

Spazi di Hilbert

Dopo aver richiamato nel Capitolo 1 le nozioni di base relative agli spazi vettoriali a dimensione finita, questo secondo Capitolo sarà dedicato agli spazi a dimensione infinita, e più precisamente agli spazi di Hilbert.

Il formalismo vettoriale non rappresenta semplicemente una “sovrastuttura” matematica (magari soltanto elegante) per molte situazioni fisiche, bensì costituisce la vera e propria struttura matematica adeguata a tradurre e descrivere il “principio di sovrapposizione” o di “linearità”.

Il principio di sovrapposizione è alla base di moltissimi processi fisici e ricordiamo che è uno dei “principi costitutivi” della meccanica quantistica. Ovviamente non tutti i fenomeni fisici obbediscono a leggi lineari; anzi, gli effetti non lineari (purtroppo assai difficili da studiare) stanno ricevendo attualmente sempre maggiore attenzione, tuttavia è ben chiaro che la schematizzazione lineare costituisce una prima (e spesso buona o quanto meno istruttiva) approssimazione dei fenomeni.

Partendo dalla serie trigonometrica di Fourier, verrà mostrata la necessità di introdurre gli spazi di funzioni L^2 e più in generale gli spazi di Hilbert, dove sarà possibile ambientare correttamente il procedimento degli sviluppi in serie, e di introdurre una generalizzazione “astratta” che porterà alla importante nozione di “set completo” di vettori.

Verranno poi studiati gli operatori lineari fra spazi di Hilbert, illustrandone le proprietà fondamentali, e introducendo classi di operatori con speciali caratteristiche, come gli operatori unitari, i proiettori, gli operatori simmetrici, gli operatori compatti. Particolare attenzione sarà rivolta ai problemi agli autovalori e alle proprietà spettrali, che rivestono enorme importanza, anche in meccanica quantistica. Questi problemi si presentano in dimensione infinita con proprietà radicalmente diverse rispetto alla dimensione finita, come verrà illustrato da numerosi esempi. Il caso specialmente interessante dell'equazione agli autovalori di Sturm–Liouville porterà ad un cenno (necessariamente sommario) ad alcune funzioni speciali spesso incontrate in fisica.

Come applicazioni immediate, potremo studiare l'evoluzione temporale di problemi periodici, delle oscillazioni elastiche (le vibrazioni di una corda o di una membrana) e della propagazione del calore, descritte rispettivamente dalle equazioni a derivate parziali di d'Alembert e di Fourier. Un'altra applicazione riguarderà le funzioni armoniche (equazione di Laplace), ovvero i problemi di potenziale.

Alcuni fra questi ultimi problemi verranno poi ripresi, con tecniche diverse, nei Capitoli successivi.

2.1 Equazione di d'Alembert. Onde stazionarie

Cominciamo a studiare il movimento di una corda perfettamente flessibile ed elastica, tesa fra i suoi estremi. Fissato l'asse x lungo la posizione di equilibrio della corda, è chiaro intanto che per descrivere completamente la situazione della corda

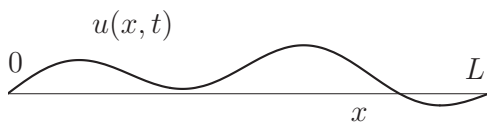


Figura 2.1 Una configurazione della corda elastica a estremi fissi.

occorre una funzione $u = u(x, t)$ di due variabili che descriva lo spostamento u (v. fig. 2.1) dalla posizione di equilibrio di ciascun punto della corda di ascissa x al tempo t . Si può immaginare che la corda sia costituita da una successione di infinite particelle uguali e molto vicine legate fra loro da forze elastiche, in modo analogo a quanto succedeva nel problema del § 1.7b (la differenza è che ora si prendono in considerazione solo movimenti *trasversali*, mentre nel § 1.7b si consideravano quelli *longitudinali*). Proprio utilizzando questa idea, si arriva a dimostrare che la funzione $u(x, t)$ che descrive le piccole vibrazioni della corda obbedisce alla seguente equazione

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \quad (2.1)$$

dove v è una costante che dipende dalle caratteristiche fisiche del problema ($v = \sqrt{\tau/\rho}$, dove τ è tensione della corda e ρ la sua densità lineare).

L'equazione (2.1) è una **equazione differenziale alle derivate parziali** e prende il nome di **equazione di d'Alembert**.

Studiamo il caso in cui gli estremi della corda (siano posti in $x = 0$ e $x = L$) vengano tenuti fissi. Questo si traduce nelle condizioni, per ogni t ,

$$u(0, t) = u(L, t) = 0 \quad (2.2)$$

che sono dette **condizioni al contorno** o **al bordo** o **agli estremi**.

Cerchiamo intanto particolari soluzioni dell'equazione di d'Alembert (cioè particolari tipi di movimento della corda) e precisamente soluzioni $u(x, t)$ che siano esprimibili come un prodotto:

$$u(x, t) = X(x) T(t)$$

in cui $X(x)$ è una funzione della sola variabile x e $T(t)$ della sola t . Questo procedimento si chiama metodo della **separazione delle variabili**. Sostituendo nella (2.1), l'equazione si scompone e diventa

$$\frac{X''}{X} = \frac{1}{v^2} \frac{T''}{T}.$$

In questa equazione il primo membro dipende solo da x , il secondo solo da t : dovendo i due membri essere identicamente uguali fra loro, l'unica possibilità è che entrambi siano uguali ad una costante C . Si ottengono così le due equazioni separate

$$\begin{cases} X'' - CX = 0 \\ T'' - C v^2 T = 0. \end{cases} \quad (2.1')$$

La più generale soluzione della prima equazione è

$$X(x) = A e^{\sqrt{C}x} + B e^{-\sqrt{C}x},$$

ma le condizioni al contorno (2.2), che ora diventano

$$X(0) = X(L) = 0$$

impongono che

$$\begin{cases} X(0) = A + B = 0 \\ X(L) = A e^{\sqrt{C}L} + B e^{-\sqrt{C}L} = 0 \end{cases}$$

e cioè

$$e^{\sqrt{C}L} = e^{-\sqrt{C}L}.$$

Allora $\sqrt{C}$ non può essere reale, e ricordando che per due numeri complessi α, β si ha

$$e^\alpha = e^\beta \iff \alpha = \beta + 2n\pi i, \quad n = 0, \pm 1, \pm 2, \dots$$

ne risulta che deve essere

$$C = -\frac{n^2\pi^2}{L^2}$$

e dunque le soluzioni possibili saranno

$$X(x) = X_n(x) = A_n \left(e^{ik_n x} - e^{-ik_n x} \right) = a_n \sin(k_n x) \quad (2.3)$$

dove si sono indicate con A_n e a_n le costanti moltiplicative arbitrarie (per ciascun n) e dove

$$k_n = \frac{n\pi}{L}, \quad C = -k_n^2. \quad (2.3')$$

Si vede così che le condizioni al contorno impongono alla costante C di assumere soltanto i valori indicati dalla condizione (2.3'), nella quale vanno assegnati i valori

$$n = 1, 2, 3, \dots$$

dato che $n = 0$ corrisponde alla soluzione identicamente nulla e il cambiamento di segno di n non produce soluzioni nuove.

Per la seconda delle equazioni (2.1') non ci sono condizioni al contorno e la soluzione più generale sarà

$$T(t) = A'_n \sin(\omega_n t) + B'_n \cos(\omega_n t) = E_n \cos(\omega_n t + \delta_n)$$

avendo posto $\omega = \omega_n = vk_n$ e indicato con $A'_n, B'_n, E_n, \delta_n$ le solite costanti arbitrarie, per ciascun n .

Fin qui abbiamo considerato il caso della corda ad estremi fissi. Naturalmente sarebbe possibile trattare in modo simile altre condizioni al contorno, ad esempio

$$\frac{\partial}{\partial x} u(0, t) = 0, \quad \frac{\partial}{\partial x} u(L, t) = 0$$

ovvero, eseguita la separazione delle variabili,

$$X'(0) = 0, \quad X'(L) = 0,$$

che corrisponde ad una situazione in cui gli estremi della corda possono spostarsi conservando però la tangente orizzontale in $x = 0$ e $x = L$. In questo caso le soluzioni per le $X_n(x)$ sarebbero

$$X_n(x) = 1, \cos(k_n x) \quad (n = 1, 2, \dots)$$

avendo ora tralasciato di scrivere le costanti arbitrarie.

Altre condizioni al bordo potrebbero essere p.es. $u(0, t) = 0, \frac{\partial u(L, t)}{\partial x} = 0$; oppure anche le “condizioni di periodicità”: v. § 2.15.

Tornando comunque per concretezza al caso $X(0) = X(L) = 0$, si sono dunque trovate infinite soluzioni a variabili separate, che – a meno del solito fattore arbitrario – sono date da

$$u_n(x, t) = \sin(k_n x) \cos(\omega_n t + \delta_n); \quad k_n = \frac{n\pi}{L}, \quad \omega_n = k_n v. \quad (2.4)$$

Queste soluzioni sono caratteristiche per la loro particolare dipendenza da x e da t e si chiamano **onde stazionarie**; esse descrivono i modi più semplici di oscillazione della corda (**modi armonici** o **normali**). Il modo n -simo corrisponde alla configurazione in cui il profilo della corda è in ogni istante una sinusoide di lunghezza d'onda $\lambda_n = 2\pi/k_n = 2L/n$, mentre al variare del tempo cambia solo l'ampiezza di tale sinusoide (v. fig. 2.2, dove è rappresentata la situazione nel caso $n = 3$); la variazione col tempo di questa ampiezza avviene con legge sinusoidale con “frequenza”

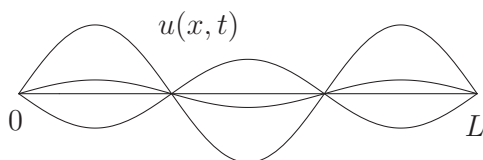


Figura 2.2 Onde stazionarie per la corda elastica a estremi fissi.

$\omega_n = k_n v = n\pi v/L$. Quando la corda si trova in uno di questi stati di vibrazione u_n , essa emette un “suono puro” cioè una nota con lunghezza d’onda fissata λ_n .

2.2 Primi problemi concernenti gli spazi a dimensione infinita

Quanto visto nel paragrafo precedente offre immediatamente lo spunto e la motivazione per vari approfondimenti.

Anzitutto, si può osservare che l’equazione che si è ottenuta per la $X(x)$, riscritta nella forma

$$\frac{d^2}{dx^2} X = -k^2 X, \quad X(0) = X(L) = 0$$

può essere vista come una equazione agli autovalori nello spazio vettoriale (di dimensione infinita!) delle funzioni $X(x)$ derivabili almeno due volte, definite nell’intervallo $0, L$ e che si annullano agli estremi di questo intervallo. È chiaro che l’operatore di cui si cercano gli autovalori e le autofunzioni è l’operatore “derivata seconda”, che è un operatore lineare su questo spazio. Da notare che sono precisamente le condizioni di annullamento al bordo a limitare gli autovalori $\lambda = -k^2$ a ben determinati valori “discreti” $k^2 = k_n^2 = n^2\pi^2/L^2$. È ovvio che dovremo riconsiderare accuratamente questo fatto (v. § 2.20).

Ma c’è un altro punto che dobbiamo subito mettere in evidenza. Poiché l’equazione di d’Alembert è un’equazione *lineare*, se u_1 e u_2 sono soluzioni dell’equazione, anche ogni combinazione lineare $\alpha_1 u_1 + \alpha_2 u_2$ lo è. Ma ora di soluzioni indipendenti ce ne sono a disposizione infinite (le soluzioni trovate prima) e quindi, al posto delle combinazioni lineari finite (come nel primo capitolo), ci aspettiamo di dover scrivere delle *serie* di funzioni. Naturalmente questo comporterà nuovi problemi, che dovremo affrontare con cura. Prima di tutto, dovremo ovviamente preoccuparci della convergenza di tali serie, o – più esattamente – di quale dovrà essere *il tipo di convergenza da richiedere*: le nozioni di convergenza puntuale, di convergenza puntuale quasi ovunque, di convergenza uniforme avranno un ruolo importante, ma vedremo che il tipo di convergenza più appropriato sarà un altro.

D’altronde la effettiva necessità di ricorrere a queste “combinazioni infinite” si intuisce facilmente pensando al tipico **problema di Cauchy**, cioè: data una configurazione iniziale della corda, determinarne il suo andamento nel tempo (il pro-

blema è analogo – a parte appunto la dimensione – a quello risolto nel § 1.7a). Poiché conosciamo l'evoluzione temporale delle onde stazionarie, sarà dunque necessario decomporre la situazione iniziale come “somma” (cioè qui come opportuna serie) delle onde stazionarie. E prima ancora sarà da chiedersi se le onde stazionarie sono in numero “sufficiente” per risolvere tutti i possibili problemi di Cauchy. Per affrontare questi problemi, è necessario cominciare rivedendo qualche risultato classico dell'analisi matematica a proposito degli sviluppi in serie tramite *funzioni trigonometriche*.

2.3 La serie di Fourier trigonometrica: le sue prime proprietà e difficoltà

Sia $f = f(x)$ una funzione, definita nell'intervallo finito $(-L, L)$, continua, oppure solo **generalmente continua** (cioè dotata al più di un numero finito di punti di discontinuità, tale che nei punti di discontinuità esistono finiti i limiti destro e sinistro; tali funzioni sono anche dette **continue a tratti**), e inoltre dotata di derivata prima pure generalmente continua. Si considerino i seguenti integrali

$$\begin{cases} \alpha_0 = \frac{1}{2L} \int_{-L}^{+L} f(x) dx \\ \alpha_n = \frac{1}{L} \int_{-L}^{+L} f(x) \cos\left(\frac{n\pi}{L}x\right) dx \\ \beta_n = \frac{1}{L} \int_{-L}^{+L} f(x) \sin\left(\frac{n\pi}{L}x\right) dx \end{cases} \quad n = 1, 2, \dots \quad (2.5)$$

e poi la serie

$$\alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos\left(\frac{n\pi}{L}x\right) + \sum_{n=1}^{\infty} \beta_n \sin\left(\frac{n\pi}{L}x\right) \quad (2.6)$$

che viene detta **serie di Fourier** (trigonometrica) della funzione $f(x)$. Si dimostra che valgono le due seguenti basilari proprietà di convergenza:

- (a) Se la funzione $f(x)$ è continua con derivata prima generalmente continua, e inoltre $f(L) = f(-L)$, allora la serie (2.6) converge uniformemente alla $f(x)$.
- (b) Se la funzione $f(x)$ è generalmente continua, con derivata prima generalmente continua, la serie (2.6) converge in ogni punto; ha per somma la $f(x)$ in tutti i punti in cui la $f(x)$ è continua, mentre in ciascun punto x_0 di discontinuità, essa converge alla “media” $\frac{1}{2} [f(x_0^+) + f(x_0^-)]$. Infine, nei punti estremi $\pm L$ la serie converge alla media $\frac{1}{2} [f(L) + f(-L)]$.

Da osservare anzitutto che serie (2.6) definisce in realtà una funzione *periodica*, che si estende “automaticamente” anche fuori dell'intervallo $(-L, L)$ con periodo $2L$. Questo spiega anche la particolare convergenza della serie nei punti $\pm L$: infatti, se $f(L) \neq f(-L)$, ripetendo periodicamente la $f(x)$ si viene a manifestare una discontinuità negli estremi dell'intervallo.

Da osservare pure che nei punti in cui la $f(x)$ presenta delle discontinuità, la convergenza della serie *non può* essere uniforme: basta infatti ricordare che – come è ben noto – una serie (o una successione) di funzioni continue che converge *uniformemente* ha per somma (o per limite) una funzione *continua*.

Ad esempio, il prolungamento periodico della funzione definita in $|x| < L$ da $f(x) = x$ risulterà discontinuo nei punti $x = \pm L$ (anzi in tutti i punti $x = \pm L, \pm 3L, \dots$): in tali punti la serie di Fourier convergerà a zero e non ci potrà essere convergenza uniforme.

Grazie ai risultati (a) e (b), si scrive

$$f(x) = \alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos\left(\frac{n\pi}{L}x\right) + \sum_{n=1}^{\infty} \beta_n \sin\left(\frac{n\pi}{L}x\right) \quad (2.7)$$

e si dice che la funzione $f(x)$ è stata sviluppata nella serie (trigonometrica) di Fourier (2.7).

Ovviamente, se p.es. la funzione $f(x)$ è pari allora si ha $\beta_n = 0$ per ogni n e la formula per gli α_n si può semplificare:

$$\alpha_0 = \frac{1}{L} \int_0^L f(x) dx \quad \alpha_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{n\pi}{L}x\right) dx.$$

Un risultato simile vale chiaramente per le funzioni dispari.

Consideriamo p.es. l'“onda quadra” (con $L = \pi$)

$$f(x) = \begin{cases} -1 & \text{per } -\pi < x < 0 \\ 1 & \text{per } 0 < x < \pi \end{cases}$$

i suoi coefficienti di Fourier sono dati da $\beta_n = 4/(n\pi)$ per n dispari e $\beta_n = 0$ per n pari e dunque si ottiene

$$f(x) = \frac{4}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x \dots \right).$$

In fig. 2.3 è riportato il grafico della somma dei primi 4, e dei primi 11 termini non nulli di questo sviluppo in serie.

Una osservazione utile a proposito della convergenza delle serie è la seguente. Si ricordi che una serie di funzioni $\sum_n f_n(x)$ è detta convergere **totalmente** se esiste una successione di costanti M_n tali che $|f_n(x)| \leq M_n$ e la serie numerica $\sum_n M_n$ è convergente; si ricordi pure che la convergenza totale implica la convergenza uniforme.

Ne segue che se i coefficienti α_n, β_n di una serie di Fourier assegnata (2.7) soddisfano la condizione

$$\sum_n |\alpha_n| + \sum_n |\beta_n| < \infty$$

allora la $f(x)$ risulta continua e soddisfa $f(L) = f(-L)$.

Si noti che, partendo dalla (2.7), e *almeno* nei casi in cui la serie converge uniformemente, le formule (2.5) possono essere ottenute direttamente: basta moltiplicare

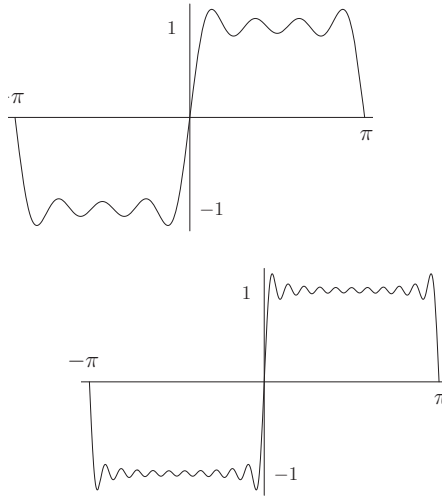


Figura 2.3 Somma dei primi 4 e dei primi 11 termini della serie di Fourier dell'onda quadra.

entrambi i membri della (2.7) per $\cos(n\pi x/L)$ e integrare termine a termine (l'ipotesi di uniformità della convergenza garantisce infatti questa possibilità; per qualche controesempio circa la possibilità di passare al limite sotto integrale v. alla fine del § 2.8); osservando poi che

$$\begin{aligned} \int_{-L}^L \cos\left(\frac{n\pi}{L}x\right) \cos\left(\frac{m\pi}{L}x\right) dx &= L\delta_{nm} \\ \int_{-L}^L \cos\left(\frac{n\pi}{L}x\right) \sin\left(\frac{m\pi}{L}x\right) dx &= 0 \end{aligned} \quad (2.8)$$

si ottiene proprio la formula (2.5) per α_m (e analogamente per α_0 e β_m).

L'insieme delle funzioni f che stiamo considerando in questo paragrafo forma evidentemente uno *spazio vettoriale*: emerge allora abbastanza chiara una qualche analogia della decomposizione (1.1) del vettore $x = x_i e^{(i)}$ in dimensione finita con lo sviluppo (2.7) della funzione f come “combinazione lineare infinita” di una “base” costituita dai vettori

$$1, \quad \cos\left(\frac{n\pi}{L}x\right), \quad \sin\left(\frac{n\pi}{L}x\right); \quad n = 1, 2, 3, \dots \quad (2.9)$$

Introducendo infatti un prodotto scalare (f, g) fra le funzioni del nostro spazio nel modo naturale seguente

$$(f, g) = \int_{-L}^{+L} f(x)g(x)dx \quad (2.10)$$

(più in generale, se le funzioni sono complesse, il prodotto scalare si dovrà definire così:

$$(f, g) = \int_{-L}^{+L} f^*(x) g(x) dx \quad (2.10')$$

(v. più avanti, §§ 2.6, 2.9)), le funzioni elencate in (2.9) risultano *ortogonali* fra loro grazie alla (2.8). Tenendo ancora conto delle (2.8), il sistema di vettori (2.9) si può anche normalizzare: basta prendere

$$\frac{1}{\sqrt{2L}}, \quad \frac{1}{\sqrt{L}} \cos\left(\frac{n\pi}{L}x\right), \quad \frac{1}{\sqrt{L}} \sin\left(\frac{n\pi}{L}x\right); \quad n = 1, 2, 3, \dots \quad (2.11)$$

e ottenere così un sistema *ortonormale*: $(f^{(m)}, f^{(n)}) = \delta_{nm}$, avendo indicato con $f^{(n)}$ la generica di queste funzioni; la (2.7) si può così scrivere nella forma

$$f = \sum_{n=0}^{\infty} a_n f^{(n)} \quad (2.12)$$

dove

$$a_0 = \sqrt{2L} \alpha_0, \quad a_{2n-1} = \sqrt{L} \beta_n, \quad a_{2n} = \sqrt{L} \alpha_n. \quad (2.13)$$

I coefficienti a_n dello sviluppo rappresentano le “componenti” di f rispetto al sistema ortonormale $f^{(n)}$ e, grazie alla (2.5), essi si possono calcolare, esattamente come in dimensione finita (v. ad esempio la (1.7)), mediante prodotti scalari:

$$a_n = (f^{(n)}, f). \quad (2.14)$$

Ma l'*analogia* con $\mathbf{C}^n$ non può essere spinta oltre. Innanzitutto, mentre in ogni spazio a dimensione finita a qualsiasi n -pla di numeri $(a_1, a_2, \dots, a_n)$ corrisponde un vettore dello spazio e *viceversa*, qui *non* è certamente vero che ad ogni successione $\{a_n\}$ corrisponda un elemento dello spazio, o comunque si ottenga una serie convergente in qualche senso.

È insomma necessario individuare contemporaneamente:

- (i) un opportuno spazio di funzioni (con proprietà abbastanza generali)
- (ii) una precisa condizione sulla successione dei coefficienti a_m in modo tale che ogni funzione di tipo (i) ammetta uno sviluppo con i coefficienti così individuati, e – *viceversa* – ogni serie costruita con questi coefficienti sia convergente (e *in che senso*) ad una funzione di tipo (i).

E resta ancora da precisare chiaramente questa nozione di “base” per uno spazio di dimensione infinita, da vedere se esistono altre possibili basi (e come si possono individuare), e come estendere il problema ambientandolo in un opportuno spazio “astratto” con caratteristiche generali.

Tuttavia, prima ancora di affrontare questa serie di questioni, quanto detto finora è già sostanzialmente sufficiente per dare un’idea di come risolvere qualche problema concreto concernente il movimento della corda elastica o la propagazione del calore su una sbarra conduttrice.

2.4 Evoluzione temporale di un'onda elastica

Premettiamo un'osservazione importante.

Sia $f(x)$ una funzione con le proprietà assegnate all'inizio del paragrafo precedente, definita però nell'intervallo $(0, L)$. Immaginiamo ora di prolungare tale funzione “con legge dispari” all'intervallo $(-L, L)$, cioè introduciamo una funzione $\bar{f}(x)$ definita in $(-L, L)$ in modo che $\bar{f}(x)$ sia una funzione *dispari* e sia tale che $\bar{f}(x) = f(x)$ quando $0 < x \leq L$. Se ora si sviluppa la $\bar{f}(x)$ in serie di Fourier (2.7) nell'intervallo $(-L, L)$, lo sviluppo che si ottiene dovrà quindi contenere i soli termini in $\sin(n\pi x/L)$ e fornirà uno sviluppo della funzione $f(x)$ in serie di soli seni nell'intervallo $(0, L)$:

$$f(x) = \sum_{n=1}^{\infty} \beta_n \sin\left(\frac{n\pi}{L}x\right). \quad (2.15)$$

Naturalmente, questo sviluppo estende automaticamente oltre l'intervallo $(-L, L)$ la funzione $\bar{f}(x)$ così prolungata, producendo una funzione periodica di periodo $2L$. La possibilità di avere uno sviluppo in $(0, L)$ della forma (2.15) è importante in relazione al problema della corda elastica a estremi fissi in 0 ed L , poiché le funzioni $\sin(n\pi x/L)$ esprimono proprio la dipendenza da x delle onde stazionarie in questo caso, come si è già visto (§ 2.1).

