

APPUNTI PER IL CORSO DI METODI
MATEMATICI DELLA FISICA
INTRODUZIONE ALL'ANALISI FUNZIONALE

tenuto dal Prof. Marco Boiti
(versione 28 maggio, 2001)

a.a. 2000-2001

Indice

1	Spazi Metrici	5
1.1	Insiemi Aperti. Insiemi Chiusi. Intorni.	6
1.2	Convergenza. Successioni di Cauchy. Completezza.	8
1.3	Completamento di uno Spazio metrico	12
1.4	Problemi	15
2	Spazi Normati. Spazi di Banach	17
2.1	Spazio Vettoriale	17
2.2	Spazio Normato. Spazio di Banach	19
2.3	Ulteriori Proprietà degli Spazi Normati	20
2.4	Spazi Normati Finito Dimensionali e Sottospazi	23
2.5	Compattezza e Dimensioni Finite	26
2.6	Operatori Lineari	29
2.7	Spazi Lineari di Operatori	32
2.8	Operatori Lineari Limitati e Continui	33
2.9	Funzionali Lineari	37
2.10	Operatori Lineari e Funzionali su Spazi Finito Dimensionali	41
2.11	Spazi Normati di Operatori. Spazio Duale	44
2.12	Problemi	45
3	Spazi con Prodotto Scalare. Spazi di Hilbert	49
3.1	Breve Orientamento sul Principale Contenuto della Teoria	49
3.2	Spazi con Prodotto Scalare, Spazio di Hilbert	50
3.3	Ulteriori Proprietà degli Spazi con Prodotto Scalare	52
3.4	Definizione Equivalente di Spazio con Prodotto Scalare	54
3.5	Completamento di uno Spazio con Prodotto Scalare	55
3.6	Complemento Ortogonale e Somma Diretta	56
3.7	Insiemi e Successioni Ortonormali	62
3.8	Serie Collegate a Successioni e Insiemi Ortonormali	65
3.9	Basi Ortonormali	68
3.10	Rappresentazione di Funzionali su Spazi di Hilbert	72
3.11	Operatori Aggiunti di Hilbert	76
3.12	Operatori Autoaggiunti, Unitari e Normali	79
3.13	Problemi	83

4	Teoremi per gli Spazi Normati e di Banach	85
4.1	Breve Orientamento sui Contenuti Principali	85
4.2	Lemma di Zorn	85
4.3	Alcune Applicazioni del Lemma di Zorn	87
4.4	Teorema di Hahn–Banach	87
4.5	Estensioni del Teorema di Hahn–Banach	90
4.6	Operatore Duale	94
4.6.1	Relazione fra l’Operatore Duale T' e l’Operatore Aggiunto di Hilbert T^*	97
4.7	Spazi Riflessivi	99
4.8	Teorema della Categoria e della Uniforme Limitatezza	102
4.9	Convergenza Forte e Debole	105
4.9.1	Esempi	108
4.10	Convergenza di Successioni di Operatori e di Funzionali	109
4.11	Teorema dell’Applicazione Aperta	111
4.12	Operatori Lineari Chiusi. Teorema del Grafico Chiuso	114
4.13	Problemi	117
5	Teoria Spettrale degli Operatori Lineari	119
5.1	Teoria Spettrale in Spazi Normati Finito Dimensionali	119
5.2	Concetti Basilari	122
5.3	Proprietà Spettrali degli Operatori Lineari Limitati	124
5.4	Ulteriori Proprietà del Risolvente e dello Spettro	127
5.5	Uso dell’Analisi Complessa nella Teoria Spettrale	130
5.6	Problemi	134
6	Operatori Lineari Compatti	135
6.1	Breve orientamento sul principale contenuto del capitolo	135
6.2	Operatori Lineari Compatti su Spazi Normati	135
6.3	Ulteriori Proprietà degli Operatori Lineari Compatti	139
6.4	Proprietà Spettrali degli Operatori Lineari Compatti	144
6.5	Ulteriori Proprietà Spettrali	149
6.6	Equazioni Operatoriali	154
6.7	Ulteriori Teoremi di Tipo Fredholm	159
6.8	Alternativa di Fredholm	164

Capitolo 1

Spazi Metrici

Uno spazio metrico è un insieme X dotato di una *metrica*. La metrica associa ad ogni coppia di elementi (*punti*) di X una *distanza*. La metrica è definita assiomaticamente, gli assiomi essendo suggeriti da alcune proprietà semplici della distanza, così com'è familiarmente definita fra punti della retta reale $\mathbb{R}$ o del piano complesso $\mathbb{C}$. Si tratta come mostrano alcuni esempi basilari di un concetto molto generale. Un'importante proprietà aggiuntiva che uno spazio metrico può possedere è la *completezza*. Un altro concetto di interesse teorico e pratico è la *separabilità* di uno spazio metrico. Gli spazi metrici separabili sono più semplici di quelli non separabili.

1.1 DEFINIZIONE (SPAZIO METRICO, METRICA)

Uno *spazio metrico* è una coppia (X, d) , dove X è un insieme e d una *metrica su X* (o *distanza su X*), cioè una funzione definita su $X \times X$ tale che per ogni $x, y, z \in X$ si abbia

M1 d è a valori reali, finito e non negativo.

M2 $d(x, y) = 0$ se e solo se $x = y$.

M3 $d(x, y) = d(y, x)$ (Simmetria)

M4 $d(x, y) \leq d(x, z) + d(z, y)$ (Disuguaglianza Triangolare)

■

Alcuni termini di uso corrente sono i seguenti. X è normalmente chiamato l'*insieme sottostante* a (X, d) . I suoi elementi sono chiamati *punti*. Per x, y fissati il numero non negativo $d(x, y)$ si chiama *distanza* fra x e y . Le proprietà da (M1) a (M4) sono gli *assiomi della metrica*. Il nome “disuguaglianza triangolare” è preso a prestito dalla geometria elementare.

Un **sottoinsieme** $(Y, \tilde{d})$ di (X, d) si ottiene prendendo un sottoinsieme $Y \subset X$ e restringendo d a $Y \times Y$; allora la metrica su Y è data dalla restrizione

$$\tilde{d} = d|_{Y \times Y}.$$

$\tilde{d}$ si chiama la metrica **indotta** su Y da d .

1.1 Insiemi Aperti. Insiemi Chiusi. Intorni.

V'è un considerevole numero di concetti ausiliari che giocano un ruolo in connessione con gli spazi metrici. Quelli di cui avremo bisogno sono inclusi in questa sezione. Perciò questa sezione contiene molti concetti, ma il lettore noterà che molti di loro divengono familiari quando vengono applicati agli spazi euclidei.

Consideriamo dapprima alcuni importanti sottoinsiemi di un dato spazio metrico $X = (X, d)$.

1.2 DEFINIZIONE (PALLA E SFERA)

Dato un punto $x_0 \in X$ ed un numero reale $r > 0$ definiamo tre tipi di insiemi

$$(a) \quad B(x_0; r) = \{x \in X \mid d(x, x_0) < r\} \quad (\text{Palla Aperta})$$

$$(b) \quad \tilde{B}(x_0; r) = \{x \in X \mid d(x, x_0) \leq r\} \quad (\text{Palla Chiusa})$$

$$(c) \quad S(x_0; r) = \{x \in X \mid d(x, x_0) = r\} \quad (\text{Sfera})$$

■

In tutti e tre i casi x_0 è chiamato il *centro* ed r il *raggio*.

Attenzione. Lavorando con gli spazi metrici è assai utile utilizzare una terminologia analoga a quella della geometria Euclidea. Tuttavia bisogna essere coscienti di quanto sia pericoloso assumere che palle e sfere in uno spazio metrico arbitrario soddisfino alle medesime proprietà soddisfatte da palle e sfere in $\mathbb{R}^3$. Ad esempio una possibile proprietà inusuale è che una sfera può essere vuota. Un'altra possibile proprietà inusuale sarà citata più in là.

1.3 DEFINIZIONE (INSIEMI APERTI, INSIEMI CHIUSI)

Un sottoinsieme M di uno spazio metrico X è detto aperto se contiene una palla centrata in ciascuno dei suoi punti. Un sottoinsieme K è detto chiuso se il suo complemento (in X) è aperto, cioè se $K^C = X - K$ è aperto. ■

Il lettore verificherà facilmente che da questa definizione segue che una palla aperta è un insieme aperto e che una palla chiusa è un insieme chiuso.

Una palla aperta $B(x_0; \varepsilon)$ di raggio ε è spesso chiamata un ε -*intorno* di x_0 . Per un **intorno** di x_0 si intende un qualunque sottoinsieme di X che contiene un ε -intorno di x_0 .

Vediamo direttamente dalla definizione che ciascun intorno di x_0 contiene x_0 ; in altre parole x_0 è un punto di ciascuno dei suoi intorni. Se N è un intorno di x_0 e $N \subset M$, allora anche M è un intorno di x_0 .

Chiamiamo x_0 un **punto interno** di un insieme $M \subset X$ se M è un intorno di x_0 . L'**interno** di M è l'insieme di tutti i punti interni a M e può essere indicato con M^0 o con $\text{Int}(M)$. $\text{Int}(M)$ è aperto ed è l'insieme aperto più grande contenuto in M .

Non è difficile mostrare che la collezione di tutti i sottoinsiemi aperti di X , che possiamo chiamare $\mathcal{T}$, ha le seguenti proprietà

T1 $\emptyset \in \mathcal{T}$, $X \in \mathcal{T}$.

T2 L'unione di membri di $\mathcal{T}$ è un membro di $\mathcal{T}$.

T3 L'intersezione di un numero finito di membri di $\mathcal{T}$ è un membro di $\mathcal{T}$.

Dimostrazione. (T1) segue dall'osservazione che $\emptyset$ non ha elementi e, ovviamente, X è aperto. Proviamo (T2). Un punto qualunque x dell'unione U degli insiemi aperti appartiene ad (almeno) uno di questi insiemi, sia M , ed M contiene una palla B di x poiché M è aperto.

Allora $B \subset U$ per definizione di unione. Ciò prova (T2). Infine se y è un punto qualunque dell'intersezione degli insiemi aperti $M_1, \dots, M_n$ allora ciascun M_j contiene una palla di y e la più piccola di queste palle è contenuta nell'intersezione. Ciò prova (T3). ■

Osserviamo che le proprietà da (T1) a (T3) sono così fondamentali che vengono di norma inserite in un contesto più generale. Precisamente si definisce come uno **spazio topologico** $(X, \mathcal{T})$ un insieme X ed una collezione $\mathcal{T}$ che soddisfa gli *assiomi* da (T1) a (T3). L'insieme $\mathcal{T}$ è chiamato una *topologia per* X . Da questa definizione segue che

Uno spazio metrico è uno spazio topologico.

Gli insiemi aperti giocano anche un ruolo in connessione con le applicazioni continue, dove la continuità è una naturale generalizzazione della continuità conosciuta dall'analisi ed è definita come segue.

1.4 DEFINIZIONE (APPLICAZIONE CONTINUA)

Siano $X = (X, d)$ e $Y = (Y, \tilde{d})$ due spazi metrici. Un'applicazione $T : X \rightarrow Y$ è detta *continua nel punto* $x_0 \in X$ se per ogni $\varepsilon > 0$ v'è un $\delta > 0$ tale che

$$\tilde{d}(Tx, Tx_0) < \varepsilon \quad \text{per ogni } x \text{ che soddisfa a} \quad d(x, x_0) < \delta.$$

T è detta *continua* se è continua in ogni punto di X . ■

È importante ed interessante che le applicazioni continue possano essere caratterizzate in termini di insiemi aperti come segue.

1.5 Teorema (Applicazioni Continue)

Un'applicazione T di uno spazio metrico X in uno spazio metrico Y è continua se e solo se l'immagine inversa di un qualunque sottoinsieme aperto di Y è un sottoinsieme aperto di X . ■

Dimostrazione. (a) Supponiamo che T sia continua. Sia $S \subset Y$ aperto ed S_0 l'immagine inversa di S . Se $S_0 = \emptyset$ è aperto. Sia $S_0 \neq \emptyset$. Per un qualunque $x_0 \in S_0$ sia $y_0 = Tx_0$. Poiché S è aperto contiene un ε -intorno N di y_0 . Poiché T è continua x_0 ha un δ -intorno N_0 che è applicato in N . Poiché $N \subset S$ abbiamo che $N_0 \subset S_0$ così che S_0 è aperto perché $x_0 \in S_0$ era arbitrario.

(b) Viceversa assumiamo che l'immagine inversa di ogni insieme aperto in Y sia un insieme aperto in X . Allora per ogni $x_0 \in X$ e qualunque ε -intorno N di Tx_0 l'immagine inversa N_0 di N è aperta, poiché N è aperto, e N_0 contiene x_0 . Quindi anche N_0 contiene un δ -intorno di x_0 che è applicato in N poiché N_0 è applicato in N . Di conseguenza, per definizione, T è continua in x_0 . Poiché $x_0 \in X$ era arbitrario ne segue che T è continua. ■

Introduciamo ora due altri concetti che sono collegati. Sia M un sottoinsieme di uno spazio metrico X . Allora un punto x_0 di X (che può o può non essere un punto di M) è chiamato un **punto di accumulazione** di M (o *punto limite di* M) se ogni intorno di x_0 contiene almeno un punto $y \in M$ distinto da x_0 . L'insieme costituito dai punti di M e dai punti di accumulazione di M è chiamato la **chiusura di** M ed è indicato con

$$\overline{M}.$$

È il più piccolo insieme chiuso che contiene M .

Prima di procedere menzioniamo un'altra proprietà inusuale delle palle in uno spazio metrico. Mentre in $\mathbb{R}^3$ la chiusura $\overline{B}(x_0; r)$ di una palla aperta $B(x_0; r)$ è la palla chiusa $\tilde{B}(x_0; r)$, in un generico spazio metrico ciò può non essere valido.

Usando il concetto di chiusura vogliamo dare una definizione che risulterà di particolare importanza nel seguito.

1.6 DEFINIZIONE (INSIEME DENSO, SPAZIO SEPARABILE)

Un sottoinsieme M di uno spazio metrico X è detto *denso in X* se

$$\overline{M} = X.$$

X è detto *separabile* se ha un sottoinsieme numerabile che è denso in X . ■

Quindi se M è denso in X ogni palla in X , per quanto piccola, conterrà punti di M ; o, in altre parole, non c'è punto $x \in X$ che non abbia un intorno che non contiene punti di M .

Vedremo nel seguito che gli spazi metrici separabili sono alquanto più semplici di quelli non separabili.

1.2 Convergenza. Successioni di Cauchy. Completezza.

Sappiamo che le successioni di numeri reali giocano un ruolo importante in analisi ed è la metrica di $\mathbb{R}$ che permette di definire il concetto basilare di convergenza di una tale successione. Lo stesso vale per le successioni di numeri complessi; in questo caso dobbiamo usare la metrica del piano complesso. In uno spazio metrico arbitrario $X = (X, d)$ la situazione è assai simile, cioè possiamo considerare una successione (x_n) di elementi $x_1, x_2, \dots$ di X ed usare la metrica d per definire la convergenza in maniera analoga a quella dell'analisi.

1.7 DEFINIZIONE (CONVERGENZA DI UNA SUCCESSIONE, LIMITE)

Una successione (x_n) in uno spazio metrico $X = (X, d)$ è detta *convergere* od *essere convergente* se v'è un $x \in X$ tale che

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0.$$

x è chiamato il *limite* di (x_n) e si scrive

$$\lim_{n \rightarrow \infty} x_n = x$$

o semplicemente

$$x_n \rightarrow x.$$

Diciamo che (x_n) *converge a x* o *ammette il limite x* . Se (x_n) non è convergente si dice che è *divergente*. ■

Come è stata usata la metrica d in questa definizione? d ha fornito la successione di numeri reali $a_n = d(x_n, x)$ la cui convergenza definisce quella di (x_n) . Quindi se $x_n \rightarrow x$, dato un $\varepsilon > 0$, esiste un $N = N(\varepsilon)$ tale che tutti gli x_n con $n > N$ giacciono in un ε -intorno $B(x; \varepsilon)$ di x .

Per evitare incomprensioni osserviamo che il limite di una successione convergente deve essere un punto dello spazio X .

Mostriamo ora che due proprietà delle successioni convergenti (unicità del limite e limitatezza), che risultano familiari dall'analisi, si mantengono in questo contesto molto più generale.

Chiamiamo un sottoinsieme non vuoto $M \subset X$ un *insieme limitato* se il suo *diametro*

$$\delta(M) = \sup_{x, y \in M} d(x, y)$$

è finito. Chiamiamo una successione (x_n) in X una **successione limitata** se l'insieme dei suoi punti è un sottoinsieme limitato di X .

Ovviamente se M è limitato allora $M \subset B(x_0; r)$, dove $x_0 \in X$ è un qualunque punto ed r è un numero reale (sufficientemente grande) e viceversa.

La nostra asserzione è allora formulata come segue.

1.8 Lemma (Limitatezza, Limite)

Sia $X = (X, d)$ uno spazio metrico. Allora

(a) Una successione convergente in X è limitata ed il suo limite è unico.

(b) Se $x_n \rightarrow x$ e $y_n \rightarrow y$ in X , allora $d(x_n, y_n) \rightarrow d(x, y)$. ■

Dimostrazione. (a) Supponiamo che $x_n \rightarrow x$. Allora prendendo $\varepsilon = 1$ possiamo trovare un N tale che $d(x_n, x) < 1$ per tutti gli $n > N$. Quindi per tutti gli n abbiamo che $d(x_n, x) < 1 + a$ dove

$$a = \max\{d(x_1, x), \dots, d(x_N, x)\}.$$

Ciò mostra che (x_n) è limitata. Assumendo che $x_n \rightarrow x$ e che $x_n \rightarrow z$ abbiamo dalla (M4)

$$0 \leq d(x, z) \leq d(x, x_n) + d(x_n, z) \rightarrow 0 + 0$$

e l'unicità $x = z$ del limite segue dalla (M2).

(b) Dalla (M4) abbiamo che

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n).$$

Da cui otteniamo

$$d(x_n, y_n) - d(x, y) \leq d(x_n, x) + d(y_n, y)$$

ed una disuguaglianza simile scambiando x_n con x e y_n con y e moltiplicando per -1 . Assieme forniscono

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \rightarrow 0$$

per $n \rightarrow \infty$. ■

Definiremo ora il concetto di completezza di uno spazio metrico, che risulterà basilare nel seguito. La completezza *non* segue dagli assiomi (M1) sino a (M4), poiché vi sono spazi metrici *incompleti*. In altre parole, la completezza è una proprietà addizionale che gli spazi metrici possono avere o non avere. Essa ha varie conseguenze che rendono gli spazi metrici completi “migliori e più semplici” di quelli incompleti.

Ricordiamo dapprima dall'analisi che una successione (x_n) di numeri reali o complessi converge sulla retta reale $\mathbb{R}$ o nel piano complesso $\mathbb{C}$ se e solamente se soddisfa il *criterio di convergenza di Cauchy*, cioè se e solo se per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$|x_m - x_n| < \varepsilon \quad \text{per tutti gli } m, n > N.$$

Qui $|x_m - x_n|$ è la distanza $d(x_m, x_n)$ da x_m a x_n sulla retta reale $\mathbb{R}$ o nel piano complesso $\mathbb{C}$. Quindi possiamo scrivere la disuguaglianza del criterio di Cauchy nella forma

$$d(x_m, x_n) < \varepsilon \quad (m, n > N).$$

Se una successione (x_n) soddisfa alla condizione del criterio di Cauchy potremo chiamarla una *successione di Cauchy*. Allora il criterio di Cauchy dice semplicemente che una successione di numeri reali o complessi converge in $\mathbb{R}$ o $\mathbb{C}$ se e solamente se è una successione di Cauchy. Sfortunatamente in spazi più generali la situazione può essere più complicata e vi possono essere successioni di Cauchy che non convergono.

1.9 DEFINIZIONE (SUCCESIONE DI CAUCHY, COMPLETEZZA)

Una successione (x_n) in uno spazio metrico $X = (X, d)$ è detta di *Cauchy* (o *fondamentale*) se per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$d(x_m, x_n) < \varepsilon \quad \text{per ogni } m, n > N. \quad (1.1)$$

Lo spazio X è detto *completo* se ogni successione di Cauchy in X converge (cioè se ha un limite che è un elemento di X). ■

A prescindere dalla completezza o meno di X la condizione (1.1) è necessaria per la convergenza di una successione. Infatti si ottiene facilmente il risultato seguente.

1.10 Teorema (Successione Convergente)

Ogni successione convergente in uno spazio metrico è una successione di Cauchy. ■

Dimostrazione. Se $x_n \rightarrow x$ allora per ogni $\varepsilon > 0$ v'è un $N = N(\varepsilon)$ tale che

$$d(x_n, x) < \frac{\varepsilon}{2} \quad \text{per tutti gli } n > N.$$

Quindi dalla disuguaglianza triangolare otteniamo per $m, n > N$

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ciò mostra che (x_n) è di Cauchy. ■

Se lo spazio X è completo la condizione di Cauchy (1.1) diventa necessaria e sufficiente per la convergenza e si parla di criterio di Cauchy per la convergenza. Il teorema che afferma la validità del criterio di Cauchy in $\mathbb{R}$ e in $\mathbb{C}$ può essere riespresso in termini di completezza come segue.

1.11 Teorema (Retta Reale, Piano Complesso)

La retta reale ed il piano complesso sono spazi metrici completi. ■

Completiamo questa sezione con tre teoremi che sono legati alla convergenza e alla completezza e che saranno necessari nel seguito.

1.12 Teorema (Chiusura, Insieme Chiuso)

Sia M un sottoinsieme non vuoto di uno spazio metrico (X, d) e $\overline{M}$ la sua chiusura così come definita nella sezione precedente. Allora

- (a) $x \in \overline{M}$ se e solo se esiste una successione (x_n) in M tale che $x_n \rightarrow x$.
- (b) M è chiuso se e solo se ogni successione convergente (x_n) di punti di M converge ad un punto di M , ossia se e solo se per ogni successione di punti $x_n \in M$ tale che $x_n \rightarrow x$ è $x \in M$. ■

Dimostrazione. (a) Sia $x \in \overline{M}$. Se $x \in M$ una successione di questo tipo è $(x, x, \dots)$. Se $x \notin M$ è un punto di accumulazione di M . Quindi per ogni $n = 1, 2, \dots$ la palla $B(x; 1/n)$ contiene un $x_n \in M$ e $x_n \rightarrow x$ perché $1/n \rightarrow 0$ per $n \rightarrow \infty$.

Viceversa se (x_n) è in M e $x_n \rightarrow x$ allora se $x \in M$ non v'è nulla da dimostrare. Se $x \notin M$ ogni intorno di x contiene punti $x_n \neq x$, così che $x \in \overline{M}$ per definizione di chiusura.

(b) M è chiuso se e solo se $M = \overline{M}$ così che (b) segue facilmente da (a). ■

1.13 Teorema (Sottospazio Completo)

Un sottospazio M di uno spazio metrico completo X è esso stesso completo se e solo se l'insieme M è chiuso in X . ■

Dimostrazione. Sia M completo. Grazie a 1.12(a) per ogni $x \in \overline{M}$ v'è una successione (x_n) in M che converge in M , il limite essendo unico per l' 1.8. Quindi $x \in M$. Questo prova che M è chiuso perché $x \in \overline{M}$ era arbitrario.

Viceversa sia M chiuso e (x_n) di Cauchy in M . Allora $x_n \rightarrow x \in X$ ciò che implica $x \in \overline{M}$ per l'1.12(a) e $x \in M$ poiché $M = \overline{M}$ per assunzione. Quindi la successione arbitraria di Cauchy (x_n) converge in M ciò che prova la completezza di M . ■

Questo teorema è molto utile e ne avremo molto spesso bisogno nel seguito.

L'ultimo dei tre teoremi annunciati mostra l'importanza della convergenza delle successioni in connessione con la continuità di un'applicazione.

1.14 Teorema (Applicazione Continua)

Un'applicazione $T : X \rightarrow Y$ di uno spazio metrico (X, d) in uno spazio metrico $(Y, \tilde{d})$ è continua in un punto $x_0 \in X$ se e solo se

$$x_n \rightarrow x_0 \quad \text{implica} \quad Tx_n \rightarrow Tx_0. \quad \blacksquare$$

Dimostrazione. Si assuma che T sia continua in x_0 . Allora per un dato $\varepsilon > 0$ v'è un $\delta > 0$ tale che

$$d(x, x_0) < \delta \quad \text{implica} \quad \tilde{d}(Tx, Tx_0) < \varepsilon.$$

Sia $x_n \rightarrow x_0$. Allora v'è un N tale che per ogni $n > N$ sia

$$d(x_n, x_0) < \delta.$$

Quindi per ogni $n > N$

$$\tilde{d}(Tx_n, Tx_0) < \varepsilon.$$

Per definizione ciò significa che $Tx_n \rightarrow Tx_0$.

Viceversa assumiamo che

$$x_n \rightarrow x_0 \quad \text{implichi} \quad Tx_n \rightarrow Tx_0$$

e proviamo che allora T è continua in x_0 . Supponiamo che ciò sia falso. Allora v'è un $\varepsilon > 0$ tale che per ogni $\delta > 0$ v'è un $x \neq x_0$ che soddisfa a

$$d(x, x_0) < \delta \quad \text{ma tale che} \quad \tilde{d}(Tx, Tx_0) \geq \varepsilon.$$

In particolare per $\delta = 1/n$ v'è un x_n che soddisfa a

$$d(x_n, x_0) < \frac{1}{n} \quad \text{ma tale che} \quad \tilde{d}(Tx_n, Tx_0) \geq \varepsilon.$$

Chiaramente $x_n \rightarrow x_0$ ma (Tx_n) non converge a Tx_0 . Ciò contraddice $Tx_n \rightarrow Tx_0$ e prova il teorema. ■

In particolare da questo teorema e dal Lemma 1.8 al punto b) segue la seguente proposizione.

1.15 Proposizione (Continuità della distanza)

La distanza $d(x, y)$ in X è continua in x and in y . ■

1.3 Completamento di uno Spazio metrico

Sappiamo che la retta razionale $\mathbb{Q}$ non è completa ma può essere “allargata” alla retta reale $\mathbb{R}$ che è completa. Questo “completamento” $\mathbb{R}$ di $\mathbb{Q}$ è tale che $\mathbb{Q}$ è denso in $\mathbb{R}$. È molto importante che un arbitrario spazio metrico incompleto possa essere “completato” in una maniera simile. Per una formulazione precisa e conveniente useremo i due concetti seguenti collegati fra loro e che hanno anche diverse altre applicazioni.

1.16 DEFINIZIONE (APPLICAZIONE ISOMETRICA, SPAZI ISOMETRICI)

Siano $X = (X, d)$ e $\tilde{X} = (\tilde{X}, \tilde{d})$ spazi metrici. Allora

- (a) Un'applicazione T di X in $\tilde{X}$ è detta *isometrica* o una *isometria* se T conserva le distanze, cioè se per ogni $x, y \in X$

$$\tilde{d}(Tx, Ty) = d(x, y),$$

dove Tx e Ty sono le immagini di x e y , rispettivamente.

- (b) Lo spazio X è detto *isometrico* allo spazio $\tilde{X}$ se esiste un'isometria biiettiva di X su $\tilde{X}$. Gli spazi X e $\tilde{X}$ sono allora chiamati *spazi isometrici*. ■

Si noti che una isometria è sempre iniettiva.

Due spazi isometrici possono differire al più per la natura dei loro punti ma sono indistinguibili dal punto di vista della metrica. In uno studio, in cui la natura dei punti non abbia importanza, i due spazi si possono considerare identici — ovvero come due copie del medesimo spazio “astratto”.

Possiamo ora formulare e provare il teorema che asserisce che ogni spazio metrico può essere completato. Lo spazio $\hat{X}$ che appare in questo teorema è chiamato **completamento** dello spazio dato X .

1.17 Teorema (Completamento)

Per uno spazio metrico $X = (X, d)$ esiste uno spazio metrico completo $\hat{X} = (\hat{X}, \hat{d})$ che ha un sottospazio W che è isometrico a X e che è denso in $\hat{X}$. Questo spazio $\hat{X}$ è unico a meno di isometrie, cioè, se $\tilde{X}$ è un qualunque spazio metrico completo che ha un sottospazio denso $\tilde{W}$ isometrico a X , allora $\tilde{X}$ e $\hat{X}$ sono isometrici. ■

Dimostrazione. La dimostrazione è piuttosto lunga ma diretta. La suddividiamo in quattro passi da (a) a (d). Costruiamo

- (a) $\hat{X} = (\hat{X}, \hat{d})$
 (b) un'isometria T di X su W , dove $\overline{W} = \hat{X}$.

Poi proviamo

- (c) la completezza di $\hat{X}$,
 (d) l'unicità di $\hat{X}$ a meno di isometrie.

Parlando rozzamente possiamo dire che il nostro compito è quello di assegnare dei limiti convenienti a quelle successioni di Cauchy in X che non convergono. Tuttavia non dobbiamo introdurre “troppi” limiti, ma dobbiamo tener conto che certe successioni “alla fine divengono arbitrariamente vicine le une alle altre”. Questa idea intuitiva può essere espressa matematicamente in termini di una conveniente relazione di equivalenza [vedi (1.2) qui di seguito].

(a) *Costruzione di $\widehat{X} = (\widehat{X}, \widehat{d})$.* Siano (x_n) e (x'_n) successioni di Cauchy in X . Definiamo (x_n) *equivalente* a (x'_n) e scriviamo $(x_n) \sim (x'_n)$ se

$$\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0. \quad (1.2)$$

Sia $\widehat{X}$ l'insieme delle classi di equivalenza $\widehat{x}, \widehat{y}, \dots$ di successioni di Cauchy così ottenute. Scriviamo $(x_n) \in \widehat{x}$ per indicare che (x_n) è un membro di $\widehat{x}$ (un *rappresentante* della classe $\widehat{x}$). Poniamo ora

$$\widehat{d}(\widehat{x}, \widehat{y}) = \lim_{n \rightarrow \infty} d(x_n, y_n) \quad (1.3)$$

dove $(x_n) \in \widehat{x}$ e $(y_n) \in \widehat{y}$. Mostriamo che questo limite esiste. Abbiamo

$$d(x_n, y_n) \leq d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n);$$

quindi otteniamo

$$d(x_n, y_n) - d(x_m, y_m) \leq d(x_n, x_m) + d(y_m, y_n)$$

e una disuguaglianza simile con m ed n scambiati. Da entrambe

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_m, y_n). \quad (1.4)$$

Poiché (x_n) e (y_n) sono di Cauchy possiamo rendere il membro a destra piccolo a piacere. Ciò implica che il limite in (1.3) esiste perché $\mathbb{R}$ è completo.

Dobbiamo anche mostrare che il limite in (1.3) è indipendente dalla particolare scelta del rappresentante. Infatti se $(x_n) \sim (x'_n)$ e $(y_n) \sim (y'_n)$ allora per la (1.2)

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n) \rightarrow 0$$

per $n \rightarrow \infty$, che implica l'asserzione che

$$\lim_{n \rightarrow \infty} d(x_n, y_n) = \lim_{n \rightarrow \infty} d(x'_n, y'_n).$$

Proviamo che $\widehat{d}$ in (1.3) è una metrica in $\widehat{X}$. Ovviamente $\widehat{d}$ soddisfa (M1) così come $\widehat{d}(\widehat{x}, \widehat{x}) = 0$ e (M3). Inoltre

$$\widehat{d}(\widehat{x}, \widehat{y}) = 0 \implies (x_n) \sim (y_n) \implies \widehat{x} = \widehat{y}$$

fornisce (M2) e (M4) per $\widehat{d}$ segue da

$$d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n)$$

per $n \rightarrow \infty$.

(b) *Costruzione di un isometria $T : X \rightarrow W \subset \widehat{X}$.* A ciascun $b \in X$ associamo la classe $\widehat{b} \in \widehat{X}$ che contiene la successione costante di Cauchy $(b, b, \dots)$. Ciò definisce un'applicazione $T : X \rightarrow W$ sul sottospazio $W = T(X) \subset \widehat{X}$. L'applicazione T è data da $b \mapsto \widehat{b} = Tb$, dove $(b, b, \dots) \in \widehat{b}$. Vediamo che T è un'isometria perché (1.3) diviene semplicemente

$$\widehat{d}(\widehat{b}, \widehat{c}) = d(b, c);$$

qui $\hat{c}$ è la classe di (y_n) dove $y_n = c$ per tutti gli n . Una qualunque isometria è iniettiva e $T : X \rightarrow W$ è surgettiva perché $T(X) = W$. Quindi W e X sono isometrici.

Mostriamo che W è denso in $\hat{X}$. Consideriamo un qualunque $\hat{x} \in \hat{X}$. Sia $(x_n) \in \hat{x}$. Per ogni $\varepsilon > 0$ v'è un N tale che

$$d(x_n, x_N) < \frac{\varepsilon}{2} \quad (n > N).$$

Sia $(x_N, x_N, \dots) \in \hat{x}_N$. Allora $\hat{x}_N \in W$. Per la (1.3)

$$\hat{d}(\hat{x}, \hat{x}_N) = \lim_{n \rightarrow \infty} d(x_n, x_N) \leq \frac{\varepsilon}{2} < \varepsilon.$$

Ciò mostra che ogni ε -intorno dell'arbitrario $\hat{x} \in \hat{X}$ contiene un elemento di W . Quindi W è denso in $\hat{X}$.

(c) *Completezza di $\hat{X}$* . Sia $(\hat{x}_n)$ una qualunque successione di Cauchy in $\hat{X}$. Poiché W è denso in $\hat{X}$ per ogni $\hat{x}_n$ v'è un $\hat{z}_n \in W$ tale che

$$\hat{d}(\hat{x}_n, \hat{z}_n) < \frac{1}{n}. \quad 4$$

Quindi per la disuguaglianza triangolare

$$\begin{aligned} \hat{d}(\hat{z}_m, \hat{z}_n) &\leq \hat{d}(\hat{z}_m, \hat{x}_m) + \hat{d}(\hat{x}_m, \hat{x}_n) + \hat{d}(\hat{x}_n, \hat{z}_n) \\ &< \frac{1}{m} + \hat{d}(\hat{x}_m, \hat{x}_n) + \frac{1}{n} \end{aligned}$$

e ciò è minore di ogni dato $\varepsilon > 0$ per m ed n sufficientemente grandi perché $(\hat{x}_m)$ è di Cauchy. Quindi $(\hat{z}_m)$ è di Cauchy. Poiché $T : X \rightarrow W$ è isometrica e $\hat{z}_m \in W$ la successione (z_m) dove $z_m = T^{-1}\hat{z}_m$ è di Cauchy in X . Sia $\hat{x} \in \hat{X}$ la classe a cui (z_m) appartiene. Mostriamo che $\hat{x}$ è il limite di $(\hat{x}_n)$. Per la (4)

$$\begin{aligned} \hat{d}(\hat{x}_n, \hat{x}) &\leq \hat{d}(\hat{x}_n, \hat{z}_n) + \hat{d}(\hat{z}_n, \hat{x}) \\ &< \frac{1}{n} + \hat{d}(\hat{z}_n, \hat{x}). \end{aligned} \quad (1.5)$$

Poiché $(z_n) \in \hat{x}$ e $\hat{z}_n \in W$, così che $(z_n, z_n, \dots) \in \hat{z}_n$, grazie alla definizione di distanza in $\hat{X}$ data nella (1.3), la disuguaglianza (1.5) diviene

$$\hat{d}(\hat{x}_n, \hat{x}) < \frac{1}{n} + \lim_{m \rightarrow \infty} d(z_n, z_m)$$

ed il membro a destra diviene più piccolo di un qualunque dato $\varepsilon > 0$ per n sufficientemente grandi. Quindi la successione arbitraria di Cauchy $(\hat{x}_n)$ in $\hat{X}$ ha il limite $\hat{x} \in \hat{X}$ e $\hat{X}$ è completo.

(d) *Unicità di $\hat{X}$ a meno di isometrie*. Sia $(\tilde{X}, \tilde{d})$ un altro spazio metrico completo con un sottospazio $\tilde{W}$ denso in $\tilde{X}$ e isometrico a X . Grazie alla proprietà transitiva della isometria W e $\tilde{W}$ sono isometrici e si può utilizzare questa isometria per definire una applicazione biiettiva T di $\hat{X}$ in $\tilde{X}$. Precisamente per ogni $\hat{x} \in \hat{X}$ consideriamo una successione $(\hat{x}_n)$ in W tale che $\hat{x}_n \rightarrow \hat{x}$. Se $(\tilde{x}_n)$ è la successione corrispondente nella isometria in $\tilde{W}$ vi sarà un $\tilde{x} \in \tilde{X}$ tale che $\tilde{x}_n \rightarrow \tilde{x}$ e poniamo allora $T\hat{x} = \tilde{x}$.

Consideriamo ora una coppia $\hat{x}, \hat{y} \in \hat{X}$ e la coppia corrispondente $\tilde{x}, \tilde{y} \in \tilde{X}$ nell'applicazione T . Abbiamo

$$\hat{d}(\hat{x}, \hat{y}) = \lim_{n \rightarrow \infty} \hat{d}(\hat{x}_n, \hat{y}_n)$$

e

$$\tilde{d}(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} \tilde{d}(\tilde{x}_n, \tilde{y}_n)$$

dove le successioni appartengono rispettivamente a $\widetilde{W}$ e a W e sono definite come indicato sopra. Poiché W e $\widetilde{W}$ sono isometrici è $\widehat{d}(\widehat{x}_n, \widehat{y}_n) = \tilde{d}(\tilde{x}_n, \tilde{y}_n)$ e le distanze in $\widetilde{X}$ e $\widehat{X}$ sono le medesime. Quindi $\widetilde{X}$ e $\widehat{X}$ sono isometrici. ■

1.4 Problemi

1. Sia ℓ^∞ lo spazio delle successioni $x = (\xi_j)_{j \in \mathbb{N}}$ limitate di numeri complessi, ossia tali che

$$|\xi_j| \leq c_x \quad (j = 1, 2, \dots)$$

per ogni x . Mostrare che

$$d(x, y) = \sup_{j \in \mathbb{N}} |\xi_j - \eta_j|$$

dove $y = (\eta_j) \in \ell^\infty$ definisce una metrica su ℓ^∞ .

2. Sia $C[a, b]$ l'insieme di tutte le funzioni a valori reali $x(t), y(t), \dots$ definite e continue nella variabile t sull'intervallo chiuso $J = [a, b]$. Mostrare che

$$d(x, y) = \max_{t \in J} |x(t) - y(t)|$$

definisce una metrica su $C[a, b]$.

3. Sia ℓ^p ($p \geq 1$) lo spazio delle successioni $x = (\xi_j)_{j \in \mathbb{N}}$ tali che

$$\sum_{j=1}^{\infty} |\xi_j|^p < \infty.$$

Mostrare che

$$d(x, y) = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p}$$

dove $y = (\eta_j) \in \ell^p$ definisce una metrica su ℓ^p .

Suggerimento: Utilizzare la disuguaglianza di Minkowski

$$\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p \right)^{1/p} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{1/p}.$$

4. Sia la distanza $D(A, B)$ fra due sottoinsiemi non vuoti A e B di uno spazio metrico (X, d) definita essere

$$D(A, B) = \inf_{\substack{a \in A \\ b \in B}} d(a, b).$$

Mostrare che $D(A, B)$ non definisce una metrica sull'insieme dei sottoinsiemi non vuoti di X .

5. Mostrare che lo spazio ℓ^∞ non è separabile.

Soluzione: Sia $y = (\eta_1, \eta_2, \dots)$ una successione di 0 e di 1. Allora $y \in \ell^\infty$ e due successioni differenti di questo tipo distano 1. Essi possono essere messi in corrispondenza biunivoca coi reali nell'intervallo $[0, 1]$, giacché un qualunque reale nell'intervallo $[0, 1]$ è rappresentabile in forma binaria come

$$\frac{\eta_1}{2^1} + \frac{\eta_2}{2^2} + \dots + \frac{\eta_n}{2^n} + \dots$$

Quindi non sono numerabili. Consideriamo ora le palle di raggio $1/3$ centrate in queste particolari successioni. Esse sono non numerabili ed hanno a due a due intersezione nulla. Quindi se esistesse un insieme M numerabile e denso in ℓ^∞ ...

6. Mostrare che lo spazio ℓ^p ($p \geq 1$) è separabile.

Soluzione: Sia M l'insieme delle successioni di ℓ^p della forma

$$y = (\eta_1, \eta_2, \dots, \eta_n, 0, 0, \dots)$$

con n intero positivo e le η_j razionali. M è numerabile. Si tratta di mostrare che è denso in ℓ^p . Sia dato un $x = (\xi_j) \in \ell^p$ arbitrario. Per ogni $\varepsilon > 0$ esiste un n tale che

$$\sum_{j=n+1}^{\infty} |\xi_j|^p < \frac{\varepsilon^p}{2}.$$

Scegliamo un $y \in M$ tale che

$$\sum_{j=1}^n |\xi_j - \eta_j|^p < \frac{\varepsilon^p}{2}.$$

Quindi ...

7. Mostrare che ℓ^∞ è completo.

Suggerimento: Sia (x_m) una successione di Cauchy di ℓ^∞ , dove $(x_m) = (\xi_1^{(m)}, \xi_2^{(m)}, \dots)$.

Mostrare che le successioni di numeri $(\xi_j^{(m)})$ per ogni j fissato sono di Cauchy ed quindi convergenti, ossia $\lim_{m \rightarrow \infty} \xi_j^{(m)} = \xi_j$. Mostrare infine che $x_m \rightarrow x$, dove $x = (\xi_1, \xi_2, \dots) \in \ell^\infty$.

8. Mostrare che ℓ^p è completo.

Suggerimento: Seguire la medesima via seguita nel caso precedente.

9. Mostrare che $C[a, b]$ è completo.

10. Mostrare che l'insieme dei polinomi considerati come funzioni di t sull'intervallo $[a, b]$ e con metrica definita come su $C[a, b]$ non è completo.

11. Se (x_n) e (x'_n) in (X, d) soddisfano a $d(x_n, x'_n) \rightarrow 0$ e (x_n) ammette limite, mostrare che anche (x'_n) converge ed ha il medesimo limite.

12. Se (x_n) e (x'_n) in (X, d) ammettono il medesimo limite, mostrare che $d(x_n, x'_n) \rightarrow 0$.

Capitolo 2

Spazi Normati. Spazi di Banach

Si ottengono degli spazi metrici particolarmente utili ed importanti se si considera uno spazio vettoriale e si definisce in esso una metrica a mezzo di una *norma*. Lo spazio risultante è chiamato *spazio normato*. Se uno spazio vettoriale normato è completo viene chiamato *spazio di Banach*. La teoria degli spazi normati, in particolare degli spazi di Banach, e la teoria degli operatori lineari definiti su di essi costituiscono la parte maggiormente sviluppata dell'analisi funzionale.

2.1 Spazio Vettoriale

2.1 DEFINIZIONE (SPAZIO VETTORIALE)

Uno *spazio vettoriale* (o *spazio lineare*) su un campo K è un insieme non vuoto X di elementi $x, y, \dots$ (chiamati *vettori*) dotato di due operazioni algebriche. Queste operazioni sono chiamate *somma vettoriale* e *moltiplicazione di vettori per scalari*, cioè per elementi di K .

La **somma vettoriale** associa ad ogni coppia ordinata (x, y) di vettori un vettore $x + y$ chiamato la *somma* di x e y , in tal modo che siano soddisfatte le seguenti proprietà. La somma vettoriale è commutativa ed associativa, cioè per tutti i vettori si ha che

$$\begin{aligned}x + y &= y + x \\x + (y + z) &= (x + y) + z;\end{aligned}$$

inoltre esiste un vettore 0 , chiamato *vettore nullo*, e per ogni vettore x un vettore $-x$ tali che per tutti i vettori si ha che

$$\begin{aligned}x + 0 &= x \\x + (-x) &= 0.\end{aligned}$$

Ciò X rispetto alla somma vettoriale è un gruppo additivo abeliano.

La **moltiplicazione per scalari** associa ad ogni vettore x e scalare α un vettore αx (scritto anche $x\alpha$) chiamato il *prodotto* di α e x , in tal modo che per tutti i vettori x, y e scalari α, β si ha che

$$\begin{aligned}\alpha(\beta x) &= (\alpha\beta)x \\1x &= x\end{aligned}$$

e le leggi distributive

$$\begin{aligned}\alpha(x+y) &= \alpha x + \beta y \\ (\alpha+\beta)x &= \alpha x + \beta x.\end{aligned}$$

■

Dalla definizione vediamo che la somma vettoriale è un'applicazione $X \times X \rightarrow X$, mentre la moltiplicazione per scalari è un'applicazione $K \times X \rightarrow X$.

K è chiamato il **campo scalare** (o *campo dei coefficienti*) dello spazio vettoriale X , e X è chiamato uno **spazio vettoriale reale** se $K = \mathbb{R}$ (il campo dei numeri reali) ed uno **spazio vettoriale complesso** se $K = \mathbb{C}$ (il campo dei numeri complessi).

L'uso dello 0 sia per lo scalare 0 che per il vettore nullo non dovrebbe, in generale, creare confusione. Se fosse desiderabile per ragioni di chiarezza, si può indicare il vettore nullo con **0**.

Il lettore può provare che per tutti i vettori e gli scalari

$$\begin{aligned}0x &= \mathbf{0} \\ \alpha\mathbf{0} &= \mathbf{0}\end{aligned}$$

e

$$(-1)x = -x.$$

Un **sottospazio** di uno spazio vettoriale X è un sottoinsieme non vuoto Y di X tale che per ogni $y_1, y_2 \in Y$ e tutti gli scalari α, β si ha che $\alpha y_1 + \beta y_2 \in Y$. Quindi Y stesso è uno spazio vettoriale, le due operazioni algebriche essendo quelle indotte da X .

Un speciale sottospazio di X è il *sottospazio improprio* $Y = X$. Ogni altro sottospazio di X ($\neq \{0\}$) è chiamato *proprio*.

Un altro sottospazio speciale di un qualunque spazio vettoriale X è $Y = \{0\}$.

Una **combinazione lineare** dei vettori $x_1, \dots, x_m$ di uno spazio vettoriale X è un'espressione della forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m$$

dove i coefficienti $\alpha_1, \dots, \alpha_m$ sono scalari qualunque.

Per ogni sottoinsieme non vuoto $M \subset X$ l'insieme di tutte le combinazioni lineari di vettori di M è chiamato l'involuppo o lo **span** di M e si scrive

$$\text{span } M.$$

Ovviamente è un sottospazio Y di M e diciamo che Y è **generato** da M .

Introduciamo ora due concetti fra di loro collegati che verranno usati molto spesso nel seguito.

2.2 DEFINIZIONE (INDIPENDENZA LINEARE, DIPENDENZA LINEARE)

L'indipendenza e la dipendenza lineare di un dato insieme M di vettori $x_1, \dots, x_r$ ($r \geq 1$) in uno spazio vettoriale X sono definite a mezzo dell'equazione

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_r x_r = 0, \tag{2.1}$$

dove gli $\alpha_1, \dots, \alpha_r$ sono scalari. Chiaramente l'equazione (2.1) vale per $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$. Se questa è la sola r -pla di scalari per cui la (2.1) è valida l'insieme M è detto *linearmente indipendente*. M è detto *linearmente dipendente* se M non è linearmente indipendente, cioè se (2.1) è anche valida per una r -pla di scalari non tutti zero.

Un sottoinsieme arbitrario M di X è detto *linearmente indipendente* se ogni sottoinsieme finito non vuoto di M è linearmente indipendente. M è detto *linearmente dipendente* se non è linearmente indipendente. ■

Una motivazione per questa terminologia proviene dal fatto che se $M = \{x_1, \dots, x_r\}$ è linearmente dipendente almeno un vettore di M può essere scritto come combinazione lineare degli altri; per esempio se (2.1) vale con un $\alpha_r \neq 0$ allora M è linearmente dipendente e possiamo risolvere (2.1) rispetto a x_r e ottenere

$$x_r = \beta_1 x_1 + \dots + \beta_{r-1} x_{r-1} \quad (\beta_j = -\alpha_j / \alpha_r).$$

Possiamo usare i concetti di dipendenza ed indipendenza lineare per definire la dimensione di uno spazio vettoriale.

2.3 DEFINIZIONE (BASE DI HAMEL)

Se X è uno spazio vettoriale qualunque e B è un sottoinsieme linearmente indipendente di X che genera X , allora B è chiamato una **base** (o **base di Hamel**) per X . ■

Uno spazio vettoriale X è detto *finito dimensionale* se ammette una base B di n vettori linearmente indipendenti. In questo caso è facile dimostrare che ogni altra base contiene n vettori indipendenti. n è perciò un numero caratteristico di X ed è chiamato la **dimensione** di X e si scrive $n = \dim X$. Per definizione $X = 0$ è finito dimensionale e $\dim X = 0$.

Se $\dim X = n$ una n -pla qualunque di vettori $\{e_1, \dots, e_n\}$ di X linearmente indipendenti costituisce una **base per** X (o una *base in* X) ed ogni $x \in X$ ha una rappresentazione unica come combinazione lineare di questi vettori, ossia

$$x = \alpha_1 e_1 + \dots + \alpha_n e_n.$$

Se X non è finito dimensionale si dice *infinito dimensionale*.

Anche nel caso infinito dimensionale ogni $x \in X$ non nullo ha una rappresentazione unica come combinazione lineare di (in numero finito!) elementi di B con coefficienti scalari non tutti nulli.

Ogni spazio vettoriale $X \neq \{0\}$ ha una base.

Per spazi vettoriali arbitrari infinito dimensionali la prova richiede l'uso del lemma di Zorn ed è rinviata a dopo che avremo introdotto questo lemma per altri propositi.

Menzioniamo il fatto che anche le basi di un dato spazio vettoriale X infinito dimensionale hanno lo stesso numero cardinale. Una prova richiederebbe alcuni strumenti piuttosto avanzati della teoria degli insiemi. Anche nel caso infinito dimensionale questo numero è chiamato la **dimensione** di X .

Più in là avremo bisogno del seguente semplice teorema.

2.4 Teorema (Dimensioni di un Sottospazio)

Sia X uno spazio vettoriale n -dimensionale. Allora ogni sottospazio proprio Y di X ha dimensioni minori di n . ■

Dimostrazione. Se $n = 0$ allora $X = \{0\}$ e non ha sottospazi propri. Se $\dim Y = 0$ allora $Y = \{0\}$ e $X \neq Y$ implica che $\dim X \geq 1$. Chiaramente $\dim Y \leq \dim X = n$. Se $\dim Y$ fosse n allora Y avrebbe una base di n elementi, che sarebbe anche una base per X perché $\dim X = n$, così che $X = Y$. Ciò mostra che un qualunque insieme di vettori linearmente indipendenti in Y deve avere meno di n elementi e quindi $\dim Y < n$. ■

2.2 Spazio Normato. Spazio di Banach

In molti casi uno spazio vettoriale X può essere al medesimo tempo uno spazio metrico perché una metrica d è definita su X . Tuttavia se non v'è relazione fra la struttura algebrica e la metrica non possiamo aspettarci una teoria utile ed applicabile che combini entrambi

i concetti. Per garantire una tale relazione fra gli aspetti “algebrici” e “geometrici” di X definiamo su X una metrica d in un modo speciale. Prima introduciamo un concetto ausiliario, quello di *norma*, che usa le operazioni algebriche dello spazio vettoriale. Poi utilizziamo la norma per ottenere una metrica d del tipo desiderato. Questa idea conduce al concetto di *spazio normato*.

2.5 DEFINIZIONE (SPAZIO NORMATO, SPAZIO DI BANACH)

Uno *spazio normato* X è uno spazio vettoriale dotato di una norma. Uno *spazio di Banach* è uno spazio normato completo (completo nella metrica definita dalla norma). La **norma** su un spazio vettoriale (reale o complesso) X è una funzione a valori reali su X il cui valore ad ogni $x \in X$ è indicato con

$$\|x\| \quad (\text{si legga “norma di } x\text{”})$$

ed ha le proprietà

$$(N1) \quad \|x\| \geq 0$$

$$(N2) \quad \|x\| = 0 \iff x = 0$$

$$(N3) \quad \|\alpha x\| = |\alpha| \|x\|$$

$$(N4) \quad \|x + y\| \leq \|x\| + \|y\| \quad (\text{Disuguaglianza Triangolare});$$

dove x e y sono vettori arbitrari in X e α è uno scalare qualunque.

Una norma su X definisce una metrica d su X che è data da

$$d(x, y) = \|x - y\| \quad (x, y \in X)$$

ed è chiamata la *metrica indotta dalla norma*. Lo spazio normato appena definito si indica con $(X, \|\cdot\|)$ o semplicemente con X . ■

2.6 Lemma (Invarianza per Traslazioni)

Una metrica d indotta da una norma in uno spazio normato X soddisfa a

$$\begin{aligned} d(x + a, y + a) &= d(x, y) \\ d(\alpha x, \alpha y) &= |\alpha| d(x, y) \end{aligned}$$

per tutti gli $x, y, a \in X$ ed ogni scalare α . ■

Dimostrazione. Abbiamo

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

e

$$d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = |\alpha| \|x - y\| = |\alpha| d(x, y). \quad \blacksquare$$

2.3 Ulteriori Proprietà degli Spazi Normati

Per definizione un **sottospazio** Y di uno spazio normato X è un sottospazio di X considerato come uno spazio vettoriale con una norma ottenuta restringendo la norma su X ad Y . Questa norma su Y è detta *indotta* dalla norma su X . Se Y è chiuso in X allora Y è chiamato un **sottospazio chiuso** di X .

Per definizione un **sottospazio** Y di uno spazio di Banach X è un sottospazio di X considerato come uno spazio normato. Quindi *non* richiediamo che Y sia completo.

A questo riguardo è utile il Teorema 1.13 perché fornisce immediatamente il seguente teorema.

2.7 Teorema (Sottospazio di uno Spazio di Banach)

Un sottospazio Y di uno spazio di Banach X è completo se e solo se l'insieme Y è chiuso in X . ■

Convergenza di successioni e concetti collegati in uno spazio normato seguono facilmente dalle corrispondenti definizioni 1.7 e 1.9 per gli spazi metrici e dal fatto che ora $d(x, y) = \|x - y\|$.

- (i) Una successione (x_n) in uno spazio normato X è *convergente* se X contiene un x tale che

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0.$$

Allora scriviamo $x_n \rightarrow x$ e chiamiamo x il *limite* di (x_n) .

- (ii) Una successione (x_n) in uno spazio normato è di *Cauchy* se per ogni $\varepsilon > 0$ esiste un N tale che

$$\|x_m - x_n\| < \varepsilon \quad \text{per tutti gli } m, n > N.$$

Le successioni erano disponibili anche in un generico spazio metrico. In uno spazio normato possiamo fare un passo avanti ed usare le serie.

Le **serie infinite** possono ora essere definite in un modo analogo a quello dell'analisi. Infatti se (x_k) è una successione in uno spazio normato X , possiamo associare a (x_k) la successione (s_n) di *somme parziali*

$$s_n = x_1 + x_2 + \cdots + x_n$$

dove $n = 1, 2, \dots$. Se (s_n) è convergente

$$s_n \rightarrow s \quad \text{cioè} \quad \|s_n - s\| \rightarrow 0$$

allora la *serie infinita* o, per brevità, la *serie*

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots$$

è detta *convergere* od *essere convergente*, s è chiamata la *somma* della serie e si scrive

$$s = \sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots.$$

Se $\|x_1\| + \|x_2\| + \cdots$ converge la serie è detta **assolutamente convergente**. Tuttavia in uno spazio normato la assoluta convergenza implica la convergenza se e solo se X è completo.

Il concetto di convergenza di una serie può essere usato per definire una “base”. Se uno spazio normato X contiene una successione (e_n) con la proprietà che per ogni $x \in X$ v'è un'unica successione di scalari (α_n) tali che

$$\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \rightarrow 0$$

allora (e_n) è chiamata una **base di Schauder** (o *base*) per X . La serie

$$\sum_{k=1}^{\infty} \alpha_k e_k$$

che ha la somma x è allora chiamata l'*espansione* di x rispetto a (e_n) e si scrive

$$x = \sum_{k=1}^{\infty} \alpha_k e_k.$$

Se uno spazio normato X ha una base di Schauder allora è separabile. La dimostrazione è semplice e viene lasciata al lettore. Sorprendentemente l'inverso non è vero, cioè uno spazio di Banach separabile non ha necessariamente una base di Schauder.

Consideriamo ora il problema di completare uno spazio normato.

2.8 Teorema (Completamento)

Sia $X = (X, \|\cdot\|)$ uno spazio normato. Allora esiste uno spazio di Banach $\widehat{X}$ ed una isometria A da X su un sottospazio W di $\widehat{X}$ che è denso in $\widehat{X}$. Lo spazio $\widehat{X}$ è unico a meno di isometrie. ■

Dimostrazione. Il teorema 1.17 implica l'esistenza di uno spazio metrico completo $\widehat{X} = (\widehat{X}, \widehat{d})$ e di una isometria $A : X \rightarrow W = A(X)$ dove W è denso in $\widehat{X}$ ed $\widehat{X}$ è unico a meno di isometrie. Conseguentemente per provare il teorema dobbiamo fare di $\widehat{X}$ uno spazio vettoriale ed introdurre in esso una norma conveniente.

Per definire su $\widehat{X}$ le due operazioni algebriche di uno spazio vettoriale consideriamo due qualunque $\widehat{x}, \widehat{y} \in \widehat{X}$ e due loro rappresentanti qualunque $(x_n) \in \widehat{x}$ e $(y_n) \in \widehat{y}$. Si ricordi che $\widehat{x}$ e $\widehat{y}$ sono classi di equivalenza di successioni di Cauchy in X . Poniamo $z_n = x_n + y_n$. Allora (z_n) è di Cauchy in X perché

$$\|z_n - z_m\| = \|x_n + y_n - (x_m + y_m)\| \leq \|x_n - x_m\| + \|y_n - y_m\|.$$

Definiamo la somma $\widehat{z} = \widehat{x} + \widehat{y}$ di $\widehat{x}$ e $\widehat{y}$ come la classe di equivalenza di cui (z_n) è il rappresentante; così $(z_n) \in \widehat{z}$. Questa definizione è indipendente dalla particolare scelta delle successioni di Cauchy appartenenti a $\widehat{x}$ e a $\widehat{y}$. Infatti la definizione di equivalenza fra successioni di Cauchy introdotta nella sezione 1.3 mostra che se $(x_n) \sim (x'_n)$ e $(y_n) \sim (y'_n)$ allora $(x_n + y_n) \sim (x'_n + y'_n)$ perché

$$\|x_n + y_n - (x'_n + y'_n)\| \leq \|x_n - x'_n\| + \|y_n - y'_n\|.$$

Analogamente definiamo il prodotto $\alpha \widehat{x} \in \widehat{X}$ di uno scalare α e di un elemento $\widehat{x}$ come la classe di equivalenza di cui (αx_n) è un rappresentante. Nuovamente la definizione è indipendente dalla particolare scelta del rappresentante di $\widehat{x}$. L'elemento zero di $\widehat{X}$ è la classe di equivalenza di tutte le successioni di Cauchy che convergono a zero. Non è difficile verificare che queste due operazioni algebriche hanno tutte le proprietà richieste dalla definizione, così che $\widehat{X}$ è uno spazio vettoriale. Dalla definizione segue che su W le operazioni di spazio vettoriale indotte da $\widehat{X}$ coincidono con quelle indotte da X a mezzo di A .

Definiamo su $\widehat{X}$ la norma ponendo $\|\widehat{x}\|_1 = \widehat{d}(0, \widehat{x})$ per ogni $\widehat{x} \in \widehat{X}$. È ovvio che $\|\cdot\|_1$ soddisfa (N1) e (N2). Gli altri due assiomi (N3) e (N4) seguono da quelli per $\|\cdot\|$ con un processo di limite. Infatti per verificare (N3) è sufficiente notare che, se (x_n) è un rappresentante della classe d'equivalenza che definisce $\widehat{x}$ e se $\widehat{x}_n$ è il corrispondente di x_n nell'isometria A di X su W , abbiamo

$$\widehat{x}_n \rightarrow \widehat{x}, \quad \alpha \widehat{x}_n \rightarrow \alpha \widehat{x}$$

e quindi per la continuità della distanza

$$\|\hat{x}\|_1 = \lim_{n \rightarrow \infty} \hat{d}(0, \hat{x}_n), \quad \|\alpha \hat{x}\|_1 = \lim_{n \rightarrow \infty} \hat{d}(0, \alpha \hat{x}_n).$$

Grazie all'isometria A di X su W ed al fatto che (N3) è soddisfatto in X si ottiene

$$\|\alpha \hat{x}\|_1 = \lim_{n \rightarrow \infty} \hat{d}(0, \alpha \hat{x}_n) = \lim_{n \rightarrow \infty} d(0, \alpha x_n) = |\alpha| \lim_{n \rightarrow \infty} d(0, x_n) = |\alpha| \lim_{n \rightarrow \infty} \hat{d}(0, \hat{x}_n)$$

cioè $\|\alpha \hat{x}\|_1 = |\alpha| \|\hat{x}\|_1$. Analoga dimostrazione si può fare per verificare (N4). ■

2.4 Spazi Normati Finito Dimensionali e Sottospazi

Rozzamente parlando il seguente lemma stabilisce che nel caso di vettori linearmente indipendenti non è possibile trovare una combinazione lineare che coinvolge grandi scalari ed al medesimo tempo rappresenta un piccolo vettore.

2.9 Lemma (Combinazione Lineare)

Sia $\{x_1, \dots, x_n\}$ un insieme di vettori linearmente indipendenti in uno spazio normato X (di dimensioni qualunque). Allora esiste un numero $c > 0$ tale che per ogni scelta degli scalari $\alpha_1, \dots, \alpha_n$ abbiamo che

$$\|\alpha_1 x_1 + \dots + \alpha_n x_n\| \geq c(|\alpha_1| + \dots + |\alpha_n|) \quad (c > 0). \quad (2.2)$$

■

Dimostrazione. Scriviamo $s = |\alpha_1| + \dots + |\alpha_n|$. Se $s = 0$ tutti gli $|\alpha_i|$ sono zero e allora la (2.2) è soddisfatta per ogni c . Sia $s > 0$. Allora la (2.2) è equivalente alla disuguaglianza che si ottiene dalla (2.2) dividendo per s e scrivendo $\beta_j = \alpha_j/s$, cioè

$$\|\beta_1 x_1 + \dots + \beta_n x_n\| \geq c \quad \left(\sum_{j=1}^n |\beta_j| = 1 \right). \quad (2.3)$$

È quindi sufficiente provare l'esistenza di un $c > 0$ tale che la (2.3) vale per ogni n -pla di scalari $\beta_1, \dots, \beta_n$ con $\sum_{j=1}^n |\beta_j| = 1$.

