

17

Spazi di Hilbert

17.1 Spazi normati

Consideriamo le funzioni continue (a valori reali o complessi) su un intervallo chiuso $[a, b]$. Se f e g sono funzioni continue, chiaramente anche $\alpha f + \beta g$, dove α e β sono scalari, lo è: le funzioni continue formano uno spazio vettoriale. Si può parlare di lunghezza, o norma, di una funzione continua, e ci sono molti modi di farlo. La “norma del sup”, o norma uniforme, è così definita

$$\|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)|. \quad (17.1)$$

È facile convincersi che questa nozione di lunghezza soddisfa le proprietà (i), (ii) e (iii) della norma presentate nella sezione 2.4. Le proprietà (i) e (ii) sono facilmente verificate e lo stesso può dirsi per la disuguaglianza triangolare. Questa disuguaglianza per numeri reali è ovvia, $|f(x) + g(x)| \leq |f(x)| + |g(x)|$, e da questa discende

$$\sup_{x \in [a, b]} |f(x) + g(x)| \leq \sup_{x \in [a, b]} |f(x)| + \sup_{x \in [a, b]} |g(x)|$$

Dunque, $C[a, b]$, lo spazio delle funzioni continue su $[a, b]$ con norma (17.1), è uno spazio normato.

Come abbiamo visto nella sezione 2.4, dalla nozione di lunghezza si può facilmente passare alla nozione di distanza. Se definiamo

$$d(f, g) = \|f - g\|_{\infty},$$

si vede che sono soddisfatte le proprietà (i), (ii), (iii) e (iv) della distanza. Se si ha una nozione di distanza, si ha una nozione metrica di vicinanza tra punti, e se si ha una nozione di vicinanza, si può parlare di convergenza di una successione (f_n) di punti nel modo seguente: la successione (f_n) converge a f se la distanza tra f_n e f tende a zero, cioè se

$$\lim_{n \rightarrow \infty} \|f - f_n\|_{\infty} = 0$$

N. B. È giunto il tempo di riprendere e approfondire le idee che sono state introdotte nella sezione 2.4.

Ora, il problema è che, in generale, non è detto che il limite di una successione di elementi che appartengono ad una classe sia ancora un elemento di quella classe. L'esempio paradigmatico sono i razionali: in generale, successioni di razionali non convergono necessariamente a numeri razionali. È per questa ragione che Cauchy introdusse la nozione di convergenza che prende il suo nome: per ogni $\epsilon > 0$ deve esistere un intero positivo N tale che per tutti $m, n > N$, $d(f_n, f_m) < \epsilon$. Si può dimostrare che ogni successione di Cauchy (f_n) in $C[a, b]$ converge ad un elemento di $C[a, b]$ (si veda i complementi 17.11.2).

Uno spazio normato, in cui ogni successione di Cauchy converge ad un elemento dello spazio stesso, è detto completo. Uno spazio di questo tipo è anche noto anche come *spazio di Banach*. Per esempio, $C[a, b]$, munito di norma uniforme, è uno *spazio vettoriale normato completo*, cioè uno spazio di Banach.

La completezza di uno spazio normato è regolata dal seguente teorema:

Uno spazio vettoriale normato è completo se e solo se ogni serie assolutamente convergente nello spazio è convergente, cioè se

$$\sum_{k=1}^{\infty} \|v_k\| < \infty \quad \Rightarrow \quad \sum_{k=1}^{\infty} v_k < \infty$$

(17.2)

Si osservi che il significato geometrico intuitivo di questo teorema è quello espresso dalla figura 2.10 del capitolo 2 (per una dimostrazione di questo fatto, si veda la 6 dei complementi 17.11.2).

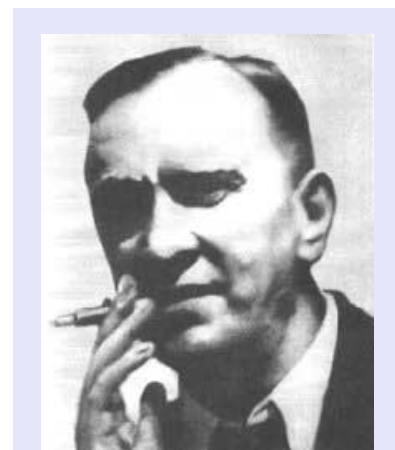
17.2 Spazi con prodotto scalare

Uno spazio di Banach, pur avendo una struttura per molti versi analoga a quella dello spazio euclideo finito-dimensionale, è privo di una nozione di prodotto scalare tra vettori. Ricordiamo dal capitolo 2 come sono fatti gli spazi vettoriali muniti di prodotto scalare ripresentando le caratteristiche salienti dello spazio euclideo complesso \mathbb{C}^n , lo spazio dei vettori $\mathbf{z} = (z_1, z_2, \dots, z_n)$, con componenti z_i complesse. Tra coppie di vettori qualunque \mathbf{z}, \mathbf{w} in questo spazio è definito il prodotto scalare

$$\langle \mathbf{z} | \mathbf{w} \rangle = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n \quad (17.3)$$

che gode delle seguenti proprietà:

- (i) $\langle \mathbf{z} | \mathbf{w} \rangle = \overline{\langle \mathbf{w} | \mathbf{z} \rangle}$ (simmetria coniugata)
- (ii) $\langle \mathbf{u} | \alpha \mathbf{z} + \beta \mathbf{w} \rangle = \alpha \langle \mathbf{u} | \mathbf{z} \rangle + \beta \langle \mathbf{u} | \mathbf{w} \rangle$ (linearità nel secondo argomento)



Stefan Banach (1892–1945) è stato un matematico polacco tra i più influenti del XX secolo. Noto al grande pubblico (si fa per dire) per il paradosso di Banach-Tarski, secondo cui esiste una decomposizione finita di una sfera tale che ricomponendone i pezzi si possono formare due sfere identiche a quella di partenza (e, naturalmente, la procedura può essere iterata).

