Analisi Funzionale

note delle lezioni

Alessio Martini

Politecnico di Torino
A.A. 2023/2024

Versione: 26 gennaio 2024, 9:53

Indice

0 Notazione **1**

1 Richiami di algebra lineare **2**
1.1 Spazi vettoriali 2
1.2 Basi e dimensione 5
1.3 Applicazioni lineari 7

2 Richiami di topologia in spazi metrici **10**
2.1 Spazi metrici 10
2.2 Successioni e completezza 11
2.3 Convergenza puntuale e convergenza uniforme 15
2.4 Topologia in spazi metrici 16
2.5 Funzioni continue, uniformemente continue e lipschitziane 19

3 Richiami di teoria della misura **22**

4 Spazi normati e spazi di Banach **29**
4.1 Norme e spazi normati 29
4.2 Norme equivalenti 32
4.3 Completezza e chiusura 36
4.4 Serie convergenti e assolutamente convergenti 37
4.5 Compattezza in spazi normati 38
4.6 Il teorema di Stone–Weierstrass 40

5 Spazi di successioni e spazi L^p **44**
5.1 Spazi di successioni 44
5.2 Spazi L^p 53

6 Spazi di Hilbert **62**
6.1 Prodotti scalari e spazi con prodotto scalare 62
6.2 Ortogonalità 68
6.3 Basi ortonormali 76
6.4 Serie di Fourier 84

7 Operatori limitati **89**
7.1 Norma operatoriale e operatori limitati 89
7.2 Isometrie lineari 99
7.3 Estensione di operatori limitati 102
7.4 Il principio di uniforme limitatezza 104
7.5 Operatori limitati invertibili 107

8 Duale di uno spazio normato **114**
8.1 Funzionali lineari e spazio duale 114
8.2 Duale di uno spazio di Hilbert 115
8.3 Duale degli spazi ℓ^p e L^p 117
8.4 Forme sesquilineari e teorema di Lax–Milgram 121
8.5 Estensione di funzionali: il teorema di Hahn–Banach 126

9 Biduale e convergenze deboli **133**
9.1 Biduale e spazi riflessivi 133
9.2 Convergenza debole 135
9.3 Convergenza debole∗ sul duale 140

10 Aggiunto e spettro in spazi di Hilbert **147**
10.1 Aggiunto di un operatore limitato 147
10.2 Operatori autoaggiunti, unitari, normali e proiezioni ortogonali . 152
10.3 Spettro di un operatore limitato 156

11 Operatori compatti e teoria di Fredholm **164**
11.1 Operatori compatti 164
11.2 Teoria di Fredholm 173
11.3 Spettro di operatori compatti 177
11.4 Il teorema spettrale 179

0 Notazione

Con $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ denotiamo l'insieme dei numeri naturali (incluso lo zero), e con $\mathbb{N}_+$ l'insieme $\mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$ degli interi positivi. Con $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ indichiamo l'insieme dei numeri interi, mentre $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ sono gli insiemi dei numeri razionali, reali e complessi.

Per un numero complesso $z = a+ib$, con $a, b \in \mathbb{R}$, denotiamo con $\mathfrak{Re}\, z$ e $\mathfrak{Im}\, z$ le sue parti reale e immaginaria a e b; scriviamo inoltre $\overline{z}$ per il suo coniugato $a - ib$, mentre il modulo di z è dato da $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

In questo corso lavoreremo prevalentemente con spazi vettoriali reali o complessi, cioè spazi vettoriali il cui campo degli scalari è $\mathbb{R}$ oppure $\mathbb{C}$. La teoria nel caso reale è in gran parte analoga a quella nel caso complesso. Per evitare ripetizioni, nella presentazione utilizzeremo spesso il simbolo $\mathbb{F}$ come "segnaposto" per il campo degli scalari, sottintendendo che possa essere rimpiazzato indifferentemente con $\mathbb{R}$ oppure con $\mathbb{C}$ (in maniera coerente).

1 Richiami di algebra lineare

In questo capitolo si riportano brevemente alcune definizioni e alcuni risultati di base dell'algebra lineare (senza dimostrazioni), che verranno diffusamente utilizzati nel seguito.

Ricordiamo che con $\mathbb{F}$ denotiamo il campo dei numeri reali $\mathbb{R}$ oppure il campo dei numeri complessi $\mathbb{C}$.

1.1 Spazi vettoriali

Definizione 1.1. Uno *spazio vettoriale* su $\mathbb{F}$ è un insieme V dotato di due operazioni

$$V \times V \ni (x, y) \mapsto x + y \in V, \qquad \mathbb{F} \times V \ni (\alpha, x) \mapsto \alpha x \in V$$

dette *somma* e *prodotto scalare-vettore*, con le seguenti proprietà:

1. V è un *gruppo abeliano* rispetto alla somma, cioè:
   (a) $\forall x, y, z \in V : (x + y) + z = x + (y + z)$
   (proprietà associativa),
   (b) $\forall x, y \in V : x + y = y + x$
   (proprietà commutativa),
   (c) $\exists! 0 \in V : \forall x \in V : x + 0 = 0 + x = x$
   (vettore nullo: elemento neutro della somma),
   (d) $\forall x \in V : \exists! -x \in V : x + (-x) = (-x) + x = 0$
   (esistenza dell'inverso rispetto alla somma);
2. il prodotto scalare-vettore è compatibile con la somma di V e le operazioni del campo $\mathbb{F}$:
   (a) $\forall \alpha \in \mathbb{F} : \forall x, y \in V : \alpha(x + y) = \alpha x + \alpha y$
   (proprietà distributiva rispetto alla somma di vettori),
   (b) $\forall \alpha, \beta \in \mathbb{F} : \forall x \in V : (\alpha + \beta)x = \alpha x + \beta x$
   (proprietà distributiva rispetto alla somma di scalari),
   (c) $\forall \alpha, \beta \in \mathbb{F} : \forall x \in V : (\alpha\beta)x = \alpha(\beta x)$
   (proprietà associativa mista),
   (d) $\forall x \in V : 1x = x$
   (il prodotto per lo scalare 1 è l'identità su V).

Gli elementi di V sono detti *vettori*, mentre gli elementi di $\mathbb{F}$ sono detti *scalari*.

Esempio 1.2. Ecco alcuni esempi di spazi vettoriali.

(1) Per ogni $n \in \mathbb{N}$, l'insieme $\mathbb{F}^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{F}\}$ delle n-uple di elementi di $\mathbb{F}$ è uno spazio vettoriale su $\mathbb{F}$ con le *operazioni componente per componente*:

$$\begin{aligned} (x_1, \dots, x_n) + (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n), \\ \alpha(x_1, \dots, x_n) &= (\alpha x_1, \dots, \alpha x_n). \end{aligned} \tag{1.1}$$

per ogni $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{F}^n$ e $\alpha \in \mathbb{F}$.

(2) Se $I \subseteq \mathbb{R}$ è un intervallo, l'insieme $C_{\mathbb{F}}(I) = C(I, \mathbb{F})$ delle funzioni continue $f : I \to \mathbb{F}$ è uno spazio vettoriale su $\mathbb{F}$, con le *operazioni puntuali*:

$$(f+g)(t) = f(t) + g(t), \qquad (\alpha f)(t) = \alpha f(t) \qquad \forall t \in I,$$

per ogni $f, g \in C_{\mathbb{F}}(I)$ e $\alpha \in \mathbb{F}$. Scriviamo semplicemente $C(I)$ invece di $C_{\mathbb{F}}(I)$ quando il campo $\mathbb{F}$ è chiaro dal contesto.

(3) Più in generale, dato uno spazio topologico Ω, possiamo considerare lo spazio vettoriale $C_{\mathbb{F}}(\Omega) = C(\Omega, \mathbb{F})$ delle funzioni continue $f : \Omega \to \mathbb{F}$, con le operazioni puntuali. Di nuovo, scriviamo $C(\Omega)$ invece di $C_{\mathbb{F}}(\Omega)$ se $\mathbb{F}$ è chiaro dal contesto.

(4) L'insieme $\mathcal{P} = \mathbb{F}[X]$ dei polinomi in un'indeterminata X a coefficienti in $\mathbb{F}$ è uno spazio vettoriale su $\mathbb{F}$, con le usuale operazioni di somma di polinomi e prodotto scalare-polinomio.

(5) Se V e W sono spazi vettoriali su $\mathbb{F}$, il loro *prodotto diretto*

$$V \times W = \{(v, w) : v \in V,\, w \in W\}$$

è uno spazio vettoriale con le operazioni componente per componente:

$$(v, w) + (v', w') = (v + v', w + w'), \qquad \alpha(v, w) = (\alpha v, \alpha w)$$

per ogni $(v, w), (v', w') \in V \times W$, $\alpha \in \mathbb{F}$.

(6) Se V è uno spazio vettoriale su $\mathbb{C}$, allora V si può anche pensare come uno spazio vettoriale su $\mathbb{R}$, restringendo l'operazione di prodotto scalare-vettore (dato che $\mathbb{R}$ è un sottocampo di $\mathbb{C}$).

(7) Siano S un insieme e V uno spazio vettoriale su $\mathbb{F}$. L'insieme $\mathcal{F}(S, V)$ di tutte le funzioni $f : S \to V$ è uno spazio vettoriale su $\mathbb{F}$ con le *operazioni puntuali*:

$$(f+g)(t) = f(t) + g(t), \qquad (\alpha f)(t) = \alpha f(t) \qquad \forall t \in S, \tag{1.2}$$

per ogni $f, g \in \mathcal{F}(S, V)$ e $\alpha \in \mathbb{F}$.

Definizione 1.3. Sia V uno spazio vettoriale su $\mathbb{F}$. Si dice *sottospazio vettoriale* di V ogni sottoinsieme U di V che, dotato della restrizione delle operazioni su V, sia a sua volta uno spazio vettoriale.

Proposizione 1.4 (caratterizzazione dei sottospazi vettoriali)**.** *Sia V uno spazio vettoriale su $\mathbb{F}$. Sono equivalenti:*

(i) U è un sottospazio vettoriale di V;

(ii) U è un sottoinsieme non vuoto di V tale che

$$\forall x, y \in U : \forall \alpha, \beta \in \mathbb{F} : \alpha x + \beta y \in U$$

(cioè U è "chiuso per combinazioni lineari").

Osservazione 1.5. Se V è uno spazio vettoriale su $\mathbb{F}$ e $x, y \in V$, un vettore della forma $\alpha x + \beta y$ si dice *combinazione lineare* dei vettori x e y con coefficienti $\alpha, \beta \in \mathbb{F}$. Vedremo in seguito (☞ definizione 1.9) una generalizzazione al caso di più vettori.

Esempio 1.6. Ecco alcuni esempi di sottospazi vettoriali.

(1) Per ogni spazio vettoriale V, gli insiemi $\{0\}$ e V sono sottospazi vettoriali di V (rispettivamente, il più piccolo e il più grande sottospazio vettoriale di V).

(2) I sottospazi vettoriali del piano $\mathbb{R}^2$, oltre a $\{0\}$ e $\mathbb{R}^2$ stesso, sono le rette passanti per l'origine.

(3) Per ogni $d \in \mathbb{N}$, l'insieme $\mathcal{P}_d = \{p \in \mathcal{P} : \deg p \leq d\}$ dei polinomi di grado al più d è un sottospazio vettoriale di $\mathcal{P}$.

(4) Sia Ω uno spazio topologico. Lo spazio $C(\Omega, \mathbb{F})$ delle funzioni continue da Ω in $\mathbb{F}$ è un sottospazio vettoriale dello spazio $\mathcal{F}(\Omega, \mathbb{F})$ di tutte le funzioni da Ω in $\mathbb{F}$.

(5) Sia $I \subseteq \mathbb{R}$. L'insieme $\{p|_I : p \in \mathcal{P}\}$ delle restrizioni a I dei polinomi è un sottospazio vettoriale dello spazio $C(I, \mathbb{F})$ delle funzioni continue su I.

(6) Sia V uno spazio vettoriale su $\mathbb{F}$. Se U_1 e U_2 sono sottospazi vettoriali di V, allora anche l'intersezione $U_1 \cap U_2$ e la *somma*
$$U_1 + U_2 \coloneqq \{x + y : x \in U_1,\ y \in U_2\}$$
sono sottospazi vettoriali di V. Invece, in generale, l'unione $U_1 \cup U_2$ <u>non</u> è un sottospazio vettoriale di V (ad esempio, in $\mathbb{R}^2$, l'unione di due rette distinte passanti per l'origine non è una retta). C'è tuttavia una relazione fra unione e somma (☞ osservazione 1.10 in seguito).

Definizione 1.7. Sia V uno spazio vettoriale. Un sottoinsieme A di V si dice *convesso* se
$$\forall x, y \in A : \forall \theta \in [0, 1] : (1 - \theta)x + \theta y \in A$$
(cioè A è "chiuso per combinazioni convesse").

Osservazione 1.8. Siano V uno spazio vettoriale e $x, y \in V$. L'insieme
$$\{(1 - \theta)x + \theta y : \theta \in [0, 1]\}$$
delle combinazioni convesse di x e y non è altro che il segmento di retta di estremi x e y. Dunque un sottoinsieme A di V è convesso se e solo se, per ogni $x, y \in A$, il segmento di estremi x e y è interamente contenuto in A. Le combinazioni convesse sono un particolare tipo di combinazioni lineari; in particolare, ogni sottospazio vettoriale di V è convesso (vedi proposizione 1.4), ma ovviamente il viceversa in generale non vale.

1.2 Basi e dimensione

Definizione 1.9. Sia V uno spazio vettoriale su $\mathbb{F}$. Siano $k \in \mathbb{N}_+$, $v_1, \dots, v_k \in V$ e $A \subseteq V$.

(a) Si dice *combinazione lineare* di $v_1, \dots, v_k$ (con coefficienti $\alpha_1, \dots, \alpha_k \in \mathbb{F}$) il vettore $\alpha_1 v_1 + \dots + \alpha_k v_k$.

(b) I vettori $v_1, \dots, v_k$ si dicono *linearmente indipendenti* se l'unica loro combinazione lineare nulla è quella con coefficienti nulli, cioè:

$$\forall \alpha_1, \dots, \alpha_k \in \mathbb{F} : (\alpha_1 v_1 + \dots + \alpha_k v_k = 0 \implies \alpha_1 = \dots = \alpha_k = 0)\,;$$

$v_1, \dots, v_k$ si dicono *linearmente dipendenti* in caso contrario.

(c) L'insieme A si dice *linearmente indipendente* se, per ogni $s \in \mathbb{N}_+$ e $w_1, \dots, w_s \in A$ distinti, i vettori $w_1, \dots, w_s$ sono linearmente indipendenti. In caso contrario, A si dice *linearmente dipendente*.

(d) Lo *spazio vettoriale generato da* A è l'insieme

$$\operatorname{span} A = \left\{ \sum_{j=1}^{s} \alpha_j w_j : s \in \mathbb{N},\, \alpha_1, \dots, \alpha_s \in \mathbb{F},\, w_1, \dots, w_s \in A \right\}$$

di tutte le combinazioni lineari di elementi di A (inclusa la combinazione vuota per $s = 0$, cioè il vettore nullo). Scriviamo anche $\operatorname{span}_{\mathbb{F}} A$ al posto di $\operatorname{span} A$ in caso si voglia esplicitare il campo $\mathbb{F}$ degli scalari.

(e) Si dice *base* di V ogni sottoinsieme $B \subseteq V$ linearmente indipendente tale che $\operatorname{span} B = V$.

Osservazione 1.10. Si può verificare che, se V è uno spazio vettoriale e $A \subseteq V$, allora $\operatorname{span} A$ è il più piccolo sottospazio vettoriale che contiene A. Inoltre, per due sottospazi vettoriali U_1, U_2 di V, la loro somma (☞ esempio 1.6(6)) si può esprimere nella forma

$$U_1 + U_2 = \operatorname{span}(U_1 \cup U_2).$$

Più in generale, se $U_1, \dots, U_n$ sono sottospazi di V, allora

$$\mathop{\Large +}_{j=1}^{n} U_j \coloneqq U_1 + \dots + U_n = \operatorname{span}(U_1 \cup \dots \cup U_n).$$

Osservazione 1.11. Siano U_1, U_2 sottospazi vettoriali di V. Se $x \in U_1 + U_2$, allora per definizione possiamo decomporre $x = x_1 + x_2$ con $x_1 \in U_1$ e $x_2 \in U_2$; se tale decomposizione è unica per ogni $x \in U_1 + U_2$, diciamo che i sottospazi U_1 e U_2 sono *in somma diretta*, e scriviamo anche $U_1 \oplus U_2$ al posto di $U_1 + U_2$. Si può verificare che

$$U_1, U_2 \text{ sono in somma diretta} \iff U_1 \cap U_2 = \{0\}.$$

In maniera simile, diciamo che i sottospazi $U_1, \dots, U_n$ sono in somma diretta se la decomposizione di ogni $x \in U_1 + \dots + U_n$ come $x = x_1 + \dots + x_n$ con $x_j \in U_j$ per $j = 1, \dots, n$ è unica. Induttivamente, si ha che

$$U_1, \dots, U_n \text{ sono in somma diretta} \iff \begin{cases} (U_1 + \dots + U_{n-1}) \cap U_n = \{0\}, \text{ e} \\ U_1, \dots, U_{n-1} \text{ sono in somma diretta.} \end{cases}$$

Osservazione 1.12. Se $B = \{e_j\}_{j\in J}$ è una base dello spazio vettoriale V (indicizzata iniettivamente, cioè $e_j \neq e_k$ per $j \neq k$), allora possiamo scrivere ogni vettore $v \in V$ in maniera unica come combinazione lineare (finita) di elementi della base: in altre parole,

$$v = \sum_{j\in J} \alpha_j e_j \tag{1.3}$$

dove i coefficienti $\alpha_j \in \mathbb{F}$ sono univocamente determinati da v, e $\alpha_j \neq 0$ per al più un numero finito di indici $j \in J$ (in modo che la somma a membro destro di (1.3) sia sempre una somma finita, una volta ignorati i termini con $\alpha_j = 0$).

Teorema 1.13 (esistenza ed equicardinalità delle basi)**.** *Sia V uno spazio vettoriale su $\mathbb{F}$. Allora:*

(i) V ha una base.
Più precisamente, dati sottoinsiemi $A \subseteq G \subseteq V$, dove A è linearmente indipendente e $\operatorname{span} G = V$, *esiste una base B di V tale che $A \subseteq B \subseteq G$. (In particolare, ogni sottoinsieme linearmente indipendente A di V si può* completare *a una base B di V.)*

(ii) Tutte le basi di V hanno lo stesso numero di elementi (cioè la stessa cardinalità).

Osservazione 1.14. Nei testi elementari di algebra lineare, il teorema precedente si trova normalmente enunciato sotto l'ipotesi aggiuntiva che V sia *finitamente generato*, cioè che esista un sottoinsieme finito $S \subseteq V$ tale che $V = \operatorname{span} S$; in tal caso, ogni base di V ha un numero finito di elementi. La dimostrazione nel caso generale (senza assumere che V sia finitamente generato) richiede tecniche dimostrative più sofisticate (lemma di Zorn).

Definizione 1.15. Sia V uno spazio vettoriale su $\mathbb{F}$. Se V ha una base con n elementi per qualche $n \in \mathbb{N}$, diciamo che V ha *dimensione* n e scriviamo $\dim V = n$. Se V non ha basi con un numero finito di elementi, diciamo che V ha *dimensione infinita* e scriviamo $\dim V = \infty$. Scriviamo anche $\dim_{\mathbb{F}} V$ al posto di $\dim V$ in caso si voglia esplicitare il campo $\mathbb{F}$ degli scalari.

Esempio 1.16. Ecco alcuni esempi di basi e dimensioni di spazi vettoriali.

(1) Per ogni $n \in \mathbb{N}$, $\dim_{\mathbb{R}} \mathbb{R}^n = n$. Una base di $\mathbb{R}^n$ è la *base canonica* $\{e_1, \dots, e_n\}$, ove $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ è la n-upla che ha un 1 in posizione j e 0 in tutte le altre componenti.

(2) Allo stesso modo si vede che, per ogni $n \in \mathbb{N}$, $\dim_{\mathbb{C}} \mathbb{C}^n = n$, e che una base di $\mathbb{C}^n$ è la base canonica $\{e_1, \dots, e_n\}$ definita come sopra.

(3) Per ogni $n \in \mathbb{N}$, $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$. Una base di $\mathbb{C}^n$ pensato come spazio vettoriale su $\mathbb{R}$ è l'insieme $\{e_1, ie_1, \dots, e_n, ie_n\}$.

(4) Per ogni $d \in \mathbb{N}$, $\dim \mathcal{P}_d = d+1$. Una base di $\mathcal{P}_d$ è l'insieme $\{1, X, \dots, X^d\}$ dei monomi monici di grado al più d. Si ha inoltre $\dim \mathcal{P} = \infty$, e una base di $\mathcal{P}$ è l'insieme $\{X^n\}_{n\in\mathbb{N}}$ di tutti i monomi monici.

(5) Se S è un insieme di n elementi per qualche $n \in \mathbb{N}$, allora $\dim \mathcal{F}(S, \mathbb{F}) = n$. Se invece S è un insieme infinito e $V \neq \{0\}$ è uno spazio vettoriale su $\mathbb{F}$, allora $\dim \mathcal{F}(S, V) = \infty$.

(6) Se $I \subseteq \mathbb{R}$ è un intervallo di lunghezza positiva, allora $\dim C_{\mathbb{F}}(I) = \infty$.

Osservazione 1.17. Se U è un sottospazio vettoriale di V, allora $\dim U \leq \dim V$: infatti una base di U si può completare a una base di V (☞ teorema 1.13). Se $\dim V < \infty$ e U è un sottospazio *proprio* (cioè $U \neq V$), allora vale anche la disuguaglianza stretta $\dim U < \dim V$. Tuttavia, se $\dim V = \infty$, allora esistono sottospazi propri U di V con la stessa dimensione di V.

Osservazione 1.18. Utilizzando la teoria dei cardinali infiniti, sarebbe possibile dare una definizione più precisa di dimensione per gli spazi vettoriali di dimensione infinita, distinguendo ad esempio fra spazi con una base di cardinalità numerabile $\aleph_0$ e spazi con una base di cardinalità del continuo $2^{\aleph_0}$. Ad esempio, potremmo scrivere che $\dim \mathcal{P} = \aleph_0$, dato che l'insieme $\{X^n\}_{n\in\mathbb{N}}$ dei monomi monici è una base numerabile di $\mathcal{P}$; invece, è possibile dimostrare che $\dim C_{\mathbb{F}}(\mathbb{R}) = 2^{\aleph_0}$. Nel seguito non ci interesseremo di queste distinzioni e scriveremo semplicemente $\dim V = \infty$ quando V ha dimensione infinita.

1.3 Applicazioni lineari

Definizione 1.19. Siano V, W spazi vettoriali su $\mathbb{F}$.

(a) Si dice *applicazione lineare* da V a W ogni funzione $T : V \to W$ tale che

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

per ogni $x, y \in V$ e $\alpha, \beta \in \mathbb{F}$; qui usiamo la notazione Tx con lo stesso significato di $T(x)$.

(b) Denotiamo con $\mathcal{L}(V, W)$ l'insieme delle applicazioni lineari da V a W; per brevità, scriviamo anche $\mathcal{L}(V)$ invece di $\mathcal{L}(V, V)$.

(c) Il *nucleo* $\operatorname{Ker} T$ di un'applicazione lineare $T \in \mathcal{L}(V, W)$ è l'insieme

$$\operatorname{Ker} T = T^{-1}(\{0\}) = \{x \in V : Tx = 0\}.$$

(d) L'*immagine* $\operatorname{Im} T$ di un'applicazione lineare $T \in \mathcal{L}(V, W)$ è l'insieme

$$\operatorname{Im} T = T(V) = \{Tx : x \in V\}.$$

A volte si scrive $\mathcal{R}(T)$ invece di $\operatorname{Im} T$.

Proposizione 1.20 (proprietà delle applicazioni lineari)**.** *Siano V, W, X spazi vettoriali su $\mathbb{F}$.*

(i) $\mathcal{L}(V, W)$ è un sottospazio vettoriale di $\mathcal{F}(V, W)$.

(ii) Per ogni $T \in \mathcal{L}(V, W)$ e $S \in \mathcal{L}(W, X)$, la loro composizione $ST := S \circ T$ è un elemento di $\mathcal{L}(V, X)$.

(iii) Se $T \in \mathcal{L}(V, W)$ è invertibile, allora $T^{-1} \in \mathcal{L}(W, V)$.

Sia ora $T \in \mathcal{L}(V, W)$. Allora:

(iv) $T0 = 0$;

(v) *per ogni sottospazio vettoriale U di V, l'immagine $T(U)$ di U è un sottospazio vettoriale di W;*

(vi) *per ogni sottospazio vettoriale Z di W, la controimmagine $T^{-1}(Z)$ di Z è un sottospazio vettoriale di V;*

(vii) *il nucleo* $\operatorname{Ker} T$ *di T è un sottospazio vettoriale di V;*

(viii) *l'immagine* $\operatorname{Im} T$ *di T è un sottospazio vettoriale di W;*

(ix) *T è iniettiva se e solo se* $\operatorname{Ker} T = \{0\}$*;*

(x) *T è suriettiva se e solo se* $\operatorname{Im} T = W$*;*

(xi) $\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V$*;*
in particolare, $\dim \operatorname{Im} T \leq \dim V$*, con uguaglianza se T è iniettiva.*

Osservazione 1.21. Nel caso $\dim V = \dim W < \infty$, dalla proposizione 1.20(xi) si deduce che, per ogni $T \in \mathcal{L}(V, W)$,

$$T \text{ è iniettiva} \iff T \text{ è suriettiva} \iff T \text{ è biiettiva.}$$

Infatti T è iniettiva se e solo se $\dim \operatorname{Ker} T = 0$; per la proposizione 1.20(xi) e l'ipotesi di dimensione finita, questo si verifica se e solo se $\dim \operatorname{Im} T = \dim V$, cioè (per l'ipotesi di uguaglianza di dimensioni), se e solo se $\dim \operatorname{Im} T = \dim W$; per l'ipotesi di dimensione finita, questo succede se e solo se $\operatorname{Im} T = W$, cioè se e solo se T è suriettiva. Si noti che in questa discussione l'ipotesi di dimensione finita è essenziale (☞ osservazione 11.22).

Proposizione 1.22 (applicazioni lineari e basi)**.** *Siano V, W spazi vettoriali su $\mathbb{F}$. Sia B una base di V. Allora, per ogni funzione $f : B \to W$, esiste un'unica applicazione lineare $T \in \mathcal{L}(V, W)$ tale che $Tx = f(x)$ per ogni $x \in B$.*

In altre parole, possiamo costruire un'applicazione lineare da V a W assegnando liberamente i suoi valori su una base di V, e questi valori determinano univocamente l'applicazione lineare.

Osservazione 1.23. Se V è uno spazio vettoriale di dimensione $n \in \mathbb{N}$ su $\mathbb{F}$ e $B = \{v_1, \dots, v_n\}$ è una sua base, possiamo definire una mappa lineare biiettiva $\Phi_B : \mathbb{F}^n \to V$ ponendo

$$\Phi_B(x_1, \dots, x_n) = x_1 v_1 + \dots + x_n v_n \tag{1.4}$$

(☞ osservazione 1.12). L'inversa $\Phi_B^{-1} : V \to \mathbb{F}^n$ è l'unica mappa lineare che manda gli elementi della base $\{v_1, \dots, v_n\}$ negli elementi della base canonica $\{e_1, \dots, e_n\}$ di $\mathbb{F}^n$ (☞ esempio 1.16). La scelta di una base di V corrisponde dunque a una scelta di *coordinate lineari* su V, cioè ci permette di identificare V a $\mathbb{F}^n$ dove $n = \dim V$.

Osservazione 1.24. Siano V, W spazi vettoriali di dimensioni $n, m \in \mathbb{N}$ rispettivamente, e siano $B = \{v_1, \dots, v_n\}$ e $C = \{w_1, \dots, w_m\}$ basi di V e W. Allora ogni mappa lineare $T \in \mathcal{L}(V, W)$ è univocamente determinata dai suoi valori $Tv_1, \dots, Tv_k$ sulla base B di V (☞ proposizione 1.22). Se scriviamo questi valori in termini della base C di W (☞ osservazione 1.12), cioè

$$Tv_k = a_{1,k} w_1 + \dots + a_{m,k} w_m, \qquad k = 1, \dots, n,$$

la matrice $A = (a_{j,k})_{\substack{j=1,\dots,m\\k=1,\dots,n}} \in \mathbb{F}^{m\times n}$ dei coefficienti di queste combinazioni lineari determina univocamente T. Diciamo A la *matrice associata* a T nelle basi B e C; da (1.4) si vede facilmente che allora

$$T\Phi_B x = \Phi_C A x \qquad \forall x \in \mathbb{F}^n,$$

dove gli elementi $x \in \mathbb{F}^n$ sono pensati come vettori colonna e Ax è il prodotto matrice-vettore. Fissate basi B di V e C di W, abbiamo dunque una corrispondenza biunivoca fra mappe lineari $T \in \mathcal{L}(V,W)$ e matrici $A \in \mathbb{F}^{m\times n}$; si può verificare che tale corrispondenza è lineare (cioè manda combinazioni lineari di mappe nelle combinazioni lineari delle rispettive matrici), e inoltre manda la composizione di mappe nel prodotto di matrici. In particolare,

$$\dim \mathcal{L}(V,W) = \dim \mathbb{F}^{n\times m} = nm = \dim V \dim W$$

per spazi di dimensione finita V e W.

2 Richiami di topologia in spazi metrici

In questo capitolo si richiamano alcune definizioni e diversi risultati fondamentali della teoria degli spazi metrici, che sono alla base della teoria sviluppata in seguito. La maggior parte delle dimostrazioni è omessa.

Come sempre, con $\mathbb{F}$ denotiamo il campo dei numeri reali $\mathbb{R}$ oppure il campo dei numeri complessi $\mathbb{C}$.

2.1 Spazi metrici

Definizione 2.1. Si dice *spazio metrico* un insieme M dotato di una funzione $d : M \times M \to [0, \infty)$ tale che:

1. $\forall x, y \in M : d(x, y) = 0 \iff x = y$
   (proprietà di separazione);
2. $\forall x, y \in M : d(x, y) = d(y, x)$
   (simmetria);
3. $\forall x, y, z \in M : d(x, z) \leq d(x, y) + d(y, z)$
   (disuguaglianza triangolare).

Una tale funzione d è detta *distanza* o *metrica* su M. Gli elementi di M sono anche detti *punti* di M.

Notazione. Nel seguito scriveremo "lo spazio metrico (M, d)" quando vogliamo specificare simultaneamente l'insieme M e la metrica d su M.

Esempio 2.2. Ecco alcuni esempi di distanze e spazi metrici.

(1) Sia $n \in \mathbb{N}_+$. L'insieme $\mathbb{F}^n$ si può dotare di diverse metriche. Per ogni $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{F}^n$, definiamo:

- $d_2(x, y) = \sqrt{\sum_{j=1}^n |x_j - y_j|^2}$
  (metrica euclidea);
- $d_1(x, y) = \sum_{j=1}^n |x_j - y_j|$
  (metrica di Manhattan);
- $d_\infty(x, y) = \max\{|x_j - y_j| : j = 1, \ldots, n\}$
  (metrica della scacchiera);
- più in generale, per $p \in [1, \infty)$, si può definire la metrica
  $d_p(x, y) = \left(\sum_{j=1}^n |x_j - y_j|^p\right)^{1/p}$.

Si noti che, nel caso $n = 1$, tutte le metriche d_p su $\mathbb{F}$ coincidono con la metrica euclidea $d_2(x, y) = |x - y|$.

Il fatto che d_p sia una metrica per ogni $p \in [1, \infty]$ non è completamente banale da dimostrare; in particolare, la dimostrazione della disuguaglianza triangolare richiede un po' di lavoro quando $p \neq 1, \infty$. Discuteremo in seguito in dettaglio la dimostrazione di una variante in "dimensione infinita" di questa disuguaglianza, nel contesto degli spazi di successioni (☞ sezione 5.1).

(2) Sia $[a, b]$ un intervallo chiuso e limitato in $\mathbb{R}$. L'insieme $C_{\mathbb{F}}[a, b]$ si può dotare della *metrica dell'estremo superiore*, detta anche *metrica della convergenza uniforme*, definita da

$$d_\infty(f, g) = \sup_{t \in [a,b]} |f(t) - g(t)| \tag{2.1}$$

per ogni $f, g \in C_{\mathbb{F}}[a, b]$. Si noti che, per il teorema di Weierstrass, in realtà vale anche

$$d_\infty(f, g) = \max_{t \in [a,b]} |f(t) - g(t)|.$$

(3) Se (M, d) è uno spazio metrico ed E è un sottoinsieme di M, allora la restrizione di d a $E \times E$ è una metrica su E (detta *metrica indotta* su E da (M, d)), che per brevità denotiamo ancora con d; dunque (E, d) è uno spazio metrico a sua volta.

(4) Se (M, d_M) e (N, d_N) sono spazi metrici, possiamo definire una metrica d sul prodotto cartesiano $\Omega = M \times N$ ponendo

$$d((x, y), (x', y')) = \max\{d_M(x, x'), d_N(y, y')\} \qquad \forall x, x' \in M,\ y, y' \in N.$$

Lo spazio (Ω, d) è detto *spazio metrico prodotto* degli spazi (M, d_M) e (N, d_N). Nel caso $(M, d_M) = (N, d_N) = (\mathbb{R}, d_2)$, lo spazio metrico prodotto non è altro che $\mathbb{R}^2$ con la metrica d_∞ discussa sopra.

2.2 Successioni e completezza

Definizione 2.3. Sia X un insieme. Una *successione* a valori in X è una funzione $s : \mathbb{N} \to X$. Per ogni $k \in \mathbb{N}$, il valore $s(k)$ è detto il *termine k-esimo* della successione s. Se $s(k) = x_k$ per ogni $k \in \mathbb{N}$, scriviamo anche $(x_k)_{k\in\mathbb{N}}$ oppure $(x_k)_{k=0}^\infty$ oppure $(x_0, x_1, x_2, \dots)$ al posto di s; per brevità a volte si scrive anche $(x_k)_k$ omettendo il dominio $\mathbb{N}$ della successione.

Osservazione 2.4. Alcuni testi utilizzano la notazione $\{x_k\}_{k\in\mathbb{N}}$ con parentesi graffe per denotare la successione $(x_k)_{k\in\mathbb{N}}$. Qui preferiamo una notazione diversa, in modo da distinguere fra la successione $(x_k)_{k\in\mathbb{N}}$, che è una funzione $S \to X$, e l'insieme dei suoi valori $\{x_k\}_{k\in\mathbb{N}}$, che è un sottoinsieme di X. Ad esempio, le successioni

$$((-1)^k)_{k\in\mathbb{N}} = (1, -1, 1, -1, \dots), \qquad ((-1)^{k+1})_{k\in\mathbb{N}} = (-1, 1, -1, 1, \dots)$$

sono due successioni diverse con lo stesso insieme di valori

$$\{(-1)^n\}_{k\in\mathbb{N}} = \{(-1)^{k+1}\}_{k\in\mathbb{N}} = \{-1, 1\}.$$

Osservazione 2.5. A volte può essere comodo considerare "successioni" $(x_k)_{k=1}^\infty$ indicizzate sull'insieme $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ anziché su $\mathbb{N}$. Più in generale si possono considerare "successioni" $(x_k)_{k=k_0}^\infty$ indicizzate sull'insieme $\{k \in \mathbb{Z} : k \geq k_0\} = \{k_0, k_0 + 1, k_0 + 2, \dots\}$ per qualche $k_0 \in \mathbb{Z}$. *Mutatis mutandis*, la discussione che segue si applica anche a queste successioni più generali.

Definizione 2.6. Sia X un insieme e sia $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X. Una *sottosuccessione* di $(x_n)_{n\in\mathbb{N}}$ è una successione della forma $(x_{\sigma(k)})_{k\in\mathbb{N}}$ per qualche $\sigma : \mathbb{N} \to \mathbb{N}$ strettamente crescente.

Definizione 2.7. Sia (M, d) uno spazio metrico. Sia $(x_n)_{n\in\mathbb{N}}$ una successione a valori in M e sia $x \in M$.

(a) Diciamo che la successione $(x_n)_{n\in\mathbb{N}}$ *tende* a x (o anche che *converge* a x), e scriviamo

$$x_n \xrightarrow[n\to\infty]{} x \qquad \text{oppure} \qquad \lim_{n\to\infty} x_n = x$$

se per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che $d(x_n, x) < \epsilon$ per ogni $n > N$. In tal caso, il punto x si dice *limite* della successione $(x_n)_{n\in\mathbb{N}}$.

(b) Diciamo che la successione $(x_n)_{n\in\mathbb{N}}$ è *di Cauchy* se per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che $d(x_n, x_m) < \epsilon$ per ogni $n, m > N$.

Osservazione 2.8. Siano (M, d) uno spazio metrico, $(x_n)_n$ una successione a valori in M e $x \in M$. Dalla definizione di limite segue subito che

$$x_n \xrightarrow[n\to\infty]{} x \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} d(x_n, x) = 0.$$

Notazione. Al posto di

$$x_n \xrightarrow[n\to\infty]{} x$$

scriviamo anche "$x_n \to x$ per $n \to \infty$", oppure semplicemente "$x_n \to x$", se non c'è ambiguità.

Proposizione 2.9. *Sia (M, d) uno spazio metrico. Siano $(x_n)_{n\in\mathbb{N}}$ una successione a valori in M e $x \in M$.*

(i) Il limite di $(x_n)_n$, se esiste, è unico.

(ii) Se $(x_n)_n$ converge a x, allora ogni sottosuccessione di $(x_n)_n$ converge a x.

(iii) Se $(x_n)_n$ converge, allora $(x_n)_n$ è di Cauchy.

Osservazione 2.10. In un arbitrario spazio metrico (M, d), non è detto che una successione di Cauchy $(x_n)_n$ converga. Ad esempio, la successione $(1/(1+n))_{n\in\mathbb{N}}$ nello spazio metrico $(\mathbb{R} \setminus \{0\}, d_2)$ è di Cauchy, ma non converge ad un punto di $\mathbb{R} \setminus \{0\}$.

Definizione 2.11. Uno spazio metrico (M, d) si dice *completo* se ogni successione di Cauchy a valori in M converge a un punto di M.

Esempio 2.12. Ecco alcuni esempi di spazi metrici completi e non completi.

(1) $(\mathbb{Q}, d_2)$ non è completo. (Ad esempio, la successione a valori in $\mathbb{Q}$ data da $(\lfloor 10^n\sqrt{2}\rfloor/10^n)_{n\in\mathbb{N}}$ è di Cauchy ma non converge in $\mathbb{Q}$.)

(2) Per ogni $n \in \mathbb{N}_+$, $(\mathbb{F}^n, d_2)$ è completo, ove $\mathbb{F} = \mathbb{R}$ oppure $\mathbb{F} = \mathbb{C}$.

(3) Se $[a, b]$ è un intervallo chiuso e limitato di $\mathbb{R}$, allora $(C_{\mathbb{F}}[a, b], d_\infty)$ è completo; questo fatto è discusso in dettaglio in seguito (☞ teorema 2.14).

La dimostrazione della completezza di $\mathbb{F}^n$ (☞ esempio 2.12(2)) si può ricondurre, ragionando componente per componente, a quella di $\mathbb{F}$, e a sua volta il caso di $\mathbb{F} = \mathbb{C}$ si riconduce a quello di $\mathbb{F} = \mathbb{R}$ identificando $\mathbb{C}$ con $\mathbb{R}^2$; infine, la completezza di $(\mathbb{R}, d_2)$ si può dedurre dal teorema di Bolzano–Weierstrass.

La stessa idea del "ragionare componente per componente" è codificata nel seguente risultato, relativo alle successioni negli spazi metrici prodotto discussi nell'esempio 2.2(4).

Proposizione 2.13. *Siano (M, d_M) e (N, d_N) spazi metrici. Sia $(M \times N, d)$ lo spazio metrico prodotto.*

(i) Una successione $((x_n, y_n))_{n\in\mathbb{N}}$ a valori in $M \times N$ converge a $(x, y) \in M \times N$ se e solo se $x_n \to x$ in M e $y_n \to y$ in N.

(ii) Una successione $((x_n, y_n))_{n\in\mathbb{N}}$ a valori in $M \times N$ è di Cauchy se e solo se $(x_n)_{n\in\mathbb{N}}$ è di Cauchy in M e $(y_n)_{n\in\mathbb{N}}$ è di Cauchy in N.

(iii) Se M e N sono completi, anche il prodotto $M \times N$ è completo.

Discutiamo ora l'esempio 2.12(3) in dettaglio.

Teorema 2.14. *Sia $[a, b]$ un intervallo chiuso e limitato di $\mathbb{R}$. Allora lo spazio metrico $(C_{\mathbb{F}}[a, b], d_\infty)$ è completo.*

Dimostrazione. Sia $(f_n)_{n\in\mathbb{N}}$ una successione di Cauchy in $C_{\mathbb{F}}[a, b]$.
Passo 1. Per ogni $t \in [a, b]$, la successione $(f_n(t))_{n\in\mathbb{N}}$ è di Cauchy in $\mathbb{F}$.

Sappiamo per ipotesi che $(f_n)_{n\in\mathbb{N}}$ è di Cauchy in $C_{\mathbb{F}}[a, b]$, cioè

$$\forall\epsilon > 0 : \exists N \in \mathbb{N} : \forall n, m > N : d_\infty(f_n, f_m) < \epsilon. \tag{2.2}$$

D'altra parte, per ogni $t \in [a, b]$, da (2.1) deduciamo la disuguaglianza

$$|f_n(t) - f_m(t)| \le d_\infty(f_n, f_m).$$

Pertanto, possiamo rimpiazzare $d_\infty(f_n, f_m)$ in (2.2) a maggior ragione con la quantità $|f_n(t) - f_m(t)|$, deducendo che

$$\forall\epsilon > 0 : \exists N \in \mathbb{N} : \forall n, m > N : |f_n(t) - f_m(t)| < \epsilon,$$

cioè $(f_n(t))_n$ è una successione di Cauchy in $\mathbb{F}$.
Passo 2. Siccome $\mathbb{F}$ è completo, per ogni $t \in [a, b]$ la successione $(f_n(t))_n$ converge e possiamo definire

$$f(t) = \lim_{n\to\infty} f_n(t) \in \mathbb{F}.$$

In questo modo definiamo una funzione $f : [a, b] \to \mathbb{F}$.
Passo 3. $d_\infty(f_n, f) \to 0$ per $n \to \infty$.

Sia $\epsilon > 0$. Siccome $(f_n)_n$ è di Cauchy in $C_{\mathbb{F}}[a, b]$, esiste $N \in \mathbb{N}$ tale che, per ogni $t \in [a, b]$,

$$|f_n(t) - f_m(t)| \le d_\infty(f_n, f_n) < \epsilon/2$$

per ogni $n, m > N$. Passando al limite per $m \to \infty$, si ottiene

$$|f_n(t) - f(t)| \le \epsilon/2$$

per ogni $t \in [a, b]$ e $n > N$; quindi, prendendo l'estremo superiore su $t \in [a, b]$, deduciamo che

$$d_\infty(f_n, f) = \sup_{t\in[a,b]} |f_n(t) - f(t)| \le \epsilon/2 < \epsilon$$

per ogni $n > N$.
Passo 4. $f : [a, b] \to \mathbb{F}$ è continua.

Sia $t \in [a,b]$. Sia $\epsilon > 0$. Per il passo 3, troviamo $N \in \mathbb{N}$ tale che $d_\infty(f_n, f) < \epsilon/3$ per ogni $n > N$. Siccome f_{N+1} è continua in t, troviamo $\delta > 0$ tale che $|f_{N+1}(t) - f_{N+1}(t')| < \epsilon/3$ per ogni $t' \in [a,b]$ con $|t - t'| < \delta$. In conclusione, per ogni $t' \in [a,b]$, se $|t - t'| < \delta$, allora

$$\begin{aligned} |f(t') - f(t)| &\leq |f(t') - f_{N+1}(t')| + |f_{N+1}(t') - f_{N+1}(t)| + |f_{N+1}(t) - f(t)| \\ &\leq 2d_\infty(f, f_{N+1}) + |f_{N+1}(t') - f_{N+1}(t)| \\ &< 2\epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

Questo dimostra che f è continua in t. □

L'argomento utilizzato nella dimostrazione del teorema 2.14 sarà utilizzato in seguito in molteplici varianti, dunque è utile fare ora alcune puntualizzazioni.

Anzitutto, la continuità delle funzioni f_n nella dimostrazione gioca un ruolo solo nel passo 4, dove si vuole dimostrare la continuità del limite f. Invece, i primi tre passi hanno senso anche per funzioni f_n più generali; in questa prima parte della dimostrazione, inoltre, il fatto che il dominio delle f_n sia un intervallo $[a,b]$ non ha alcun ruolo, e potrebbe essere sostituito da un altro insieme.

Più specificamente, dato un insieme S, possiamo definire, per due funzioni $f, g : S \to \mathbb{F}$,

$$d_\infty(f,g) := \sup_{t \in S} |f(t) - g(t)|. \tag{2.3}$$

Non è difficile verificare che, con questa definizione, per ogni $f, g, h \in \mathcal{F}(S, \mathbb{F})$,

$$\begin{gathered} d_\infty(f,g) = 0 \iff f = g, \\ d_\infty(f,g) = d_\infty(g,f), \\ d_\infty(f,h) \leq d_\infty(f,g) + d_\infty(g,h). \end{gathered} \tag{2.4}$$

Tuttavia, in generale d_∞ non è una distanza sull'insieme $\mathcal{F}(S, \mathbb{F})$ nel senso della definizione 2.1, perché può succedere che $d_\infty(f,g) = \infty$. Per evitare questo problema, ci si può restringere al sottoinsieme $\mathcal{F}_b(S, \mathbb{F})$ delle *funzioni limitate* su S, cioè

$$\mathcal{F}_b(S, \mathbb{F}) = \left\{ f \in \mathcal{F}(S, \mathbb{F}) : \sup_{t \in S} |f(t)| < \infty \right\}. \tag{2.5}$$

Allora $d_\infty(f,g) < \infty$ per ogni $f, g \in \mathcal{F}_b(S, \mathbb{F})$, e le proprietà (2.4) ci dicono che $(\mathcal{F}_b(S, \mathbb{F}), d_\infty)$ è uno spazio metrico.

Essenzialmente ripetendo, *mutatis mutandis*, i primi tre passi della dimostrazione del teorema 2.14, si ottiene allora il seguente risultato.

Teorema 2.15. *Sia S un insieme. Allora lo spazio metrico $(\mathcal{F}_b(S, \mathbb{F}), d_\infty)$ è completo.*

Notiamo che, per il teorema di Weierstrass, si ha $C_\mathbb{F}[a,b] \subseteq \mathcal{F}_b([a,b], \mathbb{F})$ per ogni intervallo $[a,b]$ chiuso e limitato di $\mathbb{R}$. Il passo 4 della dimostrazione del teorema 2.14 si può dunque interpretare come un risultato di "chiusura" del sottoinsieme $C_\mathbb{F}[a,b]$ di $\mathcal{F}_b([a,b], \mathbb{F})$: il limite in $\mathcal{F}_b([a,b], \mathbb{F})$ di una successione a valori in $C_\mathbb{F}[a,b]$ è ancora un elemento di $C_\mathbb{F}[a,b]$. Come discutiamo nella prossima sezione, questo risultato è una manifestazione del fatto che la *convergenza uniforme* di funzioni preserva proprietà come la continuità.

2.3 Convergenza puntuale e convergenza uniforme

Definizione 2.16. Sia S un insieme. Siano $f_n : S \to \mathbb{F}$ per ogni $n \in \mathbb{N}$, e $f : S \to \mathbb{F}$.

(a) Diciamo che la successione di funzioni $(f_n)_{n\in\mathbb{N}}$ *converge puntualmente* a f, e scriviamo $f_n \to f$ puntualmente, se

$$\lim_{n\to\infty} f_n(t) = f(t) \qquad \forall t \in S.$$

(b) Diciamo che la successione di funzioni $(f_n)_{n\in\mathbb{N}}$ *converge uniformemente* a f, e scriviamo $f_n \to f$ uniformemente, o anche $f_n \rightrightarrows f$, se

$$\lim_{n\to\infty} d_\infty(f_n, f) = 0,$$

ove $d_\infty(f, g)$ è definita in (2.3).

Osservazione 2.17. Esplicitando le definizioni di limite, si vede che $f_n \to f$ puntualmente se e solo se

$$\forall t \in S : \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n > N : |f_n(t) - f(t)| < \epsilon;$$

invece, $f_n \to f$ uniformemente se e solo se

$$\forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n > N : \forall t \in S : |f_n(t) - f(t)| < \epsilon.$$

La differenza fra le due condizioni è che, nella convergenza puntuale, il valore di N può dipendere sia da ϵ che dal punto t, mentre nella convergenza uniforme si può trovare un N che dipende solo da ϵ ma non da t (cioè si ha un controllo uniforme in t della rapidità di convergenza).

Osservazione 2.18. Dalla precedente osservazione si vede che la convergenza uniforme implica la convergenza puntuale. Tuttavia l'implicazione opposta in generale non vale. Ad esempio, siano $f_n : [0, 1] \to \mathbb{R}$ definite da $f_n(t) = t^n$ per ogni $n \in \mathbb{N}$ e $t \in [0, 1]$. Allora $f_n \to f$ puntualmente, ove

$$f(t) = \begin{cases} 0 & \text{se } t \in [0, 1), \\ 1 & \text{se } t = 1. \end{cases}$$

Tuttavia la convergenza $f_n \to f$ non è uniforme, dato che

$$d_\infty(f_n, f) = \sup_{t\in[0,1]} |f_n(t) - f(t)| \geq \sup_{t\in[0,1)} |t^n - 0| = \sup_{t\in[0,1)} t^n = 1$$

e quindi $d_\infty(f_n, f) \not\to 0$ per $n \to \infty$.

La convergenza uniforme di funzioni è utile in quanto permette di "scambiare l'ordine dei limiti" in molte costruzioni dell'analisi. Ad esempio, valgono le seguenti proprietà.

Proposizione 2.19. *Sia Ω uno spazio topologico. Sia $C(\Omega)$ l'insieme delle funzioni continue $f : \Omega \to \mathbb{F}$ (☞ esempio 1.2(3)).*

(i) Se $(f_n)_{n\in\mathbb{N}}$ è una successione in $C(\Omega)$ e $f_n \rightrightarrows f$, allora $f \in C(\Omega)$.

Supponiamo ora che $I \subseteq \mathbb{R}$ sia un intervallo. Sia $C^1(I)$ l'insieme delle funzioni $f : I \to \mathbb{F}$ derivabili con derivata continua.

(ii) *Se $(f_n)_{n\in\mathbb{N}}$ è una successione in $C^1(I)$ tale che $f_n \rightrightarrows f$ e $f_n' \rightrightarrows g$, allora $f \in C^1(I)$ e $f' = g$.*

Infine, sia $[a, b]$ un intervallo chiuso e limitato in $\mathbb{R}$.

(iii) *Se $(f_n)_{n\in\mathbb{N}}$ è una successione in $C[a, b]$ e $f_n \rightrightarrows f$, allora*

$$\lim_{n\to\infty} \int_a^b f_n(t)\, dt = \int_a^b f(t)\, dt.$$

Osservazione 2.20. La dimostrazione del punto (i) della precedente proposizione è una variante del passo 4 della dimostrazione del teorema 2.14. Per quanto riguarda il punto (iii), utiilizzando la teoria dell'integrazione secondo Lebesgue è possibile indebolire l'ipotesi di convergenza uniforme delle funzioni mantenendo la convergenza degli integrali (si veda il teorema della convergenza dominata, ☞ teorema 3.7(vi)).

2.4 Topologia in spazi metrici

Definizione 2.21. Sia (M, d) uno spazio metrico.

(a) Siano $x \in M$ e $r > 0$. La *palla aperta* e la *palla chiusa* di centro x e raggio r sono gli insiemi

$$B(x, r) = \{y \in M : d(x, y) < r\}, \qquad \overline{B}(x, r) = \{y \in M : d(x, y) \leq r\}.$$

Utilizziamo anche le notazioni $B_d(x, r)$ e $B_M(x, r)$ al posto di $B(x, r)$ quando sia utile specificare la distanza d o lo spazio metrico M.

(b) Un sottoinsieme E di M si dice *limitato* se esistono $x_0 \in M$ e $r > 0$ tali che $E \subseteq B(x_0, r)$; in caso contrario, E si dice *illimitato*.

(c) Una successione $(x_n)_{n\in\mathbb{N}}$ a valori in M si dice *limitata* o *illimitata* a seconda che l'insieme $\{x_n : n \in \mathbb{N}\}$ dei suoi valori sia un sottoinsieme limitato o illimitato di M.

(d) Un sottoinsieme U di M si dice *intorno* di un punto $x \in M$ se esiste $r > 0$ tale che $B(x, r) \subseteq U$.

(e) Un sottoinsieme A di M si dice *aperto* se è A è intorno di ogni suo punto; in altre parole, se, per ogni $x \in A$, esiste $r > 0$ tale che $B(x, r) \subseteq A$.

(f) Un sottoinsieme C di M si dice *chiuso* se $M \setminus C$ è aperto.

(g) Sia $E \subseteq M$. La *chiusura* di E è l'insieme

$$\overline{E} = \bigcap \{C \subseteq M : C \text{ chiuso}, E \subseteq C\}$$

(l'intersezione di tutti i chiusi contenenti E). Scriviamo anche $\overline{E}^M$ o $\overline{E}^{(M,d)}$ al posto di $\overline{E}$ quando sia utile specificare lo spazio metrico M ambiente.

(h) Sia $E \subseteq M$. La *parte interna* di E è l'insieme

$$\mathring{E} = \bigcup \{A \subseteq M : A \text{ aperto}, A \subseteq E\}$$

(l'unione di tutti gli aperti contenuti in E).

(i) Sia $E \subseteq M$. La *frontiera* di E è l'insieme $\partial E = \overline{E} \setminus \mathring{E}$.

(j) Un sottoinsieme E di M si dice *denso* in M se $\overline{E} = M$.

(k) Un sottoinsieme K di M si dice *compatto* se ogni successione $(x_n)_{n\in\mathbb{N}}$ a valori in K ha una sottosuccessione che converge a un punto di K.

(l) Un sottoinsieme E di M si dice *relativamente compatto* se $\overline{E}$ è compatto.

Osservazione 2.22. La definizione di successione limitata/illimitata data sopra (☞ definizione 2.21(c)) è coerente con quella usuale per successioni a valori reali, discussa nei primi corsi di analisi. Si presti attenzione al fatto che, nonostante la somiglianza dei termini usati, "essere limitata" e "avere limite" sono due proprietà diverse per una successione: ad esempio, la successione $((-1)^n)_n$ è limitata in $\mathbb{R}$, ma $\lim_{n\to\infty}(-1)^n$ non esiste.

Richiamiamo, senza dimostrazione, una serie di risultati fondamentali di topologia degli spazi metrici, che useremo diffusamente nel seguito.

Proposizione 2.23. *Sia (M, d) uno spazio metrico.*

(i) Le palle aperte sono sottoinsiemi aperti di M e le palle chiuse sono sottoinsiemi chiusi di M.

(ii) La famiglia dei sottoinsiemi aperti di M è una topologia *su M, cioè ha le seguenti proprietà:*

(a) M e $\emptyset$ sono sottoinsiemi aperti di M;

(b) se $\mathcal{A}$ è una famiglia di sottoinsiemi aperti di M, allora la loro unione $\bigcup \mathcal{A}$ è un sottoinsieme aperto di M;

(c) se A e B sono sottoinsiemi aperti di M, allora $A \cap B$ è un sottoinsieme aperto di M.

(iii) Analogamente, la famiglia dei sottoinsiemi chiusi di M ha le seguenti proprietà:

(a) M e $\emptyset$ sono sottoinsiemi chiusi di M;

(b) se $\mathcal{C}$ è una famiglia non vuota di sottoinsiemi chiusi di M, allora la loro intersezione $\bigcap \mathcal{C}$ è un sottoinsieme chiuso di M;

(c) se C e D sono sottoinsiemi chiusi di M, allora $C \cup D$ è un sottoinsieme chiuso di M.

(iv) Una successione $(x_n)_{n\in\mathbb{N}}$ a valori in M converge a un punto $x \in M$ se e solo se, per ogni sottoinsieme aperto A di M contenente x, esiste $N \in \mathbb{N}$ tale che $x_n \in A$ per ogni $n > N$.

(v) Ogni successione di Cauchy a valori in M è una successione limitata. In particolare, ogni successione convergente è limitata.

(vi) *Un sottoinsieme C di M è chiuso se e solo se, per ogni successione $(x_n)_{n\in\mathbb{N}}$ a valori in C, se $x_n \to x$ in M allora $x \in C$.*

(vii) *Sia $E \subseteq M$. La chiusura $\overline{E}$ di E è il più piccolo sottoinsieme chiuso di M che contiene E. In particolare, E è chiuso se e solo se $\overline{E} = E$. Inoltre*

$$\begin{aligned}\overline{E} &= \{x \in M : x = \lim_{n\to\infty} x_n \text{ per qualche } (x_n)_n \text{ a valori in } E\} \\ &= \{x \in M : d(x, E) = 0\},\end{aligned}$$

ove $d(x, E)$ è la distanza del punto x dall'insieme E, *definita da*

$$d(x, E) \coloneqq \inf\{d(x, y) : y \in E\}. \tag{2.6}$$

(viii) *Sia $E \subseteq M$. La parte interna $\mathring{E}$ di E è il più grande sottoinsieme aperto di M contenuto in E. In particolare, E è aperto se e solo se $E = \mathring{E}$. Inoltre*

$$M \setminus \mathring{E} = \overline{M \setminus E}, \qquad M \setminus \overline{E} = (M \setminus E)^{\circ}.$$

(ix) *Sia $E \subseteq M$. Allora E è denso in M se e solo se, per ogni $x \in M$, esiste una successione $(x_n)_n$ a valori in E tale che $x_n \to x$.*

(x) *Sia $K \subseteq M$. Allora K è compatto se e solo se ogni* ricoprimento aperto *di K (cioè una famiglia di aperti la cui unione contiene K) ha un sottoricoprimento finito.*

(xi) *Se $K \subseteq M$ è compatto, allora K è chiuso e limitato.*

L'implicazione inversa nell'ultimo enunciato non vale in generale. Tuttavia vale negli spazi euclidei, ed è legata, come la proprietà di completezza, al teorema di Bolzano–Weierstrass.

Teorema 2.24 (Heine–Borel)**.** *Sia $\mathbb{F} = \mathbb{R}$ oppure $\mathbb{F} = \mathbb{C}$. Nello spazio metrico $(\mathbb{F}^n, d_2)$, un sottoinsieme di $\mathbb{F}^n$ è compatto se e solo se è chiuso e limitato.*

Definizione 2.25. Uno spazio metrico (M, d) si dice *separabile* se M ha un sottoinsieme denso al più numerabile.

Esempio 2.26. Ecco alcuni esempi di spazi metrici separabili.

(1) $(\mathbb{R}^n, d_2)$ è separabile: il sottoinsieme $\mathbb{Q}^n$ è denso e numerabile.

(2) $(\mathbb{C}^n, d_2)$ è separabile: il sottoinsieme $(\mathbb{Q} + i\mathbb{Q})^n$ è denso e numerabile, ove $\mathbb{Q} + i\mathbb{Q} = \{x + iy : x, y \in \mathbb{Q}\}$.

(3) Se $[a, b] \subseteq \mathbb{R}$ è un intervallo chiuso e limitato, allora $(C_{\mathbb{F}}[a, b], d_\infty)$ è separabile: questa è una conseguenza del teorema di Stone–Weierstrass, discusso nel seguito (☞ teorema 4.30).

(4) Se (M, d) è uno spazio metrico separabile e $E \subseteq M$, allora anche (E, d) è separabile. Infatti, se $\{x_n\}_{n\in\mathbb{N}}$ è un denso numerabile in (M, d), per ogni $n, k \in \mathbb{N}$ tali che $d(E, x_n) < 2^{-k}$, scegliamo $y_{n,k} \in E \cap B(x_n, 2^{-k})$; non è difficile verificare che l'insieme $\{y_{n,k} : n, k \in \mathbb{N},\, d(E, x_n) < 2^{-k}\}$ è allora un denso al più numerabile in (E, d).

(5) Ogni spazio metrico compatto M è separabile. Infatti, per ogni $n \in \mathbb{N}$, esiste (☞ proposizione 2.23(x)) un ricoprimento finito di M fatto di palle di raggio 2^{-n}; l'insieme E_n dei centri di tali palle è finito e ogni punto di M è a distanza al più 2^{-n} da un punto di E_n. L'unione $E = \bigcup_{n \in \mathbb{N}} E_n$ è allora un sottoinsieme denso al più numerabile di M.

2.5 Funzioni continue, uniformemente continue e lipschitziane

Definizione 2.27. Siano (M, d_M) e (N, d_N) spazi metrici. Una funzione $f : M \to N$ si dice:

(a) *continua in un punto* $x \in M$ se, per ogni $\epsilon > 0$, esiste $\delta > 0$, tale che $d_N(f(x), f(x')) < \epsilon$ per ogni $x' \in B_M(x, \delta)$;

(b) *continua* se f è continua in ogni punto $x \in M$;

(c) *uniformemente continua* se, per ogni $\epsilon > 0$, esiste $\delta > 0$ tale che si abbia $d_N(f(x), f(x')) < \epsilon$ per ogni $x, x' \in M$ con $d_M(x, x') < \delta$;

(d) un *omeomorfismo* se f è continua e invertibile, e l'inversa $f^{-1} : N \to M$ è a sua volta continua.

Della continuità è possibile dare una caratterizzazione topologica.

Proposizione 2.28 (continuità e aperti)**.** *Siano (M, d_M) e (N, d_N) spazi metrici e $f : M \to N$. Sono equivalenti:*

(i) $f : M \to N$ è continua;

(ii) $f^{-1}(A)$ è aperto in M per ogni sottoinsieme aperto A di N;

(iii) $f^{-1}(C)$ è chiuso in M per ogni sottoinsieme chiuso C di N.

Inoltre la continuità può essere caratterizzata in termini di successioni.

Proposizione 2.29 (continuità e successioni)**.** *Siano (M, d_M) e (N, d_N) spazi metrici, $f : M \to N$ e $x \in M$. Sono equivalenti:*

(i) f è continua nel punto x;

(ii) per ogni successione $(x_n)_n$ a valori in M, se $x_n \to x$ in M, allora $f(x_n) \to f(x)$ in N.

C'è anche un'utile relazione fra funzioni uniformemente continue e successioni di Cauchy.

Proposizione 2.30 (unif. continuità e successioni di Cauchy)**.** *Siano (M, d_M) e (N, d_N) spazi metrici. Sia $f : M \to N$ uniformemente continua. Se $(x_n)_n$ è una successione di Cauchy in M, allora $(f(x_n))_n$ è una successione di Cauchy in N.*

Dimostrazione. Sia $\epsilon > 0$. Siccome f è uniformemente continua, esiste $\delta > 0$ tale che, per ogni $x, x' \in M$,

$$d_N(f(x), f(x')) < \epsilon \qquad \text{ogniqualvolta} \qquad d_M(x, x') < \delta. \tag{2.7}$$

Siccome $(x_n)_n$ è di Cauchy in M e $\delta > 0$, applicando la definizione di successione di Cauchy (definizione 2.7) con δ al posto di ϵ, esiste $N \in \mathbb{N}$ tale che

$$d_M(x_n, x_m) < \delta \qquad \text{per ogni } n, m > N. \tag{2.8}$$

Ma allora, combinando (2.7) e (2.8), per ogni $n, m > N$ si ha anche

$$d_N(f(x_n), f(x_m)) < \epsilon. \tag{2.9}$$

Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $(f(x_n))_n$ è di Cauchy in N. □

Osservazione 2.31. Siano (M, d_M) e (N, d_N) spazi metrici e $f : M \to N$. Allora f è continua se e solo se

$$\forall x \in M : \forall \epsilon > 0 : \exists \delta > 0 : \forall x' \in M : (d_M(x, x') < \delta \Rightarrow d_N(f(x), f(x')) < \epsilon)\,,$$

mentre f è uniformemente continua se e solo se

$$\forall \epsilon > 0 : \exists \delta > 0 : \forall x \in M : \forall x' \in M : (d_M(x, x') < \delta \Rightarrow d_N(f(x), f(x')) < \epsilon)\,.$$

Dal confronto è chiaro che in generale l'uniforme continuità è una proprietà più forte della continuità: nel caso della continuità, la quantità δ può dipendere sia da x che da ϵ, mentre nella continuità uniforme δ dipende solo da ϵ. Dunque ogni funzione uniformemente continua è continua. Il viceversa in generale non vale, a meno di assumere qualche ipotesi sullo spazio di partenza.

Teorema 2.32 (Heine–Cantor)**.** *Siano (M, d_M) e (N, d_N) spazi metrici, con M compatto. Se $f : M \to N$ è continua, allora f è uniformemente continua.*

Introduciamo una nozione più restrittiva rispetto alla continuità uniforme.

Definizione 2.33. Siano (M, d_M) e (N, d_N) spazi metrici. Sia $L \geq 0$. Una funzione $f : M \to N$ si dice *L-lipschitziana* (o *lipschitziana di costante L*) se

$$\forall x, x' \in M : d_N(f(x), f(x')) \leq L d_M(x, x').$$

Diciamo che f è *lipschitziana* se f è L-lipschitziana per qualche $L \geq 0$.

Osservazione 2.34. Se f è L-lipschitziana, allora f è uniformemente continua: dato $\epsilon > 0$, basta prendere $\delta = \epsilon/L$ per soddisfare la condizione nella definizione di uniforme continuità. L'implicazione opposta in generale non vale.

Esempio 2.35. Consideriamo $\mathbb{R}$ e i suoi sottoinsiemi come spazi metrici con la metrica euclidea.

(1) La funzione $f : [0, 1] \to \mathbb{R}$ definita da $f(t) = \sqrt{t}$ è uniformemente continua (per il teorema di Heine–Cantor), ma non lipschitziana: infatti, se f fosse L-lipschitziana per qualche $L \geq 0$, si avrebbe

$$\sqrt{t} = |f(t) - f(0)| \leq L|t - 0| = Lt$$

per ogni $t \in [0, 1]$, da cui $L \geq t^{-1/2}$ per ogni $t \in (0, 1]$, e passando al limite per $t \to 0^+$ si avrebbe un assurdo.

(2) La funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(t) = t$ è 1-lipschitziana, dato che

$$|f(t) - f(t')| = |t - t'| \leq 1 \cdot |t - t'|$$

per ogni $t, t' \in \mathbb{R}$.

(3) La funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(t) = t^2$ è continua, ma non uniformemente continua. Infatti, se f fosse uniformemente continua, preso $\epsilon = 1$, dovrebbe esistere $\delta > 0$ tale che $|f(t) - f(t')| < 1$ ogniqualvolta $|t - t'| < \delta$; in particolare, preso $t' = t + \delta/2$, si avrebbe $|t - t'| = \delta/2 < \delta$ e

$$1 > |f(t) - f(t')| = |t^2 - (t + \delta/2)^2| = |2t + \delta/2|\delta/2$$

per ogni $t \in \mathbb{R}$, e passando al limite per $t \to +\infty$ si otterrebbe un assurdo.

Un'altra conseguenza importante della compattezza è la seguente estensione del teorema di Weierstrass dell'analisi reale.

Teorema 2.36 (Weierstrass)**.** *Siano M uno spazio metrico compatto e $f : M \to \mathbb{R}$ una funzione continua. Allora f è limitata ed esistono $x, y \in M$ tali che $f(x) = \max f$ e $f(y) = \min f$.*

Come conseguenza del teorema di Weierstrass, si ha in particolare $C_{\mathbb{F}}(M) \subseteq \mathcal{F}_b(M, \mathbb{F})$ per ogni spazio metrico compatto M. Utilizzando il teorema 2.15 e la proposizione 2.19(i), deduciamo la seguente estensione del teorema 2.14.

Teorema 2.37. *Sia M uno spazio metrico compatto. Allora $(C_{\mathbb{F}}(M), d_\infty)$ è uno spazio metrico completo.*

Richiamiamo infine un importante risultato di compattezza per sottoinsiemi dello spazio metrico $(C_{\mathbb{F}}(M), d_\infty)$, di cui omettiamo la dimostrazione.

Teorema 2.38 (Ascoli–Arzelà)**.** *Sia M uno spazio metrico compatto. Sia $(f_n)_n$ una successione a valori in $C_{\mathbb{F}}(M)$. Assumiamo che:*

(a) la successione $(f_n)_n$ è limitata*:*

$$\sup_{n \in \mathbb{N}} \sup_{t \in M} |f_n(t)| < \infty;$$

(b) la successione $(f_n)_n$ è equicontinua*:*

$$\forall t \in M : \forall \epsilon > 0 : \exists \delta > 0 : \forall n \in \mathbb{N} : \forall t' \in M :$$
$$(d_M(t, t') < \delta \Rightarrow |f_n(t) - f_n(t')| < \epsilon).$$

Allora $(f_n)_n$ ha una sottosuccessione convergente in $(C_{\mathbb{F}}(M), d_\infty)$ a una qualche $f \in C_{\mathbb{F}}(M)$.

3 Richiami di teoria della misura

Si riportano qui, molto schematicamente e senza dimostrazioni, i principali risultati e definizioni di teoria della misura e dell'integrazione utilizzati nel seguito.

Definizione 3.1. Sia M un insieme.

(a) $\mathcal{P}(M)$ denota l'*insieme delle parti* di M, cioè la famiglia di tutti i sottoinsiemi di M.

(b) Una *σ-algebra* su M è una famiglia $\mathcal{M} \subseteq \mathcal{P}(M)$ di sottoinsiemi di M con le seguenti proprietà:

- $\emptyset, M \in \mathcal{M}$;
- se $A \in \mathcal{M}$ allora anche $M \setminus A \in \mathcal{M}$;
- se $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$, allora anche $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

(c) Sia $\mathcal{M}$ una σ-algebra su M. Una *misura* (nonnegativa) su $(M, \mathcal{M})$ è una funzione $\mu : \mathcal{M} \to [0, \infty]$ con le seguenti proprietà:

- $\mu(\emptyset) = 0$;
- se $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ e $A_n \cap A_m = \emptyset$ per ogni $n, m \in \mathbb{N}$ con $n \neq m$, allora $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

(d) Se $\mathcal{M}$ è una σ-algebra su M e μ è una misura su $(M, \mathcal{M})$, diciamo $(M, \mathcal{M}, \mu)$ uno *spazio di misura*.

(e) Uno spazio di misura $(M, \mathcal{M}, \mu)$ si dice *finito* se $\mu(M) < \infty$.

(f) Uno spazio di misura $(M, \mathcal{M}, \mu)$ si dice *σ-finito* se esiste una sottofamiglia numerabile $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ tale che $M = \bigcup_{n\in\mathbb{N}} A_n$ e $\mu(A_n) < \infty$ per ogni $n \in \mathbb{N}$.

Definizione 3.2. Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura.

(a) Un sottoinsieme $A \subseteq M$ si dice *misurabile* se $A \in \mathcal{M}$.

(b) Un insieme $A \in \mathcal{M}$ si dice *μ-trascurabile* se $\mu(A) = 0$.

(c) Diciamo che una proprietà di punti di M vale *μ-quasi ovunque* se l'insieme dei punti dove tale proprietà non vale è μ-trascurabile. In particolare, per due funzioni f, g con dominio M, diciamo che "$f = g$ μ-quasi ovunque" se l'insieme $\{t \in M : f(t) \neq g(t)\}$ è μ-trascurabile.

(d) Una funzione $f : M \to [0, \infty]$ si dice *misurabile* se $f^{-1}((a, \infty]) \in \mathcal{M}$ per ogni $a > 0$.

(e) Una funzione $f : M \to \mathbb{C}$ si dice *misurabile* se $f^{-1}(A) \in \mathcal{M}$ per ogni sottoinsieme aperto A di $\mathbb{C}$.

Osservazione 3.3. Le due definizioni di funzione misurabile date sopra sono compatibili: in altre parole, se $f : M \to [0, \infty)$, allora f è misurabile nel senso di (d) se e solo se è misurabile nel senso di (e).

Proposizione 3.4. *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura.*

(i) *Se $A \in \mathcal{M}$, allora la* funzione caratteristica $\mathbf{1}_A$ *di A, definita da*

$$\mathbf{1}_A(t) = \begin{cases} 1 & \text{se } t \in A, \\ 0 & \text{altrimenti}, \end{cases}$$

è misurabile.

(ii) *Se $f, g : M \to \mathbb{C}$ sono misurabili, allora $f + g$ e $f \cdot g$ sono misurabili. Lo stesso vale per funzioni misurabili $f, g : M \to [0, \infty]$, ove si conviene che $0 \cdot \infty = 0$.*

(iii) *Se $f : M \to \mathbb{C}$ è misurabile e $\phi : \mathbb{C} \to \mathbb{C}$ è continua, allora anche $\phi \circ f$ è misurabile.*

(iv) *Sia $(f_n)_{n\in\mathbb{N}}$ una successione di funzioni misurabili a valori in $[0, \infty]$. Sia $f : M \to [0, \infty]$ data da una delle seguenti espressioni:*

$$\begin{aligned} f(t) &= \sup_{n\in\mathbb{N}} f_n(t) \qquad \forall t \in M, \\ f(t) &= \inf_{n\in\mathbb{N}} f_n(t) \qquad \forall t \in M, \\ f(t) &= \limsup_{n\to\infty} f_n(t) \qquad \forall t \in M, \\ f(t) &= \liminf_{n\to\infty} f_n(t) \qquad \forall t \in M. \end{aligned}$$

Allora f è misurabile.

(v) *Sia $(f_n)_{n\in\mathbb{N}}$ una successione di funzioni misurabili a valori in $\mathbb{C}$. Sia $f : M \to \mathbb{C}$ data da $f(t) = \lim_{n\to\infty} f_n(t)$ per ogni $t \in M$. Allora f è misurabile.*

(vi) *Sia S^+ l'insieme delle funzioni semplici misurabili nonnegative, dato da*

$$S^+ = \left\{ \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j} : n \in \mathbb{N},\ A_1, \dots, A_n \in \mathcal{M},\ \alpha_1, \dots, \alpha_n \in [0, \infty) \right\}.$$

Per ogni $f : M \to [0, \infty]$ misurabile, esiste una successione $(f_n)_n$ a valori in S^+ e crescente (cioè $f_n \leq f_m$ per $n \leq m$) tale che $f(t) = \lim_{n\to\infty} f_n(t)$ per ogni $t \in M$.

Teorema 3.5 (integrale di Lebesgue)**.** *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura. Esiste un unico modo di associare ad ogni funzione misurabile $f : M \to [0, \infty]$ un elemento $\int_M f \, d\mu \in [0, \infty]$, detto* integrale *della funzione f, in modo tale che:*

(i) *per ogni $A \in \mathcal{M}$, $\int_M \mathbf{1}_A \, d\mu = \mu(A)$;*

(ii) *per ogni $f, g : M \to [0, \infty]$ misurabili e $\alpha, \beta \in [0, \infty)$,*

$$\int_M (\alpha f + \beta g) \, d\mu = \alpha \int_M f \, d\mu + \beta \int_M g \, d\mu.$$

(iii) (convergenza monotona) *se* $(f_n)_{n\in\mathbb{N}}$ *è una successione* <u>*crescente*</u> *(cioè* $f_n \leq f_m$ *per* $n \leq m$*) di funzioni misurabili a valori in* $[0,\infty]$ *e* $f : M \to [0,\infty]$ *è data da* $f(t) = \sup_{n\in\mathbb{N}} f_n(t)$ *per ogni* $t \in M$*, allora*

$$\int_M f\,d\mu = \sup_{n\in\mathbb{N}} \int_M f_n\,d\mu.$$

Valgono inoltre le seguenti proprietà:

(iv) *Se* $f, g : M \to [0,\infty]$ *sono misurabili e* $f = g$ μ*-quasi ovunque, allora* $\int_M f\,d\mu = \int_M g\,d\mu$*.*

(v) *Se* $f : M \to [0,\infty]$ *è misurabile, allora* $\int_M f\,d\mu = 0$ *se e solo se* $f = 0$ μ*-quasi ovunque.*

(vi) *Se* $f : M \to [0,\infty]$ *è misurabile e* $\int_M f\,d\mu < \infty$*, allora* $f < \infty$ μ*-quasi ovunque.*

(vii) (monotonia) *Se* $f, g : M \to [0,\infty]$ *sono misurabili e* $f \leq g$*, allora si ha* $\int_M f\,d\mu \leq \int_M g\,d\mu$*.*

(viii) (lemma di Fatou) *Sia* $(f_n)_{n\in\mathbb{N}}$ *una successione di funzioni misurabili a valori in* $[0,\infty]$*, e sia* $f : M \to [0,\infty]$ *data da* $f(t) = \liminf_{n\to\infty} f_n(t)$ *per ogni* $t \in M$*. Allora*

$$\int_M f\,d\mu \leq \liminf_{n\to\infty} \int_M f_n\,d\mu.$$

Notazione. Per una funzione $u : M \to \mathbb{R}$, denotiamo con u_+ e u_- la *parte positiva* e la *parte negativa* di u, definite da

$$u_+(t) := \max\{u(t), 0\}, \qquad u_-(t) := \max\{-u(t), 0\}$$

per ogni $t \in M$.

Definizione 3.6. Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura.

(a) Una funzione misurabile $f : M \to \mathbb{C}$ è detta *sommabile* se $\int_M |f|\,d\mu < \infty$.

(b) Sia $f : M \to \mathbb{C}$ una funzione sommabile, e decomponiamo

$$f = \mathfrak{Re}\, f + i\,\mathfrak{Im}\, f = (\mathfrak{Re}\, f)_+ - (\mathfrak{Re}\, f)_- + i(\mathfrak{Im}\, f)_+ - i(\mathfrak{Im}\, f)_- \tag{3.1}$$

come combinazione lineare di funzioni nonnegative. Definiamo l'*integrale* $\int_M f\,d\mu \in \mathbb{C}$ di f ponendo

$$\begin{aligned}\int_M f\,d\mu = &\int_M (\mathfrak{Re}\, f)_+\,d\mu - \int_M (\mathfrak{Re}\, f)_-\,d\mu \\ &+ i\int_M (\mathfrak{Im}\, f)_+\,d\mu - i\int_M (\mathfrak{Im}\, f)_-\,d\mu.\end{aligned}$$

Notazione. Scriviamo anche $\int_M f(t)\,d\mu(t)$ al posto di $\int_M f\,d\mu$, quando sia comodo esplicitare nella notazione la variabile t di integrazione.

Teorema 3.7 (proprietà dell'integrale)**.** *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura.*

(i) *Se $f : M \to \mathbb{C}$ è sommabile, allora*
$$\left| \int_M f \, d\mu \right| \leq \int_M |f| \, d\mu$$
e inoltre
$$\int_M f \, d\mu = \int_M \mathfrak{Re} \, f \, d\mu + i \int_M \mathfrak{Im} \, f \, d\mu, \qquad \overline{\int_M f \, d\mu} = \int_M \overline{f} \, d\mu.$$
In particolare, se f è a valori reali, allora $\int_M f \, d\mu \in \mathbb{R}$.

(ii) *Se $f, g : M \to \mathbb{C}$ sono misurabili, g è sommabile e $|f| \leq |g|$, allora anche f è sommabile.*

(iii) *Se $f, g : M \to \mathbb{C}$ sono entrambe sommabili e $f = g$ μ-quasi ovunque, allora $\int_M f \, d\mu = \int_M g \, d\mu$.*

(iv) *Se $f, g : M \to \mathbb{R}$ sono sommabili e $f \leq g$ μ-quasi ovunque, allora $\int_M f \, d\mu \leq \int_M g \, d\mu$.*

(v) *Se $f, g : M \to \mathbb{C}$ sono sommabili e $\alpha, \beta \in \mathbb{C}$, allora $\alpha f + \beta g$ è sommabile e*
$$\int_M (\alpha f + \beta g) \, d\mu = \alpha \int_M f \, d\mu + \beta \int_M g \, d\mu.$$

(vi) (convergenza dominata) *Sia $(f_n)_{n \in \mathbb{N}}$ una successione di funzioni misurabili, che converge puntualmente μ-quasi ovunque a una funzione $f : M \to \mathbb{C}$. Se esiste una funzione sommabile g tale che $|f_n| \leq |g|$ μ-quasi ovunque per ogni $n \in \mathbb{N}$, allora sia le funzioni f_n che la funzione f sono sommabili e*
$$\int_M f \, d\mu = \lim_{n \to \infty} \int_M f_n \, d\mu.$$

Teorema 3.8 (misura su sottoinsiemi)**.** *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura. Sia $E \in \mathcal{M}$.*

(i) *L'insieme $\mathcal{M}|_E \coloneqq \mathcal{M} \cap \mathcal{P}(E)$ è una σ-algebra su E. Inoltre la restrizione di μ a $\mathcal{M}|_E$ è una misura su $(E, \mathcal{M}|_E)$, che diciamo la misura* indotta *da $(M, \mathcal{M}, \mu)$ su E e denotiamo ancora con μ.*

(ii) *Se $(M, \mathcal{M}, \mu)$ è σ-finito, allora anche $(E, \mathcal{M}|_E, \mu)$ lo è.*

(iii) *Se f è una funzione misurabile su $(M, \mathcal{M}, \mu)$, allora la sua restrizione $f|_E$ è misurabile su $(E, \mathcal{M}|_E, \mu)$.*

(iv) *Se g è una funzione misurabile su $(E, \mathcal{M}|_E, \mu)$, allora la sua* estensione a zero *g_0 su M, data da*
$$g_0(t) = \begin{cases} g(t) & \text{se } t \in E, \\ 0 & \text{altrimenti,} \end{cases}$$
è misurabile su $(M, \mathcal{M}, \mu)$. Inoltre, g è sommabile su $(E, \mathcal{M}|_E, \mu)$ se e solo se g_0 è sommabile su $(M, \mathcal{M}, \mu)$. Infine, se g è nonnegativa o sommabile,
$$\int_E g \, d\mu = \int_M g_0 \, d\mu.$$

Teorema 3.9 (misura prodotto). *Siano $(M_1, \mathcal{M}_1, \mu_1)$ e $(M_2, \mathcal{M}_2, \mu_2)$ due spazi di misura σ-finiti. Sia $M = M_1 \times M_2$.*

(i) Sull'insieme M esiste un'unica σ-algebra $\mathcal{M}$, detta σ-algebra prodotto *di $\mathcal{M}_1$ e $\mathcal{M}_2$, con la proprietà di essere la più piccola σ-algebra su M contenente gli insiemi della forma $A_1 \times A_2$ per ogni $A_1 \in \mathcal{M}_1$ e $A_2 \in \mathcal{M}_2$.*

(ii) Esiste inoltre un'unica misura μ su $(M, \mathcal{M})$, detta misura prodotto *di μ_1 e μ_2, tale che*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

per ogni $A_1 \in \mathcal{M}_1$ e $A_2 \in \mathcal{M}_2$.

Per lo spazio di misura $(M, \mathcal{M}, \mu)$, detto spazio di misura prodotto *degli spazi $(M_1, \mathcal{M}_1, \mu_1)$ e $(M_2, \mathcal{M}_2, \mu_2)$, valgono le seguenti proprietà.*

(iii) Lo spazio di misura $(M, \mathcal{M}, \mu)$ è σ-finito.

(iv) Se f_1 e f_2 sono funzioni misurabili su $(M_1, \mathcal{M}_1, \mu_1)$ e $(M_2, \mathcal{M}_2, \mu_2)$, allora il loro prodotto tensore *$f_1 \otimes f_2$, definito da*

$$f_1 \otimes f_2(t_1, t_2) = f_1(t_1) f_2(t_2) \qquad \forall (t_1, t_2) \in M,$$

è misurabile su $(M, \mathcal{M}, \mu)$.