Supponiamo che ciò sia falso. Ossia supponiamo che per ogni $c > 0$ esista una n -pla di scalari $\beta_1, \dots, \beta_n$ con $\sum_{j=1}^n |\beta_j| = 1$ tali che $\|\beta_1 x_1 + \dots + \beta_n x_n\| < c$. In corrispondenza della successione $c = 1/m$ esiste allora una successione (y_m) di vettori

$$y_m = \beta_1^{(m)} x_1 + \dots + \beta_n^{(m)} x_n \quad \left(\sum_{j=1}^n |\beta_j^{(m)}| = 1 \right)$$

tali che

$$\|y_m\| < \frac{1}{m}$$

e quindi tali che

$$\|y_m\| \rightarrow 0 \quad \text{come } m \rightarrow \infty.$$

Ora ragioniamo come segue. Poiché $\sum_{j=1}^n |\beta_j^{(m)}| = 1$ abbiamo che $|\beta_j^{(m)}| \leq 1$. Quindi per ogni fisso j la successione

$$(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \dots)$$

è limitata. Di conseguenza per il teorema di Bolzano–Weierstrass $(\beta_1^{(m)})$ ammette una successione estratta convergente. Sia β_1 il limite di questa successione estratta e sia $(y_{1,m})$ la corrispondente successione estratta di (y_m) . Per il medesimo argomento $(y_{1,m})$ ammette una successione estratta $(y_{2,m})$ per cui la corrispondente successione estratta di scalari $(\beta_2^{(m)})$ converge; sia β_2 il limite. Continuando in questo modo dopo n passi si ottiene una successione estratta $(y_{n,m}) = (y_{n,1}, y_{n,2}, \dots)$ di (y_m) i cui termini sono della forma

$$y_{n,m} = \sum_{j=1}^n \gamma_j^{(m)} x_j \quad \left(\sum_{j=1}^n |\gamma_j^{(m)}| = 1 \right)$$

con scalari $\gamma_j^{(m)}$ che soddisfano il limite $\gamma_j^{(m)} \rightarrow \beta_j$ per $m \rightarrow \infty$. Quindi per $m \rightarrow \infty$

$$y_{n,m} \rightarrow y = \sum_{j=1}^n \beta_j x_j$$

dove $\sum |\beta_j| = 1$, così che non tutti i β_j possono essere zero. Poiché $\{x_1, \dots, x_n\}$ è un insieme linearmente indipendente abbiamo che $y \neq 0$. D'altro lato $y_{n,m} \rightarrow y$ implica $\|y_{n,m}\| \rightarrow \|y\|$ per la continuità della norma. Poiché $\|y_m\| \rightarrow 0$ per assunzione e $(y_{n,m})$ è una successione estratta di (y_m) dobbiamo avere $\|y_{n,m}\| \rightarrow 0$. Quindi $\|y\| = 0$ così che $y = 0$ per (N2). Ciò contraddice $y \neq 0$ ed il lemma è provato. ■

Come prima applicazione del lemma proviamo il seguente teorema basilare.

2.10 Teorema (Completezza)

Ogni sottospazio finito dimensionale Y di uno spazio normato X è completo. In particolare ogni spazio normato finito dimensionale è completo. ■

Dimostrazione. Consideriamo un'arbitraria successione di Cauchy (y_m) in Y , il cui limite sarà indicato con y . Sia $\dim Y = n$ e $\{e_1, \dots, e_n\}$ una base qualunque di Y . Allora ciascun y_m ha un'unica rappresentazione della forma

$$y_m = \alpha_1^{(m)} e_1 + \dots + \alpha_n^{(m)} e_n.$$

Poiché (y_m) è una successione di Cauchy, per ogni $\varepsilon > 0$ v'è un N tale che $\|y_m - y_r\| < \varepsilon$ quando $m, r > N$. Da ciò e dal Lemma 2.9 abbiamo che per qualche $c > 0$

$$\varepsilon > \|y_m - y_r\| = \left\| \sum_{j=1}^n (\alpha_j^{(m)} - \alpha_j^{(r)}) e_j \right\| \geq c \sum_{j=1}^n |\alpha_j^{(m)} - \alpha_j^{(r)}|,$$

dove $m, r > N$. Dividendo per c si ottiene

$$|\alpha_j^{(m)} - \alpha_j^{(r)}| \leq \sum_{j=1}^n |\alpha_j^{(m)} - \alpha_j^{(r)}| < \frac{\varepsilon}{c} \quad (m, r > N).$$

Ciò mostra che ciascuna delle n successioni

$$(\alpha_j^{(m)}) = (\alpha_j^{(1)}, \alpha_j^{(2)}, \dots) \quad j = 1, \dots, n$$

è di Cauchy in $\mathbb{R}$ o in $\mathbb{C}$. Quindi converge; sia α_j il limite. Usando allora gli n limiti $\alpha_1, \dots, \alpha_n$ definiamo

$$y = \alpha_1 e_1 + \dots + \alpha_n e_n.$$

Chiaramente $y \in Y$. Inoltre

$$\|y_m - y\| = \left\| \sum_{j=1}^n (\alpha_j^{(m)} - \alpha_j) e_j \right\| \leq \sum_{j=1}^n |\alpha_j^{(m)} - \alpha_j| \|e_j\|.$$

A destra $\alpha_j^{(m)} \rightarrow \alpha_j$. Quindi $\|y_m - y\| \rightarrow 0$, cioè $y_m \rightarrow y$. Ciò mostra che (y_m) è convergente in Y . Poiché (y_m) era una successione di Cauchy in Y , ciò prova che Y è completo. ■

Da questo teorema e dal Teorema 1.13 si deriva il seguente teorema.

2.11 Teorema (Chiusura)

Ogni sottospazio finito dimensionale Y di uno spazio normato X è chiuso in X . ■

Avremo bisogno di questo teorema in numerose occasioni nel seguito.

Un'altra proprietà interessante di uno spazio vettoriale finito dimensionale X è che tutte le norme su X portano alla medesima topologia per X , cioè gli insiemi aperti sono gli stessi, a prescindere dalla particolare scelta della norma in X . I dettagli sono i seguenti.

2.12 DEFINIZIONE (NORME EQUIVALENTI)

Una norma $\|\cdot\|_1$ su uno spazio vettoriale X è detta *equivalente* alla norma $\|\cdot\|_2$ su X se esistono dei numeri positivi a e b tali che per ogni $x \in X$ si ha

$$a\|x\|_2 \leq \|x\|_1 \leq b\|x\|_2. \quad (2.4)$$

■

Questo concetto è motivato dal seguente fatto.

Norme equivalenti su X definiscono la medesima topologia per X .

Infatti ciò segue dalla (2.4) e dal fatto che ogni insieme aperto non vuoto è un'unione di palle aperte. Lasciamo i dettagli di una prova formale al lettore, che può anche mostrare che le successioni di Cauchy in $(X, \|\cdot\|_1)$ e $(X, \|\cdot\|_2)$ sono le stesse.

Usando il Lemma 2.9 possiamo ora provare il seguente teorema (che *non* vale per gli spazi infinito dimensionali).

2.13 Teorema (Norme Equivalenti)

Su uno spazio vettoriale finito dimensionale X ogni norma $\|\cdot\|_1$ è equivalente a qualsiasi altra norma $\|\cdot\|_2$. ■

Dimostrazione. Sia $\dim X = n$ e $\{e_1, \dots, e_n\}$ una base qualunque di X . Allora ogni $x \in X$ ha un'unica rappresentazione

$$x = \alpha_1 e_1 + \dots + \alpha_n e_n.$$

Per il Lemma 2.9 esiste una costante positiva c tale che

$$\|x\|_1 \geq c(|\alpha_1| + \dots + |\alpha_n|).$$

D'altra parte la disuguaglianza triangolare dà

$$\|x\|_2 \leq \sum_{j=1}^n |\alpha_j| \|e_j\|_2 \leq k \sum_{j=1}^n |\alpha_j| \quad k = \max_j \|e_j\|_2.$$

Assieme danno $a\|x\|_2 \leq \|x\|_1$ dove $a = c/k > 0$. L'altra disuguaglianza in (2.4) si ottiene scambiando il ruolo di $\|\cdot\|_1$ e $\|\cdot\|_2$ nelle considerazioni precedenti. ■

Questo teorema ha un notevole importanza pratica. Ad esempio implica che la convergenza o la divergenza di una successione in uno spazio vettoriale finito dimensionale non dipende dalla particolare scelta della norma su questo spazio.

2.5 Compattezza e Dimensioni Finite

Alcune altre proprietà basilari degli spazi normati finito dimensionali e dei loro sottospazi sono legate al concetto di compattezza. Quest'ultima si definisce come segue.

2.14 DEFINIZIONE (COMPATTEZZA)

Uno spazio metrico X è detto *compatto*¹ se ogni successione in X ha una successione estratta convergente. Un sottoinsieme M di X è detto *compatto* se M è compatto considerato come sottospazio di X , cioè se ogni successione in M ha una successione estratta convergente ad un limite che è un elemento di M . ■

Una proprietà generale degli insiemi compatti è espressa nel seguente Lemma.

2.15 Lemma (Compattezza)

Un sottoinsieme compatto M di uno spazio metrico X è chiuso e limitato. ■

Dimostrazione. Per ogni $x \in \overline{M}$ v'è una successione (x_n) in M tale che $x_n \rightarrow x$; cf. 1.12(a). Poiché M è compatto $x \in M$. Quindi M è chiuso perché $x \in \overline{M}$ era arbitrario. Proviamo che M è limitato. Si noti che se $\sup_{x,y \in M} d(x,y) = \infty$ è anche $\sup_{x,y \in M} d(x,b) = \infty$ per un qualunque b fisso appartenente a X . Infatti se così non fosse avremmo $\sup d(x,y) \leq \sup d(x,b) + \sup d(b,y) < \infty$ che non è possibile. Quindi se M fosse non limitato conterebbe una successione (y_n) tale che $d(y_n, b) > n$ e quindi tale che $\lim_{n \rightarrow \infty} d(y_n, b) = \infty$. Questa successione non potrebbe avere una successione estratta convergente perché allora $d(y_n, b)$ ammetterebbe una successione estratta convergente, ciò che è impossibile perché è divergente. ■

L'inverso di questo lemma è in generale falso.

Dimostrazione. Per provare questo importante fatto consideriamo la successione (e_n) in l^2 , dove $e_n = (\delta_{nj})$ ha l' n -mo termine 1 e tutti gli altri termini 0. Questa successione è limitata perché $\|e_n\| = 1$. Poiché è $\|e_n - e_m\| = \sqrt{2}$ per ogni n e m , la successione non ha punti di accumulazione. Quindi i suoi termini costituiscono un insieme di punti che è chiuso perché non ha punti di accumulazione. Per la medesima ragione questo insieme non è compatto. ■

Tuttavia per uno spazio normato finito dimensionale abbiamo il seguente Teorema.

2.16 Teorema (Compattezza)

In uno spazio normato finito dimensionale X un qualsiasi sottoinsieme $M \subset X$ è compatto se e solo se M è chiuso e limitato. ■

¹Più precisamente sequenzialmente compatto; questo è il tipo più importante di compattezza in analisi. Menzioniamo che ci sono due altri tipi di compattezza, ma per gli spazi metrici i tre concetti divengono identici.

Dimostrazione. La compattezza implica la chiusura e la limitatezza per il Lemma 2.15. Proviamo l'inverso. Sia M chiuso e limitato. Sia $\dim X = n$ e $\{e_1, \dots, e_n\}$ una base per X . Consideriamo una qualunque successione (x_m) in M . Ciasun x_m ha una rappresentazione

$$x_m = \xi_1^{(m)} e_1 + \dots + \xi_n^{(m)} e_n.$$

Poiché M è limitato lo è anche (x_m) , cioè $\|x_m\| \leq k$ per tutti gli m . Per il Lemma 2.9

$$k \geq \|x_m\| = \left\| \sum_{j=1}^n \xi_j^{(m)} e_j \right\| \geq c \sum_{j=1}^n |\xi_j^{(m)}|$$

dove $c > 0$. Quindi la successione di numeri $(\xi_j^{(m)})$ (j fisso) è limitata e, per il teorema di Bolzano–Weierstrass, ha un punto di accumulazione ξ_j ; qui $1 \leq j \leq n$. Come nella prova del Lemma 2.9 concludiamo che (x_m) ha una successione estratta (z_m) che converge a $z = \sum \xi_j e_j$. Poiché M è chiuso $z \in M$. Ciò mostra che la successione arbitraria (x_m) in M ha una successione estratta che converge in M . Quindi M è compatto. ■

La nostra discussione mostra il seguente. In $\mathbb{R}^n$ (o in ogni altro spazio normato finito dimensionale) i sottoinsiemi compatti sono precisamente i sottoinsiemi chiusi e limitati, così che questa proprietà (chiusura e limitatezza) può essere usata per *definire* la compattezza. Tuttavia questo non può più essere fatto nel caso degli spazi normati infinito dimensionali.

Una sorgente di altri risultati interessanti è il seguente lemma di F. Riesz.

2.17 Lemma (F. Riesz)

Siano Y e Z sottospazi di uno spazio normato X (di una dimensione qualunque) e supponiamo che Y sia chiuso e sia un sottospazio proprio di Z . Allora per ogni numero reale θ nell'intervallo $(0, 1)$ v'è un $z \in Z$ tale che

$$\|z\| = 1, \quad \|z - y\| \geq \theta \text{ per ogni } y \in Y. \quad \blacksquare$$

Dimostrazione. Consideriamo un qualsiasi $v \in Z - Y$ ed indichiamo la sua distanza da Y con a , cioè

$$a = \inf_{y \in Y} \|v - y\|. \quad (2.5)$$

Chiaramente $a > 0$ perché Y è chiuso. Prendiamo ora un qualunque $\theta \in (0, 1)$. Per la definizione di estremo inferiore v'è un $y_0 \in Y$ tale che

$$a \leq \|v - y_0\| \leq \frac{a}{\theta} \quad (2.6)$$

(si noti che $a/\theta > a$ perché $0 < \theta < 1$). Sia

$$z = c(v - y_0) \quad \text{dove} \quad c = \frac{1}{\|v - y_0\|}.$$

Allora $\|z\| = 1$ e mostriamo che $\|z - y\| \geq \theta$ per ogni $y \in Y$. Abbiamo che

$$\begin{aligned} \|z - y\| &= \|c(v - y_0) - y\| \\ &= c\|v - y_0 - c^{-1}y\| \\ &= c\|v - y_1\| \end{aligned}$$

dove

$$y_1 = y_0 + c^{-1}y.$$

La forma di y_1 mostra che $y_1 \in Y$. Quindi $\|v - y_1\| \geq a$ per definizione di a . Usando (2.5) e (2.6) otteniamo

$$\|z - y\| = c\|v - y_1\| \geq ca = \frac{a}{\|v - y_0\|} \geq \frac{a}{a/\theta} = \theta.$$

Poiché $y \in Y$ era arbitrario ciò completa la prova. ■

2.18 Teorema (Dimensioni Finite)

Se uno spazio normato X ha la proprietà che la palla chiusa unitaria $M = \{x : \|x\| \leq 1\}$ è compatta allora X è finito dimensionale. ■

Dimostrazione. Assumiamo che M sia compatto ma che $\dim X = \infty$ e mostriamo che ciò porta ad una contraddizione. Scegliamo un qualunque x_1 di norma 1. Questo x_1 genera uno sottospazio unidimensionale X_1 di X , che è chiuso (cf. 2.11) ed è un sottospazio proprio di X perché $\dim X = \infty$. Per il lemma di Riesz v'è un $x_2 \in X$ di norma 1 tale che

$$\|x_2 - x_1\| \geq \theta = \frac{1}{2}.$$

Gli elementi x_1, x_2 generano un sottospazio X_2 bidimensionale proprio e chiuso di X . Per il lemma di Riesz v'è un x_3 di norma 1 tale che per tutti gli $x \in X_2$ abbiamo che

$$\|x_3 - x\| \geq \frac{1}{2}.$$

In particolare

$$\begin{aligned} \|x_3 - x_1\| &\geq \frac{1}{2}, \\ \|x_3 - x_2\| &\geq \frac{1}{2}. \end{aligned}$$

Procedendo per induzione otteniamo una successione (x_n) di elementi $x_n \in M$ tali che

$$\|x_m - x_n\| \geq \frac{1}{2} \quad (m \neq n).$$

Ovviamente (x_n) non può avere una successione estratta convergente. Ciò contraddice la compattezza di M . Quindi la nostra assunzione $\dim X = \infty$ è falsa e $\dim X < \infty$. ■

Questo teorema ha varie applicazioni. Lo utilizzeremo come uno strumento basilare in connessione con i cosiddetti operatori compatti.

Gli insiemi compatti sono importanti perché hanno un “buon comportamento”; essi ammettono numerose proprietà basilari simili a quelle degli insiemi finiti e che non sono soddisfatte dagli insiemi non compatti. In connessione con le applicazioni continue una proprietà fondamentale è che gli insiemi compatti hanno immagini compatte.

2.19 Teorema (Applicazioni Continue)

Siano X e Y spazi metrici e $T : X \rightarrow Y$ un'applicazione continua. Allora l'immagine di un sottoinsieme compatto M di X sotto T è compatto. ■

Dimostrazione. Per la definizione di compattezza è sufficiente mostrare che ogni successione (y_n) nell'immagine $T(M) \subset Y$ contiene una successione estratta che converge in $T(M)$. Poiché $y_n \in T(M)$ abbiamo che $y_n = Tx_n$ per qualche $x_n \in M$. Poiché M è compatto (x_n) contiene una successione estratta (x_{n_k}) che converge in M . L'immagine di (x_{n_k}) è una successione estratta di (y_n) che converge in $T(M)$ per l'1.14 perché T è continua. Quindi $T(M)$ è compatto. ■

Da questo teorema concludiamo che la seguente proprietà, ben nota dall'analisi per le funzioni continue, si estende agli spazi metrici.

2.20 Corollario (Massimo e Minimo)

Un'applicazione continua T di un sottoinsieme M compatto di uno spazio metrico X in $\mathbb{R}$ assume un massimo ed un minimo in qualche punto di M . ■

Dimostrazione. $T(M) \subset \mathbb{R}$ è compatto per il Teorema 2.19 e chiuso e limitato per il Lemma 2.15 [applicato a $T(M)$], così che $\inf T(M) \in T(M)$, $\sup T(M) \in T(M)$ e le immagini inverse di questi due punti consistono dei punti di M in cui Tx è, rispettivamente, minimo e massimo. ■

2.6 Operatori Lineari

Nel caso degli spazi vettoriali ed in particolare degli spazi normati un'applicazione è chiamata un **operatore**.

Di speciale interesse sono gli operatori che “conservano” le due operazioni algebriche degli spazi vettoriali, nel senso della seguente definizione.

2.21 DEFINIZIONE (OPERATORI LINEARI)

Un *operatore lineare* T è un operatore tale che

- (i) il dominio $\mathcal{D}(T)$ di T è uno spazio vettoriale e l'immagine $\mathcal{R}(T)$ giace in uno spazio vettoriale sul medesimo campo,
- (ii) per tutti gli $x, y \in \mathcal{D}(T)$ e scalari α

$$\begin{aligned} T(x + y) &= Tx + Ty \\ T(\alpha x) &= \alpha Tx \end{aligned} \tag{2.7}$$

■

Si osservi la **notazione**; scriviamo Tx invece di $T(x)$; questa semplificazione è standard in analisi funzionale. Inoltre **per tutto il seguito useremo la seguente notazione**.

$\mathcal{D}(T)$ indica il dominio di T .

$\mathcal{R}(T)$ indica l'immagine di T .

$\mathcal{N}(T)$ indica lo spazio nullo di T .

Per definizione lo **spazio nullo** di T è l'insieme di tutti gli $x \in \mathcal{D}(T)$ tali che $Tx = 0$. (Un'altra parola per lo spazio nullo è “kernel”. Non adotteremo questo termine perché dobbiamo riservarlo ad un altro scopo nella teoria delle equazioni integrali.)

Dobbiamo anche dire qualcosa sull'uso delle frecce in connessione con gli operatori. Sia $\mathcal{D}(T) \subset X$ ed $\mathcal{R}(T) \subset Y$, dove X e Y sono spazi vettoriali entrambi reali o complessi. Allora T è un operatore *da* (o applicazione *di*) $\mathcal{D}(T)$ **su** $\mathcal{R}(T)$ che scriviamo

$$T : \mathcal{D}(T) \rightarrow \mathcal{R}(T),$$

o da $\mathcal{D}(T)$ in Y , che scriviamo

$$T : \mathcal{D}(T) \rightarrow Y.$$

Se $\mathcal{D}(T)$ è tutto lo spazio X , allora – e solo allora – scriviamo

$$T : X \rightarrow Y.$$

Chiaramente (2.7) è equivalente a

$$T(\alpha x + \beta y) = \alpha T x + \beta T y. \quad (2.8)$$

Prendendo $\alpha = 0$ in (2.7) otteniamo la seguente formula di cui avremo bisogno molte volte nel seguito

$$T0 = 0. \quad (2.9)$$

2.22 DEFINIZIONE (OPERATORE IDENTITÀ)

L'operatore identità $I_X : X \rightarrow X$ è definito da $I_X x = x$ per tutti gli $x \in X$. Scriviamo anche semplicemente I per I_X ; così $Ix = x$. ■

2.23 DEFINIZIONE (OPERATORE ZERO)

L'operatore zero $0 : X \rightarrow Y$ è definito da $0x = 0$ per tutti gli $x \in X$. ■

2.24 Teorema (Immagine e Spazio Nullo)

Sia T un operatore lineare. Allora

- (a) L'immagine $\mathcal{R}(T)$ è uno spazio vettoriale.
- (b) Se $\dim \mathcal{D}(T) = n < \infty$, allora $\dim \mathcal{R}(T) \leq n$.
- (c) Lo spazio nullo $\mathcal{N}(T)$ è uno spazio vettoriale. ■

Dimostrazione. (a) Prendiamo due qualunque $y_1, y_2 \in \mathcal{R}(T)$ e mostriamo che $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$ per due scalari qualunque α e β . Poiché $y_1, y_2 \in \mathcal{R}(T)$ abbiamo che $y_1 = Tx_1$ e $y_2 = Tx_2$ per qualche $x_1, x_2 \in \mathcal{D}(T)$. Anche $\alpha x_1 + \beta x_2 \in \mathcal{D}(T)$ perché $\mathcal{D}(T)$ è uno spazio vettoriale. La linearità di T dà

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = \alpha y_1 + \beta y_2.$$

Quindi $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$. Poiché $y_1, y_2 \in \mathcal{R}(T)$ erano arbitrari e così lo erano gli scalari ciò prova che $\mathcal{R}(T)$ è uno spazio vettoriale.

(b) Scegliamo $n + 1$ elementi $y_1, \dots, y_{n+1}$ di $\mathcal{R}(T)$ in una maniera arbitraria. Allora abbiamo $y_1 = Tx_1, \dots, y_{n+1} = Tx_{n+1}$ per qualche $x_1, \dots, x_{n+1}$ in $\mathcal{D}(T)$. Poiché $\dim \mathcal{D}(T) = n$ questo insieme $\{x_1, \dots, x_{n+1}\}$ deve essere linearmente dipendente. Quindi

$$\alpha_1 x_1 + \dots + \alpha_{n+1} x_{n+1} = 0$$

per degli scalari non tutti nulli. Poiché T è lineare e $T0 = 0$ applicando T ad entrambi membri si ottiene

$$T(\alpha_1 x_1 + \dots + \alpha_{n+1} x_{n+1}) = \alpha_1 y_1 + \dots + \alpha_{n+1} y_{n+1} = 0.$$

Ciò mostra che l'insieme $\{y_1, \dots, y_{n+1}\}$ è linearmente dipendente perché gli α_j non sono tutti nulli. Ricordando che questo sottoinsieme di $\mathcal{R}(T)$ era stato scelto in una maniera

arbitraria ne concludiamo che $\mathcal{R}(T)$ non ammette sottoinsiemi linearmente indipendenti di $n + 1$ o più elementi. Per definizione ciò significa che $\dim \mathcal{R}(T) \leq n$.

(c) Prendiamo due qualunque $x_1, x_2 \in \mathcal{N}(T)$. Allora $Tx_1 = Tx_2 = 0$. Poiché T è lineare per scalari qualunque α, β abbiamo che

$$T(\alpha_1 x_1 + \beta_2 x_2) = 0.$$

Ciò mostra che $\alpha_1 x_1 + \beta_2 x_2 \in \mathcal{N}(T)$. Quindi $\mathcal{N}(T)$ è uno spazio vettoriale. ■

La seguente conseguenza immediata della parte (b) della dimostrazione è degna di nota.

Gli operatori lineari conservano la dipendenza lineare.

Occupiamoci ora dell'inverso di un operatore lineare. Ricordiamo dapprima che un'applicazione $T : \mathcal{D}(T) \rightarrow Y$ è detta **iniettiva** o **biunivoca** se punti differenti nel dominio hanno immagini differenti, cioè se per ogni $x_1, x_2 \in \mathcal{D}(T)$

$$x_1 \neq x_2 \implies Tx_1 \neq Tx_2 \quad (2.10)$$

o in maniera equivalente se

$$Tx_1 = Tx_2 \implies x_1 = x_2. \quad (2.11)$$

In questo caso esiste l'applicazione

$$\begin{array}{ll} T^{-1} : \mathcal{R}(T) \rightarrow \mathcal{D}(T) \\ y_0 \mapsto x_0 \end{array} \quad (y_0 = Tx_0) \quad (2.12)$$

che applica ogni $y_0 \in \mathcal{R}(T)$ su quel $x_0 \in \mathcal{D}(T)$ per cui $Tx_0 = y_0$. L'applicazione T^{-1} è chiamata l'**inversa** di T .

Dalla (2.12) abbiamo chiaramente che

$$\begin{array}{ll} T^{-1}Tx = x & \text{per tutti gli } x \in \mathcal{D}(T) \\ TT^{-1}x = x & \text{per tutti gli } x \in \mathcal{R}(T). \end{array}$$

In connessione con gli operatori lineari sugli spazi vettoriali la situazione è la seguente. L'inverso di un operatore lineare esiste se e solo se lo spazio nullo dell'operatore consiste solamente del vettore nullo. Più precisamente abbiamo il seguente utile criterio che utilizzeremo molto frequentemente.

2.25 Teorema (Operatore Inverso)

Siano X e Y spazi vettoriali entrambi reali o complessi. Sia $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare con dominio $\mathcal{D}(T) \subset X$ e immagine $\mathcal{R}(T) \subset Y$. Allora

(a) *L'inverso $T^{-1} : \mathcal{R}(T) \rightarrow \mathcal{D}(T)$ esiste se e solo se*

$$Tx = 0 \implies x = 0.$$

(b) *Se T^{-1} esiste, è un operatore lineare.*

(c) *Se $\dim \mathcal{D}(T) = n < \infty$ e T^{-1} esiste allora $\dim \mathcal{R}(T) = \dim \mathcal{D}(T)$.* ■

Dimostrazione. (a) Supponiamo che $Tx = 0$ implichi $x = 0$. Sia $Tx_1 = Tx_2$. Poiché T è lineare

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0$$

così che $x_1 - x_2 = 0$ per ipotesi. Quindi $Tx_1 = Tx_2$ implica $x_1 = x_2$ e T^{-1} esiste per la (2.11). Viceversa se T^{-1} esiste allora la (2.11) vale. Dalla (2.11) con $x_2 = 0$ e dalla (2.9) otteniamo

$$Tx_1 = T0 = 0 \implies x_1 = 0.$$

Ciò completa la dimostrazione di (a).

(b) Assumiamo che T^{-1} esiste e mostriamo che T^{-1} è lineare. Il dominio di T^{-1} è $\mathcal{R}(T)$ ed è uno spazio vettoriale per il Teorema 2.24(a). Consideriamo degli $x_1, x_2 \in \mathcal{D}(T)$ qualunque e le loro immagini

$$y_1 = Tx_1 \quad \text{e} \quad y_2 = Tx_2.$$

Allora

$$x_1 = T^{-1}y_1 \quad \text{e} \quad x_2 = T^{-1}y_2.$$

T è lineare e così per degli scalari qualunque α e β abbiamo

$$\alpha y_1 + \beta y_2 = \alpha Tx_1 + \beta Tx_2 = T(\alpha x_1 + \beta x_2).$$

Poiché $x_j = T^{-1}y_j$ ciò implica che

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1}y_1 + \beta T^{-1}y_2$$

e prova che T^{-1} è lineare.

(c) Abbiamo che $\dim \mathcal{R}(T) \leq \dim \mathcal{D}(T)$ per il Teorema 2.24(b) e $\dim \mathcal{D}(T) \leq \dim \mathcal{R}(T)$ per il medesimo teorema applicato a T^{-1} . ■

Menzioniamo infine una formula utile per l'inverso della composizione di operatori lineari.

2.26 Lemma (Inverso del Prodotto)

Siano $T : X \rightarrow Y$ e $S : Y \rightarrow Z$ operatori lineari biiettivi, dove X, Y, Z sono spazi vettoriali. Allora l'inverso $(ST)^{-1} : Z \rightarrow X$ del prodotto (composizione) ST esiste e

$$(ST)^{-1} = T^{-1}S^{-1}. \quad (2.13)$$

■

Dimostrazione. La dimostrazione è lasciata al lettore. ■

2.7 Spazi Lineari di Operatori

Consideriamo due spazi vettoriali qualunque X e Y (entrambi reali o complessi) e l'insieme

$$L(X, Y)$$

costituito da tutti gli operatori lineari da X in Y , cioè ciascuno di tali operatori è definito su tutto X e la sua immagine giace in Y . Vogliamo mostrare che $L(X, Y)$ stesso può essere dotato della struttura di spazio vettoriale.

Il tutto è molto semplice. $L(X, Y)$ diviene uno spazio vettoriale se, in maniera del tutto naturale, definiamo la somma $T_1 + T_2$ di due operatori $T_1, T_2 \in L(X, Y)$ come

$$(T_1 + T_2)x = T_1x + T_2x$$

e il prodotto αT di $T \in L(X, Y)$ per uno scalare α come

$$(\alpha T)x = \alpha Tx.$$

2.8 Operatori Lineari Limitati e Continui

Siamo ora interessati a definire una classe particolare di operatori lineari che ammettono norma e che costituiscono quindi essi stessi uno spazio normato.

2.27 DEFINIZIONE (OPERATORI LINEARI LIMITATI)

Siano X e Y spazi normati e $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare, dove $\mathcal{D}(T) \subset X$. L'operatore T è detto *limitato* se esiste un numero reale c tale che per tutti gli $x \in \mathcal{D}(T)$

$$\|Tx\| \leq c\|x\|. \quad (2.14)$$

■

In (2.14) la norma a sinistra è quella di Y e la norma a destra quella di X . Per semplicità abbiamo indicato col medesimo simbolo $\|\cdot\|$ entrambe le norme, senza pericolo di confusione. La formula (2.14) mostra che un operatore limitato applica insiemi limitati in $\mathcal{D}(T)$ in insiemi limitati in Y . Ciò motiva il termine “operatore limitato”.

Attenzione. Si noti che il presente uso della parola “limitato” è differente da quello in analisi, dove una funzione limitata è una funzione la cui immagine è un insieme limitato.

Qual'è il più piccolo c tale che la (2.14) è ancora valida per tutti gli $x \in \mathcal{D}(T)$ che non siano nulli? [Possiamo escludere $x = 0$ perché $Tx = 0$ per $x = 0$.] Dividendo si ottiene

$$\frac{\|Tx\|}{\|x\|} \leq c \quad (x \neq 0)$$

e ciò mostra che c deve essere almeno altrettanto grande che l'estremo superiore dell'espressione a sinistra considerata su $\mathcal{D}(T) - \{0\}$. Quindi il minimo possibile c nella (2.14) è questo estremo superiore. Questa quantità è indicata con $\|T\|$; così

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}. \quad (2.15)$$

$\|T\|$ è chiamato la **norma** dell'operatore T . Se $\mathcal{D}(T) = \{0\}$ definiamo $\|T\| = 0$; in questo caso (relativamente ininteressante) $T = 0$ perché $T0 = 0$.

Si noti che la (2.14) con $c = \|T\|$ diventa

$$\|Tx\| \leq \|T\|\|x\|. \quad (2.16)$$

Applicheremo questa formula piuttosto frequentemente.

Naturalmente dovremmo giustificare l'uso del termine “norma” nel presente contesto. Questo viene fatto nel seguente lemma.

2.28 Lemma (Norma)

Sia T un operatore lineare limitato. Allora

(a) *Una formula alternativa per la norma di T è*

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\|=1}} \|Tx\|. \quad (2.17)$$

(b) *La norma definita dalla (2.15) soddisfa (N1) sino a (N4).*

■

Dimostrazione. (a) Utilizzando la proprietà (N3) della norma in Y e la linearità di T otteniamo dalla (2.15)

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{1}{\|x\|} \|Tx\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \left\| T \left(\frac{1}{\|x\|} x \right) \right\| = \sup_{\substack{y \in \mathcal{D}(T) \\ \|y\|=1}} \|Ty\|.$$

Scrivendo x invece di y a destra abbiamo la (2.17).

(b) (N1) è ovvio e così $\|0\| = 0$. Da $\|T\| = 0$ abbiamo che $Tx = 0$ per tutti gli $x \in \mathcal{D}(T)$, così che $T = 0$. Quindi (N2) vale. Inoltre (N3) è ottenuto da

$$\sup_{\|x\|=1} \|\alpha T\| = \sup_{\|x\|=1} |\alpha| \|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\|$$

dove $x \in \mathcal{D}(T)$. Infine (N4) segue da

$$\sup_{\|x\|=1} \|(T_1 + T_2)x\| = \sup_{\|x\|=1} \|T_1x + T_2x\| \leq \sup_{\|x\|=1} \|T_1x\| + \sup_{\|x\|=1} \|T_2x\|;$$

qui $x \in \mathcal{D}(T)$. ■

Si noti che l'operatore identità $I : X \rightarrow X$ e l'operatore zero $0 : X \rightarrow X$ su uno spazio normato X sono operatori limitati ed hanno rispettivamente norma $\|I\| = 1$ e $\|0\| = 0$.

Osservazione. Nel caso in cui T sia una matrice $n \times n$ di elementi (τ_{jk}) e la norma nello spazio vettoriale di definizione sia quella euclidea la sua norma risulta essere $\|T\|^2 = \max_k \sum_{j=1}^n |\tau_{jk}|^2$. Se ne lascia per esercizio la dimostrazione al lettore.

Dal punto (b) del Lemma 2.28 otteniamo immediatamente il risultato cercato.

2.29 Teorema (Spazio $B(X, Y)$)

Lo spazio vettoriale $B(X, Y)$ di tutti gli operatori limitati lineari da uno spazio normato X in uno spazio normato Y è esso stesso uno spazio normato con norma definita da

$$\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx\|. \quad (2.18) \quad \blacksquare$$

Esaminiamo ora alcune proprietà specifiche importanti degli operatori lineari limitati.

2.30 Teorema (Dimensioni Finite)

Se uno spazio normato X è finito dimensionale allora ogni operatore lineare su X è limitato. ■

Dimostrazione. Sia $\dim X = n$ e $\{e_1, \dots, e_n\}$ una base per X . Prendiamo un qualunque $x = \sum \xi_j e_j$ e consideriamo un qualunque operatore lineare T su X . Poiché T è lineare

$$\|Tx\| = \left\| \sum \xi_j T e_j \right\| \leq \sum |\xi_j| \|T e_j\| \leq \max_k \|T e_k\| \sum |\xi_j|$$

(somme da 1 a n). All'ultima somma applichiamo il Lemma 2.9 con $\alpha_j = \xi_j$ e $x_j = e_j$. Allora otteniamo

$$\sum |\xi_j| \leq \frac{1}{c} \left\| \sum \xi_j e_j \right\| = \frac{1}{c} \|x\|.$$

Assieme danno

$$\|Tx\| \leq \gamma \|x\| \quad \text{dove} \quad \gamma = \frac{1}{c} \max_k \|T e_k\|.$$

Da ciò e dalla (2.14) vediamo che T è limitato. ■

Consideriamo ora alcune importanti proprietà degli operatori lineari limitati.

Gli operatori sono applicazioni, così che ad essi si applica la definizione di continuità. È un fatto fondamentale che per gli operatori *lineari* continuità e limitatezza divengono concetti equivalenti. I dettagli sono i seguenti.

Sia $T : \mathcal{D}(T) \rightarrow Y$ un operatore qualunque non necessariamente lineare, dove $\mathcal{D}(T) \subset X$ e X e Y sono spazi normati. Per la definizione 1.4 l'operatore T è *continuo in un* $x_0 \in \mathcal{D}(T)$ se per ogni $\varepsilon > 0$ v'è un $\delta > 0$ tale che

$$\|Tx - Tx_0\| < \varepsilon \quad \text{per tutti gli } x \in \mathcal{D}(T) \text{ per cui } \|x - x_0\| < \delta.$$

T è *continuo* se T è continuo in ogni $x \in \mathcal{D}(T)$.

Ora se T è lineare abbiamo il rimarchevole teorema seguente.

2.31 Teorema (Continuità e limitatezza)

Sia $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare, dove $\mathcal{D}(T) \subset X$ e siano X e Y spazi normati. Allora

(a) *T è continuo se e solamente se T è limitato.*

(b) *Se T è continuo in un singolo punto allora è continuo.* ■

Dimostrazione. **(a)** Assumiamo che T sia limitato. Per $T = 0$ l'affermazione è banale. Sia $T \neq 0$. Allora $\|T\| \neq 0$. Consideriamo un qualunque $x_0 \in \mathcal{D}(T)$. Sia dato un $\varepsilon > 0$ arbitrario. Allora poiché T è lineare per ogni $x \in \mathcal{D}(T)$ tale che

$$\|x - x_0\| < \delta \quad \text{dove} \quad \delta = \frac{\varepsilon}{\|T\|}$$

otteniamo

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\| \|x - x_0\| < \|T\| \delta = \varepsilon.$$

Poiché $x_0 \in \mathcal{D}(T)$ era arbitrario ciò mostra che T è continuo.

Viceversa assumiamo che T sia continuo in un arbitrario $x_0 \in \mathcal{D}(T)$. Allora dato un $\varepsilon > 0$ arbitrario v'è un $\delta > 0$ tale che

$$\|Tx - Tx_0\| < \varepsilon \quad \text{per tutti gli } x \in \mathcal{D}(T) \text{ per cui } \|x - x_0\| < \delta. \quad (2.19)$$

Prendiamo ora un qualunque $y \neq 0$ in $\mathcal{D}(T)$ e poniamo

$$x = x_0 + \frac{\delta}{2\|y\|}y. \quad \text{Allora} \quad x - x_0 = \frac{\delta}{2\|y\|}y.$$

Quindi $\|x - x_0\| = \delta/2$ così che possiamo usare la (2.19). Poiché T è lineare abbiamo

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\| T \left(\frac{\delta}{\|y\|}y \right) \right\| = \frac{\delta}{\|y\|} \|Ty\|$$

e (2.19) implica

$$\frac{\delta}{\|y\|} \|Ty\| < \varepsilon. \quad \text{Così} \quad \|Ty\| < \frac{\varepsilon}{\delta} \|y\|.$$

Ciò può essere scritto $\|Ty\| \leq c\|y\|$, dove $c = \varepsilon/\delta$. Quindi poiché c dipende da x_0 e non da y ne segue che T è limitato.

(b) La continuità di T in un punto implica la limitatezza di T per la seconda parte della dimostrazione di (a), che a sua volta implica la continuità di T per l'(a). ■

2.32 Corollario (Continuità, Spazio Nullo)

Sia T un operatore lineare limitato. Allora

(a) $x_n \rightarrow x$, dove $x_n, x \in \mathcal{D}(T)$, implica $Tx_n \rightarrow Tx$.

(b) Lo spazio nullo $\mathcal{N}(T)$ è chiuso. ■

Dimostrazione. (a) segue dal Teorema 2.31(a) e 1.14 o direttamente dalla (2.16) perché per $n \rightarrow \infty$

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq \|T\| \|x_n - x\| \rightarrow 0.$$

(b) Per ogni $x \in \overline{\mathcal{N}(T)}$ v'è una successione (x_n) in $\mathcal{N}(T)$ tale che $x_n \rightarrow x$; cf. 1.12(a). Quindi $Tx_n \rightarrow Tx$ per la parte (a) di questo corollario. Anche $Tx = 0$ poiché $Tx_n = 0$ così che $x \in \mathcal{N}(T)$. Poiché $x \in \overline{\mathcal{N}(T)}$ era arbitrario, $\mathcal{N}(T)$ è chiuso. ■

È lasciata al lettore la semplice prova di un'altra utile formula

$$\|T_1 T_2\| \leq \|T_1\| \|T_2\|, \quad \|T^n\| \leq \|T\|^n \quad (n \in \mathbb{N}) \quad (2.20)$$

valida per operatori lineari limitati $T_2 : X \rightarrow Y$, $T_1 : Y \rightarrow Z$ e $T : X \rightarrow X$, dove X, Y, Z sono spazi normati.

Due operatori T_1 e T_2 sono definiti **uguali**, scrivendo

$$T_1 = T_2,$$

se hanno medesimo dominio $\mathcal{D}(T_1) = \mathcal{D}(T_2)$ e se $T_1 x = T_2 x$ per tutti gli $x \in \mathcal{D}(T_1) = \mathcal{D}(T_2)$.

La **restrizione** di un operatore $T : \mathcal{D}(T) \rightarrow Y$ ad un sottoinsieme $B \subset \mathcal{D}(T)$ è indicato con

$$T|_B$$

ed è l'operatore definito da

$$T|_B : B \rightarrow Y, \quad T|_B x = Tx \text{ per tutti gli } x \in B.$$

Un'**estensione** di T ad un insieme $M \supset \mathcal{D}(T)$ è un operatore

$$\tilde{T} : M \rightarrow Y \quad \text{tale che} \quad \tilde{T}|_{\mathcal{D}(T)} = T.$$

Se $\mathcal{D}(T)$ è un sottoinsieme proprio di M , allora un dato T ha molte estensioni. Di interesse pratico sono quelle estensioni che conservano alcune proprietà basilari, per esempio la linearità (se T è lineare) o la limitatezza (se $\mathcal{D}(T)$ giace in uno spazio normato e T è limitato). Il seguente importante teorema è tipico a questo riguardo. Concerne l'estensione di un operatore lineare limitato T alla chiusura $\overline{\mathcal{D}(T)}$ del dominio tale che l'operatore esteso sia nuovamente limitato e lineare e abbia anche la stessa norma. Ciò include il caso dell'estensione da un insieme denso in uno spazio normato X a tutto X . Include anche il caso dell'estensione da uno spazio normato X al suo completamento.

2.33 Teorema (Estensione Limitata Lineare)

Sia

$$T : \mathcal{D}(T) \rightarrow Y$$

un operatore limitato lineare, dove $\mathcal{D}(T)$ giace in uno spazio normato X ed Y è uno spazio di Banach. Allora T ha un'unica estensione continua alla chiusura di $\mathcal{D}(T)$

$$\tilde{T} : \overline{\mathcal{D}(T)} \rightarrow Y.$$

Inoltre l'estensione $\tilde{T}$ è un operatore limitato lineare di norma

$$\|\tilde{T}\| = \|T\|. \quad \blacksquare$$

Dimostrazione. Consideriamo un qualunque $x \in \overline{\mathcal{D}(T)}$. Per il Teorema 1.12(a) v'è una successione (x_n) in $\mathcal{D}(T)$ tale che $x_n \rightarrow x$. Poiché T è lineare e limitato abbiamo che

$$\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| \leq \|T\| \|x_n - x_m\|.$$

Ciò mostra che (Tx_n) è di Cauchy perché (x_n) converge. Per ipotesi Y è completo così che (Tx_n) converge, ossia

$$Tx_n \rightarrow y \in Y.$$

Quindi se l'estensione $\tilde{T}$ esiste deve essere

$$\tilde{T}x = y.$$

Mostriamo che questa definizione non è ambigua, è cioè indipendente dalla particolare successione scelta in $\mathcal{D}(T)$ convergente a x . Supponiamo che $x_n \rightarrow x$ e $z_n \rightarrow x$. Allora $x_n - z_n \rightarrow 0$. Poiché T è lineare e limitato abbiamo che

$$\|Tx_n - Tz_n\| = \|T(x_n - z_n)\| \leq \|T\| \|x_n - z_n\|$$

e le due successioni (Tx_n) e (Tz_n) hanno il medesimo limite. Ciò prova che $\tilde{T}$ è univocamente definito per ogni $x \in \overline{\mathcal{D}(T)}$.

Chiaramente $\tilde{T}$ è lineare e $\tilde{T}x = Tx$ per ogni $x \in \mathcal{D}(T)$, così che $\tilde{T}$ è un'estensione di T . Ora usiamo

$$\|Tx_n\| \leq \|T\| \|x_n\|$$

e lasciamo $n \rightarrow \infty$. Allora $Tx_n \rightarrow y = \tilde{T}x$. Poiché $x \mapsto \|x\|$ definisce un'applicazione continua otteniamo che

$$\|\tilde{T}x\| \leq \|T\| \|x\|.$$

Quindi $\tilde{T}$ è limitato e $\|\tilde{T}\| \leq \|T\|$. Naturalmente $\|\tilde{T}\| \geq \|T\|$ perché la norma essendo definita mediante un estremo superiore non può decrescere in un'estensione. Assieme danno $\|\tilde{T}\| = \|T\|$. $\blacksquare$

2.9 Funzionali Lineari

Un **funzionale** è un operatore la cui immagine giace sulla linea reale $\mathbb{R}$ o nel piano complesso $\mathbb{C}$. Inizialmente l'*analisi funzionale* era l'analisi dei funzionali. Questi ultimi appaiono così frequentemente che viene usata una notazione specifica. Indichiamo i funzionali con le lettere minuscole $f, g, h, \dots$, il dominio di f con $\mathcal{D}(f)$, l'immagine con $\mathcal{R}(f)$ ed il valore di f in x con $f(x)$, con le parentesi.

I funzionali sono operatori cosicché si applicano le definizioni precedenti. Avremo in particolare bisogno delle seguenti due definizioni perché la maggioranza dei funzionali che considereremo saranno lineari e limitati.

2.34 DEFINIZIONE (FUNZIONALE LINEARE)

Un *funzionale lineare* f è un operatore lineare con dominio in uno spazio vettoriale X e immagine nel campo scalare K di X ; così

$$f : \mathcal{D}(f) \rightarrow K,$$

dove $K = \mathbb{R}$ se X è reale e $K = \mathbb{C}$ se X è complesso. ■

2.35 DEFINIZIONE (FUNZIONALI LIMITATI LINEARI)

Un *funzionale limitato lineare* f è un operatore limitato lineare (cf. Def. 2.27) con immagine nel campo scalare dello spazio normato X in cui giace il dominio $\mathcal{D}(f)$. Perciò esiste un numero reale c tale che per tutti gli $x \in \mathcal{D}(f)$

$$|f(x)| \leq c\|x\|. \quad (2.21)$$

Inoltre la *norma* di f è (cf. (2.15) nella Sez. 2.8)

$$\|f\| = \sup_{\substack{x \in \mathcal{D}(f) \\ x \neq 0}} \frac{|f(x)|}{\|x\|} \quad (2.22)$$

o

$$\|f\| = \sup_{\substack{x \in \mathcal{D}(f) \\ \|x\|=1}} |f(x)|. \quad (2.23)$$

■

La formula (2.16) nella Sez. 2.8 ora implica che

$$|f(x)| \leq \|f\| \|x\|, \quad (2.24)$$

ed il Teorema 2.31 viene riformulato nel modo seguente.

2.36 Teorema (Continuità e Limitatezza)

Un *funzionale lineare* f con dominio $\mathcal{D}(f)$ in uno spazio normato è continuo se e solo se f è limitato. ■

Il lemma seguente valido sia nel caso finito ed che infinito dimensionale trova utile applicazione nel seguito. Un lemma simile per spazi normati arbitrari sarà dato più in là, nella sezione 4.12.

2.37 Lemma (Vettore Nullo)

Sia X uno spazio vettoriale. Se $x_0 \in X$ ha la proprietà che $f(x_0) = 0$ per tutti i funzionali lineari f su X allora $x_0 = 0$. ■

Dimostrazione. Supponiamo per assurdo che $x_0 \neq 0$. Allora se X è finito dimensionale x_0 si può esprimere come combinazione lineare dei vettori di base, ossia $x_0 = \sum_{j=1}^n \xi_j e_j$, dove gli scalari ξ_j non sono tutti nulli. Sia ad esempio $\xi_1 \neq 0$. Se X è infinito dimensionale è necessario ammettere che esso sia dotato di una base di Hamel B , ciò che dimostreremo solamente nel seguito utilizzando il Lemma di Zorn. Debbono esistere quindi vettori indipendenti in numero finito, che chiameremo ancora e_j , appartenenti a B che generano x_0 e possiamo ripetere l'argomentazione precedente. Consideriamo ora il funzionale lineare f_1 che sui vettori della base B assume i valori

$$\begin{aligned} f_1(e_1) &= 1 \\ f_1(b) &= 0 \quad \text{per ogni } b \in B, \quad b \neq e_1. \end{aligned}$$

Esso viene esteso a tutti i vettori $x \in X$ utilizzando la linearità ed il fatto che ogni x può essere espresso come combinazione lineare di un numero finito di vettori della base B . È ora

$$f_1(x_0) = \xi_1 \neq 0,$$

ma ciò contraddice l'ipotesi e quindi dall'assurdo segue che $x_0 = 0$. ■

È di basilare importanza che l'insieme di tutti i funzionali lineari definiti su uno spazio vettoriale X possa esso stesso essere fatto spazio vettoriale. Questo spazio è denotato con X^* ed è chiamato **spazio duale algebrico**¹ di X . Le operazioni algebriche di spazio vettoriale sono definite in maniera naturale nel modo seguente. La *somma* $f_1 + f_2$ di due funzionali f_1 e f_2 è il funzionale s il cui valore ad ogni $x \in X$ è

$$s(x) = (f_1 + f_2)(x) = f_1(x) + f_2(x);$$

il *prodotto* αf di uno scalare α e di un funzionale f è il funzionale p il cui valore in $x \in X$ è

$$p(x) = (\alpha f)(x) = \alpha f(x).$$

Si noti che ciò concorda con il modo usuale di sommare funzioni e di moltiplicarle per costanti.

Possiamo fare ancora un passo innanzi e considerare il duale algebrico $(X^*)^*$ di X^* , i cui elementi sono i funzionali lineari definiti su X^* . Indichiamo $(X^*)^*$ con X^{**} e lo chiamiamo lo **spazio biduale algebrico** di X .

Perché consideriamo X^{**} ? Il punto è che possiamo ottenere una relazione interessante ed importante fra X e X^{**} . Scegliamo la notazione

Spazio	Generico elemento	Valore in un punto
X	x	—
X^*	f	$f(x)$
X^{**}	g	$g(f)$

Possiamo ottenere un $g \in X^{**}$, che è un funzionale lineare definito su X^* , scegliendo un $x \in X$ *fisso* e ponendo

$$g(f) = g_x(f) = f(x) \quad (x \in X \text{ fisso, } f \in X^* \text{ variabile}). \quad (2.25)$$

L'indice x serve a ricordare che abbiamo ottenuto g coll'uso di un certo $x \in X$. Il lettore deve tenere ben presente che qui f è la variabile mentre x è fisso. Tenendo ciò in mente non dovrebbe avere difficoltà a capire la nostra presente considerazione.

g_x come definito dalla (2.25) è lineare. Ciò può essere visto dalla

$$g_x(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha g_x(f_1) + \beta g_x(f_2).$$

Quindi g_x è un elemento di X^{**} , per definizione di X^{**} .

A ciascun $x \in X$ corrisponde un $g_x \in X^{**}$. Ciò definisce un'applicazione

$$C : X \rightarrow X^{**}$$

$$x \longmapsto g_x.$$

C è chiamata l'**applicazione canonica** di X in X^{**} .

¹Si noti che questa definizione non involve una norma. Il cosiddetto *spazio duale* X' consistente di tutti i funzionali *limitati* lineari su X sarà considerato nella Sez. 2.11.

C è lineare perché il suo dominio è uno spazio vettoriale ed abbiamo

$$\begin{aligned}(C(\alpha x + \beta y))(f) &= g_{\alpha x + \beta y}(f) \\ &= f(\alpha x + \beta y) \\ &= \alpha f(x) + \beta f(y) \\ &= \alpha g_x(f) + \beta g_y(f) \\ &= \alpha(Cx)(f) + \beta(Cy)(f).\end{aligned}$$

C è anche iniettiva. Infatti se $Cx_0 = 0$ abbiamo che per tutti gli $f \in X^*$

$$(Cx_0)(f) = g_{x_0}(f) = f(x_0) = 0.$$

Ciò implica $x_0 = 0$ per il Lemma precedente e quindi dal Teorema 2.25 segue la iniettività della C .

C è chiamato l'*immersione (embedding) canonica* di X in X^{**} . Per comprendere e motivare questo termine spieghiamo dapprima il concetto di "isomorfismo", che è di interesse generale.

Nel nostro lavoro ci occupiamo di diversi spazi. Comune a tutti loro è il fatto che essi consistono di un insieme, chiamiamolo X , e di una "struttura" definita su X . Per uno spazio metrico questa è una metrica. Per uno spazio vettoriale le due operazioni algebriche formano la struttura. Per uno spazio normato la struttura consiste di queste due operazioni algebriche e della norma.

Dati due spazi X e $\tilde{X}$ dello stesso tipo (ad esempio appunto due spazi metrici o vettoriali o normati) è di interesse sapere quando essi possano essere considerati "essenzialmente identici", cioè quando essi si possano considerare coincidenti per quanto riguarda la loro struttura e differenti al più per la natura dei loro punti, ossia quando essi possano essere considerati due realizzazioni del medesimo oggetto "astratto".

Una risposta matematicamente precisa è data dall'introduzione del concetto di **isomorfismo**. Per definizione si tratta di un'applicazione biettiva di X su $\tilde{X}$ che conserva la struttura.

Corrispondentemente un *isomorfismo* T di uno spazio metrico $X = (X, d)$ su uno spazio metrico $\tilde{X} = (\tilde{X}, \tilde{d})$ è un'applicazione biettiva che conserva la distanza, cioè per tutti gli $x, y \in X$

$$\tilde{d}(Tx, Ty) = d(x, y).$$

$\tilde{X}$ è allora detto *isomorfo* ad X . Ciò non ci è nuovo ma si tratta semplicemente di un altro nome per l'isometria introdotta nella Def. 1.16. Nuovo è il seguito.

Un *isomorfismo* T di uno spazio vettoriale X su uno spazio vettoriale $\tilde{X}$ sul medesimo campo è un'applicazione biettiva che conserva le due operazioni algebriche dello spazio vettoriale; così per tutti gli $x, y \in X$ e scalari α

$$T(x + y) = Tx + Ty, \quad T(\alpha x) = \alpha Tx,$$

cioè $T : X \rightarrow \tilde{X}$ è un operatore biettivo lineare. $\tilde{X}$ è allora detto *isomorfo* ad X ed X e $\tilde{X}$ sono detti *spazi vettoriali isomorfi*.

Isomorfismi fra spazi normati sono isomorfismi fra spazi vettoriali che conservano anche le norme.

Al momento possiamo utilizzare il concetto di isomorfismo fra spazi vettoriali come segue.

Poiché C è lineare e iniettiva è un isomorfismo di X sull'immagine $\mathcal{R}(C) \subset X^{**}$.

Se X è isomorfo ad un sottospazio di uno spazio vettoriale Y diciamo che X è **immergibile** (**embeddable**) in Y . Quindi X è immergibile in X^{**} e C è anche chiamato l'*immersione* (*embedding*) canonica di X in X^{**} .

Se C è surgettiva così che $\mathcal{R}(C) = X^{**}$ allora X è detto **algebricamente riflessivo**. Proveremo nella prossima sezione che se X è finito dimensionale allora X è algebricamente riflessivo.

Una discussione simile che involve le norme e che conduce al concetto di *riflessività* di uno spazio *normato* sarà presentato nella Sez. 4.7 dopo avere sviluppato gli strumenti necessari (in particolare il famoso teorema di Hahn–Banach).

2.10 Operatori Lineari e Funzionali su Spazi Finito Dimensionali

Siano X e Y spazi vettoriali finito dimensionali sul medesimo campo e $T : X \rightarrow Y$ un operatore lineare. Scegliamo una base $E = \{e_1, \dots, e_n\}$ per X ed una base $B = \{b_1, \dots, b_r\}$ per Y con i vettori elencati secondo un ordine definito che manteniamo fisso. Allora ogni $x \in X$ ha un'unica rappresentazione

$$x = \xi_1 e_1 + \dots + \xi_n e_n. \quad (2.26)$$

Poiché T è lineare x ha l'immagine

$$y = Tx = T \left(\sum_{k=1}^n \xi_k e_k \right) = \sum_{k=1}^n \xi_k T e_k. \quad (2.27)$$

Poiché la rappresentazione (2.26) è unica otteniamo il nostro primo risultato.

T è unicamente determinato se le immagini $y_k = T e_k$ degli n vettori di base sono assegnate.

Poiché y e $y_k = T e_k$ sono in Y essi hanno un'unica rappresentazione della forma

$$y = \sum_{j=1}^r \eta_j b_j \quad (2.28)$$

$$T e_k = \sum_{j=1}^r \tau_{jk} b_j. \quad (2.29)$$

Sostituendo in (2.27) si ottiene

$$y = \sum_{j=1}^r \eta_j b_j = \sum_{k=1}^n \xi_k T e_k = \sum_{k=1}^n \xi_k \sum_{j=1}^r \tau_{jk} b_j = \sum_{j=1}^r \left(\sum_{k=1}^n \tau_{jk} \xi_k \right) b_j.$$

Poiché le b_j formano un insieme linearmente indipendente, i coefficienti di ciascun b_j a destra e a sinistra devono essere gli stessi, cioè

$$\eta_j = \sum_{k=1}^n \tau_{jk} \xi_k \quad j = 1, \dots, r. \quad (2.30)$$

Ciò fornisce il nostro risultato successivo.

L'immagine $y = Tx = \sum \eta_j b_j$ di $x = \sum \xi_k e_k$ può essere ottenuto dalla (2.30).

Si noti la posizione inusuale dell'indice di somma j di τ_{jk} in (2.29), che è necessaria per arrivare alla posizione usuale dell'indice di somma nella (2.30).

I coefficienti nella (2.30) formano una matrice

$$T_{EB} = (\tau_{jk})$$

con r righe ed n colonne. Se sono assegnate una base E di X ed una base B di Y con gli elementi dati secondo un ordine definito (che è arbitrario ma fisso), allora la matrice T_{EB} è univocamente determinata dall'operatore lineare T . Diciamo che la matrice T_{EB} **rappresenta** l'operatore T rispetto a queste basi.

Introducendo i vettori colonna $\tilde{x} = (\xi_k)$ e $\tilde{y} = (\eta_j)$ possiamo scrivere la (2.30) in notazione matriciale

$$\tilde{y} = T_{EB}\tilde{x}. \quad (2.31)$$

Analogamente anche (2.29) può essere scritta in forma matriciale

$$Te = T_{EB}^\top b \quad (2.32)$$

dove Te è il vettore colonna con componenti $Te_1, \dots, Te_n$ (che sono essi stessi vettori) e b è il vettore colonna di componenti $b_1, \dots, b_r$ e dove dobbiamo usare il trasposto $T_{EB}^\top$ di T_{EB} perché nella (2.29) sommiamo su j che è il primo indice.

Le nostre considerazioni mostrano che un operatore lineare T determina un'unica matrice rappresentante T rispetto ad una data base per X ed ad una data base per Y , dove i vettori di ciascuna delle due basi sono assunti essere dati secondo un ordine fisso. Viceversa ogni matrice con r righe ed n colonne determina un operatore lineare che essa rappresenta rispetto a basi date per X e Y .

Ritorniamo ora ai **funzionali lineari** su X , dove $\dim X = n$ e $\{e_1, \dots, e_n\}$ è una base per X come prima. Questi funzionali costituiscono lo spazio algebrico duale X^* di X come sappiamo dalla sezione precedente. Per ogni tale funzionale e per ogni $x = \sum \xi_j e_j \in X$ abbiamo

$$f(x) = f\left(\sum_{j=1}^n \xi_j e_j\right) = \sum_{j=1}^n \xi_j f(e_j) = \sum_{j=1}^n \xi_j \alpha_j \quad (2.33)$$

dove

$$\alpha_j = f(e_j) \quad j = 1, \dots, n \quad (2.34)$$

ed f è unicamente determinata dai suoi valori α_j sugli n vettori di base di X .

Viceversa ogni n -pla di scalari $\alpha_1, \dots, \alpha_n$ determina un funzionale lineare su X per le (2.33), (2.34). In particolare prendiamo le n -ple

$$\begin{aligned} &(1, 0, 0, \dots, 0, 0) \\ &(0, 1, 0, \dots, 0, 0) \\ &\vdots \\ &(0, 0, 0, \dots, 0, 1). \end{aligned}$$

Per le (2.33), (2.34) ciò fornisce n funzionali che denotiamo $f_1, \dots, f_n$ e che hanno valori

$$f_k(e_j) = \delta_{jk} = \begin{cases} 0 & \text{se } j \neq k \\ 1 & \text{se } j = k; \end{cases} \quad (2.35)$$

cioè f_k ha il valore 1 al k -mo vettore di base e il valore 0 agli altri $n - 1$ vettori di base. δ_{jk} è chiamato la *delta di Kroneker*. $\{f_1, \dots, f_n\}$ è chiamato la **base duale** della base $\{e_1, \dots, e_n\}$ per X . Ciò è giustificato dal seguente teorema.

2.38 Teorema (Dimensioni di X^*)

Sia X uno spazio vettoriale n -dimensionale ed $E = \{e_1, \dots, e_n\}$ una base per X . Allora $F = \{f_1, \dots, f_n\}$ data dalla (2.35) è una base per l'algebrico duale X^* di X e $\dim X^* = \dim X = n$. ■

Dimostrazione. F è un insieme linearmente indipendente perché

$$\sum_{k=1}^n \beta_k f_k(x) = 0 \quad (x \in X) \quad (2.36)$$

per $x = e_j$ dà

$$\sum_{k=1}^n \beta_k f_k(e_j) = \sum_{k=1}^n \beta_k \delta_{jk} = \beta_j = 0,$$

così che tutte le β_k in (2.36) sono zero. Mostriamo che ogni $f \in X^*$ può essere rappresentata come una combinazione lineare degli elementi di F in una maniera unica. Scriviamo come in (2.34) $f(e_j) = \alpha_j$. Per la (2.33)

$$f(x) = \sum_{j=1}^n \xi_j \alpha_j$$

per ogni $x \in X$. D'altra parte per la (2.35) otteniamo che

$$f_j(x) = f_j(\xi_1 e_1 + \dots + \xi_n e_n) = \xi_j.$$

Dal confronto

$$f(x) = \sum_{j=1}^n \alpha_j f_j(x).$$

Quindi la rappresentazione unica di un arbitrario funzionale lineare f su X in termini dei funzionali $f_1, \dots, f_n$ è

$$f = \alpha_1 f_1 + \dots + \alpha_n f_n. \quad \blacksquare$$

Usando questo teorema ed il lemma 2.37 possiamo ottenere il seguente teorema.