Il problema classico che ora affrontiamo è il problema di Cauchy, cioè quello di determinare, date le opportune informazioni sullo stato iniziale della corda, la configurazione $u(x, t)$ della corda stessa negli istanti successivi (la *evoluzione temporale* dell'onda), e anche di individuare quali sono le frequenze che la corda emette (la *analisi in frequenza*). Il problema si può risolvere completamente assegnando le *condizioni iniziali* della corda e cioè assegnando *due* funzioni che hanno il seguente significato:

$$\begin{cases} f(x) = \text{posizione della corda all'istante zero} = u(x, 0) \\ g(x) = \text{“velocità” di ciascun punto della corda all'istante zero} = \frac{\partial u(x, t)}{\partial t} \Big|_{t=0} \end{cases}. \quad (2.16)$$

Tenendo conto di quanto detto nei precedenti paragrafi, cerchiamo di risolvere questo problema partendo da una soluzione più generale dell'equazione di d'Alembert (2.1) espressa tramite una opportuna serie di onde stazionarie, cioè ¹

$$u(x, t) = \sum_{n=1}^{\infty} a_n u_n(x, t) = \sum_{n=1}^{\infty} a_n \sin(k_n x) \cos(\omega_n t + \delta_n) \quad (2.17)$$

dove $k_n = n\pi/L$, $\omega_n = vk_n$. Si riconosce subito che, per ogni fissato t , questa serie è proprio una serie di Fourier.

¹ Va sottolineato che il procedimento qui adottato è esatto, anche se richiederebbe ora una giustificazione più rigorosa. Il problema è essenzialmente quello di assicurarsi che è lecito eseguire le operazioni di derivazione della serie termine a termine. V. più avanti, i §§ 2.15, 2.28 e anche § 5.7 (i).

A questo punto basterà determinare i coefficienti incogniti a_n e δ_n che compaiono nella serie: e questo si ottiene imponendo che al tempo $t = 0$ siano soddisfatte le condizioni (2.16). Il procedimento generale, del resto facilmente deducibile, è meglio illustrato da un esempio semplice. Consideriamo il caso in cui nelle condizioni iniziali (2.16) la funzione g è nulla, mentre $f \neq 0$, cioè il caso in cui la corda viene spostata dalla sua posizione di equilibrio e quindi lasciata ripartire liberamente. Ponendo $t = 0$ nella derivata $u_t = \partial u / \partial t$ della (2.17) si ottiene

$$g(x) = \sum_{n=1}^{\infty} -a_n \omega_n \sin \delta_n \sin(k_n x)$$

che è lo sviluppo di Fourier (con coefficienti $\beta_n = -a_n \omega_n \sin \delta_n$) della funzione g . Poiché ora $g = 0$, si ha evidentemente $a_n \sin \delta_n = 0$ per ogni n e quindi si può porre $\delta_n = 0$ per tutti gli n per cui $a_n \neq 0$. Rimane quindi

$$u(x, t) = \sum_{n=1}^{\infty} a_n \sin(k_n x) \cos(\omega_n t) \quad (2.18)$$

Supponiamo ora di assegnare esplicitamente anche la funzione iniziale $f(x)$: sia per esempio (v. fig. 2.4)

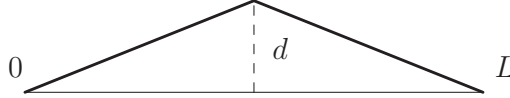


Figura 2.4 Una configurazione iniziale per la corda elastica.

$$f(x) = \begin{cases} \frac{2d}{L}x & \text{per } 0 \leq x \leq \frac{L}{2} \\ \frac{2d}{L}(L-x) & \text{per } \frac{L}{2} \leq x \leq L. \end{cases}$$

Imponendo nella (2.18) $f(x) = u(x, 0)$, i coefficienti a_n coincidono proprio con i coefficienti β_n dello sviluppo di Fourier della f (che – per quanto visto all'inizio di questo paragrafo – deve essere prolungata con *legge dispari* all'intervallo $(-L, L)$) in serie di soli seni (2.15). Essi quindi si calcolano facilmente e si ottiene:

$$a_n = \begin{cases} 0 & \text{per } n \text{ pari} \\ \frac{8d}{\pi^2 n^2} & \text{per } n = 1, 5, \dots, 4m+1, \dots \\ -\frac{8d}{\pi^2 n^2} & \text{per } n = 3, 7, \dots, 4m+3, \dots \end{cases}$$

Questo risultato intanto mostra che le lunghezze d'onda che vengono emesse dalla corda in questo caso sono i sottomultipli dispari della lunghezza $\lambda_1 = 2L$; il modo

di oscillazione con la lunghezza d'onda più alta (e quindi la frequenza più bassa) si chiama “armonica principale” o “fondamentale”, e le altre “armoniche secondarie”. Inoltre si vede che l'intensità con cui viene emessa l' n -sima armonica, essendo proporzionale al quadrato dell'ampiezza con cui tale armonica compare nello sviluppo (2.18), risulta proporzionale a $1/n^4$, e quindi decresce rapidamente all'aumentare di n . La (2.18), inserendovi gli a_n ora trovati, fornisce l'evoluzione temporale cercata. Ad esempio si verifica immediatamente che negli istanti $t = 0, 2L/v, 4L/v, \dots$, si ha $u(x, t) = u(x, 0) = f(x)$, cioè la corda riassume la sua posizione iniziale; mentre per $t = L/v, 3L/v, \dots$, si ha $u(x, t) = -u(x, 0) = -f(x)$.

Resta poi da notare che, *qualsiasi siano* le condizioni iniziali $f(x)$ e $g(x)$, la soluzione presenta una periodicità nel tempo con periodo $2L/v$, cioè $u(x, t + 2L/v) = u(x, t)$.

Una situazione simile è ottenuta, p.es., partendo dalle equazioni di Maxwell. Come è noto, queste equazioni per il campo elettrico associato a un'onda elettromagnetica nel vuoto conducono alle equazioni di d'Alembert in tre dimensioni

$$\Delta E_i = \frac{1}{c^2} \frac{\partial}{\partial t^2} E_i, \quad i = 1, 2, 3; \quad \Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

a cui deve obbedire ogni componente del campo elettrico stesso $\mathbf{E} = \mathbf{E}(x, y, z, t)$ (come ogni componente del campo magnetico). Supponiamo che il campo si trovi fra due pareti piane indefinite perfettamente riflettenti; assumendo opportunamente gli assi, siano $x = 0$ e $x = L$ le equazioni di questi piani e si assuma che il campo elettrico abbia componenti $\mathbf{E} = (0, 0, E)$, con E funzione solo di x , oltre che del tempo t ; la precedente equazione diventa allora

$$\frac{\partial^2 E}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} \quad E = E(x, t)$$

con le condizioni al contorno $E(0, t) = E(L, t) = 0$. Si ottiene quindi un problema identico a quello già studiato per le vibrazioni della corda elastica.

2.5 L'equazione del calore

Un'altra importante equazione alle derivate parziali è l'**equazione del calore** (o **della diffusione**, chiamata anche **equazione di Fourier**); si tratta di un'equazione che descrive tipicamente fenomeni fisici *irreversibili*. Per un esempio concreto, si consideri una sbarretta conduttrice di calore; se $u(x, t)$ indica la temperatura presente all'istante t nel punto x della sbarretta, si dimostra che l'equazione che regola l'andamento di tale temperatura è

$$\frac{\partial^2 u}{\partial x^2} = C \frac{\partial u}{\partial t} \quad (2.19)$$

dove $C > 0$ è una costante caratteristica del mezzo conduttore. Si supponga che gli estremi della sbarretta siano posti in $x = 0$ e $x = L$, e siano mantenuti a temperatura costante T_0 (si può sempre assumere $T_0 = 0$: basta sostituire nella (2.19) u con $u - T_0$). Se si esegue la separazione delle variabili, come nei §§ 2.1, 2.4, si trova, sostituendo nell'equazione, che $X_n(x) = \sin(\pi n x/L)$, esattamente come nei §§ 2.1, 2.4, mentre le $T_n(t)$ soddisfano l'equazione

$$\frac{dT_n}{dt} = -\frac{\pi^2 n^2}{CL^2} T_n,$$

da cui $T_n(t) \propto e^{-\alpha n^2 t}$ con $\alpha = \pi^2/CL^2$. La soluzione generale che si ottiene sotto forma di serie è quindi

$$u(x, t) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{\pi n}{L} x\right) e^{-\alpha n^2 t}. \quad (2.18')$$

Per determinare i coefficienti a_n (e quindi risolvere il problema di Cauchy) basta osservare che l'equazione contiene la derivata prima rispetto al tempo, e quindi l'evoluzione temporale risulta determinata da *una sola* condizione iniziale $f(x) = u(x, 0)$. Ponendo $t = 0$ nella (2.18') si vede che i coefficienti a_n sono, come nel paragrafo precedente, i coefficienti di Fourier della condizione iniziale $f(x)$ rispetto al set $X_n(x)$ (anche in questo caso, la $f(x)$ va prolungata con legge dispari all'intervallo $(-L, L)$). Per altre considerazioni, v. anche il § 2.15.

Anche l'equazione di Schrödinger per una particella quantistica vincolata sul segmento $(0, L)$ (buca di potenziale rettangolare infinita), che ha la forma

$$\frac{\partial^2 u}{\partial x^2} = -i \frac{\partial u}{\partial t} \quad u(0) = u(L) = 0 \quad (2.20)$$

porta a soluzioni analoghe alle precedenti (cambia radicalmente l'andamento temporale!).

Come già anticipato, i risultati di questo e del precedente paragrafo saranno in seguito opportunamente ripresi e studiati in un contesto più ampio e "adeguato".

2.6 Prodotto scalare e norma: definizione generale

Torniamo ora al problema, indicato nella seconda parte del § 2.3, di chiarire in un contesto preciso l'applicabilità della tecnica dello sviluppo in serie di Fourier, di cui abbiamo appena visto qualche utilizzazione interessante. Il punto di partenza sta proprio nella definizione di prodotto scalare (2.10) (oppure (2.10')), che in realtà è stata introdotta in modo del tutto naturale. Ma per poter utilizzare a questo proposito i risultati generali che ci fornisce la matematica, è utile ricordare alcune considerazioni "astratte" e quindi indipendenti dal particolare caso in esame.

Sia V un qualsiasi spazio vettoriale (di dimensione finita o infinita): si ricorda che in V è definito un **prodotto scalare** se è definita una legge di $V \times V \rightarrow \mathbf{C}$, che ad ogni coppia di vettori $x, y \in V$ associa un numero complesso che si indica con il simbolo (x, y) in modo che siano soddisfatte le seguenti proprietà per ogni $x, y, z \in V$ e $\lambda \in \mathbf{C}$:

- (p1) $(x, y)^* = (y, x)$
- (p2) $(x, x) \geq 0$ e inoltre $(x, x) = 0$ se e solo se $x = 0$
- (p3) $(x, \lambda y) = \lambda (x, y)$ e quindi: $(\lambda x, y) = \lambda^* (x, y)$
- (p4) $(x, y + z) = (x, y) + (x, z)$

(si è indicato semplicemente con 0 il vettore “zero” dello spazio V).

Si dice invece che in V è definita una **norma** se è definita una legge $V \rightarrow \mathbf{R}^+$ ($\mathbf{R}^+$ indica l'insieme dei numeri reali ≥ 0), che ad ogni $x \in V$ associa un numero reale positivo o nullo, che si indica con $\|x\|$, in modo che siano soddisfatte le seguenti proprietà, per ogni $x, y \in V$ e $\lambda \in \mathbf{C}$:

- (n1) $\|x\| \geq 0$ e inoltre $\|x\| = 0$ se e solo se $x = 0$
- (n2) $\|\lambda x\| = |\lambda| \|x\|$
- (n3) $\|x + y\| \leq \|x\| + \|y\|$.

Vogliamo ora mostrare che il prodotto scalare *implica* una norma: si ponga infatti

$$\|x\| = \sqrt{(x, x)}. \quad (2.21)$$

È immediato vedere che con tale definizione le proprietà (n1) e (n2) sono soddisfatte grazie alle (p1, p2, p3); per provare che vale anche la (n3) (detta **disuguaglianza triangolare**), si deve prima dimostrare l'importante **disuguaglianza di Schwarz**:

$$|(x, y)| \leq \|x\| \|y\| \quad (2.22)$$

che è conseguenza delle sole proprietà formali del prodotto scalare e della definizione (2.21). Per provare la (2.22) si osservi che, detto λ un qualsiasi numero *reale*, si ha ovviamente $\|x + \lambda (x, y) y\|^2 \geq 0$ e quindi, usando anche la (p4),

$$\|x + \lambda (x, y) y\|^2 = \|x\|^2 + 2\lambda |(x, y)|^2 + \lambda^2 |(x, y)|^2 \|y\|^2 \geq 0$$

che, dovendo essere vera per ogni λ , implica

$$|(x, y)|^4 - |(x, y)|^2 \|x\|^2 \|y\|^2 \leq 0$$

da cui segue la (2.22). Si ha allora

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

cioè appunto la disuguaglianza triangolare (n3). Questo prova che la posizione (2.21) è corretta. $\square$

Si può dimostrare che una norma può essere dedotta da un prodotto scalare *se e solo se* tale norma soddisfa la seguente **regola del parallelogramma**:

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

per ogni x, y .

Naturalmente, le definizioni date a suo tempo (1.3) e (1.4) di prodotto scalare e di norma per i vettori di $\mathbf{C}^n$ soddisfano a tutte le proprietà formali richieste sopra. Ma anche la posizione (2.10') soddisfa alle proprietà richieste² e quindi correttamente definisce un prodotto scalare. Resta allora definita anche una norma per le funzioni considerate, e sarà verificata la disuguaglianza di Schwarz, che in questo caso diventa

$$\left| \int_{-L}^L f^*(x) g(x) dx \right| \leq \sqrt{\int_{-L}^L |f(x)|^2 dx} \sqrt{\int_{-L}^L |g(x)|^2 dx}. \quad (2.23)$$

È anche possibile introdurre diverse definizioni di prodotto scalare: p.es. introducendo una *funzione peso* (o *densità*) $\rho = \rho(x) > 0$ si può definire

$$(f, g)_\rho = \int_I f^*(x) g(x) \rho(x) dx$$

che infatti soddisfa tutte le proprietà richieste per il prodotto scalare (e la norma da esso indotta).

Talvolta l'introduzione di un prodotto scalare modificato può essere un utile accorgimento: per esempio due funzioni non ortogonali rispetto al prodotto scalare ordinario possono risultare ortogonali rispetto ad un prodotto scalare opportunamente modificato: v. qualche esempio ai §§ 2.21, 2.23(b).

2.7 Il concetto di norma come “distanza”

Dati ad esempio due vettori x, y dello spazio $\mathbf{R}^3$, la quantità $\|x - y\|$ esprime, da un punto di vista geometrico, la “distanza” fra i due vettori. Preso un vettore x_0 , l'insieme dei vettori x tali che $\|x - x_0\| < \delta$ costituisce un “intorno sferico” di x_0 di raggio δ . Il concetto di norma fornisce dunque una struttura *metrica*, cioè un concetto di distanza fra gli elementi di uno spazio vettoriale, qualunque esso sia.

Bisogna però notare che in uno stesso spazio vettoriale V è possibile definire più tipi di norme, cioè varie leggi $V \rightarrow \mathbf{R}^+$ che soddisfano i requisiti (n1, n2, n3) assegnati nel paragrafo precedente. Per esempio in $\mathbf{C}^n$, oltre alla norma già usata (1.4)

² Ciò è senz'altro vero per funzioni *continue*; in caso contrario occorre una precisazione che riguarda le proprietà (p2) (ed (n1)), e che faremo nel § 2.8.

che ha la proprietà di essere indotta da un prodotto scalare, anche la seguente legge

$$x \rightarrow |x_1| + |x_2| + \cdots + |x_n|$$

definisce una norma, diversa dalla precedente (e non ottenibile da un prodotto scalare).

Analogamente, nello spazio vettoriale delle funzioni considerate nel § 2.3 a proposito della serie di Fourier, oltre alla norma

$$\|f\| = \left[\int_I |f(x)|^2 dx \right]^{1/2} \quad (2.24)$$

indotta dal prodotto scalare (2.10'), si può considerare anche la norma

$$\|f\|' = \sup_{x \in I} |f(x)| \quad (2.25)$$

spesso indicata con $\|f\|_\infty$, dove I è l'intervallo in cui varia la x . È importante sot-

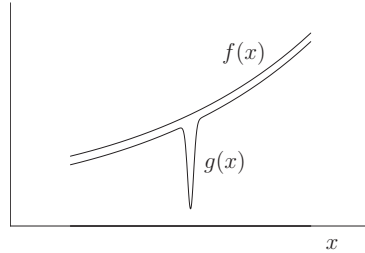


Figura 2.5 Funzioni “lontane” nella norma del *sup*, “vicine” nella norma L^1 (2.24') o L^2 (2.24).

tolineare che, come mostra la fig. 2.5, due funzioni f , g possono essere “vicine” rispetto alla prima norma

$$\|f - g\| = \left[\int_I |f(x) - g(x)|^2 dx \right]^{1/2}$$

ed essere invece “lontane” rispetto alla seconda

$$\|f - g\|' = \sup_{x \in I} |f(x) - g(x)|.$$

È a questo punto essenziale osservare che in generale la serie di Fourier di una funzione *non* è “vicina” alla funzione stessa nel senso della norma (2.25), che – come noto – individua la *convergenza uniforme*; più esattamente, la successione delle somme parziali della serie di Fourier *non approssima bene* nel senso di tale norma la funzione data: basta infatti ricordare quanto succede nei punti di discontinuità della funzione stessa. Come enunceremo in modo completo e preciso più avanti, si ha invece che la serie di Fourier approssima veramente la funzione nel senso della

norma (2.24), che d'altronde è proprio quella che segue dal prodotto scalare (2.10') introdotto in modo del tutto naturale nel calcolo dei coefficienti della serie, come si è visto. Tutto questo suggerisce che la giusta “ambientazione” della teoria delle serie di Fourier è quella che coinvolge le funzioni per le quali è definibile la norma (2.24), cioè le funzioni con modulo quadrato integrabile.

Accanto alla norma (2.24), citiamo anche un'altra scelta utile (che incontreremo più avanti) e che indichiamo qui con $\|f\|_1$:

$$\|f\|_1 = \int_I |f(x)| dx. \quad (2.24')$$

Un altro importante concetto, legato a queste considerazioni, è quello di **spazio completo** (o **chiuso**). Ricordiamo intanto che in uno spazio V in cui è definita una norma, una successione di elementi $x_n \in V$ si dice **successione di Cauchy** se $\|x_n - x_m\| \rightarrow 0$; lo spazio si dirà completo o chiuso (rispetto alla data norma) se ogni successione di Cauchy converge ad un elemento x dello spazio stesso (sempre nel senso della norma: $\|x_n - x\| \rightarrow 0$). L'esempio più ovvio di spazio *non* completo è dato dall'insieme dei numeri razionali, con la norma “naturale” $|x_1 - x_2|$: è ben noto infatti che ci sono successioni di Cauchy di numeri razionali che convergono ad un numero *irrazionale*. Invece, la proprietà caratteristica dei numeri reali è appunto quella di costituire uno spazio completo, o – meglio ancora – di essere il **completamento** (o **chiusura**) dello spazio dei razionali.

Consideriamo ora lo spazio delle funzioni *continue* definite in un intervallo compatto qualsiasi I , con la norma (2.25), che ora si può anzi scrivere

$$\|f\| = \max_{x \in I} |f(x)|. \quad (2.25')$$

Questo spazio è completo, infatti una successione di funzioni continue che sia successione di Cauchy rispetto alla norma (2.25') è una successione *uniformemente* convergente, e pertanto – come è noto e altrove già ricordato – essa converge (uniformemente) ad una funzione continua e quindi contenuta nello spazio. Invece, lo *stesso* spazio di funzioni però dotato della norma (2.24) *non* è completo, e nemmeno lo è con la norma (2.24'): è facile trovare esempi di successioni di Cauchy rispetto a queste norme che però convergono (puntualmente e nel senso della norma (2.24) o della (2.24')) a funzioni discontinue; p.es. la successione di funzioni continue ($n = 1, 2, \dots$)

$$f_n(x) = \begin{cases} -1 & \text{per } -1 \leq x \leq -1/n \\ nx & \text{per } -1/n \leq x \leq 1/n \\ 1 & \text{per } 1/n \leq x \leq 1 \end{cases}$$

converge puntualmente e in senso (2.24') ad una funzione discontinua.

Ricordiamo infine la seguente proprietà generale: se una successione x_n è convergente, cioè $\|x_n - x\| \rightarrow 0$, allora essa è di Cauchy. Infatti, per la disuguaglianza triangolare, si ha

$$\|x_n - x_m\| = \|x_n - x + x - x_m\| \leq \|x_n - x\| + \|x - x_m\|.$$

2.8 Alcune osservazioni sulla integrazione delle funzioni

Le considerazioni fatte nei paragrafi precedenti ci conducono in modo naturale ad introdurre una *norma* da definire attraverso un integrale come nella (2.24). Ma si deve osservare che, affinché la (2.24) definisca una vera norma, cioè soddisfi in particolare al requisito (n1) (v. § 2.6), occorre anzitutto che $\|f\| = 0$ implichi $f = 0$. In realtà, invece, per una funzione che sia diversa da zero solo in un numero finito o numerabile di punti, l'integrale (2.24) risulta uguale a zero. Per poter allora considerare la definizione (2.24) come una vera norma, basta *convenire* di considerare *equivalenti* (e cioè di “identificare”) le funzioni **quasi ovunque** uguali fra loro, ovvero le funzioni che differiscono fra loro soltanto in un insieme finito o numerabile di punti (o – come si dice più esattamente – in un insieme di punti di *misura nulla* secondo Lebesgue).

Nel seguito questa convenzione verrà sempre *sottintesa*. Naturalmente identiche osservazioni valgono anche per la norma (2.24').

Come si vede, la nozione di integrale a cui occorre far riferimento in questo contesto è quella dell'*integrale secondo Lebesgue*. Come noto, l'integrale di Lebesgue è una estensione dell'*integrale elementare*³ (quello dei plurirettangoli inscritti e circoscritti, o di Riemann), nel senso che esso permette di definire l'integrale *anche* per alcune funzioni che *non* risultano integrabili secondo il metodo elementare. L'esempio tipico è dato dalla funzione di Dirichlet f_D , così definita nell'intervallo $[0, 1]$:

$$f_D(x) = \begin{cases} 0 & \text{se } x \text{ è irrazionale} \\ 1 & \text{se } x \text{ è razionale} \end{cases}$$

che *non* è integrabile nel senso dell'integrazione elementare, ma è integrabile secondo Lebesgue, e in particolare ha integrale nullo: in effetti essa differisce dalla funzione nulla soltanto in un insieme numerabile di punti (si ricordi che l'insieme dei numeri razionali è *numerabile*), e dunque è da considerare equivalente alla funzione nulla.

Alcune importanti proprietà dell'integrale di Lebesgue che ci serviranno nel seguito sono le seguenti:

- (a) Se $F(x)$ è una funzione integrabile, ed $f(x)$ un'altra funzione tale che $|f(x)| \leq F(x)$ allora anche f è integrabile.
- (b) f è integrabile se e solo se $|f|$ è integrabile.

È da osservare che tali proprietà sono in generale false se riferite all'integrale elementare (per esempio, per la funzione di Dirichlet f_D si ha $f_D = |f_D| \leq F = 1$ in $[0, 1]$); tuttavia tornano ad essere vere se ci si limita alle funzioni *f generalmente continue*.

³ Per maggiori dettagli, v. per esempio i testi di Rudin [30], di Royden [29] o di Fano [3]. In particolare, dovremo sempre sottintendere – come d'abitudine – che le funzioni che verranno considerate siano “misurabili” secondo Lebesgue. Naturalmente, questa non è una grave limitazione, dato che gli esempi che si possono costruire di insiemi o funzioni non misurabili sono piuttosto “astrusi”, non solo dal punto di vista fisico, ma anche da quello matematico, poiché per costruire questi esempi occorre utilizzare l'*assioma della scelta*.

In realtà, occorre osservare che la proprietà (b) risulta vera qui “per definizione”: infatti la *definizione* stessa di integrabilità secondo Lebesgue richiede che siano separatamente integrabili la “parte positiva”

$$f^+(x) = \begin{cases} f(x) & \text{se } f(x) > 0 \\ 0 & \text{se } f(x) < 0 \end{cases}$$

e la “parte negativa” $f^-(x)$ (definita analogamente) della $f(x)$ ⁴, dunque è $f = f^+ - f^-$ e $|f| = f^+ + f^-$, e la proprietà (b) segue automaticamente.

È tuttavia importante far presente che è possibile introdurre una nozione “meno restrittiva” di integrabilità: consideriamo, ad esempio, la funzione (definita e continua su tutto $\mathbf{R}$)

$$f(x) = \frac{\sin x}{x}$$

si vede, anche solo osservandone il grafico, che l'integrale $\int_{-\infty}^{+\infty} |f(x)| dx$ si comporta come la serie armonica $\sum_n (1/n)$ e dunque diverge, mentre esiste finito l'integrale definito mediante il seguente limite (ora il comportamento è all'incirca come la serie armonica a segni alternati):

$$\lim_{\substack{R \rightarrow +\infty \\ R' \rightarrow +\infty}} \int_{-R'}^{+R} f(x) dx \quad (2.26)$$

dove i limiti per $R \rightarrow \infty$ e per $R' \rightarrow \infty$ vanno eseguiti *indipendentemente*.

La (2.26) definisce il cosiddetto **integrale improprio** della funzione $f(x)$.

