Capitolo 4

Date: 2020-05-08 13:03 +0200 Revision: 273 : 251bdd4389bd

Spazi di Hilbert

4.1 Spazi prehilbertiani

Vogliamo ora parlare di una categoria di spazi molto importanti dal punto di vista delle applicazioni alla meccanica quantistica: i cosiddetti spazi di Hilbert, in cui si può definire un prodotto scalare che arricchisce notevolmente la struttura di spazio vettoriale normato, e per certi aspetti rende tali spazi geometricamente simili agli spazi vettoriali finito-dimensionali.

Definizione 4.1 *Sia X spazio vettoriale su $\mathbb{R}$. Si dice* forma bilineare *su X una applicazione $g : X \times X \to \mathbb{R}$ che sia lineare in entrambe le variabili, ossia che*

$$g(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 g(x, y_1) + \alpha_2 g(x, y_2) \tag{4.1}$$

$$g(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 g(x_1, y) + \alpha_2 g(x_2, y) \tag{4.2}$$

$$\tag{4.3}$$

per ogni $x, x_1, x_2, y, y_1, y_2 \in X$ e $\alpha_1, \alpha_2 \in \mathbb{R}$.

Definizione 4.2 *Una forma bilineare g su X si dice* simmetrica *se*

$$g(x, y) = g(y, x) \qquad \textit{per ogni } x, y \in X \tag{4.4}$$

Nel caso di uno spazio vettoriale sul campo dei numeri complessi, le forme $g : X \times X \to \mathbb{C}$ si definiscono in modo leggermente diverso: alla linearità sulla prima variabile si sostituisce la *semilinearità* (detta anche *antilinearità*).

Definizione 4.3 *Sia X spazio vettoriale su $\mathbb{C}$. Si dice* forma sesquilineare *su X una applicazione $g : X \times X \to \mathbb{C}$ che sia antilineare nella prima variabile e lineare nella seconda, ossia che*

$$g(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1^* g(x_1, y) + \alpha_2^* g(x_2, y) \tag{4.5a}$$

$$g(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 g(x, y_1) + \alpha_2 g(x, y_2) \tag{4.5b}$$

per ogni $x, x_1, x_2, y, y_1, y_2 \in X$ e $\alpha_1, \alpha_2 \in \mathbb{C}$.

Definizione 4.4 *Una forma bilineare g su X si dice* hermitiana *se*

$$g(x, y) = g(y, x)^* \qquad \textit{per ogni } x, y \in X \tag{4.6}$$

OSSERVAZIONE: L'avere imposto la condizione di hermiticità nel caso di campo complesso garantisce che $g(x, x)$ sia reale per ogni $x \in X$. Per consistenza bisogna allora imporre l'antilinearità in una delle due variabili di g. I fisici scelgono la prima, i matematici la seconda.

Notiamo che la definizione di forma simmetrica nel caso reale si ottiene dalla definizione di forma hermitiana nel caso complesso limitando il campo degli scalari ad $\mathbb{R}$, in cui la coniugazione complessa in $\mathbb{C}$ diventa la semplice identità in $\mathbb{R}$. Pertanto possiamo, senza perdita di generalità, trattare solamente il caso di forma hermitiana in campo complesso, precisando se necessario le semplificazioni che si hanno nel caso reale.

Definizione 4.5 *Una forma sesquilineare g su X si dice* non degenere *se*

$$g(x, y) = 0 \quad \textit{per ogni } x \in X \quad \Longrightarrow \quad y = 0\,. \tag{4.7}$$

In caso contrario la forma è detta degenere

Esempio: Supponiamo X uno spazio vettoriale di dimensione finita n sul campo dei complessi, ed $(e_k)_{1\le k\le n}$ una sua base. Allora, se:

$$x = \sum_{j=1}^{n} x_j e_j\ , \qquad y = \sum_{k=1}^{n} y_k e_k\ ,$$

per una qualsiasi forma sesquilineare g, usando le proprietà di linearità e antilinearità :

$$g(x, y) = \sum_{j,k=1}^{n} x_j^* y_k\, g(e_j, e_k) = \sum_{j,k=1}^{n} x_j^* y_k\, G_{jk} \tag{4.8}$$

e la forma sesquilineare è completamente determinata dai valori assunti sui vettori di base, cioè dalla matrice $G_{jk} := g(e_j, e_k)$. Se g è hermitiana, si ha $G_{jk} = G^*_{kj}$, e se g è non degenere la matrice G è invertibile. Viceversa, assegnata una base ed una matrice quadrata $n \times n$ di numeri complessi o reali, posso costruire una forma sesquilineare tramite la formula (4.8) e l'hermiticità della matrice e la sua invertibilità comportano rispettivamente l'hermiticità e la non degenerazione della forma sesquilineare. Sostanzialmente, sfruttando l'equivalenza di uno spazio vettoriale finito-dimensionale con lo spazio $\mathbb{C}^n$, le proprietà di una forma sesquilineare sono riconducibili alle caratteristiche di una matrice.

Teorema 4.1 (identità di polarizzazione) *Per ogni forma sesquilineare (non necessariamente hermitiana) in $\mathbb{C}$ vale*

$$g(x, y) = \frac{1}{4}\left[g(x+y, x+y) - g(x-y, x-y) - \mathrm{i}g(x+\mathrm{i}y, x+\mathrm{i}y) + \mathrm{i}g(x-\mathrm{i}y, x-\mathrm{i}y)\right]\ . \tag{4.9}$$

Per ogni forma bilineare e simmetrica in $\mathbb{R}$ vale

$$g(x, y) = \frac{1}{2}\left[g(x+y, x+y) - g(x, x) - g(y, y)\right]\ . \tag{4.10}$$

Dimostrazione: Verifica diretta. Notare che, per la forma reale, è richiesta anche la simmetria per la validità della formula. C.V.D.

Dalle identità di polarizzazione segue il fatto notevole che una forma sesquilineare sia completamente determinata dai valori che assume per $x = y$:

Teorema 4.2
(a) Una forma sesquilineare complessa è completamente determinata dai suoi valori diagonali $g(x, x)$ al variare di $x \in X$.
(b) g è hermitiana se e solo se $g(x, x) \in \mathbb{R}$ per ogni $x \in X$.

Analogamente, per una forma bilineare e simmetrica:

Teorema 4.3 *Una forma bilineare e simmetrica reale è completamente determinata dai suoi valori diagonali $g(x, x)$ al variare di $x \in X$.*

Definizione 4.6 *Una forma sesquilineare hermitiana è detta* definita positiva *se*

$$g(x, x) \ge 0 \qquad \forall\, x \in X\ , \qquad g(x, x) = 0 \quad \Longleftrightarrow \quad x = 0\,. \tag{4.11}$$

Una forma sesquilineare hermitiana definita positiva è anche detta prodotto scalare hermitiano. *Una forma bilineare simmetrica definita positiva in uno spazio reale è invece detta* prodotto scalare euclideo. *Indicheremo con $(x|y) = g(x, y)$ il prodotto scalare.*

OSSERVAZIONE: Una forma definita positiva, sia essa sesquilineare hermitiana o bilineare simmetrica, è necessariamente non degenere.

OSSERVAZIONE: Se $\mathbb{K} = \mathbb{C}$, la positività di g implica l'hermiticità, in quanto un numero positivo è reale, mentre se $\mathbb{K} = \mathbb{R}$ l'ipotesi di simmetria non è deducibile dalla positività, per cui deve essere richiesta esplicitamente.

Definizione 4.7 *Uno spazio vettoriale H sul quale è definito un prodotto scalare hermitiano si chiama* spazio pre-hilbertiano.

Il prodotto scalare elementare tra due vettori dello spazio euclideo ordinario, definito come il prodotto dei moduli (o norme) dei vettori per il coseno dell'angolo compreso è un'utile guida intuitiva per la nozione astratta che stiamo presentando, e ce ne serviremo frequentemente.

Vogliamo ora provare che, se $(\cdot|\cdot)$ è un prodotto scalare su H, la quantità

$$||x|| := (x|x)^{1/2} \tag{4.12}$$

definisce una norma su H. È banale verificare che è $||\lambda x|| = |\lambda|\,||x||$ per ogni $\lambda \in \mathbb{K}$, $x \in H$. Non è immediata la sub-additività (o disuguaglianza triangolare). Per dimostrarla, proviamo prima un'importante risultato.

Teorema 4.4 (disuguaglianza di Schwarz) *Sia H uno spazio vettoriale pre-hilbertiano. Allora*

$$|(x|y)| \leq (x|x)^{1/2}\,(y|y)^{1/2} \qquad \textit{per ogni } x, y \in H\ . \tag{4.13}$$

Inoltre vale l'uguaglianza se e solo se i vettori x e y sono uno multiplo dell'altro: $x = \lambda y$ con $\lambda \in \mathbb{K}$.

Dimostrazione: Se $y = 0$ la disuguaglianza è banalmente soddisfatta, poiché entrambi i membri sono nulli. Se invece $y \neq 0$, consideriamo al variare di $\alpha \in \mathbb{K}$ l'espressione $(x - \alpha y|x - \alpha y)$; questa è sempre ≥ 0, e vale 0 se e solo se $x = \alpha y$ per un $\alpha \in \mathbb{K}$ (fig. 4.1a). Inoltre, pensando al prodotto scalare elementare, tale espressione rappresenta il quadrato del vettore $x - \alpha y$, e dovrebbe essere minima per quell'α tale che αy sia la proiezione ortogonale di x sulla retta di y (fig. 4.1b), cioè con

$$\alpha y = (\hat{y}|x)\hat{y} = (\frac{y}{||y||}|x)\frac{1}{||y||}y\ , \quad \hat{y} = \frac{y}{||y||} \qquad \Longrightarrow \qquad \alpha = \frac{(y|x)}{(y|y)} \tag{4.14}$$

Si ha ora

$$0 \leq (x - \alpha y|x - \alpha y) = (x|x) - \alpha(x|y) - \alpha^*(y|x) + |\alpha|^2(y|y)$$

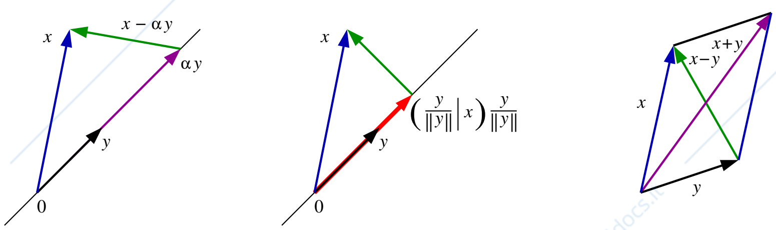
e, con il valore di α determinato appena prima nell'eq. (4.14),

$$0 \leq (x|x) - \frac{(y|x)(x|y)}{(y|y)} - \frac{(x|y)(y|x)}{(y|y)} + \frac{|(x|y)|^2}{(y|y)^2}(y|y)\ .$$

Osservando che $(x|y)(y|x) = (x|y)(x|y)^* = |(x|y)|^2$ e moltiplicando per $(y|y)$, troviamo

$$0 \leq (x|x)(y|y) - |(x|y)|^2 \qquad \Longrightarrow \qquad |(x|y)|^2 \leq (x|x)(y|y)$$

da cui segue la tesi estraendo le radici quadrate. L'affermazione sull'uguaglianza si lascia come esercizio. C.V.D.