(v) Se $f : M \to \mathbb{C}$ è misurabile, allora le sue sezioni *$f(t_1, \cdot) : M_2 \to \mathbb{C}$ e $f(\cdot, t_2) : M_1 \to \mathbb{C}$ sono misurabili per ogni $t_1 \in M_1$ e $t_2 \in M_2$. Lo stesso vale per funzioni misurabili a valori in $[0, \infty]$.*

(vi) (teorema di Fubini–Tonelli per funzioni nonnegative) *Se $f : M \to [0, \infty]$ è una funzione misurabile, allora $\int_{M_2} f(\cdot, t_2)\, d\mu_2(t_2)$ è misurabile su M_1, $\int_{M_1} f(t_1, \cdot)\, d\mu_1(t_1)$ è misurabile su M_2, e*

$$\begin{aligned} \int_M f\, d\mu &= \int_{M_2} \int_{M_1} f(t_1, t_2)\, d\mu_1(t_1)\, d\mu_2(t_2) \\ &= \int_{M_1} \int_{M_2} f(t_1, t_2)\, d\mu_2(t_2)\, d\mu_1(t_1). \end{aligned}$$

In particolare, se $\int_M f\, d\mu < \infty$, allora $\int_{M_2} f(\cdot, t_2)\, d\mu_2(t_2) < \infty$ μ_1-quasi ovunque e $\int_{M_1} f(t_1, \cdot)\, d\mu_1(t_1) < \infty$ μ_2-quasi ovunque.

(vii) (teorema di Fubini–Tonelli per funzioni a valori complessi) *Se $f : M \to \mathbb{C}$ è sommabile, allora le sue sezioni $f(t_1, \cdot) : M_2 \to \mathbb{C}$ e $f(\cdot, t_2) : M_1 \to \mathbb{C}$ sono sommabili per quasi ogni $t_1 \in M_1$ e $t_2 \in M_2$. Inoltre $\int_{M_2} f(\cdot, t_2)\, d\mu_2(t_2)$ è sommabile su M_1, $\int_{M_1} f(t_1, \cdot)\, d\mu_1(t_1)$ è sommabile su M_2, e*

$$\begin{aligned} \int_M f\, d\mu &= \int_{M_2} \int_{M_1} f(t_1, t_2)\, d\mu_1(t_1)\, d\mu_2(t_2) \\ &= \int_{M_1} \int_{M_2} f(t_1, t_2)\, d\mu_2(t_2)\, d\mu_1(t_1). \end{aligned}$$

In queste note, utilizzeremo prevalentemente due classi di misure. La prima è la misura di Lebesgue su (sottoinsiemi di) $\mathbb{R}^n$, che include ed estende la classica teoria dell'integrazione secondo Riemann.

Teorema 3.10 (misura di Lebesgue)**.** *Sia $n \in \mathbb{N}_+$.*

(i) *Su $\mathbb{R}^n$ esiste una σ-algebra $\mathcal{M}_{\mathbb{R}^n}$, detta* σ-algebra di Borel*, con la proprietà di essere la più piccola σ-algebra su $\mathbb{R}^n$ che contiene tutti i sottoinsiemi aperti di $\mathbb{R}^n$. Gli elementi di tale σ-algebra sono detti sottoinsiemi* boreliani *di $\mathbb{R}^n$.*

(ii) *Esiste un'unica misura λ_n su $(\mathbb{R}^n, \mathcal{M}_{\mathbb{R}^n})$, detta* misura di Lebesgue*, con le proprietà che*
$$\lambda_n([0,1]^n) = 1$$
e che λ_n sia invariante per traslazioni, cioè
$$\lambda_n(A + x) = \lambda_n(A)$$
per ogni $A \in \mathcal{M}_{\mathbb{R}^n}$ e $x \in \mathbb{R}^n$, dove $A + x = \{y + x : y \in A\}$.

Valgono inoltre le seguenti proprietà.

(iii) *Lo spazio di misura $(\mathbb{R}^n, \mathcal{M}_{\mathbb{R}^n}, \lambda_n)$ è σ-finito.*

(iv) *Ogni funzione continua $f : \mathbb{R}^n \to \mathbb{C}$ è misurabile rispetto alla σ-algebra di Borel su $\mathbb{R}^n$.*

(v) *Per ogni boreliano $B \in \mathcal{M}_{\mathbb{R}^n}$ e ogni $\epsilon > 0$ esiste un aperto $A \subseteq \mathbb{R}^n$ tale che $B \subseteq A$ e $\lambda_n(A \setminus B) < \epsilon$.*

(vi) *Per ogni $n, m \in \mathbb{N}_+$, se $\mathbb{R}^{n+m}$ è pensato come il prodotto $\mathbb{R}^n \times \mathbb{R}^m$, allora la σ-algebra di Borel su $\mathbb{R}^{n+m}$ è la σ-algebra prodotto delle σ-algebre di Borel su $\mathbb{R}^n$ e $\mathbb{R}^m$. Inoltre la misura di Lebesgue λ_{n+m} è la misura prodotto delle misure di Lebesgue λ_n e λ_m.*

(vii) *Se $[a,b]$ è un intervallo chiuso e limitato di $\mathbb{R}$ e $f : [a,b] \to \mathbb{R}$ è continua, allora*
$$\int_{[a,b]} f \, d\lambda_1 = \int_a^b f(t) \, dt,$$
dove l'integrale a secondo membro è l'integrale di Riemann.

La seconda classe di misure con cui lavoreremo è quella delle "misure del conteggio" su insiemi numerabili, che ci permettono di trattare le serie numeriche come integrali.

Teorema 3.11 (misura del conteggio)**.** *Sia M un insieme al più numerabile. Definiamo $\sharp : \mathcal{P}(M) \to [0, \infty]$ ponendo, per ogni $A \subseteq M$,*
$$\sharp(A) = \begin{cases} \infty & \text{se } A \text{ è un insieme infinito,} \\ n & \text{se } A \text{ è un insieme finito di } n \text{ elementi, per qualche } n \in \mathbb{N}. \end{cases}$$

Allora:

(i) *$\sharp$ è una misura su $(M, \mathcal{P}(M))$, detta* misura del conteggio *su M.*

(ii) *Lo spazio $(M, \mathcal{P}(M), \sharp)$ è σ-finito.*

(iii) *Ogni funzione* $f : M \to [0, \infty]$ *oppure* $f : M \to \mathbb{C}$ *è misurabile rispetto a* $(M, \mathcal{P}(M), \sharp)$.

(iv) *L'unico sottoinsieme misurabile di misura nulla è* $\emptyset$. *In particolare, una proprietà vale* $\sharp$*-quasi ovunque e se solo se vale ovunque.*

(v) *Sia* $M = \{x_0\}_{n=0}^N$, *dove* $N \in \mathbb{N} \cup \{\infty\}$, *un'enumerazione iniettiva di* M *(cioè* $x_n \neq x_m$ *per* $n \neq m$*). Allora, per ogni* $f : M \to [0, \infty]$,

$$\int_M f \, d\sharp = \sum_{x \in M} f(x) = \sum_{n=0}^N f(x_n).$$

Lo stesso vale per ogni $f : M \to \mathbb{C}$ *sommabile.*

4 Spazi normati e spazi di Banach

4.1 Norme e spazi normati

Ricordiamo che $\mathbb{F}$ denota il campo dei numeri reali $\mathbb{R}$ oppure il campo dei numeri complessi $\mathbb{C}$.

Definizione 4.1. Sia X uno spazio vettoriale su $\mathbb{F}$. Si dice *norma* su X una funzione $\|\cdot\| : X \to [0,\infty)$ tale che, per ogni $x, y \in X$ e $\alpha \in \mathbb{F}$:

(a) $\|x\| = 0$ se e solo se $x = 0$
(proprietà di separazione);

(b) $\|\alpha x\| = |\alpha|\|x\|$
(1-omogeneità);

(c) $\|x + y\| \leq \|x\| + \|y\|$
(disuguaglianza triangolare).

Uno spazio vettoriale $(X, \|\cdot\|)$ dotato di una norma si dice *spazio normato*.

Definizione 4.2. Se $(X, \|\cdot\|)$ è uno spazio normato, definiamo la *distanza* $d : X \times X \to [0,\infty)$ *indotta dalla norma* $\|\cdot\|$ ponendo

$$d(x, y) = \|x - y\|$$

per ogni $x, y \in X$.

Proposizione 4.3. *Sia $(X, \|\cdot\|)$ uno spazio normato. La distanza d indotta dalla norma $\|\cdot\|$ è una distanza sull'insieme X (nel senso della definizione 2.1).*

Dimostrazione. Per esercizio. □

Dunque ogni spazio normato $(X, \|\cdot\|)$ si può pensare come spazio metrico, con la distanza indotta dalla norma. In particolare agli spazi normati si applica tutta la teoria degli spazi metrici discussa nel capitolo 2.

Definizione 4.4. Si dice *spazio di Banach* uno spazio normato completo.

Esempio 4.5. Ecco alcuni esempi di spazi normati.

(1) Siano $n \in \mathbb{N}_+$ e $p \in [1,\infty]$. Allora $(\mathbb{F}^n, \|\cdot\|_p)$ è uno spazio normato, ove

$$\|x\|_p = \begin{cases} \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} & \text{se } p < \infty, \\ \max\{|x_j| : j = 1, \dots, n\} & \text{se } p = \infty \end{cases}$$

per ogni $x \in \mathbb{F}^n$. La distanza indotta è la distanza d_p discussa nell'esempio 2.2(1). Nel caso $p = 2$, dagli esempi 2.12(2) e 2.26 sappiamo che $(\mathbb{F}^n, d_2)$ è uno spazio metrico completo e separabile, dunque $(\mathbb{F}^n, \|\cdot\|_2)$ è uno spazio di Banach separabile; in effetti, le stesse proprietà valgono anche per $p \neq 2$ (☞ esempio 4.9 e corollario 4.11).

(2) Se M è uno spazio metrico compatto, allora $(C_{\mathbb{F}}(M), \|\cdot\|_\infty)$ è uno spazio normato, ove $\|\cdot\|_\infty$ è la *norma dell'estremo superiore* definita da

$$\|f\|_\infty = \sup_{x\in M} |f(x)| \tag{4.1}$$

per ogni $f \in C_{\mathbb{F}}(M)$. La distanza indotta è la distanza d_∞ discussa nell'esempio 2.2(2). Dal teorema 2.37 sappiamo che $(C_{\mathbb{F}}(M), d_\infty)$ è uno spazio metrico completo, dunque $(C_{\mathbb{F}}(M), \|\cdot\|_\infty)$ è uno spazio di Banach. Inoltre, nel caso $M = [a, b]$ sia un intervallo chiuso e limitato di $\mathbb{R}$, vedremo nel corollario 4.31 che $(C_{\mathbb{F}}[a, b], \|\cdot\|_\infty)$ è uno spazio di Banach separabile.

(3) Se $(X, \|\cdot\|)$ è uno spazio normato e Y è un sottospazio vettoriale di X, allora la restrizione a Y della norma $\|\cdot\|$ è una norma su Y (detta *norma indotta* da $(X, \|\cdot\|)$ su Y), che per brevità denotiamo ancora con $\|\cdot\|$, e dunque $(Y, \|\cdot\|)$ è a sua volta uno spazio normato. Se d è la distanza su X indotta dalla norma $\|\cdot\|$, chiaramente la distanza indotta su Y dalla norma $\|\cdot\|$ è la restrizione di d a $Y \times Y$, cioè coincide con la metrica indotta su Y da (X, d) nel senso dell'esempio 2.2(3).

(4) Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ spazi normati. Ricordiamo che il prodotto diretto $X \times Y$ è uno spazio vettoriale con le operazioni componente per componente (☞ esempio 1.2(5)). Se definiamo

$$\|(x, y)\|_{X\times Y} = \max\{\|x\|_X, \|y\|_Y\}$$

per ogni $(x, y) \in X \times Y$, non è difficile verificare che $\|\cdot\|_{X\times Y}$ è una norma su $X \times Y$, e che la distanza indotta da $\|\cdot\|_{X\times Y}$ su $X \times Y$ coincide con la metrica prodotto delle distanze indotte da $\|\cdot\|_X$ e $\|\cdot\|_Y$ sui fattori (☞ esempio 2.2(4)). Chiamiamo $(X \times Y, \|\cdot\|_{X\times Y})$ lo *spazio normato prodotto* degli spazi $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$. Osserviamo che, per la proposizione 2.13(iii), se X e Y sono spazi di Banach, anche $X \times Y$ è uno spazio di Banach.

Proposizione 4.6. *Sia $(X, \|\cdot\|)$ uno spazio normato.*

(i) La norma $\|\cdot\| : X \to [0, \infty)$ è 1-lipschitziana (dunque uniformemente continua):
$$|\|x\| - \|y\|| \leq \|x - y\| \qquad \forall x, y \in X.$$

(ii) Le operazioni di somma e prodotto scalare-vettore sono continue:

(a) se $x_n \to x$ in X e $\alpha_n \to \alpha$ in $\mathbb{F}$, allora $\alpha_n x_n \to \alpha x$ in X;

(b) se $x_n \to x$ e $y_n \to y$ in X, allora $x_n + y_n \to x + y$ in X.

(iii) Per ogni $x \in X$ e $r > 0$, le palle $B(x, r)$ e $\overline{B}(x, r)$ sono convesse (nel senso della definizione 1.7).

(iv) Un sottoinsieme E di X è limitato se e solo se

$$\sup_{x\in E} \|x\| < \infty.$$

Dunque, una successione $(x_n)_{n\in\mathbb{N}}$ a valori in X è limitata se e solo se

$$\sup_{n\in\mathbb{N}} \|x_n\| < \infty.$$

Dimostrazione. (i). Per la disuguaglianza triangolare applicata ai vettori $x - y$ e y,

$$\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|,$$

pertanto

$$\|x\| - \|y\| \leq \|x - y\|.$$

Scambiando il ruolo di x e y si deduce anche

$$\|y\| - \|x\| \leq \|y - x\| = \|x - y\|,$$

dunque

$$|\|x\| - \|y\|| = \max\{\|x\| - \|y\|, \|y\| - \|x\|\} \leq \|x - y\|,$$

come desiderato.

(ii). Osserviamo che, grazie alla caratterizzazione della continuità per successioni (☞ proposizione 2.29) e alla caratterizzazione delle successioni convergenti in uno spazio prodotto (☞ proposizione 2.13(i)), gli enunciati (a) e (b) sono effettivamente espressioni della continuità delle mappe $\mathbb{F} \times X \ni (\alpha, x) \mapsto \alpha x \in X$ e $X \times X \ni (x, y) \mapsto x + y \in X$, cioè delle operazioni di prodotto scalare-vettore e di somma in X.

Dimostriamo (a). Se $\alpha_n \to \alpha$ in $\mathbb{F}$, si ha $M = \sup_{n \in \mathbb{N}} |\alpha_n| < \infty$. Dunque

$$\begin{aligned} 0 \leq \|\alpha_n x_n - \alpha x\| &\leq \|\alpha_n x_n - \alpha_n x\| + \|\alpha_n x - \alpha x\| \\ &= |\alpha_n| \|x_n - x\| + |\alpha_n - \alpha| \|x\| \leq M \|x_n - x\| + \|x\| |\alpha_n - \alpha|. \end{aligned} \tag{4.2}$$

Se inoltre $x_n \to x$ in X, passando al limite per $n \to \infty$, si ha $\|x_n - x\| \to 0$ e $|\alpha_n - \alpha| \to 0$, dunque $M\|x_n - x\| + \|x\| |\alpha_n - \alpha| \to 0$ per l'Algebra dei Limiti, ma allora il teorema dei carabinieri e la disuguaglianza (4.2) implicano che $\|\alpha_n x_n - \alpha x\| \to 0$, cioè $\alpha_n x_n \to \alpha x$.

In maniera simile si dimostra (b). Se $x_n \to x$ e $y_n \to y$ in X, allora $\|x_n - x\| \to 0$ e $\|y_n - y\| \to 0$, pertanto, per la disuguaglianza triangolare e l'Algebra dei Limiti,

$$0 \leq \|(x_n + y_n) - (x + y)\| = \|(x_n - x) + (y_n - y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0$$

e il teorema dei carabinieri ci permette di concludere che anche $\|(x_n + y_n) - (x + y)\| \to 0$, cioè $x_n + y_n \to x + y$, come desiderato.

(iii). Se $y, z \in B(x, r)$, allora $\|y - x\| < r$ e $\|z - x\| < r$. Dunque, per ogni $\theta \in [0, 1]$,

$$\begin{aligned} \|((1-\theta)y+\theta z)-x\| = \|(1-\theta)(y-x)+\theta(z-x)\| &\leq \|(1-\theta)(y-x)\|+\|\theta(z-x)\| \\ &= (1 - \theta)\|y - x\| + \theta\|z - x\| < (1 - \theta)r + \theta r = r, \end{aligned}$$

da cui segue che $(1 - \theta)y + \theta z \in B(x, r)$. Per la definizione 1.7, questo mostra che $B(x, r)$ è convessa. In maniera simile (rimpiazzando $<$ con $\leq$) si dimostra che $\overline{B}(x, r)$ è convessa.

(iv). Sia E un sottoinsieme di X. Se

$$M = \sup_{x \in E} \|x\| < \infty,$$

allora $d(x, 0) = \|x\| < M + 1$ per ogni $x \in E$, pertanto $E \subseteq B(0, M + 1)$ e dunque, per la definizione 2.21(b), E è limitato.

Viceversa, supponiamo che E sia limitato. Allora, per la definizione 2.21(b), esistono $x_0 \in X$ e $r > 0$ tali che $E \subseteq B(x_0, r) = \{x \in X : \|x - x_0\| < r\}$. In particolare, per ogni $x \in E$,

$$\|x\| = \|(x - x_0) + x_0\| \leq \|x - x_0\| + \|x_0\| < r + \|x_0\|,$$

da cui deduciamo che

$$\sup_{x \in E} \|x\| \leq r + \|x_0\| < \infty,$$

come desiderato. □

4.2 Norme equivalenti

Definizione 4.7. Sia X uno spazio vettoriale su $\mathbb{F}$. Siano $\|\cdot\|$ e $\|\cdot\|'$ norme su X. Diciamo che $\|\cdot\|$ e $\|\cdot\|'$ sono *equivalenti* se esistono costanti $A, B \in (0, \infty)$ tali che

$$\|x\| \leq A\|x\|' \qquad \text{e} \qquad \|x\|' \leq B\|x\| \qquad \text{per ogni } x \in X. \tag{4.3}$$

Osservazione 4.8. La relazione di equivalenza fra norme è riflessiva, simmetrica e transitiva (in altre parole, è una "relazione di equivalenza" propriamente detta); si lascia per esercizio la facile verifica.

Esempio 4.9. Su $\mathbb{F}^n$, le norme $\|\cdot\|_p$ per $p \in [1, \infty]$ definite nell'esempio 4.5(1) sono tutte fra loro equivalenti. Infatti, per ogni $x = (x_1, \dots, x_n) \in \mathbb{F}^n$ e $p \in [1, \infty)$,

$$\|x\|_\infty^p = \max\{|x_j|^p : j = 1, \dots, n\} \leq \sum_{j=1}^{p} |x_j|^p \leq n\|x\|_\infty^p,$$

cioè

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p}\|x\|_\infty; \tag{4.4}$$

questo mostra che $\|\cdot\|_\infty$ è equivalente a $\|\cdot\|_p$ per ogni $p \in [1, \infty)$. Per l'osservazione 4.8, deduciamo che tutte le norme $\|\cdot\|_p$ con $p \in [1, \infty]$ sono fra loro equivalenti.

Proposizione 4.10. *Siano $\|\cdot\|$ e $\|\cdot\|'$ norme sullo spazio vettoriale X. Allora $\|\cdot\|$ e $\|\cdot\|'$ sono equivalenti se e solo se la funzione identità $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$* è bilipschitziana *(cioè lipschitziana con inversa lipschitziana).*

Dimostrazione. Supponiamo che $\|\cdot\|$ e $\|\cdot\|'$ siano equivalenti. Siano $A, B \in (0, \infty)$ le costanti per cui valgono le disuguaglianze (4.3). In particolare, per ogni $x, y \in X$,

$$\|\mathrm{id}_X(x) - \mathrm{id}_X(y)\|' = \|x - y\|' \leq B\|x - y\|,$$

dunque $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ è B-lipschitziana. Scambiando il ruolo di $\|\cdot\|$ e $\|\cdot\|'$, si vede in maniera analoga che $\mathrm{id}_X^{-1} = \mathrm{id}_X : (X, \|\cdot\|') \to (X, \|\cdot\|)$ è A-lipschitziana. Dunque $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ è bilipschitziana, come desiderato.

Viceversa, supponiamo che $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ sia bilipschitziana. Dunque esiste $L > 0$ tale che $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ e $\mathrm{id}_X^{-1} = \mathrm{id}_X$:

$(X, \|\cdot\|') \to (X, \|\cdot\|)$ sono entrambe L-lipschitziane. In particolare, per ogni $x \in X$,

$$\|x\|' = \|x - 0\|' = \|\mathrm{id}_X(x) - \mathrm{id}_X(0)\|' \leq L\|x - 0\| = L\|x\|,$$
$$\|x\| = \|x - 0\| = \|\mathrm{id}_X(x) - \mathrm{id}_X(0)\| \leq L\|x - 0\|' = L\|x\|'$$

dunque (4.3) è soddisfatta con $A = B = L$, e pertanto $\|\cdot\|$ e $\|\cdot\|'$ sono equivalenti. □

Corollario 4.11. *Siano $\|\cdot\|$ e $\|\cdot\|'$ norme equivalenti sullo spazio vettoriale X. Allora:*

(i) Le norme $\|\cdot\|$ e $\|\cdot\|'$ inducono la stessa topologia su X. In particolare, nozioni topologiche come

- *convergenza di successioni,*
- *continuità di funzioni,*
- *sottoinsiemi aperti, chiusi, compatti, densi*

rimangono invariate se si rimpiazza la norma $\|\cdot\|$ con la norma equivalente $\|\cdot\|'$ su X.

(ii) Un sottoinsieme E di X è limitato in $(X, \|\cdot\|)$ se e solo se E è limitato in $(X, \|\cdot\|')$.

(iii) Una successione $(x_n)_n$ a valori in X è di Cauchy in $(X, \|\cdot\|)$ se e solo se è di Cauchy in $(X, \|\cdot\|')$. In particolare, $(X, \|\cdot\|)$ è di Banach se e solo se $(X, \|\cdot\|')$ è di Banach.

Dimostrazione. (i). Dalla proposizione 4.10 segue che $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ è bilipschitziana, dunque id_X è un *omeomorfismo* (continua con inversa continua); pertanto un sottoinsieme A di X è aperto in $(X, \|\cdot\|)$ se e solo se $\mathrm{id}_X(A) = A$ è aperto in $(X, \|\cdot\|')$ (vedere la proposizione 2.28) e quindi le topologie su $(X, \|\cdot\|)$ e $(X, \|\cdot\|')$ sono la stessa.

(ii). Siano $A, B \in (0, \infty)$ le costanti in (4.3). Se E è un sottoinsieme limitato di $(X, \|\cdot\|)$, allora

$$\sup_{x \in E} \|x\| < \infty$$

per la proposizione 4.6(iv), pertanto anche

$$\sup_{x \in E} \|x\|' \leq \sup_{x \in E} (B\|x\|) = B \sup_{x \in E} \|x\| < \infty$$

e dunque E è limitato in $(X, \|\cdot\|')$. Scambiando il ruolo di $\|\cdot\|$ e $\|\cdot\|'$ si dimostra l'implicazione opposta.

(iii). Siccome $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ è bilipschitziana, per l'osservazione 2.34 sia la mappa $\mathrm{id}_X : (X, \|\cdot\|) \to (X, \|\cdot\|')$ che la sua inversa $\mathrm{id}_X : (X, \|\cdot\|') \to (X, \|\cdot\|)$ sono uniformemente continue, e dalla proposizione 2.30 applicata con $f = \mathrm{id}_X$ deduciamo che le successioni di Cauchy in $(X, \|\cdot\|)$ e in $(X, \|\cdot\|')$ sono le stesse. □

Esempio 4.12. Abbiamo già osservato nell'esempio 4.5(1) che $(\mathbb{F}^n, \|\cdot\|_2)$ è uno spazio di Banach. Siccome su $\mathbb{F}^n$ le norme $\|\cdot\|_p$ per $p \in [1, \infty]$ sono tutte equivalenti (☞ esempio 4.9), dal corollario 4.11 deduciamo che anche $(\mathbb{F}^n, \|\cdot\|_p)$ è uno spazio di Banach per ogni $p \in [1, \infty]$.

Il fatto che le norme $\|\cdot\|_p$ su $\mathbb{F}^n$ siano tutte fra loro equivalenti è in realtà un caso particolare di un risultato molto più generale, che vale in ogni spazio vettoriale di dimensione finita.

Teorema 4.13 (equivalenza delle norme)**.** *Sia X uno spazio vettoriale su $\mathbb{F}$. Siano $\|\cdot\|$ e $\|\cdot\|'$ norme su X. Se $\dim X < \infty$, allora $\|\cdot\|$ e $\|\cdot\|'$ sono equivalenti.*

Dimostrazione. Passo 1. Scegliendo una base di X, costruiamo una mappa lineare biiettiva tra X e $\mathbb{F}^n$, dove $n = \dim X$ (☞ osservazione 1.23). Tramite tale mappa, possiamo "trapiantare" su $\mathbb{F}^n$ le norme su X di cui vogliamo mostrare l'equivalenza. Quindi senza perdita di generalità possiamo assumere $X = \mathbb{F}^n$.

Passo 2. Possiamo assumere che una delle due norme, diciamo $\|\cdot\|'$, è la norma euclidea $\|\cdot\|_2$ su $\mathbb{F}^n$. Infatti, per l'osservazione 4.8, se tutte le norme su $\mathbb{F}^n$ sono equivalenti alla norma euclidea, allora sono anche equivalenti fra loro.

Passo 3. La norma $\|\cdot\|$ è lipschitziana come mappa da $(\mathbb{F}^n, \|\cdot\|_2)$ a $[0,\infty)$. Infatti, per ogni $x = (x_1, \dots, x_n) \in \mathbb{F}^n$, possiamo scrivere $x = \sum_j x_j e_j$, ove $e_1, \dots, e_n$ è la base canonica di $\mathbb{F}^n$; dunque

$$\|x\| = \left\| \sum_j x_j e_j \right\| \leq \sum_j |x_j| \, \|e_j\| \leq L\|x\|_\infty \leq L\|x\|_2,$$

ove $L = \sum_j \|e_j\| \in (0,\infty)$, e si è usata la disuguaglianza (4.4) fra le norme $\|x\|_\infty$ e $\|x\|_2$; di conseguenza, per ogni $x, y \in \mathbb{F}^n$,

$$|\|x\| - \|y\|| \leq \|x - y\| \leq L\|x - y\|_2,$$

ove si è usata la proposizione 4.6(i). Questo mostra che $\|\cdot\|$ è L-lipschitziana rispetto alla norma $\|\cdot\|_2$ su $\mathbb{F}^n$.

Passo 4. Sia $S = \{x \in \mathbb{F}^n : \|x\|_2 = 1\}$ la sfera unitaria in $(\mathbb{F}^n, \|\cdot\|_2)$. Per il teorema 2.24, S è un sottoinsieme compatto di $(\mathbb{F}^n, \|\cdot\|_2)$. Dal passo 3 sappiamo che $\|\cdot\|$ è lipschitziana, dunque continua, su $(\mathbb{F}^n, \|\cdot\|_2)$; pertanto, per il teorema 2.36, $\|\cdot\|$ ha un massimo e un minimo su S.

Passo 5. Si ha

$$m := \min\{\|x\| : x \in S\} > 0.$$

In caso contrario, ci sarebbe $x \in S$ con $\|x\| = 0$; siccome $\|\cdot\|$ è una norma, ne dedurremmo che $x = 0$, il che è assurdo perché $0 \notin S$.

Passo 6. Per ogni $x \in \mathbb{F}^n$, se $x \neq 0$, si ha

$$\|x/\|x\|_2\|_2 = \|x\|_2/\|x\|_2 = 1$$

per la 1-omogeneità della norma; dunque $x/\|x\|_2 \in S$, ma allora $\|x/\|x\|_2\| \geq m$ e quindi

$$\|x\| = \|x\|_2 \|x/\|x\|_2\| \geq m\|x\|_2.$$

Questo dimostra che

$$\|x\|_2 \leq \frac{1}{m}\|x\|$$

per ogni $x \in \mathbb{F}^n$ con $x \neq 0$, e chiaramente la stessa disuguaglianza vale quando $x = 0$. Inoltre, dal passo 3, sappiamo che

$$\|x\| \leq L\|x\|_2$$

per ogni $x \in \mathbb{F}^n$. Le due disuguaglianze mostrano che le norme $\|\cdot\|$ e $\|\cdot\|_2$ sono equivalenti (con costanti L e $1/m$), come desiderato. □

Corollario 4.14. *Sia X uno spazio vettoriale di dimensione finita. Allora ogni norma $\|\cdot\|$ su X induce la stessa topologia, e rispetto a tale norma:*

(i) $(X, \|\cdot\|)$ *è uno spazio di Banach separabile;*

(ii) *ogni sottoinsieme chiuso e limitato di X è compatto.*

Dimostrazione. Come nella dimostrazione del teorema 4.13, utilizzando coordinate lineari (☞ osservazione 1.23), non è restrittivo supporre che $X = \mathbb{F}^n$.

Per il teorema 4.13, qualunque norma $\|\cdot\|$ su $\mathbb{F}^n$ è equivalente alla norma euclidea $\|\cdot\|_2$, dunque induce la stessa topologia (☞ corollario 4.11). Inoltre, siccome le proprietà (i) e (ii) valgono per $\mathbb{F}^n$ con la norma euclidea (☞ esempio 4.5(1) e teorema 2.24), le stesse proprietà valgono anche con la norma equivalente $\|\cdot\|$ (☞ corollario 4.11). □

La discussione precedente mostra che, se vogliamo un esempio di norme non equivalenti, dobbiamo lavorare su spazi vettoriali di dimensione infinita.

Esempio 4.15. Dall'esempio 4.5(2), sappiamo che $X = C[0,1]$ è uno spazio normato con la norma

$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|.$$

Definiamo anche

$$\|f\|_* = \sup_{t\in[0,1]} |(1+t)f(t)|, \qquad \|f\|_\dagger = \sup_{t\in[0,1]} |tf(t)|$$

per ogni $f \in X$. Allora:

1. $\|\cdot\|_*$ e $\|\cdot\|_\dagger$ sono norme su X. (La verifica è lasciata per esercizio.)

2. $\|\cdot\|_*$ e $\|\cdot\|_\infty$ sono norme equivalenti su X. Infatti $1 \le |1+t| \le 2$ per ogni $t \in [0,1]$; quindi, per ogni $f \in X$ e $t \in [0,1]$,

$$|f(t)| \le |(1+t)f(t)| \le 2|f(t)|$$

e passando all'estremo superiore per $t \in [0,1]$ si ottiene

$$\|f\|_\infty \le \|f\|_* \le 2\|f\|_\infty.$$

3. $\|\cdot\|_\dagger$ e $\|\cdot\|_\infty$ non sono equivalenti. Infatti, per ogni $n \in \mathbb{N}$, definiamo $f_n \in X$ ponendo

$$f_n(t) = \max\{0, 1-nt\} \qquad \text{per ogni } t \in [0,1].$$

Allora $1 \ge \|f_n\|_\infty \ge |f(0)| = 1$, cioè $\|f_n\|_\infty = 1$ per ogni $n \in \mathbb{N}$. D'altra parte, $f_n(t) = 0$ per ogni $t > 1/n$, pertanto

$$\|f_n\|_\dagger = \sup_{t\in[0,1]} |tf_n(t)| = \sup_{t\in[0,1/n]} |tf_n(t)| \le 1/n$$

e quindi $\|f_n\|_\dagger \to 0$ per $n \to \infty$. Questo mostra che $f_n \to 0$ in $(X, \|\cdot\|_\dagger)$, mentre $f_n \not\to 0$ in $(X, \|\cdot\|_\infty)$; per il corollario 4.11, concludiamo che le norme $\|\cdot\|_\dagger$ e $\|\cdot\|_\infty$ non sono equivalenti.

4.3 Completezza e chiusura

Proposizione 4.16. *Siano* $(X, \|\cdot\|)$ *uno spazio normato e* Y *un sottospazio vettoriale di* X*.*

(i) Se $(Y, \|\cdot\|)$ *è completo, allora* Y *è chiuso in* X*.*

(ii) Se $\dim Y < \infty$*, allora* Y *è chiuso in* X*.*

(iii) Se $(X, \|\cdot\|)$ *è uno spazio di Banach, allora* Y *è chiuso in* X *se e solo se* $(Y, \|\cdot\|)$ *è completo.*

Dimostrazione. (i). Per la caratterizzazione dei sottoinsiemi chiusi data nella proposizione 2.23(vi), dobbiamo dimostrare che, se $(x_n)_n$ è una successione a valori in Y che converge in $(X, \|\cdot\|)$ a un punto $x \in X$, allora $x \in Y$.

Siccome $x_n \to x$ in X, per la proposizione 2.9(iii) la successione $(x_n)_n$ è di Cauchy in $(X, \|\cdot\|)$, cioè

$$\forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n, m > N : \|x_n - x_m\| < \epsilon. \tag{4.5}$$

D'altra parte, $(x_n)_n$ è a valori in Y, e la norma su Y è la restrizione della norma su X, quindi (4.5) ci dice anche che $(x_n)_n$ è una successione di Cauchy in $(Y, \|\cdot\|)$. Siccome per ipotesi $(Y, \|\cdot\|)$ è completo, ne deduciamo che $(x_n)_n$ converge in $(Y, \|\cdot\|)$ a qualche punto $y \in Y$. In altre parole,

$$\|x_n - y\| \to 0 \qquad \text{per } n \to \infty. \tag{4.6}$$

Siccome la norma su X e su Y è la stessa e $y \in Y \subseteq X$, (4.6) ci dice anche che $x_n \to y$ in $(X, \|\cdot\|)$. Sappiamo già tuttavia che $x_n \to x$ in $(X, \|\cdot\|)$. Per l'unicità del limite, ne concludiamo che $x = y$ e che quindi $x \in Y$, come desiderato.

(ii). Dal corollario 4.14 sappiamo che $(Y, \|\cdot\|)$ è completo, dunque possiamo applicare (i).

(iii). Grazie a (i), rimane solo da dimostrare che, se Y è chiuso in X, allora $(Y, \|\cdot\|)$ è completo.

Supponiamo dunque che Y sia chiuso in X, e sia $(y_n)_n$ una successione di Cauchy in $(Y, \|\cdot\|)$. Siccome la norma su X e Y è la stessa e $Y \subseteq X$, ne deduciamo che $(y_n)_n$ è anche una successione di Cauchy in $(X, \|\cdot\|)$. Siccome per ipotesi $(X, \|\cdot\|)$ è di Banach, cioè completo, $(x_n)_n$ converge in $(X, \|\cdot\|)$ a un punto $x \in X$, cioè

$$\|x_n - x\| \to 0 \qquad \text{per } n \to \infty. \tag{4.7}$$

D'altra parte, Y è chiuso in X e $(x_n)_n$ è a valori in Y, pertanto, per la proposizione 2.23(vi), da $x_n \to x$ deduciamo che anche $x \in Y$. Siccome la norma su X e Y è la stessa, da (4.7) deduciamo che $x_n \to x$ in $(Y, \|\cdot\|)$.

Abbiamo dunque dimostrato che ogni successione di Cauchy $(x_n)_n$ in $(Y, \|\cdot\|)$ converge, cioè che $(Y, \|\cdot\|)$ è completo. □

Esempio 4.17. Sia $X = C[a, b]$ per qualche intervallo $[a, b]$ chiuso e limitato di $\mathbb{R}$. Dall'esempio 4.5(2) sappiamo che $(X, \|\cdot\|_\infty)$ è uno spazio di Banach. Sia

$$Y = \{p|_{[a,b]} : p \in \mathcal{P}\},$$

dove $\mathcal{P}$ è l'insieme dei polinomi (☞ esempio 1.2(4)). Allora Y è un sottospazio vettoriale di X (la verifica è lasciata per esercizio), ma $(Y, \|\cdot\|_\infty)$ non è di Banach.

Infatti, come vedremo nel teorema 4.30, Y è denso in X, cioè $\overline{Y}^X = X$. Se Y fosse chiuso in X, si avrebbe allora $Y = \overline{Y}^X = X$, il che è assurdo perché Y è un sottoinsieme proprio di X (non tutte le funzioni continue sono polinomi). Siccome Y non è chiuso in X, dalla proposizione 4.16 deduciamo che $(Y, \|\cdot\|_\infty)$ non è completo, dunque non è uno spazio di Banach.

4.4 Serie convergenti e assolutamente convergenti

Definizione 4.18. Siano $(X, \|\cdot\|)$ uno spazio normato, $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X, e $x \in X$. Diciamo che:

(a) la serie $\sum_{n=0}^\infty x_n$ *converge* a x se

$$\lim_{N\to\infty} \sum_{n=0}^N x_n = x$$

(in tal caso scriviamo anche

$$\sum_{n=0}^\infty x_n = x$$

e chiamiamo x la *somma della serie*);

(b) la serie $\sum_{n=0}^\infty x_n$ *converge assolutamente* se

$$\sum_{n=0}^\infty \|x_n\| < \infty.$$

Osservazione 4.19. Nel caso $X = \mathbb{R}$ con la norma euclidea, sappiamo che non tutte le serie convergenti sono assolutamente convergenti (ad esempio, $\sum_{n=1}(-1)^n/n$ converge ma non assolutamente), ma che la convergenza assoluta di una serie ne implica la convergenza. Quest'ultima implicazione si può dimostrare in più generali spazi normati sotto l'ipotesi di completezza.

Proposizione 4.20. *Siano $(X, \|\cdot\|)$ uno spazio di Banach e $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X. Se la serie $\sum_{n=0}^\infty x_n$ converge assolutamente, allora la serie converge a qualche $x \in X$.*

Dimostrazione. Sia $y_n = \sum_{k=0}^n x_k$ per ogni $n \in \mathbb{N}$; in altre parole, $(y_n)_n$ è la successione delle somme parziali della serie. Vogliamo dimostrare che $(y_n)_n$ converge in X. Siccome X è di Banach, è sufficiente mostrare che $(y_n)_n$ è una successione di Cauchy.

Per ogni $n, m \in \mathbb{N}$, se $n \geq m$ si ha

$$\|y_n - y_m\| = \left\| \sum_{k=m+1}^n x_k \right\| \leq \sum_{k=m+1}^n \|x_k\| \leq \sum_{k=m+1}^\infty \|x_k\|. \tag{4.8}$$

D'altra parte, siccome la serie $\sum_{n=0}^\infty x_n$ converge assolutamente, la serie delle norme $\sum_{k=0}^\infty \|x_k\|$ converge; dunque, per l'Algebra dei Limiti,

$$\sum_{k=m+1}^\infty \|x_k\| = \sum_{k=0}^\infty \|x_k\| - \sum_{k=0}^m \|x_k\| \to 0 \qquad \text{per } m \to \infty \tag{4.9}$$

(le code della serie convergente $\sum_{k=0}^{\infty} \|x_k\|$ tendono a zero); dunque, per ogni $\epsilon > 0$, esiste $N \in \mathbb{N}$ tale che

$$\sum_{k=m+1}^{\infty} \|x_k\| < \epsilon \qquad \text{per ogni } m > N. \tag{4.10}$$

Da (4.8) e (4.10) deduciamo allora che

$$\|y_n - y_m\| < \epsilon$$

per ogni $n \geq m > N$; scambiando il ruolo di m e n, si vede che la stessa disuguaglianza vale per ogni $n, m > N$. Siccome $\epsilon > 0$ era arbitrario, questo dimostra che $(y_n)_n$ è una successione di Cauchy in $(X, \|\cdot\|)$. □

Un'altra conseguenza della convergenza assoluta, nota per serie numeriche, è che la somma della serie non dipende dall'ordine in cui i termini vengono sommati. Anche questo fatto, di cui omettiamo la dimostrazione, si generalizza al contesto degli spazi normati.

Proposizione 4.21. *Siano $(X, \|\cdot\|)$ uno spazio normato e $(x_n)_n$ una successione a valori in X. Se $\sum_{n=0}^{\infty} x_n = x$ e la convergenza è assoluta, allora $\sum_{n=0}^{\infty} x_{\sigma(n)} = x$ per ogni biiezione $\sigma : \mathbb{N} \to \mathbb{N}$.*

Osservazione 4.22. È noto dall'analisi reale che, se una serie $\sum_{n=0}^{\infty} a_n$ a termini reali $a_n \in \mathbb{R}$ converge, ma non converge assolutamente, allora è possibile riordinarne i termini in modo da farla convergere a qualunque altro numero reale; in altre parole, nel caso di serie a termini reali, la convergenza assoluta equivale alla *convergenza incondizionata* (cioè l'invarianza della somma per riordinamento dei termini). La proposizione 4.21 ci dice che l'implicazione "convergenza assoluta ⇒ convergenza incondizionata" vale per serie convergenti in qualunque spazio normato X. Studiando gli spazi di Hilbert, vedremo tuttavia che l'implicazione inversa non vale in generale, dato che troveremo serie che convergono incondizionatamente senza convergere assolutamente (☞ osservazione 6.45).

4.5 Compattezza in spazi normati

Sia X uno spazio normato. Dalla proposizione 2.23(xi) sappiamo che ogni sottoinsieme compatto di X è chiuso e limitato. Per il corollario 4.14, l'implicazione opposta (ogni sottoinsieme chiuso e limitato di X è compatto) vale se $\dim X < \infty$. Vediamo ora che invece, in dimensione infinita, l'implicazione opposta fallisce.

Ricordiamo da (2.6) la definizione della distanza fra un punto e un sottoinsieme di uno spazio metrico.

Lemma 4.23 (lemma di Riesz). *Siano $(X, \|\cdot\|)$ uno spazio normato e $Y \subseteq X$ un sottospazio chiuso proprio (cioè $Y \neq X$). Allora esiste $w \in X$ tale che $\|w\| = 1$ e $d(w, Y) \geq 1/2$.*

Dimostrazione. Siccome Y è un sottoinsieme chiuso proprio di X, se prendo $x \in X \setminus Y$ si ha $d(x, Y) > 0$ per la proposizione 2.23(vii). Per la definizione (2.6) di $d(x, Y)$, esiste $z \in Y$ tale che

$$d(x, Y) \leq \|x - z\| < 2d(x, Y), \tag{4.11}$$

e in particolare $\|x-z\| > 0$. Sia $w = (x-z)/\|x-z\|$. Allora $\|w\| = 1$ e inoltre, per ogni $y \in Y$,

$$\|w-y\| = \left\| \frac{x-z}{\|x-z\|} - y \right\| = \frac{1}{\|x-z\|} \|x - (z + \|x-z\|y)\|; \tag{4.12}$$

ora $z + \|x-z\|y \in Y$, perché $z, y \in Y$ e Y è un sottospazio, pertanto

$$\|x - (z + \|x-z\|y)\| \geq d(x, Y) \tag{4.13}$$

e quindi, da (4.11)-(4.12)-(4.13) deduciamo che

$$\|w-y\| > \frac{d(x,Y)}{2d(x,Y)} = \frac{1}{2}.$$

Passando all'estremo inferiore su $y \in Y$ nella precedente disuguaglianza, otteniamo che $d(w, Y) \geq 1/2$, come desiderato. □

Osservazione 4.24. Una variante della precedente dimostrazione permette, per ogni dato $\epsilon > 0$, di trovare un $w \in X$ con $\|w\| = 1$ e $d(w, Y) \geq 1/(1+\epsilon)$.

Teorema 4.25. *Sia $(X, \|\cdot\|)$ uno spazio normato di dimensione infinita. Allora $\overline{B}(0,1)$ e $S(0,1) := \{x \in X : \|x\| = 1\}$ non sono compatti.*

Dimostrazione. Siccome $S(0,1) \subseteq \overline{B}(0,1)$, è sufficiente dimostrare che esiste una successione a valori in S che non ammette sottosuccessioni convergenti.

Più precisamente, costruiamo in maniera induttiva una successione $(x_n)_{n\in\mathbb{N}}$ a valori in $S(0,1)$ tale che

$$\|x_n - x_m\| \geq 1/2 \qquad \text{per ogni } n, m \in \mathbb{N}. \tag{4.14}$$

Ovviamente ogni sottosuccessione di una tale successione ha ancora la stessa proprietà (4.14); inoltre una successione che verifichi (4.14) non può essere di Cauchy, dunque non può convergere.

Per $n = 0$, prendiamo un qualunque $x_0 \in X$ con $\|x_0\| = 1$ (esiste perché $\dim X = \infty$).

Supponiamo ora di avere costruito per qualche $n \in \mathbb{N}$ i punti $x_0, x_1, \dots, x_n$ della successione, in modo che $\|x_j - x_k\| \geq 1/2$ per ogni $j, k = 0, \dots, n$ e $\|x_j\| = 1$ per $j = 0, \dots, n$. Allora $Y := \operatorname{span}\{x_0, \dots, x_n\}$ è un sottospazio vettoriale di X di dimensione finita, quindi è chiuso (per la proposizione 4.16(ii)) e proprio (perché $\dim X = \infty$). Dunque per il lemma 4.23 troviamo $x_{n+1} \in X$ con $\|x_{n+1}\| = 1$ e $d(x_{n+1}, Y) \geq 1/2$, e in particolare $\|x_{n+1} - x_j\| \geq d(x_{n+1}, Y) \geq 1/2$ per $j = 0, \dots, n$, perché $y_j \in Y$. Abbiamo così costruito un nuovo punto x_{n+1} della successione, mantenendo le proprietà richieste.

Procedendo induttivamente, si costruisce così una successione $(x_n)_{n\in\mathbb{N}}$ a valori in $S(0,1)$ che soddisfa (4.14). □

Corollario 4.26. *Sia $(X, \|\cdot\|)$ uno spazio normato. Sono equivalenti:*

(i) *ogni successione limitata in X ha una sottosuccessione convergente;*

(ii) *tutti i sottoinsiemi chiusi e limitati di X sono compatti in X;*

(iii) *$\overline{B}(0,1)$ è compatta in X;*

(iv) $\dim X < \infty$.

Dimostrazione. (i) $\Rightarrow$ (ii). Sia C chiuso e limitato in X. Allora ogni successione a valori in C è limitata, quindi per (i) ha una sottosuccessione convergente, e siccome C è chiuso il limite della successione è in C. Per l'arbitrarietà della successione, questo mostra che C è compatto.

(ii) $\Rightarrow$ (i). Sia $(x_n)_n$ una successione limitata. Allora esiste una palla chiusa $B = \overline{B}(\bar{x}, r)$ tale che $\{x_n\}_n \subseteq B$. Per (ii), la palla B è compatta, quindi $(x_n)_n$ ha una sottosuccessione convergente a un punto di B.

(ii) $\Rightarrow$ (iii). La palla $\overline{B}(0, 1)$ è un sottoinsieme limitato di X, dunque è compatta per (ii).

(iii) $\Rightarrow$ (iv). Questo segue dal teorema 4.25.

(iv) $\Rightarrow$ (ii). Questo segue dal corollario 4.14 □

Osservazione 4.27. Vedremo in seguito, con il teorema di Banach–Alaoglu (☞ teorema 9.31 e corollario 9.32), che in dimensione infinita si può a volte recuperare un analogo della proprietà (i) del corollario 4.26 utilizzando opportune nozioni di "convergenze deboli".

Osservazione 4.28. Sia $[a, b]$ un intervallo chiuso e limitato di $\mathbb{R}$. Siccome lo spazio $(C[a, b], \| \cdot \|_\infty)$ ha dimensione infinita, dal corollario 4.26 deduciamo che non tutte le successioni limitate di funzioni continue su $[a, b]$ hanno una sottosuccessione uniformemente convergente. Questo spiega la necessità di un'ulteriore condizione (la equicontinuità) nelle ipotesi del teorema di Ascoli–Arzelà (☞ teorema 2.38).

4.6 Il teorema di Stone–Weierstrass

Come discusso (☞ osservazione 4.28), per lo spazio di Banach $(C[a, b], \| \cdot \|_\infty)$ delle funzioni continue su un intervallo chiuso e limitato non vale l'analogo del teorema di Heine–Borel (☞ teorema 2.24) sulla caratterizzazione dei sottoinsiemi compatti di $(\mathbb{F}^n, \| \cdot \|_2)$. Una proprietà che invece rimane vera in $C[a, b]$ è la separabilità. Dedurremo questo risultato dal teorema di Stone–Weierstrass, che permette di approssimare arbitrarie funzioni continue su un intervallo $[a, b]$ con polinomi, nel senso della convergenza uniforme.

Premettiamo una caratterizzazione degli spazi normati separabili che sarà utile in seguito.

Proposizione 4.29. *Sia $(X, \| \cdot \|)$ uno spazio normato. Sono fatti equivalenti:*

(i) X è separabile;

(ii) esiste un sottoinsieme $E \subseteq X$ al più numerabile tale che $X = \overline{\operatorname{span} E}$.

Dimostrazione. (i) $\Rightarrow$ (ii). Se X è separabile, esiste un sottoinsieme $B \subseteq X$ al più numerabile tale che $X = \overline{B}$. Allora, a maggior ragione $X = \overline{\operatorname{span} B}$, e basta prendere $E = B$.

(ii) $\Rightarrow$ (i). Supponiamo esiste un insieme $E \subseteq X$ al più numerabile con $X = \overline{\operatorname{span} E}$. Se E è finito, allora X ha dimensione finita, e allora dal corollario 4.14 sappiamo che X è separabile.

Supponiamo invece che $E = \{e_n\}_{n\in\mathbb{N}}$ sia numerabile. Definiamo allora, per ogni $n \in \mathbb{N}$, lo spazio $X_n = \operatorname{span}\{e_0, \dots, e_n\}$. Allora ciascun X_n ha dimensione finita, quindi (con la norma indotta da X) è separabile, e possiamo trovare un sottoinsieme $B_n \subseteq X_n$ denso e al più numerabile. Allora anche l'unione $B = \bigcup_{n\in\mathbb{N}} B_n$ è al più numerabile, e inoltre

$$\overline{B} \supseteq \overline{B_n} \supseteq X_n \qquad \forall n \in \mathbb{N},$$

da cui deduciamo anche

$$\overline{B} \supseteq \bigcup_{n\in\mathbb{N}} X_n = \bigcup_{n\in\mathbb{N}} \operatorname{span}\{e_0, \dots, e_n\} = \operatorname{span} E$$

e infine, prendendo nuovamente la chiusura,

$$\overline{B} \supseteq \overline{\operatorname{span} E} = X,$$

cioè B è il sottoinsieme denso al più numerabile di X cercato. □

Procediamo con la dimostrazione del teorema di Stone–Weierstrass. Ricordiamo che, per una successione $(f_n)_n$ a valori in $C[a,b]$ e $f \in C[a,b]$,

$$f_n \rightrightarrows f \qquad \iff \qquad \|f_n - f\|_\infty \to 0$$

(☞ definizione 2.16), cioè la convergenza uniforme equivale alla convergenza nello spazio normato $(C[a,b], \|\cdot\|_\infty)$.

Teorema 4.30 (Stone–Weierstrass)**.** *Sia $[a,b]$ un intervallo chiuso e limitato in $\mathbb{R}$. Se $f \in C[a,b]$, allora esiste una successione di polinomi $(p_n)_{n\in\mathbb{N}}$ in $\mathcal{P}$ tale che $p_n|_{[a,b]} \rightrightarrows f$.*

Dimostrazione. Passo 1. Possiamo assumere $[a,b] = [0,1]$.

La riduzione da $[a,b]$ a $[0,1]$ si effettua tramite il cambio di variabili affine $\Phi : [0,1] \to [a,b]$ dato da $\Phi(t) = a + (b-a)t$. Tale cambio di variabili conserva la classe dei polinomi e la convergenza uniforme di successioni di funzioni.

Passo 2. Possiamo anche assumere $f(0) = f(1) = 0$.

Possiamo definire un polinomio di primo grado $q(t) = f(0) + t(f(1) - f(0))$ con gli stessi valori di f nei punti 0 e 1; pertanto la funzione $f_* = f - q|_{[0,1]}$ è continua su $[0,1]$ e $f_*(0) = f_*(1) = 0$. Se possiamo approssimare uniformemente f_* con una successione di polinomi p_n su $[0,1]$, allora

$$\|f - (p_n + q)|_{[0,1]}\|_\infty = \|f_* - p_n|_{[0,1]}\|_\infty \to 0,$$

cioè f è approssimata uniformemente dalla successione di polinomi $p_n + q$.

Passo 3. Per ogni $n \in \mathbb{N}_+$, $C_n := \int_{-1}^1 (1-x^2)^n \, dx > 1/\sqrt{n}$.

Notiamo che $\frac{d}{dx}[(1-x^2)^n - (1-nx^2)] = 2nx[1-(1-x^2)^{n-1}] \geq 0$ per ogni $x \in [0,1]$. Di conseguenza $x \mapsto (1-x^2)^n - (1-nx^2)$ è crescente su $[0,1]$, e siccome si annulla in $x = 0$, ne concludiamo che

$$1 - nx^2 \leq (1-x^2)^n$$

per ogni $x \in [0,1]$; per parità, la stessa disuguaglianza in effetti vale per ogni $x \in [-1,1]$. Ma allora

$$C_n \geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} (1-nx^2)\, dx = \frac{2}{\sqrt{n}} \int_0^1 (1-s^2)\, ds = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}}.$$

Passo 4. Per $n \in \mathbb{N}_+$, poniamo $Q_n(t) = \frac{1}{C_n}(1-t^2)^n$. Allora $Q_n \in \mathcal{P}$ e

$$Q_n|_{[-1,1]} \geq 0, \qquad \int_{-1}^{1} Q_n(t)\,dt = 1.$$

Inoltre, per ogni $\delta \in (0,1)$,

$$|Q_n(t)| = C_n^{-1}(1-t^2)^n \leq \sqrt{n}(1-\delta^2)^n \qquad \text{se } \delta \leq |t| \leq 1.$$

Passo 5. Sia $f \in C_{\mathbb{F}}[0,1]$ con $f(0) = f(1) = 0$. Estendiamo f a tutto $\mathbb{R}$ ponendo $f(t) = 0$ per $t \in \mathbb{R} \setminus [0,1]$. Allora $f \in C_{\mathbb{F}}(\mathbb{R})$. Definiamo, per ogni $n \in \mathbb{N}_+$ e $t \in \mathbb{R}$,

$$p_n(t) = f * Q_n(t) = \int_{\mathbb{R}} f(s)Q_n(t-s)\,ds = \int_0^1 f(s)Q_n(t-s)\,ds.$$

Allora possiamo scrivere $Q_n(t-s) = \sum_{k=0}^{2n} q_{n,k}(s)t^k$ per qualche $q_{n,k} \in \mathcal{P}$, dunque

$$p_n(t) = \sum_{k=0}^{2n} \left(\int_0^1 f(s)q_{n,k}(s)\,ds \right) t^k$$

e quindi anche $p_n \in \mathcal{P}$.

Inoltre, se $t \in [0,1]$, si ha

$$p_n(t) = \int_{\mathbb{R}} f(t-s)Q_n(s)\,ds = \int_{-1}^{1} f(t-s)Q_n(s)\,ds,$$

ove si è usato che f si annulla fuori da $[0,1]$. Siccome $\int_{-1}^1 Q_n(s)\,ds = 1$, si ha

$$p_n(t) - f(t) = \int_{-1}^{1} [f(t-s) - f(t)]Q_n(s)\,ds;$$

dunque, per ogni $\delta \in (0,1)$,

$$\begin{aligned}
|p_n(t) - f(t)| &\leq \int_{-1}^{1} |f(t-s) - f(t)|Q_n(s)\,ds \\
&= \int_{|s|<\delta} + \int_{\delta \leq |s| \leq 1} \\
&\leq \sup_{|s|<\delta} |f(t-s) - f(t)| \int_{-1}^{1} Q_n(s)\,ds \\
&\qquad + \sup_{\delta \leq |s| \leq 1} Q_n(s) \int_{-1}^{1} |f(t-s) - f(t)|\,dt \\
&\leq \sup_{|s|<\delta} |f(t-s) - f(t)| + 4 \sup_{s \in \mathbb{R}} |f(s)| \sqrt{n}(1-\delta^2)^n,
\end{aligned}$$

ove si sono usate le proprietà di Q_n enunciate nel passo 4. In conclusione, per ogni $\delta \in (0,1)$,

$$\|p_n|_{[0,1]} - f\|_\infty \leq \sup_{t \in [0,1], |s|<\delta} |f(t-s) - f(t)| + 4 \sup_{s \in [0,1]} |f(s)| \sqrt{n}(1-\delta^2)^n.$$

ove si è ancora usato che f si annulla fuori da $[0,1]$.

Notiamo ora che $f|_{[0,1]}$ è limitata per il teorema di Weierstrass (☞ teorema 2.36), dunque $\sup_{s\in[0,1]} |f(s)| \leq M$ per qualche $M \geq 0$. Inoltre $f|_{[-2,2]}$ è uniformemente continua per il teorema di Heine–Cantor (☞ teorema 2.32). Pertanto, dato $\epsilon > 0$, esiste $\delta \in (0,1)$ tale che $|f(x) - f(y)| < \epsilon/2$ per ogni $x, y \in [-2,2]$ con $|x - y| < \delta$. Dalla precedente stima deduciamo allora che

$$\|p_n|_{[0,1]} - f\|_\infty \leq \epsilon/2 + 4M\sqrt{n}(1-\delta^2)^n;$$

siccome $\delta \in (0,1)$, si ha $4M\sqrt{n}(1-\delta^2)^n \to 0$ per $n \to \infty$, dunque esiste $N \in \mathbb{N}$ abbastanza grande per cui $4M\sqrt{n}(1-\delta^2)^n < \epsilon/2$ per ogni $n > N$, ma allora

$$\|p_n|_{[0,1]} - f\|_\infty < \epsilon/2 + \epsilon/2 = \epsilon$$

per ogni $n > N$. Questo dimostra che $p_n|_{[0,1]} \rightrightarrows f$ per $n \to \infty$. □

Corollario 4.31. *Sia $[a,b]$ un intervallo chiuso e limitato in $\mathbb{R}$. Lo spazio normato $(C[a,b], \|\cdot\|_\infty)$ è separabile.*

Dimostrazione. Il teorema di Stone–Weierstrass (☞ teorema 4.30) ci dice che l'insieme

$$S = \{p|_{[a,b]} : p \in \mathcal{P}\}$$

è denso in $C[a,b]$, cioè $\overline{S} = C[a,b]$. D'altra parte, se $p_n(t) = t^n$ è il monomio monico di grado n, allora

$$\mathcal{P} = \operatorname{span}\{p_n\}_{n\in\mathbb{N}}$$

(☞ esempio 1.16(4)) e quindi anche

$$S = \operatorname{span}\{p_n|_{[a,b]}\}_{n\in\mathbb{N}}.$$

La separabilità di $C[a,b]$ segue dunque dalla proposizione 4.29 applicata all'insieme $E = \{p_n|_{[a,b]}\}_{n\in\mathbb{N}}$. □

5 Spazi di successioni e spazi L^p

5.1 Spazi di successioni

Ricordiamo che $\mathbb{F}$ denota il campo $\mathbb{R}$ dei numeri reali oppure il campo $\mathbb{C}$ dei numeri complessi.

Scriviamo $\mathbb{F}^{\mathbb{N}}$ per denotare l'insieme delle successioni indicizzate su $\mathbb{N}$ a valori in $\mathbb{F}$. In simboli,
$$\mathbb{F}^{\mathbb{N}} = \{\underline{x} = (x_k)_{k\in\mathbb{N}} : x_k \in \mathbb{F}\ \forall k \in \mathbb{N}\}.$$
Ricordiamo dalla definizione 2.3 che una successione a valori in $\mathbb{F}$ non è altro che una funzione da $\mathbb{N}$ a $\mathbb{F}$. Pertanto, con la notazione dell'esempio 1.2(7), si ha $\mathbb{F}^{\mathbb{N}} = \mathcal{F}(\mathbb{N},\mathbb{F})$. In particolare, $\mathbb{F}^{\mathbb{N}}$ è naturalmente dotato di una struttura di spazio vettoriale su $\mathbb{F}$, con le *operazioni componente per componente*
$$\underline{x} + \underline{y} = (x_k + y_k)_{k\in\mathbb{N}}, \qquad \alpha\underline{x} = (\alpha x_k)_{k\in\mathbb{N}} \tag{5.1}$$
per ogni $\underline{x} = (x_k)_{k\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$, $\underline{y} = (y_k)_{k\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ e $\alpha \in \mathbb{F}$. Queste operazioni non sono altro che le operazioni puntuali (1.2) su $\mathcal{F}(\mathbb{N},\mathbb{F})$, riscritte nella notazione delle successioni. In particolare, l'elemento neutro della somma in $\mathbb{F}^{\mathbb{N}}$ è la successione costante nulla, denotata con $\underline{0}$.

Osservazione 5.1. Preferiamo utilizzare la notazione delle successioni, anziché quella delle funzioni, quando lavoriamo con $\mathbb{F}^{\mathbb{N}}$, al fine di enfatizzarne la relazione con gli spazi $\mathbb{F}^n$ per $n \in \mathbb{N}_+$, discussi nell'esempio 1.2(1); si confrontino, ad esempio, le formule (1.1) e (5.1) per le relative operazioni. Possiamo dunque pensare a $\mathbb{F}^{\mathbb{N}}$ come a un analogo di $\mathbb{F}^n$, in cui, invece di considerare vettori con n componenti numeriche, si considerano vettori con *infinite* componenti. In effetti, sulla base dell'esempio 1.16, si ha $\dim \mathbb{F}^n = n$ e $\dim \mathbb{F}^{\mathbb{N}} = \infty$.

Abbiamo visto (☞ esempio 4.5(1)) che su $\mathbb{F}^n$ per $n \in \mathbb{N}_+$ possiamo introdurre una famiglia di norme $\|\cdot\|_p$, per $p \in [1,\infty]$. Una costruzione analoga si può fare su $\mathbb{F}^{\mathbb{N}}$; tuttavia, in questo caso, le quantità $\|\cdot\|_p$ che definiremo non saranno norme su tutto lo spazio $\mathbb{F}^{\mathbb{N}}$, ma solo su opportuni sottospazi. Cominciamo per semplicità a discutere il caso $p = \infty$.

Definizione 5.2. Per ogni $\underline{x} = (x_k)_{k\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$, definiamo
$$\|\underline{x}\|_\infty = \sup_{k\in\mathbb{N}} |x_k|.$$
Denotiamo inoltre con ℓ^∞ il sottoinsieme di $\mathbb{F}^{\mathbb{N}}$ costituito dalle *successioni limitate*, cioè poniamo
$$\ell^\infty = \{\underline{x} \in \mathbb{F}^{\mathbb{N}} : \|\underline{x}\|_\infty < \infty\}.$$

Osservazione 5.3. Notiamo che, se $\underline{x} \in \mathbb{F}^{\mathbb{N}}$ è una successione illimitata (ad esempio, $\underline{x} = (2^n)_{n\in\mathbb{N}}$), allora $\|\underline{x}\|_\infty = \infty$ (☞ proposizione 4.6(iv)). In particolare, la funzione $\|\cdot\|_\infty$ su $\mathbb{F}^{\mathbb{N}}$ può assumere il valore ∞, e pertanto non è una norma su $\mathbb{F}^{\mathbb{N}}$ nel senso della definizione 4.1. Per comodità, con un abuso di linguaggio, chiameremo tuttavia $\|\cdot\|_\infty$ la *norma infinito*; questo uso è almeno in parte giustificato dal seguente risultato.

Proposizione 5.4. *ℓ^∞ è un sottospazio vettoriale di $\mathbb{F}^{\mathbb{N}}$ e $\|\cdot\|_\infty$ è una norma su ℓ^∞.*

Dimostrazione. Chiaramente $\underline{0}$ è in ℓ^∞. Inoltre, se $\underline{x}, \underline{y} \in \ell^\infty$, allora

$$\begin{aligned} \|\underline{x}+\underline{y}\|_\infty = \sup_{k\in\mathbb{N}} |x_k + y_k| &\leq \sup_{k\in\mathbb{N}}(|x_k| + |y_k|) \\ &\leq \sup_k |x_k| + \sup_k |y_k| = \|\underline{x}\|_\infty + \|\underline{y}\|_\infty < \infty, \end{aligned} \tag{5.2}$$

pertanto anche $\underline{x}+\underline{y} \in \ell^\infty$. Similmente, se $\alpha \in \mathbb{F}$ e $\underline{x} \in \ell^\infty$, allora

$$\|\alpha\underline{x}\|_\infty = \sup_{k\in\mathbb{N}} |\alpha x_k| = \sup_{k\in\mathbb{N}}(|\alpha|\,|x_k|) = |\alpha| \sup_{k\in\mathbb{N}} |x_k| = |\alpha|\,\|\underline{x}\|_\infty < \infty, \tag{5.3}$$

da cui segue che anche $\alpha\underline{x} \in \ell^\infty$. Per la proposizione 1.4, concludiamo che ℓ^∞ è un sottospazio vettoriale di $\mathbb{F}^\mathbb{N}$.

Vediamo ora che $\|\cdot\|_\infty$ ristretta a ℓ^∞ è una norma nel senso della definizione 4.1. Per definizione di ℓ^∞ si ha che $\|\cdot\|_\infty$ su ℓ^∞ prende valori in $[0,\infty)$. Inoltre, chiaramente $\|\underline{0}\|_\infty = 0$; viceversa, se $\|\underline{x}\|_\infty = 0$ per qualche $\underline{x} = (x_k)_{k\in\mathbb{N}} \in \ell^\infty$, allora $\sup_k |x_k| = 0$, pertanto $x_k = 0$ per ogni $k \in \mathbb{N}$ e quindi $\underline{x} = \underline{0}$. Infine, da (5.2) e (5.3) deduciamo la disuguaglianza triangolare e la 1-omogeneità. □

Come vedremo, lo spazio normato $(\ell^\infty, \|\cdot\|_\infty)$ è completo, dunque uno spazio di Banach. Conviene discutere di questa proprietà insieme alla discussione di certi sottospazi di ℓ^∞.

Definizione 5.5. Denotiamo con c_0 il sottoinsieme di $\mathbb{F}^\mathbb{N}$ formato dalle *successioni infinitesime*, cioè poniamo

$$c_0 = \left\{ \underline{x} = (x_k)_{k\in\mathbb{N}} \in \mathbb{F}^\mathbb{N} : \lim_{k\to\infty} x_k = 0 \right\}.$$

Inoltre, denotiamo con c_{00} il sottoinsieme di $\mathbb{F}^\mathbb{N}$ formato dalle *successioni definitvamente nulle*, cioè poniamo

$$c_{00} = \{\underline{x} = (x_k)_{k\in\mathbb{N}} \in \mathbb{F}^\mathbb{N} : \exists N \in \mathbb{N} : \forall k > N : x_k = 0\}.$$

Proposizione 5.6. *Valgono le seguenti proprietà.*

(i) $(\ell^\infty, \|\cdot\|_\infty)$ *è uno spazio di Banach.*

(ii) c_0 *è un sottospazio vettoriale chiuso di* $(\ell^\infty, \|\cdot\|_\infty)$. *Dunque* $(c_0, \|\cdot\|_\infty)$ *è a sua volta uno spazio di Banach.*

(iii) c_{00} *è un sottospazio vettoriale denso e non chiuso di* $(c_0, \|\cdot\|_\infty)$.

Dimostrazione. (i). Confrontando la definizione 5.2 con la formula (2.5), notiamo che $\ell^\infty = \mathcal{F}_b(\mathbb{N}, \mathbb{F})$, e che la norma $\|\cdot\|_\infty$ induce la distanza d_∞ di (2.3). La completezza di ℓ^∞ è dunque un caso particolare del teorema 2.15.

(ii). Siccome ogni successione convergente è limitata, si ha $c_0 \subseteq \ell^\infty$. Inoltre, chiaramente $\underline{0}$ è infinitesima, e dall'Algebra dei Limiti, si deduce che combinazioni lineari di successioni infinitesime sono infinitesime; questo, per la proposizione 1.4, dimostra che c_0 è un sottospazio vettoriale di ℓ^∞.

Rimane da dimostrare che c_0 è chiuso in ℓ^∞; infatti, una volta stabilito che c_0 è chiuso, il fatto che $(c_0, \|\cdot\|_\infty)$ sia uno spazio di Banach segue da (i) e dalla proposizione 4.16. Sulla base della proposizione 2.23(vi), per verificare che c_0 è chiuso in ℓ^∞, è sufficiente dimostrare che, se $(\underline{x}^{(n)})_n$ è una successione a valori in c_0 che converge in ℓ^∞ a un punto $\underline{x} \in \ell^\infty$, allora anche $\underline{x} \in c_0$.

Sia $\epsilon > 0$. Siccome $\underline{x}^{(n)} \to \underline{x}$ in ℓ^∞ per $n \to \infty$, posso trovare $N \in \mathbb{N}$ tale che $\|\underline{x}^{(n)} - \underline{x}\|_\infty < \epsilon/2$ per ogni $n > N$. Sia $\bar{n} = N + 1$. Siccome $\underline{x}^{\bar{n}} \in c_0$, si ha $\lim_{k\to\infty} x_k^{\bar{n}} = 0$, dunque esiste $K \in \mathbb{N}$ tale che $|x_k^{(\bar{n})}| < \epsilon/2$ per ogni $k > K$. Ma allora, per ogni $k > K$,

$$|x_k| \leq |x_k - x_k^{(\bar{n})}| + |x_k^{(\bar{n})}| \leq \|\underline{x} - \underline{x}^{(\bar{n})}\|_\infty + |x_k^{(\bar{n})}| < \epsilon/2 + \epsilon/2 = \epsilon;$$

per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $|x_k| \to 0$ per $k \to \infty$, cioè $\underline{x}$ è infinitesima e dunque appartiene a c_0.

(iii). Chiaramente ogni successione definitivamente nulla è infinitesima, cioè $c_{00} \subseteq c_0$. Inoltre, $\underline{0}$ è definitivamente nulla, e non è difficile verificare (lasciamo la verifica per esercizio) che combinazioni lineari di successioni definitivamente nulle sono a loro volta definitivamente nulle. Pertanto c_{00} è un sottospazio vettoriale di c_0.

Rimane solo da dimostrare che c_{00} è denso in c_0. Infatti, una volta stabilita la densità, si ha $\overline{c_{00}} = c_0$; se per assurdo c_{00} fosse chiuso, si avrebbe $c_{00} = \overline{c_{00}} = c_0$, ma questo è assurdo perché perché c_{00} è un sottospazio proprio di c_0 (esistono successioni infinitesime che non sono definitivamente nulle, come ad esempio $(1/(1+k))_{k\in\mathbb{N}}$).

Per verificare che c_{00} è denso in c_0, sulla base della proposizione 2.23(ix), è sufficiente dimostrare che ogni $\underline{x} \in c_0$ è limite in $(c_0, \|\cdot\|_\infty)$ di una successione a valori in c_{00}. Sia dunque $\underline{x} \in c_0$, e consideriamo, per ogni $n \in \mathbb{N}$, la troncata n-sima $\underline{x}^{(n)}$ di $\underline{x}$, definita da

$$x_k^{(n)} = \begin{cases} x_k & \text{se } k \leq n, \\ 0 & \text{se } k > n \end{cases} \tag{5.4}$$

per ogni $k \in \mathbb{N}$. Per costruzione, $\underline{x^{(n)}} \in c_{00}$. Inoltre, sempre per costruzione, le componenti della differenza $\underline{x} - \underline{x^{(n)}}$ con indice $k \leq n$ sono tutte nulle, pertanto

$$\|\underline{x} - \underline{x^{(n)}}\|_\infty = \sup_{k\in\mathbb{N}} |x_k - x_k^{(n)}| = \sup_{k>n} |x_k - 0| = \sup_{k>n} |x_k| \to 0$$

per $n \to \infty$, dato che $|x_k| \to 0$ per $k \to \infty$. Questo mostra che $\underline{x^{(n)}} \to \underline{x}$ rispetto alla norma $\|\cdot\|_\infty$, come desiderato. □

Passiamo ora a discutere l'analogo su $\mathbb{F}^\mathbb{N}$ delle norme $\|\cdot\|_p$ su $\mathbb{F}^n$ (☞ esempio 4.5(1)) nel caso $p < \infty$.

Definizione 5.7. Sia $p \in [1, \infty)$. Per ogni $\underline{x} = (x_k)_{k\in\mathbb{N}} \in \mathbb{F}^\mathbb{N}$, definiamo

$$\|\underline{x}\|_p = \left(\sum_{k=0}^{\infty} |x_k|^p \right)^{1/p}, \tag{5.5}$$

ove si conviene che $\|\underline{x}\|_p = \infty$ quando la serie nel membro destro di (5.5) diverge. Denotiamo inoltre con ℓ^p il sottoinsieme di $\mathbb{F}^\mathbb{N}$ costituito dalle *successioni p-sommabili*, cioè poniamo

$$\ell^p = \{\underline{x} \in \mathbb{F}^\mathbb{N} : \|\underline{x}\|_p < \infty\}.$$

Osservazione 5.8. È utile fare qualche commento sulla definizione di successioni p-sommabili.

- Le successioni 1-sommabili sono anche dette semplicemente successioni *sommabili*. In altre parole, una successione $\underline{x} = (x_n)_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ è sommabile se e solo se la serie $\sum_{n=0}^{\infty} x_n$ è assolutamente convergente in $\mathbb{F}$ (☞ definizione 4.18(b)), cioè se e solo se la serie dei moduli $\sum_{n\in\mathbb{N}} |x_n|$ converge.

- Più in generale, per $p \in [1,\infty)$, una successione $\underline{x} = (x_n)_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ è p-sommabile se e solo se la successione $(|x_n|^p)_{n\in\mathbb{N}}$ delle potenze p-sime dei moduli è sommabile.

- Quando $p \in \mathbb{N}_+$, si ha $|x_n|^p = |x_n^p|$, dunque $(x_n)_n$ è p-sommabile se e solo se $(x_n^p)_n$ è sommabile.

- In particolare, nel caso $p = 2$, gli elementi di ℓ^2, cioè le successioni 2-sommabili, sono anche dette successioni *a quadrato sommabile*.

Osservazione 5.9. Anche nel caso $p \in [1,\infty)$, la funzione $\|\cdot\|_p$ può assumere valore ∞ su $\mathbb{F}^{\mathbb{N}}$, quindi non è propriamente una norma su $\mathbb{F}^{\mathbb{N}}$ nel senso della definizione 4.1. Ciononostante useremo spesso il nome di *norma p-esima* (o *norma p*, o *norma ℓ^p*) per riferirci ad essa.

Proposizione 5.10. *Per ogni $p \in [1,\infty)$, ℓ^p è un sottospazio vettoriale di $\mathbb{F}^{\mathbb{N}}$. Inoltre, per ogni $\underline{x} \in \mathbb{F}^{\mathbb{N}}$, si ha $\|\underline{x}\|_p = 0$ se e solo se $\underline{x} = \underline{0}$.*

Dimostrazione. Per dimostrare che ℓ^p è un sottospazio, verifichiamo le proprietà indicate nella proposizione 1.4.

Chiaramente $\|\underline{0}\|_p = 0$, dunque $\underline{0} \in \ell^p$ e ℓ^p non è vuoto.

Inoltre, se $\underline{x} \in \ell^p$ e $\alpha \in \mathbb{F}$, si ha

$$\|\alpha\underline{x}\|_p^p = \sum_{n=0}^{\infty} |\alpha x_n|^p = \sum_{n=0}^{\infty} |\alpha|^p |x_n|^p = |\alpha|^p \sum_{n=0}^{\infty} |x_n|^p,$$

da cui segue che

$$\|\alpha\underline{x}\|_p = |\alpha| \|\underline{x}\|_p < \infty \tag{5.6}$$

siccome $\underline{x} \in \ell^p$, e quindi anche $\alpha\underline{x} \in \ell^p$.

Infine, siano $\underline{x}, \underline{y} \in \ell^p$. Notiamo che, per ogni $k \in \mathbb{N}$,

$$|x_k + y_k| \le |x_k| + |y_k| \le 2\max\{|x_k|, |y_k|\},$$

pertanto

$$|x_k + y_k|^p \le (2\max\{|x_k|, |y_k|\})^p = 2^p \max\{|x_k|^p, |y_k|^p\} \le 2^p(|x_k|^p + |y_k|^p). \tag{5.7}$$

Dunque

$$\|\underline{x}+\underline{y}\|_p^p = \sum_{k=0}^{\infty} |x_k + y_k|^p \le 2^p \left(\sum_{k=0}^{\infty} |x_k|^p + \sum_{k=0}^{\infty} |y_k|^p \right) = 2^p(\|\underline{x}\|_p^p + \|\underline{y}\|_p^p) < \infty \tag{5.8}$$

e quindi $\underline{x} + \underline{y} \in \ell^p$.

Questo dimostra che ℓ^p è un sottospazio vettoriale di $\mathbb{F}^{\mathbb{N}}$. Inoltre sappiamo già che $\|\underline{0}\|_p = 0$. Viceversa, sia $\underline{x} \in \mathbb{F}^{\mathbb{N}}$ tale che $\|\underline{x}\|_p = 0$; allora

$$0 = \|\underline{x}\|_p^p = \sum_{k=0} |x_k|^p.$$

Siccome gli addendi di quest'ultima somma sono tutti nonnegativi, se la somma si annulla tutti gli addendi devono essere nulli. Ne concludiamo che $|x_k|^p = 0$ per ogni $k \in \mathbb{N}$, dunque $x_k = 0$ per ogni $k \in \mathbb{N}$, cioè $\underline{x} = \underline{0}$, come desiderato. ☐

Per giustificare il nome di "norma p" dato a $\|\cdot\|_p$, anche nel caso $p \in [1, \infty)$ vorremmo dimostrare che $\|\cdot\|_p$ è una norma su ℓ^p (☞ definizione 4.1), come fatto per $p = \infty$ nella proposizione 5.4. In effetti, l'enunciato della proposizione 5.10 ci dà la proprietà di separazione per $\|\cdot\|_p$. Inoltre, esaminandone la dimostrazione, vediamo che (5.6) dimostra la 1-omogeneità; tuttavia, (5.8) dà una proprietà più debole della disuguaglianza triangolare.

In effetti, nel caso $p = 1$, si potrebbe rimpiazzare (5.7) con la stima più semplice $|x_k + y_k| \leq |x_k| + |y_k|$ per ottenere, al posto di (5.8), la disuguaglianza triangolare per $\|\cdot\|_1$. La dimostrazione della disuguaglianza triangolare per la norma $\|\cdot\|_p$ per $p \in (1, \infty)$ è più sottile, e richiede un po' di lavoro preparatorio.

Definizione 5.11. Sia $p \in [1, \infty]$. L'elemento $q \in [1, \infty]$ tale che

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{5.9}$$

è detto *esponente coniugato* di p (qui si intende che $1/\infty = 0$). In tal caso, diciamo anche che p e q sono *esponenti coniugati*.

Osservazione 5.12. Da (5.9) è chiaro che, dato $p \in [1, \infty]$, esiste un unico $q \in [1, \infty]$ coniugato a p. Inoltre p è coniugato a q se e solo se q è coniugato a p. In particolare, 1 è coniugato a ∞, mentre 2 è coniugato a 2.

Proposizione 5.13 (disuguaglianza di Young)**.** *Siano $p, q \in (1, \infty)$ esponenti coniugati. Allora, per ogni $a, b \in [0, \infty)$,*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Dimostrazione (cenno). Dalla definizione di esponenti coniugati si ricava facilmente che

$$p - 1 = \frac{1}{q - 1}.$$

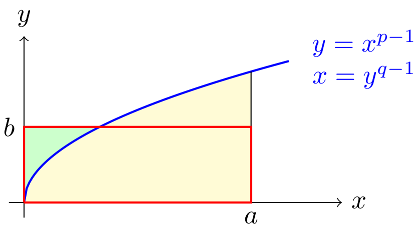
In particolare, per ogni $x, y > 0$ si ha

$$y = x^{p-1} \iff x = y^{q-1}.$$

Nella figura 1, l'area gialla rappresenta l'area della regione compresa fra la curva $y = x^{p-1}$, l'asse delle ascisse e le rette $x = 0$ e $x = a$, cioè

$$\int_0^a x^{p-1}\,dx = \frac{a^p}{p};$$

Figura 1: disuguaglianza di Young (proposizione 5.13)



similmente, l'area verde rappresenta l'area della regione compresa fra la curva $x = y^{q-1}$, l'asse delle ordinate e le rette $y = 0$ e $y = b$, data da

$$\int_0^b y^{q-1}\,dy = \frac{b^q}{q}.$$

Dalla figura si vede che tali regioni includono il rettangolo di base a e altezza b tracciato in rosso. Dunque la somma delle aree è maggiore o uguale dell'area ab del rettangolo. □

Proposizione 5.14 (disuguaglianza di Hölder). *Siano $p, q \in [1,\infty]$ esponenti coniugati. Siano $\underline{x} \in \ell^p$ e $\underline{y} \in \ell^q$. Allora*

$$\sum_{k=0}^{\infty} |x_k y_k| \le \|\underline{x}\|_p \|\underline{y}\|_q. \tag{5.10}$$

In particolare, la serie $\sum_{k=0}^{\infty} x_k y_k$ converge assolutamente in $\mathbb{F}$ e

$$\left|\sum_{k=0}^{\infty} x_k y_k\right| \le \|\underline{x}\|_p \|\underline{y}\|_q. \tag{5.11}$$

Dimostrazione. La disuguaglianza (5.11) è un'immediata conseguenza di (5.10) e della teoria delle serie numeriche; basta dunque dimostrare (5.10).

Nel caso $p = 1$ e $q = \infty$, si ha

$$\sum_{k=0}^{\infty} |x_k y_k| \le \sum_{k=0}^{\infty} |x_k| \sup_{h\in\mathbb{N}} |y_h| = \|\underline{y}\|_\infty \sum_{k=0}^{\infty} |x_k| = \|\underline{x}\|_1 \|\underline{y}\|_\infty;$$

in maniera analoga si procede quando $p = \infty$ e $q = 1$. Nel seguito, possiamo dunque supporre che $p, q \in (1,\infty)$.

Consideriamo anzitutto il caso in cui $\|\underline{x}\|_p = 1$ e $\|\underline{y}\|_q = 1$. In questo caso, per la disuguaglianza di Young (☞ proposizione 5.13),

$$\sum_{k=0}^{\infty} |x_k y_k| \le \sum_{k=0}^{\infty} \left(\frac{|x_k|^p}{p} + \frac{|y_k|^q}{q} \right) = \frac{1}{p}\|\underline{x}\|_p^p + \frac{1}{q}\|\underline{y}\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 = \|\underline{x}\|_p \|\underline{y}\|_q,$$

dove abbiamo usato che $\|\underline{x}\|_p = \|\underline{y}\|_q = 1$ e che p e q sono esponenti coniugati.

Supponiamo ora che $\|\underline{x}\|_p = 0$ oppure $\|\underline{y}\|_q = 0$. In questo caso, per la proposizione 5.10, si ha $\underline{x} = \underline{0}$ oppure $\underline{y} = \underline{0}$, da cui segue che anche $x_k y_k = 0$ per ogni $k \in \mathbb{N}$; pertanto, in questo caso, entrambi i membri di (5.10) si annullano e la disuguaglianza è banalmente verificata.

