Note del corso di
Fondamenti di Analisi Matematica

Anno Accademico 2017-18
Corso di Laurea Triennale in Fisica
Università di Roma Tor Vergata

Gerardo Morsella*

18 marzo 2020

*Università di Roma Tor Vergata, Dipartimento di Matematica, viale della Ricerca Scientifica 1, I-00133 Roma, Italy, E-mail: morsella@mat.uniroma2.it

Indice

Introduzione

Il titolo, in realtà un po' generico, di questo corso è dovuto a motivi storici legati all'organizzazione del corso di laurea in Fisica di Tor Vergata, che si è andata modificando lungo gli anni. Questo ha significato in particolare che si facesse meno pressante la necessità di un corso di "dimostrazioni" che gli studenti più interessati agli aspetti matematici potessero affiancare ai corsi di "calcolo" dei primi anni, che attualmente già offrono presentazioni decisamente rigorose. La maggiore libertà che ne è conseguita ha permesso che si potessero affrontare in questo corso argomenti di Analisi più avanzati, in grado di introdurre gli studenti più motivati ai metodi moderni della Fisica Matematica in generale, e della studio rigorso della Meccanica Quantistica in particolare. L'artefice di questo mutamento è stato il prof. John Roberts, che ha tenuto il corso fino all'anno accademico 2011/12. Non mi sono quindi ovviamente azzardato ad andare contro la tradizione.

Lo scopo del corso è pertanto quello di fornire un'introduzione a tutta una serie di strumenti matematici che si sono rivelati di fondamentale importanza nello studio rigoroso dei problemi che nascono dalla Meccanica Quantistica e dalla Teoria Quantistica dei Campi. Tali strumenti sono costituiti essenzialmente dalla teoria degli spazi di Hilbert infinito-dimensionali e degli operatori (limitati e non) su di essi, il che conduce in modo assolutamente naturale allo studio delle C*-algebre, commutative e non. Queste intervengono in primo luogo come la giusta cornice in cui inquadrare la teoria spettrale degli operatori – di ovvio interesse per la Meccanica Quantistica –, ma ci si rende poi presto conto del fatto che il loro ruolo è molto più generale, e consente da una parte una formulazione assiomatica chiara della Meccanica Quantistica stessa, e dall'altra fornisce una grande mole di risultati sulla sua struttura che sarebbe difficile ottenere altrimenti. Questo è ancora più vero se si considera la Teoria Quantistica dei Campi, dove l'impiego di metodi algebrici si è rivelato pressoché imprescindibile nello studio di problemi strutturali, e recentemente ha permesso di ottenere nuovi risultati anche nello studio di modelli concreti. Quindi, sebbene ovviamente questi non rientrino tra i temi trattati nel corso, ho ritenuto che fornire agli studenti una introduzione alle C*-algebre potesse tornar loro utile se in futuro vorranno approfondire la conoscenza di questi campi, che sono attualmente oggetto di ricerca in rapido sviluppo.

Sfortunamente la preparazione matematica di base degli studenti di Fisica delle moderne lauree triennali non contempla i fondamenti dell'Analisi Funzionale, cioè le nozioni di base su spazi normati, metrici e topologici, o la teoria dell'integrazione alla Lebesgue. Pertanto anche questi argomenti, preliminari rispetto a quelli citati sopra, sono oggetto del corso, sebbene limitatamente agli enunciati dei teoremi fondamentali per quanto riguarda la teoria dell'integrazione. Si assume invece come prerequisito la teoria delle funzioni di variabile complessa, che gli studenti di Fisica apprendono all'inizio del terzo anno. Inoltre, per cercare di convincere i miei studenti che tutta la fatica fatta per impadronirsi di questi strumenti non è completamente sprecata, ho incluso nel programma alcune applicazioni alla Meccanica Quantistica: la sua formulazione assiomatica, la caratterizzazione dei sistemi con un numero finito di gradi di libertà in termini dell'algebra di Weyl, e lo studio delle hamiltoniane della particella libera e dell'oscillatore armonico.

Ovviamente esistono molti ottimi testi su tutti questi temi, e alcuni di essi – quelli su cui mi sono maggiormente basato per preparare le lezioni – sono citati nella bibliografia, assolutamente non esaustiva, per non dire decisamente minimale, di queste note. Ma non è facile indicarne uno dove si trovino tutti insieme, e che non contenga al tempo stesso troppe altre cose, rischiando così di risultare dispersivo per uno studente alle prime armi. Il tentativo di fornire ai miei studenti un unico riferimento è stato quindi il motivo principale che mi ha spinto a scrivere queste note, sperando inoltre che risultino loro utili per orientarsi nel grande mare della letteratura, e li incoraggino e incuriosiscano ad approfondirne la conoscenza.

Ho infine inserito in coda ad ogni capitolo un certo numero di esercizi, proposti anche durante le lezioni. Nella maggior parte dei casi la loro soluzione, spesso non difficile, fornisce la dimostrazione di risultati ausiliari a quelli dimostrati nelle note, o ad essi complementari. Quelli più impegnativi sono segnalati da un asterisco (*). Gli studenti sono caldamente incoraggiati a cercare di risolverli *tutti*: non si ripeterà mai abbastanza che l'unico modo per imparare la Matematica, e la Scienza in generale, è farla.

Capitolo 1

Spazi normati, metrici e topologici

Come è noto dal corso di Meccanica Quantistica, gli stati di un sistema quantistico possono essere descritti da certe funzioni ψ di $3n$ variabili reali (se n è il numero di particelle del sistema), le *funzioni d'onda*, e l'insieme di tali funzioni forma uno spazio vettoriale (su $\mathbb{C}$), cioè la somma di due funzioni d'onda, o il prodotto di una funzione d'onda per un numero complesso, sono ancora funzioni d'onda. In tale spazio vettoriale, detto *spazio di Hilbert*, si ha una nozione naturale, e fisicamente significativa, di "grandezza", di una funzione, definita come

$$\|\psi\| = \left(\int_{\mathbb{R}^{3n}} |\psi(x)|^2 d^{3n}x \right)^{1/2}. \tag{1.1}$$

Noi studieremo in dettaglio gli spazi di Hilbert nel capitolo 3. Più in generale, in moltissimi problemi di Analisi avanzata si è condotti in modo naturale a considerare spazi vettoriali (reali o complessi) infinito-dimensionali di funzioni (o di oggetti più generali) in cui è definita una nozione di "grandezza" degli elementi che ha proprietà formali analoghe a quelle di cui gode la (1.1).

Sorge quindi la necessità di identificare gli aspetti comuni a tutti questi casi, il che porta a introdurre la nozione astratta di spazio normato, che inizieremo a studiare nella sezione 1.1, insieme con i concetti più generali di spazio metrico e topologico, anch'essi astrazioni di situazioni naturali in molte questioni di Analisi, alcune delle quali, connesse alla teoria spettrale degli operatori su spazi di Hilbert, saranno oggetto di studio per noi nel capitolo 4. Vedremo poi, nella sezione 1.2, come si generalizzano a spazi topologici le nozioni di limite e continuità di successioni e funzioni, e nella sezione 1.3 inizieremo lo studio delle più semplici funzioni tra spazi normati, gli operatori limitati. Infine discuteremo le generalizzazioni di concetti fondamentali quali quello di completezza, sezione 1.4, e di compattezza, sezione 1.5.

1.1 Nozioni di base

In tutto quello che segue, indicheremo con $\mathbb{K}$ uno dei campi $\mathbb{R}$ o $\mathbb{C}$.

Definizione 1.1.1. Sia X un $\mathbb{K}$-spazio vettoriale. Una *norma* su X è un'applicazione $\|\cdot\| : X \to [0, +\infty)$ tale che, per ogni $x, y \in X$ e $\alpha \in \mathbb{K}$:

(i) $\|x\| = 0$ se e solo se $x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$ (*omogeneità della norma*);

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

La coppia $(X, \|\cdot\|)$ è detta uno *spazio normato.*

Con un abuso di linguaggio comune, che adotteremo anche per altre situazioni simili che si presenteranno in seguito, diremo anche semplicemente "X è uno spazio normato", senza specificare quale norma si consideri su di esso, quando non ci sia dubbio su quest'ultima.

Le proprietà (i)-(iii) della definizione, come si vede, dicono come la norma si comporta rispetto alle operazioni che definiscono la struttura di spazio vettoriale di X. Notiamo anche esplicitamente che dalla (ii) di sopra segue chiaramente che $\|0\| = \|0 \cdot x\| = 0$, e dunque il vero contenuto della (i) è che vale il viceversa: se $\|x\| = 0$ allora necessariamente $x = 0$.

Esempi 1.1.2. (a) Considerato $X = \mathbb{K}^n$, è noto che ponendo

$$\|x\|_2 := |x| := \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2}, \qquad x = (x_1, \ldots, x_n) \in \mathbb{K}^n,$$

si ottiene una norma su X, detta *norma euclidea.*

(b) Sempre su $X = \mathbb{K}^n$, è immediato verificare che si ottengono altre norme ponendo

$$\begin{aligned} \|x\|_1 &:= \sum_{k=1}^{n} |x_k|, \\ \|x\|_\infty &:= \max_{k=1,\ldots,n} |x_k|, \end{aligned} \qquad x = (x_1, \ldots, x_n) \in \mathbb{K}^n.$$

(c) Sia T un insieme arbitrario, e si consideri lo spazio

$$X := \ell^\infty(T, \mathbb{K}) := \{f : T \to \mathbb{K} \,:\, f \text{ è limitata su } T\},$$

che è, come si verifica facilmente, un $\mathbb{K}$-spazio vettoriale con le operazioni definite puntualmente:

$$(f + g)(t) := f(t) + g(t), \quad (\alpha f)(t) = \alpha f(t), \qquad f, g \in X,\ \alpha \in \mathbb{K},\ t \in T.$$

Definendo allora

$$\|f\|_\infty := \sup_{t \in T} |f(t)|, \qquad f \in X,$$

si ottiene una norma su X. Infatti: (i) se $\sup_{t\in T} |f(t)| = 0$ allora $f(t) = 0$ per ogni $t \in T$, e dunque $f = 0$ come elemento di X; (ii) dalla definizione di αf si ha $\|\alpha f\|_\infty = \sup_{t\in T} |\alpha f(t)| = |\alpha| \sup_{t\in T} |f(t)| = |\alpha|\|f\|_\infty$; (iii) dalla definizione di $f + g$ e di $\|\cdot\|_\infty$ segue, per ogni $t \in T$, $|(f + g)(t)| = |f(t) + g(t)| \leq |f(t)| + |g(t)| \leq \|f\|_\infty + \|g\|_\infty$, e dunque, prendendo l'estremo superiore su $t \in T$ del primo membro, $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Poiché considereremo con più frequenza il caso complesso, per semplicità porremo $\ell^\infty(T) := \ell^\infty(T, \mathbb{C})$.

(d) Sia $[a,b] \subset \mathbb{R}$ un intervallo compatto, e si consideri lo spazio $X = C([a,b],\mathbb{K})$ delle funzioni continue su $[a,b]$ a valori in $\mathbb{K}$, con le operazioni di spazio vettoriale definite puntualmente come nell'esempio precedente. Per il teorema di Weierstrass si ha allora che X è un sottospazio vettoriale di $\ell^\infty([a,b],\mathbb{K})$, e dunque la norma $\|\cdot\|_\infty$, ristretta a X, fornisce chiaramente una norma su X. Anche in questo caso si porrà $C([a,b]) := C([a,b],\mathbb{C})$.

(e) Sempre su $X = C([a,b],\mathbb{K})$ con le operazioni puntuali, si può definire

$$\|f\|_1 := \int_a^b |f(t)|dt, \qquad f \in X,$$

e si ottiene ancora una norma. L'unica proprietà di verifica non immediata è la (i): sia $\|f\|_1 = 0$ e supponiamo, per assurdo, che esista $t_0 \in [a,b]$ tale che $f(t_0) \neq 0$. Allora, essendo f continua esisterà un $\delta > 0$ tale che $|f(t)| > |f(t_0)|/2$ per ogni $t \in (t_0-\delta, t_0+\delta) \cap [a,b]$. Da questo segue, per la monotonia dell'integrale, e indicando con $|I|$ la lunghezza del generico intervallo $I \subset \mathbb{R}$,

$$\|f\|_1 = \int_a^b |f(t)|dt \geq \int_{(t_0-\delta,t_0+\delta)\cap[a,b]} |f(t)|dt \geq \frac{|f(t_0)|}{2}\big|(t_0-\delta, t_0+\delta)\cap[a,b]\big| > 0,$$

che contraddice l'ipotesi $\|f\|_1 = 0$.

(f) Ancora su $X = C([a,b],\mathbb{K})$ possiamo definire una terza norma come

$$\|f\|_2 := \left(\int_a^b |f(t)|^2 dt\right)^{1/2}, \qquad f \in X.$$

La verifica che $\|\cdot\|_2$ gode della proprietà (i) è del tutto analoga a quella dell'esempio precedente. La verifica di (ii) è immediata. Per verificare (iii), utilizzeremo fatti che dimostreremo (in modo indipendente da quanto fatto qui), nel capitolo 3 sugli spazi di Hilbert. In particolare, limitandosi al caso, per noi più importante, in cui $\mathbb{K} = \mathbb{C}$ (il caso di funzioni reali è del tutto analogo), si vede facilmente che l'applicazione $(f,g) \in X \times X \mapsto \int_a^b \overline{f(t)}g(t)\,dt$ è un prodotto scalare, definizione 3.1.1. Da questo segue, per il teorema 3.1.2, la disuguaglianza di Cauchy-Schwarz:

$$\left|\int_a^b \overline{f(t)}g(t)\,dt\right| \leq \|f\|_2\|g\|_2, \qquad f,g \in X.$$

Date allora $f,g \in X$, e ricordando che per $z \in \mathbb{C}$ si ha $|\mathrm{Re}\, z| \leq |z|$, si ottiene

$$\begin{aligned}\|f+g\|_2^2 &= \int_a^b \big(|f(t)|^2 + |g(t)|^2 + 2\mathrm{Re}\,(\overline{f(t)}g(t))\big)dt \\ &= \|f\|_2^2 + \|g\|_2^2 + 2\mathrm{Re}\int_a^b \overline{f(t)}g(t)\,dt \leq (\|f\|_2 + \|g\|_2)^2.\end{aligned}$$

Questa norma coincide con quella dell'equazione (1.1) con cui abbiamo iniziato. Vedremo però, nell'esempio 1.4.3(d), che $C([a,b],\mathbb{C})$, con questa norma, *non è* uno spazio di Hilbert.

In $\mathbb{R}^n$, la norma euclidea $|\cdot|$ è utilizzata principalmente per calcolare la distanza tra due punti $x, y \in \mathbb{R}^n$, definita come la lunghezza del vettore che li unisce: $d(x,y) := |x-y| = (\sum_{k=1}^n |x_k - y_k|^2)^{1/2}$. Anche in questo caso, come in quello della norma, risulta utile, per trattare casi più complicati, generalizzare questa nozione astraendone le proprietà fondamentali.

Definizione 1.1.3. Sia X un insieme. Una *metrica* su X è un'applicazione $d : X \times X \to [0, +\infty)$ tale che, per ogni $x, y, z \in X$:

(i) $d(x,y) = 0$ se e solo se $x = y$;

(ii) $d(x,y) = d(y,x)$ (*simmetria della metrica*);

(iii) $d(x,y) \leq d(x,z) + d(z,y)$ (*disuguaglianza triangolare*).

La coppia (X, d) è detta uno *spazio metrico.*

Come ci si può aspettare da quanto detto prima della definizione, una classe generale di esempi di spazi metrici si ottiene considerando, per un dato spazio normato $(X, \|\cdot\|)$, la metrica su X definita da

$$d(x,y) := \|x-y\|, \qquad x, y \in X. \tag{1.2}$$

Che la d così definita sia effettivamente una metrica è di facile verifica: (i) $d(x,y) = 0 \Leftrightarrow \|x-y\| = 0 \Leftrightarrow x = y$ per la (i) della definizione 1.1.1; (ii) $d(x,y) = \|x-y\| = \|(-1)(y-x)\| = \|y-x\| = d(y,x)$ per la (ii) della definizione 1.1.1; (iii) $d(x,y) = \|x-y\| = \|(x-z)+(z-y)\| \leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y)$ per la (iii) della definizione 1.1.1. Nel caso particolare $X = \mathbb{R}^n$, la metrica indotta dalla norma euclidea sarà detta *metrica euclidea.* Viceversa, non ogni metrica si ottiene da una norma tramite la (1.2), e anzi a priori l'insieme X su cui è definita non è necessariamente uno spazio vettoriale. Un semplice esempio in questo senso è fornito nell'esercizio 1.1 alla fine del capitolo.

Dato uno spazio metrico (X, d), una nozione di fondamentale importanza è quella di *palla aperta* di centro $x \in X$ e raggio $\delta > 0$, che è il sottoinsieme di X definito da

$$B_\delta(x) := \{y \in X \,:\, d(x,y) < \delta\}.$$

Usando tali insiemi si può dare un senso preciso alla nozione intuitiva di "vicinanza" tra punti di X, o di "intorno" di un punto di X. Come vedremo, però, metriche diverse su uno stesso spazio X possono dare luogo a nozioni di "vicinanza" coincidenti. È allora utile fare un ulteriore passaggio di astrazione, e introdurre la nozione di *topologia.* Nella definizione seguente $\mathscr{P}(X)$ denota l'*insieme delle parti di* X, cioè l'insieme di tutti i suoi sottoinsiemi.

Definizione 1.1.4. Sia (X, d) uno spazio metrico. Un insieme $A \subset X$ è detto *aperto* se è vuoto, o se per ogni $x \in A$ esiste un $\delta > 0$ tale che $B_\delta(x) \subset A$. La collezione $\tau \subset \mathscr{P}(X)$ di tutti gli insiemi aperti è detta la *topologia indotta dalla metrica d.*

La definizione precedente è chiaramente una generalizzazione ad un arbitrario spazio metrico della nozione di sottoinsieme aperto di $\mathbb{R}^n$ rispetto alla metrica euclidea. La topologia di $\mathbb{R}^n$ indotta dalla metrica euclidea sarà detta la *topologia usuale* di $\mathbb{R}^n$. Le proprietà fondamentali degli insiemi aperti di un generico spazio metrico sono date dal seguente risultato.

Proposizione 1.1.5. *Sia (X, d) uno spazio metrico, e τ la topologia indotta da d. Si ha:*

(i) $\emptyset, X \in \tau$;

(ii) *se $A_\alpha \in \tau$ per ogni $\alpha \in I$, I insieme di indici arbitrario, allora $\bigcup_{\alpha\in I} A_\alpha \in \tau$ (unioni arbitrarie di aperti sono aperte);*

(iii) *se $A_1, \ldots, A_n \in \tau$, allora $A_1 \cap \cdots \cap A_n \in \tau$ (intersezioni finite di aperti sono aperte).*

Dimostrazione. (i) $\emptyset \in \tau$ per definizione, e $X \in \tau$ poichè contiene tutte le palle aperte.

(ii) Se tutti gli A_α sono vuoti, anche la loro unione lo è, e quindi è aperta. Altrimenti, se $x \in \bigcup_\alpha A_\alpha$, deve essere $x \in A_\alpha$ per qualche $\alpha \in I$, e dunque esiste una palla aperta $B_\delta(x) \subset A_\alpha \subset \bigcup_\alpha A_\alpha$.

(iii) Se esiste $x \in A_1 \cap \cdots \cap A_n$, devono esistere palle aperte $B_{\delta_j}(x) \subset A_j$ per ogni $j = 1, \ldots, n$. Ma posto allora $\delta := \min_{j=1,\ldots,n} \delta_j$ si avrà $B_\delta(x) \subset B_{\delta_j}(x)$ per ogni j, e dunque $B_\delta(x) \subset A_1 \cap \cdots \cap A_n$. □

Nelle proprietà (i)-(iii) della proposizione precedente non appare mai esplicitamente la metrica d. È allora naturale porre la definizione seguente.

Definizione 1.1.6. Sia X un insieme. Una *topologia* su X è una famiglia τ di sottoinsiemi di X per la quale valgano le proprietà (i)-(iii) della proposizione 1.1.5. Gli elementi $A \in \tau$ sono detti *(insiemi) aperti di* X, e la coppia (X, τ) è detta uno *spazio topologico*.

La nozione di spazio topologico è chiaramente piuttosto astratta, ma vedremo nel capitolo 4 che spazi topologici la cui topologia non è definita da una metrica sorgono in modo naturale da problemi più concreti, che costituiscono il fondamento della teoria spettrale degli operatori su uno spazio di Hilbert, di ovvio interesse per la Fisica. Per questo motivo studieremo ora gli aspetti fondamentali della teoria degli spazi topologici.

Accanto agli insiemi aperti, si definiscono ovviamente quelli *chiusi*: $C \subset X$ è chiuso se $C^c := X \setminus C$ è aperto. Le proprietà fondamentali degli insiemi chiusi, "duali" rispetto a quelle degli insiemi aperti, sono date dall'esercizio 1.4 (la cui soluzione richiede l'esercizio 1.3). Semplici esempi di insiemi aperti e chiusi in uno spazio metrico sono dati nell'esercizio 1.5.

Un'altra nozione fondamentale quando si considerano spazi topologici è quella di *intorno* di un punto.

Definizione 1.1.7. Siano X uno spazio topologico e $x \in X$. Un insieme $U \subset X$ è detto un *intorno di* x se esiste un aperto $A \subset X$ tale che $x \in A \subset U$.

Indicheremo con $\mathscr{I}_x$ la famiglia di tutti gli intorni di $x \in X$. Chiaramente un insieme aperto $A \subset X$ è intorno di ogni suo punto, e viceversa se A è intorno di ogni suo punto è aperto, in quanto è unione di una famiglia $\{A_x : x \in A\}$ di insiemi aperti tali che $x \in A_x \subset A$. Una sottofamiglia $\mathscr{B}_x \subset \mathscr{I}_x$ è detta una *base di intorni* di x se per ogni U intorno di x esiste un $B \in \mathscr{B}_x$ tale che $B \subset U$. Può essere in generale difficile descrivere esplicitamente tutti gli intorni di un punto in una data topologia, ma spesso è facile descrivere una base di intorni. Ad esempio se la topologia è indotta da una metrica, basi di intorni notevoli sono date dall'esercizio 1.6.

La definizione seguente identifica una sottoclasse di spazi topologici con la quale avremo a che fare pressoché esclusivamente.

Definizione 1.1.8. Uno spazio topologico X è detto *di Hausdorff* se per ogni coppia di punti $x, y \in X$ con $x \neq y$, esistono intorni $U \in \mathscr{I}_x$ e $V \in \mathscr{I}_y$ tali che $U \cap V = \emptyset$.

Cioè brevemente: X è di Hausdorff se ogni coppia di punti distinti ha intorni disgiunti.

Proposizione 1.1.9. *Sia (X, d) uno spazio metrico. La topologia indotta da d è di Hausdorff.*

Dimostrazione. Siano $x, y \in X$ distinti. Allora $d(x, y) > 0$, e si può quindi trovare $\delta > 0$ tale che $\delta < \frac{1}{2}d(x, y)$. Da questo segue $B_\delta(x) \cap B_\delta(y) = \emptyset$: se esistesse $z \in B_\delta(x) \cap B_\delta(y)$, si avrebbe, per la disuguaglianza triangolare,

$$d(x, y) \leq d(x, z) + d(z, y) < 2\delta < d(x, y),$$

il che è assurdo. Essendo allora $B_\delta(x) \in \mathscr{I}_x$, $B_\delta(y) \in \mathscr{I}_y$ come conseguenza dell'esercizio 1.5, si ha la tesi. □

Non tutti gli spazi topologici sono di Hausdorff (e quindi non tutte le topologie sono indotte da una metrica): un semplice esempio è dato dall'esercizio 1.7.

Un'ulteriore nozione importante è quella di chiusura di un'insieme.

Definizione 1.1.10. Siano X uno spazio topologico e $B \subset X$ un insieme. La *chiusura di B* è l'insieme $\bar{B}$ definito come l'intersezione di tutti gli insiemi chiusi di X contenenti B.

Chiaramente si ha $B \subset \bar{B}$, e $\bar{B}$, essendo intersezione di chiusi, è chiuso. Dunque $C \subset X$ è chiuso se e solo se $\bar{C} \subset C$. La proposizione seguente fornisce un'utile caratterizzazione della chiusura di un insieme.

Proposizione 1.1.11. *Sia X uno spazio topologico. Dato $B \subset X$, si ha $x \in \bar{B}$ se e solo se per ogni U intorno di x si ha $U \cap B \neq \emptyset$.*

Dimostrazione. Sia $x \in \bar{B}$ e si supponga, per assurdo, che esista un intorno U di x disgiunto da B. Ovviamente si può suppore che U sia aperto, passando, se occorre, ad un suo sottoinsieme. Ma allora U^c è un chiuso che contiene B ma non x, il che è assurdo per la definizione di $\bar{B}$.

Viceversa supponiamo che per ogni intorno U di x si abbia $U \cap B \neq \emptyset$, e sia, sempre per assurdo, $C \supset B$ un chiuso che non contiene x. Ne segue che x appartiene all'aperto $U = C^c$ che è quindi un intorno di x disgiunto da B, contro l'ipotesi. □

Ovviamente la nozione "duale" a quella di chiusura è quella di *interno* di un insieme $B \subset X$, che è l'insieme $\mathring{B}$ definito come l'unione di tutti gli aperti $A \subset B$.

Definizione 1.1.12. Siano X uno spazio topologico e $B \subset X$ un insieme. Un punto $x \in X$ è detto un *punto di accumulazione* per B se per ogni U intorno di x si ha $(U \cap B) \setminus \{x\} \neq \emptyset$.

Detto in breve: x è di accumulazione per B se ogni suo intorno contiene punti di B diversi da x stesso. È chiaro, in base alla proposizione 1.1.11, che un punto di accumulazione per B sta in $\bar{B}$, ma il viceversa non è vero in generale: ad esempio se si considera, in $\mathbb{R}^2$ con la topologia usuale, l'insieme $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} \cup \{(2, 0)\}$, il punto $(2, 0)$ sta in B e quindi in $\bar{B}$, ma non è di accumulazione poichè ad esempio il suo intorno $B_{1/2}((2, 0))$ ha intersezione vuota con $B \setminus \{(2, 0)\}$. I punti di un insieme B (e quindi della sua chiusura) che non sono di accumulazione, vengono detti *punti isolati* di B.

1.2 Limiti e continuità

Uno dei maggiori vantaggi garantiti dall'introduzione di spazi topologici, è quello di consentire una formulazione molto generale della nozione di limite di successioni e di funzioni. Per iniziare a familiarizzarci con questi concetti, cominciamo con il caso più semplice, quello di limite di una successione, cioè di una funzione $x : \mathbb{N} \to X$ (X spazio topologico), che si indica comunemente con il simbolo $(x_j)_{j \in \mathbb{N}}$ o, per brevità, semplicemente con (x_j).

Definizione 1.2.1. Siano X uno spazio topologico e $(x_j)_{j \in \mathbb{N}} \subset X$ una successione. Si dice che $x \in X$ è un limite della successione (x_j), e si scrive

$$x \in \lim_{j \to +\infty} x_j, \quad \text{o anche} \quad x_j \to x,$$

se per ogni intorno U di x esiste $j_U \in \mathbb{N}$ tale che, per ogni $j \geq j_U$, si ha $x_j \in U$.

Si noti che in generale una successione può avere più di un limite. Vedremo però, come caso particolare della proposizione 1.2.7, che se X è uno spazio di Hausdorff una successione ha al più un limite, nel qual caso si scrive $x = \lim_{j \to +\infty} x_j$. In particolare quindi questo sarà vero per successioni in uno spazio metrico (X, d). In questo caso, inoltre, usando il fatto che gli insiemi $B_\delta(x)$, $\delta > 0$, formano una base di intorni di x (esercizio 1.6), si vede facilmente che $x_j \to x$ se e solo se $d(x_j, x) \to 0$ come successione in $\mathbb{R}$: se $x_j \to x$, dato $\delta > 0$ deve esistere $j_\delta = j_{B_\delta(x)} \in \mathbb{N}$ tale che se $j \geq j_\delta$ si ha $x_j \in B_\delta(x)$, e cioè $d(x_j, x) < \delta$, da cui $d(x_j, x) \to 0$; viceversa se $d(x_j, x) \to 0$, dato un intorno U di x esisterà $\delta > 0$ tale che $B_\delta(x) \subset U$ e $j_\delta \in \mathbb{N}$ tale che se $j \geq j_\delta$ allora $d(x_j, x) < \delta$, e basta allora prendere $j_U = j_\delta$. Il risultato seguente caratterizza le chiusure degli insiemi in uno spazio metrico tramite limiti di successioni.

Teorema 1.2.2. *Siano (X, d) uno spazio metrico, e $B, C \subset X$ insiemi. Si ha*

(i) $x \in \bar{B}$ se e solo se esiste una successione $(x_j) \subset B$ tale che $x_j \to x$;

(ii) *C è chiuso se e solo se per ogni successione $(x_j) \subset C$ che sia convergente ad un $x \in X$, si ha $x \in C$.*

Dimostrazione. (i) Sia $x \in \bar{B}$. Poiché $B_{1/j}(x)$ è un intorno di x, per la proposizione 1.1.11 esisterà $x_j \in B_{1/j}(x) \cap B$, da cui $d(x_j, x) < 1/j$, e quindi $x_j \to x$ per l'osservazione precedente. Viceversa se $(x_j) \subset B$ converge a x, dato U intorno di x si ha $x_j \in U \cap B$ per $j \geq j_U$, e dunque $x \in \bar{B}$.

(ii) C è chiuso se e solo se $\bar{C} \subset C$, e questo, per il punto (i), è equivalente a dire che se $(x_j) \subset C$ è una successione convergente a un $x \in X$, si ha $x \in C$. □

Una caratterizzazione analoga vale per i punti di accumulazione di un insieme in uno spazio metrico (esercizio 1.8).

Poiché allora dare una topologia su X significa dare la collezione degli insiemi aperti o, equivalentemente, quella degli insiemi chiusi, loro complementari, si vede che il punto (ii) del teorema precedente implica che la topologia di uno spazio metrico è fissata dalle successioni convergenti, o, in altre parole, che la conoscenza delle successioni convergenti di uno spazio metrico equivale alla conoscenza della sua topologia. Questo non vale invece per uno spazio topologico generale: il punto, come si vede dalla dimostrazione, è che in uno spazio metrico ogni punto ha una base numerabile di intorni, cosa in generale falsa in uno spazio topologico, come mostra l'esempio seguente.

Esempio 1.2.3. Sia $X = [0, 1]$, e su di esso si consideri, invece della topologia usuale, la topologia definita dichiarando che un insieme $C \subset [0, 1]$ è chiuso se e solo se è al più numerabile, o $C = [0, 1]$. Considerato allora l'insieme (non chiuso) $B = [0, 1)$, si avrà $\bar{B} = [0, 1]$ (è l'unico chiuso che contiene B). Ma se $(x_n) \subset B$ è una successione, $U := \{x_n : n \in \mathbb{N}\}^c$ è un aperto (è il complementare di un insieme numerabile, e dunque chiuso) contenente il punto $1 \in \bar{B}$, ed è quindi un suo intorno, ed essendo $x_n \notin U$ per ogni $n \in \mathbb{N}$, (x_n) non converge a 1.

Per ottenere un risultato analogo al teorema 1.2.2 per uno spazio topologico arbitrario è necessario generalizzare la nozione di successione, che è quanto ci accingiamo a fare.

Definizione 1.2.4. Sia I un insieme. Una relazione $\leq$ tra coppie elementi di I è detta un *ordine parziale* se per ogni $\alpha, \beta, \gamma \in I$ si ha

(i) $\alpha \leq \alpha$ (*riflessività*)

(ii) $\alpha \leq \beta$ e $\beta \leq \alpha$ implica $\alpha = \beta$ (*antisimmetria*);

(iii) $\alpha \leq \beta$ e $\beta \leq \gamma$ implica $\alpha \leq \gamma$ (*transitività*).

L'ordine parziale $\leq$ è inoltre detto *diretto* se

(iv) per ogni $\alpha, \beta \in I$ esiste $\gamma \in I$ tale che $\alpha, \beta \leq \gamma$.

La coppia $(I, \leq)$ sarà detta un *insieme parzialmente ordinato* se $\leq$ è un ordine parziale, e un *insieme diretto* se $\leq$ è anche diretto.

Notiamo esplicitamente che in generale, dati elementi $\alpha, \beta \in I$, non è detto che debba valere una delle relazioni $\alpha \leq \beta$ o $\beta \leq \alpha$.

Esempi 1.2.5. (a) $I = \mathbb{N}$ con l'usuale relazione d'ordine tra numeri interi è chiaramente un insieme diretto.

(b) Sia X uno spazio topologico, e, dato $x \in X$ si definisca una relazione $\leq$ nell'insieme $\mathscr{I}_x$ degli intorni di x dichiarando che $U \leq V$ se e solo se $V \subset U$. Si ha allora che $\mathscr{I}_x$ è, con questa relazione, un insieme diretto: la verifica delle proprietà (i)-(iii) è immediata, e per quanto riguarda la (iv) basta notare che dati $U, V \in \mathscr{I}_x$ la loro intersezione $U \cap V$ è ancora un intorno di x ed è contenuta sia in U che in V, e dunque $U, V \leq U \cap V$. In questo esempio esistono chiaramente coppie di elementi $U, V \in \mathscr{I}_x$ che non sono in relazione tra loro.

Un altro esempio interessante è fornito dall'esercizio 1.9(a).

Definizione 1.2.6. Sia X uno spazio topologico. Un *net* in X è una funzione $x : I \to X$, con I insieme diretto, che si indica comunemente con il simbolo $(x_\alpha)_{\alpha \in I}$, o semplicemente (x_α). Si dice che $x \in X$ è un limite del net (x_α), o che il net (x_α) è convergente a x, e si scrive

$$x \in \lim_{\alpha \in I} x_\alpha, \quad \text{o anche} \quad x_\alpha \to x,$$

se per ogni intorno U di x esiste $\alpha_U \in I$ tale che, per ogni $\alpha \geq \alpha_U$, si ha $x_\alpha \in U$.

È facile vedere che se $I = \mathbb{N}$ si ritrova la nozione di convergenza di una successione, definizione 1.2.1, e che, un net $(x_\alpha)_{\alpha \in I}$ (con I insieme diretto generico) in uno spazio metrico (X, d) converge a $x \in X$ se e solo se $d(x_\alpha, x) \to 0$ come net in $\mathbb{R}$ (esercizio 1.10). Un esempio notevole di net convergente è dato dall'esercizio 1.9(b).

Proposizione 1.2.7. *Siano X uno spazio topologico di Hausdorff, e $(x_\alpha)_{\alpha \in I} \subset X$ un net. Se $x, y \in X$ sono limiti di (x_α), allora $x = y$.*

Dimostrazione. Sia, per assurdo, $x \neq y$. Essendo X di Hausdorff, esistono $U \in \mathscr{I}_x$, $V \in \mathscr{I}_y$ disgiunti, e poiché $x_\alpha \to x$ e $x_\alpha \to y$ esisteranno $\alpha_U, \alpha_V \in I$ tale che se $\alpha \geq \alpha_U$ allora $x_\alpha \in U$, e se $\alpha \geq \alpha_V$ allora $x_\alpha \in V$. Ma dal fatto che I è diretto segue che esiste $\bar{\alpha} \geq \alpha_U, \alpha_V$, e dunque $x_{\bar{\alpha}} \in U \cap V$, il che è assurdo. □

Il prossimo teorema mostra che i net svolgono, per uno spazio topologico generale, lo stesso ruolo che le successioni svolgono per gli spazi metrici.

Teorema 1.2.8. *Siano X uno spazio topologico, e $B, C \subset X$ insiemi. Si ha*

(i) $x \in \bar{B}$ se e solo se esiste un net $(x_\alpha) \subset B$ tale che $x_\alpha \to x$;

(ii) C è chiuso se e solo se per ogni net $(x_\alpha) \subset C$ che sia convergente ad un $x \in X$, si ha $x \in C$.

Dimostrazione. (i) Sia $x \in \bar{B}$. Per la proposizione 1.1.11, per ogni $U \in \mathscr{I}_x$ esisterà $x_U \in U \cap B$. Ricordando allora l'esempio 1.2.5(b), si vede che $(x_U)_{U \in \mathscr{I}_x}$ è un net, e per esso si ha $x_U \to x$: se $V \geq U$ vale $x_V \in V \subset U$. Viceversa se $(x_\alpha) \subset B$ converge a x, dato U intorno di x si ha $x_\alpha \in U \cap B$ per $\alpha \geq \alpha_U$, e dunque $x \in \bar{B}$.

(ii) Segue da (i) come nel caso degli spazi metrici, teorema 1.2.2. □

Dunque per definire una topologia su uno spazio X è sufficiente, invece che dare la classe degli insiemi aperti (o chiusi), dare la classe dei net convergenti.

Passiamo ora a considerare il caso generale di limiti di funzioni.

Definizione 1.2.9. Siano X, Y spazi topologici, $D \subset X$ un sottoinsieme, $x_0 \in X$ un punto di accumulazione per D, e $f : D \to Y$ una funzione. Si dice che il punto $\ell \in Y$ è un limite di f in x_0, e si scrive

$$\ell \in \lim_{x \to x_0} f(x),$$

se per ogni intorno U di ℓ in Y esiste un intorno V di x_0 in X tale che $f\big((D \setminus \{x_0\}) \cap V\big) \subset U$.

Usando la nozione di net data sopra, si può formulare un'analogo, per il concetto di limite appena definito, del "teorema ponte" noto dai corsi di base di Calcolo.

Proposizione 1.2.10. *Siano X, Y, D, x_0 e f come nella definizione di sopra. Si ha $\ell \in \lim_{x\to x_0} f(x)$ se e solo se per ogni net $(x_\alpha) \subset D \setminus \{x_0\}$ tale che $x_\alpha \to x_0$, si ha $f(x_\alpha) \to \ell$.*

Dimostrazione. Sia $\ell \in \lim_{x\to x_0} f(x)$, e siano U, V intorni di ℓ, x_0 rispettivamente come nella definizione 1.2.9. Se allora $(x_\alpha) \subset D \setminus \{x_0\}$ è un net tale che $x_\alpha \to x_0$, esisterà α_V tale che $x_\alpha \in V \cap (D \setminus \{x_0\})$ per ogni $\alpha \geq \alpha_V$, e dunque $f(x_\alpha) \in U$, da cui $f(x_\alpha) \to \ell$.

Viceversa se, per assurdo, $\ell \notin \lim_{x\to x_0} f(x)$, negando la definizione si vede che deve esistere un intorno U di ℓ in Y tale che per ogni intorno V di x_0 in X esista un $x_V \in (D \setminus \{x_0\}) \cap V$ per il quale $f(x_V) \notin U$. Ma allora, analogamente alla dimostrazione del teorema 1.2.8, $(x_V)_{V \in \mathscr{I}_{x_0}}$ è un net tale che $x_V \to x_0$, ma chiaramente $f(x_V) \not\to \ell$ (poiché non entra mai nell'intorno U), contro l'ipotesi. □

Da questo risultato e dalla proposizione 1.2.7 si ha il corollario seguente.

Corollario 1.2.11. *Il limite di una funzione a valori in uno spazio di Hausdorff, se esiste, è unico.*

Nel caso particolare in cui f è una funzione tra spazi metrici, la nozione di limite assume una forma più familiare dai corsi di Calcolo, in termini di "epsilon e delta".

Proposizione 1.2.12. *Siano (X, d), (Y, ρ) spazi metrici, $D \subset X$ con $x_0 \in X$ come punto di accumulazione, $f : D \to Y$. Si ha $\lim_{x\to x_0} f(x) = \ell$ se e solo se per ogni $\varepsilon > 0$ esiste $\delta_\varepsilon > 0$ tale che se $x \in D$ e $0 < d(x, x_0) < \delta_\varepsilon$, allora $\rho(f(x), \ell) < \varepsilon$.*

Dimostrazione. Sia $\lim_{x\to x_0} f(x) = \ell$. Data allora una palla $B_\varepsilon(\ell) \subset Y$, che è un intorno di ℓ, esisterà un intorno V di x_0 tale che $f(x) \in B_\varepsilon(\ell)$ per ogni $x \in (D \setminus \{x_0\}) \cap V$, ma poiché le palle formano un base di intorni in uno spazio metrico, esisterà $B_{\delta_\varepsilon}(x_0) \subset V$, e pertanto se $x \in D$ e $0 < d(x, x_0) < \delta_\varepsilon$, allora $x \in (D \setminus \{x_0\}) \cap B_{\delta_\varepsilon}(x_0) \subset (D \setminus \{x_0\}) \cap V$, e dunque $\rho(f(x), \ell) < \varepsilon$.

Viceversa dato un intorno U di ℓ in Y, esisterà una palla $B_\varepsilon(\ell) \subset U$. Preso allora $\delta_\varepsilon > 0$ come nell'enunciato e posto $V = B_{\delta_\varepsilon}(x_0)$, è chiaro che se $x \in (D \setminus \{x_0\}) \cap V$ allora $f(x) \in B_\varepsilon(\ell) \subset U$. □

In base a quanto visto sopra circa il ruolo delle successioni nel definire la topologia degli spazi metrici, non è sorprendente che nella situazione della proposizione appena dimostrata (f funzione tra spazi metrici), il "teorema ponte" 1.2.10 si può formulare usando successioni invece che net arbitrari (esercizio 1.11).

Data la nozione di limite per una funzione tra spazi topologici, la nozione di continuità si definisce in modo del tutto ovvio.

Definizione 1.2.13. Siano X, Y spazi topologici. Una funzione $f : X \to Y$ si dirà *continua in* $x_0 \in X$ se $f(x_0) \in \lim_{x \to x_0} f(x)$ (in particolare $f(x_0) = \lim_{x \to x_0} f(x)$ se Y è di Hausdorff). f si dirà continua in X se è continua in ogni $x_0 \in X$.

Si vede facilmente, con dimostrazione analoga a quella per funzioni su $\mathbb{R}$, che composizioni di funzioni continue sono continue (esercizio 1.12). Da questo ad esempio segue subito che le usuali operazioni algebriche con funzioni continue da uno spazio topologico a $\mathbb{C}$ producono funzioni continue: somme e prodotti di funzioni continue sono continue, l'inversa di una funzione continua è continua dove non si annulla ecc.

Il prossimo teorema fornisce un'importante e utile caratterizzazione delle funzioni continue.

Teorema 1.2.14. *Siano X, Y spazi topologici. Una funzione $f : X \to Y$ è continua, se e solo se per ogni aperto $A \subset Y$ la sua controimmagine*

$$f^{-1}(A) := \{x \in X \,:\, f(x) \in A\}$$

è un aperto di X.

Dimostrazione. Sia $f : X \to Y$ continua, e sia $A \subset Y$ un aperto. Per dimostrare che $f^{-1}(A) \subset X$ è aperto basta dimostrare che è intorno di ogni suo punto. Sia allora $x \in f^{-1}(A)$. Essendo $f(x) \in A$ e A aperto, A sarà intorno di $f(x)$ e quindi, per la continuità di f, esisterà un intorno $V \subset X$ di x tale che $f(V) \subset A$, o, equivalentemente, $x \in V \subset f^{-1}(A)$, e dunque $f^{-1}(A)$ è intorno di x.

Viceversa supponiamo che la controimmagine tramite f di ogni aperto di Y sia un aperto di X. Bisogna mostrare che dato un $x \in X$, f è continua in x. Se U è un intorno di $f(x)$, per definizione esiste un $A \subset U$ aperto che contiene $f(x)$. Ma allora $V := f^{-1}(A) \subset X$ è un aperto che contiene x, ed è dunque un suo intorno, e inoltre è tale che $f(V) \subset A \subset U$, da cui la continuità di f in x. □

Nell'esercizio 1.13 è data l'ovvia formulazione "duale" di questo risultato, con i chiusi al posto degli aperti.

1.3 Operatori limitati e norme equivalenti

Come visto, gli spazi topologici più "semplici" sono gli spazi normati. Tra tutte le funzioni tra due spazi normati (complessi) X e Y, una classe particolarmente naturale, individuata dal fatto che i suoi elementi rispettano la struttura vettoriale, è quella delle funzioni lineari, dette anche *operatori lineari*, cioè delle applicazioni $T : X \to Y$ tali che $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$ per ogni $x, y \in X$, $\alpha, \beta \in \mathbb{C}$, già considerate nel corso di Geometria. Tali operatori per noi saranno molto importanti, se non altro per il ruolo fondamentale che svolgono in Meccanica Quantistica. Per essi, la continuità si esprime in modo particolarmente semplice.

Definizione 1.3.1. Siano X, Y spazi normati. Un operatore lineare $T : X \to Y$ è detto *limitato*, se

$$\|T\| := \sup_{\|x\| \leq 1} \|Tx\| < +\infty.$$

Teorema 1.3.2. *Siano X, Y spazi normati e $T : X \to Y$ lineare. Sono equivalenti:*

(i) T è limitato;

(ii) T è continuo nel vettore $0 \in X$;

(iii) T è continuo su X.

Dimostrazione. (i) $\Rightarrow$ (ii). Dato $x \in X \setminus \{0\}$, si ha

$$\|Tx\| = \left\| T\left(\frac{x}{\|x\|}\right) \right\| \|x\| \leq \|T\| \|x\|,$$

avendo usato la linearità di T e l'omoegenità della norma nel primo passaggio, e il fatto che $x/\|x\|$ ha norma 1 nel secondo. Dato allora $\varepsilon > 0$ e posto $\delta_\varepsilon := \varepsilon/\|T\|$, se $\|x - 0\| = \|x\| < \delta_\varepsilon$ si ha $\|Tx - T0\| = \|Tx\| \leq \|T\|\|x\| < \varepsilon$, da cui, ricordando la proposizione 1.2.12, la continuità di T in 0.

(ii) $\Rightarrow$ (i). Per la continuità di T in 0, dato $\varepsilon > 0$, esisterà $\delta > 0$ tale che se $\|x\| < \delta$, allora $\|Tx\| < \varepsilon$. Allora qualunque sia $x \in X$ tale che $\|x\| \leq 1$, si avrà

$$\|Tx\| = \left\| T\left(\frac{x\delta}{2}\right) \right\| \frac{2}{\delta} < \frac{2\varepsilon}{\delta},$$

dove nel secondo passaggio si è usato il fatto che $\|\frac{x\delta}{2}\| = \delta/2 < \delta$. Ne segue $\|T\| < 2\varepsilon/\delta$.

(ii) $\Rightarrow$ (iii). Siano ε e δ come sopra. Dato un qualunque $x_0 \in X$, se $\|x - x_0\| < \delta$ si avrà $\|Tx - Tx_0\| = \|T(x - x_0)\| < \varepsilon$ per la linearità di T, e dunque T è continuo in x_0.

(iii) $\Rightarrow$ (ii). Ovvio. □

Conformandoci, in questo come in altri casi, all'uso della letteratura anglosassone, indicheremo con $B(X, Y)$ l'insieme di tutti gli operatori lineari limitati da X a Y (*bounded* è la parola inglese per "limitato"), e, per semplicità, porremo $B(X) := B(X, X)$ (operatori limitati di uno spazio normato X in se stesso).