Per sottolineare la differenza fra la nozione di integrabilità secondo Lebesgue e quella più “debole” secondo questa nozione di integrale improprio, le funzioni integrabili secondo Lebesgue si dicono anche **sommabili**. Dunque $f(x) = \frac{\sin x}{x}$ ammette integrale improprio su $\mathbf{R}$, ma non è sommabile. Nel § 3.13 calcoleremo questo integrale con le tecniche della variabile complessa.

L'insieme delle funzioni sommabili su un insieme I si indica con $L^1(I)$, e dunque la (2.24') è precisamente la definizione della *norma* $L^1(I)$.

La proprietà rilevante dello spazio $L^1(I)$ è che esso è *completo* rispetto a tale norma (**teorema di Riesz**), ed anzi è il *completamento* (o *chiusura*), sempre rispetto a tale norma, del sottospazio $C^0(I)$ delle funzioni continue.

Infine, particolarmente importante, anche nelle applicazioni, è il seguente

Teorema di Lebesgue o Teorema della convergenza dominata.

Sia $f_n(x)$ una successione di funzioni sommabili che converge puntualmente quasi ovunque per $n \rightarrow \infty$ ad una funzione $f(x)$; se esiste una funzione $F(x)$ sommabile, tale che, per ogni n , sia

$$|f_n(x)| \leq F(x), \quad (2.27)$$

⁴ Se f è complessa, lo stesso procedimento si applica alla parte reale e alla parte immaginaria della f .

allora:

(i) anche f è sommabile;

(ii) si ha

$$\lim_{n \rightarrow \infty} \int_I f_n(x) dx = \int_I f(x) dx \quad (2.28)$$

(analogo risultato vale anche se – invece di una successione $f_n(x)$ – si ha una famiglia $f_\varepsilon(x)$ dipendente da un parametro continuo $\varepsilon \rightarrow 0$).

Questo teorema costituisce una importante generalizzazione del noto risultato dell'analisi elementare, che assicura la possibilità di passare al limite sotto l'integrale come nella (2.28) nelle ipotesi che la successione f_n converga *uniformemente* in un intervallo di lunghezza *finita*.

Un paio di controesempi per mostrare che se qualcuna di queste ipotesi non è verificata, allora la possibilità di passare al limite per $n \rightarrow \infty$ sotto l'integrale non è garantita:

i) Se

$$f_n(x) = \begin{cases} n \sin nx & \text{per } 0 \leq x \leq \pi/n \\ 0 & \text{per } \pi/n \leq x \leq \pi \end{cases} \quad (n = 1, 2, \dots)$$

si ha che $f_n(x) \rightarrow 0$ puntualmente per ogni $x \in [0, \pi]$, ma $\int_0^\pi f_n(x) dx = 2$.

ii) Se

$$f_n(x) = \begin{cases} 1/n & \text{per } -n \leq x \leq n \\ 0 & \text{per } |x| \geq n \end{cases} \quad (n = 1, 2, \dots)$$

si ha che $f_n(x) \rightarrow 0$ uniformemente per ogni $x \in \mathbf{R}$, ma $\int_{\mathbf{R}} f_n(x) dx = 2$.

Ovviamente, nessuna di queste due successioni può essere maggiorata da una funzione $F(x)$ sommabile.

2.9 Lo spazio $L^2(I)$

Sia I un qualsiasi intervallo o sottoinsieme di $\mathbf{R}$ (misurabile, eventualmente anche di misura infinita, per esempio l'intera retta); si indica con $L^p(I)$ l'insieme delle funzioni f tali che esiste finito l'integrale

$$\int_I |f(x)|^p dx.$$

Ovviamente, si possono considerare anche funzioni in $\mathbf{R}^n$. Per semplicità, ci limiteremo in generale al caso $n = 1$; l'estensione a funzioni di $n > 1$ variabili non presenta difficoltà.

Come si è visto precedentemente, l'insieme delle funzioni *a quadrato sommabile* $L^2(I)$ merita particolare attenzione. Proviamo intanto che L^2 è uno spazio vettoriale: siano infatti $f, g \in L^2(I)$, cioè

$$\int_I |f|^2 dx < +\infty, \quad \int_I |g|^2 dx < +\infty;$$

osservando che

$$|f+g|^2 \leq |f+g|^2 + |f-g|^2 = 2|f|^2 + 2|g|^2$$

segue, per la proprietà (a) del paragrafo precedente, che anche $f+g \in L^2(I)$. È immediato poi osservare che anche $\alpha f \in L^2(I)$ per ogni numero complesso α . Nello spazio $L^2(I)$ si può inoltre definire un prodotto scalare nel modo già visto

$$(f, g) = \int_I f^*(x) g(x) dx \quad (2.29)$$

e l'integrale esiste finito per qualunque coppia di funzioni $f, g \in L^2(I)$ grazie alle proprietà (a), (b) del paragrafo precedente. Infatti

$$|f^* g| = |f| |g| \leq \frac{1}{2} |f|^2 + \frac{1}{2} |g|^2$$

dove si è fatto uso della ovvia disuguaglianza

$$0 \leq (|f| - |g|)^2 = |f|^2 + |g|^2 - 2|f||g|$$

dunque $f^* g \in L^1(I)$ (e così anche ovviamente fg e $|f|^2$).

Si può notare che se I è un intervallo finito, allora se $f \in L^2(I)$ si ha anche $f \in L^1(I)$, infatti

$$\|f\|_{L^1} = \int_I |f(x)| dx = (1, |f|) \leq \text{mis}(I) \|f\|_{L^2}$$

ma non è vero il viceversa, p.es. $f = 1/\sqrt{x} \in L^1(0, 1)$ ma $\notin L^2(0, 1)$. Se I ha misura infinita, $f \in L^2$ non implica $f \in L^1$: p.es. $f(x) = x/(1+x^2) \in L^2(\mathbf{R})$ ma $\notin L^1(\mathbf{R})$.

Vale inoltre il seguente

Teorema di Riesz-Fischer. *Lo spazio $L^2(I)$ è completo rispetto alla norma*

$$\|f\|^2 = \int_I |f|^2 dx \quad (2.30)$$

indotta dal prodotto scalare (2.29).

Inoltre $L^2(I)$ è il *completamento*, rispetto alla norma (2.30), del sottospazio delle funzioni continue $C^0(I)$. Naturalmente, la convergenza cui ora ci si riferisce è quella *in media di ordine due*, come già osservato, cioè

$$\|f_n - f\|^2 = \int_I |f_n - f|^2 dx \rightarrow 0.$$

Enunciamo, infine, una generalizzazione della disuguaglianza di Schwarz e cioè la cosiddetta

Disuguaglianza di Schwarz-Hölder: *Siano $f \in L^p(I)$ e $g \in L^q(I)$, dove p e q sono numeri tali che $1/p + 1/q = 1$. Allora il prodotto fg è sommabile e vale la*

disuguaglianza

$$\int_I |f g| dx \leq \left[\int_I |f|^p dx \right]^{1/p} \left[\int_I |g|^q dx \right]^{1/q}. \quad (2.31)$$

2.10 Lo spazio di Hilbert: definizione generale. Teorema di Fourier in L^2

Uno **spazio di Hilbert** è uno spazio vettoriale in cui è definito un prodotto scalare e che è completo rispetto alla norma indotta da tale prodotto scalare⁵.

Quindi gli spazi $L^2(I)$ sono spazi di Hilbert. Ma anche gli spazi a dimensione finita $\mathbf{C}^n$ (o $\mathbf{R}^n$) sono spazi di Hilbert, dato che sono automaticamente spazi completi (usualmente però, quando si parla di spazi di Hilbert, ci si riferisce a spazi di dimensione infinita).

Consideriamo ora lo spazio $L^2(I)$ nel caso che I sia l'intervallo $(-L, L)$. Osserviamo intanto che l'insieme delle funzioni che soddisfano i requisiti dati all'inizio del § 2.3, e per le quali avevamo introdotto lo sviluppo in serie di Fourier, è ovviamente un sottospazio delle funzioni a quadrato sommabile $L^2(-L, L)$. Siamo ora in grado di generalizzare questo risultato con la seguente formulazione precisa del classico teorema sulla sviluppabilità in serie (trigonometrica) in L^2 :

Teorema di Fourier. Sia $f = f(x) \in L^2(-L, L)$ e siano a_n i prodotti scalari $a_n = (f_n, f)$ fra $f(x)$ e le funzioni $f_n(x)$ del set (2.11); allora la serie (2.12) che così si costruisce risulta convergente alla funzione f nella norma di L^2 (cioè la successione delle somme parziali approssima la funzione f “in media di ordine due”).

Abbiamo dunque *ampliato* (rispetto a quanto visto in § 2.3) lo spazio delle funzioni per le quali è possibile calcolare lo sviluppo in serie tramite funzioni trigonometriche (2.11), *confermato* le formule per il calcolo dei coefficienti di Fourier a_n , e soprattutto abbiamo *precisato* qual è il tipo di convergenza che è corretto richiedere.

Grazie alla proprietà descritta da questo teorema, il sistema (2.11) viene detto **sistema completo** in $L^2(-L, L)$ (*si noti bene* che, accanto alla parola “sistema”, l'aggettivo “completo” porta al concetto di “base”, ben diverso dal significato che ha accanto alla parola “spazio”, v. § 2.7, dove ammette come sinonimo “chiuso”).

Ma è anche chiaro che questa fondamentale proprietà del set (2.11) deve essere “formalizzata”, allo scopo di estenderla ad altri sistemi, e a spazi di Hilbert qualsiasi. Svilupperemo dunque nei prossimi paragrafi una teoria fondata solo sulle proprietà “astratte” dello spazio di Hilbert, che ci consentirà di introdurre una nozione generale e astratta di sistema completo e di serie di Fourier di vettori.

⁵ Se manca la completezza, si parla di spazio **pre-Hilbertiano**; se invece lo spazio è completo, ma la norma non proviene da un prodotto scalare, si ha uno **spazio di Banach**. Tutti gli spazi L^p sono spazi di Banach, ma solo L^2 è spazio di Hilbert.

2.11 Sistemi indipendenti e ortonormali

Sia H un generico spazio di Hilbert e sia $x_1, x_2, \dots, x_n, \dots$ un insieme (finito o infinito) di vettori linearmente indipendenti contenuti in H . Se lo spazio ha dimensione infinita (ed è il caso che qui ci interessa), è evidentemente possibile trovare infiniti vettori indipendenti⁶.

Può essere utile il seguente criterio, che sarebbe facile dimostrare: condizione necessaria e sufficiente affinché gli n vettori $x_1, x_2, \dots, x_n$ siano linearmente indipendenti è che il determinante

$$G(x_1, x_2, \dots, x_n) = \det \begin{pmatrix} (x_1, x_1) & (x_1, x_2) & \cdots & (x_1, x_n) \\ (x_2, x_1) & (x_2, x_2) & \cdots & \\ \vdots & & \ddots & \\ (x_n, x_1) & \cdots & & (x_n, x_n) \end{pmatrix}$$

detto **determinante di Gram**, sia diverso da zero.

Per esempio, vettori ortogonali sono (ovviamente!) indipendenti.

Si ha il seguente importante risultato: comunque sia dato un insieme $x_1, x_2, \dots, x_n, \dots$ di vettori indipendenti, si può sempre ricavare un altro insieme $e_1, e_2, \dots, e_n, \dots$ di vettori *ortonormali* in modo che, per ogni m , i vettori $e_1, e_2, \dots, e_m$ siano combinazioni lineari dei soli primi m vettori $x_1, x_2, \dots, x_m$.

Per costruire tale *sistema ortonormale* si può seguire il procedimento di **ortonormalizzazione di Schmidt**: si comincia ponendo

$$e_1 = \frac{x_1}{\|x_1\|}.$$

Si sottrae poi ad x_2 la componente “parallela” a e_1 , cioè

$$e'_2 = x_2 - (e_1, x_2)e_1$$

e si pone

$$e_2 = \frac{e'_2}{\|e'_2\|}.$$

I vettori e_1 ed e_2 sono ortonormali, come si verifica subito. Il procedimento si ripete analogamente: il termine generale è

$$e'_k = x_k - \sum_{i=1}^{k-1} (e_i, x_k) e_i, \quad e_k = \frac{e'_k}{\|e'_k\|}$$

(si noti che $e'_k \neq 0$, altrimenti $x_1, x_2, \dots, x_k$ sarebbero linearmente dipendenti).

⁶ Per l'esattezza, parlare di un insieme infinito (numerabile) di vettori indipendenti significa questo: comunque si fissi un intero n , i primi vettori $x_1, x_2, \dots, x_n$ sono linearmente indipendenti fra loro.

2.12 Serie di Fourier “astratta”

Sia $\{e_n\}$ un qualsiasi sistema *ortonormale* in uno spazio di Hilbert “astratto” H . Fissato comunque $x \in H$ ed un intero n , vogliamo determinare n numeri complessi $a_1, a_2, \dots, a_n$ in modo tale che la “distanza”

$$\left\| x - \sum_{i=1}^n a_i e_i \right\|$$

sia minima. Cioè vogliamo trovare quale combinazione lineare dei primi n vettori “approssima” meglio il vettore x . Si ha

$$\begin{aligned} 0 \leq \left\| x - \sum_{i=1}^n a_i e_i \right\|^2 &= \|x\|^2 - \sum_{i=1}^n a_i (x, e_i) - \sum_{i=1}^n a_i^* (e_i, x) + \sum_{i=1}^n |a_i|^2 = \\ &= \|x\|^2 - \sum_{i=1}^n |(e_i, x)|^2 + \sum_{i=1}^n |a_i - (e_i, x)|^2 \end{aligned} \quad (2.32)$$

e tale quantità è chiaramente minima quando

$$a_i = (e_i, x) \quad (2.33)$$

cioè scegliendo come a_i proprio la “componente” di x sul vettore e_i .

In questo modo, con la (2.33), il vettore $\sum_{i=1}^n a_i e_i$ rappresenta quella che si può chiamare la proiezione ortogonale di x sul sottospazio generato da $e_1, e_2, \dots, e_n$. Sempre con la scelta (2.33) e usando $\|x - \sum a_i e_i\|^2 \geq 0$, la (2.32) dà ancora

$$\|x\|^2 \geq \sum_{i=1}^n |(e_i, x)|^2 = \sum_{i=1}^n |a_i|^2$$

che è verificata *per ogni* n e coinvolge solo quantità reali positive, con il primo membro indipendente da n . Quindi si può prendere il limite $n \rightarrow \infty$ al secondo (e terzo) membro, la serie converge e si ottiene

$$\|x\|^2 \geq \sum_{i=1}^{\infty} |(e_i, x)|^2 = \sum_{i=1}^{\infty} |a_i|^2 \quad (2.34)$$

che si chiama **disuguaglianza di Bessel**.

I coefficienti a_i definiti dalla (2.33) si chiamano **coefficienti di Fourier** del vettore x *rispetto al sistema* $\{e_i\}$: la ragione di questa denominazione è chiara grazie, ad esempio, all’analogia con la (2.14).

Si consideri ora la seguente “serie di vettori”

$$\sum_{i=1}^{\infty} a_i e_i = \sum_{i=1}^{\infty} (x, e_i) e_i$$

costruita con i coefficienti di Fourier del vettore x , che si chiama **serie di Fourier** del vettore x rispetto al sistema $\{e_i\}$: naturalmente “sommare” questa serie vuol dire cercare il limite $\bar{x}$ (beninteso: nel senso della norma assegnata in H) della successione delle somme parziali

$$\bar{x}_n = \sum_{i=1}^n a_i e_i$$

e l’esistenza di questo limite è assicurata dal fatto che la successione $\bar{x}_n$ è di Cauchy: infatti (se $n > m$)

$$\|\bar{x}_n - \bar{x}_m\|^2 = \left\| \sum_{i=m+1}^n a_i e_i \right\|^2 = \sum_{i=m+1}^n |a_i|^2 \rightarrow 0$$

grazie alla convergenza della serie numerica $\sum |a_i|^2$ (v. la (2.34)).

Il vettore $\bar{x}$ sarà dunque la *migliore approssimazione* di x mediante una serie degli $\{e_i\}$, ma *non* coinciderà in generale col vettore di partenza x . Ciò dipenderà da come è stato scelto all’inizio il sistema ortonormale $\{e_i\}$: ad esempio, se prendiamo $H = L^2(I)$ dove I è l’intervallo $(-L, L)$ e come sistema ortonormale l’insieme $\{L^{-1/2} \sin(\pi n x / L)\}$, è chiaro che lo sviluppo in serie di Fourier di soli seni di una funzione $f \in L^2(-L, L)$ convergerà alla funzione stessa *solo* se essa è dispari, mentre ne darà in generale la sola “parte dispari”⁷ $\frac{1}{2}[f(x) - f(-x)]$, e addirittura darà zero se la funzione è pari.

Come altro semplice esempio, sia, in $L^2(-1, 1)$,

$$f(x) = \begin{cases} 0 & \text{per } -1 < x < 0 \\ 1 & \text{per } 0 < x < 1 \end{cases} :$$

se si prende il set ortonormale $\{1/\sqrt{2}, \sin n\pi x \ (n = 1, 2, \dots)\}$, lo sviluppo di questa $f(x)$ coincide con la funzione stessa; scegliendo invece il set ortonormale $\{1/\sqrt{2}, \cos n\pi x \ (n = 1, 2, \dots)\}$, la migliore approssimazione $\bar{f}(x)$ della $f(x)$ sarà la funzione costante $\bar{f}(x) = 1/2$ (non occorrono calcoli: basta osservare che $f(x) = 1/2 +$ una funzione dispari!).

Tornando al caso generale, diremo che $\{e_i\}$ è un **sistema** (o **set**) **ortonormale completo** in H se, per ogni $x \in H$, la serie di Fourier di x rispetto a tale sistema converge nella norma di H ad x stesso, cioè se $\|\bar{x}_n - x\| \rightarrow 0$, e si scriverà

$$x = \sum_{i=1}^{\infty} a_i e_i, \quad a_i = (e_i, x). \quad (2.35)$$

I sistemi completi generalizzano dunque, in dimensione infinita, il concetto di **base** per gli spazi a dimensione finita.

⁷ Ogni funzione f definita in un intervallo simmetrico rispetto all’origine si può scomporre nella somma di una funzione pari f_p e di una dispari f_d : $f(x) = \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)] \equiv f_p + f_d$.

Come si è già detto, grazie al teorema di Fourier (§ 2.10), il set (2.11) è un esempio di set ortonormale completo per lo spazio $H = L^2(-L, L)$. Dobbiamo ora vedere altre importanti e utili proprietà *caratterizzanti* i sistemi completi.

2.13 Proprietà dei sistemi completi

Dimostriamo ora che la completezza di un sistema *ortonormale* $\{e_n\}$ è equivalente ad altre proprietà, che spesso risulteranno più agevoli a verificare negli esempi concreti. Precisamente proveremo che le seguenti quattro proprietà sono *equivalenti* fra loro (cioè basta verificarne una sola perché siano verificate anche le altre).

(i) Il sistema $\{e_n\}$ è completo secondo la definizione del paragrafo precedente, cioè – in breve – ogni vettore $x \in H$ coincide con la sua serie di Fourier (2.35) rispetto a tale sistema.

(ii) L'insieme delle combinazioni lineari finite

$$\sum_{i=1}^n a_i e_i$$

al variare di n e dei coefficienti a_i è denso in H , cioè ogni vettore $x \in H$ può essere “approssimato” in norma quanto si vuole da una combinazione lineare finita degli $\{e_n\}$. Fra queste combinazioni, quella che approssima meglio ogni prefissato vettore $x \in H$ è quella i cui coefficienti a_i sono dati da $a_i = (e_i, x)$.

(iii) Per ogni $x \in H$ vale la seguente identità:

$$\|x\|^2 = \sum_{n=1}^{\infty} |(e_n, x)|^2 \quad (2.36)$$

detta **identità di Parseval**.

(iv) L'unico vettore ortogonale a tutti gli e_n è il vettore nullo, cioè: sia x_0 un vettore tale che per ogni n si abbia $(e_n, x_0) = 0$, allora $x_0 = 0$.

Dimostrazione:

(i) $\Rightarrow$ (ii): In realtà la (ii) è un altro modo di enunciare la (i). Infatti, se il sistema è completo, per ogni $x \in H$, la somma parziale $\bar{x}_n$ della sua serie di Fourier approssima in norma x quanto si vuole, e la migliore approssimazione è proprio quella che si ottiene con $a_i = (e_i, x)$.

(ii) $\Rightarrow$ (iii): Ricordando la (2.32) con la scelta (2.33), si ha per $n \rightarrow \infty$ che il primo membro della (2.32) tende a zero, e quindi, direttamente dal secondo membro, segue la (2.36).

(iii) $\Rightarrow$ (iv): La (2.36) applicata al vettore x_0 dà $x_0 = 0$, infatti

$$\|x_0\|^2 = \sum_{n=1}^{\infty} |(e_n, x_0)|^2 = 0.$$

(iv) $\Rightarrow$ (i): Fissato comunque $x \in H$, si consideri

$$\bar{x} = \sum_{n=1}^{\infty} (e_n, x) e_n.$$

Come già visto, tale serie è convergente e $\bar{x} \in H$. Per ogni intero j si ha

$$(e_j, \bar{x} - x) = \left(e_j, \sum_{i=1}^{\infty} (e_i, x) e_i \right) - (e_j, x) = \sum_{i=1}^{\infty} (e_i, x) \delta_{ji} - (e_j, x) = 0$$

cioè $\bar{x} - x$ è ortogonale a tutti gli e_j , quindi $\bar{x} = x$ e cioè x coincide con la sua serie di Fourier. $\square$

Nota. Nella dimostrazione precedente si è scritto

$$\left(e_j, \sum_{i=1}^{\infty} \dots \right) = \sum_{i=1}^{\infty} (e_j, \dots)$$

“scambiando” il prodotto scalare con la serie. Poiché la linearità del prodotto scalare vale – per definizione! – per le combinazioni lineari *finite*, occorre provare che il passaggio sopra eseguito è lecito. Si ha in effetti:

Lemma (*continuità del prodotto scalare*): Sia x_n una successione di vettori convergente ad $x \in H$, cioè $x_n \rightarrow x$ ovvero $\|x_n - x\| \rightarrow 0$, allora, per ogni $y \in H$, si ha

$$(y, x) = \lim_{n \rightarrow \infty} (y, x_n). \quad (2.37)$$

Dimostrazione: Immediata: infatti, usando la disuguaglianza di Schwarz

$$|(y, x) - (y, x_n)| = |(y, x - x_n)| \leq \|y\| \|x - x_n\| \rightarrow 0. \quad \square$$

Questo giustifica la dimostrazione precedente: basta infatti prendere come successione x_n la successione delle somme parziali della serie di Fourier di x .

Questo Lemma fornisce pure un'altra condizione di “passaggio al limite sotto integrale” in $H = L^2(I)$: se $g, f_n \in L^2(I)$ e $f_n \rightarrow f$ in senso $L^2(I)$, ossia $\|f_n - f\|_{L^2} \rightarrow 0$, la (2.37) diventa

$$\int_I g^* f_n dx \rightarrow \int_I g^* f dx.$$

Se in particolare I ha misura finita si può prendere $g = 1$ e si ottiene

$$\int_I f_n dx \rightarrow \int_I f dx.$$

Inoltre, il Lemma giustifica la possibilità di scambiare l'integrazione con la serie, cioè di integrare termine a termine (v. § 2.3) una serie di Fourier come la (2.7): è infatti sufficiente *la sola ipotesi della convergenza in norma L^2* .

Si considerino ora due vettori $x, y \in H$, con le rispettive serie di Fourier rispetto al set ortonormale completo $\{e_n\}$:

$$x = \sum_{i=1}^{\infty} a_i e_i, \quad y = \sum_{i=1}^{\infty} b_i e_i.$$

Usando il precedente Lemma, si ha

$$\begin{aligned} (x, y) &= \lim_{n \rightarrow \infty} \left(\sum_{i=1}^n a_i e_i, y \right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i^* (e_i, y) = \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i^* b_i = \sum_{i=1}^{\infty} a_i^* b_i = \sum_{i=1}^{\infty} (x, e_i) (e_i, y) \end{aligned} \quad (2.36')$$

che si chiama **identità di Parseval generalizzata** (infatti la (2.36) si ottiene da questa ponendo $x = y$). Pertanto, anche tale identità è una proprietà equivalente alla completezza del sistema ortonormale $\{e_n\}$.

Un esempio (molto istruttivo!) di set ortonormale completo $\{f_n\}$ è

$$\begin{aligned} f_1 &= (e_1 - e_2)/\sqrt{2} \\ f_2 &= (e_1 + e_2 - 2e_3)/\sqrt{6} \\ &\vdots \\ f_n &= (e_1 + e_2 + \dots + e_n - n e_{n+1})/\sqrt{n(n+1)} \\ &\vdots \end{aligned} \quad (2.38)$$

infatti la condizione (iv), cioè $(f_n, x_0) = 0, \forall n$, implica che i coefficienti di Fourier a_i rispetto al set e_i del vettore $x_0 \in H$ devono essere tutti uguali, e dunque tutti nulli dovendo $\sum_i |a_i|^2 < \infty$. Le combinazioni lineari finite $\sum_{n=1}^m \beta_n f_n$ dei vettori del set f_n hanno chiaramente la proprietà $\sum_i a_i = 0$. Dunque, grazie alla prima parte della condizione (ii), il sottospazio delle combinazioni finite $z = \sum_{i=1}^n a_i e_i$ tali che $\sum_{i=1}^n a_i = 0$ è – sorprendentemente – *denso* in H ; a maggior ragione, sarà denso in H il sottospazio più generale $V^{(0)}$ dei vettori

$$z = \sum_{i=1}^{\infty} a_i e_i \quad \text{tali che} \quad \sum_{i=1}^{\infty} a_i = 0. \quad (2.38')$$

Ci si può rendere conto di questo fatto su un esempio semplice: la successione dei vettori $z_n = e_1 - (e_2 + \dots + e_{n+1})/n$ soddisfa $\sum_i a_i = 0$ ma approssima e_1 quanto si vuole, infatti

$$\|z_n - e_1\|^2 = \frac{1}{n}.$$

In altre parole, si può dire che il sottospazio $V^{(0)} \subset H$ è un sottospazio *vettoriale* in H , denso in H , ma *non* è un sottospazio *di Hilbert*, poiché non è un sottospazio *completo* (forse sarebbe meglio per chiarezza dire qui *chiuso*; ancora una volta, non si confonda l'aggettivo “completo” riferito a sottospazio con lo stesso aggettivo

riferito a set di vettori, v. § 2.10). E si potrà anche dire che H è il *completamento* (o *chiusura*) di $V^{(0)}$.