2.39 Teorema (Riflessività Algebrica)

Ogni spazio vettoriale finito dimensionale è algebricamente riflessivo. ■

Dimostrazione. L'applicazione canonica $C : X \rightarrow X^{**}$ considerata nella sezione precedente è lineare e iniettiva. Quindi dal Teorema 2.25 segue che l'applicazione C ha un inverso lineare $C^{-1} : \mathcal{R}(C) \rightarrow X$, dove $\mathcal{R}(C)$ è l'immagine di C . Abbiamo anche per il medesimo teorema che $\dim \mathcal{R}(C) = \dim X$. Ora per il Teorema 2.38

$$\dim X^{**} = \dim X^* = \dim X.$$

Confrontando, $\dim \mathcal{R}(C) = \dim X^{**}$. Quindi $\mathcal{R}(C) = X^{**}$ perché, se $\mathcal{R}(C)$ fosse un sottospazio proprio di X^{**} , per il Teorema 2.4 avrebbe dimensioni minori di X^{**} , che è impossibile. Ciò prova la riflessività algebrica. ■

2.11 Spazi Normati di Operatori. Spazio Duale

In che caso lo spazio normato $B(X, Y)$ delle applicazioni lineari limitate di X in Y è uno spazio di Banach? Questa è una domanda centrale a cui si risponde nel seguente teorema. È rimarchevole che le ipotesi del teorema non coinvolgano X ; ossia X può essere o non essere completo.

2.40 Teorema (Completezza)

Se Y è uno spazio di Banach allora $B(X, Y)$ è uno spazio di Banach. ■

Dimostrazione. Consideriamo una successione arbitraria di Cauchy (T_n) in $B(X, Y)$ e mostriamo che (T_n) converge ad un operatore $T \in B(X, Y)$. Poiché (T_n) è di Cauchy, per ogni $\varepsilon > 0$ v'è un N tale che

$$\|T_n - T_m\| < \varepsilon \quad (n, m > N).$$

Per tutti gli $x \in X$ ed $n, m > N$ otteniamo così [cf. (2.16) nella Sez. 2.8]

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\| \|x\| < \varepsilon \|x\|. \quad (2.37)$$

Ora per ogni x fisso e per un dato $\tilde{\varepsilon}$ possiamo scegliere un $\varepsilon = \varepsilon_x$ tale che $\varepsilon_x \|x\| < \tilde{\varepsilon}$. Allora dalla (2.37) abbiamo $\|T_n x - T_m x\| < \tilde{\varepsilon}$ e vediamo che $(T_n x)$ è di Cauchy in Y . Poiché Y è completo $(T_n x)$ converge, ossia $T_n x \rightarrow y$. Chiaramente il limite $y \in Y$ dipende dalla scelta di $x \in X$. Ciò definisce un operatore $T : X \rightarrow Y$, dove $y = Tx$. L'operatore T è lineare perché

$$\lim T_n(\alpha x + \beta z) = \lim(\alpha T_n x + \beta T_n z) = \alpha \lim T_n x + \beta \lim T_n z.$$

Proviamo che T è limitato e che $T_n \rightarrow T$, ossia che $\|T_n - T\| \rightarrow 0$.

Poiché la (2.37) vale per ogni $m > N$ e $T_m x \rightarrow Tx$ possiamo fare $m \rightarrow \infty$. Usando la continuità della norma allora otteniamo dalla (2.37) per ogni $n > N$ e per tutti gli $x \in X$

$$\|T_n x - Tx\| = \|T_n x - \lim_{m \rightarrow \infty} T_m x\| = \lim_{m \rightarrow \infty} \|T_n x - T_m x\| \leq \varepsilon \|x\|. \quad (2.38)$$

Ciò mostra che $(T_n - T)$ con $n > N$ è un operatore limitato. Poiché T_n è limitato $T = T_n - (T_n - T)$ è limitato, ossia $T \in B(X, Y)$. Inoltre se in (2.38) prendiamo l'estremo superiore per tutti gli x di norma 1 otteniamo

$$\|T_n - T\| \leq \varepsilon \quad (n > N).$$

Quindi $\|T_n - T\| \rightarrow 0$. ■

Questo teorema ha importanti conseguenze rispetto allo spazio duale X' di X , che è definito come segue.

2.41 DEFINIZIONE (SPAZIO DUALE X')

Sia X uno spazio normato. Allora l'insieme di tutti i funzionali lineari limitati su X costituisce uno spazio normato con norma definita da

$$\|f\| = \sup_{\substack{x \in X \\ \|x\|=1}} |f(x)| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|} \quad (2.39)$$

[cf. (2.22) e (2.23) nella Sez. 2.9] che è chiamato la *spazio duale*² di X ed è indicato con X' . ■

²Altri termini sono *duale*, *spazio aggiunto* e *spazio coniugato*. Si ricordi dalla Sez. 2.9 che lo spazio duale *algebrico* X^* è lo spazio vettoriale di *tutti* i funzionali lineari su X .

Poiché un funzionale lineare su X applica X in $\mathbb{R}$ o $\mathbb{C}$ (i campi scalari di X) e poiché $\mathbb{R}$ e $\mathbb{C}$ presi colla metrica usuale sono completi vediamo che X' è $B(X, Y)$ con lo spazio completo $Y = \mathbb{R}$ o $\mathbb{C}$. Quindi il Teorema 2.40 è applicabile ed implica il basilare teorema seguente.

2.42 Teorema (Spazio Duale)

Lo spazio duale X' di uno spazio normato X è uno spazio di Banach (lo sia o no X). ■

Costituisce un principio fondamentale dell'analisi funzionale che lo studio degli spazi sia spesso combinato con quello dei loro duali.

In particolare, ricordando la discussione sullo spazio algebrico biduale X^{**} nella Sez. 2.9, ci possiamo chiedere se sia utile considerare $X'' = (X')'$ lo spazio biduale di X . La risposta è positiva ma dobbiamo posporre questa discussione alla Sez. 4.7 dove sviluppiamo gli strumenti necessari per ottenere dei risultati sostanziosi in questa direzione.

2.12 Problemi

1. Mostrare che nello spazio delle n -ple $x = (\xi_1, \xi_2, \dots, \xi_n)$ reali o complesse possibili norme sono

$$\begin{aligned} \|x\|_p &= (|\xi_1|^p + |\xi_2|^p + \dots + |\xi_n|^p)^{1/p} & (1 \leq p < +\infty) \\ \|x\|_\infty &= \max(|\xi_1| + |\xi_2| + \dots + |\xi_n|). \end{aligned}$$

2. Mostrare che lo spazio ℓ^p ($1 \leq p < +\infty$) delle n -ple infinite $x = (\xi_1, \xi_2, \dots, \xi_n, \dots)$, per le quali $\sum_{j=1}^{\infty} |\xi_j|^p < \infty$, ammette norma

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p}.$$

Suggerimento: Conviene considerare il caso $p = 1$ a parte e per $p \neq 1$ introdurre q l'esponente coniugato di p definito da

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

Per dimostrare la disuguaglianza triangolare conviene dapprima dimostrare la disuguaglianza ausiliaria

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$$

valida per $\alpha \geq 0$, $\beta \geq 0$ qualunque ed utilizzarla per dimostrare la disuguaglianza di Hölder

$$\sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} \left(\sum_{k=1}^{\infty} |\eta_k|^q \right)^{1/q}.$$

Successivamente si dimostra la disuguaglianza di Minkowski

$$\sum_{j=1}^{\infty} |\xi_j + \eta_j|^{1/p} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} + \left(\sum_{k=1}^{\infty} |\eta_k|^p \right)^{1/p}$$

e quindi la disuguaglianza triangolare.

3. Mostrare che lo spazio ℓ^∞ delle n -ple infinite $x = (\xi_1, \xi_2, \dots, \xi_n, \dots)$, per le quali $\|x\| = \sup_{j \in \mathbb{N}} |\xi_j| < \infty$, ammette norma

$$\|x\| = \sup_{j \in \mathbb{N}} |\xi_j|.$$

4. Mostrare che lo spazio $C[a, b]$ delle funzioni $u(t)$ continue nell'intervallo $[a, b]$ ammette norma

$$\|u\| = \max_{t \in [a, b]} |u(t)|.$$

5. Mostrare che gli spazi ℓ^p ed ℓ^∞ sono spazi di Banach.

Suggerimento: Mostrare che le componenti j -sime $\xi_j^{(n)}$ di una successione di Cauchy $x^{(n)}$ costituiscono una successione di Cauchy, che converge alla componente j -sima del limite cercato della successione $x^{(n)}$.

6. Mostrare che ℓ^∞ è uno spazio di Banach.

7. Mostrare che lo spazio ℓ^∞ non è separabile.

Suggerimento: Si consideri il sottoinsieme M di ℓ^∞ costituito dagli elementi $y = (\eta_1, \eta_2, \dots, \eta_n, \dots)$ con $\eta_j = 0$ o $\eta_j = 1$. Si osservi che la distanza fra due qualunque elementi del sottoinsieme è 1 e che mediante la formula

$$\hat{y} = \frac{\eta_1}{2} + \frac{\eta_2}{2^2} + \dots + \frac{\eta_n}{2^n} + \dots$$

si stabilisce una corrispondenza surgettiva fra M ed i numeri reali $\hat{y}$ dell'intervallo $[0, 1]$. Poiché quindi l'insieme M non è numerabile ...

8. Mostrare che lo spazio ℓ^p è separabile.

Suggerimento: Mostrare che lo spazio numerabile M di ℓ^p costituito dagli elementi $y = (\eta_1, \eta_2, \dots, \eta_n, 0, 0, \dots)$ con n qualunque e η_j razionale o complesso razionale è denso in ℓ^p .

9. Mostrare che la chiusura $\overline{Y}$ di un sottospazio vettoriale Y di uno spazio normato X è ancora un sottospazio vettoriale.

10. Se $\dim Y < \infty$ nel Lemma di Riesz mostrare che si può anche scegliere $\theta = 1$.

11. Sia X lo spazio dei polinomi $x(t)$ in $J = [0, 1]$ con norma

$$\|x\| = \max_{t \in J} |x(t)|$$

e sia $T : X \rightarrow X$ l'operatore differenziazione $Tx(t) = x'(t)$. Mostrare che T è lineare non è limitato.

Suggerimento: Considerare l'azione di T sui polinomi $x_n(t) = t^n$.

12. Si consideri l'operatore integrale $T : x(t) \in C[0, 1] \mapsto y(t) \in C[0, 1]$ definito da

$$y(t) = \int_0^1 k(t, \tau) x(\tau) d\tau$$

con $k(t, \tau)$ continua nel quadrato $J \times J$. Mostrare che T è lineare e limitato e che $\|T\| \leq k_0 \|x\|$ con $k_0 = \max_{(t, \tau) \in J \times J} |k(t, \tau)|$.

13. Sia $X = C[a, b]$. Mostrare che

$$f(x) = \int_a^b x(t) dt$$

è un funzionale lineare limitato e che $\|f\| = (b - a)$.

14. Sia $X = C[a, b]$. Mostrare che

$$f(x) = x(t_0),$$

dove t_0 è un punto fisso di $[a, b]$, è un funzionale lineare limitato e che $\|f\| = 1$.

Capitolo 3

Spazi con Prodotto Scalare. Spazi di Hilbert

Gli spazi con prodotto scalare sono, come vedremo, degli speciali spazi normati. Storicamente sono più vecchi degli spazi normati generali. La loro teoria è più ricca e conserva molti degli aspetti dello spazio euclideo, un concetto centrale essendo l'ortogonalità. In effetti gli spazi con prodotto scalare sono la generalizzazione più naturale dello spazio euclideo.

Questi spazi sono stati sino ad ora gli spazi più utili nelle applicazioni pratiche dell'analisi funzionale.

3.1 Breve Orientamento sul Principale Contenuto della Teoria

Uno *spazio con prodotto scalare* X è uno spazio vettoriale dotato di *prodotto scalare* $\langle x, y \rangle$. Quest'ultimo generalizza il prodotto scalare di vettori nello spazio tridimensionale ed è usato per definire

(I) una *norma* $\| \cdot \|$ con $\|x\| = \langle x, x \rangle^{1/2}$.

(II) l'ortogonalità con $\langle x, y \rangle = 0$.

Uno *spazio di Hilbert* H è uno spazio con prodotto scalare completo.

La teoria degli spazi con prodotto scalare e degli spazi di Hilbert è più ricca di quella degli spazi normati generali e degli spazi di Banach. Ci occuperemo

(i) della rappresentazione di H come somma diretta di un sottospazio chiuso e del suo *complemento ortogonale*,

(ii) degli *insiemi e successioni ortogonali* e corrispondenti rappresentazioni degli elementi di H ,

(iii) della *rappresentazione di Riesz* dei funzionali limitati lineari mediante il prodotto scalare,

(iv) dell'*operatore aggiunto di Hilbert* T^* di un operatore limitato lineare.

Vedremo che gli insiemi e le successioni ortogonali sono veramente interessanti solo se sono totali e che gli operatori aggiunti di Hilbert possono essere usati per definire classi di operatori (*autoaggiunti, unitari, normali*) che sono di grande importanza nelle applicazioni.

3.2 Spazi con Prodotto Scalare, Spazio di Hilbert

3.1 DEFINIZIONE (SPAZI CON PRODOTTO SCALARE)

Uno *spazio con prodotto scalare* (o *pre-hilbertiano*) è uno spazio vettoriale X dotato di prodotto scalare. Un **prodotto scalare** su X è un'applicazione che associa ad ogni coppia di vettori x e y uno scalare che viene scritto $\langle x, y \rangle$ ed è chiamato *prodotto scalare* (o *prodotto interno*) di x e y , tale che per tutti i vettori x, y, z e scalari α si ha

$$(IP1) \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

$$(IP2) \quad \langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

$$(IP3) \quad \langle x, y \rangle = \overline{\langle y, x \rangle}$$

$$(IP4) \quad \begin{aligned} \langle x, x \rangle &\geq 0 \\ \langle x, x \rangle &= 0 \iff x = 0. \end{aligned} \quad \blacksquare$$

In (IP3) la barra indica il complesso coniugato. Di conseguenza se X è uno spazio vettoriale *reale* abbiamo semplicemente che

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

Un prodotto scalare su X definisce una *norma* su X data da

$$\|x\| = \sqrt{\langle x, x \rangle} \quad (\geq 0) \quad (3.1)$$

e quindi una *metrica* su X data da

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}. \quad (3.2)$$

La prova che (3.1) soddisfa agli assiomi da (N1) a (N4) di una norma sarà data all'inizio della prossima sezione.

Quindi gli *spazi con prodotto scalare sono spazi normati*.

3.2 DEFINIZIONE (SPAZIO DI HILBERT)

Uno spazio con prodotto scalare che sia completo (completo nella metrica definita dal prodotto scalare; cf. (3.2)) si dice *spazio di Hilbert*. $\blacksquare$

Da (IP1) a (IP3) otteniamo le formule

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle \quad (3.3)$$

$$\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle \quad (3.4)$$

$$\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle \quad (3.5)$$

che useremo molto spesso. (3.3) mostra che il prodotto scalare è lineare nel primo fattore. Poiché in (3.5) abbiamo a destra i numeri complessi coniugati $\overline{\alpha}$ e $\overline{\beta}$ diciamo che il prodotto scalare è *coniugato lineare* nel secondo fattore. Riferendosi ad entrambe le proprietà diciamo che il prodotto scalare è *sesquilineare*. Ciò significa “ $1\frac{1}{2}$ volte lineare” ed è motivato dal fatto che “coniugato lineare” è anche noto come “semilineare”, un termine meno suggestivo che non useremo.

Il lettore può mostrare con un semplice calcolo diretto che la norma in uno spazio con prodotto scalare soddisfa l'importante **uguaglianza del parallelogramma**

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \quad (3.6)$$

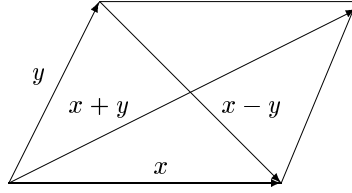


Figura 3.1: Uguaglianza del parallelogramma

Questo nome è suggerito dalla geometria elementare, come si vede dalla figura 3.1, se ricordiamo che la norma generalizza il concetto elementare di lunghezza di un vettore. È del tutto rimarchevole che una tale equazione continui a valere nel nostro caso molto più generale.

Concludiamo che se una norma non soddisfa la (3.6) essa non può essere ottenuta da un prodotto scalare con l'uso della (3.1). Queste norme effettivamente esistono. Possiamo perciò dire che *non tutti gli spazi normati sono spazi dotati di prodotto scalare*.

Definiamo ora il concetto di ortogonalità che è basilare in tutta la teoria. Sappiamo che se il prodotto scalare di due vettori nello spazio tridimensionale è zero allora i vettori sono ortogonali, cioè sono perpendicolari o almeno uno di essi è il vettore nullo. Ciò suggerisce e motiva la seguente definizione.

3.3 DEFINIZIONE (ORTOGONALITÀ)

Un elemento x di uno spazio X con prodotto scalare è detto *ortogonale* ad un elemento $y \in X$ se

$$\langle x, y \rangle = 0.$$

Diciamo anche che x e y sono *ortogonali* e scriviamo $x \perp y$. Analogamente per sottoinsiemi $A, B \subset X$ scriviamo $x \perp A$ se $x \perp a$ per tutti gli $a \in A$ ed $A \perp B$ se $a \perp b$ per tutti gli $a \in A$ e per tutti i $b \in B$. ■

Per due elementi ortogonali x, y è facile ottenere la **relazione di Pitagora**

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Più in generale se $\{x_1, \dots, x_n\}$ è un insieme ortogonale allora

$$\|x_1 + \dots + x_n\|^2 = \|x_1\|^2 + \dots + \|x_n\|^2.$$

Infatti $\langle x_j, x_k \rangle = 0$ se $j \neq k$ e di conseguenza

$$\left\| \sum_j x_j \right\|^2 = \left\langle \sum_j x_j, \sum_k x_k \right\rangle = \sum_j \sum_k \langle x_j, x_k \rangle = \sum_j \langle x_j, x_j \rangle = \sum_j \|x_j\|^2$$

(la somma va da 1 a n).

Infine menzioniamo il seguente fatto interessante. Il prodotto scalare può essere espresso esplicitamente in funzione della norma data dalla (3.1). Infatti il lettore può verificare con un calcolo diretto che per un prodotto scalare reale abbiamo

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \quad (3.7)$$

e per un prodotto scalare complesso

$$\begin{aligned}\Re\langle x, y \rangle &= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \\ \Im\langle x, y \rangle &= \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)\end{aligned}\quad (3.8)$$

dove $\Re\langle x, y \rangle$ e $\Im\langle x, y \rangle$ indicano la parte reale ed immaginaria. La formula (3.8) è talvolta chiamata *identità di polarizzazione*.

3.3 Ulteriori Proprietà degli Spazi con Prodotto Scalare

Prima di tutto dovremmo verificare che la (3.1) della sezione precedente definisce una norma.

(N1) e (N2) seguono dalla (IP4). Inoltre (N3) è ottenuto coll'uso di (IP2) e (IP3); infatti

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x \rangle = \alpha \bar{\alpha} \langle x, x \rangle = |\alpha|^2 \|x\|^2.$$

Infine (N4) è incluso nel seguente Lemma.

3.4 Lemma (Disuguaglianza di Schwarz, Disuguaglianza Triangolare)

Un prodotto scalare e la corrispondente norma soddisfano la disuguaglianza di Schwarz e la disuguaglianza triangolare secondo quanto segue.

(a) Abbiamo

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (\text{Disuguaglianza di Schwarz}) \quad (3.9)$$

dove il segno uguale vale se e solo se $\{x, y\}$ è un insieme linearmente dipendente.

(b) Questa norma soddisfa anche

$$\|x + y\| \leq \|x\| + \|y\| \quad (\text{Disuguaglianza Triangolare}) \quad (3.10)$$

dove il segno d'uguale vale se e solo se $y = 0$ o $x = cy$ (c reale e ≥ 0). ■

Dimostrazione. (a) Se $y = 0$ allora la (3.9) vale perché $\langle x, 0 \rangle = 0$. Sia $y \neq 0$. Per ogni scalare α abbiamo

$$\begin{aligned}0 &\leq \|x - \alpha y\|^2 = \langle x - \alpha y, x - \alpha y \rangle \\ &= \langle x, x \rangle - \bar{\alpha} \langle x, y \rangle - \alpha [\langle y, x \rangle - \bar{\alpha} \langle y, y \rangle].\end{aligned}$$

Vediamo che l'espressione nella parentesi $[\dots]$ è zero se scegliamo $\bar{\alpha} = \langle y, x \rangle / \langle y, y \rangle$. La disuguaglianza risultante è

$$0 \leq \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle} \langle x, y \rangle = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2},$$

dove abbiamo usato $\langle y, x \rangle = \overline{\langle x, y \rangle}$. Moltiplicando per $\|y\|^2$, trasferendo l'ultimo termine a sinistra e prendendo la radice si ottiene la (3.9).

L'uguaglianza vale in questa derivazione se e solo se $y = 0$ o $0 = \|x - \alpha y\|^2$, quindi $x - \alpha y = 0$, così che $x = \alpha y$, che mostra la dipendenza lineare.

¹Si noti che questa condizione per l'uguaglianza è perfettamente "simmetrica" in x e y perché $x = 0$ è incluso in $x = cy$ (per $c = 0$) ed è pure $y = kx$, $k = 1/c$ (per $c > 0$).

(b) Proviamo la (3.10). Abbiamo

$$||x + y||^2 = \langle x + y, x + y \rangle = ||x||^2 + \langle x, y \rangle + \langle y, x \rangle + ||y||^2 = ||x||^2 + 2\operatorname{Re}\langle x, y \rangle + ||y||^2.$$

Poiché $\operatorname{Re}\langle x, y \rangle \leq |\langle x, y \rangle|$ e a sua per la disuguaglianza di Schwarz

$$|\langle x, y \rangle| \leq ||x|| ||y||$$

otteniamo

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

Prendendo la radice di entrambi i membri abbiamo la (3.10).

L'uguaglianza in questa derivazione vale se e solo se

$$\operatorname{Re}\langle x, y \rangle = ||x|| ||y||.$$

Da ciò e dalla (3.9)

$$\operatorname{Re}\langle x, y \rangle = ||x|| ||y|| \geq |\langle x, y \rangle|. \quad (3.11)$$

Poiché la parte reale di un numero complesso non può essere maggiore del suo modulo dobbiamo avere l'uguaglianza. il che implica la dipendenza lineare per la parte (a), cioè $y = 0$ o $x = cy$. Mostriamo che c è reale e ≥ 0 . Dalla (3.11) col segno d'uguale abbiamo $\Re\langle x, y \rangle = |\langle x, y \rangle|$. Ma se la parte reale di un numero complesso è uguale al suo modulo la parte immaginaria deve essere zero. Quindi $\langle x, y \rangle = \Re\langle x, y \rangle \geq 0$ per la (3.11) e $c \geq 0$ segue da

$$0 \leq \langle x, y \rangle = \langle cy, y \rangle = c||y||^2. \quad \blacksquare$$

La disuguaglianza di Schwarz (3.9) è molto importante e sarà utilizzata nelle dimostrazioni più e più volte. Un'altra proprietà frequentemente usata è la continuità del prodotto scalare.

3.5 Lemma (Continuità del Prodotto Scalare)

Se in uno spazio con prodotto scalare $x_n \rightarrow x$ e $y_n \rightarrow y$ allora $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ e quindi per il Teorema 1.14 il prodotto scalare $\langle x, y \rangle$ è continuo in x e in y . ■

Dimostrazione. Sottraendo ed aggiungendo un termine, usando la disuguaglianza triangolare ed infine la disuguaglianza di Schwarz otteniamo

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\ &\leq |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\ &\leq ||x_n|| ||y_n - y|| + ||x_n - x|| ||y|| \rightarrow 0 \end{aligned}$$

poiché $y_n - y \rightarrow 0$ e $x_n - x \rightarrow 0$ per $n \rightarrow \infty$. ■

3.4 Definizione Equivalente di Spazio con Prodotto Scalare

Uno spazio pre-Hilbertiano (o con prodotto scalare) può essere definito in maniera equivalente come segue.

3.6 DEFINIZIONE (SPAZIO PRE-HILBERTIANO)

Uno spazio normato X (reale o complesso) è chiamato *spazio pre-hilbertiano* se la sua norma soddisfa all'uguaglianza del parallelogramma

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \quad (3.12)$$

■

L'equivalenza di questa definizione con la definizione 3.1 è provata dal seguente teorema e suo corollario.

3.7 Teorema (Spazio pre-Hilbertiano Reale)

Definiamo in uno spazio normato reale X , che soddisfa all'uguaglianza del parallelogramma,

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2). \quad (3.13)$$

Allora $\langle x, y \rangle$ soddisfa a tutte le proprietà richieste al prodotto scalare, da (IP1) a (IP4). ■

Dimostrazione. (IP3) e (IP4) sono evidenti. Per ottenere la (IP1) osserviamo dapprima che dalla (3.13) segue

$$\langle x, z \rangle + \langle y, z \rangle = \frac{1}{4}(\|x + z\|^2 - \|x - z\|^2 + \|y + z\|^2 - \|y - z\|^2).$$

Se ora notiamo che

$$\begin{aligned} x \pm z &= \left(\frac{x + y}{2} \pm z \right) + \frac{x - y}{2} \\ y \pm z &= \left(\frac{x + y}{2} \pm z \right) - \frac{x - y}{2} \end{aligned}$$

otteniamo utilizzando l'uguaglianza del parallelogramma

$$\langle x, z \rangle + \langle y, z \rangle = \frac{1}{2} \left(\left\| \frac{x + y}{2} + z \right\|^2 - \left\| \frac{x + y}{2} - z \right\|^2 \right)$$

ed infine per la (3.13)

$$\langle x, z \rangle + \langle y, z \rangle = 2 \left\langle \frac{x + y}{2}, z \right\rangle. \quad (3.14)$$

Se prendiamo $y = 0$ otteniamo $\langle x, z \rangle = 2 \left\langle \frac{x}{2}, z \right\rangle$, perché $\langle 0, z \rangle = 0$ per la (3.13). Quindi dalla (3.14) otteniamo (IP1). Dalla (IP1) segue che per m intero positivo qualunque

$$\langle mx, z \rangle = m \langle x, z \rangle$$

e quindi la (IP2) vale per un numero razionale positivo qualunque $\alpha = \frac{m}{n}$, perché

$$\left\langle \frac{x}{n}, z \right\rangle = \frac{n}{n} \left\langle \frac{x}{n}, z \right\rangle = \frac{1}{n} \langle x, z \rangle.$$

Poiché dalla definizione (3.13) segue che $\langle -x, z \rangle = -\langle x, z \rangle$, la (IP2) vale anche per un numero razionale negativo qualunque.

In uno spazio normato $\|\alpha x + y\|^2$ e $\|\alpha x - y\|^2$ sono continui in α e quindi per la (3.13) anche $\langle \alpha x, y \rangle$ è continuo in α . Poiché ogni reale α si può ottenere come limite di una successione di numeri razionali, anche (IP2) è provata. ■

3.8 Corollario (Spazio pre-Hilbertiano Complesso)

Definiamo in uno spazio normato complesso X , che soddisfa all'uguaglianza del parallelogramma,

$$\langle x, y \rangle = \langle x, y \rangle_1 + i \langle x, iy \rangle_1 \quad (3.15)$$

dove

$$\langle x, y \rangle_1 = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2). \quad (3.16)$$

Allora $\langle x, y \rangle$ soddisfa a tutte le proprietà richieste al prodotto scalare, da (IP1) a (IP4). ■

Dimostrazione. Poiché X è anche uno spazio pre-hilbertiano reale $\langle x, y \rangle_1$ e $\langle x, iy \rangle_1$ e quindi $\langle x, y \rangle$ soddisfano a (IP1) e a (IP2) per α reale. Per la (3.16) abbiamo $\langle y, x \rangle_1 = \langle x, y \rangle_1$, $\langle ix, iy \rangle_1 = \langle x, y \rangle_1$ e quindi $\langle y, ix \rangle_1 = \langle -i iy, ix \rangle_1 = -\langle iy, x \rangle_1 = -\langle x, iy \rangle_1$. Perciò

$$\langle y, x \rangle = \langle y, x \rangle_1 + i \langle y, ix \rangle_1 = \langle x, y \rangle_1 - i \langle x, iy \rangle_1 = \overline{\langle x, y \rangle}$$

e vale (IP3). Analogamente abbiamo

$$\langle ix, y \rangle = \langle ix, y \rangle_1 + i \langle ix, iy \rangle_1 = -\langle x, iy \rangle_1 + i \langle x, y \rangle_1 = i \langle x, y \rangle$$

e quindi abbiamo provato la (IP2) per α complesso. Infine vale la (IP4) perché

$$\langle x, x \rangle_1 = \|x\|^2 \quad \text{e} \quad \langle x, ix \rangle_1 = \frac{1}{4}(|1 + i|^2 - |1 - i|^2)\|x\|^2 = 0. \quad \blacksquare$$

3.5 Completamento di uno Spazio con Prodotto Scalare

Mostriamo che ogni spazio con prodotto scalare può essere completato. Il completamento è uno spazio di Hilbert ed è unico a meno di isomorfismi. La definizione di isomorfismo è qui la seguente (come suggerito dalla nostra discussione nella Sez. 2.9).

Un **isomorfismo** T di uno spazio con prodotto scalare X su uno spazio con prodotto scalare $\tilde{X}$ sullo stesso campo è un operatore lineare biiettivo $T : X \rightarrow \tilde{X}$ che conserva il prodotto scalare, cioè tale che per tutti gli $x, y \in X$

$$\langle Tx, Ty \rangle = \langle x, y \rangle,$$

dove per semplicità indichiamo con lo stesso simbolo il prodotto scalare su X e su $\tilde{X}$. $\tilde{X}$ è allora detto *isomorfo* ad X e X ed $\tilde{X}$ sono detti *spazi con prodotto scalare isomorfi*. Si osservi che la biattività e la linearità garantiscono che T è un isomorfismo fra spazi vettoriali di X su $\tilde{X}$, così che T conserva l'intera struttura di spazio con prodotto scalare. T è anche un'isometria di X su $\tilde{X}$ perché le distanze in X e $\tilde{X}$ sono determinate dalle norme definite a mezzo dei prodotti scalari in X ed $\tilde{X}$.

Il teorema sul completamento di spazi con prodotto scalare può ora esser formulato come segue.

3.9 Teorema (Completamento)

Per un qualunque spazio con prodotto scalare X esiste uno spazio di Hilbert H ed un isomorfismo A da X su un sottospazio denso $W \subset H$. Lo spazio H è unico a meno di isomorfismi. ■

Dimostrazione. Per il Teorema 2.8 esiste uno spazio di Banach H ed un'isometria A da X su un sottospazio W di H che è denso in H . Essendo X e W isometrici anche W soddisfa all'uguaglianza del parallelogramma. Poiché ogni elemento di H si può esprimere come limite di una successione di elementi di W , grazie alla continuità della norma, ne deduciamo che pure H soddisfa all'uguaglianza del parallelogramma ed è quindi uno spazio pre-hilbertiano. Concludiamo che H , essendo uno spazio di Banach, è uno spazio di Hilbert.

Il Teorema 2.8 garantisce anche che H è unico a meno di isometrie, cioè due completamenti H e $\tilde{H}$ di X sono collegati da un isometria $T : H \rightarrow \tilde{H}$. Ragionando come nel caso di A concludiamo che T deve essere un isomorfismo dello spazio di Hilbert H sullo spazio di Hilbert $\tilde{H}$. ■

Un **sottospazio** Y di uno spazio con prodotto scalare X è definito come un sottospazio vettoriale di X preso con il prodotto scalare su X ristretto a $Y \times Y$.

Analogamente un **sottospazio** Y di uno spazio di Hilbert H è definito come un sottospazio di H , considerato come uno spazio con prodotto scalare. Si noti che Y non è necessariamente uno spazio di Hilbert perché Y può non essere completo. Infatti dai Teoremi 2.7 e 2.10 ricaviamo immediatamente le affermazioni (a) e (b) del seguente teorema.

3.10 Teorema (Sottospazio)

Sia Y un sottospazio di uno spazio di Hilbert H . Ne segue

- (a) *Y è completo se e solo se Y è chiuso in H .*
- (b) *Se Y è finito dimensionale allora Y è completo.*
- (c) *Se H è separabile lo è anche Y . Più generalmente ogni sottoinsieme di uno spazio con prodotto scalare separabile è separabile.* ■

La semplice dimostrazione di (c) è lasciata al lettore.

3.6 Complemento Ortogonale e Somma Diretta

In uno spazio metrico X la *distanza* δ di un elemento $x \in X$ da un sottospazio non vuoto $M \subset X$ è definita essere

$$\delta = \inf_{\tilde{y} \in M} d(x, \tilde{y}) \quad (M \neq \emptyset).$$

In uno spazio normato ciò diventa

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\|. \quad (3.17)$$

Vedremo che è importante sapere se v'è un $y \in M$ tale che

$$\delta = \|x - y\|, \quad (3.18)$$

ossia, parlando intuitivamente, se v'è un punto $y \in M$ che sia il più vicino ad un dato x e se, esistendo questo punto, esso sia unico. Questo è un *problema di esistenza ed unicità*.

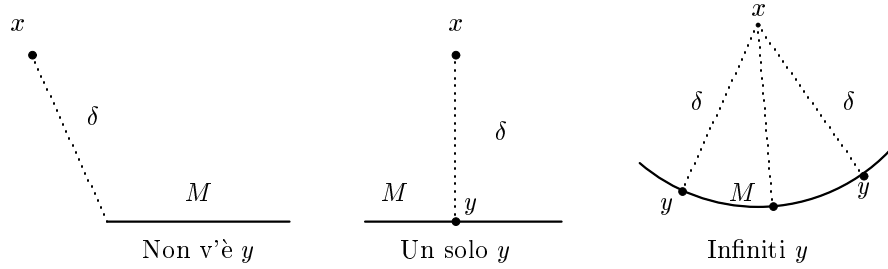


Figura 3.2: Vettore minimizzante

Esso è di fondamentale importanza sia teorica che applicativa, ad esempio in connessione con l'approssimazione delle funzioni.

La figura 3.2 illustra il fatto che anche in uno spazio molto semplice come lo spazio euclideo $\mathbb{R}^3$ vi può non essere alcun y che soddisfa la (3.18) od anche uno solo o più di uno. Possiamo aspettarci che in altri spazi, in particolare negli spazi multidimensionali, vi possano essere a questo riguardo situazioni anche più complicate. Per un generico spazio normato questo è il caso (come si può vedere), ma per gli spazi di Hilbert la situazione rimane relativamente semplice. Questo fatto è sorprendente ed ha diverse conseguenze teoriche e pratiche. Ed è una delle ragioni per cui la teoria degli spazi di Hilbert è più semplice di quella dei generici spazi di Banach.

Per considerare il problema di esistenza ed unicità per gli spazi di Hilbert e per formulare il teorema chiave (3.11, qui di seguito) abbiamo bisogno di due concetti collegati, che sono di interesse generale.

Il **segmento** che congiunge due dati elementi x e y di uno spazio vettoriale X è definito come l'insieme di tutti gli $z \in X$ della forma

$$z = (1 - \alpha)x + \alpha y \quad (\alpha \in \mathbb{R}, 0 \leq \alpha \leq 1).$$

Un sottoinsieme M di X è detto **convesso** se per ogni $x, y \in M$ il segmento che congiunge x e y è contenuto in M .

Per esempio ogni sottospazio Y di X è convesso e l'intersezione di spazi convessi è convesso.

Possiamo ora fornire il principale strumento di questa sezione.

3.11 Teorema (Vettore Minimizzante)

Sia X uno spazio con prodotto scalare ed $M \neq \emptyset$ un sottoinsieme convesso completo (nella metrica indotta dal prodotto scalare). Allora per ogni dato $x \in X$ esiste un unico $y \in M$ tale che

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|. \quad (3.19)$$

■

Dimostrazione. (a) Esistenza. Per definizione di estremo inferiore v'è una successione (y_n) in M tale che

$$\delta_n \rightarrow \delta \quad \text{dove} \quad \delta_n = \|x - y_n\|. \quad (3.20)$$

Mostriamo che (y_n) è di Cauchy. Ponendo $y_n - x = v_n$ abbiamo che $\|v_n\| = \delta_n$ e che

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2 \left\| \frac{1}{2}(y_n + y_m) - x \right\| \geq 2\delta$$

perché M è convesso, così che $\frac{1}{2}(y_n + y_m) \in M$. Inoltre abbiamo $y_n - y_m = v_n - v_m$. Quindi per l'uguaglianza del parallelogramma

$$\begin{aligned} \|y_n - y_m\|^2 &= \|v_n - v_m\|^2 = -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \\ &\leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2) \end{aligned}$$

e (3.20) implica che (y_n) è di Cauchy. Poiché M è completo (y_n) converge, ossia $y_n \rightarrow y \in M$. Poiché $y \in M$ abbiamo che $\|x - y\| \geq \delta$. Ma per la (3.20)

$$\|x - y\| \leq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \rightarrow \delta.$$

Ciò mostra che $\|x - y\| = \delta$.

(b) Unicità. Assumiamo che entrambi $y \in M$ e $y_0 \in M$ soddisfino

$$\|x - y\| = \delta \quad \text{e} \quad \|x - y_0\| = \delta$$

e mostriamo che allora $y_0 = y$. Per l'uguaglianza del parallelogramma

$$\begin{aligned} \|y - y_0\|^2 &= \|(y - x) - (y_0 - x)\|^2 \\ &= 2\|y - x\|^2 + 2\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2^2 \left\| \frac{1}{2}(y + y_0) - x \right\|^2. \end{aligned}$$

A destra $\frac{1}{2}(y + y_0) \in M$, così che

$$\left\| \frac{1}{2}(y + y_0) - x \right\| \geq \delta.$$

Ciò implica che il membro a destra è minore od uguale a $2\delta^2 + 2\delta^2 - 4\delta^2 = 0$. Quindi abbiamo la disequaglianza $\|y - y_0\|^2 \leq 0$, ossia $y_0 = y$. ■

Passando dagli spazi convessi arbitrari ai sottospazi otteniamo il lemma che generalizza l'idea familiare della geometria elementare che l'unico punto y in un sottospazio dato Y più vicino ad un dato x viene trovato "tracciando una perpendicolare da x a Y ".

3.12 Lemma (Ortogonalità)

Nel teorema 3.11 sia M un sottospazio completo Y e sia $x \in X$ fissato. Allora $z = x - y$ è ortogonale a Y . ■

Dimostrazione. Se $z \perp Y$ fosse falso vi sarebbe un $y_1 \in Y$ tale che

$$\langle z, y_1 \rangle = \beta \neq 0. \quad (3.21)$$

Chiaramente $y_1 \neq 0$ perché altrimenti $\langle z, y_1 \rangle = 0$. Inoltre per ogni scalare α

$$\begin{aligned} \|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \bar{\alpha} \langle z, y_1 \rangle - \alpha [\langle y_1, z \rangle - \bar{\alpha} \langle y_1, y_1 \rangle] \\ &= \langle z, z \rangle - \bar{\alpha} \beta - \alpha [\bar{\beta} - \bar{\alpha} \langle y_1, y_1 \rangle]. \end{aligned}$$

L'espressione in parentesi è zero se scegliamo

$$\bar{\alpha} = \frac{\bar{\beta}}{\langle y_1, y_1 \rangle}.$$

Dalla (3.19) abbiamo $\|z\| = \|x - y\| = \delta$ così che la nostra equazione ora ci dà

$$\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1, y_1 \rangle} < \delta^2.$$

Ma questo è impossibile perché abbiamo

$$z - \alpha y_1 = x - y_2 \quad \text{dove} \quad y_2 = y + \alpha y_1 \in Y,$$

così che $\|z - \alpha y_1\| \geq \delta$ per definizione di δ . Quindi (3.21) non può essere valida ed il lemma è dimostrato. ■

È sovente utile rappresentare uno spazio di Hilbert come somma diretta di due sottospazi. Questa decomposizione risulta essere particolarmente semplice e conveniente se viene realizzata utilizzando l'ortogonalità. Per comprendere la situazione ed il problema introduciamo dapprima il concetto di somma diretta. Questo concetto ha senso per qualunque spazio vettoriale ed è definito come segue.

3.13 DEFINIZIONE (SOMMA DIRETTA)

Uno spazio vettoriale X è detto *somma diretta* di due sottospazi Y e Z di X e si scrive

$$X = Y \oplus Z$$

se ciascun $x \in X$ ha un'unica rappresentazione

$$x = y + z \quad y \in Y, z \in Z.$$

Allora Z è chiamato un *complemento algebrico* di Y in X e viceversa, e Y, Z è chiamato una *coppia complementare di sottospazi in X* . ■

Ad esempio $Y = \mathbb{R}$ è un sottospazio del piano euclideo $\mathbb{R}^2$. Chiaramente Y ha infiniti diversi complementi algebrici in $\mathbb{R}^2$, ciascuno dei quali è una retta reale. Ma il più conveniente è un complemento che è perpendicolare.

Analogamente nel caso di uno spazio generale di Hilbert H l'interesse principale riguarda le rappresentazioni di H come somma diretta di un sottospazio chiuso Y e del suo **complemento ortogonale**

$$Y^\perp = \{z \in H : z \perp Y\},$$

che è l'insieme di tutti i vettori ortogonali a Y . Questo è il maggior risultato in questa sezione, che è qualche volta chiamato il *teorema della proiezione* per ragioni che spiegheremo dopo la dimostrazione.

3.14 Teorema (Somma Diretta)

Sia Y un qualunque sottospazio di uno spazio di Hilbert H . Allora

$$H = \overline{Y} \oplus Z \quad Z = Y^\perp. \quad (3.22)$$

■

Dimostrazione. Poiché H è completo e $\overline{Y}$ è chiuso, $\overline{Y}$ è completo per il Teorema 1.13. Poiché $\overline{Y}$ è convesso il Teorema 3.11 ed il Lemma 3.12 implicano che per ogni $x \in H$ v'è un $y \in \overline{Y}$ tale che

$$x = y + z \quad z \in \overline{Y}^\perp. \quad (3.23)$$

Notiamo ora che $\overline{Y}^\perp = Y^\perp$ perché se $v \perp Y$, per la continuità del prodotto scalare, è anche $v \perp \overline{Y}$.

Rimane quindi da provare l'unicità della decomposizione (3.22). Assumiamo sia

$$x = y + z = y_1 + z_1$$

dove $y, y_1 \in \overline{Y}$ e $z, z_1 \in Z$. Allora $y - y_1 = z_1 - z$. Poiché $y - y_1 \in \overline{Y}$ mentre $z_1 - z \in Z = Y^\perp = \overline{Y}^\perp$ vediamo che $y - y_1 \in \overline{Y} \cap \overline{Y}^\perp = \{0\}$. Ciò implica $y = y_1$. Quindi anche $z = z_1$. ■

In genere questo teorema si utilizza nel caso in cui Y è chiuso. È allora

$$H = Y \oplus Y^\perp$$

e la decomposizione ortogonale di x si scrive

$$x = y + z \quad y \in Y, \quad z \in Y^\perp. \quad (3.24)$$

y in (3.24) è chiamato, con terminologia presa a prestito dalla geometria elementare, la **proiezione ortogonale** di x su Y (o brevemente la *proiezione* di x su Y).

L'equazione (3.24) definisce un'applicazione

$$P : H \rightarrow Y \\ x \mapsto y = Px.$$

P è chiamato la **proiezione** (ortogonale) o (*operatore di proiezione*) di H su Y .

L'operatore lineare P è limitato con norma $\|P\| = 1$. Infatti, poiché y e z sono ortogonali, abbiamo

$$\|Px\| = \|y\| \leq \sqrt{\|y\|^2 + \|z\|^2} = \|x\| \quad \text{per ogni } x \in H,$$

cioè P è limitato con norma $\|P\| \leq 1$. Per $x = y$ è $\|Py\| = \|y\|$ e quindi $\|P\| = 1$.

Si noti che la restrizione di P a Y è l'operatore identità su Y essendo

$$Py = y, \quad \text{per ogni } y \in Y$$

e che P è **idempotente**, ossia

$$P^2 = P,$$

essendo

$$P^2x = P(Px) = Py = y = Px, \quad \text{per ogni } x \in H.$$

Inoltre

$$Pz = 0, \quad \text{per ogni } z \in Z = Y^\perp.$$

Un discorso perfettamente analogo si può ripetere per l'operatore $I - P$ che è il proiettore ortogonale su $Z = Y^\perp$.

Riportiamo ora alcune osservazioni sui complementi ortogonali, che saranno utili nel seguito. Sia $M \neq \emptyset$ un sottoinsieme (non necessariamente un sottospazio) di uno spazio con prodotto scalare X e sia $M^\perp$ il suo complemento ortogonale, ossia l'insieme

$$M^\perp = \{x \in X : \langle x, v \rangle = 0, \forall v \in M\}.$$

Si noti che $M^\perp$ è uno spazio vettoriale poiché $x, y \in M^\perp$ implica per tutti i $v \in M$ e per tutti gli scalari α, β

$$\langle \alpha x + \beta y, v \rangle = \alpha \langle x, v \rangle + \beta \langle y, v \rangle = 0$$

e quindi $\alpha x + \beta y \in M^\perp$.

$M^\perp$ è inoltre chiuso come il lettore può provare utilizzando la continuità del prodotto scalare.

Scriviamo $M^{\perp\perp}$ per indicare $(M^\perp)^\perp$. In generale abbiamo

$$\overline{M} \subset M^{\perp\perp}. \quad (3.25)$$

Infatti se

$$x \in M \implies x \perp M^\perp \implies x \in (M^\perp)^\perp$$

cioè $M \subset M^{\perp\perp}$. Inoltre se $x \in \overline{M}$ allora esiste una successione (x_n) di M tale che $x_n \rightarrow x$. Per la continuità del prodotto scalare, poiché $x_n \perp M^\perp$ anche $x \perp M^\perp$ e quindi $x \in M^{\perp\perp}$.

Nel caso in cui M sia un sottospazio possiamo formulare il seguente lemma.

3.15 Lemma (Doppio Complemento Ortogonale)

Se Y è un sottospazio di uno spazio di Hilbert H allora

$$\overline{Y} = Y^{\perp\perp}. \quad (3.26)$$

■

Dimostrazione. $\overline{Y} \subset Y^{\perp\perp}$ per la (3.25). Mostriamo che $\overline{Y} \supset Y^{\perp\perp}$. Sia $x \in Y^{\perp\perp}$. Allora per il teorema 3.14 possiamo scrivere per x la decomposizione ortogonale $x = y + z$, dove $y \in \overline{Y}$ e $z \in Y^\perp$. Ma $x \in Y^{\perp\perp}$ per assunzione ed $y \in Y^{\perp\perp}$ perché per la (3.25) $\overline{Y} \subset Y^{\perp\perp}$ e quindi, poiché $Y^{\perp\perp}$ è uno spazio vettoriale, anche $z = x - y \in Y^{\perp\perp}$. Essendo $z \in Y^\perp$ abbiamo che $z \perp z$ e quindi $z = 0$ così che $x = y$, ossia $x \in \overline{Y}$. Poiché $x \in Y^{\perp\perp}$ era arbitrario ciò prova che $\overline{Y} \supset Y^{\perp\perp}$. ■

Il Teorema 3.14 ed il Lemma conseguente 3.15 implicano facilmente una caratterizzazione degli insiemi negli spazi di Hilbert il cui span è denso, che è la seguente.

3.16 Lemma (Insieme Denso)

Per ogni sottoinsieme $M \neq \emptyset$ di uno spazio di Hilbert H lo span V di M è denso in H se e solo se $M^\perp = \{0\}$. ■

Dimostrazione. (a) Sia $x \in M^\perp$ ed assumiamo che $V = \text{span } M$ sia denso in H . Allora $x \in \overline{V} = H$. Per il teorema 1.12(a) v'è una successione (x_n) in V tale che $x_n \rightarrow x$. Poiché $x \in M^\perp$ e $M^\perp \perp V$ abbiamo $\langle x_n, x \rangle = 0$. La continuità del prodotto scalare (cf. Lemma 3.5) implica che $\langle x_n, x \rangle \rightarrow \langle x, x \rangle$. In conclusione $\langle x, x \rangle = \|x\|^2 = 0$, ossia $x = 0$. Poiché $x \in M^\perp$ era arbitrario ciò mostra che $M^\perp = \{0\}$.

(b) Viceversa supponiamo che $M^\perp = \{0\}$. Se $x \perp V$ allora $x \perp M$ così che $x \in M^\perp$ e $x = 0$. Quindi $V^\perp = \{0\}$. Per il Teorema 3.14 abbiamo che $H = \overline{V} \oplus V^\perp$. Ma essendo $V^\perp = \{0\}$ otteniamo $H = \overline{V}$, cioè V è denso in H . ■

3.7 Insiemi e Successioni Ortonormali

3.17 DEFINIZIONE (INSIEMI E SUCCESSIONI ORTOGONALI)

Un *insieme ortogonale* M in uno spazio con prodotto scalare X è un sottoinsieme $M \subset X$ i cui elementi sono a due a due ortogonali. Un *insieme ortonormale* $M \subset X$ è un insieme ortogonale in X i cui elementi hanno norma 1, ossia per tutti gli $x, y \in M$

$$\langle x, y \rangle = \begin{cases} 0 & \text{if } x \neq y \\ 1 & \text{if } x = y. \end{cases} \quad (3.27)$$

Se un insieme ortogonale o ortonormale M è numerabile possiamo ordinarlo in una successione (x_n) e chiamarlo, rispettivamente, *successione ortogonale* o *ortonormale*.

Più in generale un insieme con indice, o *famiglia*, (x_α) , $\alpha \in I$, è chiamata ortogonale se $x_\alpha \perp x_\beta$ per tutti gli $\alpha, \beta \in I$, $\alpha \neq \beta$. La famiglia è chiamata *ortonormale* se è ortogonale e tutti gli x_α hanno norma 1, così che per tutti gli $\alpha, \beta \in I$ abbiamo

$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ 1 & \text{if } \alpha = \beta. \end{cases}, \quad (3.28)$$

dove $\delta_{\alpha\beta}$ è la delta di Kronecker. ■

Consideriamo ora alcune semplici proprietà degli insiemi ortogonali e ortonormali.

3.18 Lemma (Indipendenza Lineare)

Un *insieme ortonormale* è linearmente indipendente. ■

Dimostrazione. Sia $\{e_1, \dots, e_n\}$ ortonormale e consideriamo l'equazione

$$\alpha_1 e_1 + \dots + \alpha_n e_n = 0.$$

Moltiplicando per un e_j fisso si ottiene

$$\left\langle \sum_k \alpha_k e_k, e_j \right\rangle = \sum_k \alpha_k \langle e_k, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j = 0$$

ciò che prova l'indipendenza lineare per ogni insieme finito o infinito ortonormale. ■

Un grande vantaggio delle successioni ortonormali rispetto alle successioni arbitrarie linearmente indipendenti è il seguente. Se sappiamo che un dato x può essere rappresentato come una combinazione lineare di alcuni elementi di una successione ortonormale, allora la ortonormalità rende l'effettiva determinazione dei coefficienti molto facile. Infatti se $\{e_1, \dots, e_n\}$ è una successione ortonormale in uno spazio con prodotto scalare X e se abbiamo che $x \in \text{span}\{e_1, \dots, e_n\}$, dove n è fisso, allora per la definizione di span

$$x = \sum_{k=1}^n \alpha_k e_k, \quad (3.29)$$

e se calcoliamo il prodotto scalare per un e_j fisso otteniamo

$$\langle x, e_j \rangle = \left\langle \sum_{k=1}^n \alpha_k e_k, e_j \right\rangle = \sum_{k=1}^n \alpha_k \langle e_k, e_j \rangle = \alpha_j.$$

Con questi coefficienti la (3.29) diventa

$$x = \sum_{k=1}^n \langle x, e_k \rangle e_k. \quad (3.30)$$

L'utilizzo di un insieme ortonormale risulta vantaggioso anche quando di un elemento $\bar{x} \in \text{span}\{e_1, \dots, e_{n+1}\}$ si conosca già la proiezione x sul sottospazio $\text{span}\{e_1, \dots, e_n\}$ espressa secondo (3.30). E allora

$$\bar{x} = x + \alpha_{n+1} e_{n+1}$$

con $\alpha_{n+1} = \langle \bar{x}, e_{n+1} \rangle$ e quindi per esprimere $\bar{x}$ come combinazione lineare degli e_k rimane da calcolare solamente questo coefficiente mentre gli altri rimangono invariati.

3.19 Teorema (Diseguaglianza di Bessel)

Sia (e_k) una successione ortonormale in uno spazio con prodotto scalare X . Allora per ogni $x \in X$, il vettore $y \in Y_n = \text{span}\{e_1, \dots, e_n\}$ che ha distanza minima da x è

$$y = \sum_{k=1}^n \langle x, e_k \rangle e_k, \quad (3.31)$$

e vale

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2 \quad (\text{Diseguaglianza di Bessel}). \quad (3.32)$$

I prodotti scalari $\langle x, e_k \rangle$ nella (3.32) sono chiamati i coefficienti di Fourier di x rispetto alla successione ortonormale (e_k) . ■

Dimostrazione. Per un generico $\tilde{y} \in Y_n$

$$\tilde{y} = \sum_{k=1}^n \alpha_k e_k$$

abbiamo, grazie all'ortonormalità degli e_k ,

$$\begin{aligned} \|x - \tilde{y}\|^2 &= \left\langle x - \sum_{k=1}^n \alpha_k e_k, x - \sum_{k=1}^n \alpha_k e_k \right\rangle \\ &= \|x\|^2 - \sum_{k=1}^n \alpha_k \langle e_k, x \rangle - \sum_{k=1}^n \bar{\alpha}_k \langle x, e_k \rangle + \sum_{k=1}^n |\alpha_k|^2 \\ &= \|x\|^2 - \sum_{k=1}^n |\langle x, e_k \rangle|^2 + \sum_{k=1}^n |\langle x, e_k \rangle - \alpha_k|^2. \end{aligned}$$

Quindi il vettore di Y_n che ha distanza minima da x è quello definito in (3.31).

Poiché $\|x - y\|^2 \geq 0$ abbiamo per ogni $n = 1, 2, \dots$

$$\sum_{k=1}^n |\langle x, e_k \rangle|^2 \leq \|x\|^2. \quad (3.33)$$

Questa somma ha termini non negativi e perciò forma una successione monotona non decrescente. Questa successione converge perché è limitata da $\|x\|^2$. Quindi la (3.33) implica la disuguaglianza di Bessel (3.32). ■

Si noti che se X è finito dimensionale allora ogni insieme ortonormale in X , essendo linearmente indipendente per il Lemma 3.18, deve essere finito. Quindi in questo caso in (3.32) abbiamo una somma finita.

Osserviamo infine che, grazie al Lemma 3.12, $z = x - y \perp y$. Ciò si può anche mostrare direttamente. Notiamo dapprima che per l'ortonormalità

$$\|y\|^2 = \left\langle \sum_{k=1}^n \langle x, e_k \rangle e_k, \sum_{m=1}^n \langle x, e_m \rangle e_m \right\rangle = \sum_{k=1}^n |\langle x, e_k \rangle|^2. \quad (3.34)$$

Quindi, usando questa formula, otteniamo

$$\begin{aligned} \langle z, y \rangle &= \langle x - y, y \rangle = \langle x, y \rangle - \langle y, y \rangle \\ &= \left\langle x, \sum_{k=1}^n \langle x, e_k \rangle e_k \right\rangle - \|y\|^2 \\ &= \sum_{k=1}^n \overline{\langle x, e_k \rangle} \langle x, e_k \rangle - \sum_{k=1}^n |\langle x, e_k \rangle|^2 \\ &= 0 \end{aligned}$$

ossia $z \perp y$.

Abbiamo visto che le successioni ortonormali sono molto convenienti da utilizzare. Rimane il problema pratico di come ottenere una successione ortonormale se è data un'arbitraria successione linearmente indipendente. Ciò si ottiene mediante un procedimento costruttivo, il **processo di Gram-Schmidt** per ortonormalizzare una successione linearmente indipendente (x_j) in uno spazio con prodotto scalare. La successione ortonormale risultante (e_n) ha la proprietà che per ogni n

$$\text{span}\{e_1, \dots, e_n\} = \text{span}\{x_1, \dots, x_n\}.$$

Il processo è il seguente.

Primo passo. Il primo elemento di (e_n) è

$$e_1 = \frac{1}{\|x_1\|} x_1.$$

Secondo passo. x_2 può esser scritto

$$x_2 = \langle x_2, e_1 \rangle e_1 + v_2.$$

Allora

$$v_2 = x_2 - \langle x_2, e_1 \rangle e_1$$

non è il vettore nullo perché (x_j) è linearmente indipendente; inoltre $v_2 \perp e_1$ perché $\langle v_2, e_1 \rangle = 0$ così che possiamo prendere

$$e_2 = \frac{1}{\|v_2\|} v_2.$$

Passo n-mo. Il vettore

$$v_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k \quad (3.35)$$

non è nullo ed è ortogonale a $e_1, \dots, e_{n-1}$. Quindi otteniamo

$$e_n = \frac{1}{\|v_n\|} v_n. \quad (3.36)$$

Si noti che la somma che viene sottratta nel membro a destra della (3.35) è la proiezione di x sullo $\text{span}\{e_1, \dots, e_n\}$. In altre parole ad ogni passo sottraiamo a x_n la sua “componente” nella direzione dei vettori precedentemente ortonormalizzati. Ciò dà v_n che è poi moltiplicato per $1/\|v_n\|$ in modo da ottenere un vettore di norma uno. v_n per n qualunque non può essere un vettore nullo. Infatti se n fosse il più piccolo indice per cui $v_n = 0$ allora la (3.35) mostrerebbe che x_n sarebbe una combinazione lineare degli $e_1, \dots, e_{n-1}$ e quindi una combinazione lineare degli $x_1, \dots, x_{n-1}$ contraddicendo l'assunzione che $\{x_1, \dots, x_n\}$ è linearmente indipendente.

3.8 Serie Collegate a Successioni e Insiemi Ortonormali

In questa sezione consideriamo dapprima le serie collegate alle successioni ortonormali ed poi estendiamo i risultati ottenuti agli insiemi ortonormali non numerabili.

Data una successione ortonormale (e_n) in uno spazio di Hilbert H possiamo considerare la serie della forma

$$\sum_{k=1}^{\infty} \alpha_k e_k \quad (3.37)$$

dove gli $\alpha_1, \alpha_2, \dots$ sono scalari qualunque. In accordo con la definizione data nella Sez. 2.3 una tale serie *converge* ed ha la *somma* s se esiste un $s \in H$ tale che la successione (s_n) delle somme parziali

$$s_n = \alpha_1 e_1 + \dots + \alpha_n e_n$$

converge a s , ossia $\|s - s_n\| \rightarrow 0$ per $n \rightarrow \infty$.

3.20 Teorema (Convergenza)

Sia (e_n) una successione ortonormale in uno spazio di Hilbert H . Allora

(a) *La serie (3.37) converge se e solo se la seguente serie*

$$\sum_{k=1}^{\infty} |\alpha_k|^2 \quad (3.38)$$

converge.

- (b) Se la (3.37) converge e si indica con x la sua somma, allora i coefficienti α_k sono i coefficienti di Fourier $\langle x, e_k \rangle$ di x e si può scrivere

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k. \quad (3.39)$$

- (c) Per qualsiasi $x \in H$ la serie (3.37) con $\alpha_k = \langle x, e_k \rangle$ converge.

- (d) Dato un qualsiasi $x \in H$ condizione necessaria e sufficiente perché sia

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

è che sia

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2. \quad (3.40)$$

■

Dimostrazione. (a) Siano

$$s_n = \alpha_1 e_1 + \cdots + \alpha_n e_n \quad \text{e} \quad \sigma_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2.$$

Allora a causa dell'ortonormalità per qualsiasi m e $n > m$

$$\begin{aligned} \|s_n - s_m\|^2 &= \|\alpha_{m+1} e_{m+1} + \cdots + \alpha_n e_n\|^2 \\ &= |\alpha_{m+1}|^2 + \cdots + |\alpha_n|^2 = \sigma_n - \sigma_m. \end{aligned}$$

Quindi (s_n) è di Cauchy in H se e solo se (σ_n) è di Cauchy in $\mathbb{R}$. Poiché H e $\mathbb{R}$ sono completi ne segue la prima affermazione del teorema.

(b) Sia α_k un generico ma fisso coefficiente della serie. Prendendo il prodotto scalare di un qualunque s_n con $n \geq k$ e di e_k ed usando l'ortonormalità abbiamo che

$$\langle s_n, e_k \rangle = \alpha_k \quad \text{per ogni } n \geq k.$$

Per ipotesi $s_n \rightarrow x$. Poiché il prodotto scalare è continuo

$$\alpha_k = \langle s_n, e_k \rangle \rightarrow \langle x, e_k \rangle$$

come si voleva dimostrare.

- (c) Dalla disuguaglianza di Bessel nel Teorema 3.19 segue la convergenza della serie

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2.$$

Da ciò e da (a) concludiamo che (c) deve esser valido.

- (d) Se è $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ allora per la norma di x possiamo scrivere

$$\|x\|^2 = \left\langle x, \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \right\rangle,$$

da cui utilizzando la continuità del prodotto scalare otteniamo la (3.40). Se viceversa è $\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$ consideriamo

$$\left\| x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \right\|^2 = \left\langle x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k, x - \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \right\rangle.$$

Utilizzando la continuità del prodotto scalare e l'ortonormalità della successione $\{e_n\}$ otteniamo

$$\left\| x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$$

e quindi $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ come volevasi dimostrare. ■

Se una famiglia ortonormale (e_κ) , $\kappa \in I$, in uno spazio con prodotto scalare X non è numerabile (perché l'insieme di indici I non è numerabile) possiamo ancora formare i coefficienti di Fourier $\langle x, e_\kappa \rangle$ di un $x \in X$. In questo caso possiamo provare il rimarchevole Teorema seguente.

3.21 Teorema (Coefficienti di Fourier)

Un qualsiasi x in uno spazio con prodotto scalare X può avere, rispetto ad una famiglia ortonormale (e_κ) , $\kappa \in I$, in X , al più una quantità numerabile di coefficienti di Fourier $\langle x, e_\kappa \rangle$ diversi da zero.

