Spazi ed operatori lineari

Corso di Fisica Matematica 3, a.a. 2018-2019
Dipartimento di Matematica, Università di Milano

16/5/2019

1 Spazi lineari

Nel seguito dovremo considerare degli spazi lineari, che possono essere pensati come una generalizzazione (in particolare, al caso infinito-dimensionale e all'ambito funzionale) degli spazi lineari incontrati nello studio dell'algebra lineare, cioè degli spazi vettoriali di dimensione finita; allo stesso modo, dovremo considerare degli operatori lineari in questi spazi lineari; detti operatori possono essere pensati come una generalizzazione delle matrici.[1]

1.1 Definizione

Uno **spazio lineare** S è un insieme di elementi dotati di una operazione di somma tra di loro ed una di prodotto per un numero (reale o complesso; si parla allora di spazio lineare reale o di spazio lineare complesso, e noi considereremo questo secondo caso) che soddisfano le proprietà seguenti:
(i) Se $f \in S$, $\alpha \in \mathbf{C}$, allora $\alpha f \in S$, cioè S è chiuso rispetto alla moltiplicazione per numeri complessi;
(ii) Se $f, g \in S$, allora $f + g \in S$, cioè S è chiuso rispetto alla somma *finita* di suoi elementi;
(iii) Per ogni $f, g, h \in S$, $f + g = g + f$ e $(f + g) + h = f + (g + h)$, cioè l'operazione di somma in S è commutativa ed associativa;
(iv) Esiste un elemento neutro per la somma, cioè un elemento $0 \in S$ tale che $f + 0 = f \ \forall f \in S$, appartenente a S;
(v) Per ogni $f \in S$, esiste un $\widehat{f} \in S$ tale che $f + \widehat{f} = 0$.

D'ora in poi, capiterà di scrivere per brevità solmente "spazio"; con questo si intenderà uno spazio lineare.

[1] Questa dispensa è un riassunto di due dispense originariamente scritte per il corso di Fisica Matematica 2 svariati anni fa. Al di là di un compattamento, il materiale non è stato modificato; quindi si trova ad avere una impostazione un po' diversa da quella che sarebbe più naturale per il corso di Fisica Matematica 3. Una sorgente, semplice ma completa, per questo materiale è costituita dall'ottimo testo di Cicogna.

Uno spazio lineare è detto *completo* se ogni successione f_k di elementi di S che ammette un limite converge ad un elemento di S; è facile esibire esempi di spazi lineari non completi.

Esempio 1. L'insieme delle funzioni continue sulla retta equipaggiato con le operazioni di somma e di prodotto per numeri complessi definite da

$$(f+g)(x) \ = \ f(x) \, + \, g(x) \ ; \quad (\alpha f)(x) \ = \ \alpha \, f(x) \ ,$$

è uno spazio lineare, ma come ben noto esistono successioni di funzioni continue che convergono ad una funzione discontinua, ad esempio

$$f_k(x) \ = \ \begin{cases} 0 & \text{per } x \le 0 \\ k\,x & \text{per } 0 < x \le 1/k \\ 1 & \text{per } x > 1/k \ . \end{cases}$$

1.2 Prodotto scalare

Un **prodotto scalare** in S è una funzione $(.,.) : S \times S \to \mathbf{C}$ da coppie di elementi di S in $\mathbf{C}$ che soddisfa, per ogni $f, g, h \in S$ ed ogni $\alpha, \beta \in \mathbf{C}$, le seguenti proprietà:

$$\begin{array}{ll} (i) & (g,f) \ = \ (f,g)^* \ , \\ (ii) & (f,f) \ \ge \ 0 \ , \\ (iii) & (f, \alpha g) \ = \ \alpha \, (f,g) \ , \\ (iv) & (f, g+h) \ = \ (f,g) \ + \ (f,h) \ , \\ (v) & (\alpha f, g) \ = \ \alpha^* \, (f,g) \ . \end{array}$$

Notiamo che la (i) implica che $(f,f) \in \mathbf{R}$; ha dunque senso richiedere la (ii); inoltre (i) e (iii) implicano che valga anche la (v).[2]

La (ii) ci permette di definire la norma di f come

$$|f| \ := \ \sqrt{(f,f)} \ . \tag{1}$$

Il prodotto scalare è detto *non degenere* se $(f,f) = 0$ se e solo se $f = 0$, ossia se l'unico elemento di norma zero è lo zero.

Gli elementi f e g sono detti *ortogonali* tra loro se $(f,g) = 0$; per un prodotto scalare non degenere, esiste un unico elemento ortogonale a tutti gli elementi di S cioè $f_0 = 0$.

L'elemento $(g,f)g$ è anche detto la *proiezione* di f su g (esattamente come nel caso finito dimensionale).

Se esiste un insieme di elementi $\Phi = \{\varphi_1, \varphi_2, ...\}$ di S tali che per ogni $f \in S$ si ha $f = \sum_i c_i \varphi_i$ (si noti che in generale la somma è infinita), allora Φ è una *base* di S, e $c_i = (\varphi_i, f)$ sono le componenti di f rispetto a questa base.

Lemma 1. *Il prodotto scalare soddisfa necessariamente la disuguaglianza di Cauchy-Schwarz*

$$|(f,g)| \ \le \ |f| \cdot |g| \ . \tag{2}$$

[2] Una funzione che soddisfi la (iii) e la (v) è detta *sesquilineare*.

Dimostrazione. Infatti, sia $x \in R$ e consideriamo $h = f - x(g,f)g \in S$; allora

$$\begin{aligned} (h,h) &= (f,f) + x^2|(g,f)|^2(g,g) - x(g,f)(f,g) - x(g,f)^*(g,f) \\ &= (f,f) \; - \; 2\,x\;|(f,g)|^2 \; + \; x^2\,|(f,g)|^2\;(g,g) \; := \; P(x)\;. \end{aligned}$$

D'altra parte, sappiamo che deve essere $(h,h) \geq 0$.

Perché $P(x) = ax^2 + 2bx + c$ non abbia nessuna (o abbia una sola) radice reale deve essere $\Delta = b^2 - ac < 0$ (o $\Delta = 0$). Nel nostro caso

$$\Delta \; = \; |(f,g)|^2 \; \left[|(f,g)|^2 \; - \; (g,g)\;(f,f)\right] \; ;$$

e quindi deve, a meno che non sia $|(f,g)| = 0$ (nel qual caso la (2) è banalmente vera), essere

$$|(f,g)|^2 \; - \; (g,g)\;(f,f) \; \leq \; 0 \; ,$$

che è equivalente alla (2). △

Corollario 1. Per f,g qualsiasi in S si ha

$$|f-g| \; \leq |f| + |g| \; . \tag{3}$$

Dimostrazione. Infatti, dal Lemma 1 abbiamo

$$\begin{aligned} |f-g|^2 &= (f-g,f-g) = (f,f) + (g,g) - (f,g) - (g,f) \\ &\leq (f,f) + (g,g) + 2|(f,g)| = |f|^2 + |g|^2 + 2|(f,g)| \; ; \end{aligned}$$

applicando ancora la disuguaglianza di Cauchy-Schwarz all'ultimo termine, otteniamo

$$|f-g|^2 \; \leq \; |f|^2 \; + \; |g|^2 \; + \; 2\,|f| \cdot |g| \; = \; (|f| + |g|)^2 \; ,$$

e prendendo la radice di ambo i membri si ha la (3). △

Il prodotto scalare permette di introdurre lo *spazio duale* S^* dei funzionali lineari L_f su S definiti come il prodotto scalare con un dato elemento f di S; evidentemente $S^* \simeq S$, e possiamo vedere il prodotto scalare come una funzione $(.,.) : S^* \times S \to \mathbf{C}$.

1.3 La notazione di Dirac

Una notazione particolarmente conveniente (in particolare quando si studiano operatori lineari, ancor più se hermitiani, v. nel seguito) è stata introdotta da Dirac.

Come indicato più sopra, il prodotto scalare (f,g) può essere visto come l'applicazione di un elemento – il prodotto scalare con $f \in S$ – dello spazio duale S^* su un elemento $g \in S$.

Indichiamo allora gli elementi di S con un *ket* $|g\rangle$, e quelli di S^* con dei *bra* $\langle f|$; la combinazione dei due produce allora il *bra-ket*, o *bracket*, $\langle f|g\rangle = (f,g)$.