Dato un sistema di vettori $\{w_n\}$ non ortogonali, ci si può chiedere se tale sistema è o no, completo, anche senza eseguirne preliminarmente la sua ortonormalizzazione. Naturalmente non saranno più vere né la (2.33) né la (2.35), e nemmeno la proprietà (iii) (che non sarebbero vere neppure in dimensione finita!); la completezza del set w_n può tuttavia essere definita mediante la prima parte della proprietà (ii), che si dovrà ora enunciare in questa forma:

per ogni $x \in H$ e per ogni $\varepsilon > 0$, si possono trovare un intero N ed N opportuni coefficienti α_i in modo che

$$\|x - \sum_{i=1}^N \alpha_i w_i\| < \varepsilon$$

anche se non c'è un procedimento generale per trovare i coefficienti α_i .

Tenendo presente il procedimento di ortonormalizzazione di Schmidt, è facile quanto importante vedere che resta vera, anche per un set "obliquo" w_n , l'equivalenza della completezza con la proprietà (iv).

Un esempio di set completo non ortogonale è dato dal set di vettori $w_n = \{e_1 - e_2, e_1 - e_3, \dots, e_1 - e_n, \dots\}$, dove e_n ($n = 1, 2, \dots$) è un set ortonormale completo (infatti ...). Usando un esempio come quello visto poco sopra, si vede che il vettore $z_n = (1/n)(w_1 + \dots + w_{n+1})$ è una combinazione di vettori w_n che approssima e_1 , infatti $\|z_n - e_1\|^2 = 1/n$. Da notare che ortonormalizzando il set di vettori w_n si ritrova il set f_n (2.38).

Un altro caso istruttivo: sia $H = L^2(I)$ e $v_n(x)$ un set ortonormale completo in H ; sia $\varphi(x)$ una funzione (continua e limitata, per semplicità). Allora $w_n(x) = \varphi(x)v_n(x)$ è un set completo (in generale, non ortogonale) se $\varphi(x) \neq 0$ oppure anche se $\varphi(x)$ ammette qualche zero isolato: infatti dalla condizione (iv) di completezza $0 = (w_n, f_0) = (v_n, \varphi^* f_0)$ segue che $f_0(x)$ è la funzione nulla (quasi ovunque!).

Un importante esempio di set completo non ortogonale verrà dato nel prossimo paragrafo.

Si può notare come quanto detto in questo paragrafo si traduca in fatti elementari e ben noti se ci si riduce allo spazio $\mathbf{R}^3$ (o a spazi a dimensione finita). Per esempio, la (2.36) è in $\mathbf{R}^2$ il teorema di Pitagora, mentre nello spazio $H = L^2(I)$ diventa

$$\int_I |f(x)|^2 dx = \sum_{n=1}^{\infty} |a_n|^2.$$

2.14 Spazi separabili e lo spazio ℓ^2

Si deve anzitutto notare che *non* in tutti gli spazi di Hilbert è presente un sistema completo⁸. Si chiamano **separabili** quegli spazi (e sono gli unici che ci interes-

⁸ Si possono infatti costruire esempi di spazi di Hilbert in cui, comunque si scelga un insieme numerabile di vettori, le combinazioni lineari di questi vettori *non* risultano dense nello spazio.

sano) nei quali esiste un set completo. È allora separabile, per esempio, lo spazio $L^2(-L, L)$, nel quale il set (2.11) è completo.

Nello stesso spazio è completo anche il sistema di vettori ortogonali

$$v_k(x) = e^{i\pi kx/L} = e^{i\omega_k x} \quad ; \quad \omega_k = k \frac{2\pi}{2L}, \quad k = 0, \pm 1, \pm 2, \dots \quad (2.39)$$

che può essere normalizzato ponendo

$$u_k = (2L)^{-1/2} v_k$$

avendo osservato che

$$(v_h, v_k) = 2L \delta_{hk}.$$

Per provarne la completezza, basta applicare la proprietà (iv) , oppure semplicemente osservare che mediante le combinazioni $(u_k \pm u_{-k})/\sqrt{2}$ si riottiene il set (2.11). Lo sviluppo in serie di una $f(x)$ rispetto al set (2.39) avrà la forma

$$f(x) = \sum_{k=-\infty}^{+\infty} c_k e^{i\omega_k x} = \sum_{k=-\infty}^{+\infty} c'_k u_k(x)$$

dove i coefficienti c_k sono dati da (basta calcolare (u_h, f) e ricordare la continuità del prodotto scalare)

$$c_k = \frac{1}{2L} \int_{-L}^L e^{-i\omega_k x} f(x) dx = \frac{1}{\sqrt{2L}} (u_k, f) = \frac{1}{\sqrt{2L}} c'_k$$

mentre l'identità di Parseval diventa qui

$$\|f\|^2 = 2L \sum_{k=-\infty}^{+\infty} |c_k|^2 = \sum_{k=-\infty}^{+\infty} |c'_k|^2. \quad (2.39')$$

Ancora, tenendo conto di quanto visto all'inizio del § 2.4, si conclude che anche il sistema $\{\sin \pi n x / L\}$ è completo nello spazio $L^2(0, L)$. Con una traslazione $x \rightarrow x - L'$ in cui $L' = L/2$ (e osservando poi che L' è generico), lo si può riportare all'intervallo $(-L, L)$ e ottenere così quest'altro set completo per $L^2(-L, L)$:

$$v_n(x) = \begin{cases} \frac{1}{\sqrt{L}} \cos \frac{\pi n_d x}{2L} = \frac{1}{\sqrt{L}} \cos \left((h - \frac{1}{2}) \frac{\pi x}{L} \right), & n_d = 2h - 1 = 1, 3, 5, \dots \\ \frac{1}{\sqrt{L}} \sin \frac{\pi n_p x}{2L} = \frac{1}{\sqrt{L}} \sin \left(h \frac{\pi x}{L} \right), & n_p = 2h = 2, 4, \dots \end{cases} \quad (2.40)$$

che è *diverso* dal set (2.11): provare infatti a disegnare i primi elementi del set (2.11) e di questo set $v_n(x)$; notare p.es. che tutti gli elementi del set (2.11) tranne il primo (la costante) hanno valor medio nullo su $(-L, L)$, mentre per le funzioni (2.40)

Per motivo analogo (basta prolungare le funzioni con *legge pari* all'intervallo $(-L, L)$), è completo in $L^2(0, L)$ il set $\{1, \cos(\pi n x / L)\}$.

Se I è un qualsiasi intervallo (di lunghezza finita), tutti i sistemi sopra ricordati sono – opportunamente traslati – altrettanti set completi in $L^2(I)$.

Dalla proprietà (ii) dei set completi segue dunque che il sottospazio vettoriale delle funzioni continue, e anche quello delle funzioni C^∞ , è *denso* in L^2 , ovvero che L^2 è il *completamento* di C^∞ rispetto alla norma L^2 .

Un altro importante sistema completo (ma non ortonormale) su un insieme compatto I è fornito dalle potenze $\{x^n\}$ ($n = 0, 1, 2, \dots$). Per provare questa proprietà si può anzitutto ricordare che, come ora osservato, ogni $f \in L^2(I)$ può essere approssimata in norma L^2 da una funzione continua g , e ricorrere poi ad un classico *teorema di Weierstrass*, che afferma che l'insieme dei polinomi è *denso rispetto alla norma* (2.25) nello spazio delle funzioni continue; per concludere basta osservare che se g e h sono “vicine” nella norma (2.25) (su un compatto I), lo sono anche nella norma $L^2(I)$:

$$\|g - h\|_{L^2(I)}^2 = \int_I |g - h|^2 dx \leq \max_{x \in I} |g(x) - h(x)|^2 \text{mis}(I).$$

Dunque anche i polinomi sono densi (nella norma $L^2(I)$) in $L^2(I)$.

Un altro fatto sorprendente è che risulta completo in $L^2(I)$ anche il set delle potenze $\{x^n\}$ con $n \geq N$ dove N è un qualsiasi numero intero > 0 . Infatti, utilizzando la proprietà (iv) della completezza, la condizione $(x^n, f_0) = 0, \forall n \geq N$ si può riscrivere $(x^m, x^N f_0) = 0, \forall m \geq 0$, ma la completezza di $\{x^m\}$ per $m \geq 0$ implica $x^N f_0(x) = 0$, dunque ...

Se poi I è tutta la retta $\mathbf{R}$, oppure una semiretta, si può dimostrare (ritroveremo più avanti questo fatto, v. § 2.23 (1)-(2)) che sono set completi in $L^2(\mathbf{R})$ e in $L^2(0, +\infty)$ insiemi di funzioni del tipo

$$e^{-x^2/2} H_n(x) \quad \text{e rispettivamente} \quad e^{-x} L_n(x), \quad n = 0, 1, 2, \dots$$

dove $H_n(x)$ e $L_n(x)$ sono opportuni polinomi di grado n , detti rispettivamente **polinomi di Hermite** e **polinomi di Laguerre** (importanti anche perché descrivono soluzioni dell'*equazione di Schrödinger* per l'oscillatore armonico e rispettivamente per l'atomo di idrogeno: v. più avanti il § 2.23).

È facile estendere questi risultati sulla esistenza di set completi anche a funzioni L^2 di più variabili reali (v. ad esempio § 2.22 e anche il § 6.13, a proposito degli spazi “prodotto tensore”).

In conclusione, tutti gli spazi L^2 sono separabili; inoltre in tutti gli spazi L^2 il sottospazio delle funzioni C^∞ è denso. Negli spazi $L^2(\mathbf{R})$ e $L^2(0, +\infty)$, è denso anche il sottospazio delle funzioni C^∞ che “decrescono esponenzialmente” all'infinito.

Sia ora H un qualsiasi spazio di Hilbert separabile ed $\{e_n\}$ un set ortonormale completo in H . Come già detto, ogni vettore $x \in H$ è individuato dalla successione delle sue componenti di Fourier a_n rispetto ad e_n :

$$x \rightarrow (a_1, a_2, \dots, a_n, \dots), \quad (2.41)$$

e, grazie all'identità di Parseval, questa successione gode della proprietà:

$$\sum_{n=1}^{\infty} |a_n|^2 < +\infty. \quad (2.42)$$

Questo suggerisce di considerare lo spazio costituito da tutte le successioni di numeri complessi $\{a_n\}$ che soddisfano la (2.42): questo spazio si indica con ℓ^2 e verificheremo tra poco che anch'esso è uno spazio di Hilbert, con la seguente definizione di prodotto scalare: se $a \equiv \{a_n\}$ e $b \equiv \{b_n\}$ sono due successioni di ℓ^2 , si definisce

$$(a, b) = \sum_{n=1}^{\infty} a_n^* b_n. \quad (2.43)$$

Lo spazio ℓ^2 gioca nei confronti degli spazi di Hilbert separabili lo stesso ruolo svolto da $\mathbf{C}^n$ nei confronti degli spazi vettoriali di dimensione n : cioè la corrispondenza data dalla (2.41) fra il generico elemento $x \in H$ e una successione numerica $a \in \ell^2$ individua un **isomorfismo** fra qualsiasi spazio di Hilbert separabile H e ℓ^2 . Infatti, la corrispondenza (2.41) è in realtà *biunivoca*, cioè è anche vero che ogni successione $a \in \ell^2$ individua un $x \in H$: i calcoli eseguiti nel § 2.12 per costruire il vettore $\bar{x}$ = limite delle migliori approssimazioni del vettore generico x tramite combinazioni degli $\{e_n\}$ mostrano infatti che la condizione (necessaria e sufficiente) sui coefficienti a_n affinché la serie $\sum a_n e_n$ sia convergente in H è precisamente la condizione “ ℓ^2 ” (2.42) (cf. la fine del § 2.3).

E ancora, l'identità di Parseval (2.36) ci dice che la norma, calcolata secondo la definizione data in H , coincide con la norma calcolata in ℓ^2 della successione numerica $\{a_n\}$ individuata da x tramite la (2.41).

Resta da provare che ℓ^2 è uno spazio di Hilbert: dimostriamo intanto che se $a \equiv \{a_n\}$ e $b \equiv \{b_n\}$ sono due successioni di ℓ^2 , anche $a + b \equiv \{a_n + b_n\} \in \ell^2$: infatti

$$|a_n + b_n|^2 \leq 2|a_n|^2 + 2|b_n|^2$$

e quindi anche $\sum |a_n + b_n|^2 < +\infty$; è d'altronde immediato che $\alpha a \in \ell^2$ per ogni $\alpha \in \mathbf{C}$, $a \in \ell^2$. Inoltre, la definizione (2.43) di prodotto scalare è corretta: la serie è anzi assolutamente convergente, poiché

$$|a_n^* b_n| = |a_n| |b_n| \leq \frac{1}{2} |a_n|^2 + \frac{1}{2} |b_n|^2$$

e le proprietà (p1, ..., p4) del § 2.6 sono evidentemente soddisfatte. Infine si potrebbe dimostrare senza difficoltà che lo spazio ℓ^2 è anche *completo* rispetto alla norma (indotta dal prodotto scalare (2.43))

$$\|a\|_{\ell^2} = \left[\sum_{n=1}^{\infty} |a_n|^2 \right]^{1/2}. \quad (2.44)$$

Un esempio naturale di set ortonormale completo in ℓ^2 è dato dal “set canonico”:

$$e_1 \equiv (1, 0, 0, \dots), e_2 \equiv (0, 1, 0, \dots), \dots, e_n \equiv (\underbrace{0, \dots, 0}_n, 1, 0, \dots), \dots \quad (2.45)$$

Si può anche introdurre lo spazio ℓ^1 , formato dalle successioni di numeri complessi $a = \{a_n\} (n = 1, 2, \dots)$ tali che

$$\sum_{n=1}^{\infty} |a_n| < \infty .$$

Si tratta di uno spazio di Banach, non di Hilbert, perché la norma $\|a\|_{\ell^1} = \sum_{n=1}^{\infty} |a_n|$ non è ottenibile tramite un prodotto scalare. Si vede facilmente che se $\{a_n\} \in \ell^1$ allora anche $\in \ell^2$ (ma non viceversa!). Nonostante le notazioni, non c'è alcun collegamento diretto fra gli spazi ℓ^1 ed $L^1(I)$.

2.15 Proprietà ed applicazioni delle serie di Fourier

La tecnica degli sviluppi in serie di Fourier trova innumerevoli applicazioni in diversi contesti.

Un esempio particolarmente semplice è il seguente.

Si consideri una massa m legata ad una molla di costante elastica K e sottoposta ad una forza $f(t)$ periodica di periodo $2T$; l'equazione del moto per lo spostamento $x = x(t)$ della massa dalla posizione di riposo è

$$m\ddot{x} + Kx = f(t) .$$

Supponendo per semplicità che la funzione $f(t)$ sia dispari, il suo sviluppo in serie di Fourier è (ovviamente qui si sostituisce la variabile x con il tempo t)

$$f(t) = \sum_{n=1}^{\infty} f_n \sin \omega_n t \quad \omega_n = n \frac{2\pi}{2T} .$$

Cerchiamo una soluzione $x(t)$ anch'essa periodica con lo stesso periodo $2T$, dunque del tipo

$$x(t) = \alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos \omega_n t + \sum_{n=1}^{\infty} \beta_n \sin \omega_n t$$

sostituendo nell'equazione del moto⁹ si trova subito che $\alpha_0 = \alpha_n = 0$, mentre

$$(-m\omega_n^2 + K)\beta_n = f_n .$$

Dunque, se per tutti gli n si ha $K \neq m\omega_n^2$, si trova la soluzione

$$x(t) = \sum_{n=1}^{\infty} \frac{f_n}{K - m\omega_n^2} \sin \omega_n t .$$

⁹ Anche qui valgono le osservazioni già fatte nello studio dell'evoluzione della corda elastica; v. la nota nel § 2.4.

Se invece per qualche n_0 si ha $\omega_{n_0}^2 = K/m$ e $f_{n_0} \neq 0$, l'equazione per β_{n_0} non ha soluzione (è il ben noto fenomeno della risonanza). Ci si può meravigliare che il metodo dello sviluppo di Fourier non riesca a risolvere il caso della presenza di una risonanza (e nemmeno trovi le soluzioni dell'equazione omogenea): la spiegazione sta semplicemente nel fatto che, avendo usato la serie di Fourier e quindi cercato soluzioni periodiche con un fissato periodo, tutte le soluzioni non periodiche (la soluzione nel caso di risonanza) e quelle con periodo diverso (le soluzioni dell'equazione omogenea) vengono perdute.

Se nel precedente problema della molla inseriamo anche la presenza di uno smorzamento viscoso:

$$m\ddot{x} + \beta\dot{x} + Kx = f(t)$$

si può ripetere lo stesso procedimento, con l'unica differenza che ora nella soluzione $x(t)$ saranno presenti anche componenti in $\cos \omega_n t$. Alternativamente, può essere utile sviluppare $f(t)$ e $x(t)$ in serie rispetto al set completo¹⁰ $v_k = e^{-i\omega_k t}$ (2.39).

Il set (2.39) risulta anche particolarmente utile ad esempio quando si vuole studiare la risposta di circuiti elettrici sottoposti a tensioni periodiche: in tal caso le $v_k = v_k(t)$ rappresentano i termini "alternati puri" di frequenza ω_k ; il segnale periodico verrà espresso come serie di questi; v. più avanti in questo paragrafo, e anche i §§ 2.28, 4.2 e 4.10.

Tornando allo studio della evoluzione temporale nel problema della corda elastica e dell'equazione del calore (§§ 2.4, 2.5), varie osservazioni possono essere utili. Anzitutto si vede ora che le condizioni iniziali possono essere tutte prese in $L^2(0, L)$; inoltre, la (2.18) mostra, per quanto detto nel paragrafo precedente, che la soluzione $u(x, t)$ del problema della corda resta – per ogni t fissato – in L^2 .

Nel caso poi dell'equazione del calore (2.19), si può anche osservare che la serie nella (2.18') risulta convergente, per $t > 0$, anche *totalmente* e dunque *uniformemente*, e pertanto si ha che $u(x, t)$ è anche una funzione continua (oltre che L^2 , ovviamente). Ma si può dire di più.

Ricordiamo intanto un teorema elementare sulla derivazione delle successioni (e quindi anche delle serie) di funzioni: se una successione di funzioni derivabili converge puntualmente ad una $f = f(x)$ e se la *successione delle derivate* è *uniformemente* convergente, allora la successione delle derivate converge alla derivata $f'(x)$ della f (si noti che *non* basterebbe l'ipotesi di convergenza uniforme della serie di partenza: p.es. $f_n(x) = \frac{1}{n} \sin nx$ converge uniformemente a 0, ma $f'_n = \cos nx$ non converge nemmeno puntualmente!).

Tornando all'equazione del calore, questa osservazione permette di concludere che $u(x, t)$ diventa "immediatamente" (cioè appena $t > 0$) infinitamente derivabile sia rispetto ad x che rispetto a t – nonostante che alla condizione iniziale $u(x, 0)$ si richieda solo che di essere L^2 . Questo segue dalla convergenza totale e quindi uniforme delle serie ottenute derivando termine a termine: infatti, qualsiasi derivata eseguita termine a termine nella (2.18') conduce a una serie di Fourier i cui coeffi-

¹⁰ È ovviamente irrilevante scrivere $e^{i\omega_k t}$ oppure $e^{-i\omega_k t}$, dato che $k \in \mathbf{Z}$; nel seguito, per coerenza con la scelta che sarà fatta per la trasformata di Fourier, scriveremo generalmente $e^{-i\omega_k t}$.

cienti sono del tipo $a_n n^k e^{-n^2 t}$ (si è posto per semplicità $C = 1$, $L = \pi$) e dunque è garantita, per ogni $t > 0$, la convergenza uniforme (e anche quella L^2 , ovviamente).

Si può pure verificare che la soluzione $u(x, t)$ trovata nel § 2.5 tende a zero, in norma $L^2(0, \pi)$, per $t \rightarrow +\infty$: infatti si ha

$$\|u(x, t)\|^2 = \frac{\pi}{2} \sum_{n=1}^{\infty} |a_n|^2 e^{-2n^2 t} \leq \frac{\pi}{2} e^{-2t} \sum_{n=1}^{\infty} |a_n|^2 = e^{-2t} \|u(x, 0)\|^2 \rightarrow 0$$

anzi, più precisamente, la “velocità di decrescenza a zero” di tale norma sarà $\simeq e^{-2n_1^2 t}$ se a_{n_1} è primo coefficiente non nullo dello sviluppo in serie di Fourier di $u(x, 0)$.

Possiamo dare ora anche una motivazione più precisa del procedimento della “separazione delle variabili” adottato per la soluzione delle equazioni di d’Alembert e del calore. Considerando infatti, per esempio, l’equazione di d’Alembert (2.1) e scegliendo di nuovo le condizioni al contorno (2.2), basta chiedere che la soluzione sia, per ogni $t \in \mathbf{R}$, una funzione di $L^2(0, L)$: si può allora scrivere la più generale soluzione come serie

$$u(x, t) = \sum_{n=1}^{+\infty} a_n(t) \sin(nx)$$

($L = \pi$, $v = 1$), con coefficienti $a_n(t)$ dipendenti dal tempo. Sostituendo nell’equazione si trova $\ddot{a}_n = -n^2 a_n$ e si ritrovano le soluzioni già note.

È anche possibile introdurre diverse condizioni agli estremi: p.es., le condizioni dette “di periodicità” agli estremi, cioè (si è scelto qui per comodità $L = 2\pi$)

$$u(0, t) = u(2\pi, t), \quad u_x(0, t) = u_x(2\pi, t);$$

ora il set completo appropriato è $v_n = e^{-inx}$, $n \in \mathbf{Z}$ (2.39), e la soluzione va scritta nella forma

$$u(x, t) = \sum_{n=-\infty}^{+\infty} a_n(t) e^{-inx}.$$

Il metodo si presta ad essere generalizzato: si consideri per esempio l’equazione del calore con un termine “non omogeneo” $F(x) \in L^2(0, 2\pi)$ (con $C = 1$, $L = 2\pi$):

$$u_t = u_{xx} + F(x)$$

con le condizioni di periodicità agli estremi, e con la condizione iniziale assegnata $u(x, 0) = f(x) \in L^2(0, 2\pi)$. Sviluppando

$$F(x) = \sum_{n=-\infty}^{\infty} F_n e^{-inx} \quad \text{e} \quad f(x) = \sum_{n=-\infty}^{\infty} f_n e^{-inx}$$

e sostituendo, si trova per $a_n(t)$ l’equazione

$$\dot{a}_n = -n^2 a_n + F_n$$

che si risolve facilmente tenendo conto della condizione iniziale, che implica $a_n(0) = f_n$. La soluzione è

$$u(x, t) = \sum_{n=-\infty}^{\infty} \left(f_n e^{-n^2 t} + \frac{F_n}{n^2} (1 - e^{-n^2 t}) \right) e^{-inx}.$$

Del teorema sopra ricordato circa la derivazione delle serie di funzioni anticipiamo una versione “avanzata” più generale, che verrà precisata più avanti: v. § 2.28; in breve: la possibilità di derivare termine a termine una serie di funzioni è ancora garantita se la serie *delle derivate* è convergente *in senso* L^2 . Una importante conseguenza è, per esempio, che se si ha una serie di Fourier del tipo

$$f(x) = \sum_n a_n e^{-inx} \quad \text{con} \quad |a_n| \leq \frac{c}{|n|^\alpha} \quad \text{e} \quad \alpha > k + \frac{1}{2}$$

almeno per “grandi” n , $|n| > n_0$, dove c è una costante e k un intero, allora f è derivabile k volte con derivata in L^2 , ma la k -esima derivata potrà essere non continua. (Ovviamente un risultato analogo vale per sviluppi rispetto a set completi costituiti da funzioni come $\cos nx$ e/o $\sin nx$).

2.16 Trasformazioni lineari

La definizione di **trasformazione** (od **operatore**) **lineare** T fra due spazi di Hilbert H ed H'

$$T : H \rightarrow H'$$

è, almeno al principio, del tutto ovvia e simile al caso degli operatori fra spazi a dimensione finita:

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y); \quad \alpha, \beta \in \mathbf{C}, \quad x, y \in H.$$

In realtà si vedrà immediatamente che la dimensione infinita produce notevoli (forse non inattese) differenze e – naturalmente – qualche difficoltà.