Se, per un qualsiasi $x \in X$ considerato fisso, si ordinano gli e_κ con $\langle x, e_\kappa \rangle \neq 0$ in una successione ortonormale $(e_1, e_2, \dots)$ si possono considerare le quantità considerate nei paragrafi precedenti. In particolare vale la disuguaglianza di Bessel

$$\sum_{\kappa \in I} |\langle x, e_\kappa \rangle|^2 \leq \|x\|^2$$

e se X è uno spazio di Hilbert la serie

$$\sum_{\kappa \in I} \langle x, e_\kappa \rangle e_\kappa \tag{3.41}$$

converge. Inoltre la sua somma non dipende dall'ordine secondo cui gli e_κ con coefficiente di Fourier diverso da zero sono inseriti nella successione. ■

Dimostrazione. Per ciascun fissato $m = 1, 2, \dots$ il numero dei coefficienti di Fourier tali che $|\langle x, e_\kappa \rangle| > 1/m$ deve essere finito, perché, qualora ciò non fosse, vi sarebbe una successione ortonormale di e_κ per la quale la serie $\sum_{\kappa \in I} |\langle x, e_\kappa \rangle|^2$ sarebbe divergente in contraddizione colla disuguaglianza di Bessel (3.32) del Teorema 3.19. La convergenza della serie (3.41) segue dal Teorema 3.20. Sia ora (w_n) un riordinamento di (e_n) . Per definizione questo significa che v'è un'applicazione biiettiva $n \rightarrow m(n)$ di $\mathbb{N}$ in se stesso tale che i termini corrispondenti delle due successioni sono uguali, cioè $w_{m(n)} = e_n$. Poniamo

$$\alpha_n = \langle x, e_n \rangle, \quad \beta_m = \langle x, w_m \rangle$$

e

$$x_1 = \sum_{n=1}^{\infty} \alpha_n e_n, \quad x_2 = \sum_{m=1}^{\infty} \beta_m w_m.$$

Allora per il Teorema 3.20(b),

$$\alpha_n = \langle x, e_n \rangle = \langle x_1, e_n \rangle, \quad \beta_m = \langle x, w_m \rangle = \langle x_2, w_m \rangle.$$

Poiché $e_n = w_{m(n)}$ otteniamo

$$\begin{aligned} \langle x_1 - x_2, e_n \rangle &= \langle x_1, e_n \rangle - \langle x_2, w_{m(n)} \rangle \\ &= \langle x, e_n \rangle - \langle x, w_{m(n)} \rangle = 0 \end{aligned}$$

e analogamente $\langle x_1 - x_2, w_m \rangle = 0$. Ciò implica

$$\begin{aligned} \|x_1 - x_2\|^2 &= \langle x_1 - x_2, \sum_{n=1}^{\infty} \alpha_n e_n - \sum_{m=1}^{\infty} \beta_m w_m \rangle \\ &= \sum_{n=1}^{\infty} \bar{\alpha}_n \langle x_1 - x_2, e_n \rangle - \sum_{m=1}^{\infty} \bar{\beta}_m \langle x_1 - x_2, w_m \rangle = 0. \end{aligned}$$

Di conseguenza $x_1 - x_2 = 0$. Poiché il riordino (w_m) di (e_n) era arbitrario ciò completa la dimostrazione. ■

3.9 Basi Ortonormali

Gli insiemi ortonormali veramente interessanti negli spazi con prodotto scalare e negli spazi di Hilbert sono quelli che consistono di un numero “sufficientemente grande” di elementi perché ogni elemento dello spazio possa essere rappresentato o approssimato con sufficiente accuratezza utilizzando questi insiemi ortonormali. Per gli spazi finito dimensionali n -dimensionali la situazione è semplice; tutto ciò di cui abbiamo bisogno è un insieme ortonormale di n elementi. Il problema è quello di stabilire come si debba tenere conto degli spazi infinito dimensionali. I concetti principali sono i seguenti.

3.22 DEFINIZIONE (BASE ORTONORMALE)

Si dice *base ortonormale* in uno spazio normato X un insieme ortonormale $M \subset X$ (successione o famiglia) il cui span è denso in X , ossia tale che

$$\overline{\text{span } M} = X. \quad \blacksquare$$

È importante notare che, a meno che X non sia finito dimensionale, la base qui definita non è una base nel senso dell'algebra così come è definita in 2.3, dove X è considerato esclusivamente come uno spazio vettoriale.

Ogni spazio di Hilbert $H \neq \{0\}$ ammette una base ortonormale.

Per uno spazio finito dimensionale H ciò è chiaro. Per uno spazio H infinito dimensionale separabile (cf. 1.6) ciò segue dal procedimento di Gram-Schmidt per induzione (ordinaria). Per uno spazio H non separabile una prova (non costruttiva) segue dal lemma di Zorn, come vedremo nella Sez. 4.2, dove introduciamo e spieghiamo il lemma per altri fini.

Tutti le basi ortonormali in un dato spazio di Hilbert $H \neq \{0\}$ hanno la medesima cardinalità. Quest'ultima è chiamata la dimensione di Hilbert o la dimensione ortogonale di H . (Se $H = \{0\}$ questa dimensione è definita come 0).

Per un H finito dimensionale l'affermazione è evidente perché allora la dimensione di Hilbert è la dimensione nel senso dell'algebra. Per un H infinito dimensionale separabile l'affermazione seguirà facilmente dal Teorema 3.25 (qui di seguito), mentre per un generico H la dimostrazione richiede degli strumenti più avanzati dalla teoria degli insiemi.

Il seguente teorema mostra che una base ortonormale non può essere aumentata a costituire un insieme ortonormale più esteso aggiungendo dei nuovi elementi.

3.23 Teorema (Criterio del complemento ortogonale)

Sia M un insieme ortonormale di uno spazio di Hilbert H . Allora M è una base se e solo se $M^\perp = \{0\}$. ■

Dimostrazione. Segue immediatamente dal Lemma 3.16. ■

Un altro importante criterio perché un insieme ortonormale M sia una base si desume dal seguente teorema.

3.24 Teorema (Criterio di Parseval)

Un insieme ortonormale M in uno spazio di Hilbert H è una base se e solo se per tutti gli $x \in H$ vale la relazione di Parseval

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2, \quad (\text{relazione di Parseval}) \quad (3.42)$$

dove nel caso M sia non numerabile la somma si intende estesa su tutti i coefficienti di Fourier non nulli di x rispetto a M . ■

Dimostrazione. (a) Se M non è una base per il Teorema 3.23 v'è un $x \perp M$ in X non nullo. Poiché $x \perp M$ in (3.42) abbiamo $\langle x, e_k \rangle = 0$ per tutti i k , così che il membro a sinistra in (3.42) è nullo mentre $\|x\|^2 \neq 0$. Ciò mostra che la (3.42) non vale. Quindi se la (3.42) vale per tutti gli $x \in M$ allora M deve essere una base in H .

(b) Viceversa si assuma che M sia una base in H . Si consideri un qualunque $x \in H$ ed i suoi coefficienti di Fourier non nulli ordinati in una successione $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \dots$, o scritti in un definito ordine se sono in numero finito. Definiamo ora y mediante

$$y = \sum_k \langle x, e_k \rangle e_k \quad (3.43)$$

notando che nel caso di una serie infinita la convergenza segue dal Teorema 3.20. Mostriamo che $x - y \perp M$. Per ogni e_j che appare in (3.43) abbiamo, usando l'ortonormalità,

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_k \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0.$$

Inoltre per ogni $v \in M$ non contenuto in (3.43) abbiamo $\langle x, v \rangle = 0$, così che

$$\langle x - y, v \rangle = \langle x, v \rangle - \sum_k \langle x, e_k \rangle \langle e_k, v \rangle = 0 - 0 = 0.$$

Quindi $x - y \perp M$, cioè $x - y \in M^\perp$. Poiché M è una base in H abbiamo da 3.16 che $M^\perp = \{0\}$. Di conseguenza $x - y = 0$, cioè $x = y$. Quindi per il punto (d) del Teorema 3.20 abbiamo

$$\|x\|^2 = \sum_k |\langle x, e_k \rangle|^2.$$

Ciò completa la dimostrazione. ■

Passiamo ora a considerare gli spazi di Hilbert che sono separabili. Per la Def. 1.6 un tale spazio contiene un sottoinsieme numerabile che è denso nello spazio. Gli spazi di Hilbert separabili sono più semplici di quelli non separabili perché non possono contenere insiemi ortonormali non numerabili secondo quanto indicato dal seguente teorema.

3.25 Teorema (Spazi di Hilbert Separabili)

Sia H uno spazio di Hilbert. Ne segue

- (a) se H è separabile ogni insieme ortonormale in H è numerabile,
 (b) se H contiene una successione ortonormale che è una base in H allora H è separabile. ■

Dimostrazione. (a) Sia H separabile e sia B un qualunque insieme numerabile denso in H ed M un qualunque insieme ortonormale. Allora due qualunque elementi distinti x e y di M hanno distanza $\sqrt{2}$ perché

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle = 2.$$

Quindi gli intorni sferici N_x di x e N_y di y di raggio $\sqrt{2}/3$ sono disgiunti. Poiché B è denso in H v'è un $b_x \in B$ in N_x ed un $b_y \in B$ in N_y e $b_x \neq b_y$ perché $N_x \cap N_y = \emptyset$. Quindi se M fosse non numerabile avremmo un insieme non numerabile di intorni sferici a due a due disgiunti (per ciascun $x \in M$ uno di essi), così che B sarebbe non numerabile contraddicendo la separabilità di H . Da ciò concludiamo che M deve essere numerabile.

(b) Sia (e_k) una base ortonormale in H ed A l'insieme di tutte le combinazioni lineari

$$\gamma_1^{(n)} e_1 + \cdots + \gamma_n^{(n)} e_n \quad n = 1, 2, \dots$$

dove $\gamma_k^{(n)} = a_k^{(n)} + i b_k^{(n)}$ e $a_k^{(n)}$ e $b_k^{(n)}$ sono razionali (e $b_k^{(n)} = 0$ se H è reale). Chiaramente A è numerabile. Proviamo che A è denso in H mostrando che per ogni $x \in H$ ed $\varepsilon > 0$ v'è un $v \in A$ tale che $\|x - v\| < \varepsilon$.

Poiché la successione (e_k) è una base in H v'è un n tale che $Y_n = \text{span}\{e_1, \dots, e_n\}$ contiene un punto

$$y = \sum_{k=1}^n \alpha_k e_k$$

la cui distanza da x è minore di $\varepsilon/2$. Quindi abbiamo

$$\left\| x - \sum_{k=1}^n \alpha_k e_k \right\| < \frac{\varepsilon}{2}.$$

Poiché i razionali sono densi su $\mathbb{R}$ per ciascun α_k v'è un $\gamma_k^{(n)}$ (con parte reale ed immaginaria razionale) tale che

$$\left| \alpha_k - \gamma_k^{(n)} \right| < \frac{\varepsilon}{2n}$$

ed abbiamo allora

$$\left\| \sum_{k=1}^n [\alpha_k - \gamma_k^{(n)}] e_k \right\| < \sum_{k=1}^n \left| \alpha_k - \gamma_k^{(n)} \right| < \frac{\varepsilon}{2}.$$

Quindi $v \in A$ definito da

$$v = \sum_{k=1}^n \gamma_k^{(n)} e_k$$

soddisfa

$$\begin{aligned} \|x - v\| &= \left\| x - \sum \gamma_k^{(n)} e_k \right\| \\ &\leq \left\| x - \sum \langle x, e_k \rangle e_k \right\| + \left\| \sum \langle x, e_k \rangle e_k - \sum \gamma_k^{(n)} e_k \right\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Ciò prova che A è denso in H e quindi, poiché A è numerabile, H è separabile. ■

Per concludere questa sezione vogliamo sottolineare che la nostra presente discussione ha alcune conseguenze di importanza basilare che possono esser formulate in termini di isomorfismo di spazi di Hilbert.

Il fatto più straordinario in questa discussione è che due spazi astratti di Hilbert sul medesimo campo sono distinguibili solo per le loro dimensioni di Hilbert, una situazione che generalizza quella degli spazi euclidei. Questo è il significato del seguente teorema.

3.26 Teorema (Isomorfismo e Dimensione di Hilbert)

Due spazi di Hilbert H e $\tilde{H}$, entrambi reali o complessi, sono isomorfi se e solo se hanno la medesima dimensione di Hilbert. ■

Dimostrazione. (a) Se H è isomorfo ad $\tilde{H}$ e $T : H \rightarrow \tilde{H}$ è un isomorfismo e quindi

$$\langle Tx, Ty \rangle = \langle x, y \rangle,$$

allora gli elementi ortonormali in H hanno immagini ortonormali sotto T . Sia quindi M una base ortonormale di H e mostriamo che l'insieme ortonormale immagine $\tilde{M} = T(M)$ è una base in $\tilde{H}$ della medesima cardinalità. Sia $V = \text{span } M$. Poiché T è lineare $T(V) = \text{span } \tilde{M}$, poiché T è limitato e quindi continuo $T(\overline{V}) = \overline{T(V)}$ ed infine, essendo $\overline{V} = H$ perché M è una base, poiché T è biiettivo concludiamo che $\overline{T(V)} = T(H) = \tilde{H}$, cioè T applica in modo biiettivo ogni base ortonormale in H in una base ortonormale in $\tilde{H}$. Quindi H e $\tilde{H}$ hanno le medesime dimensioni di Hilbert.

(b) Viceversa supponiamo che H e $\tilde{H}$ abbiano le medesime dimensioni di Hilbert. Il caso $H = \{0\}$ e $\tilde{H} = \{0\}$ è banale. Sia $H \neq \{0\}$. Allora $\tilde{H} \neq \{0\}$ ed due basi qualunque ortonormali, M in H ed $\tilde{M}$ in $\tilde{H}$, hanno la medesima cardinalità, così che possiamo indicizzarle col medesimo insieme di indici $\{\kappa\}$ e scrivere $M = (e_\kappa)$ e $\tilde{M} = (\tilde{e}_\kappa)$.

Per mostrare che H e $\tilde{H}$ sono isomorfi costruiamo un isomorfismo di H su $\tilde{H}$. Per ogni $x \in H$ abbiamo

$$x = \sum_{\kappa} \langle x, e_\kappa \rangle e_\kappa \quad (3.44)$$

dove il membro a destra è una somma finita o una serie infinita (cf. 3.21) e $\sum_{\kappa} |\langle x, e_\kappa \rangle|^2 < \infty$ per la disuguaglianza di Bessel. Introduciamo la serie

$$\tilde{x} = Tx = \sum_{\kappa} \langle x, e_\kappa \rangle \tilde{e}_\kappa. \quad (3.45)$$

Essa è convergente per il 3.20 e quindi definisce un $\tilde{x} \in \tilde{H}$. L'operatore T è lineare perché il prodotto scalare è lineare rispetto al primo fattore. T è isometrico perché usando dapprima (3.45) e poi (3.44) otteniamo

$$\|\tilde{x}\|^2 = \|Tx\|^2 = \sum_{\kappa} |\langle x, e_\kappa \rangle|^2 = \|x\|^2.$$

Da ciò e da (3.7), (3.8) nella Sez. 3.2 vediamo che T conserva il prodotto scalare. Inoltre l'isometria implica l'iniettività. Infatti se $Tx = Ty$ allora

$$\|x - y\| = \|T(x - y)\| = \|Tx - Ty\| = 0,$$

così che $x = y$ e T è iniettivo per il 2.25.

Mostriamo infine che T è surgettivo. Dato un qualunque

$$\tilde{x} = \sum_{\kappa} \alpha_{\kappa} \tilde{e}_{\kappa}$$

in $\tilde{H}$ abbiamo che $\sum |\alpha_{\kappa}|^2 < \infty$ per la disuguaglianza di Bessel. Quindi

$$\sum_{\kappa} \alpha_{\kappa} e_{\kappa}$$

è una somma finita o una serie che converge ad un $x \in H$ per il 3.20 ed $\alpha_{\kappa} = \langle x, e_{\kappa} \rangle$ per il medesimo teorema. Abbiamo perciò $\tilde{x} = Tx$ per la (3.45). Poiché $\tilde{x} \in \tilde{H}$ era arbitrario ciò mostra che T è surgettivo. ■

3.10 Rappresentazione di Funzionali su Spazi di Hilbert

È di importanza pratica conoscere la forma generale dei funzionali lineari limitati su vari spazi. Ciò è stato messo in rilievo e spiegato nella Sez. 2.11. Per spazi di Banach generici queste formule e la loro derivazione possono a volte essere complicate. Tuttavia per uno spazio di Hilbert la situazione è sorprendentemente semplice.

3.27 Teorema (Riesz. Funzionali su Spazi di Hilbert)

Ogni funzionale lineare limitato f su uno spazio di Hilbert H può essere rappresentato in termini di un prodotto scalare e precisamente come segue

$$f(x) = \langle x, z \rangle \quad (3.46)$$

dove z dipende da f , è univocamente determinato da f ed ha norma

$$\|z\| = \|f\|. \quad (3.47)$$

■

Dimostrazione. Proviamo che

(a) f ha una rappresentazione (3.46),

(b) z in (3.46) è unico,

(c) vale la formula (3.47).

I dettagli sono i seguenti.

(a) Se $f = 0$ allora (3.46) e (3.47) sono validi se prendiamo $z = 0$. Sia $f \neq 0$. Per motivare l'idea della dimostrazione chiediamoci che proprietà deve avere z se la rappresentazione (3.46) esiste. Prima di tutto $z \neq 0$ perché altrimenti $f = 0$. Inoltre $\langle x, z \rangle = 0$ per tutti gli x per cui $f(x) = 0$, cioè per tutti gli x nello spazio nullo $\mathcal{N}(f)$ di f . Quindi $z \perp \mathcal{N}(f)$. Ciò suggerisce di considerare $\mathcal{N}(f)$ ed il suo complemento ortogonale $\mathcal{N}(f)^{\perp}$.

$\mathcal{N}(f)$ è uno spazio vettoriale per il 2.24 ed è chiuso per il 2.32. Inoltre $f \neq 0$ implica $\mathcal{N}(f) \neq H$, così che $\mathcal{N}(f)^\perp \neq \{0\}$ per il teorema della proiezione 3.14. Quindi $\mathcal{N}(f)^\perp$ contiene un $z_0 \neq 0$. Poniamo

$$v = f(x)z_0 - f(z_0)x$$

dove $x \in H$ è arbitrario. Applicando f otteniamo

$$f(v) = f(x)f(z_0) - f(z_0)f(x) = 0.$$

Ciò mostra che $v \in \mathcal{N}(f)$. Poiché $z_0 \perp \mathcal{N}(f)$ abbiamo

$$\begin{aligned} 0 &= \langle v, z_0 \rangle = \langle f(x)z_0 - f(z_0)x, z_0 \rangle \\ &= f(x)\langle z_0, z_0 \rangle - f(z_0)\langle x, z_0 \rangle. \end{aligned}$$

Notando che $\langle z_0, z_0 \rangle = \|z_0\|^2 \neq 0$ possiamo risolvere rispetto a $f(x)$. Il risultato è

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle} \langle x, z_0 \rangle.$$

Questo può essere riscritto nella forma (3.46) dove

$$z = \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0.$$

Poiché $x \in H$ era arbitrario la (3.46) è provata.

(b) Proviamo che z nella (3.46) è unico. Supponiamo che per tutti gli $x \in H$

$$f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle.$$

Allora $\langle x, z_1 - z_2 \rangle = 0$ per tutti gli x . Scegliendo in particolare $x = z_1 - z_2$ abbiamo

$$\langle x, z_1 - z_2 \rangle = \langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0.$$

Quindi $z_1 - z_2 = 0$, ossia l'unicità.

(c) Infine proviamo la (3.47). Se $f = 0$ allora $z = 0$ e la (3.47) è valida. Sia $f \neq 0$. Allora $z \neq 0$. Dalla (3.46) con $x = z$ e dalla (2.24) nella Sez. 2.9 otteniamo

$$\|z\|^2 = \langle z, z \rangle = f(z) \leq \|f\| \|z\|.$$

Dividendo per $\|z\| \neq 0$ si ottiene $\|z\| \leq \|f\|$. Rimane da mostrare che $\|f\| \leq \|z\|$. Dalla (3.46) e dalla disuguaglianza di Schwarz vediamo che

$$|f(x)| = |\langle x, z \rangle| \leq \|x\| \|z\|.$$

Ciò implica

$$\|f\| = \sup_{\|x\|=1} |\langle x, z \rangle| \leq \|z\|. \quad \blacksquare$$

Si noti che la corrispondenza $f \longleftrightarrow z$ fra H ed il suo duale H' è biunivoca, conserva la norma ma è coniugata lineare.

L'idea utilizzata nella dimostrazione dell'unicità nella parte (b) merita di essere formulata nella forma di un lemma, che mette in evidenza come il prodotto scalare sia in grado di discriminare fra due vettori diversi.

3.28 Lemma (Eguaglianza)

Se $\langle v_1, w \rangle = \langle v_2, w \rangle$ per tutti i w in uno spazio con prodotto scalare X allora $v_1 = v_2$. In particolare $\langle v_1, w \rangle = 0$ per tutti i $w \in X$ implica $v_1 = 0$. ■

Dimostrazione. Per ipotesi per tutti i w

$$\langle v_1 - v_2, w \rangle = \langle v_1, w \rangle - \langle v_2, w \rangle = 0.$$

Per $w = v_1 - v_2$ ciò dà $\|v_1 - v_2\|^2 = 0$. Quindi $v_1 - v_2 = 0$. In particolare $\langle v_1, w \rangle = 0$ con $w = v_1$ dà $\|v_1\|^2 = 0$, così che $v_1 = 0$. ■

L'importanza pratica dei funzionali lineari limitati sugli spazi di Hilbert risulta in larga misura dalla semplicità della rappresentazione di Riesz (3.46).

Inoltre (3.46) è molto importante nella teoria degli operatori sugli spazi di Hilbert. In particolare per quanto riguarda l'operatore aggiunto di Hilbert T^* di un operatore lineare limitato T che definiremo nella prossima sezione. A questo scopo abbiamo bisogno di una preparazione che è anche di interesse generale. Cominciamo con la seguente definizione.

3.29 DEFINIZIONE (FORME SESQUILINEARI)

Siano X e Y spazi vettoriali sul medesimo campo K ($\mathbb{R}$ o $\mathbb{C}$). Allora una *forma sesquilineare* (o *funzionale sesquilineare*) h su $X \times Y$ è un'applicazione

$$h : X \times Y \rightarrow K$$

tale che per tutti gli $x, x_1, x_2 \in X$ e $y, y_1, y_2 \in Y$ e tutti gli scalari α, β

$$h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y) \quad (3.48)$$

$$h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2) \quad (3.49)$$

$$h(\alpha x, y) = \alpha h(x, y) \quad (3.50)$$

$$h(x, \beta y) = \overline{\beta} h(x, y). \quad (3.51)$$

Quindi h è *lineare* nel primo argomento e *coniugato lineare* nel secondo argomento. Se X e Y sono reali ($K = \mathbb{R}$) allora (3.51) diviene semplicemente

$$h(x, \beta y) = \beta h(x, y)$$

e h è chiamata *bilineare* perché è lineare in entrambi gli argomenti.

Se X e Y sono spazi normati e se esiste un numero reale c tale che per tutti x, y

$$|h(x, y)| \leq c \|x\| \|y\|, \quad (3.52)$$

allora h è detta *limitata* ed il numero

$$\|h\| = \sup_{\substack{x \in X - \{0\} \\ y \in Y - \{0\}}} \frac{|h(x, y)|}{\|x\| \|y\|} = \sup_{\substack{\|x\|=1 \\ \|y\|=1}} |h(x, y)| \quad (3.53)$$

è chiamato la *norma* di h . ■

Ad esempio il prodotto scalare è sesquilineare e limitato.

Si noti che dalla (3.52) e dalla (3.53) abbiamo che

$$|h(x, y)| \leq \|h\| \|x\| \|y\|. \quad (3.54)$$

Entrambi i termini “forma” e “funzionale” sono comuni. Forse è preferibile usare il termine “forma” in questo caso a due variabili e riservare il termine “funzionale” al caso ad una variabile. È ciò che noi faremo nel seguito.

È molto interessante che dal Teorema 3.27 si possa ottenere una generale rappresentazione delle forme sesquilineari su spazi di Hilbert come segue.

3.30 Teorema (Rappresentazione di Riesz)

Siano H_1, H_2 spazi di Hilbert e

$$h : H_1 \times H_2 \rightarrow K$$

una forma sesquilineare limitata. Allora h ha la rappresentazione

$$h(x, y) = \langle Sx, y \rangle \quad (3.55)$$

dove $S : H_1 \rightarrow H_2$ è un operatore lineare limitato. S è unicamente determinato da h ed ha norma

$$\|S\| = \|h\|. \quad (3.56)$$

■

Dimostrazione. Consideriamo $\overline{h(x, y)}$. È lineare in y a causa della barra. Per poter applicare il Teorema 3.27 manteniamo x fisso. Allora questo teorema ci dà una rappresentazione in cui y è la variabile, ossia

$$\overline{h(x, y)} = \langle y, z \rangle.$$

Quindi

$$h(x, y) = \langle z, y \rangle. \quad (3.57)$$

Qui $z \in H_2$ è unico ma dipende naturalmente dal nostro $x \in H_1$ fisso. Ne segue che la (3.57) con la variabile x definisce un operatore

$$S : H_1 \rightarrow H_2 \quad \text{dato da} \quad z = Sx.$$

Sostituendo $z = Sx$ nella (3.57) abbiamo la (3.55).

S è lineare. Infatti il suo dominio è lo spazio vettoriale H_1 e dalla (3.55) e dalla sesquilinearità otteniamo

$$\begin{aligned} \langle S(\alpha x_1 + \beta x_2), y \rangle &= h(\alpha x_1 + \beta x_2, y) \\ &= \alpha h(x_1, y) + \beta h(x_2, y) \\ &= \alpha \langle Sx_1, y \rangle + \beta \langle Sx_2, y \rangle \\ &= \langle \alpha Sx_1 + \beta Sx_2, y \rangle \end{aligned}$$

per tutti gli y in H_2 , così che per il Lemma 3.28

$$S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2.$$

S è limitato. Infatti lasciando da parte il caso banale $S = 0$ abbiamo dalla (3.53) e dalla (3.55)

$$\|h\| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \geq \sup_{\substack{x \neq 0 \\ Sx \neq 0}} \frac{|\langle Sx, Sx \rangle|}{\|x\| \|Sx\|} = \sup_{x \neq 0} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

Ciò prova la limitatezza. Inoltre è $\|h\| \geq \|S\|$.

Otteniamo ora la (3.56) notando che $\|h\| \leq \|S\|$ segue dall'applicazione della disuguaglianza di Schwarz

$$\|h\| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \leq \sup_{x \neq 0} \frac{\|Sx\| \|y\|}{\|x\| \|y\|} = \|S\|.$$

S è unico. Infatti se assumiamo che esista un operatore lineare $T : H_1 \rightarrow H_2$ tale che per tutti gli $x \in H_1$ e $y \in H_2$ si abbia

$$h(x, y) = \langle Sx, y \rangle = \langle Tx, y \rangle,$$

vediamo che per il Lemma 3.28 $Sx = Tx$ per tutti gli $x \in H_1$. Quindi $S = T$ per definizione. ■

3.11 Operatori Aggiunti di Hilbert

I risultati della sezione precedente ci permettono ora di introdurre l'operatore aggiunto di Hilbert di un operatore lineare limitato su uno spazio di Hilbert. Questo operatore è stato suggerito dai problemi sulle matrici e sulle equazioni differenziali e integrali. Vedremo che esso aiuta anche a definire tre importanti classi di operatori (chiamati *operatori auto-aggiunti*, *unitari* e *normali*), che sono state ampiamente studiate perché giocano un ruolo chiave in varie applicazioni.

3.31 DEFINIZIONE (OPERATORE AGGIUNTO DI HILBERT T^*)

Sia $T : H_1 \rightarrow H_2$ un operatore lineare limitato, dove H_1 e H_2 sono spazi di Hilbert. Allora l'*operatore aggiunto di Hilbert* T^* di T è l'operatore

$$T^* : H_2 \rightarrow H_1$$

tale che² per tutti gli $x \in H_1$ e $y \in H_2$

$$\langle Tx, y \rangle = \langle x, T^*y \rangle. \quad (3.58)$$

■

Naturalmente dobbiamo prima mostrare che questa definizione ha senso, dobbiamo cioè provare che per un dato T un tale T^* esiste.

3.32 Teorema (Esistenza)

L'operatore aggiunto di Hilbert T^ di T esiste, è unico ed è un operatore lineare limitato con norma*

$$\|T^*\| = \|T\|. \quad (3.59)$$

■

Dimostrazione. La formula

$$h(y, x) = \langle y, Tx \rangle \quad (3.60)$$

²Possiamo indicare i prodotti scalari su H_1 e H_2 col medesimo simbolo perché i fattori mostrano a quale spazio il prodotto scalare si riferisce.

definisce una forma sesquilineare su $H_2 \times H_1$ perché il prodotto scalare è sesquilineare e T è lineare. La linearità coniugata della forma si verifica direttamente come segue

$$\begin{aligned} h(y, \alpha x_1 + \beta x_2) &= \langle y, T(\alpha x_1 + \beta x_2) \rangle \\ &= \langle y, \alpha T x_1 + \beta T x_2 \rangle \\ &= \overline{\alpha} \langle y, T x_1 \rangle + \overline{\beta} \langle y, T x_2 \rangle \\ &= \overline{\alpha} h(y, x_1) + \overline{\beta} h(y, x_2). \end{aligned}$$

h è limitato. Infatti per la disuguaglianza di Schwarz

$$|h(y, x)| = |\langle y, T x \rangle| \leq \|y\| \|T x\| \leq \|y\| \|T\| \|x\|.$$

Ciò implica anche $\|h\| \leq \|T\|$. Inoltre abbiamo $\|h\| \geq \|T\|$ da

$$\|h\| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|\langle y, T x \rangle|}{\|y\| \|x\|} \geq \sup_{\substack{x \neq 0 \\ T x \neq 0}} \frac{|\langle T x, T x \rangle|}{\|T x\| \|x\|} = \|T\|.$$

Dal confronto

$$\|h\| = \|T\|. \quad (3.61)$$

Il Teorema 3.30 fornisce una rappresentazione di Riesz per h ; scrivendo T^* per S abbiamo

$$h(y, x) = \langle T^* y, x \rangle \quad (3.62)$$

e sappiamo da questo teorema che $T^* : H_2 \rightarrow H_1$ è un operatore lineare limitato unicamente determinato con norma [cf. (3.61)]

$$\|T^*\| = \|h\| = \|T\|.$$

Ciò prova la (3.59). Inoltre $\langle y, T x \rangle = \langle T^* y, x \rangle$ confrontando (3.60) e (3.62), così che abbiamo la (3.58) prendendo i coniugati e concludiamo che T^* è effettivamente l'operatore che stavamo cercando. ■

Nel nostro studio delle proprietà degli operatori aggiunti di Hilbert sarà utile fare uso del seguente lemma.

3.33 Lemma (Operatore Nullo)

Siano X e Y spazi con prodotto scalare e $Q : X \rightarrow Y$ un operatore lineare. Ne segue

(a) $Q = 0$ se e solo se $\langle Q x, y \rangle = 0$ per tutti gli $x \in X$ e $y \in Y$.

(b) Se $Q : X \rightarrow X$, dove X è complesso, e $\langle Q x, x \rangle = 0$ per tutti gli $x \in X$ allora $Q = 0$. ■

Dimostrazione. (a) $Q = 0$ significa $Q x = 0$ per tutti gli x ed implica

$$\langle Q x, y \rangle = \langle 0, y \rangle = 0.$$

Viceversa $\langle Q x, y \rangle = 0$ per tutti gli x e y implica $Q x = 0$ per tutti gli x per il 3.28, così che $Q = 0$ per definizione.

(b) Per ipotesi $\langle Q v, v \rangle = 0$ per ogni $v = \alpha x + y \in X$, ossia

$$\begin{aligned} 0 &= \langle Q(\alpha x + y), \alpha x + y \rangle \\ &= |\alpha|^2 \langle Q x, x \rangle + \langle Q y, y \rangle + \alpha \langle Q x, y \rangle + \overline{\alpha} \langle Q y, x \rangle. \end{aligned}$$

I primi due termini a destra sono zero per ipotesi. $\alpha = 1$ dà

$$\langle Qx, y \rangle + \langle Qy, x \rangle = 0.$$

$\alpha = i$ dà

$$\langle Qx, y \rangle - \langle Qy, x \rangle = 0.$$

Sommando $\langle Qx, y \rangle = 0$ e $Q = 0$ segue dalla (a). ■

Nella parte (b) di questo lemma è essenziale che X sia complesso. Infatti la conclusione può non essere valida se X è reale. Un controesempio è la rotazione Q del piano $\mathbb{R}^2$ di un angolo retto. Q è lineare e $Qx \perp x$, quindi $\langle Qx, x \rangle = 0$ per tutti gli $x \in \mathbb{R}^2$, ma $Q \neq 0$.

Possiamo elencare e provare alcune proprietà generali degli operatori aggiunti di Hilbert che si usano frequentemente nelle applicazioni di questi operatori.

3.34 Teorema (Proprietà degli Operatori Aggiunti di Hilbert)

Siano H_1 e H_2 spazi di Hilbert, $S : H_1 \rightarrow H_2$ e $T : H_1 \rightarrow H_2$ operatori lineari limitati ed α uno scalare qualunque. Allora abbiamo

$$\langle T^*y, x \rangle = \langle y, Tx \rangle \quad (x \in H_1, y \in H_2) \quad (3.63)$$

$$(S + T)^* = S^* + T^* \quad (3.64)$$

$$(\alpha T)^* = \overline{\alpha} T^* \quad (3.65)$$

$$(T^*)^* = T \quad (3.66)$$

$$\|T^*T\| = \|TT^*\| = \|T\|^2 \quad (3.67)$$

$$T^*T = 0 \iff T = 0 \quad (3.68)$$

e assumendo $S : H_1 \rightarrow H_2$ e $T : H_2 \rightarrow H_3$

$$(TS)^* = S^*T^*. \quad (3.69)$$

■

Dimostrazione. (a) Dalla (3.58) abbiamo la (3.63)

$$\langle T^*y, x \rangle = \overline{\langle x, T^*y \rangle} = \overline{\langle Tx, y \rangle} = \langle y, Tx \rangle.$$

(b) Per la (3.58) per tutti gli x e y

$$\begin{aligned} \langle x, (S + T)^*y \rangle &= \langle (S + T)x, y \rangle \\ &= \langle Sx, y \rangle + \langle Tx, y \rangle \\ &= \langle x, S^*y \rangle + \langle x, T^*y \rangle \\ &= \langle x, (S^* + T^*)y \rangle. \end{aligned}$$

Quindi $(S + T)^*y = (S^* + T^*)y$ per tutti gli y per il 3.28 che è la (3.64) per definizione.

(c) La formula (3.65) non deve essere confusa colla formula $T^*(\alpha x) = \alpha T^*x$. È ottenuta dal seguente calcolo e dalla successiva applicazione del Lemma 3.33(a) a $Q = (\alpha T)^* - \overline{\alpha} T^*$.

$$\begin{aligned} \langle (\alpha T)^*y, x \rangle &= \langle y, (\alpha T)x \rangle \\ &= \langle y, \alpha(T)x \rangle \\ &= \overline{\alpha} \langle y, Tx \rangle \\ &= \overline{\alpha} \langle T^*y, x \rangle \\ &= \langle \overline{\alpha} T^*y, x \rangle. \end{aligned}$$

(d) $(T^*)^*$ è scritto T^{**} ed è uguale a T perché per tutti gli $x \in H_1$ ed $y \in H_2$ abbiamo dalla (3.63) ed (3.58)

$$\langle (T^*)^* x, y \rangle = \langle x, T^* y \rangle = \langle Tx, y \rangle$$

e la (3.66) segue dal Lemma 3.33(a) per $Q = (T^*)^* - T$.

(e) Notiamo che $T^*T : H_1 \rightarrow H_1$ ma $TT^* : H_2 \rightarrow H_2$. Per la disuguaglianza di Schwarz

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leq \|T^*Tx\| \|x\| \leq \|T^*T\| \|x\|^2.$$

Prendendo l'estremo superiore per tutti gli x di norma 1 otteniamo $\|T\|^2 \leq \|T^*T\|$. Applicando la (2.20), Sez. 2.8, e la (3.59) abbiamo così

$$\|T\|^2 \leq \|T^*T\| \leq \|T^*\| \|T\| = \|T\|^2.$$

Quindi $\|T^*T\| = \|T\|^2$. Sostituendo T con T^* ed usando di nuovo la (3.59) abbiamo anche

$$\|T^{**}T^*\| = \|T^*\|^2 = \|T\|^2.$$

Ora $T^{**} = T$ per la (3.66), così che la (3.67) è provata.

(f) Dalla (3.67) otteniamo immediatamente la (3.68).

(g) Un'applicazione ripetuta della (3.58) dà

$$\langle x, (ST)^* y \rangle = \langle (ST)x, y \rangle = \langle Tx, S^* y \rangle = \langle x, T^* S^* y \rangle.$$

Quindi $(ST)^* y = T^* S^* y$ per il 3.28 che è la (3.69) per definizione. ■

3.12 Operatori Autoaggiunti, Unitari e Normali

Classi di operatori lineari limitati di grande importanza pratica possono essere definiti utilizzando l'operatore aggiunto di Hilbert.

3.35 DEFINIZIONE (OPERATORI AUTOAGGIUNTI, UNITARI E NORMALI)

Un operatore lineare limitato $T : H \rightarrow H$ su uno spazio di Hilbert H è detto

<i>autoaggiunto</i> o <i>hermitiano</i> se	$T^* = T$	■
<i>anti-hermitiano</i> se	$T^* = -T$	
<i>unitario</i> se T è biiettivo e	$T^* = T^{-1}$	
<i>normale</i> se	$TT^* = T^*T$.	

L'operatore aggiunto di Hilbert T^* di T è definito dalla (3.58), Sez. 3.11, ossia

$$\langle Tx, y \rangle = \langle x, T^* y \rangle.$$

Se T è autoaggiunto vediamo che la formula diviene

$$\langle Tx, y \rangle = \langle x, Ty \rangle. \quad (3.70)$$

Se T è *autoaggiunto*, *anti-hermitiano* o *unitario*, T è *normale*.

Ciò può essere immediatamente visto dalla definizione. Naturalmente un operatore normale non è necessariamente autoaggiunto o unitario.

I termini usati nella Def. 3.35 sono anche usati in connessione con le matrici. Vogliamo spiegare le ragioni di ciò e menzionare alcune importanti relazioni.

Esempio (Matrici). Consideriamo $\mathbb{C}^n$ col prodotto scalare definito da

$$\langle x, y \rangle = x^\top \overline{y}, \quad (3.71)$$

dove x e y sono scritti come vettori colonna e $\top$ significa trasposto; allora $x^\top = (\xi_1, \dots, \xi_n)$ ed usiamo l'ordinaria moltiplicazione fra matrici.

Sia $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ un operatore lineare (che è limitato per il Teorema 2.30). Essendo data una base per $\mathbb{C}^n$ possiamo rappresentare T ed il suo aggiunto di Hilbert T^* con due matrici quadrate a n righe, siano A e B rispettivamente.

Usando la (3.71) e la regola familiare $(Bx)^\top = x^\top B^\top$ per il trasposto di un prodotto otteniamo

$$\langle Tx, y \rangle = (Ax)^\top \overline{y} = x^\top A^\top \overline{y}$$

e

$$\langle x, T^*y \rangle = x^\top \overline{B^*y}.$$

Per la (3.58), Sez. 3.11, i membri a sinistra sono uguali per tutti gli $x, y \in \mathbb{C}^n$. Quindi dobbiamo avere $A^\top = \overline{B}$. Di conseguenza

$$B = \overline{A}^\top.$$

Il risultato è il seguente.

Se è data una base per $\mathbb{C}^n$ ed un operatore lineare su $\mathbb{C}^n$ è rappresentato da una certa matrice, allora il suo operatore aggiunto di Hilbert è rappresentato dal trasposto complesso coniugato di questa matrice.

Pertanto una matrice quadrata $A = (\alpha_{jk})$ è detta

$$\begin{aligned} \text{hermitiana se } \overline{A}^\top &= A & (\text{quindi } \overline{\alpha}_{kj} &= \alpha_{jk}) \\ \text{anti-hermitiana se } \overline{A}^\top &= -A & (\text{quindi } \overline{\alpha}_{kj} &= -\alpha_{jk}) \\ \text{unitaria se } \overline{A}^\top &= A^{-1} \\ \text{normale se } A\overline{A}^\top &= \overline{A}^\top A. \end{aligned}$$

Invece una matrice quadrata reale $A = (\alpha_{jk})$ è detta

$$\begin{aligned} \text{(reale) simmetrica se } A^\top &= A & (\text{quindi } \alpha_{kj} &= \alpha_{jk}) \\ \text{(reale) anti-simmetrica se } A^\top &= -A & (\text{quindi } \alpha_{kj} &= -\alpha_{jk}) \\ \text{ortogonale se } A^\top &= A^{-1}. \end{aligned}$$

Quindi una matrice reale hermitiana è una matrice (reale) simmetrica. Una matrice reale anti-hermitiana è una matrice (reale) anti-simmetrica. Una matrice reale unitaria è una matrice ortogonale. ■

Ritorniamo agli operatori lineari su un arbitrario spazio di Hilbert e enunciamo un importante e piuttosto semplice criterio per la proprietà di autoaggiunto.

3.36 Teorema (Autoaggiunto)

Sia $T : H \rightarrow H$ un operatore lineare limitato su uno spazio di Hilbert H . Ne segue

- (a) *Se T è autoaggiunto $\langle Tx, x \rangle$ è reale per tutti gli $x \in H$.*
- (b) *Se H è complesso e $\langle Tx, x \rangle$ è reale per tutti gli $x \in H$ l'operatore T è autoaggiunto. ■*

Dimostrazione. (a) Se T è autoaggiunto allora per tutti gli x

$$\overline{\langle Tx, x \rangle} = \langle x, Tx \rangle = \langle Tx, x \rangle.$$

Quindi $\langle Tx, x \rangle$ è uguale al suo complesso coniugato, cioè è reale.

(b) Se $\langle Tx, x \rangle$ è reale per tutti gli x allora

$$\langle Tx, x \rangle = \overline{\langle Tx, x \rangle} = \overline{\langle x, T^*x \rangle} = \langle T^*x, x \rangle$$

e $T - T^* = 0$ per il Lemma 3.33(b) poiché H è complesso. ■

Nella parte (b) del teorema è essenziale che H sia complesso. Ciò risulta chiaro dal fatto che per un H reale il prodotto scalare ha valori reali, ciò che rende $\langle Tx, x \rangle$ reale senza bisogno di alcuna ulteriore assunzione sull'operatore lineare T .

Prodotti (composizioni) di operatori autoaggiunti appaiono spesso nelle applicazioni, così che il seguente teorema risulterà essere utile.

3.37 Teorema (Autoaggiunto di un Prodotto)

Il prodotto di due operatori lineari limitati autoaggiunti S e T su uno spazio di Hilbert H è autoaggiunto se e solo gli operatori commutano

$$ST = TS. \quad \blacksquare$$

Dimostrazione. Per la (3.69) nella Sez. 3.11 e per le ipotesi

$$(ST)^* = T^*S^* = TS.$$

Quindi

$$ST = (ST)^* \iff ST = TS.$$

Ciò completa la dimostrazione. ■

Successioni di operatori autoaggiunti capitano in diversi problemi e per esse abbiamo il seguente teorema.

3.38 Teorema (Successioni di Operatori Autoaggiunti)

Sia (T_n) una successione di operatori lineari limitati autoaggiunti $T_n : H \rightarrow H$ su uno spazio di Hilbert H . Supponiamo che (T_n) converga, ossia,

$$T_n \rightarrow T, \quad \text{cioè} \quad \|T_n - T\| \rightarrow 0,$$

dove $\|\cdot\|$ è la norma sullo spazio $B(H, H)$; cf. Sez. 2.8. Allora l'operatore limite T è un operatore lineare limitato autoaggiunto su H . ■

Dimostrazione. Dobbiamo mostrare che $T^* = T$ ossia che $\|T - T^*\| = 0$ od anche, equivalentemente, per la continuità della norma che $\lim_{n \rightarrow \infty} \|T_n - T^*\| = 0$. Per provarlo basta osservare che, essendo gli operatori T_n autoaggiunti, per il 3.34 ed il 3.32 abbiamo

$$\|T_n - T^*\| = \|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|.$$

Quindi poiché $\lim_{n \rightarrow \infty} \|T_n - T\| = 0$ è $\|T - T^*\| = 0$ e $T^* = T$. ■

Questi teoremi ci danno un'idea delle proprietà basilari degli operatori lineari autoaggiunti. Essi saranno anche utili nel nostro lavoro successivo, in particolare nella teoria spettrale di questi operatori, dove discuteremo ulteriori proprietà.

Ritorniamo ora agli operatori unitari e consideriamo alcune delle loro proprietà basilari.

3.39 Teorema (Operatori Unitari)

Siano gli operatori $U : H \rightarrow H$ e $V : H \rightarrow H$ unitari su uno spazio di Hilbert H . Allora

- (a) U è isometrico; quindi $\|Ux\| = \|x\|$ per tutti gli $x \in H$;
- (b) $\|U\| = 1$, purché $H \neq \{0\}$,
- (c) $U^{-1} (= U^*)$ è unitario,
- (d) UV è unitario,
- (e) U è normale.

Inoltre

- (f) Un operatore lineare limitato T su uno spazio di Hilbert complesso H è unitario se e solo se T è isometrico e surgettivo. ■

Dimostrazione. (a) può essere visto da

$$\|Ux\|^2 = \langle Ux, Ux \rangle = \langle x, U^*Ux \rangle = \langle x, Ix \rangle = \|x\|^2.$$

- (b) segue immediatamente da (a).
- (c) Poiché U è biiettivo lo è anche U^{-1} e per il 3.34

$$(U^{-1})^* = U^{**} = U = (U^{-1})^{-1}.$$

- (d) UV è biiettiva e 3.34 e 2.26 danno

$$(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}.$$

- (e) segue da $U^{-1} = U^*$ e $UU^{-1} = U^{-1}U = I$.

(f) Supponiamo che T sia isometrico e surgettivo. L'isometria implica l'iniettività, così che T è biiettivo. Mostriamo che $T^* = T^{-1}$. Per l'isometria

$$\langle T^*Tx, x \rangle = \langle Tx, Tx \rangle = \langle x, x \rangle = \langle Ix, x \rangle.$$

Quindi

$$\langle (T^*T - I)x, x \rangle = 0$$

e $T^*T - I = 0$ per il Lemma 3.33(b), così che $T^*T = I$. Da ciò

$$TT^* = TT^*(TT^{-1}) = T(T^*T)T^{-1} = TIT^{-1} = I.$$

In conclusione $TT^* = T^*T = I$. Quindi $T^* = T^{-1}$ e T è unitario. L'inverso è immediato perché T è isometrico per (a) e surgettivo per definizione. ■

Si noti che un operatore isometrico non è necessariamente unitario perché potrebbe non essere surgettivo. Un esempio è l'operatore di traslazione a destra $T : l^2 \rightarrow l^2$ dato da

$$(\xi_1, \xi_2, \xi_3, \dots) \mapsto (0, \xi_1, \xi_2, \xi_3, \dots)$$

dove $x = (\xi_j) \in l^2$.

3.13 Problemi

1. Mostrare che ℓ^2 ammette il prodotto scalare

$$\langle x, y \rangle = \sum_{j=1}^{\infty} \xi_j \bar{\eta}_j$$

e che, dotato di tale prodotto scalare, ℓ^2 è uno spazio di Hilbert.

2. Mostrare che ℓ^p con $p \neq 2$ non ammette prodotto scalare.

Suggerimento: Si considerino i due elementi di ℓ^p , $x = (1, 1, 0, 0, \dots)$ ed $y = (1, -1, 0, 0, \dots)$ e si mostri che essi non soddisfano all'eguaglianza del parallelogramma.

3. Mostrare che $C[a, b]$ non ammette un prodotto scalare.

Capitolo 4

Teoremi Fondamentali per gli Spazi Normati e di Banach

Questo capitolo contiene, si può dire, le basi di una teoria più avanzata degli spazi normati e di Banach senza le quali l'utilità di questi spazi e le loro applicazioni sarebbero piuttosto limitate. I quattro teoremi importanti in questo capitolo sono il teorema di Hahn–Banach, il teorema della uniforme limitatezza, il teorema dell'applicazione aperta ed il teorema del grafico chiuso.

4.1 Breve Orientamento sui Contenuti Principali

1. *Teorema di Hahn–Banach* 4.8 (e varianti 4.9, 4.10). Questo è un teorema di estensione per i funzionali lineari sugli spazi vettoriali. Garantisce che uno spazio normato è sufficientemente dotato di funzionali lineari, perché si ottenga una teoria adeguata degli spazi duali così come una teoria soddisfacente degli operatori aggiunti (Sez. 4.6, 4.7).

2. *Teorema di uniforme limitatezza* 4.24 di Banach e Steinhaus. Questo teorema fornisce le condizioni sufficienti perché la successione $(\|T_n\|)$ sia limitata, essendo i T_n operatori lineari limitati da uno spazio di Banach in uno spazio normato. Ha diverse applicazioni (semplici e più profonde) in analisi, ad esempio in connessione con le serie di Fourier, la convergenza debole, la sommabilità di successioni, l'integrazione numerica, etc.

3. *Teorema dell'applicazione aperta* 4.39. Questo teorema stabilisce che un operatore lineare limitato T da uno spazio di Banach su un altro spazio di Banach è un'applicazione aperta, che cioè applica insiemi aperti su insiemi aperti. Quindi se T è biiettivo T^{-1} è continuo (“teorema dell'inverso limitato”).

4. *Teorema del grafico chiuso* 4.42. Questo teorema dà le condizioni sotto le quali un operatore lineare chiuso (cf. 4.40) è limitato. Gli operatori lineari chiusi sono importanti in fisica ed in altre applicazioni.

4.2 Lemma di Zorn

Avremo bisogno del lemma di Zorn nella dimostrazione del teorema fondamentale di Hahn–Banach. Il lemma ha a che fare con gli insiemi parzialmente ordinati.

4.1 DEFINIZIONE (INSIEME PARZIALMENTE ORDINATO)

Un *insieme parzialmente ordinato* è un insieme M su cui è definito un *ordine parziale*, cioè una relazione binaria che si scrive $\leq$ e soddisfa le condizioni

- (PO1) $a \leq a$ per ogni $a \in M$. (Riflessività)
 (PO2) Se $a \leq b$ e $b \leq a$ allora $a = b$. (Antisimmetria)
 (PO3) Se $a \leq b$ e $b \leq c$ allora $a \leq c$. (Transitività)

■

La parola “parzialmente” mette in rilievo il fatto che M può contenere elementi a e b per cui non vale né $a \leq b$ né $b \leq a$. Allora a e b sono chiamati *elementi inconfrontabili*. Al contrario due elementi a e b sono chiamati *elementi confrontabili* se soddisfano a $a \leq b$ o a $b \leq a$ (o ad entrambi).

La relazione d'ordine $\leq$ sopra introdotta è anche detta relazione d'ordine debole, mentre se $a \leq b$ e $a \neq b$ si specifica che la relazione d'ordine fra a e b è forte e si scrive $a < b$.

4.2 DEFINIZIONE (CATENA)

Un *insieme totalmente ordinato* o *catena* è un insieme parzialmente ordinato tale che due elementi qualunque dell'insieme sono confrontabili. In altre parole una catena è un insieme parzialmente ordinato che non ha elementi inconfrontabili.

■

4.3 DEFINIZIONE (LIMITE SUPERIORE)

Un *limite superiore* di un sottoinsieme W di un insieme parzialmente ordinato M è un elemento $u \in M$ tale che

$$x \leq u \quad \text{per ogni } x \in W.$$

■

A seconda degli insiemi M e W che si considerano un tale u può esistere o non esistere.

4.4 DEFINIZIONE (ELEMENTO MASSIMALE)

Un elemento m di M è detto *elemento massimale* se

$$\text{non esiste } x \in M \text{ tale che } m < x.$$

■

Di nuovo M può avere o non avere elementi massimali. Si noti inoltre che un elemento massimale non è necessariamente un limite superiore, perché sarà $m \geq x$ solamente per gli elementi x di M confrontabili con m , che non sono necessariamente tutti gli elementi di M .

Possiamo ora formulare il lemma di Zorn, che consideriamo come un'assioma.¹

4.5 Lemma (Zorn)

Sia $M \neq \emptyset$ un insieme parzialmente ordinato. Supponiamo che ogni catena $C \subset M$ abbia un limite superiore. Allora M ha almeno un elemento massimale.

■

Nemmeno per un analista è indispensabile conoscere la dimostrazione del lemma di Zorn, che si può fare a partire dall'assioma della scelta. È invece essenziale comprenderne bene l'enunciato e saperlo utilizzare. Il lemma di Zorn ha applicazioni numerose e molto importanti in analisi; in particolare è uno strumento indispensabile per stabilire alcuni teoremi di esistenza. Può essere considerato un principio di induzione transfinita che estende cioè il principio di induzione ad insiemi non numerabili.

¹ Il nome “lemma” è per ragioni storiche. Il lemma di Zorn può essere derivato dall'**assioma della scelta**, che stabilisce che per ogni dato insieme E esiste un'applicazione c (“funzione di scelta”) dall'insieme potenza $\mathcal{P}(E)$ in E tale che se $B \subset E$, $B \neq \emptyset$, allora $c(B) \in B$. Viceversa questo assioma segue dal lemma di Zorn, così che il lemma di Zorn e l'assioma della scelta possono essere considerati come assiomi equivalenti.

4.3 Alcune Applicazioni del Lemma di Zorn

4.6 Teorema (Base di Hamel)

Ogni spazio vettoriale $X \neq \{0\}$ ha una base di Hamel. (Cf. Sez. 2.1.) ■

Dimostrazione. Sia M l'insieme di tutti i sottoinsiemi linearmente indipendenti di X . Poiché $X \neq \{0\}$ v'è un elemento $x \neq 0$ e $\{x\} \in M$, così che $M \neq \emptyset$. L'inclusione fra insiemi definisce un ordine parziale su M .

Cominciamo col dimostrare che l'unione dei sottoinsiemi costituenti una catena, grazie alla relazione d'ordine totale esistente in una catena, è un sottoinsieme linearmente indipendente. Dobbiamo mostrare che ogni suo sottoinsieme finito $\{v_1, v_2, \dots, v_n\}$ è linearmente indipendente. Vi sono degli elementi della catena $B_1, B_2, \dots, B_n$ tali che $v_i \in B_i$ per ogni $i = 1, 2, \dots, n$ e poiché C è una catena $B_i \subseteq B_j$ o $B_j \subseteq B_i$ per ciascuna coppia di elementi e poichè i B sono in numero finito uno di essi deve contenere tutti gli altri e quindi tutti i v , che sono perciò indipendenti.

Quindi ogni catena $C \subset M$ ha un estremo superiore e, precisamente, l'unione di tutti i sottoinsiemi di X che sono elementi di C . Per il lemma di Zorn M ha un elemento massimale B . Mostriamo che B è una base di Hamel per X . Sia $Y = \text{span } B$. Allora Y è un sottospazio di X e $Y = X$ perché altrimenti $B \cup \{z\}$, $z \in X$, $z \notin Y$ sarebbe un insieme linearmente indipendente contenente B come un sottoinsieme proprio, in contrasto con la massimalità di B . ■

4.7 Teorema (Base Ortonormale)

Ogni spazio di Hilbert $H \neq \{0\}$ ammette una base ortonormale. (Cf. Sez. 3.9.) ■

Dimostrazione. Sia M l'insieme di tutti i sottoinsiemi ortonormali di H . Poiché $H \neq \{0\}$ v'è un elemento $x \neq 0$ ed un sottoinsieme ortonormale di H è $\{y\}$ dove $y = \|x\|^{-1}x$. Quindi $M \neq \emptyset$. L'inclusione fra insiemi definisce un ordine parziale su M . Possiamo utilizzare la relazione d'ordine totale esistente in una catena per dedurre, in maniera analoga a quanto fatto nel teorema 4.6, che l'unione dei sottoinsiemi costituenti una catena è un sottoinsieme ortonormale. Quindi ogni catena $C \subset M$ ha un estremo superiore e precisamente l'unione di tutti i sottoinsiemi di H che sono elementi di C . Per il lemma di Zorn M ha un elemento massimale F . Proviamo che F è una base in H . Supponiamo che ciò sia falso. Allora per il Teorema 3.23 esiste un $z \in H$ non nullo tale che $z \perp F$. Quindi $F_1 = F \cup \{e\}$, dove $e = \|z\|^{-1}z$, è ortonormale ed F è un sottoinsieme proprio di F_1 . Ciò contraddice la massimalità di F . ■

4.4 Teorema di Hahn–Banach

Il teorema di Hahn–Banach è un teorema di estensione per i funzionali lineari. Vedremo nella prossima sezione che il teorema garantisce che uno spazio normato è riccamente fornito di funzionali lineari limitati, ciò che rende possibile un'adeguata teoria degli spazi duali, che è una parte essenziale della teoria generale degli spazi normati. In questo modo il teorema di Hahn–Banach diventa uno dei teoremi più importanti in connessione con gli operatori lineari limitati.

Generalmente parlando, in un *problema di estensione*, si considera un oggetto matematico (per esempio un'applicazione) definito su un sottoinsieme Z di un dato insieme X e si vuole estendere l'oggetto da Z all'intero insieme X in modo tale che alcune proprietà basilari dell'oggetto continuino a valere per l'oggetto esteso.

Nel teorema di Hahn–Banach l'oggetto da estendere è un funzionale lineare f che è definito su un sottospazio Z di uno spazio vettoriale X reale ed ha una certa proprietà di

limitatezza che è formulata in termini di un **funzionale sublineare**. Per definizione esso è un funzionale p a valori reali su uno spazio vettoriale, che è **subadditivo**, ossia tale che

$$p(x + y) \leq p(x) + p(y) \quad \text{per tutti gli } x, y \in X, \quad (4.1)$$

e **positivo-omogeneo**, ossia tale che

$$p(\alpha x) = \alpha p(x) \quad \text{per tutti gli } \alpha \geq 0 \text{ in } \mathbb{R} \text{ e } x \in X. \quad (4.2)$$

(Si noti che la *norma* di uno spazio normato è un funzionale di questo tipo.)

Assumeremo che il funzionale f che deve essere esteso sia maggiorato da un tale funzionale p definito su X ed estenderemo f da Z a X senza perdere la linearità e la maggiorazione, così che il funzionale esteso $\tilde{f}$ su X è ancora lineare ed ancora maggiorato da p . Questo è il punto centrale del teorema. Una generalizzazione del teorema agli spazi vettoriali complessi e normati è riportata nella sezione successiva.

4.8 Teorema (Hahn–Banach. Estensione di Funzionali Lineari)

Sia X uno spazio vettoriale reale e p un funzionale sublineare su X . Inoltre sia f un funzionale lineare definito su un sottospazio Z di X e tale che

$$f(x) \leq p(x) \quad \text{per tutti gli } x \in Z. \quad (4.3)$$

Allora f ammette una estensione lineare $\tilde{f}$ da Z a X tale che

$$\tilde{f}(x) \leq p(x) \quad \text{per tutti gli } x \in X, \quad (4.4)$$

■

Dimostrazione. Procedendo passo a passo proveremo

(a) L'insieme E di tutte le estensioni lineari g di f dal sottospazio $\mathcal{D}(f)$ al sottospazio $\mathcal{D}(g)$ che soddisfano $g(x) \leq p(x)$ per ogni $x \in \mathcal{D}(g)$ possono essere parzialmente ordinate ed il lemma di Zorn garantisce l'esistenza di un elemento massimale $\tilde{f}$ di E .

(b) $\tilde{f}$ è definito sull'intero spazio X .

(c) Una relazione ausiliaria che è stata usata in (b).

Cominciamo con la parte (a).

(a) Sia E l'insieme di tutte le estensioni lineari g di f che soddisfano la condizione

$$g(x) \leq p(x) \quad \text{per tutti gli } x \in \mathcal{D}(g).$$

Chiaramente $E \neq \emptyset$ perché $f \in E$. Su E possiamo definire un ordine parziale con

$$g \leq h \quad \text{significa} \quad h \text{ è un estensione di } g,$$

cioè per definizione $\mathcal{D}(h) \supset \mathcal{D}(g)$ ed $h(x) = g(x)$ per ogni $x \in \mathcal{D}(g)$.

Per qualsiasi catena $C \subset E$ definiamo ora $\hat{g}$ con

$$\hat{g}(x) = g(x) \quad \text{se } x \in \mathcal{D}(g) \quad (g \in C).$$

$\hat{g}$ è un funzionale lineare con dominio

$$\mathcal{D}(\hat{g}) = \bigcup_{g \in C} \mathcal{D}(g),$$

che è uno spazio vettoriale perché C è una catena. La definizione di $\hat{g}$ non è ambigua. Infatti per un $x \in \mathcal{D}(g_1) \cap \mathcal{D}(g_2)$ con $g_1, g_2 \in C$ abbiamo $g_1(x) = g_2(x)$ perché C è una catena,

così che $g_1 \leq g_2$ o $g_2 \leq g_1$. Chiaramente $g \leq \hat{g}$ per tutti gli $g \in C$. Quindi $\hat{g}$ è un estremo superiore di C . Poiché $C \subset E$ era arbitrario il lemma di Zorn implica che E ha un elemento massimale $\tilde{f}$. Per definizione di E questo è una estensione lineare di f che soddisfa

$$\tilde{f}(x) \leq p(x) \quad x \in \mathcal{D}(\tilde{f}). \quad (4.5)$$

(b) Mostriamo ora che $\mathcal{D}(\tilde{f})$ è tutto X . Supponiamo che ciò sia falso. Allora possiamo scegliere un $y_1 \in X - \mathcal{D}(\tilde{f})$ e considerare il sottospazio Y_1 di X generato da $\mathcal{D}(\tilde{f})$ ed y_1 . Si noti che $y_1 \neq 0$ perché $0 \in \mathcal{D}(\tilde{f})$. Qualsiasi $x \in Y_1$ può essere scritto

$$x = y + \alpha y_1 \quad y \in \mathcal{D}(\tilde{f}).$$

Questa rappresentazione è unica. Infatti $y + \alpha y_1 = \tilde{y} + \beta y_1$ con $\tilde{y} \in \mathcal{D}(\tilde{f})$ implica $y - \tilde{y} = (\beta - \alpha)y_1$, dove $y - \tilde{y} \in \mathcal{D}(\tilde{f})$ mentre $y_1 \notin \mathcal{D}(\tilde{f})$, così che la sola soluzione è $y - \tilde{y} = 0$ e $\beta - \alpha = 0$. Ciò significa l'unicità.

Un funzionale g_1 su Y_1 è definito da

$$g_1(y + \alpha y_1) = \tilde{f}(y) + \alpha c \quad (4.6)$$

dove c è una qualsiasi costante reale. Non è difficile vedere che g_1 è lineare. Inoltre per $\alpha = 0$ abbiamo $g_1(y) = \tilde{f}(y)$. Quindi g_1 è una *estensione* di $\tilde{f}$ ed è *propria*, perché $\mathcal{D}(\tilde{f})$ è un sottoinsieme proprio di $\mathcal{D}(g_1)$. Conseguentemente se possiamo provare che, per un opportuno c , l'estensione $g_1 \in E$ è tale che

$$g_1(x) \leq p(x) \quad \text{per tutti gli } x \in \mathcal{D}(g_1), \quad (4.7)$$

ciò contraddirà la massimalità di $\tilde{f}$, così che $\mathcal{D}(\tilde{f}) \neq X$ è falso e $\mathcal{D}(\tilde{f}) = X$ è vero.