(a) (b) (c)

Dimostriamo ora la sub-additività della norma.

Teorema 4.5 (disuguaglianza di Minkowski) *Sia H uno spazio vettoriale pre-hilbertiano. Allora*

$$||x+y|| \le ||x|| + ||y|| \quad \text{per ogni } x, y \in H \ . \tag{4.15}$$

Inoltre vale l'uguaglianza se e solo se $y = 0$, oppure $x = \lambda y$ con $\lambda \ge 0$ reale positivo.

Dimostrazione:

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

Usando la disuguaglianza di Schwarz si ha infine

$$||x+y||^2 \le ||x||^2 + ||y||^2 + 2||x||\,||y|| = (||x|| + ||y||)^2$$

e si conclude. L'uguaglianza vale solo se $\mathrm{Re}\,(x|y) = |(x|y)| = ||x||\,||y||$ come visto sopra, e l'ultima uguaglianza (escluso il caso $y = 0$) si ha se e solo se $x = \lambda y$ per qualche $\lambda \in \mathbb{K}$ (dal teorema di Schwarz); ma allora

$$(x|y) = (\lambda y|y) = \lambda^*(y|y) = \lambda^*||y||^2 \ , \tag{4.16}$$

ed allora $\mathrm{Re}\,(x|y) = |(x|y)|$ se e solo se $\mathrm{Re}\,(\lambda)||y||^2 = |\lambda^*|\,||y||^2$ se e solo se $\lambda \ge 0$. C.V.D.

La norma indotta dal prodotto scalare gode di una interessante proprietà:

Teorema 4.6 (identità del parallelogramma) *Sia H uno spazio vettoriale pre-hilbertiano. Allora*

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2 \quad \text{per ogni } x, y \in H \ . \tag{4.17}$$

Dimostrazione: Verifica diretta. C.V.D.

Quest'identità è così chiamata perché dice che la somma dei quadrati delle diagonali di un parallelogramma è uguale alla somma dei quadrati dei lati (fig. (4.1)c). Tale identità caratterizza gli spazi a prodotto scalare fra gli spazi normati: si può dimostrare che in uno spazio normato, la cui norma soddisfa l'identità del parallelogramma, si può definire un prodotto scalare che dà tale norma. Come si intuisce, la formula che fornisce tale prodotto scalare è proprio l'identità di polarizzazione, sostituendo a secondo membro $g(a,a) \to ||a||^2$.

Teorema 4.7 (continuità del prodotto scalare) *Sia H uno spazio pre-hilbertiano. Il prodotto scalare è una funzione continua rispetto alla topologia indotta dal prodotto scalare stesso. In altre parole, l'applicazione*

$$\begin{aligned} g : H \times H &\to \mathbb{K} \\ (x,\ y) &\mapsto (x|y) \end{aligned}$$

è continua.

Dimostrazione: Per la disuguaglianza di Schwarz abbiamo

$$\begin{aligned}\big|(x|y) - (x_0|y_0)\big| &= \big|(x_0 + x - x_0|y_0 + y - y_0) - (x_0|y_0)\big| = \big|(x_0|y - y_0) + (x - x_0|y_0) + (x - x_0|y - y_0)\big| \\ &\overset{(4.13)}{\le} ||x_0||\,||y-y_0|| + ||x-x_0||\,||y_0|| + ||x-x_0||\,||y-y_0||\end{aligned}$$

che può essere reso piccolo a piacere quando $||x - x_0|| \to 0$ e $||y - y_0|| \to 0$. C.V.D.

Corollario 4.8 *Sia $(x_n)_{n\in\mathbb{N}} \subset H$ una successione di vettori che tende al vettore $x \in H$. Allora*

$$x_n \xrightarrow{n\to\infty} x \qquad \Longrightarrow \qquad \lim_{n\to\infty}(x_n|y) = (x|y) \quad e \ \lim_{n\to\infty}(y|x_n) = (y|x) \quad \text{per ogni } y \in H \ .$$

Uno spazio pre-hilbertiano risulta quindi essere anche uno spazio normato, metrico e topologico, e sorge quindi naturale chiedersi se sia completo. Essendo il prodotto scalare una qualsiasi forma sesquilineare hermitiana e definita positiva, la completezza non è in generale assicurata, per cui è una proprietà aggiuntiva che eventualmente deve essere richiesta. La completezza di uno spazio a prodotto scalare è una proprietà molto importante, a cui si riserva un nome speciale:

Definizione 4.8 *Uno spazio pre-hilbertiano $\mathcal{H}$ completo nella metrica indotta dal prodotto scalare è detto* spazio di Hilbert

Una conseguenza immediata della continuità del prodotto scalare è data dal fatto che uno spazio pre-hilbertiano può essere sempre completato:

Teorema 4.9 *Ogni spazio pre-hilbertiano H ammette come completamento uno spazio di Hilbert $\mathcal{H}$.*

Dimostrazione: La costruzione del completamento ricalca quella del completamento di uno spazio metrico e di uno spazio normato: si considerano come elementi del completamento le classi di equivalenza delle successioni di Cauchy, le quali vengono dotate del prodotto interno definito da

$$([x]|[y]) := \lim_{n\to\infty} (x_n|y_n) ,$$

ove (x_n) e (y_n) sono succesioni di Cauchy rappresentanti di $[x]$ ed $[y]$ rispettivamente. È immediato verificare che i numeri $(x_n|y_n)$ formano una successione di Cauchy in $\mathbb{C}$ che pertanto converge. Poiché il prodotto scalare risulta una funzione continua dei suoi argomenti, la definizione sopra è indipendente dalla scelta delle successioni $(x_n) \in [x]$ e $(y_n) \in [y]$, e inoltre, sempre per continuità , valgono tutte le proprietà del prodotto scalare. Pertanto la definizione è ben posta e la norma e la distanza in $\mathcal{H}$ sono esprimibili tramite il prodotto scalare così definito. C.V.D.

Esempi di spazi (pre-)hilbertiani
(1) $\mathcal{H} = \mathbb{R}^n$ con $(x|y) = \sum_{k=1}^n x_k y_k$.
(2) $\mathcal{H} = \mathbb{C}^n$ con $(x|y) = \sum_{k=1}^n x_k^* y_k$.
(3) Dati $a, b \in \mathbb{R}$, $a < b$, $H = C([a, b], \mathbb{K})$ con prodotto scalare

$$(f|g) = \int_a^b f^*(x) g(x)\, \mathrm{d}x$$

(4) a, b come sopra, $H = C^m([a, b], \mathbb{K})$ con prodotto scalare

$$(f|g) = \sum_{j=0}^{m} \int_a^b f^{(m)*}(x) g^{(m)}(x)\, \mathrm{d}x$$

(5) Lo spazio l^2 delle successioni a quadrato sommabile

$$l^2 := \{(x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{C} , \quad \sum_{n\in\mathbb{N}} |x_n|^2 < \infty\} \tag{4.18}$$

è spazio di Hilbert rispetto al prodotto scalare

$$(x|y) = \sum_{n\in\mathbb{N}} x_n^* y_n .$$

L'esistenza del prodotto scalare (cioè della convergenza della serie qui sopra) è garantita dalla disuguaglianza di Hölder con $p = 2 = q$, che in questo caso coincide con la disuguaglianza di Schwarz.

La norma indotta da questo prodotto scalare è proprio la norma-2 $||\cdot||_2$. Lo spazio l^2 risulta completo in questa norma, ed è quindi uno spazio di Hilbert. Esso è anche separabile. Si può dimostrare che tutti gli spazi di Hilbert separabili sono isometrici ad l^2.

(6) Lo spazio $\mathcal{L}^2$ delle (classi di equivalenza di) funzioni a quadrato sommabile

$$L^2(\mathbb{R}^n) := \{f : \mathbb{R}^n \in \mathbb{C} , \quad \int_{\mathbb{R}^n} |f(x)|^2\, \mathrm{d}x < \infty\} \tag{4.19}$$

$\mathcal{L}^2$ è spazio di Hilbert rispetto al prodotto scalare

$$(f|g) = \int_{\mathbb{R}^n} f^*(x) g(x)\, \mathrm{d}x .$$

L'esistenza del prodotto scalare (cioè della sommabilità dell'integrale qui sopra) è garantita dalla disuguaglianza di Hölder con $p = 2 = q$, che anche in questo caso coincide con la disuguaglianza di Schwarz.

La norma indotta da questo prodotto scalare è proprio la norma-2 $||\cdot||_2$. Lo spazio $\mathcal{L}^2$ risulta completo in questa norma, ed è quindi uno spazio di Hilbert. Anche $\mathcal{L}^2$ è separabile.

4.2 Ortogonalità

Nel seguito di questo capitolo studieremo spazi a prodotto scalare, quindi sarà sottointeso che consideriamo spazi pre-hilbertiani.

La disuguaglianza di Schwarz permette di introdurre delle nozioni di carattere geometrico in uno spazio pre-hilbertiano. Consideriamo due vettori non nulli x e y, allora i "versori" $\hat{x} = x/||x||$ ed $\hat{y} = y/||y||$ individuano due direzioni in tale spazio, e, nel caso di uno spazio con prodotto scalare reale, la disuguaglianza di Schwarz implica che

$$-1 \leq \frac{(x|y)}{||x||\,||y||} \leq 1 \tag{4.20}$$

con le uguaglianze verificate quando i due vettori sono collineari, cioè proporzionali l'uno all'altro. In analogia con lo spazio euclideo $\mathbb{R}^3$, questo permette di definire l'angolo θ compreso tra i due vettori:

$$\cos(\theta) := \frac{(x|y)}{||x||\,||y||} \qquad \theta \in [0, \pi]\ . \tag{4.21}$$

Con questa definizione, abbiamo proprio che i due vettori x ed y sono paralleli quando $\theta = 0$, antiparalleli quando $\theta = \pi$. Estendendo l'analogia, possiamo stabilire che due vettori sono ortogonali quando $\theta = \pi/2$, cioè $\cos\theta = 0$. La proprietà (4.20) non vale per un prodotto scalare hermitiano (che assume valori complessi), ma possiamo comunque definire anche in questo caso il concetto di ortogonalità:

Definizione 4.9 *Due vettori di uno spazio pre-hilbertiano sono detti* ortogonali *se è nullo il loro prodotto scalare:*

$$x \perp y \qquad \iff \qquad (x|y) = 0\ . \tag{4.22}$$

OSSERVAZIONE: L'ortogonalità è una relazione simmetrica, grazie alla simmetria o hermiticità del prodotto scalare: $x \perp y \iff y \perp x$.

