

Appunti di Metodi Matematici

Fabio Ortolani

19 febbraio 2013

Indice

1	Funzioni olomorfe	1
1.1	Il piano complesso	1
1.1.1	Una curiosità: i quaternioni	7
1.2	Funzioni di variabile complessa	12
1.2.1	L'esponenziale	13
1.2.2	Funzioni trigonometriche	14
1.2.3	Funzioni iperboliche	16
1.2.4	Differenziabilità ed olomorfismo	17
1.2.5	Condizioni di olomorfismo	19
1.2.6	Curve e domini	23
1.2.7	Integrali di contorno	26
1.3	Teorema di Cauchy	31
1.3.1	Rappresentazione integrale di Cauchy	35
1.3.2	Rappresentazioni integrali	39
1.3.3	Teorema di Morera	42
1.3.4	La serie di Taylor.	43
1.3.5	La serie di Laurent.	45
1.3.6	Zeri e punti singolari	49
1.3.7	Comportamento locale di una funzione analitica	52
1.4	Residui.	54
1.4.1	Teorema dei residui	57
1.4.2	Calcolo dei residui	58
1.4.3	Residuo all'infinito	60
1.4.4	Indicatore logaritmico	62
1.4.5	Calcolo di integrali definiti col metodo dei residui	64
1.4.6	Poli semplici in prossimità del cammino di integrazione.	71
1.5	Cenni sulle funzioni polidrome.	79
1.5.1	Logaritmo.	79
1.5.2	La radice quadrata.	83
1.5.3	Calcolo di integrali che coinvolgono funzioni polidrome.	85
1.6	Esercizi	91
1.6.1	Soluzioni	95
2	Topologia.	113
2.1	Spazi topologici.	113
2.1.1	Insiemi aperti	114
2.1.2	Insiemi chiusi	116
2.1.3	Topologia indotta e topologia generata	118

2.1.4	Punti interni, esterni, di frontiera	119
2.1.5	Sistemi fondamentali di intorni	121
2.2	Spazi metrici	125
2.3	Proprietà topologiche	131
2.3.1	Insiemi compatti	131
2.3.2	Spazi di Hausdorff	132
2.3.3	Densità e separabilità	137
2.4	Continuità	140
2.4.1	Continuità e compattezza	145
2.5	Successioni	147
2.5.1	Completezza	153
2.6	Esercizi	162
2.6.1	Soluzioni	164
3	Spazi lineari.	171
3.1	Strutture algebriche	171
3.1.1	Spazi vettoriali	173
3.1.2	Sottospazi	176
3.1.3	Dipendenze lineari e basi	177
3.2	Applicazioni lineari	186
3.3	Somme dirette	192
3.3.1	Dualità	199
3.4	Spazi normati	203
3.4.1	Norme	203
3.4.2	Spazi di Banach	207
3.4.3	Serie infinite di vettori	210
3.4.4	Applicazioni lineari tra spazi normati	213
3.4.5	Norma operatoriale	219
3.4.6	Norme equivalenti	221
3.5	Spazi L^p	226
3.5.1	Disuguaglianza di Hölder	228
3.5.2	Disuguaglianza di Minkowski	232
3.5.3	Spazi l^p	235
3.5.4	Appendice: \limsup e \liminf	238
3.6	Esercizi	241
3.6.1	Soluzioni	246
4	Spazi di Hilbert.	259
4.1	Prodotti scalari	259
4.1.1	Forme sesquilineari	259
4.1.2	Prodotto scalare	263
4.2	Geometria di uno spazio di Hilbert	268
4.2.1	Ortogonalità	268

4.2.2	Decomposizione in sottospazi ortogonali	270
4.2.3	Duale topologico	274
4.3	Sistemi ortonormali.	277
4.3.1	Basi hilbertiane	283
4.3.2	Serie di Fourier.	288
4.3.3	Polinomi ortogonali.	296
4.4	Esercizi	305
4.4.1	Soluzioni	307
5	Operatori	315
5.1	Applicazioni lineari limitate tra spazi normati.	315
5.1.1	Convergenza forte di applicazioni continue.	320
5.1.2	Estensione continua di applicazioni continue.	327
5.2	Serie operatoriali.	329
5.2.1	Serie esponenziale	333
5.2.2	Serie di Neumann	335
5.2.3	Serie di potenze di un operatore	336
5.3	Operatore aggiunto	338
5.3.1	Operatori autoaggiunti	342
5.3.2	Operatori unitari	343
5.3.3	Operatori di proiezione	346
5.4	Trasformate di Fourier	350
5.4.1	Convoluzione di funzioni.	353
5.4.2	Funzioni a decrescenza rapida.	357
5.4.3	Formula di inversione.	361
5.4.4	Trasformate di Fourier in L^2	365
5.5	Esercizi	371
5.5.1	Soluzioni	374

Capitolo 1.

Funzioni olomorfe

1.1 Il piano complesso

L'estensione dal sistema dei numeri reali al sistema dei numeri complessi è molto importante sia in matematica che in fisica. L'area della fisica dove i numeri complessi sono praticamente indispensabili è la meccanica quantistica, in quanto la funzione d'onda associata ad un sistema fisico è generalmente a valori complessi. In matematica il loro uso è diffuso in quasi tutte le aree.

Non vogliamo trattare le teorie algebriche che introducono i numeri complessi in maniera astratta, ma li consideriamo come una estensione naturale dei numeri reali, basata sull'esistenza di un numero, l'**unità immaginaria**, che indicheremo sempre con la lettera i e che soddisfa la regola algebrica:

$$i^2 = -1. \quad (1.1)$$

La qualifica “immaginaria” denota la natura puramente matematica di tale numero, in quanto nessuna esperienza fisica fornisce come risultato di una misura il numero i , ma sempre uno o più numeri reali. Combinando aritmeticamente (cioè tramite le operazioni elementari: addizione, sottrazione, moltiplicazione e divisione) l'unità immaginaria con gli ordinari numeri reali otteniamo l'insieme dei numeri complessi. Possiamo allora assumere che un **numero complesso** z è univocamente determinato da una coppia di numeri reali x, y ed useremo la notazione (**forma cartesiana**):

$$z = x + iy, \quad x, y \in \mathbb{R}. \quad (1.2)$$

Con tale rappresentazione le operazioni con i numeri complessi sono eseguite con le usuali regole aritmetiche, tenendo però sempre presente la convenzione (1.1) precedente. Non ci sono novità sostanziali nella struttura algebrica dei numeri complessi rispetto a quella dei numeri reali. L'unità moltiplicativa rimane l'unità reale 1, l'elemento neutro additivo coincide lo zero reale 0. Dati due numeri complessi:

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2,$$

la somma e la moltiplicazione risultano:

$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2) + i(y_1 + y_2), \\ z_1 \cdot z_2 &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \end{aligned}$$

Lasciamo come esercizio per il lettore la determinazione della differenza e della divisione tra due numeri complessi.

I numeri reali x e y della rappresentazione (1.2) sono detti rispettivamente **parte reale** e **parte immaginaria** di z e indicati anche con:

$$\begin{cases} \operatorname{Re} z \equiv x \\ \operatorname{Im} z \equiv y. \end{cases} \quad (1.3)$$

Dato un numero complesso $z = x + iy$, il numero complesso:

$$\bar{z} = x - iy, \quad (1.4)$$

viene detto **compleSSo coniugato** di z , o semplicemente **coniugato** e indicato a volte con z^* . Algebricamente abbiamo:

$$\begin{cases} \operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \\ \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}). \end{cases} \quad (1.5)$$

L'insieme di tutti i numeri complessi viene indicato tradizionalmente con il simbolo \mathbb{C} (in analogia con la notazione \mathbb{R} per i numeri reali). Chiaramente \mathbb{C} risulta isomorfo a \mathbb{R}^2 e i numeri x e y della forma cartesiana possono essere visti come coordinate cartesiane ortogonali di un punto del piano, per cui si è soliti visualizzare i numeri complessi con un piano, detto **piano complesso** o di **Gauss**¹, oppure di **Argand**². In tale piano possiamo identificare i numeri reali (del tipo $x + i0$) con l'asse x (**asse reale**) e i numeri **immaginari puri** (del tipo $0 + iy$) con l'asse y (**asse immaginario**).

¹ Carl Friedrich Gauss (Braunschweig, 30 aprile 1777 – Gottinga, 23 febbraio 1855) è stato un matematico, astronomo e fisico tedesco, che ha fornito contributi determinanti all'analisi matematica, teoria dei numeri, calcolo numerico, geometria differenziale, geodesia, magnetismo e ottica. Talvolta definito “il più grande matematico della modernità” in contrasto ad Archimede, che egli stesso considerava il più grande dei matematici per quanto riguardava “l'antichità”, e il “principe della matematica”, Gauss è ricordato tra i più importanti matematici della storia, avendo contribuito in modo decisivo all'evoluzione delle scienze matematiche, fisiche e naturali.



² Jean-Robert Argand (Ginevra, 18 luglio 1768 – Parigi, 13 agosto 1822) è stato un matematico svizzero non professionista. Nel 1806, mentre era il gestore di una libreria a Parigi, pubblicò a proprie spese un libro in cui veniva esposta l'idea dell'interpretazione geometrica dei numeri complessi, oggi nota come diagramma di Argand. Tra i suoi contributi occorre anche ricordare una dimostrazione (non completamente corretta) del Teorema fondamentale dell'algebra; Argand sembra essere stato il primo a trattare il caso in cui i coefficienti di un polinomio possono anche essere numeri complessi.



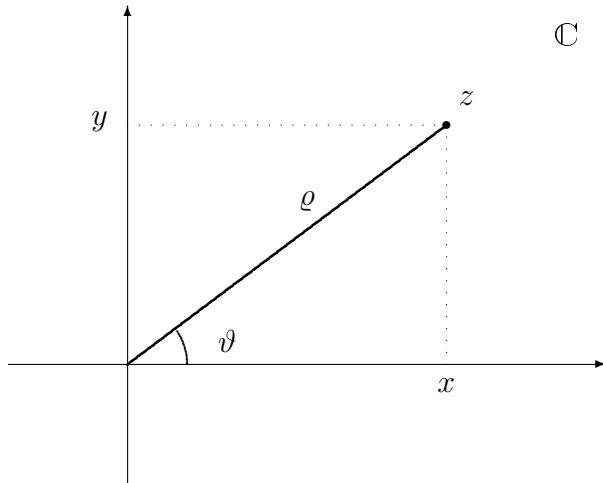


Figura 1.1: Il piano complesso.

La visualizzazione geometrica dei numeri complessi permette di introdurre in maniera naturale la **forma polare** di un numero complesso, considerando le coordinate polari nel piano (Fig. 1.1):

$$\begin{cases} x = \varrho \cos \vartheta \\ y = \varrho \sin \vartheta \end{cases} \quad -\pi < \vartheta \leq \pi, \quad \varrho \geq 0, \quad (1.6)$$

$$\begin{cases} \varrho = \sqrt{x^2 + y^2} \\ \vartheta = \arctan(x, y). \end{cases} \quad -\pi < \vartheta \leq \pi. \quad (1.7)$$

e quindi:

$$z = \varrho (\cos \vartheta + i \sin \vartheta). \quad (1.8)$$

ϱ viene detto il **modulo** del numero complesso z , e viene indicato con $|z|$, mentre ϑ è detto **argomento** o **fase** di z . Risulta chiaro che la fase è in realtà determinata a meno di un multiplo intero di 2π per la periodicità delle funzioni trigonometriche, e l'argomento propriamente detto risulta:

$$\arg z = \vartheta + 2k\pi, \quad k = 0, \pm 1, \pm 2, \dots \quad (1.9)$$

se ϑ lo consideriamo compreso tra $-\pi$ e π . Generalmente si indica la scelta $-\pi < \arg z \leq \pi$ (oppure $0 \leq \arg z < 2\pi$) come determinazione principale dell'argomento.

Osservazione. Si noti che la funzione $\arctan(x, y)$ non coincide esattamente con la funzione arcotangente della trigonometria $\arctan(\frac{y}{x})$, intesa come funzione inversa della tangente. La funzione $\arctan(x, y)$ è “sensibile” al quadrante in cui si trova il punto (x, y) e restituisce un angolo all’interno di un intervallo di ampiezza 2π , mentre $\arctan(\frac{y}{x})$

restituisce un angolo in un intervallo di ampiezza π . Ad esempio:

$$\begin{aligned}\arctan(1, 1) &= \frac{\pi}{4} + 2n\pi, & \arctan(-1, -1) &= -\frac{3\pi}{4} + 2n\pi \\ \arctan\left(\frac{1}{1}\right) &= \arctan\left(\frac{-1}{-1}\right) = \tan^{-1}(1) = \frac{\pi}{4} + n\pi\end{aligned}$$

Notiamo inoltre che l'unico punto in cui non è definita $\arctan(x, y)$ è l'origine $(0, 0)$, mentre formalmente $\arctan\left(\frac{y}{x}\right)$ non è definita per $x = 0$ (anche se il limite per $x \rightarrow 0$ esiste quando $y \neq 0$).

Il modulo di un numero complesso gode di proprietà simili al modulo o valore assoluto di un numero reale. Una scrittura alternativa per il modulo di $z = x + iy$ è:

$$|z| = |x + iy| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} \quad (1.10)$$

ed è facile verificare che:

$$|z_1 z_2| = |z_1| |z_2| \quad (1.11)$$

$$|z_1 + z_2| \leq |z_1| + |z_2|, \quad (\textbf{diseguaglianza triangolare}) \quad (1.12)$$

Il modulo costituisce anche un caso particolare del concetto di norma (di un vettore in uno spazio normato) che vedremo più avanti.

Osservazione. Un numero reale x si può sempre pensare come numero complesso con una parte immaginaria nulla e le definizioni di modulo, sia come numero reale, sia come numero complesso, sono consistenti tra loro:

$$|x| = \sqrt{x^2} = \sqrt{x^2 + 0^2} = |x + i0|, \quad x \in \mathbb{R}$$

Ciò porta alla nomenclatura di **valore assoluto** di un numero complesso per indicare il suo modulo, come estensione del concetto di valore assoluto introdotto nei numeri reali.

Tornando alla rappresentazione polare e ricordando gli sviluppi in serie delle funzioni trigonometriche (convergenti per ogni valore di ϑ):

$$\begin{aligned}\cos \vartheta &= \sum_{k=0}^{\infty} (-1)^k \frac{\vartheta^{2k}}{(2k)!} \\ \sin \vartheta &= \sum_{k=0}^{\infty} (-1)^k \frac{\vartheta^{2k+1}}{(2k+1)!}\end{aligned}$$

abbiamo formalmente:

$$\begin{aligned}\cos \vartheta + i \sin \vartheta &= \sum_{k=0}^{\infty} (-1)^k \frac{\vartheta^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{\vartheta^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} (i)^{2k} \frac{\vartheta^{2k}}{(2k)!} + \sum_{k=0}^{\infty} (i)^{2k+1} \frac{\vartheta^{2k+1}}{(2k+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(i\vartheta)^n}{n!} = e^{i\vartheta}\end{aligned}$$

Vedremo in seguito come definire in generale l'esponenziale di un numero complesso e la convergenza della corrispondente serie. La relazione sopra, che può essere assunta come definizione di $e^{i\vartheta}$ con ϑ reale, è nota come **formula di Eulero**³:

$$e^{i\vartheta} = \cos \vartheta + i \sin \vartheta, \quad (\vartheta \text{ reale}) \quad (1.13)$$

e sono garantite le usuali proprietà algebriche della funzione esponenziale. Per un numero complesso abbiamo quindi la seguente scrittura in **forma polare**:

$$z = \varrho e^{i\vartheta}. \quad (1.14)$$

Si può facilmente verificare che:

$$\bar{z} = \varrho e^{-i\vartheta}. \quad (1.15)$$

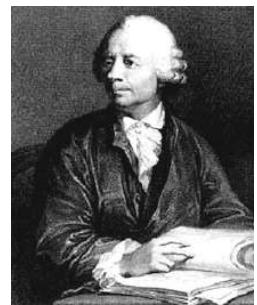
Formalmente, data una espressione complessa, l'operazione di coniugazione equivale a cambiare ogni occorrenza dell'unità immaginaria col suo opposto:

$$i \longleftrightarrow -i.$$

Mentre la rappresentazione cartesiana (1.2) risulta molto comoda quando si eseguono addizioni e sottrazioni tra numeri complessi, la forma polare risulta invece utile per eseguire moltiplicazioni e divisioni.

$$\left\{ \begin{array}{l} z_1 = \varrho_1 e^{i\vartheta_1} \\ z_2 = \varrho_2 e^{i\vartheta_2} \end{array} \right. \implies \left\{ \begin{array}{l} z_1 z_2 = \varrho_1 \varrho_2 e^{i(\vartheta_1+\vartheta_2)} \\ \frac{z_1}{z_2} = \frac{\varrho_1}{\varrho_2} e^{i(\vartheta_1-\vartheta_2)} \end{array} \right.$$

³ **Leonhard Euler** (Basilea, 15 aprile 1707 – San Pietroburgo, 18 settembre 1783) noto in Italia come Eulero, è stato un matematico e fisico svizzero. È considerato il più importante matematico dell'Illuminismo. Allievo di Johann Bernoulli, è noto per essere tra i matematici scientificamente più prolifici di tutti i tempi. Ha fornito contributi storicamente cruciali in svariate aree: analisi infinitesimale, funzioni speciali, meccanica razionale, meccanica celeste, teoria dei numeri, teoria dei grafi. Complessivamente esistono 886 pubblicazioni di Eulero. Buona parte della simbologia matematica tutt'ora in uso è stata introdotta da Eulero, per esempio Σ come simbolo per la sommatoria, $f(x)$ per indicare una funzione, la lettera π per indicare il numero pi greco.



La formula di Eulero può essere invertita fornendo le seguenti relazioni:

$$\boxed{\begin{aligned}\cos \vartheta &= \frac{e^{i\vartheta} + e^{-i\vartheta}}{2} \\ \sin \vartheta &= \frac{e^{i\vartheta} - e^{-i\vartheta}}{2i}\end{aligned}} \quad (1.16)$$

Inoltre abbiamo la seguente relazione (**formula di De Moivre**⁴):

$$\boxed{\cos(n\vartheta) + i \sin(n\vartheta) = e^{in\vartheta} = (e^{i\vartheta})^n = (\cos \vartheta + i \sin \vartheta)^n} \quad (1.17)$$

per cui, sviluppando la potenza del binomio ed eguagliando tra loro le parti reali e le parti immaginarie otteniamo facilmente le espressioni per i seni e i coseni dei multipli di un angolo. Ad esempio:

$$\begin{aligned}\cos(3\vartheta) + i \sin(3\vartheta) &= (\cos \vartheta + i \sin \vartheta)^3 \\ &= \cos^3 \vartheta - 3 \cos \vartheta \sin^2 \vartheta + i (3 \cos^2 \vartheta \sin \vartheta - \sin^3 \vartheta), \\ \cos 3\vartheta &= \cos^3 \vartheta - 3 \cos \vartheta \sin^2 \vartheta = 4 \cos^3 \vartheta - 3 \cos \vartheta, \\ \sin 3\vartheta &= 3 \cos^2 \vartheta \sin \vartheta - \sin^3 \vartheta = -4 \sin^3 \vartheta + 3 \sin \vartheta.\end{aligned}$$

Avendo identificato l'insieme dei numeri complessi con i punti del piano complesso le usuali nozioni topologiche ed insiemistiche in \mathbb{R}^2 si trasportano in maniera ovvia su \mathbb{C} . Notiamo inoltre che la distanza euclidea tra due punti $(x_1, y_1), (x_2, y_2)$ diviene semplicemente il modulo della differenza tra i corrispondenti numeri complessi:

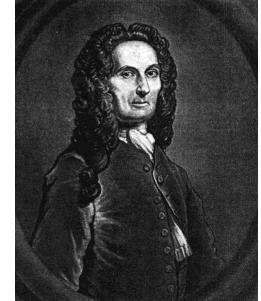
$$|z_1 - z_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \quad z_1 = x_1 + i y_1, \quad z_2 = x_2 + i y_2.$$

La topologia in \mathbb{C} permette quindi i concetti di punto di accumulazione, punto interno ed esterno ad un insieme, come pure i concetti di convergenza per una successione o per una serie.

Osservazione. Risulta chiaro, dalla relazione fondamentale della trigonometria, che $e^{i\vartheta}$ (con ϑ reale) rappresenta un numero complesso a modulo uno, o unitario:

$$\boxed{|e^{i\vartheta}| = \sqrt{\cos^2 \vartheta + \sin^2 \vartheta} = 1, \quad \vartheta \in \mathbb{R}} \quad (1.18)$$

⁴ **Abraham de Moivre** (Vitry-le-François, 26 maggio 1667 – Londra, 27 novembre 1754) è stato un matematico francese, amico di Isaac Newton. È noto per la formula di de Moivre (che collega i numeri complessi con la trigonometria), i suoi lavori sulla distribuzione normale e la teoria della probabilità, nonché scoprì (anche se in forma incompleta) l'approssimazione di Stirling. Quest'ultima venne usata nel 1733 da de Moivre per derivare la variabile casuale normale come una approssimazione della variabile casuale binomiale. Successivamente de Moivre accreditò a James Stirling i miglioramenti apportati alla formula.



ed in particolare si ha:

$$1 = e^{i0}, \quad i = e^{i\frac{\pi}{2}}, \quad -1 = e^{i\pi}, \quad -i = e^{i\frac{3\pi}{2}}$$

Notiamo inoltre che, dato un numero complesso z , la moltiplicazione per un fattore di fase $e^{i\theta}$, con θ reale, $ze^{i\theta}$, ruota il punto z di un angolo θ nel piano complesso attorno all'origine.

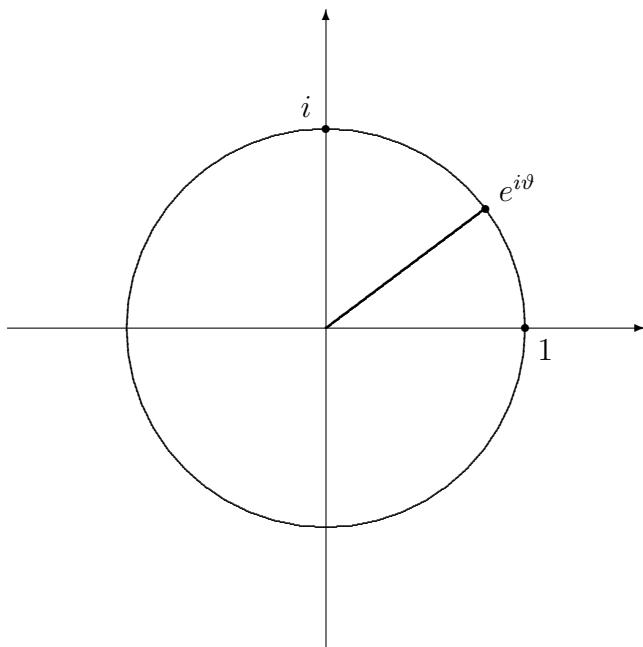


Figura 1.2: Il cerchio unitario.

1.1.1 Una curiosità: i quaternioni

Ci potremmo chiedere a questo punto il perchè di tutte queste puntualizzazioni e definizioni; sembra alla fine che abbiamo fornito solo delle ricette di cucina utilizzando cose abbastanza ovvie. In realtà una ragione per tutto questo è data dal fatto che si può procedere ulteriormente nel considerare tipi ancora più generali di “campi di numeri”, in cui alcune delle proprietà algebriche elementari cessano di valere (l’unità immaginaria, ad esempio, viola la positività del quadrato di un numero). Diventa allora molto importante formalizzare la trattazione in maniera adeguata in modo che sia chiaro che cosa è vero e che cosa non lo è. Ad esempio, la successiva estensione del concetto di numero oltre ai

numeri complessi è data dai **quaternioni**, introdotti da Hamilton⁵ nel 1843 ed applicati con successo alla meccanica nello spazio tridimensionale. La costruzione dei quaternioni è basata sull'assunzione di tre indipendenti unità immaginarie, indicate di solito con i , j e k , che violano la proprietà commutativa della moltiplicazione (caratteristica dei numeri reali e complessi) e soddisfano le regole algebriche:

$$\begin{aligned} i^2 &= j^2 = k^2 = -1, \\ ij &= -ji = k, \quad jk = -kj = i, \quad ki = -ik = j \end{aligned} \tag{1.19}$$

Un quaternione q è allora una quantità della forma:

$$q = q_0 + q_1 i + q_2 j + q_3 k, \tag{1.20}$$

con q_0 , q_1 , q_2 e q_3 numeri reali. Questa costruzione ci permette di identificare i quaternioni con lo spazio \mathbb{R}^4 , e viene usata la notazione \mathbb{H} per indicarne anche la struttura algebrica. La componente q_0 viene identificata con la parte reale del quaternione (o parte scalare, con una terminologia fisica), mentre la parte rimanente è detta parte immaginaria (o parte vettoriale dai fisici). Hamilton le identificava rispettivamente come quaternione scalare e quaternione proprio.

Possiamo introdurre una operazione di **coniugazione**, in cui i segni di i , j e k sono cambiati:

$$\bar{q} = q_0 - q_1 i - q_2 j - q_3 k. \tag{1.21}$$

Il **modulo** $|q|$ di un quaternione è un numero reale non negativo, definito come la radice quadrata di:

$$|q|^2 \equiv \bar{q}q = q\bar{q} = q_0^2 + q_1^2 + q_2^2 + q_3^2, \tag{1.22}$$

e chiaramente si annulla solo se $q = 0$.

Ci possiamo ora chiedere quali delle proprietà tipiche dei numeri complessi sono valide anche per i quaternioni. La risposta non è così ovvia ed occorre procedere con verifiche opportune. Per rendere la cosa facile è necessario formalizzare bene la struttura matematica e la risposta sostanzialmente è data dalla sola perdita della commutatività moltiplicativa, cioè in generale $qq' \neq q'q$. Però viene mantenuta una proprietà moltiplicativa importante:

$$|q q'| = |q| |q'|, \tag{1.23}$$

⁵ Sir William Rowan Hamilton (Dublino, Irlanda, 4 agosto 1805 – Dublino, Irlanda, 2 settembre 1865) fu un matematico accreditato della scoperta dei quaternioni, la prima algebra non commutativa. Egli inventò anche nuovi metodi in meccanica e per questo la funzione che rappresenta l'energia e regola la dinamica viene chiamata Hamiltoniana in suo onore. Lo studio della dinamica di Hamilton fu significativo per lo sviluppo dell'operatore hamiltoniano in meccanica quantistica. Hamilton mostrò il suo immenso talento in età precoce, come notò John Breinkley, astronomo e vescovo di Cloyne, nel 1823, quando Hamilton aveva diciotto anni: “Questo giovane, non dico sarà, ma è, il matematico migliore della sua età.”



verificabile in base alle definizioni precedenti.

La non commutatività della moltiplicazione pone qualche problema nella definizione della divisione tra quaternioni. In effetti questo problema è stato proprio quello che ha portato Hamilton alla costruzione dei quaternioni. Egli aveva notato che l'isomorfismo tra i numeri complessi e il piano permetteva di costruire una struttura aritmetica (le quattro operazioni: addizione, sottrazione, moltiplicazione e divisione) tra vettori bidimensionali e voleva generalizzare tali operazioni elementari per i vettori dello spazio tridimensionale. Ebbe l'intuizione che tale possibilità esisteva per vettori a quattro dimensioni mentre passeggiava con la moglie e scrisse l'equazione fondamentale dei quaternioni:

$$i^2 = j^2 = k^2 = i j k = -1 \quad (1.24)$$

equivalente alle relazioni (1.19) su un lato del ponte Brougham (noto ora come Broom Bridge) a Dublino.

La proprietà (1.22) permette la costruzione in maniera univoca del **reciproco** di un quaternione:

$$q^{-1} = \frac{\bar{q}}{|q|^2} \quad (1.25)$$

per cui:

$$q^{-1} q = q q^{-1} = 1$$

e permette di definire una divisione tramite il prodotto con il reciproco. Occorre però distinguere tra una divisione “sinistra”, $q^{-1} q'$, ed una divisione “destra”, $q' q^{-1}$, in generale distinte tra loro per la non commutatività della moltiplicazione (non ha quindi senso la divisione ordinaria q'/q , in cui non risulta specificato se si tratta di una divisione destra o di una divisione sinistra).

I quaternioni ammettono non solo una rappresentazione tramite lo spazio \mathbb{R}^4 , ma anche tramite matrici. In effetti moltiplicando le matrici:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.26)$$

per $-\sqrt{-1}$, otteniamo tre matrici che soddisfano le relazioni fondamentali dei quaternioni. Abbiamo usato la notazione $\sqrt{-1}$ per distinguere l'unità immaginaria dei numeri complessi dal quaternione proprio i . In questa rappresentazione un quaternione è rappresentato dalla matrice complessa:

$$q = \begin{pmatrix} q_0 - \sqrt{-1} q_3 & -\sqrt{-1} q_1 - q_2 \\ -\sqrt{-1} q_1 + q_2 & q_0 + \sqrt{-1} q_3 \end{pmatrix} \quad (1.27)$$

con l'algebra derivante dall'usuale somma e moltiplicazione tra matrici (con la conseguente non validità della proprietà commutativa). Le matrici (1.26) sono note come matrici di **Pauli**⁶.

Ricordando che la moltiplicazione per un numero complesso a modulo unitario corrisponde ad una rotazione nel piano, per cui l'unità immaginaria esprime una rotazione di 90° , i quaternioni i, j, k possono essere visti come rotazioni piane di 90° in uno spazio quadridimensionale. Si può dare anche una rappresentazione dei quaternioni tramite matrici reali a quattro dimensioni:

$$\begin{aligned} q &= q_0 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + q_1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &\quad + q_2 \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} + q_3 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{1.28} \\ &= \begin{pmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \end{aligned}$$

L'uso dei quaternioni suscitò notevoli controversie. Alcuni dei sostenitori di Hamilton si opposero veementemente allo sviluppo di settori emergenti quali l'algebra lineare e il calcolo vettoriale (sviluppato principalmente dai fisici), affermando che i quaternioni offrivano una notazione migliore. A Dublino costituisce ancora una materia d'esame ed in alcune università americane costituisce l'unico argomento di matematica avanzata.

⁶ **Wolfgang Ernst Pauli** (Vienna, 25 aprile 1900 – Zurigo, 15 dicembre 1958) è stato un fisico austriaco. Fu fra i padri fondatori della meccanica quantistica. Suo è il principio di esclusione nel 1924 che gli valse il Premio Nobel nel 1945, secondo il quale due elettroni in un atomo non possono avere tutti i numeri quantici uguali. Nato da Wolfgang Joseph Pauli e Berta Camilla Schütz, Pauli, sin dal nome era destinato alla strada della fisica: il suo secondo nome, infatti, gli è stato dato dal padre in onore di Ernst Mach. Nel 1926, dopo la pubblicazione da parte di Heisenberg del formalismo matriciale della teoria quantistica, Pauli lo utilizzò per derivare lo spettro dell'atomo di idrogeno. Nel 1927 introdusse le matrici di Pauli come basi per le operazioni sugli spin; questo lavoro influenzò Dirac per formulare l'omonima equazione sugli elettroni relativistici. Nel 1931 propose l'esistenza di una particella senza carica e senza massa per spiegare lo spettro del decadimento beta: nel 1934 Fermi incorporò nella sua teoria sul decadimento tale particella che chiamò neutrino. La prima evidenza sperimentale della sua esistenza avvenne nel 1959. Nel 1940 dimostra il teorema sulla statistica degli spin, che stabilisce che una particella a spin semi-intero è un fermione, mentre una a spin intero un bosone.



Dalla fine del XIX secolo l'analisi vettoriale ha rimpiazzato i quaternioni permettendo di esprimere gli stessi concetti in maniera più semplice ed intuitiva. Oggi sappiamo che i quaternioni sono una struttura molto particolare, che non offre molte generalizzazioni in altre dimensioni (se si escludono gli **ottanioni** in dimensione otto) e, pur mantenendo un posto al caldo nel cuore dei matematici, non sono utilizzati in teorie generali. Attualmente i quaternioni vengono utilizzati principalmente nella rappresentazione di rotazioni e direzioni nello spazio tridimensionale. Hanno applicazione nella grafica 3D computerizzata, nella teoria del controllo, nell'elaborazione dei segnali, nel controllo dell'assetto in fisica e in astrodinamica. Ad esempio, è comune per i veicoli spaziali un sistema di controllo dell'assetto comandato mediante quaternioni. La ragione è che la combinazione di molte trasformazioni descritte da quaternioni è più rapida e stabile numericamente della combinazione di molte trasformazioni matriciali.

1.2 Funzioni di variabile complessa

Definire una **funzione** su un insieme $D \subseteq \mathbb{C}$ equivale a dare una legge che ad ogni punto $z \in D$ (l'insieme D costituisce il **dominio** di definizione della funzione) associa un **unico** numero complesso w :

$$f : D \longrightarrow \mathbb{C}, \quad w = f(z). \quad (1.29)$$

Secondo tale definizione una funzione risulta **univoca**, o ad un sol valore, detta anche **monodroma** (vedremo in seguito la nozione di funzione **multivoca** o a molti valori, detta anche **polidroma**). Se si separano in una funzione f e nel suo argomento z la parte reale ed immaginaria possiamo scrivere:

$$f(z) = u(x, y) + i v(x, y), \quad z = x + i y, \quad (1.30)$$

con u, v funzioni reali degli argomenti reali x e y . Risulta evidente che la teoria delle funzioni di variabile complessa rientra nella teoria più generale delle trasformazioni da \mathbb{R}^2 a \mathbb{R}^2 (alla coppia (x, y) associamo la coppia (u, v)). Noi saremo interessati in particolare ad una classe più ristretta, quella delle funzioni analitiche, che soddisfano opportune richieste di regolarità o, più specificatamente, di differenziabilità. Spiegheremo tra breve cosa ciò significhi ma possiamo già dire che, sebbene le condizioni richieste impongano limitazioni, si ottiene una teoria elegante e estremamente potente in pratica.

Risulta semplice introdurre i concetti di limite e continuità per funzioni complesse, deducendoli dai corrispondenti concetti per le trasformazioni da \mathbb{R}^2 a \mathbb{R}^2 .

Def. 1.1 (Limite) Se f è definita in un intorno di un punto $a \in \mathbb{C}$, escluso al più il punto a , diremo che $A \in \mathbb{C}$ è il **limite** di f per z che tende ad a e scriviamo:

$$\lim_{z \rightarrow a} f(z) = A \quad (1.31)$$

se per ogni intorno U_A del punto A esiste un intorno V_a di a tale che per ogni $z \in U_A$, $z \neq a$, si ha $f(z) \in U_A$. Equivalentemente, per ogni $\epsilon > 0$ esiste un $\delta > 0$ tale che:

$$0 < |z - a| < \delta \implies |f(z) - A| < \epsilon. \quad (1.32)$$

Notiamo che, in generale, il numero reale δ dipenderà da ϵ e dal punto a : $\delta = \delta(a, \epsilon)$.

Lasciamo al lettore l'estensione del concetto di limite ai casi $a = \infty$ o $A = \infty$ (in questo caso si può parlare di divergenza della funzione). Risulta facile verificare che f tende a A se e solo se convergono separatamente le parti reale ed immaginaria di f alle corrispondenti parti reale ed immaginaria di A . Grazie alla evidente somiglianza della definizione di limite con quella del caso reale, i teoremi elementari sui limiti si estendono automaticamente al caso complesso ed eviteremo di enunciarli e dimostrarli. Se si riesaminano tutti i risultati studiati per funzioni reali ci si rende conto che erano

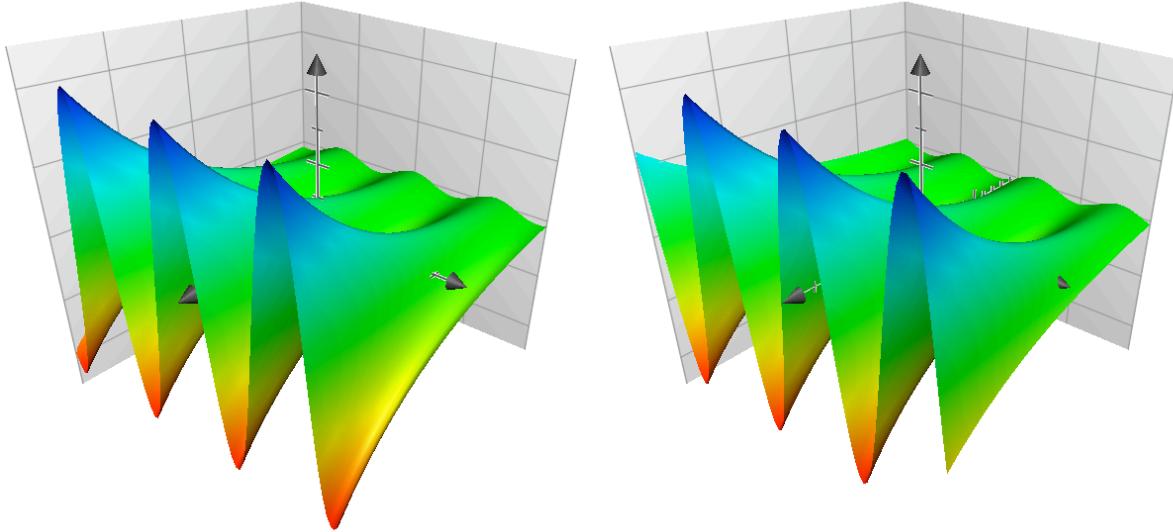


Figura 1.3: Parte reale e parte immaginaria della funzione esponenziale complessa e^z .

basati sull'utilizzo del “valore assoluto” e delle sue proprietà, che può essere sostituito esattamente dal “modulo” nel caso complesso.

Def. 1.2 Sia f una funzione definita in un intorno di un punto $a \in \mathbb{C}$. Diremo che f è **continua** in a se:

$$\lim_{z \rightarrow a} f(z) = f(a) . \quad (1.33)$$

Per le stesse ragioni precedenti, i teoremi elementari sulle funzioni continue noti nel caso reale si estendono pari pari al caso complesso.

1.2.1 L'esponenziale

Generalmente le funzioni di variabile complessa elementari sono estensioni delle corrispondenti funzioni di variabile reale e definite in modo tale che valgano le stesse proprietà note nel caso reale. Per definire tali funzioni spesso si usa lo sviluppo in serie di potenze noto nel caso reale, sostituendo formalmente la variabile reale con una variabile complessa. Occorre quindi provare che lo sviluppo formale ottenuto è convergente in un opportuno dominio (generalmente un disco nel piano complesso), e con le funzioni elementari spesso si ha convergenza in tutto il piano complesso. Un esempio è fornito dalla funzione esponenziale. Nel caso reale sappiamo che:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} , \quad x \in \mathbb{R} ,$$

e la sua estensione al caso complesso è definita da:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad z \in \mathbb{C}. \quad (1.34)$$

Affermare la convergenza di tale serie significa che la successione delle **ridotte**:

$$S_N(z) = \sum_{n=0}^N \frac{z^n}{n!}$$

definisce una successione di punti nel piano complesso che risulta convegente ad un punto definito come e^z . In effetti la successione S_N soddisfa il criterio di convergenza di Cauchy:

$$|S_N - S_M| = \left| \sum_{n=M+1}^N \frac{z^n}{n!} \right| \leq \sum_{n=M+1}^N \frac{|z|^n}{n!} \xrightarrow[N, M \rightarrow \infty]{} 0$$

(abbiamo assunto $N > M$) in quanto la serie (in \mathbb{R}):

$$e^{|z|} = \sum_{n=0}^{\infty} \frac{|z|^n}{n!},$$

è convergente per ogni $|z|$. Poichè il criterio di Cauchy è necessario e sufficiente in \mathbb{R}^2 (e quindi in \mathbb{C}), la serie esponenziale (1.34) è convergente per ogni z .

Si può scomporre l'esponenziale nella sua parte reale ed immaginaria ($z = x + iy$):

$$e^z = e^{x+iy} = e^x e^{iy} = e^x \cos y + i e^x \sin y,$$

da cui:

$$|e^z| = |e^x \cos y + i e^x \sin y| = e^x = e^{\operatorname{Re} z}.$$

1.2.2 Funzioni trigonometriche

Analogamente si definiscono il seno ed il coseno di un numero complesso:

$$\begin{aligned} \sin z &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}, \\ \cos z &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \end{aligned} \quad (1.35)$$

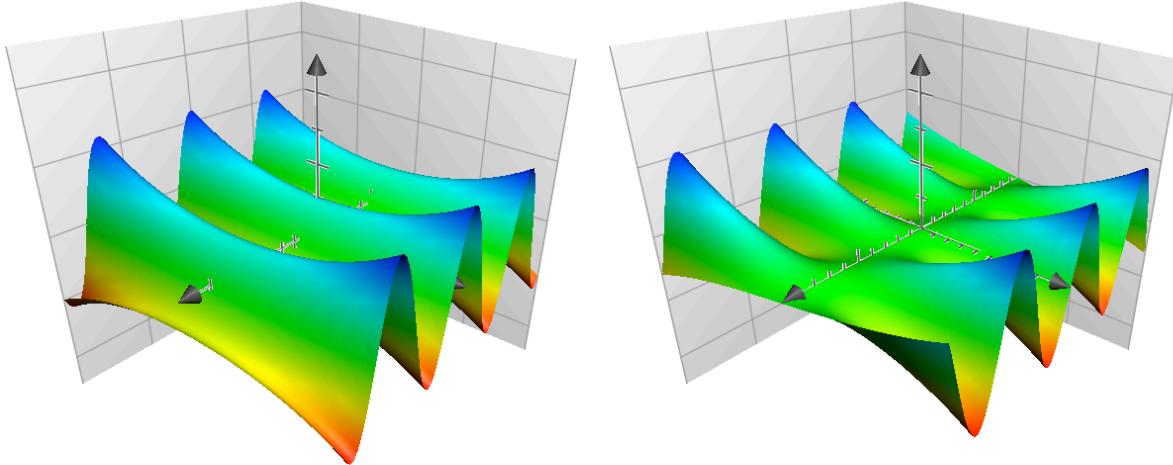


Figura 1.4: Parte reale e parte immaginaria della funzione trigonometrica complessa $\cos z$.

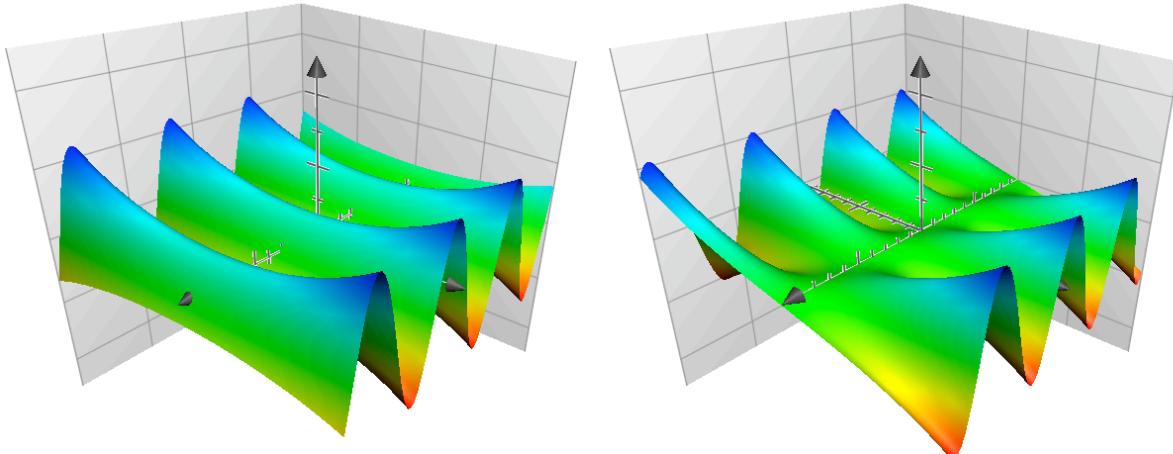


Figura 1.5: Parte reale e parte immaginaria della funzione trigonometrica complessa $\sin z$.

che risultano serie convergenti. Dalla manipolazione algebrica di tali serie è facile dedurre una generalizzazione delle formule di Eulero:

$$e^{iz} = \cos z + i \sin z, \quad (1.36)$$

$$\begin{aligned} \cos z &= \frac{e^{iz} + e^{-iz}}{2}, \\ \sin z &= \frac{e^{iz} - e^{-iz}}{2i}, \end{aligned} \quad (1.37)$$

ed è immediato verificare la relazione fondamentale della trigonometria:

$$\sin^2 z + \cos^2 z = 1, \quad (1.38)$$

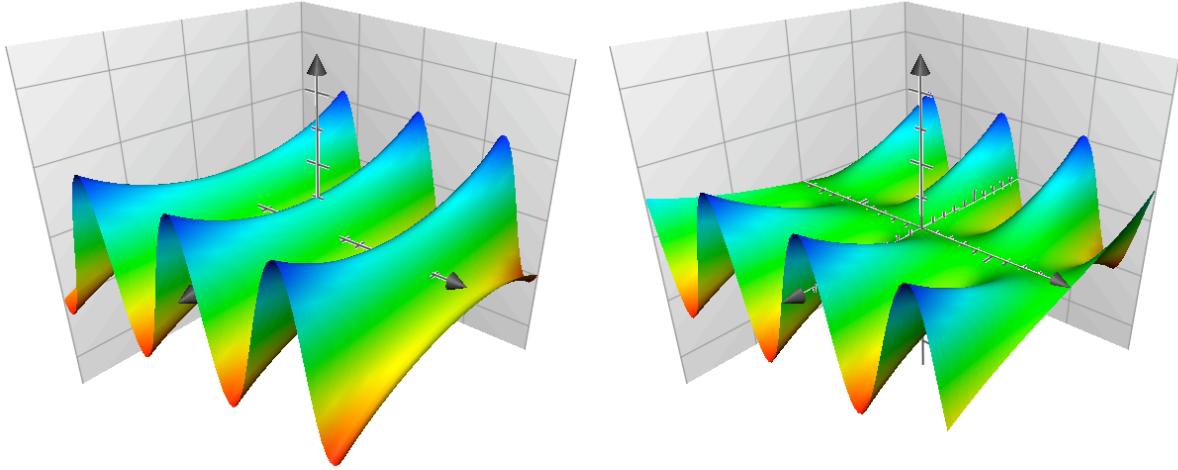


Figura 1.6: Parte reale e parte immaginaria della funzione iperbolica complessa $\cosh z$.

con tutte le conseguenze algebriche che essa comporta. Analogamente si possono estendere al caso complesso tutte le altre funzioni trigonometriche, tangente, cotangente, ecc.

Mediante gli sviluppi in serie, in maniera non sempre agevole, è possibile dimostrare tutte le proprietà algebriche fondamentali dell'esponenziale e delle funzioni trigonometriche.

1.2.3 Funzioni iperboliche

Anche le funzioni iperboliche possono essere estese al piano complesso mediante una estensione del loro sviluppo in serie oppure, più semplicemente, tramite le relazioni:

$$\boxed{\begin{aligned} \cosh z &= \frac{e^z + e^{-z}}{2}, \\ \sinh z &= \frac{e^z - e^{-z}}{2}, \end{aligned}} \quad (1.39)$$

da cui è facile verificare la relazione fondamentale:

$$\boxed{\cosh^2 z - \sinh^2 z = 1,} \quad (1.40)$$

e le relazioni con le funzioni trigonometriche:

$$\left\{ \begin{array}{l} \sinh(i z) = i \sin z, \\ \cosh(i z) = \cos z, \end{array} \right. \quad \left\{ \begin{array}{l} \sin(i z) = i \sinh z, \\ \cos(i z) = \cosh z. \end{array} \right. \quad (1.41)$$

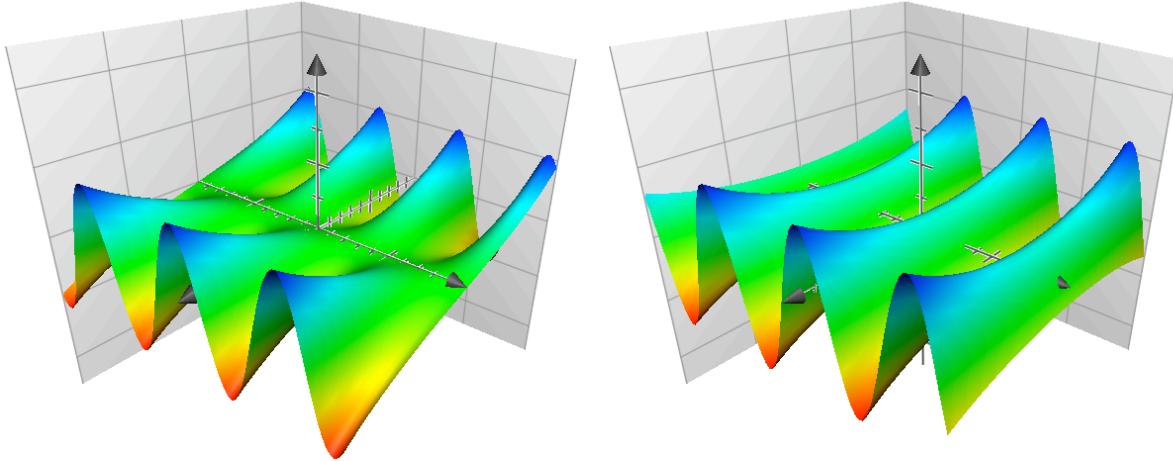


Figura 1.7: Parte reale e parte immaginaria della funzione iperbolica complessa $\sinh z$.

Ponendo $z = x + iy$, abbiamo anche, ad esempio:

$$\sin z = \sin x \cosh y + i \cos x \sinh y,$$

$$\cos z = \cos x \cosh y - i \sin x \sinh y,$$

$$\sinh z = \sinh x \cos y + i \cosh x \sin y,$$

$$\cosh z = \cosh x \cos y + i \sinh x \sin y.$$

Osservazione. Dalle espressioni sopra vediamo che le funzioni seno e coseno mantengono la caratteristica di periodicità lungo la direzione dell'asse reale ($(f(z) = f(z + 2\pi n), n$ intero), ma divergono esponenzialmente nella direzione dell'asse immaginario ($y \rightarrow \infty$). Viceversa, le corrispondenti funzioni iperboliche sono caratterizzate da una periodicità lungo la direzione immaginaria ($f(z) = f(z + i2\pi n), n$ intero) e divergono esponenzialmente lungo la direzione reale ($x \rightarrow \infty$).

Per le funzioni inverse dell'esponenziale o delle funzioni trigonometriche e iperboliche occorrono particolari cautele. Si tratta in realtà di invertire delle trasformazioni non banali da \mathbb{R}^2 a \mathbb{R}^2 , ed in generale non è possibile farlo in maniera univoca.

1.2.4 Differenziabilità ed olomorfismo

Poniamoci ora il problema della differenziabilità di una funzione complessa di variabile complessa. Poichè questa può essere vista come una applicazione da \mathbb{R}^2 a \mathbb{R}^2 ricordiamo dapprima il concetto di differenziabilità per funzioni di due variabili reali, a valori in un insieme arbitrario.

Def. 1.3 Sia Ω un aperto di \mathbb{R}^2 e sia $f(x, y)$ una funzione definita su Ω .

Si dice che f è **differenziabile** in un punto $(x_0, y_0) \in \Omega$ se esiste una forma lineare

del tipo $\alpha h + \beta k$, con α, β nello stesso insieme dei valori assunti dalla funzione f , detta **differenziale**, nelle variabili h e k , tale che:

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \alpha h + \beta k + \omega(h, k)\sqrt{h^2 + k^2}, \quad (1.42)$$

con:

$$\omega(h, k) \xrightarrow[h, k \rightarrow 0]{} 0.$$

Cioè la differenza tra l'incremento $f(x_0 + h, y_0 + k) - f(x_0, y_0)$ e l'espressione lineare $\alpha h + \beta k$ è un infinitesimo di ordine superiore a $\sqrt{h^2 + k^2}$ per $(h, k) \rightarrow (0, 0)$.

Se f è differenziabile nel punto (x_0, y_0) sappiamo che α e β sono determinate in maniera unica e coincidono con le derivate parziali:

$$\begin{cases} \alpha = \frac{\partial f}{\partial x}(x_0, y_0) \\ \beta = \frac{\partial f}{\partial y}(x_0, y_0) \end{cases}$$

ma notiamo che, viceversa, l'esistenza delle derivate parziali di f nel punto (x_0, y_0) non è in generale sufficiente affinché la funzione sia differenziabile in tale punto, ma occorre qualcosa in più, ad esempio che le derivate parziali esistano in tutto un intorno del punto e che siano continue nel punto (x_0, y_0) . In effetti l'esistenza pura e semplice delle derivate parziali fornisce indicazioni di differenziabilità solo lungo le direzioni degli assi coordinati ($k = 0$, oppure $h = 0$), lungo una sola dimensione alla volta, mentre la condizione (1.42) è globale a due dimensioni.

Veniamo ora al caso particolare delle funzioni complesse di una variabile complessa $z = x + iy$ (x, y reali). La richiesta di differenziabilità (1.42) è perfettamente lecita (α e β risultano quantità complesse), ma saremo interessati alle funzioni differenziabili con una richiesta ulteriore sulla forma dell'espressione lineare $\alpha h + \beta k$, cioè che in essa h e k compaiano solo nella combinazione $w = h + ik$.

Def. 1.4 Una funzione complessa f definita su un aperto Ω di \mathbb{C} è detta **olomorfa** in un punto $z_0 \in \Omega$ se esiste $\gamma \in \mathbb{C}$ tale che:

$$f(z_0 + w) - f(z_0) = \gamma w + \omega(w)|w|, \quad (1.43)$$

con:

$$\omega(w) \xrightarrow[w \rightarrow 0]{} 0.$$

Chiaramente, ponendo $z_0 = x_0 + iy_0$, $w = h + ik$, la condizione di olomorfismo in un punto implica la differenziabilità di f , intesa come funzione delle variabili reali x e y , ma

con una condizione ulteriore sui coefficienti α e β , ottenuta da un confronto tra le relazioni (1.42) e (1.43):

$$\begin{cases} \alpha = \gamma \\ \beta = i\gamma \end{cases} \quad (1.44)$$

cioè α e β devono essere espressi tramite un'unica quantità γ . Considerando l'espressione (1.43) con $w = h + ik \neq 0$, abbiamo:

$$\frac{f(z_0 + w) - f(z_0)}{w} = \gamma + \frac{|w|}{w} \omega(w),$$

il rapporto $|w|/w$ è un numero complesso a modulo unitario, per cui risulta evidente che la richiesta di olomorfismo è equivalente a richiedere l'esistenza del limite del rapporto incrementale nel punto z_0 , cioè la derivabilità di f in tale punto:

$$\boxed{\gamma = \lim_{w \rightarrow 0} \frac{f(z_0 + w) - f(z_0)}{w} = f'(z_0) = \frac{df}{dz}(z_0).} \quad (1.45)$$

Notiamo che il limite (1.45) è in effetti un limite doppio, nel senso che sia la parte reale di w che quella immaginaria devono tendere a zero. La richiesta è quindi che tale limite esista e sia indipendente dal modo in cui w tende a zero.

La definizione di olomorfismo si trasporta naturalmente da un punto a tutto un sottoinsieme del piano complesso. Diremo che f è **olomorfa** in un sottoinsieme D di \mathbb{C} se è olomorfa in ogni punto $z \in D$.

1.2.5 Condizioni di olomorfismo

Vediamo ora di determinare delle condizioni pratiche che ci permettano di verificare l'esistenza della derivata per un funzione $f(z)$ in un punto $z = x + iy$. La condizione (1.44) ci fornisce la risposta. Dobbiamo avere sostanzialmente la derivabilità di f rispetto a x e y e le derivate devono verificare:

$$\begin{cases} \alpha = \frac{\partial f}{\partial x} = \gamma \\ \beta = \frac{\partial f}{\partial y} = i\gamma \end{cases}$$

cioè:

$$\boxed{\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0.} \quad (1.46)$$

Separando la parte reale e la parte immaginaria di f :

$$f(z) = u(x, y) + i v(x, y), \quad z = x + iy, \quad (1.47)$$

con $u(x, y)$ e $v(x, y)$ funzioni **reali** delle variabili reali x, y , ed operando la medesima separazione nella (1.46), otteniamo le condizioni di olomorfismo di **Cauchy-Riemann**:

$$\boxed{\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \end{cases}} \quad (1.48)$$

Come sottoprodotto delle relazioni precedenti abbiamo anche (valida solo per funzioni olomorfe):

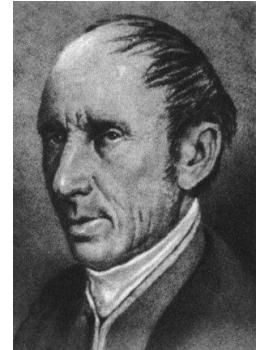
$$\frac{d f}{d z} = \frac{\partial f}{\partial x}. \quad (1.49)$$

Per capire meglio il significato delle relazioni (1.48) di Cauchy⁷-Riemann⁸ forniamo qualche ulteriore manipolazione. Considerando le relazioni:

$$\begin{cases} z = x + i y \\ \bar{z} = x - i y \end{cases} \quad \begin{cases} x = \frac{1}{2}(z + \bar{z}) \\ y = \frac{1}{2i}(z - \bar{z}) \end{cases} \quad (1.50)$$

possiamo considerarle formalmente come un semplice cambio algebrico di variabili nel piano \mathbb{R}^2 (non è un vero cambio di coordinate nel piano a causa dei fattori complessi i)

⁷ **Augustin-Louis Cauchy** (Parigi, 21 agosto 1789 – Sceaux, 23 maggio 1857) è stato un matematico e ingegnere francese. Ha avviato il progetto della formulazione e dimostrazione rigorosa dei teoremi dell'analisi infinitesimale basato sull'utilizzo delle nozioni di limite e di continuità. Ha dato anche importanti contributi alla teoria delle funzioni di variabile complessa e alla teoria delle equazioni differenziali. La sistematicità e il livello di questi suoi lavori lo collocano tra i padri dell'analisi matematica. Notiamo che, sebbene Cauchy abbia “insegnato” il rigore ai matematici, i suoi risultati si basavano spesso sull'intuizione, senza essere rigorosi, incorrendo talvolta in contestazioni e controesempi da parte dei colleghi.



⁸ **Georg Friedrich Bernhard Riemann** (Breselenz, Germania, 17 Settembre 1826 – Selasca, ora Verbania, Italia, 20 Luglio 1866) è stato un matematico tedesco. Egli fornì importanti contributi all'analisi e alla geometria differenziale, alcuni dei quali spianarono la strada per lo sviluppo della relatività generale. Rimane non dimostrata l'Ipotesi di Riemann legata alla struttura dei numeri primi: “tutti gli zeri complessi della funzione Zeta hanno parte reale $1/2$ ”. L'ipotesi di Riemann costituisce l'ottavo problema di Hilbert, tra quelli elencati da Hilbert nel 1900 come oggetti di studio del XX secolo, ed è l'unico rimasto irrisolto alla fine del secolo, per cui è stato inserito tra i “Problemi per il Millennio”.



ed otteniamo formalmente la corrispondente trasformazione per le derivate:

$$\begin{cases} \frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) \\ \frac{\partial f}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) \end{cases} \quad (1.51)$$

Le condizioni di olomorfismo si possono esprimere formalmente in forma compatta come:

$$\boxed{\frac{\partial f}{\partial \bar{z}} = 0.} \quad (1.52)$$

Possiamo quindi capire subito se una funzione è un buon candidato per essere olomorfa: non deve dipendere esplicitamente da \bar{z} , ma solo da z . In altri termini le variabili reali x e y devono comparire esplicitamente solo nella combinazione $z = x + iy$.

Risulta chiaro che le operazioni aritmetiche elementari non “distruggono” la proprietà di olomorfismo. Se due funzioni $f(z)$ e $g(z)$ sono olomorfe (cioè derivabili) in un punto z_0 , anche la loro somma $f(z) + g(z)$, la differenza $f(z) - g(z)$, e il loro prodotto $f(z)g(z)$, sono funzioni olomorfe in z_0 . Se $g(z_0) \neq 0$, anche il rapporto $f(z)/g(z)$ è olomorfo in z_0 (notiamo che se $g(z_0) = 0$, il rapporto non è definito in z_0 , cioè z_0 non appartiene al dominio del rapporto).

La definizione data di funzione olomorfa viene spesso assunta come definizione di funzione **analitica**. In realtà il concetto di analiticità è diverso e riconducibile alla sviluppabilità in serie. Per la precisione diremo che una funzione $f(z)$ a valori reali o complessi, di una variabile reale o complessa z , definita in un aperto Ω , è **analitica** in Ω se, per tutti i punti $z_0 \in \Omega$, $f(z)$ è sviluppabile in serie di potenze di $(z - z_0)$, cioè esiste un reale $\rho(z_0) > 0$ e una serie di potenze con raggio di convergenza $\rho(z_0)$ tale che:

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad \text{per } |z - z_0| < \rho(z_0). \quad (1.53)$$

In realtà dimostreremo che tutte le funzioni olomorfe su un aperto Ω sono anche analitiche (sviluppabili in serie). Viceversa, ogni funzione analitica in Ω risulterà anche olomorfa. Quindi, per le funzioni di una variabile complessa, si ha equivalenza tra i concetti di olomorfismo e analiticità. Il concetto di analiticità si può porre anche per funzioni di variabile reale, mentre quello di olomorfismo è definito solo per funzioni di variabile complessa (a valori complessi). Nel caso di funzioni di più variabili complesse la cosa si complica ed occorre fare molta attenzione, ma poiché ciò esula dai nostri scopi, non tratteremo qui il problema. Vediamo ora qualche semplice esempio di funzione olomorfa.

Esempio 1.1 Polinomi. La più semplice, sebbene banale, funzione olomorfa è la funzione costante:

$$p_0(z) = \text{cost.}$$

la cui derivata esiste ed è ovviamente nulla.

Consideriamo ora la funzione

$$p_1(z) = z.$$

Chiaramente si ha

$$\begin{aligned} \frac{p_1(z + \Delta z) - p_1(z)}{\Delta z} &= \frac{z + \Delta z - z}{\Delta z} = 1, \\ \frac{dp_1(z)}{dz} &= 1. \end{aligned}$$

Consideriamo ora una potenza z^n ; abbiamo

$$\frac{(z + \Delta z)^n - z^n}{\Delta z} = \binom{n}{1} z^{n-1} + \sum_{k=2}^n \binom{n}{k} (\Delta z)^{k-1} z^{n-k} \xrightarrow[\Delta z \rightarrow 0]{} n z^{n-1},$$

con manipolazioni algebriche perfettamente identiche a quelle che dimostrano la derivata della funzione reale di variabile reale x^n . D'altra parte z^n verifica le condizioni di Cauchy-Riemann. Infatti, posto $z = x + iy$:

$$\begin{aligned} z^n &= (x + iy)^n, \\ \frac{\partial(z^n)}{\partial x} &= n(x + iy)^{n-1}, \\ \frac{\partial(z^n)}{\partial y} &= i n(x + iy)^{n-1}, \\ \frac{\partial(z^n)}{\partial x} + i \frac{\partial(z^n)}{\partial y} &= n(x + iy)^{n-1}(1 + i^2) = 0. \end{aligned}$$

Possiamo così dire che ogni polinomio in z , essendo combinazione lineare di monomi, esprime una funzione olomorfa in tutto il piano complesso.

— ○ —

Esempio 1.2 Potenze negative. Consideriamo la funzione:

$$f(z) = \frac{1}{z} \quad z \neq 0,$$

se $z = x + iy$ allora:

$$\begin{aligned} f(z) &= \frac{1}{x + iy} = \frac{x - iy}{x^2 + y^2}, \\ \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} &= -\frac{1}{(x + iy)^2}(1 + i^2) = 0, \end{aligned}$$

per cui f risulta olomorfa per $z \neq 0$, e:

$$\frac{df}{dz} = \frac{\partial f}{\partial x} = -\frac{1}{(x + iy)^2} = -\frac{1}{z^2}, \quad z \neq 0.$$

— ○ —

1.2.6 Curve e domini

Dopo il concetto di differenziabilità, o meglio, quello di derivabilità per una funzione vogliamo vedere quello di integrabilità. Poichè tratteremo principalmente l'integrazione lungo una curva nel piano complesso, vediamo prima il concetto di curva. Sostanzialmente una curva è definita (ed usata) tramite una parametrizzazione, ma dobbiamo fare una piccola precisazione in quanto, quando ci riferiamo ad una curva, siamo interessati principalmente all'aspetto geometrico, mentre una parametrizzazione contiene informazioni ulteriori, non strettamente necessarie. Fisicamente una parametrizzazione descrive sia la traiettoria (la **curva** in senso stretto), sia il modo in cui tale traiettoria viene percorsa (l'equazione oraria o **cammino**).

Def. 1.5 Si chiama **cammino** γ una applicazione continua di un intervallo $[\alpha, \beta]$ dell'asse reale in \mathbb{C} . I punti $a = \gamma(\alpha)$ e $b = \gamma(\beta)$ sono detti **estremi**. Il cammino è detto **chiuso** se $\gamma(\alpha) = \gamma(\beta)$.

Un cammino è definito da una particolare parametrizzazione continua tramite una variabile reale che varia su un intervallo (che assumiamo finito). La curva rappresenta il luogo dei punti descritto dalla equazione $z = \gamma(t)$ e per ridurre la sua dipendenza dalla scelta della parametrizzazione occorre stabilire quando due cammini descrivono lo stesso luogo di punti. Diremo che due cammini:

$$\gamma_1 : [\alpha_1, \beta_1] \longrightarrow \mathbb{C}, \quad \gamma_2 : [\alpha_2, \beta_2] \longrightarrow \mathbb{C},$$

sono **equivalenti** (e scriveremo $\gamma_1 \sim \gamma_2$) se esiste una funzione continua, suriettiva e strettamente crescente:

$$\tau : [\alpha_1, \beta_1] \xrightarrow{\text{su}} [\alpha_2, \beta_2],$$

tale che $\gamma_1(t) = \gamma_2(\tau(t))$ per ogni $t \in [\alpha_1, \beta_1]$. In altri termini una parametrizzazione si ottiene dall'altra mediante un opportuno cambio di variabile. Si verifica facilmente che tale relazione tra cammini è una relazione di equivalenza matematica e porta al concetto astratto di curva come classe di equivalenza. In pratica una curva è individuata solamente dal suo percorso geometrico (la traiettoria) e non dal modo in cui viene percorsa.

Def. 1.6 Si chiama **curva** una classe di cammini equivalenti secondo la relazione di equivalenza precedente \sim .

Una curva è quindi un insieme di punti ordinati con continuità dai valori crescenti di un parametro. Non ha importanza in che modo viene percorsa, ma ha comunque importanza il verso di percorrenza. Abbiamo infatti richiesto che il cambio di variabile sia espresso tramite una funzione crescente, che mantiene l'ordine dei punti lungo la curva, e risulta quindi implicito un orientamento della curva. La continuità richiesta nella parametrizzazione esprime il concetto che la curva può essere tracciata senza staccare la penna dal foglio.

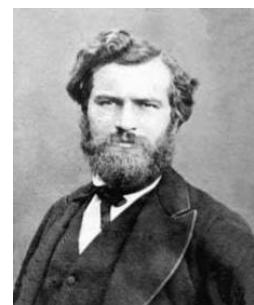
Un cammino è detto **semplice** se la parametrizzazione γ è continua e biiettiva all'interno dell'intervallo di variabilità del parametro ($\alpha < t < \beta$, con le convenzioni precedenti). In pratica la semplicità equivale a richiedere che la curva non si autointersechi, cioè non presenti nodi o altri punti che si ottengono per due o più valori distinti del parametro (a parte gli estremi nel caso di curve chiuse). Un cammino semplice e chiuso viene detto di **Jordan** (la biettività è persa solo sugli estremi). Un cammino è detto **differenziabile**, oppure **regolare**, se la parametrizzazione γ ammette una derivata $\gamma'(t)$ continua e non nulla in ogni punto. Vogliamo cioè che in ogni punto della curva sia ben definita una direzione tangente alla curva. Fisicamente vogliamo escludere il caso in cui percorriamo la curva “fermandoci” in un suo punto (senza velocità non conosciamo la direzione del moto), oppure che vi sia più di una direzione tangente. Un cammino è detto **rettificabile** se $\gamma(t) = x(t) + iy(t)$ ammette quasi ovunque una derivata $\gamma'(t)$ con valore assoluto integrabile, cioè esiste:

$$L_\gamma = \int_\alpha^\beta |\gamma'(t)| dt = \int_\alpha^\beta \sqrt{[x'(t)]^2 + [y'(t)]^2} dt, \quad (1.54)$$

che definisce la **lunghezza** del cammino. Risulta chiaro che se un cammino è differenziabile a tratti è anche rettificabile. La lunghezza non dipende dalla parametrizzazione scelta (lasciamo al lettore la dimostrazione) ed è quindi una proprietà della curva, cioè comune a tutti i cammini equivalenti tra loro. Le medesime condizioni e definizioni si estendono alle curve, definendo in maniera ovvia i concetti di **curva di Jordan**⁹, **regolare**, **rettificabile**, con una precisazione ulteriore per il concetto di regolarità. Per non violare la richiesta di differenziabilità nel cambio di parametrizzazione dobbiamo richiedere che il cambio di variabile τ risulti una funzione differenziabile con continuità e con derivata strettamente positiva.

Saranno utili in seguito anche alcune caratterizzazioni che potremo dare per un sottoinsieme dei numeri complessi. In particolare considereremo regioni del piano complesso che sono aperte e **connesse per archi**, cioè due punti qualunque $a, b \in D$ possono essere collegati da un cammino di estremità a e b tutto contenuto in D . La connessione per archi risulta un caso particolare del concetto di insieme connesso, cioè costituito essenzialmente da una sola parte.

⁹ **Marie Camille Ennemond Jordan** (Lione, 5 gennaio 1838 – Milano, 22 gennaio 1922) è stato uno fra i più importanti matematici francesi. Jordan lavorò in un'ampia varietà di settori, contribuendo praticamente ad ogni argomento matematico studiato nel suo periodo: nella sua opera troviamo lavori sui gruppi finiti, sull'algebra lineare e sull'algebra multilineare, sulla teoria dei numeri, sulla topologia dei poliedri, sulle equazioni differenziali, sulla meccanica.



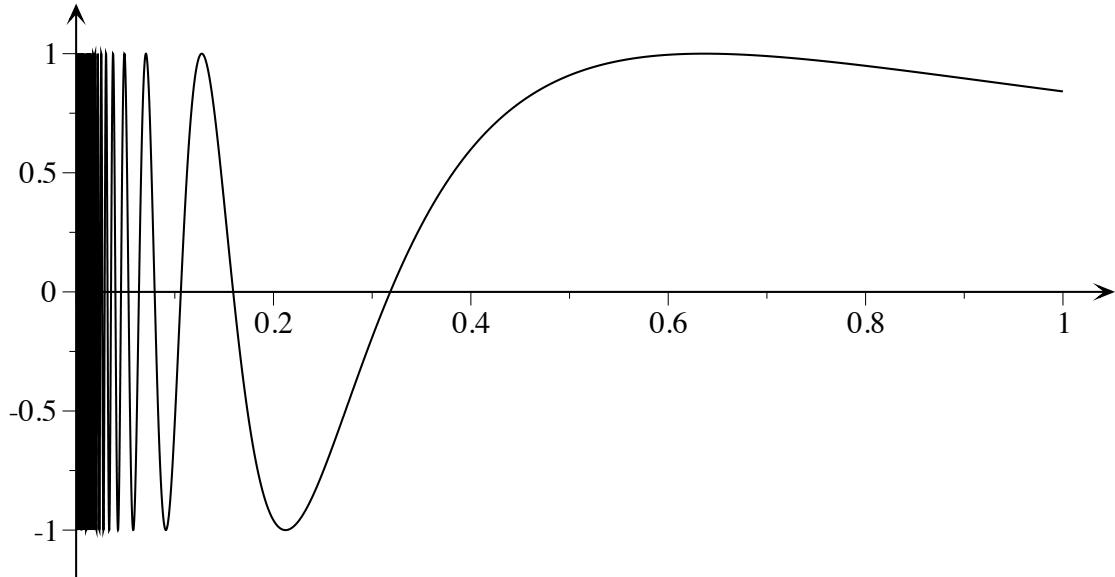


Figura 1.8: Insieme connesso ma non connesso per archi.

Def. 1.7 Un insieme M è detto **connesso** se non può essere ripartito in due parti non vuote N_1 e N_2 , cioè non esistono due parti N_1 , N_2 , non vuote, tali che $M = N_1 \cup N_2$ con $\overline{N_1} \cap N_2 = \emptyset$ e $N_1 \cap \overline{N_2} = \emptyset$.

Osservazione. Notiamo che si richiede che nessuna delle due parti abbia punti di aderenza per l'altra, ma non resta esclusa la possibilità che le chiusure si intersechino, cioè $\overline{N_1} \cap \overline{N_2} \neq \emptyset$. Notiamo inoltre che un insieme connesso per archi è connesso, ma il reciproco è generalmente falso. Un esempio è dato dal sottoinsieme di \mathbb{R}^2 (Fig. 1.8):

$$M = \{(0, y) ; |y| \leq 1\} \cup \left\{ \left(x, \sin \frac{1}{x} \right) ; x > 0 \right\},$$

connesso, ma non connesso per archi. Le due nozioni sono però equivalenti nel caso di insiemi aperti.

Diremo inoltre che un sottoinsieme $D \subseteq \mathbb{C}$ è **semplicemente connesso** se il suo bordo ∂D è un insieme connesso (in caso contrario lo diremo **molteplicemente connesso**). Si veda la figura 1.9 per un esempio di dominio semplicemente connesso ed un esempio di dominio molteplicemente connesso.

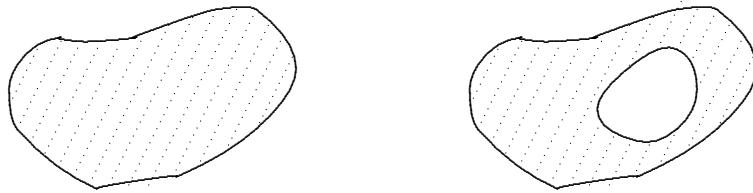


Figura 1.9: Regioni semplicemente e molteplicemente connesse.

1.2.7 Integrali di contorno

Veniamo ora alla integrazione vera e propria. Sia $f(z)$ una funzione della variabile complessa $z = x + iy$ e consideriamo una curva γ nel piano complesso, che supponiamo regolare, nel senso che può essere espressa mediante una equazione parametrica:

$$z = z(t) = x(t) + iy(t) \quad t_1 \leq t \leq t_2 \quad (1.55)$$

dove t è un parametro reale e $x(t)$, $y(t)$ sono funzioni univoche, reali, continue e derivabili, con derivate prime continue (possiamo considerare più in generale, curve regolari a tratti).

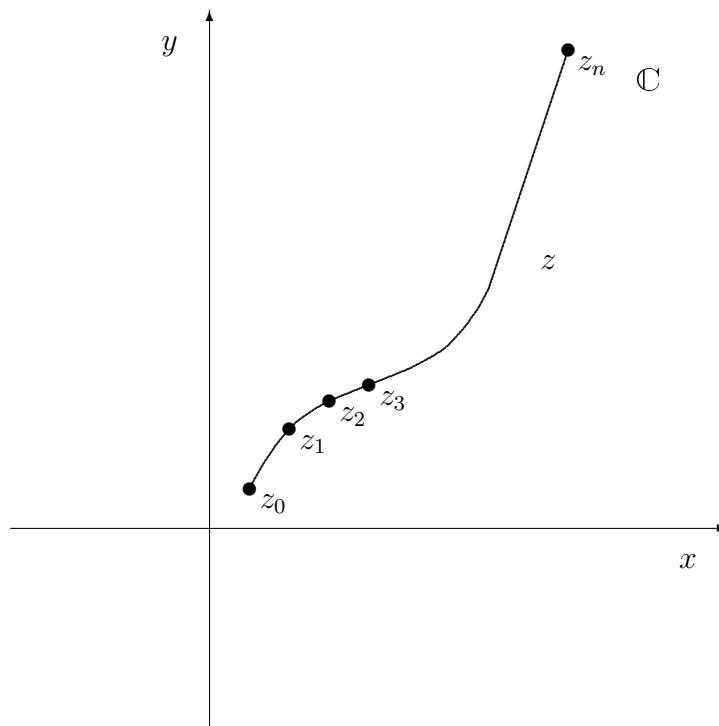


Figura 1.10: Integrazione.

Come mostrato in figura 1.10, dapprima suddividiamo l'arco $a b$ (se a, b sono i punti terminali della curva) in n intervalli, introducendo $n + 1$ punti sulla curva:

$$z_0 = a, z_1, \dots z_{n-1}, z_n = b.$$

Quindi su ogni arco $z_{k-1} z_k$ scegliamo un ulteriore punto ξ_k e consideriamo la somma

$$I_n = \sum_{k=1}^n f(\xi_k)(z_k - z_{k-1}),$$

e valutiamo il limite per $n \rightarrow \infty$ con la condizione:

$$|z_k - z_{k-1}| \rightarrow 0 \quad \text{per ogni } k.$$

Questo limite, nel caso esista, e che sia indipendente dal modo in cui sono scelti i punti z_k e ξ_k , è detto **integrale di contorno** di $f(z)$ lungo la curva γ e indicato con:

$$I = \int_{\gamma} f(z) dz. \quad (1.56)$$

A volte si scrive anche $\int_a^b f(z) dz$ ma la notazione è imprecisa poichè il valore dell'integrale dipende in generale dal percorso, cioè dalla curva scelta per congiungere i punti a e b . Nel caso che la curva γ sia chiusa, i punti terminali coincidono ($a \equiv b$), si scrive:

$$I = \oint_{\gamma} f(z) dz. \quad (1.57)$$

Separando $f(z)$ e z nelle loro parti reale ed immaginaria, l'integrale curvilineo può essere espresso anche come ($f = u + i v$, $z = x + i y$):

$$I = \int_{\gamma} (u dx - v dy) + i \int_{\gamma} (v dx + u dy), \quad (1.58)$$

e questo a sua volta può considerarsi come un integrale ordinario rispetto ad una variabile reale t con cui è parametrizzata la curva

$$\begin{cases} dx = \frac{dx}{dt} dt \\ dy = \frac{dy}{dt} dt \end{cases} \quad dz = \frac{dz}{dt} dt,$$

$$I = \int_{t_a}^{t_b} \left(u \frac{dx}{dt} - v \frac{dy}{dt} \right) dt + i \int_{t_a}^{t_b} \left(v \frac{dx}{dt} + u \frac{dy}{dt} \right) dt = \int_{t_a}^{t_b} f(z(t)) \frac{dz(t)}{dt} dt.$$

Osservazione. In ultima analisi possiamo dire che l'integrale curvilineo è definito tramite integrazioni su un intervallo reale. Le formule date che definiscono l'integrazione lungo un

cammino sono ben poste, nel senso che non dipendono dalla parametrizzazione della curva. La definizione iniziale serve solo per mettere in evidenza il fatto che l'integrale dipende dalla curva, intesa come luogo di punti, e non dal modo in cui viene percorsa. Scegliendo un altro cammino equivalente per la medesima curva si ottiene lo stesso risultato in base al teorema del cambio di variabile nell'integrazione noto dai corsi di analisi. Si lascia al lettore interessato la verifica come esercizio.

Notiamo inoltre che una curva è un insieme ordinato di punti, cioè implicito nella sua costruzione è inteso anche un orientamento. Questo è particolarmente importante nel caso di curve chiuse in cui non esplicitiamo un punto iniziale e un punto finale particolare.

Le proprietà fondamentali degli integrali di contorno sono analoghe a quelle degli integrali ordinari, cioè:

$$\int_{\gamma} af(z)dz = a \int_{\gamma} f(z)dz \quad \text{con } a \text{ costante,} \quad (1.59)$$

$$\int_{\gamma} [f(z) + g(z)] dz = \int_{\gamma} f(z)dz + \int_{\gamma} g(z)dz, \quad (1.60)$$

$$\int_{\gamma_1 \cup \gamma_2} f(z)dz = \int_{\gamma_1} f(z)dz + \int_{\gamma_2} f(z)dz, \quad (1.61)$$

con le due curve γ_1 e γ_2 senza parti in comune (escluso eventualmente gli estremi o, in ogni caso, intersezioni a misura nulla, cioè punti singoli). Se invece una curva γ è percorsa in senso inverso (diremo che percorriamo la curva $-\gamma$ o γ^-), avremo:

$$\int_{-\gamma} f(z)dz = - \int_{\gamma} f(z)dz. \quad (1.62)$$

Notiamo che se $z = \gamma(t)$, con $\alpha \leq t \leq \beta$, è una parametrizzazione della curva γ , una possibile parametrizzazione della curva $-\gamma$ è data da $z = \gamma(\beta + \tau(\alpha - \beta))$ con $0 \leq \tau \leq 1$.

Possiamo considerare anche un cambio di variabile (complessa) tramite una trasformazione:

$$z = z(\zeta), \quad (1.63)$$

e scrivere:

$$\int_{\gamma} f(z)dz = \int_{\gamma'} f(z(\zeta)) \frac{dz(\zeta)}{d\zeta} d\zeta, \quad (1.64)$$

dove le curve γ e γ' sono una immagine dell'altra tramite la trasformazione (1.63): $\gamma = z(\gamma')$. Ovviamente la trasformazione (1.63) deve essere un valido cambio di coordinate nel piano, cioè invertibile con Jacobiano non nullo, in pratica la funzione $z(\zeta)$ deve essere una funzione olomorfa con derivata non nulla (almeno lungo la curva γ').

Supponiamo ora che in una regione R (aperta) del piano complesso che contiene la curva γ la funzione $f(z)$ possa essere espressa come la derivata di un'altra funzione $F(z)$ (che risulterà quindi olomorfa in R). In tal caso $F(z)$ è detta **primitiva** di $f(z)$:

$$f(z) = \frac{dF(z)}{dz}, \quad (1.65)$$

e, parametrizzando la curva γ , abbiamo:

$$f(z(t)) \frac{dz(t)}{dt} = \frac{dF}{dz}(z(t)) \frac{dz(t)}{dt} = \frac{d}{dt} F(z(t)),$$

Se la funzione $f(z)$ è continua possiamo affermare:

$$\boxed{\int_{\gamma} f(z) dz = \int_{\gamma} \frac{dF(z)}{dz} dz = \int_{t_a}^{t_b} \frac{dF(z(t))}{dt} dt = F(b) - F(a),} \quad (1.66)$$

che esprime il cosiddetto teorema fondamentale del calcolo integrale (in questo caso l'integrale non dipende dal percorso, ma solo dai punti estremi a e b). In particolare, prendendo un percorso chiuso γ :

$$\boxed{\oint_{\gamma} f(z) dz = \oint_{\gamma} \frac{dF(z)}{dz} dz = 0} \quad (1.67)$$

essendo gli estremi coincidenti $a = b$. Inoltre, se esistono tutte le derivate e queste sono continue:

$$\begin{aligned} \frac{d}{dz} (f(z)g(z)) &= f(z) \frac{dg(z)}{dz} + \frac{df(z)}{dz} g(z), \\ \boxed{\int_{\gamma} \frac{df(z)}{dz} g(z) dz} &= f(z) g(z) \Big|_{z=a}^{z=b} - \int_{\gamma} \frac{dg(z)}{dz} f(z) dz. \end{aligned} \quad (1.68)$$

Notiamo che abbiamo richiesto la continuità delle funzioni e delle loro derivate ma, come vedremo, nel caso di funzioni olomorfe la continuità delle derivate è conseguenza dell'olomorfismo stesso e può essere omessa come ipotesi.

È spesso molto utile poter avere delle maggiorazioni di certi integrali di contorno. Consideriamo l'integrale:

$$I = \int_{\gamma} f(z) dz,$$

con γ curva regolare a tratti e $|f(z)|$ limitata su γ . Allora l'integrale stesso risulta il limite, per $n \rightarrow \infty$, della somma:

$$I_n = \sum_{k=1}^n (z_k - z_{k-1}) f(\xi_k),$$

$$|I_n| \leq \sum_{k=1}^n |f(\xi_k)| |z_k - z_{k-1}| \leq \sup_{\gamma} |f(z)| \sum_{k=1}^n |z_k - z_{k-1}| \leq \sup_{\gamma} |f(z)| L_{\gamma},$$

con L_{γ} lunghezza della curva γ . Otteniamo così, per $n \rightarrow \infty$:

$$\boxed{\left| \int_{\gamma} f(z) dz \right| \leq \sup_{\gamma} |f(z)| L_{\gamma}} \quad (1.69)$$

detta **disuguaglianza di Darboux**¹⁰.

¹⁰ **Jean Gaston Darboux** (Nîmes, Francia, 14 agosto 1842 – Parigi, 23 febbraio 1917) fu un matematico francese. Contribuì in modo significativo alla geometria e all’analisi matematica, in particolare nel campo delle equazioni differenziali. Fu il biografo di Henri Poincaré. Ricevette la laurea dalla École Normale Supérieure nel 1866. Nel 1884, fu eletto alla Académie des Sciences e nel 1900 fu nominato presidente permanente dell’Accademia.



1.3 Teorema di Cauchy

Abbiamo visto che quando dobbiamo integrare una funzione $f(z)$ lungo un percorso in una regione in cui $f(z)$ ammette una primitiva il valore dell'integrale dipende solo dai punti estremi. In generale una generica funzione di variabile complessa non possiede una primitiva univoca, e l'integrale di contorno dipende inevitabilmente dal cammino scelto che congiunge i due punti. D'altra parte le funzioni analitiche hanno una proprietà fondamentale: all'interno di regioni opportune, queste possiedono una primitiva univoca (a meno di una costante additiva), ed il risultato della loro integrazione non dipende dal percorso. Tutto ciò è una conseguenza del teorema di Cauchy, che gioca un ruolo centrale nella teoria delle funzioni olomorfe. Si possono dare diverse versioni del teorema, ma tutte affermano in ogni caso l'annularsi dell'integrale di una funzione olomorfa su un percorso chiuso, e si distinguono solo per la forma topologica della regione aperta del piano complesso che contiene la curva (con implicazioni sulle possibili forme della curva stessa). Per i nostri scopi possiamo dare la seguente formulazione:

Teo. 1.1 (Teorema integrale di Cauchy) *Sia $f(z)$ olomorfa in una regione aperta semplicemente连通的 D, allora se γ è una curva chiusa semplice, regolare a tratti, contenuta in D, si ha:*

$$\oint_{\gamma} f(z) dz = 0. \quad (1.70)$$

Notiamo che l'affermazione è molto semplice e generale con ipotesi minime e, come spesso avviene in questi casi, la sua dimostrazione è molto laboriosa. In effetti questa procede dal particolare al generale, coinvolgendo la verifica nei casi particolari di risultati che vedremo in seguito come conseguenza del teorema stesso, per cui la omettiamo, rimandando chi fosse interessato ai testi più importanti sulle funzioni analitiche. Nella sua forma originale il teorema richiedeva una ipotesi supplementare, cioè non solo che $f(z)$ fosse derivabile, ma anche che tale derivata fosse continua (vedi esercizio 1.6). Successivamente **Goursat**¹¹ si è accorto che la derivata di una funzione olomorfa è sempre continua ed ha mostrato che la continuità della derivata non è necessaria. In base a ciò il teorema è noto anche come teorema di **Cauchy-Goursat**. Sostanzialmente Goursat è

¹¹ **Edouard Jean-Baptiste Goursat** (Lanzac, 21 maggio 1858 – Parigi, 25 novembre 1936) è stato un matematico francese, ricordato principalmente per il suo *Cours d'analyse mathématique* (1902-13). Questo corso stabilì uno standard per l'insegnamento dell'analisi matematica, in particolare dell'analisi complessa. Goursat è anche ricordato per aver dato una dimostrazione del Teorema di Cauchy (*Démonstration du théorème de Cauchy*, 1884) che superava i problemi formali della dimostrazione standard fino ad allora conosciuta.



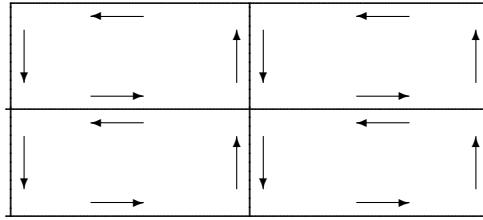


Figura 1.11: Procedimento di Goursat.

riuscito a dimostrare il risultato in un caso molto semplice, quando la curva γ è un semplice rettangolo, con una tecnica molto ingegnosa ed elegante, usando solo la condizione di olomorfismo.

Ripercorriamo velocemente le sue idee. Se la curva γ è un rettangolo R nel piano complesso possiamo dividere ogni lato a metà e costruire quattro nuovi rettangoli (con dei lati in comune) R_1, R_2, R_3, R_4 , orientati concordemente con il rettangolo originario (vedi figura 1.11). L'integrale originale può essere composto come somma degli integrali sui quattro rettangoli (i tratti interni sono percorsi due volte in sensi opposti):

$$\begin{aligned} J(R) &= \oint_R f(z) dz = \sum_{k=1}^4 \oint_{R_k} f(z) dz = \sum_{k=1}^4 J(R_k) \\ |J(R)| &\leq \sum_{k=1}^4 |J(R_k)| \leq 4 \max\{|J(R_k)|\} = 4|J(R_1)| \end{aligned}$$

assumendo il primo rettangolo R_1 quello con contributo maggiore in valore assoluto. Notiamo che se L è il perimetro del rettangolo originale R , ognuno dei quattro sottorettangoli ha perimetro $L/2$. Possiamo procedere iterativamente nella suddivisione ripetendo il ragionamento per R_1 e selezionare il sottorettangolo con contributo maggiore, e così via, generando una sequenza di rettangoli, ognuno contenuto nel precedente, R_1, R_2, \dots, R_n , con perimetro $2^{-n} L$, e:

$$|J(R)| \leq 4^n |J(R_n)|$$

La sequenza di rettangoli converge chiaramente ad un rettangolo degenere di ampiezza nulla, cioè ad un punto z_0 , interno al rettangolo originario R o sul bordo stesso ∂R (i centri dei vari rettangoli formano una successione convergente a z_0) e se $z \in R_n$ abbiamo:

$$|z - z_0| \leq 2^{-n} L, \quad z \in R_n \tag{1.71}$$

Finora abbiamo solo supposto che i vari integrali esistano, per cui è sufficiente che la funzione $f(z)$ sia continua. È a questo punto che entra in gioco l'ipotesi di olomorfismo (e niente altro) nel punto z_0 . Tramite una verifica diretta abbiamo che, per ogni rettangolo:

$$\oint_{R_n} dz = 0, \quad \oint_{R_n} z dz = 0$$

per cui:

$$J(R_n) = \oint_{R_n} f(z) dz = \oint_{R_n} \left(f(z) - f(z_0) - (z - z_0) \frac{df}{dz}(z_0) \right) dz$$

L'olomorfismo di $f(z)$ in z_0 ci dice che l'integrando precedente è un infinitesimo di ordine superiore a $z - z_0$, per cui, per ogni $\epsilon > 0$:

$$\left| f(z) - f(z_0) - (z - z_0) \frac{df}{dz}(z_0) \right| \leq |z - z_0| \epsilon$$

quando $|z - z_0| \leq \delta_\epsilon$ con un opportuno δ_ϵ . Scegliendo n abbastanza grande in modo che $2^{-n}L < \delta_\epsilon$ possiamo applicare la diseguaglianza di Darboux:

$$\begin{aligned} |J(R_n)| &\leq \sup_{z \in R_n} \left| f(z) - f(z_0) - (z - z_0) \frac{df}{dz}(z_0) \right| L_{R_n} \\ &\leq \epsilon \sup_{z \in R_n} |z - z_0| 2^{-n} L \leq \epsilon (2^{-n} L)^2 \\ |J(R)| &\leq \epsilon L^2 \end{aligned} \tag{1.72}$$

cioè, essendo ϵ arbitrario:

$$J(R) = \oint_R f(z) dz = 0 \tag{1.73}$$

Sfruttando questo risultato è stato poi possibile formulare il teorema di Cauchy per qualsiasi curva chiusa contenuta in una regione circolare (o ellittica) per giungere poi al caso generale da noi considerato utilizzando le conseguenze del teorema stesso in tali regioni. Attualmente si preferisce seguire un ragionamento analogo utilizzando triangoli al posto dei rettangoli, permettendo immediatamente la formulazione del teorema di Cauchy per regioni convesse (un insieme è **convesso** se contiene tutto il segmento congiungente due punti qualsiasi dell'insieme).

Una conseguenza immediata del teorema integrale di Cauchy è che l'integrale di contorno tra due punti in un dominio di olomorfismo semplicemente connesso non dipende dal percorso scelto che collega i due punti. Siano infatti C_1 e C_2 due curve arbitrarie che collegano due estremi fissi a e b (vedi figura 1.12). Allora l'unione dei due cammini, di cui uno percorso in senso inverso, forma una curva chiusa e:

$$0 = \oint_{C_1 \cup \{-C_2\}} f(z) dz = \int_{C_1} f(z) dz - \int_{C_2} f(z) dz$$

In questo caso ha senso parlare di $\int_a^b f(z) dz$ in quanto l'integrale non dipende dal percorso.

La formulazione data del teorema richiede l'olomorfismo in una regione semplicemente connessa. Questa condizione può essere indebolita prendendo in considerazione anche

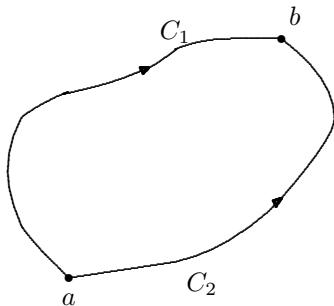


Figura 1.12: Indipendenza dal cammino.

regioni multiplamente connesse che si rendono semplicemente connesse mediante la formazione di barriere, linee di taglio. In tal caso occorre però fare attenzione alla orientazione della curva γ su cui si integra, o meglio, alla orientazione relativa delle varie curve che compongono la curva totale. Le varie curve devono avere un orientamento relativo tra loro tale che, da una loro medesima parte (in genere si intende a sinistra per curve orientate positivamente) racchiudano una regione connessa in cui la funzione è olomorfa.

Osservazione. Una curva chiusa semplice divide il piano in due regioni separate, una delle quali si può pensare interna, o a sinistra, mentre l'altra esterna o a destra, rispetto all'orientamento prescelto sulla curva (**teorema di Jordan**).

Si può così generalizzare il teorema di Cauchy dicendo che se $f(z)$ è olomorfa in una regione D (che possiamo supporre connessa) e γ è una curva chiusa in D , orientata in modo tale che il suo interno sia tutto contenuto nella regione di olomorfismo della funzione, allora:

$$\oint_{\gamma} f(z) dz = 0.$$

L'idea sostanziale di tale estensione è data dal fatto che mediante opportuni tagli la regione considerata può essere resa semplicemente connessa e la curva γ resa equivalente ad una curva chiusa semplice, regolare a tratti, compresa in tale regione semplicemente connessa. Cerchiamo di chiarire (senza entrare nella dimostrazione generale) con una figura di esempio. Sia D un dominio aperto come in figura 1.13 (a) in cui una funzione $f(z)$ risulta olomorfa e $\gamma = \gamma_1 \cup \gamma_2$. L'orientamento è tale che le curve γ_1, γ_2 hanno in comune alla loro sinistra una regione tutta contenuta in D . Modifichiamo le curve γ_1 e γ_2 aprendole e aggiungendo dei tratti orientati che congiungono le due curve formando un'unica curva chiusa Γ (vedi figura 1.13 (b)). Sostanzialmente abbiamo scelto due punti vicini A, A' su γ_1 , ottenendo l'arco C_1 , due punti vicini B e B' su γ_2 , con il conseguente arco C_2 , e aggiunto i segmenti AB e $B'A'$. Senza variare gli orientamenti precedenti abbiamo ottenuto un'unica curva Γ chiusa, regolare a tratti, che racchiude una regione

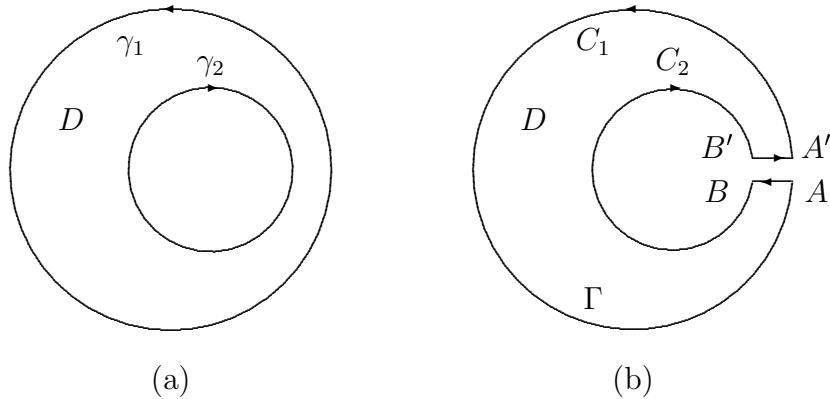


Figura 1.13: Modifica del percorso di integrazione.

semplicemente connessa in cui $f(z)$ è analitica:

$$0 = \oint_{\Gamma} f(z) dz = \left(\int_{C_1} + \int_{AB} + \int_{C_2} + \int_{B'A'} \right) f(z) dz.$$

Facendo ora tendere A e B verso A' e B' rispettivamente, gli integrali lungo i segmenti si annullano tra loro e gli integrali sugli archi tendono agli integrali sulle curve chiuse γ_1 , γ_2 :

$$\left(\oint_{\gamma_1} + \oint_{\gamma_2} \right) f(z) dz = \oint_{\gamma} f(z) dz = 0.$$

1.3.1 Rappresentazione integrale di Cauchy

Dal teorema di Cauchy è possibile derivare una formula integrale che è molto importante per lo sviluppo della teoria, nonché per una ampia varietà di applicazioni in problemi fisici.

Teo. 1.2 (Rappresentazione integrale di Cauchy) Sia $f(z)$ una funzione olomorfa in una regione aperta semplicemente connessa D , e sia γ una curva semplice chiusa, regolare a tratti, contenuta in D , allora, se z non appartiene a γ , si ha:

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{z' - z} dz' = \begin{cases} f(z) & \text{se } z \text{ è interno a } \gamma \\ 0 & \text{se } z \text{ è esterno a } \gamma \end{cases} \quad (1.74)$$

dove l'integrazione lungo γ si deve prendere in senso antiorario (positivo).

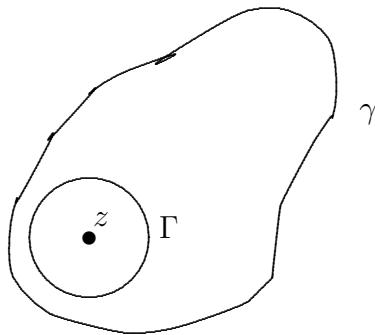


Figura 1.14: Rappresentazione di Cauchy.

Dim. 1.2 Se z è esterno alla curva γ il risultato è immediato in quanto la funzione integranda (come funzione di z'):

$$g(z') = \frac{f(z')}{z' - z}$$

risulta olomorfa in tutto l'interno della curva γ , sia sulla curva, sia in un intorno opportuno della curva, quindi in una regione semplicemente连nessa che contiene γ e, per il teorema integrale di Cauchy:

$$\oint_{\gamma} g(z') dz' = 0,$$

in accordo con la (1.74).

Se invece z è interno alla curva γ consideriamo il rapporto:

$$\frac{f(z') - f(z)}{z' - z}$$

Poiché f è olomorfa, essa è anche continua (conseguenza della derivabilità) quindi, fissato ϵ , abbiamo:

$$|f(z') - f(z)| < \epsilon \text{ quando } |z' - z| < \delta(\epsilon),$$

con $\delta(\epsilon)$ opportuno. Sia allora Γ un cerchio contenuto in D , di centro z e raggio $r < \delta(\epsilon)$ (possiamo assumere il raggio r abbastanza piccolo in modo che il cerchio sia interno a γ). Parametrizziamo il cerchio in direzione antioraria:

$$z'(\theta) = z + re^{i\theta} \quad 0 \leq \theta \leq 2\pi,$$

allora, per $z' \in \Gamma$:

$$\left| \frac{f(z') - f(z)}{z' - z} \right| = \frac{|f(z') - f(z)|}{r} < \frac{\epsilon}{r},$$

per cui, utilizzando la diseguaglianza di Darboux:

$$\left| \oint_{\Gamma} \frac{f(z') - f(z)}{z' - z} dz' \right| \leq \frac{\epsilon}{r} 2\pi r = 2\pi\epsilon$$

$$\lim_{\epsilon \rightarrow 0} \oint_{\Gamma} \frac{f(z') - f(z)}{z' - z} dz' = 0,$$

(si noti che Γ dipende da ϵ tramite $r < \delta(\epsilon)$). Possiamo allora dire che:

$$\oint_{\Gamma} \frac{f(z)}{z' - z} dz' = f(z) \int_0^{2\pi} \frac{1}{re^{i\theta}} ire^{i\theta} d\theta = 2\pi i f(z)$$

$$\lim_{\epsilon \rightarrow 0} \oint_{\Gamma} \frac{f(z')}{z' - z} dz' = \lim_{\epsilon \rightarrow 0} \oint_{\Gamma} \frac{f(z') - f(z)}{z' - z} dz' + \lim_{\epsilon \rightarrow 0} \oint_{\Gamma} \frac{f(z)}{z' - z} dz' = 2\pi i f(z).$$

D'altra parte, compiendo ragionamenti analoghi a quelli del paragrafo precedente, l'integrale sulla circonferenza Γ è uguale all'integrale sulla curva γ :

$$\oint_{\gamma} \frac{f(z')}{z' - z} dz' = \oint_{\Gamma} \frac{f(z')}{z' - z} dz'$$

in quanto la funzione integranda risulta olomorfa in ogni punto z' di D diverso da z . In questo modo il parametro ϵ scompare e abbiamo:

$$\oint_{\gamma} \frac{f(z')}{z' - z} dz' = 2\pi i f(z), \quad (1.75)$$

che conclude la prova del teorema.

— ■ —

Notiamo come qualsiasi percorso possa essere modificato a piacere (senza però variare le sue caratteristiche topologiche) con continuità, con l'unica condizione che le sue variazioni avvengano rimanendo sempre all'interno della regione di olomorfismo della funzione integranda.

Osservazione. Il teorema non dice nulla quando il punto z appartiene alla curva γ . In effetti se $z \in \gamma$ l'integrando presenta in generale una singolarità (a meno che $f(z')$ non si annulli per $z' \rightarrow z$) e l'integrale in generale non esiste. Sarebbe come integrare $1/x$ su un intervallo contenente l'origine $x = 0$.

Osservazione. Poiché la maggior parte dei risultati sono basati sulla formula integrale di Cauchy, per evitare ambiguità, d'ora in poi adottiamo la convenzione che ogni integrazione

lungo un percorso chiuso avvenga in senso antiorario, a meno di specificare il contrario. Più in generale sceglieremo sempre il verso sulla curva chiusa in modo tale che la regione interna si trovi alla sinistra della curva. Ogni integrazione in senso opposto comporta un cambiamento di segno.

La (1.75) è nota anche sotto il nome di **formula integrale di Cauchy**, ed esprime un risultato veramente potente, ricco di conseguenze. Una prima conseguenza della formula integrale è che la derivata di una funzione olomorfa in un aperto è a sua volta derivabile, cioè olomorfa. Infatti possiamo scrivere:

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{z' - z} dz', \quad (1.76)$$

dove γ è una curva regolare chiusa che racchiude il punto z al suo interno, e nessuna altra singolarità di $f(z')$. Vediamo ora come sia possibile derivare la (1.76) quante volte si voglia sotto il segno di integrale, sfruttando la continuità di $f(z')$:

$$\begin{aligned} \frac{f(z + \Delta z) - f(z)}{\Delta z} &= \frac{1}{2\pi i} \oint_{\gamma} \left\{ \frac{f(z')}{z' - z - \Delta z} - \frac{f(z')}{z' - z} \right\} \frac{dz'}{\Delta z} \\ &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z') dz'}{(z' - z)^2} \\ &\quad + \frac{1}{2\pi i} \oint_{\gamma} \left\{ \frac{f(z')}{z' - z - \Delta z} - \frac{f(z')}{z' - z} - \frac{f(z') \Delta z}{(z' - z)^2} \right\} \frac{dz'}{\Delta z}, \end{aligned}$$

Inoltre:

$$\begin{aligned} &\left| \oint_{\gamma} \left\{ \frac{f(z')}{\Delta z(z' - z - \Delta z)} - \frac{f(z')}{\Delta z(z' - z)} - \frac{f(z')}{(z' - z)^2} \right\} dz' \right| \\ &= \left| \oint_{\gamma} \frac{f(z')}{(z' - z)^2(z' - z - \Delta z)} dz' \right| |\Delta z| \xrightarrow[\Delta z \rightarrow 0]{} 0, \end{aligned}$$

in quanto z non appartiene a γ e l'integrando è limitato. Ciò prova che anche per la derivata di $f(z)$ vale una rappresentazione integrale:

$$\frac{df}{dz}(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{(z' - z)^2} dz', \quad (1.77)$$

e si può proseguire nella derivazione sotto il segno di integrale, per cui $\frac{df}{dz}$ è a sua volta derivabile, olomorfa e in generale:

$$\boxed{\frac{d^n f}{dz^n}(z) = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(z')}{(z' - z)^{n+1}} dz' \quad (z \text{ interno a } \gamma)} \quad (1.78)$$

Questo risultato è molto importante e non ha una sua controparte nelle funzioni di variabile reale. Nel caso di variabili reali la derivabilità non implica affatto che le derivate siano a loro volta continue e derivabili ulteriormente. Per le funzioni di variabile complessa è sufficiente avere la derivabilità una volta per avere la continuità e derivabilità ad ogni ordine con conseguenze di regolarità altissime. Le funzioni olomorfe sono quindi molto “dolci” e nel loro dominio di olomorfismo non presentano irregolarità di alcun genere. (salti, cuspidi, ecc.)

1.3.2 Rappresentazioni integrali

La rappresentazione integrale di Cauchy è un esempio del concetto più generale di **rap-presentazione integrale**. Si dice in generale che si ha una rappresentazione integrale per una funzione $f(z)$ se si può scrivere:

$$f(z) = \int_{\gamma} K(z, z') g(z') dz', \quad (1.79)$$

con γ curva semplice, che assumiamo di lunghezza finita. Nell'integrando (1.79) la variabile z gioca il ruolo di un parametro ed in generale la rappresentazione sarà valida solo quando $z \in D$, con D opportuno dominio del piano complesso. La funzione $K(z, z')$ prende il nome di **kernel** o **nucleo** integrale della rappresentazione.

Quando il kernel assume la forma:

$$K(z, z') = \frac{1}{z' - z},$$

si parla di **rappresentazione spettrale** della funzione $f(z)$:

$$f(z) = \int_{\gamma} \frac{g(z')}{z' - z} dz', \quad (1.80)$$

e la funzione $g(z')$ viene detta **funzione spettrale**. Abbiamo a proposito il seguente teorema:

Teo. 1.3 Qualsiasi funzione $f(z)$ per la quale si possa dare una rappresentazione spettrale del tipo (1.80), dove γ è una curva regolare (aperta o chiusa) di lunghezza finita, e $g(z')$ è una funzione continua su γ , è una funzione olomorfa per ogni $z \notin \gamma$.

Osservazione. Se γ è una curva chiusa e $g(z')$ olomorfa, ritroviamo sostanzialmente la rappresentazione integrale di Cauchy, che viene a definire due diverse funzioni, una per z interna alla curva γ , ed una per z esterna alla curva γ .

Dim. 1.3 La dimostrazione ripercorre sostanzialmente quanto già visto a proposito della rappresentazione integrale di Cauchy per la derivata di una funzione olomorfa, mostrando che si può derivare la (1.80) rispetto a z sotto il segno di integrale. Consideriamo la quantità:

$$\Delta = \left| \frac{f(z + \Delta z) - f(z)}{\Delta z} - \int_{\gamma} \frac{g(z')}{(z' - z)^2} dz' \right|,$$

con z e $z + \Delta z$ non appartenenti a γ . Operando le medesime manipolazioni algebriche la (1.80) comporta che:

$$\Delta = \left| \Delta z \int_{\gamma} \frac{g(z')}{(z' - z)^2 (z' - z - \Delta z)} dz' \right|,$$

e per ipotesi la funzione integranda risulta continua e quindi limitata su γ (a sua volta di lunghezza finita). Pertanto l'integrale risulta limitato e il fattore esplicito Δz comporta che:

$$\Delta \xrightarrow[\Delta z \rightarrow 0]{} 0,$$

e $f(z)$ risulta derivabile, cioè olomorfa, quando $z \notin \gamma$:

$$\frac{df(z)}{dz} = \int_{\gamma} \frac{g(z')}{(z' - z)^2} dz', \quad z \notin \gamma. \quad (1.81)$$

— ■ —

Iterando il procedimento possiamo derivare ulteriormente sotto il segno di integrale ottenendo in generale:

$$\frac{d^n f}{dz^n} = n! \int_{\gamma} \frac{g(z')}{(z' - z)^{n+1}} dz', \quad z \notin \gamma. \quad (1.82)$$

Osservazione. Rimarchiamo ulteriormente che non richiediamo in generale che la funzione $g(z')$ sia olomorfa, ma semplicemente che essa sia continua sulla curva γ .

Tornando ora alla rappresentazione integrale generale (1.79), possiamo dimostrare che è lecito derivare sotto il segno di integrale.

Teo. 1.4 Se per una funzione $f(z)$ vale una rappresentazione integrale:

$$f(z) = \int_{\gamma} k(z, z') g(z') dz', \quad z \in D \quad (1.83)$$

con γ curva regolare di lunghezza finita e z appartenente ad un dominio D , che assumiamo aperto e semplicemente connesso, e valgono le seguenti ipotesi:

- a) per $z \in D$ la funzione $k(z, z')$ è una funzione olomorfa di z per ogni z' appartenente a γ ;
- b) per ogni $z \in D$ la funzione $k(z, z') g(z')$ è una funzione continua di z' , quando $z' \in \gamma$;
allora la funzione $f(z)$ è olomorfa in D e:

$$\frac{df(z)}{dz} = \int_{\gamma} \frac{\partial k(z, z')}{\partial z} g(z') dz', \quad z \in D. \quad (1.84)$$

Dim. 1.4 L'ipotesi a) permette di scrivere una formula integrale di Cauchy per $k(z, z')$:

$$k(z, z') = \frac{1}{2\pi i} \oint_C \frac{k(t, z')}{(t - z)} dt,$$

con C curva chiusa contenuta in D e che racchiude il punto z al suo interno. Introducendo tale espressione nella (1.83) e scambiando l'ordine delle integrazioni (cosa lecita per la continuità della funzione integranda, come per gli integrali ordinari – un integrale al contorno può sempre essere ricondotto ad un integrale ordinario), si ha:

$$f(z) = \frac{1}{2\pi i} \oint_C dt \frac{\int_{\gamma} k(t, z') g(z') dz'}{t - z}.$$

La funzione:

$$\int_{\gamma} k(t, z') g(z') dz'$$

è chiaramente una funzione continua di t , e l'ipotesi b) del teorema ci permette di poter derivare sotto il segno di integrale nell'integrazione lungo C (abbiamo un caso particolare del teorema precedente con una rappresentazione spettrale):

$$\frac{df}{dz} = \frac{1}{2\pi i} \oint_C dt \frac{\int_{\gamma} k(t, z') g(z') dz'}{(t - z)^2}.$$

Scambiando di nuovo l'ordine di integrazione e notando che (sempre in virtù del teorema precedente):

$$\frac{\partial k(z, z')}{\partial z} = \frac{1}{2\pi i} \oint_C dt \frac{k(t, z')}{(t - z)^2},$$

otteniamo l'equazione (1.84) che ci fornisce esplicitamente la derivabilità e quindi l'olomorfismo di $f(z)$ nel dominio D .

1.3.3 Teorema di Morera

Veniamo ora all'importante teorema di Morera¹² che fornisce sostanzialmente l'inverso del teorema di Cauchy.

Teo. 1.5 (Morera) *Se l'integrale*

$$\oint_{\gamma} f(z) dz$$

di una funzione continua in una regione aperta e connessa D si annulla per ogni curva chiusa, semplice e regolare a tratti, γ interna alla regione D , allora $f(z)$ è olomorfa in D .

Dim. 1.5 L'annullarsi:

$$\oint_{\gamma} f(z) dz = 0,$$

per ogni curva γ qualsiasi, implica che l'integrale di $f(z)$ non dipende dal percorso, per cui, scelto un punto $z_0 \in D$ possiamo definire:

$$F(z) = \int_{z_0}^z f(z') dz',$$

con z variabile in D . Vediamo ora che $F(z)$ è olomorfa con derivata $f(z)$, per cui essendo la derivata di una funzione olomorfa, avremo che f è olomorfa. Abbiamo:

$$\begin{aligned} \frac{F(z + \Delta z) - F(z)}{\Delta z} &= \frac{1}{\Delta z} \left(\int_{z_0}^{z + \Delta z} - \int_{z_0}^z \right) f(z') dz' \\ &= \frac{1}{\Delta z} \int_z^{z + \Delta z} f(z') dz' = \frac{1}{\Delta z} \int_z^{z + \Delta z} (f(z) + f(z') - f(z)) dz' \\ &= f(z) + \frac{1}{\Delta z} \int_z^{z + \Delta z} (f(z') - f(z)) dz'. \end{aligned}$$

¹² **Giacinto Morera** (Novara, 18 luglio 1856 – Torino, 8 febbraio 1909) è stato un matematico italiano. Si laureò a Torino in ingegneria ed in matematiche pure. Docente di meccanica all'Università di Torino, sviluppò un teorema sulla teoria delle funzioni di variabile complessa, il Teorema di Morera.



Scegliendo un tratto rettilineo tra z e $z + \Delta z$, per la disuguaglianza di Darboux, abbiamo:

$$\left| \frac{1}{\Delta z} \int_z^{z+\Delta z} [f(z') - f(z)] dz' \right| \leq \max_{z' \in [z, z+\Delta z]} |f(z') - f(z)| \xrightarrow[\Delta z \rightarrow 0]{} 0,$$

per la continuità di f , per cui il limite del rapporto incrementale di F esiste e possiamo scrivere:

$$\frac{dF}{dz}(z) = f(z)$$

per cui $F(z)$ e $f(z)$ sono olomorfe.

— ■ —

Osservazione. Questo teorema non è esattamente l'inverso del teorema di Cauchy in quanto non abbiamo alcuna ipotesi sull'aperto D che deve essere solamente un aperto connesso, mentre nel teorema di Cauchy richiediamo che la regione sia semplicemente connessa.

1.3.4 La serie di Taylor.

Le funzioni olomorfe hanno la prerogativa molto importante della sviluppabilità in serie di Taylor¹³. Supponiamo infatti che $f(z)$ sia una funzione olomorfa all'interno di un disco D centrato in $z = z_0$, e raggio ρ . Allora se $z \in D$, abbiamo:

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{z' - z} dz',$$

dove γ è una circonferenza centrata in z_0 , contenuta nel disco D , e che racchiuda anche z al suo interno. Allora possiamo scrivere:

$$\begin{aligned} \frac{1}{z' - z} &= \frac{1}{(z' - z_0) - (z - z_0)} = \frac{1}{z' - z_0} \frac{1}{1 - \frac{(z - z_0)}{(z' - z_0)}} \\ &= \frac{1}{z' - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{z' - z_0} \right)^n \end{aligned}$$

¹³ **Brook Taylor** (Edmonton, Inghilterra, 18 agosto 1685 – Londra, Inghilterra, 29 dicembre 1731) è stato un matematico inglese, il quale aggiunse alla matematica una nuova branca oggi conosciuta come calcolo delle differenze finite, inventò l'integrazione per parti e scoprì la famosa formula nota come espansione di Taylor.