(c) Per $\alpha = 0$ abbiamo $x \in \mathcal{D}(\tilde{f})$ e quindi $g_1(x) = \tilde{f}(x)$ e la (4.7) è automaticamente soddisfatta. Per $\alpha \neq 0$ la condizione (4.7) è equivalente alle due seguenti condizioni

$$\tilde{f}\left(\frac{y}{\alpha}\right) + c \leq p\left(\frac{y}{\alpha} + y_1\right) \quad \text{per } \alpha > 0 \quad (4.8)$$

$$\tilde{f}\left(-\frac{y}{\alpha}\right) - c \leq p\left(-\frac{y}{\alpha} - y_1\right) \quad \text{per } \alpha < 0. \quad (4.9)$$

Consideriamo un qualunque y e z in $\mathcal{D}(\tilde{f})$. Dalla (4.5) ed (4.1) otteniamo

$$\begin{aligned} \tilde{f}(y) - \tilde{f}(z) &= \tilde{f}(y - z) \leq p(y - z) \\ &= p(y + y_1 - y_1 - z) \\ &\leq p(y + y_1) + p(-y_1 - z). \end{aligned}$$

portando l'ultimo termine a sinistra ed il termine $\tilde{f}$ a destra abbiamo

$$-p(-y_1 - z) - \tilde{f}(z) \leq p(y + y_1) - \tilde{f}(y), \quad (4.10)$$

dove y_1 è fisso. Poiché y non appare a sinistra e z non appare a destra l'ineguaglianza continua a valere se prendiamo l'estremo superiore sugli $z \in \mathcal{D}(\tilde{f})$ a sinistra (chiamiamolo m_0) e l'estremo inferiore sugli $y \in \mathcal{D}(\tilde{f})$ a destra, chiamiamolo m_1 . Allora $m_0 \leq m_1$ e per un c tale che $m_0 \leq c \leq m_1$ abbiamo dalla (4.10)

$$-p(-y_1 - z) - \tilde{f}(z) \leq c \quad \text{per tutti gli } z \in \mathcal{D}(\tilde{f}) \quad (4.11)$$

$$c \leq p(y + y_1) - \tilde{f}(y) \quad \text{per tutti gli } y \in \mathcal{D}(\tilde{f}). \quad (4.12)$$

Sosostituendo in (4.11) z con y/α ed in (4.12) y con y/α otteniamo rispettivamente la (4.9) e (4.8) ed il teorema è dimostrato. ■

Possiamo fare a meno del lemma di Zorn? Questa domanda è di interesse in particolare perché il lemma non dà un metodo di costruzione. Se nella (4.6) prendiamo f invece di $\tilde{f}$ otteniamo per ciascun c reale una estensione lineare g_1 di f al sottospazio Z_1 generato da $\mathcal{D}(f) \cup \{y_1\}$ e possiamo scegliere c in modo tale che $g_1(x) \leq p(x)$ per tutti gli $x \in Z_1$, come si può vedere dalla parte (c) della dimostrazione con $\tilde{f}$ sostituito con f . Se $X = Z_1$ abbiamo concluso. Se $X \neq Z_1$ possiamo prendere un $y_2 \in X - Z_1$ e ripetere il processo per estendere f a Z_2 generato da Z_1 e y_2 , etc. Ciò fornisce una successione di sottospazi ciascuno contenente il precedente e tali che f può essere esteso da ciascuno al successivo e l'estensione g_j soddisfa $g_j(x) \leq p(x)$ per tutti gli $x \in Z_j$. Se

$$X = \bigcup_{j=1}^n Z_j$$

abbiamo concluso dopo n passi e se

$$X = \bigcup_{j=1}^{\infty} Z_j$$

possiamo usare l'ordinaria induzione. Tuttavia se X non ha una tale rappresentazione abbiamo bisogno del lemma di Zorn qui presentato.

Naturalmente per spazi speciali l'intera situazione può diventare più semplice. Gli spazi di Hilbert sono di questo tipo grazie alla rappresentazione di Riesz 3.27. Discuteremo questo fatto nella prossima sezione.

4.5 Estensioni del Teorema di Hahn–Banach agli Spazi Vettoriali Complessi e agli Spazi Normati

Il teorema di Hahn–Banach 4.8 concerne gli spazi vettoriali *reali*. Il seguente teorema è una generalizzazione che include gli spazi vettoriali complessi.

4.9 Teorema (Hahn–Banach. Generalizzato)

Sia X uno spazio vettoriale reale o complesso e p un funzionale a valori reali su X che sia subadditivo, cioè tale che per tutti gli $x, y \in X$ soddisfi

$$p(x + y) \leq p(x) + p(y) \quad (4.13)$$

(come nel Teorema 4.8) e per ogni scalare α

$$p(\alpha x) = |\alpha|p(x). \quad (4.14)$$

Inoltre sia f un funzionale lineare definito su un sottospazio Z di X e tale da soddisfare

$$|f(x)| \leq p(x) \quad \text{per tutti gli } x \in Z. \quad (4.15)$$

Allora f ha una estensione lineare $\tilde{f}$ da Z ad X che soddisfa

$$|\tilde{f}(x)| \leq p(x) \quad \text{per tutti gli } x \in X. \quad (4.16)$$

■

Dimostrazione. Osserviamo preliminarmente che la condizione (4.13) calcolata per $y = 0$ e per $y = -x$ fornisce rispettivamente

$$0 \leq p(0)$$

e

$$p(0) \leq p(x) + p(-x).$$

Quindi, usando la (4.14) per $\alpha = -1$, otteniamo che per ogni x è $p(x) \geq 0$.

Consideriamo ora separatamente i due casi, quello in cui X è uno spazio vettoriale reale e quello complesso.

(a) Spazio vettoriale reale. Se X è reale la situazione è semplice. Allora (4.15), essendo $p(x) \geq 0$, implica $f(x) \leq p(x)$ per tutti gli $x \in Z$. Quindi per il teorema di Hahn-Banach 4.8 v'è una estensione lineare $\tilde{f}$ da Z a X tale che

$$\tilde{f}(x) \leq p(x) \quad \text{per tutti gli } x \in X. \quad (4.17)$$

Da questa e dalla (4.14) otteniamo

$$-\tilde{f}(x) = \tilde{f}(-x) \leq p(-x) = |-1|p(x) = p(x),$$

ossia $-\tilde{f}(x) \leq p(x)$. Assieme alla (4.17) ciò prova la (4.16).

(b) Spazio vettoriale complesso. Sia X complesso. Allora anche Z è uno spazio vettoriale complesso. Quindi f è a valori complessi e possiamo scrivere

$$f(x) = f_1(x) + if_2(x) \quad x \in Z$$

dove f_1 e f_2 sono a valori reali. Per il momento consideriamo X e Z come spazi vettoriali reali e li indichiamo con X_r e Z_r rispettivamente; ciò significa semplicemente che restringiamo la moltiplicazione per scalari ai numeri reali (invece che ai numeri complessi). Poiché f è lineare su Z e f_1 e f_2 sono a valori reali, f_1 e f_2 sono funzionali lineari su Z_r . Inoltre $f_1(x) \leq |f(x)|$ perché la parte reale di un numero complesso non può essere maggiore del suo modulo. Quindi per la (4.15)

$$f_1(x) \leq p(x) \quad \text{per tutti gli } x \in Z.$$

Per il punto **(a)** v'è un'estensione $\tilde{f}_1$ di f_1 da Z_r a X_r tale che

$$\tilde{f}_1(x) \leq p(x) \quad \text{per tutti gli } x \in X. \quad (4.18)$$

Ciò per quanto riguarda f_1 . Occupiamoci ora di f_2 . Ritornando a Z ed usando $f = f_1 + if_2$, abbiamo per ogni $x \in Z$

$$i[f_1(x) + if_2(x)] = if(x) = f(ix) = f_1(ix) + if_2(ix).$$

Le parti reali dei due membri devono essere uguali

$$f_2(x) = -f_1(ix) \quad x \in Z. \quad (4.19)$$

Quindi se per tutti gli $x \in X$ poniamo

$$\tilde{f}(x) = \tilde{f}_1(x) - i\tilde{f}_1(ix) \quad x \in X, \quad (4.20)$$

vediamo dalla (4.19) che $\tilde{f}(x) = f(x)$ su Z . Ciò mostra che $\tilde{f}$ è un'estensione di f da Z a X . Ci rimane da dimostrare che

- (i) $\tilde{f}$ è un funzionale lineare sullo spazio vettoriale *complesso* X ,
(ii) $\tilde{f}$ soddisfa la (4.16) su X .

Che la (i) sia valida può essere visto dal seguente calcolo che usa la (4.20) e la linearità della f_1 sullo spazio vettoriale reale X_r , dove $a + ib$ con a e b reali è un qualunque scalare complesso,

$$\begin{aligned}\tilde{f}((a + ib)x) &= \tilde{f}_1(ax + ibx) - i\tilde{f}_1(iax - bx) \\ &= a\tilde{f}_1(x) + b\tilde{f}_1(ix) - i[a\tilde{f}_1(ix) - b\tilde{f}_1(x)] \\ &= (a + ib)[\tilde{f}_1(x) - i\tilde{f}_1(ix)] \\ &= (a + ib)\tilde{f}(x).\end{aligned}$$

Proviamo la (ii). Per un qualsiasi x tale che $\tilde{f}(x) = 0$ ciò vale perché $p(x) \geq 0$. Sia x tale che $\tilde{f}(x) \neq 0$. Allora possiamo scrivere usando la forma polare delle quantità complesse

$$\tilde{f}(x) = |\tilde{f}(x)|e^{i\theta} \quad \text{e così} \quad |\tilde{f}(x)| = \tilde{f}(x)e^{-i\theta} = \tilde{f}(e^{-i\theta}x).$$

Poiché $|\tilde{f}(x)|$ è reale, l'ultima espressione è reale e quindi uguale alla sua parte reale. Quindi per la (2)

$$|\tilde{f}(x)| = \tilde{f}(e^{-i\theta}x) = \tilde{f}_1(e^{-i\theta}x) \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x).$$

Ciò completa la dimostrazione. ■

Sebbene il teorema di Hahn–Banach direttamente non dica nulla sulla continuità, un'applicazione del teorema fra le principali ha a che fare con i funzionali lineari limitati. Ciò ci riporta agli spazi normati che costituiscono il nostro interesse principale. Infatti il Teorema 4.9 implica il seguente teorema basilare.

4.10 Teorema (Hahn–Banach. Spazi Normati)

Sia f un funzionale lineare limitato su un sottospazio Z di uno spazio normato X . Allora esiste un funzionale lineare limitato $\tilde{f}$ su X che è un'estensione di f a X e che ha la stessa norma

$$\|\tilde{f}\|_X = \|f\|_Z \tag{4.21}$$

dove

$$\|\tilde{f}\|_X = \sup_{\substack{x \in X \\ \|x\|=1}} |\tilde{f}(x)|, \quad \|f\|_Z = \sup_{\substack{x \in Z \\ \|x\|=1}} |f(x)|$$

(e $\|f\|_Z = 0$ nel caso banale $Z = \{0\}$). ■

Dimostrazione. Se $Z = \{0\}$ allora $f = 0$ e l'estensione è $\tilde{f} = 0$. Sia $Z \neq \{0\}$. Vogliamo usare il Teorema 4.9. Quindi dobbiamo prima scoprire una p conveniente. Per tutti gli $x \in Z$ abbiamo

$$|f(x)| \leq \|f\|_Z \|x\|.$$

Questa equazione è della forma (4.15) dove

$$p(x) = \|f\|_Z \|x\|. \tag{4.22}$$

Vediamo che p è definito su tutto X . Inoltre p soddisfa la (4.13) su X poichè per la disuguaglianza triangolare

$$p(x+y) = \|f\|_Z \|x+y\| \leq \|f\|_Z (\|x\| + \|y\|) = p(x) + p(y).$$

p inoltre soddisfa la (4.14) su X perché

$$p(\alpha x) = \|f\|_Z \|\alpha x\| = |\alpha| \|f\|_Z \|x\| = |\alpha| p(x).$$

Quindi possiamo ora applicare il Teorema 4.9 e concludere che esiste un funzionale lineare $\tilde{f}$ su X che è un'estensione di f e che soddisfa

$$|\tilde{f}(x)| \leq p(x) = \|f\|_Z \|x\| \quad x \in X.$$

Prendendo l'estremo superiore su tutti gli $x \in X$ di norma 1 otteniamo la disuguaglianza

$$\|\tilde{f}\|_X = \sup_{\substack{x \in X \\ \|x\|=1}} |\tilde{f}(x)| \leq \|f\|_Z.$$

Poiché sotto un'estensione la norma non può decrescere abbiamo anche $\|\tilde{f}\|_X \geq \|f\|_Z$. Confrontando otteniamo la (4.21) ed il teorema è dimostrato. ■

In casi speciali la situazione può diventare molto semplice. Gli spazi di Hilbert sono di questo tipo. Infatti se Z è un sottospazio chiuso di uno spazio di Hilbert $X = H$, allora f ha una rappresentazione di Riesz 3.27, ossia

$$f(x) = \langle x, z \rangle \quad z \in Z$$

dove $\|z\| = \|f\|$. Naturalmente poichè il prodotto scalare è definito su tutto H ciò fornisce immediatamente una estensione lineare $\tilde{f}$ di f da Z a H , ed $\tilde{f}$ ha la stessa norma di f perché $\|\tilde{f}\| = \|z\| = \|f\|$ per il Teorema 3.27. Quindi in questo caso l'estensione è immediata.

Dal Teorema 4.10 deriveremo ora un altro utile risultato che, in parole povere, mostra che lo spazio duale X' di uno spazio normato X consiste di un numero sufficiente di funzionali lineari limitati da poter distinguere fra punti di X . Ciò diventerà essenziale in connessione con gli operatori aggiunti e la cosiddetta convergenza debole.

4.11 Teorema (Funzionali Lineari Limitati)

Sia X uno spazio normato e sia $x_0 \neq 0$ un elemento qualunque di X . Allora esiste un funzionale lineare limitato $\tilde{f}$ su X tale che

$$\|\tilde{f}\| = 1, \quad \tilde{f}(x_0) = \|x_0\|. \quad \blacksquare$$

Dimostrazione. Consideriamo il sottospazio Z di X consistente di tutti gli elementi $x = \alpha x_0$ dove α è uno scalare. Su Z definiamo un funzionale lineare con

$$f(x) = f(\alpha x_0) = \alpha \|x_0\|. \quad (4.23)$$

f è limitato ed ha norma $\|f\| = 1$ perché

$$|f(x)| = |f(\alpha x_0)| = |\alpha| \|x_0\| = \|x\|.$$

Il Teorema 4.10 implica che f ha una estensione lineare $\tilde{f}$ da Z a X di norma $\|\tilde{f}\| = \|f\| = 1$. Dalla (4.23) vediamo che $\tilde{f}(x_0) = f(x_0) = \|x_0\|$. ■

Da questo Teorema segue il seguente Corollario, che evidenzia come i funzionali lineari limitati siano in grado di discriminare fra due punti diversi.

4.12 Corollario (Norma, Vettore Nullo)

Per ogni x in uno spazio normato X abbiamo

$$\|x\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|}. \quad (4.24)$$

Quindi se x_0 è tale che $f(x_0) = 0$ per tutti gli $f \in X'$ allora $x_0 = 0$. ■

Dimostrazione. Dal Teorema 4.11 abbiamo, scrivendo x per x_0 ,

$$\sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} \geq \frac{|\tilde{f}(x)|}{\|\tilde{f}\|} = \frac{\|x\|}{1} = \|x\|$$

e da $|f(x)| \leq \|f\| \|x\|$ otteniamo

$$\sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} \leq \|x\|.$$

Dal confronto segue la (4.24). ■

4.6 Operatore Duale

Ad un operatore lineare limitato $T : X \rightarrow Y$ su uno spazio normato X possiamo associare il cosiddetto operatore duale (o coniugato) T' di T . Una motivazione per T' viene dalla sua utilità nella risoluzione delle equazioni che coinvolgono operatori; tali equazioni intervengono per esempio in fisica ed in altre applicazioni. Nella presente sezione definiamo l'operatore duale T' e consideriamo alcune delle sue proprietà, inclusa la sua relazione con l'operatore aggiunto di Hilbert T^* definito nella Sez. 3.11. È importante notare che la nostra presente discussione dipende dal Teorema di Hahn–Banach (attraverso il Teorema 4.11) e che senza di esso non andremmo molto lontano.

Consideriamo un operatore lineare limitato $T : X \rightarrow Y$, dove X e Y sono spazi normati, e vogliamo definire l'operatore duale T' di T . A questo scopo partiamo da un qualsiasi funzionale lineare limitato g su Y . Chiaramente g è definito per ogni $y \in Y$. Ponendo $y = Tx$ otteniamo un funzionale su X , che chiamiamo f ,

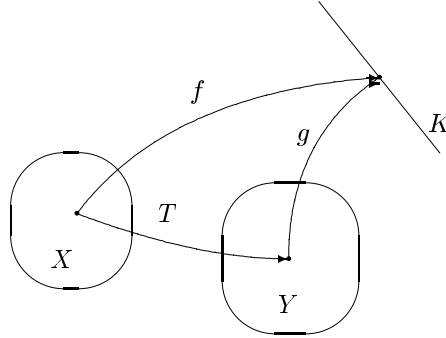
$$f(x) = g(Tx) \quad x \in X. \quad (4.25)$$

f è quello che in inglese si chiama il funzionale “pull back” di g , è cioè il funzionale che si ottiene rispingendo g all'indietro su X , utilizzando l'applicazione T (vedi figura 4.1). f è lineare dal momento che g e T sono lineari. f è limitato perché

$$|f(x)| = |g(Tx)| \leq \|g\| \|Tx\| \leq \|g\| \|T\| \|x\|.$$

Prendendo l'estremo superiore su tutti gli $x \in X$ di norma *uno* otteniamo la disuguaglianza

$$\|f\| \leq \|g\| \|T\|. \quad (4.26)$$

Figura 4.1: Pull back di g tramite T

Ciò mostra che $f \in X'$, dove X' è lo spazio duale di X definito in 2.41. Per ipotesi $g \in Y'$. Conseguentemente per $g \in Y'$ *variabile* la formula (4.25) definisce un operatore da Y' in X' , che è chiamato l'*operatore duale* di T ed è indicato con T' . Abbiamo così

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ X' & \xleftarrow{T'} & Y' \end{array} \quad (4.27)$$

Si porti particolare attenzione al fatto che T' è definito su Y' mentre l'operatore dato T è definito su X . Possiamo riassumere enunciando la seguente definizione.

4.13 DEFINIZIONE (OPERATORE DUALE T')

Sia $T : X \rightarrow Y$ un operatore lineare limitato, dove X e Y sono spazi normati. Allora l'*operatore duale* (o *coniugato*) $T' : Y' \rightarrow X'$ di T è definito da

$$f(x) = (T'g)(x) = g(Tx) \quad (g \in Y') \quad (4.28)$$

dove X' e Y' sono rispettivamente gli spazi duali di X e di Y . ■

Il nostro primo obiettivo è di provare che l'operatore duale ha la medesima norma dell'operatore stesso. Questa proprietà è basilare, come vedremo più innanzi. Per la dimostrazione avremo bisogno del Teorema 4.11, che è stato derivato dal teorema di Hahn–Banach. Il teorema di Hahn–Banach è perciò vitale per costruire una teoria soddisfacente degli operatori duali, che a loro volta sono una parte essenziale della teoria generale degli operatori lineari.

4.14 Teorema (Norma dell'Operatore Duale)

L'operatore duale T' nella Def. 4.13 è lineare, limitato e

$$\|T'\| = \|T\|. \quad (4.29)$$

■

Dimostrazione. L'operatore T' è lineare perché il suo dominio Y' è uno spazio vettoriale ed otteniamo facilmente

$$\begin{aligned} (T'(\alpha g_1 + \beta g_2))(x) &= (\alpha g_1 + \beta g_2)(Tx) \\ &= (\alpha g_1)(Tx) + (\beta g_2)(Tx) \\ &= \alpha(T'g_1)(x) + \beta(T'g_2)(x). \end{aligned}$$

Proviamo la (4.29). Dalla (4.28) abbiamo $f = T'g$ e per la (4.26) ne segue che

$$\|T'g\| = \|f\| \leq \|g\| \|T\|.$$

Prendendo l'estremo superiore su tutte le $g \in Y'$ di norma uno otteniamo la disuguaglianza

$$\|T'\| \leq \|T\|.$$

Quindi per ottenere la (4.29) dobbiamo ora provare che $\|T'\| \geq \|T\|$. Il Teorema 4.11 implica che per ogni $x \in X$ tale che $Tx \neq 0$ v'è un $\tilde{g} \in Y'$ tale che

$$\|\tilde{g}\| = 1 \quad \text{e} \quad \tilde{g}(Tx) = \|Tx\|. \quad (4.30)$$

Possiamo quindi scrivere, per $\|x\| \neq 0$,

$$\begin{aligned} \|T'\| &= \sup_{\substack{g \in Y' \\ \|g\| \neq 0}} \frac{\|T'g\|}{\|g\|} \geq \frac{\|T'\tilde{g}\|}{\|\tilde{g}\|} = \frac{\|\tilde{f}\| \|x\|}{\|x\|} \\ &\geq \frac{|\tilde{f}(x)|}{\|x\|} = \frac{|\tilde{g}(Tx)|}{\|x\|} = \frac{\|Tx\|}{\|x\|}, \end{aligned}$$

dove $\tilde{f} = T'\tilde{g}$ e sono state utilizzate le proprietà di $\tilde{g}$ in (4.30). Poiché x è generico, prendendo l'estremo superiore di entrambi i membri per ogni $x \neq 0$ otteniamo

$$\|T'\| \geq \|T\|$$

e quindi $\|T'\| = \|T\|$. ■

Illustriamo la presente discussione con le matrici considerate come rappresentanti di operatori.

Esempio (Matrici). Abbiamo mostrato nella Sez. 2.10 che, se X è uno spazio vettoriale n -dimensionale, una volta scelta una base $E = \{e_1, \dots, e_n\}$ per X , i punti x di X possono essere rappresentati da vettori colonna $x = (\xi_1, \dots, \xi_n)^t$ e gli operatori lineari $T: X \rightarrow X$ da matrici $T_E = (\tau_{jk})$ in modo tale che, se $y = Tx$ è rappresentato dal vettore colonna $y = (\eta_1, \dots, \eta_n)^t$ e se le matrici sono moltiplicate righe per colonne, allora

$$y = T_E x, \quad \text{od in componenti} \quad \eta_j = \sum_{k=1}^n \tau_{jk} \xi_k, \quad (4.31)$$

dove $j = 1, \dots, n$. Sia $F = \{f_1, \dots, f_n\}$ la base duale di E (cf. Sez. 2.10) ossia tale che

$$f_j(e_i) = \delta_{ij}. \quad (4.32)$$

Questa è una base per X' (che è anche esso uno spazio n -dimensionale). Sia ora $f = T'g$ ed abbiano g ed $f \in X'$ la rappresentazione

$$g = \alpha_1 f_1 + \dots + \alpha_n f_n \quad (4.33)$$

$$f = \beta_1 f_1 + \dots + \beta_n f_n. \quad (4.34)$$

Ricordando che le componenti ξ_k per e_i valgono $\xi_k = \delta_{ki}$ ($k = 1, 2, \dots, n$), abbiamo dalla (4.31)

$$Te_i = \sum_{j=1}^n \tau_{ji} e_j.$$

Dalla definizione di operatore duale otteniamo

$$f(e_i) = g(Te_i) = \sum_{j=1}^n \tau_{ji} g(e_j).$$

Dalla rappresentazione di g ed f in (4.33) e (4.34) abbiamo, utilizzando la (4.32),

$$\begin{aligned} g(e_j) &= \alpha_j \\ f(e_i) &= \beta_i \end{aligned}$$

e quindi

$$\beta_i = \sum_{j=1}^n \tau_{ji} \alpha_j.$$

Possiamo ora interpretare questa equazione come l'espressione in componenti di $f = T'_E g$. Osservando che nel membro a destra si somma rispetto al primo indice (così che si somma su tutti gli elementi di una *colonna* di T_E) abbiamo il seguente risultato:

Se T è rappresentato da una matrice T_E allora l'operatore duale T' è rappresentato dal trasposto di T_E . ■

Lavorando con gli operatori aggiunti sono utili le formule seguenti dalla (4.35) alla (4.37). La corrispondente prova è lasciata al lettore. Siano $S, T \in B(X, Y)$; cf. Sez. 2.8. Allora

$$(S + T)' = S' + T' \quad (4.35)$$

$$(\alpha T)' = \alpha T'. \quad (4.36)$$

Siano X, Y, Z spazi normati e $T \in B(X, Y)$ e $S \in B(Y, Z)$. Allora per l'operatore duale del prodotto ST abbiamo

$$(ST)' = T' S'. \quad (4.37)$$

Se $T \in B(X, Y)$ e T^{-1} esiste e $T^{-1} \in B(Y, X)$ allora anche $(T')^{-1}$ esiste, $(T')^{-1} \in B(X', Y')$ e

$$(T')^{-1} = (T^{-1})'. \quad (4.38)$$

4.6.1 Relazione fra l'Operatore Duale T' e l'Operatore Aggiunto di Hilbert T^* .

Mostriamo che una tale relazione esiste nel caso di un operatore lineare limitato $T : X \rightarrow Y$ se X e Y sono spazi di Hilbert, cioè $X = H_1$ e $Y = H_2$. In questo caso abbiamo

$$\begin{aligned} H_1 &\xrightarrow{T} H_2 \\ H'_1 &\xleftarrow{T'} H'_2 \end{aligned} \quad (4.39)$$

dove come prima l'operatore duale T' dell'operatore dato T è definito da

$$f = T'g \quad (4.40)$$

$$f(x) = g(Tx) \quad (f \in H'_1, g \in H'_2). \quad (4.41)$$

La novità è che poiché f e g sono funzionali su spazi di Hilbert hanno rappresentazioni di Riesz (cf. 3.27), ossia

$$f(x) = \langle x, x_0 \rangle \quad (x_0 \in H_1) \quad (4.42)$$

$$g(y) = \langle y, y_0 \rangle \quad (y_0 \in H_2), \quad (4.43)$$

e dal Teorema 3.27 sappiamo anche che x_0 e y_0 sono determinati univocamente, rispettivamente, da f e g . Ciò definisce gli operatori

$$A_1 : H'_1 \rightarrow H_1 \quad \text{con} \quad A_1 f = x_0, \quad (4.44)$$

$$A_2 : H'_2 \rightarrow H_2 \quad \text{con} \quad A_2 g = y_0. \quad (4.45)$$

Dal teorema 3.27 vediamo che A_1 e A_2 sono biiettivi e isometrici perché $\|A_1 f\| = \|x_0\| = \|f\|$ e analogamente per A_2 . Inoltre gli operatori sono coniugati lineari (cf. Sez. 3.2). Infatti se scriviamo $f_1(x) = \langle x, x_1 \rangle$ e $f_2(x) = \langle x, x_2 \rangle$ abbiamo per tutti gli x e scalari α, β

$$\begin{aligned} (\alpha f_1 + \beta f_2)(x) &= \alpha f_1(x) + \beta f_2(x) \\ &= \alpha \langle x, x_1 \rangle + \beta \langle x, x_2 \rangle \\ &= \langle x, \overline{\alpha} x_1 + \overline{\beta} x_2 \rangle. \end{aligned}$$

Per definizione di A_1 ciò mostra che A_1 è coniugato lineare

$$A_1(\alpha f_1 + \beta f_2) = \overline{\alpha} A_1 f_1 + \overline{\beta} A_1 f_2.$$

Per A_2 la dimostrazione è simile.

La composizione dà l'operatore

$$T^* = A_1 T' A_2^{-1} : H_2 \rightarrow H_1 \quad \text{definito da} \quad T^* y_0 = x_0. \quad (4.46)$$

T^* è lineare perché coinvolge *due* applicazioni lineari coniugate in aggiunta all'operatore lineare T' . Proviamo che T^* è proprio l'operatore aggiunto di Hilbert di T . Ciò è semplice perché dalle (4.40)-(4.43) abbiamo

$$\langle Tx, y_0 \rangle = g(Tx) = f(x) = \langle x, x_0 \rangle$$

e dalle (4.44), (4.45)

$$\langle x, T^* y_0 \rangle = \langle x, A_1 T' A_2^{-1} y_0 \rangle = \langle x, A_1 T' g \rangle = \langle x, A_1 f \rangle = \langle x, x_0 \rangle.$$

Quindi si ottiene la (3.58) della Sez. 3.11 a meno della notazione. Il nostro risultato è il seguente.

La formula (4.46) rappresenta l'operatore aggiunto di Hilbert T^ di un operatore lineare T su uno spazio di Hilbert in termini dell'operatore duale T' di T .*

Si noti inoltre che $\|T^*\| = \|T\|$ (Teorema 3.32) segue ora immediatamente dalla (4.29) e dall'isometria di A_1 e A_2 .

Per completare questa discussione dovremmo anche elencare alcune delle principali differenze fra l'operatore duale T' di $T : X \rightarrow Y$ e l'operatore aggiunto di Hilbert $T^* : H_1 \rightarrow H_2$, dove X e Y sono spazi normati e H_1 e H_2 sono spazi di Hilbert.

T' è definito sul duale dello spazio che contiene l'immagine di T , mentre T^* è definito direttamente sullo spazio che contiene l'immagine di T . Questa proprietà di T^* ci permette di definire importanti classi di operatori mediante l'uso dei loro operatori aggiunti di Hilbert (cf. 3.35).

Per T' abbiamo per la (4.36)

$$(\alpha T)' = \alpha T'$$

ma per T^* abbiamo per la 3.34

$$(\alpha T)^* = \overline{\alpha} T^*.$$

Nel caso finito dimensionale T' è rappresentato dalla trasposta della matrice che rappresenta T , mentre T^* è rappresentato dalla complesso coniugato trasposta di questa matrice.

4.7 Spazi Riflessivi

La riflessività *algebrica* di spazi *vettoriali* è stata discussa nella Sez. 2.9. La *riflessività* di spazi *normati* è l'argomento della presente sezione.

Consideriamo uno spazio normato X , il suo spazio duale X' così come definito in 2.41 ed inoltre lo spazio duale $(X')'$ di X' . Questo spazio è indicato con X'' ed è chiamato lo **spazio biduale** di X .

Definiamo un funzionale g_x su X' scegliendo un $x \in X$ fisso e ponendo

$$g_x(f) = f(x) \quad (f \in X' \text{ variabile}). \quad (4.47)$$

Questa definizione sembra la stessa di quella data per la riflessività algebrica, ma si noti che ora f è limitato. E anche g_x risulta essere limitato, poiché abbiamo il seguente lemma basilare.

4.15 Lemma (Norma di g_x)

Per ogni x fisso in uno spazio normato X il funzionale g_x definito dalla (4.47) è un funzionale lineare limitato su X' , così che $g_x \in X''$ ed ha norma

$$\|g_x\| = \|x\|. \quad (4.48)$$

■

Dimostrazione. La linearità di g_x è nota dalla Sez. 2.9 e la (4.48) segue dalla (4.47) e dal Corollario 4.12

$$\|g_x\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|g_x(f)|}{\|f\|} = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} = \|x\|. \quad (4.49)$$

■

Ad ogni $x \in X$ corrisponde un unico funzionale lineare limitato $g_x \in X''$ dato dalla (4.47). Ciò definisce un'applicazione

$$\begin{aligned} C : X &\rightarrow X'' \\ x &\mapsto g_x. \end{aligned} \quad (4.50)$$

C è chiamato l'**applicazione canonica** di X in X'' . Mostriamo che C è lineare ed iniettiva e conserva la norma. Ciò può essere espresso in termini di un isomorfismo di spazi normati come definito nella Sez. 2.11.

4.16 Lemma (Applicazione Canonica)

L'applicazione canonica C data dalla (4.50) è un isomorfismo dello spazio normato X sullo spazio normato $\mathcal{R}(C)$, l'immagine di C . ■

Dimostrazione. La linearità di C può essere verificata come nella Sez.2.9 perché

$$g_{\alpha x + \beta y}(f) = f(\alpha x + \beta y) = \alpha f(y) + \beta f(y) = \alpha g_x(f) + \beta g_y(f).$$

In particolare $g_x - g_y = g_{x-y}$. Quindi per la (4.48) otteniamo

$$\|g_x - g_y\| = \|g_{x-y}\| = \|x - y\|.$$

Ciò mostra che C è isometrico; esso conserva la norma. L'isometria implica l'iniettività. Possiamo anche vederlo direttamente dalla nostra formula. Infatti se $x \neq y$ allora $g_x \neq g_y$ per l'assioma (N2) nella Sez. 2.2. Quindi C è biiettiva se considerata come un'applicazione sulla sua immagine. ■

X è detto **immersibile (embeddable)** in uno spazio normato Z se X è isomorfo ad un sottospazio di Z . Ciò è simile alla Sez. 2.9, ma si noti che qui abbiamo a che fare con un isomorfismo di spazi normati, cioè con un isomorfismo di spazi vettoriali che conserva la norma (cf. Sez. 2.11). Il Lemma 4.16 mostra che X è immersibile in X'' e C è chiamato *l'immersione canonica (embedding canonico)* di X in X'' .

In generale C non sarà surgettivo, così che l'immagine $\mathcal{R}(C)$ sarà un sottospazio *proprio* di X'' . Il caso surgettivo quando $\mathcal{R}(C)$ è tutto X'' è sufficientemente importante da dargli un nome.

4.17 DEFINIZIONE (RIFLESSIVITÀ)

Uno spazio normato X è detto *riflessivo* se

$$\mathcal{R}(C) = X''$$

dove $C : X \rightarrow X''$ è l'applicazione canonica data dalla (4.50) e dalla (4.47). ■

Se X è riflessivo è isomorfo (quindi isometrico) a X'' per il Lemma 4.16. È interessante il fatto che l'inverso in generale non valga.

Inoltre la completezza non implica la riflessività ma viceversa abbiamo il seguente teorema.

4.18 Teorema (Completezza)

Se uno spazio normato X è riflessivo allora è completo (quindi uno spazio di Banach). ■

Dimostrazione. Poiché X'' è lo spazio duale di X' esso è completo per il Teorema 2.42. La riflessività di X significa che $\mathcal{R}(C) = X''$. La completezza di X ora segue da quella di X'' per il Lemma 4.16. ■

Ogni spazio normato X finito dimensionale è riflessivo. Infatti se $\dim X < \infty$ allora ogni funzionale lineare su X è limitato (cf. 2.30), così che $X' = X^*$ e la riflessività algebrica di X implica allora il seguente teorema.

4.19 Teorema (Dimensioni Finite)

Ogni spazio normato finito dimensionale è riflessivo. ■

Anche gli spazi di Hilbert risultano essere sempre riflessivi.

4.20 Teorema (Spazio di Hilbert)

Ogni spazio di Hilbert è riflessivo. ■

Dimostrazione. Proviamo la surgettività dell'applicazione canonica $C : H \rightarrow H''$ mostrando che per ogni $g \in H''$ v'è un $x \in H$ tale che $g = Cx$. Come preparazione definiamo $A : H' \rightarrow H$ con $Af = z$ dove z è dato dalla rappresentazione di Riesz $f(x) = \langle x, z \rangle$ in 3.27. Dal Teorema 3.27 sappiamo che A è biiettivo ed isometrico. A è coniugato lineare come si vede ricordando

che il prodotto scalare è coniugato lineare a destra. Ora H' è completo per il Teorema 2.42 e diviene uno spazio di Hilbert quando si introduce il prodotto scalare definito da

$$\langle f_1, f_2 \rangle = \langle Af_2, Af_1 \rangle.$$

Si noti il diverso ordine di f_1, f_2 nei due membri. (IP1) sino a (IP4) nella Sez. 3.2 sono facilmente verificati. In particolare (IP2) segue dalla linearità coniugata di A

$$\langle \alpha f_1, f_2 \rangle = \langle Af_2, A(\alpha f_1) \rangle = \langle Af_2, \bar{\alpha} Af_1 \rangle = \alpha \langle f_1, f_2 \rangle_1.$$

Sia $g \in H''$ arbitrario. Sia la sua rappresentazione di Riesz

$$g(f) = \langle f, f_0 \rangle = \langle Af_0, Af \rangle.$$

Ricordiamo ora che $f(x) = \langle x, z \rangle$ dove $z = Af$. Scrivendo $Af_0 = x$ abbiamo così

$$\langle Af_0, Af \rangle = \langle x, z \rangle = f(x).$$

Confrontando $g(f) = f(x)$, ossia $g = Cx$ per la definizione di C . Poiché $g \in H''$ era arbitrario C è surgettivo e quindi H è riflessivo. ■

A volte separabilità e non separabilità possono giocare un ruolo nella prova che certi spazi non sono riflessivi. Per questo studio è utile il seguente Lemma.

4.21 Lemma (Esistenza di un Funzionale)

Sia Y un sottospazio proprio chiuso di uno spazio normato X . Sia $x_0 \in X - Y$ arbitrario e

$$\delta = \inf_{y \in Y} \|y - x_0\| \quad (4.51)$$

la distanza da x_0 a Y . Allora esiste un $\tilde{f} \in X'$ tale che

$$\|\tilde{f}\| = 1, \quad \tilde{f}(y) = 0 \text{ per tutti gli } y \in Y, \quad \tilde{f}(x_0) = \delta. \quad (4.52)$$

■

Dimostrazione. Per dimostrare il Lemma si considera il sottospazio $Z \subset X$ generato da Y e x_0 , si definisce su Z un funzionale lineare limitato f con

$$f(z) = f(y + \alpha x_0) = \alpha \delta \quad y \in Y, \quad (4.53)$$

si mostra che f soddisfa alla (4.52) ed infine si estende f a X utilizzando il Teorema 4.10. I dettagli sono i seguenti.

Ogni $z \in Z = \text{span}(Y \cup \{x_0\})$ ha un'unica rappresentazione

$$z = y + \alpha x_0 \quad y \in Y.$$

Ciò è usato nella (4.53). La linearità di f si vede facilmente. Inoltre poiché Y è chiuso $\delta > 0$ così che $f \neq 0$. Ora $\alpha = 0$ dà $f(y) = 0$ per tutti gli $y \in Y$. Per $\alpha = 1$ ed $y = 0$ abbiamo $f(x_0) = \delta$.

Mostriamo che f è limitato. $\alpha = 0$ dà $f(z) = 0$. Sia $\alpha \neq 0$. Usando la (4.51) e notando che $-(1/\alpha)y \in Y$ otteniamo

$$\begin{aligned} |f(z)| &= |\alpha| \delta = |\alpha| \inf_{\tilde{y} \in Y} \|\tilde{y} - x_0\| \\ &\leq |\alpha| \left\| -\frac{1}{\alpha} y - x_0 \right\| \\ &= \|y + \alpha x_0\|, \end{aligned}$$

cioè $|f(z)| \leq \|z\|$. Quindi f è limitato e $\|f\| \leq 1$.

Mostriamo che $\|f\| \geq 1$. Per definizione di estremo inferiore Y contiene una successione (y_n) tale che $\|y_n - x_0\| \rightarrow \delta$. Sia $z_n = y_n - x_0$. Allora abbiamo $f(z_n) = -\delta$ per la (4.53) con $\alpha = -1$. Inoltre

$$\|f\| = \sup_{\substack{z \in Z \\ z \neq 0}} \frac{|f(z)|}{\|z\|} \geq \frac{|f(z_n)|}{\|z_n\|} = \frac{\delta}{\|z_n\|} \rightarrow \frac{\delta}{\delta} = 1$$

per $n \rightarrow \infty$. Quindi $\|f\| \geq 1$, così che $\|f\| = 1$. Per il teorema di Hahn–Banach 4.10 per gli spazi normati possiamo estendere f a X senza aumentarne la norma. ■

Usando questo lemma possiamo ottenere il teorema desiderato.

4.22 Teorema (Separabilità)

Se lo spazio duale X' di uno spazio normato X è separabile allora X stesso è separabile. ■

Dimostrazione. Assumiamo che X' sia separabile. Allora anche la sfera unitaria $U' = \{f \mid \|f\| = 1\} \subset X'$ contiene un sottoinsieme denso numerabile. Sia esso (f_n) . Poiché $f_n \in U'$ abbiamo

$$\|f_n\| = \sup_{\|x\|=1} |f_n(x)| = 1.$$

Per definizione di estremo superiore possiamo trovare dei punti $x_n \in X$ di norma 1 tali che

$$|f_n(x_n)| \geq \frac{1}{2}.$$

Sia Y la chiusura di $\text{span}(x_n)$. Allora Y è separabile perché Y ha un sottoinsieme denso numerabile, cioè l'insieme di tutte le combinazioni lineari dei x_n con coefficienti la cui parte reale ed immaginaria sono razionali.

Mostriamo che $Y = X$. Supponiamo $Y \neq X$. Allora poiché Y è chiuso per il Lemma 4.21 esiste un $\tilde{f} \in X'$ con $\|\tilde{f}\| = 1$ ed $\tilde{f}(y) = 0$ per tutti gli $y \in Y$. Poiché $x_n \in Y$ abbiamo $\tilde{f}(x_n) = 0$ e per tutti gli n

$$\begin{aligned} \frac{1}{2} &\leq |f_n(x_n)| = |f_n(x_n) - \tilde{f}(x_n)| \\ &= |(f_n - \tilde{f})(x_n)| \\ &\leq \|f_n - \tilde{f}\| \|x_n\| \end{aligned}$$

dove $\|x_n\| = 1$. Quindi $\|f_n - \tilde{f}\| \geq \frac{1}{2}$, ma ciò contraddice l'ipotesi che (f_n) sia denso in U' perché $\tilde{f}$ è esso stesso in U' , essendo $\|\tilde{f}\| = 1$. ■

Da questo teorema segue immediatamente la seguente proposizione.

Se X è uno spazio normato separabile e riflessivo allora anche il suo duale X' è separabile.

Infatti se uno spazio normato X è riflessivo, X'' è isomorfo a X per il Teorema 4.16, così che la separabilità di X implica la separabilità di X'' e quindi per il Teorema 4.22 anche la separabilità di X' .

4.8 Teorema della Categoria e della Uniforme Limitatezza

Il teorema di uniforme limitatezza (o *principio di uniforme limitatezza*) è spesso considerato come una delle pietre d'angolo dell'analisi funzionale negli spazi normati, le altre essendo

il teorema di Hahn–Banach, il teorema dell'applicazione aperta ed il teorema del grafico chiuso. Al contrario del teorema di Hahn–Banach gli altri tre di questi quattro teoremi richiedono la completezza. Infatti essi caratterizzano alcune delle più importanti proprietà degli spazi di Banach che in generale gli spazi normati possono non avere.

È molto interessante notare che otterremo tutti e tre i teoremi da una sorgente comune. Più precisamente proveremo il cosiddetto teorema della categoria di Baire e deriveremo da esso sia il teorema di uniforme limitatezza (in questa sezione) che il teorema della applicazione aperta (nella Sez. 4.11). Quest'ultimo fornirà allora facilmente il teorema del grafico chiuso (nella Sez. 4.12).

4.23 Teorema (Categoria di Baire. Spazi Metrici Completi)

Se uno spazio metrico $X \neq \emptyset$ è completo e

$$X = \bigcup_{k=1}^{\infty} A_k \quad \text{con tutti gli } A_k \text{ chiusi,} \quad (4.54)$$

allora almeno un A_k contiene un sottoinsieme aperto non vuoto. ■

Dimostrazione. Supponiamo per assurdo che lo spazio metrico completo $X \neq \emptyset$ sia tale che

$$X = \bigcup_{k=1}^{\infty} M_k \quad (4.55)$$

con ciascun M_k chiuso non contenente un sottoinsieme aperto non vuoto. Costruiremo una successione di Cauchy (p_k) il cui limite p (che esiste per la completezza) non è in alcun M_k , contraddicendo perciò la rappresentazione (4.55).

Per ipotesi M_1 non contiene un insieme aperto non vuoto. Ma X sì (ad esempio X stesso). Questo implica che $M_1^C \neq \emptyset$. Quindi il complemento $M_1^C = X - M_1$ di M_1 è non vuoto ed aperto. Possiamo così scegliere un punto p_1 in M_1^C ed una palla chiusa di cui è centro tale che

$$B_1 = B(p_1; r_1) \subset M_1^C \quad r_1 \leq \frac{1}{2}.$$

Per ipotesi anche M_2 non contiene un insieme aperto non vuoto. Quindi non contiene la palla aperta $\overset{\circ}{B}_1$. Ciò implica che $M_2^C \cap \overset{\circ}{B}_1$ è non vuoto ed aperto, così che possiamo scegliere una palla chiusa B_2 in questo insieme tale che

$$B_2 = B(p_2; r_2) \subset M_2^C \cap \overset{\circ}{B}_1 \quad r_2 \leq \frac{1}{2^2}.$$

Per induzione otteniamo così una successione di palle chiuse

$$B_k = B(p_k; r_k) \quad r_k \leq \frac{1}{2^k}$$

tali che $B_k \cap M_k = \emptyset$ e

$$B_{k+1} \subset B_k \quad k = 1, 2, \dots$$

Poiché $r_k \leq 2^{-k}$ la successione (p_k) dei centri è di Cauchy e converge, ossia $p_k \rightarrow p \in X$ perché X è completo per ipotesi. Inoltre per ogni m e $n > m$ abbiamo $B_n(p_n; r_n) \subset B_m(p_m; r_m)$ così che

$$\begin{aligned} d(p_m, p) &\leq d(p_m, p_n) + d(p_n, p) \\ &\leq r_m + d(p_n, p) \end{aligned}$$

e per $n \rightarrow \infty$

$$d(p_m, p) \leq r_m.$$

Ossia $p \in B_m$ per ogni m . Poiché $B_m \subset M_m^C$, vediamo ora che $p \notin M_m$ per ogni m , così che $p \notin X$. Ciò contraddice $p \in X$. Il teorema di Baire è dimostrato. ■

Notiamo che l'inverso del teorema di Baire in generale non è vero.

Dal teorema di Baire otterremo ora facilmente il teorema della uniforme limitatezza. Questo teorema stabilisce che se X è uno spazio di Banach ed una successione di operatori $T_n \in B(X, Y)$ è limitata in ogni punto $x \in X$, allora la successione è uniformemente limitata. In altre parole la limitatezza puntiforme implica la limitatezza in un senso più forte, ossia la uniforme limitatezza. (Il numero reale c_x nella (4.56), qui di seguito, varia in generale con x , un fatto che indichiamo coll'indice x ; il punto essenziale è che c_x non dipende da n .)

4.24 Teorema (Uniforme Limitatezza)

Sia (T_n) una successione di operatori lineari limitati $T_n : X \rightarrow Y$ da uno spazio di Banach X in uno spazio normato Y tale che $(\|T_n x\|)$ sia limitata per ogni $x \in X$, cioè

$$\|T_n x\| \leq c_x \quad n = 1, 2, \dots \quad (4.56)$$

dove c_x è un numero reale. Allora la successione di norme $\|T_n\|$ è limitata, ossia v'è un c tale che

$$\|T_n\| \leq c \quad n = 1, 2, \dots \quad (4.57)$$

■

Dimostrazione. Per ogni $k \in \mathbb{N}$ sia $A_k \subset X$ l'insieme di tutti gli x tali che

$$\|T_n x\| \leq k \quad \text{per tutti gli } n.$$

A_k è chiuso. Infatti per ogni $x \in \overline{A_k}$ v'è una successione (x_j) in A_k che converge a x . Questo significa che per ogni n fisso abbiamo $\|T_n x_j\| \leq k$ ed otteniamo $\|T_n x\| \leq k$ perché T_n è continuo e così la norma in Y (cf. Sez. 2.2). Quindi $x \in A_k$ e A_k è chiuso.

Per la (4.56) ciascun $x \in X$ appartiene a qualche A_k . Quindi

$$X = \bigcup_{k=1}^{\infty} A_k.$$

Poiché X è completo il teorema di Baire implica che qualche A_k contiene una palla aperta, sia

$$B_0 = B(x_0; r) \subset A_{k_0}. \quad (4.58)$$

Sia $x \in X$ arbitrario e non nullo. Poniamo

$$z = x_0 + \gamma x \quad \gamma = \frac{r}{2\|x\|}. \quad (4.59)$$

Allora $\|z - x_0\| < r$, così che $z \in B_0$. Per la (4.58) e dalla definizione di A_{k_0} abbiamo quindi che $\|T_n z\| \leq k_0$ per tutti gli n . Inoltre $\|T_n x_0\| \leq k_0$ perché $x_0 \in B_0$. Dalla (4.59) otteniamo

$$x = \frac{1}{\gamma}(z - x_0).$$

Ciò fornisce per tutti gli n

$$\|T_n x\| = \frac{1}{\gamma} \|T_n(z - x_0)\| \leq \frac{1}{\gamma} (\|T_n z\| + \|T_n x_0\|) \leq \frac{4}{r} \|x\| k_0.$$

Quindi per tutti gli n

$$\|T_n\| = \sup_{\|x\|=1} \|T_n x\| \leq \frac{4}{r} k_0,$$

che è della forma (4.57) con $c = 4k_0/r$. ■

4.9 Convergenza Forte e Debole

Sappiamo che in analisi si definiscono diversi tipi di convergenza (ordinaria, condizionale, assoluta ed uniforme). Ciò fornisce maggiore flessibilità nella teoria e nelle applicazioni delle successioni e delle serie. Nell'analisi funzionale la situazione è simile e vi sono una ancora maggiore varietà di possibilità che risultano di interesse pratico. Nella presente sezione ci occupiamo principalmente della “convergenza debole”. Questo è un concetto basilare. Lo presentiamo ora perché la teoria della convergenza debole fa un uso essenziale del teorema di uniforme limitatezza discusso nella sezione precedente. Infatti è questa una delle maggiori applicazioni di questo teorema.

La convergenza di successioni di elementi in uno spazio normato è stata definita nella Sez. 2.3 e d'ora innanzi sarà chiamata *convergenza forte* per distinguerla dalla “convergenza debole”.

Definiamo quindi dapprima la convergenza forte.

4.25 DEFINIZIONE (CONVERGENZA FORTE)

Una successione (x_n) in uno spazio normato X è detta *convergente fortemente* (o *convergente in norma*) se v'è un $x \in X$ tale che

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0.$$

Si scrive

$$\lim_{n \rightarrow \infty} x_n = x$$

o semplicemente

$$x_n \rightarrow x.$$

x è chiamato il *limite forte* di (x_n) e diciamo che (x_n) *converge fortemente a x* . ■

La convergenza debole è definita in termini di funzionali lineari limitati su X come segue.

4.26 DEFINIZIONE (CONVERGENZA DEBOLE)

Una successione (x_n) in uno spazio normato X è detta *debolmente convergente* se v'è un x tale che per ogni $f \in X'$

$$\lim_{n \rightarrow \infty} f(x_n) = f(x).$$

Si scrive

$$x_n \xrightarrow{w} x$$

o $x_n \rightharpoonup x$. L'elemento x è chiamato il *limite debole* di (x_n) e diciamo che (x_n) *converge debolmente a x* . ■

Si noti che la convergenza debole significa la convergenza per la successione di numeri $a_n = f(x_n)$ per ogni $f \in X'$.

Il concetto illustra un principio basilare dell'analisi funzionale e precisamente il fatto che la studio degli spazi è spesso collegato a quello dei loro spazi duali.

Per applicare la convergenza debole abbiamo bisogno di conoscere alcune proprietà basilari, che enunciamo nel prossimo lemma. Il lettore noterà che nella prova utilizziamo il teorema di Hahn–Banach (sia 4.12 che 4.15) ed il teorema di uniforme limitatezza. Ciò mostra l'importanza di questi teoremi in connessione con la convergenza debole.

4.27 Lemma (Convergenza Debole)

Sia (x_n) una successione convergente debolmente in uno spazio normato X , ossia $x_n \xrightarrow{w} x$. Allora

- (a) *Il limite debole x di (x_n) è unico.*
- (b) *Ogni sottosuccessione di (x_n) converge debolmente a x .*
- (c) *La successione $(\|x_n\|)$ è limitata.* ■

Dimostrazione. (a) Supponiamo che $x_n \xrightarrow{w} x$ e che del pari $x_n \xrightarrow{w} y$. Allora $f(x_n) \rightarrow f(x)$ e del pari $f(x_n) \rightarrow f(y)$. Poiché $(f(x_n))$ è una successione di numeri il suo limite è unico. Quindi $f(x) = f(y)$, cioè per ogni $f \in X'$ abbiamo

$$f(x) - f(y) = f(x - y) = 0.$$

Ciò implica $x - y = 0$ per il Corollario 4.12 e mostra che il limite debole è unico.

(b) segue dal fatto che $(f(x_n))$ è una successione convergente di numeri, così che ogni sottosuccessione di $(f(x_n))$ converge ed ha il medesimo limite.

(c) Poiché $(f(x_n))$ è una successione convergente di numeri essa è limitata, cioè $|f(x_n)| \leq c_f$ per tutti gli n , dove c_f è una costante che dipende da f ma non da n . Usando l'applicazione canonica $C : X \rightarrow X''$ (Sez. 4.7) possiamo definire $g_n \in X''$ con

$$g_n(f) = f(x_n) \quad f \in X'.$$

(Scriviamo g_n invece di g_{x_n} per evitare indici di indici.) Allora per tutti gli n

$$|g_n(f)| = |f(x_n)| \leq c_f,$$

cioè la successione $(|g_n(f)|)$ è limitata per ogni $f \in X'$. Poiché X' è completo per il 2.42 è applicabile il teorema della uniforme limitatezza 4.24 e ciò implica che $(\|g_n\|)$ è limitata. Ora $\|g_n\| = \|x_n\|$ per il 4.15 e quindi (c) è provato. ■

Il lettore può forse meravigliarsi del fatto che la convergenza debole non giochi un ruolo in analisi. La semplice ragione è che negli spazi normati finito dimensionali la distinzione fra convergenza debole e forte scompare completamente. Proviamo questo fatto e giustifichiamo anche i termini “forte” e “debole”.

4.28 Teorema (Convergenza Forte e Debole)

Sia (x_n) una successione in uno spazio normato X . Ne segue

- (a) *La convergenza forte implica la convergenza debole col medesimo limite.*
- (b) *L'inverso di (a) non è in generale vero.*
- (c) *Se $\dim X < \infty$ allora la convergenza debole implica la convergenza forte.* ■

Dimostrazione. (a) Per definizione $x_n \rightarrow x$ significa $\|x_n - x\| \rightarrow 0$ ed implica che per ogni $f \in X'$

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \|x_n - x\| \rightarrow 0.$$

Ciò mostra che $x_n \xrightarrow{w} x$.

(b) può essere visto da una successione ortonormale (e_n) in uno spazio di Hilbert H . Infatti ogni $f \in H'$ ha una rappresentazione di Riesz $f(x) = \langle x, z \rangle$. Quindi $f(e_n) = \langle e_n, z \rangle$. Ora la disuguaglianza di Bessel è (cf. 3.19)

$$\sum_{n=1}^{\infty} |\langle e_n, z \rangle|^2 \leq \|z\|^2.$$

Quindi la serie a sinistra converge, così che i suoi termini devono tendere a zero per $n \rightarrow \infty$. Ciò implica

$$f(e_n) = \langle e_n, z \rangle \rightarrow 0.$$

Poiché $f \in H'$ era arbitrario vediamo che $e_n \xrightarrow{w} 0$. Tuttavia (e_n) non converge fortemente perché

$$\|e_m - e_n\|^2 = \langle e_m - e_n, e_m - e_n \rangle = 2 \quad (m \neq n).$$

(c) Supponiamo che $x_n \xrightarrow{w} x$ e che $\dim X = k$. Sia $\{e_1, \dots, e_k\}$ una base per X e sia

$$x_n = \alpha_1^{(n)} e_1 + \dots + \alpha_k^{(n)} e_k$$

e

$$x = \alpha_1 e_1 + \dots + \alpha_k e_k.$$

Per ipotesi $f(x_n) \rightarrow f(x)$ per ogni $f \in X'$. Prendiamo in particolare $f_1, \dots, f_k$ definiti da

$$f_j(e_j) = 1, \quad f_j(e_m) = 0 \quad (m \neq j).$$

(Ricordiamo che questa è la base duale di $\{e_1, \dots, e_k\}$). Allora

$$f_j(x_n) = \alpha_j^{(n)}, \quad f_j(x) = \alpha_j.$$

Quindi $f_j(x_n) \rightarrow f_j(x)$ implica $\alpha_j^{(n)} \rightarrow \alpha_j$. Da ciò otteniamo facilmente che

$$\begin{aligned} \|x_n - x\| &= \left\| \sum_{j=1}^k (\alpha_j^{(n)} - \alpha_j) e_j \right\| \\ &\leq \sum_{j=1}^k |\alpha_j^{(n)} - \alpha_j| \|e_j\| \rightarrow 0 \end{aligned}$$

per $n \rightarrow \infty$. Ciò mostra che (x_n) converge fortemente a x . ■

È interessante notare che esistono anche spazi infinito dimensionali tali che i concetti di convergenza forte e debole sono equivalenti. Un esempio è l^1 .

4.29 Teorema (Convergenza Debole)

Se in uno spazio normato

- (a) *la successione $(\|x_n\|)$ è limitata,*
 (b) *per ogni elemento f di un sottoinsieme $M \subset X'$ per cui è $\overline{\text{span } M} = X'$ si ha $f(x_n) \rightarrow f(x)$*

allora $x_n \xrightarrow{w} x$. ■

Dimostrazione. Consideriamo un qualunque $f \in X'$ e mostriamo che $f(x_n) \rightarrow f(x)$, che significa la convergenza debole per definizione.

Per (a) abbiamo $\|x_n\| \leq c$ per tutti gli n e $\|x\| \leq c$ per un c sufficientemente largo. Poiché per l'ipotesi (b) $\overline{\text{span } M} = X'$, per ogni $f \in X'$ v'è una successione (f_j) in $\text{span } M$ tale che $f_j \rightarrow f$. Quindi per ogni dato $\varepsilon > 0$ possiamo trovare un j tale che

$$\|f_j - f\| < \frac{\varepsilon}{3c}.$$

Inoltre poiché $f_j \in \text{span } M$ per l'ipotesi (b) v'è un N tale che per tutti gli $n > N$

$$|f_j(x_n) - f_j(x)| < \frac{\varepsilon}{3}.$$

Usando queste due disuguaglianze e applicando il teorema della disuguaglianza triangolare otteniamo per $n > N$

$$\begin{aligned} |f(x_n) - f(x)| &\leq |f(x_n) - f_j(x_n)| + |f_j(x_n) - f_j(x)| + |f_j(x) - f(x)| \\ &< \|f - f_j\| \|x_n\| + \frac{\varepsilon}{3} + \|f_j - f\| \|x\| \\ &< \frac{\varepsilon}{3c} c + \frac{\varepsilon}{3} + \frac{\varepsilon}{3c} c = \varepsilon. \end{aligned}$$

Poiché $f \in X'$ era arbitrario ciò mostra che la successione (x_n) converge debolmente a x . ■

In conclusione consideriamo la convergenza debole in due spazi particolarmente importanti.

4.9.1 Esempi**4.30 Teorema (Spazio di Hilbert)**

In uno spazio di Hilbert $x_n \xrightarrow{w} x$ se e solo se $\langle x_n, z \rangle \rightarrow \langle x, z \rangle$ per tutti gli z nello spazio. ■

Dimostrazione. È evidente per il 3.27. ■

4.31 Teorema (Spazio l^p .)

Nello spazio l^p , dove $1 < p < +\infty$, abbiamo $x_n \xrightarrow{w} x$ se e solo se

- (a) *La successione $(\|x_n\|)$ è limitata.*
 (b) *Per ogni j fisso abbiamo $\xi_j^{(n)} \rightarrow \xi_j$ per $n \rightarrow \infty$, dove $x_n = (\xi_j^{(n)})$ e $x = (\xi_j)$.* ■

Dimostrazione. Lo spazio duale di l^p è l^q , dove $1/p + 1/q = 1$. Una base di Schauder di l^q è (e_n) dove $e_n = (\delta_{nj})$ ha 1 al posto n -mo e zero altrove. $\text{span}(e_n)$ è denso in l^q , cioè è una base, così che la conclusione segue dal teorema 4.29. ■

4.10 Convergenza di Successioni di Operatori e di Funzionali

Per le successioni di operatori $T_n \in B(X, Y)$ tre tipi di convergenza risultano essere di rilievo sia teorico che pratico. Essi sono

1. Convergenza in norma su $B(X, Y)$,
2. Convergenza forte di $(T_n x)$ in Y ,
3. Convergenza debole di $(T_n x)$ in Y .

Le definizioni e la terminologia sono le seguenti.

4.32 DEFINIZIONE (CONVERGENZA DI SUCCESSIONI DI OPERATORI)

Siano X e Y spazi normati. Una successione di operatori (T_n) di operatori $T_n \in B(X, Y)$ è detta

1. **convergente uniformemente** se (T_n) converge in norma su $B(X, Y)$,
2. **convergente fortemente** se $(T_n x)$ converge fortemente in Y per ogni $x \in X$,
3. **convergente debolmente** se $(T_n x)$ converge debolmente in Y per ogni $x \in X$.

In formule questo significa che v'è un operatore T tale che, rispettivamente,

$$\|T_n - T\| \rightarrow 0 \quad (4.60)$$

$$\|T_n x - T x\| \rightarrow 0 \quad \text{per tutti gli } x \in X \quad (4.61)$$

$$|f(T_n x) - f(T x)| \rightarrow 0 \quad \text{per tutti gli } x \in X \text{ e per tutti gli } f \in Y'. \quad (4.62)$$

T è chiamato, rispettivamente, l'*operatore limite uniforme, forte e debole*. ■

Abbiamo rilevato nella sezione precedente che anche in analisi, in una situazione molto più semplice, l'uso di diversi concetti di convergenza dà una maggiore flessibilità. Ciò nonostante il lettore può essere ancora sconcertato dai molti concetti di convergenza che abbiamo appena introdotto. Si può chiedere se è proprio necessario avere tre concetti di convergenza per le successioni di operatori. La risposta è che molti degli operatori che appaiono nei problemi pratici sono dati come una sorta di limite di operatori più semplici. È importante sapere che cosa si intende per "una sorta di" e di sapere quali proprietà dell'operatore limite sono implicate dalle proprietà della successione. Inoltre all'inizio di una ricerca non sempre uno sa in che senso il limite esiste; quindi è utile avere una molteplicità di possibilità. A volte in un problema specifico si è dapprima capaci di stabilire la convergenza in un senso molto "dolce", così che uno ha almeno qualcosa da cui partire, e solo successivamente è in grado di sviluppare degli strumenti per provare la convergenza in un senso più forte, ciò che garantisce "migliori" proprietà dell'operatore limite. Questa è una situazione tipica nelle equazioni alle derivate parziali.

Non è difficile mostrare che

$$(4.60) \implies (4.61) \implies (4.62)$$

(il limite essendo lo stesso), ma l'inverso in generale non è vero, come si può vedere da alcuni esempi.

I **funzionali lineari** sono operatori lineari (con immagine nei campi scalari $\mathbb{R}$ o $(\mathbb{C})$), così che (4.60), (4.61) e (4.62) si applicano immediatamente. Tuttavia (4.61) e (4.62) ora

diventano equivalenti per la seguente ragione. Abbiamo $T_n x \in Y$, ma ora abbiamo $f_n(x) \in \mathbb{R}$ (o $\mathbb{C}$). Quindi la convergenza in (4.61) e (4.62) ora ha luogo in uno spazio finito dimensionale (unidimensionale) $\mathbb{R}$ (o $\mathbb{C}$) e l'equivalenza di (4.61) e (4.62) segue dal Teorema 4.28(c). I due concetti rimanenti sono chiamati *convergenza forte* e *debole** (si legga “convergenza debole star”).