(iii) $\langle \mathbf{z} | \mathbf{z} \rangle \geq 0$ (=0 sse $\mathbf{z} = 0$) (positività)

Si osservi che (i) e (ii) implicano

$$\langle \alpha \mathbf{z} + \beta \mathbf{w} | \mathbf{u} \rangle = \bar{\alpha} \langle \mathbf{z} | \mathbf{u} \rangle + \bar{\beta} \langle \mathbf{w} | \mathbf{u} \rangle$$

Il prodotto scalare in \mathbb{C}^n è dunque una forma definita su coppie di vettori di \mathbb{C}^n che è lineare in un argomento e antilineare nell'altro. Una forma di questo tipo è detta *hermitiana* o *sesquilineare*.

Se c'è un prodotto scalare tra vettori, c'è anche una nozione di lunghezza: la lunghezza del vettore $\mathbf{z} = (z_1, z_2, \dots, z_n)$ è data dal *teorema di Pitagora*

$$\|\mathbf{z}\| = \sqrt{|z_1|^2 + \dots + |z_n|^2} = \sqrt{\langle \mathbf{z} | \mathbf{z} \rangle}$$

e soddisfa le proprietà (i), (ii) e (iii) della norma, in particolare la disuguaglianza triangolare $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, che stabilisce che la somma delle lunghezze di due lati di un triangolo è maggiore o uguale della lunghezza del terzo lato.

La disuguaglianza triangolare si verifica facilmente a partire dalla (17.3). Ma c'è anche un modo più complicato per verificarla, che a posteriori risulta più profondo e utile. Questo modo si realizza in due passi. Primo passo: a partire dalle proprietà (i) (ii) e (iii) del prodotto scalare e dalla *definizione*

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}, \quad (17.4)$$

si dimostra la *disuguaglianza di Cauchy-Schwarz*

$$|\langle \mathbf{u} | \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|. \quad (17.5)$$

Secondo passo: si usa la disuguaglianza di Cauchy-Schwarz per dimostrare la disuguaglianza triangolare. Si vedano le dimostrazioni 3 e 4 dei complementi 17.11.2. Ma qual è il vantaggio di questo modo più complicato di ottenere la disuguaglianza triangolare? Un guadagno notevole di generalità!

Un qualunque spazio vettoriale in cui è definito un prodotto scalare, soddisfacente le proprietà (i), (ii) e (iii), risulta automaticamente uno spazio normato, con norma definita dalla (17.4). Infatti, le proprietà (i) e (ii) della norma sono banalmente verificate e la disuguaglianza triangolare vale in quanto conseguenza della disuguaglianza di Cauchy-Schwarz. *Qualunque riferimento alla dimensione finita dello spazio è stato così eliminato* e il fatto che uno spazio con prodotto scalare diventi uno spazio normato è indipendente dalla sua dimensione; in altre parole, vale anche per spazi di dimensione infinita.

Se lo spazio normato così ottenuto è completo (è cioè uno spazio di Banach rispetto alla norma indotta dal prodotto scalare), è

detto *spazio di Hilbert*; se questo non accade, cioè se lo spazio non è completo rispetto alla norma indotta dal prodotto scalare, è usualmente chiamato *spazio pre-hilbertiano*. Uno spazio di Hilbert è di solito denotato con il simbolo \mathcal{H} .

Teoremi dimostrati per un generico spazio di Hilbert varranno automaticamente per spazi di Hilbert concreti, di cui, nel seguito vedremo esempi particolarmente importanti.

Ecco un esempio di teorema che vale per uno spazio di Hilbert generico e che si dimostra facilmente a partire dalle proprietà del prodotto scalare e dalla disuguaglianza di Cauchy-Schwarz:

TEOREMA DI CONTINUITÀ DEL PRODOTTO SCALARE. Se $\{\mathbf{u}_n\}$ e $\{\mathbf{v}_n\}$ sono due successioni di vettori di \mathcal{H} che convergono in norma,

$$\lim_{n \rightarrow \infty} \mathbf{u}_n = \mathbf{u} \quad \lim_{n \rightarrow \infty} \mathbf{v}_n = \mathbf{v},$$

allora

$$\lim_{n \rightarrow \infty} \langle \mathbf{u}_n | \mathbf{v}_n \rangle = \langle \mathbf{u} | \mathbf{v} \rangle$$

17.3 Lo spazio ℓ^2

Lo spazio \mathbb{C}^n , con prodotto scalare dato da (17.3) è uno spazio di Hilbert finito-dimensionale. L'esempio più semplice di spazio di Hilbert infinito-dimensionale è lo spazio delle successioni a quadrato sommabile, vale a dire delle successioni infinite $\mathbf{z} = (z_1, z_2, \dots)$, con componenti z_k complesse, tali che

$$\|\mathbf{z}\| \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} |z_k|^2 < \infty \quad (17.6)$$

Questo spazio è usualmente chiamato ℓ^2 . In questo spazio il prodotto scalare è così definito

$$\langle \mathbf{z} | \mathbf{w} \rangle = \sum_{k=1}^{\infty} \bar{z}_k w_k \quad (17.7)$$

La serie a secondo membro converge in virtù della disuguaglianza di Cauchy-Schwarz (esercizio).