Questa notazione è standard nei testi dedicati alla Meccanica Quantistica. Nel seguito la troveremo anche noi comoda e la useremo: una funzione complicata dipendente dal parametro k verrà indicata semplicemente con $|k\rangle$.

1.4 Metrica

Uno spazio lineare può essere dotato di una **metrica** ρ; questa è una funzione a valori reali, non negativa e simmetrica[3] (distanza), definita su coppie di elementi di S, che soddisfa la *disuguaglianza triangolare*

$$\rho(f,g) \ \leq \ \rho(f,h) \ + \ \rho(h,g) \ . \tag{4}$$

Quando ρ è tale che $\rho(f,g)=0$ se e solo se $f=g$, la metrica è detta essere *non degenere*.

Uno spazio dotato di metrica è anche detto essere uno *spazio metrico*.

Se lo spazio lineare S è dotato di un prodotto scalare, possiamo definire una metrica canonicamente associata a questo[4] come

$$\rho(f,g) \ = \ |f-g| \ = \ \sqrt{(f-g,f-g)} \ . \tag{5}$$

E' possibile avere spazi lineari dotati di metrica ma non di un prodotto scalare.

Lemma 2. *La metrica (5) soddisfa la disuguaglianza triangolare (4).*

Dimostrazione. Si tratta di una immediata applicazione della disuguaglianza di Cauchy-Schwarz[5]. Notiamo innanzi tutto che, con $h \in S$ qualsiasi, $(f-g) = (f-h)+(h-g)$, e dunque $|f-g| = |(f-h)-(g-h)|$. Applicando ora la (3) al membro di destra, otteniamo

$$|f-g| \ \leq \ |f-h| \ + \ |g-h| \ .$$

Ricordando la (5), questa è proprio la (4). $\triangle$

1.5 Spazi di Hilbert e spazi di Banach

Uno spazio lineare provvisto di un prodotto scalare e completo è detto uno **spazio di Hilbert**; il prodotto scalare induce una metrica naturale $\rho(f,g) = |f-g| = \sqrt{(f-g,f-g)}$ e quindi ogni spazio di Hilbert è metrico.

Uno spazio lineare completo provvisto di una norma (ma non necessariamente di un prodotto scalare) è uno **spazio di Banach**; la norma induce una metrica $\rho(f,g) = ||f-g||$ e quindi ogni spazio di Banach è metrico.

Ogni spazio di Hilbert in cui si definisca la norma associata al prodotto scalare, $||f|| = \sqrt{(f,f)}$, diviene uno spazio di Banach. Non tutti gli spazi di Banach sono anche spazi di Hilbert; ad esempio lo spazio L^p delle funzioni per cui l'integrale di $|f(x)|^p$ sulla retta converge è di Banach per ogni p, ma è uno spazio di Hilbert solo per $p=2$.

[3] Cioé tale che $\rho(f,g) = \rho(g,f) \geq 0$, $\forall f,g \in S$.

[4] Naturalmente, potremmo anche definire una diversa metrica in S.

[5] In effetti, a volte la stessa disuguaglianza di Cauchy-Schwarz è detta disuguaglianza triangolare.

1.6 Lo spazio $L^2[a,b]$

Consideriamo l'intervallo reale $[a,b]$, con $L = b - a > 0$. Definiamo lo spazio delle funzioni a quadrato sommabile sull'intervallo considerato,

$$L^2[a,b] \; = \; \left\{ f : [a,b] \to \mathbf{C} \;:\; \int_a^b |f(x)|^2 \, dx \; < \; \infty \right\} . \tag{6}$$

La somma tra funzioni ed il prodotto di una funzione per un numero complesso saranno definiti da

$$(f+g)(x) \; = \; f(x) \; + \; g(x) \; , \quad (\alpha f)(x) \; = \; \alpha \, [f(x)] \; , \tag{7}$$

ed il prodotto scalare sarà definito come

$$(f,g) \; = \; \int_a^b f^*(x) \cdot g(x) \; dx \; . \tag{8}$$

E' immediato controllare che questo soddisfa le condizioni per essere un prodotto scalare.

La norma (indotta dal prodotto scalare) delle funzioni in $L^2[a,b]$ è $||f|| = |f|$, vale a dire

$$||f||^2 \; = \; (f,f) \; = \; \int_a^b |f(x)|^2 \; dx \; . \tag{9}$$

La metrica canonicamente definita dal prodotto scalare in $L^2[a,b]$ è quindi

$$\rho(f,g) \; = \; \sqrt{(f-g,f-g)} \; = \; \sqrt{\int_a^b |f(x) - g(x)|^2 \; dx} \; .$$

Notiamo infine che quando $a = -\infty$, $b = +\infty$ si usa, anziché la notazione $L^2[-\infty,+\infty]$, semplicemente $L^2[\mathbf{R}]$ o anche L^2 *tout court*.

Esercizio 1. Mostrare che $S = L^2[a,b]$ è uno spazio lineare, e che (8) è un prodotto scalare; determinare se quest'ultimo è degenere o meno. Rispondere alle stesse domande nel caso $S = L^2[a,b] \cap \mathcal{C}^0[a,b]$.

Esercizio 2. Mostrare direttamente che il prodotto scalare (8) soddisfa la disuguaglianza (2), che $\rho(f,g) \geq 0$, e che $\rho(f,g) = 0$ se e solo se f e g differiscono solo su un insieme di misura nulla, cioè se $f = g$ quasi ovunque.

Problema 1. Indichiamo con $\mathcal{C}([a,b])$ l'insieme delle funzioni assolutamente continue a valori in $\mathbf{C}$. Discutere se $f \in L^2[R] \cap \mathcal{C}(R)$ implica $\lim_{x\to\pm\infty} f(x) = 0$.

Problema 2. Proporre un esempio di funzione che sia in $L^2[R]$ e che *non* soddisfi $\lim_{x\to\pm\infty} f(x) = 0$.

Problema 3. Si considerino sullo spazio $L^2[a,b]$ la norma $||f||_1 = |f| = \sqrt{(f,f)}$ e la norma $||f||_2 = \sup|f(x)|$, e le metriche associate $\rho_1(f,g)$ e $\rho_2(f,g)$. Si determini se esistono funzioni tali che $\rho_1(f,g) = 0$ e $\rho_2(f,g) \neq 0$.

Problema 4*. Lo spazio delle funzioni che soddisfano $f \in L^2[\mathbf{R}]$ e $f' \in L^2[\mathbf{R}]$ è detto $H^1[\mathbf{R}]$, o semplicemente H^1; discutere se $f \in H^1[\mathbf{R}]$ implica $\lim_{x\to\pm\infty} f(x) = 0$.

1.7 Lo spazio $L^2[a,b;w]$

Le definizioni precedenti si generalizzano introducendo una funzione *peso* $w(x)$.

Consideriamo ancora l'intervallo reale $[a,b]$, con $L = b - a > 0$, ed una funzione reale positiva $w : [a,b] \to R$, $w(x) > 0$ (se $L = \infty$, potrà essere $\lim w(x) = 0$ per $x \to \pm\infty$), il cui integrale sull'intervllo $[a,b]$ sia finito[6].

Definiamo lo spazio delle funzioni a quadrato sommabile rispetto al peso w sull'intervallo considerato,

$$L^2[a,b;w] \;=\; \left\{ f : [a,b] \to \mathbf{C} \;:\; \int_a^b |f(x)|^2\, w(x)\, dx \;<\; \infty \right\} . \tag{10}$$

Il prodotto scalare sarà ora definito come

$$(f,g) \;=\; \int_a^b \left[f^*(x) \cdot g(x) \right] \; w(x) \; dx \; . \tag{11}$$

E' possibile mostrare che lo spazio $L^2[a,b;w]$ è completo (si tratta del teorema di Riesz-Fischer); naturalmente questo si applica anche al caso dello spazio L^2 usuale (cioè con $w(x) \equiv 1$).

Esercizio 3. Mostrare che $S = L^2[a,b;w]$ è uno spazio lineare, e che (11) è un prodotto scalare; determinare se quest'ultimo è degenere o meno.

Esercizio 4. Mostrare direttamente che il prodotto scalare (11) soddisfa la disuguaglianza (2), che $\rho(f,g) \geq 0$, e che $\rho(f,g) = 0$ se e solo se f e g differiscono solo su un insieme di misura nulla, cioè se $f = g$ quasi ovunque.