Proposizione 1.3.3. *Siano X, Y, Z spazi normati. Allora*

(i) $B(X, Y)$ è uno spazio vettoriale con le operazioni definite puntualmente:

$$(T + S)x := Tx + Sx, \quad (\alpha T)x := \alpha Tx, \qquad T, S \in B(X, Y),\ x \in X,\ \alpha \in \mathbb{C},$$

e $T \in B(X, Y) \mapsto \|T\|$ è una norma su di esso, detta norma operatoriale*;*

(ii) definita, per ogni $T \in B(X,Y)$, $S \in B(Y,Z)$, la composizione $S \circ T : X \to Z$ tramite $(S \circ T)x := S(Tx)$, $x \in X$, si ha $\|S \circ T\| \leq \|S\|\|T\|$.

Dimostrazione. (i) È chiaro che $T + S$ e αT definiti nell'enunciato sono ancora operatori lineari da X a Y. Per dimostrare che $B(X,Y)$ è uno spazio vettoriale bisogna quindi dimostrare che sono limitati. Nel fare questo dimostreremo anche che $\|\cdot\|$ è una norma. Per ogni $x \in X$, $\|x\| \leq 1$, si ha, per la disuguaglianza triangolare della norma su Y e per la definizione di $\|T\|$,

$$\|(T+S)x\| = \|Tx + Sx\| \leq \|Tx\| + \|Sx\| \leq \|T\| + \|S\|,$$

e dunque $\|T + S\| = \sup_{\|x\|\leq 1} \|(T + S)x\| \leq \|T\| + \|S\| < +\infty$ e $T + S \in B(X,Y)$. Analogamente, sempre con $x \in X$, $\|x\| \leq 1$,

$$\|(\alpha T)x\| = \|\alpha\, Tx\| = |\alpha|\|Tx\|,$$

da cui, prendendo l'estremo superiore su tutti i vettori $x \in X$ con $\|x\| \leq 1$, $\|\alpha T\| = |\alpha|\|T\| < +\infty$ e $\alpha T \in B(X,Y)$. Infine se $\|T\| = 0$, dalla diseguaglianza $\|Tx\| \leq \|T\|\|x\|$, $x \in X$, mostrata sopra, segue subito che $T = 0$.

(ii) Analogamente a sopra, la tesi segue subito dalla diseguaglianza

$$\|(S \circ T)x\| = \|S(Tx)\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|,$$

valida per $x \in X$, $\|x\| \leq 1$. □

Nel contesto del punto (ii), la composizione di S e T si indicherà anche semplicemente con ST. In particolare, si ottiene dalla proposizione precedente che $B(X)$ è un'*algebra normata*, cioè uno spazio normato in cui sia definito anche un prodotto bilineare $(S,T) \in B(X) \times B(X) \mapsto ST \in B(X)$ che, rispetto alla norma, soddisfi la proprietà (ii) dell'enunciato (cf. definizione 4.1.3(a)).

Nel caso particolare, ma importante, in cui $Y = \mathbb{C}$, lo spazio $X^* := B(X,\mathbb{C})$ viene detto il *duale topologico* di X, e gli $f \in X^*$ vengono detti i *funzionali lineari (continui)* su X.

Esempio 1.3.4. Siano $X = C([a,b])$, $Y = \mathbb{C}$. Dato $c \in (a,b)$, consideriamo il funzionale $\delta_c : X \to Y$ definito da $\delta_c f := f(c)$, detto *delta di Dirac* (centrata in c). Usando su X la norma $\|\cdot\|_\infty$ (esempio 1.1.2(d)), si ha $\delta_c \in B(X,Y)$. Infatti,

$$|\delta_c f| = |f(c)| \leq \|f\|_\infty, \qquad f \in X,$$

da cui, prendendo $\|f\|_\infty \leq 1$, $\|\delta_c\| \leq 1$. Anzi si ha proprio $\|\delta_c\| = 1$, come si vede osservando che chiaramente esiste una $f \in X$ tale che $\|f\|_\infty \leq 1$ e $f(c) = \delta_c f = 1$. Il funzionale δ_c non è invece limitato se su X si considera la norma $\|\cdot\|_1$ (esempio 1.1.2(e)). Se infatti si considerano le funzioni $(f_n)_{n\in\mathbb{N}} \subset X$ il cui grafico è il triangolo isoscele di base $[c - 1/n, c + 1/n]$ e altezza n, si avrà chiaramente $\|f_n\|_1 = \int_a^b |f_n(t)|\, dt = 1$ e $\delta_c f = f(c) = n$, da cui $\|\delta_c\| = \sup_{\|f\|_1\leq 1} |\delta_c f| = +\infty$ (in sostanza, si possono avere funzioni continue i cui integrali sono limitati ma che in un dato punto possono essere arbitrariamente grandi).

Dunque uno stesso operatore può essere limitato o meno a seconda di quali norme si considerano sugli spazi dominio o codominio. Un altro esempio a proposito è fornito dall'esercizio 1.15.

È chiaro dalla definizione che, grosso modo, la norma $\|T\|$ misura il massimo degli allungamenti dei vettori di X sotto l'azione dell'operatore T. Non è allora sorprendente che, nel caso di operatori $T : \mathbb{C}^n \to \mathbb{C}^n$ (o, equivalentemente, da uno spazio finito-dimensionale in se stesso), la norma si può calcolare in termini degli autovalori di T (quindi in particolare un tale T è automaticamente continuo, esercizio 1.16). Vedremo nel capitolo 4 che questo risultato si generalizza a operatori tra spazi infinito-dimensionali, e oltre (teorema 4.1.13(iii)), a patto di dare la giusta definizione di spettro per tali operatori (che non coinciderà con l'insieme degli autovalori).

Nelle applicazioni può accadere che ciò che interessa veramente è non tanto la norma, ma la topologia che essa definisce, e anzi è possibile che norme diverse inducano su uno spazio la stessa topologia. In tal caso esse si dicono *topologicamente equivalenti* (o semplicemente equivalenti). Il risultato seguente fornisce un criterio utile allo scopo di verificare tale proprietà.

Proposizione 1.3.5. *Sia X uno spazio vettoriale. Due norme $\|\cdot\|_\alpha$, $\|\cdot\|_\beta$ su X sono topologicamente equivalenti se e solo se esistono constanti $c, C > 0$ tali che*

$$c\|x\|_\alpha \leq \|x\|_\beta \leq C\|x\|_\alpha, \qquad \forall\, x \in X. \tag{1.3}$$

Dimostrazione. La disuguaglianza di destra della (1.3) è equivalente al fatto che l'applicazione identica $\mathbb{1} : X \to X$, definita da $\mathbb{1}(x) := x$, $x \in X$, è limitata (con norma $\leq C$), se pensata come applicazione da $(X, \|\cdot\|_\alpha)$ a $(X, \|\cdot\|_\beta)$. Questo a sua volta è equivalente, per il teorema 1.3.2, al fatto che tale applicazione è continua, e dunque, per il teorema 1.2.14, al fatto che per ogni $A \subset X$ che sia aperto rispetto alla topologia τ_β definita da $\|\cdot\|_\beta$, si abbia che $A = \mathbb{1}^{-1}(A)$ è aperto nella topologia τ_α definita da $\|\cdot\|_\alpha$, cioè al fatto che $\tau_\beta \subset \tau_\alpha$. Analogamente si vede che la prima disuguaglianza della (1.3) è equivalente a $\tau_\alpha \subset \tau_\beta$. Quindi la (1.3) è equivalente a $\tau_\alpha = \tau_\beta$, cioè la tesi. □

Utilizzeremo la notazione abbreviata $\|\cdot\|_\beta \lesssim \|\cdot\|_\alpha$ per indicare che vale la disuguaglianza a destra della (1.3), e la notazione $\|\cdot\|_\beta \lnsim \|\cdot\|_\alpha$ per indicare che vale la disuguaglianza a destra della (1.3) ma non quella a sinistra (e quindi che le due norme non sono equivalenti). In tal caso si dice anche che la norma $\|\cdot\|_\alpha$ è *(strettamente) più forte* della $\|\cdot\|_\beta$.

Esempio 1.3.6. Sullo spazio $X = C([a, b])$ abbiamo definito le norme $\|\cdot\|_\infty$, $\|\cdot\|_1$ e $\|\cdot\|_2$, negli esempi 1.1.2(d), (e), (f) rispettivamente. Mostriamo ora che si ha

$$\|\cdot\|_1 \lnsim \|\cdot\|_2 \lnsim \|\cdot\|_\infty.$$

Si hanno le disuguaglianze ovvie

$$\|f\|_1 = \int_a^b |f(t)|\, dt \leq \|f\|_\infty (b - a),$$

$$\|f\|_2 = \left(\int_a^b |f(t)|^2\, dt \right)^{1/2} \leq \|f\|_\infty (b - a)^{1/2},$$

e, usando la disuguaglianza di Cauchy-Schwarz, teorema 3.1.2,

$$\|f\|_1 = \int_a^b |f(t)| \cdot 1\, dt \le \|f\|_2 \|1\|_2 = \|f\|_2 (b-a)^{1/2},$$

e dunque

$$\|f\|_1 \le (b-a)^{1/2} \|f\|_2 \le (b-a)\|f\|_\infty, \qquad \forall f \in X,$$

da cui $\|\cdot\|_1 \lesssim \|\cdot\|_2 \lesssim \|\cdot\|_\infty$. Mostriamo poi che $\|\cdot\|_\infty \not\lesssim \|\cdot\|_2$. A tale scopo assumiamo, per semplicità, che $[a,b] = [0,1]$, e consideriamo la successione di funzioni

$$f_n(t) := \begin{cases} 0 & t \in [1/n, 1] \\ 1 - nt & t \in [0, 1/n] \end{cases} \qquad n \in \mathbb{N}.$$

Si ha allora

$$\|f_n\|_2^2 = \int_0^{1/n} (1-nt)^2\, dt = \frac{1}{3n} \to 0 \qquad \text{per } n \to +\infty,$$

mentre $\|f_n\|_\infty = 1$ per ogni $n \in \mathbb{N}$, per cui non può esistere $C > 0$ tale che $\|f_n\|_\infty \le C\|f_n\|_2$ per ogni $n \in \mathbb{N}$ (o, equivalentemente, $\mathbb{1} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_\infty)$ non può essere continua). Infine, allo scopo di mostrare che $\|\cdot\|_2 \not\lesssim \|\cdot\|_1$, consideriamo la successione di funzioni

$$g_n(t) := \begin{cases} n & t \in [0, 1/n^2] \\ \frac{1}{\sqrt{t}} & t \in [1/n^2, 1] \end{cases} \qquad n \in \mathbb{N},$$

ottenuta troncando all'altezza n la funzione $g(t) = \frac{1}{\sqrt{t}}$ che è integrabile in $[0,1]$, ma il cui quadrato non è integrabile. Infatti si ha, per $n \to +\infty$,

$$\begin{aligned} \|g_n\|_2^2 &= \int_0^{1/n^2} n^2\, dt + \int_{1/n^2}^1 \frac{dt}{t} = 1 + \int_{1/n^2}^1 \frac{dt}{t} \to +\infty, \\ \|g_n\|_1 &= \int_0^{1/n^2} n\, dt + \int_{1/n^2}^1 \frac{dt}{\sqrt{t}} = \frac{1}{n} + \int_{1/n^2}^1 \frac{dt}{\sqrt{t}} \to \int_0^1 \frac{dt}{\sqrt{t}} < +\infty, \end{aligned}$$

e dunque anche in questo caso non può esistere $C > 0$ tale che $\|g_n\|_2 \le C\|g_n\|_1$ per ogni $n \in \mathbb{N}$.

Un altro esempio di norme non equivalenti è fornito dall'esercizio 1.17. Il teorema seguente mostra che l'esistenza di norme non equivalenti distingue drasticamente gli spazi infinito-dimensionali da quelli finito-dimensionali.

Teorema 1.3.7. *Sia X uno spazio vettoriale finito-dimensionale. Tutte le norme su X sono equivalenti.*

Dimostrazione. Sia $(e_i)_{i=1,\dots,n}$ una base di X. Allora ogni $x \in X$ si scrive in modo unico come $x = \sum_{i=1}^n \xi_i e_i$, $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{C}^n$, ed è chiaro che

$$\|x\|_1 := \sum_{i=1}^n |\xi_i|$$

è una norma su X. Basta allora dimostrare che se $\|\cdot\|$ è una norma su X, essa è equivalente a $\|\cdot\|_1$. Si ha

$$\Big\|\sum_{i=1}^n \xi_i e_i\Big\| \le \sum_{i=1}^n |\xi_i|\|e_i\| \le \Big(\max_{i=1,\dots,n}\|e_i\|\Big)\|x\|_1, \qquad \forall\, x\in X,$$

e dunque $\|\cdot\| \lesssim \|\cdot\|_1$ (prendendo $C := \max_{i=1,\dots,n}\|e_i\|$). Per concludere bisogna allora dimostrare che esiste $c>0$ tale che $\|x\|_1 \le c\|x\|$ per ogni $x\in X$, cioè, supponendo $x\neq 0$ (se $x=0$ la disuguaglianza è ovvia) e dividendo per $\|x\|_1$,

$$\Big\|\frac{x}{\|x\|_1}\Big\| \ge \frac{1}{c} > 0 \qquad \forall x\in X,\ x\neq 0,$$

o, equivalentemente,

$$\inf_{\|x\|_1=1}\|x\| = \inf_{\substack{\xi\in\mathbb{C}^n\\ \|\xi\|_1=1}}\Big\|\sum_{i=1}^n \xi_i e_i\Big\| > 0.$$

Ma l'insieme $S := \{\xi\in\mathbb{C}^n : \|\xi\|_1 = 1\}$ è chiuso e limitato in $\mathbb{C}^n$ (con la topologia usuale): chiuso in quanto controimmagine dell'insieme chiuso $\{1\}\subset\mathbb{R}$ tramite l'applicazione continua $\xi\in\mathbb{C}^n \mapsto \|\xi\|_1 = \sum_{i=1}^n|\xi_i|$, e limitato poiché $\|\xi\|_1 = 1$ implica $|\xi_i|\le 1$ per ogni $i=1,\dots,n$. Inoltre l'applicazione $\xi\mapsto\|\sum_{i=1}^n\xi_i e_i\|$ è pure continua su $\mathbb{C}^n$ poiché

$$\bigg|\Big\|\sum_{i=1}^n \xi_i e_i\Big\| - \Big\|\sum_{i=1}^n \eta_i e_i\Big\|\bigg| \le \Big\|\sum_{i=1}^n (\xi_i-\eta_i) e_i\Big\| \le C\|\xi-\eta\|_1,$$

avendo usato nella prima disuguaglianza l'esercizio 1.18 e nella seconda il fatto che $\|\cdot\|\lesssim\|\cdot\|_1$ visto sopra. Dunque per il teorema di Weierstrass

$$\inf_{\substack{\xi\in\mathbb{C}^n\\ \|\xi\|_1=1}}\Big\|\sum_{i=1}^n \xi_i e_i\Big\| = \min_{\xi\in S}\Big\|\sum_{i=1}^n \xi_i e_i\Big\| > 0$$

e si conclude. □

Dunque in particolare le norme $\|\cdot\|_\infty$, $\|\cdot\|_1$ e $\|\cdot\|_2$ su $\mathbb{C}^n$ sono equivalenti. Per esse, è anche facile calcolare le costanti che intervengono nella (1.3) (esercizio 1.19). Un altro esempio di norme equivalenti, questo su spazi anche infinito-dimensionali, è dato nell'esercizio 1.20, in cui viene anche introdotta l'importante nozione di *somma diretta* di due spazi normati.

Un'altra conseguenza interessante del teorema, e dell'esercizio 1.16, è che un operatore lineare con dominio finito-dimensionale è automaticamente limitato (esercizio 1.21).

1.4 Completezza

Come noto dai corsi di Calcolo, una proprietà essenziale dell'insieme $\mathbb{R}$ dei numeri reali, che lo distingue nettamente da quello dei razionali, è la completezza, cioè il fatto che ogni successione di Cauchy in $\mathbb{R}$ è convergente. Grazie a essa, ad esempio, è possibile definire la radice quadrata di un numero reale non negativo. Essendo la nozione di convergenza di una successione in uno spazio metrico una diretta generalizzazione di quella di successioni reali, non è sorprendente che la nozione di completezza giochi un ruolo fondamentale anche nella teoria generale degli spazi metrici.

Definizione 1.4.1. Sia (X, d) uno spazio metrico. Una successione $(x_n)_{n\in\mathbb{N}} \subset X$ è detta una *successione di Cauchy*, se per ogni $\varepsilon > 0$ esiste $n_\varepsilon \in \mathbb{N}$ tale che se $n, m \geq n_\varepsilon$ allora $d(x_n, x_m) < \varepsilon$.

Si vede facilmente, come per le successioni reali, che ogni successione convergente in uno spazio metrico è di Cauchy (esercizio 1.22). Come per $\mathbb{R}$, la completezza di uno spazio metrico è il fatto che valga anche il viceversa.

Definizione 1.4.2. Uno spazio metrico (X, d) si dice *completo* se ogni sua successione di Cauchy è convergente. Uno spazio normato $(X, \|\cdot\|)$ si dice uno *spazio di Banach* se è completo nella metrica indotta dalla norma.

È noto, sempre dai corsi di Calcolo, che $\mathbb{R}^n$ e $\mathbb{C}^n$ con la norma euclidea sono spazi di Banach. Da questo, e dal teorema 1.3.7 segue che ogni spazio normato finito-dimensionale è di Banach (esercizio 1.23). Un semplice esempio di spazio metrico non completo è fornito nell'esercizio 1.24. Di seguito vediamo alcuni esempi più interessanti.

Esempi 1.4.3. (a) Lo spazio $(\ell^\infty(T), \|\cdot\|_\infty)$, definito nell'esempio 1.1.2(c), è uno spazio di Banach. Sia infatti $(f_n) \subset \ell^\infty(T)$ una successione di Cauchy. Poiché allora

$$|f_n(t) - f_m(t)| \leq \|f_n - f_m\|_\infty \qquad \forall\, t \in T,$$

la successione $(f_n(t))_{n\in\mathbb{N}} \subset \mathbb{C}$ è di Cauchy in $\mathbb{C}$ per ogni $t \in T$, ed essendo $\mathbb{C}$ completo esisterà

$$f(t) := \lim_{n\to+\infty} f_n(t) \qquad \forall\, t \in T.$$

Si ottiene in tal modo una funzione $f : T \to \mathbb{C}$. Per concludere bisogna allora dimostrare che $f \in \ell^\infty(T)$ e che $\|f_n - f\|_\infty \to 0$. A tale scopo, fissato $\varepsilon > 0$ sia $n_\varepsilon \in \mathbb{N}$ come nella definizione di successione di Cauchy. Allora in base a quanto appena visto si ha, per ogni $n > n_\varepsilon$,

$$|f_n(t) - f(t)| = \lim_{m\to+\infty} |f_n(t) - f_m(t)| \leq \limsup_{m\to+\infty} \|f_n - f_m\|_\infty \leq \varepsilon^\dagger \qquad \forall\, t \in T,$$

poichè definitivamente $m > n_\varepsilon$ per $m \to +\infty$. Da questo segue $\|f_n - f\|_\infty \to 0$ per $n \to +\infty$, e inoltre, fissato $n \in \mathbb{N}$,

$$|f(t)| \leq |f(t) - f_n(t)| + |f_n(t)| \leq \|f - f_n\|_\infty + \|f_n\|_\infty < +\infty \qquad \forall t \in T,$$

e quindi f è limitata, cioè $f \in \ell^\infty(T)$.

† Per richiami sulle nozioni di lim sup e lim inf di una successione, e sulle loro principali proprietà, si veda l'esercizio 1.25.

(b) Dall'esempio precedente segue subito che anche $(C([a,b]), \|\cdot\|_\infty)$ è uno spazio di Banach. Basta infatti osservare che $\|f_n - f\|_\infty \to 0$ equivale a dire che la successione di funzioni (f_n) converge uniformemente in $[a,b]$ alla funzione f. Se allora $f \in \ell^\infty([a,b])$ è il limite della successione di Cauchy $(f_n) \subset C([a,b]) \subset \ell^\infty([a,b])$ costruito sopra, dai teoremi sulla convergenza uniforme visti nei corsi di Calcolo è noto che $f \in C([a,b])$.

(c) Sia I un insieme arbitrario, e si indichi con $\mathscr{P}_0(I)$ l'insieme dei suoi sottoinsiemi finiti (o parti finite). Su $\mathscr{P}_0(I)$ si può definire un ordinamento parziale diretto dichiarando che $F \leq G$ se e solo se $F \subset G$. Dato allora uno spazio normato $(X, \|\cdot\|)$ e una famiglia $(x_\alpha)_{\alpha\in I}$ di vettori di X (che, malgrado la notazione, *non è* un net poiché non stiamo assumendo che I sia diretto), si può definire la sua somma $\sum_{\alpha\in I} x_\alpha$ come il limite, se esiste, delle somme finite $\sum_{\alpha\in F} x_\alpha$, $F \in \mathscr{P}_0(I)$, pensate come un net in X su $\mathscr{P}_0(I)$ (con l'ordinamento parziale appena definito):

$$\sum_{\alpha\in I} x_\alpha := \lim_{F\in\mathscr{P}_0(I)} \sum_{\alpha\in F} x_\alpha.$$

Ciò posto, consideriamo gli insiemi

$$\ell^1(I) := \Big\{ x = (x_\alpha)_{\alpha\in I} \subset \mathbb{C} \,:\, \|x\|_1 := \sum_{\alpha\in I} |x_\alpha| < +\infty \Big\},$$

$$\ell^2(I) := \Big\{ x = (x_\alpha)_{\alpha\in I} \subset \mathbb{C} \,:\, \|x\|_2 := \Big(\sum_{\alpha\in I} |x_\alpha|^2 \Big)^{1/2} < +\infty \Big\}.$$

Si verifica facilmente (esercizio 1.26) che $(\ell^1(I), \|\cdot\|_1)$, $(\ell^2(I), \|\cdot\|_2)$ sono spazi normati, naturali generalizzazioni infinito-dimensionali di $(\mathbb{C}^n, \|\cdot\|_1)$, $(\mathbb{C}^n, \|\cdot\|_2)$. Mostriamo che sono di Banach. Sia $(x_n)_{n\in\mathbb{N}} \subset \ell^p(I)$, $p = 1, 2$, una successione di Cauchy. Osserviamo esplicitamente, per evitare confusione, che ogni elemento x_n di una tale successione è a sua volta una famiglia, a priori infinita, di numeri complessi $x_n = (x_{n\alpha})_{\alpha\in I} \in \ell^p(I)$. Per dimostrare che esiste $x \in \ell^p(I)$ limite della successione (x_n) seguiamo la falsariga dell'esempio (a) di sopra. Grazie all'esercizio 1.27, si ha

$$|x_{n\alpha} - x_{m\alpha}| \leq \|x_n - x_m\|_p \qquad \forall\, \alpha \in I,$$

e dunque, per la completezza di $\mathbb{C}$, esisterà $x_\alpha := \lim_{n\to+\infty} x_{n\alpha} \in \mathbb{C}$ per ogni $\alpha \in I$. Posto allora $x := (x_\alpha)_{\alpha\in I}$, e scelti $\varepsilon > 0$ e $n_\varepsilon \in \mathbb{N}$ come nella definizione di successione di Cauchy, si avrà, per ogni $n > n_\varepsilon$ e per ogni $F \in \mathscr{P}_0(I)$,

$$\sum_{\alpha\in F} |x_{n\alpha} - x_\alpha|^p = \lim_{m\to+\infty} \sum_{\alpha\in F} |x_{n\alpha} - x_{m\alpha}|^p \leq \limsup_{m\to+\infty} \|x_n - x_m\|_p^p \leq \varepsilon^p,$$

da cui segue

$$\|x_n - x\|_p^p = \lim_F \sum_{\alpha\in F} |x_{n\alpha} - x_\alpha|^p \leq \varepsilon^p,$$

e cioè $\|x_n - x\|_p \to 0$. Inoltre osservando che, per ogni $F \in \mathscr{P}_0(I)$, $(x_\alpha)_{\alpha\in F} \mapsto (\sum_{\alpha\in F} |x_\alpha|^p)^{1/p}$ è la norma $\|\cdot\|_p$ su $\mathbb{C}^{\sharp F}$, si ha

$$\Big(\sum_{\alpha\in F} |x_\alpha|^p \Big)^{1/p} \leq \Big(\sum_{\alpha\in F} |x_\alpha - x_{n\alpha}|^p \Big)^{1/p} + \Big(\sum_{\alpha\in F} |x_{n\alpha}|^p \Big)^{1/p} \leq \|x - x_n\|_p + \|x_n\|_p,$$

da cui $\|x\|_p \leq \|x - x_n\|_p + \|x_n\|_p < +\infty$ e quindi $x \in \ell^p(I)$.

(d) Mostriamo che gli spazi $(C([0,1]), \|\cdot\|_1)$, $(C([0,1]), \|\cdot\|_2)$ non sono di Banach. A tale scopo consideriamo la successione di funzioni $(f_n)_{n\in\mathbb{N}} \subset C([0,1])$ definita da

$$f_n(t) := \begin{cases} n & t \in [0, e^{-n}] \\ -\log t & t \in [e^{-n}, 1] \end{cases} \qquad n \in \mathbb{N},$$

cioè il troncamento ad altezza n della funzione $t \mapsto -\log t$, che è integrabile (in senso improprio) e a quadrato integrabile in $[0,1]$, ma ovviamente non sta in $C([0,1])$. Mostriamo che tale successione è di Cauchy rispetto alla norma $\|\cdot\|_2$, e quindi anche rispetto alla norma $\|\cdot\|_1$ grazie a quanto visto nell'esempio 1.3.6 (cioè al fatto che $\|\cdot\|_1 \lesssim \|\cdot\|_2$). Per ogni $n, p \in \mathbb{N}$, si ha

$$\begin{aligned}
\|f_{n+p} - f_n\|_2^2 &= \int_0^1 |f_{n+p} - f_n|^2 && (f_{n+p}(t) = f_n(t) = -\log t \text{ per } t \in [e^{-n}, 1]\,) \\
&= \int_0^{e^{-n}} |f_{n+p} - f_n|^2 \\
&\le \int_0^{e^{-n}} (|f_{n+p}| + |f_n|)^2 && (|f_{n+p}(t)| + |f_n(t)| \le -2\log t\,) \\
&\le 4 \int_0^{e^{-n}} \log^2 t \, dt && (\text{per parti due volte}) \\
&= \big[t(\log^2 t - 2\log t + 2)\big]_0^{e^{-n}} \\
&= e^{-n}(n^2 + 2n + 2) \to 0 && \text{per } n \to +\infty,
\end{aligned}$$

e dunque $\|f_{n+p} - f_n\|_2$ può essere resa arbitrariamente piccola prendendo n sufficientemente grande. Basta allora dimostrare che la successione (f_n) non converge rispetto a $\|\cdot\|_1$, poiché in tal caso non converge nemmeno rispetto a $\|\cdot\|_2$, sempre per quanto visto nell'esempio 1.3.6. Sia allora $f \in C([0,1])$, e mostriamo che $\|f_n - f\|_1 \not\to 0$. A tale scopo fissiamo $\bar{n} > \|f\|_\infty$. Allora per ogni $n > \bar{n}$ si ha

$$\begin{aligned}
\|f_n - f\|_1 &= \int_0^1 |f_n - f| \ge \int_0^{e^{-\bar{n}}} |f_n - f| \ge \int_0^{e^{-\bar{n}}} (|f_n| - |f|) \\
&\ge \int_0^{e^{-\bar{n}}} (f_n(t) - \|f\|_\infty)\, dt \\
&= \int_0^{e^{-n}} n\, dt - \int_{e^{-n}}^{e^{-\bar{n}}} \log t\, dt - \|f\|_\infty e^{-\bar{n}} \\
&= ne^{-n} - \big[t\log t - t\big]_{e^{-n}}^{e^{-\bar{n}}} - \|f\|_\infty e^{-\bar{n}} \\
&= e^{-\bar{n}}(\bar{n} + 1 - \|f\|_\infty) - e^{-n},
\end{aligned}$$

e l'ultimo membro converge, per $n \to +\infty$, a $e^{-\bar{n}}(\bar{n} + 1 - \|f\|_\infty) > 0$ (per la scelta di $\bar{n}$). Si potrebbe pensare che, per ottenere spazi completi, basti aggiungere a $C([0,1])$ le funzioni impropriamente integrabili o a quadrato integrabile. Si può in realtà dimostrare che nemmeno questo è sufficiente, mentre vedremo che spazi completi di funzioni integrabili si costruiscono con relativa facilità ricorrerendo alla teoria dell'integrazione alla Lebesgue (teorema 2.3.4). E anzi questa è una delle motivazioni più forti per andare oltre la teoria dell'integrazione di Riemann e considerare quella di Lebesgue.

Altri esempi sono forniti nell'esercizio 1.28. Un'ulteriore importante classe di spazi di Banach si ottiene considerando operatori limitati a valori in uno spazio di Banach.

Proposizione 1.4.4. *Siano X uno spazio normato e Y uno spazio di Banach. Allora $B(X,Y)$ è uno spazio di Banach (con la norma operatoriale).*

Dimostrazione. La dimostrazione è basata ancora sulla stessa idea degli esempi 1.4.3(a) e (c). Sia dunque $(T_n)_{n\in\mathbb{N}} \subset B(X,Y)$ una successione di Cauchy. Poiché $\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|$ per ogni $x \in X$, $n, m \in \mathbb{N}$, e poiché per ipotesi Y è di Banach, esiste in Y il limite

$$Tx := \lim_{n\to+\infty} T_n x \qquad \forall\, x \in X.$$

Si verifica facilmente che l'applicazione $x \in X \mapsto Tx \in Y$ così definita è lineare:

$$\begin{aligned} T(\alpha_1 x_1 + \alpha_2 x_2) &= \lim_{n\to+\infty} T_n(\alpha_1 x_1 + \alpha_2 x_2) \\ &= \lim_{n\to+\infty} \alpha_1 T_n x_1 + \alpha_2 T_n x_2 = \alpha_1 T x_1 + \alpha_2 T x_2. \end{aligned}$$

Inoltre dato $\varepsilon > 0$ si avrà, per n sufficientemente grande,

$$\|T_n x - Tx\| = \lim_{m\to+\infty} \|T_n x - T_m x\| \leq \limsup_{m\to+\infty} \|T_n - T_m\|\|x\| \leq \varepsilon\|x\| \qquad \forall\, x \in X,$$

da cui $\|T_n - T\| \to 0$. Infine, fissato $n \in \mathbb{N}$,

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| \leq (\|T - T_n\| + \|T_n\|)\|x\| \qquad \forall\, x \in X,$$

e quindi $T \in B(X,Y)$. □

In particolare, se X è uno spazio di Bananch, $B(X)$ è un'*algebra di Banach*, cioè un'algebra normata completa rispetto alla metrica indotta dalla norma.

Un sottoinsieme D di uno spazio topologico X si dice *denso* in X se $\bar{D} = X$. Capita assai spesso nelle applicazioni che sia necessario definire un operatore su uno spazio normato, ma che sia "facile" farlo soltanto su un suo opportuno sottospazio denso. Si pone dunque il problema di estendere tale operatore a tutto lo spazio. Il fondamentale teorema seguente garantisce che questo è possibile se si sa dimostrare che l'operatore in questione è limitato (sul sottospazio su cui è definito) e che prende valori in uno spazio di Banach.

Teorema 1.4.5. *Siano X uno spazio normato, e Y uno spazio di Banach. Dato un sottospazio vettoriale D denso in X e un operatore limitato $T \in B(D,X)$, esiste un unico operatore $\bar{T} \in B(X,Y)$ che estende T (cioè tale che $\bar{T}x = Tx$ per ogni $x \in D$). Inoltre si ha $\|\bar{T}\| = \|T\|$.*

Dimostrazione. Dimostriamo l'esistenza di $\bar{T}$. Dato $x \in X = \bar{D}$, esisterà (per il teorema 1.2.2) una successione $(x_n) \subset D$ convergente a x, e quindi di Cauchy in X. Essendo $\|T\| < \infty$, dalla disuguaglianza

$$\|Tx_n - Tx_m\| \leq \|T\|\|x_n - x_m\|, \qquad n, m \in \mathbb{N},$$

segue allora che $(Tx_n)_{n\in\mathbb{N}}$ è una successione di Cauchy in Y di Banach, e dunque che esiste

$$\bar{T}x := \lim_{n\to+\infty} Tx_n, \qquad x \in X.$$

Tale limite è inoltre indipendente dalla successione (x_n) scelta: se infatti $(\tilde{x}_n) \subset D$ è un'altra successione convergente a $x \in X$, si avrà $\|Tx_n - T\tilde{x}_n\| \leq \|T\|\|x_n - \tilde{x}_n\| \to 0$. Dunque il limite di sopra definisce effettivamente un'applicazione $\bar{T} : X \to Y$ che estende T, in quanto se $x \in D$ per calcolare $\bar{T}x$ si può chiaramente prendere la successione $x_n = x$ per ogni $n \in \mathbb{N}$, e quindi $\bar{T}x = Tx$. Inoltre $\bar{T}$ così definita è lineare: se $x, y \in X$ e $\alpha, \beta \in \mathbb{C}$, prese successioni $(x_n), (y_n) \subset D$ tali che $x_n \to x$, $y_n \to y$, si ha $\alpha x_n + \beta y_n \in D$ per ogni $n \in \mathbb{N}$ (D è un sottospazio) e $\alpha x_n + \beta y_n \to \alpha x + \beta y$, e pertanto, per definizione di $\bar{T}$,

$$\bar{T}(\alpha x + \beta y) = \lim_{n\to+\infty} T(\alpha x_n + \beta y_n) = \lim_{n\to+\infty} \alpha\, Tx_n + \beta\, Ty_n = \alpha\, \bar{T}x + \beta\, \bar{T}y.$$

Infine si ha $\bar{T} \in B(X, Y)$: se $(x_n) \subset D$ converge a $x \in X$ si ha anche $\|x_n\| \to \|x\|$, $\|Tx_n\| \to \|\bar{T}x\|$ (esercizio 1.18), e quindi

$$\|\bar{T}x\| = \lim_{n\to+\infty} \|Tx_n\| \leq \lim_{n\to+\infty} \|T\|\|x_n\| = \|T\|\|x\|, \qquad x \in X,$$

da cui $\|\bar{T}\| \leq \|T\|$. Essendo poi $\|\bar{T}x\| = \|Tx\|$ per ogni $x \in D$ si ha anche $\|\bar{T}\| \geq \|T\|$, e quindi $\|\bar{T}\| = \|T\|$.

L'unicità di $\bar{T}$ è praticamente ovvia: se $\tilde{T} \in B(X, Y)$ è un'altra estensione di T, dato $x \in X$ e $(x_n) \subset D$ convergente a x sarà, per la continuità di $\tilde{T}$,

$$\tilde{T}x = \lim_{n\to+\infty} \tilde{T}x_n = (\tilde{T} \text{ estende } T) = \lim_{n\to+\infty} Tx = \bar{T}x,$$

cioè $\tilde{T} = \bar{T}$. □

In realtà, nel teorema precedente, l'ipotesi "Y spazio di Banach" non è veramente restrittiva, poiché ogni spazio normato si può pensare come un sottospazio denso di un opportuno spazio di Banach, e anzi, più in generale, un qualunque spazio metrico può essere immerso come un sottoinsieme denso in un opportuno spazio completo. Questo è precisato dall'importante teorema seguente, per formulare il quale introduciamo la nozione di *isometria* tra due spazi metrici (X, d) e (Y, ρ): è questa un'applicazione $\phi : X \to Y$ che rispetta le distanze, cioè tale che $\rho(\phi(x_1), \phi(x_2)) = d(x_1, x_2)$ per ogni $x_1, x_2 \in X$. È chiaro che un'isometria è iniettiva. Se è suriettiva (e quindi biunivoca) X e Y, come spazi metrici, si possono identificare.

Teorema 1.4.6 (Completamento di spazi metrici e normati)**.** *Dato uno spazio metrico (X, d), esistono uno spazio metrico completo $(\bar{X}, \bar{d})$, detto il* completamento di (X, d), *e un'isometria $j : X \to \bar{X}$, tali che $\overline{j(X)} = \bar{X}$ (cioè $j(X)$ è denso in $\bar{X}$). Inoltre il completamento di (X, d) è unico, nel senso che se $(\tilde{X}, \tilde{d})$ è un altro spazio completo e $k : X \to \tilde{X}$ è un'altra isometria tali che $\overline{k(X)} = \tilde{X}$, esiste un'isometria suriettiva $\phi : \bar{X} \to \tilde{X}$ tale che $\phi \circ j = k$.*

Infine, se $(X, \|\cdot\|)$ è uno spazio normato, il suo completamento (come spazio metrico con la metrica d indotta dalla norma) $\bar{X}$ ha una struttura di spazio vettoriale, la metrica $\bar{d}$ è indotta da una norma $\|\cdot\|^-$ e $j : X \to \bar{X}$ è lineare. Lo spazio di Banach $(\bar{X}, \|\cdot\|^-)$ è detto allora il completamento dello spazio normato $(X, \|\cdot\|)$.

Dimostrazione. Esistenza. Definiremo $\bar{X}$ come un opportuno sottoinsieme chiuso dello spazio di Banach $\ell^\infty(X)$ delle funzioni limitate da X a $\mathbb{C}$, dotato della metrica $\bar{d}$ indotta dalla norma $\|\cdot\|_\infty$. A tale scopo fissiamo un $a \in X$ e consideriamo, per ogni $x \in X$, l'applicazione $\varphi_x : X \to \mathbb{C}$ definita da

$$\varphi_x(\xi) := d(\xi, x) - d(\xi, a), \qquad \xi \in X.$$

Si ha $\varphi_x \in \ell^\infty(X)$ per ogni $x \in X$: infatti dall'esercizio 1.29 (conseguenza della disuguaglianza triangolare) segue

$$|\varphi_x(\xi)| = |d(\xi, x) - d(\xi, a)| \leq d(x, a) \qquad \forall \xi \in X,$$

e quindi $\|\varphi_x\|_\infty \leq d(x, a)$. Si può allora definire un'applicazione $j : X \to \ell^\infty(X)$ ponendo $j(x) := \varphi_x$, $x \in X$. Tale applicazione è isometrica. Si ha infatti, per ogni $x, y \in X$,

$$\bar{d}(j(x), j(y)) = \|\varphi_x - \varphi_y\|_\infty = \sup_{\xi \in X} |\varphi_x(\xi) - \varphi_y(\xi)| = \sup_{\xi \in X} |d(x, \xi) - d(y, \xi)| \leq d(x, y),$$

avendo usato ancora l'esercizio 1.29 nell'ultimo passaggio, ed essendo poi

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

si ha proprio $\bar{d}(j(x), j(y)) = \sup_{\xi \in X} |\varphi_x(\xi) - \varphi_y(\xi)| = d(x, y)$. Per concludere, basta allora porre $\bar{X} := \overline{j(X)}$ (chiusura in $\ell^\infty(X)$), che è completo, rispetto alla metrica ottenuta restringendo a esso la metrica di $\ell^\infty(X)$, grazie all'esercizio 1.30.

Unicità. Vedere il punto (e) dell'esercizio 1.31, la cui soluzione si basa sulle idee della dimostrazione del teorema 1.4.5.

Infine, per il completamento di uno spazio normato, vedere il punto (f) dell'esercizio 1.31. □

I punti (a)-(d) dell'esercizio 1.31 forniscono una costruzione alternativa del completamento, basata sulla stessa procedura che si usa per definire i numeri reali a partire da successioni di Cauchy di numeri razionali. Ma, per l'unicità del completamento, le due costruzioni danno luogo a spazi isometrici.

1.5 Compattezza

Un'altra proprietà fondamentale di $\mathbb{R}^n$, anch'essa nota dai corsi di Calcolo, è il fatto che le successioni contenute in insiemi chiusi e limitati (detti anche compatti) hanno sempre sottosuccessioni convergenti (a un punto dell'insieme), il che si rivela di grande utilità per dimostrare l'esistenza di "soluzioni" di certe equazioni costruendo opportunamente successioni di soluzioni "approssimate". Tale proprietà in realtà caratterizza gli insiemi chiusi e limitati in $\mathbb{R}^n$ (teorema di Bolzano-Weierstrass). È quindi naturale il desiderio di identificare gli insiemi che ne godono anche in spazi più generali, e vedremo nel capitolo **??** che tali insiemi giocano un ruolo fondamentale nella teoria spettrale degli operatori su uno spazio di Hilbert. Purtroppo però una generalizzazione diretta fallisce drammaticamente in dimensione infinita, come mostra il seguente risultato.

Proposizione 1.5.1. *Sia X uno spazio normato infinito-dimensionale. La palla unitaria (chiusa) di X,*

$$B_X := \{x \in X \,:\, \|x\| \leq 1\},$$

contiene successioni che non ammettono sottosuccessioni convergenti.

Dimostrazione. Basta dimostrare che esiste una successione $(x_j)_{j\in\mathbb{N}} \subset X$ tale che $\|x_j\| = 1$ e $\|x_j - x_k\| \geq 1/2$ per ogni $j, k \in \mathbb{N}$, $j \neq k$. Infatti una tale successione non potrà ammettere sottosuccessioni di Cauchy, e quindi tantomeno sottosuccessioni convergenti. Supponiamo allora induttivamente di aver costruito i primi k elementi $x_1, \ldots, x_k$ di tale successione, e poniamo $V_k := \langle x_1, \ldots, x_k \rangle$, il sottospazio vettoriale di X generato da $x_1, \ldots, x_k$ (cioè il sottospazio delle combinazioni lineari $\sum_{i=1}^k \alpha_i x_i$). Poiché ovviamente V_k ha dimensione finita ($\leq k$), esisterà un $y \in X \setminus V_k$. Posto allora

$$\operatorname{dist}(y, V_k) := \inf_{x \in V_k} \|y - x\|,$$

si avrà $\operatorname{dist}(y, V_k) > 0$: in caso contrario, esisterebbe, per le proprietà dell'estremo inferiore, una successione $(\xi_n) \subset V_k$ convergente a y, ma essendo V_k chiuso (esercizio 1.32), questo implicherebbe $y \in V_k$, contro la scelta fatta di y. Dunque, ancora per le proprietà dell'estremo inferiore, si può scegliere $y_k \in V_k$ tale che $\|y - y_k\| < 2 \operatorname{dist}(y, V_k)$. Posto allora

$$x_{k+1} := \frac{y - y_k}{\|y - y_k\|},$$

si ha chiaramente $\|x_{k+1}\| = 1$, e pertanto per concludere il passo induttivo, e quindi la dimostrazione, basta verificare che $\|x_{k+1} - x_j\| \geq 1/2$ per ogni $j = 1, \ldots, k$. Si ha infatti

$$\|x_{k+1} - x_j\| = \frac{\|y - (y_k + \|y - y_k\| x_j)\|}{\|y - y_k\|} > \frac{\operatorname{dist}(y, V_k)}{2 \operatorname{dist}(y, V_k)} = \frac{1}{2},$$

dove nel secondo passaggio si è usato il fatto che $y_k + \|y - y_k\| x_j \in V_k$. □

Esiste tuttavia un'altra caratterizzazione, puramente topologica, dei sottoinsiemi chiusi e limitati di $\mathbb{R}^n$, che fa intervenire la nozione di *ricoprimento aperto* di un insieme C, cioè di una famiglia di aperti $(U_\alpha)_{\alpha \in I}$, I insieme arbitrario di indici, tale che $C \subset \bigcup_{\alpha \in I} U_\alpha$. Se per un sottoinsieme $J \subset I$ risulta ancora $C \subset \bigcup_{\alpha \in J} U_\alpha$, la sottofamiglia $(U_\alpha)_{\alpha \in J}$ sarà detta un *sottoricoprimento* di $(U_\alpha)_{\alpha \in I}$. Ovviamente tali nozioni hanno senso in uno spazio topologico arbitrario.

Teorema 1.5.2 (Heine-Borel). *Un insieme $C \subset \mathbb{R}^n$ è chiuso e limitato se e solo se da ogni suo ricoprimento aperto si può estrarre un sottoricoprimento finito.*

Dimostrazione. Sia $C \subset \mathbb{R}^n$ chiuso e limitato, e per assurdo supponiamo che esista un suo ricoprimento aperto $(U_\alpha)_{\alpha \in I}$ che non ammette sottoricoprimenti finiti. Sia $Q \subset \mathbb{R}^n$ un n-cubo chiuso di lato l contenente C. Prendendo i punti medi dei lati di Q, lo si può suddividere in 2^n n-cubi chiusi di lato $l/2$, privi di punti interni in comune. Tra questi almeno uno, sia esso Q_1, sarà tale che l'insieme chiuso e limitato $C \cap Q_1$ non potrà essere ricoperto da un numero finito di aperti del ricoprimento $(U_\alpha)_{\alpha \in I}$ (e quindi in particolare $C \cap Q_1 \neq \emptyset$). Iterando questo ragionamento, si costruisce un successione

$$Q =: Q_0 \supset Q_1 \supset \cdots \supset Q_k \supset Q_{k+1} \supset \ldots$$

di cubi di lato $l/2^k$ tali che $C \cap Q_k \neq \emptyset$ non possa essere ricoperto da un numero finito di aperti di $(U_\alpha)_{\alpha \in I}$, per ogni $k \in \mathbb{N}$. Scelto allora $x_k \in C \cap Q_k$, $k \in \mathbb{N}$, per il teorema di Bolzano-Weierstrass la successione $(x_k)_{k \in \mathbb{N}} \subset C$ avrà un punto limite $x_0 \in \bigcap_{k=1}^{+\infty} C \cap Q_k$. Ma poiché $(U_\alpha)_{\alpha \in I}$ è un ricoprimento di C, esisterà $\bar{\alpha} \in I$ tale che $x_0 \in U_{\bar{\alpha}}$, e allora $C \cap Q_k \subset U_{\bar{\alpha}}$ per k sufficientemente grande, il che è assurdo per la costruzione degli Q_k.

Viceversa sia $C \subset \mathbb{R}^n$ tale che da ogni suo ricoprimento aperto si possa estrarre un sottoricoprimento finito. Se per assurdo C non fosse chiuso esisterebbe $x_0 \in \bar{C} \setminus C$, e posto allora $U_k := \mathbb{R}^n \setminus \bar{B}_{1/k}(x_0)$ si avrebbe $\bigcup_{k=1}^{+\infty} U_k = \mathbb{R}^n \setminus \{x_0\} \supset C$, ma, per ogni $m \in \mathbb{N}$, $\bigcup_{k=1}^{m} U_k = \mathbb{R}^n \setminus \bar{B}_{1/m}(x_0) \not\supset C$, cioè $(U_k)_{k \in \mathbb{N}}$ sarebbe un ricoprimento aperto di C privo di sottoricoprimenti finiti. Se invece C non fosse limitato chiaramente $(B_k(0))_{k \in \mathbb{N}}$ sarebbe un ricoprimento aperto di C privo di sottoricoprimenti finiti. □

Il teorema di Heine-Borel motiva la definizione seguente.