Lo spazio ℓ^2 è completo ed è quindi uno spazio di Hilbert. La dimostrazione di questo fatto è data nei complementi 17.11.2 (dimostrazione 5).

17.4 La notazione di Dirac

In meccanica quantistica, lo stato di un sistema fisico è rappresentato da un raggio in uno spazio di Hilbert complesso e separabile (daremo la definizione di spazio separabile nel seguito) o, equivalentemente, da un punto nello spazio di Hilbert proiettivo del sistema.



David Hilbert (1862–1943) è stato un matematico tedesco, probabilmente il più influente a cavallo tra il XIX e il XX secolo. Noto il per il suo approccio formalistico al problema dei fondamenti della matematica, si occupò anche di fisica, in particolare arrivò alle equazioni della relatività generale quasi parallelamente ad Einstein. L'azione, da cui seguono queste equazioni per principio di minimo, è detta "azione di Einstein-Hilbert". Sembra che abbia detto: "la fisica è troppo difficile per essere lasciata ai fisici".

Con in mente le applicazioni della teoria degli spazi di Hilbert alla meccanica quantistica, Dirac introdusse la notazione dei bra e dei ket per descrivere i vettori di uno spazio di Hilbert. Questa notazione può essere descritta nel seguente modo.

Incominciamo col complicarci la vita e rappresentiamo un vettore \mathbf{u} di \mathcal{H} con il simbolo $|\mathbf{u}\rangle$. Lo chiameremo vettore “ket”, ma resta sempre lo stesso vettore \mathbf{u} di prima. Adesso rappresentiamo lo stesso vettore \mathbf{u} con il simbolo $\langle\mathbf{u}|$, che chiameremo vettore “bra”, ma resta sempre, di nuovo, lo stesso vettore di prima. Introduciamo infine le seguenti regole per manipolare i simboli di bra e ket

$$|\alpha\mathbf{u} + \beta\mathbf{v}\rangle = \alpha|\mathbf{u}\rangle + \beta|\mathbf{v}\rangle \quad \langle\alpha\mathbf{u} + \beta\mathbf{v}| = \bar{\alpha}\langle\mathbf{u}| + \bar{\beta}\langle\mathbf{v}|$$

Il prodotto scalare (“braket”¹) risulta, simbolicamente, il prodotto dei simboli bra e ket,

$$\langle\mathbf{u}|\mathbf{v}\rangle \stackrel{\text{def}}{=} \langle\mathbf{u}|\mathbf{v}\rangle$$

La notazione rende così manifesto che il prodotto scalare è una forma hermitiana (lineare in un argomento e antilineare nell’altro).

Si potrebbe dare alla notazione simbolica anche un significato matematico e identificare i “bra” con il duale \mathcal{H}^* di \mathcal{H} (cioè lo spazio dei funzionali lineari continui su \mathcal{H}). Ma si tratta di una sofisticazione matematica inutile: poiché per il teorema di Riesz, \mathcal{H}^* è naturalmente identificato con \mathcal{H} stesso, è fuorviante pensare che per rappresentare lo stato di un sistema quantistico siano necessari due tipi distinti di vettori, i “bra” e i “ket” o che il prodotto scalare sia tra un prodotto misterioso di un “bra” e di un “ket”. In breve, la notazione di Dirac va presa per quello che è (indipendentemente da che cosa ne pensasse Dirac stesso): una notazione conveniente e utile, ma niente di più.

La notazione di Dirac è particolarmente conveniente per denotare l’operatore di proiezione ortogonale su un vettore \mathbf{v} con

$$\frac{|\mathbf{v}\rangle\langle\mathbf{v}|}{\langle\mathbf{v}|\mathbf{v}\rangle} = |\mathbf{e}_v\rangle\langle\mathbf{e}_v|, \quad \mathbf{e}_v = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

per cui la proiezione di \mathbf{u} lungo \mathbf{v} risulta

$$|\mathbf{e}_v\rangle\langle\mathbf{e}_v|\mathbf{u}\rangle.$$

Con la notazione di Dirac, risulta particolarmente trasparente che il quadrato del proiettore sia il proiettore stesso: $|\mathbf{e}_v\rangle\langle\mathbf{e}_v||\mathbf{e}_v\rangle\langle\mathbf{e}_v| = |\mathbf{e}_v\rangle\langle\mathbf{e}_v|$, in quanto $\langle\mathbf{e}_v|\mathbf{e}_v\rangle \stackrel{\text{def}}{=} \langle\mathbf{e}_v|\mathbf{e}_v\rangle = 1$.