4.33 DEFINIZIONE (CONVERGENZA FORTE E DEBOLE* DI UNA SUCCESSIONE DI FUNZIONALI)
Sia (f_n) una successione di funzionali lineari limitati su uno spazio normato X . Allora

(a) La *convergenza forte* di (f_n) significa che v'è un $f \in X'$ tale che $\|f_n - f\| \rightarrow 0$. Si scrive

$$f_n \rightarrow f.$$

(b) La *convergenza debole** di (f_n) significa che v'è un $f \in X'$ tale che $f_n(x) - f(x) \rightarrow 0$ per tutti gli $x \in X$. Si scrive²

$$f_n \xrightarrow{w^*} f.$$

f in (a) e (b) è chiamato, rispettivamente, il *limite forte* e *debole** di (f_n) . ■

Ritornando agli operatori $T_n \in B(X, Y)$ ci chiediamo che cosa si possa dire sull'operatore limite $T : X \rightarrow Y$ in (4.60)-(4.62).

Se la convergenza è uniforme $T \in B(X, Y)$; altrimenti $\|T_n - T\|$ non avrebbe senso. Se la convergenza è forte o debole T è ancora lineare ma può non essere limitato se X non è completo.

Tuttavia se X è completo abbiamo il seguente lemma basilare.

4.34 Lemma (Convergenza Forte di Operatori)

Sia $T_n \in B(X, Y)$ dove X è uno spazio di Banach ed Y uno spazio normato. Se (T_n) converge fortemente con limite T allora $T \in B(X, Y)$. ■

Dimostrazione. La linearità di T segue facilmente dalla linearità di T_n . Poiché $T_n x \rightarrow Tx$ per ogni $x \in X$ la successione $(T_n x)$ è limitata per ogni x ; cf. 1.8. Poiché X è completo $(\|T_n\|)$ è limitato per il teorema della uniforme limitatezza, ossia $\|T_n\| \leq c$ per tutti gli n . Da ciò segue che $\|T_n x\| \leq \|T_n\| \|x\| \leq c \|x\|$. Ciò implica $\|Tx\| \leq c \|x\|$. ■

Un utile criterio per la convergenza forte di operatori è il seguente.

4.35 Teorema (Convergenza Forte di Operatori)

Sia data una successione (T_n) di operatori $T_n \in B(X, Y)$, dove X è normato ed Y è uno spazio di Banach. Allora se

(a) la successione $(\|T_n\|)$ è limitata e

(b) la successione $(T_n x)$ converge in Y per ogni x in un sottoinsieme M tale che $\text{span } M$ è denso in X ,

la successione (T_n) è fortemente convergente ■

²Questo concetto è alquanto più importante di quello di *convergenza debole* di (f_n) , che per il 4.26 significa $g(f_n) \rightarrow g(f)$ per tutti i $g \in X''$. La convergenza debole implica la convergenza debole* come si può vedere usando l'applicazione canonica definita nella Sez. 4.7.

Dimostrazione. Sia c tale che $\|T_n\| \leq c$ per tutti gli n . Consideriamo un qualunque $x \in X$ e mostriamo che $(T_n x)$ converge fortemente in Y . Sia $\varepsilon > 0$ dato. Poiché $\text{span } M$ è denso in X v'è un $y \in \text{span } M$ tale che

$$\|x - y\| < \frac{\varepsilon}{3c}.$$

Poiché $y \in \text{span } M$ la successione $(T_n y)$ è di Cauchy per (b). Quindi v'è un N tale che

$$\|T_n y - T_m y\| < \frac{\varepsilon}{3} \quad (m, n > N).$$

Usando queste due disuguaglianze ed applicando la disuguaglianza triangolare vediamo facilmente che $(T_n x)$ è di Cauchy in Y perché per $m, n > N$ otteniamo

$$\begin{aligned} \|T_n x - T_m x\| &\leq \|T_n x - T_n y\| + \|T_n y - T_m y\| + \|T_m y - T_m x\| \\ &< \|T_n\| \|x - y\| + \frac{\varepsilon}{3} + \|T_m\| \|x - y\| \\ &< c \frac{\varepsilon}{3c} + \frac{\varepsilon}{3} + c \frac{\varepsilon}{3c} = \varepsilon. \end{aligned}$$

Poiché Y è completo $(T_n x)$ converge in Y . Poiché $x \in X$ era arbitrario ciò prova la convergenza forte di (T_n) . ■

4.36 Corollario (Funzionali)

Sia data una successione (f_n) di funzionali lineari limitati su uno spazio normato X . Se

- (a) la successione $(\|f_n\|)$ è limitata e
- (b) la successione $(f_n(x))$ è di Cauchy per ogni x in un sottoinsieme M di X tale che $\overline{\text{span } M} = X$,

allora (f_n) è convergente debolmente*, il limite essendo un funzionale lineare limitato su X . ■

4.11 Teorema dell'Applicazione Aperta

Vi sono applicazioni tali che l'immagine di un qualunque insieme aperto è un insieme aperto. Ricordando la nostra discussione sull'importanza degli insiemi aperti nella Sez. 1.1 comprendiamo che le applicazioni aperte sono di interesse generale. Più specificamente il teorema dell'applicazione aperta stabilisce le condizioni per le quali un operatore lineare limitato è un'applicazione aperta. Come nel teorema di uniforme limitatezza abbiamo nuovamente bisogno della completezza ed il presente teorema mostra un'altra ragione per cui gli spazi di Banach sono più soddisfacenti degli spazi normati incompleti. Il teorema fornisce anche le condizioni per le quali l'inverso di un operatore lineare limitato è limitato. La dimostrazione del teorema dell'applicazione aperta si baserà sul teorema di Baire enunciato e spiegato nella Sez. 4.8.

Incominciamo coll'introdurre il concetto dell'applicazione aperta.

4.37 DEFINIZIONE (APPLICAZIONE APERTA)

Siano X e Y spazi metrici. Allora $T : \mathcal{D}(T) \rightarrow Y$ con dominio $\mathcal{D}(T) \subset X$ è chiamato un'applicazione aperta se per ogni insieme aperto in $\mathcal{D}(T)$ l'immagine è un insieme aperto in Y . ■

Si noti che se l'applicazione non è surgettiva bisogna fare attenzione a distinguere fra le asserzioni che l'applicazione è aperta come un'applicazione dal suo dominio

- (a) in Y ,
- (b) sulla sua immagine.

(b) è più debole di (a). Per esempio se $X \subset Y$ l'applicazione $x \mapsto x$ di X in Y è aperta se e solo se X è un sottoinsieme aperto di Y , mentre l'applicazione $x \mapsto x$ di X sulla sua immagine (che è X) è aperta in ogni caso.

Inoltre per evitare confusioni dovremmo ricordare che, per il Teorema 1.5, un'applicazione continua $T : X \rightarrow Y$ ha la proprietà caratteristica che per ogni insieme aperto in Y l'immagine inversa è un insieme aperto in X . Ciò *non* implica che T applica insiemi aperti in X su insiemi aperti in Y . Per esempio l'applicazione $\mathbb{R} \rightarrow \mathbb{R}$ data da $t \mapsto \sin t$ è continua ma applica $(0, 2\pi)$ su $[-1, 1]$.

Per dimostrare il teorema dell'applicazione aperta è necessario dapprima provare il seguente lemma.

4.38 Lemma (Palla Unitaria Aperta)

Un operatore lineare limitato T da uno spazio di Banach X su uno spazio di Banach Y , ossia tale che $T \in B(X, Y)$ e $Y = T(X)$, ha la proprietà che l'immagine $T(B_0)$ della palla unitaria aperta $B_0 = B(0; 1) \subset X$ contiene una palla aperta di centro $0 \in Y$. ■

Dimostrazione. Procedendo passo a passo proviamo

- (a) La chiusura dell'immagine della palla aperta $B_1 = B(0; 2^{-1})$, ossia $\overline{T(B_1)}$ contiene una palla aperta B^* non necessariamente centrata in $0 \in Y$.
- (b) $\overline{T(B_0)}$ contiene una palla aperta $B(0; \epsilon)$ di centro $0 \in Y$.
- (c) $T(B_0)$ contiene la palla aperta $B(0; \epsilon/2)$ di centro $0 \in Y$.

I dettagli sono i seguenti.

- (a) In connessione con i sottoinsiemi $A \subset X$ scriveremo αA (α uno scalare) e $A + w$ ($w \in X$) per indicare

$$\alpha A = \{x \in X \mid x = \alpha a, a \in A\} \quad (4.63)$$

$$A + w = \{x \in X \mid x = a + w, a \in A\} \quad (4.64)$$

ed analogamente per i sottoinsiemi di Y .

Consideriamo la palla aperta $B_1 = B(0; 1/2) \subset X$. Un qualunque $x \in X$ fisso è in kB_1 per k reale sufficientemente grande ($k > 2\|x\|$). Quindi

$$X = \bigcup_{k=1}^{\infty} kB_1.$$

Poiché T è lineare e surgettivo

$$Y = T(X) = T\left(\bigcup_{k=1}^{\infty} kB_1\right) = \bigcup_{k=1}^{\infty} kT(B_1) = \bigcup_{k=1}^{\infty} k\overline{T(B_1)}. \quad (4.65)$$

Si noti che prendendo la chiusura non aggiungiamo alcun altro punto all'unione perché l'unione era già l'intero spazio Y . Poiché Y è completo è non magro in se stesso per il teorema

4.23 di Baire. Quindi notando che (4.65) è simile a (4.54) del Teorema 4.23 concludiamo che almeno un $kT(B_1)$ deve contenere una palla aperta. Ciò implica che anche $\overline{T(B_1)}$ contiene una palla aperta, sia $B^* = B(y_0; \varepsilon) \subset \overline{T(B_1)}$. Ne segue che

$$B^* - y_0 = B(0; \varepsilon) \subset \overline{T(B_1)} - y_0. \quad (4.66)$$

(b) Proviamo che $B^* - y_0 \subset \overline{T(B_0)}$ dove B_0 è la palla unitaria aperta. Questo risultato si ottiene mostrando che [cf. (4.66)]

$$\overline{T(B_1)} - y_0 \subset \overline{T(B_0)}. \quad (4.67)$$

Sia $y \in \overline{T(B_1)} - y_0$. Allora $y + y_0 \in \overline{T(B_1)}$ e ricordiamo che anche $y_0 \in \overline{T(B_1)}$. Per il teorema 1.12(a) vi sono due successioni u_n e v_n in $T(B_1)$ tali che

$$u_n \rightarrow y + y_0, \quad v_n \rightarrow y_0.$$

Queste due successioni sono immagine per l'operatore T di due successioni w_n e z_n in B_1 , ossia è

$$u_n = Tw_n, \quad v_n = Tz_n.$$

Poiché B_1 ha raggio $1/2$ ne segue che

$$\|w_n - z_n\| \leq \|w_n\| + \|z_n\| < \frac{1}{2} + \frac{1}{2} = 1,$$

così che $w_n - z_n \in B_0$. Da

$$T(w_n - z_n) = Tw_n - Tz_n = u_n - v_n \rightarrow y$$

vediamo che $y \in \overline{T(B_0)}$. Poiché $y \in \overline{T(B_1)} - y_0$ era arbitrario ciò prova la (4.67). Dalla (4.66) abbiamo così

$$B^* - y_0 = B(0; \varepsilon) \subset \overline{T(B_0)}. \quad (4.68)$$

(c) Infine proviamo che

$$V_1 = B(0; \varepsilon/2) \subset T(B_0).$$

Introduciamo le seguenti palle aperte

$$\begin{aligned} B_n &= B(0; 2^{-n}) \subset X \\ V_n &= B(0; \varepsilon/2^n) \subset X. \end{aligned}$$

Poiché T è lineare $\overline{T(B_n)} = 2^{-n}\overline{T(B_0)}$. Dalla (4.68) otteniamo così

$$V_n = B(0; \varepsilon/2^n) \subset \overline{T(B_n)}. \quad (4.69)$$

Sia ora $y \in V_1$ e costruiamo una successione (z_n) che converge ad un punto $x \in B_0$ tale che $y = Tx$. Dalla (4.69) con $n = 1$ abbiamo $V_1 \subset \overline{T(B_1)}$. Quindi $y \in \overline{T(B_1)}$. Per il 1.12(a) vi deve essere un $v \in T(B_1)$ vicino a y , ossia $\|y - v\| < \varepsilon/4$. Ora $v \in T(B_1)$ implica $v = Tx_1$ per qualche $x_1 \in B_1$. Quindi

$$\|y - Tx_1\| < \frac{\varepsilon}{4}.$$

Da ciò e dalla (4.69) per $n = 2$ vediamo che $y - Tx_1 \in V_2 \subset \overline{T(B_2)}$. Come prima concludiamo che v'è un $x_2 \in B_2$ tale che

$$\|(y - Tx_1) - Tx_2\| < \frac{\varepsilon}{8}.$$

Quindi $(y - Tx_1) - Tx_2 \in V_3 \subset \overline{T(B_3)}$ e così via. All' n -mo passo possiamo scegliere un $x_n \in B_n$ tale che

$$\left\| y - \sum_{k=1}^n Tx_k \right\| < \frac{\varepsilon}{2^{n+1}} \quad (n = 1, 2, \dots). \quad (4.70)$$

Sia $z_n = x_1 + \dots + x_n$. Poiché $x_k \in B_k$ abbiamo $\|x_k\| < 1/2^k$. Ciò fornisce per $n > m$

$$\|z_n - z_m\| \leq \sum_{k=m+1}^n \|x_k\| < \sum_{k=m+1}^{\infty} \frac{1}{2^k} \rightarrow 0$$

per $m \rightarrow \infty$. Quindi (z_n) è di Cauchy. (z_n) converge, ossia $z_n \rightarrow x$ perché X è completo. Anche $x \in B_0$ perché B_0 ha raggio 1 e

$$\|x\| \leq \sum_{k=1}^{\infty} \|x_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1. \quad (4.71)$$

Poiché T è continuo $Tz_n \rightarrow Tx$ e la (4.70) mostra che $Tx = y$. Quindi $y \in T(B_0)$. Si noti infine, anche se ciò è inessenziale per la dimostrazione, che la successione $(z_n) \subset B_0$. ■

4.39 Teorema (Applicazione Aperta e Inverso Limitato)

Un operatore lineare limitato T da uno spazio di Banach X su uno spazio di Banach Y è un'applicazione aperta. Inoltre se T è biiettivo T^{-1} è continuo, ossia limitato. ■

Dimostrazione. Proviamo che per ogni insieme aperto $A \subset X$ l'immagine $T(A)$ è aperta in Y . Lo facciamo mostrando che per ogni $y = Tx \in T(A)$ l'insieme $T(A)$ contiene una palla aperta di centro $y = Tx$.

Sia $y = Tx \in T(A)$. Poiché A è aperto contiene una palla aperta di centro x . Quindi $A - x$ contiene una palla aperta di centro 0; sia r il raggio della palla e poniamo $k = 1/r$. Allora $k(A - x)$ contiene la palla unitaria aperta $B(0; 1)$. Il Lemma 4.38 ora implica che $T(k(A - x)) = k[T(A) - Tx]$ contiene una palla aperta di centro 0 e così pure $T(A) - Tx$. Quindi $T(A)$ contiene una palla aperta di centro $Tx = y$. Poiché $y \in T(A)$ era arbitrario $T(A)$ è aperto.

Infine se $T^{-1} : Y \rightarrow X$ esiste è continuo per il Teorema 1.5 perché T è aperto. Poiché T^{-1} è lineare per il Teorema 2.25 esso è limitato per il Teorema 2.31. ■

4.12 Operatori Lineari Chiusi. Teorema del Grafico Chiuso

Non tutti gli operatori lineari di importanza pratica sono limitati. In meccanica quantistica ed in altre applicazioni si ha bisogno molto frequentemente di operatori non limitati. Tuttavia praticamente tutti gli operatori lineari che l'analista deve usare sono quasi sempre operatori lineari cosiddetti chiusi. In questa sezione definiamo gli operatori lineari chiusi su spazi normati e consideriamo alcune delle loro proprietà, in particolare in connessione coll'importante teorema del grafico chiuso che stabilisce le condizioni sufficienti per le quali un operatore lineare chiuso su uno spazio di Banach è limitato.

Incominciamo colla seguente definizione.

4.40 DEFINIZIONE (OPERATORI LINEARI CHIUSI)

Siano X e Y spazi normati e $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare con dominio $\mathcal{D}(T) \subset X$. Allora T è chiamato un *operatore lineare chiuso* se il suo *grafico*

$$\mathcal{G}(T) = \{(x, y) \mid x \in \mathcal{D}(T), y = Tx\}$$

è chiuso nello spazio normato $X \times Y$, dove le due operazioni algebriche di uno spazio vettoriale sono definite in $X \times Y$ nel modo usuale, cioè

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) \\ \alpha(x, y) &= (\alpha x, \alpha y)\end{aligned}$$

(α uno scalare) e la norma su $X \times Y$ è definita da

$$\|(x, y)\| = \|x\| + \|y\|. \quad (4.72)$$

■

Dimostriamo il seguente utile criterio che esprime una proprietà che è spesso presa come definizione di chiusura di un operatore lineare.

4.41 Teorema (Operatore Lineare Chiuso)

Sia $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare, dove $\mathcal{D}(T) \subset X$ ed X e Y sono spazi normati. Allora T è chiuso se e solo se ha la seguente proprietà. Se $x_n \rightarrow x$, dove $x_n \in \mathcal{D}(T)$ e $Tx_n \rightarrow y$ allora $x \in \mathcal{D}(T)$ e $Tx = y$. ■

Dimostrazione. Sia T chiuso. Se $x_n \rightarrow x$ con $x_n \in \mathcal{D}(T)$ e $Tx_n \rightarrow y$ ne segue che la successione $z_n = (x_n, Tx_n) \in \mathcal{G}(T)$ converge in $X \times Y$ a $z = (x, y) \in \overline{\mathcal{G}(T)}$. Ma $\mathcal{G}(T)$ è chiuso e quindi $x \in \mathcal{D}(T)$ e $Tx = y$.

Se per ogni successione convergente $x_n \rightarrow x$, tale che $x_n \in \mathcal{D}(T)$ e $Tx_n \rightarrow y$, vale $x \in \mathcal{D}(T)$ e $Tx = y$, ne segue che ogni successione $z_n = (x_n, Tx_n) \in \mathcal{G}(T)$ convergente in $X \times Y$ converge necessariamente a $z = (x, Tx) \in \mathcal{G}(T)$ e quindi $\mathcal{G}(T)$ è chiuso e quindi T è chiuso. ■

Si noti bene che questa proprietà è differente dalla seguente proprietà di un operatore lineare limitato. Se un operatore lineare T è limitato e quindi continuo e (x_n) è una successione in $\mathcal{D}(T)$ che converge in $\mathcal{D}(T)$ allora anche (Tx_n) converge; cf. 1.14. Ciò non è necessariamente valido per un operatore lineare chiuso, cioè Tx_n potrebbe non convergere. Tuttavia se T è chiuso e due successioni (x_n) e $(\bar{x}_n)$ nel dominio di T convergono al medesimo limite e se le corrispondenti successioni (Tx_n) e $(T\bar{x}_n)$ convergono entrambe allora queste ultime hanno il medesimo limite.

Osserviamo infine che se X e Y sono completi, anche $X \times Y$ con norma definita da (4.72) è completo. Sia infatti (z_n) di Cauchy in $X \times Y$, dove $z_n = (x_n, y_n)$. Allora per ogni $\varepsilon > 0$ v'è un N tale che

$$\|z_n - z_m\| = \|x_n - x_m\| + \|y_n - y_m\| < \varepsilon \quad (m, n > N). \quad (4.73)$$

Quindi (x_n) e (y_n) sono di Cauchy, rispettivamente, in X ed in Y e convergono, cioè $x_n \rightarrow x$ e $y_n \rightarrow y$ perché X e Y sono completi. Ciò implica che $z_n \rightarrow z = (x, y)$ poiché dalla (4.73) per $m \rightarrow \infty$ abbiamo $\|z_n - z\| \leq \varepsilon$ per $n > N$. Poiché la successione di Cauchy (z_n) era arbitraria $X \times Y$ è completo.

Sotto quali condizioni un operatore lineare chiuso è limitato? Una risposta è data dal seguente importante teorema.

4.42 Teorema (Grafico Chiuso)

Siano X e Y spazi di Banach e $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare chiuso con dominio $\mathcal{D}(T) \subset X$. Allora se $\mathcal{D}(T)$ è chiuso in X l'operatore T è limitato. ■

Dimostrazione. Per ipotesi $\mathcal{G}(T)$ è chiuso in $X \times Y$ e $\mathcal{D}(T)$ è chiuso in X . Quindi $\mathcal{G}(T)$ e $\mathcal{D}(T)$ sono completi per il 1.13. Consideriamo ora l'applicazione

$$\begin{aligned} P &: \mathcal{G}(T) \rightarrow \mathcal{D}(T) \\ (x, Tx) &\mapsto x. \end{aligned}$$

P è lineare ed è limitato perché

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|.$$

P è biiettivo; infatti l'applicazione inversa è

$$\begin{aligned} P^{-1} &: \mathcal{D}(T) \rightarrow \mathcal{G}(T) \\ x &\mapsto (x, Tx). \end{aligned}$$

Poiché $\mathcal{G}(T)$ e $\mathcal{D}(T)$ sono completi possiamo applicare il teorema dell'inverso limitato 4.39 e vedere che P^{-1} è limitato, ossia $\|(x, Tx)\| \leq b\|x\|$ per qualche b e tutti gli $x \in \mathcal{D}(T)$. Quindi T è limitato perché

$$\|Tx\| \leq \|Tx\| + \|x\| = \|(x, Tx)\| \leq b\|x\|$$

per tutti gli $x \in \mathcal{D}(T)$. ■

Con degli opportuni esempi si può verificare la seguente affermazione.

La chiusura non implica, se $\mathcal{D}(T)$ non è chiuso, la limitatezza di un operatore lineare. Viceversa la limitatezza non implica la chiusura.

La nostra presente discussione sembra indicare che in connessione con gli operatori non limitati la determinazione dei domini ed i problemi di estensione possono giocare un ruolo basilare. L'affermazione che abbiamo appena enunciato è piuttosto di natura negativa. Dal lato positivo abbiamo invece il seguente lemma.

4.43 Lemma (Operatore Chiuso)

Sia $T : \mathcal{D}(T) \rightarrow Y$ un operatore lineare limitato con dominio $\mathcal{D}(T) \subset X$, dove X e Y sono spazi normati. Ne segue

(a) Se $\mathcal{D}(T)$ è un sottoinsieme chiuso di X allora T è chiuso.

(b) Se T è chiuso e Y è completo allora $\mathcal{D}(T)$ è un sottoinsieme chiuso di X . ■

Dimostrazione. (a) Se (x_n) è in $\mathcal{D}(T)$ e converge, ossia $x_n \rightarrow x$ ed è tale che anche (Tx_n) converge allora $x \in \overline{\mathcal{D}(T)} = \mathcal{D}(T)$ perché $\mathcal{D}(T)$ è chiuso e $Tx_n \rightarrow Tx$ perché T è continuo. Quindi T è chiuso per il Teorema 4.41.

(b) Per $x \in \overline{\mathcal{D}(T)}$ v'è una successione (x_n) in $\mathcal{D}(T)$ tale che $x_n \rightarrow x$; cf. 1.12. Poiché T è limitato la successione Tx_n è di Cauchy ed è $Tx_n \rightarrow y$ perché Y è completo. Poiché T è chiuso $x \in \mathcal{D}(T)$ per il 4.41 (e $Tx = y$). Quindi $\mathcal{D}(T)$ è chiuso perché $x \in \overline{\mathcal{D}(T)}$ era arbitrario. ■

4.13 Problemi

1. Siano $T_n : \ell^2 \rightarrow \ell^2$ definiti da

$$T_n (\xi_1, \xi_2, \dots, \xi_n, \xi_{n+1}, \dots) = (\underbrace{0, 0, \dots, 0}_{n \text{ zeri}}, \xi_n, \xi_{n+1}, \dots).$$

Mostrare che i T_n sono lineari e limitati con norma $\|T_n\| = 1$. Mostrare che la successione T_n converge a 0 fortemente, ma non uniformemente.

2. Siano $T_n : \ell^2 \rightarrow \ell^2$ definiti da

$$T_n (\xi_1, \xi_2, \dots) = (\underbrace{0, 0, \dots, 0}_{n \text{ zeri}}, \xi_1, \xi_2, \dots).$$

Mostrare che i T_n sono lineari e limitati. Mostrare che la successione T_n converge a 0 debolmente, ma non fortemente.

Suggerimento: Si utilizzi per un qualunque funzionale lineare limitato f su ℓ^2 la rappresentazione di Riesz

$$f(x) = \langle x, z \rangle = \sum_{j=1}^{\infty} \xi_j \bar{\zeta}_j.$$

Allora

$$f(T_n x) = \sum_{k=1}^{\infty} \xi_k \bar{\zeta}_{n+k}$$

e quindi ... Per mostrare che la successione T_n non può convergere fortemente si consideri $\|T_m x - T_n x\|$ per $x = (1, 0, 0, \dots)$.

Capitolo 5

Teoria Spettrale degli Operatori Lineari in Spazi Normati

Breve orientamento sul principale contenuto del capitolo

Cominciamo con gli spazi finito dimensionali. La teoria spettrale in questo caso è essenzialmente la teoria degli autovalori delle matrici (Sez. 5.1) ed è molto più semplice di quella degli operatori negli spazi infinito dimensionali. I problemi degli autovalori delle matrici suggeriscono anche parte dell'approccio generale ed alcuni dei concetti della teoria spettrale negli spazi normati infinito dimensionali come definiti nella Sez. 5.2, sebbene il caso infinito dimensionale sia molto più complicato di quello finito dimensionale.

Importanti proprietà dello spettro degli operatori lineari limitati su spazi normati e spazi di Banach sono discussi nelle Sez. 5.3 e 5.4.

Assunzioni generali

Escludiamo lo spazio vettoriale banale $\{0\}$ ed assumiamo che tutti gli spazi siano complessi a meno che non sia diversamente stabilito.

5.1 Teoria Spettrale in Spazi Normati Finito Dimensionali

Sia X uno spazio normato finito dimensionale e $T : X \rightarrow X$ un operatore lineare. La teoria spettrale di tali operatori è più semplice di quella degli operatori definiti su spazi infinito dimensionali. Infatti dalla Sez. 2.10 sappiamo che possiamo rappresentare T con delle matrici (che dipendono dalla scelta delle basi per X) e mostreremo che la teoria spettrale di T è essenzialmente la teoria degli autovalori delle matrici. Perciò cominciamo colle matrici.

Osserviamo che la presente sezione è algebrica, ma faremo ben presto uso della norma a partire dalla prossima sezione.

Per una matrice data (reale o complessa) quadrata a n -righe $A = (\alpha_{jk})$ i concetti di autovalori e di autovettori sono definiti in termini dell'equazione

$$Ax = \lambda x \tag{5.1}$$

come segue.

5.1 DEFINIZIONE (SPETTRO, INSIEME RISOLVENTE DI UNA MATRICE)

Un *autovalore* di una matrice quadrata $A = (\alpha_{jk})$ è un numero λ tale che la (5.1) ha una soluzione $x \neq 0$. Questo x è chiamato un *autovettore* di A corrispondente all'autovalore λ . Gli

autovettori corrispondenti a questo autovalore λ ed il vettore nullo formano un sottospazio vettoriale di X che è chiamato l'*autospazio* di A corrispondente a questo autovalore λ . L'insieme $\sigma(A)$ di tutti gli autovalori di A è chiamato lo *spettro* di A . Il suo complemento $\rho(A) = \mathbb{C} - \sigma(A)$ nel piano complesso è chiamato l'*insieme risolvente* di A . ■

Cosa possiamo dire sull'esistenza degli autovalori di una matrice in generale?

Per rispondere a questa domanda osserviamo dapprima che la (5.1) può essere riscritta

$$(A - \lambda I)x = 0 \quad (5.2)$$

dove I è la matrice unità a n -righe. Questo è un sistema omogeneo di n equazioni lineari in n incognite $\xi_1, \dots, \xi_n$, le componenti di x . Il determinante dei coefficienti è $\det(A - \lambda I)$ e deve essere zero affinché la (5.2) abbia una soluzione $x \neq 0$. Ciò dà l'**equazione caratteristica** di A

$$\det(A - \lambda I) = \begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} - \lambda & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \lambda \end{vmatrix} = 0. \quad (5.3)$$

$\det(A - \lambda I)$ è chiamato il **determinante caratteristico** di A . Sviluppandolo otteniamo un polinomio in λ di grado n , il **polinomio caratteristico** di A . L'equazione (5.3) è chiamata l'*equazione caratteristica* di A .

Il nostro risultato è il teorema basilare seguente.

5.2 Teorema (Autovalori di una Matrice)

Gli autovalori di una matrice quadrata a n -righe $A = (\alpha_{jk})$ sono dati dalle soluzioni dell'equazione caratteristica (5.3) di A . Quindi A ha almeno un autovalore (e al più n autovalori numericamente differenti). ■

La seconda affermazione vale perché, per il cosiddetto teorema fondamentale dell'algebra ed il teorema della fattorizzazione, un polinomio di grado n e con coefficienti in $\mathbb{C}$ ha una radice in $\mathbb{C}$ (ed al più n radici numericamente differenti). Si noti che le radici possono essere complesse anche se A è reale.

Come possiamo applicare il nostro risultato ad un operatore lineare $T : X \rightarrow X$ su uno spazio normato X di dimensioni n ? Sia $e = \{e_1, \dots, e_n\}$ una base qualunque per X e $T_e = (\alpha_{jk})$ la matrice che rappresenta T rispetto a questa base (i cui elementi sono dati in un ordine fissato). Allora gli autovalori della matrice T_e sono chiamati **autovalori dell'operatore T** ed analogamente per lo *spettro* e per l'*insieme risolvente*. Ciò è giustificato dal seguente teorema.

5.3 Teorema (Autovalori di un Operatore)

Tutte le matrici che rappresentano un dato operatore lineare $T : X \rightarrow X$ su uno spazio normato finito dimensionale X rispetto alle varie basi per X hanno i medesimi autovalori. ■

Dimostrazione. Dobbiamo vedere che cosa capita nel passaggio da una base per X ad un'altra. Siano $e = (e_1, \dots, e_n)$ ed $e' = (e'_1, \dots, e'_n)$ due basi qualunque per T scritte come vettori riga. Per definizione di base ciascun e_j è una combinazione lineare degli e'_k e viceversa. Possiamo scrivere cioè

$$e' = eC \quad \text{o} \quad e'^T = C^T e^T \quad (5.4)$$

dove C è una matrice quadrata a n -righe non singolare. Ogni x ha un'unica rappresentazione rispetto a ciascuna delle due basi, ossia

$$x = e\xi = \sum \xi_j e_j = e'\xi' = \sum \xi'_k e'_k$$

dove $\xi = (\xi_j)$ e $\xi' = (\xi'_k)$ sono vettori colonna. Da questa e dalla (5.4) abbiamo $e\xi = e'\xi' = eC\xi'$. Quindi

$$\xi = C\xi'. \quad (5.5)$$

Analogamente da $Tx = y = e\eta = e'\eta'$, dove $\eta = (\eta_j)$ e $\eta' = (\eta'_j)$, abbiamo

$$\eta = C\eta'. \quad (5.6)$$

Di conseguenza, se A e A' denotano le matrici che rappresentano T rispetto ad e ed e' , allora

$$\eta = A\xi \quad \text{e} \quad \eta' = A'\xi'$$

e da ciò e dalle (5.5) e (5.6)

$$CA'\xi' = C\eta' = \eta = A\xi = AC\xi'.$$

Moltiplicando a sinistra per C^{-1} otteniamo la legge di trasformazione

$$A' = C^{-1}AC \quad (5.7)$$

con C determinato dalle basi secondo la (5.4) (e indipendente da T). Usando la (5.7) e $\det(C^{-1})\det C = 1$ possiamo mostrare ora che i determinanti caratteristici di A e A' sono uguali

$$\begin{aligned} \det(A' - \lambda I) &= \det(C^{-1}AC - \lambda C^{-1}IC) \\ &= \det(C^{-1}(A - \lambda I)C) \\ &= \det(C^{-1})\det(A - \lambda I)\det C \\ &= \det(A - \lambda I). \end{aligned} \quad (5.8)$$

L'eguaglianza degli autovalori di A e di A' segue ora dal Teorema 5.2. ■

Menzioniamo di passaggio che possiamo anche esprimere i nostri risultati in termini del seguente concetto, che è di interesse generale. Una matrice $n \times n$ A' è detta *similare* ad una matrice $n \times n$ A se esiste una matrice C non singolare tale che vale la (5.7). A e A' sono allora chiamate *matrici similari*. In termini di questo concetto la nostra dimostrazione mostra quanto segue.

- (i) *Due matrici rappresentanti il medesimo operatore lineare T su uno spazio normato finito dimensionale X relative a due basi qualunque per X sono similari.*
- (ii) *Matrici similari hanno gli stessi autovalori.*

Inoltre i Teoremi 5.2 e 5.3 implicano il seguente teorema.

5.4 Teorema (Esistenza degli Autovalori)

Un operatore lineare su uno spazio normato complesso finito dimensionale $X \neq \{0\}$ ha almeno un autovalore. ■

In generale non possiamo dire di più.

Inoltre la (5.8) con $\lambda = 0$ dà $\det A' = \det A$. Quindi il valore del determinante rappresenta una proprietà intrinseca dell'operatore T , così che possiamo parlare senza ambiguità della quantità $\det T$.

5.2 Concetti Basilari

Nella sezione precedente gli spazi erano finito dimensionali. In questa sezione consideriamo spazi normati di dimensione qualunque e vedremo che negli spazi infinito dimensionali la teoria spettrale diventa molto più complicata.

Sia $X \neq \{0\}$ uno spazio normato complesso, $T : X \rightarrow X$ un operatore lineare e $\mathcal{R}(T) \subset X$ la sua immagine. A T associamo l'operatore

$$T_\lambda = T - \lambda I \quad (5.9)$$

dove λ è un numero complesso ed I è l'operatore identità su X . Indichiamo con $\mathcal{R}(T_\lambda)$ l'immagine di T_λ . Se l'operatore lineare $T_\lambda : X \rightarrow \mathcal{R}(T_\lambda)$ ammette inverso lo indichiamo con $R_\lambda(T)$, cioè

$$R_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1} \quad (5.10)$$

e lo chiamiamo l'*operatore risolvente* di T o, semplicemente, il **risolvente** di T . Invece di $R_\lambda(T)$ scriviamo semplicemente R_λ se è chiaro a quale operatore T ci si riferisce nella specifica discussione.

Il nome “risolvente” è appropriato perché $R_\lambda(T)$ aiuta a risolvere l'equazione $T_\lambda x = y$. Infatti $x = T_\lambda^{-1} y = R_\lambda(T)y$ purché $R_\lambda(T)$ esista.

Più importante ancora è il fatto che lo studio delle proprietà di R_λ risulta basilare per una comprensione dell'operatore T stesso. Naturalmente molte proprietà di T_λ e R_λ dipendono da λ . La *teoria spettrale* ha a che fare con queste proprietà. Ad esempio saremo interessati all'insieme di tutti i λ nel piano complesso tali che R_λ esista. La limitatezza di R_λ è un'altra proprietà che risulterà essenziale. Ci chiederemo anche per quali λ il dominio di R_λ è denso in X e ciò solo per indicare alcuni aspetti.

Notiamo inoltre che $R_\lambda(T)$ è un operatore lineare per il Teorema 2.25(b).

Per lo studio di T , T_λ e R_λ avremo bisogno dei seguenti concetti che sono basilari nella teoria spettrale.

5.5 DEFINIZIONE (VALORI REGOLARI, INSIEME RISOLVENTE, SPETTRO)

Sia $X \neq \{0\}$ uno spazio normato complesso e $T : X \rightarrow X$ un operatore lineare. Un *valore regolare* di T è un numero complesso tale che

R1 $R_\lambda(T)$ esiste,

R2 $R_\lambda(T)$ è limitato,

R3 $\overline{\mathcal{R}(T_\lambda)} = X$, ossia $R_\lambda(T)$ è definito su un insieme che è denso in X .

L'*insieme risolvente* $\rho(T)$ di T è l'insieme di tutti i valori regolari λ di T . Il suo complemento $\sigma(T) = \mathbb{C} - \rho(T)$ nel piano complesso $\mathbb{C}$ è chiamato lo *spettro* di T ed un $\lambda \in \sigma(T)$ è chiamato un *valore spettrale* di T . Inoltre lo spettro $\sigma(T)$ è partizionato in tre insiemi disgiunti come segue.

Lo **spettro puntuale** o *spettro discreto* $\sigma_p(T)$ è l'insieme tale che $R_\lambda(T)$ non esiste. Un $\lambda \in \sigma_p(T)$ è chiamato un *autovalore* di T .

Lo **spettro continuo** $\sigma_c(T)$ è l'insieme tale che $R_\lambda(T)$ esiste e soddisfa (R3) ma non (R2), cioè $R_\lambda(T)$ è definito in un insieme che è denso in X , ma non è limitato.

Lo **spettro residuale** $\sigma_r(T)$ è l'insieme tale che $R_\lambda(T)$ esiste (e può essere limitato o no) ma non soddisfa (R3), cioè il dominio di $R_\lambda(T)$ non è denso in X . ■

Per evitare incomprensioni notiamo che alcuni degli insiemi in questa definizione possono essere vuoti. Si tratta di un problema di esistenza che dovremo discutere. Per esempio $\sigma_c(T) = \sigma_r(T) = \emptyset$ nel caso finito dimensionale come sappiamo dalla Sez. 5.1.

Le condizioni stabilite nella Definizione precedente possono essere riassunte nella seguente tabella

Soddisfatto			Non soddisfatto	λ appartiene a
(R1)	(R2)	(R3)		$\rho(T)$
(R1)		(R3)	(R1)	$\sigma_p(T)$
(R1)			(R2)	$\sigma_c(T)$
(R1)			(R3)	$\sigma_r(T)$

Per aiutare la comprensione di questi concetti cominciamo con alcune osservazioni generali.

Notiamo dapprima che i quattro insiemi della tabella sono disgiunti e che la loro unione è l'intero piano complesso

$$\begin{aligned}\mathbb{C} &= \rho(T) \cup \sigma(T) \\ &= \rho(T) \cup \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T).\end{aligned}$$

Inoltre se il risolvente $R_\lambda(T)$ esiste è lineare per il Teorema 2.25 come s'è detto prima. Questo teorema mostra anche che $R_\lambda(T) : \mathcal{R}(T_\lambda) \rightarrow X$ esiste se e solo se $T_\lambda x = 0$ implica $x = 0$, cioè se lo spazio nullo di T_λ è $\{0\}$ ($\mathcal{R}(T_\lambda)$ denota l'immagine di T_λ).

Quindi se $T_\lambda x = (T - \lambda I)x = 0$ per qualche $x \neq 0$ allora $\lambda \in \sigma_p(T)$ per definizione, cioè λ è un autovalore di T . Allora x è chiamato un **autovettore** di T (o **autofunzione** di T se X è uno spazio funzionale) corrispondente all'autovalore λ . Il sottospazio di X consistente dello 0 e di tutti gli autovettori di T corrispondenti ad un autovalore λ di T è chiamato l'*autospazio* di T corrispondente all'autovalore λ .

Vediamo che la nostra presente definizione di autovalore è in armonia con quella data alla Sez. 5.1. Vediamo anche che lo spettro di un operatore lineare in uno spazio finito dimensionale è un *puro spettro puntuale*, cioè sia lo spettro continuo che quello residuale sono vuoti, come s'è detto prima, così che ogni valore spettrale è un autovalore.

Una motivazione per la partizione di $\sigma(T) - \sigma_p(T)$ in $\sigma_c(T)$ e $\sigma_r(T)$ è data dal fatto che $\sigma_r(T) = \emptyset$ per l'importante classe degli operatori autoaggiunti sugli spazi di Hilbert.

Nella presente discussione il teorema dell'inverso limitato 4.39 contribuisce come segue. Se $T : X \rightarrow X$ è limitato e lineare ed X è completo e se per qualche λ il risolvente $R_\lambda(T)$ esiste ed è definito su tutto lo spazio X allora per questo λ il risolvente è limitato.

Inoltre i seguenti fatti (necessari più in là) possono anche essere utili per una migliore comprensione dei presenti concetti.

5.6 Lemma (Dominio del Risolvente)

Sia X uno spazio di Banach complesso, $T : X \rightarrow X$ un operatore lineare e $\lambda \in \rho(T)$. Si assuma che (a) T sia chiuso o che (b) T sia limitato. Allora $R_\lambda(T)$ è definito su tutto lo spazio X . ■

Dimostrazione. (a) Poiché T è chiuso lo è anche T_λ per il 4.41. Poiché il grafico di R_λ si ottiene da quello di T_λ scambiando coordinate con ascisse in $X \times X$, R_λ è chiuso. R_λ è limitato per (R2). Quindi il suo dominio $\mathcal{D}(R_\lambda)$ è chiuso per il 4.43(b) applicato a R_λ , così che (R3) implica $\mathcal{D}(R_\lambda) = \overline{\mathcal{D}(R_\lambda)} = X$.

(b) Poiché X è chiuso T è chiuso per il 4.43(a) e l'enunciato segue dalla parte (a) di questa dimostrazione. ■

Riportiamo infine una proprietà basilare degli autovettori.

5.7 Teorema (Indipendenza Lineare)

Gli autovettori $x_1, \dots, x_n$ corrispondenti ad autovalori differenti $\lambda_1, \dots, \lambda_n$ di un operatore lineare T in uno spazio vettoriale X costituiscono un insieme linearmente indipendente. ■

Dimostrazione. Assumiamo che $\{x_1, \dots, x_n\}$ sia linearmente dipendente e deriviamo una contraddizione. Sia x_m il primo dei vettori che sia una combinazione lineare dei precedenti, ossia

$$x_m = \alpha_1 x_1 + \dots + \alpha_{m-1} x_{m-1}. \quad (5.11)$$

Allora $\{x_1, \dots, x_{m-1}\}$ è linearmente indipendente. Applicando $T - \lambda_m I$ ad entrambi i membri della (5.11) otteniamo

$$\begin{aligned} (T - \lambda_m I)x_m &= \sum_{j=1}^{m-1} \alpha_j (T - \lambda_m I)x_j \\ &= \sum_{j=1}^{m-1} \alpha_j (\lambda_j - \lambda_m) x_j. \end{aligned}$$

Poiché x_m è un autovettore corrispondente a λ_m il membro a sinistra è zero. Poiché i vettori a destra formano un insieme linearmente indipendente dobbiamo avere

$$\alpha_j (\lambda_j - \lambda_m) = 0, \quad \text{quindi} \quad \alpha_j = 0 \quad (j = 1, \dots, m-1)$$

giacché $\lambda_j - \lambda_m \neq 0$. Ma allora $x_m = 0$ per la (5.11). Ciò contraddice il fatto che $x_m \neq 0$ perché x_m è un autovettore e completa la dimostrazione. ■

5.3 Proprietà Spettrali degli Operatori Lineari Limitati

Le proprietà dello spettro dipendono dal tipo di spazio su cui l'operatore è definito e dal tipo di operatore che si considera. Questa situazione suggerisce di studiare separatamente larghe classi di operatori con proprietà spettrali comuni ed in questa sezione incominciamo con gli operatori lineari limitati T su uno spazio di Banach complesso X . Così $T \in B(X, X)$, dove X è completo; cf. Sez. 2.11.

Il nostro primo teorema è un teorema chiave in varie parti della teoria, come vedremo.

5.8 Teorema (Inverso)

Sia $T \in B(X, X)$, dove X è uno spazio di Banach. Se $\|T\| < 1$ allora $(I - T)^{-1}$ esiste ed è un operatore lineare limitato su tutto lo spazio X e

$$(I - T)^{-1} = \sum_{j=0}^{\infty} T^j = I + T + T^2 + \dots \quad (5.12)$$

[dove la serie a destra è convergente secondo la norma su $B(X, X)$]. ■

Dimostrazione. Abbiamo $\|T^j\| \leq \|T\|^j$ per la (2.20), Sez. 2.8. Ricordiamo che la serie geometrica $\sum \|T\|^j$ converge per $\|T\| < 1$. Quindi la serie nella (5.12) è assolutamente convergente per $\|T\| < 1$. Poiché X è completo lo è pure $B(X, X)$ per il Teorema 2.40. L'assoluta convergenza perciò implica la convergenza come sappiamo dalla Sez. 2.3.

Indichiamo la somma della serie nella (5.12) con S . Rimane da mostrare che $S = (I - T)^{-1}$. A questo scopo calcoliamo

$$\begin{aligned} & (I - T)(I + T + \cdots + T^m) \\ &= (I + T + \cdots + T^m)(I - T) \\ &= I - T^{m+1}. \end{aligned} \quad (5.13)$$

Lasciamo ora $m \rightarrow \infty$. Allora $T^{m+1} \rightarrow 0$ perché $\|T\| < 1$. Otteniamo così

$$(I - T)S = S(I - T) = I. \quad (5.14)$$

Ciò mostra che $S = (I - T)^{-1}$. ■

Come prima applicazione di questo teorema proviamo l'importante fatto che lo spettro di un operatore lineare limitato è un insieme chiuso nel piano complesso. (mostreremo che $\sigma \neq \emptyset$ in 5.18.).

5.9 Teorema (Spettro Chiuso)

L'insieme risolvente $\rho(T)$ di un operatore lineare limitato T su uno spazio di Banach complesso X è aperto; quindi lo spettro $\sigma(T)$ è chiuso. ■

Dimostrazione. Se $\rho(T) = \emptyset$ è aperto. (In effetti $\rho(T) \neq \emptyset$ come vedremo nel Teorema 5.11.) Sia $\rho(T) \neq \emptyset$. Per un $\lambda_0 \in \rho(T)$ fisso ed un $\lambda \in \mathbb{C}$ qualsiasi abbiamo

$$\begin{aligned} T - \lambda I &= T - \lambda_0 I - (\lambda - \lambda_0)I \\ &= (T - \lambda_0 I)[I - (\lambda - \lambda_0)(T - \lambda_0 I)^{-1}]. \end{aligned}$$

Indicando l'operatore in parentesi $[\cdots]$ con V possiamo scriverlo nella forma

$$T_\lambda = T_{\lambda_0} V \quad \text{dove} \quad V = I - (\lambda - \lambda_0)R_{\lambda_0}. \quad (5.15)$$

Poiché $\lambda_0 \in \rho(T)$ e T è limitato il Lemma 5.6(b) implica che $R_{\lambda_0} = T_{\lambda_0}^{-1} \in B(X, X)$. Inoltre il Teorema 5.8 mostra che V ha un inverso

$$V^{-1} = \sum_{j=0}^{\infty} [(\lambda - \lambda_0)R_{\lambda_0}]^j = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^j \quad (5.16)$$

in $B(X, X)$ per tutti i λ tali che $\|(\lambda - \lambda_0)R_{\lambda_0}\| < 1$, ossia

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}. \quad (5.17)$$

Poiché $T_{\lambda_0}^{-1} = R_{\lambda_0} \in B(X, X)$ vediamo da ciò e dalla (5.15) che per ogni λ che soddisfa la (5.17) l'operatore T_λ ha un inverso

$$R_\lambda = T_\lambda^{-1} = (T_{\lambda_0} V)^{-1} = V^{-1} R_{\lambda_0}. \quad (5.18)$$

Quindi (5.17) rappresenta un intorno di λ_0 consistente di valori regolari λ di T . Poiché $\lambda_0 \in \rho(T)$ era arbitrario $\rho(T)$ è aperto così che il complemento $\sigma(T) = \mathbb{C} - \rho(T)$ è chiuso. ■

È di grande importanza osservare che in questa dimostrazione abbiamo anche ottenuto un rappresentazione basilare del risolvente con una serie di potenze di λ . Infatti dalle (5.16)-(5.18) otteniamo immediatamente il seguente teorema.

5.10 Teorema (Rappresentazione del Risolvente)

Per X e T come nel Teorema 5.9 ed ogni $\lambda_0 \in \rho(T)$ il risolvente $R_\lambda(T)$ ha la rappresentazione

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^{j+1}, \quad (5.19)$$

la serie essendo assolutamente convergente per ogni λ nel disco aperto dato da [cf. (5.17)]

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|} \quad (5.20)$$

nel piano complesso. Questo disco è un sottoinsieme di $\rho(T)$. ■

Questo teorema ci permetterà di applicare l'analisi complessa alla teoria spettrale, come vedremo nella Sez. 5.5.

Come ulteriore conseguenza del Teorema 5.8 proviamo l'importante fatto che per un operatore lineare limitato lo spettro è un insieme limitato nel piano complesso.

5.11 Teorema (Spettro)

Lo spettro $\sigma(T)$ di un operatore limitato $T : X \rightarrow X$ su uno spazio di Banach complesso X è compatto e giace nel disco dato da

$$|\lambda| \leq \|T\|. \quad (5.21)$$

Quindi l'insieme risolvente $\rho(T)$ di T è non vuoto. [in 5.18 mostreremo che $\sigma(T) \neq \emptyset$.] ■

Dimostrazione. Sia $\lambda \neq 0$. Dal Teorema 5.8 otteniamo la rappresentazione

$$R_\lambda = (T - \lambda I)^{-1} = -\frac{1}{\lambda} \left(I - \frac{1}{\lambda} T \right)^{-1} = -\frac{1}{\lambda} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda} T \right)^j \quad (5.22)$$

dove per il Teorema 5.8 la serie converge per tutti i λ tali che

$$\left\| \frac{1}{\lambda} T \right\| = \frac{\|T\|}{|\lambda|} < 1 \quad \text{cioè} \quad |\lambda| > \|T\|.$$

Il medesimo teorema mostra anche che un qualunque tale λ è in $\rho(T)$. Quindi lo spettro $\sigma(T) = \mathbb{C} - \rho(T)$ deve giacere nel disco (5.21), così che $\sigma(T)$ è limitato. Inoltre $\sigma(T)$ è chiuso per il Teorema 5.9. Quindi $\sigma(T)$ è compatto. ■

Poiché dal teorema appena dimostrato sappiamo che per un operatore lineare limitato T su uno spazio di Banach lo spettro è limitato sembra naturale chiedersi qual'è il disco minimo attorno all'origine che contiene l'intero spettro. Questa domanda suggerisce il seguente concetto.

5.12 DEFINIZIONE (RAGGIO SPETTRALE)

Il *raggio spettrale* $r_\sigma(T)$ di un operatore $T \in B(X, X)$ su uno spazio di Banach complesso X è il raggio

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$$

del più piccolo disco chiuso centrato nell'origine del piano complesso di λ che contiene $\sigma(T)$. ■

Dalla (5.21) è ovvio che per il raggio spettrale di un operatore lineare limitato su uno spazio di Banach complesso abbiamo

$$r_\sigma(T) \leq \|T\| \quad (5.23)$$

e nella Sez. 5.5 proveremo che

$$r_\sigma(T) = \lim_{n \rightarrow \infty} (\|T^n\|)^{1/n}. \quad (5.24)$$

5.4 Ulteriori Proprietà del Risolvente e dello Spettro

Alcune ulteriori proprietà del risolvente interessanti e basilari al medesimo tempo sono espresse nel seguente Teorema.

5.13 Teorema (Equazione del Risolvente, Commutatività)

Siano X uno spazio di Banach complesso, $T \in B(X, X)$ e $\lambda, \mu \in \rho(T)$ [cf. 5.5]. Allora

(a) Il risolvente R_λ di T soddisfa alla identità di Hilbert o equazione del risolvente

$$R_\mu - R_\lambda = (\mu - \lambda)R_\mu R_\lambda \quad [\lambda, \mu \in \rho(T)]. \quad (5.25)$$

(b) R_λ commuta con un qualunque $S \in B(X, X)$ che commuta con T .

(c) Abbiamo

$$R_\lambda R_\mu = R_\mu R_\lambda \quad [\lambda, \mu \in \rho(T)]. \quad (5.26)$$

■

Dimostrazione. (a) Per il 5.6 l'immagine di T_λ è tutto X . Quindi $I = T_\lambda R_\lambda$ dove I è l'operatore identità su X . È anche $I = R_\mu T_\mu$. Di conseguenza

$$\begin{aligned} R_\mu - R_\lambda &= R_\mu (T_\lambda R_\lambda) - (R_\mu T_\mu) R_\lambda \\ &= R_\mu (T_\lambda - T_\mu) R_\lambda \\ &= R_\mu [T - \lambda I - (T - \mu I)] R_\lambda \\ &= (\mu - \lambda) R_\mu R_\lambda. \end{aligned}$$

(b) Per ipotesi $ST = TS$. Quindi $ST_\lambda = T_\lambda S$. Usando $I = T_\lambda R_\lambda = R_\lambda T_\lambda$ otteniamo così

$$R_\lambda S = R_\lambda ST_\lambda R_\lambda = R_\lambda T_\lambda S R_\lambda = S R_\lambda.$$

(c) R_μ commuta con T per (b). Quindi R_λ commuta con R_μ per (b). ■

Il nostro prossimo risultato sarà l'importante teorema dell'applicazione spettrale e parliamo con una motivazione suggerita dalla teoria degli autovalori delle matrici.

Se λ è un autovalore di una matrice quadrata A allora $Ax = \lambda x$ per qualche $x \neq 0$. Applicando A si ottiene

$$A^2 x = A \lambda x = \lambda A x = \lambda^2 x.$$

Continuando in questa maniera abbiamo per ogni intero positivo m

$$A^m x = \lambda^m x;$$

ossia se λ è un autovalore di A allora λ^m è un autovalore di A^m . Più generalmente allora

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0$$

è un autovalore della matrice

$$p(A) = \alpha_n A^n + \alpha_{n-1} A^{n-1} + \cdots + \alpha_0 I.$$

Si può mostrare che si ottengono così tutti gli autovalori della matrice $p(A)$.

È assai notevole che questa proprietà si estenda agli spazi di Banach complessi di dimensioni qualunque, come dimostreremo. Nella dimostrazione useremo il fatto che un operatore lineare limitato ha uno spettro non vuoto. Questo lo mostreremo più in là con i metodi dell'analisi complessa.

Una notazione conveniente per formulare il teorema è

$$p(\sigma(T)) = \{\mu \in \mathbb{C} \mid \mu = p(\lambda), \lambda \in \sigma(T)\}, \quad (5.27)$$

cioè $p(\sigma(T))$ è l'insieme di tutti i numeri complessi μ tali che $\mu = p(\lambda)$ per qualche $\lambda \in \sigma(T)$. Useremo anche $p(\rho(T))$ con un significato simile.

5.14 Teorema (Applicazione Spettrale per i Polinomi)

Sia X uno spazio di Banach complesso, $T \in B(X, X)$ e

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0 \quad (\alpha_n \neq 0).$$

Allora

$$\sigma(p(T)) = p(\sigma(T)); \quad (5.28)$$

ossia lo spettro $\sigma(p(T))$ dell'operatore

$$p(T) = \alpha_n T^n + \alpha_{n-1} T^{n-1} + \cdots + \alpha_0 I$$

consiste precisamente di tutti quei valori che il polinomio p assume sullo spettro $\sigma(T)$ di T . ■

Dimostrazione. Assumiamo che $\sigma(T) \neq \emptyset$; ciò sarà dimostrato in 5.18. Il caso $n = 0$ è banale; allora $p(\sigma(T)) = \{\alpha_0\} = \sigma(p(T))$. Sia $n > 0$. Nella parte (a) proviamo che

$$\sigma(p(T)) \subset p(\sigma(T)) \quad (5.29)$$

e nella parte (b) che

$$p(\sigma(T)) \subset \sigma(p(T)), \quad (5.30)$$

così che otteniamo la (5.28). I dettagli sono i seguenti.

(a) Sia $\mu \in \sigma(p(T))$ e supponiamo per assurdo che $\mu \notin p(\sigma(T))$. Allora non esiste alcun $\lambda \in \sigma(T)$ per cui sia $\mu = p(\lambda)$.

Per semplicità scriviamo $S = p(T)$ e

$$S_\mu = p(T) - \mu I.$$

Poiché X è complesso il polinomio dato da $s_\mu(\lambda) = p(\lambda) - \mu$ deve fattorizzare completamente in termini lineari, ossia

$$s_\mu(\lambda) = p(\lambda) - \mu = \alpha_n(\lambda - \gamma_1)(\lambda - \gamma_2) \cdots (\lambda - \gamma_n), \quad (5.31)$$

dove $\gamma_1, \dots, \gamma_n$ sono gli zeri di s_μ (che dipendono naturalmente da μ), ossia

$$s_\mu(\gamma_j) = p(\gamma_j) - \mu = 0$$

per ogni γ_j . Per l'ipotesi assurda fatta nessun $\lambda \in \sigma(T)$ può essere radice di $s_\mu(\lambda)$ e quindi ciascun γ_j è in $\rho(T)$ e ciascun $T - \gamma_j I$ ha un inverso limitato che, per il 5.6, è definito su tutto X . In corrispondenza alla (5.31) abbiamo

$$S_\mu = p(T) - \mu I = \alpha_n(T - \gamma_1 I)(T - \gamma_2 I) \cdots (T - \gamma_n I).$$

e quindi per la (2.13) nella Sez. 2.6 S_μ^{-1} esiste e vale

$$S_\mu^{-1} = \frac{1}{\alpha_n}(T - \gamma_n I)^{-1} \cdots (T - \gamma_2 I)^{-1}(T - \gamma_1 I)^{-1}.$$

Quindi in questo caso $\mu \in \rho(p(T))$ in contraddizione con l'ipotesi assurda fatta. Da ciò concludiamo che

$$\sigma(p(T)) \subset p(\sigma(T)).$$

(b) Dimostriamo ora la (5.30)

$$p(\sigma(T)) \subset \sigma(p(T)).$$

Lo facciamo mostrando che

$$\kappa \in p(\sigma(T)) \implies \kappa \in \sigma(p(T)). \quad (5.32)$$

Sia $\kappa \in p(\sigma(T))$. Per definizione ciò significa che

$$\kappa = p(\beta) \quad \text{per qualche} \quad \beta \in \sigma(T).$$

Ci sono ora due possibilità

(A) $T - \beta I$ non ha inverso.

(B) $T - \beta I$ ha inverso.

Consideriamo separatamente questi due casi.

(A) Da $\kappa = p(\beta)$ abbiamo $p(\beta) - \kappa = 0$. Quindi β è uno zero del polinomio

$$s_\kappa(\lambda) = p(\lambda) - \kappa.$$

Ne segue che possiamo scrivere

$$s_\kappa(\lambda) = p(\lambda) - \kappa = (\lambda - \beta)g(\lambda),$$

dove $g(\lambda)$ denota il prodotto degli altri $n - 1$ fattori lineari e di α_n . In corrispondenza a questa rappresentazione abbiamo

$$S_\kappa(T) = p(T) - \kappa I = (T - \beta I)g(T). \quad (5.33)$$

Siccome i fattori di $g(T)$ commutano tutti con $(T - \beta I)$ abbiamo anche

$$S_\kappa = g(T)(T - \beta I). \quad (5.34)$$

Supponiamo per assurdo che $\kappa \in \rho(p(T))$ e che quindi S_κ abbia un inverso. Allora la (5.33) e la (5.34) ci danno

$$I = (T - \beta I)g(T)S_\kappa^{-1} = S_\kappa^{-1}g(T)(T - \beta I)$$

che mostra che $(T - \beta I)$ ha un inverso in contraddizione con la nostra ipotesi. Dalla contraddizione segue che $\kappa \in \sigma(p(T))$. Poiché $\kappa \in p(\sigma(T))$ era arbitrario ciò prova la (5.32) sotto l'ipotesi che $(T - \beta I)$ non abbia un inverso.

(B) Supponiamo come prima che sia $\kappa = p(\beta)$ per qualche $\beta \in \sigma(T)$, ma ora assumiamo che l'inverso $(T - \beta I)^{-1}$ esista. Allora per l'immagine di $(T - \beta I)$ dobbiamo avere

$$\mathcal{R}(T - \beta I) \neq X \quad (5.35)$$

perché altrimenti $(T - \beta I)^{-1}$ sarebbe limitato per il teorema dell'inverso limitato 4.39 applicato a $T - \beta I$, così che $\beta \in \rho(T)$, che contraddirebbe $\beta \in \sigma(T)$. Dalla (5.33) e dalla (5.35) otteniamo

$$\mathcal{R}(S_\kappa) \neq X.$$

Ciò mostra che $\kappa \in \sigma(p(T))$, perché $\kappa \in \rho(p(T))$ implicherebbe che $\mathcal{R}(S_\kappa) = X$ per il Lemma 5.6(b) applicato a $p(T)$. Ciò prova la (5.32) sotto l'ipotesi che $T - \beta I$ abbia un inverso. Il teorema è dimostrato. ■

5.5 Uso dell'Analisi Complessa nella Teoria Spettrale

Uno strumento importante nella teoria spettrale è l'analisi complessa. Connessioni tra le due aree possono essere ottenute mediante gli integrali complessi curvilinei e le serie di potenze. Noi useremo solamente le serie di potenze. In questo modo saremo capaci di mantenere la discussione ad un livello più elementare ed avremo bisogno solamente dei pochi concetti e fatti basilari che seguono.

Uno spazio metrico è detto *connesso* se non è l'unione di due insiemi aperti non vuoti disgiunti. Un sottoinsieme di uno spazio metrico è detto *connesso* se è connesso quando è riguardato come sottospazio.

Per un **dominio** G nel piano complesso $\mathbb{C}$ intendiamo un sottoinsieme aperto connesso G di $\mathbb{C}$.

Si può mostrare che un sottoinsieme aperto G di $\mathbb{C}$ è connesso se e solo se ciascuna coppia di punti di G può essere congiunta da una linea spezzata consistente di un numero finito di segmenti di retta con tutti i loro punti appartenenti a G . (In molti libri sull'analisi complessa ciò è usato come definizione di connessione.)

Il risolvete R_λ è un operatore che dipende dal parametro complesso λ . Ciò ci suggerisce di introdurre il nostro presente approccio nel modo seguente.

Per una *funzione a valori vettoriali* o **funzione operatoriale** di una variabile complessa λ intendiamo un'applicazione

$$\begin{aligned} S : \Lambda &\rightarrow B(X, X) \\ \lambda &\mapsto S_\lambda \end{aligned} \quad (5.36)$$

dove Λ è un sottoinsieme qualunque del piano complesso di λ . (Scriviamo S_λ invece di $S(\lambda)$ per avere una notazione simile ad R_λ , perché nel seguito considereremo $S_\lambda = R_\lambda$ e $\Lambda = \rho(T)$.)

5.15 DEFINIZIONE (OLOMORFISMO LOCALE, OLOMORFISMO)

Sia Λ un sottoinsieme aperto di $\mathbb{C}$ ed X uno spazio di Banach complesso. Allora S_λ nella (5.36) è detto essere *localmente olomorfo* su Λ se per ogni $\lambda_0 \in \Lambda$ ha una rappresentazione in serie di potenze

$$S_\lambda = \sum_{j=0}^{\infty} S_j (\lambda - \lambda_0)^j$$

con $S_j \in B(X, X)$ e raggio di convergenza non nullo. S_λ è detto *olomorfo* su Λ se S_λ è localmente olomorfa su Λ e Λ è un dominio. S è detto *olomorfo nel punto* $\lambda_0 \in \mathbb{C}$ se S_λ è olomorfo in un ε -intorno di λ_0 . ■

Questa definizione tiene conto della seguente situazione. L'insieme risolvente $\rho(T)$ di un operatore lineare limitato T è aperto (per il 5.9) ma non sempre può essere un dominio, così che in generale è l'unione di domini disgiunti (insiemi aperti connessi disgiunti). Vedremo che il risolvente è olomorfo in ogni punto di $\rho(T)$. Quindi in qualsiasi caso è localmente olomorfo su $\rho(T)$ (definendo così una funzione operatoriale olomorfa su ciascuno di questi domini) ed è olomorfo su $\rho(T)$ se e solo se $\rho(T)$ è connesso, così che $\rho(T)$ è un singolo dominio.