2 Operatori lineari

Un operatore $A : S \to S$ è *lineare* se, per ogni $\alpha \in \mathbf{C}$, ed ogni $f, g \in S$, si ha

$$A(f+g) \;=\; A(f) \;+\; A(g) \;, \quad A(\alpha f) \;=\; \alpha \; A(f) \; . \tag{12}$$

Esempi di operatori lineari in spazi di funzioni di una variabile (con proprietà opportune, ad esempio di derivabilità) sono operatori differenziali o integrali, quali

$$\begin{aligned} A(f) &= \frac{df}{dx} \;, \\ A(f) &= \sum_{k=0}^{n} \alpha_k \, \frac{d^k f}{dx^k} \;, \\ A(f) &= \int_a^b K(x,y)\, f(y) \; dy \; . \end{aligned}$$

Nel seguito, diremo spesso per brevità "operatori", intendendo in effetti operatori lineari.

[6] Ovviamente possiamo sempre ridurre questo al valore unitario.

2.1 Generalità

Due operatori sono uguali se forniscono lo stesso risultato quando sono applicati su qualsiasi elemento di S, cioè

$$A \; = \; B \quad \Leftrightarrow \; A(f) \; = \; B(f) \;\; \forall f \in S \; .$$

La somma ed il prodotto di operatori lineari sono definiti da

$$(A+B)(f) \; = \; A(f) \; + \; B(f) \; , \quad (A \cdot B)(f) \; = \; A[B(f)] \; . \tag{13}$$

Da queste seguono anche le proprietà associativa e distributiva (ovviamente vere nel caso di operatori differenziali)

$$A \cdot B \cdot C \; = \; (A \cdot B) \cdot C \; = \; A \cdot (B \cdot C) \; , \quad A \cdot (B + C) \; = \; A \cdot B \; + \; A \cdot C \; .$$

Allora lo spazio degli operatori lineari $A : S \to S$ è esso stesso uno spazio lineare, spesso indicato con $\mathcal{L}(S)$.

Il *commutatore* di due operatori è

$$[A, B] \; := A \cdot B \; - \; B \cdot A \; . \tag{14}$$

L'operatore *aggiunto* (o hermitiano coniugato) di A è quell'operatore A^+ per cui, per ogni $f, g \in S$, si ha[7]

$$(A^+ f, g) \; = \; (f, Ag) \; . \tag{15}$$

E' facile vedere che

$$(AB)^+ \; = \; B^+ A^+ \; .$$

Infatti,

$$((AB)^+ f, g) \; := \; (f, ABg) \; = \; (A^+ f, Bg) \; = \; (B^+ A^+ f, g) \; .$$

Esempio 2. Sia S lo spazio delle funzioni $f : \mathbf{R} \to \mathbf{C}$ continue e di quadrato sommabile insieme alle loro derivate (ciò implica che le funzioni vadano a zero a $\pm\infty$), con il prodotto scalare (8) in cui $a \to -\infty$, $b \to +\infty$; consideriamo $A = (d/dx)$. Allora

$$\begin{aligned} (f, Ag) \quad &= \quad \int_{-\infty}^{+\infty} f^*(x)\,(dg/dx)\,dx \; = \; f^*(x) g(x) \Big|_{-\infty}^{+\infty} \; - \; \int_{-\infty}^{+\infty} (df^*/dx)\,g(x)\,dx \\ &= \quad - \int_{-\infty}^{+\infty} (df^*/dx)\,g(x)\,dx \; := \; (A^+ f, g) \; ; \end{aligned}$$

quindi $A^+ = -(d/dx)$.

Esercizio 5. Calcolare, per S lo spazio considerato nell'esempio precedente, l'aggiunto A^+ di $A = i(d/dx)$ e di $A = (d^2/dx^2)$.

Esercizio 6. Dimostrare che, con $\alpha \in \mathbf{C}$, valgono le seguenti proprietà:

$$\begin{array}{ll} (i) & (A^+)^+ \; = \; A \; , \\ (ii) & (\alpha A)^+ \; = \; \alpha^* \, A^+ \; , \\ (iii) & (A \cdot B)^+ \; = \; B^+ \cdot A^+ \; . \end{array}$$

[7] In effetti, questa relazione definisce A^+.

2.2 Operatori hermitiani ed unitari

Se $A = A^+$ diciamo che A è *hermitiano* o *autoaggiunto*; se $A = -A^+$ diciamo che A è *anti-hermitiano* o *anti-autoaggiunto.*

Se A ammette sia un inverso sinistro che un inverso destro, questi coincidono e si parla di operatore inverso A^{-1}. Se $A^+ = A^{-1}$, cioè $A^+A = AA^+ = I$, allora A è detto un operatore *unitario.*

Un operatore unitario conserva il prodotto scalare; infatti, se A è unitario allora si ha

$$(Af, Ag) \;=\; (f, A^+Ag) \;=\; (f, g) \quad \forall f, g \in S. \tag{16}$$

Esercizio 7. Possimo affermare il viceversa, cioè che un operatore che conserva il prodotto scalare – cioè soddisfa la (16) – è unitario (secondo la nostra definizione)?

2.3 Operatori di proiezione

Un operatore P che soddisfi $P^2 = P$ è detto essere un *operatore di proiezione.* Un esempio di operatore di proiezione (ben noto dall'algebra lineare) è l'operatore di proiezione in un sottospazio di $\mathbf{R}^n$.

Più in generale, sia $S_0 \subseteq S$ un sottospazio lineare dello spazio lineare S. Un operatore P_0 che soddisfi $P_0^2 = P_0$ e per cui $P_0 f = f$ su ogni $f \in S_0$ è un operatore di proiezione su S_0.

Se S_0 è generato da $(s_1, ..., s_n)$, allora consideriamo $\Pi_0 = \sum_k (s_k, f) s_k$; si controlla facilmente che questo è un operatore di proiezione su S_0.

Esercizio 8. Verificare che Π_0 definito qui sopra è effettivamente un operatore di proiezione.

Esercizio 9. Determinare se le proprietà $P_0^2 = P_0$ e $P_0 f = f$ per $f \in S_0$ definiscono un unico operatore P_0.

Esercizio 10. Mostrare che se P_1 e P_2 sono operatori di proiezione, la loro somma $P = P_1 + P_2$ è un operatore di proiezione se e solo se $P_1 \cdot P_2 = P_2 \cdot P_1 = 0$.

2.4 Operatori limitati e continui

Un operatore $A : S \to S$ è *continuo* in $f_0 \in S$ se per ogni successione f_n tale che $\lim |f_n - f_0| = 0$ si ha $\lim |Af_n - Af_0| = 0$. Se A è continuo in un punto f_0 di S (ad esempio in $f_0 = 0$), esso è continuo ovunque; questo segue facilmente dalle proprietà degli operatori lineari. Sottolineamo che la norma utilizzata nella definizione è quella indotta dal prodotto scalare.

Un operatore lineare $A : S \to S$ si dice *limitato* se il rapporto $|Af|/|f|$ (con $f \neq 0$) si mantiene limitato al variare di $f \neq 0$ in S, ossia se

$$\sup_{f \neq 0 \in S} \frac{|Af|}{|f|} \;=\; K \;<\; \infty \;.$$

Se A è limitato, il K definito nella formula precedente è detto *norma* di A, ed abbiamo $|A| = K$; per ogni $f \in S$, $|Af| \leq |A| \cdot |f|$.

If effetti, continuità e limitatezza sono (beninteso, per operatori lineari) equivalenti. Vale infatti il seguente risultato.

Lemma 3. *Se A è limitato, esso è anche continuo, e viceversa.*

Dimostrazione. Se A è limitato e f_n converge a f_0, allora

$$|Af_n - Af_0| \ = \ |A(f_n - f_0)| \ \leq \ |A| \cdot |f_n - f_0| \ \to \ 0 \ .$$

Viceversa, sia A continuo e supponiamo che sia non limitato; allora deve esistere (per ogni n) un f_n tale che

$$|Af_n| \ > \ n \ |f_n| \ . \tag{17}$$

Consideriamo $g_n = f_n/(n|f_n|)$, che evidentemente converge a zero. Dunque, essendo A continuo, la successione Ag_n deve anche convergere a zero. Ma la (17) fornisce invece

$$|Ag_n| \ = \ \frac{1}{n\,|f_n|}\,|Af_n| \ > \ 1 \ .$$

Dunque abbiamo una contraddizione, e non possono esistere degli f_n con la proprietà (17) per ogni n. $\triangle$

3 Autovalori ed autofunzioni

Sia A un operatore lineare su S. Se, con $\lambda \in \mathbf{C}$ e $f \neq 0 \in S$, si ha

$$A(f) \ = \ \lambda \ f \ , \tag{18}$$

diciamo che f è una *autofunzione* di A con *autovalore* λ.