Definizione 1.5.3. Uno spazio topologico X è detto *compatto* se da ogni suo ricoprimento aperto si può estrarre un sottoricoprimento finito.

In particolare tale definizione si applica a un sottoinsieme C di uno spazio topologico X, quando su C si consideri la *topologia relativa*, i cui aperti sono le intersezioni con C degli aperti di X (è immediato verificare che questa è effettivamente una topologia su C).

Per convincersi che questa sia la "giusta" (nel senso di utile) nozione di compattezza per spazi topologici, bisogna verificare che implichi una qualche forma della proprietà di Bolzano-Weierstrass. Ricordando che per spazi topologici generali il ruolo delle successioni è svolto dai net, siamo quindi condotti a introdurre la seguente nozione di sottonet.

Definizione 1.5.4. Sia $(x_\alpha)_{\alpha \in I}$ un net nello spazio topologico X. Un suo *sottonet* è un net $(y_\beta)_{\beta \in J}$, per il quale esista un'applicazione $f : J \to I$ tale che $y_\beta = x_{f(\beta)}$, e per ogni $\alpha \in I$ esista $\beta_\alpha \in J$ tale che $\beta \geq \beta_\alpha$ implichi $f(\beta) \geq \alpha$.

Quindi gli elementi $f(\beta)$, $\beta \in J$, sono definitivamente più grandi di ogni $\alpha \in I$. Che questa sia la corretta generalizzazione del concetto di sottosuccessione è mostrato nell'esercizio 1.33.

Teorema 1.5.5 (Bolzano-Weierstrass generalizzato)**.** *Uno spazio topologico X è compatto se e solo se ogni suo net ammette un sottonet convergente.*

Dimostrazione. Sia X compatto, e dato un net $(x_\alpha)_{\alpha\in I} \subset X$ si definiscano gli insiemi aperti $U_\alpha := \overline{\{x_{\alpha'} : \alpha' \geq \alpha\}}^c$, $\alpha \in I$. Si ha allora che $(U_\alpha)_{\alpha\in I}$ non è un ricoprimento aperto di X: se lo fosse, per la compattezza di X esisterebbero $\alpha_1, \dots, \alpha_n \in I$ tali che $X = \bigcup_{i=1}^n U_{\alpha_i}$, il che equivale a $\bigcap_{i=1}^n \overline{\{x_{\alpha'} : \alpha' \geq \alpha_i\}} = \emptyset$, ma essendo I diretto si può prendere $\alpha \geq \alpha_1, \dots \alpha_n$, e dunque $x_\alpha \in \bigcap_{i=1}^n \overline{\{x_{\alpha'} : \alpha' \geq \alpha_i\}}$. Pertanto esiste $x \in (\bigcup_{\alpha\in I} U_\alpha)^c = \bigcap_{\alpha\in I} \overline{\{x_{\alpha'} : \alpha' \geq \alpha\}}$ il che, per la proposizione 1.1.11, implica che per ogni U intorno di x e per ogni $\alpha \in I$ esista $x_{\alpha'} \in U$ con $\alpha' \geq \alpha$, cioè che x sia un punto limite di (x_α), e dunque che esista un sottonet di (x_α) convergente a x (esercizio 1.33).

Viceversa sia X uno spazio topologico in cui ogni net ammetta un sottonet convergente, e, per assurdo, sia $(U_\alpha)_{\alpha\in I}$ un ricoprimento aperto di X privo di sottoricoprimenti finiti. Dunque per ogni $F \in \mathscr{P}_0(I)$ esisterà $x_F \in (\bigcup_{\alpha\in F} U_\alpha)^c$. Ricordando allora che $\mathscr{P}_0(I)$ è diretto per inclusione (esempio 1.4.3(c)), il net $(x_F)_{F\in\mathscr{P}_0(I)}$ avrà un punto limite $x \in X$, e sarà $x \in U_{\bar\alpha}$ per qualche $\bar\alpha \in I$. Dunque $U_{\bar\alpha}$ è un intorno di x, e pertanto dovrà esistere $F \geq \{\bar\alpha\}$ tale che $x_F \in U_{\bar\alpha} \subset \bigcup_{\alpha\in F} U_\alpha$, il che contraddice la costruzione di x_F. $\square$

Non è sorprendente, anche se di dimostrazione non immediata, il fatto che uno spazio metrico è compatto se e solo se da ogni sua *successione* si può estrarre una *sottosuccessione* convergente. (Si noti che il teorema precedente assicura soltanto che da una successione in uno spazio metrico compatto si può estrarre un sottonet convergente.)

Negli esercizi 1.34-1.38 sono enunciate diverse importanti proprietà degli spazi compatti e delle funzioni continue su di essi. In particolare l'esercizio 1.34 costituisce la generalizzazione del noto teorema di Weierstrass sulle funzioni reali continue su un sottoinsieme chiuso e limitato di $\mathbb{R}^n$.

Esercizi

1.1 Su $X := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$, la palla unitaria in $\mathbb{R}^2$, si definisca

$$d(x,y) := \begin{cases} |x-y| & \text{se } x \text{ e } y \text{ sono allineati con l'origine,} \\ |x|+|y| & \text{altrimenti} \end{cases} \qquad x, y \in X$$

($|\cdot|$ norma euclidea in $\mathbb{R}^2$). Dimostrare che (X, d) è uno spazio metrico.

1.2 Sia X un insieme di cardinalità n. Mostrare che $\mathscr{P}(X)$, l'insieme delle parti di X, ha cardinalità 2^n.

1.3 Siano X un insieme e B_α, $\alpha \in I$, una collezione di suoi sottoinsiemi (I insieme arbitrario di indici). Dimostrare la *dualità di De Morgan*:

$$\Big(\bigcap_{\alpha\in I} B_\alpha\Big)^c = \bigcup_{\alpha\in I} B_\alpha^c.$$

1.4 Sia X uno spazio topologico. Dimostrare che

(a) $\emptyset$, X sono chiusi;

(b) se C_α, $\alpha \in I$, sono chiusi, allora $\bigcap_{\alpha\in I} C_\alpha$ è chiuso;

(c) se $C_1, \dots, C_n$ sono chiusi, allora $C_1 \cup \dots \cup C_n$ è chiuso.

1.5 Siano (X,d) uno spazio metrico, $x \in X$ e $\delta > 0$. Dimostrare che le palle

$$B_\delta(x) := \{y \in X \,:\, d(x,y) < \delta\}, \qquad \bar{B}_\delta(x) := \{y \in X \,:\, d(x,y) \le \delta\}$$

sono, rispettivamente, aperta e chiusa (nella topologia indotta da d).

1.6 Siano (X,d) uno spazio metrico, e $x \in X$. Verificare che le famiglie di insiemi

$$\mathscr{B}_x^1 := \{B_\delta(x) \,:\, \delta > 0\}, \quad \mathscr{B}_x^2 := \{\bar{B}_\delta(x) \,:\, \delta > 0\}, \quad \mathscr{B}_x^3 := \{B_{\delta_n}(x) \,:\, n \in \mathbb{N}\}\ (\delta_n \to 0),$$

sono basi di intorni di x.

1.7 Verificare che la famiglia di sottoinsiemi di $\mathbb{R}$

$$\tau_\iota := \{(a, +\infty) \,:\, a \in \mathbb{R}\} \cup \{\emptyset, \mathbb{R}\}$$

definisce una topologia su $\mathbb{R}$ che non è di Hausdorff.

1.8 Siano (X,d) uno spazio metrico, $x \in X$ e $B \subset X$. Mostrare che x è un punto di accumulazione per B se e solo se esiste una successione $(x_j) \subset B \setminus \{x\}$ tale che $x_j \to x$.

1.9 Siano $[a,b] \subset \mathbb{R}$ un intervallo, e Δ l'insieme delle decomposizioni di $[a,b]$:

$$\delta = \{a = x_0 < x_1 < \dots x_n = b\}, \qquad n \in \mathbb{N}.$$

Si definisca su Δ una relazione $\le$ dicendo che $\delta_1 \le \delta_2$ se $\delta_1 \subset \delta_2$ (cioè δ_2 è ottenuta da δ_1 aggiungendo dei punti). Data inoltre $f : [a,b] \to \mathbb{R}$ limitata, si ponga

$$\sigma_\delta := \sum_{k=1}^n \Big(\inf_{[x_{k-1},x_k]} f\Big)(x_k - x_{k-1}), \qquad \Sigma_\delta := \sum_{k=1}^n \Big(\sup_{[x_{k-1},x_k]} f\Big)(x_k - x_{k-1}).$$

Dimostrare che:

(a) $(\Delta, \le)$ è un insieme parzialmente ordinato diretto;

(b) il net $(\sigma_\delta)_{\delta\in\Delta} \subset \mathbb{R}$ (risp. $(\Sigma_\delta)_{\delta\in\Delta}$) è non decrescente (risp. non crescente), cioè se $\delta_1 \le \delta_2$ allora $\sigma_{\delta_1} \le \sigma_{\delta_2}$ (risp. $\Sigma_{\delta_1} \ge \Sigma_{\delta_2}$);

(c) f è integrabile secondo Riemann se e solo se $\lim_\delta \sigma_\delta = \lim_\delta \Sigma_\delta = \int_a^b f$ nel senso dei net in $\mathbb{R}$ (con la topologia usuale). (Sugg.: f è integrabile se e solo se per ogni $\varepsilon > 0$ esiste δ tale che $\Sigma_\delta - \sigma_\delta < \varepsilon$.)

1.10 Siano (X, d) uno spazio metrico e $(x_\alpha)_{\alpha\in I}$ un net. Verificare che sono equivalenti

(a) $x_\alpha \to x$;

(b) $d(x_\alpha, x) \to 0$ come net in $\mathbb{R}$ (con la topologia usuale);

(c) per ogni $\varepsilon > 0$ esiste $\alpha_\varepsilon \in I$ tale che se $\alpha \geq \alpha_\varepsilon$ allora $d(x_\alpha, x) < \varepsilon$.

In particolare, se $I = \mathbb{N}$ si ritrova la nozione di convergenza di una successione in uno spazio metrico.

1.11 Siano X, Y spazi metrici, $D \subset X$, x_0 punto di accumulazione per D e $f : D \to Y$. Verificare che $\lim_{x\to x_0} f(x) = \ell$ se e solo se per ogni successione $(x_n) \subset D \setminus \{x_0\}$ tale che $x_n \to x$, si ha $f(x_n) \to \ell$.

1.12 Siano X, Y, Z spazi topologici, e $f : X \to Y$, $g : Y \to Z$ funzioni. Mostrare che se f è continua in $x_0 \in X$ e g è continua in $f(x_0) \in Y$ allora $g \circ f : X \to Z$ è continua in x_0.

1.13 Siano X, Y spazi topologici e $f : X \to Y$ funzione. Verificare che f è continua in X se e solo se per ogni chiuso $C \subset Y$, $f^{-1}(C)$ è un chiuso di X.

*1.14 Su $X = [0, 1]$ si definisca una topologia dichiarando che $C \subset [0, 1]$ è chiuso se e solo se è al più numerabile, o $C = X$. Dimostrare che $f : X \to \mathbb{R}$ è continua, se si dota l'immagine $\mathbb{R}$ della topologia usuale, se e solo se è costante.

1.15 Siano $X = C^1([a, b])$, $Y = C^0([a, b])$ e $D : X \to Y$ definito da $Df := f'$. Mostrare che:

(a) ponendo
$$\|f\|_{(1)} := \|f\|_\infty + \|f'\|_\infty, \qquad f \in C^1([a, b]),$$
si definisce una norma su X;

(b) se su X si mette la norma $\|\cdot\|_{(1)}$ e su Y la norma $\|\cdot\|_\infty$, D è limitato e $\|D\| = 1$ (sugg.: considerare le funzioni $f(t) = e^{\alpha t}$ e fare $\alpha \to +\infty$);

(c) se si mette la norma $\|\cdot\|_\infty$ sia su X che su Y, D non è limitato (sugg.: considerare le funzioni $f_n(t) := \frac{\sin nt}{n}$).

1.16 Sia $T : \mathbb{C}^n \to \mathbb{C}^n$, con $\mathbb{C}^n$ dotato della norma euclidea, e si indichi con $\sigma(T)$ lo spettro di T. Dimostrare:

(a) se $T = \mathrm{diag}[\lambda_1, \dots, \lambda_n]$ è diagonale, allora $\|T\| = \max_j |\lambda_j|$;

(b) se T è hermitiana ($T^* = T$), allora $\|T\| = \max_{\lambda\in\sigma(T)} |\lambda|$ (sugg.: se T è hermitiana, $T = UDU^*$ con D diagonale e $U^*U = UU^* = 1$);

(c) per T generica, $\|T\| = \max_{\lambda\in\sigma(T^*T)} |\lambda|^{1/2}$.

1.17 Dimostrare che su $X = C^1([a, b])$ la norma $\|f\|_{(1)} = \|f\|_\infty + \|f'\|_\infty$ è strettamente più forte di $\|\cdot\|_\infty$ (cioè esiste $C > 0$ tale che $\|f\|_\infty \leq C\|f\|_{(1)}$, ma non il viceversa).

1.18 Sia $(X, \|\cdot\|)$ normato. Verificare che $|\|x\| - \|y\|| \leq \|x - y\|$ per ogni $x, y \in X$.

1.19 Mostrare che su $X = \mathbb{C}^n$
$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty, \qquad x \in \mathbb{C}^n.$$

1.20 Siano $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ spazi normati, e si consideri $X \times Y$ con la struttura di spazio vettoriale somma diretta (algebrica):

$$\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2) := (\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2), \quad \alpha_i \in \mathbb{C}, x_i \in X, y_i \in Y.$$

Verificare che ponendo

$$\begin{aligned} \|(x,y)\|_\infty &:= \max\{\|x\|_X, \|y\|_Y\}, \\ \|(x,y)\|_1 &:= \|x\|_X + \|y\|_Y, \\ \|(x,y)\|_2 &:= \big(\|x\|_X^2 + \|y\|_Y^2\big)^{1/2}, \end{aligned}$$

si ottengono norme su $X \times Y$ che sono topologicamente equivalenti. Lo spazio $X \times Y$, quando dotato di una di queste norme, si indica anche con $X \oplus Y$ e viene chiamato *(spazio normato) somma diretta* degli spazi normati X e Y.

1.21 Siano X, Y spazi normati con X finito-dimensionale, e $T : X \to Y$ lineare. Dimostrare che T è limitato.

1.22 Siano (X, d) uno spazio metrico e $(x_n)_{n\in\mathbb{N}} \subset X$ convergente. Mostrare che (x_n) è di Cauchy.

1.23 Mostrare che uno spazio normato $(X, \|\cdot\|)$ con X finito-dimensionale è di Banach.

1.24 Verificare che $\mathbb{R}$, con la metrica $d(x,y) = |e^x - e^y|$, non è completo (sugg.: considerare $x_n = -n$.)

1.25 Sia $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ una successione, e si definiscano

$$\limsup_{n\to+\infty} a_n := \inf_{k\in\mathbb{N}} \sup_{n\geq k} a_n, \qquad \liminf_{n\to+\infty} a_n := \sup_{k\in\mathbb{N}} \inf_{n\geq k} a_n.$$

Si dimostri:

(a) $\liminf\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} \inf\limits_{k\geq n} a_k$, $\limsup\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} \sup\limits_{k\geq n} a_k$;

(b) $\liminf\limits_{n\to+\infty} a_n \leq \limsup\limits_{n\to+\infty} a_n$;

(c) $\liminf\limits_{n\to+\infty} a_n = l'$ (risp. $\limsup\limits_{n\to+\infty} a_n = l''$) se e solo se

$$\forall\,\varepsilon > 0 \quad \begin{cases} \exists\, n_\varepsilon \in \mathbb{N} \,:\, n \geq n_\varepsilon \Rightarrow & a_n > l' - \varepsilon \text{ (risp. } a_n < l'' + \varepsilon), \\ \forall \bar{n} \in \mathbb{N} \;\; \exists\, n \geq \bar{n} \,: & a_n < l' + \varepsilon \text{ (risp. } a_n > l'' - \varepsilon); \end{cases}$$

(d) $\exists \lim\limits_{n\to+\infty} a_n = l$ se e solo se $\limsup\limits_{n\to+\infty} a_n = \liminf\limits_{n\to+\infty} a_n = l$.

1.26 Verificare che $(\ell^1(I), \|\cdot\|_1)$, $(\ell^2(I), \|\cdot\|_2)$ sono spazi normati.

1.27 Sia $(x_\alpha)_{\alpha\in I} \subset \mathbb{R}_+$, I insieme di indici arbitrario. Mostrare che

$$\sum_{\alpha\in I} x_\alpha = \sup_{F\in\mathscr{P}_0(I)} \sum_{\alpha\in F} x_\alpha.$$

1.28 Mostrare che $(C^1([a,b]), \|\cdot\|_{(1)})$ è di Banach, mentre $(C^1([a,b]), \|\cdot\|_\infty)$ non lo è.

1.29 Sia (X, d) metrico. Verificare che $|d(x, y) - d(y, z)| \leq d(x, z)$ per ogni $x, y, z \in X$.

1.30 Siano (X, d) uno spazio metrico completo e $C \subset X$ chiuso. Mostrare che, indicando con d_C la restrizione di d a $C \times C$ (detta la *metrica indotta* da d su C), (C, d_C) è uno spazio metrico completo.

*1.31 Sia (X, d) uno spazio metrico. Dimostrare:

(a) la relazione definita sull'insieme delle successioni di Cauchy di X da

$$(x_n) \sim (y_n) \Leftrightarrow \lim_{n \to +\infty} d(x_n, y_n) = 0$$

è di equivalenza;

(b) indicata con $[(x_n)]$ la classe di equivalenza della successione di Cauchy $(x_n) \subset X$ secondo la relazione del punto (a), e con $\bar{X}$ l'insieme delle classi di equivalenza, l'applicazione $\bar{d} : \bar{X} \times \bar{X} \to \mathbb{R}_+$:

$$\bar{d}([(x_n)], [(y_n)]) := \lim_{n \to +\infty} d(x_n, y_n), \qquad [(x_n)], [(y_n)] \in \bar{X},$$

è ben definita, cioè il limite esiste e non dipende dai rappresentanti scelti per le classi, ed è una metrica su $\bar{X}$ (sugg.: per l'esistenza del limite, usare l'esercizio 1.29);

(c) indicata, per ogni $x \in X$, con $j(x) \in \bar{X}$ la classe di equivalenza della successione costante $x_n = x$, $n \in \mathbb{N}$, l'applicazione $j : X \to \bar{X}$ così definita è isometrica e $j(X)$ è denso in $\bar{X}$;

(d) $(\bar{X}, \bar{d})$ è uno spazio metrico completo, detto il *completamento* di (X, d) (sugg.: sia $(\bar{x}_n) \subset \bar{X}$ di Cauchy, e sia $(x_n) \subset X$ tale che $\bar{d}(j(x_n), \bar{x}_n) < 1/n$...);

(e) se $(\tilde{X}, \tilde{d})$ è uno spazio metrico completo e $k : X \to \tilde{X}$ è un'isometria con $k(X)$ denso in $\tilde{X}$, esiste un'isometria suriettiva $\phi : \bar{X} \to \tilde{X}$ tale che $\phi(j(x)) = k(x)$ per ogni $x \in X$ (*unicità del completamento*) (sugg.: si definisca prima $\phi : j(X) \to k(X)$ ponendo $\phi(j(x)) := k(x)$, allora ϕ è isometrica, poi $j(X)$ è denso in $\bar{X}$ e quindi...);

(f) sia $(X, \|\cdot\|)$ normato, e sul suo completamento $(\bar{X}, \bar{d})$ (considerando X come spazio metrico con la metrica indotta dalla norma), si ponga

$$\begin{aligned} \bar{x} + \bar{y} &:= \lim_{n \to +\infty} j(x_n + y_n), \\ \alpha \bar{x} &:= \lim_{n \to +\infty} j(\alpha x_n), \qquad\qquad \alpha \in \mathbb{C}, \bar{x}, \bar{y} \in \bar{X}; \\ \|\bar{x}\|^- &:= \lim_{n \to +\infty} \|x_n\|, \end{aligned}$$

dove $(x_n), (y_n) \subset X$ sono tali che $j(x_n) \to \bar{x}$, $j(y_n) \to \bar{y}$; allora le operazioni di spazio vettoriale e la norma su $\bar{X}$ sono ben defnite, $j : X \to \bar{X}$ è lineare, e $(\bar{X}, \|\cdot\|^-)$ è uno spazio di Banach la cui norma $\|\cdot\|^-$ è indotta da $\bar{d}$.

1.32 Siano X uno spazio normato e $V \subset X$ un sottospazio finito dimensionale. Verficare che V è chiuso.

1.33 Sia $(x_\alpha)_{\alpha \in I}$ un net in uno spazio topologico X. Un $x \in X$ è un *punto limite* per (x_α) se per ogni intorno U di x ed ogni $\alpha \in I$ esiste $\alpha' \geq \alpha$ tale che $x_{\alpha'} \in U$. Dimostrare:

(a) se esiste un sottonet $(y_\beta)_{\beta\in J}$ di (x_α) convergente a $x \in X$, allora x è un punto limite per (x_α);

(b) se $x \in X$ è un punto limite di (x_α), allora definendo un ordine parziale diretto su $J := \mathscr{I}_x \times I$ tramite

$$(U', \alpha') \geq (U, \alpha) \Leftrightarrow U' \subset U, \alpha' \geq \alpha,$$

e definendo $f : J \to I$ tramite $f(U, \alpha) := \alpha'$, dove $\alpha' \in I$ è tale che $\alpha' \geq \alpha$ e $x_{\alpha'} \in U$, si ha che $(x_{f(\beta)})_{\beta\in J}$ è un sottonet di (x_α) convergente a x.

In sostanza: (x_α) ammette un sottonet convergente a x se e solo se x è un suo punto limite.

1.34 Siano X, Y spazi topologici con X compatto, e $f : X \to Y$ continua. Mostrare che allora $f(X)$ è compatto in Y (con la topologia relativa).

1.35 Sia X compatto e $C \subset X$ chiuso. Mostrare che allora C è compatto (con la topologia relativa).

1.36 Sia X spazio di Hausdorff e $K \subset X$ compatto. Mostrare che allora K è chiuso.

1.37 Siano X compatto, Y di Hausdorff e $f : X \to Y$ continua e biunivoca. Mostrare che allora $f^{-1} : Y \to X$ è continua.

1.38 Sia X spazio normato di dimensione infinita, e $K \subset X$ compatto. Mostrare che K ha interno vuoto.

Capitolo 2

Richiami di teoria dell'integrazione alla Lebesgue

Il difetto principale della teoria di integrazione di Riemann è che non permette di effettuare il passaggio al limite sotto il segno di integrale con sufficiente generalità. In particolare, non garantisce che se una successione (limitata) di funzioni continue (e quindi integrabili) converge puntualmente, allora la funzione limite è integrable e il suo integrale è il limite degli integrali delle funzioni della successione data, cosa che sarebbe assai naturale aspettarsi. Questo problema è alla radice del fatto, visto nell'esempio 1.4.3(d) del capitolo precedente, che gli spazi di funzioni continue non sono completi nelle norme $\|\cdot\|_p$ ($p = 1, 2$) il che, nelle applicazioni concrete, è piuttosto sgradevole, come si capirà meglio in seguito.

La necessità di disporre di una teoria dell'integrazione che non presentasse tali svantaggi è stato il principale motivo che ha indotto Lebesgue a cercare di generalizzare la teoria di Riemann. L'idea centrale di Lebesgue è stata quella di definire l'integrale come limite di somme ottenute suddividendo il codominio della funzione da integrare, invece che il suo dominio come nella definizione di Riemann. Poiché la controimmagine di un intervallo tramite una funzione continua non è necessariamente un intervallo, per poter procedere in questo modo è necessario preliminarmente definire la "lunghezza", o "misura" di insiemi sufficientemente generali. Questo conduce a una *teoria della misura* per sottoinsiemi di $\mathbb{R}$ (o di $\mathbb{R}^n$), la cui formulazione è però tale da consentire una naturale e immediata generalizzazione a spazi metrici o topologici, o a insiemi ancora più generali.

In questo capitolo esporremo dunque i principali risultati delle teoria della misura (sez. 2.1) e dell'integrazione (sez. 2.2) alla Lebesgue, per lo più limitandoci, per brevità, ai soli enunciati, e alle dimostrazioni di quelli riguardanti gli spazi di funzioni integrabili (sez. 2.3), che sono di maggiore interesse dal punto di vista dell'Analisi Funzionale. Il capitolo è quindi inteso più che altro come un modo di stabilire un linguaggio e delle notazioni, e come un riferimento rapido a risultati utilizzati nella parte principale del corso per gli studenti che conoscano già gli aspetti fondamentali della teoria. Non bisogna pertanto aspettarsi di poterne comprendere a fondo le idee e le tecniche leggendo solo quanto segue. Chi volesse appronfondire questi temi è invitato a consultare i molti ottimi testi sull'argomento, a cominciare da [Rud1], su cui è basata essenzialmente la nostra esposizione.

2.1 Teoria della misura

Dovendo essere una generalizzazione delle nozioni di lunghezza, area o volume di semplici sottoinsiemi di $\mathbb{R}^n$ ($n = 1, 2, 3$), quali intervalli, rettangoli ecc., una misura sarà, a grandi linee, una applicazione a valori non negativi definita su certi sottoinsiemi di un insieme X dato. Cominciamo dunque con l'introdurre opportune famiglie di sottoinisiemi di X, che serviranno da domini per le misure che considereremo. Per il momento non supporremo alcuna struttura sull'insieme X, che sarà dunque assolutamente arbitrario.

Definizione 2.1.1. Sia X un insieme. Una famiglia $\mathfrak{R} \subset \mathscr{P}(X)$ è detta un *anello (di insiemi)* se

(i) $A, B \in \mathfrak{R} \Rightarrow A \cup B \in \mathfrak{R}$;

(ii) $A, B \in \mathfrak{R} \Rightarrow A \setminus B \in \mathfrak{R}$.

Un anello $\mathfrak{R}$ è detto un'*algebra (di insiemi)* se $X \in \mathfrak{R}$.

Un'algebra $\mathfrak{M} \subset \mathscr{P}(X)$ è detta una *σ-algebra* se per ogni successione $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{M}$ si ha $\bigcup_{k\in\mathbb{N}} A_k \in \mathfrak{M}$.

Equivalentemente, $\mathfrak{M} \subset \mathscr{P}(X)$ è una σ-algebra se vale: $X \in \mathfrak{M}$; $A \in \mathfrak{M} \Rightarrow A^c \in \mathfrak{M}$; $(A_k) \subset \mathfrak{M} \Rightarrow \bigcup_k A_k \in \mathfrak{M}$.

Precisiamo ora la nozione di misura accennata sopra.

Definizione 2.1.2. Sia X un insieme.

(i) Una *misura su un anello* $\mathfrak{R} \subset \mathscr{P}(X)$ è un'applicazione $\mu : \mathfrak{R} \to [0, +\infty]$ tale che

$$A, B \in \mathfrak{R},\ A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B) \quad \text{(additività di } \mu\text{)}.$$

(ii) Una *misura su una σ-algebra* $\mathfrak{M} \subset \mathscr{P}(X)$ è un'applicazione $\mu : \mathfrak{M} \to [0, +\infty]$ tale che

$$(A_k)_{k\in\mathbb{N}} \subset \mathfrak{M},\ A_k \cap A_j = \emptyset \text{ se } k \neq j \Rightarrow \mu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = \sum_{k=0}^{+\infty} \mu(A_k)$$

(σ-additività di μ).

Come vedremo nella prossima sezione, allo scopo di definire l'integrale di Lebesgue è necessario disporre di una misura su una σ-algebra. L'introduzione della nozione di misura additiva su un anello è dovuta al fatto che in molti casi interessanti, mentre è complicato descrivere esplicitamente una σ-algebra rilevante e definire direttamente una misura σ-additiva su di essa, è relativamente facile ottenere una misura su un anello opportuno. Quest'ultima si estende poi a una misura σ-additiva utilizzando il teorema di estensione di Lebesgue 2.1.8 di seguito. In particolare questo è il procedimento che, come vedremo subito, si usa per definire la misura di Lebesgue in $\mathbb{R}^n$ a partire dal "volume" dei parallelepipedi n-dimensionali.

Definizione 2.1.3. Un *intervallo* $I \subset \mathbb{R}^n$ è un insieme della forma $I = J_1 \times \cdots \times J_n$, con $J_i \subset \mathbb{R}$ intervallo limitato (aperto, semiaperto o chiuso, eventualmente vuoto), $i = 1, \ldots, n$. La sua misura è

$$m(I) := \prod_{i=1}^{n} |J_i|,$$

dove $|J_i|$ è la lunghezza dell'intervallo J_i.

Un *insieme elementare, o plurintervallo*, $E \subset \mathbb{R}^n$ è un insieme della forma $E = \bigcup_{k=1}^{r} I_k$, con $I_k \subset \mathbb{R}^n$ intervallo, $k = 1, \ldots, r$.

La dimostrazione della proposizione seguente, che non daremo, è probabilmente già nota dai corsi di Calcolo, in quanto alla base della teoria dell'integrazione di Riemann in più dimensioni, ed è comunque di facile intuizione almeno nei casi $n = 1, 2, 3$, in cui basta fare qualche disegno.

Proposizione 2.1.4. *(i) La famiglia $\mathfrak{E} \subset \mathscr{P}(\mathbb{R}^n)$ degli insiemi elementari è un anello.*

(ii) Per ogni $E \in \mathfrak{E}$ esistono $I_1, \ldots, I_r \subset \mathbb{R}^n$ intervalli disgiunti tali che $E = \bigcup_{k=1}^{r} I_k$.

(iii) Dato $E \in \mathfrak{E}$ e intervalli disgiunti $I_1, \ldots, I_r \subset \mathbb{R}^n$ tali che $E = \bigcup_{k=1}^{r} I_k$, il numero

$$m(E) := \sum_{k=1}^{r} m(I_k)$$

è indipendente dalla decomposizione di E in intervalli disgiunti, ed è detto la misura *di E.*

(iv) m è additiva su $\mathfrak{E}$ (e quindi è effettivamente una misura).

(v) m è regolare*, cioè per ogni $E \in \mathfrak{E}$ e $\varepsilon > 0$ esistono $A, C \in \mathfrak{E}$, A aperto e C chiuso, tali che*

$$C \subset E \subset A, \qquad m(A) - \varepsilon < m(E) < m(C) + \varepsilon.$$

Osservazione 2.1.5. Data una funzione monotona non decrescente e continua a destra $F : \mathbb{R} \to \mathbb{R}$, per ogni $x_0 \in \mathbb{R}$ esiste finito il limite $F(x_0^-) := \lim_{x \to x_0^-} F(x)$, e si può dunque definire un'applicazione non negativa m_F sull'insieme degli intervalli (limitati) di $\mathbb{R}$ tramite

$$\begin{aligned} m_F((a,b]) &:= F(b) - F(a), \\ m_F([a,b)) &:= F(b^-) - F(a^-), \\ m_F((a,b)) &:= F(b^-) - F(a), \\ m_F([a,b]) &:= F(b) - F(a^-). \end{aligned}$$

Nel caso particolare $F(\lambda) = \lambda$, $m_F(I)$ è proprio la lunghezza dell'intervallo I. Si vede allora facilmente che, analogamente alla m della proposizione precedente, m_F si estende a una misura regolare sugli insiemi elementari di $\mathbb{R}$.

La proposizione 2.1.4 rappresenta il primo passo – la definizione di una misura su un anello – nella strategia di estensione delineata sopra per ottenere la misura di Lebesgue. Per eseguire il secondo passo abbiamo bisogno ancora di alcune definizioni.

Definizione 2.1.6 (misura esterna di Lebesgue)**.** Sia μ una misura regolare sugli insiemi elementari $\mathfrak{E} \subset \mathscr{P}(\mathbb{R}^n)$ che sia finita (cioè $\mu(E) < +\infty$ per ogni $E \in \mathfrak{E}$). La *misura esterna* μ^* associata a μ è un'applicazione $\mu^* : \mathscr{P}(\mathbb{R}^n) \to [0, +\infty]$ definita da

$$\mu^*(A) := \inf\Big\{ \sum_{k=1}^{+\infty} \mu(E_k) \;:\; A \subset \bigcup_{k=1}^{+\infty} E_k,\, E_k \in \mathfrak{E} \text{ aperti}\Big\}, \qquad A \subset \mathbb{R}^n.$$

Dati insiemi $A, B \subset \mathbb{R}^n$, ricordiamo che la loro *differenza simmetrica* è l'insieme $A\Delta B := (A \setminus B) \cup (B \setminus A)$.

Definizione 2.1.7 (insiemi misurabili secondo Lebesgue)**.** Sia μ una misura regolare finita sugli insiemi elementari $\mathfrak{E} \subset \mathscr{P}(\mathbb{R}^n)$. Posto

$$\mathfrak{M}_F(\mu) := \{A \subset \mathbb{R}^n \;:\; \forall \varepsilon > 0,\, \exists E \in \mathfrak{E} \;:\; \mu^*(A\Delta E) < \varepsilon\},$$

la famiglia degli *insiemi misurabili secondo Lebesgue* è definita da

$$\mathfrak{M}(\mu) := \Big\{A \subset \mathbb{R}^n \;:\; \exists (A_k)_{k\in\mathbb{N}} \subset \mathfrak{M}_F(\mu),\, A = \bigcup_{k\in\mathbb{N}} A_k\Big\}.$$

Teorema 2.1.8 (estensione di Lebesgue)**.** *La famiglia $\mathfrak{M}(\mu)$ è una σ-algebra, e $\mu^*|_{\mathfrak{M}(\mu)}$ è l'unica misura σ-additiva che estende μ a $\mathfrak{M}(\mu)$. Tale misura si denota ancora con μ.*

La dimostrazione del teorema, che omettiamo, consiste essenzialmente nel verificare che $d(A, B) := \mu^*(A\Delta B)$ è una metrica su (classi di equivalenza di elementi di) $\mathfrak{E}$, e che $\mathfrak{M}(\mu)$ è il completamento di $\mathfrak{E}$ rispetto a essa.

Applicando il teorema di estensione 2.1.8 alla misura m su $\mathfrak{E}$ fornita dalla proposizione 2.1.4 si ottiene la *misura di Lebesgue* m su $\mathbb{R}^n$, e applicandolo alle misure m_F dell'osservazione 2.1.5 si ottengono le *misure di Lebesgue-Stieltjes* su $\mathbb{R}$, che interverranno in seguito nella discussione del teorema spettrale per operatori autoaggiunti.

Allo scopo di descrivere un po' più in dettaglio gli insiemi misurabili di quanto non faccia la definizione, introduciamo una classe notevole di σ-algebre in spazi topologici.

Definizione 2.1.9. Sia X uno spazio topologico. La σ-algebra $\mathfrak{B}(X)$ degli *insiemi boreliani* di X è la più piccola σ-algebra su X contenente tutti gli aperti, cioè l'intersezione di tutte le σ-algebre su X che contengono tutti gli aperti.

La verifica delle affermazioni della precedente definizione è lasciata al lettore (esercizio 2.1).

La seguente proposizione segue facilmente dalla proprietà di regolarità di una misura sugli insiemi elementari.

Proposizione 2.1.10. *Sia μ una misura regolare finita su $\mathfrak{E} \subset \mathscr{P}(\mathbb{R}^n)$. Si ha $\mathfrak{B}(\mathbb{R}^n) \subset \mathfrak{M}(\mu)$ e per ogni $A \in \mathfrak{M}(\mu)$ esistono $F, G \in \mathfrak{B}(\mathbb{R}^n)$, con F unione numerabile di chiusi e G intersezione numerabile di aperti, tali che*

$$F \subset A \subset G, \qquad \mu(A \setminus F) = 0 = \mu(G \setminus A).$$

Insiemi del tipo degli insiemi F e G della proposizione precedente vengono detti insiemi F_σ e G_δ rispettivamente. Dunque un insieme misurabile secondo Lebesgue differisce da un F_σ e da un G_δ per un insieme di misura nulla. In particolare, tutti gli aperti e tutti i chiusi (nonché tutti gli F_σ e gli G_δ, tutte le intersezioni numerabili di F_σ e le unioni numerabili di G_δ, ecc.) sono misurabili.

L'esercizio 2.2 fornisce esempi di insiemi misurabili secondo Lebesgue, ma non secondo Peano-Jordan, con interno vuoto e misura positiva, o densi (in un intervallo) e con misura arbitrariamente piccola. (Ricordiamo che $A \subset \mathbb{R}^n$ è misurabile secondo Peano-Jordan se posto

$$m_i(A) := \sup\{m(E) \,:\, E \in \mathfrak{E},\, E \subset A\}, \quad m_e(A) := \inf\{m(E) \,:\, E \in \mathfrak{E},\, E \supset A\},$$

si ha $m_i(A) = m_e(A)$.)

D'altra parte esistono insiemi non misurabili secondo Lebesgue, ma, come ci si può facilmente aspettare dalle osservazioni appena fatte, tali insiemi non sono "costruibili" utilizzando le normali procedure della teoria degli insiemi, e la dimostrazione della loro esistenza richiede l'utilizzo dell'assioma della scelta.

Concludiamo la sezione menzionando che una *misura di Borel regolare* su uno spazio di Hausdorff X è una misura μ su $\mathfrak{B}(X)$ tale che per ogni $B \in \mathfrak{B}(X)$ e ogni $\varepsilon > 0$ esistano un compatto $C \subset B$ e un aperto $A \supset B$ tali che

$$\mu(B \setminus C) < \varepsilon, \qquad \mu(A \setminus B) < \varepsilon.$$

Da questo segue facilmente che μ soddisfa le analoghe delle proprietà nella tesi della proposizione 2.1.10.

2.2 Teoria dell'integrazione

La teoria di Lebesgue, richiamata nella sezione precedente, permette di dare un senso alla misura di sottoinsiemi molto generali di $\mathbb{R}^n$. Ci si può dunque aspettare che si possa utilizzare per ottenere una teoria dell'integrazione molto più flessibile di quella di Riemann, basata essenzialmente sugli insiemi elementari. Esporremo ora i risultati fondamentali di tale teoria. Poiché questi dipendono in realtà soltanto dalle proprietà formali di una misura (cioè essenzialmente dalla σ-additività), procederemo in modo astratto, assumendo che sia data una misura su un insieme privo di altre strutture. Otterremo così una teoria che copre, oltre alla misura di Lebesgue, molti altri casi interessanti e che ci saranno utili in seguito.

Cominciamo anche qui con alcune definizioni.

Definizione 2.2.1. Uno *spazio di misura* è una terna $(X, \mathfrak{M}, \mu)$, dove X è un insieme, $\mathfrak{M} \subset \mathscr{P}(X)$ una σ-algebra e μ una misura definita su $\mathfrak{M}$. Gli insiemi $A \in \mathfrak{M}$ sono detti *insiemi misurabili.*

Esempi 2.2.2. (a) Se m è la misura di Lebesgue su $\mathbb{R}^n$ e $\mathfrak{M}(m)$ è la σ-algebra degli insiemi misurabili secondo Lebesgue, teorema 2.1.8, $(\mathbb{R}^n, \mathfrak{M}(m), m)$ è uno spazio di misura. Analogamente per le misure di Lebesgue-Stieltjes m_F.

(b) Sia X un insieme. Dato $x \in X$ si definisca $\delta_x : \mathscr{P}(X) \to [0,1]$ tramite

$$\delta_x(A) := \begin{cases} 1 & \text{se } x \in A, \\ 0 & \text{se } x \notin A, \end{cases} \qquad A \subset X.$$

È chiaro che δ_x è una misura su X, detta *misura di Dirac (concentrata in x)*. Dunque $(X, \mathscr{P}(X), \delta_x)$ è uno spazio di misura. Nel caso $X = \mathbb{R}$, si ha che δ_x è la misura di Lebesgue-Stieltjes associata alla funzione (non decrescente e continua a destra) $F = \chi_{[x,+\infty)}$ (esercizio 2.3).

(c) Sia X un insieme. Si definisca $\mu_\sharp : \mathscr{P}(X) \to [0,+\infty]$ tramite

$$\mu_\sharp(A) := \begin{cases} \sharp A & \text{se } A \text{ è finito}, \\ +\infty & \text{se } A \text{ è infinito}, \end{cases} \qquad A \subset X.$$

Anche in questo caso si verifica subito che $\mu_\sharp$ è una misura, detta la *misura che conta* in X, e pertanto $(X, \mathscr{P}(X), \mu_\sharp)$ è uno spazio di misura.

In tutto il resto di questa sezione, salvo diverso avviso, $(X, \mathfrak{M}, \mu)$ denoterà uno spazio di misura arbitrario.

Definizione 2.2.3. Una funzione $f : X \to Y$, Y spazio topologico, è detta *misurabile* se $f^{-1}(B) = \{x \in X \,:\, f(x) \in B\} \in \mathfrak{M}$ per ogni $B \in \mathfrak{B}(Y)$.

Si noti l'analogia tra la nozione di funzione misurabile e quella di funzione continua tra spazi topologici, nella formulazione del teorema 1.2.14. Si osservi anche che la misurabilità di una funzione non dipende realmente dalla misura μ, ma solo dalla σ-algebra $\mathfrak{M}$ su cui essa è definita.

I casi per noi di gran lunga più interessanti sono quelli in cui $Y = \mathbb{R}, \mathbb{C}$, dotati ovviamente della topologia usuale, o $Y = \tilde{\mathbb{R}} := [-\infty, \infty]$ dotato della topologia i cui aperti sono unioni arbitarie di intervalli della forma (a,b), $(a,+\infty]$ o $[-\infty, b)$, $a, b \in \mathbb{R}$.

Introduciamo alcune notazioni brevi molto utili quando si ha a che fare con funzioni misurabili. Data una funzione $f : X \to Y$ e $B \subset Y$ si porrà $\{f \in B\} := f^{-1}(B)$, e in particolare se $Y = \tilde{\mathbb{R}}$ scriveremo $\{f > a\}$ al posto di $f^{-1}((a,+\infty])$ (e analogamente per $\{f < a\}$, $\{f \geq a\}$, $\{f \leq a\}$).

Le principali proprietà delle funzioni misurabili sono raccolte nel teorema seguente. La dimostrazione di alcune di esse è lasciata al lettore (esercizio 2.4).

Teorema 2.2.4. *Valgono le seguenti affermazioni:*

(i) $f : X \to Y$ *è misurabile se e solo se* $f^{-1}(A) \in \mathfrak{M}$ *per ogni aperto* $A \subset Y$*;*

(ii) $f : X \to \tilde{\mathbb{R}}$ *è misurabile* $\Leftrightarrow \{f > a\} \in \mathfrak{M}$ *per ogni* $a \in \mathbb{R} \Leftrightarrow \{f \geq a\} \in \mathfrak{M}$ *per ogni* $a \in \mathbb{R} \Leftrightarrow \{f < a\} \in \mathfrak{M}$ *per ogni* $a \in \mathbb{R} \Leftrightarrow \{f \leq a\} \in \mathfrak{M}$ *per ogni* $a \in \mathbb{R}$*;*

(iii) se $f_n : X \to \tilde{\mathbb{R}}$ *è misurabile per ogni* $n \in \mathbb{N}$*, allora* $\sup_n f_n$*,* $\inf_n f_n$ *sono misurabili;*

(iv) *se* Y*,* Z *sono spazi topologici, e* $f : X \to Y$ *è misurabile e* $g : Y \to Z$ *è continua,* $g \circ f : X \to Z$ *è misurabile;*

(v) *se* $f_1, f_2 : X \to \mathbb{R}$ *sono misurabili, allora* $f := (f_1, f_2) : X \to \mathbb{R}^2$ *è misurabile;*

(vi) $f : X \to \mathbb{C}$ *è misurabile se e solo se* $\operatorname{Re} f, \operatorname{Im} f : X \to \mathbb{R}$ *sono misurabili;*

(vii) *se* $f, g : X \to \mathbb{C}$ *sono misurabili, allora* $f + g, f \cdot g : X \to \mathbb{C}$ *sono misurabili.*

(viii) *se* $f_n : X \to \tilde{\mathbb{R}}$ *(o* $f_n : X \to \mathbb{C}$*) è misurabile per ogni* $n \in \mathbb{N}$*, allora* $E := \{x \in X \,:\, \exists f(x) := \lim_{n \to +\infty} f(x)\} \in \mathfrak{M}$ *e* $f : E \to \tilde{\mathbb{R}}$ *(*$f : E \to \mathbb{C}$*) è misurabile (rispetto alla* σ*-algebra* $\mathfrak{M}_E := \{F \cap E \,:\, F \in \mathfrak{M}\}$ *su* E*);*

(ix) *se* $f : X \to \tilde{\mathbb{R}}$ *(o* $f : X \to \mathbb{C}$*) è misurabile anche* $|f|$ *lo è;*

La (i) implica in particolare che se X è uno spazio topologico e $\mathfrak{M} = \mathfrak{B}(X)$, le $f : X \to Y$ continue sono misurabili. Inoltre, grazie alla (ii), per verificare che una $f : X \to \tilde{\mathbb{R}}$ è misurabile, basta verificare la proprietà $f^{-1}(B) \in \mathfrak{M}$ per una famiglia di insiemi $B \subset \tilde{\mathbb{R}}$ molto più piccola di quella di tutti i boreliani. È anche possibile generalizzare (viii) al caso di successioni di funzioni a valori in un generico spazio topologico (esercizio 2.5).

Da questa proposizione si vede inoltre che, come nel caso degli insiemi, tutte le operazioni fondamentali dell'Analisi applicate a funzioni misurabili danno luogo a funzioni misurabili. Osserviamo però esplicitamente che le proprietà (iii) e (viii) non si estendono al caso dei net di funzioni misurabili (com'è facile attendersi a partire dal fatto che una σ-algebra è stabile per operazioni su famiglie soltanto numerabili di insiemi).

Introduciamo ora una classe particolare di funzioni misurabili, che giocano, nella teoria di Lebesgue, il ruolo che le funzioni a gradino giocano in quella di Riemann.

Definizione 2.2.5. Una funzione misurabile $s : X \to \mathbb{C}$ è detta una *funzione semplice* se l'insieme dei suoi valori $s(X) \subset \mathbb{C}$ è finito.

Equivalentemente $s : X \to \mathbb{C}$ è misurabile semplice se è della forma

$$s = \sum_{k=1}^{n} c_k \chi_{E_k},$$

dove $\{c_1, \dots, c_n\} = s(X)$ sono i valori assunti da s, $E_k := \{s = c_k\} \in \mathfrak{M}$ è l'insieme su cui s ha valore c_k, e χ_{E_k} è la funzione caratteristica dell'insieme E_k (cioè la funzione che vale 1 su E_k e 0 altrove; si noti che χ_E è misurabile se e solo se $E \in \mathfrak{M}$). In tal caso, si ha chiaramente che $E_h \cap E_k = \emptyset$ se $h \neq k$, e $X = \bigcup_{k=1}^{n} E_k$, cioè gli insiemi E_k, $k = 1, \dots, n$, formano una *partizione misurabile* di X. È poi chiaro che somme e prodotti di funzioni semplici sono semplici.

Il risultato seguente è il fondamento di tutta la teoria dell'integrazione di Lebesgue.

