

Filippo Cesi

Rudimenti di analisi infinito dimensionale

Versione 0.3 (Armadillo)

(a planet dust diversion)

Precisazione. Trattandosi di una versione alpha, questa (ehm) “opera” conterrà sicuramente molti errori, imprecisioni, omissioni, dimostrazioni lasciate in sospeso, esercizi incompleti, punti interrogativi, geroglifici o simboli di alfabeti ancora sconosciuti, macchie di sugo, ecc. Il Lettore, nel pieno delle sue facoltà, dichiara di accettare l’opera “così com’è” e dichiara (altresì) che mai e poi mai avrà intenzione di rivalersi sull’Autore per eventuali danni materiali, morali o di altro genere in cui potrà incorrere a causa di informazioni errate, incomplete o anche palesemente malevole contenute nell’opera stessa.

Indice

1	Prerequisiti	13
1.1	Algebra elementare	13
1.1.1	Identità algebriche elementari	13
1.1.2	Divisione di polinomi	14
1.1.3	Risoluzione di equazioni di grado superiore al secondo	14
1.2	Algebra lineare	16
1.2.1	Il metodo della riduzione delle righe	16
1.2.2	Autovalori, autovettori e diagonalizzazione di una matrice	22
1.2.3	Funzioni di matrici	25
1.3	Analisi reale	25
1.3.1	Regola di Leibniz per la derivata di un prodotto	25
1.3.2	Integrali impropri su \mathbb{R}	26
1.3.3	Confronto fra integrali e somme	29
1.3.4	Integrali gaussiani	29
1.3.5	Funzioni continue, uniformemente continue, Lipschitz	29
1.3.6	Formule di Taylor	30
1.3.7	Derivare sotto il segno di integrale	32
2	Spazi metrici	35
2.1	Definizione ed esempi	35
2.2	Convergenza, completezza	42
2.3	Teorema della categoria di Baire	45
2.4	Compattezza	47
2.5	Funzioni continue	50
2.6	Completamento di uno spazio metrico	52
2.6.1	I numeri p -adici	53
3	Spazi vettoriali e di Banach	61
3.1	Definizioni	61
3.2	Esempi	61
3.3	Esempi: Spazi finito dimensionali	65
3.3.1	Le norme $\ \cdot\ _p$	65
3.3.2	Disuguaglianze di Hölder e di Minkowski	67
3.4	Concetti “metrici” negli spazi vettoriali normati	69
3.5	Spazi infinito dimensionali: spazi di successioni	71
3.5.1	Lo spazio delle successioni limitate: ℓ_∞	73
3.5.2	Lo spazio delle successioni convergenti a zero: ℓ_0	76
3.5.3	Gli spazi ℓ_p con $p \in [1, \infty)$	77
3.5.4	Lo spazio delle successioni finite: ℓ_f	79
3.5.5	Problemi	79

3.6	Spazi infinito dimensionali: spazi di funzioni	81
3.6.1	Gli spazi $C[a, b]$, $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, $C_c(\mathbb{R})$	81
3.6.2	Gli spazi $C_p[a, b]$, $C_p(\mathbb{R})$ in cui $p \in [1, \infty)$	83
3.6.3	Spazi di funzioni su \mathbb{R}^n . Multiindici	85
3.6.4	Lo spazio $C_c^\infty(\mathbb{R}^n)$ e lo spazio di Schwartz $\mathcal{S}(\mathbb{R}^n)$	86
3.6.5	Altri spazi vettoriali di funzioni	91
3.6.6	Problemi	91
3.7	Indipendenza lineare	94
3.7.1	Come si dimostra che un insieme di vettori è linearmente indipendente?	95
3.8	Insiemi completi di vettori e basi	96
3.9	Separabilità	100
3.10	Completezza. Spazi di Banach	103
3.10.1	Spazi di successioni	103
3.10.2	Spazi di funzioni	105
3.10.3	Riassunto: strategia per dimostrare la completezza	107
3.10.4	Lo spazio di Schwartz come spazio metrico completo	108
3.11	Bagliori di teoria della misura e spazi L_p	109
3.11.1	La misura di Jordan	110
3.11.2	σ -algebre e la misura di Lebesgue	113
3.11.3	L'integrale di Lebesgue e gli spazi L_p	115
4	Spazi di Hilbert	119
4.1	Definizioni	119
4.1.1	Esempi di spazi euclidei	121
4.1.2	La regola del parallelogramma	122
4.1.3	Problemi	123
4.2	Complemento ortogonale	124
4.3	Il teorema della proiezione	126
4.4	Sistemi ortogonali, completezza, basi	130
4.5	Disuguaglianza di Bessel	133
4.5.1	C'è un unico spazio di Hilbert separabile	140
4.6	Riassunto delle precedenti puntate	141
4.7	Proiezioni ortogonali	141
4.7.1	Problemi	142
4.8	Basi ortonormali negli spazi L_2	145
4.8.1	$L_2(\mathbb{R})$ e i polinomi di Hermite	145
5	Funzionali lineari e distribuzioni	147
5.1	Funzionali lineari continui	147
5.2	Lo spazio duale	154
5.2.1	Lo spazio duale è uno spazio vettoriale normato	154
5.2.2	Identificazione di alcuni spazi duali importanti	157
5.2.3	Lo spazio duale di uno spazio di Hilbert	161
5.3	Convergenza debole	162
6	Distribuzioni	165
6.1	Funzioni continue o differenziabili a tratti	166
6.2	Lo spazio delle distribuzioni	169
6.3	La delta di Dirac	173
6.4	La parte principale di $1/x^n$	177
6.5	Operazioni sulle distribuzioni	181

6.6	Alcune proprietà di $P(1/x^n)$	191
6.7	Problemi svolti	194
6.8	La distribuzione $\delta(b(x))$ (da non confondere con $\delta_0(b!)$)	197
6.8.1	Problemi	200
6.9	Alcune identità “notevoli” fra distribuzioni	202
6.10	Distribuzioni su \mathbb{R}^n	207
6.11	Applicazioni	209
6.11.1	Il potenziale elettrostatico	209
6.11.2	Potenziali ritardati	212
7	Operatori lineari	213
7.1	Definizione ed esempi	213
7.1.1	Esempi	215
7.2	Somme e prodotti di operatori lineari	218
7.3	Operatore inverso	220
7.4	Operatore aggiunto (di Hilbert)	222
7.4.1	Proiettori ortogonali	226
7.5	Spettro	226
7.5.1	Il caso finito dimensionale (\mathbb{C}^n).	226
7.5.2	Il caso generale	227
7.6	Operatori autoaggiunti	228
7.7	Operatori compatti	230
8	Serie di Fourier	235
8.1	$L_2[-\pi, \pi]$ e i polinomi trigonometrici	235
8.1.1	Serie di Fourier sull'intervallo $[a, b]$	237
8.1.2	Serie di Fourier sull'intervallo $[-l, l]$	238
8.1.3	Serie di Fourier con solo coseni o seni	238
8.1.4	Problemi	239
8.1.5	Completezza dei polinomi trigonometrici (schema di dimostrazione)	241
8.1.6	Relazione fra la serie di Fourier di f e quella di f'	242
8.1.7	Serie di Fourier complessa	245
8.2	Convergenza puntuale della serie di Fourier	247
8.2.1	Derivata destra e sinistra di una funzione DAT	248
8.2.2	Il teorema sulla convergenza puntuale	250
8.3	Quale sviluppo scegliere?	255
8.4	Equazione del calore su un intervallo finito	256
8.4.1	Condizioni al bordo di Neumann	256
8.4.2	Condizioni al bordo di Dirichlet	259
8.5	Convergenza uniforme della serie di Fourier	259
9	Trasformata di Fourier	261
9.1	L'idea	261
9.2	Parità, traslazioni, dilatazioni	262
9.3	Definizione e proprietà elementari	263
9.3.1	Funzioni che si incontrano spesso	264
9.3.2	Il metodo grafico	268
9.4	Regolarità e andamento all'infinito	271
9.4.1	Un importante sottospazio invariante	274
9.4.2	Problemi	275
9.5	Formula di inversione e Teorema di Plancherel	279

9.6	Convoluzione	284
9.7	Equazione del calore su un intervallo infinito	288
9.8	Brevi considerazioni probabilistiche	291
9.8.1	La funzione di distribuzione cumulativa	291
9.8.2	Somma di variabili casuali indipendenti	292
9.8.3	Funzione caratteristica	294
9.8.4	Il teorema del limite centrale “classico”	296
9.9	Problemi	297

Elenco dei simboli usati

- ☆ Esercizio non ultra-facilissimo.
- ★ Esercizio non facilissimo.
- ✱ Esercizio che quasi si potrebbe dire difficilino (quasi).
- ✿ Esercizio difficilino (senza quasi).
- ☞ Esercizio svolto (più o meno) nei Rudimenti.
- ⊗ Materiale addizionale (il docente può decidere di ometterlo).

\subset $A \subset B$ significa che A è un sottoinsieme di B (non necessariamente proprio, quindi include il caso $A = B$).

\subsetneq $A \subsetneq B$ significa che A è un sottoinsieme proprio di B .

\sqcup Unione disgiunta. Unione fra 2 insiemi che hanno intersezione vuota.

F^c Complemento dell'insieme F .

\overline{F} Chiusura dell'insieme F .

F° Parte interna di F .

∂F Bordo di F .

V^* Spazio duale dello spazio vettoriale normato V [5.3].

T^* Aggiunto dell'operatore lineare T [7.23].

V^\perp Complemento ortogonale di V .

$[x]$ Parte intera di x , il più grande intero minore o uguale a x .

$\lceil x \rceil$ il più piccolo intero maggiore o uguale a x .

$|x|$ Se $x \in \mathbb{R}_n$ (oppure $x \in \mathbb{C}^n$) $|x| := \|x\|_2 = (|x_1|^2 + \dots + |x_n|^2)^{1/2}$.

$\|x\|_p$ Se $p \in [1, \infty)$, $\|x\|_p := (|x_1|^p + |x_2|^p + \dots)^{1/p}$.

$\|x\|_\infty$ $\sup_k |x_k|$.

$\|f\|_p$ Se $p \in [1, \infty)$, $\|f\|_p := (\int |f(x)|^p dx)^{1/p}$.

$\|f\|_u$ $\sup_x |f(x)|$.

$\mathbb{1}_A$ Funzione indicatrice (o funzione caratteristica) dell'insieme A . Si usa anche il simbolo χ_A .

$\text{Arg } z$ $\arg_{(-\pi, \pi]} z$ è l'argomento principale di z .

- $B_r(x)$ Palla di centro x e raggio r .
- $C(\mathbb{R})$ Insieme delle funzioni continue su \mathbb{R} .
- $C(X)$ Insieme delle funzioni continue su uno spazio metrico X .
- $C_b(X)$ Insieme delle funzioni continue e limitate su uno spazio metrico X .
- $C_0(\mathbb{R})$ Insieme delle funzioni continue su \mathbb{R} che tendono a zero all'infinito.
- $C_c(X)$ Insieme delle funzioni continue su X a supporto compatto.
- $C_p(\mathbb{R})$ ($p \in [1, \infty)$). Insieme delle funzioni continue con norma $\|\cdot\|_p$ finita.
- $C^k(\mathbb{R})$ Insieme delle funzioni reali con k derivate continue.
- $C^\infty(\mathbb{R})$ Insieme delle funzioni reali infinitamente differenziabili.
- $C_c^\infty(\mathbb{R})$ Insieme delle funzioni reali infinitamente differenziabili a supporto compatto.
- \mathbb{C} Il campo dei numeri complessi.
- $\mathbb{C}[x]$ Insieme dei polinomi a coefficienti complessi.
- $\text{cis}(\vartheta)$ $\cos \vartheta + i \sin \vartheta$.
- δ_{jk} Delta di Kroenecker.
- δ_a Delta di Dirac nel punto a (5.5).
- ∂^α Derivata parziale in \mathbb{R}^n rispetto al multiindice α .
- D Derivata.
- φ_g Funzionale lineare corrispondente alla funzione g (5.3).
- $H(x)$ Funzione a gradino di Heaviside [(6.8)].
- \mathcal{K} Spazio delle funzioni fondamentali [6.10].
- \mathcal{K}^* Distribuzioni sullo spazio delle funzioni fondamentali.
- $\xrightarrow{\mathcal{K}}$ Convergenza nello spazio \mathcal{K} [6.11].
- $\xrightarrow{\mathcal{K}^*}$ Convergenza nello spazio \mathcal{K}^* .
- $\text{Ker } A$ Nucleo dell'operatore lineare A .
- ℓ_∞ Insieme delle successioni reali limitate.
- ℓ_0 Insieme delle successioni reali che tendono a zero.
- ℓ_p ($p \in [1, \infty)$). Insieme delle successioni reali con norma $\|\cdot\|_p$ finita.
- ℓ_f Insieme delle successioni reali con un numero finito di elementi non nulli.
- $\ell_f(\mathbb{Q})$ Insieme delle successioni a elementi razionali con un numero finito di elementi non nulli.
- $\mathcal{L}(V, Z)$ Insieme degli operatori lineari limitati da V in Z [7.5].
- $\text{Lip}(\mathbb{R})$ Insieme delle funzioni Lipschitz su \mathbb{R} [1.22].
- $\text{Mat}_{m \times n}(R)$ Insieme delle matrici $m \times n$ a coefficienti nell'anello R .
- $\text{Mat}_n(R)$ Anello delle matrici $n \times n$ coefficienti nell'anello R .
 - \mathbb{N} Insieme dei numeri naturali (che, da che mondo è mondo, partono da zero!).
 - \mathbb{N}^* Insieme dei numeri naturali positivi. $\mathbb{N} \setminus \{0\}$.
 - Pf Operatore di parità applicato alla funzione f . [Sezione 9.2].
 - $P(\mathbb{R})$ Insieme dei polinomi a coefficienti reali. Indicato anche con $\mathbb{R}[x]$.
 - $P(1/x^n)$ Parte principale della distribuzione $1/x^n$.
 - π_W Proiezione ortogonale sul sottospazio chiuso W .