L'equazione (18) si scrive anche come

$$(A \ - \ \lambda\, I) \ f \ = \ 0 \ ,$$

dove I è l'operatore identità, $I(f) = f \ \forall f \in S$. Gli autovalori generalizzati (e le autofunzioni generalizzate) di rango k sono definiti dall'equazione

$$(A \ - \ \lambda\, I)^k \ f \ = \ 0 \ . \tag{19}$$

L'insieme degli autovalori di un operatore è detto essere il suo *spettro*. Notiamo che mentre nel caso delle matrici lo spettro è un insieme discreto di punti (di cardinalità non superiore alla dimensione della matrice), nel caso di operatori lineari generali esso può essere un insieme continuo, ovvero essere composto sia da insiemi continui (spettro continuo) che da un insieme discreto di punti (spettro discreto).[8]

Lemma 4. *Per un operatore hermitiano $A = A^+$ gli autovalori sono reali; autofunzioni corrispondenti ad autovalori diversi sono ortogonali tra loro.*

[8] Un operatore può avere uno spettro composto da uno spettro continuo e da uno spettro discreto.

Dimostrazione. Infatti, sia $f \neq 0$ una autofunzione con autovalore λ; allora, usando $A^+ = A$,

$$(f, Af) = (A^+ f, f) = (f, A^+ f)^* = (f, Af)^* = (f, \lambda f)^* = (\lambda f, f) = \lambda^* (f, f) \ ;$$

d'altra parte abbiamo anchd

$$(f, Af) = (f, \lambda f) = \lambda (f, f) \ ,$$

e quindi $\lambda^* = \lambda$.

Siano ora f e g autofunzioni con autovalore rispettivamente λ e μ. Allora

$$(f, Ag) \ = \ (f, \mu g) \ = \ \mu \, (f, g) \ ;$$

d'altra parte, usando il fatto (appena dimostrato) che $\lambda^* = \lambda$, abbiamo anche

$$(f, Ag) \ = \ (A^+ f, g) \ = \ (Af, g) \ = \ \lambda^* \, (f, g) \ = \ \lambda \, (f, g) \ .$$

Dunque, $\lambda(f, g) = \mu(f, g)$; dato che per ipotesi $\lambda \neq \mu$, deve necessariamente essere $(f, g) = 0$. $\triangle$

Osservazione. Se un autovalore λ è multiplo con molteplicità k, allora possiamo scegliere, tramite la procedura di ortogonalizzazione, k autofunzioni f_k corrispondenti all'autovalore λ (cioè tali che $Af_k = \lambda f_k$), indipendenti e mutuamente ortogonali. La procedura per costruirle è quella di Gram-Schmidt, studiata in algebra lineare, e non verrà discussa qui. $\odot$

Esercizio 11. Possiamo fare delle affermazioni simili a quelle del Lemma 4 (e quali) nel caso di operatori anti-hermitiani? Cosa possiamo dire, sulla falsariga del Lemma 4, riguardo agli autovalori di operatori unitari?

Esercizio 12. Fornire esempi di operatori che siano al tempo stesso unitari ed autoaggiunti, ovvero unitari ed anti-autoaggiunti. Cosa possiamo dire sullo spettro di questi operatori?

Lemma 5. *Se $H = A^+ A$ per qualche operatore A, allora gli autovalori di H sono reali e non negativi.*

Dimostrazione. E' evidente che H è hermitiano; quindi i suoi autovalori sono reali. D'altra parte, se $Hf = \lambda f$, abbiamo

$$(f, Hf) \ = \ (f, \lambda f) \ = \ \lambda \, (f, f) \ ,$$

ed anche

$$(f, Hf) \ = \ (f, A^+ A f) \ = \ (Af, Af) \ .$$

Ne segue che possiamo scrivere

$$\lambda \ = \ \frac{(Af, Af)}{(f, f)} \ = \ \frac{|Af|^2}{|f|^2} \ \geq \ 0 \ ,$$

come affermato. $\triangle$

Esempio 3. Il lemma ci permette di affermare che l'equazione differenziale

$$- \psi''(x) \ + \ (1 + x^2) \, \psi(x) \ = \ \lambda \ \psi(x)$$

ammette soluzioni solo per $\lambda \geq 0$. Infatti questa equazione si scrive come $H\psi = \lambda\psi$, con

$$H \ = \ -d^2/dx^2 \ + \ (1 + x^2) \ .$$

A sua volta questo operatore si può scrivere come $H = A^+ A$ con $A = (x - d/dx)$; in effetti, abbiamo $A^+ = (x + d/dx)$, e quindi

$$A^+ A(f) = A^+(xf - f') = x^2 f - xf' + f + xf' - f'' = (x^2 + 1)f - f'' = Hf \ .$$

Il Lemma 5 appena dimostrato ci assicura che tutti gli autovalori di H sono non negativi, quindi che l'equazione proposta ammette soluzione non nulla solo per $\lambda \geq 0$ (naturalmente qui λ non è arbitrario; risulta che esiste un'infinità numerabile di valori di λ per cui si ha soluzione). Questo esempio è rilevante in Meccanica Quantistica.

Osservazione. E' forse opportuno sottolineare come l'insieme degli autovalori di un operatore dipenda dallo spazio su cui è definito l'operatore, e non solo dalla sua espressione algebrica. Del resto, l'operatore stesso dipende dal suo insieme di definizione e non solo dal modo in cui è scritto. Dunque, lo "stesso operatore" che agisce in due spazi diversi (in realtà, operatori diversi scritti nello stesso modo) o in due domini diversi dello stesso spazio, avrà in generale autovalori (ed autofunzioni) diversi. $\odot$

Esempio 4. Consideriamo l'operatore di derivata seconda, $A = d^2/dx^2$; consideriamo ora gli spazi lineari S_1 ed S_2, sottospazi di $L^2[0, 2\pi]$ definiti come le funzioni della forma

$$S_\alpha \ = \ \{ f \in L^2[0, \pi] \ , \ f = \sum_{k=1}^{\infty} \ c_k \ \sin(\alpha k x) \} \quad (\alpha = 1, 2)$$

e quindi periodiche di periodo 2π per $\alpha = 1$ e di periodo π per $\alpha = 2$.

E' immediato verificare che le autofunzioni sono in ogni caso proprio le funzioni "di base" $f_k = \sin(\alpha k x)$, e che queste soddisfano

$$A(f_k) \ = \ -\alpha^2 k^2 \ f_k \ .$$

Dunque per $\alpha = 1$ (cioè per A definito in S_1) gli autovalori sono tutti i quadrati degli interi (con segno negativo), $\lambda_k = -k^2$, mentre per $\alpha = 2$ (cioè per A definito in S_2) gli autovalori sono tutti i quadrati degli interi pari (ancora con segno negativo), $\lambda_k = -(2k)^2$.

Esercizio 13. Verificare che gli insiemi S_α definiti nell'esempio 4 siano in effetti degli spazi lineari, per $\alpha \in \mathbf{N}$ arbitrario. Qual è la condizione di appartenenza ad $L^2[0, 2\pi]$ per funzioni scritte come nella definizione di S_α fornita nell'esempio ?

Esercizio 14. Determinare autovalori ed autofunzioni per l'operatore $A = \partial^2/\partial x^2 + \partial^2/\partial y^2$ definito sulle funzioni di $L^2[Q]$, dove $Q = [0, 2\pi] \times [0, 2\pi]$, della forma

$$f(x, y) \ = \ \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} c_{km} \ \sin(kx) \ \sin(my) \ .$$

Esercizio 15. Detto $S \subset L^2[Q]$ l'insieme di funzioni considerato nell'esercizio precedente, verificare che S sia uno spazio lineare. Qual è la condizione di appartenenza ad $L^2[Q]$ per funzioni scritte come nell'esercizio precedente ?

4 Base completa in spazi di Hilbert

Come nel caso degli spazi vettoriali di dimensione finita, quando si opera in spazi lineari è conveniente introdurre una *base*. In questo caso, bisogna usare qualche cautela nel generalizzare le proprietà familiari dall'algebra lineare, sostanzialmente perché si ha a che fare in generale con spazi di dimensione infinita.