Rimane da considerare il caso in cui $\|\underline{x}\|_p \neq 0$ e $\|\underline{y}\|_q \neq 0$. In questo caso, possiamo normalizzare $\underline{x}$ e $\underline{y}$, cioè definire $\underline{x}' = \underline{x}/\|\underline{x}\|_p$ e $\underline{y}' = \underline{y}/\|\underline{y}\|_q$. Da (5.6) segue allora che $\|\underline{x}'\|_p = 1 = \|\underline{y}'\|_q$. Pertanto possiamo applicare a $\underline{x}'$ e $\underline{y}'$ il caso particolare già dimostrato della disuguaglianza di Hölder e ottenere che

$$\sum_{k=0}^{\infty} \left| \frac{x_k}{\|\underline{x}\|_p} \frac{y_k}{\|\underline{y}\|_q} \right| \leq 1;$$

moltiplicando ambo i membri di quest'ultima disuguaglianza per $\|\underline{x}\|_p \|\underline{y}\|_q$ si ottiene infine (5.10). □

Proposizione 5.15 (disuguaglianza di Minkowski)**.** *Sia $p \in [1, \infty]$. Allora, per ogni $\underline{x}, \underline{y} \in \mathbb{F}^{\mathbb{N}}$,*

$$\|\underline{x} + \underline{y}\|_p \leq \|\underline{x}\|_p + \|\underline{y}\|_p. \tag{5.12}$$

Dimostrazione. Come già osservato, nel caso $p = \infty$ la disuguaglianza è dimostrata in (5.2). Inoltre, nel caso $p = 1$, procedendo come in (5.8) ma usando la stima $|x_k + y_k| \leq |x_k| + |y_k|$ al posto di (5.7), si ottiene

$$\|\underline{x} + \underline{y}\|_1 = \sum_{k=0}^{\infty} |x_k + y_k| \leq \sum_{k=0}^{\infty} |x_k| + \sum_{k=0}^{\infty} |y_k| = \|\underline{x}\|_1 + \|\underline{y}\|_1.$$

Nel seguito possiamo dunque assumere che $p \in (1, \infty)$.

Nel caso $\|\underline{x}\|_p = \infty$ oppure $\|\underline{y}\|_p = \infty$, la disuguaglianza (5.12) vale banalmente perché il secondo membro è ∞. In maniera simile, se $\|\underline{x} + \underline{y}\|_p = 0$, la disuguaglianza (5.12) è banalmente verificata.

Possiamo dunque supporre che $\|\underline{x}\|_p < \infty$ e $\|\underline{y}\|_p < \infty$, cioè $\underline{x}, \underline{y} \in \ell^p$, e che inoltre $\|\underline{x} + \underline{y}\|_p \neq 0$. Dalla proposizione 5.10 deduciamo che anche $\underline{x} + \underline{y} \in \ell^p$, dunque $\|\underline{x} + \underline{y}\|_p \in (1, \infty)$.

Scriviamo ora

$$\begin{aligned} \|\underline{x} + \underline{y}\|_p^p = \sum_{k=0}^{\infty} |x_k + y_k|^p &= \sum_{k=0}^{\infty} |x_k + y_k|^{p-1} |x_k + y_k| \\ &\leq \sum_{k=0}^{\infty} |x_k + y_k|^{p-1} |x_k| + \sum_{k=0}^{\infty} |x_k + y_k|^{p-1} |y_k|. \end{aligned} \tag{5.13}$$

Sia $q = p/(p-1) \in (1, \infty)$ l'esponente coniugato di p. Allora la successione $(|x_k + y_k|^{p-1})_{k \in \mathbb{N}}$ è in ℓ^q, dato che

$$\|(|x_k + y_k|^{p-1})_{k \in \mathbb{N}}\|_q^q = \sum_{k=0}^{\infty} \left(|x_k + y_k|^{p-1}\right)^q = \sum_{k=0}^{\infty} |x_k + y_k|^p = \|\underline{x} + \underline{y}\|_p^p < \infty. \tag{5.14}$$

Applicando la disuguaglianza di Hölder (☞ proposizione 5.14) con esponenti q e p alle successioni $(|x_k + y_k|^{p-1})_{k \in \mathbb{N}}$ e $\underline{x}$, si ottiene

$$\begin{aligned} \sum_{k=0}^{\infty} |x_k + y_k|^{p-1} |x_k| &\leq \|(|x_k + y_k|^{p-1})_{k \in \mathbb{N}}\|_q \|\underline{x}\|_p \\ &= \|\underline{x} + \underline{y}\|_p^{p/q} \|\underline{x}\|_p = \|\underline{x} + \underline{y}\|_p^{p-1} \|\underline{x}\|_p, \end{aligned}$$

ove si è usata l'identità (5.14) e il fatto che $p/q = p-1$. In maniera simile, scambiando il ruolo di $\underline{x}$ e $\underline{y}$, si ottiene che

$$\sum_{k=0}^{\infty} |x_k + y_k|^{p-1} |y_k| \le \|\underline{x} + \underline{y}\|_p^{p-1} \|\underline{y}\|_p.$$

Ma allora, da (5.13) deduciamo che

$$\|\underline{x} + \underline{y}\|_p^p \le \|\underline{x} + \underline{y}\|_p^{p-1} \|\underline{y}\|_p + \|\underline{x} + \underline{y}\|_p^{p-1} \|\underline{y}\|_p.$$

Siccome $\|\underline{x} + \underline{y}\|_p \in (0, \infty)$, dividendo per $\|\underline{x} + \underline{y}\|_p^{p-1}$ entrambi i membri deduciamo (5.12). □

Corollario 5.16. *Per ogni $p \in [1, \infty)$, $\| \cdot \|_p$ è una norma su ℓ^p.*

Dimostrazione. Verifichiamo le proprietà nella definizione di norma (☞ definizione 4.1). Per definizione dello spazio ℓ^p (☞ definizione 5.7), la funzione $\| \cdot \|_p$ assume valori in $[0, \infty)$ su ℓ^p. Inoltre, la proposizione 5.10 dà la proprietà di separazione, mentre la 1-omogeneità è dimostrata in (5.6); infine, la disuguaglianza triangolare segue dalla disuguaglianza di Minkowski (☞ proposizione 5.15). □

Insieme alla proposizione 5.4, quest'ultimo risultato completa la dimostrazione che $(\ell^p, \| \cdot \|_p)$ è uno spazio normato per ogni $p \in [1, \infty]$. Discutiamo ora alcune inclusioni fra gli spazi di successioni introdotti finora.

Proposizione 5.17. *Valgono le seguenti proprietà.*

(i) *Per ogni $p \in [1, \infty]$, $\underline{x} \in \mathbb{F}^{\mathbb{N}}$ e $k \in \mathbb{N}$, si ha*

$$|x_k| \le \|\underline{x}\|_p. \tag{5.15}$$

(ii) *Per ogni $p, q \in [1, \infty]$, se $p < q$ allora*

$$\|\underline{x}\|_q \le \|\underline{x}\|_p \tag{5.16}$$

per ogni $\underline{x} \in \mathbb{F}^{\mathbb{N}}$. In particolare $\ell^p \subseteq \ell^q$, e inoltre l'inclusione è propria.

(iii) *Per ogni $p \in [1, \infty)$, si hanno le inclusioni $c_{00} \subseteq \ell^p \subseteq c_0$. Tali inclusioni sono proprie. Inoltre c_{00} è denso in $(\ell^p, \| \cdot \|_p)$ e ℓ^p è denso in $(c_0, \| \cdot \|_\infty)$.*

Dimostrazione. (i). La disuguaglianza (5.15) è ovvia nel caso $p = \infty$, dato che $\|\underline{x}\|_\infty = \sup_{k \in \mathbb{N}} |x_k|$. Nel caso $p \in [1, \infty)$, si ha

$$\|\underline{x}\|_p^p = \sum_{n=0}^{\infty} |x_n|^p \ge |x_k|^p$$

per ogni $k \in \mathbb{N}$ (una somma a termini nonnegativi è maggiore o uguale di ciascun suo addendo), e la disuguaglianza (5.15) segue prendendo la potenza $1/p$ di ambo i membri.

(ii). Nel caso $q = \infty$, la disuguaglianza (5.16) segue da (i), prendendo l'estremo superiore su k nel primo membro della disuguaglianza (5.15).

Supponiamo invece $q < \infty$. Allora

$$\|\underline{x}\|_q^q = \sum_{k=0}^{\infty} |x_k|^q = \sum_{k=0}^{\infty} |x_k|^{q-p}|x_k|^p \leq \sum_{k=0}^{\infty} \|\underline{x}\|_p^{q-p}|x_k|^p = \|\underline{x}\|_p^{q-p}\|\underline{x}\|_p^p = \|\underline{x}\|_p^q,$$

ove si è usata (5.15) nella disuguaglianza intermedia, insieme al fatto che $q-p > 0$; prendendo la potenza $1/q$ di entrambi i membri, si deduce (5.16).

Dalla disuguaglianza (5.16) segue immediatamente l'inclusione $\ell^p \subseteq \ell^q$ (se il secondo membro di (5.16) è finito, dev'esserlo anche il primo). L'inclusione è propria perché $((1+k)^{-1/p})_{k\in\mathbb{N}} \in \ell^q \setminus \ell^p$.

(iii). Chiaramente le successioni definitivamente nulle sono p-sommabili per ogni $p \in [1,\infty)$, quindi $c_{00} \subseteq \ell^p$, e l'inclusione è propria (ad esempio, $(2^{-n})_{n\in\mathbb{N}}$ è p-sommabile e mai nulla).

Sia ora $\underline{x} \in \ell^p$. Allora $\sum_{k=0}^{\infty} |x_k|^p < \infty$, pertanto $|x_k|^p \to 0$ per $k \to \infty$, ma allora anche $x_k \to 0$ per $k \to \infty$, cioè $\underline{x} \in c_0$ (se una serie converge, il termine generale è infinitesimo). Per l'arbitrarietà di $\underline{x} \in \ell^p$, questo dimostra l'inclusione $\ell^p \subseteq c_0$.

Ora, preso $q \in (p,\infty)$, le inclusioni precedentemente dimostrate ci danno la catena di inclusioni $\ell^p \subseteq \ell^q \subseteq c_0$; siccome da (ii) sappiamo che la prima inclusione nella catena è propria, ne segue che l'inclusione $\ell^p \subseteq c_0$ è propria a sua volta.

Rimangono da dimostrare le proprietà di densità. Abbiamo appena verificato la catena di inclusioni

$$c_0 \supseteq \ell^p \supseteq c_{00}.$$

Prendendo la chiusura in $(c_0, \|\cdot\|_\infty)$ di questi spazi si ottiene

$$c_0 = \overline{c_0}^{c_0} \supseteq \overline{\ell^p}^{c_0} \supseteq \overline{c_{00}}^{c_0} = c_0,$$

ove l'ultima uguaglianza segue dalla proposizione 5.6(iii). Dunque in quest'ultima catena tutte le inclusioni sono uguaglianze e in particolare $\overline{\ell^p}^{c_0} = c_0$, cioè ℓ^p è denso in c_0.

Per dimostrare che c_{00} è denso in $(\ell^p, \|\cdot\|_p)$, procediamo come nella dimostrazione della proposizione 5.6(iii). In altre parole, dato $\underline{x} \in \ell^p$, per ogni $n \in \mathbb{N}$ consideriamo la troncata n-sima $\underline{x^{(n)}} \in c_{00}$ di $\underline{x}$, definita come in (5.4). Allora

$$\|\underline{x} - \underline{x^{(n)}}\|_p^p = \sum_{k=0}^{n} |x_k - x_k|^p + \sum_{k=n+1}^{\infty} |x_k - 0|^p = \sum_{k=n+1}^{\infty} |x_k|^p \to 0 \tag{5.17}$$

per $n \to \infty$; nell'ultimo passaggio usiamo il fatto che la serie $\sum_{k=0}^{\infty} |x_k|^p$ è convergente (dato che $\underline{x} \in \ell^p$) e che le code di una serie convergente sono infinitesime (come in (4.9)). Da (5.17) abbiamo che $\|\underline{x} - \underline{x^{(n)}}\|_p \to 0$ per $n \to \infty$, cioè $\underline{x^{(n)}} \to \underline{x}$ in ℓ^p per $n \to \infty$. Siccome $\underline{x^{(n)}} \in c_{00}$ per ogni $n \in \mathbb{N}$, per l'arbitrarietà di $\underline{x} \in \ell^p$ questo dimostra (☞ proposizione 2.23(ix)) la densità di c_{00} in ℓ^p. □

Anche per gli spazi ℓ^p con $p < \infty$ vale un risultato analogo alla proposizione 5.6(i), ovverosia la fondamentale proprietà di completezza.

Proposizione 5.18. *Sia $p \in [1,\infty)$. Allora $(\ell^p, \|\cdot\|_p)$ è uno spazio di Banach.*

Dimostrazione. Sia $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione di Cauchy in ℓ^p.

Passo 1. Per ogni $k \in \mathbb{N}$, $(x_k^{(n)})_{n\in\mathbb{N}}$ è una successione di Cauchy in $\mathbb{F}$. Questo segue dalla disuguaglianza

$$|x_k^{(n)} - x_k^{(m)}| \leq \|\underline{x}^{(n)} - \underline{x}^{(m)}\|_p,$$

per ogni $k, m, n \in \mathbb{N}$, che a sua volta è conseguenza di (5.15) applicata alla successione $\underline{x}^{(n)} - \underline{x}^{(m)}$.

Passo 2. Siccome $\mathbb{F}$ è completo (☞ esempio 2.12(2)), per ogni $k \in \mathbb{N}$ la successione di Cauchy $(x_k^{(n)})_{n\in\mathbb{N}}$ converge in $\mathbb{F}$ e possiamo definire $\underline{x} \in \mathbb{F}^{\mathbb{N}}$ ponendo

$$x_k = \lim_{n\to\infty} x_k^{(n)} \tag{5.18}$$

per ogni $k \in \mathbb{N}$.

Passo 3. $\underline{x} \in \ell^p$. Infatti, da (5.18) segue che $|x_k|^p = \lim_{n\to\infty} |x_k^{(n)}|^p$ per ogni $k \in \mathbb{N}$. Pertanto dal lemma di Fatou (☞ teorema 3.5(viii)) deduciamo che

$$\|\underline{x}\|_p^p = \sum_{k=0}^{\infty} |x_k|^p \leq \liminf_{n\to\infty} \sum_{k=0}^{\infty} |x_k^{(n)}|^p \leq \sup_{n\in\mathbb{N}} \|\underline{x}^{(n)}\|_p^p < \infty,$$

dove l'ultima disuguaglianza è dovuta al fatto che la successione $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ è limitata in ℓ^p (☞ proposizione 2.23(v)).

Passo 4. $\underline{x}^{(n)} \to \underline{x}$ in ℓ^p per $n \to \infty$. Infatti, da (5.18) deduciamo che $\lim_{m\to\infty} |x_k^{(m)} - x_k^{(n)}|^p = |x_k - x_k^{(n)}|^p$ per ogni $n, k \in \mathbb{N}$. Di conseguenza, per il lemma di Fatou (☞ teorema 3.5(viii)), per ogni $n \in \mathbb{N}$,

$$\begin{aligned} \|\underline{x} - \underline{x}^{(n)}\|_p^p &= \sum_{k=0}^{\infty} |x_k - x_k^{(n)}|^p \\ &\leq \liminf_{m\to\infty} \sum_{k=0}^{\infty} |x_k^{(m)} - x_k^{(n)}|^p = \liminf_{m\to\infty} \|\underline{x}^{(m)} - \underline{x}^{(n)}\|_p^p. \end{aligned} \tag{5.19}$$

Sia $\epsilon > 0$. Siccome $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ è di Cauchy in ℓ^p, esiste $N \in \mathbb{N}$ tale che $\|\underline{x}^{(m)} - \underline{x}^{(n)}\|_p < \epsilon/2$ per ogni $n, m > N$. Di conseguenza, da (5.19) deduciamo che, per ogni $n > N$,

$$\|\underline{x} - \underline{x}^{(n)}\|_p \leq \liminf_{m\to\infty} \|\underline{x}^{(m)} - \underline{x}^{(n)}\|_p \leq \epsilon/2 < \epsilon.$$

Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $\|\underline{x} - \underline{x}^{(n)}\|_p \to 0$ per $n \to \infty$, cioè $\underline{x}^{(n)} \to \underline{x}$ in ℓ^p. □

Osservazione 5.19. A partire dai risultati di densità delle proposizioni 5.6(iii) e 5.17(iii), è possibile derivare informazioni sulla separabilità degli spazi di successioni qui introdotti. Presenteremo questi risultati in seguito (☞ proposizione 8.42), nel contesto delle conseguenze del teorema di Hahn–Banach.

5.2 Spazi L^p

Introduciamo brevemente (omettendo molti dettagli e dimostrazioni) la teoria degli spazi L^p, che generalizza quella degli spazi ℓ^p introdotti sopra. La discussione è fondamentalmente basata sulle nozioni di teoria della misura richiamate nel capitolo 3.

Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura (☞ definizione 3.1). In altre parole, $\mathcal{M}$ è una σ-algebra sull'insieme M e μ è una misura su $(M, \mathcal{M})$. Per evitare casi banali, assumiamo che $\mu(M) > 0$.

Definizione 5.20. Una funzione $f : M \to \mathbb{F}$ si dice *misurabile* se $f^{-1}(A) \in \mathcal{M}$ per ogni aperto $A \subseteq \mathbb{F}$. Denotiamo con $\mathscr{L}^0(M, \mathcal{M}, \mu)$, o più brevemente con $\mathscr{L}^0$, l'insieme di tutte le funzioni $f : M \to \mathbb{F}$ misurabili.

La teoria delle funzioni misurabili (☞ proposizione 3.4) dà il seguente risultato.

Proposizione 5.21. *$\mathscr{L}^0$ è un sottospazio vettoriale di $\mathcal{F}(M, \mathbb{F})$. In altre parole, $\mathscr{L}^0$ è uno spazio vettoriale con le operazioni puntuali (☞ esempio 1.2(7)).*

Lo spazio $\mathscr{L}^0$ gioca il ruolo di "spazio ambiente" che $\mathbb{F}^{\mathbb{N}}$ gioca nella definizione degli spazi di successioni. Definiamo ora alcuni sottospazi di $\mathscr{L}^0$.

Definizione 5.22. Per ogni $f \in \mathscr{L}^0$, poniamo

$$\|f\|_\infty = \operatorname{ess\,sup} |f| := \inf\{\lambda \in \mathbb{R} : \mu\{x \in M : |f(x)| > \lambda\} = 0\} \tag{5.20}$$

e, per ogni $p \in [1, \infty)$,

$$\|f\|_p = \left(\int_M |f|^p \, d\mu \right)^{1/p}.$$

Definiamo inoltre, per ogni $p \in [1, \infty]$,

$$\mathscr{L}^p = \mathscr{L}^p(M, \mathcal{M}, \mu) = \{f \in \mathcal{L}^0 : \|f\|_p < \infty\}.$$

Gli elementi di $\mathscr{L}^\infty$ sono detti funzioni *essenzialmente limitate* su M; per $p \in [1, \infty)$, gli elementi di $\mathscr{L}^p$ sono detti funzioni *p-sommabili*.

Osservazione 5.23. Consideriamo il caso in cui $(M, \mathcal{M}, \mu)$ sia lo spazio di misura $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp)$, ove $\mathcal{P}(\mathbb{N})$ è l'insieme delle parti di $\mathbb{N}$ e $\sharp$ è la misura del conteggio su $\mathbb{N}$ (☞ teorema 3.11). In questo caso, si ha $\mathscr{L}^0 = \mathcal{F}(\mathbb{N}, \mathbb{F}) = \mathbb{F}^{\mathbb{N}}$, e inoltre l'integrale di $\underline{x} \in \mathbb{F}^{\mathbb{N}}$ rispetto alla misura $\sharp$ non è altro che la somma $\sum_{k=0}^\infty x_k$ dei valori di $\underline{x}$. Da questo si vede che $\mathscr{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp) = \ell^p$ per ogni $p \in [1, \infty]$.

Come nel caso degli spazi ℓ^p, è possibile dimostrare che gli $\mathscr{L}^p$ sono sottospazi vettoriali di $\mathscr{L}^0$. Inoltre dalla definizione segue subito che $\|\cdot\|_p$ è 1-omogenea, cioè

$$\|\alpha f\|_p = |\alpha| \|f\|_p$$

per ogni $\alpha \in \mathbb{F}$ e $f \in \mathscr{L}^0$. Infine, $\|\cdot\|_p$ soddisfa la disuguaglianza triangolare; più precisamente, come nel caso degli spazi ℓ^p, è possibile dimostrare le seguenti disuguaglianze.

Teorema 5.24 (disuguaglianza di Hölder)**.** *Siano $p, q \in [1, \infty]$ esponenti coniugati. Allora, per ogni $f \in \mathscr{L}^p$ e $g \in \mathscr{L}^q$,*

$$\int_M |fg| \, d\mu \le \|f\|_p \|g\|_q.$$

In particolare, $fg \in \mathscr{L}^1$ e

$$\left| \int_M fg \, d\mu \right| \le \|f\|_p \|g\|_q.$$

Teorema 5.25 (disuguaglianza di Minkowski). *Sia $p \in [1, \infty]$. Allora, per ogni $f, g \in \mathscr{L}^0$,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Per poter concludere che $\|\cdot\|_p$ è una norma su $\mathscr{L}^p$, mancherebbe solo da verificare la proprietà di separazione (☞ definizione 4.1). Tuttavia questa proprietà in generale fallisce: infatti si ha solamente che, per ogni $f \in \mathscr{L}^0$ e $p \in [1, \infty]$,

$$\|f\|_p = 0 \qquad \Longleftrightarrow \qquad f = 0 \ \mu\text{-quasi ovunque} \tag{5.21}$$

(☞ teorema 3.5(v)). In particolare, se esistono sottoinsiemi μ-trascurabili non vuoti di M, allora esistono funzioni f misurabili che non sono nulle, ma sono solo quasi ovunque nulle, e per queste funzioni si ha $\|f\|_p = 0$ senza che f sia identicamente nulla. Dunque, in generale, $\|\cdot\|_p$ non è una norma su $\mathscr{L}^p$, ma solo una *seminorma*.

Per ovviare a questo problema (dato che vogliamo poter utilizzare la teoria degli spazi normati anche in questo caso), la soluzione comunemente adottata è quella di "non distinguere" fra funzioni che sono uguali μ-quasi ovunque (cioè che differiscono solo su insiemi di misura nulla). In altre parole, invece di lavorare con funzioni $f : M \to \mathbb{F}$, lavoreremo con *classi di equivalenza* di funzioni.

Definizione 5.26. Diciamo che due funzioni $f, g \in \mathscr{L}^0$ sono equivalenti (rispetto alla misura μ), e scriviamo $f \sim g$, se $f = g$ μ-quasi ovunque (cioè se $\mu\{x \in M : f(x) \neq g(x)\} = 0$).

È facile verificare che $\sim$ è una relazione di equivalenza (riflessiva, simmetrica e transitiva) su $\mathscr{L}^0$. Per $f \in \mathscr{L}^0$, denotiamo con $[f]$ la classe di equivalenza di f, cioè l'insieme

$$[f] = \{g \in \mathscr{L}^0 : g \sim f\}$$

delle funzioni equivalenti ad f. L'insieme di tutte le classi di equivalenza rispetto a $\sim$ è detto *insieme quoziente* rispetto a $\sim$.

Definizione 5.27. Per ogni $p \in [1, \infty]$, poniamo

$$L^p(M, \mathcal{M}, \mu) = \mathscr{L}^p(M, \mathcal{M}, \mu)/{\sim} \coloneqq \{[f] : f \in \mathscr{L}^p(M, \mathcal{M}, \mu)\}.$$

Si può verificare che la relazione di equivalenza $\sim$ è compatibile con le operazioni puntuali su $\mathscr{L}^0$ e con le seminorme $\|\cdot\|_p$.

Lemma 5.28. *Siano $f, g, \tilde{f}, \tilde{g} \in \mathscr{L}^0$ e $p \in [1, \infty]$. Se $f \sim \tilde{f}$ e $g \sim \tilde{g}$, allora:*

(i) $f + g \sim \tilde{f} + \tilde{g}$;

(ii) $f \cdot g \sim \tilde{f} \cdot \tilde{g}$;

(iii) $\|f\|_p = \|\tilde{f}\|_p$.

Di conseguenza:

- le operazioni puntuali su $\mathscr{L}^p$ passano al quoziente e definiscono una struttura di spazio vettoriale su $L^p(M, \mathcal{M}, \mu)$:

$$[f] + [g] \coloneqq [f + g], \qquad \alpha[f] \coloneqq [\alpha f]$$

per ogni $[f], [g] \in L^p(M, \mathcal{M}, \mu)$ e $\alpha \in \mathbb{F}$;

- la seminorma $\|\cdot\|_p$ su $\mathscr{L}^p$ passa al quoziente e definisce una norma su $L^p(M, \mathcal{M}, \mu)$:
$$\|[f]\|_p \coloneqq \|f\|_p$$
per ogni $[f] \in L^p(M, \mathcal{M}, \mu)$.

Infatti, avendo identificato in $L^p(M, \mathcal{M}, \mu)$ funzioni fra loro equivalenti, la proprietà (5.21) implica la proprietà di separazione per $\|\cdot\|_p$ su $L^p(M, \mathcal{M}, \mu)$.

In altre parole, $(L^p(M, \mathcal{M}, \mu), \|\cdot\|_p)$ è uno spazio normato. Vale in effetti una proprietà più forte.

Teorema 5.29. *Per ogni $p \in [1, \infty]$, $(L^p(M, \mathcal{M}, \mu), \|\cdot\|_p)$ è uno spazio di Banach.*

Osservazione 5.30. Il passaggio dall'insieme di funzioni $\mathscr{L}^p(M, \mathcal{M}, \mu)$ all'insieme $L^p(M, \mathcal{M}, \mu)$ di classi di equivalenza di funzioni appena discusso (☞ definizione 5.27) è fondamentale da un punto di vista tecnico per poter applicare la teoria degli spazi normati in questo contesto. D'altra parte, per evitare di appesantire la notazione, nella pratica si tende a non distinguere fra la funzione f e la sua classe di equivalenza $[f]$, a meno che non sia strettamente necessario. Ad esempio, con un abuso di linguaggio, si può dire "una funzione f in $L^p(M, \mathcal{M}, \mu)$" invece di "una funzione f la cui classe di equivalenza $[f]$ è in $L^p(M, \mathcal{M}, \mu)$". Queste imprecisioni normalmente non creano problemi e di fatto sono sistematicamente usate nella letteratura matematica.

Osservazione 5.31. La dimostrazione del teorema 5.29 richiede argomenti un po' più sofisticati di quelli usati per gli spazi ℓ^p. In particolare, nel caso $p < \infty$, nel contesto degli spazi $L^p(M, \mathcal{M}, \mu)$ in generale non è vera la disuguaglianza
$$|f(t) - g(t)| \le \|f - g\|_p$$
per ogni $t \in M$, o anche solo per μ-quasi ogni $t \in M$; questo impedisce di procedere come nel passo 1 delle dimostrazioni del teorema 2.14 e della proposizione 5.18, nel senso che non è possibile dedurre la convergenza puntuale dalla convergenza in norma, o la proprietà di Cauchy puntuale dalla proprietà di Cauchy in norma. L'argomento più sofisticato usato per la dimostrazione del teorema 5.29 contiene come "sottoprodotti" i seguenti risultati.

Proposizione 5.32. *Sia $p \in [1, \infty)$. Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura. Siano $(f_n)_n$ una successione a valori in $L^p(M, \mathcal{M}, \mu)$ e $f \in L^p(M, \mathcal{M}, \mu)$.*

(i) Se $f_n \to f$ in $L^p(M, \mathcal{M}, \mu)$ per $n \to \infty$, allora $(f_n)_n$ ha una sottosuccessione che converge puntualmente μ-quasi ovunque a f.

(ii) Se $f_n \to f$ puntualmente μ-quasi ovunque, ed esiste $g \in L^p(M, \mathcal{M}, \mu)$ tale che $|f_n| \le g$ μ-quasi ovunque per ogni $n \in \mathbb{N}$, allora $f_n \to f$ in $L^p(M, \mathcal{M}, \mu)$ per $n \to \infty$.
(teorema di convergenza dominata in L^p)

Notazione. Invece della notazione $L^p(M, \mathcal{M}, \mu)$, quando non ci sia pericolo di confusione possiamo anche utilizzare le notazioni più brevi $L^p(M, \mu)$ oppure $L^p(M)$.

In particolare, quando I è un intervallo di $\mathbb{R}$, considereremo normalmente I come spazio di misura con la misura di Lebesgue (☞ teorema 3.10) ristretta ad I, e la notazione $L^p(I)$ indicherà gli spazi L^p corrispondenti.

Quando $I = [a, b]$ è un intervallo limitato, scriviamo anche $L^p(a, b)$ anziché $L^p(I)$. Si noti che lavorare con l'intervallo chiuso $[a, b]$ o l'intervallo aperto (a, b) nella pratica non fa differenza, dato che gli elementi di L^p sono definiti "a meno di insiemi di misura nulla" e quindi possiamo trascurare gli estremi a e b dell'intervallo (l'insieme $\{a, b\}$ degli estremi è Lebesgue-trascurabile). In effetti, per $p < \infty$, l'espressione della norma $\|\cdot\|_p$ di un elemento $f \in L^p(a, b)$ è

$$\|f\|_p^p = \int_a^b |f(t)|^p \, dt. \tag{5.22}$$

Notiamo che lo spazio $C[a, b]$ delle funzioni continue $f : [a, b] \to \mathbb{F}$ (☞ esempio 1.2(2)) si può pensare come un sottospazio di $L^\infty(a, b)$; infatti, le funzioni continue su $[a, b]$ sono funzioni limitate (☞ teorema 2.36), quindi a maggior ragione sono essenzialmente limitate (☞ definizione 5.22). Ad essere precisi, gli elementi di $C[a, b]$ e di $L^\infty(a, b)$ hanno natura diversa (i primi sono funzioni, i secondi sono classi di equivalenza di funzioni). Tuttavia, se due funzioni continue $f, g \in C[a, b]$ coincidono quasi ovunque, allora in realtà $f = g$ ovunque (l'insieme dove due funzioni continue differiscono è un aperto, e un aperto che ha misura di Lebesgue nulla è vuoto). In altre parole, per funzioni continue $f, g \in C[a, b]$ si ha $f = g$ se e solo se $[f] = [g]$, e quindi non c'è pericolo a identificare $f \in C[a, b]$ con la sua classe di equivalenza $[f]$. In questo senso, possiamo scrivere $C[a, b] \subseteq L^\infty(a, b)$.

Per un motivo analogo, per una funzione continua $f \in C[a, b]$ si ha

$$\sup_{t \in [a,b]} |f(t)| = \operatorname*{ess\,sup}_{t \in [a,b]} |f(t)|;$$

in altre parole, la norma $\|\cdot\|_\infty$ su $L^\infty(a, b)$ (definita come in (5.20)), quando ristretta a $C[a, b]$ coincide con la norma $\|\cdot\|_\infty$ su $C[a, b]$ (definita come in (4.1)); dunque l'uso della stessa notazione $\|\cdot\|_\infty$ in questo caso non è ambiguo. Siccome $(C[a, b], \|\cdot\|_\infty)$ è uno spazio di Banach (☞ esempio 4.5(2)), dalla proposizione 4.16 deduciamo il seguente risultato.

Proposizione 5.33. *Sia $[a, b]$ un intervallo chiuso e limitato di $\mathbb{R}$. Allora $(C[a, b], \|\cdot\|_\infty)$ è un sottospazio chiuso proprio di $(L^\infty(a, b), \|\cdot\|_\infty)$.*

Osserviamo ora che fra gli spazi $L^p(a, b)$ si hanno inclusioni "opposte" rispetto a quelle discusse per gli spazi ℓ^p nella proposizione 5.17(ii). Questo vale più in generale per gli spazi $L^p(M, \mu)$ quando $\mu(M) < \infty$, come conseguenza della disuguaglianza di Hölder.

Proposizione 5.34. *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura* finito*, cioè tale che $\mu(M) < \infty$. Siano $p, q \in [1, \infty]$. Se $p \le q$, allora $L^q(M, \mu) \subseteq L^p(M, \mu)$, e inoltre*

$$\|f\|_p \le \mu(M)^{1/p - 1/q} \|f\|_q$$

per ogni $f \in L^q(M, \mu)$.

Dimostrazione. Si applichi la disuguaglianza di Hölder (☞ teorema 5.24) alla coppia di funzioni $|f|^p$ e $\mathbf{1}_M$,usando come esponenti q/p e il suo coniugato. □

Osservazione 5.35. Si noti che l'ipotesi di misura finita nella proposizione 5.34 è essenziale per la validità delle inclusioni sopra enunciate. In effetti, nel caso $M = \mathbb{R}$ con la misura di Lebesgue, non c'è alcuna inclusione fra gli spazi $L^p(\mathbb{R})$ per p distinti. Inoltre, nel caso $M = [a, b]$, l'inclusione $L^q(a, b) \subseteq L^p(a, b)$ per $1 \leq p < q \leq \infty$ è propria. Lasciamo le facili verifiche per esercizio.

Dalle proposizioni 5.33 e 5.34 deduciamo la catena di inclusioni

$$C[a, b] \subseteq L^\infty(a, b) \subseteq L^p(a, b)$$

per ogni $p \in [1, \infty)$. Tuttavia, mentre $C[a, b]$ è chiuso in $(L^\infty(a, b), \|\cdot\|_\infty)$, lo stesso non vale per $(L^p(a, b), \|\cdot\|_p)$ con $p < \infty$.

La non-chiusura di $C[a, b]$ in $L^p(a, b)$ per $p < \infty$ è una conseguenza del seguente risultato, in cui si mostra che la funzione caratteristica $\mathbf{1}_{[c,d]}$ di un intervallo $[c, d] \subseteq (a, b)$, che chiaramente non è continua, si può approssimare in norma $\|\cdot\|_p$ con funzioni continue su $[a, b]$. Per la precisione, possiamo prendere le funzioni approssimanti nell'insieme

$$C_c(a, b) = \{f \in C[a, b] : \operatorname{supp} f \subseteq (a, b)\} \tag{5.23}$$

delle funzioni continue *a supporto compatto*[1] contenuto in (a, b), dove il supporto $\operatorname{supp} f$ di $f \in C[a, b]$ è definito da

$$\operatorname{supp} f = \overline{\{t \in [a, b] : f(t) \neq 0\}}. \tag{5.24}$$

È facile verificare che $C_c(a, b)$ è un sottospazio vettoriale di $C[a, b]$.

Lemma 5.36. *Sia $p \in [1, \infty)$. Siano $a, b, c, d \in \mathbb{R}$ tali che $a < c < d < b$. Allora esiste una successione di funzioni $f_n \in C_c(a, b)$ tali che $f_n \to \mathbf{1}_{[c,d]}$ in $(L^p(a, b), \|\cdot\|_p)$ per $n \to \infty$.*

Dimostrazione. Vogliamo costruire una successione di funzioni continue su $[a, b]$ che approssima la funzione caratteristica $\mathbf{1}_{[c,d]}$ dell'intervallo $[c, d]$. Per ogni $n \in \mathbb{N}_+$, definiamo $f_n \in C[a, b]$ ponendo

$$f_n(t) = \begin{cases} 1 & \text{se } c \leq t \leq d, \\ 1 - n(t - d) & \text{se } d \leq t \leq d + 1/n, \\ 1 - n(c - t) & \text{se } c - 1/n \leq t \leq c, \\ 0 & \text{altrimenti} \end{cases}$$

per ogni $t \in [a, b]$. In altre parole, la funzione f_n vale 1 su $[c, d]$, vale 0 fuori dall'intervallo $(c-1/n, d+1/n)$, e nei rimanenti intervalli $[c-1/n, c]$ e $[d, d+1/n]$ si "raccorda" in maniera lineare affine in modo da risultare continua. Si noti che, per n abbastanza grande, $\operatorname{supp} f_n = [c - 1/n, d + 1/n] \subseteq (a, b)$ e quindi $f_n \in C_c(a, b)$.

[1] In generale, se Ω è uno spazio topologico e $f \in C(\Omega)$, il *supporto* $\operatorname{supp} f$ di f è la chiusura in Ω dell'insieme $\{x \in \Omega : f(x) \neq 0\}$. Si denota con $C_c(\Omega)$ il sottoinsieme di $C(\Omega)$ delle funzioni *a supporto compatto*, cioè delle $f \in C(\Omega)$ tali che $\operatorname{supp} f$ è compatto; è facile verificare che $C_c(\Omega)$ è un sottospazio vettoriale di $C(\Omega)$. Nel caso in cui $\Omega = (a, b)$ è un intervallo aperto e limitato con la topologia euclidea, si vede facilmente che ogni funzione $f \in C(a, b)$ a supporto compatto si estende in maniera unica a una funzione $f \in C[a, b]$ che vale zero agli estremi dell'intervallo; in questo modo, $C_c(a, b)$ si può pensare come un sottospazio di $C[a, b]$, e le definizioni (5.23) e (5.24) risultano corenti con la teoria generale qui descritta.

Inoltre non è difficile verificare che

$$f_n \to \mathbf{1}_{[c,d]} \qquad \text{puntualmente su } [a,b] \text{ per } n \to \infty.$$

Inoltre $|f_n(t)| \leq 1$ per ogni $t \in [a,b]$. Il teorema di convergenza dominata (☞ proposizione 5.32(ii)) allora implica che, per ogni $p \in [1,\infty)$,

$$\|f_n - \mathbf{1}_{[c,d]}\|_p^p = \int_a^b |f_n(t) - \mathbf{1}_{[c,d]}(t)|^p \, dt \to 0 \qquad \text{per } n \to \infty,$$

cioè $f_n \to \mathbf{1}_{[c,d]}$ in $L^p(a,b)$ per $n \to \infty$. □

Dal lemma 5.36 deduciamo che $\mathbf{1}_{[c,d]}$ appartiene alla chiusura di $C[a,b]$ in $(L^p(a,b), \|\cdot\|_p)$ (☞ proposizione 2.23(vii)) per ogni $p \in [1,\infty)$; siccome $\mathbf{1}_{[c,d]}$ non è continua su $[a,b]$ (ha discontinuità di salto in c e in d), ne concludiamo che $C[a,b]$ non è chiuso in $(L^p(a,b), \|\cdot\|_p)$ per $p \in [1,\infty)$.

Utilizzando tecniche più raffinate di teoria della misura, il risultato di approssimazione descritto nel lemma 5.36 si può generalizzare, arrivando a concludere che ogni funzione in $L^p(a,b)$ è approssimabile in norma $\|\cdot\|_p$ con funzioni continue.

Teorema 5.37. *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura. Sia $p \in [1,\infty)$.*

(i) L'insieme
$$\operatorname{span}\{\mathbf{1}_E : E \in \mathcal{M},\ \mu(E) < \infty\} \tag{5.25}$$
è denso in $L^p(M)$.

(ii) Supponiamo che $M = \Omega$ sia un sottoinsieme aperto di $\mathbb{R}^n$ con la misura di Lebesgue λ_n. L'insieme
$$\operatorname{span}\{\mathbf{1}_A : A \subseteq \Omega \text{ aperto},\ \lambda_n(A) < \infty\} \tag{5.26}$$
è denso in $L^p(\Omega)$.

(iii) Supponiamo che $M = I$ sia un intervallo di $\mathbb{R}$ con la misura di Lebesgue. L'insieme
$$\operatorname{span}\{\mathbf{1}_{[c,d]} : [c,d] \text{ intervallo chiuso e limitato},\ [c,d] \subseteq I\} \tag{5.27}$$
è denso in $L^p(I)$.

(iv) Supponiamo che $M = [a,b]$ sia un intervallo chiuso e limitato in $\mathbb{R}$ con la misura di Lebesgue. Lo spazio $C_c(a,b)$ è denso in $(L^p(a,b), \|\cdot\|_p)$ per ogni $p \in [1,\infty)$. Di conseguenza, anche $C[a,b]$ è denso in $(L^p(a,b), \|\cdot\|_p)$ per ogni $p \in [1,\infty)$.

Dimostrazione (cenno). (i). Grazie alla decomposizione (3.1), ci si riduce a dimostrare che ogni $f \in L^p(M)$ nonnegativa appartiene alla chiusura in $L^p(M)$ dell'insieme (5.25). A sua volta, questo fatto segue dal risultato di approssimazione della proposizione 3.4(vi) utilizzando convergenza dominata (☞ proposizione 5.32(ii)).

(ii). Grazie a (i) è sufficiente dimostrare che, per ogni E boreliano di misura finita, la funzione $\mathbf{1}_E$ appartiene alla chiusura in $L^p(\Omega)$ dell'insieme 5.26. D'altra parte, per il teorema 3.10(v), per ogni $\epsilon > 0$ troviamo un aperto $A \subseteq \Omega$ con $E \subseteq A$ e $\lambda_n(A \setminus E) < \epsilon^p$, dunque $\|\mathbf{1}_E - \mathbf{1}_A\|_p \leq \epsilon$.

(iii). A meno di rimuoverne gli estremi (che sono Lebesgue-trascurabili), non è restrittivo supporre che I sia un intervallo aperto. Grazie a (ii) è allora sufficiente dimostrare che, per ogni aperto $A \subseteq I$ di misura finita, la funzione $\mathbf{1}_A$ appartiene alla chiusura in $L^p(I)$ dell'insieme (5.27). In effetti, non è restrittivo supporre che A sia limitato: infatti, per ogni aperto $A \subseteq I$, se poniamo $A_n = A \cap (-n, n)$ per ogni $n \in \mathbb{N}_+$, allora gli A_n sono aperti limitati e $\|\mathbf{1}_A - \mathbf{1}_{A_n}\|_p \to 0$ per convergenza dominata (☞ proposizione 5.32(ii)). Per una ragione simile, non è restrittivo supporre anche che $\overline{A} \subseteq I$: se ad esempio $I = (a, b)$ è limitato, possiamo approssimare $\mathbf{1}_A$ con $\mathbf{1}_{A \cap (a+1/n, b-1/n)}$ per $n \to \infty$.

D'altra parte, ogni aperto $A \subseteq \mathbb{R}$ è unione al più numerabile di intervalli aperti disgiunti (le componenti connesse di A). Se A è limitato lo sono a maggior ragione tali intervalli, e se $\overline{A} \subseteq I$, anche le chiusure di tali intervalli sono contenute in I. Se $A = \bigcup_{j=0}^N (c_j, d_j)$ è unione finita di tale intervalli, allora $\mathbf{1}_A = \sum_{j=0}^N \mathbf{1}_{[c_j, d_j]}$ appartiene all'insieme (5.27), e abbiamo finito (si noti che $\mathbf{1}_{[c_j, d_j]} = \mathbf{1}_{(c_j, d_j)}$ in $L^p(I)$). Se invece $A = \bigcup_{j=0}^\infty (c_j, d_j)$ è unione numerabile di tali intervalli, allora ponendo $A_n = \bigcup_{j=0}^n (c_j, d_j)$, di nuovo per convergenza dominata abbiamo $\|\mathbf{1}_A - \mathbf{1}_{A_n}\|_p \to 0$.

(iv). Grazie a (iii), è sufficiente dimostrare che, per ogni intervallo chiuso e limitato $[c, d] \subseteq (a, b)$, la funzione $\mathbf{1}_{[c,d]}$ è nella chiusura di $C_c(a, b)$ in $L^p(a, b)$. Questo tuttavia è conseguenza del lemma 5.36. □

Osservazione 5.38. Confrontando il teorema 5.37(iv) e la proposizione 5.17(iii), si può notare un'analogia fra gli spazi $C_c(a, b)$ e c_{00}, nelle loro rispettive relazioni con gli spazi $L^p(a, b)$ e ℓ^p. In effetti, se $\mathbb{N}$ è dotato della metrica indotta da $\mathbb{R}$, i sottoinsiemi compatti di $\mathbb{N}$ sono tutti e soli i sottoinsiemi finiti di $\mathbb{N}$; in questo senso, gli elementi di c_{00} sono esattamente le "funzioni a supporto compatto" da $\mathbb{N}$ a $\mathbb{F}$.

Come conseguenza del teorema 5.37(iv) e del teorema di Stone–Weierstrass (☞ teorema 4.30), ricaviamo il seguente risultato.

Corollario 5.39. *Sia $[a, b]$ un intervallo limitato di $\mathbb{R}$. Per ogni $p \in [1, \infty)$, lo spazio $(L^p(a, b), \|\cdot\|_p)$ è separabile.*

Dimostrazione. Per il corollario 4.31, esiste un sottoinsieme E di $C[a, b]$ numerabile e denso rispetto alla norma $\|\cdot\|_\infty$ su $C[a, b]$. Affermiamo che E è anche denso in $L^p(a, b)$ rispetto alla norma $\|\cdot\|_p$.

Infatti, dalla disuguaglianza fra le norme (☞ proposizione 5.34)

$$\|f\|_p = \left(\int_a^b |f(t)|^p \, dt \right)^{1/p} \leq (b - a)^{1/p} \|f\|_\infty,$$

deduciamo che ogni successione convergente in norma $\|\cdot\|_\infty$ converge anche in norma $\|\cdot\|_p$. Dalla caratterizzazione della chiusura di un insieme in termini di limiti di successioni (☞ proposizione 2.23(vii)), deduciamo allora che

$$\overline{E}^{(L^p(a,b), \|\cdot\|_p)} \supseteq \overline{E}^{(C[a,b], \|\cdot\|_\infty)} = C[a, b],$$

dove l'ultima uguaglianza è data dalla densità di E in $C[a,b]$. Prendendo nuovamente le chiusure in $L^p(a,b)$, si deduce infine

$$\overline{E}^{(L^p(a,b),\|\cdot\|_p)} \supseteq \overline{C[a,b]}^{(L^p(a,b),\|\cdot\|_p)} = L^p(a,b),$$

dove l'ultima uguaglianza è conseguenza del teorema 5.37(iv). □

Evidentemente la dimostrazione del corollario 5.39 non si estende al caso $p = \infty$, dato che $C[a,b]$ non è denso in $L^\infty(a,b)$. In effetti, è possibile dimostrare (☞ osservazione 8.43) che $L^\infty(a,b)$ <u>non</u> è separabile.

6 Spazi di Hilbert

6.1 Prodotti scalari e spazi con prodotto scalare

Per introdurre la definizione di prodotto scalare, discutiamo separatamente i due casi in cui il campo $\mathbb{F}$ degli scalari è $\mathbb{R}$ oppure $\mathbb{C}$.

Definizione 6.1. Sia H uno spazio vettoriale su $\mathbb{R}$. Un *prodotto scalare* su H è una mappa $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ con le seguenti proprietà:

(a) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{R}$ (lineare nella prima variabile);

(b) $\langle x, y \rangle = \langle y, x \rangle$ per ogni $x, y \in H$ (simmetrica);

(c) $\langle x, x \rangle \geq 0$ per ogni $x \in H$, e $\langle x, x \rangle = 0$ se e solo se $x = 0$ (definita positiva).

Osservazione 6.2. Dalle proprietà (a) e (b) della definizione 6.1, deduciamo:

(a') $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{R}$ (lineare nella seconda variabile).

Insieme, le proprietà (a) e (a') si esprimono dicendo che $\langle \cdot, \cdot \rangle$ è *bilineare*. Con questa terminologia, si può equivalentemente riformulare la definizione di prodotto scalare (su uno spazio vettoriale reale H): un prodotto scalare è una *forma bilineare simmetrica definita positiva* su H. Qui la parola "forma" indica che il prodotto scalare è una mappa a valori scalari.

Vediamo ora come la definizione 6.1 va modificata quando si lavora con spazi vettoriali complessi anziché reali.

Definizione 6.3. Sia H uno spazio vettoriale su $\mathbb{C}$. Un *prodotto scalare* su H è una mappa $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ con le seguenti proprietà:

(a) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{C}$ (lineare nella prima variabile);

($\bar{\text{b}}$) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ per ogni $x, y \in H$ (hermitiana);

(c) $\langle x, x \rangle \geq 0$ per ogni $x \in H$, e $\langle x, x \rangle = 0$ se e solo se $x = 0$ (definita positiva).

Osservazione 6.4. Dalle proprietà (a) e ($\bar{\text{b}}$) della definizione 6.3, deduciamo:

($\bar{\text{a}}$') $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{C}$ (<u>anti</u>lineare nella seconda variabile).

Insieme, le proprietà (a) e ($\bar{\text{a}}$') si esprimono dicendo che $\langle \cdot, \cdot \rangle$ è *sesquilineare* (☞ definizione 8.19). Con questa terminologia, si può equivalentemente riformulare la definizione di prodotto scalare (su uno spazio vettoriale complesso H): un prodotto scalare è una *forma sesquilineare hermitiana definita positiva* su H.

Osservazione 6.5. Si noti che $\overline{\alpha} = \alpha$ per ogni numero reale $\alpha \in \mathbb{R}$. Pertanto, se si rimpiazza $\mathbb{C}$ con $\mathbb{R}$, le proprietà nella definizione 6.3 si riducono a quelle nella definizione 6.1. Possiamo dunque riscrivere la definizione di prodotto scalare in maniera unificata (valida sia per spazi vettoriali reali che complessi). In altre parole, se H è uno spazio vettoriale su $\mathbb{F}$ (dove $\mathbb{F}$ può essere indifferentemente $\mathbb{R}$ oppure $\mathbb{C}$), un prodotto scalare $\langle \cdot, \cdot \rangle$ su H soddisfa le seguenti proprietà:

(a) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{F}$;

(a') $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$ per ogni $x, y, z \in H$ e $\alpha, \beta \in \mathbb{F}$;

(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ per ogni $x, y \in H$;

(c) $\langle x, x \rangle \geq 0$ per ogni $x \in H$, e $\langle x, x \rangle = 0$ se e solo se $x = 0$.

Possiamo dunque continuare la discussione trattando in maniera unificata il caso in cui $\mathbb{F}$ sia $\mathbb{R}$ oppure $\mathbb{C}$.

Definizione 6.6. Uno spazio vettoriale H su $\mathbb{F}$ dotato di un prodotto scalare $\langle \cdot, \cdot \rangle$ è detto *spazio pre-hilbertiano* (o *spazio con prodotto scalare*).

Esempio 6.7. Ecco alcuni esempi di spazi pre-hilbertiani.

(1) Il *prodotto scalare euclideo* (o *prodotto scalare standard*) su $\mathbb{F}^n$ è dato da

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_j}$$

per ogni $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{F}^n$. È facile verificare che $\langle \cdot, \cdot \rangle$ è un prodotto scalare. Infatti, le proprietà di bilinearità e simmetria (risp. sesquilinearità e hermitianità) sono immediatamente verificate. Inoltre, per ogni $x \in \mathbb{F}^n$,

$$\langle x, x \rangle = \sum_{j=1}^{n} |x_j|^2 = \|x\|_2^2 \geq 0 \tag{6.1}$$

e la proprietà di separazione della norma euclidea $\| \cdot \|_2$ su $\mathbb{F}^n$ implica che $\langle x, x \rangle = 0 \iff \|x\|_2 = 0 \iff x = 0$; dunque $\langle \cdot, \cdot \rangle$ è definita positiva.

(2) Sullo spazio ℓ^2 un prodotto scalare è dato da

$$\langle \underline{x}, \underline{y} \rangle = \sum_{k=0}^{\infty} x_k \overline{y_k} \tag{6.2}$$

per ogni $\underline{x}, \underline{y} \in \ell^2$. Si noti che, applicando la disuguaglianza di Hölder (☞ proposizione 5.14) con $p = q = 2$, si deduce che la serie in (6.2) converge assolutamente per ogni $\underline{x}, \underline{y} \in \ell^2$, dunque $\langle \cdot, \cdot \rangle : \ell^2 \times \ell^2 \to \mathbb{F}$ è ben definito. La verifica di bilinearità e simmetria (risp. sesquilinearità e hermitianità) è immediata. Inoltre

$$\langle \underline{x}, \underline{x} \rangle = \sum_{k=0}^{\infty} |x_k|^2 = \|\underline{x}\|_2^2 \geq 0 \tag{6.3}$$

e la proprietà di separazione della norma $\| \cdot \|_2$ su ℓ^2 implica che $\langle \underline{x}, \underline{x} \rangle = 0 \iff \|\underline{x}\|_2 = 0 \iff \underline{x} = \underline{0}$; dunque $\langle \cdot, \cdot \rangle$ è definita positiva.

(3) Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura. Sullo spazio $L^2(M, \mathcal{M}, \mu)$ un prodotto scalare è dato da
$$\langle f, g \rangle = \int_M f \, \overline{g} \, d\mu$$
per ogni $f, g \in L^2(M, \mathcal{M}, \mu)$. La verifica che $\langle \cdot, \cdot \rangle$ è un prodotto scalare è analoga a quella per ℓ^2 in (2). In particolare, si ha
$$\langle f, f \rangle = \int_M |f|^2 \, d\mu = \|f\|_2^2 \tag{6.4}$$
per ogni $f \in L^2(M, \mathcal{M}, \mu)$.

(4) Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano e V un sottospazio vettoriale di H. Allora la restrizione di $\langle \cdot, \cdot \rangle$ a $V \times V$ è un prodotto scalare su V (detto *prodotto scalare indotto* da $(H, \langle \cdot, \cdot \rangle)$ su V), che per brevità denotiamo ancora con $\langle \cdot, \cdot \rangle$. Dunque $(V, \langle \cdot, \cdot \rangle)$ è uno spazio pre-hilbertiano.

In molti degli esempi visti sopra di prodotto scalare $\langle \cdot, \cdot \rangle$ su particolari spazi vettoriali H, abbiamo osservato che l'espressione $\langle x, x \rangle$ si può riscrivere come $\|x\|^2$ per un'opportuna norma $\| \cdot \|$ su H (si vedano le formule (6.1), (6.3) e (6.4)). Questo fatto suggerisce la seguente definizione.

Definizione 6.8. Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano. La *norma indotta dal prodotto scalare* è la mappa $\| \cdot \| : H \times H \to [0, \infty)$ definita da
$$\|x\| = \sqrt{\langle x, x \rangle} \tag{6.5}$$
per ogni $x \in H$.

Osservazione 6.9. Siccome $\langle \cdot, \cdot \rangle$ è definito positivo, si ha $\langle x, x \rangle \geq 0$ per ogni $x \in H$ e quindi la radice quadrata nel membro destro di (6.5) è un ben definito elemento di $[0, \infty)$.

Vogliamo ora verificare che la mappa $\|\cdot\|$ della definizione 6.8 è effettivamente una norma su H (nel senso della definizione 4.1). A questo scopo, dimostriamo preliminarmente una disuguaglianza.

Proposizione 6.10 (disuguaglianza di Cauchy–Schwarz)**.** *Siano $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano e $\| \cdot \|$ la norma indotta dal prodotto scalare. Allora*
$$|\langle x, y \rangle| \leq \|x\| \|y\| \tag{6.6}$$
per ogni $x, y \in H$.

Dimostrazione. La disuguaglianza è ovvia se $x = 0$ oppure $y = 0$, dato che ambo i membri si annullano. Possiamo dunque assumere che $x \neq 0 \neq y$.

Siano $x, y \in H$ e $\alpha \in \mathbb{F}$. Siccome $\langle \cdot, \cdot \rangle$ è bilineare (risp. sesquilineare), simmetrico (risp. hermitiano) e definito positivo, si ha
$$\begin{aligned} 0 \leq \langle x - \alpha y, x - \alpha y \rangle &= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\alpha} \langle x, y \rangle + \alpha \overline{\alpha} \langle y, y \rangle \\ &= \|x\|^2 + |\alpha|^2 \|y\|^2 - 2 \, \mathfrak{Re}(\overline{\alpha} \langle x, y \rangle), \end{aligned} \tag{6.7}$$
ove si è usato che $2 \, \mathfrak{Re}(\overline{\alpha} \langle x, y \rangle) = \overline{\alpha} \langle x, y \rangle + \overline{\overline{\alpha} \langle x, y \rangle} = \overline{\alpha} \langle x, y \rangle + \alpha \langle y, x \rangle$.

Notiamo ora che $\|y\|^2 = \langle y, y\rangle > 0$ perché $y \neq 0$ e $\langle \cdot, \cdot \rangle$ è definito positivo. Scegliendo allora $\alpha = \langle x, y\rangle / \|y\|^2$, da (6.7) deduciamo

$$0 \leq \|x\|^2 + \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2 - 2\frac{|\langle x, y\rangle|^2}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2},$$

cioè

$$\frac{|\langle x, y\rangle|^2}{\|y\|^2} \leq \|x\|^2;$$

moltiplicando ambo i membri per $\|y\|^2$ e prendendone la radice quadrata, otteniamo (6.6). □

Proposizione 6.11. *Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano. La norma $\|\cdot\|$ indotta dal prodotto scalare è una norma su H.*

Dimostrazione. Abbiamo già discusso (☞ osservazione 6.9) che la mappa $\|\cdot\| : H \times H \to [0, \infty)$ data dalla definizione 6.8 è ben definita. Verifichiamo ora le proprietà nella definizione 4.1.

Per ogni $x \in H$, siccome $\|x\|^2 = \langle x, x\rangle$ per definizione, e $\langle \cdot, \cdot \rangle$ è definito positivo, si ha $\|x\| = 0 \iff \|x\|^2 = 0 \iff \langle x, x\rangle = 0 \iff x = 0$. Dunque $\|\cdot\|$ soddisfa la proprietà di separazione.

Si ha inoltre, per ogni $\alpha \in \mathbb{F}$ e $x \in H$,

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x\rangle} = \sqrt{|\alpha|^2\|x\|^2} = |\alpha|\|x\|,$$

ove si è usata la bilinearità (risp. sesquilinearità) del prodotto scalare. Questo dimostra la 1-omogeneità di $\|\cdot\|$.

Infine, per ogni $x, y \in H$, siccome $\langle \cdot, \cdot \rangle$ è bilineare (risp. sesquilineare),

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\,\mathfrak{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 \end{aligned} \tag{6.8}$$

ove si è usato che $2\,\mathfrak{Re}\langle x, y\rangle = \langle x, y\rangle + \overline{\langle x, y\rangle} = \langle x, y\rangle + \langle y, x\rangle$ (per la simmetria/hermitianità) e che $|\langle x, y\rangle| \leq \|x\|\|y\|$ (per la disuguaglianza di Cauchy–Schwarz, ☞ proposizione 6.10). Prendendo le radici quadrate di primo e ultimo membro di (6.8) si deduce che

$$\|x+y\| \leq \|x\| + \|y\|,$$

cioè la disuguaglianza triangolare. □

Sulla base del precedente risultato, ogni spazio pre-hilbertiano $(H, \langle \cdot, \cdot \rangle)$ si può anche pensare come spazio normato, con la norma $\|\cdot\|$ indotta dal prodotto scalare. In particolare, agli spazi pre-hilbertiani si applica tutta la teoria degli spazi normati discussa in precedenza (e a maggior ragione quella degli spazi metrici).

Definizione 6.12. Uno *spazio di Hilbert* è uno spazio pre-hilbertiano completo. In altre parole, uno spazio di Hilbert è uno spazio pre-hilbertiano che, con la norma indotta dal prodotto scalare, è uno spazio di Banach.

Esempio 6.13. Ecco alcuni esempi e non-esempi di spazi di Hilbert.

(1) Gli spazi pre-hilbertiani $(\mathbb{F}^n, \langle \cdot, \cdot \rangle)$, $(\ell^2, \langle \cdot, \cdot \rangle)$ e $(L^2(M, \mathcal{M}, \mu), \langle \cdot, \cdot \rangle)$ discussi nell'esempio 6.7 sono tutti spazi di Hilbert. Infatti, abbiamo già visto (☞ esempio 4.5(1), proposizione 5.18, teorema 5.29) che, con le rispettive norme indotte, gli spazi $(\mathbb{F}^n, \| \cdot \|_2)$, $(\ell^2, \| \cdot \|_2)$ e $(L^2(M, \mathcal{M}, \mu), \| \cdot \|_2)$ sono spazi di Banach.

(2) Ogni spazio pre-hilbertiano di dimensione finita è uno spazio di Hilbert. Infatti, già sappiamo che ogni spazio normato di dimensione finita è uno spazio di Banach (☞ corollario 4.14).

(3) Lo spazio c_{00} con il prodotto scalare

$$\langle \underline{x}, \underline{y} \rangle = \sum_{k=0}^{\infty} x_k \overline{y_k}$$

indotto da $(\ell^2, \langle \cdot, \cdot \rangle)$ <u>non</u> è uno spazio di Hilbert. Infatti c_{00} è un sottospazio denso e non chiuso di ℓ^2 (☞ proposizione 5.6(iii)), pertanto $(c_{00}, \langle \cdot, \cdot \rangle)$ non è completo (☞ proposizione 4.16).

(4) Sia $[a, b]$ un intervallo chiuso e limitato di $\mathbb{R}$. Lo spazio $C[a, b]$ con il prodotto scalare integrale

$$\langle f, g \rangle = \int_a^b f(t)\, \overline{g(t)}\, dt$$

indotto da $(L^2(a, b), \langle \cdot, \cdot \rangle)$ <u>non</u> è uno spazio di Hilbert. Infatti $C[a, b]$ è un sottospazio denso e non chiuso di $L^2(a, b)$ (☞ teorema 5.37(iv)), pertanto $(C[a, b], \langle \cdot, \cdot \rangle)$ non è completo (☞ proposizione 4.16).

Proposizione 6.14. *Sia* $(H, \langle \cdot, \cdot \rangle)$ *uno spazio pre-hilbertiano su* $\mathbb{F}$.

(i) Per ogni $x, y \in H$,

$$4\langle x, y \rangle = \begin{cases} \|x + y\|^2 - \|x - y\|^2 & \textit{se } \mathbb{F} = \mathbb{R}, \\ \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 & \textit{se } \mathbb{F} = \mathbb{C} \end{cases}$$

(identità di polarizzazione).

(ii) Per ogni $x, y \in H$,

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{6.9}$$

(identità del parallelogramma).

Dimostrazione. (i). Come già visto in (6.8), per ogni $x, y \in H$,

$$\|x + y\|^2 = \|x\|^2 + 2\,\mathfrak{Re}\langle x, y \rangle + \|y\|^2, \tag{6.10}$$

e scambiando y con $-y$ si ottiene anche

$$\|x-y\|^2 = \|x\|^2 - 2\,\mathfrak{Re}\langle x, y\rangle + \|y\|^2. \tag{6.11}$$

Sottraendo la seconda identità alla prima, si ottiene dunque

$$\|x+y\|^2 - \|x-y\|^2 = 4\,\mathfrak{Re}\langle x, y\rangle.$$

Nel caso $\mathbb{F} = \mathbb{R}$, si ha $\mathfrak{Re}\langle x, y\rangle = \langle x, y\rangle$ e abbiamo finito. Se invece $\mathbb{F} = \mathbb{C}$, sostituendo y con iy, si ottiene anche

$$\|x+iy\|^2 - \|x-iy\|^2 = 4\,\mathfrak{Re}\langle x, iy\rangle = 4\,\mathfrak{Re}(-i\langle x, y\rangle) = 4\,\mathfrak{Im}\langle x, y\rangle.$$

Pertanto,

$$4\langle x, y\rangle = 4\,\mathfrak{Re}\langle x, y\rangle + i\,\mathfrak{Im}\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|,$$

come desiderato.

(ii). Sommando membro a membro (6.10) e (6.11), si ottiene l'identità desiderata. □

Osservazione 6.15. L'identità del parallelogramma (☞ Proposizione 6.14(ii)) corrisponde al seguente fatto geometrico: in un parallelogramma, la somma dei quadrati dei lati è uguale alla somma dei quadrati delle diagonali. La formula (6.9) esprime questa proprietà per il parallelogramma di vertici 0, x, $x+y$ e y.

Corollario 6.16. *Sia $(H, \langle\cdot,\cdot\rangle)$ uno spazio pre-hilbertiano. Allora la mappa $H \times H \ni (x, y) \mapsto \langle x, y\rangle \in \mathbb{F}$ è continua.*

Dimostrazione. L'identità di polarizzazione 6.14(i) permette di scrivere $\langle x, y\rangle$ come un polinomio nelle espressioni

$$\|x+y\|, \quad \|x-y\|, \quad \|x+iy\|, \quad \|x-iy\| \tag{6.12}$$

(le ultime due sono rilevanti solo se $\mathbb{F} = \mathbb{C}$). Per l'Algebra delle Funzioni Continue, è dunque sufficiente dimostrare che tali espressioni sono funzioni continue di (x, y). Siccome la norma $\|\cdot\|$ è continua su H (☞ proposizione 4.6(i)), e le operazioni di somma e prodotto scalare-vettore sono continue (☞ proposizione 4.6(ii)), per composizione di funzioni deduciamo anche la continuità delle espressioni in (6.12). □

Osservazione 6.17. In base alla proposizione 6.14(ii), se una norma $\|\cdot\|$ è indotta da un prodotto scalare, allora deve soddisfare l'identità del parallelogramma. In altre parole, la validità dell'identità del parallelogramma è una condizione necessaria affinché una norma sia indotta da un prodotto scalare. In realtà, è possibile dimostrare che tale condizione è anche sufficiente.

Esempio 6.18. Vediamo come si può utilizzare la non-validità dell'identità del parallelogramma per escludere che certe norme siano indotte da un prodotto scalare.

(1) Per $p \neq 2$, la norma $\|\cdot\|_p$ su ℓ^p non è indotta da un prodotto scalare. Infatti, se prendiamo $\underline{x} = (1, 0, 0, 0, \dots)$ e $\underline{y} = (0, 1, 0, 0, \dots)$, allora $\underline{x} \pm \underline{y} = (1, \pm 1, 0, 0, \dots)$, ed è facile verificare che

$$\|\underline{x}\|_p = \|\underline{y}\|_p = 1, \qquad \|\underline{x}+\underline{y}\|_p = \|\underline{x}-\underline{y}\|_p = 2^{1/p}.$$

In particolare

$$\|\underline{x} + \underline{y}\|_p^2 + \|\underline{x} - \underline{y}\|_p^2 = 2 \cdot 2^{1/p} \neq 4 = 2\|\underline{x}\|_p^2 + 2\|\underline{y}\|_p^2$$

perché $p \neq 2$.

(2) La norma $\| \cdot \|_\infty$ su $C[a,b]$ non è indotta da un prodotto scalare. Infatti, non è difficile costruire $f, g \in C[a,b]$ con supporti disgiunti e tali che $\|f\|_\infty = \|g\|_\infty = 1$. Di conseguenza, si ha anche $\|f+g\|_\infty = \|f-g\|_\infty = 1$ e quindi

$$\|f + g\|_\infty^2 + \|f - g\|_\infty^2 = 2 \neq 4 = 2\|f\|_\infty^2 + 2\|g\|_\infty^2.$$

6.2 Ortogonalità

Definizione 6.19. Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano.

(a) Siano $x, y \in H$. Diciamo che x è *ortogonale* a y, e scriviamo $x \perp y$, se $\langle x, y \rangle = 0$.

(b) Se $A, B \subseteq H$, scriviamo $A \perp B$ per indicare che $x \perp y$ per ogni $x \in A$ e $y \in B$.

(c) Se $x \in H$ e $A \subseteq H$, scriviamo anche $x \perp A$ al posto di $\{x\} \perp A$.

Osservazione 6.20. Per la simmetria/hermitianità del prodotto scalare, si ha $\langle x, y \rangle = \overline{\langle y, x \rangle}$, quindi $x \perp y$ se e solo se $y \perp x$. Dalla bilinearità/sesquilinearità deduciamo anche che $x \perp 0$ per ogni $x \in H$. Infine, dalla definita positività deduciamo che $x \not\perp x$ se $x \neq 0$.

Definizione 6.21. Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano.

(a) Un insieme $S \subseteq H$ si dice *ortogonale* se $x \perp y$ per ogni $x, y \in S$ con $x \neq y$.

(b) Un insieme $S \subseteq H$ si dice *ortonormale* se S è ortogonale e $\|x\| = 1$ per ogni $x \in S$.

Osservazione 6.22. Siano $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano e $\{x_i\}_{i \in I} \subseteq H$ indicizzato iniettivamente (cioè $x_i \neq x_j$ se $i \neq j$). Allora $\{x_i\}_{i \in I}$ è ortonormale se e solo se

$$\langle x_i, x_j \rangle = \delta_{i,j} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

Il simbolo $\delta_{i,j}$ usato sopra è detto *delta di Kronecker.*

Notazione. Nel seguito, scriveremo "l'insieme ortogonale $\{x_i\}_{i \in I}$" o "l'insieme ortonormale $\{x_i\}_{i \in I}$" sottintendendo che l'indicizzazione è iniettiva.

Proposizione 6.23. *Sia* $(H, \langle \cdot, \cdot \rangle)$ *uno spazio pre-hilbertiano.*

(i) Se $S \subseteq H \setminus \{0\}$ *è ortogonale, allora è linearmente indipendente.*

(ii) (teorema di Pitagora) *Sia* $n \in \mathbb{N}_+$. *Se* $\{x_1, \dots, x_n\} \subseteq H$ *è ortogonale, allora*

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

(iii) Sia $n \in \mathbb{N}_+$. Sia $\{x_1, \dots, x_n\} \subseteq H$ ortonormale. Allora

$$\left\langle \sum_{j=1}^n \alpha_j x_j, \sum_{j=1}^n \beta_j x_j \right\rangle = \sum_{j=1}^n \alpha_j \overline{\beta_j}, \qquad \left\| \sum_{j=1}^n \alpha_j x_j \right\| = \sqrt{\sum_{j=1}^n |\alpha_j|^2}$$

per ogni $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{F}$.

Dimostrazione. (i). Supponiamo che

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 0$$

per qualche $n \in \mathbb{N}_+$, $x_1, \dots, x_n \in S$ distinti e $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. Allora, per ogni $k = 1, \dots, n$,

$$0 = \langle \alpha_1 x_1 + \dots + \alpha_n x_n, x_k \rangle = \sum_{j=1}^n \alpha_j \langle x_j, x_k \rangle = \alpha_k \|x_k\|^2, \tag{6.13}$$

dove si è usata la linearità di $\langle \cdot, \cdot \rangle$ nella prima componente, e il fatto che $x_j \perp x_k$ per $j \neq k$, dato che S è ortogonale. Siccome $0 \notin S$ per ipotesi, si ha $\|x_k\| \neq 0$ (dato che $\|\cdot\|$ è una norma, ☞ proposizione 6.11), ma allora da (6.13) deduciamo che $\alpha_k = 0$. Siccome questo vale per ogni $k = 1, \dots, n$, abbiamo dimostrato che $\alpha_1 = \dots = \alpha_k = 0$. Questo mostra che S è linearmente indipendente (☞ definizione 1.9(c)).

(ii). Si ha

$$\left\| \sum_{j=1}^n x_j \right\|^2 = \left\langle \sum_{j=1}^n x_j, \sum_{k=1}^n x_k \right\rangle = \sum_{j=1}^n \sum_{k=1}^n \langle x_j, x_k \rangle = \sum_{j=1}^n \langle x_j, x_j \rangle = \sum_{j=1}^n \|x_j\|^2,$$

dove si sono usati la bilinearità/sesquilinearità del prodotto scalare e il fatto che $x_j \perp x_k$ se $j \neq k$.

(iii). Si ha

$$\left\langle \sum_{j=1}^k \alpha_j x_j, \sum_{k=1}^n \beta_k x_k \right\rangle = \sum_{j=1}^k \sum_{k=1}^n \alpha_j \overline{\beta_k} \langle x_j, x_k \rangle = \sum_{j=1}^k \sum_{k=1}^n \alpha_j \overline{\beta_k} \delta_{j,k} = \sum_{j=1}^n \alpha_j \overline{\beta_j},$$

dove si sono usati la bilinearità/sesquilinearità del prodotto scalare e l'osservazione 6.22. Questo dimostra la prima identità, e la seconda si ottiene dalla prima prendendo $\beta_j = \alpha_j$ per $j = 1, \dots, n$. □

Teorema 6.24 (algoritmo di ortonormalizzazione di Gram–Schmidt)**.** *Siano $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano, $n \in \mathbb{N}_+$, e $\{x_1, \dots, x_n\} \subseteq H$ un sottoinsieme linearmente indipendente. Allora esiste un insieme $\{e_1, \dots, e_n\} \subseteq H$ ortonormale tale che*

$$\operatorname{span}\{x_1, \dots, x_j\} = \operatorname{span}\{e_1, \dots, e_j\} \qquad \text{per } j = 1, \dots, n.$$

Dimostrazione. Procediamo per induzione su $n \in \mathbb{N}_+$.

Nel caso base ($n = 1$) basta porre $e_1 = x_1/\|x_1\|$.

Supponiamo ora $n > 1$. Applicando l'ipotesi induttiva all'insieme linearmente indipendente $\{x_1, \dots, x_{n-1}\}$, otteniamo un insieme $\{e_1, \dots, e_{n-1}\}$ ortonormale tale che

$$\operatorname{span}\{x_1, \dots, x_j\} = \operatorname{span}\{e_1, \dots, e_j\} \qquad \text{per } j = 1, \dots, n-1.$$

Definiamo ora

$$\tilde{e}_n = x_n - \sum_{j=1}^{n-1} \langle x_n, e_j \rangle e_j. \tag{6.14}$$

Per indipendenza lineare, si ha $x_n \notin \operatorname{span}\{x_1, \dots, x_{n-1}\} = \operatorname{span}\{e_1, \dots, e_{n-1}\}$, dunque da (6.14) si vede che $\tilde{e}_n \neq 0$. Possiamo dunque definire $e_n = \tilde{e}_n / \|\tilde{e}_n\|$.

Per costruzione, si ha $\|e_n\| = 1$. Inoltre, per ogni $k = 1, \dots, n-1$,

$$\begin{aligned} \langle e_n, e_k \rangle = \frac{1}{\|\tilde{e}_n\|} \langle \tilde{e}_n, e_k \rangle &= \frac{1}{\|\tilde{e}_n\|} \left[\langle x_n, e_k \rangle - \sum_{j=1}^{n-1} \langle x_n, e_j \rangle \langle e_j, e_k \rangle \right] \\ &= \frac{1}{\|\tilde{e}_n\|} \left[\langle x_n, e_k \rangle - \langle x_n, e_k \rangle \right] \\ &= 0, \end{aligned}$$

dove si è usata la bilinearità/sesquilinearità del prodotto scalare e il fatto che $\langle e_j, e_k \rangle = \delta_{j,k}$ per $j, k = 1, \dots, n-1$, dato che $\{e_1, \dots, e_{n-1}\}$ è ortonormale.

Dunque, e_n è un vettore di norma unitaria, che è ortogonale a ciascuno degli $e_1, \dots, e_{n-1}$. Aggiungendo allora e_n all'insieme ortonormale $\{e_1, \dots, e_{n-1}\}$ si ottiene un nuovo insieme ortonormale $\{e_1, \dots, e_n\}$. Infine, dalle definizioni di e_n ed $\tilde{e}_n$ segue facilmente che

$$\begin{aligned} \operatorname{span}\{e_1, \dots, e_n\} &= \operatorname{span}\{e_1, \dots, e_{n-1}, \tilde{e}_n\} \\ &= \operatorname{span}\{e_1, \dots, e_{n-1}, x_n\} = \operatorname{span}\{x_1, \dots, x_{n-1}, x_n\}, \end{aligned}$$

dove nell'ultimo passaggio si è usato il fatto già noto che $\operatorname{span}\{e_1, \dots, e_{n-1}\} = \operatorname{span}\{x_1, \dots, x_{n-1}\}$. □

Definizione 6.25. Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano.

(a) Sia $A \subseteq H$. Il *complemento ortogonale* di A è l'insieme

$$A^\perp = \{y \in H : y \perp A\}.$$

(b) Per ogni $x \in H$, scriviamo anche $x^\perp$ al posto di $\{x\}^\perp$.

Proposizione 6.26. *Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano. Siano $A, B \subseteq H$.*

(i) $A^\perp = \bigcap_{x \in A} x^\perp$.

(ii) $A^\perp$ *è un sottospazio vettoriale chiuso di* H, *e inoltre* $A \cap A^\perp \subseteq \{0\}$.

(iii) Se $A \subseteq B$ *allora* $A^\perp \supseteq B^\perp$.

(iv) $(\operatorname{span} A)^\perp = \overline{A}^\perp = A^\perp$.

(v) $0^\perp = H$ *e* $H^\perp = \{0\}$.

(vi) $A^{\perp\perp} \supseteq A$.

Dimostrazione. (i). Per ogni $y \in H$, si ha $y \in A^\perp$ se e solo se y è ortogonale a tutti gli elementi di A, cioè se e solo se $y \in x^\perp$ per ogni $x \in A$.

(ii). Notiamo anzitutto che, se $x \in A \cap A^\perp$, allora $x \perp x$ e quindi $x = 0$ (☞ osservazione 6.20). Questo mostra che $A \cap A^\perp \subseteq \{0\}$.

Rimane da dimostrare che $A^\perp$ è un sottospazio vettoriale chiuso di H. Ricordiamo che $A^\perp = \bigcap_{x \in A} x^\perp$ per (i). Siccome l'intersezione di sottospazi vettoriali è un sottospazio vettoriale (☞ proposizione 1.4), e inoltre l'intersezione di sottoinsiemi chiusi è un chiuso (☞ proposizione 2.23(iii)), è sufficiente verificare che $x^\perp$ è un sottospazio vettoriale chiuso di H.

D'altra parte $x^\perp = \{y \in H : y \perp x\} = \{y \in H : \langle x, y\rangle = 0\} = f^{-1}(\{0\})$ è la controimmagine dell'insieme $\{0\}$ tramite la mappa $f : H \ni x \mapsto \langle x, y\rangle \in \mathbb{F}$. Siccome tale mappa è continua (☞ corollario 6.16) e $\{0\}$ è chiuso in $\mathbb{C}$, ne concludiamo che $x^\perp$ è chiuso in H (☞ proposizione 2.28). Inoltre, la mappa $f : H \to \mathbb{F}$ è lineare (per la definizione di prodotto scalare), dunque $x^\perp = f^{-1}(\{0\}) = \operatorname{Ker} f$ (☞ definizione 1.19(c)), e quindi $x^\perp$ è un sottospazio vettoriale di H (☞ proposizione 1.20(vii)).

(iii). Usando l'espressione dimostrata in (i), siccome $A \subseteq B$ si ha

$$A^\perp = \bigcap_{x \in A} x^\perp \supseteq \bigcap_{x \in B} x^\perp = B^\perp,$$

dato che, intersecando più insiemi, l'intersezione può solo rimpicciolire.

(iv). Siccome $A \subseteq \overline{A}$ e $A \subseteq \operatorname{span} A$, da (iii) segue immediatamente che $(\operatorname{span} A)^\perp \subseteq A^\perp$ e $\overline{A}^\perp \subseteq A^\perp$. Rimangono da dimostrare le inclusioni opposte.

D'altra parte, se $y \in A^\perp$ e $x \in \operatorname{span} A$, allora possiamo scrivere $x = \sum_{j=1}^n \alpha_j x_j$ per qualche $n \in \mathbb{N}$, $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ e $x_1, \dots, x_n \in A$ (☞ definizione 1.9(d)). Siccome $y \in A^\perp$, si ha allora $y \perp x_j$ per ogni $j = 1, \dots, n$, dunque $x_1, \dots, x_n \in y^\perp$; siccome quest'ultimo è un sottospazio vettoriale per (ii), ne deduciamo che anche $x = \sum_{j=1}^n \alpha_j x_j \in y^\perp$, cioè $y \perp x$. Dato che x è un arbitrario elemento di $\operatorname{span} A$, ne deduciamo che $y \in (\operatorname{span} A)^\perp$. Siccome d'altra parte y è un arbitrario elemento di $A^\perp$, ne deduciamo che $A^\perp \subseteq (\operatorname{span} A)^\perp$.

Similmente, siano $y \in A^\perp$ e $x \in \overline{A}$. Allora esiste una successione $(x_n)_{n\in\mathbb{N}}$ a valori in A tale che $x_n \to x$ in H per $n \to \infty$. Siccome $x_n \in A$ e $y \in A^\perp$, si ha dunque $x_n \perp y$, cioè $x_n \in y^\perp$ per ogni $n \in \mathbb{N}$, cioè $(x_n)_n$ è a valori in $y^\perp$. Siccome da (ii) sappiamo che $y^\perp$ è chiuso, ne concludiamo che anche il limite x della successione sta in $y^\perp$, cioè $y \perp x$. Dato che x è un arbitrario elemento di $\overline{A}$, ne deduciamo che $y \in \overline{A}^\perp$. Siccome d'altra parte y è un arbitrario elemento di $A^\perp$, ne deduciamo che $A^\perp \subseteq \overline{A}^\perp$.

(v). Siccome $x \perp 0$ per ogni $x \in H$ (☞ osservazione 6.20), si ha $0^\perp = H$. D'altra parte, l'unico vettore ortogonale a tutti gli elementi di H è il vettore nullo (☞ osservazione 6.20: un vettore non può essere ortogonale a se stesso a meno che non sia nullo), dunque $H^\perp = \{0\}$.

(vi). Sia $x \in A$. Allora, per ogni $y \in A^\perp$, si ha $x \perp y$; pertanto x è ortogonale a tutti gli elementi di $A^\perp$, cioè $x \in A^{\perp\perp}$. □

Vediamo ora che il complemento ortogonale di un sottospazio vettoriale Y di uno spazio pre-hilbertiano si può caratterizzare in termini della distanza dall'insieme Y, definita come in (2.6).

Proposizione 6.27. *Siano $(H, \langle\cdot,\cdot\rangle)$ uno spazio pre-hilbertiano e Y un sottospazio vettoriale di H. Per ogni $x \in H$, le seguenti condizioni sono equivalenti:*

(i) $x \in Y^\perp$;

(ii) $\|x - y\| \geq \|x\|$ *per ogni* $y \in Y$;

(iii) $d(x, Y) = \|x\|$.

Dimostrazione. (i) $\Rightarrow$ (ii). Per ogni $x \in Y^\perp$ e $Y \in Y$, si ha $x \perp y$, dunque da (6.11) deduciamo

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 \geq \|x\|^2.$$

(ii) $\Rightarrow$ (iii). Da (ii) e (2.6) deduciamo

$$d(x, Y) = \inf_{y \in Y} \|x - y\| \geq \|x\|,$$

e in realtà vale l'uguaglianza, perché possiamo prendere $y = 0$, dato che Y è un sottospazio e quindi $0 \in Y$.

(iii) $\Rightarrow$ (i). Per ogni $y \in Y$,

$$\|x\|^2 \leq d(x, Y)^2 \leq \|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\,\mathfrak{Re}\langle x, y\rangle,$$

dove abbiamo usato (iii), (2.6) e (6.11); ne deduciamo che

$$2\,\mathfrak{Re}\langle x, y\rangle \leq \|y\|^2. \tag{6.15}$$

Siccome Y è un sottospazio, per ogni $y \in Y$ e $t \in (0, \infty)$, si ha anche $t\langle x, y\rangle y \in Y$; rimpiazzando y con $t\langle x, y\rangle y$ in (6.15) otteniamo

$$2t|\langle x, y\rangle|^2 \leq t^2 |\langle x, y\rangle|^2 \|y\|^2,$$

e dividendo per t ambo i membri,

$$2|\langle x, y\rangle|^2 \leq t|\langle x, y\rangle|^2 \|y\|^2.$$

Prendendo il limite per $t \to 0^+$, si ottiene infine $|\langle x, y\rangle| = 0$, cioè $x \perp y$. Siccome $y \in Y$ era arbitrario, ne deduciamo $x \in Y^\perp$, come desiderato. □

Ricordiamo (☞ definizione 1.7) la definizione di sottoinsieme convesso di uno spazio vettoriale.

Teorema 6.28 (proiezione sui convessi chiusi)**.** *Sia $(H, \langle\cdot,\cdot\rangle)$ uno spazio di Hilbert. Sia $C \subseteq H$ un sottoinsieme convesso, chiuso e non vuoto di H. Allora, per ogni $x \in X$, esiste un unico $y \in C$, detto* proiezione di x su C, *che minimizza la distanza di x da C, cioè tale che*

$$\|x - y\| = d(x, C).$$

Dimostrazione. Unicità. Preso $x \in H$, supponiamo che esistano $y, y' \in C$ che minimizzano entrambi la distanza di x da C, cioè

$$\|x - y\| = \|x - y'\| = d(x, C). \tag{6.16}$$

Siccome C è convesso, si ha anche $(y+y')/2 \in C$, dunque

$$d(x,C)^2 \leq \|x-(y+y')/2\|^2 = \frac{1}{4}\|(x-y)+(x-y')\|^2, \tag{6.17}$$

dove si sono usate (2.6) e la 1-omogeneità della norma. Applicando l'identità del parallelogramma (☞ proposizione 6.14(ii)) ai vettori $x-y$ e $x-y'$, otteniamo dunque

$$\begin{aligned} d(x,C)^2 &\leq \frac{1}{4}\left[2\|x-y\|^2 + 2\|x-y'\|^2 - \|(x-y)-(x-y')\|^2\right] \\ &= d(x,C)^2 - \frac{1}{4}\|y-y'\|^2, \end{aligned} \tag{6.18}$$

dove si è usata (6.16). Ne deduciamo $\|y-y'\|^2 \leq 0$, dunque $\|y-y'\| = 0$ e $y = y'$.