La prima “sorpresa” viene quando si osserva che – a differenza di quanto avviene in dimensione finita – può succedere che non sia semplice definire una trasformazione lineare su *tutti* i vettori dello spazio. Ad esempio se $H = L^2(I) = H'$ e T è l’operatore di derivazione d/dx , è chiaro che T non può essere definito su tutto H poiché non tutte le funzioni di H sono derivabili, e inoltre non sempre $f \in L^2$ implica che anche $df/dx \in L^2$ (p.es., $f = \sqrt{x}$ in $L^2(0, 1)$). Ma l’aspetto forse più sorprendente di questa particolarità consiste nel fatto che può capitare che la trasformazione T sia in realtà perfettamente definita su *tutti* i vettori di un sistema completo dello spazio (questo avviene, per esempio, proprio per l’operatore d/dx considerato sopra, che è certamente definito su tutti i sistemi completi visti finora negli spazi L^2) senza che questo permetta di estenderne la definizione a tutti gli altri vettori. Anche questa è una notevole differenza rispetto al caso della dimensione

finita, dove era sufficiente definire una trasformazione sui vettori di base per poterla poi estendere “per linearità” a tutti i vettori.

Per capire meglio questo fatto consideriamo quest’altro esempio: sia $H = H'$ e $\{e^{(n)}\}$ un sistema ortonormale completo in H , sia T l’operatore così definito¹¹

$$T e^{(n)} = n e^{(n)}, \quad n = 1, 2, \dots \quad (2.46)$$

Sia x il vettore, appartenente ad H , le cui componenti di Fourier rispetto ad $\{e^{(n)}\}$ sono date da $a_n = 1/n$ (la successione $\{a_n = 1/n\} \in \ell^2$), cioè

$$x = \sum_{n=1}^{\infty} \frac{1}{n} e^{(n)}.$$

Grazie alla linearità dell’operatore T si può immediatamente calcolare come T trasforma tutte le somme parziali *finite* di tale serie:

$$T \left(\sum_{n=1}^N \frac{1}{n} e^{(n)} \right) = \sum_{n=1}^N \frac{1}{n} T e^{(n)} = \sum_{n=1}^N e^{(n)},$$

ma si vede che, mentre – da un lato – la successione delle somme parziali tende al vettore x , dall’altro la successione delle loro trasformate sotto l’operatore T *non converge*, e quindi non dà la possibilità di *estendere* ad x la definizione della trasformazione T . Nel prossimo paragrafo analizzeremo meglio questa situazione.

L’insieme dei vettori su cui è definita una trasformazione lineare T costituisce un sottospazio vettoriale dello spazio H ; tale sottospazio D si chiama **dominio** di T . Gli esempi mostrati ora indicano casi in cui $D \neq H$; naturalmente in altri casi si potrà avere $D = H$, come vedremo su molti esempi.

Si chiama invece **immagine** o **codominio** o **range** di T il sottospazio D' di H' costituito da tutti i vettori “trasformati” $T x$. Come avviene per il dominio, può anche avvenire che l’immagine di un operatore T contenga tutti i vettori di un set completo, e tuttavia *non* coincida con l’intero spazio: p.es. l’immagine dell’operatore di H in sé definito da $T e^{(n)} = e^{(n)}/n$ ($n = 1, 2, \dots$) contiene tutti gli $\{e_n\}$ ma *non* contiene il vettore $y = \sum e^{(n)}/n$ (infatti $\dots$).

2.17 Continuità, limitatezza e norma di un operatore. Estensione per continuità

È innanzitutto chiaro che, in dimensione finita, se si fa tendere a zero un vettore x , facendone tendere a zero le sue componenti, anche il vettore $T x$, qualunque sia la trasformazione lineare (cioè la matrice) T , tende a zero. La stessa cosa però *non* è in generale vera in dimensione infinita, cioè se $x \rightarrow 0$ (nel senso della norma:

¹¹ In meccanica quantistica l’operatore “energia” dell’oscillatore armonico è un operatore di questo tipo, v. i §§ 2.23(1) e 6.15.

$\|x\| \rightarrow 0$) non è detto che $Tx \rightarrow 0$. Ad esempio, se $e^{(n)}$ è il set canonico (2.45) nello spazio ℓ^2 , la successione di vettori $x^{(n)} = e^{(n)}/n$ tende evidentemente a zero per $n \rightarrow \infty$, poiché $\|x^{(n)}\| = 1/n$, mentre se prendiamo come trasformazione T l'operatore $Te^{(n)} = ne^{(n)}$ considerato nel paragrafo precedente, la successione $Tx^{(n)} = e^{(n)}$ non tende a zero, anzi non è convergente (non è di Cauchy).

Una trasformazione lineare T si dice **continua** in x_0 ($x_0 \in D$) se

$$x \rightarrow x_0 \quad \text{cioè} \quad \|x - x_0\| \rightarrow 0 \quad \Rightarrow \quad Tx \rightarrow Tx_0 \quad \text{cioè} \quad \|Tx - Tx_0\| \rightarrow 0$$

(dove naturalmente $x \in D$). È assai utile notare che, in realtà, affinché una trasformazione lineare T sia continua ovunque è sufficiente che essa sia continua in un punto, ad esempio *nello zero*, cioè è sufficiente verificare che

$$x \rightarrow 0 \quad \Rightarrow \quad Tx \rightarrow 0.$$

Infatti, se x e x_0 sono elementi qualsiasi del dominio D e $x \rightarrow x_0$, si ha $(x - x_0) \rightarrow 0$ e quindi, se T è continua nello zero, anche $Tx - Tx_0 = T(x - x_0) \rightarrow 0$.

Vale la pena di notare che il concetto di continuità di una trasformazione è fondato sul concetto di “intorno” (e cioè di “vicinanza”: una trasformazione è continua se trasforma elementi “vicini” in elementi “vicini”). Come sappiamo, nel nostro caso è proprio la norma che fornisce questa struttura “topologica” allo spazio.

Oltre all'esempio dato all'inizio di questo paragrafo, anche l'operatore d/dx non è continuo: la successione in $L^2(0, \pi)$ (v. anche il § 2.15) $f_n = \frac{1}{n} \sin nx \rightarrow 0$ in norma L^2 per $n \rightarrow \infty$, ma Tf_n non è convergente.

Il concetto di continuità di un operatore è importante in relazione alla questione vista nel paragrafo precedente. Infatti, sappiamo che ogni vettore x è il limite della successione delle somme parziali della sua serie di Fourier rispetto ad un dato sistema completo $e^{(n)}$:

$$x_n = \sum_{k=1}^n a_k e^{(k)}, \quad x_n \rightarrow x.$$

Allora, se T è un operatore continuo, sarà anche $Tx_n \rightarrow Tx$ e, in questo caso, si potrà scrivere

$$Tx = T\left(\sum_{k=1}^{\infty} a_k e^{(k)}\right) = \sum_{k=1}^{\infty} a_k T e^{(k)}. \quad (2.47)$$

Dimostreremo alla fine di questo paragrafo che la continuità di un operatore T è *equivalente* ad un'altra importante proprietà, che viene detta **limitatezza** di T e che è in genere più facilmente verificabile, come si vedrà presto; tale proprietà si definisce come segue.

Un operatore lineare T si dice **limitato** se la quantità $\|Tx\|/\|x\|$ al variare di x nel dominio D di T si mantiene limitata, ovvero – in altre parole – se

$$\sup_{x \in D} \frac{\|Tx\|}{\|x\|} = K < +\infty \quad (2.48)$$

(va ovviamente escluso in questo rapporto il caso $x = 0$; ricordare che $T(0) = 0$, infatti $T(0) = T(x - x) = T(x) - T(x) = 0$).

Se un operatore T è limitato, la quantità finita K definita dalla (2.48) si chiama **norma** di T e si indica con $\|T\|$: si vede subito infatti che si tratta di una vera norma, nel senso che tale definizione, applicata allo spazio vettoriale $\mathcal{L}(H)$ degli operatori lineari limitati di H , verifica le condizioni assegnate nel § 2.6. Molti esempi verranno visti nel seguito.

Naturalmente, per ogni vettore $x \in H$, vale la maggiorazione

$$\|Tx\| \leq \|T\| \|x\|. \quad (2.49)$$

Qualche esempio di calcolo della norma di un operatore (altri esempi saranno incontrati più avanti):

i) Se T è l'operatore di traslazione in $L^2(\mathbf{R})$ definito da $Tf(x) = f(x - a)$, dove a è un numero reale, si ha $\|Tf\| = \|f(x - a)\| = \|f(x)\|$, $\forall f(x) \in L^2(\mathbf{R})$. Dunque $\|T\| = 1$.

ii) Sia e_n ($n = 1, 2, \dots$) un set ortonormale completo e T l'operatore così definito sui vettori $x = \sum_n a_n e_n$:

$$Tx = \sum_{n=1}^{\infty} a_{2n} e_n = a_2 e_1 + a_4 e_2 + \dots$$

Si ha subito $\|Tx\| \leq \|x\|$; questo implica $\|T\| \leq 1$, ma prendendo p. es. $\tilde{x} = e_2$ si ha $\|T\tilde{x}\| = \|\tilde{x}\|$, dunque $\|T\| = 1$.

iii) Se T è definito in ℓ^2 da $T(a) = (a_1/2, 0, a_3/3, 0, \dots, a_{2n-1}/(n+1), 0, \dots)$ dove $a = (a_1, a_2, \dots)$, si ha

$$\|T(a)\|^2 = (|a_1|^2/4 + |a_3|^2/9 + \dots) \leq (1/4)(|a_1|^2 + |a_3|^2 + \dots) \leq \|a\|^2/4$$

da cui $\|T\| \leq 1/2$. Ma prendendo p. es. $\tilde{x} = e_1$ si ha $\|T\tilde{x}\| = (1/2)\|\tilde{x}\|$, dunque $\|T\| = 1/2$.

In questi esempi, la norma è stata ottenuta come “max”, infatti si è sempre trovato un vettore $\tilde{x}$ tale che $\|T\tilde{x}\|/\|\tilde{x}\| = \|T\|$. Ma ecco un esempio in cui la norma si ottiene veramente come “sup”.

iv) Si consideri in $L^2(0, a)$ l'operatore

$$Tf(x) = xf(x) \quad (2.50)$$

si ha:

$$\|Tf\|^2 = \int_0^a x^2 |f(x)|^2 dx \leq a^2 \int_0^a |f(x)|^2 dx = a^2 \|f\|^2.$$

Dunque $\|T\| \leq a$. Per avere una $\tilde{f}$ tale che $\|T\tilde{f}\| \simeq a\|\tilde{f}\|$ occorre prendere una $\tilde{f}$ con supporto il più possibile vicino al punto $x = a$ (non si può prendere una $\tilde{f}$ con supporto nel solo punto $x = a$: questa sarebbe (equivalente a) la funzione nulla!);

occorre dunque prendere una famiglia di funzioni del tipo

$$f_\varepsilon(x) = \begin{cases} 0 & \text{per } 0 < x < a - \varepsilon \\ 1 & \text{per } a - \varepsilon < x < a \end{cases}$$

e si può verificare in questo modo che $\|T f_\varepsilon\|/\|f_\varepsilon\| \rightarrow a$ per $\varepsilon \rightarrow 0$.

Lo stesso operatore (2.50) è invece non limitato in $L^2(\mathbf{R})$, e ha dominio denso in $L^2(\mathbf{R})$ (ma non coincidente con l'intero $L^2(\mathbf{R})$, infatti ...).

v) Più in generale, per gli “operatori di moltiplicazione” $Tf(x) = h(x)f(x)$ con $f(x) \in L^2(I)$ e dove $h = h(x)$ è una funzione assegnata, si ha $\|T\| = \sup_{x \in I} |h(x)|$, infatti

$$\|Tf\|^2 = \int_I |h(x)f(x)|^2 dx \leq \sup_{x \in I} |h(x)|^2 \|f\|^2$$

e dunque ...¹². Analogamente, T è non limitato se $\sup_{x \in I} |h(x)| = +\infty$. P.es., $Tf = (1/x)f$ è non limitato (ma ha dominio denso, infatti ...) in $L^2(0, 1)$, mentre $\|T\| = 1$ in $L^2(1, +\infty)$.

vi) L'operatore d/dx è non limitato: basta ricordare che non è continuo, oppure, osservare che, p.es. in $L^2(0, \pi)$, si ha $\|(d/dx)(\sin nx)\|/\|\sin nx\| = n$.

Torniamo ora al caso in cui un operatore lineare viene definito soltanto sugli elementi di un set ortonormale completo $\{e^{(n)}\}$ di uno spazio H , e si vuole vedere se è possibile estenderne la definizione a tutto lo spazio. Come già detto, se H avesse dimensione finita, l'operatore risulterebbe automaticamente definito su tutti i vettori dello spazio. In caso contrario, invece, utilizzando la linearità dell'operatore, possiamo intanto definirlo certamente sull'insieme (denso!) F delle combinazioni lineari finite degli $e^{(n)}$. Si può ora procedere “a vista”, controllando direttamente se, presa una qualsiasi successione di vettori $\{x_n\} \subset F$, la successione Tx_n è tale che il rapporto $\|Tx_n\|/\|x_n\|$ si mantiene limitato. Più formalmente, si calcola l'estremo superiore

$$\sup_{x \in F} \frac{\|Tx\|}{\|x\|}$$

come nella (2.48), ma ora restringendosi ai soli vettori del sottoinsieme F ; se tale estremo superiore è *finito*, e sia K_F , allora se ne può concludere che T si può *prolungare* (o *estendere*) “per continuità” su tutti gli elementi di H . Infatti, se x è un elemento qualsiasi in H (e $\notin F$), c'è senz'altro una successione di vettori $x_n \in F$ che tende in norma ad x (per esempio la successione delle “troncate” della serie di Fourier di x rispetto al set $\{e^{(n)}\}$): basta ora osservare che la successione Tx_n è una successione di Cauchy, infatti

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

essendo pure di Cauchy la successione x_n in quanto convergente. Se ne conclude (per la completezza dello spazio di Hilbert) che anche la Tx_n deve essere convergente

¹² Si noti che qui *non* è possibile usare la disuguaglianza di Schwarz, perché né $|h(x)|^2$ né $|f(x)|^2$ sono in generale funzioni L^2 . E *non* è vero che $\int_I |hf|^2 dx \leq \left| \int_I hf dx \right|^2$!

e il vettore limite y di tale successione si prende come *definizione* del “trasformato” Tx del vettore x :

$$Tx_n \rightarrow y := T(x).$$

In questo modo l'operatore T risulta definito su tutto H , e si potrebbe pure dimostrare che, se si approssima il vettore x con qualsiasi altra successione $z_n \rightarrow x$, si otterrebbe lo stesso risultato, cioè anche $Tz_n \rightarrow y$ (quindi la definizione di Tx *non dipende* dalla successione scelta per approssimare x). In altre parole, l'operatore così prolungato risulta continuo e, in particolare, il vettore Tx può essere calcolato mediante la formula (2.47), dove gli a_k sono ovviamente i coefficienti di Fourier di x rispetto al set $\{e^{(k)}\}$. Inoltre si ha che l'operatore prolungato su tutto H ha ancora norma K_F :

$$\sup_{x \in F} \frac{\|Tx\|}{\|x\|} = \sup_{x \in H} \frac{\|Tx\|}{\|x\|} = \|T\|.$$

Dimostriamo ora:

Teorema. *Continuità e limitatezza di un operatore lineare sono proprietà equivalenti.*

Dimostrazione.

(i) Sia T continuo. Supponiamo, per assurdo, che non sia limitato: allora, per ogni numero n , si potrebbe trovare un vettore x_n tale che

$$\|Tx_n\| > n\|x_n\|.$$

Si consideri ora la successione

$$y_n = \frac{x_n}{n\|x_n\|}.$$

Poiché $y_n \rightarrow 0$, dovrebbe anche essere $Ty_n \rightarrow 0$ – essendo T continuo – mentre la precedente disuguaglianza dà

$$\|Ty_n\| = \frac{1}{n\|x_n\|} \|Tx_n\| > 1.$$

(ii) Se T è limitato, si ha

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\| \|x - x_0\| \rightarrow 0$$

e quindi T è continuo. □

2.18 Una applicazione concernente il problema della corda elastica.

Come semplice applicazione di quanto visto nei precedenti paragrafi, studiamo il problema dell'evoluzione temporale di una corda elastica mantenuta ad estremi fissi.

Supponiamo che nelle condizioni iniziali sia

$$g = \frac{\partial u}{\partial t} \Big|_{t=0} = 0$$

mentre $f(x) = u(x, 0)$ sia una funzione generica (ovviamente, come già detto, $f \in L^2$; le notazioni sono quelle del § 2.4). Supponiamo però – per pura comodità – che ora la corda abbia lunghezza $2L$ e abbia gli estremi in $x = \pm L$. Detto I l'intervallo $(-L, L)$, si tratta, come si è già visto, di sviluppare la funzione $f \in L^2(I)$ in serie di Fourier rispetto al sistema ortonormale completo costituito dalle “onde stazionarie”, che in questo caso ha la forma $v_n(x)$ data nella (2.40):

$$u(x, 0) = f(x) = \sum_{n=1}^{\infty} a_n v_n(x) = \sum_{n \text{ dispari}} \alpha_n \cos\left(\frac{\pi n}{2L} x\right) + \sum_{n \text{ pari}} \beta_n \sin\left(\frac{\pi n}{2L} x\right).$$

Prima di cercare l'evoluzione temporale di questa configurazione, consideriamo l'operatore lineare T di $L^2(I)$ in se stesso così definito:

$$T f(x) = -f(-x).$$

È immediato provare che T è limitato (anzi: $\|T\| = 1$) e quindi continuo, e pertanto è valida la relazione (2.47), che ora diventa

$$\begin{aligned} T f(x) &= \sum_{n=1}^{\infty} a_n T(v_n(x)) = \\ &= \sum_{n \text{ dispari}} (-\alpha_n) \cos\left(\frac{\pi n}{2L} x\right) + \sum_{n \text{ pari}} \beta_n \sin\left(\frac{\pi n}{2L} x\right) = \sum_{n=1}^{\infty} (-1)^n a_n v_n(x). \end{aligned}$$

Tornando al problema precedente, l'evoluzione temporale dell'onda risulta

$$u(x, t) = \sum_{n \text{ dispari}} \alpha_n \cos\left(\frac{\pi n}{2L} x\right) \cos(\omega_n t) + \sum_{n \text{ pari}} \beta_n \sin\left(\frac{\pi n}{2L} x\right) \cos(\omega_n t)$$

dove $\omega_n = \pi n v / 2L$.

Si può concludere allora, per esempio, che negli istanti $t = 2L/v, 6L/v, \dots$ si ha $u(x, t) = -f(-x)$.

2.19 Operatore aggiunto. Operatori unitari. Proiettori

(a) Operatore aggiunto

La definizione dell'operatore aggiunto T^+ di un operatore lineare T è formalmente identica a quella introdotta in dimensione finita tramite la (1.12'): dato l'operatore $T : H \rightarrow H'$, si dice **trasformazione aggiunta** (o semplicemente **aggiunto**)

l'operatore

$$T^+ : H' \rightarrow H$$

(indicato anche con $T^\dagger$) tale che, per ogni x, y' nel rispettivo dominio, si ha

$$(y', Tx) = (T^+ y', x). \quad (2.51)$$

Rinviando al § 2.27 per una discussione più completa, ci limitiamo qui a dire che, se l'operatore T è limitato, allora il suo aggiunto esiste sempre ed è anch'esso un operatore lineare e limitato, anzi si ha $\|T\| = \|T^+\|$, infatti

$$\|T^+ x\|^2 = (T^+ x, T^+ x) = (x, TT^+ x) \leq \|x\| \|T\| \|T^+ x\|$$

avendo usato la disuguaglianza di Schwarz e la limitatezza di T ; se ne ricava $\|T^+ x\| \leq \|T\| \|x\|$, dunque T^+ è limitato e $\|T^+\| \leq \|T\|$. Scambiando i ruoli di T e T^+ si arriva al risultato. Si ha inoltre che il dominio di T^+ coincide con tutto H' .

Si verifica anche immediatamente che

$$(T^+)^+ = T \quad \text{e} \quad (TS)^+ = S^+ T^+$$

dove T ed S sono due qualsiasi operatori limitati.

Scegliendo ora per semplicità $H = H'$, e preso un set ortonormale completo $\{e_n\}$ in H , i prodotti scalari $(e_i, T e_j) = T_{ij}$ si chiamano, come in dimensione finita, elementi di matrice dell'operatore T . Da notare per inciso che – utilizzando gli elementi T_{ij} e almeno nel caso che T sia limitato – si può generalizzare alla dimensione infinita la (1.10): posto $x' = Tx$, le componenti x'_i del vettore “trasformato” x' si possono infatti esprimere tramite le componenti x_i di x nella forma

$$x'_i = \sum_{j=1}^{\infty} T_{ij} x_j.$$

Per quanto riguarda l'operatore aggiunto T^+ , si verifica subito che si ha (come nella (1.12))

$$(T^+)_{ij} = (e_i, T^+ e_j) = (T e_i, e_j) = (e_j, T e_i)^* = T_{ji}^*.$$

Maggiore cautela è richiesta invece se T non è limitato, soprattutto per quanto riguarda la determinazione del dominio D' dell'aggiunto T^+ (v. anche il § 2.27); in ogni caso, comunque, la proprietà caratterizzante è ancora la (2.51).

Per esempio, si vuole trovare l'aggiunto dell'operatore di derivazione $T = d/dx$ in $L^2(\alpha, \beta)$: si ha intanto (naturalmente nel sottinsieme delle funzione derivabili con derivata in L^2)

$$(f, Tg) = \left(f, \frac{dg}{dx}\right) = \int_{\alpha}^{\beta} f^* g' dx = [f^* g]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} f'^* g dx = [\dots] + \left(-\frac{df}{dx}, g\right);$$

occorre dunque eliminare il termine $[\dots]$, che non può essere messo nella forma di prodotto scalare (h, g) . Per esempio si può chiedere che il dominio di T sia costituito

dalle funzioni derivabili che *si annullano agli estremi* α, β : allora il dominio di T^+ è dato dalle funzioni derivabili limitate agli estremi. Ma si può invece chiedere che il dominio di T e quello di T^+ siano dati *entrambi* dalle funzioni derivabili che soddisfano le *condizioni di periodicità agli estremi*, cioè $f(\alpha) = f(\beta)$ (si noti che tutti questi domini sono densi in L^2). Restringendosi a quest'ultimo dominio, che indicheremo con D_0 , si può allora dire che in D_0 si ha $T^+ = -T$.

Se poi si considera l'operatore¹³

$$T = i \frac{d}{dx} \quad (2.52)$$

ne segue che in D_0 sussiste l'identità

$$(f, Tg) = (Tf, g) \quad \forall f, g \in D_0. \quad (2.53)$$

Un operatore T per il quale si può trovare un opportuno dominio D_0 , denso in H , contenuto nel dominio di T , in cui è soddisfatta la (2.53), si dice **simmetrico o hermitiano** (in D_0).

Se poi avviene che il dominio D di un operatore T coincide con il dominio D' di T^+ e che in tale dominio vale la (2.53), allora si scrive

$$T = T^+ \quad (2.54)$$

e si dice che T è **autoaggiunto**. È chiaro allora che la (2.53) è in generale un po' più debole della (2.54); esse sono equivalenti se T è continuo: in tal caso infatti è $D = H$. Si può dire tuttavia che, per esempio, l'operatore (2.52) “si comporta come autoaggiunto” nel dominio D_0 sopra definito¹⁴.

Si può anche dimostrare che *un operatore hermitiano il cui dominio è l'intero spazio di Hilbert H è limitato*.

(b) Trasformazioni unitarie

Sia $U : H \rightarrow H$; la proprietà caratteristica di conservare i prodotti scalari:

$$(Ux, Uy) = (x, y) \quad (2.55)$$

che si può anche scrivere $U^+U = \mathbf{I}$, ora, in dimensione infinita, *non* è più sufficiente a garantire l'invertibilità dell'operatore U . Precisamente, la (2.55) assicura (oltre alla limitatezza dell'operatore, anzi $\|U\| = 1$) soltanto l'iniettività di U , infatti

$$Ux = Uy \Rightarrow 0 = \|U(x-y)\| = \|x-y\| \Rightarrow x = y$$

ma non la sua surgettività (per un esempio, v. più sotto la (2.57)). Aggiungendo alla (2.55) l'ulteriore condizione (che, in dimensione finita, è automaticamente verificata, v. § 1.4)

$$\text{Imm } U = H \quad (2.56)$$

¹³ È proporzionale all'operatore “impulso” o “quantità di moto” in meccanica quantistica.

¹⁴ Per maggiori dettagli, v. per esempio i testi di Abbati e Cirelli [23], di Akhiezer e Glazman [27], di Reed e Simon [19], o di Yoshida [31].

(dove $\text{Imm}U$ indica l'immagine di U), allora U risulta anche surgettivo e dunque invertibile, e si ricava – come in dimensione finita –

$$U^{-1} = U^+.$$

Riassumendo, un operatore si dice unitario se soddisfa entrambe le (2.55-56). È anche utile osservare che la (2.55) può essere sostituita dalla condizione più “debole” della sola conservazione delle norme:

$$\|Ux\| = \|x\|; \quad (2.55')$$

infatti si prova facilmente che, grazie alla linearità di U , la (2.55') implica la (2.55).