Con il concetto di ortogonalità ritroviamo un fondamento della geometria euclidea, il *teorema di Pitagora*. Infatti, se x ed y sono ortogonali tra loro, è immediato verificare che

$$(x|y) = 0 \qquad \Longrightarrow \qquad ||x+y||^2 = ||x-y||^2 = ||x||^2 + ||y||^2\ , \tag{4.23}$$

dove interpretiamo geometricamente la norma $||x-y||$ o $||x+y||$ come l'ipotenusa del triangolo rettangolo di cateti $||x||$ e $||y||$.

Definizione 4.10 *Una famiglia* $\{e_\lambda : \lambda \in \Lambda\}$ *di vettori di uno spazio pre-hilbertiano è detta* sistema ortogonale *se i vettori della famiglia sono a due a due ortogonali:*

$$\lambda, \mu \in \Lambda\ , \quad \lambda \neq \mu \quad \Longrightarrow \quad (e_\lambda|e_\mu) = 0\ .$$

La famiglia è detta sistema ortonormale *se è ortogonale ed ogni elemento della famiglia ha norma unitaria:* $||e_\lambda|| = 1$ *per ogni* $\lambda \in \Lambda$*, ovvero*

$$(e_\lambda|e_\mu) = \delta_{\lambda\mu} = \begin{cases} 1 & (\lambda = \mu) \\ 0 & (\lambda \neq \mu)\ . \end{cases}$$

Teorema 4.10 *Ogni famiglia ortogonale che non contenga il vettore nullo è linearmente indipendente.*

Dimostrazione: Sia $\sum_{\lambda \in F} \alpha_\lambda e_\lambda = 0$ con $\alpha_\lambda \in \mathbb{K}$, F sottoinsieme finito di Λ. Fissato $\mu \in F$, si moltiplichi scalarmente (a sinistra) l'uguaglianza precedente per e_μ. Si ottiene

$$0 = (e_\mu| \sum_{\lambda \in F} \alpha_\lambda e_\lambda) = \sum_{\lambda \in F} \alpha_\lambda (e_\mu|e_\lambda) = \alpha_\mu (e_\mu|e_\mu)$$

perché solo il termine con $\lambda = \mu$ non è nullo. Ma $(e_\mu|e_\mu) = ||e_\mu||^2 > 0$ poiché per ipotesi $e_\mu \neq 0$. Quindi $\alpha_\mu = 0$, e questo per qualsiasi $\mu \in F$. C.V.D.

Definizione 4.11 *Sia $S \subset H$ sottoinsieme non vuoto. Il* complemento ortogonale *di S è l'insieme dei vettori di S ortogonali a tutti i vettori di S:*

$$S^\perp := \{x \in H : (s|x) = 0 \text{ per ogni } s \in S\} \,. \tag{4.24}$$

Si verifica subito che $S \cap S^\perp \subset \{0\}$, poiché se $x \in S$ e $x \in S^\perp$ allora x deve essere ortogonale a se stesso, e questo vale solo per il vettore nullo. Anche se l'insieme S non ha strutture particolari, il suo ortogonale ce le ha:

Proposizione 4.11
(i) Sia $S \subset H$ sottoinsieme non vuoto. Allora $S^\perp$ è un sottospazio vettoriale chiuso: $S^\perp = \overline{S^\perp} < H$.
(ii) Se $S \subset T \subset H$ allora $S^\perp \supset T^\perp$.
(iii) Se $\langle S\rangle$ è il sottospazio generato da $S \subset H$, allora $\langle S\rangle^\perp = S^\perp$.
(iv) $S^\perp = \overline{S}^{\,\perp}$.
(v) $S \subset (S^\perp)^\perp$.

Dimostrazione: (i) Per la linearità del prodotto scalare, se $y_1, y_2 \in S^\perp$, cioè sono entrambi ortogonali ad ogni $s \in S$, allora anche $\alpha_1 y_1 + \alpha_2 y_2$ è ortogonale ad ogni s e quindi appartiene ad $S^\perp$, quindi $S^\perp < H$. Se $x \in \overline{S^\perp}$ allora esiste una successione di vettori $(x_n)_{n\in\mathbb{N}}$ di $S^\perp$ che converge ad x. Quindi, per la continuità del prodotto scalare, abbiamo

$$(s|x) = (s|\lim_{n\to\infty} x_n) = \lim_{n\to\infty}(s|x_n) = 0 \quad \text{ per ogni } s \in S$$

e quindi $x \in S^\perp$, ossia $S^\perp$ è chiuso.

(ii) ovvia: se $x \in T^\perp$ allora x è ortogonale a tutti i vettori di T, quindi a tutti i vettori di $S \subset T$ quindi sta in $S^\perp$.

(iii) Siccome $S \subset \langle S\rangle$ allora per (ii) $\langle S\rangle^\perp \subset S^\perp$. Dimostriamo ora che $S^\perp \subset \langle S\rangle^\perp$. Sia quindi $x \in S^\perp$, cioè ortogonale ad ogni $s \in S$. Allora, per la linearità del prodotto scalare, x è ortogonale ad ogni combinazione lineare finita di vettori di s, e quindi sta in $\langle S\rangle^\perp$.

(iv) Siccome $S \subset \overline{S}$ allora per (ii) $\overline{S}^{\,\perp} \subset S^\perp$. Sia poi $y \in S^\perp$. Vogliamo dimostrare che y è ortogonale ad ogni elemento $x \in \overline{S}$. Consideriamo una successione $(x_n)_{n\in\mathbb{N}} \subset S$, $x_n \xrightarrow{n\to\infty} x$. Allora, per la continuità del prodotto scalare,

$$(x|y) = (\lim_{n\to\infty} x_n|y) = \lim_{n\to\infty}(x_n|y) = 0\,.$$

(v) Sia $s \in S$. Allora s è ortogonale a tutti gli elementi di $S^\perp$, quindi $s \in (S^\perp)^\perp$. C.V.D.

Pensando al punto (v) del precedente teorema, ci chiediamo chi possa essere $(S^\perp)^\perp$, l'ortogonale dell'ortogonale dell'insieme S. Senz'altro contiene S, e siccome è uno spazio vettoriale, dovrà anche contenere $\langle S\rangle$, lo spazio generato da S. Inoltre deve essere chiuso e questo basta:

$$S^{\perp\perp} = \overline{\langle S\rangle}\,. \tag{4.25}$$

Teorema 4.12 (del sottospazio) *Sia $M < \mathcal{H}$ un sottospazio di uno spazio di Hilbert $\mathcal{H}$. Allora*
(i) M è completo se e solo se è chiuso;
(ii) se M è finito dimensionale, allora M è completo.

Dimostrazione: (i) Sia M chiuso. Sia $(x_n)_{n\in\mathbb{N}}$ una successione di Cauchy in M. Poiché $\mathcal{H}$ è completo, esiste $x \in \mathcal{H}$ tale che $x_n \to x$. Poiché M è chiuso $x_n \to x$ implica $x \in M$. Quindi ogni successione di Cauchy in M è convergente in M.

Viceversa, sia M completo. Per ogni $x \in \overline{M}$ c'è una successione $x_n \to x$ con $x_n \in M$ per ogni $n \in \mathbb{N}$. Poiché (x_n) è una successione di Cauchy in M completo ne segue che $x \in M$.

(ii) Ogni sottospazio finito-dimensionale di dimensione n è isomorfo ed omeomorfo a $\mathbb{K}^n$ che è completo. C.V.D.

In particolare, tutti gli spazi pre-hilbertiani di dimensione finita sono completi.

Una proprietà utile negli spazi vettoriali con prodotto scalare a dimensione finita è costituita dal fatto che se M è un sottospazio di uno spazio vettoriale H, allora si può considerare il sottospazio complemento ortogonale $M^\perp$ e tutto lo spazio H può decomporsi nella somma diretta di M e $M^\perp$. Ci proponiamo ora di generalizzare tale risultato al caso di uno spazio a dimensione infinita. La differenza cruciale tra i due casi è costituita dal fatto che in uno spazio a dimensione finita ogni sottospazio risulta anche chiuso, mentre questo non è più garantito a dimensioni infinite, anche se ci troviamo in uno spazio di Hilbert, cioè completo. La proprietà di chiusura è importante e la decomposizione è garantita quando sappiamo che il sottospazio è chiuso.