Esistenza. Per la definizione (2.6) di $d(x,C)$, siccome $C \neq \emptyset$, esiste una successione $(y_n)_{n\in\mathbb{N}}$ a valori in C tale che la successione delle distanze $(\|x-y_n\|)_{n\in\mathbb{N}}$ è decrescente e

$$d(x,C) = \inf_{n\in\mathbb{N}} \|x-y_n\| = \lim_{n\to\infty} \|x-y_n\|. \tag{6.19}$$

Mostriamo ora che la successione $(y_n)_{n\in\mathbb{N}}$ è di Cauchy in H. Per ogni $n,m \in \mathbb{N}$, siccome C è convesso, si ha $(y_n+y_m)/2 \in C$; ragionando come in (6.17)-(6.18), usando l'identità del parallelogramma, deduciamo che

$$d(x,C)^2 \leq \frac{1}{4}\left[2\|x-y_m\|^2 + 2\|x-y_n\|^2 - \|y_m-y_n\|^2\right],$$

cioè

$$\|y_m-y_n\|^2 \leq 2(\|x-y_m\|^2 - d(x,C)^2) + 2(\|x-y_n\|^2 - d(x,C)^2). \tag{6.20}$$

D'altra parte, da (6.19) deduciamo che

$$\lim_{n\to\infty} (\|x-y_n\|^2 - d(x,C)^2) = 0.$$

Pertanto, per ogni $\epsilon > 0$, esiste $N \in \mathbb{N}$ tale che, per ogni $n > N$, $\|x-y_n\|^2 - d(x,C)^2 < \epsilon^2/4$. Da (6.20) deduciamo allora che, per ogni $n,m > N$,

$$\|y_m-y_n\|^2 < 2\epsilon^2/4 + 2\epsilon^2/4 = \epsilon^2,$$

cioè $\|y_m-y_n\| < \epsilon$. Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $(y_n)_n$ è una successione di Cauchy.

Siccome H è uno spazio di Hilbert, H è completo e quindi la successione di Cauchy $(y_n)_n$ converge in H a un punto $y \in H$. Siccome $y_n \in C$ per ogni $n \in \mathbb{N}$, e C è un sottoinsieme chiuso di H, ne deduciamo che $y \in C$ (☞ proposizione 2.23(vi)). Siccome $y_n \to y$, da (6.19) e dalla continuità della norma e delle operazioni di spazio vettoriale (☞ proposizione 4.6) deduciamo allora che

$$d(x,C) = \|x-y\|,$$

cioè $y \in C$ minimizza la distanza di x da C, come desiderato. □

Osservazione 6.29. Nella dimostrazione data sopra dell'*unicità* della proiezione, non si usano le ipotesi che C sia chiuso o che H sia completo. È tuttavia fondamentale l'ipotesi di convessità di C. Se, ad esempio, si prende come C una circonferenza nel piano $H = \mathbb{R}^2$ e come x il centro della circonferenza, ovviamente tutti i punti di C minimizzano la distanza di x da C, quindi in questo caso l'unicità fallisce, ma d'altra parte la circonferenza C non è convessa. La chiusura di C e la completezza di H sono invece ipotesi essenziali per la dimostrazione dell'esistenza.

Osservazione 6.30. L'esempio della circonferenza nell'osservazione 6.29 mostra che la convessità di C non è strettamente necessaria per l'esistenza di punti minimizzanti la distanza; ad esempio, la convessità si può sostituire con la *compattezza* di C. Infatti, se C è compatto, presa una successione minimizzante $(y_n)_n$ come nella dimostrazione sopra, si può usare la compattezza di C per estrarne una sottosuccessione convergente, e il limite di tale sottosuccessione è il punto cercato. Il fatto che la compattezza di C permetta di mostrare l'esistenza di massimi e minimi non stupisce (si pensi al teorema di Weierstrass, ☞ teorema 2.36). Viceversa, è particolarmente interessante che il teorema 6.28 si applichi anche a insiemi C ben lontani dall'essere compatti (purché convessi), come i sottospazi vettoriali chiusi non banali di H.

Ricordiamo che i sottospazi vettoriali di uno spazio di Hilbert H sono convessi (☞ osservazione 1.8). Il teorema 6.28 si applica dunque in particolare al caso in cui C sia un sottospazio vettoriale chiuso di H.

Definizione 6.31. Sia H uno spazio di Hilbert e Y un sottospazio vettoriale chiuso di H. Per ogni $x \in H$, l'unico punto di Y che minimizza la distanza di x da Y (nel senso del teorema 6.28) è detto *proiezione ortogonale* di x su Y e denotato con $P_Y x$. In altre parole, la proiezione ortogonale $P_Y x$ di x su Y è caratterizzata dalla seguente proprietà:

$$\forall y \in Y : (y = P_Y x \iff d(x, Y) = \|x - y\|) .$$

L'esistenza della proiezione ortogonale su un sottospazio chiuso ci permette di dimostrare il seguente fondamentale risultato di decomposizione ortogonale in spazi di Hilbert.

Teorema 6.32 (decomposizione ortogonale)**.** *Siano H uno spazio di Hilbert e Y un suo sottospazio vettoriale chiuso. Per ogni $x \in H$ esistono un unico $y \in Y$ e un unico $z \in Y^\perp$ tali che*

$$x = y + z. \tag{6.21}$$

Inoltre, in tal caso,

$$y = P_Y x \qquad e \qquad \|x\|^2 = \|y\|^2 + \|z\|^2 \geq \|y\|^2. \tag{6.22}$$

Dimostrazione. Unicità. Supponiamo esistano $y, y' \in Y$ e $z, z' \in Y^\perp$ tali che

$$x = y + z = y' + z'.$$

Allora $y - y' = z' - z$. D'altra parte, siccome Y e $Y^\perp$ sono sottospazi di H (☞ proposizione 6.26(ii)), si ha $y - y' \in Y$ e $z' - z \in Y^\perp$, dunque

$$y - y' = z' - z \in Y \cap Y^\perp = \{0\}$$

(☞ proposizione 6.26(ii)). Pertanto $y - y' = z' - z = 0$, cioè $y = y'$ e $z = z'$.

Esistenza e (6.22). Dato $x \in H$, poniamo $y = P_Y x$ e $z = x - y$. Allora chiaramente (6.21) vale, e inoltre $y \in Y$ e

$$\|z\| = \|x - y\| = d(x, Y)$$

(☞ definizione 6.31).

Vogliamo ora dimostrare che $z \in Y^\perp$. In effetti, per ogni $w \in Y$, si ha

$$\|z - w\| = \|(x - y) - w\| = \|x - (y + w)\| \geq d(x, Y) = \|z\|,$$

dove si è usato che $y + w \in Y$ (perché Y è un sottospazio vettoriale) e la definizione di $d(x, Y)$. Siccome questa disuguaglianza vale per ogni $w \in Y$, dalla proposizione 6.27 deduciamo che $z \in Y^\perp$.

In conclusione, abbiamo trovato $y \in Y$ e $z \in Y^\perp$ tali che (6.21) vale, e inoltre $y = P_Y x$ per costruzione. Infine, siccome $y \in Y$ e $z \in Y^\perp$, si ha $y \perp z$, dunque da (6.21) e dal teorema di Pitagora (☞ proposizione 6.23(ii)) deduciamo che

$$\|x\|^2 = \|y\|^2 + \|z\|^2 \geq \|y\|^2,$$

cioè le proprietà in (6.22) sono verificate. □

Corollario 6.33. *Sia H uno spazio di Hilbert. Sia Y un sottospazio vettoriale chiuso di H. Sia $A \subseteq H$. Allora:*

(i) $H = Y \oplus Y^\perp$.

(ii) $Y^{\perp\perp} = Y$.

(iii) *Per ogni $x \in H$, si ha*

$$x = P_Y x + P_{Y^\perp} x, \qquad \|x\|^2 = \|P_Y x\|^2 + \|P_{Y^\perp} x\|^2 \geq \|P_Y x\|^2. \tag{6.23}$$

(iv) *La mappa di proiezione ortogonale $P_Y : H \to Y$, che associa a ogni $x \in H$ la sua proiezione ortogonale $P_Y x \in Y$, è lineare e 1-lipschitziana.*

(v) *Per ogni $x \in H$, si ha $d(x, Y) = \|P_{Y^\perp} x\|$.*

(vi) $A^{\perp\perp} = \overline{\operatorname{span} A}$.

Dimostrazione. (i). Questo fatto è una riscrittura dell'esistenza e unicità della decomposizione ortogonale data dal teorema 6.32 (☞ osservazione 1.11).

(ii). Sappiamo già che $Y \subseteq Y^{\perp\perp}$ (☞ proposizione 6.26(vi)), quindi rimane da verificare l'inclusione opposta. Sia $x \in Y^{\perp\perp}$. Applicando il teorema 6.32 a x, possiamo scrivere $x = y+z$, ove $y \in Y$ e $z \in Y^\perp$. In particolare, $z = x-y \in Y^{\perp\perp}$ (dato che $x \in Y^{\perp\perp}$, $y \in Y \subseteq Y^{\perp\perp}$ e $Y^{\perp\perp}$ è un sottospazio vettoriale); siccome sappiamo già che $z \in Y^\perp$, ne concludiamo che $z \in Y^\perp \cap Y^{\perp\perp} = \{0\}$ (☞ proposizione 6.26(ii)), cioè $z = 0$ e $x = y \in Y$. Per l'arbitrarietà di $x \in Y^{\perp\perp}$, questo dimostra l'inclusione $Y^{\perp\perp} \subseteq Y$, come desiderato.

(iii). Applicando il teorema 6.32 al vettore x e al sottospazio Y, otteniamo la decomposizione $x = y + z$ con $y = P_Y x \in Y$ e $z \in Y^\perp$. D'altra parte, anche $Y^\perp$ è un sottospazio vettoriale chiuso di H (☞ proposizione 6.26(ii)), quindi possiamo applicare il teorema 6.32 al vettore x e al sottospazio $Y^\perp$, ottenendo la decomposizione $x = u+w$ con $u = P_{Y^\perp} x \in Y^\perp$ e $w \in Y^{\perp\perp}$. D'altra parte, da

(ii) sappiamo che $Y^{\perp\perp} = Y$, pertanto $w \in Y$. In altre parole, abbiamo ottenuto due decomposizioni

$$x = y + z = w + u$$

con $y, w \in Y$ e $z, u \in Y^{\perp}$; l'unicità della decomposizione ortogonale (☞ teorema 6.32) implica allora che $w = y = P_Y x$ e $z = u = P_{Y^\perp} x$. Le proprietà (6.23) sono dunque una riscrittura di (6.21) e (6.22).

(iv). Dimostriamo che $P_Y : H \to Y$ è lineare. Siano $x, x' \in H$, $\alpha, \alpha' \in \mathbb{F}$. Per il teorema 6.32, possiamo decomporre

$$x = y + z, \qquad x' = y' + z'$$

con $y, y' \in Y$, $z, z' \in Y^{\perp}$, e si ha $y = P_Y x$, $y' = P_Y x'$. Di conseguenza, si ha anche

$$\alpha x + \alpha' x' = (\alpha y + \alpha' y') + (\alpha z + \alpha' z'), \tag{6.24}$$

e inoltre $\alpha y + \alpha' y' \in Y$ e $\alpha z + \alpha' z' \in Y^{\perp}$, perché Y e $Y^{\perp}$ sono sottospazi vettoriali. In altre parole, (6.24) è una decomposizione ortogonale di $\alpha x + \alpha' x'$, dunque l'unicità della decomposizione ortogonale (☞ teorema 6.32) implica che

$$P_Y(\alpha x + \alpha' x') = \alpha y + \alpha' y' = \alpha P_Y x + \alpha' P_Y x'.$$

Questo prova la linearità della mappa P_Y.

Rimane da verificare la 1-lipschitzianità. Si ha infatti, per ogni $x, x' \in H$,

$$\|P_Y x - P_Y x'\| = \|P_Y(x - x')\| \leq \|x - x'\|,$$

dove nel primo passaggio si è usata la linearità di P_Y, e nel secondo la disuguaglianza in (6.23).

(v). Per definizione di proiezione ortogonale (☞ definizione 6.31),

$$d(x, Y) = \|x - P_Y x\| = \|P_{Y^\perp} x\|,$$

dove si è usata (6.23).

(vi). Applicando (ii) con $Y = \overline{\operatorname{span} A}$, si ottiene $\overline{\operatorname{span} A} = \overline{\operatorname{span} A}^{\perp\perp}$. D'altra parte, dalla proposizione 6.26(iv) sappiamo che $\overline{\operatorname{span} A}^{\perp} = A^{\perp}$ e quindi anche $\overline{\operatorname{span} A}^{\perp\perp} = A^{\perp\perp}$. □

6.3 Basi ortonormali

Riprendiamo la discussione dei sottoinsiemi ortonormali di uno spazio prehilbertiano H cominciata con la definizione 6.21 e l'osservazione 6.22. I risultati ad essi relativi visti finora (☞ proposizione 6.23 e teorema 6.24) si concentrano sul caso di insiemi ortonormali finiti; tuttavia, se H ha dimensione infinita, è possibile trovare anche insiemi ortonormali infiniti.

Esempio 6.34. Ecco alcuni esempi di insiemi ortonormali infiniti.

(1) Sia $H = \ell^2$ con il prodotto scalare dell'esempio 6.7(2). Per ogni $n \in \mathbb{N}$, definiamo $\underline{e^{(n)}} \in c_{00} \subseteq \ell^2$ ponendo

$$e_k^{(n)} = \delta_{n,k}$$

per ogni $k \in \mathbb{N}$. Allora, per ogni $n, m \in \mathbb{N}$,

$$\langle \underline{e^{(n)}}, \underline{e^{(m)}} \rangle = \sum_{k=0}^{\infty} e_k^{(n)} \overline{e_k^{(m)}} = \sum_{k=0}^{\infty} \delta_{n,k}\delta_{m,k} = \delta_{n,m}.$$

Dunque $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ è un sottoinsieme ortonormale di ℓ^2 (☞ osservazione 6.22).

(2) Assumiamo che $\mathbb{F} = \mathbb{C}$. Sia $H = L^2(-\pi, \pi)$ con il prodotto scalare dell'esempio 6.7(3). Per ogni $n \in \mathbb{Z}$, definiamo $\phi_n \in C[-\pi, \pi] \subseteq L^2(-\pi, \pi)$ ponendo

$$\phi_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}$$

per ogni $t \in [-\pi, \pi]$. Allora, per ogni $n, m \in \mathbb{Z}$,

$$\langle \phi_n, \phi_m \rangle = \int_{-\pi}^{\pi} \phi_n(t)\overline{\phi_m(t)}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t} = \delta_{n,m}.$$

Dunque $\{\phi_n\}_{n\in\mathbb{Z}}$ è un sottoinsieme ortonormale di $L^2(-\pi, \pi)$ (☞ osservazione 6.22).

Sappiamo già (☞ proposizione 6.23(i)) che ogni sottoinsieme ortonormale di uno spazio pre-hilbertiano è linearmente indipendente. Ci chiediamo ora se sia possibile utilizzare sottoinsiemi ortonormali come "basi".

Osservazione 6.35. La costruzione del sottoinsieme ortonormale $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ di ℓ^2 dell'esempio 6.34(1) è chiaramente analoga a quella della base canonica di $\mathbb{F}^n$ (☞ esempio 1.16). Tuttavia, nel caso di ℓ^2, si ha

$$\operatorname{span}\{\underline{e^{(n)}}\}_{n\in\mathbb{N}} = c_{00}. \tag{6.25}$$

Infatti, l'inclusione $\subseteq$ è dovuta al fatto che $\underline{e^{(n)}} \in c_{00}$ per ogni $n \in \mathbb{N}$, e che c_{00} è chiuso per combinazioni lineari (si ricordi che $\operatorname{span} A$ è l'insieme delle combinazioni lineari <u>finite</u> di elementi di A, ☞ definizione 1.9). Per l'inclusione opposta, basta osservare che, per ogni $\underline{x} = (x_k)_{k\in\mathbb{N}} \in c_{00}$, esiste $N \in \mathbb{N}$ tale che $x_k = 0$ per ogni $k > N$, pertanto

$$\underline{x} = (x_0, x_1, \dots, x_N, 0, 0, 0, \dots) = \sum_{k=0}^{N} x_k \underline{e^{(k)}} \in \operatorname{span}\{\underline{e^{(k)}}\}_{k\in\mathbb{N}}.$$

Siccome c_{00} è incluso propriamente in ℓ^2 (☞ proposizione 5.17(iii)), da (6.25) deduciamo che $\operatorname{span}\{\underline{e^{(n)}}\}_{n\in\mathbb{N}} \neq \ell^2$. In altre parole, $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ non è una base di ℓ^2 nel senso dell'algebra lineare (☞ definizione 1.9). Vedremo tuttavia che, nel contesto degli spazi di Hilbert, è possibile dare un senso anche a opportune "combinazioni lineari infinite" di elementi di $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$, e in questo senso sarà possibile considerare $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ come "base" di ℓ^2 (☞ esempio 6.43).

Proposizione 6.36 (disuguaglianza di Bessel). *Siano $(H, \langle\cdot,\cdot\rangle)$ uno spazio pre-hilbertiano e $x \in H$.*

(i) Sia $\{e_0, \dots, e_N\}$ un sottoinsieme ortonormale di H, per qualche $N \in \mathbb{N}$. Allora

$$\sum_{n=0}^{N} |\langle x, e_n \rangle|^2 \leq \|x\|^2. \tag{6.26}$$

(ii) Sia $\{e_n\}_{n\in\mathbb{N}}$ un sottoinsieme ortonormale di H. Allora

$$\sum_{n=0}^{\infty} |\langle x, e_n\rangle|^2 \leq \|x\|^2. \tag{6.27}$$

In particolare, $(\langle x, e_n\rangle)_{n\in\mathbb{N}} \in \ell^2$.

Dimostrazione. (i). Sia $x_N = \sum_{j=0}^{N} \langle x, e_n\rangle e_n$. Allora, per ogni $n = 0, \dots, N$,

$$\langle x - x_N, e_n\rangle = \langle x, e_n\rangle - \langle x_N, e_n\rangle = \langle x, e_n\rangle - \langle x, e_n\rangle = 0,$$

dove per il calcolo di $\langle x_N, e_n\rangle$ si è usata la proposizione 6.23(iii) applicata all'insieme ortonormale $\{e_0, \dots, e_N\}$. Pertanto $x - x_N \perp e_n$ per ogni $n = 0, \dots, N$, cioè

$$x - x_N \in \{e_0, \dots, e_N\}^\perp = (\operatorname{span}\{e_0, \dots, e_N\})^\perp \subseteq x_N^\perp$$

ove si sono usati la proposizione 6.26 e il fatto che $x_N \in \operatorname{span}\{e_0, \dots, e_N\}$. Di conseguenza, $x - x_N \perp x_N$, e quindi, per il teorema di Pitagora (☞ proposizione 6.23),

$$\|x\|^2 = \|x - x_N\|^2 + \|x_N\|^2 = \|x - x_N\|^2 + \sum_{n=0}^{N} |\langle x, e_n\rangle|^2, \tag{6.28}$$

dove per l'ultima uguaglianza si è nuovamente usato che $\{e_0, \dots, e_N\}$ è un insieme ortonormale. Siccome $\|x - x_N\|^2 \geq 0$, da (6.28) si ottiene la disugaglianza cercata.

(ii). Per ogni $N \in \mathbb{N}$, l'insieme $\{e_0, \dots, e_N\}$ è un insieme ortonormale. Pertanto, per (i), la disuguaglianza (6.26) vale per ogni $N \in \mathbb{N}$, e passando al limite per $N \to \infty$ si ottiene la disuguaglianza (6.27).

Siccome $\|x\|^2 < \infty$, la serie a primo membro di (6.27) converge, dunque la successione $(\langle x, e_n\rangle)_{n\in\mathbb{N}}$ è a quadrato sommabile, cioè appartiene a ℓ^2. □

Possiamo ora mostrare che, in uno spazio di Hilbert, è possibile considerare anche "combinazioni lineari infinite" di elementi di un insieme ortonormale, purché i coefficienti della combinazione formino una successione a quadrato sommabile.

Teorema 6.37. *Siano $(H, \langle\cdot,\cdot\rangle)$ uno spazio di Hilbert e $\{e_n\}_{n\in\mathbb{N}}$ un insieme ortonormale in H. Sia $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$. Allora la serie*

$$\sum_{n=0}^{\infty} \alpha_n e_n \tag{6.29}$$

converge in H se e solo se $\underline{\alpha} \in \ell^2$. Inoltre, in tal caso,

$$\left\| \sum_{n=0}^{\infty} \alpha_n e_n \right\| = \|\underline{\alpha}\|_2 \tag{6.30}$$

e

$$\left\langle \sum_{n=0}^{\infty} \alpha_n e_n, e_k \right\rangle = \alpha_k \qquad \text{per ogni } k \in \mathbb{N}. \tag{6.31}$$

Dimostrazione. Supponiamo anzitutto che la serie (6.29) converga in H, e sia $x = \sum_{n=0}^{\infty} \alpha_n e_n = \lim_{N\to\infty} \sum_{n=0}^{N} \alpha_n e_n$ la somma della serie. Allora, per ogni $k \in \mathbb{N}$, per la continuità del prodotto scalare,

$$\langle x, e_k \rangle = \lim_{N\to\infty} \left\langle \sum_{n=0}^{N} \alpha_n e_n, e_k \right\rangle = \lim_{N\to\infty} \sum_{n=0}^{N} \alpha_n \delta_{n,k} = \alpha_k,$$

il che dimostra (6.31). Dunque, dalla disuguaglianza di Bessel (☞ proposizione 6.36(ii)) deduciamo che $\underline{\alpha} = (\langle x, e_k \rangle)_k \in \ell^2$.

Viceversa, supponiamo invece che $\underline{\alpha} \in \ell^2$. Per ogni $n \in \mathbb{N}$, poniamo $x_n = \sum_{k=0}^{n} \alpha_k e_k$. Siccome H è completo, per dimostrare che la serie (6.29) converge in H, è sufficiente dimostrare che la successione $(x_n)_{n\in\mathbb{N}}$ delle somme parziali è di Cauchy. Ora, per ogni $n, m \in \mathbb{N}$, se $n \geq m$, si ha

$$\|x_n - x_m\|^2 = \left\| \sum_{k=m+1}^{n} \alpha_k e_k \right\|^2 = \sum_{k=m+1}^{n} |\alpha_k|^2 \leq \sum_{k=m+1}^{\infty} |\alpha_k|^2. \tag{6.32}$$

Siccome $\underline{\alpha} \in \ell^2$, la serie numerica $\sum_{k=0}^{\infty} |\alpha_k|^2$ converge, da cui deduciamo che $\lim_{m\to\infty} \sum_{k=m+1}^{\infty} |\alpha_k|^2 = 0$ (le code di una serie convergente sono infinitesime). Pertanto, per ogni $\epsilon > 0$, esiste $N \in \mathbb{N}$ tale che $\sum_{k=m+1}^{\infty} |\alpha_k|^2 < \epsilon^2$ per ogni $m > N$. Da (6.32) deduciamo allora che, per ogni $n \geq m > N$,

$$\|x_n - x_m\| < \epsilon;$$

scambiando il ruolo di n e m, si vede che la stessa disuguaglianza vale in realtà per ogni $n, m > N$. Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $(x_n)_n$ è di Cauchy.

Per la completezza di H, deduciamo allora che $\sum_{k=0}^{\infty} \alpha_k e_k = \lim_{n\to\infty} x_n = x$ per qualche $x \in H$. Ora, per ogni $n \in \mathbb{N}$, siccome $\{e_0, \dots, e_n\}$ è ortonormale, per la proposizione 6.23 si ha

$$\|x_n\|^2 = \sum_{k=0}^{n} |\alpha_k|^2$$

e passando al limite per $n \to \infty$, per la continuità della norma, si ottiene (6.30). □

Osservazione 6.38. Sotto le ipotesi del teorema 6.37, non è detto che la serie (6.29) converga assolutamente (☞ definizione 4.18). Infatti si ha

$$\sum_{n=0}^{\infty} \|\alpha_n e_n\| = \sum_{n=0}^{\infty} |\alpha_n| = \|\underline{\alpha}\|_1,$$

pertanto la convergenza assoluta di (6.29) in H avviene se e solo se $\underline{\alpha} \in \ell^1$. Data l'inclusione propria $\ell^1 \subsetneq \ell^2$ (☞ proposizione 5.17(ii)), l'ipotesi $\underline{\alpha} \in \ell^2$ del teorema 6.37 è dunque più debole di quanto servirebbe per garantire la convergenza assoluta. Sotto l'ipotesi più forte $\underline{\alpha} \in \ell^1$, la convergenza della serie (6.29) si potrebbe direttamente dedurre dalla proposizione 4.20. Il fatto che la serie (6.29) converga sotto l'ipotesi più debole $\underline{\alpha} \in \ell^2$ è dunque un risultato più raffinato, che utilizza in maniera fondamentale il fatto che $\{e_n\}_{n\in\mathbb{N}}$ è un sistema ortonormale.

I risultati precedenti ci permettono di trovare una formula per la proiezione ortogonale su un sottospazio chiuso.

Proposizione 6.39. *Sia $(H, \langle\cdot,\cdot\rangle)$ uno spazio di Hilbert. Sia Y un sottospazio vettoriale chiuso di H.*

(i) Supponiamo che $Y = \operatorname{span}\{e_0, \dots, e_N\}$, dove $\{e_0, \dots, e_N\}$ è un insieme ortonormale, per qualche $N \in \mathbb{N}$. Allora, per ogni $x \in H$,

$$P_Y x = \sum_{n=0}^{N} \langle x, e_n\rangle e_n.$$

(ii) Supponiamo che $Y = \overline{\operatorname{span}\{e_n\}_{n\in\mathbb{N}}}$, dove $\{e_n\}_{n\in\mathbb{N}}$ è un insieme ortonormale. Allora, per ogni $x \in H$,

$$P_Y x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n, \tag{6.33}$$

dove la serie converge in H.

Dimostrazione. Discutiamo solo la dimostrazione di (ii); la dimostrazione di (i) è analoga e più semplice.

Siccome $\{e_n\}_{n\in\mathbb{N}}$ è ortonormale, per la disuguaglianza di Bessel (☞ proposizione 6.36(ii)) sappiamo che $(\langle x, e_n\rangle)_{n\in\mathbb{N}} \in \ell^2$, pertanto per il teorema 6.37 la serie nel membro destro di (6.33) converge in H. Sia $y \in H$ la somma della serie; vogliamo dimostrare che $y = P_Y x$.

Notiamo anzitutto che $y = \lim_{N\to\infty} \sum_{n=0}^{N} \langle x, e_n\rangle e_n$ e che $\sum_{n=0}^{N} \langle x, e_n\rangle e_n \in \operatorname{span}\{e_n\}_{n\in\mathbb{N}}$ per ogni $N \in \mathbb{N}$, pertanto $y \in \overline{\operatorname{span}\{e_n\}_{n\in\mathbb{N}}} = Y$ (☞ proposizione 2.23(vii)). Per l'unicità della decomposizione ortogonale (☞ teorema 6.32), per concludere che $y = P_Y x$ è sufficiente dunque dimostrare che $x - y \in Y^\perp$.

D'altra parte, per ogni $k \in \mathbb{N}$, da (6.31) segue che $\langle y, e_k\rangle = \langle x, e_k\rangle$, quindi

$$\langle x - y, e_k\rangle = \langle x, e_k\rangle - \langle y, e_k\rangle = 0.$$

Questo mostra che $x - y \in \{e_k\}_{k\in\mathbb{N}}^\perp = \overline{\operatorname{span}\{e_k\}_{k\in\mathbb{N}}}^\perp = Y^\perp$ (☞ proposizione 6.26(iv)), come desiderato. □

Definizione 6.40. Sia H uno spazio di Hilbert.

(a) Un insieme ortonormale $E \subseteq H$ si dice *completo* se $H = \overline{\operatorname{span} E}$.

(b) Diciamo *base ortonormale* di H un insieme ortonormale completo al più numerabile.

Osservazione 6.41. Con la terminologia della definizione 6.40, la proposizione 6.39 ci dà dunque una formula per la proiezione ortogonale P_Y su un sottospazio chiuso Y dello spazio di Hilbert, in termini di una base ortonormale di Y.

Osservazione 6.42. Sia H uno spazio di Hilbert. Se $E \subseteq H$ è un insieme ortonormale finito, allora $\operatorname{span} E$ ha dimensione finita e quindi $\overline{\operatorname{span} E} = \operatorname{span} E$ (☞ proposizione 4.16(ii)). In particolare, se E è una base ortonormale finita di H, allora $H = \operatorname{span} E$, e quindi E è anche una base di H nel senso della definizione 1.9 (sappiamo già che ogni insieme ortonormale è linearmente indipendente, ☞

proposizione 6.23). Tuttavia, come mostra il prossimo esempio, la situazione è diversa quando E è infinito; dunque, in generale, una base ortonormale nel senso della definizione 6.40 non è necessariamente una base nel senso della definizione 1.9, e di caso in caso occorre fare attenzione al significato della parola "base".

Esempio 6.43. Sia $H = \ell^2$. L'insieme ortonormale $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ costruito nell'esempio 6.34(1) è una base ortonormale di ℓ^2, detta *base ortonormale canonica* di ℓ^2. Infatti $\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ è numerabile, e inoltre

$$\operatorname{span}\{\underline{e^{(n)}}\}_{n\in\mathbb{N}} = c_{00}$$

(☞ osservazione 6.35), dunque

$$\overline{\operatorname{span}\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}}^{\ell^2} = \overline{c_{00}}^{\ell^2} = \ell^2$$

(☞ proposizione 5.17(iii)).

Vediamo ora alcune proprietà fondamentali delle basi ortonormali.

Proposizione 6.44. *Sia H uno spazio di Hilbert. Sia $\{e_n\}_{n\in\mathbb{N}}$ un insieme ortonormale numerabile. Sono fatti equivalenti:*

(i) $\{e_n\}_{n\in\mathbb{N}}$ *è completo.*

(ii) $\{e_n\}_{n\in\mathbb{N}}^{\perp} = \{0\}$.

(iii) $x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n$ *per ogni* $x \in H$*, dove la serie converge in* H.

(iv) $\|x\|^2 = \sum_{n=0}^{\infty} |\langle x, e_n\rangle|^2$ *per ogni* $x \in H$.

(v) $\langle x, y\rangle = \sum_{n=0}^{\infty} \langle x, e_n\rangle \overline{\langle y, e_n\rangle}$ *per ogni* $x, y \in H$.

Un risultato analogo vale per un insieme ortonormale $\{e_n\}_{n=0}^{N}$ *finito, qualunque sia* $N \in \mathbb{N}$.

Dimostrazione. Consideriamo solo il caso di un insieme ortonormale numerabile; il caso finito è analogo e più semplice.

(i) ⇒ (iii). Da (i) sappiamo che $H = \overline{\operatorname{span}\{e_n\}_{n\in\mathbb{N}}}$ (☞ definizione 6.40), pertanto dalla proposizione 6.39 applicata con $Y = H$ segue che

$$x = P_H x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n,$$

ove la serie converge in H.

(iii) ⇒ (v). Per ogni $x \in H$, da (iii) sappiamo che $x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n = \lim_{N\to\infty} \sum_{n=0}^{N} \langle x, e_n\rangle e_n$ con convergenza in H. Usando la continuità e la linearità del prodotto scalare nel primo argomento, deduciamo allora che, per ogni $y \in H$,

$$\begin{aligned} \langle x, y\rangle = \left\langle \lim_{N\to\infty} \sum_{n=0}^{N} \langle x, e_n\rangle e_n, y \right\rangle &= \lim_{N\to\infty} \left\langle \sum_{n=0}^{N} \langle x, e_n\rangle e_n, y \right\rangle \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \langle x, e_n\rangle \langle e_n, y\rangle = \sum_{n=0}^{\infty} \langle x, e_n\rangle \overline{\langle y, e_n\rangle}, \end{aligned}$$

come desiderato; nell'ultimo passaggio si è usata la simmetria/hermitianità del prodotto scalare.

(v) $\Rightarrow$ (iv). Si applichi (v) con $y = x$.

(iv) $\Rightarrow$ (ii). Sia $x \in \{e_n\}_{n\in\mathbb{N}}^{\perp}$. Allora $\langle x, e_n\rangle = 0$ per ogni $n \in \mathbb{N}$; dunque, applicando (iv),

$$\|x\|^2 = \sum_{n=0}^{\infty} |\langle x, e_n\rangle|^2 = 0$$

e quindi $x = 0$.

(ii) $\Rightarrow$ (i). Da $\{e_n\}_{n\in\mathbb{N}}^{\perp} = \{0\}$ e dal corollario 6.33(vi) segue che

$$\overline{\operatorname{span}\{e_n\}_{n\in\mathbb{N}}} = \{e_n\}_{n\in\mathbb{N}}^{\perp\perp} = \{0\}^{\perp} = H$$

(☞ proposizione 6.26(v)), pertanto $\{e_n\}_{n\in\mathbb{N}}$ è completo (☞ definizione 6.40). □

Osservazione 6.45. Notiamo che la somma della serie $\sum_{n=0}^{\infty} \langle x, e_n\rangle e_n$ nella proposizione 6.44(iii) non dipende dall'ordine degli addendi. Infatti, se $\sigma : \mathbb{N} \to \mathbb{N}$ è biiettiva, possiamo applicare la proposizione 6.44 all'insieme ortonormale $\{e_{\sigma(n)}\}_{n\in\mathbb{N}}$ (cioè lo stesso insieme ortonormale indicizzato in un ordine diverso), ottenendo che

$$x = \sum_{n=0}^{\infty} \langle x, e_{\sigma(n)}\rangle e_{\sigma(n)};$$

dunque la somma della serie è sempre x, indipendentemente dalla biiezione σ. Questa proprietà di invarianza della somma rispetto a permutazioni degli addendi (detta *convergenza incondizionata* della serie) è vera nonostante in generale la serie non converga assolutamente (☞ osservazione 6.38), e quindi non si possa applicare la proposizione 4.21.

Teorema 6.46. *Uno spazio di Hilbert H ha una base ortonormale (nel senso della definizione 6.40) se e solo se H è separabile.*

Dimostrazione. Se H ha una base ortonormale B (al più numerabile), si ha $H = \overline{\operatorname{span} B}$ (☞ definizione 6.40), e la separabilità di H segue dalla proposizione 4.29.

Viceversa, supponiamo che H sia separabile. Sia $A \subseteq H$ denso e numerabile. Possiamo allora trovare un sottoinsieme $B \subseteq A$ linearmente indipendente tale che $\operatorname{span} B = \operatorname{span} A$ (☞ teorema 1.13); in particolare, B è finito o numerabile, cioè $B = \{x_n\}_{n=0}^{N}$ per $N \in \mathbb{N} \cup \{\infty\}$. Applicando induttivamente l'algoritmo di Gram–Schmidt a B (☞ teorema 6.24), si costruisce un insieme ortonormale $\{e_n\}_{n=0}^{N}$ tale che

$$\operatorname{span}\{e_n\}_{n=0}^{N} = \operatorname{span}\{x_n\}_{n=0}^{N} = \operatorname{span} A,$$

e quindi, passando alle chiusure,

$$\overline{\operatorname{span}\{e_n\}_{n=0}^{N}} = \overline{\operatorname{span} A} = H,$$

siccome A è denso. Questo dimostra che l'insieme ortonormale $\{e_n\}_{n=0}^{N}$ è completo (☞ definizione 6.40). □

Definizione 6.47. Due spazi pre-hilbertiani $(H_1, \langle\cdot,\cdot\rangle_{H_1})$ e $(H_2, \langle\cdot,\cdot\rangle_{H_2})$ si dicono *isometricamente isomorfi* se esiste un'applicazione lineare invertibile $\Phi : H_1 \to H_2$ tale che
$$\langle \Phi(x), \Phi(y)\rangle_{H_2} = \langle x, y\rangle_{H_1}$$
per ogni $x, y \in H_1$.

Corollario 6.48. *Sia H uno spazio di Hilbert separabile.*

(i) Se $\dim H = \infty$, allora H è isometricamente isomorfo a $(\ell^2, \langle\cdot,\cdot\rangle)$.

(ii) Se $\dim H = n \in \mathbb{N}_+$, allora H è isometricamente isomorfo a $(\mathbb{F}^n, \langle\cdot,\cdot\rangle)$.

Dimostrazione. Discutiamo solo il caso $\dim H = \infty$; l'altro caso è analogo e più semplice.

Sia $\{e_n\}_{n\in\mathbb{N}}$ una base ortonormale di H (☞ teorema 6.46). Definiamo allora $\Xi : H \to \ell^2$ ponendo $\Xi(x) = (\langle x, e_n\rangle)_{n\in\mathbb{N}}$. Dalla proposizione 6.44(iv) si ha
$$\|\Xi(x)\|_2^2 = \|x\|^2 < \infty \tag{6.34}$$
per ogni $x \in H$, pertanto $\Xi(x) \in \ell^2$ per ogni $x \in H$, dunque Ξ è ben definita. Inoltre, per la linearità del prodotto scalare nella prima variabile, si ha
$$\Xi(\alpha x + \beta y) = (\langle \alpha x + \beta y, e_n\rangle)_{n\in\mathbb{N}} = (\alpha\langle x, e_n\rangle + \beta\langle y, e_n\rangle)_{n\in\mathbb{N}} = \alpha\Xi(x) + \beta\Xi(y)$$
per ogni $x, y \in H$ e $\alpha, \beta \in \mathbb{F}$, dunque $\Xi : H \to \ell^2$ è lineare.

Dall'identità (6.34) segue anche che, per ogni $x \in H$, se $\Xi(x) = 0$ allora $x = 0$; in altre parole, $\operatorname{Ker}\Xi = \{0\}$ e quindi Ξ è iniettiva. Inoltre, se $\underline{\alpha} \in \ell^2$, per il teorema 6.37 possiamo definire $x = \sum_{n=0}^\infty \alpha_n e_n \in H$, e si ha $\langle x, e_n\rangle = \alpha_n$ per ogni $n \in \mathbb{N}$, cioè $\Xi(x) = \underline{\alpha}$; questo dimostra che Ξ è suriettiva. Quindi $\Xi : H \to \ell^2$ è biiettiva e invertibile.

Infine, dalla proposizione 6.44(v) si ha
$$\langle \Xi(x), \Xi(y)\rangle = \sum_{n=0}^\infty \langle x, e_n\rangle \overline{\langle y, e_n\rangle} = \langle x, y\rangle$$
per ogni $x, y \in H$. Ne concludiamo (☞ definizione 6.47) che H e ℓ^2 sono isometricamente isomorfi. □

Osservazione 6.49. Il corollario 6.48 mostra che, a meno di isomorfismo isometrico, gli spazi di Hilbert separabili sono interamente classificati dalla loro dimensione, nel senso che tutti gli spazi di Hilbert separabili di dimensione infinita sono isomorfi a ℓ^2 e tutti gli spazi di Hilbert di dimensione n sono isomorfi a $\mathbb{F}^n$. In altre parole, a meno di isomorfismo isometrico, c'è un solo spazio di Hilbert per ogni dimensione. In questo senso, la teoria degli spazi di Hilbert è molto più "rigida" di quella degli spazi di Banach, dove a parità di dimensione si trovano spazi non isomorfi.

Osservazione 6.50. Con piccole variazioni, la discussione sopra si potrebbe estendere in modo da trattare anche il caso di spazi di Hilbert non separabili. In effetti, è possibile mostrare che *ogni* spazio di Hilbert ha un insieme ortonormale completo $\{e_i\}_{i\in I}$, dove l'insieme degli indici I può essere più che numerabile; inoltre, per tale insieme ortonormale completo, valgono formule analoghe a quelle della proposizione 6.44, ove però occorre dare un opportuno significato a somme della forma $\sum_{i \in I} \langle x, e_i \rangle e_i$ con una quantità più che numerabile di addendi. La corrispondente generalizzazione del corollario 6.48 permette allora di classificare, a meno di isomorfismo isometrico, gli spazi di Hilbert in base alla cardinalità di un loro insieme ortonormale completo.

6.4 Serie di Fourier

Sappiamo dal corollario 5.39 che, per ogni intervallo limitato $[a, b]$ di $\mathbb{R}$, lo spazio di Hilbert $L^2(a, b)$ è separabile. Il teorema 6.46 ci garantisce dunque, in astratto, l'esistenza di basi ortonormali di $L^2(a, b)$. Tuttavia, nella pratica, per poter utilizzare questo strumento, è utile avere a disposizione esempi concreti di tali basi ortonormali di $L^2(a, b)$ — un po' come, nel caso di ℓ^2, abbiamo a disposizione la base ortonormale canonica $\{\underline{e}^{(n)}\}_{n \in \mathbb{N}}$.

In questa sezione vogliamo brevemente discutere di alcune basi ortonormali comunemente usate per gli spazi $L^2(a, b)$. Un esempio di insieme ortonormale su $L^2(-\pi, \pi)$, nel caso $\mathbb{F} = \mathbb{C}$, è stato discusso nell'esempio 6.34(2); uno dei risultati qui presentati è proprio il fatto che tale sistema ortonormale è completo.

Per brevità, nel seguente enunciato scriviamo $\cos(nt)$ per denotare la funzione $t \mapsto \cos(nt)$, e analogo significato hanno le notazioni $\sin(nt)$ e e^{int}.

Proposizione 6.51. *Valgono le seguenti proprietà.*

(i) *L'insieme* $B_0 = \{\sqrt{1/\pi}\} \cup \{\sqrt{2/\pi} \cos(nt)\}_{n=1}^{\infty}$ *è una base ortonormale di* $L^2(0, \pi)$.

(ii) *L'insieme* $B_1 = \{\sqrt{2/\pi} \sin(nt)\}_{n=1}^{\infty}$ *è una base ortonormale di* $L^2(0, \pi)$.

(iii) *L'insieme* $B = \{\sqrt{1/2\pi}\} \cup \{\sqrt{1/\pi} \cos(nt)\}_{n=1}^{\infty} \cup \{\sqrt{1/\pi} \sin(nt)\}_{n=1}^{\infty}$ *è una base ortonormale di* $L^2(-\pi, \pi)$.

(iv) *Nel caso* $\mathbb{F} = \mathbb{C}$, *l'insieme* $E = \{\sqrt{1/2\pi} e^{int}\}_{n \in \mathbb{Z}}$ *è una base ortonormale di* $L^2(-\pi, \pi)$.

Dimostrazione (cenno). (i). Usando le formule di Werner

$$\cos(nt)\cos(mt) = \frac{1}{2}\left[\cos((n+m)t) + \cos((n-m)t)\right] \qquad \forall n, m \in \mathbb{Z}\ \forall t \in \mathbb{R}, \tag{6.35}$$

insieme al fatto che

$$\int_0^\pi \cos(nt)\, dt = \begin{cases} 0 & \text{se } n \in \mathbb{Z} \setminus \{0\}, \\ \pi & \text{se } n = 0, \end{cases}$$

non è difficile verificare che B_0 è un insieme ortonormale in $L^2(0, \pi)$. Rimane da verificarne la completezza. Siccome sappiamo già che $\overline{C[0, \pi]}^{L^2(0,\pi)} = L^2(0, \pi)$ (☞ teorema 5.37(iv)), è sufficiente verificare che $C[0, \pi] \subseteq \overline{\operatorname{span} B_0}^{L^2(0,\pi)}$.

D'altra parte, $\arccos : [-1, 1] \to [0, \pi]$ è un omeomorfismo. Dunque, se $f \in C[0, \pi]$, allora $f \circ \arccos \in C[-1, 1]$. Per il teorema di Stone–Weierstrass (☞ teorema 4.30) esiste una successione $(p_j)_{j \in \mathbb{N}}$ di polinomi tali che $p_j|_{[-1,1]} \rightrightarrows f \circ \arccos$ per $j \to \infty$. Ma allora, se $\tilde{p}_j := p_j \circ \cos|_{[0,\pi]}$, si ha

$$\|\tilde{p}_j - f\|_{L^2(0,\pi)} \le \pi^{1/2} \|\tilde{p}_j - f\|_\infty = \pi^{1/2} \|p_j|_{[-1,1]} - f \circ \arccos\|_\infty \to 0$$

per $j \to \infty$, cioè $\tilde{p}_j \to f$ in $L^2(0,\pi)$. Rimane solo da verificare che le funzioni $\tilde{p}_j = p_j \circ \cos|_{[0,\pi]}$ sono in $\operatorname{span} B_0$; una volta verificato questo, possiamo allora concludere (☞ proposizione 2.23(vii)) che $f \in \overline{\operatorname{span} B_0}^{L^2(0,\pi)}$.

D'altra parte, per ogni polinomio $p \in \mathcal{P}$, si ha $p(x) = \sum_{k=0}^N \alpha_k x^k$ per qualche $N \in \mathbb{N}$ e $\alpha_0, \dots, \alpha_N \in \mathbb{F}$. Pertanto, se $\tilde{p} = p \circ \cos|_{[0,\pi]}$, si ha

$$\tilde{p}(t) = p(\cos t) = \sum_{k=0}^N \alpha_k (\cos t)^k \in \operatorname{span} B_0;$$

per giustificare l'ultimo passaggio, è sufficiente usare iterativamente le formule di Werner (6.35) per dimostrare, per induzione su $k \in \mathbb{N}$, che

$$(\cos t)^k = \sum_{l=0}^k \lambda_{k,l} \cos(lt)$$

per opportuni coefficienti $\lambda_{k,l} \in \mathbb{R}$.

(ii). Per la verifica dell'ortonormalità si può procedere in maniera simile a (i), utilizzando le formule di Werner

$$\sin(nt)\sin(mt) = \frac{1}{2}\left[\cos((n-m)t) - \cos((n+m)t)\right] \qquad \forall n,m \in \mathbb{Z}\ \forall t \in \mathbb{R}. \tag{6.36}$$

Per la completezza, siccome sappiamo che

$$\overline{C_c(0,\pi)}^{L^2(0,\pi)} = L^2(0,\pi)$$

(☞ teorema 5.37(iv)), è sufficiente verificare che $C_c(0,\pi) \subseteq \overline{\operatorname{span} B_1}^{L^2(0,\pi)}$.

Sia allora $f \in C_c(0,\pi)$. Allora possiamo definire $g \in C_c(0,\pi)$ ponendo $g(t) = f(t)/\sin(t)$ per $t \in (0,\pi)$. In particolare $g \in C[0,\pi]$ e quindi, come nella dimostrazione di (i), possiamo costruire una successione $(g_n)_{n\in\mathbb{N}}$ a valori in $\operatorname{span} B_0$ tale che $g_n \to g$ in $L^2(0,\pi)$. Se ora definiamo $f_n : [0,\infty] \to \mathbb{F}$ ponendo $f_n(t) = g_n(t)\sin(t)$ per $t \in [0,\pi]$, si ha anche

$$\begin{aligned}\|f - f_n\|_{L^2(0,\pi)}^2 &= \int_0^\pi |g(t) - g_n(t)|^2 \sin^2 t \, dt \\ &\le \int_0^\pi |g(t) - g_n(t)|^2 \, dt = \|g - g_n\|_{L^2(0,\pi)}^2 \to 0 \qquad \text{per } n \to \infty,\end{aligned}$$

pertanto anche $f_n \to f$ in $L^2(0,\pi)$. Per concludere, è sufficiente verificare che $f_n \in \operatorname{span} B_1$ per ogni $n \in \mathbb{N}$; infatti, in tal caso, potremmo dedurre (☞ proposizione 2.23(vii)) che $f \in \overline{\operatorname{span} B_1}^{L^2(0,\pi)}$.

In altre parole, è sufficiente verificare che, per ogni $h \in \operatorname{span} B_0$, si ha $h(t)\sin t \in \operatorname{span} B_1$. Scrivendo $h(t)$ come combinazione lineare di termini della forma $\cos(nt)$, il problema si riduce a verificare che $\cos(nt)\sin t \in \operatorname{span} B_1$ per ogni $n \in \mathbb{N}$. D'altra parte, quest'ultimo fatto è un'immediata conseguenza delle formule di Werner

$$\cos(nt)\sin(mt) = \frac{1}{2}\left[\sin((n+m)t) - \sin((n-m)t)\right] \qquad \forall n,m \in \mathbb{Z}\ \forall t \in \mathbb{R}. \tag{6.37}$$

(iii). Per la verifica dell'ortonormalità di B si procede come per (i) e (ii); in effetti, con considerazioni di parità, ci si può ridurre ai calcoli già effettuati per B_0 e B_1.

Per la completezza, osserviamo che ogni $f \in L^2(-\pi,\pi)$ si decompone in parte pari e parte dispari, cioè $f = f_p + f_d$, ove $f_p(t) = (f(t) + f(-t))/2$ e $f_d(t) = (f(t) - f(-t))/2$. Inoltre, $f_p \perp f_d$ (il prodotto $f_p \overline{f_d}$ è dispari e quindi ha integrale nullo su $(-\pi,\pi)$), pertanto, per il teorema di Pitagora,

$$\|f\|^2_{L^2(-\pi,\pi)} = \|f_p\|^2_{L^2(-\pi,\pi)} + \|f_d\|^2_{L^2(-\pi,\pi)} = 2\left[\|f_p\|^2_{L^2(0,\pi)} + \|f_d\|^2_{L^2(0,\pi)}\right]$$

(per l'ultimo passaggio si è usato che $|f_p|^2$ e $|f_d|^2$ sono entrambe pari e quindi l'integrale su $(-\pi,\pi)$ è il doppio di quello su $(0,\pi)$). In particolare si ha $f_p|_{[0,\pi]}, f_d|_{[0,\pi]} \in L^2(0,\pi)$.

Applicando (i) a $f_p|_{[0,\pi]}$ e (ii) a $f_d|_{[0,\pi]}$, possiamo costruire una successione $(g_n)_n$ a valori in $\operatorname{span} B_0$ e una successione $(h_n)_n$ a valori in $\operatorname{span} B_1$ tali che $g_n \to f_p|_{[0,\pi]}$ e $h_n \to f_d|_{[0,\pi]}$ in $L^2(0,\pi)$ per $n \to \infty$. Ora, siccome $g_n(t)$ è una combinazione lineare di funzioni della forma $\cos(kt)$, la stessa combinazione lineare ne dà un'estensione a una funzione pari $\tilde{g}_n : [-\pi,\pi] \to \mathbb{F}$; similmente, siccome $h_n(t)$ è una combinazione lineare di funzioni della forma $\sin(kt)$, la stessa combinazione lineare ne dà un'estensione a una funzione dispari $\tilde{h}_n : [-\pi,\pi] \to \mathbb{F}$. In particolare, $\tilde{g}_n$ e $\tilde{h}_n$ sono combinazioni lineari di $\cos(kt)$ e $\sin(kt)$, quindi $\tilde{g}_n, \tilde{h}_n \in \operatorname{span} B$. Inoltre

$$\begin{aligned} \|f_p - \tilde{g}_n\|^2_{L^2(-\pi,\pi)} &= 2\|f_p|_{[0,\pi]} - g_n\|^2_{L^2(0,\pi)} \to 0 \quad \text{per } n \to \infty, \\ \|f_d - \tilde{h}_n\|^2_{L^2(-\pi,\pi)} &= 2\|f_d|_{[0,\pi]} - h_n\|^2_{L^2(0,\pi)} \to 0 \quad \text{per } n \to \infty. \end{aligned}$$

Pertanto, se poniamo $f_n = \tilde{g}_n + \tilde{h}_n$, si ha anche $f_n \to f_p + f_d = f$ in $L^2(-\pi,\pi)$ per $n \to \infty$, come desiderato.

(iv). L'ortonormalità di E è già stata verificata nell'esempio 6.34(2); rimane da verificarne la completezza. Ora, dalle identità

$$\cos(nt) = \frac{e^{int} + e^{-int}}{2}, \qquad \sin(nt) = \frac{e^{int} - e^{-int}}{2i},$$

e

$$e^{int} = \cos(nt) + i\sin(nt) = \cos(|n|t) + i\epsilon_n \sin(|n|t),$$

ove $\epsilon_n = \operatorname{sign}(n)$, si vede facilmente che $\operatorname{span}_{\mathbb{C}} E = \operatorname{span}_{\mathbb{C}} B$ (ogni combinazione lineare di esponenziali si scrive come combinazione lineare di seni e coseni, e viceversa). Pertanto

$$\overline{\operatorname{span}_{\mathbb{C}} E}^{L^2(-\pi,\pi)} = \overline{\operatorname{span}_{\mathbb{C}} B}^{L^2(-\pi,\pi)} = L^2(-\pi,\pi)$$

per (iii). □

Osservazione 6.52. Applicando la proposizione 6.44 alla base ortonormale E della proposizione 6.51, si ottiene che ogni $f \in L^2(-\pi,\pi)$ si scrive nella forma

$$f(t) = \sum_{n=-\infty}^{\infty} \alpha_n e^{int} \tag{6.38}$$

ove la serie converge in $L^2(-\pi,\pi)$ e i coefficienti α_n sono dati da

$$\alpha_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt. \tag{6.39}$$

Si ha inoltre

$$\|f\|^2_{L^2(-\pi,\pi)} = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}|\alpha_n|^2. \tag{6.40}$$

La serie nel membro destro di (6.38) è detta *serie di Fourier* di f e i coefficienti in (6.39) sono i *coefficienti di Fourier* di f; inoltre (6.40) è detta *identità di Parseval*. In altre parole, la teoria degli spazi di Hilbert, applicata alla base ortonormale E, dimostra che ogni $f \in L^2(-\pi,\pi)$ si sviluppa in serie di Fourier (con convergenza in L^2 della serie). Considerazioni simili valgono per le altre basi ortonormali della proposizione 6.51: la base ortonormale B corrisponde allo svilippo in *serie di Fourier reale*, mentre B_0 e B_1 corrispondono agli sviluppi in *serie di coseni* e *serie di seni*.

Osservazione 6.53. Utilizzando opportunamente traslazioni e dilatazioni, a partire dalle basi ortonormali discusse nella proposizione 6.51 si possono costruire analoghe basi ortonormali di $L^2(a,b)$ per ogni intervallo limitato $[a,b]\subseteq\mathbb{R}$.

Vediamo ora come le costruzioni precedenti si possono combinare per costruire serie di Fourier multi-dimensionali.

Proposizione 6.54. *Siano $(M_1,\mathcal{M}_1,\mu_1)$ e $(M_2,\mathcal{M}_2,\mu_2)$ spazi di misura σ-finiti. Sia $(M,\mathcal{M},\mu)$ lo spazio di misura prodotto. Siano $\{\phi_n\}_{n\in\mathbb{N}}$ e $\{\psi_m\}_{m\in\mathbb{N}}$ basi ortonormali di $L^2(M_1)$ e $L^2(M_2)$ rispettivamente. Allora $\{\phi_n\otimes\psi_m\}_{(n,m)\in\mathbb{N}^2}$ è una base ortonormale di $L^2(M)$, dove $\phi_n\otimes\psi_m(t,s) \coloneqq \phi_n(t)\psi_m(s)$ è il* prodotto tensore *di ϕ_m e ψ_n.*

Dimostrazione (cenno). Utilizzando il teorema di Fubini–Tonelli, si vede che

$$\begin{aligned}\langle \phi_n\otimes\psi_m, \phi_{n'}\otimes\psi_{m'}\rangle_{L^2(M)} &= \int_{M_2}\int_{M_1}\phi_n(t)\psi_m(s)\overline{\phi_{n'}(t)\psi_{m'}(s)}\,d\mu_1(t)\,d\mu_2(s)\\ &= \langle\phi_n,\phi_{n'}\rangle_{L^2(M_1)}\langle\psi_m,\psi_{m'}\rangle_{L^2(M_2)} = \delta_{n,n'}\delta_{m,m'},\end{aligned}$$

da cui segue che $\{\phi_n\otimes\psi_m\}_{(n,m)\in\mathbb{N}^2}$ è un insieme ortonormale in $L^2(M)$.

Per mostrare che è completo, per la proposizione 6.44(ii) è sufficiente dimostrare che $\{\phi_n\otimes\psi_m\}^{\perp}_{(n,m)\in\mathbb{N}^2} = \{0\}$ in $L^2(M)$. Sia dunque $f\in L^2(M)$ tale che $f\perp\phi_n\otimes\psi_m$ per ogni $n,m\in\mathbb{N}$. Allora, per ogni $m,n\in\mathbb{N}$, per il teorema di Fubini–Tonelli,

$$\begin{aligned}0 &= \int_{M_2}\int_{M_1} f(t,s)\,\overline{\phi_n(t)\,\psi_m(s)}\,d\mu_1(t)\,d\mu_2(s)\\ &= \int_{M_2}\left(\int_{M_1} f(t,s)\,\overline{\phi_n(t)}\,d\mu_1(t)\right)\overline{\psi_m(s)}\,d\mu_2(s).\end{aligned} \tag{6.41}$$

Si noti che, con lo stesso argomento utilizzato per dimostrare la buona definizione degli operatori integrali (☞ esempio 7.8(6)), dato che $f\in L^2(M)$ e $\psi_m\in L^2(M_1)$, si ha anche $\int_{M_1} f(t,\cdot)\,\overline{\phi_n(t)}\,d\mu_1(t)\in L^2(M_2)$.

Fissato $n \in \mathbb{N}$, l'identità (6.41) dice che la funzione $\int_{M_1} f(t,\cdot)\,\overline{\phi_n(t)}\,d\mu_1(t) \in L^2(M_2)$ è ortogonale a ψ_m in $L^2(M_2)$ per ogni $m \in \mathbb{N}$. Per la completezza di $\{\psi_m\}_m$ in $L^2(M_2)$, questo implica che

$$\int_{M_1} f(t,s)\,\overline{\phi_n(t)}\,d\mu_1(t) = 0 \tag{6.42}$$

per μ_2-quasi ogni $s \in M_2$. In altre parole, se N_n è l'insieme degli $s \in M_2$ per cui vale (6.42), allora N_n ha misura piena in M_2 (cioè $M_2 \setminus N_n$ è trascurabile). Se $N = \bigcap_{n\in\mathbb{N}} N_n$, allora anche N è un insieme di misura piena (il complementare è un'unione numerabile di insiemi trascurabili, dunque è trascurabile).

Inoltre, per ogni $s \in N$, l'identità (6.42) vale per ogni $n \in \mathbb{N}$; in altre parole, per ogni $s \in N$, la funzione $f(\cdot,s) \in L^2(M_1)$ è ortogonale a ϕ_n in $L^2(M_1)$ per ogni $n \in \mathbb{N}$; per la completezza di $\{\phi_n\}_n$ in $L^2(M_1)$, ne concludiamo che $f(t,s) = 0$ per μ_1-quasi ogni $t \in M_1$ e ogni $s \in N$. Siccome N ha misura piena in M_2, questo significa che $f = 0$ μ-quasi ovunque. □

Esempio 6.55. Utilizzando la proposizione 6.54, possiamo costruire basi ortonormali di $L^2((a,b)\times(c,d))$ prendendo il "prodotto tensore" di basi ortonormali di $L^2(a,b)$ e $L^2(c,d)$. Ad esempio, nel caso $(a,b) = (c,d) = (-\pi,\pi)$ e $\mathbb{F} = \mathbb{C}$, sappiamo (☞ proposizione 6.51(iv)) che $\{t \mapsto e^{int}\}_{n\in\mathbb{Z}}$ è una base ortonormale di $L^2(-\pi,\pi)$. Dalla proposizione 6.54 segue allora che $\{(t,s) \mapsto e^{i(nt+ms)}\}_{n,m\in\mathbb{Z}}$ è una base ortonormale di $L^2((-\pi,\pi)^2)$. Lo sviluppo di una funzione in $L^2((-\pi,\pi)^2)$ rispetto a questa base corrisponde a uno "sviluppo in serie di Fourier in due variabili". Chiaramente la costruzione si può iterare, ottenendo basi ortonormali per spazi di funzioni di più variabili.

7 Operatori limitati

7.1 Norma operatoriale e operatori limitati

Ricordiamo che $\mathbb{F}$ denota il campo $\mathbb{R}$ dei numeri reali oppure il campo $\mathbb{C}$ dei numeri complessi.

Definizione 7.1. Siano X e Y spazi vettoriali su $\mathbb{F}$. Un *operatore lineare* da X a Y è un'applicazione lineare $T : X \to Y$ (☞ definizione 1.19). Se $X = Y$, chiamiamo T un operatore lineare *su* X.

Osservazione 7.2. Come già discusso (☞ definizione 1.19 e proposizione 1.20), l'insieme $\mathcal{L}(X, Y)$ degli operatori lineari da X a Y è un sottospazio vettoriale di $\mathcal{F}(X, Y)$; in altre parole, $\mathcal{L}(X, Y)$ è uno spazio vettoriale con le operazioni puntuali. Inoltre, quando $X = Y$, scriviamo $\mathcal{L}(X)$ anziché $\mathcal{L}(X, X)$.

Definizione 7.3. Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ spazi normati su $\mathbb{F}$.

(a) Se $T \in \mathcal{L}(X, Y)$, definiamo la *norma operatoriale* $\|T\|_{\mathrm{op}}$ di T ponendo

$$\|T\|_{\mathrm{op}} = \inf\{C \in [0, \infty) : \|Tx\|_Y \leq C\|x\|_X \;\forall x \in X\}.$$

Scriviamo anche $\|T\|_{X \to Y}$ anziché $\|T\|_{\mathrm{op}}$, quando sia utile esplicitare gli spazi normati di partenza e arrivo di T.

(b) Un operatore $T \in \mathcal{L}(X, Y)$ si dice *limitato* se $\|T\|_{\mathrm{op}} < \infty$. In altre parole, $T \in \mathcal{L}(X, Y)$ è limitato se esiste $C \in [0, \infty)$ tale che $\|Tx\|_Y \leq C\|x\|_X$ per ogni $x \in X$. Se $T \in \mathcal{L}(X, Y)$ non è limitato, si dice *illimitato*.

(c) L'insieme degli operatori lineari limitati da X a Y si denota con $\mathcal{B}(X, Y)$. Scriviamo anche $\mathcal{B}(X)$ al posto di $\mathcal{B}(X, X)$.

Osservazione 7.4. Pur avendo chiamato $\|\cdot\|_{\mathrm{op}}$ "norma operatoriale", questo non significa automaticamente che $\|\cdot\|_{\mathrm{op}}$ sia una norma nel senso della definizione 4.1. In effetti, in generale $\|\cdot\|_{\mathrm{op}}$ su $\mathcal{L}(X, Y)$ può assumere il valore ∞ (questo succede quando esistono operatori illimitati da X a Y), e in tal caso $\|\cdot\|_{\mathrm{op}}$ non è una norma su $\mathcal{L}(X, Y)$. Vedremo a breve (☞ proposizione 7.10) che $\mathcal{B}(X, Y)$ è un sottospazio vettoriale di $\mathcal{L}(X, Y)$ e che $\|\cdot\|_{\mathrm{op}}$ ristretta a $\mathcal{B}(X, Y)$ è effettivamente una norma su $\mathcal{B}(X, Y)$. Questo lieve abuso di linguaggio nell'uso del termine "norma" è analogo a quello già discusso per le "norme p-esime" $\|\cdot\|_p$ su $\mathbb{F}^{\mathbb{N}}$ (☞ osservazioni 5.3 e 5.9).

Ci sono numerose caratterizzazioni equivalenti della norma operatoriale; il seguente enunciato ne discute alcune.

Lemma 7.5. *Siano X, Y spazi normati e $T \in \mathcal{L}(X, Y)$. Allora*

$$\|T\|_{\mathrm{op}} = \sup\{\|Tx\|_Y : x \in X,\ \|x\|_X \leq 1\}. \tag{7.1}$$

Se poi $X \neq \{0\}$, si ha anche

$$\begin{aligned}\|T\|_{\mathrm{op}} &= \sup\{\|Tx\|_Y : x \in X, \|x\|_X = 1\} \\ &= \sup\left\{\frac{\|Tx\|_Y}{\|x\|_X} : x \in X \setminus \{0\}\right\}.\end{aligned}$$

Dimostrazione. Siano

$$\begin{aligned}
A &= \{C \in [0,\infty) : \|Tx\|_Y \leq C\|x\|_X \ \forall x \in X\},\\
B_1 &= \{\|Tx\|_Y : x \in X, \|x\|_X \leq 1\},\\
B_2 &= \{\|Tx\|_Y : x \in X, \|x\|_X = 1\},\\
B_3 &= \left\{ \frac{\|Tx\|_Y}{\|x\|_X} : x \in X \setminus \{0\} \right\}.
\end{aligned}$$

Notiamo che, se $X = \{0\}$, allora anche $T = 0$ e quindi $A = [0,\infty)$, $B_1 = \{0\}$ e $\sup B_1 = 0 = \inf A$. Questo dimostra (7.1) e conclude la dimostrazione quando $X = \{0\}$.

Supponiamo invece che $X \neq \{0\}$. Vediamo ora che

$$B_1 \supseteq B_2 = B_3. \tag{7.2}$$

Le inclusioni $B_1 \supseteq B_2$ e $B_2 \subseteq B_3$ sono ovvie. Per l'inclusione $B_3 \subseteq B_2$, basta osservare che $\|Tx\|_Y/\|x\|_x$ si può scrivere nella forma $\|T\tilde{x}\|_Y$ con $\tilde{x} = x/\|x\|_X$, e che $\|\tilde{x}\|_X = 1$ se $x \neq 0$.

Dalle relazioni (7.2) si ha

$$\sup B_1 \geq \sup B_2 = \sup B_3 \geq 0;$$

l'ultima disuguaglianza segue dal fatto che l'insieme $B_2 = B_3$ non è vuoto (siccome $X \neq \{0\}$) e per definizione i suoi elementi sono nonnegativi.

Dimostriamo ora la disuguaglianza $\sup B_1 \leq \sup B_2$. Sia $x \in X$ con $\|x\|_X \leq 1$. Se $x = 0$, allora $\|Tx\|_Y = 0 \leq \sup B_2$. Se invece $x \neq 0$, allora possiamo scrivere $x = \|x\|_X \tilde{x}$, ove $\tilde{x} = x/\|x\|_X$ ha norma 1; pertanto

$$\|Tx\|_Y = \|T(\|x\|_X \tilde{x})\|_Y = \|x\|_X \|T\tilde{x}\|_Y \leq \|T\tilde{x}\|_Y \in B_2$$

perché $\|x\|_X \leq 1$ e $\|\tilde{x}\|_X = 1$; in particolare, ne concludiamo di nuovo che $\|Tx\|_Y \leq \sup B_2$. Abbiamo quindi dimostrato che ogni elemento di B_1 è minore o uguale di B_2, e quindi $\sup B_1 \leq \sup B_2$.

In conclusione,

$$\sup B_1 = \sup B_2 = \sup B_3,$$

e rimane solo da dimostrare che questi sup sono uguali a $\|T\|_{\mathrm{op}} = \inf A$.

Notiamo ora che A è l'insieme dei maggioranti di B_3. Infatti, se $C \in A$, allora è evidente dalle definizioni che C maggiora ogni elemento di B_3. Viceversa, se C è un maggiorante di B_3, allora $C \geq \|Tx\|_Y/\|x\|_X$ per ogni $x \in X \setminus \{0\}$, e quindi $\|Tx\|_Y \leq C\|x\|_X$ per ogni $x \in X \setminus \{0\}$; quest'ultima disuguaglianza tuttavia vale banalmente anche per $x = 0$ (qui usiamo che $C \geq 0$, perché $X \neq \{0\}$ e quindi $B_3 \neq \emptyset$), il che dimostra che $C \in A$.

Per definizione, quando B_3 è superiormente limitato, $\sup B_3$ è il minimo dei maggioranti di B_3, pertanto $\sup B_3 = \min A = \inf A$ in tal caso. Se invece B_3 è superiormente illimitato, allora $A = \emptyset$ e quindi di nuovo $\inf A = \infty = \sup B_3$. □

Vediamo ora che ci sono diverse caratterizzazioni equivalenti degli operatori limitati.

Proposizione 7.6. *Siano X, Y spazi normati e $T \in \mathcal{L}(X,Y)$. Sono fatti equivalenti:*

(i) T *è un operatore limitato;*

(ii) T *è lipschitziano;*

(iii) T *è uniformemente continuo;*

(iv) T *è continuo;*

(v) T *è continuo in* 0.

Inoltre, se T *è limitato, allora* T *è* $\|T\|_{\mathrm{op}}$*-lipschitziano e*

$$\|Tx\|_Y \leq \|T\|_{\mathrm{op}} \|x\|_X \tag{7.3}$$

per ogni $x \in X$.

Dimostrazione. Nel caso $X = \{0\}$, si ha $T = 0$ e $\|T\|_{\mathrm{op}} = 0$, quindi tutte le proprietà enunciate sono banalmente vere. Possiamo dunque supporre $X \neq \{0\}$.

(i) $\Rightarrow$ (ii). Dal lemma 7.5 sappiamo che

$$\frac{\|Tx\|_Y}{\|x\|_X} \leq \|T\|_{\mathrm{op}} < \infty$$

per ogni $x \in X \setminus \{0\}$, da cui segue (7.3); quest'ultima disuguaglianza vale banalmente anche per $x = 0$. In particolare, siccome T è lineare,

$$\|Tx - Tx'\|_Y \leq \|T(x - x')\|_Y \leq \|T\|_{\mathrm{op}} \|x - x'\|_X$$

per ogni $x, x' \in X$; da questo segue che T è $\|T\|_{\mathrm{op}}$-lipschitziano (☞ definizione 2.33).

(ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v). Queste implicazioni sono ovvie, in quanto valgono per qualunque funzione fra spazi metrici (☞ sezione 2.5).

(v) $\Rightarrow$ (i). Siccome $T : X \to Y$ continuo in 0 e $T0 = 0$, preso $\epsilon = 1$ nella definizione 2.27(a), troviamo $\delta > 0$ tale che, per ogni $x \in X$,

$$\|Tx\|_Y < 1 \qquad \text{ogniqualvolta } \|x\|_X < \delta. \tag{7.4}$$

Ora, se $x \in X$ e $\|x\|_X \leq 1$, allora $\|(\delta/2)x\|_X \leq \delta/2 < \delta$, e quindi, applicando (7.4) con $(\delta/2)x$ al posto di x,

$$(\delta/2)\|Tx\|_Y = \|T((\delta/2)x)\|_Y < 1,$$

cioè

$$\|Tx\|_Y < 2/\delta$$

per ogni $x \in X$ con $\|x\|_X \leq 1$. Passando all'estremo superiore in x, per il lemma 7.5 otteniamo

$$\|T\|_{\mathrm{op}} \leq 2/\delta < \infty,$$

dunque T è limitato. □

Osservazione 7.7. Come discusso nella dimostrazione della proposizione 7.6, le implicazioni (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) valgono per qualunque funzione fra spazi metrici; tuttavia, per funzioni arbitrarie, tali implicazioni non si possono invertire (☞ esempio 2.35). L'aspetto interessante del precedente risultato è che tali implicazioni si possono invertire sotto l'ipotesi che T sia una mappa lineare.

Esempio 7.8. Ecco alcuni esempi e non-esempi di operatori limitati.

(1) Se $X \neq \{0\}$ è uno spazio normato, allora l'*operatore identità* $\mathrm{id}_X : X \to X$, definito da $\mathrm{id}_X x = x$ per ogni $x \in X$, è lineare e soddisfa $\|\mathrm{id}_X x\|_X = \|x\|_X$ per ogni $x \in X$, pertanto $\|\mathrm{id}_X\|_{\mathrm{op}} = 1$ e $\mathrm{id}_X \in \mathcal{B}(X)$. Se invece $X = \{0\}$, è sempre vero che $\mathrm{id}_X \in \mathcal{B}(X)$, ma qui $\mathrm{id}_X = 0$ e quindi $\|\mathrm{id}_X\|_{\mathrm{op}} = 0$.

(2) Sia X lo spazio normato $(C(M), \|\cdot\|_\infty)$ discusso nell'esempio 4.5(2), dove M è uno spazio metrico compatto. Sia $Y = \mathbb{F}$ con la norma euclidea. Per ogni $p \in M$, definiamo $V_p : C(M) \to \mathbb{F}$ ponendo

$$V_p f = f(p)$$

per ogni $f \in C(M)$; V_p è detto *operatore di valutazione nel punto* p. Dato che su $C(M)$ utilizziamo le operazioni puntuali, è immediato verificare che V_p è un operatore lineare:

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

per ogni $\alpha, \beta \in \mathbb{F}$ e $f, g \in C(M)$. Inoltre chiaramente

$$|V_p f| = |f(p)| \leq \sup_{x \in M} |f(x)| = \|f\|_\infty,$$

da cui segue che $\|V_p\|_{\mathrm{op}} \leq 1$ e quindi $V_p \in \mathcal{B}(C(M), \mathbb{F})$. In effetti, applicando V_p alla funzione costante $\mathbf{1}_M \in C(M)$ si ha

$$|V_p \mathbf{1}_M| = 1 = \|\mathbf{1}_M\|_\infty,$$

pertanto $\|V_p\|_{\mathrm{op}} = 1$.

(3) Sia $[a,b]$ un intervallo chiuso e limitato di $\mathbb{R}$. Allora l'insieme $C^1[a,b]$ delle funzioni di classe C^1 (derivabili e con derivata continua) su $[a,b]$ è un sottospazio vettoriale di $C[a,b]$. Prendiamo come X lo spazio normato $C^1[a,b]$ con la norma $\|\cdot\|_\infty$ indotta da $C[a,b]$; inoltre prendiamo $Y = \mathbb{F}$ con la norma euclidea. Notiamo che, per ogni $f \in C^1[a,b]$, la sua derivata f' è in $C[a,b]$. Per $p \in [a,b]$, possiamo allora definire $S_p : C^1[a,b] \to \mathbb{F}$ ponendo

$$S_p f = V_p(f') = f'(p)$$

per ogni $f \in C^1[a,b]$, dove V_p è l'operatore di valutazione V_p definito in (2). Non è difficile verificare che $S_p \in \mathcal{L}(C^1[a,b], \mathbb{F})$: questo segue dalla linearità di V_p e dell'operazione di derivazione $f \mapsto f'$. D'altra parte, $S_p : C^1[a,b] \to \mathbb{F}$ non è limitato. Infatti, sia $f_n \in C^1[a,b]$ data da $f_n(t) = \sin(n(t-p))$ per ogni $t \in [a,b]$ e $n \in \mathbb{N}_+$. Allora $f_n'(t) = n\cos(n(t-p))$ e quindi

$$S_p(f_n) = f_n'(p) = n$$

per ogni $n \in \mathbb{N}_+$, mentre

$$\|f_n\|_\infty = \sup_{t \in [a,b]} |\sin(n(t-p))| \leq 1,$$

da cui segue che

$$|S_p(f_n)|/\|f_n\|_\infty \geq n \to \infty \qquad \text{per } n \to \infty,$$

e quindi $\|S_p\|_{\mathrm{op}} = \sup\{|S_p(f)|/\|f\|_\infty : f \neq 0\} = \infty$.

(4) Sia $(H, \langle \cdot, \cdot \rangle)$ uno spazio pre-hilbertiano. Sia $y \in H$. La mappa $\langle \cdot, y \rangle : x \mapsto \langle x, y \rangle$ è una mappa lineare $H \to \mathbb{F}$ (☞ definizioni 6.1(a) e 6.3(a)), e inoltre sappiamo che è continua (☞ corollario 6.16), quindi $\langle \cdot, y \rangle \in \mathcal{B}(H, \mathbb{F})$ (☞ proposizione 7.6). In effetti, dalla disuguaglianza di Cauchy–Schwarz (☞ proposizione 6.10) deduciamo che

$$|\langle x, y \rangle| \leq \|y\| \|x\|$$

per ogni $x \in H$, pertanto $\|\langle \cdot, y \rangle\|_{\mathrm{op}} \leq \|y\|$. Ora, se $y = 0$, ovviamente $\|\langle \cdot, y \rangle\|_{\mathrm{op}} = \|y\| = 0$. D'altra parte, se $y \neq 0$, prendendo $x = y$, si ha anche

$$\|\langle \cdot, y \rangle\|_{\mathrm{op}} \geq |\langle y, y \rangle| / \|y\| = \|y\|^2 / \|y\| = \|y\|$$

e quindi si ha ancora $\|\langle \cdot, y \rangle\|_{\mathrm{op}} = \|y\|$.

(5) Siano $[a, b]$ e $[c, d]$ due intervalli chiusi e limitati in $\mathbb{R}$. Sia $K \in C([a, b] \times [c, d])$. Definiamo l'*operatore integrale* $T_K : C[c, d] \to C[a, b]$ con *nucleo integrale* K ponendo

$$T_K f(x) = \int_c^d K(x, y) f(y)\, dy \tag{7.5}$$

per ogni $f \in C[c, d]$ e $x \in [a, b]$. Non è difficile verificare (per passaggio al limite sotto il segno d'integrale, utilizzando convergenza dominata (☞ teorema 3.7(vi)) e la continuità e limitatezza di K (☞ teorema 2.36)) che, per ogni $f \in C[c, d]$, l'espressione (7.5) effettivamente definisce una funzione continua $x \mapsto T_K f(x)$ su $[a, b]$, dunque $T_K : C[c, d] \to C[a, b]$ è ben definito. Inoltre, per la linearità dell'integrale, si vede che T_K è lineare: per ogni $\alpha, \beta \in \mathbb{F}$, $f, g \in C[c, d]$ e $x \in [a, b]$,

$$\begin{aligned} T_K(\alpha f + \beta g)(x) &= \int_c^d K(x, y)(\alpha f(y) + \beta g(y))\, dy \\ &= \alpha \int_c^d K(x, y) f(y) + \beta \int_c^d K(x, y) g(y)\, dy = \alpha T_K f(x) + \beta T_K g(x) \end{aligned} \tag{7.6}$$

pertanto $T_K(\alpha f + \beta g) = \alpha T_K f + \beta T_K g$. Infine, per ogni $f \in C[c, d]$ e $x \in [a, b]$, si ha

$$|T_K f(x)| \leq \int_c^d |K(x, y)| |f(y)|\, dy \leq \|f\|_\infty \int_c^d |K(x, y)|\, dy$$

e quindi

$$\|T_K f\|_\infty = \sup_{x \in [a, b]} |T_K f(x)| \leq \|f\|_\infty \sup_{x \in [a, b]} \int_c^d |K(x, y)|\, dy.$$

Questo dimostra che

$$\|T_K\|_{\mathrm{op}} \leq \sup_{x \in [a, b]} \int_c^d |K(x, y)|\, dy \leq (d - c) \|K\|_\infty < \infty,$$

perché K è una funzione limitata, e quindi $T_K \in \mathcal{B}(C[c, d], C[a, b])$.

(6) Con la notazione dell'esempio precedente, supponiamo ora soltanto $K \in L^2((a,b) \times (c,d))$. Verifichiamo allora che l'espressione (7.5) definisce un operatore integrale $T_K : L^2(c,d) \to L^2(a,b)$. Siccome $K \in L^2((a,b) \times (c,d))$, si ha

$$\|K\|^2_{L^2((a,b)\times(c,d))} = \int_a^b \int_c^d |K(x,y)|^2 \, dx \, dy < \infty,$$

pertanto, per il teorema di Fubini–Tonelli (☞ teorema 3.9(vi)),

$$\int_c^d |K(x,y)|^2 \, dy < \infty \qquad \text{per q.o. } x \in [a,b],$$

cioè $K(x,\cdot) \in L^2(c,d)$ per q.o. $x \in [a,b]$; per tali $x \in [a,b]$ la disuguaglianza di Hölder (☞ teorema 5.24) implica che, per ogni $f \in L^2(c,d)$, la funzione $y \mapsto K(x,y)f(y)$ è sommabile su $[c,d]$, dunque l'integrale in (7.5) è definito e

$$|T_K f(x)| = \left| \int_c^d K(x,y) f(y) \, dy \right| \leq \|K(x,\cdot)\|_{L^2(c,d)} \|f\|_{L^2(c,d)}.$$

Prendendo i quadrati e integrando in x, si ottiene allora

$$\begin{aligned} \|T_K f\|^2_{L^2(a,b)} &= \int_a^b |T_K f(x)|^2 \, dx \leq \int_a^b \|K(x,\cdot)\|^2_{L^2(c,d)} \|f\|^2_{L^2(c,d)} \, dx \\ &= \|f\|^2_{L^2(c,d)} \int_a^b \int_c^d |K(x,y)|^2 \, dy \, dx = \|K\|^2_{L^2((a,b)\times(c,d))} \|f\|^2_{L^2(c,d)} < \infty, \end{aligned} \tag{7.7}$$

dunque $T_K f \in L^2(a,b)$ per ogni $f \in L^2(c,d)$. Come in (7.6) si verifica che $T_K : L^2(c,d) \to L^2(a,b)$ è lineare. Infine la disuguaglianza (7.7) ci dice anche che

$$\|T_K f\|_{L^2(a,b)} \leq \|K\|_{L^2((a,b)\times(c,d))} \|f\|_{L^2(c,d)}$$

per ogni $f \in L^2(c,d)$, quindi

$$\|T_K\|_{\mathrm{op}} \leq \|K\|_{L^2((a,b)\times(c,d))} < \infty$$

e $T_K \in \mathcal{B}(L^2(c,d), L^2(a,b))$.

(7) Procedendo in maniera perfettamente analoga all'esempio (6), si può verificare più in generale che, se $(M_1, \mathcal{M}_1, \mu_1)$ e $(M_2, \mathcal{M}_2, \mu_2)$ sono spazi di misura σ-finiti e $(M, \mathcal{M}, \mu)$ è lo spazio di misura prodotto (☞ teorema 3.9), allora, per ogni $K \in L^2(M,\mu)$, l'espressione

$$T_K f(x) = \int_{M_2} K(x,y) f(y) \, d\mu_2(y)$$

definisce un operatore $T_K \in \mathcal{B}(L^2(M_2,\mu_2), L^2(M_1,\mu_1))$, detto operatore integrale con nucleo integrale K, e

$$\|T_K\|_{\mathrm{op}} \leq \|K\|_{L^2(M,\mu)}.$$

(8) Sia $\underline{w} \in \ell^\infty$. Definiamo allora la mappa $D_{\underline{w}} : \ell^2 \to \ell^2$ ponendo

$$D_{\underline{w}}\underline{x} = \underline{w} \cdot \underline{x} := (w_k x_k)_{k\in\mathbb{N}}$$

per ogni $\underline{x} \in \ell^2$ (qui $\underline{w} \cdot \underline{x}$ denota il *prodotto componente per componente* delle successioni $\underline{w}$ e $\underline{x}$). È possibile verificare che la mappa $D_{\underline{w}} : \ell^2 \to \ell^2$ è ben definita e lineare, e inoltre

$$\|D_{\underline{w}}\|_{\mathrm{op}} = \|\underline{w}\|_\infty < \infty,$$

dunque $D_{\underline{w}} \in \mathcal{B}(\ell^2)$. Chiamiamo $D_{\underline{w}}$ l'*operatore di moltiplicazione* per $\underline{w}$ su ℓ^2.