4.1 Definizione e proprietà fondamentali

Scegliamo un insieme numerabile $\Psi = \{\psi_1, \psi_2, ...\}$ di elementi (includendo la possibilità che sia un insieme infinito, purché appunto numerabile) di uno spazio di Hilbert S, che chiameremo anche un *sistema*. Nel caso questi soddisfino condizioni di ortonormalità,

$$(\psi_i, \psi_j) \ = \ \delta_{ij} \ , \tag{20}$$

parleremo di *sistema ortonormale*.[9]

Per ogni $f \in S$, definiamo le componenti di f rispetto al sistema Ψ come

$$c_j \ := \ (\psi_j, f) \ ; \tag{21}$$

si tratta, come già visto, (delle ampiezze) delle proiezioni di f sulle ψ_j.

Lemma 6. *Se Ψ è un sistema ortonormale in S, allora vale per qualsiasi $f \in S$ la* disuguaglianza di Bessel

$$\sum_{j=1}^{n} |c_j|^2 \ \le \ |f|^2 \ . \tag{22}$$

Dimostrazione. Infatti, consideriamo $h = f - \sum c_j \psi_j$. Abbiamo da una parte $|h|^2 \ge 0$; e dall'altra, usando le relazioni (20) e (21),

$$\begin{aligned} |h|^2 &= |f - \sum_j c_j \psi_j|^2 \\ &= |f|^2 - \sum_j [c_j (f, \psi_j) + c_j^* (\psi_j, f)] \ + \ \sum_i \sum_j c_i^* c_j (\psi_i, \psi_j) \\ &= |f|^2 - \sum_j [c_j c_j^* + c_j^* c_j] \ + \ \sum_i \sum_j c_i^* c_j (\psi_i, \psi_j) \\ &= |f|^2 \ + \ \sum_j |c_j|^2 \ - \ 2 \sum_j |c_j|^2 \ = \ |f|^2 \ - \ \sum_j |c_j|^2 \ . \end{aligned}$$

[9] Come ben noto, se Ψ è composto di n vettori indipendenti (cioè è tale che il determinante di Gram, ossia il determinante della matrice G di elementi $G_{ij} = (\psi_i, \psi_j)$, sia non zero), allora è sempre possibile estrarne un sistema ortonormale $\{\varphi_1, ..., \varphi_n\}$ tramite la procedura di ortonormalizzazione di Gram-Schmidt.

La (22) resta perciò dimostrata. $\triangle$

Osservazione. Notiamo che nel caso $n = \infty$ la disuguaglianza di Bessel comporta che la somma $\sum |c_j|^2$ sia convergente, ed anzi che la successione delle somme parziali

$$\chi_n \ := \ \sum_{j=1}^{n} |c_j|^2 \tag{23}$$

sia di Cauchy. [10] $\odot$

Se per ogni $f \in S$ la disuguaglianza di Bessel è saturata, ossia si ha

$$\sum_j |(\psi_j, f)|^2 \ = \ |f|^2 \ , \tag{24}$$

allora si dice che Ψ è un *sistema completo*; in questo caso si dice anche che Ψ constituisce una *base* (ortonormale e completa) per S. La (24) è anche nota come *identità di Parseval.*

Non ogni spazio di Hilbert ammette una base; uno spazio di Hilbert che ammette una base si dice *separabile.* In questa dispensa considereremo solo spazi di Hilbert separabili[11], e nel seguito assumeremo sempre che il sistema Ψ sia infinito (i risultati validi in questo caso lo saranno anche nel caso finito, ma certe difficoltà su cui ci soffermeremo non sarebbero presenti, ed alcuni limiti non avrebbero senso, per sistemi finiti).

Se Ψ è una base, la serie

$$F(x) \ = \ \sum_{k=1}^{\infty} (\psi_k, f) \ \psi_k(x) \tag{25}$$

è detta la *serie di Fourier* (o la *rappresentazione di Fourier*) per f.[12]

Lemma 7. *Se Ψ è un sistema completo in S, allora $\forall f \in S$*

$$\lim_{n\to\infty} \ |f \ - \ \sum_{j=1}^{n} c_j \, \psi_j| \ = \ 0 \ . \tag{26}$$

Dimostrazione. Si veda la dimostrazione del lemma 1. $\triangle$

Quando la (26) è verificata, diciamo che la successione delle somme parziali

$$F_n \ := \ \sum_{j=1}^{n} c_j \, \psi_j \tag{27}$$

[10] Ricordiamo che questo significa che per ogni $\varepsilon > 0$ esiste un numero $N(\varepsilon)$ tale che, ogni qual volta sia n che m sono superiori ad $N(\varepsilon)$, si ha $|\chi_n - \chi_m| < \varepsilon$.

[11] Più precisamente, considereremo solo spazi $L^2[a,b]$ o $L^2[R]$, i quali sono separabili; o loro sottospazi (ottenuti intersecando con altri spazi).

[12] Storicamente, questo nome ha origine per le serie trigonometriche (od esponenziali) che vedremo tra poco, ma è passato ad indicare lo sviluppo su una qualsiasi base ortonormale.

converge in media ad f. Dunque il lemma 2 afferma che se Ψ è un insieme completo, allora F_n converge in media ad f. Va sottolineato che la (26) *non* implica di per sé che la successione F_n converga ad f[13]; si vedano in questo senso il Lemma 11 e la successiva osservazione.

La convergenza in media ad un dato elemento $f \in S$ implica l'identità del prodotto scalare con qualsiasi elemento di S. Questa proprietà, enunciata più precisamente nel successivo lemma 3, è anche nota come *continuità del prodotto scalare*, e costituisce una possibile caratterizzazione alternativa dei sistemi completi; si vedano a questo proposito i Lemmi 9 e 10. (Il Lemma 11 è allora un corollario di questi, anche se abbiamo preferito darne una dimostrazione indipendente.)

Lemma 8. *Se la successione F_n di elementi di S converge in media ad $f \in S$, allora per ogni $g \in S$ si ha*

$$\lim_{n\to\infty} (F_n, g) \ = \ (f, g) \ .$$

Dimostrazione. E' sufficiente usare la disuguaglianza di Cauchy-Schwarz $|(f, g)| \leq |f| \cdot |g|$, e considerare

$$|(F_n, g) - (f, g)| \ = \ |(F_n - f, g)| \ \leq \ |F_n - f| \cdot |g| \ ;$$

per ipotesi $|F_n - f|$ converge a zero, e quindi si ha convergenza a zero anche per $|(F_n, g) - (f, g)|$. △

Lemma 9. *Se Ψ è una base di S, allora l'unico elemento $f_0 \in S$ per cui $(\psi_k, f_0) = 0$ per tutti i k è $f_0 = 0$.*

Dimostrazione. Utilizzando la (24) per $f = f_0$, abbiamo che $(\psi_k, f_0) = 0$ per ogni k implica necessariamente $|f|^2 = 0$. △

Lemma 10. *Se Ψ è un sistema ortonormale in S tale che l'unico elemento $f_0 \in S$ per cui $(\psi_k, f_0) = 0$ per tutti i k è $f_0 = 0$, allora Ψ è un sistema ortonormale completo.*

Dimostrazione. Scegliamo un qualsiasi $f \in S$, e sia $\bar{f} = \sum_k c_k \psi_k$, dove $c_k = (\psi_k, f)$ (questa serie è convergente per quanto visto in precedenza). Per qualsiasi j abbiamo $(\psi_j, f - \bar{f}) = (\psi_j, f) - (\psi_j, \bar{f})$; ricordando che $(\psi_j, \psi_k) = \delta_{jk}$ il secondo termine diviene

$$\left(\psi_j, \sum_k c_k \psi_k\right) \ = \ \sum_k c_k \ (\psi_j, \psi_k) \ = \ c_j \ ,$$

e quindi la relazione precedente fornisce $(\psi_j, f - \bar{f}) = c_j - c_j = 0$. △

[13] Cioè, se abbiamo uno spazio di funzioni su $\mathbf{R}$, non implica che F_n converga ad f nel senso della convergenza puntuale, ovvero che $F_n(x)$ converga ad $f(x)$ per ogni $x \in \mathbf{R}$.