Teorema 4.13 *Sia $\mathcal{H}$ uno spazio di Hilbert ed M un suo sottospazio chiuso. Allora per ogni $x \in \mathcal{H}$ esiste un unico elemento $y_M \in M$ di minima distanza da x, cioè tale che*

$$d(x, M) := \inf_{y \in M} ||x - y|| = ||x - y_M|| \, . \tag{4.26}$$

Dimostrazione: Se $x \in M$, allora banalmente $y_M = x$ e $d(x, M) = 0$. Altri vettori di M sono a distanza non nulla da x. In generale, l'insieme dei numeri $||x-y||$ con $y \in M$ è limitato inferiormente da 0, quindi la distanza $d = d(x, M)$ è ben definita. Per definizione di inf, esiste una successione $(y_n)_{n\in\mathbb{N}}$ in M tale che $||x-y_n|| \xrightarrow{n\to\infty} d$. Allora

$$\begin{aligned}
||y_n - y_m||^2 &= ||y_n - x - (y_m - x)||^2 \\
&\overset{(4.17)}{=} 2||y_n - x||^2 + 2||y_m - x||^2 - || - 2x + y_n + y_m||^2 \\
&= 2||y_n - x||^2 + 2||y_m - x||^2 - 4||x - \frac{y_n + y_m}{2}||^2 \\
&\leq 2||y_n - x||^2 + 2||y_m - x||^2 - 4d^2 \\
&\xrightarrow{m,n\to\infty} 2d^2 + 2d^2 - 4d^2 = 0
\end{aligned}$$

dove abbiamo utilizzato l'identità del parallelogramma (4.17) ed il fatto che $(y_n + y_m)/2 \in M$ e quindi che $||x - (y_n + y_m)/2|| \geq d$. Pertanto (y_n) è una successione di Cauchy e poiché M è chiuso (e quindi completo) (y_n) converge ad un elemento $y_M \in M$, il quale realizza la minima distanza:

$$y_M = \lim_{n\to\infty} y_n \in M \, , \qquad ||x - y_M|| = \lim_{n\to\infty} ||x - y_n|| = d(x, M) \, .$$

Mostriamo ora che $z := x - y_M$ è ortogonale ad M. Sia y un elemento arbitrario di M, e $\alpha \in \mathbb{K}$. Allora $y_M + \alpha y \in M$ e si ha

$$\begin{aligned}
&d^2 \leq ||x - (y_M + \alpha y)||^2 = ||z - \alpha y||^2 = d^2 - 2\mathrm{Re}\left(\alpha(z|y)\right) + |\alpha|^2||y||^2 \\
&\Longrightarrow \qquad |\alpha|^2||y||^2 \geq 2\mathrm{Re}\left(\alpha(z|y)\right) \, .
\end{aligned}$$

Per ogni $\alpha \in \mathbb{R}$ si ha

$$\alpha^2||y||^2 \geq 2\alpha\mathrm{Re}\,(z|y) \qquad \Longrightarrow \qquad \mathrm{Re}\,(z|y) = 0 \, .$$

Per ogni $\alpha = -\mathrm{i}\beta \in \mathbb{I}$ immaginario si ha

$$\beta^2||y||^2 \geq 2\beta\mathrm{Im}\,(z|y) \qquad \Longrightarrow \qquad \mathrm{Im}\,(z|y) = 0$$

e quindi $(z|y) = 0$ per ogni $y \in M$, quindi $z \in M^\perp$.

Unicità : da quando detto sopra, abbiamo decomposto $x = y_M + z$ con $y_M \in M$ e $z \in M^\perp$. Ma $M \cap M^\perp = \{0\}$, quindi M ed $M^\perp$ sono in somma diretta, pertanto la decomposizione di qualsiasi $x \in M$ (che qui abbiamo dimostrato esistere) è unica. C.V.D.

Abbiamo così dimostrato anche l'importante

Teorema 4.14 *Sia $\mathcal{H}$ uno spazio di Hilbert e M un suo sottospazio chiuso. Allora vale*

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

$$x = x_\parallel + x_\perp$$

cioè ogni vettore $x \in H$ è decomposto in maniera univoca come somma (ortogonale) di un vettore $x_\parallel \in M$ e un vettore $x_\perp \in M^\perp$.

Dimostrazione: Fatta prima. C.V.D.

La componente $x_\parallel$ è detta *proiezione ortogonale* di x sul sottospazio M (e di conseguenza $x_\perp$ è la proiezione ortogonale sul complemento ortogonale $M^\perp$) e definisce un operatore lineare

$$P_M : \mathcal{H} \to \mathcal{H}\ , \qquad P_M(x) = x_\parallel \tag{4.27}$$

detto *proiettore ortogonale* su M.

Corollario 4.15 *Sia $\mathcal{H}$ spazio di Hilbert ed $M < \mathcal{H}$. Allora M è chiuso se e solo se $(M^\perp)^\perp = M$.*

4.3 Sistemi ortonormali

In questa sezione ci proponiamo di estendere il concetto di base introdotto precedentemente. Ricordiamo che un insieme di vettori è una base se e solo se ogni vettore dello spazio vettoriale si può esprimere univocamente come combinazione lineare *finita* dei vettori di base. Se lo spazio vettoriale è normato, quindi se ha una topologia, possiamo considerare serie infinite di vettori, e pertanto ci possiamo chiedere se ci siano delle famiglie di vettori indipendenti tali che ogni vettore dello spazio si possa esprimere univocamente come combinazione lineare *infinita* della famiglia indipendente.

La possibilità di ammettere C.L. infinite permette di ridurre notevolmente la cardinalità della "base", e quindi permette di semplificare la decomposizione dei vettori e dello spazio vettoriale. Questo schema di lavoro è particolarmente utile nel contesto degli spazi di Hilbert, e grazie al prodotto scalare sarà conveniente considerare sistemi ortonormali. Un sistema può essere finito o infinito. Il caso interessante si ha quando i vettori sono infiniti. Vediamo una prima conseguenza della definizione per un sistema numerabile.

Lemma 4.16 *Sia $(e_n)_{n\in\mathbb{N}}$ un sistema ortonormale numerabile in uno spazio di Hilbert $\mathcal{H}$. Allora la serie vettoriale*

$$\sum_{n\in\mathbb{N}} c_n e_n$$

è convergente se e solo se

$$\sum_{n\in\mathbb{N}} |c_n|^2 < \infty\ .$$

In questo caso, la somma della serie vettoriale

$$x_c := \sum_{n\in\mathbb{N}} c_n e_n \in \mathcal{H}\ ,$$

*non dipende dall'ordine dei termini, inoltre $c_n = (e_n|x_c)$ e vale l'*identità di Parseval

$$||x_c||^2 = \sum_{n\in\mathbb{N}} |c_n|^2\ . \tag{4.28}$$

Dimostrazione: Consideriamo le successioni delle ridotte

$$S_N := \sum_{k=0}^{N} c_k e_k\ , \qquad \sigma_N := \sum_{k=0}^{N} |c_k|^2$$

e consideriamo il criterio di Cauchy. Dati M, N con $M < N$, si ha

$$\begin{aligned} ||S_N - S_M||^2 &= \left|\left|\sum_{k=M+1}^{N} c_k e_k\right|\right|^2 = \left(\sum_{j=M+1}^{N} c_j e_j \middle| \sum_{k=M+1}^{N} c_k e_k\right) \\ &= \sum_{j,k=M+1}^{N} c_j^* c_k (e_j|e_k) = \sum_{k=M+1}^{N} c_k^* c_k = \sum_{k=M+1}^{N} |c_k|^2 = \sigma_N - \sigma_M \end{aligned}$$

per cui la convergenza di una serie è garantita dalla convergenza dell'altra serie (ricordiamo che sia lo spazio di Hilbert $\mathcal{H}$, sia $\mathbb{C}$ sono completi, per cui vale il criterio di Cauchy).

In maniera analoga, ripetendo il calcolo appena fatto con $M = 0$, e poi prendendo il limite per $N \to \infty$,

$$||x_c||^2 = \left|\left|\sum_{k=0}^{\infty} c_k e_k\right|\right|^2 = \lim_{N\to\infty} ||S_N||^2 = \lim_{N\to\infty} \sigma_N = \sum_{k=0}^{\infty} |c_k|^2 \ .$$

Definita la somma della serie vettoriale x_c, possiamo ricavarne il coefficiente n-esimo facendo il prodotto scalare di x_c (a destra) con e_n (a sinistra). Infatti, per ogni $N \geq n$ si ha

$$(e_n|\sum_{k=0}^{N} c_k e_k) = \sum_{k=0}^{N} c_k(e_n|e_k) = \sum_{k=0}^{N} c_k \delta_{kn} = c_n$$

e si dimostra l'affermazione facendo il limite per $N \to \infty$ e sfruttando la continuità del prodotto scalare.

Mostriamo ora che la somma della serie vettoriale non dipende dall'ordine dei termini. Siano

$$x = \sum_{n=0}^{\infty} c_k e_k \ , \qquad y = \sum_{k=0}^{\infty} c_{p(k)} e_{p(k)} \ ,$$

con $p : \mathbb{N} \to \mathbb{N}$ una biiezione in $\mathbb{N}$, cioè una permutazione degli indici. Notiamo che insieme $\{e_{p(k)} : k \in \mathbb{N}\}$ è esattamente il sistema ortonormale di partenza, in quanto in un insieme non si distingue l'ordine degli elementi. Inoltre le serie numeriche assolutamente convergenti non dipendono dall'ordine dei termini. Quindi

$$||x||^2 \overset{(4.28)}{=} \sum_{k=0}^{\infty} |c_k|^2 = \sum_{k=0}^{\infty} |c_{p(k)}|^2 \overset{(4.28)}{=} ||y||^2$$

Consideriamo ora

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

I termini $||x||^2$ e $||y||^2$ li abbiamo appena calcolati. Calcoliamo il prodotto $(x|y)$. Sfruttando la continuità del prodotto scalare si ha

$$(x|y) = \lim_{N\to\infty} \sum_{k=0}^{N} c_k^*(e_k|y) \ .$$

Ogni indice k risulta la permutazione di un unico indice j_k tale che $k = p(j_k)$, per cui

$$(e_k|y) = \lim_{M\to\infty} \sum_{j=0}^{M} c_{p(j)}(e_k|e_{p(j)}) = c_{p(j_k)} = c_k \ .$$

che conferma quanto già dimostrato sopra nel ricavare c_n tramite prodotto scalare con e_n. Pertanto

$$(x|y) = \sum_{k=0}^{\infty} c_k^* c_k = \sum_{k=0}^{\infty} |c_k|^2 = ||x||^2 \ .$$

Analogamente

$$(y|x) = \sum_{k=0}^{\infty} |c_k|^2 = ||x||^2$$

e quindi

$$||x - y||^2 = 0 \qquad \Longrightarrow \qquad x = y \ .$$

C.V.D.

Nel corso della dimostrazione abbiamo visto che, con un sistema ortonormale numerabile $(e_n)_{n\in\mathbb{N}}$, posto

$$x = \sum_{n=0}^{\infty} c_n e_n \qquad \text{si ha} \qquad c_n = (e_n|x) \ .$$

Ci chiediamo ora cosa possiamo dire con un sistema ortonormale generale $\{e_\lambda : \lambda \in \Lambda\}$ con Λ non necessariamente numerabile. L'equazione precedente può essere generalizzata e vale

Teorema 4.17 (disuguaglianza di Bessel) *Sia $\{e_\lambda : \lambda \in \Lambda\}$ un sistema ortonormale (non necessariamente numerabile) in uno spazio di Hilbert $\mathcal{H}$ e sia $x \in \mathcal{H}$. Allora i numeri $(e_n|x)$ sono non nulli al più per una infinità numerabile di indici $\lambda \in \Lambda_x$ e vale la* disuguaglianza di Bessel

$$\sum_{\lambda \in \Lambda_x} |(e_\lambda|x)|^2 \leq ||x||^2 . \tag{4.29}$$

dove la somma è estesa all'infinità numerabile di valori dell'indice $\lambda \in \Lambda_x$. Inoltre la serie (numerabile)

$$\sum_{\lambda \in \Lambda_x} (e_\lambda|x) e_\lambda$$

risulta convergente, indipendente dall'ordine dei termini non nulli, e converge alla proiezione ortogonale di x sulla chiusura del sottospazio generato dal sistema $(e_\lambda)_{\lambda \in \Lambda}$.