(9) Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ spazi normati, e sia $(X \times Y, \|\cdot\|_{X\times Y})$ il loro prodotto (☞ esempio 4.5(4)). È facile verificare che le *proiezioni* $\pi_X : X \times Y \to X$ e $\pi_Y : X \times Y \to Y$ sui due fattori, definite da

$$\pi_X(x, y) = x, \qquad \pi_Y(x, y) = y$$

per ogni $(x, y) \in X \times Y$, sono operatori lineari. Si ha inoltre, per ogni $(x, y) \in X \times Y$,

$$\begin{aligned}\|\pi_X(x, y)\|_X &= \|x\|_X \le \max\{\|x\|_X, \|y\|_Y\} = \|(x, y)\|_{X\times Y},\\ \|\pi_Y(x, y)\|_Y &= \|y\|_Y \le \max\{\|x\|_X, \|y\|_Y\} = \|(x, y)\|_{X\times Y},\end{aligned}$$

pertanto $\|\pi_X\|_{\mathrm{op}} \le 1$ e $\|\pi_Y\|_{\mathrm{op}} \le 1$, e di conseguenza $\pi_X \in \mathcal{B}(X \times Y, X)$ e $\pi_Y \in \mathcal{B}(X \times Y, Y)$. In effetti, si ha anche

$$\begin{aligned}\|\pi_X(x, 0)\|_X &= \|x\|_X = \|(x, 0)\|_{X\times Y},\\ \|\pi_Y(0, y)\|_Y &= \|y\|_Y = \|(0, y)\|_{X\times Y}\end{aligned}$$

per ogni $x \in X$ e $y \in Y$; da questo è facile dedurre che

$$\|\pi_X\|_{\mathrm{op}} = \begin{cases} 1 & \text{se } X \neq \{0\},\\ 0 & \text{se } X = \{0\},\end{cases} \qquad \|\pi_Y\|_{\mathrm{op}} = \begin{cases} 1 & \text{se } Y \neq \{0\},\\ 0 & \text{se } Y = \{0\}.\end{cases}$$

Nell'esempio che abbiamo discusso (☞ esempio 7.8(3)) di operatori non limitati, lo spazio di partenza $C^1[a, b]$ ha dimensione infinita; in effetti, il prossimo risultato mostra che, quando invece lo spazio di partenza ha dimensione finita, la discussione della limitatezza di un operatore lineare si banalizza.

Proposizione 7.9. *Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ spazi normati. Se $\dim X < \infty$, allora $\mathcal{B}(X, Y) = \mathcal{L}(X, Y)$, cioè tutti gli operatori lineari da X a Y sono limitati.*

Dimostrazione. Ovviamente $\mathcal{B}(X, Y) \subseteq \mathcal{L}(X, Y)$, dunque ciò che dobbiamo dimostrare è l'inclusione opposta.

Sia $n = \dim X$, e sia $\{e_1, \dots, e_n\}$ una base di X. Allora la mappa

$$(\alpha_1, \dots, \alpha_n) \mapsto \alpha_1 e_1 + \dots + \alpha_n e_n \tag{7.8}$$

è un isomorfismo lineare da $\mathbb{F}^n$ a X (ovverosia una mappa lineare biiettiva); in particolare, possiamo definire una norma $\|\cdot\|$ su X ponendo

$$\|\alpha_1 e_1 + \dots + \alpha_n e_n\| = |\alpha_1| + \dots + |\alpha_n| = \|(\alpha_1, \dots, \alpha_n)\|_1 \tag{7.9}$$

per ogni $(\alpha_1, \dots, \alpha_n) \in \mathbb{F}^n$, ovverosia "trapiantando" su X tramite l'isomorfismo (7.8) la norma $\|\cdot\|_1$ su $\mathbb{F}^n$. Per il teorema di equivalenza delle norme (☞ teorema 4.13), siccome $\dim X < \infty$, le norme $\|\cdot\|_X$ e $\|\cdot\|$ sono equivalenti, dunque esiste $M \in (0, \infty)$ tale che

$$M^{-1}\|x\| \le \|x\|_X \le M\|x\| \tag{7.10}$$

per ogni $x \in X$.

Allora, per ogni $x \in X$, se sviluppiamo $x = \alpha_1 e_1 + \dots + \alpha_n e_n$ rispetto alla base $\{e_1, \dots, e_n\}$, si ha

$$\begin{aligned}\|Tx\|_Y = \|T(\alpha_1 e_1 + \dots + \alpha_n e_n)\|_Y &\le |\alpha_1|\|Te_1\|_Y + \dots + |\alpha_n|\|Te_n\|_Y \\ &\le C\sum_{k=1}^n |\alpha_k| = C\|x\| \le CM\|x\|_X,\end{aligned}$$

dove $C := \max\{\|Te_1\|_Y, \dots, \|Te_n\|_Y\}$, e dove abbiamo usato (7.9) e (7.10). Ne concludiamo che $\|T\|_{\mathrm{op}} \le CM < \infty$ e quindi $T \in \mathcal{B}(X, Y)$. □

Proposizione 7.10. *Siano X, Y, Z spazi normati.*

(i) L'insieme $\mathcal{B}(X, Y)$ degli operatori lineari limitati da X a Y è un sottospazio vettoriale di $\mathcal{L}(X, Y)$, e $\|\cdot\|_{\mathrm{op}}$ è una norma su $\mathcal{B}(X, Y)$ nel senso della definizione 4.1.

(ii) Se $T \in \mathcal{B}(X, Y)$ e $S \in \mathcal{B}(Y, Z)$, allora $ST \in \mathcal{B}(X, Z)$ e

$$\|ST\|_{\mathrm{op}} \le \|S\|_{\mathrm{op}}\|T\|_{\mathrm{op}}. \tag{7.11}$$

Dimostrazione. (i). Per definizione, $\mathcal{B}(X, Y) \subseteq \mathcal{L}(X, Y)$; inoltre l'operatore identicamente nullo 0 da X a Y è limitato, con $\|0\|_{\mathrm{op}} = 0$, pertanto $\mathcal{B}(X, Y) \neq \emptyset$.

Mostriamo ora che $\mathcal{B}(X, Y)$ è chiuso per combinazioni lineari. Siano $T_1, T_2 \in \mathcal{B}(X, Y)$ e $\alpha_1, \alpha_2 \in \mathbb{F}$. Sappiamo già (☞ proposizione 1.20(i)) che $\alpha_1 T_1 + \alpha_2 T_2 \in \mathcal{L}(X, Y)$. Inoltre, per ogni $x \in X$,

$$\begin{aligned}&\|(\alpha_1 T_1 + \alpha_2 T_2)x\|_Y = \|\alpha_1 T_1 x + \alpha_2 T_2 x\|_Y \le |\alpha_1|\|T_1 x\|_Y + |\alpha_2|\|T_2 x\|_Y \\ &\quad \le |\alpha_1|\|T_1\|_{\mathrm{op}}\|x\|_X + |\alpha_2|\|T_2\|_{\mathrm{op}}\|x\|_X = (|\alpha_1|\|T_1\|_{\mathrm{op}} + |\alpha_2|\|T_2\|_{\mathrm{op}})\|x\|_X,\end{aligned}$$

ove si sono usate la disuguaglianza triangolare e la 1-omogeneità di $\|\cdot\|_Y$ e la disuguaglianza (7.3); di conseguenza

$$\|\alpha_1 T_1 + \alpha_2 T_2\|_{\mathrm{op}} \le |\alpha_1|\|T_1\|_{\mathrm{op}} + |\alpha_2|\|T_2\|_{\mathrm{op}} < \infty, \tag{7.12}$$

dato che T_1 e T_2 sono limitati, e pertanto anche $\alpha_1 T_1 + \alpha_2 T_2 \in \mathcal{B}(X, Y)$. Per la proposizione 1.4, possiamo concludere che $\mathcal{B}(X, Y)$ è un sottospazio vettoriale di $\mathcal{L}(X, Y)$.

Verifichiamo ora che $\|\cdot\|_{\mathrm{op}}$ ristretta a $\mathcal{B}(X, Y)$ soddisfa le proprietà della definizione 4.1. Per definizione di norma operatoriale e di operatore limitato, chiaramente $\|T\|_{\mathrm{op}} \in [0, \infty)$ per ogni $T \in \mathcal{B}(X, Y)$. Inoltre, sappiamo già che $\|0\|_{\mathrm{op}} = 0$; viceversa, se $T \in \mathcal{B}(X, Y)$ e $\|T\|_{\mathrm{op}} = 0$, allora dalla disuguaglianza

(7.3) deduciamo che $Tx = 0$ per ogni $x \in X$, cioè $T = 0$. Questo dimostra la proprietà di separazione per $\|\cdot\|_{\mathrm{op}}$. Si ha anche, per ogni $T \in \mathcal{B}(X,Y)$ e $\alpha \in \mathbb{F}$,

$$\begin{aligned}\|\alpha T\|_{\mathrm{op}} = \sup_{\|x\|_X \le 1} \|(\alpha T)x\|_Y = \sup_{\|x\|_X \le 1} \|\alpha Tx\|_Y \\ = \sup_{\|x\|_X \le 1} |\alpha| \|Tx\|_Y = |\alpha| \sup_{\|x\|_X \le 1} \|Tx\|_Y = |\alpha| \|T\|_{\mathrm{op}},\end{aligned}$$

ove si è usata la caratterizzazione (7.1) della norma operatoriale; dunque la norma operatoriale $\|\cdot\|_{\mathrm{op}}$ è 1-omogenea. Infine, la disuguaglianza (7.12) con $\alpha_1 = \alpha_2 = 1$ dà la disuguaglianza triangolare per $\|\cdot\|_{\mathrm{op}}$. Concludiamo dunque che $\|\cdot\|_{\mathrm{op}}$ è una norma su $\mathcal{B}(X,Y)$.

(ii). Siano $T \in \mathcal{B}(X,Y)$ e $S \in \mathcal{B}(Y,Z)$. Sappiamo già che $ST \in \mathcal{L}(X,Z)$ (☞ proposizione 1.20(ii)); vediamo ora che $ST : X \to Z$ è un operatore limitato. Per ogni $x \in X$,

$$\|(ST)x\|_Z = \|S(Tx)\|_Z \le \|S\|_{\mathrm{op}} \|Tx\|_Y \le \|S\|_{\mathrm{op}} \|T\|_{\mathrm{op}} \|x\|_X,$$

dove si è usata la disuguaglianza (7.3) applicata agli operatori limitati S e T; da questo segue che

$$\|ST\|_{\mathrm{op}} \le \|S\|_{\mathrm{op}} \|T\|_{\mathrm{op}} < \infty,$$

e dunque $ST \in \mathcal{B}(X,Z)$, come desiderato. □

Osservazione 7.11. La proprietà (7.11) è detta *submoltiplicatività* della norma operatoriale. È possibile mostrare che in generale non vale l'uguaglianza in (7.11), nel senso che ci sono esempi di operatori S e T per cui vale la disuguaglianza stretta.

Sulla base della proposizione 7.10(i), d'ora in poi lo spazio $\mathcal{B}(X,Y)$ degli operatori lineari limitati fra due spazi normati X e Y sarà pensato come spazio normato con la norma operatoriale $\|\cdot\|_{\mathrm{op}}$. Sappiamo già che le operazioni di spazio vettoriale (somma e prodotto scalare-vettore) in $\mathcal{B}(X,Y)$ sono continue (☞ proposizione 4.6(ii)). Come conseguenza della submoltiplicatività della norma operatoriale (☞ proposizione 7.10(ii)), vediamo ora che anche il prodotto di operatori è una mappa continua.

Corollario 7.12. *Siano X, Y, Z spazi normati. La mappa*

$$\mathcal{B}(Y,Z) \times \mathcal{B}(X,Y) \ni (S,T) \mapsto ST \in \mathcal{B}(X,Z)$$

è continua.

Dimostrazione. Come nella dimostrazione della proposizione 4.6(ii), grazie alla caratterizzazione della continuità per successioni (☞ proposizione 2.29) e alla caratterizzazione delle successioni convergenti in uno spazio prodotto (☞ proposizione 2.13(i)), è sufficiente dimostrare che, se $S_n \to S$ in $\mathcal{B}(Y,Z)$ e $T_n \to T$ in $\mathcal{B}(X,Y)$, allora $S_n T_n \to ST$ in $\mathcal{B}(X,Z)$.

Notiamo che, siccome $S_n \to S$ in $\mathcal{B}(Y,Z)$, la successione $(S_n)_n$ è limitata in $\mathcal{B}(Y,Z)$ (☞ proposizione 2.23(v)), cioè $M := \sup_{n\in\mathbb{N}} \|S_n\|_{\mathrm{op}} < \infty$ (☞ proposizione 4.6(iv)). In particolare, per ogni $n \in \mathbb{N}$,

$$\begin{aligned}\|S_n T_n - ST\|_{\mathrm{op}} &= \|S_n(T_n - T) + (S_n - S)T\|_{\mathrm{op}} \\ &\le \|S_n\|_{\mathrm{op}} \|T_n - T\|_{\mathrm{op}} + \|S_n - S\|_{\mathrm{op}} \|T\|_{\mathrm{op}} \\ &\le M \|T_n - T\|_{\mathrm{op}} + \|S_n - S\|_{\mathrm{op}} \|T\|_{\mathrm{op}},\end{aligned} \tag{7.13}$$

ove si sono usate la disuguaglianza triangolare e la submoltiplicatività della norma operatoriale. Siccome $\|S_n - S\|_{\mathrm{op}} \to 0$ e $\|T_n - T\|_{\mathrm{op}} \to 0$ per ipotesi, da (7.13) deduciamo che anche $\|S_n T_n - ST\|_{\mathrm{op}} \to 0$, come desiderato. □

Vediamo ora alcune proprietà di sottospazi vettoriali naturalmente associati a un operatore limitato.

Proposizione 7.13. *Siano X e Y spazi normati. Sia $T \in \mathcal{B}(X, Y)$.*

(i) Il nucleo $\operatorname{Ker} T$ *di T è un sottospazio vettoriale chiuso di X.*

(ii) Il grafico $\Gamma(T) := \{(x, Tx) : x \in X\}$ *di T è un sottospazio vettoriale chiuso dello spazio prodotto $X \times Y$.*

Dimostrazione. (i). Siccome $T \in \mathcal{L}(X, Y)$, sappiamo già che $\operatorname{Ker} T$ è un sottospazio vettoriale di X (☞ proposizione 1.20(vii)). Inoltre, $T : X \to Y$ è continua (☞ proposizione 7.6) e inoltre $\{0\}$ è un sottoinsieme chiuso di Y, pertanto $\operatorname{Ker} T = T^{-1}(\{0\})$ è un sottoinsieme chiuso di X (☞ proposizione 2.28).

(ii). Definiamo $S : X \times Y \to Y$ ponendo

$$S(x, y) = y - Tx$$

per ogni $(x, y) \in X \times Y$. In altre parole,

$$S = \pi_Y - T\pi_X,$$

ove $\pi_X : X \times Y \to X$ e $\pi_Y : X \times Y \to Y$ sono le proiezioni sui fattori (☞ esempio 7.8(9)). Siccome T, π_X, π_Y sono operatori limitati, dalla proposizione 7.10 segue che anche S è un operatore limitato. Pertanto, per (i), il nucleo di S è un sottospazio vettoriale chiuso di $X \times Y$. D'altra parte,

$$\begin{aligned}\operatorname{Ker} S = \{(x, y) \in X \times Y : S(x, y) = 0\} &= \{(x, y) \in X \times Y : y = Tx\} \\ &= \{(x, Tx) : x \in X\} = \Gamma(T),\end{aligned}$$

pertanto $\Gamma(T)$ è un sottospazio vettoriale chiuso di $X \times Y$, come desiderato. □

Se X e Y sono spazi normati, abbiamo visto (☞ proposizione 7.10) che $\mathcal{B}(X, Y)$ è uno spazio normato con la norma operatoriale. È naturale chiedersi sotto quali ipotesi $\mathcal{B}(X, Y)$ sia completo, cioè uno spazio di Banach.

Dalla caratterizzazione (7.1) si vede che la norma operatoriale $\|T\|_{\mathrm{op}}$ di un operatore $T \in \mathcal{B}(X, Y)$ ha una natura analoga alla norma dell'estremo superiore $\|f\|_\infty$ di una funzione continua, definita in (4.1): specificamente, si può pensare alla norma operatoriale come a una sorta di "norma infinito" della restrizione $T|_{\overline{B}_X(0,1)}$ dell'operatore alla palla unitaria di X. Non è allora sorprendente che una variante della dimostrazione del teorema 2.14 per lo spazio $C[a, b]$ si possa usare per dimostrare il seguente risultato.

Proposizione 7.14. *Siano X uno spazio normato e Y uno spazio di Banach. Allora $\mathcal{B}(X, Y)$ con la norma operatoriale è uno spazio di Banach.*

Dimostrazione (cenno). Sia $(T_n)_{n\in\mathbb{N}}$ una successione di Cauchy in $\mathcal{B}(X, Y)$.

Passo 1. Per ogni $x \in X$, la successione $(T_n x)_{n\in\mathbb{N}}$ è di Cauchy in Y. Infatti, la successione $(T_n x)_n$ è l'immagine termine a termine della successione $(T_n)_n$

tramite la mappa di valutazione $\mathcal{B}(X,Y) \ni S \mapsto Sx \in Y$, e tale mappa è $\|x\|_X$-lipschitziana, perché

$$\|Sx - S'x\|_Y = \|(S - S')x\|_Y \leq \|S - S'\|_{\mathrm{op}}\|x\|_X;$$

pertanto il fatto che $(T_n)_n$ sia di Cauchy in $\mathcal{B}(X,Y)$ implica che $(T_n x)_n$ sia di Cauchy in Y (☞ proposizione 2.30 e osservazione 2.34).

Passo 2. Siccome Y è uno spazio di Banach, per ogni $x \in X$, la successione di Cauchy $(T_n x)_n$ converge in Y e possiamo definire $Tx = \lim_{n\to\infty} T_n x$. In questo modo, abbiamo definito una mappa $T : X \to Y$.

Passo 3. $T : X \to Y$ è lineare. Infatti, se $\alpha, \alpha' \in \mathbb{F}$ e $x, x' \in X$, allora

$$\begin{aligned} T(\alpha x + \alpha' x') = \lim_{n\to\infty} T_n(\alpha x + \alpha' x') &= \lim_{n\to\infty} (\alpha T_n x + \alpha' T_n x') \\ &= \alpha \lim_{n\to\infty} T_n x + \alpha' \lim_{n\to\infty} T_n x' = \alpha T x + \alpha' T x', \end{aligned}$$

ove si sono usate la linearità dei T_n e la continuità delle operazioni di somma e prodotto scalare-vettore in Y (☞ proposizione 4.6(ii)).

Passo 4. $T : X \to Y$ è limitato. Infatti, per ogni $x \in X$, per la continuità della norma (☞ proposizione 4.6(i)),

$$\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \leq \sup_{n\in\mathbb{N}} \|T_n x\|_Y \leq \|x\|_X \sup_{n\in\mathbb{N}} \|T_n\|_{\mathrm{op}},$$

dove si è usata (7.3) applicata agli operatori limitati T_n; pertanto

$$\|T\|_{\mathrm{op}} \leq \sup_{n\in\mathbb{N}} \|T_n\|_{\mathrm{op}} < \infty,$$

dove l'ultima disuguaglianza segue dal fatto che $(T_n)_n$ è di Cauchy in $\mathcal{B}(X,Y)$, quindi è una successione limitata (☞ proposizioni 2.23(v) e 4.6(iv)).

Passo 5. $T_n \to T$ in $\mathcal{B}(X,Y)$ per $n \to \infty$. Infatti, sia $\epsilon > 0$. Siccome $(T_n)_n$ è di Cauchy in $\mathcal{B}(X,Y)$, esiste $N \in \mathbb{N}$ tale che $\|T_n - T_m\|_{\mathrm{op}} < \epsilon$ per ogni $n, m > N$. In particolare, per ogni $x \in X$, da (7.3) deduciamo che

$$\|T_n x - T_m x\|_Y \leq \epsilon \|x\|_X$$

per ogni $n, m > N$. Per fissato $n > N$, passando al limite per $m \to \infty$ nella precedente disuguaglianza e usando la continuità della norma $\|\cdot\|_Y$, otteniamo che

$$\|T_n - Tx\|_Y \leq \epsilon \|x\|_X$$

per ogni $x \in X$; pertanto

$$\|T_n - T\|_{\mathrm{op}} \leq \epsilon$$

per ogni $n > N$. Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $\|T_n - T\|_{\mathrm{op}} \to 0$ per $n \to \infty$, come desiderato. □

7.2 Isometrie lineari

Definizione 7.15. Siano X e Y spazi normati. Una *isometria lineare* da X a Y è un operatore $T \in \mathcal{L}(X,Y)$ che *preserva la norma*, cioè

$$\|Tx\|_Y = \|x\|_X \tag{7.14}$$

per ogni $x \in X$.

Proposizione 7.16. *Siano X e Y spazi normati e $T \in \mathcal{L}(X, Y)$. Sono equivalenti:*

(i) $T : X \to Y$ è un'isometria lineare;

(ii) T preserva la distanza, *cioè*

$$\|Tx - Tx'\|_Y = \|x - x'\|_X \tag{7.15}$$

per ogni $x, x' \in X$.

Inoltre, se $T \in \mathcal{L}(X, Y)$ è un'isometria lineare, allora:

(iii) T è un operatore limitato, con

$$\|T\|_{\mathrm{op}} = \begin{cases} 1 & \text{se } X \neq \{0\}, \\ 0 & \text{se } X = \{0\}. \end{cases}$$

(iv) T è iniettivo;

(v) se T è anche suriettivo, l'inverso $T^{-1} : Y \to X$ è a sua volta un'isometria lineare;

(vi) se Z è uno spazio normato e $S : Y \to Z$ è un'isometria lineare, allora $ST : X \to Z$ è un'isometria lineare.

Dimostrazione. (i) $\Leftrightarrow$ (ii). Applicando (7.14) con $x - x'$ al posto di x e usando la linearità di T, si ottiene (7.15). Viceversa, applicando (7.15) con $x' = 0$ e usando che $T0 = 0$, si ottiene (7.14).

(i) $\Rightarrow$ (iii). Da (7.14) è chiaro che $\|T\|_{\mathrm{op}} \leq 1$, dunque $T \in \mathcal{B}(X, Y)$, e che inoltre $\|T\|_{\mathrm{op}} = 1$ se $X \neq \{0\}$; se invece $X = \{0\}$ si ha $T = 0$ e $\|T\|_{\mathrm{op}} = 0$.

(i) $\Rightarrow$ (iv). Se $x \in \operatorname{Ker} T$, si ha $\|x\|_X = \|Tx\|_Y = \|0\|_Y = 0$, ove si è usata (7.14); pertanto $x = 0$. Questo dimostra che $\operatorname{Ker} T = \{0\}$ e quindi T è iniettiva (☞ proposizione 1.20(ix)).

(i) $\Rightarrow$ (v). Siccome $T \in \mathcal{L}(X, Y)$ e T è biiettiva, allora l'inversa $T^{-1} : Y \to X$ è a sua volta lineare. Inoltre, per ogni $y \in Y$, se $x = T^{-1}y$, allora $y = Tx$ e quindi

$$\|T^{-1}y\|_X = \|x\|_X = \|Tx\|_Y = \|y\|_Y,$$

dove si è usata (7.14) applicata all'isometria T; pertanto anche T^{-1} soddisfa (7.14) e quindi è un'isometria lineare a sua volta.

(i) $\Rightarrow$ (vi). Per ogni $x \in X$, applicando (7.14) a S e a T, si ottiene

$$\|(ST)x\|_Z = \|S(Tx)\|_Z = \|Tx\|_Y = \|x\|_X,$$

pertanto anche ST soddisfa (7.14) e quindi è un'isometria lineare. □

Osservazione 7.17. Non tutti gli operatori lineari T con $\|T\|_{\mathrm{op}} = 1$ sono isometrie lineari. Ad esempio:

(1) Se X e Y sono spazi normati non banali, le proiezioni sui fattori $\pi_X : X \times Y \to X$ e $\pi_Y : X \times Y \to Y$ hanno entrambe norma 1 (☞ esempio 7.8(9)). Tuttavia i loro nuclei $\operatorname{Ker} \pi_X = \{(0, y) : y \in Y\}$ e $\operatorname{Ker} \pi_Y = \{(x, 0) : x \in X\}$ sono non banali, pertanto π_X e π_Y non sono iniettive.

(2) Sia X lo spazio normato $(\mathbb{F}^2, \|\cdot\|_p)$ per qualche $p \in [1,\infty]$ (☞ esempio 4.5(1)). Allora la mappa $T : \mathbb{F}^2 \to \mathbb{F}^2$, data da

$$T(x,y) = (x, y/2)$$

per ogni $(x,y) \in \mathbb{F}^2$, è lineare e iniettiva, e inoltre non è difficile verificare che $\|T\|_{\mathrm{op}} = 1$. D'altra parte,

$$\|T(0,1)\|_p = \|(0,1/2)\|_p = 1/2 \neq 1 = \|(0,1)\|_p,$$

pertanto T non è un isometria.

Osservazione 7.18. Se $T : X \to X$ è un'isometria lineare da uno spazio X in sé, e $\dim X < \infty$, allora T è suriettiva (dunque biiettiva): infatti, in tal caso, data l'iniettività di T, si ha

$$\dim \operatorname{Im} T = \dim X$$

(☞ proposizione 1.20(xi)) e quindi $\operatorname{Im} T = X$ per via della dimensione finita. In generale, tuttavia, non è detto che un'isometria lineare sia suriettiva. Ad esempio:

(1) Per ogni $p \in [1,\infty]$, la mappa $T : (\mathbb{F}, \|\cdot\|_p) \to (\mathbb{F}^2, \|\cdot\|_p)$, definita da $Tx = (x,0)$ per ogni $x \in \mathbb{F}$, è un'isometria lineare non suriettiva.

(2) Per ogni $p \in [1,\infty]$, la mappa $T : \ell^p \to \ell^p$, definita da

$$T\underline{x} = (0, x_0, x_1, x_2, \dots)$$

per ogni $\underline{x} = (x_0, x_1, x_2, \dots) \in \ell^p$, è un'isometria lineare non suriettiva da ℓ^p in sé.

Definizione 7.19. Siano X e Y spazi normati.

(a) Un *isomorfismo isometrico* da X a Y è un'isometria lineare biiettiva da X a Y.

(b) Gli spazi X e Y si dicono *isometricamente isomorfi* se esiste un isomorfismo isometrico da X a Y. In tal caso, scriviamo anche

$$X \underset{\text{isom.}}{\cong} Y.$$

Osservazione 7.20. Siano X, Y, Z spazi normati. Allora:

- $\mathrm{id}_X : X \to X$ è banalmente un isomorfismo isometrico da X in sé;
- se $T : X \to Y$ è un isomorfismo isometrico, anche $T^{-1} : Y \to X$ lo è (☞ proposizione 7.16(v));
- se $T : X \to Y$ e $S : Y \to Z$ sono isomorfismi isometrici, anche $ST : X \to Z$ lo è (☞ proposizione 7.16(vi)).

Dalle tre proprietà sopra segue che la relazione "essere isometricamente isomorfo a" tra spazi normati è riflessiva, simmetrica e transitiva, cioè è una relazione di equivalenza.

Osservazione 7.21. Se due spazi normati sono isometricamente isomorfi e uno dei due è completo, lo è anche l'altro. Questa è un'immediata conseguenza del fatto che un isomorfismo isometrico è bi-lipschitziano (☞ proposizioni 7.16 e 7.6), pertanto preserva le successioni di Cauchy e i limiti di successioni (☞ proposizioni 2.29 e 2.30).

Abbiamo già introdotto una nozione di "essere isometricamente isomorfi" tra spazi pre-hilbertiani (☞ definizione 6.47). Il prossimo risultato mostra che questa definizione è compatibile con quella (☞ definizione 7.19) data sopra per spazi normati.

Proposizione 7.22. *Siano* $(H_1, \langle\cdot,\cdot\rangle_1)$ *e* $(H_2, \langle\cdot,\cdot\rangle_2)$ *spazi pre-hilbertiani e* $T \in \mathcal{L}(H_1, H_2)$. *Sono equivalenti:*

(i) $T : H_1 \to H_2$ *è un'isometria lineare;*

(ii) T preserva il prodotto scalare, *cioè*

$$\langle Tx, Tx'\rangle_2 = \langle x, x'\rangle_1 \tag{7.16}$$

per ogni $x, x' \in X$.

Dimostrazione (cenno). Data la definizione della norma in termini del prodotto scalare (☞ definizione 6.8), l'identità (7.14) segue da (7.16) prendendo $x' = x$. L'implicazione opposta si dimostra in maniera simile mediante l'identità di polarizzazione (☞ proposizione 6.14(i)), che permette di esprimere il prodotto scalare in termini della norma. □

7.3 Estensione di operatori limitati

Teorema 7.23. *Siano* X *uno spazio normato e* Y *uno spazio di Banach. Sia* D *un sottospazio vettoriale* denso *di* X*; pensiamo* D *come spazio normato con la norma indotta da* X. *Per ogni* $T \in \mathcal{B}(D, Y)$, *esiste un unico* $\widetilde{T} \in \mathcal{B}(X, Y)$ *tale che* $\widetilde{T}|_D = T$, *e si ha* $\|\widetilde{T}\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}$; *tale operatore* $\widetilde{T}$ *è detto* estensione per continuità *di* T *a* X.

Dimostrazione. Unicità. Siano $\widetilde{T}$ e $\widetilde{T}'$ due estensioni di $T \in \mathcal{B}(D, Y)$ a X, cioè $\widetilde{T}, \widetilde{T}' \in \mathcal{B}(X, Y)$ e $\widetilde{T}|_D = \widetilde{T}'|_D = T$. Allora $(\widetilde{T} - \widetilde{T}')|_D = T - T = 0$, pertanto

$$\mathrm{Ker}(\widetilde{T} - \widetilde{T}') \supseteq D. \tag{7.17}$$

Siccome $\widetilde{T} - \widetilde{T}' \in \mathcal{B}(X, Y)$ (☞ proposizione 7.10(i)), il nucleo $\mathrm{Ker}(\widetilde{T} - \widetilde{T}')$ è un sottoinsieme chiuso di X (☞ proposizione 7.13(i)); pertanto, prendendo le chiusure in (7.17) si ottiene

$$X \supseteq \mathrm{Ker}(\widetilde{T} - \widetilde{T}') \supseteq \overline{D} = X,$$

dove si è usata la densità di D in X. Quindi $\mathrm{Ker}(\widetilde{T} - \widetilde{T}') = X$, cioè $\widetilde{T} - \widetilde{T}' = 0$ e $\widetilde{T} = \widetilde{T}'$.

Esistenza. Sia $T \in \mathcal{B}(D, Y)$. Vogliamo costruire l'estensione $\widetilde{T} \in \mathcal{B}(X, Y)$ dell'operatore T.

Passo 1. Sia $x \in X$. Siccome $\overline{D} = X$, esistono successioni $(x_n)_{n\in\mathbb{N}}$ a valori in D tali che $x_n \to x$ in X per $n \to \infty$ (☞ proposizione 2.23(vii)).

Passo 2. Sia $x \in X$. Per ogni successione $(x_n)_{n\in\mathbb{N}}$ a valori in D tali che $x_n \to x$, la successione $(Tx_n)_{n\in\mathbb{N}}$ converge in Y. Infatti $(x_n)_n$ è di Cauchy (☞ proposizione 2.9(iii)) e $T : D \to Y$ è uniformemente continuo (☞ proposizione 7.6), pertanto anche $(Tx_n)_n$ è di Cauchy in Y (☞ proposizione 2.30) e quindi converge, siccome Y è completo per ipotesi.

Passo 3. Sia $x \in X$. Se $(x_n)_{n\in\mathbb{N}}$ è come nel passo 2, allora $\lim_{n\to\infty} Tx_n$ dipende solo da x e non dalla successione $(x_n)_n$ che vi converge. Infatti, sia $(x'_n)_{n\in\mathbb{N}}$ un'altra successione a valori in D che converge a x in X. Allora la successione $(\tilde{x}_n)_{n\in\mathbb{N}}$, definita da

$$\tilde{x}_{2n} = x_n, \qquad \tilde{x}_{2n+1} = x'_n \qquad \text{per ogni } n \in \mathbb{N},$$

è ancora una successione a valori in D, e $\tilde{x}_n \to x$ per $n \to \infty$; pertanto, applicando a questa successione il passo 2, deduciamo che $(T\tilde{x}_n)_n$ converge in Y. D'altra parte, si ha

$$T\tilde{x}_{2n} = Tx_n, \qquad T\tilde{x}_{2n+1} = Tx'_n \qquad \text{per ogni } n \in \mathbb{N},$$

dunque $(Tx_n)_n$ e $(Tx'_n)_n$ sono sottosuccessioni di $(T\tilde{x}_n)_n$, e quindi convergono tutte allo stesso limite.

Passo 4. Per ogni $x \in X$, definiamo

$$\widetilde{T}x = \lim_{n\to\infty} Tx_n, \tag{7.18}$$

dove $(x_n)_n$ è una qualunque successione a valori in D tale che $x_n \to x$ in X. Osserviamo che $\widetilde{T}x \in Y$ è ben definito da (7.18) per i passi 1, 2 e 3 (il limite esiste e non dipende dalla successione), per ogni $x \in X$. In particolare, abbiamo costruito una mappa $\widetilde{T} : X \to Y$.

Passo 5. $\widetilde{T}|_D = T$. Infatti, se $x \in D$, posso prendere come $(x_n)_n$ la successione costante uguale a x, pertanto

$$\widetilde{T}x = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Tx = Tx.$$

Passo 6. $\widetilde{T} : X \to Y$ è lineare. Siano $x, x' \in X$ e $\alpha, \alpha' \in \mathbb{F}$. Per il passo 1, esistono successioni $(x_n)_n$ e $(x'_n)_n$ a valori in D con $x_n \to x$ e $x'_n \to x'$ per $n \to \infty$. Allora $\alpha x_n + \alpha' x'_n \in D$ per ogni $n \in \mathbb{N}$ (perché D è un sottospazio vettoriale) e $\alpha x_n + \alpha' x'_n \to \alpha x + \alpha' x'$ (☞ proposizione 4.6(ii)). Pertanto, possiamo usare la successione $(\alpha x_n + \alpha' x'_n)_n$ per calcolare $\widetilde{T}(\alpha x + \alpha' x')$ come in (7.18), ottenendo

$$\begin{aligned} \widetilde{T}(\alpha x + \alpha' x') &= \lim_{n\to\infty} T(\alpha x_n + \alpha' x'_n) = \lim_{n\to\infty} (\alpha T x_n + \alpha' T x'_n) \\ &= \alpha \lim_{n\to\infty} Tx_n + \alpha' \lim_{n\to\infty} Tx'_n = \alpha \widetilde{T}x + \alpha' \widetilde{T}x', \end{aligned}$$

dove abbiamo usato la linearità di T, la continuità delle operazioni (☞ proposizione 4.6(ii)) e nuovamente (7.18).

Passo 7. $\|\widetilde{T}\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}$. La disuguaglianza $\|\widetilde{T}\|_{\mathrm{op}} \geq \|T\|_{\mathrm{op}}$ segue immediatamente dalla definizione di norma operatoriale, perché $\widetilde{T}|_D = T$. Per la disuguaglianza opposta, osserviamo che, se $x \in X$, presa $(x_n)_n$ a valori in D con $x_n \to x$, si ha $\widetilde{T}x = \lim_{n\to\infty} Tx_n$ e quindi, per la continuità della norma,

$$\|\widetilde{T}x\|_Y = \lim_{n\to\infty} \|Tx_n\|_Y \leq \lim_{n\to\infty} (\|T\|_{\mathrm{op}} \|x_n\|_X) = \|T\|_{\mathrm{op}} \|x\|_X;$$

da questo segue che $\|\widetilde{T}\|_{\mathrm{op}} \leq \|T\|_{\mathrm{op}}$. □

Osservazione 7.24. Nella dimostrazione del teorema 7.23, per l'unicità dell'estensione l'ipotesi di completezza di Y non è usata (sono sufficienti la densità di D in X e la continuità dell'estensione). L'ipotesi di completezza di Y è invece essenziale per l'esistenza dell'estensione. Nelle applicazioni, capita tuttavia di voler estendere operatori $T : D \to E$ il cui codominio E non è completo; in tal caso, se il codominio è a sua volta sottospazio di uno spazio di Banach Y, è sufficente "estendere il codominio" e applicare il teorema 7.23 all'operatore $T : D \to Y$. Per comodità, enunciamo separatamente questo risultato.

Corollario 7.25. *Siano X uno spazio normato e Y uno spazio di Banach. Siano D ed E sottospazi vettoriali di X ed Y rispettivamente; dotiamo D ed E delle norme indotte. Supponiamo che D sia denso in X. Per ogni $T \in \mathcal{B}(D, E)$, esiste un unico $\widetilde{T} \in \mathcal{B}(X, Y)$ tale che $\widetilde{T}|_D = T$, e si ha $\|\widetilde{T}\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}$.*

Osservazione 7.26. Vedremo in seguito, con il teorema di Hahn–Banach (☞ teorema 8.31), che nel caso $Y = \mathbb{F}$ l'esistenza di un'estensione con la stessa norma si può dimostrare anche senza l'ipotesi di densità di D. In tal caso tuttavia l'estensione non è necessariamente unica.

7.4 Il principio di uniforme limitatezza

Teorema 7.27 (Banach–Steinhaus)**.** *Siano X uno spazio di Banach e Y uno spazio normato. Sia $\mathcal{F} \subseteq \mathcal{B}(X, Y)$ tale che*

$$\sup_{T\in\mathcal{F}} \|Tx\|_Y < \infty \qquad \textit{per ogni } x \in X. \tag{7.19}$$

Allora

$$\sup_{T\in\mathcal{F}} \|T\|_{\mathrm{op}} < \infty. \tag{7.20}$$

Osservazione 7.28. Chiaramente dalla condizione (7.20) si deduce

$$\sup_{T\in\mathcal{F}} \|Tx\|_Y \le \|x\|_X \sup_{T\in\mathcal{F}} \|T\|_{\mathrm{op}} < \infty$$

per ogni $x \in X$, cioè (7.19). Il teorema 7.27 afferma che, sorprendentemente, l'implicazione si può invertire quando X è completo.

Per dimostrare il teorema di Banach–Steinhaus faremo uso di un risultato profondo della teoria degli spazi metrici.

Teorema 7.29 (lemma di Baire)**.** *Sia (M, d) uno spazio metrico completo.*

(i) Sia $\{A_n\}_{n\in\mathbb{N}}$ una famiglia numerabile di sottoinsiemi aperti densi di M. Allora $\overline{\bigcap_{n\in\mathbb{N}} A_n} = M$, cioè anche la loro intersezione è densa.

(ii) Supponiamo che $M \neq \emptyset$. Sia $\{C_n\}_{n\in\mathbb{N}}$ una famiglia di sottoinsiemi chiusi di M tali che $\bigcup_{n\in\mathbb{N}} C_n = M$. Allora esiste $n \in \mathbb{N}$ tale che $\mathring{C}_n \neq \emptyset$.

Dimostrazione (cenno). (i). Siano $x \in M$ e $\epsilon > 0$. Costruiamo induttivamente, per ogni $n \in \mathbb{N}$, punti $x_n \in M$ e raggi $r_n \in (0, \infty)$ tali che

$$\overline{B}(x_n, r_n) \subseteq \begin{cases} A_n \cap B(x_{n-1}, r_{n-1}) & \text{se } n > 0, \\ A_0 \cap B(x, \epsilon) & \text{se } n = 0, \end{cases} \tag{7.21}$$

$$r_n \le 2^{-n}. \tag{7.22}$$

Per $n = 0$, siccome A_0 è denso in M, si ha $B(x, \epsilon) \cap A_0 \neq \emptyset$ e inoltre $B(x, \epsilon) \cap A_0$ è aperto (intersezione di aperti), quindi possiamo scegliere $x_0 \in B(x, \epsilon) \cap A_0$ e $r_0 \in (0, 1]$ tali che $B(x_0, 2r_0) \subseteq A_0 \cap B(x, \epsilon)$. In particolare le condizioni (7.21)-(7.22) sono soddisfatte per $n = 0$.

Supponiamo ora di aver costruito, per qualche $N \in \mathbb{N}_+$, punti $x_0, \dots, x_{N-1}$ e raggi $r_0, \dots, r_{N-1}$ in modo che (7.21)-(7.22) sono soddisfatte per $n = 0, \dots, N-1$. Procedendo come sopra, siccome A_N è aperto e denso in M, si ha che $A_N \cap B(x_{N-1}, r_{N-1}) \neq \emptyset$ ed esistono $x_N \in A_N \cap B(x_{N-1}, r_{N-1})$ e $r_N \in (0, 2^{-N}]$ tali che $B(x_N, 2r_N) \subseteq B_{N-1} \cap A_N$. Di conseguenza le condizioni (7.21)-(7.22) sono soddisfatte per $n = N$.

Abbiamo dunque costruito una successione $(x_n)_n$ con le proprietà cercate. Osserviamo ora che dalla proprietà (7.21) si deduce induttivamente che, per ogni $n, k \in \mathbb{N}$,

$$\overline{B}(x_n, r_n) \subseteq B(x_k, r_k) \qquad \text{se } n > k, \tag{7.23}$$

e quindi anche

$$\overline{B}(x_n, r_n) \subseteq B(x, \epsilon) \cap \bigcap_{k=0}^{n} A_k. \tag{7.24}$$

Deduciamo ora che $(x_n)_n$ è una successione di Cauchy in M. Infatti, sia $\delta > 0$. Possiamo allora trovare $N \in \mathbb{N}$ tale che $2^{-N} < \delta/2$. Per ogni $n, m > N$, da (7.23) segue che $x_n, x_m \in B(x_N, r_N)$ e quindi

$$d(x_n, x_m) \leq d(x_n, x_N) + d(x_N, x_m) < 2r_N \leq 2 \cdot 2^{-N} < \delta,$$

ove si è usata (7.22).

Siccome M è completo, la successione di Cauchy $(x_n)_n$ converge a un punto $\bar{x} \in M$. Ora, da (7.23) deduciamo che $x_n \in B(x_k, r_k)$ per ogni $n > k$, e passando al limite per $n \to \infty$ e $k \in \mathbb{N}$ fissato, si ottiene che

$$\bar{x} \in \overline{B}(x_k, r_k) \qquad \text{per ogni } k \in \mathbb{N}.$$

Questo insieme a (7.24) implica che

$$\bar{x} \in B(x, \epsilon) \cap \bigcap_{n \in \mathbb{N}} A_n$$

e quindi $B(x, \epsilon) \cap \bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$. Per l'arbitrarietà di $\epsilon > 0$, questo mostra che

$$x \in \overline{\bigcap_{n \in \mathbb{N}} A_n}.$$

Per l'arbitrarietà di $x \in M$, questo mostra che $\bigcap_{n \in \mathbb{N}} A_n$ è denso in M.

(ii). Definiamo $A_n = M \setminus C_n$ per ogni $n \in \mathbb{N}$. Siccome $\bigcup_{n \in \mathbb{N}} C_n = M$, passando ai complementari si ha $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$. D'altra parte, gli A_n sono aperti (in quanto complementari di chiusi). Se tutti gli A_n fossero densi in M, allora per (i) anche $\bigcap_{n \in \mathbb{N}} A_n$ sarebbe denso, ma questo è assurdo perché tale intersezione è vuota, mentre $M \neq \emptyset$. Di conseguenza, esiste $n \in \mathbb{N}$ tale che $\overline{A_n} \neq M$ e, passando ai complementari (☞ proposizione 2.23(viii)), si ottiene $\mathring{C}_n \neq \emptyset$. □

Procediamo ora a dimostrare il teorema di Banach–Steinhaus.

Dimostrazione del teorema 7.27. Sia $\mathcal{F} \subseteq \mathcal{B}(X, Y)$ una famiglia di operatori che soddisfa la condizione (7.19). Per ogni $n \in \mathbb{N}$, definiamo l'insieme

$$C_n = \bigcap_{T \in \mathcal{F}} \{x \in X : \|Tx\|_Y \leq n\}. \tag{7.25}$$

Notiamo che, per ogni $T \in \mathcal{F}$, la funzione $\Phi_T : X \ni x \mapsto \|Tx\|_Y \in \mathbb{R}$ è continua, in quanto composizione delle funzioni continue T e $\|\cdot\|_Y$; ma allora, per ogni $n \in \mathbb{N}$, l'insieme

$$\{x \in X : \|Tx\|_Y \leq n\} = \Phi_T^{-1}((-\infty, n])$$

è un sottoinsieme chiuso di X, dato che è la controimmagine del sottoinsieme chiuso $(-\infty, n]$ di $\mathbb{R}$ tramite la funzione continua Φ_T (☞ proposizione 2.28). Pertanto anche l'insieme C_n definito in (7.25) è un sottoinsieme chiuso di X, in quanto intersezione di chiusi (☞ proposizione 2.23(iii)).

Osserviamo ora che, per la condizione (7.19), per ogni $x \in X$ possiamo trovare $n \in \mathbb{N}$ tale che $\sup_{T \in \mathcal{F}} \|Tx\|_Y \leq n$; ma allora $x \in C_n$ per (7.25). Questo dimostra che $\bigcup_{n \in \mathbb{N}} C_n = X$.

Chiaramente $0 \in X$, quindi $X \neq \emptyset$. Possiamo allora applicare il teorema 7.29(ii) allo spazio metrico completo X e alla famiglia $\{C_n\}_n$ di chiusi, deducendone che esiste $n \in \mathbb{N}$ tale che $\mathring{C}_n \neq \emptyset$. In altre parole, esistono $\bar{x} \in C_n$ e $r > 0$ tali che $B(\bar{x}, r) \subseteq C_n$.

Sia ora $x \in X$ con $\|x\|_X \leq 1$. Allora $\bar{x}, \bar{x} + (r/2)x \in B(\bar{x}, r) \subseteq C_n$ e quindi, per la definizione (7.25) di C_n, per ogni $T \in \mathcal{F}$,

$$\|T\bar{x}\|_Y \leq n, \qquad \|T(\bar{x} + (r/2)x)\|_Y \leq n;$$

per le proprietà della norma e la linearità di T, si ha allora anche

$$\begin{aligned} \|Tx\|_Y = (2/r)\|T((r/2)x)\|_Y &= (2/r)\|T(\bar{x} + (r/2)x) - T\bar{x}\|_Y \\ &\leq (2/r)(\|T(\bar{x} + (r/2)x)\|_Y + \|T\bar{x}\|_Y) \leq 4n/r. \end{aligned}$$

Passando all'estremo superiore su $x \in X$ con $\|x\|_X \leq 1$, si ottiene

$$\|T\|_{\mathrm{op}} \leq 4n/r$$

(☞ lemma 7.5), e siccome questa stima vale per ogni $T \in \mathcal{M}$, ne deduciamo anche

$$\sup_{T \in \mathcal{F}} \|T\|_{\mathrm{op}} \leq 4n/r < \infty,$$

come desiderato. □

Vediamo ora un'importante conseguenza del teorema di Banach–Steinhaus relativa al limite puntuale (☞ definizione 2.16) di una successione di operatori limitati.

Corollario 7.30. *Siano X uno spazio di Banach e Y uno spazio normato. Sia $(T_n)_{n \in \mathbb{N}}$ una successione a valori in $\mathcal{B}(X, Y)$ tale che, per ogni $x \in X$, esiste il limite $\lim_{n \to \infty} T_n x$ in Y. Allora $\sup_{n \in \mathbb{N}} \|T_n\|_{\mathrm{op}} < \infty$. Inoltre, se definiamo $T : X \to Y$ ponendo*

$$Tx = \lim_{n \to \infty} T_n x \qquad \text{per ogni } x \in X,$$

allora $T \in \mathcal{B}(X,Y)$ *e*

$$\|T\|_{\mathrm{op}} \le \liminf_{n\to\infty} \|T_n\|_{\mathrm{op}} \le \sup_{n\in\mathbb{N}} \|T_n\|_{\mathrm{op}}.$$

Dimostrazione. Passo 1. $T : X \to Y$ è lineare. Infatti, per ogni $x, x' \in X$ e $\alpha, \alpha' \in \mathbb{F}$,

$$\begin{aligned} T(\alpha x + \alpha' x') &= \lim_{n\to\infty} T_n(\alpha x + \alpha' x') = \lim_{n\to\infty} (\alpha T_n x + \alpha' T_n x') \\ &= \alpha \lim_{n\to\infty} T_n x + \alpha' \lim_{n\to\infty} T_n x' = \alpha T x + \alpha' T x', \end{aligned}$$

dove si sono usate la linearità dei T_n e la continuità delle operazioni.

Passo 2. $\sup_{n\in\mathbb{N}} \|T_n\|_{\mathrm{op}} < \infty$. Per vedere questo, applicando il teorema 7.27 alla famiglia $\mathcal{F} = \{T_n\}_{n\in\mathbb{N}}$, è sufficiente verificare che

$$\sup_{n\in\mathbb{N}} \|T_n x\|_Y < \infty. \tag{7.26}$$

per ogni $x \in X$. D'altra parte, per ipotesi, per ogni $x \in X$ la successione $(T_n x)_{n\in\mathbb{N}}$ è convergente in Y, ma allora tale successione è limitata (☞ proposizione 2.23(v)), e quindi (7.26) è verificata (☞ proposizione 4.6(iv)).

Passo 3. Sia $K = \liminf_{n\to\infty} \|T_n\|_{\mathrm{op}}$. Dal passo 2 sappiamo che $K \le \sup_{n\in\mathbb{N}} \|T_n\|_{\mathrm{op}} < \infty$. Inoltre, esiste una sottosuccessione $(T_{n_k})_k$ tale che $K = \lim_{k\to\infty} \|T_{n_k}\|_{\mathrm{op}}$. Allora, per ogni $x \in X$ e $k \in \mathbb{N}$,

$$\|T_{n_k} x\|_Y \le \|T_{n_k}\|_{\mathrm{op}} \|x\|_X;$$

passando al limite per $k \to \infty$ e usando la continuità della norma $\|\cdot\|_Y$, si ottiene

$$\|Tx\|_Y \le K \|x\|_X$$

per ogni $x \in X$, cioè $\|T\|_{\mathrm{op}} \le K < \infty$ e $T \in \mathcal{B}(X,Y)$, come desiderato. □

Osservazione 7.31. Il corollario 7.30 è abbastanza sorprendente se confrontato con il fatto che, in generale, il limite puntuale di una successione di funzioni continue non è necessariamente una funzione continua (☞ osservazione 2.18). Il corollario 7.30 afferma che, sotto l'ipotesi aggiuntiva che tali funzioni siano lineari (e lo spazio di partenza sia completo), allora si può concludere che il limite puntuale è continuo (oltre che lineare).

7.5 Operatori limitati invertibili

Definizione 7.32. Siano X, Y spazi normati. Un operatore $T \in \mathcal{B}(X,Y)$ si dice *invertibile con inverso limitato*, oppure un *isomorfismo*, se $T : X \to Y$ è biiettivo e $T^{-1} \in \mathcal{B}(Y,X)$.

Osservazione 7.33. Siano X, Y, Z spazi normati.

(1) Chiaramente $\mathrm{id}_X \in \mathcal{B}(X)$ è un isomorfismo, con $\mathrm{id}_X^{-1} = \mathrm{id}_X$.

(2) Se $T \in \mathcal{B}(X,Y)$ è un isomorfismo, allora anche $T^{-1} \in \mathcal{B}(Y,X)$ lo è, e $(T^{-1})^{-1} = T$.

(3) Se $T \in \mathcal{B}(X,Y)$ e $S \in \mathcal{B}(Y,Z)$ sono isomorfismi, allora anche $ST \in \mathcal{B}(X,Z)$ è un isomorfismo e $(ST)^{-1} = T^{-1}S^{-1}$.

(4) Se $T \in \mathcal{B}(X)$ e $S \in \mathcal{B}(X)$ commutano ($ST = TS$) e S è un isomorfismo, allora anche S^{-1} e T commutano ($S^{-1}T = TS^{-1}$).

(5) Se $T \in \mathcal{B}(X)$ e $S \in \mathcal{B}(X)$ commutano e ST è un isomorfismo, allora anche S e T sono isomorfismi, e $S^{-1} = T(ST)^{-1}$ e $T^{-1} = S(ST)^{-1}$.

(6) Se $T \in \mathcal{B}(X,Y)$ è un isomorfismo e $X \neq \{0\}$, allora

$$1 = \|\mathrm{id}_X\|_{\mathrm{op}} = \|T^{-1}T\|_{\mathrm{op}} \leq \|T^{-1}\|_{\mathrm{op}}\|T\|_{\mathrm{op}},$$

dove abbiamo usato la submoltiplicatività della norma (☞ proposizione 7.10(ii)); di conseguenza

$$\|T^{-1}\|_{\mathrm{op}} \geq \frac{1}{\|T\|_{\mathrm{op}}}. \tag{7.27}$$

Esempio 7.34. La disuguaglianza in (7.27) può essere stretta. Ad esempio, se $X = Y$ è $(\mathbb{F}^2, \|\cdot\|_p)$ e $T \in \mathcal{B}(X)$ è la mappa dell'osservazione 7.17(2), cioè

$$T(x,y) = (x, y/2) \qquad \forall (x,y) \in \mathbb{F}^2,$$

allora $\|T\|_{\mathrm{op}} = 1$, mentre l'inversa è data da

$$T^{-1}(x,y) = (x, 2y) \qquad \forall (x,y) \in \mathbb{F}^2,$$

e si ha $\|T^{-1}\|_{\mathrm{op}} = 2$. Dunque in questo caso $\|T\|_{\mathrm{op}}\|T^{-1}\|_{\mathrm{op}} = 2 > 1$.

Esempio 7.35. Siano X, Y spazi normati. Se un operatore $T \in \mathcal{L}(X,Y)$ è biiettivo, allora anche l'inverso $T^{-1} : Y \to X$ è un operatore lineare. Tuttavia, se T è limitato, in generale non è detto che T^{-1} sia anche limitato.

(1) Se $\dim X < \infty$, sappiamo (☞ proposizione 7.9) che $\mathcal{B}(X,Y) = \mathcal{L}(X,Y)$. Quindi, se $\dim X < \infty$ e $\dim Y < \infty$, si ha banalmente che ogni operatore $T \in \mathcal{B}(X,Y)$ biiettivo è un isomorfismo.

(2) Sia $X = Y = c_{00}$ con la norma $\|\cdot\|_p$ per qualche $p \in [1,\infty]$. Sia $T : c_{00} \to c_{00}$ definito da

$$T\underline{x} = (2^{-k}x_k)_{k\in\mathbb{N}}$$

per ogni $\underline{x} = (x_k)_{k\in\mathbb{N}} \in c_{00}$. Non è difficile verificare che $T : c_{00} \to c_{00}$ è ben definito, lineare e limitato, con $\|T\underline{x}\|_p \leq \|\underline{x}\|_p$ per ogni $\underline{x} \in c_{00}$. Inoltre chiaramente T ha un inverso $T^{-1} : c_{00} \to c_{00}$ dato da

$$T^{-1}\underline{x} = (2^k x_k)_{k\in\mathbb{N}}.$$

Tuttavia, se $\underline{e^{(n)}}$ è definito come nell'esempio 6.34(1), allora

$$\|\underline{e^{(n)}}\|_p = 1, \qquad \|T^{-1}\underline{e^{(n)}}\|_p = \|2^n\underline{e^{(n)}}\|_p = 2^n$$

per ogni $n \in \mathbb{N}$, e quindi

$$\|T^{-1}\|_{\mathrm{op}} \geq \|T^{-1}\underline{e^{(n)}}\|_p / \|\underline{e^{(n)}}\|_p = 2^n \to \infty$$

per $n \to \infty$, cioè T^{-1} non è limitato.

(3) Sappiamo che c_{00} è denso in ℓ^2 (☞ proposizione 5.17(iii)); pertanto, per il teorema di estensione (☞ corollario 7.25), l'operatore T dell'esempio precedente ha un'unica estensione $\widetilde{T} \in \mathcal{B}(\ell^2)$, e non è difficile verificare che tale estensione è data da

$$\widetilde{T}\underline{x} = (2^{-k}x_k)_{k\in\mathbb{N}} \tag{7.28}$$

per ogni $\underline{x} \in \ell^2$. Osserviamo che $\widetilde{T}$ è iniettivo, dato che, per ogni $\underline{x} \in \ell^2$,

$$\underline{x} \in \operatorname{Ker}\widetilde{T} \iff 2^{-k}x_k = 0\ \forall k \in \mathbb{N} \iff x_k = 0\ \forall k \in \mathbb{N} \iff \underline{x} = \underline{0}.$$

In questo caso, tuttavia, $\widetilde{T} : \ell^2 \to \ell^2$ non è suriettivo: infatti, ad esempio, $(2^{-k})_{k\in\mathbb{N}} \in \ell^2 \setminus \operatorname{Im}\widetilde{T}$ (se esistesse $\underline{x} \in \ell^2$ tale che $\widetilde{T}\underline{x} = (2^{-k})_k$, da (7.28) dedurremmo che $\underline{x} = (1)_{k\in\mathbb{N}}$ è la successione costante 1, ma tale successione non è in ℓ^2, assurdo).

(4) Notiamo che l'operatore in (7.28) è un particolare operatore di moltiplicazione $D_{\underline{w}}$ su ℓ^2 (☞ esempio 7.8(8)). Generalizzando il ragionamento nell'esempio precedente, si può verificare che, per ogni $\underline{w} = (w_k)_{k\in\mathbb{N}} \in \ell^\infty$,

$$D_{\underline{w}} \text{ è iniettivo} \iff w_k \neq 0\ \forall k \in \mathbb{N},$$

mentre

$$D_{\underline{w}} \text{ è un isomorfismo} \iff \inf_{k\in\mathbb{N}} |w_k| > 0.$$

Nei precedenti esempi, abbiamo trovato un controesempio all'implicazione "T limitato e biiettivo $\Rightarrow T^{-1}$ limitato" lavorando con lo spazio c_{00}, che tuttavia non è completo rispetto alla norma $\|\cdot\|_p$ utilizzata. Questo rispecchia un fatto generale.

Teorema 7.36 (teorema dell'isomorfismo di Banach)**.** *Siano X e Y spazi di Banach e $T \in \mathcal{B}(X,Y)$. Se $T : X \to Y$ è biiettivo, allora $T^{-1} \in \mathcal{B}(Y,X)$, cioè T è un isomorfismo.*

Omettiamo la dimostrazione. Vediamo ora una conseguenza importante di questo risultato.

Teorema 7.37 (teorema del grafico chiuso)**.** *Siano X e Y spazi di Banach, e sia $T \in \mathcal{L}(X,Y)$. Allora $T \in \mathcal{B}(X,Y)$ se e solo se il grafico $\Gamma(T) := \{(x,Tx) : x \in X\}$ di T è un sottoinsieme chiuso di $X \times Y$.*

Dimostrazione. Sappiamo già (vedi proposizione 7.13(ii)) che, se $T \in \mathcal{B}(X,Y)$, allora $\Gamma(T)$ è chiuso in $X \times Y$. Rimane da dimostrare l'implicazione opposta.

Supponiamo dunque che $T \in \mathcal{L}(X,Y)$ e che $\Gamma(T)$ sia un sottoinsieme chiuso di $X \times Y$. Come nella dimostrazione della proposizione 7.13(ii) si vede che $\Gamma(T) = \operatorname{Ker} S$, dove $S \in \mathcal{L}(X \times Y, Y)$ è dato da $S(x,y) = y - Tx$ per ogni $(x,y) \in X \times Y$, pertanto $\Gamma(T)$ è un sottospazio vettoriale di $X \times Y$.

Sappiamo (☞ esempio 4.5(4)) che il prodotto $X \times Y$ è uno spazio di Banach, dato che X e Y lo sono. Siccome $\Gamma(T)$ è un sottospazio vettoriale chiuso di $X \times Y$, anche $\Gamma(T)$, con la norma indotta da $X \times Y$, è uno spazio di Banach (☞ proposizione 4.16).

Sia ora $R = \pi_X|_{\Gamma(T)} : \Gamma(T) \to X$ la restrizione della proiezione $\pi_X : X \times Y \to Y$. Siccome $\|\pi_X\|_{\mathrm{op}} \leq 1$, anche $\|R\|_{\mathrm{op}} \leq 1$, dunque $R \in \mathcal{B}(\Gamma(T), X)$.

Vediamo adesso che R è iniettivo. Infatti, se $(x,y) \in \operatorname{Ker} R$, allora $(x,y) \in \Gamma(T)$, pertanto $y = Tx$; inoltre $0 = R(x,y) = \pi_X(x,y) = x$, dunque anche $y = Tx = 0$, quindi $(x,y) = (0,0)$.

Infine, vediamo che $R : \Gamma(T) \to X$ è suriettivo. Se $x \in X$, allora $(x, Tx) \in \Gamma(T)$ e chiaramente $R(x,Tx) = \pi_X(x,Tx) = x$, dunque $x \in \operatorname{Im} R$.

In conclusione, $R : \Gamma(T) \to X$ è lineare, limitato e biiettivo, e $\Gamma(T)$ e X sono entrambi spazi di Banach. Dunque, per il teorema dell'isomorfismo (☞ teorema 7.36), anche $R^{-1} : X \to \Gamma(T)$ è limitato.

D'altra parte, come già osservato, per ogni $x \in X$ si ha $(x,Tx) \in \Gamma(T)$ e $R(x,Tx) = x$, pertanto $R^{-1}(x) = (x,Tx)$. Ma allora

$$\|Tx\|_Y \leq \|(x,Tx)\|_{X\times Y} = \|R^{-1}x\|_{\Gamma(T)} \leq \|R^{-1}\|_{\mathrm{op}}\|x\|_X$$

per ogni $x \in X$, da cui concludiamo che $\|T\|_{\mathrm{op}} \leq \|R^{-1}\|_{\mathrm{op}} < \infty$ e quindi $T \in \mathcal{B}(X,Y)$. □

Osservazione 7.38. Abbiamo presentato una dimostrazione del teorema del grafico chiuso basata sul teorema dell'isomorfismo. Viceversa, è possibile dimostrare il teorema dell'isomorfismo a partire dal teorema del grafico chiuso. Infatti, se $T \in \mathcal{B}(X,Y)$ è biiettivo, allora $T^{-1} : Y \to X$ è lineare, e inoltre, per ogni $(x,y) \in X \times Y$,

$$Tx = y \quad\iff\quad x = T^{-1}y,$$

cioè

$$(x,y) \in \Gamma(T) \quad\iff\quad (y,x) \in \Gamma(T^{-1}).$$

In altre parole, il grafico $\Gamma(T^{-1})$ di T^{-1} è l'immagine di $\Gamma(T)$ tramite l'omeomorfismo $X \times Y \ni (x,y) \to Y \times X$. Siccome T è limitato, $\Gamma(T)$ è chiuso in $X \times Y$ (☞ proposizione 7.13(ii)), dunque anche $\Gamma(T^{-1})$ è chiuso in $Y \times X$. Il teorema del grafico chiuso ci permette dunque di concludere che $T^{-1} \in \mathcal{B}(Y,X)$, come desiderato.

Vediamo ora un altro criterio di invertibilità per operatori limitati, che dimostriamo indipendentemente dal teorema dell'isomorfismo di Banach. Nel teorema dell'isomorfismo, per dimostrare che T^{-1} è limitato, assumiamo sia l'iniettività che la suriettività dell'operatore T; nel prossimo risultato, indeboliamo l'ipotesi di suriettività, ma rafforziamo (in senso quantitativo) quella di iniettività.

Teorema 7.39. *Siano X uno spazio di Banach e Y uno spazio normato. Sia $T \in \mathcal{B}(X,Y)$. Allora T è un isomorfismo se e solo se valgono <u>entrambe</u> le seguenti proprietà:*

(a) $\overline{\operatorname{Im} T} = Y$
(T ha immagine densa);

(b) esiste $C \in (0,\infty)$ tale che $\|x\|_X \leq C\|Tx\|_Y$ per ogni $x \in X$
(T è coercivo in norma*).*

Alla dimostrazione premettiamo un lemma.

Lemma 7.40. *Siano X uno spazio di Banach e Y uno spazio normato. Sia $T \in \mathcal{B}(X,Y)$. Se T è* coercivo in norma*, cioè esiste $C \in (0,\infty)$ tale che*

$$\|x\|_X \leq C\|Tx\|_Y$$

per ogni $x \in X$, allora $\operatorname{Im} T$ è chiusa in Y.

Dimostrazione. Sia $y \in \overline{\operatorname{Im} T}$. Allora esiste una successione $(x_n)_n$ a valori in X tale che $Tx_n \to y$ in Y. In particolare, $(Tx_n)_n$ è di Cauchy in Y. Applicando la disuguaglianza di coercività ai vettori $x_n - x_m$ si ottiene

$$\|x_n - x_m\|_X \leq C\|Tx_n - Tx_m\|_Y$$

per ogni $n, m \in \mathbb{N}$, da cui si deduce che anche $(x_n)_n$ è di Cauchy in X. Siccome X è di Banach, esiste $x \in X$ tale che $x_n \to x$ in X. Ma allora, per continuità di T,

$$Tx = \lim_{n\to\infty} Tx_n = y$$

e quindi $y \in \operatorname{Im} T$. □

Dimostrazione del teorema 7.39. Se $T \in \mathcal{B}(X, Y)$ è un isomorfismo, allora T è suriettivo, dunque (a) chiaramente vale. Inoltre, siccome $T^{-1} : Y \to X$ è limitato, si ha

$$\|x\|_X = \|T^{-1}Tx\|_X \leq \|T^{-1}\|_{\mathrm{op}}\|Tx\|_Y$$

per ogni $x \in X$, e possiamo prendere $C = 1 + \|T^{-1}\|_{\mathrm{op}}$ per verificare (b).

Viceversa, sia $T \in \mathcal{B}(X, Y)$ un operatore che soddisfa le condizioni (a) e (b). Allora, da (b) e dal lemma 7.40, deduciamo che $\operatorname{Im} T$ è chiusa in Y, ma allora da (a) segue che $\operatorname{Im} T = Y$, cioè T è suriettivo. Inoltre, sempre da (b), è chiaro che T è iniettivo (se $Tx = 0$ allora $\|x\|_X \leq C\|Tx\|_Y = 0$ e quindi $x = 0$). Pertanto $T : X \to Y$ è biiettivo e l'inverso $T^{-1} : Y \to X$ è lineare. Inoltre, per ogni $y \in Y$, da (b) deduciamo che

$$\|T^{-1}y\|_X \leq C\|TT^{-1}y\|_Y = \|y\|_Y,$$

cioè $\|T^{-1}\|_{\mathrm{op}} \leq C < \infty$ e $T^{-1} \in \mathcal{B}(Y, X)$, come desiderato. □

Osservazione 7.41. Nell'esempio 7.35(3), l'operatore $\widetilde{T} \in \mathcal{B}(\ell^2)$ ha immagine densa in ℓ^2, dato che $\operatorname{Im} \widetilde{T} \supseteq \operatorname{Im} T = c_{00}$. Tuttavia, per ogni $n \in \mathbb{N}$,

$$\|\underline{e}^{(n)}\|_2 = 1, \qquad \|T\underline{e}^{(n)}\|_2 = \|2^{-n}\underline{e}^{(n)}\|_2 = 2^{-n};$$

quindi non può esistere $C \in (0, \infty)$ tale che $\|\underline{x}\|_2 \leq C\|T\underline{x}\|_2$, cioè T non è coercivo rispetto alla norma (nonostante $\widetilde{T}$ sia iniettivo).

Vediamo ora un importante criterio di invertibilità per operatori limitati da uno spazio di Banach in sé.

Teorema 7.42 (criterio di Neumann)**.** *Sia X uno spazio di Banach. Sia $T \in \mathcal{B}(X)$ tale che $\|T\|_{\mathrm{op}} < 1$. Allora $\mathrm{id}_X - T$ è un isomorfismo e*

$$(\mathrm{id}_X - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

dove la serie converge in $\mathcal{B}(X)$. Inoltre

$$\|(\mathrm{id}_X - T)^{-1}\|_{\mathrm{op}} \leq \frac{1}{1 - \|T\|_{\mathrm{op}}}.$$

Dimostrazione. Osserviamo che, per la submoltiplicatività della norma operatoriale,

$$\sum_{n=0}^{\infty} \|T^n\|_{\mathrm{op}} \leq \sum_{n=0}^{\infty} \|T\|_{\mathrm{op}}^n = \frac{1}{1-\|T\|_{\mathrm{op}}} < \infty, \tag{7.29}$$

siccome $\|T\|_{\mathrm{op}} < 1$, dove si è usata la formula per la somma di una serie geometrica; dunque la serie di operatori $\sum_{n=0}^{\infty} T_n$ converge assolutamente in $\mathcal{B}(X)$. Siccome X è di Banach, anche $\mathcal{B}(X)$ lo è (☞ proposizione 7.14), pertanto la serie $\sum_{n=0}^{\infty} T_n$ converge in $\mathcal{B}(X)$ a un operatore $S \in \mathcal{B}(X)$ (☞ proposizione 4.20), e si ha

$$\|S\|_{\mathrm{op}} \leq \sum_{n=0}^{\infty} \|T^n\|_{\mathrm{op}} \leq \frac{1}{1-\|T\|_{\mathrm{op}}}$$

per (7.29).

Per concludere, è dunque sufficiente dimostrare che S è l'inverso di $\mathrm{id}_X - T$. Infatti, per ogni $N \in \mathbb{N}$, si ha

$$(\mathrm{id}_X - T) \sum_{n=0}^{N} T^n = \sum_{n=0}^{N} T^n (\mathrm{id}_X - T) = \sum_{n=0}^{N} T^n - \sum_{n=0}^{N} T^{n+1} = \mathrm{id}_X - T^{N+1}. \tag{7.30}$$

Inoltre, $\|T^n\|_{\mathrm{op}} \leq \|T\|_{\mathrm{op}}^n \to 0$ per $n \to \infty$, dato che $\|T\|_{\mathrm{op}} < 1$ per ipotesi, pertanto

$$T^n \to 0 \quad \text{per } n \to \infty.$$

Passando al limite per $N \to \infty$ in (7.30) e usando la continuità delle operazioni (☞ proposizione 4.6(ii) e corollario 7.12), si ottiene allora

$$(\mathrm{id}_X - T)S = S(\mathrm{id}_X - T) = \mathrm{id}_X,$$

il che dimostra che $S = (\mathrm{id}_X - T)^{-1}$, come desiderato. □

Corollario 7.43. *Sia X uno spazio di Banach.*

(i) Se $R \in \mathcal{B}(X)$ è un isomorfismo, $T \in \mathcal{B}(X)$ e

$$\|R - T\|_{\mathrm{op}} < 1/\|R^{-1}\|_{\mathrm{op}}, \tag{7.31}$$

allora T è un isomorfismo.

(ii) L'insieme

$$\mathcal{I}(X) := \{T \in \mathcal{B}(X) : T \text{ è un isomorfismo}\}$$

è un sottoinsieme aperto di $\mathcal{B}(X)$.

Dimostrazione. (i). Scriviamo

$$T = R - (R - T) = R(\mathrm{id}_X - R^{-1}(R - T)). \tag{7.32}$$

Siccome, per (7.31), si ha

$$\|R^{-1}(R - T)\|_{\mathrm{op}} \leq \|R^{-1}\|_{\mathrm{op}} \|R - T\|_{\mathrm{op}} < 1,$$

il criterio di Neumann (☞ teorema 7.42) ci garantisce che $\mathrm{id}_X - R^{-1}(R-T)$ è un isomorfismo. Siccome anche R lo è, da (7.32) vediamo che T è la composizione di due isomorfismi, quindi anche T è un isomorfismo (☞ osservazione 7.33(3)).

(ii). Se $X = \{0\}$ la tesi è ovvia (si ha $\mathcal{I}(X) = \mathcal{B}(X) = \{0\}$). Supponiamo dunque $X \neq \{0\}$. Allora da (ii) deduciamo che, per ogni $R \in \mathcal{I}(X)$, anche $B(R, 1/\|R^{-1}\|_{\mathrm{op}}) \subseteq \mathcal{I}(X)$, dunque $\mathcal{I}(X)$ è aperto in $\mathcal{B}(X)$. □

Osservazione 7.44. Nel caso $\dim X < \infty$, il risultato del corollario 7.43(ii) si può dimostrare in maniera più diretta. Si ha infatti

$$\mathcal{I}(X) = \{T \in \mathcal{B}(X) : \det T \neq 0\}.$$

Siccome $\det : \mathcal{B}(X) \to \mathbb{F}$ è continuo (il determinante è un polinomio nei coefficienti della matrice che rappresenta T in una fissata base di X), allora $\mathcal{I}(X)$ è la controimmagine dell'aperto $\mathbb{F} \setminus \{0\}$ tramite una funzione continua, quindi è aperto. Quando $\dim X = \infty$, tuttavia, non ha più senso parlare di "determinante" di un elemento $T \in \mathcal{B}(X)$; la dimostrazione del corollario 7.43(ii) tramite il criterio di Neumann dà dunque un approccio alternativo, che funziona anche in dimensione infinita.

8 Duale di uno spazio normato

8.1 Funzionali lineari e spazio duale

Ricordiamo che $\mathbb{F}$ denota il campo $\mathbb{R}$ dei numeri reali oppure il campo $\mathbb{C}$ dei numeri complessi.

Definizione 8.1. Sia X uno spazio vettoriale su $\mathbb{F}$. Un *funzionale lineare* su X è una mappa lineare $\varphi : X \to \mathbb{F}$. Lo spazio vettoriale $\mathcal{L}(X, \mathbb{F})$ dei funzionali lineari su X è detto *duale algebrico* di X.

Definizione 8.2. Sia X uno spazio normato su $\mathbb{F}$. Lo *spazio duale* (o *duale topologico*) di X è lo spazio $X' \coloneqq \mathcal{B}(X, \mathbb{F})$ dei funzionali lineari continui su X.

Osservazione 8.3. Se X è uno spazio normato, il duale $X' = \mathcal{B}(X, \mathbb{F})$ è a sua volta uno spazio normato con la norma operatoriale (☞ proposizione 7.10). In particolare, per definizione,

$$\|\varphi\|_{X'} = \|\varphi\|_{\mathrm{op}} = \inf\{C \in [0, \infty) : |\varphi(x)| \leq C\|x\|_X\}$$

per ogni $\varphi \in X'$. In effetti, siccome $\mathbb{F}$ è uno spazio di Banach (☞ esempio 2.12(2)), il duale X' è uno spazio di Banach per ogni spazio normato X (☞ proposizione 7.14).

Osservazione 8.4. Se X è uno spazio normato con $\dim X < \infty$, allora $\mathcal{L}(X, \mathbb{F}) = \mathcal{B}(X, \mathbb{F}) = X'$ (☞ proposizione 7.9), cioè tutti i funzionali lineari su X sono continui, e dunque il duale algebrico coincide con il duale topologico. Sappiamo tuttavia (☞ esempio 7.8(3)) che, quando $\dim X = \infty$, possono esistere funzionali lineari su X che non sono continui.

Esempio 8.5. Ecco alcuni esempi di funzionali lineari continui.

(1) Sia X lo spazio $(C(M), \|\cdot\|_\infty)$ dove M è uno spazio metrico compatto (☞ esempio 4.5(2)). Per ogni $p \in M$, l'operatore di valutazione $V_p : f \mapsto f(p)$ (☞ esempio 7.8(2)) è un funzionale lineare continuo su $C(M)$, cioè $V_p \in C(M)'$, e inoltre
$$\|V_p\|_{C(M)'} = \|V_p\|_{\mathrm{op}} = 1.$$

(2) Sia H uno spazio pre-hilbertiano. Per ogni $y \in H$, l'operatore $\langle \cdot, y\rangle : H \to \mathbb{F}$ è un funzionale lineare continuo su H (☞ esempio 7.8(4)), cioè $\langle \cdot, y\rangle \in H'$, e inoltre
$$\|\langle \cdot, y\rangle\|_{H'} = \|\langle \cdot, y\rangle\|_{\mathrm{op}} = \|y\|_H.$$

(3) Sia $X = \ell^p$ per qualche $p \in [1, \infty]$. Sia $q \in [1, \infty]$ l'esponente coniugato a p. Allora, per ogni $\underline{y} \in \ell^q$, possiamo definire $\varphi_{\underline{y}} : \ell^p \to \mathbb{F}$ ponendo
$$\varphi_{\underline{y}}(\underline{x}) = \sum_{n=0}^{\infty} x_n y_n \tag{8.1}$$
per ogni $\underline{x} \in \ell^p$. Infatti, per la disuguaglianza di Hölder (☞ proposizione 5.14), la serie a membro destro di (8.1) converge assolutamente per ogni

$\underline{x} \in \ell^p$, pertanto $\varphi_{\underline{y}} : \ell^p \to \mathbb{F}$ è ben definito. Inoltre, è facile verificare che $\varphi_{\underline{y}} : \ell^p \to \mathbb{F}$ è lineare e, sempre per la disuguaglianza di Hölder,

$$\left|\varphi_{\underline{y}}(\underline{x})\right| \leq \|\underline{y}\|_{\ell^q} \|\underline{x}\|_{\ell^p}$$

per ogni $\underline{x} \in \ell^p$; da questo segue che $\varphi_{\underline{y}} : \ell^p \to \mathbb{F}$ è limitato, cioè $\varphi_{\underline{y}} \in (\ell^p)'$, e inoltre

$$\|\varphi_{\underline{y}}\|_{(\ell^p)'} \leq \|\underline{y}\|_{\ell^q}.$$

(4) Più in generale, sia $X = L^p(M, \mathcal{M}, \mu)$ per qualche spazio di misura $(M, \mathcal{M}, \mu)$, e siano $p, q \in [1, \infty]$ esponenti coniugati. Allora, per ogni $g \in L^q(M, \mathcal{M}, \mu)$ possiamo definire un funzionale $\varphi_g \in L^p(M, \mathcal{M}, \mu)'$ ponendo

$$\varphi_g(f) = \int_M fg \, d\mu \tag{8.2}$$

per ogni $f \in L^p(M, \mathcal{M}, \mu)$, e per la disuguaglianza di Hölder si ha

$$\|\varphi_g\|_{L^p(M,\mathcal{M},\mu)'} \leq \|g\|_{L^q(M,\mathcal{M},\mu)}.$$

Un problema rilevante in molte applicazioni è caratterizzare il duale X' di un dato spazio normato X, cioè capire quale forma possono avere i funzionali lineari continui su X. Ad esempio, ci si può porre il problema se, per gli spazi normati considerati nell'esempio 8.5, i funzionali lineari continui ivi descritti esauriscano gli elementi dei rispettivi duali, o se ce ne siano altri. Nel seguito discuteremo questo problema nel caso degli spazi di Hilbert e degli spazi ℓ^p ed L^p, espandendo la discussione degli esempi 8.5(2)-(3)-(4) sopra.

Osservazione 8.6. Nel caso $X = \mathbb{F}^n$ con la norma euclidea, gli elementi di $X' = \mathcal{B}(\mathbb{F}^n, \mathbb{F}) = \mathcal{L}(\mathbb{F}^n, \mathbb{F})$ si possono rappresentare con matrici $1 \times n$, mentre gli elementi di $\mathbb{F}^n$ si pensano come vettori colonna, cioè matrici $n \times 1$. In questo caso, dunque, possiamo stabilire un isomorfismo tra $\mathbb{F}^n$ e $(\mathbb{F}^n)'$, che corrisponde a "trasporre" la matrice (passare da un vettore colonna a un vettore riga); invertendo l'isomorfismo, otteniamo dunque una "rappresentazione concreta" degli elementi del duale $(\mathbb{F}^n)'$ come vettori di $\mathbb{F}^n$. Le caratterizzazioni dei duali che discuteremo in seguito si possono pensare come tentativi di generalizzare questo risultato al caso di spazi di dimensione infinita; tuttavia, come vedremo, se $\dim X = \infty$, non è sempre vero che X' è isomorfo a X.

8.2 Duale di uno spazio di Hilbert

Teorema 8.7 (teorema di rappresentazione di Riesz–Fréchet)**.** *Sia H uno spazio di Hilbert su $\mathbb{F}$. Allora la mappa $\Phi : H \to H'$, definita da*

$$\Phi(y) = \langle \cdot, y \rangle \tag{8.3}$$

per ogni $y \in H$, è una isometria antilineare suriettiva*, cioè:*

(a) $\|\Phi(y)\|_{H'} = \|y\|_H$ *per ogni $y \in H$*
(Φ è un'isometria)*;*

(b) $\Phi(\alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1}\Phi(y_1) + \overline{\alpha_2}\Phi(y_2)$ *per ogni $y_1, y_2 \in H$ e $\alpha_1, \alpha_2 \in \mathbb{F}$*
(Φ è antilineare)*;*

(c) $\Phi(H) = H'$
(Φ è suriettiva).

Osservazione 8.8. Nel caso $\mathbb{F} = \mathbb{R}$, si ha $\overline{\alpha} = \alpha$, dunque la mappa $\Phi : H \to H'$ del teorema di Riesz–Fréchet è lineare e quindi è un *isomorfismo isometrico* nel senso della definizione 7.19. Nel caso $\mathbb{F} = \mathbb{C}$, tuttavia, $\Phi : H \to H'$ non è lineare (è antilineare) e quindi Φ non è un isomorfismo isometrico; diciamo invece che Φ è un *anti-isomorfismo isometrico*.

Dimostrazione del teorema 8.7. Dall'esempio 8.5(2) sappiamo già che la mappa $\Phi : H \to H'$ data da (8.3) è ben definita, e che inoltre vale (a).

La proprietà (b) è conseguenza dell'antilinearità del prodotto scalare $\langle \cdot, \cdot \rangle$ nella seconda variabile: per ogni $y_1, y_2, \in H$ e $\alpha_1, \alpha_2 \in \mathbb{F}$, si ha

$$\begin{aligned}\Phi(\alpha_1 y_1 + \alpha_2 y_2)(x) &= \langle x, \alpha_1 y_1 + \alpha_2 y_2 \rangle \\ &= \overline{\alpha_1} \langle x, y_1 \rangle + \overline{\alpha_2} \langle x, y_2 \rangle = \overline{\alpha_1} \Phi(y_1)(x) + \overline{\alpha_2} \Phi(y_2)(x)\end{aligned}$$

per ogni $x \in H$, da cui segue (b).

Rimane solo da dimostrare (c), cioè che $\Phi : H \to H'$ è suriettiva. Sia dunque $\varphi \in H'$. Se $\varphi = 0$, allora ovviamente $\varphi = \Phi(0) \in \Phi(H)$. Supponiamo invece che $\varphi \neq 0$. Allora $Y := \operatorname{Ker} \varphi$ è un sottospazio vettoriale chiuso proprio di H (☞ proposizione 7.13(i)), e $H = Y \oplus Y^\perp$ (☞ corollario 6.33), quindi $Y^\perp \neq \{0\}$.

Scegliamo allora $z \in Y^\perp \setminus \{0\}$; siccome $z \notin Y = \operatorname{Ker} \varphi$, si ha $\varphi(z) \neq 0$; a meno di sostituire z con $z/\varphi(z)$, possiamo assumere che $\varphi(z) = 1$.

Allora, per ogni $x \in Y^\perp$,

$$\varphi(x - \varphi(x)z) = \varphi(x) - \varphi(x)\varphi(z) = \varphi(x) - \varphi(x) = 0,$$

cioè $x - \varphi(x)z \in \operatorname{Ker} \varphi = Y$; siccome tuttavia $x, z \in Y^\perp$, si ha anche $x - \varphi(x)z \in Y^\perp$; dato che $Y \cap Y^\perp = \{0\}$ (☞ proposizione 6.26(ii)) ne deduciamo che $x - \varphi(x)z = 0$, cioè

$$x = \varphi(x)z \qquad \forall x \in Y^\perp.$$

Dunque $Y^\perp = \operatorname{span}\{z\}$, quindi $\{z/\|z\|\}$ è una base ortonormale di $Y^\perp$ e

$$P_{Y^\perp} x = \|z\|^{-2} \langle x, z \rangle z \tag{8.4}$$

per ogni $x \in H$ (☞ proposizione 6.39).