- \mathbb{Q} Il campo dei numeri razionali.
- $R_z(T)$ Risolvente di T in z [7.34].
- $\text{Ran } f$ Immagine di f . Se $f : A \rightarrow B$, $\text{Ran } f = f(A)$.
- $\rho(T)$ Insieme risolvente dell'operatore lineare T .
- $\text{sgn}(x)$ Segno di x .
- \mathbb{R} Il campo dei numeri reali.
- \mathbb{R}^∞ Insieme delle successioni reali.
- $\mathbb{R}[x]$ Insieme dei polinomi a coefficienti reali.
- $\text{spec } T$ Spettro dell'operatore lineare T .
- $\sigma(T)$ Spettro dell'operatore lineare T .
- $\sigma_p(T)$ Spettro puntuale dell'operatore lineare T .
- $\sigma_c(T)$ Spettro continuo dell'operatore lineare T .
- $\text{span}_{\mathbb{F}}\{X\}$ Insieme delle combinazioni lineari di elementi di X coefficienti nel campo \mathbb{F} .
- $T_a f$ Operatore di traslazione applicato alla funzione f . [Sezione 9.2].
- ϑ Distribuzione di Heavyside (6.9).
- ϑ_+ Operatore di traslazione verso destra (sezione 7.1.1).
- ϑ_- Operatore di traslazione verso sinistra (sezione 7.1.1).
- \mathbb{Z} L'anello dei numeri interi.
- $\Phi_a f$ Operatore di dilatazione/contrazione applicato alla funzione f . [Sezione 9.2].

1. Prerequisiti

W: I'm Winston Wolf, I solve problems

1.1 Algebra elementare

1.1.1 Identità algebriche elementari

1.1. (Coefficiente binomiale). Il simbolo $\binom{n}{k}$ è il coefficiente binomiale dato da

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

1.2. (Coefficiente binomiale generalizzato). Se $b \in \mathbb{R}(\mathbb{C})$ e $k \in \mathbb{N}$, definisco

$$\binom{b}{k} := \frac{b(b-1)(b-2)\cdots(b-k+1)}{k!}.$$

Esempio:

$$\binom{1/2}{4} = \frac{(1/2)(-1/2)(-3/2)(-5/2)}{4!} = -\frac{5}{128}$$

1.3. Alcune identità:

$$(x+y)^2 = x^2 + 2xy + y^2$$

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

$$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + y^n$$

$$(x^2 - y^2) = (x+y)(x-y)$$

$$(x^4 - y^4) = (x+y)(x-y)(x^2 + y^2)$$

$$(x^n - y^n) = (x-y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1})$$

$$(x^n + y^n) = (x+y)(x^{n-1} - x^{n-2}y + x^{n-3}y^2 + \cdots - xy^{n-2} + y^{n-1}) \quad n \text{ dispari}$$

1.4. Somme geometriche

$$\sum_{k=0}^n a^k = \frac{a^{n+1} - 1}{a - 1} \qquad \sum_{k=1}^n a^k = \frac{a^{n+1} - a}{a - 1}$$

1.5 Problema. Calcolare $\sum_{k=0}^{37} \binom{37}{k}$. (Sugg: usare l'identità...)

Risp: 137438953472

1.1.2 Divisione di polinomi

1.6 Problema. Siano

$$P(x) := 3x^4 + 2x^3 - x^2 + 5x + 3 \quad Q(x) := 2x^2 - 4x + 1.$$

- (1) Trovare quoziente e resto della divisione di P per Q , vale a dire trovare due polinomi A e B tali che

$$\textcircled{1} \quad P(x) = A(x)Q(x) + B(x).$$

- (2) Dimostrare che la soluzione di questo problema è unica. Supponiamo cioè che valga la $\textcircled{1}$ e contemporaneamente

$$P(x) = A'(x)Q(x) + B'(x).$$

Allora deve essere necessariamente $A = A'$ e $B = B'$.

Risp: $A(x) = \frac{3}{2}x^2 + 4x + \frac{27}{4}$, $B(x) = 28x - \frac{15}{4}$.

1.1.3 Risoluzione di equazioni di grado superiore al secondo

Le soluzioni di un'equazione di terzo grado

$$x^3 + bx^2 + cx + d = 0$$

si trovano con la formuletta che tutti conoscono

$$\begin{aligned} x_1 &= -\frac{b}{3} - \frac{\alpha}{3\beta} + \frac{\beta}{3} \\ x_2 &= -\frac{b}{3} - \frac{(1+i\sqrt{3})\alpha}{6\beta} - \frac{1}{6}(1-i\sqrt{3})\beta \\ x_3 &= -\frac{b}{3} + \frac{(1-i\sqrt{3})\alpha}{6\beta} - \frac{1}{6}(1+i\sqrt{3})\beta \end{aligned}$$

in cui

$$\begin{aligned} \alpha &:= -b^2 + 3c \\ \gamma &:= -2b^3 + 9bc - 27d \\ \beta &:= \left[\frac{1}{2} \left(\gamma + \sqrt{4\alpha^3 + \gamma^2} \right) \right]^{1/3} \end{aligned}$$

Un formula esiste anche per le equazioni di quarto grado, ma tenderei a risparmiarvela.

Se si ha motivo di sospettare che l'equazione¹

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_{n-1} x + a_n$$

abbia soluzioni intere oppure razionali con denominatori piccoli, si può provare a cercare soluzioni fra i divisori del termine noto a_n . Ovviamente non è assolutamente detto che ci siano soluzioni intere, anche nel caso in cui tutti i coefficienti a_i siano interi.

¹il coefficiente del termine di grado massimo si può sempre porre uguale ad 1

L'equazione può non avere alcuna soluzione reale, come nel caso di $x^{10} + 1 = 0$ o può avere soluzioni reali irrazionali. Ad esempio le soluzioni di

$$x^2 + 4x + 1 = 0$$

sono

$$x_1 = -2 - \sqrt{5} \quad x_2 = -2 + \sqrt{5}.$$

Vediamo invece un caso il cui il metodo del “tirare a indovinare” ha successo.

1.7 Problema. Trovare tutte le soluzioni dell'equazione

$$p(x) := x^4 - 11x^3 + 42x^2 - 64x + 32 = 0.$$

Soluzione. Se siamo fortunati troviamo qualche radice fra i divisori del termine noto 32. I divisori (con segno) di 32 sono

$$\pm 1, \pm 2, \pm 4, \pm 8, \pm 16, \pm 32.$$

Io inizierei a provare con $x = 1$.

$$x = 1 : \quad p(1) = 1^4 - 11 \cdot 1^3 + 42 \cdot 1^2 - 64 \cdot 1 + 32 = 0 !$$

Stupendo, una delle radici è $x = 1$. A questo punto possiamo

- (1) o dividere $p(x)$ per $x - 1$ ottenendo un polinomio di terzo grado,
- (2) oppure provare a vedere se qualche altro divisore di 32 è anch'esso radice.

Seguo la prima strada e ottengo

$$x^4 - 11x^3 + 42x^2 - 64x + 32 = (x - 1)(x^3 - 10x^2 + 32x - 32).$$

Devo trovare dunque le 3 radici di $q(x) := x^3 - 10x^2 + 32x - 32$. Provo di nuovo con i divisori del termine noto.

$$\begin{array}{ll} x = 1 : & q(1) = 1^3 - 10 \cdot 1^2 + 32 \cdot 1 - 32 = -9 \neq 0 & : -(\\ x = -1 : & q(-1) = (-1)^3 - 10 \cdot (-1)^2 + 32 \cdot (-1) - 32 = -75 \neq 0 & : -(\\ x = 2 : & q(2) = 2^3 - 10 \cdot 2^2 + 32 \cdot 2 - 32 = 0 & : -) \end{array}$$

Ottimo, un'altra radice è $x = 2$. Divido di nuovo $q(x)$ per $x - 2$ e ottengo

$$q(x) = (x - 2)(x^2 - 8x + 16).$$

L'ultimo fattore è un polinomio di secondo grado facilmente scomponibile

$$(x^2 - 8x + 16) = (x - 4)^2$$

che vuol dire che ottengo la radice $x = 4$ con molteplicità (algebraica) 2. Le radici di $p(x)$ sono quindi

$$(1.1) \quad \begin{array}{ll} x = 1 & \text{con molteplicità } 1 \\ x = 2 & \text{con molteplicità } 1 \\ x = 4 & \text{con molteplicità } 2. \end{array}$$

□

1.2 Algebra lineare

1.2.1 Il metodo della riduzione delle righe

Questo metodo è spesso conveniente per risolvere vari problemi connessi con le matrici, come:

- (1) risolvere un sistema di equazioni lineari
- (2) trovare il rango di una matrice
- (3) trovare una base del nucleo di una matrice
- (4) trovare una base dell'immagine di una matrice

Consideriamo la matrice

$$(1.2) \quad A = \begin{pmatrix} 3 & -1 & 4 & 7 & 11 & 0 & 5 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 2 & 0 & 2 & 8 & 9 & 2 & -2 \\ 2 & 1 & 1 & 8 & 9 & 0 & 5 \end{pmatrix}$$

La matrice A può essere pensata come un'applicazione

$$A : \mathbb{R}^7 \rightarrow \mathbb{R}^4$$

che agisce come moltiplicazione, cioè, se $x = (x_1, x_2, \dots, x_7)$, allora

$$(Ax)_i := \sum_{j=1}^7 A_{ij} x_j.$$

Il metodo della riduzione delle righe consiste nell'eseguire un certo numero di trasformazioni fino ad arrivare ad una nuova matrice B in "forma ridotta". Ad esempio, nel caso di A , una sua forma ridotta B è data da

$$(1.3) \quad B = \begin{pmatrix} \boxed{1} & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & \boxed{1} & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1/2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Gli elementi che caratterizzano la forma ridotta sono:

- (a) ogni riga è o composta di tutti zeri, oppure ha un primo elemento non nullo che è uguale ad 1. Questo "uno" è detto "uno iniziale" della riga e appare dentro un quadrato nella (1.3);
- (b) se una colonna contiene un $\boxed{1}$ (un "uno iniziale"), allora la porzione di colonna al di sotto dell' $\boxed{1}$ contiene solo zeri.

Per ridurre una matrice, cioè per passare da A a B , si effettua una serie di operazioni permesse. Le operazioni permesse sono:

- (1) moltiplicare una riga per un numero diverso da zero;
- (2) aggiungere ad una riga il multiplo di un'altra riga;
- (3) scambiare due righe.

Effettuando queste operazioni conviene fare attenzione allo scopo di evitare di introdurre frazioni quando non sono necessarie.² Riparto dalla matrice A

$$A = \begin{pmatrix} 3 & -1 & 4 & 7 & 11 & 0 & 5 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 2 & 0 & 2 & 8 & 9 & 2 & -2 \\ 2 & 1 & 1 & 8 & 9 & 0 & 5 \end{pmatrix}$$

Si procede per colonne. Inizio dalla prima colonna. Voglio che appaia un $\boxed{1}$ in alto e sotto tutti zeri. Sottraggo la terza riga dalla prima

$$\begin{array}{l} R_1 - R_3 \\ R_2 \\ R_3 \\ R_4 \end{array} \begin{bmatrix} 1 & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 2 & 0 & 2 & 8 & 9 & 2 & -2 \\ 2 & 1 & 1 & 8 & 9 & 0 & 5 \end{bmatrix}$$

Poi procedo nel modo seguente

$$\begin{array}{l} R_1 \\ R_2 \\ R_3 - 2R_1 \\ R_4 - 2R_1 \end{array} \begin{bmatrix} 1 & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 0 & 2 & -2 & 10 & 5 & 6 & -16 \\ 0 & 3 & -3 & 10 & 5 & 4 & -9 \end{bmatrix}$$

E la prima colonna sta a posto. Ora elimino il 2 e il 3 in fondo alla seconda colonna

$$\begin{array}{l} R_1 \\ R_2 \\ R_3 - 2R_2 \\ R_4 - 3R_2 \end{array} \begin{bmatrix} 1 & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 6 & 3 & 6 & -18 \\ 0 & 0 & 0 & 4 & 2 & 4 & -12 \end{bmatrix}$$

Nella terza riga trasformo il primo 6 in $\boxed{1}$

$$\begin{array}{l} R_1 \\ R_2 \\ R_3/6 \\ R_4 \end{array} \begin{bmatrix} 1 & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1/2 & 1 & -3 \\ 0 & 0 & 0 & 4 & 2 & 4 & -12 \end{bmatrix}$$

Sistemo la quarta colonna eliminando il 4 in fondo

$$(1.4) \quad \begin{array}{l} R_1 \\ R_2 \\ R_3 \\ R_4 - 4R_3 \end{array} \begin{bmatrix} \boxed{1} & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & \boxed{1} & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1/2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Fatto.³ Nella (1.4) appare *una* versione ridotta della matrice A . Non è l'unica possibile. Si potrebbe infatti a questo punto aggiungere, ad esempio, la seconda riga alla prima e ottenere un'altra versione della matrice ridotta. Ai due requisiti (a) e (b) che definiscono la forma ridotta di una matrice se ne può aggiungere un terzo. Diremo che il processo di riduzione delle righe è *completo*⁴ se, oltre alle proprietà (a) e (b) è verificata anche la proprietà

²ma spesso è inevitabile introdurre frazioni

³il lettore sarebbe certamente fuorviato qualora pensasse che i calcoli vengono sempre così semplici