Osservazione. Nella dimostrazione precedente abbiamo scambiato la somma infinita in k con il prodotto scalare. Gli assiomi sul prodotto scalare garantiscono però solo che si possa scambiare una somma finita con il prodotto scalare; è quindi necessario discutere la legittimità di questa operazione. Discutiamo questa questione nel contesto che più ci interessa, ossia nel caso di spazi $L^2[a,b]$ e del prodotto scalare standard in essi,

$$(f,g) \ = \ \int_a^b f^*(x)\, g(x)\ dx \ .$$

Ora abbiamo, riproducendo il calcolo precedente,

$$\begin{aligned} \left(\psi_j, \sum_k c_k \psi_k\right) \ &= \ \int \psi_j^*(\xi) \cdot \left(\sum_k c_k \psi_k(\xi) \right) d\xi \\ &= \ \int \left(\sum_k c_k \psi_j^*(\xi) \psi_k(\xi) \right) d\xi \\ &= \ \sum_k \left(c_k \ \int \psi_j^*(\xi) \psi_k(\xi)\, d\xi \right) \ = \ c_j \ . \end{aligned}$$

Abbiamo scambiato somma infinita ed integrale (che naturalmente rappresenta il prodotto scalare); come noto, lo scambio è giustificato se F_n è uniformemente convergente[14]. D'altra parte usando le relazioni di ortonormalità tra le ψ_j abbiamo (scegliamo $m > n$ per concretezza),

$$|F_m - F_n|^2 \ = \ |\sum_{j=n+1}^m c_j \psi_j|^2 \ = \ \sum_{j=n+1}^m |c_j|^2 \ = \ \chi_m - \chi_n \ ;$$

dal fatto, osservato in precedenza, che χ_n è una successione di Cauchy segue dunque che F_n è uniformemente convergente, e quindi la validità del calcolo precedente. $\odot$

Lemma 11. *Se Ψ è un insieme completo e la successione F_n definita in (27) con $c_j = (\psi_j, f)$ è convergente, allora essa converge ad $\bar{f} = \sum \bar{c}_j \psi_j$ tale che $\bar{c}_j - c_j = 0$ e quindi $\rho(\bar{f}, f) = 0$.*

Dimostrazione. L'eguaglianza tra $\bar{c}_j$ e c_j segue dalla definizione di F_n; infatti, è chiaro che $\bar{c}_i := (\psi_i, \bar{f}) = c_i$. Dunque f ed $\bar{f}$ hanno lo stesso prodotto scalare con qualsiasi elemento di Ψ (e pertanto di S), ovvero $(\bar{f} - f, \psi_i) = 0$; ne segue che $\bar{f} = f$.

Mostriamo anche direttamente che $\rho(\bar{f}, f) = (\bar{f} - f, \bar{f} - f) = 0$. In effetti, in generale abbiamo

$$\begin{aligned} (\bar{f} - f, \bar{f} - f) \ &= \ (\bar{f}, \bar{f}) \ + \ (f,f) \ - \ (\bar{f}, f) \ - \ (f, \bar{f}) \\ &= \ \sum_j |\bar{c}_j|^2 \ + \ \sum_j |c_j|^2 \ - \ \sum_j (\bar{c}_j^* c_j + c_j^* \bar{c}_j) \ ; \end{aligned}$$

[14] La successione F_n S è uniformemente convergente se $|F_n - F_m| < \varepsilon$ ogni qual volta sia n che m sono superiori ad un numero $N(\varepsilon)$; questa è l'equivalente della condizione di Cauchy per successioni in spazi lineari.

per $\bar{c}_j = c_j$ questa mostra immediatamente $(\bar{f} - f, \bar{f} - f) = 0$ e quindi anche $\rho(\bar{f}, f) = 0$. $\triangle$

Osservazione. Lo studente è probabilmente un po' sorpreso dalla presenza di una dimostrazione dettagliata, ed anzi di due dimostrazioni. La ragione di questo è nel desiderio di far comprendere che, nell'ambito degli spazi di Hilbert, la relazione $f = \bar{f}$ va intesa precisamente nel senso di avere lo stesso prodotto scalare con tutti gli elementi di una base (e quindi con tutti gli elementi dello spazio). D'altra parte, se pensiamo ad una realizzazione concreta dello spazio di Hilbert – ad esempio lo spazio di funzioni L^2 – allora avere lo stesso prodotto scalare con ogni altra funzione *non* richiede che sia $f(x) = \bar{f}(x)$ punto per punto: è sufficiente che la relazione sia verificata su un insieme di misura piena, ovvero *a meno di insiemi di misura nulla*. Questa stessa osservazione può esser formulata in modo leggermente diverso: la metrica ρ indotta dal prodotto scalare è ben definita (non degenere) su classi di equivalenza di funzioni in L^2, dove la relazione di equivalenza è precisamente quella di essere uguali quasi ovunque. Si noti anche che già in precedenza si è (implicitamente) fatto uso di questa relazione di equivalenza, nelle dimostrazioni di alcuni Lemmi (lo studente è invitato a controllare quali) in cui abbiamo identificato le funzioni che hanno distanza zero da f_0 con f_0 stessa. $\odot$

Completiamo la nostra discussione con una generalizzazione della (24), nota come identità di Parseval generalizzata o come *relazione di completezza*.

Lemma 12. *Se Ψ è un insieme ortonormale completo e rispetto a questo f e g hanno rispettivamente coefficienti α_i e β_i, allora $(f,g) = \sum_j \alpha_j^* \beta_j$.*

Dimostrazione. Si tratta di un semplice calcolo:

$$(f,g) \ = \ (\alpha_i \psi_i, \beta_j \psi_j) \ = \ \sum_{i,j} \alpha_i^* \beta_j (\psi_i, \psi_j) \ = \ \sum_{i,j} \alpha_i^* \beta_j \, \delta_{ij} \ . \qquad \triangle$$

Osservazione. Ricordando chi sono α_i e β_i, abbiamo anche

$$(f,g) \ = \ \sum_j \alpha_j^* \beta_j \ = \ \sum_{j=1}^{\infty} (f, \psi_j) \, (\psi_j, g) \ .$$

Nella notazione di Dirac la relazione di completezza si scrive

$$\langle f | g \rangle \ = \ \sum_{k=1}^{\infty} \langle f | k \rangle \, \langle k | g \rangle \ , \qquad (28)$$

dove abbiamo indicato con $|k\rangle$ la funzione ψ_k.

Problema 5*. Determinare quali delle proprietà enunciate per una base ortonormale restano vere per una base generale (dunque anche non ortogonale), ossia per un sistema $\Xi = \{\xi_1, \xi_2, ...\}$ di elementi di uno spazio di Hilbert S con la proprietà che ogni $f \in S$ si può scrivere come $f = \sum_i (\xi_i, f) \xi_i$.

Problema 6. Si consideri una base ortonormale Ψ di $L^2[a,b]$ ed un generico elemento $f \in S$. Si chiede di determinare quali sono i numeri c_i che rendono minimo il valore di

$$\int_a^b \left| f(x) - \sum_i c_i \psi_i(x) \right|^2 dx \ .$$

4.2 Operatori densamente definiti

Notiamo anche che un operatore può essere ben definito su una base dello spazio di Hilbert H senza esserlo su tutto lo spazio. Questo è il caso degli operatori di derivazione (ad esempio, $A = d/dx$) sullo spazio L^2, che comprende funzioni non continue e dunque tantomeno derivabili.

In effetti, essendo l'insieme degli elementi che si scrivono come combinazioni lineari degli elementi della base *denso* in H, possiamo pensare A come definito su tutto lo spazio, anche sugli elementi f che non sarebbero nel suo dominio, attraverso la sua azione sulle serie di Fourier che rappresentano f.

Osservazione. E' grazie a questa procedura che ha senso considerare l'equazione delle onde (o altre equazioni differenziali) nello spazio $L^2[a,b]$, e dunque su insiemi di funzioni che comprendono funzioni non continue e tantomeno derivabili due volte in x.[15] $\odot$

4.3 Basi di autofunzioni

Terminiamo questa lunga sezione notando che è interessante considerare il caso[16] in cui si ha un sistema completo costituito da autofunzioni di un operatore lineare A.

Infatti in questo caso si ha

$$A(f) \ = \ A \left(\sum_j c_j \psi_j \right) \ = \ \sum_j c_j \, A(\psi_j) \ = \ \sum_j \lambda_j c_j \psi_j \ ; \tag{29}$$

in altre parole, calcolare l'azione dell'operatore su qualsiasi elemento $f \in S$ risulta assolutamente banale.

Naturalmente, perché questo sia effettivamente utile bisogna che l'operatore A sia proprio quello che si vuole studiare in un determinato contesto (ad esempio, l'operatore differenziale associato ad un'equazione differenziale che vogliamo studiare o risolvere).