In particolare, per ogni $x \in H$, si ha $x = P_Y x + P_{Y^\perp} x$ (☞ corollario 6.33), e inoltre $\varphi(P_Y x) = 0$ (perché $Y = \operatorname{Ker} \varphi$); dunque, per (8.4),

$$\varphi(x) = \varphi(P_Y x) + \varphi(P_{Y^\perp} x) = 0 + \|z\|^{-2} \langle x, z \rangle \varphi(z) = \langle x, \|z\|^{-2} z \rangle,$$

dato che $\varphi(z) = 1$. Questo dimostra che $\varphi = \Phi(\|z\|^{-2} z) \in \Phi(H)$. □

Osservazione 8.9. Nel caso $H = L^2(M, \mathcal{M}, \mu)$, l'anti-isomorfismo isometrico $\Phi : L^2(M) \to L^2(M)'$ dato dal teorema di Riesz–Fréchet si scrive più concretamente come

$$\Phi(g)(f) = \langle f, g \rangle_{L^2} = \int_M f \, \overline{g} \, d\mu$$

per ogni $f, g \in L^2(M)$. Se allora definiamo $\Psi : L^2(M) \to L^2(M)'$ ponendo $\Psi(g) = \Phi(\overline{g})$, cioè

$$\Psi(g)(f) = \int_M f\, g\, d\mu, \tag{8.5}$$

si ha che $\Psi : L^2(M) \to L^2(M)'$ è lineare, e quindi Ψ è un isomorfismo isometrico tra $L^2(M)$ e il suo duale. Abbiamo dunque ottenuto, come conseguenza del teorema di Riesz–Fréchet, che

$$L^2(M, \mathcal{M}, \mu)' \underset{\text{isom.}}{\cong} L^2(M, \mathcal{M}, \mu)$$

e in particolare

$$(\ell^2)' \underset{\text{isom.}}{\cong} \ell^2$$

(☞ definizione 7.19). Nella prossima sezione, investigheremo risultati analoghi per ℓ^p e L^p con $p \neq 2$. Come già accennato negli esempi 8.5(3)-(4), vedremo che la stessa formula (8.5) qui usata per L^2 può essere anche usata per definire un "accoppiamento di dualità" fra L^p e L^q, ove p e q sono esponenti coniugati.

8.3 Duale degli spazi ℓ^p e L^p

Teorema 8.10. *Siano $p, q \in [1, \infty]$ esponenti coniugati ($1/p + 1/q = 1$). Sia $\Psi : \ell^q \to (\ell^p)'$ definita da*

$$\Psi(\underline{y}) = \varphi_{\underline{y}}$$

per ogni $\underline{y} \in \ell^q$, dove $\varphi_{\underline{y}} \in (\ell^p)'$ è definito in (8.1); in altre parole,

$$\Psi(\underline{y})(\underline{x}) = \sum_{k=0}^{\infty} x_k y_k \tag{8.6}$$

per ogni $\underline{y} \in \ell^q$ e $\underline{x} \in \ell^p$. Allora:

(i) $\Psi : \ell^q \to (\ell^p)'$ è un'isometria lineare.

(ii) Se $p \neq \infty$, allora $\Psi : \ell^q \to (\ell^p)'$ è un isomorfismo isometrico.

Dimostrazione. Sappiamo già dall'esempio 8.5(3) che, per la disuguaglianza di Hölder,

$$\|\Psi(\underline{y})\|_{(\ell^p)'} \leq \|\underline{y}\|_{\ell^q} \tag{8.7}$$

per ogni $\underline{y} \in \ell^q$.

Inoltre, è facile verificare che $\Psi : \ell^q \to (\ell^p)'$ è lineare: se $\underline{y}, \underline{y'} \in \ell^q$ e $\alpha, \alpha' \in \mathbb{F}$, allora

$$\begin{aligned}\Psi(\alpha\underline{y} + \alpha'\underline{y'})(\underline{x}) &= \sum_{k=0}^{\infty} x_k (y_k + \alpha' y'_k) \\ &= \alpha \sum_{k=0}^{\infty} x_k y_k + \alpha' \sum_{k=0}^{\infty} x_k y'_k = \alpha\Psi(\underline{y})(\underline{x}) + \alpha'\Psi(\underline{y'})(\underline{x})\end{aligned}$$

per ogni $\underline{x} \in \ell^p$, dunque

$$\Psi(\alpha\underline{y} + \alpha'\underline{y'}) = \alpha\Psi(\underline{y}) + \alpha'\Psi(\underline{y'}).$$

Consideriamo anzitutto il caso $p = \infty$, cioè $q = 1$. Preso $\underline{y} \in \ell^1$ e definito $\underline{\tilde{y}} \in \mathbb{F}^{\mathbb{N}}$ ponendo, per ogni $k \in \mathbb{N}$,

$$\tilde{y}_k = \begin{cases} |y_k|/y_k & \text{se } y_k \neq 0, \\ 0 & \text{altrimenti,} \end{cases}$$

si ha chiaramente $\|\tilde{y}\|_\infty \leq 1$; dunque $\tilde{y} \in \ell^\infty$ e

$$\varphi_{\underline{y}}(\underline{\tilde{y}}) = \sum_{k=0}^{\infty} \tilde{y}_k y_k = \sum_{k=0}^{\infty} |y_k| = \|\underline{y}\|_1,$$

da cui segue che

$$\|\underline{y}\|_1 = |\varphi_{\underline{y}}(\underline{\tilde{y}})| \leq \|\varphi_{\underline{y}}\|_{(\ell^\infty)'} \|\underline{\tilde{y}}\|_\infty \leq \|\varphi_{\underline{y}}\|_{(\ell^\infty)'}.$$

Insieme a (8.7), questa disuguaglianza mostra che in effetti

$$\|\underline{y}\|_1 = \|\varphi_{\underline{y}}\|_{(\ell^\infty)'} = \|\Psi(\underline{y})\|_{(\ell^\infty)'}$$

per ogni $\underline{y} \in \ell^1$, cioè $\Phi : \ell^1 \to (\ell^\infty)'$ è un'isometria lineare. Questo conclude la dimostrazione del teorema nel caso $p = \infty$.

Supponiamo ora che $p < \infty$. Sia $\varphi \in (\ell^p)'$. Definiamo $\underline{y} \in \mathbb{F}^{\mathbb{N}}$ ponendo, per ogni $k \in \mathbb{N}$,

$$y_k = \varphi(\underline{e^{(k)}}), \tag{8.8}$$

ove $\underline{e^{(k)}} \in c_{00} \subseteq \ell^p$ è definito nell'esempio 6.34(1). Mostriamo ora che $\underline{y} \in \ell^q$.

Se $p = 1$, cioè $q = \infty$, allora

$$|y_k| = |\varphi(\underline{e^{(k)}})| \leq \|\varphi\|_{(\ell^1)'} \|\underline{e^{(k)}}\|_\infty = \|\varphi\|_{(\ell^1)'}$$

per ogni $k \in \mathbb{N}$, pertanto $\|\underline{y}\|_\infty \leq \|\varphi\|_{(\ell^1)'} < \infty$ e $\underline{y} \in \ell^\infty$.

Se invece $1 < p < \infty$, cioè $1 < q < \infty$, definita $\underline{\tilde{y}} \in \mathbb{F}^{\mathbb{N}}$ ponendo, per ogni $k \in \mathbb{N}$,

$$\tilde{y}_k = \begin{cases} |y_k|^q/y_k & \text{se } y_k \neq 0, \\ 0 & \text{altrimenti,} \end{cases}$$

e definita, per ogni $n \in \mathbb{N}$, la troncata n-esima $\underline{z^{(n)}} \in c_{00}$ di $\underline{\tilde{y}}$, cioè

$$\underline{z^{(n)}} = (\tilde{y}_0, \tilde{y}_1, \dots, \tilde{y}_n, 0, 0, 0, \dots) = \sum_{k=0}^{n} \tilde{y}_k \underline{e^{(k)}},$$

si ha che

$$\varphi(\underline{z^{(n)}}) = \sum_{k=0}^{n} \tilde{y}_k \varphi(\underline{e^{(k)}}) = \sum_{k=0}^{n} \tilde{y}_k y_k = \sum_{k=0}^{n} |y_k|^q.$$

Pertanto

$$\sum_{k=0}^{n} |y_k|^q = |\varphi(\underline{z^{(n)}})| \leq \|\varphi\|_{(\ell^p)'} \|\underline{z^{(n)}}\|_p = \|\varphi\|_{(\ell^p)'} \left(\sum_{k=0}^{n} |\tilde{y}_k|^p \right)^{1/p}.$$

Siccome $q > 1$, si ha $|\tilde{y}_k| = |y_k|^{q-1}$ per ogni $k \in \mathbb{N}$, e inoltre $(q-1)p = q$ siccome p, q sono coniugati; pertanto $|\tilde{y}_k|^p = |y_k|^q$ e

$$\sum_{k=0}^{n} |y_k|^q \leq \|\varphi\|_{(\ell^p)'} \left(\sum_{k=0}^{n} |\tilde{y}_k|^p \right)^{1/p} = \|\varphi\|_{(\ell^p)'} \left(\sum_{k=0}^{n} |y_k|^q \right)^{1/p}.$$

Quando $\sum_{k=0}^{n} |y_k|^q \neq 0$, dividendo ambo i membri per $(\sum_{k=0}^{n} |y_k|^q)^{1/p}$ si ottiene

$$\left(\sum_{k=0}^{n} |y_k|^q \right)^{1/q} \leq \|\varphi\|_{(\ell^p)'};$$

la stessa disuguaglianza vale banalmente anche quando $\sum_{k=0}^{n} |y_k|^q = 0$. Infine, prendendo il limite per $n \to \infty$, ne deduciamo

$$\|\underline{y}\|_q \leq \|\varphi\|_{(\ell^p)'} < \infty$$

e quindi $\underline{y} \in \ell^q$.

In ogni caso, per $p < \infty$, abbiamo $\underline{y} \in \ell^q$ e

$$\|\underline{y}\|_{\ell^q} \leq \|\varphi\|_{(\ell^p)'}. \tag{8.9}$$

Possiamo allora considerare il funzionale $\varphi_{\underline{y}} = \Psi(\underline{y}) \in (\ell^p)'$, e osservare che

$$\varphi_{\underline{y}}(\underline{e^{(n)}}) = \sum_{n=0}^{\infty} y_k \delta_{k,n} = y_n = \varphi(\underline{e^{(n)}})$$

per ogni $n \in \mathbb{N}$, per definizione di $\underline{y}$. Questo dimostra che $\underline{e^{(n)}} \in \operatorname{Ker}(\varphi_{\underline{y}} - \varphi)$ per ogni $n \in \mathbb{N}$; ma allora, siccome il nucleo di una mappa lineare è un sottospazio vettoriale,

$$\operatorname{Ker}(\varphi_{\underline{y}} - \varphi) \supseteq \operatorname{span}\{\underline{e^{(n)}}\}_{n \in \mathbb{N}} = c_{00},$$

e siccome $\operatorname{Ker}(\varphi_{\underline{y}} - \varphi)$ è chiuso (☞ proposizione 7.13(i)) e c_{00} è denso in ℓ^p (☞ proposizione 5.17(iii)) ne concludiamo che $\operatorname{Ker}(\varphi_{\underline{y}} - \varphi) = \ell^p$, cioè

$$\varphi = \varphi_{\underline{y}} = \Psi(\underline{y}) \in \operatorname{Im} \Psi.$$

Per l'arbitrarietà di $\varphi \in (\ell^p)'$, ciò mostra che $\Psi : \ell^q \to (\ell^p)'$ è suriettiva.

In aggiunta, per ogni $\underline{y} \in \ell^q$, se ripetiamo l'argomento precedente applicandolo al funzionale $\varphi = \varphi_{\underline{y}} \in (\ell^p)'$, la disuguaglianza (8.9) ci dà

$$\|\underline{y}\|_{\ell^q} \leq \|\varphi_{\underline{y}}\|_{(\ell^p)'};$$

combinandola con (8.7), ne deduciamo allora che

$$\|\underline{y}\|_{\ell^q} = \|\varphi_{\underline{y}}\|_{(\ell^p)'} = \|\Psi(\underline{y})\|_{(\ell^p)'}$$

per ogni $\underline{y} \in \ell^q$, cioè $\Psi : \ell^q \to (\ell^p)'$ è un'isometria. □

Osservazione 8.11. Nel caso $p = \infty$, dal teorema precedente sappiamo solo che $\Psi(\ell^1)$ è un sottospazio vettoriale chiuso di $(\ell^\infty)'$; il fatto che $\Psi(\ell^\infty)$ sia chiuso segue dal lemma 7.40. Vedremo in seguito (come conseguenza del teorema di Hahn–Banach, ☞ corollario 8.44) che tale sottospazio è proprio, cioè

$$\Psi(\ell^1) \neq (\ell^\infty)'.$$

Osservazione 8.12. In effetti, se si tentasse di ripetere nel caso $p = \infty$ l'argomento della dimostrazione del teorema 8.10 per la suriettività di Ψ, dato $\varphi \in (\ell^\infty)'$ si potrebbe costruire $\underline{y} \in \mathbb{F}^\mathbb{N}$ come in (8.8), e ottenere come sopra

$$\|\underline{y}\|_1 \leq \|\varphi\|_{(\ell^\infty)'}, \qquad \varphi|_{c_{00}} = \varphi_{\underline{y}}|_{c_{00}}.$$

Tuttavia, a questo punto l'argomento "si inceppa", dato che c_{00} non è denso in ℓ^∞; infatti $\overline{c_{00}}^{\ell^\infty} = c_0$ (☞ proposizione 5.6), quindi non possiamo concludere che $\varphi = \varphi_{\underline{y}}$, ma solo che

$$\varphi|_{c_0} = \varphi_{\underline{y}}|_{c_0}.$$

Questa osservazione "negativa" può essere trasformata in un risultato "positivo", relativamente alla caratterizzazione del duale di c_0.

Teorema 8.13. *La mappa $\Psi : \ell^1 \to (c_0)'$, definita da*

$$\Psi(\underline{y}) = \varphi_{\underline{y}}|_{c_0}$$

per ogni $\underline{y} \in \ell^1$, ovverosia

$$\Psi(\underline{y})(\underline{x}) = \sum_{k=0}^{\infty} x_k y_k \tag{8.10}$$

per ogni $\underline{y} \in \ell^1$ e $\underline{x} \in c_0$, è un isomorfismo isometrico da ℓ^1 a $(c_0)'$.

Dimostrazione. La dimostrazione è una variante di quella del teorema 8.10, seguendo quanto indicato nell'osservazione 8.12; i dettagli sono lasciati per esercizio. □

Osservazione 8.14. I risultati precedenti sono spesso ricordati per brevità nella forma

"il duale di ℓ^p è ℓ^q (se $p < \infty$)",
"il duale di c_0 è ℓ^1".

Se da una parte questo è un buon modo per tenerli a mente, dall'altra è essenziale tenere presente che le identificazioni

$$(\ell^p)' \underset{\text{isom.}}{\cong} \ell^q, \qquad (c_0)' \underset{\text{isom.}}{\cong} \ell^1$$

passano attraverso uno specifico isomorfismo isometrico Ψ (la cui definizione è parte integrante degli enunciati sopra, ☞ formule (8.6) e (8.10)).

Discutiamo ora (senza dimostrazioni) risultati analoghi sulla caratterizzazione del duale di $L^p(M, \mathcal{M}, \mu)$ per più generali spazi di misura $(M, \mathcal{M}, \mu)$. Ricordiamo (☞ definizione 3.1(f)) che uno spazio di misura $(M, \mathcal{M}, \mu)$ si dice σ-finito se M è unione numerabile di sottoinsiemi misurabili di misura finita.

Teorema 8.15. *Siano $p, q \in [1, \infty]$ esponenti coniugati ($1/p + 1/q = 1$). Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Sia $\Psi : L^q(M) \to L^p(M)'$ definita da*

$$\Psi(g) = \varphi_g$$

per ogni $g \in L^q(M)$, dove $\varphi_g \in (L^p(M))'$ è dato da (8.2); in altre parole,

$$\Psi(g)(f) = \int_M f\, g\, d\mu \tag{8.11}$$

per ogni $g \in L^q(M)$ e $f \in L^p(M)$. Allora $\Psi : L^q(M) \to L^p(M)'$ è un'isometria lineare; se poi $p < \infty$, allora Ψ è un isomorfismo isometrico.

Osservazione 8.16. Per esponenti coniugati p e q, abbiamo dunque, più in generale,

$$(L^p(M))' \underset{\text{isom.}}{\cong} L^q(M)$$

ovverosia, in maniera imprecisa,

"il duale di $L^p(M)$ è $L^q(M)$",

ogniqualvolta $p < \infty$ e $(M, \mathcal{M}, \mu)$ è σ-finito. Anche qui valgono le cautele già indicate nell'osservazione 8.14.

Corollario 8.17. *Siano $p, q \in [1, \infty]$ esponenti coniugati. Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Allora, per ogni $f \in L^p(M)$,*

$$\begin{aligned}\|f\|_{L^p(M)} &= \sup \left\{ \left| \int_M f\, g\, d\mu \right| : g \in L^q(M),\ \|g\|_{L^q(M)} \le 1 \right\} \\ &= \sup_{g \in L^q(M) \setminus \{0\}} \frac{\left| \int_M f\, g\, d\mu \right|}{\|g\|_{L^q(M)}}.\end{aligned}$$

Dimostrazione. I termini a membro destro sono espressioni (☞ lemma 7.5) della norma di φ_f in $(L^q(M))'$, che è uguale alla norma di f in $L^p(M)$ per il teorema 8.15. ☐

Osservazione 8.18. Il corollario 8.17 mostra che la costante $C = 1$ nella disuguaglianza di Hölder

$$\left| \int_M f\, g\, d\mu \right| \le C \|f\|_{L^p} \|g\|_{L^q}$$

è ottimale, nel senso che non può essere rimpiazzata da un valore più piccolo.

8.4 Forme sesquilineari e teorema di Lax–Milgram

Ricordiamo che $\mathbb{F}$ denota il campo $\mathbb{R}$ dei numeri reali oppure il campo $\mathbb{C}$ dei numeri complessi.

Definizione 8.19. Siano X, Y spazi vettoriali su $\mathbb{F}$. Una mappa $F : X \times Y \to \mathbb{F}$ è detta *forma sesquilineare* se valgono entrambe le seguenti proprietà:

(a) $F(\cdot, y) : X \to \mathbb{F}$ è lineare per ogni $y \in Y$, cioè

$$F(\alpha x + \alpha' x', y) = \alpha F(x, y) + \alpha' F(x', y) \qquad \forall x, x' \in X\ \forall y \in Y\ \forall \alpha, \alpha' \in \mathbb{F}.$$

(b) $F(x, \cdot) : Y \to \mathbb{F}$ è antilineare per ogni $x \in X$, cioè

$$F(x, \alpha y + \alpha' y') = \overline{\alpha} F(x, y) + \overline{\alpha'} F(x, y') \qquad \forall x \in X\ \forall y, y' \in Y\ \forall \alpha, \alpha' \in \mathbb{F}.$$

Nel caso $X = Y$, diciamo anche che una tale mappa $F : X \times X \to \mathbb{F}$ è una forma sesquilineare su X.

Osservazione 8.20. Come già discusso nell'osservazione 8.8, nel caso $\mathbb{F} = \mathbb{R}$, linearità e antilinearità coincidono, dunque una mappa F che soddisfa le proprietà della definizione 8.19 è lineare sia nella prima che nella seconda variabile; in questo caso, sarebbe più appropriato parlare di *forma bilineare*. Con una lieve improprietà di linguaggio, nel seguito utilizziamo lo stesso termine "forma sesquilineare" sia nel caso $\mathbb{F} = \mathbb{R}$ che nel caso $\mathbb{F} = \mathbb{C}$, in modo da trattare i due casi in maniera uniforme.

Osservazione 8.21. Come già discusso (☞ osservazioni 6.2 e 6.4), se $X = Y$, un prodotto scalare su X è una particolare forma sesquilineare $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$.

Per le forme sesquilineari su spazi normati, vale una caratterizzazione della continuità analoga a quella vista per gli operatori lineari (☞ proposizione 7.6).

Proposizione 8.22. *Siano X, Y spazi normati e $F : X \times Y \to \mathbb{F}$ una forma sesquilineare. Le seguenti condizioni sono equivalenti:*

(i) $F : X \times Y \to \mathbb{F}$ *è continua;*

(ii) $\|F\| \coloneqq \sup\{|F(x, y)| : x \in X,\ y \in Y,\ \|x\|_X \le 1, \|y\|_Y \le 1\} < \infty$.

In tal caso si ha anche

$$|F(x, y)| \le \|F\|\, \|x\|_X \|y\|_Y \tag{8.12}$$

per ogni $x \in X$ e $y \in Y$.

Dimostrazione. Supponiamo che $F : X \times Y \to \mathbb{F}$ sia continua. Notiamo che, per linearità, si ha $F(0, 0) = 0$. Dunque, siccome F è continua in $(0, 0)$, prendendo $\epsilon = 1$ nella definizione di continuità in un punto (☞ definizione 2.27(a)), troviamo $\delta > 0$ tale che $|F(x, y)| < 1$ per ogni $(x, y) \in X \times Y$ con $\|(x, y)\|_{X\times Y} = \max\{\|x\|_X, \|y\|_Y\} < \delta$. Se ora $(x, y) \in X \times Y$ e $\|x\|_X \le 1$ e $\|y\| \le 1$, posto $\tilde{x} = (\delta/2)x$ e $\tilde{y} = (\delta/2)y$, si ha $\|\tilde{x}\|_X < \delta$ e $\|\tilde{y}\|_Y < \delta$, dunque

$$|F(x, y)| = (4/\delta^2)|F(\tilde{x}, \tilde{y})| < 4/\delta^2,$$

e passando all'estremo superiore su x e y si ottiene $\|F\| \le 4/\delta^2 < \infty$.

Supponiamo viceversa che $\|F\| < \infty$. Dimostriamo anzitutto la stima (8.12). Notiamo che, per ogni $(x, y) \in X \times Y$,

$$F(x, 0) = F(0, y) = 0$$

per (anti)linearità; pertanto (8.12) è banalmente soddisfatta quando $x = 0$ oppure $y = 0$. Supponiamo invece che $x \neq 0$ e $y \neq 0$. Allora, posto $\tilde{x} = x/\|x\|_X$ e $\tilde{y} = y/\|y\|_Y$, si ha $\|\tilde{x}\|_X = \|\tilde{y}\|_Y = 1$, pertanto $|F(\tilde{x}, \tilde{y})| \le \|F\|$ e

$$|F(x, y)| = \|x\|_X \|y\|_Y |F(\tilde{x}, \tilde{y})| \le \|x\|_X \|y\|_Y \|F\|,$$

cioè (8.12).

Dimostriamo ora che $F : X \times Y \to \mathbb{F}$ è continua. Notiamo preliminarmente che, se $(x, y), (x', y') \in X \times Y$, si ha

$$\begin{aligned} |F(x, y) - F(x', y')| &= |F(x - x', y) + F(x', y - y')| \\ &\le \|F\|(\|x - x'\|_X \|y\|_Y + \|x'\|_X \|y - y'\|_Y), \end{aligned} \tag{8.13}$$

dove si è usata (8.12). Dunque, fissato $(x, y) \in X \times Y$ e presa una successione $((x_n, y_n))_{n\in\mathbb{N}}$ a valori in $X \times Y$ tale che $(x_n, y_n) \to (x, y)$ in $X \times Y$, si ha anche $x_n \to x$ in X e $y_n \to y$ in Y (☞ proposizione 2.13(i)), pertanto anche $\|x_n\|_X \to \|x\|_X$ e $\|y_n\|_Y \to \|y\|_Y$ per la continuità delle norme; usando (8.13), si deduce allora

$$\begin{aligned}|F(x, y) - F(x_n, y_n)| &\leq \|F\|(\|x - x_n\|_X \|y\|_Y + \|x_n\|_X \|y - y_n\|_Y) \\ &\to \|F\|(0 \cdot \|y\|_Y + \|x\|_X \cdot 0) = 0 \end{aligned} \tag{8.14}$$

per $n \to \infty$, cioè $F(x_n, y_n) \to F(x, y)$. Questo mostra che F è continua nel punto (x, y) (☞ proposizione 2.29); per l'arbitrarietà di $(x, y) \in X \times Y$, deduciamo che $F : X \times Y \to \mathbb{F}$ è continua. □

Proposizione 8.23. *Siano X uno spazio normato e H uno spazio pre-hilbertiano. Per ogni $A \in \mathcal{L}(X, H)$ definiamo $F_A : X \times H \to \mathbb{F}$ ponendo*

$$F_A(x, y) = \langle Ax, y\rangle_H \qquad \forall x \in X\ \forall y \in H. \tag{8.15}$$

Allora $F_A : X \times H \to \mathbb{F}$ è una forma sesquilineare e

$$\|F_A\| = \|A\|_{\mathrm{op}}.$$

In particolare la forma F_A è continua se e solo se l'operatore A è limitato.

Dimostrazione. Per ogni $y \in H$, usando il fatto che $\langle \cdot, y\rangle_H : H \to \mathbb{F}$ e $A : X \to H$ sono lineari, si ha che $F_A(\cdot, y) = \langle A\cdot, y\rangle_H : X \to \mathbb{F}$ è lineare (in quanto composizione di mappe lineari); inoltre, per ogni $x \in X$, la mappa $F_A(x, \cdot) = \langle Ax, \cdot\rangle_H : Y \to \mathbb{F}$ è antilineare per definizione di prodotto scalare. Dunque F_A è sesquilineare.

Ricordiamo ora dall'esempio 8.5(2) che

$$\|z\|_H = \|\langle \cdot, z\rangle_H\|_{\mathrm{op}} = \sup_{\|y\|_H \leq 1} |\langle y, z\rangle_H| = \sup_{\|y\|_H \leq 1} |\langle z, y\rangle_H|. \tag{8.16}$$

per ogni $z \in H$. In particolare,

$$\begin{aligned}\|F_A\| &= \sup_{\|x\|_X \leq 1,\, \|y\|_H \leq 1} |F_A(x, y)| \\ &= \sup_{\|x\|_X \leq 1,\, \|y\|_H \leq 1} |\langle Ax, y\rangle_H| = \sup_{\|x\|_X \leq 1} \|Ax\|_H = \|A\|_{\mathrm{op}},\end{aligned}$$

dove si è applicata (8.16) con $z = Ax$.

Di conseguenza, $\|F_A\| < \infty$ se e solo se $\|A\|_{\mathrm{op}} < \infty$, cioè F_A è continua se e solo se A è limitato (☞ proposizioni 7.6 e 8.22). □

Definizione 8.24. Siano X uno spazio normato e H uno spazio pre-hilbertiano. La forma $F_A : X \times H \to \mathbb{F}$ definita in (8.15) è detta *forma sesquilineare associata* all'operatore $A \in \mathcal{L}(X, H)$.

Nella proposizione 8.23 abbiamo costruito, per ogni operatore $A \in \mathcal{B}(X, H)$, una forma sesquilineare continua F_A. Vediamo ora che, se H è uno spazio di Hilbert, la corrispondenza si può invertire.

Proposizione 8.25. *Siano X uno spazio normato e H uno spazio di Hilbert. Per ogni forma sesquilineare continua $F : X \times H \to \mathbb{F}$, esiste un unico operatore $A \in \mathcal{B}(X, H)$ tale che $F = F_A$.*

Dimostrazione. Unicità. Se $A, B \in \mathcal{B}(X, H)$ sono tali che $F_A = F_B = F$, allora

$$\langle (A - B)x, y\rangle_H = \langle Ax, y\rangle_H - \langle Bx, y\rangle_H = F_A(x, y) - F_B(x, y) = 0$$

per ogni $x \in X$ e $y \in H$; pertanto $(A - B)x \in H^\perp = \{0\}$, cioè $(A - B)x = 0$ per ogni $x \in X$, e quindi $A = B$.

Esistenza. Sia $x \in X$. Allora la mappa $F(x, \cdot) : H \to \mathbb{F}$ è antilineare e continua, quindi la mappa $\psi_x := \overline{F(x, \cdot)} : H \to \mathbb{F}$ è lineare e continua, cioè $\psi_x \in H'$. Sia ora $\Phi : H \to H'$ l'anti-isomorfismo isometrico (8.3) del teorema di Riesz–Fréchet, e definiamo $A : X \to H$ ponendo

$$Ax = \Phi^{-1}(\psi_x)$$

per ogni $x \in X$. In particolare, $\Phi(Ax) = \psi_x$, cioè

$$\langle Ax, y\rangle_H = \overline{\langle y, Ax\rangle_H} = \overline{\psi_x(y)} = F(x, y) \tag{8.17}$$

per ogni $x \in X$ e $y \in H$.

Osserviamo ora che $A : X \to H$ è composizione della mappa $x \mapsto \psi_x = \overline{F(x, \cdot)}$, che è antilineare (perché $x \mapsto F(x, \cdot)$ è lineare) e della mappa Φ^{-1}, che è pure antilineare (in quanto inversa di una mappa antilineare); ma allora $A : X \to H$ è lineare.

Confrontando (8.17) con (8.15), si vede che $F = F_A$. Siccome F è continua, dalla proposizione 8.23 deduciamo che A è limitato, come desiderato. □

Definizione 8.26. Sia X uno spazio normato. Una forma sesquilineare $F : X \times X \to \mathbb{F}$ si dice *coerciva* se esiste $m \in (0, \infty)$ tale che

$$F(x, x) \geq m\|x\|_X^2$$

per ogni $x \in X$.

Proposizione 8.27. *Sia H uno spazio di Hilbert e $A \in \mathcal{B}(H)$. Se la forma sesquilineare $F_A : H \times H \to \mathbb{F}$ associata ad A è coerciva, allora A è un isomorfismo.*

Dimostrazione. Se F_A è coerciva, esiste $m \in (0, \infty)$ tale che

$$F_A(x, x) \geq m\|x\|_H^2 \qquad \forall x \in H. \tag{8.18}$$

Allora, per ogni $x \in H \setminus \{0\}$,

$$m\|x\|_H^2 \leq F_A(x, x) = \langle Ax, x\rangle \leq \|Ax\|_H \|x\|_H$$

per (8.18) e la disuguaglianza di Cauchy–Schwarz (☞ proposizione 6.10); dividendo per $\|x\|_H$, ne deduciamo

$$m\|x\|_H \leq \|Ax\|_H$$

e quest'ultima disuguaglianza vale anche banalmente per $x = 0$, dunque vale per ogni $x \in H$. In altre parole, A è coercivo in norma (☞ lemma 7.40).

Inoltre, se $y \in (\operatorname{Im} A)^\perp$, si ha $y \perp Ay$, dunque $F_A(y,y) = \langle Ay, y\rangle_H = 0$ e quindi, per (8.18), deduciamo che $y = 0$. Dunque $(\operatorname{Im} A)^\perp = \{0\}$, cioè $\overline{\operatorname{Im} A} = (\operatorname{Im} A)^{\perp\perp} = H$ (☞ corollario 6.33(vi) e proposizione 6.26(v)), cioè A ha immagine densa.

Per uno dei criteri di invertibilità (☞ teorema 7.39), siccome A è coercivo per la norma e ha immagine densa, e inoltre H è completo, deduciamo che $A \in \mathcal{B}(H)$ è un isomorfismo. □

Osservazione 8.28. Dalla dimostrazione della Proposizione 8.27 si vede che la coercività della forma F_A associata a un operatore $A \in \mathcal{B}(H)$ implica la coercività in norma dell'operatore A. L'implicazione inversa tuttavia non vale in generale. Infatti, per il teorema 7.39, la coercività in norma di A è verificata ogniqualvolta A è invertibile in $\mathcal{B}(H)$. D'altra parte, la condizione di coercività (8.18) implica in particolare che $F_A(x,x) \geq 0$ per ogni $x \in H$; se prendiamo ad esempio $A = -\mathrm{id}_H$, si ha $F_A(x,x) = -\|x\|^2 < 0$ per ogni $x \in H \setminus \{0\}$, quindi (se H è non banale) F_A non è coerciva in questo caso, sebbene A sia coercivo in norma. In altre parole, la condizione di coercività per una forma sesquilineare include anche una proprietà di "positività" che è assente dalla condizione di coercività in norma per operatori.

Teorema 8.29 (Lax–Milgram)**.** *Sia H uno spazio di Hilbert. Sia $F : H \times H \to \mathbb{F}$ una forma sesquilineare continua e coerciva. Allora, per ogni $\varphi \in H'$, esiste un unico $y \in H$ tale che*

$$\varphi = F(\cdot, y). \tag{8.19}$$

Dimostrazione. Sia $F^* : H \times H \to \mathbb{F}$ definita da

$$F^*(x,y) = \overline{F(y,x)}$$

per ogni $x, y \in H$. È immediato verificare che $F^* : H \times H \to \mathbb{F}$ è a sua volta una forma sesquilineare continua e coerciva. Dunque (☞ proposizioni 8.25 e 8.27) esiste un isomorfismo $A \in \mathcal{B}(X)$ tale che $F^* = F_A$. In particolare,

$$F(x,y) = \overline{F^*(y,x)} = \overline{F_A(y,x)} = \overline{\langle Ay, x\rangle_H} = \langle x, Ay\rangle_H$$

per ogni $x, y \in H$.

Pertanto, l'equazione (8.19) equivale a $\varphi = \langle \cdot, Ay\rangle_H$, cioè $\varphi = \Phi(Ay)$, dove $\Phi : H \to H'$ è l'anti-isomorfismo isometrico (8.3) del teorema di Riesz–Fréchet. Siccome sia $\Phi : H \to H'$ che $A : H \to H$ sono invertibili, possiamo riscrivere quest'ultima equazione come

$$y = A^{-1}\Phi^{-1}(\varphi),$$

da cui è chiaro che, per ogni $\varphi \in H'$, esiste un unico $y \in H$ che la soddisfa. □

Osservazione 8.30. Il teorema di Lax–Milgram ha importanti applicazioni per la dimostrazione dell'esistenza di soluzioni deboli di certe equazioni differenziali. Ad esempio, si consideri l'equazione differenziale ordinaria del secondo ordine

$$-u'' + Vu = h$$

ove u è l'incognita, V è un coefficiente e h è il dato, pensati come funzioni su $\mathbb{R}$; usando integrazione per parti, si può riscrivere l'equazione in forma debole come

$$\langle \phi', u'\rangle + \langle \phi, Vu\rangle = \langle \phi, h\rangle,$$

ove $\langle \cdot, \cdot \rangle$ è il prodotto scalare di $L^2(\mathbb{R})$, e ϕ varia in un'opportuna classe di funzioni test. Ponendo $F(\phi, u) = \langle \phi', u' \rangle + \langle \phi, Vu \rangle$ e $\varphi = \langle \cdot, h \rangle$, ci si ritrova allora con un'equazione della forma (8.19). Con un'opportuna scelta di uno spazio di Hilbert H di funzioni su $\mathbb{R}$ con cui lavorare, e sotto opportune ipotesi su V, è possibile mostrare che la forma F è continua e coerciva; il teorema 8.29 garantisce quindi l'esistenza di una soluzione $u \in H$ per ogni dato h tale che il funzionale φ associato sia in H'. Maggiori dettagli richiederebbero la teoria degli spazi di Sobolev, che vanno oltre gli argomenti coperti da queste note.

8.5 Estensione di funzionali: il teorema di Hahn–Banach

Teorema 8.31 (Hahn–Banach)**.** *Sia X uno spazio normato. Sia V un sottospazio vettoriale di X; dotiamo V della norma indotta da X. Sia $\varphi \in V'$. Allora esiste $\widetilde{\varphi} \in X'$ tale che $\widetilde{\varphi}|_V = \varphi$ e*

$$\|\widetilde{\varphi}\|_{X'} = \|\varphi\|_{V'}. \tag{8.20}$$

Osservazione 8.32. Nel caso V sia denso in X, questo risultato è contenuto nel teorema di estensione (☞ teorema 7.23) che abbiamo già discusso (dato che $X' = \mathcal{B}(X, \mathbb{F})$ e $\mathbb{F}$ è completo); inoltre, in tal caso l'estensione $\widetilde{\varphi}$ è unica. In generale, se V non è denso in X, l'estensione $\widetilde{\varphi}$ di φ non è unica.

Osservazione 8.33. Se $\widetilde{\varphi}|_V = \varphi$, chiaramente $\|\widetilde{\varphi}\|_{X'} \geq \|\varphi\|_{V'}$ (si tratta di norme operatoriali, ☞ osservazione 8.3); la disuguaglianza opposta è il contenuto non banale di (8.20).

Dimostrazione (idea, caso $\mathbb{F} = \mathbb{R}$). Se $V = X$ non c'è nulla da dimostrare, quindi possiamo supporre che V sia un sottospazio vettoriale proprio di X. Inoltre, se $\varphi = 0$, chiaramente possiamo prendere $\widetilde{\varphi} = 0$ come estensione. Quindi possiamo assumere che $\varphi \neq 0$ e, a meno di rimpiazzare φ con $\varphi/\|\varphi\|_{V'}$, possiamo anche assumere $\|\varphi\|_{\mathrm{op}} = \|\varphi\|_{V'} = 1$.

Passo 1. Dimostriamo il teorema quando $X = \mathrm{span}(V \cup \{y\})$ per qualche $y \in X \setminus V$, cioè quando V ha *codimensione 1* in X.

In questo caso, si ha $X = V \oplus \mathbb{R}y$, cioè ogni $x \in X$ si scrive in maniera unica come

$$x = v + \alpha y,$$

dove $v \in V$ e $\alpha \in \mathbb{R}$. Per ogni dato $\lambda \in \mathbb{R}$, possiamo dunque definire un'estensione $\widetilde{\varphi} \in \mathcal{L}(X, \mathbb{R})$ di φ ponendo

$$\widetilde{\varphi}(v + \alpha y) = \varphi(v) + \alpha\lambda \qquad \forall v \in V \ \forall \alpha \in \mathbb{R}. \tag{8.21}$$

Il problema è scegliere $\lambda \in \mathbb{R}$ in modo tale che l'estensione $\widetilde{\varphi}$ soddisfi (8.20); siccome $\|\varphi\|_{V'} = 1$, per l'osservazione 8.33 il problema si riduce a scegliere $\lambda \in \mathbb{R}$ in modo tale che

$$\|\widetilde{\varphi}\|_{\mathrm{op}} \leq 1. \tag{8.22}$$

D'altra parte, siccome $\widetilde{\varphi}$ è definito da (8.21), la stima (8.22) equivale a

$$|\varphi(v) + \alpha\lambda| \leq \|v + \alpha y\|_X \qquad \forall v \in V, \ \forall \alpha \in \mathbb{R},$$

cioè

$$-\varphi(v) - \|v + \alpha y\|_X \leq \alpha\lambda \leq -\varphi(v) + \|v + \alpha y\|_X \qquad \forall v \in V, \ \forall \alpha \in \mathbb{R}.$$

Per $\alpha = 0$, tale stima è senz'altro verificata, dato che $\|\varphi\|_{\mathrm{op}} = 1$ e $\|v\|_X = \|v\|_V$. Per $\alpha \neq 0$, tramite riscalamento ci si riconduce alla stima equivalente

$$-\varphi(v) - \|v+y\|_X \leq \lambda \leq -\varphi(v) + \|v+y\|_X \qquad \forall v \in V. \tag{8.23}$$

Affinché esista $\lambda \in \mathbb{R}$ che soddisfi (8.23), è necessario e sufficiente che

$$\sup\{-\varphi(v) - \|v+y\|_X : v \in V\} \leq \inf\{-\varphi(v) + \|v+y\|_X : v \in V\},$$

cioè

$$-\varphi(v) - \|v+y\|_X \leq -\varphi(w) + \|w+y\|_X \qquad \forall v, w \in V,$$

cioè

$$\varphi(w-v) \leq \|w+y\|_X + \|v+y\|_X \qquad \forall v, w \in V.$$

Quest'ultima stima tuttavia è senz'altro verificata, perché $\|\varphi\|_{\mathrm{op}} = 1$ e quindi

$$\begin{aligned}\varphi(v-w) \leq |\varphi(w-v)| \leq \|w-v\|_V = \|w-v\|_X& \\ = \|(w+y)-(v+y)\|_X \leq \|w+y\|_X + \|v+y\|_X&\end{aligned}$$

per la disuguaglianza triangolare.

Pertanto possiamo scegliere $\lambda \in \mathbb{R}$ in modo da soddisfare (8.23) e quindi l'estensione $\widetilde{\varphi}$ definita da (8.21) soddisfa (8.22), come desiderato.

Passo 2. Dimostriamo il caso generale.

Sia $\mathcal{F}$ l'insieme delle coppie (W, ψ), dove W è un sottospazio vettoriale di X contenente V, e $\psi \in W'$ è un'estensione di φ tale che $\|\psi\|_{\mathrm{op}} \leq 1$. Ovviamente $(V, \varphi) \in \mathcal{F}$, quindi $\mathcal{F}$ non è vuoto. Definiamo una relazione d'ordine $\leq$ su $\mathcal{F}$ stabilendo che $(W_1, \psi_1) \leq (W_2, \psi_2)$ se $W_1 \subseteq W_2$ e $\psi_2|_{W_1} = \psi_1$.

Vogliamo applicare il lemma di Zorn a $(\mathcal{F}, \leq)$. A tal fine, dobbiamo verificare che, preso un sottoinsieme $\{(W_\alpha, \psi_\alpha)\}_{\alpha \in A}$ non vuoto e totalmente ordinato di $\mathcal{F}$, tale sottoinsieme ha un maggiorante in $\mathcal{F}$. Infatti, posto $W = \bigcup_{\alpha \in A} W_\alpha$, si ha che W è un sottospazio vettoriale di X contenente V (in quanto unione di sottospazi W_α di X contenenti V e totalmente ordinati per inclusione). Inoltre, le mappe $\psi_\alpha : W_\alpha \to \mathbb{R}$ sono una l'estensione dell'altra, dunque si "incollano" in un'unica mappa $\psi : W \to \mathbb{R}$, che si verifica facilmente essere lineare (dato che le ψ_α sono lineari) e soddisfare $|\psi(x)| \leq \|x\|_X$ per ogni $x \in W = \bigcup_{\alpha \in A} W_\alpha$, cioè $\|\psi\|_{\mathrm{op}} \leq 1$ (dato che $\|\psi_\alpha\|_{\mathrm{op}} \leq 1$ per ogni $\alpha \in A$). Pertanto si ha $(W, \psi) \in \mathcal{F}$, e inoltre $(W_\alpha, \psi_\alpha) \leq (W, \psi)$ per ogni $\alpha \in A$, cioè (W, ψ) è un maggiorante di $\{(W_\alpha, \psi_\alpha)\}_{\alpha \in A}$ in $(\mathcal{F}, \leq)$.

Per il lemma di Zorn, possiamo allora concludere che $(\mathcal{F}, \leq)$ ha un elemento massimale (W_0, ψ_0). Se $W_0 = X$, allora ψ_0 è l'estensione $\widetilde{\varphi}$ di φ cercata. Supponiamo dunque per assurdo che W_0 sia un sottospazio vettoriale proprio di X. Allora esiste $y \in X \setminus W_0$, e possiamo applicare il passo 1 per estendere ψ_0 a un funzionale $\widetilde{\psi_0}$ su $\widetilde{W_0} = \mathrm{span}(W_0 \cup \{y\})$ tale che $\|\widetilde{\psi_0}\|_{\mathrm{op}} \leq 1$. Ma allora $\widetilde{W_0} \supseteq W_0 \supseteq V$ e $\widetilde{\psi_0}|_V = \psi_0|_V = \varphi$, cioè $(\widetilde{W_0}, \widetilde{\psi_0}) \in \mathcal{F}$ e $(W_0, \psi_0) \leq (\widetilde{W_0}, \widetilde{\psi_0})$; siccome $(W_0, \psi_0) \neq (\widetilde{W_0}, \widetilde{\psi_0})$, questo contraddice che (W_0, ψ_0) sia massimale in $(\mathcal{F}, \leq)$, assurdo. $\square$

Dimostrazione (idea, caso $\mathbb{F} = \mathbb{C}$). Ricordiamo che ogni spazio vettoriale su $\mathbb{C}$ si può anche pensare come spazio vettoriale su $\mathbb{R}$, restringendo il prodotto scalare-vettore (☞ esempio 1.2(6)); analogamente, ogni spazio normato su $\mathbb{C}$ si può pensare come spazio normato su $\mathbb{R}$, con la stessa norma.

Ora, siccome $\varphi : V \to \mathbb{C}$ è $\mathbb{C}$-lineare, $\mathfrak{Re}\,\varphi : V \to \mathbb{R}$ è $\mathbb{R}$-lineare, e inoltre

$$|\mathfrak{Re}\,\varphi(x)| \le |\varphi(x)| \le \|\varphi\|_{\mathrm{op}} \|x\|_V \qquad \forall x \in V,$$

pertanto $\|\mathfrak{Re}\,\varphi\|_{\mathrm{op}} \le \|\varphi\|_{\mathrm{op}} < \infty$ e $\mathfrak{Re}\,\varphi : V \to \mathbb{R}$ è un funzionale $\mathbb{R}$-lineare continuo su V. Applicando il teorema di Hahn–Banach per spazi normati reali, possiamo dunque estendere $\mathfrak{Re}\,\varphi$ a un funzionale $\mathbb{R}$-lineare $\psi : X \to \mathbb{R}$ con $\|\psi\|_{\mathrm{op}} \le \|\mathfrak{Re}\,\varphi\|_{\mathrm{op}} \le \|\varphi\|_{\mathrm{op}}$. Definiamo ora $\widetilde{\varphi} : X \to \mathbb{C}$ ponendo

$$\widetilde{\varphi}(x) = \psi(x) - i\psi(ix) \qquad \forall x \in X.$$

Allora $\widetilde{\varphi} : X \to \mathbb{C}$ è chiaramente $\mathbb{R}$-lineare (in quanto somma di composizioni di mappe $\mathbb{R}$-lineari), e inoltre

$$\widetilde{\varphi}(ix)\psi(ix) - i\psi(i^2 x) = \psi(ix) - i\psi(-x) = i(\psi(x) - i\psi(ix)) = i\widetilde{\varphi}(x)$$

per ogni $x \in X$; da questo deduciamo che $\widetilde{\varphi} : X \to \mathbb{C}$ è $\mathbb{C}$-lineare.

Inoltre, se $x \in V$, anche $ix \in V$ (dato che V è un $\mathbb{C}$-sottospazio di X), e inoltre φ è $\mathbb{C}$-lineare, dunque

$$\psi(x) = \mathfrak{Re}\,\varphi(x), \qquad \psi(ix) = \mathfrak{Re}\,\varphi(ix) = \mathfrak{Re}(i\varphi(x)) = -\,\mathfrak{Im}\,\varphi(x)$$

da cui segue che

$$\widetilde{\varphi}(x) = \psi(x) - i\psi(ix) = \mathfrak{Re}\,\varphi(x) + i\,\mathfrak{Im}\,\varphi(x) = \varphi(x);$$

in altre parole $\widetilde{\varphi}$ è un'estensione di φ.

Infine, per ogni $x \in X$, possiamo trovare $\alpha \in \mathbb{C}$ con $|\alpha| = 1$ tale che $\widetilde{\varphi}(\alpha x) = \alpha\widetilde{\varphi}(x) \in \mathbb{R}$, quindi

$$|\widetilde{\varphi}(x)| = |\widetilde{\varphi}(\alpha x)| = |\mathfrak{Re}\,\widetilde{\varphi}(\alpha x)| = |\psi(\alpha x)| \le \|\psi\|_{\mathrm{op}} \|\alpha x\|_X \le \|\varphi\|_{\mathrm{op}} \|x\|_X;$$

questo dimostra che $\|\widetilde{\varphi}\|_{\mathrm{op}} \le \|\varphi\|_{\mathrm{op}}$, e quindi $\widetilde{\varphi}$ è l'estensione di φ cercata. □

Osservazione 8.34. Nel caso X sia uno spazio di Hilbert, il risultato di estensione del teorema 8.20 si può dimostrare in maniera più semplice. Anzitutto, usando il teorema di estensione (☞ teorema 7.23), un funzionale $\varphi \in V'$ si estende in maniera unica a un funzionale continuo su $\overline{V}'$ con la stessa norma, che denotiamo ancora con φ. A questo punto, si può costruire un'estensione $\widetilde{\varphi}$ di $\varphi \in \overline{V}'$ ponendo $\widetilde{\varphi} = \varphi \circ P_{\overline{V}}$, dove $P_{\overline{V}} : X \to \overline{V}$ è la mappa di proiezione ortogonale (☞ definizione 6.31). Chiaramente $P_{\overline{V}} x = x$ per $x \in \overline{V}$, quindi $\widetilde{\varphi}|_{\overline{V}} = \varphi$. Siccome $\|P_{\overline{V}}\|_{\mathrm{op}} \le 1$ (☞ corollario 6.33(iv)), per la submoltiplicatività della norma operatoriale si ha $\|\widetilde{\varphi}\|_{X'} \le \|\varphi\|_{V'} \|P_{\overline{V}}\|_{\mathrm{op}} \le \|\varphi\|_{V'}$, da cui segue (8.20) (☞ osservazione 8.33).

Osservazione 8.35. Ci si può chiedere se un approccio analogo a quello discusso nell'osservazione 8.34 sia praticabile per più generali spazi normati. Tuttavia, per un arbitrario spazio normato X (anche se completo), non è detto che esista una "proiezione" lineare da X a $\overline{V}$ di norma ≤ 1. Questo è ciò che rende il teorema di Hahn–Banach un risultato non banale *già in dimensione finita.* In effetti, il cuore della dimostrazione è il passo 1 (estensione a uno spazio "di una dimensione più grande"), e il passo 2 è semplicemente una "iterazione transfinita" del passo 1; quando X ha dimensione finita (o, più in generale, quando $\dim(X/V) < \infty$), per questa iterazione, anziché usare il lemma di Zorn, si potrebbe semplicemente procedere per induzione.

Discutiamo ora alcune conseguenze importanti del teorema di Hahn–Banach. Il primo risultato ci dice, fra le altre cose, che il duale di uno spazio normato non banale è a sua volta non banale.

Corollario 8.36. *Sia X uno spazio normato.*

(i) Per ogni $x \in X \setminus \{0\}$, esiste $\varphi \in X'$ tale che $\|\varphi\|_{X'} = 1$ e $\varphi(x) = \|x\|_X$.

(ii) Per ogni $x \in X$,

$$\|x\|_X = \max\{|\varphi(x)| : \varphi \in X',\ \|\varphi\|_{X'} \le 1\}$$

e, se $X \neq \{0\}$, si ha anche

$$\|x\|_X = \max\{|\varphi(x)| : \varphi \in X',\ \|\varphi\|_{X'} = 1\}.$$

(iii) Se $x_1, x_2 \in X$ e $x_1 \neq x_2$, allora esiste $\varphi \in X'$ tale che $\varphi(x_1) \neq \varphi(x_2)$; in altre parole, i funzionali $\varphi \in X'$ separano i punti *di X.*

(iv) Se $X \neq \{0\}$, allora $X' \neq \{0\}$.

Dimostrazione. (i). Sia $V = \operatorname{span}\{x\} = \{\alpha x : \alpha \in \mathbb{F}\}$ e sia $\psi : V \to \mathbb{F}$ l'unica mappa lineare tale che $\psi(x) = \|x\|_X$ (☞ proposizione 1.22). Allora, per ogni $\alpha \in \mathbb{F}$,

$$|\psi(\alpha x)| = |\alpha||\psi(x)| = |\alpha|\|x\|_X = \|\alpha x\|_X,$$

da cui segue che $\|\psi\|_{\mathrm{op}} = 1$ e $\psi \in V'$. Per il teorema di Hahn–Banach (☞ teorema 8.31) ψ si estende a un funzionale $\varphi \in X'$ con $\|\varphi\|_{X'} = 1$, e chiaramente $\varphi(x) = \psi(x) = \|x\|_X$.

(ii). Se $X = \{0\}$, si ha anche $X' = \{0\}$ e il risultato è banalmente vero. Supponiamo invece $X \neq \{0\}$. Preso $x \in X$, per ogni $\varphi \in X'$ con $\|\varphi\|_{X'} \le 1$ si ha

$$|\varphi(x)| \le \|\varphi\|_{X'}\|x\|_X \le \|x\|_X,$$

dunque senz'altro valgono le disuguaglianze

$$\sup\{|\varphi(x)| : \varphi \in X',\ \|\varphi\|_{X'} = 1\} \le \sup\{|\varphi(x)| : \varphi \in X',\ \|\varphi\|_{X'} \le 1\} \le \|x\|_X.$$

D'altra parte, il funzionale φ costruito in (i) soddisfa $\|\varphi\|_{X'} = 1$ e $|\varphi(x)| = \|x\|_X$, dunque le disuguaglianze sopra sono tutte uguaglianze e gli estremi superiori sono massimi.

(iii). Se $x_1 \neq x_2$, allora $x_1 - x_2 \neq 0$; il funzionale $\varphi \in X'$ dato da (i) con $x = x_1 - x_2$ soddisfa allora $\varphi(x_1) - \varphi(x_2) = \varphi(x_1 - x_2) = \|x_1 - x_2\|_X \neq 0$, cioè $\varphi(x_1) \neq \varphi(x_2)$.

(iv). Preso $x \in X \setminus \{0\}$, il funzionale $\varphi \in X'$ dato da (i) ha norma non nulla, dunque $\varphi \neq 0$ e $X' \neq \{0\}$. □

Osservazione 8.37. Se X è uno spazio pre-hilbertiano e $x \in X \setminus \{0\}$, il funzionale $\varphi = \langle \cdot, \tilde{x} \rangle \in X'$ associato al "versore" $\tilde{x} = x/\|x\|_X$ (il vettore di norma 1 nella stessa direzione di x) soddisfa $\|\varphi\|_{X'} = \|\tilde{x}\|_X = 1$ e $\varphi(x) = \langle x, x/\|x\|_X \rangle = \|x\|_X$ (☞ esempio 8.5(2)). Il corollario 8.36(i), che si applica a un arbitrario spazio normato X, può essere dunque pensato come la costruzione di un "versore" $\varphi \in X'$ che realizza la norma di x.

Osservazione 8.38. Sappiamo già (☞ osservazione 8.3 e lemma 7.5) che per ogni $\varphi \in X'$ vale

$$\|\varphi\|_{X'} = \sup\{|\varphi(x)| : x \in X,\ \|x\|_X \leq 1\} \tag{8.24}$$

e, se $X \neq \{0\}$,

$$\|\varphi\|_{X'} = \sup\{|\varphi(x)| : x \in X,\ \|x\|_X = 1\}. \tag{8.25}$$

Il corollario 8.36(ii) afferma dunque che vale anche una relazione "duale" con i ruoli delle norme $\|\cdot\|_{X'}$ e $\|\cdot\|_X$ scambiati. In effetti quest'ultima relazione è più precisa, dato che nel corollario 8.36(ii) abbiamo massimi anziché estremi superiori, mentre al contrario in (8.24)-(8.25) non è detto che esista $x \in X$ che "realizza la norma" di φ.

Il prossimo risultato si può pensare come una generalizzazione del corollario 8.36(i); invece di un funzionale φ che "realizza la norma" di un vettore x (cioè la distanza di x da 0), qui costruiamo un funzionale che "realizza la distanza da un sottospazio" (☞ formula (2.6)).

Corollario 8.39. *Sia X uno spazio normato. Sia V un sottospazio vettoriale di X.*

(i) Per ogni $x \in X$ tale che $d(x, V) > 0$, esiste $\varphi \in X'$ tale che $\|\varphi\|_{X'} = 1$, $\varphi|_V = 0$ e $\varphi(x) = d(x, V)$.

(ii) Se V è un sottospazio vettoriale chiuso proprio di X, esiste $\varphi \in X'$ con $\|\varphi\|_{X'} = 1$ e $\varphi|_V = 0$.

Dimostrazione. (i). Sia $W = \operatorname{span}(V \cup \{x\}) = V \oplus \mathbb{F}x$. Definiamo $\psi : W \to \mathbb{F}$ ponendo

$$\psi(v + \alpha x) = \alpha d(x, V) \qquad \forall \alpha \in \mathbb{F}\ \forall v \in V.$$

Allora $\psi \in \mathcal{L}(W, \mathbb{F})$ e, per ogni $v \in V$ e $\alpha \in \mathbb{F} \setminus \{0\}$,

$$|\psi(v + \alpha x)| = |\alpha| d(x, V) \leq |\alpha|\, \|x - (-v/\alpha)\|_X = \|\alpha x + v\|_X,$$

dove si è usato che $d(x, V) \leq \|x - (-v/\alpha)\|_X$, dato che $-v/\alpha \in V$ (☞ formula (2.6)). Una disuguaglianza analoga vale banalmente anche per $\alpha = 0$, pertanto $\|\psi\|_{\mathrm{op}} \leq 1$ e $\psi \in W'$.

In effetti, per ogni $\epsilon > 0$, preso $v \in V$ tale che $\|x - v\|_X \leq (1 + \epsilon) d(x, V)$, si ha

$$\psi(x - v) = d(x, V) \geq (1 + \epsilon)^{-1} \|x - v\|_X,$$

da cui

$$\|\psi\|_{\mathrm{op}} \geq \frac{|\psi(x - v)|}{\|x - v\|_X} \geq (1 + \epsilon)^{-1};$$

siccome $\epsilon > 0$ è arbitrario, passando al limite per $\epsilon \to 0^+$, si ottiene

$$\|\psi\|_{W'} = \|\psi\|_{\mathrm{op}} \geq 1$$

e quindi in effetti $\|\psi\|_{W'} = 1$.

Possiamo allora applicare il teorema di Hahn–Banach (☞ teorema 8.31) ed estendere ψ a un funzionale $\varphi \in X'$ con $\|\varphi\|_{X'} = \|\psi\|_{W'} = 1$. Allora

$$\varphi(x) = \psi(x) = d(x, V), \qquad \varphi(v) = \psi(v) = 0 \quad \forall v \in V,$$

cioè φ ha le proprietà cercate.

(ii). Siccome V è chiuso e proprio, esiste $x \in X \setminus V$, e si ha $d(x, V) > 0$ (☞ proposizione 2.23(vii)). Basta allora applicare (i) a V e x. ☐

Osservazione 8.40. Nel caso X sia uno spazio di Hilbert, si ha $d(x, V) = d(x, \overline{V}) = \|P_{\overline{V}^\perp} x\|_X$; in tal caso, il funzionale $\varphi = \langle \cdot, P_{\overline{V}^\perp} x / \|P_{\overline{V}^\perp} x\|_X \rangle$ ha le proprietà enunciate nel corollario 8.39(i).

Infine, ecco una conseguenza relativa alla separabilità degli spazi normati.

Corollario 8.41. *Sia X uno spazio normato. Se X' è separabile, allora anche X è separabile.*

Dimostrazione. Possiamo assumere $X \neq \{0\}$. Se X' è separabile, anche la sfera unitaria

$$S = \{\varphi \in X' : \|\varphi\|_{X'} = 1\}$$

è separabile (☞ esempio 2.26(4)). Dunque esiste un sottoinsieme $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq S$ tale che $\overline{\{\varphi_n\}_{n\in\mathbb{N}}}^{X'} = S$.

Per ogni $n \in \mathbb{N}$, siccome $\|\varphi_n\|_{X'} = 1$, esiste $x_n \in X$ con $\|x_n\|_X = 1$ e

$$|\varphi_n(x_n)| \geq 1/2 \tag{8.26}$$

(☞ formula (8.25)). Sia $W = \overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}}^X$. Se $W = X$, allora X è separabile (☞ proposizione 4.29) e abbiamo finito.

Supponiamo per assurdo che invece W sia un sottospazio vettoriale chiuso proprio di X. Per il corollario 8.39(ii), esiste allora $\psi \in S$ con $\psi|_W = 0$. In particolare $\operatorname{Ker}\psi \supseteq W \supseteq \{x_n\}_{n\in\mathbb{N}}$, cioè

$$\psi(x_n) = 0$$

per ogni $n \in \mathbb{N}$. D'altra parte, siccome $\{\varphi_n\}_{n\in\mathbb{N}}$ è denso in S, esiste una successione $\{\varphi_{n_k}\}_{k\in\mathbb{N}}$ tale che $\varphi_{n_k} \to \psi$ in X' per $k \to \infty$, ma allora

$$|\phi_{n_k}(x_{n_k})| = |\phi_{n_k}(x_{n_k}) - \psi(x_{n_k})| \leq \|\phi_{n_k} - \psi\|_{X'} \|x_{n_k}\|_X = \|\phi_{n_k} - \psi\|_{X'} \to 0$$

per $k \to \infty$, da cui segue che $\lim_{k\to\infty} |\phi_{n_k}(x_{n_k})| = 0$; questo tuttavia contraddice (8.26), assurdo. □

L'implicazione "X' separabile $\Rightarrow$ X separabile" del corollario 8.41 in generale non si può invertire, come mostra il prossimo risultato.

Proposizione 8.42. *Valgono le seguenti proprietà.*

(i) *ℓ^∞ non è separabile.*

(ii) *Se $p \in [1, \infty)$, ℓ^p è separabile.*

(iii) *c_0 è separabile.*

Dimostrazione. (i). Se per assurdo ℓ^∞ fosse separabile, anche il sottoinsieme

$$\{0,1\}^{\mathbb{N}} = \{\underline{x} \in \ell^\infty : x_k \in \{0,1\} \ \forall k \in \mathbb{N}\}$$

sarebbe separabile (☞ esempio 2.26(4)) rispetto alla distanza indotta dalla norma di ℓ^∞. Sia dunque $\{\underline{x}^{(n)}\}_{n\in\mathbb{N}}$ un denso numerabile in $\{0,1\}^{\mathbb{N}}$. Costruiamo allora un elemento $\underline{y} \in \{0,1\}^{\mathbb{N}}$ ponendo, per ogni $k \in \mathbb{N}$,

$$y_k = \begin{cases} 1 & \text{se } x_k^{(k)} = 0, \\ 0 & \text{se } x_k^{(k)} = 1. \end{cases}$$

Si ha allora, per ogni $n \in \mathbb{N}$,

$$\|\underline{y} - \underline{x^{(n)}}\|_\infty \geq |y_n - x_n^{(n)}| = |1 - 0| = 1,$$

e quindi non può esistere una successione di elementi di $\{\underline{x^{(n)}}\}_{n\in\mathbb{N}}$ che converge a $\underline{y}$ in norma $\|\cdot\|_\infty$. Siccome $\underline{y} \in \{0,1\}^{\mathbb{N}}$, questo contraddice che $\{\underline{x^{(n)}}\}_{n\in\mathbb{N}}$ è denso in $\{0,1\}^{\mathbb{N}}$, assurdo.

(ii) e (iii). Queste sono conseguenze della proposizione 4.29 e del fatto che $c_{00} = \operatorname{span}\{\underline{e^{(n)}}\}_{n\in\mathbb{N}}$ è denso in c_0 e in ℓ^p per $p < \infty$ (☞ formula (6.25) e proposizioni 5.6(iii) e 5.17(iii)). □

Osservazione 8.43. I risultati della proposizione 8.42 possono in alcuni casi essere generalizzati agli spazi $L^p(M, \mathcal{M}, \mu)$.

(1) L'argomento nella dimostrazione della proposizione 8.42(i) si può generalizzare a $L^\infty(M, \mathcal{M}, \mu)$, sotto l'ipotesi che esista una famiglia numerabile $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ di insiemi misurabili a due a due disgiunti con $\mu(A_n) > 0$ per ogni $n \in \mathbb{N}$. In tal caso, si può ripetere la dimostrazione sopra sostituendo a $\{0,1\}^{\mathbb{N}}$ l'insieme

$$\{f \in L^\infty(M) : \forall n \in \mathbb{N} : f|_{A_n} \equiv 1 \text{ oppure } f|_{A_n} \equiv 0\},$$

così concludendo che $L^\infty(M, \mathcal{M}, \mu)$ non è separabile.

D'altra parte, se M è un insieme finito, allora $\dim L^\infty(M, \mathcal{M}, \mu) < \infty$ e quindi in questo caso $L^\infty(M, \mathcal{M}, \mu)$ è separabile.

(2) Come discusso nel corollario 5.39, nel caso $M = [a, b]$ sia un intervallo limitato di $\mathbb{R}$ con la misura di Lebesgue, $L^p(a, b)$ è separabile per $p < \infty$.

Corollario 8.44. *Il duale di ℓ^∞ non è isomorfo a ℓ^1. In particolare, l'isometria lineare $\Psi : \ell^1 \to (\ell^\infty)'$ del teorema 8.10 non è suriettiva.*

Dimostrazione. Siccome ℓ^∞ non è separabile (☞ proposizione 8.42(i)), per il corollario 8.41 anche $(\ell^\infty)'$ non è separabile. D'altra parte, ℓ^1 è separabile (☞ proposizione 8.42(ii)), quindi $(\ell^\infty)'$ e ℓ^1 non sono isomorfi. □

9 Biduale e convergenze deboli

9.1 Biduale e spazi riflessivi

Definizione 9.1. Sia X uno spazio normato. Il *biduale* X'' dello spazio X è il duale del duale di X.

Osservazione 9.2. Se $p, q \in (1, \infty)$ sono esponenti coniugati, allora sappiamo (☞ osservazione 8.14) che

$$(\ell^p)' \underset{\text{isom.}}{\cong} \ell^q, \qquad (\ell^q)' \underset{\text{isom.}}{\cong} \ell^p,$$

da cui è possibile dedurre che

$$(\ell^p)'' \underset{\text{isom.}}{\cong} (\ell^q)' \underset{\text{isom.}}{\cong} \ell^p.$$

Possiamo chiederci se una simile relazione fra uno spazio X e il suo biduale X'' si verifichi per più generali spazi normati.

Proposizione 9.3. *Sia X uno spazio normato. Definiamo $J : X \to X''$ ponendo*

$$J(x)(\varphi) = \varphi(x) \qquad \forall \varphi \in X' \ \forall x \in X. \tag{9.1}$$

Allora $J : X \to X''$ è un'isometria lineare, detta immersione canonica di X nel suo biduale.

Dimostrazione. Verifichiamo anzitutto che $J : X \to X''$ è ben definita. In altre parole, dobbiamo verificare che, se $x \in X$, allora $J(x) \in X'' = \mathcal{B}(X', \mathbb{F})$.

Chiaramente, per ogni $\varphi \in X'$, si ha $J(x)(\varphi) = \varphi(x) \in \mathbb{F}$, dunque $J(x) : X' \to \mathbb{F}$ è ben definita. Inoltre, $J(x)$ è lineare: se $\alpha_1, \alpha_2 \in \mathbb{F}$ e $\varphi_1, \varphi_2 \in X'$, si ha

$$\begin{aligned} J(x)(\alpha_1\varphi_1 + \alpha_2\varphi_2) &= (\alpha_1\varphi_1 + \alpha_2\varphi_2)(x) \\ &= \alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) = \alpha_1 J(x)(\varphi_1) + \alpha_2 J(x)(\varphi_2), \end{aligned}$$

dove abbiamo usato (9.1) e la definizione delle operazioni su X'. Infine, per ogni $\varphi \in X'$,

$$|J(x)(\varphi)| = |\varphi(x)| \leq \|\varphi\|_{X'} \|x\|_X,$$

da cui segue che

$$\|J(x)\|_{\text{op}} \leq \|x\|_X < \infty \tag{9.2}$$

e quindi $J(x) \in \mathcal{B}(X', \mathbb{F}) = X''$.

Verifichiamo ora che $J : X \to X''$ è lineare. Infatti, per ogni $\alpha_1, \alpha_2 \in \mathbb{F}$, $x_1, x_2 \in X$ e $\varphi \in X'$, si ha

$$\begin{aligned} J(\alpha_1 x_1 + \alpha_2 x_2)(\varphi) &= \varphi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1\varphi(x_1) + \alpha_2\varphi(x_2) \\ &= \alpha_1 J(x_1)(\varphi) + \alpha_2 J(x_2)(\varphi) = (\alpha_1 J(x_1) + \alpha_2 J(x_2))(\varphi), \end{aligned}$$

dove abbiamo usato (9.1), la linearità di φ, e la definizione delle operazioni su X''; per l'arbitrarietà di $\varphi \in X'$, questo implica che

$$J(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 J(x_1) + \alpha_2 J(x_2),$$

dunque J è lineare.

Verifichiamo infine che $J : X \to X''$ è un'isometria lineare (☞ definizione 7.15). Preso $x \in X$, sappiamo già da (9.2) che $\|J(x)\|_{X''} = \|J(x)\|_{\mathrm{op}} \le \|x\|_X$; rimane da dimostrare la disuguaglianza opposta. D'altra parte, se $x = 0$, per linearità si ha anche $J(x) = 0$ e il risultato è banalmente vero. Se invece $x \neq 0$, per il corollario 8.36(i) del teorema di Hahn–Banach, esiste $\varphi \in X'$ con $\|\varphi\|_{X'} = 1$ e $\varphi(x) = \|x\|_X$; pertanto

$$\|J(x)\|_{\mathrm{op}} \ge |J(x)(\varphi)|/\|\varphi\|_{X'} = |\varphi(x)| = \|x\|_X,$$

come desiderato. □

Dunque, in base alla proposizione 9.3, ogni spazio normato X è isometricamente isomorfo a un sottospazio vettoriale del biduale X'', cioè l'immagine $J(X)$ dell'immersione canonica.

Definizione 9.4. Uno spazio normato X si dice *riflessivo* se l'immersione canonica $J : X \to X''$ è suriettiva.

Osservazione 9.5. Se X è riflessivo, allora $X \underset{\text{isom.}}{\cong} X'' = (X')'$, dunque (☞ osservazione 8.3) X è necessariamente uno spazio di Banach.

Osservazione 9.6. Se $\dim X < \infty$, allora $\dim X'' = \dim X' = \dim X$ (☞ osservazione 8.6); dato che $J : X \to X''$ è lineare e iniettiva, si ha allora $\dim \operatorname{Im} J = \dim X = \dim X''$ (☞ proposizione 1.20(xi)), pertanto $\operatorname{Im} J = X''$, cioè J è suriettiva. Dunque ogni spazio normato di dimensione finita è riflessivo.

Proposizione 9.7. *Ogni spazio di Hilbert H è riflessivo.*

Dimostrazione (cenno). Sia $L \in H''$. Sia $\Phi : H \to H'$ l'anti-isomorfismo isometrico (8.3) del teorema di Riesz–Fréchet. Allora è facile verificare che la mappa $\overline{L \circ \Phi} : H \to \mathbb{F}$ è lineare e continua, cioè $\overline{L \circ \Phi} \in H'$. Quindi, per il teorema di Riesz–Fréchet (☞ teorema 8.7), esiste $x \in H$ tale che

$$\Phi(x) = \overline{L \circ \Phi}. \tag{9.3}$$

Pertanto, per ogni $y \in H$,

$$L(\Phi(y)) = L \circ \Phi(y) = \overline{\Phi(x)(y)} = \langle x, y \rangle = \Phi(y)(x) = J(x)(\Phi(y)), \tag{9.4}$$

dove abbiamo usato (9.3), (8.3) e (9.1). Siccome $\Phi : H \to H'$ è suriettiva, per l'arbitrarietà di $y \in H$ deduciamo da (9.4) che

$$L = J(x) \in J(X).$$

Per l'arbitrarietà di $L \in H''$, questo dimostra che J è suriettiva. □

In maniera simile, ragionando con la mappa Ψ di (8.6)-(8.11) invece che con Φ, è possibile dimostrare il seguente risultato, di cui omettiamo la dimostrazione.

Proposizione 9.8. *Valgono i seguenti risultati.*

(i) Sia $(M, \mathcal{M}, \mu)$ uno uno spazio di misura σ-finito. Se $p \in (1, \infty)$, allora $L^p(M, \mathcal{M}, \mu)$ è riflessivo.

(ii) Se $p \in (1, \infty)$, lo spazio ℓ^p è riflessivo.

(iii) Gli spazi ℓ^∞, ℓ^1 e c_0 non sono riflessivi.

Osservazione 9.9. Siccome (☞ osservazione 8.14)

$$(c_0)'' \underset{\text{isom.}}{\cong} (\ell^1)' \underset{\text{isom.}}{\cong} \ell^\infty$$

e ℓ^∞ non è separabile, mentre c_0 lo è (☞ 8.42), si esclude che c_0 sia riflessivo. La non-riflessività di ℓ^1 e ℓ^∞ si può dedurre dal seguente risultato, di cui omettiamo la dimostrazione.

Proposizione 9.10. *Sia X uno spazio di Banach. Allora X è riflessivo se e solo se X' è riflessivo.*

Osservazione 9.11. Ci si può chiedere se i risultati negativi della proposizione 9.8(iii) abbiano un analogo per $L^p(M, \mathcal{M}, \mu)$ per più generali spazi di misura.

(1) Se M è un insieme finito, si ha $\dim L^p(M, \mathcal{M}, \mu) < \infty$ e quindi sia $L^1(M)$ che $L^\infty(M)$ sono banalmente riflessivi in questo caso (☞ osservazione 9.6).

(2) Se invece $M = [a, b]$ è un intervallo limitato di $\mathbb{R}$, sappiamo (☞ osservazione 8.43) che $L^1(a, b)$ è separabile, mentre $L^\infty(a, b) \underset{\text{isom.}}{\cong} (L^1(a, b))'$ non lo è, pertanto nemmeno $(L^\infty(a, b))' \underset{\text{isom.}}{\cong} (L^1(a, b))''$ è separabile (☞ proposizione 8.41). Da questo deduciamo che $L^1(a, b)$ non è riflessivo, e pertanto (☞ proposizione 9.10) nemmeno $L^\infty(a, b)$ è riflessivo.

9.2 Convergenza debole

Siccome ogni spazio normato $(X, \|\cdot\|_X)$ è uno spazio metrico, abbiamo già una nozione di convergenza per successioni a valori in X, e si ha

$$x_n \xrightarrow[n\to\infty]{} x \quad \text{in } X \qquad \iff \qquad \|x_n - x\|_X \xrightarrow[n\to\infty]{} 0$$

(☞ osservazione 2.8). Questa nozione di convergenza è anche detta *convergenza in norma*.

Introduciamo ora una diversa nozione di convergenza su X, definita in termini del duale X' di X.

Definizione 9.12. Sia X uno spazio normato. Siano $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X e $x \in X$. Diciamo che la successione $(x_n)_n$ *converge debolmente* a x in X, e scriviamo

$$x_n \underset{n\to\infty}{\rightharpoonup} x \quad \text{in } X,$$

se

$$\varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x) \qquad \forall \varphi \in X'.$$

Notazione. Al posto di "$x_n \underset{n\to\infty}{\rightharpoonup} x$", scriviamo anche "$x_n \rightharpoonup x$ per $n \to \infty$", oppure semplicemente "$x_n \rightharpoonup x$", se non c'è ambiguità.

Proposizione 9.13. *Sia X uno spazio normato. Sia $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X. Siano $x, x' \in X$.*

(i) (unicità del limite debole) *Se* $x_n \rightharpoonup x$ *e* $x_n \rightharpoonup x'$ *in* X*, allora* $x = x'$*.*

(ii) (conv. in norma $\Rightarrow$ conv. debole) *Se* $x_n \to x$ *in* X*, allora* $x_n \rightharpoonup x$ *in* X*.*

Dimostrazione. (i). Per ogni $\varphi \in X'$, per la definizione di convergenza debole, si ha

$$\varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x), \qquad \varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x')$$

in $\mathbb{F}$; per l'unicità del limite in $\mathbb{F}$ ne deduciamo che $\varphi(x) = \varphi(x')$ per ogni $\varphi \in X'$. Siccome i funzionali $\varphi \in X'$ separano i punti di X (☞ corollario 8.36(iii)), concludiamo che $x = x'$.

(ii). Se $x_n \to x$, siccome ogni $\varphi \in X' = \mathcal{B}(X, \mathbb{F})$ è una mappa continua, ne deduciamo (☞ proposizione 2.29) che $\varphi(x_n) \to \varphi(x)$ in $\mathbb{F}$; per l'arbitrarietà di $\varphi \in X'$, si ha dunque $x_n \rightharpoonup x$. □

Osservazione 9.14. Sia $X = \mathbb{F}^d$ con la norma euclidea (☞ esempio 4.5(1)). Siano $x^{(n)} = (x_1^{(n)}, \dots, x_d^{(n)})$, $x = (x_1, \dots, x_d) \in \mathbb{F}^d$. Allora

$$\|x^{(n)} - x\|_2 = \sqrt{\sum_{j=1}^{d} |x_j^{(n)} - x_j|^2}.$$

Da questa relazione si vede facilmente che

$$\|x^{(n)} - x\|_2 \xrightarrow[n\to\infty]{} 0$$

se e solo se

$$|x_j^{(n)} - x_j| \xrightarrow[n\to\infty]{} 0 \qquad \text{per ogni } j = 1, \dots, d,$$

cioè la convergenza in norma in $\mathbb{F}^d$ equivale alla *convergenza componente per componente* (un risultato analogo si trova nella proposizione 2.13 sugli spazi prodotto).

D'altra parte, per ogni $j = 1, \dots, d$, il funzionale di proiezione sulla j-esima componente

$$\psi_j : \mathbb{F}^d \ni x \mapsto x_j \in \mathbb{F}$$

è lineare e continuo su $\mathbb{F}^d$ (infatti $|\psi_j(x)| = |x_j| \leq \|x\|_2$). Applicando la definizione 9.12 con $\varphi = \psi_j$ per $j = 1, \dots, d$, ne deduciamo che la convergenza debole in $\mathbb{F}^d$ implica la convergenza componente per componente. Siccome quest'ultima è equivalente alla convergenza in norma, la quale a sua volta implica la convergenza debole (☞ proposizione 9.13(ii)), ne concludiamo che <u>in $\mathbb{F}^d$ la convergenza in norma e la convergenza debole sono equivalenti</u>.

Usando l'equivalenza delle norme (☞ teorema 4.13), questo fatto si estende a ogni spazio normato X di dimensione finita. Potremo dunque vedere una differenza fra la convergenza in norma e la convergenza debole solo in spazi di dimensione infinita.

Osservazione 9.15. Sia $X = \ell^2$. Siano $\underline{e^{(n)}} \in \ell^2$ definiti come nell'esempio 6.34(1) per ogni $n \in \mathbb{N}$. Vediamo ora che la successione $(\underline{e^{(n)}})_{n\in\mathbb{N}}$ converge debolmente a $\underline{0}$ in ℓ^2. Infatti, per ogni $\varphi \in (\ell^2)'$, per il teorema di Riesz–Fréchet (☞ teorema 8.7) possiamo scrivere $\varphi = \langle \cdot, \underline{y} \rangle$ per qualche $\underline{y} \in \ell^2$. In particolare

$$\varphi(\underline{e^{(n)}}) = \langle \underline{e^{(n)}}, \underline{y} \rangle = \sum_{k=0}^{\infty} e_k^{(n)} \overline{y_k} = \overline{y_n} \to 0$$

per $n \to \infty$, dato che $\underline{y} \in \ell^2 \subseteq c_0$ e quindi $y_n \to 0$ per $n \to \infty$. Per l'arbitrarietà di $\varphi \in (\ell^2)'$, questo dimostra la convergenza debole di$(\underline{e^{(n)}})_{n\in\mathbb{N}}$ a $\underline{0}$. D'altra parte

$$\|\underline{e^{(n)}} - \underline{0}\|_2 = \|\underline{e^{(n)}}\|_2 = 1 \not\to 0$$

per $n \to \infty$. Abbiamo dunque dimostrato che

$$\underline{e^{(n)}} \not\to \underline{0}, \qquad \text{ma} \qquad \underline{e^{(n)}} \rightharpoonup \underline{0}$$

in ℓ^2 per $n \to \infty$. Questo mostra che l'implicazione della proposizione 9.13(ii) in generale non si può invertire, e che quindi effettivamente "convergenza in norma" e "convergenza debole" in generale sono nozioni diverse.

Stabilire se una successione converge debolmente oppure no applicando direttamente la definizione può non essere facile, dato che, in principio, richiede di "testare" la convergenza su un qualunque elemento φ del duale (☞ definizione 9.12). Le caratterizzazioni dei duali di particolari spazi normati stabilite nel capitolo 8 risultano dunque particolarmente utili in questo contesto, dato che permettono di rendere più "concreta" la nozione di convergenza debole.

Ad esempio, abbiamo già visto nell'osservazione 9.15 che il teorema di Riesz–Fréchet risulta utile per discutere la convergenza debole in spazi di Hilbert, dato che tale teorema fornisce una descrizione esplicita di tutti gli elementi del duale (☞ teorema 8.7). In effetti, procedendo come nell'osservazione 9.15, si può facilmente dimostrare il seguente risultato.

Proposizione 9.16. *Sia H uno spazio di Hilbert. Siano $(x_n)_{n\in\mathbb{N}}$ una successione a valori in H e $x \in H$. Allora*

$$x_n \xrightharpoonup[n\to\infty]{} x \text{ in } H \qquad \iff \qquad \langle x_n, y\rangle \xrightarrow[n\to\infty]{} \langle x, y\rangle \quad \forall y \in H.$$

In maniera simile, utilizzando la caratterizzazione dei duali degli spazi di successioni e degli spazi L^p (☞ teoremi 8.10, 8.13, 8.15) si ottiene il seguente risultato.

Proposizione 9.17. *Valgono le seguenti proprietà.*

(i) *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Siano $p, q \in [1, \infty]$ esponenti coniugati con $p < \infty$. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^p(M)$ e $f \in L^p(M)$. Allora*

$$f_n \xrightharpoonup[n\to\infty]{} f \text{ in } L^p(M) \qquad \iff \qquad \int_M f_n\, g\, d\mu \xrightarrow[n\to\infty]{} \int_M f\, g\, d\mu \quad \forall g \in L^q(M).$$

(ii) *Siano $p, q \in [1, \infty]$ esponenti coniugati con $p < \infty$. Siano $(\underline{x^{(n)}})_{n\in\mathbb{N}}$ una successione a valori in ℓ^p e $\underline{x} \in \ell^p$. Allora*

$$\underline{x^{(n)}} \xrightharpoonup[n\to\infty]{} \underline{x} \text{ in } \ell^p \qquad \iff \qquad \sum_{k=0}^{\infty} x_k^{(n)} y_k \xrightarrow[n\to\infty]{} \sum_{k=0}^{\infty} x_k y_k \quad \forall \underline{y} \in \ell^q.$$

(iii) *Siano $(\underline{x^{(n)}})_{n\in\mathbb{N}}$ una successione a valori in c_0 e $\underline{x} \in c_0$. Allora*

$$\underline{x^{(n)}} \xrightharpoonup[n\to\infty]{} \underline{x} \text{ in } c_0 \qquad \iff \qquad \sum_{k=0}^{\infty} x_k^{(n)} y_k \xrightarrow[n\to\infty]{} \sum_{k=0}^{\infty} x_k y_k \quad \forall \underline{y} \in \ell^1.$$

In ogni spazio normato, sappiamo che una successione convergente in norma è limitata (☞ proposizione 2.23(v)). Vediamo ora che, come conseguenza del teorema di Banach–Steinhaus, una proprietà simile vale per la convergenza debole.

Proposizione 9.18. *Sia X uno spazio normato. Siano $(x_n)_{n\in\mathbb{N}}$ una successione a valori in X e $x \in X$. Se $x_n \rightharpoonup x$ in X per $n \to \infty$, allora la successione $(x_n)_{n\in\mathbb{N}}$ è limitata in X e*

$$\|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X \leq \sup_{n\in\mathbb{N}} \|x_n\|_X.$$

Dimostrazione. Sia $J : X \to X''$ l'immersione canonica (☞ proposizione 9.3). Allora, per ogni $\varphi \in X'$,

$$J(x_n)(\varphi) = \varphi(x_n) \to \varphi(x) = J(x)(\varphi)$$

per $n \to \infty$, dove abbiamo usato che $x_n \rightharpoonup x$ (☞ definizione 9.12). In altre parole, gli operatori $J(x_n) \in \mathcal{B}(X', \mathbb{F})$ convergono puntualmente all'operatore $J(x) \in \mathcal{B}(X', \mathbb{F})$. Siccome X' è uno spazio di Banach (☞ osservazione 8.3), possiamo allora applicare il corollario 7.30 del teorema di Banach–Steinhaus con $T_n = J(x_n)$ e $T = J(x)$, ottenendo che

$$\sup_{n\in\mathbb{N}} \|J(x_n)\|_{X''} < \infty \quad \text{e} \quad \|J(x)\|_{X''} \leq \liminf_{n\to\infty} \|J(x_n)\|_{X''} \leq \sup_{n\in\mathbb{N}} \|J(x_n)\|_{X''}.$$

Siccome $J : X \to X''$ è un'isometria lineare (☞ proposizione 9.3), ne deduciamo che

$$\sup_{n\in\mathbb{N}} \|x_n\|_X < \infty \quad \text{e} \quad \|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X \leq \sup_{n\in\mathbb{N}} \|x_n\|_X,$$

come desiderato. □

La proposizione precedente, che afferma che la limitatezza è una condizione necessaria per la convergenza debole di una successione, si può "completare" a una caratterizzazione della convergenza debole, enunciata nel seguente risultato.

