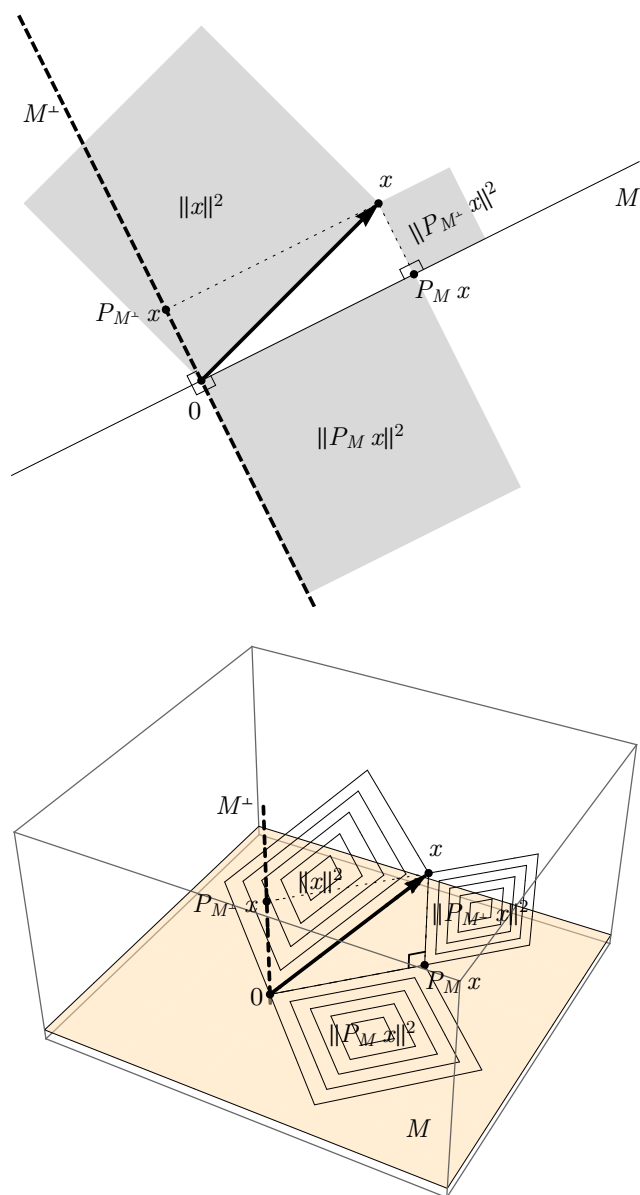


Figura 1.16: Decomposizione ortogonale di uno spazio di Hilbert nella somma diretta di due sottospazi. Sopra, il caso in cui M ed M^\perp sono rette, ed i quadrati costruiti sui cateti e l'ipotenusa del triangolo rettangolo di vertici l'origine, x e $P_M x$, che forniscono l'interpretazione geometrica della formula (1.49) che si può chiamare "Teorema di Pitagora". Sotto, M è un piano e M^\perp è una retta

Esiste quindi una sottosuccessione λ_{n_k} che converge a uno scalare $\bar{\lambda}$ (non si intende il coniugato di λ). Ma allora anche x_{n_k} converge a un qualche \bar{x} perché differenza di successioni convergenti:

$$x_{n_k} = z_{n_k} - \lambda_{n_k} \bar{y} \longrightarrow \bar{z} - \bar{\lambda} \bar{y} =: \bar{x}.$$

Poiché $x_{n_k} \in M$ e M è chiuso, anche $\bar{x} \in M$. Dall'uguaglianza $\bar{z} = \bar{x} + \bar{\lambda} \bar{y}$ segue infine che $\bar{z} \in \text{span } M \cup \{\bar{y}\}$, come desiderato. \square

Svilupperemo alcuni temi del lemma precedente nel prossimo capitolo (da p. 74), ma per ora ci basta questa conseguenza:

Proposizione 1.6.2. *Sia V uno spazio normato e M un sottospazio a dimensione finita. Allora M è chiuso in V .*

Notare che non si chiede che V abbia dimensione finita.

Dimostrazione. Per induzione sulla dimensione di M . Se la dimensione è zero, cioè $M = \{0\}$, allora chiaramente M è chiuso. Il passo induttivo usa semplicemente il lemma precedente. \square

Sia ora H uno spazio di Hilbert, $x \in H$, e M un sottospazio a dimensione finita, generato dalla base u_1, u_2, \dots, u_n . Sappiamo che esiste unico il punto \bar{y} di M più vicino a x , ossia la proiezione ortogonale di x su M . Come possiamo *calcolarla*? Il modo più semplice è usare la caratterizzazione tramite l'ortogonalità Prop. 1.5.6 ($x - \bar{y} \in M^\perp$) o Prop. 1.4.4:

$$\langle x - \bar{y}, y \rangle = 0 \quad \forall y \in M$$

(notare che questa caratterizzazione non richiede che lo spazio H sia completo). È sufficiente imporre l'uguaglianza $\langle x - \bar{y}, y \rangle = 0$ quando y è un elemento della base di M :

$$\langle x - \bar{y}, u_k \rangle = 0 \quad \forall k = 1, 2, \dots, n.$$

Essendo \bar{y} un elemento di M , è esprimibile come $\bar{y} = \lambda_1 u_1 + \dots + \lambda_n u_n$, dove $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ sono scalari incogniti (Figura 1.17). Sostituendo otteniamo le condizioni

$$\left\langle x - \sum_{j=1}^n \lambda_j u_j, u_k \right\rangle = 0 \quad \forall k = 1, 2, \dots, n,$$

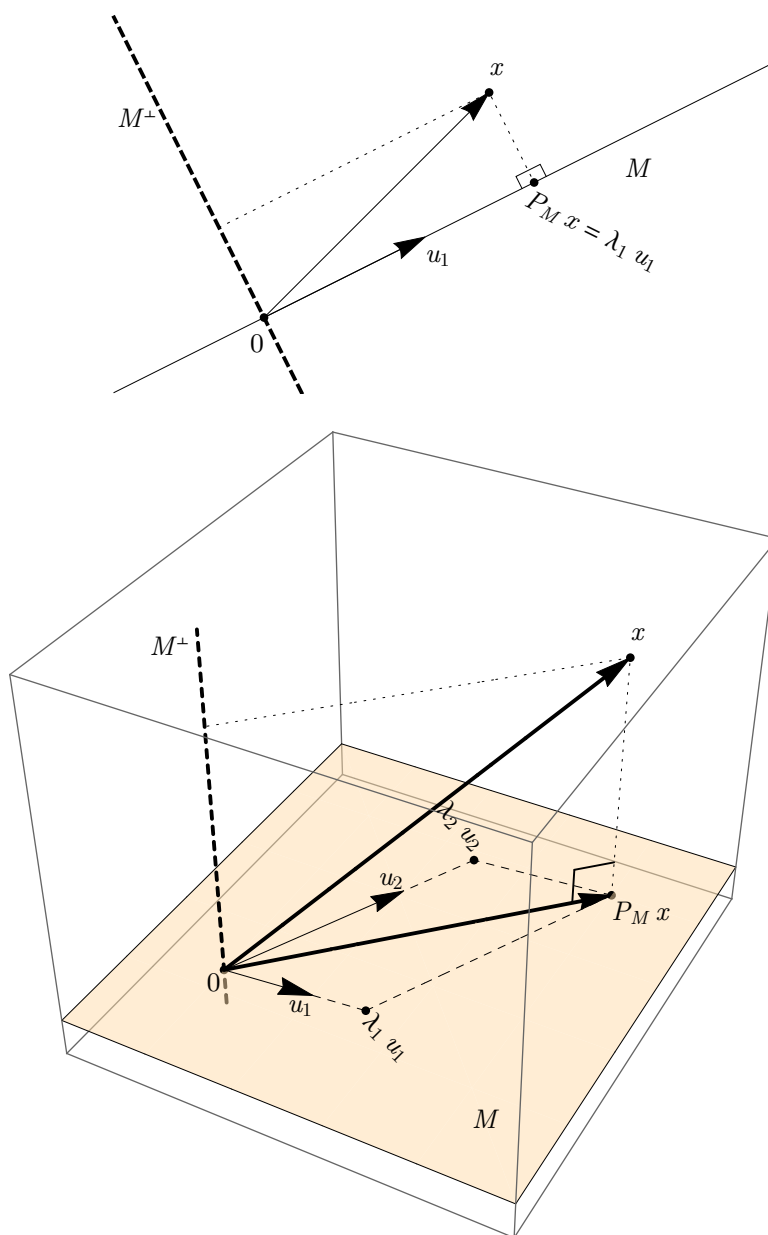


Figura 1.17: La proiezione ortogonale $P_M x$ di un vettore x di uno spazio di Hilbert su un sottospazio M generato da una base u_1, \dots, u_n . Sopra, il caso in cui $n = 1$ e H è bidimensionale. Sotto, $n = 2$ dentro un H tridimensionale e base $\{u_1, u_2\}$ di M non ortonormale. Confrontare col caso ortonormale di Fig. 1.18

da cui, distribuendo i prodotti scalari e riordinando:

$$\sum_{j=1}^n \lambda_j \langle u_j, u_k \rangle = \langle x, u_k \rangle \quad \forall k = 1, 2, \dots, n. \quad (1.51)$$

Possiamo interpretare queste equazioni (1.51) come n equazioni lineari nelle incognite $\lambda_1, \dots, \lambda_n$, con coefficienti $\langle u_j, u_k \rangle$ e termini noti $\langle x, u_k \rangle$. La matrice dei coefficienti deve essere invertibile, perché sappiamo per via geometrica che il problema ha sempre soluzione unica. Il problema di calcolare \bar{y} si riconduce quindi all'inversione di una matrice.

Esercizio 1.6.3. Mostrare che la matrice dei coefficienti $a_{j,k} := \langle u_j, u_k \rangle$ è hermitiana (simmetrica nel caso reale) e definita positiva. In particolare, l'ipotesi che H sia completo non è necessaria.

Il caso che più ci interessa è quello in cui la base di M è *ortonormale*, cioè tale che

$$\langle u_j, u_k \rangle = \begin{cases} 1 & \text{se } j = k, \\ 0 & \text{se } j \neq k. \end{cases}$$

In queste ipotesi la matrice dei coefficienti delle equazioni (1.51) è la matrice identità, e le equazioni sono già risolte:

$$\lambda_k = \langle x, u_k \rangle \quad \forall k = 1, 2, \dots, n.$$

Ribadiamo che l'ipotesi che H sia completo non è necessaria. Riassumendo in un enunciato:

Proposizione 1.6.4 (proiezione su un sottospazio a dimensione finita). *Se V è uno spazio con prodotto scalare e u_1, \dots, u_n è una base ortonormale del sottospazio M , esiste unico il punto $\bar{y} \in M$ più vicino a x , ossia la proiezione ortogonale di x su M , e si scrive come*

$$\bar{y} = P_M x = \sum_{k=1}^n \langle x, u_k \rangle u_k. \quad (1.52)$$

Il caso $n = 1$ costituisce una giustificazione a posteriori della formula della proiezione di x sulla retta di y fatta nella Figura 1.3 pag. 12.

Notare che il coefficiente di u_k è $\langle x, u_k \rangle$ e quindi dipende solo da u_k , non dagli altri elementi della base. Questa fatto non si estende alle basi non

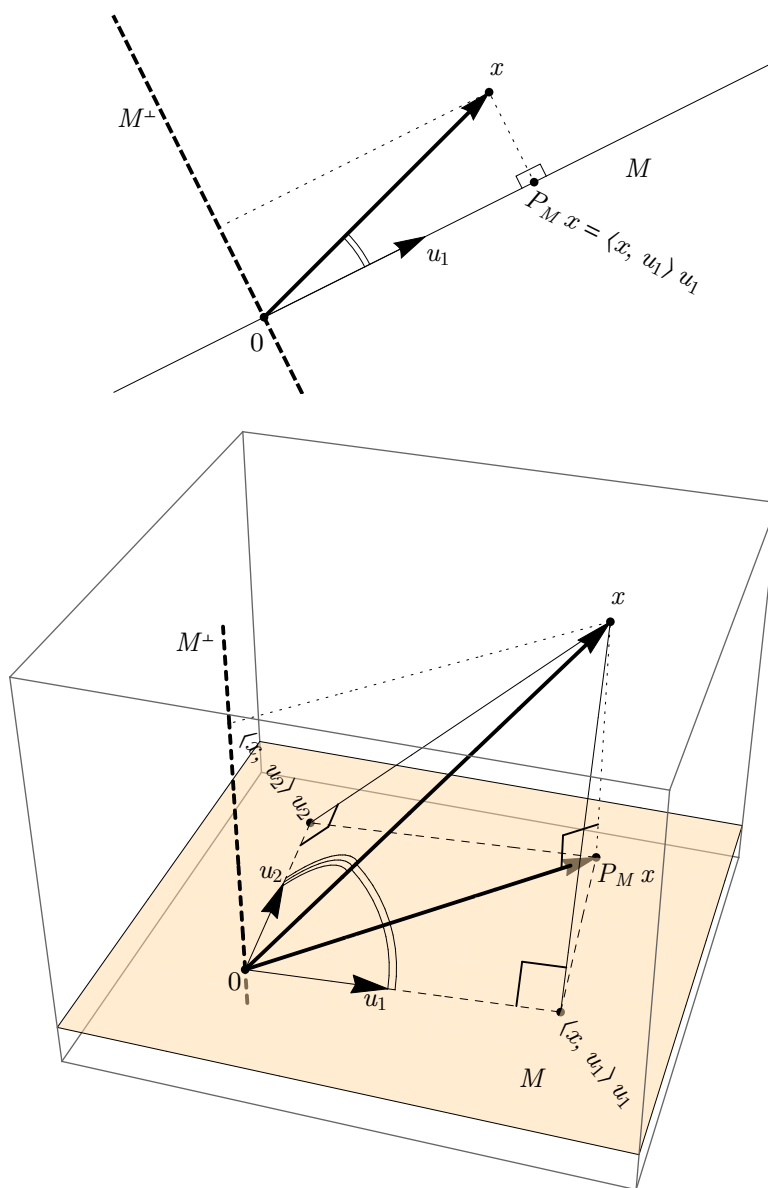


Figura 1.18: Interpretazione geometrica della formula della proiezione ortogonale $P_M x$ di un vettore x di uno spazio di Hilbert su un sottospazio M generato da una base *ortonormale* u_1, \dots, u_n (Prop. 1.6.4; confrontare con la Fig. 1.17). Sopra, il caso in cui $n = 1$ e H è bidimensionale. Sotto, $n = 2$ dentro un H tridimensionale. Il coefficiente $\langle x, u_k \rangle$ di u_k si può interpretare come la norma di x per il coseno dell'angolo fra x e u_k

ortonormali: se cambio gli altri elementi della base, posso aspettarmi che cambi anche il coefficiente di u_k .

La norma quadra di $P_M x$ si può calcolare così:

$$\begin{aligned} \|\bar{y}\|^2 &= \|P_M x\|^2 = \left\| \sum_{k=1}^n \langle x, u_k \rangle u_k \right\|^2 = \sum_{k=1}^n \|\langle x, u_k \rangle u_k\|^2 = \\ &= \sum_{k=1}^n |\langle x, u_k \rangle|^2 \|u_k\|^2 = \sum_{k=1}^n |\langle x, u_k \rangle|^2, \end{aligned} \quad (1.53)$$

grazie al fatto che i vettori della somma sono a due a due ortogonali, e quindi i doppi prodotti sono tutti nulli (“teorema di Pitagora”, formula (1.2) pag. 8).

Corollario 1.6.5. *Sia H uno spazio di Hilbert, siano u_1, \dots, u_n dei vettori ortonormali, e $x \in H$. Allora*

$$\sum_{k=1}^n |\langle x, u_k \rangle|^2 \leq \|x\|^2. \quad (1.54)$$

La disuguaglianza (1.54) è la prima versione di quella che è chiamata *disuguaglianza di Bessel*.

Dimostrazione. Sia $M = \text{span}\{u_1, \dots, u_n\}$. Dalle formule (1.53) e (1.50) segue che

$$\sum_{k=1}^n |\langle x, u_k \rangle|^2 = \left\| \sum_{k=1}^n \langle x, u_k \rangle u_k \right\|^2 = \|P_M x\|^2 \leq \|x\|^2.$$

□

Ci proponiamo di estendere la formula della proiezione ortogonale a un sottospazio generato da *infiniti* vettori ortonormali. Però prima dobbiamo chiarire cosa si intende per somma di infiniti vettori, che non è del tutto ovvia.

Esercizio 1.6.6. Se H ha dimensione finita e u_1, \dots, u_n formano una base ortonormale, dati $x = \sum_k \alpha_k u_k$, $y = \sum_k \beta_k u_k$, trovare una formula per $\langle x, y \rangle$ in termini dei coefficienti α_k, β_k . Confrontare con l’esempio 1.1.7 di p. 13.

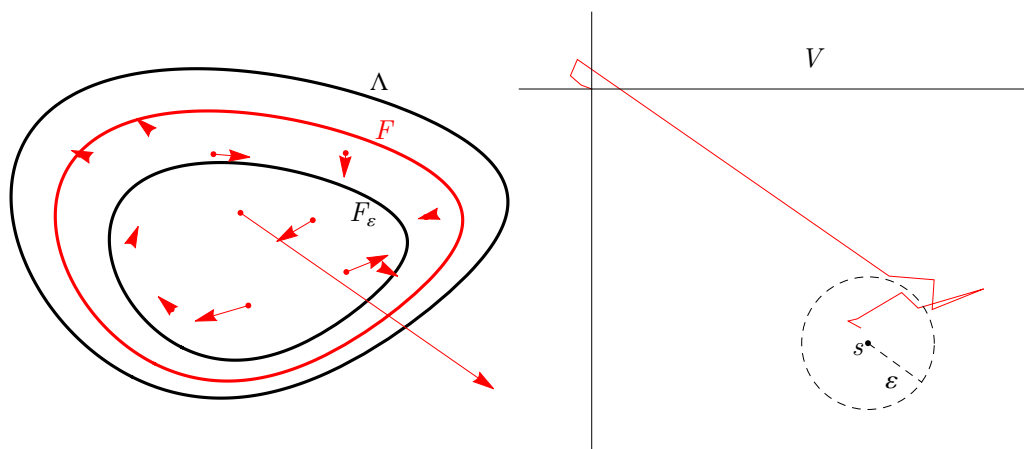


Figura 1.19: La definizione 1.7.1 di sommabilità coinvolge da una parte un insieme astratto di indici Λ , raffigurato a sinistra, e dall'altra lo spazio normato V , rappresentato come un piano \mathbb{R}^2 qui a destra. Il vettore somma s ha un intorno di raggio ε , nel quale giungono tutte le somme dei vettori corrispondenti agli indici che stanno in insiemi finiti F che siano sovrainsiemi di un certo F_ε

1.7 Somme infinite di vettori

Daremo per scontato di sapere cos'è $\sum_{\lambda \in \Lambda} x_\lambda$ quando Λ è un insieme *finito* di indici e gli x_λ sono elementi di uno spazio vettoriale.

Definizione 1.7.1 (Sommabilità). Sia V uno spazio normato, Λ un insieme di indici, $\{x_\lambda : \lambda \in \Lambda\} \subset V$ una famiglia di vettori indicizzata da Λ , e $s \in V$. Diremo che la famiglia è *sommabile con somma s* se per ogni $\varepsilon > 0$ esiste un sottinsieme *finito* $F_\varepsilon \subseteq \Lambda$ di indici tale che per ogni altro insieme finito F tale che $F_\varepsilon \subseteq F \subseteq \Lambda$ si abbia che

$$\left\| s - \sum_{\lambda \in F} x_\lambda \right\| < \varepsilon.$$

Scriveremo in tal caso

$$s = \sum_{\lambda \in \Lambda} x_\lambda$$

(si veda la figura 1.19).

Esercizio 1.7.2. La somma di una famiglia sommabile è unica. Cosa succede se $\Lambda = \emptyset$? Siano V, W spazi normati, $A: V \rightarrow W$ lineare continua, $x_\lambda \in V$ famiglia sommabile con somma s . Che dire della famiglia $Ax_\lambda \in W$?

Esercizio 1.7.3. Se $\Lambda_1 \cap \Lambda_2 = \emptyset$ allora $\sum_{\lambda \in \Lambda_1 \cup \Lambda_2} x_\lambda = \sum_{\lambda \in \Lambda_1} x_\lambda + \sum_{\lambda \in \Lambda_2} x_\lambda$.

Esercizio 1.7.4. Se $\Lambda_0 \subseteq \Lambda$ e $x_\lambda = 0 \forall \lambda \in \Lambda \setminus \Lambda_0$ allora $\sum_{\lambda \in \Lambda} x_\lambda = \sum_{\lambda \in \Lambda_0} x_\lambda$ (gli addendi nulli non contano).

Esercizio 1.7.5. Se $\varphi: \Lambda_1 \rightarrow \Lambda_2$ è biiettiva allora $\sum_{\mu \in \Lambda_2} x_\mu = \sum_{\lambda \in \Lambda_1} x_{\varphi(\lambda)}$ (cambio di indici nella somma).

Esercizio 1.7.6. Se $V = \mathbb{C}^n$ e $x_\lambda = (x_\lambda(1), x_\lambda(2), \dots, x_\lambda(n)) \in V$, allora la somma vettoriale $\sum x_\lambda$ è sommabile se e solo se sono sommabili tutte le componenti, e le componenti della somma sono le somme delle componenti.

Definizione 1.7.7 (condizione di Cauchy). Sia V uno spazio normato, Λ un insieme di indici, $\{x_\lambda : \lambda \in \Lambda\} \subset V$ una famiglia di vettori indicizzata da Λ . Diremo che la famiglia è *di Cauchy per la sommabilità* se per ogni $\varepsilon > 0$ esiste un sottinsieme *finito* $G_\varepsilon \subseteq \Lambda$ di indici tale che per ogni altro insieme finito $G \subseteq \Lambda$ tale che $G_\varepsilon \cap G = \emptyset$ si abbia che

$$\left\| \sum_{\lambda \in G} x_\lambda \right\| < \varepsilon.$$

(si veda la figura 1.20).

Proposizione 1.7.8 (la sommabilità implica Cauchy). *Se la famiglia x_λ è sommabile, allora è di Cauchy.*

Dimostrazione. Supponiamo che x_λ sia sommabile con somma s e fissiamo $\varepsilon > 0$. Esiste $F_{\varepsilon/2}$ finito tale che per ogni $F \supset F_{\varepsilon/2}$ finito vale la disuguaglianza

$$\left\| s - \sum_{\lambda \in F} x_\lambda \right\| < \frac{\varepsilon}{2}.$$

Poniamo $G_\varepsilon = F_{\varepsilon/2}$. Sia G finito disgiunto da G_ε . Allora sia G_ε che $G \cup G_\varepsilon$ sono soprainsiemi finiti di G_ε , e quindi

$$\left\| \sum_{\lambda \in G} x_\lambda \right\| = \left\| \sum_{\lambda \in G \cup G_\varepsilon} x_\lambda - s + s - \sum_{\lambda \in G_\varepsilon} x_\lambda \right\| \leq$$

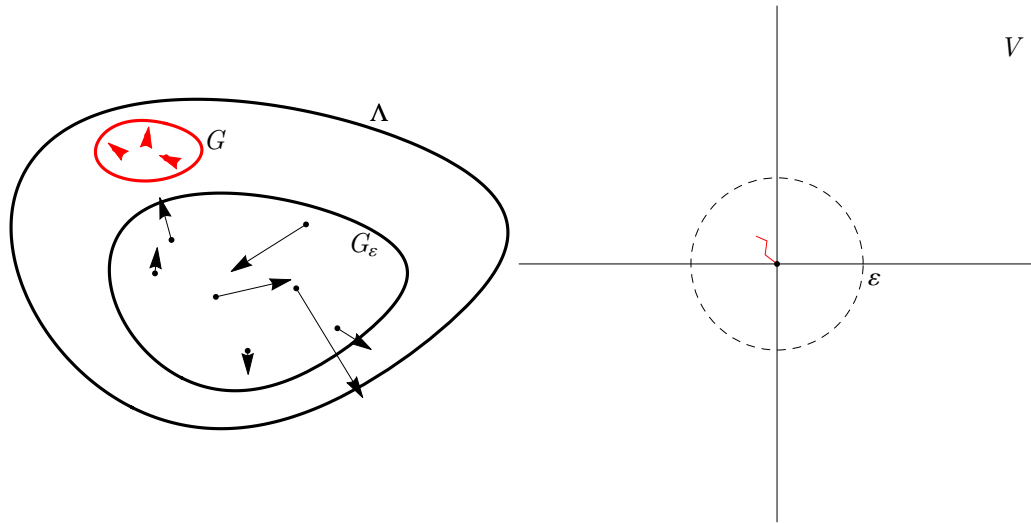


Figura 1.20: La definizione 1.7.7 di condizione di Cauchy per la sommabilità richiede che le somme dei vettori che stanno in un qualsiasi insieme finito G disgiunto da un certo G_ε non escono da una palla di centro l'origine di V e raggio ε

$$\begin{aligned} &\leq \left\| \sum_{\lambda \in G \cup F_{\varepsilon/2}} x_\lambda - s \right\| + \left\| s - \sum_{\lambda \in G \cup F_{\varepsilon/2}} x_\lambda \right\| < \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

La definizione di proprietà di Cauchy è pertanto verificata. \square

Proposizione 1.7.9 (i vettori non nulli sono al più numerabili). *Se la famiglia x_λ ha la proprietà di Cauchy, allora l'insieme $\{\lambda \in \Lambda : x_\lambda \neq 0\}$ è finito o numerabile.*

Dimostrazione. Sia $\varepsilon > 0$, sia G_ε come richiesto da Cauchy, e $\lambda_0 \in \Lambda \setminus G_\varepsilon$. Allora $G = \{\lambda_0\}$ è disgiunto da G_ε , e pertanto

$$\|x_{\lambda_0}\| = \left\| \sum_{\lambda \in \{\lambda_0\}} x_\lambda \right\| < \varepsilon.$$

Ne segue che

$$\|x_{\lambda_0}\| \geq \varepsilon \Rightarrow \lambda_0 \in G_\varepsilon.$$

Questo vuol dire che, grosso modo, i vettori “grandi” sono tutti in G_ε , o, ancora, che fuori da G_ε i vettori sono tutti “piccoli”. Osserviamo ora che

$$\begin{aligned} \{\lambda \in \Lambda : x_\lambda \neq 0\} &= \{\lambda \in \Lambda : \|x_\lambda\| > 0\} = \\ &= \{\lambda \in \Lambda : \exists n \in \mathbb{N} \text{ tale che } \|x_\lambda\| \geq 1/n\} = \\ &= \bigcup_{n \in \mathbb{N}} \{\lambda \in \Lambda : \|x_\lambda\| \geq 1/n\} \subseteq \\ &\subseteq \bigcup_{n \in \mathbb{N}} G_{1/n}. \end{aligned}$$

L'ultimo membro, essendo l'unione numerabile di insiemi finiti, è al più numerabile. Altrettanto è di conseguenza il primo membro. \square

Proposizione 1.7.10 (Cauchy implica sommabilità). *In uno spazio V normato completo, se la famiglia x_λ ha la proprietà di Cauchy, allora è sommabile.*

La figura 1.21 dovrebbe aiutare a raccapezzarsi. Supponiamo che valga la proprietà di Cauchy. Sia $\Lambda_0 := \{\lambda \in \Lambda : x_\lambda \neq 0\}$. Per la proposizione precedente, Λ_0 è al più numerabile. Se Λ_0 è finito è chiaro che la famiglia è sommabile. Ci mettiamo quindi nel caso non banale in cui Λ_0 sia numerabile, cioè che esista una applicazione $\varphi: \mathbb{N} \rightarrow \Lambda_0$ biettiva. Consideriamo la successione delle somme parziali

$$s_n := \sum_{k=1}^n x_{\varphi(k)}.$$

Dico che s_n è una successione di Cauchy in V . Sia infatti $\varepsilon > 0$ e sia G_ε finito come nella definizione 1.7.7. Poniamo $N_\varepsilon := \max\{n \in \mathbb{N} : \varphi(n) \in G_\varepsilon\} = \max \varphi^{-1}(G_\varepsilon)$ (se per caso $\varphi^{-1}(G_\varepsilon)$ è vuoto si ponga $N_\varepsilon = 1$), di modo che $n > N_\varepsilon$ implichi che $\varphi(n) \notin G_\varepsilon$. Siano $n > m > N_\varepsilon$. Allora $G := \{\varphi(m+1), \varphi(m+2), \dots, \varphi(n)\}$ è disgiunto da G_ε , e quindi

$$\begin{aligned} \|s_n - s_m\| &= \left\| \sum_{k=1}^n x_{\varphi(k)} - \sum_{k=1}^m x_{\varphi(k)} \right\| = \left\| \sum_{k=m+1}^n x_{\varphi(k)} \right\| = \\ &= \left\| \sum_{\lambda \in G} x_\lambda \right\| < \varepsilon. \end{aligned}$$

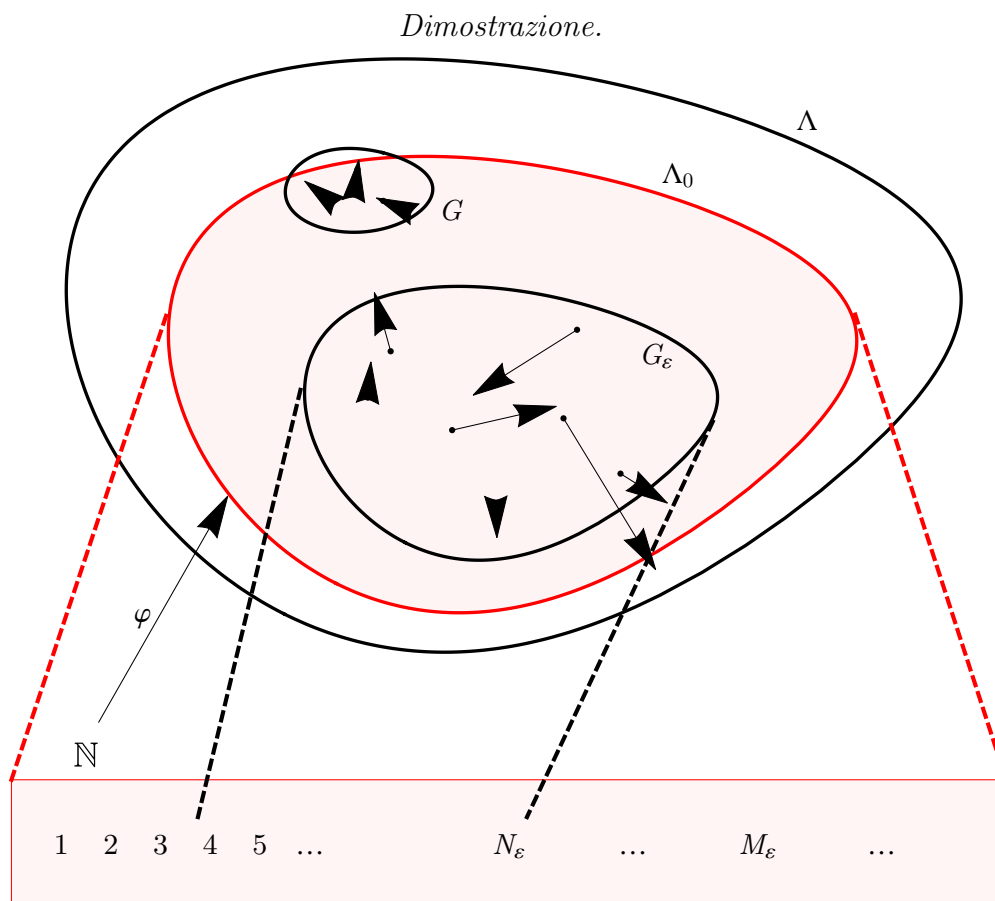


Figura 1.21: La dimostrazione che condizione di Cauchy implica la sommabilità (Proposizione 1.7.10) coinvolge diversi sottinsiemi dell'insieme Λ di indici. Λ_0 raccoglie gli indici con vettori non nulli, ed è in corrispondenza biunivoca φ con \mathbb{N} , raffigurato sotto. La contrimmagine di G_ε comprende un numero finito di interi, dei quali N_ε è il più grande. Un qualsiasi numero finito di vettori corrispondenti a numeri $> N_\varepsilon$ cadrà in un G disgiunto da G_ε , e quindi la loro somma sarà $< \varepsilon$. Non sono raffigurati gli insiemi $F_\varepsilon \subseteq F$ usati nella dimostrazione, per non affollare troppo la figura

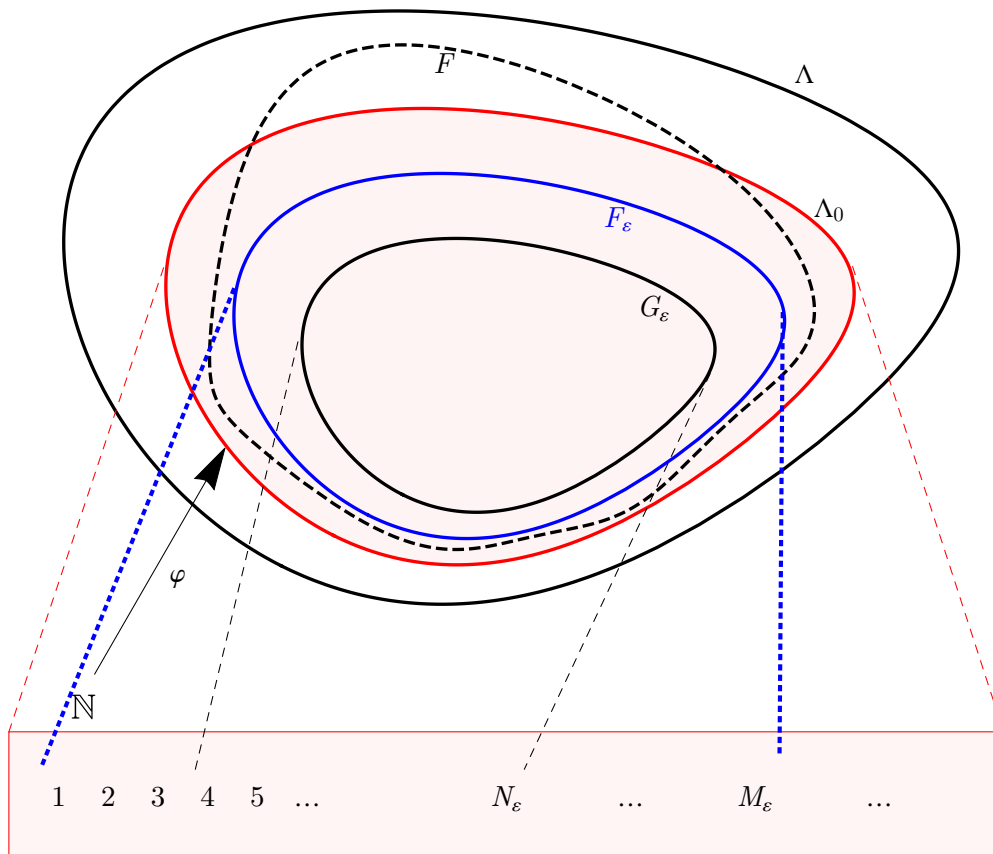


Figura 1.22: La seconda parte della dimostrazione che condizione di Cauchy implica la sommabilità (Proposizione 1.7.10) coinvolge diversi altri sottinsiemi dell'insieme Λ di indici. Le somme parziali al di là di M_ε della serie di vettori distano da s meno di ε . F_ε raccoglie tutti gli indici fino corrispondenti agli interi consecutivi da 1 a M_ε . La somma dei vettori di una soprafamiglia $F \supseteq F_\varepsilon$ disterà poco da s perché differisce dalla somma parziale per una somma di vettori fuori da G_ε

Poiché V è completo, la successione s_n converge a un qualche $s \in V$. Ci proponiamo di dimostrare che la famiglia x_λ è sommabile con somma precisamente s .

Sia $\varepsilon > 0$ e $M_\varepsilon \in \mathbb{N}$ tale che $M_\varepsilon > N_\varepsilon$ e inoltre per ogni $n \geq M_\varepsilon$ si abbia che $\|s_n - s\| < \varepsilon$. Poniamo

$$F_\varepsilon := \{\varphi(n) : n \in \mathbb{N}, n \leq M_\varepsilon\}.$$

È chiaro che $\Lambda_0 \cap G_\varepsilon \subseteq F_\varepsilon \subseteq \Lambda_0$. Sia $F \supseteq F_\varepsilon$ finito. Allora

$$\begin{aligned} \left\| s - \sum_{\lambda \in F} x_\lambda \right\| &= \left\| s - \sum_{\lambda \in F_\varepsilon} x_\lambda - \sum_{\lambda \in F \setminus F_\varepsilon} x_\lambda \right\| = \\ &= \left\| s - \sum_{\lambda \in F_\varepsilon} x_\lambda - \sum_{\lambda \in \Lambda_0 \cap F \setminus F_\varepsilon} x_\lambda \right\| \leq \\ &\leq \left\| s - \sum_{k=1}^{M_\varepsilon} x_{\varphi(k)} \right\| + \left\| \sum_{\lambda \in \Lambda_0 \cap F \setminus F_\varepsilon} x_\lambda \right\| = \\ &= \|s - s_{M_\varepsilon}\| + \left\| \sum_{\lambda \in \Lambda_0 \cap F \setminus F_\varepsilon} x_\lambda \right\| < \varepsilon + \varepsilon = 2\varepsilon, \end{aligned}$$

poiché $\Lambda_0 \cap F \setminus F_\varepsilon$ non interseca F_ε , e quindi a maggior ragione è disgiunto da G_ε . \square

Nel caso di somme infinite di numeri reali non negativi c'è una definizione alternativa, più semplice, di somma, che è compatibile con quella del caso vettoriale:

Proposizione 1.7.11. *Sia $V = \mathbb{R}$ e sia $x_\lambda \geq 0$ per ogni $\lambda \in \Lambda \neq \emptyset$. Poniamo*

$$s = \sup \left\{ \sum_{\lambda \in H} x_\lambda : H \text{ finito} \subseteq \Lambda \right\}$$

Allora la famiglia x_λ è sommabile se e solo se $s < +\infty$, e in tal caso la somma è proprio s .

Dimostrazione. Supponiamo che x_λ sia sommabile con somma $s' \in \mathbb{R}$. Sia $\varepsilon > 0$ e F_ε con le proprietà della definizione. Sia H un sottinsieme finito di Λ . Allora $H \cup F_\varepsilon$ è un soprainsieme finito di F_ε e quindi

$$\sum_{\lambda \in H} x_\lambda \leq \sum_{\lambda \in H \cup F_\varepsilon} x_\lambda = s' + \sum_{\lambda \in H \cup F_\varepsilon} x_\lambda - s' \leq s' + \varepsilon.$$

Valendo questo per ogni H finito, deduciamo che $s \leq s' + \varepsilon < +\infty$. Inoltre, essendo questo vero per ogni $\varepsilon > 0$, si ricava che $s \leq s'$. Per dimostrare che $s \geq s'$ basta osservare che

$$s \geq \sum_{\lambda \in F_\varepsilon} x_\lambda = s' - \left(s' - \sum_{\lambda \in F_\varepsilon} x_\lambda \right) \geq s' - \varepsilon \quad \forall \varepsilon > 0.$$

Viceversa, supponiamo che $s < +\infty$. Per definizione di estremo superiore, esiste G_ε finito tale che

$$\sum_{\lambda \in G_\varepsilon} x_\lambda \geq s - \varepsilon.$$

Sia G disgiunto da G_ε . Allora

$$\begin{aligned} \left\| \sum_{\lambda \in G} x_\lambda \right\| &= \sum_{\lambda \in G} x_\lambda = \sum_{\lambda \in G \cup G_\varepsilon} x_\lambda - \sum_{\lambda \in G_\varepsilon} x_\lambda \leq s - \sum_{\lambda \in G_\varepsilon} x_\lambda \leq \\ &\leq s - (s - \varepsilon) = \varepsilon. \end{aligned}$$

Quindi la famiglia x_λ ha la proprietà di Cauchy e pertanto è sommabile. \square

Nel caso non infrequente in cui x_λ siano tutti reali ≥ 0 , possiamo dunque *definire* la loro somma come

$$\sum_{\lambda \in \Lambda} x_\lambda := \sup \left\{ \sum_{\lambda \in H} x_\lambda : H \text{ finito } \subseteq \Lambda \right\} \in [0, +\infty].$$

La famiglia è sommabile se e solo se la somma è finita.

Esercizio 1.7.12. Dimostrare che se $\sum \|x_\lambda\| < +\infty$ allora $\sum \|x_\lambda\|^2 < +\infty$.

Il criterio seguente, che possiamo chiamare della “convergenza assoluta”, lega la sommabilità di vettori alla sommabilità delle norme, però in una sola direzione:

Proposizione 1.7.13 (la convergenza assoluta implica Cauchy). *Sia V uno spazio normato e x_λ una famiglia di vettori. Se*

$$\sum_{\lambda \in \Lambda} \|x_\lambda\| < +\infty,$$

allora la famiglia ha la proprietà di Cauchy per la sommabilità.

Dimostrazione. Poniamo $r = \sum_{\lambda \in \Lambda} \|x_\lambda\|$. Per ipotesi $0 \leq r < +\infty$. Sia $\varepsilon > 0$. Per definizione di estremo superiore, esiste G_ε finito $\subseteq \Lambda$ tale che

$$\sum_{\lambda \in G_\varepsilon} \|x_\lambda\| > r - \varepsilon.$$

Se G è finito e disgiunto da G_ε , abbiamo

$$\begin{aligned} \left\| \sum_{\lambda \in G} x_\lambda \right\| &\leq \sum_{\lambda \in G} \|x_\lambda\| = \sum_{\lambda \in G \cup G_\varepsilon} \|x_\lambda\| - \sum_{\lambda \in G_\varepsilon} \|x_\lambda\| \leq \\ &\leq r - \sum_{\lambda \in G_\varepsilon} \|x_\lambda\| < r - (r - \varepsilon) = \varepsilon. \end{aligned}$$

□

Proposizione 1.7.14 (disuguaglianza triangolare infinita). *Se $\sum x_\lambda$ è sommabile, allora*

$$\left\| \sum_{\lambda \in \Lambda} x_\lambda \right\| \leq \sum_{\lambda \in \Lambda} \|x_\lambda\| \leq +\infty. \quad (1.55)$$

Dimostrazione. Siano $s := \sum x_\lambda$, $\sigma := \sum \|x_\lambda\|$ ed $\varepsilon > 0$. Per definizione di somme infinite, esiste $F_\varepsilon \subseteq \Lambda$ finito tali che

$$\left\| s - \sum_{\lambda \in F_\varepsilon} x_\lambda \right\| < \varepsilon.$$

Usando la disuguaglianza triangolare nella versione finita possiamo scrivere

$$\|s\| - \varepsilon < \left\| \sum_{\lambda \in F_\varepsilon} x_\lambda \right\| \leq \sum_{\lambda \in F_\varepsilon} \|x_\lambda\| \leq \sigma.$$

Valendo questo per ogni $\varepsilon > 0$, deduciamo che $\|s\| \leq \sigma$, come volevasi dimostrare. □

Proposizione 1.7.15 (la convergenza assoluta equivale alla sommabilità per somme reali). *Quando $V = \mathbb{R}$, condizione necessaria e sufficiente affinché la famiglia $x_\lambda \in \mathbb{R}$ sia sommabile è che*

$$\sum_{\lambda \in \Lambda} |x_\lambda| < +\infty.$$

Se questo succede, allora la somma s è data da

$$s = \sum_{\lambda \in \Lambda, x_\lambda \geq 0} x_\lambda - \sum_{\lambda \in \Lambda, x_\lambda < 0} -x_\lambda.$$

Proposizione 1.7.16 (la convergenza assoluta equivale alla sommabilità in dimensione finita). *Quando V è \mathbb{R}^n o \mathbb{C}^n , condizione necessaria e sufficiente affinché la famiglia $x_\lambda \in \mathbb{R}$ sia sommabile è che*

$$\sum_{\lambda \in \Lambda} \|x_\lambda\| < +\infty.$$

Esiste anche un altro modo, totalmente equivalente, di approssimare le somme infinite quando $V = \mathbb{R}$, o \mathbb{C} : quello di *integrale rispetto alla misura $\#$ del conteggio*, introdotta con la formula (1.7) a pag. 15:

$$\sum_{\lambda \in \Lambda} x_\lambda := \int_{\Lambda} x \, d\#,$$

che ha senso come numero reale (o complesso) se l'applicazione $\lambda \mapsto x_\lambda$ sta in $\ell^1(\Lambda)$, che equivale alla convergenza assoluta. L'equivalenza di questo approccio viene lasciata al lettore, perché dipende da come ha impostato la definizione di integrale rispetto a una misura positiva astratta.

In dimensione infinita la convergenza assoluta implica la sommabilità, ma non è equivalente, cioè *ci sono famiglie sommabili x_λ tali che*

$$\sum_{\lambda \in \Lambda} \|x_\lambda\| = +\infty.$$

Esempio 1.7.17. Nello spazio ℓ^p delle successioni $x: \mathbb{N} \rightarrow \mathbb{C}$ con la norma $\|x\|_p := (\sum_{n \in \mathbb{N}} |x(n)|^p)^{1/p}$ con $1 < p < +\infty$ consideriamo la seguente successione di vettori:

$$\begin{aligned} x_1 &= (1, 0, 0, 0, 0, \dots), \\ x_2 &= (0, \frac{1}{2}, 0, 0, 0, \dots), \\ x_3 &= (0, 0, \frac{1}{3}, 0, 0, \dots), \\ x_4 &= (0, 0, 0, \frac{1}{4}, 0, \dots), \\ &\dots \\ x_n(k) &= \begin{cases} \frac{1}{k} & \text{se } k = n, \\ 0 & \text{se } k \neq n. \end{cases} \end{aligned}$$

Si vede che $\|x_n\|_p = 1/n$ e quindi $\sum_{n \in \mathbb{N}} \|x_n\|_p = \sum_{n \in \mathbb{N}} 1/n = +\infty$. Però la famiglia di vettori $n \mapsto x_n$ è sommabile con somma

$$s = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{n}, \dots).$$

Infatti, dato $\varepsilon > 0$, sia N_ε tale che $\sum_{n>N_\varepsilon} 1/n^p < \varepsilon^p$. Poniamo $F_\varepsilon := \{1, 2, 3, \dots, N_\varepsilon\} \subset \mathbb{N}$. Se $F_\varepsilon \subseteq F \subset \mathbb{N}$ allora

$$\left\| s - \sum_{n \in F} x_n \right\|_p^p = \sum_{n \in \mathbb{N} \setminus F} \frac{1}{n^p} \leq \sum_{n > N_\varepsilon} \frac{1}{n^p} < \varepsilon^p.$$

Cosa succede se $p = +\infty$ o $p = 1$?

Esempio 1.7.18. Sia $V = \mathcal{C}([0, 1])$ lo spazio delle funzioni continue $[0, 1] \rightarrow \mathbb{R}$ con la norma del sup $\|f\| := \sup\{|f(t)| : 0 \leq t \leq 1\}$. Per $n \in \mathbb{N}$ consideriamo la funzione f_n definita per interpolazione lineare fra i valori

$$f(0) = f(1) = f\left(\frac{1}{n}\right) = f\left(\frac{1}{n+1}\right) = 0, \quad f\left(\frac{\frac{1}{n} + \frac{1}{n+1}}{2}\right) = \frac{1}{n}.$$

Vedi Figura 1.23. Si ha $\|f_n\| = \sup|f_n| = 1/n$, per cui $\sum_n \|f_n\| = \sum_n 1/n = +\infty$. Però f_n è sommabile, come si verifica per esercizio, e le somme parziali convergono uniformemente.

Nell'importante caso in cui l'insieme Λ di indici coincide con \mathbb{N} , è noto il concetto elementare di *convergenza "semplice" di serie*, cioè di *limite delle somme parziali*:

$$s = \lim_{N \rightarrow +\infty} \sum_{k=1}^N x_k.$$

Bisogna tenere ben distinto questo concetto da quello di sommabilità, benché ci sia una parziale sovrapposizione. L'esempio da capire bene è la *serie armonica a segni alterni*, che notoriamente converge nel senso delle somme parziali:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots = \lim_{N \rightarrow +\infty} \sum_{n=1}^N \frac{(-1)^{n+1}}{n}.$$

Si sa che la somma dei valori assoluti dei termini diverge:

$$\sum_{n=1}^N \left| \frac{(-1)^{n+1}}{n} \right| = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots = +\infty,$$

e quindi la convergenza non è assoluta, ma "semplice". Essendo $V = \mathbb{R}$, non c'è neppure sommabilità.

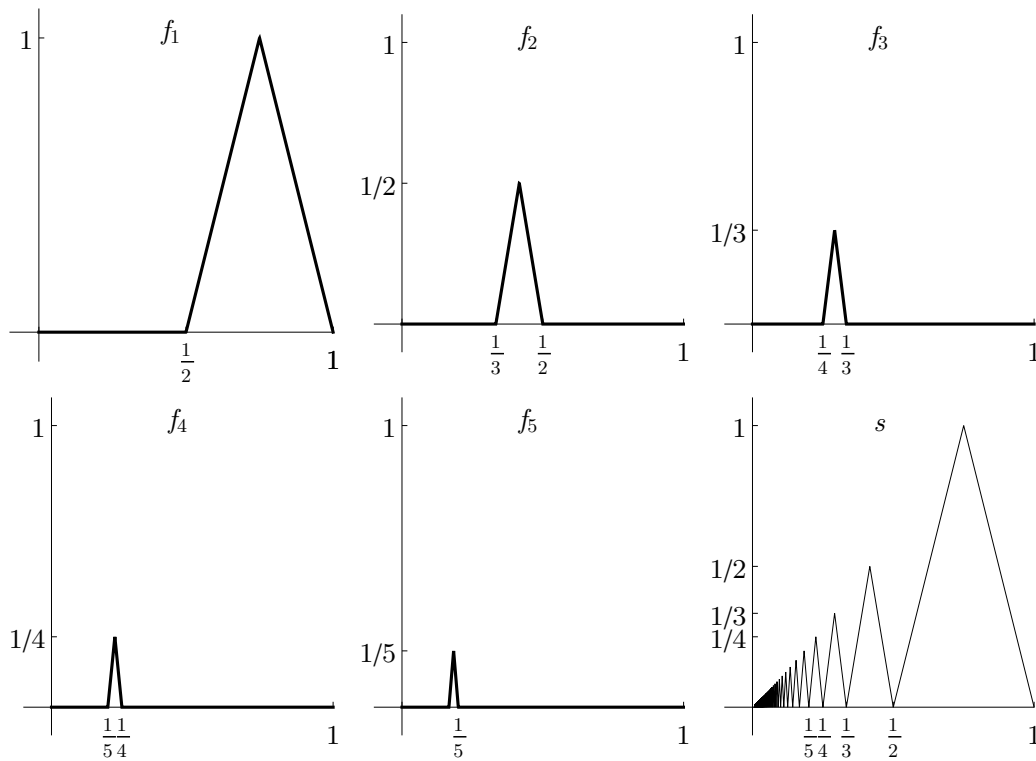


Figura 1.23: Una famiglia di funzioni continue $f_n: [0, 1] \rightarrow \mathbb{R}$ per le quali $\sum_n \sup|f_n| = +\infty$ ma che è sommabile. Le singole f_n sono funzioni con grafico a triangolo (tenda indiana, “tent function”), mentre la somma s ha infiniti triangoli (accampamento), Le cime delle tende sono allineate?

Una ulteriore via con cui possiamo convincerci che la serie armonica a segni alterni non è sommabile è che *la somma della serie dipende dall'ordine degli addendi*: per ogni $\alpha \in [-\infty, +\infty]$ esiste una permutazione degli indici, cioè una $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ biiettiva, tale che le somme parziali di $\sum_n (-1)^{n+1}/n$ convergano semplicemente ad α . In altri termini, per la convergenza semplice non vale la proprietà commutativa (per infiniti addendi): anche se φ è biiettiva, non è detto che $\sum_n a_n$ coincida con $\sum_n a_{\varphi(n)}$, se stiamo parlando di convergenza semplice.

È facile invece verificare (esercizio 1.7.5 p. 45) che se $\lambda \mapsto x_\lambda$ è una famiglia sommabile con somma s e $\varphi : \Lambda \rightarrow \Lambda$ è biiettiva, allora anche $\lambda \mapsto x_{\varphi(\lambda)}$ è sommabile con la stessa somma. In altre parole, la sommabilità e la relativa somma non dipendono da un eventuale ordinamento presente sull'insieme di indici Λ , oppure che *la somma nel senso della sommabilità è infinitamente commutativa*.

La sommabilità invece implica la convergenza semplice, con la stessa somma. Abbiamo già usato questa idea nella dimostrazione che la proprietà di Cauchy implica la sommabilità (Proposizione 1.7.10).

1.8 Sistemi ortonormali infiniti

Definizione 1.8.1. Una famiglia di vettori $\{u_\lambda : \lambda \in \Lambda\}$ in uno spazio con prodotto scalare si dice *sistema ortogonale* se i vettori sono a due a due ortogonali. Se poi hanno anche tutti norma 1 si parla di *sistema ortonormale*.

Le applicazioni della sommabilità che ci interesseranno riguardano famiglie ortogonali in spazi di Hilbert, per le quali vale il seguente criterio molto semplice.