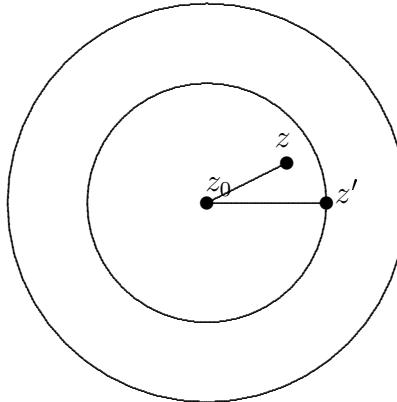


Figura 1.15: Serie di Taylor.

dove lo sviluppo in serie geometrico è giustificato in quanto (vedi figura 1.15)

$$\left| \frac{z - z_0}{z' - z_0} \right| < 1.$$

Pertanto

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} dz' \sum_{n=0}^{\infty} \frac{f(z')}{(z' - z_0)^{n+1}} (z - z_0)^n,$$

e poiché ogni serie di potenze è uniformemente convergente all'interno del suo raggio di convergenza, possiamo integrare termine a termine ottenendo uno sviluppo in serie di Taylor per $f(z)$:

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

$$a_n = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{(z' - z_0)^{n+1}} dz' = \frac{1}{n!} \frac{d^n f}{dz^n}(z_0),$$

(1.85)

ricordando la formula integrale di Cauchy per le derivate.

Sorge ora spontanea una questione. Quale è il raggio di convergenza della serie? La risposta è fornita direttamente dalla formula (1.85) su cui si basa l'espansione in serie di Taylor. Sostanzialmente tale formula cessa di valere quando il cerchio γ incontra una singolarità (cioè un punto di non olomorfismo) per $f(z)$. Notiamo inoltre che nella espressione (1.85) per i coefficienti possiamo rilassare la condizione che la curva γ sia una circonferenza centrata in z_0 , e questa può essere deformata in una qualsiasi curva chiusa a patto di non includere punti singolari della funzione ed avere sempre z_0 al suo interno.

Il risultato mostra come ogni funzione olomorfa sia quindi anche analitica (sviluppabile in serie di potenze), e viceversa, è una semplice conseguenza del teorema di Morera vedere come ogni funzione analitica sia anche olomorfa. Infatti se f è sviluppata in serie di

potenze in un intorno dell'origine (non è restrittivo considerare l'origine), allora all'interno del disco di convergenza:

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad |z| < \rho,$$

e la convergenza risulta uniforme, per cui integrando la serie su una curva γ chiusa e interna al disco di convergenza, possiamo integrare termine a termine:

$$\oint_{\gamma} f(z) dz = \oint_{\gamma} \left\{ \sum_{n=0}^{\infty} a_n z^n \right\} dz = \sum_{n=0}^{\infty} a_n \oint_{\gamma} z^n dz = 0,$$

in quanto z^n è olomorfa per ogni n . $f(z)$ verifica quindi il teorema di Morera e all'interno del disco di convergenza risulta olomorfa.

Ricordiamo che una funzione viene detta **olomorfa** in un punto z_0 se in tale punto essa risulta derivabile, e in tal caso diremo anche che essa è analitica in tale punto. L'insieme dei punti del dominio in cui una funzione è olomorfa è detto **dominio di analiticità** della funzione. Se il dominio di analiticità coincide con tutto il piano complesso, la funzione viene detta **intera** (chiaramente essa risulta sviluppabile in serie di potenze in un intorno di ogni punto con raggio di convergenza infinito).

Un punto in cui la funzione risulta olomorfa si dice anche punto **regolare** della funzione. Se una funzione non è olomorfa in un punto, tale punto viene detto punto **singolare** della funzione. Può capitare che una funzione sia regolare in tutto un intorno di un dato punto escluso il punto stesso. In questo caso il punto viene detto punto **singolare isolato** per la funzione.

1.3.5 La serie di Laurent.

Abbiamo visto che se una funzione è olomorfa in un punto z_0 e in tutto un suo intorno circolare, allora è sviluppabile in serie di Taylor attorno a quel punto. Può capitare spesso che una funzione $f(z)$ sia olomorfa in tutta una regione anulare, cioè a forma di corona circolare, attorno a un punto z_0 (il centro della corona), senza per questo essere necessariamente olomorfa per $z = z_0$. In tal caso vale ancora uno sviluppo in serie di potenze di $z - z_0$ per $f(z)$, ma in generale con potenze sia positive che negative, cioè:

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n (z - z_0)^n, \tag{1.86}$$

anche con $n < 0$.

Osservazione. La serie (1.86) è da intendersi come somma di due serie distinte, la serie costituita da potenze positive e quella formata da potenze negative. La parte di potenze negative si può scrivere:

$$\sum_{n<0} a_n (z - z_0)^n = \sum_{n=1}^{\infty} a_{-n} \left(\frac{1}{z - z_0} \right)^n,$$

e sarà convergente all'esterno di un opportuno raggio di convergenza (anzi, in tal caso, uniformemente convergente).

La serie (1.86) risulterà convergente (uniformemente) all'interno di una corona circolare, centrata in z_0 e di raggi r_1 e r_2 . Il risultato notevole è il seguente.

Teo. 1.6 (Serie di Laurent) *Se $f(z)$ è olomorfa in una regione a corona circolare di centro z_0 e compresa tra due raggi r_1 , r_2 ,*

$$r_1 < |z - z_0| < r_2,$$

*allora $f(z)$ è sviluppabile in **serie di Laurent** attorno a z_0 , cioè esprimibile mediante la serie uniformemente convergente:*

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n (z - z_0)^n, \quad (1.87)$$

in ogni punto z interno alla corona, con:

$$a_n = \frac{1}{2\pi i} \oint_{\Gamma} \frac{f(z')}{(z' - z_0)^{n+1}} dz', \quad (1.88)$$

e Γ è una qualsiasi curva chiusa semplice interna alla corona circolare e che contiene z_0 al suo interno.

Notiamo che la relazione (1.88) per esprimere i coefficienti della serie di Laurent¹⁴ è la stessa dello sviluppo in serie di Taylor. Nel caso del teorema di Taylor avevamo in più l'espressione del coefficiente n -esimo, per $n > 0$, anche tramite la derivata n -esima di $f(z)$ nel punto z_0 (ciò non è più vero per i coefficienti della serie di Laurent).

Dim. 1.6 Sia z interno alla corona circolare e sia C la curva in figura 1.16 composta dalle due circonferenze $-\Gamma_1$ e Γ_2 , dal cerchio $-\gamma$ e dai tratti di segmenti orientati. C racchiude una regione di olomorfismo semplicemente连nessa per $f(z)$ per cui:

$$\oint_C \frac{f(z')}{z' - z} dz' = 0.$$

¹⁴ **Pierre Alphonse Laurent** (Parigi, 18 luglio 1813 – Parigi, 2 settembre 1854) è stato un matematico francese meglio noto come lo scopritore della serie di Laurent, un'espansione di una funzione in una serie di potenze infinita, generalizzando l'espansione in serie di Taylor. Il suo risultato venne sottoposto a giudizio per il Grand Prize dell'Académie des Sciences nel 1843, ma la sua tesi venne consegnata dopo la data di scadenza così che non venne pubblicata e mai considerata per il premio. Il suo lavoro venne pubblicato solo dopo la sua morte.

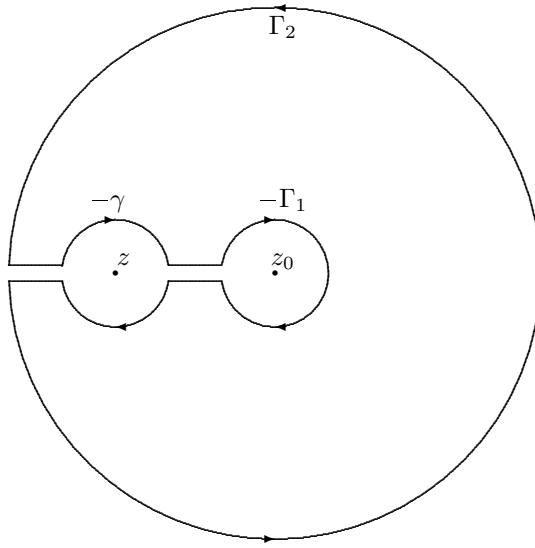


Figura 1.16: Serie di Laurent.

Eliminando i segmenti,

$$\begin{aligned} \oint_{\Gamma_2} \frac{f(z')}{z' - z} dz' - \oint_{\Gamma_1} \frac{f(z')}{z' - z} dz' - \oint_{\gamma} \frac{f(z')}{z' - z} dz' &= 0 \\ \oint_{\gamma} \frac{f(z')}{z' - z} dz' &= \oint_{\Gamma_2} \frac{f(z')}{z' - z} dz' - \oint_{\Gamma_1} \frac{f(z')}{z' - z} dz' \\ f(z) &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{z' - z} dz' = \frac{1}{2\pi i} \oint_{\Gamma_2} \frac{f(z')}{z' - z} dz' - \frac{1}{2\pi i} \oint_{\Gamma_1} \frac{f(z')}{z' - z} dz' \end{aligned}$$

Il primo integrale può essere espresso in potenze positive di $z - z_0$ esattamente come nel caso della serie di Taylor, in quanto:

$$z' - z = z' - z_0 - (z - z_0),$$

$$\frac{1}{z' - z} = \frac{1}{z' - z_0} \frac{1}{1 - \frac{z - z_0}{z' - z_0}} = \frac{1}{z' - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{z' - z_0} \right)^n = \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{(z' - z_0)^{n+1}},$$

grazie al fatto che, con $z' \in \Gamma_2$, abbiamo:

$$\left| \frac{z - z_0}{z' - z_0} \right| < 1.$$

Pertanto:

$$\frac{1}{2\pi i} \oint_{\Gamma_2} \frac{f(z')}{z' - z} dz' = \sum_{n \geq 0} a_n (z - z_0)^n$$

$$a_n = \frac{1}{2\pi i} \oint_{\Gamma_2} \frac{f(z')}{(z' - z_0)^{n+1}} dz',$$

in perfetta analogia con la serie di Taylor.

Per l'integrale su Γ_1 abbiamo invece:

$$\left| \frac{z' - z_0}{z - z_0} \right| < 1,$$

e possiamo allora sviluppare il denominatore:

$$\begin{aligned} \frac{1}{z' - z} &= -\frac{1}{z - z'} = -\frac{1}{z - z_0 - (z' - z_0)} \\ &= -\frac{1}{(z - z_0)} \frac{1}{1 - \frac{z' - z_0}{z - z_0}} = -\frac{1}{z - z_0} \sum_{n=0}^{\infty} \left(\frac{z' - z_0}{z - z_0} \right)^n \\ &= -\sum_{n=0}^{\infty} \frac{(z' - z_0)^n}{(z - z_0)^{n+1}} = -\sum_{n<0} \frac{1}{(z' - z_0)^{n+1}} (z - z_0)^n. \end{aligned}$$

In questo modo si ottiene una serie uniformemente convergente per cui si può integrare termine a termine e si ha:

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\Gamma_1} \frac{f(z')}{z' - z} dz' &= -\frac{1}{2\pi i} \sum_{n<0} \left[\oint_{\Gamma_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz' \right] (z - z_0)^n = -\sum_{n<0} a_n (z - z_0)^n \\ a_n &= \frac{1}{2\pi i} \oint_{\Gamma_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz'. \end{aligned}$$

Otteniamo così la serie di Laurent:

$$f(z) = \sum_{n=-\infty}^{n=+\infty} a_n (z - z_0)^n,$$

e poiché le curve di integrazione nella regione di olomorfismo possono essere variate a piacere, possiamo unificare Γ_1 e Γ_2 in una unica curva chiusa Γ (che racchiuda z_0 al suo interno):

$$a_n = \frac{1}{2\pi i} \oint_{\Gamma} \frac{f(z')}{(z' - z_0)^{n+1}} dz'.$$

— ■ —

Osservazione. Dalla dimostrazione precedente risulta chiaro il significato dei raggi r_1 e r_2 che delimitano la corona circolare. r_1 delimita un disco nel piano complesso all'esterno del quale la serie di potenze negative risulta convergente, e non è proibito che risulti $r_1 = 0$ (la funzione è olomorfa in un intorno di z_0 , escluso al più z_0 stesso). r_2 individua un disco all'interno del quale la serie di potenze positive risulta convergente, e non è escluso il caso $r_2 = \infty$ nel caso che la funzione sia olomorfa ovunque (escluso al più z_0 o un intorno di z_0).

1.3.6 Zeri e punti singolari

Diamo una nomenclatura relativa a punti particolari di una funzione.

Zeri

Se una funzione $f(z)$ si annulla in un punto $z = z_0$ allora tale punto è detto **zero** di $f(z)$. Diremo che tale zero è di **ordine** n se f è olomorfa in z_0 e:

$$f(z_0) = \frac{df}{dz}(z_0) = \dots = \frac{d^{n-1}f}{dz^{n-1}}(z_0) = 0, \quad (1.89)$$

ma:

$$\frac{d^n f}{dz^n}(z_0) \neq 0. \quad (1.90)$$

In questo caso se f è olomorfa in un intorno di z_0 (compreso z_0) possiamo svilupparla in serie di Taylor con i primi n coefficienti a_0, a_1, \dots, a_{n-1} nulli (essendo proporzionali alle derivate):

$$\begin{aligned} f(z) &= a_n(z - z_0)^n + a_{n+1}(z - z_0)^{n+1} + \dots \\ &= (z - z_0)^n \sum_{k=0}^{\infty} a_{n+k}(z - z_0)^k = (z - z_0)^n g(z), \end{aligned}$$

con $g(z)$ regolare e non nulla per $z = z_0$, e per continuità non nulla in tutto un intorno di z_0 . Questo comporta che esiste tutto un intorno di z_0 in cui, a parte il punto $z = z_0$, la funzione $f(z)$ è non nulla. Il punto z_0 risulta quindi uno zero isolato. Se non possiamo individuare l'ordine dello zero, cioè, pur essendo $f(z)$ olomorfa in un intorno, tutte le derivate sono nulle, allora la sviluppabilità in serie di Taylor impone che $f(z)$ sia identicamente nulla in tale intorno.

Questo implica che gli zeri di una funzione analitica sono isolati, cioè formano un insieme discreto (privo di punti di accumulazione), all'interno del dominio di olomorfismo della funzione stessa. Se un punto z_0 è un punto di accumulazione di zeri per una funzione $f(z)$, allora esso sarà necessariamente un punto non di olomorfismo della funzione, per

cui è un punto singolare. L'unica eccezione permessa è data dalla funzione identicamente nulla (chiaramente olomorfa) nel suo dominio di definizione.

Una conseguenza importante di questo fatto è che se due funzioni olomorfe $f_1(z)$ e $f_2(z)$ coincidono su un insieme di punti il quale abbia anche un solo punto di accumulazione interno al campo di olomorfismo di entrambe, esse sono necessariamente identiche. Infatti l'insieme degli zeri della differenza $f_1(z) - f_2(z)$ presenta un punto di accumulazione regolare, e ciò è possibile solo se $f_1(z) - f_2(z)$ è identicamente nulla. Questa considerazione, apparentemente banale, è molto importante per poter estendere in maniera analitica una funzione su un dominio più vasto quando questa, per motivi tecnici, è definita originariamente su una regione limitata (ad esempio, tramite uno sviluppo in serie all'interno di un disco di raggio finito).

Punti singolari isolati

Supponiamo ora che $f(z)$ abbia una singolarità isolata in un punto $z = z_0$ e sia analitica all'interno di un disco centrato in quel punto, escluso quindi z_0 . Possiamo allora sviluppare in serie di Laurent la funzione attorno al punto z_0 :

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n + \frac{a_{-1}}{z - z_0} + \frac{a_{-2}}{(z - z_0)^2} + \dots$$

È chiaro dall'espansione che se $f(z)$ è singolare in z_0 , almeno uno dei coefficienti a_{-n} è non nullo. Se il coefficiente a_{-n} è non nullo, mentre tutti i successivi (relativi alle potenze negative) sono nulli:

$$a_{-(n+1)} = a_{-(n+2)} = \dots = a_{-(n+k)} = \dots = 0,$$

allora il punto z_0 è detto un **polo di ordine n** , e in tal caso la somma:

$$\frac{a_{-1}}{z - z_0} + \frac{a_{-2}}{(z - z_0)^2} + \dots + \frac{a_{-n}}{(z - z_0)^n}$$

è detta **parte principale** di $f(z)$ in $z = z_0$. Se $n = 1$, cioè $a_{-1} \neq 0$ mentre tutti gli altri coefficienti delle potenze negative sono nulli, allora z_0 è detto **polo semplice**. Se $f(z)$ ha un polo di ordine n possiamo scrivere:

$$\begin{aligned} f(z) &= \sum_{k \geq 0} a_k(z - z_0)^k + \frac{a_{-1}}{z - z_0} + \frac{a_{-2}}{(z - z_0)^2} + \dots + \frac{a_{-n}}{(z - z_0)^n} \\ &= \frac{1}{(z - z_0)^n} \left[a_{-n} + \dots + a_{-1}(z - z_0)^{n-1} + \sum_{k \geq 0} a_k(z - z_0)^{n+k} \right] \\ &= \frac{1}{(z - z_0)^n} \sum_{k=0}^{\infty} a_{-n+k}(z - z_0)^k \\ &= \frac{h(z)}{(z - z_0)^n}, \end{aligned}$$

con $h(z)$ regolare e non nulla in z_0 (e per continuità in tutto un intorno). Risulta quindi facile verificare se una funzione ha un polo di ordine n , in quanto in tal caso la funzione reciproca $1/f(z)$ ha uno zero di ordine n nel medesimo punto.

Una funzione che sia analitica in una regione (aperta) del piano complesso, escluso al più un insieme di punti dove la funzione ha dei poli, è detta **meromorfa** in tale regione.

Quando nella espansione di Laurent attorno ad un punto singolare isolato si presentano un numero infinito di coefficienti non nulli a_{-k} per le potenze negative, il punto z_0 viene detto una **singolarità essenziale** isolata. La caratteristica più importante di una singolarità essenziale è fornita dal seguente teorema di Weierstrass¹⁵ (che non dimostriamo).

Teo. 1.7 (Weierstrass) Se una funzione $f(z)$ ha una singolarità essenziale isolata in un punto z_0 , allora per qualsiasi $\epsilon > 0$ e $\delta > 0$, e per ogni numero complesso a si ha:

$$|f(z) - a| < \epsilon, \quad (1.91)$$

per qualche punto z con $|z - z_0| < \delta$.

Esprimendo tale risultato in maniera diversa possiamo dire che in un qualsiasi intorno di una singolarità essenziale isolata la funzione varia così rapidamente e ampiamente che assume qualsiasi valore (complesso).

Punto all'infinito

Fino ad ora ci siamo limitati a considerare il comportamento di una funzione analitica in un generico punto $z = z_0$, supposto al finito. Tuttavia le considerazioni fatte sono immediatamente estensibili allo studio del comportamento di una funzione analitica $f(z)$ nell'intorno del punto all'infinito. Infatti, operando la sostituzione:

$$z = \frac{1}{\zeta}, \quad (1.92)$$

e definendo:

$$\varphi(\zeta) = f\left(\frac{1}{\zeta}\right), \quad (1.93)$$

¹⁵ **Karl Theodor Wilhelm Weierstrass** (Ostenfelde, Germania, 31 ottobre 1815 – Berlino, Germania, 19 febbraio 1897) è stato un matematico tedesco spesso chiamato padre dell'analisi moderna. Si occupò di definire rigorosamente i fondamenti dell'analisi, dando per primo l'esempio di una funzione continua ovunque ma non derivabile. Il suo nome è legato al Teorema di Weierstrass, al Teorema di Bolzano-Weierstrass e al criterio di Weierstrass per la convergenza uniforme delle serie.



Weierstrass

possiamo studiare il punto all'infinito per $f(z)$ esaminando il punto $\zeta = 0$ per $\varphi(\zeta)$. Diremo così che $z = \infty$ è uno zero di ordine n , un polo di ordine n , o una singolarità essenziale isolata, se $\zeta = 0$ è uno zero di ordine n , un polo di ordine n , o una singolarità essenziale isolata, rispettivamente, per $\varphi(\zeta)$.

1.3.7 Comportamento locale di una funzione analitica

Consideriamo una funzione $f(z)$ analitica in un aperto D contenente z_0 e sia Γ una circonferenza di raggio r contenuta in D con centro z_0 . La circonferenza può essere parametrizzata come:

$$z(\theta) = z_0 + re^{i\theta} \quad 0 \leq \theta \leq 2\pi,$$

e possiamo scrivere:

$$f(z_0) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{f(z)}{z - z_0} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} ire^{i\theta} d\theta,$$

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta, \quad (\textbf{proprietà della media})$$

(1.94)

La relazione sopra dice sostanzialmente che il valore di una funzione analitica in un punto regolare egualia il valore medio della funzione su una circonferenza centrata nel punto e interna al suo dominio di analiticità.

Come conseguenza abbiamo che:

$$|f(z_0)| \leq \max_{z \in \Gamma} |f(z)|,$$

(1.95)

per cui in qualsiasi punto interno al dominio non si può avere un massimo locale per il modulo di $f(z)$ a meno che $f(z)$ non sia costante (**principio del massimo modulo**).

Si può mostrare anche che $|f(z)|$ non può avere un minimo in un punto regolare interno z_0 al dominio di analiticità, a meno che non sia $f(z_0) = 0$ o f costante. Infatti, se $f(z_0) \neq 0$ il reciproco $\frac{1}{f(z)}$ risulta regolare in z_0 e $\left| \frac{1}{f(z)} \right|$ non può avere un massimo locale.

Un risultato analogo vale sia per la parte reale che la parte immaginaria di una funzione analitica, come si può vedere considerando le funzioni

$$e^{f(z)}, \quad e^{-f(z)}, \quad e^{if(z)}, \quad e^{-if(z)}$$

che hanno lo stesso dominio di analiticità di $f(z)$, e abbiamo

$$|e^{\pm f(z)}| = e^{\pm \operatorname{Re} f(z)} \quad |e^{\pm if(z)}| = e^{\mp \operatorname{Im} f(z)}.$$

Un'altra conseguenza della formula integrale di Cauchy è costituito dal seguente teorema dovuto a Liouville¹⁶:

Teo. 1.8 (Liouville) Una funzione *interna* (olomorfa in tutto \mathbb{C}) e *limitata* deve essere costante.

Infatti, per la derivata abbiamo

$$\left| \frac{df(z)}{dz} \right| = \left| \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z')}{(z' - z)^2} dz' \right| \leq \max_{z' \in \gamma} \frac{|f(z')|}{R} \leq \frac{C}{R},$$

dove γ è una circonferenza di raggio R , centrata in z , per cui, per $R \rightarrow \infty$:

$$\frac{df}{dz} = 0, \quad f(z) = \text{cost.} .$$

¹⁶ **Joseph Liouville** (Saint-Omer, 24 marzo 1809 – Parigi, 8 settembre 1882) è stato un matematico francese. Liouville ha pubblicato in diversi ambiti della matematica, tra cui la teoria dei numeri, l'analisi complessa, la geometria differenziale. Resta ancora celebre il teorema di Liouville, risultato oggi ampiamente studiato in analisi complessa.



1.4 Residui.

Sia $f(z)$ una funzione analitica in una regione D (aperta) escluso un punto z_0 interno a D , dove $f(z)$ può avere una singolarità isolata. Allora, se γ è una curva chiusa semplice, regolare a tratti, contenuta in D , che racchiude z_0 , l'integrale (γ è supposta orientata in senso antiorario, positivo):

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz,$$

che si annulla se z_0 è un punto regolare, non si annulla in generale. Tale quantità definisce il **residuo** di $f(z)$ nel punto z_0 :

$$\text{Res } f(z)|_{z=z_0} \equiv \frac{1}{2\pi i} \oint_{\gamma} f(z) dz,$$

(1.96)

(ovviamente l'integrale non dipende dalla scelta della curva γ che circonda il punto, purché essa sia semplice, cioè compia un solo giro attorno al punto z_0 , sia contenuta all'interno del dominio di analiticità della funzione e non racchiuda altre singolarità all'infuori di $z = z_0$).

Osservazione. La definizione di residuo viene generalmente data nel caso che il punto z_0 sia un punto singolare isolato per la funzione $f(z)$, ma possiamo accettare la definizione anche per un punto regolare, per il quale il residuo è nullo in conseguenza del teorema di Cauchy.

Il valore del residuo di una funzione in un punto singolare determina, in un certo senso, l'importanza della singolarità. Vediamo un esempio fisico che ci fornisce una interessante analogia.

Consideriamo il moto di un fluido in un piano e sia:

$$\mathbf{V} = \mathbf{V}(x, y)$$

il campo di velocità del fluido (l'assunzione di moto in un piano sottointende che la velocità non abbia componente in direzione ortogonale al piano e non dipenda dalla coordinata ortogonale al piano). Allora è noto dalla dinamica dei fluidi che in ogni regione priva di sorgenti o pozzi il moto è **solenoidale**, cioè a divergenza nulla:

$$\nabla \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} = 0,$$

inoltre, se in una regione non sono presenti vortici, il moto è **irrotazionale**, cioè:

$$\nabla \times \mathbf{V} = 0.$$

In un moto piano il rotore del campo di velocità ha solo una componente (ortogonale al piano) per cui:

$$\frac{\partial V_x}{\partial y} - \frac{\partial V_y}{\partial x} = 0.$$

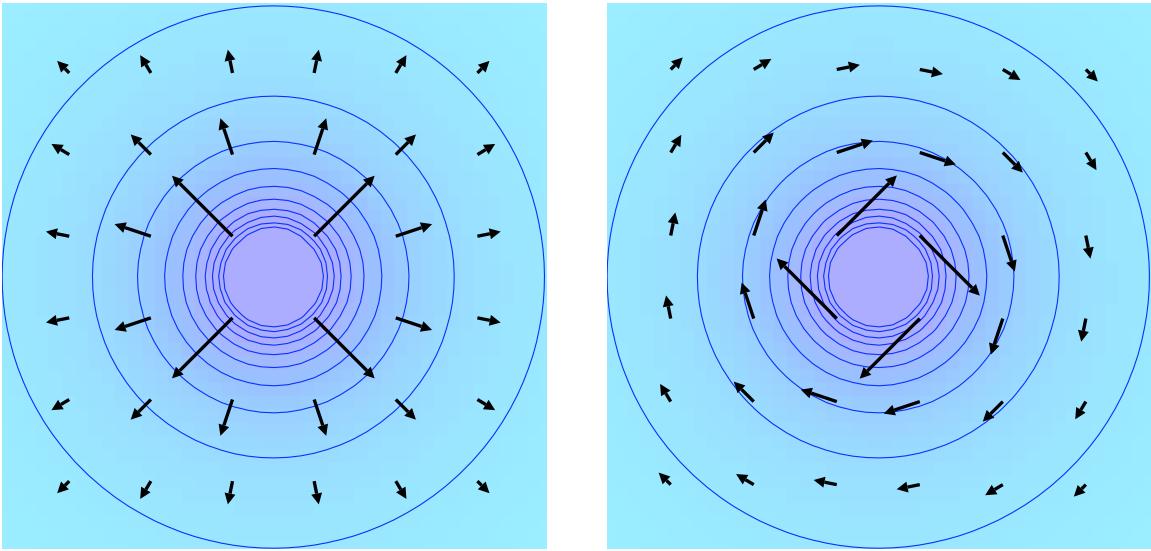


Figura 1.17: Campi di velocità corrispondenti alla funzioni $1/z$ (sinistra, con residuo reale) e i/z (destra, con residuo immaginario puro).

Se poniamo:

$$u(x, y) = V_x(x, y), \quad v(x, y) = -V_y(x, y),$$

è immediato riconoscere che le funzioni u e v soddisfano le condizioni di Cauchy-Riemann e sono identificabili con la parte reale e la parte immaginaria di una funzione complessa f , della variabile complessa $z = x + iy$, olomorfa nella regione in cui il moto del fluido è solenoidale e irrotazionale. Per descrivere il moto la funzione $f(z)$ e il campo di velocità \mathbf{V} sono perfettamente equivalenti.

Dove la funzione $f(z)$ risulta regolare non possono esistere sorgenti, pozzi, o vortici. Un punto che sia sorgente o vortice corrisponde ad una singolarità della funzione $f(z)$ (ad esempio un polo). Se un punto $z = z_0$ è una singolarità isolata allora questa può essere una sorgente o un vortice. L'intensità della sorgente o vortice è una informazione contenuta nel residuo della funzione in tale punto. Se γ è una curva chiusa che racchiude z_0 (e nessuna altra singolarità), allora il residuo R di $f(z)$ nel punto z_0 è dato dalla relazione:

$$\oint_{\gamma} f(z) dz = 2\pi i R.$$

Separando le parti reali ed immaginarie di tale relazione abbiamo:

$$\oint_{\gamma} (u dx - v dy) = -2\pi \operatorname{Im} R,$$

$$\oint_{\gamma} (u dy + v dx) = 2\pi \operatorname{Re} R,$$

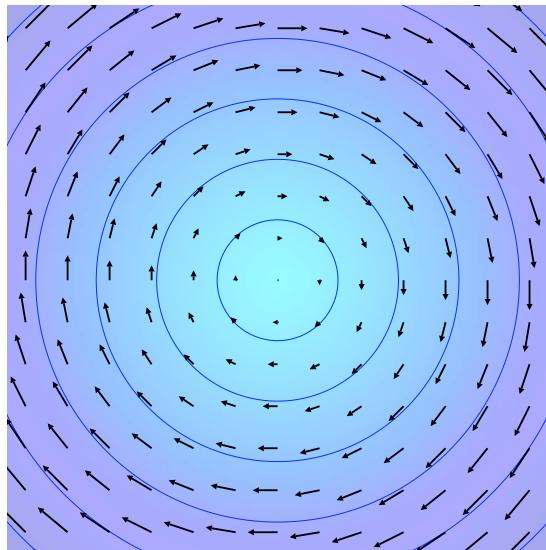


Figura 1.18: Campo di velocità corrispondente alla funzione non olomorfa $f(z) = i\bar{z}$. (vortice forzato, rigido).

e quindi, riesprimendo tali integrali al contorno tramite il corrispondente campo di velocità:

$$\oint_{\gamma} (V_x dx + V_y dy) = \oint_{\gamma} \mathbf{V} \cdot \tau ds = -2\pi \Im R,$$

$$\oint_{\gamma} (V_x dy - V_y dx) = \oint_{\gamma} \mathbf{V} \cdot \mathbf{n} ds = 2\pi \Re R,$$

con τ e \mathbf{n} i versori rispettivamente tangente e normale (rivolto verso l'esterno) alla curva γ . Tali relazioni mostrano che $\Im R$ è proporzionale all'intensità del vortice posto in $z = z_0$ (con il segno che individua il verso di rotazione), mentre la parte reale del residuo è proporzionale alla intensità di emissione della sorgente o di assorbimento del pozzo (a seconda del segno di $\Re R$) situato in $z = z_0$.

Nella figura 1.17 viene mostrato un esempio di moto fluido piano corrispondente alle funzioni $1/z$ e i/z . Nel primo caso abbiamo una sorgente di fluido nell'origine, mentre nel secondo caso abbiamo un esempio di vortice. Tale tipo di vortice viene detto *libero*, caratterizzato da una velocità tangenziale inversamente proporzionale alla distanza dal centro:

$$V_\theta = \frac{1}{r},$$

e rappresenta molto bene il vortice che si crea in un lavandino quando si toglie il tappo. La sua caratteristica è quella di avere una densità di momento angolare uniforme:

$$r V_\theta = \text{const.}$$

In contrapposizione abbiamo un altro tipo di vortice, detto **forzato** (figura 1.18), o rigido, che si ottiene ad esempio imprimendo ad un secchio pieno di fluido un moto rotatorio attorno all'asse. Questo movimento, in condizioni di equilibrio, viene trasferito al fluido con un flusso rotatorio caratterizzato da una velocità tangenziale proporzionale alla distanza dal centro di rotazione (tipica del moto rotatorio di un corpo rigido):

$$V_\theta = r .$$

In questo caso la corrispondente funzione complessa $f(z) = i\bar{z}$ risulta non analitica (ovunque).

1.4.1 Teorema dei residui

Supponiamo ora di volere integrare una funzione $f(z)$ lungo una curva chiusa γ che racchiuda, invece di una sola singolarità isolata, un certo numero m (**finito**) di singolarità, sempre isolate, di $f(z)$. Possiamo procedere nello stesso modo usato per derivare l'espansione di Laurent. Racchiudiamo ogni singolarità z_j ($j = 1, \dots, m$) all'interno di una circonferenza γ_j che contenga solo z_j (e nessuna altra singolarità) e colleghiamo ogni circonferenza con la curva γ mediante coppie di segmenti infinitesimalmente separati che facciamo tendere a coincidere. In pratica deformiamo la curva γ , trasformandola in una sequenza di circonferenze γ_j , senza “attraversare” punti singolari di $f(z)$. Fissando il verso di integrazione su ogni circonferenza γ_j concordemente con la curva γ arriviamo al risultato:

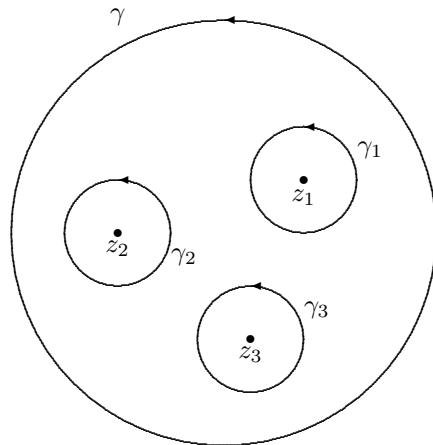


Figura 1.19: Teorema dei residui.

$$\oint_{\gamma} f(z) dz = \sum_{j=1}^m \oint_{\gamma_j} f(z) dz .$$

Dalla definizione di residuo abbiamo quindi

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{j=1}^m \operatorname{Res} f(z) \Big|_{z=z_j} \quad (1.97)$$

Tale relazione esprime l'importante **teorema dei residui**. Malgrado l'apparente banalità di tale teorema, la sua portata è in realtà straordinaria, come si avrà occasione di rilevare in seguito. È sul calcolo dei residui che si basa la possibilità di valutare un grande numero di integrali definiti che sarebbero assolutamente inattaccabili con i metodi elementari del calcolo integrale acquisiti nei corsi di analisi.

Osservazione. È importante che le singolarità siano tutte isolate per poterle racchiudere ognuna, da sola, all'interno di una circonferenza. Questo non è possibile se abbiamo un punto di accumulazione di singolarità in quanto una qualsiasi circonferenza centrata in tale punto conterrebbe inevitabilmente infinite singolarità.

1.4.2 Calcolo dei residui

Il problema di calcolare degli integrali di contorno per una funzione che abbia solo delle singolarità isolate si riduce quindi al calcolo dei residui di tale funzione. Vediamo quindi come eseguire il calcolo di tali residui, a seconda del tipo di singolarità (evitando di calcolare gli integrali che definiscono i residui stessi).

Polo di ordine n

Consideriamo dapprima il caso in cui $z = z_0$ sia un polo di ordine n per $f(z)$. In tal caso, in un intorno di z_0 , possiamo scrivere:

$$f(z) = \frac{g(z)}{(z - z_0)^n}, \quad (1.98)$$

con $g(z)$ funzione regolare e non nulla in z_0 ; allora:

$$\operatorname{Res} f(z) \Big|_{z=z_0} = \frac{1}{2\pi i} \oint_{\gamma} \frac{g(z)}{(z - z_0)^n} dz$$

con γ che racchiude solo il polo z_0 . Ricordando la rappresentazione integrale di Cauchy per le derivate, abbiamo:

$$\operatorname{Res} f(z) \Big|_{z=z_0} = \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} g(z) \Big|_{z=z_0},$$

$$\operatorname{Res} f(z) \Big|_{z=z_0} = \frac{1}{(n-1)!} \lim_{z \rightarrow z_0} \frac{d^{n-1}}{dz^{n-1}} [(z - z_0)^n f(z)] \quad (\text{polo di ordine } n). \quad (1.99)$$

L'equazione (1.99) fornisce un metodo per valutare i residui in singolarità di tipo polare, basato su semplici operazioni di derivazione e di limite.

Polo semplice

Nel caso, particolarmente importante, in cui il punto $z = z_0$ sia un polo semplice, per cui:

$$f(z) = \frac{g(z)}{z - z_0}, \quad (1.100)$$

la (1.99) diviene semplicemente:

$$\text{Res } f(z)|_{z=z_0} = \lim_{z \rightarrow z_0} (z - z_0)f(z). \quad (1.101)$$

Generalmente un polo semplice deriva dall'annullarsi di un denominatore, cioè:

$$f(z) = \frac{p(z)}{q(z)} \quad q(z_0) = 0, \quad (1.102)$$

con $p(z)$ regolare in z_0 , mentre $q(z)$ ha uno zero semplice in z_0 ; allora la (1.100) porta a:

$$\begin{aligned} \text{Res } f(z)|_{z=z_0} &= \lim_{z \rightarrow z_0} \frac{(z - z_0)p(z)}{q(z)} \\ &= \lim_{z \rightarrow z_0} \frac{(z - z_0)}{q(z) - q(z_0)} p(z) = \frac{p(z_0)}{q'(z_0)} \quad (\text{polo semplice}) \end{aligned} \quad (1.103)$$

che risulta una formula molto pratica.

Metodo della serie di Laurent

Se $f(z)$ ha una singolarità isolata in z_0 è olomorfa in una regione circolare (a corona circolare) attorno al punto z_0 ed è sviluppabile in serie di Laurent (la corona circolare degenera in un disco privato del solo centro). Esaminando l'espressione per i coefficienti a_n osserviamo in particolare che:

$$a_{-1} = \frac{1}{2\pi i} \oint_{\gamma} f(z) dz, \quad (1.104)$$

quindi il coefficiente a_{-1} dello sviluppo in serie di Laurent fornisce direttamente il residuo della funzione:

$$a_{-1} = \text{Res } f(z)|_{z=z_0}. \quad (1.105)$$

La precedente relazione può costituire una definizione alternativa per il residuo di una funzione $f(z)$ in un punto singolare isolato.

1.4.3 Residuo all'infinito

Abbiamo visto che anche all'infinito può essere studiato il comportamento di una funzione olomorfa ma il problema di un eventuale residuo all'infinito merita una trattazione a parte. Assumiamo che, data una funzione $f(z)$, il punto all'infinito, $z = \infty$, sia “isolato” rispetto alle singolarità al finito della funzione $f(z)$. In altre parole, il punto all'infinito non è un punto di accumulazione di singolarità, ed è quindi possibile costruire una curva γ_∞ , semplice, chiusa, che racchiude tutte le eventuali singolarità al finito della funzione $f(z)$. Tale curva (che possiamo assumere una circonferenza centrata nell'origine con raggio sufficientemente elevato) può essere interpretata come una curva attorno al punto $z = \infty$ e permette di definire un residuo anche per il punto all'infinito. La curva va intesa però percorsa in senso orario (in modo da avere il punto all'infinito alla sua sinistra, che risulta al suo “interno”). Avremo quindi, per definizione:

$$\text{Res } f(z)|_{z=\infty} = \frac{1}{2\pi i} \oint_{-\gamma_\infty} f(z) dz = -\frac{1}{2\pi i} \oint_{\gamma_\infty} f(z) dz, \quad (1.106)$$

dove, mantenendo per γ_∞ l'ordinario orientamento antiorario, abbiamo messo in evidenza con i segni, il verso di percorrenza della curva.

Il modo più semplice per valutare il precedente integrale (1.106) è quello di operare il cambio di variabile complessa:

$$z = \frac{1}{\zeta}$$

mediante cui il punto $z = \infty$ viene a corrispondere col punto $\zeta = 0$ e la curva γ_∞ si trasforma in una curva γ_0 che racchiude l'origine $\zeta = 0$, percorsa però in verso opposto (γ_∞ si trasforma in realtà in $-\gamma_0$), per cui, considerando che:

$$\frac{dz}{d\zeta} = -\frac{1}{\zeta^2}.$$

abbiamo (con un “triplo” segno negativo):

$$\text{Res } f(z)|_{z=\infty} = -\frac{1}{2\pi i} \oint_{\gamma_0} f\left(\frac{1}{\zeta}\right) \frac{d\zeta}{\zeta^2},$$

mantenendo sempre convenzionalmente il verso antiorario come orientamento anche per la curva γ_0 . Pertanto il calcolo è ricondotto alla valutazione di un residuo nell'origine della nuova variabile, $\zeta = 0$:

$$\text{Res } f(z)|_{z=\infty} = -\text{Res} \left\{ f\left(\frac{1}{\zeta}\right) \frac{1}{\zeta^2} \right\} \Big|_{\zeta=0}. \quad (1.107)$$

Osservazione. È opportuno a questo punto notare esplicitamente che esistono delle funzioni che, pur essendo regolari all'infinito, hanno tuttavia un residuo non nullo all'infinito.

L'esempio classico è fornito dalla semplice funzione:

$$f(z) = \frac{1}{z},$$

il cui unico punto singolare è il punto $z = 0$, un polo del primo ordine, con residuo unitario, in base alla (1.105). Il punto all'infinito è un punto regolare (posto $z = 1/\zeta$, il punto $\zeta = 0$ è un punto regolare di $f(1/\zeta) = \zeta$), ma si ha dalla (1.107):

$$\text{Res} \left. \frac{1}{z} \right|_{z=\infty} = -1.$$

In generale si può dire che, se nell'intorno del punto all'infinito una funzione $f(z)$ ammette uno sviluppo in serie di Laurent:

$$f(z) = \sum_{k=-\infty}^{\infty} a_k z^k = \cdots + \frac{a_{-3}}{z^3} + \frac{a_{-2}}{z^2} + \frac{a_{-1}}{z} + a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots$$

(assumiamo tale sviluppo valido in una corona circolare $r < |z| < \infty$, con r opportuno), allora:

$$f\left(\frac{1}{\zeta}\right) \frac{1}{\zeta^2} = \cdots + a_{-3} \zeta + a_{-2} + \frac{a_{-1}}{\zeta} + \frac{a_0}{\zeta^2} + \frac{a_1}{\zeta^3} + \cdots$$

e quindi, dalla (1.107):

$$\text{Res} f(z) \Big|_{z=\infty} = -a_{-1}.$$

Il residuo di una generica funzione $f(z)$ nel punto $z = \infty$ coincide quindi con l'opposto del coefficiente di $1/z$ nello sviluppo asintotico di $f(z)$ nell'intorno di $z = \infty$.

Tenendo conto dell'eventuale residuo all'infinito, possiamo affermare il seguente risultato.

Teo. 1.9 Se una funzione olomorfa $f(z)$ ha solo singolarità isolate (senza punti di accumulazione al finito e all'infinito), la somma di tutti i residui, tenuto conto anche dell'eventuale residuo non nullo all'infinito, è uguale a zero.

Dim. 1.9 La funzione $f(z)$ è olomorfa in tutto il piano complesso ad esclusione di un numero finito di singolarità isolate (deve risultare isolato anche il punto all'infinito, sia che esso risulti singolare oppure regolare). Consideriamo allora una generica curva semplice chiusa γ , orientata in senso positivo, che non passi per nessuna singolarità di $f(z)$. Una parte delle singolarità di $f(z)$ cadranno quindi all'interno della curva γ e una parte all'esterno (compreso il punto all'infinito). Per il teorema dei residui possiamo dire che:

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz = \sum_{\text{interni}} \text{Res} f(z), \quad (1.108)$$

ma possiamo pensare la curva come percorsa in senso negativo rispetto all'esterno, e che circonda la parte esterna del piano, incluso il punto all'infinito, per cui, tenendo conto del verso di percorrenza, includendo nel calcolo anche il punto all'infinito:

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz = - \sum_{\text{esterni}} \operatorname{Res} f(z). \quad (1.109)$$

Allora la somma di tutti i residui (compreso il residuo all'infinito), interni ed esterni alla curva si annulla:

$$\sum_{\text{totale}} \operatorname{Res} f(z) = 0. \quad (1.110)$$

— ■ —

1.4.4 Indicatore logaritmico

Consideriamo ora la funzione:

$$L(z) = \frac{f'(z)}{f(z)}, \quad (1.111)$$

con $f(z)$ olomorfa in un aperto D ($f'(z)$ indica la derivata di $f(z)$ per cui $L(z)$ è la derivata logaritmica di $f(z)$). Osserviamo che, se in un punto z_0 la funzione $f(z)$ ha uno zero di ordine m , in un suo intorno sarà esprimibile come:

$$f(z) = (z - z_0)^m g(z),$$

con $g(z)$ olomorfa, regolare e non nulla in $z = z_0$, e per la derivata abbiamo:

$$f'(z) = m(z - z_0)^{m-1} g(z) + (z - z_0)^m g'(z).$$

Pertanto la funzione $L(z)$ avrà, nell'intorno del punto $z = z_0$, un andamento del tipo:

$$L(z) = \frac{m}{z - z_0} + \frac{g'(z)}{g(z)},$$

cioè presenta un polo semplice con residuo m (il secondo termine è regolare per $z = z_0$).

Se la funzione $f(z)$ ha in un punto $z = z_0$ un polo di ordine n , in un suo intorno abbiamo:

$$f(z) = \frac{h(z)}{(z - z_0)^n},$$

con $h(z)$ olomorfa e non nulla in $z = z_0$. Allora:

$$f'(z) = -n \frac{h(z)}{(z - z_0)^{n+1}} + \frac{h'(z)}{(z - z_0)^n},$$

e, nell'intorno di $z = z_0$:

$$L(z) = -\frac{n}{(z - z_0)} + \frac{h'(z)}{h(z)},$$

presenta un polo semplice con residuo $-n$ (il secondo temine è regolare per $z = z_0$).

Se ora consideriamo l'integrale di $L(z)$ lungo una curva chiusa γ nel dominio di olo-morfismo D di $f(z)$, che racchiuda un numero finito di zeri e di poli di $f(z)$, si ottiene il cosiddetto **indicatore logaritmico** della funzione $f(z)$ rispetto alla curva γ :

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f'(z)}{f(z)} dz = (m_1 + m_2 + \dots + m_k) - (n_1 + n_2 + \dots + n_l), \quad (1.112)$$

dove supponiamo che all'interno della curva γ cadano k zeri di $f(z)$ con molteplicità m_1, m_2, \dots, m_k , e l poli di ordine n_1, n_2, \dots, n_l . Se conveniamo di contare uno zero di ordine m come m zeri ed un polo di ordine n come n poli possiamo dare la seguente proposizione.

Teo. 1.10 (indicatore logaritmico) *Data una curva chiusa γ , che non passi per nessuno zero e nessuna singolarità di una funzione $f(z)$, e tale che al suo interno siano situati un certo numero di zeri ed un certo numero di poli, ma nessuna singolarità essenziale, allora l'integrale della derivata logaritmica di $f(z)$ lungo la curva γ egualia la differenza tra il numero degli zeri e quello dei poli interni alla curva, moltiplicata per $2\pi i$ (vedi equazione (1.112)).*

Osservazione. Il ragionamento fatto per un punto al finito può essere ripetuto, con le dovute precauzioni, anche per il punto all'infinito, $z = \infty$. Se il punto all'infinito è uno “zero” di ordine m per $f(z)$, avremo asintoticamente:

$$f(z) = \frac{g(z)}{z^m},$$

$$L(z) = -\frac{m}{z} + \frac{g'(z)}{g(z)},$$

con $g(z)$ regolare, non nulla e finita per $z = \infty$, per cui $L(z)$, pur essendo regolare all'infinito, ha residuo m (si veda la (1.107)), molteplicità dello zero.

Analogamente se $z = \infty$ è un polo di ordine n allora asintoticamente:

$$f(z) = z^n h(z),$$

$$L(z) = \frac{n}{z} + \frac{h'(z)}{h(z)},$$

con $h(z)$ regolare, non nulla e finita per $z = \infty$, per cui $L(z)$ è regolare, ma presenta un residuo $-n$ all'infinito.

Una semplice conseguenza della proposizione sull'indicatore logaritmico si ottiene osservando che se una funzione $f(z)$ è razionale (quindi senza singolarità essenziali) anche la

funzione $L(z)$ è razionale e, includendo l'eventuale residuo non nullo all'infinito, abbiamo:

$$\sum_{\text{totale}} \operatorname{Res} L(z) = 0.$$

Pertanto, considerando l'intero piano complesso, incluso il punto all'infinito, una funzione razionale $f(z)$ ha tanti zeri quanti poli:

$$\sum_{\text{zeri } z_s} m_s = \sum_{\text{poli } z_p} n_p. \quad (1.113)$$

Se consideriamo in particolare un polinomio di grado n , che ha solamente un polo di ordine n in $z = \infty$, abbiamo che il numero degli zeri deve essere uguale a n (contando le eventuali molteplicità). Abbiamo riottenuto il teorema fondamentale dell'algebra (una equazione algebrica di grado n ha esattamente n radici, in generale complesse).

1.4.5 Calcolo di integrali definiti col metodo dei residui

Ci proponiamo ora di calcolare alcuni integrali definiti senza esplicitare una funzione primitiva della funzione integranda, ma deducendo il valore dell'integrale tramite una somma di residui relativi ai punti singolari di una funzione olomorfa scelta in maniera opportuna. Non esiste però un metodo generale che permetta di trattare il problema e ci limiteremo a considerare qualche classe di integrali, indicando, per ognuna di esse, quale procedimento permetta di ricondurre l'integrazione ad un calcolo di residui. Assumeremo anche in generale integrandi relativamente semplici in quanto siamo interessati a fornire una tecnica di calcolo. Ovviamente, caso per caso, occorre verificare le ipotesi dei teoremi che vengono usati per poter estendere le tecniche a casi meno semplici.

Funzioni di seno e coseno.

Consideriamo un integrale della forma:

$$I_1 = \int_0^{2\pi} R(\cos t, \sin t) dt,$$

(1.114)

con $R(x, y)$ funzione razionale (rapporto di due polinomi) priva di singolarità sul cerchio $x^2 + y^2 = 1$ (per non avere singolarità lungo il percorso di integrazione).

Poniamo:

$$z = e^{it}, \quad 0 \leq t < 2\pi,$$

quando t varia tra 0 e 2π il punto z descrive una circonferenza unitaria nel piano complesso. Abbiamo:

$$\begin{aligned}\cos t &= \frac{1}{2} (e^{it} + e^{-it}) = \frac{1}{2} \left(z + \frac{1}{z} \right) \\ \sin t &= \frac{1}{2i} (e^{it} - e^{-it}) = \frac{1}{2i} \left(z - \frac{1}{z} \right) \\ dz &= ie^{it} dt = iz dt \\ dt &= \frac{1}{iz} dz,\end{aligned}$$

Quindi I_1 può pensarsi come ottenuto dalla parametrizzazione di un integrale di contorno sul cerchio $|z| = 1$ di una opportuna funzione $\tilde{R}(z)$:

$$I_1 = \oint_{|z|=1} \frac{1}{iz} R \left(\frac{1}{2} \left(z + \frac{1}{z} \right), \frac{1}{2i} \left(z - \frac{1}{z} \right) \right) dz = \oint_{|z|=1} \tilde{R}(z) dz. \quad (1.115)$$

L'integrando $\tilde{R}(z)$, essendo R una funzione razionale, è ancora una funzione razionale in z , pertanto olomorfa, escluso al più un numero finito di poli, che possono essere interni od esterni alla circonferenza (ma non sulla circonferenza $|z| = 1$, in quanto ciò comporterebbe una singolarità di $R(x, y)$ per $x^2 + y^2 = 1$, e l'integrale non esisterebbe, contrariamente alle ipotesi fatte). Pertanto, per il teorema dei residui:

$$I_1 = 2\pi i \sum_{|z_p|<1} \text{Res } \tilde{R}(z) \Big|_{z=z_p}, \quad (1.116)$$

dove la somma è estesa a tutti i poli z_p contenuti all'interno della circonferenza di raggio unitario (si veda l'esercizio 1.10 come esempio).

Integrazione su tutto l'asse reale

Consideriamo integrali della forma:

$$I_2 = \int_{-\infty}^{+\infty} R(x) dx, \quad (1.117)$$

con $R(x)$ funzione razionale di x senza singolarità per x reale. Affinchè tale integrale risulti convergente si deve avere:

$$\lim_{|x| \rightarrow \infty} xR(x) = 0. \quad (1.118)$$

L'integrale I_2 si può pensare come il limite per $L \rightarrow \infty$:

$$I_2 = \lim_{L \rightarrow \infty} \int_{-L}^{+L} R(x) dx = \lim_{L \rightarrow \infty} I_2(L). \quad (1.119)$$

Consideriamo allora la complessificazione di $R(x)$ sostituendo l'argomento reale x con una variabile complessa $z = x + iy$; $R(z)$ risulta una funzione razionale (rapporto di due polinomi) e presenterà un numero finito di singolarità di tipo polare in punti z_k al di fuori dell'asse reale. La condizione di convergenza nulla all'infinito si estende ovviamente su tutto il piano complesso:

$$\lim_{|z| \rightarrow \infty} z R(z) = 0. \quad (1.120)$$

Sia Γ_L il cammino chiuso nel piano complesso come in figura 1.20, ottenuto aggiungendo al segmento lungo l'asse reale (tra $-L$ e L) una semicirconferenza C_L di raggio L nel semipiano superiore. Se L è sufficientemente grande la curva racchiude tutti i poli (sono

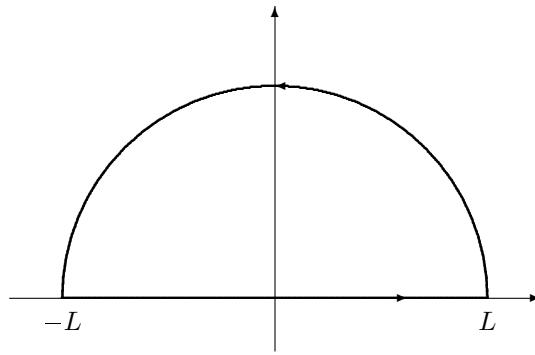


Figura 1.20: Integrale I_2 .

in numero finito) del semipiano superiore $\Im z > 0$. Pertanto:

$$\oint_{\Gamma_L} R(z) dz = 2\pi i \sum_{\Im z_k > 0} \operatorname{Res} R(z) \Big|_{z=z_k}.$$

D'altra parte:

$$\oint_{\Gamma_L} R(z) dz = I_2(L) + \int_{C_L} R(z) dz$$

Come vedremo fra poco la condizione $z R(z) \xrightarrow[|z| \rightarrow \infty]{} 0$ ci garantisce che:

$$\lim_{L \rightarrow \infty} \int_{C_L} R(z) dz = 0, \quad (1.121)$$

per cui otteniamo:

$$I_2 = \lim_{L \rightarrow \infty} \oint_{\Gamma_L} R(z) dz = 2\pi i \sum_{\Im z_k > 0} \operatorname{Res} R(z) \Big|_{z=z_k}. \quad (1.122)$$

Si veda l'esercizio 1.11 come esempio.

Vediamo ora come garantire il risultato (1.121).

Lem. 1.11 Sia $f(z)$ una funzione definita in un settore circolare:

$$z = |z| e^{i\theta}, \quad \theta_1 \leq \theta \leq \theta_2,$$

e sia C_r l'arco di circonferenza di raggio r in tale settore (figura 1.21):

$$C_r = \{ z = r e^{i\theta}, \theta_1 \leq \theta \leq \theta_2 \}.$$

Allora, se $zf(z)$ tende a zero in tale settore per $|z| \rightarrow \infty$, uniformemente rispetto all'angolo:

$$\lim_{r \rightarrow \infty} \sup_{z \in C_r} |zf(z)| = 0, \quad (1.123)$$

l'integrale di $f(z)$ sull'arco di circonferenza C_r di raggio r si annulla per $r \rightarrow \infty$:

$$\int_{C_r} f(z) dz \xrightarrow[r \rightarrow \infty]{} 0. \quad (1.124)$$

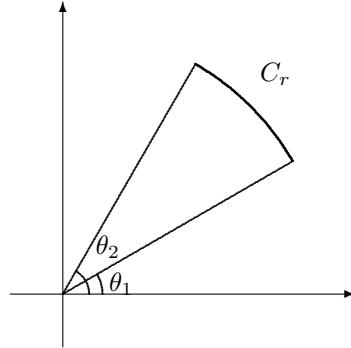


Figura 1.21: Lemma del grande cerchio.

Dim. 1.11 Posto:

$$M(r) = \sup_{z \in C_r} |f(z)|$$

l'estremo superiore di $|f(z)|$ su C_r , abbiamo, per il teorema di Darboux:

$$\left| \int_{C_r} f(z) dz \right| \leq r M(r) (\theta_2 - \theta_1),$$

ma per la (1.123) $rM(r) \xrightarrow[r \rightarrow \infty]{} 0$, da cui il risultato.

— ■ —

Osservazione. Notiamo che, se la (1.123) veniva sostituita con un analogo limite per $r \rightarrow 0$, allora (1.124) vale anche per $r \rightarrow 0$. Questo significa che $f(z)$ può anche divergere nell'origine, ma non troppo rapidamente.

Lem. 1.12 *Sia $f(z)$ una funzione definita in un settore circolare:*

$$z = |z| e^{i\theta}, \quad \theta_1 \leq \theta \leq \theta_2,$$

e sia C_r l'arco di circonferenza di raggio r in tale settore:

$$C_r = \{ z = r e^{i\theta}, \theta_1 \leq \theta \leq \theta_2 \}.$$

Allora, se $zf(z)$ tende a zero in tale settore per $|z| \rightarrow 0$, uniformemente rispetto all'angolo:

$$\lim_{r \rightarrow 0} \sup_{z \in C_r} |zf(z)| = 0, \quad (1.125)$$

l'integrale di $f(z)$ sull'arco di circonferenza C_r di raggio r si annulla per $r \rightarrow 0$:

$$\int_{C_r} f(z) dz \xrightarrow[r \rightarrow 0]{} 0. \quad (1.126)$$

I due lemmi precedenti sono a volte noti rispettivamente come **lemma del grande cerchio** e **lemma del piccolo cerchio**. Notiamo che non abbiamo richiesto che la funzione $f(z)$ sia olomorfa, per cui i lemmi sono validi in condizioni molto generali.

Lemma di Jordan

Ci proponiamo ora di studiare integrali della forma:

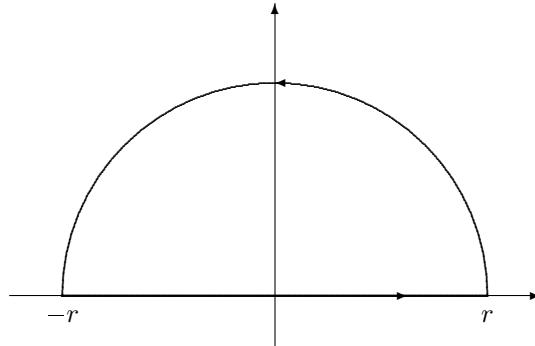
$$I_3 = \int_{-\infty}^{+\infty} f(x) e^{ix} dx,$$

(1.127)

dove supponiamo che l'estensione complessa $f(z)$ sia olomorfa nel semipiano $\Im z \geq 0$, escluso al più un numero finito di singolarità che supponiamo non essere situate sull'asse reale. In particolare vogliamo studiare l'integrale:

$$I_r = \int_{-r}^{+r} f(x) e^{ix} dx, \quad (1.128)$$

nel limite per $r \rightarrow \infty$. La convergenza di I_r non implica in generale la convergenza di I_3 , ma nel caso che I_3 sia convergente, allora $I_3 = \lim_{r \rightarrow \infty} I_r$. L'idea è quella di procedere come nel caso precedente considerando il percorso della figura 1.22, formato dal segmento reale tra $-r$ e r e dalla semicirconferenza C_r di raggio r e centro l'origine.

Figura 1.22: Integrale I_3 .

In tal caso:

$$\oint_{\Gamma_r} f(z)e^{iz} dz = \int_{-r}^{+r} f(x)e^{ix} dx + \int_{C_r} f(z)e^{iz} dz = I_r + \int_{C_r} f(z)e^{iz} dz \quad (1.129)$$

e l'integrale su Γ_r può essere valutato col teorema dei residui. Se possiamo dire che l'integrale su C_r tende ad un valore (possibilmente nullo), allora abbiamo una ricetta per studiare la convergenza dell'integrale I_r .

Ovviamente se la funzione $f(z)$ verifica il lemma precedente (1.123) nel semipiano superiore il gioco è fatto, in quanto il fattore e^{iz} non disturba:

$$|e^{iz}| = |e^{i(x+iy)}| = e^{-y} \leq 1 \quad (y \geq 0),$$

ma proprio la presenza di tale fattore (che si annulla per $\Im z = y \rightarrow +\infty$) potrebbe permettere una condizione meno restrittiva su $f(z)$.

Consideriamo l'integrale di $f(z)$ su un arco di circonferenza di raggio r , centro nell'origine tra due angoli θ_1 e θ_2 (vedi figura 1.21). Posto $M(r)$ l'estremo superiore di $|f(re^{i\theta})|$ su tale arco:

$$M(r) = \sup_{z \in C_r} |f(z)|,$$

abbiamo (supponiamo $0 \leq \theta_1 \leq \theta_2 \leq \pi$, nel semipiano superiore):

$$\begin{aligned} \left| \int_{C_r} f(z)e^{iz} dz \right| &= \left| \int_{\theta_1}^{\theta_2} f(re^{i\theta})e^{ire^{i\theta}} ire^{i\theta} d\theta \right| \\ &\leq \int_{\theta_1}^{\theta_2} |f(re^{i\theta})| re^{-r \sin \theta} d\theta \leq M(r) \int_{\theta_1}^{\theta_2} re^{-r \sin \theta} d\theta \\ &\leq M(r) \int_0^\pi re^{-r \sin \theta} d\theta = 2M(r) \int_0^{\frac{\pi}{2}} re^{-r \sin \theta} d\theta. \end{aligned}$$

Nell'ultima uguaglianza sfruttiamo la simmetria della funzione $\sin \theta$ rispetto al centro dell'intervallo $\theta = \pi/2$:

$$\sin(\pi - \theta) = \sin \theta.$$

Ora, se $0 \leq \theta \leq \frac{\pi}{2}$ (vedi figura 1.23):

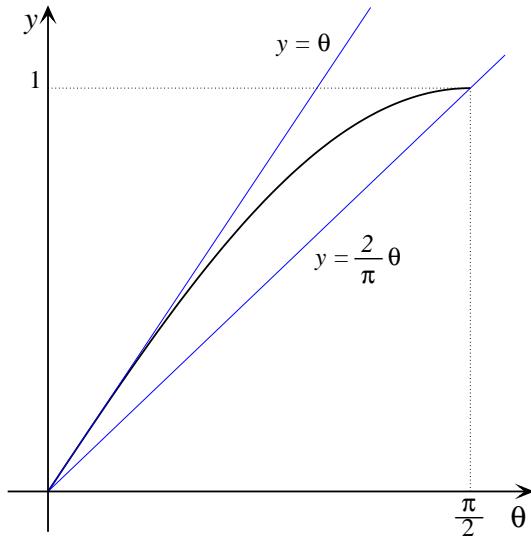


Figura 1.23: La funzione $y = \sin \theta$ è compresa tra la corda $y = \frac{2}{\pi} \theta$ e la tangente $y = \theta$.

$$\frac{2}{\pi} \leq \frac{\sin \theta}{\theta} \leq 1 \quad \Rightarrow \quad e^{-r \sin \theta} \leq e^{-\frac{2}{\pi} r \theta}$$

$$\int_0^{\frac{\pi}{2}} r e^{-r \sin \theta} d\theta \leq \int_0^{\frac{\pi}{2}} r e^{-\frac{2}{\pi} r \theta} d\theta = \frac{\pi}{2} \int_0^r e^{-x} dx = \frac{\pi}{2} (1 - e^{-r}) \leq \frac{\pi}{2}.$$

Quindi abbiamo:

$$\left| \int_{C_r} f(z) e^{iz} dz \right| \leq \pi M(r),$$

ed è quindi sufficiente che $M(r)$ tenda a zero.

Quindi, verificato che per la semicirconferenza C_r nel semipiano superiore l'estremo $M(r)$ tende a zero per $r \rightarrow \infty$, possiamo risolvere l'integrale I_3 col teorema dei residui:

$$I_3 = \lim_{r \rightarrow \infty} \oint_{\Gamma_r} f(z) e^{iz} dz = 2\pi i \sum_{\Im z_k > 0} \operatorname{Res} f(z) e^{iz} \Big|_{z=z_k}. \quad (1.130)$$

In pratica abbiamo dimostrato il seguente **lemma di Jordan**.

Lem. 1.13 (Jordan) Se $f(z)$ è una funzione definita nel semipiano superiore $\Im z > 0$

e se:

$$\lim_{|z| \rightarrow \infty} f(z) = 0, \quad (1.131)$$

uniformemente rispetto all'angolo in un settore $0 \leq \theta_1 \leq \arg z \leq \theta_2 \leq \pi$, detto C_r l'arco di circonferenza in tale settore:

$$C_r = \{ z = r e^{i\theta}, \theta_1 \leq \theta \leq \theta_2 \}.$$

si ha che:

$$\int_{C_r} f(z) e^{iz} dz \xrightarrow[r \rightarrow \infty]{} 0. \quad (1.132)$$

Analogamente si può mostrare che se:

$$\lim_{|z| \rightarrow 0} f(z) = 0, \quad (1.133)$$

allora:

$$\int_{C_r} f(z) e^{iz} dz \xrightarrow[r \rightarrow 0]{} 0. \quad (1.134)$$

Osservazione. Se, invece di avere:

$$\int_{-\infty}^{+\infty} f(x) e^{ix} dx,$$

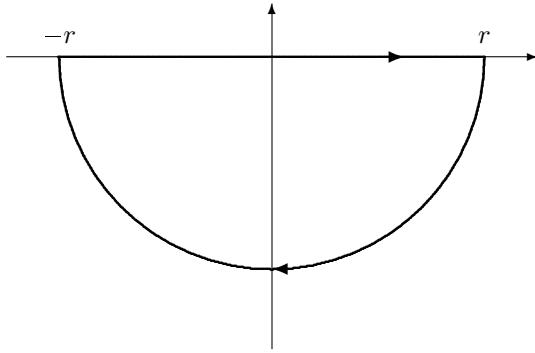
avessimo avuto:

$$I_3^* = \int_{-\infty}^{+\infty} f(x) e^{-ix} dx, \quad (1.135)$$

cioé un coefficiente negativo all'esponente, il fattore e^{-iz} con z complesso è divergente nel semipiano superiore mentre al contrario tende a zero nel semipiano inferiore ($\Im z < 0$) per $|z| \rightarrow \infty$. In tal caso occorre "chiudere" con un semicerchio nel semipiano inferiore ottenendo una curva Γ_r orientata in senso negativo (figura 1.24). Questo comporta un cambiamento di segno e il calcolo dei residui nel semipiano inferiore. Si può mostrare, con ragionamenti analoghi, che il lemma di Jordan vale ancora anche nel semipiano inferiore (col cambiamento di segno nell'esponente). In generale, in presenza di un fattore $e^{\alpha z}$ con α complesso, occorre considerare il semipiano per cui $|e^{\alpha z}| \leq 1$.

1.4.6 Poli semplici in prossimità del cammino di integrazione.

Può capitare a volte che la funzione integranda $f(z)$ possieda un punto singolare lungo il cammino di integrazione ed occorra "evitare" la singolarità con un percorso che passi "vicino". Nel caso di poli semplici esiste una formula generale per dedurre il contributo all'integrazione.

Figura 1.24: Integrale I_3^* .

Supponiamo, per semplicità, che $f(x)$ abbia un polo semplice in un punto reale $x = x_0$ e consideriamo l'integrale di $f(z)$ su un percorso semicircolare (orientato positivamente) come in figura 1.25, di raggio ϵ attorno a x_0 e vediamo di valutare:

$$\int_{\gamma_\epsilon} f(z) dz \quad \text{per } \epsilon \rightarrow 0.$$

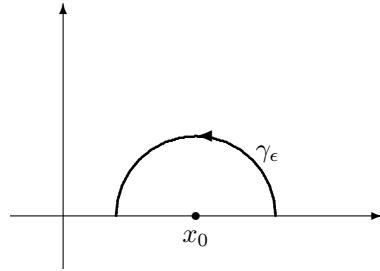


Figura 1.25: Contorno di un polo semplice.

Essendo x_0 un polo semplice possiamo scrivere (in un intorno di x_0 , e non è restrittivo supporre che tale intorno contenga la semicirconferenza γ_ϵ):

$$f(z) = \frac{a_{-1}}{z - x_0} + g(z), \quad (1.136)$$

con $g(z)$ regolare in x_0 ,

$$\lim_{z \rightarrow x_0} (z - x_0)g(z) = 0. \quad (1.137)$$

In questo modo possiamo dire che (lemma del piccolo cerchio):

$$\lim_{\epsilon \rightarrow 0} \int_{\gamma_\epsilon} g(z) dz = 0, \quad (1.138)$$

mentre, considerando su γ_ϵ la parametrizzazione $z = x_0 + \epsilon e^{i\theta}$ ($0 \leq \theta \leq \pi$):

$$\int_{\gamma_\epsilon} \frac{a_{-1}}{z - x_0} dz = a_{-1} \int_0^\pi i d\theta = \pi i a_{-1}.$$

Quindi ricordando il significato di a_{-1} :

$$\lim_{\epsilon \rightarrow 0} \int_{\gamma_\epsilon} f(z) dz = \pi i \operatorname{Res} f(z) \Big|_{z=x_0} \quad (\text{polo semplice}). \quad (1.139)$$

Osservazione. Notiamo che il ragionamento fatto vale anche per un arco di circonferenza diverso da una semicirconferenza, ma che sottende un angolo diverso da π . In tal caso il fattore π è sostituito dall'ampiezza dell'angolo sotteso.

Consideriamo un caso importante che si presenta a volte nello studio di fenomeni fisici, precisamente un integrale del tipo:

$$\int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx, \quad (1.140)$$

dove x_0 è un punto reale e supponiamo che $f(x)$ sia regolare per ogni x reale e che tenda a zero per $x \rightarrow \pm\infty$ in modo da non avere problemi di divergenza dell'integrale per $x \rightarrow \pm\infty$.

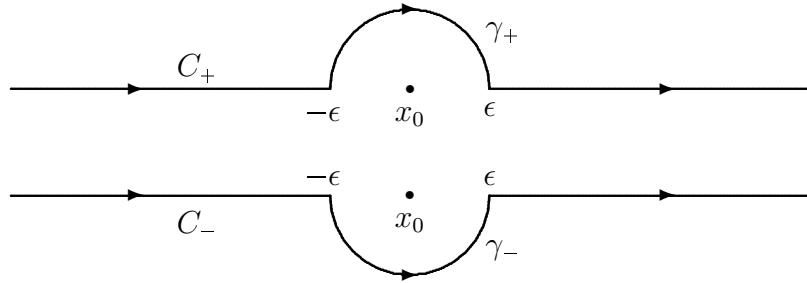


Figura 1.26: Singolarità sul cammino di integrazione.

Notiamo subito che se $f(x_0) \neq 0$ allora l'integrale (1.140) in senso stretto non esiste (l'integrando non è integrabile in un intorno del punto x_0). Si può cercare però, sfruttando le considerazioni precedenti ed assumendo che la complessificazione $f(z)$ risulti una funzione olomorfa in tutto un intorno dell'asse reale (il punto x_0 costituisce allora un polo semplice dell'integrando), di attribuire un significato a tale integrale considerando un percorso opportuno nel piano complesso che coincida sostanzialmente con l'asse reale, ma eviti la singolarità passandole “vicino a piacere”. Possiamo allora scegliere un cammino C_+ che passi sopra il punto singolare, circondandolo con una semicirconferenza γ_+ di raggio $\epsilon > 0$ (che faremo poi tendere a zero), oppure un cammino C_- che passi sotto il punto singolare (vedi figura 1.26). Abbiamo:

$$\begin{aligned} \int_{C_+} \frac{f(z)}{z - x_0} dz &= \int_{-\infty}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{+\infty} \frac{f(x)}{x - x_0} dx + \int_{\gamma_+} \frac{f(z)}{z - x_0} dz, \\ \int_{C_-} \frac{f(z)}{z - x_0} dz &= \int_{-\infty}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{+\infty} \frac{f(x)}{x - x_0} dx + \int_{\gamma_-} \frac{f(z)}{z - x_0} dz, \end{aligned}$$

Nel limite per $\epsilon \rightarrow 0$ i primi due integrali, comuni a secondo membro delle relazioni sopra, definiscono la **parte principale** (secondo Cauchy) dell'integrale (1.140):

$$\boxed{P \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx = \lim_{\epsilon \rightarrow 0} \left[\int_{-\infty}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{+\infty} \frac{f(x)}{x - x_0} dx \right].} \quad (1.141)$$

In base alle considerazioni precedenti, gli integrali sulle semicirconferenze γ_+ e γ_- , sempre nel limite per $\epsilon \rightarrow 0$, sono dati da:

$$\lim_{\epsilon \rightarrow 0} \int_{\gamma_{\pm}} \frac{f(z)}{z - x_0} dz = \mp i \pi f(x_0),$$

tenendo conto del verso di percorrenza delle semicirconferenze. Pertanto abbiamo:

$$\lim_{\epsilon \rightarrow 0} \int_{C_{\pm}} \frac{f(z)}{z - x_0} dz = P \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx \mp i \pi f(x_0). \quad (1.142)$$

Questa non è però la forma solita per esprimere tale relazione, ma possiamo operare una ulteriore trasformazione di percorso sulle curve C_{\pm} . Consideriamo l'integrale sulla curva C_+ . L'integrale (ϵ è “piccolo” ma finito) è da intendersi come un limite:

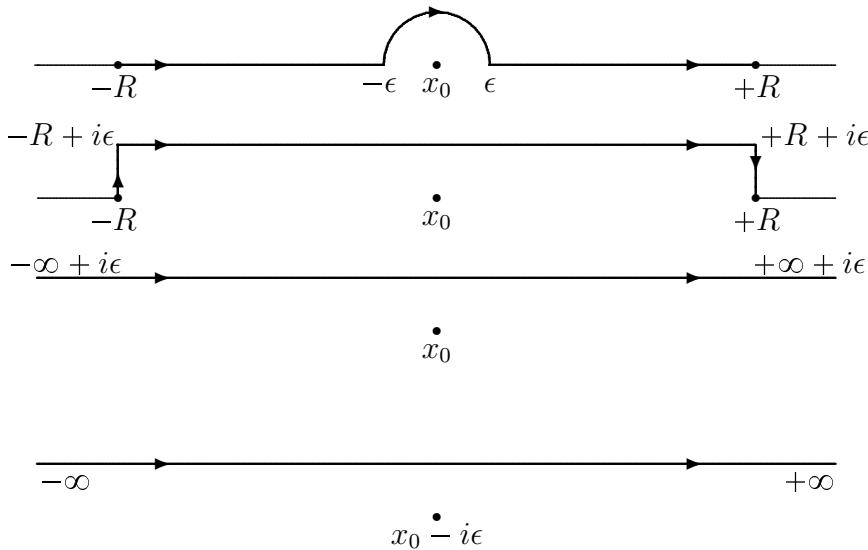


Figura 1.27: Deformazione del percorso C_+ .

$$\begin{aligned} \int_{C_+} \frac{f(z)}{z - x_0} dz &= \lim_{R \rightarrow \infty} \left[\int_{-R}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{+R} \frac{f(x)}{x - x_0} dx \right] + \int_{\gamma_+} \frac{f(z)}{z - x_0} dz \\ &= \lim_{R \rightarrow \infty} \left[\int_0^{\epsilon} \frac{f(-R + iy)}{-R - x_0 + iy} i dy + \int_{-R + i\epsilon}^{R + i\epsilon} \frac{f(z)}{z - x_0} dz + \int_{\epsilon}^0 \frac{f(R + iy)}{R - x_0 + iy} i dy \right], \end{aligned}$$

dove abbiamo trasformato la curva C_+ tra $-R$ e R in un percorso rettangolare a distanza ϵ dall'asse reale. Operando il limite per $R \rightarrow \infty$ gli integrali sui segmenti verticali (di ampiezza finita ϵ) si annullano (la funzione tende a zero in prossimità dell'asse reale per $x \rightarrow \pm\infty$) e possiamo dire:

$$\int_{C_+} \frac{f(z)}{z - x_0} dz = \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} \frac{f(z)}{z - x_0} dz = \int_{-\infty}^{+\infty} \frac{f(x + i\epsilon)}{x - x_0 + i\epsilon} dx.$$

Abbiamo supposto che la funzione sia olomorfa in tutto un intorno della retta reale, per cui risulta continua e, nel limite per $\epsilon \rightarrow 0$ possiamo sostituire $f(x + i\epsilon)$ con $f(x)$:

$$\int_{C_+} \frac{f(z)}{z - x_0} dz = \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0 + i\epsilon} dx.$$

Sostanzialmente siamo tornati ad un integrale sull'asse reale ma abbiamo “spostato” la singolarità “sotto” l'asse reale, nel punto $x_0 - i\epsilon$ (in ogni caso “fuori” dal cammino di integrazione). Un ragionamento analogo vale anche per il percorso C_- ed equivale semplicemente a scambiare ϵ con $-\epsilon$ (la singolarità viene spostata sopra l'asse reale in $x_0 + i\epsilon$). Possiamo allora riscrivere la relazione (1.142) nella maniera comunemente ricordata:

$$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0 \pm i\epsilon} dx = P \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx \mp i\pi f(x_0).$$

(1.143)

Osservazione. Poichè tale relazione è valida per una vasta classe di funzioni, si tende ad eliminare la menzione della funzione $f(x)$. Occorre però introdurre una funzione (impropria) $\delta(x - x_0)$, detta funzione delta di Dirac¹⁷, definita, in maniera puramente operativa, dalla proprietà:

$$\int_{-\infty}^{+\infty} \delta(x - x_0) f(x) dx = f(x_0), \quad (1.144)$$

¹⁷ **Paul Adrien Maurice Dirac** (Bristol, 8 agosto 1902 – Tallahassee, Florida, 20 ottobre 1984) è stato un fisico e matematico britannico, che viene annoverato tra i fondatori della fisica quantistica. Ricevette una educazione rigida e dura a causa delle tendenze autoritarie del padre, però evidenziò sin da piccolo una ottima predisposizione per la matematica. Nel 1926 sviluppò una formalizzazione della meccanica quantistica basata sull'algebra non commutativa di operatori. Nel 1933 ricevette il premio Nobel assieme a Schrödinger per “la scoperta di nuove forme della teoria atomica”. Seppur considerato uno scapolo predestinato, nel 1937 si sposò con Margit Wigner, sorella del fisico magiaro Eugene Wigner. Tra le sue passioni, vanno annoverati i viaggi e le passeggiate in montagna. Più di qualunque altro fisico contemporaneo, assegnò al concetto di “bellezza matematica” un ruolo preminente tra gli aspetti fondamentali insinuati della natura fino al punto di sostenere che “una teoria includente una bellezza matematica ha più probabilità di essere giusta e corretta di una sgradevole che venga confermata dai dati sperimentali”. In suo onore fu bandito il Premio Dirac.



valida per qualsiasi funzione $f(x)$ (con proprietà di regolarità “sufficientemente buone”). Allora possiamo riscrivere formalmente la relazione (1.143) in una forma molto utile da un punto di vista mnemonico:

$$\frac{1}{x - x_0 \pm i\epsilon} = P\left(\frac{1}{x - x_0}\right) \mp i\pi\delta(x - x_0). \quad (1.145)$$

Notiamo che in realtà non esiste alcuna funzione ordinaria che possa “agire” come la funzione δ . Essa rientra nella categoria delle **funzioni generalizzate o distribuzioni**, la cui trattazione esula per ora dai nostri scopi.

Parte principale

Vediamo ora di chiarire il significato della definizione (1.141) di parte principale di un integrale. Se la funzione è regolare (non ha singolarità) e tende a zero all’infinito, tale limite risulta ben definito. Notiamo che presi singolarmente, i due integrali che entrano nella definizione della parte principale sono in generale divergenti (logaritmicamente):

$$\begin{aligned} \int_{-\infty}^{x_0-\epsilon} \frac{f(x)}{x-x_0} dx &\sim f(x_0) \ln \epsilon \xrightarrow[\epsilon \rightarrow 0]{} \infty \\ \int_{x_0+\epsilon}^{+\infty} \frac{f(x)}{x-x_0} dx &\sim -f(x_0) \ln \epsilon \xrightarrow[\epsilon \rightarrow 0]{} \infty \end{aligned}$$

ma le due divergenze si compensano. Possiamo vedere ciò nella seguente maniera:

$$\begin{aligned} & \int_{-\infty}^{x_0-\epsilon} \frac{f(x)}{x-x_0} dx + \int_{x_0+\epsilon}^{+\infty} \frac{f(x)}{x-x_0} dx \\ &= \int_{-\infty}^{-\epsilon} \frac{f(x_0+x)}{x} dx + \int_{+\epsilon}^{+\infty} \frac{f(x_0+x)}{x} dx \\ &= \int_{\epsilon}^{\infty} \frac{f(x_0+x) - f(x_0-x)}{x} dx \end{aligned}$$

e quest’ultimo integrale non presenta in generale singolarità nell’origine in quanto:

$$\lim_{x \rightarrow 0} \frac{f(x_0+x) - f(x_0-x)}{x} = 2f'(x_0).$$

Potremmo in realtà definire la parte principale dell’integrale nella seguente maniera:

$$P \int_{-\infty}^{+\infty} \frac{f(x)}{x-x_0} dx = \int_0^{\infty} \frac{f(x_0+x) - f(x_0-x)}{x} dx \quad (1.146)$$

senza rendere esplicito un limite per $\epsilon \rightarrow 0$, ma si preferisce utilizzare la definizione (1.141) che si estende in maniera ovvia anche al caso di estremi di integrazione diversi da $\pm\infty$.

Relazioni di dispersione

Torniamo alla relazione (1.143), valida quando la funzione è regolare in tutto un intorno dell'asse reale (e tende a zero per $x \rightarrow \pm\infty$), e proviamo a richiedere condizioni più stringenti di regolarità. Supponiamo che $f(z)$ sia olomorfa anche in tutto il semipiano superiore ($\Im z > 0$) e che tenda a zero per $|z| \rightarrow \infty$, uniformemente rispetto alla fase di z . In pratica $f(z)$ soddisfa le condizioni del lemma di Jordan e $f(z)/(z - x_0)$ soddisfa le condizioni del lemma del grande cerchio. Nelle relazioni (1.142) possiamo valutare gli integrali sulle curve C_+ e C_- “richiudendole” con una semicirconferenza nel semipiano superiore (dove assumiamo $f(z)$ olomorfa ovunque):

$$\begin{aligned}\int_{C_+} \frac{f(z)}{z - x_0} dz &= \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0 + i\epsilon} dx = 0, \\ \int_{C_-} \frac{f(z)}{z - x_0} dz &= \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0 - i\epsilon} dx = 2\pi i f(x_0).\end{aligned}$$

Sostituendo nella (1.142), otteniamo l'equazione:

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = i\pi f(x_0), \quad (1.147)$$

che si può spezzare in una coppia di equazioni, dette **trasformazioni di Hilbert**¹⁸:

$$\begin{aligned}\Re f(x_0) &= \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\Im f(x)}{x - x_0} dx, \\ \Im f(x_0) &= -\frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\Re f(x)}{x - x_0} dx.\end{aligned} \quad (1.148)$$

Se invece supponiamo che $f(z)$ risulti olomorfa in tutto il semipiano inferiore ($\Im z < 0$), possiamo ancora valutare gli integrali lungo i percorsi C_+ e C_- , chiudendo il percorso con una semicirconferenza all'infinito nel semipiano inferiore (percorsa quindi in senso

¹⁸ **David Hilbert** (Königsberg, 23 gennaio 1862 – Gottinga, 14 febbraio 1943) è stato uno dei più eminenti ed influenti matematici del periodo a cavallo tra il XIX secolo e il XX secolo. Ha inventato e sviluppato una vasta classe di idee fondamentali, l'assiomatizzazione della geometria e, con la nozione degli spazi di Hilbert, ha posto le fondamenta dell'analisi funzionale. È conosciuto anche come il fondatore della teoria della prova, della logica matematica e della distinzione tra matematica e meta-matematica. Hilbert era un personaggio quantomeno singolare: era donnaiolo ed insofferente al conservatorismo della vita universitaria, alle regole e ai divieti sociali. Si racconta che durante gli anni venti, come tecnica di corteggiamento, al ristorante chiedesse alle signore più eleganti ed avvenenti di prestargli il loro boa piumato per ripararsi dagli spifferi.



negativo) e otteniamo:

$$\int_{C_+} \frac{f(z)}{z - x_0} dz = -2\pi i f(x_0),$$

$$\int_{C_-} \frac{f(z)}{z - x_0} dz = 0.$$

con un conseguente cambio di segno sostituendo nelle (1.142):

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = -i\pi f(x_0),$$

e nelle analoghe trasformazioni di Hilbert:

$$\Re f(x_0) = -\frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\Im f(x)}{x - x_0} dx,$$

$$\Im f(x_0) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\Re f(x)}{x - x_0} dx.$$

(1.149)

Le equazioni (1.148) e (1.149) sono esempi delle cosiddette **relazioni di dispersione**, che collegano la parte reale e la parte immaginaria di una grandezza fisica. Il nome deriva dal fatto che sono state introdotte la prima volta per collegare la dipendenza dalla frequenza della parte reale dell'indice di rifrazione (indice di rifrazione effettivo) a quella immaginaria (coefficiente di assorbimento) nello studio della dispersione della luce in un materiale.

Osservazione. Notiamo che, per avere la validità delle relazioni di dispersione, occorre assicurarsi che la funzione $f(z)$ sia olomorfa in tutto un semipiano (inferiore o superiore) e che tenda a zero all'infinito in tale semipiano. Fortunatamente, per le quantità di interesse fisico, questa ipotesi è spesso garantita dal principio di **causalità**, la proprietà intuitiva che la risposta ad una sollecitazione operata su un sistema sia successiva a questa (l'effetto avviene dopo la causa, oppure la causa precede l'effetto).

1.5 Cenni sulle funzioni polidrome.

La teoria delle funzioni analitiche sviluppata fino ad ora era basata sull'assunzione che le funzioni considerate, e le loro derivate fossero ad un sol valore. A prima vista sembra che le funzioni a più valori debbano essere trattate in modo completamente differente. Fortunatamente la teoria delle funzioni analitiche può essere estesa in modo da includere una vasta classe di funzioni a più valori usando una ingegnosa costruzione geometrica nota come superficie di Riemann. Vedremo ciò con qualche esempio, mediante funzioni elementari.

1.5.1 Logaritmo.

Poniamoci il problema di voler definire il logaritmo naturale di un numero complesso z . Sappiamo che nel caso reale il logaritmo di un numero è definito come l'esponente che si deve dare al numero di Nepero¹⁹ e per ottenere il numero assegnato (è la funzione inversa dell'esponenziale):

$$y = \log x \iff e^y = x.$$

Poichè sappiamo valutare l'esponenziale nel caso complesso, possiamo impostare l'equazione:

$$e^w = z, \tag{1.150}$$

con z complesso, assegnato, e w complesso, incognito, e assumiamo che questa equazione definisca w come il logaritmo di z . Ponendo:

$$w = u + iv \quad (u, v \text{ reali}),$$

e z in forma polare:

$$\begin{aligned} z &= \rho e^{i\theta}, \quad \rho = |z|, \quad 0 \leq \theta < 2\pi, \\ e^{u+iv} &= \rho e^{i\theta}, \\ \begin{cases} e^u = \rho \\ e^{iv} = e^{i\theta} \end{cases} &\iff \begin{cases} u = \log \rho = \log |z|, \\ v = \theta + 2k\pi, \quad k = 0, \pm 1, \pm 2, \dots \end{cases} \end{aligned}$$

¹⁹ **John Napier** (1550, Merchiston Castle, Scozia – 4 Aprile 1617, Edimburgo, Scozia) noto come Giovanni Nepero o semplicemente Nepero, è stato un matematico, astronomo e fisico scozzese, celebre per l'introduzione del logaritmo naturale, dei bastoncini (o ossi) di Nepero – un ingegnoso abaco di calcolo – e anche per aver sostenuto l'uso delle frazioni decimali e del punto come separatore decimale. Non era un matematico di professione, bensì un ricco proprietario terriero scozzese di nobile famiglia che riusciva a condurre i suoi poderi con efficace razionalità. Della sua vita non si hanno molte notizie e in particolare non è chiaro dove abbia potuto ricevere una buona educazione umanistica e matematica; si può solo congetturare che abbia frequentato una università europea, forse quella di Parigi.



Pertanto vediamo che non esiste una soluzione unica del problema, e risulta definita una funzione a più valori in corrispondenza ad uno stesso punto z . Siamo cioè in presenza di una funzione **polidroma**. Possiamo scrivere:

$$\log z = \log |z| + i \arg z, \quad (1.151)$$

dove la polidromia è racchiusa nelle possibili scelte per $\arg z$. Ad ogni scelta corrisponde una **determinazione** del logaritmo. Se vogliamo mettere maggiormente in evidenza tutti i possibili valori del logaritmo scriveremo anche:

$$\log z = \log |z| + i \arg z + 2\pi k i, \quad k \in \mathbb{Z}, \quad (1.152)$$

(notiamo che la funzione $\arg z$ è essa stessa una funzione polidroma definita dalla relazione:

$$z = |z| e^{i \arg z},$$

analoga a quella che definisce il logaritmo). Scegliendo $0 \leq \arg z < 2\pi$ si ottiene la cosiddetta **determinazione principale** del logaritmo:

$$\log z = \log |z| + i\theta \quad 0 \leq \theta = \arg z < 2\pi.$$

Notiamo inoltre che, con le dovute precauzioni, continuano a valere le proprietà dei logaritmi:

$$\begin{aligned} \log z_1 z_2 &= \log |z_1 z_2| + i \arg(z_1 z_2) + 2\pi k i \\ &= \log |z_1| + \log |z_2| + i \arg z_1 + i \arg z_2 + 2\pi k' i \\ &= \log z_1 + \log z_2 + 2\pi n i, \end{aligned}$$

cioè, a meno di multipli di $2\pi i$ il logaritmo del prodotto è la somma dei logaritmi.

In genere si opera una scelta per la determinazione del logaritmo, ottenendo una funzione ad un sol valore, ma in ogni caso la peculiarità del logaritmo può essere visualizzata considerando un percorso semplice chiuso C attorno all'origine (figura 1.28). Supponiamo di scegliere un punto iniziale z_0 su tale curva, di fissare una determinazione del logaritmo, cioè una soluzione dell'equazione (1.150), con $z = z_0$, e seguiamo idealmente la soluzione facendo variare z con continuità lungo la curva (scegliendo in ogni punto la soluzione più vicina alla soluzione del punto precedente) fino a tornare al punto di partenza. La funzione $\log z$ così ottenuta varia con continuità lungo la curva, ma dopo il completamento di un giro, abbiamo incrementato il logaritmo di $2\pi i$:

$$(\log z_0)_{\text{finale}} = (\log z_0)_{\text{iniziale}} + 2\pi i,$$

cioè non siamo tornati allo stesso valore.

In generale un punto del piano complesso per il quale, dopo aver fatto un giro completo attorno ad esso, una data funzione $f(z)$ non ritorna al valore iniziale, è detto **punto di**

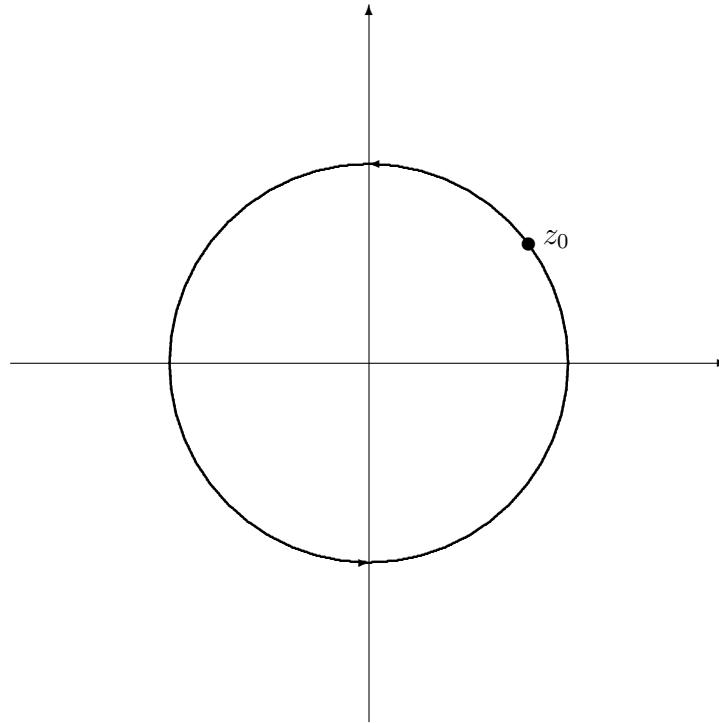


Figura 1.28: Percorso attorno all'origine.

diramazione (branch point) della funzione. Il punto risulta anche un punto singolare per la funzione.

Pertanto il punto $z = 0$ è un punto di diramazione per $f(z) = \log z$ (in effetti l'origine è un punto singolare, nemmeno nel caso reale si può definire $\log 0$). Considerando $f(1/z') = \log(1/z') = -\log z'$ troviamo analogamente che anche il punto all'infinito è un branch point. Si può verificare che non ci sono altri punti di diramazione per $\log z$.

Supponiamo ora di tracciare una curva che congiunge i due punti di diramazione $z = 0$ e $z = \infty$, cioè da $z = 0$ fino all'infinito, ed eliminiamo dal dominio della funzione tutti i punti di tale curva (generalmente si sceglie una semiretta uscente dall'origine), dicendo che abbiamo “tagliato” il piano complesso lungo un **taglio** di diramazione (**branch line**). Nel dominio così ottenuto definiamo una successione di funzioni continue ad un sol valore $f_n(z)$ che risolvono l'equazione (1.150). Ad esempio assumiamo un taglio lungo il semiasse reale negativo e definiamo le funzioni:

$$f_n(z) = f_n(r, \theta) = \log r + i(\theta + 2\pi n) \quad z = re^{i\theta} \quad -\pi < \theta < \pi.$$

In questo modo otteniamo funzioni analitiche. Questo può essere visto considerando in tale dominio la funzione analitica $\frac{1}{z}$ ed integrandola tra un punto iniziale (arbitrario) z_0 e un punto finale z :

$$g(z) = \int_{z_0}^z \frac{dz'}{z'}.$$

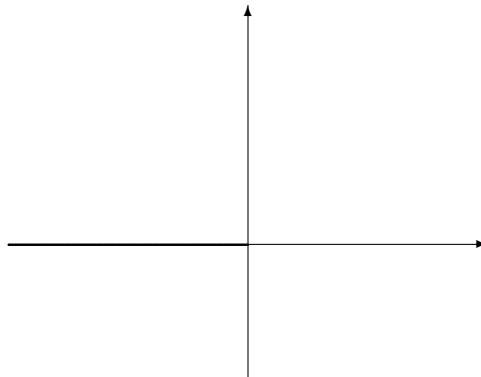


Figura 1.29: Taglio del logaritmo.

L'integrale, indipendente dal percorso (a cui impediamo l'attraversamento del semiasse reale negativo), può essere calcolato esplicitamente secondo il percorso di figura 1.30:

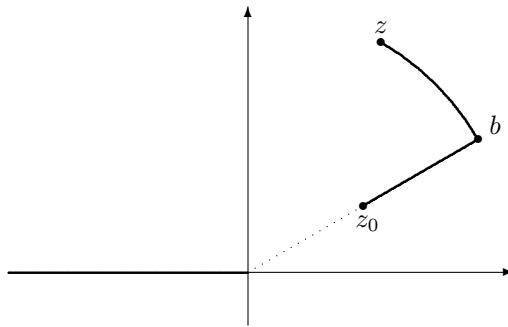


Figura 1.30: Il logaritmo.

$$\begin{aligned}
 g(z) &= \int_{z_0}^b \frac{dz'}{z'} + \int_b^z \frac{dz'}{z'} = \int_{r_0}^r dr' \frac{e^{i\theta_0}}{r' e^{i\theta_0}} + \int_{\theta_0}^{\theta} \frac{i r e^{i\theta'} d\theta'}{r e^{i\theta'}} \\
 &= \int_{r_0}^r \frac{dr'}{r'} + i \int_{\theta_0}^{\theta} d\theta' = \log r + i\theta - (\log r_0 + i\theta_0) \\
 &= f_n(z) - f_n(z_0),
 \end{aligned}$$

da cui, essendo $g(z)$ analitica (è una primitiva di $1/z$), lo è anche $f_n(z)$.

Ogni funzione $f_n(z)$ risulta discontinua lungo il taglio, nel senso che:

$$\lim_{\epsilon \rightarrow 0^+} [f_n(r, \pi - \epsilon) - f_n(r, -\pi + \epsilon)] = 2\pi i.$$

D'altra parte notiamo che:

$$\lim_{\epsilon \rightarrow 0^+} f_n(r, \pi - \epsilon) = \lim_{\epsilon \rightarrow 0^+} f_{n+1}(r, -\pi + \epsilon).$$

Questo suggerisce la seguente costruzione geometrica. Consideriamo il piano complesso come sovrapposizione di infiniti piani, ognuno dei quali corrispondente ad un valore di n ($n = 0, \pm 1, \pm 2, \dots$) e collegiamo piani adiacenti lungo il taglio, in modo che il bordo superiore del taglio sul piano n è collegato col bordo inferiore del taglio nel piano $n + 1$. Il punto di diramazione $z = 0$ è comune a tutti i piani. Pertanto attraversare il taglio equivale ad un passaggio da un piano all'altro. La superficie geometrica ottenuta da questa sovrapposizione ad elica è chiamata **superficie di Riemann** ed ogni piano è detto una **falda di Riemann** per la funzione $\log z$. In questo modo, mediante una sequenza di funzioni ad un sol valore definite su un singolo piano, otteniamo una funzione continua ad un sol valore definita sulla superficie di Riemann. L'unico punto in cui la funzione risulta non analitica è nei punti di diramazione.

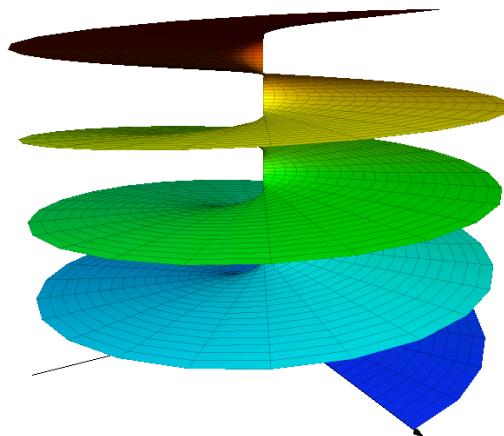


Figura 1.31: Superficie di Riemann del logaritmo.

Si può dare una classificazione dei punti di diramazione per una funzione. Un punto di diramazione è di ordine n se facendo $n + 1$ giri (ma non meno) completi attorno ad esso la funzione ritorna al valore originale. Altrimenti il punto è detto di ordine infinito (come nel caso del logaritmo). I punti di diramazione sono punti singolari per la funzione, ma di un carattere diverso da un polo o da una singolarità essenziale.

1.5.2 La radice quadrata.

Un altro esempio è fornito dalla radice di un numero complesso. Supponiamo di impostare l'equazione:

$$w^2 = z;$$

ponendo $z = |z|e^{i\theta}$, $w = |w|e^{i\varphi}$, dobbiamo avere:

$$\begin{aligned} |w|^2 e^{2i\varphi} &= |z|e^{i\theta}, \\ |w| &= \sqrt{|z|} \quad \varphi = \frac{\theta}{2} + k\pi. \end{aligned}$$

$\sqrt{z} = \sqrt{|z|} e^{i\frac{\theta}{2} + k\pi i} \quad k \in \mathbb{Z}$

(1.153)

Abbiamo quindi due valori distinti di w :

$$\begin{aligned} w &= \sqrt{|z|} e^{i\frac{\theta}{2}} \\ w &= \sqrt{|z|} e^{i\frac{\theta}{2} + i\pi} = -\sqrt{|z|} e^{i\frac{\theta}{2}} \end{aligned}$$

che corrispondono alle due determinazioni possibili della radice \sqrt{z} .

In questo caso $z = 0$ è un punto di diramazione (di ordine 1) ed è possibile scegliere una superficie di Riemann a due falde scegliendo un taglio (arbitrario tra $z = 0$ e $z = \infty$) a partire da $z = 0$. Questo può essere scelto sul semiasse reale positivo, definendo due funzioni di:

$$\begin{aligned} z &= \varrho e^{i\theta} \quad 0 < \theta < 2\pi, \\ f_1(z) &= \sqrt{\varrho} e^{i\frac{\theta}{2}} = f_1(\varrho, \theta) \\ f_2(z) &= -\sqrt{\varrho} e^{i\frac{\theta}{2}} = f_2(\varrho, \theta). \end{aligned}$$

Abbiamo:

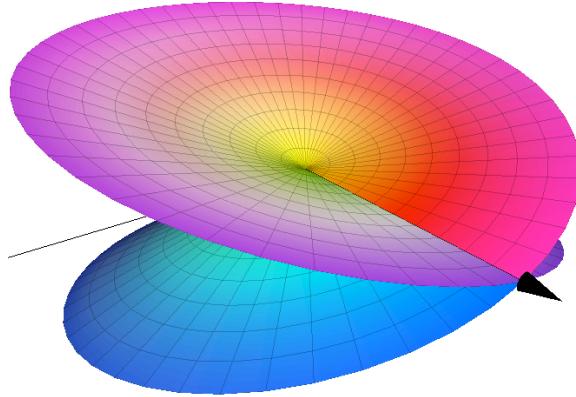


Figura 1.32: Superficie di Riemann della radice.

$$\begin{aligned} \lim_{\theta \rightarrow 0^+} f_1(\varrho, \theta) &= \lim_{\theta \rightarrow 2\pi^-} f_2(\varrho, \theta) \\ \lim_{\theta \rightarrow 2\pi^-} f_1(\varrho, \theta) &= \lim_{\theta \rightarrow 0^+} f_2(\varrho, \theta), \end{aligned}$$

e quindi scegliendo una superficie di Riemann come in figura otteniamo una funzione ad un sol valore definita per tutti i punti della superficie.

Con l'aiuto della superficie di Riemann otteniamo una descrizione unica di funzioni a molti valori. Il trucco consiste essenzialmente nel rimpiazzare una funzione a molti valori definita su un insieme semplice (il piano complesso originario) con una funzione ad un sol valore definita su una struttura geometrica complicata (la superficie di Riemann). In pratica, è spesso sufficiente focalizzare la propria attenzione su una falda particolare della superficie di Riemann, trattando tale parte della superficie come un piano complesso con un taglio. I valori della funzione corrispondono ad una particolare determinazione valida per la falda considerata. Da un punto di vista generale non c'è alcun motivo per preferire una falda rispetto ad un'altra, e il taglio può essere scelto in modo arbitrario (solo i punti terminali sono vincolati ad essere punti di diramazione). Nelle applicazioni fisiche si dà una preferenza a certi particolari tagli, per ragioni puramente fisiche, che, in generale, non hanno nulla a che fare con la matematica.

1.5.3 Calcolo di integrali che coinvolgono funzioni polidrome.

Consideriamo integrali del tipo:

$$I_4 = \int_0^\infty \frac{R(x)}{x^\alpha} dx \quad 0 < \alpha < 1, \quad (1.154)$$

e supponiamo che $R(x)$ sia una funzione razionale senza poli per $x \geq 0$ e che tenda a zero all'infinito in modo da avere garantita la convergenza dell'integrale.

$$\lim_{x \rightarrow \infty} R(x) = 0.$$

In tale integrale abbiamo la presenza del fattore x^α che comporta qualche problema nella sua generalizzazione nel campo complesso. z^α è una funzione a molti valori come la funzione \sqrt{z} discussa in precedenza.

Notiamo che per x reale abbiamo:

$$x^\alpha = e^{\alpha \log x},$$

e generalizzando al caso complesso:

$$z^\alpha = e^{\alpha \log z} = e^{\alpha \log |z| + i\alpha \theta + 2k\pi i\alpha} = |z|^\alpha e^{i\alpha(\theta + 2k\pi)},$$

i valori interi di k corrispondono a diverse determinazioni della potenza generica z^α (per $\alpha = \frac{1}{2}$ ritroviamo la definizione precedente di radice). Occorre quindi precisare quale determinazione si sceglie per z^α , e dove si opera il taglio.

Per calcolare il nostro integrale scegliamo il taglio proprio sul semiasse reale positivo su cui si integra (!). In questo modo z^α è ben definita per:

$$z = \varrho e^{i\theta}, \quad 0 < \theta < 2\pi$$

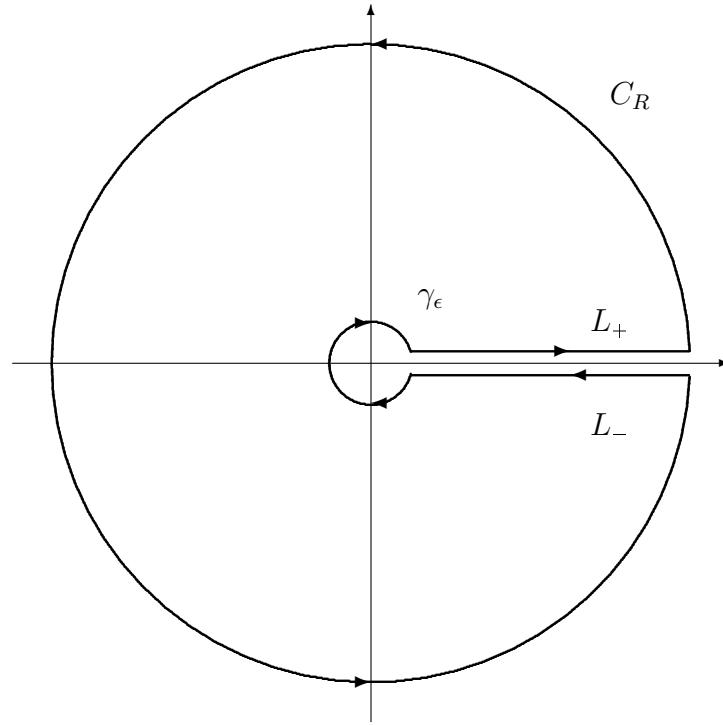


Figura 1.33: Integrazione polidroma.

(estremi esclusi) da

$$z^\alpha = |z|^\alpha e^{i\alpha\theta},$$

e consideriamo l'integrale sulla curva chiusa in figura 1.33 di $R(z)/z^\alpha$:

$$\int_{\Gamma(R,\epsilon)} \frac{R(z)}{z^\alpha} dz = 2\pi i \sum_{z \neq 0} \text{Res} \frac{R(z)}{z^\alpha}.$$

D'altra parte, essendo:

$$\lim_{|z| \rightarrow \infty} z \frac{R(z)}{z^\alpha} = 0 \quad \lim_{|z| \rightarrow 0} z \frac{R(z)}{z^\alpha} = 0,$$

($R(z)$ è razionale col denominatore di grado più alto del numeratore), gli integrali su C_R e γ_ϵ tendono a zero per $R \rightarrow \infty$ e $\epsilon \rightarrow 0$ per cui rimangono gli integrali su L_+ e L_- .

L_+ è un “pelo” sopra il taglio, per cui:

$$\frac{R(z)}{z^\alpha} = \frac{R(x)}{x^\alpha}, \quad z = xe^{i0^+}$$

mentre L_- è un “pelo” sotto:

$$\frac{R(z)}{z^\alpha} = \frac{R(x)}{x^\alpha} e^{-2\pi i\alpha}, \quad z = xe^{i2\pi^-},$$

abbiamo (in pratica sfruttiamo la discontinuità della funzione polidroma lungo il taglio):

$$\begin{aligned}
 2\pi i \sum_{z \neq 0} \operatorname{Res} \frac{R(z)}{z^\alpha} &= \left(\int_{L_+} + \int_{L_-} \right) \frac{R(z)}{z^\alpha} dz \\
 &= \int_0^\infty \frac{R(x)}{x^\alpha} dx - e^{-2\pi i \alpha} \int_0^\infty \frac{R(x)}{x^\alpha} dx \\
 &= (1 - e^{-2\pi i \alpha}) \int_0^\infty \frac{R(x)}{x^\alpha} dx \\
 I_4 &= \boxed{\frac{2\pi i}{1 - e^{-2\pi i \alpha}} \sum_{z \neq 0} \operatorname{Res} \frac{R(z)}{z^\alpha}}. \tag{1.155}
 \end{aligned}$$

Si veda l'esercizio 1.17 a titolo di esempio.

Vediamo ora integrali del tipo:

$$\boxed{I_5 = \int_0^\infty R(x) \log x dx}, \tag{1.156}$$

dove $R(x)$ è una funzione razionale senza poli per $x \geq 0$ e tale che:

$$\lim_{x \rightarrow \infty} xR(x) = 0,$$

(in modo da avere la convergenza dell'integrale).

Consideriamo lo stesso percorso $\Gamma(R, \epsilon)$ dell'integrale precedente e scegliamo la determinazione del logaritmo con l'argomento di z compreso tra 0 e 2π ($0 < \theta < 2\pi$), estremi esclusi.

$$\log z = \log |z| + i\theta, \quad 0 < \theta = \arg z < 2\pi,$$

e consideriamo l'integrale della funzione

$$R(z) \log^2(z)$$

lungo tale percorso. $\log^2 z$ non ha singolarità interne alla curva $\Gamma(R, \epsilon)$ e:

$$zR(z) \log^2 z \xrightarrow[|z| \rightarrow \infty]{} 0 \quad \text{uniformemente},$$

in quanto $R(z)$ è una funzione razionale con un denominatore di grado superiore di almeno 2 rispetto al numeratore. Inoltre, essendo $R(z)$ regolare nell'origine:

$$zR(z) \log^2 z \xrightarrow[z \rightarrow 0]{} 0.$$

Pertanto possiamo scrivere:

$$2\pi i \sum_{z \neq 0} \operatorname{Res} (R(z) \log^2 z) = \left(\int_{L_+} + \int_{L_-} \right) R(z) \log^2 z dz.$$

Ora, su L_+ , $z = xe^{i0^+}$,abbiamo:

$$\log z = \log x + i0^+ = \log x,$$

mentre su L_- :

$$\log z = \log x + i2\pi,$$

pertanto:

$$\begin{aligned} \left(\int_{L_+} + \int_{L_-} \right) R(z) \log^2(z) dz &= \int_0^\infty R(x) [\log^2 x - (\log x + 2\pi i)^2] dx \\ &= \int_0^\infty R(x) [-4\pi i \log x + 4\pi^2] dx \\ &= 4\pi^2 \int_0^\infty R(x) dx - 4\pi i \int_0^\infty R(x) \log x dx. \end{aligned}$$

Abbiamo quindi una formula risolutiva:

$$2\pi i \sum_{z \neq 0} \operatorname{Res}(R(z) \log^2 z) = 4\pi^2 \int_0^\infty R(x) dx - 4\pi i \int_0^\infty R(x) \log x dx. \quad (1.157)$$

che ci permette il calcolo dell'integrale richiesto. Con le ipotesi fatte i due integrali sopra sono reali:

$$\begin{aligned} \int_0^\infty R(x) \log x dx &= -\frac{1}{2} \Re \left\{ \sum_{z \neq 0} \operatorname{Res}(R(z) \log^2 z) \right\} \\ \int_0^\infty R(x) dx &= -\frac{1}{2\pi} \Im \left\{ \sum_{z \neq 0} \operatorname{Res}(R(z) \log^2 z) \right\}. \end{aligned}$$

Si veda per un esempio di applicazione l'esercizio 1.19.

Se invece di considerare la funzione $R(z) \log^2 z$ consideriamo $R(z) \log z$ avremmo avuto:

$$2\pi i \sum_{z \neq 0} \operatorname{Res}(R(z) \log z) = \int_0^\infty R(x)(\log x - \log x - 2\pi i) dx = -2\pi i \int_0^\infty R(x) dx$$

Per cui abbiamo un altro modo per eseguire un integrale (che non contiene un termine logaritmico):

$$\int_0^\infty R(x) dx = - \sum_{z \neq 0} \operatorname{Res}(R(z) \log z). \quad (1.158)$$

Il metodo precedente si può applicare, in certi casi, anche quando $R(x)$ presenta un polo semplice per $x = 1$. In questo caso:

$$\int_0^\infty R(x) \log x dx$$

esiste ancora in quanto $\log x$ si annulla (in maniera lineare) per $x = 1$ (zero semplice). In questo caso occorre però modificare il percorso di integrazione, in particolare il tratto L_- (infatti, a causa della determinazione scelta per il logaritmo, lungo L_- non abbiamo più uno zero semplice per $x = 1$ per cui $R(z) \log z$ risulterebbe singolare per $z = |z|e^{i2\pi} = 1$). Tale singolarità viene allora evitata con una semicirconferenza di raggio ϵ (vedi figura 1.34).

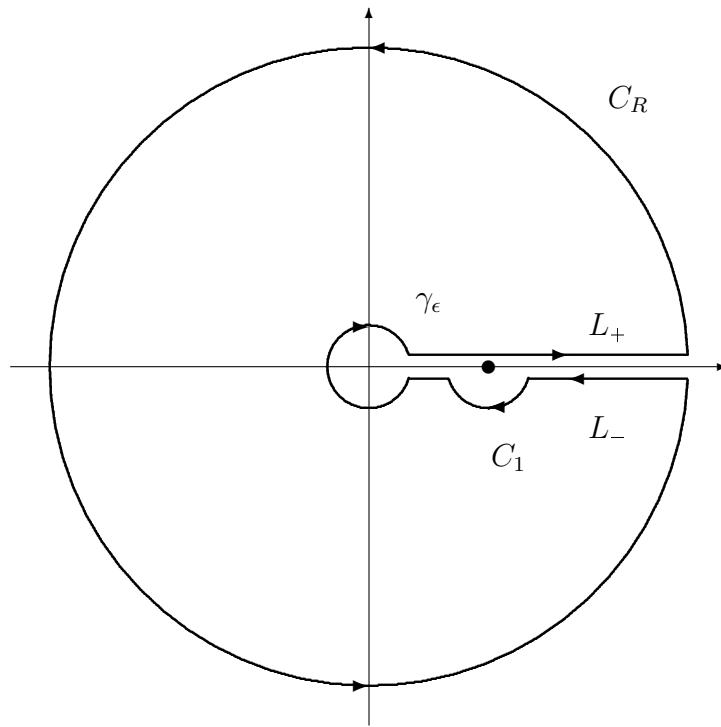


Figura 1.34: Integrazione logaritmo.

In tal caso abbiamo

$$\begin{aligned}
 & 2\pi i \sum_{z \neq 0,1} \operatorname{Res} (R(z) \log^2(z)) \int_0^\infty R(x)(\log x)^2 dx \\
 & - \left(\int_0^{1-\epsilon} + \int_{1+\epsilon}^\infty \right) R(x)(\log x + 2\pi i)^2 dx - \int_{C_1^+(\epsilon)} R(z) \log^2 z dz \\
 & = \left(\int_0^{1-\epsilon} + \int_{1+\epsilon}^\infty \right) R(x) [\log^2 x - (\log x + 2\pi i)^2] dx - \int_{C_1(\epsilon)} R(z) \log^2 z dz ,
 \end{aligned}$$

lungo $C_1(\epsilon)$ abbiamo

$$\begin{aligned} z &= 1 + \epsilon e^{i\theta} \quad \pi < \theta < 2\pi \\ \log z &= \log(1 + \epsilon e^{i\theta}) = 2\pi i + O(\epsilon) \\ \lim_{\epsilon \rightarrow 0} \int_{C_1(\epsilon)} R(z) \log^2 z dz &= -4\pi^2 \lim_{\epsilon \rightarrow 0} \int_{C_1(\epsilon)} R(z) dz = -4\pi^2 \pi i \operatorname{Res} R(z)|_{z=1}, \end{aligned}$$

da cui la formula

$$\begin{aligned} 2\pi i \sum_{z \neq 0,1} \operatorname{Res}(R(z) \log^2 z) &= 4\pi^2 P \int_0^\infty R(x) dx - 4\pi i \int_0^\infty R(x) \log x dx \\ &\quad + 4\pi^3 i \operatorname{Res}(R(z))_{z=1}. \end{aligned} \tag{1.159}$$

Dove:

$$P \int_0^\infty R(x) dx \equiv \lim_{\epsilon \rightarrow 0} \left\{ \int_0^{1-\epsilon} R(x) dx + \int_{1+\epsilon}^\infty R(x) dx \right\} \tag{1.160}$$

viene detta **parte principale** dell'integrale. Notiamo che, in presenza di una singolarità (di tipo polare) nel punto $x = 1$, l'integrale (ordinario) di $R(x)$ non esiste (cioè non converge), ma per un polo semplice il limite sopra esiste ed è finito (come esempio si consideri l'esercizio 1.20).