Vediamo ora come possiamo estendere il concetto di base tramite i sistemi ortonormali.

Definizione 4.12 *Un* sistema ortonormale *in uno spazio di Hilbert $\mathcal{H}$ è detto* completo *se è massimale, cioè se non può essere esteso ulteriormente con l'aggiunta di ulteriori vettori.*[1]

L'utilità dei sistemi ortonormali completi è espressa dal seguente teorema, che fornisce sostanzialmente delle definizioni equivalenti di sistema ortonormale completo, e trasforma la disuguaglianza di Bessel in una uguaglianza (di Parseval)

Teorema 4.18 *Sia $(e_\lambda)_{\lambda \in \Lambda}$ un sistema ortonormale in uno spazio di Hilbert $\mathcal{H}$. Allora le seguenti affermazioni sono equivalenti.*

1. *$(e_\lambda)_{\lambda \in \Lambda}$ è completo.*
2. *Il sistema ortonormale genera tutto $\mathcal{H}$, cioè $\overline{\langle (e_\lambda)_{\lambda \in \Lambda} \rangle} = \mathcal{H}$.*
3. *Ogni vettore $x \in \mathcal{H}$ può essere decomposto secondo la relazione*
$$x = \sum_{\lambda \in \Lambda_x} (e_\lambda|x) e_\lambda \tag{4.30}$$
dove la somma è estesa ai valori di $\lambda \in \Lambda_x$ (al più numerabili) per cui $(e_\lambda|x) \neq 0$.
4. *Per ogni $x \in \mathcal{H}$ vale la* relazione di Parseval
$$||x||^2 = \sum_{\lambda \in \Lambda_x} |(e_\lambda|x)|^2 \tag{4.31}$$
5. *Per ogni $x, y \in \mathcal{H}$ vale la* relazione di completezza
$$\sum_{\lambda \in \Lambda_x} (x|e_\lambda)(e_\lambda|y) = (x|y) \tag{4.32}$$

Dimostrazione: Dimostriamo soltanto alcune implicazioni. Denotiamo la chiusura dello spazio generato dal sistema ortonormale con $M := \overline{\langle (e_\lambda)_{\lambda \in \Lambda} \rangle}$.

(1)⇒(2): Se il sistema ortonormale è completo, non possiamo aggiungere ad esso alcun vettore ortogonale e quindi, decomponendo $\mathcal{H} = M \oplus M^\perp$, il complemento ortogonale deve essere nullo $M^\perp = \{0\}$ (in caso contrario potremmo allargare il sistema ortonormale con un vettore non nullo di $M^\perp$); quindi $M = \mathcal{H}$.

(2)⇒(1): Sia $M = \mathcal{H}$. Se il sistema non fosse completo (cioè massimale) potremmo trovare un nuovo vettore ortonormale agli altri, cioè un vettore non nullo ortogonale a $M = \mathcal{H}$, il che è assurdo.

[1] Alcuni autori usano anche i termini *base hilbertiana* oppure *base ortonormale* per designare un sistema ortonormale completo. Si faccia poi attenzione a non confondere la completezza di uno spazio metrico con la completezza di un insieme di vettori: sono due concetti totalmente indipendenti.

(3)⇒(4). Dalla continuità del prodotto scalare

$$||x||^2 = (x|x) = \lim_{N\to\infty} \sum_{k=0}^{N} (x|e_{\lambda_k})(e_{\lambda_k}|x) = \sum_{\lambda\in\Lambda} |(e_\lambda|x)|^2 \ .$$

(3)⇒(5). Se

$$x = \sum_{\lambda} (e_\lambda|x) e_\lambda \ , \qquad y = \sum_{\lambda} (e_\lambda|y) e_\lambda \ ,$$

allora, per la continuità del prodotto scalare

$$(x|y) = \lim_{N\to\infty} \sum_{k=0}^{N} (x|e_{\lambda_k})(e_{\lambda_k}|x) = \sum_{\lambda\in\Lambda} |(e_\lambda|x)|^2 \ .$$

(5)⇒(4): Basta considerare il caso $y = x$. C.V.D.

Ci potremmo porre ora il problema dell'esistenza di sistemi ortonormali completi. Mediante le tecniche con le quali si può dimostrare che ogni spazio vettoriale ammette una base, si può anche dimostrare che ogni spazio di Hilbert ammette un sistema ortonormale completo. Ovviamente, se lo spazio di Hilbert ha dimensioni finite, un sistema ortonormale completo risulta anche una base ordinaria.

Nel caso di spazi vettoriali a dimensioni finite avevamo anche l'importante risultato che due basi diverse hanno la stessa cardinalità . Questo risultato lo si può estendere anche ai sistemi ortonormali completi di uno spazio di Hilbert, anche a dimensione infinita, stabilendo una corrispondenza biunivoca tra i due sistemi la quale permette di affermare che due sistemi ortonormali completi hanno la stessa cardinalità .

In molte situazioni si ha a disposizione un insieme numerabile di vettori, cioè una successione di vettori, che generano un sottospazio di uno spazio di Hilbert, e può sorgere la necessità di costruire un sistema ortonormale che generi il medesimo sottospazio. Esiste a tale scopo un procedimento standard.

Teorema 4.19 (metodo di Gram-Schmidt) *Sia data una successione di vettori $(x_n)_{n\in\mathbb{N}}$ (quindi un insieme numerabile) in uno spazio di Hilbert $\mathcal{H}$. Allora esiste un sistema ortonormale numerabile $(e_k)_{k\in\mathbb{N}}$ tale che: (1) Ogni vettore x_n è combinazione lineare dei vettori $e_1, e_2, \cdots, e_n$, con $n \in \mathbb{N}$; (2) il sottospazio generato dai vettori (x_n) coincide col sottospazio generato dai vettori (e_k):*

Dimostrazione: Nel caso in cui la successione di vettori $(x_n)_{n\in\mathbb{N}}$ sia un insieme linearmente dipendente, eliminiamo da essa tutti i vettori nulli, e poi per ogni $n \in \mathbb{N}$ eliminiamo il vettore x_n se questo risulta linearmente dipendente dai vettori $\{x_j : j < n\}$. Rimettendo in successione i vettori rimasti, la successione di vettori $(x_n)_{n\in\mathbb{N}}$ è ora un insieme linearmente indipendente.

Definiamo quindi i vettori $e_k : k \in \mathbb{N}$ secondo il seguente algoritmo ricorsivo:

$$\begin{aligned}
u_0 &:= x_0 & e_0 &:= \frac{u_0}{||u_0||} \\
u_1 &:= x_1 - (e_0|x_1)e_0 & e_1 &:= \frac{u_1}{||u_1||} \\
&\vdots & &\vdots \\
u_k &:= x_k - \sum_{n=0}^{k-1} (e_n|x_k)e_n & e_k &:= \frac{u_k}{||u_k||} \\
&\vdots & &\vdots
\end{aligned}$$

In pratica, il primo vettore x_0 viene semplicemente normalizzato, mentre ai successivi x_k viene sottratta la loro proiezione ortogonale sul sottospazio generato dai vettori precedenti, in modo da restare con un vettore u_k (non nullo!) ortogonale ai precedenti il quale, normalizzato, individua e_k. In questo modo $(e_k)_{k\in\mathbb{N}}$ forma un sistema ortonormale.

Osserviamo che, per ogni $k \in \mathbb{N}$, lo spazio generato dagli $x_n : n \le k$ coincide con lo spazio generato dagli $e_n : n \le k$ (se non è evidente, dimostrarlo per induzione). Quindi u_k non è mai nullo, in quanto x_k è indipendente dagli $x_n : n < k$, quindi $||u_k|| \neq 0$ ed ha senso definire e_k. Poiché $\langle (x_n)_{n\leq k}\rangle = \langle (e_n)_{n\leq k}\rangle$, nel limite $k \to \infty$ si ricava che $\langle (x_n)_{n\in\mathbb{N}}\rangle = \langle (e_n)_{n\in\mathbb{N}}\rangle$. C.V.D.

Teorema 4.20 *Sia $\mathcal{H}$ uno spazio di Hilbert. Allora $\mathcal{H}$ è separabile se e solo se $\mathcal{H}$ ha un sistema ortonormale completo numerabile.*

Dimostrazione:

Supponiamo che $\mathcal{H}$ ammetta un sistema ortonormale completo numerabile $(e_n)_{n\in\mathbb{N}}$. Allora come insieme numerabile e denso in $\mathbb{N}$ possiamo prendere il sottoinsieme $D \subset \mathcal{H}$ formato dalle combinazioni lineari *finite* a *coefficienti razionali* (intendendo, nel caso di numeri complessi, che sia la parte reale che la parte immaginaria sono razionali). Infatti, le combinazioni lineari finite formano ovviamente un insieme denso nello spazio di Hilbert, ed ognuna di esse può essere approssimata con una corrispondente combinazione lineare a coefficienti razionali. Poiché un unione numerabile di insiemi numerabili è numerabile, otteniamo che D è numerabile e denso in $\mathcal{H}$.

Viceversa, supponiamo che $\mathcal{H}$ sia separabile, per cui esiste un suo sottoinsieme denso e numerabile $D = (x_n)_{n\in\mathbb{N}}$. Allora, mediante il metodo di Gram-Schmidt, con i vettori x_n possiamo costruire un sistema ortonormale $(e_n)_{n\in\mathbb{N}}$ che genera il medesimo sottospazio:

$$\langle (x_n)_{n\in\mathbb{N}}\rangle = \langle (e_n)_{n\in\mathbb{N}}\rangle \ .$$

D'altra parte

$$\mathcal{H} = \overline{(x_n)_{n\in\mathbb{N}}} \subset \overline{\langle (x_n)_{n\in\mathbb{N}}\rangle} = \overline{\langle (e_n)_{n\in\mathbb{N}}\rangle} \ ,$$

per cui i vettori (e_n) formano un sistema ortonormale completo. C.V.D.

La condizione di separabilità era all'inizio inclusa nella definizione stessa di spazio di Hilbert, ma con lo sviluppo della teoria ci si rese conto che molti risultati non dipendevano da tale condizione e la richiesta venne eliminata dalle assunzioni di base.