Proposizione 1.8.2 (condizione di sommabilità di famiglie ortogonali). *Siano V uno spazio con prodotto scalare e $x_\lambda \in V$ una famiglia di vettori a due a due ortogonali. Allora la famiglia ha la proprietà di Cauchy per la sommabilità se e solo se*

$$\sum_{\lambda \in \Lambda} \|x_\lambda\|^2 < +\infty.$$

Se poi la famiglia x_λ è sommabile, allora

$$\left\| \sum_{\lambda \in \Lambda} x_\lambda \right\|^2 = \sum_{\lambda \in \Lambda} \|x_\lambda\|^2. \quad (1.56)$$

La formula (1.56) generalizza la (1.2) di pagina 8, e quindi si potrebbe chiamare “teorema di Pitagora infinito”.

Dimostrazione. Facciamo uso sistematico del “teorema di Pitagora” (1.2) di pagina 8. Supponiamo che $\lambda \mapsto x_\lambda$ abbia la proprietà di Cauchy. Per ogni $\varepsilon > 0$ sia G_ε come da definizione. Allora, se G è disgiunto da $G_{\sqrt{\varepsilon}}$,

$$\varepsilon = (\sqrt{\varepsilon})^2 > \left\| \sum_{\lambda \in G} x_\lambda \right\|^2 = \sum_{\lambda \in G} \|x_\lambda\|^2,$$

e quindi anche la famiglia scalare $\lambda \mapsto \|x_\lambda\|^2 \in \mathbb{R}$ è di Cauchy.

Viceversa, supponiamo che $\lambda \mapsto \|x_\lambda\|^2 \in \mathbb{R}$ sia di Cauchy. Sia $\varepsilon > 0$ e sia $G_\varepsilon \subseteq \Lambda$ finito tale che se $G \cap G_\varepsilon = \emptyset$ allora $\sum_{\lambda \in G} \|x_\lambda\|^2 < \varepsilon$. Poniamo $G'_\varepsilon := G_{\varepsilon^2}$. Se G è disgiunto da G'_ε allora

$$\left\| \sum_{\lambda \in G} x_\lambda \right\|^2 = \sum_{\lambda \in G} \|x_\lambda\|^2 < \varepsilon^2,$$

da cui, estraendo le radici quadrate,

$$\left\| \sum_{\lambda \in G} x_\lambda \right\| < \varepsilon,$$

il che dimostra che anche $\lambda \mapsto x_\lambda$ ha la proprietà di Cauchy per la sommabilità.

Supponiamo ora che la famiglia di vettori x_λ sia sommabile con somma s . Fissiamo $\varepsilon > 0$ e sia F_ε come da definizione. Sia $F \supset F_\varepsilon$ finito. Allora, usando il prodotto notevole $a^2 - b^2 = (a - b)(a + b)$, la formula (1.2) di pag. 8 e la disuguaglianza (1.8) di pag. 15,

$$\begin{aligned} \left| \|s\|^2 - \sum_{\lambda \in F} \|x_\lambda\|^2 \right| &= \left| \|s\|^2 - \left\| \sum_{\lambda \in F} x_\lambda \right\|^2 \right| = \\ &= \left| \|s\| - \left\| \sum_{\lambda \in F} x_\lambda \right\| \right| \cdot \left(\|s\| + \left\| \sum_{\lambda \in F} x_\lambda \right\| \right) \leq \\ &\leq \left\| s - \sum_{\lambda \in F} x_\lambda \right\| \cdot \left(\|s\| + \left\| s - s + \sum_{\lambda \in F} x_\lambda \right\| \right) \leq \end{aligned}$$

$$\begin{aligned} &\leq \left\| s - \sum_{\lambda \in F} x_\lambda \right\| \cdot \left(\|s\| + \|s\| + \left\| s - \sum_{\lambda \in F} x_\lambda \right\| \right) \leq \\ &\leq \varepsilon(2\|s\| + \varepsilon). \end{aligned}$$

Poiché l'ultimo membro si può rendere piccolo a piacere, è dimostrato che la famiglia scalare $\lambda \mapsto \|x_\lambda\|^2$ è sommabile (che fosse di Cauchy già lo sapevamo) con somma $\|s\|^2$. \square

Esercizio 1.8.3. Cosa succede all'Esempio 1.7.17 quando $p = 2$? Che ne è della somma delle f_n dell'esempio 1.7.18 se ambientate in $L^2([0, 1])$ invece che in $\mathcal{C}([0, 1])$? Se non supponiamo che la famiglia x_λ sia ortogonale, la condizione $\sum \|x_\lambda\|^2 < +\infty$ implica la sommabilità?

Esercizio 1.8.4. Riassumere in uno schema le implicazioni fra i seguenti concetti: sommabilità, condizione di Cauchy, convergenza assoluta, convergenza semplice delle somme parziali, somma delle norme quadre finita. Distinguere le situazioni in cui i vettori sono in dimensione finita, infinita, reali positivi.

La proposizione seguente verrà surclassata da un teorema successivo, ma la anticipiamo perché la dimostrazione è particolarmente semplice.

Proposizione 1.8.5. *Siano V uno spazio con prodotto scalare, $\{u_\lambda : \lambda \in \Lambda\}$ un sistema ortonormale, ed $x \in V$. Allora la somma*

$$\sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda \tag{1.57}$$

ha la proprietà di Cauchy per la sommabilità.

Dimostrazione. Sia F un sottinsieme finito di Λ e M_F il sottospazio vettoriale generato dai versori u_λ al variare di λ in F . La proposizione 1.6.4 p. 41 fornisce la formula per il punto \bar{y} di M_F più vicino a x , e inoltre $\bar{y} \perp x - \bar{y}$:

$$\begin{aligned} \sum_{\lambda \in F} \|\langle x, u_\lambda \rangle u_\lambda\|^2 &= \left\| \sum_{\lambda \in F} \langle x, u_\lambda \rangle u_\lambda \right\|^2 = \|\bar{y}\|^2 \leq \\ &\leq \|\bar{y}\|^2 + \|x - \bar{y}\|^2 = \|x\|^2 < +\infty. \end{aligned}$$

Poiché l'ultimo membro non dipende da $F \subseteq \Lambda$ finito, abbiamo che

$$\sum_{\lambda \in \Lambda} \|\langle x, u_\lambda \rangle u_\lambda\|^2 \leq \|x\|^2 < +\infty,$$

per cui la somma $\sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda$ verifica la condizione necessaria e sufficiente per la proprietà di Cauchy delle somme ortogonali (Proposizione 1.8.2). \square

Definizione 1.8.6. Dato uno spazio di Hilbert H e un sistema ortonormale $\{u_\lambda : \lambda \in \Lambda\}$, indicheremo con \hat{x} la funzione $\Lambda \rightarrow \mathbb{C}$ definita da $\hat{x}(\lambda) := \langle x, u_\lambda \rangle$. Lo scalare $\hat{x}(\lambda)$ sarà detto il *coefficiente di Fourier di x rispetto al sistema ortonormale*. In genere terremo sottinteso il sistema ortonormale nelle notazioni.

Proposizione 1.8.7. *Sia dato uno spazio di Hilbert H e un sistema ortonormale $\{u_\lambda : \lambda \in \Lambda\}$. Allora per ogni $x \in H$ la funzione \hat{x} appartiene a $\ell^2(\Lambda)$ e l'applicazione $x \mapsto \hat{x}$ è lineare, continua e suriettiva da H a $\ell^2(\Lambda)$. Più precisamente, vale la disuguaglianza $\|\hat{x}\|_{\ell^2} \leq \|x\|_H$.*

Dimostrazione. Usando le notazioni della dimostrazione precedente, vale la disuguaglianza

$$\begin{aligned} \|\hat{x}\|_{\ell^2}^2 &= \int_{\Lambda} |\hat{x}|^2 d\# = \sum_{\lambda \in \Lambda} |\hat{x}(\lambda)|^2 = \sup_F \sum_{\lambda \in F} |\langle x, u_\lambda \rangle|^2 = \\ &= \sup_F \sum_{\lambda \in F} \|\langle x, u_\lambda \rangle u_\lambda\|_H^2 = \sup_F \left\| \sum_{\lambda \in F} \langle x, u_\lambda \rangle u_\lambda \right\|_H^2 = \sup_F \|P_{M_F} x\|_H^2 \leq \\ &\leq \|x\|_H^2, \end{aligned}$$

da cui segue che l'applicazione $x \mapsto \hat{x}$ ha valori in $\ell^2(\Lambda)$, e, dato che è ovviamente lineare, anche la sua continuità.

Per vedere che è anche suriettiva, prendiamo $a \in \ell^2(\Lambda)$ e definiamo il vettore s come

$$s := \sum_{\lambda \in \Lambda} a_\lambda u_\lambda.$$

Questa somma è sommabile perché i vettori sono a due a due ortogonali e

$$\sum_{\lambda \in \Lambda} \|a_\lambda u_\lambda\|^2 = \sum_{\lambda \in \Lambda} |a_\lambda|^2 = \int_{\Lambda} |a|^2 d\# = \|a\|_{\ell^2}^2 < +\infty.$$

Prendiamo $\lambda_0 \in \Lambda$. Il funzionale $L_{\lambda_0} : H \rightarrow \mathbb{C}$ definito da $L_{\lambda_0} z = \langle z, u_{\lambda_0} \rangle$ è lineare e continuo. Quindi conserva la sommabilità e commuta con la somma infinita:

$$\langle s, u_{\lambda_0} \rangle = L_{\lambda_0} s = L_{\lambda_0} \sum_{\lambda \in \Lambda} a_\lambda u_\lambda = \sum_{\lambda \in \Lambda} a_\lambda L_{\lambda_0} u_\lambda = \sum_{\lambda \in \Lambda} a_\lambda \langle u_\lambda, u_{\lambda_0} \rangle.$$

Quest'ultima somma infinita ha addendi tutti nulli tranne (forse) quello con $\lambda = \lambda_0$, per cui

$$\langle s, u_{\lambda_0} \rangle = a_{\lambda_0} \cdot 1 = a_{\lambda_0}.$$

Questo mostra che $a = \hat{s}$. □

La disuguaglianza

$$\sum_{\lambda \in \Lambda} |\langle x, u_\lambda \rangle|^2 \leq \|x\|_H^2,$$

è anche chiamata *disuguaglianza di Bessel*, e generalizza quella del corollario 1.6.5 di pagina 43. Vedremo più avanti anche l'*identità* di Bessel.

Il lemma seguente riprende e precisa una parte della dimostrazione del teorema della minima distanza.

Lemma 1.8.8. *Sia H uno spazio di Hilbert, C un convesso non vuoto di H , $x \in H$, $d = \text{dist}(x, C)$ ed $\varepsilon > 0$. Allora esiste $\delta > 0$ tale che per ogni y_1, y_2*

$$\left. \begin{array}{l} y_1, y_2 \in C \\ \|x - y_1\| < d + \delta \\ \|x - y_2\| < d + \delta \end{array} \right\} \implies \|y_1 - y_2\| < \varepsilon.$$

Dimostrazione. Dati $y_1, y_2 \in C$, riprendiamo la disuguaglianza (1.19) di pag. 21, che deriva dall'identità del parallelogrammo e dal fatto che $(y_1 + y_2)/2 \in C$:

$$\|y_1 - y_2\|^2 \leq 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4d^2.$$

Sia $\delta > 0$ per ora generico e supponiamo che $\|x - y_1\| < d + \delta$ e $\|x - y_2\| < d + \delta$. Risulta

$$\|y_1 - y_2\|^2 < 2(d + \delta)^2 + 2(d + \delta)^2 - 4d^2 = 4(d + \delta)^2 - 4d^2.$$

Condizione sufficiente affinché $\|y_1 - y_2\| < \varepsilon$ è che $4(d + \delta)^2 - 4d^2 \leq \varepsilon^2$, che equivale a

$$(d + \delta)^2 \leq d^2 + \frac{\varepsilon^2}{4},$$

cioè a

$$\delta \leq \sqrt{d^2 + \frac{\varepsilon^2}{4}} - d.$$

Poiché $\sqrt{d^2 + \varepsilon^2/4} - d > 0$, esistono certamente $\delta > 0$ con la proprietà richiesta, e $\delta = \sqrt{d^2 + \varepsilon^2/4} - d$ è uno di questi. □

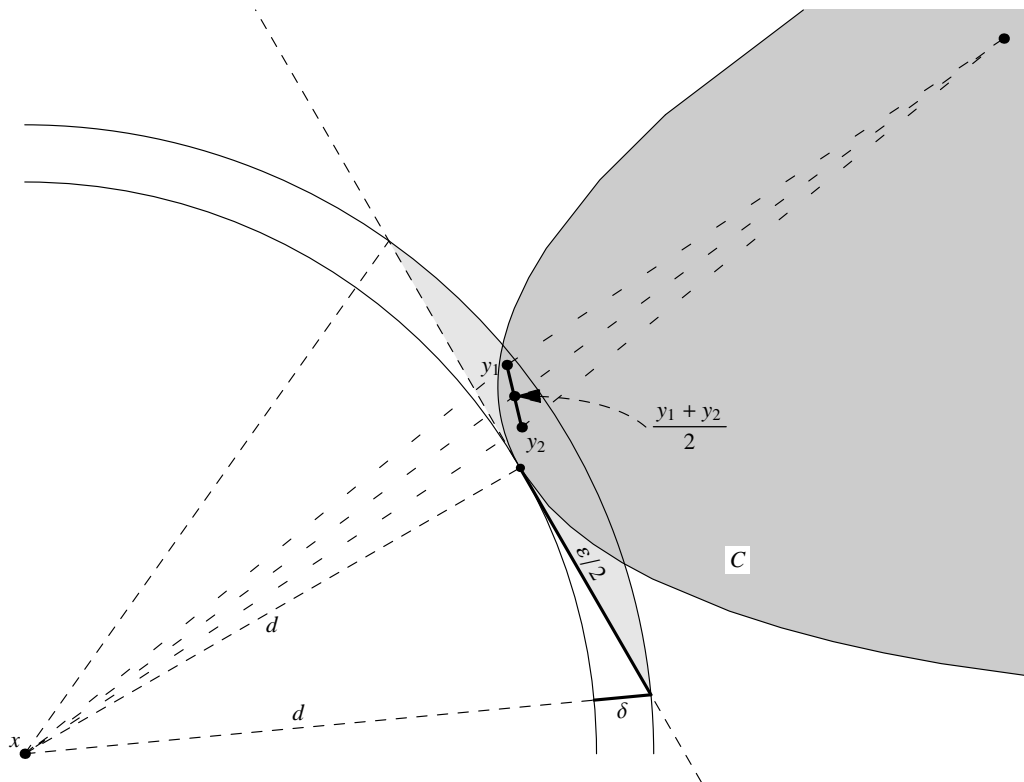


Figura 1.24: Illustrazione del lemma 1.8.8. Due punti del convesso C che si trovino nella corona circolare di spessore δ distano fra loro non più di ϵ . La relazione fra δ e ϵ si può ricavare dal triangolo rettangolo di cateti $d, \epsilon/2$ e ipotenusa $d + \delta$

Con un ragionamento geometrico sulla Figura 1.24 ci si può convincere che la scelta di $\delta = \sqrt{d^2 + \varepsilon^2/4} - d$ è ottimale.

Siamo finalmente pronti a generalizzare la formula della proiezione su un sottospazio a dimensione finita (Proposizione 1.6.4 p. 41).

Teorema 1.8.9 (proiezione ortogonale su un sottospazio chiuso qualsiasi). *Siano H uno spazio di Hilbert, $\{u_\lambda : \lambda \in \Lambda\}$ un sistema ortonormale, M la chiusura del sottospazio vettoriale generato dagli u_λ , ed $x \in H$. Allora la proiezione ortogonale di x su M è data dalla somma*

$$P_M x = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda, \quad (1.58)$$

intesa nel senso delle somme infinite di vettori.

Dimostrazione. Sia $\varepsilon > 0$, $d = \text{dist}(x, M)$, e sia $\delta > 0$ dato dalla proposizione precedente in corrispondenza con ε . Dato che $P_M x$ è un elemento di M , esiste una combinazione lineare finita z di vettori del sistema ortonormale che si avvicina a $P_M x$ meno di δ :

$$\|P_M x - \underbrace{(a_1 u_{\lambda_1} + \cdots + a_n u_{\lambda_n})}_{=:z}\| < \delta.$$

Per la disuguaglianza triangolare

$$\|x - z\| \leq \underbrace{\|x - P_M x\|}_{=d} + \underbrace{\|P_M x - z\|}_{<\delta} < d + \delta.$$

Poniamo $F_\varepsilon := \{\lambda_1, \dots, \lambda_n\}$, e sia F un insieme finito tale che $F_\varepsilon \subseteq F \subseteq \Lambda$. Sia M_F il sottospazio vettoriale generato da $\{u_\lambda : \lambda \in F\}$. In particolare $z \in M_F$. La proiezione ortogonale di x su M_F per definizione dista da x non più di quanto ne disti z :

$$\|x - P_{M_F} x\| \leq \|x - z\| < d + \delta.$$

D'altra parte $\|x - P_M x\| = d < d + \delta$. Ponendo $y_1 = P_M x$ e $y_2 = P_{M_F} x$, otteniamo che

$$\begin{cases} y_1 = P_M x \in M \\ y_2 = P_{M_F} x \in M_F \subseteq M \\ \|x - y_1\| < d + \delta \\ \|x - y_2\| < d + \delta \end{cases}$$

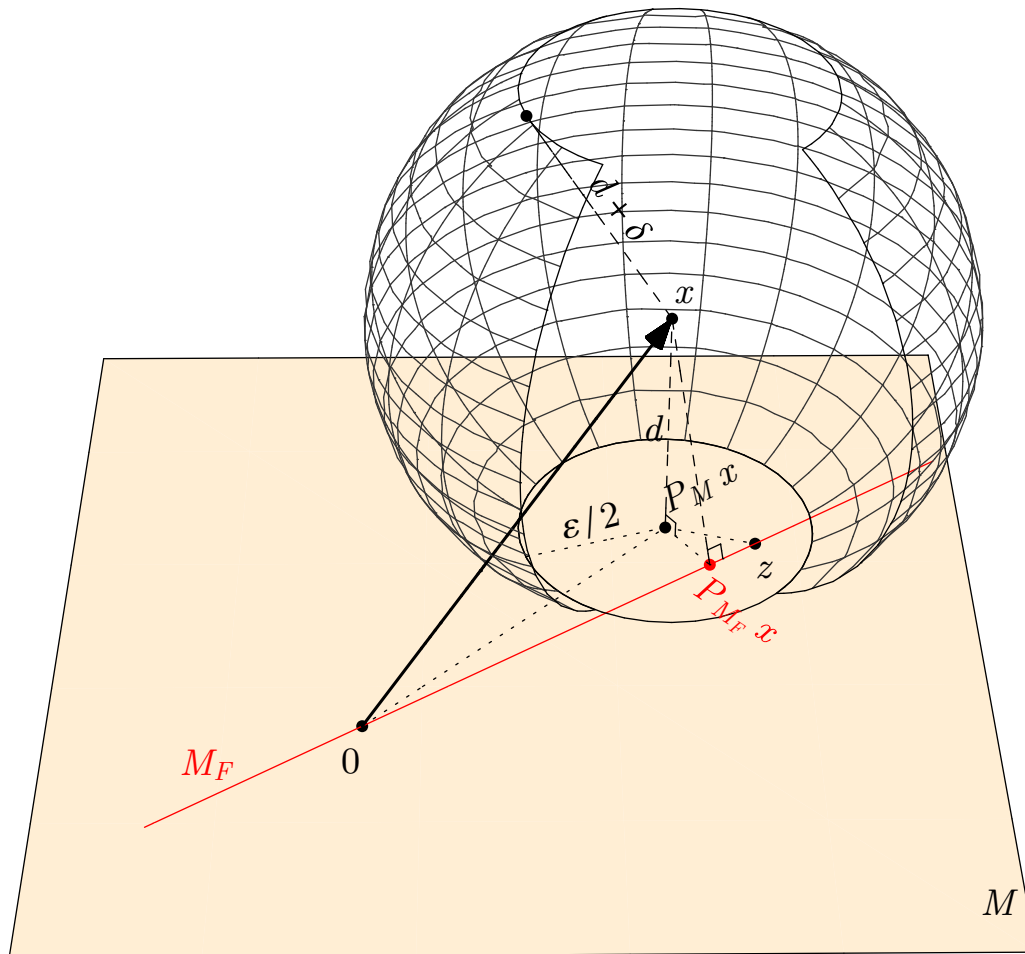


Figura 1.25: Nello spazio euclideo tridimensionale si possono seguire visivamente i vari passi della dimostrazione del teorema 1.8.9 della proiezione ortogonale. Il sottospazio M ha qui dimensione 2, M_F di dimensione 1. L'unica difficoltà è che le uniche possibili scelte per F sono $F = F_\varepsilon$ (quella fatta in figura) e $F = \Lambda$. Se vivessimo in dimensione 4 ci potremmo concedere di prendere $F_\varepsilon \subsetneq F \subsetneq \Lambda$

e quindi, per la proposizione precedente,

$$\|P_M x - P_{M_F} x\| < \varepsilon.$$

Essendo $\{u_\lambda : \lambda \in F\}$ una base ortonormale di M_F , vale la formula della proiezione ortogonale su un sottospazio a dimensione finita (Proposizione 1.6.4 p. 41):

$$P_{M_F} x = \sum_{\lambda \in F} \langle x, u_\lambda \rangle u_\lambda.$$

Abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste un F_ε finito $\subset \Lambda$ tale che per ogni F finito $\supseteq F_\varepsilon$ si ha che

$$\left\| P_M x - \sum_{\lambda \in F} \langle x, u_\lambda \rangle u_\lambda \right\| < \varepsilon.$$

Per definizione di sommabilità concludiamo che

$$P_M x = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda.$$

□

1.9 Basi hilbertiane

Teorema 1.9.1 (caratterizzazioni delle basi hilbertiane). *Siano H uno spazio di Hilbert e $\{u_\lambda : \lambda \in \Lambda\}$ un sistema ortonormale. Allora si equivalgono le condizioni seguenti:*

1. *il sistema ortonormale $\{u_\lambda : \lambda \in \Lambda\}$ è massimale nell'insieme di tutti i sistemi ortonormali in H , ordinato per inclusione;*
2. *$\text{span}\{u_\lambda : \lambda \in \Lambda\}$ è denso in H ;*
3. *per ogni $x \in H$ si ha che $x = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda$, nel senso delle somme infinite di vettori;*
4. *per ogni $x \in H$ vale l'uguaglianza $\|x\|_H = \|\hat{x}\|_{\ell^2(\Lambda)}$, cioè l'applicazione $x \mapsto \hat{x}$ è un'isometria di spazi normati;*
5. *per ogni $x, y \in H$ vale l'uguaglianza $\langle x, y \rangle_H = \langle \hat{x}, \hat{y} \rangle_{\ell^2(\Lambda)}$, cioè l'applicazione $x \mapsto \hat{x}$ è un'isometria di spazi con prodotto scalare.*

6. l'applicazione $x \mapsto \hat{x}$ è iniettiva da H a $\ell^2(\Lambda)$ (oltre a essere lineare, continua e suriettiva).

Dimostrazione. Chiamiamo M la chiusura del sottospazio vettoriale generato dagli u_λ .

Dimostriamo che $1 \Rightarrow 2$, o, che è lo stesso, la contronominale non $2 \Rightarrow$ non 1. Supponiamo pertanto che $\text{span}\{u_\lambda : \lambda \in \Lambda\}$ non sia denso in H , cioè che M sia un sottospazio *proprio* di H . Per il teorema di decomposizione ortogonale, $H = M \oplus M^\perp$. L'ortogonale M^\perp non può essere costituito dal solo zero. Prendiamo un $z \in M^\perp \setminus \{0\}$, che possiamo supporre avere norma 1. Questo z è un versore che è ortogonale a tutti i vettori di M . In particolare, $z \perp u_\lambda$ e $z \neq u_\lambda$ per ogni λ . Se aggiungiamo z alla famiglia $\{u_\lambda : \lambda \in \Lambda\}$ otteniamo una nuova famiglia ortonormale $\{z\} \cup \{u_\lambda : \lambda \in \Lambda\}$ (indicizzata opportunamente), che contiene *propriamente* quella di partenza. Concludiamo che il sistema $\{u_\lambda : \lambda \in \Lambda\}$ non è massimale per inclusione.

Dimostriamo che $2 \Rightarrow 3$. Supponiamo che $\text{span}\{u_\lambda : \lambda \in \Lambda\}$ sia denso in H . In particolare $M = H$, e quindi la proiezione ortogonale su M è semplicemente l'identità. Dal teorema 1.8.9 sulla proiezione ortogonale otteniamo che $x = P_H x = P_M x = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda$ per ogni $x \in H$.

Dimostriamo che $3 \Rightarrow 4$. Supponiamo che per ogni $x \in H$ valga $x = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda$. Poiché la somma è di vettori a due a due ortogonali, possiamo applicare il teorema di Pitagora generalizzato 1.8.2 e le norme quadre si distribuiscono sulla somma:

$$\begin{aligned} \|x\|_H^2 &= \left\| \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle u_\lambda \right\|_H^2 = \sum_{\lambda \in \Lambda} \|\langle x, u_\lambda \rangle u_\lambda\|_H^2 = \sum_{\lambda \in \Lambda} |\langle x, u_\lambda \rangle|^2 = \\ &= \sum_{\lambda \in \Lambda} |\hat{x}(\lambda)|^2 = \int_\Lambda |\hat{x}|^2 d\# = \\ &= \|\hat{x}\|_{\ell^2(\Lambda)}^2. \end{aligned}$$

Quindi l'applicazione $x \mapsto \hat{x}$ conserva le norme nei rispettivi spazi.

Dimostriamo che $4 \Rightarrow 5$. Supponiamo che l'applicazione $x \mapsto \hat{x}$ sia una isometria di spazi normati, cioè che $\|x\|_H = \|\hat{x}\|_{\ell^2(\Lambda)}$ per ogni $x \in H$. Ripeschiamo l'identità di polarizzazione (1.9) di pagina 16:

$$\langle x, y \rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4} + i \frac{\|ix + y\|^2 - \|ix - y\|^2}{4},$$

che vale in ogni spazio (complesso) con prodotto scalare, e quindi sia in H che in $\ell^2(\Lambda)$. Sfruttando il fatto che l'applicazione $x \mapsto \hat{x}$ è lineare sui complessi,

possiamo scrivere

$$\begin{aligned}
 \langle x, y \rangle_H &= \frac{\|x + y\|_H^2 - \|x - y\|_H^2}{4} + i \frac{\|ix + y\|_H^2 - \|ix - y\|_H^2}{4} = \\
 &= \frac{\|\widehat{x + y}\|_{\ell^2}^2 - \|\widehat{x - y}\|_{\ell^2}^2}{4} + i \frac{\|\widehat{ix + y}\|_{\ell^2}^2 - \|\widehat{ix - y}\|_{\ell^2}^2}{4} = \\
 &= \frac{\|\hat{x} + \hat{y}\|_{\ell^2}^2 - \|\hat{x} - \hat{y}\|_{\ell^2}^2}{4} + i \frac{\|i\hat{x} + \hat{y}\|_{\ell^2}^2 - \|i\hat{x} - \hat{y}\|_{\ell^2}^2}{4} = \\
 &= \langle \hat{x}, \hat{y} \rangle_{\ell^2}
 \end{aligned}$$

In generale, un'applicazione lineare che conservi la norma conserva necessariamente anche l'eventuale prodotto scalare che la genera. Il caso di spazi reali è analogo.

Dimostriamo che $5 \Rightarrow 1$, oppure, che è lo stesso, la contronominale non $1 \Rightarrow$ non 5 . Se la famiglia $\{u_\lambda : \lambda \in \Lambda\}$ non è massimale per inclusione fra le famiglie ortonormali, vuol dire che esiste almeno un versore $z \in H$ che è ortogonale a tutti gli u_λ . Ma allora $\hat{z}(\lambda) = \langle z, u_\lambda \rangle = 0$ per ogni λ , ossia $\hat{z} = 0 \in \ell^2(\Lambda)$. Quindi

$$\langle z, z \rangle_H = 1 \neq 0 = \langle 0, 0 \rangle_{\ell^2} = \langle \hat{z}, \hat{z} \rangle_{\ell^2}$$

e quindi l'applicazione $x \mapsto \hat{x}$ non conserva il prodotto scalare.

Dimostriamo che $4 \Rightarrow 6$. Supponiamo che $\|x\|_H = \|\hat{x}\|_{\ell^2(\Lambda)}$ per ogni x . In particolare, da $\hat{x} = 0$ segue che $x = 0$, cioè il nucleo dell'applicazione lineare $x \mapsto \hat{x}$ è composto dal solo zero. Concludiamo che l'applicazione è iniettiva.

Dimostriamo infine che $6 \Rightarrow 1$, oppure la contronominale non $1 \Rightarrow$ non 6 . Come prima, se $\{u_\lambda : \lambda \in \Lambda\}$ non è massimale, esiste almeno un versore $z \in H$ tale che $\hat{z} = 0$, dal che si deduce che l'applicazione lineare $x \mapsto \hat{x}$ non è iniettiva. \square

Quando valgono le condizioni del teorema, l'uguaglianza

$$\sum_{\lambda \in \Lambda} |\langle x, u_\lambda \rangle|^2 = \|x\|_H^2$$

è anche chiamata *identità di Bessel*, mentre l'uguaglianza

$$\langle x, y \rangle_H = \sum_{\lambda \in \Lambda} \langle x, u_\lambda \rangle \overline{\langle y, u_\lambda \rangle} = \langle \hat{x}, \hat{y} \rangle_{\ell^2}$$

è detta *identità di Parseval*.

Definizione 1.9.2. Dato uno spazio di Hilbert H , un sistema ortonormale $\{u_\lambda : \lambda \in \Lambda\}$ sarà detto una *base hilbertiana*, o sistema ortonormale completo, o insieme ortonormale massimale, se verifica le condizioni del precedente Teorema 1.9.1.

Una base hilbertiana di uno spazio di Hilbert permette di identificarlo con uno spazio standardizzato ℓ^2 . Ogni vettore dello spazio si potrà scrivere, in modo unico, come una combinazione lineare finita o infinita dei vettori della base.

Il concetto di base hilbertiana va tenuto distinto dal concetto di base vettoriale: dato uno spazio vettoriale V , una *base vettoriale* di V è una famiglia di vettori $\{x_\lambda : \lambda \in \Lambda\}$ tale che ogni vettore $x \in V$ si può scrivere *in modo unico* come combinazione lineare *finita* di vettori della base. Le basi vettoriali sono utili in dimensione finita, però diventano problematiche in dimensione infinita. Spero di non aver frainteso le mie letture se dico che una base vettoriale di uno spazio normato completo, ossia di uno spazio Banach, può essere o finita o più che numerabile, ma non può essere numerabile, e inoltre nel caso che la base sia più che numerabile la sua esistenza può essere dimostrata solo usando l'assioma di scelta. Si pensi per esempio allo spazio $\mathcal{C}([0, 1])$ delle funzioni continue su $[0, 1]$ con la norma del sup: non ne è nota nessuna base vettoriale concreta. Lo spazio dei polinomi a coefficienti reali o complessi è uno spazio vettoriale con una base vettoriale concreta numerabile (i monomi), ma non è completa per nessuna norma, per quanto ne so.

L'importanza delle basi hilbertiane è che nei casi di spazi concreti importanti, come L^2 con la misura di Lebesgue, sono ben note delle basi hilbertiane concrete e numerabili, con cui si possono fare dei calcoli significativi.

Proposizione 1.9.3. *Uno spazio di Hilbert ha una base hilbertiana finita o numerabile se e solo se è separabile, cioè se ammette un sottinsieme numerabile denso.*

Dimostrazione. Supponiamo che lo spazio di Hilbert H abbia una base hilbertiana numerabile $\{u_n : n \in \mathbb{N}\}$. Allora l'insieme delle combinazioni lineari finite a coefficienti razionali, o complessi razionali, è densa in H , e quindi H è separabile.

Viceversa, supponiamo di conoscere un sottinsieme numerabile denso $\{x_n : n \in \mathbb{N}\}$. Il procedimento di ortonormalizzazione di Gram-Schmidt produce un sistema ortonormale, che è anche completo, e verrà al più numerabile. \square

Quindi gli spazi di Hilbert separabili a dimensione infinita sono tutti isometrici fra loro e a $\ell^2(\mathbb{N})$. In generale, uno spazio di Hilbert è determinato a meno di isometria dalla cardinalità di una sua base hilbertiana.

Usando l'assioma di scelta si può dimostrare che ogni spazio di Hilbert, anche non separabile, ammette una base hilbertiana. La formulazione equivalente dell'assioma di scelta che usiamo qui è la seguente:

Teorema 1.9.4 (principio di massimalità di Hausdorff). *Ogni insieme parzialmente ordinato e non vuoto contiene un sottinsieme totalmente ordinato massimale.*

Teorema 1.9.5. *Ogni sistema ortonormale in uno spazio di Hilbert si può estendere a una base hilbertiana.*

Dimostrazione. Per semplicità qui tratteremo i sistemi ortonormali come insiemi di vettori, invece che come famiglie indicizzate, grazie al fatto che i vettori sono necessariamente tutti distinti. Se proprio vogliamo, un insieme può sempre venire indicizzato da se stesso.

Sia B un sistema ortonormale nello spazio di Hilbert H , e sia \mathcal{C} l'insieme di tutti i sistemi ortonormali in H che contengono B . Questo \mathcal{C} è parzialmente ordinato per inclusione. Per il principio di massimalità di Hausdorff, esiste un sottinsieme $\Omega \subseteq \mathcal{C}$ che è totalmente ordinato, cioè presi due sistemi ortonormali $A_1, A_2 \in \Omega$, uno dei due è sicuramente contenuto nell'altro, e inoltre Ω è massimale per inclusione, cioè non posso aggiungergli alcun sistema ortonormale in modo che rimanga totalmente ordinato. Gli elementi di Ω sono insiemi di vettori: raccogliamo in un insieme S tutti i vettori che appartengono a un sistema o a un altro di Ω .

Completate il lettore la dimostrazione, non difficile per chi ne ha viste di simili, che S è una base hilbertiana di H . \square

1.10 Basi hilbertiane esplicite

Lo spazio $\ell^2(\Lambda)$ ha una base hilbertiana esplicita indicizzata da Λ stesso: $u_{\lambda_0}(\lambda)$ vale 1 se $\lambda = \lambda_0$ e vale 0 altrimenti. Per esempio, se $\Lambda = \mathbb{N}$ possiamo pensare a u_n come alla infinitupla $(0, 0, \dots, 0, 1, 0, \dots)$, dove l'1 sta nella posizione n . Se $\Lambda = \{1, 2, \dots, n\}$ siamo riportati alla usuale base canonica di \mathbb{R}^n o \mathbb{C}^n .

Per lo spazio $L^2(\mathbb{R})$ è nota fin dal 1910 una base chiamata delle *ondine* (*wavelets*) di Haar. Queste si fondano sull'ondina madre $\varphi: \mathbb{R} \rightarrow \mathbb{R}$

$$\psi(t) := \begin{cases} 1 & \text{se } 0 \leq t < 1/2, \\ -1 & \text{se } 1/2 \leq t < 1, \\ 0 & \text{altrimenti,} \end{cases} \quad (1.59)$$

dalla quale, per traslazioni e omotetie si generano le ondine figlie

$$\psi_{n,m}(t) := 2^{n/2} \psi(2^n t - m).$$

La base delle ondine di Haar è $\{\psi_{n,m} : n, m \in \mathbb{Z}\}$, avente $\mathbb{Z} \times \mathbb{Z}$ come insieme di indici. Una variante di questa base aggiunge una seconda ondina madre

$$\varphi := \chi_{[0,1]}, \quad (1.60)$$

la quale genera per traslazione le figlie $\varphi_n(t) := \varphi(t - n)$. La nuova base è $\{\varphi_n : n \in \mathbb{Z}\} \cup \{\psi_{n,m} : n, m \in \mathbb{Z}, n \geq 0\}$. La base precedente, che usa le sole $\psi_{n,m}$, è più concisa ed elegante da scrivere, ma per rappresentare funzioni a gradino semplicissime come la φ ha bisogno di combinazioni lineari infinite di elementi della base. Si può osservare che le ondine $\psi_{n,m}$ hanno tutte integrale nullo, per cui una funzione di L^2 con integrale non nullo potrà essere approssimata dalle $\psi_{n,m}$ nella norma di L^2 ma mai nella norma di L^1 .

Per lo spazio $L^2([0, 2\pi])$ c'è la base trigonometrica originale di Fourier:

$$\left\{ \frac{1}{\sqrt{2\pi}} \sin nt : n = 1, 2, 3, \dots \right\} \cup \left\{ \frac{1}{\sqrt{2\pi}} \cos nt : n = 0, 1, 2, 3, \dots \right\}.$$

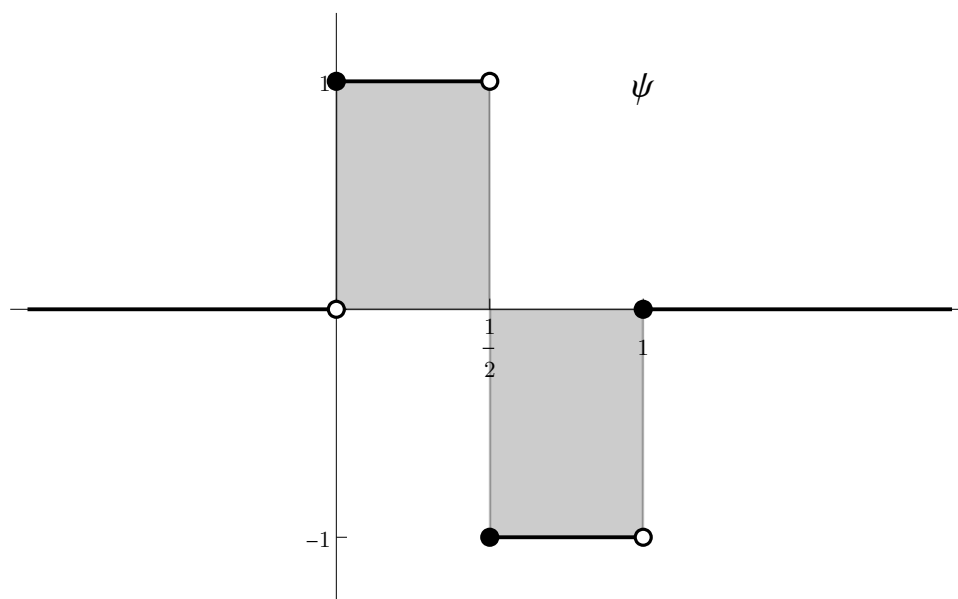


Figura 1.26: Grafico dell'ondina madre ψ di Haar, formula (1.59). È una funzione a gradino, ha tre discontinuità di salto, ha integrale nullo ed è un versore di $L^2(\mathbb{R})$

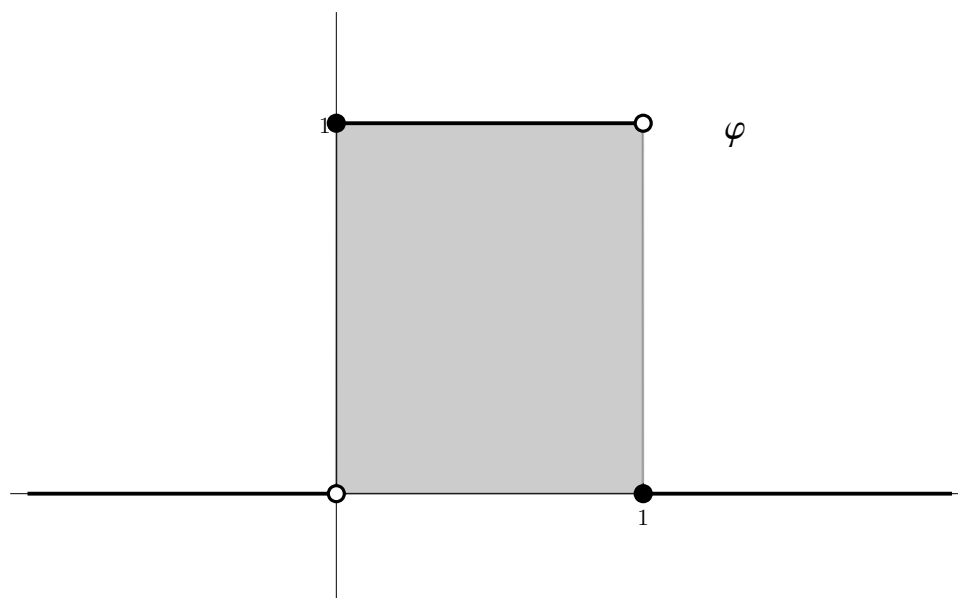


Figura 1.27: Grafico dell'ondina madre ausiliaria ϕ di Haar, formula (1.60). È una funzione a gradino, ha due discontinuità di salto, ha integrale 1 ed è un versore di $L^2(\mathbb{R})$

Capitolo 2

Insiemi compatti

2.1 Compatti in dimensione infinita

Definizione 2.1.1. Uno spazio topologico si dice *compatto* se ogni suo ricoprimento aperto ammette un sottoricoprimento finito.

La palla aperta di centro x e raggio r in uno spazio metrico X sarà indicata con $B(x, r)$ o con $B_X(x, r)$.

Definizione 2.1.2. Uno spazio metrico X è detto *totalmente limitato*, o *precompatto*, se per ogni $\varepsilon > 0$ si può ricoprire X con un numero finito di palle aperte di raggio ε , cioè esistono $x_1, \dots, x_n \in X$ tali che $X \subseteq \bigcup_{k=1}^n B(x_k, \varepsilon)$.

Proposizione 2.1.3. *Uno spazio metrico X è compatto se e solo se vale una delle condizioni equivalenti:*

1. *ogni successione in X ammette una sottosuccessione convergente (compattezza per successioni);*
2. *X è completo e totalmente limitato.*

Esercizio 2.1.4. Un sottinsieme di uno spazio metrico completo è precompatto se e solo se la sua chiusura è compatta; un precompatto è limitato; una funzione continua manda compatti in compatti; una funzione uniformemente continua manda successioni di Cauchy in successioni di Cauchy, completi in completi, precompatti in precompatti; una funzione lineare continua fra spazi normati manda successioni di Cauchy in successioni di Cauchy, completi in completi, precompatti in precompatti.

Sappiamo che i sottinsiemi chiusi e limitati di \mathbb{R}^n o di \mathbb{C}^n sono compatti. Questa proprietà *non* si estende a spazi normati di dimensione infinita. L'esempio più semplice è il seguente:

Esempio 2.1.5. Sia dato uno spazio di Hilbert con una base hilbertiana $\{u_n : n \in \mathbb{N}\}$ numerabile. Gli elementi u_n della base formano una successione limitata, perché fatta da versori, ma che non ha alcuna sottosuccessione convergente, perché ogni vettore dista $\sqrt{2}$ da ogni altro.

L'idea si generalizza a spazi normati generici:

Lemma 2.1.6. *Sia V uno spazio normato ed M un sottospazio vettoriale chiuso strettamente contenuto in V . Allora esiste un versore $\bar{z} \in V$ che dista almeno $1/2$ da ogni elemento di M .*

Dimostrazione. Si faccia riferimento alla Figura 2.1. Sia $\bar{y} \in V \setminus M$. Poiché M è chiuso, esiste un intorno di \bar{y} che non incontra M , e quindi $d := \text{dist}(\bar{y}, M) > 0$. Fissiamo $\varepsilon > 0$. Per definizione di distanza, esiste $\bar{x} \in M$ tale che $\|\bar{x} - \bar{y}\| < d + \varepsilon$, e poniamo

$$\bar{z} = \frac{\bar{y} - \bar{x}}{\|\bar{y} - \bar{x}\|},$$

che per costruzione è un versore. Sia $x \in M$ qualsiasi. Allora

$$\begin{aligned} \|x - \bar{z}\| &= \left\| x - \frac{\bar{y} - \bar{x}}{\|\bar{y} - \bar{x}\|} \right\| = \left\| \frac{\|\bar{y} - \bar{x}\|x - \bar{y} + \bar{x}}{\|\bar{y} - \bar{x}\|} \right\| = \\ &= \frac{1}{\underbrace{\|\bar{y} - \bar{x}\|}_{< d + \varepsilon}} \underbrace{\left\| (\|\bar{y} - \bar{x}\|x + \bar{x}) - \bar{y} \right\|}_{\substack{\geq d \\ \in M}} > \\ &> \frac{d}{d + \varepsilon}. \end{aligned}$$

Si può concludere prendendo $\varepsilon = d$. L'enunciato si poteva analogamente dimostrare più in generale con un qualsiasi numero positivo < 1 al posto di $1/2$. \square

Proposizione 2.1.7. *Sia V uno spazio normato di dimensione infinita. Allora in V esiste una successione limitata che non ha alcuna sottosuccessione di Cauchy.*

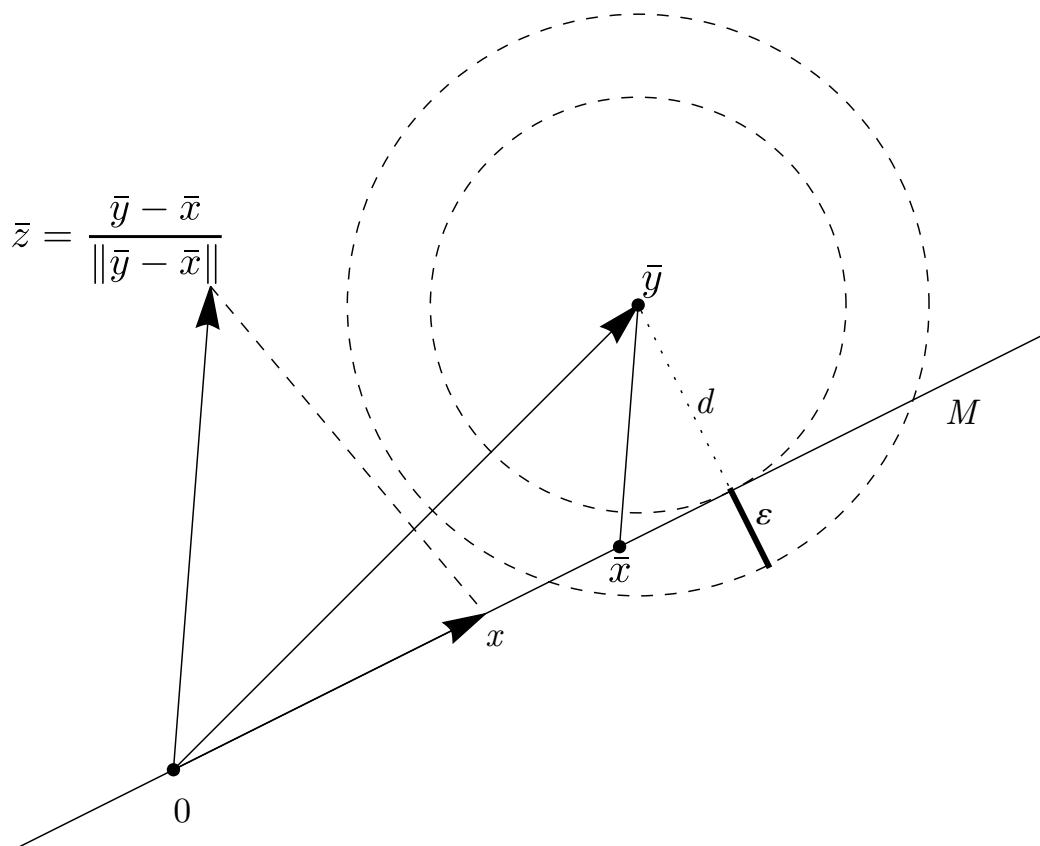


Figura 2.1: Illustrazione della dimostrazione del lemma 2.1.6. Dato un vettore \bar{y} fuori da M , si può ricavare un vettore \bar{z} che dista da M quanto si vuole

Dimostrazione. Per ipotesi esiste una successione x_n di vettori fra loro linearmente indipendenti, e sia M_n il sottospazio generato da x_1, \dots, x_n . Poniamo $\bar{y}_1 = x_1/\|x_1\|$. Essendo a dimensione finita, M_n è chiuso in M_{n+1} (Prop. 1.6.2 p. 39). Per il lemma precedente, esiste un versore $\bar{y}_{n+1} \in M_{n+1}$ che dista da M_n almeno $1/2$. Questa successione \bar{y}_n è limitata, perché i vettori hanno tutti norma 1, ma non ha alcuna sottosuccessione di Cauchy, perché ogni vettore dista da ogni altro almeno $1/2$. \square

Corollario 2.1.8. *In uno spazio normato a dimensione infinita, un sottinsieme aperto non vuoto contiene sempre una successione che non ha sottosuccessioni di Cauchy, e ogni compatto ha parte interna vuota. Se uno spazio normato ha dimensione infinita, non è omeomorfo a uno spazio \mathbb{C}^n o \mathbb{R}^n .*

Dimostrazione. Sia A un sottinsieme che contiene tutta una palla di centro x_0 e raggio $\varepsilon > 0$. Sia \bar{y}_n la successione di versori con distanze reciproche $\geq 1/2$, che abbiamo introdotto nella dimostrazione della proposizione precedente. Allora la successione di vettori $y_n = x_0 + \varepsilon y_n/2$ giace in A e gli elementi distano reciprocamente almeno $\varepsilon/4$, per cui non ha sottosuccessioni convergenti. Concludiamo che A non è compatto. \square

Lemma 2.1.9. *Sia V uno spazio normato, M un sottospazio vettoriale chiuso e $\bar{y} \in V \setminus M$. Sia $M' = \text{span } M \cup \{\bar{y}\} = M \oplus \mathbb{C}\bar{y}$ e siano $\varphi_1: M' \rightarrow M$ e $\varphi_2: M' \rightarrow \mathbb{C}$ tali che $x = \varphi_1(x) + \varphi_2(x)\bar{y}$ per ogni $x \in M'$. Allora φ_1 e φ_2 sono lineari e continue.*

Dimostrazione. La linearità delle due proiezioni è ben nota, per cui rimane da vedere la continuità. Come nel Lemma 2.1.6, abbiamo che $d := \text{dist}(\bar{y}, M) > 0$, come in Figura 2.2. Prendiamo $x \in M'$. Possiamo scrivere

$$x = \varphi_1(x) + \varphi_2(x)\bar{y}.$$

Se $\varphi_2(x) \neq 0$ allora, dividendo per $\varphi_2(x)$ e prendendo le norme,

$$\left\| \frac{x}{\varphi_2(x)} \right\| = \left\| \frac{\varphi_1(x)}{\varphi_2(x)} + \bar{y} \right\| = \left\| \bar{y} - \overbrace{\left(-\frac{\varphi_1(x)}{\varphi_2(x)} \right)}^{\in M} \right\| \geq d.$$

da cui

$$|\varphi_2(x)| \leq \frac{1}{d} \|x\|,$$

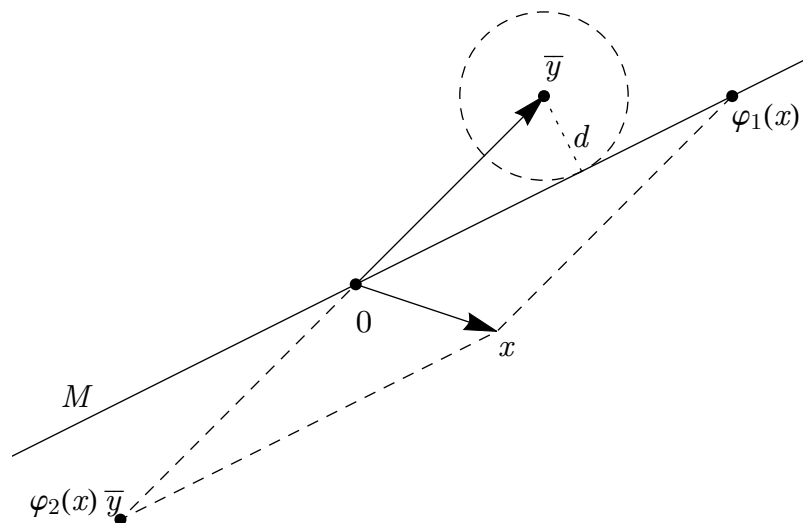


Figura 2.2: Continuità delle proiezioni

che dimostra che φ_2 è continua. Per φ_1 basta osservare che è la differenza di due funzioni continue, oppure, più per disteso, che

$$\begin{aligned} \|\varphi_1(x)\| &= \|x - \varphi_1(x)\bar{y}\| \leq \\ &\leq \|x\| + |\varphi_1(x)| \cdot \|\bar{y}\| \leq \\ &\leq \|x\| + \frac{1}{d}\|x\| \cdot \|\bar{y}\| = \\ &= \left(1 + \frac{1}{d}\right)\|x\|, \end{aligned}$$

per cui anche φ_1 è continua. □

Lemma 2.1.10. *I sottospazi a dimensione finita di uno spazio normato sono sempre completi, come spazi metrici.*

Dimostrazione. Se lo spazio normato ambiente V è completo, possiamo usare il fatto che i sottospazi chiusi di uno spazio metrico completo sono completi. Se non supponiamo che V sia completo, ragioniamo per induzione sulla dimensione del sottospazio M . Quando la dimensione del sottospazio M è zero, questo è un singolo, che è chiaramente completo. Supponiamo che

la proprietà sia vera ogniqualvolta M abbia dimensione n . Prendiamo un $\bar{y} \in V \setminus M$, $M' = \text{span } M \cup \{\bar{y}\}$ e φ_1, φ_2 le proiezioni sulle componenti, come nel lemma precedente. Sia $z_n \in M'$ una successione di Cauchy. Poiché φ_1, φ_2 sono (uniformemente) continue, anche $\varphi_1(x_n)$ e $\varphi_2(x_n)$ sono di Cauchy rispettivamente in M e negli scalari, e quindi convergono. Concludiamo che converge anche $z_n = \varphi_1(x_n) + \varphi_2(x_n)\bar{y}$. \square

Proposizione 2.1.11. *Ogni spazio normato a dimensione finita ha un omeomorfismo lineare su un \mathbb{C}^n o un \mathbb{R}^n . In particolare, in uno spazio normato a dimensione finita i chiusi e limitati sono compatti.*

Dimostrazione. Sia M uno spazio normato, per esempio sui complessi, con base finita u_1, \dots, u_n . Possiamo definire $\varphi: \mathbb{C}^n \rightarrow M$ come $\varphi(\alpha_1, \dots, \alpha_n) = \alpha_1 u_1 + \dots + \alpha_n u_n$. È chiaro che φ è lineare e biettiva. È anche continua, rispetto per esempio alla norma euclidea su \mathbb{C}^n , perché somma di funzioni continue. Lo spazio di partenza \mathbb{C}^n è notoriamente completo, ma per il lemma precedente anche quello di arrivo lo è. Per il teorema della mappa aperta concludiamo che φ è un omeomorfismo.