Teorema 2.2.6. *Sia* $f : X \to \mathbb{C}$ *misurabile. Esiste una successione* $(s_n)_{n \in \mathbb{N}}$ *di funzioni misurabili semplici che converge puntualmente a* f *su* X*.*

Cenno di dimostrazione. Sia $f \geq 0$. Definendo gli insiemi misurabili

$$E_{n,k}. := \left\{ \frac{k-1}{2^n} \leq f < \frac{k}{2^n} \right\}, \quad F_n := \{f \geq n\}, \qquad n \in \mathbb{N},\, k = 1, \dots, n2^n,$$

e ponendo

$$s_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{n,k}} + n\chi_{F_n}, \qquad n \in \mathbb{N},$$

si ottiene una successione $(s_n)_{n\in\mathbb{N}}$ di funzioni misurabili semplici tali che $s_1(x) \leq s_2(x) \leq \dots \leq f(x)$ e $\lim_{n\to+\infty} s_n(x) = f(x)$ per ogni $x \in X$.

Nel caso generale, posto $f_1 := \operatorname{Re} f$, $f_2 := \operatorname{Im} f$ si considerino le funzioni misurabili non negative $f_i^+ := \max\{f_i, 0\}$, $f_i^- := -\min\{f_i, 0\}$ (parte positiva e negativa di f_i rispettivamente), $i = 1, 2$. Si ha allora $f = (f_1^+ - f_1^-) + i(f_2^+ - f_2^-)$, e definite le funzioni misurabili semplici $s_{i,n}^\pm$ a partire da $f_i^\pm$ come sopra, e posto $s_n := (s_{1,n}^+ - s_{1,n}^-) + i(s_{2,n}^+ - s_{2,n}^-)$, ne segue che s_n è semplice e $s_n(x) \to f(x)$ per ogni $x \in X$. □

Come si vede, l'idea della dimostrazione precedente è quella di partizionare il codominio di $f \geq 0$ in intervalli di ampiezza $1/2^n$ e di approssimare quindi f con una funzione costante sulla loro controimmagine tramite f. Come già notato all'inizio del capitolo, affinché tale ribaltamento di prospettiva rispetto a quella di Riemann sia sensato e utile è necessario che gli insiemi risultanti, che possono essere assai complicati anche se f è molto regolare, siano misurabili per una classe di funzioni f sufficientemente ampia, il che è garantito dalla notevole generalità della nozione di misurabilità di Lebesgue.

A questo punto siamo nella condizione di dare la definizione di integrale, cominciando da quello delle funzioni semplici, e poi usando questo per definire, per approssimazione, quello di funzioni via via più generali.

Definizione 2.2.7. (a) L'integrale di una funzione misurabile semplice $s = \sum_{k=1}^n c_k \chi_{E_k}$ è

$$\int_X s\, d\mu := \sum_{k=1}^n c_k \mu(E_k).$$

(b) L'integrale di una funzione misurabile non negativa $f : X \to [0, +\infty]$ è il numero in $[0, +\infty]$ definito da

$$\int_X f\, d\mu := \sup\left\{ \int_X s\, d\mu \;:\; s \text{ semplice, } 0 \leq s \leq f \right\}.$$

(c) L'integrale di una funzione misurabile $f : X \to \tilde{\mathbb{R}}$ è il numero in $\tilde{\mathbb{R}}$ definito da

$$\int_X f\, d\mu := \int_X f^+\, d\mu - \int_X f^-\, d\mu,$$

a patto che almeno uno dei due integrali a secondo membro sia finito.

(d) Una funzione misurabile $f : X \to \mathbb{C}$ è detta *integrabile (secondo Lebesgue)* se $\int_X |f| d\mu < +\infty$, e il suo integrale è il numero complesso definito da

$$\int_X f\, d\mu := \int_X \mathrm{Re}\, f\, d\mu + i \int_X \mathrm{Im}\, f\, d\mu.$$

L'insieme delle funzioni integrabili è denotato $L^1(X, \mu)$ (o $L^1(X)$, o semplicemente L^1, quando non ci sia rischio di confusione).

Notiamo che se $f \in L^1(X, \mu)$, $f = f_1 + i f_2$ (f_1, f_2 reali), si ha $f_i^\pm \leq |f_i| \leq |f|$, $i = 1, 2$, e dunque, per il punto (ii) della proposizione 2.2.9 seguente, $\int_X f_i^\pm d\mu \leq \int_X |f| d\mu < +\infty$, per cui la definizione di $\int_X f d\mu$ data al punto (d) è sensata.

Dato un insieme $E \in \mathfrak{M}$, si porrà anche $\int_E f\, d\mu := \int_X \chi_E f\, d\mu$. Quando non ci sia ambiguità, si potrà utilizzare la notazione semplificata $\int_X f$ al posto di $\int_X f d\mu$, mentre quando sia necessario specificare l'espressione di f si utilizzerà la notazione $\int_X f(x)\, d\mu(x)$. In particolare se $E \subset \mathbb{R}^n$ è misurabile secondo Lebesgue si scriverà $\int_E f(x)\, d^n x$ invece che $\int_E f(x)\, dm(x)$ per indicare l'integrale fatto rispetto alla misura di Lebesgue m su $\mathbb{R}^n$, e nel caso $n = 1$ e $E = [a, b]$ (o $(a, b), (a, b], [a, b)$) si userà la notazione classica $\int_a^b f(x)\, dx$ (o solo $\int_a^b f$).

Esempio 2.2.8. Su $X = [0, 1]$, con la misura di Lebesgue m, la funzione $f = \chi_\mathbb{Q}$ è misurabile, in quanto funzione caratteristica di $\mathbb{Q} \cap [0, 1]$ che è unione numerabile di insiemi costituiti da un solo punto, che sono quindi chiusi. Inoltre f è semplice (è una funzione caratteristica) e il suo integrale è dunque

$$\int_0^1 \chi_\mathbb{Q} = 0 \cdot m([0, 1] \setminus \mathbb{Q}) + 1 \cdot m([0, 1] \cap \mathbb{Q}) = 0,$$

poiché $m(\{x\}) = 0$ per ogni $x \in [0, 1]$ ($\{x\} = [x, x]$ è un intervallo) e m è σ-additiva. Ricordiamo che f è il tipico esempio di funzione non integrabile secondo Riemann.

È chiaro che gli insiemi di misura nulla sono in qualche modo trascurabili nella teoria dell'integrazione: ad esempio se $f, g \geq 0$ sono funzioni misurabili che differiscono soltanto su un insieme di misura nulla si vede facilmente dalla definizione che $\int_X f = \int_X g$. È allora conveniente introdurre la seguente terminologia: si dice che una proprietà $P(x)$, dipendente da $x \in X$, è vera *quasi ovunque (rispetto a μ)* (abbreviato q.o.), se l'insieme $\{x \in X : P(x)$ è falsa$\}$ è di misura nulla.

Le principali proprietà dell'integrale di Lebesgue sono enunciate nella proposizione seguente. Omettiamo o lasciamo al lettore (esercizio 2.6) la loro dimostrazione.

Proposizione 2.2.9. *Siano f, g funzioni misurabili su X (a valori in $\tilde{\mathbb{R}}$ o $\mathbb{C}$). Si ha:*

(i) se $f, g \in L^1$ e $\alpha, \beta \in \mathbb{C}$, allora $\alpha f + \beta g \in L^1$ e $\int_X (\alpha f + \beta g) = \alpha \int_X f + \beta \int_X g$ (linearità dell'integrale);

(ii) se $f \leq g$, allora $\int_X f \leq \int_X g$ (monotonia dell'integrale);

(iii) $|\int_X f| \leq \int_X |f|$;

(iv) se $f \geq 0$ e $\int_X f\, d\mu = 0$, allora $f = 0$ q.o.;

(v) *se* $f \in L^1(X,\mu)$ *e* $\int_E f\,d\mu = 0$ *per ogni* $E \in \mathfrak{M}$*, allora* $f = 0$ *q.o.;*

(vi) *se* $\int_X |f| d\mu < +\infty$*, allora* $|f| < +\infty$ *q.o.*

Enunciamo ora, senza dimostrarli, i fondamentali teoremi di passaggio al limite sotto il segno di integrale, che garantiscono che le debolezze della teoria di Riemann, ricordate all'inizio del capitolo, sono superate.

Teorema 2.2.10 (convergenza monotona di Levi)**.** *Siano* $f_n, f : X \to [0,+\infty]$*,* $n \in \mathbb{N}$*, funzioni misurabili tali che* $f_1(x) \leq f_2(x) \leq \cdots \leq f_n(x) \leq \ldots$ *e* $\lim_{n\to+\infty} f_n(x) = f(x)$ *per quasi ogni* $x \in X$*. Allora*

$$\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu.$$

Lemma 2.2.11 (Fatou)**.** *Siano* $f_n : X \to [0,+\infty]$*,* $n \in \mathbb{N}$*, funzioni misurabili. Posto*

$$f(x) := \liminf_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} \inf_{k\geq n} f_k(x), \qquad x \in X,$$

si ha

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

Si noti che la f dell'enunciato precedente è misurabile per il teorema 2.2.4.

Teorema 2.2.12 (convergenza dominata di Lebesgue)**.** *Siano* $f_n, f : X \to \mathbb{C}$*,* $n \in \mathbb{N}$*, funzioni misurabili tali che*

(i) $\lim_{n\to+\infty} f_n(x) = f(x)$ *per q.o.* $x \in X$*;*

(ii) *esiste* $g \in L^1(X,\mu)$ *tale che* $|f_n(x)| \leq g(x)$ *per q.o.* $x \in X$ *e per ogni* $n \in \mathbb{N}$*.*

Allora

$$\lim_{n\to+\infty} \int_X |f_n - f|\,d\mu = 0.$$

Notiamo esplicitamente che da questo segue, per la proposizione 2.2.9(iii),

$$\lim_{n\to+\infty} \int_X f_n\,d\mu = \int_X f\,d\mu.$$

Altre semplici conseguenze sono un risultato di integrazione di serie termine a termine e uno di derivazione sotto il segno di integrale (esercizi 2.7 e 2.8).

Un altro risultato interessante è la seguente semplice caratterizzazione delle funzioni integrabili secondo Riemann, che, stranamente, è possibile solo nella teoria di Lebesgue (il che costituisce l'ennesimo svantaggio della teoria di Riemann).

Teorema 2.2.13 (Vitali-Lebesgue)**.** *Una funzione limitata* $f : [a,b] \to \mathbb{R}$ *è integrabile secondo Riemann se e solo se l'insieme dei suoi punti di discontinuità ha misura nulla secondo Lebesgue. Inoltre in tal caso* $f \in L^1([a,b])$*, e gli integrali secondo Riemann e secondo Lebesgue di* f *coincidono.*

Concludiamo questa sezione enunciando la generalizzazione del teorema fondamentale del calcolo per gli integrali di Lebesgue.

Teorema 2.2.14 (fondamentale del calcolo)**.** *Data* $f \in L^1([a,b])$ *e posto*

$$F(x) := \int_a^x f(t)\,dt, \qquad x \in [a,b], \tag{2.1}$$

si ha che esiste $F'(x) = f(x)$ *per q.o.* $x \in [a,b]$.

Viceversa se $F : [a,b] \to \mathbb{C}$ *è tale che esista* $F'(x)$ *per ogni* $x \in [a,b]$, *e* $F' \in L^1([a,b])$, *allora*

$$F(x) - F(a) = \int_a^x F'(t)\,dt, \qquad x \in [a,b]. \tag{2.2}$$

Osservazioni. (a) Non si può sperare che dalla (2.1) segua $F'(x) = f(x)$ per ogni $x \in [a,b]$, come si vede considerando $f = \chi_{\mathbb{Q}}$, per la quale F è identicamente nulla.

(b) Si possono costruire funzioni continue $g : [a,b] \to \mathbb{R}$ tali che $\exists\, g'(x) = 0$ per q.o. $x \in [a,b]$, ma non costanti (e non derivabili *per ogni* $x \in [a,b]$). Pertanto le "primitive" di una $f \in L^1([a,b])$ (come ad esempio la F della (2.1)) non differiscono tra loro soltanto per una costante, come per le funzioni continue.

(c) Questo mostra anche che per la validità della (2.2) non basta assumere soltanto che F' esista quasi ovunque.

2.3 Spazi L^p

Anche in questa sezione, salvo diverso avviso, $(X, \mathfrak{M}, \mu)$ denoterà un generico spazio di misura.

Definizione 2.3.1. Sia $p \geq 1$. Indicheremo con $L^p(X,\mu)$ lo spazio delle funzioni misurabili $f : X \to \mathbb{C}$ tali che

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p} < +\infty.$$

In base alla proposizione 2.2.9(iv), se $\|f\|_p = 0$ allora $f = 0$ q.o. D'altra parte si è già osservato che gli insiemi di misura nulla sono trascurabili nell'integrazione, e pertanto d'ora in poi due funzioni f e g uguali quasi ovunque verranno identificate. A stretto rigore, allora, gli elementi di L^p saranno classi di equivalenza di funzioni uguali quasi ovunque, anche se, per semplicità, e ove questo non crei ambiguità, continueremo a pensarli come funzioni.

Tenendo presente quest'ultima osservazione, si ha il risultato seguente.

Proposizione 2.3.2. $(L^p(X,\mu), \|\cdot\|_p)$ *è uno spazio normato per ogni* $p \geq 1$.

Dimostrazione. Il caso $p = 1$ segue immediatamente dalla disuguaglianza $|f+g| \leq |f| + |g|$ e dalla monotonia dell'integrale.

Per $p = 2$, l'unica cosa da verificare è che date $f, g \in L^2$ si abbia $f + g \in L^2$ e valga la disuguaglianza triangolare $\|f+g\|_2 \leq \|f\|_2 + \|g\|_2$. A tale proposito, si ha

$|f+g|^2 \leq (|f|+|g|)^2 \leq (2\max\{|f|,|g|\})^2 = 4\max\{|f|^2,|g|^2\} \leq 4(|f|^2+|g|^2) \in L^1$, e dunque $f+g \in L^2$. Similmente, $|\bar{f}g| \leq \max\{|f|,|g|\}^2 \leq |f|^2+|g|^2 \in L^1$, e pertanto l'applicazione $(f,g) \in L^2 \times L^2 \mapsto \int_X \bar{f}g\,d\mu$ definisce un prodotto scalare, definizione 3.1.1, e usando allora 2.2.9(iii) si conclude come nell'esempio 1.1.2(f).

Omettiamo la dimostrazione nei casi $p \neq 1,2$. Il lettore interessato può trovarla ad esempio su [Rud2]. □

Nel seguito avremo spesso a che fare anche con il caso $p=\infty$.

Definizione 2.3.3. Una funzione misurabile $f : X \to \mathbb{C}$ è detta *essenzialmente limitata* se esiste $M \geq 0$ tale che l'insieme $\{|f|>M\}$ ha misura nulla. La quantità

$$\|f\|_\infty := \inf\{M \geq 0 \,:\, \mu(\{|f|>M\})=0\}$$

è detta *estremo superiore essenziale* di f. L'insieme delle funzioni essenzialmente limitate si denota con $L^\infty(X,\mu)$.

È facile verificare, sempre identificando tra loro funzioni uguali quasi ovunque, che $(L^\infty(X,\mu),\|\cdot\|_\infty)$ è uno spazio normato (esercizio 2.9).

Come già accennato, il risultato seguente è di importanza fondamentale, e, a posteriori, motiva praticamente da solo la teoria dell'integrazione di Lebesgue.

Teorema 2.3.4 (Riesz-Fisher)**.** *$(L^p(X,\mu),\|\cdot\|_p)$ è uno spazio di Banach per ogni $p \in [1,+\infty]$.*

Dimostrazione. Sia $p \in [1,+\infty)$. Data una successione di Cauchy $(f_n)_{n\in\mathbb{N}} \subset L^p$, si fissi $n_1 \in \mathbb{N}$ tale che per ogni $n,m \geq n_1$ si abbia $\|f_n - f_m\|_p < 1/2$. Sarà poi possibile trovare $n_2 \geq n_1$ tale che per ogni $n,m \geq n_2$ si abbia $\|f_n - f_m\|_p < 1/4$, e d'altra parte $\|f_{n_1} - f_{n_2}\|_p \leq 1/2$. Procedendo induttivamente in questo modo si determina allora una sottosuccessione $(f_{n_k})_{k\in\mathbb{N}}$ tale che

$$\|f_{n_k} - f_{n_{k+1}}\|_p < \frac{1}{2^k}, \qquad k \in \mathbb{N}.$$

Si definiscano poi le funzioni a valori in $[0,+\infty]$

$$\left.\begin{aligned} g_r(x) &:= \sum_{k=1}^{r} |f_{n_{k+1}}(x) - f_{n_k}(x)|, \\ g(x) &:= \lim_{r\to+\infty} g_r(x) = \sum_{k=1}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|, \end{aligned}\right\} \qquad x \in X.$$

Essendo $\|\cdot\|_p$ una norma si avrà dunque

$$\|g_r\|_p \leq \sum_{k=1}^{r} \|f_{n_{k+1}} - f_{n_k}\|_p < \sum_{k=1}^{r} \frac{1}{2^k} < 1,$$

e quindi, per il lemma di Fatou 2.2.11,

$$\|g\|_p^p = \int_X |g|^p\,d\mu = \int_X \lim_{r\to+\infty} |g_r|^p\,d\mu \leq \liminf_{r\to+\infty} \int_X |g_r|^p\,d\mu = \liminf_{r\to+\infty} \|g_r\|_p^p \leq 1,$$

da cui $g(x) < +\infty$ per q.o. $x \in X$ (proposizione 2.2.9(vi)). Ne segue che è possibile definire una funzione $f : X \to \mathbb{C}$ tramite

$$f(x) := f_{n_1}(x) + \sum_{k=1}^{+\infty} \big(f_{n_{k+1}}(x) - f_{n_k}(x)\big),$$

in quanto la serie nel membro di destra è assolutamente convergente per q.o. $x \in X$, e si può poi porre $f(x) = 0$ per gli x in cui non converge. Poiché allora la somma di una serie di funzioni misurabili è misurabile per il teorema 2.2.4, la f così definita è misurabile in base all'esercizio 2.10. Ma allora, osservando che la serie che definisce f è telescopica, si avrà, per q.o. $x \in X$,

$$\begin{aligned} f(x) &= f_{n_1}(x) + \lim_{r\to+\infty} \sum_{k=1}^{r-1} \big(f_{n_{k+1}}(x) - f_{n_k}(x)\big) \\ &= f_{n_1}(x) + \lim_{r\to+\infty} \big(f_{n_r}(x) - f_{n_1}(x)\big) = \lim_{r\to+\infty} f_{n_r}(x), \end{aligned}$$

e dato allora $\varepsilon > 0$ e $n_\varepsilon \in \mathbb{N}$ tale che se $m, n > n_\varepsilon$ allora $\|f_n - f_m\|_p < \varepsilon$, si otterrà, ancora grazie al lemma di Fatou,

$$\int_X |f - f_m|^p \, d\mu \le \liminf_{r\to+\infty} \int_X |f_{n_r} - f_m|^p \, d\mu \le \varepsilon^p, \quad m > n_\varepsilon,$$

da cui $f = (f - f_m) + f_m \in L^p$ e $\|f - f_m\|_p \to 0$.

Lasciamo al lettore (esercizio 2.11) la dimostrazione del caso $p = \infty$. $\square$

Osserviamo esplicitamente che dalla dimostrazione si vede che da una successione di Cauchy (e quindi in particolare da una successione convergente) in L^p si può estrarre una sottosuccessione convergente q.o.

In base all'esercizio 2.12, dal teorema di Riesz-Fisher si ottiene un'altra dimostrazione, rispetto a quella data nell'esempio 1.4.3(c), della completezza di $\ell^p(I)$. Ulteriori semplici relazioni tra spazi ℓ^p e L^p con diversi valori di p sono date dall'esercizio 2.13.

L'ultimo risultato che vogliamo illustrare, e che ci sarà utile in seguito, afferma sostanzialmente che, se μ è una misura di Borel regolare (finita) su uno spazio metrico, le funzioni di $L^p(X, \mu)$, che a priori possono essere molto irregolari, si possono comunque approssimare, nella norma $\|\cdot\|_p$, con funzioni continue. Questo non è sorprendente, se si pensa al ruolo che svolgono gli insiemi aperti nelle nozioni di continuità e di misura regolare.

Teorema 2.3.5. *Siano (X, d) uno spazio metrico e μ una misura di Borel regolare su di esso finita (cioè tale che $\mu(X) < +\infty$). Allora lo spazio $C_b(X)$ delle funzioni continue limitate su X è denso (nella topologia indotta da $\|\cdot\|_p$) in $L^p(X, \mu)$ per ogni $p \in [1, +\infty)$.*

Dimostrazione. Osserviamo per iniziare che $g \in C_b(X)$ è misurabile e che dalla disuguaglianza $\int_X |g|^p \, d\mu \le \|g\|_\infty^p \mu(X) < +\infty$ segue $C_b(X) \subset L^p(X, \mu)$. Bisogna dunque dimostrare che per ogni $f \in L^p(X, \mu)$ e ogni $\varepsilon > 0$ esiste una $g \in C_b(X)$ tale che $\|g - f\|_p < \varepsilon$. Lo faremo in diversi passi, prendendo f via via più generali.

Per cominciare, sia $f = \chi_C$ la funzione caratterstica di un insieme chiuso $C \subset X$, e sia

$$g_n(x) := \frac{1}{1 + n\mathrm{dist}(x, C)}, \qquad x \in X,\, n \in \mathbb{N}.$$

È chiaro che $0 \leq g_n \leq 1$ e $g_n(x) = 1$ per ogni $x \in C$, e, in base all'esercizio 2.14, si ha $g_n \in C_b(X)$, $g_n(x) \to 0$ per ogni $x \notin C$. Cioè $g_n \to \chi_C$ puntualmente in X, ed essendo, per ogni $n \in \mathbb{N}$, $|g_n - \chi_C|^p \leq 2^p \in L^1(X, \mu)$ (se la misura è finita le funzioni costanti sono integrabili), per il teorema di convergenza dominata 2.2.12 si avrà $\|g_n - \chi_C\|_p \to 0$, e quindi $\|g_n - \chi_C\|_p < \varepsilon$ per n sufficientemente grande.

Sia poi $f = \chi_A$ con $A \in \mathfrak{B}(X)$. Dato $\varepsilon > 0$ esisterà, per la regolarità di μ, un chiuso $C \subset A$ tale che $\mu(A \setminus C) < (\varepsilon/2)^p$, e d'altra parte, per quanto appena visto, $\|g_n - \chi_C\|_p < \varepsilon/2$ per n sufficientemente grande. Ne segue

$$\begin{aligned}
\|g_n - \chi_A\|_p &\leq \|g_n - \chi_C\|_p + \|\chi_C - \chi_A\|_p \\
&= \|g_n - \chi_C\|_p + \left(\int_X |\chi_C - \chi_A|^p\right)^{1/p} \\
&= \|g_n - \chi_C\|_p + \left(\int_X \chi_{A\setminus C}\right)^{1/p} \\
&= \|g_n - \chi_C\|_p + \mu(A \setminus C)^{1/p} < \varepsilon.
\end{aligned}$$

Da questo si ottiene la tesi anche per ogni s funzione misurabile semplice: ogni tale funzione è combinanzione lineare di funzioni caratteristiche di insiemi boreliani, e basta allora definire $g \in C_b(X)$ come la combinanzione lineare, con gli stessi coefficienti, delle rispettive funzioni continue approssimanti definite sopra.

Presa poi $f \in L^p$ non negativa, esisterà, per la dimostrazione del teorema 2.2.6, una successione (s_n) di funzioni misurabili semplici non negative che converge puntualmente e monotonamente a f. Pertanto $|s_n - f|^p = (f - s_n)^p \leq f^p \in L^1$ e quindi, ancora per il teorema di convergenza dominata, si avrà $\|s_n - f\|_p \to 0$. D'altra parte, per quanto visto sopra, per ogni $n \in \mathbb{N}$ esisterà $g_n \in C_b(X)$ tale che $\|g_n - s_n\|_p < 1/n$, e dunque

$$\|g_n - f\|_p \leq \|g_n - s_n\|_p + \|s_n - f\|_p < \frac{1}{n} + \|s_n - f\|_p \to 0.$$

Infine il caso generale di $f \in L^p$ si riconduce al precedente scomponendo $f_1 := \mathrm{Re}\, f$ e $f_2 := \mathrm{Im}\, f$ in parti positive e negative come nella seconda parte della dimostrazione del teorema 2.2.6, e notando che $(f_i^{\pm})^p \leq |f_i|^p \leq |f|^p$, $i = 1, 2$. □

In particolare, dunque, se $K \subset \mathbb{R}^n$ è compatto, lo spazio $C(K)$ delle funzioni continue (complesse) su K è denso in $L^p(K)$ (con la misura di Lebesgue) per ogni $p \in [1, +\infty)$. Non è difficile dimostrare che invece non lo è per $p = \infty$ (vedere l'esercizio 2.15).

Esercizi

2.1 Sia X uno spazio topologico. Verificare che la famiglia di sottoinsiemi di X

$$\mathscr{B}(X) := \bigcap \{\mathfrak{M} : \mathfrak{M} \subset \mathscr{P}(X)\, \sigma\text{-algebra t.c. } A \in \mathfrak{M}\ \forall A \subset X \text{ aperto}\}$$

è la più piccola σ-algebra su X che contiene tutti gli insiemi aperti.

2.2 Sia $(q_n)_{n\in\mathbb{N}}$ una numerazione dei razionali in $(0,1)$ e, dato $\varepsilon > 0$, si considerino gli insiemi

$$A := \bigcup_{n=1}^{+\infty} \Big(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}}\Big) \cap (0,1)$$

e $K = [0,1] \setminus A$. Dimostrare:

(a) A, K sono Lebesgue-misurabili e $m(A) \le \varepsilon$, $m(K) \ge 1 - \varepsilon$;

(b) A è denso in $(0,1)$ e K ha interno vuoto;

*(c) K non è misurabile secondo Peano-Jordan (sugg.: si ha $m_e(K) \ge 1 - \varepsilon$: se così non fosse, si potrebbe trovare un plurintervallo aperto $E \subset \mathbb{R}$ tale che $K \subset E$ e $m(E) < 1 - \varepsilon$, ma allora $[0,1] \setminus E \subset A$...)

2.3 Mostrare che la misura di Lebesgue-Stieltjes associata alla funzione non decrescente a continua a destra $F = \chi_{[x,+\infty)}$ è la misura di Dirac concentrata in $x \in \mathbb{R}$.

2.4 Dimostrare le proprietà (iii), (iv) e (vi)-(ix) del teorema 2.2.4.

2.5 Siano $(X, \mathfrak{M}, \mu)$ uno spazio di misura, Y uno spazio topologico, $f_n : X \to Y$, $n \in \mathbb{N}$, funzioni misurabili tali che per ogni $x \in X$ esista $f(x) := \lim_{n\to+\infty} f_n(x)$. Mostrare che allora $f : X \to Y$ è misurabile.

2.6 Dimostrare le proprietà (iv)-(vi) della proposizione 2.2.9.

2.7 Siano $(X, \mathfrak{M}, \mu)$ uno spazio di misura, e $(f_n) \subset L^1(X,\mu)$. Verificare che se $\sum_{n=0}^{+\infty} \|f_n\|_1 < +\infty$, allora $\sum_{n=0}^{+\infty} f_n(x)$ converge per q.o. $x \in X$, definisce una funzione in $L^1(X,\mu)$ e

$$\int_X \Big(\sum_{n=0}^{+\infty} f_n\Big) d\mu = \sum_{n=0}^{+\infty} \int_X f_n \, d\mu.$$

2.8 Siano $(X, \mathfrak{M}, \mu)$ uno spazio di misura, $A \subset \mathbb{R}^n$ un aperto, e $f : X \times A \to \mathbb{C}$ una funzione tale che

(a) $f(\cdot, y) \in L^1(X,\mu)$ per ogni $y \in A$;

(b) esista $\frac{\partial f}{\partial y_j}(x,y)$ per ogni $(x,y) \in X \times A$;

(c) esista $g \in L^1(X,\mu)$ tale che $\left|\frac{\partial f}{\partial y_j}(x,y)\right| \le g(x)$ per ogni $(x,y) \in X \times A$.

Allora la funzione $F : A \to \mathbb{C}$ definita da $F(y) := \int_X f(x,y)\, d\mu(x)$, $y \in A$, è tale che esiste

$$\frac{\partial F}{\partial y_j}(y) = \int_X \frac{\partial f}{\partial y_j}(x,y)\, d\mu(x), \qquad y \in A.$$

2.9 Sia $(X, \mathfrak{M}, \mu)$ uno spazio di misura. Verificare:

(a) se $f : X \to \mathbb{C}$ è misurabile, la funzione $t \mapsto \mu(\{|f| > t\})$ è continua a destra;

(b) se f è essenzialmente limitata, $\|f\|_\infty = \min\{t \geq 0 \,:\, \mu(\{|f| > t\}) = 0\}$;

(c) $\|\cdot\|_\infty$ è una norma su $L^\infty(X, \mu)$;

2.10 Siano $(X, \mathfrak{M}, \mu)$ uno spazio di misura, $Y \in \mathfrak{M}$ tale che $\mu(Y^c) = 0$ e g una funzione misurabile su Y (a valori in $\tilde{\mathbb{R}}$ o $\mathbb{C}$). Mostrare che se f è l'estensione di g a X definita ponendo $f(x) = 0$ se $x \in Y^c$, allora f è misurabile.

*2.11 Mostrare che $(L^\infty(X, \mu), \|\cdot\|_\infty)$ è uno spazio di Banach. (Sugg.: data $(f_n) \subset L^\infty$ di Cauchy, sia E l'unione, su $k, m, n \in \mathbb{N}$, degli insiemi $A_k := \{|f_k| > \|f_k\|_\infty\}$, $B_{m,n} := \{|f_m - f_n| > \|f_m - f_n\|_\infty\}$; allora $\mu(E) = 0$ e $\ell^\infty(E^c)$ è completo.)

2.12 Sia I un insieme. Dimostrare che $\int_I f d\mu_\sharp = \sum_{\alpha \in I} f(\alpha)$ per ogni f per cui ha senso, e che quindi $L^p(I, \mu_\sharp) = \ell^p(I)$ per ogni $p \in [1, +\infty]$ ($\mu_\sharp$ è la misura che conta in I). Che succede se al posto di $\mu_\sharp$ si usa δ_{x_0}?

2.13 Mostrare che $\ell^1(\mathbb{N}) \subset \ell^2(\mathbb{N})$ e che, se $\mu(X) < \infty$, $L^2(X, \mu) \subset L^1(X, \mu)$. Mostrare che, usando la misura di Lebesgue, $L^2(\mathbb{R}) \not\subset L^1(\mathbb{R})$, $L^1(\mathbb{R}) \not\subset L^2(\mathbb{R})$.

2.14 Dato uno spazio metrico (X, d), e posto $\mathrm{dist}(x, B) := \inf_{y \in B} d(x, y)$, $x \in X$, $B \subset X$, verificare:

(a) $x \in X \mapsto \mathrm{dist}(x, B)$ è continua;

(b) se C è chiuso e $x \notin C$, allora $\mathrm{dist}(x, C) > 0$.

2.15 Siano (X, d) uno spazio metrico, e μ una misura di Borel su X tale che, per ogni $x \in X$, $\delta > 0$, si abbia $\mu(B_\delta(x)) > 0$. Mostrare che $C_b(X)$ è chiuso in $L^\infty(X, \mu)$ (sugg.: ricordare la dimostrazione di completezza di L^∞). Dare esempi in cui $C_b(X) = L^\infty(X, \mu)$ e in cui $C_b(X) \subsetneq L^\infty(X, \mu)$.

Capitolo 3

Teoria elementare degli spazi di Hilbert

Gli spazi di Hilbert (infinito-dimensionali) che, come noto, svolgono un ruolo fondamentale in Meccanica Quantistica, sono la generalizzazione più diretta degli spazi euclidei finito-dimensionali (come $\mathbb{R}^n$ o $\mathbb{C}^n$), e ne ereditano la maggior parte delle proprietà geometriche, come ad esempio il fatto, che vedremo nella sezione 3.2, che si può sempre trovare un complementare di un sottospazio (chiuso), cosa falsa in uno spazio di Banach generico. Esiste inoltre una nozione di base ortonormale per uno spazio di Hilbert che generalizza in molti aspetti quella degli spazi finito-dimensionali (sezione 3.3). Tutto ciò fa sì che la geometria degli spazi di Hilbert sia tutto sommato "banale", mentre ciò che è veramente interessante e ricco di risultati è lo studio degli operatori definiti su di essi, che noi inizieremo, con gli aspetti più elementari, nella sezione 3.4 di questo capitolo e proseguiremo, in forma astratta e più generale, nei successivi due.

3.1 Definizione e proprietà elementari

Definizione 3.1.1. Sia H uno spazio vettoriale complesso. Una *forma hermitiana* su H è un'applicazione $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ tale che:

(i) per ogni $x \in H$, l'applicazione $y \in H \mapsto \langle x, y\rangle$ è lineare;

(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ per ogni $x, y \in H$.

Una forma hermitiana è detta *semidefinita positiva* se vale inoltre

(iii) $\langle x, x\rangle \geq 0$ per ogni $x \in H$,

mentre è detta un *prodotto scalare* se

(iii') $\langle x, x\rangle > 0$ per ogni $x \in H$, $x \neq 0$.

In quest'ultimo caso, la coppia $(H, \langle\cdot,\cdot\rangle)$ è detta uno *spazio prehilbertiano.*

Osserviamo esplicitamente che dalla (i) e (ii) segue che $x \mapsto \langle x, y\rangle$ è *antilineare* per ogni $y \in H$ (cioè $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \bar{\alpha}_1\langle x_1, y\rangle + \bar{\alpha}_2\langle x_2, y\rangle$ per ogni $x_1, x_2, y \in H$, $\alpha_1, \alpha_2 \in \mathbb{C}$). Dunque una forma hermitiana è in particolare ciò che si dice una forma *sesquilineare.* Inoltre chiaramente $\langle 0, 0\rangle = 0$, e quindi un prodotto scalare è una forma hermitiana semidefinita positiva per cui $x = 0$ è l'unico vettore tale che $\langle x, x\rangle = 0$.

Teorema 3.1.2 (Disuguagliazna di Cauchy-Schwarz)**.** *Data una forma hermitiana semidefinita positiva $\langle \cdot, \cdot\rangle$ su H, si ha*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}, \qquad x, y \in H. \tag{3.1}$$

Inoltre, se $\langle \cdot, \cdot\rangle$ è un prodotto scalare, l'uguaglianza vale se e solo se x e y sono linearmente dipendenti.

Dimostrazione. Se $\langle x, y\rangle = 0$ la tesi è banalmente verificata. Sia allora $\langle x, y\rangle \neq 0$. Dato $\lambda \in \mathbb{R}$, si avrà

$$\begin{aligned} 0 &\leq \big\langle \lambda y + \langle x, y\rangle x, \lambda y + \langle x, y\rangle x\big\rangle \\ &= \lambda^2\langle y, y\rangle + \lambda\langle x, y\rangle\langle y, x\rangle + \lambda\overline{\langle x, y\rangle}\langle x, y\rangle + \overline{\langle x, y\rangle}\langle x, y\rangle\langle x, x\rangle \\ &= \lambda^2\langle y, y\rangle + 2\lambda|\langle x, y\rangle|^2 + |\langle x, y\rangle|^2\langle x, x\rangle. \end{aligned} \tag{3.2}$$

Osservando che l'ultimo membro è un polinomio quadratico in λ, si conclude che affinché esso sia non negativo deve aversi

$$|\langle x, y\rangle|^4 - \langle y, y\rangle|\langle x, y\rangle|^2\langle x, x\rangle \leq 0,$$

che, dividendo per $\langle x, y\rangle \neq 0$, fornisce la (3.1).

È poi chiaro che se x, y sono linearmente dipendenti si ha l'uguaglianza nella (3.1). Viceversa, se $\langle \cdot, \cdot\rangle$ è un prodotto scalare e vale l'uguaglianza nella (3.1), dalla (3.2) si ha ancora, per ogni $\lambda \in \mathbb{R}$,

$$\begin{aligned} \big\langle \lambda y + \langle x, y\rangle x, \lambda y + \langle x, y\rangle x\big\rangle &= \lambda^2\langle y, y\rangle + 2\lambda\langle x, x\rangle\langle y, y\rangle + \langle y, y\rangle\langle x, x\rangle^2 \\ &= \Big(\lambda\sqrt{\langle y, y\rangle} + \sqrt{\langle y, y\rangle}\langle x, x\rangle\Big)^2 \end{aligned}$$

e prendendo allora $\lambda = -\langle x, x\rangle$ si ottiene $\big\langle \lambda y + \langle x, y\rangle x, \lambda y + \langle x, y\rangle x\big\rangle = 0$ che, per la (iii') della definizione 3.1.1, implica $\lambda y + \langle x, y\rangle x = 0$, cioè x, y linearmente dipendenti. □

Corollario 3.1.3. *Sia $(H, \langle \cdot, \cdot\rangle)$ uno spazio prehilbertiano. Allora*

$$\|x\| := \sqrt{\langle x, x\rangle}, \qquad x \in H,$$

è una norma su H, detta la norma indotta dal prodotto scalare.

Dimostrazione. È chiaro dalla definizione 3.1.1 che $\|x\| = 0$ solo se $x = 0$ e che $\|\cdot\|$ è omogenea. Dalla disuguaglianza di Cauchy-Schwarz si ha poi, per ogni $x, y \in H$,

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

e quindi la disuguaglianza triangolare. □

Dunque ogni spazio prehilbertiano è, in modo naturale, uno spazio normato. Questo ci mette nella posizione di poter dare la definizione di spazio di Hilbert.

Definizione 3.1.4. Uno *spazio di Hilbert* è uno spazio prehilbertiano che sia completo nella norma indotta dal prodotto scalare.

Prima di passare agli esempi, scriviamo un paio di identità che ci saranno molto utili in seguito. Sia $\langle\cdot,\cdot\rangle$ una forma sesquilineare (non necessariamente hermitiana) su H, e $R_4 := \{\pm 1, \pm i\}$ l'insieme delle radici quarte dell'unità (in $\mathbb{C}$). Usando anche in questo caso, per brevità, la notazione $\|x\|^2 = \langle x, x\rangle$ (che in generale non sarà nemmeno un numero non negativo, se $\langle\cdot,\cdot\rangle$ non è semidefinita positiva), dato $\omega \in R_4$ si ha

$$\|x+\omega y\|^2 = \|x\|^2 + \omega\langle x, y\rangle + \bar{\omega}\langle y, x\rangle + \|y\|^2, \qquad x, y \in H, \tag{3.3}$$

e moltiplicando questa equazione per $\bar{\omega}$ e sommando al variare di $\omega \in R_4$ si ottiene, notando che $\sum_{\omega\in R_4}\bar{\omega} = 0 = \sum_{\omega\in R_4}\bar{\omega}^2$, l'*identità di polarizzazione*

$$\langle x, y\rangle = \frac{1}{4}\sum_{\omega\in R_4}\bar{\omega}\|x+\omega y\|^2, \qquad x, y \in H \tag{3.4}$$

che, nel caso particolare di uno spazio prehilbertiano, permette di ricostruire il prodotto scalare dalla conoscenza della norma. Similmente, sommando tra loro le (3.3) per $\omega = \pm 1$, si ottiene l'*identità del parallelogramma*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \qquad x, y \in H, \tag{3.5}$$

il cui nome è dovuto al fatto che lega le lunghezze dei lati del parallelogramma definito dai vettori x e y con le lunghezze delle sue diagonali ($x+y$ e $x-y$). Si noti che la (3.5) è un'identità soddisfatta dalla norma definita dal prodotto scalare. Si può dimostrare che in effetti essa caratterizza le norme che provengono da un prodotto scalare.

Esempi 3.1.5. (a) L'esempio più semplice di spazio di Hilbert è $\mathbb{C}^n$ con il prodotto scalare standard definito da $\langle x, y\rangle := \sum_{k=1}^n \bar{x}_k y_k$, a cui è associata la norma euclidea $\|x\|_2 = (\sum_{k=1}^n |x_k|^2)^{1/2}$, in cui, com'è noto, $\mathbb{C}^n$ è completo.

(b) Sullo spazio $\ell^2(I)$, I insieme arbitrario, dell'esempio 1.4.3(c) si può definire un prodotto scalare ponendo, in analogia all'esempio precedente:

$$\langle x, y\rangle := \sum_{\alpha\in I}\bar{x}_\alpha y_\alpha, \qquad x = (x_\alpha), y = (y_\alpha) \in \ell^2(I).$$

Per convincersi che questo è effettivamente ben definito, cioè che esiste il limite che definisce il secondo membro, osserviamo che per ogni sottoinsieme finito $F \subset I$, ponendo

$$\langle x, y\rangle_F := \sum_{\alpha\in F}\bar{x}_\alpha y_\alpha, \qquad x, y \in \ell^2(I),$$

si definisce una forma hermitiana semidefinita positiva $\langle\cdot,\cdot\rangle_F$ su $\ell^2(I)$, per la quale vale dunque l'identità di polarizzazione:

$$\langle x, y\rangle_F = \frac{1}{4}\sum_{\omega\in R_4}\bar{\omega}\sum_{\alpha\in F}|x_\alpha + \omega y_\alpha|^2.$$

Il secondo membro di quest'ultima equazione converge, come net su $\mathscr{P}_0(I)$, a $\frac{1}{4}\sum_{\omega\in R_4}\bar{\omega}\|x+\omega y\|_2^2$ per definizione della norma su $\ell^2(I)$. Dunque esiste $\langle x,y\rangle = \lim_{F\in\mathscr{P}_0(I)}\langle x,y\rangle_F$ per ogni $x,y\in\ell^2(I)$, e si verifica poi facilmente, ancora usando $\langle\cdot,\cdot\rangle_F$ e passando al limite, che è una forma hermitiana definita positiva, ed è poi un prodotto scalare in quanto chiaramente $\langle x,x\rangle = \|x\|_2^2 > 0$ per ogni $x\neq 0$. Poiché allora, come visto nell'esempio 1.4.3(c), $(\ell^2(I),\|\cdot\|_2)$ è completo, $(\ell^2(I),\langle\cdot,\cdot\rangle)$ è uno spazio di Hilbert, che, nel caso $I=\mathbb{N}$, rappresenta lo spazio dei vettori di stato di un sistema quantistico nella meccanica delle matrici di Heisenberg, Bohr e Jordan.

(c) Dato uno spazio di misura $(X,\mathfrak{M},\mu)$, si definisce su $L^2(X,\mu)$ un prodotto scalare ponendo

$$\langle\psi,\varphi\rangle := \int_X \bar{\psi}\varphi\,d\mu, \qquad \psi,\varphi\in L^2(X,\mu).$$

Che questo sia ben definito (cioè che $\bar{\psi}\varphi\in L^1$) è stato visto nella dimostrazione della proposizione 2.3.2, ed è poi chiaro che è un prodotto scalare, la cui norma indotta è proprio $\|\psi\|_2 = (\int_X|\psi|^2)^{1/2}$. Pertanto dal teorema di Riesz-Fisher 2.3.4 si ottiene che $L^2(X,\mu)$ è uno spazio di Hilbert. Nel caso particolare $X=\mathbb{R}^3$ (con la misura di Lebesgue), come già osservato all'inizio del primo capitolo, questo è lo spazio dei vettori di stato della meccanica ondulatoria di Schrödinger. Uno dei maggiori problemi agli albori della Meccanica Quantistica, noto come *problema delle rappresentazioni*, fu proprio quello di capire se le descrizioni di Heisenberg e di Schrödinger fossero o meno equivalenti. Una parte della soluzione di von Neumann fu proprio la dimostrazione che i due spazi di Hilbert relativi, $\ell^2(\mathbb{N})$ e $L^2(\mathbb{R}^3)$, sono in realtà "lo stesso". Noi vedremo quest'ultimo fatto in questo capitolo, e studieremo il problema delle rappresentazioni nella sua interezza, in linguaggio moderno, nel capitolo **??**.

Altri esempi di spazi di Hilbert, costruiti a partire da spazi di Hilbert dati, si trovano negli esercizi 3.1, 3.2, dove è introdotta l'importante nozione di *somma diretta* $\bigoplus_{\alpha\in I}H_\alpha$ di una famiglia arbitraria $(H_\alpha)_{\alpha\in I}$ di spazi di Hilbert. In particolare si vede che $\ell^2(I)=\bigoplus_{\alpha\in I}H_\alpha$ con $H_\alpha=\mathbb{C}$ per ogni $\alpha\in I$.

Si è visto, col teorema 1.4.6, che uno spazio normato può sempre essere pensato come un sottospazio denso di uno spazio di Banach, il suo completamento. Nel caso che la norma provenga da un prodotto scalare, il prossimo risultato che vogliamo dimostrare garantisce che anche quella del completamento gode della stessa proprietà, e che quindi il completamento è uno spazio di Hilbert.

Prima di formulare questo in dettaglio, osserviamo che se H,K sono spazi pre-hilbertiani, un'applicazione lineare $j:H\to K$ è isometrica se e solo se conserva i prodotti scalari, cioè $\langle j(x),j(y)\rangle = \langle x,y\rangle$ per ogni $x,y\in H$. Infatti se j è isometrica si ha, dall'identità di polarizzazione,

$$\langle j(x),j(y)\rangle = \frac{1}{4}\sum_{\omega\in R_4}\bar{\omega}\|j(x+\omega y)\|^2 = \frac{1}{4}\sum_{\omega\in R_4}\bar{\omega}\|x+\omega y\|^2 = \langle x,y\rangle,$$

e viceversa se j conserva i prodotti scalari

$$\|j(x)\|^2 = \langle j(x),j(x)\rangle = \langle x,x\rangle = \|x\|^2.$$

È inoltre facile vedere che un'applicazione che conserva i prodotti scalari è automaticamente lineare (esercizio 3.3).

Se $j : H \to K$ è un'isometria suriettiva, gli spazi H e K si dicono (isometricamente) isomorfi, e si scrive $H \cong K$.