In fisica invece, ed in particolare in meccanica quantistica, si utilizza pesantemente tale condizione, per cui generalmente viene richiesta, anche se a volte è sottointesa. Ad esempio, si assume che lo spazio degli stati fisici quantistici costituisca uno spazio di Hilbert separabile.

SONC in l^2

Per ogni $n \in \mathbb{N}$ denotiamo con $e^{(n)} = (e_k^{(n)})_{k\in\mathbb{N}}$ la successione che presenta il numero 1 in posizione n-esima (partendo da $k = 0$) e 0 altrove:

$$\begin{aligned} e^{(0)} &= (1, 0, 0, 0, \cdots) \\ e^{(1)} &= (0, 1, 0, 0, \cdots) \\ e^{(2)} &= (0, 0, 1, 0, \cdots) \\ &\cdots \\ e^{(n)} &= (0, \cdots, 0, 1, 0, \cdots) \\ &\cdots \end{aligned}$$

Alternativamente, $e_k^{(n)} = \delta_{nk}$. Ognuna di queste successioni sta in $l^p : 0 \leq p \leq \infty$. Formalmente, per ogni successione $x = (x_k)_{k\in\mathbb{N}}$ potremmo scrivere $(x_k) = \sum_{n\in\mathbb{N}} x_k e^{(k)}$, ma trattandosi di una somma infinita, essa assume significato solo se si precisa in quale topologia va intesa. Ad esempio, se non richiediamo la limitatezza dei numeri x_k, tale espressione non ha evidentemente senso. Ma anche lavorando in l^∞ dotato della sup-norma, tale espressione non ha senso. Negli spazi l^p, con $1 \leq p < \infty$, abbiamo invece che

$$\lim_{N\to\infty} \sum_{n=0}^{N} x_k e^{(k)} \to x \tag{4.33}$$

nella norma-p. Tuttavia l'insieme $E = \{e^{(n)} : n \in \mathbb{N}\} \subset l^p$ non costituisce una base di l^p, in quanto ogni successione $x \in l^p$ con un numero infinito di coefficienti x_k non nulli non si può esprimere come C.L. finita di elementi di E.

Tuttavia, in l^2 l'insieme E costituisce un sistema ortonormale completo. Che sia un insieme ortonormale è evidente. Che sia completo lo si può dedurre dal punto (3) del teorema 4.18. Infatti

$$\begin{aligned} (e^{(n)}|x) &= \sum_{k\in\mathbb{N}} e_k^{(n)*} x_k = \sum_{k\in\mathbb{N}} \delta_{nk} x_k = x_n \\ \Longrightarrow \quad \sum_{k\in\mathbb{N}} (e^{(n)}|x) e^{(n)} &= \sum_{k\in\mathbb{N}} x_k e^{(k)} = x \,. \end{aligned}$$

Alternativamente, possiamo dimostrare che E è massimale: se esistesse $y \in l^2$ non nullo ed ortogonale a tutti gli elementi di E avremmo

$$0 = (e^{(n)}|y) = y_n \quad \text{per ogni} \quad n \in \mathbb{N} \qquad \Longrightarrow \qquad y = 0$$

il che contraddice l'ipotesi.

SONC in $L^2_{2\pi}(\mathbb{R},\mathbb{C})$

Nello spazio $L^2_{2\pi}(\mathbb{R},\mathbb{C})$ delle funzioni da $\mathbb{R}$ in $\mathbb{C}$ periodiche di periodo $\lambda = 2\pi$ $(k = 1)$ a quadrato integrabile possiamo introdurre il prodotto scalare

$$(f|g) := \int_{-\pi}^{\pi} f^*(x) g(x) \,\mathrm{d}x \tag{4.34}$$

ed abbiamo già visto, nella parte sulle serie di Fourier, che le funzioni trigonometriche $x \mapsto \mathrm{e}^{\mathrm{i}nx} : n \in \mathbb{Z}$, dette anche *onde piane*, formano un sistema ortogonale, che si può rendere ortonormale moltiplicando tali funzioni per il fattore $1/\sqrt{2\pi}$ (oppure ridefinendo il prodotto scalare aggiungendo un fattore $1/(2\pi)$).

Vogliamo dimostrare che l'insieme di queste funzioni

$$E = \{e^{(n)} : \mathbb{R} \to \mathbb{C}\,,\ n \in \mathbb{Z}\}\,, \qquad e^{(n)}(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}nx}$$

forma un sistema ortonormale completo in $L^2_{2\pi}(\mathbb{R},\mathbb{C})$.

L'ortonormalità, come già detto, è una semplice verifica. La completezza si può dimostrare verificando che una qualsiasi funzione $f \in L^2_{2\pi}$ ortogonale a tutti gli elementi di E deve essere quasi ovunque nulla, e quindi appartenere alla classe di equivalenza della funzione nulla. Assumiamo dunque che $f \perp E$, quindi che $f \perp e^{(n)}$ per ogni $n \in \mathbb{Z}$. Questo significa che tutti i coefficienti di Fourier di f sono nulli:

$$f_n := (e^{(n)}|f) = \int_{-\pi}^{\pi} \frac{\mathrm{e}^{-\mathrm{i}nx}}{\sqrt{2\pi}} f(x) \,\mathrm{d}x = 0$$

($f_n = \sqrt{2\pi} c_n$ con i c_n coefficienti di Fourier definiti nell'eq. (2.39)). Consideriamo ora la primitiva F di f

$$F(x) = \int_{-\pi}^{x} f(t) \,\mathrm{d}t + C$$

che è senz'altro una funzione continua, ed i cui coefficienti di Fourier sono legati a quelli di f. Infatti per $n \neq 0$ abbiamo

$$\begin{aligned} F_n := (e^{(n)}|F) &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}nx} \left[C + \int_{-\pi}^{x} \mathrm{d}t\, f(t) \right] = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}t\, f(t) \int_{t}^{\pi} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}nx} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}t\, f(t) \frac{\mathrm{e}^{-\mathrm{i}n\pi} - \mathrm{e}^{-\mathrm{i}nt}}{-\mathrm{i}n} = \frac{1}{\mathrm{i}n} \left[f_n + (-1)^n f_0 \right] = 0 \end{aligned}$$

dove abbiamo usato il teorema di Fubini-Tonelli per poter scambiare l'ordine di integrazione. Per quanto riguarda il coefficiente F_0, vale

$$\begin{aligned} F_0 &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}x\, \left[C + \int_{-\pi}^{x} \mathrm{d}t\, f(t) \right] = C + \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}t\, f(t) \int_{t}^{\pi} \mathrm{d}x \\ &= C + \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}t\, f(t)[\pi - t] = C + \pi f_0 - \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{d}t\, t f(t) \,. \end{aligned}$$

Se scegliamo C in modo da annullare F_0, allora abbiamo che $F_n = 0$ per ogni n. Inoltre abbiamo che

$$F(\pi) = \int_{-\pi}^{\pi} f(x)\,\mathrm{d}x + C = \sqrt{2\pi} f_0 + C = C = F(-\pi)$$

Ci troviamo quindi con una funzione continua, periodica, i cui coefficienti di Fourier sono tutti nulli. Allora per il teorema di convergenza puntuale (2.1), $F(x) = 0$ per ogni x. Ma allora abbiamo che, per ogni intervallo $[a, b] \subset [-\pi, \pi]$,

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) = 0\,,$$

ossia l'integrale di f si annulla su ogni intervallo di $[-\pi, \pi]$, ed è noto dalla teoria dell'integrazione che questo implica l'annullarsi quasi ovunque di f.

Ovviamente, quando svolto per le funzioni periodiche di periodo 2π vale per gli spazi L^2 di funzioni periodiche di periodo λ qualsiasi. In questo caso un insieme ortonormale completo è dato dalle onde piane

$$e^{(n)}(x) = \frac{1}{\sqrt{\lambda}} \mathrm{e}^{\mathrm{i}knx} \quad \text{ove} \quad k = \frac{2\pi}{\lambda}\,. \tag{4.35}$$

Poiché questo sistema ortonormale completo è numerabile, tutti gli spazi L^2_λ sono spazi di Hilbert separabili.

SONC in $L^2([a, b], \mathbb{C})$

Dati $a, b \in \mathbb{R}$, $a < b$, Possiamo estendere senza problemi allo spazio $L^2([a, b], \mathbb{C})$, lo spazio delle funzioni $f : [a, b] \to \mathbb{C}$ a quadrato integrabile, quanto detto per lo spazio delle funzioni periodiche di periodo $\lambda = b - a$. Infatti ogni funzione $f : [a, b] \to \mathbb{C}$, estesa per periodicità con periodo $b - a$, individua un'unica funzione periodica[2] $p_f : \mathbb{R} \to \mathbb{C}$ tale che la sua restrizione all'intervallo $[a, b]$ sia uguale ad f: $p_f|_{[a,b]} = f$. Viceversa, ogni funzione periodica $p : \mathbb{R} \to \mathbb{C}$ di periodo $b - a$ individua univocamente la funzione $f_p : [a, b] \to \mathbb{C}$ restrizione di p ad $[a, b]$. Il prodotto scalare in $L^2([a, b], \mathbb{C})$ coincide con quello in L^2_{b-a}, quindi le funzioni (4.35) forniscono un sistema ortonormale completo per lo spazio di Hilbert $L^2([a, b], \mathbb{C})$.

Poiché questo sistema ortonormale completo è numerabile, tutti gli spazi $L^2([a, b])$ sono spazi di Hilbert separabili.

Nel capitolo sulle serie di Fourier abbiamo visto che i coefficienti di Fourier si possono definire per tutte le funzioni L^1_λ periodiche (o con supporto un intervallo finito $[a, a + \lambda]$). Nulla però garantisce la convergenza della corrispondente serie di Fourier. Neanche la successione delle somme parziali

$$[S_N f](x) := \sum_{n=-N}^{N} c_n(f) \mathrm{e}^{\mathrm{i}nkx}$$

converge in qualche norma ad una qualche funzione per una generica $f \in \mathcal{L}^1_\lambda$. Tuttavia, se c'è convergenza in norma-1, il limite è proprio f:

Teorema 4.21 *Siano $f, g \in L^1([a, b])$ o in L^1_λ con $\lambda = b - a > 0$. Se*

$$||S_N f - g||_1 \xrightarrow{N\to\infty} 0\,,$$

allora $f(x) = g(x)$ quasi ovunque.