17.5 Basi ortonormali

Una successione (limitata o infinita) di vettori $\{\mathbf{e}_n\}$ in uno spazio di Hilbert \mathcal{H} per cui $\langle\mathbf{e}_n|\mathbf{e}_m\rangle = 0$ quando $n \neq m$ si chiama *sistema*

¹ Risulta così svelato l’arcano della terminologia “bra” e “ket”: “braket” in inglese vuol dire “parentesi” (e ci sono ovviamente due tipi di parentesi, quelle che hanno la gobba a sinistra e quelle che la hanno a destra). Questa terminologia, che riflette anche un sottile humor, è di Dirac.

ortogonale. Se inoltre si ha $\langle \mathbf{e}_n | \mathbf{e}_n \rangle = \|\mathbf{e}_n\| = 1$, il sistema si chiama *ortonormale*. Riassumendo, per tale sistema si ha

$$\langle \mathbf{e}_n | \mathbf{e}_m \rangle = \delta_{nm}. \quad (17.8)$$

Se un sistema ortogonale è formato da un numero finito n di vettori, questi vettori sono linearmente indipendenti. Infatti, considerando il prodotto scalare dell'elemento \mathbf{e}_i , $i = 1, \dots, n$ con $a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \dots + a_n\mathbf{e}_n = 0$, si deduce che $a_i = 0$, per $i = 1, \dots, n$. Quindi, in uno spazio di Hilbert di dimensione finita N , un sistema ortonormale è costituito da non più di N elementi.

Vogliamo adesso chiarire quali requisiti deve avere un sistema ortonormale per essere l'analogo infinito-dimensionale di una *base* ortonormale, concetto che si assume noto per gli spazi di dimensione finita. In primo luogo, osserviamo che una scrittura del tipo

$$\mathbf{u} = \sum_{n=1}^{\infty} c_n \mathbf{e}_n \quad (17.9)$$

va intesa nel senso della convergenza in norma della serie a secondo membro, significa cioè

$$\lim_{N \rightarrow \infty} \left\| \mathbf{u} - \sum_{n=1}^N c_n \mathbf{e}_n \right\| = 0.$$

In secondo luogo, osserviamo che se la serie converge, allora

$$c_n = \langle \mathbf{e}_n | \mathbf{u} \rangle \quad n = 1, 2, \dots$$

Questo si verifica moltiplicando scalarmente per \mathbf{e}_n ambo i membri della (17.9) e sfruttando le condizioni di ortogonalità dei vettori della base:²

$$\langle \mathbf{e}_m | \mathbf{u} \rangle = \langle \mathbf{e}_m | \sum_{n=1}^{\infty} c_n \mathbf{e}_n \rangle = \sum_{n=1}^{\infty} c_n \langle \mathbf{e}_m | \mathbf{e}_n \rangle = c_m$$

I coefficienti c_n sono dunque le *coordinate* di \mathbf{u} rispetto al sistema ortogonale $\{\mathbf{e}_n\}$, nello stesso senso in cui questo vale per uno spazio di dimensione finita. Questo fatto può essere espresso in forma compatta riscrivendo la (17.9) con la notazione di Dirac:

$$|\mathbf{u}\rangle = \sum_n |\mathbf{e}_n\rangle \langle \mathbf{e}_n | \mathbf{u} \rangle \quad (17.10)$$

È importante osservare che la precedente relazione vale sotto l'ipotesi che $|\mathbf{u}\rangle$ sia uguale alla serie a secondo membro, ma non sempre questo è vero. Per esempio, i vettori \mathbf{i} e \mathbf{j} dello spazio euclideo tridimensionale formano un sistema ortogonale, ma se decomponiamo rispetto a questo sistema un vettore \mathbf{u} che non giace nel piano da essi generato, la somma $(\mathbf{u} \cdot \mathbf{i})\mathbf{i} + (\mathbf{u} \cdot \mathbf{j})\mathbf{j}$ (cioè il secondo membro della

² A rigore, si dovrebbe procedere con cautela e scrivere

$$\langle \mathbf{e}_m | \mathbf{u} \rangle = \lim_{N \rightarrow \infty} \langle \mathbf{e}_m | \sum_{n=1}^N c_n \mathbf{e}_n \rangle.$$

Per il teorema di continuità del prodotto scalare (sezione 17.2), il limite può essere scambiato con la somma, quindi

$$\langle \mathbf{e}_m | \mathbf{u} \rangle = \lim_{N \rightarrow \infty} \sum_{n=1}^N c_n \langle \mathbf{e}_m | \mathbf{e}_n \rangle = c_m$$

(17.10), non ci restituisce il vettore \mathbf{u} , ma la sua proiezione ortogonale sul piano generato da \mathbf{i} e \mathbf{j} .