Teorema 3.1.6. *Dato uno spazio prehilbertiano $(X, \langle\cdot,\cdot\rangle_X)$, esistono uno spazio di Hilbert $(H, \langle\cdot,\cdot\rangle_H)$, detto il* completamento di X*, e un'isometria $j : X \to H$ tali che $\overline{j(X)} = H$. Il completamento di X è unico nello stesso senso del teorema 1.4.6.*

Dimostrazione. Sia H il completamento di X pensato come spazio normato, fornito dalla seconda parte del teorema 1.4.6, e $j : X \to H$ la relativa isometria tale che $\overline{j(X)} = H$. Dati allora $x, y \in H$ esisteranno successioni $(x_n), (y_n) \subset X$ tali che $j(x_n) \to x$, $j(y_n) \to y$. Sarà allora naturale definire un prodotto scalare su H tramite:

$$\langle x, y\rangle_H := \lim_{n\to+\infty} \langle x_n, y_n\rangle_X.$$

Affinché questo sia ben definito, è necessario che il limite esista e non dipenda dalle successioni (x_n), (y_n) scelte. Per quanto riguarda l'esistenza, si ha, per ogni $n, m \in \mathbb{N}$,

$$\begin{aligned}|\langle x_n, y_n\rangle_X - \langle x_m, y_m\rangle_X| &\le |\langle x_n - x_m, y_n\rangle_X| + |\langle x_m, y_n - y_m\rangle_X| \\ &\le \|x_n - x_m\|_X\|y_n\|_X + \|x_m\|_X\|y_n - y_m\|_X \\ &= \|j(x_n) - j(x_m)\|_H\|j(y_n)\|_H + \|j(x_m)\|_H\|j(y_n) - j(y_m)\|_H,\end{aligned}$$

avendo usato la disuguaglianza di Cauchy-Schwarz nel secondo passaggio e l'isometria di j nel terzo. Poiché allora le successioni convergenti $(j(x_n)), (j(y_n))$ sono in particolare limitate e di Cauchy, ne segue che $(\langle x_n, y_n\rangle_X)_{n\in\mathbb{N}} \subset \mathbb{C}$ è di Cauchy, e quindi convergente. Per quanto riguarda l'indipendenza dalle successioni scelte, siano $(x'_n), (y'_n) \subset X$ altre due successioni tali che $j(x'_n) \to x$, $j(y'_n) \to y$. Si avrà allora, con passaggi analoghi a quelli appena fatti,

$$|\langle x_n, y_n\rangle_X - \langle x'_n, y'_n\rangle_X| \le \|j(x_n) - j(x'_n)\|_H\|j(y_n)\|_H + \|j(x'_n)\|_H\|j(y_n) - j(y'_n)\|_H.$$

Poiché allora le successioni convergenti $(j(x'_n)), (j(y_n))$ sono limitate, e $\|j(x_n) - j(x'_n)\|_H \to 0$, $\|j(y_n) - j(y'_n)\|_H \to 0$, si ha $\lim_{n\to+\infty}\langle x_n, y_n\rangle_X = \lim_{n\to+\infty}\langle x'_n, y'_n\rangle_X$, e dunque $\langle x, y\rangle_H$ è ben definito. È poi immediato verificare, usando il fatto che $\langle\cdot,\cdot\rangle_X$ è un prodotto scalare e passando al limite, che $\langle\cdot,\cdot\rangle_H$ così definita è una forma hermitiana semidefinita positiva. Che sia poi anche un prodotto scalare deriva dal fatto che la norma che esso induce è proprio la norma $\|\cdot\|_H$ del completamento:

$$\langle x, x\rangle_H = \lim_{n\to+\infty} \langle x_n, x_n\rangle_X = \lim_{n\to+\infty} \|x_n\|_X^2 = \lim_{n\to+\infty} \|j(x_n)\|_H^2 = \|x\|_H^2.$$

Infine la dimostrazione dell'unicità del completamento è analoga a quella del caso degli spazi metrici e normati. □

In base a questo risultato, il teorema 2.3.5 nel caso $p = 2$ si può riesprimere dicendo che, se X è uno spazio metrico e μ una misura di Borel regolare finita su X, $L^2(X, \mu)$ è il completamento delle funzioni continue e limitate su X. Un altro esempio di uno spazio prehilbertiano del cui completamento si può dare una rappresentazione "concreta" è dato dall'esercizio 3.4.

3.2 Ortogonali

Il vantaggio più importante derivante dall'introduzione di un prodotto scalare, è quello di poter definire l'ortogonalità tra vettori, il che consente di estendere agli spazi di Hilbert la maggior parte delle costruzioni geometriche che si fanno negli spazi finito-dimensionali.

Siano H uno spazio di Hilbert, e S un suo sottoinsieme. L'*ortogonale* di S è l'insieme

$$S^\perp := \{x \in H \,:\, \langle x, y\rangle = 0, \forall\, y \in S\}.$$

Indicheremo inoltre con $\langle S\rangle$ il sottospazio vettoriale di H generato da S, cioè l'insieme

$$\langle S\rangle := \Big\{\sum_{i=1}^n \alpha_i x_i \,:\, x_i \in S,\, \alpha_i \in \mathbb{C},\, n \in \mathbb{N}\Big\}$$

delle combinazioni lineari (finite) di elementi di S. La sua chiusura $\overline{\langle S\rangle}$ sarà detta il *sottospazio chiuso generato da* S (in base all'esercizio 3.5 questo è effettivamente un sottospazio).

Proposizione 3.2.1. *Siano H uno spazio di Hilbert e $S \subset H$. Si ha:*

(i) $S \cap S^\perp = \{0\}$;

(ii) $S^\perp$ *è un sottospazio chiuso di* H;

(iii) se $S_1 \subset S_2$, *allora* $S_2^\perp \subset S_1^\perp$;

(iv) $\overline{\langle S\rangle}^{\,\perp} = S^\perp = \langle S\rangle^\perp$.

Dimostrazione. (i) Se $x \in S \cap S^\perp$, allora $\|x\|^2 = \langle x, x\rangle = 0$, da cui $x = 0$.

(ii) Dati $x_1, x_2 \in S^\perp$ e $\alpha_1, \alpha_2 \in \mathbb{C}$, chiaramente $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \bar\alpha_1\langle x_1, y\rangle + \bar\alpha_2\langle x_2, y\rangle = 0$ per ogni $y \in S$, e dunque $\alpha_1 x_1 + \alpha_2 x_2 \in S^\perp$. Inoltre se $(x_n) \subset S^\perp$ è una successione convergente a $x \in H$, per la continuità dell'applicazione $x \mapsto \langle x, y\rangle$, $y \in H$ (esercizio 3.6), si avrà, per ogni $y \in S$, $\langle x, y\rangle = \lim_{n\to+\infty}\langle x_n, y\rangle = 0$, e dunque $x \in S^\perp$.

(iii) Dati $x \in S_2^\perp$ e $y \in S_1 \subset S_2$, si ha ovviamente $\langle x, y\rangle = 0$ e quindi $x \in S_1^\perp$.

(iv) Dalle inclusioni ovvie $S \subset \langle S\rangle \subset \overline{\langle S\rangle}$ e dal punto (iii) segue $\overline{\langle S\rangle}^{\,\perp} \subset \langle S\rangle^\perp \subset S^\perp$, e dunque basta dimostrare $S^\perp \subset \overline{\langle S\rangle}^{\,\perp}$. Sia allora $x \in S^\perp$. Per ogni $y \in \overline{\langle S\rangle}$ esisteranno $y_{n,i} \in S$ e $\alpha_{n,i} \in \mathbb{C}$, $n \in \mathbb{N}$, $i = 1, \dots, r_n$, tali che $y = \lim_{n\to+\infty}\sum_{i=1}^{r_n}\alpha_{n,i}y_{n,i}$, e pertanto

$$\langle x, y\rangle = \lim_{n\to+\infty}\sum_{i=1}^{r_n}\alpha_{n,i}\langle x, y_{n,i}\rangle = 0,$$

da cui $x \in \overline{\langle S\rangle}^{\,\perp}$. □

Siamo ora nella posizione di dimostrare il risultato che è alla base di tutta la teoria della struttura geometrica degli spazi di Hilbert.

Teorema 3.2.2 (della proiezione)**.** *Siano H uno spazio di Hilbert e $K \subset H$ un sottospazio chiuso. Per ogni $x \in H$ esistono unici $y \in K$ e $y_\perp \in K^\perp$ tali che $x = y + y_\perp$.*

Dimostrazione. Unicità. Dati altri due vettori $\tilde{y} \in K$, $\tilde{y}_\perp \in K^\perp$ tali che

$$x = y + y_\perp = \tilde{y} + \tilde{y}_\perp,$$

si avrà $y - \tilde{y} = \tilde{y}_\perp - y_\perp$ e pertanto tale vettore è in $K \cap K^\perp$ ($K^\perp$ è un sottospazio per la proposizione 3.2.1(ii)), che, per 3.2.1(i) è uguale a $\{0\}$, e dunque $y = \tilde{y}$ e $y_\perp = \tilde{y}_\perp$.

Esistenza. Dato $x \in H$, sia $d := \operatorname{dist}(x, K) = \inf_{z \in K} \|x - z\|$, e si scelga una successione $(y_n)_{n \in \mathbb{N}} \subset K$ tale che

$$d^2 \leq \|x - y_n\|^2 \leq d^2 + \frac{1}{n}, \qquad n \in \mathbb{N}.$$

Mostriamo che tale successione è di Cauchy. Si avrà:

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

avendo usato, nel secondo passaggio, l'identità del parallelogramma (3.5). Osservando allora che $(y_n + y_m)/2 \in K$, si ottiene

$$\begin{aligned}
\|y_n - y_m\|^2 &= 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2 \\
&\leq 2\left(d^2 + \frac{1}{n} + d^2 + \frac{1}{m}\right) - 4d^2 = \frac{1}{n} + \frac{1}{m}.
\end{aligned}$$

Poiché allora l'ultimo membro si può rendere arbitrariamente piccolo, ne segue l'esistenza di $\lim_{n \to +\infty} y_n =: y \in K$ (che è chiuso per ipotesi). Per concludere basta allora dimostrare che $y_\perp := x - y \in K^\perp$. A tale scopo si osservi che da quanto appena visto segue che $\|x - y\|^2 = d^2$, e dunque, per ogni $z \in K$, la funzione $\lambda \in \mathbb{R} \mapsto \|x - (y + \lambda z)\|^2 \geq d^2$ ha un minimo per $\lambda = 0$. D'altra parte si ha

$$\|x - (y + \lambda z)\|^2 = \|x - y\|^2 + \lambda^2 \|z\|^2 - 2\lambda \operatorname{Re} \langle x - y, z \rangle,$$

e dunque affinché tale polinomio quadratico abbia un minimo per $\lambda = 0$ è necessario che $\operatorname{Re} \langle x - y, z \rangle = 0$. Potendosi ripetere l'argomento con $iz \in K$ al posto di $z \in K$, si ottiene anche $\operatorname{Im} \langle x - y, z \rangle = -\operatorname{Re} \langle x - y, iz \rangle = 0$, e quindi $y_\perp \in K^\perp$. □

Si vede dalla dimostrazione che, per l'esistenza del vettore $y \in K$ che realizza $\operatorname{dist}(x, K)$, è sufficiente che K sia un chiuso *convesso* (cioè tale che se $x_1, x_2 \in K$, tutti i punti $\lambda x_1 + (1 - \lambda) x_2$, $\lambda \in [0, 1]$, del segmento che unisce x_1 e x_2, sono in K).

Corollario 3.2.3. *Dato un sottospazio chiuso K di uno spazio di Hilbert H:*

(i) esiste un unico operatore $P \in B(H)$*, detto il* proiettore ortogonale *su* K*, tale che* $Px \in K$*,* $(\mathbb{1} - P)x \in K^\perp$ *per ogni* $x \in H$*;*

(ii) un operatore $P \in B(H)$ *è il proiettore ortogonale su* $K = \{Px : x \in H\}$ *se e solo se*

$$P^2 = P \qquad e \qquad \langle Px, y\rangle = \langle x, Py\rangle, \quad x, y \in H. \tag{3.6}$$

Dimostrazione. (i) In base al teorema precedente, dato $x \in H$ esiste unico $y \in K$, tale che $x - y \in K^\perp$, ed è allora chiaro che $P : H \to H$ è univocamente determinato da $Px := y \in K$, e si ha $(\mathbb{1} - P)x \in K^\perp$. Inoltre P è lineare: dati $x_1, x_2 \in H$ e considerati i corrispondenti $y_1 = Px_1, y_2 = Px_2 \in K$, si ha chiaramente, per ogni $\alpha_1, \alpha_2 \in \mathbb{C}$, $\alpha_1 x_1 + \alpha_2 x_2 - (\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1(x_1 - y_1) + \alpha_2(x_2 - y_2) \in K^\perp$, e dunque per l'unicità del teorema della proiezione, $P(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 P x_1 + \alpha_2 P x_2$. Infine P è limitato, poiché

$$\|x\|^2 = \langle x, x\rangle = \langle Px + (\mathbb{1} - P)x, Px + (\mathbb{1} - Px)\rangle = \|Px\|^2 + \|(\mathbb{1} - P)x\|^2,$$

avendo usato, nell'ultima uguaglianza, il fatto che $\langle Px, (\mathbb{1} - P)x\rangle = 0$. Da questo segue dunque $\|Px\| \le \|x\|$ per ogni $x \in H$, da cui $\|P\| \le 1$.

(ii) Se $P \in B(H)$ è il proiettore ortogonale su un sottospazio chiuso $K \subset H$, è chiaro che $\{Px : x \in H\} = \operatorname{ran} P \subset K$, e viceversa se $x \in K$, potendosi scrivere $x = x + 0$ con $0 \in K^\perp$, si avrà, sempre per l'unicità del teorema della proiezione, $Px = x$, e dunque $K = \operatorname{ran} P$. Questo mostra anche che $P^2 x = P(Px) = Px$ per ogni $x \in H$, cioè $P^2 = P$. Infine, usando ancora $\langle Px, (\mathbb{1} - P)y\rangle = 0$ per ogni $x, y \in H$,

$$\langle Px, y\rangle = \langle Px, Py + (\mathbb{1} - P)y\rangle = \langle Px, Py\rangle = \langle Px + (\mathbb{1} - P)x, Py\rangle = \langle x, Py\rangle.$$

Viceversa se $P \in B(H)$ soddisfa la (3.6), $K := \operatorname{ran} P$ è un sottospazio chiuso di H: se $(Px_n)_{n\in\mathbb{N}} \subset K$ converge a $y \in H$ si ha, per la continuità di P, $Py = \lim_{n\to+\infty} P^2 x_n = \lim_{n\to+\infty} Px_n = y$, da cui $y \in K$. Inoltre, per ogni $x, y \in H$,

$$\langle Px, (\mathbb{1} - P)y\rangle = \langle x, P(\mathbb{1} - P)y\rangle = \langle x, (P - P^2)y\rangle = 0,$$

cioè $(\mathbb{1} - P)y \in K^\perp$ per ogni $y \in H$, e quindi P è il proiettore ortogonale su K. □

Osservazioni. (a) Per simmetria, è chiaro che se P è il proiettore sul sottospazio chiuso $K \subset H$, $Q := \mathbb{1} - P$, che, come si vede subito, soddisfa le (3.6), è il proiettore su $K^\perp$.

(b) Un operatore lineare $P : H \to H$ che soddisfa le (3.6) è in realtà automaticamente limitato. Si ha infatti, per ogni $x \in H$,

$$\begin{aligned}\|Px\|^2 + \|(\mathbb{1} - P)x\|^2 &= \langle Px, Px\rangle + \langle(\mathbb{1} - P)x, (\mathbb{1} - P)x\rangle \\ &= 2\langle Px, Px\rangle + \langle x, x\rangle - \langle Px, x\rangle - \langle x, Px\rangle \\ &= 2\langle x, P^2 x\rangle + \langle x, x\rangle - 2\langle x, Px\rangle = \|x\|^2,\end{aligned}$$

da cui appunto $\|P\| \le 1$.

È evidente che un proiettore P ha i soli autovalori 1, con autospazio K, e 0, con autospazio $K^\perp$. In Meccanica Quantistica, pertanto, i proiettori rappresentano le osservabili che possono assumere soltanto i valori 0 e 1, dette anche *questioni* sul sistema quantistico in esame. È intuitivamente chiaro che ogni altra osservabile si può esprimere in termini di questioni (quelle relative al fatto che l'osservabile assuma o meno un dato valore). La controparte matematica di questo fatto è data dal teorema spettrale, che, come vedremo nel capitolo **??**, afferma in sostanza che ogni operatore (autoaggiunto) si scrive come un opportuno "integrale" di una famiglia di proiettori.

Notiamo inoltre che da quanto visto sopra segue che, dato un sottospazio chiuso $K \subset H$, H si identifica con lo spazio di Hilbert somma diretta $K \oplus K^\perp$ (esercizio 3.7). In particolare quindi, in uno spazio di Hilbert H, ogni sottospazio chiuso K ammette un complementare (nel senso appunto di un altro sottospazio chiuso L tale che $H \cong K \oplus L$). Si può dimostrare che tale proprietà caratterizza gli spazi di Hilbert tra quelli di Banach.

L'ultima conseguenza del teorema della proiezione che vogliamo illustrare, è una descrizione molto esplicita dei funzionali limitati su uno spazio di Hilbert H. Come visto nell'esercizio 3.6, ci sono infatti elementi naturali di H^*, indotti, tramite il prodotto scalare, dai vettori di H: i funzionali f_y, $y \in H$, definiti da $f_y(x) := \langle y, x\rangle$, $x \in H$. Il prossimo teorema afferma che tutti i funzionali limitati su H si ottengono in questo modo.

Teorema 3.2.4 (rappresentazione di Riesz)**.** *Per ogni $f \in H^*$, esiste un unico $y \in H$ tale che $f = f_y$. Inoltre $\|y\| = \|f\|$.*

Un altro modo di esprimere questo risultato è dire che l'applicazione $y \in H \mapsto f_y \in H^*$ è un'isometria (antilineare) suriettiva, cioè un isomorfismo isometrico.

Dimostrazione. Sia $f \in H^*$, e si ponga $N := \ker f = \{x \in H \,:\, f(x) = 0\} = f^{-1}(\{0\})$, che è un sottospazio chiuso di H (in quanto controimmagine del chiuso $\{0\} \subset \mathbb{R}$ tramite f continuo). Se $N = H$ basta prendere $y = 0$. Supponiamo allora che $N \subsetneq H$, e sia $z \in H \setminus N$, $z \neq 0$. Detto allora P il proiettore ortogonale su N, si ha $(\mathbb{1} - P)z \neq 0$, e ogni $x \in H$ si scrive in modo unico come

$$x = \alpha(\mathbb{1} - P)z + w, \qquad \alpha \in \mathbb{C},\ w \in N.$$

Posto infatti $\alpha = \frac{f(x)}{f((\mathbb{1}-P)z)}$ (è $f((\mathbb{1} - P)z) \neq 0$ poiché $(\mathbb{1} - P)z \in N^\perp$), si ha

$$f\left(x - \alpha(\mathbb{1} - P)z\right) = f(x) - \frac{f(x)}{f((\mathbb{1} - P)z)} f((\mathbb{1} - P)z) = 0.$$

Posto allora $y := \frac{\overline{f((\mathbb{1}-P)z)}}{\|(\mathbb{1}-P)z\|^2}(\mathbb{1} - P)z$, si ottiene

$$\langle y, x\rangle = \langle y, \alpha(1 - P)z\rangle = \frac{f(x)}{\|(\mathbb{1} - P)z\|^2}\langle(\mathbb{1} - P)z, (\mathbb{1} - P)z\rangle = f(x), \qquad x \in H,$$

e dunque $f = f_y$. Se poi fosse $f = f_{\tilde y}$ per un altro $\tilde y \in H$, si avrebbe $\langle y - \tilde y, x\rangle = 0$ per ogni $x \in H$, da cui, prendendo $x = y - \tilde y$, segue $y = \tilde y$. È infine chiaro che $\|f_y\| \leq \|y\|$, e da $f_y(y) = \|y\|^2$ segue allora $\|f\| = \|f_y\| = \|y\|$. $\square$

3.3 Basi ortonormali

Definizione 3.3.1. Sia H uno spazio di Hilbert. Un *sistema ortonormale* è una famiglia $(e_\alpha)_{\alpha\in I} \subset H$ tale che

$$\langle e_\alpha, e_\beta\rangle = \delta_{\alpha\beta}, \qquad \alpha, \beta \in I.$$

Un sistema ortonormale massimale, cioè non contenuto propriamente in nessun altro sistema ortonormale, è detto una *base ortonormale*.

Esempi 3.3.2. (a) In $H = L^2([0, 2\pi])$ le funzioni

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \qquad x \in [0, 2\pi],\ n \in \mathbb{Z},$$

formano un sistema ortonormale, infatti

$$\langle e_n, e_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)x}\,dx = \delta_{m,n}, \qquad m, n \in \mathbb{Z}.$$

In realtà $(e_n)_{n\in\mathbb{Z}}$ è una base ortonormale di $L^2([0, 2\pi])$. Questo segue ad esempio dalla teoria delle serie di Fourier. Per noi sarà conseguenza del teorema di Stone-Weierstrass (teorema 4.3.1 ed esercizio 4.25).

(b) In $H = \ell^2(I)$ la famiglia di vettori $(\delta_\alpha)_{\alpha\in I}$, definiti da $\delta_\alpha = (\delta_{\alpha\beta})_{\beta\in I}$, è una base ortonormale, detta la *base ortonormale canonica* . È infatti chiaro che $\langle\delta_\alpha, \delta_\beta\rangle = \delta_{\alpha\beta}$, $\alpha, \beta \in I$, e dato poi $x = (x_\alpha)_{\alpha\in I} \in \ell^2(I)$ che sia ortogonale a tutti i δ_α, si ha

$$0 = \langle x, \delta_\alpha\rangle = \sum_{\beta\in I} \bar{x}_\beta \delta_{\alpha\beta} = x_\alpha, \qquad \alpha \in I,$$

e dunque $x = 0$ e $(\delta_\alpha)_{\alpha\in I}$ è un sistema ortonormale massimale.

Daremo ora alcune caratterizzazioni delle basi ortonormali, mostrando in particolare che ogni $x \in H$ si scrive come una opportuna "combinazione lineare infinita" di elementi della base. Cominciamo con alcuni risultati preliminari.

Lemma 3.3.3. *Siano H uno spazio di Hilbert, e $(e_\alpha)_{\alpha\in I} \subset H$ un sistema ortonormale. Allora:*

(i) per ogni $x \in H$ si ha

$$\sum_{\alpha\in I} |\langle e_\alpha, x\rangle|^2 \leq \|x\|^2 \qquad \textit{(disuguaglianza di Bessel)};$$

(ii) esiste un operatore isometrico $U : \ell^2(I) \to H$ tale che

$$Ux = \sum_{\alpha\in I} x_\alpha e_\alpha, \qquad x = (x_\alpha)_{\alpha\in I} \in \ell^2(I);$$

(iii) la somma $\sum_{\alpha\in I} x_\alpha e_\alpha$ converge in H se e solo se $(x_\alpha)_{\alpha\in I} \in \ell^2(I)$.

Dimostrazione. (i) Si ha, per ogni $F \subset I$ finito,

$$\begin{aligned}0 &\leq \left\| x - \sum_{\alpha\in F} \langle e_\alpha, x\rangle e_\alpha \right\|^2 \\ &= \|x\|^2 - \sum_{\alpha\in F} \langle e_\alpha, x\rangle\langle x, e_\alpha\rangle - \sum_{\alpha\in F} \overline{\langle e_\alpha, x\rangle}\langle e_\alpha, x\rangle + \sum_{\alpha,\beta\in F} \overline{\langle e_\alpha, x\rangle}\langle e_\beta, x\rangle\langle e_\alpha, e_\beta\rangle \\ &= \|x\|^2 - \sum_{\alpha\in F} |\langle e_\alpha, x\rangle|^2,\end{aligned}$$

e la disuguaglianza richiesta si ottiene passando al limite su F.

(ii) La base ortonormale canonica $(\delta_\alpha)_{\alpha\in I} \subset \ell^2(I)$ è chiaramente una base di Hamel[†] per lo spazio $c_c(I)$ denso in $\ell^2(I)$ (vedere l'esercizio 3.4). Pertanto esiste un unico operatore lineare $U : c_c(I) \to H$ tale che $U\delta_\alpha = e_\alpha$, $\alpha \in I$, e si avrà chiaramente $Ux = \sum_{\alpha\in I} x_\alpha e_\alpha$, per ogni $x = (x_\alpha)_{\alpha\in I} \in c_c(I)$, essendo la somma a secondo membro finita. Inoltre, sempre grazie al fatto che le somme sono finite,

$$\|Ux\|^2 = \left\| \sum_{\alpha\in I} x_\alpha e_\alpha \right\|^2 = \sum_{\alpha,\beta\in I} \bar{x}_\alpha x_\beta \langle e_\alpha, e_\beta\rangle = \sum_{\alpha\in I} |x_\alpha|^2 = \|x\|^2.$$

Dunque $U : c_c(I) \to H$ è isometrico, e quindi per il teorema 1.4.5 si estende univocamente a un operatore limitato, che denoteremo ancora con U, da $\ell^2(I)$ a H. Inoltre, in base all'esercizio 3.9, per ogni $x = (x_\alpha)_{\alpha\in I} \in \ell^2(I)$ si ha $x = \lim_F \sum_{\alpha\in F} x_\alpha\delta_\alpha$, e quindi, per la continuità di U,

$$Ux = \lim_F U\left(\sum_{\alpha\in F} x_\alpha\delta_\alpha\right) = \lim_F \sum_{\alpha\in F} x_\alpha e_\alpha = \sum_{\alpha\in I} x_\alpha e_\alpha,$$

e inoltre, essendo U isometrico su $c_c(I)$,

$$\|Ux\|^2 = \lim_F \left\| U\left(\sum_{\alpha\in F} x_\alpha\delta_\alpha\right)\right\|^2 = \lim_F \sum_{\alpha\in F} |x_\alpha|^2 = \|x\|^2,$$

cioè l'isometria di U su $\ell^2(I)$.

(iii) Se $\sum_{\alpha\in I} x_\alpha e_\alpha = x \in H$, si ha $\langle x, e_\alpha\rangle = x_\alpha$, e dunque $(x_\alpha)_{\alpha\in I} \in \ell^2(I)$ per (i). Il viceversa segue immediatamente da (ii). □

Teorema 3.3.4. *Siano H uno spazio di Hilbert, e $(e_\alpha)_{\alpha\in I} \subset H$ un sistema ortonormale. Sono equivalenti:*

(i) $(e_\alpha)_{\alpha\in I}$ è una base ortonormale;

[†] Ricordiamo che dato uno spazio vettoriale X, una base di Hamel (o semplicemente una base) per X è una famiglia $(e_\alpha)_{\alpha\in I} \subset X$ di vettori linearmente indipendenti, tali che ogni $x \in X$ si scriva come combinazione lineare (finita) degli $(e_\alpha)_{\alpha\in I}$.

(ii) per ogni $x \in H$,

$$x = \sum_{\alpha \in I} \langle e_\alpha, x \rangle e_\alpha;$$

(iii) l'isometria $U : \ell^2(I) \to H$ definita nel lemma 3.3.3(ii) è suriettiva;

(iv) per ogni $x, y \in H$,

$$\langle x, y \rangle = \sum_{\alpha \in I} \langle x, e_\alpha \rangle \langle e_\alpha, y \rangle;$$

(v) per ogni $x \in H$,

$$\|x\|^2 = \sum_{\alpha \in I} |\langle x, e_\alpha \rangle|^2 \qquad \textit{(identità di Parseval)};$$

(vi) per ogni $x \in H$, $\langle e_\alpha, x \rangle = 0$ per ogni $\alpha \in I$ implica $x = 0$.

Dimostrazione. (i) $\Rightarrow$ (ii). Per il lemma 3.3.3, la somma $\sum_{\alpha \in I} \langle e_\alpha, x \rangle e_\alpha$ converge in H. Posto allora $y := x - \sum_{\alpha \in I} \langle e_\alpha, x \rangle e_\alpha$, si ha chiaramente $\langle e_\alpha, y \rangle = 0$ per ogni $\alpha \in I$, e dunque, per la massimalità del sistema ortonormale $(e_\alpha)_{\alpha \in I}$, deve essere $y = 0$.

(ii) $\Rightarrow$ (iii). Dato $x \in H$, e posto $\xi := (\langle e_\alpha, x \rangle)_{\alpha \in I}$, si ha $\xi \in \ell^2(I)$ per il lemma 3.3.3(i), e chiaramente $U\xi = x$.

(iii) $\Rightarrow$ (iv). Se U è suriettivo è invertibile, e $U^{-1}x = (\langle e_\alpha, x \rangle)_{\alpha \in I}$. Dunque, essendo U, e quindi U^{-1}, isometrico,

$$\langle x, y \rangle = \langle U^{-1}x, U^{-1}y \rangle = \sum_{\alpha \in I} \langle x, e_\alpha \rangle \langle e_\alpha, y \rangle.$$

(iv) $\Rightarrow$ (v). Basta porre $y = x$.

(v) $\Rightarrow$ (vi). Ovvio.

(vi) $\Rightarrow$ (i). Ovvio. □

Dunque, come annunciato, ogni vettore dello spazio di Hilbert si scrive come una somma (a priori infinita) convergente di multipli dei vettori di una base ortonormale, analogamente a quanto succede in spazi finito-dimensionali. Non è difficile vedere che tali somme sono in realtà al più numerabili (esercizio 3.10), cioè sono delle serie. L'esercizio 3.11 illustra la differenza tra la nozione di base ortonormale e quella classica di base di Hamel.

Vogliamo ora dimostrare che ogni spazio di Hilbert ammette una base ortonormale. Per far questo, abbiamo bisogno di alcune nozioni e di un risultato fondamentale di teoria degli insiemi.

Definizione 3.3.5. Sia $(I, \leq)$ un insieme parzialmente ordinato. Allora:

(i) un *maggiorante* per un sottoinsieme $J \subset I$ è un $\alpha \in I$ tale che $\alpha \geq \beta$ per ogni $\beta \in J$;

(ii) un elemento $\alpha \in I$ è detto *massimale* se $\alpha \leq \beta$, $\beta \in I$, implica $\beta = \alpha$;

(iii) $(I, \leq)$ è detto *linearmente ordinato* se per ogni $\alpha, \beta \in I$ si ha $\alpha \leq \beta$ o $\beta \leq \alpha$.

Ad esempio, è chiaro che l'insieme $\mathbb{N}$, con l'ordine naturale, è linearmente ordinato e privo di elementi massimali, mentre l'insieme $\mathscr{P}(X)$ dei sottoinsiemi di un insieme X fissato, con l'ordine definito dall'inclusione, $A \leq B \Leftrightarrow A \subset B$, non è linearmente ordinato e ha un unico elemento massimale, l'insieme X stesso.

Il prossimo risultato, che non dimostreremo, è una conseguenza diretta dell'assioma della scelta, e anzi è ad esso equivalente.

Lemma 3.3.6 (Zorn). *Sia $(I, \leq)$ un insieme parzialmente ordinato in cui ogni sottoinsieme $L \subset I$ linearmente ordinato ammette un maggiorante. Allora I ha un elemento massimale.*

Dimostrazione. La omettiamo. Il lettore interessato può trovarla ad esempio su [Ped, thm. 1.1.6]. □

Il lemma di Zorn interviene in Analisi Funzionale in moltissime dimostrazioni di esistenza "astratte", come esemplificato dal teorema seguente.

Teorema 3.3.7. *In ogni spazio di Hilbert esiste una base ortonormale.*

Dimostrazione. Consideriamo l'insieme $\mathscr{E}$ i cui elementi sono i sistemi ortonormali in H, con l'ordinamento parziale definito dall'inclusione: $(e_\alpha)_{\alpha \in I} \leq (f_\beta)_{\beta \in J}$ se $\{e_\alpha : \alpha \in I\} \subset \{f_\beta : \beta \in J\}$. Un qualunque sottoinsieme linearmente ordinato $\mathscr{L} \subset \mathscr{E}$ ha un maggiorante. Infatti la famiglia $(e^{\mathscr{L}}_\alpha)_{\alpha \in I_{\mathscr{L}}} \subset H$ ottenuta facendo l'unione di tutti i sistemi ortonormali appartenenti a $\mathscr{L}$ è essa stessa un sistema ortonormale: essendo $\mathscr{L}$ linearmente ordinato, due vettori $e^{\mathscr{L}}_\alpha$, $e^{\mathscr{L}}_\beta$, $\alpha, \beta \in I_{\mathscr{L}}$, appartengono a uno stesso sistema ortonormale in $\mathscr{L}$, e dunque

$$\langle e^{\mathscr{L}}_\alpha, e^{\mathscr{L}}_\beta \rangle = \delta_{\alpha\beta}.$$

È poi chiaro che $(e^{\mathscr{L}}_\alpha)_{\alpha \in \mathscr{L}}$ maggiora tutti gli elementi di $\mathscr{L}$. Ne segue, per il lemma di Zorn, che $\mathscr{E}$ ha un elemento massimale, che è dunque una base ortonormale di H. □

Dunque, in base al teorema precedente e al teorema 3.3.4(iii), ogni spazio di Hilbert H si può identificare, fissata una base ortonormale, con un $\ell^2(I)$ per un insieme I opportuno. Si può poi dimostrare che la cardinalità di I è univocamente determinata, il che permette di definire la nozione di dimensione (hilbertiana) di H.

Il difetto del teorema 3.3.7, come di tutti i risultati basati sull'assioma della scelta, è che non fornisce una costruzione esplicita di una base ortonormale, ed è quindi di scarsa utilità pratica. D'altra parte, rinunciare a utilizzare l'assioma della scelta implicherebbe un notevole impoverimento dei risultati dell'Analisi Funzionale.

Fortunatamente, gli spazi di Hilbert di interesse principale in Meccanica Quantistica appartengono a una classe più ristretta, che ora introdurremo, e per tali spazi è disponibile un procedimento "costruttivo" in grado di fornire un gran numero di basi ortonormali.

Definizione 3.3.8. Uno spazio topologico è detto *separabile* se ammette un sottoinsieme numerabile denso.

È chiaro, ad esempio, che $\mathbb{R}$ è separabile, in quanto $\mathbb{Q} \subset \mathbb{R}$ è denso e numerabile. Da questo segue subito che $\mathbb{C}^n$ è pure separabile: basta approssimare parti reali e immaginare di ogni componente di $z \in \mathbb{C}^n$ con numeri razionali, e ricordare che un prodotto cartesiano finito (o numerabile) di insiemi numerabili è numerabile.

Teorema 3.3.9. *Uno spazio di Hilbert è separabile se e solo se ammette una base ortonormale numerabile.*

Dimostrazione. Sia H uno spazio di Hilbert separabile, e $\{y_1, y_2, y_3, \dots\} \subset H$ una successione densa. Da questa si può estrarre un sottoinsieme $\{x_1, x_2, \dots\}$, al piú numerabile, di vettori linearmente indipendenti tali che $\langle x_1, x_2, \dots\rangle = \langle y_1, y_2, \dots\rangle$. A tale scopo basta infatti porre $x_1 := y_1$, $x_2 := y_{n_2}$, con y_{n_2} il primo vettore della successione $\{y_2, y_3, \dots\}$ linearmente indipendente da x_1, $x_3 := y_{n_3}$, con y_{n_3} il primo vettore della successione $\{y_{n_2+1}, y_{n_2+2}, \dots\}$ linearmente indipendente da x_1, x_2, e così via. È allora chiaro che ogni combinazione lineare (finita) degli $\{y_1, y_2, \dots\}$ si può scrivere come combinazione lineare degli $\{x_1, x_2, \dots\}$, e dunque effettivamente $\langle x_1, x_2, \dots\rangle = \langle y_1, y_2, \dots\rangle$. Supponiamo che la famiglia $\{x_1, x_2, \dots\}$ sia infinita numerabile (il caso in cui sia finita si tratta analogamente). Applicando a essa il procedimento di Gram-Schmidt, già noto nel caso di un numero finito di vettori, e che si estende immutato anche al caso infinito (esercizio 3.12), si ottiene un sistema ortonormale $(e_n)_{n\in\mathbb{N}}$ tale che, per ogni $n \in \mathbb{N}$, $\langle e_1, \dots, e_n\rangle = \langle x_1, \dots, x_n\rangle$. Da questo segue

$$\overline{\langle e_1, e_2, \dots\rangle} = \overline{\langle x_1, x_2, \dots\rangle} = \overline{\langle y_1, y_2, \dots\rangle} \supset \overline{\{y_1, y_2, \dots\}} = H,$$

e dunque, per la proposizione 3.2.1(iv),

$$\{e_1, e_2, \dots\}^\perp = \overline{\langle e_1, e_2, \dots\rangle}^\perp = H^\perp = \{0\},$$

cioè $(e_n)_{n\in\mathbb{N}}$ è una base ortonormale in H.

Viceversa se $(e_n)_{n\in\mathbb{N}} \subset H$ è una base ortonormale, l'insieme delle combinazioni lineari finite di tali vettori a coefficienti in $\mathbb{Q} + i\mathbb{Q}$ è numerabile e denso in H, che è quindi separabile. La facile verifica di questi fatti è lasciata al lettore (esercizio 3.13). □

Dunque $\ell^2(\mathbb{N})$ è il prototipo di ogni spazio di Hilbert separabile. Gli esercizi 3.15 e 3.16 forniscono esempi notevoli, e fisicamente interessanti, di basi ortonormali numerabili in $L^2(\mathbb{R})$ e $L^2([0, +\infty))$ rispettivamente. In particolare questo mostra che $L^2(\mathbb{R}) \cong \ell^2(\mathbb{N})$, come annunciato nell'esempio 3.1.5(c).

Questi esempi illustrano anche il fatto spesso utile che, grazie al procedimento di Gram-Schmidt, è sempre possibile, se H è separabile, trovare una base ortonormale contenuta in un sottospazio denso di vettori "interessanti", come ad esempio il sottospazio di $L^2(\mathbb{R})$ costituito dalle funzioni C^∞.

3.4 Aggiunti

Come visto fin qui, dal punto di vista geometrico gli spazi di Hilbert godono di proprietà che sono una diretta generalizzazione di quelle tipiche degli spazi finito-dimensionali. Se da una parte questo fa sì che il loro studio non rivesta particolare interesse, dall'altra permette invece di sviluppare una teoria molto ricca e appronfondita degli operatori lineari tra spazi di Hilbert, il cui culmine è il teorema spettrale (teorema **??**).

Cominceremo ora questa analisi con un risultato, di utilizzo piuttosto frequente, che stabilisce una corrispondenza tra operatori limitati tra spazi di Hilbert e certe forme sesquilineari, che ora definiamo.

Definizione 3.4.1. Siano H, K spazi di Hilbert. Ricordiamo che un'applicazione $s : H \times K \to \mathbb{C}$ è una *forma sesquilineare* se per ogni $x \in H$, l'applicazione $s(x, \cdot) : K \to \mathbb{C}$ è lineare e per ogni $y \in K$, l'applicazione $s(\cdot, y) : H \to \mathbb{C}$ è antilineare. La forma sesquilineare s è detta *limitata*, se esiste $C > 0$ tale che

$$|s(x, y)| \leq C\|x\|\|y\|, \qquad x \in H, y \in K.$$

Lemma 3.4.2. *Siano H, K spazi di Hilbert. Una forma sesquilineare $s : H \times K \to \mathbb{C}$ è limitata se e solo se esiste $T \in B(K, H)$ tale che*

$$s(x, y) = \langle x, Ty \rangle, \qquad x \in H, y \in K. \tag{3.7}$$

In tal caso, $T \in B(K, H)$ è univocamente determinato, e si ha $\|T\| \leq C$, dove $C > 0$ è la costante che appare nella definizione 3.4.1.

Dimostrazione. Se vale la (3.7), si ha, per la disuguaglianza di Cauchy-Schwarz e la limitatezza di T,

$$|s(x, y)| = |\langle x, Ty \rangle| \leq \|T\|\|x\|\|y\|, \qquad x \in H, y \in K,$$

e dunque s è limitata.

Se viceversa s è limitata, per ogni $y \in K$ il funzionale lineare su H $\overline{s(\cdot, y)}$ è pure chiaramente limitato, e dunque per il teorema di rappresentazione di Riesz esisterà $Ty \in H$ univocamente determinato tale che $\overline{s(x, y)} = \langle Ty, x \rangle$ per ogni $x \in H$, che è la (3.7). Dall'unicità di Ty e dalla linearità di $y \mapsto s(x, y)$ segue poi chiaramente che $y \mapsto Ty$ è lineare e che T è univocamente determinato, e infine, sempre per il teorema di Riesz,

$$\|Ty\| = \left\|\overline{s(\cdot, y)}\right\| = \sup_{\|x\| \leq 1} |s(x, y)| \leq C\|y\|, \qquad y \in K,$$

da cui $T \in B(K, H)$ e $\|T\| \leq C$. □

Come prima applicazione di questo lemma, mostreremo che esiste un'ulteriore operazione sugli operatori limitati su uno spazio di Hilbert, oltre a quelle di combinazione lineare e composizione (viste nella proposizione 1.3.3), che generalizza la coniugazione complessa.

Teorema 3.4.3. *Siano H, K spazi di Hilbert, e $T \in B(H, K)$. Esiste un unico operatore $T^* \in B(K, H)$, detto l'*aggiunto *di T, tale che*

$$\langle Tx, y \rangle = \langle x, T^*y \rangle, \qquad x \in H, y \in K. \tag{3.8}$$

Inoltre l'applicazione $T \in B(H, K) \mapsto T^* \in B(K, H)$ gode delle seguenti proprietà:

(i) *antilinearità:* $(\alpha T + \beta S)^* = \bar{\alpha} T^* + \bar{\beta} S^*$ *per ogni* $T, S \in B(H, K)$, $\alpha, \beta \in \mathbb{C}$;

(ii) *antimoltiplicatività:* $(ST)^* = T^* S^*$ *per ogni* $T \in B(H, K)$, $S \in B(K, L)$ *(*L *spazio di Hilbert);*

(iii) *involutività:* $T^{**} = T$ *per ogni* $T \in B(H, K)$;

(iv) *isometria:* $\|T^*\| = \|T\|$ *per ogni* $T \in B(H, K)$;

(v) *identità* C^**:* $\|T^* T\| = \|T\|^2$ *per ogni* $T \in B(H, K)$.

Dimostrazione. Si consideri la forma sesquilineare $s : H \times K \to \mathbb{C}$ definita da

$$s(x, y) := \langle Tx, y \rangle, \qquad x \in H, y \in K.$$

Si ha chiaramente $|s(x, y)| \leq \|T\| \|x\| \|y\|$, e dunque s è limitata. Quindi per il lemma precedente esiste un unico $T^* \in B(K, H)$ tale che $s(x, y) = \langle x, T^* y \rangle$, cioè la (3.8), e inoltre $\|T^*\| \leq \|T\|$.

Le proprietà (i) e (ii) seguono subito dall'unicità di T^*: ad esempio si ha, per ogni $T \in B(H, K), S \in B(K, L), x \in H, z \in K$,

$$\langle STx, z \rangle = \langle Tx, S^* z \rangle = \langle x, T^* S^* z \rangle,$$

dove la prima uguagliaza segue dalla definizione di S^* e la seconda da quella di T^*, e dunque per l'appunto $(ST)^* = T^* S^*$.

La (iii) si ottiene poi scambiando i ruoli di T e T^* nella (3.8), e da questo, insieme a quanto visto sopra, segue $\|T\| = \|T^{**}\| \leq \|T^*\|$, e dunque la (iv).

Infine si ha

$$\|T\|^2 = \sup_{\|x\| \leq 1} \|Tx\|^2 = \sup_{\|x\| \leq 1} \langle Tx, Tx \rangle = \sup_{\|x\| \leq 1} \langle x, T^* Tx \rangle$$

ed essendo, per ogni $x \in H$, $|\langle x, T^* T x \rangle| \leq \|T^* T\| \|x\|^2$ per la disuguaglianza di Cauchy-Schwarz, ne segue $\|T\|^2 \leq \|T^* T\|$. D'altra parte $\|T^* T\| \leq \|T\|^2$ per la proposizione 1.3.3(ii) e per la (iv), e quindi si ottiene la (v). □

Osserviamo esplicitamente che da questo e dal corollario 3.2.3 si ha che $P \in B(H)$ è un proiettore ortogonale se e solo se $P^2 = P = P^*$. Analogamente un operatore $V : H \to K$ è isometrico se e solo se

$$\langle x, y \rangle = \langle Vx, Vy \rangle = \langle x, V^* V y \rangle, \qquad x, y \in H,$$

e dunque se e solo se $V^* V = \mathbb{1}$.

Come vedremo nel prossimo capitolo, l'esistenza dell'operazione di aggiunzione rende $B(H)$ il prototipo di una struttura algebrico-topologica dalle proprietà assai interessanti e ricche di conseguenze.

Esercizi

3.1 Siano $(H, \langle\cdot,\cdot\rangle_H)$, $(K, \langle\cdot,\cdot\rangle_K)$ spazi di Hilbert. Sullo spazio vettoriale somma diretta

$$H \oplus K := \{(x, y) \, : x \in H, y \in K\},$$

i cui elementi saranno denotati con $x \oplus y := (x, y)$, si definisca

$$\langle x_1 \oplus y_1, x_2 \oplus y_2\rangle := \langle x_1, x_2\rangle_H + \langle y_1, y_2\rangle_K, \qquad x_i \oplus y_i \in H \oplus K.$$

Verificare che $(H \oplus K, \langle\cdot,\cdot\rangle)$ è uno spazio di Hilbert, detto *somma diretta (hilbertiana)* di H e K.

*3.2 Siano $(H_\alpha, \langle\cdot,\cdot\rangle_\alpha)$, $\alpha \in I$, spazi di Hilbert, e si definisca

$$\bigoplus_{\alpha\in I} H_\alpha := \Big\{(x_\alpha)_{\alpha\in I} \, : \, x_\alpha \in H_\alpha \, \forall\alpha \in I, \sum_{\alpha\in I} \|x_\alpha\|_\alpha^2 < \infty\Big\}.$$

Verificare:

(a) $\bigoplus_{\alpha\in I} H_\alpha$ è uno spazio vettoriale con le operazioni definite da

$$(x_\alpha)_{\alpha\in I} + (y_\alpha)_{\alpha\in I} := (x_\alpha + y_\alpha)_{\alpha\in I}, \qquad \lambda(x_\alpha)_{\alpha\in I} := (\lambda x_\alpha)_{\alpha\in I};$$

(b) $\bigoplus_{\alpha\in I} H_\alpha$ è uno spazio di Hilbert con il prodotto scalare definito da

$$\big\langle (x_\alpha)_{\alpha\in I}, (y_\alpha)_{\alpha\in I}\big\rangle := \sum_{\alpha\in I} \langle x_\alpha, y_\alpha\rangle_\alpha.$$

Tale spazio è detto la *somma diretta* della famiglia $(H_\alpha)_{\alpha\in I}$.

3.3 Siano X, Y spazi prehilbertiani, e $j : X \to Y$ un'applicazione che preserva i prodotti scalari. Mostrare che j è lineare (sugg.: calcolare $\|j(x + y) - j(x) - j(y)\|^2, \|j(\alpha x) - \alpha j(x)\|^2$).

3.4 Mostrare che $\ell^2(I)$ è il completamento dello spazio prehilbertiano

$$c_c(I) := \{x = (x_\alpha)_{\alpha\in I} \, : \, x_\alpha \in \mathbb{C},\, x_\alpha \neq 0 \text{ solo per un insieme finito di } \alpha \in I\}$$

con il prodotto scalare $\langle x, y\rangle = \sum_{\alpha\in I} \bar{x}_\alpha y_\alpha$.

3.5 Siano X uno spazio normato, $S \subset X$ un sottoinsieme, e $\langle S\rangle$ il sottospazio da esso generato. Mostrare che $\overline{\langle S\rangle}$ è un sottospazio chiuso di X.

3.6 Sia H uno spazio di Hilbert. Mostrare che, per ogni $y \in H$, il funzionale lineare $x \in H \mapsto \langle y, x\rangle \in \mathbb{C}$ è limitato.