Come già detto, la sola condizione (2.55), *non* è sufficiente a garantire la surgettività (e dunque l'unitarietà): basta considerare, per esempio, l'operatore lineare, detto anche “operatore di salita”, definito da

$$T e_n = e_{n+1}, \quad n = 1, 2, 3, \dots \quad (2.57)$$

dove e_n è un set ortonormale completo. Questo operatore soddisfa la (2.55), dunque conserva i prodotti scalari ed è iniettivo (il suo nucleo è infatti $\text{Ker}(T) = \{0\}$), ma *non* è surgettivo. Da notare che l'aggiunto di T , che è

$$T^+ e_n = e_{n-1} \quad \text{se } n > 1, \quad \text{e } T^+ e_1 = 0 \quad (2.57')$$

(operatore “di discesa”) è invece surgettivo ma *non* iniettivo (il suo nucleo è unidimensionale, $\text{Ker}(T) = \{\alpha e_1\}$, $\alpha \in \mathbf{C}$) e $T^+ T = \mathbf{I}$ ma $T T^+ \neq \mathbf{I}$; si ha anche $\|T\| = \|T^+\| = 1$.

Sono, per esempio, unitari lo stesso operatore (2.57) se $n = 0, \pm 1, \pm 2, \dots$, oppure gli operatori di traslazione $U_a: f(x) \rightarrow g(x) = f(x-a)$ in $L^2(\mathbf{R})$, o la trasformazione considerata nel § 2.18. Una proprietà *caratteristica* degli operatori unitari, che è semplice controllare, è quella di trasformare set ortonormali completi in set ancora ortonormali completi.

In altre parole, gli operatori unitari possono essere visti come “cambiamenti di base” che ne conservano l'ortonormalità. Esattamente come in dimensione finita (v. il § 1.4), ci si può chiedere come si trasforma un generico operatore T passando alla nuova base: il diagramma (1.16) e la (1.16') sono ancora validi, dunque si ha

$$T' = UTU^+.$$

Da qui segue, in particolare, che se A è l'aggiunto di T , allora $A' = UTU^+$ è l'aggiunto di T' ; se T è simmetrico in un dominio D_0 , allora T' è simmetrico nel dominio UD_0 dei vettori trasformati Ux (con $x \in D_0$). Se T è un proiettore (v. il prossimo punto (c)) che proietta su un sottospazio H_1 , anche T' lo è e proietta sul sottospazio trasformato UH_1 .

Più in generale, se S è un operatore *invertibile* si può considerare la trasformazione $T \rightarrow T'' = STS^{-1}$. Una trasformazione di questo tipo (v. ancora il § 1.4) non conserva l'ortonormalità delle basi, né p.es. le proprietà di autoaggiunzione degli

operatori. Tuttavia, se T ammette un autovettore x con autovalore λ (v. il prossimo paragrafo), allora T'' ammette l'autovettore Sx con autovalore λ . Se T soddisfa un'equazione algebrica $a_n T^n + \dots + a_0 \mathbf{I} = 0$, anche T'' soddisfa la stessa equazione.

(c) Proiettori

Sia V_1 un sottospazio *vettoriale* di H : se V_1 è *denso* in H , allora chiaramente non esiste in H alcun vettore ortogonale a tutti i vettori $x_1 \in V_1$ e il completamento (o chiusura) $\overline{V_1}$ di V_1 coincide con H (per esempio, questo avviene se $H = L^2(I)$ e V_1 è il sottospazio delle funzioni continue). Se invece V_1 non è denso, si considerino tutti i vettori $x_2 \in H$ ortogonali a tutti gli $x_1 \in V_1$: è facile verificare che l'insieme di tali vettori x_2 forma un sottospazio *di Hilbert* (cioè è anche automaticamente *completo* (o chiuso): basta applicare la continuità del prodotto scalare, v. la (2.37), infatti ...). Si indichi con H_2 tale sottospazio. Se V_1 non era un sottospazio di Hilbert (cioè non era completo), sia $H_1 = \overline{V_1}$ il suo completamento (o chiusura): H_1 risulta costituito da tutti i vettori ortogonali agli $x_2 \in H_2$, dunque H_1 e H_2 sono due spazi di Hilbert ortogonali e "complementari". Si scrive allora H come **somma diretta**

$$H = H_1 \oplus H_2 \quad (2.58)$$

e si dice che H_2 è il **complemento ortogonale** di H_1 (si scrive anche $H_2 = H_1^\perp$, e naturalmente pure $H_1 = H_2^\perp$).

Si può ora dimostrare che per ogni vettore $x \in H$ è sempre possibile eseguire, in modo unico, la decomposizione, esattamente come in dimensione finita (cfr. § 1.8)

$$x = x_1 + x_2; \quad x_1 \in H_1, \quad x_2 \in H_2. \quad (2.59)$$

Ad esempio, se si sceglie in H un set ortonormale completo in modo che una parte dei suoi elementi $\{e_{n'}\}$ sia un set completo per H_1 e la rimanente $\{e_{n''}\}$ sia un set completo per H_2 , allora $\{e_{n'}\} \cup \{e_{n''}\}$ è set completo per H , le componenti $(e_{n'}, x)$ sono le componenti di Fourier di x_1 , e le componenti $(e_{n''}, x)$ sono quelle di x_2 . Ed è anche facile dimostrare che la decomposizione (2.59) è unica, infatti se fosse $x = x_1 + x_2 = x'_1 + x'_2$ si otterrebbe $x_1 - x'_1 + x_2 - x'_2 = 0$ da cui

$$0 = (x_1 - x'_1, x_1 - x'_1 + x_2 - x'_2) = \|x_1 - x'_1\|^2.$$

Una volta eseguita la decomposizione (2.59), è naturale definire gli operatori di proiezione (o proiettori):

$$P_1 x = x_1, \quad P_2 x = x_2 \quad \text{e} \quad P_1 + P_2 = \mathbf{I}, \quad P_1 P_2 = P_2 P_1 = 0. \quad (2.60)$$

È immediato vedere che si tratta di operatori lineari, che ogni proiettore è continuo (anzi è $\|P\| = 1$), e inoltre valgono le proprietà

$$P^+ = P, \quad P^2 = P. \quad (2.61)$$

È facile anche verificare che un operatore definito su tutto H è un proiettore se e solo se soddisfa le (2.61). I proiettori qui definiti si chiamano più propriamente proiettori

ortogonali (sono proiettori “obliqui” quelli che soddisfano soltanto la seconda delle (2.61)).

Ad esempio, se $I = (-L, L)$, lo spazio $H = L^2(I)$ è somma diretta dei due sottospazi $H_1 = L_p^2(I)$ e $H_2 = L_d^2(I)$ costituiti rispettivamente dalle funzioni pari e dispari; un set completo in $H_1 = L_p^2(I)$ è dato dalle funzioni $1, \cos(n\pi x/L)$, uno in $H_2 = L_d^2(I)$ dalle funzioni $\sin(n\pi x/L)$.

Una utile e semplice proprietà dei proiettori: se P' e P'' sono due proiettori su due sottospazi rispettivamente H' e H'' , si ha che $P' + P''$ è un proiettore se e solo se $P'P'' = P''P' = 0$, ovvero se H' e H'' sono ortogonali: $(x', x'') = 0, \forall x' \in H', x'' \in H''$. Naturalmente, se questa ipotesi è soddisfatta, $P' + P''$ proietta sul sottospazio $H' \oplus H''$.

2.20 Autovalori ed autovettori. Spettro di un operatore

Sia $T : H \rightarrow H$. L'equazione agli autovalori si imposta esattamente come in dimensione finita: si tratta cioè di trovare soluzioni *non nulle* $x \in D$ dell'equazione

$$Tx = \lambda x; \quad \lambda \in \mathbf{C}, \quad x \neq 0. \quad (2.62)$$

La situazione ora è però totalmente diversa: nulla infatti si può dire, in generale, circa l'esistenza di autovettori di un operatore in dimensione infinita. Questa difficoltà *non* dipende dall'eventualità che l'operatore sia o no limitato. Ecco alcuni esempi significativi; v. più avanti per altri esempi:

i) L'operatore (2.57) è limitato (estendibile per continuità) e $\|T\| = 1$. L'equazione agli autovalori, con $x = \sum_n a_n e_n$, è

$$Tx = \sum a_n e_{n+1} = a_1 e_2 + a_2 e_3 + \dots = \lambda(a_1 e_1 + a_2 e_2 + \dots)$$

da cui si ottiene un sistema di infinite equazioni $0 = \lambda a_1, a_1 = \lambda a_2, \dots$ che ammette la sola soluzione $x = 0$.

ii) Per l'operatore (2.57') invece, le analoghe equazioni danno il sorprendente risultato che c'è un autovettore per ogni λ tale che $|\lambda| < 1$.

iii) Lo stesso operatore (2.57) ma ora con $n = 0, \pm 1, \pm 2, \dots$: si vede facilmente che non ha autovettori (si ricordi sempre che deve essere $\sum_n |a_n|^2 < \infty$!).

iv) Nemmeno l'operatore (unitario) di traslazione $T : f(x) \rightarrow f(x-a)$ in $L^2(\mathbf{R})$, $a \in \mathbf{R}$, ammette alcuna autofunzione: infatti, osservando che $\|Tf\| = \|f\|$, si ha che gli eventuali autovalori λ devono soddisfare $|\lambda| = 1$, ne segue che le autofunzioni di T dovrebbero soddisfare $|f(x-a)| = |f(x)|$; ma questo implica che $|f(x)|$ è una funzione periodica e dunque $f \notin L^2(\mathbf{R})$.

v) Anche l'operatore di moltiplicazione (2.50) non ha autofunzioni: infatti, l'equazione $(x - \lambda)f(x) = 0$ è risolta solo da funzioni $f(x)$ diverse da zero solo al più in un punto, dunque $\dots$, indipendentemente dal fatto che I sia o no un in-

tervallo finito (ma v. i §§4.12, 5.4, 5.6). Più in generale, un operatore di moltiplicazione $Tf(x) = h(x)f(x)$ ammette autofunzioni se la funzione $h(x)$ è *costante* in qualche sottointervallo: le autofunzioni sono le funzioni con supporto in questo sottointervallo.

vi) Si può trovare una “realizzazione” concreta dell’operatore assegnato nell’esempio iii): scegliendo $H = L^2(0, 2\pi)$ e $e_n = e^{-inx}$, l’operatore diventa $Te_n = e^{-ix}e_n$, e dunque semplicemente $Tf(x) = e^{-ix}f(x)$ su tutte le $f(x) \in L^2(0, 2\pi)$. Ovviamente l’equazione agli autovalori $Tf = \lambda f$ non ha soluzioni, confermando il risultato visto in iii).

vii) L’operatore definito da $Tx = v(w, x)$, dove $v, w \in H$ sono due vettori fissati, ammette l’autovalore $\lambda = 0$ con autovettori tutti i vettori x ortogonali a w , e l’autovalore $\lambda = (w, v)$ con autovettore $x = v$. Se $v = w$ e $\|v\| = 1$, T è il proiettore sul sottospazio unidimensionale generato da v .

viii) Ogni operatore di proiezione (§2.19(c)) ammette ovviamente un set completo di autovettori (i vettori indicati con $\{e'_n\}$ ed $\{e''_n\}$, rispettivamente con autovalori 1 e 0). Anche un operatore lineare T così definito su un set ortonormale completo $\{e_n\}$

$$Te_n = c_n e_n \quad (2.63)$$

dove c_n è una qualsiasi successione di numeri complessi, possiede un set ortonormale completo di autovettori (gli $\{e_n\}$ stessi), con autovalori $\lambda_n = c_n$ che possono essere o no degeneri a seconda di come sono scelti i numeri c_n . Si ha anche $\|T\| = \sup_n |c_n|$ e dunque T può essere o no limitato a seconda di come sono scelti i numeri c_n .

Sussiste comunque il seguente (semplice ma importante):

Teorema. *Se T è simmetrico e se possiede qualche autovettore, si ha che:*

- (i) *suoi autovalori sono reali;*
- (ii) *autovettori relativi ad autovalori diversi sono ortogonali.*

Dimostrazione.

(i) Sia $Tx = \lambda x$, allora $\lambda(x, x) = (x, Tx) = (Tx, x) = \lambda^*(x, x)$, essendo T simmetrico, quindi $\lambda = \lambda^*$.

(ii) Siano λ_1 e λ_2 due autovalori diversi: $Tx_1 = \lambda_1 x_1$, $Tx_2 = \lambda_2 x_2$. Allora

$$(x_2, Tx_1) = \lambda_1(x_2, x_1), \quad (Tx_2, x_1) = \lambda_2(x_2, x_1);$$

ma $(Tx_2, x_1) = (x_2, Tx_1)$, quindi, sottraendo membro a membro:

$$(\lambda_1 - \lambda_2)(x_2, x_1) = 0 \text{ da cui } (x_2, x_1) = 0, \text{ essendo } \lambda_1 \neq \lambda_2. \quad \square$$

Naturalmente questa dimostrazione poteva essere data anche nel caso di operatore T hermitiano in spazi a dimensione *finita*. La dimostrazione data nel §1.6 era stata in quel caso preferita perché aveva permesso di giungere alla ulteriore conclusione, valida in dimensione finita, che gli autovettori formano una *base* per lo spazio.

Un altro esempio già incontrato di operatore (in dimensione infinita) che possiede un set completo di autofunzioni è dato dall’operatore $T = d^2/dx^2$ con la condizione di *annullamento* agli estremi dell’intervallo: il set delle sue autofunzioni

$\{\sin n\pi x/L\}$ è infatti completo in $L^2(0, L)$ (ovvero il set (2.40) in $L^2(-L, L)$). Gli insiemi (2.11) e (2.39) sono autofunzioni di T scegliendo invece la condizione di *periodicità agli estremi*, cioè:

$$f(-L) = f(L), \quad f'(-L) = f'(L)$$

e sono set completi in $L^2(-L, L)$.

Si noti pure che – come si verifica facilmente – l'operatore d^2/dx^2 risulta *simmetrico* sia nel dominio delle funzioni che si annullano agli estremi, sia scegliendo le condizioni di periodicità. Ma si tratta comunque di operatori *diversi*: p.es., se si vuole risolvere l'equazione $f'' = g$ in $L^2(0, \pi)$ (dove $g(x)$ è data e $f(x)$ è l'incognita) con condizioni di annullamento agli estremi, un set completo appropriato sarà $\{\sin nx\}$; sviluppando f e g in serie di Fourier si trova

$$f'' = \sum_{n=1}^{\infty} (-n^2 f_n) \sin nx = \sum_{n=1}^{\infty} g_n \sin nx = g(x)$$

e si vede che la soluzione $f(x)$ esiste unica, con coefficienti $f_n = -g_n/n^2$, per ogni $g(x)$. Invece, considerando lo stesso problema, nello stesso spazio, ma con condizioni di periodicità agli estremi 0 e π , un set completo appropriato è $\{e^{-2inx}\}$; sviluppando si ha

$$f'' = 4 \sum_{n=-\infty}^{\infty} (-n^2 f_n) e^{-2inx} = \sum_{n=-\infty}^{\infty} g_n e^{-2inx} = g(x)$$

ma ora il problema ha soluzione *soltanto se* $g_0 = 0$ (dovendo essere $-4n^2 f_n = g_n$, $n = 0, \pm 1, \pm 2, \dots$), e la soluzione, quando esiste, *non* è unica, ma è data a meno di una costante additiva f_0 .

Una utile osservazione può essere la seguente: sia $T = -d^2/dx^2$ in $L^2(0, \pi)$ con condizioni di annullamento agli estremi $x = 0$ e $x = \pi$. L'equazione del calore (§ 2.5) può essere scritta come

$$\frac{du}{dt} = T u, \quad u = u(x, t)$$

e si vede che la soluzione (2.18') è esattamente l'equivalente “infinito-dimensionale” dell'evoluzione temporale trovata nel § 1.7 per i sistemi dinamici a dimensione finita, e precisamente della (1.24'): le autofunzioni $\{\sin nx\}$ (set completo!) di T prendono il posto degli autovettori $\mathbf{c}^{(k)}$ della matrice A nella (1.23'), e la condizione iniziale $u(x, 0)$ si sviluppa come nella (1.24).

Enunceremo nel prossimo paragrafo un teorema che fornisce una classe di operatori con proprietà simili all'operatore d^2/dx^2 e nel § 2.30 vedremo un'altra speciale classe di operatori che ammette un set ortonormale completo di autovettori.

Altri risultati utili sono i seguenti.

(1) Se un operatore T soddisfa un'equazione algebrica (dove $\mathbf{I}$ è l'operatore identità)

$$\mathcal{P}(T) = a_n T^n + \dots + a_0 \mathbf{I} = 0$$

si vede facilmente che i suoi autovalori λ sono le soluzioni dell'equazione $\mathcal{P}(\lambda) = 0$. Inoltre T possiede effettivamente autovettori: infatti, detta λ_1 una soluzione di $\mathcal{P}(\lambda) = 0$, si ha $\mathcal{P}(T) = (T - \lambda_1) \mathcal{Q}(T) = 0$ e dunque, per qualsiasi $v \in H$, si ha $(T - \lambda_1) \mathcal{Q}(T) v = 0$; allora, se v è tale che $\mathcal{Q}(T)v \neq 0$, si deduce che $w := \mathcal{Q}(T)v$ è autovettore di T con autovalore λ_1 ; se invece $\mathcal{Q}(T)v = 0$, il problema si riconduce ad un'equazione di grado $n - 1$.

(2) Se T ed S sono due operatori tali che $TS = ST$, e T ha un autovalore λ , allora il sottospazio V_λ dei corrispondenti autovettori è un sottospazio *invariante* sotto S , cioè $S|_{V_\lambda} : V_\lambda \rightarrow V_\lambda$, infatti, per ogni $v_\lambda \in V_\lambda$ (cfr. con § 1.6),

$$T S v_\lambda = S T v_\lambda = \lambda S v_\lambda \quad \Rightarrow \quad S v_\lambda \in V_\lambda.$$

Se, in particolare, λ è *non degenera*, ne segue che v_λ è anche autovettore di S .

Se invece V_λ ha dimensione *finita*, allora per la restrizione $S|_{V_\lambda}$ valgono i teoremi per le matrici (a dimensione finita), e dunque se, per esempio, tale restrizione è una matrice normale allora esiste in V_λ una base di autovettori *anche* per S , etc.

Una estensione del concetto di autovalore è quella di **spettro** di un operatore: un numero (reale o complesso) σ si dice che appartiene allo spettro dell'operatore T se l'operatore $(T - \sigma \mathbf{I})$ *non ammette inverso limitato*: cioè se *non* è invertibile, oppure se ammette un inverso *non* limitato.

Anzitutto, se σ è un autovalore di T , si ha che $T - \sigma \mathbf{I}$ non è iniettivo e dunque non invertibile; ne segue che ogni autovalore di un operatore appartiene al suo spettro. Se si considera p.es. l'operatore lineare definito da $T e_n = (1/n) e_n$ ($n = 1, 2, \dots$) oltre ai suoi autovalori $\lambda_n = 1/n$, appartiene al suo spettro anche $\sigma = 0$, infatti l'inverso T^{-1} esiste ma non è limitato.

È anche evidente che in dimensione finita lo spettro di un operatore coincide con l'insieme dei suoi autovalori. Si vede invece facilmente, per esempio, che lo spettro dell'operatore $T f(x) = x f(x)$ in $L^2(I)$ è dato dai numeri reali $\sigma \in \bar{I} =$ chiusura dell'insieme I : infatti l'operatore inverso è $(T - \sigma \mathbf{I})^{-1} f(x) = f(x)/(x - \sigma)$ e risulta non limitato se (e solo se) $\sigma \in \bar{I}$: in tal caso $T - \sigma \mathbf{I}$ è iniettivo e il suo range (o immagine) è *denso* in $L^2(I)$, ma *non* coincidente con $L^2(I)$.

Invece, come altro esempio, l'operatore (2.57) pur essendo iniettivo, ha range non denso in H (essendo dato dal sottospazio ortogonale a e_1), dunque non è invertibile (nemmeno con inverso non limitato), e si può concludere così che $\sigma = 0$ appartiene al suo spettro.

Si chiama **risolvente** di T l'insieme complementare dello spettro: dunque se $\lambda \in \mathbb{C}$ è nel risolvente, l'equazione $(T - \lambda)x = v$, dove $v \in H$ è assegnato e x il vettore incognito, ha una unica soluzione $x = (T - \lambda)^{-1}v$. Se invece λ è nello spettro, l'equazione ha soluzione solo se $v \in \text{Imm}(T - \lambda \mathbf{I})$ e in questo caso la soluzione è unica se λ non è autovalore di T (altrimenti si possono aggiungere le soluzioni z dell'equazione omogenea $(T - \lambda)z = 0$).

Si può anche dimostrare che lo spettro è un insieme chiuso (e dunque il risolvente un insieme aperto).

2.21 Problema di Sturm-Liouville

Si consideri l'equazione

$$\frac{1}{\rho(x)} \left[\frac{d}{dx} \left(p(x) \frac{d}{dx} \right) - q(x) \right] u + \lambda u = 0 \quad (2.64)$$

per la funzione $u = u(x)$, dove $p(x)$, $q(x)$, $\rho(x)$ sono funzioni assegnate “abbastanza regolari”, come preciseremo più avanti, in un intervallo finito $\alpha \leq x \leq \beta$; si supponga inoltre che la funzione $u(x)$ debba soddisfare alle condizioni al bordo, per esempio di annullamento agli estremi ¹⁵

$$u(\alpha) = u(\beta) = 0. \quad (2.65)$$

Si tratta evidentemente di un problema agli autovalori, detto **Problema di Sturm-Liouville**: ci si deve attendere che solo per particolari valori di λ la $u(x)$ possa soddisfare la (2.64) con le condizioni (2.65). L'operatore che si sta studiando, cioè

$$T = -\frac{1}{\rho(x)} \left[\frac{d}{dx} \left(p(x) \frac{d}{dx} \right) - q(x) \right] \quad (2.66)$$

è un operatore (non limitato) nello spazio $L^2(\alpha, \beta)$ e generalizza l'operatore $-d^2/dx^2$ (che infatti si ottiene ponendo $\rho(x) = p(x) = 1$, $q(x) = 0$) incontrato nei primi paragrafi di questo capitolo e nel paragrafo precedente.

Supponiamo che le funzioni p, q, ρ verifichino le seguenti condizioni

$$\begin{aligned} p(x) &> 0, \text{ derivabile e con derivata continua;} \\ q(x) &\geq 0, \rho(x) > 0 \text{ entrambe continue.} \end{aligned} \quad (2.67)$$

Intanto, grazie alla condizione di positività di $\rho(x)$, è possibile definire correttamente il prodotto scalare (v. § 2.6)

$$(f, g)_\rho = \int_\alpha^\beta f^*(x) g(x) \rho(x) dx. \quad (2.68)$$

Si controlla subito facilmente che l'operatore (2.66) è *simmetrico*, rispetto a questo prodotto scalare, nel sottinsieme (denso in $L^2(\alpha, \beta)$) delle funzioni derivabili almeno due volte in L^2 e soddisfacenti le condizioni agli estremi (2.65). Ne segue (v. § 2.20) che se $u_1(x)$ e $u_2(x)$ sono due autofunzioni di T (e quindi due soluzioni delle (2.64,65)) relative ad autovalori diversi, esse risultano ortogonali rispetto alla funzione peso $\rho(x)$:

$$(u_1, u_2)_\rho = \int_\alpha^\beta u_1^*(x) u_2(x) \rho(x) dx = 0. \quad (2.69)$$

¹⁵ Si ottengono risultati simili anche imponendo altre condizioni al bordo, per esempio $u'(\alpha) = u'(\beta) = 0$, oppure $u(\alpha) = u'(\beta) = 0$, o condizioni di periodicità. V. anche il paragrafo precedente.

Possiamo infine enunciare i seguenti risultati:

a) Gli autovalori del problema di Sturm-Liouville (2.64, 65) sono tutti reali e maggiori di zero.

b) Le corrispondenti autofunzioni $\{u_n(x)\}$, che risultano ortogonali fra loro rispetto al prodotto scalare (2.68), formano un sistema completo nello spazio, relativamente alla norma indotta da tale prodotto scalare.

Ciò significa che ogni funzione f tale che

$$\int_{\alpha}^{\beta} |f|^2 \rho \, dx < +\infty$$

può essere sviluppata in serie di Fourier di tali autofunzioni u_n :

$$f = \sum_{n=1}^{\infty} a_n u_n \quad (2.70)$$

nel senso che, per $N \rightarrow \infty$,

$$\int_{\alpha}^{\beta} \left| f - \sum_{n=1}^N a_n u_n \right|^2 \rho \, dx \rightarrow 0,$$

cioè la serie di Fourier converge in media del secondo ordine rispetto alla funzione densità ρ .

Per quanto riguarda il calcolo dei coefficienti a_n , si ottiene, grazie anche alla (2.69), la regola naturale

$$(u_m, f)_{\rho} = \sum_n a_n (u_m, u_n)_{\rho} = a_m (u_m, u_m)_{\rho}.$$

Si noti ad esempio che l'equazione di Schrödinger indipendente dal tempo unidimensionale per una particella quantistica soggetta ad un potenziale $V(x)$ è del tipo

$$\left[-\frac{d^2}{dx^2} + V(x) \right] u = \lambda u$$

cioè appunto del tipo di Sturm-Liouville. Le corrispondenti autofunzioni formano quindi un sistema completo: risultato importante in fisica, perché assicura la possibilità di esprimere qualsiasi “stato” (v. § 1.11) come sovrapposizione (cioè come serie di Fourier) degli “autostati” dell'operatore energia.