Nel caso finito-dimensionale, un sistema ortonormale è una base, quando “non si perde niente”, cioè quando l’operatore $\sum_n |\mathbf{e}_n\rangle\langle\mathbf{e}_n|$ non proietta il vettore su un sottospazio, ma è l’identità su tutto lo spazio. Questa caratterizzazione si estende in modo naturale al caso infinito-dimensionale³

Un sistema ortogonale $\{\mathbf{e}_n\}$ è completo in \mathcal{H} o, equivalentemente, è una base ortonormale in \mathcal{H} , quando QUALUNQUE vettore $\mathbf{u} \in \mathcal{H}$ può essere espresso come

$$\mathbf{u} = \sum_n |\mathbf{e}_n\rangle\langle\mathbf{e}_n | \mathbf{u} \rangle. \quad (17.11)$$

Abbiamo detto che l’uguaglianza (17.10) va intesa come convergenza in norma,

$$\lim_{N \rightarrow \infty} \left\| \mathbf{u} - \sum_{n=1}^N |\mathbf{e}_n\rangle\langle\mathbf{e}_n | \mathbf{u} \rangle \right\| = 0.$$

Calcoliamo adesso il quadrato della norma della differenza tra il vettore \mathbf{u} e la somma parziale N -esima della serie:

$$\begin{aligned} \left\| \mathbf{u} - \sum_{n=1}^N |\mathbf{e}_n\rangle\langle\mathbf{e}_n | \mathbf{u} \rangle \right\|^2 &= \langle \mathbf{u} - \sum_{n=1}^N \mathbf{e}_n \langle \mathbf{e}_n | \mathbf{u} \rangle | \mathbf{u} - \sum_{n=1}^N \mathbf{e}_n \langle \mathbf{e}_n | \mathbf{u} \rangle \rangle \\ &= \langle \mathbf{u} | \mathbf{u} \rangle - 2 \sum_{n=1}^N |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 + \sum_{n=1}^N |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 \\ &= \|\mathbf{u}\|^2 - \sum_{n=1}^N |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 \end{aligned} \quad (17.12)$$

Da cui segue che si ha convergenza in norma per $N \rightarrow \infty$ se e solo se

$$\sum_{n=1}^{\infty} |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 = \|\mathbf{u}\|^2. \quad (17.13)$$

Questa uguaglianza è nota come *identità di Parseval*. Perciò

Condizione necessaria e sufficiente affinché un sistema ortonormale $\{\mathbf{e}_n\}$ in \mathcal{H} sia una base, è che l’identità di Parseval sia soddisfatta da ogni vettore \mathbf{u} in \mathcal{H} .

(17.14)

Si osservi che, indipendentemente dal fatto che la serie converga a \mathbf{u} , il secondo membro della (17.12) è sempre positivo. Passando al limite $N \rightarrow \infty$ della (17.12), si ottiene la disuguaglianza

$$\sum_{n=1}^{\infty} |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 \leq \|\mathbf{u}\|^2, \quad (17.15)$$

³ In uno spazio di dimensione finita, decidere che un sistema ortogonale è una base è una semplice questione di conteggio: se lo spazio ha dimensione N , un sistema ortogonale di N elementi è automaticamente una base. Nel caso infinito il conteggio non funziona. Non basta infatti che un sistema abbia infiniti elementi per essere una base. Per esempio, le funzioni

$$\frac{1}{\sqrt{2\pi}} \cos nx$$

per $-\pi < x \leq \pi$, sono un sistema ortonormale, per la (16.3), ma non formano una base. Questo sistema non è completo (“si perde qualcosa”): con combinazioni lineari, anche infinite, di queste funzioni, si possono formare solo funzioni pari (“si perdono le funzioni dispari”).

nota come *disuguaglianza di Bessel*, che vale dunque sotto la sola ipotesi che il sistema $\{\mathbf{e}_n\}$ sia ortonormale (e non necessariamente una base).

AVVERTENZA. L'esistenza di un sistema ortonormale completo *numerabile* non può essere data per scontata. In effetti, non è conseguenza delle proprietà del prodotto scalare e della completezza della norma da esso indotta. Se uno spazio di Hilbert ha una base ortogonale completa numerabile è detto *separabile*. Nel seguito assumeremo tacitamente che questa condizione sia soddisfatta e identificheremo la nozione di spazio di Hilbert con quella di spazio di Hilbert separabile.

SISTEMA ORTOGONALE CHIUSO. Un'altra nozione importante è la nozione di sistema ortogonale *chiuso*, condizione che si realizza quando non esiste alcun vettore non nullo nello spazio di Hilbert che sia ortogonale a tutti gli elementi del sistema. In altre parole, un sistema $\{\mathbf{e}_n\}$ è chiuso se dalle relazioni

$$\langle \mathbf{e}_n | \mathbf{u} \rangle = 0, \quad n = 1, 2, 3, \dots$$

segue che $\mathbf{u} = 0$.

Chiusura e completezza di un sistema ortonormale sono identificate dal seguente teorema (si veda la dimostrazione 7 dei complementi 17.11.2):

Condizione necessaria e sufficiente affinché un sistema ortonormale sia una base è che sia chiuso.	(17.16)
--	---------

17.6 Teorema di Riesz-Fisher

Abbiamo visto che, fissato un sistema ortonormale $\{\mathbf{e}_n\}$, ad un vettore \mathbf{u} è associata la successione delle sue coordinate

$$c_n = \langle \mathbf{e}_n | \mathbf{u} \rangle, \quad n = 1, 2, 3, \dots \quad (17.17)$$

Questa successione è tale, per la disuguaglianza di Bessel,

$$\sum_{n=1}^{\infty} |\langle \mathbf{e}_n | \mathbf{u} \rangle|^2 \leq \|\mathbf{u}\|^2,$$

che la serie

$$\sum_{n=1}^{\infty} |c_n|^2 \quad (17.18)$$

è convergente.