3.7 Siano H uno spazio di Hilbert e $K \subset H$ un sottospazio chiuso. Mostrare che H è isometricamente isomorfo a $K \oplus K^\perp$ (cioè esiste $j : H \to K \oplus K^\perp$ isometria lineare suriettiva).

3.8 Sia $H = \ell^2(\mathbb{N})$.

(a) Se $x = (x_n) \in H$, verificare che la serie di potenze $\sum_{n=0}^{+\infty} x_n z^n$ ha raggio di convergenza $r \geq 1$.

(b) Dato $\lambda \in \mathbb{C}$, $|\lambda| < 1$, verificare che posto $f(x) := \sum_{n=0}^{+\infty} x_n \lambda^n$, $x \in H$, si ha $f \in H^*$.

(c) Determinare $y \in H$ tale che $f(x) = \langle y, x\rangle$ per ogni $x \in H$, e calcolare $\|f\|$.

3.9 Sia $(\delta_\alpha)_{\alpha\in I} \subset \ell^2(I)$ la base ortonormale canonica: $\delta_\alpha = (\delta_{\alpha\beta})_{\beta\in I}$. Senza usare il teorema 3.3.4, dimostrare che per ogni $x = (x_\alpha)_{\alpha\in I} \in \ell^2(I)$ si ha $x = \sum_{\alpha\in I} x_\alpha\delta_\alpha$, cioè $\lim_{F\in\mathscr{P}_0(I)} \|x - \sum_{\alpha\in F} x_\alpha\delta_\alpha\| = 0$.

3.10 Mostrare che, se $x \in H$ e $(e_\alpha)_{\alpha\in I} \subset H$ è un sistema ortonormale, l'insieme $\{\alpha \in I : \langle e_\alpha, x\rangle \neq 0\}$ è al più numerabile (sugg.: si ha $\sharp\{\alpha \in I : |\langle e_\alpha, x\rangle| \geq 1/n\} \leq n^2 \sum_{\alpha\in I} |\langle e_\alpha, x\rangle|^2$).

*3.11 Dimostrare che una base ortonormale di uno spazio di Hilbert H infinito-dimensionale non è una base di Hamel, e che una base di Hamel di H ha cardinalità più che numerabile.

3.12 Siano H uno spazio di Hilbert, e $(x_n)_{n\in\mathbb{N}} \subset H$ una successione di vettori linearmente indipendenti. Definita una successione $(f_n)_{n\in\mathbb{N}}$ tramite il *procedimento di ortogonalizzazione di Gram-Schmidt*:

$$f_1 := x_1, \quad f_2 := x_2 - \frac{\langle x_2, f_1\rangle}{\langle f_1, f_1\rangle} f_1, \quad f_3 := x_3 - \frac{\langle x_3, f_1\rangle}{\langle f_1, f_1\rangle} f_1 - \frac{\langle x_3, f_2\rangle}{\langle f_2, f_2\rangle} f_2, \quad \ldots$$

dimostrare che i vettori $e_n := f_n/\|f_n\|$, $n \in \mathbb{N}$, sono un sistema ortonormale in H tale che, per ogni $n \in \mathbb{N}$, $\langle e_1, \ldots, e_n\rangle = \langle x_1, \ldots, x_n\rangle$.

3.13 Completare la seconda parte della dimostrazione del teorema 3.3.9.

*3.14 Sia $f \in L^1(\mathbb{R})$ tale che $e^{\delta|x|} f \in L^1(\mathbb{R})$ per qualche $\delta > 0$. Mostrare allora che la trasformata di Fourier

$$\hat{f}(p) = \int_{\mathbb{R}} f(x) e^{-ipx}\, dx, \qquad p \in \mathbb{R}$$

di f si estende ad una funzione analitica nella striscia $\{p \in \mathbb{C} : |\mathrm{Im}\, p| < \delta\}$ e che si ha lo sviluppo di MacLaurin

$$\hat{f}(p) = \sum_{n=0}^{+\infty} \frac{(-ip)^n}{n!} \int_{\mathbb{R}} f(x) x^n\, dx, \qquad |p| < \delta$$

(sugg.: l'estensione ovvia di $\hat{f}$ soddisfa le equazioni di Cauchy-Riemann per l'esercizio 2.8, oppure si può scambiare l'integrale con lo sviluppo in serie di e^{-ipx} per l'esercizio 2.7).

3.15 Le *funzioni di Hermite* $\psi_n \in L^2(\mathbb{R})$, $n \in \mathbb{N}$, sono definite applicando il procedimento di Gram-Schmidt alle funzioni $x \in \mathbb{R} \mapsto x^n e^{-x^2/2}$ e pertanto hanno la forma $\psi_n(x) = H_n(x) e^{-x^2/2}$, con gli H_n polinomi di grado n, detti *polinomi di Hermite*. Dimostrare:

(a) $\{x^n e^{-x^2/2} : n \in \mathbb{N}\}^\perp = \{0\}$ (sugg.: se $\psi \in \{x^n e^{-x^2/2} : n \in \mathbb{N}\}^\perp$, applicare l'esercizio 3.14 a $f = \psi e^{-x^2/2}$);

(b) $(\psi_n)_{n\in\mathbb{N}}$ è una base ortonormale in $L^2(\mathbb{R})$;

(c) vale la *formula di Rodrigues*:

$$H_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}\pi^{1/4}} e^{x^2} \frac{d^n}{dx^n}(e^{-x^2})$$

(sugg.: basta verificare, integrando per parti, che se H_n è dato dalla formula di sopra, allora $\langle x^k e^{-x^2/2}, H_n e^{-x^2/2} \rangle = 0$, $k = 0, 1, \ldots, n-1$, e $\|H_n e^{-x^2/2}\|_2 = 1$);

(d) si ha

$$\frac{1}{\sqrt{2}}\left(x - \frac{d}{dx}\right)\psi_n = \sqrt{n+1}\,\psi_{n+1}, \qquad \frac{1}{\sqrt{2}}\left(x + \frac{d}{dx}\right)\psi_n = \begin{cases} \sqrt{n}\,\psi_{n-1}, & n \geq 1, \\ 0, & n = 0, \end{cases}$$

$$\left(-\frac{d^2}{dx^2} + x^2\right)\psi_n = (2n+1)\psi_n.$$

3.16 Dimostrare che applicando il procedimento di Gram-Schmidt in $L^2([0,+\infty))$ alle funzioni $x \mapsto x^n e^{-x}$ si ottengono funzioni $x \mapsto L_n(x)e^{-x}$, con gli L_n polinomi di grado n detti *polinomi di Laguerre*, che formano una base ortonormale in $L^2([0,+\infty))$.

Capitolo 4

Algebre di Banach e C*-algebre commutative

Come noto dai corsi di Algebra Lineare, il teorema spettrale fornisce un'utilissima descrizione delle matrici hermitiane o, equivalentemente, degli operatori hermitiani su uno spazio di Hilbert finito-dimensionale. È inoltre noto che il teorema spettrale permette di definire un *calcolo funzionale* di un tale operatore A, cioè di associare ad ogni funzione f sullo spettro di A un nuovo operatore $f(A)$. E anzi l'esistenza di tale calcolo funzionale è in realtà equivalente al teorema spettrale (si veda anche l'esercizio 4.31).

Passando a operatori su spazi infinito-dimensionali, le algebre di Banach forniscono il contesto matematico naturale entro il quale formulare la "corretta" generalizzazione della nozione di *spettro* di una matrice, che, come vedremo nella sezione 4.1, non si riduce a quella di "insieme degli autovalori". È inoltre possibile definire, per un elemento di un'algebra di Banach, un calcolo funzionale (che non discuteremo) per un'opportuna classe di funzioni, quelle analitiche in un intorno dello spettro. Per ottenere però un calcolo funzionale definito sulla famiglia molto più vasta di *tutte* le funzioni continue sullo spettro è necessario restringersi a una classe particolare di algebre di Banach, le C*-algebre, che costituiscono una diretta generalizzazione degli operatori limitati su uno spazio di Hilbert (sezione 4.4).

Vedremo poi nel prossimo capitolo come il teorema spettrale per gli operatori autoaggiunti limitati su uno spazio di Hilbert, sebbene si possa dimostrare con metodi più elementari, segua in modo assolutamente naturale dal calcolo funzionale delle C*-algebre e trovi in esso la sua giusta collocazione.

4.1 Teoria spettrale per algebre di Banach

Cominciamo con alcune definizioni di carattere puramente algebrico.

Definizione 4.1.1. (a) Un'*algebra* $\mathscr{A}$ è uno spazio vettoriale in cui sia definita un'applicazione bilineare e associativa $(a, b) \in \mathscr{A} \times \mathscr{A} \mapsto ab \in \mathscr{A}$, detta *prodotto*. L'algebra $\mathscr{A}$ è detta *con identità* (o *con unità*) se esiste $\mathbb{1} \in \mathscr{A}$ tale che $\mathbb{1} \cdot a = a \cdot \mathbb{1} = a$, ed è detta *commutativa* (o *abeliana*) se $ab = ba$ per ogni $a, b \in \mathscr{A}$.

(b) Una **-algebra* è un'algebra $\mathscr{A}$ in cui sia definita un'applicazione $a \in \mathscr{A} \mapsto a^* \in \mathscr{A}$, detta *involuzione*, soddisfacente le seguenti proprietà:

(i) antilinearità: $(\alpha a + \beta b)^* = \bar{\alpha} a^* + \bar{\beta} b^*$ per ogni $a, b \in \mathscr{A}$, $\alpha, \beta \in \mathbb{C}$;

(ii) antimoltiplicatività: $(ab)^* = b^*a^*$ per ogni $a, b \in \mathscr{A}$;

(iii) involutività: $a^{**} = a$ per ogni $a \in \mathscr{A}$;

Se $\mathscr{A}$ è un'algebra, una sua *sottoalgebra* è un sottospazio vettoriale $\mathscr{B} \subset \mathscr{A}$ chiuso per il prodotto (cioè tale che $a, b \in \mathscr{B} \Rightarrow ab \in \mathscr{B}$), e se $\mathscr{A}$ è una *-algebra, una *-sottoalgebra è una sottoalgebra chiusa anche per l'involuzione.

Data un'algebra con identità $\mathscr{A}$, un elemento $a \in \mathscr{A}$ è detto invertibile se esiste $a^{-1} \in \mathscr{A}$ tale che $aa^{-1} = \mathbb{1} = a^{-1}a$. Se poi $\mathscr{A}$ è anche una *-algebra, un elemento $a \in \mathscr{A}$ è detto *normale* se $a^*a = aa^*$, *autoaggiunto* se $a^* = a$, e *positivo* se esiste $b \in \mathscr{A}$ tale che $a = b^*b$. Inoltre $p \in \mathscr{A}$ è un *proiettore* se è autoaggiunto e $p^2 = p$, e infine $u \in \mathscr{A}$ è *unitario* se $u^*u = \mathbb{1} = uu^*$.

Un ruolo importante nella teoria è ovviamente giocato dalle applicazioni che rispettano le operazioni algebriche ora introdotte.

Definizione 4.1.2. (a) Siano $\mathscr{A}$, $\mathscr{B}$ algebre. Un *omomorfismo (di algebre)* da $\mathscr{A}$ a $\mathscr{B}$ è un'applicazione lineare $\rho : \mathscr{A} \to \mathscr{B}$ che rispetta il prodotto: $\rho(ab) = \rho(a)\rho(b)$, $a, b \in \mathscr{A}$. Se $\mathscr{A}$ e $\mathscr{B}$ hanno unità $\mathbb{1}_\mathscr{A}$ e $\mathbb{1}_\mathscr{B}$, ρ è detto *unitale* se $\rho(\mathbb{1}_\mathscr{A}) = \mathbb{1}_\mathscr{B}$. Un omomorfismo biunivoco è detto un *isomorfismo*.

(b) Siano $\mathscr{A}$, $\mathscr{B}$ *-algebre. Uno **-omomorfismo* da $\mathscr{A}$ a $\mathscr{B}$ è un omomorfismo di algebre $\rho : \mathscr{A} \to \mathscr{B}$ che rispetta anche l'involuzione: $\rho(a^*) = \rho(a)^*$, $a \in \mathscr{A}$. Uno *-omomorfismo biunivoco è detto uno **-isomorfismo*.

Nel seguito, noi saremo interessati pressoché esclusivamente ad algebre con una struttura topologica addizionale, definita da una norma compatibile con le operazioni algebriche, nel senso seguente.

Definizione 4.1.3. (a) Un'*algebra di Banach* è uno spazio di Banach $(\mathscr{A}, \|\cdot\|)$ tale che $\mathscr{A}$ sia un'algebra, e si abbia $\|ab\| \leq \|a\|\|b\|$, $a, b \in \mathscr{A}$.

(b) Una **-algebra di Banach* è un'algebra di Banach $(\mathscr{A}, \|\cdot\|)$ tale che $\mathscr{A}$ sia una *-algebra, e l'involuzione sia isometrica: $\|a^*\| = \|a\|$, $a \in \mathscr{A}$.

(c) Una *C*-algebra* è una *-algebra di Banach $(\mathscr{A}, \|\cdot\|)$ in cui valga l'*identità C**: $\|a^*a\| = \|a\|^2$ per ogni $a \in \mathscr{A}$.

Se $\mathscr{A}$ è una C*-algebra, una sua C*-sottoalgebra sarà una *-sottoalgebra $\mathscr{B} \subset \mathscr{A}$ chiusa in norma.

Esempi 4.1.4. (a) $\mathscr{A} = \mathbb{C}$ con la norma data dal modulo di un numero complesso, le usuali operazioni, e l'involuzione definita dal complesso coniugato ($z^* := \bar{z}$ per ogni $z \in \mathbb{C}$), è chiaramente il più semplice esempio di C*-algebra commutativa con identità (poiché $|z|^2 = z^*z = |z^*z|$). Come subito si verifica, gli elementi invertibili sono i numeri complessi non nulli, ogni $z \in \mathbb{C}$ è normale (poiché l'algebra è commutativa!), gli elementi autoaggiunti sono i numeri reali e quelli positivi i reali non negativi, gli unici proiettori sono i numeri 0 e 1, e gli unitari sono i numeri complessi di modulo 1.

(b) Sia X uno spazio compatto di Hausdorff. Lo spazio di Banach $(C(X), \|\cdot\|_\infty)$ diventa una C*-algebra commutativa con identità definendo il prodotto e l'involuzione come

$$(f \cdot g)(x) := f(x)g(x), \qquad f^*(x) := \overline{f(x)}, \qquad x \in X.$$

Si verifica infatti facilmente che il prodotto è bilineare, commutativo e con identità $\mathbb{1}(x) := 1$, $x \in X$, e che l'involuzione è antilineare, antimoltiplicativa, involutiva e isometrica. Inoltre essendo $|f(x)g(x)| \leq \|f\|_\infty \|g\|_\infty$ per ogni $x \in X$, si ha anche $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$, e infine l'identità C*:

$$\|f^*f\|_\infty = \sup_{x\in X} |f^*(x)f(x)| = \sup_{x\in X} |f(x)|^2 = \|f\|_\infty^2.$$

È poi chiaro che le $f \in C(X)$ invertibili sono quelle che non si annullano mai, quelle autoaggiunte (positive) sono quelle a valori reali (non negativi), e che gli unitari sono le funzioni a valori nel cerchio unitario $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. Infine i proiettori sono le f che assumono valori in $\{0, 1\}$, cioè le funzioni caratteristiche continue. Il che mostra in particolare che non tutte le C*-algebre ammettono proiettori non banali (cioè diversi da 0 e da $\mathbb{1}$): se $X = [0, 1]$ è chiaro che nessuna funzione caratteristica (diversa da $\mathbb{1}$ e dalla funzione identicamente nulla) può stare in $C([0, 1])$ (mentre ad esempio ce ne sono esattamente due non banali in $C([0, 1] \cup [2, 3])$: $\chi_{[0,1]}$ e $\chi_{[2,3]}$). Si noti anche che l'esempio precedente è il caso particolare in cui X è ridotto a un punto.

(c) Sia $(X, \mathfrak{M}, \mu)$ uno spazio di misura. Definendo prodotto e involuzione come nell'esempio precedente, si verifica facilmente, analogamente a quanto lì fatto, che $(L^\infty(X, \mu), \|\cdot\|_\infty)$ diventa una C*-algebra commutativa con identità. Si noti che in questo caso i proiettori di $L^\infty(X, \mu)$ sono le funzioni caratteristiche *misurabili*, e quindi in generale ce n'è grande abbondanza.

(d) Sia H uno spazio di Hilbert, e $\mathscr{A} \subset B(H)$ un sottospazio chiuso in norma tale che $S, T \in \mathscr{A} \Rightarrow S^*, ST \in \mathscr{A}$. È allora chiaro, dal teorema 3.4.3 e dai successivi commenti, che $\mathscr{A}$ è una C*-algebra (non commutativa, in generale), e anzi una C*-sottoalgebra della C*-algebra $B(H)$, e che i suoi proiettori sono precisamente i proiettori ortogonali di $B(H)$ in essa contenuti. Notiamo anche che mentre un'isometria $V \in B(H)$ è automaticamente un unitario se H è finito-dimensionale (cioè $V^*V = \mathbb{1} \Rightarrow VV^* = \mathbb{1}$), questo non è necessariamente vero se H è infinito-dimensionale (esercizio 4.2).

Vedremo in seguito (teorema 4.4.1 dell'isomorfismo di Gelfand, e teorema **??** di Gelfand-Naimark) che gli esempi 4.1.4(b) e (d) sono il prototipo, rispettivamente, delle C*-algebre commutative (con identità) e non commutative. Un esempio di *-algebra di Banach che non è una C*-algebra è dato dall'esercizio 4.3.

Introduciamo ora le nozioni fondamentali della teoria spettrale nelle algebre di Banach.

Definizione 4.1.5. Sia $\mathscr{A}$ un'algebra di Banach con identità. L'*insieme risolvente* di un $a \in \mathscr{A}$ è il sottoinsieme di $\mathbb{C}$

$$\rho(a) := \{\lambda \in \mathbb{C} : a - \lambda\mathbb{1} \text{ è invertibile in } \mathscr{A}\}.$$

Lo *spettro* di $a \in \mathscr{A}$ è l'insieme $\sigma(a) := \mathbb{C} \setminus \rho(a) = \{\lambda \in \mathbb{C} : a - \lambda\mathbb{1} \text{ non è invertibile in } \mathscr{A}\}$.

Per brevità, quando non crei confusione, scriveremo a volte $a - \lambda$ invece di $a - \lambda\mathbb{1}$.

Esempi 4.1.6. (a) Nell'algebra di Banach $\mathscr{A} = C(X)$, con X spazio compatto di Hausdorff, si ha, per ogni $f \in C(X)$, $\sigma(f) = f(X)$, l'insieme dei valori assunti da f su X. Per verificarlo basta dimostrare che $0 \in \sigma(f) \Leftrightarrow 0 \in f(X)$: infatti chiaramente $\sigma(f - \lambda\mathbb{1}) = \sigma(f) - \lambda = \{\mu - \lambda : \mu \in \sigma(f)\}$ (esercizio 4.4), e di conseguenza $\lambda \in \sigma(f) \Leftrightarrow 0 \in \sigma(f - \lambda\mathbb{1}) \Leftrightarrow 0 \in (f - \lambda\mathbb{1})(X) \Leftrightarrow \lambda \in f(X)$. Mostriamo allora che effettivamente $0 \in \sigma(f) \Leftrightarrow 0 \in f(X)$. Se $0 \in f(X)$ esisterà $x_0 \in X$ tale che $f(x_0) = 0$ e dunque f non è invertibile, per cui $0 \in \sigma(f)$. Viceversa se f non è invertibile, e cioè $0 \in \sigma(f)$, se per assurdo $0 \notin f(X)$ la funzione $g = \frac{1}{f}$ sarebbe continua su X, e quindi un inversa di f in $C(X)$.

(b) Nell'algebra di Banach $\mathscr{A} = M_n(\mathbb{C}) = B(\mathbb{C}^n)$ delle matrici complesse $n \times n$ (ricordiamo che su uno spazio finito-dimensionale ogni operatore lineare è limitato, esercizio 1.21), data $a \in \mathscr{A}$, $a - \lambda\mathbb{1}$ non è invertibile se e solo se $\det(a - \lambda\mathbb{1}) = 0$, e dunque lo spettro $\sigma(a)$ come definito sopra coincide con lo spettro come definito nei corsi di Algebra Lineare, cioè con l'insieme degli autovalori di a, cioè ancora con l'insieme $\{\lambda \in \mathbb{C} : \exists x \in \mathbb{C}^n \setminus \{0\} : Ax = \lambda x\} = \{\lambda \in \mathbb{C} : \ker(a - \lambda\mathbb{1}) \neq \{0\}\}$. Notiamo che in uno spazio finito-dimensionale ($\mathbb{C}^n$) si ha

$$\ker(a - \lambda\mathbb{1}) = \{0\} \Leftrightarrow \operatorname{ran}(a - \lambda\mathbb{1}) = \mathbb{C}^n,$$

e dunque la condizione $\ker(a - \lambda\mathbb{1}) = \{0\}$ equivale all'invertibilità di $a - \lambda\mathbb{1}$. In uno spazio di Hilbert H a dimensione infinita può invece accadere, come mostra l'esempio seguente, che $\ker(a - \lambda\mathbb{1}) = \{0\}$ e $\operatorname{ran}(a - \lambda\mathbb{1}) \subsetneq H$, e dunque $a - \lambda\mathbb{1}$, non essendo suriettivo, non sarà invertibile, e quindi $\lambda \in \sigma(a)$ pur non essendo λ un autovalore di a.

(c) Sia $H = L^2([0,1])$ (con la misura di Lebesgue), e consideriamo l'operatore $A : H \to H$ definito da

$$(A\psi)(x) = x\psi(x), \qquad x \in [0,1],\ \psi \in L^2([0,1]).$$

Essendo $\|A\psi\|_2^2 = \int_0^1 |x\psi(x)|^2 dx \leq \int_0^1 |\psi(x)|^2 dx = \|\psi\|_2^2$ si ha $A \in B(H)$ e $\|A\| \leq 1$. Mostriamo che A, in quanto elemento dell'algebra di Banach (con identità) $B(H)$, ha spettro $\sigma(A) = [0,1]$. Se infatti $\lambda \in [0,1]$ si ha che la funzione $\psi_0(x) = 1$, $x \in [0,1]$, non appartiene a $\operatorname{ran}(A - \lambda\mathbb{1})$: se così non fosse, dovrebbe esistere $\psi \in L^2([0,1])$ tale che

$$1 = [(A - \lambda\mathbb{1})\psi](x) = (x - \lambda)\psi(x), \qquad \text{per q.o. } x \in [0,1],$$

da cui $\psi(x) = \frac{1}{x-\lambda}$ per q.o. $x \in [0,1]$, ma tale funzione non è chiaramente in $L^2([0,1])$. Dunque $A - \lambda\mathbb{1}$, non essendo suriettivo, non è invertibile e $\lambda \in \sigma(A)$. Viceversa se $\lambda \notin [0,1]$, data $\varphi \in L^2([0,1])$ per determinare, se esiste, $\psi = (A - \lambda\mathbb{1})^{-1}\varphi$ occorre risolvere l'equazione

$$\varphi(x) = [(A - \lambda\mathbb{1})\psi](x) = (x - \lambda)\psi(x), \qquad \text{per q.o. } x \in [0,1],$$

che ha chiaramente come soluzione unica $\psi(x) = \frac{\varphi(x)}{x-\lambda}$ per q.o. $x \in [0,1]$, e si ha

$$\|\psi\|_2^2 = \int_0^1 \frac{|\varphi(x)|^2}{|x-\lambda|^2} dx \leq \frac{1}{\operatorname{dist}(\lambda, [0,1])^2} \|\varphi\|_2^2 < +\infty$$

(essendo $[0,1]$ chiuso e $\lambda \notin [0,1]$, si ha $\operatorname{dist}(\lambda, [0,1]) > 0$). Dunque $A - \lambda\mathbb{1}$ è invertibile e $\|(A - \lambda\mathbb{1})^{-1}\| \leq 1/\operatorname{dist}(\lambda, [0,1])$, per cui $(A - \lambda\mathbb{1})^{-1} \in B(H)$ e $\lambda \notin \sigma(A)$.
D'altra parte è facile vedere che A non ha autovalori, cioè $\ker(A - \lambda\mathbb{1}) = \{0\}$ per ogni $\lambda \in \mathbb{C}$. Infatti se così non fosse dovrebbe esistere $\psi \in L^2([0,1])$ tale che $(A - \lambda\mathbb{1})\psi(x) = (x-\lambda)\psi(x) = 0$, da cui $\psi(x) = 0$ per q.o. $x \in [0,1] \setminus \{\lambda\}$ che è un insieme il cui complementare ha misura nulla in $[0,1]$, e dunque $\psi = 0$ come elemento di $L^2([0,1])$.

L'operatore nell'ultimo esempio è un caso particolare dell'importante classe degli *operatori di moltiplicazione*, che come vedremo sono i prototipi di tutti gli operatori autoaggiunti, e il cui spettro è studiato nell'esercizio 4.5.

Come appena visto, lo spettro di un operatore su uno spazio di Hilbert può differire drasticamente dall'insieme dei suoi autovalori. Conviene allora introdurre una terminologia apposita. Dato $A \in B(H)$, H spazio di Hilbert, il suo *spettro puntuale* sarà

$$\sigma_p(A) := \{\lambda \in \mathbb{C} : \ker(A - \lambda\mathbb{1}) \neq \{0\}\},$$

cioè appunto l'insieme dei suoi autovalori. È chiaro che $\sigma_p(A) \subset \sigma(A)$, ma, come visto, può essere $\sigma_p(A) = \emptyset$. È altresì chiaro che una tale nozione non ha senso per elementi di un'algebra di Banach arbitraria.

Nello studio degli aspetti principali della teoria spettrale nelle algebre di Banach, riveste un ruolo fondamentale la teoria delle funzioni analitiche a valori in spazi di Banach, la cui definizione è la naturale generalizzazione di quelle a valori complessi.

Definizione 4.1.7. Siano $A \subset \mathbb{C}$ un aperto e X uno spazio di Banach. Una funzione $f : A \to X$ è detta *analitica* (o *olomorfa*) in A se per ogni $z \in A$ esiste

$$f'(z) := \lim_{h \to 0} \frac{f(h+z) - f(z)}{h}$$

nella topologia di X indotta dalla norma.

I principali teoremi della teoria delle funzioni analitiche a valori in $\mathbb{C}$ (che supponiamo nota ai lettori), con le relative dimostrazioni, si generalizzano immediatamente alle funzioni analitiche a valori in spazi di Banach, e pertanto non li studieremo qui in dettaglio. Ci limiteremo a dare gli enunciati di due teoremi che ci saranno utili a breve.

Teorema 4.1.8 (Liouville)**.** *Sia X uno spazio di Banach. Una $f : \mathbb{C} \to X$ analitica e limitata (cioè tale che $\|f(z)\| \leq M$ per ogni $z \in \mathbb{C}$, con $M > 0$ costante opportuna) è costante.*

Teorema 4.1.9 (dello sviluppo di Taylor)**.** *Siano $f : A \subset \mathbb{C} \to X$ analitica. Allora f è indefinitamente derivabile in senso complesso in A e per ogni $z_0 \in A$ si ha lo sviluppo di Taylor*

$$f(z) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

con raggio di convergenza $r \geq \operatorname{dist}(z_0, \partial A)$.

Ovviamente, per una serie di potenze $\sum_{n=0}^{+\infty} x_n(z - z_0)^n$ a coefficienti $x_n \in X$ spazio di Banach, il raggio di convergenza r si calcola, come per le serie a coefficienti complessi, tramite la formula di Hadamard

$$r = \frac{1}{\limsup_{n \to +\infty} \|x_n\|^{1/n}}.$$

Lemma 4.1.10. *Siano $\mathscr{A}$ un'algebra di Banach con identità, e $a \in \mathscr{A}$. Allora:*

(i) $\rho(a) \subset \mathbb{C}$ è aperto e la funzione $r_a : \rho(a) \to \mathscr{A}$ definita da $r_a(\lambda) := (\lambda\mathbb{1} - a)^{-1}$, $\lambda \in \rho(a)$, è analitica;

(ii) $\sigma(a) \subset \{z \in \mathbb{C} : |z| \leq \|a\|\}$ e $\sigma(a) \neq \emptyset$.

Dimostrazione. (i) Osserviamo per cominciare che dato $b \in \mathscr{A}$, $\|b\| < 1$, la *serie di Neumann*

$$\sum_{n=0}^{+\infty} b^n$$

(dove ovviamente $b^0 := \mathbb{1}$) converge assolutamente (poiché $\sum_{n=0}^{+\infty} \|b^n\| \leq \sum_{n=0}^{+\infty} \|b\|^n$, serie geometrica convergente) e dunque converge in $\mathscr{A}$ (esercizio 4.8(a)), e la sua somma è $(\mathbb{1} - b)^{-1}$, in quanto

$$(\mathbb{1} - b) \sum_{n=0}^{N} b^n = \sum_{n=0}^{N} b^n (\mathbb{1} - b) = \mathbb{1} - b^{N+1} \to \mathbb{1} \quad \text{per } N \to +\infty.$$

Dato allora $\lambda_0 \in \rho(a)$, si avrà

$$\lambda \mathbb{1} - a = (\lambda - \lambda_0)\mathbb{1} + (\lambda_0 \mathbb{1} - a) = (\lambda_0 \mathbb{1} - a)(\mathbb{1} - (\lambda_0 - \lambda) r_a(\lambda_0)),$$

e dunque se $|\lambda - \lambda_0| < \|r_a(\lambda_0)\|^{-1}$ per l'osservazione appena fatta, applicata a $b = (\lambda_0 - \lambda) r_a(\lambda_0)$, esisterà

$$(\lambda \mathbb{1} - a)^{-1} = \left[\sum_{n=0}^{+\infty} (\lambda_0 - \lambda)^n r_a(\lambda_0)^n \right] r_a(\lambda_0) = \sum_{n=0}^{+\infty} (\lambda_0 - \lambda)^n r_a(\lambda_0)^{n+1}. \tag{4.1}$$

In particolare ne segue che $\rho(a)$, che contiene un intorno di raggio $\|r_a(\lambda_0)\|^{-1}$ del suo generico punto λ_0, è aperto, e che $r_a(\lambda) = (\lambda \mathbb{1} - a)^{-1}$, che in tale intorno si esprime tramite la serie di potenze (4.1), è analitica.

(ii) Dato $\lambda \in \mathbb{C}$, $|\lambda| > \|a\|$, $\lambda \mathbb{1} - a = \lambda(\mathbb{1} - \lambda^{-1} a)$ è invertibile in $\mathscr{A}$, ancora per l'osservazione precedente, applicata a $b = \lambda^{-1} a$. Dunque $\lambda \in \rho(a)$ e $\sigma(a) \subset \{z \in \mathbb{C} : |z| \leq \|a\|\}$. Inoltre, sempre per $|\lambda| > \|a\|$,

$$\|r_a(\lambda)\| = \|\lambda^{-1}(\mathbb{1} - \lambda^{-1} a)^{-1}\| = |\lambda|^{-1} \left\| \sum_{n=0}^{+\infty} \lambda^{-n} a^n \right\| \leq \frac{|\lambda|^{-1}}{1 - |\lambda|^{-1}\|a\|} = \frac{1}{|\lambda| - \|a\|}.$$

Poiché allora l'ultimo membro è infinitesimo per $|\lambda| \to +\infty$, se fosse $\sigma(a) = \emptyset$ la funzione analitica $r_a : \rho(a) = \mathbb{C} \to \mathscr{A}$ sarebbe limitata, e dunque, per il teorema di Liouville, costante, e poiché, come appena visto, è infinitesima all'infinito, si concluderebbe che $(\lambda \mathbb{1} - a)^{-1} = 0$ per ogni $\lambda \in \mathbb{C}$, il che è assurdo, essendo 0 chiaramente non invertibile. $\square$

Dal lemma appena dimostrato segue in particolare che $\sigma(a)$ è un sottoinsieme compatto (non vuoto) di $\mathbb{C}$ per ogni $a \in \mathscr{A}$. L'elemento $r_a(\lambda) = (\lambda \mathbb{1} - a)^{-1}$ è detto il *risolvente* di $a \in \mathscr{A}$ (nel punto $\lambda \in \rho(a)$).

Definizione 4.1.11. Sia $\mathscr{A}$ un'algebra di Banach con identità. Il *raggio spettrale* di un elemento $a \in \mathscr{A}$ è il numero

$$r(a) := \sup\{|\lambda| \,:\, \lambda \in \sigma(a)\}.$$

Dal lemma precedente si ha allora $r(a) \leq \|a\|$. In realtà si può dare una formula esplicita per il calcolo del raggio spettrale, che sarà per noi fondamentale.

Teorema 4.1.12 (del raggio spettrale di Gelfand)**.** *Sia $\mathscr{A}$ un'algebra di Banach con identità. Si ha, per ogni $a \in \mathscr{A}$,*

$$r(a) = \lim_{n \to +\infty} \|a^n\|^{1/n}.$$

Dimostrazione. Siano $a \in \mathscr{A}$, $\lambda \in \mathbb{C}$. Posto $b := \sum_{k=0}^{n-1} \lambda^{n-1-k} a^k$, un semplice calcolo (identico a quello che si fa con i numeri reali) mostra che

$$\lambda^n - a^n = (\lambda - a)b = b(\lambda - a). \tag{4.2}$$

Da questo segue che se $\lambda \in \sigma(a)$ allora $\mu := \lambda^n \in \sigma(a^n)$: se infatti così non fosse, $\lambda^n - a^n$ sarebbe invertibile, e quindi dalla (4.2) si avrebbe

$$\mathbb{1} = [(\lambda^n - a^n)^{-1} b](\lambda - a) = (\lambda - a)[b(\lambda^n - a^n)^{-1}],$$

cioè, grazie all'esercizio 4.6 l'invertibilità di $\lambda - a$. Si ha dunque $\{|\lambda|^n \,:\, \lambda \in \sigma(a)\} \subset \{|\mu| \,:\, \mu \in \sigma(a^n)\}$, e pertanto, usando anche il lemma 4.1.10, $r(a)^n \leq r(a^n) \leq \|a^n\|$ per ogni $n \in \mathbb{N}$, e, ancora, $r(a) \leq \liminf_{n\to+\infty} \|a^n\|^{1/n}$.

Si consideri poi, per $|z| < \|a\|^{-1}$,

$$f(z) = (\mathbb{1} - za)^{-1} = \sum_{n=0}^{+\infty} z^n a^n. \tag{4.3}$$

Avendosi, per $z \neq 0$, $f(z) = z^{-1}(z^{-1} - a)^{-1} = z^{-1} r_a(z^{-1})$ ed essendo r_a analitica in $\rho(a)$, si vede che f è analitica nell'aperto $A := \{z \in \mathbb{C} \setminus \{0\} \,:\, z^{-1} \in \rho(a)\} \cup \{0\}$ (in generale più grande di $\{z \,:\, |z| < \|a\|^{-1}\}$). Dunque per il teorema dello sviluppo di Taylor, la serie (4.3) ha raggio di convergenza

$$r = \left(\limsup_{n\to+\infty} \|a^n\|^{1/n} \right)^{-1} \geq \operatorname{dist}(0, \partial A).$$

Se allora $r(a) = 0$ si ha $A = \mathbb{C}$ e $\limsup_{n\to+\infty} \|a^n\|^{1/n} = 0 = r(a)$. Se invece $r(a) > 0$ si ha, poiché $0 \in A$ e $\partial A \subset A^c$, $\operatorname{dist}(0, \partial A) = \operatorname{dist}(0, A^c)$, e poiché $z \in A^c$ se e solo se $z \neq 0$ e $\lambda := z^{-1} \in \sigma(a)$,

$$\operatorname{dist}(0, A^c) = \inf_{z \in A^c} |z| = \inf_{\lambda \in \sigma(a)\setminus\{0\}} |\lambda|^{-1} = r(a)^{-1}.$$

In definitiva, si sono ottenute le disuguaglianze

$$\limsup_{n\to+\infty} \|a^n\|^{1/n} \leq r(a) \leq \liminf_{n\to+\infty} \|a^n\|^{1/n},$$

che implicano la tesi (esercizio 4.9). $\square$

Gli elementi $a \in \mathscr{A}$ per cui $r(a) = 0$ sono detti *topologicamente nilpotenti* (ricordiamo che un operatore lineare a è nilpotente se $a^n = 0$ per qualche $n \in \mathbb{N}$). Un esempio classico è mostrato nell'esercizio 4.10, che fornisce anche un'interessante applicazione della teoria fin qui svolta alla risoluzione di certe equazioni integrali.

Come già accennato, la teoria spettrale per elementi di un'algebra di Banach culmina nella definizione di un *calcolo funzionale analitico*, cioè di un'omomorfismo $f \in H(\sigma(a)) \mapsto f(a) \in \mathscr{A}$, dove $H(\sigma(a))$ è l'algebra delle funzione analitiche in un intorno dello spettro di $a \in \mathscr{A}$, e dove $f(a)$ è definita tramite l'*integrale di Dunford-Schwartz*

$$f(a) = \frac{1}{2\pi i} \int_\gamma f(z)(z\mathbb{1} - a)^{-1}\, dz$$

(γ circuito attorno a $\sigma(a)$), chiara generalizzazione dell'integrale di Cauchy.

Noi non intraprenderemo questo studio, e ci dedicheremo invece allo studio della teoria spettrale nel contesto più ristretto delle C*-algebre, che, per elementi a normali, conduce alla formulazione di un calcolo funzionale sull'assai più ampia algebra $C(\sigma(a))$ di tutte le funzioni continue sullo spettro. Il primo passo in questa direzione è fornito dal teorema seguente, essenzialmente una conseguenza del teorema del raggio spettrale.

Teorema 4.1.13. *Sia $\mathscr{A}$ una C*-algebra con identità. Allora, dato $a \in \mathscr{A}$, si ha*

(i) $\sigma(a^*) = \overline{\sigma(a)}$;

(ii) se a è invertibile, $\sigma(a^{-1}) = \sigma(a)^{-1}$;

(iii) se a è normale, $r(a) = \|a\|$;

(iv) se $a = p$ è un proiettore, $\sigma(p) \subset \{0, 1\}$;

(v) se $a = u$ è unitario, $\sigma(u) \subset \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$.

(vi) se $a = a^*$, $\sigma(a) \subset \mathbb{R}$;

Dimostrazione. (i) È chiaro che a è invertibile se e solo se a^* è invertibile, e in tal caso $(a^*)^{-1} = (a^{-1})^*$. Dunque

$$\lambda \in \rho(a) \Leftrightarrow \exists (\lambda - a)^{-1} \Leftrightarrow \exists ((\lambda - a)^*)^{-1} = (\bar{\lambda} - a^*)^{-1} \Leftrightarrow \bar{\lambda} \in \rho(a^*).$$

(ii) Essendo a invertibile è $0 \notin \sigma(a)$. Dato allora $\lambda \neq 0$, si hanno le ovvie identità

$$\begin{aligned} \lambda - a &= -\lambda a(\lambda^{-1} - a^{-1}) = -(\lambda^{-1} - a^{-1})\lambda a, \\ \lambda^{-1} - a^{-1} &= -(\lambda - a)\lambda^{-1} a^{-1} = -\lambda^{-1} a^{-1} (\lambda - a), \end{aligned}$$

da cui segue, ragionando in modo analogo a quanto fatto all'inizio della dimostrazione del teorema 4.1.12, $\lambda \in \rho(a) \Leftrightarrow \lambda^{-1} \in \rho(a^{-1})$.

(iii) Si ha

$$\begin{aligned} \|a^2\|^2 &= \|(a^2)^* a^2\| = \|(a^* a)(a^* a)\| \\ &= \|(a^* a)^*(a^* a)\| = \|a^* a\|^2 = \|a\|^4 \end{aligned}$$

avendo usato l'identità C* nella prima, quarta e quinta uguaglianza, e la normalità di a nella seconda. Dunque $\|a^2\| = \|a\|^2$, e, per induzione, $\|a^{2^m}\| = \|a\|^{2^m}$ per ogni $m \in \mathbb{N}$ (infatti $\|a^{2^{m+1}}\| = \|(a^{2^m})^2\| = \|a^{2^m}\|^2 = \|a\|^{2^{m+1}}$ usando il fatto che a^{2^m} è normale). Ma allora dal teorema del raggio spettrale segue

$$\|a\| = \lim_{m \to +\infty} \|a^{2^m}\|^{1/2^m} = r(a).$$

(iv) Esercizio 4.11.

(v) Si ha

$$1 = \|\mathbb{1}\| = \|u^* u\| = \|u\|^2,$$

avendo usato l'esercizio 4.1 e l'identità C*. Dunque $\|u\| = \|u^*\| = 1$ e $\sigma(u) \subset \{z \in \mathbb{C} : |z| \leq 1\}$. D'altra parte, essendo u invertibile, per il punto (ii) si ha anche $\sigma(u)^{-1} = \sigma(u^{-1}) = \sigma(u^*) \subset \{z \in \mathbb{C} : |z| \leq 1\}$, da cui $\sigma(u) \subset \{z \in \mathbb{C} : |z| \geq 1\}$ e quindi $\sigma(u) \subset \mathbb{T}$.

(vi) In base all'esercizio 4.12(a), si può considerare l'elemento $u := e^{ia} = \sum_{n=0}^{+\infty} \frac{i^n}{n!} a^n \in \mathscr{A}$, e si ha, per la continuità dell'involuzione, $u^* = \sum_{n=0}^{+\infty} \frac{(-i)^n}{n!} a^n = e^{-ia}$, e dunque, per l'esercizio 4.12(b), $uu^* = e^{ia-ia} = \mathbb{1} = u^*u$, cioè u è unitario. Valgono inoltre le identità

$$u - e^{i\lambda}\mathbb{1} = (e^{i(a-\lambda)} - \mathbb{1})e^{i\lambda} = (a - \lambda)b = b(a - \lambda),$$

dove si è posto $b := \sum_{n=1}^{+\infty} \frac{i^n}{n!} e^{i\lambda}(a - \lambda)^{n-1}$. Da questo segue che se $\lambda \notin \mathbb{R}$, cioè $e^{i\lambda} \notin \mathbb{T}$, per il punto (v) $u - e^{i\lambda}\mathbb{1}$ è invertibile, e quindi anche $a - \lambda$ lo è, e pertanto $\lambda \in \rho(a)$. □

Una conseguenza semplice ma notevole del teorema appena dimostrato è che la norma di una C*-algebra è determinata univocamente dalla struttura algebrica: infatti dato $a \in \mathscr{A}$, essendo a^*a autoaggiunto, e quindi normale, si ha $\|a\| = \|a^*a\|^{1/2} = r(a^*a)^{1/2}$, e d'altra parte il raggio spettrale di un elemento x è definito in modo puramente algebrico (senza cioè far intervenire la norma), come l'estremo superiore dei moduli dei $\lambda \in \mathbb{C}$ tali che $\lambda - x$ non sia invertibile. Quindi ogni C*-algebra è tale rispetto a un'unica norma (mentre, come visto, su uno spazio di Banach possono esistere molte norme equivalenti).

È questo un primo esempio della notevole rigidità della struttura delle C*-algebre, che ha molte importanti e interessanti conseguenze.

4.2 Caratteri e ideali

Il teorema del raggio spettrale visto alla fine della sezione precedente è uno degli ingredienti fondamentali nella dimostrazione del principale teorema della teoria delle algebre di Banach commutative, il già menzionato teorema dell'isomorfismo di Gelfand 4.4.1, secondo il quale ogni C*-algebra commutativa con unità è isometricamente *-isomorfa alla C*-algebra delle funzioni continue su un opportuno spazio compatto di Hasudorff.

Come primo passo verso la dimostrazione di questo teorema, e allo scopo di comprendere come si possa costruire in modo naturale uno spazio compatto a partire da una C*-algebra commutativa, consideriamo proprio il caso dell'algebra $\mathscr{A} = C(X)$, con X spazio compatto di Hausdorff. Ovviamente in questo caso lo spazio compatto di cui il teorema di Gelfand asserisce l'esistenza dovrà coincidere con X stesso. Si pone quindi il problema di associare ai punti $x \in X$ opportuni oggetti che siano descrivibili in termini puramente algebrici. L'osservazione centrale a questo proposito è che ad ogni $x \in X$ possiamo associare il funzionale $\delta_x \in \mathscr{A}^*$, $\delta_x(f) := f(x)$, che oltre ad essere lineare è anche *moltiplicativo*:

$$\delta_x(fg) = f(x)g(x) = \delta_x(f)\delta_x(g), \qquad f, g \in \mathscr{A}.$$

Si ottiene in tal modo un'applicazione $x \in X \mapsto \delta_x$ nell'insieme dei funzionali moltiplicativi di $\mathscr{A}$. Usando inoltre un risultato classico di topologia noto come lemma di Urysohn (per il quale rimandiamo ad esempio a [Rud2, Thm. 2.12]), si ottiene che $C(X)$ *separa i punti di* X, cioè per ogni $x, y \in X$ distinti, esiste $f \in C(X)$ tale che $f(x) \neq f(y)$ (la facile dimostrazione di questo fatto nel caso in cui X è uno spazio metrico è lasciata al lettore, esercizio 4.13). Questo ha come conseguenza che l'applicazione $x \mapsto \delta_x$ è biunivoca da X sull'insieme dei funzionali moltiplicativi non nulli di $\mathscr{A} = C(X)$ (esercizio 4.14).

Tutto ciò motiva la seguente fondamentale definzione.

Definizione 4.2.1. Sia $\mathscr{A}$ un'algebra di Banach commutativa con unità. Un *carattere* di $\mathscr{A}$ è un funzionale $\omega : \mathscr{A} \to \mathbb{C}$ lineare, moltiplicativo e non nullo. L'insieme $\Omega(\mathscr{A})$ dei caratteri di $\mathscr{A}$, dotato della topologia in cui gli aperti sono unioni arbitarie di insiemi della forma

$$B(a_1, \dots, a_n; V_1, \dots, V_n) := \{\omega \in \Omega(\mathscr{A}) \,:\, \omega(a_1) \in V_1, \dots, \omega(a_n) \in V_n\},$$

con $a_i \in \mathscr{A}$, e $V_i \subset \mathbb{C}$ aperto, $i = 1, \dots, n$, $n \in \mathbb{N}$, è detto lo *spettro di Gelfand* di $\mathscr{A}$.

La semplice verifica del fatto che quella data nella definizione sia effettivamente una topologia di Hausdorff su $\Omega(\mathscr{A})$ è lasciata al lettore, così come una sua possibile definizione alternativa in termini di net (esercizio 4.15). Si può poi dimostrare che $\Omega(\mathscr{A})$, con questa topologia, è uno spazio compatto [Ped, Thm. 4.2.3]. Nel caso di $\mathscr{A} = C(X)$ considerato sopra, si vede inoltre facilmente che l'applicazione $x \in X \mapsto \delta_x \in \Omega(C(X))$ è continua insieme con la sua inversa (esercizio 4.16), cioè è quello che si dice un *omeomorfismo* degli spazi topologici X e $\Omega(C(X))$, che possono quindi essere identificati.