1.6 Esercizi

Esercizio 1.1 Esprimere i seguenti numeri complessi in forma polare esponenziale:

$$-1, \quad -i, \quad 1-i, \quad \frac{\sqrt{3}}{2} + \frac{1}{2}i.$$

— ○ —

Esercizio 1.2 Esprimere ciascuno dei seguenti numeri complessi nella forma cartesiana $a+ib$:

$$e^{3-i\frac{\pi}{2}}, \quad \frac{1}{1-i}, \quad (1-i)^3, \quad |4+3i|.$$

— ○ —

Esercizio 1.3 Trovare le radici complesse delle seguenti equazioni:

$$z^3 = -1, \quad z^4 = 1, \quad z^4 - 2z^2 + 2 = 0.$$

— ○ —

Esercizio 1.4 Verificare i seguenti risultati:

$$\overline{z+w} = \overline{z} + \overline{w}$$

$$|z| - |w| \leq |z-w|$$

$$\operatorname{Re} z \leq |z|$$

$$|\overline{w}z + w\overline{z}| \leq 2|wz|$$

— ○ —

Esercizio 1.5 Si consideri nel piano complesso il rettangolo:

$$R = \{z = x + iy; \quad a < x < b, \quad c < y < d\}$$

con a, b, c, d reali e si calcoli l'integrale sul bordo ∂R , orientato in senso antiorario, delle seguenti funzioni:

$$\begin{aligned} f_1(z) &= 1, \\ f_2(z) &= z, \\ f_3(z) &= z^2, \\ f_4(z) &= |z|^2. \end{aligned}$$

— ○ —

Esercizio 1.6 Mostrare la validità del teorema di Cauchy assumendo l'ipotesi di continuità della derivata $\frac{df}{dz}$ (suggerimento: si usi il teorema di Green).

— ○ —

Esercizio 1.7 Determinare lo sviluppo in serie di Laurent delle seguenti funzioni, relativamente ai punti indicati:

$$f_1(z) = \frac{e^{2z}}{(z-1)^3}; \quad z_0 = 1,$$

$$f_2(z) = (z-3) \sin\left(\frac{1}{z+2}\right); \quad z_0 = -2,$$

$$f_3(z) = \frac{z - \sin z}{z^3}; \quad z_0 = 0$$

— ○ —

Esercizio 1.8 Determinare lo sviluppo in serie di Laurent della funzione:

$$f(z) = \frac{z}{(z+1)(z+2)},$$

attorno al punto $z_0 = -2$.

— ○ —

Esercizio 1.9 Espandere:

$$f(z) = \frac{1}{(z+1)(z+3)}$$

in serie di Laurent valida nei seguenti casi:

- 1) $1 < |z| < 3$
- 2) $|z| > 3$
- 3) $0 < |z+1| < 2$
- 4) $|z| < 1$

— ○ —

Esercizio 1.10 Verificare che:

$$\int_0^{2\pi} \frac{dt}{a + \sin t} = \frac{2\pi}{\sqrt{a^2 - 1}} \quad a > 1$$

— ○ —

Esercizio 1.11 Mostrare che:

$$\int_0^\infty \frac{dx}{1+x^6} = \frac{\pi}{3}.$$

— ○ —

Esercizio 1.12 Verificare che:

$$I = \int_0^\infty \frac{\cos x}{x^2 + 1} dx = \frac{\pi}{2e}.$$

— ○ —

Esercizio 1.13 (Tema d'esame proposto il 30 giugno 2003) Si calcoli il seguente integrale:

$$F(\alpha) = \int_0^\infty \frac{\cos(\alpha x)}{x^2 + \alpha^2} dx, \quad \alpha \in \mathbb{R} \setminus \{0\}.$$

— ○ —

Esercizio 1.14 (Tema d'esame proposto il 4 settembre 2002) Calcolare il seguente integrale:

$$I_\alpha = \int_{-\infty}^\infty \frac{e^{i\alpha x}}{(1 - ix)(1 + 2ix)} dx, \quad \alpha \in \mathbb{R}.$$

— ○ —

Esercizio 1.15 Verificare che:

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2},$$

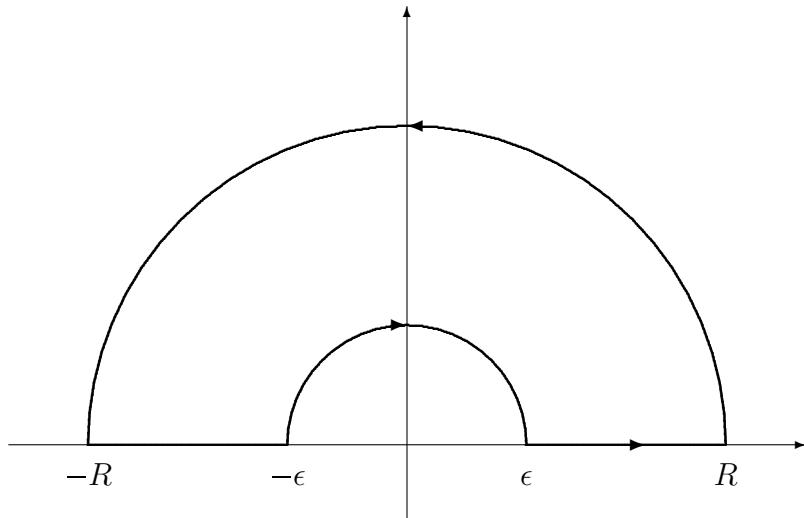
utilizzando il cammino in figura 1.35 e la funzione $\frac{e^{iz}}{z}$.

Figura 1.35: Percorso di integrazione per l'esercizio 1.15.

— ○ —

Esercizio 1.16 (Tema d'esame proposto il 13 gennaio 2005) Calcolare la parte principale del seguente integrale:

$$\begin{aligned} I &= P \int_{-\infty}^{\infty} \frac{\sin x}{(x-1)(x-i)} dx \\ &\equiv \lim_{\epsilon \rightarrow 0} \left\{ \int_{-\infty}^{1-\epsilon} \frac{\sin x}{(x-1)(x-i)} dx + \int_{1+\epsilon}^{\infty} \frac{\sin x}{(x-1)(x-i)} dx \right\}. \end{aligned}$$

— ○ —

Esercizio 1.17 Verificare:

$$I = \int_0^{\infty} \frac{dx}{x^{\alpha}(1+x)} = \frac{\pi}{\sin \pi \alpha} \quad 0 < \alpha < 1,$$

e, in particolare:

$$\int_0^{\infty} \frac{dx}{\sqrt{x}} \frac{1}{1+x} = \pi$$

— ○ —

Esercizio 1.18 (Tema d'esame proposto il 13 giugno 2003) Calcolare il seguente integrale:

$$I = \int_0^{\infty} \frac{dx}{(x+i)\sqrt{x}}.$$

— ○ —

Esercizio 1.19 Mostrare che:

$$\begin{aligned} \int_0^{\infty} \frac{\log x}{(1+x)^3} dx &= -\frac{1}{2} \\ \int_0^{\infty} \frac{1}{(1+x)^3} dx &= \frac{1}{2}. \end{aligned}$$

— ○ —

Esercizio 1.20 Si verifichi:

$$I = \int_0^{\infty} \frac{\log x}{x^2-1} dx = \frac{\pi^2}{4}.$$

Verificare anche il risultato (non ovvio):

$$P \int_0^{\infty} \frac{dx}{x^2-1} = 0.$$

— ○ —

1.6.1 Soluzioni

Soluzione 1.1 Abbiamo:

$$\begin{aligned} -1 &= \cos(\pi) + i \sin(\pi) = e^{i\pi}, \\ -i &= \cos(\pi/2) - i \sin(\pi/2) = e^{-i\frac{\pi}{2}} = e^{i\frac{3\pi}{2}}, \\ 1-i &= \sqrt{2} \frac{1-i}{\sqrt{2}} = \sqrt{2} e^{-i\frac{\pi}{4}} = \sqrt{2} e^{i\frac{7\pi}{4}}, \\ \frac{\sqrt{3}}{2} + \frac{1}{2}i &= \cos(\pi/6) + i \sin(\pi/6) = e^{i\frac{\pi}{6}}. \end{aligned}$$

— ○ —

Soluzione 1.2

$$\begin{aligned} e^{3-i\frac{\pi}{2}} &= e^3 [\cos(\pi/2) - i \sin(\pi/2)] = -i e^3, \\ \frac{1}{1-i} &= \frac{1+i}{(1-i)(i+i)} = \frac{1+i}{2} = \frac{1}{2} + i \frac{1}{2}, \\ (1-i)^3 &= 1 - 3i + 3i^2 - i^3 = -2 - 2i, \\ |4+3i| &= \sqrt{16+9} = 5. \end{aligned}$$

— ○ —

Soluzione 1.3 Consideriamo l'equazione $z^3 = -1$. Ponendo:

$$z = \rho e^{i\theta}, \quad 0 \leq \theta < 2\pi,$$

abbiamo:

$$\rho^3 e^{3i\theta} = -1 = e^{i\pi},$$

per cui si deve avere:

$$\begin{aligned} \rho^3 &= 1, \quad \rho = 1, \\ 3\theta &= \pi + 2k\pi, \quad k = 0, \pm 1, \pm 2, \dots \\ z &= e^{i\frac{\pi}{3}}, \quad e^{i\pi} = -1, \quad e^{i\frac{5\pi}{3}}, \end{aligned}$$

che costituiscono le tre soluzioni distinte dell'equazione.

Analogamente, per l'equazione $z^4 = 1$:

$$\begin{aligned} z &= \rho e^{i\theta}, \quad 0 \leq \theta < 2\pi, \\ \rho^4 e^{i4\theta} &= 1 = e^{i2k\pi}, \quad k = 0, \pm 1, \dots \\ \rho &= 1, \quad \theta = k\frac{\pi}{2}, \\ z &= \pm 1, \quad \pm i. \end{aligned}$$

Per l'ultima equazione, $z^4 - 2z^2 + 2 = 0$,abbiamo:

$$\begin{aligned} z^2 &= 1 \pm \sqrt{1-2} = 1 \pm i = \sqrt{2} e^{\pm i \frac{\pi}{4}}, \\ z &= 2^{1/4} e^{i \frac{\pi}{8}}, 2^{1/4} e^{-i \frac{\pi}{8}}, 2^{1/4} e^{i \frac{7\pi}{8}}, 2^{1/4} e^{-i \frac{7\pi}{8}}. \end{aligned}$$

Le soluzioni corrispondono ai vertici di un rettangolo.

— ○ —

Soluzione 1.4 Ponendo $z = x + iy$, $w = u + iv$:

$$\begin{aligned} \overline{z+w} &= \overline{(x+u) + i(y+v)} = x+u - i(y+v) = (x-iy) + (u-iv) = \overline{z} + \overline{w}, \\ |z| &= |z-w+w| \leq |z-w| + |w| \implies |z| - |w| \leq |z-w|, \\ |w| &= |w-z+z| \leq |z-w| + |z| \implies |w| - |z| \leq |z-w|, \\ \Re z = x &\leq |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = |z|, \\ |\overline{w}z + w\overline{z}| &= |(u-iv)(x+iy) + (u+iv)(x-iy)| = |2ux + 2vy| = 2|ux + vy|, \\ (ux + vy)^2 &= u^2x^2 + v^2y^2 + 2uvxy \leq u^2x^2 + v^2y^2 + u^2y^2 + v^2x^2 = (u^2 + v^2)(x^2 + y^2), \\ |ux + vy| &\leq \sqrt{u^2 + v^2} \sqrt{x^2 + y^2}. \end{aligned}$$

— ○ —

Soluzione 1.5 L'integrale di una funzione $f(z)$ sul bordo ∂R si compone di quattro integrazioni distinte sui lati del rettangolo (indichiamo i vertici con A, B, C, D):

$$\begin{aligned} \oint_{\partial R} f(z) dz &= \int_{AB} f(z) dz + \int_{BC} f(z) dz + \int_{CD} f(z) dz + \int_{CA} f(z) dz \\ &= \int_a^b f(x+ic) dx + \int_c^d f(b+iy) i dy \\ &\quad + \int_b^a f(x+id) dx + \int_d^c f(a+iy) i dy \\ &= \int_a^b [f(x+ic) - f(x+id)] dx \\ &\quad + i \int_c^d [f(b+iy) - f(a+iy)] dy \end{aligned}$$

Quindi:

$$\begin{aligned}
\oint_{\partial R} f_1(z) dz &= \int_a^b (1 - 1) dx + i \int_c^d (1 - 1) dy = 0 \\
\oint_{\partial R} f_2(z) dz &= \int_a^b (x + ic - x - id) dx + i \int_c^d (b + iy - a - iy) dy \\
&= -i(d - c)(b - a) + i(b - a)(d - c) = 0 \\
\oint_{\partial R} f_3(z) dz &= \int_a^b [(x + ic)^2 - (x + id)^2] dx \\
&\quad + i \int_c^d [(b + iy)^2 - (a + iy)^2] dy \\
&= \int_a^b [2i(c - d)x + d^2 - c^2] dx \\
&\quad + i \int_c^d [b^2 - a^2 + 2i(b - a)y] dy \\
&= i(c - d)(b^2 - a^2) + (b - a)(d^2 - c^2) \\
&\quad + i(b^2 - a^2)(d - c) - (b - a)(d^2 - c^2) \\
&= 0 \\
\oint_{\partial R} f_4(z) dz &= \int_a^b [(x^2 + c^2) - (x^2 + d^2)] dx \\
&\quad + i \int_c^d [(y^2 + b^2) - (y^2 + a^2)] dy \\
&= \int_a^b (c^2 - d^2) dx + i \int_c^d (b^2 - a^2) dy \\
&= (b - a)(c^2 - d^2) + i(b^2 - a^2)(d - c) \\
&= (b - a)(d - c)(-c - d + ia + ib)
\end{aligned}$$

Soluzione 1.6 Il teorema di Green²⁰ afferma sostanzialmente che se due funzioni (reali) $p(x, y)$ e $q(x, y)$, definite su una regione semplicemente connessa $D \subset \mathbb{R}^2$, sono continue e derivabili con derivate continue in tutto l'insieme D , compresa la frontiera ∂D , che assumiamo descritta da una curva semplice, chiusa e regolare a tratti, allora:

$$\oint_{\partial D} p(x, y) dx + q(x, y) dy = \iint_D \left[\frac{\partial q}{\partial x}(x, y) - \frac{\partial p}{\partial y}(x, y) \right] dx dy. \quad (1.161)$$

dove la curva ∂D che delimita D è orientata in senso antiorario.

Utilizzando tale risultato diventa molto semplice verificare il teorema di Cauchy. Separando in una funzione olomorfa $f(z)$ la parte reale u e la parte immaginaria v :

$$z = x + iy, \quad f(z) = u(x, y) + iv(x, y)$$

abbiamo che u e v sono continue con derivate continue per l'ipotesi fatta sulla funzione complessa f , per cui:

$$\begin{aligned} \oint_{\partial D} f(z) dz &= \oint_{\partial D} u dx - v dy + i \oint_{\partial D} v dx + u dy \\ &= \iint_D \left(-\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} \right) dx dy + i \iint_D \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy = 0 \end{aligned}$$

²⁰ **George Green** (Sneinton, Nottinghamshire, Inghilterra, 14 luglio 1793 – Nottingham, Nottinghamshire, Inghilterra, 31 maggio 1841), matematico e fisico inglese, scrisse nel 1828 un saggio “An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism” che introdusse molti concetti importanti, tra cui un teorema simile al moderno teorema di Green, l’idea della funzione potenziale usata in fisica, ed il concetto di ciò che ora è nota come funzione di Green. La storia della sua vita è notevole in quanto fu praticamente un autodidatta. Figlio di un mugnaio nella cittadina di Sneinton, ora un sobborgo di Nottingham, nell’infanzia frequentò la scuola solo per un anno (all’età di otto e nove anni) e nell’età adulta lavorò nel mulino a vento del padre (il mulino è ora un museo scientifico dedicato a George Green). Non è chiaro come riuscisse a studiare con profitto la matematica ed essere aggiornato sui progressi della fisica del tempo, in quanto la sua cittadina non aveva praticamente alcuna attività culturale e l’unica persona che conosceva la matematica era il direttore della “High School” di Nottingham. Il suo saggio del 1828 fu sostenuto dalla sottoscrizione di 51 suoi concittadini (che probabilmente non comprendevano affatto quanto c’era scritto). Fu però acquistato dal ricco matematico e possidente terriero Edward Bromhead, che incoraggiò Green nei suoi studi e gli permise di entrare come studente all’Università di Cambridge all’età di 40 anni. La sua carriera accademica fu eccellente e si occupò con successo di problematiche fisiche e matematiche, anticipando concetti e metodologie che saranno usate in seguito. Einstein, durante una visita a Nottingham, affermò che Green era avanti di 20 anni rispetto al suo tempo. Purtroppo Green morì poco tempo dopo aver conseguito la laurea.



in virtù delle condizioni di Cauchy-Riemann.

Questa semplice dimostrazione è stata formulata nel 1848 da Stokes²¹, che generalizzò il teorema di Green a superfici dello spazio. Notiamo che, in ultima analisi, il contenuto concettuale dei teoremi di Green, di Stokes e della divergenza di Gauss è lo stesso, sono tutti una generalizzazione del teorema fondamentale del calcolo integrale.

— ○ —

Soluzione 1.7 $f_1(z) = \frac{e^{2z}}{(z-1)^3}; \quad z_0 = 1$

Conviene porre $z-1 = u$, $z = 1+u$,

$$\begin{aligned} \frac{e^{2z}}{(z-1)^3} &= e^2 \frac{e^{2u}}{u^3} = \frac{e^2}{u^3} \left\{ 1 + 2u + \frac{(2u)^2}{2!} + \frac{(2u)^3}{3!} + \dots \right\} \\ &= \frac{e^2}{(z-1)^3} + \frac{2e^2}{(z-1)^2} + \frac{2e^2}{(z-1)} + \frac{4e^2}{3} + \frac{2e^2}{3}(z-1) + \dots, \end{aligned}$$

$z = 1$ è un polo di ordine 3 (polo triplo).

$f_2(z) = (z-3) \sin\left(\frac{1}{z+2}\right); \quad z_0 = -2$.

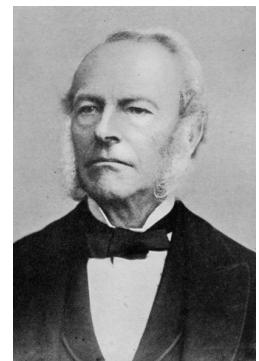
$$z+2 = u; \quad z = u-2$$

$$\begin{aligned} (z-3) \sin\left(\frac{1}{z+2}\right) &= (u-5) \sin\frac{1}{u} = (u-5) \left\{ \frac{1}{u} - \frac{1}{3!u^3} + \frac{1}{5!u^5} - \dots \right\} \\ &= 1 - \frac{5}{u} - \frac{1}{6u^2} + \frac{5}{6u^3} + \frac{1}{120u^4} - \frac{1}{24u^5} + \dots \\ &= 1 - \frac{5}{z+2} - \frac{1}{6(z+2)^2} + \frac{5}{6(z+2)^3} + \frac{1}{120(z+2)^4} - \dots, \end{aligned}$$

Il punto $z = -2$ è una singolarità essenziale e la serie converge per ogni valore di $z \neq -2$.

$f_3(z) = \frac{z - \sin z}{z^3}; \quad z_0 = 0$.

²¹ **George Gabriel Stokes** (Skreen, County Sligo, Irlanda, 13 agosto 1819 – Cambridge, Inghilterra, 1 febbraio 1903) è stato un fisico e matematico irlandese. Possiamo dire che si occupava di tematiche poco amate dai fisici-matematici dell'epoca che affrontava con abilità sia sperimentalmente che teoricamente. Ha stabilito la legge della resistenza incontrata da una sfera in moto in un fluido viscoso (legge di Stokes) e studiato il fenomeno della fluorescenza prodotta da raggi ultravioletti. A suo nome è stata intitolata l'unità di misura della viscosità dinamica (Stokes, simbolo St). Il teorema di Stokes è un esempio notevole della sua abilità matematica.



$$\begin{aligned}\frac{z - \sin z}{z^3} &= \frac{1}{z^3} \left\{ z - \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots \right) \right\} \\ &= \frac{1}{6} - \frac{z^2}{5!} + \frac{z^4}{7!} + \dots,\end{aligned}$$

$z = 0$ è una singolarità removibile e la serie (effettivamente una serie di Taylor) è convergente per ogni z .

— ○ —

Soluzione 1.8

$$\begin{aligned}z + 2 &= u \quad z = u - 2 \\ \frac{z}{(z+1)(z+2)} &= \frac{u-2}{(u-1)u} = \frac{2-u}{u} \frac{1}{1-u} \\ &= \frac{2-u}{u} \{1 + u + u^2 + u^3 + \dots\} \\ &= \frac{1}{u} \{2 - u + 2u - u^2 + 2u^2 - u^3 + 2u^3 - u^4 + \dots\} \\ &= \frac{1}{u} \{2 + u + u^2 + u^3 + u^3 + \dots\} \\ &= \frac{2}{u} + 1 + u + u^2 + \dots \\ &= \frac{2}{z+2} + 1 + (z+2) + (z+2)^2 + \dots = \frac{2}{z+2} - \frac{1}{z+1}\end{aligned}$$

$z = -2$ è un polo di ordine 1 (polo semplice) e la serie converge per $0 < |z+2| < 1$.

— ○ —

Soluzione 1.9

$$f(z) = \frac{1}{(z+1)(z+3)} = \frac{1}{2} \frac{1}{(z+1)} - \frac{1}{2} \frac{1}{(z+3)}$$

1. Se $|z| > 1$:

$$\begin{aligned}\frac{1}{2} \frac{1}{z+1} &= \frac{1}{2z \left(1 + \frac{1}{z}\right)} \\ &= \frac{1}{2z} \left(1 - \frac{1}{z} + \frac{1}{z^2} - \frac{1}{z^3} + \dots\right) \\ &= \frac{1}{2z} - \frac{1}{2z^2} + \frac{1}{2z^3} - \dots,\end{aligned}$$

Inoltre, se $|z| < 3$:

$$\frac{1}{2(z+3)} = \frac{1}{6} \frac{1}{1 + \frac{z}{3}} = \frac{1}{6} \left(1 - \frac{z}{3} + \frac{z^2}{9} - \frac{z^3}{27} + \dots \right)$$

pertanto, se $1 < |z| < 3$:

$$f(z) = \dots - \frac{1}{2z^4} + \frac{1}{2z^3} - \frac{1}{2z^2} + \frac{1}{2z} - \frac{1}{6} + \frac{z}{18} - \frac{z^2}{54} + \dots$$

Notiamo che $z = 0$ non è una singolarità di $f(z)$.

2. Se $|z| > 3$ possiamo scrivere:

$$\begin{aligned} \frac{1}{2(z+3)} &= \frac{1}{2z} \frac{1}{\left(1 + \frac{3}{z}\right)} = \frac{1}{2z} \left(1 - \frac{3}{z} + \frac{9}{z^2} - \frac{27}{z^3} + \dots \right) \\ &= \frac{1}{2z} - \frac{3}{2z^2} + \frac{9}{2z^3} - \frac{27}{2z^4} + \dots, \end{aligned}$$

pertanto:

$$f(z) = \frac{1}{z^2} - \frac{4}{z^3} + \frac{13}{z^4} - \dots$$

3. Poniamo $z + 1 = u$:

$$\begin{aligned} \frac{1}{(z+1)(z+3)} &= \frac{1}{u(u+2)} = \frac{1}{2u} \frac{1}{1 + \frac{u}{2}} \\ &= \frac{1}{2u} \left(1 - \frac{u}{2} + \frac{u^2}{4} - \frac{u^3}{8} + \dots \right) \\ &= \frac{1}{2u} - \frac{1}{4} + \frac{u}{8} - \frac{u^2}{16} + \dots \\ &= \frac{1}{2(z+1)} - \frac{1}{4} + \frac{(z+1)}{8} - \frac{(z+1)^2}{16} + \dots \end{aligned}$$

4. $|z| < 1$:

$$\begin{aligned} \frac{1}{2(z+1)} &= \frac{1}{2} (1 - z + z^2 - z^3 + \dots) = \frac{1}{2} - \frac{z}{2} + \frac{z^2}{2} - \frac{z^3}{2} + \dots, \\ f(z) &= \frac{1}{3} - \frac{4}{9}z + \frac{13}{27}z^2 - \frac{40}{81}z^3 + \dots, \end{aligned}$$

che è in realtà una serie di Taylor ($z = 0$ è un punto regolare).

— ○ —

Soluzione 1.10 Abbiamo:

$$\begin{aligned} \int_0^{2\pi} \frac{dt}{a + \sin t} &= \oint_{|z|=1} \frac{1}{a + \frac{1}{2i} \left(z - \frac{1}{z} \right)} \frac{dz}{iz} \\ &= \oint_{|z|=1} \frac{2dz}{z^2 + 2iaz - 1} \end{aligned}$$

L'integrando ha due poli semplici in:

$$z_{\pm} = -i(a \pm \sqrt{a^2 - 1}),$$

Notando che se $a > 1$ allora:

$$\begin{aligned} a^2 - 2a < a^2 - 2, \quad (a-1)^2 < a^2 - 1, \quad a-1 < \sqrt{a^2 - 1} \\ a - \sqrt{a^2 - 1} < 1, \end{aligned}$$

abbiamo che il polo z_- è situato all'interno della circonferenza unitaria (z_+ è chiaramente esterno), e il residuo dell'integrando vale:

$$R_- = \frac{2}{2z + 2ia} \Big|_{z=z_-} = \frac{1}{i\sqrt{a^2 - 1}}$$

da cui il risultato.

— ○ —

Soluzione 1.11 Abbiamo:

$$\int_0^\infty \frac{dx}{1+x^6} = \frac{1}{2} \int_{-\infty}^\infty \frac{dx}{1+x^6} = \pi i \sum_{\Im z_k > 0} \operatorname{Res} \frac{1}{1+z^6} \Big|_{z=z_k}.$$

Le singolarità sono le soluzioni dell'equazione:

$$z^6 = -1 = e^{i\pi}, \quad z_k = e^{i(2k+1)\pi/6}, \quad k = 0, \dots, 5.$$

Sono poli semplici con residuo:

$$R_k = \frac{1}{6z_k^5} = \frac{1}{6} e^{-i(2k+1)5\pi/6}$$

e quindi:

$$\begin{aligned} \int_0^\infty \frac{dx}{1+x^6} &= \frac{\pi}{6} i \left\{ e^{-i5\pi/6} + e^{-i15\pi/6} + e^{-i25\pi/6} \right\} \\ &= \frac{\pi}{6} i (-i - 2i \sin(\pi/6)) = \frac{\pi}{6} i (-2i) \\ &= \frac{\pi}{3}. \end{aligned}$$

— ○ —

Soluzione 1.12 Notando che l'integrandi è una funzione pari, abbiamo:

$$\begin{aligned}
 I &= \frac{1}{2} \int_{-\infty}^{\infty} \frac{\cos x}{x^2 + 1} dx = \lim_{r \rightarrow \infty} \frac{1}{2} \int_{-r}^r \frac{\cos x}{x^2 + 1} dx \\
 &= \lim_{r \rightarrow \infty} \frac{1}{2} \operatorname{Re} \int_{-r}^r \frac{e^{ix}}{x^2 + 1} dx = \operatorname{Re} \pi i \operatorname{Res} \left(\frac{e^{iz}}{z^2 + 1} \right) \Big|_{z=i} \\
 &= \operatorname{Re} \pi i \frac{e^{iz}}{2z} \Big|_{z=i} = \pi \frac{e^{-1}}{2} = \frac{\pi}{2e}.
 \end{aligned}$$

— ○ —

Soluzione 1.13 La funzione

$$f(z) = \frac{e^{i\alpha z}}{z^2 + \alpha^2}$$

ha due poli semplici in $z = \pm i|\alpha|$ e soddisfa il lemma di Jordan in $\Im z > 0$ per $\alpha > 0$ e in $\Im z < 0$ per $\alpha < 0$. Integriamo la funzione sul percorso di figura quando $\alpha > 0$:

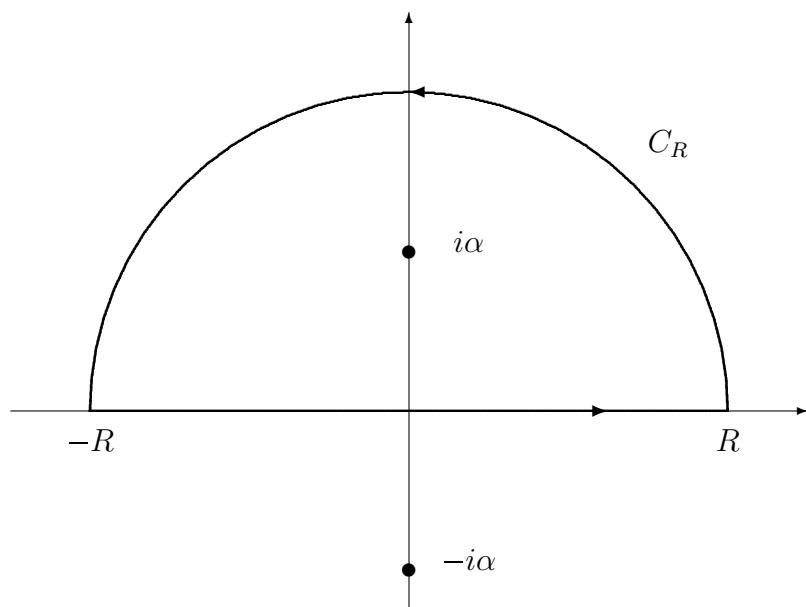


Figura 1.36: Percorso nel semipiano superiore.

$$\begin{aligned}
& \int_{-R}^0 \frac{e^{i\alpha x}}{x^2 + \alpha^2} dx + \int_0^R \frac{e^{i\alpha x}}{x^2 + \alpha^2} dx = 2\pi i \operatorname{Res} f(z) \Big|_{z=i\alpha} \\
& 2 \int_0^R \frac{\cos(\alpha x)}{x^2 + \alpha^2} dx = 2\pi i \operatorname{Res} f(z) \Big|_{z=i\alpha} \\
& F(\alpha) = i\pi \operatorname{Res} f(z) \Big|_{z=i\alpha}, \quad \alpha > 0, \\
& \operatorname{Res} f(z) \Big|_{z=i\alpha} = \lim_{z \rightarrow i\alpha} (z - i\alpha) \frac{e^{i\alpha z}}{(z + i\alpha)(z - i\alpha)} = \frac{e^{-\alpha^2}}{2i\alpha} \\
& F(\alpha) = \frac{\pi}{2} \frac{e^{-\alpha^2}}{\alpha}, \quad \alpha > 0.
\end{aligned}$$

Poichè $F(-\alpha) = F(\alpha)$ abbiamo:

$$F(\alpha) = \frac{\pi}{2} \frac{e^{-\alpha^2}}{|-\alpha|}, \quad \alpha < 0,$$

e globalmente:

$$F(\alpha) = \frac{\pi}{2} \frac{e^{-\alpha^2}}{|\alpha|}, \quad \alpha \neq 0.$$

Il caso $\alpha < 0$ può essere visto anche utilizzando un percorso nel semipiano inferiore come in figura 1.37.

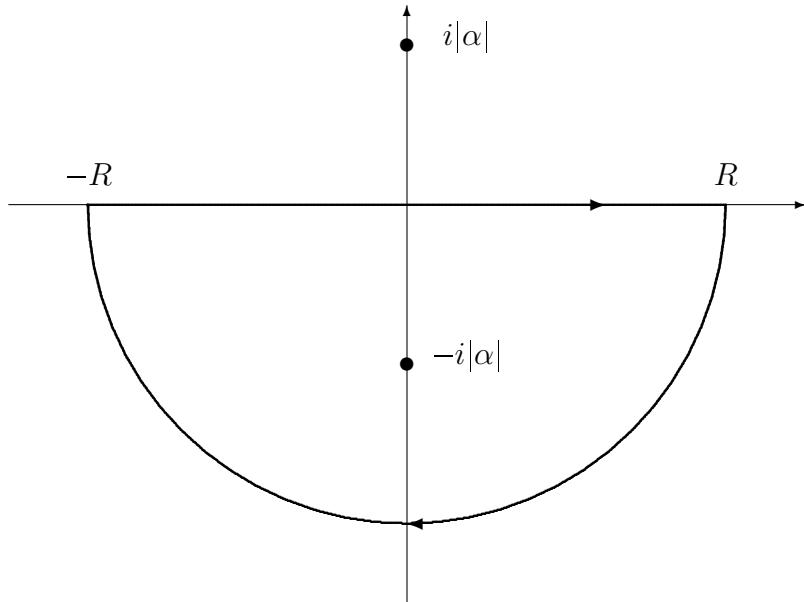


Figura 1.37: Percorso nel semipiano inferiore.

$$2 \int_0^R \frac{\cos(\alpha x)}{x^2 + \alpha^2} = -2\pi i \operatorname{Res} f(z) \Big|_{z=-i|\alpha|},$$

$$F(\alpha) = -\frac{\pi}{2} \frac{e^{-\alpha^2}}{\alpha}, \quad \alpha < 0.$$

— ○ —

Soluzione 1.14 La funzione

$$f(z) = \frac{e^{i\alpha z}}{(1-iz)(1+2iz)} = \frac{e^{i\alpha z}}{-i(z+i)2i(z-\frac{i}{2})} = \frac{e^{i\alpha z}}{2(z+i)(z-\frac{i}{2})}$$

ha poli semplici in $z_1 = -i$ e $z_2 = \frac{i}{2}$. Per $\alpha > 0$ soddisfa il lemma di Jordan in $\Im z > 0$, per $\alpha < 0$ in $\Im z < 0$. Per $\alpha = 0$ soddisfa il lemma del grande cerchio. Applicando il teorema dei residui e il lemma di Jordan si ottiene (per $\alpha > 0$, figura 1.38):

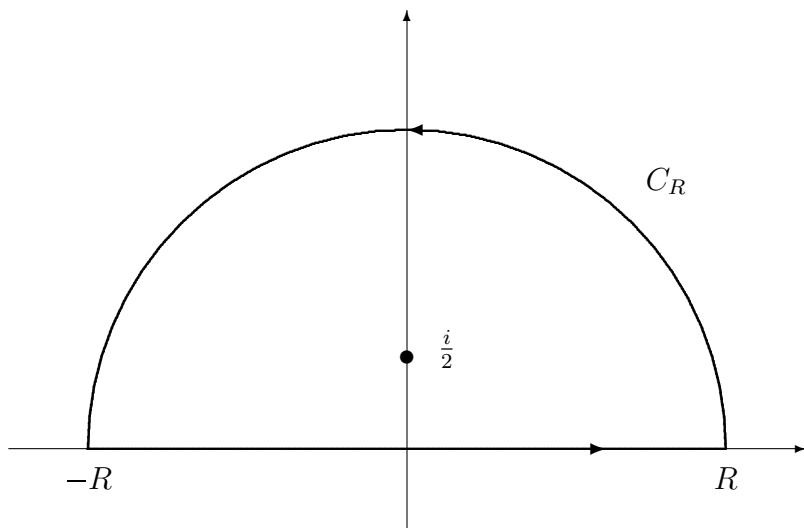
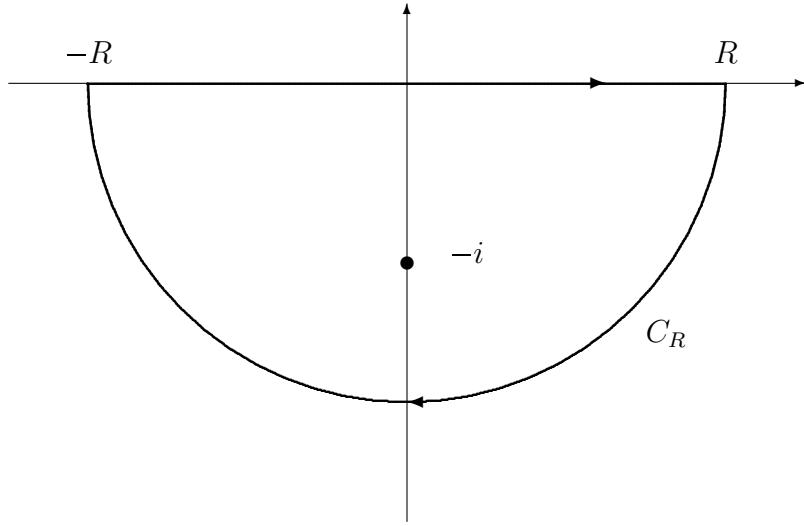


Figura 1.38: Esercizio 1.14, $\alpha > 0$.

$$\begin{aligned} I_\alpha &= 2\pi i \operatorname{Res} f(z) \Big|_{z=\frac{i}{2}} \\ \operatorname{Res} f(z) \Big|_{z=\frac{i}{2}} &= \lim_{z \rightarrow \frac{i}{2}} \frac{e^{i\alpha z}}{2(z+i)} = \frac{e^{-\frac{\alpha}{2}}}{3i} \end{aligned}$$

Nel caso $\alpha < 0$ consideriamo invece il percorso della figura 1.39:

$$\begin{aligned} I_\alpha &= -2\pi i \operatorname{Res} f(z) \Big|_{z=-i} \\ \operatorname{Res} f(z) \Big|_{z=-i} &= \lim_{z \rightarrow -i} \frac{e^{i\alpha z}}{2(z-\frac{i}{2})} = -\frac{e^\alpha}{3i}, \end{aligned}$$

Figura 1.39: Esercizio 1.14, $\alpha < 0$.

Riassumendo (per $\alpha = 0$ la scelta tra i due percorsi è indifferente):

$$I_\alpha = \begin{cases} \frac{2\pi}{3} e^{-\frac{\alpha}{2}} & \alpha \geq 0, \\ \frac{2\pi}{3} e^\alpha & \alpha \leq 0. \end{cases}$$

— ○ —

Soluzione 1.15 Abbiamo, sfruttando la parità dell'integrando:

$$\begin{aligned} I &= \int_0^\infty \frac{\sin x}{x} dx = \lim_{\substack{\epsilon \rightarrow 0 \\ r \rightarrow \infty}} \int_\epsilon^r \frac{\sin x}{x} dx \\ &= \frac{1}{2} \lim_{\substack{\epsilon \rightarrow 0 \\ r \rightarrow \infty}} \operatorname{Im} \int_{\epsilon < |x| < r} \frac{e^{ix}}{x} dx \end{aligned}$$

La funzione $\frac{e^{iz}}{z}$ verifica il lemma di Jordan per $|z| \rightarrow \infty$ e presenta un polo semplice nell'origine, per cui considerando la curva $\Gamma_{\epsilon,r}$ della figura 1.35, non abbiamo singolarità interne al cammino di integrazione e, nel limite $\epsilon \rightarrow 0$, $r \rightarrow \infty$:

$$\begin{aligned} 0 &= \oint_{\Gamma_{\epsilon,r}} \frac{e^{iz}}{z} dz \rightarrow \lim_{\substack{\epsilon \rightarrow 0 \\ r \rightarrow \infty}} \int_{\epsilon < |x| < r} \frac{e^{ix}}{x} dx - \pi i \operatorname{Res} \left. \frac{e^{iz}}{z} \right|_{z=0} \\ &\quad \lim_{\substack{\epsilon \rightarrow 0 \\ r \rightarrow \infty}} \int_{\epsilon < |x| < r} \frac{e^{ix}}{x} dx = \pi i e^{iz} \Big|_{z=0} = \pi i \\ I &= \frac{\pi}{2}. \end{aligned}$$

— ○ —

Soluzione 1.16 Possiamo scrivere:

$$2i I = P \int_{-\infty}^{\infty} \frac{e^{ix}}{(x-1)(x-i)} dx - P \int_{-\infty}^{\infty} \frac{e^{-ix}}{(x-1)(x-i)} dx = I_+ - I_-$$

$$I_{\pm} = \lim_{\substack{R \rightarrow \infty \\ \epsilon \rightarrow 0}} \left\{ \int_{-R}^{1-\epsilon} + \int_{1+\epsilon}^R \right\} \frac{e^{\pm ix}}{(x-1)(x-i)} dx$$

Considerando gli integrandi come funzioni (complesse) della variabile complessa z :

$$f_{\pm}(z) = \frac{e^{\pm iz}}{(z-1)(z-i)},$$

questi presentano due poli semplici in $z = 1$ e $z = i$ e:

$$\text{Res } \frac{e^{\pm iz}}{(z-1)(z-i)} \Big|_{z=i} = \frac{e^{\mp 1}}{i-1}$$

$$\text{Res } \frac{e^{\pm iz}}{(z-1)(z-i)} \Big|_{z=1} = \frac{e^{\pm i}}{1-i} = -\frac{e^{\pm i}}{i-1}$$

Per l'integrale I_+ consideriamo il percorso in figura 1.40:

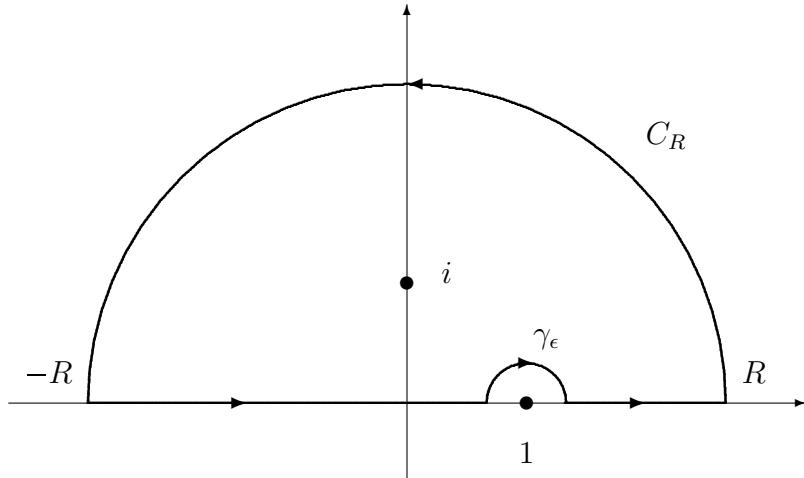


Figura 1.40: Esercizio 1.16.

$$I_+ = 2i\pi \left(\frac{e^{-1}}{i-1} \right) + i\pi \left(-\frac{e^i}{i-1} \right)$$

Per l'integrale I_- consideriamo invece il percorso di figura 1.41:

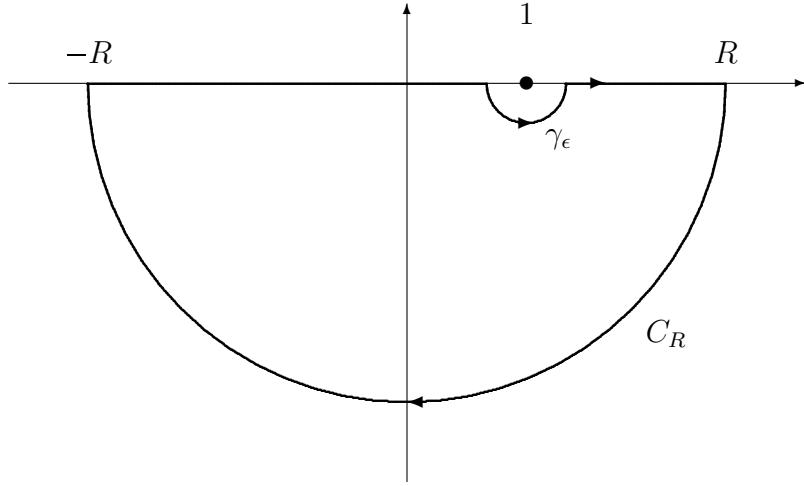


Figura 1.41: Esercizio 1.16.

$$\begin{aligned}
 I_- &= -i\pi \left(-\frac{e^{-i}}{i-1} \right) = \frac{i\pi e^{-i}}{i-1} \\
 2iI &= \frac{2i\pi e^{-1}}{i-1} - \frac{i\pi e^i}{i-1} - \frac{i\pi e^{-i}}{i-1} = \frac{2i\pi e^{-1}}{i-1} - \frac{2i\pi}{i-1} \left(\frac{e^i + e^{-i}}{2} \right) \\
 I &= \frac{\pi}{i-1} (e^{-1} - \cos(1)) .
 \end{aligned}$$

— ○ —

Soluzione 1.17 Ripetendo il ragionamento discusso in generale:

$$\begin{aligned}
 I &= \frac{2\pi i}{1 - e^{-2\pi i \alpha}} \operatorname{Res} \left. \frac{1}{z^\alpha (1+z)} \right|_{z=-1} = \frac{2\pi i}{1 - e^{-2\pi i \alpha}} \left. \frac{1}{z^\alpha} \right|_{z=-1} \\
 &= \frac{2\pi i}{1 - e^{-2\pi i \alpha}} \frac{1}{e^{\pi i \alpha}} = \frac{2\pi i}{e^{\pi i \alpha} - e^{-\pi i \alpha}} = \frac{\pi}{\sin(\pi \alpha)}
 \end{aligned}$$

e, in particolare, con $\alpha = \frac{1}{2}$:

$$\int_0^\infty \frac{dx}{\sqrt{x}} \frac{1}{1+x} = \pi$$

— ○ —

Soluzione 1.18 La funzione

$$f(z) = \frac{1}{(z+i)\sqrt{z}}$$

ha un polo semplice in $z = -i$ e un punto di diramazione in $z = 0$. Scegliendo la determinazione principale per \sqrt{z} e il taglio di diramazione lungo l'asse reale positivo, si può considerare il percorso come in figura 1.42. Per il teorema dei residui:

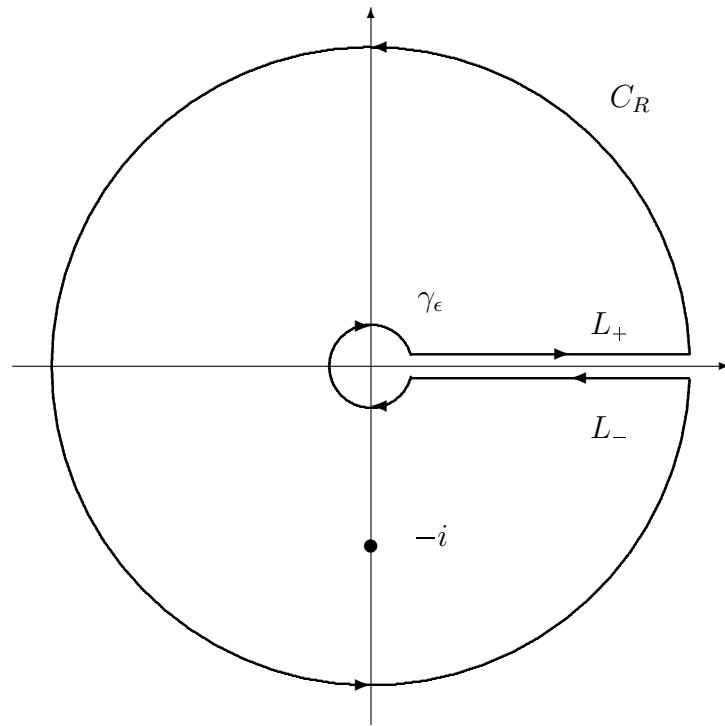


Figura 1.42: Esercizio 1.18.

$$\int_{\gamma_R} f(z) dz + \int_{\gamma_\epsilon} f(z) dz + \int_{L^+} f^+(x) dx + \int_{L^-} f^-(x) dx = 2\pi i \operatorname{Res} f(z) \Big|_{z=-i}$$

$$f^+(x) = \frac{1}{(x+i)\sqrt{x}}, \quad f^-(x) = \frac{1}{(x+i)(-\sqrt{x})}$$

$$f^+(x) - f^-(x) = \frac{2}{(x+i)\sqrt{x}}$$

$$\operatorname{Res} f(z) \Big|_{z=-i} = \lim_{z \rightarrow -i} \frac{1}{\sqrt{z}} = e^{-i \frac{3}{4}\pi}$$

$$\int_{\gamma_R} f(z) dz + \int_{\gamma_\epsilon} f(z) dz + \int_{\epsilon}^R [f^+(x) - f^-(x)] dx = 2\pi e^{-i \frac{\pi}{4}}$$

Poichè

$$|zf(z)| \xrightarrow[R \rightarrow \infty]{\epsilon \rightarrow 0} 0$$

risultano verificati i lemmi del grande cerchio e del piccolo cerchio e:

$$2 \int_0^\infty \frac{dx}{(x+i)\sqrt{x}} = 2\pi e^{-i\frac{\pi}{4}}$$

$$I = \pi e^{-i\frac{\pi}{4}} = \pi \frac{(1-i)}{\sqrt{2}}.$$

— ○ —

Soluzione 1.19 Utilizzando la formula risolutiva (1.157) abbiamo:

$$4\pi^2 \int_0^\infty \frac{1}{(1+x)^3} dx - 4\pi i \int_0^\infty \frac{\log x}{(1+x)^3} dx = 2\pi i \operatorname{Res} \left(\frac{\log^2 z}{(1+z)^3} \right)_{z=-1}$$

$$= \pi i \left. \frac{d^2}{dz^2} \log^2 z \right|_{z=-1} = 2\pi i \left. \frac{d}{dz} \frac{\log z}{z} \right|_{z=-1}$$

$$= 2\pi i \left. \frac{1-\log z}{z^2} \right|_{z=-1}$$

$$= 2\pi i(1-\pi i) = 2\pi^2 + 2\pi i$$

Da cui i risultati richiesti.

— ○ —

Soluzione 1.20 L'integrando è singolare in $x = \pm 1$ (escludendo la singolarità nell'origine del logaritmo, comunque integrabile). I punti $x = \pm 1$ sono poli semplici e, applicando il procedimento discusso in generale:

$$2\pi i \operatorname{Res} \left. \frac{\log^2 z}{z^2-1} \right|_{z=-1} = 4\pi^2 P \int_0^\infty \frac{1}{x^2-1} dx - 4\pi i \int_0^\infty \frac{\log x}{x^2-1} dx$$

$$+ 4\pi^3 i \operatorname{Res} \left. \frac{1}{z^2-1} \right|_{z=+1}$$

$$2\pi i \left. \frac{\log^2 z}{z-1} \right|_{z=-1} = 4\pi^2 P \int_0^\infty \frac{1}{x^2-1} dx - 4\pi i \int_0^\infty \frac{\log x}{x^2-1} dx$$

$$+ 4\pi^3 i \left. \frac{1}{z+1} \right|_{z=+1}$$

$$\pi^3 i = 4\pi^2 P \int_0^\infty \frac{1}{x^2-1} dx - 4\pi i \int_0^\infty \frac{\log x}{x^2-1} dx + 2\pi^3 i$$

$$-\pi^3 i = 4\pi^2 P \int_0^\infty \frac{1}{x^2-1} dx - 4\pi i \int_0^\infty \frac{\log x}{x^2-1} dx$$

da cui:

$$\int_0^\infty \frac{\log x}{x^2 - 1} dx = \frac{\pi^2}{4},$$

$$P \int_0^\infty \frac{1}{x^2 - 1} dx = 0.$$

— ○ —

Capitolo 2.

Topologia.

2.1 Spazi topologici.

La topologia può essere definita, in prima approssimazione, come lo studio dei concetti di intorno di un punto, passaggio al limite e continuità. Queste nozioni esprimono sostanzialmente il concetto di “vicinanza” tra oggetti matematici. Uno studio preliminare della topologia in astratto è importante per stabilire un linguaggio corretto per la formulazione dei risultati matematici. Le definizioni rigorose dei concetti di insieme aperto, insieme chiuso, di intorno, di limite e continuità, risultano molto importanti nel caso di spazi a dimensione infinita, più di quanto non lo siano negli spazi finito dimensionali. Possiamo individuare ad esempio due ragioni.

1) Nel caso di \mathbb{C}^n esiste essenzialmente un solo modo in cui una successione di vettori o una successione di matrici possono tendere ad un limite. Le relazioni:

$$\begin{aligned}\lim_{m \rightarrow \infty} \mathbf{v}^{(m)} &= \mathbf{v}, \quad \mathbf{v}^{(m)}, \mathbf{v} \in \mathbb{C}^n, \\ \lim_{m \rightarrow \infty} A^{(m)} &= A, \quad A^{(m)}, A \text{ matrici complesse } n \times n,\end{aligned}$$

possono significare indifferentemente una convergenza numerica delle singole componenti dei vettori o delle matrici:

$$\begin{aligned}\lim_{m \rightarrow \infty} v_i^{(m)} &= v_i, \quad i = 1, 2, \dots, n, \\ \lim_{m \rightarrow \infty} A_{ij}^{(m)} &= A_{ij}, \quad i, j = 1, 2, \dots, n,\end{aligned}$$

oppure, usando la norma dei vettori:

$$\begin{aligned}\lim_{m \rightarrow \infty} |\mathbf{v}^{(m)} - \mathbf{v}| &= 0, \\ \lim_{m \rightarrow \infty} |(A^{(m)} - A)\mathbf{v}| &= 0, \quad \text{per ogni } \mathbf{v} \in \mathbb{C}^n.\end{aligned}$$

Invece nel caso generale ciò non è più vero. Possono esistere diversi modi *non equivalenti tra di loro* in cui si può tendere ad un limite. Ad esempio quando affermiamo che una successione di funzioni f_n converge ad una funzione limite f possiamo intendere un semplice limite puntuale:

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \forall x,$$

oppure una convergenza in media:

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} |f_n(x) - f(x)| dx,$$

a seconda del contesto e delle necessità, e i due modi non sono, in generale, equivalenti tra loro. Stabilire un contesto univoco per la definizione di limite definisce la topologia in cui si opera.

2) Nel caso di spazi finito dimensionali l'analogia con lo spazio ordinario \mathbb{R}^3 è ancora abbastanza stretta, per cui si “vede” abbastanza bene il significato della frase “**x** vicino a **y**”. Invece in spazi a dimensioni infinite risulta difficile farsi una idea geometrica intuitiva dei concetti in questione ed occorre appellarsi al ragionamento ed a definizioni rigorose.

2.1.1 Insiemi aperti

Alla base della topologia si trova il concetto di insieme aperto. Tramite gli insiemi aperti vengono stabiliti infatti i concetti di continuità e di limite. Solitamente si introducono dapprima le nozioni fondamentali degli spazi metrici (in cui cioè si definisce una distanza tra i punti), quindi, basandosi sulle proprietà della distanza, si definiscono gli insiemi aperti. Si può del resto agire altrimenti, senza introdurre una metrica, ma definire in un dato insieme un sistema di insiemi aperti in maniera assiomatica richiedendo solo le proprietà essenziali. Questo metodo garantisce una maggiore libertà d’azione e conduce alla classe degli **spazi topologici** nei confronti dei quali gli spazi metrici rappresentano un caso particolare.

Def. 2.1 (Spazio topologico) Sia X un insieme non vuoto (il supporto della topologia), una **topologia** su X è una famiglia \mathcal{A} di sottoinsiemi $A \subseteq X$, detti **aperti** con le seguenti caratteristiche:

i) L’insieme vuoto e tutto il supporto sono aperti:

$$\emptyset \in \mathcal{A}, \quad X \in \mathcal{A}. \quad (2.1)$$

ii) L’unione di una collezione **arbitraria** di insiemi aperti è un insieme aperto:

$$A_i \in \mathcal{A}, \quad i \in \mathcal{I} \quad \implies \quad \bigcup_{i \in \mathcal{I}} A_i \in \mathcal{A}, \quad (2.2)$$

con \mathcal{I} insieme di indici.

iii) L’intersezione di una collezione **finita** di insiemi aperti è un insieme aperto:

$$A_i \in \mathcal{A}, \quad i = 1, 2, \dots, N \text{ (finito)} \quad \implies \quad \bigcap_{i=1}^N A_i \in \mathcal{A}. \quad (2.3)$$

L’insieme X con la topologia \mathcal{A} in esso assegnata, cioè la coppia (X, \mathcal{A}) , viene detto **spazio topologico**.

Osservazione. Osserviamo che, in base alla definizione assiomatica data, il concetto di aperto non è necessariamente una proprietà dell'insieme in sè, ma dipende da come sono definiti globalmente tutti i possibili insiemi aperti. Preso singolarmente non possiamo dire a priori se un insieme è aperto oppure no, tale proprietà dipende dalle sue relazioni con gli altri insiemi aperti, cioè dal contesto topologico in cui operiamo. Le proprietà essenziali degli insiemi aperti sono individuate nelle operazioni insiemistiche di unione e intersezione. Una topologia è una collezione “chiusa” rispetto a tali operazioni.

Tramite l'appartenenza di diversi elementi di X al medesimo sottoinsieme aperto o ad aperti distinti si formalizzano le nozioni intuitive di “vicinanza” o “lontananza” tra essi, senza ricorrere ad una valutazione di natura metrica.

Esempio 2.1 Consideriamo nell'insieme \mathbb{R} dei numeri reali (la retta reale) la nozione (quasi intuitiva) di insieme aperto. Dato un sottinsieme A dei numeri reali, verifichiamo la sua apertura se, preso un qualsiasi punto $p \in A$, possiamo trovare un intervallo (sottinteso aperto) del tipo:

$$]a, b[= \{x \in \mathbb{R} ; a < x < b\}, \quad a < b,$$

con $a < p < b$, tutto contenuto in A : $]a, b[\subseteq A$. Questa verifica sembra definire un aperto in maniera intrinseca, ma proprio le proprietà di tali insiemi hanno fornito il modello per una loro astrazione. Possiamo dare una definizione ugualmente elementare di insiemi aperti della retta reale: un aperto è un qualsiasi sottoinsieme della retta che risulti unione (arbitraria) di intervalli del tipo $]a, b[$, oppure l'insieme vuoto \emptyset . È immediato controllare la validità delle proprietà della definizione 2.1. Riguardo all'ultima proprietà osserviamo che l'intersezione di infiniti intervalli aperti del tipo:

$$-a < x < a \quad \text{con } a \text{ positivo ma arbitrario,}$$

risulta costituito dal solo punto $\{x = 0\}$ che non è un insieme aperto. Si può allora capire che è necessario limitare l'intersezione ad un numero *finito* di aperti per avere ancora un aperto.

— ○ —

Esempio 2.2 Se $X \neq \emptyset$, possiamo definire come aperto un qualsiasi sottoinsieme di X (compreso X stesso e l'insieme vuoto \emptyset). La topologia risultante viene detta **discreta** oppure **massimale**. La denominazione discreta deriva dal fatto che con questa topologia ogni punto è separato dagli altri ed è “vicino” solo a sè stesso.

— ○ —

Esempio 2.3 Sia X un insieme arbitrario e definiamo come soli aperti l'insieme vuoto \emptyset e X stesso:

$$\mathcal{A} = \{\emptyset, X\}.$$

Otteniamo in questo modo la topologia detta **banale** o **minimale** e si ottiene uno spazio di “punti incollati” (la topologia non riesce a distinguere tra un punto e l'altro, tutti i punti

sono “vicini” tra loro, non è possibile trovare due punti diversi appartenenti ad aperti distinti).

— ○ —

I due esempi precedenti non forniscono delle topologie veramente utili in pratica, la loro importanza è solo teorica, di esistenza. Comportano che è sempre possibile definire una topologia in un qualsiasi insieme non vuoto e costituiscono degli “estremi” entro cui delimitare qualsiasi topologia. Il seguente esempio ci permette di affermare che possiamo definire una topologia anche in un insieme discreto.

Esempio 2.4 Sia $X = \{a, b, c\}$ costituito da soli tre elementi e definiamo la famiglia di aperti:

$$\mathcal{A} = \{\emptyset, X, \{a\}, \{a, b\}\} .$$

Le proprietà della definizione 2.1 sono verificabili direttamente.

— ○ —

2.1.2 Insiemi chiusi

Una volta definiti gli aperti possiamo definire gli insiemi chiusi in maniera banale come i loro complementari.

Def. 2.2 Gli insiemi del tipo $C = X \setminus A$, complementari di insiemi aperti A in uno spazio topologico X sono detti **chiusi**:

$$C \text{ chiuso} \iff X \setminus C \text{ aperto.} \quad (2.4)$$

Dagli assiomi sugli aperti, utilizzando le **relazioni di dualità**:

- 1) il complementare dell’unione è uguale all’intersezione dei complementari:

$$X \setminus \left(\bigcup_{\alpha} W_{\alpha} \right) = \bigcap_{\alpha} (X \setminus W_{\alpha}); \quad (2.5)$$

- 2) il complementare dell’intersezione è uguale all’unione dei complementari:

$$X \setminus \left(\bigcap_{\alpha} W_{\alpha} \right) = \bigcup_{\alpha} (X \setminus W_{\alpha}); \quad (2.6)$$

si ottengono facilmente le proprietà dei chiusi (notiamo che nelle relazioni di dualità non viene richiesta alcuna restrizione alle unioni ed intersezioni e le relazioni sono valide per collezioni arbitrarie di insiemi W_{α}).

Teo. 2.1 Sia X uno spazio topologico. Allora la famiglia \mathcal{C} costituita dagli insiemi chiusi in X gode delle seguenti proprietà:

i) L'insieme vuoto \emptyset e tutto X sono chiusi:

$$\emptyset \in \mathcal{C}, \quad X \in \mathcal{C}. \quad (2.7)$$

ii) L'intersezione di una collezione *arbitraria* di insiemi chiusi è un insieme chiuso:

$$C_i \in \mathcal{C}, \quad i \in \mathcal{I} \implies \bigcap_{i \in \mathcal{I}} C_i \in \mathcal{C}, \quad (2.8)$$

con \mathcal{I} insieme di indici.

iii) L'unione di un numero *finito* di insiemi chiusi è un insieme chiuso:

$$C_i \in \mathcal{C}, \quad i = 1, 2, \dots, N \text{ (finito)} \implies \bigcup_{i=1}^N C_i \in \mathcal{C}. \quad (2.9)$$

Osservazione. Risulta evidente che la topologia poteva essere definita tramite i chiusi e le loro proprietà, definendo in seguito gli aperti come i loro complementari e deducendone le proprietà mediante le relazioni di dualità.

Riguardo alla necessità che per garantire la chiusura occorre limitare l'unione di chiusi ad un numero finito, si consideri il seguente esempio sulla retta reale. L'unione di infiniti intervalli chiusi del tipo:

$$\left\{x ; \frac{1}{n} \leq x \leq 1\right\}, \quad n \in \mathbb{N}, \quad n > 0,$$

risulta l'intervallo $0 < x \leq 1$, non chiuso.

Notiamo che non è proibito avere insiemi contemporaneamente sia aperti che chiusi. Nel caso della topologia discreta tutti i sottoinsiemi di X sono sia aperti che chiusi.

Esempio 2.5 Riprendendo l'esempio 2.4 i chiusi sono costituiti dagli insiemi:

$$\mathcal{C} = \{\emptyset, X, \{b, c\}, \{c\}\},$$

e si possono verificare direttamente le relative proprietà. Rimangono esclusi dalla famiglia degli aperti e da quella dei chiusi i seguenti sottoinsiemi:

$$\{b\}, \{a, c\}.$$

2.1.3 Topologia indotta e topologia generata

Una volta definita una topologia in un insieme X , un qualsiasi suo sottoinsieme $Y \subseteq X$ può essere considerato uno spazio topologico definendo in Y la **topologia indotta** da X in questo modo:

Def. 2.3 Un sottoinsieme $O \subseteq Y$ è **aperto in Y** se esiste un aperto A di X tale che $O = A \cap Y$.

e, come conseguenza, abbiamo che $W \subseteq Y$ è chiuso in Y se e solo se esiste un chiuso C di X tale che $W = C \cap Y$. È semplice verificare che gli insiemi aperti in Y così definiti verificano le condizioni degli aperti per cui Y diviene uno spazio topologico a tutti gli effetti.

Mediante il concetto di topologia indotta si vuole mettere in evidenza la compatibilità tra la topologia di Y e quella di X . Notiamo però che Y stesso risulta contemporaneamente sia aperto che chiuso ($Y = X \cap Y$) nella topologia indotta senza essere necessariamente aperto o chiuso nella topologia di X . Analogamente non possiamo affermare in generale che se O è aperto in Y esso risulti (come sottoinsieme di X) aperto anche nella topologia di X .

Supponiamo ora che su un medesimo insieme di supporto X non vuoto siano definite due topologie \mathcal{T}_1 e \mathcal{T}_2 , cioè due collezioni di insiemi che verificano le richieste della definizione 2.1. Risultano allora definiti due spazi topologici (X, \mathcal{T}_1) e (X, \mathcal{T}_2) . Si pone immediatamente il problema della confrontabilità delle due topologie.

Def. 2.4 Si dice che la topologia \mathcal{T}_2 è **più forte** o **più fine** di \mathcal{T}_1 se ogni aperto di \mathcal{T}_1 è anche un aperto di \mathcal{T}_2 . Possiamo anche dire in tal caso che \mathcal{T}_1 è **più debole** di \mathcal{T}_2 .

In pratica ciò significa che la topologia più forte ha un numero di aperti maggiore di quella più debole. Confrontando le collezioni di aperti abbiamo insiemisticamente $\mathcal{T}_1 \subseteq \mathcal{T}_2$. Non è detto in generale che due topologie siano confrontabili tra loro, comunque la collezione di tutte le topologie possibili in X risulta parzialmente ordinata in modo naturale (la topologia \mathcal{T}_1 precede \mathcal{T}_2 se \mathcal{T}_1 è più debole di \mathcal{T}_2 , cioè $\mathcal{T}_1 \subseteq \mathcal{T}_2$). Tale collezione di topologie possibili ha però un *elemento massimale*, identificabile con la topologia in cui tutti gli insiemi sono aperti (topologia discreta) che risulta ovviamente più forte di qualsiasi altra topologia, ed un *elemento minimale*, cioè la topologia banale in cui sono aperti solo il vuoto e X (topologia dei punti incollati).

Teo. 2.2 L'intersezione di un insieme qualsiasi di topologie:

$$\mathcal{T} = \bigcap_{\alpha} \mathcal{T}_{\alpha}$$

in X è ancora una topologia e risulta più debole di ciascuna topologia \mathcal{T}_{α} .

Dim. 2.2 È chiaro che l'intersezione $\bigcap_{\alpha} \mathcal{T}_{\alpha}$ contiene sia X che il vuoto \emptyset . Inoltre, dal fatto che ciascuna topologia \mathcal{T}_{α} è chiusa rispetto a unioni arbitrarie e ad intersezioni finite, risulta che anche l'intersezione \mathcal{T} gode di tali proprietà.

— ■ —

Come conseguenza abbiamo che se \mathcal{B} è una famiglia qualsiasi di sottoinsiemi di X allora esiste una topologia minimale in X contenente \mathcal{B} (l'esistenza di topologie contenenti \mathcal{B} è garantita dal fatto che nella topologia discreta tutti gli insiemi sono aperti), e tale topologia coincide con l'intersezione di tutte le topologie contenenti \mathcal{B} .

Def. 2.5 Sia \mathcal{B} una famiglia di insiemi in uno spazio X . L'intersezione di tutte le topologie contenenti \mathcal{B} è detta **topologia generata** dal sistema \mathcal{B} e la indicheremo con $\tau(\mathcal{B})$.

2.1.4 Punti interni, esterni, di frontiera

Mediante gli aperti è possibile caratterizzare i punti relativamente ad un insieme. Sia quindi X uno spazio topologico, A un suo sottoinsieme, e diamo le seguenti definizioni.

Def. 2.6 Un punto $x \in A$ è detto **interno** ad A se esiste un aperto O contenente x e tutto contenuto in A : $x \in O \subseteq A$. L'insieme dei punti interni ad A viene detto **interno** di A e indicato con $\text{Int}(A)$ oppure A° .

Def. 2.7 Un punto x è detto di **aderenza** ad A se per ogni aperto O contenente x si ha $O \cap A \neq \emptyset$. L'insieme dei punti di aderenza costituisce la **chiusura** di A e viene indicato con A^- .

Con tali definizioni si comincia a formalizzare la nozione di vicinanza. Intuitivamente infatti i punti interni sono quelli che hanno come vicini *solo* punti dell'insieme e sono separati, cioè *lontani*, dal complementare. Quelli di aderenza risultano invece i punti che hanno *vicino* a sè elementi dell'insieme. Notiamo che un punto di aderenza per A non deve necessariamente appartenere ad A .

Come conseguenza delle definizioni poste abbiamo che:

Teo. 2.3 $\text{Int}(A)$ rappresenta l'aperto più grande (massimale rispetto alla relazione d'ordine di inclusione insiemistica) contenuto in A , e coincide con l'unione di tutti gli aperti contenuti in A :

$$A^\circ = \bigcup_{O \subseteq A} O, \quad (O \text{ aperto}). \quad (2.10)$$

inoltre se A è aperto si ha $A = A^\circ$, cioè A è aperto se e solo se tutti i suoi punti sono interni.

La chiusura A^- risulta invece l'insieme chiuso più piccolo (minimale rispetto alla relazione d'ordine di inclusione insiemistica) contenente A , e coincide con l'intersezione di tutti i chiusi contenenti A :

$$A^- = \bigcap_{C \supseteq A} C, \quad (C \text{ chiuso}). \quad (2.11)$$

e se A è chiuso allora $A = A^-$, cioè A è chiuso se e solo se tutti i suoi punti di aderenza appartengono a A .

Dim. 2.3 La caratterizzazione dell'interno è immediata. Infatti se x è interno ad A ($x \in A^\circ$), allora esiste un aperto O contenente x e contenuto in A per cui x appartiene all'unione di tutti gli aperti contenuti in A . Viceversa, se x appartiene a tale unione ($x \in \bigcup_{O \subseteq A} O$), x apparterrà ad almeno un aperto contenuto in A , per cui x è interno ad A . Questo prova la relazione (2.10), ed in particolare A° risulta aperto. Chiaramente si ha sempre $A^\circ \subseteq A$, e se A è aperto allora esso stesso è un aperto contenuto in A e quindi contenuto in A° (equazione 2.10), per cui $A = A^\circ$.

Per quanto riguarda la caratterizzazione della chiusura, consideriamo un punto x di aderenza per A e sia C un qualsiasi chiuso contenente A . Se, per assurdo x non appartenesse a C , allora x apparterrebbe al complementare $X \setminus C$, aperto e ad intersezione vuota con A , e ciò contraddice l'ipotesi che x sia di aderenza. Viceversa se x appartiene a tutti i chiusi C contenenti A , consideriamo un aperto O contenente x . Se, per assurdo, fosse $O \cap A = \emptyset$, allora il complementare di O risulterebbe un chiuso contenente A ma non x , che contraddice l'ipotesi fatta su x . Pertanto $O \cap A \neq \emptyset$ e x è di aderenza. Ciò prova la relazione (2.11) e che A^- è un insieme chiuso. Chiaramente per ogni insieme A si ha $A \subseteq A^-$, e se A è chiuso egli stesso forma uno dei chiusi contenenti A , per cui (vedi l'equazione 2.11) $A^- \subseteq A$, e $A = A^-$.

— ■ —

Altre caratteristiche relative tra un punto ed un insieme danno luogo alle seguenti definizioni.

Def. 2.8 Sia A un insieme in uno spazio topologico X . Allora un punto $x \in X$ è detto di **accumulazione** per A se per ogni aperto O contenente x si ha $(O \setminus \{x\}) \cap A \neq \emptyset$. L'insieme dei punti di accumulazione per un insieme A viene detto **derivato** di A e lo indicheremo con $D(A)$.

Def. 2.9 Se $x \in A$ e x non è di accumulazione per A allora x è detto **punto isolato** di A . In caso contrario, se $x \in A$ e x è di accumulazione per A , viene detto **non isolato**.

Osservazione. Notiamo che un punto di accumulazione risulta anche di aderenza. L'unica distinzione nella differenza tra le definizioni di punto di aderenza e punto di accumulazione è di togliere x dall'aperto che lo contiene, per cui si prescinde dall'appartenenza o meno di x all'insieme. Un punto di accumulazione deve avere *vicino* punti dell'insieme diversi da sè stesso. Chiaramente aggiungendo ai punti di accumulazione i punti dell'insieme stesso otteniamo tutti i punti di aderenza per cui se A è un insieme possiamo dire:

$$A^- = A \bigcup D(A). \quad (2.12)$$

Inoltre possiamo dire che A è chiuso se contiene tutti i suoi punti di accumulazione:

$$A = A^- \iff D(A) \subseteq A. \quad (2.13)$$

Def. 2.10 Sia A un insieme in uno spazio topologico X . Allora un punto $x \in X$ è detto **esterno** per A se è interno al complementare di A . x è detto **di frontiera** per A se x non è né interno né esterno per A .

Mediante gli aperti si possono definire gli intorni di un punto.

Def. 2.11 Sia X uno spazio topologico e $x \in X$. Definiamo un **intorno** di x un qualsiasi insieme U che contiene un aperto O della topologia contenente il punto x : $x \in O = O^\circ \subseteq U$.

Sostanzialmente un intorno di x è un qualsiasi insieme contenente x al suo interno. Non imponiamo che un intorno sia aperto o chiuso, ma solo il fatto che un intorno debba contenere un aperto, cioè che x risulti un punto interno. Ciò permette spesso di poter usare gli intorni alla stessa maniera degli aperti, senza preoccuparsi del fatto che un intorno possa essere non aperto. Nelle definizioni precedenti avremmo potuto usare intorni al posto di aperti. Potremmo ad esempio definire che un punto è interno ad un insieme se esiste un intorno del punto tutto contenuto nell'insieme. Quando sarà necessario si specificherà che l'intorno è aperto, e in tal caso si indicherà in pratica un aperto contenente il punto.

2.1.5 Sistemi fondamentali di intorni

Come si è visto, assegnare una topologia nello spazio X vuol dire assegnarvi un sistema di insiemi aperti. Nei problemi concreti è comodo assegnare non tutta la topologia ma solo una famiglia di insiemi aperti in base alla quale si determina la collezione di tutti gli insiemi aperti. Questo porta alla definizione di base di uno spazio topologico come una famiglia \mathcal{B} di insiemi aperti tale che ogni aperto di X possa esprimersi come unione di insiemi di \mathcal{B} . Volendo essere più formali possiamo dare la seguente definizione equivalente:

Def. 2.12 Una famiglia \mathcal{B} di insiemi è una **base per una topologia** su X se valgono le seguenti condizioni:

i) Il vuoto è un elemento della base:

$$\emptyset \in \mathcal{B}. \quad (2.14)$$

ii) Il supporto X si ottiene dalla base:

$$\bigcup_{B \in \mathcal{B}} B = X. \quad (2.15)$$

iii) Se B_1 e B_2 appartengono alla famiglia \mathcal{B} , la loro intersezione si ottiene dalla base stessa:

$$B_1 \cap B_2 = \bigcup_{B \in \overline{\mathcal{B}}} B \quad (2.16)$$

per un opportuno sottoinsieme $\overline{\mathcal{B}} \subseteq \mathcal{B}$.

Le basi di una topologia costituiscono spesso lo strumento principe per considerare proprietà relative ai singoli punti $x \in X$, per cui forniscono dei sistemi fondamentali di intorni per lo studio delle proprietà topologiche.

Mostriamo ora che la definizione è ben posta, cioè che una base secondo la definizione 2.12 definisce effettivamente una topologia.

Teo. 2.4 Sia X un insieme e \mathcal{B} una base per una topologia su X . Poniamo

$$\mathcal{U}_{\mathcal{B}} = \{U \subseteq X ; U \text{ unione di elementi di } \mathcal{B}\},$$

cioè:

$$U \in \mathcal{U}_{\mathcal{B}} \iff U = \bigcup_{B \in \mathcal{B}_U} B, \quad (2.17)$$

con $\mathcal{B}_U \subseteq \mathcal{B}$ collezione di insiemi della base. Allora $\mathcal{U}_{\mathcal{B}}$ è una topologia su X , coincidente con la topologia $\tau(\mathcal{B})$ generata da \mathcal{B} :

$$\tau(\mathcal{B}) = \mathcal{U}_{\mathcal{B}}. \quad (2.18)$$

Dim. 2.4 L'insieme $\mathcal{U}_{\mathcal{B}}$ definisce una topologia su X . Infatti:

– Il vuoto $\emptyset \in \mathcal{B}$, per cui $\emptyset \in \mathcal{U}_{\mathcal{B}}$, mentre, per la proprietà ii) della definizione 2.12 di base, X stesso risulta unione di elementi di \mathcal{B} , per cui $X \in \mathcal{U}_{\mathcal{B}}$.

– Se $\mathcal{V} \subseteq \mathcal{U}_{\mathcal{B}}$ allora risulta

$$\bigcup_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V}} \left(\bigcup_{B \in \mathcal{B}_V} B \right) = \bigcup_{B \in \tilde{\mathcal{B}}} B,$$

con \mathcal{B}_V collezione di insiemi della base che compongono V , e:

$$\tilde{\mathcal{B}} = \bigcup_{V \in \mathcal{V}} \mathcal{B}_V,$$

per cui la proprietà di chiusura rispetto all'unione è verificata.

– La proprietà di chiusura rispetto alla intersezione finita si dimostra facilmente per induzione utilizzando la proprietà iii) della definizione 2.12 di base.

Risulta inoltre evidente che $\tau(\mathcal{B}) = \mathcal{U}_{\mathcal{B}}$. Infatti, essendo $\mathcal{U}_{\mathcal{B}}$ una topologia contenente gli insiemi di \mathcal{B} , è sicuramente più fine della topologia minimale $\tau(\mathcal{B})$: $\tau(\mathcal{B}) \subseteq \mathcal{U}_{\mathcal{B}}$. D'altra parte, se $U \in \mathcal{U}_{\mathcal{B}}$ allora U è unione di elementi di \mathcal{B} e deve appartenere a qualsiasi topologia contenente \mathcal{B} , per cui $\mathcal{U}_{\mathcal{B}} \subseteq \tau(\mathcal{B})$.

— ■ —

Sostanzialmente una topologia può essere definita specificando una famiglia di insiemi che verificano la definizione 2.12. Mediante tali insiemi di base, che risultano ovviamente aperti, costruiamo qualsiasi altro aperto. Ad esempio sulla retta reale gli intervalli aperti $]a, b[= \{x \in \mathbb{R}; a < x < b\}$ costituiscono la base naturale per la topologia ordinaria dei numeri reali.

Se abbiamo due basi $\mathcal{B}_1, \mathcal{B}_2$ per due topologie su X allora le due basi si dicono **equivalenti** se generano la stessa topologia $\mathcal{U}_{\mathcal{B}_1} = \mathcal{U}_{\mathcal{B}_2}$.

Teo. 2.5 Sia X un insieme e $\mathcal{B}_1, \mathcal{B}_2$ basi per due topologie su X . Le due basi sono equivalenti se e solo se:

- i) Per ogni $x \in X$ ed ogni $B_1 \in \mathcal{B}_1$ con $x \in B_1$ esiste un aperto $B_2 \in \mathcal{B}_2$ tale che $x \in B_2 \subseteq B_1$.
- ii) Per ogni $x \in X$ ed ogni $B_2 \in \mathcal{B}_2$ con $x \in B_2$ esiste un aperto $B_1 \in \mathcal{B}_1$ tale che $x \in B_1 \subseteq B_2$.

Dim. 2.5 Assumiamo dapprima come ipotesi che le due basi siano equivalenti e sia $x \in B_1 \in \mathcal{B}_1$. Allora $B_1 \in \mathcal{U}_{\mathcal{B}_1} = \mathcal{U}_{\mathcal{B}_2}$, per cui:

$$B_1 = \bigcup_{B_2 \in \tilde{\mathcal{B}}_2} B_2,$$

per un opportuno sottoinsieme $\tilde{\mathcal{B}}_2 \subseteq \mathcal{B}_2$, per cui, dato $x \in B_1$, possiamo selezionare un elemento $B_2(x) \in \tilde{\mathcal{B}}_2$ che verifica la proprietà i) del teorema. Analogamente si dimostra la proprietà ii).

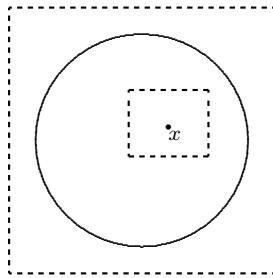


Figura 2.1: Equivalenza di topologie.

Viceversa supponiamo valide le proprietà i) e ii) e sia $B_1 \in \mathcal{B}_1$. Allora, per la proprietà i), per ogni $x \in B_1$ esiste $B_2(x) \in \mathcal{B}_2$ tale che $B_2(x) \subseteq B_1$. Possiamo dire che:

$$B_1 \subseteq \bigcup_{x \in B_1} B_2(x) \subseteq B_1,$$

cioè B_1 è unione di insiemi della base \mathcal{B}_2 :

$$B_1 = \bigcup_{x \in B_1} B_2(x),$$

da cui $B_1 \in \mathcal{U}_{\mathcal{B}_2}$ e $\mathcal{U}_{\mathcal{B}_1} \subseteq \mathcal{U}_{\mathcal{B}_2}$. Nello stesso modo la proprietà ii) del teorema implica che $\mathcal{U}_{\mathcal{B}_2} \subseteq \mathcal{U}_{\mathcal{B}_1}$ e quindi $\mathcal{U}_{\mathcal{B}_1} = \mathcal{U}_{\mathcal{B}_2}$.

— ■ —

Esempio 2.6 In \mathbb{R}^2 possiamo definire una base come la collezione di dischi aperti:

$$\{(x, y) \in \mathbb{R}^2 ; (x - x_0)^2 + (y - y_0)^2 < r^2\},$$

oppure possiamo definire una base tramite i rettangoli aperti:

$$\{(x, y) \in \mathbb{R}^2 ; x_1 < x < x_2, y_1 < y < y_2\}.$$

Le due basi sono evidentemente equivalenti in quanto ogni cerchio lo possiamo pensare interno ad un rettangolo ed ogni rettangolo lo possiamo pensare incluso all'interno di una circonferenza (vedi figura 2.1).

— ○ —

2.2 Spazi metrici.

Vediamo ora un caso importante di spazi topologici, i cosiddetti spazi metrici, nei quali ha senso parlare di distanza tra i punti.

Def. 2.13 Sia X un insieme in cui sia definita una applicazione:

$$d : X \times X \longrightarrow \mathbb{R},$$

che goda delle seguenti proprietà ($x, y, z \in X$):

i) È una funzione positiva:

$$0 \leq d(x, y) < \infty. \quad (2.19)$$

ii) È strettamente definita positiva:

$$d(x, y) = 0 \iff x = y. \quad (2.20)$$

iii) È una funzione simmetrica:

$$d(x, y) = d(y, x). \quad (2.21)$$

iv) Verifica la diseguaglianza triangolare:

$$d(x, y) \leq d(x, z) + d(z, y). \quad (2.22)$$

Allora l'applicazione d è detta **distanza** e si dice che nello spazio X è definita una **metrica**. La coppia (X, d) definisce uno **spazio metrico**, o più semplicemente che X è uno **spazio metrico**.

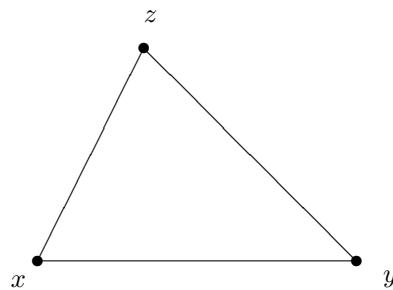


Figura 2.2: Diseguaglianza triangolare.

Esempio 2.7 Nello spazio \mathbb{R}^n la distanza usuale è definita da:

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}. \quad (2.23)$$

Una distanza alternativa può essere invece:

$$d(x, y) = \max_{1 \leq j \leq n} |x_j - y_j|. \quad (2.24)$$

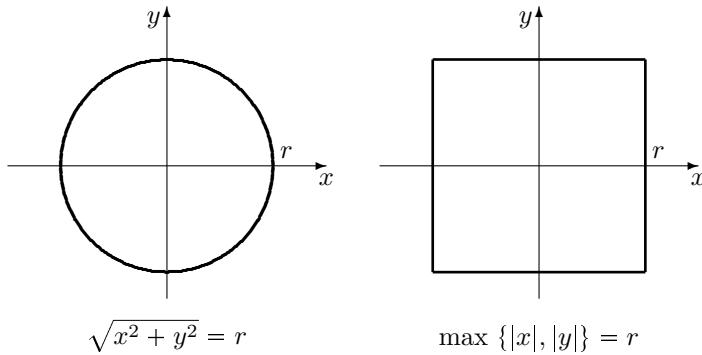


Figura 2.3: Distanze in \mathbb{R}^2 .

Notiamo che con la distanza (2.23) in \mathbb{R}^2 abbiamo l'usuale rappresentazione geometrica delle circonferenze di dato raggio, mentre la distanza (2.24) da luogo a “circonferenze” di forma quadrata coi bordi paralleli agli assi coordinati (figura 2.3).

— ○ —

Esempio 2.8 Sia X un insieme qualsiasi e consideriamo l'insieme $\mathcal{F}(X)$ delle funzioni limitate su X :

$$\mathcal{F}(X) = \left\{ f : X \longrightarrow \mathbb{R}; \sup_{x \in X} |f(x)| < \infty \right\}, \quad (2.25)$$

e definiamo in $\mathcal{F}(X)$ la distanza:

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|, \quad (2.26)$$

(la condizione di limitatezza imposta nella costruzione di $\mathcal{F}(X)$ ci garantisce l'esistenza di tale estremo superiore per ogni scelta di f e g). L'unica proprietà non immediata è la diseguaglianza triangolare. Dalle proprietà del valore assoluto in \mathbb{R} sappiamo che:

$$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)|$$

per ogni $x \in X$, per cui considerando l'estremo superiore al variare di $x \in X$ otteniamo la diseguaglianza triangolare per la distanza in $\mathcal{F}(X)$:

$$\sup_{x \in X} |f(x) - g(x)| \leq \sup_{x \in X} |f(x) - h(x)| + \sup_{x \in X} |h(x) - g(x)|.$$

— ○ —

Esempio 2.9 Sia X un insieme arbitrario (non vuoto) e definiamo in X la metrica **discreta**:

$$d(x, y) = \begin{cases} 1 & \text{se } x \neq y, \\ 0 & \text{se } x = y. \end{cases} \quad (2.27)$$

Questa definizione verifica banalmente le proprietà di distanza, ma in pratica distingue solamente i punti tra di loro.

— ○ —

Notiamo che dalla proprietà iv) della definizione 2.13 abbiamo anche:

$$|d(x, z) - d(z, y)| \leq d(x, y) \quad \forall x, y, z. \quad (2.28)$$

Infatti:

$$\begin{aligned} d(x, z) \leq d(x, y) + d(y, z) &\implies d(x, z) - d(z, y) \leq d(x, y), \\ d(z, y) \leq d(z, x) + d(x, y) &\implies d(z, y) - d(x, z) \leq d(x, y), \end{aligned}$$

e si formalizza un risultato di base della geometria: “ogni lato di un triangolo è minore della somma e maggiore della differenza degli altri due lati”.

La proprietà dei triangoli iv) può essere estesa per induzione a sequenze (finite) di punti nello spazio (cioè a un poligono):

$$\begin{aligned} d(x, y) \leq \sum_{j=0}^n d(x_j, x_{j+1}), \quad \text{con } x_0 = x, \quad x_{n+1} = y \\ \text{e } x_k \text{ arbitrari, } k = 1, \dots, n. \end{aligned}$$

Vediamo ora come costruire una topologia tramite la distanza. Definiamo la sfera (aperta) centrata in un punto x e raggio r , e tramite le sfere, definiamo gli insiemi aperti:

Def. 2.14 Sia $x \in X$ con X spazio metrico, r un numero reale (positivo). Definiamo la **sfera (aperta) di centro x e raggio r** l’insieme:

$$S(x, r) = \{y \in X : d(x, y) < r\}, \quad (2.29)$$

e diciamo che un sottoinsieme $O \subseteq X$ (metrico) è **aperto** se risulta vuoto oppure unione di sfere del tipo $S(x, r)$:

$$O = \bigcup_{\alpha \in \mathcal{A}} S(x_\alpha, r_\alpha), \quad (2.30)$$

con \mathcal{A} opportuno insieme di indici.

Osservazione. Abbiamo ristretto i possibili raggi a numeri positivi. Potevamo anche lasciare arbitrario il raggio, facendo rientrare nella definizione di sfera anche l'insieme vuoto, ma vogliamo garantirci che quando diremo “esiste una sfera” questa sia non vuota (conterrà almeno il suo centro) ed abbia un raggio finito e positivo. Questo implica l'inserzione esplicita del vuoto stesso nella costruzione della topologia costruita con le sfere.

Occorre mostrare che la definizione posta ha senso, cioè che la famiglia di aperti di cui sopra definisce una topologia. Ovviamente l'insieme vuoto \emptyset risulta aperto mentre X stesso si può ottenere considerando l'unione di tutte le possibili sfere con un raggio prefissato (ad esempio 1):

$$X = \bigcup_{x \in X} S(x, 1).$$

Se consideriamo una collezione arbitraria di aperti O_α , essendo ognuno di essi una unione di sfere, anche la loro unione è ancora una unione di sfere, per cui la proprietà dell'unione è banalmente soddisfatta.

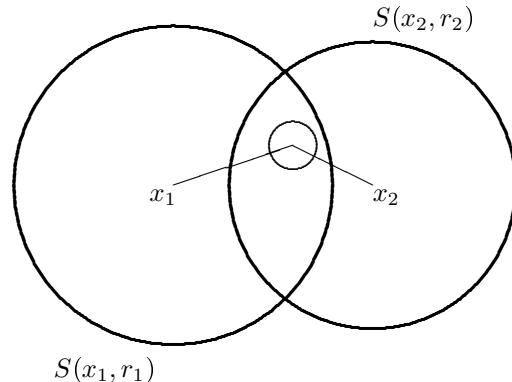


Figura 2.4: $S_1 \cap S_2 = \bigcup_{x \in S_1 \cap S_2} S(x, \rho(x))$.

Per mostrare la proprietà dell'intersezione cominciamo col mostrare che l'intersezione di due sfere è un aperto (cioè a sua volta unione di sfere). Siano $S_1 = S(x_1, r_1)$, $S_2 = S(x_2, r_2)$ due sfere a intersezione non vuota (in caso contrario abbiamo che il vuoto è un insieme aperto), e sia $x \in S_1 \cap S_2$, allora $d(x, x_1) < r_1$, $d(x, x_2) < r_2$, ed esiste un numero positivo $\rho(x)$ tale che

$$d(x, x_1) + \rho(x) < r_1,$$

$$d(x, x_2) + \rho(x) < r_2,$$

allora tutta la sfera $S(x, \rho(x))$ è sicuramente contenuta nell'intersezione $S_1 \cap S_2$, infatti, se $z \in S(x, \rho(x))$:

$$\begin{aligned} d(x_1, z) &\leq d(x_1, x) + d(x, z) < d(x_1, x) + \rho(x) < r_1, \\ d(x_2, z) &\leq d(x_2, x) + d(x, z) < d(x_2, x) + \rho(x) < r_2. \end{aligned}$$

L'intersezione risulta allora unione delle sfere così costruite:

$$S_1 \cap S_2 = \bigcup_{x \in S_1 \cap S_2} S(x, \rho(x))$$

ed è quindi un aperto secondo la definizione data.

Supponiamo ora di avere un numero finito di aperti O_j , $j = 1, \dots, N$, ognuno dei quali costruito come unione di sfere:

$$O_j = \bigcup_{\alpha \in \mathcal{A}_j} S(x_\alpha, r_\alpha).$$

Abbiamo

$$\begin{aligned} O_1 \cap O_2 &= \left(\bigcup_{\alpha_1 \in \mathcal{A}_1} S(x_{\alpha_1}, r_{\alpha_1}) \right) \cap \left(\bigcup_{\alpha_2 \in \mathcal{A}_2} S(x_{\alpha_2}, r_{\alpha_2}) \right) \\ &= \bigcup_{\alpha_1, \alpha_2} \left(S(x_{\alpha_1}, r_{\alpha_1}) \cap S(x_{\alpha_2}, r_{\alpha_2}) \right), \end{aligned}$$

ma abbiamo visto che l'intersezione di due sfere è un aperto, per cui anche l'intersezione $O_1 \cap O_2$ risulta un aperto. Per induzione otteniamo che l'intersezione di un numero finito N di aperti è ancora un aperto.

In pratica abbiamo verificato che le sfere aperte costituiscono una base per la topologia di uno spazio metrico. Potremmo ridefinire un insieme aperto se per ogni suo punto x esiste una sfera aperta centrata nel punto e tutta contenuta nell'insieme. Infatti se x appartiene ad un aperto O come definito sopra, allora x apparterrà ad una sfera centrata in un punto y e raggio r e contenuta nell'aperto O :

$$d(x, y) < r,$$

allora scelto un numero reale r_1 , positivo e inferiore a $r - d(x, y)$:

$$r_1 < r - d(x, y),$$

tutta la sfera $S(x, r_1)$ risulta contenuta in $S(y, r)$ e nell'insieme O .

Con la definizione di distanza ordinaria (esempio 2.7) tra elementi in \mathbb{R}^n ritroviamo l'usuale topologia degli aperti in \mathbb{R}^n , mentre con la distanza discreta (esempio 2.9) determiniamo la topologia discreta.

Abbiamo puntualizzato nella definizione di sfera (2.29) come questa sia aperta (ancora prima di definire gli aperti) per distinguerla dalla definizione di **sfera chiusa**:

$$S_c(x, r) = \{y \in X : d(x, y) \leq r\}, \quad (2.31)$$

che risulta un insieme chiuso. Sia infatti z appartenente al complementare di $S_c(x, r)$; allora $d(z, x) > r$, e, scelto un numero reale positivo $r_1 < d(z, x) - r$, la sfera (aperta) di raggio r_1 e centro z è tutta contenuta nel complementare di $S_c(x, r)$. Infatti se $y \in S(z, r_1)$ abbiamo

$$\begin{aligned} d(z, x) &\leq d(y, x) + d(z, y) < d(y, x) + r_1, \\ \implies d(y, x) &> d(z, x) - r_1 > r \implies y \notin S_c(x, r). \end{aligned}$$

Esempio 2.10 Notiamo che possono esistere delle metriche in cui vi sono degli insiemi che sono contemporaneamente sia aperti che chiusi. Consideriamo ad esempio la metrica discreta dell'esempio 2.9. Abbiamo

$$\begin{aligned} S(x, r) &= \begin{cases} X & r > 1 \\ \{x\} & r \leq 1 \end{cases} \\ S_c(x, r) &= \begin{cases} X & r \geq 1 \\ \{x\} & r < 1 \end{cases} \end{aligned}$$

e qualsiasi insieme in X risulta contemporaneamente aperto e chiuso (essendo il suo complementare aperto).

Tale esempio mostra anche che non è detto che $S(x, r)^- = S_c(x, r)$, oppure che $S_c(x, r)^\circ = S(x, r)$. Infatti con la metrica discreta abbiamo:

$$\begin{aligned} S(x, 1) &= \{x\} \\ S_c(x, 1) &= X \\ S(x, 1)^- &= \{x\} \\ S_c(x, 1)^\circ &= X \end{aligned}$$

— ○ —

La metrica permette di dare un significato al concetto di insieme limitato. Infatti possiamo definire un insieme come **limitato** se esiste una sfera (di raggio finito) che contiene tutto l'insieme.

2.3 Proprietà topologiche

Vediamo quali proprietà possono presentare i possibili sottoinsiemi o tutto lo spazio in conseguenza della scelta di una topologia.

2.3.1 Insiemi compatti

Def. 2.15 Sia Y un insieme di uno spazio topologico X . Una collezione di insiemi $\mathcal{R} = \{A_\alpha, \alpha \in \mathcal{A}\}$ è detto un **ricoprimento** di Y se

$$Y \subseteq \bigcup_{\alpha \in \mathcal{A}} A_\alpha,$$

e il ricoprimento è detto **finito** se l'insieme degli indici \mathcal{A} è finito.

Se il ricoprimento è formato con insiemi aperti in X allora il ricoprimento è detto **ricoprimento aperto**. Analogamente possiamo definire un **ricoprimento chiuso**.

Se $\mathcal{R} = \{A_\alpha, \alpha \in \mathcal{A}\}$ è un ricoprimento di Y , $\mathcal{R}' \subseteq \mathcal{R}$, e \mathcal{R}' è esso stesso un ricoprimento di Y , si dice che \mathcal{R}' è un **sottoricoprimento**.

Def. 2.16 Sia X uno spazio topologico e $Y \subseteq X$. Y è detto **compatto** se da ogni ricoprimento aperto:

$$Y \subseteq \bigcup_{\alpha} O_\alpha, \quad O_\alpha \text{ aperto},$$

è possibile estrarre un sottoricoprimento finito:

$$Y \subseteq \bigcup_{i=1}^N O_{\alpha_i}.$$

In particolare se questo è vero per tutto lo spazio topologico X diremo che X è uno **spazio topologico compatto**. In questo caso l'inclusione sopra è in realtà un'uguaglianza (non abbiamo sovrabbondanza):

$$X = \bigcup_{\alpha} O_\alpha \implies X = \bigcup_{i=1}^N O_{\alpha_i}, \quad O_\alpha \text{ aperto}.$$

Abbiamo immediatamente il seguente risultato:

Teo. 2.6 Sia X uno spazio topologico compatto. Allora ogni sottoinsieme chiuso C di X è compatto.

Dim. 2.6 Supponiamo di avere un ricoprimento aperto di C :

$$C \subseteq \bigcup_{\alpha} O_{\alpha},$$

gli insiemi O_{α} , assieme al complementare di C , $X \setminus C$, formano un ricoprimento aperto di X . È allora possibile estrarre un sottoricoprimento finito (che non è restrittivo assumere che contenga anche $X \setminus C$) di X :

$$X = \bigcup_{i=1}^N (O_{\alpha_i}) \bigcup (X \setminus C).$$

Eliminando $X \setminus C$ (sicuramente C non è sottoinsieme del suo complementare) otteniamo un ricoprimento finito di C .

Il viceversa (cioè che un insieme compatto sia anche chiuso) non è immediato e vedremo in seguito sotto quali condizioni verificare ciò.

— ■ —

2.3.2 Spazi di Hausdorff

Osserviamo ora che se siamo in uno spazio metrico M , possiamo definire una topologia, ma la struttura dello spazio risulta molto più ricca di un semplice spazio topologico. Ad esempio, grazie alla nozione di distanza, possiamo sempre separare nettamente tra loro i punti. Se abbiamo $x, y \in M$ con $x \neq y$ allora esistono due sfere aperte $S(x, \epsilon), S(y, \eta)$ a intersezione vuota:

$$S(x, \epsilon) \bigcap S(y, \eta) = \emptyset.$$

Essendo i punti distinti la loro distanza è non nulla, $d(x, y) > 0$, e basta allora scegliere ϵ e η positivi e tali che $\epsilon + \eta < d(x, y)$. In questo modo è impossibile trovare un punto z appartenente all'intersezione delle due sfere perché in tal caso si avrebbe:

$$d(x, y) \leq d(x, z) + d(z, y) < \epsilon + \eta,$$

in contraddizione con la scelta operata.

Questa semplice proprietà di separazione tra i punti può essere generalizzata come una proprietà (individuata da Hausdorff¹) che uno spazio topologico può o non può avere.

¹ **Felix Hausdorff** (Breslau, Germania, 8 novembre 1868 – Bonn, Germania, 26 gennaio 1942) è stato un matematico tedesco. Si laureò presso l'Università di Leipzig e ottenne il dottorato di ricerca nel 1891. Ivi insegnò matematica fino al 1910, finché non ottenne la stessa cattedra all'Università di Bonn. In questo periodo fornì grandi contributi alla teoria dei gruppi. Tra i fondatori della topologia moderna, contribuì in modo significativo anche alla teoria degli insiemi e all'analisi funzionale. Definì e studiò gli insiemi parzialmente ordinati, gli spazi di Hausdorff e le dimensioni di Hausdorff. Elaborò il principio di massimalità di Hausdorff e la soluzione a ciò che ora viene definito il problema del momento di Hausdorff. Pubblicò anche testi letterari e filosofici con lo pseudonimo di “Paul Mongré”.



Def. 2.17 (Assioma di Hausdorff) Uno spazio topologico X è detto **spazio separato** o **spazio di Hausdorff**, se per ogni $x, y \in X$ con $x \neq y$ esistono un intorno O_x di x e un intorno O_y di y disgiunti tra loro:

$$O_x \cap O_y = \emptyset, \quad x \in O_x, \quad y \in O_y.$$

Vediamo con qualche esempio che la proprietà di Hausdorff può non essere verificata da una scelta di topologia.

Esempio 2.11 Sulla retta reale definiamo come aperti le semirette:

$$S_y = \{x \in \mathbb{R} ; x > y\}.$$

Le proprietà degli aperti sono banalmente verificate, ma non è possibile trovare due aperti disgiunti. Se consideriamo due punti distinti x, y , con $x < y$ ad esempio, scelto z in posizione intermedia, $x < z < y$, allora y appartiene a S_z , intorno aperto di y non contenente x (ma non possiamo trovare un intorno aperto di x che non contenga y).

Questo esempio mostra tuttavia una debole separazione tra i punti, nel senso che, dati due punti distinti x, y , è possibile trovare un intorno di x che non contiene y , oppure un intorno di y che non contiene x . Si parla in questo caso di uno spazio topologico di Kolmogorov².

— ○ —

Esempio 2.12 Sia X un insieme infinito e definiamo gli aperti come la famiglia:

$$\mathcal{O} = \{\emptyset, X, (X \setminus F), \text{ con } F \text{ sottoinsieme finito di } X\},$$

cioè un insieme è aperto se si ottiene da X togliendo un numero finito di punti. Potremmo dire che abbiamo definito un insieme chiuso come un insieme discreto formato da un numero finito di punti. Chiaramente sono verificate le proprietà dei chiusi e gli aperti, loro complementari, definiscono una topologia.

In tale topologia, dati due punti distinti x e y , è possibile determinare un intorno di x che non contiene y , e un intorno di y che non contiene x (proprietà che definisce uno

² Andrej Nikolaevič Kolmogorov (Tambov, 25 aprile 1903 – Mosca, 20 ottobre 1987) è stato un matematico russo che compì importanti progressi in diversi campi accademici: tra questi la teoria delle probabilità, la topologia, la logica intuizionista, la turbolenza, la meccanica classica e la complessità computazionale.



spazio topologico di Frèchet³): $O_x = X \setminus \{y\}$, $O_y = X \setminus \{x\}$. Non è possibile però fare in modo che tali intorni siano disgiunti.

— ○ —

Osservazione. In uno spazio topologico di Frèchet un punto singolo costituisce un insieme chiuso. Infatti se $x \neq y$ esiste un intorno O_y di y non contenente x , per cui O_y è tutto contenuto nel complementare dell'insieme $\{x\}$. Tale complementare è quindi aperto (ogni punto è interno). Ricordiamo che un intorno di un punto deve contenere un aperto contenente il punto. Vale anche il viceversa, cioè se la topologia è tale che l'insieme costituito da un solo punto $\{x\}$ è chiuso per ogni x allora vale l'assioma di Fréchet. Infatti se $y \neq x$ il complementare di $\{x\}$ è un intorno di y non contenente x e il complementare di $\{y\}$ è un intorno di x non contenente y .

La proprietà di Hausdorff, che per quanto visto sopra è verificata dagli spazi metrici, garantisce il seguente risultato che completa sostanzialmente il teorema 2.6.

Teo. 2.7 *Sia X uno spazio di Hausdorff, allora ogni sottoinsieme K compatto risulta anche chiuso.*

Dim. 2.7 Basta mostrare che il complementare di K è aperto, cioè che per ogni $x \notin K$ esiste un aperto (o un intorno) U_x contenente x e a intersezione vuota con K .

Essendo X di Hausdorff per ogni $y \in K$ esistono due insiemi aperti e disgiunti U_y e V_y contenenti x e y rispettivamente (x lo teniamo fisso e costruiamo gli aperti U_y e V_y al variare di $y \in K$):

$$U_y \cap V_y = \emptyset.$$

La famiglia di insiemi $\{V_y ; y \in K\}$ è un ricoprimento aperto di K (non contenente x). K è compatto per cui possiamo considerare un sottoricoprimento finito $\{V_{y_i} ; i = 1, \dots, n\}$ di K non contenente x :

$$K \subseteq \bigcup_{i=1}^n V_{y_i}.$$

Considerando i corrispondenti insiemi U_{y_i} , $i = 1, \dots, n$, e la loro intersezione:

$$U_x = \bigcap_{i=1}^n U_{y_i},$$

³ **Maurice René Frèchet** (Maligny, 2 settembre 1878 – Parigi, 4 giugno 1973) è stato un matematico francese. Diede fondamentali contributi in ambito topologico e nello studio di spazi astratti. Compì i suoi studi all'École di Parigi, conseguendo ottimi risultati in ambito accademico e stupendo gli eruditi dell'epoca con le sue brillanti ed ardite congetture. Porta il suo nome la Variabile casuale di Fréchet.



U_x risulta un aperto contenente x e a intersezione nulla con K :

$$U_x \cap K \subseteq U_x \cap \left(\bigcup_i V_{y_i} \right) = \emptyset.$$

Per l'arbitrarietà di $x \notin K$, K risulta chiuso.

— ■ —

Da tale dimostrazione, considerando gli insiemi aperti U_x (contenente x) e il ricoprimento di K :

$$V = \bigcup_{i=1}^n V_{y_i},$$

aperto e chiaramente disgiunto da U_x possiamo dedurre anche il seguente risultato.

Teo. 2.8 Sia X uno spazio di Hausdorff, allora per ogni punto x e compatto K disgiunti, cioè $x \notin K$, esistono due aperti U_x, V disgiunti, contenenti x e K rispettivamente:

$$U_x \cap V = \emptyset, \quad x \in U_x, \quad K \subseteq V.$$

Questa separazione può essere estesa a coppie di compatti:

Teo. 2.9 Sia X uno spazio di Hausdorff, allora dati due compatti disgiunti H, K , $H \cap K = \emptyset$, esistono due aperti U e V disgiunti e contenenti H e K rispettivamente:

$$U \cap V = \emptyset, \quad H \subseteq U, \quad K \subseteq V.$$

Dim. 2.9 Per ogni $x \in H$ possiamo trovare U_x e V_x aperti tali che (teorema 2.8):

$$x \in U_x, \quad K \subseteq V_x, \quad U_x \cap V_x = \emptyset.$$

La collezione degli insiemi U_x al variare di x in H è quindi un ricoprimento per H da cui possiamo estrarre un sottoricoprimento finito:

$$H \subseteq \bigcup_{j=1}^n U_{x_j} = U.$$

Considerando ora i corrispondenti aperti V_{x_j} , ognuno contenente K , abbiamo che la loro intersezione:

$$V = \bigcap_{j=1}^n V_{x_j},$$

è disgiunta da U e contiene K .

— ■ —

Siamo ora in grado di mostrare un importante risultato relativo agli spazi metrici a dimensioni finite:

Teo. 2.10 *Un insieme in \mathbb{R}^n è compatto se e solo se è chiuso e limitato.*

Osservazione. La nozione di limitatezza è legata alla nozione di distanza e un insieme può essere limitato in una metrica ma non in un'altra. La limitatezza è una nozione metrica e non topologica.

Dim. 2.10 Consideriamo un compatto K in \mathbb{R}^n . Sappiamo che \mathbb{R}^n è uno spazio metrico con la distanza ordinaria (esempio 2.7) per cui risulta uno spazio topologico separato. In uno spazio di Hausdorff ogni insieme compatto risulta anche chiuso (teorema 2.7). Occorre mostrare che è anche limitato.

Supponiamo per assurdo che non sia limitato, allora per ogni m intero positivo possiamo trovare un elemento $x_m \in K$ con $d(0, x_m) = |x_m| \geq m$. Consideriamo l'insieme $S = \{x_m, m = 1, 2, \dots\}$. Esso è un insieme chiuso in quanto il suo complementare è aperto: infatti all'interno di ogni sfera (di raggio perciò finito) centrata in un punto non appartenente a S possiamo trovare al più un numero finito di punti di S e quindi possiamo restringere ulteriormente la sfera in modo che non contenga alcun punto di S . Il centro della sfera risulta quindi un punto interno al complementare di S . Essendo S chiuso e sottoinsieme del compatto K , anche S è compatto. Se ora consideriamo il ricoprimento di S formato mediante sfere centrate nei punti di S , e ognuna con un proprio raggio ϵ_j :

$$S \subseteq \bigcup_{x_j \in S} S(x_j, \epsilon_j),$$

ogni sfera potrà contenere al massimo un numero finito di punti di S e quindi non è possibile ricoprire S con un numero finito di tali sfere. Questo assurdo porta a concludere che K deve essere limitato, oltre che chiuso.

Supponiamo ora di avere un insieme $K \subseteq \mathbb{R}^n$ chiuso e limitato e vediamo che deve essere compatto. Intanto se è limitato deve essere contenuto in un poli-intervallo (o poli-rettangolo) chiuso:

$$K \subseteq B = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n ; a_j \leq x_j \leq b_j\}.$$

Se mostriamo che B è compatto anche K , essendo chiuso (in \mathbb{R}^n e quindi anche in B secondo la topologia indotta su B), è compatto.

Sia quindi:

$$\mathcal{O} = \bigcup_{\alpha} O_{\alpha},$$

un ricoprimento arbitrario di B , e supponiamo per assurdo che non esista un sottoricoprimento finito. Bisecando ogni spigolo del polirettangolo (siamo in n dimensioni) possiamo

costruire 2^n sottorettangoli di cui almeno uno di questi, B_1 , non è ricoperto da un sottoricoprimento finito. Siano $a_j^{(1)}, b_j^{(1)}$ le coordinate degli estremi di tale sottorettangolo ($j = 1, \dots, n$), allora:

$$a_j \leq a_j^{(1)} < b_j^{(1)} \leq b_j ; \quad b_j^{(1)} - a_j^{(1)} = \frac{b_j - a_j}{2} ,$$

ed è chiaro che $B_1 \subseteq B$. Possiamo reiterare il ragionamento bisecando ulteriormente B_1 , costruendo B_2 non compatto e di lati con ampiezza:

$$b_j^{(2)} - a_j^{(2)} = \frac{b_j - a_j}{2^2} .$$

Proseguendo su questa via possiamo costruire una successione decrescente (secondo l'inclusione) di polirettangoli $B_m \supseteq B_{m+1}$:

$$B_m = \{x \in \mathbb{R}^n ; a_j^{(m)} \leq x_j \leq b_j^{(m)}\} , \quad b_j^{(m)} - a_j^{(m)} = \frac{b_j - a_j}{2^m} ,$$

di dimensioni sempre più ridotte. In particolare otteniamo una successione $a_j^{(m)}$, $m = 1, 2, \dots$, monotona crescente (e limitata da b_j), ed una successione monotona decrescente $b_j^{(m)}$, $m = 1, 2, \dots$, convergenti ad un certo c_j comune:

$$a_j^{(m)} \nearrow c_j \swarrow b_j^{(m)} , \quad a_j^{(m)} \leq c_j \leq b_j^{(m)} ,$$

Si viene quindi a definire un punto $c \in B$ che sarà contenuto in un aperto O_α del ricoprimento, e quindi in una sfera di raggio opportuno ϵ_α , $S(c, \epsilon_\alpha) \subseteq O_\alpha$. Tale sfera, essendo c il limite comune delle successioni di estremi, conterrà uno almeno (e quindi tutti i successivi) dei polirettangoli B_m , che risultano ricoperti da un solo O_α . Ciò contraddice il fatto che B_m non sia compatto, per cui B deve essere compatto.

— ■ —

2.3.3 Densità e separabilità.

Spesso risulta complicato dimostrare direttamente una proprietà per tutti gli elementi di un insieme ma può risultare più facile la verifica per “quasi tutti” gli elementi e poi trasportarla su tutto l’insieme. Il concetto di “quasi tutti” gli elementi può essere formalizzato nel concetto di densità.

Def. 2.18 Siano A e B due insiemi in uno spazio topologico X con $A \subseteq B$. Allora A è detto **denso** in B se $B \subseteq A^-$.

Sostanzialmente questo significa che per ogni punto $b \in B$ possiamo trovare un elemento $a \in A$ “vicino” topologicamente a b , cioè in ogni intorno di b possiamo trovare un elemento $a \in A$ (b è un punto di aderenza per A).

Spesso gli elementi su cui verificare una proprietà sono caratterizzati da interi, cioè costituiscono un insieme numerabile. Questo porta al concetto di separabilità (da non confondere con la “separazione” discussa in precedenza).

Def. 2.19 Uno spazio topologico X è detto **separabile** se esiste un sottoinsieme denso in X e numerabile:

$$\{x_n \in X, n = 1, 2, \dots\}^- = X. \quad (2.32)$$

Topologicamente abbiamo visto come le basi di aperti permettano di determinare tutta la topologia dello spazio. Una base di aperti è in generale un insieme infinito e può essere comodo averla indicizzata tramite interi. Questo è strettamente connesso con la separabilità dello spazio, per lo meno in uno spazio metrico.

Teo. 2.11 Sia X uno spazio metrico. Allora X è separabile se e solo se esiste una base di aperti numerabile.

Dim. 2.11 Consideriamo prima l'affermazione diretta e poi l'inversa.

Separabilità \Rightarrow base numerabile. Sia X separabile, allora abbiamo un insieme numerabile $Y = \{x_j\}$ denso in X . Consideriamo le sfere di centro x_j e raggio $1/n$:

$$S_{j,n} = S(x_j, \frac{1}{n}), \quad j, n = 1, 2, \dots, \quad (2.33)$$

(le coppie di numeri interi formano un insieme numerabile) e mostriamo che queste formano una base di aperti. Per fare ciò basta mostrare che qualsiasi sfera $S(x, \delta)$ (mediante le sfere costruiamo la topologia in uno spazio metrico) è unione di sfere del tipo (2.33). Sia y appartenente ad una sfera $S(x, \delta)$, $d(x, y) < \delta$, allora, essendo Y denso (in X), in ogni intorno di y , in particolare una sfera $S(y, \eta)$ di raggio arbitrario η , posso trovare un elemento $x_j \in Y$:

$$d(x_j, y) < \eta,$$

e $y \in S(x_j, \eta)$. Scegliendo $\eta = 1/n$ in modo da rimanere sempre all'interno della sfera originaria $S(x, \delta)$, anche con la sfera $S(x_j, \eta)$:

$$2\eta + d(x, y) = \frac{2}{n} + d(y, x) < \delta,$$

ottengo $x_j \in S(x, \delta)$, $y \in S_{j,n}$, con $S_{j,n} \subseteq S(x, \delta)$. Questo dimostra che all'interno di $S(x, \delta)$ troviamo delle sfere $S_{j,n}$ in corrispondenza ad ogni punto e che:

$$S(x, \delta) = \bigcup_{S_{j,n} \subseteq S(x, \delta)} S_{j,n}.$$

Base numerabile \Rightarrow separabilità. Supponiamo l'esistenza di una base numerabile di aperti $\mathcal{O} = \{O_n\}$. Scegliamo un punto x_n in ogni aperto O_n (escludendo l'insieme vuoto che possiamo assumere coincidente con O_0). Otteniamo in questo modo un insieme $Y = \{x_n\}$ denso in X . Infatti se $x \in X$ e W è un intorno aperto di x , W è unione di aperti della base numerabile

$$W = \bigcup_i O_{n_i},$$

e in corrispondenza ad ogni O_{n_i} troviamo un elemento $x_{n_i} \in Y$. x è quindi di aderenza per Y .

— ■ —

Notiamo che la metrica è stata usata solo nella dimostrazione dell'affermazione diretta (separabilità \Rightarrow base numerabile). Per la validità dell'affermazione inversa (base numerabile \Rightarrow separabilità) tale richiesta può essere omessa.

2.4 Continuità

Veniamo ora ad una delle motivazioni principali per la costruzione di una topologia, il concetto di continuità. Riprendiamo la definizione di funzione continua da \mathbb{R} in \mathbb{R} che possiamo esprimere nel seguente modo.

Sia $f : \mathbb{R} \rightarrow \mathbb{R}$, allora f è continua in x_0 se per ogni $\varepsilon > 0$ esiste $\delta(\varepsilon) > 0$ tale che:

$$|x - x_0| < \delta(\varepsilon) \implies |f(x) - f(x_0)| < \varepsilon.$$

Possiamo esprimere più in generale questa definizione in termini di spazi topologici.