⁴la terminologia me la sono appena inventata

- (c) se una colonna contiene un “uno iniziale”, cioè un $\boxed{1}$, tutti gli altri elementi della colonna (anche quelli che si trovano sopra l' $\boxed{1}$) sono nulli

Il procedimento di riduzione completa delle righe determina una matrice ridotta in maniera univoca. Nel caso della matrice (1.4) per ottenere la riduzione completa dobbiamo cancellare il -1 che si trova il colonna 2 e i due elementi non nulli che si trovano sopra l' $\boxed{1}$ in colonna 4. Il risultato è:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & \frac{5}{2} & -3 & 11 \\ 0 & 1 & -1 & 0 & 0 & -2 & 7 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Determinazione della dimensione del nucleo e dell'immagine di A

Riscrivo la matrice ottenuta aggiungendo le variabili x_1, \dots, x_7 . Se nella colonna i -sima compare un $\boxed{1}$ inquadretto la variabile corrispondente x_i

$$(1.5) \quad \begin{bmatrix} \boxed{1} & -1 & 2 & -1 & 2 & -2 & 7 \\ 0 & \boxed{1} & -1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1/2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \boxed{x_1} & \boxed{x_2} & x_3 & \boxed{x_4} & x_5 & x_6 & x_7 \end{bmatrix}$$

Sia

$$\begin{aligned} \text{Ker } A &:= \text{il nucleo di } A \subset \mathbb{R}^7 \\ \text{Ran } A &:= \text{l'immagine di } A \subset \mathbb{R}^4. \end{aligned}$$

Dalla (1.5) ottengo che, poiché ci sono 3 $\boxed{1}$, vale a dire 3 righe non nulle, allora

$$(1.6) \quad \begin{aligned} 3 &= \text{numero di righe linearmente indipendenti} = \\ &= \text{numero di colonne linearmente indipendenti} = \dim \text{Ran } A \end{aligned}$$

D'altra parte sappiamo che

$$\dim \text{Ker } A + \dim \text{Ran } A = \text{dimensione del dominio di } A = \dim \mathbb{R}^7 = 7,$$

per cui

$$\dim \text{Ran } A = 3 \quad \dim \text{Ker } A = 4.$$

Base dell'immagine di A.

Una base dell'immagine di A è data dai vettori colonna nella matrice originale A che corrispondono alle variabili $\boxed{x_i}$. Nel nostro caso le colonne 1, 2, 4 di A .

$$\begin{aligned} \text{Base di Ran } A &\Rightarrow \begin{aligned} u_1 &= (3, 0, 2, 2) \\ u_2 &= (-1, 1, 0, 1) \\ u_3 &= (7, 2, 8, 8). \end{aligned} \end{aligned}$$

Base del nucleo di A (ovvero, come risolvere un sistema omogeneo).

Il nucleo di A è l'insieme dei vettori $x \in \mathbb{R}^7$ tali che $Ax = 0$. Quindi determinare il nucleo di A equivale a trovare la soluzione generale del sistema omogeneo $Ax = 0$. Per risolvere questo sistema si può usare al posto di A la matrice ridotta B data dalla (1.5). Il metodo è il seguente:

- (1) le variabili senza quadretto sono *libere*. A queste si possono assegnare parametri liberi r, s, t, q e si portano a secondo membro;
- (2) le variabili inquadrette $\boxed{x_i}$ invece vengono determinate *in ordine inverso* in funzione di quelle libere.

Quindi, utilizzando la (1.5), ottengo

$$(1.7) \quad x_3 = r \quad x_5 = s \quad x_6 = t \quad x_7 = q$$

e, a partire da queste, ottengo le $\boxed{x_i}$:

$$\begin{aligned} x_4 &= -\frac{s}{2} - t + 3q \\ x_2 &= r - 2x_3 - s - q = r + 2t - 7q \\ x_1 &= x_2 - 2r + x_4 - 2s + 2t - 7q = -r - \frac{5}{2}s + 3t - 11q \end{aligned}$$

Riordino, metto tutto insieme e scrivo la

$$\begin{array}{l} \text{soluzione generale del} \\ \text{система lineare omogeneo } Ax = 0 \end{array} \Rightarrow \begin{array}{l} x_1 = -r - \frac{5}{2}s + 3t - 11q \\ x_2 = r + 2t - 7q \\ x_3 = r \\ x_4 = -\frac{s}{2} - t + 3q \\ x_5 = s \\ x_6 = t \\ x_7 = q \end{array}$$

L'ultimo passo, quello della costruzione di una base del nucleo di A , a questo punto è quasi a costo zero. Ho 4 parametri liberi e scrivo 4 vettori v_1, \dots, v_4 nel modo seguente

$$\begin{array}{l} \text{in } v_1 \text{ prendo } r = 1 \text{ e gli altri parametri uguale a zero} \\ \text{in } v_2 \text{ prendo } s = 1 \text{ e gli altri parametri uguale a zero} \\ \vdots \end{array}$$

Questa non è l'unica ricetta possibile (ovviamente) ma funziona. In particolare il valore 1 è arbitrario e può essere sostituito con un qualsiasi numero diverso da 0. Ad esempio, dato che s appare sempre diviso per 2, porrò $s = 2$ per calcolare v_2 . Dunque ci siamo:

$$\begin{array}{l} \text{base di Ker } A \end{array} \Rightarrow \begin{array}{l} v_1 := (-1, 1, 1, 0, 0, 0, 0) \\ v_2 := (-5, 0, 0, -1, 0, 0, 0) \\ v_3 := (3, 2, 0, -1, 0, 1, 0) \\ v_4 := (-11, -7, 0, 3, 0, 0, 1) \end{array}$$

Precisazione

L'autore non ha difficoltà ad ammettere che la matrice A che appare nella (1.2) è stata scelta con una certa... benevolenza allo scopo di ottenere calcoli semplici. Ci si potrà

facilmente rendere conto che per una matrice i cui elementi vengono scelti “a caso”, il processo di riduzione delle righe può risultare più complicato. Tuttavia, con un po’ di pratica, il procedimento diventa abbastanza rapido. Costruisco, ad esempio, una matrice i cui elementi sono interi presi a caso con distribuzione uniforme fra -5 e 5

$$\begin{bmatrix} 0 & 2 & -1 & 0 & -3 & 4 & -2 \\ 2 & -3 & -4 & -4 & 4 & -1 & -1 \\ 4 & 1 & -5 & -1 & -2 & 2 & 4 \\ 1 & -2 & 1 & 2 & 1 & -4 & -1 \end{bmatrix}$$

Si vede subito ad occhio che la versione completamente ridotta di questa matrice è

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \frac{16}{9} & -\frac{10}{3} & \frac{257}{27} \\ 0 & 1 & 0 & 0 & -\frac{4}{9} & \frac{1}{3} & \frac{91}{27} \\ 0 & 0 & 1 & 0 & \frac{19}{9} & -\frac{10}{3} & \frac{236}{27} \\ 0 & 0 & 0 & 1 & -\frac{17}{9} & \frac{5}{3} & -\frac{169}{27} \end{bmatrix}$$

però, appunto, come dicevo, è necessaria un po’ di pratica.⁵ Ci può porre il seguente problema:

1.8 Problema. Sia A una matrice $m \times n$ i cui elementi sono interi scelti a caso con distribuzione uniforme fra $-k$ e k . Qual è la probabilità che il procedimento di riduzione completa delle righe produca una matrice i cui elementi sono tutti numeri interi?

Soluzione di un sistema lineare non omogeneo

Se A è una matrice $m \times n$ e b è un vettore in \mathbb{R}^m , possiamo costruire il sistema di equazioni lineari (non omogeneo) $Ax = b$, che, scritto esplicitamente, diventa

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.8}$$

Denotiamo con C_i la i -sima colonna della matrice A

$$C_i := \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

Il sistema (1.8) può essere riscritto come

$$x_1 C_1 + x_2 C_2 + \cdots + x_n C_n = b,$$

quindi dire che il sistema $Ax = b$ ammette almeno una soluzione equivalente a dire che b è una combinazione lineare delle colonne di A . Per risolvere questo sistema si procede più o meno come nel caso di un sistema omogeneo. Si costruisce la matrice $[A|b]$ con $n + 1$ colonne, fatta da A e dal vettore b . Poi si applica il metodo della

⁵oppure un programma di manipolazione simbolica come Maple o Mathematica

riduzione delle righe alla matrice $[A|b]$. Facciamo un esempio: sia A la matrice già usata in precedenza, definita nella (1.2) e sia

$$b = (5, 1, 3, 5).$$

Allora abbiamo

$$(1.9) \quad [A|b] = \left[\begin{array}{cccccc|c} 3 & -1 & 4 & 7 & 11 & 0 & 5 \\ 0 & 1 & -1 & 2 & 1 & 0 & 1 \\ 2 & 0 & 2 & 8 & 9 & 2 & -2 \\ 2 & 1 & 1 & 8 & 9 & 0 & 5 \end{array} \right]$$

Operiamo con la riduzione delle righe e otteniamo⁶

$$(1.10) \quad \left[\begin{array}{cccccc|c} \boxed{1} & -1 & 2 & -1 & 2 & -2 & 7 & | & 2 \\ 0 & \boxed{1} & -1 & 2 & 1 & 0 & 1 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & \frac{1}{2} & 1 & -3 & | & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 \end{array} \right]$$

La soluzione generale del sistema, a questo punto si ottiene in modo analogo al caso del sistema omogeneo. Fare attenzione che i numeri che appaiono nell'ultima colonna si trovano già "al secondo membro" quindi non vanno cambiati di segno.

Contrariamente al caso omogeneo in cui può avere una soluzione oppure infinite soluzioni, un sistema lineare non omogeneo può non avere soluzioni. Questo avviene quando nella matrice ridotta appare una riga del tipo

$$(1.11) \quad [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ | \ 1]$$

che corrisponde all'equazione

$$0 = 1.$$

Esaminiamo dei casi di matrici già ridotte. Partiamo dalla (1.10)

$$(1.12) \quad \left[\begin{array}{cccccc|c} \boxed{1} & -1 & 2 & -1 & 2 & -2 & 7 & | & 2 \\ 0 & \boxed{1} & -1 & 2 & 1 & 0 & 1 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & \frac{1}{2} & 1 & -3 & | & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 \\ \boxed{x_1} & \boxed{x_2} & x_3 & \boxed{x_4} & x_5 & x_6 & x_7 & & \end{array} \right]$$

In questo caso abbiamo quattro parametri liberi e nessuna "riga impossibile" come la (1.11), quindi il sistema ammette infinite soluzioni. A volte, per dire che ci sono infinite soluzioni dipendenti da 4 parametri, vale a dire che l'insieme delle soluzioni è uno spazio quadridimensionale, si dice che il sistema ha ∞^4 soluzioni. Il sistema

$$(1.13) \quad \left[\begin{array}{cccc|c} \boxed{1} & -1 & 2 & -1 & 2 & | & 2 \\ 0 & 0 & \boxed{1} & 0 & 2 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 & | & -1 \\ 0 & 0 & 0 & 0 & 0 & | & 1 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \\ \boxed{x_1} & x_2 & \boxed{x_3} & \boxed{x_4} & x_5 & & \end{array} \right]$$

⁶basta applicare all'ultima colonna la stessa sequenza di operazioni usate in precedenza per la matrice A

invece non ha soluzioni a causa della quarta riga. Il sistema

$$(1.14) \quad \left[\begin{array}{cccc|c} \boxed{1} & -1 & 2 & -1 & 2 \\ 0 & \boxed{1} & -1 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline \boxed{x_1} & \boxed{x_2} & \boxed{x_3} & \boxed{x_4} & \end{array} \right]$$

ha un'unica soluzione, perché non ci sono righe impossibili né parametri liberi. Il sistema

$$(1.15) \quad \left[\begin{array}{cccc|c} \boxed{1} & -1 & 2 & -1 & 2 \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ \hline \boxed{x_1} & x_2 & \boxed{x_3} & x_4 & \end{array} \right]$$

non ha soluzioni a causa della riga 3.

1.2.2 Autovalori, autovettori e diagonalizzazione di una matrice

1.9 Problema. Trovare autovalori e autovettori della matrice A e diagonalizzare A se è possibile

$$A = \begin{pmatrix} 5 & 3 & -2 & -1 \\ -3 & 1 & 6 & -3 \\ -2 & 0 & 8 & -4 \\ -5 & -3 & 10 & -3 \end{pmatrix}$$

Soluzione. Si costruisce la matrice

$$(1.16) \quad \lambda I - A = \begin{pmatrix} \lambda - 5 & -3 & 2 & 1 \\ 3 & \lambda - 1 & -6 & 3 \\ 2 & 0 & \lambda - 8 & 4 \\ 5 & 3 & -10 & \lambda + 3 \end{pmatrix}$$

e se ne calcola il determinante, detto anche polinomio caratteristico

$$p(\lambda) := \det(\lambda I - A) = \begin{vmatrix} \lambda - 5 & -3 & 2 & 1 \\ 3 & \lambda - 1 & -6 & 3 \\ 2 & 0 & \lambda - 8 & 4 \\ 5 & 3 & -10 & \lambda + 3 \end{vmatrix} = \lambda^4 - 11\lambda^3 + 42\lambda^2 - 64\lambda + 32$$

Gli autovalori sono le soluzioni dell'equazione.

$$p(\lambda) := \lambda^4 - 11\lambda^3 + 42\lambda^2 - 64\lambda + 32 = 0.$$

Dalla soluzione del Problema 1.7 otteniamo gli autovalori di A

$$(1.17) \quad \begin{array}{lll} \lambda_1 = 1 & \text{con molteplicità } 1 & \Rightarrow 1 \text{ autovettore} \\ \lambda_2 = 2 & \text{con molteplicità } 1 & \Rightarrow 1 \text{ autovettore} \\ \lambda_3 = 4 & \text{con molteplicità } 2 & \Rightarrow 1 \text{ o } 2 \text{ autovettori} \end{array}$$

Per trovare gli autovettori sostituiamo ciascun autovalore al posto di x e risolviamo il sistema

$$(1.18) \quad (\lambda_i I - A)v = 0.$$

Prendiamo il primo autovalore $\lambda_1 = 1$. Ponendo

$$v = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

la (1.18) diventa

$$(1.19) \quad \begin{pmatrix} -4 & -3 & 2 & 1 \\ 3 & 0 & -6 & 3 \\ 2 & 0 & -7 & 4 \\ 5 & 3 & -10 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Risolvendo col metodo (ad esempio) di riduzione delle righe si trova che w può essere preso come parametro libero e la soluzione si può scrivere come

$$(1.20) \quad x = \frac{s}{3} \quad y = \frac{s}{3} \quad z = \frac{2s}{3} \quad w = s.$$

Ponendo quindi $s = 3$ otteniamo un autovettore corrispondente all'autovalore $\lambda = 1$, dato da

$$u_1 = (1, 1, 2, 3)$$

- secondo autovalore $\lambda_2 = 2$.