Come preannunciato lo strumento chiave per l'applicazione dell'analisi complessa alla teoria spettrale sarà il Teorema 5.10. In particolare da esso e dalla definizione 5.15 sopra data possiamo derivare immediatamente il seguente teorema.

5.16 Teorema (Olomorfia del Risolvente)

Il risolvente $R_\lambda(T)$ di un operatore lineare limitato $T : X \rightarrow X$ su uno spazio di Banach complesso è olomorfo in ogni punto λ_0 dell'insieme risolvente $\rho(T)$ di T . Quindi è localmente olomorfo su $\rho(T)$. ■

Inoltre $\rho(T)$ è l'insieme più grande su cui il risolvente di T è localmente olomorfo. Infatti il risolvente non può essere continuato analiticamente in punti dello spettro. Ciò si può vedere dalla (5.38) del seguente teorema.

5.17 Teorema (Risolvente)

Se $T \in B(X, X)$ dove X è uno spazio di Banach complesso e $\lambda_0 \in \rho(T)$ allora

$$\|R_{\lambda_0}(T)\| \geq \frac{1}{\delta(\lambda_0)} \quad \text{dove} \quad \delta(\lambda_0) = \inf_{s \in \sigma(T)} |\lambda_0 - s| \quad (5.37)$$

è la distanza di λ_0 dallo spettro $\sigma(T)$. Quindi

$$\|R_{\lambda_0}(T)\| \rightarrow \infty \quad \text{come} \quad \delta(\lambda_0) \rightarrow 0. \quad (5.38)$$

Dimostrazione. Dal Teorema 5.10 abbiamo che per ogni $\lambda_0 \in \rho(T)$ il disco (5.20) è un sottoinsieme di $\rho(T)$. Quindi il raggio r_0 del massimo disco aperto centrato in λ_0 e contenuto in $\rho(T)$ deve essere

$$r_0 \geq \frac{1}{\|R_{\lambda_0}\|}.$$

La distanza di λ_0 dallo spettro deve essere a sua volta

$$\delta(\lambda_0) \geq r_0$$

e quindi ne segue la (5.37) come enunciato. ■

È di grande importanza teorica e pratica che lo spettro di un operatore lineare limitato T su uno spazio di Banach complesso non possa mai essere un insieme vuoto.

5.18 Teorema (Spettro)

Se $X \neq \{0\}$ è uno spazio di Banach complesso e $T \in B(X, X)$ allora $\sigma(T) \neq \emptyset$. ■

Dimostrazione. Per ipotesi $X \neq \{0\}$. Se $T = 0$ allora $\sigma(T) = \{0\} \neq \emptyset$. Sia $T \neq 0$. Allora $\|T\| \neq 0$. La serie (5.22), Sez. 5.3, è

$$R_\lambda = -\frac{1}{\lambda} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda} T\right)^j \quad (|\lambda| > \|T\|). \quad (5.39)$$

Poiché questa serie converge per $1/|\lambda| < 1/\|T\|$ essa converge assolutamente per $1/|\lambda| < 1/2\|T\|$, ossia per $|\lambda| > 2\|T\|$. Per questi λ per la formula della somma di una serie geometrica

$$\|R_\lambda\| \leq \frac{1}{|\lambda|} \sum_{j=0}^{\infty} \left\| \frac{1}{\lambda} T \right\|^j = \frac{1}{|\lambda| - \|T\|} \leq \frac{1}{\|T\|} \quad (|\lambda| > 2\|T\|). \quad (5.40)$$

Mostriamo che l'assunzione $\sigma(T) = \emptyset$ porta ad una contraddizione. $\sigma(T) = \emptyset$ implica $\rho(T) = \mathbb{C}$. Quindi R_λ è olomorfo per tutti i λ , per il 5.16. Di conseguenza per un $x \in X$ fisso ed un $f \in X'$ fisso la funzione h definita da

$$h(\lambda) = f(R_\lambda x)$$

e sviluppabile in serie di potenze nell'intorno di ogni λ ed è quindi olomorfa su $\mathbb{C}$, ossia h è una funzione intera. Poiché l'olomorfia implica la continuità h è continua e così limitata sul disco compatto $|\lambda| \leq 2\|T\|$. Ma h è anche limitata per $|\lambda| \geq 2\|T\|$ perché $\|R_\lambda\| < 1/\|T\|$ per la (5.40) e

$$\begin{aligned} |h(\lambda)| &= |f(R_\lambda x)| \leq \|f\| \|R_\lambda x\| \leq \|f\| \|R_\lambda\| \|x\| \\ &\leq \|f\| \|x\| / \|T\|. \end{aligned}$$

Quindi h è limitata su $\mathbb{C}$ e perciò costante, per il *teorema di Liouville*, che stabilisce che una funzione intera che è limitata su tutto il piano complesso è una costante. Poiché $x \in X$ ed $f \in X'$ in h erano arbitrari $h = \text{cost}$ implica che R_λ è indipendente da λ e quindi anche $R_\lambda^{-1} = T - \lambda I$. Ma ciò è impossibile ed il teorema è dimostrato. ■

Possiamo infine provare il seguente teorema.

5.19 Teorema (Raggio Spettrale)

Se T è un operatore lineare limitato su uno spazio di Banach complesso allora per il raggio spettrale $r_\sigma(T)$ di T abbiamo

$$r_\sigma(T) = \lim_{n \rightarrow \infty} \|T^n\|^{1/n} = \inf_{k \geq 1} \{\|T^k\|^{1/k}\}. \quad (5.41)$$

■

Dimostrazione. Sia $r = \inf_{k \geq 1} \{\|T^k\|^{1/k}\}$. È evidentemente per ogni $n \geq 1$

$$r \leq \|T^n\|^{1/n}. \quad (5.42)$$

Dato un $\epsilon > 0$ qualunque per definizione di inf esiste un m tale che $\|T^m\|^{1/m} \leq r + \epsilon$. Per un n arbitrario scriviamo $n = pm + q$ dove $0 \leq q \leq m - 1$. Allora, poiché $\|AB\| \leq \|A\| \|B\|$, otteniamo

$$\|T^n\|^{1/n} \leq \|T^m\|^{p/n} \|T\|^{q/n} \leq (r + \epsilon)^{mp/n} \|T\|^{q/n}.$$

Ad m fisso per $n \rightarrow \infty$ è $q/n \rightarrow 0$ e quindi $mp/n \rightarrow 1$. Perciò abbiamo che la limitazione superiore di $\|T^n\|^{1/n}$ tende a $r + \epsilon$ per $n \rightarrow \infty$ e quindi esiste un N tale che per $n > N$ sia

$$\|T^n\|^{1/n} \leq (r + \epsilon) + \epsilon.$$

Da questa disuguaglianza e dalla (5.42) segue che $\lim_{n \rightarrow \infty} \|T^n\|^{1/n}$ esiste ed è uguale ad r . Mostriamo ora che

$$r_\sigma(T) \leq \lim_{n \rightarrow \infty} \|T^n\|^{1/n}. \quad (5.43)$$

Abbiamo $\sigma(T^n) = [\sigma(T)]^n$ per il teorema dell'applicazione spettrale 5.14, così che

$$r_\sigma(T^n) = [r_\sigma(T)]^n.$$

Dalla (5.23) nella Sez. 5.3 applicata a T^n invece che a T vediamo che

$$r_\sigma(T^n) \leq \|T^n\|.$$

Dal confronto

$$r_\sigma(T) = (r_\sigma(T^n))^{1/n} \leq (\|T^n\|)^{1/n}$$

per ogni n . Quindi segue la (5.43).

Mostriamo infine che

$$r_\sigma(T) \geq \lim_{n \rightarrow \infty} \|T^n\|^{1/n}. \quad (5.44)$$

Consideriamo la serie

$$R_\lambda = -\frac{1}{\lambda} \sum_{k=0}^{\infty} \left(\frac{T}{\lambda}\right)^k \quad (5.45)$$

Si può dimostrare in maniera del tutto analoga a quella utilizzata per le serie di potenze di variabile reale che essa converge all'esterno del disco centrato nell'origine di raggio $\max \lim_{n \rightarrow \infty} \|T^n\|^{1/n}$. Quindi poiché $\lim_{n \rightarrow \infty} \|T^n\|^{1/n}$ esiste abbiamo che $\max \lim_{n \rightarrow \infty} \|T^n\|^{1/n} = \lim_{n \rightarrow \infty} \|T^n\|^{1/n} = r$ e la serie converge per $|\lambda| > r$. Sia ora S_n l' n -sima somma parziale della serie (5.45) ed osserviamo che

$$T_\lambda S_n = S_n T_\lambda = I - \left(\frac{T}{\lambda}\right)^{n+1}$$

Poiché la serie (5.45) è convergente per $|\lambda| > r$ deve essere $\lim_{n \rightarrow \infty} \|\lambda^{-n} T^n\| = 0$ per $|\lambda| > r$ e quindi la somma della serie (5.45) è proprio l'operatore risolvente. Ciò prova la (5.44). Allora la (5.43) e la (5.44) implicano la (5.41). ■

5.6 Problemi

1. Sia $X = C[0, 1]$ e sia $T : X \rightarrow X$ definito da

$$Tx(t) = v(t)x(t)$$

con $v(t) \in X$. Mostrare che T è lineare e limitato e studiarne lo spettro.

Suggerimento: È $\|T\| = \|v\|$. $T_\lambda = v(t) - \lambda$ ammette inverso limitato $R_\lambda = (v(t) - \lambda)^{-1}$ per tutti i λ , che non appartengono all'intervallo chiuso $\mathcal{R}(v)$ e quindi $\sigma(T) = \sigma_p(T) = \mathcal{R}(v)$.

2. Sia $\{\alpha_j\}_{j \in \mathbb{N}}$ un insieme numerabile di reali denso nell'intervallo $[0, 1]$. Si consideri l'operatore $T : \ell^2 \rightarrow \ell^2$ definito da $y = Tx$ dove se $x = (\xi_j)$ risulta $y = (\alpha_j \xi_j)$. Mostrare che T è lineare e limitato e studiarne lo spettro.

Suggerimento: Poiché $\sum |\alpha_j \xi_j|^2 \leq \sum |\xi_j|^2$ abbiamo $\|T\| \leq 1$. Essendo

$$(T - \lambda I)x = ((\alpha_1 - \lambda)\xi_1, (\alpha_2 - \lambda)\xi_2, \dots)$$

l'inverso se esiste è

$$(T - \lambda I)^{-1}x = ((\alpha_1 - \lambda)^{-1}\xi_1, (\alpha_2 - \lambda)^{-1}\xi_2, \dots).$$

Se $\lambda \notin [0, 1]$ l'inverso R_λ esiste ed è definito su tutto ℓ^2 . Infatti, se $\delta(\lambda) = \min_{s \in [0, 1]} |\lambda - s|$ è la distanza di λ dall'intervallo, abbiamo

$$\sum \frac{|\xi_j|^2}{|\alpha_j - \lambda|^2} \leq \frac{1}{\delta(\lambda)^2} \sum |\xi_j|^2.$$

Se $\lambda \in [0, 1]$ allora R_λ sicuramente non esiste per $\lambda = \alpha_j$. Quindi poiché $\sigma(T)$ è chiuso segue che $\sigma(T) = [0, 1]$. Se $\lambda \neq \alpha_j$ allora R_λ certamente esiste per tutti gli $x \in \ell^2$ della forma $(\xi_1, \xi_2, \dots, \xi_n, 0, 0, \dots)$ che costituiscono un insieme denso in ℓ^2 . Quindi $\sigma_p(T) = \{\alpha_j\}_{j \in \mathbb{N}}$ e $\sigma_c(T) = \sigma(T) - \sigma_p(T)$.

3. Si consideri l'operatore $T : \ell^p \rightarrow \ell^p$ definito da $y = Tx$ dove se $x = (\xi_j)$ risulta $y = (\xi_{j+1})$. Mostrare che T è lineare e limitato e studiarne lo spettro.

Suggerimento: È $\|T\| = 1$ e quindi se $|\lambda| > 1$ segue che $\lambda \in \rho(T)$. Se $|\lambda| < 1$ gli autovalori di T , ossia tali che $T_\lambda x = 0$ sono

$$x = (\alpha, \alpha\lambda, \alpha\lambda^2, \dots)$$

con $\alpha \in \mathbb{C}$. Per $|\lambda| < 1$ essi appartengono a ℓ^p e quindi $\sigma(T) = \{\lambda \mid |\lambda| \leq 1\}$. Per $|\lambda| = 1$ invece $x \notin \ell^p$ e quindi l'inverso R_λ esiste. Poiché $\mathcal{R}(T) = X$ l'operatore R_λ per $|\lambda| = 1$ non può essere limitato e quindi $\sigma_p(T)$ è il disco aperto centrato nell'origine e di raggio 1 e $\sigma_c(T)$ è la circonferenza centrata nell'origine e di raggio 1.

Capitolo 6

Operatori Lineari Compatti su Spazi Normati e loro Spettro

Gli operatori lineari compatti sono molto importanti nelle applicazioni. Ad esempio essi giocano un ruolo centrale nella teoria delle equazioni integrali ed in vari problemi di fisica matematica.

La loro teoria è servita come un modello nei primi lavori di analisi funzionale. Le loro proprietà assomigliano da vicino a quelle degli operatori negli spazi finito dimensionali. La teoria spettrale per un operatore compatta può essere trattata in maniera quasi esaustiva nel senso che la famosa teoria di Fredholm sulle equazioni integrali lineari può essere estesa alle equazioni lineari funzionali $Tx - \lambda x = y$ con un parametro complesso λ . Questa teoria generalizzata è chiamata *teoria di Riesz-Schauder*.

6.1 Breve orientamento sul principale contenuto del capitolo

La nozione di *compattezza* di un operatore lineare (Def. 6.3) è stata suggerita dalle equazioni integrali. È stata la proprietà risultata essenziale nella teoria di Fredholm. Discuteremo le importanti proprietà generali degli operatori lineari compatti nelle Sezioni 6.2 e 6.3 e le loro proprietà spettrali nelle Sezioni 6.4 e 6.5. La *teoria di Riesz-Schauder* si basa sulle Sezioni 6.4 e 6.5 ed i risultati ottenuti per le equazioni operatoriali sono presentati nelle Sezioni dalla 6.6 alla 6.8. Sono incluse le applicazioni alle equazioni integrali nella Sezione 6.8.

6.2 Operatori Lineari Compatti su Spazi Normati

Abbiamo già definito cosa si intende per sottoinsieme compatto di uno spazio normato (Cf. Def. 2.14). Conviene ora introdurre il concetto di sottoinsieme relativamente compatto.

6.1 DEFINIZIONE (RELATIVAMENTE COMPATTO)

Sia X uno spazio normato. Un sottoinsieme M di X si dice *relativamente compatto* se la sua chiusura $\overline{M}$ è compatta. ■

È utile nel seguito il seguente lemma.

6.2 Lemma (Criterio di relativamente compatto)

Sia X uno spazio normato. Se per ogni successione di un sottoinsieme M di X esiste una successione estratta che converge in X allora M è relativamente compatto. ■

Dimostrazione. Consideriamo una qualunque successione (x_n) contenuta in $\overline{M}$ e mostriamo che ammette una successione estratta che converge in $\overline{M}$. Per definizione di chiusura esiste una successione (z_n) di M tale che $d(x_n, z_n) < 1/n$. Per l'ipotesi fatta la successione (z_n) ammette una successione estratta (z_{n_k}) che converge ad un certo z in X e quindi in $\overline{M}$. Ora

$$d(x_{n_k}, z) \leq d(x_{n_k}, z_{n_k}) + d(z_{n_k}, z)$$

e per $k \rightarrow \infty$ otteniamo $d(x_{n_k}, z) \rightarrow 0$ e quindi la successione estratta (x_{n_k}) converge in $\overline{M}$. ■

Possiamo ora introdurre gli operatori lineari compatti.

6.3 DEFINIZIONE (OPERATORE LINEARE COMPATTO)

Siano X ed Y spazi normati. Un operatore $T : X \rightarrow Y$ è chiamato un *operatore lineare compatto* (o un *operatore lineare completamente continuo*) se T è lineare e se per ogni sottoinsieme limitato M di X l'immagine di $T(M)$ è relativamente compatta. ■

Molti operatori lineari in analisi sono compatti. Una teoria sistematica degli operatori lineari compatti è emersa dallo studio delle *equazioni integrali* della forma

$$(T - \lambda I)x(s) = y(s) \quad \text{dove} \quad Tx(s) = \int_a^b k(s, t)x(t)dt \quad (6.1)$$

dove $\lambda \in \mathbb{C}$ è un parametro,¹ y ed il *kernel* k sono funzioni date (soggette a certe condizioni), ed x è la funzione incognita. Queste equazioni giocano un ruolo anche nella teoria delle equazioni differenziali ordinarie ed alle derivate parziali. D. Hilbert (1912) scoprì il fatto sorprendente che i risultati essenziali relativi alla risolubilità di (6.1) (“teoria di Fredholm”) non dipende dall’esistenza della rappresentazione integrale di T in (6.1) ma solamente dall’assunzione che T sia un operatore lineare compatto. F. Riesz (1918) nel suo famoso articolo del 1918 pose la teoria di Fredholm in una forma astratta assiomatica.

Il termine “compatto” è suggerito dalla definizione. Il termine più antico “completamente continuo” può essere motivato dal seguente lemma, che mostra che un operatore lineare compatto è continuo, mentre l’inverso in generale non è vero.

6.4 Lemma (Continuità)

Siano X ed Y spazi normati. Allora

(a) Ogni operatore lineare compatto $T : X \rightarrow Y$ è limitato e quindi continuo.

(b) Se $\dim X = \infty$ l’operatore identità $I : X \rightarrow X$ (che è continuo) non è compatto. ■

Dimostrazione. (a) La sfera unitaria $U = \{x \in X \mid \|x\| = 1\}$ è limitata. Poiché T è compatto, $\overline{T(U)}$ è compatto ed è limitato per il 2.15, così che

$$\sup_{\|x\|=1} \|Tx\| < \infty.$$

Quindi T è limitato e 2.31 mostra che è continuo.

(b) Naturalmente la palla unitaria chiusa $M = \{x \in X \mid \|x\| \leq 1\}$ è limitata. Se $\dim X = \infty$, allora 2.18 implica che M non può essere compatto; quindi $I(M) = M = \overline{M}$ non è relativamente compatto. ■

¹Assumiamo che $\lambda \neq 0$. Allora (6.1) è detta *di secondo tipo*. Con $\lambda = 0$ è detta *di primo tipo*. Le due corrispondenti teorie sono molto differenti, per ragioni che non possono essere spiegate in poche parole.

Dalla definizione di compattezza di un insieme (cf. 2.14) otteniamo facilmente un utile criterio per gli operatori.

6.5 Teorema (Criterio di compattezza)

Siano X ed Y spazi normati e $T : X \rightarrow Y$ un operatore lineare. Allora T è compatto se e solo se applica ogni successione limitata (x_n) di X in una successione (Tx_n) di Y che ammette una successione estratta convergente. ■

Dimostrazione. Se T è compatto e (x_n) è limitata, allora la chiusura di (Tx_n) in Y è compatta e la Definizione 2.14 mostra che (Tx_n) contiene una successione estratta convergente.

Viceversa, si assuma che ogni successione limitata (x_n) contenga una successione estratta (x_{n_k}) tale che (Tx_{n_k}) converga in Y . Si consideri un qualunque sottoinsieme limitato $B \subset X$ e sia (y_n) una qualunque successione in $T(B)$. Allora $y_n = Tx_n$ per qualche $x_n \in B$ e (x_n) è limitata poiché $\overline{B}$ è limitato. Per ipotesi (Tx_n) contiene una successione estratta convergente. Quindi $\overline{T(B)}$ è compatto per il lemma 6.2 perché (y_n) in $T(B)$ era arbitraria. Per definizione ciò mostra che T è compatto. ■

Da questo teorema segue in maniera pressoché ovvia che la somma $T_1 + T_2$ di due operatori lineari compatti $T_i : X \rightarrow Y$ è compatta. Analogamente αT_1 , con α un qualunque scalare, è compatto. Quindi abbiamo il seguente risultato.

Gli operatori lineari compatti da X in Y formano uno spazio vettoriale.

Inoltre il Teorema 6.5 implica che nel caso finito dimensionale si verificano alcune semplificazioni.

6.6 Teorema (Dominio o immagine finito dimensionale)

Siano X ed Y spazi normati e $T : X \rightarrow Y$ un operatore lineare. Allora

(a) *Se T è limitato e $\dim T(X) < \infty$, l'operatore T è compatto.*

(b) *Se $\dim X < \infty$, l'operatore T è compatto.* ■

Dimostrazione. (a) Sia (x_n) una qualunque successione limitata in X . Allora la disuguaglianza $\|Tx_n\| \leq \|T\| \|x_n\|$ mostra che l'insieme (Tx_n) è limitato. Quindi $\overline{(Tx_n)}$ è chiuso e limitato. Essendo $\dim T(X) < \infty$ ne deduciamo per il Teorema 2.16 che $\overline{(Tx_n)}$ è compatto e quindi (Tx_n) è relativamente compatto. Poiché (x_n) era una successione limitata arbitraria in X , l'operatore T è compatto grazie a 6.5.

(b) segue dalla (a) per l'osservazione che $\dim X < \infty$ implica la limitatezza di T per il 2.30 e $\dim T(X) \leq \dim X$ per il 2.24(b). ■

Menzioniamo che un operatore $T \in B(X, Y)$ con $\dim T(X) < \infty$ (cf. 6.6(a)) è spesso chiamato un *operatore di rango finito*.

Il seguente teorema stabilisce le condizioni sotto le quali il limite di una successione di operatori lineari compatti è compatto. Il teorema costituisce anche un importante strumento che permette di dimostrare la compattezza di un dato operatore mostrando che esso è il limite uniforme di una successione di operatori lineari compatti.

6.7 Teorema (Successione di operatori lineari compatti)

Sia (T_n) una successione di operatori lineari compatti da uno spazio normato X in uno spazio di Banach Y . Se (T_n) converge uniformemente nel senso degli operatori, cioè, se $\|T_n - T\| \rightarrow 0$ (cf. Sez. 4.10), allora l'operatore limite T è compatto. ■

Dimostrazione. Usando il “metodo della diagonalizzazione” mostriamo che per ogni successione limitata (x_m) in X l'immagine (Tx_m) ammette una successione estratta convergente e poi applichiamo il Teorema 6.5.

Poiché T_1 è compatto, (x_m) ammette una successione estratta $(x_{1,m})$ tale che $(T_1 x_{1,m})$ è di Cauchy. Analogamente, $(x_{1,m})$ ammette una successione estratta $(x_{2,m})$ tale che $(T_2 x_{2,m})$ è di Cauchy. Continuando in questa maniera vediamo che la “successione diagonale” $(y_m) = (x_{m,m})$ è una successione estratta di (x_m) tale che per ogni intero positivo fissato n la successione $(T_n y_m)_{m \in \mathbb{N}}$ è di Cauchy. (x_m) è limitata, ossia $\|x_m\| \leq c$ per tutti gli m . Quindi $\|y_m\| \leq c$ per tutti gli m . Sia $\varepsilon > 0$. Poiché $T_m \rightarrow T$ c'è un $n = p$ tale che $\|T - T_p\| < \varepsilon/3c$. Poiché $(T_p y_m)_{m \in \mathbb{N}}$ è di Cauchy v'è un N tale che

$$\|T_p y_j - T_p y_k\| < \frac{\varepsilon}{3} \quad (j, k > N).$$

Di qui otteniamo per $j, k > N$

$$\begin{aligned} \|T y_j - T y_k\| &\leq \|T y_j - T_p y_j\| + \|T_p y_j - T_p y_k\| + \|T_p y_k - T y_k\| \\ &\leq \|T - T_p\| \|y_j\| + \frac{\varepsilon}{3} + \|T_p - T\| \|y_k\| \\ &< \frac{\varepsilon}{3c} c + \frac{\varepsilon}{3} + \frac{\varepsilon}{3c} c = \varepsilon. \end{aligned}$$

Ciò mostra che $(T y_m)$ è di Cauchy e converge poiché Y è completo. Ricordando che (y_m) è una successione estratta di una successione limitata arbitraria (x_m) vediamo che il Teorema 6.5 implica la compattezza dell'operatore T . ■

Si noti che il Teorema diviene falso se si sostituisce la convergenza uniforme di operatori con la convergenza forte di operatori $\|T_n x - T x\| \rightarrow 0$. Ciò può esser visto considerando la successione $T_n : \ell^2 \rightarrow \ell^2$ definita da $T_n x = (\xi_1, \dots, \xi_n, 0, 0, \dots)$, dove $x = (\xi_j) \in \ell^2$. Poiché T_n è lineare e limitato, T_n è compatto per il 6.6(a). Chiaramente $T_n x \rightarrow x = Ix$, ma I non è compatto perché $\dim \ell^2 = \infty$; cf. 6.4(b).

Un'altra proprietà interessante e basilare degli operatori lineari compatti è che essi trasformano successioni debolmente convergenti in successioni fortemente convergenti.

6.8 Teorema (Convergenza debole)

Siano X ed Y spazi normati e $T : X \rightarrow Y$ un operatore lineare compatto. Supponiamo che (x_n) in X converga debolmente e precisamente sia $x_n \xrightarrow{w} x$. Allora $T x_n$ converge fortemente in Y ed ha limite $y = T x$. ■

Dimostrazione. Scriviamo $y_n = T x_n$ e $y = T x$. Mostriamo dapprima che

$$y_n \xrightarrow{w} y \quad (6.2)$$

e poi che

$$y_n \rightarrow y. \quad (6.3)$$

Sia g un qualsiasi funzionale lineare limitato su Y . Definiamo un funzionale f su X ponendo

$$f(z) = g(Tz) \quad (z \in X).$$

f è lineare. f è limitata perché T è compatto e quindi limitato ed è

$$|f(z)| = |g(Tz)| \leq \|g\| \|Tz\| \leq \|g\| \|T\| \|z\|.$$

Per definizione $x_n \xrightarrow{w} x$ implica $f(x_n) \rightarrow f(x)$ e quindi per definizione $g(T x_n) \rightarrow g(T x)$, cioè $g(y_n) \rightarrow g(y)$. Poiché g era arbitraria ciò prova (6.2).

Proviamo (6.3). Assumiamo per assurdo che (6.3) non valga. Allora (y_n) ammette una successione estratta (y_{n_k}) tale che

$$\|y_{n_k} - y\| \geq \eta \quad (6.4)$$

per qualche $\eta > 0$. Poiché (x_n) converge debolmente, (x_n) è limitata per il 4.27(c) ed a maggior ragione (x_{n_k}) . La compattezza di T ora implica per il 6.5 che (Tx_{n_k}) ammette una successione estratta convergente, diciamo $(\tilde{y}_j)$. Sia $\tilde{y}_j \rightarrow \tilde{y}$. A fortiori $\tilde{y}_j \xrightarrow{w} \tilde{y}$. Quindi $\tilde{y} = y$ per la (6.2) e 4.27(b). Di conseguenza

$$\|\tilde{y}_j - y\| \rightarrow 0.$$

Ma

$$\|\tilde{y}_j - y\| \geq \eta > 0$$

per la (6.4). Dall'assurdo segue la validità della (6.3). ■

6.3 Ulteriori Proprietà degli Operatori Lineari Compatti

In questa sezione proviamo che un operatore lineare compatto su uno spazio normato ha un'immagine separabile ed un operatore duale compatto. Queste proprietà saranno necessarie nello studio dello spettro degli operatori lineari compatti, che inizia nella prossima sezione.

Noi fondiamo le nostre considerazioni su due concetti collegati che sono di interesse generale in connessione con la compattezza degli insiemi.

6.9 DEFINIZIONE (ε -RETE, LIMITATEZZA TOTALE)

Sia B un sottoinsieme di uno spazio metrico X e sia dato $\varepsilon > 0$. Un insieme $M_\varepsilon \subset X$ è detto una ε -rete per B se per ogni punto $z \in B$ v'è un punto di M_ε ad una distanza da z minore di ε . L'insieme B è detto *totalmente limitato* se per ogni $\varepsilon > 0$ esiste una ε -rete finita $M_\varepsilon \subset X$ per B , dove per "finita" si intende che M_ε è un insieme finito (ossia consiste di un numero finito di punti). ■

Di conseguenza la totale limitatezza di B significa che per ogni dato $\varepsilon > 0$ l'insieme B è contenuto nell'unione di un numero finito di palle aperte di raggio ε .

Possiamo comprendere il significato e l'importanza dei concetti appena definiti dal seguente lemma, che gioca inoltre un ruolo chiave nelle dimostrazioni di questa sezione.

6.10 Lemma (Limitatezza totale)

Sia B un sottoinsieme di uno spazio metrico X . Allora

- (a) *Se B è relativamente compatto, B è totalmente limitato.*
- (b) *Se B è totalmente limitato e X è completo, B è relativamente compatto.*
- (c) *Se B è totalmente limitato, per ogni $\varepsilon > 0$ ammette un' ε -rete finita $M_\varepsilon \subset B$.*
- (d) *Se B è totalmente limitato, B è separabile.* ■

Dimostrazione. **(a)** Assumiamo che B sia relativamente compatto e mostriamo che, per ogni assegnato $\varepsilon_0 > 0$, esiste una ε_0 -rete finita per B . Se $B = \emptyset$, allora $\emptyset$ è una ε_0 -rete per B . Se $B \neq \emptyset$ prendiamo un qualunque $x_1 \in B$. Se $d(x_1, z) < \varepsilon_0$ per tutti gli $z \in B$, allora $\{x_1\}$ è una ε_0 -rete per B . Altrimenti sia $x_2 \in B$ tale che $d(x_1, x_2) \geq \varepsilon_0$. Se per tutti gli $z \in B$

$$d(x_j, z) < \varepsilon_0 \quad (j = 1 \text{ o } 2) \quad (6.5)$$

allora $\{x_1, x_2\}$ è una ε_0 -rete per B . Altrimenti continuiamo selezionando un $x_3 \in B$ e così via. Asseriamo che esiste un intero positivo n tale che l'insieme $\{x_1, x_2, \dots, x_n\}$ ottenuto dopo n passi di tal fatta è una ε_0 -rete per B . Infatti, se non ci fosse un tale n , la nostra costruzione fornirebbe una successione (x_j) che soddisfa alla disuguaglianza

$$d(x_j, x_k) \geq \varepsilon_0 \quad \text{per } j \neq k.$$

Naturalmente (x_j) non potrebbe avere una successione estratta che sia di Cauchy. Quindi (x_j) non potrebbe ammettere una successione estratta che converga in X . Ma ciò contraddice la relativa compattezza di B perché (x_j) giace in B per costruzione. Quindi deve esistere una ε_0 -rete finita per B . Poiché $\varepsilon_0 > 0$ era arbitrario concludiamo che B è totalmente limitato.

(b) Sia B totalmente limitato e X completo. Consideriamo una qualunque successione (x_n) in B e mostriamo che essa ammette una successione estratta che converge in X , cosicché B per il Lemma 6.2 risulterà relativamente compatto. Per ipotesi B ammette una ε -rete finita per $\varepsilon = 1$. Quindi B è contenuta nell'unione di un numero finito di palle di raggio 1. Fra queste palle possiamo sceglierne una B_1 che contiene un numero infinito di termini di (x_n) (contando le ripetizioni). Sia $(x_{1,n})$ la successione estratta di (x_n) che giace in B_1 . Analogamente, per ipotesi, B è anche contenuta nell'unione di un numero finito di palle di raggio $\varepsilon = 1/2$. Fra queste palle scegliamo una palla B_2 che contiene una successione estratta $(x_{2,n})$ della successione estratta $(x_{1,n})$. Continuiamo induttivamente scegliendo $\varepsilon = 1/3, 1/4, \dots$ e ponendo $y_n = x_{n,n}$. Allora per ogni dato $\varepsilon > 0$ v'è un N (che dipende da ε) tale che tutti gli y_n con $n > N$ giacciono in una palla di raggio ε . Quindi (y_n) è di Cauchy e converge in X , ossia $y_n \rightarrow y \in X$, perché X è completo.

(c) Il caso $B = \emptyset$ è ovvio. Sia $B \neq \emptyset$. Per ipotesi, dato un $\varepsilon > 0$, esiste una ε_1 -rete finita $M_{\varepsilon_1} \subset X$ per B , dove $\varepsilon_1 = \varepsilon/2$. Quindi B è contenuto nell'unione di un numero finito di palle di raggio ε_1 con elementi di M_{ε_1} come centri. Siano $B_1, \dots, B_n$ le palle che intersecano B e siano $x_1, \dots, x_n$ i loro centri. Selezioniamo un punto $z_j \in B \cap B_j$. Vedi Figura 6.1. Allora $M_\varepsilon = \{z_1, \dots, z_n\} \subset B$ è una ε -rete per B , perché per ogni $z \in B$ c'è un B_j contenente z e

$$d(z, z_j) \leq d(z, x_j) + d(x_j, z_j) < \varepsilon_1 + \varepsilon_1 = \varepsilon.$$

(d) Supponiamo che B sia totalmente limitato. Allora per (c) l'insieme B per ogni $\varepsilon_n = 1/n$, $n = 1, 2, \dots$ contiene una ε_n -rete finita M_{ε_n} per se stesso. L'unione M di tutte queste reti è numerabile. M è denso in B . Infatti per ogni dato $\varepsilon > 0$ c'è un n tale che $1/n < \varepsilon$; quindi per ogni $z \in B$ c'è un $a \in M_{1/n} \subset M$ tale che $d(z, a) < \varepsilon$. Ciò mostra che B è separabile. ■

La limitatezza totale implica la limitatezza. L'inverso in generale non è vero.

La prima affermazione è pressoché ovvia. La seconda segue dall'osservazione che la palla unitaria chiusa $U = \{x \mid \|x\| \leq 1\} \subset \ell^2$ è limitata ma non totalmente limitata, perché ℓ^2 è infinito dimensionale e completo, così che U non è compatto (cf. 2.18) e quindi non totalmente compatto per il 6.10(b).

Usando questo Lemma possiamo ora provare facilmente il seguente Teorema.

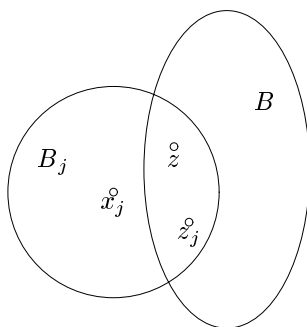


Figura 6.1: Notazioni utilizzate in 6.10

6.11 Teorema (Separabilità dell'immagine)

Siano X e Y spazi normati. L'immagine $\mathcal{R}(T)$ di un operatore lineare compatto $T : X \rightarrow Y$ è separabile. ■

Dimostrazione. Consideriamo la palla $B_n = B(0; n) \subset X$. Poiché T è compatto, l'immagine $C_n = T(B_n)$ è relativamente compatta. C_n è separabile per il Lemma 6.10. La norma di un qualunque $x \in X$ è finita, cosicché $\|x\| < n$ e quindi $x \in B_n$ per n sufficientemente grande. Conseguentemente

$$(a) \quad X = \bigcup_{n=1}^{\infty} B_n, \quad (b) \quad T(X) = \bigcup_{n=1}^{\infty} T(B_n) = \bigcup_{n=1}^{\infty} C_n. \quad (6.6)$$

Poiché C_n è separabile, esso ammette un sottoinsieme D_n numerabile e denso e l'unione

$$D = \bigcup_{n=1}^{\infty} D_n$$

è numerabile. (6.6(b)) mostra che D è denso nell'immagine $\mathcal{R}(T) = T(X)$. ■

Nel prossimo teorema mostriamo che un operatore lineare compatto su uno spazio normato X può essere esteso al completamento di X , rimanendo un operatore lineare compatto.

6.12 Teorema (Estensione compatta)

Un operatore lineare compatto $T : X \rightarrow Y$ da uno spazio normato X in uno spazio di Banach Y ammette un'estensione lineare compatta $\hat{T} : \hat{X} \rightarrow Y$, dove $\hat{X}$ è il completamento di X . (Cf 2.8)

Dimostrazione. Possiamo considerare X come un sottospazio di $\hat{X}$; cf. Teorema 2.8. Poiché T è limitato (cf. 6.4), esso ammette un'estensione lineare limitata $\hat{T} : \hat{X} \rightarrow Y$; cf. 2.33. Mostriamo che la compattezza di T implica che anche $\hat{T}$ è compatto. A questo scopo consideriamo una successione arbitraria limitata $(\hat{x}_n)$ in $\hat{X}$ e mostriamo che $(\hat{T}\hat{x}_n)$ ammette una successione estratta convergente.

Poiché X è denso in $\hat{X}$ v'è una successione (x_n) in X tale che $\hat{x}_n - x_n \rightarrow 0$. Chiaramente anche (x_n) è limitata. Poiché T è compatto, (Tx_n) ammette una successione estratta convergente (Tx_{n_k}) , sia

$$Tx_{n_k} \rightarrow y_0 \in Y. \quad (6.7)$$

Ora $\hat{x}_n - x_n \rightarrow 0$ implica $\hat{x}_{n_k} - x_{n_k} \rightarrow 0$. Poiché $\hat{T}$ è lineare e limitato, esso è continuo. Otteniamo perciò (cf. 1.14)

$$\hat{T}\hat{x}_{n_k} - Tx_{n_k} = \hat{T}(\hat{x}_{n_k} - x_{n_k}) \rightarrow \hat{T}0 = 0.$$

Per la (6.7) ciò implica che $\hat{T}\hat{x}_{n_k} \rightarrow y_0$. Abbiamo così provato che la successione limitata arbitraria $(\hat{x}_n)$ ammette una successione estratta $(\hat{x}_{n_k})$ tale che $(\hat{T}\hat{x}_{n_k})$ converge. Ciò per il 6.5. prova la compattezza di $\hat{T}$. ■

Mostreremo più avanti in questo capitolo che gli operatori lineari compatti appaiono in equazioni operatoriali di grande importanza pratica e teorica. Una teoria generale sulla risolubilità di queste equazioni deve, infatti, fare un uso essenziale di questi operatori. In questo contesto particolarmente cruciale è il fatto che l'operatore duale di un operatore lineare compatto è esso stesso compatto. Proviamo quindi ciò e nella prossima sezione inizieremo la discussione sullo spettro di questi operatori.

6.13 Teorema (Operatore duale)

Sia $T : X \rightarrow Y$ un operatore lineare. Se T è compatto, lo è anche il suo operatore duale $T' : Y' \rightarrow X'$. Qui X e Y sono spazi normati e X' e Y' sono gli spazi duali di X e Y (cf. Def. 2.41 e Sez. 4.6) ■

Dimostrazione. Consideriamo un qualunque sottoinsieme B di Y' che sia limitato, ossia

$$\|g\| \leq c \quad \text{per ogni } g \in B,$$

e mostriamo che l'immagine $T'(B) \subset X'$ è totalmente limitata, così che $T'(B)$ è relativamente compatto per il 6.10(b), perché X' è completo (cf. 2.42).

Dobbiamo quindi provare che per ogni $\varepsilon > 0$ l'immagine ammette una ε -rete finita. Poiché T è compatto, l'immagine $T(S)$ della sfera unitaria

$$S = \{x \in X \mid \|x\| = 1\}$$

è relativamente compatta. Quindi $T(S)$ è totalmente limitata per il 6.10(a). Dal 6.10(c) segue che esiste una ε_1 -rete finita $M \subset T(S)$ per $T(S)$ con $\varepsilon_1 = \varepsilon/4c$. Ciò significa che S contiene dei punti $x_1, \dots, x_n$ tali che per ogni $x \in S$ vale

$$\|Tx - Tx_j\| \leq \frac{\varepsilon}{4c} \quad \text{per qualche } j. \quad (6.8)$$

Definiamo un operatore lineare $A : Y' \rightarrow \mathbb{C}^n$ (o $\mathbb{R}^n$) con

$$Ag = (g(Tx_1), g(Tx_2), \dots, g(Tx_n)). \quad (6.9)$$

Abbiamo

$$\|Ag\|_0^2 = \sum_{j=1}^n |g(Tx_j)|^2,$$

dove $\|\bullet\|_0$ è la norma in $\mathbb{R}^n$. Poiché g è limitato per ipotesi e T è limitato per il 6.4(a) possiamo scrivere

$$\|Ag\|_0^2 \leq \|g\|^2 \|T\|^2 \sum_{j=1}^n \|x_j\|^2 = \|g\|^2 \|T\|^2 n,$$

$$\begin{array}{ccc}
X & \xrightarrow{T} & Y \\
T'(B) \subset X' & \xleftarrow{T'} & Y' \supset B \\
& & \downarrow A \\
& & \mathbb{C}^n \text{ (o } \mathbb{R}^n)
\end{array}$$

Figura 6.2: Notazioni utilizzate in 6.13

cioè A è limitato e di conseguenza, essendo $\mathbb{R}^n$ di dimensioni finite, per il 6.6 compatto. Poiché B è limitato, $A(B)$ è relativamente compatto. Quindi $A(B)$ è totalmente limitato per il 6.10(a). Per il 6.10(c) esso contiene una ε_2 -rete finita $\{Ag_1, \dots, Ag_m\}$ per se stesso, dove $\varepsilon_2 = \varepsilon/4$. Ciò significa che ogni $g \in B$ soddisfa a

$$\|Ag - Ag_k\|_0 < \frac{\varepsilon}{4} \quad \text{per qualche } k. \quad (6.10)$$

Se mostriamo che $\{T'g_1, \dots, T'g_m\}$ è la desiderata ε -rete per $T'(B)$ la dimostrazione del teorema risulta completata.

Dalla (6.9) e (6.10) vediamo immediatamente che per ogni j e per ogni $g \in B$ v'è un k tale che

$$\begin{aligned}
|g(Tx_j) - g_k(Tx_j)|^2 &\leq \sum_{j=1}^n |g(Tx_j) - g_k(Tx_j)|^2 \\
&= \|A(g - g_k)\|^2 < \left(\frac{\varepsilon}{4}\right)^2.
\end{aligned} \quad (6.11)$$

Sia $x \in S$ arbitrario. Allora v'è un j per cui (6.8) vale. Sia $g \in B$ arbitrario. Allora v'è un k per cui (6.10) vale e (6.11) vale per questo k ed ogni j . Otteniamo così

$$\begin{aligned}
|g(Tx) - g_k(Tx)| &\leq |g(Tx) - g(Tx_j)| + |g(Tx_j) - g_k(Tx_j)| \\
&\quad + |g_k(Tx_j) - g_k(Tx)| \\
&< \|g\| \|Tx - Tx_j\| + \frac{\varepsilon}{4} + \|g_k\| \|Tx_j - Tx\| \\
&\leq c \frac{\varepsilon}{4c} + \frac{\varepsilon}{c} + c \frac{\varepsilon}{4c} < \varepsilon.
\end{aligned}$$

Poiché ciò vale per ogni $\|x\| = 1$ e poiché per definizione $g(Tx) = (T'g)(x)$ abbiamo infine

$$\|T'g - T'g_k\| = \sup_{\|x\|=1} |(T'g - T'g_k)(x)| = \sup_{\|x\|=1} |g(Tx) - g_k(Tx)| < \varepsilon.$$

Ciò mostra che $\{T'g_1, \dots, T'g_m\}$ è una ε -rete per $T'(B)$. Poiché ε era arbitrario, $T'(B)$ è totalmente limitato e quindi relativamente compatto per il 6.10(b). Poiché B era un sottoinsieme limitato qualunque di Y' , ciò prova la compattezza di T' secondo la definizione 6.3. ■

6.4 Proprietà Spettrali degli Operatori Lineari Compatti su Spazi Normati

In questa sezione e nella prossima considereremo le proprietà spettrali di un operatore lineare compatto $T : X \rightarrow X$ su uno spazio normato X . A questo scopo useremo nuovamente l'operatore

$$T_\lambda = T - \lambda I \quad (\lambda \in \mathbb{C}) \quad (6.12)$$

ed i concetti basilari della teoria spettrale introdotti nella sezione 5.2.

La teoria spettrale degli operatori lineari compatti è una generalizzazione relativamente semplice della teoria degli autovalori della matrici finite (Sez. 5.1) ed assomiglia sotto molti aspetti a questo caso finito dimensionale. Ciò si può vedere dal seguente sommario delle sezioni 6.4 e 6.5 che includiamo al fine di aiutare il lettore ad orientarsi nella trattazione seguente. Nel sommario sono anche indicati i numeri dei corrispondenti teoremi.

6.14 Sommario

Un operatore lineare compatto $T : X \rightarrow X$ su uno spazio normato X ammette le seguenti proprietà.

L'insieme degli autovalori di T è numerabile (eventualmente finito od anche vuoto). (Cf. 6.15)

$\lambda = 0$ è il solo possibile punto di accumulazione di questo insieme. (Cf. 6.15)

Ogni valore spettrale $\lambda \neq 0$ è un autovalore. (Cf. 6.24) Se X è infinito dimensionale allora $0 \in \sigma(T)$.

Per $\lambda \neq 0$ la dimensione di ogni autospazio di T è finito. (Cf. 6.17)

Per $\lambda \neq 0$ gli spazi nulli di $T_\lambda, T_\lambda^2, T_\lambda^3, \dots$ sono finito dimensionali (cf. 6.17, 6.18) e le immagini di questi operatori sono chiusi (cf. 6.19, 6.20)

C'è un numero r (che dipende da λ , dove $\lambda \neq 0$) tale che

$$X = \mathcal{N}(T_\lambda^r) \oplus T_\lambda^r(X)$$

(cf. 6.25); inoltre lo spazio nullo soddisfa a

$$\mathcal{N}(T_\lambda^r) = \mathcal{N}(T_\lambda^{r+1}) = \mathcal{N}(T_\lambda^{r+2}) = \dots$$

e l'immagine soddisfa a

$$T_\lambda^r(X) = T_\lambda^{r+1}(X) = T_\lambda^{r+2}(X) = \dots$$

(cf. 6.23); se $r > 0$ sono soddisfatte le inclusioni proprie

$$\mathcal{N}(T_\lambda^0) \subsetneq \mathcal{N}(T_\lambda) \subsetneq \dots \subsetneq \mathcal{N}(T_\lambda^r)$$

e

$$T_\lambda^0(X) \supsetneq T_\lambda(X) \supsetneq \dots \supsetneq T_\lambda^r(X).$$

■

Il nostro primo teorema riguarda gli autovalori. Ci dice che lo spettro puntuale di un operatore lineare compatto non è complicato. Il teorema risulta anche più potente di quanto non risulti a prima vista. Infatti nella prossima sezione mostreremo che ogni valore spettrale $\lambda \neq 0$ che ammette un operatore lineare compatto è necessariamente un autovalore. Ciò mostra che in larga misura lo spettro di un operatore lineare compatto assomiglia allo spettro di un operatore in uno spazio finito dimensionale.

6.15 Teorema (Autovalori)

L'insieme degli autovalori di un operatore lineare compatto $T : X \rightarrow X$ su uno spazio normato X è numerabile (eventualmente finito od anche vuoto) ed il solo possibile punto di accumulazione è $\lambda = 0$. ■

Dimostrazione. Ovviamente è sufficiente mostrare che per ogni reale $k > 0$ l'insieme di tutti i $\lambda \in \sigma_p(T)$ tali che $|\lambda| \geq k$ è finito

Supponiamo il contrario per qualche $k_0 > 0$. Allora v'è una successione (λ_n) di infiniti autovalori distinti tali che $|\lambda_n| \geq k_0$ e per cui esiste un $x_n \neq 0$ tale che $Tx_n = \lambda_n x_n$. L'insieme di tutti gli x_n è linearmente indipendente per il Teorema 5.7. Sia $M_n = \text{span}\{x_1, \dots, x_n\}$. Allora ogni $x \in M_n$ ammette un'unica rappresentazione

$$x = \alpha_1 x_1 + \dots + \alpha_n x_n.$$

Applichiamo $T - \lambda_n I$ a x ed usiamo $Tx_j = \lambda_j x_j$

$$(T - \lambda_n I)x = \alpha_1 (\lambda_1 - \lambda_n) x_1 + \dots + \alpha_{n-1} (\lambda_{n-1} - \lambda_n) x_{n-1}.$$

Vediamo che x_n non compare nel membro a destra. Quindi

$$(T - \lambda_n I)x \in M_{n-1} \quad \text{per tutti gli } x \in M_n. \quad (6.13)$$

Gli M_n , poiché sono finito dimensionali, sono chiusi (cf. 2.11) e, grazie all'indipendenza lineare degli x_j , M_{n-1} è incluso propriamente in M_n . Per il lemma di Riesz 2.17 v'è una successione (y_n) tale che

$$y_n \in M_n, \quad \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2} \text{ per tutti gli } x \in M_{n-1}.$$

Mostriamo che

$$\|Ty_n - Ty_m\| \geq \frac{1}{2}k_0 \quad (6.14)$$

cosicché (Ty_n) non ammette una successione estratta convergente perché $k_0 > 0$. Ciò contraddice la compattezza di T perché (y_n) è limitata.

Sia ad esempio $m < n$. Introduciamo

$$\tilde{x} = \lambda_n y_n - Ty_n + Ty_m. \quad (6.15)$$

Mostriamo che $\tilde{x} \in M_{n-1}$. Poiché $m \leq n-1$ è $y_m \in M_m \subset M_{n-1} = \text{span}\{x_1, \dots, x_{n-1}\}$. Quindi $Ty_m \in M_{n-1}$ perché $Tx_j = \lambda_j x_j$. Per la (6.13)

$$\lambda_n y_n - Ty_n = -(T - \lambda_n I)y_n \in M_{n-1}$$

e quindi $\tilde{x} \in M_{n-1}$. Così anche $x = \lambda_n^{-1} \tilde{x} \in M_{n-1}$ e di conseguenza

$$\|\lambda_n y_n - \tilde{x}\| = |\lambda_n| \|y_n - x\| \geq \frac{1}{2} |\lambda_n| \geq \frac{1}{2} k_0$$

perché $|\lambda_n| \geq k_0$. Da ciò e dalla (6.15) abbiamo la (6.14). Quindi l'ipotesi che esiste un numero infinito di autovalori che soddisfano a $|\lambda_n| \geq k_0$ per qualche $k_0 > 0$ conduce ad una contraddizione ed il teorema è provato. ■

Questo teorema mostra che, se un operatore lineare compatto su uno spazio normato ha infiniti autovalori, essi possono essere ordinati in una successione che converge a zero.

La composizione di un operatore lineare compatto con un operatore lineare limitato fornisce un operatore lineare compatto. Questo fatto è provato nel seguente lemma, che ha numerose interessanti applicazioni ed in particolare verrà utilizzato per provare una proprietà basilare degli operatori compatti (6.18 qui di seguito).

6.16 Lemma (Compattezza del prodotto)

Sia $T : X \rightarrow X$ un operatore lineare compatto ed $S : X \rightarrow X$ un operatore lineare limitato su uno spazio normato X . Allora TS e ST sono compatti. ■

Dimostrazione. Sia $B \subset X$ un qualunque insieme limitato. Poiché S è un operatore limitato, $S(B)$ è un insieme limitato e l'insieme $T(S(B)) = TS(B)$ è relativamente compatto perché T è compatto. Quindi TS è un operatore lineare compatto.

Proviamo che anche ST è compatto. Sia (x_n) una qualunque successione limitata in X . Allora (Tx_n) ammette una successione estratta convergente (Tx_{n_k}) per il 6.5 e (STx_{n_k}) converge per il 1.14. Quindi ST è compatto per il 6.5. ■

In accordo con l'affermazione fatta all'inizio di questo capitolo, che la teoria spettrale degli operatori lineari compatti non si discosta di molto da quella degli operatori lineari sugli spazi finito dimensionali, mostriamo che l'autospazio di un qualunque autovalore non nullo di un operatore lineare compatto è finito dimensionale.

6.17 Teorema (Spazio nullo)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X . Allora per ogni $\lambda \neq 0$ lo spazio nullo $\mathcal{N}(T_\lambda)$ di $T_\lambda = T - \lambda I$ è finito dimensionale. ■

Dimostrazione. Mostriamo che la palla chiusa unitaria M in $\mathcal{N}(T_\lambda)$ è compatta e quindi applichiamo il Teorema 2.18.

Sia (x_n) in M . Allora (x_n) è limitata ($\|x_n\| \leq 1$) e (Tx_n) ammette una successione estratta (Tx_{n_k}) convergente per il 6.5. Ora $x_n \in M \subset \mathcal{N}(T_\lambda)$ implica $T_\lambda x_n = Tx_n - \lambda x_n = 0$, cosicché $x_n = \lambda^{-1}Tx_n$ perché $\lambda \neq 0$. Di conseguenza anche $(x_{n_k}) = (\lambda^{-1}Tx_{n_k})$ converge. Il limite è in M perché M è chiuso. Quindi M è compatto per la Definizione 2.14 giacché la successione (x_n) in M era arbitraria. Ciò prova per il 2.18 che $\dim \mathcal{N}(T_\lambda) < \infty$. ■

6.18 Corollario (Spazi nulli)

Nell'ipotesi del Teorema 6.17

$$\dim \mathcal{N}(T_\lambda^n) < \infty \quad n = 1, 2, \dots \quad (6.16)$$

e

$$\{0\} = \mathcal{N}(T_\lambda^0) \subset \mathcal{N}(T_\lambda) \subset \mathcal{N}(T_\lambda^2) \subset \dots \quad (6.17)$$

Dimostrazione. Poiché T_λ è lineare, esso applica 0 in 0. Quindi $T_\lambda^n x = 0$ implica $T_\lambda^{n+1} x = 0$ e segue la 6.17.

Proviamo la 6.16. Per il teorema binomiale

$$\begin{aligned} T_\lambda^n &= (T - \lambda I)^n = \sum_{k=0}^n \binom{n}{k} T^k (-\lambda)^{n-k} \\ &= (-\lambda)^n I + T \sum_{k=1}^n \binom{n}{k} T^{k-1} (-\lambda)^{n-k}. \end{aligned}$$

Ciò può essere scritto

$$T_\lambda^n = W - \mu I \quad \mu = -(-\lambda)^n,$$

dove $W = TS = ST$ e S indica la somma del membro a destra. T è compatto ed S è limitato perché T è limitato per il 6.4(a). Quindi W è compatto per il Lemma 6.16, cosicché otteniamo la (6.16) applicando il Teorema 6.17 a $W - \mu I$. ■

Considereremo ora le immagini di $T_\lambda, T_\lambda^2, \dots$ per un operatore lineare compatto T e $\lambda \neq 0$ qualunque. A questo riguardo dovremmo prima ricordare che per un operatore lineare limitato lo spazio nullo è sempre chiuso ma l'immagine non necessariamente è chiusa (cf. 2.32(b)). Tuttavia, se T è compatto, allora T_λ per ogni $\lambda \neq 0$ ha un'immagine chiusa e lo stesso vale per $T_\lambda, T_\lambda^2, \dots$. Proviamolo dapprima per T_λ . L'estensione del risultato a T_λ^n per ogni $n \in \mathbb{N}$ sarà immediata.

6.19 Teorema (Immagine)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X . Allora per ogni $\lambda \neq 0$ l'immagine di $T_\lambda = T - \lambda I$ è chiusa. ■

Dimostrazione. La dimostrazione è indiretta. Assumiamo che l'immagine $T_\lambda(X)$ non sia chiusa e deduciamone una contraddizione procedendo come segue.

- a) Consideriamo un y appartenente alla chiusura di $T_\lambda(X)$ ma non a $T_\lambda(X)$ e costruiamo una successione $(T_\lambda x_n)$ con $x_n \notin \mathcal{N}(T_\lambda)$ che converge ad y . Costruiamo una successione (z_n) contenuta in $\mathcal{N}(T_\lambda)$ tale che $\|x_n - z_n\| < 2\delta_n$, ove $\delta_n > 0$ è la distanza di x_n da $\mathcal{N}(T_\lambda)$.
- b) Mostriamo che $a_n \rightarrow \infty$, dove $a_n = \|x_n - z_n\|$.
- c) Otteniamo una contraddizione considerando la successione (w_n) , dove $w_n = a_n^{-1}(x_n - z_n)$.

I dettagli sono i seguenti

- a) Per l'ipotesi assurda che $T_\lambda(X)$ non sia chiuso, v'è un $y \in \overline{T_\lambda(X)}$, $y \notin T_\lambda(X)$ ed una successione (x_n) in X tale che

$$y_n = T_\lambda x_n \rightarrow y. \quad (6.18)$$

Poiché $T_\lambda(X)$ è uno spazio vettoriale, $0 \in T_\lambda(X)$. Ma $y \notin T_\lambda(X)$ e quindi $y \neq 0$. Ciò implica che per n sufficientemente grande $y_n \neq 0$, ossia $y_n \notin \mathcal{N}(T_\lambda)$. Senza perdita di generalità possiamo supporre che ciò valga per ogni n . Poiché $\mathcal{N}(T_\lambda)$ è chiuso la distanza δ_n di x_n da $\mathcal{N}(T_\lambda)$ è positiva, cioè

$$\delta_n = \inf_{z \in \mathcal{N}(T_\lambda)} \|x_n - z\| > 0.$$

Per definizione di estremo inferiore c'è una successione (z_n) in $\mathcal{N}(T_\lambda)$ tale che

$$a_n = \|x_n - z_n\| < 2\delta_n. \quad (6.19)$$

- b) Mostriamo che

$$a_n = \|x_n - z_n\| \rightarrow \infty \quad (n \rightarrow \infty). \quad (6.20)$$

Supponiamo che la (6.20) non sia vera. Allora $(x_n - z_n)$ ammette una successione estratta limitata. Poiché T è compatto, segue dal 6.5 che $(T(x_n - z_n))$ ammette

una successione estratta convergente. Ora da $T_\lambda = T - \lambda I$ abbiamo, per $\lambda \neq 0$, $I = \lambda^{-1} (T - T_\lambda)$. Usando $T_\lambda z_n = 0$ (si ricordi che $z_n \in \mathcal{N}(T_\lambda)$) si ottiene che

$$\begin{aligned} x_n - z_n &= \frac{1}{\lambda} (T - T_\lambda) (x_n - z_n) \\ &= \frac{1}{\lambda} [T(x_n - z_n) - T_\lambda x_n]. \end{aligned}$$

$(T(x_n - z_n))$ ammette una successione estratta convergente e $(T_\lambda x_n)$ converge per la (6.18); quindi $x_n - z_n$ ammette una successione estratta convergente, sia $x_{n_k} - z_{n_k} \rightarrow v$. Poiché T è compatto, T è continuo e così T_λ . Quindi per il Teorema 1.14

$$T_\lambda (x_{n_k} - z_{n_k}) \rightarrow T_\lambda v.$$

Ora $T_\lambda z_{n_k} = 0$ perché $z_n \in \mathcal{N}(T_\lambda)$, cosicché per la (6.18) abbiamo anche

$$T_\lambda (x_{n_k} - z_{n_k}) = T_\lambda x_{n_k} \rightarrow y.$$

In conclusione $T_\lambda v = y$, ossia $y \in T_\lambda(X)$ in contraddizione con l'ipotesi $y \notin T_\lambda(X)$ fatta all'inizio del punto **a)** della dimostrazione. Poiché questa contraddizione segue dall'ipotesi fatta all'inizio di questo punto **b)** che la (6.20) non valesse, abbiamo dimostrato la (6.20).

c) Usando ancora a_n così come definito nella (6.19) e ponendo

$$w_n = \frac{1}{a_n} (x_n - z_n) \tag{6.21}$$

abbiamo $\|w_n\| = 1$. Poiché $a_n \rightarrow \infty$ mentre $T_\lambda z_n = 0$ e $(T_\lambda x_n)$ converge, ne segue che

$$T_\lambda w_n = \frac{1}{a_n} T_\lambda x_n \rightarrow 0. \tag{6.22}$$

Usando nuovamente $I = \lambda^{-1} (T - T_\lambda)$ otteniamo

$$w_n = \frac{1}{\lambda} (T w_n - T_\lambda w_n). \tag{6.23}$$

Poiché T è compatto e (w_n) è limitata, $(T w_n)$ ammette una successione estratta convergente. Inoltre $(T_\lambda w_n)$ converge per la (6.22) e quindi la (6.23) mostra che (w_n) ammette una successione estratta convergente, ossia

$$w_{n_k} \rightarrow w. \tag{6.24}$$

Confrontando con la (6.22) ne deduciamo che $T_\lambda w = 0$ e quindi $w \in \mathcal{N}(T_\lambda)$. Poiché $z_n \in \mathcal{N}(T_\lambda)$ anche

$$u_n = z_n + a_n w \in \mathcal{N}(T_\lambda).$$

Quindi per la distanza fra x_n e u_n deve valere la disuguaglianza

$$\|x_n - u_n\| \geq \delta_n.$$

Esplicitando u_n ed utilizzando (6.21) e (6.19) otteniamo così

$$\begin{aligned}\delta_n &\leq \|x_n - z_n - a_n w\| \\ &= \|a_n w_n - a_n w\| \\ &= a_n \|w_n - w\| \\ &< 2\delta_n \|w_n - w\|.\end{aligned}$$

Dividendo per $2\delta_n > 0$ abbiamo $\frac{1}{2} < \|w_n - w\|$. Ciò contraddice la (6.24) e quindi il teorema è dimostrato. ■

6.20 Corollario (Immagini)

Nelle ipotesi del Teorema 6.19 l'immagine di T_λ^n è chiusa per ogni $n = 0, 1, 2, \dots$. Inoltre

$$X = T_\lambda^0(X) \supset T_\lambda(X) \supset T_\lambda^2(X) \supset \dots \quad \blacksquare$$

Dimostrazione. La prima affermazione segue dal Teorema 6.19 notando che W nella dimostrazione del 6.18 è compatto. La seconda affermazione segue per induzione. Infatti abbiamo $T_\lambda^0(X) = I(X) = X \supset T_\lambda(X)$ ed applicando T_λ a $T_\lambda^{n-1}(X) \supset T_\lambda^n(X)$ otteniamo $T_\lambda^n(X) \supset T_\lambda^{n+1}(X)$. ■

6.5 Ulteriori Proprietà Spettrali degli Operatori Lineari Compatti

Dalla sezione precedente sappiamo che, dato un operatore lineare compatto T su uno spazio normato X , per ogni $\lambda \neq 0$ gli spazi nulli $\mathcal{N}(T_\lambda^n)$, $n = 1, 2, \dots$, sono finito dimensionali e soddisfano alla relazione di inclusione $\mathcal{N}(T_\lambda^n) \subset \mathcal{N}(T_\lambda^{n+1})$ e le immagini $T_\lambda^n(X)$ sono chiuse e soddisfano a $T_\lambda^n(X) \supset T_\lambda^{n+1}(X)$.

Possiamo dire di più. Precisamente a partire da un medesimo $n = r$ questi spazi nulli e queste immagini sono uguali.