Def. 2.20 Siano X e Y due spazi topologici e

$$f : X \rightarrow Y.$$

Sia $x_0 \in X$, allora f è detta **continua** in x_0 se per ogni aperto V contenente $f(x_0)$ esiste un aperto U contenente x_0 tale che $U \subseteq f^{-1}(V)$, o equivalentemente $f(U) \subseteq V$, con l'**immagine** e la **retroimmagine** di insiemi definite rispettivamente da:

$$f(U) = \{y \in Y ; y = f(x) \text{ con } x \in U\}, \quad (2.34)$$

$$f^{-1}(V) = \{x \in X ; f(x) \in V\}. \quad (2.35)$$

Osservazione. Notiamo che, in base alle definizioni di immagine e retroimmagine di un insieme, queste esistono sempre indipendentemente dalla invertibilità della trasformazione, ma in generale, se $A \subseteq X$, allora $A \subseteq f^{-1}(f(A))$ che non coincide necessariamente con A , mentre è sempre vero che $A = f(f^{-1}(A))$, con $A \subseteq Y$. Ovviamente si ha anche:

$$A \subseteq B \subseteq X \implies f(A) \subseteq f(B) \subseteq Y,$$

$$A \subseteq B \subseteq Y \implies f^{-1}(A) \subseteq f^{-1}(B) \subseteq X,$$

ma sulle implicazioni in senso contrario possiamo affermare solamente:

$$f^{-1}(A) \subseteq f^{-1}(B) \subseteq X \implies A \subseteq B \subseteq Y.$$

La definizione precedente è equivalente a dire che:

Teo. 2.12 f è continua in x_0 se e solo se per ogni intorno W di $f(x_0)$, la sua retroimmagine $f^{-1}(W)$ è un intorno di x_0 .

Dim. 2.12 Supponiamo f continua in x_0 e sia W un intorno di $f(x_0)$, allora esiste un aperto V contenente $f(x_0)$, con $V \subseteq W$. Pertanto:

$$f^{-1}(W) \supseteq f^{-1}(V) \supseteq U,$$

con U aperto contenente x_0 , per cui $f^{-1}(W)$ è un intorno di x_0 . Viceversa, possiamo scegliere come intorno W un aperto V e abbiamo $f^{-1}(V)$ intorno di x_0 , per cui esiste un aperto U contenente x_0 tale che $f^{-1}(V) \supseteq U$.

— ■ —

Def. 2.21 f è detta **continua** se è continua in tutti i suoi punti.

Abbiamo diversi modi per verificare la continuità.

Teo. 2.13 Siano X, Y spazi topologici e $f : X \rightarrow Y$. Allora le seguenti affermazioni sono equivalenti:

- i) f è continua.
- ii) Per ogni aperto O , con $O \subseteq Y$, si ha che $f^{-1}(O)$ è un aperto di X .
- iii) Per ogni chiuso C , con $C \subseteq Y$, si ha che $f^{-1}(C)$ è un chiuso di X .
- iv) Per ogni insieme $A \subseteq X$ si ha che $A^- \subseteq f^{-1}(f(A)^-)$, oppure equivalentemente: $f(A^-) \subseteq f(A)^-$.

Dim. 2.13 Vediamo con ordine le equivalenze tra le varie affermazioni.

i) \Rightarrow ii) Sia f continua e O un aperto in Y . Se $x \in f^{-1}(O)$ si ha che $f(x) \in O$, per cui esiste un aperto U contenente x tale che $f(U) \subseteq O$, cioè $U \subseteq f^{-1}(f(U)) \subseteq f^{-1}(O)$ che risulta quindi aperto.

ii) \Rightarrow i) Deriva direttamente dalla definizione di continuità.

ii) \Leftrightarrow iii) Sono equivalenti in maniera ovvia, considerando che un chiuso è il complementare di un aperto e che:

$$X \setminus f^{-1}(A) = f^{-1}(Y \setminus A), \quad \forall A \subseteq Y.$$

iii) \Rightarrow iv) L'insieme $f(A)^-$ è chiuso, quindi $f^{-1}(f(A)^-)$ è chiuso e contiene A , $A \subseteq f^{-1}(f(A)^-) = f^{-1}(f(A)^-)^-$, allora secondo la definizione di chiusura come il più piccolo chiuso contenente A , abbiamo $A^- \subseteq f^{-1}(f(A)^-)$.

iv) \Rightarrow iii) Se C è un chiuso di Y allora considerando vera l'ipotesi per l'insieme $A = f^{-1}(C)$, abbiamo:

$$\begin{aligned} A^- &\subseteq f^{-1}(f(A)^-) \\ f^{-1}(C)^- &\subseteq f^{-1}(C^-) = f^{-1}(C) \implies \\ f^{-1}(C)^- &= f^{-1}(C). \end{aligned}$$

Per cui $f^{-1}(C)$ è chiuso.

L'equivalenza nella seconda parte dell'ultima affermazione deriva dalle considerazioni fatte nell'osservazione precedente. Infatti:

$$\begin{aligned} A^- &\subseteq f^{-1}(f(A)^-) \implies f(A^-) \subseteq f(f^{-1}(f(A)^-)) = f(A)^-, \\ f(A^-) &\subseteq f(A)^- \implies A^- \subseteq f^{-1}(f(A^-)) \subseteq f^{-1}(f(A)^-), \end{aligned}$$

ragionando con gli insiemi A^- e $f(A)^-$.

— ■ —

Osservazione. Osserviamo che l'immagine diretta di insiemi aperti non è necessariamente aperta. Consideriamo ad esempio la seguente funzione:

$$f : \mathbb{R} \longrightarrow [0, \infty[, \quad f(x) = x^2.$$

Abbiamo che l'immagine dell'intervallo aperto $] -1, 1[$ non è aperto:

$$f :] -1, 1[\longrightarrow [0, 1[.$$

L'insieme di partenza è aperto, quello di arrivo no.

Osservazione. Abbiamo visto come, dati due spazi topologici, possiamo verificare se una funzione è continua o meno. Viceversa, data una funzione $f : X \longrightarrow Y$, con Y spazio topologico, possiamo costruire in X una topologia tale che f risulti continua. Basta prendere in X la topologia generata dalle retroimmagini degli aperti in Y .

Un esempio di ciò si ha nella costruzione di una topologia per il prodotto cartesiano di due spazi topologici.

Esempio 2.13 Siano X_1, X_2 due spazi topologici e consideriamo il prodotto cartesiano:

$$X_1 \times X_2 = \{(x_1, x_2); x_1 \in X_1, x_2 \in X_2\}. \quad (2.36)$$

Possiamo definire allora le funzioni $p_j : X_1 \times X_2 \longrightarrow X_j$, $j = 1, 2$, definite da:

$$p_j(x_1, x_2) = x_j, \quad j = 1, 2, \quad (2.37)$$

che danno le componenti dei punti nel prodotto cartesiano. Le retroimmagini di insiemi aperti in X_1 e X_2 (sono delle strisce) costituiscono una collezione di insiemi che generano

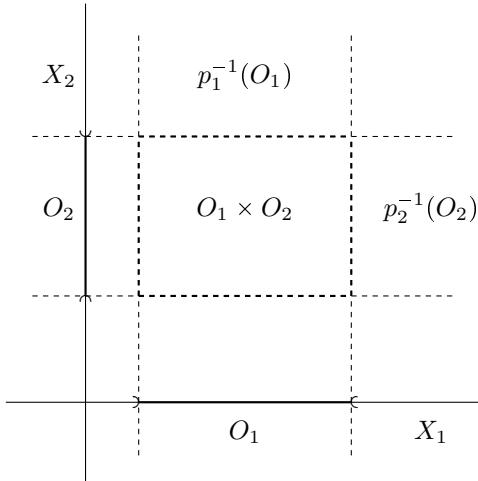


Figura 2.5: Prodotto cartesiano di topologie.

una topologia nel prodotto cartesiano. Risulta chiaro che in tale topologia le funzioni p_j sono continue. Questa topologia può essere vista come generata dai “rettangoli aperti” del tipo $O_1 \times O_2$. Ad esempio in \mathbb{R}^2 un aperto è una unione di rettangoli $\{a_1 < x_1 < b_1, a_2 < x_2 < b_2\}$.

In maniera analoga al prodotto cartesiano di due spazi topologici si può definire una topologia nel prodotto cartesiano di una arbitraria collezione di spazi topologici X_α , con $\alpha \in \mathcal{A}$ insieme di indici, ma in generale non risulta espressa tramite “polirettangoli”, se non nel caso che gli spazi topologici X_α siano in numero finito. Il motivo consiste nel fatto che possiamo intersecare le strisce aperte solo in numero finito per avere un aperto, cioè possiamo considerare “intervalli aperti” solo lungo un numero finito di dimensioni.

— ○ —

Osservazione. Considerando l’esempio 2.13 con due copie di uno spazio metrico X , la cui topologia è sostanzialmente determinata dalla metrica, possiamo vedere che la distanza d risulta una funzione continua da $X \times X$ a \mathbb{R} .

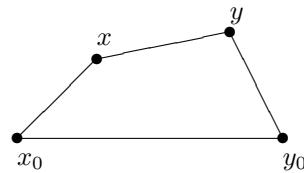


Figura 2.6: Continuità della distanza.

Siano infatti (x_0, y_0) , (x, y) due punti nel prodotto cartesiano $X \times X$. Abbiamo,

applicando la diseguaglianza triangolare:

$$\begin{aligned} d(x, y) &\leq d(x, x_0) + d(x_0, y) \leq d(x, x_0) + d(x_0, y_0) + d(y_0, y), \\ d(x, y) - d(x_0, y_0) &\leq d(x, x_0) + d(y_0, y), \end{aligned}$$

e analogamente (partendo da $d(x_0, y_0)$):

$$d(x_0, y_0) - d(x, y) \leq d(x_0, x) + d(y, y_0).$$

Cioè:

$$|d(x, y) - d(x_0, y_0)| \leq d(x, x_0) + d(y, y_0). \quad (2.38)$$

Allora per ogni intorno V di $d(x_0, y_0) \in \mathbb{R}$, che possiamo assumere un intervallo aperto di raggio ϵ , possiamo scegliere come intorno U di (x_0, y_0) l'insieme aperto:

$$U = S(x_0, \frac{\epsilon}{2}) \times S(y_0, \frac{\epsilon}{2}),$$

ottenendo la continuità della distanza come funzione.

Ricordiamo ora che se abbiamo tre spazi topologici X , Y e Z e due funzioni $f : X \rightarrow Y$, $g : Y \rightarrow Z$, possiamo costruire la funzione composta $g \circ f : X \rightarrow Z$, definita da:

$$(g \circ f)(x) = g(f(x)). \quad (2.39)$$

Se f e g sono entrambe continue allora anche la funzione composta $g \circ f$ è continua.

Sia infatti V un aperto di Z , allora, essendo g continua abbiamo che $g^{-1}(V)$ è un aperto in Y , e, per la continuità di f , $f^{-1}(g^{-1}(V)) = (g \circ f)^{-1}(V)$ è un aperto di X .

Def. 2.22 Sia $f : X \rightarrow Y$. Se $f(x) = f(x')$ implica che $x = x'$ allora diremo che f è **iniettiva** o 1 – 1 (a volte è detta anche semplicemente **in**).

Se $f(X) = Y$ diremo che f è **suriettiva** o semplicemente **su**.

Se f è 1 – 1 e suriettiva è detta **bijettiva** o una **biiezione**, e, in tal caso, esiste un'unica funzione inversa f^{-1} tale che $f^{-1} \circ f = id_X$, e $f \circ f^{-1} = id_Y$:

$$f^{-1}(f(x)) = x \quad \forall x \in X; \quad f(f^{-1}(y)) = y \quad \forall y \in Y. \quad (2.40)$$

Con id_X e id_Y denotiamo le applicazioni identiche in X e Y rispettivamente.

Osservazione. Se f è continua ed esiste l'inversa, non è detto che l'inversa sia continua. Notiamo inoltre che $f : X \rightarrow Y$ è invertibile se possiamo applicare f^{-1} a qualsiasi elemento di Y .

Consideriamo ad esempio la semplice identità id in \mathbb{R} , ma pensata tra \mathbb{R} con la topologia discreta (come spazio di partenza) e \mathbb{R} con la topologia ordinaria (come spazio di arrivo):

$$id : x \mapsto x. \quad (2.41)$$

Con l'identità ogni insieme coincide con la sua immagine e con la sua retroimmagine (indipendentemente dalla topologia). Con le topologie introdotte l'identità risulta continua (la retroimmagine di un aperto è aperto in quanto, nella topologia discreta, tutti gli insiemi sono aperti), ma la sua inversa (che è ancora l'identità) non è continua in quanto nella topologia ordinaria un qualsiasi insieme non è necessariamente aperto.

Def. 2.23 Siano X, Y due spazi topologici. Una biiezione di X su Y è detta un **omeomorfismo** se sia f che f^{-1} sono continue.

Due spazi topologici sono detti **omeomorfi** se esiste fra loro un omeomorfismo.

In pratica occorre poter mutare ogni aperto di uno spazio in un aperto dell'altro spazio e i due spazi sono, da un punto di vista topologico, la stessa cosa.

2.4.1 Continuità e compattezza

Consideriamo ora una importante connessione tra continuità e compattezza.

Teo. 2.14 Siano X e Y due spazi topologici e $f : X \rightarrow Y$ una funzione continua e suriettiva tra i due spazi. Allora se X è compatto anche Y è compatto.

Dim. 2.14 Consideriamo un ricoprimento aperto di Y :

$$Y = \bigcup_{\alpha} O_{\alpha},$$

allora le retroimmagini $f^{-1}(O_{\alpha})$ formano un ricoprimento aperto di X . Possiamo selezionarne quindi un numero finito:

$$X = \bigcup_{i=1}^N f^{-1}(O_{\alpha_i}),$$

ed essendo f suriettiva e $f(f^{-1}(O_{\alpha})) = O_{\alpha}$, otteniamo che gli aperti O_{α_i} formano un sottoricoprimento finito di Y che risulta compatto.

— ■ —

Come conseguenza abbiamo:

Cor. 2.15 Se $f : X \rightarrow Y$ è continua e X è compatto, allora l'immagine $f(X)$ è un insieme compatto di Y .

Osservazione. Le funzioni continue sui compatti trasportano le informazioni di compattezza in avanti (dallo spazio di partenza allo spazio di arrivo). Invece le proprietà di chiusura o apertura sono trasportate all'indietro (la retroimmagine di un aperto (chiuso) tramite una funzione continua è un aperto (chiuso)).

Anche la proprietà di Hausdorff ha conseguenze sulla continuità. In particolare come conseguenza del corollario 2.15 e del precedente risultato abbiamo il teorema di Weierstrass:

Teo. 2.16 *Se f è una funzione reale e continua su un compatto allora f ammette massimo e minimo.*

Dim. 2.16 È una conseguenza immediata del fatto che f muta compatti in compatti e un compatto in \mathbb{R} è un insieme chiuso e limitato, quindi dotato di massimo e minimo.

— ■ —

Teo. 2.17 *Se f è una trasformazione continua tra uno spazio topologico compatto X e uno spazio di Hausdorff Y allora f trasforma insiemi chiusi in insiemi chiusi.*

Dim. 2.17 Sia C chiuso in X , allora C risulta compatto (vedi teorema 2.6) e il suo trasformato $f(C)$ è compatto (corollario 2.15). Poichè siamo in uno spazio di Hausdorff $f(C)$ è chiuso.

— ■ —

2.5 Successioni.

Def. 2.24 Sia X uno spazio topologico e $\{x_n, n = 1, 2, \dots\}$ una successione di elementi di X . Diremo che x_n **converge** verso x se per ogni intorno U di x esiste un intero $n(U)$ tale che $x_n \in U$ per ogni $n \geq n(U)$, e scriveremo:

$$x = \lim_{n \rightarrow \infty} x_n, \quad (2.42)$$

oppure:

$$x_n \rightarrow x \text{ per } n \rightarrow \infty; \quad (2.43)$$

$$x_n \xrightarrow[n \rightarrow \infty]{} x. \quad (2.44)$$

In pratica significa che da ogni intorno di x risultano esclusi solo un numero finito di elementi della successione, e gli elementi della successione si “avvicinano” sempre più a x . Il concetto di limite è quindi strettamente legato alla topologia dello spazio. Quando esiste il limite siamo abituati a pensare che tale limite sia unico, ma tale proprietà è una prerogativa garantita solo negli spazi *separati* di Hausdorff.

Teo. 2.18 Sia X uno spazio topologico di Hausdorff. Se x_n è una successione convergente, allora il limite è unico.

Dim. 2.18 Infatti se $x_n \rightarrow x$ e contemporaneamente $x_n \rightarrow y$ (cioè sia x che y verificano la definizione 2.24 di limite), allora da un certo punto in poi i punti x_n si troveranno sia in un intorno arbitrario di x che in un intorno arbitrario di y , ma se $x \neq y$ ciò non è possibile perché possiamo scegliere gli intorni disgiunti. Pertanto si deve avere $x = y$.

— ■ —

Il concetto di convergenza di una successione è molto simile al concetto di continuità per una funzione. Abbiamo in effetti il seguente risultato.

Teo. 2.19 Se $f : X \rightarrow Y$ è una trasformazione fra due spazi topologici X e Y ed è continua in un punto x , allora per ogni successione x_n convergente a x abbiamo che $f(x_n)$ è una successione in Y convergente a $f(x)$:

$$\lim_{n \rightarrow \infty} f(x_n) = f \left(\lim_{n \rightarrow \infty} x_n \right), \quad (2.45)$$

(possiamo “scambiare” l’ordine operativo tra il limite e la funzione).

Dim. 2.19 Sia U un intorno di $f(x)$. Per la continuità $f^{-1}(U)$ è un intorno di x , per cui $x_n \in f^{-1}(U)$ per $n \geq n(U)$, e $f(x_n) \in U$.

— ■ —

In generale, per spazi topologici arbitrari, il viceversa non è vero, cioè la convergenza per successioni non implica la continuità. Per avere ciò occorre aggiungere delle ipotesi opportune sulla struttura topologica degli intorni di un punto. Essendo interessati alla convergenza e continuità in un singolo punto x , consideriamo il concetto di base di intorni di un punto.

Def. 2.25 Sia X uno spazio topologico e $x \in X$. Chiameremo **base di intorni** di x una famiglia di intorni di x tale che ogni intorno (arbitrario) di x contiene un intorno della famiglia.

Osservazione. Se X è uno spazio metrico, allora ogni punto x possiede una base di intorni numerabile. Basta considerare le sfere

$$S_n = S(x, \frac{1}{n}).$$

Supponiamo che un punto x ammetta una base di intorni $\{U_n\}$ numerabile. Allora è possibile costruire la base di intorni $\{V_n\}$ in modo tale che sia ordinata rispetto all'inclusione (in senso decrescente):

$$V_1 \supseteq V_2 \supseteq V_3 \supseteq \dots.$$

Basta definire:

$$\begin{aligned} V_1 &= U_1, \\ V_2 &= U_1 \bigcap U_2, \\ &\vdots \\ V_n &= U_1 \bigcap U_2 \bigcap \cdots \bigcap U_n, \\ &\vdots \end{aligned}$$

L'esistenza di una base numerabile di intorni è la proprietà topologica necessaria per far discendere la continuità generale dalla continuità per convergenza di successioni.

Teo. 2.20 Siano X, Y spazi topologici. $x \in X$ possiede una base di intorni numerabile e sia $f : X \rightarrow Y$ una applicazione fra i due spazi. Se è vero che per ogni successione x_n convergente a x si ha che $f(x_n)$ converge a $f(x)$:

$$x_n \rightarrow x \implies f(x_n) \rightarrow f(x), \quad (2.46)$$

allora f è continua in x .

Osservazione. Sostanzialmente nel caso di spazi topologici con basi di intorni numerabili si può studiare la continuità per mezzo della convergenza di successioni.

Dim. 2.20 Se, per assurdo, f non è continua in x esiste un intorno U di $f(x)$ tale che $f^{-1}(U)$ non è un intorno di x . Questo implica che, avendo costruito un base numerabile, decrescente, di intorni $V_j \supseteq V_{j+1}$ di x , $f^{-1}(U)$ non contiene nessun intorno V_j , e per ogni V_j possiamo trovare $x_j \in V_j$ con $x_j \notin f^{-1}(U)$, cioè $f(x_j) \notin U$. Questo comporta che $f(x_j)$ non converge a $f(x)$ (altrimenti U dovrebbe contenere $f(x_j)$ da un certo punto in poi). D'altra parte x_j converge verso x . Infatti se W è un intorno arbitrario di x esso conterrà un intorno della base V_j e tutti i successivi (essendo essi decrescenti), quindi conterrà x_j e tutti i successivi punti della successione. L'ipotesi del teorema asserisce di conseguenza che $f(x_j)$ converge a $f(x)$. Da tale contraddizione abbiamo che f deve essere continua in x .

— ■ —

Tramite l'esistenza di basi di intorni numerabili possiamo dare un'altra caratterizzazione alla chiusura di un insieme A .

Teo. 2.21 *Sia X uno spazio topologico tale che ogni suo punto ammetta una base di intorni numerabile e sia A un sottoinsieme di X . Allora vale*

$$A^- = \left\{ x \in X ; x = \lim_{n \rightarrow \infty} x_n, \text{ con } x_n \in A \right\}. \quad (2.47)$$

Dim. 2.21 Ricordiamo che abbiamo definito la chiusura come l'insieme dei punti di aderenza:

$$A^- = \{x \in X ; \text{ogni intorno } U_x \text{ di } x \text{ contiene punti di } A\}.$$

Se $x \in A^-$ allora per ogni U intorno di x esiste $y \in A \cap U$. Consideriamo gli intorni V_j della base e sia $y_j \in A$ il corrispondente punto in V_j . Chiaramente otteniamo una successione $y_j \in A$ convergente a x .

Viceversa sia $x_n \rightarrow x$, con $x_n \in A$. Allora per ogni intorno U di x abbiamo $x_j \in U$ da un certo punto in poi. Ma $x_j \in A$ per cui x è di aderenza per A . Notiamo che questo è sempre vero, indipendentemente dall'esistenza della base numerabile. Il limite di una successione convergente è sempre un punto di aderenza per la successione e l'insieme di appartenenza della successione.

— ■ —

Mediante le successioni si può costruire una nozione leggermente diversa di compattezza, quella di sequenziale compattezza.

Def. 2.26 *Un sottoinsieme K di uno spazio topologico X è detto **sequenzialmente compatto** se da ogni successione di elementi di K è possibile estrarre una sottosuccessione convergente.*

Osservazione. Notiamo che non è necessario che la sottosuccessione sia convergente nell'insieme K , è importante che esista il limite in X . Se K è chiuso allora possiamo dire che il limite appartiene a K .

Nella pratica, quando abbiamo a disposizione una metrica, le due definizioni, di compattezza e di sequenziale compattezza, sono equivalenti (a parte la chiusura). Infatti abbiamo il seguente risultato.

Teo. 2.22 *Sia X uno spazio metrico. Allora un sottoinsieme K è compatto se e solo se K è chiuso e sequenzialmente compatto.*

Dim. 2.22 Sia K compatto. Essendo X metrico e quindi di Hausdorff, K risulta chiuso. Supponiamo per assurdo che K non sia sequenzialmente compatto, allora esiste una successione di elementi x_n di K che non ha alcuna sottosuccessione convergente. Tale successione deve avere quindi infiniti elementi diversi tra loro (in caso contrario otterremo una sottosuccessione stazionaria e convergente). Se prendo un generico punto x dello spazio deve esistere un suo intorno (aperto) contenente al più un numero finito di punti della successione (non possiamo convergere con una sottosuccessione da nessuna parte, per cui x non verifica la definizione di limite). Tramite gli intorni di ogni punto così costruiti formiamo quindi un ricoprimento aperto di K . Ogni collezione finita di aperti può contenere quindi al massimo un numero finito di elementi della successione e quindi (essendo la successione in K), non può formare un ricoprimento di K . Ciò contraddice l'ipotesi per cui K deve risultare sequenzialmente compatto.

Viceversa sia K chiuso e sequenzialmente compatto. Mostriamo innanzitutto che K è separabile costruendo un suo sottoinsieme denso e numerabile. Anticipiamo prima un risultato che approfondiremo in seguito. Se abbiamo una successione x_n convergente in uno spazio metrico, cioè esiste il limite:

$$x = \lim_{n \rightarrow \infty} x_n,$$

allora necessariamente si ha (criterio di Cauchy):

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) \xrightarrow[m, n \rightarrow \infty]{} 0.$$

Consideriamo allora un punto $x_0 \in K$ e sia:

$$\epsilon_0 = \sup_{x \in K} d(x, x_0).$$

L'estremo superiore esiste ed è finito. Infatti se non lo fosse potremmo scegliere $y_1 \in K$ tale che:

$$d(y_1, x_0) > 1,$$

poi y_2 in modo tale che:

$$d(y_2, x_0) > d(y_1, x_0) + 1,$$

e così via:

$$d(y_n, x_0) > d(y_{n-1}, x_0) + 1,$$

per cui:

$$d(y_n, y_{n-1}) \geq |d(y_n, x_0) - d(y_{n-1}, x_0)| > 1,$$

e, in generale:

$$d(y_n, x_0) > d(y_m, x_0) + (n - m), \quad n > m,$$

$$d(y_n, y_m) > (n - m), \quad \forall n > m.$$

Allora per ogni sottosuccessione estratta da $\{y_n\}$ si avrebbe:

$$d(y_{n_j}, y_{n_k}) > |n_j - n_k| \geq 1,$$

e la sottosuccessione non può convergere. Ciò costituisce una contraddizione per cui ϵ_0 esiste ed è finito. In pratica la sfera $S(x_0, \epsilon_0)$ è la sfera di ampiezza minore che racchiude tutto K (“toccando” K sul bordo, vedi figura 2.7).



Figura 2.7: Costruzione di un sottoinsieme denso di un compatto.

Appurata l'esistenza di ϵ_0 , scegliamo un punto $x_1 \in K$ tale che:

$$\epsilon_0 \geq d(x_1, x_0) \geq \frac{\epsilon_0}{2}$$

Dati i due punti x_0, x_1 consideriamo la funzione (limitata):

$$d_1(x) = \min\{d(x, x_0), d(x, x_1)\} \leq d(x, x_0) \leq \epsilon_0,$$

per cui, posto:

$$\epsilon_1 = \sup_{x \in K} d_1(x) \leq \epsilon_0,$$

possiamo selezionare x_2 tale che:

$$\epsilon_1 \geq d_1(x_2) \geq \frac{\epsilon_1}{2}.$$

Di nuovo abbiamo racchiuso l'insieme K nell'unione delle due sfere centrate in x_0 e x_1 di raggio comune e minimo possibile ϵ_1 (figura 2.7).

Procedendo per ricorsione costruisco una successione di punti x_n , e una successione di reali ϵ_n tale che:

$$\begin{aligned}\epsilon_{n-1} &\geq \min_{0 \leq j \leq n-1} d(x_n, x_j) \geq \frac{\epsilon_{n-1}}{2}, \\ \epsilon_n &= \sup_{x \in K} \left(\min_{0 \leq j \leq n} d(x, x_j) \right) \leq \epsilon_{n-1}.\end{aligned}$$

La successione ϵ_n è decrescente per cui convergente. Essa deve tendere a zero. Infatti, se ciò non fosse vero, esisterebbe $\eta > 0$ tale che:

$$\epsilon_{n-1} \geq \min_{0 \leq j \leq n-1} d(x_n, x_j) \geq \frac{\epsilon_{n-1}}{2} > \eta,$$

e non sarebbe possibile estrarre una sottosuccessione convergente dalla successione x_n . Ciò è assurdo per cui $\epsilon_n \rightarrow 0$.

Allora mediante i punti x_n possiamo “approssimare” qualsiasi $x \in K$. Dato ϵ e scelto $\epsilon_n < \epsilon$ abbiamo, per ogni $x \in K$:

$$\min_{0 \leq j \leq n} d(x, x_j) \leq \epsilon_n < \epsilon,$$

per cui esiste x_j ($j \leq n$) tale che:

$$d(x, x_j) < \epsilon.$$

K risulta quindi separabile. Essendo K uno spazio metrico, ammette una base di aperti numerabile $\{V_j\}$, con V_j aperto in K e:

$$K = \bigcup_j V_j.$$

Supponiamo ora di avere un ricoprimento aperto di K :

$$\begin{aligned}K &\subseteq \bigcup_\lambda O_\lambda, \\ K &= \bigcup_\lambda O'_\lambda, \quad O'_\lambda = O_\lambda \cap K.\end{aligned}$$

Sappiamo che ogni aperto O'_λ è unione di elementi della base numerabile $\{V_j\}$. Costruiamo l'insieme di indici:

$$\mathcal{J} = \{j ; \exists \lambda, V_j \subseteq O'_\lambda\},$$

allora:

$$K = \bigcup_\lambda O'_\lambda = \bigcup_{j \in \mathcal{J}} V_j.$$

Infatti ogni V_j , con $j \in \mathcal{J}$, è contenuto, per la scelta fatta, in almeno uno degli aperti O'_λ , e quindi nella loro unione. D'altra parte, ogni O'_λ si ottiene come unione di aperti della base V_j , e gli aperti che concorrono a tale unione sono contenuti in O'_λ , per cui deve esistere un sottoinsieme \mathcal{J}' di \mathcal{J} , tale che:

$$O'_\lambda = \bigcup_{j \in \mathcal{J}'} V_j, \quad \mathcal{J}' \subseteq \mathcal{J},$$

e l'unione di tutti gli aperti O'_λ è contenuta nell'unione degli aperti V_j , $j \in \mathcal{J}$.

Possiamo ora rinumerare l'insieme numerabile di aperti $\{V_j ; j \in \mathcal{J}\}$, ottenendo sempre (non è più, in generale, una base di aperti):

$$\bigcup_j V_j = \bigcup_\lambda O'_\lambda = K;$$

Per ogni j possiamo ora scegliere $O'_{\lambda(j)}$ tale che:

$$V_j \subseteq O'_{\lambda(j)} \subseteq O_{\lambda(j)},$$

per cui abbiamo estratto un sottoricoprimento numerabile:

$$K = \bigcup_j V_j \subseteq \bigcup_j O_{\lambda(j)}.$$

Vediamo ora che possiamo estrarre un sottoricoprimento finito. Se per assurdo non fosse possibile, possiamo costruire una successione x_n tale che:

$$\begin{aligned} x_1 &\in K, \quad x_1 \notin O_{\lambda(1)}, \\ x_2 &\in K, \quad x_2 \notin O_{\lambda(1)} \bigcup O_{\lambda(2)}, \\ &\vdots \\ x_n &\in K, \quad x_n \notin \bigcup_{j=0}^n O_{\lambda(j)}, \end{aligned}$$

realizzando una successione in K tale che ogni $O_{\lambda(j)}$ contiene solo un numero finito di punti. Non è quindi possibile estrarre una sottosuccessione convergente in K . Di nuovo un assurdo per cui deve esistere un sottoricoprimento finito e K è compatto.

— ■ —

2.5.1 Completezza.

Come abbiamo visto nella dimostrazione del teorema 2.22, se abbiamo una successione convergente in uno spazio metrico, la distanza tra gli elementi della successione deve tendere a zero. Tale proprietà è detta di Cauchy.

Def. 2.27 Sia X uno spazio metrico. Allora una successione $\{x_n\}$ è detta **di Cauchy** se per ogni $\epsilon > 0$ esiste un n_ϵ tale che:

$$d(x_n, x_m) < \epsilon \quad \forall n, m \geq n_\epsilon. \quad (2.48)$$

Se abbiamo una successione x_n convergente a x sappiamo che per ogni ϵ possiamo determinare n_ϵ tale che se $n \geq n_\epsilon$:

$$d(x_n, x) < \frac{\epsilon}{2},$$

per cui, se $n, m \geq n_\epsilon$:

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \epsilon.$$

Il viceversa non è vero in generale, e quando le proprietà dello spazio metrico X sono tali da verificare ciò ci troviamo in presenza di uno spazio completo.

Def. 2.28 Uno spazio metrico X è detto **completo** se ogni successione di Cauchy ammette un limite in X .

In pratica in uno spazio completo la condizione di Cauchy diventa un criterio di convergenza non solo necessario ma anche sufficiente. Esempi di spazi completi sono \mathbb{R}^n e \mathbb{C}^n in cui è noto che il criterio di Cauchy è il metodo principale per verificare la convergenza di una successione. Vediamo subito una semplice applicazione del concetto di completezza.

Def. 2.29 Sia X uno spazio metrico e $f : X \rightarrow X$ una trasformazione in tale spazio. f è detta una **contrazione** se esiste una costante α , con $0 \leq \alpha < 1$, tale che:

$$d(f(x), f(y)) \leq \alpha d(x, y). \quad (2.49)$$

Se ci troviamo in uno spazio metrico completo possiamo garantire allora che:

Teo. 2.23 In uno spazio metrico completo ogni contrazione ammette un unico **punto fisso**, cioè un punto x tale che:

$$f(x) = x. \quad (2.50)$$

Dim. 2.23 L'unicità è immediata. Supponiamo esistano due punti x, y tali che $f(x) = x$, $f(y) = y$, allora:

$$d(f(x), f(y)) \leq \alpha d(x, y),$$

$$d(x, y) \leq \alpha d(x, y),$$

con $\alpha < 1$ si giunge ad un assurdo se $x \neq y$, per cui il punto fisso è unico (indipendentemente dalla completezza).

Vediamo ora l'esistenza della soluzione dell'equazione (2.50). Sia x_0 un qualsiasi punto (un "guess" iniziale per la soluzione), e costruiamo la successione:

$$x_n = f(x_{n-1}), \quad n = 1, 2, \dots$$

Allora abbiamo:

$$\begin{aligned} d(x_n, x_m) &= d(f(x_{n-1}), f(x_{m-1})) \leq \alpha d(x_{n-1}, x_{m-1}) \\ &\leq \alpha^2 d(x_{n-2}, x_{m-2}) \leq \alpha^m d(x_{n-m}, x_0), \quad (n > m), \\ d(x_q, x_0) &\leq \sum_{j=0}^{q-1} d(x_{q-j}, x_{q-j-1}) \leq \sum_{j=0}^{q-1} \alpha^{q-j-1} d(x_1, x_0) \\ &= \left(\sum_{j=0}^{q-1} \alpha^j \right) d(x_1, x_0) = \frac{1 - \alpha^q}{1 - \alpha} d(x_1, x_0), \\ d(x_n, x_m) &\leq \alpha^m \frac{1 - \alpha^{n-m}}{1 - \alpha} d(x_1, x_0) \leq \frac{\alpha^m}{1 - \alpha} d(x_1, x_0), \end{aligned}$$

e, notando che $1 - \alpha > 0$, $\alpha^m \rightarrow 0$ per $m \rightarrow \infty$, abbiamo che la successione x_n è di Cauchy. Essendo lo spazio completo, abbiamo l'esistenza del limite, $x_n \rightarrow x$, ed essendo ovvio che una contrazione è anche una funzione continua:

$$\begin{aligned} x_n &= f(x_{n-1}) \\ &\downarrow \quad \downarrow \quad n \rightarrow \infty \\ x &= f(x), \end{aligned}$$

provando il risultato.

— ■ —

Osservazione. Dalla relazione:

$$d(x_n, x_m) \leq \alpha^m \frac{1 - \alpha^{n-m}}{1 - \alpha} d(x_1, x_0)$$

mandando n all'infinito a fissato m , abbiamo:

$$\lim_{n \rightarrow \infty} d(x_n, x_m) \leq \frac{\alpha^m}{1 - \alpha} d(x_1, x_0),$$

e, poichè la distanza è essa stessa una funzione continua:

$$d(x, x_m) \leq \frac{\alpha^m}{1 - \alpha} d(f(x_0), x_0),$$

ottenendo una stima (anche se rozza) della rapidità di convergenza. Un'altra stima la otteniamo con un ragionamento già fatto, notando che:

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{j=1}^p d(x_{n+j}, x_{n+j-1}) \leq \sum_{j=1}^p \alpha^j d(x_n, x_{n-1}) \\ &\leq \frac{\alpha}{1-\alpha} d(x_n, x_{n-1}), \end{aligned}$$

da cui, passando al limite per $p \rightarrow \infty$:

$$d(x, x_n) \leq \frac{\alpha}{1-\alpha} d(x_n, x_{n-1}).$$

In uno spazio metrico, possiamo raffinare ulteriormente il concetto di continuità.

Def. 2.30 Siano X e Y due spazi metrici dotati di distanze d_X e d_Y rispettivamente. Diremo che $f : X \rightarrow Y$ è **uniformemente continua** se per ogni $\epsilon > 0$ esiste $\delta(\epsilon)$ tale che:

$$d_Y(f(x), f(y)) < \epsilon$$

ogni volta che

$$d_X(x, y) < \delta(\epsilon),$$

con $\delta(\epsilon)$ indipendente dalla scelta dei punti x e y in cui si verifica la continuità.

Osservazione. Notiamo che la nozione di continuità uniforme è un concetto metrico, non topologico, in quanto esprime una proprietà della misura degli intorni coinvolti nella definizione di continuità.

Teo. 2.24 Siano X e Y spazi metrici, con X compatto e $f : X \rightarrow Y$ continua. Allora f è uniformemente continua.

Dim. 2.24 Se f non fosse uniformemente continua allora esisterebbe un $\epsilon > 0$ e due successioni x_n, y_n tali che:

$$d_Y(f(x_n), f(y_n)) > \epsilon, \quad d_X(x_n, y_n) < \frac{1}{n}.$$

Ora, per la compattezza di X e la sua sequenziale compattezza, esiste una sottosuccessione $x_{n_j} \rightarrow x$. Allora anche $y_{n_j} \rightarrow x$:

$$d_X(y_{n_j}, x) \leq d_X(y_{n_j}, x_{n_j}) + d_X(x_{n_j}, x) < \frac{1}{n_j} + d_X(x_{n_j}, x).$$

Siccome f è continua, si deve avere $f(x_{n_j}) \rightarrow f(x)$ e $f(y_{n_j}) \rightarrow f(x)$, che contraddice la relazione iniziale $d_Y(f(x_n), f(y_n)) > \epsilon$.

— ■ —

La nozione di uniforme continuità è importante ai fini del **prolungamento continuo** di applicazioni continue fra due spazi metrici.

Teo. 2.25 Siano X e Y due spazi metrici, D un sottoinsieme denso di X e f una applicazione continua da D a Y . Se Y è completo e f uniformemente continua, allora esiste una unica applicazione \tilde{f} da X a Y , uniformemente continua, detta **estensione continua**, che prolunga f , cioè:

$$f(x) = \tilde{f}(x) \quad \forall x \in D. \quad (2.51)$$

Dim. 2.25 Sia $x \in X$, allora esiste una successione $x_n \in D$ tale che $x_n \rightarrow x$ (x potrebbe non appartenere a D). Ora, per la continuità uniforme, $f(x_n)$ risulta di Cauchy:

$$d_Y(f(x_n), f(x_m)) < \epsilon \quad \text{se} \quad d_X(x_n, x_m) < \delta(\epsilon),$$

(con $\delta(\epsilon)$ indipendente da x_n e x_m), che è verificata se n e m sono abbastanza grandi. Essendo Y completo $f(x_n)$ converge ad un limite $\xi \in Y$:

$$f(x_n) \longrightarrow \xi.$$

Consideriamo ora un'altra possibile successione in D convergente verso x , $y_n \rightarrow x$, allora $f(y_n)$ è di Cauchy e converge verso un limite η . Abbiamo (per la continuità della distanza):

$$d_Y(\xi, \eta) = \lim_{n \rightarrow \infty} d_Y(f(x_n), f(y_n)).$$

Siccome $x_n \rightarrow x$, e $y_n \rightarrow x$, la distanza $d_X(x_n, y_n)$ può essere resa piccola a piacere e di conseguenza, per l'uniforme continuità di f anche la distanza $d_Y(f(x_n), f(y_n))$ può essere resa piccola a piacere. Pertanto $\xi = \eta$ indipendente dalla scelta della successione convergente a x . Abbiamo definito quindi in maniera univoca una funzione $\tilde{f}(x)$ per ogni $x \in X$:

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} f(x_n), \quad \text{con} \quad x_n \xrightarrow{n \rightarrow \infty} x.$$

Se $x \in D$ possiamo considerare una successione stazionaria $\{x_n = x\}$, per la quale $f(x_n) = f(x) \longrightarrow f(x)$, per cui \tilde{f} è una estensione di f .

Vediamo ora che \tilde{f} è uniformemente continua. Sia $\epsilon > 0$, allora tramite f (uniformemente continua in D) possiamo trovare un $\delta(\epsilon)$ tale che:

$$d_X(x, y) < \delta(\epsilon), \quad x, y \in D \implies d_Y(f(x), f(y)) < \epsilon.$$

Siano allora $x, y \in X$, e $x_n, y_n \in D$, tali che:

$$d_X(x, y) < \frac{\delta(\epsilon)}{3}, \quad d_X(x, x_n) < \frac{\delta(\epsilon)}{3}, \quad d_X(y, y_n) < \frac{\delta(\epsilon)}{3};$$

allora:

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n) < \delta(\epsilon) \implies d_Y(f(x_n), f(y_n)) < \epsilon,$$

e passando al limite per $n \rightarrow \infty$:

$$d_Y(\tilde{f}(x), \tilde{f}(y)) \leq \epsilon,$$

per cui \tilde{f} risulta uniformemente continua.

L'unicità dell'estensione è garantita dalla continuità uniforme di f e di \tilde{f} . Infatti se $x \in X$, esiste una successione $x_n \in D$ tale che $x_n \rightarrow x$, e qualunque estensione (continua) \tilde{f} deve verificare $f(x_n) = \tilde{f}(x_n) \rightarrow \tilde{f}(x)$ unico per l'unicità del limite.

— ■ —

Osservazione. Una applicazione non si sa estendere con continuità se Y non è completo. Sia X qualsiasi con D denso in X e $Y = D$, con $f(x) = x$. Se potessimo prolungare con continuità giungeremmo ad un assurdo, malgrado l'identità sia uniformemente continua. Se esiste \tilde{f} allora per ogni $x \in X$ esiste una successione $x_n \rightarrow x$ in X e $x_n = f(x_n) = \tilde{f}(x_n) \rightarrow \tilde{f}(x) \in Y = D$, ma allora $x = \tilde{f}(x) \in D$ che è assurdo.

Osservazione. Se Y è completo, ma f non è uniformemente continua, può succedere che il prolungamento non esista. Si consideri il seguente esempio:

$$\begin{aligned} X &= [0, 1], \quad D =]0, 1[, \quad Y = \mathbb{R}, \\ f(x) &= \frac{1}{x}, \quad x \in D, \end{aligned}$$

che non può essere estesa a tutto l'inervallo $[0, 1]$.

Il fatto che una successione di Cauchy possa non convergere può lasciare perplessi essendo abituati a considerare tale condizione un criterio di convergenza fondamentale. In effetti si può sempre fare in modo che il criterio di Cauchy sia sufficiente aggiungendo, in un certo senso, allo spazio anche i punti limite delle successioni di Cauchy. Per vedere ciò introduciamo prima il concetto di isometria.

Def. 2.31 Una applicazione f da uno spazio metrico X ad uno spazio metrico Y è detta **isometrica** se lascia invariate le distanze:

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2), \quad (2.52)$$

e i due spazi sono detti **isometrici** quando esiste una applicazione isometrica tra loro.

Chiaramente l'applicazione isometrica fra i due spazi risulta una applicazione iniettiva e continua, anzi uniformemente continua e, da un punto di vista metrico (e quindi anche topologico) lo spazio X e la sua immagine $f(X)$ sono perfettamente identificabili fra loro.

Def. 2.32 Sia X uno spazio metrico non completo, diremo che lo spazio metrico completo \tilde{X} è un **completamento** di X , se esiste una applicazione isometrica di X in \tilde{X} , tale che l'immagine di X è densa in \tilde{X} .

Teo. 2.26 Ogni spazio metrico X non completo ammette un completamento \tilde{X} . Inoltre due qualsiasi completamenti sono tra loro omeomorfi tramite una isometria, cioè il completamento è unico a meno di isometrie.

Dim. 2.26 Vediamo subito l'unicità del completamento. Siano X_1 e X_2 i sottoinsiemi densi dei completamenti \tilde{X}_1 , \tilde{X}_2 , isometrici a X .

Allora siano $f_j : X \rightarrow X_j$, $j = 1, 2$ le due corrispondenti applicazioni isometriche (che sono su sulle loro immagini X_j). Esse sono invertibili e:

$$\begin{aligned} f_2 \circ f_1^{-1} &: X_1 \rightarrow X_2, \\ f_1 \circ f_2^{-1} &: X_2 \rightarrow X_1 \end{aligned}$$

sono una l'inversa dell'altra e isometriche. Ora $f_2 \circ f_1^{-1}$ è prolungabile univocamente in una applicazione $T : \tilde{X}_1 \rightarrow \tilde{X}_2$ continua e quindi isometrica. Analogamente $f_1 \circ f_2^{-1}$ è prolungabile univocamente in una applicazione $S : \tilde{X}_2 \rightarrow \tilde{X}_1$ continua e isometrica.

Sia ora $x \in \tilde{X}_1$, allora esiste una successione x_j in X_1 con $x_j \rightarrow x$ in \tilde{X}_1 . Vale:

$$(S \circ f_2 \circ f_1^{-1})(x_j) = x_j,$$

e passando al limite per $j \rightarrow \infty$ si ha:

$$S(T(x)) = x.$$

Analogamente:

$$T(S(z)) = z, \quad \forall z \in \tilde{X}_2.$$

Le distanze sono preservate per cui T e S sono le isometrie (una l'inversa dell'altra) richieste fra \tilde{X}_1 e \tilde{X}_2 che rendono i due spazi omeomorfi.

Vediamo ora come individuare un completamento. Si consideri l'insieme delle successioni di Cauchy in X . Su questo insieme poniamo una relazione di equivalenza:

$$\{x_n\} \sim \{y_n\} \quad \text{se} \quad \lim_{n \rightarrow \infty} d(x_n, y_n) = 0.$$

È chiaro che si tratta di una relazione di equivalenza che divide l'insieme delle successioni di Cauchy in classi di equivalenza $[\{x_n\}]$. Sia \tilde{X} l'insieme di queste classi. Possiamo costruire una corrispondenza ϕ fra X e un sottoinsieme di \tilde{X} :

$$\phi(x) = [\{x\}],$$

che associa sostanzialmente a x la successione stazionaria $\{x_n = x\}$ e tutte le successioni equivalenti. La corrispondenza ϕ è 1-1 perchè:

$$\phi(x) = \phi(y) \Rightarrow [\{x\}] = [\{y\}] \Rightarrow x = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = y.$$

Dotiamo ora \tilde{X} di una metrica. Siano $\{x_n\}$ e $\{y_n\}$ due successioni di Cauchy. Abbiamo:

$$d(x_n, y_n) \leq d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n),$$

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_m, y_n).$$

Essendo le successioni di Cauchy, anche la successione, in \mathbb{R} , $\{d(x_n, y_n)\}$ è di Cauchy e quindi convergente (\mathbb{R} è completo), cioè esiste:

$$\tilde{d} = \lim d(x_n, y_n).$$

Se $\{x_n\} \sim \{x'_n\}$ e $\{y_n\} \sim \{y'_n\}$, allora:

$$d(x_n, y_n) \leq d(x_n, x'_n) + d(x'_n, y'_n) + d(y'_n, y_n),$$

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y'_n, y_n).$$

Per cui \tilde{d} è indipendente dalle successioni scelte all'interno delle classi di equivalenza, ma dipende solo dalle classi $[\{x_n\}]$ e $[\{y_n\}]$. Definiamo allora:

$$\tilde{d}([\{x_n\}], [\{y_n\}]) = \lim_{n \rightarrow \infty} d(x_n, y_n),$$

e lo spazio \tilde{X} risulta metrico. Infatti possiamo verificare le proprietà metriche:

$$\tilde{d}([\{x_n\}], [\{y_n\}]) \geq 0.$$

$$\begin{aligned} \tilde{d}([\{x_n\}], [\{y_n\}]) = 0 &\implies \lim d(x_n, y_n) = 0 \\ &\implies \{x_n\} \sim \{y_n\} \implies [\{x_n\}] = [\{y_n\}]. \end{aligned}$$

$$\tilde{d}([\{x_n\}], [\{y_n\}]) = \tilde{d}([\{y_n\}], [\{x_n\}]).$$

$$\begin{aligned} \tilde{d}([\{x_n\}], [\{y_n\}]) &= \lim d(x_n, y_n) \leq \lim d(x_n, z_n) + \lim d(z_n, y_n) \\ &= \tilde{d}([\{x_n\}], [\{z_n\}]) + \tilde{d}([\{z_n\}], [\{y_n\}]). \end{aligned}$$

Considerando $X_1 = \phi(X)$, è chiaro che ϕ è una isometria:

$$\tilde{d}(\phi(x), \phi(y)) = \lim_{n \rightarrow \infty} d(x, y) = d(x, y).$$

Vediamo ora che X_1 è denso in \tilde{X} . Sia $[\{x_n\}] \in \tilde{X}$, abbiamo:

$$\tilde{d}([\{x_n\}], \phi(x_m)) = \lim_{n \rightarrow \infty} d(x_n, x_m).$$

Essendo $\{x_n\}$ di Cauchy, dato $\epsilon > 0$, possiamo imporre $d(x_n, x_m) < \epsilon$ per $n, m \geq N(\epsilon)$, da cui:

$$\tilde{d}([\{x_n\}], \phi(x_m)) \leq \epsilon,$$

(ogni successione di Cauchy è approssimata dai suoi elementi).

Rimane da vedere la completezza di \tilde{X} . Sia \tilde{x}_n una successione di Cauchy in \tilde{X} . Per ogni n intero e positivo possiamo scegliere $x_n \in X$ tale che:

$$\tilde{d}(\tilde{x}_n, \phi(x_n)) < \frac{1}{n}.$$

Gli elementi x_n formano essi stessi una successione di Cauchy:

$$\begin{aligned} d(x_n, x_m) &= \tilde{d}(\phi(x_n), \phi(x_m)) \leq \tilde{d}(\phi(x_n), \tilde{x}_n) + \tilde{d}(\tilde{x}_n, \tilde{x}_m) + \tilde{d}(\tilde{x}_m, \phi(x_m)) \\ &< \frac{1}{n} + \frac{1}{m} + \tilde{d}(\tilde{x}_n, \tilde{x}_m). \end{aligned}$$

E abbiamo che $\tilde{x} = [\{x_n\}]$ costituisce proprio il limite della successione \tilde{x}_n :

$$\tilde{d}(\tilde{x}_n, \tilde{x}) \leq \tilde{d}(\tilde{x}_n, \phi(x_n)) + \tilde{d}(\phi(x_n), \tilde{x}) < \frac{1}{n} + \lim_{k \rightarrow \infty} d(x_n, x_k) \xrightarrow{n \rightarrow \infty} 0.$$

Essendo \tilde{x}_n una successione di Cauchy arbitraria, \tilde{X} risulta completo.

— ■ —

2.6 Esercizi

Esercizio 2.1 Si dia un esempio di spazio metrico che contiene due sfere $S(x, r_1)$ e $S(y, r_2)$, con $x \neq y$ e $r_1 > r_2$, tali che la sfera di raggio minore contenga la sfera di raggio maggiore:

$$S(x, r_1) \subset S(y, r_2).$$

— ○ —

Esercizio 2.2 Sia (X, d) uno spazio metrico. Si dimostri che:

a) la funzione

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)},$$

definisce un'altra metrica su X che rende X limitato;

b) ogni spazio metrico è omeomorfo ad uno spazio metrico limitato.

— ○ —

Esercizio 2.3 Si consideri lo spazio delle successioni limitate:

$$E = \left\{ x = \{x_n\}_{n=0}^{\infty} ; \sup_k |x_k| < \infty \right\},$$

a) Si dimostri che E è uno spazio metrico completo con la distanza:

$$d(x, y) = \sup_k |x_k - y_k|.$$

b) Sia dato il sottoinsieme $S \subset E$ formato dalle successioni x i cui elementi x_k possono assumere solo i valori interi positivi $0, 1, \dots, 9$. Si calcoli la minima e la massima distanza tra due punti distinti di S e si dimostri che S non è denso in E .

— ○ —

Esercizio 2.4 Siano X, Y due spazi metrici e $\{f_n\}_{n \in \mathbb{N}}$ una successione di funzioni continue da X a Y tale che f_n converge uniformemente ad una funzione $f : X \rightarrow Y$. Si dimostri che:

a) f è continua;

b) se $x_n \rightarrow x$ in X quando $n \rightarrow \infty$, allora $f_n(x_n) \rightarrow f(x)$ in Y .

— ○ —

Esercizio 2.5 Si consideri lo spazio delle funzioni continue $C([a, b])$ su un intervallo $[a, b]$ e a valori reali (o complessi).

a) Si dimostri che la seguente funzione definisce una metrica su di esso:

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|,$$

e che lo spazio è completo.

b) Si stabilisca se il seguente sottoinsieme è aperto o chiuso e se è denso in $C([a, b])$:

$$M = \{ f \in C([a, b]) ; f(x_0) = k \},$$

essendo x_0 un punto dell'intervallo $[a, b]$ e k un numero reale (complesso) fissati.

— ○ —

Esercizio 2.6 Si consideri lo spazio metrico completo $C^0([-1, 1])$ delle funzioni continue dall'intervallo $[-1, 1]$ a valori complessi, con la distanza:

$$d(f, g) = \sup_{-1 \leq x \leq 1} |f(x) - g(x)|.$$

a) Si determini se la successione di funzioni:

$$f_n(x) = \sqrt{\frac{1}{n^2} + x^2}, \quad n = 1, 2, \dots$$

è di Cauchy. Si dica poi se è convergente e, in caso affermativo, se ne determini il limite.

b) Si determini se il sottospazio $C^1([-1, 1])$ delle funzioni derivabili con derivata prima continua nell'intervallo $[-1, 1]$ è chiuso.

— ○ —

2.6.1 Soluzioni

Soluzione 2.1 Si consideri come spazio metrico la semiretta positiva dell'asse reale con l'ordinaria distanza tra numeri reali:

$$\mathbb{R}^+ = \{x \in \mathbb{R}; x \geq 0\},$$

$$d(x, y) = |x - y|.$$

Considerando le due sfere:

$$S(0, 2) = \left\{ x \in \mathbb{R}^+; d(x, 0) < 2 \right\} = \left\{ 0 \leq x < 2 \right\},$$

$$S(1, 3/2) = \left\{ x \in \mathbb{R}^+; d(x, 1) < \frac{3}{2} \right\} = \left\{ 0 \leq x < \frac{5}{2} \right\},$$

abbiamo effettivamente $S(0, 2) \subset S(1, 3/2)$.

— ○ —

Soluzione 2.2 a) La funzione d' definisce una metrica in quanto:

- è definita positiva:

$$d'(x, y) \geq 0, \quad d'(x, y) = 0 \Leftrightarrow x = y,$$

poichè il denominatore è sempre una quantità strettamente positiva, mentre il numeratore è maggiore o uguale a zero e si annulla solo quando $x = y$;

- è simmetrica:

$$d'(x, y) = d'(y, x),$$

in quanto rapporto di funzioni simmetriche;

- soddisfa la diseguaglianza triangolare. Infatti, essendo d una metrica, abbiamo:

$$\begin{aligned} d(x, y) &\leq d(x, z) + d(z, y), \\ 1 + \frac{1}{d(x, y)} &\geq 1 + \frac{1}{d(x, z) + d(z, y)}, \\ \frac{1}{1 + \frac{1}{d(x, y)}} &\leq \frac{1}{1 + \frac{1}{d(x, z) + d(z, y)}}, \\ \frac{d(x, y)}{d(x, y) + 1} &\leq \frac{d(x, z) + d(z, y)}{d(x, z) + d(z, y) + 1} \\ &= \frac{d(x, z)}{d(x, z) + d(z, y) + 1} + \frac{d(z, y)}{d(x, z) + d(z, y) + 1} \\ &\leq \frac{d(x, z)}{d(x, z) + 1} + \frac{d(z, y)}{d(z, y) + 1}, \end{aligned}$$

$$d'(x, y) \leq d'(x, z) + d'(z, y).$$

Infine osserviamo che:

$$d'(x, y) < 1 \quad \forall x, y \in X,$$

per cui tutto lo spazio è contenuto nella sfera unitaria.

È sufficiente mostrare che gli spazi (X, d) e (X, d') sono tra loro omeomorfi tramite la trasformazione identica:

$$f : x \mapsto x,$$

chiaramente invertibile e coincidente con la sua inversa. Essa è un omeomorfismo se sia f che f^{-1} sono continue. Ragionando con gli aperti questo equivale a dire che un insieme aperto nello spazio topologico (X, d) è aperto anche nello spazio topologico (X, d') e viceversa. Le due topologie sono costruite a partire dalle sfere aperte con centro un punto qualunque $x_0 \in X$ e raggio arbitrario:

$$S(x_0, r) = \{x \in X ; d(x_0, x) < r\},$$

$$S'(x_0, r) = \{x \in X ; d'(x_0, x) < r\},$$

e le topologie sono equivalenti se, data una sfera $S(x_0, r) \in (X, d)$, possiamo determinare due sfere $S'(x_0, r'_1)$ $S'(x_0, r'_2)$ nello spazio topologico (X, d') tali che:

$$S'(x_0, r'_1) \subseteq S(x_0, r) \subseteq S'(x_0, r'_2).$$

Notiamo innanzitutto che $d'(x, y) \leq d(x, y)$, quindi, se $x \in S(x_0, r)$, allora $d'(x_0, x) \leq d(x_0, x) < r$ e $x \in S'(x_0, r)$. È quindi sufficiente prendere $r'_2 = r$.

D'altra parte, se $x \in S'(x_0, r'_1)$, allora:

$$d(x_0, x) = \frac{d'(x_0, x)}{1 - d'(x_0, x)} < \frac{r'_1}{1 - r'_1}.$$

È quindi sufficiente imporre:

$$r = \frac{r'_1}{1 - r'_1},$$

$$r'_1 = \frac{r}{1 + r}.$$

— ○ —

Soluzione 2.3 $d(x, y)$ definisce una metrica in quanto, per ogni $x, y, z \in E$:

$$d(x, y) \geq 0, \quad d(x, y) = 0 \iff x = y,$$

$$d(x, y) = \sup_k |x_k - y_k| = \sup_k |y_k - x_k| = d(y, x),$$

$$\begin{aligned} d(x, y) &= \sup_k |x_k - y_k| = \sup_k |x_k - z_k + z_k - y_k| \\ &\leq \sup_k |x_k - z_k| + \sup_k |z_k - y_k| = d(x, z) + d(z, y). \end{aligned}$$

Per dimostrare che E è completo consideriamo in esso una successione di Cauchy $\{x^{(n)}\}_{n \in \mathbb{N}}$:

$$d(x^{(m)}, x^{(n)}) < \epsilon, \quad \forall m, n > \bar{n},$$

(\bar{n} dipende solo da ϵ). Per ogni k , la successione di numeri (reali o complessi) $\{x_k^{(n)}\}_{n \in \mathbb{N}}$ (k fissato), è di Cauchy in quanto:

$$|x_k^{(m)} - x_k^{(n)}| \leq \sup_k |x_k^{(m)} - x_k^{(n)}| = d(x^{(m)}, x^{(n)}) < \epsilon \quad m, n > \bar{n}.$$

Ne segue che, essendo \mathbb{R} (o \mathbb{C}) spazi completi, esiste per ogni k il limite:

$$\lim_{n \rightarrow \infty} x_k^{(n)} = x_k,$$

che definisce una successione $\{x_k\}_{k=0}^{\infty}$. Inoltre, prendendo il limite per $n \rightarrow \infty$, abbiamo anche:

$$|x_k^{(m)} - x_k| \leq \epsilon, \quad \forall m > \bar{n},$$

e, prendendo l'estremo superiore al variare di k :

$$d(x^{(m)}, x) \leq \epsilon, \quad m > \bar{n}$$

questo comporta che x è una successione limitata ($x \in E$) e:

$$x = \lim_{m \rightarrow \infty} x^{(m)},$$

nella topologia dello spazio E .

Consideriamo ora il sottoinsieme S . Se $x, y \in S$, ovvero x_k, y_k possono assumere solo un numero finito di valori, lo stesso vale per le differenze $x_k - y_k$, e deve esistere almeno un \bar{k} tale che:

$$d(x, y) = \sup_k |x_k - y_k| = |x_{\bar{k}} - y_{\bar{k}}|.$$

Inoltre $|x_{\bar{k}} - y_{\bar{k}}|$ assume valore minimo 1 (dobbiamo avere $x \neq y$) e valore massimo 9:

$$d_{min} = 1, \quad d_{max} = 9.$$

Ricordiamo che S è denso in E se e solo se ogni intorno di un punto arbitrario $x_0 \in E$ contiene almeno un punto di S . Il fatto che la distanza minima tra i punti di S distinti è finita comporta che S non può essere denso in E : prendiamo ad esempio $x_0 = (1/2, 0, 0, \dots)$ e come intorno la sfera $S(x_0, 1/4)$ che non contiene alcun punto di S . Notiamo, per inciso, che S è un insieme con cardinalità infinita e non numerabile in quanto i suoi elementi possono essere posti in corrispondenza biunivoca con l'intervallo reale $[0, 1]$ (scrivendo un numero in forma decimale, la sequenza dei decimali definisce una successione in S).

— ○ —

Soluzione 2.4

a) Ricordiamo che $f_n \rightarrow f$ per $n \rightarrow \infty$ uniformemente se:

$$\sup_{x \in X} d_Y(f_n(x), f(x)) < \epsilon \quad \forall n > \bar{n},$$

e che la convergenza uniforme implica quella puntuale:

$$f_n(x) \rightarrow f(x) \quad \forall x \in X.$$

Per la diseguaglianza triangolare si ha:

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(x_0)) + d_Y(f_n(x_0), f(x_0)).$$

D'altra parte, se $n > \bar{n}$ abbiamo:

$$d_Y(f(x), f_n(x)) < \epsilon, \quad d_Y(f_n(x_0), f(x_0)) < \epsilon,$$

in quanto f_n converge puntualmente (e uniformemente) a f , e:

$$d_Y(f_n(x), f_n(x_0)) < \epsilon \quad \text{se } d_X(x, x_0) < \delta,$$

(δ può dipendere dal punto x_0 e da ϵ), essendo f_n continua per ipotesi. Pertanto $d_Y(f(x), f(x_0))$ può essere resa piccola a piacere e f è continua.

b) Per la diseguaglianza triangolare si ha:

$$d_Y(f_n(x_n), f(x)) \leq d_Y(f_n(x_n), f(x_n)) + d(f(x_n), f(x)).$$

Ora, poichè f_n converge puntualmente e uniformemente a f abbiamo:

$$d_Y(f_n(x_n), f(x_n)) < \epsilon \quad \text{se } n > \bar{n} \text{ (indipendente da } x_n\text{)},$$

e, essendo f continua:

$$d_Y(f(x_n), f(x)) < \epsilon \quad \text{se } d_X(x_n, x) < \delta,$$

con δ dipendente da ϵ e x in generale. Risulta quindi verificata l'affermazione.
— ○ —

Soluzione 2.5 $d(f, g)$ definisce una metrica poichè per ogni $f, g, h \in C([a, b])$:

$$d(f, g) \geq 0, \quad d(f, g) = 0 \iff f = g;$$

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)| = \sup_{x \in [a, b]} |g(x) - f(x)| = d(g, f),$$

$$\begin{aligned} d(f, g) &= \sup_{x \in [a, b]} |f(x) - g(x)| = \sup_{x \in [a, b]} |f(x) - h(x) + h(x) - g(x)| \\ &\leq \sup_{x \in [a, b]} |f(x) - h(x)| + \sup_{x \in [a, b]} |h(x) - g(x)| = d(f, h) + d(h, g). \end{aligned}$$

Per dimostrare che $C([a, b])$ è completo consideriamo in esso una successione di Cauchy $\{f_n\}_{n \in \mathbb{N}}$:

$$d(f_m, f_p) < \epsilon, \quad \forall m, p > \bar{n}.$$

Per ogni $x \in [a, b]$, la successione di numeri reali (o complessi) $\{f_n(x)\}_{n \in \mathbb{N}}$ è di Cauchy:

$$|f_m(x) - f_p(x)| \leq \sup_{t \in [a, b]} |f_m(t) - f_p(t)| = d(f_m, f_p) < \epsilon \quad \forall m, p > \bar{n}.$$

Ne segue che essendo \mathbb{R} (o \mathbb{C}) completi, $f_n(x)$ è convergente e definisce una funzione f :

$$f(x) = \lim_{n \rightarrow \infty} f_n(x).$$

Operando il limite per $p \rightarrow \infty$ abbiamo anche che:

$$|f_m(x) - f(x)| \leq \epsilon, \quad \forall m > \bar{n}$$

(\bar{n} non dipende da x) e prendendo l'estremo superiore nell'intervallo $[a, b]$:

$$d(f_m, f) \leq \epsilon \quad \forall m > \bar{n},$$

pertanto f_n converge uniformemente a f grazie alla metrica introdotta in $C([a, b])$ e la funzione $f(x)$ è limitata. Resta da mostrare la continuità. Abbiamo:

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|. \end{aligned}$$

D'altra parte, per ogni $x \in [a, b]$ abbiamo $|f(x) - f_n(x)| < \epsilon$ se $n > \bar{n}$ (indipendente da x) e $|f_n(x) - f_n(x_0)| < \epsilon$ quando $|x - x_0| < \delta_\epsilon$, essendo $f_n(x)$ una funzione continua su un intervallo chiuso $[a, b]$ (e quindi uniformemente continua).

Considerando la parte (b) ricordiamo che se una successione f_n di funzioni in $C([a, b])$ converge uniformemente (ovvero nella topologia definita dalla metrica qui usata) ad una funzione $f \in C([a, b])$ allora essa converge anche puntualmente: $f_n(x) \rightarrow f(x)$ per ogni $x \in [a, b]$.

Per vedere se M è aperto o chiuso, consideriamo una successione di Cauchy $\{f_n\}_{n \in \mathbb{N}}$ di funzioni in M , ovvero tali che $f_n(x_0) = k$. Per quanto visto in (a), tale successione converge ad una funzione $f \in C([a, b])$. D'altra parte la successione numerica $f_n(x_0)$ è stazionaria, $f_n(x_0) = k$ per ogni n , e quindi necessariamente anche la funzione limite soddisfa $f(x_0) = k$, cioè f appartiene a M che è quindi un insieme chiuso (contiene i suoi punti di accumulazione).

L'insieme M non può essere denso in $C([a, b])$, in quanto possiamo considerare una funzione $g \in C([a, b])$ con $g(x_0) \neq k$ e questa non può appartenere alla chiusura $M^- = M$ (non può essere limite di funzioni in M).

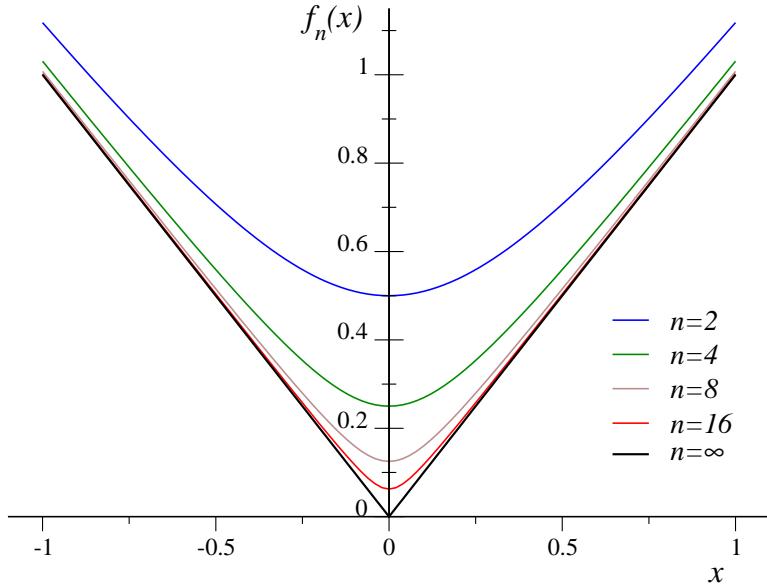


Figura 2.8: Esercizio 2.6.

Soluzione 2.6 La successione è di Cauchy in quanto (vedi figura 2.8):

$$d(f_n, f_m) = \sup_{-1 \leq x \leq 1} \left| \sqrt{\frac{1}{n^2} + x^2} - \sqrt{\frac{1}{m^2} + x^2} \right| = \left| \frac{1}{n} - \frac{1}{m} \right| \xrightarrow{n, m \rightarrow \infty} 0.$$

È quindi convergente perché lo spazio $C^0([-1, 1])$ è completo. In effetti il limite puntuale coincide con il limite uniforme e la successione converge alla funzione:

$$f(x) = \sqrt{x^2} = |x|.$$

La successione del punto (a) è una successione di Cauchy di funzioni derivabili con derivata continua. Questa successione converge però ad una funzione che non è derivabile in $x = 0$. Il sottospazio in questione non è quindi chiuso.

— ○ —

Capitolo 3.

Spazi lineari.

3.1 Strutture algebriche.

Vogliamo ora generalizzare le buone proprietà degli spazi \mathbb{R}^n e \mathbb{C}^n al caso di spazi vettoriali astratti, definiti assiomaticamente. In particolare, gli spazi vettoriali che definiremo possono avere un numero arbitrario di dimensioni. Questa generalizzazione è di grande interesse per le applicazioni fisiche in quanto è spesso necessario considerare spazi a dimensioni infinite per la descrizione matematica di un sistema. Un esempio è fornito dalla trattazione quantistica dell'atomo di idrogeno, che pur costituendo un sistema molto semplice, implica la costruzione di uno spazio vettoriale infinito-dimensionale.

Per giungere alla definizione di spazio vettoriale ripercorriamo la costruzione delle varie strutture algebriche astratte, senza soffermarci troppo su di esse.

Def. 3.1 (Gruppo) Sia X un insieme in cui sia definita una **operazione binaria** (detta anche **legge di composizione**) \cdot che ad ogni coppia di elementi $a, b \in X$ associa un elemento di X indicato con $a \cdot b$, o semplicemente $a b$:

$$\cdot : (a, b) \longrightarrow a \cdot b. \quad (3.1)$$

Allora l'insieme X si chiama **gruppo** se l'operazione binaria verifica le condizioni:

i) Vale la **legge associativa**:

$$\forall a, b, c \in X, \quad (a \cdot b) \cdot c = a \cdot (b \cdot c). \quad (3.2)$$

ii) Esiste un **elemento neutro** $u \in X$ tale che:

$$a \cdot u = u \cdot a = a, \quad \forall a \in X. \quad (3.3)$$

iii) Ogni elemento $a \in X$ ammette l'**elemento inverso** a^{-1} tale che:

$$a \cdot a^{-1} = u = a^{-1} \cdot a. \quad (3.4)$$

Per mettere in evidenza la legge di composizione binaria il gruppo viene a volte indicato con la coppia (X, \cdot) .

Def. 3.2 Un gruppo X si dice **abeliano** se vale la **legge commutativa**, cioè se:

$$a \cdot b = b \cdot a \quad \forall a, b \in X. \quad (3.5)$$

Osservazione. Quando X è un gruppo abeliano, spesso l'operazione binaria viene detta **addizione** e indicata con $+$. In tal caso l'elemento neutro viene detto **zero** e indicato con 0 . L'inverso viene indicato con $-a$ (detto **opposto**) e si pone (per convenzione):

$$a - b = a + (-b).$$

Il gruppo X viene anche detto, in questo caso, **gruppo additivo**.

Def. 3.3 Una applicazione $f : X \rightarrow Y$ da un gruppo X ad un altro Y si chiama **omomorfismo** se preserva l'operazione:

$$f(a \cdot b) = f(a) \cdot f(b), \quad \forall a, b \in X. \quad (3.6)$$

Se f è anche biettiva allora f viene detta un **isomorfismo** e i due gruppi X e Y sono detti **isomorfi**.

Def. 3.4 Se $f : X \rightarrow Y$ è un omomorfismo, allora il **nucleo** di f è definito come l'insieme degli elementi di X che si trasformano, tramite f , nell'elemento neutro u di Y :

$$\mathcal{N}(f) = \{a \in X ; f(a) = u\}. \quad (3.7)$$

Def. 3.5 (Anello) Sia \mathbb{A} un insieme non vuoto in cui siano definite due operazioni binarie, una detta **addizione** (o a volte **somma**) $+$, e una detta **moltiplicazione** (o anche **prodotto**) \cdot :

$$+ : \mathbb{A} \times \mathbb{A} \rightarrow \mathbb{A}, \quad (a, b) \mapsto a + b, \quad (3.8)$$

$$\cdot : \mathbb{A} \times \mathbb{A} \rightarrow \mathbb{A}, \quad (a, b) \mapsto a \cdot b = ab. \quad (3.9)$$

\mathbb{A} viene detto **anello** se sono verificate le proprietà:

- i) \mathbb{A} è un gruppo abeliano rispetto alla addizione.
- ii) Vale la **proprietà associativa** per la moltiplicazione:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \forall a, b, c \in \mathbb{A}. \quad (3.10)$$

iii) Vale la **proprietà distributiva** della moltiplicazione rispetto alla addizione:

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c & \forall a, b, c \in \mathbb{A}. \\ (b + c) \cdot a &= b \cdot a + c \cdot a \end{aligned} \quad (3.11)$$

Spesso l'anello viene indicato con la notazione $(\mathbb{A}, +, \cdot)$ per mettere in evidenza le operazioni binarie della sua struttura.

Def. 3.6 Un anello \mathbb{A} è detto **commutativo** se la proprietà commutativa vale anche per la moltiplicazione:

$$a \cdot b = b \cdot a \quad \forall a, b \in \mathbb{A}. \quad (3.12)$$

Un anello \mathbb{A} è detto **unitario** se esiste un elemento neutro, detto **unità** e indicato con 1, per la moltiplicazione:

$$\exists 1 \in \mathbb{A}; \quad 1 \cdot x = x \cdot 1 = x \quad \forall x \in \mathbb{A}. \quad (3.13)$$

Def. 3.7 (Campo) Un anello **commutativo** e **unitario** \mathbb{K} è definito **campo** se ogni elemento $a \in \mathbb{K}$, con a diverso dall'elemento neutro additivo (lo **zero**), $a \neq 0$, ammette un inverso a^{-1} rispetto alla moltiplicazione:

$$\forall a \in \mathbb{K}, a \neq 0 \quad \exists a^{-1} \in \mathbb{K}, \quad a \cdot a^{-1} = a^{-1} \cdot a = 1. \quad (3.14)$$

L'inverso moltiplicativo a^{-1} di a viene spesso detto **reciproco** e indicato anche con $\frac{1}{a}$, per distinguerlo dall'inverso additivo, detto **opposto** e indicato con $-a$.

Osservazione. Esempi importanti di campi sono l'insieme dei numeri reali \mathbb{R} e l'insieme dei numeri complessi \mathbb{C} . In pratica saranno gli unici campi che useremo in seguito.

3.1.1 Spazi vettoriali

Dopo queste premesse giungiamo alla definizione vera e propria di spazio vettoriale.

Def. 3.8 (Spazio vettoriale) Sia \mathbb{K} un campo, i cui elementi sono detti **scalari**, V un gruppo abeliano additivo, i cui elementi sono detti **vettori**. Se esiste una applicazione, detta **moltiplicazione per uno scalare** (per la quale non useremo alcun simbolo particolare):

$$\begin{aligned} \mathbb{K} \times V &\longrightarrow V, \\ (k, x) &\longrightarrow kx, \end{aligned} \quad (3.15)$$

tale che:

- i) vale la **proprietà distributiva** della moltiplicazione per uno scalare rispetto alla somma tra vettori:

$$k(u + v) = ku + kv, \quad \forall k \in \mathbb{K}, u, v \in V; \quad (3.16)$$

- ii) vale la **proprietà distributiva** della moltiplicazione per uno scalare rispetto alla somma tra scalari:

$$(k + h)u = ku + hu, \quad \forall k, h \in \mathbb{K}, u \in V; \quad (3.17)$$

- iii) vale la **proprietà associativa**:

$$(kh)u = k(hu), \quad \forall k, h \in \mathbb{K}, u \in V; \quad (3.18)$$

- iv) l'unità 1 di \mathbb{K} risulta elemento neutro anche per la moltiplicazione per uno scalare:

$$1u = u, \quad \forall u \in V; \quad (3.19)$$

allora V viene detto **spazio vettoriale** o **spazio lineare** sul campo \mathbb{K} . Se $\mathbb{K} = \mathbb{R}$ lo spazio vettoriale viene detto **spazio vettoriale reale**, mentre se $\mathbb{K} = \mathbb{C}$, lo spazio è detto **spazio vettoriale complesso**.

Abbiamo usato le stesse notazioni per indicare sia la somma tra vettori che la somma tra scalari, come pure le stesse notazioni per la moltiplicazione tra scalari e la moltiplicazione per uno scalare. Risulta evidente che si tratta di operazioni distinte, ma compatibili tra loro, e risulterà sempre chiaro dal contesto di quale operazione si tratta.

Osservazione. Notiamo che nella definizione di spazio vettoriale non si fa uso del fatto che \mathbb{K} sia un anello commutativo dotato di inverso per ogni elemento non nullo. In effetti si potrebbero richiedere esattamente le stesse proprietà con un anello unitario \mathbb{K} invece di un campo. In questo modo si otterrebbe una struttura nota come **\mathbb{K} -modulo**, che rappresenta una generalizzazione del concetto di spazio vettoriale (gli scalari appartengono ad un anello invece che ad un campo).

Osservazione. La moltiplicazione di un vettore per uno scalare è definita con lo scalare a sinistra del vettore, ma è convenzione comune dare significato anche al prodotto con lo scalare a destra del vettore assumendo:

$$vh \equiv hv \quad h \in \mathbb{K}, v \in V,$$

ottenendo in ogni caso il **multiplo** del vettore v secondo il fattore scalare di proporzionalità h .

Vediamo qualche esempio di spazio vettoriale.

Esempio 3.1 Sia $V = \mathbb{R}^n$ formato dalle n -uple di numeri reali (x_1, \dots, x_n) , $x_j \in \mathbb{R}$, $j = 1, 2, \dots, n$. \mathbb{R}^n diviene uno spazio vettoriale reale definendo la somma tra vettori e la moltiplicazione per uno scalare tramite le relazioni:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

$$\alpha(x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n), \quad \alpha \in \mathbb{R}.$$

— ○ —

Esempio 3.2 Sia \mathbb{C}^n formato dalle n -uple di numeri complessi (z_1, \dots, z_n) , esso diviene uno spazio vettoriale complesso definendo la somma tra vettori e la moltiplicazione per uno scalare formalmente nella stessa maniera di \mathbb{R}^n :

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

$$\alpha(z_1, z_2, \dots, z_n) = (\alpha z_1, \alpha z_2, \dots, \alpha z_n), \quad \alpha \in \mathbb{C}.$$

Se lo scalare α viene ristretto a \mathbb{R} , mantenendo sempre come vettori gli elementi di \mathbb{C}^n , si ottiene ancora uno spazio vettoriale, reale anziché complesso (ma di dimensione doppia).

— ○ —

Esempio 3.3 Sia Y un insieme qualsiasi (non vuoto) e consideriamo l'insieme delle funzioni da Y a \mathbb{R} :

$$\mathbb{R}^Y = \{f : Y \longrightarrow \mathbb{R}\}.$$

Esso diviene uno spazio vettoriale reale se lo strutturiamo con le usuali operazioni algebriche tra funzioni:

$$(f + g)(y) = f(y) + g(y),$$

$$(\alpha f)(y) = \alpha f(y), \quad \alpha \in \mathbb{R}.$$

— ○ —

Osservazione. Con la notazione B^A , con A e B insiemi, si denota l'insieme di tutte le applicazioni da A a B , e la notazione deriva dal fatto che se A e B sono finiti, anche l'insieme delle possibili applicazioni da A a B è finito, e la sua cardinalità $\#$ è proprio espressa dalla relazione:

$$\#(B^A) = (\#B)^{\#A}$$

Esempio 3.4 Sia $\mathbb{K} = \mathbb{R}$ oppure \mathbb{C} , a scelta. Definiamo X come l'insieme delle sequenze (infinte) $(x_1, x_2, \dots, x_n, \dots)$ di elementi di \mathbb{K} in cui solo un numero finito di elementi x_j è non nullo. Con le definizioni abituali di somma e moltiplicazione per uno scalare, X diviene uno spazio vettoriale su \mathbb{K} . Si possono verificare come esercizio tutte le proprietà di spazio vettoriale.

— ○ —

Essendo uno spazio vettoriale un gruppo abeliano additivo, il suo elemento neutro viene indicato con il simbolo 0 , oppure, quando si vuole mettere in evidenza la sua natura vettoriale, con i simboli $\mathbf{0}$ o $\vec{0}$, distinguendolo dall'elemento neutro additivo del campo \mathbb{K} , indicato sempre con 0 .

Osservazione. Risulta banale verificare che uno spazio costituito dal solo vettore nullo, $\{0\}$, costituisce uno spazio vettoriale, denotato come **spazio nullo**, ma in seguito, quando considereremo uno spazio vettoriale, assumeremo generalmente uno spazio non nullo, a meno di non menzionare esplicitamente il fatto che sia uno spazio nullo.

Le definizioni di campo di scalari e spazio vettoriale comportano alcune conseguenze di natura algebrica intuitive.

- A) Gli elementi neutri (sia additivi che moltiplicativi) sono unici e per ogni coppia di vettori x, y esiste un unico vettore z tale che:

$$x = y + z, \quad (3.20)$$

e vale:

$$z = x - y. \quad (3.21)$$

- B) Per ogni vettore x e per ogni scalare α :

$$0x = \vec{0}, \quad \alpha\vec{0} = \vec{0}. \quad (3.22)$$

- C) La relazione (α scalare, x vettore):

$$\alpha x = 0, \quad (3.23)$$

implica che o $\alpha = 0$, oppure $x = 0$.

È ovvio inoltre che la somma di un numero qualsiasi (finito) di elementi (sia per gli scalari che per i vettori) è univocamente definita indipendentemente dall'ordine con il quale eseguire le operazioni.

3.1.2 Sottospazi

La struttura di spazio vettoriale può indurre tale proprietà anche su un sottoinsieme, che può divenire spazio vettoriale se strutturato con le medesime operazioni, e per il quale si parla di sottospazio vettoriale. Non è però necessario verificare tutte le proprietà di spazio vettoriale per il sottoinsieme, ma è sufficiente garantire la chiusura dell'insieme rispetto alle operazioni di addizione tra vettori e di moltiplicazione per uno scalare. Tutte le proprietà di spazio vettoriale risultano automaticamente verificate in quanto valide per lo spazio totale. Possiamo allora dare la seguente definizione equivalente di sottospazio:

Def. 3.9 Un sottoinsieme L di uno spazio vettoriale X su un campo \mathbb{K} è un **sottospazio vettoriale** (o **sottospazio lineare**) se verifica le proprietà:

1. chiusura (algebrica) rispetto alla addizione fra vettori:

$$x, y \in L \implies x + y \in L; \quad (3.24)$$

2. chiusura (algebrica) rispetto alla moltiplicazione per uno scalare:

$$\alpha \in \mathbb{K}, x \in L \implies \alpha x \in L. \quad (3.25)$$

Osservazione. La chiusura algebrica di cui sopra non è da confondere con la chiusura in senso topologico. Non abbiamo ancora introdotto alcuna topologia in uno spazio lineare.

Le proprietà di spazio vettoriale per il sottoinsieme L sono garantite in pratica se L è chiuso rispetto alle combinazioni lineari (finite), cioè possiamo inglobare le condizioni (3.24) e (3.25) in una unica condizione:

$$\alpha, \beta \in \mathbb{K}, x, y \in L \implies \alpha x + \beta y \in L. \quad (3.26)$$

È immediato verificare che se abbiamo due sottospazi vettoriali L_1 e L_2 di un medesimo spazio vettoriale V , la loro intersezione (non vuota in quanto contiene almeno il vettore nullo) risulta un sottospazio. Al massimo l'intersezione contiene solo il vettore nullo, e diremo che i due sottospazi sono **disgiunti**, ma risulta in ogni caso un sottospazio.

Per quanto riguarda l'unione, in generale l'unione di sottospazi non è più un sottospazio, ma si può sempre costruire un sottospazio che la contenga.

3.1.3 Dipendenze lineari e basi.

Tramite le operazioni definite nello spazio vettoriale si possono costruire combinazioni lineari di vettori, generando nuovi vettori dipendenti dai precedenti. Sorge allora il problema di determinare dei criteri per stabilire se un vettore può essere o non essere esprimibile come combinazione lineare di altri vettori assegnati. Questo porta alla definizione rigorosa del concetto di dipendenza o indipendenza tra vettori.

Def. 3.10 Sia X uno spazio vettoriale sul campo \mathbb{K} . Un insieme non vuoto $\Sigma \subseteq X$ di vettori non nulli è detto **linearmente indipendente** se l'annullarsi di una combinazione lineare finita di vettori arbitrari di Σ implica che i corrispondenti coefficienti della combinazione sono nulli ($x_j \in \Sigma, \alpha_j \in \mathbb{K}, j = 1, \dots, N$):

$$\sum_{j=1}^N \alpha_j x_j = 0 \implies \alpha_j = 0, j = 1, \dots, N. \quad (3.27)$$

Cioè l'annullarsi dei coefficienti di una combinazione lineare non è solo una condizione sufficiente ma anche necessaria per ottenere il vettore nullo. In caso contrario abbiamo un insieme di vettori dipendenti fra di loro.

Def. 3.11 Sia X uno spazio vettoriale sul campo \mathbb{K} . Un insieme non vuoto $\Sigma \subseteq X$ di vettori non nulli è detto **linearmente dipendente** se non è linearmente indipendente, cioè se esiste un numero finito di vettori $x_j \in \Sigma$, $j = 1, \dots, N$, tali che:

$$\sum_{j=1}^N \alpha_j x_j = 0, \quad \text{con } \alpha_j \neq 0, \quad j = 1, \dots, N. \quad (3.28)$$

Osservazione. Notiamo che non abbiamo posto alcuna condizione sulla cardinalità dell'insieme Σ , questo può quindi essere finito, infinito, numerabile o non numerabile, ma in ogni caso la verifica della dipendenza o indipendenza lineare coinvolge sempre un numero finito N di vettori.

Chiaramente in un sistema di vettori linearmente indipendenti questi sono non nulli e distinti fra loro. Infatti, se $x_1 = x_2$, abbiamo $x_1 - x_2 = 0$, e quindi troviamo una combinazione lineare nulla con coefficienti non nulli: $\alpha_1 = 1$, $\alpha_2 = -1$.

Considerando i possibili sistemi di vettori linearmente indipendenti possiamo distinguere due situazioni:

1. esistono dei sistemi di vettori linearmente indipendenti formati da un numero arbitrariamente grande di vettori;
2. oppure in ogni sistema di vettori linearmente indipendenti il numero di vettori che ne fanno parte è limitato, cioè il numero massimo di vettori linearmente indipendenti è finito.

Nel secondo caso diremo che lo spazio è **finito** dimensionale e il numero massimo n di vettori linearmente indipendenti viene chiamato dimensionalità dello spazio. Ciò significa che i gradi di libertà sono in numero finito ed ogni elemento dello spazio può essere espresso tramite un numero finito di parametri. Per convenzione assumiamo inoltre che lo spazio nullo abbia dimensione zero.

Nel primo caso abbiamo invece a che fare con uno spazio vettoriale detto a **infinte dimensioni**.

Def. 3.12 Sia X uno spazio vettoriale su un campo \mathbb{K} . Un insieme di vettori $\Sigma \subseteq X$ è detto una **base** se è un insieme linearmente indipendente massimale, ovvero tale che non può essere ingrandito con l'aggiunta di ulteriori vettori (senza perdere la proprietà di indipendenza lineare).

Teo. 3.1 Un sistema di vettori Σ è una base per uno spazio vettoriale X se e solo se Σ è formato da vettori linearmente indipendenti ed ogni $x \in X$ è esprimibile in maniera univoca come combinazione lineare finita di elementi di Σ :

$$x = \sum_{\alpha \in \mathcal{F}} c_\alpha x_\alpha, \quad x_\alpha \in \Sigma, \quad \mathcal{F} \text{ finito}, \quad (3.29)$$

Dim. 3.1 Se Σ è una base allora è formata per definizione da vettori linearmente indipendenti, e se aggiungiamo un vettore non nullo $x \in X$, $x \notin \Sigma$, $\{x\} \cup \Sigma$ non è più linearmente indipendente. Pertanto esistono degli scalari d, d_α , non tutti nulli ed in numero finito, tali che:

$$d x + \sum_{\alpha} d_\alpha x_\alpha = 0, \quad x_\alpha \in \Sigma.$$

La combinazione lineare è, in base alle definizioni di dipendenza e indipendenza lineare, finita, e il coefficiente d di x è non nullo (in caso contrario avremmo solo una combinazione lineare nulla di $x_\alpha \in \Sigma$, indipendenti tra loro, a coefficienti obbligatoriamente nulli). Pertanto:

$$x = \sum_{\alpha} \frac{(-d_\alpha)}{d} x_\alpha = \sum_{\alpha} c_\alpha x_\alpha.$$

Chiaramente, se $x \in \Sigma$, la relazione (3.29) è banalmente verificata (un solo termine, $x = x$), mentre se $x = 0$, i coefficienti $c_\alpha = 0$ verificano immediatamente e univocamente la relazione.

La combinazione lineare (3.29) che esprime x è inoltre unica. Assumiamo l'esistenza di due combinazioni lineari finite che esprimano il medesimo elemento x :

$$\begin{aligned} x &= \sum_{\alpha \in \mathcal{A}} a_\alpha x_\alpha && (\mathcal{A} \text{ finito}) \\ &= \sum_{\beta \in \mathcal{B}} b_\beta x_\beta && (\mathcal{B} \text{ finito}), \end{aligned}$$

con $x_\alpha, x_\beta \in \Sigma$, $\alpha \in \mathcal{A}$, $\beta \in \mathcal{B}$. Gli insiemi di vettori di Σ , $\{x_\alpha, \alpha \in \mathcal{A}\}$, $\{x_\beta, \beta \in \mathcal{B}\}$, sono finiti e considerando l'unione dei due insiemi (che risulta ancora finito) possiamo estendere le somme su un medesimo insieme finito di indici $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$:

$$x = \sum_{\gamma \in \mathcal{C}} a_\gamma x_\gamma = \sum_{\gamma \in \mathcal{C}} b_\gamma x_\gamma,$$

dove abbiamo posto $a_\gamma = 0$ se $\gamma \notin \mathcal{A}$, e analogamente $b_\gamma = 0$ se $\gamma \notin \mathcal{B}$, per cui:

$$0 = \sum_{\gamma \in \mathcal{C}} a_\gamma x_\gamma - \sum_{\gamma \in \mathcal{C}} b_\gamma x_\gamma = \sum_{\gamma \in \mathcal{C}} (a_\gamma - b_\gamma) x_\gamma.$$

Ma i vettori $x_\gamma \in \Sigma$ sono linearmente indipendenti per cui $a_\gamma = b_\gamma$ e le espansioni del tipo (3.29) sono univocamente determinate da x .

Viceversa, se Σ è linearmente indipendente ed ogni vettore $x \in X$ si può sviluppare univocamente come:

$$x = \sum_{\alpha} c_{\alpha} x_{\alpha},$$

con al più un numero finito di termini non nulli, allora Σ è massimale, perché l'aggiunta di un qualsiasi elemento non nullo e distinto da Σ distrugge l'indipendenza dell'insieme.

— ■ —

Osservazione. Notiamo che, in base alle definizioni poste, lo sviluppo (3.29) è una somma finita, quindi senza problemi di convergenza. Non possiamo considerare somme infinite, cioè serie, fino a quando non introduciamo una topologia nello spazio mediante la quale poter stabilire la convergenza. In seguito, dopo avere introdotto una topologia, potremo modificare il concetto di base, ammettendo anche un numero infinito di termini nelle somme.

I coefficienti scalari c_{α} dello sviluppo (3.29) sono detti **componenti** di x lungo x_{α} .

Sorge ora il problema dell'esistenza di una base e, nel caso di spazi vettoriali a dimensione finita, se basi distinte abbiano la stessa cardinalità. Nel caso di uno spazio vettoriale a dimensioni finite il problema è facilmente risolubile. Ricordiamo che X ha dimensione n se il numero massimo di vettori linearmente indipendenti è n (finito).

Teo. 3.2 Sia X uno spazio vettoriale finito dimensionale di dimensione n . Allora X ammette almeno una base formata da n vettori ed ogni altra base ha lo stesso numero n di elementi. Inoltre se $\{e_1, e_2, \dots, e_k\}$, con $k < n$, è un sistema di vettori linearmente indipendenti, si possono aggiungere $n - k$ vettori opportuni e_{k+1}, \dots, e_n , formando una base $\{e_1, \dots, e_n\}$.

Dim. 3.2 È immediato vedere che esiste una base di n elementi. Essendo lo spazio di dimensione n , esiste, per definizione, un insieme $\{e_1, e_2, \dots, e_n\}$ di vettori linearmente indipendenti. Essendo n il massimo numero di vettori linearmente indipendenti nello spazio, questo forma automaticamente una base.

Chiaramente una qualsiasi base di X non può avere più di n elementi, altrimenti tale base costituirebbe un insieme di vettori linearmente indipendenti che contraddice l'ipotesi fatta sulla dimensionalità di X . Se avessimo anche una base $\{f_1, f_2, \dots, f_m\}$ con $m < n$, allora, per il teorema precedente, potremmo sviluppare i vettori della base precedente nella nuova base:

$$e_j = \sum_{k=1}^m \alpha_{jk} f_k, \quad j = 1, \dots, n,$$

con coefficienti α_{jk} non tutti nulli. Considerando ora il sistema omogeneo di m equazioni nelle n incognite c_j :

$$\sum_{j=1}^n c_j \alpha_{jk} = 0, \quad k = 1, \dots, m,$$

essendo $m < n$, tale sistema ammette una soluzione non nulla, cioè c_j non tutti nulli. Ciò porta ad un assurdo in quanto otterremmo una combinazione nulla dei vettori e_j (linearmente indipendenti), con coefficienti non nulli:

$$\sum_{j=1}^n c_j e_j = \sum_{j=1}^n \sum_{k=1}^m c_j \alpha_{jk} f_k = 0.$$

per cui si deve avere necessariamente $n = m$, e tutte le basi hanno lo stesso numero di elementi. Si forma quindi una base scegliendo una qualsiasi n -upla di vettori linearmente indipendenti.

Il fatto che un sistema di vettori linearmente indipendenti possa essere completato per dare una base è ora ovvio. Se tale sistema è formato da k vettori, con $k < n$, non essendo un sistema massimale, esiste un vettore dello spazio linearmente indipendente che possiamo aggiungere al sistema, e possiamo proseguire aggiungendo un vettore dopo l'altro fino ad ottenere una base (il procedimento necessariamente finisce quando abbiamo n vettori linearmente indipendenti).

— ■ —

Nel caso infinito dimensionale non possiamo invocare lo stesso ragionamento (basato sostanzialmente sull'algebra lineare) e le cose si complicano. L'esistenza di una base è ancora garantita, ma trovarla può non essere semplice.

Teo. 3.3 *Ogni spazio vettoriale non nullo X ammette una base. Ogni sistema di vettori linearmente indipendenti, se non forma una base può essere completato in modo da formare una base.*

Dim. 3.3 La dimostrazione è basata su una proprietà generale delle relazioni d'ordine nota come lemma di Zorn¹. Ricordiamo che, dato un insieme S non vuoto, una **relazione d'ordine parziale** \leq è una relazione definita nell'insieme S che verifica le proprietà:

¹ **Max August Zorn** (6 Giugno 1906, Krefeld, Germania – 9 Marzo 1993, Bloomington, Indiana, Stati Uniti) era un matematico americano di origine tedesca esperto in algebra, teoria dei gruppi e analisi numerica. È famoso per il lemma di Zorn, scoperto nel 1935, un risultato di notevole importanza della teoria degli insiemi e applicabile ad un'ampia classe di costruzioni matematiche. Il lemma viene detto anche Lemma di Kuratowski-Zorn: in effetti esso fu scoperto da Kazimierz Kuratowski nel 1922 e riscoperto indipendentemente in seguito da Max Zorn.



i) **riflessiva**:

$$x \leq x, \quad \forall x \in S; \quad (3.30)$$

ii) **antisimmetrica**:

$$x \leq y, y \leq x, \quad x, y \in S, \quad \Rightarrow \quad x = y; \quad (3.31)$$

iii) **transitiva**:

$$x \leq y, y \leq z, \quad x, y, z \in S, \quad \Rightarrow \quad x \leq z. \quad (3.32)$$

Se in S risulta definita una relazione d'ordine parziale allora S è detto **parzialmente ordinato** o **semi-ordinato**.

La qualifica parziale deriva dal fatto che non è richiesto che due elementi di S siano necessariamente in relazione tra loro. Un sottoinsieme $S' \subseteq S$ è detto **linearmente ordinato** se due qualsiasi elementi di S' sono in relazione (cioè confrontabili) tra loro:

$$x, y \in S' \quad \Rightarrow \quad x \leq y \quad \text{oppure} \quad y \leq x. \quad (3.33)$$

Se $S' \subseteq S$, si dice che S' è **limitato superiormente** se esiste $b \in S$, detto **limite superiore**, tale che $a \leq b$ per ogni $a \in S'$. Un limite superiore s per S' è detto **estremo superiore** per S' se ogni altro limite superiore b verifica $s \leq b$.

Un elemento $m \in S'$ è detto un **massimale** per S' se le ipotesi in contemporanea, $p \in S'$ e $m \leq p$, implicano $m = p$, cioè non esiste un elemento di S' distinto da m che sia successivo a m secondo la relazione d'ordine.

Osservazione. Notiamo che non è detto che un elemento massimale (secondo tale definizione) sia anche un limite superiore o un estremo superiore, in quanto potrebbero esistere degli elementi di S' non confrontabili col massimale. Un limite superiore, e in particolare un estremo superiore, deve essere confrontabile con tutti gli elementi di S' .

Inoltre possono esistere più elementi massimali. Il massimale non è quindi da confondere con il **massimo**, cioè un elemento di S' che risulta anche estremo superiore, cioè confrontabile con tutti gli elementi di S' e successivo ad essi secondo la relazione d'ordine. Quando esiste il massimo questo è unico per la proprietà antisimmetrica della relazione d'ordine (se abbiamo m e m' massimi appartenenti a S' , allora necessariamente $m \leq m'$, e $m' \leq m$, per cui $m = m'$).

Con una relazione d'ordine parziale vale il seguente risultato.

Lem. 3.4 (Lemma di Zorn) Se in un insieme S (non vuoto) è definita una relazione d'ordine parziale, ed ogni sottoinsieme linearmente ordinato risulta limitato superiormente, allora esiste un elemento **massimale** $m \in S$ per tutto l'insieme S .

A seconda di come si imposta assiomaticamente la teoria degli insiemi, il lemma di Zorn può essere assunto come postulato, oppure derivabile in ultima analisi dall'assioma della scelta di Zermelo²:

“Data una famiglia non vuota di insiemi non vuoti esiste una funzione che ad ogni insieme della famiglia fa corrispondere un suo elemento.”

Per non addentrarci nei meandri dei fondamenti della matematica, assumiamo il lemma di Zorn come assioma equivalente a quello di Zermelo.

Siamo ora in grado di provare l'esistenza di una base, notando che tra insiemi la relazione di inclusione \subseteq costituisce proprio una relazione d'ordine. Consideriamo allora come S l'insieme di tutti i sistemi Σ linearmente indipendenti, con la relazione di inclusione (tra insiemi). Sicuramente S non è vuoto in quanto ogni sistema costituito da un solo vettore non nullo è linearmente indipendente. Se consideriamo un sottoinsieme $S' \subseteq S$ linearmente ordinato, esso è limitato superiormente dal sistema:

$$\bigcup_{\Sigma \in S'} \Sigma,$$

che si può provare facilmente essere linearmente indipendente, quindi un elemento di S . Allora possiamo applicare il lemma di Zorn ed asserire l'esistenza di un sistema Σ_m , massimale per S . Questo costituisce quindi una base in quanto non può più essere aumentato senza distruggere la sua indipendenza lineare (in caso contrario non sarebbe un massimale).

Analogamente si può dimostrare la seconda affermazione sul completamento di un qualsiasi sistema di vettori linearmente indipendenti, considerando l'insieme S formato da tutti i sistemi linearmente indipendenti contenenti il sistema dato.

— ■ —

Osservazione. Notiamo che la dimostrazione non offre alcuno spunto costruttivo per la determinazione della base, che può costituire un notevole problema in molti casi. È come il problema dei numeri trascendenti, si sa che sono la maggioranza dei numeri reali, ma è difficilissimo provare che un numero sia trascendente.

² **Ernst Friedrich Ferdinand Zermelo** (Berlino, 27 luglio 1871 – Friburgo, 21 maggio 1953) è stato un matematico e filosofo tedesco, reso celebre dai suoi contributi allo sviluppo della teoria assiomatica degli insiemi.



Esempio 3.5 Gli esempi 3.1 (\mathbb{R}^n) e 3.2 (\mathbb{C}^n) sono esempi di spazi vettoriali finito dimensionali e una base si ottiene considerando le n -uple:

$$e_j = (0, 0, \dots, 0, 1, 0, \dots, 0),$$

con l'elemento non nullo posto nella posizione j , $j = 1, 2, \dots, n$. Le combinazioni lineari dei vettori e_j a coefficienti reali generano tutto \mathbb{R}^n , mentre le combinazioni lineari complesse generano \mathbb{C}^n .

Pensando \mathbb{C}^n come spazio vettoriale reale una base è invece fornita dalle n -uple:

$$e_{2j-1} = (0, 0, \dots, 0, 1, 0, \dots, 0),$$

$$e_{2j} = (0, 0, \dots, 0, i, 0, \dots, 0), \quad i = \sqrt{-1},$$

sempre con l'elemento non nullo in posizione j , $j = 1, \dots, n$. Si ottengono $2n$ vettori linearmente indipendenti le cui combinazioni lineari a coefficienti reali esprimono qualsiasi elemento di \mathbb{C}^n .

— ○ —

Esempio 3.6 L'esempio 3.3 può essere finito dimensionale se la cardinalità di Y è finita, oppure infinito dimensionale se la cardinalità di Y è infinita. In questo ultimo caso risulta difficile stabilire una base per tutto lo spazio, mentre nel primo caso si può operare in analogia con \mathbb{R}^n .

— ○ —

Esempio 3.7 Nell'esempio 3.4 dell'insieme delle sequenze con un numero finito di elementi non nulli, una base può essere costruita mediante le sequenze:

$$e_j = (0, 0, \dots, 0, 1, 0, \dots),$$

sempre con l'unico elemento non nullo in posizione j , $j = 1, 2, \dots$. Considerando le combinazioni lineari (finite) di tali vettori si ottengono tutti i vettori dello spazio (sequenze con un numero finito di termini non nulli).

— ○ —

Sia ora S un sottoinsieme non vuoto di uno spazio vettoriale X . Se S non ha la struttura di spazio vettoriale, cioè S non è un sottospazio vettoriale, possiamo ricercare il minimo sottospazio lineare di X che contenga S , ottenendo il cosiddetto sottospazio lineare generato da S :

Def. 3.13 Sia $S \subseteq X$, con X spazio vettoriale sul campo \mathbb{K} . Definiamo **sottospazio lineare generato** da S il minimo sottospazio $\mathcal{L}(S)$, intersezione di tutti i sottospazi contenenti S :

$$\mathcal{L}(S) = \bigcap_{L \ni S} L \quad L \text{ sottospazio lineare di } X. \quad (3.34)$$

Una definizione alternativa è quella di definire $\mathcal{L}(S)$ come l'insieme delle combinazioni lineari **finite** di elementi di S :

$$\mathcal{L}(S) = \left\{ x \in X ; x = \sum_{j=1}^N c_j x_j, c_j \in \mathbb{K}, x_j \in S, N \in \mathbb{N} \right\}. \quad (3.35)$$

Infatti l'insieme delle combinazioni lineari finite è sicuramente un sottospazio lineare contenente S e rientra tra i sottospazi L di cui si opera l'intersezione. D'altra parte ogni sottospazio contenente S deve (essendo algebricamente chiuso) contenere una qualsiasi combinazione lineare finita di elementi di S . La caratterizzazione (3.35) risulta quindi il minimo sottospazio contenente S .

Se l'insieme S è finito, $S = \{x_1, x_2, \dots, x_n\}$ (n finito), allora è ovvio che il sottospazio generato $\mathcal{L}\{x_1, x_2, \dots, x_n\}$ risulta finito-dimensionale e contiene al più n vettori linearmente indipendenti. Se S è infinito, ma in S possiamo trovare al più un numero finito n di vettori linearmente indipendenti, allora $\mathcal{L}(S)$ risulta ovviamente a dimensioni finite.

Per mettere in evidenza la generazione del sottospazio mediante combinazioni lineari di elementi di S si usa anche la notazione $\text{Span}(S)$ per l'insieme (3.35).

3.2 Applicazioni lineari.

Introduciamo ora le trasformazioni tra spazi vettoriali che mantengono la struttura algebrica.

Def. 3.14 Siano X e Y due spazi vettoriali sul medesimo campo \mathbb{K} , allora una applicazione:

$$A : X \longrightarrow Y,$$

che associa ad ogni $x \in X$ un (unico) elemento $Ax \in Y$ è **lineare** oppure un **omomorfismo** se verifica l'additività e l'omogeneità:

$$A(x + y) = Ax + Ay \quad \forall x, y \in X, \quad (3.36)$$

$$A(\alpha x) = \alpha Ax \quad \forall \alpha \in \mathbb{K}, x \in X. \quad (3.37)$$

Indicheremo inoltre con $\mathcal{L}(X, Y)$ l'insieme delle applicazioni lineari da X a Y . Possiamo riassumere le relazioni precedenti in una richiesta unica:

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y \in X. \quad (3.38)$$

Nel caso particolare di una applicazione lineare da X al campo stesso \mathbb{K} (che può essere visto come spazio vettoriale unidimensionale):

$$\rho : X \longrightarrow \mathbb{K},$$

$$\rho(x + y) = \rho(x) + \rho(y), \quad \forall x, y \in X, \quad (3.39)$$

$$\rho(\alpha x) = \alpha \rho(x), \quad \forall \alpha \in \mathbb{K}, x \in X, \quad (3.40)$$

parliamo più propriamente di **funzionale lineare**. Anche in questo caso la condizione di linearità può essere espressa in una forma più compatta:

$$\rho(\alpha x + \beta y) = \alpha \rho(x) + \beta \rho(y), \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y \in X. \quad (3.41)$$

Esempio 3.8 In \mathbb{R}^n consideriamo una n -upla di numeri reali fissati: a_1, a_2, \dots, a_n . Allora l'applicazione:

$$\rho(x) = \sum_{j=1}^n a_j x_j, \quad x = (x_1, x_2, \dots, x_n),$$

risulta lineare, anzi un funzionale lineare.

Nel caso di spazi vettoriali complessi può aver senso definire un'altro tipo di trasformazioni, quelle antilineari.

Def. 3.15 Una trasformazione $A : X \rightarrow Y$, tra due spazi vettoriali complessi, è detta **antilineare** se verifica le proprietà:

$$A(x + y) = Ax + Ay, \quad \forall x, y \in X, \quad (3.42)$$

$$A(\alpha x) = \overline{\alpha} Ax \quad \forall \alpha \in \mathbb{C}, x \in X, \quad (3.43)$$

dove $\overline{\alpha}$ denota il complesso coniugato di α .

Nel caso che Y coincida con \mathbb{C} , allora il funzionale $\rho : X \rightarrow \mathbb{C}$, è detto **coniugato omogeneo** o semplicemente **antilineare** se:

$$\rho(x + y) = \rho(x) + \rho(y), \quad \forall x, y \in X, \quad (3.44)$$

$$\rho(\alpha x) = \overline{\alpha} \rho(x), \quad \forall \alpha \in \mathbb{C}, x \in X. \quad (3.45)$$

Esempio 3.9 Considerando lo spazio vettoriale \mathbb{C}^n , in analogia e con le stesse notazioni dell'esempio 3.8, il funzionale:

$$\rho(x) = \sum_{j=1}^n a_j \overline{x_j},$$

risulta un funzionale coniugato omogeneo (possiamo assumere in generale $a_j \in \mathbb{C}$).

— ○ —

Notiamo che una trasformazione lineare (o antilineare) potrebbe essere definita anche solo su un sottoinsieme $D \subseteq X$ di uno spazio vettoriale X , che viene detto **dominio** e indicato con $\mathcal{D}(A)$, ma le condizioni di linearità richiedono che D abbia una struttura di spazio vettoriale, cioè dati x, y in D , anche la somma $x + y$ e i loro multipli $\alpha x, \beta y$, appartengono al dominio della trasformazione (in caso contrario possiamo estendere la sua definizione, proprio tramite le relazioni di linearità o antilinearità, a tutto il sottospazio generato da D). Non è quindi restrittivo supporre che una trasformazione lineare sia definita su tutto uno spazio vettoriale X , cioè che il dominio di definizione sia uno spazio vettoriale, comprendendo in questo modo anche il caso in cui il dominio sia un sottospazio proprio (che ha, in ogni caso, una struttura di spazio vettoriale) di uno spazio vettoriale più vasto.

A volte risulta però necessario specificare il dominio come sottospazio (proprio o improprio) di uno spazio vettoriale X , in particolare se si intendono studiare le caratteristiche del dominio, come sottoinsieme di X , oppure si vogliono confrontare tra loro due trasformazioni tra i medesimi spazi vettoriali. Per dire che due trasformazioni coincidono, non è sufficiente verificare che la loro azione (algebrica) coincide, ma devono coincidere anche

i domini di definizione:

$$A = B \iff \begin{cases} \mathcal{D}(A) = \mathcal{D}(B), \\ Ax = Bx, \quad \forall x \in \mathcal{D}(A). \end{cases} \quad (3.46)$$

Questo risulta particolarmente importante in spazi a dimensione infinita. Se consideriamo spazi a dimensione finita, una trasformazione lineare si può pensare equivalente ad una trasformazione lineare associata ad una matrice di scalari, e può essere estesa senza particolari problemi a tutto lo spazio.

Nel caso particolare di applicazioni lineari in cui lo spazio di arrivo coincide sostanzialmente con lo spazio di partenza parliamo di operatori lineari agenti in tale spazio. Più propriamente:

Def. 3.16 Sia X uno spazio vettoriale. Una applicazione lineare tra un sottospazio D di X , e X stesso:

$$A : D \longrightarrow X, \quad (3.47)$$

è detta **operatore lineare** dello spazio X . Il sottospazio D dei vettori per i quali è definita l'applicazione costituisce il **dominio** $\mathcal{D}(A)$ dell'operatore A , per cui possiamo scrivere:

$$A : \mathcal{D}(A) \longrightarrow X, \quad \mathcal{D}(A) \subseteq X \quad (3.48)$$

Osservazione. Abbiamo formalmente distinto la definizione di operatore lineare da quella di applicazione lineare, ma compiremo spesso un abuso di linguaggio chiamando operatori lineari anche le applicazioni lineari tra spazi vettoriali distinti. Risulterà chiaro dal contesto se si tratta di veri operatori (definiti su un unico spazio), oppure più in generale, di applicazioni lineari. Vedremo tra poco che in realtà possiamo sempre pensare due spazi distinti come sottospazi di un medesimo spazio vettoriale più vasto e la distinzione tra operatore e applicazione lineare viene sostanzialmente meno.

Data una trasformazione lineare $A : X \longrightarrow Y$, l'insieme dei punti $y \in Y$ immaginati di un qualche punto $x \in X$ è detto **range** di A e indicato con $\mathcal{R}(A)$:

$$\mathcal{R}(A) = \{y \in Y ; y = Ax, x \in X\}. \quad (3.49)$$

Essendo A lineare è facile vedere che $\mathcal{R}(A)$ forma un sottospazio lineare. La dimensione del sottospazio $\mathcal{R}(A)$ è detta **rango** di A .

Chiaramente una trasformazione lineare risulta suriettiva sul suo range:

$$A : \mathcal{D}(A) \xrightarrow{\text{su}} \mathcal{R}(A), \quad (3.50)$$

e A è detta **suriettiva** o **su** se $\mathcal{R}(A) = Y$.

Per una applicazione lineare A risulta un sottospazio (sempre a causa della linearità) anche il suo **nucleo**, detto anche **kernel**:

$$\mathcal{N}(A) = \{x \in X ; Ax = 0\}. \quad (3.51)$$

Se il nucleo è non banale, cioè esiste $x \neq 0$ tale che $Ax = 0$, allora A è detta **singolare**, mentre è detta **non singolare** se il kernel risulta coincidente con il sottospazio nullo, cioè:

$$Ax = 0 \implies x = 0. \quad (3.52)$$

Una applicazione lineare iniettiva, cioè tale che:

$$Ax_1 = Ax_2 \implies x_1 = x_2, \quad (3.53)$$

risulta necessariamente non singolare e viceversa. Il nucleo $\mathcal{N}(A)$ ci fornisce quindi informazioni sulla iniettività della trasformazione.

Se A è iniettiva allora per ogni $y \in \mathcal{R}(A)$ esiste un unico $x \in X$ tale che $y = Ax$, cioè risulta definita univocamente una applicazione $B : \mathcal{R}(A) \rightarrow X$, tale che:

$$BAx = x \quad \forall x \in X. \quad (3.54)$$

Tale applicazione risulta anch'essa 1 – 1, infatti:

$$\begin{aligned} By = 0 \quad y \in \mathcal{R}(A) &\implies y = Ax, \quad x \in X, \quad 0 = By = BAx = x, \\ &\implies y = Ax = 0. \end{aligned}$$

Allora, se $y \in \mathcal{R}(A)$, $y = Ax$, $x \in X$, allora B , definita dalla (3.54), verifica:

$$ABy = ABAx = Ax = y,$$

per cui si ha anche:

$$ABy = y \quad \forall y \in \mathcal{R}(A). \quad (3.55)$$

È chiaro che se $\mathcal{R}(A) = Y$, cioè A è suriettiva (oltre che iniettiva), allora B risulta definita su tutto Y , A è **invertibile**, e B è detta l'**inversa** di A e indicata con A^{-1} . Essendo A lineare anche A^{-1} è lineare e A risulta un **isomorfismo** tra i due spazi vettoriali.

Osservazione. Notiamo che solo dopo aver asserito l'iniettività e suriettività della trasformazione lineare A abbiamo potuto definire una trasformazione inversa A^{-1} , con:

$$\mathcal{D}(A^{-1}) = Y, \quad \mathcal{R}(A^{-1}) = X, \quad (3.56)$$

tale che:

$$\begin{aligned} A^{-1}Ax &= x \quad \forall x \in X, \\ AA^{-1}y &= y \quad \forall y \in Y. \end{aligned} \quad (3.57)$$

In generale, se non abbiamo informazioni sulla iniettività e suriettività di A , per avere l'invertibilità, è necessario determinare una trasformazione A^{-1} (identificabile poi con l'inversa) che verifichi **entrambe** le relazioni (3.57) (e di conseguenza abbiamo l'iniettività e suriettività di A e A^{-1}).

Nel caso che X e Y coincidano con un medesimo spazio a dimensioni finite, per cui le trasformazioni sono rappresentabili da matrici quadrate, è sufficiente verificare una sola delle due equazioni (3.57), per avere l'invertibilità. Infatti in questo caso possiamo dire che il determinante di A (e di A^{-1}) è non nullo, la trasformazione A è non singolare e vale anche l'altra relazione. A dimensioni infinite non possiamo invocare tale ragionamento, tipico dell'algebra lineare, il determinante è privo di significato in generale.

Considerando spazi vettoriali a dimensione finita, possiamo stabilire un confronto tra essi, basato sulla dimensionalità.

Teo. 3.5 Due spazi vettoriali (sul medesimo campo) finito-dimensionalì sono isomorfi se e solo se hanno la stessa dimensione.