Consideriamo ora una successione c_1, c_2, c_3, \dots tale che la serie (17.18) sia convergente. Esiste un vettore \mathbf{u} per cui valgono le (17.17)? La risposta è affermativa ed è data dal seguente teorema:

TEOREMA DI RIESZ-FISHER. Sia $\{\mathbf{e}_n\}$ una base ortonormale in un spazio di Hilbert \mathcal{H} . Se $\{c_n\}$ è una successione di numeri tale che la serie $\sum_{n=1}^{\infty} |c_n|^2$ converge, allora esiste allora esiste uno ed un solo vettore $\mathbf{u} \in \mathcal{H}$ tale che

$$\mathbf{u} = \sum_{n=1}^{\infty} c_n \mathbf{e}_n \quad e \quad c_n = \langle \mathbf{e}_n | \mathbf{u} \rangle$$

(Si veda la dimostrazione 8 dei complementi 17.11.2.)

17.7 Universalità dello spazio ℓ^2

Una conseguenza importante del teorema di Riesz-Fisher è che tutti gli spazi di Hilbert infinito dimensionali⁴ sono *unitariamente equivalenti* allo spazio ℓ^2 delle successioni a quadrato sommabile.

Prima di tutto, chiariamo la terminologia. Due spazi di Hilbert \mathcal{H} e \mathcal{H}' sono detti unitariamente equivalenti se esiste una corrispondenza biunivoca tra i vettori \mathbf{u} di \mathcal{H} e i vettori \mathbf{u}' di \mathcal{H}' tale che:

- (1) se a \mathbf{u} corrisponde \mathbf{u}' e a \mathbf{v} corrisponde \mathbf{v}' , al vettore $\alpha \mathbf{u} + \beta \mathbf{v}$ di \mathcal{H} corrisponde il vettore $\alpha \mathbf{u}' + \beta \mathbf{v}'$ di \mathcal{H}' qualunque siano i numeri α e β ;
- (2) se a \mathbf{u} corrisponde \mathbf{u}' , allora

$$\|\mathbf{u}\|_{\mathcal{H}} = \|\mathbf{u}'\|_{\mathcal{H}'}$$

Adesso rendiamo esplicito il senso dell'affermazione iniziale. Nella sezione 17.3 abbiamo mostrato che lo spazio ℓ^2 delle successioni a quadrato sommabile è uno spazio di Hilbert infinito-dimensionale. Una sistema ortonormale in ℓ^2 è dato dai vettori $\hat{\mathbf{e}}_n$, $n = 1, 2, 3, \dots$ le cui componenti sono tutte 0 eccetto per l' n -esima, che vale 1:

$$\hat{\mathbf{e}}_1 = (1, 0, 0, 0, \dots), \quad \hat{\mathbf{e}}_2 = (0, 1, 0, 0, \dots), \quad \hat{\mathbf{e}}_3 = (0, 0, 1, 0, \dots), \quad \dots$$

Tale sistema è ovviamente chiuso e quindi completo; dunque $\{\hat{\mathbf{e}}_n\}$ è una base ortonormale in ℓ^2 . Se \mathcal{H} uno spazio di Hilbert e $\{\mathbf{e}_n\}$ una base ortonormale in esso, si consideri la corrispondenza

$$\mathbf{e}_1 \leftrightarrow \hat{\mathbf{e}}_1, \quad \mathbf{e}_2 \leftrightarrow \hat{\mathbf{e}}_2, \quad \mathbf{e}_3 \leftrightarrow \hat{\mathbf{e}}_3, \quad \dots$$

e la si estenda per linearità

$$\mathbf{u} = \sum_{n=1}^{\infty} c_n \mathbf{e}_n \leftrightarrow \hat{\mathbf{u}} = \sum_{n=1}^{\infty} c_n \hat{\mathbf{e}}_n = (c_1, c_2, c_3, \dots)$$

Allora a \mathbf{u} in \mathcal{H} è associato il vettore in ℓ^2

$$\hat{\mathbf{u}} = \sum_{n=1}^{\infty} c_n \hat{\mathbf{e}}_n = (c_1, c_2, c_3, \dots), \quad c_n = \langle \mathbf{e}_n | \mathbf{u} \rangle$$

⁴ Occorrerebbe aggiungere "separabili", ma nella sezione precedente abbiamo convenuto di identificare la nozione di spazio di Hilbert con quella di spazio di Hilbert separabile.

e

$$\|\mathbf{u}\|_{\mathcal{H}}^2 = \sum_{n=1}^{\infty} |c_n|^2 = \|\hat{\mathbf{u}}\|_{\ell^2}^2$$

segue dall'identità di Parseval. Viceversa, per il teorema di Riesz-Fisher, a

$$\hat{\mathbf{u}} = \sum_{n=1}^{\infty} c_n \hat{\mathbf{e}}_n = (c_1, c_2, c_3, \dots)$$

in ℓ^2 è associato il vettore \mathbf{u} in \mathcal{H} . La corrispondenza tra \mathcal{H} e ℓ^2 è dunque biunivoca e isometrica, che è quanto si voleva dimostrare.