Sia K un sottinsieme chiuso e limitato di M . Allora $\varphi(K)$ è limitato in \mathbb{C}^n perché φ è lineare e continua, ed è chiuso perché φ è omeomorfismo. Ma allora $\varphi(K)$ è compatto, e infine $K = \varphi^{-1}(\varphi(K))$ è compatto in M perché immagine di un compatto tramite una funzione continua. \square

Se cerchiamo dei sottinsiemi compatti di uno spazio normato, possiamo annoverarci tutti i sottinsiemi chiusi e limitati che giacciono in un sottospazio a dimensione finita. Vedremo che esistono anche (pre)compatti che si spalmano su infinite dimensioni, sia pure con parte interna vuota.

2.2 Infinite estrazioni

Definizione 2.2.1. Data una successione $n \mapsto a_n$, chiameremo sua *sottosuccessione* ogni successione della forma $n \mapsto a_{\varphi(n)}$ con $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente.

È tradizionale scrivere n_k invece che $\varphi(k)$, e $k \mapsto a_{n_k}$ come sottosuccessione, ma la notazione con indici di indici non è adatta alle sotto-sotto-sotto- \dots -successioni con cui avremo a che fare.

È noto che se una successione converge, allora tutte le sue sottosuccessioni convergono, allo stesso limite. Più in generale, se a_n converge, per essere

sicuri che $a_{\varphi(n)}$ pure converga è sufficiente che la φ sia strettamente crescente e sia definita su un insieme del tipo $\{n \in \mathbb{N} : n \geq n_0\}$: potremmo chiamare questa una pseudosottosuccessione. Inoltre ogni successione limitata di scalari, o di vettori in uno spazio normato a dimensione finita, ammette sottosuccessioni convergenti. Se una successione è di Cauchy, anche tutte le sue sottosuccessioni sono di Cauchy.

Esercizio 2.2.2. Se $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ è strettamente crescente, allora $\varphi(n) \geq n$ per ogni $n \in \mathbb{N}$ e $\lim_{n \rightarrow +\infty} \varphi(n) = +\infty$. Cosa si può dire di una $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}$ strettamente crescente?

Esercizio 2.2.3. La composizione di funzioni strettamente crescenti è strettamente crescente.

Proposizione 2.2.4. *Sia data una matrice infinita $A_n(m)$, $n, m \in \mathbb{N}$, a valori reali o complessi, e limitata, cioè esista un $M > 0$ tale che $|A_n(m)| \leq M$ per ogni $n, m \in \mathbb{N}$. Allora esiste una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che per ogni $m \in \mathbb{N}$ la successione $n \mapsto A_{\varphi(n)}(m)$ converge.*

Visualizziamo la matrice infinita $A_n(m)$ prendendo come n l'indice di riga e m l'indice di colonna:

$$\begin{array}{cccccc}
 A_1(1) & A_1(2) & A_1(3) & A_1(4) & A_1(5) & \dots \\
 A_2(1) & A_2(2) & A_2(3) & A_2(4) & A_2(5) & \dots \\
 A_3(1) & A_3(2) & A_3(3) & A_3(4) & A_3(5) & \dots \\
 A_4(1) & A_4(2) & A_4(3) & A_4(4) & A_4(5) & \dots \\
 A_5(1) & A_5(2) & A_5(3) & A_5(4) & A_5(5) & \dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
 \end{array}$$

Se la matrice è limitata, la proposizione ci dice che possiamo cancellare delle righe della matrice in modo tale che, in quello che sopravvive, *ogni singola*

colonna sia convergente, ognuna a un qualche suo limite:

$$\begin{array}{cccccc}
 A_1(1) & A_1(2) & A_1(3) & A_1(4) & A_1(5) & \dots \\
 \\
 A_3(1) & A_3(2) & A_3(3) & A_3(4) & A_3(5) & \dots \\
 A_4(1) & A_4(2) & A_4(3) & A_4(4) & A_4(5) & \dots \\
 \\
 A_6(1) & A_6(2) & A_6(3) & A_6(4) & A_6(5) & \dots \\
 \\
 \\
 \\
 A_9(1) & A_9(2) & A_9(3) & A_9(4) & A_9(5) & \dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \\
 \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
 x_1 & x_2 & x_3 & x_4 & x_5 &
 \end{array}$$

Il valore $\varphi(n)$ dice che la riga n -esima nella matrice sopravvissuta (e ricompattata) proveniva dalla riga $\varphi(n)$ -esima nella matrice di partenza. Nell'esempio qui sopra $\varphi(1) = 1$, $\varphi(2) = 3$, $\varphi(3) = 4$, $\varphi(4) = 6$, $\varphi(5) = 9$.

Potremmo anche cancellare colonne e far convergere le righe, ma, per non confondere le idee, nel séguito imposteremo le cose in modo da far convergere sempre le colonne.

Dimostrazione. La successione scalare $n \mapsto A_n(1)$ è limitata. Esiste pertanto una $\varphi_1: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che la sottosuccessione $n \mapsto A_{\varphi_1(n)}(1)$ converge.

Anche la successione $n \mapsto A_{\varphi_1(n)}(2)$ è limitata, per cui esiste una $\varphi_2: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto A_{\varphi_1 \circ \varphi_2(n)}(2)$ converge. D'altra parte $n \mapsto A_{\varphi_1 \circ \varphi_2(n)}(1)$ è una sottosuccessione di $n \mapsto A_{\varphi_1(n)}(1)$, e quindi pure converge.

Anche la successione $n \mapsto A_{\varphi_1 \circ \varphi_2(n)}(3)$ è limitata, per cui esiste una $\varphi_3: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3(n)}(3)$ converge. D'altra parte $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3(n)}(m)$ è una sottosuccessione di $n \mapsto A_{\varphi_1 \circ \varphi_2(n)}(m)$, e quindi converge per $m \in \{1, 2\}$.

Anche la successione $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3(n)}(4)$ è limitata, per cui esiste una $\varphi_4: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3 \circ \varphi_4(n)}(4)$ converge. D'altra parte $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3 \circ \varphi_4(n)}(m)$ è una sottosuccessione di $n \mapsto A_{\varphi_1 \circ \varphi_2 \circ \varphi_3(n)}(m)$, e quindi converge per $m \in \{1, 2, 3\}$.

Procedendo in questo modo, è possibile associare ad ogni $k \in \mathbb{N}$ un'applicazione $\varphi_k: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente con la proprietà che per ogni $k \in \mathbb{N}$ la successione $n \mapsto A_{\varphi_1 \circ \dots \circ \varphi_k(n)}(m)$ converge per ogni $m \in \{1, 2, \dots, k\}$.

Per ogni $n \in \mathbb{N}$ definiamo ora

$$\varphi(n) := \varphi_1 \circ \varphi_2 \cdots \circ \varphi_n(n).$$

Notare che il numero di funzioni che vengono composte varia insieme con n . È chiaro che $\varphi: \mathbb{N} \rightarrow \mathbb{N}$. Verifichiamo che φ è strettamente crescente:

$$\begin{aligned} \varphi(n+1) &= \varphi_1 \circ \varphi_2 \cdots \circ \varphi_n \circ \varphi_{n+1}(n+1) = \\ &= \underbrace{\varphi_1 \circ \varphi_2 \cdots \circ \varphi_n}_{\text{strettamente cresc.}} \left(\underbrace{\varphi_{n+1}(n+1)}_{> \varphi_{n+1}(n)} \right) > \\ &> \varphi_1 \circ \varphi_2 \cdots \circ \varphi_n \left(\underbrace{\varphi_{n+1}(n)}_{\geq n} \right) \geq \\ &\geq \varphi_1 \circ \varphi_2 \cdots \circ \varphi_n(n) = \\ &= \varphi(n). \end{aligned}$$

Si è usato il fatto elementare che una $\psi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente necessariamente è tale che $\psi(n) \geq n$ per ogni n .

Verifichiamo che per ogni $k \in \mathbb{N}$ esiste una $\psi: \{k, k+1, k+2, \dots\} \rightarrow \mathbb{N}$ strettamente crescente tale che

$$\varphi(n) = \varphi_1 \circ \varphi_2 \cdots \circ \varphi_k(\psi(n)) \quad \forall n > k.$$

Per questo basta porre $\psi(k) = k$ e $\psi(n) := \varphi_{k+1} \circ \dots \circ \varphi_n(n)$ per $n > k$. Che questa ψ sia strettamente crescente si dimostra come fatto per φ .

Fissato $k \in \mathbb{N}$, concludiamo che $n \mapsto A_{\varphi(n)}(k)$ converge, in quanto (pseudo)sottosuccessione di $n \mapsto A_{\varphi_1 \circ \dots \circ \varphi_k(n)}(k)$ che converge per quanto detto prima. \square

Esercizio: la proposizione 2.2.4 vale anche se sostituiamo la parola “converge” con “è di Cauchy”. Si può anche supporre che ognuna delle colonne sia limitata separatamente, cioè che l'insieme $\{A_n(m) : n \in \mathbb{N}\}$ sia limitato per ogni m fissato.

Esercizio 2.2.5. Per ogni $k \in \mathbb{N}$ sia $\varphi_k: \mathbb{Z} \rightarrow \mathbb{Z}$ strettamente crescente. Poniamo $\varphi(n) := \varphi_1 \circ \varphi_2 \circ \dots \circ \varphi_n(n)$. Questa φ è strettamente crescente?

2.3 Il teorema di Ascoli-Arzelà

Dotiamo lo spazio $\mathcal{C}([0, 1])$ delle funzioni continue $[0, 1] \rightarrow \mathbb{R}$ della norma della convergenza uniforme $\|f\| = \|f\|_\infty := \sup_{[0,1]} |f|$. Gli insiemi precompatti in questo spazio si possono caratterizzare in modo abbastanza semplice.

Teorema 2.3.1 (di Ascoli-Arzelà). *Un sottinsieme K di $\mathcal{C}([0, 1])$ è precompatto se e solo se valgono le due condizioni seguenti:*

1. K è equilimitato, cioè esiste un $M > 0$ tale che $\|f\| \leq M$ per ogni $f \in K$;
2. K è equi-uniformemente continuo, cioè per ogni $\varepsilon > 0$ esiste $\delta_\varepsilon > 0$ tale che per ogni $x, y \in [0, 1]$ e per ogni $f \in K$ si abbia che $|x - y| < \delta_\varepsilon \implies |f(x) - f(y)| < \varepsilon$.

Dimostrazione. Supponiamo che K sia precompatto. È ovvio che K è limitato. Poiché K è anche totalmente limitato, fissiamo $\varepsilon > 0$ e siano $g_1, \dots, g_n \in K$ tali che $K \subseteq B(g_1, \varepsilon) \cup \dots \cup B(g_n, \varepsilon)$. Siano $\delta_1, \dots, \delta_n > 0$ i moduli di uniforme continuità delle rispettive funzioni g_1, \dots, g_n . Posto $\delta_\varepsilon := \min\{\delta_1, \dots, \delta_n\}$ è chiaro che $\delta_\varepsilon > 0$ e che

$$\forall k = 1, \dots, n \quad \forall x, y \in [0, 1] \quad |x - y| < \delta_\varepsilon \implies |g_k(x) - g_k(y)| < \varepsilon.$$

Sia ora $f \in K$. Esiste un indice $k_{\varepsilon, f} \in \{1, \dots, n\}$ tale che $\|f - g_{k_{\varepsilon, f}}\| < \varepsilon$. Ma allora

$$\begin{aligned} \forall x, y \in [0, 1] \quad |x - y| < \delta_\varepsilon \leq \delta_{k_{\varepsilon, f}} &\implies |f(x) - f(y)| \leq \\ &\leq \underbrace{|f(x) - g_{k_{\varepsilon, f}}(x)|}_{\leq \|f - g_{k_{\varepsilon, f}}\|} + \underbrace{|g_{k_{\varepsilon, f}}(x) - g_{k_{\varepsilon, f}}(y)|}_{< \varepsilon \text{ per unif. cont.}} + \underbrace{|g_{k_{\varepsilon, f}}(y) - f(y)|}_{\leq \|f - g_{k_{\varepsilon, f}}\|} < \\ &< \|f - g_{k_{\varepsilon, f}}\| + \varepsilon + \|f - g_{k_{\varepsilon, f}}\| \leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon. \end{aligned}$$

Questo dimostra che K è equiuniformemente continuo.

Viceversa, supponiamo di sapere che K è equilimitato ed equiuniformemente continuo, coi parametri dell'enunciato. Data una successione $f_n \in K$, vogliamo dimostrare che esiste una sottosuccessione che converge uniformemente.

Sia $T := \{x(k) : k \in \mathbb{N}\}$ una successione densa in $[0, 1]$. La matrice scalare infinita $(n, k) \mapsto A_n(k) := f_n(x(k))$ ha elementi limitati. Sappiamo quindi

che esiste $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ tale che la successione $n \mapsto f_{\varphi(n)}(x(k))$ converge per ogni k fissato. Dimostriamo che $n \mapsto f_{\varphi(n)}$ è di Cauchy per la convergenza uniforme. Sia $\varepsilon > 0$ e per ogni $x \in T$ sia $N_{x,\varepsilon} \in \mathbb{N}$ tale che

$$\forall n, m \in \mathbb{N} \quad n, m > N_{x,\varepsilon} \implies |f_{\varphi(n)}(x) - f_{\varphi(m)}(x)| < \varepsilon. \quad (2.1)$$

Sia $\delta_\varepsilon > 0$ dato dall'equiuniforme continuità. Per la densità di T , esiste un sottinsieme *finito* $T_\varepsilon \subset T$ tale che per ogni $x \in [0, 1]$ c'è un $x_\varepsilon \in T_\varepsilon$ tale che $|x - x_\varepsilon| < \delta$. Poniamo

$$N_\varepsilon := \max\{N_{x,\varepsilon} : x \in T_\varepsilon\} < +\infty.$$

Siano ora $n, m > N_\varepsilon$. Per ogni $x \in [0, 1]$ stimiamo

$$\begin{aligned} |f_{\varphi(n)}(x) - f_{\varphi(m)}(x)| &\leq \overbrace{|f_{\varphi(n)}(x) - f_{\varphi(n)}(x(\varepsilon))|}^{< \varepsilon \text{ per l'equiunif. cont.}} + \\ &+ \underbrace{|f_{\varphi(n)}(x(\varepsilon)) - f_{\varphi(m)}(x(\varepsilon))|}_{< \varepsilon \text{ per la cond. di Cauchy (2.1)}} + \underbrace{|f_{\varphi(m)}(x(\varepsilon)) - f_{\varphi(m)}(x)|}_{< \varepsilon \text{ per l'equiunif. cont.}} < 3\varepsilon. \end{aligned}$$

Poiché $x \in [0, 1]$ è generico, abbiamo che $\|f_{\varphi(n)} - f_{\varphi(m)}\| < 3\varepsilon$. È pertanto dimostrato che la successione $n \mapsto f_{\varphi(n)}$ è di Cauchy per la convergenza uniforme. \square

Usando il teorema di Ascoli-Arzelà si possono esibire facilmente numerosi insiemi precompatti nello spazio delle funzioni continue.

Esercizio 2.3.2. Dimostrare che è (pre?)compatto l'insieme delle funzioni $[0, 1] \rightarrow \mathbb{R}$ che sono limitate da una costante $M > 0$ e lipschitziane di costante $L > 0$. Dimostrare che questo compatto non è contenuto in nessun sottospazio a dimensione finita.

Per fissare le idee abbiamo ambientato il teorema di Ascoli-Arzelà nello spazio delle funzioni da $[0, 1]$ a \mathbb{R} , ma la dimostrazione si generalizza facilmente: al posto di $[0, 1]$ possiamo prendere un qualsiasi spazio metrico X che sia (pre)compatto e come spazio di arrivo va bene un qualsiasi spazio normato a dimensione finita.

Esercizio 2.3.3. Data una successione di funzioni $f_n: \mathbb{R} \rightarrow \mathbb{R}$ equilimitata, dimostrare che esiste una sottosuccessione f_{n_k} tale che $\lim_{k \rightarrow +\infty} f_{n_k}(x)$ esiste per ogni $x \in \mathbb{Q}$. Se le funzioni sono continue, la sottosuccessione necessariamente converge uniformemente su tutto \mathbb{R} ? E cosa succede se sono equiuniformemente continue?

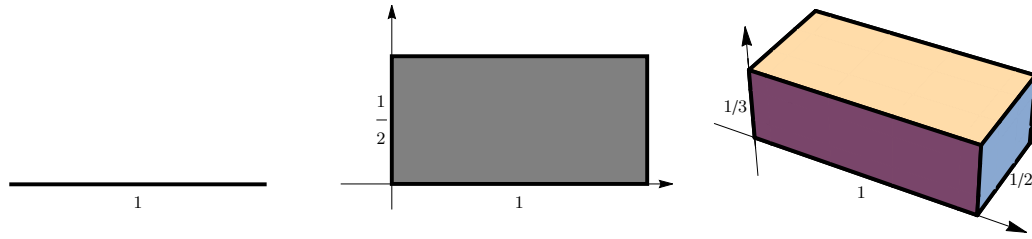


Figura 2.3: Il cubo di Hilbert in $\ell^2(\mathbb{N})$ proiettato ortogonalmente sui sottospazi $\text{span}\{u_1\}$, $\text{span}\{u_1, u_2\}$, $\text{span}\{u_1, u_2, u_3\}$

2.4 Il cubo di Hilbert

Definizione 2.4.1. Chiamiamo *cubo di Hilbert* l'insieme delle successioni $x: \mathbb{N} \rightarrow \mathbb{R}$ tali che $x(m) \in [0, 1/m]$ per ogni $m \in \mathbb{N}$. Oppure anche una sua variante, per esempio l'insieme delle $x: \mathbb{N} \rightarrow \mathbb{C}$ tali che $|x(m)| \leq 1/m$.

Il cubo di Hilbert, che indicheremo qui con C , è contenuto nello spazio di Banach $\ell^p(\mathbb{N})$, se $1 < p < +\infty$. Verifichiamo che C è un compatto in queste topologie. Innanzitutto C è chiuso in $\ell^p(\mathbb{N})$. Infatti se $n \mapsto x_n$ è una successione a valori in C e che converge a un $\bar{x} \in \ell^p(\mathbb{N})$, prendendo la componente m -esima possiamo scrivere

$$|x_n(m) - \bar{x}(m)|^p \leq \sum_{k \in \mathbb{N}} |x_n(k) - \bar{x}(k)|^p = \|x_n - \bar{x}\|_p^p \rightarrow 0,$$

da cui $\lim_{n \rightarrow +\infty} x_n(m) = \bar{x}(m)$. In generale, la convergenza in ℓ^p implica la convergenza puntuale, componente per componente. Poiché $x_n(m) \in [0, 1/m]$ per ogni n, m , passando al limite per $n \rightarrow +\infty$ risulta che anche $\bar{x}(m) \in [0, 1/m]$ per ogni m , il che mostra che $\bar{x} \in C$.

Prendiamo poi una successione $n \mapsto x_n \in C$, che possiamo visualizzare come matrice infinita:

$$\begin{aligned} x_1 &= (x_1(1), x_1(2), x_1(3), x_1(4), x_1(5), \dots), \\ x_2 &= (x_2(1), x_2(2), x_2(3), x_2(4), x_2(5), \dots), \\ x_3 &= (x_3(1), x_3(2), x_3(3), x_3(4), x_3(5), \dots), \\ x_4 &= (x_4(1), x_4(2), x_4(3), x_4(4), x_4(5), \dots), \\ &\dots \end{aligned}$$

Teniamo presente che la prima colonna dopo i segni = è compresa in $[0, 1]$, la seconda in $[0, 1/2]$, la terza in $[0, 1/3]$ eccetera. Dimostriamo che $n \mapsto x_n$ ha una sottosuccessione convergente. La matrice a due indici $x_n(m)$ è limitata. Quindi esiste $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che per ogni $m \in \mathbb{N}$ esiste finito $\bar{x}(m) = \lim_{n \rightarrow +\infty} x_{\varphi(n)}(m)$. Come prima, $\bar{x}(m) \in [0, 1/m]$ per ogni m , per cui $\bar{x} \in \ell^p(\mathbb{N})$. Rimane da dimostrare che $\|x_n - \bar{x}\|_p \rightarrow 0$ per $n \rightarrow +\infty$.

Fissiamo $\varepsilon > 0$. Poiché la serie numerica $\sum_{m \in \mathbb{N}} m^{-p}$ converge, le sue “code” si possono rendere piccole a piacere, per cui esiste $M_\varepsilon \in \mathbb{N}$ tale che $\sum_{m > M_\varepsilon} m^{-p} < \varepsilon^p/2$.

A causa della convergenza puntuale di ogni singola colonna, per ogni $m \in \mathbb{N}$ esiste $N_{\varepsilon, m} \in \mathbb{N}$ tale che per ogni $n \geq N_{\varepsilon, m}$ vale $|x_{\varphi(n)}(m) - \bar{x}(m)| < \varepsilon/(2M_\varepsilon)^{1/p}$ (il fattore M_ε al denominatore è cruciale, come vedremo). Poniamo

$$N_\varepsilon = \max\{N_{\varepsilon, 1}, N_{\varepsilon, 2}, \dots, N_{\varepsilon, M_\varepsilon}\} < +\infty.$$

Allora, se $n \geq N_\varepsilon$ possiamo fare la stima seguente, che separa la serie nei primi M_ε termini, ognuno dei quali si magiora con la definizione di limite, e la coda, che è piccola indipendentemente da n perché siamo nel cubo di Hilbert:

$$\begin{aligned} \|x_{\varphi(n)} - \bar{x}\|_p^p &= \sum_{m \in \mathbb{N}} |x_{\varphi(n)}(m) - \bar{x}(m)|^p = \\ &= \sum_{m=1}^{M_\varepsilon} |x_{\varphi(n)}(m) - \bar{x}(m)|^p + \sum_{m=M_\varepsilon+1}^{+\infty} \underbrace{|x_{\varphi(n)}(m) - \bar{x}(m)|^p}_{\in [-1/m, 1/m]} < \\ &< \sum_{m=1}^{M_\varepsilon} |x_{\varphi(n)}(m) - \bar{x}(m)|^p + \sum_{m=M_\varepsilon+1}^{+\infty} (1/m)^p = \\ &= \sum_{m=1}^{M_\varepsilon} \frac{\varepsilon^p}{2M_\varepsilon} + \sum_{m=M_\varepsilon+1}^{+\infty} m^{-p} < \\ &< M_\varepsilon \frac{\varepsilon^p}{2M_\varepsilon} + \frac{\varepsilon^p}{2} = \varepsilon^p. \end{aligned}$$

Se non avessimo preso la precauzione di dividere ε^p per M_ε avremmo una maggiorazione con un termine del tipo $M_\varepsilon \varepsilon^p$, di cui non si sa dire gran che, essendo il prodotto di un numero grande per uno piccolo.

Il cubo di Hilbert è un esempio particolarmente semplice di un compatto in dimensione infinita che non è contenuto in alcun sottospazio a dimensione finita.

Ogni spazio di Hilbert di dimensione infinita contiene una copia del cubo di Hilbert. Sia infatti $\{u_n : n \in \mathbb{N}\}$ un sistema ortonormale numerabile (non necessariamente massimale). Allora l'insieme dei vettori della forma $\sum_{n \in \mathbb{N}} x(n)u_n$ con x che varia nel cubo di Hilbert è un compatto.

E se volessimo un compatto in $\ell^1(\mathbb{N})$ o in $\ell^\infty(\mathbb{N})$? Intendo non banale, non contenuto in un sottospazio di dimensione finita.

Esercizio 2.4.2. Piccola variante del cubo di Hilbert: l'insieme delle successioni $x: \mathbb{N} \rightarrow \mathbb{C}$ tali che $|x(n)| \leq 1/n$ per ogni $n \in \mathbb{N}$. Mostrare che anche questa variante è compatta in ℓ^p se $1 < p < +\infty$.

Esercizio 2.4.3. Sia V uno spazio di Banach, e $u_n \in V$ una successione di versori. Sia C l'insieme dei vettori del tipo $\sum_{n \in \mathbb{N}} \alpha_n u_n$, con $\alpha_n \in \mathbb{C}$, $|\alpha_n| \leq 1/n^2$ per ogni n . Mostrare che C è compatto.

Esercizio 2.4.4. Esibire un compatto non banale di $\ell^1(\mathbb{N})$.

Capitolo 3

Operatori compatti e aggiunti

3.1 Operatori compatti

Definizione 3.1.1. Dati due spazi normati X, Y e un operatore lineare $A: X \rightarrow Y$, diremo che A è *compatto* se l'immagine tramite A di ogni sottinsieme limitato di X è precompatta in Y .

Esercizio 3.1.2. Affinché un operatore sia compatto è sufficiente che sia precompatta l'immagine di una singola palla di raggio > 0 , non importa di che centro, chiusa o aperta.

Esercizio 3.1.3. L'operatore identità $X \rightarrow X$ è compatto se e solo se X ha dimensione finita.

Proposizione 3.1.4. *Un operatore è compatto se e solo se da ogni successione $n \mapsto x_n$ limitata in X posso estrarre una sottosuccessione $k \mapsto x_{n_k}$ tale che $k \mapsto Ax_{n_k}$ sia di Cauchy in Y .*

Proposizione 3.1.5. *Un operatore compatto è continuo.*

Dimostrazione. Quando $A: X \rightarrow Y$ è compatto manda limitati in precompatti. In particolare manda limitati in limitati, il che per un operatore lineare è condizione necessaria e sufficiente per essere continuo. \square

Proposizione 3.1.6. *Somma di operatori compatti è compatta.*

Dimostrazione. Siano $A, B: X \rightarrow Y$ compatti. Sia x_n limitata in X . Poiché A è compatto esiste una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto Ax_{\varphi(n)}$ è di Cauchy. Poiché B è compatta e $n \mapsto x_{\varphi(n)}$ è limitata, esiste

una $\psi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto Bx_{\varphi \circ \psi(n)}$ è di Cauchy. È chiaro che anche la successione $n \mapsto (A+B)x_{\varphi \circ \psi(n)} = Ax_{\varphi \circ \psi(n)} + Bx_{\varphi \circ \psi(n)}$ è di Cauchy, in quanto somma di successioni di Cauchy. \square

Proposizione 3.1.7. *Siano X, Y, Z spazi normati e $A: X \rightarrow Y$ e $B: Y \rightarrow Z$ lineari continui. Se uno dei due è compatto, allora anche la composizione BA è compatta.*

Dimostrazione. Si può fare riferimento al diagramma seguente:

$$X \xrightarrow{A} Y \xrightarrow{B} Z$$

Supponiamo che A sia compatto. Sia x_n limitata in X . Esiste una sottosuccessione tale che $k \mapsto Ax_{n_k}$ è di Cauchy. Ma allora anche $k \mapsto BAx_{n_k}$ è di Cauchy, perché gli operatori lineari (uniformemente) continui mandano successioni di Cauchy in successioni di Cauchy. Quindi la composizione AB è compatta.

Supponiamo che B sia compatto. Sia x_n limitata in X . Poiché A è continua, anche Ax_n è limitata, in Y . Poiché B è compatto, esiste una sottosuccessione tale che $k \mapsto BAx_{n_k}$ è di Cauchy. Di nuovo AB è compatta. \square

L'insieme $\mathcal{L}(X, Y)$ degli operatori lineari continui da uno spazio normato X in un altro Y è esso stesso uno spazio normato con la norma operatoriale:

$$\|A\|_{\mathcal{L}(X, Y)} = \sup_{x \in X \setminus \{0\}} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{x \in X, \|x\| \leq 1} \|Ax\|_Y = \sup_{x \in X, \|x\|=1} \|Ax\|_Y.$$

È noto che quando Y è completo, cioè di Banach, anche $\mathcal{L}(X, Y)$ lo è.

Proposizione 3.1.8. *L'insieme degli operatori compatti $X \rightarrow Y$ è un sotto-spazio vettoriale chiuso di $\mathcal{L}(X, Y)$.*

Dimostrazione. Sappiamo già che l'insieme degli operatori compatti è stabile per somma, ed è chiaro che è stabile anche per prodotto per uno scalare. Resta da vedere che è chiuso.

Sia A_n una successione di operatori compatti, e A continuo, tali che $A_n \rightarrow A$ nella norma operatoriale. Vogliamo dimostrare che anche A è compatto. Sia x_n una successione limitata in X , per esempio $\|x_n\| \leq c < +\infty$ per

ogni n . Applichiamo le A_n alla successione x_m disponendo i valori in una matrice a valori in Y :

$$\begin{array}{cccccc} A_1x_1 & A_2x_1 & A_3x_1 & A_4x_1 & \dots \\ A_1x_2 & A_2x_2 & A_3x_2 & A_4x_2 & \dots \\ A_1x_3 & A_2x_3 & A_3x_3 & A_4x_3 & \dots \\ A_1x_4 & A_2x_4 & A_3x_4 & A_4x_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Poiché ciascuna delle A_n è compatta, da ogni colonna della matrice, cioè da ogni successione $m \mapsto A_n x_m$, possiamo estrarre una sottosuccessione di Cauchy. Col procedimento diagonale possiamo estrarre una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che la successione $m \mapsto A_n x_{\varphi(m)}$ è di Cauchy per ogni $n \in \mathbb{N}$.

Dimostriamo che anche la successione $m \mapsto Ax_{\varphi(m)}$ è di Cauchy. Fissiamo $\varepsilon > 0$. Poiché $A_n \rightarrow A$, esiste $N_\varepsilon \in \mathbb{N}$ tale che $\|A_n - A\| < \varepsilon/(4c)$ per ogni $n > N_\varepsilon$. Fissiamo un $n_0 > N_\varepsilon$. La successione $m \mapsto A_{n_0} x_{\varphi(m)}$ è di Cauchy. Quindi esiste $M_\varepsilon \in \mathbb{N}$ tale che per ogni $m_1, m_2 \in \mathbb{N}$ si ha che

$$m_1, m_2 > M_\varepsilon \implies \|A_{n_0} x_{\varphi(m_1)} - A_{n_0} x_{\varphi(m_2)}\| < \frac{\varepsilon}{2}.$$

Allora se $m_1, m_2 > M_\varepsilon$ possiamo scrivere

$$\begin{aligned} \|Ax_{\varphi(m_1)} - Ax_{\varphi(m_2)}\| &\leq \|Ax_{\varphi(m_1)} - A_{n_0}x_{\varphi(m_1)}\| + \\ &\quad + \|A_{n_0}x_{\varphi(m_1)} - A_{n_0}x_{\varphi(m_2)}\| + \\ &\quad + \|A_{n_0}x_{\varphi(m_2)} - Ax_{\varphi(m_2)}\| = \\ &= \|(A - A_{n_0})x_{\varphi(m_1)}\| + \\ &\quad + \|A_{n_0}x_{\varphi(m_1)} - A_{n_0}x_{\varphi(m_2)}\| + \\ &\quad + \|(A_{n_0} - A)x_{\varphi(m_2)}\| \leq \\ &\leq \|A - A_{n_0}\| \cdot \|x_{\varphi(m_1)}\| + \\ &\quad + \|A_{n_0}x_{\varphi(m_1)} - A_{n_0}x_{\varphi(m_2)}\| + \\ &\quad + \|A_{n_0} - A\| \cdot \|x_{\varphi(m_2)}\| \leq \\ &\leq 2c\|A_{n_0} - A\| + \|A_{n_0}x_{\varphi(m_1)} - A_{n_0}x_{\varphi(m_2)}\| < \\ &< 2c\frac{\varepsilon}{4c} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Resta verificato che $m \mapsto Ax_{\varphi(m)}$ è di Cauchy. \square

Definizione 3.1.9. Siano X, Y spazi vettoriali e $A: X \rightarrow Y$ lineare. Si dice che A ha *rango finito* se l'immagine di A ha dimensione finita.

Proposizione 3.1.10. *Siano X, Y normati e $A: X \rightarrow Y$ continua tale che l'immagine di A abbia dimensione finita. Allora A è compatto.*

Dimostrazione. Se l'immagine di A ha dimensione finita, allora è omeomorfo ad un \mathbb{R}^n o un \mathbb{C}^n (prop. 2.1.11 p. 76). Quindi A , essendo continua, manda un limitato di X in un limitato di uno spazio a dimensione finita, che è precompatto perché ha chiusura compatta. \square

Corollario 3.1.11. *Se $A_n: X \rightarrow Y$ è una successione di operatori lineari a rango finito fra due spazi normati, e $A_n \rightarrow A$ nella norma operatoriale, allora A è compatto.*

Proposizione 3.1.12. *Siano X, Y normati e $A: X \rightarrow Y$ compatto. Allora l'immagine di A è separabile, cioè ha un sottinsieme numerabile denso.*

Dimostrazione. Sia \mathcal{U} la palla unitaria chiusa di X . Poiché A è compatto, $A(\mathcal{U})$ è totalmente limitato. Preso $\varepsilon_n = 1/n$ esiste un sottinsieme finito $U_n \subset \mathcal{U}$ tale che $A(\mathcal{U}) \subset \bigcup_{x \in U_n} B(Ax, \varepsilon_n)$. Sia $U = \bigcup_{n \in \mathbb{N}} U_n$. È chiaro che U è (al più) numerabile e che $A(U)$ è denso in $A(\mathcal{U})$. Per omogeneità, $X = \bigcup_{n \in \mathbb{N}} n\mathcal{U}$, per cui $A(X) = \bigcup_{n \in \mathbb{N}} nA(\mathcal{U})$. Quindi l'insieme dei multipli interi dei vettori di $A(U)$ è (al più) numerabile e denso in $A(X)$. \square

Proposizione 3.1.13. *Sia X uno spazio normato, H uno spazio di Hilbert, $A: X \rightarrow H$ un operatore lineare compatto. Allora esiste una successione $A_n: X \rightarrow H$ di operatori lineari di rango finito tali che $A_n \rightarrow A$ rispetto alla norma operatoriale.*

Dimostrazione. Per escludere la situazione banale supponiamo che A non abbia già esso stesso rango finito. Essendo A compatto, l'immagine $A(X)$ è separabile, e la sua chiusura M in H è uno spazio di Hilbert con una base hilbertiana numerabile $\{u_n : n \in \mathbb{N}\}$. Sia P_n la proiezione ortogonale di H sul sottospazio di dimensione finita generato da $\{u_1, \dots, u_n\}$, e sia P_M la proiezione ortogonale di H su M . In formule, $P_n y = \sum_{k=1}^n \langle y, u_k \rangle u_k$, $P_M y = \sum_{n \in \mathbb{N}} \langle y, u_n \rangle u_n$. Sappiamo che $\|P_n\| = 1$ (formula (1.50) p. 35) e che per ogni $y \in M$ fissato $P_n y \rightarrow y$, cioè P_n tende a P_M puntualmente, sebbene in norma operatoriale $\|P_n - P_M\| = 1$ non tenda a zero. Poniamo $A_n = P_n A$. Chiaramente gli A_n hanno rango finito. Notare che su M la

proiezione P_M agisce come l'identità, per cui $P_M A = A$. Vogliamo dimostrare che $\|A_n - A\| \rightarrow 0$.

Fissiamo un $\varepsilon > 0$. Sia \mathcal{U} la palla unitaria chiusa di X . L'immagine $A(\mathcal{U})$ è un precompatto, e quindi esiste un insieme finito $U_\varepsilon \subset \mathcal{U}$ tale che $A(\mathcal{U}) \subseteq \bigcup_{x \in U_\varepsilon} B(Ax, \varepsilon/3)$. Poiché $A_n x = P_n A x \rightarrow P_M A x = Ax$ per $n \rightarrow +\infty$, per ogni $x \in X$ esiste un indice $N_{x,\varepsilon} \in \mathbb{N}$ tale che

$$\forall n \in \mathbb{N} \quad n > N_{x,\varepsilon} \quad \Rightarrow \quad \|P_n A x - Ax\| < \frac{\varepsilon}{3}.$$

Sia $N_\varepsilon = \max\{N_{x,\varepsilon} : x \in U_\varepsilon\} < +\infty$. Sia $x \in \mathcal{U}$. Esiste un $x_\varepsilon \in U_\varepsilon$ tale che $\|Ax - Ax_\varepsilon\| < \varepsilon/3$. Allora per ogni $n > N_\varepsilon \geq N_{x_\varepsilon,\varepsilon}$ possiamo scrivere

$$\begin{aligned} \|(A_n - A)x\| &= \|P_n A x - Ax\| \leq \\ &\leq \|P_n A x - P_n A x_\varepsilon\| + \|P_n A x_\varepsilon - Ax_\varepsilon\| + \|Ax_\varepsilon - Ax\| = \\ &= \|P_n(Ax - Ax_\varepsilon)\| + \|P_n A x_\varepsilon - Ax_\varepsilon\| + \|Ax_\varepsilon - Ax\| \leq \\ &\leq \|P_n\| \cdot \|Ax - Ax_\varepsilon\| + \|P_n A x_\varepsilon - Ax_\varepsilon\| + \|Ax_\varepsilon - Ax\| = \\ &= 2\|Ax - Ax_\varepsilon\| + \|P_n A x_\varepsilon - Ax_\varepsilon\| < \\ &< 2\frac{\varepsilon}{3} + \|P_n A x_\varepsilon - Ax_\varepsilon\| < \\ &< \frac{2}{3}\varepsilon + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Valendo questo per ogni $x \in \mathcal{U}$, abbiamo che

$$n > N_\varepsilon \quad \Rightarrow \quad \|A_n - A\| = \sup_{x \in \mathcal{U}} \|(A_n - A)x\| < \varepsilon,$$

cioè $A_n \rightarrow A$ in norma operatoriale, come desiderato. \square

Esempio 3.1.14 (operatore integrale). Sia $X = \mathcal{C}([0, 1])$ lo spazio delle funzioni continue $[0, 1] \rightarrow \mathbb{R}$ con la norma della convergenza uniforme. Per $f \in X$ poniamo $(Af)(t) := \int_0^t f(x) dx$. Chiaramente $A: X \rightarrow X$ ed è lineare. Dimostriamo che è compatto. Consideriamo la palla unitaria chiusa $\mathcal{U} \subset X$ e prendiamo $f \in \mathcal{U}$. Allora

$$\begin{aligned} \|Af\| &= \sup_{t \in [0,1]} |(Af)(t)| = \sup_{t \in [0,1]} \left| \int_0^t f(x) dx \right| \leq \\ &\leq \sup_{t \in [0,1]} \int_0^t |f(x)| dx \leq \sup_{t \in [0,1]} \int_0^1 1 dx = 1. \end{aligned}$$

Inoltre le funzioni di $A(\mathcal{U})$ sono equilipschitziane: se $0 \leq t_1 \leq t_2 \leq 1$

$$\begin{aligned} |(Af)(t_2) - (Af)(t_1)| &= \left| \int_0^{t_2} f(x) dx - \int_0^{t_1} f(x) dx \right| = \left| \int_{t_1}^{t_2} f(x) dx \right| \leq \\ &\leq \int_{t_1}^{t_2} |f(x)| dx \leq \int_{t_1}^{t_2} 1 dx \leq |t_2 - t_1|, \end{aligned}$$

e quindi equiuniformemente continue. Per il teorema di Ascoli-Arzelà, $A\mathcal{U}$ è precompatto.

Esempio 3.1.15. Nello spazio $\ell^p(\mathbb{N})$, con $1 < p < +\infty$, consideriamo l'applicazione A che manda il vettore $x = (x(1), x(2), x(3), \dots, x(n), \dots)$ in $Ax = (x(1), x(2)/2, x(3)/3, \dots, x(n)/n, \dots)$. Vista come matrice infinita, A sarebbe la matrice diagonale con $1, 1/2, 1/3, \dots$ sulla diagonale. L'operatore A è chiaramente lineare. Se x appartiene alla palla unitaria chiusa \mathcal{U} di $\ell^p(\mathbb{N})$, cioè se $\|x\|_{\ell^p} \leq 1$, allora ogni componente di Ax si può stimare con

$$|(Ax)(n)| = |x(n)/n| = \frac{1}{n} \left(|x(n)|^p \right)^{1/p} \leq \frac{1}{n} \left(\sum_{k \in \mathbb{N}} |x(k)|^p \right)^{1/p} \leq \frac{1}{n}.$$

Quindi Ax appartiene al cubo di Hilbert (la variante dell'esercizio 2.4.2 p. 84), che è compatto in $\ell^p(\mathbb{N})$. Ne deduciamo che $A: \ell^p(\mathbb{N}) \rightarrow \ell^p(\mathbb{N})$ è compatto.

3.2 Duale ed operatori aggiunti

Definizione 3.2.1. Dato uno spazio normato X , diremo *duale* di X l'insieme dei funzionali lineari continui da X a valori scalari. Il duale di X è indicato con X^* .

Per indicare elementi di X^* useremo simboli come x^* , che va inteso come un unico simbolo indivisibile, e non ha niente a che fare con l'eventuale vettore $x \in X$ che compaia nello stesso discorso. L'azione di x^* sul vettore x verrà indicata con $x^*(x)$ o con x^*x , e non con $\langle x^*, x \rangle$, che in questo corso è riservata al prodotto scalare (definizione 1.1.1 p. 5).

Il duale X^* è pure uno spazio vettoriale, ed è normato con la norma operatoriale di $\mathcal{L}(X, \mathbb{C})$, che qui talvolta indicheremo brevemente con $\|\cdot\|_*$:

$$\|x^*\|_* = \sup_{\|x\|_X=1} |x^*(x)|.$$

Poiché lo spazio di arrivo \mathbb{C} è completo, il duale è completo, anche se X stesso non lo è.

Dal teorema di Hahn-Banach segue che per ogni $x \in X$ esiste $x^* \in X^*$ tale che $\|x^*\|_* = 1$ e $x^*(x) = \|x\|_X$. Da qui traiamo l'ulteriore relazione che lega le norme dello spazio e del duale:

$$\|x\|_X = \max_{\|x^*\|_* = 1} |x^*(x)|. \quad (3.1)$$

Per chi fosse restio a invocare l'assioma di scelta, si può ricordare che la dimostrazione del teorema di Hahn-Banach lo richiede solo nel caso che X sia non separabile. Inoltre per gli spazi di Hilbert, anche non separabili, si può prendere come x^* semplicemente il funzionale $y \mapsto \langle y, x/\|x\| \rangle$.

Per il teorema di Hahn-Banach, dati dei vettori $x_1, x_2, \dots, x_n \in X$ linearmente indipendenti, esistono $x_1^*, x_2^*, \dots, x_n^* \in X^*$ tali che $x_j^*(x_k)$ vale 1 se $j = k$ e 0 altrimenti, cioè $x_j^*(x_k)$ è la delta di Kronecker $\delta_{j,k}$. Se la dimensione di X è infinita, pure il duale ha dimensione infinita.

Definizione 3.2.2. Se X ha dimensione finita n e x_1, x_2, \dots, x_n è una base, allora gli $x_1^*, x_2^*, \dots, x_n^* \in X^*$ tali che $x_j^*(x_k) = \delta_{j,k}$ sono unici e formano una base di X^* , chiamata *base duale*.

Se decomponiamo un $x \in X$ come combinazione lineare $x = \sum_{k=1}^n \alpha_k x_k$ troviamo che i coefficienti α_k sono forniti dalla base duale applicata ad x :

$$x_j^*(x) = x_j^*\left(\sum_{k=1}^n \alpha_k x_k\right) = \sum_{k=1}^n \alpha_k x_j^*(x_k) = \sum_{k=1}^n \alpha_k \delta_{j,k} = \alpha_j,$$

cioè

$$x = \sum_{k=1}^n x_k^*(x) x_k. \quad (3.2)$$

Il duale di X è completo, anche quando X non lo è.

Definizione 3.2.3. Dati due spazi normati X, Y e un operatore lineare $A: X \rightarrow Y$, chiameremo *operatore aggiunto* di A , indicato con A^* , l'operatore che associa ad ogni funzionale $y^* \in Y^*$ il funzionale composto $y^*A \in X^*$.

Nella notazione A^* , l'asterisco è un operatore che agisce su A . Invece quando scrivo x^* , questo è un simbolo indivisibile, indipendente da x .

La situazione è chiarita dai diagrammi seguenti:

$$\begin{array}{ccc}
 X \xrightarrow{A} Y & & X \xrightarrow{A} Y \\
 \searrow^{y^*} \swarrow_{A^*(y^*)} & \xrightarrow{A^*} & \downarrow y^* \\
 & & C
 \end{array}
 \qquad
 \begin{array}{ccc}
 X & \xrightarrow{A} & Y \\
 & & \\
 X^* & \xleftarrow{A^*} & Y^*
 \end{array}
 \tag{3.3}$$

A sinistra abbiamo un diagramma commutativo. Componendo A e y^* abbiamo $y^* \circ A$ che si può interpretare o come y^* che agisce su A , oppure A che agisce su y^* . L'aggiunto A^* esprime l'azione che A esercita su y^* , producendo un funzionale in X^* . Se applichiamo il tutto nel vettore x abbiamo le “trasposizioni” delle lettere A, y , o permutazioni della terna A, y, x :

$$y^* Ax = (A^* y^*)x.$$

Faremo altre trasposizioni più avanti. Nell'espressione $y^* Ax$ si possono sottintendere le parentesi, perché è una normale composizione associativa di funzioni: si parte da x , gli si applica A , al risultato si applica y^* . In $(A^* y^*)x$ non si può togliere la parentesi, perché per esempio non ha senso $y^*(x)$.

Esercizio 3.2.4. Se si usasse la notazione $\langle x^*, x \rangle$ per l'azione del funzionale x^* sul vettore x , come si scriverebbe la relazione di aggiunzione?

L'aggiunto A^* manda Y^* in X^* , e quindi bisogna fare attenzione alla direzione delle frecce nella parte destra dell'equazione (3.3). L'aggiunto è lineare: $(A^*(y_1^* + y_2^*))x = (y_1^* + y_2^*)Ax = y_1^*Ax + y_2^*Ax = (A^*y_1^*)x + (A^*y_2^*)x = (A^*y_1^* + A^*y_2^*)x$ e $(A^*(\alpha y^*))x = (\alpha y^*)Ax = \alpha y^*Ax = (\alpha(A^*y^*))x$ per ogni $x \in X$ e ogni $\alpha \in \mathbb{C}$, da cui $A^*(y_1^* + y_2^*) = A^*y_1^* + A^*y_2^*$ e $A^*(\alpha y^*) = \alpha A^*y^*$.

Esercizio 3.2.5. Qual è l'aggiunto dell'operatore identità $\text{id}_X: X \rightarrow X$?

Proposizione 3.2.6. *In dimensione finita, la matrice che rappresenta l'aggiunto A^* rispetto alle basi duali è la trasposta della matrice che rappresenta A .*

Dimostrazione. Sia X con base x_1, \dots, x_n , Y con base y_1, \dots, y_m , ed $A: X \rightarrow Y$ di matrice $a_{j,k}$, di modo che

$$Ax_k = \sum_j a_{j,k} y_j \quad \forall k = 1, \dots, n. \tag{3.4}$$

Rispetto alle basi duali l'aggiunto avrà matrice $a_{j,k}^*$:

$$A^*y_k^* = \sum_j a_{j,k}^* x_j^* \quad \forall k = 1, \dots, m. \quad (3.5)$$

Per trovare la relazione fra $a_{j,k}$ e $a_{j,k}^*$ applichiamo il primo membro della (3.5) a un vettore $x \in X$ e usiamo la formula (3.2):

$$\begin{aligned} (A^*y_k^*)(x) &= y_k^*Ax = y_k^*A \sum_j x_j^*(x)x_j = y_k^* \sum_j x_j^*(x)Ax_j = \\ &= y_k^* \sum_j x_j^*(x) \sum_\ell a_{\ell,j}y_\ell = \sum_j x_j^*(x) \sum_\ell a_{\ell,j}y_k^*y_\ell = \\ &= \sum_j x_j^*(x) \sum_\ell a_{\ell,j}\delta_{k,\ell} = \sum_j x_j^*(x)a_{k,j} = \\ &= \left(\sum_j a_{k,j}x_j^* \right)(x). \end{aligned}$$

Il primo e l'ultimo membro sono funzioni di x che coincidono sempre. Quindi

$$A^*y_k^* = \sum_j a_{k,j}x_j^* \quad \forall k. \quad (3.6)$$

Confrontando questa con la (3.5) otteniamo che $a_{j,k}^* = a_{k,j}$. \square

Proposizione 3.2.7. *Se $A: X \rightarrow Y$ è lineare continuo, anche l'aggiunto è continuo da Y^* in X^* . L'applicazione che manda A nell'aggiunto A^* è una isometria lineare da $\mathcal{L}(X, Y)$ a $\mathcal{L}(Y^*, X^*)$.*

Dimostrazione. Usiamo la caratterizzazione della norma operatoriale in termini della sfera unitaria:

$$\begin{aligned} \|A^*\|_{\mathcal{L}(Y^*, X^*)} &= \sup_{\|y^*\|_* = 1} \|A^*y^*\|_{X^*} = \\ &= \sup_{\|y^*\|_* = 1} \sup_{\|x\|_X = 1} |(A^*y^*)x| = \sup_{\|y^*\|_* = 1} \sup_{\|x\|_X = 1} |y^*Ax| \leq \\ &\leq \sup_{\|y^*\|_* = 1} \sup_{\|x\|_X = 1} \|y^*\|_* \cdot \|Ax\|_Y = \\ &= \sup_{\|x\|_X = 1} \|Ax\|_Y = \|A\|_{\mathcal{L}(X, Y)} < +\infty. \end{aligned}$$

Anzi, invertendo l'ordine dei due sup e usando la formula (3.1) si può rendere la catena di tutte uguglianze:

$$\begin{aligned}
 \|A^*\|_{\mathcal{L}(Y^*, X^*)} &= \sup_{\|y^*\|_* = 1} \|A^*y^*\|_* = \\
 &= \sup_{\|y^*\|_* = 1} \sup_{\|x\|_X = 1} |(A^*y^*)x| = \sup_{\|y^*\|_* = 1} \sup_{\|x\|_X = 1} |y^*Ax| = \\
 &= \sup_{\|x\|_X = 1} \sup_{\|y^*\|_* = 1} |y^*Ax| = \sup_{\|x\|_X = 1} \max_{\|y^*\|_* = 1} |y^*(Ax)| = \\
 &= \sup_{\|x\|_X = 1} \|Ax\|_Y = \|A\|_{\mathcal{L}(X, Y)}.
 \end{aligned}$$

L'applicazione $A \mapsto A^*$ è additiva perché per ogni $x \in X$, $y^* \in Y^*$

$$\begin{aligned}
 ((A + B)^*(y^*))(x) &= y^*((A + B)x) = y^*(Ax + Bx) = y^*Ax + y^*Bx = \\
 &= (A^*(y^*))(x) + (B^*(y^*))(x) = \\
 &= (A^*(y^*) + B^*(y^*))(x) = \\
 &= ((A^* + B^*)(y^*))(x),
 \end{aligned}$$

da cui $(A + B)^* = A^* + B^*$. Infine

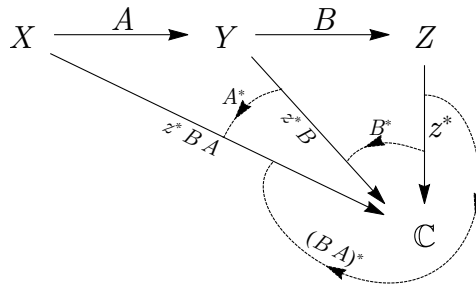
$$\begin{aligned}
 ((\alpha A)^*(y^*))(x) &= y^*((\alpha A)x) = y^*(\alpha Ax) = \alpha y^*(Ax) = \\
 &= \alpha (A^*(y^*))(x) = ((\alpha A^*)(y^*))(x)
 \end{aligned}$$

da cui

$$(\alpha A)^* = \alpha A^*. \quad (3.7)$$

□

Esercizio 3.2.8. Cosa ne è di $(BA)^*$, l'aggiunto di una composizione? Si può far riferimento al diagramma seguente:



Esercizio 3.2.9. Consideriamo l'operatore integrale $(Af)(x) := \int_0^x f(t) dt$ come nell'esempio 3.1.14 p. 89. Mostrare che A manda $L^p([0, 1])$ in sé per ogni $p \in [1, +\infty]$ e trovare una formula integrale per l'aggiunto A^* , usando i risultati noti sul duale di L^p .