Dim. 3.5 Siano X e Y i due spazi isomorfi tramite una applicazione invertibile A , e sia $\{e_1 \dots e_n\}$ una base di X . Allora i vettori $\{f_1 = A e_1, \dots, f_n = A e_n\}$ generano $\mathcal{R}(A) = Y$. Essi sono linearmente indipendenti in quanto A è 1-1. X ed Y hanno quindi la stessa dimensione n .

Viceversa se X ed Y hanno la stessa dimensione, siano $\{e_1, \dots, e_n\}$, $\{f_1, \dots, f_n\}$ due basi per X e Y rispettivamente, e definiamo la trasformazione lineare A tramite la richiesta che $A e_j = f_j$, $j = 1, \dots, n$:

$$\begin{aligned} x &= \sum_{j=1}^n x_j e_j, \quad x_j \in \mathbb{K}, \\ Ax &= \sum_{j=1}^n x_j A e_j = \sum_{j=1}^n x_j f_j. \end{aligned} \tag{3.58}$$

Allora chiaramente $\mathcal{R}(A) = Y$ e $Ax = 0 \Rightarrow x = 0$, per cui A è un isomorfismo.

— ■ —

L'insieme di tutte le trasformazioni lineari tra due spazi vettoriali X e Y viene indicato con $\mathcal{L}(X, Y)$:

$$\mathcal{L}(X, Y) = \{ A : X \longrightarrow Y, A \text{ lineare} \}. \tag{3.59}$$

Se non specifichiamo il contrario assumiamo che tutte le applicazioni siano definite su tutto lo spazio X , dominio comune di tutte le trasformazioni.

Date due trasformazioni A e B , definite ognuna sul proprio dominio (sottospazi di un medesimo spazio vettoriale X), e col medesimo codominio Y , possiamo definire l'applica-

zione somma $A + B$ e la moltiplicazione per uno scalare α :

$$(A + B)x = Ax + Bx \quad \forall x \in \mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(A + B), \quad (3.60)$$

$$(\alpha A)x = \alpha(Ax) \quad \forall x \in \mathcal{D}(A). \quad (3.61)$$

Chiaramente tali operazioni introducono una struttura vettoriale nell'insieme delle applicazioni lineari e, in particolare, possiamo affermare che $\mathcal{L}(X, Y)$ forma uno spazio vettoriale sul medesimo campo \mathbb{K} degli spazi X e Y .

3.3 Somme dirette.

Abbiamo visto che l'intersezione di due sottospazi vettoriali L_1 ed L_2 costituisce a sua volta un sottospazio vettoriale (eventualmente nullo), ma generalmente l'unione non è un sottospazio. Possiamo però determinare in generale il sottospazio generato dall'unione stessa. Risulta importante il caso in cui i due (o più) sottospazi sono “disgiunti” o meglio “indipendenti”, per i quali si può parlare di somma diretta.

Def. 3.17 Sia X uno spazio vettoriale e E, F due suoi sottospazi. Diremo che X è la **somma diretta** di E e F , e scriveremo:

$$X = E \oplus F, \quad (3.62)$$

se ogni elemento $x \in X$ può esprimersi **univocamente** come somma di un elemento $e \in E$ e di un elemento $f \in F$:

$$\forall x \in X \exists! e \in E, f \in F; \quad x = e + f. \quad (3.63)$$

In questo caso diremo anche che E ed F sono ciascuno un **sottospazio supplementare** dell'altro.

Osservazione. Il fatto che e ed f debbano essere univocamente determinati da x implica necessariamente che i due sottospazi siano disgiunti, cioè che la loro intersezione concida con lo spazio nullo: $E \cap F = \{0\}$. Infatti se, per assurdo, avessimo un vettore x non nullo nell'intersezione, avremmo:

$$x = 0 + x = x + 0 = \frac{x}{2} + \frac{x}{2},$$

con le possibili scelte (tra infinite scelte):

$$\begin{cases} e = 0 \\ f = x \end{cases} \quad \begin{cases} e = x \\ f = 0 \end{cases} \quad \begin{cases} e = \frac{x}{2} \\ f = \frac{x}{2} \end{cases}$$

che distruggono l'univocità.

Viceversa se abbiamo due sottospazi E ed F disgiunti: $E \cap F = \{0\}$, e ogni $x \in X$ lo possiamo decomporre come:

$$x = e + f, \quad e \in E, f \in F,$$

allora la decomposizione è unica. Se infatti abbiamo due decomposizioni:

$$x = e + f = e' + f', \quad e, e' \in E, \quad f, f' \in F,$$

allora le differenze:

$$e - e' = f' - f,$$

appartengono a entrambi i sottospazi e devono essere nulle:

$$e - e' = 0 = f' - f.$$

Pertanto possiamo dire che $X = E \oplus F$.

Due sottospazi E, F a intersezione nulla sono anche detti **linearmente indipendenti**, intendendo con ciò la proprietà che l'annullarsi della somma di due vettori appartenenti ad E ed F rispettivamente comporta l'annullarsi dei vettori stessi. Infatti, se E e F sono disgiunti e:

$$e + f = 0, \quad e \in E, \quad f \in F,$$

allora $e = -f$ e sia e che f appartengono allo stesso sottospazio, intersezione dei due sottospazi. Essendo il vettore nullo l'unico elemento dell'intersezione, devono essere nulli sia e che f . Viceversa se:

$$e + f = 0, \quad e \in E, \quad f \in F \implies e = f = 0,$$

allora se $x \in E \cap F$, la somma $x - x = 0$, pensando $x \in E$ e $-x \in F$, comporta $x = 0$.

Ovviamente la decomposizione di uno spazio nella somma diretta di due sottospazi può essere estesa alla somma diretta di più sottospazi E_1, \dots, E_n :

$$X = E_1 \oplus E_2 \oplus \dots \oplus E_n, \tag{3.64}$$

allorquando ogni vettore x dello spazio può essere decomposto in maniera unica come somma di vettori nei sottospazi:

$$x = \sum_{j=1}^n x_j, \quad x_j \in E_j, \quad j = 1, \dots, n, \tag{3.65}$$

e i vari sottospazi devono essere disgiunti tra loro.

Se invece abbiamo degli spazi vettoriali E_j , $j = 1, \dots, n$, non necessariamente sottospazi di un medesimo spazio vettoriale, ma definiti sul medesimo campo \mathbb{K} , possiamo costruire uno spazio vettoriale che sia sostanzialmente la somma diretta degli spazi E_j . Definiamo infatti nel prodotto cartesiano:

$$X = E_1 \times E_2 \times \dots \times E_n$$

una struttura di spazio vettoriale mediante le operazioni:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n), \tag{3.66}$$

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n). \tag{3.67}$$

Lo spazio X è detto (in analogia con le definizioni precedenti) **somma diretta** degli spazi E_j , e indicato di nuovo con:

$$\bigoplus_{j=1}^n E_j = E_1 \oplus E_2 \oplus \cdots \oplus E_n,$$

in quanto è immediato vedere che esiste una corrispondenza biunivoca tra E_j e il sottospazio:

$$E'_j = \{(0, \dots, 0, x_j, 0, \dots, 0) ; x_j \in E_j\}, \quad j = 1, \dots, n.$$

In questo modo possiamo costruire nuovi spazi vettoriali in maniera gerarchica. Risulta immediato verificare che se E_1, \dots, E_n hanno dimensione finita, lo stesso vale per la loro somma diretta, e la dimensione risulta la somma delle dimensioni dei singoli spazi:

$$\dim \left(\bigoplus_{j=1}^n E_j \right) = \sum_{j=1}^n \dim E_j. \quad (3.68)$$

Esempio 3.10 Sia $X = \mathbb{R}^{\mathbb{R}}$ lo spazio vettoriale formato dalle funzioni reali di variabile reale, e definiamo i sottospazi:

$$\begin{aligned} V_p &= \{f : \mathbb{R} \rightarrow \mathbb{R} ; f(x) = f(-x)\}, \\ V_d &= \{f : \mathbb{R} \rightarrow \mathbb{R} ; f(x) = -f(-x)\}. \end{aligned}$$

V_p e V_d sono costituiti dalle funzioni pari e dispari rispettivamente, e sono chiaramente disgiunti (solo la funzione identicamente nulla è contemporaneamente pari e dispari). Ora, se f è una funzione arbitraria, possiamo decomporla come:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2},$$

ottenendo la decomposizione dello spazio vettoriale nei suoi sottospazi:

$$\mathbb{R}^{\mathbb{R}} = V_p \oplus V_d.$$

— ○ —

Abbiamo visto come verificare se uno spazio vettoriale sia esprimibile come somma diretta di suoi sottospazi vettoriali, o come costruire la somma diretta di spazi (e sottospazi) vettoriali. Sorge ora la questione dell'esistenza o meno di sottospazi la cui somma diretta genera lo spazio vettoriale. Più precisamente ci poniamo la questione: dato uno spazio vettoriale X e un suo sottospazio Y , è possibile determinare un sottospazio Z , disgiunto da Y , che ne sia il supplementare, cioè:

$$X = Y \oplus Z.$$

La risposta è affermativa, e vediamo quindi come costruire tale sottospazio.

Sia Y un sottospazio di uno spazio vettoriale X . Cominciamo col costruire lo **spazio quoziente** X/Y . Nello spazio X introduciamo una relazione di equivalenza \sim definita da $(x, x' \in X)$:

$$x \sim x' \iff x - x' \in Y. \quad (3.69)$$

Questa relazione è di equivalenza in quanto, grazie alla struttura di gruppo additivo del sottospazio Y , sono verificate le proprietà:

riflessiva:

$$x - x = 0 \in Y \implies x \sim x;$$

simmetrica:

$$x \sim x' \implies x - x' \in Y \implies x' - x \in Y \implies x' \sim x;$$

transitiva:

$$\begin{aligned} x \sim x', x' \sim x'' &\implies x - x' \in Y, x' - x'' \in Y \\ &\implies x - x'' = (x - x') + (x' - x'') \in Y \\ &\implies x \sim x''. \end{aligned}$$

Lo spazio vettoriale X risulta allora decomposto in classi di equivalenza distinte e si definisce **insieme quoziente** X/Y la collezione di tali classi:

$$X/Y = \{\tilde{x} = [x]_{\sim} ; x \in X\} \quad (3.70)$$

Possiamo dotare l'insieme quoziente di una struttura di spazio vettoriale definendo la somma e la moltiplicazione per uno scalare:

$$[x]_{\sim} + [y]_{\sim} = [x + y]_{\sim}, \quad (3.71)$$

$$\alpha [x]_{\sim} = [\alpha x]_{\sim}. \quad (3.72)$$

Occorre però verificare che le definizioni (3.71) e (3.72), basate sulle operazioni tra i rappresentativi interni alle classi, siano ben poste, cioè indipendenti dalla scelta dei rappresentativi. Ciò è assicurato dalla struttura di sottospazio vettoriale di Y . Se $x \sim x'$, $y \sim y'$, e $\alpha \in \mathbb{K}$, allora:

$$\begin{aligned} (x + y) - (x' + y') &= (x - x') + (y - y') \in Y \\ &\implies (x + y) \sim (x' + y'), \\ \alpha x - \alpha x' &\in Y \implies \alpha x \sim \alpha x'. \end{aligned}$$

Le proprietà di spazio vettoriale dell'insieme quoziente sono facilmente verificabili e abbiamo così costruito lo **spazio vettoriale quoziente**. Notiamo che il sottospazio Y costituisce da solo una classe di equivalenza che definisce l'elemento nullo dello spazio quoziente

(ogni elemento di Y è equivalente allo zero). Inoltre abbiamo anche una corrispondenza lineare π tra X e X/Y :

$$\begin{aligned} \pi : X &\longrightarrow X/Y, \\ x &\longmapsto \pi(x) = [x]_{\sim}, \quad x \in X, \end{aligned} \tag{3.73}$$

una sorta di ‘proiezione’ che associa a x la sua classe di appartenenza $[x]_{\sim}$. La corrispondenza è chiaramente suriettiva in quanto ogni classe di equivalenza ammette un elemento rappresentativo in X , la cui immagine secondo π è la classe stessa. La linearità è conseguenza immediata della consistenza della struttura vettoriale introdotta nello spazio quoziante ($\alpha, \beta \in \mathbb{K}, x, y \in X$):

$$\pi(\alpha x + \beta y) = [\alpha x + \beta y]_{\sim} = \alpha [x]_{\sim} + \beta [y]_{\sim} = \alpha \pi(x) + \beta \pi(y).$$

Inoltre π si annulla sul sottospazio Y che costituisce il suo nucleo:

$$\pi(x) = 0 \iff x \in Y.$$

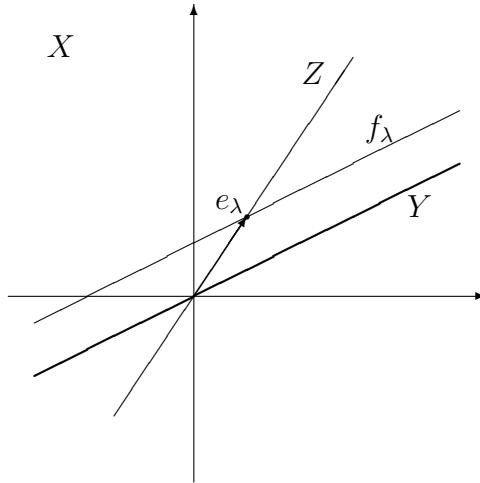


Figura 3.1: Esistenza del supplementare.

Siamo ora in grado di costruire un supplementare Z del sottospazio Y . Essendo X/Y uno spazio vettoriale, questo ammette una base $\{f_{\lambda}, \lambda \in \Lambda\}$, formata ovviamente da classi di equivalenza. Allora per ogni $\lambda \in \Lambda$ scegliamo un $e_{\lambda} \in X$ nella corrispondente classe di equivalenza, cioè:

$$e_{\lambda} \in \pi^{-1}(f_{\lambda}), \quad \pi(e_{\lambda}) = f_{\lambda}.$$

I vettori e_{λ} sono linearmente indipendenti in quanto:

$$\begin{aligned} \sum_{\lambda} c_{\lambda} e_{\lambda} = 0 &\implies \sum_{\lambda} c_{\lambda} \pi(e_{\lambda}) = 0 \\ &\implies \sum_{\lambda} c_{\lambda} f_{\lambda} = 0 \implies c_{\lambda} = 0 \quad \forall \lambda. \end{aligned}$$

Osservazione. Notiamo che tale ragionamento mostra come tramite le immagini inverse di una applicazione lineare possiamo costruire vettori linearmente indipendenti. Mediante l'immagine diretta è necessaria invece l'iniettività.

Mediante i vettori e_λ possiamo generare un sottospazio vettoriale:

$$Z = \mathcal{L} \{e_\lambda ; \lambda \in \Lambda\},$$

che costituisce proprio un supplementare di Y . Sia infatti $x \in X$ e consideriamo la sua classe di equivalenza $\pi(x)$, allora in maniera univoca abbiamo lo sviluppo:

$$\pi(x) = \sum_{\lambda} c_{\lambda} f_{\lambda} = \sum_{\lambda} c_{\lambda} \pi(e_{\lambda}).$$

Una volta determinate le componenti c_{λ} , abbiamo:

$$\pi \left(x - \sum_{\lambda} c_{\lambda} e_{\lambda} \right) = 0 \implies x - \sum_{\lambda} c_{\lambda} e_{\lambda} = y \in Y.$$

e quindi, essendo $\sum_{\lambda} c_{\lambda} e_{\lambda} = z \in Z$, abbiamo la decomposizione di x :

$$x = y + z, \quad y \in Y, z \in Z.$$

Inoltre i due sottospazi Y e Z sono disgiunti. Infatti se $x \in Y \cap Z$, si ha che $\pi(x) = 0$ perché $x \in Y$, e

$$x = \sum_{\lambda} c_{\lambda} e_{\lambda},$$

in quanto $x \in Z$. Ma allora:

$$0 = \pi(x) = \sum_{\lambda} c_{\lambda} \pi(e_{\lambda}) = \sum_{\lambda} c_{\lambda} f_{\lambda} \implies c_{\lambda} = 0,$$

per cui $x = 0$, e i due spazi sono vettorialmente disgiunti. Concludendo, Z è proprio il sottospazio cercato:

$$X = Y \oplus Z.$$

Ovviamente il supplementare non è unico in quanto dipende in ultima analisi dalla scelta del vettore e_{λ} per ogni classe di equivalenza f_{λ} di base per lo spazio quoziante. Tale scelta determina la “direzione” dei vettori che individuano il sottospazio Z .

Sostanzialmente abbiamo mostrato il seguente risultato:

Teo. 3.6 *Ogni sottospazio Y di uno spazio vettoriale X ammette un sottospazio supplementare Z tale che:*

$$X = Y \oplus Z.$$

Vediamo ora un'altra applicazione del concetto di insieme quoziente. Supponiamo di avere una trasformazione lineare tra due spazi vettoriali E ed F :

$$A : E \longrightarrow F,$$

e sia N un sottospazio di E tale che:

$$Ax = 0 \quad \forall x \in N,$$

cioè N è un sottospazio del nucleo di A . Allora esiste una **unica** applicazione lineare $\hat{A} : E/N \longrightarrow F$, tale che:

$$A = \hat{A} \circ \pi$$

dove π è la corrispondenza (3.73), vista in precedenza, tra un elemento di E e la sua classe di equivalenza in E/N .

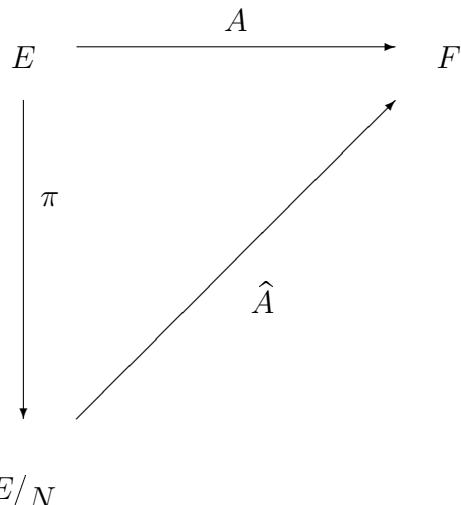


Figura 3.2: Decomposizione della applicazione A

Sia infatti $\xi \in E/N$, allora esiste $x \in E$ tale che $\pi(x) = \xi$, e definiamo:

$$\hat{A}\xi = Ax.$$

La definizione è ben posta, perchè se consideriamo x' nella stessa classe di equivalenza di x , allora $x - x' \in N$:

$$A(x - x') = 0 \implies Ax' = Ax = \hat{A}\xi.$$

Pertanto, per ogni $x \in E$:

$$\hat{A}\pi(x) = Ax.$$

L'unicità dell'applicazione \hat{A} è una conseguenza immediata della suriettività dell'applicazione π . Infatti se esistesse un'altra applicazione \hat{B} tale che:

$$Ax = \hat{A}\pi(x) = \hat{B}\pi(x) \quad \forall x \in E$$

allora $\hat{A}\xi = \hat{B}\xi$ per ogni $\xi \in E/N$.

Se in più imponiamo che N coincida esattamente con il nucleo di A , $N = \mathcal{N}(A)$, allora l'applicazione \hat{A} risulta anche iniettiva:

$$\hat{A}\xi = 0, \quad \xi = \pi(x) \implies Ax = 0 \implies x \in \mathcal{N}(A) \implies \xi = 0.$$

Abbiamo costruito in questo modo un isomorfismo tra l'insieme quoziente $E/\mathcal{N}(A)$ e $\mathcal{R}(A)$, e di conseguenza, ricordando il teorema precedente di esistenza del supplementare, un isomorfismo tra un sottoinsieme di E , supplementare di $\mathcal{N}(A)$, e $\mathcal{R}(A)$. Nel caso che E sia finito dimensionale abbiamo immediatamente il seguente risultato.

Teo. 3.7 Siano E ed F spazi vettoriali, con E finito dimensionale, e sia $A : E \rightarrow F$ una applicazione lineare tra i due spazi. Allora, indicando con \dim la dimensione di uno spazio o sottospazio lineare:

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = \dim E. \quad (3.74)$$

3.3.1 Dualità.

Abbiamo visto che nel caso di applicazioni lineari tra uno spazio vettoriale X ed il suo campo di definizione \mathbb{K} , queste sono dette funzionali lineari. L'insieme di tali funzionali costituisce uno spazio vettoriale detto duale.

Def. 3.18 Sia X uno spazio vettoriale sul campo \mathbb{K} . L'insieme di tutti i funzionali lineari $f : X \rightarrow \mathbb{K}$, costituisce uno spazio vettoriale detto **duale**, o meglio **duale algebrico**, di X , e indicato con X^* :

$$X^* = \mathcal{L}(X, \mathbb{K}). \quad (3.75)$$

Osservazione. La qualifica algebrico è stata messa per distinguere tale spazio duale dal cosiddetto duale topologico, che definiremo una volta introdotta una topologia nello spazio vettoriale.

Una volta definito lo spazio duale, essendo questo uno spazio vettoriale posso considerare il suo duale $(X^*)^* = X^{**}$, e così via. In pratica la catena si arresta molto presto, perché lo spazio di partenza X può essere identificato con un sottospazio dello spazio X^{**} (da cui la parola duale).

Teo. 3.8 Sia X uno spazio vettoriale sul campo \mathbb{K} . Allora esiste una corrispondenza iniettiva e lineare $J : X \rightarrow X^{**}$, tale che:

$$(J(x))(f) = f(x), \quad \forall f \in X^*. \quad (3.76)$$

Dim. 3.8 Sia f un funzionale lineare su X per cui $f(x) \in \mathbb{K}$ è ben definito per ogni $x \in X$. Allora, a fissato x , l'applicazione $T_x : X^* \rightarrow \mathbb{K}$:

$$f \xrightarrow{T_x} f(x),$$

risulta lineare ($\alpha, \beta \in \mathbb{K}, f, g \in X^*$):

$$\begin{aligned} (\alpha f + \beta g)(x) &= \alpha f(x) + \beta g(x) \implies \\ T_x(\alpha f + \beta g) &= \alpha T_x f + \beta T_x g, \end{aligned}$$

e $T_x \in X^{**}$. Pertanto, considerando ora x variabile, l'equazione (3.76) definisce un elemento del biduale:

$$J(x) = T_x.$$

Chiaramente la corrispondenza $x \rightarrow J(x)$ è lineare. Per ogni $f \in X^*$, e per ogni $\alpha, \beta \in \mathbb{K}$, $x, y \in X$, si ha:

$$\begin{aligned} (J(\alpha x + \beta y))(f) &= f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \\ &= \alpha (J(x))(f) + \beta (J(y))(f) = (\alpha J(x) + \beta J(y))(f). \end{aligned}$$

La corrispondenza J risulta anche iniettiva, infatti:

$$\begin{aligned} J(x_1) = J(x_2) &\iff J(x_1)(f) = J(x_2)(f) \quad \forall f \in X^* \\ &\iff f(x_1) = f(x_2) \quad \forall f \in X^* \\ &\iff f(x_1 - x_2) = 0 \quad \forall f \in X^*. \end{aligned}$$

Ma questo equivale a dire che $x_1 = x_2$. Infatti se $x_1 - x_2 \neq 0$, possiamo costruire il sottospazio unidimensionale non nullo:

$$L = \{y \in X ; y = \lambda(x_1 - x_2), \lambda \in \mathbb{K}\},$$

considerare un suo supplemento M , $X = L \oplus M$, e definire un funzionale lineare $f_0 \in X^*$:

$$f_0(\lambda(x_1 - x_2) + m) = \lambda, \quad m \in M,$$

che non si annulla in $(x_1 - x_2)$ (in effetti $f_0(x_1 - x_2) = 1$). Pertanto la trasformazione J risulta iniettiva e permette di identificare X con il sottospazio $J(X)$, immagine di X in X^{**} .

— ■ —

Osservazione. Spesso l'espressione $f(x)$, con $f \in X^*$ e $x \in X$, viene indicata con la notazione:

$$\langle f, x \rangle_{X^* X},$$

per mettere in evidenza l'aspetto di dualità tra i due spazi vettoriali e poterla pensare sia come applicazione lineare tra X e \mathbb{K} , cioè come elemento di X^* :

$$X \longrightarrow \mathbb{K},$$

$$x \longrightarrow \langle f, x \rangle_{X^* X},$$

che come applicazione lineare tra X^* e \mathbb{K} , cioè come elemento di X^{**} :

$$X^* \longrightarrow \mathbb{K},$$

$$f \longrightarrow \langle f, x \rangle_{X^* X}.$$

È facile vedere che:

Teo. 3.9 Sia X uno spazio vettoriale finito dimensionale di dimensione n . Allora anche X^* è finito dimensionale di dimensione n .

Dim. 3.9 Sia $\{e_1, \dots, e_n\}$ una base per X , per cui ogni $x \in X$ può essere sviluppato nella base:

$$x = \sum_{k=1}^n x_k e_k. \quad (3.77)$$

Se $f \in X^*$ è un funzionale lineare, allora (tralasciando i suffissi relativi alla dualità):

$$\langle f, x \rangle = \sum_{k=1}^n x_k \langle f, e_k \rangle,$$

e f risulta definito dai valori assunti sugli elementi di base. Consideriamo allora i funzionali lineari ϵ_j , definiti sulla base $\{e_k\}$ dalle relazioni:

$$\epsilon_j(e_k) = \langle \epsilon_j, e_k \rangle = \delta_{jk}, \quad j, k = 1, \dots, n.$$

Chiaramente abbiamo, usando lo sviluppo (3.77):

$$\epsilon_j(x) = x_j,$$

$$f(x) = \sum_{k=1}^n f(e_k) \epsilon_k(x),$$

$$f = \sum_{k=1}^n f(e_k) \epsilon_k = \sum_{k=1}^n \langle f, e_k \rangle \epsilon_k.$$

I funzionali ϵ_j generano quindi X^* (essendo f arbitraria), ed essendo linearmente indipendenti ($\alpha_j \in \mathbb{K}$):

$$\begin{aligned} \sum_j \alpha_j \epsilon_j = 0 &\implies \\ 0 = \left\langle \sum \alpha_j \epsilon_j, x \right\rangle &= \sum_{j=1}^n \alpha_j \langle \epsilon_j, x \rangle = \sum_{j=1}^n \alpha_j x_j, \quad \forall x_j \in \mathbb{K} \\ \implies \alpha_j &= 0, \quad j = 1, \dots, n. \end{aligned}$$

essi risultano una base per X^* , con la stessa cardinalità finita delle basi di X .

— ■ —

Chiaramente, nel caso finito dimensionale, anche X^{**} ha la stessa dimensionalità di X^* e quindi di X . Pertanto nel caso finito dimensionale X e X^{**} sono isomorfi e perfettamente identificabili tra loro.

3.4 Spazi normati.

Vogliamo ora arricchire la struttura degli spazi vettoriali introducendo in essi una topologia. La struttura di spazio vettoriale e la struttura topologica non devono essere arbitrarie, ma compatibili tra loro. Le due strutture devono essere tali da interferire in modo costruttivo ed armonico nello sviluppo della teoria. In questo modo si ottiene uno spazio lineare topologico.

Def. 3.19 Un insieme X è detto **spazio lineare topologico** se:

- i) X è uno spazio lineare.
- ii) X è uno spazio topologico di Hausdorff.
- iii) La somma $x + y$ di due vettori $x, y \in X$ è una funzione *continua* di x e y .
- iv) Il prodotto cx di un vettore $x \in X$ per uno scalare c è una funzione *continua* di c e x .

In particolare X si dirà reale o complesso se lo spazio vettoriale risulta reale o complesso rispettivamente.

Non discuteremo in generale gli spazi lineari topologici, ma introdurremo subito negli spazi vettoriali delle particolari topologie, sostanzialmente derivate da una metrica e discuteremo le loro proprietà, ma ovviamente molti risultati, specialmente se di natura prettamente topologica, saranno in generale validi in un qualsiasi spazio vettoriale topologico. Inoltre, se non specificheremo il campo degli scalari su cui è definito uno spazio vettoriale, sottointenderemo il campo dei numeri complessi.

3.4.1 Norme.

Introduciamo il concetto di norma, mediante la quale renderemo lo spazio vettoriale uno spazio metrico, e quindi topologico. La norma corrisponde al concetto intuitivo di lunghezza di un vettore. Quando, in fisica, si introducono in maniera pratica le grandezze vettoriali, si dice che sono caratterizzate da tre grandezze: intensità, direzione e verso. L'intensità, o modulo, del vettore si generalizza in maniera astratta nel concetto di norma. La norma di un vettore costituisce anche una generalizzazione del valore assoluto di un numero reale o complesso, e i campi dei numeri reali e complessi costituiscono degli esempi elementari di spazi normati, metrici, e quindi topologici.

Def. 3.20 Uno spazio lineare reale o complesso X è detto **normato** se ad ogni elemento $x \in X$ è associato un numero reale (detto **norma**) di x :

$$x \longrightarrow \|x\| \in \mathbb{R}, \quad (3.78)$$

tale che:

i) $\|x\|$ è una funzione non negativa:

$$\|x\| \geq 0 \quad \forall x \in X. \quad (3.79)$$

ii) $\|x\|$ è strettamente definita positiva:

$$\|x\| = 0 \iff x = 0. \quad (3.80)$$

iii) La norma è omogenea in valore assoluto:

$$\|\alpha x\| = |\alpha| \|x\| \quad \forall \alpha \in \mathbb{K}, \quad x \in X. \quad (3.81)$$

iv) Vale la diseguaglianza triangolare:

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in X. \quad (3.82)$$

Osservazione. La condizione ii) della definizione di norma è cruciale per poter definire una topologia di Hausdorff nello spazio vettoriale. Se tale condizione è eliminata e sostituita dalla:

$$x = 0 \implies \|x\| = 0,$$

che è una conseguenza della condizione iii) con $\alpha = 0$, allora la funzione $\|x\|$ è detta **seminorma** e indicata generalmente con $p(x)$.

Ovviamente la diseguaglianza triangolare può essere estesa, procedendo ricorsivamente, ad una somma di più vettori $x_k \in X$, $k = 1, \dots, n$:

$$\left\| \sum_{k=1}^n x_k \right\| \leq \sum_{k=1}^n \|x_k\|. \quad (3.83)$$

Inoltre, sempre dalla diseguaglianza triangolare, è possibile dedurre la validità della relazione:

$$|\|x\| - \|y\|| \leq \|x - y\|, \quad (3.84)$$

valida per ogni coppia di vettori x, y nello spazio normato.

Esempio 3.11 Lo spazio \mathbb{C}^n è ovviamente uno spazio normato, con la norma ordinaria definita da:

$$\|x\| = \sqrt{\sum_{j=1}^n |x_j|^2}, \quad x = (x_1, \dots, x_n) \in \mathbb{C}^n \quad (3.85)$$

Esempio 3.12 Lo spazio $C([a, b])$ delle funzioni (reali o complesse) continue nell'intervallo $[a, b]$, diviene uno spazio normato con la norma definita da:

$$\|f\| = \sup_{t \in [a, b]} |f(t)|,$$

detta anche norma uniforme.

— ○ —

Esempio 3.13 Nello spazio \mathbb{R}^3 si può definire una seminorma nel modo seguente:

$$p(x) = \sqrt{x_1^2 + x_2^2}, \quad x = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

È allora chiaro che tutti i vettori del tipo $(0, 0, x_3)$ hanno seminorma nulla.

— ○ —

Teo. 3.10 Se $p(x)$ è una seminorma in uno spazio vettoriale X , l'insieme N dei vettori x tali che $p(x) = 0$ forma un sottospazio vettoriale. È possibile definire una norma nello spazio vettoriale quoziante X/N assegnando ad ogni classe di equivalenza $[x]$ la funzione numerica:

$$\|[x]\| = p(x), \quad x \in [x]. \quad (3.86)$$

Dim. 3.10 Se $x, y \in N$, si ha:

$$p(x) = p(y) = 0.$$

Ma allora, se α è uno scalare:

$$\begin{aligned} p(\alpha x) &= |\alpha| p(x) = 0, \\ 0 &\leqslant p(x+y) \leqslant p(x) + p(y) = 0. \end{aligned}$$

Pertanto αx e $x+y$ appartengono a N che risulta un sottospazio.

Vediamo ora che la definizione (3.86) è ben posta, cioè che definisce una funzione della classe, indipendente dalla scelta del rappresentante di cui si calcola la seminorma. Siano infatti x, y appartenenti alla stessa classe di equivalenza, per cui la loro differenza appartiene a N :

$$p(x-y) = p(y-x) = 0;$$

allora abbiamo:

$$\begin{aligned} p(y) &= p(x+y-x) \leqslant p(x) + p(y-x) = p(x), \\ p(x) &= p(y+x-y) \leqslant p(y) + p(x-y) = p(y), \end{aligned}$$

per cui $p(x) = p(y)$ e la definizione è ben posta.

La funzione $\|[x]\|$ eredita immediatamente le proprietà i), iii), iv) della definizione di norma e seminorma. Rimane da verificare la proprietà della positività stretta. Abbiamo che:

$$\|[x]\| = p(x) = 0 \iff [x] = N,$$

ma N costituisce proprio l'elemento nullo dello spazio quoziante X/N .

— ■ —

Mediante una norma è possibile introdurre in uno spazio vettoriale una metrica. Definiamo infatti la **distanza** tra due vettori x, y in uno spazio normato X tramite la formula:

$$d(x, y) = \|x - y\|, \quad (3.87)$$

ed è immediato verificare le proprietà della definizione di distanza. Lo spazio normato diviene immediatamente uno spazio topologico, metrico, e di Hausdorff.

Osservazione. Notiamo che mediante una norma è sempre possibile definire una distanza, ma non è vero il viceversa, in quanto uno spazio metrico può non essere lineare. Inoltre la condizione di omogenità:

$$d(\alpha x, \alpha y) = |\alpha| d(x, y),$$

che vale in conseguenza della definizione (3.87), non è in genere richiesta da una distanza.

Con la metrica e la conseguente topologia indotta nello spazio vettoriale questo risulta in effetti uno spazio lineare topologico.

Teo. 3.11 *Uno spazio normato X risulta uno spazio lineare topologico.*

Dim. 3.11 Le condizioni di spazio vettoriale e di spazio topologico di Hausdorff sono già verificate per costruzione (la topologia è indotta dalla metrica). Rimangono da verificare le condizioni di compatibilità tra la struttura vettoriale e la struttura topologica, cioè la continuità della somma tra vettori e della moltiplicazione per uno scalare.

Se consideriamo gli intorni di due punti x_0 e y_0 :

$$\begin{aligned} S(x_0, \epsilon) &= \{x \in X ; \|x - x_0\| < \epsilon\}, \\ S(y_0, \epsilon) &= \{y \in X ; \|y - y_0\| < \epsilon\}, \end{aligned}$$

e $x \in S(x_0, \epsilon)$, $y \in S(y_0, \epsilon)$, abbiamo:

$$\|(x + y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\| < 2\epsilon.$$

da cui la continuità della somma.

Inoltre, se $x \in S(x_0, \epsilon)$, e $|\alpha - \alpha_0| < \epsilon$:

$$\begin{aligned}\|\alpha x - \alpha_0 x_0\| &= \|\alpha x - \alpha_0 x + \alpha_0 x - \alpha_0 x_0\| \\ &= \|(\alpha - \alpha_0)(x - x_0 + x_0) + \alpha_0(x - x_0)\| \\ &\leq |\alpha - \alpha_0| \|x - x_0\| + |\alpha - \alpha_0| \|x_0\| + |\alpha_0| \|x - x_0\| \\ &< \epsilon^2 + (\|x_0\| + |\alpha_0|)\epsilon,\end{aligned}$$

che, essendo $\|x_0\| + |\alpha_0| \geq 0$, può essere reso piccolo a piacere scegliendo opportunamente ϵ . Otteniamo quindi anche la continuità della moltiplicazione per uno scalare.

— ■ —

L'introduzione di una topologia tramite la norma comporta immediatamente che la norma stessa, intesa come funzione tra X e \mathbb{R} , risulta una funzione continua. Infatti, se x e x_0 appartengono a X , allora:

$$|\|x\| - \|x_0\|| \leq \|x - x_0\|,$$

e $f(x) = \|x\|$ è vicino a $f(x_0) = \|x_0\|$ (nella topologia dei reali) ogni volta che x è vicino a x_0 (nella topologia di X). Notiamo che la continuità risulta inoltre uniforme ($\delta(\epsilon) = \epsilon$, indipendentemente da x_0 e x) su tutto lo spazio.

3.4.2 Spazi di Banach.

Avendo introdotto una topologia tramite la norma dei vettori, sorge ora immediata la questione della completezza dello spazio. Si definisce così il concetto di spazio di Banach³.

Def. 3.21 Uno spazio lineare **normato completo** è detto **spazio di Banach**.

Esempio 3.14 Lo spazio $C([a, b])$ dell'esempio 3.12, con la **norma uniforme** definita da:

$$\|f\|_\infty = \sup_{t \in [a, b]} |f(t)|, \quad (3.88)$$

³ **Stefan Banach** (Cracovia, 30 marzo 1892 – Leopoli, 31 agosto 1945) è stato un matematico polacco, uno degli animatori della Scuola matematica di Lwow nella Polonia tra le due guerre. Egli era sostanzialmente un autodidatta in matematica e il suo genio fu scoperto accidentalmente da Hugo Steinhaus. Banach viene considerato il fondatore dell'analisi funzionale, argomento del quale avviò una trattazione sistematica sviluppando i risultati precedenti sulle equazioni integrali di Vito Volterra, Eric Ivar Fredholm e David Hilbert. Egli diede anche importanti contributi alla teoria degli spazi vettoriali topologici, alla teoria della misura, alla teoria degli insiemi e alla teoria dei polinomi ortogonali.



costituisce uno spazio di Banach. Notiamo infatti che la norma (3.88), e la corrispondente distanza:

$$d(f, g) = \sup_{t \in [a, b]} |f(t) - g(t)|,$$

tra funzioni continue nell'intervallo (chiuso) $[a, b]$, equivalgono ad imporre la convergenza uniforme delle successioni di funzioni, ed è un risultato noto dall'analisi che il limite di una successione uniformemente convergente di funzioni continue, nell'intervallo $[a, b]$, è una funzione continua nell'intervallo $[a, b]$.

Potevamo dotare invece lo spazio $C([a, b])$ con la norma:

$$\|f\|_1 = \int_a^b |f(t)| dt, \quad (3.89)$$

ottenendo ancora uno spazio normato, ma con una topologia diversa dalla precedente, e lo spazio non risulta più completo.

Infatti, consideriamo ad esempio l'intervallo $[-1, 1]$ e la successione di funzioni continue:

$$f_n(t) = \begin{cases} -1 & -1 \leq t < -\frac{1}{n} \\ nt & -\frac{1}{n} \leq t \leq +\frac{1}{n} \\ +1 & +\frac{1}{n} < t \leq +1 \end{cases} \quad n = 1, 2, \dots$$

f_n risulta di Cauchy (con la norma $\|\cdot\|_1$):

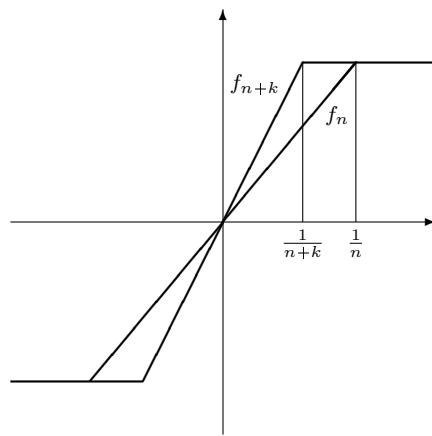


Figura 3.3: Una successione in $C([-1, 1])$.

$$\|f_{n+k} - f_n\|_1 = \frac{1}{n} - \frac{1}{n+k} = \frac{k}{n(n+k)} < \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0 \quad \forall k,$$

ma il limite puntuale $f(t)$:

$$f(t) = \begin{cases} -1 & -1 \leq t < 0 \\ 0 & t = 0 \\ +1 & 0 < t \leq 1 \end{cases}$$

che eventualmente verificherebbe la definizione di limite della successione f_n , non è una funzione continua, per cui la successione di Cauchy non converge nello spazio $C([a, b])$.

Notiamo che nella topologia della norma uniforme la successione f_n non risulta neanche di Cauchy:

$$\|f_{n+k} - f_n\|_\infty = \frac{k}{n+k}$$

(non tende a zero in maniera indipendente da k per $n \rightarrow \infty$).
— ○ —

Abbiamo visto che uno spazio metrico non completo può sempre essere completato considerando l'insieme delle successioni di Cauchy. Uno spazio normato risulta metrico, per cui esiste un completamento anche per uno spazio normato. In più, essendo lo spazio originale normato, è possibile rendere normato anche il completamento.

Teo. 3.12 *Uno spazio normato non completo X ammette sempre come completamento uno spazio di Banach \tilde{X} .*

Dim. 3.12 Abbiamo visto in precedenza come costruire il completamento tramite l'insieme delle successioni di Cauchy in X , con una relazione di equivalenza tra le successioni. In questo modo abbiamo ottenuto uno spazio metrico \tilde{X} con una distanza \tilde{d} dedotta dalla distanza in X . Tale costruzione rimane valida anche in uno spazio normato. Nel completamento risulta però definita una distanza (con relativa topologia), e non una norma, ma si può intuire che questa possa essere definita estendendo la norma definita in X . Per avere che tale spazio sia uno spazio di Banach occorre mostrare che:

- 1) \tilde{X} è uno spazio lineare;
- 2) \tilde{X} è uno spazio normato completo secondo la topologia definita dalla sua norma.

La prima affermazione è praticamente immediata introducendo una ovvia struttura vettoriale nello spazio \tilde{X} . Se $\{x_n\}$ e $\{y_n\}$ sono due successioni di Cauchy in X allora anche la successione somma $\{x_n + y_n\}$ è di Cauchy, in virtù della disuguaglianza triangolare:

$$\|(x_n + y_n) - (x_m + y_m)\| \leq \|x_n - x_m\| + \|y_n - y_m\|.$$

e si può vedere facilmente che tale nozione di somma è in realtà una nozione di classe, cioè:

$$\{x_n\} \sim \{x'_n\}, \quad \{y_n\} \sim \{y'_n\} \quad \implies \quad \{x_n + y_n\} \sim \{x'_n + y'_n\}.$$

Analogamente, possiamo provare che $\{\alpha x_n\}$, con α scalare, è di Cauchy al pari di $\{x_n\}$, e definisce una operazione di classe. Le operazioni così introdotte verificano facilmente tutte le proprietà di spazio vettoriale per cui otteniamo uno spazio lineare.

Per mostrare che \tilde{X} è uno spazio normato, notiamo innanzitutto che la distanza d in X , definita tramite la norma (equazione (3.87)), verifica le proprietà di omogeneità ed invarianza per traslazione:

$$d(\alpha x, \alpha y) = |\alpha| d(x, y), \quad (3.90)$$

$$d(x + z, y + z) = d(x, y), \quad (3.91)$$

valide per ogni $x, y, z \in X$ e $\alpha \in \mathbb{C}$, e per continuità, tali proprietà vengono a valere anche per la distanza \tilde{d} in \tilde{X} . Allora possiamo introdurre la seguente norma nel completamento \tilde{X} :

$$\|x\|_{\sim} = \tilde{d}(x, 0), \quad x \in \tilde{X}. \quad (3.92)$$

Sicuramente tale definizione estende la norma di X , infatti se $x \in X$, ed è identificato col corrispondente elemento di \tilde{X} , abbiamo:

$$\|x\|_{\sim} = \tilde{d}(x, 0) = d(x, 0) = \|x\|.$$

$\|x\|_{\sim}$ è una funzione chiaramente positiva e le proprietà della norma sono ovviamente verificate ($x, y \in \tilde{X}, \alpha \in \mathbb{K}$):

$$\begin{aligned} \|x\|_{\sim} = \tilde{d}(x, 0) = 0 &\iff x = 0 \\ \|\alpha x\|_{\sim} = \tilde{d}(\alpha x, 0) = |\alpha| \tilde{d}(x, 0) &= |\alpha| \|x\|_{\sim}, \\ \|x + y\|_{\sim} = \tilde{d}(x + y, 0) &\leq \tilde{d}(x + y, y) + \tilde{d}(y, 0) \\ &= \tilde{d}(x, 0) + \tilde{d}(y, 0) = \|x\|_{\sim} + \|y\|_{\sim}. \end{aligned}$$

Inoltre, la relazione:

$$\|x - y\|_{\sim} = \tilde{d}(x - y, 0) = \tilde{d}(x, y),$$

implica immediatamente che la topologia indotta dalla norma coincide con la topologia indotta dalla metrica \tilde{d} , e \tilde{X} è uno spazio di Banach, con la topologia indotta proprio dalla sua norma.

— ■ —

3.4.3 Serie infinite di vettori

Mediante una topologia abbiamo ovviamente ben definito il concetto di convergenza, ed in particolare possiamo definire la convergenza di una **serie** di vettori in uno spazio normato:

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots + x_n + \cdots$$

tramite la convergenza della successione delle ridotte della serie:

$$s = \sum_{k=1}^{\infty} x_k \iff \|s - s_n\| \xrightarrow{n \rightarrow \infty} 0, \quad (3.93)$$

con:

$$s_n = \sum_{k=1}^n x_k. \quad (3.94)$$

Ovviamente, se siamo in uno spazio di Banach, il criterio di Cauchy diviene sufficiente per garantire la convergenza, detta **convergenza in norma** per mettere in evidenza la sottostante struttura topologica.

Per una serie possiamo anche parlare di **assoluta convergenza**.

Def. 3.22 Una serie:

$$x_1 + x_2 + \cdots + x_k + \cdots,$$

è detta **assolutamente convergente** se è convergente la serie numerica in \mathbb{R} :

$$\|x_1\| + \|x_2\| + \cdots + \|x_k\| + \cdots.$$

Se ci troviamo in uno spazio di Banach, la diseguaglianza triangolare garantisce che una serie assolutamente convergente converge in norma. Infatti se consideriamo il criterio di Cauchy per la serie, abbiamo (assumendo $n > m$):

$$\left\| \sum_{k=1}^n x_k - \sum_{k=1}^m x_k \right\| = \left\| \sum_{k=m+1}^n x_k \right\| \leq \sum_{k=m+1}^n \|x_k\|,$$

che tende a zero in \mathbb{R} per $n, m \rightarrow \infty$.

Osservazione. Possiamo osservare che, in uno spazio di Banach, l'assoluta convergenza di una serie di vettori comporta che il limite della serie non dipende dall'ordine dei termini (in perfetta analogia col caso di serie numeriche assolutamente convergenti in \mathbb{R}). Siano infatti:

$$\sum_{k=1}^{\infty} x_k, \quad \sum_{k=1}^{\infty} x_{p(k)},$$

due serie (assolutamente convergenti) ottenute l'una dall'altra cambiando l'ordine dei termini (indicata con una permutazione p degli indici). Chiaramente le serie numeriche delle norme sono indipendenti dall'ordine:

$$\sum_{k=1}^{\infty} \|x_{p(k)}\| = \sum_{k=1}^{\infty} \|x_k\| < \infty,$$

e le due serie sono entrambe convergenti. L'assoluta convergenza ci dice che, per ogni $\epsilon > 0$, esiste un intero m tale che:

$$\sum_{k=m+1}^{\infty} \|x_k\| < \epsilon.$$

Sia n abbastanza grande in modo che i vettori x_1, x_2, \dots, x_m siano un sottoinsieme dei vettori riordinati $x_{p(1)}, x_{p(2)}, \dots, x_{p(n)}$. Se $M > m$ e $N > n$, abbiamo allora:

$$\left\| \sum_{k=1}^N x_{p(k)} - \sum_{k=1}^M x_k \right\| \leq \sum_{k=m+1}^{\infty} \|x_k\| < \epsilon,$$

e, prendendo il limite per $M \rightarrow \infty$:

$$\left\| \sum_{k=1}^N x_{p(k)} - \sum_{k=1}^{\infty} x_k \right\| \leq \epsilon \quad \text{per } N > n,$$

cioè le due serie coincidono:

$$\sum_{k=1}^{\infty} x_{p(k)} = \sum_{k=1}^{\infty} x_k. \tag{3.95}$$

Il legame tra la convergenza assoluta e la convergenza (semplice, vettoriale) per le serie di vettori può in realtà costituire un criterio di convergenza alternativo a quello di Cauchy.

Teo. 3.13 Uno spazio normato X è completo, cioè di Banach, se e solo se ogni serie assolutamente convergente risulta convergente.

Dim. 3.13 Abbiamo già visto che se X è uno spazio di Banach, cioè il criterio di Cauchy è sufficiente per la convergenza, allora ogni serie assolutamente convergente risulta anche convergente (semplicemente) in X . Viceversa occorre mostrare che l'assunzione che ogni serie assolutamente convergente risulta convergente comporta che ogni successione di Cauchy converge nello spazio X .

L'idea di base è fondata sul fatto che, come ogni serie è in realtà una successione delle sue ridotte, ogni successione x_n può essere interpretata come una serie:

$$x_n = x_1 + (x_2 - x_1) + (x_3 - x_2) + \cdots + (x_n - x_{n-1}),$$

ma occorre “controllare” i singoli termini per avere una serie assolutamente convergente. Assumiamo quindi di avere una successione $x_n \in X$, $n = 1, 2, \dots$, che verifichi il criterio di Cauchy:

$$\|x_n - x_m\| \xrightarrow[n, m \rightarrow \infty]{} 0.$$

Possiamo allora dire che esiste n_1 intero tale che:

$$\|x_n - x_{n_1}\| < \frac{1}{2}, \quad \forall n \geq n_1.$$

Determiniamo quindi $n_2 > n_1$ tale che:

$$\|x_n - x_{n_2}\| < \frac{1}{2^2}, \quad \forall n \geq n_2,$$

e così via, possiamo costruire una sottosuccessione x_{n_k} :

$$\|x_n - x_{n_k}\| < \frac{1}{2^k}, \quad \forall n \geq n_k,$$

con $n_k > n_{k-1}$, e in particolare abbiamo:

$$\|x_{n_{k+1}} - x_{n_k}\| < \frac{1}{2^k}.$$

La serie delle differenze $x_{n_{k+1}} - x_{n_k}$, $k = 1, 2, \dots$, risulta assolutamente convergente:

$$\sum_{k=1}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1,$$

e, per ipotesi, esiste la serie vettoriale:

$$\sum_{k=1}^{\infty} (x_{n_{k+1}} - x_{n_k}) = \lim_{k \rightarrow \infty} (x_{n_{k+1}} - x_{n_1}),$$

cioè esiste il limite in X :

$$x = \lim_{k \rightarrow \infty} x_{n_k} \in X.$$

Chiaramente x , sempre per la condizione di Cauchy, risulta anche il limite della successione x_n , cioè:

$$\|x_n - x\| \leq \|x_n - x_{n_k}\| + \|x_{n_k} - x\|$$

può essere reso piccolo a piacere.

— ■ —

3.4.4 Applicazioni lineari tra spazi normati.

Con l'introduzione di una topologia abbiamo la possibilità di trasformazioni continue tra spazi normati. In particolare siamo interessati alle trasformazioni lineari. Se A è una trasformazione lineare tra due spazi normati X e Y (sul medesimo campo reale o complesso):

$$A : X \longrightarrow Y,$$

possiamo dire che A è continua (secondo le topologie indotte negli spazi X e Y dalle corrispondenti norme vettoriali) in un punto $x_0 \in X$, se per ogni $\epsilon > 0$ esiste $\delta(\epsilon)$ tale che:

$$\|x - x_0\| < \delta(\epsilon) \implies \|Ax - Ax_0\| < \epsilon.$$

In linea di principio δ può dipendere anche dal punto x_0 in cui si verifica la continuità, ma la linearità dell'applicazione rende δ indipendente dal punto x_0 , e la continuità stessa risulta una proprietà globale indipendente dal punto, come sarà messo in evidenza tra poco.

Osservazione. Ovviamente le norme sono da intendersi nel corrispondente spazio normato. Se $x \in X$, $\|x\|$ esprime la norma nello spazio X , mentre $\|Ax\|$ esprime la norma dello spazio Y , che, anche in caso di coincidenza fra i due spazi, può risultare definita in modo diverso. Risulterà chiaro in generale dal contesto quale norma debba essere considerata, senza bisogno di specifiche ulteriori.

Per le trasformazioni lineari tra spazi normati è utile introdurre anche un'altra proprietà.

Def. 3.23 Una applicazione lineare $A : X \rightarrow Y$ tra due spazi normati è detta **limitata** se esiste una costante reale (finita) C , tale che:

$$\|Ax\| \leq C\|x\| \quad \forall x \in X. \quad (3.96)$$

Notiamo che la costante C (ovviamente non negativa) della (3.96) deve essere indipendente da x .

Osservazione. La nozione di limitatezza per una trasformazione lineare è diversa dalla nozione usuale di funzione limitata. A causa della linearità non ha senso dire che, come funzione, una applicazione lineare possa essere limitata. La condizione di omogeneità:

$$A(\lambda x) = \lambda Ax,$$

comporta la non limitatezza ordinaria, per cui si è modificata la definizione, e vedremo come questa risulta molto utile in pratica. Sostanzialmente la condizione (3.96) ci dice che la “pendenza” della trasformazione lineare è limitata. Questo fatto risulta naturale in spazi vettoriali di dimensioni finite dove una trasformazione lineare è rappresentata da una matrice di scalari che forniscono le possibili “pendenze” lungo le varie direzioni. Essendo le direzioni indipendenti in numero finito esiste sicuramente una pendenza massima e finita. Se la dimensionalità degli spazi vettoriali è infinita questo ragionamento cessa di valere e la limitatezza diventa una caratteristica che una trasformazione lineare può avere o non avere.

Osservazione. Ovviamente i concetti di continuità e limitatezza sono validi anche nel caso di funzionali lineari $f : X \rightarrow \mathbb{C}$, in quanto \mathbb{C} stesso può essere visto come spazio normato con la norma definita dall'operazione di modulo o valore assoluto.

Il concetto di limitatezza e di continuità per trasformazioni lineari sono in realtà coincidenti, infatti abbiamo il seguente risultato generale:

Teo. 3.14 Siano X e Y spazi normati e sia $A : X \rightarrow Y$ una applicazione lineare tra i due spazi. Allora le seguenti affermazioni sono equivalenti:

1. A è continua (in ogni punto).
2. A è continua nell'origine.
3. A mappa insiemi limitati in insiemi limitati.
4. A è limitata.
5. Esiste (finito) l'estremo superiore:

$$\sup_{\|x\| \leq 1} \|Ax\| = M < \infty. \quad (3.97)$$

Inoltre, se A è continua:

$$\begin{aligned} \sup_{\|x\| \leq 1} \|Ax\| &= \sup_{\|x\|=1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \\ &= \inf\{C \in \mathbb{R} ; \|Ax\| \leq C\|x\| \forall x \in X\}. \end{aligned} \quad (3.98)$$

Notiamo che in uno spazio normato possiamo dire che un insieme è limitato se è contenuto in una sfera centrata nell'origine.

Dim. 3.14 Vediamo le varie equivalenze, una alla volta.

- 1) \Rightarrow 2). Se A è continua in ogni punto, a maggior ragione lo sarà nell'origine.
- 2) \Rightarrow 1). A risulta continua nell'origine, per cui, se $\epsilon > 0$, esiste δ tale che:

$$\begin{aligned} \|Ax\| &< \epsilon \quad \forall x \text{ con } \|x\| < \delta \quad \Rightarrow \\ \|Ax - Ax_0\| &= \|A(x - x_0)\| < \epsilon \quad \forall x \text{ con } \|x - x_0\| < \delta, \end{aligned}$$

e A risulta continua in qualsiasi punto x_0 (inoltre δ non dipende da x_0 , essendo stato scelto in riferimento all'origine).

- 2) \Rightarrow 3). Sia B un insieme limitato in X , per cui esiste $R > 0$ tale che:

$$\|b\| \leq R \quad \forall b \in B.$$

Ma possiamo scalare ogni vettore $b \in B$ all'interno di una sfera (aperta) di raggio piccolo a piacere:

$$\left\| \frac{\delta}{2R} b \right\| \leq \frac{\delta}{2} < \delta,$$

rendendo piccola a piacere, per la continuità nell'origine, la sua immagine:

$$\left\| A \left(\frac{\delta}{2R} b \right) \right\| < \epsilon \quad \Rightarrow \quad \|Ab\| < \frac{2\epsilon R}{\delta}.$$

Ciò mostra la limitatezza dell'immagine $A(B)$.

3 \Rightarrow 4). Sia:

$$B = \{x \in X ; \|x\| \leq 1\},$$

allora, per ipotesi, esiste un raggio $R > 0$ per una sfera contenente l'immagine di B :

$$\|x\| \leq 1 \quad \Rightarrow \quad \|Ax\| \leq R.$$

Sia ora x qualsiasi e non nullo. Lo posso sempre ridurre in B :

$$\left\| \frac{x}{\|x\|} \right\| = 1 \quad \Rightarrow \quad \left\| A \left(\frac{x}{\|x\|} \right) \right\| \leq R \quad \Rightarrow \quad \|Ax\| \leq R\|x\|,$$

per cui A è limitata. Se $x = 0$ la relazione è banalmente verificata per ogni R positivo.

4) \Rightarrow 2). Sia C una costante tale che:

$$\|Ax\| \leq C\|x\| \quad \forall x.$$

Allora:

$$\|x\| < \delta \quad \Rightarrow \quad \|Ax\| < C\delta,$$

che può essere resa piccola a piacere, per cui A è continua nell'origine.

4) \Rightarrow 5). Come conseguenza della limitatezza tramite una costante $C \geq 0$, abbiamo:

$$0 \leq \|Ax\| \leq C \quad \forall x \text{ tale che } \|x\| \leq 1.$$

Pertanto l'insieme numerico $\{\|Ax\| ; \|x\| \leq 1\}$ è limitato ed esiste l'estremo superiore.

5) \Rightarrow 4). Sappiamo che $\|Ax\| \leq M$ se $\|x\| \leq 1$. Allora se x è arbitrario e non nullo:

$$\left\| \frac{x}{\|x\|} \right\| = 1 \quad \Rightarrow \quad \left\| A \frac{x}{\|x\|} \right\| \leq M \quad \Rightarrow \quad \|Ax\| \leq M\|x\|,$$

(di nuovo se $x = 0$ la verifica è banale).

Abbiamo così chiuso il ciclo delle cinque equivalenze. Assumendo A continua, vediamo di ottimizzare la costante C della definizione di limitatezza. Dalla dimostrazione precedente 4) \Rightarrow 5) risulta chiaro che $M \leq C$, per ogni costante C che verifica la condizione di limitatezza:

$$M \leq \inf\{C ; \|Ax\| \leq C\|x\|\}.$$

Dalla dimostrazione 5) \Rightarrow 4) notiamo invece che M costituisce una delle possibili costanti C che verificano la limitatezza:

$$\|Ax\| \leq M\|x\|,$$

pertanto:

$$\sup_{\|x\| \leq 1} \|Ax\| = \inf\{C \in \mathbb{R} ; \|Ax\| \leq C\|x\| \forall x \in X\}.$$

Possiamo restringere la ricerca dell'estremo superiore fra i vettori di lunghezza unitaria. Poniamo:

$$M = \sup_{\|x\| \leq 1} \|Ax\|, \quad M' = \sup_{\|x\|=1} \|Ax\|,$$

ed è immediato che $M' \leq M$ (essendo M' un estremo su un insieme più ristretto). D'altra parte se $x \neq 0$ e $\|x\| \leq 1$:

$$\left\| A \frac{x}{\|x\|} \right\| = \frac{\|Ax\|}{\|x\|} \geq \|Ax\|,$$

(abbiamo aumentato il denominatore), e considerando l'estremo superiore della parte sinistra (dove A è applicato ad un vettore unitario):

$$M' \geq \|Ax\| \quad \forall x \text{ con } \|x\| \leq 1,$$

e quindi, operando ora l'estremo superiore della parte destra, $M' \geq M$, e:

$$\sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\|.$$

E risulta ovvio che (essendo $\left\| \frac{x}{\|x\|} \right\| = 1$):

$$\sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|},$$

completando in questo modo la verifica delle relazioni (3.98).

— ■ —

La condizione che A è continua se e solo se muta insiemi limitati in insiemi limitati viene spesso usata per verificare la non continuità, nel senso che si cerca una successione di vettori x_n limitata, cioè $\|x_n\| \leq C$ per ogni n , per la quale risulta $\|Ax_n\|$ divergente.

Esempio 3.15 Vediamo un esempio di applicazione lineare che può essere continua oppure non continua, a seconda della scelta di un parametro. Sia X l'insieme dei polinomi (complessi) in una variabile reale:

$$X = \{p : \mathbb{R} \longrightarrow \mathbb{C} ; p(x) \text{ polinomio nella variabile } x\}.$$

Chiaramente l'insieme dei polinomi forma uno spazio vettoriale (a dimensioni infinite se non restringiamo il grado dei polinomi). Dotiamo X della seguente norma uniforme nell'intervallo $[0, 1]$:

$$\|p\| = \sup_{0 \leq t \leq 1} |p(t)|.$$

Essa verifica facilmente le proprietà della norma, ed in particolare:

$$\|p\| = 0 \quad \Rightarrow \quad p(t) = 0 \quad \forall t,$$

in quanto l'annullarsi identicamente nell'intervallo $[0, 1]$ comporta, per il principio di identità dei polinomi, l'annullarsi identicamente ovunque. Notiamo che la restrizione della variabile indipendente all'intervallo $[0, 1]$ vale solo nel calcolo della norma, mentre il polinomio è definito ovunque sull'asse reale.

Sia ora $A : X \rightarrow \mathbb{C}$ l'applicazione definita da:

$$A p = p(x_0),$$

con x_0 prefissato. La trasformazione A è chiaramente lineare:

$$A(\alpha p + \beta q) = (\alpha p + \beta q)(x_0) = \alpha p(x_0) + \beta q(x_0) = \alpha A p + \beta A q,$$

con $\alpha, \beta \in \mathbb{C}$, e p, q polinomi arbitrari.

Abbiamo:

$$\frac{\|A p\|}{\|p\|} = \frac{|p(x_0)|}{\left(\sup_{0 \leq x \leq 1} |p(x)| \right)},$$

ed è chiaro che, se x_0 appartiene all'intervallo $[0, 1]$, l'espressione sopra risulta limitata da una costante unitaria $C = 1$, e la trasformazione A è continua. Se invece $|x_0| > 1$, non è possibile limitare tale espressione. Basta considerare i monomi:

$$p_n(x) = x^n, \quad \|p_n\| = \sup_{0 \leq x \leq 1} x^n = 1,$$

e l'espressione:

$$\frac{\|A p_n\|}{\|p_n\|} = |x_0|^n,$$

può essere resa grande a piacere ($|x_0| > 1$) scegliendo n in maniera opportuna, e l'applicazione A non risulta continua.

Lasciamo al lettore lo studio del caso $-1 \leq x_0 < 0$.

— ○ —

Esempio 3.16 Sia ancora X l'insieme dei polinomi dell'esempio precedente, con la medesima norma uniforme nell'intervallo $[0, 1]$. Definiamo ora l'operatore lineare:

$$A : X \rightarrow X,$$

$$(A p)(x) = x p(x).$$

Allora:

$$\|A p\| = \sup_{[0,1]} |x p(x)| \leq \sup_{[0,1]} |p(x)| = \|p\|,$$

per cui A è limitato e continuo.

— ○ —

Esempio 3.17 Sempre nel medesimo insieme X dei polinomi con la norma uniforme nell'intervallo $[0, 1]$, consideriamo l'operatore lineare:

$$A : X \longrightarrow X,$$

$$(A p)(x) = \frac{dp}{dx}(x),$$

che associa ad un polinomio la sua derivata. Otteniamo ancora un polinomio, ma la trasformazione, pur essendo lineare, non è continua. Basta considerare di nuovo i monomi $p_n(x) = x^n$, $n = 1, 2, \dots$, per mostrare la non limitatezza:

$$\|p_n\| = 1, \quad \|Ap_n\| = \sup_{[0,1]} |n x^{n-1}| = n.$$

— ○ —

3.4.5 Norma operatoriale

L'insieme degli operatori lineari e continui tra due spazi normati X e Y , viene indicato con $\mathcal{L}_c(X, Y)$:

$$\mathcal{L}_c(X, Y) = \{A : X \longrightarrow Y ; A \text{ continuo}\} \quad (3.99)$$

oppure con $\mathcal{B}(X, Y)$, dall'inglese “bounded”, per mettere in evidenza il carattere di limitatezza delle applicazioni. Chiaramente $\mathcal{L}_c(X, Y)$ risulta uno spazio vettoriale, sottospazio di $\mathcal{L}(X, Y)$, inoltre può essere dotato di una norma definendo la **norma operatoriale**:

$$\begin{aligned} \|A\| &= \inf\{C \in \mathbb{R} ; \|Ax\| \leq C\|x\| \quad \forall x \in X\} \\ &= \sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \end{aligned} \quad (3.100)$$

La norma $\|A\|$ è sostanzialmente la migliore costante utilizzabile per maggiorare uniformemente $\|Ax\|$:

$$\|Ax\| \leq \|A\| \|x\|. \quad (3.101)$$

Teo. 3.15 Siano X e Y due spazi normati. Allora:

1. $\mathcal{L}_c(X, Y)$ è uno spazio normato, con la norma definita dalla relazione (3.100).
2. Se Y è uno spazio di Banach, anche $\mathcal{L}_c(X, Y)$ è uno spazio di Banach.

Dim. 3.15 Per la prima affermazione si tratta di verificare che la corrispondenza (3.100) verifica le proprietà di una norma. Chiaramente $\|A\| \geq 0$ e, (per ogni $A, B \in \mathcal{L}_c(X, Y)$, $\lambda \in \mathbb{C}$):

$$\begin{aligned} \|A\| = 0 &\iff \|Ax\| = 0, \quad \forall x \in X \\ &\iff Ax = 0, \quad \forall x \in X \iff A = 0, \\ \|\lambda A\| &= \sup_{\|x\| \leq 1} \|\lambda Ax\| = |\lambda| \sup_{\|x\| \leq 1} \|Ax\| = |\lambda| \|A\|, \\ \|(A + B)x\| &= \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \\ &\leq (\|A\| + \|B\|) \|x\| \quad \forall x \in X \\ &\implies \|A + B\| \leq \|A\| + \|B\|. \end{aligned}$$

Per la seconda affermazione dobbiamo vedere che $\mathcal{L}_c(X, Y)$ è completo, nell'ipotesi che Y sia completo (indipendentemente dalla completezza o meno del dominio X).

Sia A_n una successione di Cauchy di applicazioni continue, allora per ogni $\epsilon > 0$ esiste $n(\epsilon)$ tale che:

$$\|A_n - A_m\| < \epsilon \quad \forall n, m \geq n(\epsilon).$$

Quindi, per ogni $x \in X$:

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \epsilon \|x\|, \tag{3.102}$$

e la successione $A_n x$ è di Cauchy in Y , per cui converge (Y è completo per ipotesi) ad un punto $y \in Y$:

$$y = \lim_{n \rightarrow \infty} A_n x.$$

Il limite y dipende linearmente da x per cui definisce una trasformazione $A : X \rightarrow Y$. A è lineare come conseguenza del fatto che il limite della somma è uguale alla somma dei limiti:

$$A_n x + A_n x' \xrightarrow[n \rightarrow \infty]{} Ax + Ax',$$

cioè abbiamo la continuità della somma nella topologia indotta dalla norma. Analogamente la moltiplicazione per uno scalare:

$$\lambda A_n x \xrightarrow[n \rightarrow \infty]{} \lambda Ax,$$

conseguenza di nuovo della compatibilità richiesta tra la struttura topologica e la struttura algebrica.

La relazione (3.102), valida per ogni $x \in X$, passando al limite per $m \rightarrow \infty$, ci fornisce inoltre l'informazione che l'operatore $A_n - A$ è continuo:

$$\|A_n x - Ax\| \leq \epsilon \|x\|,$$

per cui, essendo A_n continuo, anche $A = A_n - (A_n - A)$ è continuo (e limitato), cioè $A \in \mathcal{L}_c(X, Y)$. Inoltre:

$$\|A_n - A\| \leq \epsilon$$

per cui A è proprio il limite della successione A_n nella topologia di $\mathcal{L}_c(X, Y)$.

— ■ —

Osservazione. La condizione:

$$\|A_n - A\| \xrightarrow{n \rightarrow \infty} 0, \quad (3.103)$$

esprime il concetto di **convergenza in norma** degli operatori, e si dice che A è il **limite in norma** della successione A_n . A volte viene anche detta **convergenza uniforme** degli operatori, essendo la norma esprimibile tramite un estremo superiore. Tale nozione di convergenza è spesso molto stringente, per cui si introdurranno in seguito nozioni di convergenza meno restrittive per successioni di operatori.

Notiamo che la nozione di norma operatoriale si può porre solo per le trasformazioni lineari continue (o limitate). Pertanto solo l'insieme $\mathcal{L}_c(X, Y)$ forma uno spazio vettoriale normato, mentre l'insieme più generale $\mathcal{L}(X, Y)$ ha solo la struttura di spazio vettoriale, pur contenendo al suo interno, come sottospazio, $\mathcal{L}_c(X, Y)$.

3.4.6 Norme equivalenti.

Su un medesimo spazio vettoriale possiamo introdurre diverse norme, ognuna delle quali determina una propria metrica e una corrispondente topologia. Se le due topologie sono uguali (ogni aperto di una topologia è un aperto anche dell'altra) allora le due norme sono dette equivalenti.

Def. 3.24 Sia X uno spazio vettoriale. Due norme su X sono dette **equivalenti** se inducono la medesima topologia.

Ci si può porre il problema di individuare un criterio per stabilire se due norme sul medesimo spazio vettoriale X sono equivalenti. Siano $\|\cdot\|_1, \|\cdot\|_2$ le due norme e indichiamo con X_1, X_2 rispettivamente i due spazi topologici risultanti, col medesimo supporto X . Per vedere se un aperto in X_1 è aperto anche in X_2 e viceversa, consideriamo la trasformazione identica:

$$id : X_1 \longrightarrow X_2,$$

$$id(x) = x.$$

e le sue proprietà di continuità. Chiaramente id è lineare, iniettiva e suriettiva, cioè invertibile (l'inversa coincide con l'identità stessa). Ricordando che la continuità è verificabile tramite l'apertura delle retroimmagini di aperti, e che tramite l'identità ogni insieme è immagine e retroimmagine di se stesso, possiamo dire che:

id è continua se e solo se ogni aperto in X_2 è aperto in X_1 .

id^{-1} è continua se e solo se ogni aperto in X_1 è aperto in X_2 .

Ma la continuità è equivalente alla limitatezza per cui id è continua se e solo se esiste una costante C tale che:

$$\|x\|_2 = \|id(x)\|_2 \leq C \|x\|_1, \quad \forall x \in X.$$

Analogamente id^{-1} è continua se e solo se esiste una costante C' tale che:

$$\|x\|_1 = \|id^{-1}(x)\|_1 \leq C' \|x\|_2, \quad \forall x \in X.$$

Ovviamente le costanti C e C' sono non nulle, essendo le norme non identicamente nulle. In definitiva:

Teo. 3.16 Due norme $\|\cdot\|_1$, $\|\cdot\|_2$, definite sul medesimo spazio vettoriale X , sono equivalenti, cioè inducono la stessa topologia, se e solo se esistono due costanti D e C positive, tali che:

$$D \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1, \quad \forall x \in X. \quad (3.104)$$

Come conseguenza immediata di ciò abbiamo che in uno spazio a dimensione finita tutte le norme sono equivalenti. Notiamo infatti che se $\|\cdot\|_1$ è equivalente a $\|\cdot\|_2$ e $\|\cdot\|_2$ è equivalente a $\|\cdot\|_3$ allora $\|\cdot\|_1$ è equivalente a $\|\cdot\|_3$ (si lascia la dimostrazione al lettore come esercizio).

Teo. 3.17 In uno spazio vettoriale finito-dimensionale X è sempre possibile definire una norma, e tutte le norme sono equivalenti tra loro.

Dim. 3.17 Sia $\{e_1, e_2, \dots, e_n\}$ una base di X ($n = \dim X$), e costruiamo una norma di riferimento. Se:

$$x = \sum_{j=1}^n x_j e_j, \quad x_j \in \mathbb{C},$$

poniamo:

$$\|x\|_2 = \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}}.$$

Sia ora $\|\cdot\|$ una norma qualsiasi definita in X . Allora:

$$\begin{aligned} \|x\| &= \left\| \sum_{j=1}^n x_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|e_j\| \\ &\leq \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n \|e_j\|^2 \right)^{\frac{1}{2}} = C \|x\|_2, \end{aligned}$$

dove abbiamo fatto uso della diseguaglianza di Schwarz⁴ in \mathbb{R}^n (il prodotto scalare di due vettori è minore del prodotto delle lunghezze dei vettori).

Dobbiamo ora provare l'altra diseguaglianza della equivalenza tra norme, cioè che esiste C' tale che $\|x\|_2 \leq C'\|x\|$. A questo proposito costruiamo la funzione:

$$\phi : \mathbb{C}^n \longrightarrow \mathbb{R},$$

$$\phi(x_1, \dots, x_n) = \left\| \sum_{j=1}^n x_j e_j \right\|.$$

ϕ risulta una funzione continua. Infatti:

$$\begin{aligned} |\phi(x_1, \dots, x_n) - \phi(y_1, \dots, y_n)| &= \left\| \left\| \sum_{j=1}^n x_j e_j \right\| - \left\| \sum_{j=1}^n y_j e_j \right\| \right\| \\ &\leq \left\| \sum_{j=1}^n (x_j - y_j) e_j \right\| \leq \sum_{j=1}^n |x_j - y_j| \|e_j\| \\ &\leq \left(\sum_{j=1}^n |x_j - y_j|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n \|e_j\|^2 \right)^{\frac{1}{2}} \leq K \left(\sum_{j=1}^n |x_j - y_j|^2 \right)^{\frac{1}{2}}, \end{aligned}$$

da cui la continuità in \mathbb{C}^n (la continuità è inoltre uniforme). Consideriamo ora in \mathbb{C}^n la superficie sferica di raggio unitario:

$$\sum_{j=1}^n |x_j|^2 = 1,$$

che costituisce un insieme chiuso e limitato, cioè compatto. Ma una funzione continua su un insieme compatto ammette massimo e minimo, per cui esiste in particolare un punto di minimo ($\tilde{x}_1, \dots, \tilde{x}_n$) su tale superficie, e per ogni punto sulla sfera unitaria:

$$\phi(x_1, \dots, x_n) \geq \phi(\tilde{x}_1, \dots, \tilde{x}_n) = \mu.$$

⁴ **Karl Hermann Amandus Schwarz** (Hersdorf, 25 gennaio 1843 – Berlino, 30 novembre 1921) è stato un matematico tedesco, noto per i suoi contributi all'analisi complessa. Schwarz fu inizialmente studente di chimica a Berlino, e successivamente avvicinato alla matematica da Kummer e Weierstrass. La sua ricerca ha riguardato l'analisi funzionale, la geometria differenziale ed il calcolo delle variazioni.



Il valore μ del minimo non può essere nullo:

$$\begin{aligned}\phi(\tilde{x}_1, \dots, \tilde{x}_n) = 0 &\iff \left\| \sum_{j=1}^n \tilde{x}_j e_j \right\| = 0 \\ &\iff \tilde{x}_j = 0, \quad j = 0, \dots, n,\end{aligned}$$

e la n -upla nulla non appartiene alla superficie sferica unitaria. Pertanto essendo ϕ l'espressione della norma $\|x\|$ come funzione delle componenti di x , abbiamo:

$$\|x\| \geq \mu > 0 \quad \text{se} \quad \|x\|_2 = 1.$$

Sia ora x arbitrario e non nullo:

$$\left\| \frac{x}{\|x\|_2} \right\| \geq \mu \implies \frac{\|x\|}{\|x\|_2} \geq \mu \implies \|x\|_2 \leq \frac{1}{\mu} \|x\|,$$

e abbiamo completato l'equivalenza tra le norme.

— ■ —

Una conseguenza immediata di questo risultato è che in uno spazio vettoriale finito-dimensionale la nozione di continuità non dipende dalla norma scelta. Inoltre:

Teo. 3.18 Se X e Y sono spazi vettoriali a dimensioni finite allora ogni applicazione lineare $A : X \rightarrow Y$ risulta continua.

Dim. 3.18 Sia infatti $\{e_1, e_2, \dots, e_n\}$ una base nello spazio X . La continuità è conseguenza di:

$$\begin{aligned}x &= \sum_{j=1}^n x_j e_j \implies Ax = \sum_{j=1}^n x_j A e_j \implies \\ \|Ax\| &\leq \sum_{j=1}^n |x_j| \|A e_j\| \leq \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n \|A e_j\|^2 \right)^{\frac{1}{2}} = K \|x\|_2.\end{aligned}$$

— ■ —

Osservazione. La diseguaglianza di Schwarz in \mathbb{R}^n può essere vista come una conseguenza della banale diseguaglianza:

$$2\alpha\beta \leq \alpha^2 + \beta^2, \tag{3.105}$$

valida per ogni $\alpha, \beta \in \mathbb{R}$, e derivante dalla non negatività del quadrato della differenza tra due numeri. Infatti, se abbiamo due n -uple di numeri reali, abbiamo:

$$\begin{aligned} \left(\sum_{j=1}^n a_j b_j \right)^2 &= \sum_{j=1}^n a_j b_j \sum_{k=1}^n a_k b_k = \sum_{j=1}^n a_j^2 b_j^2 + \sum_{1 \leq j < k \leq n} 2a_j b_k a_k b_j \\ &\leq \sum_{j=1}^n a_j^2 b_j^2 + \sum_{1 \leq j < k \leq n} [(a_j b_k)^2 + (a_k b_j)^2] \\ &= \sum_{j=1}^n a_j^2 b_j^2 + \sum_{j \neq k=1}^n a_j^2 b_k^2 = \left(\sum_{j=1}^n a_j^2 \right) \left(\sum_{k=1}^n b_k^2 \right), \end{aligned}$$

dove abbiamo usato la diseguaglianza (3.105) con $\alpha = a_j b_k$ e $\beta = a_k b_j$. Pertanto, estraendo la radice quadrata, otteniamo la diseguaglianza di Schwarz usata nelle dimostrazioni precedenti:

$$\left| \sum_{j=1}^n a_j b_j \right| \leq \left(\sum_{j=1}^n a_j^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^n b_k^2 \right)^{\frac{1}{2}}. \quad (3.106)$$

3.5 Spazi L^p .

Vogliamo ora discutere un importante esempio di spazi normati, quello delle funzioni p -sommabili.

Sia Ω un aperto di \mathbb{R}^n (eventualmente coincidente con tutto \mathbb{R}^n) e sia p positivo, $0 < p < \infty$. Consideriamo la totalità delle funzioni $f : \Omega \rightarrow \mathbb{C}$, misurabili e tali che:

$$\int_{\Omega} |f(x)|^p dx < \infty, \quad (3.107)$$

tal insieme, se strutturato con la usuale somma tra funzioni e con la moltiplicazione per uno scalare, diviene uno spazio vettoriale. Una funzione misurabile che verifica la condizione (3.107) è detta **p -sommabile**. L'unica difficoltà nella verifica della struttura di spazio vettoriale consiste nel mostrare che la somma di due funzioni p -sommabili è ancora p -sommabile. Ciò diviene evidente, per $p \geq 1$, considerando la validità della diseguaglianza:

$$|f(x) + g(x)|^p \leq 2^{p-1} (|f(x)|^p + |g(x)|^p), \quad (3.108)$$

conseguenza della convessità della funzione t^p per $t \geq 0$. Infatti (vedi figura 3.4), se a e b

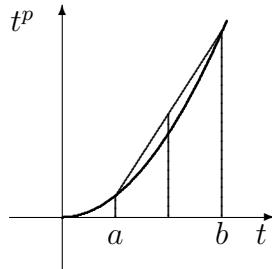


Figura 3.4: Convessità della potenza.

sono numeri positivi, la convessità della potenza p -esima comporta che tale funzione nel punto medio sia inferiore alla media dei valori agli estremi:

$$\left(\frac{a+b}{2} \right)^p \leq \frac{a^p + b^p}{2}$$

da cui la diseguaglianza (3.108) quando $p \geq 1$. Una diseguaglianza meno stringente, ma valida anche per $0 \leq p < 1$, è utile allo stesso scopo è:

$$|f(x) + g(x)|^p \leq 2^p (|f(x)|^p + |g(x)|^p), \quad (3.109)$$

conseguenza semplicemente dalla monotonicità della potenza p -esima (a, b positivi):

$$\begin{aligned} \frac{a+b}{2} &\leq \max\{a, b\} \quad \implies \\ \left(\frac{a+b}{2} \right)^p &\leq (\max\{a, b\})^p \leq a^p + b^p. \end{aligned}$$

In ogni caso le diseguaglianze (3.108) e (3.109) ci garantiscono che se $|f(x)|^p$ e $|g(x)|^p$ sono integrabili su Ω , allora lo è pure $|f(x) + g(x)|^p$: la somma di due funzioni p -sommabili è p -sommabile. In quanto segue saremo interessati principalmente al caso $p \geq 1$, per il quale riusciremo a costruire uno spazio di Banach.

Lavorando con le funzioni p -sommabili si opera principalmente mediante l'integrazione sull'insieme Ω e sappiamo che un integrale non dipende dal valore assunto dall'integrandi in un punto particolare. Ciò è particolarmente vero se si intende l'integrazione secondo Lebesgue⁵, per cui ad esempio due funzioni f e g danno luogo allo stesso integrale (3.107) se differiscono fra loro solo in un insieme di misura nulla. Conviene pertanto identificare le funzioni uguali fra loro quasi dappertutto introducendo una relazione di equivalenza tra le funzioni sommabili:

$$f \sim g \iff f(x) = g(x) \text{ q.d.}, \quad (3.110)$$

dove la qualifica q.d. è una abbreviazione per “quasi dappertutto”, e indica che $f(x) = g(x)$ per ogni $x \in \Omega \setminus N$ con N insieme a misura nulla. La relazione \sim suddivide l'insieme delle funzioni p -sommabili in classi disgiunte e due funzioni appartenenti a classi distinte sono effettivamente diverse ai fini dell'integrazione. È allora chiaro che:

$$\begin{aligned} \int_{\Omega} |f(x)|^p dx &= \int_{\Omega \setminus N} |f(x)|^p dx + \int_N |f(x)|^p dx \\ &= \int_{\Omega \setminus N} |g(x)|^p dx = \int_{\Omega} |g(x)|^p dx, \end{aligned}$$

e l'integrale (3.107) risulta in realtà una funzione della classe $[f]$ cui appartiene f . L'insieme delle classi di equivalenza può essere dotato di una struttura di spazio vettoriale e viene indicato con $L^p(\Omega)$. La struttura è uguale a quella di uno spazio quoziante vista in precedenza. Infatti l'insieme delle funzioni p -sommabili e nulle quasi ovunque:

$$\mathcal{N} = \{f : \Omega \rightarrow \mathbb{C} ; f = 0 \text{ q.d.}\}$$

forma un sottospazio vettoriale e $f \sim g$ se e solo se $f - g \in \mathcal{N}$.

Essendo l'integrale (3.107) una funzione di classe definiamo:

$$\|[f]\|_p = \|f\|_p = \left(\int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}}. \quad (3.111)$$

⁵ **Henri Léon Lebesgue** (Beauvais, Francia, 28 giugno 1875 – Parigi, Francia, 26 luglio 1941) è stato un matematico francese, famoso soprattutto per i suoi contributi alla moderna teoria dell'integrazione, per i quali è considerato il più grande matematico francese. La teoria dell'integrazione di Lebesgue fu pubblicata per la prima volta nella sua tesi, “Intègrale, longueur, aire” (“Integrale, lunghezza, area”), all'Università di Nancy nel 1902.



Abbiamo immediatamente che:

$$\|[f]\|_p = 0 \implies [f] = 0.$$

Cioè la classe $[f]$ coincide con il sottospazio \mathcal{N} , elemento nullo dello spazio quoziante. In genere si omette la notazione della classe $[f]$ e si indicano gli elementi con f sottointendendo, quando serve, l'intera classe di equivalenza contenente f . A volte la notazione L^p viene usata per indicare l'insieme delle funzioni p -sommabili, senza la divisione in classi di equivalenza, e si evince dal contesto di quale struttura si parla, cioè se un risultato è una proprietà di classe o di una singola funzione.

La nozione di spazio L^p può essere estesa anche al caso $p = \infty$, definendo lo spazio $L^\infty(\Omega)$, come l'insieme delle funzioni misurabili su Ω e limitate quasi ovunque:

$$\{f : \Omega \rightarrow \mathbb{C} ; f \text{ misurabile e limitata q.d.}\}, \quad (3.112)$$

suddiviso a sua volta in classi di equivalenza. Dire che f è limitata quasi dappertutto significa che esiste una costante K tale che $|f(x)| \leq K$ quasi ovunque. In questo modo possiamo definire la funzione di classe:

$$\|f\|_\infty = \inf \{K \in \mathbb{R} ; |f(x)| \leq K \text{ q.d.}\}. \quad (3.113)$$

Con le dovute cautele possiamo anche scrivere formalmente:

$$\|f\|_\infty = \sup_{x \in \Omega} |f(x)|,$$

dove però tale estremo superiore è da intendersi quasi ovunque, escludendo l'insieme a misura nulla in cui f può non essere limitata. Il significato esatto è da intendersi tramite la relazione (3.113) e viene detto **estremo superiore essenziale**:

$$\|f\|_\infty = \text{ess sup}_{x \in \Omega} |f(x)|,$$

e in ogni caso abbiamo:

$$|f(x)| \leq \|f\|_\infty \quad \text{q.d.} . \quad (3.114)$$

3.5.1 Disuguaglianza di Hölder

Vogliamo ora mostrare che la quantità $\|f\|_p$ definisce in effetti una norma e che $L^p(\Omega)$ risulta uno spazio normato, anzi di Banach. Notiamo che la suddivisione in classi di equivalenza è necessaria per avere la positività stretta della norma. In caso contrario l'espressione (3.107) definisce solo una seminorma. La proprietà non banale da provare è la disuguaglianza triangolare. A tale scopo premettiamo un utile lemma e un'importante disuguaglianza.

Lem. 3.19 Siano a, b due numeri reali non negativi, $a, b \geq 0$, p, q reali, con $1 < p, q < \infty$, tali che:

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (3.115)$$

Allora vale la disegualanza:

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \quad (3.116)$$

e l'uguaglianza vale solo se $a^p = b^q$.

Dim. 3.19 Consideriamo la funzione:

$$f(a) = \frac{1}{p} a^p + \frac{1}{q} b^q - ab,$$

con a variabile, b fisso, e

$$\frac{1}{p} + \frac{1}{q} = 1 \iff q = \frac{p}{p-1} \iff p = \frac{q}{q-1}.$$

Compiendo un rapido studio di tale funzione, abbiamo:

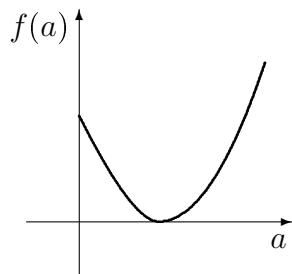


Figura 3.5: Disegualanza di Holder.

$$f'(a) = a^{p-1} - b,$$

$$f'(a) = 0 \iff b = a^{p-1} \iff b^q = a^p,$$

$$f''(a) = (p-1)a^{p-2} \geq 0 \quad \forall a \geq 0,$$

$$f(b^{\frac{q}{p}}) = \left(\frac{1}{p} + \frac{1}{q}\right) b^q - b^{\frac{q}{p}+1} = b^q - b^q = 0,$$

per cui $f(a) \geq 0$ per ogni a non negativo, e risulta dimostrata la disegualanza (3.116) con il corrispondente caso di uguaglianza.

— ■ —

Siamo ora in grado di mostrare la cosiddetta **disuguaglianza di Hölder**⁶.

Teo. 3.20 (Disuguaglianza di Hölder) Siano p, q , con $1 \leq p, q \leq \infty$, tali che:

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (3.117)$$

Allora, se $f \in L^p(\Omega)$ e $g \in L^q(\Omega)$, il prodotto $f g \in L^1(\Omega)$, e vale la disuguaglianza:

$$\|f g\|_1 \leq \|f\|_p \|g\|_q. \quad (3.118)$$

Dim. 3.20 Per il lemma precedente 3.19 sappiamo che, per ogni $x \in \Omega$:

$$|f(x) g(x)| \leq \frac{1}{p} |f(x)|^p + \frac{1}{q} |g(x)|^q,$$

se $1 < p < \infty$. Ma questa disuguaglianza implica la sommabilità del prodotto $f g$, cioè $f g \in L^1(\Omega)$. Inoltre, integrando su Ω otteniamo:

$$\begin{aligned} \int_{\Omega} |f(x) g(x)| dx &\leq \frac{1}{p} \int_{\Omega} |f(x)|^p dx + \frac{1}{q} \int_{\Omega} |g(x)|^q dx, \\ \|f g\|_1 &\leq \frac{1}{p} (\|f\|_p)^p + \frac{1}{q} (\|g\|_q)^q, \end{aligned}$$

che possiamo applicare alle funzioni normalizzate a uno:

$$\left\| \frac{f}{\|f\|_p} \frac{g}{\|g\|_q} \right\|_1 \leq \frac{1}{p} + \frac{1}{q} = 1$$

da cui la disuguaglianza di Hölder (3.118).

Se $p = 1$, e quindi $q = \infty$, come pure nel caso $p = \infty$, $q = 1$, l'affermazione risulta ovvia in base alla discussione precedente sulla relazione (3.114).

— ■ —

⁶ **Otto Ludwig Hölder** (Stoccarda, 22 dicembre 1859 – Lipsia, 29 agosto 1937) è stato un matematico tedesco. È famoso per aver enunciato la disuguaglianza di Hölder. Otto Hölder ha completato i suoi studi nell'università di Berlino, conseguendo poi il dottorato di ricerca all'università di Tubinga nel 1882. Il titolo della sua tesi dottorale fu “Beiträge zur Potentialtheorie” (“Contributi a una teoria potenziale”). Lavorò all'università di Lipsia fino al 1899, anno nel quale si ritirò in pensione.



Osservazione. Spesso la disuguaglianza di Hölder viene scritta in forma integrale:

$$\int_{\Omega} |f(x)g(x)| dx \leq \left(\int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}} \left(\int_{\Omega} |g(x)|^q dx \right)^{\frac{1}{q}}.$$

e nel caso $p = q = 2$ è nota anche col nome di **disuguaglianza integrale di Schwarz**.

La disuguaglianza di Hölder può essere generalizzata al prodotto di più funzioni sommabili.

Cor. 3.21 Siano date n funzioni $f_j \in L^{p_j}(\Omega)$, con $0 < p_j \leq \infty$, $j = 1, \dots, n$. Allora $\prod_{j=1}^n f_j \in L^r(\Omega)$, dove:

$$\frac{1}{r} = \sum_{j=1}^n \frac{1}{p_j}, \quad (3.119)$$

e vale la disuguaglianza:

$$\left\| \prod_{j=1}^n f_j \right\|_r \leq \prod_{j=1}^n \|f_j\|_{p_j}. \quad (3.120)$$

Dim. 3.21 Possiamo procedere ricorsivamente. Se $f_1 \in L^{p_1}(\Omega)$, $f_2 \in L^{p_2}(\Omega)$, allora, determinato r_2 in modo tale che:

$$\frac{1}{r_2} = \frac{1}{p_1} + \frac{1}{p_2},$$

possiamo considerare le funzioni:

$$g_1 = |f_1|^{r_2} \in L^{q_1}, \quad q_1 = \frac{p_1}{r_2},$$

$$g_2 = |f_2|^{r_2} \in L^{q_2}, \quad q_2 = \frac{p_2}{r_2},$$

$$\frac{1}{q_1} + \frac{1}{q_2} = 1, \quad q_1, q_2 \geq 1,$$

e il prodotto $g_1 g_2 \in L^1(\Omega)$, con:

$$\|g_1 g_2\|_1 \leq \|g_1\|_{q_1} \|g_2\|_{q_2},$$

pertanto il prodotto $f_1 f_2 \in L^{r_2}(\Omega)$, e:

$$\| |f_1 f_2|^{r_2} \|_1 \leq \| |f_1|^{r_2} \|_{q_1} \| |f_2|^{r_2} \|_{q_2},$$

$$(\|f_1 f_2\|_{r_2})^{r_2} \leq (\|f_1\|_{p_1})^{r_2} (\|f_2\|_{p_2})^{r_2},$$

$$\|f_1 f_2\|_{r_2} \leq \|f_1\|_{p_1} \|f_2\|_{p_2}.$$

Abbiamo così dimostrato il risultato per $n = 2$. Procedendo ora per induzione, assumiamo valido il corollario per $n - 1$ funzioni, per cui il prodotto $f_1 \cdots f_{n-1} \in L^{r^{n-1}}(\Omega)$, con:

$$\frac{1}{p_1} + \cdots + \frac{1}{p_{n-1}} = \frac{1}{r_{n-1}},$$

$$\frac{1}{r_{n-1}} + \frac{1}{p_n} = \frac{1}{r},$$

di conseguenza il prodotto $f_1 \cdots f_n \in L^r(\Omega)$, e:

$$\left\| \prod_{j=1}^n f_j \right\|_r = \left\| \left(\prod_{j=1}^{n-1} f_j \right) f_n \right\|_r \leq \left\| \prod_{j=1}^{n-1} f_j \right\|_{r_{n-1}} \|f_n\|_{p_n} \leq \prod_{j=1}^n \|f_j\|_{p_j}.$$

— ■ —

Osservazione. Notiamo che non abbiamo dovuto imporre la condizione $p_j \geq 1$, $j = 1, \dots, n$, e abbiamo generalizzato la diseguaglianza di Hölder anche quando $0 < p_j < 1$.

3.5.2 Diseguaglianza di Minkowski

Siamo ora in grado di mostrare la diseguaglianza triangolare:

$$\boxed{\|f + g\|_p \leq \|f\|_p + \|g\|_p,} \quad (3.121)$$

valida per ogni $f, g \in L^p(\Omega)$, con $1 \leq p \leq \infty$, e che prende il nome di **diseguaglianza di Minkowski**⁷.

Se $p = \infty$, oppure $p = 1$, la diseguaglianza è banale e deriva dal fatto che il modulo della somma di due numeri è minore o uguale alla somma dei due moduli. Sia pertanto $1 < p < \infty$ e $f, g \in L^p(\Omega)$, per cui sappiamo che $f + g \in L^p(\Omega)$, e:

$$\int_{\Omega} |f(x) + g(x)|^p dx = \int_{\Omega} |f(x) + g(x)| |f(x) + g(x)|^{p-1} dx.$$

Ma se $|f + g| \in L^p(\Omega)$ allora $|f + g|^{p-1} \in L^q(\Omega)$, con:

$$q(p-1) = p, \quad \frac{1}{q} + \frac{1}{p} = 1.$$

⁷ **Hermann Minkowski** (Aleksotas, 22 giugno 1864 – Göttingen, 12 gennaio 1909) è stato un matematico tedesco. Nato in Lituania da una famiglia di origine ebraica, Hermann Minkowski frequentò in Germania l’Università di Berlino e l’Università di Königsberg, dove conseguì la laurea nel 1885. A Zurigo, fu uno degli insegnanti di Albert Einstein. Egli sviluppò la teoria geometrica dei numeri ed utilizzò metodi geometrici per risolvere impegnativi problemi della teoria dei numeri, della fisica matematica e della teoria della relatività.



Allora, possiamo applicare la disuguaglianza di Hölder:

$$\begin{aligned}
 \|f + g\|_p^p &\leq \int_{\Omega} |f(x)| |f(x) + g(x)|^{p-1} dx + \int_{\Omega} |g(x)| |f(x) + g(x)|^{p-1} dx \\
 &\leq \|f\|_p \| |f + g|^{p-1} \|_q + \|g\|_p \| |f + g|^{p-1} \|_q \\
 &= (\|f\|_p + \|g\|_p) \left(\int_{\Omega} |f(x) + g(x)|^{(p-1)q} dx \right)^{\frac{1}{q}} \\
 &= (\|f\|_p + \|g\|_p) \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{\frac{p-1}{p}} \\
 &= (\|f\|_p + \|g\|_p) \|f + g\|_p^{p-1}.
 \end{aligned}$$

Da cui, semplificando il fattore comune, la disuguaglianza triangolare, nota in questo contesto come disuguaglianza di Minkowski, specialmente nella sua forma integrale:

$$\left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{\frac{1}{p}} \leq \left(\int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}} + \left(\int_{\Omega} |g(x)|^p dx \right)^{\frac{1}{p}}. \quad (3.122)$$

Il risultato precedente ci garantisce che gli spazi L^p sono spazi vettoriali normati, ma possiamo dire qualcosa in più, essi sono completi.

Teo. 3.22 Sia Ω un aperto di \mathbb{R}^n , e $1 \leq p \leq \infty$. Allora $L^p(\Omega)$ è uno spazio di Banach.

Dim. 3.22 Consideriamo prima il caso $p = \infty$. Sia u_n una successione di Cauchy in $L^\infty(\Omega)$. Allora, quasi ovunque abbiamo:

$$\begin{aligned}
 |u_n(x)| &\leq \|u_n\|_\infty, \quad \text{q.d.} \\
 |u_n(x) - u_m(x)| &\leq \|u_n - u_m\|_\infty \quad \text{q.d.}
 \end{aligned}$$

pertanto, quasi ovunque, $u_n(x)$ converge uniformemente ad una funzione $u(x)$. Come limite quasi ovunque di funzioni misurabili, anche il limite $u(x)$ è misurabile e siccome $\|u_n\|_\infty$ converge (forma una successione di Cauchy in \mathbb{R}), allora anche il limite puntuale $u(x)$ risulta una funzione limitata quasi ovunque, cioè $u \in L^\infty(\Omega)$, con:

$$\|u_n - u\|_\infty \xrightarrow{n \rightarrow \infty} 0.$$

Sia ora $1 \leq p < \infty$, e u_n una successione di Cauchy in $L^p(\Omega)$. Allora $\|u_n - u_m\|_p$ può essere resa piccola a piacere e operiamo nel seguente modo. Scegliamo n_1 in modo tale che:

$$\|u_n - u_{n_1}\|_p \leq \frac{1}{2} \quad \forall n \geq n_1,$$

Poi determiniamo $n_2 > n_1$ tale che:

$$\|u_n - u_{n_2}\|_p \leq \frac{1}{2^2} \quad \forall n \geq n_2,$$

e così via:

$$\|u_n - u_{n_k}\|_p \leq \frac{1}{2^k} \quad \forall n \geq n_k.$$

Otteniamo allora una successione u_{n_j} che verifica:

$$\|u_{n_{j+1}} - u_{n_j}\|_p \leq \frac{1}{2^j}, \quad j = 1, 2, \dots,$$

Poniamo ora:

$$v_m(x) = \sum_{j=1}^m |u_{n_{j+1}}(x) - u_{n_j}(x)|,$$

per cui $v_m \in L^p(\Omega)$ e la sua norma è limitata:

$$\|v_m\|_p \leq \sum_{j=1}^m \frac{1}{2^j} < 1 \quad \forall m.$$

Per il lemma di Fatou⁸, sappiamo che:

$$\int_{\Omega} \liminf_{m \rightarrow \infty} v_m^p(x) dx \leq \liminf_{m \rightarrow \infty} \int_{\Omega} v_m^p(x) dx \leq 1.$$

La successione v_m^p , monotona crescente e formata da funzioni positive (per cui il limite inferiore coincide con il limite), converge quindi quasi ovunque ad una funzione (finita) appartenente a $L^p(\Omega)$. Considerando la successione:

$$u_{n_{m+1}} = u_{n_1} + (u_{n_2} - u_{n_1}) + \dots + (u_{n_{m+1}} - u_{n_m}),$$

vista come serie, è assolutamente convergente quasi ovunque, per cui è puntualmente convergente quasi ovunque ad una funzione misurabile u . Ora abbiamo anche (sempre per il lemma di Fatou):

$$\begin{aligned} \int_{\Omega} \liminf_{j \rightarrow \infty} |u_{n_j}(x) - u_{n_k}(x)|^p dx &\leq \liminf_{j \rightarrow \infty} (\|u_{n_j} - u_{n_k}\|_p)^p, \\ (\|u - u_{n_k}\|_p)^p &\leq \liminf_j (\|u_{n_j} - u_{n_k}\|_p)^p. \end{aligned}$$

⁸ **Pierre Joseph Louis Fatou** (Lorient, Francia, 28 febbraio 1878 – PorNICHEt, Francia, 10 agosto 1929) è stato un matematico francese. Egli è noto soprattutto per i suoi lavori nel campo dei sistemi dinamici e dell'analisi complessa. Egli fu tra i primi a definire un frattale che sarà in seguito noto come “insieme di Mandelbrot”.



Essendo u_n di Cauchy, per ogni $\epsilon > 0$, esiste $N(\epsilon)$ tale che se $n, m \geq N(\epsilon)$:

$$\|u_n - u_m\|_p < \epsilon,$$

allora, se $n_k \geq N(\epsilon)$:

$$\|u - u_{n_k}\|_p \leq \epsilon,$$

Pertanto, $u - u_{n_k} \in L^p(\Omega)$, come pure $u = u_{n_k} + u - u_{n_k}$. Ma, se $m \geq N(\epsilon)$:

$$\|u - u_m\|_p \leq \|u - u_{n_k}\|_p + \|u_{n_k} - u_m\|_p \leq 2\epsilon,$$

quindi l'intera successione u_m , e non solo la sottosuccessione u_{n_k} , converge a $u \in L^p(\Omega)$ nella topologia di $L^p(\Omega)$, che risulta completo, cioè di Banach.

— ■ —

3.5.3 Spazi l^p

A completamento degli spazi L^p , introduciamo una versione “discretizzata”, considerando al posto delle funzioni le successioni (una successione può essere vista come una funzione definita su numeri naturali), con opportune condizioni. Sia $p \in \mathbb{R}^+$, $1 \leq p < \infty$ e definiamo:

$$l^p = \left\{ x = \{x_n\}_{n \in \mathbb{N}} ; x_n \in \mathbb{C} \forall n, \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}. \quad (3.123)$$

Gli elementi della successione devono dar luogo ad una serie (assolutamente) convergente quando se ne considerano le potenze p -esime e diremo che una tale successione è p -sommabile. Tale insieme diviene uno spazio normato completo con la norma:

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}, \quad x \in l^p. \quad (3.124)$$

Sostanzialmente abbiamo sostituito l'integrale con la somma di una serie. Per mostrare che lo spazio l^p è uno spazio di Banach potremmo ripercorrere lo stesso tipo di ragionamento fatto per gli spazi $L^p(\Omega)$, sostituendo sostanzialmente gli integrali con somme, ma, a titolo di utile esercizio, verifichiamo le proprietà di tali spazi in maniera più diretta.

Risulta ovvio che se $x \in l^p$, $\|x\|$ è una applicazione ben definita e positiva e si annulla solo per la successione identicamente nulla $x = 0$. Analogamente è immediato verificare l'omogeneità (in valore assoluto) rispetto alla moltiplicazione per uno scalare λ :

$$\|\lambda x\|_p = |\lambda| \|x\|_p.$$

La proprietà cruciale è data dalla diseguaglianza triangolare:

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p, \quad a, b \in l^p,$$

che ci garantisce anche la struttura di spazio vettoriale di l^p (cioè la chiusura algebrica di l^p rispetto all'operazione di somma).

Consideriamo quindi due successioni p -sommabili $a = \{a_n\}$, $b = \{b_n\}$, e poniamo (assumiamo due successioni non nulle):

$$\begin{aligned}\alpha &= \|a\|_p, & a' &= \left\{ \frac{1}{\alpha} |a_n| \right\}, \\ \beta &= \|b\|_p, & b' &= \left\{ \frac{1}{\beta} |b_n| \right\},\end{aligned}$$

per cui $\|a'\|_p = \|b'\|_p = 1$. Allora:

$$\begin{aligned}|a_n + b_n|^p &\leqslant (|a_n| + |b_n|)^p = (\alpha a'_n + \beta b'_n)^p = (\alpha + \beta)^p \left(\frac{\alpha}{\alpha + \beta} a'_n + \frac{\beta}{\alpha + \beta} b'_n \right)^p \\ &= (\alpha + \beta)^p [\lambda a'_n + (1 - \lambda) b'_n]^p,\end{aligned}$$

con:

$$\lambda = \frac{\alpha}{\alpha + \beta}, \quad 0 < \lambda < 1.$$

La convessità della funzione $t \mapsto t^p$ ci fornisce l'ulteriore disegualanza:

$$|a_n + b_n|^p \leqslant (\alpha + \beta)^p [\lambda a'^p_n + (1 - \lambda) b'^p_n],$$

e sommando su n :

$$\begin{aligned}(\|a + b\|_p)^p &= \sum_n |a_n + b_n|^p \leqslant \sum_n (\alpha + \beta)^p [\lambda a'^p_n + (1 - \lambda) b'^p_n] \\ &= (\alpha + \beta)^p [\lambda (\|a'\|_p)^p + (1 - \lambda) (\|b'\|_p)^p] = (\alpha + \beta)^p \\ &= (\|a\|_p + \|b\|_p)^p.\end{aligned}$$

Con questo risulta verificata la struttura di spazio vettoriale normato per l^p . Notiamo che abbiamo provato la disegualanza di Minkowski (per le somme) evitando la disegualanza di Hölder, che in realtà continua a valere (non ripetiamo la dimostrazione perfettamente analoga al caso integrale):

$$\|a \cdot b\|_1 \leqslant \|a\|_p \|b\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p, q > 1, \quad a \in l^p, \quad b \in l^q, \quad (3.125)$$

con:

$$a \cdot b = \{a_n b_n\}_{n \in \mathbb{N}} \in l^1.$$

Per vedere la completezza di l^p consideriamo una successione di Cauchy $\{x^{(n)}\}_{n \in \mathbb{N}}$ in l^p , cioè una successione a due indici, $x_{n,k}$ che verifica:

$$\sum_k |x_{n,k}|^p < \infty, \quad \left(\sum_k |x_{m,k} - x_{n,k}|^p \right)^{1/p} \xrightarrow[n,m \rightarrow \infty]{} 0.$$

Dobbiamo mostrare l'esistenza del limite di $x_{n,k}$ per $n \rightarrow \infty$ e che tale limite definisce (al variare di k) una successione ancora p -sommabile. Assegnato un $\epsilon > 0$, la condizione di Cauchy garantisce l'esistenza di un intero n_ϵ tale che:

$$\left(\sum_{k=1}^{\infty} |x_{m,k} - x_{n,k}|^p \right)^{1/p} < \epsilon, \quad \forall m, n > n_\epsilon$$

e quindi, essendo tutti i termini della serie positivi, abbiamo in particolare:

$$\begin{aligned} \left(\sum_{k=1}^N |x_{m,k} - x_{n,k}|^p \right)^{1/p} &< \epsilon, \quad \forall m, n > n_\epsilon, \quad \forall N \in \mathbb{N}, \\ |x_{m,k} - x_{n,k}| &< \epsilon, \quad \forall m, n > n_\epsilon, \quad \forall k \in \mathbb{N}. \end{aligned}$$

La completezza di \mathbb{C} implica allora l'esistenza, per ogni k , del limite:

$$x_k = \lim_{n \rightarrow \infty} x_{n,k}, \quad \forall k \in \mathbb{N},$$

con la convergenza uniforme rispetto a k (n_ϵ dipende solo da ϵ). Eseguendo il limite per $m \rightarrow \infty$ nelle relazioni precedenti, abbiamo:

$$\begin{aligned} |x_k - x_{n,k}| &\leq \epsilon, \quad \forall n > n_\epsilon, \quad \forall k \in \mathbb{N}. \\ \left(\sum_{k=1}^N |x_k - x_{n,k}|^p \right)^{1/p} &\leq \epsilon, \quad \forall n > n_\epsilon, \quad \forall N \in \mathbb{N}, \\ \left(\sum_{k=1}^N |x_k|^p \right)^{1/p} &\leq \left(\sum_{k=1}^N |x_{n,k}|^p \right)^{1/p} + \left(\sum_{k=1}^N |x_k - x_{n,k}|^p \right)^{1/p} \\ &\leq \left(\sum_{k=1}^N |x_{n,k}|^p \right)^{1/p} + \epsilon, \quad \forall n > n_\epsilon, \quad \forall N \in \mathbb{N}, \end{aligned}$$

e quindi (operando il limite per $N \rightarrow \infty$) abbiamo che la successione limite x_k definisce un elemento di l^p , limite nella norma $\|\cdot\|_p$ della successione di Cauchy $x^{(n)}$:

$$\begin{aligned} \|x\|_p &= \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^{\infty} |x_{n,k}|^p \right)^{1/p} + \epsilon < \infty \quad \forall n > n_\epsilon, \\ \|x - x^{(n)}\|_p &= \left(\sum_{k=1}^{\infty} |x_k - x_{n,k}|^p \right)^{1/p} \leq \epsilon, \quad \forall n > n_\epsilon. \end{aligned}$$

Come per gli spazi L^p , possiamo generalizzare anche al caso $p = \infty$, definendo lo spazio delle successioni limitate l^∞ :

$$l^\infty = \left\{ x = \{x_n\}_{n \in \mathbb{N}} ; x_n \in \mathbb{C} \quad \forall n, \quad \sup_n |x_n| < \infty \right\}, \quad (3.126)$$

strutturato con la norma uniforme:

$$\|x\|_\infty = \sup_n |x_n|, \quad x \in l^\infty. \quad (3.127)$$

Il lettore può verificare (vedi esercizio 3.7) che anche l^∞ costituisce un esempio di spazio di Banach. Ovviamente la disuguaglianza di Hölder si estende anche al caso $p = \infty$, $q = 1$, oppure $p = 1$, $q = \infty$, in quanto moltiplicando termine a termine una serie assolutamente convergente per una successione limitata, abbiamo ancora una serie assolutamente convergente.

3.5.4 Appendice: \limsup e \liminf

I concetti di \limsup e \liminf possono risultare non familiari, per cui li ricordiamo. Sia $A \subset \mathbb{R}$ un insieme contenente infiniti elementi. Se A ammette dei punti di accumulazione, allora il **maximile**, o **limite sup**, di A è definito come:

$$\limsup A = \sup\{x \in \mathbb{R} ; x \text{ di accumulazione per } A\}, \quad (3.128)$$

e analogamente si definisce il **minimile**, o **limite inf**, per A :

$$\liminf A = \inf\{x \in \mathbb{R} ; x \text{ di accumulazione per } A\}. \quad (3.129)$$

Se A è limitato (e infinito) allora l'insieme dei suoi punti di accumulazione è chiaramente non vuoto e la definizione ha senso. Nel caso che A sia illimitato superiormente estendiamo la definizione accettando l'infinito come punto di accumulazione, e definendo $\limsup A = \infty$. Se A è limitato superiormente, e non ha punti di accumulazione (per cui essendo infinito deve essere illimitato inferiormente), poniamo $\limsup A = -\infty$. Chiaramente estensioni analoghe si pongono per il \liminf . Sostanzialmente si accettano come punti di accumulazione anche i punti all'infinito, e qualsiasi insieme non finito ammette un massimo limite ed un minimo limite.

Possiamo dare la seguente caratterizzazione, che non dimostriamo, per il \limsup (e analoga per \liminf):

Teo. 3.23 Sia A un insieme contenente infiniti punti, $b = \limsup A$. Allora, per ogni $\epsilon > 0$, l'insieme $\{a \in A ; a > b + \epsilon\}$ è finito (o vuoto), mentre l'insieme $\{a \in A ; a > b - \epsilon\}$ è infinito.

Con ovvie modifiche nel caso di estensioni all'infinito. La definizione viene spesso usata quando A consiste in una successione di numeri reali a_n . In tal caso la definizione viene leggermente estesa per poter comprendere anche le successioni stazionarie (o in generale quando non si hanno infiniti elementi distinti nella successione). Si considerano cioè tutti i possibili punti b che sono limite di una qualche sottosuccessione a_{n_k} , comprendendo come limiti eventuali anche $+\infty$ e $-\infty$, poi si ricercano l'estremo superiore e l'estremo

inferiore, definendo il massimo limite ed il minimo limite, rispettivamente. Questi punti limite b comprendono quindi tutti i punti di accumulazione della successione, intesa come insieme, ma si accettano anche i punti limite di sottosuccessioni stazionarie che non sono necessariamente di accumulazione per la successione. Ricordiamo che otteniamo una sottosuccessione selezionando una successione di interi n_k , $k = 1, 2, \dots$, con $n_k \rightarrow \infty$ per $k \rightarrow \infty$, e considerando poi la successione a_{n_k} . Allora possiamo dire che b è un punto limite per la successione a_n , se per ogni N ed ogni ϵ , esiste un $n > N$ con $|b - a_n| < \epsilon$. L'estremo superiore dell'insieme di tali punti limite (includendo anche l'infinito) è detto $\limsup_{n \rightarrow \infty} a_n$, mentre l'estremo inferiore è detto $\liminf_{n \rightarrow \infty} a_n$. Osserviamo che tale caratterizzazione significa che $(b, +\infty)$ è un punto di accumulazione per le coppie $(a_n, n) \in \mathbb{R} \times \mathbb{N}$. In questo modo ritroviamo una definizione solo in termini di punti di accumulazione.

Per una successione il massimo limite ed il minimo limite possono essere valutati tramite le relazioni:

$$\limsup_{n \rightarrow \infty} a_n = \lim_{k \rightarrow \infty} \left(\sup_{m \geq k} a_m \right), \quad (3.130)$$

$$\liminf_{n \rightarrow \infty} a_n = \lim_{k \rightarrow \infty} \left(\inf_{m \geq k} a_m \right). \quad (3.131)$$

Risulta chiaro che (comprendendo i casi di limiti infiniti) il massimo limite e il minimo limite di una successione esistono sempre (al massimo divergono).

Chiaramente, se una successione è convergente, l'unico punto limite è il limite stesso della successione, per cui possiamo dire che una successione è convergente se il \limsup e \liminf coincidono con un numero finito.

Possiamo dire che valgono le seguenti proprietà (tutte valide nel caso di insiemi limitati, alcune anche con insiemi non limitati):

$$\begin{aligned} \limsup_{n \rightarrow \infty} (a_n + b_n) &\leq \limsup_{n \rightarrow \infty} a_n + \limsup_{n \rightarrow \infty} b_n, \\ \limsup_{n \rightarrow \infty} a_n b_n &\leq (\limsup_{n \rightarrow \infty} a_n) (\limsup_{n \rightarrow \infty} b_n), \\ \limsup_{n \rightarrow \infty} (c a_n) &= c \limsup_{n \rightarrow \infty} a_n \quad c > 0, \\ \limsup_{n \rightarrow \infty} (c a_n) &= c \liminf_{n \rightarrow \infty} a_n \quad c < 0, \end{aligned}$$

Il cosiddetto lemma di Fatou viene espresso proprio tramite i concetti di maxlimite e minlimite.

Lem. 3.24 (Lemma di Fatou) Se f_n è una successione di funzioni sommabili, $f_n \in L^1$, e non negative, $f_n(x) \geq 0$, e se:

$$\liminf_{n \rightarrow \infty} \|f_n\|_1 < \infty,$$

allora:

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x),$$

è una funzione sommabile e:

$$\|f\|_1 \leq \liminf_{n \rightarrow \infty} \|f_n\|_1.$$

Notiamo che nulla viene detto sulla differenza $\|f - f_n\|_1$, e sulla sua eventuale convergenza.

3.6 Esercizi

Esercizio 3.1 Data una base $\{e_1, e_2\}$ di vettori in \mathbb{R}^2 , si definisca la norma euclidea di un vettore qualunque $x = x_1 e_1 + x_2 e_2$ come:

$$\|x\|_0 = \sqrt{x_1^2 + x_2^2}.$$

Si definiscano inoltre le seguenti applicazioni da \mathbb{R}^2 a \mathbb{R} :

$$\begin{aligned}\|x\|_1 &= |x_1| + |x_2|, \\ \|x\|_2 &= \max \{|x_1|, |x_2|\}.\end{aligned}$$

a) Si dimostri che sia $\|\cdot\|_1$ che $\|\cdot\|_2$ definiscono una norma e si rappresenti nel piano cartesiano la sfera unitaria centrata nell'origine per tutte e tre le metriche:

$$S_j(0, 1) = \{x \in \mathbb{R}^2 ; \|x\|_j < 1\}, \quad j = 0, 1, 2.$$

b) Ricordando che tutte le metriche in \mathbb{R}^2 sono equivalenti si calcolino le costanti C_j, C'_j ($j = 1, 2$), per cui:

$$\|x\|_0 \leq C_j \|x\|_j, \quad \|x\|_j \leq C'_j \|x\|_0.$$

— ○ —

Esercizio 3.2 Sia X uno spazio vettoriale su cui sono definite due norme $m(x)$ e $n(x)$. Si dimostri che sono norme anche le seguenti applicazioni:

- a) $p(x) = \max\{m(x), n(x)\}$
- b) $q(x) = \alpha m(x) + \beta n(x)$ dove $\alpha, \beta \geq 0$ non simultaneamente nulle
- c) $r(x) = \sqrt{m(x)^2 + n(x)^2}$.