Possiamo ancora enunciare altri risultati di questa teoria.

(i) Gli autovalori $\lambda_1, \lambda_2, \dots, \lambda_n, \dots$ sono crescenti e $\lambda_n \rightarrow +\infty$.

(ii) Se λ_1 è il più piccolo degli autovalori, la corrispondente autofunzione $u_1(x)$ si annulla solo nei punti $x = \alpha$ e $x = \beta$.

(iii) **Teorema di monotonia.** Mantenendo inalterate le p, q, ρ ma riducendo l'intervallo da $[\alpha, \beta]$ ad $[\alpha', \beta']$ con $\alpha < \alpha' < \beta' < \beta$, aumentano tutti gli autovalori. La stessa cosa succede aumentando p e/o q ; oppure diminuendo la ρ .

(iv) **Teorema di separazione.** Dette u_n, u_m due autosoluzioni con autovalori λ_n, λ_m rispettivamente, se $\lambda_n > \lambda_m$ allora fra due zeri consecutivi di u_m cade almeno uno zero di u_n .

(v) **Teorema di oscillazione.** Supponendo di aver ordinato gli autovalori λ_n in ordine crescente, la n -sima autosoluzione u_n ha $n - 1$ zeri nell'intervallo aperto (α, β) , ossia $n + 1$ includendo gli estremi $x = \alpha$ e $x = \beta$.

È immediato constatare come tutti questi risultati sono verificati nel caso particolare dell'operatore $-d^2/dx^2$.

Prima di passare ad altre considerazioni sul problema di Sturm-Liouville, accenniamo ad un altro caso in cui è assicurata l'esistenza di un *sistema completo* di autovettori. Si consideri l'operatore di $L^2(0, 1)$ in sé così definito:

$$Tf = \int_0^1 K(x, t) f(t) dt$$

dove $f \in L^2(0, 1)$ e $K(x, t)$ è una funzione assegnata di $L^2(0, 1) \times L^2(0, 1)$, cioè tale che

$$\int_0^1 \int_0^1 |K(x, t)|^2 dx dt < +\infty.$$

L'equazione agli autovalori $Tf = \lambda f$ è, come si vede, un'equazione integrale (l'"incognita" f compare infatti anche sotto il segno di integrale). Si riesce a provare che, se $K(x, t) = K^*(t, x)$, allora l'equazione ha un sistema completo di autosoluzioni f_n e che i corrispondenti autovalori sono reali, decrescenti e tendono a zero (v. anche il § 2.30).

2.22 L'equazione di d'Alembert in due dimensioni

Casi notevoli di equazioni di Sturm-Liouville si ottengono da molte equazioni differenziali alle derivate parziali con il metodo della separazione delle variabili e con assegnate condizioni agli estremi. Ne vedremo qualche esempio speciale nel prossimo paragrafo. Un esempio semplice verrà dato anche dall'equazione di d'Alembert in due dimensioni x, y :

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \quad u = u(x, y; t) \quad (2.71)$$

che è l'equazione cui obbediscono le vibrazioni di una superficie elastica. Separata, come nel § 2.1, la dipendenza da t :

$$u(x, y; t) = w(x, y) T(t)$$

si ottiene per $w(x, y)$ l'equazione agli autovalori

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) w = -k^2 w. \quad (2.72)$$

Se supponiamo ora che la superficie abbia un contorno rettangolare fisso, si debbono aggiungere le seguenti condizioni al contorno (i vertici del rettangolo siano $(0,0)$, $(a,0)$, (a,b) , $(0,b)$):

$$w(0,y) = w(a,y) = w(x,0) = w(x,b) = 0.$$

Separando ulteriormente

$$w(x,y) = X(x)Y(y)$$

l'equazione diventa

$$\frac{X''}{X} + \frac{Y''}{Y} = -k^2$$

che si spezza nelle due equazioni usuali (che sono due semplici equazioni di Sturm-Liouville!)

$$X'' = -k_1^2 X, \quad Y'' = -k_2^2 Y; \quad k^2 = k_1^2 + k_2^2$$

con le rispettive condizioni

$$X(0) = X(a) = 0, \quad Y(0) = Y(b) = 0.$$

Le autosoluzioni sono allora (a meno del solito fattore arbitrario)

$$w_{n_1 n_2}(x,y) = \sin(k_{n_1} x) \sin(k_{n_2} y), \quad n_1, n_2 = 1, 2, 3, \dots$$

$$k_{n_1} = \frac{\pi}{a} n_1, \quad k_{n_2} = \frac{\pi}{b} n_2; \quad k_{n_1 n_2}^2 = \pi^2 \left(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} \right).$$

Si verifica subito che tali funzioni $w_{n_1 n_2}(x,y)$ sono un *sistema completo* nello spazio di Hilbert $L^2(D)$, dove D è il dominio rettangolare in esame $[0,a] \times [0,b]$, con il prodotto scalare definito naturalmente da

$$(f,g) = \int_0^a \int_0^b f^*(x,y) g(x,y) dx dy.$$

Si noti in particolare che l'autovalore k^2 può risultare *degenere*: per esempio, nel caso $a = b$, all'autovalore $k_{n_1 n_2}^2 = (\pi/a)^2 (n_1^2 + n_2^2)$ corrispondono, per ogni $n_1 \neq n_2$, le *due* soluzioni indipendenti $w_{n_1 n_2}$ e $w_{n_2 n_1}$.

Le *linee nodali* di ciascuna soluzione $w_{n_1 n_2}$, cioè il luogo dei punti in cui $w_{n_1 n_2} = 0$, sono date da segmenti che congiungono i lati opposti del rettangolo considerato parallelamente ai suoi lati e che lo dividono esattamente in $n_1 \times n_2$ rettangoli uguali fra loro.

La parte dipendente dal tempo, corrispondentemente a ciascuna $w_{n_1 n_2}$, sarà naturalmente $T(t) = \cos(\omega_{n_1 n_2} t + \delta_{n_1 n_2})$, con $\omega_{n_1 n_2} = v k_{n_1 n_2}$.

Il caso di superficie con contorno circolare sarà considerato nel prossimo paragrafo (punto 4).

2.23 Equazione di Sturm-Liouville con punti singolari. Alcune funzioni speciali

Le autosoluzioni dell'equazione di Sturm-Liouville sono funzioni regolari (continue e limitate) se α e β sono finiti e se sono verificate le condizioni (2.67) sulle funzioni $p(x)$, $q(x)$ e $\rho(x)$. I punti in cui una o più di queste condizioni non sono verificate si chiamano **punti singolari** dell'equazione e sono punti in cui la soluzione può presentare singolarità. In generale, si riesce a dimostrare che, imponendo opportune condizioni di regolarità alle soluzioni, le autosoluzioni formano un set completo anche in questo caso. Accenneremo ad alcuni esempi particolarmente importanti di questa situazione¹⁶.

(1) Polinomi e funzioni di Hermite

L'equazione

$$\left(\frac{d^2}{dx^2} - x^2 + \lambda\right)u(x) = 0, \quad -\infty < x < +\infty \quad (2.73)$$

presenta come punti singolari i punti $x = \pm\infty$. Si vede facilmente che le soluzioni della (2.73) si comportano “asintoticamente” (per grandi x) come $e^{\pm x^2/2}$. Si può provare che si hanno soluzioni non divergenti all'infinito soltanto se λ assume i valori “discreti” $\lambda = \lambda_n = 2n + 1$, $n = 0, 1, 2, \dots$, ed in tal caso le soluzioni corrispondenti hanno la forma

$$u_n(x) = H_n(x) e^{-x^2/2}, \quad n = 0, 1, 2, \dots$$

dove $H_n(x)$ è un polinomio di grado n . Questi polinomi sono detti **polinomi di Hermite** (e le $u_n(x)$ *funzioni di Hermite*); si prova anche che tali autovalori λ_n sono non degeneri e che per le soluzioni così trovate rimangono vere tutte le proprietà date nel § 2.21; in particolare si ha che il set $u_n(x)$ è *completo* in $L^2(\mathbf{R})$. Si ha inoltre che i polinomi di Hermite sono *pari* per n pari e *dispari* per n dispari.

Poiché l'operatore $-d^2/dx^2 + x^2$ è il doppio dell'operatore “energia” dell'oscillatore armonico in meccanica quantistica (di massa $m = 1$, costante elastica $k = 1$ e avendo posto $\hbar = 1$), se ne ricava che gli autovalori dell'energia dell'oscillatore armonico quantistico sono dati da $\frac{1}{2}\lambda_n = n + \frac{1}{2}$.

(2) Polinomi di Legendre e armoniche sferiche

L'equazione

$$(\Delta - W(r) + \lambda)u = 0 \quad (2.74)$$

dove W è una funzione assegnata della sola variabile $r = \sqrt{x^2 + y^2 + z^2}$, scritta in coordinate sferiche r, θ, φ , diventa

$$\left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} - W(r) + \lambda\right]u = 0. \quad (2.75)$$

¹⁶ Per maggiori dettagli, v. per esempio i testi di Weinberger [22] o di Morse e Feshbach [18].

Posto $u = R(r)\Theta(\theta)\Phi(\varphi)$, l'equazione si separa in tre equazioni di Sturm-Liouville. Nell'equazione per $\Phi(\varphi)$ la condizione che va imposta per φ è ovviamente

$$\Phi(\varphi + 2\pi) = \Phi(\varphi)$$

con questa, l'equazione per Φ si risolve immediatamente:

$$\Phi_m(\varphi) = e^{im\varphi}, \quad m = 0, \pm 1, \pm 2, \dots$$

L'equazione per Θ diventa

$$\left[\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} \right] \Theta + \mu \Theta = 0 \quad (2.76)$$

dove naturalmente μ è la costante introdotta dalla separazione delle equazioni ed m un numero intero fissato dalla soluzione $\Phi_m(\varphi)$. La (2.76) presenta i punti $\theta = 0, \theta = \pi$ come punti singolari. Si dimostra che si hanno soluzioni *regolari* solo se $\mu = \ell(\ell + 1)$, con $\ell = 0, 1, 2, \dots$, e se $|m| \leq \ell$, e le autosoluzioni corrispondenti $\Theta_\ell^{(m)}(\theta)$ sono polinomi in $\sin \theta$ e $\cos \theta$. Le funzioni

$$Y_{\ell m}(\theta, \varphi) = \Theta_\ell^{(m)}(\theta) e^{im\varphi}$$

si chiamano **armoniche sferiche**. Si noti che le $e^{im\varphi}$ sono ortogonali fra loro per valori diversi di m , mentre, con lo stesso m , le $\Theta_\ell^{(m)}$ sono ortogonali fra loro, per valori diversi di ℓ , rispetto alla *funzione peso* $\sin \theta$. Quindi le $Y_{\ell m}$ risultano ortogonali fra loro proprio rispetto all'*angolo solido* $d\Omega = \sin \theta d\theta d\varphi$. Esse forniscono un set completo nello spazio delle funzioni $f = f(\theta, \varphi) \in L^2(S)$, dove S è la superficie della sfera. In particolare, le funzioni $\Theta_\ell^{(0)}(\theta)$ si chiamano **polinomi di Legendre** e sono un altro esempio di *sistema completo* in $L^2(0, \pi)$.

(3) Polinomi di Laguerre

La terza delle equazioni in cui si è separata la (2.75) riguarda la funzione $R(r)$ ed è l'unica che contiene la funzione $W(r)$: essa è

$$\left[\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) - \frac{\ell(\ell + 1)}{r^2} - W(r) + \lambda \right] R = 0 \quad (2.77)$$

dove ora ℓ è un numero intero fissato da quanto detto nel precedente esempio. Scegliendo (proporzionale al potenziale coulombiano)

$$W(r) = -\frac{2}{r} \quad (2.77')$$

i punti singolari sono $r = 0$ e $r = \infty$. Le autosoluzioni regolari che si ottengono, imponendo in particolare la condizione

$$\lim_{r \rightarrow \infty} R(r) = 0,$$

sono del tipo

$$R_n^{(\ell)}(r) = e^{-\alpha r} L_n^{(\ell)}(r) \quad (n = 1, 2, \dots)$$

dove $\alpha > 0$ è una costante e gli $L_n^{(\ell)}(r)$ sono dei polinomi di grado $n - 1$ nella variabile r . Se in particolare è $\ell = 0$, i polinomi $L_n^{(0)}(r)$ si chiamano **polinomi di Laguerre**. Il sistema $R_n^{(0)}(r)$ è un *sistema completo* in $L^2(0, \infty)$ rispetto alla funzione peso $\rho = r^2$.

Anche l'operatore $\Delta - W(r)$ nella (2.74) è collegato (a meno di costanti moltiplicative) all'energia in meccanica quantistica; in particolare gli autovalori che si trovano nel caso del potenziale (2.77') sono proporzionali a $1/n^2$, a loro volta proporzionali agli autovalori dell'energia dell'atomo di idrogeno (v. anche il § 6.14).

(4) Funzioni di Bessel

Si scriva l'equazione di d'Alembert in due dimensioni (previa separazione della variabile t : v. (2.71-72))

$$(\Delta + k^2)w = 0$$

in coordinate polari r, φ ; si ottiene

$$\left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + k^2 \right] w(r, \varphi) = 0. \quad (2.78)$$

Se tale equazione si riferisce al problema di una superficie elastica a contorno circolare fisso, si deve imporre la condizione al contorno

$$w(a, \varphi) = 0$$

dove a è il raggio del contorno circolare. Separando

$$w = R(r) \Phi(\varphi)$$

e scrivendo ora la soluzione dell'equazione per φ nella forma più conveniente (si ricordi la condizione $\Phi(\varphi + 2\pi) = \Phi(\varphi)$)

$$\Phi = \sin(m\varphi + \delta_m), \quad m = 0, 1, 2, \dots$$

si ottiene per R

$$\left[\frac{1}{r} \frac{d}{dr} \left(r \frac{d}{dr} \right) - \frac{m^2}{r^2} + k^2 \right] R = 0$$

e infine, posto $x = kr$, $f(x) = R(x/k)$:

$$f'' + \frac{1}{x} f' + \left(1 - \frac{m^2}{x^2} \right) f = 0 \quad (2.79)$$

che si chiama **equazione di Bessel**. I suoi punti singolari sono $x = 0$, $x = \infty$. Più in generale l'equazione può essere studiata anche per valori non interi di m . Per ogni

valore intero di $m = 0, 1, 2, \dots$, la soluzione regolare in $x = 0$ si chiama **funzione di Bessel di prima specie** e si indica col simbolo $J_m(x)$.

Le $J_m(x)$ hanno andamento oscillante smorzato con infiniti zeri non equidistanti. Inoltre $J_0(0) = 1$ mentre $J_m(0) = 0$ per $m = 1, 2, \dots$; le $J_m(x)$ sono espresse da serie di potenze del tipo

$$J_m(x) = x^m \sum_{n=0}^{\infty} a_n x^{2n}$$

dove le serie convergono per ogni x .

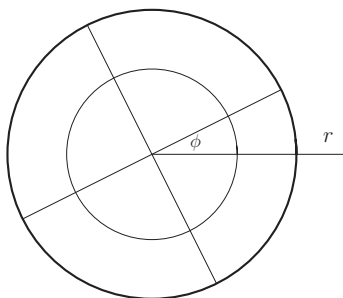


Figura 2.6 Linee nodali di una soluzione dell'equazione di d'Alembert per il cerchio.

Per quanto riguarda la condizione al contorno, essa diventa nel nostro caso $J_m(ka) = 0$ e quindi ka deve coincidere con uno degli zeri della $J_m(x)$. Questo determina k , cioè l'autovalore. Le linee nodali di ogni autosoluzione sono di due tipi e si ottengono cercando i punti in cui $\sin(m\phi + \delta_m) = 0$ oppure $J_m(kr) = 0$ nell'interno del cerchio considerato. La prima equazione fornisce m diametri del cerchio, la seconda dei cerchi concentrici al primo aventi i raggi r_i tali che kr_i coincida con gli (eventuali) zeri della J_m contenuti fra 0 e ka (v. fig. 2.6, dove sono indicate le linee nodali relative al caso $m = 2$ e con ka coincidente con il terzo zero della funzione $J_2(x)$).

Zeri delle funzioni di Bessel

$J_0 = 0$	per $x =$	—	$\pm 2.41 \dots$	$\pm 5.52 \dots$	$\dots$
$J_1 = 0$	per $x =$	0	$\pm 3.83 \dots$	$\pm 7.02 \dots$	$\dots$
$J_2 = 0$	per $x =$	0	$\pm 5.14 \dots$	$\pm 8.42 \dots$	$\dots$

2.24 Equazione di Laplace e funzioni armoniche. Problemi di Dirichlet e di Neumann

Passiamo ora a considerare l'equazione alle derivate parziali per la funzione incognita $u = u(x, y)$

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (2.80)$$

che si chiama **equazione di Laplace** e le cui soluzioni si chiamano **funzioni armoniche**.

Le funzioni armoniche descrivono, per esempio, fra molte altre grandezze fisiche, il potenziale elettrostatico piano in assenza di cariche.

Risolviamo ora esplicitamente il problema di trovare il potenziale all'interno di una circonferenza di centro l'origine e raggio $r = 1$, essendo noto il potenziale al bordo $F(\varphi)$, ovvero risolviamo la (2.80) con la condizione

$$u(r = 1, \varphi) = F(\varphi) \quad (2.81)$$

avendo introdotto coordinate polari r, φ e assumendo che $F(\varphi) \in L^2(0, 2\pi)$. Scritta l'equazione nelle coordinate r, φ

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = 0$$

e separate le variabili

$$u(r, \varphi) = R(r) \Phi(\varphi)$$

si trovano immediatamente per $\Phi(\varphi)$ le solite soluzioni

$$\Phi(\varphi) = e^{im\varphi} \quad m = 0, \pm 1, \pm 2, \dots$$

mentre l'equazione per $R(r)$, che diventa

$$R'' + \frac{1}{r} R' - \frac{m^2}{r^2} R = 0$$

è facilmente risolta, per ogni $m \geq 0$, a meno di una costante moltiplicativa, da

$$R_m(r) = r^m.$$

In realtà l'equazione per $R(r)$ ha anche le soluzioni singolari nell'origine r^{-m} se $m \neq 0$, e $\log r$ se $m = 0$, che nel problema che stiamo ora considerando vanno escluse in quanto divergenti (naturalmente, se il problema si pone per esempio in una corona circolare, occorre tener conto anche di queste soluzioni). La soluzione sarà allora da ricercare fra le funzioni del tipo

$$u(r, \varphi) = \alpha_0 + \sum_{m=1}^{\infty} r^m (\alpha_m e^{im\varphi} + \beta_m e^{-im\varphi}) \quad (2.82)$$

dove, per soluzioni $u(r, \varphi)$ reali, si ha

$$\alpha_0 = \alpha_0^*, \quad \alpha_m = \beta_m^*.$$

Dobbiamo ora imporre la condizione al contorno (2.81): per $r = 1$ dovrà essere:

$$u(1, \varphi) = \alpha_0 + \sum_{m=1}^{\infty} (\alpha_m e^{im\varphi} + \beta_m e^{-im\varphi}) = F(\varphi)$$

e dunque i coefficienti $\alpha_0, \alpha_m, \beta_m$ sono i coefficienti dello sviluppo in serie di Fourier della funzione reale assegnata $F(\varphi)$ rispetto al set completo $e^{im\varphi}$:

$$\alpha_0 = \frac{1}{2\pi} \int_0^{2\pi} F(\theta) d\theta, \quad \alpha_m = \frac{1}{2\pi} \int_0^{2\pi} F(\theta) e^{-im\theta} d\theta = \beta_m^*. \quad (2.83)$$

Da notare in particolare che, per $r = 0$, si ha dalla (2.82) che il potenziale al centro è dato da $u(0, \varphi) = \alpha_0 =$ valor medio dei valori al bordo $F(\varphi)$, come è ben noto. Sostituendo le (2.83) nella (2.82), si trova

$$\begin{aligned} 2\pi u(r, \varphi) &= \int_0^{2\pi} F(\theta) d\theta + 2 \operatorname{Re} \int_0^{2\pi} d\theta F(\theta) \sum_{m=1}^{\infty} r^m e^{im(\varphi-\theta)} = \\ &= \int_0^{2\pi} d\theta F(\theta) \left(1 + 2 \operatorname{Re} \sum_{m=1}^{\infty} z^m \right) \end{aligned}$$

avendo posto $z = r e^{i(\varphi-\theta)}$. Essendo

$$\sum_{m=1}^{\infty} z^m = \frac{1}{1-z} - 1 = \frac{z}{1-z}, \quad |z| = r < 1 \quad (2.84)$$

e inoltre (come si verifica con semplici passaggi)

$$1 + 2 \operatorname{Re} \frac{z}{1-z} = \frac{1-r^2}{1+r^2-2r\cos(\varphi-\theta)}$$

ne segue la formula finale, detta **formula di Poisson**

$$u(r, \varphi) = \frac{1-r^2}{2\pi} \int_0^{2\pi} \frac{F(\theta)}{1+r^2-2r\cos(\varphi-\theta)} d\theta \quad (2.85)$$

che risolve completamente il problema posto all'inizio.

Il risultato si generalizza immediatamente a circonferenze di raggio a arbitrario:

$$u(r, \varphi) = \frac{a^2 - r^2}{2\pi} \int_0^{2\pi} \frac{F(\theta)}{a^2 + r^2 - 2ra\cos(\varphi-\theta)} d\theta \quad (2.85')$$

Anche il problema di risolvere la (2.80) per regioni a contorno rettangolare può essere risolto con la separazione delle variabili $u(x, y) = X(x)Y(y)$ e mediante serie di Fourier. P.es., se il rettangolo è $[0, \pi] \times [0, 1]$ e si impone che la $u(x, y)$ sui lati "verticali" sia nulla: $u(0, y) = u(\pi, y) = 0$, si ricava facilmente che la soluzione dovrà avere la forma

$$u(x, y) = \sum_{n=1}^{\infty} \sin nx (a_n e^{ny} + b_n e^{-ny})$$

dove le a_n, b_n sono determinate dalle altre condizioni al bordo da assegnare sui lati "orizzontali" $u(x, 0)$ e $u(x, 1)$.

Il problema di trovare una funzione armonica u all'interno di una regione S essendo noto il suo valore F lungo il bordo ℓ si chiama **problema di Dirichlet**. Si chiama invece **problema di Neumann** il problema di determinare una funzione armonica u essendo noto il valore della sua *derivata normale* $G = \partial u / \partial n$ lungo il bordo.

Il procedimento visto in questo paragrafo può essere applicato, con le opportune modifiche, anche al problema di Neumann: naturalmente la funzione u risulterà determinata a meno di una costante additiva.

Si consideri p.es. il caso del cerchio: assegnata la condizione al bordo

$$G = G(\varphi) = \left. \frac{\partial u}{\partial r} \right|_{r=1}$$

cioè il valore della derivata radiale del potenziale sulla circonferenza, si dovrà uguagliare la derivata radiale $\partial u / \partial r$ dell'espressione (2.82) calcolata per $r = 1$ allo sviluppo in serie della $G(\varphi)$ rispetto al set $e^{im\varphi}$. Poiché nella derivata viene a mancare il termine costante, si vede che dovrà essere nullo il coefficiente g_0 dello sviluppo di $G(\varphi)$, cioè dovrà essere

$$g_0 = \frac{1}{2\pi} \int_0^{2\pi} G(\varphi) d\varphi = 0.$$

Ma in questo caso la condizione è automaticamente verificata, poiché g_0 è precisamente il flusso del campo elettrico attraverso la circonferenza, che è nullo grazie al teorema di Gauss (non ci sono cariche all'interno).

Il risultato è generale: la soluzione del problema di Neumann richiede che i valori della derivata normale assegnati lungo il bordo ℓ non possono essere arbitrari, ma devono soddisfare la condizione

$$\oint_{\ell} \frac{\partial u}{\partial n} d\ell = 0$$

ovvero, grazie ancora al teorema di Gauss:

$$\oint_{\ell} \frac{\partial u}{\partial n} d\ell = \iint_S \Delta u dS = 0 \quad (2.86)$$

dove S è la parte di piano racchiusa dalla linea ℓ , che risulta ovviamente soddisfatta se in S si ha $\Delta u = 0$.

2.25 Equazioni alle derivate parziali. Il metodo di d'Alembert

Abbiamo incontrato fin qui vari esempi di equazioni differenziali lineari alle derivate parziali del secondo ordine. La più generale equazione di questo tipo sarà evidentemente

$$L[u] = A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + F u = 0$$

dove $u = u(x, y)$ e $A, B, \dots$ sono costanti reali. A seconda che sia

$$B^2 - 4AC > 0, = 0, < 0$$

l'equazione si dice rispettivamente di tipo **iperbolico, parabolico, ellittico**.