Proposizione 9.19. *Sia X uno spazio normato. Sia $\Delta \subseteq X'$ tale che*

$$\overline{\operatorname{span} \Delta}^{X'} = X'. \tag{9.5}$$

Siano $(x_n)_n$ una successione a valori in X e $x \in X$. Sono equivalenti:

(i) $x_n \xrightharpoonup[n\to\infty]{} x$ *in X;*

(ii) la successione $(x_n)_n$ è limitata in X e $\varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x)$ per ogni $\varphi \in \Delta$.

Dimostrazione. (i) ⇒ (ii). La limitatezza di $(x_n)_n$ segue dalla proposizione 9.18, e il fatto che $\varphi(x_n) \to \varphi(x)$ per ogni $\varphi \in X'$ (dunque a maggior ragione per ogni $\varphi \in \Delta$) è la definizione di convergenza debole (☞ definizione 9.12).

(ii) ⇒ (i). Definiamo l'insieme

$$V = \left\{ \varphi \in X' : \varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x) \right\}. \tag{9.6}$$

In base alla definizione 9.12, verificare che $x_n \rightharpoonup x$ in X è equivalente a dimostrare che $V = X'$. Chiaramente per ipotesi $\Delta \subseteq V$.

Vediamo ora che V è un sottospazio vettoriale di X'. Infatti, chiaramente $0 \in V$. Inoltre, se $\varphi_1, \varphi_2 \in V$ e $\alpha_1, \alpha_2 \in \mathbb{F}$, si ha

$$\varphi_1(x_n) \to \varphi_1(x), \qquad \varphi_2(x_n) \to \varphi_2(x)$$

per $n \to \infty$, dunque anche

$$(\alpha_1\varphi_1 + \alpha_2\varphi_2)(x_n) = \alpha_1\varphi_1(x_n) + \alpha_2\varphi_2(x_n) \to \alpha_1\varphi_1(x) + \alpha_2\varphi_2(x),$$

dove si sono usate la definizione delle operazioni su X' e l'algebra dei limiti. Da questo segue che $\alpha_1\varphi_1 + \alpha_2\varphi_2 \in V$. Per la proposizione 1.4 deduciamo che V è un sottospazio vettoriale di X'.

Infine, dimostriamo che V è un sottoinsieme chiuso di X'. Siano $(\varphi_k)_{k\in\mathbb{N}}$ una successione a valori in V e $\varphi \in X'$ tali che $\varphi_k \to \varphi$ in X' per $k \to \infty$; vogliamo dimostrare che $\varphi \in V$. Osseviamo che, per ogni $k, n \in \mathbb{N}$,

$$\begin{aligned}|\varphi(x_n) - \varphi(x)| &\leq |\varphi(x_n) - \varphi_k(x_n)| + |\varphi_k(x_n) - \varphi_k(x)| + |\varphi_k(x) - \varphi(x)| \\ &\leq M\|\varphi - \varphi_k\|_{\mathrm{op}} + |\varphi_k(x_n) - \varphi_k(x)|, \end{aligned} \tag{9.7}$$

dove $M \coloneqq \sup_{n\in\mathbb{N}} \|x_n\|_X + \|x\|_X < \infty$, dato che $(x_n)_n$ è limitata. Preso ora $\epsilon > 0$, siccome $\varphi_k \xrightarrow[k\to\infty]{} \varphi$ in X', esiste $K \in \mathbb{N}$ tale che $\|\varphi_k - \varphi\|_{\mathrm{op}} < \epsilon/(2M)$ per ogni $k > K$. Fissato un tale $k > K$, siccome $\varphi_k \in V$, si ha $\varphi_k(x_n) \xrightarrow[n\to\infty]{} \varphi_k(x)$, dunque esiste $N \in \mathbb{N}$ per cui $|\varphi_k(x_n) - \varphi_k(x)| < \epsilon/2$ per ogni $n > N$. Ma allora, per ogni $n > N$, da (9.7) deduciamo che

$$|\varphi(x_n) - \varphi(x)| < M\epsilon/(2M) + \epsilon/2 = \epsilon.$$

Siccome $\epsilon > 0$ è arbitrario, ne segue che $\varphi(x_n) \to \varphi(x)$, cioè $\varphi \in V$. Dalla proposizione 2.23(vi) deduciamo che V è chiuso in X'.

Siccome $V \supseteq \Delta$ e inoltre V è un sottospazio vettoriale di X', deduciamo che $V \supseteq \operatorname{span} \Delta$; ma allora, siccome V è un sottoinsieme chiuso di X', si ha anche $V \supseteq \overline{\operatorname{span} \Delta} = X'$, dove si è usata l'ipotesi (9.5). In conclusione $V = X'$, come desiderato. □

Dalla precedente proposizione, utilizzando opportune scelte del sottoinsieme Δ del duale, deduciamo ulteriori caratterizzazioni della convergenza debole nel caso degli spazi di successioni e degli spazi L^p.

Corollario 9.20. *Sia $p \in (1, \infty)$.*

(i) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in ℓ^p e $\underline{x} \in \ell^p$. Allora

$$\underline{x}^{(n)} \xrightharpoonup[n\to\infty]{} \underline{x} \text{ in } \ell^p \iff \begin{cases} \sup_{n\in\mathbb{N}} \|\underline{x}^{(n)}\|_p < \infty, \\ x_k^{(n)} \xrightarrow[n\to\infty]{} x_k \qquad \forall k \in \mathbb{N}. \end{cases}$$

(ii) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in c_0 e $\underline{x} \in c_0$. Allora

$$\underline{x}^{(n)} \xrightharpoonup[n\to\infty]{} \underline{x} \text{ in } c_0 \iff \begin{cases} \sup_{n\in\mathbb{N}} \|\underline{x}^{(n)}\|_\infty < \infty, \\ x_k^{(n)} \xrightarrow[n\to\infty]{} x_k \qquad \forall k \in \mathbb{N}. \end{cases}$$

(iii) *Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^p(M)$ e $f \in L^p(M)$. Allora*

$$f_n \xrightarrow[n\to\infty]{} f \text{ in } L^p(M) \iff \begin{cases} \sup_{n\in\mathbb{N}} \|f_n\|_{L^p(M)} < \infty, \\ \int_E f_n \, d\mu \xrightarrow[n\to\infty]{} \int_E f \, d\mu \\ \quad \text{per ogni } E \in \mathcal{M} \text{ con } \mu(E) < \infty. \end{cases}$$

(iv) *Sia I un intervallo in $\mathbb{R}$, dotato della misura di Lebesgue. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^p(I)$ e $f \in L^p(I)$. Allora*

$$f_n \xrightarrow[n\to\infty]{} f \text{ in } L^p(I) \iff \begin{cases} \sup_{n\in\mathbb{N}} \|f_n\|_{L^p(I)} < \infty, \\ \int_a^b f_n(t) \, dt \xrightarrow[n\to\infty]{} \int_a^b f(t) \, dt \\ \quad \text{per ogni intervallo limitato } [a,b] \subseteq I. \end{cases}$$

Dimostazione (cenno). (i) si deduce applicando la proposizione 9.19 all'insieme $\Delta = \Psi(\{\underline{e}^{(n)}\}_{n\in\mathbb{N}})$, dove $\Psi : \ell^q \to (\ell^p)'$ è l'isomorfismo del teorema 8.10 e $q \in (1, \infty)$ è l'esponente coniugato; dato che

$$\operatorname{span}\{\underline{e}^{(n)}\}_{n\in\mathbb{N}} = c_{00}$$

è denso in ℓ^q (☞ osservazione 6.35 e proposizione 5.17(iii)), e Ψ è un isomorfismo, l'insieme $\operatorname{span} \Delta$ è denso in $(\ell^p)'$.

La dimostrazione di (ii) è perfettamente analoga, usando il teorema 8.13 invece del teorema 8.10.

Anche (iii) è dimostrato in maniera simile. Qui si usa l'isomorfismo $\Psi : L^q(M) \to (L^p(M))'$ del teorema 8.15, e il fatto che l'insieme

$$\operatorname{span}\{\mathbf{1}_E : E \in \mathcal{M},\ \mu(E) < \infty\}$$

è denso in $L^q(M)$ (☞ teorema 5.37(i)).

Infine, per (iv), si utilizza il fatto che l'insieme

$$\operatorname{span}\{\mathbf{1}_{[a,b]} : [a,b] \subseteq I \text{ intervallo limitato}\}$$

è denso in $L^q(I)$ (☞ teorema 5.37(iii)). □

9.3 Convergenza debole∗ sul duale

Sia X uno spazio di Banach. Il suo duale X' è a sua volta uno spazio di Banach, quindi possiamo applicare ad esso le nozioni di convergenza di successioni discusse finora: specificamente, per una successione $(\varphi_n)_n$ a valori in X' e un elemento $\varphi \in X'$, possiamo parlare di *convergenza in norma*,

$$\varphi_n \xrightarrow[n\to\infty]{} \varphi \quad \text{in } X' \qquad \iff \qquad \|\varphi_n - \varphi\|_{X'} \xrightarrow[n\to\infty]{} 0,$$

oppure di *convergenza debole*,

$$\varphi_n \xrightarrow[n\to\infty]{} \varphi \quad \text{in } X' \qquad \iff \qquad \Lambda(\varphi_n) \xrightarrow[n\to\infty]{} \Lambda(\varphi) \quad \forall \Lambda \in X''.$$

Introduciamo ora una terza nozione di convergenza per successioni in X'.

Definizione 9.21. Sia X uno spazio di Banach. Siano $(\varphi_n)_{n\in\mathbb{N}}$ una successione a valori in X' e $\varphi \in X'$. Diciamo che la successione $(\varphi_n)_n$ *converge debolmente∗* a φ in X', e scriviamo

$$\varphi_n \xrightarrow[n\to\infty]{*} \varphi \quad \text{in } X',$$

se

$$\varphi_n(x) \xrightarrow[n\to\infty]{} \varphi(x) \qquad \forall x \in X.$$

Notazione. Al posto di "$\varphi_n \xrightarrow[n\to\infty]{*} \varphi$", scriviamo anche "$\varphi_n \overset{*}{\rightharpoonup} \varphi$ per $n \to \infty$", oppure semplicemente "$\varphi_n \overset{*}{\rightharpoonup} \varphi$", se non c'è ambiguità.

Osservazione 9.22. In base alla definizione 9.12, la convergenza debole∗ $\varphi_n \overset{*}{\rightharpoonup} \varphi$ di una successione di funzionali non è altro che la convergenza puntuale (☞ definizione 2.16) delle funzioni $\varphi_n : X \to \mathbb{F}$ alla funzione $\varphi : X \to \mathbb{F}$.

Proposizione 9.23. *Sia X uno spazio di Banach. Sia $(\varphi_n)_{n\in\mathbb{N}}$ una successione a valori in φ. Siano $\varphi, \varphi' \in X'$.*

(i) (unicità del limite debole∗) *Se $\varphi_n \overset{*}{\rightharpoonup} \varphi$ e $\varphi_n \overset{*}{\rightharpoonup} \varphi'$ in X', allora $\varphi = \varphi'$.*

(ii) (conv. debole ⇒ conv. debole∗) *Se $\varphi_n \rightharpoonup \varphi$ in X', allora $\varphi_n \overset{*}{\rightharpoonup} \varphi$ in X'. L'implicazione inversa vale quando X è riflessivo.*

Dimostrazione. (i). Per ogni $x \in X$, per la definizione di convergenza debole∗, si ha

$$\varphi_n(x) \xrightarrow[n\to\infty]{} \varphi(x), \qquad \varphi_n(x) \xrightarrow[n\to\infty]{} \varphi'(x)$$

in $\mathbb{F}$; per l'unicità del limite in $\mathbb{F}$ ne deduciamo che $\varphi(x) = \varphi'(x)$ per ogni $x \in X$, cioè $\varphi = \varphi'$.

(ii). Sia $J : X \to X''$ l'immersione canonica. Se $\varphi_n \rightharpoonup \varphi$ in X', per definizione di convergenza debole si ha

$$\Lambda(\varphi_n) \xrightarrow[n\to\infty]{} \Lambda(\varphi)$$

per ogni $\Lambda \in X''$. In particolare, per ogni $x \in X$, se prendiamo $\Lambda = J(x)$, otteniamo che

$$\varphi_n(x) = J(x)(\varphi_n) \xrightarrow[n\to\infty]{} J(x)(\varphi) = \varphi(x),$$

dove si è usata la definizione (9.1) di J; per l'arbitrarietà di $x \in X$, abbiamo dimostrato che $\varphi_n \overset{*}{\rightharpoonup} \varphi$ in X'.

Se X è riflessivo, cioè $J : X \to X''$ è suriettiva, possiamo invertire il ragionamento sopra. Infatti, se $\varphi_n \overset{*}{\rightharpoonup} \varphi$ in X', allora

$$\varphi_n(x) \xrightarrow[n\to\infty]{} \varphi(x)$$

per ogni $x \in X$. Siccome $J : X \to X''$ è suriettiva, ogni $\Lambda \in X''$ ha la forma $\Lambda = J(x)$ per qualche $x \in X$, dunque

$$\Lambda(\varphi_n) = J(x)(\varphi_n) = \varphi(x_n) \xrightarrow[n\to\infty]{} \varphi(x) = J(x)(\varphi) = \Lambda(\varphi);$$

per l'arbitrarietà di $\Lambda \in X''$, abbiamo dimostrato che $\varphi_n \rightharpoonup \varphi$ in X'. □

Osservazione 9.24. La proposizione 9.23(ii) mostra che la convergenza debole$*$ su X' può differire dalla convergenza debole già discussa soltanto quando X non è riflessivo.

Abbiamo visto (☞ teorema 8.15) che, se $(M, \mathcal{M}, \mu)$ è uno spazio di misura σ-finito, lo spazio $L^\infty(M)$ si può identificare con il duale di $L^1(M)$ tramite l'isomorfismo isometrico $\Psi : L^\infty(M) \to (L^1(M))'$ dato da (8.11). Possiamo allora parlare di "convergenza debole$*$ in $L^\infty(M)$", nel senso che possiamo "trapiantare" su $L^\infty(M)$ la nozione di convergenza debole$*$ in $(L^1(M))'$ tramite l'isomorfismo Ψ; in altre parole, per una successione $(f_n)_n$ di funzioni in $L^\infty(M)$,

$$f_n \xrightarrow[n\to\infty]{*} f \text{ in } L^\infty(M) \quad\iff\quad \Psi(f_n) \xrightarrow[n\to\infty]{*} \Psi(f) \text{ in } (L^1(M))'.$$

Questo vale in particolare nel caso di ℓ^∞ e $(\ell^1)'$, cioè possiamo parlare di "convergenza debole$*$ in ℓ^∞": per una successione $(\underline{x}^{(n)})_n$ a valori in ℓ^∞,

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^\infty \quad\iff\quad \Psi(\underline{x}^{(n)}) \xrightarrow[n\to\infty]{*} \Psi(\underline{x}) \text{ in } (\ell^1)',$$

dove $\Psi : \ell^\infty \to (\ell^1)'$ è l'isomorfismo isometrico del teorema 8.10.

In maniera simile, parliamo di "convergenza debole$*$ in ℓ^1" mediante l'identificazione di ℓ^1 con il duale di c_0 (☞ teorema 8.13): dunque, per una successione $(\underline{x}^{(n)})_n$ a valori in ℓ^1,

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^1 \quad\iff\quad \Psi(\underline{x}^{(n)}) \xrightarrow[n\to\infty]{*} \Psi(\underline{x}) \text{ in } (c_0)',$$

dove $\Psi : \ell^1 \to (c_0)'$ è l'isomorfismo isometrico dato da (8.10).

Esplicitando in questi casi la definizione di convergenza debole$*$, possiamo ottenere il seguente risultato, analogo e complementare alla proposizione 9.17.

Proposizione 9.25. *Valgono le seguenti proprietà.*

(i) Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^\infty(M)$ e $f \in L^\infty(M)$. Allora

$$f_n \xrightarrow[n\to\infty]{*} f \text{ in } L^\infty(M) \quad\iff\quad \int_M f_n\, g\, d\mu \xrightarrow[n\to\infty]{} \int_M f\, g\, d\mu \quad \forall g \in L^1(M).$$

(ii) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in ℓ^∞ e $\underline{x} \in \ell^\infty$. Allora

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^\infty \quad\iff\quad \sum_{k=0}^\infty x_k^{(n)} y_k \xrightarrow[n\to\infty]{} \sum_{k=0}^\infty x_k y_k \quad \forall \underline{y} \in \ell^1.$$

(iii) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in ℓ^1 e $\underline{x} \in \ell^1$. Allora

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^1 \quad\iff\quad \sum_{k=0}^\infty x_k^{(n)} y_k \xrightarrow[n\to\infty]{} \sum_{k=0}^\infty x_k y_k \quad \forall \underline{y} \in c_0.$$

Osservazione 9.26. Confrontando la proposizione 9.25(iii) con il caso $p = 1$ della proposizione 9.17(ii), si può apprezzare la differenza fra le convergenze debole e debole$*$ in ℓ^1: nel primo caso, la convergenza di una successione è "testata" su ogni $\underline{y} \in \ell^\infty$, nel secondo caso solo su $\underline{y} \in c_0$. Questo è coerente con l'implicazione fra le convergenze debole e debole$*$ dimostrata nella proposizione 9.23(ii); in effetti, è possibile trovare esempi di successioni in ℓ^1 che mostrano che l'implicazione non si può invertire in generale.

Osservazione 9.27. Come fatto per $L^\infty(M)$, si potrebbe parlare di "convergenza debole$*$ in $L^q(M)$" anche per $q \in (1,\infty)$, identificando $L^q(M)$ con $(L^p(M))'$ tramite l'isomorfismo isometrico del teorema 8.15 (qui $p \in (1,\infty)$ è l'esponente coniugato di q). Tuttavia, dato che $L^p(M)$ è riflessivo (☞ proposizione 9.8), in questo caso la convergenza debole e la convergenza debole$*$ su $L^q(M)$ coincidono (☞ proposizione 9.23(ii)).

Enunciamo ora gli analoghi per la convergenza debole$*$ delle proposizioni 9.18 e 9.19. Le dimostrazioni sono simili e lasciate per esercizio.

Proposizione 9.28. *Sia X uno spazio di Banach. Siano $(\varphi_n)_{n\in\mathbb{N}}$ una successione a valori in X' e $\varphi \in X'$. Se $\varphi_n \overset{*}{\rightharpoonup} \varphi$ in X' per $n \to \infty$, allora la successione $(\varphi_n)_{n\in\mathbb{N}}$ è limitata in X' e*

$$\|\varphi\|_{X'} \leq \liminf_{n\to\infty} \|\varphi_n\|_{X'} \leq \sup_{n\in\mathbb{N}} \|\varphi_n\|_{X'}.$$

Proposizione 9.29. *Sia X uno spazio di Banach. Sia $\Delta \subseteq X$ tale che*

$$\overline{\operatorname{span} \Delta}^X = X.$$

Siano $(\varphi_n)_n$ una successione a valori in X' e $\varphi \in X'$. Sono equivalenti:

(i) $\varphi_n \xrightarrow[n\to\infty]{*} \varphi$ *in* X';

(ii) la successione $(\varphi_n)_n$ è limitata in X' e $\varphi_n(x) \xrightarrow[n\to\infty]{} \varphi(x)$ per ogni $x \in \Delta$.

Applicando la proposizione 9.29 con opportune scelte di $\Delta \subseteq X$, possiamo ottenere le seguenti caratterizzazioni della convergenza debole$*$ in casi particolari, che complementano quelle per la convergenza debole del corollario 9.20.

Corollario 9.30. *Valgono le seguenti proprietà.*

(i) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in ℓ^∞ e $\underline{x} \in \ell^\infty$. Allora

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^\infty \iff \begin{cases} \sup_{n\in\mathbb{N}} \|\underline{x}^{(n)}\|_\infty < \infty, \\ x_k^{(n)} \xrightarrow[n\to\infty]{} x_k \qquad \forall k \in \mathbb{N}. \end{cases}$$

(ii) Siano $(\underline{x}^{(n)})_{n\in\mathbb{N}}$ una successione a valori in ℓ^1 e $\underline{x} \in \ell^1$. Allora

$$\underline{x}^{(n)} \xrightarrow[n\to\infty]{*} \underline{x} \text{ in } \ell^1 \iff \begin{cases} \sup_{n\in\mathbb{N}} \|\underline{x}^{(n)}\|_1 < \infty, \\ x_k^{(n)} \xrightarrow[n\to\infty]{} x_k \qquad \forall k \in \mathbb{N}. \end{cases}$$

(iii) Sia $(M, \mathcal{M}, \mu)$ uno spazio di misura σ-finito. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^\infty(M)$ e $f \in L^\infty(M)$. Allora

$$f_n \xrightarrow[n\to\infty]{*} f \text{ in } L^\infty(M) \iff \begin{cases} \sup_{n\in\mathbb{N}} \|f_n\|_{L^\infty(M)} < \infty, \\ \int_E f_n \, d\mu \xrightarrow[n\to\infty]{} \int_E f \, d\mu \\ \quad \text{per ogni } E \in \mathcal{M} \text{ con } \mu(E) < \infty. \end{cases}$$

(iv) *Sia I un intervallo in $\mathbb{R}$, dotato della misura di Lebesgue. Siano $(f_n)_{n\in\mathbb{N}}$ una successione a valori in $L^\infty(I)$ e $f \in L^\infty(I)$. Allora*

$$f_n \xrightarrow[n\to\infty]{*} f \text{ in } L^\infty(I) \iff \begin{cases} \sup_{n\in\mathbb{N}} \|f_n\|_{L^\infty(I)} < \infty, \\ \int_a^b f_n(t)\,dt \xrightarrow[n\to\infty]{} \int_a^b f(t)\,dt \\ \quad \text{per ogni intervallo limitato } [a,b] \subseteq I. \end{cases}$$

Dimostriamo ora un risultato profondo, che si può interpretare come un risultato di "compattezza debole∗" per la palla unitaria chiusa del duale di uno spazio di Banach. Ricordiamo che, per il corollario 4.26, l'analogo risultato è falso in dimensione infinita se si considera la convergenza in norma anziché quella debole∗.

Teorema 9.31 (Banach–Alaoglu)**.** *Sia X uno spazio di Banach separabile. Sia $(\varphi_n)_{n\in\mathbb{N}}$ una successione limitata in X'. Allora $(\varphi_n)_n$ ha una sottosuccessione convergente debolmente∗ a qualche $\varphi \in X'$.*

Dimostrazione. Siccome $(\varphi_n)_{n\in\mathbb{N}}$ è limitata in X', si ha $M \coloneqq \sup_{n\in\mathbb{N}} \|\varphi_n\|_{\mathrm{op}} < \infty$; dunque, per ogni $x \in X$, si ha anche

$$\sup_{n\in\mathbb{N}} |\varphi_n(x)| \leq M\|x\|_X < \infty, \tag{9.8}$$

cioè la successione $(\varphi_n(x))_{n\in\mathbb{N}}$ è limitata in $\mathbb{F}$.

Siccome X è separabile, esiste un sottoinsieme denso $\Delta = \{x_n\}_{n\in\mathbb{N}}$ numerabile di X. Costruiamo ora induttivamente, per ogni $N \in \mathbb{N}$, una successione $(\varphi_n^{(N)})_{n\in\mathbb{N}}$ a valori in X' con le seguenti proprietà:

- $(\varphi_n^{(N)})_n$ è una sottosuccessione di $(\varphi_n)_n$ e, se $N > 0$, anche di $(\varphi_n^{(N-1)})_n$;
- la successione $(\varphi_n^{(N)}(x_N))_n$ converge in $\mathbb{F}$.

Per $N = 0$, osserviamo che per (9.8) la successione $(\varphi_n(x_0))_{n\in\mathbb{N}}$ è limitata in $\mathbb{F}$, dunque (per il teorema di Bolzano–Weierstrass) ha una sottosuccessione $(\varphi_{n_k}(x_0))_{k\in\mathbb{N}}$ convergente in $\mathbb{F}$; possiamo allora porre $\varphi_k^{(0)} = \varphi_{n_k}$ per ogni $k \in \mathbb{N}$.

Supponiamo ora, per $N > 0$, di avere costruito $(\varphi_n^{(0)})_n, \dots, (\varphi_n^{(N-1)})_n$. Notiamo che $(\varphi_n^{(N-1)})_n$ è una sottosuccessione di (φ_n), quindi anche $(\varphi_n^{(N-1)}(x_N))_n$ è una sottosuccessione di $(\varphi_n(x_N))_n$; dalla stima (9.8) deduciamo dunque che $(\varphi_n^{(N-1)}(x_N))_n$ è limitata in $\mathbb{F}$, ma allora (per il teorema di Bolzano–Weierstrass) ha una sottosuccessione $(\varphi_{n_k}^{(N-1)}(x_N))_{k\in\mathbb{N}}$ convergente in $\mathbb{F}$, e possiamo porre $\varphi_k^{(N)} = \varphi_{n_k}^{(N-1)}$ per ogni $k \in \mathbb{N}$.

Osserviamo ora che, siccome ogni successione $(\varphi_n^{(N)})_n$ è sottosuccessione della successione precedente $(\varphi_n^{(N-1)})_n$, abbiamo anche che

- $(\varphi_n^{(N)})_n$ è una sottosuccessione di $(\varphi_n^{(M)})_n$ per ogni $N \geq M \geq 0$.

Poniamo ora $\psi_n = \varphi_n^{(n)}$ per ogni $n \in \mathbb{N}$. Allora $(\psi_n)_{n\in\mathbb{N}}$ è una sottosuccessione di $(\varphi_n)_{n\in\mathbb{N}}$, e inoltre $(\psi_n)_{n\geq N}$ è una sottosuccessione di $(\varphi_n^{(N)})_{n\geq N}$ per ogni $N \in \mathbb{N}$. In particolare, per ogni $N \in \mathbb{N}$, $(\psi_n(x_N))_{n\geq N}$ è una sottosuccessione di $(\varphi_n^{(N)}(x_N))_{n\geq N}$, che è convergente in $\mathbb{F}$. Da questo deduciamo che

- $(\psi_n(x_N))_n$ converge in $\mathbb{F}$ per ogni $N \in \mathbb{N}$.

In altre parole $(\psi_n(x))_n$ converge in $\mathbb{F}$ per ogni $x \in \Delta$. Utilizzando la linearità delle ψ_n e l'algebra dei limiti, si deduce allora che $(\psi_n(x))_n$ converge in $\mathbb{F}$ per ogni $x \in V := \operatorname{span} \Delta$.

Possiamo dunque definire $\psi : V \to \mathbb{F}$ ponendo

$$\psi(x) = \lim_{n\to\infty} \psi_n(x) \qquad \forall x \in V. \tag{9.9}$$

Utilizzando la linearità delle ψ_n e l'algebra dei limiti, si verifica facilmente che anche $\psi \in \mathcal{L}(V, \mathbb{F})$. Ricordando poi che $(\psi_n)_n$ è una sottosuccessione di $(\varphi_n)_n$, si ha anche

$$\sup_n \|\psi_n\|_{\mathrm{op}} \leq \sup_n \|\varphi_n\|_{\mathrm{op}} = M < \infty. \tag{9.10}$$

Quindi, come in (9.8),

$$|\psi(x)| \leq \sup_{n\in\mathbb{N}} |\psi_n(x)| \leq M\|x\|_X$$

per ogni $x \in V$, il che dimostra che $\|\psi\|_{\mathrm{op}} \leq M < \infty$ (qui dotiamo il sottospazio V della norma indotta da X), e dunque $\psi \in \mathcal{B}(V, \mathbb{F})$. Siccome Δ è denso in X, a maggior ragione $V = \operatorname{span} \Delta$ è denso in X. Per il teorema di estensione (☞ teorema 7.23) concludiamo che ψ si estende per continuità a un funzionale $\varphi \in \mathcal{B}(X, \mathbb{F}) = X'$.

In conclusione, per (9.10) la successione $(\psi_n)_n$ è limitata in X', e inoltre, per (9.9), si ha

$$\varphi(x) = \lim_{n\to\infty} \psi_n(x) \qquad \forall x \in \Delta.$$

Siccome $\operatorname{span} \Delta$ è denso in X, possiamo applicare la proposizione 9.29 e dedurre che $\psi_n \overset{*}{\rightharpoonup} \psi$ in X' per $n \to \infty$. Dunque $(\psi_n)_n$ è la sottosuccessione di $(\varphi_n)_n$ cercata. □

Corollario 9.32. *Sia X uno spazio di Banach separabile e riflessivo. Sia $(x_n)_{n\in\mathbb{N}}$ una successione limitata in X. Allora $(x_n)_n$ ha una sottosuccessione convergente debolmente a qualche $x \in X$.*

Dimostrazione. Siccome X è separabile, anche il biduale X'' è separabile (dato che è isomorfo a X per l'ipotesi di riflessività); quindi, per il corollario 8.41 del teorema di Hahn–Banach, anche X' è separabile. Si può dunque applicare il teorema di Banach–Alaoglu (☞ teorema 9.31) allo spazio di Banach separabile X'.

Sia $(x_n)_n$ una successione limitata in X. Siccome l'immersione canonica $J : X \to X''$ è un'isometria, anche $(J(x_n))_n$ è limitata in X''. Per il teorema di Banach–Alaoglu (applicato a X'), esiste allora una sottosuccessione $(J(x_{n_k}))_{k\in\mathbb{N}}$ che converge a qualche $\Lambda \in X''$. Siccome $J : X \to X''$ è suriettiva (X è riflessivo), esiste $x \in X$ tale che $\Lambda = J(x)$. Ma allora, per ogni $\varphi \in X'$,

$$\varphi(x_{n_k}) = J(x_{n_k})(\varphi) \to J(x)(\varphi) = \varphi(x)$$

per $k \to \infty$, dove si è usato che $J(x_{n_k}) \overset{*}{\rightharpoonup} J(x)$ in X' e la definizione (9.1) di J. Per l'arbitrarietà di $\varphi \in X'$, deduciamo che $x_{n_k} \rightharpoonup x$ in X, come desiderato. □

Un esempio di applicazione del corollario 9.32 è il seguente risultato di minimizzazione di funzionali non lineari, che siano semicontinui inferiormente rispetto alla convergenza debole.

Proposizione 9.33. *Sia X uno spazio di Banach riflessivo e separabile. Sia $F : X \to \mathbb{R}$ tale che:*

(a) *esiste $c \in \mathbb{R}$ tale che l'insieme $\{x \in X : F(x) \leq c\}$ è non vuoto e limitato in X;*

(b) *per ogni successione $(x_n)_n$ a valori in X e $x \in X$, se $x_n \rightharpoonup x$ in X, allora $F(x) \leq \liminf_{n\to\infty} F(x_n)$.*

Allora F ha un punto di minimo in X.

Dimostrazione. Sia $B = \{x \in X : F(x) \leq c\}$. Siccome $B \neq \emptyset$, si ha $\inf F = \inf(F|_B) \leq c$.

Possiamo allora trovare in B una successione minimizzante per F; in altre parole, esiste una successione $(x_n)_n$ a valori in B tale che $(F(x_n))_n$ è decrescente e $\lim_{n\to\infty} F(x_n) = \inf F$. Siccome B è limitato in X, la successione $(x_n)_n$ è limitata; quindi possiamo applicare il corollario 9.32 del teorema di Banach–Alaoglu e dedurre che, a meno di sostituire $(x_n)_n$ con una sua sottosuccessione, si ha $x_n \rightharpoonup x$ per qualche $x \in X$. Ma allora

$$F(x) \leq \liminf_{n\to\infty} F(x_n) = \inf F,$$

da cui segue che $F(x) = \min F$. Quindi x è il punto di minimo cercato. □

10 Aggiunto e spettro in spazi di Hilbert

10.1 Aggiunto di un operatore limitato

Ricordiamo che $\mathbb{F}$ denota il campo dei numeri reali $\mathbb{R}$ oppure il campo dei numeri complessi $\mathbb{C}$.

Proposizione 10.1. *Siano H_1 e H_2 spazi di Hilbert su $\mathbb{F}$. Sia $A \in \mathcal{B}(H_1, H_2)$. Allora esiste un unico operatore $B \in \mathcal{B}(H_2, H_1)$ tale che*

$$\langle Bx, y\rangle_{H_1} = \langle x, Ay\rangle_{H_2} \qquad \forall x \in H_2 \ \forall y \in H_1. \tag{10.1}$$

Inoltre tale operatore B soddisfa $\|B\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}$.

Dimostrazione. Siano $A \in \mathcal{B}(H_1, H_2)$ e $B \in \mathcal{B}(H_2, H_1)$. Siano $F_A : H_1 \times H_2 \to \mathbb{F}$ e $F_B : H_2 \times H_1 \to \mathbb{F}$ le forme sesquilineari ad essi associate (☞ definizione 8.24). Si ha dunque, per ogni $x \in H_2$ e $y \in H_1$,

$$F_A(y, x) = \langle Ay, x\rangle_{H_2}, \qquad F_B(x, y) = \langle Bx, y\rangle_{H_1}$$

e quindi anche

$$F_A^*(x, y) := \overline{F_A(y, x)} = \langle x, Ay\rangle_{H_2}; \tag{10.2}$$

notiamo che $F_A^* : H_2 \times H_1 \to \mathbb{F}$ è a sua volta una forma sesquilineare.

Possiamo allora riscrivere la condizione (10.1) equivalentemente come

$$F_A^* = F_B.$$

In particolare, dato $A \in \mathcal{B}(H_1, H_2)$, per la proposizione 8.25 esiste un unico $B \in \mathcal{B}(H_2, H_1)$ la cui forma associata F_B sia uguale a F_A^*, cioè tale da soddisfare la condizione (10.1). Per tale operatore B si ha inoltre, per la proposizione 8.23,

$$\begin{aligned}\|B\|_{\mathrm{op}} = \|F_B\| = \|F_A^*\| &= \sup_{\|x\|_{H_1} \le 1, \|y\|_{H_2} \le 1} |F_A^*(x, y)| \\ &= \sup_{\|x\|_{H_1} \le 1, \|y\|_{H_2} \le 1} |F_A(y, x)| = \|F_A\| = \|A\|_{\mathrm{op}},\end{aligned}$$

come desiderato. □

Definizione 10.2. Siano H_1 e H_2 spazi di Hilbert. Sia $A \in \mathcal{B}(H_1, H_2)$. L'unico operatore $B \in \mathcal{B}(H_2, H_1)$ che soddisfa (10.1) è detto *aggiunto* dell'operatore A e si denota con A^*.

Osservazione 10.3. Supponiamo che $\dim H_1 = n < \infty$ e $\dim H_2 = m < \infty$. Siano $\{e_1, \ldots, e_n\}$ e $\{f_1, \ldots, f_m\}$ basi ortonormali di H_1 e H_2. Allora, per ogni $A \in \mathcal{B}(H_1, H_2)$, da (10.1) deduciamo che

$$\langle Ae_j, f_k\rangle_{H_2} = \overline{\langle f_k, Ae_j\rangle_{H_2}} = \overline{\langle A^* f_k, e_j\rangle_{H_1}} \tag{10.3}$$

per $j = 1, \ldots, n$ e $k = 1, \ldots, m$. Siccome

$$(\langle Ae_j, f_k\rangle_{H_2})_{k,j}, \qquad (\langle A^* f_k, e_j\rangle_{H_2})_{j,k}$$

sono le matrici che rappresentano A e A^* nelle basi scelte (☞ osservazione 1.24), la relazione (10.3) afferma che la matrice dell'operatore aggiunto A^* è la *trasposta coniugata* della matrice di A. (Nel caso $\mathbb{F} = \mathbb{R}$, pertanto, la matrice di A^* è semplicemente la trasposta della matrice di A.)

Esempio 10.4. Ecco alcuni esempi di aggiunti di operatori fra spazi di Hilbert.

(1) Se H è uno spazio di Hilbert, si ha $\mathrm{id}_H^* = \mathrm{id}_H$. Infatti, per ogni $x, y \in H$,

$$\langle \mathrm{id}_H x, y\rangle = \langle x, y\rangle = \langle x, \mathrm{id}_H y\rangle.$$

(2) Sia $H = \ell^2$. Sia $S : \ell^2 \to \ell^2$ l'operatore di *shift verso sinistra*, definito da

$$S\underline{x} = (x_{k+1})_{k\in\mathbb{N}} \qquad \forall \underline{x} = (x_k)_{k\in\mathbb{N}} \in \ell^2;$$

in altre parole,

$$S(x_0, x_1, x_2, x_3, \dots) = (x_1, x_2, x_3, \dots).$$

È facile verificare che $S \in \mathcal{B}(\ell^2)$ e $\|S\|_{\mathrm{op}} = 1$. Per calcolare l'aggiunto $S^* \in \mathcal{B}(\ell^2)$, osserviamo che è univocamente determinato dalla condizione (10.1), cioè

$$\langle S^*\underline{x}, \underline{y}\rangle = \langle \underline{x}, S\underline{y}\rangle = \sum_{k=0}^{\infty} x_k \overline{y_{k+1}} = \sum_{k=1}^{\infty} x_{k-1}\overline{y_k} \tag{10.4}$$

per ogni $\underline{x}, \underline{y} \in \ell^2$. Confrontando prima e ultima espressione in questa serie di uguaglianze, vediamo che allora, per ogni $\underline{x} \in \ell^2$ e $k \in \mathbb{N}$,

$$(S^*\underline{x})_k = \begin{cases} 0 & \text{se } k = 0, \\ x_{k-1} & \text{se } k > 0 \end{cases} \tag{10.5}$$

(più precisamente, si può dedurre (10.5) valutando l'identità (10.4) in $\underline{y} = \underline{e^{(k)}}$ per ogni $k \in \mathbb{N}$). In altre parole,

$$S^*(x_0, x_1, x_2, \dots) = (0, x_0, x_1, x_2, \dots),$$

cioè S^* è l'operatore di *shift verso destra con aggiunta di zero.*

(3) Per $\underline{w} \in \ell^\infty$, sia $D_{\underline{w}} \in \mathcal{B}(\ell^2)$ l'operatore di moltiplicazione per $\underline{w}$ definito nell'esempio 7.8(8). Per calcolarne l'aggiunto $D_{\underline{w}}^* \in \mathcal{B}(\ell^2)$, osserviamo che, per (10.1), per ogni $\underline{x}, \underline{y} \in \ell^2$,

$$\langle D_{\underline{w}}^*\underline{x}, \underline{y}\rangle = \langle \underline{x}, D_{\underline{w}}\underline{y}\rangle = \sum_{k=0}^{\infty} x_k \overline{w_k y_k} = \sum_{k=0}^{\infty} \overline{w_k} x_k \overline{y_k} = \langle D_{\overline{\underline{w}}}\underline{x}, \underline{y}\rangle,$$

dove $\overline{\underline{w}} = (\overline{w_k})_{k\in\mathbb{N}}$ è il coniugato componente per componente di $\underline{w}$. Chiaramente

$$\|\overline{\underline{w}}\|_\infty = \sup_{k\in\mathbb{N}} |\overline{w_k}| = \sup_{k\in\mathbb{N}} |w_k| = \|\underline{w}\|_\infty < \infty,$$

dunque $D_{\overline{\underline{w}}} \in \mathcal{B}(\ell^2)$ è l'aggiunto di $D_{\underline{w}}$.

(4) Siano $[a, b]$ e $[c, d]$ intervalli limitati in $\mathbb{R}$. Per $K \in L^2((a, b) \times (c, d))$, sia $T_K \in \mathcal{B}(L^2(c, d), L^2(a, b))$ l'operatore integrale con nucleo integrale K definito nell'esempio 7.8(6). Vogliamo calcolare l'aggiunto $T_K^* \in$

$\mathcal{B}(L^2(a,b), L^2(c,d))$. Osserviamo che, per (10.1), per ogni $f \in L^2(a,b)$ e $g \in L^2(c,d)$,

$$\begin{aligned}\langle T_K^* f, g\rangle_{L^2(c,d)} = \langle f, T_K g\rangle_{L^2(a,b)} &= \int_a^b f(x)\,\overline{T_K g(x)}\,dx \\ &= \int_a^b f(x)\,\overline{\int_c^d K(x,y)\,g(y)\,dy}\,dx \\ &= \int_a^b \int_c^d f(x)\,\overline{K(x,y)}\,\overline{g(y)}\,dy\,dx \\ &= \int_c^d \int_a^b \overline{K(x,y)}\,f(x)\,dx\,\overline{g(y)}\,dy,\end{aligned} \tag{10.6}$$

dove si sono usati la definizione (7.5) di operatore integrale e il teorema di Fubini–Tonelli (☞ teorema 3.9). Nell'ultimo termine, riconosciamo nell'integrale interno l'espressione di un operatore integrale con nucleo integrale

$$K^*(y,x) \coloneqq \overline{K(x,y)}; \tag{10.7}$$

in effetti dal teorema di Fubini–Tonelli è chiaro che

$$\|K^*\|_{L^2((c,d)\times(a,b))} = \|K\|_{L^2((a,b)\times(c,d))} < \infty,$$

quindi $T_{K^*} \in \mathcal{B}(L^2(a,b), L^2(c,d))$ e (10.6) si può riscrivere come

$$\langle T_K^* f, g\rangle_{L^2(c,d)} = \int_c^d T_{K^*} f(y)\,\overline{g(y)}\,dy = \langle T_{K^*} f, g\rangle_{L^2(c,d)}$$

per ogni $f \in L^2(a,b)$ e $g \in L^2(c,d)$. Ne concludiamo che

$$T_K^* = T_{K^*},$$

cioè l'aggiunto di un operatore integrale è a sua volta un operatore integrale, e i loro nuclei integrali sono legati da (10.7).

Discutiamo ora alcune proprietà fondamentali della nozione di aggiunto.

Proposizione 10.5. *Siano H_1, H_2, H_3 spazi di Hilbert.*

(i) *Se $A \in \mathcal{B}(H_1,H_2)$, allora $A^{**} = A$* ($A \mapsto A^*$ è involutiva)*.*

(ii) *Se $A, B \in \mathcal{B}(H_1,H_2)$ e $\alpha, \beta \in \mathbb{F}$, allora $(\alpha A + \beta B)^* = \overline{\alpha}A^* + \overline{\beta}B^*$* ($A \mapsto A^*$ è antilineare)*.*

(iii) *Se $A \in \mathcal{B}(H_1,H_2)$ e $B \in \mathcal{B}(H_2,H_3)$, allora $(BA)^* = A^*B^*$* ($A \mapsto A^*$ è antimoltiplicativa)*.*

(iv) *Se $A \in \mathcal{B}(H_1,H_2)$, allora $\|A^*A\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}^2$* (condizione C*)*.*

(v) *Se $A \in \mathcal{B}(H_1,H_2)$ è un isomorfismo, allora anche $A^* \in \mathcal{B}(H_2,H_1)$ lo è, e inoltre $(A^*)^{-1} = (A^{-1})^*$.*

(vi) *Se $A \in \mathcal{B}(H_1, H_2)$, allora $\mathrm{Ker}(A^*) = (\mathrm{Im}\, A)^\perp$ e $\overline{\mathrm{Im}\, A} = \mathrm{Ker}(A^*)^\perp$. Di conseguenza, si ha la decomposizione ortogonale*

$$H_2 = \mathrm{Ker}(A^*) \oplus \overline{\mathrm{Im}\, A}.$$

(vii) *Sia $A \in \mathcal{B}(H_1)$ e sia $V \subseteq H_1$ un sottospazio vettoriale chiuso. Supponiamo che V sia invariante per A e A^*, cioè $A(V) \subseteq V$ e $A^*(V) \subseteq V$. Allora*

$$(A|_V)^* = A^*|_V,$$

dove $A|_V, A^*|_V \in \mathcal{B}(V)$ sono le restrizioni di A e A^* a V, pensato come spazio di Hilbert con il prodotto scalare indotto da H_1.

Dimostrazione. (i). Per definizione di aggiunto, per ogni $x \in H_1$ e $y \in H_2$,

$$\langle A^{**}x, y\rangle_{H_2} = \langle x, A^*y\rangle_{H_1} = \langle Ax, y\rangle_{H_2},$$

dunque $F_{A^{**}} = F_A$ e $A^{**} = A$.

(ii). Per ogni $x \in H_2$ e $y \in H_1$,

$$\begin{aligned}\langle(\alpha A + \beta B)^*x, y\rangle_{H_1} &= \langle x, (\alpha A + \beta B)y\rangle_{H_2} = \overline{\alpha}\langle x, Ay\rangle_{H_2} + \overline{\beta}\langle x, By\rangle_{H_2} \\ &= \overline{\alpha}\langle A^*x, y\rangle_{H_1} + \overline{\beta}\langle B^*x, y\rangle_{H_1} = \langle(\overline{\alpha}A^* + \overline{\beta}B^*)x, y\rangle_{H_1},\end{aligned}$$

dunque $(\alpha A + \beta B)^* = \overline{\alpha}A^* + \overline{\beta}B^*$.

(iii). Per ogni $x \in H_3$ e $y \in H_1$,

$$\langle(BA)^*x, y\rangle_{H_1} = \langle x, BAy\rangle_{H_3} = \langle B^*x, Ay\rangle_{H_2} = \langle A^*B^*x, y\rangle_{H_1},$$

dunque $(BA)^* = A^*B^*$.

(iv). Per la submoltiplicatività della norma operatoriale (☞ 7.10(ii)) e il fatto che $\|A^*\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}$ (☞ proposizione 10.1), si ha

$$\|A^*A\|_{\mathrm{op}} \leq \|A^*\|_{\mathrm{op}}\|A\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}^2.$$

Rimane da dimostrare la disuguaglianza opposta. D'altra parte, se $x \in H_1$ e $\|x\|_{H_1} \leq 1$, allora

$$\|Ax\|_{H_2}^2 = \langle Ax, Ax\rangle_{H_2} = \langle A^*Ax, x\rangle_{H_1} \leq \|A^*Ax\|_{H_1}\|x\|_{H_1} \leq \|A^*Ax\|_{H_1},$$

dove si sono usate la definizione di aggiunto e la disuguaglianza di Cauchy–Schwarz (☞ proposizione 6.10). Dunque (☞ lemma 7.5)

$$\begin{aligned}\|A\|_{\mathrm{op}}^2 = \left(\sup_{\|x\|_{H_1} \leq 1} \|Ax\|_{H_2}\right)^2 &= \sup_{\|x\|_{H_1} \leq 1} \|Ax\|_{H_2}^2 \\ &\leq \sup_{\|x\|_{H_1} \leq 1} \|A^*Ax\|_{H_1} = \|A^*A\|_{\mathrm{op}},\end{aligned}$$

come desiderato.

(v). Siccome $A \in \mathcal{B}(H_1, H_2)$ è un isomorfismo, abbiamo $A^{-1} \in \mathcal{B}(H_2, H_1)$; quindi

$$\begin{aligned}(A^{-1})^*A^* &= (AA^{-1})^* = (\mathrm{id}_{H_2})^* = \mathrm{id}_{H_2}, \\ A^*(A^{-1})^* &= (A^{-1}A)^* = (\mathrm{id}_{H_1})^* = \mathrm{id}_{H_1},\end{aligned}$$

dove si sono usati (iii) e l'esempio 10.4(1). Questo dimostra che $A^* : H_2 \to H_1$ è invertibile e $(A^*)^{-1} = (A^{-1})^* \in \mathcal{B}(H_1, H_2)$.

(vi). Per la definizione di aggiunto, si ha

$$\langle A^* x, y\rangle_{H_1} = \langle x, Ay\rangle_{H_2} \tag{10.8}$$

per ogni $x \in H_2$ e $y \in H_1$. Pertanto, se $x \in \operatorname{Ker}(A^*)$, da (10.8) deduciamo che $x \perp Ay$ per ogni $y \in H_1$, cioè $x \in (\operatorname{Im} A)^\perp$. Viceversa, se $x \in (\operatorname{Im} A)^\perp$, allora da (10.8) segue che $\langle A^* x, y\rangle_{H_1} = 0$ per ogni $y \in H$, e prendendo $y = A^* x$ deduciamo che $\|A^* x\|_{H_1} = 0$, cioè $A^* x = 0$ e $x \in \operatorname{Ker}(A^*)$.

Questo dimostra l'uguaglianza $\operatorname{Ker}(A^*) = (\operatorname{Im} A)^\perp$. Prendendo i complementi ortogonali di ambo i membri, si ha allora

$$\operatorname{Ker}(A^*)^\perp = (\operatorname{Im} A)^{\perp\perp} = \overline{\operatorname{Im} A},$$

dove si è usato il corollario 6.33(vi). Infine, applicando il corollario 6.33(i) con $Y = \operatorname{Ker}(A^*)$, si ottiene

$$H_2 = \operatorname{Ker}(A^*) \oplus \operatorname{Ker}(A^*)^\perp = \operatorname{Ker}(A^*) \oplus \overline{\operatorname{Im} A},$$

come desiderato.

(vii). Dalla definizione di aggiunto sappiamo che

$$\langle A^* x, y\rangle_{H_1} = \langle x, Ay\rangle_{H_1}$$

per ogni $x, y \in H_1$. Siccome V è invariante per A e A^*, se $x, y \in V$ si ha anche $A^* x = A^*|_V x \in V$ e $Ay = A|_V y \in V$; pertanto si ha

$$\langle A^*|_V x, y\rangle_V = \langle x, A|_V y\rangle_V$$

per ogni $x, y \in V$, e per la definizione di aggiunto concludiamo che $A^*|_V$ è l'aggiunto di $A|_V$. □

Corollario 10.6. *Siano H_1, H_2 spazi di Hilbert. La mappa $A \mapsto A^*$ è un anti-isomorfismo isometrico da $\mathcal{B}(H_1, H_2)$ a $\mathcal{B}(H_2, H_1)$.*

Dimostrazione. L'antilinearità è dimostrata nella proposizione 10.5(ii), e il fatto che sia un'isometria, cioè che $\|A^*\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}$, è nella proposizione 10.1. Infine, per l'involutività (☞ proposizione 10.5(i)), la mappa $\mathcal{B}(H_1, H_2) \ni A \mapsto A^* \in \mathcal{B}(H_2, H_1)$ è invertibile, con inversa $\mathcal{B}(H_2, H_1) \ni B \mapsto B^* \in \mathcal{B}(H_1, H_2)$. □

Corollario 10.7 (criterio di invertibilità in spazi di Hilbert)**.** *Siano H_1, H_2 spazi di Hilbert. Sia $T \in \mathcal{B}(H_1, H_2)$. Allora T è un isomorfismo se e solo se valgono entrambe le seguenti proprietà:*

(a) $\operatorname{Ker}(T^*) = \{0\}$
(l'aggiunto T^* è iniettivo);

(b) esiste $C \in (0, \infty)$ tale che $\|x\|_{H_1} \leq C\|Tx\|_{H_2}$ per ogni $x \in H_1$
(T è coercivo in norma*).*

Dimostrazione. Grazie alle proposizioni 10.5(vi) e 6.26, si ha

$$\operatorname{Ker}(T^*) = \{0\} \quad \Longleftrightarrow \quad \overline{\operatorname{Im} A} = H.$$

Pertanto, le proprietà (a) e (b) sono equivalenti a quelle elencate in uno dei criteri di invertibilità precedentemente discussi (☞ teorema 7.39), che in base a tale criterio sono a loro volta equivalenti al fatto che T sia un isomorfismo. □

10.2 Operatori autoaggiunti, unitari, normali e proiezioni ortogonali

Utilizzando la nozione di aggiunto, introduciamo alcune classi importanti di operatori su uno spazio di Hilbert.

Definizione 10.8. Sia H uno spazio di Hilbert. Un operatore $T \in \mathcal{B}(H)$ si dice:

(a) *autoaggiunto*, se $T^* = T$
(T è uguale al suo aggiunto);

(b) *normale*, se $T^*T = TT^*$
(T commuta con il suo aggiunto);

(c) *unitario*, se $T^*T = TT^* = \mathrm{id}_H$
(l'aggiunto T^* è l'inverso di T).

L'ultima definizione si generalizza al caso di operatori fra due diversi spazi di Hilbert.

Definizione 10.9. Siano H_1, H_2 spazi di Hilbert. Un operatore $T \in \mathcal{B}(H_1, H_2)$ si dice *unitario* se

$$T^*T = \mathrm{id}_{H_1} \qquad \text{e} \qquad TT^* = \mathrm{id}_{H_2}.$$

Osservazione 10.10. Dalla definizione 10.8 è chiaro che, per un operatore $T \in \mathcal{B}(H)$ su uno spazio di Hilbert H, valgono le implicazioni

$$T \text{ autoaggiunto} \implies T \text{ normale} \impliedby T \text{ unitario}.$$

Le implicazioni opposte in generale non valgono.

Esempio 10.11. Ecco alcuni esempi e non-esempi di operatori autoaggiunti, normali e unitari.

(1) Ricordiamo dall'esempio 10.4(2) che, se $S \in \mathcal{B}(\ell^2)$ è l'operatore di shift verso sinistra, il suo aggiunto S^* è l'operatore di shift verso destra con aggiunta di zero. In particolare,

$$S^*S(x_0, x_1, x_2, x_3, \dots) = S^*(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots),$$

mentre

$$SS^*(x_0, x_1, x_2, x_3, \dots) = S^*(0, x_0, x_1, x_2, x_3, \dots) = (x_0, x_1, x_2, x_3, \dots)$$

per ogni $\underline{x} = (x_k)_k \in \ell^2$. Dunque

$$SS^* = \mathrm{id}_{\ell^2} \neq S^*S. \tag{10.9}$$

Questo mostra che S non è un operatore normale, e a maggior ragione non è né autoaggiunto né unitario.

(2) Ricordiamo dall'esempio 10.4(3) che, se $\underline{w} \in \ell^\infty$, allora l'aggiunto dell'operatore di moltiplicazione $D_{\underline{w}} \in \mathcal{B}(\ell^2)$ è dato da

$$D_{\underline{w}}^* = D_{\overline{\underline{w}}}. \tag{10.10}$$

Ora, non è difficile verificare che la mappa

$$\ell^\infty \ni \underline{w} \mapsto D_{\underline{w}} \in \mathcal{B}(\ell^2) \tag{10.11}$$

è lineare; siccome $\|D_{\underline{w}}\|_{\mathrm{op}} = \|\underline{w}\|_\infty$ (☞ esempio 7.8(8)), in effetti la mappa (10.11) è un'isometria lineare (☞ definizione 7.15), dunque è iniettiva (☞ proposizione 7.16(iv)). In aggiunta, si può dimostrare che tale mappa è moltiplicativa, nel senso che

$$D_{\underline{v}\cdot\underline{w}} = D_{\underline{v}} D_{\underline{w}} \tag{10.12}$$

per ogni $\underline{v}, \underline{w} \in \ell^\infty$, dove

$$\underline{v} \cdot \underline{w} := (v_k w_k)_{k\in\mathbb{N}}$$

è il prodotto componente per componente di $\underline{v}$ e $\underline{w}$.

Siccome il prodotto componente per componente è commutativo, dall'identità (10.12) segue che tutti gli operatori di moltiplicazione $D_{\underline{w}}$ commutano fra loro. In particolare, per (10.10), ogni operatore di moltiplicazione commuta con il proprio aggiunto, cioè $D_{\underline{w}}$ è un operatore normale per ogni $\underline{w} \in \ell^\infty$.

Inoltre, siccome la mappa (10.11) è iniettiva, dall'identità (10.10) si deduce che $D_{\underline{w}}$ è autoaggiunto se e solo se $\overline{\underline{w}} = \underline{w}$, cioè se e solo se $\underline{w}$ è a valori reali.

Infine, $D_{\underline{w}}$ è unitario se e solo se $D_{\underline{w}}^* D_{\underline{w}} = D_{\underline{w}} D_{\underline{w}}^* = \mathrm{id}_{\ell^2} = D_{\underline{1}}$, dove $\underline{1}$ è la successione costante 1; ragionando come sopra, deduciamo che $D_{\underline{w}}$ è unitario se e solo se $\overline{\underline{w}} \cdot \underline{w} = \underline{1}$, cioè $|w_k|^2 = \overline{w_k} w_k = 1$ per ogni $k \in \mathbb{N}$, cioè $|w_k| = 1$ per ogni $k \in \mathbb{N}$.

Osservazione 10.12. Se H_1 e H_2 sono spazi di Hilbert, allora, per ogni $T \in \mathcal{B}(H_1, H_2)$, l'operatore T^*T è autoaggiunto. Infatti, per la proposizione 10.5 si ha

$$(T^*T)^* = T^*T^{**} = T^*T.$$

Vediamo ora una caratterizzazione alternativa degli operatori unitari.

Proposizione 10.13. *Siano H_1 e H_2 spazi di Hilbert. Un operatore $T \in \mathcal{B}(H_1, H_2)$ è unitario se e solo se $T : H_1 \to H_2$ è un isomorfismo isometrico.*

Dimostrazione. Supponiamo che $T \in \mathcal{B}(H_1, H_2)$ sia unitario. Allora per definizione T è invertibile, con inverso $T^* \in \mathcal{B}(H_2, H_1)$. Rimane solo da dimostrare che T è un'isometria. D'altra parte,

$$\|Tx\|_{H_2}^2 = \langle Tx, Tx\rangle_{H_2} = \langle T^*Tx, x\rangle_{H_1} = \langle x, x\rangle_{H_1} = \|x\|_{H_1}^2$$

per ogni $x \in H_1$, dove si è usato che $T^*T = \mathrm{id}_{H_1}$; dunque $T : H_1 \to H_2$ è un'isometria lineare (☞ definizione 7.15), come desiderato.

Supponiamo viceversa che $T : H_1 \to H_2$ sia un isomorfismo isometrico. Allora (☞ definizione 7.19 e osservazione 7.20) T è invertibile con inverso limitato. Inoltre T preserva il prodotto scalare (☞ proposizione 7.22), dunque

$$\langle T^*Tx, y\rangle_{H_1} = \langle Tx, Ty\rangle_{H_2} = \langle x, y\rangle_{H_1}$$

per ogni $x, y \in H_1$, da cui deduciamo che $T^*T = \mathrm{id}_{H_1}$. Moltiplicando ambo i membri a destra per T^{-1}, si ottiene infine $T^* = T^{-1}$, cioè T è normale. □

Infine, vediamo come la nozione di aggiunto permetta di caratterizzare in maniera algebrica un'ulteriore classe di operatori limitati di cui abbiamo già discusso.

Proposizione 10.14. *Siano H uno spazio di Hilbert e $P \in \mathcal{B}(H)$. Sono fatti equivalenti:*

(i) P è la mappa di proiezione ortogonale P_Y su un qualche sottospazio vettoriale chiuso Y di H.

(ii) $P^2 = P = P^*$.

Inoltre, in tal caso, $Y = \operatorname{Im} P = \{x \in H : Px = x\}$ e $Y^\perp = \operatorname{Ker} P$.

Dimostrazione. (i) $\Rightarrow$ (ii). Supponiamo che $P = P_Y$ per qualche sottospazio chiuso Y di H. Per definizione della proiezione ortogonale ($P_Y x$ è il punto di Y che minimizza la distanza di x da Y) chiaramente

$$Py = P_Y y = y \qquad \forall y \in Y. \tag{10.13}$$

In particolare, per ogni $y \in Y$, si ha $y = Py \in \operatorname{Im} P$; dunque $Y \subseteq \operatorname{Im} P$; siccome P_Y è a valori in Y per costruzione, se ne conclude $\operatorname{Im} P = Y$. Inoltre, per ogni $x \in H$, si ha $Px \in \operatorname{Im} P = Y$, ma allora $P^2 x = PPx = Px$ per (10.13) applicata con $y = Px$. Questo dimostra che $P^2 = P$.

Inoltre, per ogni $x, x' \in H$, si ha

$$\begin{aligned}\langle Px, x'\rangle &= \langle P_Y x, P_Y x' + P_{Y^\perp} x'\rangle \\ &= \langle P_Y x, P_Y x'\rangle = \langle P_Y x + P_{Y^\perp} x, P_Y x'\rangle = \langle x, Px'\rangle,\end{aligned}$$

dove abbiamo usato la decomposizione ortogonale di x e x' (☞ corollario 6.33) e il fatto che $Y \perp Y^\perp$. Da questa identità deduciamo che $P^* = P$, come desiderato.

(ii) $\Rightarrow$ (i). Sia $Y = \{x \in H : Px = x\}$. Allora $Y = \operatorname{Ker}(P - \mathrm{id}_H)$ è un sottospazio vettoriale chiuso di H (☞ proposizione 7.13(i)). Inoltre, siccome $P^2 = P$, si ha $\operatorname{Im} P \subseteq Y$. D'altra parte, per ogni $x \in Y$, si ha $x = Px \in \operatorname{Im} P$, pertanto $Y \subseteq \operatorname{Im} P$ e quindi $Y = \operatorname{Im} P$.

Siccome inoltre $P^* = P$, dalla proposizione 10.5(vi) deduciamo che $Y^\perp = (\operatorname{Im} P)^\perp = \operatorname{Ker} P$. In particolare, per ogni $x \in H$, considerata la decomposizione ortogonale $x = P_Y x + P_{Y^\perp} x$ rispetto al sottospazio Y (☞ teorema 6.32), si ha

$$Px = PP_Y x + PP_{Y^\perp x} = P_Y x + 0 = P_Y x,$$

ove si è usato che $P|_Y = \mathrm{id}_Y$ (per definizione di Y) e $P|_{Y^\perp} = 0$ (perché $Y^\perp = \operatorname{Ker} P$). Dunque $P = P_Y$, come desiderato. □

Abbiamo visto (☞ corollario 7.43(ii)) che l'insieme degli operatori limitati invertibili su uno spazio di Banach X è un sottoinsieme aperto dell'insieme $\mathcal{B}(X)$ di tutti gli operatori limitati. Vediamo ora che le classi di operatori discusse sopra determinano invece dei sottoinsiemi chiusi.

Proposizione 10.15. *Sia H uno spazio di Hilbert. I seguenti sottoinsiemi di $\mathcal{B}(H)$ sono chiusi (topologicamente):*

(i) l'insieme degli operatori normali;

(ii) l'insieme degli operatori autoaggiunti;

(iii) l'insieme degli operatori unitari;

(iv) l'insieme delle proiezioni ortogonali.

Dimostrazione (cenno). Sulla base della definizione 10.8 e della proposizione 10.14, gli operatori T appartenenti a questi quattro sottoinsiemi sono caratterizzati rispettivamente dalle seguenti condizioni:

(i) $T^*T = TT^*$;

(ii) $T^* = T$;

(iii) $T^*T = TT^* = \mathrm{id}_H$;

(iv) $T^2 = T = T^*$.

Usando la continuità delle operazioni di prodotto e aggiunto su $\mathcal{B}(H)$ (☞ corollario 7.12 e corollario 10.6), si vede che ciascuna di queste condizioni è preservata per passaggio al limite di successioni in $\mathcal{B}(H)$, dunque ciascuno dei corrispondenti insiemi di operatori è chiuso in $\mathcal{B}(H)$ (☞ proposizione 2.23(vi)). □

Nel caso $\mathbb{F} = \mathbb{C}$, gli operatori autoaggiunti si possono caratterizzare sulla base dei valori "diagonali" della forma sesquilineare associata.

Proposizione 10.16. *Supponiamo che $\mathbb{F} = \mathbb{C}$. Siano H uno spazio di Hilbert e $T \in \mathcal{B}(H)$. Allora T è autoaggiunto se e solo se*

$$\langle Tx, x\rangle \in \mathbb{R} \qquad \forall x \in H. \tag{10.14}$$

Dimostrazione. Se T è autoaggiunto, si ha

$$\langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$$

per ogni $x \in H$, da cui segue che $\langle Tx, x\rangle \in \mathbb{R}$.

Viceversa, supponiamo che T soddisfi (10.14). Allora, per ogni $x, y \in H$,

$$\langle T(x+y), x+y\rangle = \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle;$$

siccome i termini $\langle T(x+y), x+y\rangle, \langle Tx, x\rangle, \langle Ty, y\rangle$ sono reali, per differenza si ha anche

$$\langle Tx, y\rangle + \langle Ty, x\rangle \in \mathbb{R},$$

dunque

$$\mathfrak{Im}\langle Tx, y\rangle = -\,\mathfrak{Im}\langle Ty, x\rangle.$$

Prendendo iy al posto di y, deduciamo pure

$$\mathfrak{Re}\langle Tx, y\rangle = \mathfrak{Re}\langle Ty, x\rangle.$$

Combinando le ultime due identità, si vede che

$$\langle Ty, x\rangle = \overline{\langle Tx, y\rangle} = \langle y, Tx\rangle,$$

e per l'arbitrarietà di $x, y \in H$ deduciamo che T è autoaggiunto. □

Infine, discutiamo alcune importanti proprietà degli operatori normali.

Proposizione 10.17. *Sia H uno spazio di Hilbert. Sia $T \in \mathcal{B}(H)$ un operatore normale. Allora:*

(i) $\|Tx\|_H = \|T^*x\|_H$ *per ogni* $x \in H$*;*

(ii) $\operatorname{Ker} T = \operatorname{Ker}(T^*)$*;*

(iii) $T - \lambda \mathrm{id}_H$ *è normale per ogni* $\lambda \in \mathbb{F}$*.*

In particolare,

(iv) $\operatorname{Ker}(T - \lambda \mathrm{id}_H) = \operatorname{Ker}(T^* - \overline{\lambda}\mathrm{id}_H)$ *per ogni* $\lambda \in \mathbb{C}$*.*

Dimostrazione. (i). Sia $x \in H$. Allora

$$\|Tx\|_H^2 - \|T^*x\|_H^2 = \langle Tx, Tx\rangle - \langle T^*x, T^*x\rangle = \langle T^*Tx, x\rangle - \langle TT^*x, x\rangle = 0,$$

dove si è usato che $T^*T = TT^*$ perché T è normale.

(ii). Da (i) segue che, per ogni $x \in H$, $Tx = 0$ se e solo se $T^*x = 0$.

(iii). Per ogni $\lambda \in \mathbb{F}$, posto $I = \mathrm{id}_H$, si ha (☞ proposizione 10.5)

$$\begin{aligned}(T - \lambda I)^*(T - \lambda I) &= (T^* - \overline{\lambda} I)(T - \lambda I) = T^*T - \lambda T^* - \overline{\lambda} T + |\lambda|^2 I,\\ (T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda} I) = TT^* - \lambda T^* - \overline{\lambda} T + |\lambda|^2 I,\end{aligned}$$

e le due espressioni sono uguali perché $T^*T = TT^*$.

(iv). Segue applicando (ii) all'operatore $T - \lambda \mathrm{id}_H$, che è normale per (iii). □

10.3 Spettro di un operatore limitato

In questa sezione, $H \neq \{0\}$ è uno spazio di Hilbert su $\mathbb{F}$. Poniamo per brevità $I = \mathrm{id}_H$.

Definizione 10.18. Sia $T \in \mathcal{B}(H)$. Lo *spettro* di T è l'insieme

$$\sigma(T) = \{\lambda \in \mathbb{F} : T - \lambda I \text{ non è invertibile in } \mathcal{B}(H)\}.$$

L'*insieme risolvente* $\rho(T)$ di T è il complementare $\mathbb{F} \setminus \sigma(T)$ dello spettro.

Osservazione 10.19. Se $\dim H < \infty$, allora (☞ proposizione 1.201.21)

$$\begin{aligned}T - \lambda I \text{ non è invertibile} &\overset{(\dagger)}{\iff} T - \lambda I \text{ non è iniettivo}\\ &\iff \operatorname{Ker}(T - \lambda I) \neq \{0\} \iff \lambda \text{ è un autovalore di } T\end{aligned}$$

(gli elementi non nulli di $\operatorname{Ker}(T - \lambda I)$ sono gli autovettori di T di autovalore λ). In dimensione finita, lo spettro è l'insieme degli autovalori.

In dimensione infinita, tuttavia, in (†) si ha solo l'implicazione $\Longleftarrow$. In altre parole, ci sono altre possibili "ostruzioni" all'invertibilità oltre alla mancanza di iniettività. In questo caso, lo spettro ha una struttura più complicata.

Definizione 10.20. Sia $T \in \mathcal{B}(H)$.

(a) Lo *spettro puntuale* di T è l'insieme

$$\sigma_p(T) = \{\lambda \in \mathbb{F} : \mathrm{Ker}(T - \lambda I) \neq \{0\}\}.$$

In altre parole, $\sigma_p(T)$ è l'insieme degli autovalori di T.

(b) Lo *spettro residuo* di T è l'insieme

$$\sigma_r(T) = \{\lambda \in \mathbb{F} : \mathrm{Ker}(T - \lambda I) = \{0\},\ \text{ma } \overline{\mathrm{Im}(T - \lambda I)} \neq H\}.$$

(c) Lo *spettro continuo* di T è l'insieme

$$\sigma_c(T) = \{\lambda \in \mathbb{F} : \mathrm{Ker}(T-\lambda I) = \{0\}, \overline{\mathrm{Im}(T - \lambda I)} = H, \text{ ma } \mathrm{Im}(T-\lambda I) \neq H\}.$$

Osservazione 10.21. Dalla definizione, è chiaro che $\sigma_p(T), \sigma_r(T), \sigma_c(T)$ sono a due a due disgiunti, e inoltre

$$\sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T).$$

Infatti, se $\lambda \in \sigma_p(T)$ allora $T - \lambda I$ non è iniettivo, mentre se $\lambda \in \sigma_r(T) \cup \sigma_c(T)$ allora $T - \lambda I$ non è suriettivo, e in ogni caso T non è invertibile. Viceversa, se $\lambda \notin \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T)$, allora $T-\lambda I$ è iniettivo e suriettivo, dunque invertibile e quindi $(T - \lambda I)^{-1} \in \mathcal{B}(H)$ per il teorema dell'isomorfismo (☞ teorema 7.36), cioè $\lambda \notin \sigma(T)$.

Discutiamo ora alcune proprietà fondamentali dello spettro.

Proposizione 10.22. *Sia* $T \in \mathcal{B}(H)$.

(i) Lo spettro $\sigma(T)$ *è un sottoinsieme chiuso della palla* $\overline{B}(0, \|T\|_{\mathrm{op}})$ *in* $\mathbb{F}$.

(ii) Se $\mathbb{F} = \mathbb{C}$, *lo spettro* $\sigma(T)$ *è non vuoto.*

Dimostrazione. (i). L'enunciato da dimostrare è equivalente al fatto che l'insieme risolvente $\rho(T)$ è aperto in $\mathbb{F}$ e contiene il complementare $\mathbb{F} \setminus \overline{B}(0, \|T\|_{\mathrm{op}})$ della palla.

Ora, se $\lambda \in \rho(T)$, allora $T - \lambda I$ è un isomorfismo. Pertanto, per ogni $\lambda' \in \mathbb{F}$ con $|\lambda - \lambda'| < 1/\|T - \lambda I\|_{\mathrm{op}}$, si ha

$$\|(T - \lambda' I) - (T - \lambda I)\|_{\mathrm{op}} = \|(\lambda - \lambda')I\|_{\mathrm{op}} = |\lambda - \lambda'| < 1/\|T - \lambda I\|_{\mathrm{op}};$$

dal corollario 7.43(i) segue allora che anche $T - \lambda' I$ è un isomorfismo, cioè $\lambda' \in \rho(T)$. Questo dimostra che $B(\lambda, 1/\|T - \lambda I\|_{\mathrm{op}}) \subseteq \rho(T)$ per ogni $\lambda \in \rho(T)$, e quindi che $\rho(T)$ è aperto in $\mathbb{F}$.

Similmente, sia $\lambda \in \mathbb{F} \setminus \overline{B}(0, \|T\|_{\mathrm{op}})$, cioè tale che $|\lambda| > \|T\|_{\mathrm{op}}$. Allora $T - \lambda I = -\lambda(I - T/\lambda)$ e $\|T/\lambda\|_{\mathrm{op}} < 1$, dunque per il criterio di Neumann (☞ teorema 7.42) deduciamo che $I - T/\lambda$ è un isomorfismo, ma allora anche $T - \lambda I = -\lambda(I - T/\lambda)$ lo è (☞ osservazione 7.33(3)), cioè $\lambda \in \rho(T)$.

(ii). Se $\dim H < \infty$, il risultato si deduce dall'esistenza di radici in $\mathbb{C}$ del polinomio caratteristico $\lambda \mapsto \det(T - \lambda I)$, che a sua volta è conseguenza del teorema fondamentale dell'algebra. Nel caso generale, si utilizza un argomento più sofisticato basato sull'analisi complessa, di cui omettiamo i dettagli. □

Vediamo ora come lo spettro si comporta rispetto ad alcune "operazioni" sull'insieme $\mathcal{B}(H)$.

Proposizione 10.23. *Sia $T \in \mathcal{B}(H)$.*

(i) $\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$.

(ii) *T è invertibile in $\mathcal{B}(H)$ se e solo se $0 \notin \sigma(T)$. Inoltre, in tal caso, si ha $\sigma(T^{-1}) = \{\lambda^{-1} : \lambda \in \sigma(T)\}$.*

(iii) $\sigma(\alpha I) = \{\alpha\}$ *per ogni $\alpha \in \mathbb{F}$.*

(iv) *Supponiamo $\mathbb{F} = \mathbb{C}$ oppure $n \leq 1$. Se $p(z) = \sum_{j=0}^n a_j z^j \in \mathbb{F}[z]$ è un polinomio, posto $p(T) \coloneqq \sum_{j=0}^n a_j T^j$, si ha $\sigma(p(T)) = p(\sigma(T))$.* (teorema della mappa spettrale)

Dimostrazione. (i). Basta dimostrare che $\rho(T^*) = \{\overline{\lambda} : \lambda \in \rho(T)\}$. Infatti, se $\lambda \in \rho(T)$, allora $T - \lambda I$ è un isomorfismo, dunque anche $(T - \lambda I)^* = T^* - \overline{\lambda} I$ lo è (☞ proposizione 10.5), e quindi $\overline{\lambda} \in \rho(T)$. Viceversa, se $\mu \in \rho(T^*)$, applicando quanto appena dimostrato all'operatore T^* si ottiene $\overline{\mu} \in \rho(T^{**}) = \rho(T)$, cioè $\mu = \overline{\lambda}$ per qualche $\lambda \in \rho(T)$, come desiderato.

(ii). Per definizione di $\sigma(T)$, è chiaro che T è un isomorfismo se e solo se $0 \notin \sigma(T)$. Supponiamo ora che T sia un isomorfismo. Allora anche T^{-1} lo è, dunque $0 \notin \sigma(T)$ e $0 \notin \sigma(T^{-1})$. Inoltre, per ogni $\lambda \in \mathbb{F} \setminus \{0\}$,

$$T - \lambda I = -\lambda T(T^{-1} - \lambda^{-1} I),$$

dunque (☞ osservazione 7.33(3)) $T - \lambda I$ è un isomorfismo se e solo se $T^{-1} - \lambda^{-1} I$ lo è, e viceversa $\lambda \in \sigma(T)$ se e solo se $\lambda^{-1} \in \sigma(T^{-1})$, come desiderato.

(iii). Per ogni $\lambda \in \mathbb{F}$, $\alpha I - \lambda I = (\alpha - \lambda) I$ è invertibile in $\mathcal{B}(H)$ se e solo se $\alpha \neq \lambda$.

(iv). Se $p(z) = a_0$ è costante, l'enunciato si riduce al fatto che $\sigma(a_0 I) = \{a_0\}$, già discusso in (iii).

Supponiamo p non sia costante; possiamo dunque assumere $n > 0$ e $a_n \neq 0$. Sia $\mu \in \mathbb{F}$ e definiamo il polinomio $q(z) = p(z) - \mu$. Allora q si fattorizza come

$$q(z) = a_n (z - \lambda_1) \cdots (z - \lambda_n),$$

dove $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ sono le radici di q (nel caso $\mathbb{F} = \mathbb{C}$, questo è dovuto al teorema fondamentale dell'algebra; se invece $\mathbb{F} = \mathbb{R}$, per ipotesi $n = 1$ e la fattorizzazione è banalmente vera). Pertanto

$$p(T) - \mu I = q(T) = a_n (T - \lambda_1 I) \cdots (T - \lambda_n I).$$

In particolare, si ha $\mu \in \sigma(p(T))$ se e solo se $a_n (T - \lambda_1) \cdots (T - \lambda_n)$ non è un isomorfismo, ma questo si verifica se e solo se uno dei fattori $T - \lambda_j I$ non è un isomorfismo (☞ osservazione 7.33(3)-(5)), cioè se e solo se $\lambda_j \in \sigma(T)$ per qualche $j = 1, \dots, n$. Siccome $p^{-1}(\mu) = q^{-1}(0) = \{\lambda_1, \dots, \lambda_n\}$, questo dimostra il risultato. □

Dimostriamo ora alcune proprietà dello spettro di certe classi di operatori.

Proposizione 10.24. *Sia $T \in \mathcal{B}(H)$.*

(i) Se T è normale, allora $\sigma_r(T) = \emptyset$.

(ii) Se T è autoaggiunto, allora $\sigma(T) \subseteq \mathbb{R}$.

(iii) Se T è unitario, allora $\sigma(T) \subseteq \{\lambda \in \mathbb{F} : |\lambda| = 1\}$.

Dimostrazione. (i). Notiamo che $\sigma_r(T) \subseteq \sigma(T) \setminus \sigma_p(T)$ (☞ osservazione 10.21).

Sia ora $\lambda \in \sigma(T) \setminus \sigma_p(T)$. Allora $T - \lambda I$ è iniettivo (☞ definizione 10.20). Siccome T è normale, lo è anche $T - \lambda I$, pertanto

$$\operatorname{Ker}((T - \lambda I)^*) = \operatorname{Ker}(T - \lambda I) = \{0\}$$

(☞ proposizione 10.17). Ma allora

$$\overline{\operatorname{Im}(T - \lambda I)} = \operatorname{Ker}((T - \lambda I)^*)^\perp = \{0\}^\perp = H$$

(☞ proposizioni 10.5(vi) e 6.26). Ne deduciamo (☞ definizione 10.20) che $\lambda \notin \sigma_r(T)$. Dunque $\sigma_r(T) = \emptyset$.

(ii). Se $\mathbb{F} = \mathbb{R}$ il risultato è banale, dunque assumiamo $\mathbb{F} = \mathbb{C}$.

Sia $\lambda = \alpha + i\beta \in \mathbb{C}$, dove $\alpha, \beta \in \mathbb{R}$. Allora

$$T - \lambda I = (T - \alpha I) - i\beta I = S - i\beta I$$

ove $S = T - \alpha I$ è pure autoaggiunto (☞ proposizione 10.5). Dunque, per ogni $x \in H$,

$$\|(T - \lambda I)x\|^2 = \langle (S - i\beta I)x, (S - i\beta I)x\rangle = \|Sx\|^2 + |\beta|^2\|x\|^2 + 2\,\mathfrak{Re}(i\beta\langle Sx, x\rangle).$$

Osserviamo ora che $\langle Sx, x\rangle \in \mathbb{R}$ perché S è autoaggiunto (☞ proposizione 10.16), quindi la parte reale nell'espressione sopra svanisce e si ottiene

$$\|(T - \lambda I)x\|^2 = \|Sx\|^2 + |\beta|^2\|x\|^2 \geq |\beta|^2\|x\|^2. \tag{10.15}$$

Lo stesso argomento con $\overline{\lambda} = \alpha - i\beta$ al posto di λ ci dà anche

$$\|(T - \lambda I)^* x\|^2 = \|(T - \overline{\lambda} I)x\|^2 \geq |\beta|^2\|x\|^2. \tag{10.16}$$

Se $\beta = \mathfrak{Im}\,\lambda \neq 0$, da (10.16) si vede che $\operatorname{Ker}((T - \lambda I)^*) = \{0\}$, e da (10.15) abbiamo che $T - \lambda I$ è coercivo per la norma; pertanto applicando il corollario 10.7 all'operatore $T - \lambda I$ deduciamo che quest'ultimo è un isomorfismo, cioè $\lambda \notin \sigma(T)$. Dunque gli elementi di $\sigma(T)$ devono avere parte immaginaria nulla, cioè $\sigma(T) \subseteq \mathbb{R}$.

(iii). Un operatore unitario T è un isomorfismo isometrico (☞ proposizione 10.13), dunque $\|T\|_{\mathrm{op}} = 1$ (☞ proposizione 7.16(iii)) e quindi $\sigma(T) \subseteq \overline{B}(0, 1)$ (☞ proposizione 10.22(i)). D'altra parte anche T^{-1} è un isomorfismo isometrico, e quindi anche $\sigma(T^{-1}) \subseteq \overline{B}(0, 1)$. Dalla proposizione 10.23(ii) deduciamo dunque che, se $\lambda \in \sigma(T)$, allora si ha sia $|\lambda| \leq 1$ che $|\lambda^{-1}| \leq 1$, cioè $|\lambda| = 1$. □

Esempio 10.25. Vediamo alcuni esempi di calcolo dello spettro.

(1) Per ogni $\underline{w} \in \ell^\infty$ sia $D_{\underline{w}} \in \mathcal{B}(\ell^2)$ l'operatore di moltiplicazione per $\underline{w}$ (☞ esempio 7.8(8)). Allora, per ogni $\lambda \in \mathbb{F}$,

$$D_{\underline{w}} - \lambda I = D_{\underline{w} - \lambda \underline{1}},$$

dove $\underline{1} = (1, 1, 1, \dots)$ (☞ esempio 10.11(2)). In particolare, dai criteri di iniettività e invertibilità per operatori diagonali dell'esempio 7.35(4), deduciamo che

$$\lambda \in \sigma_p(D_{\underline{w}}) \iff D_{\underline{w}-\lambda\underline{1}} \text{ non iniettivo} \iff \exists k \in \mathbb{N} : w_k = \lambda,$$

ovverosia

$$\sigma_p(D_{\underline{w}}) = \{w_k : k \in \mathbb{N}\};$$

inoltre

$$\lambda \in \sigma(D_{\underline{w}}) \iff D_{\underline{w}-\lambda\underline{1}} \text{ non invertibile} \iff \inf_{k\in\mathbb{N}} |w_k - \lambda| = 0,$$

ovverosia

$$\sigma(D_{\underline{w}}) = \{\lambda \in \mathbb{F} : d(\lambda, \{w_k\}_{k\in\mathbb{N}}) = 0\} = \overline{\{w_k : k \in \mathbb{N}\}}.$$

Infine, siccome $D_{\underline{w}}$ è un operatore normale (☞ esempio 10.11(2)), necessariamente

$$\sigma_r(D_{\underline{w}}) = \emptyset$$

(☞ proposizione 10.24(i)), pertanto

$$\sigma_c(D_{\underline{w}}) = \overline{\{w_k : k \in \mathbb{N}\}} \setminus \{w_k : k \in \mathbb{N}\}.$$

(2) Sia $S \in \mathcal{B}(\ell^2)$ l'operatore di shift verso sinistra (☞ esempio 10.4(2)). Sappiamo che S^* è l'operatore di shift verso destra con aggiunta di zero, e che S non è un operatore normale (☞ esempio 10.11(1)). Cerchiamo lo spettro di S.