Isomorfismo tra l^2 e gli spazi hilbertiani separabili

Siccome la completezza di un sistema ortonormale è equivalente, grazie al teorema 4.18, alla decomposizione (4.30), ogni vettore x di uno spazio di Hilbert separabile è univocamente determinato dalla successione dei suoi prodotti scalari con i vettori di base $(e_n|x)$, la quale è una successione di l^2, come mostra l'identità di Parseval (4.31). Possiamo quindi affermare che tutti gli spazi di Hilbert separabili sono in un certo senso equivalenti ad l^2 e pertanto equivalenti tra loro. Più precisamente:

[2] Se $f(b) \neq f(a)$ è sufficiente alterare f nel punto b, quindi in un insieme di misura nulla, affinché $f(b) = f(a)$.

Teorema 4.22 *Sia $\mathcal{H}$ spazio di Hilbert separabile di dimensione infinita. Allora $\mathcal{H}$ è isomorfo ed isometrico ad l^2.*

Dimostrazione: Il teorema 4.20 ci garantisce che $\mathcal{H}$ ammette un sistema ortonormale completo $(e_n)_{n\in\mathbb{N}}$. La corrispondenza

$$A : \mathcal{H} \to l^2 \ , \quad x \mapsto (x_0, x_1, \cdots, x_n, \cdots) \quad : \quad x_n := (e_n|x)$$

è ben definita, in quanto il teorema 4.18 stabilisce che la successione $(x_n)_{n\in\mathbb{N}}$ appartiene ad l^2. Tale corrispondenza è anche lineare, per la linearità destra del prodotto scalare. È iniettiva, in quanto $A(x) = 0 \in l^2$ può valere solo se $x = 0 \in \mathcal{H}$. È suriettiva perchè ad ogni successione (x_n) di l^2 corrisponde un vettore $x = \sum_{n\in\mathbb{N}} x_n e_n$ in $\mathcal{H}$. Infine la corrispondenza è un'isometria, grazie all'identità di Parseval per la quale $||x||_{\mathcal{H}} = ||(x_n)||_2$. C.V.D.

4.3.1 Polinomi ortogonali

Chiaramente uno spazio di Hilbert può ammettere più di un sistema ortonormale completo. Già in dimensione finita $n \geq 2$ abbiamo infiniti sistemi ortonormali completi differenti tra loro. Negli spazi di funzioni $L^2([a, b])$ è particolarmente utile avere dei sistemi ortonormali completi costituiti da polinomi. Infatti, nelle applicazioni pratiche si usa spesso sviluppare una funzione in serie di potenze di un qualche parametro, sperando che la serie risulti convergente in qualche senso. Le approssimazioni ottenute troncando la serie sono delle espressioni polinomiali in tale parametro. Lavorando con funzioni appartenenti ad uno spazio di Hilbert risulta allora naturale chiedersi se è possibile costruire un sistema ortonormale completo formata da polinomi. Ogni polinomio è combinazione lineare finita di monomi, per cui la tecnica usualmente adottata è quella di considerare i monomi e da questi mediante una procedura di ortonormalizzazione di Gram-Schmidt, ottenere delle funzioni polinomiali con buone proprietà ai fini dell'integrazione.

Teorema 4.23 *Il sistema ortonormale costruito, mediante il processo di ortonormalizzazione di Gram-Schmidt, dai monomi $\{1, x, x^2, \cdots, x^n, \cdots\}$ costituisce un sistema ortonormale completo in $L^2([a, b])$ con a e b reali e finiti.*

Dimostrazione: Innanzitutto l'insieme dei monomi è un sistema linearmente indipendente, per il principio di identità dei polinomi:

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N = 0 \quad \forall\, x \in [a, b] \qquad \Longleftrightarrow \qquad a_0 = a_1 = \cdots = a_N = 0 \ .$$

I sottospazi generati dai monomi (l'insieme formato dai polinomi) e dal sistema di polinomi ortonormali generati col procedimento di Gram-Schmidt coincidono, per cui è sufficiente mostrare che il loro complemento ortogonale è nullo. L'idea è questa:
- si assume che un elemento $u \in L^2([a, b])$ sia ortogonale a tutti i monomi;
- si considera il prodotto scalare $(e^{(n)}|u)$ con tutte le onde piane $e^{(n)}$ di periodo $(b - a)/n$;
- ogni onda piana è una serie di potenze che converge non solo puntualmente, ma anche in norma-1, grazie al teorema di convergenza dominata di Lebesgue;
- l'ortogonalità a tutti i monomi implica l'ortogonalità a tutte le onde piane;
- siccome il sistema delle onde piane è completo, allora $u = 0$. C.V.D.

Naturalmente gli elementi di un sistema ortonormale sono definiti a meno di un fattore di fase (numero complesso di modulo unitario). Inoltre a volte può essere conveniente normalizzare i vettori del sistema non imponendo norma unitaria ma qualche altra condizione suggerita dal problema particolare che si sta affrontando. In questi casi si hanno sistemi ortogonali completi, che però hanno caratteristiche del tutto equivalenti ai sistemi ortonormali.

Polinomi di Legendre

Il sistema ortogonale completo di polinomi nello spazio di Hilbert $L^2([-1, 1])$ che si ottiene con il metodo descritto nella sezione precedente porta ai famosi *polinomi di Legendre* $P_l : [-1, 1] \to \mathbb{R}$, $l \in \mathbb{N}$, ai quali non si impone la normalizzazione secondo il prodotto scalare ma la condizione $P_l(1) = 1$. I primi polinomi di Legendre sono

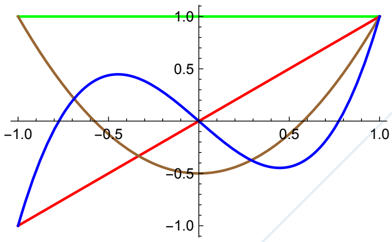
$$P_0(x) = 1 \qquad P_1(x) = x$$
$$P_2(x) = \frac{1}{2}(3x^2 - 1) \qquad P_3(x) = \frac{1}{2}(5x^3 - 3x)$$
$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3) \qquad P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

La teoria dei polinomi ortogonali è molto ricca e sviluppata. Qui ci limiteremo ad osservare alcune delle proprietà di questi sistemi di funzioni. Innanzitutto esistono forme chiuse per i polinomi di Legendre, ad esempio

$$P_l(x) = \frac{1}{2^l l!} \frac{\mathrm{d}^l[(x^2 - 1)^l]}{\mathrm{d}x^l} \tag{4.36}$$

$$= \frac{1}{2^l l!} \sum_{n=0}^{l} (-1)^{l-n} \binom{l}{n} \frac{(2n)!}{(2n - l)!} x^{2n-l} \,, \tag{4.37}$$

dove le potenze negative sono eliminate ponendo, per convenzione, $(-k)! = \infty$ per ogni $k > 0$. Si vede anche che i polinomi di Legendre hanno parità definita uguale a $(-1)^l$.



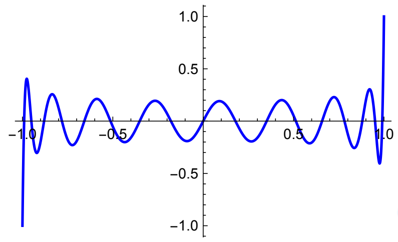


Figura 4.1: A sinistra: i primi polinomi di Legendre $P_l(x)$: $l = 0$ (verde), $l = 1$ (rosso), $l = 2$ (marrone), $l = 3$ (blu). A destra: il polinomio per $l = 17$.

I polinomi di Legendre soddisfano l'equazione differenziale

$$(1 - x^2)P_l''(x) - 2xP_l'(x) + l(l + 1)P_l(x) = 0 \tag{4.38}$$

e si possono calcolare per ricorrenza, a partire da P_0 e P_1, mediante la formula

$$(l + 1)P_{l+1}(x) = (2l + 1)xP_l(x) - lP_{l-1}(x) \,. \tag{4.39}$$

Dal momento che per i polinomi di Legendre vale

$$(P_m|P_n) = \frac{2}{2n + 1}\delta_{mn} \,, \tag{4.40}$$

l'insieme $\{\psi_l = \sqrt{(2l + 1)/2}\, P_l : l \in \mathbb{N}\}$ è un sistema ortonormale completo in $L^2([-1, 1])$.

Polinomi di Jacobi, Gegenbauer, Chebyshev

Il concetto di polinomio ortogonale dipende direttamente dalla definizione di prodotto scalare. Ad esempio, possiamo modificare la "misura" dell'integrazione considerando lo spazio di Hilbert $L^2(]a, b[; \mathrm{d}\mu)$ (a, b finiti o infiniti) con la misura $\mathrm{d}\mu$ definita da $\mathrm{d}\mu(x) := M(x)\mathrm{d}x$, ove $x \mapsto M(x)$ è una funzione definita positiva in $]a, b[$ e tale che

$$\int_a^b x^n \, \mathrm{d}\mu(x) := \int_a^b x^n M(x) \, \mathrm{d}x < \infty \quad \text{per ogni} \quad n \in \mathbb{N} \,. \tag{4.41}$$

$\mathcal{H} = L^2(]a, b[; \mathrm{d}\mu)$ è definito come l'insieme delle funzioni a quadrato sommabile nella misura $\mathrm{d}\mu$:

$$L^2(]a, b[; \mathrm{d}\mu) := \{f :]a, b[\to \mathbb{C} \,,\, \int_a^b |f|^2 \mathrm{d}\mu < \infty\} \tag{4.42}$$

ed è spazio di Hilbert rispetto al prodotto scalare definito da

$$(f|g) := \int_a^b f^* g \, \mathrm{d}\mu = \int_a^b f^*(x) g(x) M(x) \, \mathrm{d}x \, . \tag{4.43}$$

Una classe importante di polinomi ortogonali deriva dalla scelta

$$]a,b[\;\; = \;\;]-1,1[\, , \qquad M(x) = (1-x)^\alpha (1+x)^\beta \; , \quad \alpha, \beta > -1 \, . \tag{4.44}$$

I corrispondenti polinomi ortogonali $P_n^{(\alpha,\beta)}$: $n \in \mathbb{N}$ vanno sotto il nome di *polinomi di Jacobi*, e formano un sistema ortogonale completo di $\mathcal{H}$.

In particolare, se $\alpha = \beta = \rho - 1/2$, $\rho > -1/2$, si hanno i cosiddetti *polinomi di Gegenbauer* $C_n^{(\rho)}(x)$ (con l'opportuna normalizzazione). I polinomi di Legendre sono evidentemente un caso particolare di polinomi di Gegenbauer, con $\rho = 1/2$.