La tipica equazione iperbolica è l'equazione di d'Alembert ($y = vt$):

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0.$$

Ponendo $\xi = x - vt$, $\eta = x + vt$, l'equazione diventa

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0 \quad (2.87)$$

che è la **forma canonica** dell'equazione di d'Alembert (e in generale delle equazioni iperboliche: le linee in cui $\xi = \text{cost.}$, $\eta = \text{cost.}$ si chiamano **linee caratteristiche** dell'equazione). La (2.87) è facilmente risolta:

$$u = \Phi(\xi) + \Psi(\eta) = \Phi(x - vt) + \Psi(x + vt)$$

dove Φ, Ψ sono funzioni arbitrarie. Dunque per questa via (*metodo di d'Alembert*) si vede che la più generale soluzione dell'equazione di d'Alembert è la sovrapposizione di due onde che si propagano lungo l'asse x con velocità v e $-v$. È facile collegare le funzioni Φ, Ψ con le condizioni iniziali $u(x, 0)$ e $u_t(x, 0)$ e al bordo. Infatti, p.es., se si impone la condizione ad un estremo $u(0, t) = 0$ (corda con un estremo fisso) segue che deve essere $\Phi(-vt) = -\Psi(vt)$, e questa – dovendo essere verificata per ogni t – implica che Ψ è il prolungamento “con legge dispari” della Φ . In altre parole, la soluzione diventa

$$u(x, t) = \Phi(x - vt) - \Phi(-x - vt)$$

cioè la sovrapposizione di un'onda progressiva e di una uguale onda regressiva “ribaltata”. Se si impone anche la condizione $u(L, t) = 0$ (secondo estremo fisso), si ottiene $\Phi(L - vt) = \Phi(-L - vt)$, che – dovendo nuovamente essere valida per ogni t – implica che la funzione Φ deve essere periodica di periodo $2L$, in accordo con quanto ottenuto nel § 2.4.

Un esempio tipico di equazione parabolica è dato dall'equazione del calore (2.19). L'equazione di Schrödinger (dipendente dal tempo) è del tipo

$$-\frac{\partial^2 \psi}{\partial x^2} + V(x) \psi = i \frac{\partial \psi}{\partial t}$$

ed è simile ad un'equazione parabolica anche se non rientra propriamente nella classificazione data sopra, a causa del coefficiente immaginario.

Infine, la tipica equazione ellittica è l'equazione di Laplace (2.80), già incontrata nel paragrafo precedente. Vedremo nel § 3.14 che lo studio di tale equazione può

essere affrontato anche con i metodi della variabile complessa. Altre tecniche generali di ricerca di soluzioni per tutte queste equazioni sono basate sull'uso delle trasformate integrali, come mostreremo nei Capitoli 4 e 5.

2.26 Funzionali. Teorema di Riesz

Negli ultimi paragrafi di questo Capitolo verranno presentate alcune nozioni più approfondite sugli operatori fra spazi di Hilbert.

Anzitutto, si chiama **funzionale** ogni trasformazione Φ definita da uno spazio di Hilbert H allo spazio dei numeri complessi:

$$\Phi : H \rightarrow \mathbb{C}.$$

Poiché lo spazio $\mathbb{C}$ è un particolare spazio di Hilbert, i funzionali sono dunque una speciale famiglia di operatori; risulta allora chiaro che cosa significa che un funzionale è lineare, continuo, limitato e il significato di dominio D di un funzionale. Anche l'equivalenza fra continuità e limitatezza è automaticamente valida per i funzionali, la cui norma sarà data da

$$\|\Phi\| = \sup_{x \in D} \frac{|\Phi(x)|}{\|x\|}.$$

Per esempio, sia $x_0 \in H$ un vettore fissato; per ogni $x \in H$ il prodotto scalare (x_0, x) definisce evidentemente un funzionale lineare e continuo di x : $\Phi(x) = (x_0, x)$. Si ha $\|\Phi\| = \|x_0\|$, come si vede immediatamente applicando la disuguaglianza di Schwarz e ricordando che $|(y, x)|$ è massimo quando $y \propto x$. Questo però non è soltanto un esempio particolare di funzionale continuo: si ha infatti il seguente fondamentale:

Teorema di Riesz. *Ogni funzionale lineare continuo Φ può essere espresso come un prodotto scalare con un vettore x_0 , univocamente individuato da Φ :*

$$\Phi(x) = (x_0, x). \quad (2.88)$$

Il vettore x_0 si chiama **vettore rappresentativo** di Φ , e si ha, come appena visto, $\|\Phi\| = \|x_0\|$. Esiste quindi una corrispondenza biunivoca fra i vettori di H e l'insieme dei suoi funzionali lineari e continui. Tale insieme si chiama **duale** di H e si indica usualmente con H' . Grazie al teorema di Riesz, si può dunque dire che gli spazi di Hilbert sono "autoduali".

La dimostrazione del teorema di Riesz si basa su questo procedimento: sia K il nucleo di Φ ed H_1 il suo complemento ortogonale. Si prendono due qualsiasi vettori $x_1, y_1 \in H_1$; sia $z = \alpha x_1 + \beta y_1$ la loro combinazione lineare con coefficienti $\alpha = \Phi(y_1)$, $\beta = -\Phi(x_1)$ (per ipotesi α, β sono diversi da zero), allora $\Phi(z) = 0$, dunque $z \in H_1 \cap K$, ma H_1 e K sono ortogonali e pertanto $z = 0$. Questo implica che x_1 e y_1 sono linearmente dipendenti e dunque H_1 è *unidimensionale*. Allora x_0 è una base per H_1 ; in parole diverse, il teorema di Riesz afferma che il nucleo K di

un funzionale lineare continuo ha *codimensione* 1 (la codimensione è la dimensione del sottospazio complemento ortogonale), essendo dato da tutti i vettori ortogonali a x_0 .

Se invece il funzionale non è limitato, il suo nucleo risulta *denso* in H .

Per esempio, il funzionale in ℓ^2 dato da $\Phi(a) = \sum_{n=1}^N a_n$ è limitato, e $x_0 = (1, 1, \dots, 1, 0, 0, \dots)$; invece $\Phi(a) = \sum_{n=1}^{\infty} a_n$ (definito in $\ell^2 \cap \ell^1$) non lo è: per il suo nucleo, v. le (2.38-38')

Il funzionale $\Phi(f) = f(a)$ in $L^2(I)$ (o meglio, così definito nel dominio delle funzioni continue in un intorno del punto $a \in I$) è non limitato, e il suo nucleo, dato dall'insieme delle funzioni continue $h(x)$ tali che $\Phi(h) = h(a) = 0$, è denso in $L^2(I)$, infatti si può certamente approssimare quanto si vuole in norma $L^2(I)$ qualunque $f \in L^2(I)$ con una funzione continua che si annulla in un punto (o anche in più punti e perfino in un insieme numerabile di punti ...).

Il funzionale in $L^2(\mathbf{R})$ dato da $\Phi(f) = \int_{-a}^a f(x) dx$ è limitato, mentre $\Phi(f) = \int_{-\infty}^{\infty} f(x) dx$ (definito in $L^2(\mathbf{R}) \cap L^1(\mathbf{R})$) non lo è, e le funzioni $f_0 \in L^2(\mathbf{R}) \cap L^1(\mathbf{R})$ tali che $\int_{-\infty}^{\infty} f_0(x) dx = 0$ sono dense in $L^2(\mathbf{R})$ ¹⁷.

2.27 Operatore aggiunto

Mediante il teorema di Riesz è immediato provare l'esistenza dell'**operatore aggiunto**, indicato con T^+ (oppure con $T^\dagger$), di ogni operatore *limitato* T fra due spazi

di Hilbert H e H' (cf. § 2.19). Sia infatti

$$T : H \rightarrow H'$$

una trasformazione lineare continua di H in H' . Per ogni fissato $y' \in H'$, il prodotto scalare (y', Tx) definisce, al variare di x in H , un funzionale di H :

$$(y', Tx) = \Phi_{y'}(x).$$

Si verifica subito che tale funzionale è lineare e continuo; allora, per il teorema di Riesz, esiste un vettore (ed uno solo) $y \in H$ tale che

$$\Phi_{y'}(x) = (y, x).$$

Resta così definita una corrispondenza $y' \rightarrow y$, e, posto $T^+ y' = y$, si scrive

$$(y', Tx) = (y, x) = (T^+ y', x)$$

esattamente come nella (2.51). Per questa via si conclude pure, come già anticipato (§ 2.19(a)), che l'operatore T^+ è lineare, che ha per dominio l'intero spazio H' e

¹⁷ Per verificarlo, si può osservare che (v. § 4.4) le trasformate di Fourier $\widehat{f_0}(\omega)$ di queste $f_0(x)$ sono funzioni continue e che $\int_{-\infty}^{\infty} f_0(x) dx = \widehat{f_0}(0) = 0$ dunque ...

che è anch'esso limitato ($\|T^+\| = \|T\|$). Inoltre, si ha $T^{++} = T$ e $(TS)^+ = S^+ T^+$ (dove S è un altro operatore limitato).

Se invece T non è limitato, non si può più applicare il teorema di Riesz per definire T^+ , ma si deve partire direttamente dalla (2.51), e inoltre in questo caso è necessario, come vedremo, fare l'ipotesi che il dominio D di T sia *denso* in H . Precisamente, fissato un $y' \in H$, si considera il prodotto scalare (y', Tx) al variare di x nel dominio D di T e si cerca *se esiste* un vettore $y \in H$ tale che, per ogni $x \in D$, risulti

$$(y', Tx) = (y, x). \quad (2.89)$$

In genere, non per tutti i vettori y' sarà possibile trovare un y che soddisfi tale uguaglianza (v. esempi ai §§ 2.19 (a), 2.28); certamente però quando – dato un y' – si trova in corrispondenza un y siffatto, questo è anche l'unico, infatti, se la (2.89) fosse verificata anche da un altro vettore $z \in H$:

$$(y', Tx) = (y, x) = (z, x), \quad \forall x \in D$$

sarebbe $(y - z, x) = 0$, $\forall x \in D$ da cui $y = z$ grazie all'ipotesi fatta sul dominio D , ovvero che $\overline{D} = H$. Resta così ben definita una corrispondenza $y' \rightarrow y$ e si scrive

$$y = T^+ y' \quad (2.90)$$

e questo definisce l'operatore lineare T^+ , con dominio D' costituito dai vettori y' che godono della proprietà detta. Tale dominio è comunque uno spazio vettoriale; se poi avviene che D' è denso in H , allora sarà definibile anche T^{++} , e si verifica subito che T^{++} è un'estensione di T , e si scrive $T^{++} \supset T$, cioè il dominio D'' di T^{++} contiene D e sui vettori $x \in D \subset D''$ si ha $Tx = T^{++}x$ ¹⁸.

2.28 Operatori chiusi

Si consideri in uno spazio di Hilbert H un punto arbitrario x_0 , e sia x_n una qualsiasi successione convergente a x_0 . Se T è un operatore continuo, questo implica, per definizione, che $Tx_n \rightarrow Tx_0$ per ogni successione $x_n \rightarrow x_0$. Se T non è continuo, ci sarà qualche successione $x_n \rightarrow x_0$ tale che la successione dei trasformati Tx_n non risulta convergente, ma può avvenire anche che esistano successioni x'_n e x''_n entrambe convergenti ad x_0 :

$$x'_n \rightarrow x_0, \quad x''_n \rightarrow x_0$$

(con $x'_n, x''_n, x_0 \in D = \text{dominio di } T$) che vengono trasformate da T in successioni convergenti a vettori *diversi*:

$$Tx'_n \rightarrow y', \quad Tx''_n \rightarrow y'' \neq y'.$$

¹⁸ Per maggiori dettagli, v. per esempio i testi di Abbati e Cirelli [23], di Akhiezer e Glazman [27], di Reed e Simon [19], o di Yoshida [31].

Un operatore T si dice **chiuso** se, $\forall x_0 \in D$, considerate le successioni $x_n \rightarrow x_0$ per le quali le successioni trasformate Tx_n risultano convergenti, avviene che tutte tali Tx_n convergono a Tx_0 (cioè non si hanno successioni Tx'_n e Tx''_n convergenti a vettori diversi anche se x'_n e x''_n erano convergenti allo stesso x_0).

Se poi avviene che per un certo x_0 (anche non appartenente a D), considerate tutte le successioni $x_n \rightarrow x_0$ per le quali Tx_n risulta convergere, si ha che tali successioni Tx_n convergono tutte allo stesso vettore z , allora T sarà “chiudibile”, cioè sarà possibile includere x_0 nel dominio di T definendo naturalmente (e senza ambiguità) $z := Tx_0$, e ottenere così una “estensione per chiusura” di T .

Un esempio semplice ma istruttivo: si può avere qualche dubbio se includere nel dominio dell'operatore di derivata d/dx funzioni del tipo $|x|$ o $x^{2/3}$. Si può allora “per sicurezza” decidere di definire d/dx solo nel dominio delle funzioni continue con derivata continua (ed ovviamente L^2); ma si può vedere che l'operatore risulta estendibile per chiusura anche alle funzioni continue con derivata non continua (ma L^2). Come verifica diretta, si vede p.es. che derivando termine a termine la serie di Fourier di $|x|$ in $L^2(-\pi, \pi)$ sulla base (2.9) si trova lo sviluppo dell'onda quadra; v. anche più avanti.

Un esempio di operatore non chiuso (e dunque non continuo) è l'operatore lineare T così definito sui vettori di un set ortonormale completo $\{e_n\}$

$$T e_n = e_1 \quad (n = 1, 2, \dots). \quad (2.91)$$

Sull'insieme F delle combinazioni lineari finite $x = \sum_m a_m e_m$ si ha $T(x) = (\sum_m a_m) e_1$; dunque il nucleo di T è denso in H ! (v. le (2.38-38')); notare che il set ortonormale completo $\{f_n\}$ dato nella (2.38) è contenuto in F e $T f_n = 0$, $\forall n = 1, 2, \dots$. L'operatore T è non chiuso: infatti p. es. le successioni $x'_n = e_1 + (1/n)e_n$ e $x''_n = e_1 - (e_2 + \dots + e_{n+1})/n$ tendono entrambe ad e_1 , ma $T(x'_n) \rightarrow e_1$ mentre $T(x''_n) = 0$.

L'operatore (2.91) presenta varie altre “patologie”: ad esempio si verifica subito che il vettore e_1 non appartiene al dominio dell'aggiunto T^+ : infatti poiché $(e_1, Tx) = (e_1, (\sum_m a_m) e_1) = \sum_m a_m$, è impossibile trovare un vettore y tale che $(e_1, Tx) = (y, x)$ (v. la (2.89)): dovrebbe essere $y = \sum_m e_m$!

Un altro esempio di operatore non chiuso è dato dal funzionale $\Phi(f) = f(a)$ già incontrato nel § 2.26. È infatti facile costruire successioni di funzioni (continue) $f_n(x)$ convergenti a 0 in senso L^2 ma tali che la successione numerica $f_n(a)$ non tenda a 0 ma tenda a qualsiasi valore, oppure non converga affatto.

Dalla definizione stessa di aggiunto, segue invece che l'aggiunto T^+ di qualsiasi operatore T è chiuso. Infatti, sia y'_n una successione contenuta nel dominio D' di T^+ , convergente ad un vettore y' e tale che la successione trasformata $T^+ y'_n$ risulti convergente:

$$y'_n \rightarrow y', \quad T^+ y'_n \rightarrow w.$$

Allora, per ogni x nel dominio D (denso in H , per ipotesi) di T , si ha

$$(y'_n, Tx) = (T^+ y'_n, x) \rightarrow (w, x),$$

ma anche

$$(y'_n, Tx) \rightarrow (y', Tx)$$

ovvero w soddisfa la condizione

$$(y', Tx) = (w, x) \quad \forall x \in D \quad (2.92)$$

il che significa – per la definizione stessa di aggiunto T^+ (v. le (2.89-90)) – che y' sta nel dominio D' di T^+ e che $w = T^+ y' = \lim_n T^+ y'_n$.

Ne segue, per esempio, che qualsiasi operatore della forma (2.63) è chiuso.

Per lo stesso motivo, ricordando quanto visto nel § 2.19, anche l'operatore di derivazione d/dx è un operatore chiuso (o estendibile per chiusura). In breve, ciò significa che derivando una successione (o una serie) di funzioni f_n (derivabili in L^2) ed L^2 -convergente si può ottenere una successione *non* L^2 -convergente (v. l'esempio già dato nel § 2.17: $f_n = \frac{1}{n} \sin nx$), ma *se la successione delle derivate f'_n è L^2 -convergente*, allora il limite è proprio la derivata del limite (v. anche quanto anticipato nei §§ 2.4, 2.5, 2.15).

Questa proprietà dell'operatore d/dx è importante anche perché giustifica certi passaggi dell'operatore di derivazione sotto il segno di serie (di Fourier) che si eseguono abitualmente quando, ad esempio, si cerca di risolvere per serie una equazione differenziale. Si consideri il semplice esempio seguente (oltre agli altri già incontrati e che ancora incontreremo). Si vuole trovare la soluzione “forzata” $y = y(t)$, periodica di periodo 2π , dell'equazione

$$\dot{y} + y = w(t) \quad (2.93)$$

dove $w(t)$ è una funzione assegnata, periodica di periodo 2π ; per esempio, la (2.93) può essere l'equazione di un circuito “serie RL ”, in cui $R = L = 1$, $w(t)$ è la tensione applicata e $y(t)$ la corrente che circola. Se in particolare $w(t)$ è una tensione “alternata pura” $w = w_k = e^{-ikt}$ con frequenza $\omega_k = k$, la soluzione sarà

$$y_k = \frac{1}{-ik + 1} e^{-ikt}.$$

Se $w(t)$ è una funzione periodica di $L^2(0, 2\pi)$, essa potrà essere scritta in serie di Fourier sulla base e^{-ikt} :

$$w(t) = \sum_{k=-\infty}^{+\infty} c_k e^{-ikt}$$

e la soluzione si potrà scrivere sotto forma di serie

$$y(t) = \sum_{k=-\infty}^{+\infty} \frac{c_k}{-ik + 1} e^{-ikt}. \quad (2.94)$$

Questa serie risulta convergente, e inoltre si può derivare termine a termine in modo da verificare che la (2.93) è soddisfatta. Il risultato precedente sull'operatore di derivazione ci assicura infatti la possibilità di derivare termine a termine una serie ogni volta che la serie delle derivate risulti convergente in senso L^2 .

Da notare anche che la serie (2.94) converge ad una funzione continua: un modo elegante per dimostrarlo può essere quello di provare che la serie converge

totalmente (v. § 2.3), ovvero che

$$\sum_{k=-\infty}^{+\infty} \left| \frac{c_k}{1-ik} \right| < \infty \quad (2.94')$$

ma questo è certamente vero perché la (2.94') può essere vista come il prodotto scalare in ℓ^2 fra le due successioni $\{|c_k|\}$ e $\{1/|1-ik|\}$, entrambe $\in \ell^2$.

2.29 Varie nozioni di convergenza per successioni di vettori e operatori

È già stata data (e usata varie volte!) la seguente definizione di convergenza per una successione di vettori x_n in uno spazio di Hilbert:

$$x_n \rightarrow x \quad \text{se} \quad \|x_n - x\| \rightarrow 0$$

(è sempre sottinteso per $n \rightarrow \infty$). Oltre a questa, che si chiama anche **convergenza in norma**, è possibile introdurre una diversa nozione di convergenza, che si chiama: **convergenza debole**: si dice che x_n converge debolmente a $x \in H$ se

$$\forall y \in H \quad \text{si ha} \quad (y, x_n) \rightarrow (y, x).$$

È immediato vedere che la convergenza in norma di una successione x_n ad un vettore x implica anche la sua convergenza debole allo stesso vettore. Invece, una successione può convergere debolmente ma *non* in norma: per esempio, una successione di infiniti vettori e_n di un set ortonormale (non necessariamente completo) converge debolmente a zero, dato che per ogni $y \in H$ si ha, dovendo sussistere la disuguaglianza di Bessel (2.34) (o l'uguaglianza di Parseval (2.36))

$$|(y, e_n)| \rightarrow 0$$

ma non è convergente in norma (in effetti, $\|e_n - e_m\| = \sqrt{2}$ se $n \neq m$).

Si può dimostrare che ogni spazio di Hilbert, che *per definizione* è completo rispetto alla convergenza in norma, è anche **debolmente completo**, ovvero ogni successione che sia “debolmente di Cauchy” (cioè $|(y, x_n - x_m)| \rightarrow 0$) possiede un limite $x \in H$ a cui converge debolmente.

Varie nozioni di convergenza si possono introdurre anche per le successioni di operatori T_n (considereremo solo il caso di operatori limitati):

(i) **convergenza in norma**: si dice che $T_n \rightarrow T$ in norma se

$$\|T_n - T\| \rightarrow 0;$$

(ii) **convergenza forte**: si ha convergenza in senso forte se per ogni vettore $x \in H$ avviene che

$$\|(T_n - T)x\| \rightarrow 0;$$

(iii) **convergenza debole**: si parla di convergenza in senso debole se per ogni coppia di vettori $x, y \in H$ si ha

$$(x, T_n y) \rightarrow (x, T y).$$

Si verifica subito che se $T_n \rightarrow T$ in norma, allora anche converge in senso forte e debole allo stesso T , e così pure la convergenza forte implica quella debole.

Viceversa, invece, sia per esempio P_n il proiettore che proietta sul sottospazio n -dimensionale generato dai primi n vettori di un set ortonormale completo e_n nello spazio H : dunque $P_n x = \sum_{m=1}^n a_m e_m$, dove naturalmente a_m sono le componenti di $x \in H$ rispetto al set e_n . È abbastanza naturale ipotizzare che – in qualche senso – P_n tenda all'identità $\mathbf{I}$ per $n \rightarrow \infty$. Si vede subito che P_n tende infatti all'identità in senso forte:

$$\|(\mathbf{I} - P_n)x\|^2 = \sum_{m>n} |a_m|^2 \rightarrow 0$$

grazie all'identità di Parseval. Ma la successione P_n non ha limite in norma: infatti, il limite in norma, se esistesse, dovrebbe essere ancora $\mathbf{I}$, si ha invece

$$\sup_x \frac{\|(\mathbf{I} - P_n)x\|^2}{\|x\|^2} = \sup_x \frac{\sum_{m>n} |a_m|^2}{\sum_{m=1}^{\infty} |a_m|^2} = 1$$

(basta prendere un vettore $\tilde{x}$ con le prime n componenti nulle; oppure osservare che $\mathbf{I} - P_n$ è un proiettore!).

La stessa cosa si verifica per esempio per la successione di operatori in $L^2(0, 1)$ definiti da $T_n f(x) = x^n f(x)$, che tendono a 0 in senso forte (usare il teorema di Lebesgue, v. § 2.8) ma non in norma.

Infine, se T_n è l'operatore lineare così definito

$$T_n x = (e_n, x)y$$

dove $\{e_n\}$ è un set ortonormale (non necessariamente completo) e x, y sono due vettori fissati. Si vede che $T_n \rightarrow 0$ debolmente, grazie all'identità di Parseval o alla disuguaglianza di Bessel, ma non converge in senso forte.

È interessante notare che un operatore T come nella (2.63) può essere scritto sotto forma di una *decomposizione spettrale* (cf. con la (1.34)):

$$T = \sum_{n=1}^{\infty} c_n P_n$$

in cui la serie è convergente in norma se $|c_n| \rightarrow 0$, mentre è convergente solo in senso forte se i c_n sono limitati, cioè se esiste una costante M tale che $|c_n| \leq M$.

Da notare infine che lo spazio vettoriale $\mathcal{L}(H)$ degli operatori lineari e continui di H in sé, è *completo* rispetto a tutte e tre le nozioni di convergenza ora definite.

2.30 Operatori compatti

Un operatore si dice **compatto** se, data comunque una successione di vettori x_n *debolmente* convergente ad x , si ha che la successione dei vettori trasformati $T x_n$ converge in norma a $T x$, cioè $\|T x_n - T x\| \rightarrow 0$.

Evidentemente ogni operatore compatto è anche continuo, ma non è vero il viceversa (per esempio l'identità è continua ma non compatta!).

Un esempio tipico di operatore compatto è dato da

$$T e_n = \frac{1}{n} e_n \quad n = 1, 2, \dots$$

dove e_n è un set ortonormale completo in uno spazio di Hilbert. Un altro esempio è l'operatore integrale dato alla fine del § 2.21. È pure compatto l'inverso di ogni operatore di Sturm–Liouville (purché $\lambda = 0$ non sia autovalore di quest'ultimo). Infine, un proiettore è compatto se e solo se il sottospazio su cui proietta ha dimensione *finita*: infatti se tale sottospazio avesse dimensione infinita, qualunque successione di vettori ortonormali contenuta in esso (e dunque debolmente convergente a zero) sarebbe mandata in se stessa dal proiettore.

L'importanza degli operatori compatti sta nel seguente

Teorema. *Ogni operatore compatto T autoaggiunto (o più in generale normale: $T^+T = TT^+$) possiede un set ortogonale e completo di autovettori u_n :*

$$T u_n = \lambda_n u_n.$$

Inoltre, ogni autovalore non nullo ha necessariamente degenerazione finita, e – se ci sono infiniti autovalori $\lambda_n \neq 0$ – si ha che $\lambda_n \rightarrow 0$ per $n \rightarrow \infty$ ($\lambda = 0$ può essere o non essere autovalore di T). Ed è vero anche il viceversa: se un operatore ammette un set ortogonale completo di autovettori con autovalori che soddisfano le proprietà suddette, l'operatore è compatto.

Da notare, in questo teorema, che se $\lambda = 0$ non è autovalore, allora T ha immagine *densa* in H (ma non coincidente con l'intero spazio) e T^{-1} esiste ma non è limitato (per esempio l'operatore suggerito qui sopra) e $\sigma = 0$ appartiene al suo spettro.

Metodi matematici della Fisica

Cicogna, G.

2015, X, 258 pagg. 22 figg., Hardcover

ISBN: 978-88-470-5683-1