Ponendo $\lambda = 2$ nella (1.16), la (1.18) diventa:

$$(1.21) \quad \begin{pmatrix} -3 & -3 & 2 & 1 \\ 3 & 1 & -6 & 3 \\ 2 & 0 & -6 & 4 \\ 5 & 3 & -10 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Eseguendo la riduzione delle righe sulla matrice $\lambda_2 I - A$ si ottiene ad esempio

$$(1.22) \quad \begin{pmatrix} \boxed{1} & 3 & 4 & -5 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Quindi le variabili x , y e z possono essere determinate in funzione del parametro libero $w = s$. La soluzione generale del sistema (1.21) è data da

$$x = sy = 0 \quad z = sw = s.$$

Ponendo quindi $s = 1$ otteniamo l'autovettore corrispondente a $\lambda = \lambda_2 = 2$

$$u_2 = (1, 0, 1, 1).$$

- Terzo ed ultimo autovalore: $\lambda_3 = 4$.

Ponendo $\lambda = 4$ nella (1.16), la (1.18) diventa:

$$(1.23) \quad \begin{pmatrix} -1 & -3 & 2 & 1 \\ 3 & 2 & -6 & 3 \\ 2 & 0 & -4 & 4 \\ 5 & 3 & -10 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Eseguendo la riduzione delle righe sulla matrice $\lambda_3 I - A$ si ottiene ad esempio

$$(1.24) \quad \begin{pmatrix} \boxed{1} & 0 & -2 & 2 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Quindi le variabili x e y possono essere determinate in funzione dei due parametri libero $z = s$, $w = t$. La soluzione generale del sistema (1.23) è data da

$$x = 2s - 2ty = t \qquad z = sw = t.$$

Ponendo prima $s = 1$ e $t = 0$ e poi $s = 0$ e $t = 1$ otteniamo due autovettori linearmente indipendenti corrispondenti a $\lambda = \lambda_3 = 4$

$$(1.25) \qquad u_3 = (2, 0, 1, 0)$$

$$(1.26) \qquad u_4 = (-2, 1, 0, 1)$$

- Elenco gli autovalori di A con i corrispondenti autovettori

$$\begin{array}{ll} u_1 = (1, 1, 2, 3) & \lambda_1 = 1 \\ u_2 = (1, 0, 1, 1) & \lambda_2 = 2 \\ u_3 = (2, 0, 1, 0) & \lambda_3 = 4 \\ u_4 = (-2, 1, 0, 1) & \lambda_3 = 4 \end{array}$$

- Diagonalizzazione di A . Dato che ho trovato 4 autovettori linearmente indipendenti, posso diagonalizzare A , vale a dire A può essere scritta come

$$A = U D U^{-1},$$

In cui D è la matrice diagonale che contiene gli autovalori di A

$$(1.27) \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

e U è la matrice le cui colonne sono costituite dai corrispondenti autovettori:

$$(1.28) \quad U = [u_1 | u_2 | u_3 | u_4] = \begin{pmatrix} 1 & 1 & 2 & -2 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 3 & 1 & 0 & 1 \end{pmatrix}$$

1.2.3 Funzioni di matrici

1.10 Problema. Sia

$$A = \begin{pmatrix} 0 & 4 & 3 \\ 1 & -3 & 3 \\ 2 & -4 & 5 \end{pmatrix}$$

Calcolare $\det(A^{10})$, $\text{tr}(A^{10})$, $\det(e^A)$.

Risp: 10^{10} , 9766650 , e^2

1.11 Problema. Sia

$$A = \begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}$$

Calcolare e^A e verificare che $\det(e^A) = e^{\text{tr} A}$. (Sugg: conviene diagonalizzarla).

1.3 Analisi reale

1.3.1 Regola di Leibniz per la derivata di un prodotto

1.12 Proposizione. Siano $f : \mathbb{R} \rightarrow \mathbb{R}$ e $g : \mathbb{R} \rightarrow \mathbb{R}$ due funzioni differenziabili n volte in (a, b) . Allora

$$(1.29) \quad \frac{d^n}{dx^n}(fg)(x) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x). \quad \forall x \in (a, b).$$

Dimostrazione. La dimostrazione è per induzione su n . Se $n = 1$, poiché $\binom{1}{0} = \binom{1}{1} = 1$, la proposizione afferma che $(fg)' = fg' + f'g$ cosa che sappiamo essere vera. Assumo dunque che valga (1.29). Se $f, g \in C^{n+1}(a, b)$, posso scrivere, ponendo $D = d/dx$,

$$\begin{aligned} D^{n+1}(fg)(x) &= D [D^n(fg)](x) = D \left[\sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x) \right] \\ &= \sum_{k=0}^n \binom{n}{k} D \left[f^{(k)}(x) g^{(n-k)}(x) \right] \\ &= \sum_{k=0}^n \binom{n}{k} \left[f^{(k)}(x) g^{(n-k+1)}(x) + f^{(k+1)}(x) g^{(n-k)}(x) \right] \\ &= \sum_{k=0}^n \binom{n}{k} \left[f^{(k)}(x) g^{(n-k+1)}(x) + f^{(k+1)}(x) g^{(n-k)}(x) \right] \\ &= \sum_{k=0}^n \binom{n}{k} \left[f^{(k)}(x) g^{(n-k+1)}(x) \right] + \sum_{k=0}^n \binom{n}{k} \left[f^{(k+1)}(x) g^{(n-k)}(x) \right] \end{aligned}$$

Nella seconda sommatoria possiamo cambiare $k \rightarrow k - 1$ ottenendo

$$\sum_{k=0}^n \binom{n}{k} f^{(k+1)}(x) g^{(n-k)}(x) = \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)}(x) g^{(n+1-k)}(x)$$

Di conseguenza, grazie all'identità

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

si ottiene

$$\begin{aligned}
 D^{n+1}(fg)(x) &= \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k+1)}(x) + \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)}(x) g^{(n+1-k)}(x) \\
 &= \sum_{k=1}^n \left[\binom{n}{k} + \binom{n}{k-1} \right] f^{(k)}(x) g^{(n+1-k)}(x) \\
 &\quad + \binom{n}{0} f(x) g^{(n+1)}(x) + \binom{n}{n} f^{(n+1)}(x) g(x) \\
 &= \sum_{k=1}^n \binom{n+1}{k} f^{(k)}(x) g^{(n+1-k)}(x) + f(x) g^{(n+1)}(x) + f^{(n+1)}(x) g(x) \\
 &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)}(x) g^{(n+1-k)}(x).
 \end{aligned}$$

Quindi la (1.29) è valida per $n+1$. La dimostrazione per induzione è così completata \square

1.3.2 Integrali impropri su \mathbb{R}

1.13 Notazione. In questa sezione consideriamo funzioni $f : [a, b) \rightarrow \mathbb{R}$. Il simbolo b denota un numero reale finito oppure $+\infty$. In questo modo possiamo trattare parallelamente sia gli integrali impropri su un intervallo finito sia quelli su un intervallo infinito $[a, +\infty)$. Inoltre il caso in cui f è continua su un intervallo $(a, b]$ chiuso a destra e aperto a sinistra si ottiene per simmetria, a partire dal caso precedente.

Se f è una funzione continua sull'intervallo semiaperto $[a, b)$ si dice che f è integrabile (in senso improprio) su $[a, b)$ se esiste ed è finito il limite

$$(1.30) \quad \lim_{x \rightarrow b^-} \int_a^x f(t) dt$$

e si pone

$$\int_a^b f(t) dt = \lim_{x \rightarrow b^-} \int_a^x f(t) dt.$$

La dicitura “in senso improprio” viene a volte omessa perché ovvia. Nel caso di una funzione non negativa la quantità $\int_a^x f(t) dt$ è non decrescente in x , quindi il limite nella (1.30) esiste sempre, ma può essere uguale a $+\infty$. In questo caso scriveremo

$$\begin{aligned}
 \int_a^b f(t) dt = \infty & \quad \text{se } \lim_{x \rightarrow b^-} \int_a^x f(t) dt = +\infty \\
 \int_a^b f(t) dt < \infty & \quad \text{se } \lim_{x \rightarrow b^-} \int_a^x f(t) dt \text{ è finito.}
 \end{aligned}$$

Dal criterio di Cauchy per l'esistenza del limite di una funzione si ottiene il seguente criterio di integrabilità:

1.14 Teorema. (Criterio di Cauchy per gli integrali impropri). Se f è una funzione continua su $[a, b)$, allora f è integrabile (impropriamente) su $[a, b)$ se e solo se vale

$$\forall \varepsilon > 0 \exists x_0 \in [a, b) \text{ tale che per ogni } x_1, x_2 \text{ in } [x_0, b) \text{ si ha } \left| \int_{x_1}^{x_2} f(t) dt \right| < \varepsilon.$$

Inoltre se f è integrabile su $[a, b]$, allora, per ogni $x \in [a, b)$, f è anche integrabile su $[x, b]$ e vale

$$\lim_{x \rightarrow b^-} \int_x^b f(t) dt = 0.$$

Come caso particolare otteniamo il seguente

1.15 Corollario. Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ continua e non negativa. Allora

$$\int_{-\infty}^{\infty} f(t) dt < \infty \iff \lim_{x \rightarrow \infty} \int_{|t| \geq x} f(t) dt = 0.$$

Dal Teorema 1.14 si ricava il *criterio del confronto*:

1.16 Teorema. (Criterio del confronto per gli integrali impropri). Siano f e g sono due funzioni continue su $[a, b)$ tali che

$$|f(x)| \leq g(x) \quad \text{per ogni } x \in [a, b).$$

Se g è integrabile su $[a, b)$, allora anche f e $|f|$ sono integrabili su $[a, b)$.

A volte risulta più comodo utilizzare il *confronto asintotico*:

1.17 Teorema. (Criterio del confronto asintotico per gli integrali impropri). Siano f e g sono due funzioni continue non negative su $[a, b)$ tali che

$$\lim_{x \rightarrow b^-} \frac{f(x)}{g(x)} = \ell \in [0, +\infty].$$

Allora:

(1) se $\ell \in (0, +\infty)$ si ha $\int_a^b f(t) dt < \infty \iff \int_a^b g(t) dt < \infty$;

(2) se $\ell = 0$ si ha $\int_a^b g(t) dt < \infty \implies \int_a^b f(t) dt < \infty$;

(3) se $\ell = +\infty$ si ha $\int_a^b g(t) dt = \infty \implies \int_a^b f(t) dt = \infty$.

1.18 Esempi. Vediamo alcuni esempi.

(1) Sia $\alpha \in \mathbb{R}$ e si consideri l'integrale

$$(1.31) \quad \int_0^3 \frac{dx}{x^\alpha}.$$

Dalle definizione segue

$$\int_0^3 \frac{dx}{x^\alpha} = \lim_{y \rightarrow 0^+} \int_y^3 \frac{dx}{x^\alpha}.$$

Se $\alpha = 1$ si ottiene

$$\lim_{y \rightarrow 0^+} \int_y^3 \frac{dx}{x} = \lim_{y \rightarrow 0^+} [\log |x|]_y^3 = \lim_{y \rightarrow 0^+} [\log 3 - \log y]_y^3 = +\infty,$$

mentre se $\alpha \neq 1$ si ha

$$\lim_{y \rightarrow 0^+} \int_y^3 \frac{dx}{x^\alpha} = \lim_{y \rightarrow 0^+} \left[\frac{x^{1-\alpha}}{1-\alpha} \right]_y^3 = \lim_{y \rightarrow 0^+} \left[\frac{3^{1-\alpha} - y^{1-\alpha}}{1-\alpha} \right] = \begin{cases} \frac{3^{1-\alpha}}{1-\alpha} & \text{se } \alpha < 1 \\ +\infty & \text{se } \alpha > 1. \end{cases}$$

Riassumendo, abbiamo ottenuto che l'integrale (1.31) è convergente se e solo se $\alpha < 1$. Più in generale si ha che se $c \in [a, b]$

$$(1.32) \quad \int_a^b \frac{dx}{|x-c|^\alpha} < \infty \iff \alpha < 1.$$

(2) In maniera analoga al caso precedente si ottiene, per ogni $a > 0$

$$(1.33) \quad \int_a^\infty \frac{dx}{x^\alpha} < \infty \iff \int_{-\infty}^{-a} \frac{dx}{|x|^\alpha} < \infty \iff \alpha > 1.$$

(3) Siano p e q polinomi di grado, rispettivamente $\deg p$ e $\deg q$. Se q non ha radici reali, la funzione razionale $r(x) := p(x)/q(x)$ è continua su \mathbb{R} . Inoltre

$$\lim_{x \rightarrow \pm\infty} \left| \frac{r(x)}{x^{\deg p - \deg q}} \right| = \ell \neq 0.$$