Il caso $\rho = 0$ identifica i *polinomi di Chebyshev* $T_n(x)$, per i quali $M(x) = (1-x^2)^{-1/2}$. I polinomi di Chebyshev sono dati da

$$\begin{aligned} T_0(x) &= 1 & T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 & T_3(x) &= 4x^3 - 3x \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x) \end{aligned} \tag{4.45}$$

e possono essere espressi dalla relazione

$$T_n(x) = \cos\big(n \arccos(x)\big) \; , \quad n \in \mathbb{N} \, .$$

Utilizzando la trigonometria è facile convincersi che tale espressione definisce un polinomio. I polinomi di Chebyshev soddisfano la relazione di ortogonalità

$$(T_m|T_n) = \int_{-1}^{1} T_m(x) T_n(x) \, \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \frac{\pi}{2} \delta_{mn} (1 + \delta_{n0}) \, .$$

I polinomi di Chebyshev sono usati ad esempio nella ricerca di buone interpolazioni polinomiali per una funzione

$$f(x) \simeq \sum_{n=0}^{N} a_n T_n(x) \, .$$

L'approssimazione risulta generalmente buona nel senso che risulta molto piccola la differenza massima (in valore assoluto) tra $f(x)$ e lo sviluppo in polinomi di Chebyshev, troncato ad un certo ordine N prefissato.

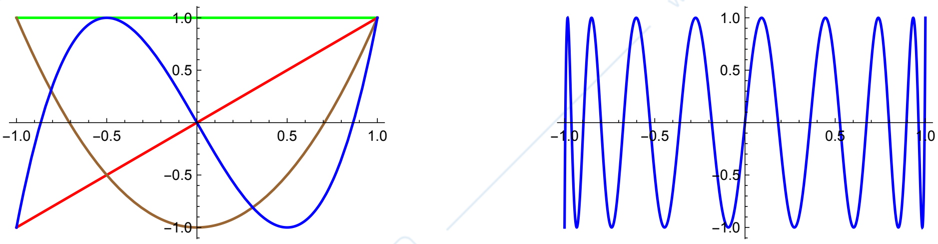


Figura 4.2: A sinistra: i primi polinomi di Chebyshev $T_n(x)$: $n = 0$ (verde), $n = 1$ (rosso), $n = 2$ (marrone), $n = 3$ (blu). A destra: il polinomio per $n = 17$.

Polinomi di Laguerre

Nel caso di un intervallo infinito, per esempio $[a,b[= [0,\infty[$, né i polinomi trigonometrici (combinazioni lineari di serie trigonometriche finite) né i polinomi usuali sono a quadrato sommabile. Anzi, non stanno in $L^p([0,\infty[; \mathrm{d}x)$ per alcun p. In questo caso l'introduzione di una misura che renda convergenti gli integrali all'infinito è necessaria. Se si usa la misura

$$M(x) = \mathrm{e}^{-x} x^\alpha \; , \quad \alpha > -1 \qquad x \in [0, \infty[\; , \tag{4.46}$$

i monomi ed i polinomi di qualsiasi ordine sono a quadrato sommabile, e la procedura di Gram-Schmidt genera (a meno della normalizzazione) i *polinomi di Laguerre*

$$L_n^{(\alpha)}(x) = \frac{1}{n!}\mathrm{e}^{-x}x^\alpha \frac{\mathrm{d}^n}{\mathrm{d}x^n}[\mathrm{e}^{-x}x^{n+\alpha}] \tag{4.47}$$

e che soddisfano l'equazione differenziale

$$x\frac{\mathrm{d}^2 L_n^{(\alpha)}(x)}{\mathrm{d}x^2} + (\alpha+1-x)\frac{\mathrm{d}L_n^{(\alpha)}(x)}{\mathrm{d}x} + nL_n^{(\alpha)}(x) = 0\,. \tag{4.48}$$

Fissato $\alpha > -1$, i polinomi di Laguerre formano un sistema ortogonale completo di $\mathcal{H} = L^2([0,\infty[;\mathrm{d}\mu^{(\alpha)})$.

Polinomi di Hermite

Il più noto sistema di polinomi ortogonali definiti su tutto l'asse reale è dato dai polinomi di Hermite. Essi sono determinati dalla solita procedura quando si adotti la misura

$$M(x) = \mathrm{e}^{-x^2}\,, \qquad x \in]-\infty,\infty[\,. \tag{4.49}$$

Si possono ottenere per ricorrenza

$$\begin{aligned} H_0(x) &= 1 & H_1(x) &= 2x \\ H_2(x) &= 4x^2-2 & H_3(x) &= 8x^3-12x \\ H_{n+1}(x) &= 2xH_n(x) - 2nH_{n-1}(x) \end{aligned} \tag{4.50}$$

e anche dalla formula

$$H_n(x) = (-1)^n \mathrm{e}^{-x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-x^2}\,. \tag{4.51}$$

I polinomi di Hermite soddisfano l'equazione differenziale

$$\frac{\mathrm{d}^2 H_n(x)}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}H_n(x)}{\mathrm{d}x} + 2nH_n(x) = 0 \tag{4.52}$$

e formano un sistema ortonormale completo di $L^2(\mathbb{R}; M\mathrm{d}x)$

4.3.2 La trasformata di Fourier in L^2

Nella sez. 2.7.1 abbiamo visto che la trasformata di Fourier (TdF) è un automorfismo (lineare e biiettivo) di $\mathcal{S}(\mathbb{R}^n)$ in sé. Anzi, introducendo il solito prodotto scalare hermitiano, lo spazio $\mathcal{S}(\mathbb{R}^n)$ diventa spazio pre-hilbertiano e la TdF diventa un'isometria, cioè una trasformazione che conserva i prodotti scalari e la norma da questo indotta. Tuttavia lo spazio di Schwartz non è completo nella norma-2, poiché possiamo trovare successioni di Cauchy che convergono a funzioni non lisce, addirittura discontinue.

Più precisamente, pensando $\mathcal{S}(\mathbb{R}^n) < L^2(\mathbb{R}^n)$, esistono successioni di Cauchy $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ in $||\cdot||_2$ che convergono a funzioni di L^2 che non stanno in $\mathcal{S}$. Tuttavia ogni successione di Cauchy in $\mathcal{S}$, essendo successione di Cauchy in L^2 che è completo, converge in L^2. E ogni funzione f di L^2 è il limite di qualche successione $(f_n)_{n\in\mathbb{N}}$ di funzioni di $\mathcal{S}$. In altre parole, $\mathcal{S}(\mathbb{R}^n)$ è denso in $L^2(\mathbb{R}^n)$. Si può dimostrare che $L^2(\mathbb{R}^n)$ è proprio il completamento (spazio di Hilbert) dello spazio pre-hilbertiano $\mathcal{S}(\mathbb{R}^n)$.

Invocando il teorema di estensione continua 3.10 possiamo estendere la trasformata di Fourier da $\mathcal{S}(\mathbb{R}^n)$ ad $L^2(\mathbb{R}^n)$, in quanto la TdF è una isometria, quindi continua ed anche uniformemente continua.

Teorema 4.24 *La TdF in $L^2(\mathbb{R}^n)$ ottenuta per estensione continua della TdF in $\mathcal{S}(\mathbb{R}^n)$*

$$\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)\,. \tag{4.53}$$

è unitaria, ossia è un automorfismo che preserva il prodotto scalare.

Dimostrazione: (a) Conservazione del prodotto scalare. Se $f, g \in L^2$, possiamo approssimarle con successioni di elementi di $\mathcal{S}$:

$$\phi_n \xrightarrow{n\to\infty} f \ , \qquad \psi_n \xrightarrow{n\to\infty} g \ , \qquad \phi_n, \psi_n \in \mathcal{S}$$

e, per definizione di estensione continua,

$$\mathcal{F}\phi_n \xrightarrow{n\to\infty} \mathcal{F}f \ , \qquad \mathcal{F}\psi_n \xrightarrow{n\to\infty} \mathcal{F}g \ , \qquad \phi_n, \psi_n \in \mathcal{S}$$

Dalla conservazione del prodotto scalare in $\mathcal{S}$, eq. (2.132)

$$(\mathcal{F}\phi_n|\mathcal{F}\psi_n) = (\phi_n|\psi_n)$$

e dalla continuità del prodotto scalare in L^2, otteniamo la conservazione del prodotto scalare in L^2:

$$(\mathcal{F}\phi|\mathcal{F}\psi) = (\phi|\psi) \ . \tag{4.54}$$

(b) L'iniettività segue subito dalla linearità di $\mathcal{F}$ e dall'isometria:

$$\mathcal{F}f = \mathcal{F}g \quad \Longleftrightarrow \quad 0 = ||\mathcal{F}f - \mathcal{F}g||_2 = ||\mathcal{F}(f-g)||_2 = ||f-g||_2 \quad \Longleftrightarrow \quad f = g \ .$$

(c) La suriettività è conseguenza della suriettività in $\mathcal{S}$: se $f \in L^2$ allora

$$\begin{aligned} f &= \lim_{n\to\infty} \phi_n \ , & \phi_n &\in \mathcal{S} \ , \\ \phi_n &= \mathcal{F}\psi_n \ , & \psi_n &\in \mathcal{S} \ . \end{aligned}$$

Le funzioni ψ_n formano una successione di Cauchy:

$$||\psi_n - \psi_m||_2 \overset{(2.132)}{=} ||\mathcal{F}\psi_n - \mathcal{F}\psi_m||_2 = ||\phi_n - \phi_m||_2 \xrightarrow{m,n\to\infty} 0 \ ,$$

quindi convergenti ad un elemento $g \in L^2$, e per continuità

$$\mathcal{F}g = \lim_{n\to\infty} \mathcal{F}\psi_n = \lim_{n\to\infty} \phi_n = f \ .$$

C.V.D.

Come in $\mathcal{S}$, L'inversa della TdF in L^2 è data da $\mathcal{F}^{-1} = \imath \circ \mathcal{F} = \mathcal{F} \circ \imath$.

Osservazione: La definizione di trasformata di Fourier tramite l'integrale (2.74) non è più valida in generale in tutto $L^2(\mathbb{R}^n)$, ma dobbiamo ricorrere alla definizione come estensione continua da $\mathcal{S}(\mathbb{R}^n)$ a $L^2(\mathbb{R}^n)$. Lo stesso discorso vale per la sua inversa $\mathcal{F}^{-1}$ definita anch'essa in generale come estensione continua della antitrasformata di funzioni in $\mathcal{S}(\mathbb{R}^n)$.