Possiamo a questo punto associare ad ogni $a \in \mathscr{A}$ una funzione su $\Omega(\mathscr{A})$ al modo seguente.

Definizione 4.2.2. Sia $\mathscr{A}$ un'algebra di Banach commutativa con unità. La *trasformata di Gelfand* è l'applicazione $\hat{\ } : \mathscr{A} \to C(\Omega(\mathscr{A}))$ definita da

$$\hat{a}(\omega) := \omega(a), \qquad a \in \mathscr{A}, \omega \in \Omega(\mathscr{A}).$$

È facile verificare che effettivamente la funzione $\hat{a} : \Omega(\mathscr{A}) \to \mathbb{C}$ così definita è continua su $\Omega(\mathscr{A})$ (esercizio 4.17). L'esercizio 4.18, per la soluzione del quale possono essere utili alcuni dei risultati dimostrati più avanti in questa sezione, mostra che la trasformata di Gelfand è una generalizzazione della serie (o trasformata) di Fourier.

Il contenuto del teorema dell'isomorfismo di Gelfand sarà dunque la dimostrazione del fatto che, se $\mathscr{A}$ è una C*-algebra commutativa con unità, la trasformata di Gelfand è uno *-isomorfismo isometrico.

Al fine di dimostrare l'isometria (e quindi l'iniettività) della trasformata di Gelfand, l'altro strumento fondamentale, oltre al teorema del raggio spettrale, è la teoria di Gelfand dei caratteri di un'algebra di Banach $\mathscr{A}$, che ne consente una descrizione in termini di opportuni sottoinsiemi di $\mathscr{A}$ che ora introdurremo.

Definizione 4.2.3. Un *ideale (bilatero)* di un'algebra $\mathscr{A}$ è una sottoalgebra $\mathscr{I} \subset \mathscr{A}$ tale che per ogni $a \in \mathscr{I}$ e $b \in \mathscr{A}$ si abbia $ab, ba \in \mathscr{I}$.

Ovviamente un'algebra $\mathscr{A}$ ha sempre i due ideali $\mathscr{I} = \{0\}$ (l'ideale banale) e $\mathscr{I} = \mathscr{A}$. Un ideale $\mathscr{I}$ è detto *proprio* se $\mathscr{I} \neq \mathscr{A}$.

Proposizione 4.2.4. *Siano $\mathscr{A}$ un'algebra con unità e $\mathscr{I} \subset \mathscr{A}$ un ideale. Allora:*

(i) $\mathscr{I}$ è proprio se e solo se $\mathbb{1} \notin \mathscr{I}$;

(ii) se $\mathscr{A}$ è un algebra di Banach e $\mathscr{I}$ è proprio, allora la sua chiusura $\bar{\mathscr{I}}$ è un ideale proprio.

Dimostrazione. (i) È ovvio che se $\mathbb{1} \notin \mathscr{I}$, $\mathscr{I}$ è proprio. Viceversa se per assurdo $\mathbb{1} \in \mathscr{I}$, per ogni $a \in \mathscr{A}$ si avrebbe $a = \mathbb{1}a \in \mathscr{I}$ e dunque $\mathscr{I} = \mathscr{A}$ non sarebbe un ideale proprio.

(ii) È immediato verificare che $\bar{\mathscr{I}}$ è un ideale (esercizio 4.19). Se per assurdo non fosse proprio, esisterebbe $(a_n)_{n\in\mathbb{N}} \subset \mathscr{I}$ convergente a $\mathbb{1}$, e dunque si avrebbe definitivamente $\|\mathbb{1} - a_n\| < 1$. Ma allora, per quanto visto all'inizio della dimostrazione del Lemma 4.1.10(i), $a_n = \mathbb{1} - (\mathbb{1} - a_n)$ sarebbe invertibile, e dunque $\mathbb{1} = a_n^{-1}a_n \in \mathscr{I}$, e $\mathscr{I}$ non sarebbe proprio per il punto (i). $\square$

Dal punto (i) della proposizione precedente segue che se $\mathscr{A}$ è un'algebra commutativa, c'è uno stretto legame tra i suoi elementi invertibili e la sua struttura di ideali, in quanto chiaramente $a \in \mathscr{A}$ non è invertibile se e solo se $\mathscr{I} := a\mathscr{A} = \mathscr{A}a$ è un ideale proprio. E poiché lo spettro di un $a \in \mathscr{A}$ dipende da quali sono gli elementi invertibili di $\mathscr{A}$ è naturale attendersi un legame tra lo spettro e gli ideali di $\mathscr{A}$, che è quanto andremo ora a discutere.

Per iniziare, ricordiamo che dato uno spazio vettoriale X con un sottospazio $M \subset X$, resta definita una relazione d'equivalenza tra elementi di X, secondo la quale $x, y \in X$ sono equivalenti se e solo se $x - y \in M$. Le classi di tale relazione d'equivalenza si indicano con $x + M := \{x + y : y \in M\}$, l'insieme X/M delle classi di equivalenza ha una struttura naturale di spazio vettoriale (esercizio 4.20(a)), detto lo *spazio quoziente* di X rispetto a M, e l'applicazione lineare $\eta : X \to X/M$, $x \mapsto x + M$, è detta la *proiezione canonica*, ed è chiaramente suriettiva e tale che $\ker \eta = M$. Se inoltre $X = \mathscr{A}$ è un algebra (con unità $\mathbb{1}$) e $M = \mathscr{I}$ è un ideale, allora il quoziente $\mathscr{A}/\mathscr{I}$ è in modo naturale un'algebra (con unità $\mathbb{1} + \mathscr{I}$, esercizio 4.20(b)).

Come è facile attendersi, nel caso in cui X è uno spazio di Banach e M un sottospazio chiuso, il quoziente è naturalmente uno spazio di Banach, e un'algebra di Banach se tale è X.

Proposizione 4.2.5. *Siano X uno spazio normato e $M \subset X$ un sottospazio chiuso. Allora:*

(i) X/M è uno spazio normato con la norma

$$\|x + M\| := \inf_{y \in M} \|x + y\|$$

e la proiezione canonica $\eta : X \to X/M$ è limitata;

(ii) se X è di Banach, anche X/M lo è;

(iii) se $X = \mathscr{A}$ è un'algebra di Banach e $M = \mathscr{I}$ è un ideale chiuso, allora $\mathscr{A}/\mathscr{I}$ è un'algebra di Banach.

Dimostrazione. (i) Se $x + M = x' + M$ si ha

$$\inf_{y \in M} \|x + y\| = \inf_{y \in M} \|x' + (x - x' + y)\| = \inf_{y' \in M} \|x' + y'\|,$$

in quanto $y \in M$ se e solo se $y' := x - x' + y \in M$. Dunque la quantità $\|x + M\|$ è ben definita (cioè dipende solo dalla classe $x + M$ e non dal rappresentante $x \in x + M$). Resta allora da dimostrare che è una norma su X/M. Essendo chiaramente $\|x + M\| = \mathrm{dist}(x, M)$ ed M chiuso, lo stesso argomento usato nella dimostrazione della proposizione 1.5.1 mostra che se $\|x + M\| = 0$ allora $x \in M$, cioè $x + M = 0 + M$ è il vettore nullo di X/M. Inoltre chiaramente $\|0(x + M)\| = \|0 + M\| = 0 = 0\|x + M\|$, mentre se $\lambda \neq 0$

$$\begin{aligned}\|\lambda(x + M)\| &= \|\lambda x + M\| = \inf_{y \in M} \|\lambda x + y\| \\ &= |\lambda| \inf_{y \in M} \|x + \lambda^{-1} y\| = |\lambda| \inf_{y' \in M} \|x + y'\| = |\lambda| \|x + M\|.\end{aligned}$$

Infine dati $x + M, y + M \in X/M$ ed $\varepsilon > 0$, esisteranno $x', y' \in M$ tali che $\|x + x'\| < \|x + M\| + \varepsilon/2$, $\|y + y'\| < \|y + M\| + \varepsilon/2$, e pertanto

$$\begin{aligned}\|(x + M) + (y + M)\| &= \|x + y + M\| = \inf_{z \in M} \|x + y + z\| \leq \|x + y + x' + y'\| \\ &\leq \|x + x'\| + \|y + y'\| < \|x + M\| + \|y + M\| + \varepsilon,\end{aligned}$$

da cui, per l'arbitrarietà di ε, $\|(x+M)+(y+M)\| \leq \|x+M\| + \|y+M\|$. La limitatezza di η segue subito dall'ovvia disuguaglianza $\|x+M\| \leq \|x\|$.

(ii) Sia $(x_n+M)_{n\in\mathbb{N}} \subset X/M$ una successione di Cauchy. Procedendo come nella dimostrazione del teorema di Riesz-Fisher 2.3.4, si può da questa estrarre una sottosuccessione $(x_{n_k}+M)_{k\in\mathbb{N}}$ tale che

$$\|(x_{n_{k+1}}+M)-(x_{n_k}+M)\| < \frac{1}{2^{k+1}}, \qquad k \in \mathbb{N}.$$

Si possono allora scegliere vettori $y_k \in (x_{n_{k+1}} - x_{n_k}) + M$ tali che $\|y_k\| < \frac{1}{2^k}$, $k \in \mathbb{N}$, e si potrà dunque definire un vettore $z := x_{n_1} + \sum_{k=1}^{+\infty} y_k$, in quanto la serie è assolutamente convergente e X è di Banach. Si ottiene pertanto

$$\begin{aligned}\|(z+M)-(x_{n_k}+M)\| &= \Big\|(z+M) - \Big[(x_{n_1}+M) + \sum_{j=1}^{k-1}(x_{n_{j+1}}+M)-(x_{n_j}+M)\Big]\Big\| \\ &= \Big\|(z+M) - \Big[(x_{n_1}+M) + \sum_{j=1}^{k-1}(x_{n_{j+1}}-x_{n_j})+M\Big]\Big\| \\ &= \Big\|\Big(z - x_{n_1} - \sum_{j=1}^{k-1} y_j\Big) + M\Big\| \leq \Big\|z - x_{n_1} - \sum_{j=1}^{k-1} y_j\Big\|,\end{aligned}$$

avendo usato, per la terza uguaglianza, il fatto che $(x_{n_{j+1}} - x_{n_j}) + M = y_j + M$. Poiché allora, per la definizione di z, l'ultimo membro è infinitesimo per $k \to +\infty$, si ottiene che la sottosuccessione $(x_{n_k}+M)$ converge a $z+M$. E si conclude allora usando il fatto che una successione di Cauchy con una sottosuccessione convergente è convergente (esercizio 4.7).

(iii) L'unica cosa da verificare è che $\|(a+\mathscr{I})(b+\mathscr{I})\| \leq \|a+\mathscr{I}\|\|b+\mathscr{I}\|$. A tale scopo, dato $\varepsilon > 0$ esisteranno $a', b' \in \mathscr{I}$ tali che $\|a+a'\| < \|a+M\| + \varepsilon$, $\|b+b'\| < \|b+M\|+\varepsilon$. Essendo allora $a'b + ab' + a'b' \in \mathscr{I}$, si avrà

$$\begin{aligned}\|(a+\mathscr{I})(b+\mathscr{I})\| &= \|ab+\mathscr{I}\| = \inf_{c\in\mathscr{I}} \|ab+c\| \leq \|ab + a'b + ab' + a'b'\| \\ &= \|(a+a')(b+b')\| \leq \|a+a'\|\|b+b'\| < (\|a+\mathscr{I}\|+\varepsilon)(\|b+\mathscr{I}\|+\varepsilon),\end{aligned}$$

e si conclude grazie all'arbitrarietà di ε. □

Utilizzando lo spazio quoziente, è possibile dare un'ulteriore caratterizzazione dei funzionali limitati su uno spazio normato, che ci sarà presto utile.

Proposizione 4.2.6. *Sia X uno spazio normato. Un funzionale lineare $f : X \to \mathbb{C}$ è limitato se e solo se* $\ker f$ *è un sottospazio chiuso di X.*

Dimostrazione. La necessità della condizione è ovvia, in quanto $\ker f = f^{-1}(\{0\})$ è la controimmagine di un sottoinsieme chiuso di $\mathbb{C}$.

Per la sufficienza, si supponga $f \neq 0$ (altrimenti la tesi è banale), e posto $N := \ker f$ si consideri l'applicazione $\theta : X/N \to \mathbb{C}$ definita da $\theta(x+N) := f(x)$. Poiché $x - y \in N$ se e solo se $f(x) = f(y)$, θ è ben definita e iniettiva. Dato poi $x_0 \in X$ tale che $f(x_0) \neq 0$, per ogni $\alpha \in \mathbb{C}$ si ha $\theta(\frac{\alpha}{f(x_0)}x_0 + N) = f(\frac{\alpha}{f(x_0)}x_0) = \alpha$, e dunque θ è un isomorfismo di spazi vettoriali. Si può allora usare per trasportare su $\mathbb{C}$ la norma di X/N, ottenendo una norma $\alpha \in \mathbb{C} \mapsto \|\theta^{-1}(\alpha)\|$. Ma essendo $\mathbb{C}$ finito-dimensionale, dal teorema 1.3.7 segue l'esistenza di una costante $C > 0$ tale che $|\alpha| \leq C\|\theta^{-1}(\alpha)\|$ per ogni $\alpha \in \mathbb{C}$, o, se $\alpha = f(x)$ (e quindi $\theta^{-1}(\alpha) = x + N$),

$$|f(x)| \leq C\|x + N\| \leq C\|x\|,$$

cioè la limitatezza di f. □

Il prossimo risultato stabilisce un fondamentale legame tra i caratteri di un'algebra di Banach e la struttura dei suoi ideali.

Teorema 4.2.7 (di Gelfand-Mazur)**.** *Sia $\mathscr{A}$ un'algebra di Banach commutativa con unità. Esiste una corrispondenza biunivoca tra i caratteri di $\mathscr{A}$ e gli ideali propri massimali di $\mathscr{A}$, definita da*

$$\omega \in \Omega(\mathscr{A}) \mapsto \ker \omega.$$

Dimostrazione. Per iniziare, osserviamo che effettivamente se $\omega \in \Omega(\mathscr{A})$, $\mathscr{I} := \ker \omega$ è un ideale proprio massimale di $\mathscr{A}$. $\mathscr{I}$ è un sottospazio per la linearità di ω, ed è un ideale per la moltiplicatività: se $a \in \mathscr{I}$ e $b \in \mathscr{A}$, allora $\omega(ab) = \omega(a)\omega(b) = 0$ e dunque $ab \in \mathscr{I}$. Essendo poi $\omega(\mathbb{1}) = 1$, si ha $\mathbb{1} \notin \mathscr{I}$, e dunque $\mathscr{I}$ è proprio. Infine se $\mathscr{K} \supset \mathscr{I}$ è un ideale di $\mathscr{A}$, $\omega(\mathscr{K})$ è un sottospazio di $\mathbb{C}$, e dunque $\omega(\mathscr{K}) = \{0\}$ o $\omega(\mathscr{K}) = \mathbb{C}$. Nel primo caso $\mathscr{K} \subset \omega^{-1}(\omega(\mathscr{K})) = \omega^{-1}(\{0\}) = \mathscr{I}$, mentre nel secondo caso dato $a \in \mathscr{A}$ esiste $b \in \mathscr{K}$ tale che $\omega(b) = \omega(a)$, da cui $a - b \in \mathscr{K}$, e pertanto $a = b + (a - b) \in \mathscr{K}$, cioè $\mathscr{K} = \mathscr{A}$. Dunque $\mathscr{I}$ è massimale.

Viceversa, dato $\mathscr{I} \subset \mathscr{A}$ ideale proprio massimale, mostriamo che $\mathscr{I} = \ker \omega$ per qualche $\omega \in \Omega(\mathscr{A})$. A tale scopo, osserviamo che essendo, per la proposizione 4.2.4(ii), $\bar{\mathscr{I}}$ un ideale proprio contenente $\mathscr{I}$ massimale, si ha che $\mathscr{I} = \bar{\mathscr{I}}$ è chiuso. Mostriamo allora che $\mathscr{A}/\mathscr{I}$ è un'algebra di Banach commutativa priva di ideali propri non banali. Sia infatti $\mathscr{K} \subset \mathscr{A}/\mathscr{I}$ un ideale. Ne segue che, per $\eta : \mathscr{A} \to \mathscr{A}/\mathscr{I}$ la proiezione canonica, $\eta^{-1}(\mathscr{K})$ è un ideale di $\mathscr{A}$: se $a \in \eta^{-1}(\mathscr{K})$ e $b \in \mathscr{A}$, $\eta(ab) = ab + \mathscr{I} = (a + \mathscr{I})(b + \mathscr{I}) \in \mathscr{K}$, cioè $ab \in \eta^{-1}(\mathscr{K})$. Inoltre chiaramente $\mathscr{I} = \eta^{-1}(\{0\}) \subset \eta^{-1}(\mathscr{K})$, e dunque, essendo $\mathscr{I}$ massimale, $\eta^{-1}(\mathscr{K}) = \mathscr{I}$ o $\eta^{-1}(\mathscr{K}) = \mathscr{A}$. Nel primo caso dunque $\mathscr{K} = \eta(\mathscr{I}) = \{0\}$ è banale, e nel secondo $\mathscr{K} = \eta(\mathscr{A}) = \mathscr{A}/\mathscr{I}$ non è proprio. Se allora $x \in \mathscr{A}/\mathscr{I}$, dato $\lambda \in \sigma(x) \neq \emptyset$, essendo $x - \lambda\mathbb{1}_{\mathscr{A}/\mathscr{I}}$ non invertibile si deve avere $x = \lambda\mathbb{1}_{\mathscr{A}/\mathscr{I}}$ (altrimenti, come già osservato, $(x - \lambda\mathbb{1}_{\mathscr{A}/\mathscr{I}})\mathscr{A}/\mathscr{I}$ sarebbe un ideale proprio non banale di $\mathscr{A}/\mathscr{I}$). Pertanto $\mathscr{A}/\mathscr{I} = \mathbb{C}\mathbb{1}_{\mathscr{A}/\mathscr{I}}$, e identificandola quindi con $\mathbb{C}$ si ha che $\eta : \mathscr{A} \to \mathbb{C}$ è un carattere tale che $\ker \eta = \mathscr{I}$.

Infine, se $\omega \in \Omega(\mathscr{A})$ è tale che $\ker \omega = \mathscr{I}$, dato $a \in \mathscr{A}$ si ha, per quanto appena visto, $a + \mathscr{I} = \eta(a)\mathbb{1}_{\mathscr{A}/\mathscr{I}} = \eta(a)\mathbb{1}_{\mathscr{A}} + \mathscr{I}$, cioè $a - \eta(a)\mathbb{1}_{\mathscr{A}} \in \mathscr{I}$, da cui $\omega(a) = \eta(a)$. □

Corollario 4.2.8. *Un carattere di un'algebra di Banach commutativa con unità è limitato.*

Dimostrazione. Come osservato nella dimostrazione del teorema di Gelfand-Mazur, $\ker \omega$ è un ideale chiuso, e dunque ω è limitato per la proposizione 4.2.6. □

È anche facile, usando i risultati precedenti, dare una stima della norma di un $\omega \in \Omega(\mathscr{A})$ (esercizio 4.23).

Un'altra importante conseguenza del teorema di Gelfand-Mazur è il seguente risultato, che permette di descrivere lo spettro di un elemento di $\mathscr{A}$ tramite i caratteri.

Teorema 4.2.9. *Sia $\mathscr{A}$ un'algebra di Banach commutativa con unità. Allora, per ogni $a \in \mathscr{A}$,*

$$\sigma(a) = \{\omega(a) \,:\, \omega \in \Omega(\mathscr{A})\}.$$

Dimostrazione. Sia $\lambda \notin \sigma(a)$. Allora, per ogni $\omega \in \Omega(\mathscr{A})$,

$$\omega(a-\lambda)\omega((a-\lambda)^{-1}) = \omega((a-\lambda)(a-\lambda)^{-1}) = \omega(\mathbb{1}) = 1,$$

e dunque $\omega(a) \neq \lambda$, cioè $\lambda \notin \{\omega(a) \,:\, \omega \in \Omega(\mathscr{A})\}$. Viceversa se $\lambda \in \sigma(a)$, essendo $a-\lambda$ non invertibile, $\mathscr{I}_\lambda = \mathscr{A}(a-\lambda)$ è un ideale proprio. Se allora si considera l'insieme di tutti gli ideali propri contenenti $\mathscr{I}_\lambda$, parzialmente ordinato per inclusione, si vede facilmente che l'unione di tutti gli ideali appartenenti a un qualunque sottoinsieme linearmente ordinato è un ideale proprio (poiché non contiene l'unità), ed è dunque un maggiorante del sottoinsieme linearmente ordinato stesso. Pertanto per il lemma di Zorn esiste un ideale proprio massimale $\mathscr{I}$ contenente $\mathscr{I}_\lambda$, e per il teorema di Gelfand-Mazur $\mathscr{I} = \ker\omega$ per qualche $\omega \in \Omega(\mathscr{A})$. Essendo dunque $a-\lambda \in \mathscr{I}$, si ottiene $\omega(a) = \lambda$, e cioè $\lambda \in \{\omega(a) \,:\, \omega \in \Omega(\mathscr{A})\}$. □

4.3 Il teorema di Stone-Weierstrass

La suriettività della trasformata di Gelfand sarà conseguenza del seguente fondamentale teorema (dovuto a Stone), che generalizza un classico risultato di Weierstrass secondo il quale ogni funzione continua su un intervallo compatto di $\mathbb{R}$ è limite di una successione di polinomi uniformemente convergente.

Teorema 4.3.1 (di Stone-Weierstrass)**.** *Siano X uno spazio compatto di Hausdorff, e $\mathscr{A} \subset C(X)$ una *-sottoalgebra contenente l'unità di $C(X)$ che separa i punti di X. Allora $\mathscr{A}$ è densa in $C(X)$.*

Ricordiamo che $\mathscr{A}$ separa i punti di X se per ogni $x, y \in X$, $x \neq y$, esiste $f \in \mathscr{A}$ tale che $f(x) \neq f(y)$.

Scomporremo la dimostrazione del teorema di Stone-Weierstrass in diversi lemmi, dai quali sarà evidente che in realtà il teorema di Stone-Weierstrass è essenzialmente una proprietà delle sottoalgebre dello spazio $C(X,\mathbb{R})$ delle funzioni continue su X a valori reali. In tutta questa sezione X denoterà sempre uno spazio compatto di Hausdorff.

Lemma 4.3.2. *Sia $\mathscr{A} \subset C(X,\mathbb{R})$ una sottoalgebra contenente l'unità di $C(X,\mathbb{R})$ che separa i punti di X. Allora per ogni $x, y \in X$, $x \neq y$, e per ogni $\lambda, \mu \in \mathbb{R}$, esiste $g \in \mathscr{A}$ tale che $g(x) = \lambda$, $g(y) = \mu$.*

Dimostrazione. Dati $x, y \in X$, $x \neq y$, sia $f \in \mathscr{A}$ tale che $f(x) \neq f(y)$. Essendo allora

$$\det\begin{pmatrix} f(x) & 1 \\ f(y) & 1 \end{pmatrix} = f(x) - f(y) \neq 0,$$

esistono, univocamente determinati, $\alpha, \beta \in \mathbb{R}$ soluzioni del sistema lineare

$$\begin{cases} \alpha f(x) + \beta = \lambda \\ \alpha f(y) + \beta = \mu \end{cases}$$

e basta quindi porre $g := \alpha f + \beta\mathbb{1} \in \mathscr{A}$. □

Il prossimo lemma è un caso particolare del teorema di Weierstrass ricordato sopra.

Lemma 4.3.3. *Esiste una successione di polinomi $(p_n)_{n\in\mathbb{N}}$ che converge uniformemente sull'intervallo $[0,1]$ alla funzione $t \mapsto \sqrt{t}$.*

Dimostrazione. Sia (p_n) la successione di funzioni definita ricorsivamente come

$$\begin{cases} p_0(t) = 0, \\ p_{n+1}(t) = p_n(t) + \frac{1}{2}(t - p_n(t)^2), \end{cases} \qquad t \in [0,1].$$

È chiaro, per induzione, che p_n è un polinomio per ogni $n \in \mathbb{N}$. Mostriamo poi, sempre per induzione, che $0 \leq p_n(t) \leq \sqrt{t}$ per ogni $t \in [0,1]$, $n \in \mathbb{N}$. L'affermazione è ovvia per $n = 0$. Inoltre dall'ipotesi induttiva e dalla definizione ricorsiva si ottiene $p_{n+1}(t) \geq 0$. Infine

$$\begin{aligned} \sqrt{t} - p_{n+1}(t) &= \sqrt{t} - p_n(t) - \frac{1}{2}(t - p_n(t)^2) = (\sqrt{t} - p_n(t))\left(1 - \frac{1}{2}(\sqrt{t} + p_n(t))\right) \\ &\geq (\sqrt{t} - p_n(t))(1 - \sqrt{t}) \geq 0, \end{aligned}$$

dove per le due disuguaglianze si è usata l'ipotesi induttiva $p_n(t) \leq \sqrt{t}$ e il fatto che $\sqrt{t} \leq 1$ per $t \in [0,1]$. Pertanto, dalla definizione ricorsiva si ottiene $p_{n+1}(t) \geq p_n(t)$ per ogni $n \in \mathbb{N}$, $t \in [0,1]$, e quindi, per monotonia, esiste il limite

$$p(t) := \lim_{n \to +\infty} p_n(t), \qquad t \in [0,1].$$

Passando allora al limite per $n \to +\infty$ nella definizione ricorsiva si ottiene che $p(t)$ deve soddisfare

$$p(t) = p(t) + \frac{1}{2}(t - p(t)^2),$$

da cui $p(t) = \sqrt{t}$. Rimane allora di dimostrare che la convergenza della successione (p_n) alla funzione $t \mapsto p(t) = \sqrt{t}$ è uniforme. A tale scopo, dato $\varepsilon > 0$ si definiscano gli insiemi

$$A_n(\varepsilon) := \{t \in [0,1] \, : \, p_n(t) > \sqrt{t} - \varepsilon\}, \qquad n \in \mathbb{N},$$

che sono aperti in $[0,1]$ per la continuità della funzione $t \mapsto p_n(t) - \sqrt{t}$. Dal fatto che $p_n \leq p_{n+1}$ segue che $A_n(\varepsilon) \subset A_{n+1}(\varepsilon)$, e dal fatto che $\lim_{n\to+\infty} p_n(t) = \sqrt{t}$ segue che $[0,1] = \bigcup_{n=0}^{+\infty} A_n(\varepsilon)$. Dalla compattezza di $[0,1]$ si ottiene allora l'esistenza di naturali $n_1 < n_2 < \cdots < n_r$ tali che $[0,1] = \bigcup_{j=1}^{r} A_{n_j}(\varepsilon) = A_{n_r}(\varepsilon)$. Pertanto per $n \geq n_r$ e per ogni $t \in [0,1]$, si ha

$$\sqrt{t} - \varepsilon < p_{n_r}(t) \leq p_n(t) \leq \sqrt{t} < \sqrt{t} + \varepsilon,$$

e cioè $\sup_{t\in[0,1]} |p_n(t) - \sqrt{t}| < \varepsilon$ per ogni $n \geq n_r$. □

L'ultima parte della dimostrazione del lemma precedente non è altro che la dimostrazione, in questo caso particolare, del seguente classico *teorema di Dini*: una successione monotona crescente di funzioni continue su un intervallo compatto (o su uno spazio compatto di Hausdorff) che converge puntualmente a una funzione continua, è uniformemente convergente.

Lemma 4.3.4. *Sia $\mathscr{A} \subset C(X,\mathbb{R})$ una sottoalgebra chiusa contenente l'unità di $C(X,\mathbb{R})$:*

(i) se $f \in \mathscr{A}$, allora $|f| \in \mathscr{A}$;

(ii) se $f, g \in \mathscr{A}$, allora $\max\{f,g\}, \min\{f,g\} \in \mathscr{A}$.

Dimostrazione. (i) Sia (p_n) la successione di polinomi di cui al lemma 4.3.3. Data allora $f \in \mathscr{A}$ si ha $p_n \circ \frac{f^2}{\|f\|_\infty^2} \in \mathscr{A}$ e

$$\Big\|p_n \circ \frac{f^2}{\|f\|_\infty^2} - \frac{|f|}{\|f\|_\infty}\Big\|_\infty = \sup_{x \in X} \Big|p_n\Big(\frac{f(x)^2}{\|f\|_\infty^2}\Big) - \sqrt{\frac{f(x)^2}{\|f\|_\infty^2}}\Big| \leq \sup_{t \in [0,1]} |p_n(t) - \sqrt{t}| \to 0$$

per $n \to +\infty$, e dunque, essendo $\mathscr{A}$ chiusa, $\frac{|f|}{\|f\|_\infty} \in \mathscr{A}$, da cui chiaramente $|f| \in \mathscr{A}$.

(ii) Usando le formule

$$\max\{f,g\} = \frac{1}{2}(f+g+|f-g|), \qquad \min\{f,g\} = \frac{1}{2}(f+g-|f-g|),$$

l'asserto segue subito da (i). □

Il lemma successivo è la versione reale del teorema di Stone-Weierstrass.

Lemma 4.3.5. *Sia $\mathscr{A} \subset C(X,\mathbb{R})$ una sottoalgebra contenente l'unità di $C(X,\mathbb{R})$ che separa i punti di X. Allora $\mathscr{A}$ è densa in $C(X,\mathbb{R})$.*

Dimostrazione. Siano $f \in C(X,\mathbb{R})$ ed $\varepsilon > 0$. Per ogni coppia di punti $y, z \in X$ esiste, per il lemma 4.3.2, una $g_{y,z} \in \mathscr{A}$ tale che $g_{y,z}(y) = f(y)$ e $g_{y,z}(z) \leq f(z) + \frac{\varepsilon}{2}$. Si definiscano allora gli insiemi

$$A_{y,z} := \{x \in X \,:\, g_{y,z}(x) < f(x) + \varepsilon\}, \qquad y, z \in X.$$

Essendo $g_{y,z} - f$ continua, $A_{y,z}$ è aperto, e $z \in A_{y,z}$ per la definizione di $g_{y,z}$. Dunque $(A_{y,z})_{z \in X}$ è un ricoprimento aperto di X compatto, e si possono pertanto trovare $z_1, \dots, z_n \in X$ tale che $X = \bigcup_{j=1}^n A_{y,z_j}$. Posto allora $h_y := \min\{g_{y,z_1}, \dots, g_{y,z_n}\}$ si avrà $h_y \in \bar{\mathscr{A}}$ per il lemma 4.3.4, $h_y(y) = \min_{j=1,\dots,n} g_{y,z_j}(y) = f(y)$, e inoltre per ogni $x \in X$,

$$h_y(x) \leq g_{y,z_j}(x) < f(x) + \varepsilon,$$

dove z_j è tale che $x \in A_{y,z_j}$. Si definiscano poi gli insiemi

$$B_y := \{x \in X \,:\, h_y(x) > f(x) - \varepsilon\}, \qquad y \in X.$$

Di nuovo B_y è aperto e $y \in B_y$ per ogni $y \in X$, e si possono quindi trovare $y_1, \dots, y_m$ tali che $X = \bigcup_{i=1}^m B_{y_i}$. Posto allora $k := \max\{h_{y_1}, \dots, h_{y_m}\}$, si avrà $k \in \bar{\mathscr{A}}$ e, per ogni $x \in X$,

$$f(x) - \varepsilon < h_{y_i}(x) \leq k(x) = \max\{h_{y_1}(x), \dots, h_{y_m}(x)\} < f(x) + \varepsilon,$$

dove y_i è tale che $x \in B_{y_i}$. Dunque $\|k - f\|_\infty < \varepsilon$, il che mostra la densità di $\bar{\mathscr{A}}$ in $C(X,\mathbb{R})$, da cui, essendo $\bar{\mathscr{A}}$ coincidente con la sua chiusura, $\bar{\mathscr{A}} = C(X,\mathbb{R})$, e cioè la tesi. □

Dimostrazione del teorema 4.3.1. Posto $\mathscr{A}_{\text{aa}} := \{f \in \mathscr{A} \,:\, f(x) \in \mathbb{R}\ \forall x \in X\}$ si ha $\mathscr{A} = \mathscr{A}_{\text{aa}} + i\mathscr{A}_{\text{aa}}$. Data infatti $f \in \mathscr{A}$ si ha $\operatorname{Re} f = \frac{1}{2}(f + f^*), \operatorname{Im} f = \frac{1}{2i}(f - f^*) \in \mathscr{A}_{\text{aa}}$. Inoltre chiaramente $\mathscr{A}_{\text{aa}}$ è una sottoalgebra di $C(X,\mathbb{R})$ contenente l'unità, e poi $\mathscr{A}_{\text{aa}}$ separa i punti di X: dati $x, y \in X$ distinti, e scelta $f \in \mathscr{A}$ tale che $f(x) \neq f(y)$, deve essere infatti $\operatorname{Re} f(x) \neq \operatorname{Re} f(y)$ o $\operatorname{Im} f(x) \neq \operatorname{Im} f(y)$. Dunque per il lemma precedente $\overline{\mathscr{A}_{\text{aa}}} = C(X,\mathbb{R})$, e allora usando l'esercizio 4.24 si conclude che $\bar{\mathscr{A}} = \overline{\mathscr{A}_{\text{aa}}} + i\overline{\mathscr{A}_{\text{aa}}} = C(X,\mathbb{R}) + iC(X,\mathbb{R}) = C(X)$. □

Due applicazioni notevoli del teorema di Stone-Weierstrass sono proposte negli esercizi 4.25 e 4.26.

4.4 C*-algebre commutative e calcolo funzionale continuo

Siamo ora nella posizione di dimostrare il teorema dell'isomorfismo di Gelfand.

Teorema 4.4.1 (dell'isomorfismo di Gelfand)**.** *Sia $\mathscr{A}$ una C*-algebra commutativa con unità. Allora la trasformata di Gelfand* $\hat{}: \mathscr{A} \to C(\Omega(\mathscr{A}))$ *è uno *-isomorfismo isometrico.*

Dimostrazione. È chiaro che la trasformata di Gelfand è lineare. Inoltre è moltiplicativa: dati $a, b \in \mathscr{A}$ si ha infatti, per ogni $\omega \in \Omega(\mathscr{A})$,

$$\widehat{ab}(\omega) = \omega(ab) = \omega(a)\omega(b) = \hat{a}(\omega)\hat{b}(\omega),$$

e dunque $\widehat{ab} = \hat{a}\hat{b}$. Per dimostrare che rispetta l'involuzione, osserviamo che dato $a \in \mathscr{A}$ e posto $a_1 := \frac{1}{2}(a + a^*), a_2 := \frac{1}{2i}(a - a^*)$, si ha $a = a_1 + ia_2$ e $a_i = a_i^*$, $i = 1, 2$, da cui segue che, per ogni $\omega \in \Omega(\mathscr{A})$, $\omega(a_i)$ è reale per i teoremi 4.1.13(vi) e 4.2.9, e quindi

$$\hat{a}^*(\omega) = \overline{\hat{a}(\omega)} = \overline{\omega(a)} = \overline{\omega(a_1 + ia_2)} = \omega(a_1 - ia_2) = \omega(a^*) = \widehat{a^*}(\omega),$$

da cui $\hat{a}^* = \widehat{a^*}$. Per quanto riguarda l'isometria, si ha

$$\|\hat{a}\|_\infty = \sup_{\omega \in \Omega(\mathscr{A})} |\hat{a}(\omega)| = \sup_{\omega \in \Omega(\mathscr{A})} |\omega(a)| = r(a) = \|a\|,$$

dove la terza uguaglianza è conseguenza del teorema 4.2.9, e la quarta del teorema 4.1.13(iii) e del fatto che in un'algebra commutativa ovviamente tutti gli elementi sono normali. Infine, per verificare la suriettività della trasformata di Gelfand, osserviamo che $\hat{\mathscr{A}}$ è una *-sottoalgebra di $C(\Omega(\mathscr{A}))$ contenente l'unità (in quanto chiaramente $\hat{\mathbb{1}}_{\mathscr{A}} = \mathbb{1}_{C(\Omega(\mathscr{A}))}$) che separa i punti di $\Omega(\mathscr{A})$, poiché per definizione $\omega_1, \omega_2 \in \Omega(\mathscr{A})$ sono distinti se e solo se esiste $a \in \mathscr{A}$ tale che $\hat{a}(\omega_1) = \omega_1(a) \neq \omega_2(a) = \hat{a}(\omega_2)$. Dunque per il teorema di Stone-Weierstrass $\hat{\mathscr{A}}$ è densa in $C(\Omega(\mathscr{A}))$, ma poiché $\hat{\mathscr{A}}$ è anche chiusa per l'esercizio 4.27 e l'isometria di $a \mapsto \hat{a}$, ne segue $\hat{\mathscr{A}} = C(\Omega(\mathscr{A}))$. □

L'ipotesi che $\mathscr{A}$ sia una C*-algebra è fondamentale per la validità del teorema, come mostrato dall'esercizio 4.28.

Il teorema appena dimostrato si può anche esprimere dicendo che l'applicazione $X \mapsto C(X)$, che ad ogni spazio compatto di Hausdorff associa l'algebra delle funzioni continue su di esso, è suriettiva sull'insieme delle C*-algebre commutative con unità. Per studiarne l'iniettività, conviene tenere presente che due C*-algebre che siano isometricamente *-isomorfe sono da considerarsi identiche, e allo stesso modo si identificheranno spazi topologici X e Y tra loro omeomorfi, cioè tali che esista una funzione biunivoca $f : X \to Y$ che sia continua insieme con la sua inversa. Più in generale, è interessante chiedersi quale sia la controparte, a livello delle algebre di funzioni continue, di un'applicazione continua tra due spazi compatti. La risposta a queste domande è fornita dal risultato seguente.

Teorema 4.4.2 (funtorialità dell'isomorfismo di Gelfand)**.** *Siano $\mathscr{A}$, $\mathscr{B}$, C*-algebre commutative con unità. Esiste una corrispondenza biunivoca $\rho \mapsto \varphi_\rho$ tra gli *-omomorfismi unitali $\rho : \mathscr{A} \to \mathscr{B}$ e le funzioni continue $\varphi : \Omega(\mathscr{B}) \to \Omega(\mathscr{A})$, definita da*

$$\varphi_\rho(\omega) := \omega \circ \rho, \qquad \omega \in \Omega(\mathscr{B}).$$

Inoltre:

(i) se φ_ρ è iniettiva, $\rho(\mathscr{A})$ è denso in $\mathscr{B}$;

(ii) se φ_ρ è suriettiva, ρ è isometrico;

*(iii) se φ_ρ è biunivoca, ρ è uno *-isomorfismo isometrico su $\mathscr{B}$, e φ_ρ è un omeomorfismo.*

Dimostrazione. Osserviamo per iniziare che se $\omega \in \Omega(\mathscr{B})$, $\omega \circ \rho : \mathscr{A} \to \mathbb{C}$ è un funzionale moltlplicativo, e che non è nullo in quanto $\omega \circ \rho(\mathbb{1}) = \omega(\mathbb{1}) = 1$, dunque effettivamente $\omega \circ \rho \in \Omega(\mathscr{A})$. Inoltre se $(\omega_\alpha) \subset \Omega(\mathscr{B})$ è un net convergente a $\omega \in \Omega(\mathscr{B})$, per definizione $\omega_\alpha(\rho(a)) \to \omega(\rho(a))$ per ogni $a \in \mathscr{A}$, e dunque $\varphi_\rho(\omega_\alpha) \to \varphi_\rho(\omega)$, il che mostra che la $\varphi_\rho : \Omega(\mathscr{B}) \to \Omega(\mathscr{A})$ definita nell'enunciato è effettivamente continua. Inoltre $\varphi_{\rho_1} = \varphi_{\rho_2}$ implica, per ogni $a \in \mathscr{A}$, $\omega \in \Omega(\mathscr{B})$,

$$\widehat{\rho_1(a)}(\omega) = \omega(\rho_1(a)) = (\varphi_{\rho_1}(\omega))(a) = (\varphi_{\rho_2}(\omega))(a) = \omega(\rho_2(a)) = \widehat{\rho_2(a)}(\omega),$$

cioè, per l'isomorfismo di Gelfand, $\rho_1(a) = \rho_2(a)$ per ogni $a \in \mathscr{A}$, da cui $\rho_1 = \rho_2$. Dunque $\rho \mapsto \varphi_\rho$ è iniettiva. Data poi $\varphi : \Omega(\mathscr{B}) \to \Omega(\mathscr{A})$ continua, per ogni $a \in \mathscr{A}$ si avrà $\hat{a} \circ \varphi \in C(\Omega(\mathscr{B}))$, e dunque, sempre per l'isomorfismo di Gelfand, esisterà un unico $\rho(a) \in \mathscr{B}$ tale che $\hat{a} \circ \varphi = \widehat{\rho(a)}$, che equivale a

$$(\varphi(\omega))(a) = (\hat{a} \circ \varphi)(\omega) = \widehat{\rho(a)}(\omega) = (\omega \circ \rho)(a), \qquad \omega \in \Omega(\mathscr{B}),$$

cioè a $\varphi(\omega) = \omega \circ \rho$. Pertanto, tenendo conto del fatto, di verifica immediata, che l'unicità implica che l'applicazione $\rho : \mathscr{A} \to \mathscr{B}$ è uno *-omomorfismo unitale, si ottiene $\varphi = \varphi_\rho$, e quindi la suriettività di $\rho \mapsto \varphi_\rho$. Per concludere verifichiamo le (i)-(iii).

(i) Chiaramente $\rho(\mathscr{A})$ è una *-sottoalgebra di $\mathscr{B}$ contenente $\mathbb{1}_\mathscr{B} = \rho(\mathbb{1}_\mathscr{A})$, e dunque $\widehat{\rho(\mathscr{A})}$ è una *-sottoalgebra di $C(\Omega(\mathscr{B}))$ contenente l'unità. Inoltre separa i punti di $\Omega(\mathscr{B})$, poiché dati $\omega_1, \omega_2 \in \Omega(\mathscr{B})$ distinti, si avrà $\omega_1 \circ \rho = \varphi_\rho(\omega_1) \neq \varphi_\rho(\omega_2) = \omega_2 \circ \rho$ per l'iniettività di φ_ρ, ed esisterà dunque $a \in \mathscr{A}$ tale che $\widehat{\rho(a)}(\omega_1) = \omega_1 \circ \rho(a) \neq \omega_2 \circ \rho(a) = \widehat{\rho(a)}(\omega_2)$. Per il teorema di Stone-Weierstrass dunque $\widehat{\rho(\mathscr{A})}$ è denso in $C(\Omega(\mathscr{B}))$, che equivale, per l'isomorfimso di Gelfand, alla densità di $\rho(\mathscr{A})$ in $\mathscr{B}$.

(ii) Si ha, per ogni $a \in \mathscr{A}$,

$$\|\rho(a)\| = r(\rho(a)) = \sup_{\omega \in \Omega(\mathscr{B})} |\omega(\rho(a))| = \sup_{\omega \in \Omega(\mathscr{B})} |\varphi_\rho(\omega)(a)| = \sup_{\omega \in \Omega(\mathscr{A})} |\omega(a)| = r(a) = \|a\|,$$

dove nella prima e nell'ultima uguaglianza si è usato 4.1.13(iii), nella seconda e nella quinta 4.2.9, nella terza la definizione di φ_ρ e nella quarta la sua suriettività.

(iii) Per (i) ρ è isometrico e dunque è uno *-omomorfismo iniettivo e $\rho(\mathscr{A})$ è chiuso in $\mathscr{B}$, ma allora per (ii) $\rho(\mathscr{A}) = \overline{\rho(\mathscr{A})} = \mathscr{B}$ e ρ è uno *-isomorfismo di $\mathscr{A}$ su $\mathscr{B}$. Infine chiaramente $\varphi_\rho^{-1} = \varphi_{\rho^{-1}}$ è continua, e dunque φ_ρ è un omeomorfismo. □

Notiamo esplicitamente che dall'ultimo punto del teorema precedente si ottiene una dimostrazione, basata sulla teoria delle C*-algebre commutative, del risultato classico contenuto nell'esercizio 1.37, nel caso particolare di una funzione biunivoca e continua tra spazi compatti di Hausdorff.

Probabilmente la conseguenza più importante dell'esistenza dell'isomorfismo di Gelfand è la possibilità, a cui abbiamo già accennato, di dare un senso alla valutazione su un elemento autoaggiunto (o normale) di una C*-algebra di una qualunque funzione continua sullo spettro dell'elemento stesso.

Prima di enunciare formalmente tale risultato, introduciamo la nozione seguente: data una C*-algebra $\mathscr{A}$ e un sottoinsieme $\mathscr{S} \subset \mathscr{A}$, definiamo la *C*-sottoalgebra generata da* $\mathscr{S}$ come

$$C^*(\mathscr{S}) := \bigcap \{\mathscr{B} \subset \mathscr{A} \text{ C*-sottoalgebra contenente } \mathscr{S}\}.$$

Si verifica facilmente che questa è effettivamente una C*-sottoalgebra di $\mathscr{A}$, e che la si può anche descrivere come la chiusura in norma dello spazio vettoriale generato dai monomi in $a, a^* \in \mathscr{S}$ (esercizio 4.29).

È anche chiaro che se $a \in \mathscr{A}$ è un elemento normale e $f : \mathbb{C} \to \mathbb{C}$ è un polinomio in λ e $\bar{\lambda}$, della forma

$$f(\lambda) = \sum_{h,k=0}^{n} c_{h,k} \lambda^h \bar{\lambda}^k, \qquad \lambda \in \mathbb{C},$$

per certi $c_{h,k} \in \mathbb{C}$, è perfettamente naturale definire la valutazione di f su a tramite le operazioni algebriche in $\mathscr{A}$:

$$f(a) := \sum_{h,k=0}^{n} c_{h,k} a^h (a^*)^k,$$

e si ottiene in questo modo un elemento $f(a) \in C^*(\{\mathbb{1}, a\})$, e anzi è chiaro che $C^*(\{\mathbb{1}, a\})$ è la chiusura in $\mathscr{A}$ dell'insieme di tali elementi.

Il nostro scopo è allora quello di dare un senso al simbolo $f(a)$ quando f è una qualunque funzione continua sullo spettro di a, e non solo un polinomio.

Teorema 4.4.3 (calcolo funzionale continuo)**.** *Siano $\mathscr{A}$ una C*-algebra con unità e $a \in \mathscr{A}$ normale. Esiste un unico *-omomorfismo isometrico e unitale $\rho : C(\sigma(a)) \to \mathscr{A}$ tale che $\rho(\iota) = a$, dove $\iota \in C(\sigma(a))$ è la funzione definita da $\iota(\lambda) = \lambda$ per ogni $\lambda \in \sigma(a)$. Inoltre $\rho(C(\sigma(a)) = C^*(\{\mathbb{1}, a\})$ e $\sigma(\rho(f)) = f(\sigma(a))$ per ogni $f \in C(\sigma(a))$ (proprietà dello* spectral mapping*).*

L'omomorfismo $\rho : C(\sigma(a)) \to \mathscr{A}$ è detto il *calcolo funzionale continuo* definito dall'elemento normale $a \in \mathscr{A}$, ed è naturale utilizzare la notazione $f(a) := \rho(f)$. Se infatti $f : \mathbb{C} \to \mathbb{C}$ è il polinomio considerato sopra, si potrà scrivere

$$f(\lambda) = \sum_{h,k=0}^{n} c_{h,k} \lambda^h \bar{\lambda}^k = \left[\sum_{h,k=0}^{n} c_{h,k} \iota^h (\iota^*)^k \right] (\lambda), \qquad \lambda \in \mathbb{C},$$

e pertanto, essendo ρ uno *-omomorfismo unitale,

$$\rho(f) = \sum_{h,k=0}^{n} c_{h,k} \rho(\iota)^h (\rho(\iota)^*)^k = \sum_{h,k=0}^{n} c_{h,k} a^h (a^*)^k = f(a).$$

Dunque il calcolo funzionale continuo permette di estendere coerentemente la nozione di "funzione di elemento di una C*-algebra" dai polinomi a tutte le funzioni continue sullo spettro.