Usando il criterio del confronto asintotico e la (1.33) otteniamo

$$(1.34) \quad \int_{\mathbb{R}} \left| \frac{p(x)}{q(x)} \right| dx < \infty \iff \deg p - \deg q \geq 2 \quad \text{se } q(x) \neq 0 \forall x \in \mathbb{R}.$$

Supponiamo invece che x_0 sia una radice reale di q di molteplicità m e che sia $p(x_0) \neq 0$.⁷ Questo implica che $r(x) = p(x)/q(x)$ diverge, per $x \rightarrow x_0$ come $g(x) := 1/(x - x_0)^m$, vale a dire

$$\lim_{x \rightarrow x_0} \left| \frac{r(x)}{g(x)} \right| = \ell \neq 0.$$

Siccome la molteplicità è almeno 1, il criterio del confronto asintotico unitamente alla (1.32) ci dice che, se $[a, b]$ è un intervallo finito si ha

$$(1.35) \quad \int_a^b \left| \frac{p(x)}{q(x)} \right| dx = \infty \quad \text{se } x_0 \in [a, b], q(x_0) = 0, p(x_0) \neq 0.$$

(4) Siano α e β due numeri reali positivi e sia $r := p/q$ una funzione razionale in cui q non ha zeri sull'asse reale. Voglio far vedere che

$$(1.36) \quad \int_{\mathbb{R}} r(x) e^{-\alpha|x|^\beta} dx < \infty.$$

Ponendo

$$f(x) = r(x) e^{-\alpha|x|^\beta} \qquad g(x) = \frac{1}{1+x^2}$$

si ha

$$\lim_{x \rightarrow \pm\infty} \frac{|f(x)|}{g(x)} = \lim_{x \rightarrow \pm\infty} \frac{p(x)(1+x^2)e^{-\alpha|x|^\beta}}{q(x)} = 0,$$

perché $p(x)(1+x^2)/q(x)$ all'infinito si comporta come una potenza, mentre il fattore $e^{-\alpha|x|^\beta}$ tende a zero, per $x \rightarrow \pm\infty$, più velocemente dell'inverso di qualsiasi potenza. Poiché l'integrale $\int_{\mathbb{R}} g(x) dx$ è finito in virtù della (1.34), dal criterio del confronto asintotico segue la (1.36).

⁷Possiamo supporre di aver semplificato il quoziente p/q eliminando i fattori comuni.

1.19 Problema. Dire per quali valori reali di a l'integrale è finito:

$$\begin{array}{lll}
 (a) \int_1^{\infty} \frac{1}{x^a} dx & (b) \int_0^1 \frac{1}{x^a} dx & (c) \int_{-\infty}^{+\infty} \frac{x}{(x^4 + x^2 + 5)^a} dx \\
 (d) \int_0^{\infty} e^{-a\sqrt{x}} dx & (e) \int_1^{\infty} e^{-x} x^a dx & (f) \int_2^{\infty} \frac{1}{(\log x)^a} dx \\
 (g) \int_0^{\infty} \frac{x^2}{1+x^a} dx & (h) \int_1^{\infty} \frac{\log x}{x^a} dx & (i) \int_2^{\infty} \frac{1}{x^a (\log x)^b} dx
 \end{array}$$

Risp: (a) $a > 1$, (b) $a < 1$, (c) $a > 1/2$, (d) $a > 0$, (e) $a \in \mathbb{R}$, (f) mai, (g) $a > 3$, (h) $a > 1$, (i) $a > 1$ e b qualsiasi oppure $a = 1$ e $b > 1$.

1.3.3 Confronto fra integrali e somme

1.20 Problema. Sia n_0 un intero e sia f una funzione continua positiva e *non crescente* su $[n_0, \infty)$. Dimostrare che

(a) per ogni intero $N > n_0$, si ha

$$\int_{n_0}^{N+1} f(x) dx \leq \sum_{k=n_0}^N f(k) \leq f(n_0) + \int_{n_0}^N f(x) dx$$

(b) la serie $\sum_{n_0}^{\infty} f(k)$ converge se e solo se l'integrale $\int_{n_0}^{\infty} f(x) dx$ è finito.

Grazie a quanto affermato nel problema precedente, possiamo usare molti risultati riguardanti la convergenza degli integrali per ottenere risultati analoghi sulla convergenza delle serie numeriche. Ad esempio dalla (1.33) si ottiene

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} < \infty \iff \alpha > 1.$$

1.3.4 Integrali gaussiani

1.21 Problema. Sapendo che $\int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi}$ (come si dimostra?), dimostrare che per ogni intero positivo n si ha

$$I_{2n} := \int_{-\infty}^{+\infty} e^{-x^2} x^{2n} dx = \sqrt{\pi} \frac{(2n-1)!!}{2^n} \quad I_{2n-1} := \int_{-\infty}^{+\infty} e^{-x^2} x^{2n-1} dx = 0$$

in cui $n!!$ è il prodotto di tutti i naturali minori o uguali a n che hanno la stessa parità di n . Es: $6!! = 2 \cdot 4 \cdot 6$, $7!! = 3 \cdot 5 \cdot 7$. Nel caso pari si ha $(2n)!! = 2^n n!$. (Sugg: scrivere $e^{-x^2} x^{2n} = (e^{-x^2} x) x^{2n-1}$ e ottenere una relazione di ricorrenza integrando per parti).

1.3.5 Funzioni continue, uniformemente continue, Lipschitz

1.22 Definizione. Una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è detta *Lipschitz* se esiste una costante $C > 0$ tale che

$$\forall x, y \in \mathbb{R} \text{ si ha } |f(y) - f(x)| \leq C |x - y|$$

L'insieme di tutte le funzioni Lipschitz su \mathbb{R} si denota con $\text{Lip}(\mathbb{R})$.

1.23 Problema. Trovare un esempio di una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ continua ma non uniformemente continua.

1.24 Problema. Trovare un esempio di una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ uniformemente continua ma non Lipschitz.

1.25 Problema. Dimostrare che la funzione reale

$$f(x) = \frac{1}{1+x^2}$$

è uniformemente continua.

1.26 Problema. Trovare una successione di funzioni $f_n : [0, 1] \rightarrow \mathbb{R}$ continue che convergono (non uniformemente) nell'intervallo $[0, 1]$, e la funzione limite $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ non è continua.

1.27 Problema. Dimostrare che la funzione reale f

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

è continua in $x = 0$.

1.28 Problema. Vero o falso? (se falso fornire un controesempio). Sia f_n una successione di funzioni $f_n : [0, 1] \rightarrow \mathbb{R}$

(a) Se $f_n(x) \rightarrow 0$ per tutti gli $x \in [0, 1]$ allora f_n converge a zero uniformemente

(b) Se le f_n sono continue e $f_n(x) \rightarrow 0$ per tutti gli $x \in [0, 1]$ allora

$$\int_0^1 f_n(x) dx \rightarrow 0$$

(c) Se le f_n sono continue e $f_n \rightarrow 0$ uniformemente, allora

$$\int_0^1 f_n(x) dx \rightarrow 0$$

(d) Se le f_n sono derivabili in $(0, 1)$ e $f_n \rightarrow 0$ uniformemente, allora $f'_n(x) \rightarrow 0$ per ogni $x \in (0, 1)$

(e) Se le f_n sono continue ed esiste $f : [0, 1] \rightarrow \mathbb{R}$ tale che $f_n \rightarrow f$ uniformemente allora f è continua

1.3.6 Formule di Taylor

Se $f \in C^n(a, b)$, $x_0 \in (a, b)$ e m è un intero positivo, $m \leq n$, possiamo scrivere lo sviluppo di Taylor fino all'ordine m centrato in x_0 per ogni $x \in (a, b)$ come

$$(1.37) \quad f(x) = \sum_{k=0}^m \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_m(x)$$

in cui, banalmente, definiamo il resto m -simo dello sviluppo come

$$(1.38) \quad R_m(x) := f(x) - \sum_{k=0}^m \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k.$$

La versione della formula di Taylor col resto di Lagrange ci dà informazioni ulteriori sul resto.

1.29 Teorema. Sia $f \in C^{m+1}(a, b)$ e $x_0 \in (a, b)$. Allora, per ogni $x \in (a, b)$ esiste $\xi(x)$ compreso fra x_0 e x tale che vale la (1.37) con

$$R_m(x) = \frac{f^{(m+1)}(\xi(x))}{(m+1)!} (x - x_0)^{m+1}.$$

L'aspetto insoddisfacente di questo risultato è che non ci dice nulla sul grado di regolarità nella dipendenza di $\xi(x)$ da x . Per questo (e altri) motivi, a volte conviene usare una versione della formula di Taylor con una diversa espressione per il resto.

1.30 Teorema. (Formula di Taylor col resto sotto forma di integrale). Sia $f \in C^n(a, b)$, $x_0 \in (a, b)$ e sia m un intero positivo con $m < n$. Allora per ogni $x \in (a, b)$ vale la (1.37) in cui $R_m(x) = (x - x_0)^{m+1} r_m(x)$ e

$$(1.39) \quad r_m(x) = \int_0^1 \frac{(1-t)^m}{m!} f^{(m+1)}(x_0 + t(x - x_0)) dt.$$

Inoltre $r_m \in C^{n-m-1}(a, b)$.

Schema di dimostrazione. Poiché R_m è la differenza fra una funzione in $C^n(a, b)$ e un polinomio, si ha che $R_m \in C^n(a, b)$. La (1.39) si può dimostrare per induzione su m . Se $m = 0$ la (1.39) diventa

$$(1.40) \quad R_0(x) = (x - x_0) \int_0^1 f'(x_0 + t(x - x_0)) dt.$$

Per verificare questa identità poniamo $y = x_0 + t(x - x_0)$ nell'integrale e otteniamo

$$(x - x_0) \int_0^1 f'(x_0 + t(x - x_0)) dt = \int_{x_0}^x f'(y) dy = f(x) - f(x_0) = R_0(x).$$

Per completare l'induzione rimane da far vedere che $(1.39)_m$ implica $(1.39)_{m+1}$. Usando quindi l'ipotesi induttiva $(1.39)_m$ e integrando per parti si ha

$$(1.41) \quad \begin{aligned} R_m(x) &= (x - x_0)^{m+1} \int_0^1 \frac{(1-t)^m}{m!} f^{(m+1)}(x_0 + t(x - x_0)) dt \\ &= (x - x_0)^{m+1} \left[-\frac{(1-t)^{m+1}}{(m+1)!} f^{(m+1)}(x_0 + t(x - x_0)) \right]_0^1 \\ &\quad + (x - x_0)^{m+2} \int_0^1 \frac{(1-t)^{m+1}}{(m+1)!} f^{(m+2)}(x_0 + t(x - x_0)) dt \\ &= (x - x_0)^{m+1} \frac{1}{(m+1)!} f^{(m+1)}(x_0) \\ &\quad + (x - x_0)^{m+2} \int_0^1 \frac{(1-t)^{m+1}}{(m+1)!} f^{(m+2)}(x_0 + t(x - x_0)) dt \end{aligned}$$

D'altra parte, dalla definizione (1.38) segue che

$$(1.42) \quad R_m(x) = R_{m+1}(x) + \frac{f^{(m+1)}(x_0)}{(m+1)!} (x - x_0)^{m+1}.$$

Confrontando le (1.41) e (1.42) si ottiene

$$R_{m+1}(x) = (x - x_0)^{m+2} \int_0^1 \frac{(1-t)^{m+1}}{(m+1)!} f^{(m+2)}(x_0 + t(x - x_0)) dt.$$

La (1.39) è dunque dimostrata per $m + 1$.

Per dimostrare l'ultima affermazione del teorema è sufficiente osservare che, nell'espressione di r_m data dalla (1.39), siamo nelle condizioni di poter derivare $n - m - 1$ volte sotto il segno di integrale. Quindi $r_m \in C^{n-m-1}$ e si ha

$$r_m^{(j)}(x) := \int_0^1 \frac{t^j (1-t)^m}{m!} f^{(m+1+j)}(x_0 + t(x-x_0)) dt \quad j = 0, 1, \dots, n-m-1.$$

□

1.3.7 Derivare sotto il segno di integrale

Sia f è una funzione di due variabili reali definita su un sottoinsieme G di \mathbb{R}^2

$$f : (t, x) \rightarrow f(t, x) \quad (t, x) \in G.$$

Indichiamo con $\partial_1 f$ e $\partial_2 f$ le derivate parziali rispetto al primo e al secondo argomento. A volte alle variabili viene assegnato un nome "standard" e in quel caso è prassi comune usare il nome della variabile come pedice per indicare il tipo di derivata parziale. Chiamando le variabili (t, x) , quindi

$$\partial_1 f(t, x) = \partial_t f(t, x) \quad \partial_2 f(t, x) = \partial_x f(t, x)$$

Attenzione però che usare ∂_t o ∂_x può creare qualche ambiguità (ad esempio, se scrivo $\partial_t f(2t, x)$).

Iniziamo ad enunciare il risultato che tratta il caso più semplice in cui è possibile derivare sotto il segno di integrale. La dimostrazione si trova in un qualsiasi libro di analisi reale.