6.21 Lemma (Spazi nulli)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X . Allora, per ogni fissato $\lambda \neq 0$, esiste un intero minimo r tale che per $n \geq r$ gli spazi nulli $\mathcal{N}(T_\lambda^n)$ sono tutti uguali. Inoltre se $r > 0$ valgono le relazioni di inclusione proprie

$$\mathcal{N}(T_\lambda^0) \subsetneq \mathcal{N}(T_\lambda) \subsetneq \dots \subsetneq \mathcal{N}(T_\lambda^{r-1}) \subsetneq \mathcal{N}(T_\lambda^r). \quad \blacksquare$$

Dimostrazione. Scriviamo per semplicità $\mathcal{N}_n = \mathcal{N}(T_\lambda^n)$. La dimostrazione si svolge in due passi successivi.

a) Assumiamo che non esista m per cui $\mathcal{N}_m = \mathcal{N}_{m+1}$ e deduciamone una contraddizione. Come strumento essenziale useremo il Lemma di Riesz 2.17.

b) Mostriamo che $\mathcal{N}_m = \mathcal{N}_{m+1}$ implica $\mathcal{N}_n = \mathcal{N}_{n+1}$ per ogni $n > m$.

I dettagli sono i seguenti.

a) Sappiamo che $\mathcal{N}_m \subset \mathcal{N}_{m+1}$ per il Corollario 6.18. Assumiamo per ipotesi assurda che non esista m per cui $\mathcal{N}_m = \mathcal{N}_{m+1}$. Allora $\mathcal{N}_{n-1}$ è un sottospazio proprio di $\mathcal{N}_n$ per

ogni n . Poiché questi spazi nulli sono chiusi il Lemma di Riesz 2.17 ci garantisce l'esistenza di una successione (y_n) tale che

$$y_n \in \mathcal{N}_n, \quad \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2} \quad \text{per ogni } x \in \mathcal{N}_{n-1}. \quad (6.25)$$

È sufficiente mostrare che dato un m per ogni $n > m$

$$\|Ty_n - Ty_m\| \geq \frac{1}{2} |\lambda| \quad (6.26)$$

per arrivare ad una contraddizione. Infatti, in questo caso (Ty_n) non può ammettere una successione estratta convergente perché $|\lambda| > 0$ e ciò, giacché (y_n) è limitata, è in contraddizione con la compattezza di T .

Possiamo scrivere, utilizzando l'identità banale $T = T_\lambda + \lambda I$,

$$Ty_n - Ty_m = \lambda y_n - \tilde{x} \quad \text{con} \quad \tilde{x} = T_\lambda y_m + \lambda y_m - T_\lambda y_n. \quad (6.27)$$

Sia $m < n$. Mostriamo che $\tilde{x} \in \mathcal{N}_{n-1}$. Poiché $m \leq n-1$ abbiamo manifestamente che $\lambda y_m \in \mathcal{N}_m \subset \mathcal{N}_{n-1}$. Inoltre $y_m \in \mathcal{N}_m$ implica

$$0 = T_\lambda^m y_m = T_\lambda^{m-1} (T_\lambda y_m),$$

cioè $T_\lambda y_m \in \mathcal{N}_{m-1} \subset \mathcal{N}_{n-1}$. Analogamente $y_n \in \mathcal{N}_n$ implica $T_\lambda y_n \in \mathcal{N}_{n-1}$. Utilizzando entrambi i risultati ne deduciamo che $\tilde{x} \in \mathcal{N}_{n-1}$. Anche $x = \lambda^{-1} \tilde{x} \in \mathcal{N}_{n-1}$ cosicché per la (6.25)

$$\|\lambda y_n - \tilde{x}\| = |\lambda| \|y_n - x\| \geq \frac{1}{2} |\lambda|.$$

Da questa e dalla (6.27) abbiamo la (6.26). Quindi la nostra assunzione che non esista m per cui $\mathcal{N}_m = \mathcal{N}_{m+1}$ è falsa e deve esistere un m per cui $\mathcal{N}_m = \mathcal{N}_{m+1}$.

- b)** Proviamo che $\mathcal{N}_m = \mathcal{N}_{m+1}$ implica $\mathcal{N}_n = \mathcal{N}_{n+1}$ per tutti gli $n > m$. Supponiamo che ciò non sia vero. Allora $\mathcal{N}_n$ è un sottospazio proprio di $\mathcal{N}_{n+1}$ per un qualche $n > m$. Consideriamo un $x \in \mathcal{N}_{n+1} - \mathcal{N}_n$. Per definizione

$$T_\lambda^{n+1} x = 0 \quad \text{ma} \quad T_\lambda^n x \neq 0.$$

Poiché $n > m$ abbiamo $n - m > 0$. Poniamo $z = T_\lambda^{n-m} x$. Allora

$$T_\lambda^{m+1} z = T_\lambda^{n+1} x = 0 \quad \text{ma} \quad T_\lambda^m z = T_\lambda^n x \neq 0.$$

Quindi $z \in \mathcal{N}_{m+1}$ e $z \notin \mathcal{N}_m$ in contraddizione con $\mathcal{N}_m = \mathcal{N}_{m+1}$ e quindi il teorema è dimostrato ed r risulta essere il minimo fra gli m per cui $\mathcal{N}_m = \mathcal{N}_{m+1}$. ■

6.22 Lemma (Immagini)

Nelle ipotesi del Lemma 6.21, per ogni fissato $\lambda \neq 0$, esiste un intero minimo q tale che per $n \geq q$ le immagini $T_\lambda^n(X)$ sono tutte uguali. Inoltre se $q > 0$ valgono le relazioni di inclusione proprie

$$T_\lambda^0(X) \supsetneq T_\lambda(X) \supsetneq \dots \supsetneq T_\lambda^{q-1}(X) \supsetneq T_\lambda^q(X). \quad \blacksquare$$

Dimostrazione. La dimostrazione è nuovamente indiretta e procede lungo linee analoghe alla precedente. Scriviamo per semplicità $\mathcal{R}_n = T_\lambda^n(X)$. Assumiamo per ipotesi assurda che non esista s per cui $\mathcal{R}_s = \mathcal{R}_{s+1}$. Allora $\mathcal{R}_n$ è un sottospazio proprio di $\mathcal{R}_{n+1}$ per ogni n . Poiché queste immagini sono chiuse per il Corollario 6.20, il Lemma di Riesz 2.17 ci garantisce l'esistenza di una successione (y_n) tale che

$$y_n \in \mathcal{R}_n, \quad \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2} \quad \text{per ogni } x \in \mathcal{R}_{n+1}. \quad (6.28)$$

Sia $n > m$. Possiamo scrivere, utilizzando l'identità banale $T = T_\lambda + \lambda I$,

$$Ty_m - Ty_n = \lambda y_m + T_\lambda y_m - Ty_n. \quad (6.29)$$

A destra $\lambda y_m \in \mathcal{R}_m$, $y_m \in \mathcal{R}_m$, così che $T_\lambda y_m \in \mathcal{R}_{m+1}$ e poiché $n > m$ anche $Ty_n \in \mathcal{R}_n \subset \mathcal{R}_{m+1}$. Quindi (6.29) è della forma

$$Ty_m - Ty_n = \lambda(y_m - x) \quad x \in \mathcal{R}_{m+1}.$$

Di conseguenza per la (6.28), dato m , per ogni $n > m$

$$\|Ty_m - Ty_n\| = |\lambda| \|y_m - x\| \geq \frac{1}{2} |\lambda| > 0. \quad (6.30)$$

Poiché (y_n) è limitata e T è compatto, (Ty_n) ammette una successione estratta convergente. Ciò contraddice la (6.30) e prova che $\mathcal{R}_s = \mathcal{R}_{s+1}$ per qualche s . Inoltre, poiché $\mathcal{R}_s = \mathcal{R}_{s+1}$ significa che T_λ applica $\mathcal{R}_s$ in se stesso, ripetute applicazioni di T_λ danno $\mathcal{R}_n = \mathcal{R}_{n+1}$ per ogni $n > s$. Quindi la prima parte del lemma risulta dimostrata e q è il minimo fra gli s per cui $\mathcal{R}_s = \mathcal{R}_{s+1}$. Se $q > 0$ la relazione di inclusione propria segue dal Corollario 6.20. ■

Inoltre si dimostra i due numeri interi q ed r introdotti nei due Lemmi precedenti coincidono ottenendo l'importante Teorema che segue.

6.23 Teorema (Spazi nulli ed immagini)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X . Allora, per ogni fissato $\lambda \neq 0$, esiste un intero minimo r tale che

$$\mathcal{N}(T_\lambda^r) = \mathcal{N}(T_\lambda^{r+1}) = \mathcal{N}(T_\lambda^{r+2}) = \dots \quad (6.31)$$

e

$$T_\lambda^r(X) = T_\lambda^{r+1}(X) = T_\lambda^{r+2}(X) = \dots \quad (6.32)$$

Inoltre, se $r > 0$, valgono le relazioni di inclusione proprie

$$\mathcal{N}(T_\lambda^0) \subsetneq \mathcal{N}(T_\lambda) \subsetneq \dots \subsetneq \mathcal{N}(T_\lambda^{r-1}) \subsetneq \mathcal{N}(T_\lambda^r) \quad (6.33)$$

e

$$T_\lambda^0(X) \supsetneq T_\lambda(X) \supsetneq \dots \supsetneq T_\lambda^{r-1}(X) \supsetneq T_\lambda^r(X). \quad (6.34)$$

■

Dimostrazione. Il Lemma 6.21 dà (6.31) e (6.33). Il Lemma 6.22 dà (6.32) e (6.34) con q invece di r . Dobbiamo mostrare solamente che $q = r$. Proviamo nella parte a) che $q \geq r$ e nella parte b) che $q \leq r$. Come sopra scriviamo per semplicità $\mathcal{N}_n = \mathcal{N}(T_\lambda^n)$ e $\mathcal{R}_n = T_\lambda^n(X)$.

a) Abbiamo $\mathcal{R}_{q+1} = \mathcal{R}_q$ per il Lemma 6.26. Ciò significa che $T_\lambda(\mathcal{R}_q) = \mathcal{R}_q$. Quindi

$$y = \mathcal{R}_q \implies y = T_\lambda x \text{ per qualche } x \in \mathcal{R}_q. \quad (6.35)$$

Mostriamo che

$$T_\lambda x = 0, \quad x \in \mathcal{R}_q \implies x = 0. \quad (6.36)$$

Supponiamo che la (6.36) non valga. Allora $T_\lambda x_1 = 0$ per qualche elemento non nullo $x_1 \in \mathcal{R}_q$. Ora la (6.35) con $y = x_1$ dà $x_1 = T_\lambda x_2$ per qualche $x_2 \in \mathcal{R}_q$. Analogamente $x_2 = T_\lambda x_3$ per qualche $x_3 \in \mathcal{R}_q$ e così via. Otteniamo perciò per ogni n attraverso successive successioni

$$0 \neq x_1 = T_\lambda x_2 = \dots = T_\lambda^{n-1} x_n \quad \text{mentre} \quad 0 = T_\lambda x_1 = T_\lambda^n x_n.$$

Quindi $x_n \notin \mathcal{N}_{n-1}$ mentre $x_n \in \mathcal{N}_n$. Per il 6.21 abbiamo che $\mathcal{N}_{n-1} \subset \mathcal{N}_n$ ed il presente risultato mostra che questa inclusione è propria per ogni n , data l'arbitrarietà di n . Ciò contraddice 6.21 e prova (6.36).

Ricordiamo che $\mathcal{R}_q = \mathcal{R}_{q+1}$ per il 6.22 e proviamo che $\mathcal{N}_q = \mathcal{N}_{q+1}$, dimostrando così per il 6.21 che $q \geq r$, giacché r è il più piccolo intero per cui vale l'uguale.

Abbiamo per il 6.18 che $\mathcal{N}_q \subset \mathcal{N}_{q+1}$. Proviamo che $\mathcal{N}_q \supset \mathcal{N}_{q+1}$, cioè proviamo che $T_\lambda^{q+1} x = 0$ implica $T_\lambda^q x = 0$. Supponiamo che ciò sia falso. Allora per qualche x_0

$$y = T_\lambda^q x_0 \neq 0 \quad \text{mentre} \quad T_\lambda y = T_\lambda^{q+1} x_0 = 0.$$

Quindi $y \in \mathcal{R}_q$, $y \neq 0$ e $T_\lambda y = 0$. Ma ciò contraddice (6.36) con $x = y$ e prova $\mathcal{N}_{q+1} \subset \mathcal{N}_q$. Quindi $\mathcal{N}_{q+1} = \mathcal{N}_q$ e $q \geq r$.

b) Proviamo ora che $q \leq r$. Se $q = 0$ ciò è vero. Sia $q \geq 1$. Proviamo che $q \leq r$ mostrando che $\mathcal{N}_{q-1}$ è un sottospazio proprio di $\mathcal{N}_q$. Ciò implica $q \leq r$ perché r , grazie a 6.21, è il più piccolo intero n tale che $\mathcal{N}_n = \mathcal{N}_{n+1}$.

Per definizione di q in 6.22 l'inclusione $\mathcal{R}_q \subset \mathcal{R}_{q-1}$ è propria. Sia $y \in \mathcal{R}_{q-1} - \mathcal{R}_q$. Poiché $y \in \mathcal{R}_{q-1}$ v'è un x tale che $y = T_\lambda^{q-1} x$. Inoltre $T_\lambda y \in \mathcal{R}_q = \mathcal{R}_{q+1}$ implica che $T_\lambda y = T_\lambda^{q+1} z$ per qualche z . Poiché $T_\lambda^q z \in \mathcal{R}_q$ mentre $y \notin \mathcal{R}_q$ abbiamo

$$T_\lambda^{q-1} (x - T_\lambda z) = y - T_\lambda^q z \neq 0.$$

Quindi $x - T_\lambda z \notin \mathcal{N}_{q-1}$. Ma $x - T_\lambda z \in \mathcal{N}_q$ perché

$$T_\lambda^q (x - T_\lambda z) = T_\lambda y - T_\lambda y = 0.$$

Ciò prova che $\mathcal{N}_{q-1} \neq \mathcal{N}_q$, cosicché $\mathcal{N}_{q-1}$ è un sottospazio proprio di $\mathcal{N}_q$. Quindi $q \leq r$ e $q = r$, giacché nella parte a) della dimostrazione avevamo mostrato che $q \geq r$. ■

L'importante proprietà di caratterizzazione dello spettro di un operatore lineare compatto esposta nel teorema seguente si deduce con facilità da questo teorema, quando ci si limiti a considerare l'operatore su uno spazio di Banach. Il caso generale è considerato in 6.31.

6.24 Teorema (Autovalori)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio di Banach X . Allora ogni valore spettrale $\lambda \neq 0$ di T (se esiste²) è un autovalore di T . ■

²Si può mostrare che T può non avere autovalori. Se T è un operatore lineare compatto *auto-aggiunto* su uno spazio di Hilbert complesso $H \neq \{0\}$ allora ammette sempre almeno un autovalore.

Dimostrazione. Se $\mathcal{N}(T_\lambda) \neq \{0\}$ allora λ è un autovalore di T . Supponiamo che $\mathcal{N}(T_\lambda) = \{0\}$ con $\lambda \neq 0$. Allora $T_\lambda x = 0$ implica $x = 0$ e $T_\lambda^{-1} : T_\lambda(X) \rightarrow X$ esiste per il 2.25. Poiché

$$\{0\} = \mathcal{N}(I) = \mathcal{N}(T_\lambda^0) = \mathcal{N}(T_\lambda),$$

abbiamo $r = 0$ per il 6.23. Quindi $X = T_\lambda^0(X) = T_\lambda(X)$ sempre per il 6.23. Ne segue che T_λ è biiettiva, T_λ^{-1} è limitata per il teorema 4.39 dell'inverso limitato giacché X è completo e $\lambda \in \rho(T)$ per definizione. ■

La validità di questo Teorema può essere estesa al caso in cui lo spazio normato X non sia necessariamente di Banach, come vedremo al Teorema 6.31.

Il valore $\lambda = 0$ è stato escluso in molti teoremi di questo capitolo ed è quindi naturale ora chiedersi che cosa si possa dire di $\lambda = 0$ nel caso di un operatore lineare compatto $T : X \rightarrow X$ su uno spazio normato complesso X . Se X è finito dimensionale allora T ammette una rappresentazione mediante matrici ed è chiaro che 0 può appartenere o no a $\sigma(T) = \sigma_p(T)$. Cioè se $\dim X < \infty$ possiamo avere $0 \notin \sigma(T)$ e quindi $0 \in \rho(T)$. Tuttavia se $\dim X = \infty$ possiamo avere $0 \in \sigma(T)$ e tutti e tre i casi

$$0 \in \sigma_p(T), \quad 0 \in \sigma_c(T), \quad 0 \in \sigma_r(T)$$

sono possibili, come si può far vedere portando degli esempi espliciti.

Un'altra interessante ed importante applicazione del Teorema 6.23 ci permette di rappresentare X come la somma diretta di due sottospazi *chiusi*, precisamente lo spazio nullo e l'immagine di T_λ^r .

6.25 Teorema (Somma diretta)

Siano X, T, λ ed r come nel Teorema 6.23. Allora³ X può essere rappresentato nella forma

$$X = \mathcal{N}(T_\lambda^r) \oplus T_\lambda^r(X). \quad (6.37)$$

■

Dimostrazione. Consideriamo un qualunque $x \in X$. Dobbiamo mostrare che x ha una rappresentazione unica della forma

$$x = (x - x_0) + x_0 \quad (x - x_0 \in \mathcal{N}_r, x_0 \in \mathcal{R}_r). \quad (6.38)$$

dove $\mathcal{N}_n = \mathcal{N}_n(T_\lambda^n)$ e $\mathcal{R}_n = T_\lambda^n(X)$, come prima. Si noti che $T_\lambda^r x \in \mathcal{R}_r$. Ora $\mathcal{R}_r = \mathcal{R}_{2r}$ per il Teorema 6.23 e quindi per qualche $x_1 \in X$ deve essere $T_\lambda^r x = T_\lambda^{2r} x_1$. Sia $x_0 = T^r x_1$. Allora $x_0 \in \mathcal{R}_r$ e

$$T_\lambda^r x_0 = T_\lambda^{2r} x_1 = T_\lambda^r x.$$

Ciò mostra che $T_\lambda^r(x - x_0) = 0$. Quindi $x - x_0 \in \mathcal{N}_r$ e ne segue la decomposizione (6.38). Ciò prova la (6.37) purché la (6.38) sia unica.

Mostriamo l'unicità. In aggiunta alla (6.38) sia

$$x = (x - \tilde{x}_0) + \tilde{x}_0 \quad (x - \tilde{x}_0 \in \mathcal{N}_r, \tilde{x}_0 \in \mathcal{R}_r).$$

³Se X è uno spazio vettoriale, allora per ogni sottospazio $Y \subset X$ esiste un sottospazio $Z \subset X$ tale che $X = Y \oplus Z$; cf. la sezione 3.6. Se X è uno spazio normato (in particolare uno spazio di Banach) e $Y \subset X$ è uno sottospazio *chiuso*, allora può non esistere un sottospazio *chiuso* $Z \subset X$ tale che $X = Y \oplus Z$. Se X è uno spazio di Hilbert, allora per ogni sottospazio chiuso Y si ha $X = Y \oplus Z$, dove $Z = Y^\perp$ è chiuso (cf. 3.13 e 3.14). Si noti che gli sottospazi in (6.37) sono chiusi.

Sia $v_0 = x_0 - \tilde{x}_0$. Allora $v_0 \in \mathcal{R}_r$ giacché $\mathcal{R}_r$ è uno spazio vettoriale. Quindi $v_0 = T_\lambda^r v$ per qualche $v \in X$. Inoltre

$$v_0 = x_0 - \tilde{x}_0 = (x - \tilde{x}_0) - (x - x_0)$$

e quindi $v_0 \in \mathcal{N}_r$ e $T_\lambda^r v_0 = 0$. Dunque

$$T_\lambda^{2r} v = T_\lambda^r v_0 = 0$$

e $v \in \mathcal{N}_{2r} = \mathcal{N}_r$ (cf. 6.23). Ciò implica che

$$v_0 = T_\lambda^r v = 0,$$

cioè $v_0 = x_0 - \tilde{x}_0 = 0$ e la rappresentazione (6.38) è unica e la somma $\mathcal{N}_r + \mathcal{R}_r$ è diretta. ■

6.6 Equazioni Operatoriali che coinvolgono Operatori Lineari Compatti

I. Fredholm (1903) ha studiato le equazioni integrali lineari ed in un suo famoso lavoro suggerì una teoria sulla risolubilità di un tipo di equazioni che coinvolgono un operatore lineare compatto. Noi introdurremo lo studente a questa teoria che fu successivamente sviluppata principalmente da F. Riesz (1918) con un contributo importante da parte di J. Schauder (1930).

Considereremo un operatore lineare compatto $T : X \rightarrow X$ su uno spazio normato X , il suo operatore duale $T' : X' \rightarrow X'$ così come definito in 4.13, l'equazione operatoriale

$$Tx - \lambda x = y \quad (y \in X \text{ dato}, \lambda \neq 0), \quad (6.39)$$

l'equazione operatoriale omogenea associata

$$Tx - \lambda x = 0 \quad (\lambda \neq 0), \quad (6.40)$$

e due equazioni simili che coinvolgono l'operatore duale, precisamente l'equazione operatoriale duale

$$T'f - \lambda f = g \quad (g \in X' \text{ dato}, \lambda \neq 0), \quad (6.41)$$

e l'equazione operatoriale duale omogenea associata

$$T'f - \lambda f = 0 \quad (\lambda \neq 0). \quad (6.42)$$

Il parametro $\lambda \in \mathbb{C}$ è arbitrario, non zero, ed è considerato fisso. Studieremo l'esistenza, rispettivamente, delle soluzioni x e f .

Il seguente riassunto dei risultati, che otterremo nel seguito, mostra come queste quattro equazioni siano strettamente interconnesse fra di loro quando se ne studi la risolubilità.

6.26 Sommario

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e $T' : X' \rightarrow X'$ l'operatore duale di T . Sia $\lambda \neq 0$. Allora:

- *L'equazione operatoriale non omogenea (6.39) è normalmente risolubile, cioè ha una soluzione x se e solo se $f(y) = 0$ per tutte le soluzioni f dell'equazione operatoriale duale omogenea associata (6.42). Quindi se $f = 0$ è la sola soluzione di (6.42), allora l'equazione (6.39) è risolubile per ogni y . (Cf. 6.27)*

- L'equazione operatoriale duale non omogenea (6.41) ha una soluzione f se e solo se $g(x) = 0$ per tutte le soluzioni x dell'equazione operatoriale omogenea associata (6.40). Quindi se $x = 0$ è la sola soluzione di (6.40), allora l'equazione (6.41) è risolubile per ogni g (Cf. 6.29)
- L'equazione operatoriale non omogenea (6.39) ha una soluzione x per ogni $y \in X$ se e solo se $x = 0$ è la sola soluzione dell'equazione operatoriale omogenea associata (6.40) (Cf. 6.30a)
- L'equazione operatoriale duale non omogenea (6.41) ha una soluzione f per ogni $g \in X'$ se e solo se $f = 0$ è la sola soluzione dell'equazione operatoriale duale omogenea associata (6.42) (Cf. 6.30b)
- L'equazione operatoriale omogenea (6.40) e la sua duale (6.42) hanno il medesimo numero di soluzioni linearmente indipendenti (Cf. 6.33)
- T_λ soddisfa all'alternativa di Fredholm (Cf. 6.35) ■

Il nostro primo teorema fornisce una condizione necessaria e sufficiente per la risolubilità dell'equazione operatoriale non omogenea (6.39).

6.27 Teorema (Soluzione dell'equazione operatoriale non omogenea)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia $\lambda \neq 0$. Allora l'equazione operatoriale non omogenea (6.39) ha una soluzione x se e solo se y è tale che

$$f(y) = 0 \tag{6.43}$$

per ogni $f \in X'$ che soddisfa all'equazione operatoriale duale omogenea associata (6.42).

Quindi se (6.42) ammette la sola soluzione banale $f = 0$, allora l'equazione (6.39) è risolubile per ogni dato $y \in X$. ■

Dimostrazione. (a) Supponiamo che l'equazione operatoriale (6.39) ammetta una soluzione $x = x_0$, cioè sia

$$y = Tx_0 - \lambda x_0 = T_\lambda x_0.$$

Sia f una qualunque soluzione dell'equazione duale omogenea associata (6.42). Allora abbiamo

$$f(y) = f(Tx_0 - \lambda x_0) = f(Tx_0) - \lambda f(x_0).$$

Ora $f(Tx_0) = (T'f)(x_0)$ per definizione di operatore duale (cf. 4.13). Quindi per (6.42)

$$f(y) = (T'f)(x_0) - \lambda f(x_0) = 0.$$

(b) Viceversa, assumiamo che y in (6.39) soddisfi a (6.43) per ogni soluzione dell'equazione duale omogenea associata (6.42) e mostriamo che allora l'equazione operatoriale (6.39) ha una soluzione.

Supponiamo per assurdo che (6.39) non abbia soluzioni. Allora non esiste x per cui $y = T_\lambda x$. Quindi $y \notin T_\lambda(X)$. Poiché $T_\lambda(X)$ è chiuso per 6.19, la distanza δ da y a $T_\lambda(X)$ è positiva. Per il Lemma 4.21 esiste un $\tilde{f}$ tale che $\tilde{f}(y) = \delta$ e $\tilde{f}(z) = 0$ per ogni $z \in T_\lambda(X)$. Abbiamo per x qualunque

$$0 = \tilde{f}(T_\lambda x) = \tilde{f}(Tx) - \lambda \tilde{f}(x) = (T'\tilde{f})(x) - \lambda \tilde{f}(x).$$

Quindi $\tilde{f}$ è una soluzione dell'equazione duale omogenea associata (6.42). Per assunzione essa soddisfa a (6.43), cioè $\tilde{f}(y) = 0$. Ma ciò contraddice $\tilde{f}(y) = \delta > 0$. Conseguentemente (6.39) deve avere una soluzione. Ciò prova la prima asserzione del teorema, che implica immediatamente la seconda. ■

Questo teorema suggerisce di introdurre il seguente concetto. Sia

$$Ax = y \quad (y \text{ dato}), \quad (6.44)$$

dove $A : X \rightarrow X$ è un operatore lineare limitato su uno spazio normato X . Supponiamo che (6.44) abbia una soluzione $x \in X$ se e solo se y soddisfa a $f(y) = 0$ per ogni soluzione $f \in X'$ dell'equazione

$$A'f = 0, \quad (6.45)$$

dove A' è l'operatore duale di A . Allora l'equazione (6.44) è detta **normalmente risolubile**.

Il Teorema 6.27 mostra che l'equazione (6.39) con un operatore lineare compatto T e $\lambda \neq 0$ è normalmente risolubile.

Per l'equazione (6.41) v'è un analogo del Teorema 6.27, che otterremo come conseguenza del Lemma che segue. Il numero reale positivo c introdotto nel Lemma può dipendere da λ , che è considerato dato. Si noti che la (6.46) del Lemma vale per *qualche* soluzione – chiamata *soluzione di norma minima* – ma non necessariamente per *tutte* le soluzioni. Quindi il Lemma *non* implica l'esistenza di $R_\lambda = T_\lambda^{-1}$.

6.28 Lemma (Soluzione a norma minima)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia dato $\lambda \neq 0$. Allora se per un dato y l'equazione operatoriale non omogenea (6.39) è risolubile, fra tutte le soluzioni ve n'è una $\tilde{x}$ che ha norma minima. Inoltre esiste un numero reale $c > 0$, indipendente da y , tale che per ogni y per cui l'equazione (6.39) è risolubile è

$$\|\tilde{x}(y)\| \leq c\|y\|, \quad (6.46)$$

dove abbiamo scritto $\tilde{x}(y)$ per indicare esplicitamente che la soluzione a norma minima $\tilde{x}$ dipende dal particolare y considerato. ■

Dimostrazione. Suddividiamo la dimostrazione in due parti.

(a) Mostriamo che se (6.39) ammette delle soluzioni, allora l'insieme di queste soluzioni contiene una soluzione di norma minima, che chiamiamo $\tilde{x}$.

(b) Mostriamo che esiste un $c > 0$ che soddisfa alla (6.46) per una soluzione $\tilde{x}$ di norma minima corrispondente ad un qualunque $y = T_\lambda \tilde{x}$ per cui (6.39) ammette soluzione.

I dettagli sono i seguenti.

(a) Sia x_0 una soluzione di (6.39). Se x è una qualunque altra soluzione di (6.39), la differenza $z = x - x_0$ soddisfa a (6.40). Quindi ogni soluzione di (6.39) può esser scritta come

$$x = x_0 + z \quad \text{dove} \quad z \in \mathcal{N}(T_\lambda)$$

e viceversa per ogni $z \in \mathcal{N}(T_\lambda)$ la somma $x_0 + z$ è soluzione della (6.39). Per un x_0 fissato la norma di x dipende da z . Scriviamo

$$p(z) = \|x_0 + z\| \quad \text{e} \quad k = \inf_{z \in \mathcal{N}(T_\lambda)} p(z).$$

Per definizione di estremo inferiore $\mathcal{N}(T_\lambda)$ contiene una successione tale che

$$p(z_n) = \|x_0 + z_n\| \rightarrow k \quad (n \rightarrow \infty). \quad (6.47)$$

Poiché T_λ è compatto, (Tz_n) ammette una successione estratta convergente. Ma $z_n \in \mathcal{N}(T_\lambda)$ significa che $T_\lambda z_n = 0$, ossia che $Tz_n = \lambda z_n$, dove $\lambda \neq 0$. Quindi (z_n) ammette una successione estratta convergente, diciamo

$$z_{n_j} \rightarrow z_0$$

dove $z_0 \in \mathcal{N}(T_\lambda)$, giacché $\mathcal{N}(T_\lambda)$ è chiuso per il 2.32. Ed è anche

$$p(z_{n_j}) \rightarrow p(z_0)$$

giacché p è continuo. Otteniamo così dalla (6.47)

$$p(z_0) = \|x_0 + z_0\| = k.$$

Ciò mostra che se l'equazione (6.39) ammette soluzioni, l'insieme di queste soluzioni contiene una soluzione $\tilde{x} = x_0 + z_0$ di norma minima.

(b) Proviamo che esiste un $c > 0$ (indipendente da y) tale che la (6.46) vale per una soluzione $\tilde{x}$ di norma minima corrispondente ad una qualsiasi $y = T_\lambda \tilde{x}$ per cui l'equazione (6.39) è risolubile.

Supponiamo che ciò non valga, ossia che

$$\sup_{y \in T_\lambda \tilde{x}} \frac{\|\tilde{x}\|}{\|y\|} = \infty.$$

Allora esiste una successione (y_n) tale che

$$\frac{\|\tilde{x}_n\|}{\|y_n\|} \rightarrow \infty \quad (n \rightarrow \infty), \quad (6.48)$$

dove $\tilde{x}_n$ è di norma minima e soddisfa a $T_\lambda \tilde{x}_n = y_n$. Moltiplicando per α si vede che ad αy_n corrisponde αx_n come soluzione di norma minima. Possiamo quindi sempre ridurci a considerare una successione $(\tilde{x}_n)$ per cui sia $\|\tilde{x}_n\| = 1$. Allora (6.48) implica $y_n \rightarrow 0$. Poiché T è compatto e $(\tilde{x}_n)$ è limitata, $(T\tilde{x}_n)$ ammette una successione estratta convergente, diciamo $T\tilde{x}_{n_j} \rightarrow v_0$, ovvero, se scriviamo per convenienza $v_0 = \lambda \tilde{x}_0$,

$$T\tilde{x}_{n_j} \rightarrow \lambda \tilde{x}_0 \quad (j \rightarrow \infty). \quad (6.49)$$

Ora, ricordando che $\lambda \neq 0$, possiamo scrivere

$$x_{n_j} = \frac{1}{\lambda} (T\tilde{x}_{n_j} - T_\lambda \tilde{x}_{n_j})$$

e poiché $y_n = T_\lambda \tilde{x}_n$

$$x_{n_j} = \frac{1}{\lambda} (T\tilde{x}_{n_j} - y_{n_j}).$$

Quindi utilizzando la (6.49) e $y_n \rightarrow 0$ otteniamo

$$\tilde{x}_{n_j} = \frac{1}{\lambda} (T\tilde{x}_{n_j} - y_{n_j}) \rightarrow \tilde{x}_0. \quad (6.50)$$

Da qui poiché T è continuo abbiamo

$$T\tilde{x}_{n_j} \rightarrow T\tilde{x}_0$$

e quindi $T\tilde{x}_0 = \lambda\tilde{x}_0$ per la (6.49). Giacché $T_\lambda\tilde{x}_{n_j} = y_{n_j}$ vediamo che $x_{n_j} = \tilde{x}_{n_j} - \tilde{x}_0$ soddisfa a $T_\lambda x_{n_j} = y_{n_j}$. Poiché $\tilde{x}_{n_j}$ è di norma minima

$$\|x_{n_j}\| = \|\tilde{x}_{n_j} - \tilde{x}_0\| \geq \|\tilde{x}_{n_j}\| = 1.$$

Ma ciò contraddice la convergenza in (6.50). Quindi (6.48) non può esser valida e la successione di quozienti in (6.48) deve essere limitata, ossia dobbiamo avere

$$c = \sup_{y \in T_\lambda \tilde{x}} \frac{\|\tilde{x}\|}{\|y\|} < \infty.$$

dove $\tilde{x}$ è la soluzione di norma minima per l' y considerato. Ciò implica la (6.46). ■

Usando questo Lemma possiamo ora fornire un criterio per la risolubilità dell'equazione (6.41) così come s'è fatto per l'equazione (6.39) nel Teorema 6.27.

6.29 Teorema (Soluzione della equazione operatoriale duale non omogenea)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia $\lambda \neq 0$. Allora l'equazione operatoriale duale non omogenea (6.41) ha una soluzione f se e solo se g è tale che

$$g(x) = 0 \tag{6.51}$$

per tutti gli $x \in X$ che soddisfano alla equazione operatoriale omogenea (6.40).

Quindi se (6.40) ammette solamente la soluzione banale $x = 0$ allora la (6.41) con una qualunque $g \in X'$ considerata assegnata è risolubile. ■

Dimostrazione. (a) Se (6.41) ammette una soluzione f ed x soddisfa a (6.40), allora la (6.51) vale perché

$$g(x) = (T'f)(x) - \lambda f(x) = f(Tx - \lambda x) = f(0) = 0.$$

(b) Viceversa assumiamo che g soddisfi alla (6.51) per ogni soluzione x della (6.40). Mostriamo che allora la (6.41) ammette una soluzione f . Consideriamo un qualunque $y \in T_\lambda(X)$ e sia $x \in X$ tale che $y = T_\lambda x$. Possiamo definire un funzionale f_0 su $T_\lambda(X)$ come segue

$$f_0(y) = f_0(T_\lambda x) = g(x).$$

Questa definizione non è ambigua perché se $T_\lambda x_1 = T_\lambda x_2$, allora $T_\lambda(x_1 - x_2) = 0$, cosicché $x_1 - x_2$ è una soluzione della (6.40). Quindi $g(x_1 - x_2) = 0$ per assunzione, ossia $g(x_1) = g(x_2)$.

f_0 è lineare perché T_λ e g sono lineari. Mostriamo che f_0 è limitata. Il Lemma 6.28 implica che per ogni $y \in T_\lambda(X)$ almeno uno dei corrispondenti x , precisamente quello di norma minima, soddisfa a

$$\|x\| \leq c\|y\| \quad (y = T_\lambda x)$$

dove c non dipende da y . La limitatezza di f_0 può ora esser vista da

$$|f_0(y)| = |g(x)| \leq \|g\| \|x\| \leq c\|g\| \|y\| = \tilde{c}\|y\|,$$

dove $\tilde{c} = c \|g\|$. Per il teorema di Hahn-Banach 4.10 il funzionale f_0 ammette una estensione f su X , che è un funzionale lineare limitato definito su tutto X . Per definizione di f_0

$$f(Tx - \lambda x) = f(T_\lambda x) = f_0(T_\lambda x) = g(x).$$

A sinistra per definizione di operatore duale abbiamo per tutti gli $x \in X$

$$f(Tx - \lambda x) = f(Tx) - \lambda f(x) = (T'f)(x) - \lambda f(x).$$

Unitamente alla formula precedente ciò mostra che f è una soluzione dell'equazione (6.41) e prova la prima asserzione del teorema. La seconda affermazione discende facilmente dalla prima. ■

6.7 Ulteriori Teoremi di Tipo Fredholm

I principali risultati della sezione precedente riguardano i criteri di risolubilità dell'equazione operatoriale non omogenea (6.39) in termini dell'equazione operatoriale duale omogenea associata (6.42) (Teorema 6.27) e dell'equazione operatoriale duale non omogenea (6.41) in termini dell'equazione operatoriale omogenea associata (6.40) (Teorema 6.29). È naturale ricercare relazioni simili fra l'equazione operatoriale (6.39) e la sua omogenea associata (6.40) e fra l'equazione operatoriale duale (6.41) e la sua associata (6.42).

6.30 Teorema (Soluzioni dell'equazione operatoriale non omogenea)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia $\lambda \neq 0$. Allora:

- (a) *L'equazione operatoriale non omogenea (6.39) ha una soluzione x per ogni $y \in X$ se e solo se l'equazione omogenea associata (6.40) ammette la sola soluzione banale $x = 0$. In questo caso la soluzione di (6.39) è unica e T_λ ha un inverso limitato.*
- (b) *L'equazione operatoriale duale non omogenea (6.41) ha una soluzione f per ogni $g \in X'$ se e solo se l'equazione duale omogenea associata (6.42) ammette la sola soluzione banale $f = 0$. In questo caso la soluzione di (6.41) è unica.* ■

Dimostrazione. (a) Proviamo che se per ogni $y \in X$ l'equazione (6.39) è risolubile allora $x = 0$ è la sola soluzione di (6.40).

Se ciò non fosse (6.40) avrebbe una soluzione $x_1 \neq 0$. Poiché (6.39) con un y qualunque è risolubile, $T_\lambda x = x_1$ ha una soluzione $x = x_2$, cioè $T_\lambda x_2 = x_1$. Per la medesima ragione v'è un x_3 tale che $T_\lambda x_3 = x_2$ e così via. Abbiamo quindi per ogni $k = 2, 3, \dots$

$$0 \neq x_1 = T_\lambda x_2 = T_\lambda^2 x_3 = \dots = T_\lambda^{k-1} x_k$$

e

$$0 = T_\lambda x_1 = T_\lambda^k x_k.$$

Quindi $x_k \in \mathcal{N}(T_\lambda^k)$ ma $x_k \notin \mathcal{N}(T_\lambda^{k-1})$. Ciò significa che il sottospazio nullo $\mathcal{N}(T_\lambda^{k-1})$ è un sottospazio proprio di $\mathcal{N}(T_\lambda^k)$ per tutti i k . Ma ciò contraddice il Teorema 6.23. Quindi $x = 0$ deve essere la sola soluzione di (6.40).

Viceversa supponiamo che $x = 0$ sia la sola soluzione di (6.40). Allora per il Teorema 6.29 è risolubile con una g arbitraria. Ora T' è compatto (cf. 6.13), cosicché possiamo applicare la prima parte di questa dimostrazione a T' e concludere che $f = 0$ deve essere

la sola soluzione di (6.42). La risolubilità di (6.39) con y qualunque segue ora dal Teorema 6.27.

L'unicità segue dal fatto che la differenza di due soluzioni della (6.39) è una soluzione della (6.40). Chiaramente tale soluzione unica $x = T_\lambda^{-1}y$ è la soluzione di norma minima e la limitatezza di T_λ^{-1} segue dal Lemma 6.28, giacché

$$\|x\| = \|T_\lambda^{-1}y\| \leq c\|y\|.$$

Si noti che in questo caso per dimostrare la limitatezza di T_λ^{-1} non è necessario che X sia di Banach, come invece si richiedeva nel Teorema 4.39.

(b) è una conseguenza di (a) e del fatto che T' è compatto (cf. 6.13). ■

Il Teorema 6.30(a) può anche essere usato per mostrare che la validità del Teorema precedentemente dimostrato 6.24 può essere esteso dagli spazi di Banach ad un generico spazio normato.

6.31 Teorema (Autovalori)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X . Allora se T ha valori spettrali diversi da zero, essi debbono essere tutti autovalori di T . ■

Dimostrazione. Se il risolvente $R_\lambda = T_\lambda^{-1}$ non esiste, $\lambda \in \sigma_p(T)$ per definizione. Sia $\lambda \neq 0$ ed assumiamo che $R_\lambda = T_\lambda^{-1}$ esista. Allora $T_\lambda x = 0$ implica $x = 0$ per il 2.25. Ciò significa che l'equazione (6.40) ammette solamente la soluzione banale. Ora il Teorema 6.30(a) mostra che la (6.39) è risolubile per ogni y , cioè R_λ è definito su tutto X ed è limitato. Quindi $\lambda \in \rho(T)$. ■

Le equazioni omogenee (6.40) e (6.42) sono anche collegate fra loro. Mostriamo che esse hanno lo stesso numero di soluzioni linearmente indipendenti. Per la dimostrazione di questo fatto è necessario utilizzare l'esistenza di certi insiemi correlati di X e X' , che costituiscono un sistema chiamato *sistema biortogonale*.

6.32 Lemma (Sistema biortogonale)

Dato un insieme $\{f_1, \dots, f_m\}$ linearmente indipendente nello spazio duale X' di uno spazio normato X , vi sono in X elementi $z_1, \dots, z_m$ tali che

$$f_j(z_k) = \delta_{jk} = \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases} \quad (j, k = 1, \dots, m). \quad (6.52)$$

■

Dimostrazione. Poiché l'ordine delle f_j non è importante, è sufficiente provare che esiste un z_m tale che

$$f_m(z_m) = 1, \quad f_j(z_m) = 0 \quad (j = 1, \dots, m-1). \quad (6.53)$$

Per $m = 1$ ciò è vero perché $f_1 \neq 0$ per l'indipendenza lineare, cosicché $f_1(x_0) \neq 0$ per qualche x_0 e $z_1 = \alpha x_0$ con $\alpha = 1/f_1(x_0)$ dà $f_1(z_1) = 1$.

Sia ora $m > 1$ e facciamo l'ipotesi induttiva che il lemma valga per $m-1$, ossia supponiamo che X contenga degli elementi $z_1, \dots, z_{m-1}$ tali che

$$f_k(z_k) = 1, \quad f_n(z_k) = 0, \quad n \neq k \quad (k, n = 1, \dots, m-1). \quad (6.54)$$

Consideriamo l'insieme

$$M = \{x \in X \mid f_1(x) = 0, \dots, f_{m-1}(x) = 0\}$$

e mostriamo che M contiene un $\tilde{z}_m$ tale che $f_m(\tilde{z}_m) = \beta \neq 0$, di modo che la (6.53) è soddisfatta per $z_m = \beta^{-1}\tilde{z}_m$.

Se così non fosse sarebbe $f_m(x) = 0$ per tutti gli $x \in M$. Prendiamo ora un qualunque $x \in X$ e poniamo

$$\tilde{x} = x - \sum_{j=1}^{m-1} f_j(x) z_j. \quad (6.55)$$

Allora per la (6.54), per $k \leq m-1$,

$$f_k(\tilde{x}) = f_k(x) - \sum_{j=1}^{m-1} f_j(x) f_k(z_j) = f_k(x) - f_k(x) = 0.$$

Ciò mostra che $\tilde{x} \in M$ e che per l'ipotesi assurda fatta $f_m(\tilde{x}) = 0$. Abbiamo quindi

$$\begin{aligned} f_m(x) &= f_m\left(\tilde{x} + \sum_{j=1}^{m-1} f_j(x) z_j\right) \\ &= f_m(\tilde{x}) + \sum_{j=1}^{m-1} f_j(x) f_m(z_j) \\ &= \sum_{j=1}^{m-1} \alpha_j f_j(x) \quad (\alpha_j = f_m(z_j)). \end{aligned}$$

Data l'arbitrarietà di x questa è una rappresentazione di f_m come combinazione lineare di $f_1, \dots, f_{m-1}$ in contraddizione con l'indipendenza lineare di $\{f_1, \dots, f_{m-1}\}$. Quindi $f_m(x) = 0$ per tutti gli $x \in M$ è impossibile ed M deve contenere un z_m tale che la (6.53) valga ed il lemma è dimostrato. ■

Usando questo lemma possiamo ora mostrare che $\dim \mathcal{N}(T_\lambda) = \dim \mathcal{N}(T'_\lambda)$, dove $T'_\lambda = (T - \lambda I)' = T' - \lambda I$. Ciò nei termini delle equazioni operatoriali che stiamo considerando può esser formulato sotto forma del seguente teorema.

6.33 Teorema (Spazi nulli di T_λ e T'_λ)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia $\lambda \neq 0$. Allora l'equazione operatoriale omogenea (6.40) e la sua duale (6.42) hanno il medesimo numero di soluzioni linearmente indipendenti. ■

Dimostrazione. T e T' sono compatti (cf. 6.13), così che $\mathcal{N}(T_\lambda)$ e $\mathcal{N}(T'_\lambda)$ sono finito dimensionali per il 6.17 e possiamo scrivere

$$\dim \mathcal{N}(T_\lambda) = n, \quad \dim \mathcal{N}(T'_\lambda) = m.$$

Suddividiamo la dimostrazione in tre parti (a), (b) e (c) dedicate

- (a) al caso $m = n = 0$ ed alla preparazione del caso $m > 0, n > 0$,
- (b) alla dimostrazione che $n < m$ è impossibile,
- (c) alla dimostrazione che $m > n$ è impossibile.

I dettagli sono i seguenti.

- (a) Se $n = 0$ la sola soluzione di (6.40) è $x = 0$. Allora (6.41) è risolubile per ogni g assegnato (cf. 6.29). Per la parte (b) del Teorema 6.30 ciò implica che $f = 0$ è la sola soluzione di (6.42). Quindi $m = 0$. Analogamente da $m = 0$ segue che $n = 0$.

Supponiamo ora $m > 0$ e $n > 0$. Sia $\{x_1, \dots, x_n\}$ una base di $\mathcal{N}(T_\lambda)$. Chiaramente $x_1 \notin Y_1 = \text{span}\{x_2, \dots, x_n\}$. Per il Lemma 4.21 v'è un $\tilde{g}_1 \in X'$ che è zero ovunque su Y_1 e per cui $\tilde{g}_1(x_1) = \delta$, dove $\delta > 0$ è la distanza di x_1 da Y_1 . Quindi $g_1 = \delta^{-1}\tilde{g}_1$ soddisfa a $g_1(x_1) = 1$ e $g_1(x_2) = 0, \dots, g_1(x_n) = 0$. Analogamente, v'è un g_2 tale che $g_2(x_2) = 1$ e $g_2(x_j) = 0$ per $j \neq 2$ e così via. Quindi X' contiene $g_1, \dots, g_n$ tali che

$$g_k(x_j) = \delta_{jk} \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases} \quad (j, k = 1, \dots, n). \quad (6.56)$$

Analogamente, se $\{f_1, \dots, f_m\}$ è una base di $\mathcal{N}(T'_\lambda)$, allora per il Lemma 6.32 vi sono elementi $z_1, \dots, z_m$ di X tali che

$$f_j(z_k) = \delta_{jk} \quad (j, k = 1, \dots, m). \quad (6.57)$$

- (b) Mostriamo che $n < m$ è impossibile. Sia $n < m$ e definiamo $S : X \rightarrow X$ come segue

$$Sx = Tx + \sum_{j=1}^n g_j(x) z_j. \quad (6.58)$$

S è compatto perché l'operatore limitato $g_j(x)z_j$, avendo immagine unidimensionale, è compatto per il Teorema 6.6(a) ed una somma di operatori compatti è un operatore compatto. Proviamo che

$$S_\lambda x_0 = Sx_0 - \lambda x_0 = 0 \implies x_0 = 0. \quad (6.59)$$

Se $S_\lambda x_0 = 0$ abbiamo $f_k(S_\lambda x_0) = f_k(0) = 0$. Quindi per la (6.58) e (6.57) otteniamo per $k = 1, \dots, n$

$$\begin{aligned} 0 = f_k(S_\lambda x_0) &= f_k\left(T_\lambda x_0 + \sum_{j=1}^n g_j(x_0) z_j\right) \\ &= f_k(T_\lambda x_0) + \sum_{j=1}^n g_j(x_0) f_k(z_j) \\ &= (T'_\lambda f_k)(x_0) + g_k(x_0). \end{aligned} \quad (6.60)$$

Poiché $f_k \in \mathcal{N}(T'_\lambda)$, abbiamo $T'_\lambda f_k = 0$. Quindi la (6.60) fornisce

$$g_k(x_0) = 0, \quad k = 1, \dots, n. \quad (6.61)$$

Ciò implica $Sx_0 = Tx_0$ per la (6.58) e quindi $S_\lambda x_0 = T_\lambda x_0$. Giacché abbiamo supposto $S_\lambda x_0 = 0$ è $T_\lambda x_0 = 0$, ossia $x_0 \in \mathcal{N}(T_\lambda)$. Poiché $\{x_1, \dots, x_n\}$ è una base per $\mathcal{N}(T_\lambda)$ deve essere

$$x_0 = \sum_{j=1}^n \alpha_j x_j,$$

dove gli α_j sono degli opportuni scalari. Applicando g_k ed utilizzando (6.61) e (6.56) abbiamo

$$0 = g_k(x_0) = \sum_{j=1}^n \alpha_j g_k(x_j) = \alpha_k \quad (k = 1, \dots, n).$$

Quindi $x_0 = 0$. Ciò prova l'implicazione (6.59). Il Teorema 6.30(a) ora implica che $S_\lambda x = y$ è risolubile per y qualunque. Scegliamo $y = z_{n+1}$. Sia $x = v$ la soluzione corrispondente, cioè sia $S_\lambda v = z_{n+1}$. Come in (6.60) facciamo il seguente calcolo utilizzando (6.57) e (6.58)

$$\begin{aligned} 1 &= f_{n+1}(z_{n+1}) = f_{n+1}(S_\lambda v) \\ &= f_{n+1}\left(T_\lambda v + \sum_{j=1}^n g_j(v) z_j\right) \\ &= (T'_\lambda f_{n+1})(v) + \sum_{j=1}^n g_j(v) f_{n+1}(z_j) \\ &= (T'_\lambda f_{n+1})(v). \end{aligned}$$

Poiché abbiamo assunto $n < m$, abbiamo $n+1 \leq m$ e $f_{n+1} \in \mathcal{N}(T'_\lambda)$. Quindi $T'_\lambda f_{n+1} = 0$. Ciò è in contraddizione con l'equazione precedente e mostra che $n < m$ è impossibile.

(c) Mostriamo che anche $n > m$ è impossibile. Il ragionamento è simile a quello fatto nella parte (b). Sia $n > m$ e definiamo $\tilde{S} : X' \rightarrow X'$ come segue

$$\tilde{S}f = T'f + \sum_{j=1}^m f(z_j) g_j. \quad (6.62)$$

T' è compatto per il 6.13 ed $\tilde{S}$ è compatto perché l'operatore limitato $f(z_j) g_j$, avendo immagine unidimensionale, è compatto per il Teorema 6.6(a). Proviamo ora (in analogia con la (6.59)) che

$$\tilde{S}_\lambda f_0 = \tilde{S}f_0 - \lambda f_0 = 0 \quad \implies \quad f_0 = 0. \quad (6.63)$$

Se $\tilde{S}_\lambda f_0 = 0$, ponendo nella (6.62) $f = f_0$ ed utilizzando la definizione di operatore duale e finalmente la (6.56) otteniamo per ciascun $k = 1, \dots, m$

$$\begin{aligned} 0 &= (\tilde{S}_\lambda f_0)(x_k) = (T'_\lambda f_0)(x_k) + \sum_{j=1}^m f_0(z_j) g_j(x_k) \\ &= f_0(T_\lambda x_k) + f_0(z_k). \end{aligned} \quad (6.64)$$

La nostra assunzione $m < n$ implica che $x_k \in \mathcal{N}(T_\lambda)$ per $k = 1, \dots, m$ (si ricordi che $\{x_1, \dots, x_n\}$ è una base per $\mathcal{N}(T_\lambda)$). Quindi $f_0(T_\lambda x_k) = f_0(0) = 0$, così che la (6.64) fornisce

$$f_0(z_k) = 0 \quad (k = 1, \dots, m). \quad (6.65)$$

Conseguentemente $\tilde{S}f_0 = T'f_0$ per la (6.62). Da ciò e da $\tilde{S}_\lambda f_0 = 0$ segue che $T'_\lambda f_0 = \tilde{S}_\lambda f_0 = 0$. Quindi $f_0 \in \mathcal{N}(T'_\lambda)$. Poiché $\{f_1, \dots, f_m\}$ è una base per $\mathcal{N}(T'_\lambda)$

$$f_0 = \sum_{j=1}^m \beta_j f_j,$$

dove le β_j sono degli opportuni scalari. Utilizzando (6.65) e (6.57) otteniamo per ciascun $k = 1, \dots, m$

$$0 = f_0(z_k) = \sum_{j=1}^m \beta_j f_j(z_k) = \beta_k.$$

Quindi $f_0 = 0$ e l'implicazione (6.63) è provata. Il Teorema 6.30(b) implica ora che $\tilde{S}_\lambda f = g$ è risolubile per g qualunque. Scegliamo $g = g_{m+1}$. Sia $f = h$ la soluzione corrispondente, cioè sia $\tilde{S}_\lambda h = g_{m+1}$. Utilizzando (6.56), (6.62) e nuovamente (6.56) otteniamo

$$\begin{aligned} 1 = g_{m+1}(x_{m+1}) &= (\tilde{S}_\lambda h)(x_{m+1}) \\ &= (T'_\lambda h)(x_{m+1}) + \sum_{j=1}^m h(z_j) g_j(z_{m+1}) \\ &= (T'_\lambda h)(x_{m+1}) \\ &= h(T_\lambda x_{m+1}). \end{aligned}$$

La nostra assunzione $m < n$ implica $m+1 \leq n$ e quindi $x_{m+1} \in \mathcal{N}(T_\lambda)$. Quindi $h(T_\lambda x_{m+1}) = h(0) = 0$. Ciò contraddice l'equazione precedente e mostra che $m < n$ è impossibile. In conclusione, giacché per la parte (b) anche $n < m$ è impossibile, deve essere $n = m$. ■

6.8 Alternativa di Fredholm

I risultati ottenuti nelle due precedenti sezioni suggeriscono di introdurre la seguente definizione.

6.34 DEFINIZIONE (ALTERNATIVA DI FREDHOLM)

Un operatore lineare limitato $A : X \rightarrow X$ su uno spazio normato X è detto soddisfare l'*alternativa di Fredholm* se A è tale che o (I) o (II) è soddisfatto.

(I) Le equazioni non omogenee

$$Ax = y, \quad A'f = g$$

($A' : X' \rightarrow X'$ è l'operatore duale di A) ammettono soluzioni x e f , rispettivamente, per ogni dato $y \in X$ e $g \in X'$ e tali soluzioni sono uniche. Le corrispondenti equazioni omogenee

$$Ax = 0, \quad A'f = 0$$

ammettono solamente le soluzioni banali $x = 0$ e $f = 0$, rispettivamente.

(II) Le equazioni omogenee

$$Ax = 0, \quad A'f = 0$$

hanno il medesimo numero di soluzioni linearmente indipendenti

$$x_1, \dots, x_n \quad \text{e} \quad f_1, \dots, f_n \quad (n \geq 1),$$

rispettivamente. Le equazioni non omogenee

$$Ax = y, \quad A'f = g$$

non sono risolubili per tutti gli y e g . Esse ammettono una soluzione se e solo se y e g sono, rispettivamente, tali che

$$f_k(y) = 0, \quad g(x_k) = 0 \quad (k = 1, \dots, n). \quad \blacksquare$$

Questa definizione può essere usata per esprimere in maniera compatta i risultati delle due ultime sezioni.

6.35 Teorema (Alternativa di Fredholm)

Sia $T : X \rightarrow X$ un operatore lineare compatto su uno spazio normato X e sia $\lambda \neq 0$. Allora $T_\lambda = T - \lambda I$ soddisfa all'alternativa di Fredholm. $\blacksquare$

Questa formulazione in termini di alternative esclusive è particolarmente importante per le applicazioni, perché spesso, invece di mostrare direttamente l'esistenza di una soluzione, è più facile dimostrare che l'equazione omogenea ammette solamente la soluzione banale.

Abbiamo già menzionato (nella sezione 6.6) che la teoria di Riesz degli operatori lineari compatti è nata come naturale sviluppo della teoria di Fredholm delle equazioni integrali di secondo tipo

$$x(s) - \mu \int_a^b k(s, t) x(t) dt = \tilde{y}(s). \quad (6.66)$$

Daremo una breve introduzione dell'applicazione della teoria degli operatori lineari compatti alle equazioni della forma (6.66).

Ponendo $\mu = 1/\lambda$ e $\tilde{y}(s) = -y(s)/\lambda$, con $\lambda \neq 0$, abbiamo

$$Tx - \lambda x = y \quad (\lambda \neq 0), \quad (6.67)$$

dove T è definito da

$$(Tx)(s) = \int_a^b k(s, t) x(t) dt. \quad (6.68)$$

Possiamo quindi enunciare il seguente teorema.

6.36 Teorema (Alternativa di Fredholm per le equazioni integrali)

Se k in (6.66) è tale che $T : X \rightarrow X$ in (6.68) è un operatore lineare compatto su uno spazio normato X , allora l'alternativa di Fredholm vale per T_λ , ossia o l'equazione integrale non omogenea (6.66) ammette una soluzione unica per tutti gli $\tilde{y} \in X$ o la corrispondente equazione omogenea ammette un numero finito di soluzioni indipendenti non banali (cioè soluzioni $x \neq 0$). $\blacksquare$

Supponiamo che T in (6.67) sia compatto (daremo nel seguito alcune condizioni che ne garantiscono la compattezza). Allora se λ appartiene all'insieme risolvente $\rho(T)$ di T , il risolvente $R_\lambda(T) = (T - \lambda I)^{-1}$ esiste, è definito su tutto X , è limitato (cf. 6.30(a)) e fornisce la soluzione unica

$$x = R_\lambda(T)y$$

di (6.67) per ogni $y \in X$. Poiché $R_\lambda(T)$ è lineare $R_\lambda(T)0 = 0$, ciò che implica che l'equazione omogenea $Tx - \lambda x = 0$ ammette la sola soluzione banale $x = 0$. Quindi $\lambda \in \rho(T)$ corrisponde al caso (I) dell'alternativa di Fredholm.

Sia $|\lambda| > \|T\|$. Assumendo che X sia uno spazio di Banach complesso, ne deduciamo che $\lambda \in \rho(T)$ per il Teorema 5.11. Inoltre (5.40) nella sezione 5.3 fornisce

$$R_\lambda(T) = -\lambda^{-1} (I + \lambda^{-1}T + \lambda^{-2}T^2 + \dots). \quad (6.69)$$

Conseguentemente per la soluzione $x = R_\lambda(T)y$ abbiamo la rappresentazione

$$x = -\lambda^{-1} (y + \lambda^{-1}Ty + \lambda^{-2}T^2y + \dots), \quad (6.70)$$

che è chiamata la **serie di Neumann**.

Il caso (II) dell'alternativa di Fredholm si ottiene se si considera (semprecché esista) un $\lambda \in \sigma(T)$ non nullo, dove $\sigma(T)$ è lo spettro di T . Il Teorema 6.36 implica che λ sia un autovalore. La dimensione del corrispondente autospazio per il Teorema 6.17 è finita ed uguale alla dimensione del corrispondente autospazio di T'_λ , per il Teorema 6.33.

In connessione col Teorema 6.36 due spazi di particolare interesse sono

$$X = L^2[a, b] \quad \text{e} \quad X = C[a, b].$$

Per applicare il teorema è necessario fornire delle condizioni per il kernel k nella (6.66), che siano sufficienti per garantire che T sia compatto.

Se $X = L^2[a, b]$ una tale condizione è che k sia in $L^2[J \times J]$, dove $J = [a, b]$. La dimostrazione richiede l'uso della teoria della misura ed esula dallo scopo di queste note.

Nel caso $X = C[a, b]$ la continuità di k implica la compattezza di T .

Otteremo questo risultato utilizzando il classico Teorema 6.37 che segue.

Una successione (x_n) in $C[a, b]$ è detta **equicontinua** se per ogni $\varepsilon > 0$ esiste un $\delta > 0$, che dipende solo da ε , tale che per ogni x_n ed ogni coppia $s_1, s_2 \in [a, b]$ che soddisfa a $|s_1 - s_2| < \delta$ vale

$$|x_n(s_1) - x_n(s_2)| < \varepsilon.$$

Vediamo da questa definizione che ciascun x_n è uniformemente continuo su $[a, b]$ e che δ non dipende da n .

6.37 Teorema (di Ascoli. Successione equicontinua)

Una successione equicontinua limitata (x_n) in $C[a, b]$ ammette una successione estratta che converge (nella norma di $C[a, b]$). ■

Utilizzando questo teorema che non dimostreremo, possiamo dimostrare il seguente teorema che ci fornisce il risultato desiderato nel caso $X = C[a, b]$.

6.38 Teorema (Operatori integrali compatti)

Sia $J = [a, b]$ e supponiamo che k sia continuo su $J \times J$. Allora l'operatore $T : X \rightarrow X$ definito dalla (6.68), dove $X = C[a, b]$ è un operatore lineare compatto. ■

Dimostrazione. T è lineare. La limitatezza di T segue da

$$\|Tx\| = \max_{s \in J} \left| \int_a^b k(s, t) x(t) dt \right| \leq \|x\| \max_{s \in J} \int_a^b |k(s, t)| dt,$$

che è della forma $\|Tx\| \leq \tilde{c}\|x\|$. Sia (x_n) una qualunque successione limitata in X , diciamo sia $\|x_n\| \leq c$ per ogni n . Sia $Tx_n = y_n$. Allora $\|y_n\| \leq \|T\| \|x_n\|$. Quindi (y_n) è anche limitata. Mostriamo che (y_n) è equicontinua. Poiché il kernel k è continuo su $J \times J$ per ipotesi e $J \times J$ è compatto, k è uniformemente continuo su $J \times J$. Quindi, dato un $\varepsilon > 0$, esiste un $\delta > 0$ tale che per tutti i $t \in J$ e tutti gli $s_1, s_2 \in J$ che soddisfano a $|s_1 - s_2| < \delta$ abbiamo

$$|k(s_1, t) - k(s_2, t)| < \frac{\varepsilon}{(b-a)c}.$$

Conseguentemente, per s_1, s_2 come sopra e per n arbitrario otteniamo

$$\begin{aligned} |y_n(s_1) - y_n(s_2)| &= \left| \int_a^b [k(s_1, t) - k(s_2, t)] x_n(t) dt \right| \\ &< (b-a) \frac{\varepsilon}{(b-a)c} c = \varepsilon. \end{aligned}$$

Ciò prova l'equicontinuità della successione (y_n) . Il Teorema di Ascoli implica che (y_n) ammette una successione estratta. Poiché (x_n) era un'arbitraria successione limitata e $y_n = Tx_n$, la compattezza di T segue dal Teorema 6.5. ■