3.3 Aggiunti e compattezza

Teorema 3.3.1 (l'aggiunto di un compatto è compatto). *Se $A: X \rightarrow Y$ è lineare e compatto, allora anche $A^*: Y^* \rightarrow X^*$ è compatto.*

Dimostrazione. Sia \mathcal{U} la palla unitaria chiusa di X . Poiché A è compatto, l'immagine $A(\mathcal{U})$ è totalmente limitata. Quindi esiste una successione $n \mapsto u_n \in \mathcal{U}$ tale che per ogni $\varepsilon > 0$ esiste un insieme finito $F_\varepsilon \subset \mathbb{N}$ di indici tali che $A(\mathcal{U}) \subseteq \bigcup_{n \in F_\varepsilon} B(Au_n, \varepsilon)$.

Sia $y_n^* \in Y^*$ limitata, con $\|y_n^*\|_{Y^*} < c$. Voglio dimostrare che esiste una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che la successione $n \mapsto A^*y_{\varphi(n)}^*$ sia di Cauchy in X^* . Consideriamo la matrice scalare a due indici

$$(A^*y_n^*)(u_m) = y_n^*(Au_m).$$

Poiché y_n^* , u_m e Au_m sono limitate nei rispettivi spazi, l'intera matrice è limitata. Pensando a n come indice di riga e m di colonna:

$$\begin{array}{cccccc} y_1^*(Au_1) & y_1^*(Au_2) & y_1^*(Au_3) & y_1^*(Au_4) & \dots & \\ y_2^*(Au_1) & y_2^*(Au_2) & y_2^*(Au_3) & y_2^*(Au_4) & \dots & \\ y_3^*(Au_1) & y_3^*(Au_2) & y_3^*(Au_3) & y_3^*(Au_4) & \dots & \\ y_4^*(Au_1) & y_4^*(Au_2) & y_4^*(Au_3) & y_4^*(Au_4) & \dots & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{array}$$

Col procedimento diagonale possiamo estrarre una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto (A^*y_{\varphi(n)}^*)(u_m) = y_{\varphi(n)}^*(Au_m)$ sia di Cauchy in \mathbb{C} per ogni m fissato (converge anche, ma non è importante). Ci proponiamo di dimostrare che questa sottosuccessione $n \mapsto A^*y_{\varphi(n)}^*$ è di Cauchy in X^* .

La distanza fra $A^*y_{\varphi(n_1)}^*$ e $A^*y_{\varphi(n_2)}^*$ si può scrivere così:

$$\begin{aligned} \|A^*y_{\varphi(n_1)}^* - A^*y_{\varphi(n_2)}^*\|_{X^*} &= \sup_{x \in \mathcal{U}} |(A^*y_{\varphi(n_1)}^*)x - (A^*y_{\varphi(n_2)}^*)x| = \\ &= \sup_{x \in \mathcal{U}} |y_{\varphi(n_1)}^*(Ax) - y_{\varphi(n_2)}^*(Ax)|. \end{aligned}$$

Fissiamo dunque $\varepsilon > 0$, a cui associamo l' $F_\varepsilon \subset \mathbb{N}$ descritto sopra. Poiché $n \mapsto (A^*y_{\varphi(n)}^*)(u_m) = y_{\varphi(n)}^*(Au_m)$ è di Cauchy, per ogni m esiste un indice $N_{\varepsilon,m} \in \mathbb{N}$ tale che per ogni n_1, n_2

$$n_1, n_2 > N_{\varepsilon,m} \quad \Rightarrow \quad |y_{\varphi(n_1)}^*(Au_m) - y_{\varphi(n_2)}^*(Au_m)| < \varepsilon.$$

Poniamo $N_\varepsilon = \max_{m \in F_\varepsilon} N_{\varepsilon,m} < +\infty$.

Siano $n_1, n_2 > N_\varepsilon$. Prendiamo $x \in \mathcal{U}$. Esiste un $n_{x,\varepsilon} \in F_\varepsilon$ tale che $\|Ax - Au_{n_{x,\varepsilon}}\| < \varepsilon$. Poiché $n_1, n_2 > N_\varepsilon \geq N_{\varepsilon, n_{x,\varepsilon}}$,

$$\begin{aligned} |y_{\varphi(n_1)}^*(Ax) - y_{\varphi(n_2)}^*(Ax)| &\leq |y_{\varphi(n_1)}^*(Ax) - y_{\varphi(n_1)}^*(Au_{n_{x,\varepsilon}})| + \\ &\quad + |y_{\varphi(n_1)}^*(Au_{n_{x,\varepsilon}}) - y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}})| + \\ &\quad + |y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}}) - y_{\varphi(n_2)}^*(Ax)| = \\ &= |y_{\varphi(n_1)}^*(Ax - Au_{n_{x,\varepsilon}})| + \\ &\quad + |y_{\varphi(n_1)}^*(Au_{n_{x,\varepsilon}}) - y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}})| + \\ &\quad + |y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}} - Ax)| \leq \\ &\leq \|y_{\varphi(n_1)}^*\|_{X^*} \cdot \|Ax - Au_{n_{x,\varepsilon}}\|_Y + \\ &\quad + |y_{\varphi(n_1)}^*(Au_{n_{x,\varepsilon}}) - y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}})| + \\ &\quad + \|y_{\varphi(n_2)}^*\|_{X^*} \cdot \|Au_{n_{x,\varepsilon}} - Ax\|_Y \leq \\ &\leq 2c\varepsilon + |y_{\varphi(n_1)}^*(Au_{n_{x,\varepsilon}}) - y_{\varphi(n_2)}^*(Au_{n_{x,\varepsilon}})| < \\ &< 2c\varepsilon + \varepsilon = (2c + 1)\varepsilon. \end{aligned}$$

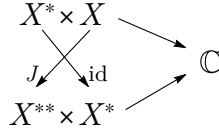
Valendo questo per ogni $x \in \mathcal{U}$ abbiamo che

$$n_1, n_2 > N_\varepsilon \quad \Rightarrow \quad \|A^*y_{\varphi(n_1)}^* - A^*y_{\varphi(n_2)}^*\|_{X^*} \leq (2c + 1)\varepsilon.$$

Questo mostra che $n \mapsto A^*y_{\varphi(n)}^*$ è di Cauchy in X^* . \square

Quando abbiamo $x \in X$ e $x^* \in X^*$, il calcolo di $x^*(x) \in \mathbb{C}$ può essere visto in due modi: tenendo fisso x^* e facendo variare x , oppure tenendo fisso x e facendo variare x^* . In questo secondo punto di vista, è x che agisce sul funzionale, e quindi è lui stesso un funzionale che manda X^* negli scalari. Più formalmente, dato $x \in X$ definiamo $J(x) \in X^{**}$ come $J(x)(x^*) := x^*(x)$. È un altro esempio di trasposizione fra i simboli x, x^* . L'applicazione J è l'immersione canonica di X nel bidual. Il diagramma seguente forse può

aiutare a capire la situazione:



L'applicazione J è lineare. Inoltre è continua: se $x \in X$

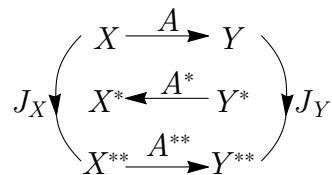
$$\begin{aligned}
 \|J(x)\|_{**} &= \sup_{\|x^*\|_*=1} |J(x)x^*| = \sup_{\|x^*\|_*=1} |x^*(x)| \leq \\
 &\leq \sup_{\|x^*\|_*=1} \|x^*\|_* \cdot \|x\|_X = \|x\|_X.
 \end{aligned}$$

Per il teorema di Hahn-Banach esiste un $h \in X^*$ con $\|h\|_* = 1$ e $h(x) = \|x\|_X$, per cui il sup è realizzato con $x^* = h$ e la disuguaglianza è un'uguaglianza. Insomma, $J: X \rightarrow X^{**}$ è una isometria. In particolare, J è iniettiva. Non sempre J è suriettiva: qualora lo sia, si dice che lo spazio X è *riflessivo*.

Esercizio 3.3.2. Sia X di dimensione finita con base x_1, \dots, x_n , e consideriamo la base duale x_1^*, \dots, x_n^* di X^* (Def. 3.2.2 p. 91) e la base biduale $x_1^{**}, \dots, x_n^{**}$ di X^{**} . Come si esprime l'immersione canonica $J: X \rightarrow X^{**}$ in termini di queste basi?

Teorema 3.3.3. *Siano X, Y spazi normati e $A: X \rightarrow Y$ lineare continua. Supponiamo che l'operatore aggiunto $A^*: Y^* \rightarrow X^*$ sia compatto. Allora anche A è compatto.*

Dimostrazione. Useremo le immersioni canoniche J_X, J_Y nel biduale di X e di Y , secondo lo schema seguente:



Il diagramma è commutativo. Infatti se $x \in X, y^* \in Y^*$ per definizione di aggiunto e di J possiamo fare un gioco delle tre carte con le lettere A, x, y :

$$\underbrace{(A^{**} J_X(x))}_{\in X^{**}} (y^*) = J_X(x) \underbrace{(A^* y^*)}_{\in X^*} = (A^* y^*)(x) = y^*(Ax) =$$

$$= \overbrace{(J_Y(Ax))}^{\in Y^{**}}(y^*) = \overbrace{(J_Y A(x))}^{\in Y^{**}}(y^*).$$

Valendo questo per ogni $y^* \in Y^*$ si ottiene che $A^{**}(J_X(x)) = J_Y(Ax)$, e valendo quest'ultimo per ogni $x \in X$ si ricava $A^{**}J_X = J_Y A$.

Supponiamo ora che A^* sia compatto. Sia x_n una successione limitata in X . Voglio dimostrare che $Ax_{\varphi(n)}$ è di Cauchy per una opportuna sottosuccessione.

Per il teorema precedente, il biaggiunto A^{**} è compatto da X^{**} in Y^{**} . Poiché J_X è una isometria, la successione $J_X(x_n)$ è pure limitata, e quindi esiste una sottosuccessione tale che $n \mapsto A^{**}J_X(x_{\varphi(n)})$ sia di Cauchy. Per la commutatività del diagramma, questa coincide con $n \mapsto J_Y(Ax_{\varphi(n)})$. Applicando l'inversa J_Y^{-1} , che è pure una isometria, abbiamo che anche $n \mapsto J_Y^{-1}J_Y(Ax_{\varphi(n)}) = Ax_{\varphi(n)}$ è di Cauchy, come volevasi dimostrare. \square

Chi volesse evitare di usare l'assioma di scelta riesce a salvare qualcosa del teorema precedente?

3.4 Operatori a rango finito

Lemma 3.4.1. *Siano X, Y spazi normati e $A: X \rightarrow Y$ un operatore lineare iniettivo. Allora si equivalgono le condizioni seguenti:*

- A^{-1} non è continuo,
- esiste una successione $n \mapsto x_n$ in X tale che $\|x_n\|_X \rightarrow +\infty$ mentre $\|Ax_n\|_Y \rightarrow 0$ per $n \rightarrow +\infty$.

Dimostrazione. Nel caso banale $X = \{0\}$ le due condizioni sono equivalenti in quanto entrambe false. Supponiamo quindi che X non sia ridotto al solo zero.

Supponiamo che l'operatore lineare A^{-1} , definito sull'immagine di A , non sia continuo. Questo significa che la sua norma operatoriale è infinita:

$$\|A^{-1}\| = \sup_{\substack{y \in A(X) \\ \|y\|_Y=1}} \|A^{-1}y\|_X = +\infty.$$

Allora esiste una successione $y_n \in A(X)$ tale che

$$\|y_n\|_Y \equiv 1 \quad \text{e} \quad \|A^{-1}y_n\|_X \rightarrow +\infty.$$

Poniamo $z_n = A^{-1}y_n$. Allora

$$\|Az_n\|_Y \equiv 1 \quad \text{e} \quad \|z_n\|_X \rightarrow +\infty.$$

Poniamo

$$\lambda_n := \|z_n\|_X \quad \text{e} \quad x_n := \frac{z_n}{\sqrt{\lambda_n}}.$$

Notare che $\lambda_n \rightarrow +\infty$ per $n \rightarrow +\infty$ e quindi

$$\begin{aligned} \|x_n\|_X &= \left\| \frac{z_n}{\sqrt{\lambda_n}} \right\| = \frac{\|z_n\|_X}{\sqrt{\lambda_n}} = \frac{\lambda_n}{\sqrt{\lambda_n}} = \sqrt{\lambda_n} \rightarrow +\infty, \\ \|Ax_n\|_Y &= \left\| \frac{Az_n}{\sqrt{\lambda_n}} \right\| = \frac{\|Az_n\|_Y}{\sqrt{\lambda_n}} = \frac{1}{\sqrt{\lambda_n}} \rightarrow 0. \end{aligned}$$

come desiderato.

Il viceversa è lasciato come esercizio. \square

Proposizione 3.4.2. *Siano X, Y spazi normati e $A: X \rightarrow Y$ un operatore lineare continuo. Allora si equivalgono:*

1. *A è compatto e l'immagine $A(X)$ è completa, come spazio metrico;*
2. *A ha rango finito.*

Dimostrazione. È noto che un operatore con rango finito è compatto e la sua immagine è completa in quanto equivalente a uno spazio \mathbb{C}^n o \mathbb{R}^n (Lemma 2.1.10 p. 75).

Supponiamo quindi che A sia compatto e che $A(X)$ sia completa. Posto $Z = A(X)$, possiamo considerare A come a valori in Z invece che in Y , e dimenticarci di Y . Per ipotesi Z è completo, cioè di Banach. Si tratta di dimostrare che Z ha dimensione finita.

Vediamo prima una dimostrazione più veloce aggiungendo l'ipotesi che anche X sia completo, cioè di Banach. L'operatore $A: X \rightarrow Z$ è lineare, continuo e suriettivo fra due spazi di Banach. Per il teorema della mappa aperta, A è anche aperto. Quindi l'immagine della palla unitaria aperta $B_X(0, 1)$ di X è un intorno aperto del vettore zero di Z . Se per assurdo Z avesse dimensione infinita, questo intorno aperto conterrebbe una successione z_n che non ammette sottosuccessioni di Cauchy (Corollario 2.1.8 p. 74). Per ogni n esiste $x_n \in B_X(0, 1)$ tale che $z_n = Ax_n$. Questa x_n è una successione limitata, la cui immagine tramite A non ha sottosuccessioni di Cauchy. Questo contraddice l'ipotesi che A sia compatto.

Il caso generale in cui X è soltanto normato richiede più lavoro. Lavoreremo sui duali e con l'operatore aggiunto di A , come nel diagramma

$$\begin{array}{ccc} X & \xrightarrow{A} & Z \subseteq Y \\ & & \\ X^* & \xleftarrow{A^*} & Z^* \end{array}$$

Affermo che $A^*: Z^* \rightarrow X^*$ è iniettivo. Preso infatti un $z^* \in Z^*$, tendo conto che A è suriettiva su Z , si ha

$$\begin{aligned} A^*z^* = 0 \in X^* & \iff \forall x \in X \quad (A^*z^*)(x) = 0 \in \mathbb{C} \\ & \iff \forall x \in X \quad z^*(Ax) = 0 \in \mathbb{C} \\ & \iff \forall z \in Z \quad z^*(z) = 0 \in \mathbb{C} \\ & \iff z^* = 0 \in Z^*, \end{aligned}$$

per cui il nucleo di A^* è costituito dal solo zero. (In generale, l'aggiunto di un suriettivo è iniettivo).

Affermo che l'inversa $(A^*)^{-1}$ è continua sull'immagine $A^*(Z^*)$. Se per assurdo fosse *discontinua*, per il lemma precedente esisterebbe una successione $z_n^* \in Z^*$ tale che per $n \rightarrow +\infty$

$$\|z_n^*\|_{Z^*} \rightarrow +\infty \quad \text{e} \quad \|A^*z_n^*\|_{X^*} \rightarrow 0.$$

Sia $z \in Z$. Esiste $x \in X$ tale che $z = Ax$. Quindi

$$|z_n^*(z)| = |z_n^*(Ax)| = |(A^*z_n^*)(x)| \leq \underbrace{\|A^*z_n^*\|_{X^*}}_{\rightarrow 0} \cdot \|x\|_X \rightarrow 0 \quad \text{per } n \rightarrow +\infty.$$

In particolare

$$\forall z \in Z \quad \sup_{n \in \mathbb{N}} |z_n^*(z)| < +\infty.$$

La successione z_n^* è dunque una famiglia di operatori lineari continui da Z in \mathbb{C} che è puntualmente limitata, cioè $\{z_n^*(z) : n \in \mathbb{N}\}$ è limitato in \mathbb{C} per ogni z fissato. Lo spazio di arrivo delle z_n^* è \mathbb{C} , che è completo. Possiamo applicare il teorema di Banach-Steinhaus, o dell'uniforme limitatezza, e dedurre che le norme operatoriali delle z_n^* sono limitate:

$$\sup_{n \in \mathbb{N}} \|z_n^*\| < +\infty.$$

Questo è assurdo perché abbiamo detto che $\|z_n^*\|$ tende all'infinito.

Riassumendo, $A^*: Z^* \rightarrow X^*$ è continua, iniettiva, e l'inversa è continua sull'immagine:

$$X^* \supseteq A^*(X^*) \begin{array}{c} \xleftarrow{A^*} \\ \xrightarrow{(A^*)^{-1}} \end{array} Z^* \begin{array}{c} \xrightarrow{\text{id}_{Z^*}} \\ \xleftarrow{\text{id}_{Z^*}} \end{array}$$

Quindi A^* è un omeomorfismo lineare di Z^* sull'immagine.

Usiamo finalmente l'ipotesi che A sia compatta, da cui segue che anche A^* è compatta. Ma allora l'operatore identità su Z^* è compatto, perché si può scrivere come $\text{id}_{Z^*} = (A^*)^{-1} \circ A^*$, composizione di un operatore compatto con uno continuo (Prop. 3.1.7 p. 86). Quindi da ogni successione limitata in Z^* si può estrarre una sottosuccessione di Cauchy. Concludiamo che Z^* , e quindi anche Z , hanno dimensione finita. \square

Esercizio 3.4.3. Se $A: X \rightarrow Y$ è suriettiva, allora A^* è iniettiva.

Capitolo 4

Compatti e aggiunti hilbertiani

4.1 Duale e bduale di uno spazio di Hilbert

Una conseguenza del teorema della proiezione ortogonale è il fatto che ogni funzionale lineare continuo su uno spazio di Hilbert a valori scalari è sempre rappresentabile come il prodotto scalare con un vettore fisso:

Teorema 4.1.1 (di rappresentazione). *Sia H uno spazio di Hilbert e sia $x^*: H \rightarrow \mathbb{C}$ un funzionale lineare continuo. Allora esiste un unico vettore $K(x^*) \in H$ tale che*

$$x^*(x) = \langle x, K(x^*) \rangle \quad \forall x \in H.$$

Dimostrazione. L'unicità è facile: se entrambi y_1, y_2 hanno la proprietà che $x^*(x) = \langle x, y_1 \rangle = \langle x, y_2 \rangle$, allora sottraendo membro a membro otteniamo

$$0 = x^*(x) - x^*(x) = \langle x, y_1 \rangle - \langle x, y_2 \rangle = \langle x, y_1 - y_2 \rangle \quad \forall x \in H.$$

In particolare, ponendo $x = y_1 - y_2$, otteniamo

$$0 = \langle y_1 - y_2, y_1 - y_2 \rangle = \|y_1 - y_2\|^2,$$

da cui $y_1 = y_2$. Per dimostrare l'esistenza, consideriamo $M = \ker x^*$. Poiché x^* è lineare e continuo, M è un sottospazio vettoriale chiuso di H , e pertanto vale la decomposizione ortogonale

$$H = M \oplus M^\perp = (\ker x^*) \oplus (\ker x^*)^\perp.$$

Nel caso banale in cui $x^* \equiv 0$ possiamo prendere direttamente $K(x^*) = 0$. Se invece x^* non è identicamente nullo, $M = \ker x^* \subsetneq H$ e quindi $M^\perp \supsetneq \{0\}$. Prendiamo un vettore $z_0 \in M^\perp \setminus \{0\}$. Normalizzandolo se necessario, possiamo supporre che $\|z_0\| = 1$. Fissato $x \in H$ calcoliamo come segue:

$$\begin{aligned} x^*(x) &= x^*(x)\|z_0\|^2 = x^*(x)\langle z_0, z_0 \rangle = \langle x^*(x)z_0, z_0 \rangle = \\ &= \langle \underbrace{x^*(x)z_0 - x^*(z_0)x}_{\in M} + x^*(z_0)x, z_0 \rangle. \end{aligned} \quad (4.1)$$

Osserviamo che il termine $x^*(x)z_0 - x^*(z_0)x$ appartiene al $\ker x^*$. Infatti, poiché x^* è lineare e $x^*(x), x^*(z_0)$ sono coefficienti scalari,

$$x^*(x^*(x)z_0 - x^*(z_0)x) = x^*(x)x^*(z_0) - x^*(z_0)x^*(x) = 0.$$

Tornando alla (4.1),

$$\begin{aligned} x^*(x) &= \langle x^*(x)z_0 - x^*(z_0)x + x^*(z_0)x, z_0 \rangle = \\ &= \langle \underbrace{x^*(x)z_0 - x^*(z_0)x}_{\in M}, \underbrace{z_0}_{\in M^\perp} \rangle + \langle x^*(z_0)x, z_0 \rangle = \\ &= 0 + \langle x^*(z_0)x, z_0 \rangle = x^*(z_0)\langle x, z_0 \rangle = \\ &= \langle x, \overline{x^*(z_0)z_0} \rangle. \end{aligned}$$

Essendo questo valido per ogni $x \in H$, possiamo porre $K(x^*) = \overline{x^*(z_0)z_0}$. \square

L'applicazione K del teorema di rappresentazione va da H^* in H e si inserisce nel diagramma seguente:

$$\begin{array}{ccc} H^* \times H & & \\ \text{id} \searrow & \searrow K & \\ H \times H & \xrightarrow{\langle \cdot, \cdot \rangle} & \mathbb{C} \end{array}$$

La K è suriettiva, perché preso un qualsiasi $x \in H$, se poniamo $f(y) := \langle y, x \rangle$ si ha che $f \in H^*$ e $K(f) = x$. Chiaramente K è additiva, dato che

$$\begin{aligned} \langle x, K(x_1^* + x_2^*) \rangle &= (x_1^* + x_2^*)(x) = x_1^*(x) + x_2^*(x) = \\ &= \langle x, K(x_1^*) \rangle + \langle x, K(x_2^*) \rangle = \\ &= \langle x, K(x_1^*) + K(x_2^*) \rangle \end{aligned}$$

per ogni $x \in H$, da cui $K(x_1^* + x_2^*) = K(x_1^*) + K(x_2^*)$. Purtroppo K non è omogenea sui complessi, perché se $x \in H$ e $\alpha \in \mathbb{C}$

$$\langle x, K(\alpha x^*) \rangle = (\alpha x^*)(x) = \alpha x^*(x) = \alpha \langle x, K(x^*) \rangle = \langle x, \bar{\alpha} K(x^*) \rangle,$$

da cui $K(\alpha x^*) = \bar{\alpha} K(x^*)$. Però K è lineare sui reali.

In ogni caso K è una isometria. Per vederlo, osserviamo innanzi tutto che grazie alla disuguaglianza di Schwartz

$$\|y\| = \max_{\|x\| \leq 1} |\langle y, x \rangle| = \max_{\|x\|=1} |\langle y, x \rangle|,$$

il massimo essendo raggiunto quando x è il versore di y (se questo è non nullo). Si tratta della versione hilbertiana della formula (3.1) di p. 91, che qui fa a meno del teorema di Hahn-Banach e dell'assioma di scelta. Quindi

$$\|K(x^*)\|_H = \sup_{\|x\| \leq 1} |\langle x, K(x^*) \rangle| = \sup_{\|x\| \leq 1} |x^*(x)| = \|x^*\|_{H^*}.$$

Grazie a K possiamo dotare il duale H^* di un prodotto scalare che genera la norma operatoriale:

$$\langle x_1^*, x_2^* \rangle_{H^*} := \langle K(x_2^*), K(x_1^*) \rangle_H$$

(notare l'inversione dell'ordine). Quindi il duale di uno spazio di Hilbert è ancora uno spazio di Hilbert, e fra di loro c'è una isometria (non lineare sui complessi).

C'è una relazione stretta fra la K e l'immersione canonica J nel biduali che abbiamo visto per spazi normati a p. 97. Siano infatti $K_*: H^* \rightarrow H$ l'immersione del duale H^* in H , $K_{**}: H^{**} \rightarrow H^*$ l'immersione del biduali H^{**} nel duale H^* (attenzione, K_* , K_{**} non c'entrano cogli aggiunti), e $J: H \rightarrow H^{**}$ l'immersione canonica nel biduali, secondo lo schema:

$$\begin{array}{ccccc} H & \xleftarrow{K_*} & H^* & \xleftarrow{K_{**}} & H^{**} \\ & \searrow & \xrightarrow{J} & \swarrow & \\ & & & & \end{array}$$

Ebbene, il diagramma commuta. Prendiamo infatti $z^{**} \in H^{**}$, $y^* \in H^*$:

$$z^{**}(y^*) = \langle y^*, \overbrace{K_{**}(z^{**})}^{\in H^*} \rangle_{H^*} = \langle \overbrace{K_*(K_{**}(z^{**}))}^{\in H}, K_*(y^*) \rangle_H =$$

$$= y^* \underbrace{(K_* K_{**}(z^{**}))}_{\in H} = \underbrace{J(K_* K_{**}(z^{**}))}_{\in H^{**}}(y^*).$$

Valendo questo per ogni $y^* \in H^*$ si ricava che $z^{**} = J(K_* K_{**}(z^{**}))$ per ogni $z^{**} \in H^{**}$, da cui infine $JK_* K_{**} = \text{id}_{H^{**}}$. Una conseguenza è che l'immersione canonica nel biduale è suriettiva nel caso di spazi di Hilbert, cioè gli spazi di Hilbert sono riflessivi.

Se H ha dimensione finita n e gli u_1, \dots, u_n sono una base *ortonormale* di H , possiamo scrivere, dati $x \in H$ e $x^* \in H^*$,

$$\begin{aligned} x^*(x) &= x^* \left(\sum_k \langle x, u_k \rangle u_k \right) = \sum_k \langle x, u_k \rangle x^*(u_k) = \sum_k \langle x, \overline{x^*(u_k)} u_k \rangle = \\ &= \left\langle x, \sum_k \overline{x^*(u_k)} u_k \right\rangle, \end{aligned}$$

di modo che vale la formula esplicita

$$K(x^*) = \sum_{k=1}^n \overline{x^*(u_k)} u_k. \quad (4.2)$$

Se u_1^*, \dots, u_n^* è la base duale di u_1, \dots, u_n (Def. 3.2.2 p. 91), abbiamo che $K(u_k^*) = u_k$ per ogni k . Qualora H sia uno spazio vettoriale su \mathbb{R} , questo significa che la matrice che rappresenta K rispetto a queste due basi è la matrice identità. Se invece lo spazio è complesso, K è non lineare, e quindi non è rappresentabile con le convenzioni delle matrici, però possiamo scrivere

$$K \left(\sum_k \alpha_k u_k^* \right) = \sum_k \overline{\alpha_k} K(u_k^*) = \sum_k \overline{\alpha_k} u_k. \quad (4.3)$$

Espresso nella base duale, il prodotto scalare su H^* prende una forma familiare (esempio 1.1.7 p. 13):

$$\begin{aligned} \left\langle \sum_k \alpha_k u_k^*, \sum_k \beta_k u_k^* \right\rangle_{H^*} &= \left\langle K \left(\sum_k \beta_k u_k^* \right), K \left(\sum_k \alpha_k u_k^* \right) \right\rangle_H = \\ &= \left\langle \sum_k \overline{\beta_k} u_k, \sum_k \overline{\alpha_k} u_k \right\rangle_H = \\ &= \sum_k \alpha_k \overline{\beta_k}. \end{aligned}$$

4.2 Aggiunto hilbertiano

Definizione 4.2.1. Siano H, G due spazi di Hilbert, $A: H \rightarrow G$ un operatore lineare continuo. L'aggiunto hilbertiano $A^*: G \rightarrow H$ è l'operatore che associa ad ogni $y \in G$ il vettore $A^*y \in H$ tale che

$$\forall x \in H \quad \langle Ax, y \rangle_G = \langle x, A^*y \rangle_H.$$

L'aggiunto hilbertiano è imparentato ma distinto dall'aggiunto che abbiamo introdotto per gli spazi normati nella sezione 3.2, p. 90. Notare che per l'aggiunto hilbertiano usiamo la stella a 5 punte \star invece dell'asterisco a 6 punte $*$. La cosa non è troppo importante, perché non capiterà spesso che i due oggetti compaiano insieme in un discorso.

La definizione è giustificata perché la funzione $f(x) := \langle Ax, y \rangle_G$ è lineare e continua $H \rightarrow \mathbb{C}$, e quindi, per il teorema di rappresentazione, esiste unico un vettore $z = K_H f \in H$ tale che $f(x) = \langle Ax, y \rangle_G = \langle x, z \rangle_H$ per ogni $x \in H$. Osserviamo anche che

$$\forall x \in H \quad \forall y \in G \quad \langle y, Ax \rangle_G = \overline{\langle Ax, y \rangle_G} = \overline{\langle x, A^*y \rangle_H} = \langle A^*y, x \rangle_H,$$

cioè si può “portare A dall'altro lato del prodotto scalare” trasformandolo in aggiunto sia che si trovi nel lato sinistro che nel lato destro. In particolare, il biaggiunto hilbertiano di A è A medesimo:

$$(A^*)^* = A.$$

L'aggiunto hilbertiano $A^*: G \rightarrow H$ è *additivo*:

$$\begin{aligned} \langle x, A^*(y_1 + y_2) \rangle_H &= \langle Ax, y_1 + y_2 \rangle_G = \langle Ax, y_1 \rangle_G + \langle Ax, y_2 \rangle_G = \\ &= \langle x, A^*y_1 \rangle_G + \langle x, A^*y_2 \rangle_G = \\ &= \langle x, A^*y_1 + A^*y_2 \rangle_G \end{aligned}$$

da cui $A^*(y_1 + y_2) = A^*y_1 + A^*y_2$; è anche *omogeneo*:

$$\langle x, A^*(\alpha y) \rangle_H = \langle Ax, \alpha y \rangle_G = \bar{\alpha} \langle Ax, y \rangle_G = \bar{\alpha} \langle x, A^*y \rangle_G = \langle x, \alpha A^*y \rangle_G,$$

da cui

$$A^*(\alpha y) = \alpha A^*y. \tag{4.4}$$

Quindi l'aggiunto hilbertiano A^* è *lineare* $G \rightarrow H$.

Esercizio 4.2.2. Qual è l'aggiunto hilbertiano dell'operatore identità $\text{id}_H: H \rightarrow H$?

La seguente proposizione è parallela alla proposizione 3.2.6, p. 92, che riguarda l'operatore aggiunto.

Proposizione 4.2.3. *In dimensione finita, la matrice che rappresenta l'aggiunto hilbertiano A^* rispetto a una coppia di basi ortonormali è la trasposta coniugata della matrice che rappresenta A .*

Dimostrazione. Si può fare un conto analogo a quello della proposizione 3.2.6 di p. 92. Sia u_1, \dots, u_n base ortonormale di H , v_1, \dots, v_m base ortonormale di G , ed $A: H \rightarrow G$ sia rappresentata dalla matrice $a_{j,k}$, di modo che

$$Au_k = \sum_j a_{j,k} v_j \quad \forall k = 1, \dots, n. \quad (4.5)$$

Rispetto alle stesse basi, l'aggiunto avrà matrice $a_{j,k}^*$:

$$A^* v_k = \sum_j a_{j,k}^* u_j \quad \forall k = 1, \dots, m. \quad (4.6)$$

Per trovare la relazione fra $a_{j,k}$ e $a_{j,k}^*$ prendiamo $x \in H$ e calcoliamo:

$$\begin{aligned} \langle A^* v_k, x \rangle_H &= \langle v_k, Ax \rangle_G = \left\langle v_k, A \sum_j \langle x, u_j \rangle_H u_j \right\rangle_G = \\ &= \left\langle v_k, \sum_j \langle x, u_j \rangle_H Au_j \right\rangle_G = \left\langle v_k, \sum_j \langle x, u_j \rangle_H \sum_\ell a_{\ell,j} v_\ell \right\rangle_G = \\ &= \sum_j \overline{\langle x, u_j \rangle_H} \sum_\ell \overline{a_{\ell,j}} \langle v_k, v_\ell \rangle_G = \sum_j \langle u_j, x \rangle_H \sum_\ell \overline{a_{\ell,j}} \delta_{k,\ell} = \\ &= \sum_j \langle u_j, x \rangle_H \overline{a_{k,j}} = \left\langle \sum_j \overline{a_{k,j}} u_j, x \right\rangle_H. \end{aligned}$$

Valendo questo per ogni $x \in H$ e ogni k traiamo che

$$A^* v_k = \sum_j \overline{a_{k,j}} u_j \quad \forall k = 1, \dots, m.$$

Confrontando con la (4.6), ricaviamo infine

$$a_{j,k}^* = \overline{a_{k,j}} \quad \forall k, j. \quad (4.7)$$

□

Proposizione 4.2.4. *Se $A: H \rightarrow G$ è lineare continuo, allora l'aggiunto hilbertiano A^* è continuo da G in H . L'applicazione che manda A nell'aggiunto hilbertiano A^* è una isometria additiva da $\mathcal{L}(H, G)$ su tutto $\mathcal{L}(G, H)$, però $(\alpha A)^* = \bar{\alpha}A^*$.*

Dimostrazione. A^* è continuo. Infatti

$$\begin{aligned} \|A^*y\|_H &= \sup_{\|x\|_H=1} |\langle A^*y, x \rangle_H| = \sup_{\|x\|_H=1} |\langle y, Ax \rangle_G| \leq \\ &\leq \sup_{\|x\|_H=1} \|y\|_G \cdot \|Ax\|_G \leq \sup_{\|x\|_H=1} \|y\|_G \cdot \|A\|_{\mathcal{L}(H,G)} \cdot \|x\|_H = \\ &= \|y\|_G \cdot \|A\|_{\mathcal{L}(H,G)} < +\infty, \end{aligned}$$

da cui anche $\|A^*\|_{\mathcal{L}(G,H)} \leq \|A\|_{\mathcal{L}(H,G)}$. Usando il fatto che $(A^*)^* = A$ si ricava che l'aggiunto hilbertiano ha la stessa norma operatoriale dell'operatore di partenza:

$$\|A^*\|_{\mathcal{L}(G,H)} = \|A\|_{\mathcal{L}(H,G)} = \sup_{\substack{\|x\|_H \leq 1 \\ \|y\|_G \leq 1}} |\langle Ax, y \rangle_G| = \sup_{\substack{\|x\|_H \leq 1 \\ \|y\|_G \leq 1}} |\langle x, A^*y \rangle_H|.$$

L'operatore di agguinzione $A \mapsto A^*$ da $\mathcal{L}(H, G)$ in $\mathcal{L}(G, H)$ è quindi una isometria, ed è additivo, cioè $(A + B)^* = A^* + B^*$, ma non è omogeneo sui complessi, in quanto $\langle (\alpha A)x, y \rangle = \alpha \langle Ax, y \rangle = \alpha \langle x, A^*y \rangle = \langle x, \bar{\alpha}A^*y \rangle$, per cui

$$(\alpha A)^* = \bar{\alpha}A^* \quad (4.8)$$

(non confondere questo con l'omogeneità di A^* su H , cioè $A^*(\alpha x) = \alpha A^*x$, equazione (4.4)). Notare che per l'aggiunto non hilbertiano si ha invece $(\alpha A)^* = \alpha A^*$ (formula (3.7) p. 94). \square

La notazione con stella a 5 punte che usiamo qui per l'aggiunto hilbertiano si confonde facilmente con quella con asterisco a 6 punte per l'aggiunto per spazi normati, ma la cosa non è grave, perché per gli spazi di Hilbert i due concetti sono identificabili. Infatti il diagramma seguente è commutativo:

$$\begin{array}{ccc} H & \xrightarrow{A} & G \\ & & \uparrow K_G \\ H & \xleftarrow{A^*} & G \\ \uparrow K_H & & \uparrow K_G \\ H^* & \xleftarrow{A^*} & G^* \end{array}$$

Per vederlo, presi $y^* \in G^*$ e $x \in H$ si può scrivere

$$\begin{aligned} \langle x, \underbrace{K_H A^* y^*}_{\in H} \rangle_H &= (A^* y^*)(x) = y^*(Ax) = \langle Ax, K_G y^* \rangle_G = \\ &= \langle x, \underbrace{A^* K_G y^*}_{\in H} \rangle_H. \end{aligned}$$

Valendo questo per ogni $x \in H$ ricaviamo che $K_H A^* y^* = A^* K_G(y^*)$, e a sua volta valendo questo per ogni $y \in H^*$ deduciamo che $K_H A^* = A^* K_G$.

Poiché K_H e K_G sono isometrie, si vede che l'aggiunto hilbertiano è compatto se e solo se l'aggiunto ordinario è compatto. Inoltre le proprietà di compattezza dell'aggiunto ordinario si trasferiscono all'aggiunto hilbertiano: *l'aggiunto hilbertiano A^* è compatto se e solo se A è compatto.*

Esercizio 4.2.5. Trovare l'aggiunto hilbertiano dell'operatore integrale $(Af)(t) := \int_0^t f(x)dx$ dell'esempio 3.1.14, p. 89, dopo aver mostrato che manda $L^2([0, 1])$ in sé.

4.3 Convergenza debole e compattezza

Definizione 4.3.1. Dato uno spazio normato X , diremo che la successione $n \mapsto x_n \in X$ converge *debolmente* a $\bar{x} \in X$, e scriveremo $x_n \rightharpoonup \bar{x}$, se per ogni funzionale lineare continuo $x^* \in X^*$ si ha che $x^*(x_n) \rightarrow x^*(\bar{x})$ per $n \rightarrow +\infty$.

Il limite debole, se esiste, è unico, e si mantiene passando a sottosuccessioni. Se $x_n \rightharpoonup \bar{x}$ e $y_n \rightharpoonup \bar{y}$ e $\alpha \in \mathbb{C}$ allora $x_n + y_n \rightharpoonup \bar{x} + \bar{y}$ e $\alpha x_n \rightharpoonup \alpha \bar{x}$. Per il teorema di Banach-Steinhaus, una successione che converge debolmente è anche limitata. Se x_n converge in norma ("fortemente") a \bar{x} , allora converge anche debolmente alla stessa cosa, cioè

$$x_n \rightarrow \bar{x} \quad \Rightarrow \quad x_n \rightharpoonup \bar{x}.$$

Se H è uno spazio di Hilbert, grazie al teorema di rappresentazione possiamo dire che $x_n \rightharpoonup \bar{x}$ se e solo se $\langle x_n, y \rangle \rightarrow \langle \bar{x}, y \rangle$ per ogni $y \in H$.

Esempio 4.3.2. Sia $\{u_n : n \in \mathbb{N}\}$ un sistema ortonormale numerabile in uno spazio di Hilbert. Allora u_n converge debolmente verso il vettore zero. Infatti sia M la chiusura del sottospazio generato dal sistema ortonormale, e sia $y \in H$. Si ha

$$+\infty > \|P_M y\|^2 = \left\| \sum_{n \in \mathbb{N}} \langle y, u_n \rangle u_n \right\|^2 = \sum_{n \in \mathbb{N}} |\langle y, u_n \rangle|^2.$$

Dalla convergenza della serie segue in particolare che $\langle u_n, y \rangle \rightarrow 0$ per $n \rightarrow +\infty$.

Teorema 4.3.3. *In uno spazio di Hilbert, da ogni successione limitata in norma si può estrarre una sottosuccessione che converge debolmente.*

Dimostrazione. Sia H uno spazio di Hilbert e x_n una successione limitata in H , con $\|x_n\| \leq c < +\infty$. Vogliamo dimostrare che esiste una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente e un $\bar{x} \in H$ tali che $x_{\varphi(n)} \rightharpoonup \bar{x}$.

Sia M la chiusura del sottospazio generato dagli x_n . Questo M è uno spazio di Hilbert separabile, cioè ammette un sottinsieme numerabile denso $\{y_m : m \in \mathbb{N}\}$. Sia P_M la proiezione ortogonale di H su M . Consideriamo la matrice di scalari

$$\begin{array}{cccccc} \langle x_1, y_1 \rangle & \langle x_1, y_2 \rangle & \langle x_1, y_3 \rangle & \langle x_1, y_4 \rangle & \dots & \\ \langle x_2, y_1 \rangle & \langle x_2, y_2 \rangle & \langle x_2, y_3 \rangle & \langle x_2, y_4 \rangle & \dots & \\ \langle x_3, y_1 \rangle & \langle x_3, y_2 \rangle & \langle x_3, y_3 \rangle & \langle x_3, y_4 \rangle & \dots & \\ \langle x_4, y_1 \rangle & \langle x_4, y_2 \rangle & \langle x_4, y_3 \rangle & \langle x_4, y_4 \rangle & \dots & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{array}$$

Ogni colonna della matrice è limitata: $|\langle x_n, y_m \rangle| \leq \|x_n\| \cdot \|y_m\| \leq c \cdot \|y_m\|$ per ogni n . Col ragionamento diagonale otteniamo una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $n \mapsto \langle x_{\varphi(n)}, y_m \rangle$ converge.

Dico che la successione $n \mapsto \langle x_{\varphi(n)}, y \rangle$ converge per ogni $y \in H$. Sia infatti $\varepsilon > 0$. Per la densità degli y_m in M esiste un indice $m_\varepsilon \in \mathbb{N}$ tale che $\|y_{m_\varepsilon} - P_M y\| < \varepsilon/(4c)$. Inoltre per ogni $m \in \mathbb{N}$ esiste $M_\varepsilon \in \mathbb{N}$ tale che

$$n_1, n_2 > M_\varepsilon \quad \Rightarrow \quad |\langle x_{\varphi(n_1)}, y_{m_\varepsilon} \rangle - \langle x_{\varphi(n_2)}, y_{m_\varepsilon} \rangle| < \frac{\varepsilon}{2}.$$

Allora, se $n_1, n_2 > M_\varepsilon$ possiamo scrivere

$$\begin{aligned} & \left| \langle x_{\varphi(n_1)}, y \rangle - \langle x_{\varphi(n_2)}, y \rangle \right| = \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, y \rangle \right| = \\ & = \left| \overbrace{\langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, y - P_M y \rangle}^{\in M} + \overbrace{\langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, P_M y \rangle}^{\in M^\perp} \right| = \\ & = \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, P_M y \rangle \right| \leq \\ & \leq \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, P_M y - y_{m_\varepsilon} \rangle \right| + \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, y_{m_\varepsilon} \rangle \right| \leq \\ & \leq \|x_{\varphi(n_1)} - x_{\varphi(n_2)}\| \cdot \|P_M y - y_{m_\varepsilon}\| + \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, y_{m_\varepsilon} \rangle \right| \leq \\ & \leq 2c \cdot \frac{\varepsilon}{4c} + \left| \langle x_{\varphi(n_1)} - x_{\varphi(n_2)}, y_{m_\varepsilon} \rangle \right| = \end{aligned}$$

$$= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Resta quindi dimostrato che $n \mapsto \langle x_{\varphi(n)}, y \rangle$ converge in \mathbb{C} per ogni $y \in H$. Poniamo

$$f(y) := \lim_{n \rightarrow +\infty} \langle y, x_{\varphi(n)} \rangle.$$

Ovviamente $f: H \rightarrow \mathbb{C}$ è lineare, ed è anche continua perché $|f(y)| \leq c\|y\|_H$. Per il teorema di rappresentazione, esiste un $\bar{x} = K(f) \in H$ tale che $f(y) = \langle y, \bar{x} \rangle$ per ogni $y \in H$. Ma allora

$$\langle x_{\varphi(n)}, y \rangle \longrightarrow \overline{f(y)} = \langle \bar{x}, y \rangle$$

per $n \rightarrow +\infty$ per ogni $y \in H$, ossia $x_{\varphi(n)} \rightharpoonup \bar{x}$. □

Notare che \bar{x} non vuol dire il coniugato di x , che avrebbe poco senso.

Proposizione 4.3.4. *Siano H, G due spazi di Hilbert e $A: H \rightarrow G$ un operatore lineare continuo. Allora si equivalgono le condizioni seguenti:*

1. A è compatto;
2. per ogni successione $x_n \in H$ che converge debolmente a un $\bar{x} \in H$, l'immagine Ax_n converge in norma a $A\bar{x}$:

$$x_n \rightharpoonup \bar{x} \quad \Rightarrow \quad Ax_n \rightarrow A\bar{x}.$$

Dimostrazione. Supponiamo che A sia compatto e sia $x_n \in H$ tale che $x_n \rightharpoonup \bar{x}$. Allora Ax_n tende debolmente a $A\bar{x}$ (anche senza compattezza). Infatti, preso un $y \in G$, abbiamo che

$$\langle Ax_n - A\bar{x}, y \rangle_G = \langle A(x_n - \bar{x}), y \rangle_G = \langle x_n - \bar{x}, A^*y \rangle_H \longrightarrow 0.$$

Supponiamo per assurdo che Ax_n non converga in norma ad $A\bar{x}$. Allora c'è una sottosuccessione che si tiene lontana da $A\bar{x}$, cioè esiste $\varepsilon > 0$ e una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che

$$\|Ax_{\varphi(n)} - A\bar{x}\| \geq \varepsilon \quad \forall n \in \mathbb{N}.$$

Poiché x_n è limitata, per la compattezza esiste una $\psi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $Ax_{\varphi \circ \psi(n)}$ è di Cauchy, e quindi converge in norma (e in

particolare anche debolmente) a un qualche $\bar{y} \in G$. Passando al limite per $n \rightarrow +\infty$ nella disuguaglianza

$$\|Ax_{\varphi \circ \psi(n)} - A\bar{x}\| \geq \varepsilon$$

otteniamo che $\|\bar{y} - A\bar{x}\| \geq \varepsilon > 0$. In particolare, $\bar{y} \neq A\bar{x}$. La sottosuccessione $Ax_{\varphi \circ \psi(n)}$ risulterebbe convergere debolmente a due limiti diversi, il che è assurdo perché il limite debole è unico.

Viceversa, supponiamo che A mandi successioni che convergono debolmente in successioni che convergono fortemente, e dimostriamo che A è compatto. Sia $x_n \in H$ una successione limitata. Sappiamo che esiste una sottosuccessione $x_{\varphi(n)}$ che converge debolmente. Per ipotesi l'immagine $Ax_{\varphi(n)}$ converge fortemente in G , come desiderato. \square

Capitolo 5

Teoria spettrale

5.1 Autovalori, autovettori e autospazi

Definizione 5.1.1. Se X è uno spazio vettoriale su \mathbb{C} (non ridotto al solo zero) e $A: X \rightarrow X$ è un'applicazione lineare, chiameremo *autovalori* e relativi *autovettori* di A rispettivamente degli scalari $\lambda \in \mathbb{C}$ e vettori non nulli $\bar{x} \in X \setminus \{0\}$ per i quali $A\bar{x} = \lambda\bar{x}$. L'insieme degli $x \in X$ tali che $Ax = \lambda x$, cioè $\ker(\lambda \text{id}_X - A)$, si dice *autospazio* relativo all'autovalore λ (per convenzione il vettore zero è nell'autospazio ma non è autovettore). Lo *spettro* di A è l'insieme degli scalari $\lambda \in \mathbb{C}$ tali che $\lambda \text{id}_X - A$ non è biiettivo.

In dimensione finita lo spettro coincide con l'insieme degli autovalori, ma in dimensione infinita $\lambda \text{id}_X - A$ può essere iniettivo senza anche essere suriettivo.

Ogni autospazio di un operatore lineare $A: X \rightarrow X$ è un sottospazio vettoriale, in quanto nucleo di un operatore del tipo $\lambda \text{id}_X - A$, che è lineare. Se X è anche normato, gli autospazi degli operatori lineari continui sono sottinsiemi chiusi di X , perché contrimmagine del chiuso $\{0\} \subseteq X$ tramite la funzione $\lambda \text{id}_X - A$, che è continua.

Notiamo che se X è normato e $\lambda \in \mathbb{C}$ è autovalore di A , allora $|\lambda| \leq \|A\|$.

Esercizio 5.1.2. Sia $X = \mathbb{R}^2$ con la norma euclidea, e sia $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Verificare che $\lambda = 1$ è l'unico autovalore, e che

$$1 = |\lambda| < \|A\| = \sqrt{\frac{3 + \sqrt{5}}{2}}.$$

Esercizio 5.1.3. Sia X uno spazio di Banach, $A: X \rightarrow X$ lineare continuo, e $\lambda \in \mathbb{C}$ tale che $|\lambda| > \|A\|$. Dimostrare che λ non è nello spettro di A . (Mostrare che la funzione $f(z) := (y + Az)/\lambda$ è una contrazione, e quindi l'equazione $\lambda x - Ax = y$ ha come unica soluzione x il punto fisso di f).

Svilupperemo una teoria spettrale quasi soltanto negli spazi di Hilbert, soprattutto per operatori autoaggiunti e compatti.

5.2 Autospazi di operatori compatti

Proposizione 5.2.1. *Sia X uno spazio normato, $A: X \rightarrow X$ un operatore compatto, λ un autovalore non nullo di A . Allora l'autospazio di λ ha dimensione finita.*

Dimostrazione. L'autospazio $Z_\lambda = \ker(\lambda \text{id}_X - A)$ è un sottospazio vettoriale chiuso di A in quanto nucleo di un operatore lineare continuo. La restrizione di A all'autospazio manda Z_λ in sé e ivi agisce come multiplo dell'identità: se $x \in Z_\lambda$ allora $Ax = \lambda x$. Prendiamo una successione $x_n \in Z_\lambda$ limitata. Poiché A è compatta, possiamo estrarre una sottosuccessione tale che $Ax_{\varphi(n)} = \lambda x_{\varphi(n)}$ sia di Cauchy. Ma $\lambda \neq 0$, e quindi $x_{\varphi(n)} = \frac{1}{\lambda} Ax_{\varphi(n)}$ stesso è di Cauchy. Nello spazio normato Z_λ ogni successione limitata ha una sottosuccessione di Cauchy. Quindi Z_λ ha dimensione finita (Prop. 2.1.7 p. 72). \square

Ogni operatore lineare da \mathbb{C}^n in sé ha almeno un autovalore, che è uno zero del polinomio caratteristico. In dimensione infinita un operatore lineare, anche compatto, non è detto che abbia autovalori.

Esempio 5.2.2. Riprendiamo l'esempio 3.1.15 p. 90 con una modifica. Nello spazio di Banach $X = \ell^p(\mathbb{N})$ con $1 < p < +\infty$ consideriamo l'operatore $A: X \rightarrow X$ che manda il vettore $x = (x(1), x(2), x(3), \dots, x(n), \dots)$ in $Ax = (0, x(1)/2, x(2)/3, x(3)/4, \dots, x(n-1)/n, \dots)$. Si vede che A è compatto, in quanto mappa la palla unitaria nel cubo di Hilbert. Questo operatore non ha nessun autovalore. Lo zero non è autovalore di A , come si vede subito. Supponiamo che $\lambda \neq 0$ sia un autovalore con x autovettore. Allora imporre che $Ax = \lambda x$ significa che

$$0 = \lambda x(1), \quad \frac{x(n-1)}{n} = \lambda x(n) \quad \text{per } n > 1$$

da cui segue per induzione che $x(n) = 0$ per ogni n .

5.3 Operatori autoaggiunti

Definizione 5.3.1. Un operatore A lineare continuo da uno spazio di Hilbert H in se stesso si dice *autoaggiunto* se $A = A^*$, cioè se

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad \forall x, y \in H.$$

Nel caso particolare in cui $H = \mathbb{C}^n$, la matrice A è autoaggiunta se è hermitiana, cioè coincide con la sua trasposta coniugata. Se $I: H \rightarrow H$ è l'operatore identità, è chiaro che $I^* = I$.

Proposizione 5.3.2. *Sia H uno spazio di Hilbert complesso e $A: H \rightarrow H$ un operatore lineare continuo e autoaggiunto. Allora*

1. *gli autovalori di A sono reali;*
2. *autovettori corrispondenti ad autovalori distinti sono fra loro ortogonali;*
3. *il nucleo di A coincide col sottospazio ortogonale all'immagine di A , cioè $\ker A = A(H)^\perp$.*

Dimostrazione. Sia $\lambda \in \mathbb{C}$ un autovalore di A con autovettore $\bar{x} \in H \setminus \{0\}$, per cui $A\bar{x} = \lambda\bar{x}$. Allora

$$\langle A\bar{x}, \bar{x} \rangle = \begin{cases} \langle \lambda\bar{x}, \bar{x} \rangle = \lambda\langle \bar{x}, \bar{x} \rangle = \lambda\|\bar{x}\|^2, \\ \langle \bar{x}, A\bar{x} \rangle = \langle \bar{x}, \lambda\bar{x} \rangle = \bar{\lambda}\langle \bar{x}, \bar{x} \rangle = \bar{\lambda}\|\bar{x}\|^2. \end{cases}$$

Uguagliando i risultati si ottiene $\lambda\|\bar{x}\|^2 = \bar{\lambda}\|\bar{x}\|^2$, da cui, dividendo per $\|\bar{x}\|^2 \neq 0$ si ricava che $\lambda = \bar{\lambda}$, cioè $\lambda \in \mathbb{R}$.