— ○ —

Esercizio 3.3 Si consideri lo spazio vettoriale \mathcal{A} delle funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ analitiche ovunque, tali cioè che la serie di Taylor nelle derivate n -esime della funzione, calcolate in zero:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

ha raggio di convergenza infinito.

a) Si dimostri che il sottoinsieme:

$$\mathcal{E} = \{f : \mathbb{R} \rightarrow \mathbb{R} ; |f^{(n)}(0)| \leq \rho_f^n, \forall n\},$$

con ρ_f costante positiva (che non dipende da n ma può essere dipendente dalla funzione f), è un sottospazio vettoriale di \mathcal{A} .

b) Si dimostri che il sottoinsieme:

$$\mathcal{P}_N = \{ f : \mathbb{R} \rightarrow \mathbb{R} ; f^{(n)}(0) = 0, n \geq N \},$$

è un sottospazio vettoriale di \mathcal{A} di dimensione finita e se ne determini una base.

— ○ —

Esercizio 3.4 Sia $C([0, 1])$ l'insieme delle funzioni continue dall'intervallo $[0, 1]$ ai valori reali.

a) Si dimostri che $C([0, 1])$ è uno spazio vettoriale su \mathbb{R} .

b) Si mostri che l'applicazione:

$$\|f\| = \sup_{x \in [0, 1]} |f(x)|,$$

definisce una norma su $C([0, 1])$ che lo rende uno spazio di Banach.

c) Si determini quali dei seguenti sottoinsiemi di $C([0, 1])$ sono sottospazi lineari chiusi:

- i) i polinomi di grado esattamente uguale a 3;
- ii) i polinomi di grado minore o uguale a 3;
- iii) le funzioni f tali che $f(0) = 2f(1)$;
- iv) le funzioni f tali che $f(0) = f(1) + 1$;
- v) le funzioni non negative;
- vi) le funzioni tali che $f(x) = f(1 - x)$.

— ○ —

Esercizio 3.5 Si consideri lo spazio di Banach $C([0, 1])$ delle funzioni continue dall'intervallo $[0, 1]$ a valori reali con la norma:

$$\|f\| = \sup_{x \in [0, 1]} |f(x)|.$$

a) Si dimostri che il sottoinsieme:

$$X = \{ f \in C([0, 1]) ; f(0) = f(1) = 0 \}$$

è un sottospazio lineare chiuso.

b) Si dimostri che $C([0, 1])$ si scrive come somma diretta di X e di \tilde{X} , con:

$$\tilde{X} = \{ f \in C([0, 1]) ; f(x) = ax + b, a, b \in \mathbb{R} \}.$$

— ○ —

Esercizio 3.6 Nello spazio di Banach $L^2([0, 1])$ si consideri l'operatore lineare:

$$V : f(x) \mapsto V(x)f(x), \quad \forall f \in L^2([0, 1]),$$

dove:

$$V(x) = \begin{cases} x & 0 \leq x \leq \frac{1}{2}, \\ 1-x & \frac{1}{2} \leq x \leq 1. \end{cases}$$

a) Si dimostri che V è limitato.

b) Si calcoli la norma di V .

— ○ —

Esercizio 3.7 Si considerino i seguenti spazi di successioni:

$$c_K = \{ a = \{a_n\}_{n \in \mathbb{N}} ; a_n \in \mathbb{C}, a_n \neq 0 \text{ solo per un numero finito di elementi} \},$$

$$c_0 = \{ a = \{a_n\}_{n \in \mathbb{N}} ; a_n \in \mathbb{C}, \lim_{n \rightarrow \infty} a_n = 0 \},$$

$$l^\infty = \{ a = \{a_n\}_{n \in \mathbb{N}} ; a_n \in \mathbb{C}, \sup_n |a_n| < \infty \},$$

$$l^p = \{ a = \{a_n\}_{n \in \mathbb{N}} ; a_n \in \mathbb{C}, \sum_n |a_n|^p < \infty \}, \quad 1 \leq p < \infty.$$

a) Si dimostri che c_K , c_0 , l^∞ sono spazi normati con la norma:

$$\|a\| = \sup_n |a_n|,$$

e si dimostri che c_K non è completo, mentre c_0 e l^∞ sono di Banach.

b) Si dimostri che, se $1 \leq p < q < \infty$:

$$c_K \subset l^p \subset l^q \subset c_0 \subset l^\infty,$$

dove le inclusioni valgono strettamente.

— ○ —

Esercizio 3.8 Si considerino i seguenti spazi vettoriali di funzioni di variabile reale a valori complessi:

$$\mathcal{K}(\mathbb{R}) = \{ f : \mathbb{R} \rightarrow \mathbb{C} ; f \text{ è limitata q.d. con supporto compatto} \},$$

$$L_0(\mathbb{R}) = \{ f : \mathbb{R} \rightarrow \mathbb{C} ; f \text{ è limitata q.d. e } f(x) \xrightarrow[|x| \rightarrow \infty]{} 0 \},$$

dove il supporto di una funzione è definito come la chiusura dell'insieme in cui la funzione è non nulla:

$$\text{Supp } f = \{x \in \mathbb{R} ; f(x) \neq 0\}^-,$$

per cui:

$$\mathcal{K}(\mathbb{R}) \subset L_0(\mathbb{R}) \subset L^\infty(\mathbb{R}).$$

a) Si dimostri che $\mathcal{K}(\mathbb{R})$ e $L_0(\mathbb{R})$ sono spazi normati con la norma:

$$\|f\|_\infty = \sup_x |f(x)|,$$

e mostrare che $\mathcal{K}(\mathbb{R})$ non è completo, mentre $L_0(\mathbb{R})$ è di Banach.

- b) i) Si dimostri che $\mathcal{K}(\mathbb{R}) \subset L^p(\mathbb{R})$, per ogni $p \geq 1$.
- ii) Si dia un esempio di una funzione $f \in L^\infty(\mathbb{R}) \setminus L_0(\mathbb{R})$ e di una funzione $g \in L_0(\mathbb{R})$ tali che $f, g \notin L^p(\mathbb{R})$ (provando in questo modo che né lo spazio delle funzioni limitate, né quello delle funzioni limitate che si annullano all'infinito sono contenuti in $L^p(\mathbb{R})$).
- iii) Si dia un esempio di una funzione $h \in L^p(\mathbb{R})$ non limitata (mostrando che $L^p(\mathbb{R})$ non può essere contenuto né in $L^\infty(\mathbb{R})$, né in $L_0(\mathbb{R})$).
- iv) Se $1 \leq p < q < \infty$, si dia un esempio di una funzione $f \in L^p(\mathbb{R})$ tale che $f \notin L^q(\mathbb{R})$, e di una funzione $g \in L^q(\mathbb{R})$ che non appartenga a $L^p(\mathbb{R})$.

— ○ —

Esercizio 3.9 Si consideri la seguente successione di funzioni da \mathbb{R} a valori complessi:

$$f_n(t) = \begin{cases} c_n, & 0 \leq t \leq n, \\ 0, & \text{altrove}, \end{cases}$$

con $n \in \mathbb{N}$ e $c_n \in \mathbb{C}$.

- a) Si determini quali condizioni devono soddisfare i valori c_n affinché f_n converga alla funzione nulla, per $n \rightarrow \infty$, nei seguenti spazi di Banach:
 - i) $L^\infty(\mathbb{R})$;
 - ii) $L^1(\mathbb{R})$;
 - iii) $L^2(\mathbb{R})$.
- b) Determinare se è possibile trovare un esempio in cui la successione f_n converge alla funzione nulla
 - i) in $L^\infty(\mathbb{R})$ ma non in $L^1(\mathbb{R})$, né in $L^2(\mathbb{R})$,
 - ii) in $L^1(\mathbb{R})$ ma non in $L^\infty(\mathbb{R})$, né in $L^2(\mathbb{R})$,

- iii) in $L^\infty(\mathbb{R})$ e in $L^2(\mathbb{R})$, ma non in $L^1(\mathbb{R})$,
 iv) in tutti e tre gli spazi.

— ○ —

Esercizio 3.10 Sia $C^1([0, 1])$ lo spazio delle funzioni continue in $[0, 1]$, a valori complessi, dotate di derivata continua in $[0, 1]$ (compresi gli estremi).

- a) Si spieghi per quale motivo le funzioni:

$$m(u) = |u(0)|, \quad n(u) = \left(\int_0^1 |u'(x)|^2 dx \right)^{1/2},$$

con $u \in C^1([0, 1])$, non definiscono una norma, mentre risulta una norma l'applicazione:

$$\|u\|_0 = \left(|u(0)|^2 + \int_0^1 |u'(x)|^2 dx \right)^{1/2}.$$

- b) Si mostri che ogni successione di funzioni convergente con la norma $\|\cdot\|_0$ converge anche uniformemente.

— ○ —

3.6.1 Soluzioni

Soluzione 3.1

a) L'applicazione $\|x\|_1$ è una norma in quanto:

$$\|x\|_1 \geq 0, \quad \|x\|_1 = 0 \implies x_1 = x_2 = 0,$$

$$\|\lambda x\|_1 = |\lambda x_1| + |\lambda x_2| = |\lambda| (|x_1| + |x_2|) = |\lambda| \|x\|_1,$$

$$\|x + y\|_1 = |x_1 + y_1| + |x_2 + y_2| \leq |x_1| + |x_2| + |y_1| + |y_2| = \|x\|_1 + \|y\|_1.$$

Analogamente, $\|x\|_2$ è una norma in quanto:

$$\|x\|_2 \geq 0, \quad \|x\|_2 = 0 \implies x_1 = x_2 = 0,$$

$$\|\lambda x\|_2 = \max \{|\lambda x_1|, |\lambda x_2|\} = |\lambda| \max \{|x_1|, |x_2|\} = |\lambda| \|x\|_2,$$

$$\begin{aligned} \|x + y\|_2 &= \max \{|x_1 + y_1|, |x_2 + y_2|\} \leq \max \{|x_1| + |y_1|, |x_2| + |y_2|\} \leq \\ &\leq \max \{|x_1|, |x_2|\} + \max \{|y_1|, |y_2|\} = \|x\|_1 + \|y\|_1. \end{aligned}$$

Le sfere unitarie sono rappresentate in figura 3.6.

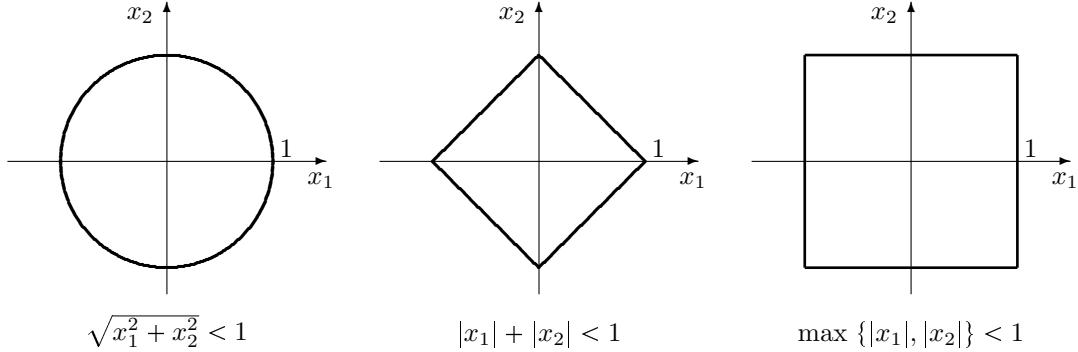


Figura 3.6: Sfere in \mathbb{R}^2 .

b) Abbiamo:

$$\|x\|_0^2 = x_1^2 + x_2^2 \leq x_1^2 + x_2^2 + 2|x_1||x_2| = (|x_1| + |x_2|)^2 = \|x\|_1^2,$$

$$\|x\|_0^2 = x_1^2 + x_2^2 \leq 2 \max \{x_1^2, x_2^2\} = 2 (\max \{|x_1|, |x_2|\})^2 = 2 \|x\|_2^2,$$

da cui $C_1 = 1$, $C_2 = \sqrt{2}$.

Viceversa, per $j = 1, 2$:

$$\begin{aligned} \|x\|_j &= \|x_1 e_1 + x_2 e_2\|_j \leq |x_1| \|e_1\|_j + |x_2| \|e_2\|_j \\ &\leq \sqrt{x_1^2 + x_2^2} \sqrt{\|e_1\|_j^2 + \|e_2\|_j^2} = \sqrt{\|e_1\|_j^2 + \|e_2\|_j^2} \|x\|_0, \end{aligned}$$

da cui $C'_j = \sqrt{\|e_1\|_j^2 + \|e_2\|_j^2} = \sqrt{2}$, essendo, per $j = 0, 1, 2$, $\|e_1\|_j = \|e_2\|_j = 1$.

— ○ —

Soluzione 3.2

a) La funzione $p(x)$ è una norma in quanto:

$$p(x) \geq 0, \quad p(x) = 0 \iff m(x) = n(x) = 0 \iff x = 0,$$

$$\begin{aligned} p(\lambda x) &= \max \{m(\lambda x), n(\lambda x)\} = \max \{|\lambda| m(x), |\lambda| n(x)\} = \\ &= |\lambda| \max \{m(x), n(x)\} = |\lambda| p(x). \end{aligned}$$

$$\begin{aligned} p(x+y) &= \max \{m(x+y), n(x+y)\} \leq \max \{m(x)+m(y), n(x)+n(y)\} \\ &\leq \max \{m(x), n(x)\} + \max \{m(y), n(y)\} = p(x) + p(y). \end{aligned}$$

b) La funzione $q(x)$ è una norma in quanto (ricordiamo che α e β sono positivi):

$$q(x) \geq 0, \quad q(x) = 0 \implies m(x) = 0 \text{ oppure } n(x) = 0$$

$$\text{oppure entrambi nulli} \implies x = 0,$$

$$\begin{aligned} q(\lambda x) &= \alpha m(\lambda x) + \beta n(\lambda x) = \alpha |\lambda| m(x) + \beta |\lambda| n(x) \\ &= |\lambda| (\alpha m(x) + \beta n(x)) = |\lambda| q(x) \end{aligned}$$

$$\begin{aligned} q(x+y) &= \alpha m(x+y) + \beta n(x+y) \leq \alpha (m(x)+m(y)) + \beta (n(x)+n(y)) \\ &= (\alpha m(x) + \beta n(x)) + (\alpha m(y) + \beta n(y)) = q(x) + q(y). \end{aligned}$$

c) Per la funzione $r(x)$ abbiamo:

$$r(x) \geq 0, \quad r(x) = 0 \iff m(x) = n(x) = 0 \iff x = 0,$$

$$\begin{aligned} r(\lambda x) &= \sqrt{m(\lambda x)^2 + n(\lambda x)^2} = \sqrt{|\lambda|^2 m(x)^2 + |\lambda|^2 n(x)^2} \\ &= |\lambda| \sqrt{m(x)^2 + n(x)^2} = |\lambda| r(x). \end{aligned}$$

Per mostrare la diseguaglianza triangolare:

$$r(x+y) \leq r(x) + r(y),$$

essendo la norma sempre una quantità positiva, è sufficiente verificare:

$$r(x+y)^2 \leq [r(x) + r(y)]^2.$$

Usando la definizione di $r(x)$ si ottiene:

$$\begin{aligned}
 r(x+y)^2 &= m(x+y)^2 + n(x+y)^2 \leq [m(x) + m(y)]^2 + [n(x) + n(y)]^2 \\
 &= m(x)^2 + m(y)^2 + 2m(x)m(y) + n(x)^2 + n(y)^2 + 2n(x)n(y) \\
 [r(x) + r(y)]^2 &= r(x)^2 + r(y)^2 + 2r(x)r(y) \\
 &= m(x)^2 + n(x)^2 + m(y)^2 + n(y)^2 \\
 &\quad + 2\sqrt{m(x)^2 + n(x)^2}\sqrt{m(y)^2 + n(y)^2}, \\
 m(x)m(y) + n(x)n(y) &\leq \sqrt{m(x)^2 + n(x)^2}\sqrt{m(y)^2 + n(y)^2}, \\
 2m(x)m(y)n(x)n(y) &\leq m(x)^2n(y)^2 + m(y)^2n(x)^2, \\
 [m(x)n(y) - m(y)n(x)]^2 &\geq 0,
 \end{aligned}$$

la quale è sempre verificata.

— ○ —

Soluzione 3.3

a) Chiaramente $\mathcal{E} \subset \mathcal{A}$ poiché se $f \in \mathcal{E}$:

$$\left| \sum_{n=0}^{\infty} f^{(n)}(0) \frac{x^n}{n!} \right| \leq \sum_{n=0}^{\infty} |f^{(n)}(0)| \frac{|x|^n}{n!} \leq \sum_{n=0}^{\infty} \frac{(\rho_f |x|)^n}{n!} = e^{\rho_f |x|}.$$

Inoltre è un sottospazio lineare in quanto, se $f, g \in \mathcal{E}$, e $\lambda \in \mathbb{R}$, allora le funzioni $f + g$ e λf sono tali che:

$$\begin{aligned}
 |f^{(n)}(0) + g^{(n)}(0)| &\leq |f^{(n)}(0)| + |g^{(n)}(0)| \leq \rho_f^n + \rho_g^n \leq (\rho_f + \rho_g)^n, \\
 |\lambda f^{(n)}(0)| &\leq |\lambda| \rho_f^n \leq \begin{cases} (|\lambda| \rho_f)^n, & \text{se } |\lambda| \geq 1, \\ \rho_f^n & \text{se } |\lambda| < 1. \end{cases}
 \end{aligned}$$

b) \mathcal{P}_N è chiaramente il sottoinsieme dei polinomi di grado al più $N - 1$, che forma un sottospazio vettoriale generato dalla base dei monomi (cardinalità N):

$$1, x, x^2, \dots, x^{N-1}.$$

— ○ —

Soluzione 3.4

a) $C([0, 1])$ è uno spazio vettoriale su \mathbb{R} in quanto la funzione somma di due funzioni continue è continua e se f è continua a valori reali, anche la funzione λf , con λ reale, è continua e a valori reali.

b) Verifichiamo che l'applicazione $\|f\|$ è una norma. Chiaramente essa è sempre positiva e:

$$\|f\| \geq 0, \quad \|f\| = 0 \implies f = 0,$$

essendo l'estremo superiore di una quantità positiva che si annulla solo se la funzione è identicamente nulla in $[0, 1]$. Se moltiplichiamo la funzione f per uno scalare:

$$\|\lambda f\| = \sup_{x \in [0,1]} |\lambda f(x)| = |\lambda| \sup_{x \in [0,1]} |f(x)| = |\lambda| \|f\|.$$

La diseguaglianza triangolare risulta facilmente soddisfatta:

$$\begin{aligned} \|f + g\| &= \sup_{x \in [0,1]} |f(x) + g(x)| \leq \sup_{x \in [0,1]} (|f(x)| + |g(x)|) \\ &\leq \sup_{x \in [0,1]} |f(x)| + \sup_{x \in [0,1]} |g(x)| = \|f\| + \|g\|. \end{aligned}$$

Il fatto che $C([0, 1])$ sia uno spazio di Banach segue dal fatto che tale spazio è completo rispetto alla metrica indotta dalla norma:

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|,$$

come abbiamo già visto nell'esercizio 2.5 nel caso di un intervallo finito arbitrario.

c) Esaminiamo i vari sottoinsiemi.

- i) Il vettore $f = 0$ non appartiene al sottoinsieme, che non risulta un sottospazio.
- ii) Le operazioni algebriche di somma e moltiplicazione per uno scalare tra polinomi di grado minore o uguale a 3 mantengono il carattere di polinomio (il grado al più diminuisce). Abbiamo quindi un sottospazio vettoriale, generato dai monomi:

$$1, x, x^2, x^3,$$

quindi finito dimensionale e chiuso.

- iii) Se abbiamo due funzioni f_1, f_2 , che soddisfano le condizioni:

$$f_1(0) = 2 f_1(1), \quad f_2(0) = 2 f_2(1),$$

allora, se α, β sono reali:

$$\alpha f_1(0) + \beta f_2(0) = 2 \alpha f_1(1) + 2 \beta f_2(1)$$

pertanto il sottoinsieme è un sottospazio. Risulta chiuso in quanto se una successione f_n di funzioni continue che soddisfano la condizione richiesta converge, anche la funzione limite soddisfa la condizione richiesta.

- iv) Il sottoinsieme non risulta un sottospazio perché non è chiuso algebricamente rispetto alla somma:

$$f_1(0) + f_2(0) = f_1(1) + f_2(1) + 2.$$

pur essendo un sottoinsieme topologicamente chiuso.

- v) Il sottoinsieme delle funzioni continue non negative non è chiuso algebricamente rispetto alla moltiplicazione per un numero reale negativo.
- vi) La proprietà di simmetria rispetto al centro dell'intervallo $[0, 1]$ si mantiene anche per le combinazioni lineari e per i limiti di successioni, formando quindi un sottospazio vettoriale topologicamente chiuso.

— ○ —

Soluzione 3.5 Se f e g appartengono a X , e si annullano agli estremi, allora anche una loro combinazione lineare si annulla agli estremi, per cui X forma un sottospazio vettoriale. Per mostrare che X è chiuso consideriamo una successione di funzioni $f_n \in X$ convergente in generale ad una funzione $f \in C([0, 1])$. f appartiene alla chiusura di X (è di aderenza), ma poiché la convergenza uniforme in $C([0, 1])$ implica la convergenza puntuale, si ha anche che le successioni di numeri $f_n(0) = 0$ e $f_n(1) = 0$ convergono a $f(0)$ e $f(1)$ rispettivamente. Pertanto si deve avere $f(0) = f(1) = 0$, cioè $f \in X$ e X è chiuso.

Per la parte (b) si verifica facilmente che ogni funzione $f \in C([0, 1])$ può essere scritta come:

$$f(x) = f_0(x) + a x + b,$$

con:

$$\begin{aligned} a &= f(1) - f(0), & b &= f(0), \\ f_0(x) &= f(x) - (a x + b) \in X. \end{aligned}$$

Ne segue che $C([0, 1]) = X \oplus \tilde{X}$. Inoltre $X \cap \tilde{X} = \{0\}$ in quanto $a x + b$ appartiene a X se e solo se $a = b = 0$.

— ○ —

Soluzione 3.6

a) L'operatore è limitato in quanto per ogni $f \in L^2([0, 1])$:

$$\|V f\|^2 = \int_0^1 |V(x)f(x)|^2 dx \leq \sup_{x \in [0, 1]} |V(x)|^2 \int_0^1 |f(x)|^2 dx = \frac{1}{4} \|f\|^2.$$

b) Per quanto visto in (a) si ha: $\|V\| \leq 1/2$. D'altro canto, per ogni $f \in L^2([0, 1])$ con $\|f\| = 1$, si ha anche $\|V f\|^2 \leq \|V\|^2 \|f\|^2 = \|V\|^2$. Consideriamo allora la seguente famiglia di funzioni in $L^2([0, 1])$:

$$f_\epsilon(x) = \begin{cases} 0 & |x - \frac{1}{2}| \geq \frac{\epsilon}{2}, \\ \frac{1}{\sqrt{\epsilon}} & |x - \frac{1}{2}| < \frac{\epsilon}{2}, \end{cases}$$

che verificano (per $\epsilon < 1$):

$$\|f_\epsilon\|^2 = \int_0^1 dx |f_\epsilon(x)|^2 = \int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}} dx \frac{1}{\epsilon} = 1.$$

Si ha:

$$\begin{aligned} \|Vf_\epsilon\|^2 &= \int_0^1 dx |V(x)f_\epsilon(x)|^2 = \int_{\frac{1-\epsilon}{2}}^{\frac{1+\epsilon}{2}} dx \frac{|V(x)|^2}{\epsilon} = \\ &= \frac{1}{\epsilon} \int_{\frac{1-\epsilon}{2}}^{\frac{1}{2}} x^2 dx + \frac{1}{\epsilon} \int_{\frac{1}{2}}^{\frac{1+\epsilon}{2}} (1-x)^2 dx = \frac{2}{\epsilon} \int_{\frac{1-\epsilon}{2}}^{\frac{1}{2}} x^2 dx = \frac{1}{4} \left(1 - \epsilon + \frac{\epsilon^2}{3}\right). \end{aligned}$$

Allora:

$$\|V\|^2 \geq \|Vf_\epsilon\|^2 = \frac{1}{4} \left(1 - \epsilon + \frac{\epsilon^2}{3}\right) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{4}$$

Quindi si ha anche $\|V\| \geq \frac{1}{2}$ e quindi $\|V\| = \frac{1}{2}$.
— ○ —

Soluzione 3.7

a) La corrispondenza che associa ad una successione a il $\sup_n |a_n|$ è chiaramente una norma negli insiemi c_K , c_0 , l^∞ , infatti:

$$\begin{aligned} \|a\| &= \sup_n |a_n| \geq 0, \\ \|a\| = 0 &\implies \sup_n |a_n| = 0 \implies a_n = 0 \forall n \implies a = 0, \\ \|\lambda a\| &= \sup_n |\lambda a_n| = |\lambda| \sup_n |a_n| = |\lambda| \|a\|, \\ \|a + b\| &= \sup_n |a_n + b_n| \leq \sup_n |a_n| + \sup_n |b_n| = \|a\| + \|b\|. \end{aligned}$$

Lo spazio c_K non è completo. Infatti, la successione $\{a^{(N)}\}_{N \in \mathbb{N}}$ di elementi di c_K con:

$$a^{(N)} = \left(1, \frac{1}{2}, \dots, \frac{1}{N}, 0, 0, \dots\right)$$

è una successione di Cauchy:

$$\|a^{(N)} - a^{(M)}\| = \max \left\{ \frac{1}{N+1}, \frac{1}{M+1} \right\} \xrightarrow[N, M \rightarrow \infty]{} 0,$$

ma la successione limite:

$$a = \left(1, \frac{1}{2}, \dots, \frac{1}{N}, \frac{1}{N+1}, \dots\right)$$

non appartiene a c_K .

c_0 e l^∞ sono invece spazi di Banach. Consideriamo infatti una successione $\{a^{(N)}\}_{N \in \mathbb{N}}$ di Cauchy di elementi di c_0 o l^∞ :

$$\|a^{(N)} - a^{(M)}\| = \sup_n |a_{N,n} - a_{M,n}| < \epsilon, \quad N, M > N_\epsilon,$$

con $\epsilon > 0$ arbitrario e N_ϵ dipendente solo da ϵ . Allora, per ogni n , la successione $\{a_{N,n}\}_{N \in \mathbb{N}}$ è di Cauchy nell'insieme \mathbb{C} , completo, e quindi convergente ad un numero complesso a_n :

$$a_n = \lim_{N \rightarrow \infty} a_{N,n}, \quad \forall n.$$

La successione $\{a_n\}_{n \in \mathbb{N}}$ risulta il limite nella topologia di c_0 o l^∞ , in quanto operando il limite per $M \rightarrow \infty$ nella condizione di Cauchy precedente abbiamo:

$$\|a^{(N)} - a\| = \sup_n |a_{N,n} - a_n| \leq \epsilon, \quad \forall N > N_\epsilon.$$

Se $a^{(N)}$ appartiene a l^∞ allora anche il limite a è una successione limitata:

$$\begin{aligned} |a_n| &\leq |a_{N,n}| + |a_n - a_{N,n}| \leq \sup_n |a_{N,n}| + \sup_n |a_{N,n} - a_n| \\ &= \|a^{(N)}\| + \|a^{(N)} - a\| \leq \|a^{(N)}\| + \epsilon < \infty, \quad N > N_\epsilon, \end{aligned}$$

mentre se $a^{(N)}$ appartiene a c_0 , anche il limite vi appartiene:

$$|a_n| \leq |a_{N,n} - a_n| + |a_{N,n}| \leq \epsilon + |a_{N,n}|, \quad N > N_\epsilon,$$

e per n sufficientemente grande $a_{N,n}$ è piccolo a piacere.

b) È ovvio che $c_K \subset l^p$ in quanto se $a \in c_K$ la serie che definisce $\|a\|_p$:

$$\|a\|_p = \left(\sum_n |a_n|^p \right)^{1/p},$$

ha un numero finito di termini non nulli, quindi convergente. Inoltre $l^p \subset c_0$ in quanto il termine n -esimo di una serie assolutamente convergente tende sicuramente a 0. L'inclusione è propria in quanto qualsiasi serie assolutamente convergente con un numero infinito di termini non è un elemento di c_K , mentre possiamo facilmente trovare una successione convergente a zero la cui serie diverge assolutamente:

$$a_n = \frac{1}{n^{1/p}}.$$

L'inclusione $c_0 \subset l^\infty$ è garantita dal fatto che una successione convergente è necessariamente limitata e l'inclusione è propria in quanto una successione stazionaria non nulla $a_n = \text{cost.} = b \neq 0$ è limitata ma non in c_0 .

Notiamo infine che se una successione $a = \{a_n\}_{n \in \mathbb{N}} \in l^p$, allora $|a_n|$ deve tendere a zero più rapidamente di $1/n^\alpha$, con $\alpha > 1/p$. Quindi, se $p < q$ si ha anche $\alpha > 1/q$, e $a \in l^q$, $l^p \subset l^q$. L'inclusione è propria in quanto, ad esempio, la successione $a_n = 1/n^\alpha$, con $1/q < \alpha < 1/p$, appartiene a l^q ma non a l^p .

— ○ —

Soluzione 3.8

a) $\mathcal{K}(\mathbb{R})$ e $L_0(\mathbb{R})$ sono chiaramente algebricamente chiusi rispetto alle operazioni di somma e moltiplicazione per uno scalare, in quanto se due funzioni appartengono a tali insiemi anche una loro combinazione lineare verifica le medesime condizioni. Pertanto sono sottospazi vettoriali e normati di $L^\infty(\mathbb{R})$.

Il sottospazio $L_0(\mathbb{R})$ risulta topologicamente chiuso, quindi di Banach. Infatti, consideriamo una successione di Cauchy $f_n \in L_0(\mathbb{R})$ convergente (in maniera uniforme) ad una funzione $f \in L^\infty(\mathbb{R})$ (essendo quest'ultimo spazio di Banach). Allora vale anche la convergenza puntuale:

$$|f(x) - f_n(x)| \xrightarrow[n \rightarrow \infty]{} 0,$$

e, sapendo che $f_n(x) \xrightarrow[|x| \rightarrow \infty]{} 0$, si ha:

$$\lim_{x \rightarrow \infty} |f(x)| = \lim_{|x| \rightarrow \infty} |f(x) - f_n(x)| \xrightarrow[n \rightarrow \infty]{} 0,$$

per cui $f \in L_0(\mathbb{R})$ che risulta chiuso.

Il sottospazio $\mathcal{K}(\mathbb{R})$ invece non è chiuso, come mostrato dal seguente contro-esempio. Si consideri la successione di funzioni, con $n \in \mathbb{N}$, in $\mathcal{K}(\mathbb{R})$:

$$g_n(x) = \begin{cases} e^{-x^2} & -n \leq x \leq n \\ 0 & \text{altrove} \end{cases}$$

che converge (nella norma uniforme) alla funzione $g(x) = e^{-x^2}$:

$$\|g_n - g\| = \sup_x |g_n(x) - g(x)| = e^{-n^2} \xrightarrow[n \rightarrow \infty]{} 0,$$

e $g \in L_0(\mathbb{R})$ ma non appartiene a $\mathcal{K}(\mathbb{R})$.

Consideriamo ora la parte b).

i) Sia $f \in \mathcal{K}(\mathbb{R})$, per cui il suo supporto sarà contenuto in un intervallo $[-K, K]$ con $K > 0$ opportuno. Allora, se $1 \leq p < \infty$:

$$\int_{\mathbb{R}} |f(x)|^p dx = \int_{-K}^K |f(x)|^p dx \leq \left(\sup_{x \in [-K, K]} |f(x)| \right)^p \int_{-K}^K dx = 2K \|f\|_\infty^p < \infty.$$

ii) Fissato un valore di p ($1 \leq p < \infty$), si prenda la funzione costante $f(x) = 1$, che chiaramente sta in $L^\infty(\mathbb{R}) \setminus L_0(\mathbb{R})$ ma non in $L^p(\mathbb{R})$ in quanto il suo integrale diverge. Si

consideri quindi la funzione in $L_0(\mathbb{R})$ ($K > 0$):

$$g(x) = \begin{cases} 1 & |x| \leq K, \\ \frac{1}{|x|^{\frac{1}{p}}} & |x| > K. \end{cases}$$

Questa funzione non appartiene a $L^p(\mathbb{R})$ in quanto:

$$\int_{\mathbb{R}} |g(x)|^p dx = 2 \int_0^K dx + 2 \int_K^{\infty} \frac{1}{x} dx = 2K + 2[\ln x]_K^{\infty} = \infty.$$

iii) Con p fissato ($1 \leq p < \infty$), e $\epsilon > 0$ (“piccolo”):

$$h(x) = \begin{cases} \frac{1}{|x|^{\frac{1-\epsilon}{p}}} & |x| \leq K, \\ 0 & |x| > K, \end{cases}$$

che non è limitata (in un intorno dell’origine) ma appartiene a $L^p(\mathbb{R})$ in quanto:

$$\int_{\mathbb{R}} |h(x)|^p dx = 2 \int_0^K \frac{1}{x^{1-\epsilon}} dx = 2 \frac{K^\epsilon}{\epsilon} < \infty.$$

iv) Notiamo che se una funzione $f \in L^p(\mathbb{R})$ per un p opportuno, allora:

- Se tende a zero all’infinito, vi deve tendere almeno come $1/|x|^\alpha$ con $\alpha > 1/p$.
- Se diverge in un punto x_0 , deve divergere al più come $1/|x - x_0|^\beta$, con $\beta < 1/p$.

Allora, se $1 \leq p < q < \infty$, la funzione:

$$f(x) = \begin{cases} \frac{1}{|x|^\beta} & |x| \leq K, \\ 0 & |x| > K, \end{cases} \quad \text{con } \frac{1}{q} < \beta < \frac{1}{p},$$

appartiene a $L^p(\mathbb{R})$ ma non a $L^q(\mathbb{R})$. Viceversa la funzione:

$$g(x) = \begin{cases} 1 & |x| \leq K, \\ \frac{1}{|x|^\alpha} & |x| > K. \end{cases} \quad \text{con } \frac{1}{q} < \alpha < \frac{1}{p},$$

appartiene a $L^q(\mathbb{R})$ ma non a $L^p(\mathbb{R})$.

— ○ —

Soluzione 3.9

a) Nei casi richiesti abbiamo:

i. in $L^\infty(\mathbb{R})$:

$$\|f_n\|_\infty = \sup_t |f_n(t)| = |c_n| \xrightarrow{n \rightarrow \infty} 0 \iff c_n \xrightarrow{n \rightarrow \infty} 0,$$

ii. in $L^1(\mathbb{R})$:

$$\|f_n\|_1 = \int_{\mathbb{R}} |f_n(t)| dt = |c_n| n \xrightarrow{n \rightarrow \infty} 0 \iff n c_n \xrightarrow{n \rightarrow \infty} 0,$$

iii. in $L^2(\mathbb{R})$:

$$\|f_n\|_2^2 = \int_{\mathbb{R}} |f_n(t)|^2 dt = |c_n|^2 n \xrightarrow{n \rightarrow \infty} 0 \iff \sqrt{n} c_n \xrightarrow{n \rightarrow \infty} 0,$$

b) Abbiamo:

i) un esempio può essere dato da:

$$c_n = \frac{1}{\sqrt{n}},$$

ii) non è possibile fornire un esempio in quanto:

$$n c_n \xrightarrow{n \rightarrow \infty} 0 \implies c_n \xrightarrow{n \rightarrow \infty} 0,$$

iii) è possibile dare il seguente esempio:

$$c_n = \frac{1}{n},$$

iv) un esempio è fornito da:

$$c_n = \frac{1}{n^2}.$$

— ○ —

Soluzione 3.10

a) Sia $m(u)$ che $n(u)$ sono positive, ma non strettamente definite positive, in quanto $m(u) = 0$ implica l'annullarsi di u solo nell'origine, $u(0) = 0$, ma non necessariamente identicamente, mentre $n(u) = 0$ è verificato per tutte le funzioni costanti nell'intervallo $[0, 1]$, ma non necessariamente nulle. Tutte le altre proprietà di norma sono verificate.

Infatti se $u, v \in C^1([0, 1])$, e $\lambda \in \mathbb{C}$:

$$\begin{aligned} m(\lambda u) &= |\lambda u(0)| = |\lambda| |u(0)| = |\lambda| m(u), \\ n(\lambda u)^2 &= \int_0^1 |\lambda u(x)|^2 dx = |\lambda|^2 \int_0^1 |u(x)|^2 dx = |\lambda|^2 n(u)^2, \\ m(u + v) &= |u(0) + v(0)| \leq |u(0)| + |v(0)| = m(u) + m(v), \\ n(u + v) &= \left(\int_0^1 |u'(x) + v'(x)|^2 dx \right)^{1/2} \\ &\leq \left(\int_0^1 |u'(x)|^2 dx \right)^{1/2} + \left(\int_0^1 |v'(x)|^2 dx \right)^{1/2} = n(u) + n(v), \end{aligned}$$

dove abbiamo usato la diseguaglianza di Minkowski per le funzioni $u'(x)$ e $v'(x)$, per cui $m(u)$ e $n(u)$ definiscono due seminorme.

L'espressione $\|u\|_0$ definisce invece una norma in quanto:

- è sempre positiva:

$$\|u\|_0 \geq 0,$$

e si annulla solo se:

$$u'(x) = 0, \quad u(0) = 0,$$

ovvero se e solo se $u(x) = \text{cost.} = u(0) = 0$;

- è assolutamente omogenea rispetto alla moltiplicazione per uno scalare $\lambda \in \mathbb{C}$:

$$\begin{aligned} \|\lambda u\|_0 &= \sqrt{|\lambda u(0)|^2 + \int_0^1 |\lambda u'(x)|^2 dx} \\ &= |\lambda| \sqrt{|u(0)|^2 + \int_0^1 |u'(x)|^2 dx} = |\lambda| \|u\|_0; \end{aligned}$$

- vale la diseguaglianza triangolare:

$$\begin{aligned} \|u + v\|_0 &= \sqrt{m(u + v)^2 + n(u + v)^2} \leq \sqrt{[m(u) + m(v)]^2 + [n(u) + n(v)]^2} \\ &\leq \sqrt{m(u)^2 + n(u)^2} + \sqrt{m(v)^2 + n(v)^2} = \|u\|_0 + \|v\|_0. \end{aligned}$$

per quanto visto in precedenza.

b) Supponiamo ora di avere una successione $\{u_n\}_{n \in \mathbb{N}}$ di funzioni in $C^1([0, 1])$ convergenti nella norma $\|\cdot\|_0$ ad una funzione $u(x)$. Essendo in uno spazio vettoriale, per linearità, possiamo supporre $u = 0$ (possiamo ridefinire la successione $v_n = u_n - u$). Allora:

$$\|u_n\|_0 \xrightarrow[n \rightarrow \infty]{} 0 \iff u_n(0) \xrightarrow[n \rightarrow \infty]{} 0, \quad \int_0^1 |u'_n(x)|^2 dx \xrightarrow[n \rightarrow \infty]{} 0.$$

Possiamo inoltre scrivere:

$$u_n(x) = u_n(0) + \int_0^x u'_n(y) dy ,$$

da cui:

$$\begin{aligned} |u_n(x)| &\leq |u_n(0)| + \left| \int_0^x u'_n(y) dy \right| \leq |u_n(0)| + \int_0^x |u'_n(y)| dy \\ &\leq |u_n(0)| + \int_0^1 |u'_n(y)| dy , \\ \sup_{x \in [0,1]} |u_n(x)| &\leq |u_n(0)| + \int_0^1 |u'_n(y)| dy \\ &\leq |u_n(0)| + \left(\int_0^1 |u'_n(y)|^2 dy \right)^{1/2} \xrightarrow[n \rightarrow \infty]{} 0 . \end{aligned}$$

L'ultima diseguaglianza deriva dal fatto che $u'_n \in L^2([0,1])$ e dalla diseguaglianza di Hölder:

$$\int_0^1 |u'_n(y)| \cdot 1 dy \leq \left(\int_0^1 |u'_n(y)|^2 dy \right)^{1/2} \left(\int_0^1 1 dy \right)^{1/2} = \left(\int_0^1 |u'_n(y)|^2 dy \right)^{1/2} .$$

— ○ —

Capitolo 4.

Spazi di Hilbert.

4.1 Prodotti scalari

Vogliamo ora parlare di una categoria di spazi molto importanti dal punto di vista delle applicazioni alla meccanica quantistica: i cosiddetti spazi di Hilbert, in cui si può definire un prodotto scalare che rende tali spazi geometricamente simili agli spazi vettoriali finito-dimensionalì.

4.1.1 Forme sesquilineari

Iniziamo discutendo il concetto generale di forma sesquilineare.

Def. 4.1 Sia E uno spazio vettoriale sul campo \mathbb{K} ($\mathbb{K} = \mathbb{R}$ o \mathbb{C}). Una applicazione:

$$q : E \times E \rightarrow \mathbb{K},$$

è detta una **forma sesquilineare** se $q(x, y)$ è lineare in y e antilineare in x :

$$q(x, \alpha y_1 + \beta y_2) = \alpha q(x, y_1) + \beta q(x, y_2), \quad (4.1)$$

$$q(\alpha x_1 + \beta x_2, y) = \bar{\alpha} q(x_1, y) + \bar{\beta} q(x_2, y), \quad (4.2)$$

con $\alpha, \beta \in \mathbb{K}$, $x, x_1, x_2, y, y_1, y_2 \in E$ e dove $\bar{\alpha}, \bar{\beta}$ denotano i complessi coniugati di α e β . Nel caso $\mathbb{K} = \mathbb{R}$ l'antilinearità è sostituita dalla linearità semplice e la forma q è detta anche semplicemente una **forma bilineare**.

Una forma sesquilineare q su E è detta **hermitiana** se:

$$q(x, y) = \overline{q(y, x)} \quad \forall x, y \in E. \quad (4.3)$$

Se $\mathbb{K} = \mathbb{R}$ una forma hermitiana è detta più semplicemente **simmetrica**.

Una forma sesquilineare q su E è detta **non degenere** se:

$$q(x, y) = 0 \quad \forall y \in E \implies x = 0. \quad (4.4)$$

In caso contrario è detta **degenera**.

Osservazione. Occorre fare attenzione alle convenzioni usate, specie se si consultano testi diversi. Possiamo capire se un testo è stato scritto da un “matematico”, oppure da un “fisico”: i matematici operano una scelta diversa, assumono la linearità di una forma sesquilineare $q(x, y)$ rispetto al primo argomento (x) e l’antilinearità rispetto al secondo argomento (y), mentre i fisici, probabilmente in seguito alle impostazioni sviluppate da Dirac, seguono le convenzioni enunciate sopra. Entrambe le convenzioni sono valide, ma è necessario mantenere coerentemente le convenzioni in tutto quello che segue.

Esempio 4.1 Supponiamo E uno spazio vettoriale finito dimensionale sul campo dei complessi, e $\{e_j\}_{1 \leq j \leq n}$ una sua base. Allora, se:

$$x = \sum_{j=1}^n x_j e_j, \quad y = \sum_{j=1}^n y_j e_j,$$

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

$$q(x, y) = \sum_{j,k=1}^n \overline{x_j} y_k q(e_j, e_k),$$

e la forma sesquilineare è completamente determinata dai valori assunti sui vettori di base, cioè dalla matrice $\alpha_{jk} = q(e_j, e_k)$. Se q è hermitiana si ha $\alpha_{jk} = \overline{\alpha_{kj}}$, e se q è non degenera la matrice α_{jk} è invertibile.

Viceversa, assegnata una base $\{e_j\}_{1 \leq j \leq n}$ ed una matrice quadrata $n \times n$ di numeri complessi o reali, posso costruire una forma sesquilineare tramite la formula:

$$q(x, y) = \sum_{j,k=1}^n \overline{x_j} y_k \alpha_{jk},$$

e l’hermiticità della matrice e la sua invertibilità comportano rispettivamente l’hermiticità e la non degenerazione della forma sesquilineare.

Sostanzialmente, sfruttando l’equivalenza di uno spazio vettoriale finito-dimensionale con lo spazio \mathbb{C}^n , le proprietà di una forma sesquilineare sono riconducibili alle caratteristiche di una matrice.

— ○ —

Teo. 4.1 Sia q una forma sesquilineare su E , spazio vettoriale sul campo \mathbb{K} .

1. Sia $\mathbb{K} = \mathbb{C}$. Allora:

- a) q è completamente determinata dai suoi valori diagonali $q(x, x)$ al variare di $x \in E$.
- b) q è hermitiana se e solo se $q(x, x) \in \mathbb{R}$ per ogni $x \in E$.

2. Sia $\mathbb{K} = \mathbb{R}$. Allora se q è simmetrica è completamente determinata dai suoi valori diagonali $q(x, x)$ al variare di $x \in E$.

Dim. 4.1 Per la prima parte (caso $\mathbb{K} = \mathbb{C}$) si può verificare direttamente (usando le proprietà di linearità e antilinearità) la validità (per ogni forma sesquilineare q) della seguente **identità di polarizzazione**:

$$\begin{aligned} q(x, y) &= \frac{1}{4} \{ q(x+y, x+y) - q(x-y, x-y) \\ &\quad - i q(x+iy, x+iy) + i q(x-iy, x-iy) \} , \end{aligned} \tag{4.5}$$

da cui si provano direttamente le affermazioni (a) e (b).

La seconda parte è conseguenza del fatto che se q è simmetrica si ha, ad esempio:

$$q(x, y) = \frac{1}{2} \{ q(x+y, x+y) - q(x, x) - q(y, y) \} , \tag{4.6}$$

identicamente per ogni $x, y \in E$.

— ■ —

Osservazione. L'affermazione (1a) non è vera in generale se $\mathbb{K} = \mathbb{R}$. Ad esempio in \mathbb{R}^2 consideriamo la forma antisimmetrica:

$$q(x, y) = x_1 y_2 - x_2 y_1 ,$$

abbiamo $q(x, x) = 0$ per ogni $x \in \mathbb{R}^2$ ma q non è identicamente nulla, come suggerito invece dalla identità di polarizzazione. Per avere una identità simile alla identità di polarizzazione deve essere appunto richiesta esplicitamente la simmetria nel caso reale.

Def. 4.2 Una forma sesquilineare **hermitiana** è detta **positiva** se verifica $q(x, x) \geq 0$ per ogni $x \in E$, e **definita positiva** se $q(x, x) > 0$ per ogni $x \neq 0$ (quindi $q(x, x) = 0 \iff x = 0$).

Sia $\mathbb{K} = \mathbb{C}$, la positività di $q(x, x)$ implica l'hermiticità, in quanto è sottinteso che un numero positivo è anche reale, mentre se $\mathbb{K} = \mathbb{R}$ l'ipotesi di simmetria non è deducibile dalla positività, per cui deve essere richiesta esplicitamente.

Teo. 4.2 Sia E uno spazio vettoriale sul campo \mathbb{K} , e q una forma sesquilineare hermitiana **positiva** su E . Allora:

i) Vale la seguente **diseguaglianza di Schwarz**:

$$|q(x, y)| \leq q(x, x)^{\frac{1}{2}} q(y, y)^{\frac{1}{2}} \quad \forall x, y \in E . \tag{4.7}$$

Se q è definita positiva vale l'uguaglianza se e solo se i vettori x e y sono l'uno multiplo dell'altro: $x = \lambda y$, $\lambda \in \mathbb{K}$.

ii) Vale la seguente diseguaglianza:

$$q(x+y, x+y)^{\frac{1}{2}} \leq q(x, x)^{\frac{1}{2}} + q(y, y)^{\frac{1}{2}} \quad \forall x, y \in E. \quad (4.8)$$

Se q è definita positiva l'uquaglianza vale se e solo se i vettori x e y sono lineamente dipendenti: $x = \lambda y$, con $\lambda \geq 0$ (assumendo y non nullo).

Dim. 4.2 Sia che si tratti di uno spazio reale o complesso la positività implica che:

$$0 \leq q(x+\lambda y, x+\lambda y) = q(x, x) + 2\operatorname{Re}\{\lambda q(x, y)\} + |\lambda|^2 q(y, y),$$

per ogni $x, y \in E$ e $\lambda \in \mathbb{K}$. Posto:

$$\lambda = \mu e^{-i \arg q(x, y)},$$

con μ reale ed arbitrario, otteniamo la seguente diseguaglianza in \mathbb{R} verificata per ogni μ :

$$q(x, x) + 2\mu |q(x, y)| + \mu^2 q(y, y) \geq 0.$$

Se $q(y, y) = 0$, la validità della precedente relazione per ogni valore reale di μ comporta che $q(x, y) = 0$, e la diseguaglianza di Schwarz è banalmente verificata. Se invece $q(y, y) \neq 0$ la condizione di positività implica la negatività del discriminante:

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

da cui la diseguaglianza di Schwarz.

Chiaramente se x e y sono proporzionali tra loro vale l'uguaglianza nella (4.7). Supponendo ora la forma definita positiva, con $y \neq 0$, se, x non è proporzionale a y , allora $x + \lambda y$ non potrà mai essere nullo, per cui si deve avere, ripetendo il ragionamento precedente:

$$q(x, x) + 2\mu |q(x, y)| + \mu^2 q(y, y) > 0,$$

per ogni μ reale e di conseguenza $\Delta < 0$, per cui non può valere l'uguaglianza nella (4.7).

Consideriamo ora la seconda diseguaglianza. Quadrando entrambi i membri, questa è equivalente alla diseguaglianza:

$$q(x+y, x+y) \leq q(x, x) + q(y, y) + 2q(x, x)^{\frac{1}{2}}q(y, y)^{\frac{1}{2}},$$

e, sviluppando algebricamente:

$$\operatorname{Re} q(x, y) \leq q(x, x)^{\frac{1}{2}}q(y, y)^{\frac{1}{2}},$$

garantita a sua volta dalla precedente disuguaglianza di Schwarz:

$$\Re e q(x, y) \leq |q(x, y)| \leq q(x, x)^{\frac{1}{2}} q(y, y)^{\frac{1}{2}},$$

Sia ora q definita positiva, per avere l'uguaglianza:

$$\Re e q(x, y) = q(x, x)^{\frac{1}{2}} q(y, y)^{\frac{1}{2}},$$

devo avere anche la validità dell'uguaglianza nella disuguaglianza di Schwarz, per cui x e y sono proporzionali tra loro, $x = \lambda y$, con λ opportuno (assumendo, senza essere restrittivi, y non nullo). Inoltre deve valere anche l'uguaglianza:

$$\Re e q(x, y) = |q(x, y)|,$$

$$\Re e \bar{\lambda} q(y, y) = |\lambda| q(y, y)$$

possibile solo se λ è reale con $\lambda \geq 0$.

Viceversa, se $x = \lambda y$, con $\lambda \geq 0$, possiamo verificare direttamente l'uguaglianza:

$$\begin{aligned} q(x + y, x + y)^{\frac{1}{2}} &= (\lambda + 1) q(y, y)^{\frac{1}{2}} \\ q(x, x)^{\frac{1}{2}} + q(y, y)^{\frac{1}{2}} &= (\lambda + 1) q(y, y)^{\frac{1}{2}}. \end{aligned}$$

— ■ —

4.1.2 Prodotto scalare

Una forma sesquilineare definita positiva costituisce quello che si chiama un prodotto scalare.

Def. 4.3 Si definisce **prodotto scalare** una forma sesquilineare (hermitiana) definita positiva. Nel caso di uno spazio vettoriale complesso la forma è detta più propriamente un **prodotto scalare hermitiano**. Nel caso di uno spazio reale la dizione esatta è **prodotto scalare euclideo**. In generale risulterà chiaro dal contesto di quale caso si tratti ed useremo la forma semplificata di prodotto scalare.

Spesso si usa indicare il prodotto scalare di due vettori x e y mediante la notazione $\langle x, y \rangle$ (che adotteremo in seguito) con le parentesi angolari (come per i funzionali lineari dello spazio duale algebrico), oppure mediante le parentesi semplici (x, y).

Un'altra notazione è $\langle x | y \rangle$, introdotta da Dirac e usata frequentemente nei testi di meccanica quantistica. La parte $|y\rangle$ indica formalmente un vettore (come la notazione

\vec{y}) detto **ket**, con proprietà algebriche lineari di composizione, analogo di un “vettore colonna”:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

in \mathbb{C}^n , mentre $\langle x|$, detto **bra**, viene identificato con il “trasposto complesso coniugato” di $|x\rangle$, analogo di un “vettore riga”:

$$(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n),$$

con conseguenti proprietà antilineari di composizione, ed il prodotto scalare è il prodotto di un bra con un ket: “braket” (dal termine inglese per “parentesi”), con proprietà analoghe al prodotto (matriciale) delle precedenti matrici “riga” e “colonna”.

In definitiva possiamo dire che un prodotto scalare definito su uno spazio vettoriale E (in generale complesso) è una applicazione:

$$E \times E \longrightarrow \mathbb{K},$$

$$(x, y) \longrightarrow \langle x, y \rangle,$$

che gode delle seguenti proprietà:

$$\boxed{\langle x, x \rangle \geq 0,} \quad \forall x \in E, \quad (4.9)$$

$$\boxed{\langle x, x \rangle = 0 \iff x = 0,} \quad (4.10)$$

$$\boxed{\langle y, x \rangle = \overline{\langle x, y \rangle}}, \quad \forall x, y \in E, \quad (4.11)$$

$$\boxed{\langle x, \alpha y \rangle = \alpha \langle x, y \rangle}, \quad \forall x, y \in E, \alpha \in \mathbb{K}, \quad (4.12)$$

$$\boxed{\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle}, \quad \forall x, y_1, y_2 \in E. \quad (4.13)$$

Notiamo che la condizione di hermiticità (4.11), combinata con le proprietà di linearità rispetto al ket, fornisce le proprietà di antilinearità rispetto al bra:

$$\langle \alpha x, y \rangle = \overline{\alpha} \langle x, y \rangle, \quad \forall x, y \in E, \alpha \in \mathbb{K}, \quad (4.14)$$

$$\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle, \quad \forall x_1, x_2, y \in E. \quad (4.15)$$

Def. 4.4 Uno spazio vettoriale E su \mathbb{K} dotato di prodotto scalare è detto **pre-hilbertiano** oppure **euclideo**.

Come conseguenza immediata del teorema 4.2 abbiamo che se q è una forma sesquilineare definita positiva, allora $q(x, x)^{\frac{1}{2}}$ definisce una norma, per cui uno spazio pre-hilbertiano risulta uno spazio normato con la **norma indotta dal prodotto scalare**:

$$\|x\| = \sqrt{\langle x, x \rangle}. \quad (4.16)$$

In effetti la (4.9) ci garantisce che $\|x\|$ è un numero reale positivo e la (4.10) che $\|x\| = 0$ solo se $x = 0$. Le proprietà di linearità e antilinearità forniscono l'omogeneità assoluta della norma. Il teorema 4.2 esprime la disuguaglianza triangolare

$$\|x + y\| \leq \|x\| + \|y\|. \quad (4.17)$$

La disuguaglianza di Schwarz diviene in uno spazio pre-hilbertiano:

$$|\langle x, y \rangle| \leq \|x\| \|y\|, \quad (4.18)$$

con l'uguaglianza valida solo quando i due vettori x, y sono linearmente dipendenti tra loro.

Uno spazio pre-hilbertiano risulta quindi anche uno spazio normato, metrico e topologico, e sorge quindi naturale chiedersi se sia completo. Essendo il prodotto scalare una qualsiasi forma sesquilineare hermitiana e definita positiva, la completezza non è in generale assicurata, per cui è una proprietà aggiuntiva che eventualmente deve essere richiesta. Quando si ha completezza si parla più propriamente di uno spazio di Hilbert.

Def. 4.5 Uno spazio vettoriale \mathcal{H} è detto **spazio di Hilbert** se verifica le seguenti condizioni:

1. \mathcal{H} è uno spazio pre-hilbertiano (ovvero in \mathcal{H} è definito un prodotto scalare $\langle x, y \rangle$ tra due arbitrari vettori x, y di \mathcal{H}).
2. \mathcal{H} è completo come spazio metrico con la metrica indotta dalla norma, a sua volta indotta dal prodotto scalare:

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}, \quad (4.19)$$

cioè \mathcal{H} ha la struttura di spazio di Banach.

Osservazione. In origine (ed ancora per i puristi) nella definizione di spazio di Hilbert veniva richiesta anche la condizione che le dimensioni dello spazio vettoriale \mathcal{H} fossero infinite. In effetti uno spazio vettoriale a dimensioni finite è identificabile con \mathbb{C}^n , in cui si introduce il prodotto scalare ordinario:

$$\langle x, y \rangle = \sum_{j=1}^n \overline{x_j} y_j,$$

e \mathbb{C}^n è completo. Per distinguere questa situazione, in un certo senso banale, si imponeva la dimensionalità infinita, ma in fisica è prevalso l'uso di prescindere da tale condizione e si considerano spazi di Hilbert anche spazi vettoriali a dimensioni finite. Ultimamente tale convenzione è adottata anche dai matematici in quanto i risultati più importanti non si avvalgono della richiesta di una dimensionalità infinita.

Esempio 4.2 Un esempio importante di spazio di Hilbert è costituito dallo spazio $L^2(\Omega)$, in cui si può definire il prodotto scalare:

$$\langle f, g \rangle = \int_{\Omega} \overline{f(x)} g(x) dx, \quad (4.20)$$

da cui si ottiene la norma in $L^2(\Omega)$:

$$\|f\|_2 = \left(\int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}} = \sqrt{\langle f, f \rangle}$$

e sappiamo che con tale norma lo spazio è completo. Notiamo che l'esistenza dell'integrale (4.20) è garantita dal teorema 3.20 con $p = q = 2$, e la disuguaglianza di Hölder (3.118) coincide con la disuguaglianza di Schwarz (4.18).

— ○ —

Esempio 4.3 Analogamente allo spazio precedente, risulta di Hilbert anche lo spazio l^2 in cui si definisce il prodotto scalare:

$$\langle x, y \rangle = \sum_{j=1}^{\infty} \overline{x_j} y_j. \quad (4.21)$$

— ○ —

Vediamo subito un aspetto importante della topologia introdotta tramite il prodotto scalare, la continuità stessa del prodotto scalare rispetto ai suoi argomenti.

Lem. 4.3 Sia E uno spazio pre-hilbertiano. Allora l'applicazione:

$$\begin{aligned} \phi : \mathcal{H} \times \mathcal{H} &\longrightarrow \mathbb{K}, \\ (x, y) &\longrightarrow \langle x, y \rangle, \end{aligned}$$

è continua.

Dim. 4.3 Per la disuguaglianza di Schwarz abbiamo:

$$\begin{aligned} |\phi(x, y) - \phi(x_0, y_0)| &= |\langle x, y \rangle - \langle x_0, y_0 \rangle| \\ &= |\langle x_0 + x - x_0, y_0 + y - y_0 \rangle - \langle x_0, y_0 \rangle| \\ &= |\langle x_0, y - y_0 \rangle + \langle x - x_0, y_0 \rangle + \langle x - x_0, y - y_0 \rangle| \\ &\leq \|x_0\| \|y - y_0\| + \|x - x_0\| \|y_0\| + \|x - x_0\| \|y - y_0\| \end{aligned}$$

che può essere reso piccolo a piacere quando $\|x - x_0\| \rightarrow 0$ e $\|y - y_0\| \rightarrow 0$.

— ■ —

Una conseguenza immediata di tale lemma è data dal fatto che uno spazio pre-hilbertiano può essere completato in uno spazio di Hilbert.

Teo. 4.4 *Ogni spazio pre-hilbertiano E ammette come completamento uno spazio di Hilbert \mathcal{H} .*

Dim. 4.4 Ricordando il risultato generale sul completamento di uno spazio metrico, sappiamo costruire il completamento mediante le classi di equivalenza di successioni di Cauchy in E . Occorre mostrare che in tale spazio \mathcal{H} è possibile definire un prodotto scalare “coerente” con il prodotto scalare definito in E . Siano x^* e y^* gli elementi del completamento individuati dalle successioni di Cauchy $\{x_n\}$ e $\{y_n\}$, e poniamo:

$$\langle x^*, y^* \rangle = \lim_{n \rightarrow \infty} \langle x_n, y_n \rangle.$$

È immediato verificare che i numeri $\langle x_n, y_n \rangle$ formano una successione di Cauchy e pertanto convergenti. Poichè il prodotto scalare risulta una funzione continua dei suoi argomenti la definizione sopra è indipendente dalla scelta delle successioni $\{x_n\}$ e $\{y_n\}$ mediante le quali “approssimiamo” gli elementi x^* e y^* , e inoltre, sempre per continuità, valgono tutte le proprietà del prodotto scalare. Pertanto la definizione è ben posta, e la norma e la distanza in \mathcal{H} sono esprimibili tramite il prodotto scalare testè definito.

— ■ —

Non ci soffermeremo oltre sulle proprietà degli spazi pre-hilbertiani, ma assumeremo in quanto segue di avere sempre uno spazio di Hilbert, anche se alcuni risultati (se non coinvolgono la completezza) possono essere validi più in generale in uno spazio pre-hilbertiano.

4.2 Geometria di uno spazio di Hilbert

La diseguaglianza di Schwarz permette di introdurre delle nozioni di carattere geometrico in uno spazio di Hilbert. Consideriamo due vettori non nulli x e y , allora i vettori $\frac{x}{\|x\|}$, $\frac{y}{\|y\|}$ individuano due direzioni (o sottospazi unidimensionali) in tale spazio, e, nel caso di uno spazio di Hilbert reale, la diseguaglianza di Schwarz comporta che:

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \|y\|} \leq +1$$

con le uguaglianze possibili quando i due vettori sono collineari, cioè proporzionali l'uno con l'altro. Questo permette di definire un angolo compreso tra i due vettori, o meglio il coseno dell'angolo tra i due vettori:

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \|y\|}. \quad (4.22)$$

Effettivamente i due vettori x e y , se indipendenti, individuano una varietà lineare a due dimensioni, cioè un piano e la relazione (4.22) definisce una valutazione dell'angolo geometrico tra i due vettori. Con la definizione (4.22) ritroviamo la definizione elementare di prodotto scalare come prodotto dei moduli per il coseno dell'angolo compreso tra i vettori:

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta, \quad (4.23)$$

Nel caso di uno spazio reale l'espressione (4.22) definisce un angolo reale che può essere scelto nell'intervallo $[0, \pi]$, mentre nel caso complesso la relazione (4.22) perde di significato (non determina un “angolo complesso”), anche se la quantità:

$$\frac{|\langle x, y \rangle|}{\|x\| \|y\|},$$

compresa rigorosamente tra 0 e 1 (compresi gli estremi), può servire a definire un angolo (nell'intervallo $[0, \pi/2]$) tra le direzioni dei due vettori.

4.2.1 Ortogonalità

La scelta di un prodotto scalare viene quindi a definire una geometria nello spazio ed in particolare introduce il concetto di ortogonalità tra vettori.

Def. 4.6 Due vettori sono detti **ortogonali** se è nullo il loro prodotto scalare:

$$x \perp y \iff \langle x, y \rangle = 0. \quad (4.24)$$

Osservazione. La definizione ha senso geometricamente quando i vettori sono non nulli, ma la possiamo considerare estesa anche al caso di vettori nulli, notando che il vettore nullo risulta in questo modo ortogonale a qualsiasi vettore dello spazio.

Con il concetto di ortogonalità ritroviamo un fondamento della geometria euclidea, il teorema di **Pitagora**¹. Infatti, se x e y sono non nulli e ortogonali tra loro, è immediato verificare che:

$$\langle x, y \rangle = 0, \quad \Rightarrow \quad \|x - y\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2,$$

dove interpretiamo geometricamente la norma $\|x - y\|$, o $\|x + y\|$, come l'ipotenusa del triangolo rettangolo di cateti $\|x\|$ e $\|y\|$ (Figura 4.1).

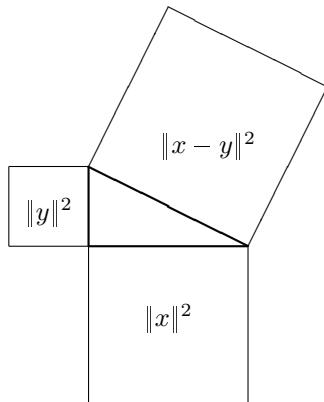


Figura 4.1: Teorema di Pitagora.

Il concetto di ortogonalità si può estendere anche a sottoinsiemi formati da più vettori.

¹ **Pitagora** (Samo, c. 575 a.C. – Metaponto, c. 495 a.C.) è stato un matematico, legislatore e filosofo greco antico secondo quanto tramandato dalla tradizione. La figura storica di Pitagora, messa in discussione da diversi studiosi, si mescola alla leggenda narrata nelle numerose Vite di Pitagora, composte nel periodo del tardo neoplatonismo e del neopitagorismo dove il filosofo viene presentato come figlio del dio Apollo, autore di miracoli e profeta, guaritore e mago. La storia di Pitagora è avvolta nel mistero, di lui sappiamo pochissimo e la maggior parte delle testimonianze che lo riguardano sono di epoca più tarda. Alcuni autori antichi o suoi contemporanei come Senofane, Eraclito ed Erodoto ci danno testimonianze tali da far pensare alla effettiva esistenza storica di Pitagora pur se inserita nella tradizione leggendaria. Quasi sicuramente Pitagora non lasciò nulla di scritto e quindi le opere attribuitegli vanno ascritte piuttosto ad autori sconosciuti che le scrissero in epoca cristiana o di poco antecedente. Il teorema per cui il filosofo è famoso era già noto agli antichi Babilonesi ma alcune testimonianze riferiscono che Pitagora ne avrebbe intuito la validità, mentre si deve a lui avere indicato come sostanza primigenia l'armonia determinata dal rapporto tra i numeri e gli accordi musicali.



Def. 4.7 Sia F un sottoinsieme di uno spazio di Hilbert \mathcal{H} . Definiamo il **complemento ortogonale** F^\perp come l'insieme dei vettori ortogonali a tutti gli elementi di F :

$$F^\perp = \left\{ y \in \mathcal{H} ; \langle y, x \rangle = 0 \forall x \in F \right\} = \bigcap_{x \in F} \left\{ y \in \mathcal{H} ; \langle y, x \rangle = 0 \right\}. \quad (4.25)$$

Ricordando che la varietà lineare $\mathcal{L}(F)$ generata da un insieme di vettori F è costituita dalle combinazioni lineari finite di elementi di F , abbiamo:

Lem. 4.5 Se F è un sottoinsieme di uno spazio di Hilbert \mathcal{H} , l'insieme F^\perp costituisce un sottospazio vettoriale chiuso e:

$$F^\perp = (F^-)^\perp = \mathcal{L}(F)^\perp = (\mathcal{L}(F)^-)^\perp. \quad (4.26)$$

Dim. 4.5 Sia $x \in \mathcal{F}$ fissato, allora l'insieme $\{y \in \mathcal{H} ; \langle y, x \rangle = 0\}$, formato dai vettori ortogonali a x forma un sottospazio (conseguenza immediata delle proprietà di linearità e antilinearità del prodotto scalare) chiuso di \mathcal{H} . La chiusura è conseguenza della continuità del prodotto scalare, in particolare della continuità della applicazione $y \rightarrow \langle y, x \rangle$, in quanto tale insieme risulta la retroimmagine, mediante una applicazione continua dell'insieme chiuso $\{0\} \subset \mathbb{K}$. La seconda parte della (4.25) nella definizione di complemento ortogonale comporta che F^\perp è chiuso in quanto intersezione di insiemi chiusi.

In generale, se $A \subseteq B$ è chiaro che $B^\perp \subseteq A^\perp$ per cui $(\mathcal{L}(F))^\perp \subseteq F^\perp$. D'altra parte se un vettore y è ortogonale ai vettori di F , per la linearità del prodotto scalare, risulta ortogonale anche a qualsiasi combinazione lineare finita di elementi di F :

$$\left\langle y, \sum_{j=1}^N \alpha_j x_j \right\rangle = \sum_{j=1}^N \alpha_j \langle y, x_j \rangle = 0,$$

per cui $F^\perp \subseteq \mathcal{L}(F)^\perp$ e $F^\perp = \mathcal{L}(F)^\perp$.

Vediamo ora le chiusure dei vari insiemi. Chiaramente $F \subseteq F^-$ per cui $(F^-)^\perp \subseteq F^\perp$. Sia ora y ortogonale ad ogni elemento di F e sia $x_0 \in F^-$. Allora esiste una successione x_n di elementi di F convergente a x_0 . Per la continuità del prodotto scalare:

$$0 = \langle y, x_n \rangle \xrightarrow{n \rightarrow \infty} \langle y, x_0 \rangle,$$

per cui y è ortogonale a x_0 , cioè $F^\perp \subseteq (F^-)^\perp$, e valgono le uguaglianze anche nella (4.26).

— ■ —

4.2.2 Decomposizione in sottospazi ortogonali

Una proprietà utile negli spazi vettoriali a dimensioni finite è costituita dal fatto che se M è un sottospazio di uno spazio vettoriale E , allora si può considerare il sottospazio

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

Teo. 4.6 *Sia \mathcal{H} uno spazio di Hilbert e M un suo sottospazio chiuso. Allora vale:*

$$\mathcal{H} = M \oplus M^\perp. \quad (4.27)$$

cioè *ogni vettore di \mathcal{H} è decomposto in maniera univoca come somma di un vettore di M e un vettore in M^\perp .*

Dim. 4.6 Notiamo subito che se esiste la decomposizione questa è unica, in quanto:

$$M \cap M^\perp = \{0\}.$$

Sia ora $x \in \mathcal{H}$. Se $x \in M$ allora $x = x + 0$ fornisce la decomposizione richiesta ($0 \in M^\perp$). Assumiamo quindi che $x \notin M$, e definiamo:

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

Vogliamo mostrare che esiste un elemento $y_M \in M$ in cui si raggiunge il valore $d(x, M)$, cioè tale che:

$$d(x, M) = \|y_M - x\|, \quad y_M \in M.$$

Per definizione stessa di estremo inferiore possiamo determinare una successione y_n in M tale che si approssimi sempre più l'estremo inferiore, cioè:

$$d(x, M) = \lim_{n \rightarrow \infty} \|y_n - x\|.$$

Allora, usando la definizione di norma derivante dal prodotto scalare, è facile verificare che:

$$\begin{aligned} & \| (y_n - x) - (y_m - x) \|^2 + \| (y_n - x) + (y_m - x) \|^2 \\ &= 2 \|y_n - x\|^2 + 2 \|y_m - x\|^2 \\ \|y_n - y_m\|^2 &= 2 \|y_n - x\|^2 + 2 \|y_m - x\|^2 - 4 \|x - \frac{y_n + y_m}{2}\|^2 \\ &\leq 2 \|y_n - x\|^2 + 2 \|y_m - x\|^2 - 4 d(x, M)^2 \xrightarrow[n, m \rightarrow \infty]{} 0. \end{aligned}$$

La successione minimizzante y_n è quindi di Cauchy, per cui (\mathcal{H} è di Banach) convergente. Essendo M chiuso il limite deve appartenere a M stesso:

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

Riusciamo quindi a dimostrare il teorema se $y_\perp = x - y_M$ è ortogonale a M . Consideriamo un qualsiasi elemento $y \in M$:

$$\begin{aligned} \|y_\perp + \lambda y\|^2 &= \|x - (y_M - \lambda y)\|^2 \geq d(x, M)^2 = \|y_\perp\|^2, \\ |\lambda|^2 \|y\|^2 + 2\Re \{\lambda \langle y_\perp, y \rangle\} &\geq 0, \quad \forall \lambda \in \mathbb{K}. \end{aligned}$$

Scegliendo λ dapprima reale, poi immaginario puro, ci si rende conto che ciò è possibile solo se $\langle y_\perp, y \rangle = 0$, per cui $y_\perp \in M^\perp$ e:

$$x = y_M + y_\perp, \quad y_M \in M, \quad y_\perp \in M^\perp.$$

Notiamo che non dobbiamo supporre che il complemento ortogonale M^\perp del sottospazio M sia a sua volta chiuso, in quanto questo è sempre verificato, indipendentemente dalla chiusura di M .

— ■ —

Dato un sottospazio chiuso M di uno spazio di Hilbert \mathcal{H} , per ogni $x \in \mathcal{H}$ risulta quindi univocamente determinato un vettore $y_M \in M$ con la proprietà di avere distanza minima dal vettore x , e tale che x si può decomporre in componenti ortogonali tra loro:

$$x = y_M + y_\perp \quad y_M \in M \quad \langle y_M, y_\perp \rangle = 0. \quad (4.28)$$

La componente y_M viene detta **proiezione ortogonale** di x sul sottospazio M (analogamente y_\perp è la proiezione ortogonale sul complemento ortogonale M^\perp) e definisce un operatore lineare P_M :

$$P_M x = y_M, \quad (4.29)$$

detto **proiettore ortogonale** su M .

Osservazione. Durante la dimostrazione del teorema abbiamo usato la relazione valida in uno spazio di Hilbert \mathcal{H} :

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2, \quad \forall a, b \in \mathcal{H}, \quad (4.30)$$

che esprime una proprietà geometrica dei parallelogrammi (la somma dei quadrati delle diagonali eguaglia la somma dei quadrati dei lati). Tale proprietà, derivabile direttamente dalla definizione di norma tramite un prodotto scalare, non è una proprietà di tutte le norme in generale, e quindi non vale in generale per uno spazio normato, ma è una caratteristica propria degli spazi di Hilbert. In effetti, se abbiamo uno spazio di Banach (pertanto normato e completo) la cui norma soddisfa la **regola del parallelogramma** (4.30), questo diviene uno spazio di Hilbert introducendo il prodotto scalare tramite

l'identità di polarizzazione nel caso di uno spazio complesso (la norma $\|x\|^2$ definisce $\langle x \bar{x} \rangle$):

$$\langle x, y \rangle = \frac{1}{4} \{ \|x + y\|^2 - \|x - y\|^2 - i \|x + iy\|^2 + i \|x - iy\|^2 \}, \quad (4.31)$$

oppure, nel caso di uno spazio reale, tramite la relazione:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2). \quad (4.32)$$

Si lascia al lettore il compito di mostrare che in questo modo si definisce in effetti un prodotto scalare, quando si assume come ipotesi la validità della regola del parallelogramma.

Teo. 4.7 Sia F un sottoinsieme di uno spazio di Hilbert \mathcal{H} , allora:

$$(F^\perp)^\perp = \mathcal{L}(F)^-. \quad \boxed{(F^\perp)^\perp = \mathcal{L}(F)^-}$$

Dim. 4.7 Chiaramente si ha:

$$F \subseteq (F^\perp)^\perp,$$

pertanto $(F^\perp)^\perp$ è un sottospazio chiuso che contiene anche tutte le combinazioni lineari di elementi di F , ed essendo chiuso, contiene anche la loro chiusura:

$$\mathcal{L}(F)^- \subseteq (F^\perp)^\perp.$$

Consideriamo $(F^\perp)^\perp$ come spazio di Hilbert (tutte le proprietà sono verificate, compresa la completezza, conseguenza del fatto che è chiuso). Allora, per il teorema precedente può essere decomposto:

$$(F^\perp)^\perp = (\mathcal{L}(F))^- \oplus K,$$

con K complemento ortogonale di $(\mathcal{L}(F))^-$. Se $k \in K$ allora k è ortogonale a $\mathcal{L}(F)^-$ e a F stesso, $k \in F^\perp$. D'altra parte $k \in (F^\perp)^\perp$ per cui k deve essere ortogonale anche a sé stesso:

$$\langle k, k \rangle = 0, \implies k = 0.$$

K risulta quindi il sottospazio nullo e:

$$(F^\perp)^\perp = (\mathcal{L}(F))^-.$$

— ■ —

Come conseguenza immediata dei risultati precedenti abbiamo il seguente corollario.

Cor. 4.8 Sia M un sottospazio chiuso di uno spazio di Hilbert \mathcal{H} , allora

$$M = (M^\perp)^\perp.$$

4.2.3 Duale topologico

Sia \mathcal{H} uno spazio di Hilbert (che assumiamo in generale complesso), $x \in \mathcal{H}$ fissato, e consideriamo l'applicazione:

$$\begin{aligned} \mathcal{H} &\longrightarrow \mathbb{C}, \\ y &\longrightarrow \langle x, y \rangle, \end{aligned}$$

questa è lineare (per la linearità del prodotto scalare), e continua, cioè limitata, per la disuguaglianza di Schwarz:

$$|\langle x, y \rangle| \leq \|x\| \|y\| = C \|y\|,$$

con $C = \|x\|$, indipendente da y . Pertanto per ogni elemento x risulta definito un elemento appartenente a $\mathcal{L}_c(\mathcal{H}, \mathbb{C}) = \mathcal{H}^*$, detto **duale** dello spazio \mathcal{H} , o, più precisamente **duale topologico**, per distinguerlo dal duale algebrico visto in precedenza, in cui non era richiesta la continuità. In generale per spazio duale si intende quello topologico (sotinsieme del duale algebrico), in quanto l'assenza della continuità non comporta alcun risultato o applicazione importante.

Sostanzialmente, tramite il prodotto scalare, abbiamo costruito una corrispondenza J tra \mathcal{H} e il suo duale \mathcal{H}^* :

$$x \xrightarrow{J} J_x \in \mathcal{H}^*$$

con J_x funzionale lineare e continuo definito da:

$$J_x(y) = \langle x, y \rangle. \quad (4.33)$$

Essendo un funzionale continuo, cioè limitato, possiamo valutare anche la sua norma (in senso operatoriale):

$$\|J_x\| = \sup_{y \neq 0} \frac{|\langle x, y \rangle|}{\|y\|}.$$

Sappiamo che, per ogni x, y , con $y \neq 0$ (disuguaglianza di Schwarz):

$$\frac{|\langle x, y \rangle|}{\|y\|} \leq \|x\|,$$

per cui:

$$\|J_x\| \leq \|x\|,$$

ma se $x = y$, o più in generale, se x e y sono proporzionali tra loro, vale l'uguaglianza. Questo significa che l'estremo superiore è raggiunto quando $y = x$ e tale limitazione è la migliore possibile al variare di y :

$$\|J_x\| = \|x\|.$$

Quindi $J : \mathcal{H} \rightarrow \mathcal{H}^*$ è una applicazione che preserva la norma, cioè **isometrica**.

Vediamo inoltre che la trasformazione J (come dipendenza da x) è **antilineare** (lineare nel caso di uno spazio di Hilbert reale):

$$\begin{aligned} J_{\alpha_1 x_1 + \alpha_2 x_2}(y) &= \langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \overline{\alpha_1} \langle x_1, y \rangle + \overline{\alpha_2} \langle x_2, y \rangle \\ &= \overline{\alpha_1} J_{x_1}(y) + \overline{\alpha_2} J_{x_2}(y), \quad \alpha_1, \alpha_2 \in \mathbb{C}, \quad x_1, x_2, y \in \mathcal{H}, \\ J_{\alpha_1 x_1 + \alpha_2 x_2} &= \overline{\alpha_1} J_{x_1} + \overline{\alpha_2} J_{x_2}, \end{aligned}$$

e **iniettiva**:

$$\begin{aligned} J_{x_1} = J_{x_2} &\implies \langle x_1, y \rangle = \langle x_2, y \rangle \quad \forall y \in \mathcal{H}, \\ &\implies \langle x_1 - x_2, y \rangle = 0 \quad \forall y \in \mathcal{H}. \end{aligned}$$

Scegliendo in particolare:

$$y = x_1 - x_2 \implies \|x_1 - x_2\|^2 = 0 \implies x_1 = x_2.$$

Ma il risultato più importante è costituito dal fatto che l'applicazione J risulta **suriettiva**.

Teo. 4.9 (Fischer-Riesz) Per ogni funzionale lineare e continuo $f \in \mathcal{H}^*$ esiste un unico elemento $x \in \mathcal{H}$ tale che:

$$f(y) = \langle x, y \rangle, \quad \forall y \in \mathcal{H}. \tag{4.34}$$

Sostanzialmente tale teorema afferma che il prodotto scalare costituisce il più generale funzionale lineare e continuo definito su uno spazio di Hilbert. Questo risultato costituisce appunto il motivo per l'uso delle medesime notazioni per indicare sia i funzionali lineari (e continui) che il prodotto scalare. Inoltre permette di identificare lo spazio duale topologico di uno spazio di Hilbert con lo spazio di Hilbert stesso. Possiamo riformulare il teorema affermando che ogni spazio di Hilbert coincide con il suo duale topologico:

$$\mathcal{H} = \mathcal{H}^*.$$

Dim. 4.9 Se $f = 0$ (funzionale identicamente nullo) basta prendere $x = 0$. Se $f \neq 0$ consideriamo il sottospazio lineare:

$$\mathcal{N} = \{y \in \mathcal{H} ; f(y) = 0\}$$

(si vede facilmente per la linearità di f che \mathcal{N} è un sottospazio) che risulta chiuso essendo la retroimmagine dell'insieme chiuso $\{0\}$ in \mathbb{C} tramite una funzione continua. Essendo f non nullo \mathcal{N} non coincide con \mathcal{H} ed esiste $x_0 \in \mathcal{N}^\perp$ con $f(x_0) \neq 0$. Sia ora $y \in \mathcal{H}$ arbitrario:

$$\begin{aligned} f\left(y - \frac{f(y)x_0}{f(x_0)}\right) &= f(y) - \frac{f(y)}{f(x_0)} f(x_0) = 0 \\ y - \frac{f(y)x_0}{f(x_0)} &\in \mathcal{N}, \\ y &= \frac{f(y)}{f(x_0)} x_0 + \left(y - \frac{f(y)x_0}{f(x_0)}\right) \in \mathcal{N}^\perp \oplus \mathcal{N}. \end{aligned}$$

Solo la componente in \mathcal{N}^\perp contribuisce al prodotto scalare di x_0 con y , per cui:

$$\begin{aligned} \langle x_0, y \rangle &= \frac{f(y)}{f(x_0)} \|x_0\|^2, \\ f(y) &= \left\langle \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0, y \right\rangle, \\ x &= \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0, \end{aligned}$$

L'unicità deriva dalla iniettività della corrispondenza J discussa in precedenza.

Questo teorema, noto anche come **teorema di rappresentazione per uno spazio di Hilbert**, è opera dei matematici Fischer² e Riesz³.

— ■ —

² **Ernst Sigismund Fischer** (Vienna, 12 Luglio 1875 – Colonia, 14 Novembre, 1954) lavorò a fianco sia di Mertens che di Minkowski alle Università di Vienna e di Zurigo, rispettivamente. In seguito diventò professore all'Università di Erlangen, dove lavorò con Emmy Noether. La sua principale area di ricerca fu l'analisi matematica, in particolare le serie ortonormali di funzioni che gettarono le basi per l'apparizione del concetto di spazio di Hilbert.



³ **Frigyes Riesz** (Györ, 22 gennaio 1880 – Budapest, 28 febbraio 1956) è stato un matematico ungherese. Fu rettore e professore dell'Università di Szeged. È il fratello maggiore del matematico Marcel Riesz. Riesz sviluppò parte del lavoro fondamentale dello sviluppo dell'analisi funzionale e la sua opera possiede diverse importanti applicazioni in fisica. Il suo lavoro si basò sulle idee introdotte da Fréchet, Lebesgue, Hilbert e altri. Ha fornito pure importanti contributi in altre aree, inclusa la teoria ergodica, ove ha dato una dimostrazione elementare del Teorema ergodico.



4.3 Sistemi ortonormali.

Discutiamo ora alcune nozioni sugli insiemi di vettori ortonormalizzati, cioè normalizzati all'unità e ortogonali tra loro.

Def. 4.8 Un insieme di vettori $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ appartenenti ad uno spazio di Hilbert \mathcal{H} è detto un **sistema ortonormale** se i vettori x_α soddisfano la relazione:

$$\langle x_\alpha, x_\beta \rangle = \begin{cases} 1 & \alpha = \beta, \\ 0 & \alpha \neq \beta, \end{cases} \quad \forall \alpha, \beta \in \mathcal{A}. \quad (4.35)$$

Osservazione. Notiamo che nella definizione l'insieme degli indici \mathcal{A} è arbitrario, e non è detto che sia numerabile, per cui la definizione rimane valida anche per famiglie di vettori dipendenti da un parametro continuo. In ogni caso un sistema ortonormale di vettori risulta linearmente indipendente. Infatti, se consideriamo una combinazione lineare finita e nulla di vettori del sistema:

$$\sum_{j=1}^N c_j x_{\alpha_j} = 0, \quad c_j \in \mathbb{C}, \quad \alpha_j \in \mathcal{A}$$

moltiplicando scalarmente per i vettori x_{α_k} , $k = 1, \dots, N$, otteniamo l'annullarsi dei coefficienti:

$$0 = \langle x_{\alpha_k}, \sum_{j=1}^N c_j x_{\alpha_j} \rangle = \sum_{j=1}^N c_j \langle x_{\alpha_k}, x_{\alpha_j} \rangle = c_k.$$

Un sistema può essere finito o infinito, il caso interessante si ha quando i vettori sono infiniti. Vediamo una prima conseguenza della definizione per un sistema numerabile (possiamo assumere $\mathcal{A} = \mathbb{N}$).

Lem. 4.10 Sia $\{x_j\}$, $j \in \mathbb{N}$, un sistema ortonormale *numerabile* in uno spazio di Hilbert \mathcal{H} , allora la serie (vettoriale):

$$\sum_{j \in \mathbb{N}} c_j x_j$$

è convergente se e solo se:

$$\sum_{j \in \mathbb{N}} |c_j|^2 < \infty,$$

e in questo caso:

$$\left\| \sum_{j \in \mathbb{N}} c_j x_j \right\|^2 = \sum_{j \in \mathbb{N}} |c_j|^2. \quad (4.36)$$

Quando c'è la convergenza la somma della serie $\sum_{j \in \mathbb{N}} c_j x_j$ non dipende dall'ordine dei termini.

Dim. 4.10 Consideriamo la successione delle ridotte:

$$s_N = \sum_{j=1}^N c_j x_j,$$

e consideriamo il criterio di Cauchy. Assumiamo M, N interi con $M > N$:

$$\begin{aligned} \|s_M - s_N\|^2 &= \left\| \sum_{j=1}^M c_j x_j - \sum_{j=0}^N c_j x_j \right\|^2 = \left\| \sum_{j=N+1}^M c_j x_j \right\|^2 \\ &= \left\langle \sum_{k=N+1}^M c_k x_k, \sum_{j=N+1}^M c_j x_j \right\rangle = \sum_{k=N+1}^M \sum_{j=N+1}^M \overline{c_k} c_j \langle x_k, x_j \rangle \\ &= \sum_{k=N+1}^M \overline{c_k} c_k = \sum_{k=N+1}^M |c_k|^2 \\ &= \sum_{j=1}^M |c_j|^2 - \sum_{j=1}^N |c_j|^2, \end{aligned}$$

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

In maniera analoga, sfruttando l'ortogonalità dei vettori x_j (teorema di Pitagora):

$$\left\| \sum_{j \in \mathbb{N}} c_j x_j \right\|^2 = \sum_{j \in \mathbb{N}} |c_j|^2.$$

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

$$x = \sum_{j=1}^{\infty} c_j x_j, \quad y = \sum_{j=1}^{\infty} c_{p(j)} x_{p(j)},$$

con $p(j)$ una permutazione degli indici. Il sistema $\{x_{p(j)}\}$ è ancora un sistema ortonormale in quanto $p : \mathbb{N} \rightarrow \mathbb{N}$ è iniettiva e suriettiva, e:

$$\|x\|^2 = \sum_{j=1}^{\infty} |c_j|^2 = \sum_{j=1}^{\infty} |c_{p(j)}|^2 = \|y\|^2$$

(le serie numeriche assolutamente convergenti non dipendono dall'ordine dei termini). Abbiamo:

$$\|x - y\|^2 = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2,$$

e, sfruttando la continuità del prodotto scalare:

$$\langle x, y \rangle = \lim_{N \rightarrow \infty} \sum_{j=1}^N \bar{c}_j \langle x_j, y \rangle.$$

Ogni indice j risulta la permutazione di un unico indice k^* , $j = p(k^*)$, per cui:

$$\begin{aligned}\langle x_j, y \rangle &= \lim_{M \rightarrow \infty} \sum_{k=1}^M c_{p(k)} \langle x_j, x_{p(k)} \rangle = c_{p(k^*)} = c_j, \\ \langle x, y \rangle &= \langle y, x \rangle = \sum_{j=1}^{\infty} |c_j|^2 = \|x\|^2, \\ \|x - y\|^2 &= 0, \quad x = y.\end{aligned}$$

— ■ —

Nel corso della dimostrazione abbiamo sostanzialmente visto che, con un sistema ortonormale numerabile, posto:

$$x = \sum_{j=1}^{\infty} c_j x_j,$$

si ha:

$$c_j = \langle x_j, x \rangle.$$

Ci chiediamo ora cosa possiamo dire con un sistema ortonormale generale $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ con \mathcal{A} non necessariamente numerabile. L'equazione precedente può essere generalizzata e vale la diseguaglianza di Bessel⁴.

Teo. 4.11 (Diseguaglianza di Bessel) *Sia $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ un sistema ortonormale (non necessariamente numerabile) in uno spazio di Hilbert \mathcal{H} e sia $x \in \mathcal{H}$. Allora i numeri $\langle x_\alpha, x \rangle$ sono non nulli al più per una infinità numerabile di indici α e vale la*

⁴ **Friedrich Wilhelm Bessel** (Minden, 22 luglio 1784 – Königsberg, 17 marzo 1846) è stato un matematico e astronomo tedesco, sistematizzatore delle funzioni di Bessel (le quali, nonostante il nome, furono scoperte da Daniel Bernoulli). Nato in Westfalia, Friedrich Bessel era contemporaneo di Carl Gauss, anch'esso matematico ed astronomo. Bessel è ricordato soprattutto perché fu il primo ad usare la parallasse per misurare la distanza di una stella. È anche dovuto a Bessel il metodo rigoroso di calcolo delle circostanze topocentriche delle eclissi e delle occultazioni. Nonostante non ebbe un'istruzione universitaria, Bessel fu una figura di primaria importanza nell'astronomia dei suoi tempi.



diseguaglianza di Bessel

$$\sum_{\alpha} |\langle x_{\alpha}, x \rangle|^2 \leq \|x\|^2, \quad (4.37)$$

dove la somma è estesa all'infinità numerabile di valori dell'indice α per cui $\langle x_{\alpha}, x \rangle \neq 0$.

Inoltre la serie:

$$\sum_{\alpha \in \mathcal{A}} x_{\alpha} \langle x_{\alpha}, x \rangle \quad (4.38)$$

(sempre con le medesime restrizioni sugli indici della somma) risulta convergente e indipendente dall'ordine dei termini non nulli, e converge alla proiezione ortogonale di x sulla chiusura del sottospazio generato dal sistema $\{x_{\alpha}\}$.

Dim. 4.11 Consideriamo un qualsiasi sottoinsieme finito di indici $\alpha_j \in \mathcal{A}$, $j = 1, \dots, N$, i corrispondenti vettori del sistema ortonormale $\{x_{\alpha_j}\}$ generano un sottospazio finito dimensionale (quindi chiuso) L . Se $x \in \mathcal{H}$ possiamo ricercare la sua proiezione ortogonale su tale sottospazio minimizzando:

$$\begin{aligned} \left\| x - \sum_{j=1}^N c_j x_{\alpha_j} \right\|^2 &= \|x\|^2 + \sum_{j=1}^N |c_j|^2 - 2\operatorname{Re} \sum_{j=1}^N \bar{c}_j \langle x_{\alpha_j}, x \rangle \\ &= \|x\|^2 - \sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2 \\ &\quad + \sum_{j=1}^N \{|\langle x_{\alpha_j}, x \rangle|^2 + |c_j|^2 - 2\operatorname{Re} \bar{c}_j \langle x_{\alpha_j}, x \rangle\} \\ &= \|x\|^2 - \sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2 \\ &\quad + \sum_{j=1}^N |c_j - \langle x_{\alpha_j}, x \rangle|^2 \geq 0. \end{aligned}$$

Da tale espressione risulta chiaro che la minima distanza si raggiunge quando:

$$c_j = \langle x_{\alpha_j}, x \rangle,$$

per cui si annulla l'ultima somma precedente, ed il corrispondente vettore:

$$x_L = \sum_{j=1}^N \langle x_{\alpha_j}, x \rangle x_{\alpha_j}, \quad \|x_L\|^2 = \sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2,$$

costituisce la proiezione ortogonale di x su tale sottospazio. La positività della distanza tra x e la sua proiezione ortogonale (valutabile con il teorema di Pitagora):

$$0 \leq \|x - x_L\|^2 = \|x\|^2 - \|x_L\|^2,$$

comporta la validità della diseguaglianza di Bessel nel caso di un sistema ortonormale finito:

$$\|x\|^2 - \sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2 \geq 0, \quad \sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2 \leq \|x\|^2.$$

Quest'ultima relazione, valida per ogni N comporta che i prodotti scalari $\langle x_\alpha, x \rangle$ siano

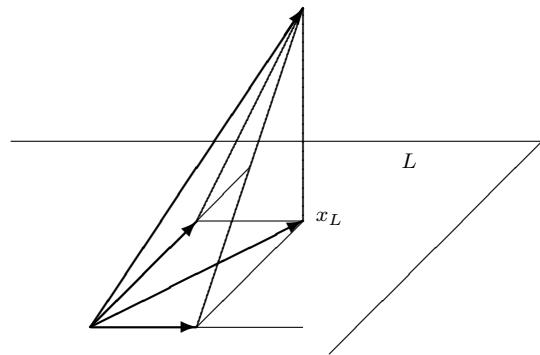


Figura 4.2: Proiezione ortogonale.

non nulli al più per una infinità numerabile di termini. Consideriamo infatti i sottoinsiemi di indici:

$$\begin{aligned} \mathcal{A}_1 &= \{\alpha \in \mathcal{A} ; |\langle x_\alpha, x \rangle| \geq 1\}, \\ \mathcal{A}_2 &= \{\alpha \in \mathcal{A} ; 1 > |\langle x_\alpha, x \rangle| \geq \frac{1}{2}\}, \\ &\vdots \\ \mathcal{A}_n &= \{\alpha \in \mathcal{A} ; \frac{1}{n-1} > |\langle x_\alpha, x \rangle| \geq \frac{1}{n}\}, \\ &\vdots \end{aligned}$$

Ognuno di tali insiemi deve essere finito altrimenti, con infiniti termini a disposizione, potremmo sceglierne un numero N (arbitrario), $\alpha_1, \dots, \alpha_N \in \mathcal{A}_n$, con:

$$\sum_{j=1}^N |\langle x_{\alpha_j}, x \rangle|^2 \geq \sum_{j=1}^N \frac{1}{n^2} = \frac{N}{n^2},$$

grande a piacere, mentre sappiamo che tale somma deve essere limitata da $\|x\|^2$.

Essendo ogni \mathcal{A}_n finito, l'insieme:

$$\mathcal{A}_0 = \{\alpha \in \mathcal{A} ; \langle x_\alpha, x \rangle \neq 0\} = \bigcup_{n=1}^{\infty} \mathcal{A}_n,$$

è numerabile ed essendo valida la disuguaglianza di Bessel per ogni sottoinsieme finito di indici, vale per tutto l'insieme numerabile \mathcal{A}_0 :

$$\sum_{\alpha \in \mathcal{A}_0} |\langle x_\alpha, x \rangle|^2 \leq \|x\|^2,$$

e ha senso considerare la serie:

$$\sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle,$$

che risulta convergente indipendentemente dall'ordine dei termini e converge ad un elemento della chiusura $(\mathcal{L}(\{x_\alpha\}))^-$. Vediamo ora che questo vettore è proprio la proiezione ortogonale su tale sottospazio mostrando che ciò che rimane è ortogonale a esso. Considerando l'identità:

$$x = \left[x - \sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle \right] + \sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle,$$

abbiamo visto che $\langle x_\alpha, x \rangle \neq 0$ solo se $\alpha \in \mathcal{A}_0$ e il vettore:

$$y = x - \sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle$$

risulta ortogonale ad ogni x_β del sistema ortonormale:

$$\left\langle x_\beta, x - \sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle \right\rangle = \langle x_\beta, x \rangle - \lim_{M \rightarrow \infty} \sum_{j=1}^M \langle x_\beta, x_{\alpha_j} \rangle \langle x_{\alpha_j}, x \rangle = 0.$$

Infatti, se $\beta \notin \mathcal{A}_0$, allora:

$$\langle x_\beta, x \rangle = 0, \quad \langle x_\beta, x_{\alpha_j} \rangle = 0,$$

mentre se $\beta \in \mathcal{A}_0$, allora nella somma precedente (per M grande) sopravvive solo l'unico termine con $\alpha_j = \beta$ e si ottiene

$$\langle x_\beta, x \rangle - \langle x_\beta, x \rangle = 0.$$

Abbiamo quindi che il vettore $y \in (\mathcal{L}(\{x_\alpha\}))^\perp = (\mathcal{L}(\{x_\alpha\})^-)^\perp$, cioè:

$$\sum_{\alpha \in \mathcal{A}_0} x_\alpha \langle x_\alpha, x \rangle = P x$$

è la proiezione ortogonale di x su $\mathcal{L}(\{x_\alpha\})^-$.

— ■ —

4.3.1 Basi hilbertiane

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

Def. 4.9 Un sistema ortonormale in uno spazio di Hilbert \mathcal{H} è detto **completo** se è massimale, cioè se non può essere esteso ulteriormente con l'aggiunta di ulteriori vettori. In questo caso esso viene detto anche **base hilbertiana** oppure **base ortonormale**.

L'utilità delle basi hilbertiane è espressa dal seguente teorema che fornisce sostanzialmente delle definizioni equivalenti di base hilbertiana, e trasforma la disuguaglianza di Bessel in una uguaglianza (di Parseval⁵)

Teo. 4.12 Sia $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ un sistema ortonormale in uno spazio di Hilbert \mathcal{H} . Allora le seguenti affermazioni sono equivalenti.

1. $\{x_\alpha\}_{\alpha \in \mathcal{A}}$ è una base ortonormale.
2. Il sistema ortonormale genera tutto \mathcal{H} , nel senso che:

$$(\mathcal{L}(\{x_\alpha, \alpha \in \mathcal{A}\}))^\perp = \mathcal{H}. \quad (4.39)$$

3. Ogni vettore $x \in \mathcal{H}$ può essere decomposto secondo la relazione:

$$x = \sum_{\alpha} \langle x_\alpha, x \rangle x_\alpha, \quad (4.40)$$

dove la somma è estesa ai valori di α (al più numerabili) per cui $\langle x_\alpha, x \rangle \neq 0$.

4. Per ogni $x \in \mathcal{H}$ vale la seguente **relazione di Parseval**:

$$\|x\|^2 = \sum_{\alpha} |\langle x_\alpha, x \rangle|^2, \quad (4.41)$$

dove la somma, come nel caso precedente, risulta estesa al più ad una infinità numerabile di termini non nulli.

5. Per ogni $x, y \in \mathcal{H}$ vale la seguente **relazione di completezza** (con le medesime convenzioni sulla somma):

$$\sum_{\alpha} \langle x, x_\alpha \rangle \langle x_\alpha, y \rangle = \langle x, y \rangle. \quad (4.42)$$

⁵ **Marc-Antoine Parseval des Chênes** (Rosières-aux-Salines, Francia, 27 aprile 1755 – Parigi, 16 agosto 1836) è stato un matematico francese, famoso per il teorema noto come Teorema di Parseval, che anticipa alcune proprietà delle trasformate di Fourier.

Dim. 4.12 Poniamo, per semplificare le notazioni:

$$M = (\mathcal{L}(\{x_\alpha\}))^\perp.$$

(1) \Rightarrow (2). Se il sistema ortonormale è completo non possiamo aggiungere ad esso alcun vettore ortogonale e quindi, decomponendo lo spazio di Hilbert:

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

il complemento ortogonale M^\perp deve essere nullo (in caso contrario potremmo allargare il sistema ortonormale), e $M = \mathcal{H}$.

(2) \Rightarrow (1). Poichè $M = \mathcal{H}$, se il sistema non fosse completo (cioè massimale) potremmo trovare un nuovo vettore ortonormale agli altri, cioè un vettore non nullo ortogonale a M , ma ciò contraddice l'ipotesi.

(2) \Rightarrow (3). La somma $\sum_\alpha \langle x_\alpha, x \rangle x_\alpha$ converge alla proiezione ortogonale di x su M , e, essendo $M = \mathcal{H}$, tale proiezione coincide con x stesso.

(3) \Rightarrow (4). Risulta ovvia per la continuità del prodotto scalare:

$$\|x\|^2 = \langle x, x \rangle = \lim_{N \rightarrow \infty} \sum_{j=1}^N \langle x_{\alpha_j}, x \rangle \langle x, x_{\alpha_j} \rangle = \sum_\alpha |\langle x_\alpha, x \rangle|^2.$$

(4) \Rightarrow (2). Decomponiamo un vettore x secondo M e il suo complemento ortogonale:

$$x = x_M + x_\perp, \quad x_M \in M, \quad x_\perp \in M^\perp,$$

$$x_M = \sum_\alpha \langle x_\alpha, x \rangle x_\alpha,$$

allora, per il teorema di Pitagora:

$$\|x\|^2 = \|x_M\|^2 + \|x_\perp\|^2 = \sum_\alpha |\langle x_\alpha, x \rangle|^2 + \|x_\perp\|^2 = \|x\|^2 + \|x_\perp\|^2,$$

da cui $\|x_\perp\|^2 = 0$, e, per l'arbitrarietà di x , $M^\perp = \{0\}$.

(3) \Rightarrow (5). Se:

$$x = \sum_\alpha \langle x_\alpha, x \rangle x_\alpha, \quad y = \sum_\alpha \langle x_\alpha, y \rangle x_\alpha,$$

allora, per la continuità del prodotto scalare:

$$\langle x, y \rangle = \lim_{N \rightarrow \infty} \sum_{j=1}^N \langle x, x_{\alpha_j} \rangle \langle x_{\alpha_j}, y \rangle.$$

(5) \Rightarrow (3). Basta considerare il caso $x = y$.

Osservazione. Notiamo che nella definizione di base hilbertiana non viene più richiesto che un vettore sia espresso come combinazione lineare finita degli elementi di base. Ciò costituisce una generalizzazione del concetto di base reso possibile dal fatto che uno spazio di Hilbert risulta topologico, e abbiamo garantito la convergenza della serie che rappresenta lo sviluppo.

Osservazione. In uno spazio di Hilbert \mathcal{H} , possiamo verificare che un sottospazio $D \subseteq H$ è denso in \mathcal{H} osservando se il suo complemento ortogonale è nullo:

$$D^\perp = \mathcal{H} \iff D^\perp = \{0\}.$$

Ci potremmo porre ora il problema dell'esistenza di basi hilbertiane. Essendo queste definite come sistemi ortonormali massimali, possiamo ripetere lo stesso tipo di ragionamento (molto astratto e non costruttivo), basato sul lemma di Zorn, che ha condotto all'esistenza di una base per un qualsiasi spazio vettoriale, ottenendo quindi come risposta che ogni spazio di Hilbert ammette una base ortonormale. Notiamo inoltre che se lo spazio di Hilbert ha dimensioni finite, una base ortonormale risulta anche una base ordinaria.

Nel caso di spazi vettoriali a dimensioni finite avevamo anche l'importante risultato che due basi diverse hanno la stessa cardinalità. Questo risultato lo si può estendere anche alle basi hilbertiane di uno spazio di Hilbert, anche a dimensione infinita. Il confronto è effettuato stabilendo una corrispondenza biunivoca (iniettiva e suriettiva) tra le due basi. La costruzione di tale corrispondenza, che può risultare alquanto laboriosa con basi non numerabili, permette di affermare quando due basi hanno la stessa cardinalità.

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

Teo. 4.13 Sia data una successione di vettori $\{x_j\}_{j=1}^\infty$ (quindi un insieme numerabile) in uno spazio di Hilbert \mathcal{H} . Allora esiste un sistema ortonormale numerabile $\{y_k\}_{k=1}^\infty$, tale che:

1. Ogni vettore x_n è combinazione lineare dei vettori y_1, y_2, \dots, y_n , con $n = 1, 2, \dots$
2. Il sottospazio generato dai vettori x_j , coincide col sottospazio generato dai vettori y_k :

$$\mathcal{L}(\{x_j\}) = \mathcal{L}(\{y_k\}). \quad (4.43)$$

Dim. 4.13 Eliminiamo dapprima dalla successione x_1, \dots, x_n, \dots i vettori che dipendono linearmente dai precedenti, per cui resteranno dei vettori (una sottosuccessione della successione originale) $x_{n_1}, \dots, x_{n_k}, \dots$ ($n_k \geq k$) linearmente indipendenti tra loro.

Consideriamo ora la seguente costruzione:

$$\begin{aligned}
 u_1 &= x_{n_1}, & y_1 &= \frac{u_1}{\|u_1\|}, \\
 u_2 &= x_{n_2} - \langle y_1, x_{n_2} \rangle y_1, & y_2 &= \frac{u_2}{\|u_2\|}, \\
 &\vdots & & \\
 u_k &= x_{n_k} - \sum_{j=1}^{k-1} \langle y_j, x_{n_k} \rangle y_j, & y_k &= \frac{u_k}{\|u_k\|}. \\
 &\vdots & &
 \end{aligned}$$

Sostanzialmente, scelto un primo vettore x_{n_1} , questo viene normalizzato fornendo direttamente y_1 , e i due vettori sono proporzionali tra loro. Successivamente, a ogni vettore x_{n_k} viene sottratta la sua proiezione ortogonale nel sottospazio generato dai precedenti vettori y_1, \dots, y_{k-1} , e il vettore residuo u_k risulta ortogonale a tale sottospazio e individua la direzione y_k . Risulta allora evidente che i vettori y_1, y_2, \dots sono ortogonali tra loro ed essendo normalizzati formano un sistema ortonormale. Inoltre, le relazioni precedenti possono essere invertite:

$$x_{n_k} = \|u_k\| y_k + \sum_{j=1}^{k-1} \langle y_j, x_{n_k} \rangle y_j = \sum_{j=1}^k \langle y_j, x_{n_k} \rangle y_j,$$

e il vettore x_{n_k} risulta combinazione lineare di y_1, y_2, \dots, y_k , da cui la validità della prima affermazione (ricordiamo che $n_k \geq k$). L'indipendenza lineare dei vettori x_{n_k} ci garantisce che i vettori residui u_k sono non nulli (e quindi normalizzabili) altrimenti x_{n_k} risulterebbe combinazione lineare dei precedenti vettori y_1, \dots, y_{k-1} , cioè $x_{n_1}, \dots, x_{n_{k-1}}$.

Inoltre il sottospazio generato da x_{n_1}, \dots, x_{n_k} coincide col sottospazio generato da y_1, \dots, y_k , per ogni k , per cui:

$$\mathcal{L}(\{x_n\}) = \mathcal{L}(\{x_{n_k}\}) = \mathcal{L}(\{y_j\}).$$

— ■ —

Osservazione. La selezione dei vettori x_{n_1}, x_{n_2}, \dots linearmente indipendenti tra loro può essere operata in “corso d’opera”. Seguendo l’ordine originario, un vettore x_n risulta dipendente dai precedenti quando il corrispondente vettore residuo u_n si annulla. In questo caso x_n può essere scartato dalla successione e si prosegue col procedimento.

Ricordando la definizione di spazio separabile, abbiamo una immediata conseguenza del procedimento di ortonormalizzazione di Gram⁶-Schmidt⁷.

Teo. 4.14 *Sia \mathcal{H} uno spazio di Hilbert. Allora \mathcal{H} è separabile se e solo se \mathcal{H} ha una base ortonormale numerabile.*

Dim. 4.14 Supponiamo che \mathcal{H} ammetta una base hilbertiana numerabile e_1, \dots, e_n, \dots , e consideriamo il sottoinsieme formato dalle combinazioni lineari finite a coefficienti razionali (intendendo, nel caso di numeri complessi, che sia la parte reale che la parte immaginaria sono razionali). Le combinazioni lineari finite formano ovviamente un insieme denso nello spazio di Hilbert, ed ognuna di esse può essere approssimata con una corrispondente combinazione lineare a coefficienti razionali. Otteniamo quindi un sottoinsieme numerabile e denso in \mathcal{H} .

Viceversa, supponiamo che \mathcal{H} sia separabile, per cui esiste un suo sottoinsieme denso e numerabile $\{x_1, x_2, \dots\}$. Con i vettori x_j , possiamo costruire un sistema ortonormale $\{y_k\}$ che genera il medesimo sottospazio generato dai vettori x_j :

$$\mathcal{L}(\{x_j\}) = \mathcal{L}(\{y_k\}).$$

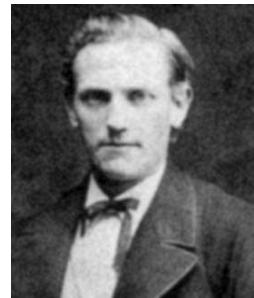
D'altra parte:

$$\mathcal{H} = \{x_j\}^\perp \subseteq (\mathcal{L}(\{x_j\}))^\perp = (\mathcal{L}(\{y_k\}))^\perp$$

per cui i vettori y_k formano un sistema ortonormale completo.

— ■ —

⁶ **Jørgen Pedersen Gram** (Nustrup, 27 Giugno 1850 – Copenhagen, 29 Aprile 1916) è stato un attuario e matematico danese. Tra i suoi articoli importanti si trovano *Sugli sviluppi in serie determinati col metodo dei minimi quadrati* e *Ricerca del numero di primi minori di un dato numero*. Il processo che porta il suo nome, il metodo di Gram-Schmidt, fu pubblicato per la prima volta nell'articolo precedente, nel 1883. La sua carriera di matematico rappresentò sempre un equilibrio tra la matematica pura e applicazioni molto pratiche. Il suo lavoro sulla probabilità e l'analisi numerica lo condussero in maniera naturale a studiare problemi astratti nella teoria dei numeri.



⁷ **Erhard Schmidt** (Dorpat, oggi Tartu, Estonia, 13 Gennaio, 1876 – 6 Dicembre, 1959) è stato un matematico tedesco. Il suo consigliere fu David Hilbert e gli fu assegnato il dottorato dall'Università Georg-August di Göttingen nel 1905. Insieme a Hilbert è ritenuto uno dei fondatori della moderna analisi funzionale astratta.



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

In fisica, e in particolare in meccanica quantistica, si utilizza pesantemente tale condizione, per cui viene richiesta (molte volte solo sottointesa) e si assume che lo spazio degli stati fisici quantistici costituisca uno spazio di Hilbert separabile.

4.3.2 Serie di Fourier.

Dato un sistema ortonormale $\{x_\alpha\}$ in uno spazio di Hilbert \mathcal{H} ed un vettore $x \in \mathcal{H}$, i numeri $\langle x_\alpha, x \rangle$ sono detti **coefficienti di Fourier**⁸ del vettore x . La terminologia deriva dalla cosiddetta **serie di Fourier** o **serie trigonometrica**:

$$\sum_{n=-\infty}^{+\infty} c_n e^{inx}, \quad x \in [-\pi, \pi], \quad (4.44)$$

mediante cui si è soliti esprimere una funzione periodica di periodo 2π .

Vogliamo ora discutere alcune proprietà della serie trigonometrica (4.44), con particolare riguardo al suo comportamento come elemento dello spazio di Hilbert $L^2(-\pi, \pi)$.

Osservazione. La scelta dell'intervallo $[-\pi, \pi]$ non è restrittiva in quanto mediante una trasformazione lineare qualsiasi intervallo $[a, b]$ può essere mappato nell'intervallo $[-\pi, \pi]$ e viceversa. La scelta fatta rende solo più semplici le notazioni in quanto è l'intervallo naturale di definizione delle funzioni trigonometriche.

Notiamo subito che le funzioni:

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad n \in \mathbb{Z}, \quad (4.45)$$

⁸ Jean Baptiste Joseph Fourier (Auxerre, 21 marzo 1768 – Parigi, 16 maggio 1830) è stato un matematico e fisico francese. Partecipò alla Rivoluzione francese, rischiando di essere ghigliottinato durante il Terrore, salvato dalla caduta di Robespierre. Succedette a Laplace nel ruolo di professore alla École Polytechnique nel 1797. Tra i suoi maggiori contributi figurano la teorizzazione della serie di Fourier e la conseguente trasformata di Fourier in matematica, e la formulazione dell'equazione generale della conduzione termica, denominata legge di Fourier, in termodinamica. Il cratere Fourier sulla Luna è intitolato a lui.



formano un sistema ortonormale in $L^2(]-\pi, \pi[)$. Infatti, abbiamo esplicitamente:

$$\begin{aligned} \langle \varphi_n, \varphi_m \rangle &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x} dx \\ &= \begin{cases} \int_{-\pi}^{\pi} \frac{1}{2\pi} dx = 1, & m = n \\ \frac{1}{2\pi i(m-n)} [e^{i(m-n)x}]_{x=-\pi}^{x=\pi} = 0, & m \neq n \end{cases} \end{aligned}$$

Sia ora $f \in L^2(]-\pi, \pi[)$, allora possiamo considerare i coefficienti di Fourier:

$$\langle \varphi_n, f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-inx} f(x) dx,$$

Questo integrale, tuttavia, ha senso anche quando $f \in L^1(]-\pi, \pi[)$. Notiamo in effetti che, essendo l'intervallo $]-\pi, \pi[$ limitato, anche la funzione costante 1 è in $L^2(]-\pi, \pi[)$, per cui $f = 1 \cdot f$ appartiene a $L^1(]-\pi, \pi[)$ e:

$$\|f\|_1 = \|1 \cdot f\|_1 \leq \|1\|_2 \|f\|_2 = \sqrt{2\pi} \|f\|_2,$$

per cui $L^2(]-\pi, \pi[) \subset L^1(]-\pi, \pi[)$.

Se $f \in L^1(]-\pi, \pi[)$ definiamo il **coefficiente di Fourier** di f di ordine n :

$$\tilde{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x) dx, \quad (4.46)$$

ed associamo a f la **serie di Fourier**:

$$\sum_{n=-\infty}^{+\infty} \tilde{f}(n) e^{inx}. \quad (4.47)$$

Osservazione. Ci possiamo porre il problema della convergenza di tale serie. Se $f \in L^2(]-\pi, \pi[)$ l'ortonormalità delle funzioni (4.45) ci garantisce che la serie è sicuramente convergente in $L^2(]-\pi, \pi[)$ e in maniera indipendente dall'ordine dei termini, e si pone la questione se tale serie converga effettivamente alla funzione f che definisce i suoi coefficienti.

In generale, se f non appartiene necessariamente a $L^2(]-\pi, \pi[)$, si pone invece anche la questione di come intendere la convergenza, ad esempio possiamo richiedere la convergenza della successione:

$$(S_n f)(x) = \sum_{k=-n}^n \tilde{f}(k) e^{ikx}, \quad (4.48)$$

oppure separatamente delle due serie:

$$\sum_{k=0}^{\infty} \tilde{f}(k) e^{ikx}, \quad \sum_{k=0}^{\infty} \tilde{f}(-k) e^{-ikx}.$$

Ci sono inoltre diverse possibilità di convergenza: puntualmente (anche solo quasi dappertutto), uniformemente, secondo la norma in L^1 , secondo la norma in L^2 , ed altre ancora. Affronteremo di seguito solo alcuni aspetti di questi problemi senza avere la pretesa di essere esaurienti.