1.31 Teorema. *Sia $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ una funzione continua. Se $\partial_2 f$ esiste ed è continua su $[a, b] \times [c, d]$, allora la funzione*

$$g(x) := \int_a^b f(t, x) dt$$

è differenziabile e si ha

$$g'(x) = \int_a^b \partial_2 f(t, x) dt.$$

Se al posto dell'intervallo di integrazione $[a, b]$ abbiamo un intervallo infinito, la situazione è un po' più delicata e il passaggio della derivata dentro l'integrale richiede un qualche tipo di ipotesi supplementare. Una possibilità è data dal seguente

1.32 Teorema. *Sia $f : \mathbb{R} \times [c, d] \rightarrow \mathbb{R}$ una funzione continua tale che:*

(a) *per ogni $x \in [c, d]$ l'integrale improprio*

$$g(x) := \int_{\mathbb{R}} f(t, x) dx$$

è convergente;

(b) *$\partial_2 f$ esiste ed è continua su $\mathbb{R} \times [c, d]$;*

(c) esiste $h : \mathbb{R} \rightarrow \mathbb{R}_+$ tale che $|\partial_2 f(t, x)| \leq h(t)$ per ogni $(t, x) \in \mathbb{R} \times [c, d]$ e

$$(1.43) \quad \int_{\mathbb{R}} h(t) dt < \infty.$$

Allora, g è differenziabile e

$$(1.44) \quad g'(x) = \int_{\mathbb{R}} \partial_2 f(t, x) dt.$$

Dimostrazione. Sia $x \in (c, d)$ e sia

$$\Delta(x, h) = \frac{g(x+h) - g(x)}{h} - \int_{\mathbb{R}} \partial_2 f(t, x) dt.$$

Devo dimostrare che $\lim_{h \rightarrow 0} \Delta(x, h) = 0$. Per ipotesi h è integrabile su \mathbb{R} , quindi, per il Corollario 1.15, per ogni $\varepsilon > 0$ esiste $L > 0$ tale che

$$(1.45) \quad \int_{-\infty}^{-L} h(t) dt + \int_L^{\infty} h(t) dt = \int_{|t| \geq L} h(t) dt < \varepsilon.$$

Allora

$$\begin{aligned} \Delta(x, h) &= \int_{\mathbb{R}} \left(\underbrace{\frac{f(t, x+h) - f(t, x)}{h} - \partial_2 f(t, x)}_{A(t, x, h)} \right) dt \\ &= \Delta_1(x, h) + \Delta_2(x, h). \end{aligned}$$

in cui

$$\Delta_1(x, h) = \int_{|t| \leq L} A(t, x, h) dt \quad \Delta_2(x, h) = \int_{|t| \geq L} A(t, x, h) dt.$$

Dal teorema di Taylor al prim'ordine segue che, per ogni h , esiste $\vartheta(h) \in (0, 1)$ tale che

$$A(x, y, h) = \partial_2 f(t, x + \vartheta(h)h) - \partial_2 f(t, x).$$

Per ipotesi $\partial_2 f$ è continua, quindi è anche uniformemente continua sul compatto $[-L, L] \times [c, d]$. Esiste dunque $\delta > 0$ tale che

$$(1.46) \quad |h| < \delta, |t| \leq L \implies |A(t, x, h)| < \frac{\varepsilon}{2L}.$$

A questo punto possiamo stimare Δ_1 come

$$|\Delta_1(x, h)| \leq \int_{|y| \leq L} |A(x, y, h)| dy \leq \frac{\varepsilon}{2L} 2L = \varepsilon.$$

Per stimare Δ_2 uso l'ipotesi (c) e ottengo

$$|A(t, x, h)| \leq |\partial_2 f(t, x + \vartheta(h)h)| + |\partial_2 f(t, x)| \leq 2h(t),$$

che, grazie alla (1.45), implica

$$|\Delta_2(x, h)| \leq \int_{|t| \geq L} 2h(t) dt \leq 2\varepsilon.$$

Abbiamo dunque ottenuto che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $|h| < \delta$, allora

$$|\Delta(x, h)| \leq |\Delta_1(x, h)| + |\Delta_2(x, h)| \leq 3\varepsilon,$$

vale dire $\lim_{h \rightarrow 0} \Delta(x, h) = 0$ che implica appunto la (1.44). \square

2. Spazi metrici

*We don't need no education
We don't need no thought control
No dark sarcasm in the classroom
Teachers leave them kids alone (P.F.)*

2.1 Definizione ed esempi

2.1 Definizione. Si definisce *spazio metrico* una coppia (X, d) , in cui X è un insieme e d è una funzione

$$d : X \times X \rightarrow \mathbb{R}$$

che associa ad ogni coppia x, y di elementi di X un numero reale non negativo $d(x, y)$ detto *distanza fra x e y* . La funzione d è chiamata *distanza* o *metrica* e deve soddisfare le seguenti condizioni per tutti gli $x, y, z \in X$:

(SM1) (positività) $d(x, y) \geq 0$

(SM2) (non degenerazione) $d(x, y) = 0$ se e solo se $x = y$

(SM3) (simmetria) $d(x, y) = d(y, x)$

(SM4) (disuguaglianza triangolare) $d(x, z) \leq d(x, y) + d(y, z)$.

2.2 Definizione. Dato uno spazio metrico (X, d) , si definisce *diametro* di X la quantità

$$(2.1) \quad \text{diam } X := \sup_{x, y \in X} d(x, y)$$

Il diametro di X è sempre non negativo e può essere finito oppure uguale a $+\infty$. Nel primo caso X è detto *limitato* e nel secondo *illimitato*.

2.3 Esempio. Consideriamo un insieme arbitrario X (finito, infinito numerabile o non numerabile) e poniamo per $x, y \in X$

$$d(x, y) := \begin{cases} 0 & \text{se } x = y \\ 1 & \text{se } x \neq y. \end{cases}$$

Tutte le coppie di punti distinti sono quindi a distanza 1 fra loro. Le proprietà (1) e (2) si verificano immediatamente. Per quanto riguarda la disuguaglianza triangolare, basta considerare i casi possibili:

se $x = z$	$d(x, z) = 0$	$d(x, y) + d(y, z) \geq 0$
se $x \neq z$ e $y = x$	$d(x, z) = 1$	$d(x, y) + d(y, z) = 0 + 1 = 1$
se $x \neq z$ e $y = z$	$d(x, z) = 1$	$d(x, y) + d(y, z) = 1 + 0 = 1$
se $x \neq z$, $y \neq x$ e $y \neq z$	$d(x, z) = 1$	$d(x, y) + d(y, z) = 1 + 1 = 2$

Questa funzione distanza è chiamata la *metrica discreta* nello spazio X .

2.4 Esempio. Sia $X = \mathbb{R}$ l'insieme dei numeri reali (oppure $X = \mathbb{C}$) e sia

$$d(x, y) := |x - y|.$$

La proprietà (SM4) (le altre sono ovvie) discende dalle proprietà della funzione modulo. Infatti si ha:

$$d(x, z) = |x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = d(x, y) + d(y, z).$$

2.5 Esempio. Sia $X = \mathbb{R}^n$ e, dati due punti di \mathbb{R}^n $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, consideriamo le funzioni

$$(2.2) \quad d_1(x, y) := \sum_{i=1}^n |x_i - y_i|$$

$$(2.3) \quad d_2(x, y) := \left[\sum_{i=1}^n |x_i - y_i|^2 \right]^{1/2}$$

$$(2.4) \quad d_\infty(x, y) := \max_{i=1, \dots, n} |x_i - y_i|.$$

Ognuna di queste funzioni è una distanza su \mathbb{R}^n .

2.6 Esempio. Consideriamo ancora $X = \mathbb{R}$ con una distanza diversa da quella usuale. Poniamo

$$\rho(x, y) = \frac{|x - y|}{1 + |x - y|}.$$

La funzione ρ soddisfa evidentemente le proprietà (SM1)–(SM3). Per verificare la (SM4) considero la funzione $g : [0, \infty) \rightarrow [0, \infty)$ definita come

$$g(t) := \frac{t}{1 + t}.$$

La funzione g è crescente, infatti

$$g'(t) = \frac{(1+t) - t}{(1+t)^2} = \frac{1}{(1+t)^2} > 0.$$

Quindi se $0 \leq t \leq u$ si ha $g(t) \leq g(u)$. D'altra parte sappiamo già che

$$|x - z| \leq |x - y| + |y - z|,$$

per cui

$$g(|x - z|) \leq g(|x - y| + |y - z|),$$

che, per definizione di g , significa

$$\frac{|x - z|}{1 + |x - z|} \leq \frac{|x - y| + |y - z|}{1 + |x - y| + |y - z|}.$$

Possiamo quindi scrivere

$$\begin{aligned} \rho(x, z) &= \frac{|x - z|}{1 + |x - z|} \leq \frac{|x - y| + |y - z|}{1 + |x - y| + |y - z|} \\ &= \frac{|x - y|}{1 + |x - y| + |y - z|} + \frac{|y - z|}{1 + |x - y| + |y - z|} \\ &\leq \frac{|x - y|}{1 + |x - y|} + \frac{|y - z|}{1 + |y - z|} = \rho(x, y) + \rho(y, z). \end{aligned}$$

La disuguaglianza triangolare è verificata, quindi ρ è una metrica su \mathbb{R} . Osservo che mentre lo spazio metrico (\mathbb{R}, d) (quello con la metrica usuale) è illimitato, lo stesso insieme con la distanza ρ diventa uno spazio metrico limitato. In particolare il diametro di (\mathbb{R}, ρ) è uguale a 1 (dimostrarlo).

2.7 Esempio. (Un esempio “infinito dimensionale”). Consideriamo lo spazio \mathbb{R}^∞ di tutte le successioni reali e, date due successioni

$$x = (x_1, x_2, \dots) \quad y = (y_1, y_2, \dots)$$

definiamo la loro distanza come

$$(2.5) \quad d(x, y) := \sum_{i=1}^{\infty} \frac{1}{2^i} \rho(x_i, y_i),$$

in cui ρ è la metrica introdotta nell'Esempio 2.6

$$\rho(x_i, y_i) = \frac{|x_i - y_i|}{1 + |x_i - y_i|}.$$

La funzione d è ben definita. Infatti $\rho(x_i, y_i) < 1$ per cui la sommatoria (2.5) è sempre convergente. È facile vedere che d è una metrica. Dimostro solo la disuguaglianza triangolare (le altre proprietà sono ovvie). Nell'Esempio 2.6 abbiamo fatto vedere che

$$\rho(x_i, z_i) \leq \rho(x_i, y_i) + \rho(y_i, z_i),$$

quindi ottengo

$$d(x, z) := \sum_{i=1}^{\infty} \frac{1}{2^i} \rho(x_i, z_i) \leq \sum_{i=1}^{\infty} \frac{1}{2^i} \rho(x_i, y_i) + \sum_{i=1}^{\infty} \frac{1}{2^i} \rho(y_i, z_i) =: d(x, y) + d(y, z).$$

2.8 Definizione. Se (X, d) è uno spazio metrico, $x \in X$ e $r > 0$, definiamo

$$B_r^X(x) := \{u \in X : d(u, x) < r\}$$

$$\bar{B}_r^X(x) := \{u \in X : d(u, x) \leq r\}$$

Questi insiemi sono chiamati rispettivamente la *palla aperta* e la *palla chiusa* di centro x e raggio r . Quando non c'è pericolo di confusione ometto il soprascritto X e indico la palla (ad esempio) aperta con $B_r(x)$.

2.9 Definizione. Sia (X, d) uno spazio metrico. Un sottoinsieme $A \subset X$ è detto *aperto in X* se per ogni $x \in A$ esiste $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset A$. Un sottoinsieme $B \subset X$ è detto *chiuso in X* se il suo complemento $B^c = X \setminus B$ è aperto.

2.10 Proposizione. *Una palla aperta in uno spazio metrico (X, d) è aperta. Una palla chiusa è chiusa.*

Dimostrazione. Sia $A = B_r^X(x_0)$. Voglio far vedere che A è aperto in X , vale a dire che comunque scelgo $x \in A$ posso sempre trovare $\varepsilon > 0$ tale che $B_\varepsilon^X(x) \subset A$. Sia infatti $x \in A$. Per definizione di A abbiamo che $d(x, x_0) < r$. Poniamo allora

$$\varepsilon := r - d(x, x_0) > 0.$$

A questo punto osservo che se $y \in B_\varepsilon^X(x)$ si ha $d(y, x) < \varepsilon$. Ma allora

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = r,$$

quindi $y \in B_r^X(x_0)$. Abbiamo dimostrato che

$$y \in B_\varepsilon^X(x) \Rightarrow y \in B_r^X(x_0)$$

cioè che $B_\varepsilon^X(x) \subset A$, dunque A è aperto.

La dimostrazione che una palla chiusa è chiusa è lasciata al lettore. □

2.11 Proposizione. *Sia (X, d) uno spazio metrico. Allora:*

- (1) *Gli insiemi \emptyset e X sono aperti*
- (2) *Se (A_1, \dots, A_n) è una collezione finita di insiemi aperti allora $\bigcap_{i=1}^n A_i$ è aperto*
- (3) *Se $(A_\alpha)_{\alpha \in I}$ è una collezione arbitraria di insiemi aperti allora $\bigcup_{\alpha \in I} A_\alpha$ è aperto.*

Dimostrazione. Il punto (1) è ovvio.

Punto (2). Sia $C := \bigcap_{i=1}^n A_i$. Per far vedere che C è aperto devo dimostrare che

$$(2.6) \quad \text{per ogni } x \in C \text{ esiste } \varepsilon > 0 \text{ tale che } B_\varepsilon(x) \subset C.$$

Sia allora $x \in C$. x appartiene a tutti gli A_i . Siccome ciascuno degli A_i è aperto so che posso trovare, per ogni i , un numero reale positivo ε_i tale che $B_{\varepsilon_i}(x) \subset A_i$. Scelgo allora

$$\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\}.$$

Poiché $\varepsilon \leq \varepsilon_i$ ottengo che

$$B_\varepsilon(x) \subset A_i \quad \forall i = 1, \dots, n,$$

il che equivale a dire che $B_\varepsilon(x) \subset \bigcap_{i=1}^n A_i = C$.