Siano $\lambda_1 \neq \lambda_2$ due autovalori di A con rispettivi autovettori $\bar{x}_1, \bar{x}_2 \in H \setminus \{0\}$. I due autovalori non sono entrambi nulli, supponiamo che $\lambda_1 \neq 0$. Allora

$$\begin{aligned} \langle \bar{x}_1, \bar{x}_2 \rangle &= \left\langle \frac{1}{\lambda_1} \lambda_1 \bar{x}_1, \bar{x}_2 \right\rangle = \frac{1}{\lambda_1} \langle \lambda_1 \bar{x}_1, \bar{x}_2 \rangle = \frac{1}{\lambda_1} \langle A\bar{x}_1, \bar{x}_2 \rangle = \\ &= \frac{1}{\lambda_1} \langle \bar{x}_1, A\bar{x}_2 \rangle = \frac{1}{\lambda_1} \langle \bar{x}_1, \lambda_2 \bar{x}_2 \rangle = \frac{\lambda_2}{\lambda_1} \langle \bar{x}_1, \bar{x}_2 \rangle. \end{aligned}$$

Portando l'ultimo membro al primo membro ricaviamo

$$\langle \bar{x}_1, \bar{x}_2 \rangle - \frac{\lambda_2}{\lambda_1} \langle \bar{x}_1, \bar{x}_2 \rangle = 0 \quad \iff \quad \frac{\lambda_1 - \lambda_2}{\lambda_1} \langle \bar{x}_1, \bar{x}_2 \rangle = 0,$$

da cui si ricava che $\langle \bar{x}_1, x_2 \rangle = 0$, cioè $x_1 \perp x_2$.

Abbiamo infine che

$$\begin{aligned}
 x \in \ker A &\iff Ax = 0 \\
 &\iff (\forall y \in H \quad \langle Ax, y \rangle = 0) \\
 &\iff (\forall y \in H \quad \langle x, Ay \rangle = 0) \\
 &\iff (\forall z \in A(H) \quad \langle x, z \rangle = 0) \\
 &\iff (\forall z \in A(H) \quad x \perp z) \\
 &\iff x \in A(H)^\perp.
 \end{aligned}$$

□

Quando A è autoaggiunto, il nucleo e l'immagine sono fra loro ortogonali. Però in dimensione infinita l'immagine può non essere chiusa, per cui non è vero in generale che H sia la somma diretta del nucleo e dell'immagine. Quello che si può dire è che

$$H = (\ker A) \oplus \text{cl } A(H),$$

dove cl indica la chiusura (usare la formula (1.48) p. 35).

Esempio 5.3.3. Riprendiamo l'esempio 3.1.15 p. 90. Nello spazio di Hilbert $H = \ell^2(\mathbb{N})$ consideriamo l'operatore A che manda il vettore $x = (x(1), x(2), x(3), \dots, x(n), \dots)$ in $Ax = (x(1), x(2)/2, x(3)/3, \dots, x(n)/n, \dots)$. Abbiamo già visto che A è compatto, in quanto mappa la palla unitaria nel cubo di Hilbert. Si verifica che è anche autoaggiunto:

$$\langle Ax, y \rangle = \sum_{n \in \mathbb{N}} \frac{x(n)}{n} \cdot \overline{y(n)} = \sum_{n \in \mathbb{N}} x(n) \cdot \frac{\overline{y(n)}}{n} = \langle x, Ay \rangle.$$

La matrice infinita che rappresenta A è diagonale con numeri reali sulla diagonale, e quindi coincide con la trasposta coniugata. I numeri $1/n$ sulla diagonale sono tutti autovalori. Il nucleo di A è il solo 0, come si vede facilmente. L'immagine $A(H)$ è densa in H , perché il suo ortogonale è $\ker A = \{0\}$. Però $A(H)$ non può essere tutto H , perché altrimenti A sarebbe un omeomorfismo compatto, il che è possibile solo in dimensione finita (Proposizione 3.4.2 p. 99).

5.4 Operatori autoaggiunti compatti

Teorema 5.4.1. *Sia H uno spazio di Hilbert (non ridotto al solo 0) e $A: H \rightarrow H$ un operatore lineare compatto e autoaggiunto. Consideriamo la forma quadratica associata*

$$T(x) := \langle Ax, x \rangle,$$

che ha valori reali. Allora

1. sulla sfera unitaria di H il valore assoluto della forma quadratica T assume valore massimo $\lambda_0 \geq 0$:

$$\lambda_0 = \max_{\|x\|=1} |T(x)|;$$

2. λ_0 coincide con la norma operatoriale di A ;
3. o λ_0 o $-\lambda_0$ è un autovalore di A , e gli autoversori x sono quelli che rendono $T(x) = \pm\lambda_0$.

Dimostrazione. La T ha valori reali perché

$$T(x) = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle} = \overline{\langle A^*x, x \rangle} = \overline{\langle Ax, x \rangle} = \overline{T(x)}.$$

Sia λ_0 l'estremo superiore di $|T|$ sulla sfera unitaria:

$$\lambda_0 := \sup_{\|x\|=1} |T(x)|.$$

Chiaramente

$$|T(x)| \leq \lambda_0 \|x\|^2 \quad \forall x \in H.$$

Se $\|x\| = 1$ allora

$$|T(x)| = |\langle Ax, x \rangle| \leq \|Ax\| \cdot \|x\| \leq \|A\| \cdot \|x\|^2 = \|A\|.$$

Quindi $\lambda_0 \leq \|A\|$. Con conti simili a quelli dell'identità di polarizzazione, possiamo derivare T da una forma bilineare:

$$\begin{aligned} T(x+y) - T(x-y) &= \langle A(x+y), x+y \rangle - \langle A(x-y), x-y \rangle = \\ &= \langle Ax, x \rangle + \langle Ax, y \rangle + \langle Ay, x \rangle + \langle Ay, y \rangle - \end{aligned}$$

$$\begin{aligned}
& - \langle Ax, x \rangle + \langle Ax, y \rangle + \langle Ay, x \rangle - \langle Ay, y \rangle = \\
& = 2\langle Ax, y \rangle + 2\langle Ay, x \rangle = 2\langle Ax, y \rangle + 2\langle y, Ax \rangle = \\
& = 2\langle Ax, y \rangle + 2\overline{\langle Ax, y \rangle} = \\
& = 4\Re\langle Ax, y \rangle.
\end{aligned}$$

Per l'identità del parallelogrammo, o semplicemente espandendo i quadrati,

$$\begin{aligned}
|T(x+y)| + |T(x-y)| & \leq \lambda_0 \|x+y\|^2 + \lambda_0 \|x-y\|^2 = \\
& = \lambda_0 (2\|x\|^2 + 2\|y\|^2) = \\
& = 2\lambda_0 (\|x\|^2 + \|y\|^2).
\end{aligned}$$

Quindi vale la seguente disuguaglianza per ogni $x, y \in H$:

$$\begin{aligned}
4\Re\langle Ax, y \rangle & = T(x+y) - T(x-y) \leq \\
& \leq |T(x+y)| + |T(x-y)| \leq \\
& = 2\lambda_0 (\|x\|^2 + \|y\|^2).
\end{aligned}$$

Prendiamo in particolare

$$\|x\| = 1 \quad \text{con} \quad Ax \neq 0 \quad \text{e poniamo} \quad y = \frac{Ax}{\|Ax\|}.$$

Sostituendo nella disuguaglianza otteniamo

$$4\Re\left\langle Ax, \frac{Ax}{\|Ax\|} \right\rangle \leq 2\lambda_0 \left(\|x\|^2 + \left\| \frac{Ax}{\|Ax\|} \right\|^2 \right),$$

che si semplifica in

$$4 \frac{1}{\|Ax\|} \Re\langle Ax, Ax \rangle \leq 2\lambda_0 (1^2 + 1^2),$$

e successivamente in

$$4 \frac{1}{\|Ax\|} \Re\|Ax\|^2 \leq 4\lambda_0.$$

ossia in

$$\|Ax\| \leq \lambda_0.$$

Valendo questo per ogni x della sfera unitaria (ovviamente anche per quelli con $Ax = 0$), ricaviamo che $\|A\| \leq \lambda_0$. Tenendo conto che avevamo già visto che $\lambda_0 \leq \|A\|$, otteniamo che

$$\lambda_0 = \|A\|.$$

Sia x_n una successione massimizzante, cioè con $\|x_n\| = 1$ e tale che $|T(x_n)| \rightarrow \|A\|$. A meno di passare a una sottosuccessione, possiamo supporre di essere in uno dei due casi seguenti:

1. $T(x_n) \rightarrow \|A\|$ oppure
2. $T(x_n) \rightarrow -\|A\|$

per $n \rightarrow +\infty$. Supponiamo di essere nel primo caso. Possiamo scrivere

$$\begin{aligned} 0 &\leq \left\| \|A\|x_n - Ax_n \right\|^2 = \\ &= \|A\|^2 \|x_n\|^2 + \|Ax_n\|^2 - 2\Re \langle \|A\|x_n, Ax_n \rangle = \\ &= \|A\|^2 + \|Ax_n\|^2 - 2\|A\| \Re \langle x_n, Ax_n \rangle = \\ &= \|A\|^2 + \|Ax_n\|^2 - 2\|A\|T(x_n) \leq \\ &\leq \|A\|^2 + \|A\|^2 \cdot \|x_n\|^2 - 2\|A\|T(x_n) = \\ &= \|A\|^2 + \|A\|^2 - 2\|A\|T(x_n) = \\ &= 2\|A\|^2 - 2\|A\|T(x_n). \end{aligned}$$

Facendo tendere $n \rightarrow +\infty$ l'ultimo membro tende a $2\|A\|^2 - 2\|A\| \cdot \|A\| = 0$. Quindi

$$\|A\|x_n - Ax_n \longrightarrow 0 \quad \text{per } n \rightarrow +\infty.$$

Usiamo finalmente l'ipotesi di compattezza: poiché x_n è limitata, esistono una $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente e un $\bar{x} \in H$ tali che

$$x_{\varphi(n)} \rightarrow \bar{x} \in H \quad \text{e} \quad Ax_{\varphi(n)} \rightarrow A\bar{x} \quad \text{per } n \rightarrow +\infty.$$

Possiamo scrivere

$$\underbrace{\|A\|x_{\varphi(n)}}_{\rightarrow \|A\|\bar{x}} = \underbrace{Ax_{\varphi(n)}}_{\rightarrow A\bar{x}} + \underbrace{\left(\|A\|x_{\varphi(n)} - Ax_{\varphi(n)} \right)}_{\rightarrow 0} \longrightarrow 0 \quad \text{per } n \rightarrow +\infty.$$

La successione $n \mapsto \|A\|x_{\varphi(n)}$ tende quindi debolmente ad $\|A\|\bar{x}$ e fortemente ad $A\bar{x}$. Per l'unicità del limite debole concludiamo che

$$A\bar{x} = \|A\|\bar{x}.$$

Non saltiamo alla conclusione che \bar{x} sia un autovettore di A , perché ci manca da dimostrare che $\bar{x} \neq 0$. Se per caso \bar{x} fosse effettivamente nullo, allora $Ax_{\varphi(n)} \rightarrow A\bar{x} = A0 = 0$ e quindi

$$|T(x_{\varphi(n)})| = |\langle Ax_{\varphi(n)}, x_{\varphi(n)} \rangle| \leq \|Ax_{\varphi(n)}\| \cdot \|x_{\varphi(n)}\| = \|Ax_{\varphi(n)}\| \rightarrow 0.$$

D'altra parte il primo membro tende a λ_0 . Quindi $\lambda_0 = \|A\| = 0$, che è possibile solo nel caso banale in cui $A \equiv 0$, quando tutti i vettori non nulli sono autovettori. Se A non è identicamente nullo, $\bar{x} \neq 0$ è un autovettore di autovalore $\|A\|$ e il suo versore realizza il massimo $|T(x)| = |\langle Ax, x \rangle| = |\langle \|A\|x, x \rangle| = \|A\| \cdot \|x\|^2 = \lambda_0$ sulla sfera unitaria.

Sia x un versore tale che $T(x) = \lambda_0 = \|A\|$. Ripetendo il conto di prima con $x_n = x$ si ottiene

$$0 \leq \left\| \|A\|x - Ax \right\|^2 \leq 2\|A\|^2 - 2\|A\|T(x) = 0,$$

da cui si evince che $Ax = \|A\|x$, cioè x è un autovettore di A con autovalore λ_0 . Viceversa, se x è un autoversore con $Ax = \|A\|x$ è banale che $T(x) = \|A\|$.

Il caso $T(x_n) \rightarrow -\|A\|$ si tratta in modo analogo, cambiando alcuni segni. Si noti che i due casi non sono esclusivi: possono esistere due sottosuccessioni, per una delle quali il limite è $\|A\|$ e per l'altra è $-\|A\|$, producendo due autovalori con autospazi distinti. \square

Teorema 5.4.2 (diagonalizzazione). *Sia H uno spazio di Hilbert e $A: H \rightarrow H$ un operatore lineare compatto e autoaggiunto (e non identicamente nullo). Allora esiste una base hilbertiana dell'immagine $A(H)$ formata da autovettori di A . Più precisamente:*

1. se $n = \dim A(H) < +\infty$ allora esistono n autovalori non nulli $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R} \setminus \{0\}$ (non necessariamente distinti) con rispettivi autoversori u_1, u_2, \dots, u_n , che formano una base ortonormale di $A(H)$, tali che

$$Ax = \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k \quad \forall x \in H.$$

2. se $A(H)$ ha dimensione infinita, esiste una successione $n \mapsto \lambda_n \in \mathbb{R} \setminus \{0\}$ di autovalori non nulli tali che $|\lambda_1| = \|A\|$, $|\lambda_{n+1}| \leq |\lambda_n|$ per ogni n , e tali che $\lambda_n \rightarrow 0$ per $n \rightarrow +\infty$, ed esiste una successione di autoversori associati $n \mapsto u_n \in H$, che formano una base hilbertiana della chiusura di $A(H)$, e tali che

$$Ax = \sum_{k \in \mathbb{N}} \lambda_k \langle x, u_k \rangle u_k \quad \forall x \in H.$$

In entrambi i casi gli autovalori non nulli di A sono tutti e soli i λ_k .

Dimostrazione. Abbiamo che $\|A\| > 0$, poiché A è non nullo. Dal teorema precedente sappiamo che $\lambda_1 = \pm\|A\| \neq 0$ è un autovalore di A e sia u_1 un autoversore associato. Poniamo $M_1 = \text{span}\{u_1\}$, che è un sottospazio chiuso di H . Verifichiamo che A manda M_1^\perp in se stesso:

$$\begin{aligned} x \in M_1^\perp &\iff \langle x, \alpha u_1 \rangle = 0 \quad \forall \alpha \in \mathbb{C} \\ &\iff \langle x, u_1 \rangle = 0 \\ &\iff \langle x, \frac{1}{\lambda_1} Au_1 \rangle = 0 \\ &\iff \langle x, Au_1 \rangle = 0 \\ &\iff \langle Ax, u_1 \rangle = 0 \\ &\iff \langle Ax, \alpha u_1 \rangle = 0 \quad \forall \alpha \in \mathbb{C} \\ &\iff Ax \perp M_1 \\ &\iff Ax \in M_1^\perp \end{aligned}$$

(controllare che le equivalenze valgano in entrambe le direzioni). La restrizione $A|_{M_1^\perp}: M_1^\perp \rightarrow M_1^\perp$ è compatta e autoaggiunta. Se per caso $A|_{M_1^\perp}$ è identicamente nulla ci fermiamo. Altrimenti, la sua norma operatoriale è non maggiore di quella di A :

$$0 < \|A|_{M_1^\perp}\| = \sup_{\substack{\|x\|=1 \\ x \in M_1^\perp}} \|Ax\| \leq \sup_{\substack{\|x\|=1 \\ x \in H}} \|Ax\| = \|A\| = |\lambda_1|.$$

Altrimenti possiamo applicare il teorema precedente alla restrizione $A|_{M_1^\perp}$ e ricavare un autovalore $\lambda_2 = \pm\|A|_{M_1^\perp}\|$ e un corrispondente autoversore $u_2 \in M_1^\perp$. Questi sono autovalori e autoversori anche della A di partenza.

Per costruzione $u_1 \perp u_2$. Sia $M_2 = \text{span}\{u_1, u_2\}$. Verifichiamo che A manda M_2^\perp in se stesso:

$$\begin{aligned}
x \in M_2^\perp &\iff \langle x, \alpha_1 u_1 + \alpha_2 u_2 \rangle = 0 \quad \forall \alpha_1, \alpha_2 \in \mathbb{C} \\
&\iff \langle x, u_1 \rangle = 0 \wedge \langle x, u_2 \rangle = 0 \\
&\iff \langle x, \frac{1}{\lambda_1} A u_1 \rangle = 0 \wedge \langle x, \frac{1}{\lambda_2} A u_2 \rangle = 0 \\
&\iff \langle x, A u_1 \rangle = 0 \wedge \langle x, A u_2 \rangle = 0 \\
&\iff \langle A x, u_1 \rangle = 0 \wedge \langle A x, u_2 \rangle = 0 \\
&\iff \langle A x, \alpha_1 u_1 + \alpha_2 u_2 \rangle = 0 \quad \forall \alpha_1, \alpha_2 \in \mathbb{C} \\
&\iff A x \perp M_2 \\
&\iff A x \in M_2^\perp.
\end{aligned}$$

La restrizione $A|_{M_2^\perp}: M_2^\perp \rightarrow M_2^\perp$ è compatta e autoaggiunta. Se per caso $A|_{M_2^\perp}$ è identicamente nulla ci fermiamo. Altrimenti, la sua norma operatoriale è non maggiore di quella di A :

$$0 < \|A|_{M_2^\perp}\| = \sup_{\substack{\|x\|=1 \\ x \in M_2^\perp}} \|A x\| \leq \sup_{\|x\|=1} \|A x\| = \|A\| = |\lambda_2|.$$

Altrimenti possiamo applicare il teorema precedente alla restrizione $A|_{M_2^\perp}$ e ricavare un autovalore $\lambda_3 = \pm \|A|_{M_2^\perp}\|$ e un corrispondente autoversore $u_3 \in M_2^\perp$.

E così via. Il procedimento può terminare arrivati a un sottospazio $M_n = \text{span}\{u_1, \dots, u_n\}$ tale che A sia identicamente nullo su M_n^\perp , mentre gli u_1, \dots, u_n sono autovettori ortonormali con autovalori non nulli $\lambda_1, \dots, \lambda_n$. In tal caso per ogni $x \in H = M_n \oplus M_n^\perp$ possiamo scrivere

$$\begin{aligned}
A x &= A(P_{M_n} x + P_{M_n^\perp} x) = A P_{M_n} x = A \sum_{k=1}^n \langle x, u_k \rangle u_k = \\
&= \sum_{k=1}^n \langle x, u_k \rangle A u_k = \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k,
\end{aligned}$$

e $A(H) = M_n$ ha dimensione finita.

L'altra possibilità è che la procedura non termini mai, generando una successione u_n di autovettori ortonormali, ciascuno col suo autovalore $\lambda_n \neq 0$

e tali che $|\lambda_{n+1}| \leq |\lambda_n|$. In particolare l'immagine $A(H)$ ha dimensione infinita, contenendo gli infiniti vettori linearmente indipendenti $\lambda_n u_n$. Come sopra poniamo $M_n = \text{span}\{u_1, \dots, u_n\}$ e sia M la chiusura dello spazio generato da tutti gli u_n . Come tutti i sistemi ortonormali numerabili, $u_n \rightarrow 0$ (esempio 4.3.2 p. 110). Dato che A è compatto ricaviamo che Au_n converge fortemente ad $A0 = 0$, per cui

$$|\lambda_n| = \|\lambda_n u_n\| = \|Au_n\| \rightarrow \|A0\| = 0.$$

Per ogni $x \in H = M_n \oplus M_n^\perp$ vale

$$x = P_{M_n^\perp} x + P_{M_n} x = P_{M_n^\perp} x + \sum_{k=1}^n \langle x, u_k \rangle u_k,$$

per cui

$$\begin{aligned} Ax &= AP_{M_n^\perp} x + AP_{M_n} x = AP_{M_n^\perp} x + \sum_{k=1}^n \langle x, u_k \rangle Au_k = \\ &= AP_{M_n^\perp} x + \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k, \end{aligned}$$

e quindi

$$\left\| Ax - \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k \right\| = \|AP_{M_n^\perp} x\| \leq \|A|_{M_n^\perp}\| \cdot \|x\| = |\lambda_{n+1}| \cdot \|x\| \rightarrow 0$$

per $n \rightarrow +\infty$. Concludiamo che per ogni $x \in H$ vale l'espansione

$$Ax = \sum_{k=1}^{+\infty} \lambda_k \langle x, u_k \rangle u_k$$

nel senso del limite delle somme parziali. La somma è sommabile anche nel senso delle somme infinite, perché i λ_n sono limitati.

Infine supponiamo che $\mu \in \mathbb{R}$ sia un autovalore di A diverso da tutti i λ_n , con autoversore $\bar{x} \in H$. Essendo $\mu \neq \lambda_n$, necessariamente $\bar{x} \perp u_n$ (proposizione 5.3.2 p. 117). Quindi

$$\mu \bar{x} = A\bar{x} = \sum_{k \in \mathbb{N}} \lambda_k \underbrace{\langle \bar{x}, u_k \rangle}_{=0} u_k = 0,$$

da cui $\mu = 0$. Non è detto che zero sia autovalore di A , come abbiamo visto nell'esempio 5.3.3 p. 118. \square

Il teorema di diagonalizzazione si può anche invertire e diventa una caratterizzazione degli operatori compatti autoaggiunti:

Proposizione 5.4.3. *Sia H uno spazio di Hilbert e $A: H \rightarrow H$ un operatore non identicamente nullo. Allora si equivalgono le condizioni seguenti:*

1. A è lineare, autoaggiunto e compatto;
2. Ax si può rappresentare nel modo seguente

$$Ax = \sum_{k \in \Lambda} \lambda_k \langle x, u_k \rangle u_k. \quad (5.1)$$

dove $\Lambda = \mathbb{N}$ oppure $\Lambda = \{1, 2, 3, \dots, n\}$, $\{u_k : k \in \Lambda\}$ è un sistema ortonormale in H , $\{\lambda_k : k \in \Lambda\} \subset \mathbb{R} \setminus \{0\}$ è una famiglia di numeri tali che $|\lambda_{k_1}| \leq |\lambda_{k_2}|$ quando $k_1 > k_2$ e, nel caso numerabile, con $\lambda_k \rightarrow 0$ per $k \rightarrow +\infty$.

Dimostrazione. L'implicazione in una direzione è il teorema precedente, per cui rimane da dimostrare l'altra. Mettiamoci nel caso più difficile in cui $\Lambda = \mathbb{N}$. La somma della formula (5.1) è sommabile, perché i λ_k sono limitati, ed è lineare in x . Poniamo come al solito $M_n = \text{span}\{u_1, \dots, u_n\}$ e sia P_{M_n} la proiezione ortogonale di H su M_n , e poniamo

$$A_n x := \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k.$$

Questa $A_n: H \rightarrow H$ è facilmente continua ed ha valori nello spazio di dimensione finita M_n , per cui è compatta. È anche autoaggiunta con verifica diretta:

$$\begin{aligned} \langle A_n x, y \rangle &= \left\langle \sum_{k=1}^n \lambda_k \langle x, u_k \rangle u_k, y \right\rangle = \sum_{k=1}^n \lambda_k \langle x, u_k \rangle \langle u_k, y \rangle = \\ &= \sum_{k=1}^n \langle x, \lambda_k \overline{\langle u_k, y \rangle} u_k \rangle = \sum_{k=1}^n \langle x, \lambda_k \langle y, u_k \rangle u_k \rangle = \\ &= \left\langle x, \sum_{k=1}^n \lambda_k \langle y, u_k \rangle u_k \right\rangle = \langle x, A_n y \rangle. \end{aligned}$$

Preso un $x \in H$ si può scrivere

$$\begin{aligned} \|(A - A_n)x\|^2 &= \|Ax - A_nx\|^2 = \left\| \sum_{k=n+1}^{+\infty} \lambda_k \langle x, u_k \rangle u_k \right\|^2 = \\ &= \sum_{k=n+1}^{+\infty} |\lambda_k|^2 \cdot |\langle x, u_k \rangle|^2 \leq |\lambda_{n+1}|^2 \sum_{k=n+1}^{+\infty} |\langle x, u_k \rangle|^2 \leq \\ &\leq |\lambda_{n+1}|^2 \cdot \|x\|^2, \end{aligned}$$

da cui deduciamo che le A_n convergono ad A nella norma operatoriale:

$$\|A - A_n\| \leq |\lambda_{n+1}| \rightarrow 0.$$

Concludiamo che A è pure compatto (corollario 3.1.11 p. 88). Facilmente si ottiene che A è autoaggiunta grazie al fatto che le A_n lo sono.

Il caso in cui $\Lambda = \{1, \dots, n\}$ è banale a questo punto. \square

5.5 L'alternativa di Fredholm

Lavoreremo in uno spazio di Hilbert H , e indicheremo con $I = \text{id}_H$ l'operatore identità su H .

Teorema 5.5.1. *Sia H uno spazio di Hilbert, $A: H \rightarrow H$ un operatore compatto, e poniamo $T = I - A$. Allora*

1. $T(H) = (\ker T^*)^\perp$, $T^*(H) = (\ker T)^\perp$, da cui in particolare $H = T(H) \oplus (\ker T^*) = T^*(H) \oplus (\ker T)$;
2. T è biiettivo se e solo se T è iniettivo se e solo se T^* è iniettivo se e solo se T^* è biiettivo;
3. $\dim \ker T = \dim \ker T^* < +\infty$.

Riguardo a iniettività e suriettività, il teorema lascia come uniche *alternative* che T abbia entrambe o nessuna. In altre parole, o 1 è autovalore di A , nel qual caso $T = I - A$ non è suriettiva, oppure 1 non è autovalore, e in tal caso $T = I - A$ è biettiva. Il risultato è semplice in dimensione finita, non altrettanto in dimensione infinita.

Dimostrazione. La dimostrazione è lunga e divisa in passi.

Passo 1. $T(H)$ è chiuso in H .

Sia $\bar{y} \in \text{cl}T(H)$. Dobbiamo dimostrare che esiste $\bar{x} \in H$ tale che $\bar{y} = T\bar{x} = \bar{x} - A\bar{x}$. Per definizione di chiusura in uno spazio metrico, esiste una successione $x_n \in H$ tale che $y_n := Tx_n = x_n - Ax_n \rightarrow \bar{y}$. Possiamo supporre che $x_n \in (\ker T)^\perp$, eventualmente rimpiazzando x_n con $P_{(\ker T)^\perp}x_n$, giacché

$$\begin{aligned} y_n &= Tx_n = T(P_{\ker T}x_n + P_{(\ker T)^\perp}x_n) = T(\overbrace{P_{\ker T}x_n}^{\in \ker T}) + T(P_{(\ker T)^\perp}x_n) = \\ &= T(P_{(\ker T)^\perp}x_n). \end{aligned}$$

Dimostriamo che x_n è limitato. Se per assurdo non lo fosse, a meno di passare a sottosuccessione potremmo assumere che $\|x_n\| \rightarrow +\infty$. Osserviamo che y_n è limitata, in quanto convergente. Dividiamo l'equazione $y_n = x_n - Ax_n$ per $\|x_n\|$:

$$\frac{y_n}{\|x_n\|} = \frac{x_n}{\|x_n\|} - A \frac{x_n}{\|x_n\|}.$$

Il primo membro tende a zero in norma, perché y_n è limitato e il denominatore tende all'infinito. La successione $x_n/\|x_n\|$ è limitata e A è compatto, per cui, a meno di passare a una ulteriore sottosuccessione, possiamo supporre che $A(x_n/\|x_n\|)$ converga in norma a un qualche $\bar{z} \in H$. Isolando il primo termine nel membro di destra, troviamo che converge fortemente a \bar{z} :

$$\frac{x_n}{\|x_n\|} = \frac{y_n}{\|x_n\|} + A \frac{x_n}{\|x_n\|} \longrightarrow 0 + \bar{z} = \bar{z}. \quad (5.2)$$

Dato che A è continua, applicandola al primo membro ricaviamo che

$$A \frac{x_n}{\|x_n\|} \rightarrow A\bar{z}.$$

Ma allora possiamo ricalcolare la formula (5.2) così:

$$\frac{x_n}{\|x_n\|} = \frac{y_n}{\|x_n\|} + A \frac{x_n}{\|x_n\|} \longrightarrow 0 + A\bar{z}. \quad (5.3)$$

Uguagliando i limiti di (5.2) e (5.3) otteniamo che $\bar{z} = A\bar{z}$, cioè $T\bar{z} = \bar{z} - A\bar{z} = 0$, ossia $\bar{z} \in \ker T$. Però avevamo preso $x_n \in (\ker T)^\perp$, che è chiuso, per cui \bar{z} deve stare pure in $(\ker T)^\perp$. Quindi $\bar{z} \in \ker T \cap (\ker T)^\perp = \{0\}$, cioè $\bar{z} = 0$. Questo è impossibile perché \bar{z} è il limite forte dei versori $x_n/\|x_n\|$. Dobbiamo ammettere che x_n è effettivamente limitata.

Estraiamo da x_n una sottosuccessione $x_{\varphi(n)}$ che converge debolmente a un qualche \bar{x} . Applicando l'operatore compatto A otteniamo che $Ax_{\varphi(n)} \rightarrow A\bar{x}$. Tornando all'uguaglianza $y_n = Tx_n = x_n - Ax_n$, facendo tendere $n \rightarrow +\infty$ la situazione è questa:

$$\underbrace{y_{\varphi(n)}}_{\rightarrow \bar{y}} = Tx_{\varphi(n)} = \underbrace{x_{\varphi(n)}}_{\rightarrow \bar{x}} - \underbrace{Ax_{\varphi(n)}}_{\rightarrow A\bar{x}} \rightarrow \bar{x} - A\bar{x} = T\bar{x}.$$

Uguagliando i limiti del primo e dell'ultimo membro ricaviamo che $\bar{y} = \bar{x} - A\bar{x} = T\bar{x}$, cioè $\bar{y} \in T(H)$.

Passo 2. $T^n(H)$ è chiusa per ogni $n \in \mathbb{N}$.

Il caso $n = 1$ è stato visto nel passo 1. Per $n = 2$ espandendo la composizione ricaviamo che $T^2 = (I - A)(I - A) = I - A - A + A^2 = I - (2I - A)A$. Quindi T^2 si può scrivere come differenza dell'identità e dell'operatore lineare $(2I - A)A$, che è compatto, in quanto composizione di un continuo con un compatto. Il passo 1, applicato con $(2I - A)A$ al posto di A , implica che anche T^2 ha immagine chiusa. Analogamente T^n si può scrivere come la differenza dell'identità con una composizione di operatori continui dei quali uno è A , ossia una combinazione a cui si può applicare il passo 1.

Passo 3. $\ker T^* = T(H)^\perp$ e $(\ker T^*)^\perp = T(H)$.

Il fatto che $\ker T^* = T(H)^\perp$ richiede soltanto che T sia lineare continuo, perché

$$\begin{aligned} x \in \ker T^* &\iff T^*x = 0 \\ &\iff (\forall y \in H \quad \langle T^*x, y \rangle = 0) \\ &\iff (\forall y \in H \quad \langle x, Ty \rangle = 0) \\ &\iff (\forall z \in T(H) \quad \langle x, z \rangle = 0) \\ &\iff x \in T(H)^\perp. \end{aligned}$$

Per il teorema della decomposizione ortogonale 1.5.7 di p. 34, poiché $T(H)$ è chiuso, $T(H) = (T(H)^\perp)^\perp = (\ker T^*)^\perp$.

Passo 4. Esiste un $\bar{k} \in \mathbb{N}$ tale che $T^{\bar{k}+1}(H) = T^{\bar{k}}(H)$.

Osserviamo che per ogni $x \in H$

$$T^{n+1}(H) \ni T^{n+1}(x) = T^n(Tx) \in T^n(H),$$

per cui $T^{n+1}(H) \subseteq T^n(H)$ per ogni n . Abbiamo la seguente catena infinita di operatori fra spazi via via sempre più piccoli andando verso destra:

$$H \xrightarrow{T} T(H) \xrightarrow{T} T^2(H) \xrightarrow{T} T^3(H) \xrightarrow{T} \dots \xrightarrow{T} T^{\bar{k}}(H) \xrightarrow{T} T^{\bar{k}+1}(H)$$

e tutti chiusi in H per il passo 2. Supponiamo per assurdo che le inclusioni siano tutte strette, cioè che $T^{n+1}(H) \subsetneq T^n(H)$ per ogni $n \in \mathbb{N}$. Per il teorema della decomposizione ortogonale possiamo scrivere

$$T^n(H) = T^{n+1}(H) \oplus M_{n+1}, \quad \text{con} \quad \{0\} \neq M_{n+1} \perp T^{n+1}(H).$$

Per ogni n possiamo quindi prendere un versore $u_n \in M_{n+1} \subset T^n(H)$. Avremo in particolare che gli u_n sono un sistema ortonormale, perché se $n < m$ allora $u_n \in M_{n+1} \perp T^{n+1}(H) \supseteq T^m(H) \ni u_m$. Sempre per $n < m$

$$\begin{aligned} Au_n - Au_m &= (u_n - Tu_n) - (u_m - Tu_m) = \\ &= \underbrace{u_n}_{\in M_{n+1}} - \underbrace{\left(\underbrace{u_m}_{\in T^m(H)} + \underbrace{Tu_n}_{\in T^{n+1}(H)} - \underbrace{Tu_m}_{\in T^{m+1}(H)} \right)}_{\in T^{n+1}(H)}, \end{aligned}$$

di cui prendiamo le norme quadre, tenendo conto che abbiamo una somma di due vettori fra loro ortogonali (teorema di Pitagora):

$$\begin{aligned} \|Au_n - Au_m\|^2 &= \|u_n\|^2 + \|u_m + Tu_n - Tu_m\|^2 = \\ &= 1 + \|u_m + Tu_n - Tu_m\|^2 \geq \\ &\geq 1. \end{aligned}$$

Quindi la successione $n \mapsto Au_n$ non ha nessuna sottosuccessione di Cauchy. Questo è assurdo perché u_n è limitata ed A è compatto.

Passo 5. Se $T: H \rightarrow H$ è biiettivo se e solo se T è iniettivo, se e solo se $T^*: H \rightarrow H$ è iniettivo, se e solo se T^* è biiettivo.

È ovvio che se T è biiettivo allora è anche iniettivo. Per il viceversa, supponiamo che T sia iniettivo ma non suriettivo. Preso l'indice \bar{k} del passo 4, anche $T^{\bar{k}}$ è iniettiva, perché composizione di funzioni iniettive. Poiché T non è suriettivo, esiste un $y_0 \in H$ tale che $Tx \neq y_0$ per ogni $x \in H$. Applicando $T^{\bar{k}}$ ai due membri della disuguaglianza $Tx - y_0 \neq 0$ otteniamo che per ogni $x \in H$ la catena di deduzioni

$$\begin{aligned} Tx \neq y_0 &\Rightarrow Tx - y_0 \neq 0 \Rightarrow T^{\bar{k}}(Tx - y_0) \neq T^{\bar{k}}(0) = 0 \Rightarrow \\ &\Rightarrow T^{\bar{k}+1}x - T^{\bar{k}}y_0 \neq 0 \Rightarrow \\ &\Rightarrow T^{\bar{k}+1}x \neq T^{\bar{k}}y_0, \end{aligned}$$

e dall'ultima segue che $T^{\bar{k}}y_0$, che è un vettore di $T^{\bar{k}}(H)$, non appartiene a $T^{\bar{k}+1}(H)$. Ma questo contraddice il passo 4.

L'aggiunto hilbertiano $T^* = I - A^*$ è anch'esso del tipo identità meno compatto. Quindi T^* è biiettivo se e solo se è iniettivo.

Supponiamo che T sia biiettivo: allora $T(H) = H$. Per il passo 3 abbiamo che $\ker T^* = T(H)^\perp = H^\perp = \{0\}$, per cui T^* è iniettivo, e quindi biiettivo.

Passo 6. $\dim \ker T = \dim \ker T^* < +\infty$.

I nuclei di T e di T^* sono gli autospazi relativi all'(eventuale) autovalore 1 degli operatori compatti A, A^* rispettivamente, e quindi sappiamo che hanno dimensione finita (proposizione 5.2.1 p. 116). Resta da dimostrare che la dimensione è la stessa. Se una delle dimensioni è nulla, l'operatore corrispondente è iniettivo, per cui anche l'altra dimensione è nulla per il passo 5. Supponiamo pertanto che le dimensioni siano entrambe positive. Siano $\{u_1, \dots, u_n\}$ una base ortonormale di $\ker T$, e $\{v_1, \dots, v_m\}$ una base ortonormale di $\ker T^*$. Salvo scambiare fra loro T con T^* , possiamo assumere che $n \leq m$. Definiamo l'operatore lineare $B: H \rightarrow H$ così:

$$Sx := Tx - \sum_{k=1}^n \langle x, u_k \rangle v_k = x - Ax - \sum_{k=1}^n \langle x, u_k \rangle v_k = x - Bx$$

(notare che dentro il prodotto scalare ci sono gli u_k , fuori i v_k), dove il B definito da

$$Bx := Ax + \sum_{k=1}^n \langle x, u_k \rangle v_k$$

è un operatore lineare compatto, in quanto somma di un compatto con un operatore di rango finito. Con l'intento di dimostrare che S è iniettivo, sia $\bar{x} \in \ker S$. Allora

$$T(H) \ni T\bar{x} = \sum_{k=1}^n \langle \bar{x}, u_k \rangle v_k \in \ker T^*.$$

Ma $\ker T^* \perp T(H)$, per cui i due membri dell'uguaglianza sono nulli. Da una parte $T\bar{x} = 0$, cioè $\bar{x} \in \ker T$. Azzerando il membro di destra otteniamo che

$$\sum_{k=1}^n \langle \bar{x}, u_k \rangle v_k = 0,$$

ma essendo questa una combinazione lineare di vettori linearmente indipendenti, tutti i coefficienti devono essere nulli:

$$\langle \bar{x}, u_k \rangle = 0 \quad \forall k = 1, \dots, n,$$

cioè $\bar{x} \in (\ker T)^\perp$. Insomma, \bar{x} appartiene sia a $\ker T$ che al suo ortogonale, per cui $\bar{x} = 0$. Effettivamente $\ker S = \{0\}$, cioè S è iniettivo. Per il passo 5 S è biiettivo. In particolare $S(H) = H$.

Supponiamo per assurdo che $n < m$ e consideriamo il vettore v_{n+1} . Poiché S è suriettivo, deve esistere un $\bar{y} \in H$ tale che $v_{n+1} = S(\bar{y})$. Ma allora

$$\begin{aligned} 1 &= \|v_{n+1}\|^2 = \langle v_{n+1}, v_{n+1} \rangle = \langle S\bar{y}, v_{n+1} \rangle = \\ &= \left\langle T\bar{y} - \sum_{k=1}^n \langle \bar{y}, u_k \rangle v_k, v_{n+1} \right\rangle = \langle T\bar{y}, v_{n+1} \rangle - \sum_{k=1}^n \langle \bar{y}, u_k \rangle \underbrace{\langle v_k, v_{n+1} \rangle}_{=0} = \\ &= \langle T\bar{y}, v_{n+1} \rangle = \langle \bar{y}, T^* \underbrace{v_{n+1}}_{\in \ker T^*} \rangle = \langle \bar{y}, 0 \rangle = 0, \end{aligned}$$

cioè $1 = 0$, che è assurdo. Concludiamo che $n = m$. \square

Teorema 5.5.2 (dell'alternativa di Fredholm). *Sia H uno spazio di Hilbert, $A: H \rightarrow H$ un operatore lineare compatto, e $\mu \in \mathbb{C} \setminus \{0\}$. Vogliamo investigare la risolubilità dell'equazione $\mu x - Ax = y$, ossia $(\mu I - A)x = y$, nell'incognita x al variare di $y \in H$. A tal proposito ci sono due alternative:*

1. *se μ è autovalore di A , allora $\bar{\mu}$ è autovalore di A^* , gli autospazi relativi hanno la stessa dimensione finita, l'equazione $(\mu I - A)x = y$ ha soluzione se e solo se $y \perp \ker(\bar{\mu} I - A^*)$, e due soluzioni in x differiscono per un elemento di $\ker(\mu I - A)$;*
2. *se μ non è autovalore di A , allora $\mu I - A$ è invertibile con inversa continua e quindi per ogni $y \in H$ esiste unica la soluzione $x = (\mu I - A)^{-1}y$.*

Il teorema esclude una terza possibilità: quella che $\mu I - A$ possa essere iniettiva ma non suriettiva, che sarebbe concepibile qualora A non fosse compatta. In altre parole, a parte lo zero, lo spettro (def. 5.1.1 p. 115) di un operatore compatto coincide con l'insieme degli autovalori.

Dimostrazione. Per riportarci al teorema precedente, dividiamo l'equazione $\mu x - Ax = y$ per μ , che è per ipotesi diverso da zero:

$$\mu x - Ax = y \iff x - \frac{1}{\mu}Ax = \frac{y}{\mu} \iff \left(I - \frac{1}{\mu}A\right)x = \frac{y}{\mu}.$$

Posto

$$T = I - \frac{1}{\mu}A,$$

a questo T possiamo applicare i risultati del teorema precedente, perché A/μ è lineare continuo. Osserviamo anche che

$$T^* = I - \frac{1}{\mu}A^*,$$

e $T^*x = 0$ se e solo se $A^*x = \bar{\mu}x$.

Mettiamoci dapprima nel caso in cui μ è autovalore di A , cioè quando T non è iniettivo. Che due soluzioni di $\mu x - Ax = y$ differiscano per un elemento di $\ker(\mu I - A)$ è ovvio. Dal teorema precedente sappiamo allora neanche T^* è iniettivo, cioè $\bar{\mu}$ è autovalore di A^* , con autospazio della stessa dimensione dell'autospazio di A , e

$$\begin{aligned} (\exists x : \mu x - Ax = y) &\iff (\exists x : Tx = y/\mu) \iff y/\mu \in T(H) \\ &\iff y \in T(H) \iff y \in (\ker T^*)^\perp \\ &\iff y \in (\ker(\bar{\mu}I - A))^{\perp}. \end{aligned}$$

Supponiamo invece che μ non sia un autovalore di A . Allora T è iniettivo, e quindi, per il teorema precedente, anche biiettivo. Per il teorema della mappa aperta, l'inversa di T , e quindi anche $(\mu I - A)^{-1}$, è continua. Allora

$$\begin{aligned} \mu x - Ax = y &\iff x - \frac{1}{\mu}Ax = \frac{y}{\mu} \iff Tx = \frac{y}{\mu} \\ &\iff x = T^{-1}\frac{y}{\mu} = \left(I - \frac{1}{\mu}A\right)^{-1}\frac{y}{\mu} = (\mu I - A)^{-1}y. \end{aligned}$$

□

Sia H uno spazio di Hilbert, $\{u_n : n \in \mathbb{N}\}$ un sistema ortonormale in H , $n \mapsto \lambda_n \in \mathbb{R} \setminus \{0\}$ una successione infinitesima, e poniamo

$$Ax = \sum_{k \in \mathbb{N}} \lambda_k \langle x, u_k \rangle u_k$$

come nella proposizione 5.4.3 p. 126. Dato $\mu \in \mathbb{C} \setminus \{0\}$ e $y \in H$ ci proponiamo di risolvere esplicitamente l'equazione

$$\mu x - Ax = y$$

nell'incognita x . Sia $M = \text{clspan}\{u_n : n \in \mathbb{N}\}$. La componente di x ortogonale a M si trova tenendo conto che $Ax \in M$:

$$\begin{aligned} \mu x - Ax = y &\implies P_{M^\perp}(\mu x - Ax) = P_{M^\perp}y \iff \\ &\iff \mu P_{M^\perp}x - \mu P_{M^\perp}Ax = P_{M^\perp}y \iff \\ &\iff \mu P_{M^\perp}x - 0 = P_{M^\perp}y \iff \\ &\iff P_{M^\perp}x = \frac{1}{\mu}P_{M^\perp}y = \frac{1}{\mu}(y - P_M y) = \frac{1}{\mu}y - \frac{1}{\mu} \sum_{k \in \mathbb{N}} \langle y, u_k \rangle u_k. \end{aligned}$$

Cerchiamo la componente di x rispetto al vettore u_n prendendo il prodotto scalare dei due membri di $\mu x - Ax = y$ con u_n :

$$\begin{aligned} \langle \mu x - Ax, u_n \rangle &= \langle y, u_n \rangle \iff \\ \langle \mu x, u_n \rangle - \langle Ax, u_n \rangle &= \langle y, u_n \rangle \iff \\ \mu \langle x, u_n \rangle - \langle x, Au_n \rangle &= \langle y, u_n \rangle \iff \\ \mu \langle x, u_n \rangle - \langle x, \lambda_n u_n \rangle &= \langle y, u_n \rangle \iff \\ \mu \langle x, u_n \rangle - \lambda_n \langle x, u_n \rangle &= \langle y, u_n \rangle \iff \\ (\mu - \lambda_n) \langle x, u_n \rangle &= \langle y, u_n \rangle. \end{aligned} \tag{5.4}$$

Se $\mu \neq \lambda_n$ allora

$$\langle x, u_n \rangle = \frac{1}{\mu - \lambda_n} \langle y, u_n \rangle.$$

Si aprono ora le due alternative del teorema di Fredholm. Se μ non è autovalore di A , allora $\mu \neq \lambda_n$ per ogni n , e quindi ricaviamo la proiezione ortogonale di x su M :

$$P_M x = \sum_{k \in \mathbb{N}} \langle x, u_k \rangle u_k = \sum_{k \in \mathbb{N}} \frac{1}{\mu - \lambda_k} \langle y, u_k \rangle u_k.$$

e abbiamo la formula esplicita per la soluzione x :

$$\begin{aligned} x &= P_M x + P_{M^\perp} x = \sum_{k \in \mathbb{N}} \frac{1}{\mu - \lambda_k} \langle y, u_k \rangle u_k + \frac{1}{\mu} y - \frac{1}{\mu} \sum_{k \in \mathbb{N}} \langle y, u_k \rangle u_k = \\ &= \frac{1}{\mu} y + \sum_{k \in \mathbb{N}} \left(\frac{1}{\mu - \lambda_k} - \frac{1}{\mu} \right) \langle y, u_k \rangle u_k = \\ &= \frac{1}{\mu} y + \frac{1}{\mu} \sum_{k \in \mathbb{N}} \frac{\lambda_k}{\mu - \lambda_k} \langle y, u_k \rangle u_k \end{aligned}$$

e per l'inversa $(\mu I - A)^{-1}$:

$$(\mu I - A)^{-1}y = \frac{1}{\mu}y + \frac{1}{\mu} \sum_{k \in \mathbb{N}} \frac{\lambda_k}{\mu - \lambda_k} \langle y, u_k \rangle.$$

Si può verificare direttamente che il membro di destra è una funzione continua di y . Anzi, è un multiplo dell'identità più un operatore compatto.

L'altra alternativa è che μ sia autovalore di A , cioè esistano degli $n \in \mathbb{N}$ tali che $\mu = \lambda_n$. In questo caso condizione necessaria e sufficiente perché esista una soluzione di $\mu x - Ax = y$ è che $y \perp \ker(\mu I - A)^* = \ker(\bar{\mu}I - A^*) = \ker(\mu I - A)$, cioè

$$\langle y, u_n \rangle = 0 \quad \text{per ogni } n \in \mathbb{N} \text{ tali che } \mu = \lambda_n.$$

Osserviamo che questa condizione si può anche ricavare dall'equazione (5.4). Due soluzioni differiscono per un elemento di $\ker(\mu I - A)$:

$$\begin{aligned} P_M x &= \sum_{k \in \mathbb{N}} \langle x, u_k \rangle u_k = \sum_{\substack{k \in \mathbb{N} \\ \lambda_k \neq \mu}} \frac{1}{\mu - \lambda_k} \langle y, u_k \rangle + \\ &+ (\text{elem. di } \text{span}\{u_k : k \in \mathbb{N}, \lambda_k = \mu\}), \end{aligned}$$

per cui le soluzioni sono della forma generale seguente:

$$\begin{aligned} x &= P_M x + P_{M^\perp} x = \\ &= \sum_{\substack{k \in \mathbb{N} \\ \lambda_k \neq \mu}} \frac{1}{\mu - \lambda_k} \langle y, u_k \rangle + (\text{elem. di } \text{span}\{u_k : k \in \mathbb{N}, \lambda_k = \mu\}) + \\ &+ \frac{1}{\mu} y - \frac{1}{\mu} \sum_{k \in \mathbb{N}} \langle y, u_k \rangle u_k = \\ &= \frac{1}{\mu} y + \frac{1}{\mu} \sum_{\substack{k \in \mathbb{N} \\ \lambda_k \neq \mu}} \frac{\lambda_k}{\mu - \lambda_k} \langle y, u_k \rangle + (\text{elem. di } \text{span}\{u_k : k \in \mathbb{N}, \lambda_k = \mu\}). \end{aligned}$$

Capitolo 6

Il problema di Sturm-Liouville

6.1 Operatore nucleo integrale

Definizione 6.1.1. Siano dati un intervallo compatto $[a, b]$ e una funzione $K: [a, b] \times [a, b] \rightarrow \mathbb{C}$, detta *nucleo integrale*. L'operatore A_K che manda una funzione $f: [a, b] \rightarrow \mathbb{C}$ in

$$(A_K f)(x) := \int_a^b K(x, y) f(y) dy$$

(quando ha senso) è detto l'*operatore nucleo integrale* associato al nucleo K .

Esempio 6.1.2. L'operatore integrale che associa ad f la funzione integrale $F(x) := \int_a^x f(y) dy$ (come nell'esempio 3.1.14 di pag. 89) si può scrivere come operatore nucleo integrale scegliendo come K la funzione caratteristica del triangolo $T = \{(x, y) : a \leq y \leq x \leq b\}$:

$$\begin{aligned} (A_K f)(x) &= \int_a^b K(x, y) f(y) dy = \int_a^b \chi_T(x, y) f(y) dy = \\ &= \int_a^x f(y) dy = F(x). \end{aligned}$$

Vedere la figura 6.1. Si constata che hanno interesse pratico anche nuclei integrali che non sono continui sul quadrato.

Si possono fare diverse ipotesi su K, f di modo che A_K lavori fra spazi noti. La mera misurabilità della funzione $y \mapsto K(x, y)f(y)$ il cui integrale

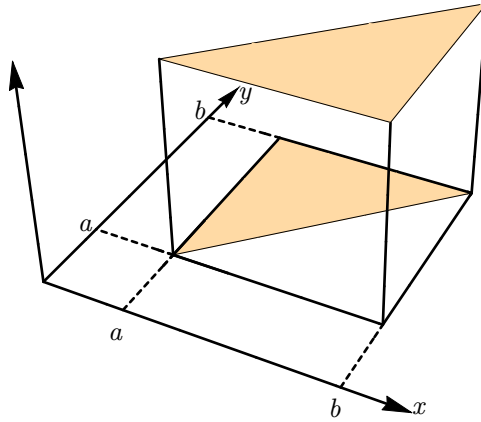


Figura 6.1: Grafico tridimensionale del nucleo integrale che riproduce l'operatore che manda $f: [a, b] \rightarrow \mathbb{R}$ in $F(x) := \int_a^x f(y) dy$, dell'esempio 6.1.2. Notare che il nucleo è discontinuo lungo la diagonale

definisce $A_K f(x)$ è garantita minimalmente se K è boreliana su $[a, b] \times [a, b]$ e f è misurabile su $[a, b]$. (I *boreliani* di uno spazio topologico sono gli elementi della più piccola σ -algebra che contiene gli aperti; per motivi di cardinalità, ci sono misurabili secondo Lebesgue che non sono boreliani). Qualora K fosse solo misurabile secondo Lebesgue, potrebbero esistere degli x (in un insieme trascurabile) per i quali la “sezione” $y \mapsto K(x, y)$ non risulti misurabile. Questo garantirebbe la misurabilità di $y \mapsto K(x, y)f(y)$ solo per quasi ogni x , che per molti scopi può anche bastare. Comunque ogni funzione misurabile secondo Lebesgue è uguale quasi ovunque a una funzione boreliana. Nella teoria spesso K è un elemento di qualche spazio L^p , e quindi è una classe di equivalenza di funzioni, all'interno della quale ci sono sicuramente funzioni boreliane. Dalla teoria della misura prodotto sappiamo che se $K_1 \equiv K_2$ quasi ovunque su $[a, b] \times [a, b]$, allora per quasi ogni x le sezioni sono misurabili e si ha che $K_1(x, y) \equiv K_2(x, y)$ per quasi ogni y . Insomma, la misurabilità delle sezioni non è un problema se partiamo da una K definita quasi ovunque su $[a, b] \times [a, b]$ e misurabile secondo Lebesgue, e il nostro scopo è definire $A_K f(x)$ per quasi ogni x .

Proposizione 6.1.3. *Se $K \in L^2([a, b] \times [a, b])$ allora A_K è un operatore lineare compatto da $L^2([a, b])$ in sé, e $\|A_K\| \leq \|K\|_{L^2([a, b])}$. Per l'aggiunto hilbertiano si ha che*

$$A_K^* = A_{K^*}, \quad \text{dove } K^*(x, y) := \overline{K(y, x)}. \quad (6.1)$$

Infine, A_K è autoaggiunto se e solo se $K = K^*$ quasi ovunque.