Se $f \in L^1(-\pi, \pi)$, consideriamo la primitiva $F(x)$ definita da:

$$F(x) = \int_{-\pi}^x f(t) dt, \quad x \in [-\pi, \pi].$$

Chiaramente F risulta una funzione continua ed il suo annullarsi identicamente nell'intervallo $[-\pi, \pi]$ comporta l'annullarsi quasi ovunque della funzione f :

$$F(x) = 0 \quad \forall x \in [-\pi, \pi] \implies f(x) = 0 \quad \text{q.d.}$$

Infatti abbiamo che, per ogni $x, y \in [-\pi, \pi]$:

$$\int_x^y f(t) dt = F(y) - F(x),$$

per cui se $F = 0$ identicamente, allora si annulla:

$$\int_B f(t) dt$$

su ogni sottoinsieme (misurabile) B di $[-\pi, \pi]$ ed è noto dalla teoria dell'integrazione che questo implica l'annullarsi quasi ovunque dell'integrando f .

La primitiva $F(x)$, essendo continua e limitata risulta anche sommabile nell'intervallo $[-\pi, \pi]$, per cui possiamo calcolare i suoi coefficienti di Fourier e legarli (quando possibile) a quelli di f . Sia $n \neq 0$, abbiamo:

$$\begin{aligned} \tilde{F}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} F(x) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \int_{-\pi}^x dt e^{-inx} f(t) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dt \int_t^{\pi} dx e^{-inx} f(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt f(t) \frac{i}{n} (e^{-in\pi} - e^{-int}) \\ &= \frac{1}{2\pi i n} \left(\int_{-\pi}^{\pi} e^{-int} f(t) dt + (-1)^{n+1} \int_{-\pi}^{\pi} f(t) dt \right) \\ &= \frac{1}{i n} \left(\tilde{f}(n) + (-1)^{n+1} \tilde{f}(0) \right). \end{aligned}$$

dove abbiamo usato il teorema di Fubini per poter scambiare l'ordine di integrazione. Per quanto riguarda il coefficiente $\tilde{F}(0)$, non abbiamo un legame altrettanto diretto:

$$\begin{aligned}\tilde{F}(0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \int_{-\pi}^x dt f(t) \\ &= \int_{-\pi}^{\pi} dt \int_t^{\pi} dx \frac{1}{2\pi} f(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt f(t) (\pi - t) \\ &= \pi \tilde{f}(0) - \frac{1}{2\pi} \int_{-\pi}^{\pi} t f(t) dt.\end{aligned}$$

Il vantaggio della funzione primitiva F è costituito dalla sua continuità. Operando con funzioni continue otteniamo con relativa facilità alcuni semplici risultati.

Lem. 4.15 *Sia $f : [-\pi, \pi] \rightarrow \mathbb{C}$ una funzione continua. Allora se $f(0) \neq 0$ esiste $n \in \mathbb{Z}$ tale che $\tilde{f}(n) \neq 0$.*

Dim. 4.15 Supponiamo di aver dimostrato il risultato quando f è a valori reali. Poichè sia la parte reale che la parte immaginaria di f sono funzioni continue a valori reali, almeno una delle due è non nulla nell'origine, ed esiste n tale che il corrispondente coefficiente è non nullo. Si vede facilmente che:

$$\begin{aligned}\widetilde{(\operatorname{Re} f)}(n) &= \frac{1}{2} \left(\tilde{f}(n) + \overline{\tilde{f}(-n)} \right) \\ \widetilde{(\operatorname{Im} f)}(n) &= \frac{1}{2i} \left(\tilde{f}(n) - \overline{\tilde{f}(-n)} \right)\end{aligned}$$

e di conseguenza (il determinante della suddetta trasformazione lineare 2×2 è non nullo) almeno un coefficiente tra $\tilde{f}(n)$ e $\tilde{f}(-n)$ deve essere non nullo.

Vediamo ora la validità del risultato nel caso di una funzione f a valori reali. Possiamo supporre $c = f(0) > 0$ (in caso contrario consideriamo $-f$). Essendo f continua esiste un intervallo $[-\delta, \delta]$ dove $f(x) \geq c/2$, e $0 < \delta < \pi$. Poniamo:

$$h(x) = 1 + \cos(x) - \cos(\delta).$$

$h(x)$ è una funzione pari e strettamente decrescente nell'intervallo $[0, \pi]$. In particolare, se $-\pi \leq x \leq \pi$:

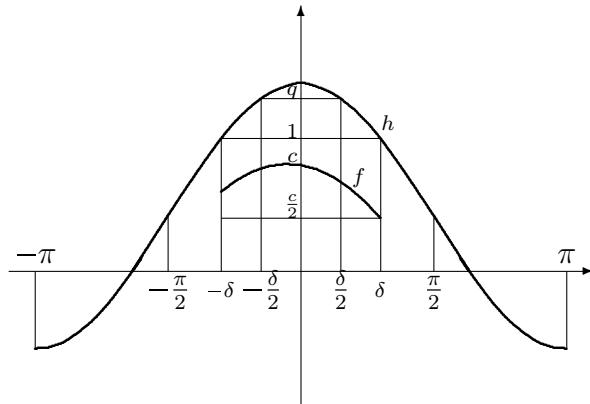
$$h(x) = h(|x|) \geq h(\pi) = -\cos(\delta) \geq -1,$$

al di fuori dell'intervallo $[-\delta, \delta]$:

$$h(x) \leq h(\delta) = 1,$$

mentre all'interno dell'intervallo $[-\delta, \delta]$:

$$h(x) > h(\delta) = 1.$$

Figura 4.3: La funzione $h(x)$.

Poniamo:

$$q = h\left(\frac{\delta}{2}\right) > 1.$$

e sia $m \in \mathbb{N}$ tale che

$$q^m > \frac{4\pi}{c\delta} \sup_{[-\pi, \pi]} |f(x)|.$$

Allora:

$$\begin{aligned} \left| \int_{-\delta \leq |x| \leq \pi} (h(x))^m f(x) dx \right| &\leq \int_{-\delta \leq |x| \leq \pi} |h(x)|^m |f(x)| dx \\ &\leq \sup_{[-\pi, \pi]} |f(x)| \int_{-\delta \leq |x| \leq \pi} dx \leq 2\pi \sup_{[-\pi, \pi]} |f(x)|, \\ \int_{-\pi}^{\pi} (h(x))^m f(x) dx &= \int_{-\delta}^{\delta} (h(x))^m f(x) dx + \int_{\delta \leq |x| \leq \pi} (h(x))^m f(x) dx \\ &\geq \int_{-\delta}^{\delta} (h(x))^m f(x) dx - \left| \int_{\delta \leq |x| \leq \pi} (h(x))^m f(x) dx \right| \\ &\geq \int_{-\delta}^{\delta} (h(x))^m f(x) dx - 2\pi \sup_{[-\pi, \pi]} |f(x)|. \end{aligned}$$

Sull'intervallo $[-\delta, \delta]$ abbiamo $(h(x))^m f(x) \geq 0$, per cui possiamo continuare la catena di disuguaglianze e:

$$\begin{aligned} \int_{-\pi}^{\pi} (h(x))^m f(x) dx &\geq \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} (h(x))^m f(x) dx - 2\pi \sup_{[-\pi, \pi]} |f(x)| \\ &\geq \frac{\delta c}{2} q^m - 2\pi \sup_{[-\pi, \pi]} |f(x)| > 0. \end{aligned}$$

Osserviamo ora che:

$$h(x) = 1 - \cos \delta + \frac{e^{ix}}{2} + \frac{e^{-ix}}{2},$$

è combinazione lineare **finita** delle funzioni $\varphi_n(x)$, e lo stesso vale per $(h(x))^m$:

$$(h(x))^m = \sum_{j=0}^p c_j e^{-in_j x}, \quad c_j \neq 0.$$

Pertanto:

$$0 \neq \frac{1}{2\pi} \int_{-\pi}^{\pi} (h(x))^m f(x) dx = \sum_j c_j \tilde{f}(n_j),$$

e qualche coefficiente $\tilde{f}(n)$ deve essere non nullo.

— ■ —

Notiamo ora che se una funzione f appartiene a $L^1(]-\pi, \pi[)$ la possiamo pensare come la restrizione di una funzione (che indicheremo ancora con f) definita su tutto \mathbb{R} e periodica di periodo 2π . In questo modo possiamo traslare la funzione (spostando quindi l'origine in un qualsiasi punto) di una quantità a :

$$(T_a f)(x) = f(x + a) = f_a(x)$$

e considerandone la restrizione all'intervallo $]-\pi, \pi[$, possiamo calcolare i suoi coefficienti di Fourier:

$$\begin{aligned} \widetilde{(T_a f)}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x + a) dx \\ &= \frac{e^{ina}}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x) dx = e^{ina} \tilde{f}(n). \end{aligned}$$

Con il trucco della traslazione possiamo generalizzare il risultato precedente a qualsiasi punto (la continuità impone però anche condizioni periodiche).

Cor. 4.16 Sia $f : [-\pi, \pi] \rightarrow \mathbb{C}$ una funzione continua che soddisfa la condizione ai bordi $f(-\pi) = f(\pi)$. Allora se $\tilde{f}(n) = 0$ per ogni $n \in \mathbb{Z}$ si ha che $f(x) = 0$ per ogni $x \in [-\pi, \pi]$.

Dim. 4.16 Possiamo prolungare f ad una funzione continua e periodica (di periodo 2π) a tutto l'asse reale, che indicheremo ancora con f . Sia a appartenente all'intervallo $[-\pi, \pi]$ e sia $T_a f$ la funzione traslata (ancora continua e periodica di periodo 2π). Essa ha tutti i coefficienti di Fourier nulli, per cui, in base al lemma precedente, deve necessariamente essere nulla nell'origine:

$$0 = (T_a f)(0) = f(a).$$

Per l'arbitrarietà di a , otteniamo l'annullarsi della funzione originaria f in tutto l'intervallo $[-\pi, \pi]$.

— ■ —

Siamo ora in grado di stabilire un importante risultato in tutto lo spazio $L^1(-\pi, \pi)$.

Teo. 4.17 Sia $f \in L^1(-\pi, \pi)$. Allora, se $\tilde{f}(n) = 0$ per ogni $n \in \mathbb{Z}$, si ha che $f(x) = 0$ q. d.

Dim. 4.17 Lavoriamo con le primitive della funzione f .

$$F(x) = \int_{-\pi}^x f(t) dt, \quad G(x) = F(x) - \tilde{F}(0),$$

Allora $G(x)$ risulta una funzione continua e:

$$\begin{aligned} G(\pi) &= F(\pi) - \tilde{F}(0) = \int_{-\pi}^{\pi} f(t) dt - \tilde{F}(0) \\ &= 2\pi \tilde{f}(0) - \tilde{F}(0) = -\tilde{F}(0) \\ &= G(-\pi), \\ \tilde{G}(0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (F(x) - \tilde{F}(0)) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x) dx - \tilde{F}(0) \\ &= \tilde{F}(0) - \tilde{F}(0) = 0, \\ \tilde{G}(n) &= \tilde{F}(n) = 0, \quad n \neq 0. \end{aligned}$$

Per il corollario precedente, si deve avere:

$$G(x) = 0, \quad \forall x \in [-\pi, \pi],$$

e, in particolare, per $x = -\pi$:

$$0 = G(-\pi) = -\tilde{F}(0).$$

Allora:

$$F(x) = G(x) = 0, \quad \forall x \in [-\pi, \pi],$$

da cui, per quanto detto in precedenza, $f(x) = 0$ quasi ovunque.

— ■ —

I risultati ottenuti non dicono nulla in generale sulla convergenza della serie di Fourier, ma sono basati sulle proprietà dei singoli coefficienti di Fourier. Se ci poniamo in $L^2(-\pi, \pi)$, sappiamo però che la serie di Fourier risulta convergente e l'ultimo teorema ci

fornisce informazioni anche sul limite. In effetti, essendo $L^2(-\pi, \pi) \subset L^1(-\pi, \pi)$, se $f \in L^2(-\pi, \pi)$ ha tutti i coefficienti di Fourier nulli, come vettore risulta $f = 0$, cioè se f è ortogonale a tutte le funzioni del sistema ortonormale $\{\varphi_n\}_{n=-\infty}^{+\infty}$ esso risulta nullo. Ma allora il complemento ortogonale di tale sistema è nullo e il sistema è completo.

Teo. 4.18 Il sistema ortonormale in $L^2(-\pi, \pi)$:

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad n \in \mathbb{Z}, \quad (4.49)$$

forma un sistema ortonormale completo.

Osservazione. Se invece dello spazio di Hilbert $L^2(-\pi, \pi)$ abbiamo lo spazio di Hilbert $L^2(a, b)$ relativo ad un intervallo limitato a, b della retta reale, il sistema ortonormale completo risulta:

$$\varphi_n(x) = \frac{1}{\sqrt{b-a}} e^{i \frac{2\pi}{b-a} n x}. \quad (4.50)$$

e i coefficienti di Fourier di una funzione $f \in L^1(a, b)$ sono espressi da:

$$\tilde{f}(n) = \frac{1}{b-a} \int_a^b e^{-i \frac{2\pi}{b-a} n x} f(x) dx. \quad (4.51)$$

Come abbiamo detto, nulla garantisce la convergenza della serie di Fourier quando, in generale, $f \in L^1(-\pi, \pi)$, e spesso si studia la convergenza della successione (4.48) $S_n f$, in cui si considerano assieme i coefficienti di Fourier $\tilde{f}(n)$ e $\tilde{f}(-n)$, ottenendo la cosiddetta **serie trigonometrica**. Quando tale successione converge in L^1 si può mostrare che converge proprio a f .

Teo. 4.19 Siano $f, g \in L^1(-\pi, \pi)$. Se

$$\|S_n f - g\|_1 \xrightarrow[n \rightarrow \infty]{} 0,$$

allora $f(x) = g(x)$ quasi dappertutto

Osservazione. Questo risultato ci dice sostanzialmente che, se $S_n f$ converge (nella topologia di L^1), il limite è proprio f , ma non ci garantisce nulla sulla convergenza.

Dim. 4.19 Considerando i coefficienti di Fourier del limite g e del termine generico della successione $S_p f$, entrambi appartenenti a $L^1(-\pi, \pi)$, possiamo dire che:

$$\begin{aligned} |\tilde{g}(n) - (\widetilde{S_p f})(n)| &= \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} e^{-inx} (g(x) - (S_p f)(x)) dx \right| \\ &\leq \frac{1}{2\pi} \|g - S_p f\|_1 \xrightarrow[p \rightarrow \infty]{} 0, \end{aligned}$$

cioè si ha una convergenza puntuale dei corrispondenti coefficienti di Fourier:

$$\tilde{g}(n) = \lim_{p \rightarrow \infty} (\widetilde{S_p f})(n).$$

D'altra parte, appena $p \geq |n|$, si ha:

$$\begin{aligned} (\widetilde{S_p f})(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} (S_p f)(x) dx \\ &= \frac{1}{2\pi} \sum_{k=-p}^p \int_{-\pi}^{\pi} \tilde{f}(k) e^{i(k-n)x} dx \\ &= \sum_{k=-p}^p \tilde{f}(k) \delta_{k,n} = \tilde{f}(n), \end{aligned}$$

pertanto per ogni n si ha:

$$\tilde{g}(n) = \tilde{f}(n).$$

La differenza $f - g \in L^1(-\pi, \pi)$ ha quindi tutti i coefficienti di Fourier nulli e, per il teorema 4.17 possiamo dire che $f(x) = g(x)$ q.d.

— ■ —

4.3.3 Polinomi ortogonali.

Una volta costruito un sistema ortonormale completo risulta facile costruirne altri, e di particolare importanza risultano i sistemi ortonormali costruiti mediante polinomi. Nelle applicazioni pratiche si usa spesso sviluppare una funzione in serie di potenze di un qualche parametro, sperando che la serie risulti convergente in qualche senso. Le approssimazioni ottenute troncando la serie sono delle espressioni polinomiali in tale parametro. Lavorando con funzioni appartenenti ad uno spazio di Hilbert risulta allora naturale chiedersi se è possibile costruire una base ortonormale formata da polinomi. Ogni polinomio è combinazione lineare finita di monomi, per cui la tecnica usualmente adottata è quella di considerare i monomi e da questi mediante una procedura di ortogonalizzazione di Gram-Schmidt, ottenere delle funzioni polinomiali con buone proprietà ai fini dell'integrazione.

Poniamoci nello spazio di Hilbert $L^2(a, b)$ con a, b intervallo limitato della retta reale. Abbiamo il seguente risultato.

Teo. 4.20 Il sistema ortonormale costruito, mediante il processo di ortogonalizzazione di Gram-Schmidt, dai monomi $\{1, x, x^2, \dots, x^n, \dots\}$ costituisce un sistema ortonormale **completo** in $L^2(a, b)$, con a e b finiti, $-\infty < a < b < \infty$.

Dim. 4.20 Notiamo innanzitutto che i monomi, intesi come elementi di $L^2([a, b])$, costituiscono un sistema di vettori linearmente indipendenti. Infatti l'annullarsi di una loro combinazione lineare finita equivale ad affermare che un polinomio risulta identicamente nullo:

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

e per il principio di identità dei polinomi, tutti i coefficienti devono essere nulli:

$$a_0 = a_1 = \cdots = a_N = 0.$$

I sottospazi generati dai monomi (l'insieme formato dai polinomi) e dal sistema di polinomi ortonormali generati col procedimento di Gram-Schmidt coincidono per cui è sufficiente mostrare che il loro complemento ortogonale è nullo. Cioè è sufficiente verificare che:

$$u \in L^2([a, b]), \quad \int_a^b x^n u(x) dx = 0 \quad \forall n \quad \implies \quad u(x) = 0 \quad \text{q.d.}$$

Assumiamo quindi l'ortogonalità di $u \in L^2([a, b])$ a qualsiasi monomio e consideriamo le onde piane $\varphi_j(x)$ (4.50) viste in precedenza, che costituiscono un sistema ortonormale completo. Abbiamo:

$$\begin{aligned} \int_a^b \overline{\varphi_j(x)} u(x) dx &= \int_a^b e^{-i \frac{2\pi}{b-a} j x} u(x) \frac{dx}{\sqrt{b-a}} \\ &= \int_a^b \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(-i \frac{2\pi}{b-a} j \right)^k \frac{x^k}{k!} u(x) \frac{dx}{\sqrt{b-a}}. \end{aligned}$$

Vediamo ora che possiamo portare il limite fuori dall'integrale determinando una maggiorazione per l'integrando uniforme in n e sommabile:

$$\begin{aligned} \left| \sum_{k=0}^n \left(-i \frac{2\pi}{b-a} j \right)^k \frac{x^k}{k!} u(x) \right| &\leq \sum_{k=0}^n \left(\frac{2\pi}{b-a} j \right)^k \frac{|x|^k}{k!} |u(x)| \\ &\leq \sum_{k=0}^{\infty} \left| \frac{2\pi}{b-a} j \right|^k \frac{|x|^k}{k!} |u(x)| = e^{\frac{2\pi}{b-a} |j x|} |u(x)|. \end{aligned}$$

La sommabilità è garantita dal fatto che sia $e^{\frac{2\pi}{b-a} |j x|}$ che $|u(x)|$ appartengono a $L^2([a, b])$, per cui, per Hölder, il loro prodotto appartiene a $L^1([a, b])$. La maggiorazione risulta abbastanza brutale, ma permette l'applicabilità del teorema di Lebesgue (scambio del limite con l'integrale), ottenendo:

$$\int_a^b \overline{\varphi_j(x)} u(x) dx = \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(-i \frac{2\pi}{b-a} j \right)^k \int_a^b \frac{x^k}{k!} u(x) \frac{dx}{\sqrt{b-a}} = 0.$$

u risulta quindi ortogonale al sistema di onde piane, completo, per cui, per quanto visto nel paragrafo precedente, $u(x) = 0$ quasi ovunque. Il sottospazio generato dai monomi, e dai

polinomi ortogonali, risulta quindi denso in $L^2(]a, b[)$ e il sistema di polinomi ortogonali risulta completo.

— ■ —

Vediamo ora che, assegnato lo spazio funzionale di Hilbert, il sistema di polinomi ortonormali è sostanzialmente unico. Consideriamo ad esempio lo spazio $L^2(a, b)$ e supponiamo di aver determinato in qualche modo due sistemi ortonormali di polinomi $\{p_n\}$ e $\{q_n\}$, con la proprietà che sia p_n che q_n hanno grado n , allora p_n e q_n coincidono a meno di una costante di proporzionalità di modulo unitario:

$$p_n(x) = \theta_n q_n(x), \quad |\theta_n| = 1. \quad (4.52)$$

Tale relazione può essere vista per induzione. Per $n = 0$ i polinomi sono costanti e la relazione risulta ovvia:

$$\begin{aligned} p_0(x) &= c_1, & q_0(x) &= c_2, \\ \|p_0\| &= |c_1|\sqrt{b-a} = 1, & \|q_0\| &= |c_2|\sqrt{b-a} = 1, \\ \frac{p_0(x)}{q_0(x)} &= \frac{c_1}{c_2}, & \left| \frac{p_0(x)}{q_0(x)} \right| &= \left| \frac{c_1}{c_2} \right| = 1 \end{aligned}$$

Supponiamo quindi la relazione (4.52) valida per $k = 0, 1, \dots, n-1$ e mostriamo che vale per $k = n$. I polinomi $\{p_0, p_1, \dots, p_n\}$ (linearmente indipendenti) generano lo stesso sottospazio generato dai monomi $\{1, x, \dots, x^n\}$, per cui possiamo dire che:

$$p_n(x) = a_n x^n + \sum_{k=0}^{n-1} \alpha_k p_k(x),$$

con a_n non nullo, essendo il grado esattamente n . L'ortonormalità permette di determinare i coefficienti α_k :

$$\begin{aligned} 0 &= \langle p_j, p_n \rangle = a_n \langle p_j, x^n \rangle + \sum_{k=0}^{n-1} \alpha_k \langle p_j, p_k \rangle \\ &= a_n \langle p_j, x^n \rangle + \alpha_j, \quad j = 0, 1, \dots, n-1, \\ \alpha_j &= -a_n \langle p_j, x^n \rangle, \quad j = 0, 1, \dots, n-1, \\ p_n(x) &= a_n \left(x^n - \sum_{k=0}^{n-1} \langle p_k, x^n \rangle p_k(x) \right). \end{aligned}$$

Analogamente il medesimo ragionamento porta a:

$$q_n(x) = b_n \left(x^n - \sum_{k=0}^{n-1} \langle q_k, x^n \rangle q_k(x) \right).$$

con $b_n \neq 0$ coefficiente in q_n del monomio di grado più elevato x^n . Per ipotesi i polinomi p_k e q_k ($k = 0, 1, \dots, n-1$) sono proporzionali tra loro mediante delle costanti θ_k a modulo unitario, per cui:

$$\begin{aligned} p_n(x) &= a_n \left(x^n - \sum_{k=0}^{n-1} \overline{\theta_k} \langle q_k, x^n \rangle \theta_k q_k(x) \right) \\ &= a_n \left(x^n - \sum_{k=0}^{n-1} \langle q_k, x^n \rangle q_k(x) \right) = \frac{a_n}{b_n} q_n(x). \end{aligned}$$

I polinomi p_n e q_n sono quindi proporzionali tra loro e la normalizzazione dei polinomi impone che la costante di proporzionalità sia a modulo unitario:

$$\left| \frac{a_n}{b_n} \right| = \frac{\|p_n\|}{\|q_n\|} = 1.$$

Polinomi di Legendre

Quanto sopra mostra sostanzialmente che non ha importanza il modo in cui si costruiscono i polinomi e questi sono distinti tra loro dal grado. Se consideriamo lo spazio $L^2([-1, 1])$ i corrispondenti polinomi ortogonali sono detti **polinomi di Legendre**⁹, e, a meno di un fattore di normalizzazione, sono di solito espressi dalla relazione:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n [(x^2 - 1)^n]}{dx^n}. \quad (4.53)$$

I primi polinomi di Legendre sono dati da:

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= \frac{3}{2}x^2 - \frac{1}{2}, \\ P_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x, \end{aligned}$$

⁹ **Adrien-Marie Legendre** (Parigi, 18 settembre 1752 – Parigi, 10 gennaio 1833) è stato un matematico francese. Discepolo di Eulero e Lagrange, ha pubblicato un lavoro ormai classico sulla geometria, *Éléments de géométrie*. Ha anche dato significativi contributi nelle equazioni differenziali, nel calcolo, nella teoria funzionale e in teoria dei numeri. Ha poi ridotto gli integrali ellittici in tre forme standard, ma la loro inversione diretta, dovuta ad Abel e a Jacobi ha reso inutile il suo lavoro. Ha anche inventato i polinomi di Legendre nel 1784 mentre studiava l'attrazione degli sferoidi. Nella teoria dei numeri, ha dimostrato l'insolubilità dell'ultimo teorema di Fermat nel caso $n = 5$ e dimostrato l'irrazionalità di π^2 .



$$P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}.$$

Con un pò di algebra si può vedere che:

$$P_n(x) = \frac{1}{2^n n!} \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} \frac{(2j)!}{(2j-n)!} x^{2j-n}, \quad (4.54)$$

dove le potenze negative sono eliminate ponendo, per convenzione, $(-k)! = \infty$.

Nel caso di un intervallo finito $[a, b]$, una possibile espressione per i polinomi ortogonali è data da (sempre a meno di un fattore di normalizzazione):

$$Q_n(x) = \left(\frac{d}{dx} \right)^n [(x-a)(x-b)]^n \quad (4.55)$$

che, a parte una costante moltiplicativa, concidono con i polinomi di Legendre quando $[a, b]$ coincide con $[-1, 1]$. Lasciamo al lettore la verifica (basata sull'integrazione per parti) della ortogonalità dei polinomi:

$$\int_a^b Q_n(x) Q_m(x) dx = 0, \quad n \neq m.$$

Eseguendo il calcolo diretto dell'integrale precedente nel caso $n = m$, si può calcolare una costante di normalizzazione per i polinomi. A meno di un fattore di fase, i polinomi ortonormalizzati risultano:

$$\phi_n(x) = \sqrt{\frac{2n+1}{2}} \frac{1}{2^n n!} \frac{d^n [(x^2 - 1)^n]}{dx^n},$$

nel caso dell'intervallo $[-1, 1]$, mentre, per un intervallo generico $[a, b]$ la normalizzazione comporta:

$$\psi_n(x) = \sqrt{\frac{2n+1}{b-a}} \frac{1}{(b-a)^n n!} \left(\frac{d}{dx} \right)^n [(x-a)(x-b)]^n.$$

Polinomi di Chebyshev

Il concetto di polinomio ortogonale dipende direttamente dalla definizione di prodotto scalare. Ad esempio, possiamo modificare la “misura” dell’integrazione considerando lo spazio di Hilbert $L^2([-1, 1], d\mu)$, con la misura μ definita da:

$$\int_{-1}^1 f(x) d\mu(x) = \int_{-1}^1 f(x) \frac{dx}{\sqrt{1-x^2}}.$$

In tale spazio il prodotto scalare risulta definito da:

$$\langle f, g \rangle = \int_{-1}^1 \overline{f(x)} g(x) \frac{dx}{\sqrt{1-x^2}},$$

e i polinomi ortogonali secondo tale prodotto scalare sono detti **polinomi di Chebyshev**¹⁰. A meno di una costante di normalizzazione questi possono essere espressi dalla relazione:

$$T_n(x) = \cos(n \cos^{-1}(x)) , \quad n = 0, 1, 2, \dots \quad (4.56)$$

Utilizzando la trigonometria è facile convincersi che tale espressione definisce un polinomio. I primi polinomi di Chebyshev risultano:

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \end{aligned}$$

e si può verificare che soddisfano la condizione di ortogonalità:

$$\int_{-1}^1 T_m(x) T_n(x) \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2} \delta_{m,n} = \begin{cases} \frac{\pi}{2} & m = n, \\ 0 & m \neq n. \end{cases} \quad (4.57)$$

I polinomi di Chebyshev sono usati ad esempio nella ricerca di buone interpolazioni polinomiali per una funzione. Nella approssimazione polinomiale di una funzione definita nell'intervallo $[-1, 1]$:

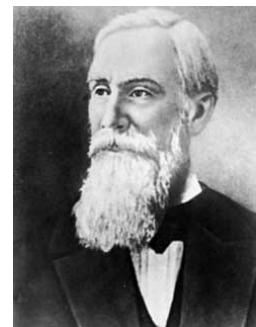
$$f(x) \approx \sum_{m=0}^n a_m T_m(x),$$

i coefficienti a_m possono essere espressi dalla relazione:

$$a_m = \frac{2}{\pi} \int_{-1}^1 T_m(x) f(x) \frac{dx}{\sqrt{1-x^2}} = \frac{2}{\pi} \int_0^\pi f(\cos \theta) \cos(m \theta) d\theta.$$

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

¹⁰ **Pafnutij L'vovič Čebyšëv** (Borovsk, 16 maggio 1821 – San Pietroburgo, 8 dicembre 1894) è stato un matematico e statistico russo. Egli è considerato uno dei padri fondatori della grande scuola matematica russa. Tra i suoi allievi presso l'Università di San Pietroburgo vanno menzionati Dmitrij Grave, Aleksandr Korkin, Aleksandr Ljapunov, Egor Zolotarëv, Andrej Markov padre e Konstantin Posse. I polinomi di Chebyshev gli devono il nome, così come esiste una famiglia di filtri elettronici analogici chiamati filtri di Chebyshev. È altresì noto per i propri lavori nell'ambito della probabilità e della statistica, dove tra l'altro riscoprì, indipendentemente da Bienaym  (di cui però divenne amico), la disuguaglianza di Chebyshev.



Polinomi di Hermite

Possiamo ora chiederci cosa succede se al posto di un intervallo finito $[a, b]$ avessimo avuto tutto l'asse reale. In $L^2(\mathbb{R})$ i polinomi non sono sommabili (e neppure quadrato-sommabili), ma possono contribuire alla costruzione di un sistema ortonormale completo. Consideriamo le funzioni:

$$\varphi_n(x) = x^n e^{-\frac{x^2}{2}}, \quad n = 0, 1, 2 \dots$$

che, grazie al fattore esponenziale, definiscono degli elementi di $L^2(\mathbb{R})$. Queste funzioni, una volta ortonormalizzate, vengono a definire un sistema ortonormale completo in $L^2(\mathbb{R})$.

Il procedimento di ortonormalizzazione di Gram-Schmidt coinvolge combinazioni lineari di φ_n e poichè il fattore gaussiano compare in tutti i termini, è possibile fattorizzarlo ottenendo dei polinomi, detti **polinomi di Hermite**¹¹, moltiplicati per il fattore gaussiano. In meccanica quantistica si ritrovano tali funzioni nella discussione dell'oscillatore armonico quantistico.

Teo. 4.21 *Il sistema ortonormale costruito ortonormalizzando le funzioni (linearmente indipendenti):*

$$\varphi_n(x) = x^n e^{-\frac{x^2}{2}}, \quad n = 0, 1, 2 \dots$$

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

La dimostrazione di tale teorema (che omettiamo) è completamente analoga al caso di $L^2([a, b])$ visto in precedenza. L'unica differenza consiste nel fatto che al posto delle onde piane “discrete”:

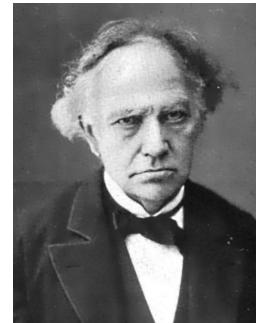
$$e^{i \frac{2\pi}{b-a} nx}, \quad n \in \mathbb{Z},$$

definite nell'intervallo $[a, b]$, occorre considerare la versione “continua”:

$$e^{ikx}, \quad k \in \mathbb{R},$$

definite in tutto l'asse reale (non sono però funzioni sommabili o quadrato-sommabili su tutto \mathbb{R}). Tutto è basato su un risultato delle trasformate di Fourier che vedremo più

¹¹ **Charles Hermite** (Dieuze, 24 dicembre 1822 – Parigi, 14 gennaio 1901) è stato un matematico francese che diede rilevanti contributi a campi quali teoria dei numeri, forme quadratiche, teoria degli invarianti, polinomi ortogonali, funzioni ellittiche e algebra. Egli fu il primo a dimostrare che la costante e , base dei logaritmi naturali, è un numero trascendente. I suoi metodi furono usati successivamente da Ferdinand von Lindemann per dimostrare il suo celebre teorema secondo il quale π è trascendente.



avanti:

$$f \in L^1(\mathbb{R}), \quad \int_{\mathbb{R}} e^{-ikx} f(x) dx = 0 \quad \forall k \in \mathbb{R} \quad \implies \quad f(x) = 0 \quad \text{q.d.}$$

e che costituisce l'analogo del teorema 4.17 visto in precedenza. Notiamo che, se $u \in L^2(\mathbb{R})$, allora, essendo $\exp(-\frac{x^2}{2}) \in L^2(\mathbb{R})$, il prodotto $u(x) \exp(-\frac{x^2}{2})$ appartiene a $L^1(\mathbb{R})$.

Il sistema di funzioni ortogonali in $L^2(\mathbb{R})$ viene indicato con:

$$H_n(x) e^{-\frac{x^2}{2}},$$

dove $H_n(x)$ è il polinomio di Hermite:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad (4.58)$$

(si vede facilmente che tale espressione dà luogo ad un polinomio).

I polinomi di grado inferiore risultano:

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

L'ortogonalità dei polinomi di Hermite è espressa dall'equazione:

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x) dx = 2^n n! \sqrt{\pi} \delta_{n,m}. \quad (4.59)$$

da cui è possibile dedurre il fattore di normalizzazione per rendere il sistema ortonormale.

Polinomi di Laguerre

Se invece di considerare tutto l'asse reale, consideriamo una semiretta reale, $L^2([0, \infty[)$, si considera generalmente il sistema ortonormale ottenuto dalle funzioni:

$$x^n e^{-x},$$

definte per $x > 0$. Il procedimento di ortonormalizzazione di Gram-Schmidt porta ai cosiddetti **polinomi di Laguerre**¹² $L_n(x)$ moltiplicati per il fattore esponenziale e^{-x} :

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x}) . \quad (4.60)$$

I primi polinomi sono dati da:

$$L_0(x) = 1 ,$$

$$L_1(x) = 1 - x ,$$

$$L_2(x) = 2 - 4x + x^2 ,$$

$$L_3(x) = 6 - 18x + 9x^2 - x^3 ,$$

$$L_4(x) = 24 - 96x + 72x^2 - 16x^3 + x^4 .$$

Si lascia al lettore, come esercizio, il calcolo della costante di normalizzazione di tali polinomi.

¹² **Edmond Nicolas Laguerre** (Bar-le-Duc, Francia, 9 aprile 1834 – Bar-le-Duc, Francia, 14 agosto 1886), matematico francese, oggi noto soprattutto per la introduzione dei polinomi che portano il suo nome. Laguerre fin dalla giovinezza ha una salute cagionevole che ostacola le sue attività. Sceglie la carriera militare e dal 1854 al 1864 come ufficiale di artiglieria viene distaccato a Mutzig, vicino a Strasburgo, presso una fabbrica di armamenti. In questo periodo riesce a continuare gli studi matematici e nel 1864 riesce a tornare come tutore all'École Polytechnique; qui rimane quasi fino alla morte. I suoi migliori risultati hanno riguardato vari aspetti della geometria e dell'analisi.



4.4 Esercizi

Esercizio 4.1 Sia $\mathcal{C}_0(\mathbb{R})$ l'insieme delle funzioni complesse di variabile reale, continue e che si annullano all'esterno di un intervallo finito. Mostrare che tale insieme forma uno spazio pre-hilbertiano con il prodotto scalare:

$$\langle f, g \rangle = \int_{\mathbb{R}} \overline{f(x)} g(x) dx.$$

— ○ —

Esercizio 4.2 Sia $E = \mathcal{C}^1([a, b])$ lo spazio delle funzioni (a valori complessi) derivabili con continuità nell'intervallo $[a, b]$, e si definisca:

$$\langle f, g \rangle = \int_a^b \overline{f'(x)} g'(x) dx,$$

dove $f'(x)$ indica la funzione derivata di $f(x)$.

- a) L'espressione $\langle \cdot, \cdot \rangle$ definisce un prodotto scalare in E ?
- b) Sia F il sottospazio di E :

$$F = \{f \in \mathcal{C}^1([a, b]) ; f(a) = 0\},$$

l'espressione $\langle \cdot, \cdot \rangle$ definisce un prodotto scalare in F ?

— ○ —

Esercizio 4.3 Mostrare che in uno spazio pre-hilbertiano vale:

$$\|x - y\| = \|x - z\| + \|z - y\|,$$

se e solo se:

$$z = \alpha x + (1 - \alpha) y,$$

con $\alpha \in [0, 1]$.

— ○ —

Esercizio 4.4 Mostrare che uno spazio di Banach complesso con una norma $\|\cdot\|$ che verifica la regola del parallelogramma:

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

diviene uno spazio di Hilbert con il prodotto scalare:

$$\langle x, y \rangle = \frac{1}{4} \{ \|x + y\|^2 - \|x - y\|^2 - i\|x + iy\|^2 + i\|x - iy\|^2 \}.$$

— ○ —

Esercizio 4.5 Si mostri che in uno spazio pre-hilbertiano vale la seguente **identità di Apollonio**¹³:

$$\|x - z\|^2 + \|z - y\|^2 = \frac{1}{2} \|x - y\|^2 + 2 \left\| z - \frac{x + y}{2} \right\|^2.$$

— ○ —

Esercizio 4.6 Trovare $a, b, c \in \mathbb{C}$ che minimizzano il valore dell'integrale:

$$\int_{-1}^{+1} |x^3 - a - b x - c x^2|^2 dx.$$

— ○ —

Esercizio 4.7 Mostrare che la seguente collezione di funzioni:

$$\sqrt{\frac{1}{\pi}}, \quad \sqrt{\frac{2}{\pi}} \cos x, \quad \sqrt{\frac{2}{\pi}} \cos 2x, \quad \sqrt{\frac{2}{\pi}} \cos 3x, \dots$$

è un sistema ortonormale completo in $L^2(0, \pi)$.

— ○ —

Esercizio 4.8 Mostrare che la seguente collezione di funzioni:

$$\sqrt{\frac{2}{\pi}} \sin x, \quad \sqrt{\frac{2}{\pi}} \sin 2x, \quad \sqrt{\frac{2}{\pi}} \sin 3x, \dots$$

è un sistema ortonormale completo in $L^2(0, \pi)$.

— ○ —

¹³ **Apollonio di Perga** (Perga, 262 a.C. – Murtina, 190 a.C.) è stato un matematico e astronomo greco antico, famoso per le sue opere sulle sezioni coniche e l'introduzione, in Astronomia, degli epicicli e deferenti. Fu Apollonio che diede alla ellisse, alla parabola e alla iperbole i nomi con i quali da allora queste curve sono identificate. Apollonio utilizzò le sue conoscenze geometriche anche per una applicazione pratica, la costruzione di una meridiana in cui l'ombra viene valutata su una superficie conica in modo da fornire una accuratezza maggiore delle meridiane con superficie piana. Il carattere innovativo della sua metodologia e della sua terminologia, specialmente nell'area delle sezioni coniche, hanno influenzato molti studiosi dei secoli successivi e tra questi Tolomeo, Pierre de Fermat, Cartesio e Isaac Newton.



4.4.1 Soluzioni

Soluzione 4.1 Notiamo innanzitutto che $\mathcal{C}_0(\mathbb{R})$ forma uno spazio vettoriale (sottospazio dell'insieme più generale delle funzioni complesse di variabile reale). Infatti se abbiamo due funzioni continue $f(x)$ e $g(x)$ che si annullano all'esterno di un intervallo finito anche la loro somma risulta una funzione continua che si annulla all'esterno di un intervallo finito. La moltiplicazione di $f(x)$ per una qualsiasi costante complessa non altera la continuità della funzione e l'annullarsi al di fuori di un intervallo finito.

L'espressione sesquilineare:

$$\langle f, g \rangle = \int_{\mathbb{R}} \overline{f(x)} g(x) dx$$

risulta ben definita in quanto il prodotto di due funzioni continue che sono non nulle solo all'interno di intervalli limitati risulta una funzione continua non nulla solo all'interno di un intervallo finito. L'integrale precedente non è quindi in realtà esteso a tutto l'asse reale, ma solo ad un intervallo finito (dipendente dalla scelta delle funzioni f e g):

$$\langle f, g \rangle = \int_a^b \overline{f(x)} g(x) dx, \quad a, b \in \mathbb{R},$$

e soddisfa tutte le proprietà di un prodotto scalare.

$$\langle f, f \rangle = \int_{\mathbb{R}} |f(x)|^2 dx \geq 0 \quad \forall f \in \mathcal{C}_0(\mathbb{R}),$$

$$\langle f, f \rangle = 0 \implies |f(x)| = 0 \text{ quasi ovunque,}$$

e, essendo f continua, si annulla ovunque: $f(x) = 0$.

$$\langle g, f \rangle = \int_{\mathbb{R}} \overline{g(x)} f(x) dx = \overline{\int_{\mathbb{R}} \overline{f(x)} g(x) dx} = \overline{\langle f, g \rangle},$$

$$\langle f, \alpha g \rangle = \int_{\mathbb{R}} \overline{f(x)} \alpha g(x) dx = \alpha \int_{\mathbb{R}} \overline{f(x)} g(x) dx = \alpha \langle f, g \rangle \quad \alpha \in \mathbb{C},$$

$$\langle f, g_1 + g_2 \rangle = \int_{\mathbb{R}} \overline{f(x)} g_1(x) dx + \int_{\mathbb{R}} \overline{f(x)} g_2(x) dx = \langle f, g_1 \rangle + \langle f, g_2 \rangle.$$

— ○ —

Soluzione 4.2 La forma sesquilineare:

$$\langle f, g \rangle = \int_a^b \overline{f'(x)} g'(x) dx,$$

risulta ben definita in quanto la continuità delle derivate garantisce l'esistenza dell'integrale, ma non è verificata la positività in senso stretto:

$$\langle f, f \rangle = \int_a^b |f'(x)|^2 dx = 0, \quad f \in \mathcal{C}^1([a, b]),$$

$$\implies f'(x) = 0 \implies f(x) = \text{costante},$$

non necessariamente nulla.

Se imponiamo la condizione ulteriore che la funzione si deve annullare all'estremo a ($f \in F$), allora l'unica costante possibile è la costante nulla, per cui $f(x) = 0$ identicamente nell'intervallo $[a, b]$. Tutte le altre proprietà del prodotto scalare sono banalmente verificate come nell'esercizio 4.1 (sostituendo alle funzioni le loro derivate), per cui E non costituisce uno spazio pre-hilbertiano, mentre F è uno spazio pre-hilbertiano.

— ○ —

Soluzione 4.3 Dobbiamo avere:

$$\begin{aligned} \|x - z + z - y\| &= \|x - z\| + \|z - y\|, \\ \|x - z + z - y\|^2 &= \|x - z\|^2 + \|z - y\|^2 + 2\|x - z\|\|z - y\|, \\ \langle x - z + z - y, x - z + z - y \rangle &= \|x - z\|^2 + \|z - y\|^2 + 2\|x - z\|\|z - y\|, \\ \Re \langle x - z, z - y \rangle &= \|x - z\|\|z - y\|, \end{aligned}$$

e, per la diseguaglianza di Schwarz:

$$\|x - z\|\|z - y\| = \Re \langle x - z, z - y \rangle \leq |\langle x - z, z - y \rangle| \leq \|x - z\|\|z - y\|.$$

Pertanto nella diseguaglianza di Schwarz vale l'uguaglianza e i due vettori $x - z, z - y$ sono proporzionali tra loro:

$$z - y = \beta(x - z),$$

Inoltre si deve avere:

$$\begin{aligned} \Re \langle x - z, \beta(x - z) \rangle &= \|x - z\|\|\beta(x - z)\| \\ \Re \beta\|x - z\|^2 &= |\beta|\|x - z\|^2 \\ \Re \beta = |\beta| &\iff \beta = \Re \beta \geq 0, \end{aligned}$$

pertanto, risolvendo rispetto a z la proporzionalità precedente:

$$z = \frac{\beta}{1 + \beta}x + \frac{1}{1 + \beta}y = \alpha x + (1 - \alpha)y,$$

con:

$$0 \leq \alpha = \frac{\beta}{1 + \beta} \leq 1.$$

— ○ —

Soluzione 4.4 Dobbiamo verificare le varie proprietà del prodotto scalare. Ponendo $y = x$ abbiamo:

$$\begin{aligned} \langle x, x \rangle &= \frac{1}{4} \{ \|x + x\|^2 - \|x - x\|^2 - i\|x + ix\|^2 + i\|x - ix\|^2 \} \\ &= \frac{1}{4} \{ 4 - i|1+i|^2 + i|1-i|^2 \} \|x\|^2 \\ &= \|x\|^2, \end{aligned}$$

pertanto:

$$\langle x, x \rangle \geq 0, \quad \langle x, x \rangle = 0 \iff x = 0.$$

Inoltre:

$$\begin{aligned} \langle y, x \rangle &= \frac{1}{4} \{ \|y + x\|^2 - \|y - x\|^2 - i \|y + ix\|^2 + i \|y - ix\|^2 \} \\ &= \frac{1}{4} \{ \|x + y\|^2 - \|x - y\|^2 - i \|i(x - iy)\|^2 + i \| - i(x + iy)\|^2 \} \\ &= \frac{1}{4} \{ \|x + y\|^2 - \|x - y\|^2 - i \|x - iy\|^2 + i \|x + iy\|^2 \} \\ &= \overline{\langle x, y \rangle}, \end{aligned}$$

Verificare la linearità è meno banale. Notiamo che la regola del parallelogramma può essere letta come:

$$\|a\|^2 + \|b\|^2 = \frac{1}{2} (\|a + b\|^2 + \|a - b\|^2),$$

con a, b vettori arbitrari, per cui:

$$\begin{aligned} \langle x, y_1 \rangle + \langle x, y_2 \rangle &= \frac{1}{4} \{ \|x + y_1\|^2 + \|x + y_2\|^2 - \|x - y_1\|^2 - \|x - y_2\|^2 \\ &\quad - i \|x + iy_1\|^2 - i \|x + iy_2\|^2 + i \|x - iy_1\|^2 + i \|x - iy_2\|^2 \} \\ &= \frac{1}{8} \{ \|2x + y_1 + y_2\|^2 + \|y_2 - y_1\|^2 - \|2x - y_1 - y_2\|^2 - \|y_1 - y_2\|^2 \\ &\quad - i \|2x + iy_1 + iy_2\|^2 - i \|iy_2 - iy_1\|^2 \\ &\quad + i \|2x - iy_1 - iy_2\|^2 + i \|iy_1 - iy_2\|^2 \} \\ &= \frac{1}{2} \left\{ \left\| x + \frac{y_1 + y_2}{2} \right\|^2 - \left\| x - \frac{y_1 + y_2}{2} \right\|^2 \right. \\ &\quad \left. - i \left\| x + i \frac{y_1 + y_2}{2} \right\|^2 + i \left\| x - i \frac{y_1 + y_2}{2} \right\|^2 \right\} \\ &= 2 \left\langle x, \frac{y_1 + y_2}{2} \right\rangle. \end{aligned}$$

Ora, ponendo $y_2 = 0$ (notando che si ha banalmente $\langle x, 0 \rangle = 0$), e $y_1 = y$, abbiamo in particolare:

$$\langle x, y \rangle = 2 \left\langle x, \frac{y}{2} \right\rangle,$$

e quindi:

$$\langle x, y_1 \rangle + \langle x, y_2 \rangle = \langle x, y_1 + y_2 \rangle.$$

Per induzione, con m, n interi:

$$\begin{aligned} m \langle x, y \rangle &= \langle x, my \rangle, \\ n \left\langle x, \frac{y}{n} \right\rangle &= \langle x, y \rangle, \\ \left\langle x, \frac{y}{n} \right\rangle &= \frac{1}{n} \langle x, y \rangle, \\ \left\langle x, \frac{m}{n} y \right\rangle &= \frac{m}{n} \langle x, y \rangle. \end{aligned}$$

Pertanto, per ogni α razionale e positivo:

$$\langle x, \alpha y \rangle = \alpha \langle x, y \rangle, \quad \alpha \in \mathbb{Q}^+,$$

e, per la continuità della norma possiamo estendere la sua validità (i razionali sono densi nei reali) a tutti i numeri reali $\alpha > 0$. Notando ora che:

$$\begin{aligned} \langle x, -y \rangle &= \frac{1}{4} \{ \|x - y\|^2 - \|x + y\|^2 - i \|x - iy\|^2 + i \|x + iy\|^2 \} \\ &= -\langle x, y \rangle, \\ \langle x, iy \rangle &= \frac{1}{4} \{ \|x + iy\|^2 - \|x - iy\|^2 - i \|x - y\|^2 + i \|x + y\|^2 \} \\ &= i \frac{1}{4} \{ -i \|x + iy\|^2 + i \|x - iy\|^2 - \|x - y\|^2 + \|x + y\|^2 \} \\ &= i \langle x, y \rangle, \end{aligned}$$

abbiamo la validità per ogni α complesso:

$$\langle x, \alpha y \rangle = \alpha \langle x, y \rangle, \quad \alpha \in \mathbb{C}.$$

Notiamo infine che, se lo spazio normato risulta completo, dalla relazione:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

abbiamo uno spazio di Hilbert (cioè completo).

— ○ —

Soluzione 4.5 L'identità di Apollonio può essere verificata direttamente, considerando che la norma è definita tramite il prodotto scalare:

$$\|x\| = \sqrt{\langle x, x \rangle},$$

oppure, come conseguenza della regola del parallelogramma:

$$\|a + b\|^2 + \|a - b\|^2 = 2 \|a\|^2 + 2 \|b\|^2,$$

considerando i vettori $a = x - z$, $b = z - y$:

$$\begin{aligned}\|x - z\|^2 + \|z - y\|^2 &= \frac{1}{2} \|x - z + z - y\|^2 + \frac{1}{2} \|x - z - z + y\|^2 \\ &= \frac{1}{2} \|x - y\|^2 + 2 \left\| z - \frac{y + z}{2} \right\|^2.\end{aligned}$$

— ○ —

Soluzione 4.6 Le funzioni:

$$\begin{aligned}f_0(x) &= 1, & f_1(x) &= x, \\ f_2(x) &= x^2, & f_3(x) &= x^3,\end{aligned}$$

definiscono degli elementi dello spazio di Hilbert $L^2(-1, 1)$ e l'integrale da minimizzare rappresenta il quadrato dalla distanza tra f_3 e un generico elemento del sottospazio (chiuso) generato da f_0 , f_1 , f_2 . La soluzione $a + b x + c x^2$ rappresenta la proiezione di f_3 su tale sottospazio.

Abbiamo:

$$\begin{aligned}\langle f_0, f_3 \rangle &= \langle f_3, f_0 \rangle = \int_{-1}^{+1} x^3 dx = 0, \\ \langle f_2, f_3 \rangle &= \langle f_3, f_2 \rangle = \int_{-1}^{+1} x^5 dx = 0, \\ \langle f_1, f_0 \rangle &= \langle f_0, f_1 \rangle = \int_{-1}^{+1} x dx = 0, \\ \langle f_1, f_2 \rangle &= \langle f_2, f_1 \rangle = \int_{-1}^{+1} x^3 dx = 0.\end{aligned}$$

Pertanto le funzioni dispari sono ortogonali alle funzioni pari per cui f_3 avrà proiezione nulla sul sottospazio generato da f_0 , f_2 e quindi:

$$a = c = 0.$$

Rimane da trovare la proiezione sul sottospazio unidimensionale generato da f_1 che sarà data da:

$$P f_3 = \frac{\langle f_1, f_3 \rangle}{\|f_1\|} \frac{f_1}{\|f_1\|} = \frac{\langle f_1, f_3 \rangle}{\|f_1\|^2} f_1 = b f_1$$

Abbiamo:

$$\begin{aligned}\langle f_1, f_3 \rangle &= \int_{-1}^{+1} x^4 dx = \frac{2}{5}, \\ \|f_1\|^2 &= \int_{-1}^{+1} x^2 dx = \frac{2}{3}, \\ b &= \frac{3}{5}.\end{aligned}$$

— ○ —

Soluzione 4.7 Le funzioni:

$$c_0(x) = \sqrt{\frac{1}{\pi}},$$

$$c_n(x) = \sqrt{\frac{2}{\pi}} \cos nx, \quad n = 1, 2, \dots$$

formano chiaramente un sistema ortonormale in quanto una verifica diretta mostra che:

$$\int_0^\pi \cos(nx) \cos(mx) dx = \frac{1}{2} \int_0^\pi [\cos(n+m)x + \cos(n-m)x] dx = 0, \quad n \neq m,$$

$$\int_0^\pi dx = \pi, \quad \int_0^\pi \cos^2(nx) dx = \frac{\pi}{2},$$

Per dimostrare la completezza in $L^2(0, \pi)$ delle funzioni c_n possiamo utilizzare il fatto che il sistema di onde piane:

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad k = 0, \pm 1, \pm 2, \dots$$

formano un sistema ortonormale completo in $L^2(-\pi, \pi)$. Infatti se $f \in L^2(0, \pi)$, possiamo definire una funzione sul segmento $[-\pi, \pi]$, ponendo per $x < 0$:

$$f(x) = f(-x),$$

che risulta chiaramente quadrato sommabile su $[-\pi, \pi]$ e sviluppabile in serie di Fourier:

$$f(x) = \sum_{k=-\infty}^{+\infty} a_k e^{ikx},$$

(la convergenza della serie è da intendersi in media, nella topologia di $L^2(-\pi, \pi)$), con:

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-ikx} f(x) dx,$$

Sfruttando la simmetria imposta rispetto all'origine, abbiamo:

$$a_k = \frac{1}{2\pi} \int_0^\pi [e^{-ikx} + e^{+ikx}] f(x) dx = \frac{1}{\pi} \int_0^\pi \cos(kx) f(x) dx = a_{-k}$$

pertanto lo sviluppo di Fourier risulta:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n [e^{+inx} + e^{-inx}] = a_0 + \sum_{n=1}^{\infty} 2 a_n \cos(nx),$$

ed è facile vedere (sfruttando la simmetria rispetto all'origine) che la convergenza in $L^2(-\pi, \pi)$ implica la convergenza della serie suddetta in $L^2(0, \pi)$.

Pertanto una qualsiasi funzione in $L^2(0, \pi)$ ammette lo sviluppo nel sistema delle funzioni ortonormali c_n :

$$f = \sum_{n=0}^{\infty} \langle c_n, f \rangle c_n,$$

che pertanto forma un sistema ortonormale completo.

— ○ —

Soluzione 4.8 Le funzioni:

$$s_n(x) = \sqrt{\frac{2}{\pi}} \sin nx, \quad n = 1, 2, \dots$$

formano chiaramente un sistema ortonormale in quanto una verifica diretta mostra che:

$$\int_0^\pi \sin(nx) \sin(mx) dx = \frac{1}{2} \int_0^\pi [\cos(n-m)x - \cos(n+m)x] dx = 0, \quad n \neq m,$$

$$\int_0^\pi \sin^2(nx) dx = \frac{\pi}{2},$$

Per dimostrare la completezza in $L^2(0, \pi)$ delle funzioni s_n possiamo utilizzare il fatto che il sistema di onde piane:

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad k = 0, \pm 1, \pm 2, \dots$$

formano un sistema ortonormale completo in $L^2(-\pi, \pi)$. Infatti se $f \in L^2(0, \pi)$, possiamo definire una funzione sul segmento $[-\pi, \pi]$, ponendo per $x < 0$:

$$f(x) = -f(-x),$$

che risulta chiaramente quadrato sommabile su $[-\pi, \pi]$ e sviluppabile in serie di Fourier:

$$f(x) = \sum_{k=-\infty}^{+\infty} b_k e^{ikx},$$

(la convergenza della serie è da intendersi in media, nella topologia di $L^2(-\pi, \pi)$), con:

$$b_k = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-ikx} f(x) dx,$$

Sfruttando la simmetria imposta rispetto all'origine, abbiamo:

$$b_0 = \frac{1}{2\pi} \int_{-\pi}^{+\pi} f(x) dx = 0,$$

$$b_k = \frac{1}{2\pi} \int_0^\pi [e^{-ikx} - e^{+ikx}] f(x) dx = -\frac{i}{\pi} \int_0^\pi \sin(kx) f(x) dx = -b_{-k}, \quad k \neq 0,$$

pertanto lo sviluppo di Fourier risulta:

$$f(x) = \sum_{n=1}^{\infty} b_n [e^{inx} - e^{-inx}] = \sum_{n=1}^{\infty} 2i b_n \sin(nx),$$

ed è facile vedere (sfruttando la simmetria rispetto all'origine) che la convergenza in $L^2(-\pi, \pi)$ implica la convergenza della serie suddetta in $L^2(0, \pi)$.

Pertanto una qualsiasi funzione in $L^2(0, \pi)$ ammette lo sviluppo nel sistema delle funzioni ortonormali s_n :

$$f = \sum_{n=1}^{\infty} \langle s_n, f \rangle s_n,$$

che pertanto forma un sistema ortonormale completo.

— ○ —

Capitolo 5.

Operatori

5.1 Applicazioni lineari limitate tra spazi normati.

Abbiamo visto che l'insieme delle applicazioni lineari continue forma uno spazio vettoriale normato. In generale, per le trasformazioni non risulta solo definita una struttura algebrica di natura vettoriale, ma possiamo considerare anche un'altra operazione, quella di composizione. Nel caso di applicazioni lineari la composizione definisce ancora una trasformazione lineare, per cui è lecito porsi la questione sulla continuità o meno della applicazione composta.

Teo. 5.1 Siano X, Y, Z spazi normati, $A \in \mathcal{L}_c(X, Y)$, $B \in \mathcal{L}_c(Y, Z)$. Allora la composizione $BA \in \mathcal{L}_c(X, Z)$, e:

$$\|BA\| \leq \|B\| \|A\|. \quad (5.1)$$

Dim. 5.1 La continuità è ovvia in quanto è noto che la composizione di funzioni continue è ancora una funzione continua. Lo stesso vale per la linearità, inoltre, se $x \in X$:

$$\|BAx\| \leq \|B\| \|Ax\| \leq \|B\| \|A\| \|x\|,$$

da cui:

$$\|BA\| \leq \|B\| \|A\|.$$

— ■ —

Consideriamo ora una semplice conseguenza del teorema del punto fisso in uno spazio di Banach. Sia X uno spazio di Banach e consideriamo l'operatore identità in tale spazio:

$$\mathbb{1} : X \rightarrow X.$$

Chiaramente l'identità $\mathbb{1}$ risulta un operatore lineare, continuo, ($\|\mathbb{1}\| = 1$), iniettivo, suriettivo e quindi invertibile, con inversa continua. Supponiamo ora che un operatore sia “vicino” all'identità nella topologia degli operatori continui, ci possiamo aspettare che abbia anch'esso le medesime proprietà. In effetti sia:

$$T : X \rightarrow X,$$

una applicazione lineare, continua e tale che:

$$\|T\| < 1,$$

allora l'operatore $\mathbb{1} - T$ è abbastanza vicino all'identità da godere delle stesse proprietà di continuità e invertibilità. Infatti l'operatore $\mathbb{1} - T$ risulta iniettivo, biettivo e quindi invertibile.

L'iniettività è conseguenza del teorema del punto fisso e del fatto che T , essendo a norma minore di uno, risulta una contrazione:

$$\|T x - T y\| \leq \|T\| \|x - y\|, \quad \|T\| < 1.$$

In uno spazio di Banach una contrazione ammette un unico punto x tale che:

$$T x = x.$$

Essendo T lineare, l'equazione precedente ammette la soluzione nulla che risulta quindi unica, e l'operatore $\mathbb{1} - T$ risulta iniettivo:

$$(\mathbb{1} - T) x = 0 \implies x = 0.$$

Il fatto che un operatore sia iniettivo non implica in generale la suriettività, ma consideriamo nel nostro caso l'equazione (nell'incognita x):

$$y = (\mathbb{1} - T) x = x - T x,$$

con $y \in X$. Allora, con y fisso, l'applicazione:

$$f(x) = y + T x,$$

risulta una contrazione:

$$\|f(x) - f(x')\| = \|T x - T x'\| \leq \|T\| \|x - x'\|,$$

per cui esiste un unico punto fisso di tale trasformazione, cioè un unico punto x tale che:

$$x = y + T x. \tag{5.2}$$

Essendo y un elemento arbitrario di X , ciò dimostra la suriettività della applicazione $\mathbb{1} - T$, che risulta invertibile. Formalmente, risolvendo l'equazione (5.2) precedente, abbiamo definito l'operatore inverso:

$$x = (\mathbb{1} - T)^{-1} y, \tag{5.3}$$

per ogni $y \in X$. Abbiamo anche:

$$\begin{aligned} \|x\| &= \|y + T x\| \leq \|y\| + \|T x\| \leq \|y\| + \|T\| \|x\|, \\ (1 - \|T\|) \|x\| &\leq \|y\|, \end{aligned} \tag{5.4}$$

e l'applicazione inversa risulta limitata (e quindi continua):

$$\|x\| = \|(\mathbb{1} - T)^{-1} y\| \leq \frac{1}{1 - \|T\|} \|y\|,$$

$$\|(\mathbb{1} - T)^{-1}\| \leq \frac{1}{1 - \|T\|}.$$

Ricordando il procedimento seguito per determinare il punto fisso (vedi pagina 154), partendo da un guess iniziale x_0 arbitrario:

$$x_1 = y + T x_0,$$

$$x_2 = y + T x_1 = y + T y + T^2 x_0,$$

⋮

$$x_n = y + T x_{n-1} = y + T y + T^2 y + \cdots + T^{n-1} y + T^n x_0$$

⋮

Sapendo che $T^n x_0$ converge al punto fisso di T , che è nullo, abbiamo dato un significato operativo di convergenza alla serie operatoriale, detta anche **serie di Neumann**¹:

$$\mathbb{1} + T + T^2 + \cdots + T^n + \cdots,$$

come espressione dell'operatore inverso $(\mathbb{1} - T)^{-1}$, quando $\|T\| < 1$. Questa serie estende il concetto di serie geometrica con ragione reale, nota dagli studi di analisi matematica, ad una serie con ragione operatoriale.

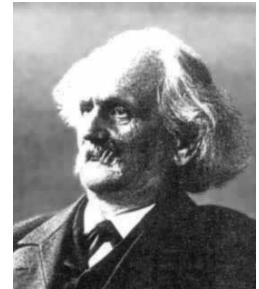
Possiamo riassumere quanto sopra nel seguente risultato:

Teo. 5.2 *Sia X uno spazio di Banach e $T : X \rightarrow X$ un operatore limitato tale che:*

$$\|T\| < 1. \quad (5.5)$$

Allora l'operatore $\mathbb{1} - T$ risulta lineare, continuo, iniettivo, biettivo, quindi invertibile, con inverso continuo. Inoltre, vale la seguente limitazione sulla norma dell'operatore

¹ **Carl Gottfried Neumann** (Königsberg, Prussia, 7 maggio 1832 – 27 marzo 1925) è stato un matematico tedesco. Neumann era figlio del mineralogista, fisico e matematico Franz Ernst Neumann (1798-1895). Neumann lavorò sul principio di Dirichlet e può essere considerato uno degli iniziatori della teoria delle equazioni integrali. Presero il suo nome la serie di Neumann, che rappresenta l'analogo della serie geometrica per matrici infinite, e la condizione al contorno di Neumann per certe equazioni differenziali ordinarie e alle derivate parziali.



inverso:

$$\|(\mathbb{1} - T)^{-1}\| \leq \frac{1}{1 - \|T\|}. \quad (5.6)$$

Esempio 5.1 Nello spazio $L^2(-\pi, \pi)$ consideriamo la seguente equazione integrale:

$$f(x) = g(x) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x - y) f(y) dy,$$

con $g \in L^2(-\pi, \pi)$ assegnata, ed $f \in L^2(-\pi, \pi)$ funzione incognita. Notiamo che, come funzione di y , $\cos(x - y)$ appartiene a $L^2(-\pi, \pi)$, e per la diseguaglianza di Hölder, il prodotto $\cos(x - y) f(y)$, con $f \in L^2(-\pi, \pi)$, è sommabile nell'intervallo $[-\pi, \pi]$. Definendo l'operatore $T : L^2(-\pi, \pi) \rightarrow L^2(-\pi, \pi)$ tramite la relazione:

$$(T f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x - y) f(y) dy,$$

abbiamo una equazione del tipo:

$$(\mathbb{1} - T) f = g.$$

Per la diseguaglianza di Hölder:

$$\begin{aligned} |(T f)(x)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |\cos(x - y) f(y)| dy \\ &\leq \frac{1}{2\pi} \|\cos(x - \cdot)\| \|f\| = \frac{1}{2\sqrt{\pi}} \|f\|, \end{aligned}$$

dove abbiamo usato le notazioni:

$$\|\cos(x - \cdot)\|^2 = \int_{-\pi}^{\pi} \cos^2(x - y) dy = \pi.$$

Pertanto:

$$\begin{aligned} \|T f\|^2 &= \int_{-\pi}^{\pi} |(T f)(x)|^2 dx \leq \frac{1}{4\pi} \|f\|^2 \int_{-\pi}^{\pi} dx = \frac{1}{2} \|f\|^2, \\ \|T f\| &\leq \frac{1}{\sqrt{2}} \|f\| \implies \|T\| \leq \frac{1}{\sqrt{2}} < 1, \end{aligned}$$

per cui T verifica le ipotesi del teorema 5.2 precedente (in realtà si può dimostrare che $\|T\| = \frac{1}{2}$). Tornando ad esaminare l'equazione integrale, possiamo scrivere:

$$\begin{aligned} f(x) &= g(x) + \frac{1}{2\pi} \cos(x) \int_{-\pi}^{\pi} \cos(y) f(y) dy \\ &\quad + \frac{1}{2\pi} \sin(x) \int_{-\pi}^{\pi} \sin(y) f(y) dy \\ &= g(x) + \alpha \cos(x) + \beta \sin(x), \end{aligned}$$

con α e β , dipendenti dalla scelta della funzione f , ma indipendenti dal punto $x \in]-\pi, \pi[$. Sostituendo questa ultima espressione nell'equazione originaria:

$$\begin{aligned} g(x) + \alpha \cos(x) + \beta \sin(x) &= \\ g(x) + \frac{1}{2\pi} \cos(x) \int_{-\pi}^{\pi} \cos(y) [g(y) + \alpha \cos(y) + \beta \sin(y)] dy & \\ + \frac{1}{2\pi} \sin(x) \int_{-\pi}^{\pi} \sin(y) [g(y) + \alpha \cos(y) + \beta \sin(y)] dy, & \end{aligned}$$

e sapendo che:

$$\begin{aligned} \int_{-\pi}^{\pi} \cos^2(y) dy &= \pi, & \int_{-\pi}^{\pi} \sin^2(y) dy &= \pi, \\ \int_{-\pi}^{\pi} \cos(y) \sin(y) dy &= 0, & & \end{aligned}$$

abbiamo:

$$\begin{aligned} \alpha \cos(x) + \beta \sin(x) &= \cos(x) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(y) g(y) dy + \frac{\alpha}{2} \right\} \\ &+ \sin(x) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin(y) g(y) dy + \frac{\beta}{2} \right\}. \end{aligned}$$

Essendo $\cos(x)$ e $\sin(x)$ funzioni linearmente indipendenti in $L^2(-\pi, \pi)$:

$$\begin{cases} \alpha = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(y) g(y) dy + \frac{\alpha}{2} \\ \beta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin(y) g(y) dy + \frac{\beta}{2} \end{cases} \quad \begin{cases} \alpha = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(y) g(y) dy \\ \beta = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(y) g(y) dy \end{cases}$$

ottenendo una espressione delle costanti α e β come dipendenti dalla funzione g . In questo modo possiamo esprimere in maniera univoca la soluzione f come dipendente dal termine noto g . Essendo la funzione g arbitraria in $L^2(-\pi, \pi)$, otteniamo una espressione esplicita dell'operatore inverso $(\mathbb{1} - T)^{-1}$:

$$\begin{aligned} f(x) &= g(x) + \frac{1}{\pi} \cos(x) \int_{-\pi}^{\pi} \cos(y) g(y) dy + \frac{1}{\pi} \sin(x) \int_{-\pi}^{\pi} \sin(y) g(y) dy \\ ((\mathbb{1} - T)^{-1} g)(x) &= g(x) + \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(x - y) g(y) dy. \end{aligned}$$

Notiamo che, accidentalmente:

$$(\mathbb{1} - T)^{-1} = \mathbb{1} + 2T,$$

per cui l'operatore T verifica l'equazione:

$$T^2 = \frac{1}{2} T.$$

Inoltre:

$$\|(\mathbb{1} - T)^{-1}\| = \|\mathbb{1} + 2T\| \leq 1 + 2\|T\| \leq 1 + \sqrt{2},$$

che risulta una stima migliore di:

$$\frac{1}{1 - \|T\|} \leq \frac{1}{1 - \frac{\sqrt{2}}{2}} = 2 + \sqrt{2}.$$

Anche per l'operatore inverso si può dare una valutazione esatta della norma, e in realtà si ha $\|\mathbb{1} + 2T\| = 2$.

— ○ —

5.1.1 Convergenza forte di applicazioni continue.

Supponiamo di avere una successione di operatori fra due spazi normati X e Y , ed assumiamo che per ogni $x \in X$, $A_n x$ converga in Y ad un vettore y , chiaramente dipendente linearmente da x . La sequenza $A_n x$ viene quindi a definire, per $n \rightarrow \infty$, un operatore lineare A fra X e Y e si dice che A_n converge fortemente a A :

Def. 5.1 Una successione di applicazioni lineari $A_n : X \rightarrow Y$ tra due spazi normati X e Y **converge fortemente** se, per ogni $x \in X$, esiste il limite in Y della successione $A_n x$, ed in questo caso il limite definisce una trasformazione lineare A :

$$A x = \lim_{n \rightarrow \infty} A_n x, \quad \forall x \in X. \quad (5.7)$$

Osservazione. Notiamo che tale nozione di convergenza risulta più debole della convergenza in norma. Anche se assumiamo che gli operatori siano continui, nulla garantisce che questi siano convergenti in norma. Viceversa la convergenza in norma di una sequenza di operatori continui implica la convergenza forte:

$$\|A_n - A\| \xrightarrow{n \rightarrow \infty} 0 \implies \|A_n x - A x\| \leq \|A_n - A\| \|x\| \xrightarrow{n \rightarrow \infty} 0, \quad \forall x \in X.$$

Il termine convergenza forte viene riservato generalmente per una convergenza vettoriale, stabilità mediante l'uso dei vettori nella topologia dello spazio di appartenenza di tali vettori.

La questione che ci poniamo ora è proprio quella di stabilire se tale convergenza possa definire una trasformazione A continua, assumendo ovviamente che tutte le trasformazioni lineari A_n siano continue. La risposta sarà affermativa se lavoriamo in spazi di Banach, cioè se X e Y sono spazi completi. Per arrivare a ciò occorrono però delle conoscenze preliminari.

Teo. 5.3 (Baire) Sia X uno spazio metrico completo e $\{F_j\}_{j=1}^{\infty}$ una collezione numerabile di sottoinsiemi chiusi di X tali che:

$$X = \bigcup_{j=1}^{\infty} F_j. \quad (5.8)$$

Allora esiste un chiuso F_{n_0} della famiglia che contiene un aperto.

Notiamo che l'affermazione non è banale perché possiamo avere degli insiemi chiusi talmente "magri" da non riuscire a contenere alcun aperto. Si pensi in \mathbb{R}^2 ad un insieme rappresentato da un segmento rettilineo, compresi gli estremi. Il risultato precedente viene anche detto teorema delle categorie di Baire².

Dim. 5.3 Consideriamo il primo chiuso F_1 . Se $F_1 = X$ allora F_1 risulta anche aperto e siamo a posto. Se $F_1 \neq X$, allora il suo complementare è aperto ed esiste un punto x_1 e una sfera $S(x_1, \delta_1)$ tale che tutta la chiusura della sfera è contenuta nel complementare:

$$S(x_1, \delta_1)^- \cap F_1 = \emptyset.$$

A questo punto si presentano due casi: o $S(x_1, \delta_1) \subset F_2$ e il risultato è provato, oppure esiste una sfera $S(x_2, \delta_2)$ contenuta in $S(x_1, \delta_1)$, la cui chiusura non interseca F_2 :

$$S(x_2, \delta_2) \subset S(x_1, \delta_1), \quad S(x_2, \delta_2)^- \cap F_2 = \emptyset,$$

e possiamo supporre (eventualmente restringendo δ_2) che:

$$\delta_2 \leq \frac{\delta_1}{2}.$$

Se $S(x_2, \delta_2) \subset F_3$, il risultato è dimostrato, altrimenti il procedimento può essere iterato per F_3, F_4, \dots , costruendo una sequenza di sfere con le proprietà:

$$S(x_j, \delta_j) \subset S(x_{j-1}, \delta_{j-1}), \quad \delta_j \leq \frac{\delta_1}{j} \quad S(x_j, \delta_j)^- \cap F_j = \emptyset.$$

Il processo deve avere termine per un certo valore n_0 , altrimenti abbiamo costruito una successione di centri delle sfere x_j che risulta di Cauchy:

$$d(x_n, x_m) < \delta_m \leq \frac{\delta_1}{m}, \quad \forall n > m,$$

² **René-Louis Baire** (Parigi, 21 gennaio 1874 – Chambéry, 5 luglio 1932) è stato un matematico francese, noto per i suoi lavori sulla continuità, sui numeri irrazionali e sul concetto di limite. Baire ha alternato la ricerca universitaria all'insegnamento nei licei, ed ha effettuato solo un numero limitato di pubblicazioni, sebbene di notevole rilevanza. La sua carriera accademica è stata resa difficile anche dalla sue cattive condizioni di salute. Diverse strutture topologiche portano il suo nome. Fra l'altro, si deve a Baire l'introduzione della nozione di continuità per funzioni a valori reali, ottenuta dalla simultaneità di semicontinuità inferiore e superiore.

e quindi, essendo X completo, convergente ad un punto $\tilde{x} \in X$ che deve appartenere alla chiusura di ogni sfera (la successione potrebbe avvicinarsi sempre più al bordo di qualche sfera, per cui abbiamo richiesto la non intersezione fra i vari F_j e le chiusure delle sfere corrispondenti):

$$\tilde{x} \in S(x_m, \delta_m)^-, \quad \forall m,$$

ed ogni sfera chiusa non interseca il corrispondente F_m . Ma allora \tilde{x} non può appartenere ad alcun F_j , e nemmeno alla loro unione, ma ciò è assurdo in quanto gli insiemi F_j formano un ricoprimento di X . Pertanto deve esistere almeno un chiuso F_{n_0} contenente una sfera aperta, cioè un aperto.

— ■ —

Siamo ora in grado di mostrare il seguente risultato noto come principio di uniforme limitatezza.

Teo. 5.4 (Principio di uniforme limitatezza) *Sia $\{A_\lambda ; \lambda \in \Lambda\}$ una famiglia di applicazioni lineari e continue tra due spazi di Banach X e Y , $A_\lambda \in \mathcal{L}_c(X, Y)$, tali che:*

$$\forall x \in X \quad \sup_{\lambda \in \Lambda} \|A_\lambda x\| = M(x) < \infty. \quad (5.9)$$

Allora anche le norme delle applicazioni sono uniformemente limitate:

$$\sup_{\lambda \in \Lambda} \|A_\lambda\| < \infty. \quad (5.10)$$

Osservazione. Di nuovo abbiamo una affermazione non banale, perché dall'uniformità (rispetto a λ) per ogni vettore abbiamo l'uniformità sulle norme operatoriali, indipendenti dal vettore a cui si applicano le trasformazioni.

Dim. 5.4 Consideriamo l'insieme ($\lambda \in \Lambda$ e $n \in \mathbb{N}$ fissi):

$$\{x \in X ; \|A_\lambda x\| \leq n\},$$

e l'applicazione:

$$\Phi(x) = \|A_\lambda x\|.$$

Φ è continua in quanto composizione di A_λ e della norma che sono applicazioni continue. L'insieme di cui sopra, retroimmagine di un chiuso, è quindi chiuso, e chiusa risulta l'intersezione:

$$F_n = \bigcap_{\lambda \in \Lambda} \{x \in X ; \|A_\lambda x\| \leq n\}.$$

Vediamo ora che gli insiemi F_n , $n \in \mathbb{N}$, formano un ricoprimento chiuso e numerabile di X . Infatti, utilizzando l'ipotesi (5.9), sia $x \in X$ arbitrario, se $n > M(x)$ allora $x \in F_n$. Essendo $M(x)$ finito, un tale n esiste, e quindi:

$$X = \bigcup_{n \in \mathbb{N}} F_n.$$

X è di Banach, pertanto metrico e completo, per il risultato precedente esiste un intero n_0 e una sfera (aperta) $S(x_0, \delta_0)$ contenuta in F_{n_0} , cioè:

$$\|x - x_0\| < \delta_0 \implies \|A_\lambda x\| \leq n_0 \quad \forall \lambda \in \Lambda.$$

Consideriamo allora un punto y in un intorno sferico dell'origine:

$$\begin{aligned} \|y\| < \delta_0 &\implies x = x_0 + y \in S(x_0, \delta_0), \\ &\implies \|A_\lambda y\| \leq \|A_\lambda x\| + \|A_\lambda x_0\| \leq 2n_0. \end{aligned}$$

Sia ora z non nullo ed arbitrario in X . Mediante una trasformazione di scala (**omotetia**) opportuna possiamo ricondurci all'intorno sferico dell'origine di raggio δ_0 :

$$\left\| \frac{z}{\|z\|} \frac{\delta_0}{2} \right\| = \frac{\delta_0}{2} < \delta_0,$$

e, per ogni $\lambda \in \Lambda$:

$$\left\| A_\lambda \frac{z}{\|z\|} \frac{\delta_0}{2} \right\| \leq 2n_0, \quad \|A_\lambda z\| \leq \frac{4n_0}{\delta_0} \|z\|, \quad \|A_\lambda\| \leq \frac{4n_0}{\delta_0}.$$

Essendo $4n_0/\delta_0$ una costante indipendente da λ abbiamo l'uniforme limitatezza delle norme operatoriali.

— ■ —

Questo risultato permette di rispondere affermativamente al quesito precedente sulla continuità del limite forte.

Teo. 5.5 *Sia A_n una successione di applicazioni lineari e continue fra due spazi di Banach X, Y : $A_n \in \mathcal{L}_c(X, Y)$. Allora, se per ogni $x \in X$ esiste il limite:*

$$\lim_{n \rightarrow \infty} A_n x \in Y,$$

questo limite definisce una applicazione lineare e continua $A \in \mathcal{L}_c(X, Y)$.

Dim. 5.5 La definizione basata sull'esistenza del limite forte:

$$A x = \lim_{n \rightarrow \infty} A_n x.$$

comporta immediatamente la linearità di A . Infatti, ad esempio:

$$\left. \begin{array}{l} A_n x + A_n y \xrightarrow{n \rightarrow \infty} A x + A y \\ A_n (x + y) \xrightarrow{n \rightarrow \infty} A (x + y) \end{array} \right\} \implies A (x + y) = A x + A y.$$

Per quanto riguarda la continuità sappiamo che se una successione ha limite allora questa è limitata (da un certo punto in poi gli elementi della successione sono vicino al limite, quelli prima sono in numero finito, per cui limitati). Allora $\|A_n x\|$ (essendo convergente) è limitata, cioè:

$$\sup_n \|A_n x\| = M(x) < \infty,$$

e, per il teorema precedente di uniforme limitatezza, esiste una costante C tale che:

$$\|A_n\| \leq C < \infty, \quad \forall n, \quad \|A_n x\| \leq C \|x\|, \quad \forall n.$$

Passando al limite per $n \rightarrow \infty$:

$$\|A_n x\| \xrightarrow{n \rightarrow \infty} \|Ax\|, \quad \|Ax\| \leq C \|x\|,$$

e A è continuo.

— ■ —

La continuità forte non implica necessariamente la convergenza in norma. Vediamolo con un esempio.

Esempio 5.2 Consideriamo come spazio vettoriale completo l'insieme l^p delle successioni p -sommabili:

$$l_p = \left\{ x = (x_1, x_2, \dots, x_n, \dots) ; x_n \in \mathbb{C}, \sum_{j=1}^{\infty} |x_j|^p < \infty \right\}, \quad (5.11)$$

con la norma:

$$\|x\|_p = \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}}. \quad (5.12)$$

In tale spazio ($1 \leq p < \infty$) definiamo, per ogni n , l'operatore di "shift" A_n :

$$A_n : (x_1, x_2, \dots) \longrightarrow (x_{n+1}, x_{n+2}, \dots).$$

e vediamo che la successione A_n converge fortemente ma non in norma.

Analizziamo la norma $\|A_n x\|$:

$$\|A_n x\| = \left(\sum_{j=n+1}^{\infty} |x_j|^p \right)^{\frac{1}{p}}.$$

Questa risulta il resto n -esimo (a parte l'esponente $\frac{1}{p}$) della serie:

$$\sum_{j=1}^{\infty} |x_j|^p,$$

che è convergente per ipotesi. Quindi, per ogni $x \in l_p$ (l'estrazione di radice p -esima è una funzione continua):

$$\sum_{j=n+1}^{\infty} |x_j|^p \xrightarrow{n \rightarrow \infty} 0, \quad A_n x \xrightarrow{n \rightarrow \infty} 0 \implies A_n \xrightarrow{n \rightarrow \infty} 0 \text{ fortemente.}$$

Notiamo inoltre che:

$$\sum_{j=n+1}^{\infty} |x_j|^p \leq \sum_{j=1}^{\infty} |x_j|^p, \quad \|A_n x\| \leq \|x\|, \implies \|A_n\| \leq 1,$$

per cui A_n è continuo (come pure l'operatore limite forte, identicamente nullo). Vediamo maggiormente in dettaglio la norma dell'operatore A_n . Sappiamo già che se $x \neq 0$:

$$\frac{\|A_n x\|}{\|x\|} \leq 1,$$

e vogliamo far vedere che esistono dei vettori x per i quali vale l'uguaglianza (a fissato n). Prendiamo un vettore non nullo x con le prime n componenti nulle. Allora si ha:

$$\|x\|^p = \sum_{j=1}^{\infty} |x_j|^p = \sum_{j=n+1}^{\infty} |x_j|^p,$$

e quindi:

$$\begin{aligned} \|A_n x\| &= \|x\|, \quad x_1 = x_2 = \dots = x_n = 0, \\ \sup_{\|x\|=1} \|A_n x\| &= 1, \quad \|A_n\| = 1. \end{aligned}$$

Sulla base di questo fatto possiamo dire che A_n non può convergere in norma. Infatti se, per assurdo, A_n convergesse in norma ad un operatore B , si avrebbe:

- i) $A_n \rightarrow B$ fortemente.
- ii) $\|A_n\| \rightarrow \|B\|$ per la continuità della norma stessa, in quanto:

$$|\|A_n\| - \|B\|| \leq \|A_n - B\|.$$

Nel nostro caso $A_n \rightarrow 0$ fortemente, per cui dovrebbe essere $B = 0$, che contraddice $\|A_n\| = \|B\| = 1$.

— ○ —

La richiesta che gli spazi siano completi può essere cruciale per le proprietà del limite forte di una successione di applicazioni lineari e continue. Possiamo vedere cosa succede se $A_n x$ è una successione convergente per ogni x appartenente ad un sottospazio D non completo, per il quale possiamo anche assumere che sia denso in uno spazio completo X :

$D^- = X$. L'informazione di convergenza vettoriale dei vettori $A_n x$ risulta troppo povera per la mancanza di completezza dell'insieme di definizione (anche se x è un punto prefissato che non varia con n , per cui non è richiesta alcuna convergenza nel dominio), in quanto viene a mancare l'informazione di uniforme limitatezza sulle norme delle applicazioni. Vediamo con un esempio cosa può succedere.

Esempio 5.3 Consideriamo di nuovo lo spazio l_p come nell'esempio precedente, e sia D l'insieme dei vettori in l_p con solo un numero finito di componenti non nulle (il numero di componenti non nulle può essere arbitrario, ma finito per ogni elemento di D):

$$D = \{x = (x_1, x_2, \dots) \in l_p ; \exists N \in \mathbb{N} \text{ tale che } x_j = 0 \forall j > N\}.$$

e in tale insieme (che risulta ovviamente un sottospazio di l_p , $1 \leq p \leq \infty$) definiamo gli operatori:

$$\begin{aligned} A_n : D &\longrightarrow l_p, \\ (x_1, x_2, \dots, x_k, \dots) &\longrightarrow n(x_{n+1}, x_{n+2}, \dots, x_{k+n}, \dots) \end{aligned}$$

Sostanzialmente abbiamo lo stesso operatore dell'esempio precedente moltiplicato per il fattore n . In modo analogo al caso precedente abbiamo quindi che A_n è continuo e:

$$\|A_n\| = n.$$

Inoltre, se $x \in D$, allora esiste $N \in \mathbb{N}$, dipendente da x , tale che $x_j = 0$ se $j > N$. Allora $A_n x = 0$ se $n \geq N$ e:

$$A_n x \xrightarrow[n \rightarrow \infty]{} 0 \quad \forall x \in D.$$

Quindi A_n converge fortemente all'operatore nullo 0 nello spazio D . Anche se D è denso in l_p , A_n può essere definito anche su tutto l_p nella stessa maniera (la definizione precedente ha senso in tutto lo spazio l_p) e risulta continuo ($\|A_n\| = n$), la successione A_n non può convergere fortemente in tutto lo spazio l_p . Se $A_n x$ convergesse per ogni $x \in l_p$, allora, per il teorema di uniforme limitatezza, $\|A_n\|$ sarebbe limitato, ma ciò non è vero per cui devono esistere dei punti $x \in l_p$ per i quali $A_n x$ non converge (e quindi per tali punti non è definita l'azione dell'operatore limite).

Per inciso possiamo notare che il sottospazio D è denso in l_p in quanto se $x = (x_1, x_2, \dots)$ è un elemento arbitrario di l_p , allora:

$$x^{(N)} = (x_1, x_2, \dots, x_N, 0, \dots, 0, \dots) \in D$$

forma una successione che “approssima” $x \in l_p$. Infatti:

$$\|x - x^{(N)}\| = \left(\sum_{j=N+1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \xrightarrow[N \rightarrow \infty]{} 0,$$

in quanto, se $x \in l_p$, la serie:

$$\sum_{j=1}^{\infty} |x_j|^p < \infty$$

è convergente. D risulta quindi denso in l_p .

— ○ —

5.1.2 Estensione continua di applicazioni continue.

Spesso, dato uno spazio lineare topologico, si riesce a definire una applicazione lineare solo su un suo sottospazio, magari denso. Ci possiamo quindi chiedere se è possibile estendere tale applicazione lineare a tutto lo spazio, e se, in caso affermativo, tale estensione è unica, lineare e continua.

In generale possiamo dire:

- Se l'operatore non è continuo non si riesce ad estendere a tutto lo spazio (al massimo lo si riesce ad estendere ad uno spazio un po più grande di quello di partenza, ma non di più).
- Se l'operatore è continuo invece l'estensione è possibile e unica.

Possiamo far riferimento al teorema generale 2.25 di estensione di una funzione uniformemente continua per determinare l'esistenza e l'unicità della estensione. Nel caso particolare di una applicazione lineare e continua questo può essere formulato nella maniera seguente.

Teo. 5.6 Siano X e Y spazi di Banach, D un sottospazio lineare di X denso in X :

$$D^- = X, \quad (5.13)$$

e A una applicazione lineare e continua tra D e Y , $A \in \mathcal{L}_c(D, Y)$. Allora esiste un'unica applicazione lineare e continua $\tilde{A} \in \mathcal{L}_c(X, Y)$, che risulta una estensione di A a tutto lo spazio X :

$$Ax = \tilde{A}x, \quad \forall x \in D. \quad (5.14)$$

Inoltre vale:

$$\|A\| = \|\tilde{A}\|. \quad (5.15)$$

Rispetto al teorema 2.25 abbiamo, come conseguenza della linearità della applicazione A , le informazioni aggiuntive sulla linearità e sulla norma operatoriale dell'estensione.

Dim. 5.6 Notiamo subito che una applicazione lineare e continua risulta immediatamente uniformemente continua. La relazione:

$$\|Ax - Ax_0\| \leq \|A\| \|x - x_0\|, \quad \forall x_0, x \in D,$$

comporta che, assegnato $\epsilon > 0$, abbiamo:

$$\|x - x_0\| < \frac{\epsilon}{\|A\|} \implies \|Ax - Ax_0\| < \epsilon,$$

indipendentemente dalla scelta del punto x_0 . Possiamo allora invocare il risultato 2.25 per garantire l'esistenza e l'unicità dell'estensione a tutto lo spazio X .

Rimane da vedere la linearità e l'invarianza della norma operatoriale. Ricordando la dimostrazione del teorema di estensione di funzioni uniformemente continue, l'estensione può essere definita mediante il limite di una successione:

$$y(x) = \lim_{n \rightarrow \infty} Ax_n \quad \text{con } x = \lim_{n \rightarrow \infty} x_n,$$

dove $x_n \in D$ è una qualsiasi successione convergente a $x \in X$. Allora la linearità dell'estensione è immediata. Se $\alpha, \beta \in \mathbb{C}$, $x, x' \in X$:

$$\begin{aligned} & \left. \begin{aligned} x_n &\xrightarrow[n \rightarrow \infty]{} x, \quad x_n \in D, \\ x'_n &\xrightarrow[n \rightarrow \infty]{} x', \quad x'_n \in D, \end{aligned} \right\} \implies \alpha x_n + \beta x'_n \xrightarrow[n \rightarrow \infty]{} \alpha x + \beta x', \\ & y(\alpha x + \beta x') = \lim_{n \rightarrow \infty} A(\alpha x_n + \beta x'_n) \\ & \quad = \alpha \lim_{n \rightarrow \infty} Ax_n + \beta \lim_{n \rightarrow \infty} Ax'_n = \alpha y(x) + \beta y(x'). \end{aligned}$$

Questo permette di definire l'applicazione lineare \tilde{A} , estensione di A :

$$\tilde{A}x = y(x), \quad x \in X.$$

Sappiamo che:

$$\|Ax_n\| \leq \|A\| \|x_n\|, \quad x_n \in D, \implies \lim_{n \rightarrow \infty} \|Ax_n\| \leq \lim_{n \rightarrow \infty} \|A\| \|x_n\|,$$

ma la norma è una funzione continua per cui se $x_n \rightarrow x \in X$:

$$\|\tilde{A}x\| \leq \|A\| \|x\|, \quad \|\tilde{A}\| \leq \|A\|.$$

Poichè \tilde{A} è una estensione di A è chiaro che:

$$\sup_{\substack{x \in X \\ x \neq 0}} \frac{\|\tilde{A}x\|}{\|x\|} \geq \sup_{\substack{x \in D \\ x \neq 0}} \frac{\|\tilde{A}x\|}{\|x\|} = \sup_{\substack{x \in D \\ x \neq 0}} \frac{\|Ax\|}{\|x\|},$$

per cui:

$$\|\tilde{A}\| \geq \|A\|, \quad \|\tilde{A}\| = \|A\|.$$

5.2 Serie operatoriali.

Alcune funzioni dell'analisi sono spesso definite tramite il loro sviluppo in serie, specialmente quando il loro argomento è un numero complesso z . Si pensi ad esempio al caso della funzione esponenziale:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad z \in \mathbb{C}.$$

Il contesto giusto in cui operare in questo caso è quello delle funzioni analitiche, cioè sviluppabili in serie di potenze. Vogliamo ora poter estendere le definizioni di tali funzioni al caso in cui i termini della serie non sono semplici variabili complesse, ma operatori lineari definiti su uno spazio di Banach. L'insieme degli operatori limitati su uno spazio di Banach X , $\mathcal{L}_c(X, X)$, forma a sua volta uno spazio di Banach. In realtà $\mathcal{L}_c(X, X)$ ha una struttura più ricca, in quanto in esso risulta definito, tramite la legge di composizione, anche il prodotto tra due operatori, che gode delle proprietà associativa e distributiva rispetto alla somma:

$$\begin{aligned} A(BC) &= (AB)C, \\ (A+B)C &= AC + BC, \\ A(B+C) &= AB + AC, \\ \forall A, B, C \in \mathcal{L}_c(X, X), \end{aligned}$$

formando in questo modo un'**algebra associativa**. Noi abbiamo già visto durante la discussione sull'invertibilità dell'operatore $\mathbb{1} - T$ un esempio di serie di potenze di un operatore, e vogliamo generalizzare la problematica.

Un risultato fondamentale sulle serie di potenze è costituito dal teorema di Cauchy-Hadamard³, che fornisce informazioni sulla convergenza. Formuliamo tale risultato nel caso di spazi di Banach.

³ **Jacques Solomon Hadamard** (Versailles, 8 dicembre 1865 – Parigi, 17 ottobre 1963) è stato un matematico francese, conosciuto principalmente per la sua dimostrazione del teorema dei numeri primi. Dopo l'affair Dreyfus, che lo vide coinvolto personalmente, diventò un attivista politico e si trasformò in uno strenuo sostenitore delle cause ebraiche. Introdusse l'idea del problema *ben posto* nella teoria delle equazioni alle derivate parziali. Inoltre diede il suo nome alla diseguaglianza dei volumi, detta diseguaglianza di Hadamard, e alla matrice di Hadamard, su cui è basata la trasformata di Hadamard, usata anche in calcoli relativi alla impostazione matriciale della meccanica quantistica.



Teo. 5.7 (Cauchy–Hadamard) Sia E uno spazio di Banach (sul campo complesso), data una successione di elementi $a_j \in E$ e due numeri complessi $z_0, z \in \mathbb{C}$, si consideri la **serie di potenze** con punto iniziale z_0 :

$$a_0 + a_1(z - z_0) + \cdots + a_n(z - z_0)^n + \cdots, \quad (5.16)$$

al variare di $z \in \mathbb{C}$. Allora, posto:

$$\rho = \left(\limsup_{n \rightarrow \infty} \|a_n\|^{1/n} \right)^{-1}, \quad (5.17)$$

la serie (5.16) converge assolutamente per ogni z tale che $|z - z_0| < \rho$, e non converge per ogni z tale che $|z - z_0| > \rho$.

Dim. 5.7 Consideriamo la serie:

$$\sum_{n=0}^{\infty} \|a_n\| |z - z_0|^n,$$

e il massimo limite:

$$M = \limsup_{n \rightarrow \infty} \|a_n\|^{1/n}.$$

Assumiamo dapprima $M < \infty$, altrimenti il teorema perde la sua utilità, venendo a stabilire solo la divergenza quasi ovunque della serie. Ricordiamo che il massimo limite M è caratterizzato dalla seguente proprietà: dati due numeri qualsiasi M_1, M_2 , con $M_1 < M < M_2$, tutti gli elementi della successione $x_n = \|a_n\|^{1/n}$, escluso un numero finito, verificano $x_j \leq M_2$, ed esiste una infinità numerabile di elementi x_j con $x_j \geq M_1$. Allora, se:

$$|z - z_0| < \rho = \frac{1}{M},$$

esiste $\epsilon > 0$ tale che:

$$|z - z_0| < \frac{1}{M + \epsilon},$$

e, per le proprietà del limite sup, per tutti gli n a partire da un opportuno $N(\epsilon)$ si ha:

$$\|a_n\|^{1/n} \leq M + \frac{\epsilon}{2}.$$

Pertanto:

$$\sum_{n=N}^{\infty} \|a_n\| |z - z_0|^n \leq \sum_{n=0}^{\infty} \left(\frac{M + \frac{\epsilon}{2}}{M + \epsilon} \right)^n = 2 \frac{M + \epsilon}{\epsilon} < \infty.$$

Sia ora:

$$|z - z_0| > \rho = \frac{1}{M},$$

allora esiste $\epsilon > 0$ tale che:

$$|z - z_0| > \frac{1}{M - \epsilon},$$

ed esistono infiniti valori di n che verificano:

$$\|a_n\|^{1/n} \geq M - \epsilon,$$

per cui:

$$\|a_n\| |z - z_0|^n > 1,$$

per infiniti valori di n per cui la serie (5.16) non risulta di Cauchy, quindi non convergente.

Ovviamente, se $M = \infty$ e $|z - z_0| > \rho = 0$, la serie diviene divergente in quanto esiste $M' < \infty$ con:

$$|z - z_0| > \frac{1}{M'},$$

ed esistono infiniti valori di n per i quali:

$$\|a_n\|^{1/n} \geq M',$$

e la serie non è di Cauchy. Quando $z = z_0$ la serie è invece ovviamente convergente in quanto tutte le potenze sono nulle.

— ■ —

Osservazione. Il valore ρ dato dalla (5.17) è detto **raggio di convergenza** della serie di potenze. La serie (5.16) definisce, per $|z - z_0| < \rho$, cioè all'interno del raggio di convergenza, una funzione $f(z)$ della variabile complessa z , e a valori in uno spazio di Banach E :

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Tale funzione risulta derivabile ovunque all'interno del raggio di convergenza, cioè esiste ovunque il limite del rapporto incrementale:

$$f'(z) = \lim_{h \rightarrow 0} \frac{f(z + h) - f(z)}{h}, \quad h \in \mathbb{C},$$

(naturalmente il limite è calcolato mantenendo il punto $z + h$ all'interno del disco di convergenza), e tale derivata può essere calcolata derivando la serie termine a termine:

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (z - z_0)^n,$$

ottenendo una serie con lo stesso raggio di convergenza in quanto:

$$\lim_{n \rightarrow \infty} n^{1/n} = \lim_{n \rightarrow \infty} e^{1/n \ln n} = 1.$$

Tale risultato è una semplice conseguenza algebrica della formula di decomposizione della differenza tra due potenze:

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} &= \sum_{n=0}^{\infty} a_n \frac{(z+h-z_0)^n - (z-z_0)^n}{h} \\ &= \sum_{n=1}^{\infty} a_n \left\{ (z+h-z_0)^{n-1} + (z-z_0)(z+h-z_0)^{n-2} \right. \\ &\quad \left. + \cdots + (z-z_0)^{n-1} \right\}. \end{aligned}$$

Se $|h| < \epsilon$, in modo tale che anche $z+h$ sia all'interno del raggio di convergenza, $|z+h-z_0| < |z-z_0| + \epsilon = \delta < \rho$:

$$\|a_n \{(z+h-z_0)^{n-1} + \cdots + (z-z_0)^{n-1}\}\| \leq n \|a_n\| \delta^{n-1},$$

otteniamo una serie uniformemente convergente, per cui possiamo scambiare il limite per $h \rightarrow 0$ con la serie, cioè derivare la serie termine a termine, ottenendo il risultato.

Di conseguenza si ha che $f'(z)$ è derivabile ulteriormente, cioè $f(z)$ ammette derivate di ogni ordine, risulta cioè di classe C^∞ , e in realtà definisce una funzione analitica, cioè sviluppabile in serie di potenze nell'intorno di ogni punto z_1 interno al cerchio di convergenza.

Analogamente alla derivata, si può costruire una primitiva, integrando formalmente termine a termine:

$$g(z) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (z-z_0)^{n+1},$$

e si ha che $g(z)$ ha il medesimo raggio di convergenza della serie originale $f(z)$, e $g'(z) = f(z)$.

Osservazione. Il teorema di Cauchy–Hadamard non dice nulla quando z si trova sul bordo del cerchio di convergenza, cioè quando $|z-z_0| = \rho$. La situazione può essere molto varia. Possiamo avere non convergenza in alcun punto della frontiera, come nel caso della serie geometrica:

$$1 + (z-z_0) + (z-z_0)^2 + \cdots + (z-z_0)^n + \cdots,$$

con raggio di convergenza $\rho = 1$, oppure possiamo avere convergenza in tutti i punti della frontiera, come nella serie:

$$\sum_{n=1}^{\infty} \frac{(z-z_0)^n}{n^2},$$

in quanto, per $|z-z_0| = 1$ (raggio di convergenza di tale serie), questa risulta assolutamente convergente:

$$\sum_{n=1}^{\infty} \frac{|z-z_0|^n}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty,$$

oppure possiamo avere convergenza in alcuni punti e non convergenza in altri, come per la serie:

$$\sum_{n=1}^{\infty} \frac{(z - z_0)^n}{n},$$

divergente quando $z - z_0 = 1$, e convergente (semplicemente e non assolutamente) quando $z - z_0 = -1$.

5.2.1 Serie esponenziale

Tutti sanno che la funzione esponenziale:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

ha raggio di convergenza infinito in quanto:

$$\lim_{n \rightarrow \infty} (n!)^{1/n} = \lim_{n \rightarrow \infty} e^{1/n(\ln n + \ln(n-1) + \dots + \ln 1)} = \infty,$$

deducibile dal fatto che (vedi figura 5.1):

$$\begin{aligned} \frac{1}{n} (\ln n + \ln(n-1) + \dots + \ln 1) &> \frac{1}{n} \int_1^n \ln(x) dx \\ &= \frac{1}{n} [x \ln x - x]_{x=1}^{x=n} = \frac{1}{n} (n \ln(n) - n + 1) \xrightarrow{n \rightarrow \infty} \infty. \end{aligned}$$

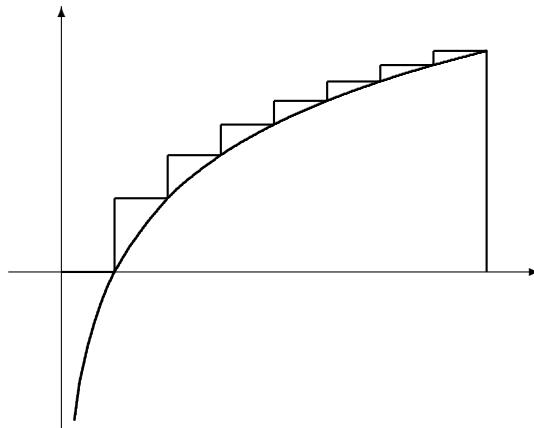


Figura 5.1: Valutazione per difetto della somma $\ln(1) + \dots + \ln(n)$.

La nozione di esponenziale può essere generalizzata. Sia X uno spazio di Banach e $A \in \mathcal{L}_c(X, X)$. Allora, per ogni $z \in \mathbb{C}$ (o anche solo \mathbb{R}) definiamo:

$$e^{zA} = \sum_{n=0}^{\infty} \frac{A^n}{n!} z^n. \quad (5.18)$$

Abbiamo infatti che:

$$\left\| \frac{A^n}{n!} \right\| \leq \frac{\|A\|^n}{n!},$$

per cui il raggio di convergenza è infinito e la **serie esponenziale** (5.18) risulta assolutamente convergente per ogni $z \in \mathbb{C}$ e definisce una funzione analitica in tutto il piano complesso e a valori in $\mathcal{L}_c(X, X)$. La serie esponenziale risulta infatti un operatore continuo, cioè limitato, con:

$$\|e^{zA}\| \leq \sum_{n=0}^{\infty} \frac{|z|^n \|A\|^n}{n!} = e^{|z|\|A\|}. \quad (5.19)$$

Notiamo che il prodotto di due esponenziali di operatori non è in generale uguale all'esponenziale della somma, ma se i due operatori commutano fra di loro allora ciò è vero.

Teo. 5.8 Siano $A, B \in \mathcal{L}_c(X, X)$ con X spazio di Banach. Se:

$$[A, B] = AB - BA = 0, \quad (5.20)$$

allora:

$$e^{zA} e^{zB} = e^{z(A+B)}. \quad (5.21)$$

Dim. 5.8 Abbiamo:

$$e^{zA} e^{zB} = \sum_{n=0}^{\infty} c_n z^n,$$

con:

$$c_n = \sum_{k=0}^n \frac{1}{k!} A^k \frac{1}{(n-k)!} B^{n-k}.$$

Se i due operatori non commutano tra di loro siamo obbligati a mantenere ordinatamente gli operatori B alla destra degli operatori A , ma se invece commutano tra loro, possiamo operare algebricamente come coi numeri complessi o reali, e ricondurci allo sviluppo algebrico di un binomio:

$$c_n = \frac{1}{n!} (A + B)^n,$$

ottenendo l'espansione in serie di potenze dell'esponenziale $e^{z(A+B)}$.

— ■ —

Osservazione. Poichè un operatore A commuta ovviamente con se stesso abbiamo che:

$$e^{z_1 A} e^{z_2 A} = e^{(z_1 + z_2) A}. \quad (5.22)$$

5.2.2 Serie di Neumann

Un altro esempio di sviluppo in serie di potenze di un operatore è dato dalla **serie di Neumann**:

$$\sum_{n=0}^{\infty} A^n z^n, \quad (5.23)$$

il cui raggio di convergenza è dato da:

$$\rho = \left(\limsup_{n \rightarrow \infty} \|A^n\|^{1/n} \right)^{-1}. \quad (5.24)$$

(Assumiamo sempre che $A \in \mathcal{L}_c(X, X)$, con X spazio di Banach)

In realtà in questo caso il \limsup risulta un limite normale in quanto:

$$r(A) = \limsup_{n \rightarrow \infty} \|A^n\|^{1/n} = \lim_{n \rightarrow \infty} \|A^n\|^{1/n} = \inf_n \|A^n\|^{1/n}. \quad (5.25)$$

Infatti, posto:

$$a = \inf_n \|A^n\|^{1/n}, \quad r(A) = \limsup_{n \rightarrow \infty} \|A^n\|^{1/n},$$

chiaramente si ha:

$$a \leq \liminf_{n \rightarrow \infty} \|A^n\|^{1/n} \leq r(A),$$

e se mostriamo che vale anche la disegualanza inversa, allora vale il risultato (5.25).

Sia $\epsilon > 0$ e, grazie alle proprietà dell'estremo inferiore, scegliamo m tale che $\|A^m\|^{\frac{1}{m}} < a + \epsilon$. Se n è arbitrario, possiamo dire che $n = mq + p$ con $0 \leq p \leq m - 1$. Allora:

$$\begin{aligned} \|A^n\|^{1/n} &= \|A^{mq} A^p\|^{1/n} \leq \|A^{mq}\|^{1/n} \|A^p\|^{1/n} \\ &\leq \|A^m\|^{q/m} \|A\|^{p/m} < (a + \epsilon)^{qm/m} \|A\|^{p/m}, \end{aligned}$$

p è limitato da m (indipendente da n), per cui i valori p/m , $p = 0, 1, \dots, m-1$ si accumulano nell'origine, mentre i valori $qm/m = qm/(qm + p)$ (in cui varia q al variare di n) si accumulano verso il valore unitario:

$$\begin{aligned} \limsup_{n \rightarrow \infty} (a + \epsilon)^{qm/m} &= \sup_{0 \leq p \leq m-1} \left\{ \limsup_{q \rightarrow \infty} (a + \epsilon)^{\frac{mq}{mq+p}} \right\} = a + \epsilon, \\ \limsup_{n \rightarrow \infty} \|A\|^{p/m} &= 1, \\ \limsup_{n \rightarrow \infty} \|A^n\|^{1/n} &\leq (a + \epsilon) \end{aligned}$$

ed essendo ϵ arbitrario:

$$r(A) \leq a,$$

dimostrando l'uguaglianza e il risultato (5.25) (chiaramente, questo prova anche l'esistenza del limite semplice, compreso tra i due estremi coincidenti).

Possiamo a questo punto generalizzare il risultato 5.2 mediante una serie di potenze dell'operatore A .

Teo. 5.9 Sia $A \in \mathcal{L}_c(X, X)$, con X spazio di Banach, e definiamo per ogni $z \in \mathbb{C}$, con:

$$|z| < \frac{1}{r(A)}, \quad (5.26)$$

dove $r(A)$ è espresso dalla relazione (5.25), la **serie di Neumann** per l'operatore A :

$$f(z) = \sum_{n=0}^{\infty} A^n z^n. \quad (5.27)$$

Allora, l'operatore lineare $\mathbb{1} - zA$ è invertibile per $|z| < 1/r(A)$, e vale:

$$f(z) = (\mathbb{1} - zA)^{-1}. \quad (5.28)$$

Dim. 5.9 Sia:

$$f_n(z) = \sum_{k=0}^n A^k z^k.$$

allora:

$$(\mathbb{1} - zA) f_n(z) = f_n(z) (\mathbb{1} - zA) = \mathbb{1} - z^{n+1} A^{n+1}.$$

Ora, $z^{n+1} A^{n+1} \rightarrow 0$ in norma per $n \rightarrow \infty$, perchè la serie di Neumann (5.27) è convergente, per cui, per $n \rightarrow \infty$ otteniamo:

$$(\mathbb{1} - zA) f(z) = f(z) (\mathbb{1} - zA) = \mathbb{1}.$$

— ■ —

Osservazione. Notiamo che $\|A^n\| \leq \|A\|^n$, per cui $r(A) \leq \|A\|$, e il raggio di convergenza:

$$\rho = \frac{1}{r(A)} \geq \frac{1}{\|A\|},$$

può risultare superiore a quanto stabilito dal risultato (5.2) basato sul teorema del punto fisso, in cui si richiede semplicemente $\|zA\| < 1$.

5.2.3 Serie di potenze di un operatore

Finora abbiamo considerato serie di potenze nella variabile complessa z , con i coefficienti appartenenti ad uno spazio di Banach, ma possiamo considerare anche le serie di potenze di un operatore A , con coefficienti complessi, generalizzando le serie di potenze complesse ordinarie.

Sia:

$$f(z) = \sum_{n=0}^{\infty} c_n z^n,$$

una serie di potenze nel piano complesso (cioè $c_n, z \in \mathbb{C}$) con raggio di convergenza ρ , determinato dal teorema di Cauchy-Hadamard. Ciò significa che per ogni z con $|z| < \rho$, la suddetta serie in \mathbb{C} è assolutamente convergente:

$$\sum_{n=0}^{\infty} |c_n| |z|^n < \infty.$$

Sappiamo che per ogni $A \in \mathcal{L}_c(X, X)$ vale $\|A^n\| \leq \|A\|^n$, per cui:

$$\left\| \sum_{n=0}^{\infty} c_n A^n \right\| \leq \sum_{n=0}^{\infty} |c_n| \|A\|^n.$$

Pertanto, se:

$$\|A\| < \rho, \quad (5.29)$$

la serie:

$$f(A) = \sum_{n=0}^{\infty} c_n A^n, \quad (5.30)$$

risulta sicuramente convergente in $\mathcal{L}_c(X, X)$, anzi assolutamente convergente, se X è uno spazio di Banach.

Un esempio è fornito dall'esponenziale stesso:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

da cui otteniamo l'esponenziale di un operatore:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Oppure considerando la serie geometrica:

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n, \quad |z| < 1,$$

ritroviamo la trasformazione inversa:

$$(\mathbb{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n, \quad \|A\| < 1.$$

5.3 Operatore aggiunto

Consideriamo ora uno spazio di Hilbert \mathcal{H} ed un operatore A definito su \mathcal{H} . Supponiamo cioè $X = Y = \mathcal{H}$. In uno spazio vettoriale finito-dimensionale un operatore lineare è sempre riconducibile ad una matrice di coefficienti ed una operazione semplice che si può eseguire su una matrice è quella di ottenere la matrice hermitiana coniugata operando le operazioni di trasposizione (scambio righe con colonne) e di coniugazione complessa. Quello che ci proponiamo è lo studio della generalizzazione di tale operazione al caso di dimensioni infinite, cioè per operatori in uno spazio di Hilbert.

Dato un operatore lineare A la proprietà fondamentale che si vuole utilizzare per la costruzione dell'aggiunto, che indicheremo con A^\dagger , oppure A^* , è espressa tramite i prodotti scalari:

$$\langle x, A y \rangle = \langle A^\dagger x, y \rangle.$$

Vogliamo studiare quando tale equazione ha senso e definisce effettivamente un operatore A^\dagger .

Poniamoci per ora in condizioni buone. Supponiamo di avere un operatore lineare e continuo $A \in \mathcal{L}_c(\mathcal{H}, \mathcal{H})$ e consideriamo, a fissato x , la corrispondenza tra \mathcal{H} e \mathbb{C} :

$$y \longrightarrow \langle x, A y \rangle. \quad (5.31)$$

Questa, essendo composizione di applicazioni continue e lineari (A e il prodotto scalare), definisce un funzionale lineare e continuo. Per il teorema di Fischer-Riesz esiste un unico elemento $x^* \in \mathcal{H}$ tale che:

$$\langle x, A y \rangle = \langle x^*, y \rangle,$$

che definisce l'operatore **aggiunto** A^\dagger :

$$A^\dagger x = x^*, \quad \langle A^\dagger x, y \rangle = \langle x, A y \rangle \quad (5.32)$$

per ogni $x \in \mathcal{H}$.

L'operatore così definito risulta lineare in conseguenza della linearità del prodotto scalare:

$$\begin{aligned} \langle \alpha x + \beta x', A y \rangle &= \bar{\alpha} \langle x, A y \rangle + \bar{\beta} \langle x', A y \rangle \\ \langle A^\dagger (\alpha x + \beta x'), y \rangle &= \bar{\alpha} \langle A^\dagger x, y \rangle + \bar{\beta} \langle A^\dagger x', y \rangle \\ \langle A^\dagger (\alpha x + \beta x'), y \rangle &= \langle \alpha A^\dagger x + \beta A^\dagger x', y \rangle, \end{aligned}$$

ed essendo ciò valido per ogni $y \in \mathcal{H}$:

$$A^\dagger (\alpha x + \beta x') = \alpha A^\dagger x + \beta A^\dagger x'.$$

Inoltre l'operatore A^\dagger risulta limitato al pari di A , infatti abbiamo:

$$\|A^\dagger x\|^2 = \langle A^\dagger x, A^\dagger x \rangle = \langle x, A A^\dagger x \rangle \leq \|x\| \|A\| \|A^\dagger x\|,$$

da cui, semplificando:

$$\|A^\dagger x\| \leq \|A\| \|x\|, \quad \|A^\dagger\| \leq \|A\|.$$

L'operatore A^\dagger risulta definito su tutto \mathcal{H} e possiamo operarne l'aggiunto ulteriore. Prendendo il complesso coniugato dell'equazione (5.32) risulta ovvio che, per operatori limitati:

$$(A^\dagger)^\dagger = A,$$

e i ruoli di A e A^\dagger possono essere scambiati, per cui $\|A\| \leq \|A^\dagger\|$ e:

$$\|A^\dagger\| = \|A\|.$$

Osservazione. Apriamo una breve parentesi per introdurre un metodo alternativo per il calcolo della norma di un operatore limitato in uno spazio di Hilbert \mathcal{H} . Se $A \in \mathcal{L}_c(\mathcal{H}, \mathcal{H})$, abbiamo:

$$\|A\| = \sup_{\|x\|, \|y\| \leq 1} |\langle x, Ay \rangle| = \sup_{\|x\| = \|y\| = 1} |\langle x, Ay \rangle| = \sup_{x, y \neq 0} \frac{|\langle x, Ay \rangle|}{\|x\| \|y\|}. \quad (5.33)$$

Infatti, per la diseguaglianza di Schwarz e la limitatezza di A :

$$|\langle x, Ay \rangle| \leq \|x\| \|Ay\| \leq \|A\| \|x\| \|y\|,$$

per cui ognuna delle espressioni della (5.33) risulta limitata superiormente da $\|A\|$. Inoltre, se z è un qualsiasi elemento dello spazio di Hilbert, dalla diseguaglianza di Schwarz abbiamo:

$$\sup_{\|x\| \leq 1} |\langle x, z \rangle| \leq \|z\|,$$

d'altra parte scegliendo:

$$x = \frac{z}{\|z\|}$$

raggiungiamo l'estremo superiore in un punto a norma unitaria e in cui vale l'uguaglianza, per cui:

$$\|z\| = \sup_{\|x\| \leq 1} |\langle x, z \rangle| = \sup_{\|x\| = 1} |\langle x, z \rangle| = \sup_{x \neq 0} \frac{|\langle x, z \rangle|}{\|x\|}.$$

Pertanto abbiamo ad esempio:

$$\|A\| = \sup_{\|y\| \leq 1} \|Ay\| = \sup_{\|y\| \leq 1} \left(\sup_{\|x\| \leq 1} |\langle x, Ay \rangle| \right) \leq \sup_{\|x\|, \|y\| \leq 1} |\langle x, Ay \rangle|,$$

ed in maniera analoga si ottiene una limitazione superiore di $\|A\|$ con qualsiasi delle espressioni (5.33) da cui la validità di tale equazione. In particolare abbiamo anche immediatamente:

$$\|A\| = \sup_{\|x\|, \|y\| \leq 1} |\langle x, Ay \rangle| = \sup_{\|x\|, \|y\| \leq 1} |\langle y, A^\dagger x \rangle| = \|A^\dagger\|.$$

Nel caso di operatori continui, l'operazione di aggiunzione definisce degli operatori continui, e si possono dare alcune semplici proprietà, oltre a quelle viste.