Come affermato poc'anzi, il caso di una base completa di autofunzioni non è raro né esotico. Consideriamo l'operatore $A = x(d/dx)$; naturalmente i monomi $\psi_k = x^{k-1}$ $(k = 1, 2, ...)$ sono autofunzioni di A, più precisamente $A\psi_k = (k-1)\psi_k$. Come ben noto, vale il seguente

[15] Ricordiamo anche che il procedimento basato sulle caratteristiche permette anch'esso di trattare situazioni in cui si hanno discontinuità (in numero finito): queste si propagano lungo le caratteristiche.

[16] Questo può sembrare un caso molto raro, ma in realtà si tratta di un caso importante (e comune) nelle applicazioni, come vedremo tra poco.

Teorema di Weierstrass. *I polinomi ψ_k formano un sistema completo nello spazio delle funzioni continue e di quadrato sommabile in $[a,b]$.*

Osservazione. L'insieme delle funzioni continue risulta essere denso nello spazio $L^2[a,b]$; la dimostrazione di questo fatto si basa sul teorema di Fejer (si veda ad esempio il testo di Kolmogorov & Fomin citato in Bibliografia). $\odot$

4.4 Complemento: Polinomi di Hermite

In effetti esistono altri esempi di sistemi completi (ed ortonormali), i cosiddetti *polinomi ortogonali*.

Discutiamo brevemente, tra questi, i polinomi di Hermite; essi rivestono una particolare importanza in Meccanica Quantistica.

Consideriamo $S = L^2[R]$ e gli operatori lineari in S definiti da (lo studente è invitato a verificare che siano effettivamente uno l'aggiunto dell'altro)

$$A \ = \ \frac{1}{\sqrt{2}} \left(x + \frac{d}{dx} \right) \ , \quad A^+ \ = \ \frac{1}{\sqrt{2}} \left(x - \frac{d}{dx} \right) \ .$$

Definiamo inoltre altri due operatori tramite questi[17],

$$N \ = \ A^+ A \ = \ \frac{1}{2} \left(x^2 - \frac{d^2}{dx^2} \right) \ ; \ \ \mathcal{H} \ = \ N \ + \ \frac{1}{2} \ .$$

I polinomi di Hermite sono definiti come

$$H_n(x) \ = \ e^{x^2/2} \left(x - \frac{d}{dx} \right)^n \ e^{-x^2/2} \ ;$$

i primi polinomi di Hermite sono dunque

$$\begin{aligned} H_0 &= 1 \ , \\ H_1 &= 2\,x \ , \\ H_2 &= 4\,x^2 \ - \ 2 \\ H_3 &= 8\,x^3 \ - \ 12\,x \ , \\ H_4 &= 16\,x^4 \ - \ 48\,x^2 \ + \ 12 \ . \end{aligned}$$

Le funzioni

$$\psi_n(x) \ = \ A_n \ e^{-x^2/2} \ H_n(x) \ ,$$

dove A_n è una costante di normalizzazione (fissata dalla condizione $|\psi_n|^2 = 1$) sono autofunzioni di N e dunque di $\mathcal{H}$; si ha

$$N(\psi_k) \ = \ k\,\psi_k \ ; \qquad \mathcal{H}(\psi_k) \ = \ (k+1/2)\,\psi_k \ .$$

[17] Come vedremo nel seguito, l'operatore $\mathcal{H}$ rappresenta l'Hamiltoniana dell'oscillatore armonico in Meccanica Quantistica.

5 Lo spazio ℓ^2

Come abbiamo visto in precedenza, una funzione $f \in S$ (spazio di Hilbert separabile con una base Ψ) è identificata dalla successione dei suoi coefficienti di Fourier $c_i = (\psi_i, f)$; e d'altra parte l'identità di Parseval (26) assicura che la somma dei moduli quadrati di questi sia finita.

Consideriamo lo spazio costituito da successioni $\mathcal{C}$ di numeri complessi aventi norma $|\mathcal{C}|$ finita, definendo la norma di una tale successione come

$$|\mathcal{C}| \ = \ |\{c_i\}| \ = \ \left(\sum_{i=1}^{\infty} |c_i|^2 \right)^{1/2} \ . \tag{30}$$

E' facile controllare che si tratta di uno spazio lineare, in cui possiamo definire il prodotto scalare tra le successioni $\mathcal{A} = \{a_i\}$ e $\mathcal{B} = \{b_i\}$ come

$$(\mathcal{A}, \mathcal{B}) \ = \ \sum_{i=1}^{\infty} a_i^* \, b_i \ . \tag{31}$$

E' chiaro che la norma (30) è proprio la norma indotta dal prodotto scalare (31).

Lo spazio delle successioni di numeri complessi con norma (30) finita si indica con ℓ^2. La corrispondenza tra elementi di uno spazio di Hilbert separabile S (di dimensione infinita[18]) ed i suoi coefficienti di Fourier (una volta scelta una base) rende manifesta la relazione tra S e lo spazio ℓ^2, che sono quindi *isomorfi*.[19]

Naturalmente il fatto che ogni spazio di Hilbert separabile di dimensione infinita sia isomorfo ad ℓ^2 implica anche che:

Proposizione. *tutti gli spazi di Hilbert separabili di dimensione infinita sono isomorfi tra di loro.*

[18] Uno spazio di Hilbert separabile di dimensione finita n è evidentemente isomorfo a $\mathbf{C}^n$ (o ad $\mathbf{R}^n$ se si tratta di uno spazio reale).

[19] Sottolineiamo che questa isomorfia non riguarda solo gli elementi degli spazi e la struttura di spazio lineare, ma anche la struttura del prodotto scalare.

Appendice. Problemi di Sturm-Liouville

I problemi di Sturm-Liouville hanno una particolare rilevanza nelle applicazioni (anche connesse alla Meccanica Quantistica).

Consideriamo equazioni differenziali ordinarie (omogenee, del secondo ordine) della forma

$$\frac{d}{dx}[p(x)\,y'(x)] \ - \ q(x)\,y(x) \ + \ \lambda\,r(x)\,y(x) \ = \ 0 \ , \tag{32}$$

in cui $p(x) > 0$, $q(x) \geq 0$ e $r(x) > 0$ sono funzioni reali note, e $y(x)$ è una funzione incognita, così come il parametro $\lambda \in \mathbf{C}$.[20]

L'equazione (32) deriva dal problema[21]

$$L[y] \ = \ \lambda \ r \ y \ ,$$

dove L è l'operatore lineare

$$L \ = \ - \left[p(x)\,\frac{d^2}{dx^2} \ + \ p'(x)\,\frac{d}{dx} \ - \ q(x) \right] \ .$$

Cerchiamo soluzioni della (32) per $x \in [a,b]$, con le condizioni al bordo[22]

$$y(a) \ = \ 0 \ = \ y(b) \ . \tag{33}$$

La (32) è del secondo ordine, e dunque avrà soluzione generale che dipende da due costanti arbitrarie; una di queste è però moltiplicativa (come segue dal fatto che l'equazione è omogenea) e quindi non può essere usata per soddisfare le (33); può invece essere utilizzata per normalizzare le soluzioni, chiedendo

$$\int_a^b r(x) \ |y(x)|^2 \ dx \ = \ 1 \ .$$

Possiamo usare l'altra costante per soddisfare una delle (33), ad esempio imponendo $y(a) = 0$. La seconda condizione potrà essere soddisfatta solo per determinati (e speciali) valori λ_n del parametro λ, detti *autovalori* (se $r(x) = 1$, si tratta degli autovalori ordinari dell'operatore L, altrimenti degli autovalori di $\widehat{L} = [1/r(x)]L$); la soluzione corrispondente è una *autofunzione* $y_n(x)$.

Lemma A1. *Gli autovalori del problema di Sturm-Liouville (32), (33) sono sempre reali.*

[20] Si tratta dunque più precisamente, per p, q ed r fissati, di una famiglia ad un parametro (complesso) di equazioni differenziali.

[21] Questo può anche essere riscritto come $\widehat{L}[y] = \lambda y$, con $\widehat{L} = [1/r(x)]L$; si ricordi che abbiamo supposto $r(x) > 0$ per ogni x.

[22] Si potrebbero considerare condizioni al bordo – e problemi di Sturm-Liouville – più generali, ma non lo faremo.