L'ovvio corollario di questo è che *tutti gli spazi di Hilbert sono unitariamente equivalenti*. Questo fatto risulta meno misterioso di quanto sembri se si tiene conto della seguente analogia. Lo spazio di Hilbert \mathcal{H} , astrattamente definito, è l'analogo infinito-dimensionale dello spazio euclideo tri-dimensionale E^3 della geometria elementare. Fissare una base ortormale $\{\mathbf{e}_n\}$ in \mathcal{H} è come fissare il sistema di versori \mathbf{i} , \mathbf{j} e \mathbf{k} in E^3 . Lo spazio ℓ^2 è come lo spazio \mathbb{R}^3 delle coordinate (x, y, z) dei vettori rispetto al sistema di riferimento individuato dai versori \mathbf{i} , \mathbf{j} e \mathbf{k} . Inoltre, la corrispondenza tra \mathcal{H} e ℓ^2 è analoga alla corrispondenza tra vettori e terne di coordinate,

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \leftrightarrow (x, y, z).$$

Infine, l'identità di Parseval è la versione infinito-dimensionale del teorema di Pitagora $\|\mathbf{r}\|^2 = x^2 + y^2 + z^2$.

17.8 Proiezioni ortogonali

Un sottospazio \mathcal{M} di \mathcal{H} è detto *chiuso* se è un sottospazio vettoriale *completo* nella norma, cioè tale che se $\{\mathbf{u}_n\}$ è una successione convergente di elementi di \mathcal{M} , anche il suo limite in norma è in \mathcal{M} . In altre parole,

$$\text{se } \mathbf{u}_n \in \mathcal{M} \text{ e } \|\mathbf{u} - \mathbf{u}_n\| \rightarrow 0 \text{ allora } \mathbf{u} \in \mathcal{M}$$

Dunque, un sottospazio chiuso è, esso stesso, uno spazio di Hilbert.

Concordemente con le nozioni usuali di geometria elementare, due elementi di uno spazio con prodotto scalare sono detto *ortogonali*, $\mathbf{u} \perp \mathbf{v}$, se $\langle \mathbf{u} | \mathbf{v} \rangle = 0$. Si dimostra facilmente che il *complemento ortogonale*

$$\mathcal{M}^\perp = \{\mathbf{u} \in \mathcal{H} \text{ tale che } \mathbf{u} \perp \mathbf{v} \text{ per qualunque } \mathbf{v} \in \mathcal{M}\}$$

di un qualunque sottoinsieme \mathcal{M} di \mathcal{H} è un sottospazio chiuso. Il seguente teorema è particolarmente importante:

TEOREMA DELLA PROIEZIONE ORTOGONALE. *Se \mathcal{M} è un sottospazio chiuso di uno spazio di Hilbert \mathcal{H} , allora ciascun $\mathbf{v} \in \mathcal{H}$ può essere unicamente rappresentato come*

$$\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \tag{17.19}$$

dove $\mathbf{v}_{\parallel} \in \mathcal{M}$ e $\mathbf{v}_{\perp} \in \mathcal{M}^{\perp}$. Inoltre, \mathbf{v}_{\parallel} (che è chiamato la proiezione ortogonale di \mathbf{v} su \mathcal{M}) soddisfa

$$\|\mathbf{v} - \mathbf{v}_{\parallel}\| < \|\mathbf{v} - \mathbf{u}\| \quad (17.20)$$

per qualunque altro elemento $\mathbf{u} \in \mathcal{M}$

Il contenuto del teorema è rappresentato nella figura 17.1. Questo fatto è geometricamente ovvio in uno spazio euclideo, sia per quel che riguarda la (17.19) sia per la (17.20), che esprime il fatto che la distanza minima tra il punto Q individuato da \mathbf{v} e il sottospazio \mathcal{M} è quella che si percorre partendo da Q e muovendosi nella direzione ortogonale a \mathcal{M} . La cosa interessante è che questo teorema è vero, pari pari in uno spazio di Hilbert qualunque. Non ne daremo la dimostrazione generale, ma, nel seguito, in un caso particolare.

Dato \mathcal{M} , l'operatore P che proietta un qualunque vettore \mathbf{v} nel vettore \mathbf{v}_{\parallel} ,

$$P\mathbf{v} = \mathbf{v}_{\parallel},$$

è chiamato proiettore ortogonale di \mathcal{H} su \mathcal{M} . Soddisfa le ovvie proprietà geometriche di una proiezione ortogonale, in particolare è idempotente

$$P^2 = P.$$

Inoltre è continuo: se $\|\mathbf{u}_n - \mathbf{u}\| \rightarrow 0$, allora $\|P(\mathbf{u}_n) - P(\mathbf{u})\| \rightarrow 0$. Si osservi che il proiettore su \mathcal{M}^{\perp} è $I - P$, dove I è l'operatore identità.