Teo. 5.10 Sia $A \in \mathcal{L}_c(\mathcal{H}, \mathcal{H})$, con \mathcal{H} spazio di Hilbert. Allora:

- i) A^\dagger è continuo e $\|A^\dagger\| = \|A\|$.
- ii) Se $B \in \mathcal{L}_c(\mathcal{H}, \mathcal{H})$, allora il prodotto AB risulta continuo e:

$$(AB)^\dagger = B^\dagger A^\dagger. \quad (5.34)$$

- iii) Il prodotto $A^\dagger A$ verifica:

$$\|A^\dagger A\| = \|A\|^2. \quad (5.35)$$

- iv) Il biaggiunto di A coincide con A :

$$A^{\dagger\dagger} = A. \quad (5.36)$$

Dim. 5.10 Abbiamo già visto la prima e la quarta proprietà. La seconda è conseguenza immediata della definizione di aggiunto:

$$\langle x, ABy \rangle = \langle A^\dagger x, By \rangle = \langle B^\dagger A^\dagger x, y \rangle.$$

Per quanto riguarda la terza affermazione, abbiamo:

$$\begin{aligned} \|Ax\|^2 &= \langle Ax, Ax \rangle = \langle A^\dagger Ax, x \rangle \leq \|A^\dagger Ax\| \|x\| \leq \|A^\dagger A\| \|x\|^2, \\ \|Ax\| &\leq \|A^\dagger A\|^{1/2} \|x\|, \quad \|A\| \leq \|A^\dagger A\|^{1/2}, \quad \|A\|^2 \leq \|A^\dagger A\|. \end{aligned}$$

D'altra parte considerandolo come prodotto di operatori:

$$\|A^\dagger A\| \leq \|A^\dagger\| \|A\| = \|A\|^2,$$

per cui:

$$\|A^\dagger A\| = \|A\|^2.$$

— ■ —

Vogliamo ora discutere la possibilità di definire l'aggiunto per un operatore **in generale non limitato**. Con x fissato, consideriamo di nuovo l'applicazione:

$$y \longrightarrow \langle x, Ay \rangle, \quad y \in \mathcal{D}(A),$$

essa è lineare ma non abbiamo in generale la garanzia della continuità a causa della non continuità dell'operatore. L'applicazione dipende però da x e potrebbe capitare che per particolari scelte di x questa risulti continua.

Assumiamo A definito su un dominio $\mathcal{D}(A)$ **denso** in \mathcal{H} (notiamo che il dominio di un operatore lineare è sempre un sottospazio vettoriale), e consideriamo l'insieme:

$$\mathcal{D}^* = \{x \in \mathcal{H} ; y \longrightarrow \langle x, Ay \rangle \text{ risulti continua per ogni } y \in \mathcal{D}(A)\}. \quad (5.37)$$

\mathcal{D}^* risulta non vuoto in quanto contiene almeno il vettore nullo (l'applicazione identicamente nulla risulta ovviamente continua). Allora, se $x \in \mathcal{D}^*$, l'applicazione:

$$\Phi : \mathcal{D}(A) \longrightarrow \mathbb{C},$$

definita da:

$$\Phi(y) = \langle x, A y \rangle,$$

risulta lineare e continua, definita su un sottoinsieme denso in \mathcal{H} per cui può essere estesa con continuità a tutto lo spazio di Hilbert \mathcal{H} . Possiamo allora invocare il teorema di Fischer-Riesz e affermare che esiste un vettore $x^* \in \mathcal{H}$ (che risulta anche unico) tale che:

$$\langle x, A y \rangle = \Phi(y) = \langle x^*, y \rangle, \quad \forall y \in \mathcal{D}(A) \quad (5.38)$$

L'unicità è conseguenza della densità del dominio. Infatti se avessimo un ulteriore vettore z^* tale che:

$$\begin{aligned} \langle x^*, y \rangle &= \langle z^*, y \rangle \quad \forall y \in \mathcal{D}(A), \\ \langle x^* - z^*, y \rangle &= 0 \quad \forall y \in \mathcal{D}(A), \end{aligned}$$

ma l'unico vettore ortogonale ad un sottoinsieme denso risulta il vettore nullo per cui $x^* = z^*$.

In conclusione se $\mathcal{D}(A)$ è denso in \mathcal{H} , abbiamo costruito una corrispondenza univoca tra \mathcal{D}^* e \mathcal{H} :

$$x \longrightarrow x^*,$$

che risulta ovviamente lineare, infatti se $z \in \mathcal{D}^*$ e $z \longrightarrow z^*$, allora, per ogni $\alpha, \beta \in \mathbb{C}$:

$$\begin{aligned} \langle \alpha x + \beta z, A y \rangle &= \overline{\alpha} \langle x, A y \rangle + \overline{\beta} \langle z, A y \rangle = \overline{\alpha} \langle x^*, y \rangle + \overline{\beta} \langle z^*, y \rangle \\ &= \langle \alpha x^* + \beta z^*, y \rangle \quad \forall y \in \mathcal{D}(A), \\ \alpha x + \beta z &\longrightarrow \alpha x^* + \beta z^*. \end{aligned}$$

Tale corrispondenza lineare definisce l'operatore **aggiunto** A^\dagger con dominio \mathcal{D}^* :

$$\mathcal{D}(A^\dagger) = \mathcal{D}^*, \quad A^\dagger x = x^*,$$

con la proprietà:

$$\langle x, A y \rangle = \langle A^\dagger x, y \rangle, \quad \forall x \in \mathcal{D}^*, \forall y \in \mathcal{D}(A).$$

Possiamo quindi riassumere il tutto nel risultato:

Teo. 5.11 Sia A un operatore lineare definito su un dominio $\mathcal{D}(A)$ denso in uno spazio

di Hilbert \mathcal{H} :

$$A : \mathcal{D}(A) \longrightarrow \mathcal{H},$$

allora, esiste un unico operatore lineare A^\dagger definito sul dominio:

$$\begin{aligned} \mathcal{D}(A^\dagger) = \mathcal{D}^* &= \{x \in \mathcal{H} ; \exists C > 0, \\ &|\langle x, Ay \rangle| \leq C \|y\| \forall y \in \mathcal{D}(A)\}, \end{aligned} \quad (5.39)$$

tale che:

$$\langle x, Ay \rangle = \langle A^\dagger x, y \rangle, \quad \forall x \in \mathcal{D}^*, \quad \forall y \in \mathcal{D}(A). \quad (5.40)$$

Osservazione. Notiamo come sia stato necessario, per poter invocare il teorema di Fischer-Riesz, richiedere la densità del dominio $\mathcal{D}(A)$. Sul dominio dell'aggiunto, \mathcal{D}^* possiamo solo dire che esiste, in quanto contiene almeno il vettore nullo, ma può anche risultare molto “piccolo”, cioè “cattivo” da un punto di vista pratico.

Notiamo inoltre la definizione alternativa del dominio dell'aggiunto \mathcal{D}^* , conseguenza del fatto che la continuità di una applicazione lineare equivale alla sua limitatezza.

5.3.1 Operatori autoaggiunti

Veniamo ora a delle definizioni importanti nello studio della meccanica quantistica. Vogliamo cioè discutere i concetti di operatore autoaggiunto, simmetrico, unitario, e il problema delle proiezioni ortogonali in uno spazio di Hilbert.

Def. 5.2 Sia A un operatore lineare densamente definito su uno spazio di Hilbert \mathcal{H} . Allora A è detto **autoaggiunto** se:

$$A = A^\dagger,$$

cioè se:

$$\mathcal{D}(A) = \mathcal{D}(A^\dagger), \quad \text{e} \quad Ax = A^\dagger x \quad \forall x \in \mathcal{D}(A).$$

A è invece detto **simmetrico** se A^\dagger è una estensione di A e scriveremo:

$$A \subseteq A^\dagger$$

intendendo con ciò:

$$\mathcal{D}(A) \subseteq \mathcal{D}(A^\dagger), \quad \text{e} \quad Ax = A^\dagger x \quad \forall x \in \mathcal{D}(A).$$

Osservazione. Possiamo dire che un operatore autoaggiunto rientra nella categoria generale degli operatori simmetrici, mentre nel caso di un operatore definito su tutto lo spazio

di Hilbert \mathcal{H} la distinzione tra operatore simmetrico ed operatore autoaggiunto perde di significato.

In pratica un operatore è simmetrico se verifica la relazione:

$$\langle x, Ay \rangle = \langle Ax, y \rangle \quad \forall x, y \in \mathcal{D}(A), \quad (5.41)$$

per cui l'aggiunto esiste e risulterà una estensione di A che, a seconda del suo dominio, può coincidere o non coincidere con A . Ponendo $x = y$ nell'equazione precedente:

$$\langle x, Ax \rangle = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle}$$

che risulta reale per ogni $x \in \mathcal{D}(A)$. Viceversa se $\langle x, Ax \rangle$ è reale per ogni $x \in \mathcal{D}(A)$ possiamo considerare la seguente relazione ($x, y \in \mathcal{D}(A)$):

$$\begin{aligned} \langle x + y, A(x + y) \rangle &= \langle A(x + y), x + y \rangle, \\ \langle x, Ax \rangle + \langle x, Ay \rangle + \langle y, Ax \rangle + \langle y, Ay \rangle &= \langle Ax, x \rangle + \langle Ax, y \rangle + \langle Ay, x \rangle + \langle Ay, y \rangle, \\ \langle x, Ay \rangle + \langle y, Ax \rangle &= \langle Ax, y \rangle + \langle Ay, x \rangle, \end{aligned}$$

Essendo ciò vero per ogni $x, y \in \mathcal{D}(A)$ allora è vero anche operando il cambiamento $y \rightarrow iy$:

$$\begin{aligned} \langle x, A(iy) \rangle + \langle iy, Ax \rangle &= \langle Ax, iy \rangle + \langle A(iy), x \rangle, \\ i\langle x, Ay \rangle - i\langle y, Ax \rangle &= i\langle Ax, y \rangle - i\langle Ay, x \rangle, \\ \langle x, Ay \rangle - \langle y, Ax \rangle &= \langle Ax, y \rangle - \langle Ay, x \rangle, \end{aligned}$$

ed operando la semi-somma con la precedente relazione:

$$\langle x, Ay \rangle = \langle Ax, y \rangle, \quad \forall x, y \in \mathcal{D}(A).$$

La validità di tale relazione comporta che se $x \in \mathcal{D}(A)$, allora $x \in \mathcal{D}(A^\dagger)$ e A è simmetrico.

Riassumendo, abbiamo mostrato la seguente affermazione:

Teo. 5.12 Sia A densamente definito su uno spazio di Hilbert \mathcal{H} . Allora A è simmetrico se e solo se $\langle x, Ax \rangle$ è reale per ogni $x \in \mathcal{D}(A)$.

5.3.2 Operatori unitari

Un altro concetto importante è quello di operatore unitario.

Def. 5.3 Un operatore lineare U definito su tutto uno spazio di Hilbert \mathcal{H} :

$$U : \mathcal{H} \longrightarrow \mathcal{H}, \quad \mathcal{D}(U) = \mathcal{H}$$

è detto **unitario** se è suriettivo e conserva il prodotto scalare:

$$\langle Ux, Uy \rangle = \langle x, y \rangle, \quad \forall x, y \in \mathcal{H}. \quad (5.42)$$

Osservazione. Un operatore unitario conserva quindi anche le norme:

$$\|Ux\|^2 = \langle Ux, Ux \rangle = \langle x, x \rangle = \|x\|^2,$$

e risulta quindi isometrico, ma nel concetto di operatore isometrico non è richiesta la suriettività. Affinchè un operatore isometrico risulti in effetti unitario occorre verificare la suriettività.

Esempio 5.4 A titolo di esempio si consideri l'operatore di “shift” definito su uno spazio di Hilbert \mathcal{H} (separabile) tramite la sua azione su un sistema ortonormale completo $\{e_j\}_{j=0}^{\infty}$:

$$Te_j = e_{j+1}, \quad j = 0, 1, \dots \quad (5.43)$$

che risulta isometrico ma non è suriettivo (il primo vettore e_0 non si ottiene mai come immagine di qualche vettore).

— ○ —

Teo. 5.13 Un operatore U definito su uno spazio di Hilbert \mathcal{H} risulta unitario se e solo se:

$$U^\dagger U = \mathbb{1} = U U^\dagger, \quad (5.44)$$

cioè U è invertibile con l'inverso coincidente con l'aggiunto:

$$U^{-1} = U^\dagger. \quad (5.45)$$

Dim. 5.13 Supponiamo U unitario, cioè definito su tutto \mathcal{H} , suriettivo e che conservi il prodotto scalare. La relazione (5.42) comporta che Ux appartiene al dominio $\mathcal{D}(U^\dagger)$ e:

$$U^\dagger Ux = x, \quad \forall x \in \mathcal{H}.$$

per cui:

$$U^\dagger U = \mathbb{1}.$$

Poichè il range di U coincide con tutto lo spazio \mathcal{H} allora la medesima relazione comporta anche che $\mathcal{D}(U^\dagger) = \mathcal{H}$, per cui, dato $x \in \mathcal{H}$ esiste $y \in \mathcal{H}$ tale che $Uy = x$ e:

$$UU^\dagger x = UU^\dagger Uy = Uy = x,$$

per cui si ha l'uguaglianza operatoriale:

$$U^\dagger U = UU^\dagger = \mathbb{1}.$$

Viceversa, supponiamo valida la relazione operatoriale (5.44). Tale equazione indica anche l'identità dei domini dei vari operatori, per cui:

$$\mathcal{D}(U) = \mathcal{D}(U^\dagger) = \mathcal{D}(\mathbb{1}) = \mathcal{H}.$$

$$U^\dagger Ux = x \quad \forall x \implies \mathcal{R}(U^\dagger) = \mathcal{H},$$

$$UU^\dagger x = x \quad \forall x \implies \mathcal{R}(U) = \mathcal{H},$$

e si ottiene la conservazione del prodotto scalare:

$$\langle Ux, Uy \rangle = \langle U^\dagger Ux, y \rangle = \langle x, y \rangle, \quad \forall x, y \in \mathcal{H}.$$

— ■ —

La proprietà di conservazione del prodotto scalare, tipica degli operatori unitari, ha una piccola conseguenza anche sulle immagini di un sottospazio.

Lem. 5.14 Se L è un sottospazio di uno spazio di Hilbert \mathcal{H} , e:

$$U : \mathcal{H} \longrightarrow \mathcal{H}$$

è unitario, allora:

$$(U(L))^\perp = U(L^\perp). \quad (5.46)$$

Dim. 5.14 Sia $z \in (U(L))^\perp$, allora, essendo U suriettivo, $z = Uy$, con $y \in \mathcal{H}$ e:

$$\langle Uy, Ux \rangle = 0 \quad \forall x \in L,$$

$$\langle y, x \rangle = 0 \quad \forall x \in L,$$

per cui $y \in L^\perp$ e $Uy = z \in U(L^\perp)$. Viceversa, assumiamo $z \in U(L^\perp)$, per cui $z = Uy$, con $y \in L^\perp$, per cui:

$$\langle y, x \rangle = 0 \quad \forall x \in L,$$

$$\langle Uy, Ux \rangle = \langle z, Ux \rangle = 0 \quad \forall x \in L,$$

cioè $z \in (U(L))^\perp$, e l'identità tra gli insiemi risulta provata.

— ■ —

5.3.3 Operatori di proiezione

Esaminiamo ora le proprietà di aggiunzione degli operatori di proiezione ortogonali definiti quando abbiamo discusso la possibilità di decomporre uno spazio di Hilbert in sottospazi ortogonali tra loro (vedi il teorema 4.6).

Se M è un sottospazio **chiuso** di uno spazio di Hilbert \mathcal{H} , allora vale la decomposizione:

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

ed ogni vettore $x \in \mathcal{H}$ può essere decomposto in maniera unica nella somma:

$$x = x' + x'', \quad x' \in M, \quad x'' \in M^\perp,$$

e la componente (univocamente determinata da x) x' definisce l'operatore di proiezione ortogonale su M :

$$P x = x',$$

definito su tutto \mathcal{H} .

Per il teorema di Pitagora la proiezione ortogonale x' ha una norma sicuramente inferiore a $\|x\|$, $\|x'\| \leq \|x\|$, per cui:

$$\|P x\| \leq \|x\| \quad \forall x \in \mathcal{H},$$

e P risulta limitato con $\|P\| \leq 1$. D'altra parte, se $x \in M$, allora la proiezione di x coincide con x stesso e:

$$P x = x \quad \forall x \in M,$$

per cui coincidono anche le norme e (a meno che P non sia identicamente nullo, cioè il sottospazio M sia formato dal solo vettore nullo):

$$\|P\| = 1.$$

Se invece $x \in M^\perp$ la sua proiezione su M risulta nulla, per cui abbiamo anche:

$$P x = 0 \quad \forall x \in M^\perp.$$

Inoltre, proiettando una proiezione su M , si riottiene ancora la proiezione, cioè risulta valida l'equazione:

$$P^2 = P.$$

e si dice che P è un operatore **idempotente**.

Poichè P è un operatore limitato, non ci sono problemi per la definizione dell'operatore aggiunto P^\dagger , che risulta anch'esso definito su tutto \mathcal{H} . Consideriamo ora due vettori x, y arbitrari e le loro proiezioni su M :

$$x' = P x, \quad y' = P y,$$

la decomposizione in sottospazi ortogonali comporta che i vettori $x - x'$ e $y - y'$ risultano ortogonali a M , per cui abbiamo:

$$\begin{aligned}\langle Px, y \rangle &= \langle x', y' + y - y' \rangle = \langle x', y' \rangle + \langle x', y - y' \rangle = \langle x', y' \rangle, \\ \langle x, Py \rangle &= \langle x' + x - x', y' \rangle = \langle x', y' \rangle + \langle x - x', y' \rangle = \langle x', y' \rangle, \\ \langle x, Py \rangle &= \langle Px, y \rangle,\end{aligned}$$

ed essendo ciò valido per ogni $x, y \in \mathcal{H}$, l'operatore di proiezione ortogonale risulta autoaggiunto:

$$P^\dagger = P.$$

Osservazione. La proprietà di autoaggiunzione e l'unitarietà della norma operatoriale sono una conseguenza dell'ortogonalità dei sottospazi che decompongono lo spazio di Hilbert. In mancanza di tale ortogonalità tali proprietà vengono a cadere, ma permangono le altre, in particolare la proprietà di idempotenza tipica delle proiezioni.

Riassumendo, dato un sottospazio chiuso M di uno spazio di Hilbert \mathcal{H} , risulta determinato un operatore limitato P definito su tutto \mathcal{H} con le proprietà:

$$\begin{aligned}P^2 &= P, \\ P^\dagger &= P,\end{aligned}$$

(5.47)

il cui range coincide con il sottospazio M stesso e:

$$\begin{aligned}Px &= x \quad \forall x \in M, \\ Px &= 0 \quad \forall x \in M^\perp.\end{aligned}\tag{5.48}$$

Queste ultime proprietà rendono unico l'operatore in quanto se P' verifica anch'esso:

$$P'x = x \quad \forall x \in M, \quad P'x = 0 \quad \forall x \in M^\perp.$$

Allora, per ogni $x \in \mathcal{H}$, decomponendolo nella somma:

$$x = x' + x'', \quad x' \in M, \quad x'' \in M^\perp,$$

abbiamo:

$$P'x = P'x' + P'x'' = x' = Px.$$

La corrispondenza tra sottospazi chiusi ed operatori idempotenti ed autoaggiunti è in realtà biunivoca, in quanto dato un operatore lineare P definito su tutto \mathcal{H} e che verifica le condizioni (5.47), possiamo costruire un sottospazio chiuso M il cui proiettore ortogonale

risulta proprio l'operatore P . Vediamo innanzitutto che P , in conseguenza delle (5.47) risulta limitato. Se $x \in \mathcal{H}$, abbiamo:

$$\begin{aligned}\|Px\|^2 &= \langle Px, Px \rangle = \langle x, P^\dagger Px \rangle \\ &= \langle x, P^2 x \rangle = \langle x, Px \rangle \leq \|x\| \|Px\|, \\ \|Px\| &\leq \|x\|.\end{aligned}$$

Definiamo ora il sottospazio M come:

$$M = \{x \in \mathcal{H} ; Px = x\}. \quad (5.49)$$

Essendo P limitato esso risulta un sottospazio chiuso, infatti se $x_n \in M$ è una successione convergente, $x_n \rightarrow x \in \mathcal{H}$, per continuità abbiamo:

$$Px_n = x_n, \implies Px = x,$$

per cui $x \in M$ e M è chiuso.

Vediamo inoltre che M coincide con il range dell'operatore P . Infatti se $x \in M$ allora, essendo l'immagine di se stesso, $x \in \mathcal{R}(P)$. Se, viceversa, $x \in \mathcal{R}(P)$ allora x è l'immagine di un elemento $y \in \mathcal{H}$:

$$x = Py \implies Px = P^2y = Py = x,$$

per cui $x \in M$ e:

$$\mathcal{R}(P) = M = M^-.$$

Vediamo ora che P si annulla sul complemento ortogonale di M . Sia $x \in M^\perp$, per cui se y è un arbitrario elemento di \mathcal{H} , $Py \in M$, e:

$$\langle y, Px \rangle = \langle Py, x \rangle = 0, \implies Px = 0.$$

L'operatore P verifica quindi le relazioni (5.48) con il sottospazio M definito da (5.49).

Osservazione. In coppia con P possiamo definire anche l'operatore:

$$Q = \mathbb{1} - P,$$

che gode anch'esso delle medesime proprietà di idempotenza ed autoaggiunzione:

$$Q^2 = (\mathbb{1} - P)^2 = \mathbb{1} - 2P + P^2 = \mathbb{1} - P = Q$$

$$Q^\dagger = (\mathbb{1} - P)^\dagger = \mathbb{1} - P = Q.$$

Inoltre:

$$PQ = P(\mathbb{1} - P) = P - P^2 = 0 = (\mathbb{1} - P)P = QP,$$

e Q risulta chiaramente il proiettore ortogonale sul sottospazio chiuso complemento ortogonale M^\perp .

Volendo estrarre delle conclusioni, abbiamo stabilito la validità della seguente proposizione.

Teo. 5.15 Ogni decomposizione di uno spazio di Hilbert \mathcal{H} in sottospazi chiusi M e M^\perp , ortogonali tra loro:

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

può essere posta in corrispondenza biunivoca (iniettiva e suriettiva) con una coppia di operatori P, Q , definiti su tutto \mathcal{H} , con le proprietà:

$$P^2 = P, \quad Q^2 = Q,$$

$$P^\dagger = P, \quad Q^\dagger = Q,$$

$$P + Q = \mathbb{1}, \quad PQ = QP = 0,$$

e che verificano:

$$Px = x, \quad Qx = 0, \quad \forall x \in M,$$

$$Px = 0, \quad Qx = x, \quad \forall x \in M^\perp.$$

Poichè un sottospazio chiuso di uno spazio di Hilbert risulta a sua volta uno spazio di Hilbert, l'affermazione può essere estesa induttivamente a più di due sottospazi.

Teo. 5.16 Ogni decomposizione di uno spazio di Hilbert \mathcal{H} in sottospazi chiusi

$$M_1, M_2, \dots, M_N,$$

ortogonali tra loro:

$$\mathcal{H} = M_1 \oplus M_2 \oplus \cdots \oplus M_N,$$

$$M_j \perp M_k \quad \forall j \neq k = 1, 2, \dots, N,$$

può essere posta in corrispondenza biunivoca (iniettiva e suriettiva) con una N -upla di operatori P_1, P_2, \dots, P_N , definiti su tutto \mathcal{H} , con le proprietà:

$$P_j P_k = \delta_{jk} P_k, \quad j, k = 1, \dots, N$$

$$P_j^\dagger = P_j, \quad j = 1, \dots, N$$

$$P_1 + P_2 + \cdots + P_N = \mathbb{1}$$

e che verificano, per $j = 1, \dots, N$:

$$P_j x = x, \quad \forall x \in M_j,$$

$$P_j x = 0, \quad \forall x \in M_j^\perp.$$

5.4 Trasformate di Fourier

La **trasformata di Fourier** è una trasformazione matematica che ad una funzione

$$f : \mathbb{R}^n \longrightarrow \mathbb{C}$$

associa una nuova funzione formalmente tramite l'operazione:

$$(\mathcal{F} f)(\xi) = \tilde{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) dx, \quad (5.50)$$

dove con $\xi \cdot x = \xi_1 x_1 + \dots + \xi_n x_n$ indichiamo il prodotto scalare ordinario di due vettori di \mathbb{R}^n . Ovviamente tale operazione ha senso solo se l'integrazione è possibile e vedremo sotto quali condizioni ciò è vero e la definizione è ben posta. L'importanza della trasformazione di Fourier risiede nelle sue proprietà algebriche e nella possibilità di estendere la sua definizione ad una opportuna classe di distribuzioni.

Assieme alla definizione precedente possiamo anche considerare la definizione (sempre quando risulta ben definita):

$$(\bar{\mathcal{F}} f)(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i\xi \cdot x} f(x) dx = \tilde{f}(-\xi). \quad (5.51)$$

che chiameremo, per distinguerla dalla precedente, la **trasformata di Fourier coniugata** oppure **antitrasformata di Fourier**. Ovviamente tutte le conclusioni che trarremo per la trasformata (diretta) risulteranno valide per la coniugata, in quanto la scelta del segno nell'esponenziale è arbitraria, come pure la scelta del fattore di normalizzazione (una volta operata una scelta convenzionale bisogna però mantenerla coerentemente).

Esempio 5.5 Consideriamo un semplice esempio (in una dimensione) che mostra alcune caratteristiche importanti della trasformazione di Fourier. Sia χ la funzione caratteristica dell'intervallo $[-1, 1]$:

$$\chi(x) = \begin{cases} 1 & \text{se } |x| \leq 1, \\ 0 & \text{se } |x| > 1, \end{cases} \quad (5.52)$$

allora l'integrale di Fourier (5.50) risulta ben definito:

$$\begin{aligned} \tilde{\chi}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x} \chi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-1}^{+1} e^{-i\xi x} dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{i}{\xi} e^{-i\xi x} \Big|_{x=-1}^{x=+1} = \frac{1}{\sqrt{2\pi}} \frac{i}{\xi} (e^{-i\xi} - e^{i\xi}) = \sqrt{\frac{2}{\pi}} \frac{\sin \xi}{\xi} \end{aligned}$$

Otteniamo cioè una funzione:

$$\tilde{\chi}(\xi) = \sqrt{\frac{2}{\pi}} \frac{\sin \xi}{\xi},$$

continua (malgrado la discontinuità della funzione caratteristica χ), limitata e tendente a zero per $\xi \rightarrow \infty$. Queste caratteristiche sono in realtà generali e riconducibili al solo fatto che la funzione caratteristica è una funzione sommabile.

— ○ —

Se consideriamo funzioni sommabili l'integrale di Fourier risulta ben definito (il fattore esponenziale è a modulo unitario) e possiamo vedere subito le conseguenze notate con l'esempio precedente.

Teo. 5.17 Sia $f \in L^1(\mathbb{R}^n)$. Allora l'integrale di Fourier $(\mathcal{F}f)(\xi)$ esiste per ogni $\xi \in \mathbb{R}^n$ e definisce una funzione $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{C}$ che gode delle seguenti proprietà:

i) \tilde{f} è limitata dalla norma in L^1 di f :

$$\|\mathcal{F}f\|_\infty \leq \frac{1}{(2\pi)^{n/2}} \|f\|_1. \quad (5.53)$$

ii) $\tilde{f} = \mathcal{F}f$ è uniformemente continua.

iii) $\tilde{f}(\xi)$ tende a 0 per $\xi \rightarrow \infty$:

$$\lim_{|\xi| \rightarrow \infty} \tilde{f}(\xi) = 0. \quad (5.54)$$

Dim. 5.17 La prima affermazione risulta una immediata conseguenza della sommabilità della funzione f :

$$|\tilde{f}(\xi)| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x)| dx = \frac{1}{(2\pi)^{n/2}} \|f\|_1.$$

Per vedere la continuità della trasformata di Fourier consideriamo la differenza:

$$\tilde{f}(\xi + h) - \tilde{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\xi \cdot x} (e^{-ih \cdot x} - 1) dx.$$

Per $h \rightarrow 0$ l'integrando si annulla per la continuità dell'esponenziale, per cui se possiamo passare il limite sotto il segno di integrale risulta verificata la continuità di \tilde{f} . Abbiamo la maggiorazione uniforme ed integrabile:

$$|f(x) e^{-i\xi \cdot x} (e^{-ih \cdot x} - 1)| \leq 2 |f(x)|,$$

pertanto possiamo applicare il teorema di convergenza dominata di Lebesgue e:

$$|\tilde{f}(\xi + h) - \tilde{f}(\xi)| \xrightarrow[h \rightarrow 0]{} 0.$$

L'uniforme continuità può essere vista ora come conseguenza della uniformità in ξ della maggiorazione ottenuta. Infatti abbiamo:

$$|\tilde{f}(\xi + h) - \tilde{f}(\xi)| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x)| |e^{-ih \cdot x} - 1| dx \xrightarrow[h \rightarrow 0]{} 0,$$

indipendentemente da ξ . In alternativa l'uniformità può essere vista come conseguenza della continuità, della limitatezza e del fatto che $\tilde{f}(\xi)$ tende a zero per $\xi \rightarrow \infty$ (in base alla successiva affermazione).

Vediamo ora l'ultima affermazione, che costituisce l'enunciato del cosiddetto **lemma di Riemann-Lebesgue**. Dalla teoria delle funzioni sommabili è noto che una funzione appartenente a $L^1(\mathbb{R})$ può essere approssimata da una funzione **semplice**, il cui codominio è cioè costituito da un numero **finito** di valori e che si può scrivere come:

$$\varphi(x) = \sum_{k=1}^N c_k \chi_{E_k}(x)$$

con χ_{E_k} funzione caratteristica di un insieme E_k a misura finita che non è restrittivo supporre un poliintervallo di \mathbb{R}^n (questa affermazione è anche nota come **teorema del sottografico**). Sostanzialmente significa che, data una funzione $f \in L^1(\mathbb{R}^n)$ e, fissato $\epsilon > 0$, esiste una funzione semplice φ con:

$$\|f - \varphi\|_1 = \int_{\mathbb{R}^n} |f(x) - \varphi(x)| dx < \epsilon.$$

La trasformata di Fourier della funzione caratteristica di un poliintervallo $E = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$:

$$\chi_E(x) = \begin{cases} 1 & x \in E, \\ 0 & x \notin E, \end{cases}$$

è facilmente calcolabile:

$$\tilde{\chi}_E(\xi) = \frac{1}{(2\pi)^{n/2}} \prod_{m=1}^n \frac{e^{-i a_m \xi_m} - e^{-i b_m \xi_m}}{i \xi_m}$$

e tende a zero per $\xi \rightarrow \infty$. Possiamo allora dire:

$$\begin{aligned} |\tilde{f}(\xi)| &= \frac{1}{(2\pi)^{n/2}} \left| \int_{\mathbb{R}^n} e^{-i \xi \cdot x} (f(x) - \varphi(x)) dx + \sum_{k=1}^N c_k \int_{\mathbb{R}^n} e^{-i \xi \cdot x} \chi_{E_k}(x) dx \right| \\ &\leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x) - \varphi(x)| dx + \sum_{k=1}^N |c_k| \frac{1}{(2\pi)^{n/2}} \left| \int_{\mathbb{R}^n} e^{-i \xi \cdot x} \chi_{E_k}(x) dx \right| \\ &= \frac{1}{(2\pi)^{n/2}} \|f - \varphi\|_1 + \sum_{k=1}^N |c_k| |\tilde{\chi}_{E_k}(\xi)|, \end{aligned}$$

e tutti i singoli termini della somma possono essere resi piccoli “a piacere”, scegliendo φ abbastanza “vicino” a f (nella topologia di $L^1(\mathbb{R}^n)$) e ξ abbastanza “grande” ($\xi \rightarrow \infty$).

— ■ —

Se ora consideriamo due funzioni $f, g \in L^1(\mathbb{R}^n)$ allora le loro trasformate di Fourier $\mathcal{F}f, \mathcal{F}g$ sono funzioni continue e limitate per cui possiamo considerare gli integrali:

$$\begin{aligned}\int_{\mathbb{R}^n} f(x) (\mathcal{F}g)(x) dx &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) g(\xi) d\xi dx \\ \int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) g(\xi) d\xi &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) g(\xi) dx d\xi\end{aligned}$$

che differiscono fra loro solo per l'ordine di integrazione. Ma l'integrandi verifica:

$$|e^{-i\xi \cdot x} f(x) g(\xi)| = |f(x)| |g(\xi)|,$$

che è sommabile su $\mathbb{R}^n \times \mathbb{R}^n$, per cui possiamo applicare il teorema di Fubini e scambiare l'ordine di integrazione. Quindi i due integrali, con la sola ipotesi di sommabilità delle due funzioni f e g , risultano uguali.

Teo. 5.18 Siano $f, g \in L^1(\mathbb{R}^n)$. Allora vale la relazione:

$$\int_{\mathbb{R}^n} f(x) (\mathcal{F}g)(x) dx = \int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) g(\xi) d\xi. \quad (5.55)$$

5.4.1 Convoluzione di funzioni.

Vogliamo ora dedicare un pò di attenzione ad una importante operazione che può essere eseguita (quando possibile) tra due funzioni, il cosiddetto prodotto di convoluzione. Date due funzioni:

$$f : \mathbb{R}^n \longrightarrow \mathbb{C}, \quad g : \mathbb{R}^n \longrightarrow \mathbb{C},$$

il **prodotto di convoluzione** $f * g$ è definito come la funzione:

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y) g(y) dy. \quad (5.56)$$

Il prodotto di convoluzione può essere visto, da un punto di vista puramente algebrico, come una operazione binaria nello spazio delle funzioni e può dare informazioni, da un punto di vista statistico, sulla correlazione tra due quantità f e g .

Vediamo ora sotto quali condizioni possiamo garantirci l'esistenza del prodotto di convoluzione e capire le sue proprietà.

Teo. 5.19 Se $f \in L^1(\mathbb{R}^n)$ e $g \in L^p(\mathbb{R}^n)$, con $1 \leq p \leq \infty$, allora la funzione:

$$y \longrightarrow f(x - y) g(y)$$

è integrabile (in y) al variare di x quasi ovunque in \mathbb{R}^n e, indicando il suo integrale

con $(f * g)(x)$, tale integrale definisce un elemento di $L^p(\mathbb{R}^n)$ per il quale vale la **disuguaglianza di Young**:

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p. \quad (5.57)$$

Dim. 5.19 Consideriamo prima il caso $p = \infty$ e sia $x \in \mathbb{R}^n$ fissato. La funzione

$$y \longrightarrow f(x - y)$$

appartiene a $L^1(\mathbb{R}^n)$, per cui, ricordando la disuguaglianza di Hölder, abbiamo che il prodotto $f(x - y)g(y)$ (sempre come funzione di y) appartiene a $L^1(\mathbb{R}^n)$, e:

$$\int_{\mathbb{R}^n} |f(x - y)g(y)| dy \leq \|f\|_1 \|g\|_\infty.$$

Essendo tale relazione valida per ogni $x \in \mathbb{R}^n$, il prodotto di convoluzione risulta limitato, cioè in $L^\infty(\mathbb{R}^n)$ e considerando l'estremo superiore al variare di x :

$$\|f * g\|_\infty \leq \|f\|_1 \|g\|_\infty.$$

Sia ora $p = 1$, consideriamo il prodotto $f(x - y)g(y)$ come funzione di entrambe le variabili x e y e l'integrale:

$$\int_{\mathbb{R}^n} dy |g(y)| \int_{\mathbb{R}^n} dx |f(x - y)| = \int_{\mathbb{R}^n} dy |g(y)| \int_{\mathbb{R}^n} dx |f(x)| = \|g\|_1 \|f\|_1,$$

chiaramente finito essendo entrambe le funzioni integrabili per ipotesi. Possiamo allora applicare il teorema di Fubini-Tonelli e affermare che il prodotto $f(x - y)g(y)$ definisce una funzione di $L^1(\mathbb{R}^n \times \mathbb{R}^n)$, l'integrale:

$$\int_{\mathbb{R}^n} dy f(x - y) g(y)$$

esiste quasi ovunque al variare di x e risulta (come funzione di x) a sua volta integrabile, per cui il prodotto di convoluzione $(f * g)(x)$ è ben definito quasi ovunque e:

$$\begin{aligned} \|f * g\|_1 &= \int_{\mathbb{R}^n} dx |(f * g)(x)| = \int_{\mathbb{R}^n} dx \left| \int_{\mathbb{R}^n} dy f(x - y) g(y) \right| \\ &\leq \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy |f(x - y) g(y)| = \|f\|_1 \|g\|_1. \end{aligned}$$

Sia ora $1 < p < \infty$. Allora, per ipotesi, $|g|^p$ è integrabile cioè appartiene a $L^1(\mathbb{R}^n)$ con:

$$\| |g|^p \|_1 = (\|g\|_p)^p.$$

Per quanto visto in precedenza, la funzione $y \rightarrow |f(x-y)| |g(y)|^p$ risulta in $L^1(\mathbb{R}^n)$ quasi ovunque al variare di x e possiamo considerare il prodotto di convoluzione:

$$(|f| * |g|^p)(x) = \int_{\mathbb{R}^n} dy |f(x-y)| |g(y)|^p,$$

ottenendo una funzione in $L^1(\mathbb{R}^n)$ con:

$$\| |f| * |g|^p \|_1 \leq \|f\|_1 \|g\|_p^p.$$

Sia ora q tale che:

$$\frac{1}{p} + \frac{1}{q} = 1,$$

abbiamo l'identità:

$$|f(x-y)| |g(y)| = |f(x-y)|^{1/q} |f(x-y)|^{1/p} |g(y)|$$

con (considerando la dipendenza da y):

$$|f(x-y)|^{1/q} \in L^q(\mathbb{R}^n), \quad |f(x-y)|^{1/p} |g(y)| \in L^p(\mathbb{R}^n).$$

Allora, per Hölder:

$$\begin{aligned} \int_{\mathbb{R}^n} dy |f(x-y)| |g(y)| &\leq \|f(x-y)^{1/q}\|_q \|f(x-y)^{1/p} g(y)\|_p \\ &= (\|f\|_1)^{1/q} (\| |f(x-y)| |g(y)|^p \|_1)^{1/p}, \end{aligned}$$

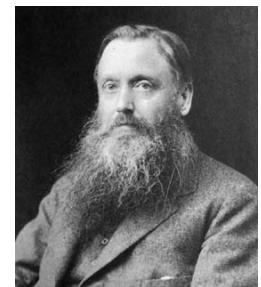
dove tutte le norme sono calcolate integrando rispetto alla variabile y . Elevando il tutto alla potenza p -esima:

$$\begin{aligned} |(f * g)(x)|^p &= \left| \int_{\mathbb{R}^n} dy f(x-y) g(y) \right|^p \leq \left(\int_{\mathbb{R}^n} dy |f(x-y)| |g(y)| \right)^p \\ &\leq \|f\|_1^{p/q} \left(\int_{\mathbb{R}^n} dy |f(x-y)| |g(y)|^p \right) = (\|f\|_1)^{p/q} (|f| * |g|^p)(x), \end{aligned}$$

ottenendo il risultato di Young⁴:

$$\begin{aligned} \|f * g\|_p &= \left(\int_{\mathbb{R}^n} dx |(f * g)(x)|^p \right)^{1/p} \leq (\|f\|_1)^{1/q} (\| |f| * |g|^p \|_1)^{1/p} \\ &\leq (\|f\|_1)^{1/q} (\|f\|_1 \| |g|^p \|_1)^{1/p} = (\|f\|_1)^{1/q} (\|f\|_1)^{1/p} \|g\|_p = \|f\|_1 \|g\|_p. \end{aligned}$$

⁴ **William Henry Young** (Londra, 20 ottobre 1863 – Losanna, 7 luglio 1942) è stato un matematico inglese. Ha lavorato principalmente sulla teoria della misura, serie di Fourier, calcolo differenziale, ed ha fornito importanti e durevoli contributi allo studio di funzioni di molte variabili complesse.



— ■ —

Osservazione. L'operazione di convoluzione è commutativa:

$$\begin{aligned}(f * g)(x) &= \int_{\mathbb{R}^n} f(x-y) g(y) dy \\ &= \int_{\mathbb{R}^n} f(z) g(x-z) dz = (g * f)(x),\end{aligned}$$

per cui i ruoli delle due funzioni f e g possono essere scambiati tra loro nel teorema precedente.

Vediamo ora una importante connessione tra la trasformata di Fourier e il prodotto di convoluzione. Se due funzioni f, g , appartengono a $L^1(\mathbb{R}^n)$, allora esistono le loro trasformate di Fourier, $\mathcal{F}f, \mathcal{F}g$, ed il loro prodotto di convoluzione $f * g$ definisce una ulteriore funzione in $L^1(\mathbb{R}^n)$ di cui possiamo fare la trasformata di Fourier:

$$\begin{aligned}(\mathcal{F}(f * g))(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} (f * g)(x) dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dx e^{-i\xi \cdot x} \int_{\mathbb{R}^n} dy f(x-y) g(y) \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy f(x-y) e^{-i\xi \cdot (x-y)} g(y) e^{-i\xi \cdot y},\end{aligned}$$

ma l'integrandino risulta chiaramente una funzione sommabile su $\mathbb{R}^n \times \mathbb{R}^n$ e possiamo applicare il teorema di Fubini cambiando l'ordine di integrazione:

$$\begin{aligned}(\mathcal{F}(f * g))(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dy e^{-i\xi \cdot y} g(y) \int_{\mathbb{R}^n} dx e^{-i\xi \cdot (x-y)} f(x-y) \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dy e^{-i\xi \cdot y} g(y) \int_{\mathbb{R}^n} dx e^{-i\xi \cdot x} f(x) \\ &= (2\pi)^{n/2} (\mathcal{F}f)(\xi) (\mathcal{F}g)(\xi).\end{aligned}$$

Possiamo allora affermare che la trasformata di Fourier riduce l'operazione di convoluzione ad un semplice prodotto.

Teo. 5.20 Se f, g appartengono a $L^1(\mathbb{R}^n)$, allora:

$$\mathcal{F}(f * g) = (2\pi)^{n/2} (\mathcal{F}f) (\mathcal{F}g). \quad (5.58)$$

5.4.2 Funzioni a decrescenza rapida.

Vediamo ora le conseguenze della regolarità della funzione f sulla trasformata di Fourier $\mathcal{F}f$. Per semplicità consideriamo il caso unidimensionale, cioè funzioni di una sola variabile reale, e sarà semplice generalizzare a funzioni di n variabili. Supponiamo di avere una funzione f continua, sommabile e che tenda a zero all'infinito. Se la funzione è derivabile e la derivata f' è continua (e si annulla all'infinito) e sommabile, possiamo farne la trasformata di Fourier:

$$(\mathcal{F}f')(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\xi x} f'(x) dx,$$

e possiamo integrare per parti, considerando che all'infinito la funzione si annulla:

$$\begin{aligned} (\mathcal{F}f')(\xi) &= \frac{1}{\sqrt{2\pi}} - \int_{-\infty}^{+\infty} (-i\xi) e^{-i\xi x} f(x) dx \\ &= i\xi (\mathcal{F}f)(\xi). \end{aligned}$$

Se la funzione f è derivabile con continuità fino all'ordine k possiamo ripetere il ragionamento anche per la derivata di ordine k , e integrando ripetutamente per parti (assumendo che ogni derivata risulti sommabile e tenda a zero all'infinito):

$$(\mathcal{F}f^{(k)})(\xi) = (i\xi)^k (\mathcal{F}f)(\xi). \quad (5.59)$$

L'operazione di derivazione diventa quindi, operando la trasformata di Fourier, una semplice operazione di moltiplicazione per un monomio. In più, notando che la trasformata di Fourier di una funzione sommabile è limitata e tende a zero all'infinito, abbiamo anche:

$$|\xi^k (\mathcal{F}f)(\xi)| \leq |(\mathcal{F}f^{(k)})(\xi)| \leq \text{const}. \quad (5.60)$$

e quindi, in particolare, $\mathcal{F}f$ deve tendere a zero più rapidamente di $1/\xi^k$.

Maggiore è la regolarità di una funzione f maggiore è la “velocità” con cui la trasformata di Fourier tende a zero all'infinito. Possiamo addirittura considerare funzioni derivabili con continuità ad ogni ordine e che siano non nulle solo all'interno di un insieme compatto di \mathbb{R}^n (per non avere problemi all'infinito). In questo caso vale la relazione:

$$\left| \xi^\alpha \int_{\mathbb{R}^n} e^{-i\xi \cdot x} \varphi(x) dx \right| \leq \|D^\alpha \varphi\|_1.$$

con α multiindice arbitrario. Questo comporta non solo che la trasformata di Fourier di una tale funzione tende a zero all'infinito ma che vi tende anche molto rapidamente, più velocemente di qualsiasi potenza. Questo ha portato alla definizione delle cosiddette **funzioni a decrescenza rapida**. Tale classe di funzioni, indicate con il simbolo $\mathcal{S}(\mathbb{R}^n)$ (o

semplicemente \mathcal{S}) dal nome del matematico Schwartz⁵ che le ha introdotte negli anni '40, è definita nel seguente modo.

Def. 5.4 Lo spazio $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$, detto delle funzioni a **decrescenza rapida** o **spazio di Schwartz**, è costituito dalle funzioni φ di classe $\mathcal{C}^\infty(\mathbb{R}^n)$ tali che per ogni multiindice α e multiindice β valga:

$$\sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty, \quad (5.61)$$

(cioè l'estremo superiore, che dipenderà da α e β esiste ed è finito).

Cioè la funzione φ e tutte le sue derivate devono tendere a zero più rapidamente di qualsiasi potenza (oltre ad essere estremamente regolari). Un esempio tipico di funzione a decrescenza rapida è costituito dalla gaussiana:

$$e^{-x^2},$$

mentre ad esempio la lorentziana:

$$\frac{1}{1+x^2},$$

pur essendo di classe \mathcal{C}^∞ e decrescendo a zero, non vi tende abbastanza rapidamente.

Osservazione. Le funzioni di classe \mathcal{C}^∞ che si annullano al di fuori di un compatto (chiuso e limitato) sono dette di classe \mathcal{C}_0^∞ risultano chiaramente anche a decrescenza rapida (il supporto compatto e la regolarità garantiscono anche l'esistenza dell'estremo superiore), inoltre possiamo affermare, insiemisticamente, che:

$$\mathcal{C}_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n). \quad (5.62)$$

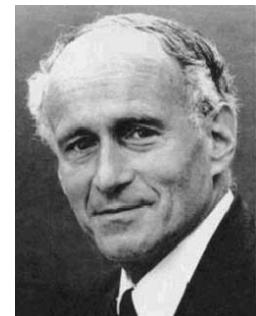
La seconda inclusione è conseguenza della decrescenza rapida. Se $\varphi \in \mathcal{S}$, allora per ogni intero N e multiindice β esiste una costante $C_{N,\beta}$ tale che:

$$(1 + |x|^2)^N |D^\beta \varphi(x)| \leq C_{N,\beta},$$

per cui:

$$|D^\beta \varphi(x)| \leq \frac{C_{N,\beta}}{(1 + |x|^2)^N},$$

⁵ Laurent Schwartz (Parigi, 5 Marzo 1915 – Parigi, 4 Luglio 2002) è stato un matematico francese. Il suo considerevole lavoro matematico, tra cui la teoria delle distribuzioni, gli fece vincere la Fields Medal nel 1950. Aldilà della sua attività scientifica, Schwartz fu un ben noto e schietto intellettuale, con simpatie verso il comunismo, ma rifiutò il totalitarismo di Stalin. In pericolo di essere classificato come ebreo sotto le leggi razziali naziste, dovette passare parte della Seconda Guerra mondiale nascondendosi sotto falso nome, prevalentemente come "Laurent Sélimartin".



che risulta sommabile operando una scelta abbastanza grande per N . Quindi non solo φ è sommabile, ma anche tutte le sue derivate.

Vediamo ora che la scelta dello spazio funzionale è ben posta ai fini della trasformazione di Fourier. Sia $f \in \mathcal{S}$ e poniamo:

$$\tilde{f}(\xi) = \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} e^{-i\xi \cdot x} f(x),$$

vogliamo mostrare che anche \tilde{f} è indefinitamente derivabile e tutte le sue derivate tendono rapidamente a zero. Consideriamo la differenziabilità:

$$\begin{aligned} \tilde{f}(\xi + \delta) - \tilde{f}(\xi) &= \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} (e^{-i\delta \cdot x} - 1) e^{-i\xi \cdot x} f(x) \\ &= \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} (e^{-i\delta \cdot x} - 1 + i\delta \cdot x) e^{-i\xi \cdot x} f(x) \\ &\quad + \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} (-i\delta \cdot x) e^{-i\xi \cdot x} f(x), \end{aligned}$$

Se riusciamo a provare che:

$$\omega(\delta) = \frac{1}{|\delta|} \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} (e^{-i\delta \cdot x} - 1 + i\delta \cdot x) e^{-i\xi \cdot x} f(x) \xrightarrow[\delta \rightarrow 0]{} 0$$

allora l'espressione:

$$-i \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} x e^{-i\xi \cdot x} f(x)$$

costituisce il differenziale di $\tilde{f}(\xi)$ e fornisce le derivate di \tilde{f} .

Se eseguiamo una maggiorazione brutale del resto della serie:

$$|e^{-i\delta \cdot x} - 1 + i\delta \cdot x| \leq \sum_{j=2}^{\infty} \frac{|\delta \cdot x|^j}{j!} = e^{|\delta \cdot x|} - 1 - |\delta \cdot x|,$$

anche se $\delta \rightarrow 0$, nell'integrazione non riusciamo a compensare la crescita esponenziale con la decrescenza rapida di f . Occorre procedere con maggiore cautela.

$$\begin{aligned} e^{i\alpha} - 1 - i\alpha &= \int_0^1 (1-t) dt \left(\frac{d}{dt} \right)^2 e^{i\alpha t} \\ &= -\alpha^2 \int_0^1 dt (1-t) e^{i\alpha t} \quad \Rightarrow \\ |e^{i\alpha} - 1 - i\alpha| &\leq |\alpha|^2 \int_0^1 dt (1-t) = \frac{|\alpha|^2}{2}, \end{aligned}$$

con α reale. Pertanto:

$$\begin{aligned}\omega(\delta) &\leq \frac{1}{2|\delta|} \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} |f(x)| (\delta \cdot x)^2 \\ &\leq \frac{|\delta|}{2} \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} |f(x)| |x|^2 \xrightarrow[|\delta| \rightarrow 0]{} 0\end{aligned}$$

essendo l'integrale finito e indipendente da δ . Abbiamo quindi verificato la differenziabilità e:

$$\frac{\partial}{\partial \xi_j} \tilde{f}(\xi) = \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} (-i x_j) e^{-i \xi \cdot x} f(x), \quad j = 1, 2, \dots, n.$$

Sostanzialmente abbiamo mostrato di poter derivare sotto il segno di integrale e che le derivate di \tilde{f} sono a loro volta delle trasformate di Fourier di funzioni in \mathcal{S} (un fattore polinomiale moltiplicativo non altera ovviamente l'appartenenza a \mathcal{S}). Il ragionamento può quindi essere iterato indefinitamente per cui $\tilde{f} \in \mathcal{C}^\infty(\mathbb{R}^n)$ e in generale abbiamo l'importante relazione:

$$D^\alpha (\mathcal{F} f) = \mathcal{F}((-i x)^\alpha f).$$

Vediamo ora cosa succede moltiplicando per una variabile ξ_j , con $j = 1, 2, \dots, n$. Mediante una integrazione per parti (f è di classe $\mathcal{C}^\infty(\mathbb{R}^n)$ e si annulla all'infinito>):

$$\begin{aligned}\xi_j \tilde{f}(\xi) &= \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} \xi_j e^{-i \xi \cdot x} f(x) \\ &= \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} \left(i \frac{d}{dx_j} e^{-i \xi \cdot x} \right) f(x) \\ &= \int_{\mathbb{R}^n} \frac{dx}{(2\pi)^{\frac{n}{2}}} e^{-i \xi \cdot x} \left(-i \frac{df(x)}{dx_j} \right),\end{aligned}$$

per cui abbiamo ancora la trasformata di Fourier di una funzione in \mathcal{S} . Iterando il ragionamento otteniamo in generale:

$$\xi^\alpha (\mathcal{F} f) = \mathcal{F}((-i D)^\alpha f).$$

Combinando le due regole ottenute:

$$\xi^\alpha D^\beta (\mathcal{F} f) = \xi^\alpha \mathcal{F}((-i x)^\beta f) = \mathcal{F}((-i D)^\alpha ((-i x)^\beta f)).$$

ed in particolare, essendo la trasformata di Fourier limitata:

$$|\xi^\alpha D^\beta (\mathcal{F} f)(\xi)| \leq M_{\alpha \beta} < \infty.$$

In questo modo otteniamo che la trasformata di Fourier di una funzione in \mathcal{S} è ancora in \mathcal{S} :

$$\mathcal{F} : \mathcal{S} \longrightarrow \mathcal{S}.$$

5.4.3 Formula di inversione.

Abbiamo visto che se $f, g \in L^1(\mathbb{R}^n)$ abbiamo:

$$\int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) g(\xi) d\xi = \int_{\mathbb{R}^n} f(\xi) (\mathcal{F}g)(\xi) d\xi.$$

Se potessimo scegliere g in maniera tale che $(\mathcal{F}g)(\xi) = \delta(\xi - x)$, avremmo una formula del tipo:

$$f(x) = \int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) g(\xi) d\xi,$$

che esprimerebbe una funzione f quando è nota la sua trasformata di Fourier $\mathcal{F}f$. Purtroppo sappiamo che non esiste una funzione del tipo richiesto, ma possiamo approssimare bene la δ di Dirac con funzioni ordinarie.

Vediamo quindi di lavorare con funzioni di classe \mathcal{S} e assumiamo che g sia del tipo:

$$g(y) = e^{iy \cdot x} \varphi(y),$$

con $\varphi \in \mathcal{S}$. Il fattore esponenziale di classe \mathcal{C}^∞ e limitato non cambia la classe di appartenenza e abbiamo:

$$(\mathcal{F}g)(\xi) = \int_{\mathbb{R}^n} \frac{dy}{(2\pi)^{\frac{n}{2}}} e^{-i\xi \cdot y} e^{iy \cdot x} \varphi(y) = (\mathcal{F}\varphi)(\xi - x),$$

e per ogni $\varphi \in \mathcal{S}$:

$$\int_{\mathbb{R}^n} d\xi (\mathcal{F}f)(\xi) e^{i\xi \cdot x} \varphi(\xi) = \int_{\mathbb{R}^n} d\xi f(\xi) (\mathcal{F}\varphi)(\xi - x).$$

Consideriamo ora la funzione $\varphi_\epsilon(y) = \varphi(\epsilon y)$ al posto di $\varphi(y)$, con ϵ positivo, ma piccolo:

$$\begin{aligned} (\mathcal{F}\varphi_\epsilon)(\xi) &= \int_{\mathbb{R}^n} \frac{dy}{(2\pi)^{\frac{n}{2}}} e^{-i\xi \cdot y} \varphi(\epsilon y) \\ &= \frac{1}{\epsilon^n} \int_{\mathbb{R}^n} \frac{d(\epsilon y)}{(2\pi)^{\frac{n}{2}}} e^{-i\frac{\xi}{\epsilon} \cdot (\epsilon y)} \varphi(\epsilon y) \\ &= \frac{1}{\epsilon^n} \int_{\mathbb{R}^n} \frac{dy}{(2\pi)^{\frac{n}{2}}} e^{-i\frac{\xi}{\epsilon} \cdot y} \varphi(y) = \frac{1}{\epsilon^n} (\mathcal{F}\varphi)\left(\frac{\xi}{\epsilon}\right), \end{aligned}$$

Quindi, se $f \in L^1(\mathbb{R}^n)$ e $\varphi \in \mathcal{S}(\mathbb{R}^n)$:

$$\begin{aligned} \int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) e^{i\xi \cdot x} \varphi(\epsilon \xi) d\xi &= \int_{\mathbb{R}^n} f(\xi) (\mathcal{F}\varphi)\left(\frac{\xi - x}{\epsilon}\right) \frac{d\xi}{\epsilon^n} \\ &= \int_{\mathbb{R}^n} f(x + \epsilon \xi) (\mathcal{F}\varphi)(\xi) d\xi, \end{aligned}$$

dove abbiamo operato un semplice cambio di variabile $\frac{\xi-x}{\epsilon} \rightarrow \xi$. Vediamo ora di far tendere ϵ a zero portando il limite dentro le integrazioni. Nel primo integrale abbiamo:

$$(\mathcal{F}f)(\xi) e^{i\xi \cdot x} \varphi(\epsilon\xi) \xrightarrow[\epsilon \rightarrow 0]{} \varphi(0) (\mathcal{F}f)(\xi) e^{i\xi \cdot x},$$

$$|(\mathcal{F}f)(\xi) e^{i\xi \cdot x} \varphi(\epsilon\xi)| \leq |(\mathcal{F}f)(\xi)| \sup_{z \in \mathbb{R}^n} |\varphi(z)|,$$

per cui possiamo far passare il limite all'interno dell'integrale se assumiamo come ipotesi che $\mathcal{F}f \in L^1(\mathbb{R}^n)$ (oltre alla ipotesi $f \in L^1(\mathbb{R}^n)$).

Nell'ultimo integrale la sola ipotesi di appartenenza a $L^1(\mathbb{R}^n)$ non è in generale sufficiente, ma abbiamo bisogno di un "briciole" di continuità sulla funzione f per poter portare il limite dentro la funzione stessa. Abbiamo:

$$|f(x + \epsilon \xi) (\mathcal{F}\varphi)(\xi)| \leq \sup_{z \in \mathbb{R}^n} |f(z)| |(\mathcal{F}\varphi)(\xi)|,$$

per cui se assumiamo f **limitata** e **continua** possiamo applicare Lebesgue, e portare il limite entro l'argomento di f . Verificate tali ipotesiabbiamo:

$$\varphi(0) \int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) e^{i\xi \cdot x} d\xi = f(x) \int_{\mathbb{R}^n} (\mathcal{F}\varphi)(\xi) d\xi.$$

Operando ora una scelta di una funzione φ , possiamo esplicitare una formula di inversione. Prendiamo come funzione in \mathcal{S} la gaussiana:

$$G(x) = e^{-\frac{x^2}{2}}, \quad x^2 = x \cdot x = \sum_{k=1}^n x_k^2, \quad (5.63)$$

e calcoliamo esplicitamente la sua trasformata di Fourier:

$$\tilde{G}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x - \frac{x^2}{2}} dx.$$

Il calcolo di tale integrale è ricondotto al calcolo unidimensionale, per cui possiamo considerare $n = 1$. Abbiamo:

$$\tilde{G}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1.$$

Poichè la gaussiana è regolare e a decrescenza rapida, possiamo derivare sotto il segno di integrale:

$$\begin{aligned} \tilde{G}'(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} (-ix) e^{-i\xi \cdot x - \frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} \left(i \frac{d}{dx} e^{-\frac{x^2}{2}} \right) dx \\ &= \frac{-i}{\sqrt{2\pi}} \int_{\mathbb{R}^n} \left(\frac{d}{dx} e^{-i\xi \cdot x} \right) e^{-\frac{x^2}{2}} dx \\ &= -\xi \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x - \frac{x^2}{2}} dx = -\xi \tilde{G}(\xi). \end{aligned}$$

Il problema di Cauchy (a variabili separabili):

$$\tilde{G}'(\xi) = -\xi \tilde{G}(\xi), \quad \tilde{G}(0) = 1,$$

ammette come unica soluzione:

$$\tilde{G}(\xi) = e^{-\frac{\xi^2}{2}}, \quad (5.64)$$

cioè la trasformata di Fourier di una gaussiana è ancora una gaussiana (con la scelta fatta dei parametri abbiamo nel nostro caso la stessa gaussiana di partenza). Inoltre abbiamo:

$$\int_{-\infty}^{\infty} \tilde{G}(\xi) d\xi = \sqrt{2\pi}.$$

Tornando a n dimensioni otteniamo la formula di inversione:

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\xi \cdot x} (\mathcal{F}f)(\xi) d\xi = (\bar{\mathcal{F}}(\mathcal{F}f))(x), \quad (5.65)$$

valida se $f \in L^1(\mathbb{R}^n)$, $\mathcal{F}f \in L^1(\mathbb{R}^n)$, f continua e limitata. Scambiando i ruoli di ξ e $-\xi$, nelle medesime ipotesiabbiamo:

$$f(x) = (\mathcal{F}(\bar{\mathcal{F}}f))(x), \quad (5.66)$$

per cui possiamo dire che in opportune condizioni la trasformata di Fourier e la sua coniugata esprimono l'una la trasformazione inversa dell'altra.

Se consideriamo una funzione $\varphi \in \mathcal{S}$ allora $\mathcal{F}\varphi \in \mathcal{S}$, φ è continua e limitata, per cui sono verificate le condizioni per la validità della formula di inversione:

$$\begin{aligned} \bar{\mathcal{F}}\mathcal{F}\varphi &= \varphi, & \forall \varphi \in \mathcal{S}. \\ \mathcal{F}\bar{\mathcal{F}}\varphi &= \varphi, \end{aligned}$$

Questo comporta che $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ risulta iniettiva e suriettiva. L'iniettività deriva da $\mathcal{F}\mathcal{F}\varphi = \varphi$:

$$\mathcal{F}\varphi = 0 \implies \varphi = \bar{\mathcal{F}}(\mathcal{F}\varphi) = \bar{\mathcal{F}}(0) = 0.$$

La suriettività deriva da $\mathcal{F}\bar{\mathcal{F}}\varphi = \varphi$. Infatti, questa relazione ci dice che φ risulta l'immagine, secondo \mathcal{F} di $\bar{\mathcal{F}}\varphi \in \mathcal{S}$.

Abbiamo visto prima una importante connessione tra la trasformazione di Fourier ed il prodotto di convoluzione. Se f, g sono funzioni sommabili, vale la relazione:

$$\mathcal{F}(f * g) = (2\pi)^{n/2} (\mathcal{F}f) (\mathcal{F}g).$$

Se consideriamo funzioni nell'insieme \mathcal{S} , ovviamente tale relazione continua a valere in quanto $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$. La relazione precedente ci dice anche che la trasformata di

Fourier del prodotto di convoluzione di due funzioni a decrescenza rapida è ancora a decrescenza rapida. Naturalmente tutte le relazioni trovate valgono anche per la trasformata coniugata (cambiando ξ in $-\xi$) per cui abbiamo anche:

$$\bar{\mathcal{F}}(f * g) = (2\pi)^{n/2} (\bar{\mathcal{F}}f)(\bar{\mathcal{F}}g).$$

Usando ora la formula di inversione (sicuramente valida per funzioni di \mathcal{S}) e la suriettività in \mathcal{S} ,abbiamo anche:

$$\begin{aligned} (\mathcal{F}f) * (\mathcal{F}g) &= \mathcal{F}\bar{\mathcal{F}}((\mathcal{F}f) * (\mathcal{F}g)) \\ &= (2\pi)^{n/2} \mathcal{F}((\bar{\mathcal{F}}\mathcal{F}f)(\bar{\mathcal{F}}\mathcal{F}g)) \\ &= (2\pi)^{n/2} \mathcal{F}(fg), \\ \mathcal{F}(fg) &= \frac{1}{(2\pi)^{n/2}} (\mathcal{F}f) * (\mathcal{F}g), \end{aligned}$$

e, per simmetria:

$$\bar{\mathcal{F}}(fg) = \frac{1}{(2\pi)^{n/2}} (\bar{\mathcal{F}}f) * (\bar{\mathcal{F}}g).$$

Possiamo riassumere tutti i risultati sin qui ottenuti per la trasformata di Fourier limitandoci al caso di funzioni in \mathcal{S} , anche se molti risultati sono validi in contesti più generali.

Teo. 5.21 La trasformata di Fourier trasforma funzioni a decrescenza rapida in funzioni a decrescenza rapida:

$$\mathcal{F} : \mathcal{S} \longrightarrow \mathcal{S}, \quad (5.67)$$

e gode delle seguenti proprietà.

i) Costituisce una trasformazione lineare ed in particolare si ha:

$$D^\alpha(\mathcal{F}\varphi) = \mathcal{F}((-i)x)^\alpha \varphi, \quad (5.68)$$

$$x^\alpha(\mathcal{F}\varphi) = \mathcal{F}((-iD)^\alpha \varphi), \quad (5.69)$$

per ogni multiindice α e $\varphi \in \mathcal{S}$.

ii) La trasformata di Fourier risulta 1 – 1 e suriettiva su \mathcal{S} , quindi invertibile e con l'inversa data dalla trasformata coniugata o antitrasformata $\bar{\mathcal{F}}$:

$$\bar{\mathcal{F}}\mathcal{F}\varphi = \varphi, \quad (5.70)$$

$$\mathcal{F}\bar{\mathcal{F}}\varphi = \varphi, \quad (5.71)$$

per ogni $\varphi \in \mathcal{S}$.

iii) Il prodotto di convoluzione di due funzioni a decrescenza rapida è ancora a decrescenza rapida e si ha:

$$\mathcal{F}(\varphi * \psi) = (2\pi)^{n/2} (\mathcal{F}\varphi)(\mathcal{F}\psi), \quad (5.72)$$

$$\overline{\mathcal{F}}(\varphi * \psi) = (2\pi)^{n/2} (\overline{\mathcal{F}}\varphi)(\overline{\mathcal{F}}\psi), \quad (5.73)$$

$$\mathcal{F}(\varphi\psi) = \frac{1}{(2\pi)^{n/2}} (\mathcal{F}\varphi) * (\mathcal{F}\psi), \quad (5.74)$$

$$\overline{\mathcal{F}}(\varphi\psi) = \frac{1}{(2\pi)^{n/2}} (\overline{\mathcal{F}}\varphi) * (\overline{\mathcal{F}}\psi), \quad (5.75)$$

per ogni $\varphi, \psi \in \mathcal{S}$.

Osservazione. Abbiamo detto che le proprietà di \mathcal{F} e $\overline{\mathcal{F}}$ sono le stesse. Possiamo vedere inoltre che applicare due volte la trasformata di Fourier (ad esempio in \mathcal{S}) riporta alla funzione originaria ma riflessa:

$$(\mathcal{F}\mathcal{F}\varphi)(\xi) = (\overline{\mathcal{F}}\mathcal{F}\varphi)(-\xi) = \varphi(-\xi).$$

5.4.4 Trasformate di Fourier in L^2 .

Abbiamo visto che se $f, g \in L^1(\mathbb{R}^n)$ allora è valida la relazione:

$$\int_{\mathbb{R}^n} (\mathcal{F}f)(\xi) g(\xi) d\xi = \int_{\mathbb{R}^n} f(\xi) (\mathcal{F}g)(\xi) d\xi,$$

e, se $\varphi, \psi \in \mathcal{S}$ possiamo usare la suriettività della trasformata di Fourier e scrivere:

$$\int_{\mathbb{R}^n} \overline{(\mathcal{F}\varphi)(\xi)} (\mathcal{F}\psi)(\xi) d\xi = \int_{\mathbb{R}^n} (\mathcal{F}\overline{(\mathcal{F}\varphi)})(\xi) \psi(\xi) d\xi.$$

D'altra parte il complesso coniugato della trasformata di Fourier è direttamente valutabile come antitrasformata della funzione complessa coniugata:

$$\overline{(\mathcal{F}\varphi)(\xi)} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\xi \cdot x} \overline{\varphi(x)} dx = (\overline{\mathcal{F}}\varphi)(\xi),$$

e di conseguenza si ottiene la relazione, detta di Plancherel⁶

$$\int_{\mathbb{R}^n} \overline{(\mathcal{F}\varphi)(\xi)} (\mathcal{F}\psi)(\xi) d\xi = \int_{\mathbb{R}^n} \overline{\varphi(x)} \psi(x) dx,$$

che esprime la conservazione del prodotto scalare in $L^2(\mathbb{R}^n)$. Sappiamo inoltre che:

$$\mathcal{C}_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n),$$

per cui possiamo vedere \mathcal{S} come sottoinsieme denso di $L^2(\mathbb{R}^n)$ ($\mathcal{C}_0^\infty(\mathbb{R}^n)$ è un sottoinsieme denso in $L^2(\mathbb{R}^n)$). La trasformata di Fourier \mathcal{F} conserva il prodotto scalare e la norma in $L^2(\mathbb{R}^n)$:

$$\langle \mathcal{F}\varphi, \mathcal{F}\psi \rangle = \langle \varphi, \psi \rangle,$$

$$\|\mathcal{F}\varphi\|_2 = \|\varphi\|_2,$$

per cui risulta continua secondo la topologia di $L^2(\mathbb{R}^n)$. Essendo la trasformazione lineare \mathcal{F} continua in un sottoinsieme denso, può essere estesa con continuità a tutto lo spazio di Hilbert $L^2(\mathbb{R}^n)$ mantenendo la stessa norma operatoriale:

$$\mathcal{F} : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n),$$

e risulta unitaria. Per mostrare quest'ultima affermazione occorre mostrare che l'estensione a tutto L^2 conserva il prodotto scalare ed è suriettiva. Se $f, g \in L^2(\mathbb{R}^n)$ possiamo approssimarle con successioni di elementi di \mathcal{S} :

$$\varphi_j \xrightarrow[j \rightarrow \infty]{L^2} f, \quad \psi_j \xrightarrow[j \rightarrow \infty]{L^2} g, \quad \varphi_j, \psi_j \in \mathcal{S},$$

e, per definizione di estensione continua:

$$\mathcal{F}\varphi_j \xrightarrow[j \rightarrow \infty]{L^2} \mathcal{F}f, \quad \mathcal{F}\psi_j \xrightarrow[j \rightarrow \infty]{L^2} \mathcal{F}g.$$

Inoltre:

$$\langle \mathcal{F}\varphi_j, \mathcal{F}\psi_j \rangle = \langle \varphi_j, \psi_j \rangle,$$

⁶ Michel Plancherel (Bussy, Fribourg, Svizzera, 16 Gennaio 1885 – Zurigo, 4 Marzo 1967) è stato un matematico svizzero. Si occupò di analisi matematica, fisica matematica, algebra ed è noto per il teorema di Plancherel nell'analisi armonica.



e la continuità del prodotto scalare comporta la conservazione del prodotto scalare:

$$\langle \mathcal{F}f, \mathcal{F}g \rangle = \langle f, g \rangle, \quad f, g \in L^2(\mathbb{R}^n). \quad (5.76)$$

La suriettività è conseguenza della suriettività in \mathcal{S} . Se $f \in L^2(\mathbb{R}^n)$:

$$f = \lim_{j \rightarrow \infty} \varphi_j, \quad \varphi_j \in \mathcal{S},$$

$$\varphi_j = \mathcal{F}\psi_j, \quad \psi_j \in \mathcal{S}.$$

le funzioni ψ_j formano una successione di Cauchy:

$$\|\psi_k - \psi_l\|_2 = \|\mathcal{F}\psi_k - \mathcal{F}\psi_l\|_2 = \|\varphi_k - \varphi_l\|_2 \xrightarrow[k,l \rightarrow \infty]{} 0,$$

quindi convergenti ad un elemento $g \in L^2(\mathbb{R}^n)$, e per continuità:

$$\mathcal{F}g = \lim_{j \rightarrow \infty} \mathcal{F}\psi_j = \lim_{j \rightarrow \infty} \varphi_j = f.$$

Possiamo quindi concludere affermando che la trasformata di Fourier \mathcal{F} costituisce una trasformazione unitaria da $L^2(\mathbb{R}^n)$ a $L^2(\mathbb{R}^n)$:

$$\mathcal{F} : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n),$$

con:

$$\mathcal{F}^\dagger = \mathcal{F}^{-1} = \overline{\mathcal{F}}. \quad (5.77)$$

Osservazione. La definizione di trasformata di Fourier tramite l'integrale (5.50) non è più valida in generale in tutto $L^2(\mathbb{R}^n)$, ma dobbiamo ricorrere alla definizione come estensione continua da \mathcal{S} a $L^2(\mathbb{R}^n)$. Lo stesso discorso vale per la sua inversa $\overline{\mathcal{F}}$ definita anch'essa in generale come estensione continua della antitrasformata di funzioni in \mathcal{S} .

Appendice: il teorema del sottografico.

Vediamo ora, per completezza, un complemento di teoria della misura, il cosiddetto teorema del sottografico.

Teo. 5.22 Sia $1 \leq p < \infty$. Le combinazioni lineari di funzioni caratteristiche di poliintervalli sono dense in $L^p(\mathbb{R}^n)$ (dove una funzione caratteristica è considerata, come elemento di $L^p(\mathbb{R}^n)$, assieme a tutta la sua classe di equivalenza).

Dim. 5.22 La verifica di tale affermazione avviene in due passi. Dapprima si vede che ogni elemento di $L^p(\mathbb{R}^n)$ è approssimabile mediante combinazioni lineari di funzioni caratteristiche di insiemi misurabili con misura finita. In seguito ogni funzione caratteristica è approssimabile con combinazioni lineari di funzioni caratteristiche di poliintervalli.

Vediamo la prima parte. Sia $u \in L^p(\mathbb{R}^n)$, u è in generale una funzione complessa, ma possiamo limitarci a considerare u reale in quanto possiamo sempre decomporre una funzione complessa nelle sue parti reale ed immaginaria, e approssimare separatamente le due funzioni reali così ottenute. Una funzione reale u a sua volta può essere decomposta nelle sue parti positiva e negativa: $u = u_+ - u_-$,

$$u_+(x) = \begin{cases} u(x) & \text{se } u(x) \geq 0 \\ 0 & \text{se } u(x) \leq 0 \end{cases}$$

$$u_-(x) = \begin{cases} 0 & \text{se } u(x) \geq 0 \\ -u(x) & \text{se } u(x) \leq 0 \end{cases}$$

per cui possiamo considerare solo funzioni positive. Supponiamo quindi:

$$u(x) \geq 0 \quad \forall x \in \mathbb{R}^n.$$

Per ogni intero k dividiamo l'intervallo di reali positivi $[0, k]$ in intervalli di lunghezza $1/2^k$:

$$\left[\frac{j}{2^k}, \frac{j+1}{2^k} \right], \quad j = 0, 1, \dots, k 2^k - 1,$$

e definiamo:

$$u_k(x) = \begin{cases} 0 & \text{se } u(x) \geq k, \\ \frac{j}{2^k} & \text{se } \frac{j}{2^k} \leq u(x) < \frac{j+1}{2^k} \end{cases}$$

In ogni caso abbiamo:

$$u_k(x) \leq u(x) \quad \forall x \in \mathbb{R}^n,$$

$$u(x) - u_k(x) \leq \frac{1}{2^k} \quad \text{se } u(x) < k,$$

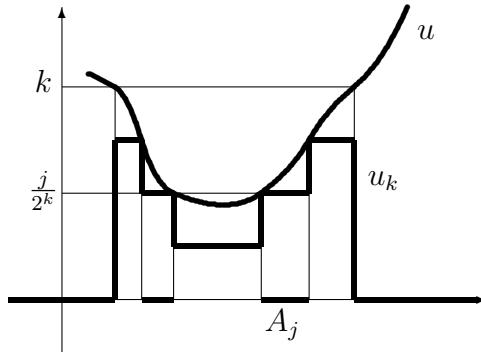


Figura 5.2: Teorema del sottografico.

per cui abbiamo una buona approssimazione nelle regioni in cui u è finita. Chiaramente u_k risulta una combinazione lineare (finita) e a coefficienti positivi di funzioni caratteristiche di insiemi misurabili A_j :

$$u_k(x) = \sum_{j=0}^{k2^k-1} \frac{j}{2^k} \chi_{A_j}(x),$$

$$A_j = \left\{ x \in \mathbb{R}^n ; \frac{j}{2^k} \leq u(x) < \frac{j+1}{2^k} \right\},$$

Questi insiemi (disgiunti) devono essere a misura finita in quanto:

$$\begin{aligned} \int_{\mathbb{R}^n} u_k(x)^p dx &= \int_{\cup_j A_j} u_k(x)^p dx = \sum_j \int_{A_j} u_k(x)^p dx \\ &= \sum_j \left(\frac{j}{2^k} \right)^p \mu(A_j) \leq \int_{\mathbb{R}^n} u(x)^p dx < \infty. \end{aligned}$$

Siccome $u \in L^p(\mathbb{R}^n)$ l'insieme:

$$\mathcal{N}_\infty = \{ x \in \mathbb{R}^n ; u(x) = \infty \}$$

risulta a misura nulla, e per ogni $x \notin \mathcal{N}_\infty$ $u(x)$ risulta finita e approssimabile con $u_k(x)$:

$$\forall x \notin \mathcal{N}_\infty \quad u_k(x) \xrightarrow[k \rightarrow \infty]{} u(x),$$

Inoltre abbiamo la maggiorazione uniforme e sommabile:

$$|u(x) - u_k(x)|^p \leq 2^p |u(x)|^p,$$

per cui possiamo applicare il teorema di Lebesgue:

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} |u(x) - u_k(x)|^p dx = 0,$$

cioè $u_k \rightarrow u$ nella topologia di $L^p(\mathbb{R}^n)$.

Vediamo ora il secondo passo. Sia χ_A la funzione caratteristica di un insieme misurabile e a misura finita A . Allora A può essere ricoperto mediante poliintervalli E_j (formando un ricoprimento numerabile), che possiamo assumere disgiunti:

$$A \subseteq \bigcup_j E_j, \quad E_j \cap E_k = \emptyset,$$

e la misura di A può essere vista come l'estremo inferiore della misura di tali poliintervalli (al variare dei poliintervalli che forniscono un ricoprimento di A):

$$\mu(A) = \inf \sum_j \mu(E_j).$$

Allora, per definizione di estremo inferiore esiste un ricoprimento $\{E_j\}$ tale che, dato $\epsilon > 0$:

$$\sum_j \mu(E_j) \geq \mu(A) > \sum_j \mu(E_j) - \epsilon,$$

per cui:

$$0 \leq \int_{\mathbb{R}^n} (\chi_{\cup_j E_j}(x) - \chi_A(x)) dx \leq \epsilon$$

Possiamo considerare un sottoinsieme finito di tali poliintervalli in quanto, sempre applicando il teorema di Lebesgue:

$$\int_{\mathbb{R}^n} \left| \chi_{\cup_j E_j}(x) - \sum_{j=1}^N \chi_{E_j}(x) \right| dx \xrightarrow[N \rightarrow \infty]{} 0.$$

per cui esiste \bar{N} tale che per $N \geq \bar{N}$:

$$\int_{\mathbb{R}^n} \left| \chi_{\cup_j E_j}(x) - \sum_{j=1}^N \chi_{E_j}(x) \right| dx \leq \epsilon,$$

e:

$$\int_{\mathbb{R}^n} \left| \chi_A(x) - \sum_{j=1}^N \chi_{E_j}(x) \right| dx \leq 2\epsilon.$$

È ora facile ottenere la convergenza in $L^p(\mathbb{R}^n)$ in quanto, assumendo le funzioni caratteristiche solo i valori 0 e 1:

$$\int_{\mathbb{R}^n} \left| \chi_A(x) - \chi_{\cup_{j=1}^N E_j}(x) \right|^p dx = \int_{\mathbb{R}^n} \left| \chi_A(x) - \chi_{\cup_{j=1}^N E_j}(x) \right| dx.$$

5.5 Esercizi

Esercizio 5.1 Sia $X = C([a, b])$ lo spazio di Banach delle funzioni continue sull'intervallo (chiuso) $[a, b]$ con la norma:

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

Su X si definisca l'operatore lineare B tale che, per ogni $f \in X$:

$$B : f(x) \mapsto (Bf)(x) = \int_a^b H(x, y) f(y) dy,$$

essendo $H(x, y) \in \mathcal{C}([a, b] \times [a, b])$.

- a) Si dimostri che B è continuo.
- b) Si calcoli e^{tB} nel caso in cui $H(x, y) = 1$.

— ○ —

Esercizio 5.2 Nello spazio di Hilbert $L^2([0, 1])$ si consideri l'operatore tale che, per ogni $f \in L^2([0, 1])$:

$$A : f(x) \longrightarrow g(x) = \int_0^x dt f(t).$$

- a) Si dimostri che A è limitato e si stimi la sua norma.
- b) Si determini A^\dagger e si calcoli $A + A^\dagger$.

— ○ —

Esercizio 5.3 In $L^2(\mathbb{R})$ si consideri per $a \in \mathbb{R}$, $a \neq 0$, l'operatore:

$$S_a : f(x) \longrightarrow f(ax), \quad f \in L^2(\mathbb{R}).$$

- a) Si dimostri che S_a è limitato e se ne calcoli la norma.
Si calcoli inoltre S_a^\dagger e si verifichi che S_a è normale, cioè commuta col suo aggiunto:

$$S_a S_a^\dagger = S_a^\dagger S_a.$$

- b) Si ponga ora, $\forall f \in L^2(\mathbb{R})$:

$$f_+(x) = \frac{1}{2}[f(x) + f(-x)],$$

$$f_-(x) = \frac{1}{2}[f(x) - f(-x)].$$

Si verifichi che $f_+(x)$ ($f_-(x)$) appartiene al sottospazio V_+ (V_-) delle funzioni pari (dispari) e che tali due sottospazi sono ortogonali tra loro, sicché $L^2(\mathbb{R}) = V_+ \oplus V_-$.

Si ponga quindi $a = -1$ e si dimostri che V_+ e V_- sono autospazi dell'operatore S_{-1} , determinandone i relativi autovalori.

— ○ —

Esercizio 5.4 Nello spazio di Hilbert:

$$\ell^2 = \{c = (c_0, c_1, c_2, \dots) : \sum_{j=0}^{\infty} |c_j|^2 < \infty\},$$

si consideri l'operatore:

$$A : (c_0, c_1, c_2, \dots) \longrightarrow (c_1, \sqrt{2}c_2, \sqrt{3}c_3, \dots).$$

- a) Si determini il dominio \mathcal{D} di A .
- b) Si determini l'operatore aggiunto di A .
- c) Si determinino autovalori ed autovettori degli operatori $B = A^\dagger A$ e $C = AA^\dagger$.

— ○ —

Esercizio 5.5 Si consideri in $L^2([0, 1])$ l'operatore A definito da:

$$(Af)(t) = tf(t) + \int_0^1 f(\tau) d\tau.$$

- a) Mostrare che l'operatore A è limitato.
- b) Determinare l'aggiunto A^\dagger di A .
- c) Determinare gli autovalori λ e le corrispondenti autofunzioni $f_\lambda(t)$ tali che:

$$\int_0^1 f(t) dt = 1.$$

— ○ —

Esercizio 5.6 In uno spazio di Hilbert H si consideri l'operatore A definito dalla relazione:

$$Ax = \alpha \langle u, x \rangle u + \beta \langle v, x \rangle v,$$

con $\alpha, \beta \in \mathbb{C}$ e u, v vettori ortonormali in H .

- a) Verificare che l'operatore A è limitato e calcolarne la norma.
- b) Calcolare esplicitamente l'operatore e^A .

— ○ —

Esercizio 5.7 Nello spazio di Hilbert $L^2(0, 2\pi)$ si consideri l'applicazione lineare A definita da:

$$(Af)(x) = \frac{1}{\pi} \int_0^{2\pi} \cos(x - y) f(y) dy$$

a) Mostrare che l'operatore A è limitato e calcolarne la norma.

b) Determinare per quali valori di $z \in \mathbb{C}$ la serie:

$$\sum_{n=0}^{\infty} z^n A^n$$

è convergente in norma.

— ○ —

Esercizio 5.8 Si calcoli, usando i teoremi dell'analisi complessa, la trasformata di Fourier della funzione:

$$f(x) = \frac{1}{(1 - ix)^2}.$$

— ○ —

Esercizio 5.9 Sia f la funzione di variabile reale:

$$f(x) = \begin{cases} \sin \omega x & |x| < T \\ 0 & |x| > T \end{cases}$$

con $T > 0$.

a) Mostrare che $f \in L^2(\mathbb{R})$.

b) Calcolare la trasformata di Fourier di f .

— ○ —

5.5.1 Soluzioni

Soluzione 5.1

a) Innanzitutto notiamo che B è ben definito poichè, se $H(x, y) \in \mathcal{C}([a, b] \times [a, b])$ allora anche $(Bf)(x)$ è continua su $[a, b]$ quando f lo è. Inoltre si ha:

$$\begin{aligned} |(Bf)(x)| &= \left| \int_a^b H(x, y) f(y) dy \right| \leq \int_a^b |H(x, y) f(y)| dy \leq \|f\| \int_a^b |H(x, y)| dy \\ &\leq (b-a) \sup_{[a,b] \times [a,b]} |H(x, y)| \|f\| = C_H \|f\| \end{aligned}$$

Per cui:

$$\|Bf\| = \sup_{[a,b]} |(Bf)(x)| \leq C_H \|f\|.$$

b) In questo caso si ha, per ogni $f \in X$:

$$\begin{aligned} (Bf)(x) &= \int_a^b f(y) dy \\ (B^2 f)(x) &= \int_a^b dz \int_a^b f(y) dy = (b-a)(Bf)(x) \\ &\vdots \\ (B^n f)(x) &= \int_a^b dz (B^{n-1} f)(x) = (b-a)^{n-1} (Bf)(x) \end{aligned}$$

da cui

$$\begin{aligned} e^{tB}(f) &= \sum_{n=0}^{\infty} \frac{(tB)^n f}{n!} = f + t(Bf) + \frac{1}{2!} t^2 (b-a)(Bf) + \frac{1}{3!} t^3 (b-a)^2 (Bf) + \dots \\ &= f + \frac{1}{b-a} \left[t(b-a) + \frac{1}{2!} t^2 (b-a)^2 + \frac{1}{3!} t^3 (b-a)^3 + \dots \right] Bf \\ &= f + \frac{e^{(b-a)t} - 1}{b-a} Bf \end{aligned}$$

ovvero

$$e^{tB} = 1 + \frac{e^{(b-a)t} - 1}{b-a} B$$

— ○ —

Soluzione 5.2

a) Si ha:

$$\begin{aligned} |g(x)| &= \left| \int_0^x dt f(t) \right| \leq \int_0^x dt |f(t)| = \langle 1, |f| \rangle \\ &\leq \|1\| \|f\| = \left(\int_0^1 dt |1|^2 \right)^{1/2} \left(\int_0^1 dt |f|^2 \right)^{1/2} = \|f\|, \end{aligned}$$

essendo 1 la funzione identicamente uguale ad uno su tutto l'intervallo. Ne segue che:

$$\|Af\|^2 = \|g\|^2 = \int_0^1 dt |g(t)|^2 \leq \int_0^1 dt \|f\|^2 = \|f\|^2,$$

da cui $\|A\| \leq 1$.

b) Si ha, per ogni $f, g \in L^2([0, 1])$:

$$\begin{aligned} \langle A^\dagger f, g \rangle &= \langle f, Ag \rangle = \int_0^1 dx \overline{f(x)} \left(\int_0^x dt g(t) \right) \\ &= \int_0^1 dx \int_0^x dt \overline{f(x)} g(t) = \int_0^1 dt \int_t^1 dx \overline{f(x)} g(t) = \int_0^1 dt \overline{\left(\int_t^1 dx f(x) \right)} g(t), \end{aligned}$$

da cui:

$$A^\dagger : f(t) \longrightarrow h(t) = \int_t^1 dx f(x)$$

Ne segue anche che:

$$A + A^\dagger : f(x) \longrightarrow \int_0^t dx f(x) + \int_t^1 dx f(x) = \int_0^1 dx f(x).$$

— ○ —

Soluzione 5.3

a) Se $f(x) \in L^2(\mathbb{R})$ si ha, operando il cambio di variabile $y = ax$:

$$\|S_a(f)\|^2 = \int_{-\infty}^{\infty} |f(ax)|^2 dx = \frac{1}{|a|} \int_{-\infty}^{\infty} |f(y)|^2 dy = \frac{1}{|a|} \|f\|^2,$$

e quindi S_a è un operatore limitato. Inoltre:

$$\frac{\|S_a(f)\|^2}{\|f\|^2} = \frac{1}{|a|}$$

e quindi

$$\|S_a\| = \sqrt{\frac{1}{|a|}}$$

Poichè S_a è limitato, S_a^\dagger è definito ovunque ed è anch'esso un operatore limitato. È definito dalla relazione, per $f, g \in L^2(\mathbb{R})$:

$$\langle S_a^\dagger f, g \rangle = \langle f, S_a g \rangle = \int_{-\infty}^{\infty} \overline{f(x)} g(ax) dx = \frac{1}{|a|} \int_{-\infty}^{\infty} \overline{f(y/a)} g(y) dy$$

da cui:

$$S_a^\dagger : f(x) \longrightarrow \frac{1}{|a|} f(x/a).$$

Inoltre si ha:

$$\begin{aligned} (S_a S_a^\dagger f)(x) &= S_a \left[\frac{1}{|a|} f(x/a) \right] = \frac{1}{|a|} f(x), \\ (S_a^\dagger S_a f)(x) &= S_a^\dagger [f(ax)] = \frac{1}{|a|} f(x), \\ S_a S_a^\dagger &= S_a^\dagger S_a, \end{aligned}$$

e l'operatore è normale.

b) $\forall f \in L^2(\mathbb{R})$ si ha:

$$\begin{aligned} f_+(-x) &= \frac{1}{2}[f(-x) + f(x)] = f_+(x), \\ f_-(-x) &= \frac{1}{2}[f(-x) - f(x)] = -f_-(x), \end{aligned}$$

ovvero $f_+(x)$ ($f_-(x)$) appartiene al sottospazio V_+ (V_-) delle funzioni pari (dispari). Tali due sottospazi sono ortogonali tra loro, poichè se $f_+ \in V_+$ e $f_- \in V_-$, il loro prodotto scalare è nullo:

$$\langle f_+, f_- \rangle = \int_{-\infty}^{\infty} \overline{f_+(x)} f_-(x) dx = 0$$

essendo $\overline{f_+(x)} f_-(x)$ una funzione dispari. Ora, ogni $f \in L^2(\mathbb{R})$ puo' essere scritta come:

$$f(x) = f_+(x) + f_-(x)$$

essendo f_\pm definite come sopra e quindi:

$$L^2(\mathbb{R}) = V_+ \oplus V_-.$$

È infine immediato verificare che V_+ e V_- sono autospazi dell'operatore S_{-1} (detto operatore parità), con autovalori $\lambda = 1$ e $\lambda = -1$ rispettivamente:

$$\begin{aligned} S_{-1}[f_+(x)] &= f_+(-x) = f_+(x), \\ S_{-1}[f_-(x)] &= f_-(-x) = -f_-(x). \end{aligned}$$

— ○ —

Soluzione 5.4

a) Un vettore $c \in \ell^2$ è nel dominio di A se e solo se $Ac \in \ell^2$. Si ha:

$$\|Ac\|^2 = \sum_{j=1}^{\infty} j|c_j|^2 < \infty \iff |c_j|^2 \sim \frac{1}{j^{2(1+\epsilon)}}, \quad \epsilon > 0, \text{ per grandi } j,$$

da cui

$$\begin{aligned} \mathcal{D}(A) &= \left\{ c = (c_0, c_1, c_2, \dots) \in \ell^2 : \sum_{j=0}^{\infty} j|c_j|^2 < \infty \right\} \\ &= \left\{ c = (c_0, c_1, c_2, \dots) \in \ell^2 : |c_j| \sim \frac{1}{j^{1+\epsilon}}, \quad \epsilon > 0, \text{ per grandi } j \right\} \end{aligned}$$

b) Sia $c = (c_0, c_1, c_2, \dots) \in \mathcal{D}(A)$ e $b = (b_0, b_1, b_2, \dots) \in \mathcal{D}(A^\dagger)$, così che:

$$\langle A^\dagger b, c \rangle = \langle b, Ac \rangle = \sum_{j=0}^{\infty} \overline{b_j} \sqrt{j+1} c_{j+1} = \sum_{k=1}^{\infty} \overline{\sqrt{k} b_{k-1}} c_k,$$

da cui:

$$(A^\dagger b)_k = \begin{cases} 0 & k = 0, \\ \sqrt{k} b_{k-1} & k = 1, 2, \dots, \end{cases}$$

$$A^\dagger (b_0, b_1, b_2, \dots) \longrightarrow (0, b_0, \sqrt{2}b_1, \dots).$$

Inoltre si ha $\mathcal{D}(A^\dagger) = \mathcal{D}(A)$.

c) Si ha:

$$B : (c_0, c_1, c_2, \dots) \longrightarrow (0, c_1, 2c_2, \dots),$$

$$C : (c_0, c_1, c_2, \dots) \longrightarrow (c_0, 2c_1, 3c_2, \dots),$$

B ha quindi autovalori: $\lambda = 0$ con autovettore $(1, 0, 0, \dots)$ e $\lambda = k$ ($k = 1, 2, \dots$) con autovettore $(0, 0, \dots, 0, 1, 0, \dots)$, dove 1 è al posto k -mo. C ha invece autovalori: $\lambda = k$ ($k = 1, 2, \dots$) con autovettore $(0, 0, \dots, 0, 1, 0, \dots)$, dove 1 è al posto $(k-1)$ -mo.

— ○ —

Soluzione 5.5

a) Si può scrivere:

$$(Af)(t) = tf + \langle 1, f \rangle \implies$$

$$\|Af\| = \|tf + \langle 1, f \rangle\| \leq \|tf\| + |\langle 1, f \rangle| \|1\| \leq \|tf\| + \|1\|^2 \|f\|$$

ed abbiamo:

$$\begin{aligned}\|t f\|^2 &= \int_0^1 t^2 |f(t)|^2 dt \leq \int_0^1 |f(t)|^2 dt = \|f\|^2, \\ \|1\|^2 &= \int_0^1 1 dt = 1.\end{aligned}$$

Pertanto:

$$\begin{aligned}\|A f\| &\leq 2 \|f\| \quad \Rightarrow \\ \|A\| &\leq 2.\end{aligned}$$

b) L'aggiunto A^\dagger è definito dalla relazione:

$$\langle f, A g \rangle = \langle A^\dagger f, g \rangle \quad \forall f, g \in L^2[0, 1],$$

ma:

$$\begin{aligned}\langle f, A g \rangle &= \langle f, t g + \langle 1, g \rangle \rangle = \langle f, t g \rangle + \langle 1, g \rangle \langle f, 1 \rangle \\ &= \langle t f, g \rangle + \overline{\langle 1, f \rangle} \langle 1, g \rangle = \langle t f, g \rangle + \langle \langle 1, f \rangle, g \rangle \\ &= \langle t f + \langle 1, f \rangle, g \rangle = \langle A f, g \rangle\end{aligned}$$

cioè l'operatore A è simmetrico ed essendo limitato è autoaggiunto $\Rightarrow A = A^\dagger$.

c) L'equazione agli autovalori è (tenendo conto della condizione di normalizzazione richiesta):

$$\begin{aligned}t f(t) + \int_0^1 f(t) dt &= \lambda f, \quad f \in L^2, f \neq 0, \\ t f + 1 &= \lambda f \quad \Rightarrow \quad f_\lambda(t) = \frac{1}{\lambda - t}\end{aligned}$$

Si vede che $f_\lambda(t) \in L^2(0, 1)$ se $\lambda \notin [0, 1]$. Imponendo effettivamente la condizione:

$$\begin{aligned}\int_0^1 f_\lambda(t) dt &= 1, \\ \int_0^1 \frac{1}{\lambda - t} dt &= \ln \frac{|\lambda|}{|\lambda - 1|} = 1, \\ \frac{\lambda}{\lambda - 1} &= \pm e, \\ \lambda &= \frac{e}{e \mp 1},\end{aligned}$$

e l'unico valore accettabile risulta:

$$\lambda = \frac{e}{e-1} > 1.$$

— ○ —

Soluzione 5.6

a) Si può scrivere $A = \alpha P_1 + \beta P_2$, con:

$$P_1 x = \langle u, x \rangle u, \quad P_2 x = \langle v, x \rangle v,$$

operatori di proiezione ortogonale lungo i vettori u e v . Si ha:

$$\begin{aligned} P_1^2 &= P_1, & P_2^2 &= P_2, \\ P_1 P_2 &= P_2 P_1 = 0, & \|P_1\| &= \|P_2\| = 1. \end{aligned}$$

L'operatore A è quindi limitato, e:

$$\|A\| \leq |\alpha| \|P_1\| + |\beta| \|P_2\| = |\alpha| + |\beta|.$$

Essendo u e v ortogonali tra loro abbiamo che $P_1 x$ e $P_2 x$ sono ortogonali tra loro:

$$\begin{aligned} \|Ax\|^2 &= |\alpha|^2 \|P_1 x\|^2 + |\beta|^2 \|P_2 x\|^2 \leq \max\{|\alpha|^2, |\beta|^2\} (\|P_1 x\|^2 + \|P_2 x\|^2) \\ &= \max\{|\alpha|^2, |\beta|^2\} \|(P_1 + P_2)x\|^2 \leq \max\{|\alpha|^2, |\beta|^2\} \|x\|^2, \end{aligned}$$

Se scegliamo $x = P_1 x$ oppure $x = P_2 x$ si ottiene

$$\begin{aligned} \|AP_1 x\| &= \|\alpha P_1 x\| = |\alpha| \|P_1 x\| = |\alpha| \|x\|, \\ \|AP_2 x\| &= \|\beta P_2 x\| = |\beta| \|P_2 x\| = |\beta| \|x\|, \end{aligned}$$

per cui:

$$\Rightarrow \|A\| = \max\{|\alpha|, |\beta|\}$$

b) Per calcolare e^A si può sfruttare lo sviluppo in serie dell'esponenziale:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = \mathbb{1} + \sum_{n=1}^{\infty} \frac{(\alpha P_1 + \beta P_2)^n}{n!}.$$

Sfruttando le proprietà degli operatori P_1 e P_2 si ha:

$$(\alpha P_1 + \beta P_2)^n = \alpha^n P_1 + \beta^n P_2 \quad n = 1, 2, \dots$$

$$\begin{aligned} e^A &= \mathbb{1} + \sum_{n=1}^{\infty} \frac{\alpha^n}{n!} P_1 + \sum_{n=1}^{\infty} \frac{\beta^n}{n!} P_2 \\ &= \mathbb{1} + (e^\alpha - 1)P_1 + (e^\beta - 1)P_2 \\ &= \mathbb{1} - P_1 - P_2 + e^\alpha P_1 + e^\beta P_2. \end{aligned}$$

— ○ —

Soluzione 5.7

a) Abbiamo:

$$(A f)(x) = \frac{1}{\pi} \int_0^{2\pi} (\cos x \cos y + \sin x \sin y) f(y) dy$$

$$c(x) = \frac{\cos x}{\sqrt{\pi}}, \quad s(x) = \frac{\sin x}{\sqrt{\pi}} \quad \|c\| = \|s\| = 1 \quad \langle c, s \rangle = 0$$

$A f = c \langle c, f \rangle + s \langle s, f \rangle$, proiettore ortogonale su $\mathcal{L}\{c, s\}$

$$\|A f\|^2 = |\langle c, f \rangle|^2 + |\langle s, f \rangle|^2 \leq \|f\|^2, \quad (\text{disuguaglianza di Bessel,})$$

per cui $\|A\| \leq 1$. Inoltre se $f = c$ o $f = s$ (cioè $f \in \mathcal{L}\{c, s\}$), allora

$$\|A f\| = \|f\| \implies \|A\| = \sup \frac{\|A f\|}{\|f\|} = 1$$

per cui $\|A\| = 1$.

b) La serie converge in norma se $\|zA\| \leq 1$ ($|z| \leq 1$). Inoltre, essendo A un proiettore ortogonale:

$$A^2 = A, \quad A^n = A, \quad n = 1, 2, \dots$$

e converge a (in questo caso):

$$\begin{aligned} (1 - zA)^{-1} &= 1 + \sum_{n=1}^{\infty} z^n A = 1 + zA \sum_{n=0}^{\infty} z^n \\ &= 1 + \frac{z}{1-z} A = \frac{1}{1-z} A + (1 - A). \end{aligned}$$

— ○ —

Soluzione 5.8

$$f(x) = \frac{1}{[-i(x+i)]^2} = -\frac{1}{(x+i)^2}.$$

La trasformata di Fourier si può calcolare in senso ordinario applicando il teorema dei Residui e il lemma di Jordan:

$$\tilde{f}(k) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-ikx}}{(x+i)^2} dx,$$

se si considera:

$$F(z) = \frac{e^{-iz}}{(z+i)^2},$$

tale funzione mostra un polo doppio in $z = -i$. Per $k < 0$ il lemma di Jordan è soddisfatto in $\Im z > 0$ ove la funzione è analitica mentre per $k > 0$ il lemma di Jordan è soddisfatto in $\Im z < 0$. Quindi:

$$\operatorname{Res} f(z) \Big|_{z=-i} = \lim_{z \rightarrow -i} \frac{d}{dz} [(z+i)^2 f(z)] = \lim_{z \rightarrow -i} \frac{d}{dz} e^{-ikz} = -ike^{-k},$$

$$\tilde{f}(k) = \begin{cases} 0, & k < 0, \\ -\frac{1}{\sqrt{2\pi}}(-2\pi i) \operatorname{Res} f(z) \Big|_{z=-i}, & k > 0, \end{cases}$$

$$\tilde{f}(k) = \begin{cases} 0 & k < 0 \\ \sqrt{2\pi}ke^{-k} & k \geq 0 \end{cases}$$

— ○ —

Soluzione 5.9 f risulta non nulla solo su un insieme chiuso e limitato e continua su tale insieme, per cui chiaramente $f \in L^2(\mathbb{R})$.

$$\begin{aligned} (\mathcal{F} f)(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx = \frac{1}{\sqrt{2\pi}} \int_{-T}^T \sin(\omega x) e^{-i\xi x} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-T}^T \frac{e^{+i(\omega-\xi)x} - e^{-i(\omega+\xi)x}}{2i} dx \\ &= \frac{1}{2i\sqrt{2\pi}} \frac{1}{i(\omega-\xi)} e^{+i(\omega-\xi)x} \Big|_{x=-T}^{x=T} - \frac{1}{2i\sqrt{2\pi}} \frac{1}{-i(\omega+\xi)} e^{-i(\omega-\xi)x} \Big|_{x=-T}^{x=T} \\ &= -\frac{1}{2\sqrt{2\pi}(\omega-\xi)} (e^{+i(\omega-\xi)T} - e^{-i(\omega-\xi)T}) \\ &\quad - \frac{1}{2\sqrt{2\pi}(\omega+\xi)} (e^{-i(\omega+\xi)T} - e^{+i(\omega+\xi)T}) \\ &= -\frac{1}{2\sqrt{2\pi}} \left\{ \frac{2i \sin(\omega-\xi)T}{\omega-\xi} - \frac{2i \sin(\omega+\xi)T}{\omega+\xi} \right\} \\ &= -\frac{i}{\sqrt{2\pi}} \left\{ \frac{\sin(\omega-\xi)T}{\omega-\xi} - \frac{\sin(\omega+\xi)T}{\omega+\xi} \right\}. \end{aligned}$$

— ○ —

Indice analitico

- addizione, 172
- aggiunto, 338, 341
- algebra, 329
- analiticità, 21
- anello, 172
 - commutativo, 173
 - unitario, 173
- Apollonio, di Perga, 306
- applicazione
 - 1-1, 144
 - antilineare, 187
 - biettiva, 144
 - iniettiva, 144
 - inversa, 189
 - invertibile, 189
 - isometrica, 158
 - limitata, 214
 - lineare, 186
 - non singolare, 189
 - singolare, 189
 - suriettiva, 144, 188
- Argand, Jean-Robert, 2
- argomento, 3
- asse reale, 2
- Baire, 321
- Baire, René-Louis, 321
- Banach, 207
- Banach, Stefan, 207
- base, 178
 - di intorni, 148
 - di una topologia, 122
 - hilbertiana, 283
 - ortonormale, 283
- Bessel, Fredrich Wilhelm, 279
- biezione, 144
- bra, 264
- branch line, 81
- branch point, 81
- cammino, 23
 - chiuso, 23
 - di Jordan, 24
 - differenziabile, 24
 - lunghezza di un, 24
 - regolare, 24
 - rettificabile, 24
- semplice, 24
- campo, 173
- Cauchy, 31, 35, 154
 - formula integrale di, 38
 - rappresentazione integrale di, 35
 - teorema di, 31
- Cauchy, Augustin-Louise, 20
- Cauchy-Goursat, 31
- Cauchy-Hadamard, 330
- Chebyshev, Pafnutij L'vovich, 301
- coefficiente di Fourier, 288, 289
- compatto, 131
- complemento ortogonale, 270
- completamento, 159
- componenti, 180
- condizioni di Cauchy-Riemann, 20
- coniugato, 2
- coniugazione, 8
- connesso, 25
- continuità, 13
- contrazione, 154
- convergenza, 147
 - assoluta, 211
 - forte, 320
 - in norma, 211, 221
 - uniforme, 221
- convesso, 33
- convoluzione, 353
 - tra funzioni, 353
- curva, 23
 - di Jordan, 24
 - regolare, 24
 - rettificabile, 24
- Darboux, Jean Gaston, 30
- De Moivre
 - formula di, 6
- De Moivre, Abraham, 6
- determinazione, 80
 - principale, 80
- differenziale, 18
- dimensione, 178
 - finita, 178
 - infinita, 178
- Dirac, 75
 - delta di, 75
- Dirac, Paul Adrien Maurice, 75

- distanza, 125, 206
- disuguaglianza
 - di Bessel, 280
 - di Darboux, 30
 - di Hölder, 230
 - di Minkowski, 232
 - di Schwarz, 223, 224, 261
 - di Young, 354
 - integrale di Schwarz, 231
 - triangolare, 4
- dominio, 12, 187, 188
- dominio di analiticità, 45
- duale
 - algebrico, 199
 - topologico, 274
- elemento
 - inverso, 171
 - neutro, 171
 - opposto, 173
- equivalenza, 23
- estensione continua, 157
- estensione di un operatore, 327
- estremi, 23
- estremo superiore, 182
- estremo superiore essenziale, 228
- Eulero, 15
- Eulero, Leonhard, 5
- fase, 3
- Fatou, 239
- Fatou, Pierre Joseph Luis, 234
- Fischer, Ernst Sigismund, 276
- Fischer-Riesz, 275
- forma
 - bilineare, 259
 - cartesiana, 1
 - polare, 3, 5
 - sesquilineare, 259
 - definita positiva, 261
 - degenera, 259
 - hermitiana, 259
 - non degenera, 259
 - positiva, 261
- formula
 - di De Moivre, 6
 - di Eulero, 5
 - integrale di Cauchy, 38
- Fourier
 - antitrasformata, 350
 - trasformata, 350
- Fourier, Jean Baptiste Joseph, 288
- Frèchet, Maurice René, 134
- funzionale
 - antilineare, 187
 - coniugato omogeneo, 187
 - lineare, 186
- funzione, 12
 - a decrescenza rapida, 357
 - analitica, 21
 - continua, 13, 140, 141
 - coseno, 14
 - coseno iperbolico, 16
 - delta di Dirac, 75
 - differenziabile, 17
 - esponenziale, 14
 - intera, 45
 - meromorfa, 51
 - monodroma, 12
 - multivoca, 12
 - olomorfa, 18, 19, 45
 - polidroma, 12, 80
 - semplice, 352
 - seno, 14
 - seno iperbolico, 16
 - sommabile, 226
 - spettrale, 39
- Gauss, Carl Friedrich, 2
- Goursat, Eduard Jean-Baptiste, 31
- Gram, Jørgen Pedersen, 287
- Green, 92, 98
 - teorema di, 98
- Green, George, 98
- gruppo, 171
 - abeliano, 172
 - additivo, 172
 - isomorfismi, 172
- Hölder, Otto Ludwig, 230
- Hamilton, Sir William Rowan, 8
- Hausdorff, 133, 147
- Hausdorff, Felix, 132
- Hermite, 303
- Hermite, Charles, 302
- Hilbert, 265
- Hilbert, David, 77
- identità di Apollonio, 306
- identità di polarizzazione, 261
- immaginario, 2
- immagine, 140
- indicatore logaritmico, 63
- insieme

- aperto, 114, 118, 127
- chiuso, 116
- chiusura, 119
- compatto, 131
- convesso, 33
- denso, 137
- derivato, 120
- interno di, 119
- limitato, 130, 182
- linearmente dipendente, 178
- linearmente indipendente, 177
- linearmente ordinato, 182
- ordinato, 182
- quoziente, 195
- sequenzialmente compatto, 149
- integrale
 - al contorno, 27
- intorno, 121
- inversa, 189
- isomorfismo, 172, 189
- Jordan, 34
 - lemma di, 70
 - teorema di, 34
- Jordan, Marie Camille Ennemond, 24
- kernel, 39, 189
- ket, 264
- Kolmogorov, Nikolaevič, 133
- $L^\infty(\Omega)$, 228
- l^∞ , 237
- $L^p(\Omega)$, 227
- l^p , 235
- Laguerre, Edmond Nicolas, 304
- Laurent, Pierre Alphonse, 46
- Lebesgue, Henri Léon, 227
- Legendre, Adrien-Marie, 299
- legge
 - commutativa, 172
 - di composizione, 171
- lemma
 - del grande cerchio, 67
 - del piccolo cerchio, 68
 - di Jordan, 70
 - di Riemann-Lebesgue, 352
 - di Zorn, 181
- limite, 12, 147
 - in norma, 221
 - inf, 238
 - sup, 238
 - superiore, 182
- Liouville, Joseph, 53
- lunghezza, 24
- massimale, 182
- massimo, 182
- maximite, 238
- meromorfa, 51
- metodo di Gram-Schmidt, 285
- metrica, 125
 - discreta, 127
- Minkowski, Hermann, 232
- minimale, 238
- modulo, 3, 8, 174
- molteplicemente connesso, 25
- moltiplicazione, 172
 - per uno scalare, 173
- monodroma
 - funzione, 12
- Morera, 42
- Morera, Giacinto, 42
- multiplo, 174
- multivoca
 - funzione, 12
- Neumann, 335, 336
- Neumann, Carl Gottfried, 317
- norma, 203
 - del prodotto scalare, 265
 - equivalenza, 221
 - operatoriale, 219
 - uniforme, 207
- nucleo, 39, 172, 189
- numero complesso, 1
- olomorfismo, 19
- omeomorfismo, 145
- omomorfismo, 172, 186
- omotetia, 323
- operatore
 - aggiunto, 338, 341
 - autoaggiunto, 342
 - di proiezione, 272
 - idempotente, 346
 - lineare, 188
 - simmetrico, 342
 - unitario, 344
- operazione
 - binaria, 171
- opposto, 172, 173
- ordine di uno zero, 49
- ortogonalità, 268
- ottanioni, 11

- Parseval des Chênes, Marc-Antoine, 283
- parte immaginaria, 2
- parte principale, 50, 74, 76, 90, 94
- parte reale, 2
- Pauli, Wolfgang Ernst, 10
- piano
 - complesso, 2
 - di Argand, 2
 - di Gauss, 2
- Pitagora, 269
- Plancherel, Michel, 366
- polidroma
 - funzione, 12, 80
- polinomi
 - di Chebyshev, 301
 - di Hermite, 302, 303
 - di Laguerre, 304
 - di Legendre, 299
- polo, 50
- polo semplice, 50
- primitiva, 28
- principio
 - del massimo modulo, 52
 - di uniforme limitatezza, 322
- prodotto, 172
 - di convoluzione, 353
- prodotto scalare, 263
 - euclideo, 263
 - hermitiano, 263
- proiettore ortogonale, 272
- proiezione ortogonale, 272
- prolungamento continuo, 157
- proprietà
 - antisimmetrica, 182
 - associativa, 171, 172, 174
 - della media., 52
 - distributiva, 173, 174
 - riflessiva, 182, 195
 - simmetrica, 195
 - transitiva, 182, 195
- punto
 - accumulazione, 120
 - aderenza, 119
 - di diramazione, 80
 - di frontiera, 121
 - esterno, 121
 - fisso, 154
 - interno, 119
 - isolato, 120
 - non isolato, 120
 - regolare, 45
- singolare, 45
- singolare isolato, 45
- quaternioni, 8
- raggio di covergenza, 331
- range, 188
- rango, 188
- rappresentazione
 - di Cauchy, 35
 - integrale, 39
 - spettrale, 39
- reciproco, 9, 173
- regola del parallelogramma, 272
- relazione
 - d'ordine, 181
 - di completezza, 283
 - di dispersione, 78
 - di equivalenza, 195
 - di Parseval, 283
 - di Plancherel, 366
- relazioni di dualità, 116
- residuo, 54, 58
- retroimmagine, 140
- ricoprimento, 131
 - aperto, 131
 - chuso, 131
 - finito, 131
- ridotte, 14
- Riemann, 79
 - falda di, 83
 - superficie di, 79, 83
- Riemann, Georg Friedrich Bernhard, 20
- Riesz, Frigyes, 276
- scalare, 173
- Schmidt, Erhard, 287
- Schwartz, Laurent, 358
- Schwarz, 223, 224, 231, 261
- Schwarz, Karl Hermann Amandus, 223
- seminorma, 204
- semplicemente connesso, 25
- serie, 210
 - assolutamente convergente, 211
 - di Fourier, 288, 289
 - di Laurent, 46
 - di Neumann, 317, 335, 336
 - di potenze], 330
 - esponenziale, 333, 334, 337
 - trigonometrica, 288, 295
- sfera aperta, 127
- sfera chiusa, 130

- singolarità essenziale, 51
- sistema ortonormale, 277
 - completo, 283
- somma, 172
 - diretta, 192, 194
- sottoricoprimento, 131
- sottospazi indipendenti, 193
- sottospazio
 - generato, 184
 - lineare, 177
 - supplementare, 192
 - vettoriale, 177
- spazi
 - isometrici, 158
 - omeomorfi, 145
- spazio
 - completo, 154
 - di Banach, 207
 - di Hilbert, 265
 - di Schwartz, 358
 - euclideo, 264
 - $L^\infty(\Omega)$, 228
 - l^∞ , 237
 - $L^p(\Omega)$, 227
 - l^p , 235
 - lineare, 174
 - lineare topologico, 203
 - metrico, 125
 - normato, 203
 - nullo, 176
 - pre-hilbertiano, 264
 - quoziente, 195
 - separabile, 138
 - separato, 133
 - topologico, 114
 - compatto, 131
 - di Frèchet, 134
 - di Hausdorff, 133
 - vettoriale, 173, 174
 - complesso, 174
 - quoziente, 195
 - reale, 174
- Stokes, 99
- Stokes, George Gabriel, 99
- successione
 - convergente, 147
 - di Cauchy, 154
 - p-sommabile, 235
- taglio, 81
- Taylor, Brook, 43
- teorema
 - dei residui, 58
 - del sottografico, 352, 368
 - di Baire, 321
 - di Cachy, 31
 - di Cauchy–Hadamard, 330
 - di Green, 98
 - di Jordan, 34
 - di Morera, 42
 - di Pitagora, 269
- topologia, 114
 - banale, 115
 - base, 122
 - basi equivalenti, 123
 - discreta, 115
 - generata, 119
 - indotta, 118
 - massimale, 115
 - minimale, 115
 - più debole, 118
 - più fine, 118
 - più forte, 118
 - trasformata di Fourier
 - coniugata, 350
 - trasformazioni
 - di Hilbert, 77
- uniforme continuità, 156
- uniforme limitatezza, 322
- unità, 173
- unità immaginaria, 1
- valore assoluto, 4
- vettore, 173
- vettori ortogonali, 268
- Weierstrass, 146
- Weierstrass, Karl Theodore Wilhelm, 51
- Young, 354
- Young, William Henry, 355
- Zemelo, Ernst Friedrich Ferdinand, 183
- Zermelo, 183
- zero, 49, 172, 173
 - ordine di, 49
- Zorn, 181
- Zorn, Max August, 181