Prima di dimostrare il teorema 4.4.3 abbiamo bisogno di un risultato sul comportamento dello spettro di un elemento al variare dell'algebra ambiente. Si noti infatti che se $\mathscr{A}$ è un'algebra di Banach con unità e $\mathscr{B} \subset \mathscr{A}$ è una sottoalgebra di Banach, la nozione di spettro di un $a \in \mathscr{B}$ dipende a priori dal fatto che lo si consideri un elemento di $\mathscr{B}$ o di $\mathscr{A}$, poiché $\lambda - a$ potrebbe non avere un inverso in $\mathscr{B}$, ma averlo in $\mathscr{A}$. Dunque indicando con $\sigma_{\mathscr{A}}(a)$ e $\sigma_{\mathscr{B}}(a)$ gli spettri di a relativamente alle due algebre $\mathscr{A}$, $\mathscr{B}$, si avrà

$$\sigma_{\mathscr{A}}(a) \subset \sigma_{\mathscr{B}}(a),$$

e l'inclusione in generale è propria, come mostrato dall'esercizio 4.30. Il fatto che questo non avvenga nel caso particolare in cui $\mathscr{A}$, $\mathscr{B}$ sono C*-algebre con la stessa identità è il contenuto del lemma seguente.

Lemma 4.4.4 (permanenza spettrale per C*-sottoalgebre)**.** *Siano $\mathscr{A}$ una C*-algebra con unità $\mathbb{1}$, e $\mathscr{B} \subset \mathscr{A}$ una C*-sottoalgebra tale che $\mathbb{1} \in \mathscr{B}$. Allora per ogni $a \in \mathscr{B}$ si ha*

$$\sigma_{\mathscr{A}}(a) = \sigma_{\mathscr{B}}(a).$$

Dimostrazione. Nel corso della dimostrazione, quando intenderemo riferici a proprietà dello spettro che non dipendono dall'algebra ambiente, utilizzeremo semplicemente il simbolo $\sigma(a)$, senza indici.

In base a quanto osservato sopra, basta dimostrare che $\sigma_{\mathscr{B}}(a) \subset \sigma_{\mathscr{A}}(a)$. Inoltre, poiché $\sigma(a - \lambda) = \sigma(a) - \lambda$, $\lambda \in \mathbb{C}$, basterà dimostrare che $0 \in \sigma_{\mathscr{B}}(a)$ implica $0 \in \sigma_{\mathscr{A}}(a)$, il che equivale a $0 \notin \sigma_{\mathscr{A}}(a) \Rightarrow 0 \notin \sigma_{\mathscr{B}}(a)$ o, ancora, al fatto che se esiste $a^{-1} \in \mathscr{A}$, allora $a^{-1} \in \mathscr{B}$.

Osserviamo inoltre che a è invertibile se e solo se aa^* e a^*a sono entrambi invertibili. Se infatti a è invertibile si ha chiaramente $(aa^*)^{-1} = (a^*)^{-1}a^{-1}$ e $(a^*a)^{-1} = a^{-1}(a^*)^{-1}$. Viceversa se aa^* e a^*a sono invertibili, allora $a(a^*(aa^*)^{-1}) = \mathbb{1}$ e $((a^*a)^{-1}a^*)a = \mathbb{1}$, da cui, in base all'esercizio 4.6, l'invertibilità di a.

Sostituendo allora al generico $a \in \mathscr{B}$ gli elementi autoaggiunti aa^* e a^*a, basterà dimostrare che se $a \in \mathscr{B}$ è autoaggiunto e se esiste $a^{-1} \in \mathscr{A}$, allora $a^{-1} \in \mathscr{B}$. Ma dall'autoaggiuntezza di a e dal teorema 4.1.13(vi) segue che $\sigma_{\mathscr{B}}(a) \subset \mathbb{R}$, e dunque per ogni $\varepsilon > 0$ esisterà $(i\varepsilon - a)^{-1} \in \mathscr{B}$. In base all'analiticità della funzione $\lambda \mapsto (\lambda - a)^{-1} \in \mathscr{A}$, lemma 4.1.10(i), si avrà allora $\lim_{\varepsilon \to 0^+}(i\varepsilon - a)^{-1} = a^{-1}$, e quindi, essendo $\mathscr{B}$ chiusa, $a^{-1} \in \mathscr{B}$. □

Dimostrazione del teorema 4.4.3. Unicità. Sia $\tilde{\rho} : C(\sigma(a)) \to \mathscr{A}$ con le stesse proprietà del ρ dell'enunciato. Allora, per quanto osservato sopra, per ogni polinomio $p : \mathbb{C} \to \mathbb{C}$ in λ, $\bar{\lambda}$, si avrà $\tilde{\rho}(p) = p(a) = \rho(p)$. D'altra parte un'immediata applicazione del teorema di Stone-Weierstrass mostra che tali polinomi sono densi in $C(\sigma(a))$ (in particolare il polinomio $\iota(\lambda) := \lambda$ separa i punti di $\sigma(a)$), e allora per la continuità di ρ e $\tilde{\rho}$ e per il teorema 1.4.5, $\tilde{\rho} = \rho$.

Esistenza. Sia $\mathscr{B} := C^*(\{\mathbb{1}, a\})$, e si consideri la funzione $\varphi : \Omega(\mathscr{B}) \to \mathbb{C}$ definita da $\varphi(\omega) := \omega(a)$. In base al teorema 4.2.9 e al lemma 4.4.4 si ha $\varphi(\Omega(\mathscr{B})) = \sigma_{\mathscr{B}}(a) = \sigma(a)$. Inoltre φ è chiaramente continua, ed è iniettiva: se $\omega_1(a) = \omega_2(a)$, allora $\omega_1(p(a)) = \omega_2(p(a))$ per ogni polinomio p, e dunque, per la densità dei polinomi in $\mathscr{B}$ osservata sopra, e per la continuità dei caratteri ω_1 e ω_2, $\omega_1(b) = \omega_2(b)$ per ogni $b \in \mathscr{B}$. Ricordando allora che $\sigma(a)$ e $\Omega(C(\sigma(a)))$ si possono identificare tramite l'omeomorfismo $\lambda \in \sigma(a) \mapsto \delta_\lambda \in \Omega(C(\sigma(a)))$ (esercizio 4.16), dal teorema 4.4.2 segue l'esistenza di uno *-isomorfismo isometrico $\rho : C(\sigma(a)) \to \mathscr{B} \subset \mathscr{A}$ tale che $\varphi = \varphi_\rho$. Si avrà allora

$$\omega(\rho(f)) = \varphi(\omega)(f) = \delta_{\omega(a)}(f) = f(\omega(a)), \qquad \omega \in \Omega(\mathscr{B}),\ f \in C(\sigma(a)),$$

da cui $\sigma(\rho(f)) = f(\sigma(a))$, e $\widehat{\rho(\iota)}(\omega) = \omega(\rho(\iota)) = \omega(a) = \hat{a}(\omega)$ per ogni $\omega \in \Omega(\mathscr{B})$, che implica, grazie all'isomorfismo di Gelfand, $\rho(\iota) = a$. Analogamente si verifica che $\rho(\mathbb{1}_{C(\sigma(a))}) = \mathbb{1}_{\mathscr{A}}$. □

Il calcolo funzionale continuo può essere considerato anche come una generalizzazione del teorema spettrale per operatori lineari su spazi finito-dimensionali: come si vede dall'esercizio 4.31, l'esistenza del calcolo funzionale per una matrice $A \in M_n(\mathbb{C})$ è equivalente alla diagonalizzabilità della matrice stessa. Nel prossimo capitolo vedremo un'altra generalizzazione del teorema spettrale, essenzialmente l'analogo "continuo" dei punti (b) e (c) dell'esercizio 4.31, basata anch'essa sull'esistenza del calcolo funzionale continuo.

È possibile anche generalizzare il calcolo funzionale a una famiglia commutativa di elementi normali di una C*-algebra, esercizio 4.32.

Concludiamo questa sezione con un accenno alla generalizzazione della teoria di Gelfand al caso di C*-algebre commutative senza unità (di cui non avremo bisogno nel seguito). Diremo che uno spazio topologico X è *localmente compatto* se ogni punto di X ha un intorno compatto. Inoltre una funzione $f : X \to \mathbb{C}$ si dirà *nulla all'infinito* se per ogni $\varepsilon > 0$ esiste un compatto $K \subset X$ tale che $|f(x)| < \varepsilon$ per ogni $x \in K^c$. Indichiamo con $C_0(X)$ lo spazio delle funzioni continue e nulle all'infinito su X, che è chiaramente una C*-algebra commutativa (senza unità) con la norma $\|\cdot\|_\infty$. Sia dunque $\mathscr{A}$ una C*-algebra commutativa senza unità, e indichiamo con $\Omega(\mathscr{A})$ l'insieme dei suoi caratteri (la cui definizione ha chiaramente senso indipendentemente dall'esistenza di un'unità). Gli esercizi 4.33-4.38 forniscono allora una traccia di dimostrazione del fatto che $\Omega(\mathscr{A})$ è uno spazio localmente compatto di Hausdorff, e che $\mathscr{A}$ è isometricamente *-isomorfa a $C_0(\Omega(\mathscr{A}))$. Tale risultato è ottenuto riducendosi al caso di C*-algebre con unità tramite l'importante costruzione dell'esercizio 4.35, detta *aggiunzione di un'unità* alla C*-algebra $\mathscr{A}$ senza unità.

4.5 Elementi positivi di una C*-algebra

Un'applicazione importante del calcolo funzionale continuo è la caratterizzazione in termini spettrali degli elementi positivi di una C*-algebra $\mathscr{A}$, che ricordiamo sono gli elementi della forma a^*a, $a \in \mathscr{A}$.

Cominciamo introducendo la nozione, puramente algebrica, di cono in uno spazio vettoriale.

Definizione 4.5.1. Sia X uno spazio vettoriale. Un sottoinsieme $C \subset X$ è un *cono* se per ogni $x, y \in C$ e ogni $t \geq 0$ si ha $x + y \in C$ e $tx \in C$. Un cono C è detto *proprio* se $x, -x \in C$ implica $x = 0$.

Ad esempio, in $\mathbb{R}^2$, il primo quadrante $\{(x, y) \in \mathbb{R}^2 \ : \ x, y \geq 0\}$ è un cono proprio, mentre il semipiano $\{(x, y) \in \mathbb{R}^2 \ : \ y \geq 0\}$ è un cono, ma non è proprio.

Teorema 4.5.2. *Sia $\mathscr{A}$ una C*-algebra con unità. Dato $a \in \mathscr{A}$, le condizioni seguenti sono equivalenti:*

(i) $a = a^*$ *e* $\sigma(a) \subset [0, +\infty)$*;*

(ii) esiste $b = b^* \in \mathscr{A}$ *tale che* $a = b^2$*;*

(iii) esiste $b \in \mathscr{A}$ *tale che* $a = b^*b$.

Inoltre l'insieme

$$\mathscr{A}_+ := \{a \in \mathscr{A} \ : \ a \text{ soddisfa (i), (ii) o (iii)}\}$$

è un cono proprio chiuso in $\mathscr{A}$, detto il cono degli elementi positivi *di $\mathscr{A}$.*

Dati $a, b \in \mathscr{A}$, useremo la notazione $a \geq b$ se $a-b \in \mathscr{A}_+$. Premettiamo alla dimostrazione del teorema un lemma algebrico interessante di per sé.

Lemma 4.5.3. *Siano $\mathscr{A}$ un'algebra con unità, e $a, b \in \mathscr{A}$. Allora $\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$.*

Dimostrazione. Ovviamente la tesi è equivalente a $\rho(ab) \setminus \{0\} = \rho(ba) \setminus \{0\}$. Sia allora $\lambda \in \rho(ab) \setminus \{0\}$, e $c := (\lambda - ab)^{-1}$. Allora

$$(\mathbb{1} + bca)(\lambda - ba) = \lambda - ba + \lambda bca - bcaba = \lambda - ba + b(c(\lambda - ab))a = \lambda$$

e analogamente $(\lambda - ba)(\mathbb{1} + bca) = \lambda$, da cui $\lambda^{-1}(\mathbb{1} + bca) = (\lambda - ba)^{-1}$, e pertanto $\lambda \in \rho(ba) \setminus \{0\}$. Scambiando i ruoli di a e b si ottiene poi $\rho(ba) \setminus \{0\} \subset \rho(ab) \setminus \{0\}$, cioè la tesi. $\square$

Dimostrazione del teorema 4.5.2. Dimostriamo per iniziare che l'insieme

$$\mathscr{A}_+ := \{a \in \mathscr{A} \,:\, a = a^*,\, \sigma(a) \subset [0, +\infty)\}$$

è un cono proprio chiuso. È chiaro che se $a \in \mathscr{A}_+$ e $t \geq 0$, allora $ta \in \mathscr{A}_+$, in quanto $\sigma(ta) = t\sigma(a)$ (poiché $\lambda - ta = t(t^{-1}\lambda - a)$), e che se $a, -a \in \mathscr{A}_+$ allora $\sigma(a) = \{0\}$, e pertanto $\|a\| = r(a) = 0$. Per verificare che la somma di elementi di $\mathscr{A}_+$ appartiene ad $\mathscr{A}_+$, osserviamo che se $a = a^*$, la condizione $\sigma(a) \subset [0, +\infty)$ implica che per ogni $\lambda \geq \|a\|$ si ha $\sigma(a) \subset [0, \lambda]$, cioè $\sigma(\lambda - a) = \lambda - \sigma(a) \subset [0, \lambda]$, da cui $\|\lambda - a\| \leq \lambda$. Viceversa se $\|\lambda - a\| \leq \lambda$ per un qualche $\lambda \geq \|a\|$, si avrà $\sigma(a), \sigma(\lambda - a) = \lambda - \sigma(a) \subset [-\lambda, \lambda]$, da cui $\sigma(a) = \lambda - (\lambda - \sigma(a)) \subset [0, \lambda] \subset [0, +\infty)$. Dati allora $a, b \in \mathscr{A}_+$ e $\lambda \geq \|a\|$, $\mu \geq \|b\|$ si ha $\lambda + \mu \geq \|a + b\|$ e

$$\|(\lambda + \mu) - (a + b)\| \leq \|\lambda - a\| + \|\mu - b\| \leq \lambda + \mu,$$

che implica, per l'osservazione precedente, $a + b \in \mathscr{A}_+$. Infine se $(a_n)_{n\in\mathbb{N}} \subset \mathscr{A}_+$ converge in norma ad $a \in \mathscr{A}$, dato $\lambda > \|a\|$ si avrà definitivamente $\lambda > \|a_n\|$, e dunque $\|\lambda - a_n\| \leq \lambda$, da cui, passando al limite per $n \to +\infty$, $\|\lambda - a\| \leq \lambda$, e pertanto $a \in \mathscr{A}_+$, il che dimostra che $\mathscr{A}_+$ è chiuso.

(i) $\Rightarrow$ (ii). La funzione $f(\lambda) := \sqrt{\lambda}$, $\lambda \geq 0$, appartiene a $C(\sigma(a))$, e dunque si può definire, tramite il calcolo funzionale, $b := f(a) =: \sqrt{a}$, e, essendo il calcolo funzionale un omomorfismo, $b^2 = a$.

(ii) $\Rightarrow$ (iii). Ovvio.

(iii) $\Rightarrow$ (i). L'elemento $a = b^*b$ è autoaggiunto. Definendo allora $|a| \in \mathscr{A}$ tramite il calcolo funzionale, si ponga

$$a_\pm := \pm\frac{1}{2}(a \pm |a|),$$

così che $a = a_+ - a_-$. Inoltre essendo le funzioni $f_\pm(\lambda) = \pm\frac{1}{2}(\lambda \pm |\lambda|)$, $\lambda \in \mathbb{R}$, a valori non negativi, per lo spectral mapping si avrà $\sigma(a_\pm) = f_\pm(\sigma(a)) \subset [0, +\infty)$. Basterà allora dimostrare che $a_- = 0$. Per questo, osserviamo che $\sqrt{a_-}$ è autoaggiunto per il calcolo funzionale, e quindi si ha

$$(b\sqrt{a_-})^*(b\sqrt{a_-}) = \sqrt{a_-}b^*b\sqrt{a_-} = \sqrt{a_-}a\sqrt{a_-} = \sqrt{a_-}(a_+ - a_-)\sqrt{a_-} = -\sqrt{a_-}a_-\sqrt{a_-} = -a_-^2,$$

dove la quarta uguaglianza è ottenuta ancora dal calcolo funzionale, in base al fatto che $\sqrt{f_-(\lambda)}f_+(\lambda) = 0$ per ogni $\lambda \in \mathbb{R}$. Dunque, sempre per la proprietà dello spectral mapping, $\sigma((b\sqrt{a_-})^*(b\sqrt{a_-})) = -\sigma(a_-^2) \subset (-\infty, 0]$. Siano poi $c, d \in \mathscr{A}$ autoaggiunti e tali che $b\sqrt{a_-} =$

$c+id$, come nella dimostrazione del teorema dell'isomorfismo di Gelfand. Allora un semplice calcolo mostra che

$$(b\sqrt{a_-})^*(b\sqrt{a_-}) + (b\sqrt{a_-})(b\sqrt{a_-})^* = 2(c^2+d^2),$$

e pertanto, da quanto visto sopra e dal fatto, visto all'inizio della dimostrazione, che $\mathscr{A}_+ = \{a \in \mathscr{A} : a = a^*,\, \sigma(a) \subset [0,+\infty)\}$ è un cono, segue che

$$\sigma\big((b\sqrt{a_-})(b\sqrt{a_-})^*\big) = \sigma\big(2(c^2+d^2) - (b\sqrt{a_-})^*(b\sqrt{a_-})\big) \subset [0,+\infty).$$

Dal lemma 4.5.3 si ottiene allora $\sigma(a_-^2) = -\sigma\big((b\sqrt{a_-})^*(b\sqrt{a_-})\big) = \{0\}$, e, in conclusione, $\|a_-\|^2 = \|a_-^2\| = r(a_-^2) = 0$. □

Esercizi

4.1 Sia $\mathscr{A}$ una C*-algebra con unità $\mathbb{1}$. Si verifichi che $\mathbb{1}^* = \mathbb{1}$ e $\|\mathbb{1}\| = 1$.

4.2 Si mostri che esiste un'isometria $U \in B(\ell^2(\mathbb{N}))$ tale che $U\delta_n = \delta_{n+1}$ ($(\delta_n)_{n\in\mathbb{N}}$ la base ortonormale canonica), e che U non è unitario. Tale U è detto lo *shift unilatero*.

4.3 Indicata con $\mathbb{B} := \{z \in \mathbb{C} : |z| < 1\}$ la palla unitaria di $\mathbb{C}$, sia $\mathscr{A} := \{f \in C(\bar{\mathbb{B}}) : f \text{ è analitica in } \mathbb{B}\}$. Si mostri che con le operazioni definite puntualmente, l'involuzione definita da $f^*(z) := \overline{f(\bar{z})}$, e la norma $\|\cdot\|_\infty$, $\mathscr{A}$ è una *-algebra di Banach, ma non una C*-algebra.

4.4 Siano $\mathscr{A}$ un'algebra di Banach con identità, $a \in \mathscr{A}$ e $\lambda \in \mathbb{C}$. Si verifichi che $\sigma(a-\lambda\mathbb{1}) = \sigma(a) - \lambda := \{\mu - \lambda : \mu \in \sigma(a)\}$.

*4.5 Siano $(X, \mathfrak{M}, \mu)$ uno spazio di misura σ-finita e $H := L^2(X,\mu)$. Data $f \in L^\infty(X,\mu)$ si definisca *l'operatore di moltiplicazione* per f:

$$(M_f\psi)(x) := f(x)\psi(x), \qquad x \in X,\, \psi \in H.$$

Si dimostri:

(a) l'applicazione $f \in L^\infty \mapsto M_f \in B(H)$ è uno *-omomorfismo isometrico di C*-algebre (sugg.: per l'isometria, si ha $\mu(\{|f| > \|f\|_\infty - 1/n\}) > 0$ per ogni $n \in \mathbb{N}$, e quindi esiste $A_n \subset \{|f| > \|f\|_\infty - 1/n\}$ tale che $0 < \mu(A_n) < +\infty$, e se $\psi_n = \chi_{A_n}$...);

(b) $\lambda \in \sigma(M_f)$ se e solo se $\mu(\{|f-\lambda| < \varepsilon\}) > 0$ per ogni $\varepsilon > 0$ (sugg.: se $\mu(\{|f-\lambda| < 1/n\}) > 0$ per ogni $n \in \mathbb{N}$, se esistesse $(\lambda 1 - M_f)^{-1} \in B(H)$, considerare $\|(\lambda 1 - M_f)^{-1}\chi_{B_n}\|^2$ con $B_n \subset \{|f-\lambda| < 1/n\}$ tale che $0 < \mu(B_n) < +\infty$);

(c) $\lambda \in \sigma_p(M_f)$ se e solo se $\mu(f^{-1}(\{\lambda\})) > 0$.

4.6 Siano $\mathscr{A}$ un'algebra con unità e $a \in \mathscr{A}$ per il quale esistano $\ell, r \in \mathscr{A}$ tali che $\ell a = \mathbb{1} = ar$. Mostrare che allora $\ell = r$ e dunque a è invertibile.

4.7 Sia $(x_n)_{n\in\mathbb{N}}$ di Cauchy in uno spazio metrico X. Si mostri che se esiste una sottosuccessione $(x_{n_k})_{k\in\mathbb{N}}$ convergente a $x \in X$, allora $x_n \to x$.

4.8 Sia X uno spazio normato. Una serie $\sum_{n=0}^\infty x_n$, $x_n \in X$, è detta *assolutamente convergente* se $\sum_{n=0}^{+\infty} \|x_n\| < +\infty$. Mostrare:

(a) se X è di Banach, ogni serie assolutamente convergente è convergente in X;

*(b) viceversa se ogni serie assolutamente convergente è convergente in X, allora X è di Banach (sugg.: data $(x_n) \subset X$ di Cauchy, esiste una sottosuccessione $(x_{n_k})_{k\in\mathbb{N}}$ tale che $\sum_k (x_{n_{k+1}} - x_{n_k})$ sia convergente, e si usa l'esercizio 4.7).

4.9 Sia $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ una successione. Si verifichi che

$$\liminf_{n\to+\infty} x_n \geq \limsup_{n\to+\infty} x_n \;\Rightarrow\; \exists \lim_{n\to+\infty} x_n = \liminf_{n\to+\infty} x_n = \limsup_{n\to+\infty} x_n.$$

4.10 Dato il triangolo $T = \{(x,y) \in \mathbb{R}^2 : 0 \leq y \leq x \leq 1\}$ e $k \in C(T)$, si definisca l'operatore di Volterra $K : C([0,1]) \to C([0,1])$:

$$(Kf)(x) := \int_0^x k(x,y)f(y)\,dy, \qquad f \in C([0,1]),\ x \in [0,1].$$

Mostrare che:

(a) effettivamente $Kf \in C([0,1])$ se $f \in C([0,1])$;

(b) esiste $M > 0$ tale che, per ogni $n \in \mathbb{N}$, $\|K^n\| \leq M^n/n!$;

(c) per ogni $\lambda \in \mathbb{C}\setminus\{0\}$ e ogni $g \in C([0,1])$, l'*equazione di Volterra di seconda specie*

$$\int_0^x k(x,y)f(y)\,dy - \lambda f(x) = g(x), \qquad x \in [0,1],$$

ha un'unica soluzione $f \in C([0,1])$.

4.11 Dimostrare il punto (iv) del teorema 4.1.13.

4.12 Sia $\mathscr{A}$ un'algebra di Banach. Mostrare:

(a) per ogni $a \in \mathscr{A}$, la serie $e^a := \sum_{n=0}^{+\infty} a^n/n!$ è convergente in $\mathscr{A}$;

(b) dati $a, b \in \mathscr{A}$ tali che $ab = ba$, si ha $e^a e^b = e^{a+b}$.

4.13 Sia X uno spazio metrico compatto. Verificare che per ogni $x \in X$ esiste $f \in C(X)$ tale che $f(x) \neq 0$ e che $C(X)$ separa i punti di X. (Sugg.: usare l'esercizio 1.29.)

*4.14 Sia X uno spazio compatto di Hausdorff e si consideri l'applicazione $x \in X \mapsto \delta_x \in C(X)^*$. Si mostri:

(a) $\delta_x \neq 0$ per ogni $x \in X$;

(b) dato per noto che $C(X)$ separa i punti di X, $x \mapsto \delta_x$ è iniettiva;

(c) $x \mapsto \delta_x$ è suriettiva sull'insieme dei funzionali moltiplicativi non nulli. (Sugg.: se per assurdo ω è un funzionale lineare moltiplicativo non nullo tale che per ogni $x \in X$ esiste $f_x \in C(X)$ con $\omega(f_x) \neq f_x(x)$, posto $g_x := f_x - \omega(f_x)$ si può trovare un ricoprimento aperto $(A_{x_j})_{j=1,\dots,n}$ di X tale che $g_{x_j}(y) \neq 0$ per ogni $y \in A_{x_j}$, $j = 1,\dots,n$, e posto allora $g := |g_{x_1}|^2 + \cdots + |g_{x_n}|^2 \in \ker\omega$ si ha $g^{-1} \in C(X)$ e $1 = gg^{-1} \in \ker\omega$.)

4.15 Siano $\mathscr{A}$ un'algebra di Banach commutativa con unità e $\Omega(\mathscr{A}) = \{\omega : \mathscr{A} \to \mathbb{C} \text{ carattere}\}$ lo spettro di Gelfand di $\mathscr{A}$. Si mostri che

(a) la famiglia τ di sottoinsiemi di $\Omega(\mathscr{A})$ che sono unioni arbitrarie di insiemi della forma

$$B(a_1, \ldots, a_n; V_1, \ldots, V_n) := \{\omega \in \Omega(\mathscr{A}) \,:\, \omega(a_1) \in V_1, \ldots, \omega(a_n) \in V_n\},$$

dove $a_i \in \mathscr{A}$, e $V_i \subset \mathbb{C}$ aperto, $i = 1, \ldots, n$, $n \in \mathbb{N}$, è una topologia su $\Omega(\mathscr{A})$;

(b) con questa topologia $\Omega(\mathscr{A})$ è uno spazio di Hausdorff;

(c) un net $(\omega_\alpha)_{\alpha \in I} \subset \Omega(\mathscr{A})$ converge a $\omega \in \Omega(\mathscr{A})$ nella topologia τ se e solo se $\omega_\alpha(a) \to \omega(a)$ per ogni $a \in \mathscr{A}$.

4.16 Sia X uno spazio compatto di Hausdorff. Dimostrare che l'applicazione $x \in X \mapsto \delta_x \in \Omega(C(X))$ è continua con inversa continua. (Sugg.: usare l'esercizio 1.37.)

4.17 Sia $\mathscr{A}$ un'algebra di Banach commutativa con unità. Si verifichi che:

(a) per ogni $a \in \mathscr{A}$ la trasformata di Gelfand $\hat{a} : \Omega(\mathscr{A}) \to \mathbb{C}$, $\omega \mapsto \omega(a)$, è continua;

(b) τ è la più debole topologia su $\Omega(\mathscr{A})$ tale che tutte le funzioni $\hat{a} : \Omega(\mathscr{A}) \to \mathbb{C}$, $a \in \mathscr{A}$, sono continue (cioè se τ' è una topologia su $\Omega(\mathscr{A})$ per la quale tali funzioni sono continue, allora $\tau \subset \tau'$).

*4.18 Si dimostri:

(a) dati $a, b \in \ell^1(\mathbb{Z})$ la serie

$$(a * b)_n := \sum_{k \in \mathbb{Z}} a_{n-k} b_k, \qquad n \in \mathbb{Z},$$

è assolutamente convergente e $\|a * b\|_1 \leq \|a\|_1 \|b\|_1$; $a * b$ è detto la *convoluzione* di $a, b \in \ell^1(\mathbb{Z})$;

(b) con il prodotto definito dalla convoluzione, e con l'involuzione definita da $a^*_n := \overline{a_{-n}}$, $a \in \ell^1(\mathbb{Z})$, si ha che $\ell^1(\mathbb{Z})$ è una *-algebra di Banach con identità;

(c) l'elemento $\zeta := (\delta_{n,1})_{n \in \mathbb{Z}} \in \ell^1(\mathbb{Z})$ è tale che, rispetto alla convoluzione, $\zeta^{-1} = (\delta_{n,-1})_{n \in \mathbb{Z}} \in \ell^1(\mathbb{Z})$, e $\langle \zeta^n \,:\, n \in \mathbb{Z} \rangle$ è denso in $\ell^1(\mathbb{Z})$;

(d) per ogni $\lambda \in \mathbb{T}$, $\omega_\lambda : \ell^1(\mathbb{Z}) \to \mathbb{C}$ definito da

$$\omega_\lambda(a) := \sum_{n \in \mathbb{Z}} a_n \lambda^n, \qquad a \in \ell^1(\mathbb{Z}),$$

è un carattere di $\ell^1(\mathbb{Z})$;

(e) per ogni $\omega \in \Omega(\ell^1(\mathbb{Z}))$ esiste un'unico $\lambda \in \mathbb{T}$ tale che $\omega = \omega_\lambda$; (sugg.: $\lambda := \omega(\zeta)$...)

(f) l'applicazione $\lambda \in \mathbb{T} \mapsto \omega_\lambda \in \Omega(\ell^1(\mathbb{Z}))$ è un omeomorfismo di spazi topologici, cioè è biunivoca, continua e con inversa continua; (sugg.: l'inversa $\omega \mapsto \omega(\zeta)$ va da uno spazio compatto a uno di Hausdorff;)

(g) identificando, come spazi topologici, $\Omega(\ell^1(\mathbb{Z}))$ con il cerchio unitario $\mathbb{T}$ tramite l'applicazione del punto (e), la trasformata di Gelfand di $a \in \ell^1(\mathbb{Z})$ si identifica con la funzione $\hat{a} \in C(\mathbb{T})$ data da

$$\hat{a}(e^{i\theta}) = \sum_{n \in \mathbb{Z}} a_n e^{in\theta}, \qquad e^{i\theta} \in \mathbb{T},$$

cioè con la serie (o trasformata) di Fourier della successione $(a_n)_{n \in \mathbb{Z}}$.

4.19 Siano $\mathscr{A}$ un'algebra normata e $\mathscr{I} \subset \mathscr{A}$ un ideale. Verificare che $\bar{\mathscr{I}}$ è un ideale.

4.20 Siano X uno spazio vettoriale e $M \subset X$ un sottospazio. Si verifichi:

(a) le operazioni

$$\begin{aligned}(x+M)+(y+M) &:= x+y+M,\\ \lambda(x+M) &:= \lambda x + M,\end{aligned} \qquad x, y \in X,\ \lambda \in \mathbb{C},$$

sono ben definite e dotano $X/M = \{x + M \,:\, x \in X\}$ di una struttura di spazio vettoriale;

(b) se $X = \mathscr{A}$ è un'algebra e $M = \mathscr{I}$ un ideale, il prodotto

$$(a+\mathscr{I})(b+\mathscr{I}) := ab + \mathscr{I}, \qquad a, b \in \mathscr{A},$$

è ben definito e rende $\mathscr{A}/\mathscr{I}$ un'algebra.

4.21 Siano H uno spazio di Hilbert e $K \subset H$ un sottospazio chiuso. Mostrare che $H/K \simeq K^{\perp}$ (isometricamente isomorfo), e che quindi H/K è uno spazio di Hilbert.

4.22 Siano $\mathscr{A}$ un'algebra di Banach con identità e $\mathscr{I} \subset \mathscr{A}$ un ideale chiuso proprio. Si verifichi che $\|\mathbb{1} + \mathscr{I}\| \geq 1$ e che $\|\mathbb{1} + \mathscr{I}\| = 1$ se $\|\mathbb{1}\| = 1$.

4.23 Siano $\mathscr{A}$ un'algebra di Banach con identità e $\omega : \mathscr{A} \to \mathbb{C}$ un carattere. Mostrare che $\|\omega\| \leq 1$, e che se $\|\mathbb{1}\| = 1$, $\|\omega\| = 1 = \omega(\mathbb{1})$. (Sugg.: in base all'esercizio 4.22, se $\mathscr{I} = \ker\omega$, $|\omega(a)| \leq \|\omega(a)\mathbb{1} + \mathscr{I}\| = \|a + \mathscr{I}\|$.)

4.24 Siano $\mathscr{A}$ una *-algebra di Banach, e $\mathscr{B} \subset \mathscr{A}$ una *-sottoalgebra. Posto $\mathscr{B}_{\text{aa}} := \{b \in \mathscr{B} \,:\, b = b^*\}$, si verifichi che $\overline{\mathscr{B}} = \overline{\mathscr{B}_{\text{aa}}} + i\overline{\mathscr{B}_{\text{aa}}}$.

4.25 Mostrare che:

(a) i polinomi trigonometrici

$$\Big\{\theta \in [0, 2\pi] \mapsto \sum_{|k| \leq n} c_k e^{ik\theta} \,:\, c_k \in \mathbb{C},\ n \in \mathbb{N}\Big\}$$

sono densi in $C_p([0, 2\pi]) := \{f \in C([0, 2\pi]) \,:\, f(0) = f(2\pi)\}$; (sugg.: $C_p([0, 2\pi]) \simeq C(\mathbb{T})$;)

(b) posto $e_n(\theta) := \frac{1}{\sqrt{2\pi}} e^{in\theta}$, $n \in \mathbb{Z}$, $(e_n)_{n\in\mathbb{Z}}$ è una base ortonormale di $L^2([0, 2\pi])$;

(c) la serie di Fourier di una $f \in L^2([0, 2\pi])$ converge a f in $L^2([0, 2\pi])$.

4.26 Mostrare che:

(a) i polinomi sono densi in $C([a, b])$ (teorema di Weierstrass);

(b) i polinomi di Legendre $(P_\ell)_{\ell\in\mathbb{N}}$, ottenuti tramite il procedimento di Gram-Schmidt in $L^2([-1, 1])$ a partire dalle funzioni $x \mapsto x^\ell$, sono una base ortonormale in $L^2([-1, 1])$;

(c) si ha

$$P_\ell(x) = \sqrt{\frac{2\ell + 1}{2}} \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell}(x^2 - 1)^\ell, \qquad x \in [-1, 1],\ \ell \in \mathbb{N}.$$

4.27 Siano X, Y spazi metrici con X completo, e $f : X \to Y$ isometrica. Mostrare che se $C \subset X$ è chiuso, allora $f(C) \subset Y$ è chiuso.

4.28 Siano $\mathscr{B}$ un'algebra di Banach con unità, e $a \in \mathscr{B}$ tale che $\sigma(a) = \{0\}$ (un tale elemento è detto *topologicamente nilpotente*), e sia $\mathscr{A} := \overline{\langle \mathbb{1}, a, a^2, \dots \rangle} \subset \mathscr{B}$ la sottoalgebra chiusa di $\mathscr{B}$ generata da a. Si mostri:

(a) $x \in \mathscr{A}$ se e solo se $x = \sum_{k=0}^{+\infty} \lambda_k a^k$ (serie convergente in $\mathscr{B}$);

(b) $\Omega(\mathscr{A}) = \{\omega_0\}$, con $\omega_0(x) = \lambda_0$;

(c) la trasformata di Gelfand $\hat{}: \mathscr{A} \to C(\Omega(\mathscr{A}))$ non è iniettiva.

4.29 Siano $\mathscr{A}$ una C*-algebra, $\mathscr{S} \subset \mathscr{A}$ un sottoinsieme, e $C^*(\mathscr{S})$ la C*-sottoalgebra di $\mathscr{A}$ generata da $\mathscr{S}$. Si verifichi che, usando la notazione $a^\sharp = a$ o a^*,

$$C^*(\mathscr{S}) = \overline{\langle a_1^\sharp \dots a_n^\sharp \, : \, a_i \in \mathscr{S}, i = 1, \dots, n, \, n \in \mathbb{N} \rangle}.$$

4.30 Siano $H = \ell^2(\mathbb{Z})$ e $U \in B(H)$ l'operatore definito da $Ue_n = e_{n+1}$, $n \in \mathbb{Z}$, con $(e_n)_{n\in\mathbb{Z}}$ la base ortonormale canonica. Si mostri:

(a) esiste un operatore unitario $V : L^2([0, 2\pi]) \to \ell^2(\mathbb{Z})$ tale che $U = VM_gV^*$, con $M_g \in B(L^2([0, 2\pi]))$ l'operatore di moltiplicazione per la funzione $g(\theta) = e^{i\theta}$, $\theta \in [0, 2\pi]$;

(b) U è unitario e $\sigma(U) = \mathbb{T}$;

(c) se $\mathscr{B} := \overline{\langle \mathbb{1}, U, U^2, \dots \rangle}$ è la sottoalgebra di Banach con unità di $B(H)$ generata da U, si ha $0 \in \sigma_{\mathscr{B}}(U)$. (Sugg: se $p \in \mathscr{B}$ è un polinomio, $\langle U^* e_n, p e_n \rangle = 0$, da cui $U^* \notin \mathscr{B}$.)

4.31 Sia $A \in M_n(\mathbb{C})$ tale che esista $\rho : C(\sigma(A)) \to M_n(\mathbb{C})$ *-omomorfismo iniettivo unitale tale che $\rho(\iota) = A$, $\iota(\lambda) = \lambda$. Mostrare:

(a) $P_\lambda := \rho(\chi_{\{\lambda\}})$, $\lambda \in \sigma(A)$, è un proiettore, e si ha $P_\lambda P_\mu = 0$ se $\lambda \neq \mu$;

(b) vale
$$\rho(f) = \sum_{\lambda \in \sigma(A)} f(\lambda) P_\lambda, \qquad f \in C(\sigma(A)),$$
da cui in particolare $\mathbb{1} = \sum_{\lambda \in \sigma(A)} P_\lambda$, $A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$;

(c) P_λ è il proiettore sull'autospazio di A associato all'autovalore $\lambda \in \sigma(A)$;

(d) A è diagonalizzabile tramite matrici unitarie.

Viceversa, se $A \in M_n(\mathbb{C})$ è diagonalizzabile tramite matrici unitarie, mostrare che esiste $\rho : C(\sigma(A)) \to M_n(\mathbb{C})$ come sopra.

4.32 Siano $\mathscr{A}$ una C*-algebra con unità, e $a_1, \dots, a_n \in \mathscr{A}$ elementi normali tali che $[a_i, a_j] = 0 = [a_i, a_j^*]$, $i, j = 1, \dots, n$. Posto $\mathscr{B} := C^*(\{\mathbb{1}, a_1, \dots, a_n\})$, si mostri:

(a) $\Omega(\mathscr{B})$ è omeomorfo a un sottoinsieme chiuso $X \subset \sigma(a_1) \times \dots \times \sigma(a_n) \subset \mathbb{C}^n$ (lo *spettro congiunto* di $a_1, \dots, a_n$);

(b) esiste un unico *-isomorfismo isometrico $\rho : C(X) \to \mathscr{B}$ tale che $\rho(\iota_k) = a_k$, dove $\iota_k \in C(X)$ è la funzione $\iota_k(\lambda) := \lambda_k$, $k = 1, \ldots, n$ (*calcolo funzionale continuo congiunto* di $a_1, \ldots, a_n$).

4.33 Sia X uno spazio di Hausdorff localmente compatto e ∞ un simbolo non appartenente a X. Mostrare che dichiarando un sottoinsieme A di $\tilde{X} := X \cup \{\infty\}$ aperto se $A \subset X$ è aperto o se $A = \tilde{X} \setminus K$ con $K \subset X$ compatto, si ottiene una topologia di Hausdorff su $\tilde{X}$ tale che $\tilde{X}$ è compatto. Viceversa mostrare che se $\tilde{X}$ è uno spazio compatto di Hausdorff e $\infty \in \tilde{X}$, allora $X := \tilde{X} \setminus \{\infty\}$ è localmente compatto. Lo spazio $\tilde{X}$ è detto la *compattificazione a un punto* di X.

4.34 Sia X uno spazio di Hausdorff localmente compatto, e $\tilde{X}$ la sua compattificazione a un punto, definita nell'esercizio precedente. Mostrare che l'applicazione $\tilde{f} \in C(\tilde{X}) \mapsto \tilde{f}|_X \in C(X)$ induce uno *-isomorfismo isometrico della C*-sottoalgebra $\{f \in C(\tilde{X}) : f(\infty) = 0\}$ su $C_0(X)$.

4.35 Sia $\mathscr{A}$ una C-algebra senza unità, e si consideri $\tilde{\mathscr{A}} := \mathscr{A} \oplus \mathbb{C}$ con la struttura di spazio vettoriale somma diretta. Si mostri:

(a) con le definizioni:

$$(a \oplus \lambda)(b \oplus \mu) := (ab + \lambda b + \mu a) \oplus (\lambda\mu), \qquad (a \oplus \lambda)^* := a^* \oplus \bar{\lambda},$$

$\tilde{\mathscr{A}}$ è una *-algebra con unità;

(b) definito l'operatore $L(a \oplus \lambda) : \mathscr{A} \to \mathscr{A}$ tramite $L(a \oplus \lambda)b := ab + \lambda b$, $b \in \mathscr{A}$, si ha $L(a \oplus \lambda) \in B(\mathscr{A})$ e $x \in \tilde{\mathscr{A}} \mapsto L(x) \in B(\mathscr{A})$ è un omomorfismo di algebre;

(c) $\|L(x)\|^2 \leq \|L(x^*x)\|$ per ogni $x \in \tilde{\mathscr{A}}$; (sugg.: un calcolo mostra $\|L(x)b\|^2 = \|b^*L(x^*x)b\|$, $x \in \tilde{\mathscr{A}}$, $b \in \mathscr{A}$;)

(d) $\|L(a \oplus 0)\| = \|a\|$, $a \in \mathscr{A}$; (sugg.: $\|L(a \oplus 0)a^*\| = \|a\|^2$;)

(e) esiste $\varepsilon > 0$ tale che $\|L(a \oplus 1)\| > \varepsilon$ per ogni $a \in \mathscr{A}$; (sugg.: altrimenti esisterebbe $a_n \in \mathscr{A}$ tale che $\|L(a_n \oplus (-1))\| \leq 1/n$, ed essendo $\|a_n - a_m\| = \|L((a_n - a_m) \oplus 0)\|$ si avrebbe $a_n \to a$ per qualche $a \in \mathscr{A}$, da cui $\|ba_n - b\| = \|L((b \oplus 0)(a_n \oplus (-1)))\| \to 0$ e similmente $\|a_n b - b\| \to 0$;)

(f) posto $\|x\| := \|L(x)\|$, $x \in \tilde{\mathscr{A}}$, si ha che $\tilde{\mathscr{A}}$ è una C*-algebra con unità.

4.36 Sia $\mathscr{A}$ una C*-algebra senza unità. Dato $\omega \in \Omega(\mathscr{A})$ si definisca $\tilde{\omega} : \tilde{\mathscr{A}} \to \mathbb{C}$ tramite $\tilde{\omega}(a \oplus \lambda) := \omega(a) + \lambda$. Si verifichi:

(a) $\tilde{\omega} \in \Omega(\tilde{\mathscr{A}})$ e $\omega \mapsto \tilde{\omega}$ è continua e iniettiva con inversa continua;

(b) $\Omega(\tilde{\mathscr{A}}) = \{\tilde{\omega} : \omega \in \Omega(\mathscr{A})\} \cup \{\omega_\infty\}$ dove $\omega_\infty(a \oplus \lambda) := \lambda$;

(c) $\Omega(\mathscr{A})$ è localmente compatto e $\Omega(\tilde{\mathscr{A}})$ ne è la compattificazione a un punto.

4.37 Siano X uno spazio compatto di Hausdorff e $\mathscr{A} \subset C(X)$ una *-sottoalgebra che separa i punti e tale che esista $x_0 \in X$ per cui $f(x_0) = 0$ per ogni $f \in \mathscr{A}$. Mostrare che $\bar{\mathscr{A}} = \{f \in C(X) : f(x_0) = 0\}$. (Sugg.: $\tilde{\mathscr{A}} := \mathscr{A} + \mathbb{C}\mathbb{1} \subset C(X)$ è densa in $C(X)$, e data $f_n + \lambda_n \mathbb{1} \to f$ con $f(x_0) = 0$, $f_n \in \mathscr{A}$, si ha $f_n \to f$.)

4.38 Sia $\mathscr{A}$ una C*-algebra commutativa senza unità. Si mostri che $\mathscr{A}$ è isometricamente *-isomorfa a $C_0(\Omega(\mathscr{A}))$. (Sugg.: si osservi che, per gli esercizi 4.33-4.36, $C_0(\Omega(\mathscr{A}))$ è isomorfo a $\{f \in C(\Omega(\tilde{\mathscr{A}})) : f(\omega_\infty) = 0\}$, si consideri $a \in \mathscr{A} \mapsto \widehat{a \oplus 0}$, con $x \mapsto \hat{x}$ isomorfismo di Gelfand di $\tilde{\mathscr{A}}$ e si usi l'esercizio 4.37.)

Bibliografia

[Arv] W. Arveson, *A Short Course on Spectral Theory*, Springer, 2002.

[Con] J. Conway, *A Course in Functional Analysis*, Springer, 1985.

[Dop] S. Doplicher, appunti del corso *Meccanica Quantistica*, Università di Roma La Sapienza, A.A. 1995-96.

[DeM] G. De Marco, *Analisi Due/1*, Zanichelli, 1992.

[KF] A. N. Kolmogorov, S. V. Fomin, *Elementi di Teoria delle Funzioni e di Analisi Funzionale*, Mir, 1980.

[Ped] G. K. Pedersen, *Analysis Now*, Springer, 1989.

[RS1] M. Reed, B. Simon, *Methods of Modern Mathematical Physics* vol. I, Springer, 1972.

[RS2] M. Reed, B. Simon, *Methods of Modern Mathematical Physics* vol. II, Springer, 1975.

[Rob] J. E. Roberts, note del corso *Fondamenti di Analisi Matematica*, Università di Roma Tor Vergata, A.A. 2011-12.

[Rud1] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, 1976.

[Rud2] W. Rudin, *Analisi Reale e Complessa*, Boringhieri, 1974.

[Str] F. Strocchi, *An Introduction to the Mathematical Structure of Quantum Mechanics*, World Scientific, 2005.