Dimostrazione. Per mostrare questo, consideriamo il membro di sinistra della (32), chiamiamolo η, ed il suo complesso coniugato η^*:

$$\begin{aligned} \eta(x) &= [p(x)\, y'(x)]' \; - \; q(x)\, y(x) \; + \; \lambda\, r(x)\, y(x) \; , \\ \eta^*(x) &= [p(x)\, (y^*)'(x)]' \; - \; q(x)\, y^*(x) \; + \; \lambda^*\, r(x)\, y^*(x) \; ; \end{aligned}$$

consideriamo ora $h(x) := [y^*(x)\eta(x) - y(x)\eta^*(x)]$, e l'integrale di h sull'intervallo $[a, b]$, che – usando la (33) per annullare i termini di bordo – risulta essere

$$\begin{aligned} \int_a^b h(x)\, dx &= \int_a^b [y^*(py')'\, dx \; + \; (\lambda - \lambda^*) \int_a^b r\, |y|^2\, dx \\ &= \left[(y^*\, p\, y')\big|_a^b \; - \; \int (y^*)'\, p\, y'\, dx \right] \\ & \quad - \left[(y\, p\, (y^*)')\big|_a^b \; - \; \int y'\, p\, (y^*)'\, dx \right] \; + \; (\lambda - \lambda^*) \int_a^b r\, |y|^2\, dx \\ &= \; + \; (\lambda - \lambda^*) \int_a^b r\, |y|^2\, dx \; . \end{aligned}$$

D'altra parte, $h(x)$ è reale (ed anzi nulla se y è soluzione della (32), nel qual caso per definizione $\eta = 0$), e quindi anche il suo integrale deve essere reale (ed anzi nullo). Anche $r(x)$ è reale, e così naturalmente $|y(x)|^2$. Ne segue che $(\lambda - \lambda^*)$ è anche reale; ma questa è (il doppio del)la parte immaginaria di λ. Quindi, deve essere $(\lambda - \lambda^*) = 0$, ossia λ reale.

Usando il fatto che l'integrale deve essere nullo si giunge naturalmente (e più direttamente) alla stessa conclusione. $\triangle$

Lemma A2. *Per ogni autovalore λ_n del problema di Sturm-Liouville (32), (33), esiste una sola autofunzione $y_n(x)$.*

Dimostrazione. Se così non fosse, esisterebbero due soluzioni $y_n^{(1)}$ ed $y_n^{(2)}$ corrispondenti alla stessa λ_n e quindi $z(x) = c_1 y_n^{(1)}(x) + c_2 y_n^{(1)}(x)$ sarebbero tutte soluzioni della (32) con $\lambda = \lambda_n$ e nulle agli estremi. Ma la soluzione generale della (32) per λ fissato dipende da due costanti arbitrarie, ed è quindi la $z(x)$, ed ha valore arbitrario in $x = a$ ed in $x = b$. Quindi abbiamo una contraddizione, ed esiste una sola $y_n(x)$. $\triangle$

Lemma A3. *Le y_n determinate dalle soluzioni (normalizzate, vedi sopra) del problema di Sturm-Liouville (32), (33) sono ortonormali rispetto al prodotto scalare in $L^2[a, b; r]$, ossia*

$$\int_a^b y_m^*(x)\, y_n(x)\; r(x)\; dx \; = \; \delta_{mn} \; .$$

Dimostrazione. Definiamo η_n ed η_n^* come sopra, considerando però $\lambda = \lambda_n$ e $y_n(x)$ anziché una generica soluzione $y(x)$, cioè

$$\eta_n(x) \; = \; [p(x)\, y_n'(x)]' \; - \; q(x)\, y_n(x) \; + \; \lambda_n\, r(x)\, y_n(x) \; .$$

Consideriamo ora la funzione $h_{mn}(x) := [y_m^*(x)\eta_n(x) - y_n(x)\eta_m^*(x)]$. Procedendo come nella dimostrazione del Lemma A1, abbiamo

$$\begin{aligned}
\int_a^b h_{mn}(x)\,dx &= \int_a^b (y_m^*\eta_n - y_n\eta_m^*)\,dx \\
&= p(x)\,[y_m^*(x)y_n'(x) - y_n(y_m^*)']_a^b - \int_a^b p[(y_m^*)'y_n' - y_n'(y_m^*)']dx \\
&\quad + \int r\,(\lambda_n y_m^* y_n \ - \ \lambda_m y_n y_m^*)\,dx \\
&= (\lambda_n - \lambda_m) \int_a^b r\,y_m^* y_n\,dx \ = \ (\lambda_n - \lambda_m)\ (y_m, y_n)\ .
\end{aligned}$$

Questo integrale deve però annullarsi se y_n, y_m solo soluzioni della (32); dunque se $\lambda_n \neq \lambda_m$ deve essere $(y_m, y_n) = 0$. △

Vale inoltre il seguente teorema (di Sturm-Liouville), che non dimostreremo; ne considereremo però un esempio, cioè una applicazione ad una concreta scelta (la più semplice possibile) di p, q, r e di a, b.

Teorema. *Sia $Y = \{y_n\}$ il sistema di autofunzioni di un problema di Sturm-Liouville (32), (33). Allora Y costituisce una base completa nello spazio $S[a,b]$ delle funzioni continue a tratti con le loro derivate di ordine uno e due che soddisfino la (33). Ogni funzione $f \in S[a,b]$ si sviluppa nella serie assolutamente convergente*

$$f(x) \ = \ \sum_k c_k\,y_k(x)\ ,$$

dove $c_k = (y_k, f)$ con $(.,.)$ il prodotto scalare di $L^2[a,b;r]$.

Esempio. Consideriamo $a = -\pi$, $b = \pi$; sia inoltre $p(x) = r(x) = 1$, $q(x) = 0$. Allora la (32) e la (33) diventano semplicemente

$$y''(x) \ + \ \lambda\,y(x) \ = \ 0\ , \quad y(-\pi) = y(\pi) = 0\ .$$

La soluzione generale dell'equazione è quindi

$$y(x) \ = \ \alpha\ \cos(\sqrt{\lambda}x) \ + \ \beta\ \sin(\sqrt{\lambda}x)\ ,$$

con α e β costanti arbitrarie. In $x = \pm\pi$ abbiamo

$$y(\pm\,\pi) \ = \ \alpha\ \cos(\sqrt{\lambda}\,\pi) \ \pm \ \beta\ \sin(\sqrt{\lambda}\,\pi)\ ,$$

e dobbiamo quindi chiedere che sia

$$\alpha\ \cos(\sqrt{\lambda}\,\pi) \ \pm \ \beta\ \sin(\sqrt{\lambda}\,\pi) \ = \ 0\ .$$

Abbiamo due famiglie di soluzioni:

$$\begin{array}{lll}
\alpha = 0\ ,\ \beta \text{ qualsiasi}\ , & \sqrt{\lambda} = n & (\lambda = n^2)\ ; \\
\alpha \text{ qualsiasi}\ ,\ \beta = 0\ , & \sqrt{\lambda} = n + \frac{1}{2} & (\lambda = (n+1/2)^2)\ .
\end{array}$$

In altre parole, normalizzando le soluzioni, abbiamo

$$\begin{array}{ll} \lambda = n^2 & y_n(x) = \sin(nx) \ , \\ \lambda = (n+1/2)^2 & \widehat{y_n}(x) = \cos[(n+1/2)x] \ . \end{array}$$

Questo problema è collegato alla cosiddetta "buca di potenziale" studiata in Meccanica Quantistica.

Bibliografia

Qualsiasi testo di analisi funzionale (inclusi molti che hanno per titolo "Analisi Reale") discute il materiale di questa dispensa, ed ovviamente molto di più. Quella che segue è una breve lista di testi a cui lo studente può rivolgersi per approfondimenti secondo un approccio non troppo diverso da quello seguito qui; in particolare i testi di Cicogna e di Kolmogorov & Fomin sono stati usati anche nella preparazione di questa dispensa.

- G. Cicogna, *Metodi Matematici della Fisica*, Springer Italia 2008
- F.W. Byron & R.W. Fuller, *Mathematics of Classical and Quantum Physics*, Dover 1992
- Ph. Dennery & A. Krzywicki, *Mathematics for Physicists*, Dover 1996
- W. Rudin, *Real and Complex Analysis*, McGraw-Hill 1970
- L. Schwartz, *Mathematics for the Physical Sciences*, Dover 2008
- M. Reed & B. Simon, *Methods of Modern Mathematical Physics – vol.I*, Academic Press 1980
- A.N. Kolmogorov & S.V. Fomin, *Elementi di Teoria delle Funzioni e di Analisi Funzionale*, MIR 1980

`G. Gaeta, 16/5/2019`