Cerchiamo anzitutto lo spettro puntuale. Per ogni $\lambda \in \mathbb{F}$ e $\underline{x} \in \ell^2$, si ha

$$(S - \lambda I)\underline{x} = (x_{k+1} - \lambda x_k)_{k\in\mathbb{N}},$$

pertanto $\underline{x} \in \operatorname{Ker}(S - \lambda I)$ se e solo se

$$x_{k+1} = \lambda x_k \quad \forall k \in \mathbb{N};$$

iterando questa relazione, deduciamo che $\underline{x} \in \operatorname{Ker}(S - \lambda I)$ se e solo se

$$x_k = \lambda^k x_0 \quad \forall k \in \mathbb{N},$$

cioè se e solo se

$$\underline{x} = x_0(1, \lambda, \lambda^2, \lambda^3, \dots). \tag{10.17}$$

Ma $(1, \lambda, \lambda^2, \lambda^3, \dots) \in \ell^2$ se e solo se

$$\sum_{k=0}^{\infty} |\lambda^k|^2 = \sum_{k=0}^{\infty} |\lambda|^{2k} < \infty$$

cioè se e solo se $|\lambda| < 1$ (si tratta di una serie geometrica di ragione $|\lambda|^2$). Pertanto, da (10.17) deduciamo che

$$\operatorname{Ker}(S - \lambda I) = \begin{cases} \operatorname{span}\{(1, \lambda, \lambda^2, \lambda^3, \dots)\} & \text{se } |\lambda| < 1, \\ \{0\} & \text{se } |\lambda| \geq 1. \end{cases}$$

Di conseguenza

$$\sigma_p(S) = \{\lambda \in \mathbb{F} : \mathrm{Ker}(S - \lambda I) \neq \{0\}\} = \{\lambda \in \mathbb{F} : |\lambda| < 1\} = B(0,1).$$

D'altra parte, $\|S\|_{\mathrm{op}} = 1$ (☞ esempio 10.4(2)), dunque (☞ proposizione 10.22) $\sigma(S) \subseteq \overline{B}(0,1)$. Ne deduciamo

$$B(0,1) = \sigma_p(S) \subseteq \sigma(S) = \overline{B}(0,1),$$

e passando alle chiusure, siccome $\sigma(S)$ è chiuso (☞ proposizione 10.22), si deduce

$$\sigma(S) = \overline{B}(0,1) = \{\lambda \in \mathbb{F} : |\lambda| \leq 1\}.$$

Dalla proposizione 10.23(i) deduciamo allora anche

$$\sigma(S^*) = \overline{B}(0,1).$$

Cerchiamo ora lo spettro puntuale di S^*. Se $\lambda \in \mathbb{F}$ e $\underline{x} \in \ell^2$, si ha

$$(S^* - \lambda I)\underline{x} = (-\lambda x_0, x_0 - \lambda x_1, x_1 - \lambda x_2, \dots),$$

dunque $\underline{x} \in \mathrm{Ker}(S^* - \lambda I)$ se e solo se

$$\lambda x_0 = 0 \qquad \text{e} \qquad x_k = \lambda x_{k+1} \quad \forall k \in \mathbb{N}. \tag{10.18}$$

Ora, se $\lambda = 0$, le equazioni (10.18) danno $x_k = 0$ per ogni $k \in \mathbb{N}$, cioè $\underline{x} = \underline{0}$; dunque $\mathrm{Ker}(S^*) = \{0\}$. Se invece $\lambda \neq 0$, la prima delle equazioni (10.18) implica che $x_0 = 0$, mentre la seconda dà $x_{k+1} = \lambda^{-1} x_k$ per ogni $k \in \mathbb{N}$; ma allora, iterando, $x_k = \lambda^{-k} x_0 = 0$ per ogni $k \in \mathbb{N}$, cioè di nuovo $\underline{x} = \underline{0}$. Questo dimostra che

$$\mathrm{Ker}(S^* - \lambda I) = \{0\} \qquad \forall \lambda \in \mathbb{F},$$

cioè

$$\sigma_p(S^*) = \emptyset.$$

Possiamo usare questa informazione per determinare le altre componenti dello spettro di S e S^*. Infatti, se $\lambda \in \sigma(S) \setminus \sigma_p(S)$, si ha

$$\mathrm{Im}(S - \lambda I)^\perp = \mathrm{Ker}(S^* - \overline{\lambda} I) = \{0\}$$

dato che $\sigma_p(S^*) = \emptyset$; ma allora $\mathrm{Im}(S - \lambda I)$ è densa in ℓ^2, pertanto $\lambda \notin \sigma_r(S)$ e quindi necessariamente $\lambda \in \sigma_c(S)$. Questo dimostra che

$$\sigma_r(S) = \emptyset, \quad \sigma_p(S) = B(0,1), \quad \sigma_c(S) = \sigma(S) \setminus \sigma_p(S) = \partial B(0,1).$$

Infine, se $\lambda \in \sigma(S^*) = \sigma(S^*) \setminus \sigma_p(S^*)$, si ha

$$\mathrm{Im}(S^* - \lambda I)^\perp = \mathrm{Ker}(S - \overline{\lambda} I)$$

e questo nucleo è non banale (cioè $\mathrm{Im}(S^* - \lambda I)$ non è densa) se e solo se $\overline{\lambda} \in \sigma_p(S) = B(0,1)$, cioè se e solo se $|\lambda| < 1$. Ne deduciamo che

$$\sigma_r(S^*) = B(0,1), \quad \sigma_p(S^*) = \emptyset, \quad \sigma_c(S^*) = \sigma(S) \setminus \sigma_r(S) = \partial B(0,1).$$

Richiamiamo infine alcune proprietà elementari relative agli autospazi di un operatore lineare. Ricordiamo che un sottospazio vettoriale $V \subseteq H$ si dice *invariante* per un operatore $T \in \mathcal{B}(H)$ se $T(V) \subseteq V$.

Proposizione 10.26. *Siano $T, S \in \mathcal{B}(H)$.*

(i) Sia $n \in \mathbb{N}_+$. Se $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ sono distinti, allora

$$\operatorname{Ker}(T - \lambda_n I) \cap \mathop{\Large +}_{j=1}^{n-1} \operatorname{Ker}(T - \lambda_j I) = \{0\}.$$

In altre parole, gli autospazi dell'operatore T sono in somma diretta, nel senso dell'osservazione 1.11.

(ii) Se $TS = ST$, allora l'immagine $\operatorname{Im} T$ e gli autospazi $\operatorname{Ker}(T - \lambda I)$ di T sono invarianti per S, per ogni $\lambda \in \mathbb{F}$.

(iii) Se T è normale, $\lambda, \mu \in \mathbb{F}$ e $\lambda \neq \mu$, allora $\operatorname{Ker}(T - \lambda I) \perp \operatorname{Ker}(T - \mu I)$.

(iv) Se V è un sottospazio vettoriale di H invariante per T, allora anche $\overline{V}$ è invariante per T, mentre $V^\perp$ è invariante per T^*.

(v) Sia $n \in \mathbb{N}_+$. Se $V_1, \dots, V_n$ sono sottospazi invarianti per T, allora anche $V_1 + \dots + V_n$ e $V_1 \cap \dots \cap V_n$ sono invarianti per T.

Dimostrazione. (i). Procediamo per induzione su $n \in \mathbb{N}_+$. Per $n = 1$, l'enunciato è banalmente vero.

Prendiamo ora $n > 1$, e supponiamo che $x \in \operatorname{Ker}(T - \lambda_n I)$ appartenga anche allo spazio $+_{j=1}^{n-1} \operatorname{Ker}(T - \lambda_j I)$. Allora possiamo scrivere

$$x = y_1 + \dots + y_{n-1},$$

dove $y_j \in \operatorname{Ker}(T - \lambda_j I)$ per $j = 1, \dots, n-1$. Applicando T alla precedente identità si ottiene

$$\lambda_n x = \lambda_1 y_1 + \dots + \lambda_{n-1} y_{n-1}.$$

Moltiplicando la prima identità per λ_1 e prendendo la differenza con la seconda, si ottiene infine

$$(\lambda_n - \lambda_1)x = (\lambda_2 - \lambda_1)y_2 + \dots + (\lambda_{n-1} - \lambda_1)y_{n-1}.$$

In questa uguaglianza, il primo membro è in $\operatorname{Ker}(T - \lambda_n I)$, mentre il secondo è in $+_{j=2}^{n-1} \operatorname{Ker}(T - \lambda_j I)$; quindi ambo i membri sono nell'intersezione di questi spazi. Applicando l'ipotesi induttiva agli $n-1$ valori distinti $\lambda_2, \dots, \lambda_n$, deduciamo che tale intersezione è $\{0\}$, pertanto $(\lambda_n - \lambda_1)x = 0$; siccome $\lambda_n \neq \lambda_1$, si ha anche $x = 0$, come desiderato.

(ii). Se $x \in \operatorname{Im} T$, allora $x = Ty$, dunque $Sx = STy = TSy \in \operatorname{Im} T$. Dunque $\operatorname{Im} T$ è invariante per S.

Similmente, se $x \in \operatorname{Ker} T$, allora $Tx = 0$, dunque $TSx = STx = S0 = 0$, cioè $Sx \in \operatorname{Ker} T$. Dunque $\operatorname{Ker} T$ è invariante per S.

Notiamo infine, per ogni $\lambda \in \mathbb{F}$, $(T - \lambda I)S = TS - \lambda S = ST - \lambda S = S(T - \lambda I)$. Applicando quanto dimostrato con $T - \lambda I$ al posto di T, si ottiene che $\operatorname{Ker}(T - \lambda I)$ è invariante per S.

(iii). Notiamo che $\mathrm{Ker}(T - \mu I) = \mathrm{Ker}((T - \mu I)^*) = \mathrm{Ker}(T^* - \overline{\mu} I)$ per la proposizione 10.17. Dunque, se $x \in \mathrm{Ker}(T - \lambda I)$ e $y \in \mathrm{Ker}(T - \mu I) = \mathrm{Ker}(T^* - \overline{\mu} I)$, si ha

$$0 = \langle (T-\lambda I)x, y\rangle = \langle x, (T-\lambda I)^* y\rangle = \langle x, (T^*-\overline{\mu} I)y\rangle + (\mu-\lambda)\langle x, y\rangle = (\mu-\lambda)\langle x, y\rangle,$$

da cui, siccome $\lambda \neq \mu$, segue che $x \perp y$.

(iv). L'invarianza di $\overline{V}$ segue da quella di V e dalla continuità di T. Inoltre, per ogni $x \in V^\perp$, se $y \in V$ si ha anche $Ty \in V$ (perché V è invariante per T), dunque

$$\langle T^* x, y\rangle = \langle x, Ty\rangle = 0$$

dato che $x \in V^\perp$; questo dimostra che $T^* x \perp y$ per ogni $y \in V$, cioè $T^* x \in V^\perp$.

(v). Questo è conseguenza dell'invarianza dei V_j e della linearità di T. □

Osservazione 10.27. Come si vede dalle dimostrazioni, le parti (i), (ii) e (v) della proposizione 10.26 rimangono vere più in generale per operatori lineari $T, S \in \mathcal{L}(X)$ su uno spazio vettoriale X, senza bisogno di una norma o un prodotto scalare su X, né della limitatezza di T e S.

11 Operatori compatti e teoria di Fredholm

11.1 Operatori compatti

Definizione 11.1. Siano X, Y spazi normati.

(a) Un operatore $T \in \mathcal{B}(X, Y)$ è detto *compatto* se, per ogni successione $(x_n)_{n\in\mathbb{N}}$ limitata in X, la successione $(Tx_n)_{n\in\mathbb{N}}$ ha una sottosuccessione convergente in Y.

(b) Denotiamo con $\mathcal{K}(X, Y)$ l'insieme degli operatori compatti da X a Y. Scriviamo anche $\mathcal{K}(X)$ invece di $\mathcal{K}(X, X)$.

Osservazione 11.2. Se $T \in \mathcal{B}(X, Y)$ e $(x_n)_n$ è una successione limitata in X, allora

$$\sup_{n\in\mathbb{N}} \|Tx_n\|_Y \leq \|T\|_{\mathrm{op}} \sup_{n\in\mathbb{N}} \|x_n\|_X < \infty,$$

cioè anche $(Tx_n)_{n\in\mathbb{N}}$ è una successione limitata in Y. Dunque, se $\dim Y < \infty$, allora $(Tx_n)_n$ ha una sottosuccessione convergente in Y (☞ corollario 4.26). Questo mostra che $\mathcal{K}(X, Y) = \mathcal{B}(X, Y)$ ogniqualvolta $\dim Y < \infty$.

Ci sono diverse caratterizzazioni equivalenti degli operatori compatti.

Proposizione 11.3. *Siano X, Y spazi normati, $T \in \mathcal{B}(X, Y)$. Sono equivalenti:*

(i) T è un operatore compatto;

(ii) $\overline{T(A)}$ è un sottoinsieme compatto di Y per ogni sottoinsieme limitato A di X;

(iii) $\overline{T(B_X(0,1))}$ è un sottoinsieme compatto di Y.

Dimostrazione. (i) ⇒ (ii). Sia $A \subseteq X$ limitato. Sia $(y_n)_{n\in\mathbb{N}}$ una successione a valori in $\overline{T(A)}$. Per ogni $n \in \mathbb{N}$, si ha $y_n \in \overline{T(A)}$, quindi esiste $x_n \in A$ tale che $\|y_n - Tx_n\|_Y < 2^{-n}$. Siccome A è limitato, $(x_n)_n$ è una successione limitata in X, quindi per (i) esiste una sottosuccessione $(Tx_{n_k})_k$ di $(Tx_n)_n$ tale che $\lim_{k\to\infty} Tx_{n_k} = y$ in Y per qualche $y \in Y$.

Siccome $Tx_{n_k} \in T(A)$ per ogni $k \in \mathbb{N}$, si ha $y \in \overline{T(A)}$. Inoltre

$$\|y_{n_k} - y\|_Y = \|y_{n_k} - Tx_{n_k}\|_Y + \|Tx_{n_k} - y\|_Y \to 0$$

per $k \to \infty$, dove usiamo che $\|y_{n_k} - Tx_{n_k}\|_Y < 2^{-n_k} \to 0$ e che $Tx_{n_k} \to y$ in Y. Di conseguenza $y_{n_k} \to y$; in altre parole, la successione $(y_n)_n$ ha una sottosuccessione $(y_{n_k})_k$ convergente a un punto $y \in \overline{T(A)}$. Siccome $(y_n)_n$ è un'arbitraria successione a valori in $\overline{T(A)}$, ne concludiamo che $\overline{T(A)}$ è compatto (☞ definizione 2.21(k)).

(ii) ⇒ (iii). Ovvio, dato che $B_X(0,1)$ è limitato in X.

(iii) ⇒ (i). Sia $(x_n)_n$ una successione limitata in X; dunque

$$M := \sup_{n\in\mathbb{N}} \|x_n\|_X < \infty.$$

Pertanto $\|x_n/(2M)\|_X \leq 1/2$, dunque $x_n/(2M) \in B_X(0,1)$ e $T(x_n/(2M)) \in \overline{T(B_X(0,1))}$ per ogni $n \in \mathbb{N}$. Siccome $\overline{T(B_X(0,1))}$ è compatto per (iii), la successione $(T(x_n/(2M))_{n\in\mathbb{N}}$ ha una sottosuccessione $(T(x_{n_k}/(2M))_k$ convergente (☞ definizione 2.21(k)), cioè $T(x_{n_k}/(2M)) \to y$ in Y per qualche $y \in \overline{T(B_X(0,1))} \subseteq Y$.

Moltiplicando per $2M$ e usando la linearità di T, si ottiene che $T(x_{n_k}) \to 2My$ in Y. Dunque $(Tx_{n_k})_k$ è la sottosuccessione di $(Tx_n)_n$ cercata. □

Corollario 11.4. *Siano X, Y spazi normati e $T \in \mathcal{K}(X,Y)$. Allora $\overline{\operatorname{Im} T}$ è separabile.*

Dimostrazione (cenno). Basta dimostrare che $\operatorname{Im} T$ è separabile. D'altra parte,

$$\operatorname{Im} T = \bigcup_{n\in\mathbb{N}_+} T(B_X(0,n)) \subseteq \bigcup_{n\in\mathbb{N}_+} \overline{T(B_X(0,n))}.$$

Siccome $B_X(0,n)$ è limitato in X e $T \in \mathcal{K}(X,Y)$, per la proposizione 11.3 l'insieme $\overline{T(B_X(0,n))}$ è compatto in Y, dunque è separabile (☞ esempio 2.26(5)). Ma allora $\operatorname{Im} T$ è contenuta in un'unione numerabile di insiemi separabili, e di conseguenza è anch'essa separabile (☞ esempio 2.26(4)). □

Proposizione 11.5. *Siano X, Y spazi normati. Sia $T \in \mathcal{B}(X,Y)$ un operatore di* rango finito*, cioè tale che $\dim \operatorname{Im} T < \infty$. Allora $T \in \mathcal{K}(X,Y)$.*

Dimostrazione. Sia $(x_n)_{n\in\mathbb{N}}$ una successione limitata in X. Allora (☞ osservazione 11.2) $(Tx_n)_n$ è a sua volta una successione limitata in Y. In effetti, $(Tx_n)_n$ è a valori in $V := \operatorname{Im} T$. Se dotiamo V della norma indotta da Y, si ha allora che $(Tx_n)_n$ è limitata in V. Siccome $\dim V < \infty$ per ipotesi, deduciamo (☞ corollario 4.26) che $(Tx_n)_n$ ha una sottosuccessione convergente in V, dunque anche in Y. □

Corollario 11.6. *Siano X, Y spazi normati. Se $\dim X < \infty$ oppure $\dim Y < \infty$, allora $\mathcal{K}(X,Y) = \mathcal{B}(X,Y)$.*

Osservazione 11.7. Nella proposizione 11.5 abbiamo dimostrato l'implicazione "T di rango finito $\Rightarrow$ T compatto". Vedremo in seguito che, in generale, tale implicazione non si può invertire.

Esempio 11.8. Sia $T : L^2(0,1) \to L^2(0,1)$ definito da

$$Tf(t) = t\int_0^1 f(s)\,ds \tag{11.1}$$

per ogni $f \in L^2(0,1)$ e quasi ogni $t \in (0,1)$. Dimostriamo che $T \in \mathcal{K}(L^2(0,1))$. Si ha infatti che T è l'operatore integrale T_K con nucleo integrale

$$K(t,s) = t \qquad \forall t,s \in (0,1),$$

e siccome

$$\|K\|^2_{L^2((0,1)\times(0,1))} = \int_0^1\int_0^1 t^2\,dt\,ds = 1/2 < \infty$$

si ha che $T \in \mathcal{B}(L^2(0,1))$ (☞ esempio 7.8(6)). Inoltre, dalla formula (11.1) si vede che $\operatorname{Im} T \subseteq \operatorname{span}\{\varphi\}$, ove $\varphi \in L^2(0,1)$ è data da $\varphi(t) = t$. Questo mostra che T ha rango finito e pertanto (☞ proposizione 11.5) è un operatore compatto.

Vediamo ora come gli operatori compatti si comportano rispetto alla composizione di operatori.

Proposizione 11.9. *Siano X, Y, Z spazi normati. Siano $T \in \mathcal{B}(X, Y)$ e $S \in \mathcal{B}(Y, Z)$. Se T è un operatore compatto, oppure S è un operatore compatto, allora ST è un operatore compatto.*

Dimostrazione. Caso 1: $T \in \mathcal{K}(X, Y)$. Se $(x_n)_n$ è una successione limitata in X, per la compattezza di T esiste una sottosuccessione $(Tx_{n_k})_k$ di $(Tx_n)_n$ tale che $Tx_{n_k} \to y$ in Y per qualche $y \in Y$. Siccome $S : Y \to Z$ è continuo, se ne deduce che $STx_{n_k} \to Sy$ in Z, cioè $(STx_{n_k})_k$ è una sottosuccessione convergente di $(STx_n)_n$. Per l'arbitrarietà di $(x_n)_n$, questo dimostra che $ST \in \mathcal{K}(X, Z)$.

Caso 2: $S \in \mathcal{K}(X, Y)$. Se $(x_n)_n$ è una successione limitata in X, per l'osservazione 11.2 anche $(Tx_n)_n$ è una successione limitata in Y. Siccome S è compatto, $(STx_n)_n$ ha una sottosuccessione convergente in Z, come desiderato. □

Proposizione 11.10. *Sia X uno spazio normato con $\dim X = \infty$.*

(i) $\mathrm{id}_X \notin \mathcal{K}(X)$.

(ii) Se Y è uno spazio normato e $T \in \mathcal{B}(X, Y)$ è un isomorfismo, allora $T \notin \mathcal{K}(X, Y)$.

Dimostrazione. (i). Siccome $\dim X = \infty$, esiste una successione $(x_n)_n$ limitata in X che non ha sottosuccessioni convergenti (☞ corollario 4.26). Siccome $(\mathrm{id}_X x_n)_n = (x_n)_n$, questa successione mostra (☞ definizione 11.1) che id_X non è compatto.

(ii). Se per assurdo T fosse compatto, siccome per ipotesi $T^{-1} \in \mathcal{B}(Y, X)$, potremmo scrivere $\mathrm{id}_X = T^{-1}T$, cioè id_X sarebbe il prodotto di un operatore limitato e di un operatore compatto; dunque id_X sarebbe compatto (☞ proposizione 11.9), ma ciò contraddice (i) perché $\dim X = \infty$. □

Esempio 11.11. Sia $S \in \mathcal{B}(\ell^2)$ l'operatore di shift verso sinistra (☞ esempio 10.4(2)). Sappiamo da (10.9) che $SS^* = \mathrm{id}_{\ell^2}$. In particolare, né S né S^* sono operatori compatti su ℓ^2: infatti, se per assurdo uno dei due lo fosse, anche il loro prodotto $SS^* = \mathrm{id}_{\ell^2}$ sarebbe compatto (☞ proposizione 11.9), ma questo è assurdo perché ℓ^2 ha dimensione infinita e quindi id_{ℓ^2} non è compatto (☞ proposizione 11.10(i)).

Teorema 11.12. *Siano X, Y spazi normati.*

(i) $\mathcal{K}(X, Y)$ è un sottospazio vettoriale di $\mathcal{B}(X, Y)$.

(ii) Se Y è uno spazio di Banach, allora $\mathcal{K}(X, Y)$ è un sottoinsieme chiuso di $(\mathcal{B}(X, Y), \|\cdot\|_{\mathrm{op}})$.

Dimostrazione. (i). Chiaramente $0 \in \mathcal{K}(X, Y)$, dato che l'operatore nullo ha rango finito ($\mathfrak{Im}\, 0 = \{0\}$). Siano ora $T, S \in \mathcal{K}(X, Y)$ e $\alpha, \beta \in \mathbb{F}$. Vogliamo dimostrare che $\alpha T + \beta S \in \mathcal{K}(X, Y)$.

Sia $(x_n)_n$ una successione limitata in X. Siccome $T \in \mathcal{K}(X, Y)$, la successione $(Tx_n)_n$ ha una sottosuccessione $(Tx_{n_k})_k$ convergente in Y. Ora, $(x_{n_k})_k$ è una sottosuccessione di $(x_n)_n$, quindi anche $(x_{n_k})_k$ è limitata in X; pertanto, siccome $S \in \mathcal{K}(X, Y)$, la successione $(Sx_{n_k})_k$ ha una sottosuccessione $(Sx_{n_{k_h}})_h$ convergente in Y. Infine, $(Tx_{n_{k_h}})_h$ è una sottosuccessione di $(Tx_{n_k})_k$, quindi anche $(Tx_{n_{k_h}})_h$ è convergente in Y. Per l'algebra dei limiti, ne concludiamo che anche $(\alpha Tx_{n_{k_h}} + \beta Sx_{n_{k_h}})_h = ((\alpha T+\beta S)x_{n_{k_h}})_h$ è convergente in Y; quest'ultima è la sottosuccessione di $((\alpha T + \beta S)x_n)_n$ cercata.

(ii). Sia $(T_n)_{n\in\mathbb{N}}$ una successione a valori in $\mathcal{K}(X,Y)$, e supponiamo che $T_n \to T$ in $\mathcal{B}(X,Y)$ per qualche $T \in \mathcal{B}(X,Y)$. Vogliamo dimostrare che $T \in \mathcal{K}(X,Y)$.

Sia dunque $(x_k)_{k\in\mathbb{N}}$ una successione limitata in X. Vogliamo mostrare che $(Tx_k)_k$ ha una sottosuccessione convergente in Y. Costruiamo induttivamente, per ogni $n \in \mathbb{N}$, una successione $(x_k^{(n)})_k$ in X con le seguenti proprietà:

- $(x_k^{(n)})_k$ è una sottosuccessione di $(x_k)_k$ e, se $n > 0$, anche di $(x_k^{(n-1)})_k$;
- $T_n x_k^{(n)} \to y_n$ in Y per $k \to \infty$, per qualche $y_n \in Y$.

Per $n = 0$, si applica semplicemente la compattezza di T_0 alla successione limitata $(x_k)_k$ in X per costruire la sottosuccessione $(x_k^{(0)})_k$. Supponendo, per $n \in \mathbb{N}_+$, di aver costruito $(x_k^{(0)})_k, \dots, (x_k^{(n-1)})_k$, si ha che $(x_k^{(n-1)})_k$ è limitata in X (in quanto sottosuccessione di $(x_k)_k$), quindi possiamo applicare la compattezza di T_n alla successione $(x_k^{(n-1)})_k$ per costruire la sottosuccessione $(x_k^{(n)})_k$.

Notiamo che, per costruzione,

- $(x_k^{(n)})_k$ è sottosuccessione di $(x_k^{(m)})_k$ per ogni $n \geq m$.

Definiamo ora $x_k^* = x_k^{(k)}$ per ogni $k \in \mathbb{N}$. Allora, $(x_k^*)_k$ è una sottosuccessione di $(x_k)_k$, e inoltre, per ogni $n \in \mathbb{N}$, si ha che $(x_k^*)_{k\geq n}$ è una sottosuccessione di $(x_k^{(n)})_{k\geq n}$. Di conseguenza, $(T_n x_k^*)_{k\geq n}$ è una sottosuccessione di $(T_n x_k^{(n)})_{k\geq n}$, ma allora

- $T_n x_k^* \to y_n$ in Y per $k \to \infty$, per ogni $n \in \mathbb{N}$.

Notiamo ora che $M \coloneqq \sup_{k\in\mathbb{N}} \|x_k^*\|_X < \infty$, dato che $(x_k^*)_k$ è una sottosuccessione di $(x_k)_k$ e dunque è limitata. Quindi, per ogni $n, h, k \in \mathbb{N}$,

$$\begin{aligned}\|Tx_h^* - Tx_k^*\|_Y &\leq \|Tx_h^* - T_n x_h^*\|_Y + \|T_n x_h^* - T_n x_k^*\|_Y + \|T_n x_k^* - Tx_k^*\|_Y \\ &\leq 2M\|T - T_n\|_{\mathrm{op}} + \|T_n x_h^* - y_n\|_Y + \|T_n x_k^* - y_n\|_Y. \end{aligned} \tag{11.2}$$

Per ogni $\epsilon > 0$, siccome $T_n \to T$ in $\mathcal{B}(X,Y)$, troviamo $N \in \mathbb{N}$ tale che $\|T - T_n\|_{\mathrm{op}} < \epsilon/(4M)$. Fissato un tale $n > N$, siccome $T_n x_k^* \to y_n$ per $k \to \infty$, troviamo $K \in \mathbb{N}$ tale che $\|T_n x_k^* - y_n\|_Y < \epsilon/4$ per ogni $k > K$. In particolare, da (11.2) deduciamo che, per ogni $h, k > K$,

$$\|Tx_h^* - Tx_k^*\|_Y < 2M\epsilon/(4M) + \epsilon/4 + \epsilon/4 = \epsilon.$$

Per l'arbitrarietà di $\epsilon > 0$, questo dimostra che $(Tx_k^*)_k$ è di Cauchy in Y. Siccome Y è completo, $(Tx_k^*)_k$ converge in Y. Dunque $(Tx_k^*)_k$ è la sottosuccessione di $(Tx_k)_k$ cercata. □

Corollario 11.13. *Siano X uno spazio normato e Y uno spazio di Banach.*

(i) $(\mathcal{K}(X,Y), \|\cdot\|_{\mathrm{op}})$ *è uno spazio di Banach.*

(ii) Se $(T_n)_n$ è una successione in $\mathcal{B}(X,Y)$ di operatori di rango finito, e $T_n \to T$ in $\mathcal{B}(X,Y)$ per $n \to \infty$, allora $T \in \mathcal{K}(X,Y)$.

Osservazione 11.14. Quando si utilizza il corollario 11.13 per dimostrare la compattezza di un operatore $T \in \mathcal{B}(X,Y)$, scrivendolo come limite di una successione $(T_n)_n$ di operatori di rango finito, è importante ricordare che non basta dimostrare la convergenza puntuale

$$T_n x \xrightarrow[n\to\infty]{} Tx \quad \text{in } Y \qquad \forall x \in X$$

della successione; occorre invece dimostrarne la convergenza in $\mathcal{B}(X,Y)$, cioè la convergenza in norma operatoriale

$$\|T_n - T\|_{\mathrm{op}} \xrightarrow[n\to\infty]{} 0.$$

Come vedremo (☞ osservazione 11.18), la sola convergenza puntuale degli operatori in generale non garantisce la compattezza del limite.

Esempio 11.15. Usiamo il precedente risultato per stabilire la compattezza di alcune classi di operatori.

(1) Per $\underline{w} \in \ell^\infty$, sia $D_{\underline{w}} \in \mathcal{B}(\ell^2)$ l'operatore di moltiplicazione per $\underline{w}$ (☞ esempio 7.8(8)). Chi chiediamo per quali $\underline{w} \in \ell^\infty$ si ha $D_{\underline{w}} \in \mathcal{K}(\ell^2)$.

Anzitutto, se $\underline{w} \in c_{00}$, esiste $K \in \mathbb{N}$ tale che $w_k = 0$ per ogni $k > K$. Pertanto, per ogni $\underline{x} \in \ell^2$, si ha anche $D_{\underline{w}}\underline{x} = \underline{w} \cdot \underline{x} = (w_k x_k)_k$, e $w_k x_k = 0$ per ogni $k > K$. Questo mostra che

$$\operatorname{Im} D_{\underline{w}} \subseteq \{\underline{y} \in \ell^2 : y_k = 0 \ \forall k > K\} = \operatorname{span}\{\underline{e^{(0)}}, \dots, \underline{e^{(K)}}\}.$$

Dunque $D_{\underline{w}}$ ha rango finito, e in particolare è compatto, ogniqualvolta $\underline{w} \in c_{00}$.

In effetti, per ogni $\underline{w} \in \ell^\infty$, si ha $D_{\underline{w}}\underline{e^{(n)}} = w_n \underline{e^{(n)}}$; quindi $\underline{e^{(n)}} = \frac{1}{w_n} D_{\underline{w}}\underline{e^{(n)}} \in \operatorname{Im} D_{\underline{w}}$ ogniqualvolta $w_n \neq 0$. Questo mostra che

$$\operatorname{Im} D_{\underline{w}} \supseteq \operatorname{span}\{\underline{e^{(n)}} : n \in \mathbb{N},\ w_n \neq 0\}$$

e in particolare $\dim \operatorname{Im} D_{\underline{w}} = \infty$ se $\sharp\{n \in \mathbb{N} : \ w_n \neq 0\} = \infty$, cioè se $\underline{w} \notin c_{00}$. Ne concludiamo che $D_{\underline{w}}$ ha rango finito se e solo se $\underline{w} \in c_{00}$.

Sia ora $\underline{w} \in c_0$. Siccome c_{00} è denso in c_0 (☞ proposizione 5.6(iii)), possiamo trovare una successione $(\underline{w^{(n)}})_{n\in\mathbb{N}}$ a valori in c_{00} tale che $\|\underline{w} - \underline{w^{(n)}}\|_\infty \to 0$ per $n \to \infty$. Siccome $w^{(n)} \in c_{00}$, l'operatore $D_{\underline{w^{(n)}}}$ è di rango finito, per ogni $n \in \mathbb{N}$. D'altra parte, sappiamo che la mappa $\underline{w} \mapsto D_{\underline{w}}$ è un'isometria lineare da ℓ^∞ a $\mathcal{B}(\ell^2)$ (☞ esempio 10.11(2)); ne concludiamo che

$$\|D_{\underline{w}} - D_{\underline{w^{(n)}}}\|_{\mathrm{op}} = \|\underline{w} - \underline{w^{(n)}}\|_\infty \to 0$$

per $n \to \infty$. Abbiamo dunque mostrato che $D_{\underline{w}}$ è limite in $\mathcal{B}(\ell^2)$ di una successione di operatori di rango finito, quindi, per il corollario 11.13, l'operatore $D_{\underline{w}}$ è compatto.

Viceversa, se $\underline{w} \notin c_0$, allora $(w_n)_n$ non è infinitesima, dunque esiste una successione strettamente crescente di indici $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ tale che

$$\delta := \inf_{k\in\mathbb{N}} |w_{n_k}| > 0.$$

Allora $(\underline{e}^{(n_k)})_{k\in\mathbb{N}}$ è una successione limitata in ℓ^2 (dato che $\|\underline{e}^{(n_k)}\|_2 = 1$ per ogni $k \in \mathbb{N}$). Tuttavia, $(D_{\underline{w}}\underline{e}^{(n_k)})_{k\in\mathbb{N}}$ non ha sottosuccessioni convergenti: infatti, per ogni $k, h \in \mathbb{N}$, se $k \neq h$,

$$\|D_{\underline{w}}\underline{e}^{(n_k)} - D_{\underline{w}}\underline{e}^{(n_h)}\|_2 = \|w_{n_k}\underline{e}^{(n_k)} - w_{n_h}\underline{e}^{(n_h)}\|_2 = \sqrt{|w_{n_k}|^2 + |w_{n_h}|^2} \geq \delta,$$

e siccome $\delta > 0$, questo mostra che $(D_{\underline{w}}\underline{e}^{(n_k)})_k$ non ha sottosuccessioni di Cauchy. Di conseguenza, $D_{\underline{w}} \notin \mathcal{K}(\ell^2)$.

In conclusione, $D_{\underline{w}} \in \mathcal{K}(\ell^2)$ se e solo se $\underline{w} \in c_0$.

(2) Siano $[a, b], [c, d]$ intervalli limitati in $\mathbb{R}$. Per $K \in L^2((a, b) \times (c, d))$, sia $T_K \in \mathcal{B}(L^2(c, d), L^2(a, b))$ l'operatore integrale con nucleo K (☞ esempio 7.8(6)). Mostriamo ora che in effetti $T_K \in \mathcal{K}(L^2(c, d), L^2(a, b))$.

Siano $\{\phi_n\}_{n\in\mathbb{N}}$ e $\{\psi_m\}_{m\in\mathbb{N}}$ basi ortonormali di $L^2(a, b)$ e $L^2(c, d)$ rispettivamente (☞ sezione 6.4). Allora (☞ proposizione 6.54) $\{\phi_n \otimes \psi_m\}_{n,m\in\mathbb{N}}$ è una base ortonormale di $L^2((a, b) \otimes (c, d))$. Siccome $K \in L^2((a, b) \times (c, d))$, possiamo sviluppare K rispetto a questa base:

$$K(x, y) = \sum_{n,m\in\mathbb{N}} \alpha_{n,m}\phi_n(x)\psi_m(y),$$

dove

$$\alpha_{n,m} = \langle K, \phi_n \otimes \psi_m\rangle_{L^2((a,b)\times(c,d))}$$

e la serie converge in $L^2((a, b) \times (c, d))$ (☞ proposizione 6.44).

Consideriamo ora, per $N \in \mathbb{N}$, la somma parziale

$$K_N(x, y) = \sum_{n,m\leq N} \alpha_{n,m}\phi_n(x)\psi_m(y)$$

della serie; allora $K_N \in L^2((a, b) \times (c, d))$ e $K_N \to K$ in L^2 per $N \to \infty$, quindi

$$\|T_K - T_{K_N}\|_{\mathrm{op}} = \|T_{K-K_N}\|_{\mathrm{op}} \leq \|K - K_N\|_{L^2((a,b)\times(c,d))} \to 0$$

per $N \to \infty$, cioè $T_{K_N} \to T_K$ in $\mathcal{B}(L^2(c, d), L^2(a, b))$.

Infine, per ogni $N \in \mathbb{N}$ e $f \in L^2(c, d)$,

$$T_{K_N}f(x) = \int_c^d K_N(x, y)f(y)\,dy = \sum_{n,m\leq N} \phi_n(x) \int_c^d \psi_m(y)\,f(y)\,dy,$$

da cui segue che

$$\operatorname{Im} T_{K_N} \subseteq \operatorname{span}\{\phi_n\}_{n\leq N}$$

e in particolare T_{K_N} ha rango finito. Ma allora $T = \lim_{N\to\infty} T_{K_N}$ è compatto (☞ corollario 11.13).

(3) L'esempio precedente si può generalizzare al caso di operatori integrali $T_K \in \mathcal{B}(L^2(M_2), L^2(M_1))$ dove (M_1, μ_1) e (M_2, μ_2) sono spazi di misura σ-finiti e $K \in L^2(M)$, dove $M = M_1 \times M_2$ è dotato della misura prodotto (☞ esempio 7.8(7)). Specificamente, se $L^2(M_1)$ e $L^2(M_2)$ sono separabili, allora si possono prendere basi ortonormali $\{\phi_n\}_{n\in\mathbb{N}}$ di $L^2(M_1)$ e $\{\psi_m\}_{m\in\mathbb{N}}$ e $L^2(M_2)$ e ripetere la dimostrazione nell'esempio (2) per mostrare che $T_K \in \mathcal{K}(L^2(M_2), L^2(M_1))$.

(4) Ricordiamo che, quando $K \in C([a,b] \times [c,d])$, l'operatore integrale con nucleo integrale K si può anche considerare come un operatore limitato $T_K : C[c,d] \to C[a,b]$ (☞ esempio 7.8(5)). In effetti, sotto queste ipotesi è possibile dimostrare che $T_K \in \mathcal{K}(C[c,d], C[a,b])$. Discutiamo brevemente l'idea della dimostrazione.

Se $(f_n)_n$ è una successione limitata in $C[c,d]$, allora dall'osservazione 11.2 sappiamo che la successione $(T_K f_n)_n$ è limitata in $C[a,b]$. Inoltre, utilizzando il fatto che $K : [a,b] \times [c,d] \to \mathbb{C}$ è uniformemente continua (☞ teorema 2.32), dall'espressione

$$T_K f_n(x) = \int_c^d K(x,y) f_n(y)\, dy \tag{11.3}$$

e dalla limitatezza di $(f_n)_n$ si può dedurre che la successione $(T_K f_n)_n$ è *uniformemente equicontinua*, cioè

$$\forall \epsilon > 0 : \exists \delta > 0 : \forall n \in \mathbb{N} : \forall x, x' \in [a,b] : \\ (|x - x'| < \delta \ \Rightarrow\ |T_K f_n(x) - T_K f_n(x')| < \epsilon)$$

(si parla di equicontinuità perché la scelta di δ dato ϵ non dipende da n, cioè si fa in maniera uniforme per tutte le funzioni $T_K f_n$ della successione; il motivo per cui in questo caso si ha equicontinuità è che la dipendenza da x dell'espressione (11.3) avviene tramite la funzione K, quindi la scelta di δ dato ϵ si basa sulla continuità della sola K e si può fare indipendentemente da n). Il teorema di Ascoli–Arzelà (☞ teorema 2.38) implica allora che $(T_K f_n)_n$ ha una sottosuccessione convergente in $C[a,b]$.

Osservazione 11.16. L'esempio 11.15(2) mostra che gli operatori integrali T_K con nucleo integrale $K \in L^2((a,b)\times(c,d))$ sono ben lungi da rappresentare tutti gli operatori in $\mathcal{B}(L^2(c,d), L^2(a,b))$. Ad esempio, se $(a,b) = (c,d)$, l'operatore identità su $L^2(a,b)$ non è compatto (☞ proposizione 11.10), quindi non si può rappresentare come operatore integrale con nucleo integrale $K \in L^2((a,b)^2)$. Analoghe considerazioni valgono per lo spazio $\mathcal{B}(C[c,d], C[a,b])$ e gli operatori integrali con nucleo integrale continuo discussi nell'esempio 11.15(4).

Abbiamo visto nel corollario 11.13 che, quando X è uno spazio normato e Y è uno spazio di Banach, se $T \in \mathcal{B}(X,Y)$ è il limite in $\mathcal{B}(X,Y)$ di una successione di operatori di rango finito, allora T è un operatore compatto. Vediamo ora che, nel caso Y sia uno spazio di Hilbert, tale implicazione si può invertire.

Teorema 11.17. *Siano X uno spazio normato e H uno spazio di Hilbert. Se $T \in \mathcal{K}(X,H)$, allora esiste una successione $(T_n)_{n\in\mathbb{N}}$ in $\mathcal{B}(X,H)$ di operatori di rango finito tale che $T_n \to T$ in $\mathcal{B}(X,H)$ per $n \to \infty$.*

Dimostrazione. Sia $T \in \mathcal{K}(X,H)$. Allora (☞ corollario 11.4) $\overline{\operatorname{Im} T}$ è un sottospazio vettoriale chiuso separabile di H; dunque, con il prodotto scalare indotto da H, lo spazio $\overline{\operatorname{Im} T}$ è uno spazio di Hilbert separabile, e $T : X \to \overline{\operatorname{Im} T}$ è compatto. A meno di rimpiazzare H con $\overline{\operatorname{Im} T}$, possiamo allora supporre che H stesso sia separabile. D'altra parte, se H ha dimensione finita, allora T è esso stesso di rango finito, e possiamo banalmente prendere $T_n = T$ per ogni $n \in \mathbb{N}$; dunque, nel seguito possiamo assumere che invece $\dim H = \infty$.

Sia dunque $\{e_n\}_{n\in\mathbb{N}}$ una base ortonormale di H (☞ teorema 6.46). Per ogni $N \in \mathbb{N}$, sia $P_N \in \mathcal{B}(H)$ la proiezione ortogonale sul sottospazio $\mathrm{span}\{e_n\}_{n\leq N}$, e sia $Q_N = \mathrm{id}_H - P_N$ la proiezione complementare su $\{e_n\}_{n\leq N}^{\perp}$ (☞ corollario 6.33(iii)). Osserviamo che, per ogni $y \in H$,

$$P_N y = \sum_{n=0}^{N} \langle y, e_n\rangle e_n \xrightarrow[N\to\infty]{} y \quad \text{in } H \tag{11.4}$$

(☞ proposizioni 6.39 e 6.44(iii)).

Ora, se poniamo $T_N = P_N T$, si ha $\mathrm{Im}\, T_N \subseteq \mathrm{Im}\, P_n = \mathrm{span}\{e_n\}_{n\leq N}$, che ha dimensione finita, dunque T_N ha rango finito per ogni $N \in \mathbb{N}$. Per concludere, è sufficiente dimostrare che $T_N \to T$ in $\mathcal{B}(X,H)$; siccome $T - T_N = Q_N T$, vogliamo in altre parole dimostrare che $\|Q_N T\|_{\mathrm{op}} \to 0$ per $N \to \infty$.

Supponiamo per assurdo che ciò non valga. Allora esiste una successione strettamente crescente $(N_k)_{k\in\mathbb{N}}$ di indici in $\mathbb{N}$ tale che

$$\delta := \inf_{k\in\mathbb{N}} \|Q_{N_k} T\|_{\mathrm{op}} > 0.$$

In particolare, possiamo trovare (☞ lemma 7.5), per ogni $k \in \mathbb{N}$, un punto $x_k \in X$ tale che

$$\|x_k\|_X \leq 1, \qquad \|Q_{N_k} T x_k\|_H \geq \delta/2. \tag{11.5}$$

Dunque, la successione $(x_k)_k$ è limitata in X; siccome T è compatto, a meno di rimpiazzare $(N_k)_k$ con una sottosuccessione, possiamo assumere che

$$Tx_k \to y \text{ in } H \text{ per } k \to \infty, \tag{11.6}$$

per qualche $y \in H$.

Si ha allora

$$\|Q_{N_k} T x_k\|_H = \|Q_{N_k}(Tx_k - y)\|_H + \|Q_{N_k} y\|_H \leq \|Tx_k - y\|_H + \|y - P_{N_k} y\|_H,$$

dove abbiamo usato che $\|Q_{N_k}\|_{\mathrm{op}} \leq 1$ (☞ corollario 6.33(iv), dato che Q_{N_k} è una proiezione ortogonale). Per (11.4) e (11.6), da questa stima deduciamo che

$$\|Q_{N_k} T x_k\|_H \to 0 \text{ per } k \to \infty;$$

questo tuttavia contraddice il fatto che

$$\inf_{k\in\mathbb{N}} \|Q_{N_k} T x_k\|_H \geq \delta/2 > 0$$

per (11.5), assurdo. □

Osservazione 11.18. Dalla dimostrazione del teorema 11.17 si vede in effetti che, per ogni $T \in \mathcal{B}(X,H)$ e $x \in H$, si ha

$$P_N T x \xrightarrow[N\to\infty]{} Tx \quad \text{in } H,$$

grazie a (11.4). In altre parole, se H è uno spazio di Hilbert separabile, ogni operatore limitato $T \in \mathcal{B}(X,H)$ è limite puntuale di una successione $(P_N T)_N$ di operatori di rango finito; per questo risultato non c'è bisogno di assumere che T sia compatto. Ciò che rende complicata la dimostrazione del teorema 11.17 è che invece siamo interessati alla convergenza in norma $\|\cdot\|_{\mathrm{op}}$ della successione, cioè a dimostrare che

$$\|P_N T - T\|_{\mathrm{op}} \xrightarrow[N\to\infty]{} 0;$$

per questa convergenza più forte l'ipotesi di compattezza di T è essenziale.

Osservazione 11.19. Ci si può chiedere se la "proprietà di approssimazione" enunciata nel teorema 11.17, cioè l'approssimabilità in norma operatoriale di un arbitrario operatore compatto $T \in \mathcal{K}(X,Y)$ con operatori di rango finito, valga più in generale quando lo spazio di arrivo Y non si assume essere uno spazio di Hilbert, ma solo uno spazio di Banach. La dimostrazione qui data utilizza in maniera fondamentale basi ortonormali e proiezioni ortogonali, specifiche degli spazi di Hilbert. Sarebbe possibile adattare la dimostrazione al caso in cui lo spazio di arrivo Y sia uno spazio di Banach che ammette una cosiddetta "base di Schauder"; tuttavia, esistono anche spazi di Banach Y per cui la suddetta proprietà di approssimazione non vale[2].

Combinando il precedente teorema con il corollario 11.13 otteniamo la seguente caratterizzazione degli operatori compatti a valori in spazi di Hilbert.

Corollario 11.20. *Siano X uno spazio normato e H uno spazio di Hilbert. Sia $T \in \mathcal{B}(X,H)$. Sono equivalenti:*

(i) $T \in \mathcal{K}(X,H)$;

(ii) esiste una successione $(T_n)_n$ in $\mathcal{B}(X,H)$ di operatori di rango finito tale che $T_n \to T$ in $\mathcal{B}(X,H)$.

Utilizzando questa caratterizzazione, possiamo facilmente dimostrare le seguenti proprietà, che mostrano come le nozioni di operatore compatto e operatore di rango finito interagiscono con la costruzione dell'aggiunto.

Proposizione 11.21. *Siano H_1, H_2 spazi di Hilbert. Sia $T \in \mathcal{B}(H_1,H_2)$.*

(i) T ha rango finito se e solo se T^* ha rango finito, e in tal caso

$$\dim \operatorname{Im} T = \dim \operatorname{Im}(T^*).$$

(ii) $T \in \mathcal{K}(H_1,H_2)$ se e solo se $T^* \in \mathcal{K}(H_2,H_1)$.

Dimostrazione. (i). Supponiamo che T abbia rango finito. Dunque $\dim \operatorname{Im} T < \infty$, e in particolare $\operatorname{Im} T$ è un sottospazio vettoriale chiuso di H_2 (☞ proposizione 4.16(ii)).

Ora, per la proposizione 10.5(vi), abbiamo la decomposizione ortogonale

$$H_2 = \operatorname{Ker}(T^*) \oplus \overline{\operatorname{Im} T} = \operatorname{Ker}(T^*) \oplus \operatorname{Im} T,$$

dove abbiamo usato che $\operatorname{Im} T$ è chiusa; dunque

$$\operatorname{Im}(T^*) = T^*(H_2) = T^*(\operatorname{Ker}(T^*) \oplus \operatorname{Im} T) = T^*(\operatorname{Im} T)$$

(dato che T^* è lineare e $T^*(\operatorname{Ker}(T^*)) = \{0\}$ per definizione di nucleo), ma allora

$$\dim \operatorname{Im}(T^*) = \dim T^*(\operatorname{Im} T) \le \dim \operatorname{Im} T < \infty$$

[2] P. Enflo, A counterexample to the approximation problem in Banach spaces, *Acta Math.* 130 (1973), 309–317.

e quindi anche T^* ha rango finito.

Dato che $T^{**} = T$ (☞ proposizione 10.5(i)), si può ripetere l'argomento scambiando i ruoli di T e T^*, dimostrando dunque che, se T^* ha rango finito, allora anche T ha rango finito e

$$\dim \operatorname{Im} T \leq \dim \operatorname{Im}(T^*);$$

le due disuguaglianze insieme dimostrano che $\dim \operatorname{Im} T = \dim \operatorname{Im}(T^*)$, come desiderato.

(ii). Per l'involutività dell'aggiunto (☞ proposizione 10.5(i)), è sufficiente dimostrare che, se T è compatto, allora anche T^* lo è.

D'altra parte, se $T \in \mathcal{K}(H_1, H_2)$, per il corollario 11.20 esiste una successione $(T_n)_n$ in $\mathcal{B}(H_1, H_2)$ di operatori di rango finito tali che $T_n \to T$ in $\mathcal{B}(H_1, H_2)$. D'altra parte, la mappa $A \mapsto A^*$ è un'isometria antilineare (☞ corollario 10.6), pertanto si ha anche $T_n^* \to T^*$ in $\mathcal{B}(H_2, H_1)$. Inoltre, da (i) sappiamo che gli operatori T_n^* sono di rango finito. Applicando nuovamente il corollario 11.20, deduciamo che anche $T \in \mathcal{K}(H_2, H_1)$. □

11.2 Teoria di Fredholm

Nel corso di questa sezione, H è uno spazio di Hilbert su $\mathbb{F}$. Scriviamo per brevità I anziché id_H.

Osservazione 11.22. Per ogni $T \in \mathcal{B}(H)$ si ha

$$\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim H$$

(☞ proposizione 1.20(xi)); in particolare, se $\dim H < \infty$, si ha

$$T \text{ iniettivo} \iff T \text{ suriettivo} \iff T \text{ biiettivo}$$

(☞ osservazione 1.21). Nel caso $\dim H = \infty$, queste equivalenze in generale non valgono: ad esempio, l'operatore di shift verso sinistra $S \in \mathcal{B}(\ell^2)$ dell'esempio 10.4(2) è suriettivo ma non iniettivo, mentre il suo aggiunto S^* è iniettivo ma non suriettivo. Tuttavia, identificheremo ora una sottoclasse di operatori limitati su H per cui queste equivalenze valgono.

Definizione 11.23. Un operatore della forma $I - K$, dove $K \in \mathcal{K}(H)$, si dice *perturbazione compatta dell'identità.*

Teorema 11.24. *Sia $T \in \mathcal{B}(H)$ una perturbazione compatta dell'identità. Allora:*

(i) $\operatorname{Im} T$ *è un sottospazio vettoriale chiuso di H. In particolare, si ha la decomposizione ortogonale*

$$H = \operatorname{Ker}(T^*) \oplus \operatorname{Im} T. \tag{11.7}$$

(ii) $\dim \operatorname{Ker} T = \dim \operatorname{Ker}(T^*) < \infty$.

Dimostrazione. Se $H = \{0\}$ il risultato è banale, quindi possiamo assumere $H \neq \{0\}$. Per ipotesi, siccome T è una perturbazione compatta dell'identità, l'operatore $K = I - T$ è compatto.

(i). Sappiamo già che $\operatorname{Im} T$ è un sottospazio vettoriale di H (☞ proposizione 1.20(viii)); ciò che va verificato è che è un sottoinsieme chiuso di H.

Osserviamo che, siccome $\operatorname{Ker} T$ è chiuso in H (☞ proposizione 7.13(i)), dalla decomposizione ortogonale

$$H = \operatorname{Ker} T \oplus (\operatorname{Ker} T)^{\perp}$$

(☞ corollario 6.33(i)) deduciamo che

$$\operatorname{Im} T = T(H) = T(\operatorname{Ker} T \oplus (\operatorname{Ker} T)^{\perp}) = T((\operatorname{Ker} T)^{\perp}), \tag{11.8}$$

perché $T(\operatorname{Ker} T) = \{0\}$. Vogliamo ora dimostrare che

$$\delta := \inf\{\|Tx\|_H : x \in (\operatorname{Ker} T)^{\perp},\ \|x\|_H = 1\} > 0. \tag{11.9}$$

Supponiamo per assurdo che $\delta = 0$. Allora possiamo trovare una successione $(x_n)_n$ tale che

$$x_n \in (\operatorname{Ker} T)^{\perp}, \qquad \|x_n\|_H = 1 \quad \text{per ogni } n \in \mathbb{N} \tag{11.10}$$

e che

$$Tx_n \xrightarrow[n\to\infty]{} 0 \quad \text{in } H. \tag{11.11}$$

Siccome $(x_n)_n$ è una successione limitata in H, e $K = I - T$ è un operatore compatto, a meno di sostituire $(x_n)_n$ con una sottosuccessione, possiamo anche assumere che

$$Kx_n \xrightarrow[n\to\infty]{} y \quad \text{in } H$$

per qualche $y \in H$. Ma allora, siccome $I = T + K$,

$$x_n = Tx_n + Kx_n \xrightarrow[n\to\infty]{} 0 + y = y \quad \text{in } H, \tag{11.12}$$

cioè $y = \lim_{n\to\infty} x_n$; dunque, per (11.11) e la continuità di T,

$$Ty = \lim_{n\to\infty} Tx_n = 0,$$

cioè $y \in \operatorname{Ker} T$. D'altra parte, da (11.10) e (11.12) segue che

$$y \in (\operatorname{Ker} T)^{\perp}, \qquad \|y\|_H = 1,$$

dato $(\operatorname{Ker} T)^{\perp}$ è chiuso e la norma è continua. Abbiamo allora $y \in \operatorname{Ker} T \cap (\operatorname{Ker} T)^{\perp} = \{0\}$, cioè $y = 0$, ma questo è assurdo dato che $\|y\|_H = 1$.

Da (11.9), riscalando, si deduce facilmente che

$$\|Tx\|_H \geq \delta \|x\|_H \qquad \forall x \in (\operatorname{Ker} T)^{\perp}.$$

In altre parole, l'operatore $T|_{(\operatorname{Ker} T)^{\perp}} : (\operatorname{Ker} T)^{\perp} \to H$ è coercivo in norma; inoltre $(\operatorname{Ker} T)^{\perp}$ è chiuso in H, dunque, dotato della norma indotta da H è uno spazio di Banach (☞ proposizione 4.16). Possiamo allora applicare il lemma 7.40 all'operatore $T|_{(\operatorname{Ker} T)^{\perp}}$, deducendone che l'immagine è chiusa. Siccome

$$\operatorname{Im}(T|_{(\operatorname{Ker} T)^{\perp}}) = T((\operatorname{Ker} T)^{\perp}) = \operatorname{Im} T$$

per (11.8), questo dimostra che $\operatorname{Im} T$ è chiusa. Di conseguenza, la decomposizione (11.7) segue dalla proposizione 10.5(vi).

(ii). Caso 1: H ha dimensione finita. Da (11.7) segue che

$$\dim H = \dim \operatorname{Ker}(T^*) + \dim \operatorname{Im} T;$$

per confronto con la relazione già nota

$$\dim H = \dim \operatorname{Ker} T + \dim \operatorname{Im} T$$

(☞ proposizione 1.20(xi)), ne deduciamo che $\dim \operatorname{Ker} T = \dim \operatorname{Ker}(T^*)$, come desiderato.

Caso 2: K ha rango finito. Siccome $K = I - T$, per ogni $x \in \operatorname{Ker} T$,

$$Kx = x - Tx = x,$$

pertanto $x = Kx \in \operatorname{Im} K$. Questo dimostra che $\operatorname{Ker} T \subseteq \operatorname{Im} K$. Allo stesso modo, siccome $K^* = I - T^*$, si deduce che $\operatorname{Ker}(T^*) \subseteq \operatorname{Im}(K^*)$.

Ora, siccome K ha rango finito, anche K^* ha rango finito (☞ proposizione 11.21(i)). Dunque, se poniamo $V = \operatorname{Im} K + \operatorname{Im}(K^*) = \operatorname{span}(\operatorname{Im} K \cup \operatorname{Im} K^*)$, si ha $\dim V < \infty$. Inoltre, si ha banalmente

$$K(V) \subseteq \operatorname{Im} K \subseteq V, \qquad K^*(V) \subseteq \operatorname{Im}(K^*) \subseteq V,$$

cioè V è invariante per K e K^*, ma allora è anche invariante per $T = I - K$ e $T^* = I - K^*$. Infine, sappiamo che

$$\operatorname{Ker} T \subseteq \operatorname{Im} K \subseteq V, \qquad \operatorname{Ker}(T^*) \subseteq \operatorname{Im}(K^*) \subseteq V,$$

da cui segue che

$$\operatorname{Ker} T = \operatorname{Ker}(T|_V), \qquad \operatorname{Ker}(T^*) = \operatorname{Ker}(T^*|_V) = \operatorname{Ker}((T|_V)^*),$$

dove abbiamo usato la proposizione 10.5(vii). A meno di sostituire T con $T|_V$, possiamo dunque assumere che H abbia dimensione finita, e ricondurci al caso già trattato sopra.

Caso 3: K è un arbitrario operatore compatto. Per il teorema 11.17, esiste $K_0 \in \mathcal{B}(H)$ di rango finito tale che $\|K - K_0\|_{\mathrm{op}} < 1$. Allora

$$T = I - K = S - K_0,$$

dove $S = I - (K - K_0)$ è invertibile in $\mathcal{B}(H)$ per il criterio di Neumann (☞ teorema 7.42); pertanto si ha anche

$$T = ST_0, \qquad \text{dove} \quad T_0 = I - S^{-1}K_0;$$

inoltre $S^{-1}K_0$ ha rango finito, dato che S è un isomorfismo e

$$\dim \operatorname{Im}(S^{-1}K_0) = \dim \operatorname{Im} K_0 < \infty.$$

Applicando il caso precedente a $T_0 = I - S^{-1}K_0$, deduciamo che

$$\dim \operatorname{Ker} T_0 = \dim \operatorname{Ker}(T_0^*) < \infty.$$

Siccome S è un isomorfismo, dalla relazione $T = ST_0$ deduciamo che $\operatorname{Ker} T = \operatorname{Ker} T_0$ e

$$\dim \operatorname{Ker} T = \dim \operatorname{Ker}(T_0).$$

D'altra parte, $T^* = T_0^* S^*$, e anche S^* è un isomorfismo (☞ proposizione 10.5(v)), da cui segue che $\operatorname{Ker} T^* = (S^*)^{-1}(\operatorname{Ker}(T_0^*))$ e

$$\dim \operatorname{Ker}(T^*) = \dim \operatorname{Ker}(T_0^*).$$

Combinando le precedenti identità si deduce quanto desiderato. □

Corollario 11.25. *Sia $T \in \mathcal{B}(H)$ una perturbazione compatta dell'identità. Allora sono equivalenti:*

(i) T è iniettivo;

(ii) T è suriettivo;

(iii) T è biiettivo;

(iv) T è un isomorfismo.

Dimostrazione. Sappiamo già che

$$\text{(iv)} \iff \text{(iii)} \iff [\ \text{(i)}\ \&\ \text{(ii)}\ ],$$

dove la prima equivalenza è dovuta al teorema dell'isomorfismo (☞ teorema 7.36), mentre la seconda vale per definizione. Rimane solo da dimostrare l'equivalenza

$$\text{(i)} \iff \text{(ii)}.$$

D'altra parte, grazie alla decomposizione (11.7), si ha che T è suriettiva se e solo se $\operatorname{Ker}(T^*) = \{0\}$. Inoltre, per il teorema 11.24(ii),

$$\operatorname{Ker}(T^*) = \{0\} \iff \dim \operatorname{Ker}(T^*) = 0 \iff \dim \operatorname{Ker} T = 0 \iff \operatorname{Ker} T = \{0\},$$

per cui $\operatorname{Ker}(T^*) = \{0\}$ se e solo se T è iniettiva. □

Corollario 11.26 (alternativa di Fredholm)**.** *Sia $T \in \mathcal{B}(H)$ una perturbazione compatta dell'identità. Allora, si può verificare uno e uno solo dei casi seguenti:*

(1) l'equazione $Tx = y$ ha un'unica soluzione $x \in H$ per ogni dato $y \in H$ (in altre parole, $T : H \to H$ è biiettivo);

(2) l'equazione $Tx = 0$ ha almeno una soluzione non nulla $x \in H$ (in altre parole, $T : H \to H$ non è iniettivo).

Osservazione 11.27. Nel caso (2) del corollario 11.26, siccome T non è iniettivo, non è nemmeno suriettivo (☞ corollario 11.25), cioè esistono anche $y \in H$ per cui l'equazione

$$Tx = y$$

non ha soluzioni $x \in H$. D'altra parte, dalla decomposizione ortogonale (11.7), si deduce che l'equazione ha soluzioni $x \in H$ ogniqualvolta

$$y \perp \operatorname{Ker}(T^*).$$

11.3 Spettro di operatori compatti

Nel corso di questa sezione, $H \neq \{0\}$ è uno spazio di Hilbert su $\mathbb{F}$. Scriviamo per brevità I anziché id_H.

Nel seguente enunciato discutiamo alcune proprietà fondamentali dello spettro di operatori compatti.

Proposizione 11.28. *Sia $T \in \mathcal{K}(H)$.*

(i) Se $\dim H = \infty$, allora $0 \in \sigma(T)$.

(ii) $0 < \dim \operatorname{Ker}(T - \lambda I) < \infty$ per ogni $\lambda \in \sigma(T) \setminus \{0\}$.

(iii) $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$.

(iv) Per ogni $t > 0$, $\sharp\{\lambda \in \sigma(T) : |\lambda| \geq t\} < \infty$.

(v) $\sigma(T)$ è finito o numerabile, e nel secondo caso 0 è l'unico punto di accumulazione di $\sigma(T)$ in $\mathbb{F}$.

Dimostrazione. (i). Se $\dim H = 0$, l'operatore compatto T non può essere un isomorfismo (☞ proposizione 11.10), dunque $0 \in \sigma(T)$ (☞ proposizione 10.23(ii)).

(ii) e (iii). Sia $\lambda \in \sigma(T) \setminus \{0\}$. Per definizione di spettro, $T - \lambda I$ non è un isomorfismo, ma allora (siccome $\lambda \neq 0$), anche

$$I - \lambda^{-1}T = -\lambda^{-1}(T - \lambda I)$$

non è un isomorfismo (☞ osservazione 7.33(3)). Siccome $I - \lambda^{-1}T$ è una perturbazione compatta dell'identità, dal corollario 11.25 segue che $I - \lambda^{-1}T$ non è iniettivo, dunque $\operatorname{Ker}(T - \lambda I) = \operatorname{Ker}(I - \lambda^{-1}T) \neq \{0\}$, e inoltre tale nucleo ha dimensione finita (☞ teorema 11.24). In particolare, siccome $\operatorname{Ker}(T - \lambda I) \neq \{0\}$, si ha $\lambda \in \sigma_p(T)$.

(iv). Supponiamo per assurdo che esista $t > 0$ tale che

$$\sharp\{\lambda \in \sigma(T) : |\lambda| \geq t\} = \infty.$$

Possiamo allora trovare una successione $(\lambda_n)_{n\in\mathbb{N}}$ di elementi distinti dello spettro $\sigma(T)$ con $|\lambda_n| \geq t$ per ogni $n \in \mathbb{N}$. Poniamo, per ogni $N \in \mathbb{N}$,

$$V_N = \sum_{n=0}^{N-1} \operatorname{Ker}(T - \lambda_n I);$$

in particolare $V_0 = \{0\}$. Per (ii) e la proposizione 10.26(i), i V_N formano una successione strettamente crescente di sottospazi vettoriali di H di dimensione finita. Inoltre, in quanto somme di autospazi, i V_N sono invarianti per T (☞ proposizione 10.26), e più precisamente

$$(T - \lambda_N I)(V_{N+1}) \subseteq (T - \lambda_N I)(V_N) + (T - \lambda_N I)(\operatorname{Ker}(T - \lambda_N I)) \subseteq V_N, \quad (11.13)$$

dove si è usato che $(T - \lambda_N I)(\operatorname{Ker}(T - \lambda_N I)) = \{0\}$.

Dato che V_N è un sottospazio proprio di V_{N+1}, il complemento ortogonale $V_N^\perp \cap V_{N+1}$ di V_N in V_{N+1} non è banale, dunque esiste

$$y_N \in V_{N+1} \cap V_N^\perp \qquad \text{con} \qquad \|y_N\| = 1$$

per ogni $N \in \mathbb{N}$. Inoltre, siccome $y_N \in V_N^\perp$, mentre $(T - \lambda_N I) y_N \in V_N$ per (11.13), la decomposizione

$$T y_N = (T - \lambda_N I) y_N + \lambda_N y_N$$

è la decomposizione ortogonale di $T y_N$ rispetto al sottospazio V_N e al suo ortogonale (☞ teorema 6.32), ma allora

$$P_{V_N^\perp} T y_N = \lambda_N y_N$$

(☞ corollario 6.33(iii)).

Pertanto, per ogni $N > M$, siccome $T y_M \in T(V_{M+1}) \subseteq V_{M+1} \subseteq V_N$, si ha

$$\|T y_N - T y_M\| \geq d(T y_N, V_N) = \|P_{V_N^\perp} T y_N\| = |\lambda_N| \geq t$$

(☞ corollario 6.33(v)). Dato che $t > 0$ è indipendente da N, M, la disugaglianza sopra mostra che la successione $(T y_N)_N$ non ha sottosuccessioni di Cauchy, mentre $(y_N)_N$ è limitata; questo contraddice la compattezza di T, assurdo.

(v). Chiaramente

$$\sigma(T) \subseteq \{0\} \cup \bigcup_{n \in \mathbb{N}} \{\lambda \in \sigma(T) : |\lambda| \geq 2^{-n}\};$$

applicando (iv), ne deduciamo che $\sigma(T)$ è contenuto in un'unione numerabile di insiemi finiti, dunque $\sigma(T)$ è al più numerabile.

Sia ora $\mu \in \mathbb{F} \setminus \{0\}$. Allora $\mathbb{F} \setminus B(0, |\mu|/2)$ è un intorno di μ in $\mathbb{F}$, che per (iv) contiene al più un numero finito di punti di $\sigma(T)$; pertanto μ non è un punto di accumulazione di $\sigma(T)$ in $\mathbb{F}$. Dunque l'unico possibile punto di accumulazione di $\sigma(T)$ è 0. □

Esempio 11.29. Sia $T_K \in \mathcal{B}(L^2(0,1))$ l'operatore integrale corrisponendente al nucleo integrale $K \in L^2((0,1) \times (0,1))$ dato da

$$K(x,y) = e^{x-y} \qquad \forall x, y \in (0,1).$$

Sappiamo (☞ esempio 11.15(2)) che $T \in \mathcal{K}(L^2(0,1))$. Cerchiamo ora lo spettro $\sigma(T_K)$. Siccome T_K è compatto e $L^2(0,1)$ ha dimensione infinita, sappiamo che $\sigma(T_K) = \{0\} \cup \sigma_p(T_K)$ (☞ proposizione 11.28), quindi basta determinare $\sigma_p(T_K)$. Inoltre, per definizione, si ha che $\lambda \in \sigma_p(T_K)$ se e solo se esiste $f \in L^2(0,1) \setminus \{0\}$ tale che $T_K f = \lambda f$.

Notiamo ora che $K(x,y) = \phi(x)\psi(y)$, dove $\phi(x) = e^x$, $\psi(y) = e^{-y}$. Pertanto

$$T_K f(x) = \int_0^1 K(x,y)\, f(y)\, dy = \phi(x) \int_0^1 f(y)\, \overline{\psi(y)}\, dy,$$

ove si è usato che ψ ha valori reali; in altre parole,

$$T_K f = \langle f, \psi \rangle \phi \qquad \forall f \in L^2(0,1). \tag{11.14}$$

In particolare, siccome $\phi \neq 0$ in $L^2(0,1)$,

$$T_K f = 0 \iff \langle f, \psi \rangle = 0 \iff f \in \psi^\perp,$$

cioè $\operatorname{Ker} T_K = \psi^\perp \neq \{0\}$ è l'autospazio di autovalore 0.

D'altra parte, sempre da (11.14), si vede che $\operatorname{Im} T_K \subseteq \operatorname{span}\{\phi\}$. Ora, se f è un autovettore di autovalore $\lambda \neq 0$, si ha $T_K f = \lambda f$, dunque $f = \lambda^{-1} T_K f \in \operatorname{Im} T_K$, pertanto un tale autovettore dev'essere multiplo di ϕ. D'altra parte,

$$T_K \phi = \langle \phi, \psi \rangle \phi,$$

quindi ϕ è effettivamente un autovettore di autovalore

$$\langle \phi, \psi \rangle = \int_0^1 e^x e^{-x} \, dx = 1.$$

Ne concludiamo che $\sigma(T_K) = \sigma_p(T_K) = \{0, 1\}$, con autospazi $\operatorname{Ker} T_K = \psi^\perp$ e $\operatorname{Ker}(T_K - I) = \operatorname{span}\{\phi\}$. Notiamo che ϕ e ψ non sono multipli l'uno dell'altro, dunque $P_{\psi^\perp} \phi \neq 0$; pertanto gli autospazi $\operatorname{Ker} T_K$ e $\operatorname{Ker}(T_K - I)$ non sono ortogonali. Questo indica che T_K non è un operatore normale (☞ proposizione 10.26(iii)).

11.4 Il teorema spettrale

Nel corso di questa sezione, $H \neq \{0\}$ è uno spazio di Hilbert su $\mathbb{F}$. Scriviamo per brevità I anziché id_H.

Proposizione 11.30. *Sia $T \in \mathcal{B}(H)$.*

(i) Se T è autoaggiunto, allora $\|T\|_{\mathrm{op}} \in \sigma(T)$ oppure $-\|T\|_{\mathrm{op}} \in \sigma(T)$.

(ii) Supponiamo che $\mathbb{F} = \mathbb{C}$ e che T sia compatto e normale. Allora esiste $\mu \in \sigma(T)$ con $|\mu| = \|T\|_{\mathrm{op}}$.

Dimostrazione. (i). Se $\|T\|_{\mathrm{op}} = 0$, allora $T = 0$ e dunque $\sigma(T) = \{0\} \ni \|T\|_{\mathrm{op}}$. Supponiamo $\|T\|_{\mathrm{op}} \neq 0$. Per il teorema della mappa spettrale (☞ proposizione 10.23(iv)),

$$\sigma(T) = \|T\|_{\mathrm{op}} \, \sigma(T/\|T\|_{\mathrm{op}}),$$

pertanto, a meno di rimpiazzare T con $T/\|T\|_{\mathrm{op}}$, possiamo supporre $\|T\|_{\mathrm{op}} = 1$.

Siccome $1 = \|T\|_{\mathrm{op}} = \sup\{\|Tx\| : x \in X, \ \|x\| = 1\}$ (☞ lemma 7.5), possiamo trovare una successione $(x_n)_n$ in H con

$$\|x_n\|_H = 1 \quad \forall n \in \mathbb{N}, \qquad \lim_{n \to \infty} \|T x_n\| = 1.$$

In particolare,

$$\begin{aligned} \|(T^2 - I) x_n\|^2 = \|T^2 x_n - x_n\|^2 &= \|T^2 x_n\|^2 + \|x_n\|^2 - 2 \,\mathfrak{Re} \langle x_n, T^2 x_n \rangle \\ &\leq \|T\|_{\mathrm{op}}^2 \|x_n\|^2 + \|x_n\|^2 - 2\|T x_n\|^2 \leq 2 - 2\|T x_n\|^2 \to 0 \end{aligned}$$

per $n \to \infty$. Dunque $(T^2 - I) x_n \to 0$, mentre $\|x_n\| = 1$ per ogni n, pertanto $T^2 - I$ non è un isomorfismo (non può avere inverso limitato). Siccome

$$T^2 - I = (T - I)(T + I),$$

uno fra $T - I$ e $T + I$ non può essere un isomorfismo (☞ osservazione 7.33(3)), quindi $1 \in \sigma(T)$ oppure $-1 \in \sigma(T)$.

(ii). Come sopra, il caso $\|T\|_{\text{op}} = 0$ è banale, dunque possiamo assumere $T \neq 0$. Siccome T è compatto, anche T^*T è compatto (☞ proposizione 11.9), e inoltre è autoaggiunto (☞ osservazione 10.12). In aggiunta, $\|T^*T\|_{\text{op}} = \|T\|_{\text{op}}^2$ (☞ proposizione 10.5(iv)), dunque $T^*T \neq 0$ perché $T \neq 0$. Applicando (i) all'operatore T^*T, otteniamo che esiste λ con $|\lambda| = \|T\|_{\text{op}}^2$ tale che $\lambda \in \sigma(T^*T)$. Siccome T^*T è compatto e $\lambda \neq 0$, dev'essere $\lambda \in \sigma_p(T^*T)$, e $V \coloneqq \operatorname{Ker}(T^*T - \lambda I)$ ha dimensione finita e non nulla (☞ proposizione 11.28)

Siccome T è normale, T commuta con T^*, dunque anche con T^*T. In particolare, $V = \operatorname{Ker}(T^*T - \lambda I)$ è invariante per T (☞ proposizione 10.26). Allora l'operatore $T|_V \in \mathcal{B}(V)$ ha spettro non vuoto (☞ proposizione 10.22(ii)), cioè ha almeno un autovalore $\mu \in \sigma_p(T|_V)$, dato che V ha dimensione finita. Se $x \in V \setminus \{0\}$ è il corrispondente autovettore, si ha $Tx = \mu x$, ma anche $T^*x = \overline{\mu}x$ (dato che T è normale, ☞ proposizione 10.17), ma allora $T^*Tx = |\mu|^2 x$. Siccome $x \in V = \operatorname{Ker}(T^*T - \lambda I)$, si ha d'altra parte $T^*Tx = \lambda x$, dunque $\lambda = |\mu|^2$, e in particolare $|\mu| = \sqrt{|\lambda|} = \|T\|_{\text{op}}$. □

Proposizione 11.31. *Sia $T \in \mathcal{K}(H)$ normale. Se $\mathbb{F} = \mathbb{R}$, assumiamo anche che T sia autoaggiunto. Allora*

$$\operatorname{Ker} T = \left(\bigcup_{\lambda \in \sigma(T) \setminus \{0\}} \operatorname{Ker}(T - \lambda I) \right)^{\perp}.$$

Dimostrazione. Poniamo

$$V = \left(\bigcup_{\lambda \in \sigma(T) \setminus \{0\}} \operatorname{Ker}(T - \lambda I) \right)^{\perp} = \bigcap_{\lambda \in \sigma(T) \setminus \{0\}} \operatorname{Ker}(T - \lambda I)^{\perp}.$$

Allora V è un sottospazio vettoriale chiuso di H, in quanto intersezione di sottospazi chiusi. Inoltre, per ogni λ, siccome T e T^* commutano (T è normale) $\operatorname{Ker}(T - \lambda I)$ è invariante sia per T che per T^*, dunque lo stesso vale per $\operatorname{Ker}(T - \lambda I)^{\perp}$ (☞ proposizione 10.26); pertanto anche l'intersezione V è invariante sia per T che per T^*.

Pensiamo V come spazio di Hilbert, con il prodotto scalare indotto da H. Siccome $T^*|_V = (T|_V)^*$ (☞ proposizione 10.5(vii)), anche $T|_V$ è un operatore compatto e normale, ed è autoaggiunto se $\mathbb{F} = \mathbb{R}$.

Supponiamo per assurdo che $T|_V \neq 0$. Allora, per la proposizione 11.30, esiste $\lambda \in \sigma(T|_V) \setminus \{0\}$, e siccome $T|_V$ è compatto, λ dev'essere un autovalore (☞ proposizione 11.28). Dunque deve esistere un autovettore $x \in V \setminus \{0\}$, cioè $Tx = \lambda x$. Ma allora $0 \neq x \in V \cap \operatorname{Ker}(T - \lambda I)$, il che è assurdo perché per costruzione $V \perp \operatorname{Ker}(T - \lambda I)$ e quindi $V \cap \operatorname{Ker}(T - \lambda I) = \{0\}$. Dunque dev'essere $T|_V = 0$, cioè $V \subseteq \operatorname{Ker} T$.

D'altra parte, per ogni $\lambda \neq 0$, siccome T è normale, si ha $\operatorname{Ker} T \perp \operatorname{Ker}(T - \lambda I)$ (☞ proposizione 10.26(iii)), quindi $\operatorname{Ker} T \subseteq \operatorname{Ker}(T - \lambda I)^{\perp}$, ma allora anche $\operatorname{Ker} T \subseteq V$. □

Teorema 11.32 (teorema spettrale)**.** *Sia $T \in \mathcal{K}(H)$ normale. Se $\mathbb{F} = \mathbb{R}$, assumiamo anche che T sia autoaggiunto. Allora:*

(i) $\sigma(T)$ è finito o numerabile, e nel secondo caso 0 è l'unico punto di accumulazione di $\sigma(T)$.

(ii) *Per ogni* $\lambda \in \sigma(T) \setminus \{0\}$*, se* $E_T(\lambda) := \operatorname{Ker}(T - \lambda I)$ *è l'autospazio di* T *di autovalore* λ*, si ha* $n_\lambda := \dim E_T(\lambda) < \infty$*.*

Sia inoltre $B_\lambda = \{e_1^\lambda, \dots, e_{n_\lambda}^\lambda\}$ *una base ortonormale di* $E_T(\lambda)$ *per ogni* $\lambda \in \sigma(T) \setminus \{0\}$*. Allora:*

(iii) $B := \bigcup_{\lambda \in \sigma(T)\setminus\{0\}} B_\lambda$ *è una base ortonormale di* $(\operatorname{Ker} T)^\perp = \overline{\operatorname{Im} T}$*.*

(iv) *Per ogni* $x \in H$*,*

$$Tx = \sum_{\lambda \in \sigma(T)\setminus\{0\}} \lambda \sum_{j=1}^{n_\lambda} \langle x, e_j^\lambda \rangle e_j^\lambda = \sum_{\lambda \in \sigma(T)\setminus\{0\}} \lambda P_{E_T(\lambda)} x, \tag{11.15}$$

con convergenza in H*.*

Osservazione 11.33. Le somme in (11.15) convergono incondizionatamente in H (☞ osservazione 6.45), cioè il loro valore non dipende dall'ordine di somma e in particolare non dipende da come si indicizza $\sigma(T) \setminus \{0\}$.

Dimostrazione del teorema 11.32. (i) e (ii) sono già noti (☞ proposizione 11.28).

(iii). Siccome T è normale, si ha

$$(\operatorname{Ker} T)^\perp = (\operatorname{Ker}(T^*))^\perp = \overline{\operatorname{Im} T}$$

(☞ proposizioni 10.17(ii) e 10.5(vi)). Inoltre, per la proposizione 11.31,

$$(\operatorname{Ker} T)^\perp = \overline{\operatorname{span}\left(\bigcup_{\lambda \in \sigma(T)\setminus\{0\}} E_T(\lambda)\right)} = \overline{\operatorname{span}\left(\bigcup_{\lambda \in \sigma(T)\setminus\{0\}} B_\lambda\right)} = \overline{\operatorname{span} B}. \tag{11.16}$$

Notiamo ora che $B_\lambda \perp B_\mu$ se $\lambda \neq \mu$, dato che T è normale (☞ proposizione 10.26(iii)); siccome ciascun B_λ è un insieme ortonormale, anche la loro unione B lo è. Da (11.16) deduciamo allora che B è un insieme ortonormale completo in $(\operatorname{Ker} T)^\perp$, cioè una sua base ortonormale (B è al più numerabile, in quanto unione al più numerabile di insiemi finiti).

(iv). Sia $x \in H$. Decomponiamo $x = y + z$, dove $y \in \operatorname{Ker} T$, $z \in (\operatorname{Ker} T)^\perp$. Allora, per ogni $\lambda \in \sigma(T) \setminus \{0\}$ e $j = 1, \dots, n_\lambda$,

$$\langle x, e_j^\lambda \rangle = \langle z, e_j^\lambda \rangle,$$

dato che $y \in \operatorname{Ker} T$ e $e_j^\lambda \in (\operatorname{Ker} T)^\perp$. Sviluppando z rispetto alla base ortonormale B di $(\operatorname{Ker} T)^\perp$, si ha allora

$$z = \sum_{\lambda \in \sigma(T)\setminus\{0\}} \sum_{j=1}^{n_\lambda} \langle x, e_j^\lambda \rangle e_j^\lambda$$

con convergenza in H (☞ proposizione 6.44(iii)). Di conseguenza, siccome T è continuo e $Ty = 0$,

$$Tx = Tz = \sum_{\lambda \in \sigma(T)\setminus\{0\}} \sum_{j=1}^{n_\lambda} \langle x, e_j^\lambda \rangle T e_j^\lambda = \sum_{\lambda \in \sigma(T)\setminus\{0\}} \lambda \sum_{j=1}^{n_\lambda} \langle x, e_j^\lambda \rangle e_j^\lambda,$$

con convergenza in H. Per concludere, osserviamo che (☞ proposizione 6.39)

$$P_{E_T(\lambda)}x = \sum_{j=1}^{n_\lambda} \langle x, e_j^\lambda \rangle e_j^\lambda,$$

dato che B_λ è una base ortonormale di $E_T(\lambda)$. □

Corollario 11.34. *Sia $T \in \mathcal{K}(H)$ normale. Se $\mathbb{F} = \mathbb{R}$, assumiamo anche che T sia autoaggiunto. Se H è separabile, allora H ha una base ortonormale fatta di autovettori di T.*

Dimostrazione. Se H è separabile, anche $\operatorname{Ker} T$ lo è, quindi (☞ teorema 6.46) esiste una base ortonormale B_0 di $\operatorname{Ker} T$. L'unione di B_0 e della base ortonormale B di $(\operatorname{Ker} T)^\perp$ data dal teorema spettrale (☞ teorema 11.32(iii)) dà la base ortonormale di H cercata. □

Osservazione 11.35. Nelle ipotesi del corollario 11.34, assumiamo inoltre che $\dim H = \infty$. Sia $\{v_n\}_{n\in\mathbb{N}}$ una base ortonormale di H costituita da autovettori di T, e sia $w_n \in \mathbb{F}$ l'autovalore corrispondente all'autovettore v_n per ogni $n \in \mathbb{N}$. Notiamo che $|w_n| \leq \|T\|_{\mathrm{op}}$ per ogni n (☞ proposizione 10.22(i)), dunque $\underline{w} = (w_n)_n \in \ell^\infty$.

Sia infine $\Xi : H \to \ell^2$ l'isomorfismo isometrico determinato dalla base ortonormale $\{v_n\}_n$, cioè

$$\Xi(x) = (\langle x, v_n \rangle)_{n\in\mathbb{N}} \qquad \forall x \in H$$

(☞ dimostrazione del corollario 6.48). Allora si ha

$$\Xi T v_n = w_n \Xi v_n = w_n \underline{e^{(n)}} = D_{\underline{w}} \underline{e^{(n)}} = D_{\underline{w}} \Xi v_n$$

per ogni $n \in \mathbb{N}$, dove $D_{\underline{w}}$ è l'operatore di moltiplicazione per $\underline{w}$ (☞ esempio 7.8(8)); siccome $\{v_n\}_n$ è una base ortonormale, ne deduciamo

$$\Xi T = D_{\underline{w}} \Xi,$$

cioè

$$\Xi T \Xi^{-1} = D_{\underline{w}}.$$

Dunque, siccome T è compatto, lo è anche $D_{\underline{w}}$, quindi $\underline{w} \in c_0$ e $\|\underline{w}\|_\infty = \|T\|_{\mathrm{op}}$.

Questa costruzione si può interpretare come la *diagonalizzazione* dell'operatore T: l'operatore $D_{\underline{w}}$ si può pensare come una "matrice diagonale con infinite righe e colonne" che agisce su "vettori con infinite componenti" in ℓ^2.

In conclusione, a meno di coniugazione con isomorfismi isometrici, ogni operatore compatto normale (autoaggiunto per $\mathbb{F} = \mathbb{R}$) su uno spazio di Hilbert separabile di dimensione infinita è della forma $D_{\underline{w}}$ per qualche $\underline{w} \in c_0$.