Dimostrazione. Siano $H = L^2([a, b])$, $G = L^2([a, b] \times [a, b])$. Notiamo che se $K \in G$ e $f \in H$, l'applicazione $(x, y) \mapsto K(x, y)f(y)$ è in L^1 del quadrato. Infatti, grazie alla disuguaglianza di Hölder

$$\begin{aligned} \int_a^b dx \int_a^b |K(x, y)f(y)| dy &\leq \\ &\leq \left(\int_a^b dx \int_a^b |K(x, y)|^2 dy \right)^{1/2} \left(\int_a^b dx \int_a^b |f(y)|^2 dy \right)^{1/2} = \\ &= \|K\|_2^2 \cdot \|f\|_2^2 \sqrt{b-a}. \end{aligned}$$

Per il teorema di Fubini-Tonelli l'integrale che definisce $(A_K f)(x)$ ha senso per quasi ogni x , e la funzione $x \mapsto (A_K f)(x)$ è misurabile e in $L^1([a, b])$.

Stimiamo la norma in L^2 di $A_K f$, applicando in modo diverso la disuguaglianza di Hölder:

$$\begin{aligned} \|A_K f\|_2^2 &= \int_a^b |(A_K f)(x)|^2 dx = \int_a^b \left| \int_a^b K(x, y)f(y) dy \right|^2 dx \leq \\ &\leq \int_a^b \left(\int_a^b |K(x, y)f(y)| dy \right)^2 dx \leq \\ &\leq \int_a^b \left(\left(\int_a^b |K(x, y)|^2 dy \right)^{1/2} \|f\|_2 \right)^2 dx = \\ &= \|f\|_2^2 \int_a^b dx \int_a^b |K(x, y)|^2 dy = \\ &= \|f\|_2^2 \cdot \|K\|_2^2 \end{aligned}$$

Quindi se $K \in G$ la A_K è un operatore lineare continuo da H in sé e $\|A_K\| \leq \|K\|_2$. L'applicazione di "nucleazione" $N: K \mapsto A_K$ è chiaramente a sua volta lineare e continua da G in $\mathcal{L}(H, H)$ e $\|N\| \leq 1$.

Calcoliamo esplicitamente $A_K f$ quando K è la funzione caratteristica del rettangolo $[x_1, x_2] \times [y_1, y_2]$:

$$\begin{aligned} A_K f(x) &= \int_a^b K(x, y)f(y) dy = \int_a^b \chi_{[x_1, x_2]}(x)\chi_{[y_1, y_2]}(y)f(y) dy = \\ &= \chi_{[x_1, x_2]}(x) \int_{y_1}^{y_2} f(y) dy. \end{aligned}$$

Vediamo che $A_K f$ è sempre un multiplo scalare della $\chi_{[x_1, x_2]}$, e quindi l'immagine di $A_K f$ ha dimensione 1.

Se K è una funzione a gradino in due variabili, cioè è una combinazione lineare finita di funzioni caratteristiche di rettangoli, allora, per la linearità della nucleazione N , A_K ha immagine di dimensione finita, e quindi è compatta.

È noto che l'insieme delle funzioni a gradino in due variabili è denso in $L^2([a, b] \times [a, b])$. Se $K \in L^2$ esiste una successione K_n di funzioni a gradino tale che $K_n \rightarrow K$ in L^2 . Ma allora $A_{K_n} \rightarrow A_K$ nella norma operatoriale, e di conseguenza A_K è compatta.

Calcoliamo l'aggiunta di A_K usando il teorema di Fubini-Tonelli e la definizione (6.1) di K^* :

$$\begin{aligned}
 \langle A_K f, g \rangle &= \int_a^b (A_K f)(x) \overline{g(x)} dx = \\
 &= \int_a^b dx \overline{g(x)} \int_a^b K(x, y) f(y) dy = \\
 &= \int_a^b dx \int_a^b K(x, y) \overline{g(x)} f(y) dy = \\
 &= \int_a^b dy \int_a^b K(x, y) \overline{g(x)} f(y) dx = \\
 &= \int_a^b dy f(y) \int_a^b K(x, y) \overline{g(x)} dx = \\
 &= \int_a^b dy f(y) \overline{\int_a^b K(x, y) g(x) dx} = \\
 &= \int_a^b dy f(y) \int_a^b K^*(y, x) g(x) dx = \\
 &= \int_a^b f(y) \overline{(A_{K^*} g)(y)} dy = \\
 &= \langle f, A_{K^*} g \rangle.
 \end{aligned}$$

Quindi $A_K^* = A_{K^*}$.

È chiaro che se $K = K^*$ allora A_K è autoaggiunta. Per il viceversa, poiché $A_K = A_K^* = A_{K^*}$ implica $0 = A_K - A_{K^*} = A_{K-K^*}$, basta far vedere che l'applicazione nucleazione $N: K \mapsto A_K$ è iniettiva. Supponiamo che

$A_K f = 0$ per ogni f . Allora $\langle A_K f, g \rangle_H = 0$ per ogni $f, g \in H$, cioè, di nuovo per Fubini-Tonelli,

$$0 = \langle A_K f, g \rangle_H = \int_a^b dx \int_a^b K(x, y) f(y) \overline{g(x)} dy = \langle K, g \otimes \bar{f} \rangle_G,$$

dove $(u \otimes v)(x, y) := u(x)v(y)$. Quindi K è ortogonale in G a tutte le funzioni della forma $\bar{g} \otimes f$, che comprendono in particolare le funzioni caratteristiche dei rettangoli. Essendo K ortogonale a un sottospazio denso in G , deduciamo che $K = 0$ quasi ovunque. \square

Il nucleo dell'esempio 6.1.2 di p. 137 è in L^p del quadrato per ogni p , ma non è continuo, e non ha la simmetria $K = K^*$. Si può verificare anche direttamente che l'operatore integrale associato non è autoaggiunto.

Un nucleo integrale K a valori reali dà un operatore A_K autoaggiunto se e solo se K è simmetrico rispetto alla diagonale $x = y$, cioè se e solo se $K(x, y) = K(y, x)$ per ogni x, y .

Esercizio 6.1.4. Dimostrare che se $p \geq 2$ e q è l'esponente coniugato, vale la disuguaglianza

$$\|A_K f\|_p^p \leq (b-a)^{\frac{p}{q}-1} \|f\|_p^p \cdot \|K\|_p^p.$$

Dedurre che se $K \in L^p([a, b] \times [a, b])$ allora A_K è un'applicazione lineare compatta da $L^p([a, b])$ in sé. (Usare la disuguaglianza di Jensen $\varphi(\int_X f d\mu) \leq \int_X \varphi \circ f d\mu$, valida quando $\mu(X) = 1$ e $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ è convessa).

Proposizione 6.1.5. *Se $K \in \mathcal{C}([a, b] \times [a, b])$ allora A_K è un'applicazione lineare compatta da $\mathcal{C}([a, b])$ in sé, ed è anche lineare compatta da $L^2([a, b])$ in $\mathcal{C}([a, b])$.*

Dimostrazione. Sia $\varepsilon > 0$. Poiché K è continua sul compatto $[a, b] \times [a, b]$, è uniformemente continua, per cui esiste $\delta > 0$ tale che

$$|x_1 - x_2| < \delta \wedge |y_1 - y_2| < \delta \quad \Rightarrow \quad |K(x_1, y_1) - K(x_2, y_2)| \leq \varepsilon.$$

Supponiamo che $|x_1 - x_2| < \delta$ e che $f: [a, b] \rightarrow \mathbb{C}$ sia continua. Allora

$$\begin{aligned} |(A_K f)(x_1) - (A_K f)(x_2)| &= \left| \int_a^b K(x_1, y) dy - \int_a^b K(x_2, y) dy \right| = \\ &= \left| \int_a^b (K(x_1, y) - K(x_2, y)) f(y) dy \right| \leq \end{aligned}$$

$$\begin{aligned}
&\leq \int_a^b \overbrace{|K(x_1, y) - K(x_2, y)|}^{< \varepsilon} \cdot |f(y)| dy \leq \\
&\leq \varepsilon \int_a^b |f(y)| dy \leq \\
&\leq (b-a)\varepsilon \|f\|_\infty.
\end{aligned}$$

Ovviamente

$$\|(A_K f)\|_\infty \leq (b-a)\|K\|_\infty \cdot \|f\|_\infty.$$

Quindi A_K mappa ogni sottinsieme limitato di $\mathcal{C}([a, b])$ in un insieme di funzioni equilimitate ed equiuniformemente continue. Per il teorema di Ascoli-Arzelà la A_K è compatta.

La dimostrazione che A_K manda $L^2([a, b])$ in $\mathcal{C}([a, b])$ ed è compatta è analoga. Infatti, con ε, δ come sopra

$$\begin{aligned}
|(A_K f)(x_1) - (A_K f)(x_2)| &\leq \varepsilon \int_a^b |f(y)| dy \leq \\
&\leq \varepsilon \left(\int_a^b |f(y)|^2 dy \right)^{1/2} \left(\int_a^b 1^2 dy \right)^{1/2} = \\
&= \varepsilon \sqrt{b-a} \|f\|_2.
\end{aligned}$$

e

$$\|(A_K f)\|_\infty \leq \|K\|_\infty \cdot \int_a^b |f(y)| dy \leq \sqrt{b-a} \|K\|_\infty \cdot \|f\|_2,$$

per cui A_K manda limitati di L^2 in precompatti di \mathcal{C}^0 . □

Esercizio 6.1.6. Se $K \in \mathcal{C}([a, b] \times [a, b])$ allora A_K è un'applicazione lineare compatta da $L^1([a, b])$ in $\mathcal{C}([a, b])$.

6.2 L'operatore differenziale di Sturm-Liouville

Sia $[a, b]$ un intervallo compatto, $\psi \in \mathcal{C}^1([a, b])$, $\varphi \in \mathcal{C}^0([a, b])$, con $\psi(x) > 0$ per ogni $x \in [a, b]$, e ψ, φ entrambe a valori reali. Definiamo l'operatore differenziale

$$Lu = -(\psi u')' + \varphi u = -\psi' u' - \psi u'' + \varphi u \quad \text{per } u \in \mathcal{C}^2([a, b]).$$

Dell'equazione differenziale $Lu = f$, con $f \in \mathcal{C}^0([a, b])$, ci proponiamo di studiare dei *problemi ai limiti*, cioè cercare soluzioni che verifichino delle condizioni nei due istanti diversi $x = a$ e $x = b$. Poiché per ipotesi $\psi(x) \neq 0$ per ogni x , l'equazione differenziale si può riscrivere in forma normale

$$u'' = -\frac{\psi'}{\psi}u' + \frac{\varphi}{\psi}u - \frac{f}{\psi}. \quad (6.2)$$

Si tratta di un'equazione differenziale lineare non omogenea a coefficienti continui. La teoria generale garantisce che ogni *problema di Cauchy* per $Lu = f$ in cui si assegnano $u(x_0)$ e $u'(x_0)$ ha soluzione unica in grande $u \in \mathcal{C}^2([a, b])$. Quando i dati iniziali sono reali, anche tutta la soluzione ha valori reali.

Proposizione 6.2.1. *Se $u, v \in \mathcal{C}^2([a, b])$ e allora*

$$\begin{aligned} \langle Lu, v \rangle &= \langle \psi u', v' \rangle + \langle \varphi u, v \rangle + \\ &+ \psi(a)u'(a)\overline{v(a)} - \psi(b)u'(b)\overline{v(b)}. \end{aligned}$$

dove i prodotti scalari sono nel senso di $L^2([a, b])$. Inoltre $\langle Lu, v \rangle$ e $\langle u, Lv \rangle$ differiscono solo per i valori delle funzioni agli estremi a, b :

$$\langle u, Lv \rangle - \langle Lu, v \rangle = \psi(b) \det \begin{pmatrix} u(b) & \overline{v(b)} \\ u'(b) & \overline{v'(b)} \end{pmatrix} - \psi(a) \det \begin{pmatrix} u(a) & \overline{v(a)} \\ u'(a) & \overline{v'(a)} \end{pmatrix}.$$

Dimostrazione. Si tratta di una semplice integrazione per parti in cui si apprezza la forma particolare di L :

$$\begin{aligned} \langle Lu, v \rangle &= \int_a^b (Lu)(x)\overline{v(x)} dx = \\ &= \int_a^b \left(-(\psi(x)u'(x))' + \varphi(x)u(x) \right) \overline{v(x)} dx = \\ &= \int_a^b \left(-(\psi(x)u'(x))' \right) \overline{v(x)} dx + \int_a^b \varphi(x)u(x)\overline{v(x)} dx = \\ &= - \left[\psi(x)u'(x)\overline{v(x)} \right]_a^b + \int_a^b \psi(x)u'(x)\overline{v'(x)} dx + \langle \varphi u, v \rangle = \\ &= \langle \psi u', v' \rangle + \langle \varphi u, v \rangle + \\ &+ \psi(a)u'(a)\overline{v(a)} - \psi(b)u'(b)\overline{v(b)}. \end{aligned}$$

Analogamente

$$\begin{aligned}
\langle u, Lv \rangle &= \int_a^b v(x) \overline{(Lv)(x)} dx = \\
&= \int_a^b u(x) \left(-(\psi(x)v'(x))' + \varphi(x)v(x) \right) dt = \\
&= \int_a^b u(x) \overline{-(\psi(x)v'(x))'} dt + \int_a^b u(x) \overline{\varphi(x)v(x)} dx = \\
&= - \left[u(x) \overline{\psi(x)v'(x)} \right]_a^b + \int_a^b u'(x) \overline{\psi(x)u'(x)} dx + \langle u, \varphi v \rangle = \\
&= \langle u', \psi v' \rangle + \langle u, \varphi v \rangle + \\
&\quad + \psi(a)u(a) \overline{v'(a)} - \psi(b)u(b) \overline{v'(b)}.
\end{aligned}$$

Ricordando di nuovo che φ, ψ sono a valori reali, si vede che $\langle u', \psi v' \rangle = \langle \psi u', v' \rangle$ e $\langle u, \varphi v \rangle = \langle \varphi u, v \rangle$. Sottraendo membro a membro:

$$\begin{aligned}
\langle u, Lv \rangle - \langle u, Lv \rangle &= \psi(a)u'(a) \overline{v(a)} - \psi(b)u'(b) \overline{v(b)} - \\
&\quad - \psi(a)u(a) \overline{v'(a)} + \psi(b)u(b) \overline{v'(b)} = \\
&= \psi(a) (u'(a) \overline{v(a)} - u(a) \overline{v'(a)}) + \\
&\quad + \psi(b) (u(b) \overline{v'(b)} - u'(b) \overline{v(b)}) = \\
&= \psi(b) \det \begin{pmatrix} u(b) & \overline{v(b)} \\ u'(b) & \overline{v'(b)} \end{pmatrix} - \psi(a) \det \begin{pmatrix} u(a) & \overline{v(a)} \\ u'(a) & \overline{v'(a)} \end{pmatrix}.
\end{aligned}$$

□

Corollario 6.2.2. *Sia X il sottospazio di $L^2([a, b])$ formato dalle $u \in \mathcal{C}^2([a, b])$ per le quali valgono le condizioni al contorno $u(a) = u(b) = 0$, oppure $u'(a) = u'(b) = 0$. Allora l'operatore L è autoaggiunto su X e*

$$\langle Lu, v \rangle = \langle u, Lv \rangle = \langle \psi u', v' \rangle + \langle \varphi u, v \rangle.$$

Se aggiungiamo anche l'ipotesi che $\varphi \geq 0$, la forma quadratica $u \mapsto \langle Lu, u \rangle$ risulta semidefinita positiva su X .

Definizione 6.2.3. Chiamiamo *wronskiano* di due funzioni $u_1, u_2 \in \mathcal{C}^1([a, b])$ la funzione

$$W(x) = \det \begin{pmatrix} u_1(x) & u_2(x) \\ u_1'(x) & u_2'(x) \end{pmatrix} = u_1(x)u_2'(x) - u_2(x)u_1'(x).$$

Chiameremo anche matrice wronskiana quella di cui si prende il determinante.

Proposizione 6.2.4. *Se $Lu_1 = 0$ e $Lu_2 = 0$ e W è il wronskiano di u_1, u_2 , allora la funzione ψW è costante. Le due funzioni u_1, u_2 sono linearmente dipendenti se e solo se $W(x_0) = 0$ per almeno un x_0 .*

Dimostrazione. Osserviamo che da $Lu_1 = 0$ segue che $\psi u_1'' = -\psi' u_1' + \varphi u_1$, e analogamente per u_2 . Calcoliamo la derivata di ψW :

$$\begin{aligned} (\psi W)' &= \psi' W + \psi W' = \psi' W + \psi \cdot (u_1 u_2' - u_2 u_1')' = \\ &= \psi' W + \psi \cdot (u_1' u_2' + u_1 u_2'' - u_2' u_1' - u_2 u_1'') = \\ &= \psi' W + u_1 \psi u_2'' - u_2 \psi u_1'' = \\ &= \psi' W + u_1 \cdot (-\psi' u_2' + \varphi u_2) - u_2 \cdot (-\psi' u_1' + \varphi u_1) = \\ &= \psi' W - \psi' u_1 u_2' + \varphi u_1 u_2 + \psi' u_2 u_1' - \varphi u_1 u_2 = \\ &= \psi' \cdot (W - u_1 u_2' + u_2 u_1') = \psi' \cdot (u_1 u_2' - u_2 u_1' - u_1 u_2' + u_2 u_1') = \\ &= 0. \end{aligned}$$

Quindi ψW è costante. Se le due funzioni u_1, u_2 sono linearmente dipendenti è ovvio che W è identicamente nullo. Viceversa, supponiamo che $W(x_0) = 0$ per un certo x_0 . Questo vuol dire che le due colonne della matrice

$$\begin{pmatrix} u_1(x_0) & u_2(x_0) \\ u_1'(x_0) & u_2'(x_0) \end{pmatrix}$$

sono linearmente dipendenti, cioè una delle due è multipla dell'altra, per esempio $u_2(x_0) = \alpha u_1(x_0)$, $u_2'(x_0) = \alpha u_1'(x_0)$ per un qualche $\alpha \in \mathbb{C}$. Ma allora le funzioni u_2 e αu_1 sono due soluzioni dell'equazione differenziale lineare omogenea del secondo ordine $Lu = 0$ che hanno stesso valore e stessa derivata prima per $x = x_0$. Per il teorema dell'esistenza e unicità della soluzione del problema di Cauchy, abbiamo che $u_2(x) = \alpha u_1(x)$ per ogni x , ossia u_1 e u_2 sono linearmente dipendenti. \square

6.3 Variazione delle costanti

Sviluppiano una formula esplicita delle soluzioni dell'equazione non omogenea $Lu = f$ col metodo della *variazione delle costanti arbitrarie*. Supponiamo

di avere a disposizione $u_1, u_2 \in \mathcal{C}^2([a, b])$, linearmente indipendenti e entrambe soluzioni dell'equazione omogenea associata $Lu = 0$. Dalla teoria generale sappiamo che $\ker L$ ha dimensione 2, e quindi ogni altra soluzione di $Lu = 0$ è combinazione lineare di u_1, u_2 , e che due soluzioni dell'equazione non omogenea $Lu = f$ differiscono per una combinazione lineare $c_1 u_1 + c_2 u_2$, con c_1, c_2 costanti. L'idea della "variazione delle costanti" è di cercare una soluzione di $Lu = f$ della forma

$$u(x) = c_1(x)u_1(x) + c_2(x)u_2(x),$$

nella quale le costanti sono diventate variabili. Avendo due funzioni c_1, c_2 con cui giocare, ci riserviamo di imporre un'ulteriore condizione oltre all'equazione $Lu = f$.

Cominciamo: calcoliamo la derivata prima di $u(x)$:

$$u'(x) = \underbrace{c_1'(x)u_1(x) + c_2'(x)u_2(x)}_{\equiv 0} + c_1(x)u_1'(x) + c_2(x)u_2'(x).$$

Il trucco cruciale del metodo è di *imporre* che $c_1' u_1 + c_2' u_2$ sia identicamente nullo:

$$c_1'(x)u_1(x) + c_2'(x)u_2(x) \equiv 0. \quad (6.3)$$

Tenendo conto di questa condizione calcoliamo la derivata seconda di u :

$$u''(x) = c_1(x)u_1''(x) + c_2(x)u_2''(x) + c_1'(x)u_1'(x) + c_2'(x)u_2'(x).$$

Combiniamo u, u', u'' in modo da ottenere Lu , e poi raccogliamo c_1, c_2 :

$$\begin{aligned} Lu &= -\psi'u' - \psi u'' + \varphi u = \\ &= -\psi' \cdot (c_1 u_1' + c_2 u_2') - \psi \cdot (c_1 u_1'' + c_2 u_2'' + c_1' u_1' + c_2' u_2') + \\ &\quad + \varphi \cdot (c_1 u_1 + c_2 u_2) = \\ &= c_1 \cdot (-\psi' u_1' - \psi u_1'' + \varphi u_1) + c_2 \cdot (-\psi' u_2' - \psi u_2'' + \varphi u_2) - \\ &\quad - \psi \cdot (c_1' u_1' + c_2' u_2') = \\ &= c_1 Lu_1 + c_2 Lu_2 - \psi \cdot (c_1' u_1' + c_2' u_2') = \\ &= -\psi \cdot (c_1' u_1' + c_2' u_2'). \end{aligned}$$

Imporre l'equazione non omogenea $Lu = f$ equivale a

$$-\psi \cdot (c_1' u_1' + c_2' u_2') = f.$$

Mettiamo insieme le due condizioni:

$$\begin{cases} c'_1 u_1 + c'_2 u_2 = 0, \\ c'_1 u'_1 + c'_2 u'_2 = -f/\psi. \end{cases}$$

Ricordiamo che qui le incognite sono c_1, c_2 e riscriviamo il sistema in forma matriciale:

$$\begin{pmatrix} u_1 & u_2 \\ u'_1 & u'_2 \end{pmatrix} \begin{pmatrix} c'_1 \\ c'_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -f/\psi \end{pmatrix}$$

Vediamo che la matrice dei coefficienti è proprio la wronskiana, che è invertibile perché per ipotesi u_1, u_2 sono linearmente indipendenti. Il determinante della matrice è il wronskiano W , che non è mai nullo. Risolviamo il sistema usando la regola di Cramer:

$$\begin{aligned} \begin{pmatrix} c'_1 \\ c'_2 \end{pmatrix} &= \begin{pmatrix} u_1 & u_2 \\ u'_1 & u'_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ -f/\psi \end{pmatrix} = \frac{1}{W} \begin{pmatrix} u'_2 & -u_2 \\ -u'_1 & u_1 \end{pmatrix} \begin{pmatrix} 0 \\ -f/\psi \end{pmatrix} = \\ &= \frac{1}{W} \begin{pmatrix} u_2 f/\psi \\ -u_1 f/\psi \end{pmatrix} = \frac{f}{\psi W} \begin{pmatrix} u_2 \\ -u_1 \end{pmatrix}, \end{aligned}$$

cioè, tornando in forma di sistema:

$$\begin{cases} c'_1 = \frac{u_2 f}{\psi W}, \\ c'_2 = -\frac{u_1 f}{\psi W}. \end{cases}$$

I secondi membri sono funzioni note. Quindi possiamo ricavare le “costanti variabili” c_1, c_2 tramite integrazione. Le costanti arbitrarie dell’integrazione corrispondono agli elementi di $\ker L$ che possiamo aggiungere a una soluzione della non omogenea $Lu = f$. Poiché siamo interessati ai problemi ai limiti, prendiamo non primitive qualsiasi di c'_1, c'_2 , ma quelle tali che

$$c_1(a) = 0, \quad c_2(b) = 0, \tag{6.4}$$

cioè

$$c_1(x) = \frac{1}{\psi(x)W(x)} \int_a^x u_2(y)f(y) dy, \tag{6.5}$$

$$c_2(x) = \frac{1}{\psi(x)W(x)} \int_x^b u_1(y)f(y) dy. \tag{6.6}$$

Notare che l'integrale di c_1 ha x come estremo superiore, quello di c_2 come estremo inferiore, e questo assorbe il segno “-” che appariva in $c_2' = -u_1 f / (\psi W)$. Abbiamo portato ψW fuori dall'integrale perché è costante. Notiamo anche che, usando le (6.4),

$$\begin{aligned} u(a) &= c_1(a)u_1(a) + c_2(a)u_2(a) = c_2(a)u_2(a) \\ u'(a) &= c_1(a)u_1'(a) + c_2(a)u_2'(a) = c_2(a)u_2'(a) \\ u(b) &= c_1(b)u_1(b) + c_2(b)u_2(b) = c_1(b)u_1(b) \\ u'(b) &= c_1(b)u_1'(b) + c_2(b)u_2'(b) = c_1(b)u_1'(b) \end{aligned}$$

cioè

$$\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} = c_2(a) \begin{pmatrix} u_2(a) \\ u_2'(a) \end{pmatrix}, \quad \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} = c_1(b) \begin{pmatrix} u_1(b) \\ u_1'(b) \end{pmatrix}. \quad (6.7)$$

Sostituiamo le formule per c_1, c_2 nella $u = c_1 u_1 + c_2 u_2$, cercando di unificare gli integrali introducendo funzioni caratteristiche:

$$\begin{aligned} u(x) &= c_1(x)u_1(x) + c_2(x)u_2(x) = \\ &= u_1(x) \frac{1}{\psi(x)W(x)} \int_a^x u_2(y)f(y) dy + u_2(x) \frac{1}{\psi(x)W(x)} \int_x^b u_1(y)f(y) dy = \\ &= \frac{1}{\psi(x)W(x)} \left(\int_a^x u_1(x)u_2(y)f(y) dy + \int_x^b u_1(y)u_2(x)f(y) dy \right) = \\ &= \frac{1}{\psi(x)W(x)} \left(\int_a^b \chi_{y \leq x}(x, y) u_1(x)u_2(y)f(y) dy + \right. \\ &\quad \left. + \int_a^b \chi_{x < y}(x, y) u_1(y)u_2(x)f(y) dy \right) = \\ &= \frac{1}{\psi(x)W(x)} \int_a^b \left(\chi_{y \leq x}(x, y) u_1(x)u_2(y) + \chi_{x < y}(x, y) u_1(y)u_2(x) \right) f(y) dy = \\ &= \int_a^b \underbrace{\frac{1}{\psi(y)W(y)} \left(\chi_{y \leq x}(x, y) u_1(x)u_2(y) + \chi_{x < y}(x, y) u_1(y)u_2(x) \right)}_{=: K(x, y)} f(y) dy = \\ &= \int_a^b K(x, y) f(y) dy = \\ &= (A_K f)(x). \end{aligned}$$

Abbiamo quindi espresso una soluzione dell'equazione non omogenea $Lu = f$ come un operatore nucleo integrale applicata al secondo membro f . Il nucleo

K si può scrivere in diversi modi:

$$\begin{aligned} W(x, y) &= \frac{1}{\psi W} \left(\chi_{y \leq x}(x, y) u_1(x) u_2(y) + \chi_{x < y}(x, y) u_1(y) u_2(x) \right) = \\ &= \frac{u_1(\max\{x, y\}) u_2(\min\{x, y\})}{\psi W} = \begin{cases} \frac{u_1(x) u_2(y)}{\psi W} & \text{se } y \leq x, \\ \frac{u_1(y) u_2(x)}{\psi W} & \text{se } y > x. \end{cases} \end{aligned}$$

Ci si può riferire alla figura 6.2 per capire la struttura di K . Sulla diagonale $y = x$ le due formule che definiscono K danno lo stesso valore, che è $u_1(x)u_2(x)/(\psi W)$. Quindi W è continuo su tutto il quadrato. È anche di classe \mathcal{C}^2 fuori dalla diagonale, ma in generale i punti della diagonale sono angolosi.

K è simmetrico rispetto alla diagonale: $K(x, y) = K(y, x)$. Se u_1, u_2 sono a valori reali, $K(x, t)$ è pure reale e simmetrica, e quindi A_K , se visto agire da $L^2([a, b])$ in sé, è compatto e autoaggiunto.

Un fatto che può lasciare perplessi è che le formule (6.5) e (6.6) che definiscono c_1, c_2 contengono primitive di $u_2 f, u_1 f$. Ora $u_1, u_2 \in \mathcal{C}^2$ però $f \in \mathcal{C}^0$, per cui mi aspetto che $c_1, c_2 \in \mathcal{C}^1$, e che la soluzione $u = c_1 u_1 + c_2 u_2$ sia soltanto di classe \mathcal{C}^1 , mentre abbiamo bisogno della derivata seconda nell'equazione. In effetti $c_1 u_1$ e $c_2 u_2$ potrebbero non avere derivata seconda, però la loro somma $c_1 u_1 + c_2 u_2$ ha derivata seconda continua! Il motivo è che la condizione (6.3) fa sì che le irregolarità separate dei due addendi $c_1 u_1$ e $c_2 u_2$ si cancellino fra di loro, e che la somma abbia derivata seconda.

Proposizione 6.3.1. *Le autofunzioni dell'operatore $A_K: L^2([a, b]) \rightarrow L^2([a, b])$ con autovalore λ non nullo sono di classe \mathcal{C}^2 e sono autofunzioni di L con autovalore $1/\lambda$:*

$$A_K u = \lambda u \quad \implies \quad Lu = \frac{1}{\lambda} u.$$

Dimostrazione. Supponiamo che f sia autofunzione di A_K con autovalore $\lambda \neq 0$:

$$A_K f = \lambda f, \quad f \in L^2([a, b]).$$

Posto $\bar{u} := A_K f$, per quanto visto nella proposizione 6.1.5 p. 141, \bar{u} è di classe \mathcal{C}^0 essendo nell'immagine di A_K con K continuo. Essendo poi $f = \frac{1}{\lambda} A_K f = \frac{1}{\lambda} \bar{u}$, in realtà anche $f \in \mathcal{C}^0([a, b])$. Ma allora $\bar{u} := A_K f$ è soluzione

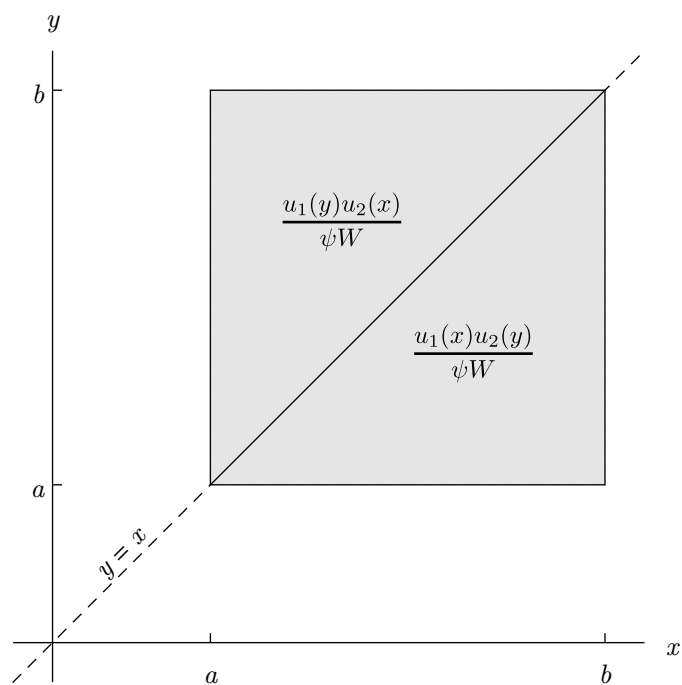


Figura 6.2: Il dominio dell'operatore integrale che risolve l'equazione non omogenea $Lu = f$. Il nucleo ha due formule diverse sopra e sotto la diagonale

di classe \mathcal{C}^2 di $L\bar{u} = f$. Deduciamo in particolare che $f = \frac{1}{\lambda}\bar{u}$ è pure di classe \mathcal{C}^2 . Inoltre da $L\bar{u} = f$ segue che $L(\lambda f) = f$, cioè

$$Lf = \frac{1}{\lambda}f.$$

Insomma, le autofunzioni di A_K fuori dal nucleo sono di classe \mathcal{C}^2 e sono anche autofunzioni di L con autovalore reciproco. \square

Esempio 6.3.2. Prendiamo $[a, b] = [0, 1]$, $\psi \equiv 1$, $\varphi \equiv 0$. L'equazione differenziale diventa $-u'' = f$. Due soluzioni indipendenti dell'omogenea associata sono $u_1 \equiv 1$, $u_2(x) \equiv x$, con wronskiano $W \equiv 1$. Il nucleo viene

$$K(x, y) = \min\{y, x\} = \begin{cases} y & \text{se } y \leq x, \\ x & \text{se } y > x. \end{cases}$$

che non è differenziabile lungo la diagonale (fig. 6.3). La soluzione generale dell'equazione non omogenea $-u'' = f$ su $[0, 1]$ si può scrivere in questa forma:

$$\begin{aligned} u(x) &= k_1 u_1(x) + k_2 u_2(x) + (A_K f)(x) = \\ &= k_1 + k_2 x + \int_0^1 f(y) \min\{y, x\} dy. \end{aligned}$$

Questa formula col nucleo integrale dà una soluzione di $-u'' = f$ anche fuori da $[0, 1]$?

Esercizio 6.3.3. Trovarne gli autovalori e autovettori dell'operatore differenziale $Lu = -u''$ su $\mathcal{C}^2([0, 1])$.

6.4 Problemi ai limiti omogenei

Riprendiamo l'operatore differenziale Lu delle sezioni precedenti

$$Lu = -(\psi u')' + \varphi u = -\psi' u' - \psi u'' + \varphi u$$

e cerchiamo di risolvere problemi ai limiti della forma

$$\begin{cases} Lu = f \\ B_a u = 0 \\ B_b u = 0 \end{cases}$$

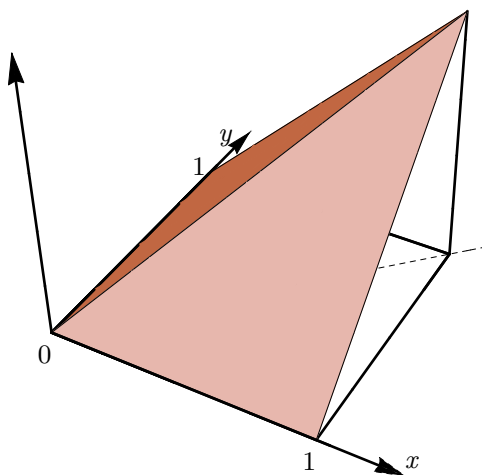


Figura 6.3: Il nucleo integrale che risolve l'equazione $-u'' = f$, come nell'esempio 3.1.14. Il nucleo è ovunque continuo, ma non è differenziabile sulla diagonale

dove $B_a, B_b: \mathcal{C}^2([a, b]) \rightarrow \mathbb{C}$ sono operatori lineari ai limiti

$$B_a u = \alpha_1 u(a) + \beta_1 u'(a), \quad B_b u = \alpha_2 u(b) + \beta_2 u'(b),$$

e per le costanti $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ valgono le ipotesi di nondegenerazione

$$(\alpha_1, \beta_1) \neq (0, 0), \quad (\alpha_2, \beta_2) \neq (0, 0).$$

Useremo anche le restrizioni di B_a, B_b al solo $\ker L$, che indicheremo con $\mathcal{B}_a, \mathcal{B}_b$, e le combinazioni $Bu = (B_a u, B_b u) \in \mathbb{C}^2$ e $\mathcal{B}u = (\mathcal{B}_a, \mathcal{B}_b) \in \mathbb{C}^2$. Può aiutare il diagramma commutativo seguente:

$$\begin{array}{ccc}
 \ker L \subset \mathcal{C}^2([a, b]) & & \\
 \searrow \mathcal{B} & & \downarrow B \\
 & & \mathbb{C}^2
 \end{array} \tag{6.8}$$

Esempi particolarmente semplici:

$$\begin{cases} Lu = 0 \\ u(a) = 0 \\ u(b) = 0 \end{cases} \quad \begin{cases} Lu = 0 \\ u'(a) = 0 \\ u'(b) = 0 \end{cases} \quad \begin{cases} Lu = 0 \\ u(a) = 0 \\ u'(b) = 0. \end{cases}$$

6.4.1 Il caso non degenere

Proposizione 6.4.1. *Le tre condizioni seguenti si equivalgono:*

1. *il problema ai limiti*

$$\begin{cases} Lu = f \\ B_a u = 0 \\ B_b u = 0 \end{cases} \quad (6.9)$$

ha soluzione unica per ogni $f \in C^0([a, b])$;

2. *il problema omogeneo*

$$\begin{cases} Lu = 0 \\ B_a u = 0 \\ B_b u = 0 \end{cases} \quad (6.10)$$

ha come unica soluzione quella identicamente nulla; altrimenti detto, per ogni soluzione di $Lu = 0$ abbiamo che $B_a u = 0$ e $B_b u = 0$ succedono contemporaneamente se e solo se $u = 0$;

3. *esistono u_1, u_2 linearmente indipendenti e soluzioni dei due problemi omogenei*

$$\begin{cases} Lu_1 = 0 \\ B_b u_1 = 0 \end{cases} \quad \begin{cases} Lu_2 = 0 \\ B_a u_2 = 0 \end{cases} \quad (6.11)$$

tali che, se combinate nel nucleo integrale K , fanno in modo che, per ogni f continua, $u = A_K f$ sia una soluzione del problema non omogeneo (6.9).

Dimostrazione. Se il problema (6.9) ha soluzione unica per ogni f , allora in particolare lo ha quando $f \equiv 0$, cioè (6.10) ha la sola soluzione nulla. In altre parole, la condizione 1 implica la 2.

Il viceversa richiede più lavoro. L'insieme delle soluzioni dell'equazione differenziale $Lu = f$ è un sottospazio affine di $C^2([a, b])$, contenente $A_K f$,

e due suoi elementi differiscono per un elemento di $\ker L$. Questo $\ker L$ ha dimensione 2. Il $\ker \mathcal{B}_a$ ha dimensione 1, ed è formato dai multipli della soluzione del problema di Cauchy

$$\begin{cases} Lu = 0 \\ u(a) = -\beta_1 \\ u'(a) = \alpha_1, \end{cases} \quad (6.12)$$

che è non nulla, giacché i due coefficienti α_1, β_1 non sono entrambi nulli. Il sottospazio $\ker \mathcal{B}_b$ è formato dai multipli della soluzione, di nuovo non nulla, del problema di Cauchy

$$\begin{cases} Lu = 0 \\ u(b) = -\beta_2 \\ u'(b) = \alpha_2. \end{cases} \quad (6.13)$$

Siamo pronti per dimostrare che la condizione 2 implica la 1. Supponiamo che il problema omogeneo (6.10) abbia come unica soluzione quella identicamente nulla. Questo equivale a dire che $\mathcal{B}: \ker L \rightarrow \mathbb{C}^2$ è iniettiva, e quindi biiettiva. Prendiamo $f \in \mathcal{C}^0([a, b])$ e consideriamo $\bar{u} = A_K f$, che è una soluzione di $Lu = f$. Non ci aspettiamo in generale che $B_a \bar{u} = B_b \bar{u} = 0$, però possiamo aggiungere a \bar{u} un elemento di $\ker L$:

$$\tilde{u} := \bar{u} - \underbrace{\mathcal{B}^{-1} B \bar{u}}_{\in \ker L}$$

in modo tale che \tilde{u} sia una soluzione del problema non omogeneo (6.9):

$$\begin{aligned} B \tilde{u} &= B(\bar{u} - \mathcal{B}^{-1} B \bar{u}) = B \bar{u} - B \underbrace{\mathcal{B}^{-1} B \bar{u}}_{\in \ker L} = B \bar{u} - \mathcal{B} \mathcal{B}^{-1} B \bar{u} = \\ &= B \bar{u} - B \bar{u} = 0. \end{aligned}$$

Tale soluzione \tilde{u} è unica perché due soluzioni differiscono per una soluzione del problema omogeneo (6.10).

Supponiamo che valga la condizione 2, la quale afferma in pratica che $\ker \mathcal{B}_a \cap \ker \mathcal{B}_b = \{0\}$. Ci proponiamo di dedurre che vale anche la condizione 3. I due sottospazi $\ker \mathcal{B}_a$ e $\ker \mathcal{B}_b$ hanno dimensione 1. Se prendo due vettori non nulli $u_2 \in \ker \mathcal{B}_a$ e $u_1 \in \ker \mathcal{B}_b$, questi saranno linearmente indipendenti. Usiamo questi due per formare il nucleo integrale K . Sappiamo

già che $u = A_K f$ risolve $Lu = f$. Grazie alle relazioni (6.7) di p. 148,

$$\begin{aligned} B_a u &= (\alpha_1 \ \beta_1) \begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} = (\alpha_1 \ \beta_1) c_2(a) \begin{pmatrix} u_2(a) \\ u_2'(a) \end{pmatrix} = c_2(a) \overbrace{B_a u_2}^{=0} = 0, \\ B_b u &= (\alpha_2 \ \beta_2) \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} = (\alpha_2 \ \beta_2) c_1(b) \begin{pmatrix} u_1(a) \\ u_1'(b) \end{pmatrix} = c_1(b) \underbrace{B_b u_1}_{=0} = 0. \end{aligned}$$

Pertanto $u = A_K f$ è soluzione del problema non omogeneo (6.9).

Supponiamo infine che valga la condizione 3. Esistendo due vettori indipendenti $u_2 \in \ker \mathcal{B}_a$ e $u_1 \in \ker \mathcal{B}_b$, i due nuclei, che hanno dimensione 1, devono incontrarsi solo nell'origine (altrimenti coinciderebbero). Quindi vale la condizione 2. \square

Se vogliamo soluzioni u di classe \mathcal{C}^2 dobbiamo prendere $f \in \mathcal{C}^0$ e quindi considerare A_K come definito su \mathcal{C}^0 a valori in \mathcal{C}^2 . Però la teoria degli operatori autoaggiunti compatti richiede una ambientazione in spazi di Hilbert.

Proposizione 6.4.2. *Se valgono le condizioni della proposizione precedente 6.4.1 e K è di quel tipo particolare, allora*

$$A_K u = \lambda u \iff \begin{cases} Lu = \frac{1}{\lambda} u \\ B_a u = 0 \\ B_b u = 0, \end{cases}$$

cioè gli autovalori non nulli dell'operatore L ristretto a $\ker B_a \cap \ker B_b$ sono tutti e soli i reciproci degli autovalori non nulli di A_K , e le autofunzioni corrispondenti sono le stesse.

Dimostrazione. Se siamo nelle condizioni della proposizione 6.4.1, allora

$$A_K u = \lambda u \iff \begin{cases} L(\lambda u) = u \\ B_a(\lambda u) = 0 \\ B_b(\lambda u) = 0, \end{cases} \iff \begin{cases} Lu = \frac{1}{\lambda} u \\ B_a u = 0 \\ B_b u = 0. \end{cases}$$

\square

Proposizione 6.4.3. *Se valgono le condizioni equivalenti della proposizione 6.4.1 e inoltre $\alpha_1, \beta_1, \alpha_2, \beta_2$ sono reali, allora*

1. l'operatore A_K è compatto e autoaggiunto su $L^2([a, b])$;
2. l'operatore differenziale L è autoaggiunto sul sottospazio X di $L^2([a, b])$ formato dalle $u \in C^2([a, b])$ per le quali valgono le condizioni al contorno $B_a u = B_b u = 0$. Se aggiungiamo anche l'ipotesi che $\varphi \geq 0$, la forma quadratica $u \mapsto \langle Lu, u \rangle$ risulta semidefinita positiva su X , e quindi gli autovalori di L sono ≥ 0 .

La decomposizione spettrale di A_K si può trovare a partire dagli autovalori e autovettori di L , senza calcolare K .

Dimostrazione. Come osservato, possiamo prendere $u_2 \in \ker \mathcal{B}_a$, $u_1 \in \ker \mathcal{B}_b$, cioè, come nelle equazioni (6.12), (6.13),

$$\begin{cases} Lu_2 = 0 \\ u_2(a) = -\beta_1 \\ u_2'(a) = \alpha_1, \end{cases} \quad \begin{cases} Lu_1 = 0 \\ u_1(b) = -\beta_2 \\ u_1'(b) = \alpha_2. \end{cases}$$

Avendo dati iniziali reali, u_1, u_2 sono a valori reali. Quindi l'operatore A_K è autoaggiunto e compatto da $L^2([a, b])$ in sé, anche se in L^2 accettiamo anche funzioni a valori complessi.

Per trattare l'operatore L generalizziamo il ragionamento del corollario 6.2.2 p. 144. Poiché i coefficienti $\alpha_1, \beta_1, \alpha_2, \beta_2$ sono reali, le condizioni $B_a u = B_b u = 0$ e $B_a v = B_b v = 0$ si possono interpretare come

$$\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} \perp \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \quad \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} \perp \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}, \\ \begin{pmatrix} v(a) \\ v'(a) \end{pmatrix} \perp \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \quad \begin{pmatrix} v(b) \\ v'(b) \end{pmatrix} \perp \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix},$$

o, equivalentemente,

$$\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} \perp \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \quad \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} \perp \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}, \\ \begin{pmatrix} \overline{v(a)} \\ \overline{v'(a)} \end{pmatrix} \perp \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \quad \begin{pmatrix} \overline{v(b)} \\ \overline{v'(b)} \end{pmatrix} \perp \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}.$$

Ma allora, trattandosi di vettori in dimensione 2 ed essendo i vettori (α_1, β_1) e (α_2, β_2) non nulli, abbiamo che

$$\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} \parallel \begin{pmatrix} \overline{v(a)} \\ \overline{v'(a)} \end{pmatrix}, \quad \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} \parallel \begin{pmatrix} \overline{v(b)} \\ \overline{v'(b)} \end{pmatrix},$$

cioè

$$\det \begin{pmatrix} u(a) & \overline{v(a)} \\ u'(a) & \overline{v'(a)} \end{pmatrix} = 0, \quad \det \begin{pmatrix} u(b) & \overline{v(b)} \\ u'(b) & \overline{v'(b)} \end{pmatrix} = 0.$$

Possiamo applicare la proposizione 6.2.1 p. 143 ed ottenere

$$\langle Lu, v \rangle = \langle u, Lv \rangle = \langle \psi u', v' \rangle + \langle \varphi u, v \rangle.$$

Da questo risulta anche chiaro che $\langle Lu, u \rangle \geq 0$ qualora $\varphi \geq 0$. Fin dall'inizio $\psi > 0$. \square

Esempio 6.4.4. Consideriamo il problema ai limiti

$$\begin{cases} -u'' = f \\ u(0) = 0 \\ u(\pi) = 0. \end{cases}$$

Nella nostra impostazione $[a, b] = [0, \pi]$, $\psi \equiv 1$, $\varphi \equiv 0$, $Lu = -u''$, $B_a u = u(0)$, $B_b u = u(\pi)$. Cominciamo a vedere se il problema omogeneo

$$\begin{cases} -u'' = 0 \\ u(0) = 0 \\ u(\pi) = 0 \end{cases}$$

ha solo la soluzione nulla. L'equazione $-u'' = 0$ equivale a $u(x) = mx + q$. Se imponiamo $u(0) = u(\pi) = 0$ abbiamo

$$\begin{cases} m \cdot 0 + q = 0 \\ m\pi + q = 0 \end{cases} \Rightarrow m = q = 0.$$

Possiamo prendere come u_1, u_2 delle soluzioni non nulle di

$$\begin{cases} -u_1'' = 0 \\ u_1(\pi) = 0 \end{cases} \quad \begin{cases} -u_2'' = 0 \\ u_2(0) = 0 \end{cases}$$

cioè per esempio

$$u_1(x) \equiv \pi - x, \quad u_2(x) \equiv x.$$

Il wronskiano vale π . Il nucleo integrale è

$$K(x, t) = \frac{1}{\pi} (\pi - \max\{t, x\}) \min\{t, x\} = \frac{1}{\pi} \begin{cases} t(\pi - x) & \text{se } 0 \leq t \leq x \leq \pi, \\ x(\pi - t) & \text{se } 0 < x < t < \pi. \end{cases}$$

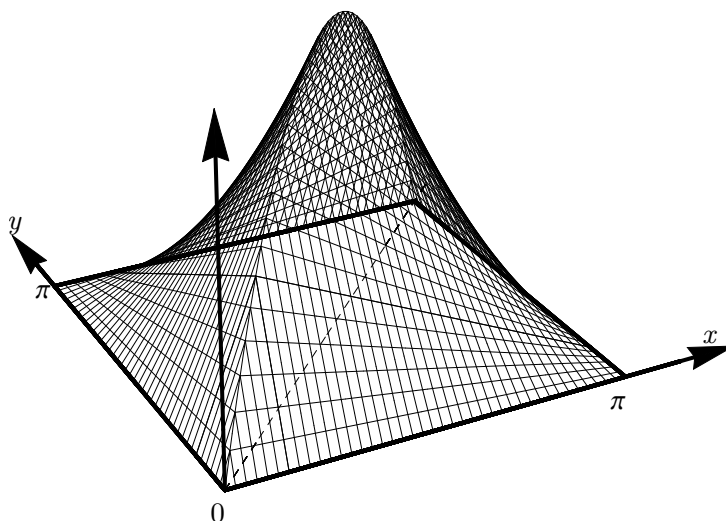


Figura 6.4: Grafico dell'operatore integrale che risolve l'equazione non omogenea $-u'' = f$ con condizioni al contorno $u(0) = u(\pi) = 0$. Si tratta di due pezzi di paraboloidi iperboliche. Confrontare con il nucleo della figura 6.3, che risolve $-u'' = f$ senza badare a condizioni al contorno

Il grafico di K è mostrato nella figura 6.4. Si può confrontare col nucleo dell'esempio 6.3.2 p. 151, dove l'equazione differenziale era la stessa, ma non si chiedevano condizioni al contorno.

Cerchiamo gli autovalori e le autofunzioni di L ristretto alle u tali che $u(0) = u(\pi) = 0$:

$$\begin{cases} -u'' = \lambda u \\ u(0) = 0 \\ u(\pi) = 0. \end{cases}$$

L'equazione lineare a coefficienti costanti $-u'' = \lambda u$ ha soluzioni della forma $u = e^{ax}$. Per trovare i coefficienti a sostituiamo nell'equazione:

$$-u'' = \lambda u \iff -a^2 e^{ax} = \lambda e^{ax} \iff a^2 = -\lambda.$$

Dalla teoria degli operatori autoaggiunti compatti sappiamo che gli autovalori λ sono necessariamente reali e di non nulli ne deve esistere almeno uno. Vediamo se possono essere negativi:

$$a^2 = -\lambda > 0 \iff a = \pm\sqrt{-\lambda}$$

$$u(x) = a_1 e^{x\sqrt{-\lambda}} + a_2 e^{-x\sqrt{-\lambda}}$$

Sostituiamo nella prima condizioni ai limiti:

$$0 = u(0) = a_1 e^{0\sqrt{-\lambda}} + a_2 e^{-0\sqrt{-\lambda}} = a_1 + a_2,$$

da cui $a_2 = -a_1$. Imponiamo la seconda condizione ai limiti:

$$0 = u(\pi) = a_1 e^{\pi\sqrt{-\lambda}} - a_1 e^{-\pi\sqrt{-\lambda}} = a_1 \underbrace{\left(\overbrace{e^{\pi\sqrt{-\lambda}}}_{>e^0=1} - \overbrace{e^{-\pi\sqrt{-\lambda}}}_{<e^0=1} \right)}_{>0},$$

da cui segue che $a_1 = a_2 = 0$. Quindi non ci sono autovalori $\lambda < 0$. Questo era prevedibile tenendo conto del fatto che $\varphi \geq 0$ e quindi $\langle Lu, u \rangle \geq 0$.

Cerchiamo autovalori $\lambda > 0$, cioè autofunzioni del tipo

$$a^2 = -\lambda < 0 \iff a = \pm i\sqrt{\lambda}$$

$$u(x) = a_1 \operatorname{sen}(x\sqrt{\lambda}) + a_2 \operatorname{cos}(x\sqrt{\lambda}).$$

Sostituendo nella condizione $u(0) = 0$ otteniamo che $a_2 = 0$. Sostituendo nella condizione $u(\pi) = 0$ otteniamo

$$0 = u(\pi) = a_1 \operatorname{sen}(\pi\sqrt{\lambda}).$$

Poiché vogliamo che $u \neq 0$, cioè $a_1 \neq 0$, la condizione diventa che l'argomento del seno sia un multiplo intero di π :

$$\exists n \in \mathbb{Z} \text{ tale che } \pi\sqrt{\lambda} = n\pi$$

cioè gli autovalori di L sono $\lambda_n = n^2$ al variare di $n \in \mathbb{N} \setminus \{0\}$, e le rispettive autofunzioni non normalizzate sono

$$v_n(x) = \operatorname{sen} nx.$$

Gli autospazi sono di dimensione 1. Per normalizzare calcoliamo

$$\begin{aligned} \|v_n\|_2^2 &= \int_0^\pi |v_n(x)|^2 dx = \int_0^\pi |\operatorname{sen} nx|^2 dx = \int_0^\pi \frac{1 - \cos 2nx}{2} dx = \\ &= \left[\frac{x}{2} - \frac{1}{4n} \operatorname{sen} 2nx \right]_0^\pi = \frac{\pi}{2} - \frac{1}{4n} \operatorname{sen} 2n\pi = \end{aligned}$$

$$= \frac{\pi}{2}.$$

Quindi gli autovettori normalizzati sono

$$u_n(x) = \sqrt{\frac{2}{\pi}} \operatorname{sen} nx \quad \text{per } n \in \mathbb{N}, n > 0.$$

La soluzione del problema ai limiti iniziale

$$\begin{cases} -u'' = f \\ u(0) = 0 \\ u(\pi) = 0 \end{cases}$$

si può scrivere in forma di serie di autofunzioni:

$$\begin{aligned} u(x) &= \sum_{n=1}^{+\infty} \frac{1}{\lambda_n} \langle f, u_n \rangle u_n(x) = \\ &= \sum_{n=1}^{+\infty} \frac{1}{n^2} \sqrt{\frac{2}{\pi}} (\operatorname{sen} nx) \int_0^\pi f(t) \sqrt{\frac{2}{\pi}} \operatorname{sen} nt \, dt = \\ &= \frac{2}{\pi} \sum_{n=1}^{+\infty} \frac{\operatorname{sen} nx}{n^2} \int_0^\pi f(t) \operatorname{sen} nt \, dt. \end{aligned}$$

La serie in linea di principio converge nel senso di L^2 , però si vede subito che quando f è continua converge totalmente, e quindi uniformemente.