Il teorema della proiezione garantisce l'unicità della decomposizione di \mathcal{H} in somma ortogonale di un suo sottospazio chiuso \mathcal{M} e del suo complemento ortogonale \mathcal{M}^{\perp} . Per esprimere questo fatto, di solito si scrive

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$$

17.9 Approssimazione in media quadratica

Sia $\{\mathbf{e}_n\}$ una base ortonormale in \mathcal{H} , si fissi un numero N (ad esempio $N = 10$) e si consideri l'insieme di tutti i vettori

$$\sum_{n=1}^N z_n \mathbf{e}_n \quad (17.21)$$

per tutte le successioni (z_1, z_2, \dots, z_N) di numeri complessi. L'insieme \mathcal{M} così ottenuto è un sottospazio finito-dimensionale di \mathcal{H} (e quindi chiuso), di dimensione N (e quindi isometricamente isomorfo a \mathbb{C}^N).

Sia \mathbf{u} un qualunque vettore di \mathcal{H} . Si osservi che in notazione di Dirac, l'operatore di proiezione P su \mathcal{M} può essere espresso come

$$P = \sum_{n=1}^N |\mathbf{e}_n\rangle\langle\mathbf{e}_n|.$$

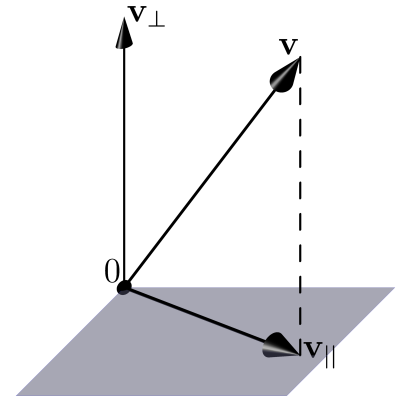


Figura 17.1: Decomposizione ortogonale di un vettore \mathbf{v} come somma della sua componente \mathbf{v}_{\parallel} in \mathcal{M} (il piano orizzontale in figura) e la sua componente verticale \mathbf{v}_{\perp} nello spazio \mathcal{M}^{\perp} (la direzione verticale in figura). La lunghezza della linea tratteggiata è la norma di $\mathbf{v} - \mathbf{v}_{\parallel}$. Il vettore \mathbf{v}_{\parallel} è il vettore nel piano \mathcal{M} che si trova alla distanza minima da \mathbf{v} . Qualunque altro vettore nel piano si trova ad una distanza maggiore da \mathbf{v} .

Infatti la proiezione ortogonale di \mathbf{u} su \mathcal{M} è

$$P\mathbf{u} = \sum_{n=1}^N |\mathbf{e}_n\rangle \langle \mathbf{e}_n | \mathbf{u} \rangle$$

Adesso ci domandiamo quale sia il vettore in \mathcal{M} che meglio approssima \mathbf{u} . Per rispondere a questa domanda usiamo il metodo dei minimi quadrati. Consideriamo la norma al quadrato della differenza D tra \mathbf{u} e un generico vettore di \mathcal{M} , cioè del tipo (17.21),

$$D = \left\| \mathbf{u} - \sum_{n=1}^N z_n \mathbf{e}_n \right\|^2$$

e determiniamo quali siano i valori dei coefficienti z_n tale che la differenza sia minima. Si ha,

$$\begin{aligned} D &= \left\langle \mathbf{u} - \sum_{n=1}^N z_n \mathbf{e}_n \mid \mathbf{u} - \sum_{n=1}^N z_n \mathbf{e}_n \right\rangle \\ &= \|\mathbf{u}\|^2 - \sum_{n=1}^N \bar{z}_n \langle \mathbf{e}_n | \mathbf{u} \rangle - \sum_{n=1}^N z_n \langle \mathbf{u} | \mathbf{e}_n \rangle + \sum_{n=1}^N |z_n|^2 \end{aligned}$$

Il minimo di D (se esiste) è in corrispondenza dell'annullarsi delle derivate prime $\partial f / \partial z_k = 0$ e $\partial f / \partial \bar{z}_k = 0$ (trattando le z e le \bar{z} come variabili indipendenti). Si ha

$$\begin{aligned} \frac{\partial D}{\partial z_k} &= -\langle \mathbf{u} | \mathbf{e}_k \rangle + \bar{z}_k = 0 \quad \Rightarrow \quad \bar{z}_k = \langle \mathbf{u} | \mathbf{e}_k \rangle \\ \frac{\partial D}{\partial \bar{z}_k} &= -\langle \mathbf{e}_k | \mathbf{u} \rangle + z_k = 0 \quad \Rightarrow \quad z_k = \langle \mathbf{e}_k | \mathbf{u} \rangle \end{aligned}$$

Si può verificare che per i valori così ottenuti si ha davvero un minimo e non un massimo o un punto sella (esercizio a casa). Il vettore in \mathcal{M} a distanza minima da \mathbf{u} è proprio

$$P\mathbf{u} = \sum_{n=1}^N |\mathbf{e}_n\rangle \langle \mathbf{e}_n | \mathbf{u} \rangle$$

in accordo con (17.20). La miglior approssimazione di un vettore in uno spazio più piccolo di quello in cui vive il vettore è dunque la proiezione del vettore stesso in quel sottospazio!

Ogni qual volta confrontiamo le nostre teorie con l'esperienza, o facciamo simulazioni al computer, siamo sempre costretti a lavorare con spazi piccoli, cioè finiti, anche se le nostre teorie fisiche sono formulate in spazi infiniti. Non dovrebbe quindi sorprendere che il teorema della proiezione sia alla base di molti metodi di approssimazione che si utilizzano in fisica e nelle scienze applicate.