Punto (3). Facile, lasciato al lettore □

2.12 Proposizione. *Sia (X, d) uno spazio metrico. Allora:*

- (1) *Gli insiemi \emptyset e X sono chiusi*
- (2) *Se (A_1, \dots, A_n) è una collezione finita di insiemi chiusi allora $\bigcup_{i=1}^n A_i$ è chiuso*
- (3) *Se $(A_\alpha)_{\alpha \in I}$ è una collezione arbitraria di insiemi chiusi allora $\bigcap_{\alpha \in I} A_\alpha$ è chiuso.*

Dimostrazione. Segue dalla Proposizione 2.11 e dalle leggi di De Morgan □

2.13 Definizione. Sia (X, d) uno spazio metrico e sia $A \subset X$. Definisco

$$(2.7) \quad A^\circ = \text{“la parte interna di } A\text{”} := \bigcup_{G \subset A, G \text{ è aperto}} G$$

$$(2.8) \quad \bar{A} = \text{“la chiusura di } A\text{”} := \bigcap_{K \supset A, K \text{ è chiuso}} K$$

$$(2.9) \quad \partial A = \text{“il bordo di } A\text{”} := \bar{A} \setminus A^\circ$$

Detto informalmente A° è “il più grande” insieme aperto contenuto in A , mentre \bar{A} è “il più piccolo” insieme chiuso che contiene A .

2.14 Proposizione. *Siano A e B sottoinsiemi di uno spazio metrico (X, d) . Allora:*

- (1) *A è aperto se e solo se $A = A^\circ$.*
- (2) *A è chiuso se e solo se $A = \bar{A}$.*
- (3) *$A^\circ = (\bar{A}^c)^c$.*
- (4) *$\bar{A} = [(A^\circ)^c]^c$.*
- (5) *$(A \cap B)^\circ = A^\circ \cap B^\circ$.*
- (6) *$\overline{A \cup B} = \bar{A} \cup \bar{B}$.*

Dimostrazione. Punto (1). Sia $\mathcal{O}(A)$ la collezione di tutti gli insiemi aperti contenuti in A . Se $A = A^\circ$ segue dalla (2.7) che

$$A = \bigcup_{G \in \mathcal{O}(A)} G,$$

quindi A è unione di aperti. Di conseguenza, per la Proposizione 2.11, A è aperto.

Supponiamo viceversa che A sia aperto. In questo caso A è un elemento della collezione $\mathcal{O}(A)$, quindi $A^\circ = A$.

Punto (2). Simile al punto (1).

Punto (3). Utilizzando le leggi di De Morgan e la definizione di chiusura ottengo

$$(\overline{A^c})^c = \left(\bigcap_{K \supset A^c, K \text{ è chiuso}} K \right)^c = \bigcup_{K \supset A^c, K \text{ è chiuso}} K^c.$$

Poiché la corrispondenza $K \rightarrow K^c$ è biunivoca possiamo “cambiare variabile” ponendo $G = K^c$, per cui si ha

$$(\overline{A^c})^c = \bigcup_{G^c \supset A^c, G^c \text{ è chiuso}} G = \bigcup_{G \subset A, G \text{ è aperto}} G = A^\circ. \quad \square$$

Punto (4). Segue dal punto (3).

Punto (5). Poiché $(A \cap B) \subset A$ abbiamo $(A \cap B)^\circ \subset A^\circ$. Analogamente $(A \cap B) \subset B$, quindi $(A \cap B)^\circ \subset B^\circ$. Abbiamo ottenuto

$$(A \cap B)^\circ \subset A^\circ \cap B^\circ.$$

Dimostro ora l’inclusione inversa $A^\circ \cap B^\circ \subset (A \cap B)^\circ$. L’insieme $A^\circ \cap B^\circ$ è un aperto contenuto in $A \cap B$. Ma $(A \cap B)^\circ$ è l’unione di *tutti* gli aperti contenuti in $A \cap B$, quindi $A^\circ \cap B^\circ \subset (A \cap B)^\circ$.

Punto (6). Usando i punti (4) e (5) e le leggi di De Morgan si ottiene

$$\begin{aligned} \overline{A \cup B} &= ([A \cup B]^\circ)^c = ([A^c \cap B^c]^\circ)^c = ([A^c]^\circ \cap [B^c]^\circ)^c \\ &= ([A^c]^\circ)^c \cup ([B^c]^\circ)^c = \overline{A} \cup \overline{B}. \quad \square \end{aligned}$$

2.15 Proposizione. *Sia (X, d) uno spazio metrico, sia $A \subset X$ e sia x un punto qualsiasi di X . Allora*

- (1) $x \in A^\circ$ se e solo se esiste un $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset A$
- (2) $x \in \overline{A}$ se e solo se per ogni $\varepsilon > 0$ si ha $B_\varepsilon(x) \cap A \neq \emptyset$
- (3) $x \in \partial A$ se e solo se per ogni $\varepsilon > 0$ si ha che la palla $B_\varepsilon(x)$ ha intersezione non vuota sia con A sia con A^c .

Dimostrazione. Punto (1). Sia

$$I(A) := \{x \in X : \exists \varepsilon > 0 \text{ tale che } B_\varepsilon(x) \subset A\}.$$

Devo far vedere che

$$I(A) = A^\circ := \bigcup_{G \subset A, G \text{ è aperto}} G.$$

A questo scopo dimostro le due inclusioni

$$(2.10) \quad I(A) \subset A^\circ \quad I(A) \supset A^\circ.$$

Sia $x \in I(A)$. Allora esiste $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset A$. Sappiamo che $B_\varepsilon(x)$ è aperto, quindi x è contenuto in un aperto che è un sottoinsieme di A . A maggior ragione

$$x \in \bigcup_{G \subset A, G \text{ è aperto}} G =: A^\circ.$$

Dunque $I(A) \subset A^\circ$. Ora procedo con l'inclusione inversa. Sia $x \in A^\circ$. Per definizione esiste G , un aperto di X , tale che $x \in G$. Ma per definizione di aperto esiste $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset G$. Siccome $G \subset A$ si ha che $B_\varepsilon(x) \subset A$. Di conseguenza $x \in I(A)$. Ho fatto vedere che $I(A) \supset A^\circ$, che, unitamente all'inclusione inversa, dice che $I(A) = A^\circ$.

Punto (2). Dalla Proposizione 2.14 sappiamo che $\bar{A} = [(A^c)^\circ]^c$. Quindi $x \in \bar{A}$ se e solo se $x \notin (A^c)^\circ$. Grazie al punto (1) questo equivale a dire che

$$\text{per ogni } \varepsilon > 0 \text{ si ha } B_\varepsilon(x) \not\subset A^c$$

o, equivalentemente,

$$\text{per ogni } \varepsilon > 0 \text{ si ha } B_\varepsilon(x) \cap A \neq \emptyset$$

Punto (3). Per definizione $x \in \partial A$ se e solo se $x \in \bar{A} \setminus A^\circ$ vale a dire $x \in \bar{A}$ e $x \notin A^\circ$.

Usando i punti (1) e (2) otteniamo

$$\begin{aligned} x \in \bar{A} &\Leftrightarrow \text{per ogni } \varepsilon > 0 \text{ si ha } B_\varepsilon(x) \cap A \neq \emptyset \\ x \notin A^\circ &\Leftrightarrow \text{per ogni } \varepsilon > 0 \text{ si ha } B_\varepsilon(x) \cap A^c \neq \emptyset \end{aligned}$$

□

2.16 Definizione. Siano (X, d) e (Y, ρ) due spazi metrici. La funzione $f : X \rightarrow Y$ è detta un'isometria se preserva la distanza, vale a dire se

$$\rho(f(x_1), f(x_2)) = d(x_1, x_2) \quad \forall x_1, x_2 \in X.$$

2.17 Osservazione. Un'isometria f da (X, d) a (Y, ρ) è banalmente iniettiva. Infatti se $f(x_1) = f(x_2)$, allora $\rho(f(x_1), f(x_2)) = 0$. Questo implica $d(x_1, x_2) = 0$, che, per le proprietà della distanza a sua volta implica $x_1 = x_2$. Un'isometria non è necessariamente suriettiva. Siano ad esempio $X = [0, 1]$ e $Y = [0, 2]$, entrambi con la distanza usuale e sia $f : X \rightarrow Y$ definita come $f(x) = x$. f è un'isometria, ma ovviamente non è suriettiva.

2.18 Problema. Dimostrare le seguenti affermazioni (se l'affermazione è vera), o esibire un controesempio (se è falsa)

- (a) Un insieme aperto in un qualsiasi spazio metrico ha cardinalità infinita (cioè possiede un numero infinito di punti).
- (b) $(A \cup B)^\circ = A^\circ \cup B^\circ$.
- (c) $\overline{A \cap B} = \overline{A} \cap \overline{B}$.
- (d) L'unione di un'infinità numerabile di insiemi chiusi è chiusa.
- (e) $\text{diam } A = \text{diam}(A^\circ)$.

2.19 Attenzione! La parte interna, il bordo e la chiusura di un insieme A sono concetti *relativi* ad uno spazio metrico ambiente (X, d) . Per evitare ambiguità bisognerebbe dire ad esempio la *chiusura di A in (X, d)* . Quando lo spazio ambiente è ovvio questa precisazione può essere omessa. Consideriamo due diversi spazi ambiente

$$\begin{aligned} X &:= \mathbb{R}^3 \\ Y &:= \{x \in \mathbb{R}^3 : x_3 = 0 \text{ e } x_2 < 1\}, \end{aligned}$$

e consideriamo un insieme A che può essere pensato come un sottoinsieme sia di X che di Y

$$A := \{x \in \mathbb{R}^3 : x_3 = 0, 0 \leq x_1 < 1 \text{ e } 0 \leq x_2 < 1\}.$$

In questo caso si ha

$$\begin{aligned} A^\circ &= \{x \in \mathbb{R}^3 : x_3 = 0, 0 < x_1 < 1 \text{ e } 0 < x_2 < 1\} && \text{in } Y \\ A^\circ &= \emptyset && \text{in } X \\ \bar{A} &= \{x \in \mathbb{R}^3 : x_3 = 0, 0 \leq x_1 \leq 1 \text{ e } 0 \leq x_2 < 1\} && \text{in } Y \\ \bar{A} &= \{x \in \mathbb{R}^3 : x_3 = 0, 0 \leq x_1 \leq 1 \text{ e } 0 \leq x_2 \leq 1\} && \text{in } X \end{aligned}$$

Anche gli attributi *aperto* e *chiuso* sono relativo allo spazio ambiente. Ad esempio l'insieme

$$\{x \in \mathbb{R}^3 : x_3 = 0, 0 < x_1 < 1 \text{ e } 0 < x_2 < 1\}$$

è aperto in Y ma non è aperto in X (e neppure chiuso), mentre l'insieme

$$\{x \in \mathbb{R}^3 : x_3 = 0, 0 \leq x_1 \leq 1 \text{ e } 0 \leq x_2 < 1\}$$

è chiuso in Y ma non in X .

2.20 Problema. Sia $X := \mathbb{R}^2$, $Y := \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\}$ e $Z := \{x = (x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 < 1\}$. Consideriamo le distanze

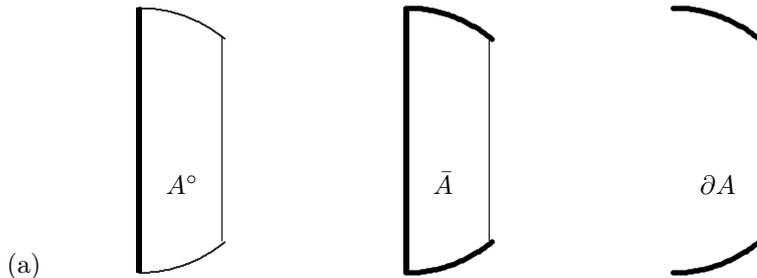
$$\begin{aligned} d_1(x, y) &:= \sum_{i=1}^2 |x_i - y_i| && d_2(x, y) := \left[\sum_{i=1}^2 |x_i - y_i|^2 \right]^{1/2} \\ d_\infty(x, y) &:= \max_{i=1,2} |x_i - y_i|. \end{aligned}$$

- Disegnare gli insiemi Y e Z .
- Disegnare la palla unitaria $B_1(0)$ negli spazi metrici (X, d_∞) , (Y, d_2) , (Z, d_1) .
- Nello spazio metrico (Y, d_2) disegnare $B_1((1, 0))$.
- Nello spazio metrico (X, d_2) , sia $A := \{x \in \mathbb{R}^2 : x_1 \in \mathbb{Q}\}$. Determinare A° , \bar{A} e ∂A .
- Nello spazio metrico (Z, d_2) , sia $A := \{x \in \mathbb{R}^2 : 0 \leq x_1 < 1, 0 \leq x_2 < 1\}$. Determinare A° , \bar{A} e ∂A .

2.21 Problema. Consideriamo lo spazio metrico (X, d) , in cui $X := \{x = (x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 < 1\}$ e d è la usuale distanza euclidea in \mathbb{R}^2 . Sia $A := \{x \in X : d(x, 0) < 2\}$.

- Disegnare in modo chiaro ed accurato gli insiemi A° , \bar{A} e ∂A .
- Dire se (X, d) è uno spazio metrico completo, giustificando la risposta.
- Fare un esempio di un insieme $C \subset X$ tale che C è chiuso in X ma C non è chiuso in \mathbb{R}^2 .

Soluzione.



- (b) Un sottoinsieme X di uno spazio metrico completo (Y, d) è completo se e solo X è chiuso in Y . Nel nostro caso (X, d) non è completo perché non è chiuso in \mathbb{R}^2 .
- (c) Si può prendere ad esempio $C := \bar{A}$ oppure $C = X$ (o mille altre cose).