6.4.2 Il caso degenere

Proposizione 6.4.5. *Le condizioni seguenti si equivalgono:*

1. *esistono infinite soluzioni del problema omogeneo*

$$\begin{cases} Lu = 0 \\ B_a u = 0 \\ B_b u = 0; \end{cases} \quad (6.14)$$

2. il problema non omogeneo

$$\begin{cases} Lu = f \\ B_b u = 0 \\ B_a u = 0 \end{cases} \quad (6.15)$$

ha infinite soluzioni se f è ortogonale in L^2 ad \bar{u} per ogni u che risolve il problema omogeneo (6.14), e non ha soluzioni altrimenti;

3. esistono basi $\{u_1, u_2\}$ di $\ker L$ nelle quali u_1 risolve il problema non omogeneo (6.14); combinata una tale base nel nucleo integrale K , il problema non omogeneo (6.15) è risolubile se e solo se $f \perp \bar{u}_1$, e in tal caso la soluzione generale è $u = \alpha u_1 + A_K f$, al variare di $\alpha \in \mathbb{C}$.

Se i coefficienti $\alpha_1, \beta_1, \alpha_2, \beta_2$ sono reali, le u_1, u_2 della condizione 3 si possono prendere a valori reali, A_K risulta autoaggiunta, e la restrizione di L a $\ker B$ è autoaggiunto.

Dimostrazione. Supponiamo che valga la condizione 1 e dimostriamo che vale la 3. Chiamiamo u_1 una soluzione non nulla del problema omogeneo (6.14). In particolare, $\ker \mathcal{B}_a$ e $\ker \mathcal{B}_b$, che sono sottospazi di dimensione 1 di $\ker L$, coincidono e sono formati dai multipli di u_1 . Prendiamo in $\ker L$ un u_2 linearmente indipendente da u_1 e costruiamo il nucleo integrale K . In particolare $\mathcal{B}_a u_2 \neq 0$. Posto $u = A_K f$, di nuovo grazie alle relazioni (6.7) di p. 148, possiamo scrivere

$$\begin{aligned} B_a u &= (\alpha_1 \ \beta_1) \begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} = (\alpha_1 \ \beta_1) c_2(a) \begin{pmatrix} u_2(a) \\ u_2'(a) \end{pmatrix} = c_2(a) \overbrace{B_a u_2}^{\neq 0}, \\ B_b u &= (\alpha_2 \ \beta_2) \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} = (\alpha_2 \ \beta_2) c_1(b) \begin{pmatrix} u_1(b) \\ u_1'(b) \end{pmatrix} = c_1(b) \underbrace{B_b u_1}_{=0} = 0. \end{aligned}$$

Dalla formula (6.6) p. 147 risulta che

$$c_2(a) = \frac{1}{\psi(a)W(a)} \int_a^b u_1(y) f(y) dy = \frac{1}{\psi(a)W(a)} \langle f, \bar{u}_1 \rangle.$$

Quindi se $f \perp \bar{u}_1$ abbiamo che

$$B_a u = \underbrace{c_2(a)}_{=0} \overbrace{B_a u_2}^{\neq 0} = 0$$

e $A_K f$ è una soluzione del problema non omogeneo, a cui possiamo aggiungere un qualsiasi multiplo di u_1 .

Viceversa, supponiamo che il problema non omogeneo (6.14) abbia una soluzione u . Le condizioni $B_a u = B_b u = 0$ e $B_a u_1 = B_b u_1 = 0$ si possono interpretare come

$$\begin{aligned} \begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} &\perp \begin{pmatrix} \overline{\alpha_1} \\ \overline{\beta_1} \end{pmatrix}, & \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} &\perp \begin{pmatrix} \overline{\alpha_2} \\ \overline{\beta_2} \end{pmatrix}, \\ \begin{pmatrix} u_1(a) \\ u'_1(a) \end{pmatrix} &\perp \begin{pmatrix} \overline{\alpha_1} \\ \overline{\beta_1} \end{pmatrix}, & \begin{pmatrix} u_1(b) \\ u'_1(b) \end{pmatrix} &\perp \begin{pmatrix} \overline{\alpha_2} \\ \overline{\beta_2} \end{pmatrix}, \end{aligned}$$

da cui segue che

$$\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} \parallel \begin{pmatrix} u_1(a) \\ u'_1(a) \end{pmatrix}, \quad \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} \parallel \begin{pmatrix} u_1(b) \\ u'_1(b) \end{pmatrix}.$$

Allora, usando la proposizione 6.2.1 p. 143, nel seguente calcolo i determinanti si annullano:

$$\begin{aligned} \langle f, \overline{u_1} \rangle &= \langle Lu, \overline{u_1} \rangle = \\ &= \langle u, L\overline{u_1} \rangle + \psi(b) \det \begin{pmatrix} u(b) & u_1(b) \\ u'(b) & u'_1(b) \end{pmatrix} - \psi(a) \det \begin{pmatrix} u(a) & u_1(a) \\ u'(a) & u'_1(a) \end{pmatrix} = \\ &= \langle u, \underbrace{L\overline{u_1}}_{=0} \rangle + 0 = 0. \end{aligned}$$

Quindi in effetti $f \perp \overline{u_1}$.

È ovvio che la condizione 3 implica la 2 e che questa a sua volta implica la 1.

Se infine i coefficienti $\alpha_1, \beta_1, \alpha_2, \beta_2$ sono reali, le u_1, u_2 della condizione 3 si possono prendere a valori reali perché si possono assegnare condizioni iniziali reali. Il nucleo K risultante è a valori reali e l'operatore A_K viene autoaggiunto. La restrizione di L a $\ker B$ è autoaggiunta con lo stesso ragionamento che nel caso non degenere (proposizione 6.4.3 p. 155). \square

Proposizione 6.4.6. *Se valgono le condizioni della proposizione precedente 6.4.5 e K è di quel tipo particolare, allora*

$$A_K u = \lambda u \iff \begin{cases} Lu = \frac{1}{\lambda} u \\ B_a u = 0 \\ B_b u = 0, \end{cases}$$

cioè gli autovalori non nulli dell'operatore L ristretto a $\ker B_a \cap \ker B_b$ sono tutti e soli i reciproci degli autovalori non nulli di A_K , e le autofunzioni corrispondenti sono le stesse.

Dimostrazione. Se siamo nelle condizioni della proposizione 6.4.5, allora

$$A_K u = \lambda u \quad \Longrightarrow \quad \begin{cases} L(\lambda u) = u \\ B_a(\lambda u) = 0 \\ B_b(\lambda u) = 0, \end{cases} \quad \Longleftrightarrow \quad \begin{cases} Lu = \frac{1}{\lambda}u \\ B_a u = 0 \\ B_b u = 0. \end{cases}$$

Viceversa, se $Lu = \frac{1}{\lambda}u$ e $Bu = (0, 0)$ necessariamente $\frac{1}{\lambda}u \perp u_1$, e quindi $\tilde{u} := A_K(\frac{1}{\lambda}u)$ risolve $L\tilde{u} = \frac{1}{\lambda}u$, $B\tilde{u} = (0, 0)$. \square

Capitolo 7

La trasformata di Fourier

Nel trattare la trasformata di Fourier useremo spesso non la misura di Lebesgue su \mathbb{R} , ma la misura di Lebesgue “normalizzata”, cioè divisa per $\sqrt{2\pi}$, perché questa porta a formule più semplici. Quando usata negli integrali, tale misura sarà indicata così:

$$\int_E f(t) \frac{dt}{\sqrt{2\pi}}, \quad \text{che significa semplicemente } \frac{1}{\sqrt{2\pi}} \int_E f(t) dt.$$

Come insiemi, gli spazi L^p rispetto alla misura normalizzata sono gli stessi che per la misura di Lebesgue, e le norme differiscono per una costante moltiplicativa.

7.1 Convoluzione

La convoluzione di due funzioni $f, g: \mathbb{R} \rightarrow \mathbb{C}$ sarà fatta rispetto alla misura normalizzata:

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)g(y) \frac{dy}{\sqrt{2\pi}}.$$

Proposizione 7.1.1 (Convoluzione di funzioni di L^1). *Siano $f, g \in L^1(\mathbb{R})$. Allora per quasi ogni $x \in \mathbb{R}$ si ha che*

$$\int_{\mathbb{R}} |f(x-y)g(y)| \frac{dy}{\sqrt{2\pi}} < +\infty, \quad (7.1)$$

e quindi per quasi ogni $x \in \mathbb{R}$ possiamo porre

$$(f * g)(x) := \int_{\mathbb{R}} f(x-y)g(y) \frac{dy}{\sqrt{2\pi}} \in \mathbb{C} \quad \text{per quasi ogni } x \in \mathbb{R}, \quad (7.2)$$

e si ha che

$$f * g \in L^1(\mathbb{R}), \quad f * g = g * f, \quad (7.3)$$

$$\int_{\mathbb{R}} (f * g)(x) \frac{dx}{\sqrt{2\pi}} = \left(\int_{\mathbb{R}} f(x) \frac{dx}{\sqrt{2\pi}} \right) \left(\int_{\mathbb{R}} g(x) \frac{dx}{\sqrt{2\pi}} \right). \quad (7.4)$$

Se $p, q \in [1, +\infty]$ sono esponenti coniugati e $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ allora la disuguaglianza (7.1) vale per ogni $x \in \mathbb{R}$, la convoluzione (7.2) ha senso per ogni $x \in \mathbb{R}$, $f * g \in L^\infty(\mathbb{R})$, e vale la commutatività $(f * g)(x) = (g * f)(x)$ per ogni x .

Dimostrazione. Consideriamo l'integrale doppio con valore assoluto, applichiamo il teorema di Fubini-Tonelli e facciamo il cambio di variabile $t = x - y$:

$$\begin{aligned} \int_{\mathbb{R} \times \mathbb{R}} |f(x-y)g(y)| \frac{dx}{\sqrt{2\pi}} \frac{dy}{\sqrt{2\pi}} &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |f(x-y)g(y)| \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(|g(y)| \int_{-\infty}^{+\infty} |f(x-y)| \frac{dx}{\sqrt{2\pi}} \right) = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(|g(y)| \int_{+\infty}^{-\infty} |f(t)| \frac{-dt}{\sqrt{2\pi}} \right) = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(|g(y)| \int_{-\infty}^{+\infty} |f(t)| \frac{dt}{\sqrt{2\pi}} \right) = \\ &= \left(\int_{-\infty}^{+\infty} |g(y)| \frac{dy}{\sqrt{2\pi}} \right) \left(\int_{-\infty}^{+\infty} |f(t)| \frac{dt}{\sqrt{2\pi}} \right) = \\ &= \|f\|_1 \cdot \|g\|_1 < +\infty. \end{aligned} \quad (7.5)$$

Ancora per il teorema di Fubini-Tonelli la sezione $y \mapsto f(x-y)g(y)$ è in $L^1(\mathbb{R})$ per quasi ogni x e nel conto precedente possiamo togliere i valori assoluti:

$$\begin{aligned} \int_{\mathbb{R}} (f * g)(x) \frac{dx}{\sqrt{2\pi}} &= \int_{\mathbb{R} \times \mathbb{R}} f(x-y)g(y) \frac{dx}{\sqrt{2\pi}} \frac{dy}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x-y)g(y) \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(g(y) \int_{-\infty}^{+\infty} f(x-y) \frac{dx}{\sqrt{2\pi}} \right) = \\ &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(g(y) \int_{+\infty}^{-\infty} f(t) \frac{-dt}{\sqrt{2\pi}} \right) = \end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \left(g(y) \int_{-\infty}^{+\infty} f(t) \frac{dt}{\sqrt{2\pi}} \right) = \\
&= \left(\int_{-\infty}^{+\infty} g(y) \frac{dy}{\sqrt{2\pi}} \right) \left(\int_{-\infty}^{+\infty} f(t) \frac{dt}{\sqrt{2\pi}} \right) = \\
&= \left(\int_{\mathbb{R}} f(x) \frac{dx}{\sqrt{2\pi}} \right) \left(\int_{\mathbb{R}} g(x) \frac{dx}{\sqrt{2\pi}} \right).
\end{aligned}$$

Di nuovo col cambio di variabile $t = x - y$

$$\begin{aligned}
(f * g)(x) &= \int_{-\infty}^{+\infty} f(x - y)g(y) \frac{dy}{\sqrt{2\pi}} = \int_{+\infty}^{-\infty} f(t)g(x - t) \frac{-dt}{\sqrt{2\pi}} = \\
&= \int_{-\infty}^{+\infty} g(x - t)f(t) \frac{dt}{\sqrt{2\pi}} = \\
&= (g * f)(x).
\end{aligned}$$

Se $p, q \in]1, +\infty[$ sono esponenti coniugati e $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ allora per la disuguaglianza di Hölder col cambio di variabile $x - y = t$, $dy = -dt$

$$\begin{aligned}
\int_{\mathbb{R}} |f(x - y)g(y)| \frac{dy}{\sqrt{2\pi}} &\leq \left(\int_{-\infty}^{+\infty} |f(x - y)|^p \frac{dy}{\sqrt{2\pi}} \right)^{1/p} \left(\int_{\mathbb{R}} |g(y)|^q \frac{dy}{\sqrt{2\pi}} \right)^{1/q} = \\
&= \left(\int_{+\infty}^{-\infty} |f(t)|^p \frac{-dt}{\sqrt{2\pi}} \right)^{1/p} \|g\|_q = \\
&= \left(\int_{-\infty}^{+\infty} |f(t)|^p \frac{dt}{\sqrt{2\pi}} \right)^{1/p} \|g\|_q = \\
&= \|f\|_p \cdot \|g\|_q < +\infty.
\end{aligned}$$

Quindi $(f * g)(x)$ ha senso per ogni x , e continua a valere la proprietà commutativa $(f * g)(x) = (g * f)(x)$.

Il caso $f \in L^1(\mathbb{R})$, $g \in L^\infty(\mathbb{R})$ è ancora più semplice. \square

Esercizio 7.1.2. Posto $f(x) = 1/\max\{1, |x|\}$, mostrare che $f \in L^2(\mathbb{R})$ ma $f * f \notin L^1(\mathbb{R})$. Qual è l'andamento all'infinito di $f * f$?

7.2 La trasformata di Fourier in L^1

Definizione 7.2.1. Data una funzione $f \in L^1(\mathbb{R})$ a valori complessi, chiameremo *trasformata di Fourier* di f la funzione $\hat{f}: \mathbb{R} \rightarrow \mathbb{C}$ definita da

$$\hat{f}(t) := \int_{\mathbb{R}} f(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}}. \quad (7.6)$$

Nelle applicazioni, la variabile indipendente x della funzione $f(x)$ ha spesso il significato di *tempo*, mentre la variabile t ha il significato di *frequenza angolare*, tant'è vero che da alcuni autori la definizione è scritta con ω al posto di t e t al posto di x :

$$\hat{f}(\omega) := \int_{\mathbb{R}} f(t) e^{-i\omega t} \frac{dt}{\sqrt{2\pi}},$$

una formula che ha il vantaggio di mettere in chiaro i diversi ruoli della variabile di f e di \hat{f} . Il valore assoluto $|\hat{f}(\omega)|$ ha il significato di *intensità* della componente di frequenza ω , e l'argomento di $\hat{f}(\omega)$ è la *fase*. Esistono tradizioni diverse anche per l'effettivo valore della trasformata di Fourier, che usano convenzioni diverse sul coefficiente all'esponentiale (in particolare e^{ixt} invece di e^{-ixt}) e sul fattore di normalizzazione. Quando si passa da un libro all'altro bisogna fare attenzione alla particolare convenzione usata. In questa trattazione useremo di solito la x come la variabile "temporale" della funzione base f da trasformare, mentre t sarà la variabile "frequenza angolare" della funzione trasformata.

Osserviamo che la definizione di $\hat{f}(t)$ è giustificata per ogni $t \in \mathbb{R}$ in quanto la funzione integranda è sommabile:

$$\begin{aligned} \int_{\mathbb{R}} |f(x) e^{-ixt}| \frac{dx}{\sqrt{2\pi}} &= \int_{\mathbb{R}} |f(x)| \cdot \overbrace{|e^{-ixt}|}^{=1} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} |f(x)| \frac{dx}{\sqrt{2\pi}} = \|f\|_1 < +\infty. \end{aligned}$$

Risulta in particolare che \hat{f} è limitata:

$$\begin{aligned} |\hat{f}(t)| &= \left| \int_{\mathbb{R}} f(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} \right| \leq \int_{\mathbb{R}} |f(x) e^{-ixt}| \frac{dx}{\sqrt{2\pi}} \leq \\ &\leq \|f\|_1 \quad \forall t \in \mathbb{R}. \end{aligned}$$

La funzione di due variabili $(x, t) \mapsto f(x)e^{-ixt}$ è misurabile secondo Lebesgue perché prodotto di una misurabile per una continua, e inoltre è in $L^1(\mathbb{R} \times [-A, A])$ per ogni $A > 0$. Per il teorema di Fubini-Tonelli, la \hat{f} , che dà l'integrale delle sezioni, risulta misurabile su ogni intervallo $[-A, A]$, e quindi è misurabile su \mathbb{R} (in realtà \hat{f} è continua, come vedremo nel teorema 7.3.3 p. 176). Due funzioni integrabili f_1, f_2 che differiscano in un insieme di misura nulla hanno evidentemente la stessa trasformata di Fourier. Concludiamo che l'applicazione $f \mapsto \hat{f}$ è lineare e continua da $L^1(\mathbb{R})$ a $L^\infty(\mathbb{R})$, e

$$\|\hat{f}\|_\infty \leq \|f\|_1. \quad (7.7)$$

Vedremo che è iniettiva, ma *non* è suriettiva su tutto L^∞ .

Una cosa da osservare è che la trasformata di Fourier, come operatore, *non è locale*, cioè cambiare la funzione f su un intervallo, per quanto piccolo, fa cambiare in generale *tutti* i valori di \hat{f} . In altre parole, i cambiamenti locali “si spalmano” su tutta la retta.

Esempio 7.2.2. Sia $f = \chi_{[0,1]}$ e calcoliamo la sua trasformata di Fourier:

$$\begin{aligned} \hat{f}(t) &= \int_{-\infty}^{+\infty} f(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \chi_{[0,1]}(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_0^1 e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-ixt}}{-it} \right]_{x=0}^{x=1} = \frac{1}{\sqrt{2\pi}} \left(\frac{e^{-it}}{-it} - \frac{e^{-i0t}}{-it} \right) = \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{-it} - 1}{-it} = \frac{\operatorname{sen} t + i(\cos t - 1)}{t\sqrt{2\pi}}. \end{aligned}$$

Per $t = 0$ bisogna fare un conto a parte, e risulta $\hat{f}(0) = 1/\sqrt{2\pi}$. Notare che questa \hat{f} ha valori complessi, è continua su tutto \mathbb{R} , anche in 0, e tende a 0 all'infinito.

Esercizio 7.2.3. Calcolare le trasformate di Fourier di $\chi_{[-1,1]}$, $x \mapsto e^x \chi_{[-1,1]}(x)$, $x \mapsto (1 - x^2) \chi_{[-1,1]}(x)$, $x \mapsto (\cos x) \chi_{[-\pi/2, \pi/2]}(x)$.

Il teorema seguente elenca delle semplici proprietà della trasformata di Fourier dovute all'invarianza per traslazione della misura di Lebesgue e dal fatto che l'applicazione $x \mapsto e^{-ixt}$ è un *carattere* del gruppo additivo \mathbb{R} , cioè un omomorfismo da $(\mathbb{R}, +)$ a (\mathbb{U}, \cdot) , dove $\mathbb{U} := \{z \in \mathbb{C} : |z| = 1\}$ è il cerchio unitario di \mathbb{C} , che è un gruppo rispetto alla moltiplicazione. Chiamerò “torsione” la moltiplicazione di una funzione $f(x)$ per il carattere e^{-ixt} , a causa dell'effetto che ha sul grafico (figura 7.3), al solo scopo di dare un nome descrittivo a delle formule.

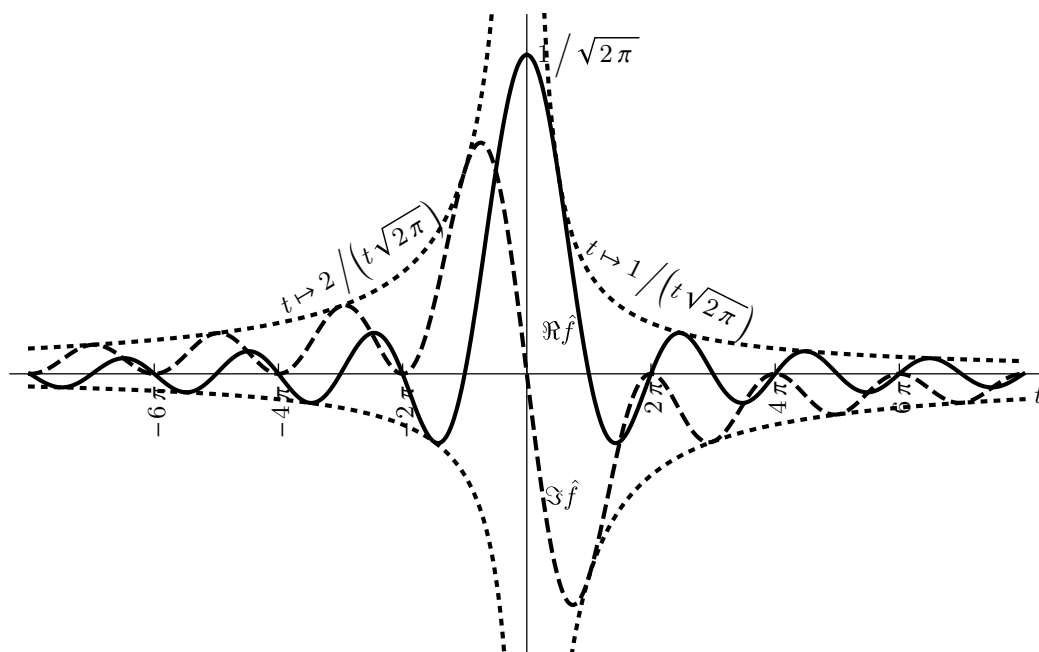


Figura 7.1: Grafico della trasformata di Fourier della funzione $f = \chi_{[0,1]}$, come nell'esempio 7.2.2: in linea continua è la parte reale $\Re \hat{f}$, tratteggiata è la parte immaginaria $\Im \hat{f}$. Coi puntini sono indicati i due inviluppi che danno l'andamento all'infinito, che è dell'ordine di $1/t$. Notare che \hat{f} è continua su tutto \mathbb{R}

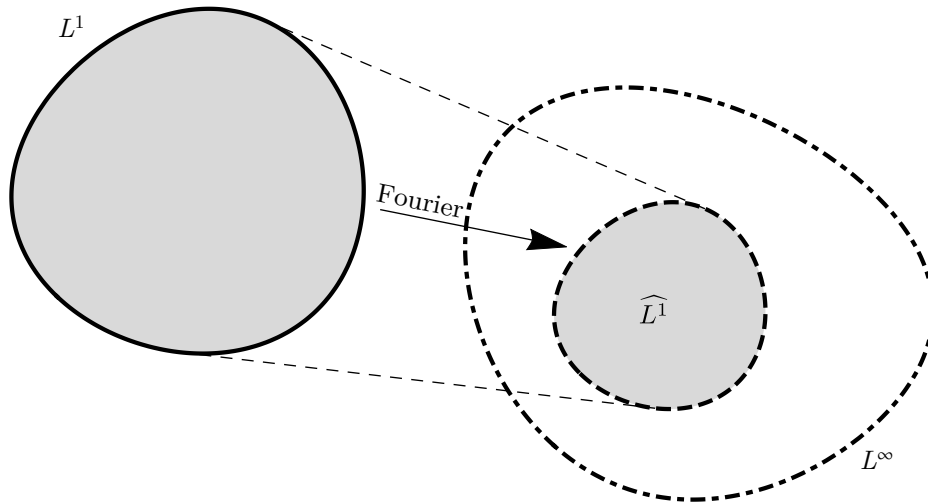


Figura 7.2: Schema di come la trasformata di Fourier si comporta rispetto agli insiemi $L^1(\mathbb{R})$ e $L^\infty(\mathbb{R})$

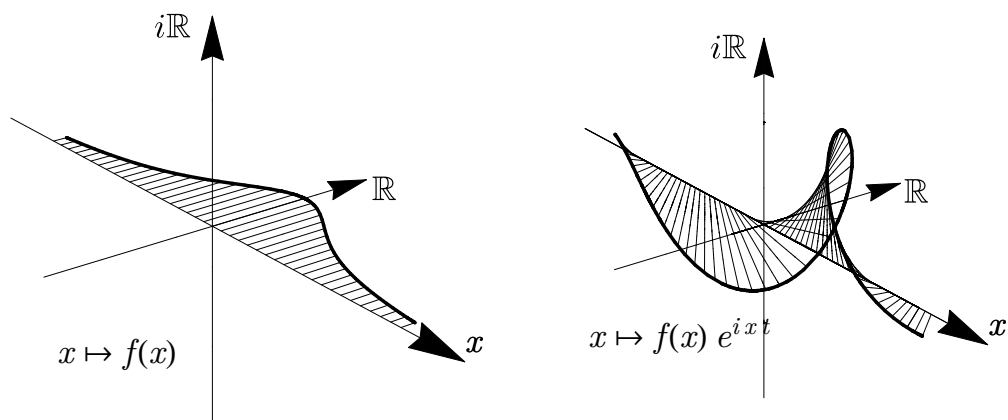


Figura 7.3: “Torsione” del grafico di una funzione $f: \mathbb{R} \rightarrow \mathbb{C}$, cioè moltiplicazione per e^{-ixt}

Teorema 7.2.4. *Supponiamo che $f \in L^1(\mathbb{R})$ e $\alpha, \lambda \in \mathbb{R}$. Allora:*

1. *posto $g(x) := f(x)e^{ix\alpha}$ abbiamo che $\hat{g}(t) = \hat{f}(t - \alpha)$ (“torsione” nello spazio diventa traslazione nelle frequenze);*
2. *posto $g(x) := f(x - \alpha)$ abbiamo che $\hat{g}(t) = \hat{f}(t)e^{-i\alpha t}$ (traslazione nello spazio diventa “torsione” nelle frequenze);*
3. *se $f, g \in L^1$ e poniamo $h = f * g$ abbiamo che $\hat{h}(t) = \hat{f}(t)\hat{g}(t)$ (la trasformata della convoluzione è il prodotto delle trasformate);*
4. *posto $g(x) := \overline{f(-x)}$ abbiamo che $\hat{g}(t) = \overline{\hat{f}(t)}$;*
5. *posto $g(x) := f(x/\lambda)$ con $\lambda > 0$ abbiamo che $\hat{g}(t) = \lambda \hat{f}(\lambda t)$ (omotetia orizzontale diventa omotetia reciproca sia orizzontale che verticale);*
6. *posto $g(x) := -ixf(x)$, supponiamo che $g \in L^1$. Allora \hat{f} è derivabile e $\hat{f}'(t) = \hat{g}(t)$ (la derivata della trasformata è la trasformata del prodotto con $-ix$).*

Dimostrazione. Posto $g(x) := f(x)e^{ix\alpha}$ calcoliamo la trasformata di g :

$$\begin{aligned}\hat{g}(t) &= \int_{\mathbb{R}} g(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} f(x)e^{ix\alpha} e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} f(x)e^{ix\alpha - ixt} \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} f(x)e^{-ix(t-\alpha)} \frac{dx}{\sqrt{2\pi}} = \\ &= \hat{f}(t - \alpha).\end{aligned}$$

Posto $g(x) := f(x - \alpha)$ calcoliamo la trasformata di g usando il cambio di variabile $x - \alpha = y$, $dx = dy$:

$$\begin{aligned}\hat{g}(t) &= \int_{\mathbb{R}} g(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} f(x - \alpha)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} f(y)e^{-i(y+\alpha)t} \frac{dy}{\sqrt{2\pi}} = \int_{\mathbb{R}} f(y)e^{-iyt} e^{-i\alpha t} \frac{dy}{\sqrt{2\pi}} = \\ &= e^{-i\alpha t} \int_{\mathbb{R}} f(y)e^{-iyt} \frac{dy}{\sqrt{2\pi}} = \\ &= e^{-i\alpha t} \hat{f}(t).\end{aligned}$$

Posto $h = f * g$ calcoliamo la trasformata di g usando il cambio di variabile $x - y = u$, $dx = du$

$$\begin{aligned}
 \hat{h}(t) &= \int_{\mathbb{R}} (f * g)(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \overbrace{f(x-y)g(y)}^{\in L^1(\mathbb{R}^2), \text{ vedi (7.5)}} \frac{dy}{\sqrt{2\pi}} \right) \overbrace{e^{-ixt}}^{\in L^\infty} \frac{dx}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dx}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x-y)g(y) e^{-ixt} \frac{dy}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x-y)g(y) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} g(y) \int_{-\infty}^{+\infty} f(x-y) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} g(y) \int_{-\infty}^{+\infty} f(u) e^{-i(u+y)t} \frac{du}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} g(y) \int_{-\infty}^{+\infty} f(u) e^{-iut} e^{-iyt} \frac{du}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} g(y) e^{-iyt} \int_{-\infty}^{+\infty} f(u) e^{-iut} \frac{du}{\sqrt{2\pi}} = \\
 &= \left(\int_{-\infty}^{+\infty} g(y) e^{-iyt} \frac{dy}{\sqrt{2\pi}} \right) \left(\int_{-\infty}^{+\infty} f(u) e^{-iut} \frac{du}{\sqrt{2\pi}} \right) = \\
 &= \hat{g}(t) \hat{f}(t).
 \end{aligned}$$

Posto $g(x) := \overline{f(-x)}$ calcoliamo la trasformata di g usando il cambio di variabile $y = -x$, $dy = -dx$

$$\begin{aligned}
 \hat{g}(t) &= \int_{\mathbb{R}} g(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \overline{f(-x)} e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\
 &= \int_{+\infty}^{-\infty} \overline{f(y)} e^{iyt} \frac{-dy}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \overline{f(y)} e^{iyt} \frac{dy}{\sqrt{2\pi}} = \\
 &= \int_{-\infty}^{+\infty} \overline{f(y) e^{-iyt}} \frac{dy}{\sqrt{2\pi}} = \overline{\int_{-\infty}^{+\infty} f(y) e^{-iyt} \frac{dy}{\sqrt{2\pi}}} = \\
 &= \overline{\hat{f}(t)}.
 \end{aligned}$$

Posto $g(x) := f(x/\lambda)$ con $\lambda > 0$ calcoliamo la trasformata di g , usando il cambio di variabile $y = x/\lambda$, $dy = dx/\lambda$

$$\begin{aligned}\hat{g}(t) &= \int_{\mathbb{R}} g(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} f(x/\lambda) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} f(y) e^{-iy\lambda t} \frac{\lambda dy}{\sqrt{2\pi}} = \lambda \int_{-\infty}^{+\infty} f(y) e^{-iy\lambda t} \frac{dy}{\sqrt{2\pi}} = \\ &= \lambda \hat{f}(\lambda t).\end{aligned}$$

Calcoliamo la derivata della trasformata:

$$\begin{aligned}\hat{f}'(t_0) &= \lim_{t \rightarrow t_0} \frac{\hat{f}(t) - \hat{f}(t_0)}{t - t_0} = \\ &= \lim_{t \rightarrow t_0} \frac{1}{t - t_0} \left(\int_{\mathbb{R}} f(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} - \int_{\mathbb{R}} f(x) e^{-ixt_0} \frac{dx}{\sqrt{2\pi}} \right) = \\ &= \lim_{t \rightarrow t_0} \frac{1}{t - t_0} \int_{\mathbb{R}} f(x) (e^{-ixt} - e^{-ixt_0}) \frac{dx}{\sqrt{2\pi}} = \\ &= \lim_{t \rightarrow t_0} \int_{\mathbb{R}} f(x) \frac{e^{-ixt} - e^{-ixt_0}}{t - t_0} \cdot \frac{dx}{\sqrt{2\pi}}.\end{aligned}$$

Stimiamo l'integrando con lo scopo di passare al limite sotto il segno di integrale:

$$\begin{aligned}\left| f(x) \frac{e^{-ixt} - e^{-ixt_0}}{t - t_0} \right| &= \left| \frac{f(x)}{t - t_0} \int_{t_0}^t -ix e^{-ixu} du \right| \leq \\ &\leq \frac{|f(x)|}{|t - t_0|} \left| \int_{t_0}^t |-ix e^{-ixu}| du \right| \leq \\ &\leq \frac{|f(x)|}{|t - t_0|} \left| \int_{t_0}^t |x| du \right| = \frac{|f(x)|}{|t - t_0|} |x| \cdot |t - t_0| = \\ &= |xf(x)|.\end{aligned}$$

Se supponiamo che $g(x) = -ixf(x)$ sia in $L^1(\mathbb{R})$, allora possiamo applicare la convergenza dominata e passare al limite sotto il segno di integrale:

$$\begin{aligned}\hat{f}'(t_0) &= \lim_{t \rightarrow t_0} \int_{\mathbb{R}} f(x) \frac{e^{-ixt} - e^{-ixt_0}}{t - t_0} \cdot \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} \lim_{t \rightarrow t_0} f(x) \frac{e^{-ixt} - e^{-ixt_0}}{t - t_0} \cdot \frac{dx}{\sqrt{2\pi}} =\end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}} f(x)(-ixe^{-ixt_0}) \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} -ixf(x)e^{-ixt_0} \frac{dx}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} g(x)e^{-ixt_0} \frac{dx}{\sqrt{2\pi}} = \hat{g}(t_0).
\end{aligned}$$

□

Esercizio 7.2.5. Se $f, f' \in L^1(\mathbb{R})$ dimostrare che la trasformata di Fourier di f' è data da $\widehat{(f')}(t) = it\hat{f}(t)$.

Esercizio 7.2.6. Se $f \in L^1(\mathbb{R})$ è a tale che $f(x) = \overline{f(-x)}$, allora \hat{f} è a valori reali. In particolare questo succede se f è a valori reali e pari. Cosa succede se f è a valori reali e dispari?

Esercizio 7.2.7. Se $f \in L^1(\mathbb{R})$ e poniamo $g(x) := f(-x)$ allora $\hat{g}(t) = \hat{f}(-t)$. Dedurre che se una funzione h è la trasformata di Fourier di una qualche funzione di L^1 , anche la “speculare” $t \mapsto h(-t)$ è la trasformata di Fourier di una funzione di L^1 . In altre parole, l’immagine di L^1 tramite la trasformata di Fourier è invariante per specularità.

7.3 Le trasformate di Fourier in L^1 sono continue

Data una funzione $f: \mathbb{R} \rightarrow \mathbb{C}$ e un $y \in \mathbb{R}$ denoteremo con f_y la *traslata* di f definita come

$$f_y(x) = f(x - y) \quad \text{per } x \in \mathbb{R}.$$

Teorema 7.3.1. *Se $1 \leq p < +\infty$ e $f \in L^p(\mathbb{R})$, l’applicazione $y \mapsto f_y$ è uniformemente continua da \mathbb{R} in $L^p(\mathbb{R})$.*

Dimostrazione. Usiamo qui per semplicità la misura di Lebesgue senza normalizzazione. Siano $f \in L^p(\mathbb{R})$ e $\varepsilon > 0$. Sappiamo che l’insieme $\mathcal{C}_c(\mathbb{R})$ delle funzioni continue a supporto compatto è densa in $L^p(\mathbb{R})$. Quindi esiste una $g \in \mathcal{C}_c(\mathbb{R})$ con supporto contenuto nell’intervallo $[-A, A]$ tale che

$$\|f - g\|_p < \varepsilon.$$

La funzione g è uniformemente continua su \mathbb{R} . Quindi esiste $\delta > 0$, che possiamo supporre minore di A , tale che

$$|s - t| < \delta \Rightarrow |g(s) - g(t)| < \frac{\varepsilon}{(3A)^{1/p}}. \quad (7.8)$$

Se $|s - t| < \delta$ allora

$$\begin{aligned}\|g_s - g_t\|_p^p &= \int_{-\infty}^{+\infty} |g_s(x) - g_t(x)|^p dx = \\ &= \int_{-\infty}^{+\infty} |g(x-s) - g(x-t)|^p dx.\end{aligned}$$

Volendo restringere il dominio di integrazione, vediamo che

$$\begin{aligned}x - s \in [-A, A] &\iff x \in [-A + s, A + s] \\ x - t \in [-A, A] &\iff x \in [-A + t, A + t],\end{aligned}$$

per cui, usando la (7.8),

$$\begin{aligned}\|g_s - g_t\|_p^p &= \int_{-A + \min\{s, t\}}^{A + \max\{s, t\}} |g(x-s) - g(x-t)|^p dx \leq \\ &\leq \left(A + \max\{s, t\} - (-A + \min\{s, t\}) \right) \left(\frac{\varepsilon}{(3A)^{1/p}} \right)^p = \\ &= (2A + \max\{s, t\} - \min\{s, t\}) \frac{\varepsilon^p}{3A} \leq \\ &= (2A + |s - t|) \frac{\varepsilon^p}{3A} \leq (2A + \delta) \frac{\varepsilon^p}{3A} < (2A + A) \frac{\varepsilon^p}{3A} = \varepsilon^p.\end{aligned}$$

Osserviamo che le norme in $L^p(\mathbb{R})$ sono invarianti per traslazione: $\|h_s\|_p = \|h\|_p$. Veniamo alle traslate di f :

$$\begin{aligned}\|f_s - f_t\|_p &\leq \|f_s - g_s\|_p + \|g_s - g_t\|_p + \|g_t - f_t\|_p = \\ &= \|(f - g)_s\|_p + \|g_s - g_t\|_p + \|(g - f)_t\|_p = \\ &= \|f - g\|_p + \|g_s - g_t\|_p + \|g - f\|_p < 3\varepsilon.\end{aligned}$$

□

Esercizio 7.3.2. La traslazione è continua anche su $L^\infty(\mathbb{R})$?

Indicheremo con $\mathcal{C}_0(\mathbb{R})$ lo spazio delle funzioni continue $\mathbb{R} \rightarrow \mathbb{C}$ che sono infinitesime all'infinito. Si tratta di un sottospazio chiuso di $\mathcal{C}_b(\mathbb{R})$, il quale è lo spazio di Banach delle funzioni continue e limitate con la norma del sup, a sua volta sottospazio chiuso di $L^\infty(\mathbb{R})$.

Teorema 7.3.3. *Se $f \in L^1(\mathbb{R})$ allora $\hat{f} \in \mathcal{C}_0(\mathbb{R})$ e $\|\hat{f}\|_\infty \leq \|f\|_1$.*

Dimostrazione. La disuguaglianza è già contenuta nella formula (7.7) p. 169. Per dimostrare la continuità di \hat{f} in $t_0 \in \mathbb{R}$ prendiamo il limite:

$$\lim_{t \rightarrow t_0} \hat{f}(t) = \lim_{t \rightarrow t_0} \int_{\mathbb{R}} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}}.$$

La funzione integranda è dominata da $|f(x)|$ che è in L^1 e non dipende da t . Quindi si può passare al limite sotto il segno di integrale:

$$\begin{aligned} \lim_{t \rightarrow t_0} \hat{f}(t) &= \lim_{t \rightarrow t_0} \int_{\mathbb{R}} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} \lim_{t \rightarrow t_0} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} f(x) e^{-it_0 x} \frac{dx}{\sqrt{2\pi}} = \hat{f}(t_0). \end{aligned}$$

Effettivamente \hat{f} è continua $\mathbb{R} \rightarrow \mathbb{C}$. Per trovare il limite all'infinito moltiplichiamo $\hat{f}(t)$ per $-e^{-i\pi} = 1$ e facciamo il cambio di variabile $y = x + \pi/t$, $dy = dx$:

$$\begin{aligned} \hat{f}(t) &= \int_{\mathbb{R}} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} = \overbrace{-e^{-i\pi}}^{=1} \int_{\mathbb{R}} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} = \\ &= - \int_{\mathbb{R}} f(x) e^{-itx} e^{-i\pi} \frac{dx}{\sqrt{2\pi}} = - \int_{\mathbb{R}} f(x) e^{-itx - i\pi} \frac{dx}{\sqrt{2\pi}} \\ &= - \int_{\mathbb{R}} f(x) e^{-it(x + \pi/t)} \frac{dx}{\sqrt{2\pi}} = - \int_{\mathbb{R}} f(y - \pi/t) e^{-ity} \frac{dy}{\sqrt{2\pi}} = \\ &= - \int_{\mathbb{R}} f(x - \pi/t) e^{-itx} \frac{dx}{\sqrt{2\pi}} = - \int_{\mathbb{R}} f_{\pi/t}(x) e^{-itx} \frac{dx}{\sqrt{2\pi}}. \end{aligned}$$

Sommando membro a membro con la definizione di $\hat{f}(t)$ otteniamo

$$\begin{aligned} 2\hat{f}(t) &= \int_{\mathbb{R}} f(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} - \int_{\mathbb{R}} f_{\pi/t}(x) e^{-itx} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} (f(x) - f_{\pi/t}(x)) e^{-itx} \frac{dx}{\sqrt{2\pi}}. \end{aligned}$$

per cui

$$\begin{aligned} |2\hat{f}(t)| &\leq \int_{\mathbb{R}} |(f(x) - f_{\pi/t}(x)) e^{-itx}| \frac{dx}{\sqrt{2\pi}} = \\ &= \|f - f_{\pi/t}\|_1 \rightarrow 0 \quad \text{per } t \rightarrow \pm\infty \end{aligned}$$

perché $f_{\pi/t} \rightarrow f_0 = f$ nella norma di L^1 , in quanto la traslazione è continua in L^1 (teorema 7.3.1 p. 175). \square

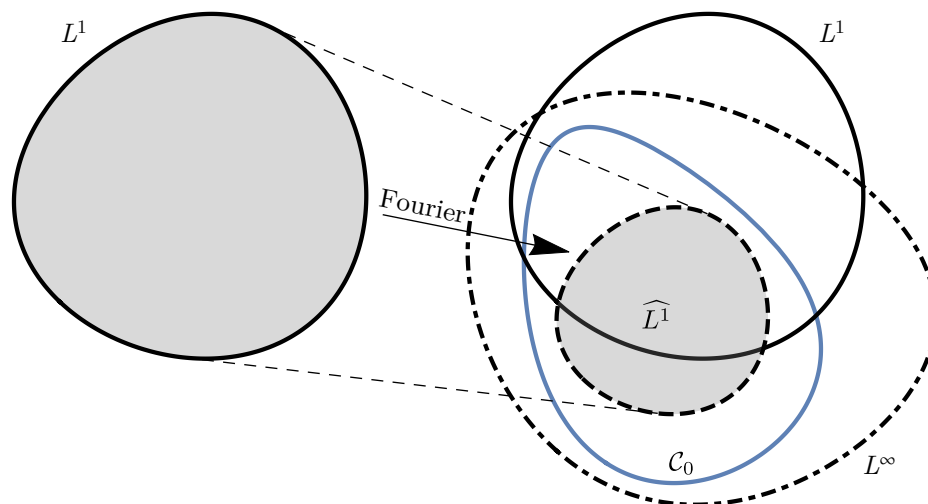


Figura 7.4: Schema di come la trasformata di Fourier si comporta rispetto agli insiemi $L^1(\mathbb{R})$ e \mathcal{C}_0 : l'immagine \widehat{L}^1 (in grigio a destra) è contenuta in $\mathcal{C}_0 \subset L^\infty$, ma non è contenuta in L^1

L'applicazione che manda f in \hat{f} è lineare continua da $L^1(\mathbb{R})$ in $\mathcal{C}_0(\mathbb{R})$.

La trasformata di Fourier *non* manda $L^1(\mathbb{R})$ in se stesso. Riprendiamo infatti la funzione $f = \chi_{[0,1]}$ dell'esempio 7.2.2 di pag. 169. La norma L^1 della sua trasformata di Fourier si può calcolare ed è infinita:

$$\begin{aligned} \|\hat{f}\|_1 &= \int_{\mathbb{R}} |\hat{f}(t)| dt = \int_{\mathbb{R}} \left| \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{-it} - 1}{-it} \right| dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \frac{e^{-it} - 1}{t} \right| dt = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \frac{\cos t - 1 - i \sin t}{t} \right| dt \geq \\ &\geq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \frac{\sin t}{t} \right| dt = +\infty. \end{aligned}$$

La figura 7.4 riassume quanto trovato finora sulle relazioni insiemistiche della trasformata di Fourier rispetto agli spazi $L^1, \mathcal{C}_0, L^\infty$.

Se parto da una f continua e introduco una discontinuità di salto modificandola in un intervallo, questo comporta un cambiamento nell'andamento asintotico di \hat{f} all'infinito.

7.4 Una famiglia di mollificatori

Sono rare le funzioni elementari $f(x)$ per le quali c'è una primitiva elementare di $x \mapsto f(x)e^{-ixt}$ e quindi di cui si può calcolare la trasformata di Fourier usando semplicemente il teorema fondamentale del calcolo. Un esempio prezioso è il seguente, a cui ci appoggeremo in modo determinante nelle dimostrazioni successive. Consideriamo la famiglia di funzioni (figure 7.5 e 7.6)

$$H_\lambda(t) = e^{-|\lambda t|} \quad \text{per } t \in \mathbb{R}, \lambda > 0.$$

Poniamo

$$h_\lambda(x) = \int_{\mathbb{R}} H_\lambda(t) e^{itx} \frac{dt}{\sqrt{2\pi}} = \hat{H}_\lambda(-x). \quad (7.9)$$

Notare che nell'esponentiale c'è un segno +, nonché l'uso delle variabili t, x : vedremo in séguito che H_λ è la trasformata di Fourier di h_λ . Qui il pedice λ non sta a significare una traslazione.

Calcoliamo:

$$\begin{aligned} h_\lambda(x) &= \int_{\mathbb{R}} e^{-|\lambda t|} e^{itx} \frac{dt}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} e^{-|\lambda t|+itx} \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^0 e^{\lambda t+itx} \frac{dt}{\sqrt{2\pi}} + \int_0^{+\infty} e^{-\lambda t+itx} \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^0 e^{(ix+\lambda)t} \frac{dt}{\sqrt{2\pi}} + \int_0^{+\infty} e^{(ix-\lambda)t} \frac{dt}{\sqrt{2\pi}} = \\ &= \frac{1}{\sqrt{2\pi}} \left[\frac{e^{(ix+\lambda)t}}{ix+\lambda} \right]_{t=-\infty}^{t=0} + \frac{1}{\sqrt{2\pi}} \left[\frac{e^{(ix-\lambda)t}}{ix-\lambda} \right]_{t=0}^{t=+\infty} = \\ &= \frac{1}{\sqrt{2\pi}} \left[e^{ixt} \frac{e^{\lambda t}}{ix+\lambda} \right]_{t=-\infty}^{t=0} + \frac{1}{\sqrt{2\pi}} \left[e^{ixt} \frac{e^{-\lambda t}}{ix-\lambda} \right]_{t=0}^{t=+\infty} = \\ &= \frac{1}{\sqrt{2\pi}} \left(e^{ix0} \frac{e^{\lambda 0}}{ix+\lambda} - 0 \right) + \frac{1}{\sqrt{2\pi}} \left(0 - e^{ix0} \frac{e^{-\lambda 0}}{ix-\lambda} \right) = \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{ix+\lambda} - \frac{1}{ix-\lambda} \right) = \frac{1}{\sqrt{2\pi}} \cdot \frac{ix-\lambda-(ix+\lambda)}{(ix+\lambda)(ix-\lambda)} = \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{-2\lambda}{-x^2-\lambda^2} = \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\lambda^2+x^2}. \end{aligned}$$

Una prima proprietà di entrambe le funzioni H_λ, h_λ è che sono a valori reali positivi. Questo non è banale, in quanto i calcoli coinvolgono funzioni

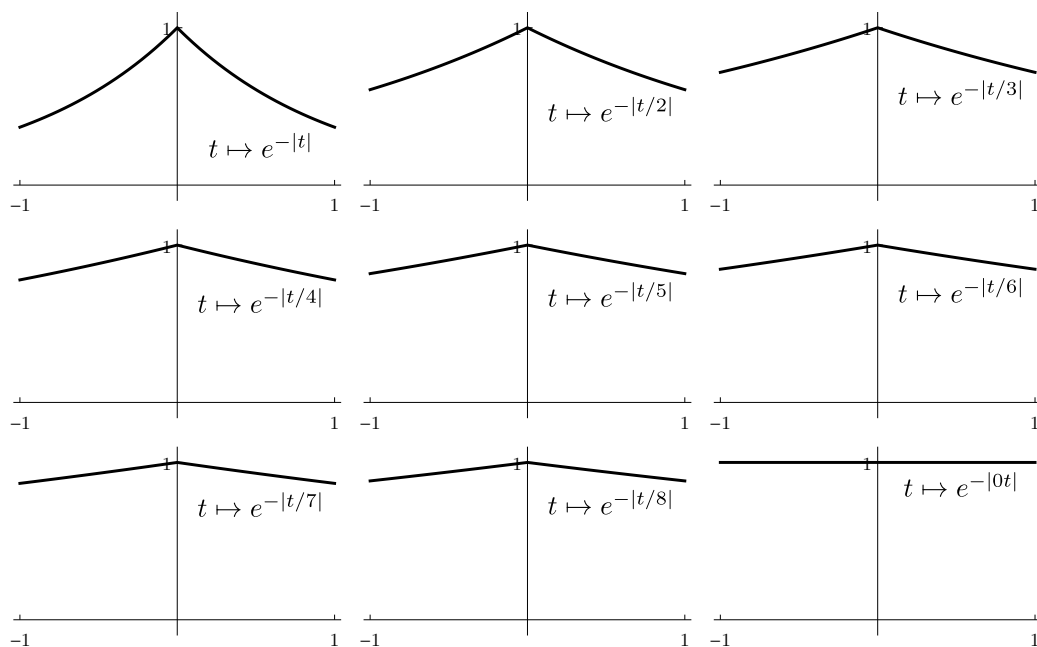


Figura 7.5: Grafici delle funzioni $t \mapsto H_\lambda(t) = e^{-|\lambda t|}$ per $\lambda = 1, 1/2, 1/3, \dots, 1/8$ separatamente e poi il caso limite $\lambda = 0$

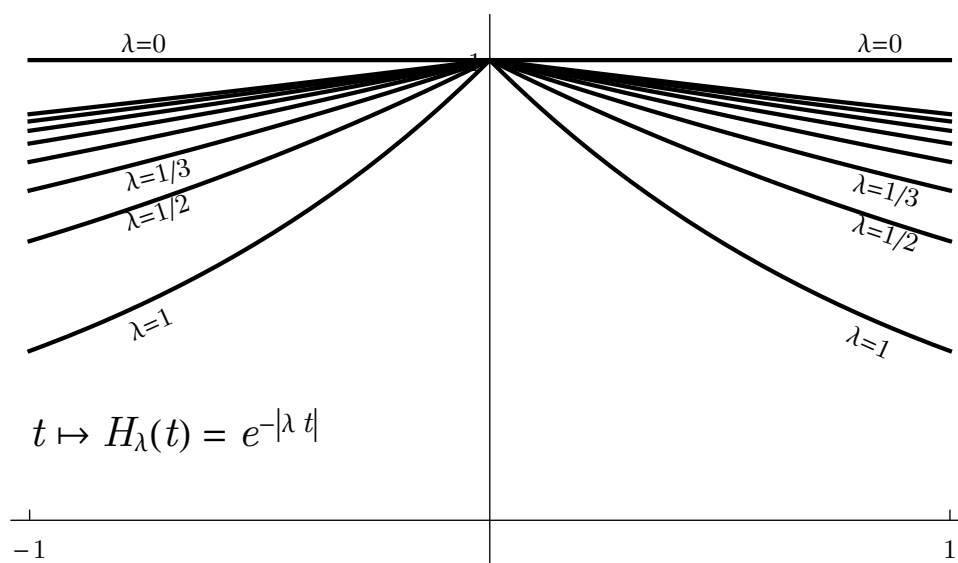


Figura 7.6: Grafici delle funzioni $t \mapsto H_\lambda(t) = e^{-|\lambda t|}$ per $\lambda = 1, 1/2, 1/3, \dots, 1/8$ e poi il caso limite $\lambda = 0$, tutte insieme

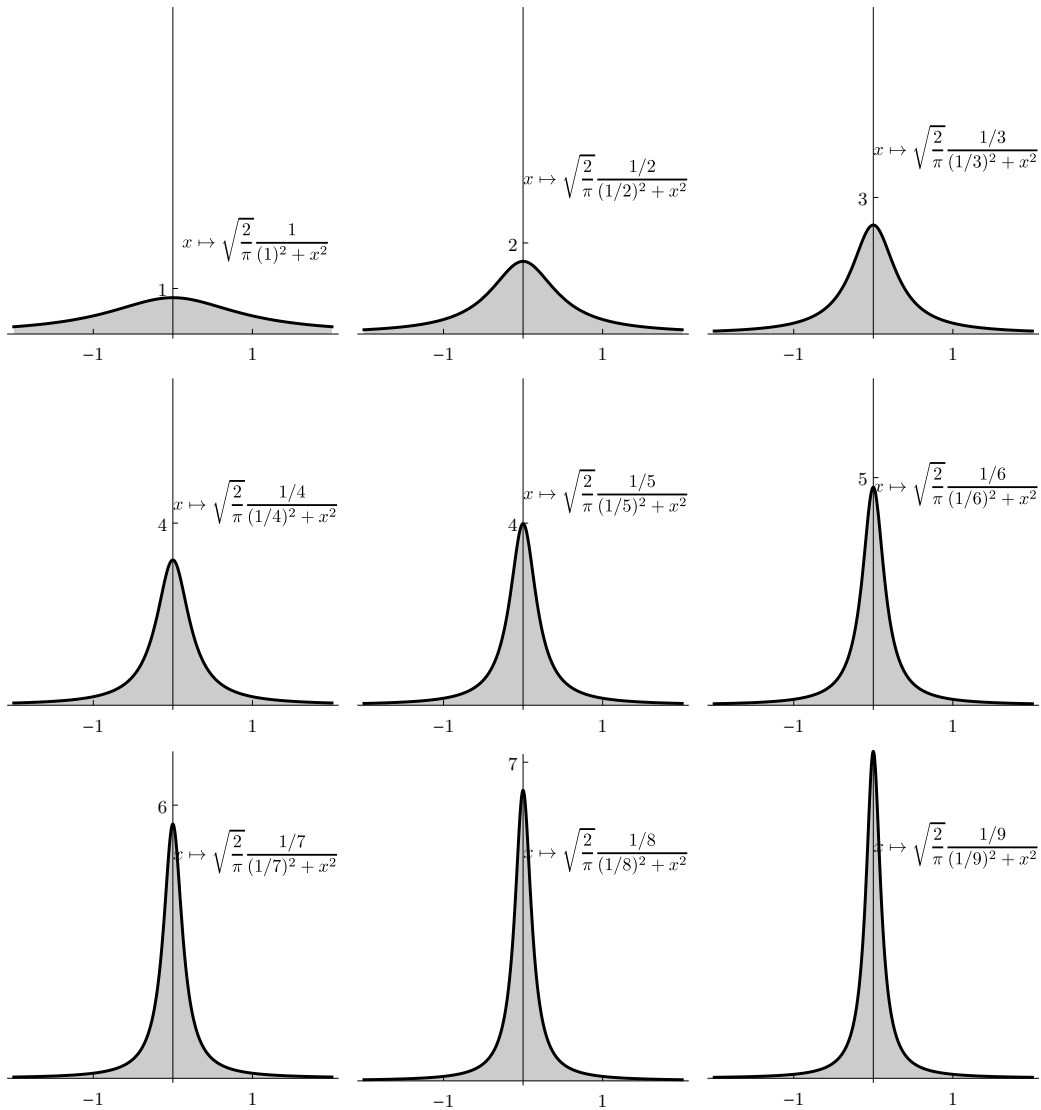


Figura 7.7: Grafici delle funzioni $x \mapsto h_\lambda(x) = \sqrt{2/\pi}\lambda/(\lambda^2 + x^2)$ per $\lambda = 1, 1/2, 1/3, \dots, 1/9$. In tutti i casi l'area, evidenziata in grigio, fra il grafico e l'asse x vale 1

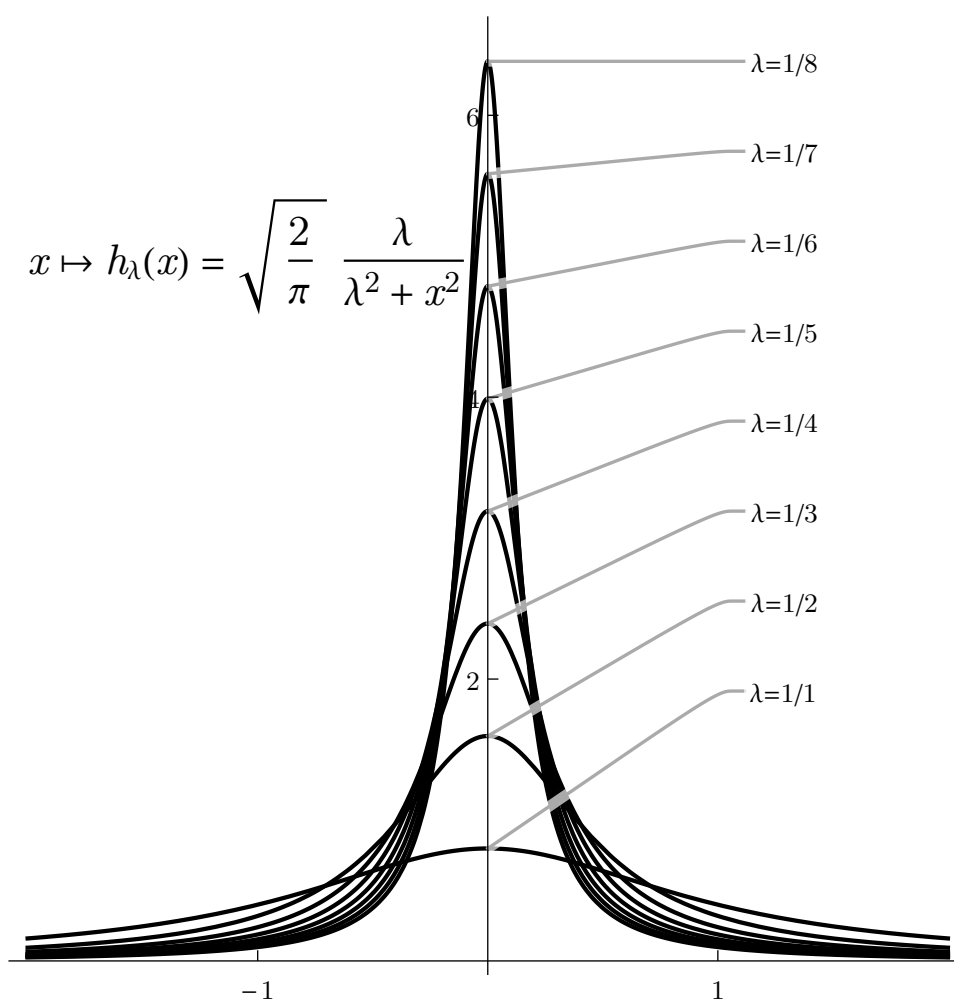


Figura 7.8: Grafici delle funzioni $x \mapsto h_\lambda(x) = \sqrt{2/\pi} \lambda / (\lambda^2 + x^2)$ per $\lambda = 1, 1/2, 1/3, \dots, 1/9$ separatamente

a valori complessi. La singola funzione $t \mapsto H_\lambda(t)$ è continua e tende a 0 esponenzialmente per $t \rightarrow \pm\infty$, per cui $H_\lambda \in L^p(\mathbb{R})$ per ogni $p \in [1, \infty]$. Una proprietà che ci verrà utile della famiglia H_λ è che è monotona decrescente rispetto a λ e che tende a 1 per $\lambda \rightarrow 0^+$:

$$0 < \lambda_1 < \lambda_2 \Rightarrow 0 < H_{\lambda_2}(t) < H_{\lambda_1}(t) \leq 1, \quad \lim_{\lambda \rightarrow 0^+} H_\lambda(x) = 1^-. \quad (7.10)$$

Le funzioni h_λ non sono monotone rispetto a λ , però godono di una simmetria di riscalamento:

$$h_\lambda(x) = \sqrt{2/\pi} \cdot \frac{1/\lambda}{1 + (x/\lambda)^2} = \frac{1}{\lambda} h_1(x/\lambda).$$

Le h_λ hanno tutte integrale 1 rispetto alla misura normalizzata: usando il cambio di variabile $x/\lambda = y$, $dx = \lambda dy$

$$\begin{aligned} \int_{\mathbb{R}} h_\lambda(x) \frac{dx}{\sqrt{2\pi}} &= \int_{\mathbb{R}} h_\lambda(x) \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} \sqrt{\frac{2}{\pi}} \cdot \frac{1/\lambda}{1 + (x/\lambda)^2} \cdot \frac{dx}{\sqrt{2\pi}} = \\ &= \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\lambda} \int_{\mathbb{R}} \frac{1}{1 + y^2} \cdot \frac{\lambda dy}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}} [\arctan y]_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} = \\ &= \sqrt{\frac{2}{\pi}} \cdot \pi \cdot \frac{1}{\sqrt{2\pi}} = 1. \end{aligned}$$

Per $\lambda \rightarrow 0^+$ abbiamo che $h_\lambda(x) \rightarrow 0^+$ se $x \neq 0$, mentre $h_\lambda(0) \rightarrow +\infty$. Le h_λ hanno quindi diverse proprietà delle cosiddette funzioni *mollificatrici* per $\lambda \rightarrow 0^+$. La convoluzione $f * h_\lambda(x_0)$ si può interpretare come la *media pesata* dei valori di $f(x)$ con x vicino a x_0 , usando il peso $h_\lambda(x_0 - x)$. Vedremo nella prossima sezione che in ipotesi ragionevoli su f le mollificate convergono in qualche senso utile alla f . Rispetto alle mollificatrici incontrate in altri corsi, notiamo che queste h_λ *non* hanno supporto compatto.

Se per $E \subset \mathbb{R}$ misurabile definiamo $\mu_\lambda(E)$ come

$$\mu_\lambda(E) = \int_E h_\lambda(x) \frac{dx}{\sqrt{2\pi}}.$$

questa risulta una misura positiva su \mathbb{R} con misura totale $\mu_\lambda(\mathbb{R}) = 1$. Nel calcolo delle probabilità questa misura viene chiamata distribuzione di Cauchy, ed è notevole perché le variabili aleatorie che seguono questa distribuzione

non hanno valor medio, o speranza matematica. Le μ_λ si prestano all'applicazione della disuguaglianza di Jensen: se μ è una misura di probabilità (cioè $\mu(X) = 1$), se $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ è convessa ed $f: X \rightarrow \mathbb{R}$ è in $L^1(\mu)$ allora

$$\varphi\left(\int_X f d\mu\right) \leq \int_X \varphi \circ f d\mu.$$

Nel nostro caso specifico, se $f: \mathbb{R} \rightarrow \mathbb{R}$ è in $L^1(\mu_\lambda)$, cioè se $fh_\lambda \in L^1(\mathbb{R})$, allora

$$\varphi\left(\int_{\mathbb{R}} f(x)h_\lambda(x)\frac{dx}{\sqrt{2\pi}}\right) \leq \int_{\mathbb{R}} \varphi(f(x))h_\lambda(x)\frac{dx}{\sqrt{2\pi}}. \quad (7.11)$$

7.5 Convergenza delle mollificate

La mollificazione di f con h_λ è la convoluzione $x \mapsto (f * h_\lambda)(x)$. Questa si può esprimere alternativamente in termini di H_λ e della trasformata \hat{f} :

Proposizione 7.5.1. *Se $f \in L^1(\mathbb{R})$ allora*

$$(f * h_\lambda)(x) = \int_{\mathbb{R}} H_\lambda(t)\hat{f}(t)e^{ixt}\frac{dt}{\sqrt{2\pi}} \quad \forall x \in \mathbb{R}.$$

Dimostrazione. Poiché $f \in L^1(\mathbb{R})$ e $h_\lambda \in L^\infty$ ha senso la convoluzione in senso puntuale, e non solo quasi ovunque. Usiamo la definizione di h_λ in termini di H_λ

$$\begin{aligned} (f * h_\lambda)(x) &= \int_{\mathbb{R}} f(x-y)h_\lambda(y)\frac{dy}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} f(x-y)\left(\int_{\mathbb{R}} H_\lambda(t)e^{ity}\frac{dt}{\sqrt{2\pi}}\right)\frac{dy}{\sqrt{2\pi}} \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y)H_\lambda(t)e^{ity}\frac{dt}{\sqrt{2\pi}}\right)\frac{dy}{\sqrt{2\pi}}. \end{aligned}$$

Per scambiare l'ordine di integrazione vediamo se la funzione $(t, y) \mapsto f(x-y)H_\lambda(t)e^{ity}$ è in $L^1(\mathbb{R}^2)$:

$$\begin{aligned} \int_{\mathbb{R} \times \mathbb{R}} |f(x-y)H_\lambda(t)e^{ity}| \frac{dt}{\sqrt{2\pi}} \frac{dy}{\sqrt{2\pi}} &= \int_{\mathbb{R} \times \mathbb{R}} |f(x-y)|e^{-\lambda|t|} \frac{dt}{\sqrt{2\pi}} \frac{dy}{\sqrt{2\pi}} = \\ &= \left(\int_{\mathbb{R}} |f(x-y)| \frac{dy}{\sqrt{2\pi}}\right) \left(\int_{\mathbb{R}} e^{-\lambda|t|} \frac{dt}{\sqrt{2\pi}}\right) = \end{aligned}$$

$$= \frac{2}{\lambda} \|f\|_1 < +\infty.$$

Quindi si può scambiare l'ordine e poi fare il cambio di variabile $x - y = u$, $dy = -du$

$$\begin{aligned} (f * h_\lambda)(x) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y) H_\lambda(t) e^{ity} \frac{dy}{\sqrt{2\pi}} \right) \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} H_\lambda(t) \left(\int_{\mathbb{R}} f(x-y) e^{ity} \frac{dy}{\sqrt{2\pi}} \right) \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} H_\lambda(t) \left(\int_{+\infty}^{-\infty} f(u) e^{it(x-u)} \frac{-du}{\sqrt{2\pi}} \right) \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} H_\lambda(t) \left(\int_{-\infty}^{+\infty} f(u) e^{itx} e^{-itu} \frac{du}{\sqrt{2\pi}} \right) \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} H_\lambda(t) e^{itx} \left(\int_{-\infty}^{+\infty} f(u) e^{-itu} \frac{du}{\sqrt{2\pi}} \right) \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} H_\lambda(t) e^{itx} \hat{f}(t) \frac{dt}{\sqrt{2\pi}}. \end{aligned}$$

□

Le mollificate di funzioni limitate convergono puntualmente alla funzione nei punti di continuità:

Teorema 7.5.2. *Se $g \in L^\infty(\mathbb{R})$ e x è un punto di continuità per g allora*

$$\lim_{\lambda \rightarrow 0^+} (g * h_\lambda)(x) = g(x).$$

Dimostrazione. Usiamo il fatto che h_λ ha integrale 1, la sua proprietà di riscaldamento e il cambio di variabile $y/\lambda = s$, $dy = \lambda ds$

$$\begin{aligned} (g * h_\lambda)(x) - g(x) &= \int_{\mathbb{R}} g(x-y) h_\lambda(y) \frac{dy}{\sqrt{2\pi}} - g(x) \overbrace{\int_{\mathbb{R}} h_\lambda(y) \frac{dy}{\sqrt{2\pi}}}^{=1} = \\ &= \int_{\mathbb{R}} g(x-y) h_\lambda(y) \frac{dy}{\sqrt{2\pi}} - \int_{\mathbb{R}} g(x) h_\lambda(y) \frac{dy}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} (g(x-y) - g(x)) h_\lambda(y) \frac{dy}{\sqrt{2\pi}} = \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}} (g(x-y) - g(x)) \frac{1}{\lambda} h_1(y/\lambda) \frac{dy}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} (g(x-\lambda s) - g(x)) \frac{1}{\lambda} h_1(s) \frac{\lambda ds}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} (g(x-\lambda s) - g(x)) h_1(s) \frac{ds}{\sqrt{2\pi}}.
\end{aligned}$$

L'ultimo integrando è dominato da $2\|g\|_{\infty}h_1(s)$, che è in $L^1(\mathbb{R})$. Quindi si può passare al limite sotto il segno di integrale per $\lambda \rightarrow 0^+$, usando l'ipotesi che g sia continua in x :

$$\begin{aligned}
\lim_{\lambda \rightarrow 0^+} ((g * h_{\lambda})(x) - g(x)) &= \lim_{\lambda \rightarrow 0^+} \int_{\mathbb{R}} (g(x-\lambda s) - g(x)) h_1(s) \frac{ds}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} \lim_{\lambda \rightarrow 0^+} (g(x-\lambda s) - g(x)) h_1(s) \frac{ds}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} (g(x-0) - g(x)) h_1(s) \frac{ds}{\sqrt{2\pi}} = \\
&= 0.
\end{aligned}$$

□

Le mollificate convergono anche globalmente nella norma di L^p :

Teorema 7.5.3. *Sia $1 \leq p < +\infty$ e $f \in L^p(\mathbb{R})$. Allora*

$$f * h_{\lambda} \in L^p(\mathbb{R}) \quad \forall \lambda > 0$$

e inoltre

$$\lim_{\lambda \rightarrow 0^+} \|f * h_{\lambda} - f\|_p = 0.$$

Dimostrazione. Poiché $h_{\lambda} \in L^q(\mathbb{R})$, dove q è l'esponente coniugato di p , la convoluzione $f * h_{\lambda}$ è definita puntualmente per ogni $x \in \mathbb{R}$. Come nel teorema precedente, ma con f al posto di g , possiamo scrivere

$$(f * h_{\lambda})(x) - f(x) = \int_{\mathbb{R}} (f(x-y) - f(x)) h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}}.$$

Possiamo applicare la disuguaglianza di Jensen (7.11) di pag. 184 con $\varphi(x) := |x|^p$, che è convessa:

$$|(f * h_{\lambda})(x) - f(x)|^p = \varphi((f * h_{\lambda})(x) - f(x)) =$$

$$\begin{aligned}
&= \varphi \left(\int_{\mathbb{R}} \underbrace{(f(x-y) - f(x)) h_{\lambda}(y)}_{\substack{\in L^1(\mathbb{R}) \text{ (Hölder)} \\ \in L^1(\mu_{\lambda})}} \frac{dy}{\sqrt{2\pi}} \right) \leq \\
&\leq \int_{\mathbb{R}} \varphi(f(x-y) - f(x)) h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} |f(x-y) - f(x)|^p h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}}.
\end{aligned}$$

Integriamo rispetto a x e scambiamo l'ordine di integrazione (le funzioni sono positive), e usiamo il fatto che $h_{\lambda}(y) = h_{\lambda}(-y)$:

$$\begin{aligned}
\|f * h_{\lambda} - f\|_p^p &= \int_{\mathbb{R}} |(f * h_{\lambda})(x) - f(x)|^p \frac{dx}{\sqrt{2\pi}} \leq \\
&\leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) - f(x)|^p h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}} \right) \frac{dx}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f_y(x) - f(x)|^p \frac{dx}{\sqrt{2\pi}} \right) h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} \overbrace{\|f_y - f\|_p^p}^{=:g(y)} h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}} = \int_{\mathbb{R}} g(y) h_{\lambda}(y) \frac{dy}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} h_{\lambda}(0-y) g(y) \frac{dy}{\sqrt{2\pi}} = \\
&= (h_{\lambda} * g)(0) = (g * h_{\lambda})(0),
\end{aligned}$$

dove abbiamo posto $g(y) := \|f_y - f\|_p^p$. In particolare

$$\|f * h_{\lambda}\|_p \leq \|f * h_{\lambda} - f\|_p + \|f\|_p \leq (h_{\lambda} * g)(0)^{1/p} + \|f\|_p < +\infty,$$

per cui $f * h_{\lambda} \in L^p(\mathbb{R})$. Sappiamo poi che la traslazione è continua in L^p , e quindi la funzione g è continua su tutto \mathbb{R} . Inoltre g è limitata, perché

$$\begin{aligned}
0 \leq g(y) &= \|f_y - f\|_p^p \leq (\|f_y\|_p + \|f\|_p)^p = (\|f\|_p + \|f\|_p)^p = \\
&= 2^p \|f\|_p^p \quad \forall y \in \mathbb{R}.
\end{aligned}$$

Possiamo quindi applicare il teorema precedente 7.5.2 e ottenere che

$$\max_{\lambda \rightarrow 0^+} \|f * h_{\lambda} - f\|_p^p \leq \lim_{\lambda \rightarrow 0^+} (g * h_{\lambda})(0) = g(0) = \|f_0 - f\|_p^p = 0.$$

□

7.6 Il teorema di inversione

Teorema 7.6.1. *Se assumiamo che $f \in L^1(\mathbb{R})$ e che anche $\hat{f} \in L^1(\mathbb{R})$, e se poniamo*

$$g(x) := \int_{\mathbb{R}} \hat{f}(t) e^{ixt} \frac{dt}{\sqrt{2\pi}} \quad \text{per } x \in \mathbb{R}, \quad (7.12)$$

allora $g \in C_0(\mathbb{R})$ e $f = g$ quasi ovunque.

Dimostrazione. La g sta in $C_0(\mathbb{R})$ perché possiamo scrivere $g(x) = \hat{f}(-x)$, e quindi, salvo la riflessione, coincide con la trasformata di Fourier di una funzione di L^1 . Prendiamo le mollificazioni $f * h_\lambda$ della f . Per la proposizione 7.5.1 p. 184 possiamo esprimere le mollificate in termini di H_λ :

$$(f * h_\lambda)(x) = \int_{\mathbb{R}} H_\lambda(t) \hat{f}(t) e^{ixt} \frac{dt}{\sqrt{2\pi}} \quad \forall x \in \mathbb{R}.$$

Le funzioni integrande sono dominate grazie all'ipotesi che $\hat{f} \in L^1(\mathbb{R})$:

$$|H_\lambda(t) \hat{f}(t) e^{ixt}| = |e^{-|\lambda t|} \hat{f}(t)| \leq |\hat{f}(t)| \in L^1(\mathbb{R}).$$

Quindi si può passare al limite sotto il segno di integrale:

$$\begin{aligned} \lim_{\lambda \rightarrow 0^+} (f * h_\lambda)(x) &= \lim_{\lambda \rightarrow 0^+} \int_{\mathbb{R}} H_\lambda(t) \hat{f}(t) e^{ixt} \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} \lim_{\lambda \rightarrow 0^+} H_\lambda(t) \hat{f}(t) e^{ixt} \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} 1 \hat{f}(t) e^{ixt} \frac{dt}{\sqrt{2\pi}} = \\ &= g(x) \quad \text{per ogni } x \in \mathbb{R}. \end{aligned}$$

D'altra parte, poiché $f \in L^1(\mathbb{R})$, per il teorema 7.5.3 p. 186 applicato con $p = 1$, le mollificate convergono in $L^1(\mathbb{R})$ a f . Per le proprietà generali della convergenza in L^1 , esiste una successione $\lambda_n \rightarrow 0^+$ per la quale c'è convergenza quasi ovunque:

$$\lim_{n \rightarrow +\infty} (f * h_{\lambda_n})(x) = f(x) \quad \text{per quasi ogni } x \in \mathbb{R}.$$

Per l'unicità del limite deduciamo che $f(x) = g(x)$ per quasi ogni $x \in \mathbb{R}$. \square

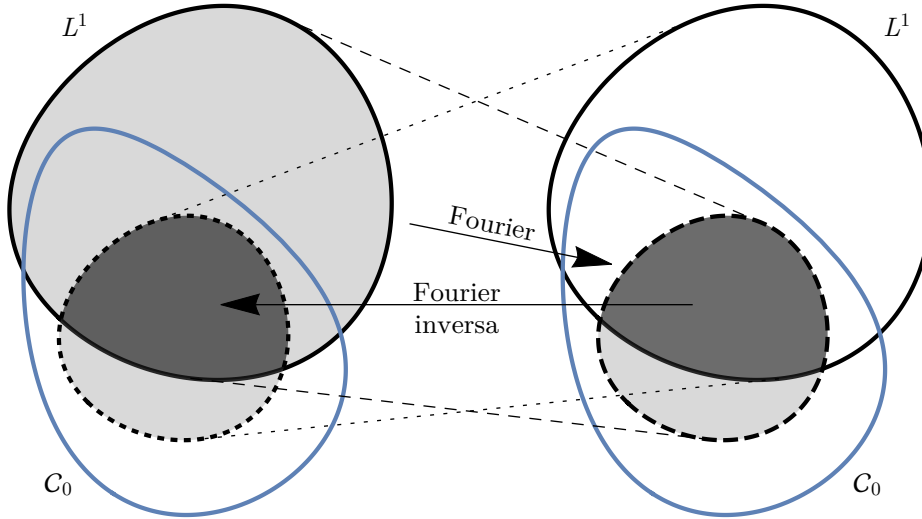


Figura 7.9: Schema di come la trasformata di Fourier e la sua inversa si comportano rispetto agli insiemi $L^1(\mathbb{R})$ e C_0

Corollario 7.6.2 (La trasformata di Fourier è quasi l'inversa di se stessa).
 Se $f, \hat{f} \in L^1(\mathbb{R})$ allora $f(x) = \hat{f}(-x)$ per quasi ogni x .

Corollario 7.6.3 (La trasformata di Fourier è iniettiva su L^1). Se $f, g \in L^1(\mathbb{R})$ e $\hat{f} = \hat{g}$ allora $f = g$ quasi ovunque.

Dimostrazione. Per la linearità, da $\hat{f} = \hat{g}$ segue che $\widehat{f - g} = \hat{f} - \hat{g} = 0$. La funzione $f - g$ è in L^1 e la sua trasformata di Fourier è pure in L^1 , essendo nulla. Quindi per quasi ogni x

$$f(x) - g(x) = \int_{\mathbb{R}} \widehat{(f - g)}(t) e^{itx} \frac{dt}{\sqrt{2\pi}} = \int_{\mathbb{R}} 0 \cdot e^{itx} \frac{dt}{\sqrt{2\pi}} = 0.$$

□

Corollario 7.6.4. Se $f \in L^1(\mathbb{R})$ è essenzialmente non continua (cioè nessuna funzione quasi ovunque uguale a f è continua), allora $\hat{f} \notin L^1(\mathbb{R})$.

A pagina 178 abbiamo verificato direttamente che la funzione $f = \chi_{[0,1]}$ dell'esempio 7.2.2 di pag. 169 ha trasformata di Fourier non in $L^1(\mathbb{R})$. Questo si può dedurre anche dal fatto che f è essenzialmente non continua.

Corollario 7.6.5. *Le funzioni h_λ e H_λ della sezione 7.4 sono una la trasformata di Fourier dell'altra.*

Dimostrazione. Sappiamo dalla formula (7.9) di p. 179 che $h_\lambda(x) = \hat{H}_\lambda(-x)$ e inoltre $H_\lambda, h_\lambda \in L^1(\mathbb{R})$. Quindi sia H_λ che \hat{H}_λ appartengono a L^1 e possiamo applicare il teorema di inversione: col cambio di variabile $y = -x$, $dy = -dx$

$$\begin{aligned} H_\lambda(t) &= \int_{\mathbb{R}} \hat{H}_\lambda(x) e^{ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} h_\lambda(-x) e^{ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{+\infty}^{-\infty} h_\lambda(y) e^{i(-y)t} \frac{-dy}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} h_\lambda(y) e^{-iyt} \frac{dy}{\sqrt{2\pi}} = \\ &= \hat{h}_\lambda(t). \end{aligned}$$

Quanto all'altra, basta ricordare che h_λ è una funzione pari: $h_\lambda(x) = h_\lambda(-x) = \hat{H}_\lambda(x)$ ancora dalla (7.9). \square

Esercizio 7.6.6. Provare a dimostrare il teorema di inversione prendendo la formula di inversione (7.12), sostituendo la definizione integrale (7.6) di \hat{f} , e poi invertendo l'ordine di integrazione.

7.7 Il teorema di Plancherel

La trasformata di Fourier si può ambientare in $L^2(\mathbb{R})$, e qui ha una simmetria molto più pulita che in $L^1(\mathbb{R})$, a costo di rinunciare in generale alla definizione puntuale in termini di integrale.

Teorema 7.7.1 (di Plancherel). *Esiste un'unica $\Phi: L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ lineare continua tale che:*

1. $\Phi(f) = \hat{f}$ per ogni $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, cioè la Φ coincide con la trasformata di Fourier già nota, dove questa ha senso;
2. $\|\Phi(f)\|_2 = \|f\|_2$ per ogni $f \in L^2(\mathbb{R})$, cioè Φ è una isometria di $L^2(\mathbb{R})$ in sé;
3. Φ è suriettiva su tutto $L^2(\mathbb{R})$;
4. l'inversa di Φ coincide con lo speculare di Φ , cioè per ogni $f \in L^2(\mathbb{R})$ si ha che

$$(\Phi^{-1}f)(x) = (\Phi f)(-x) \quad \text{per quasi ogni } x \in \mathbb{R};$$

5. fissata $f \in L^2(\mathbb{R})$ e definite, per $A > 0$, le funzioni

$$\varphi_A(t) := \int_{-A}^A f(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}}, \quad \psi_A(x) := \int_{-A}^A (\Phi f)(t) e^{+ixt} \frac{dt}{\sqrt{2\pi}},$$

abbiamo che $\varphi_A \rightarrow \Phi f$ e $\psi_A \rightarrow f$ nella norma di L^2 per $A \rightarrow +\infty$.

Dimostrazione. L'unicità viene dal fatto che l'intersezione $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ è densa in L^2 : si pensi che le funzioni continue a supporto compatto, oppure le funzioni a gradino, appartengono ad ogni $L^p(\mathbb{R})$. La Φ ha il valore fissato su questo denso per la proprietà 1 ed è continua per la 2, e pertanto è univocamente determinata su tutto $L^2(\mathbb{R})$.

Dimostriamo ora che

$$f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \implies \|\hat{f}\|_2 = \|f\|_2.$$

Prendiamo $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Poniamo $\tilde{f}(x) := \overline{f(-x)}$. Il coniugio e la simmetria speculare non cambiano le norme in L^p , per cui anche $\tilde{f} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Riprendendo il conto del teorema 7.2.4 punto 4 pagina 172, calcoliamo la trasformata di Fourier di \tilde{f} :

$$\begin{aligned} \hat{\tilde{f}}(t) &= \int_{\mathbb{R}} \tilde{f}(x) e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \overline{f(-x)} e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \int_{+\infty}^{-\infty} \overline{f(y)} e^{iyt} \frac{-dy}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \overline{f(y)} e^{iyt} \frac{dy}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} \overline{f(y) e^{-iyt}} \frac{dy}{\sqrt{2\pi}} = \overline{\int_{-\infty}^{+\infty} f(y) e^{-iyt} \frac{dy}{\sqrt{2\pi}}} = \\ &= \overline{\hat{f}(t)}. \end{aligned}$$

Consideriamo ora la convoluzione

$$g := f * \tilde{f}.$$

Per la proposizione 7.1.1 di p. 165, poiché $f, \tilde{f} \in L^1(\mathbb{R})$, abbiamo che $g \in L^1(\mathbb{R})$. Inoltre, poiché $f, \tilde{f} \in L^2(\mathbb{R})$, abbiamo che $g \in L^\infty(\mathbb{R})$. La trasformata di Fourier di g è il prodotto delle trasformate di f e \tilde{f} (proposizione 7.2.4 punto 3):

$$\hat{g}(t) = \hat{f}(t) \hat{\tilde{f}}(t) = \hat{f}(t) \overline{\hat{f}(t)} = |\hat{f}(t)|^2.$$

La convoluzione si può riportare a un prodotto scalare in $L^2(\mathbb{R})$ usando fra l'altro il cambio di variabile $-y = u$, $dy = -du$:

$$\begin{aligned} g(x) &= \int_{-\infty}^{+\infty} f(x-y)\tilde{f}(y)\frac{dy}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} f(x-y)\overline{f(-y)}\frac{dy}{\sqrt{2\pi}} = \\ &= \int_{+\infty}^{-\infty} f(x+u)\overline{f(u)}\cdot\frac{-du}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} f(x-(-u))\overline{f(u)}\frac{du}{\sqrt{2\pi}} = \\ &= \int_{-\infty}^{+\infty} f_{-x}(u)\overline{f(u)}\frac{du}{\sqrt{2\pi}} = \\ &= \langle f_{-x}, f \rangle_2. \end{aligned}$$

Poiché l'applicazione traslazione $x \mapsto f_{-x}$ è continua $\mathbb{R} \rightarrow L^2(\mathbb{R})$ (teorema 7.3.1 p. 175), la funzione g risulta *continua* su \mathbb{R} . Inoltre è anche *limitata* grazie alla disuguaglianza di Schwarz e all'invarianza della norma di L^2 per traslazioni:

$$|g(x)| = |\langle f_{-x}, f \rangle_2| \leq \|f_{-x}\|_2 \cdot \|f\|_2 = \|f\|_2 \cdot \|f\|_2 = \|f\|_2^2.$$

Riassumendo: $g \in L^\infty(\mathbb{R}) \cap \mathcal{C}(\mathbb{R})$. Mollifichiamo la g usando le h_λ della sezione 7.4 e applichiamo il risultato di convergenza puntuale dei mollificatori nei punti di continuità (proposizione 7.5.2 p. 185):

$$(g * h_\lambda)(x) \rightarrow g(x) \quad \text{per } \lambda \rightarrow 0^+ \text{ per ogni } x \in \mathbb{R},$$

da cui, prendendo $x = 0$,

$$\lim_{\lambda \rightarrow 0^+} (g * h_\lambda)(0) = g(0) = \langle f_{-0}, f \rangle_2 = \langle f, f \rangle_2 = \|f\|_2^2. \quad (7.13)$$

Un diverso modo di calcolare lo stesso limite è consentito dall'espressione delle mollificate in termini delle funzioni H_λ (proposizione 7.5.1 p. 184):

$$(g * h_\lambda)(x) = \int_{\mathbb{R}} H_\lambda(t)\hat{g}(t)e^{ixt}\frac{dt}{\sqrt{2\pi}} \quad \forall x \in \mathbb{R}.$$

Poniamo di nuovo $x = 0$:

$$(g * h_\lambda)(0) = \int_{\mathbb{R}} H_\lambda(t)\hat{g}(t)e^{i0t}\frac{dt}{\sqrt{2\pi}} = \int_{\mathbb{R}} e^{-|\lambda t|}|\hat{f}(t)|^2\frac{dt}{\sqrt{2\pi}}.$$

Vogliamo passare al limite sotto il segno di integrale per $\lambda \rightarrow 0^+$. Non possiamo applicare la convergenza dominata, perché l'unica funzione che domina

l'integrando è $|\hat{f}|^2$, che non sappiamo (ancora) se abbia integrale finito. Viene in aiuto allora il teorema della *convergenza monotona*, in quanto la funzione integranda è a valori reali ≥ 0 e cresce al decrescere di λ (formula (7.10) p. 183):

$$0 < \lambda_1 < \lambda_2 \implies 0 \leq e^{-|\lambda_2 t|} |\hat{f}(t)|^2 \leq e^{-|\lambda_1 t|} |\hat{f}(t)|^2.$$

Possiamo quindi passare al limite sotto il segno di integrale per $\lambda \rightarrow 0^+$

$$\begin{aligned} \lim_{\lambda \rightarrow 0^+} (g * h_\lambda)(0) &= \lim_{\lambda \rightarrow 0^+} \int_{\mathbb{R}} e^{-|\lambda t|} |\hat{f}(t)|^2 \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} \lim_{\lambda \rightarrow 0^+} e^{-|\lambda t|} |\hat{f}(t)|^2 \frac{dt}{\sqrt{2\pi}} = \\ &= \int_{\mathbb{R}} 1 \cdot |\hat{f}(t)|^2 \frac{dt}{\sqrt{2\pi}} = \\ &= \|\hat{f}\|_2^2. \end{aligned}$$

Uguagliando col limite della formula (7.13) otteniamo che in effetti

$$\|\hat{f}\|_2 = \|f\|_2.$$

Il grosso è fatto.

Se poniamo $\Phi(f) = \hat{f}$ per $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, abbiamo mostrato che Φ conserva la norma di L^2 . Poiché, come già ricordato, $L^1 \cap L^2$ è denso in L^2 , esiste una estensione di Φ a tutto L^2 che conserva la norma di L^2 , ed è unica.

Sia Y l'immagine di $L^1 \cap L^2$ tramite Φ , cioè

$$Y := \{\hat{f} : f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})\}.$$

Abbiamo determinato che Y è un sottospazio vettoriale di L^2 . Ci proponiamo di dimostrare che Y è denso in L^2 . Poiché L^2 è uno spazio di Hilbert, questo equivale ad affermare che $Y^\perp = \{0\}$.

Osserviamo che la funzione $t \mapsto e^{i\alpha t} H_\lambda(t)$ ("torsione" della H_λ nella terminologia della proposizione 7.2.4 punto 1 p. 172) è in ogni $L^p(\mathbb{R})$, e quindi in particolare in $L^1 \cap L^2$, per ogni $\lambda > 0$, $\alpha \in \mathbb{R}$. Calcoliamone la trasformata di Fourier di $t \mapsto e^{i\alpha t} H_\lambda(t)$, ricordando la formula (7.9) che definisce h_λ in termini di H_λ , nonché il fatto che h_λ è una funzione pari:

$$\int_{\mathbb{R}} e^{i\alpha t} H_\lambda(t) e^{-ixt} \frac{dt}{\sqrt{2\pi}} = \int_{\mathbb{R}} H_\lambda(t) e^{-ixt+i\alpha t} \frac{dt}{\sqrt{2\pi}} =$$

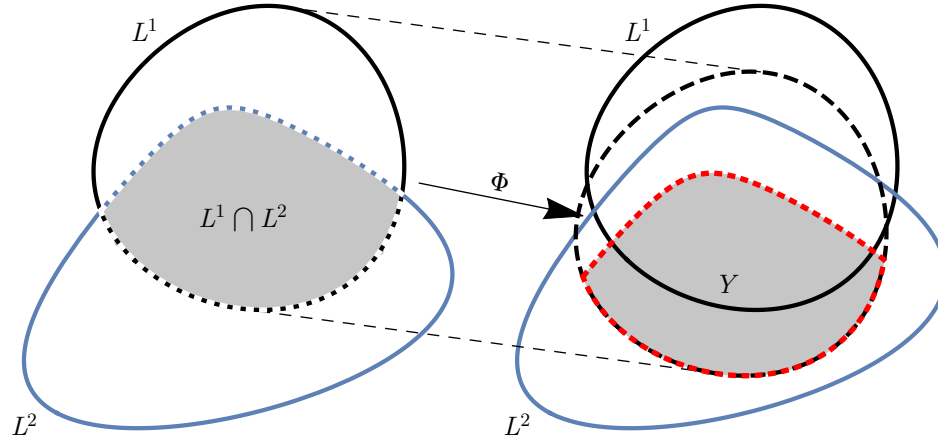


Figura 7.10: Schema di come la trasformata di Fourier si comporta rispetto agli insiemi $L^1(\mathbb{R})$ e $L^2(\mathbb{R})$. La zona in grigio a destra è $Y := \{\hat{f} : f \in L^1 \cap L^2\}$

$$\begin{aligned}
 &= \int_{\mathbb{R}} H_{\lambda}(t) e^{it(\alpha-x)} \frac{dt}{\sqrt{2\pi}} = \\
 &= h_{\lambda}(\alpha - x) = h_{\lambda}(x - \alpha) = (T_{\alpha} h_{\lambda})(x).
 \end{aligned}$$

Le funzioni $t \mapsto h_{\lambda}(x - \alpha)$ sono le traslate di h_{λ} , dove si è usata la notazione alternativa $(T_{\alpha} f)(x) := f(x - \alpha)$ per la traslazione, per non confliggere con la notazione h_{λ} . Dunque Y contiene tutte le traslate delle mollificatrici h_{λ} .

Se ora prendiamo un $w \in Y^{\perp}$ e mollifichiamo il suo coniugato, il risultato è la funzione nulla. Infatti, tenendo conto che $h_{\lambda}(x) = h_{\lambda}(-x)$:

$$\begin{aligned}
 (h_{\lambda} * \bar{w})(\alpha) &= \int_{\mathbb{R}} h_{\lambda}(\alpha - t) \overline{w(t)} \frac{dt}{\sqrt{2\pi}} = \int_{\mathbb{R}} h_{\lambda}(t - \alpha) \overline{w(t)} \frac{dt}{\sqrt{2\pi}} = \\
 &= \int_{\mathbb{R}} (T_{\alpha} h_{\lambda})(t) \overline{w(t)} \frac{dt}{\sqrt{2\pi}} = \\
 &= \underbrace{\langle T_{\alpha} h_{\lambda}, w \rangle}_2 = 0
 \end{aligned}$$

per ogni $\lambda > 0$. D'altra parte le mollificate di \bar{w} tendono a \bar{w} nella norma di L^2 , grazie al teorema 7.5.3 p. 186. Per l'unicità del limite in L^2 , concludiamo che \bar{w} è nullo (quasi ovunque), e quindi anche w .

La trasformazione Φ è dunque una isometria da $L^2(\mathbb{R})$ su un sottinsieme di $L^2(\mathbb{R})$ che contiene Y , il quale è denso in $L^2(\mathbb{R})$. Per dimostrare che Φ è

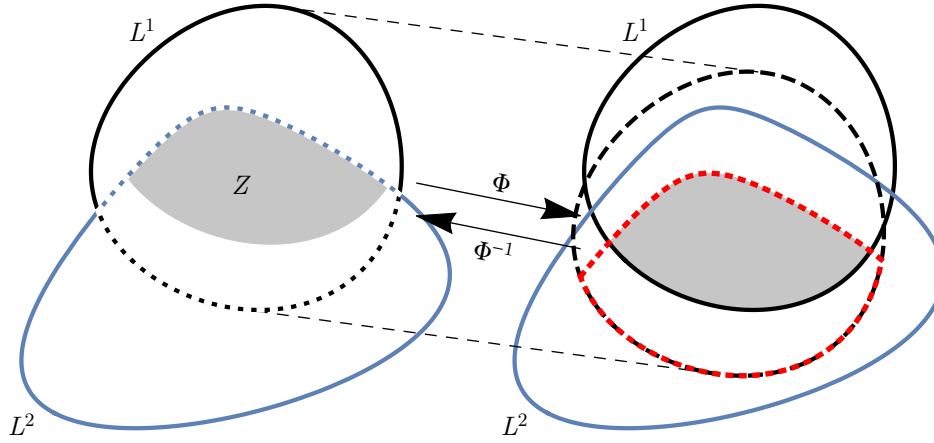


Figura 7.11: Schema di come la trasformata di Fourier e la sua inversa si comportano rispetto agli insiemi $L^1(\mathbb{R})$ e $L^2(\mathbb{R})$. La zona in grigio a sinistra è $Z := \{f \in L^1 \cap L^2 : \hat{f} \in L^1 \cap L^2\}$, e quella a destra è $\Phi(Z)$. Fra queste due zone si può applicare il teorema di inversione stabilito per L^1

suriettiva su tutto L^2 , prendiamo un $g \in L^2$ e dimostriamo che esiste $f \in L^1$ tale che $g = \hat{f}$. Per la densità di Y esiste una successione $f_n \in L^1 \cap L^2$ tale che $\Phi f_n = \hat{f}_n \rightarrow g$ in L^2 . In particolare, \hat{f}_n è una successione di Cauchy in L^2 . Essendo Φ^{-1} una isometria, anche $\Phi^{-1}\hat{f}_n = \Phi^{-1}\Phi f_n = f_n$ è di Cauchy. Quindi esiste $f \in L^2$ tale che $f_n \rightarrow f$ in L^2 . Ma allora, dato che Φ è continua, $\hat{f}_n = \Phi f_n \rightarrow \Phi f$. Per l'unicità del limite $\Phi f = g$.

Consideriamo l'insieme $Z := \{f \in L^1 \cap L^2 : \hat{f} \in L^1 \cap L^2\}$. Mostriamo che Z è denso in $L^2(\mathbb{R})$. Sappiamo già che ogni $f \in L^2$ può essere approssimata a piacere in L^2 con una funzione $g \in L^1 \cap L^2$. Per il teorema 7.5.3 p. 186, le mollificate $g * h_\lambda$ sono anche loro in $L^1 \cap L^2$. Le trasformate di Fourier delle mollificate si possono scrivere come

$$\widehat{g * h_\lambda}(t) = \hat{g}(t)\hat{h}_\lambda(t) = \hat{g}(t)H_\lambda(t),$$

e quindi sono in $L^1 \cap L^2$ perché prodotto di \hat{g} (che è in $\mathcal{C}_0(\mathbb{R})$ e quindi limitata) con H_λ , che appartiene a ogni $L^p(\mathbb{R})$. Pertanto le mollificate $g * h_\lambda$ sono in Z , e, di nuovo per il teorema 7.5.3, sono vicine a piacere a g in L^2 quando λ è piccolo.

Grazie al teorema di inversione 7.6.1, quando $g \in \Phi(Z)$ abbiamo una espressione esplicita di Φ^{-1} :

$$(\Phi^{-1}g)(x) = \int_{\mathbb{R}} g(t)e^{+ixt} \frac{dt}{\sqrt{2\pi}} = (\Phi g)(-x).$$

Poiché $\Phi(Z)$ è denso in L^2 , la relazione

$$(\Phi^{-1}g)(x) = (\Phi g)(-x) \quad \text{per quasi ogni } x \in \mathbb{R}$$

si estende a tutte le $g \in L^2(\mathbb{R})$ perché i due membri sono continui in L^2 rispetto a g .

Per dimostrare il punto 5, prese $f, \Phi f \in L^2$ osserviamo che

$$\begin{aligned} \varphi_A(t) &:= \int_{-A}^A f(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} f(x)\chi_{[-A,A]}(x)e^{-ixt} \frac{dx}{\sqrt{2\pi}} = \\ &= \Phi(\chi_{[-A,A]}f)(t), \\ \psi_A(x) &:= \int_{-A}^A (\Phi f)(t)e^{+ixt} \frac{dt}{\sqrt{2\pi}} = \int_{\mathbb{R}} (\Phi f)(t)\chi_{[-A,A]}(t)e^{+ixt} \frac{dt}{\sqrt{2\pi}} = \\ &= \Phi^{-1}(\chi_{[-A,A]}\hat{f})(x). \end{aligned}$$

La funzione $\chi_{[-A,A]}f$ è in $L^1 \cap L^2$, usando anche la disuguaglianza di Hölder:

$$\begin{aligned} \|\chi_{[-A,A]}f\|_1 &= \int_{\mathbb{R}} |\chi_{[-A,A]}(x)f(x)| \frac{dt}{\sqrt{2\pi}} \leq \\ &\leq \left(\int_{\mathbb{R}} |\chi_{[-A,A]}(x)|^2 \frac{dt}{\sqrt{2\pi}} \right)^{1/2} \cdot \left(\int_{\mathbb{R}} |f(x)|^2 \frac{dt}{\sqrt{2\pi}} \right)^{1/2} = \\ &= \sqrt{2A}\|f\|_2 < +\infty, \\ \|\chi_{[-A,A]}f\|_2^2 &= \int_{\mathbb{R}} |\chi_{[-A,A]}(x)f(x)|^2 \frac{dt}{\sqrt{2\pi}} = \int_{-A}^A |f(x)|^2 \frac{dt}{\sqrt{2\pi}} \leq \\ &\leq \int_{\mathbb{R}} |f(x)|^2 \frac{dt}{\sqrt{2\pi}} = \|f\|_2^2 < +\infty. \end{aligned}$$

Analogamente $\chi_{[-A,A]}\Phi f$ è in $L^1 \cap L^2$. Inoltre $\chi_{[-A,A]}f \rightarrow f$ in L^2 , grazie alla convergenza dominata (l'integranda è maggiorata da $|f|^2 \in L^1$):

$$\lim_{A \rightarrow +\infty} \|f - \chi_{[-A,A]}f\|_2^2 = \lim_{A \rightarrow +\infty} \int_{\mathbb{R}} |f(x) - \chi_{[-A,A]}(x)f(x)|^2 \frac{dt}{\sqrt{2\pi}} =$$

$$\begin{aligned}
&= \int_{\mathbb{R}} \lim_{A \rightarrow +\infty} |f(x) - \chi_{[-A,A]}(x)f(x)|^2 \frac{dt}{\sqrt{2\pi}} = \\
&= \int_{\mathbb{R}} 0 \frac{dt}{\sqrt{2\pi}} = 0
\end{aligned}$$

Poiché Φ è continuo da L^2 in sé,

$$\varphi_A = \Phi(\chi_{[-A,A]}f) \rightarrow \Phi f \quad \text{in } L^2 \text{ per } A \rightarrow +\infty.$$

Dal teorema di inversione in L^1 sappiamo che

$$\psi_A = \Phi^{-1}(\chi_{[-A,A]}\Phi f).$$

Poiché Φ^{-1} è continuo da L^2 in sé e $\chi_{[-A,A]}\Phi f \rightarrow \Phi f$ in L^2 ,

$$\psi_A = \Phi^{-1}(\chi_{[-A,A]}\Phi f) \rightarrow \Phi^{-1}\Phi f = f \quad \text{in } L^2 \text{ per } A \rightarrow +\infty.$$

□

Esempio 7.7.2. Abbiamo già visto nell'esempio 7.2.2 a pag. 169 che la funzione $f = \chi_{[0,1]}$ ha una trasformata di Fourier che non sta in $L^1(\mathbb{R})$:

$$\hat{f}(t) = \begin{cases} \frac{\sin t + i(\cos t - 1)}{t\sqrt{2\pi}} & \text{se } t \neq 0, \\ \frac{1}{\sqrt{2\pi}} & \text{se } t = 0. \end{cases}$$

Però f e anche \hat{f} sono in $L^2(\mathbb{R})$. Definiamo la funzione ψ_A come

$$\psi_A(x) := \int_{-A}^A \hat{f}(t)e^{ixt} \frac{dt}{\sqrt{2\pi}}.$$

La ψ_A non è calcolabile elementarmente, però si verifica facilmente che ψ_A ha valori reali, per simmetria. La figura 7.12 mostra i grafici di alcune approssimanti ψ_A . Si nota che nelle vicinanze ai punti $x = 0$ e $x = 1$, che sono i punti di discontinuità di $f = \chi_{[0,1]}$, le ψ_A presentano delle oscillazioni più grandi che altrove. Questo è chiamato “fenomeno di Gibbs”.

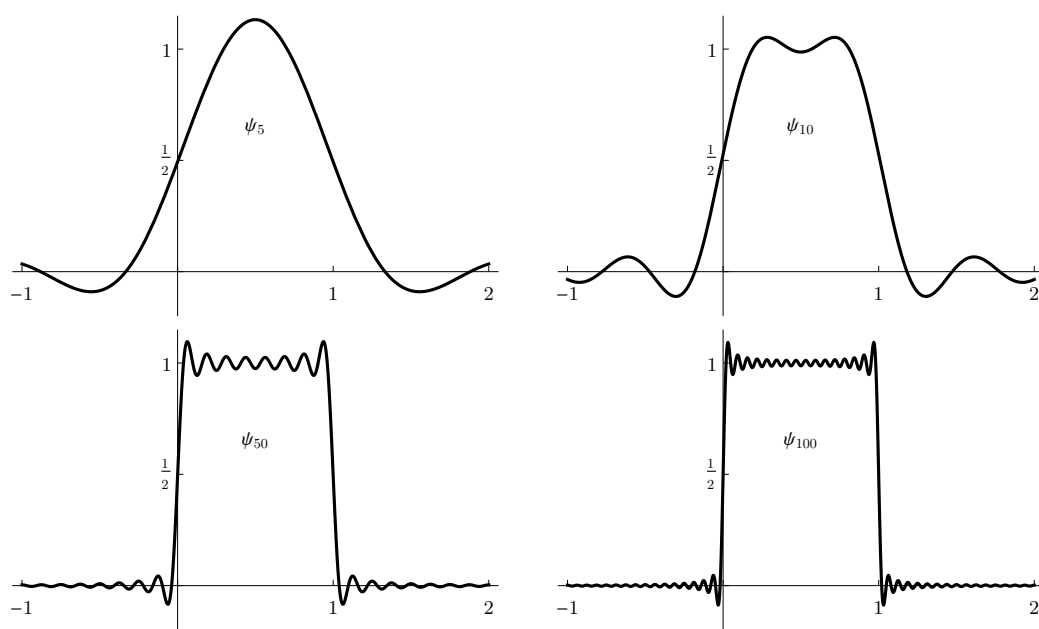


Figura 7.12: Approssimanti ψ_A della trasformata di Fourier inversa della trasformata di Fourier di $\chi_{[0,1]}$

Bibliografia

- [1] Franco Conti. Appunti delle lezioni di Istituzioni di Analisi Superiore. Pisa, anno accademico 1980–81
- [2] Giuseppe De Marco. Analisi Due, prima edizione. Decibel-Zanichelli. ISBN 88-08-01215-8
- [3] Pascual Jordan, John von Neumann. On inner products in linear, metric spaces. *Annals of Mathematics*, Vol. 36, No. 3, July 1935. Disponibile alla pagina <http://www.jstor.org/stable/1968653>
- [4] Discussione sull’etimologia della “identità di polarizzazione”: <https://math.stackexchange.com/questions/79542/polarization-etymology-question> (visitato il 2022-06-03)
- [5] Walter Rudin. Real and complex analysis. Third Edition. McGraw-Hill, 1986