2.22 Problema. Dimostrare con un controesempio che non è sempre vero che la chiusura di una palla aperta coincide con la palla chiusa corrispondente. Trovare quindi uno spazio metrico (X, d) , un punto $x \in X$ e un valore di $r > 0$ tale che $\overline{B_r(x)} \neq \bar{B}_r(x)$.

Soluzione. Si trovano facilmente infiniti controesempi. Eccone uno: sia $X = \mathbb{Z}$ con la distanza usuale. Si ha

$$B_1(0) = \{0\} \quad \overline{B_1(0)} = \overline{\{0\}} = \{0\} \quad \bar{B}_1(0) = \{-1, 0, 1\}.$$

2.23 Problema. Sia d la usuale distanza euclidea in \mathbb{R}^2 . Consideriamo lo spazio metrico

$$X := \{x \in \mathbb{R}^2 : d(x, (0, 0)) \leq 1\} \cup \{x \in \mathbb{R}^2 : d(x, (3, 0)) < 1\}.$$

- (a) Disegnare X e la palla $B_{1/2}^X((0, 1))$.
- (b) Sia $A := \{x \in X : 5/2 < x_1 < 7/2\}$. Si disegni \bar{A} evidenziando il bordo di A , ∂A .
- (c) Trovare un sottoinsieme F di X che non sia nè X nè l'insieme vuoto e tale che F sia contemporaneamente aperto e chiuso in X .

2.2 Convergenza, completezza

2.24 Definizione. Sia (X, d) uno spazio metrico e sia x un punto di X . Una successione $(x_k)_{k=1}^\infty$ di elementi di X si dice *convergente ad x* se

$$\lim_{k \rightarrow \infty} d(x_k, x) = 0,$$

vale a dire, detto in modo più esplicito, se

$$\forall \varepsilon > 0 \exists N \text{ intero positivo tale che } \forall k \geq N \text{ si ha } d(x_k, x) < \varepsilon.$$

In questo caso si scrive, simbolicamente,

$$\lim_{k \rightarrow \infty} x_k = x \quad \text{oppure} \quad x_k \rightarrow x.$$

2.25 Problema. Dimostrare che in uno spazio metrico il limite è unico, vale a dire se $x_n \rightarrow x$ e $x_n \rightarrow y$ allora $x = y$.

2.26 Definizione. Sia (X, d) uno spazio metrico e sia A un sottoinsieme di X . Un punto $x \in X$ si dice *punto di aderenza di A* se esiste una successione $(x_k)_{k=1}^\infty$ di elementi di A tale che $x_k \rightarrow x$.

2.27 Definizione. Sia (X, d) uno spazio metrico e sia A un sottoinsieme di X . Un punto $x \in X$ si dice *punto di accumulazione (o punto limite) di A* se esiste una successione $(x_k)_{k=1}^\infty$ di elementi di A tale che $x_k \neq x$ e $x_k \rightarrow x$.

2.28 Proposizione. Sia (X, d) uno spazio metrico e sia A un sottoinsieme di X . Allora \bar{A} coincide con l'insieme dei punti di aderenza di A .

Dimostrazione. Chiamo Y l'insieme dei punti di aderenza di A . Voglio far vedere che $Y = \bar{A}$. Dimostro prima che $\bar{A} \subset Y$. Supponiamo che $x \in \bar{A}$. Per la Proposizione 2.15 questo significa che

$$(2.11) \quad \text{per ogni } \varepsilon > 0 \text{ si ha } B_\varepsilon(x) \cap A \neq \emptyset.$$

Ma allora per ogni intero positivo n si ha $B_{1/n}(x) \cap A \neq \emptyset$. Quindi per ogni n posso trovare un punto $x_n \in B_{1/n}(x) \cap A$. Poiché $d(x, x_n) < 1/n$, ottengo che $x_n \rightarrow x$. Ho trovato una successione di elementi di A che tende ad x . Per definizione x è un punto di aderenza di A . Ho dimostrato che $\bar{A} \subset Y$.

Voglio dimostrare ora l'inclusione inversa $Y \subset \bar{A}$. Suppongo che $x \in Y$, vale a dire che x sia un punto di aderenza di A . Per definizione posso trovare una successione (x_n) di elementi di A tale che $x_n \rightarrow x$. Per definizione di limite questo significa che

$$\forall \varepsilon > 0 \exists N > 0 \text{ tale che } \forall n \geq N \ d(x, x_n) < \varepsilon,$$

vale a dire che tutti gli x_n con $n \geq N$ appartengono alla palla $B_\varepsilon(x)$. In particolare $x_N \in B_\varepsilon(x)$. Ma x_N è anche un elemento di A , dunque

$$\forall \varepsilon > 0 \exists N \text{ tale che } x_N \in B_\varepsilon(x) \cap A.$$

Di conseguenza $B_\varepsilon(x) \cap A \neq \emptyset$ il che implica, per la Proposizione 2.15, che $x \in \bar{A}$. Abbiamo dimostrato che $Y = \bar{A}$. \square

2.29 Proposizione. *Sia (X, d) uno spazio metrico e sia A un sottoinsieme di X . Allora A è chiuso se e solo se vale l'implicazione*

$$(2.12) \quad (x_n \in A, x_n \rightarrow x) \Rightarrow x \in A.$$

In altre parole un insieme A è chiuso se e solo se contiene tutti i limiti delle successioni convergenti che si possono costruire con elementi di A .

Dimostrazione. Supponiamo che A sia chiuso. Voglio far vedere che vale la (2.12). Suppongo quindi di avere una successione (x_n) di elementi di A che converge a x . Per definizione questo significa che x è un punto di aderenza di A . Ma per la Proposizione 2.28 $x \in \bar{A}$. Osservo infine che se A è chiuso si ha $\bar{A} = A$ e dunque $x \in A$. Quindi vale la (2.12).

Dimostro ora l'implicazione inversa, vale a dire se vale la (2.12) allora A è chiuso. Sia x un punto di aderenza di A . Per definizione esiste una successione (x_n) di elementi di A tale che $x_n \rightarrow x$. Grazie alla (2.12) $x \in A$. Quindi tutti i punti di aderenza di A appartengono ad A , di conseguenza $A = \bar{A}$ il che implica che A è chiuso. \square

2.30 Definizione. Sia (X, d) uno spazio metrico e siano A e B due sottoinsiemi di X tali che $A \subset B$. L'insieme A si dice *denso in B* se $\bar{A} \supset B$, vale a dire se per ogni $x \in B$ e per ogni $\varepsilon > 0$ posso trovare un elemento $y \in A$ tale che $d(x, y) < \varepsilon$.

2.31 Esempio. L'insieme dei razionali \mathbb{Q} è denso in \mathbb{R} (rispetto alla metrica usuale). Infatti se x è un numero reale e $\varepsilon > 0$ c'è sempre un razionale¹ nell'intervallo $(x - \varepsilon, x + \varepsilon)$.

2.32 Definizione. Sia (X, d) uno spazio metrico. Una successione $(x_k)_{k=1}^\infty$ di elementi di X si dice *di Cauchy* se per ogni $\varepsilon > 0$ esiste un intero positivo N tale che per ogni $k, n \geq N$ si ha $d(x_k, x_n) < \varepsilon$.

¹ce ne sono infiniti

2.33 Proposizione. *Ogni successione convergente è di Cauchy.*

Dimostrazione. La dimostrazione è identica al caso delle successioni reali.

2.34 Definizione. Uno spazio metrico (X, d) si dice *completo* se ogni successione di Cauchy è convergente.

2.35 Proposizione. *Gli spazi metrici \mathbb{R} , \mathbb{R}^n , \mathbb{C} , \mathbb{C}^n sono completi rispetto alla metrica usuale.*

Dimostrazione. La completezza di \mathbb{R} è nota dai corsi di analisi. La completezza degli altri spazi segue facilmente da quella di \mathbb{R} . Consideriamo il caso di \mathbb{R}^n . Sia

$$(2.13) \quad x^{(1)}, x^{(2)}, x^{(3)}, \dots$$

una successione di Cauchy di punti in \mathbb{R}^n . Ognuno di questi punti è un'ennupla, quindi

$$x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)}).$$

Poiché la successione (2.13) è Cauchy, per definizione si ha che

$$(2.14) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che se } j, k \geq N \text{ si ha } d(x^{(j)}, x^{(k)}) < \varepsilon.$$

Consideriamo ora la successione reale costituita dalle coordinate i -sime

$$(2.15) \quad x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, \dots$$

Siccome

$$d(x^{(j)}, x^{(k)}) := \sqrt{\sum_{i=1}^n (x_i^{(j)} - x_i^{(k)})^2} \geq |x_i^{(j)} - x_i^{(k)}|$$

dalla (2.14) segue che la successione reale (2.15) è anch'essa di Cauchy. Poiché \mathbb{R} è completo, la successione (2.15) è convergente, quindi esiste $y_i \in \mathbb{R}$ tale che

$$\lim_{k \rightarrow \infty} x_i^{(k)} = y_i \quad \forall i = 1, \dots, n.$$

Di conseguenza si ottiene che $d(x^{(k)}, y) \rightarrow 0$, ovvero che la successione (2.13) è convergente a y □

2.36 Osservazione. Non tutti gli spazi metrici sono completi. Un modo banale per ottenere uno spazio metrico non completo è considerare un qualsiasi spazio metrico (X, d) e scegliere un sottoinsieme A di X che non sia chiuso. Allora lo spazio metrico (A, d) non è completo. Ad esempio si prenda $X = \mathbb{R}$ e $A = (0, 1]$. È ovvio che A non è completo poiché la successione $x_n = 1/n$ è di Cauchy ma non è convergente. D'altra parte non è detto in generale che un sottoinsieme chiuso costituisca automaticamente uno spazio metrico completo. Sia $X = \mathbb{Q}$ e $B = [0, 1] \cap \mathbb{Q}$. L'insieme B considerato come sottoinsieme di \mathbb{Q} è chiuso ma non è completo. È chiaro che, se considero B un sottoinsieme non di \mathbb{Q} ma di \mathbb{R} allora B non è chiuso ma la sua chiusura in \mathbb{R} è uguale a $\bar{B} = [0, 1]$. Le cose diventano più semplici se ci si limita a considerare sottoinsiemi di spazi metrici completi. In questo caso la nozione di completezza coincide con quella di chiusura, come afferma la Proposizione seguente.

2.37 Proposizione. *Sia (X, d) uno spazio metrico completo e sia $Y \subset X$. Lo spazio metrico (Y, d) è completo se e solo se Y è chiuso in X .*

Dimostrazione. Sia (X, d) uno spazio metrico completo e sia Y un sottoinsieme di X . Se (Y, d) è completo vale l'implicazione

$$(2.16) \quad x_n \in Y \text{ e } (x_n) \text{ è di Cauchy} \quad \Rightarrow \quad \exists y \in Y \text{ tale che } x_n \rightarrow y.$$

Voglio far vedere che Y è chiuso, dimostrando che Y contiene tutti i suoi punti di aderenza. Se z è un punto di aderenza di Y allora esiste una successione (x_n) di elementi di Y tale che $x_n \rightarrow z$. Sappiamo che una successione convergente è necessariamente di Cauchy, quindi (x_n) è di Cauchy. Ma allora, per la (2.16) posso trovare $y \in Y$ tale che $x_n \rightarrow y$. Per l'unicità del limite deve essere $y = z$, quindi $z \in Y$. Dunque Y contiene tutti i suoi punti di aderenza, il che vuol dire che Y è chiuso.

Suppongo viceversa che Y sia chiuso e dimostro che (Y, d) è completo. Per far ciò devo far vedere che tutte le successioni di Cauchy di elementi di Y sono convergenti ad un elemento di Y . Sia dunque (x_n) una successione di Cauchy di elementi di Y . Poiché (X, d) è completo so che esiste $x \in X$ tale che $x_n \rightarrow x$. Quindi x è un punto di aderenza di Y . Ma stiamo assumendo che Y sia chiuso, quindi Y contiene tutti i suoi punti di aderenza, dunque $x \in Y$. Ho così dimostrato che Y è completo. \square

2.38 Problema. Dimostrare che se (x_n) è una successione di Cauchy in uno spazio metrico (X, d) che ammette una sottosuccessione (x_{n_k}) convergente a $x \in X$, allora (x_n) è anch'essa convergente a x .

2.39 Problema. Sia $X = \mathbb{R}^2$ e poniamo $d(x, y)$ uguale al numero degli indici i tali che $x_i \neq y_i$.

- (a) Disegnare $B_{1.5}((0, 0))$, $B_{0.5}((0, 0))$.
- (b) Dire se la successione $x_n = (1/n, 0)$ è convergente a zero.
- (c) Sia $A := \{x = (x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$. Determinare A° .
- (d) Elencare tutti gli insiemi densi in X .

2.40 Problema. Dimostrare che \mathbb{Q} con la distanza usuale non è completo trovando esplicitamente una successione di numeri razionali che converge ad un numero non razionale.

2.41 Problema. Sia $\rho(s, t)$ la metrica discreta su \mathbb{R} . Dati due punti di \mathbb{R}^2 , $x = (x_1, x_2)$ e $y = (y_1, y_2)$ definisco

$$d(x, y) = \rho(x_1, y_1) + |x_2 - y_2|.$$

- (a) Dimostrare che d è una metrica su \mathbb{R}^2 .
- (b) Disegnare $B_{1/2}((0, 0))$.
- (c) Sia $A := \{x \in \mathbb{R}^2 : 0 < x_1 < 1, 0 < x_2 < 1\}$. Determinare \bar{A} .

✂ 2.3 Teorema della categoria di Baire

Il teorema della categoria di Baire è alla base di alcuni risultati fondamentali dell'analisi funzionale, come il teorema dell'applicazione aperta, il teorema sull'operatore inverso 7.21 e il teorema del grafo chiuso. Iniziamo con un risultato preliminare che può essere pensato come un lemma del teorema della categoria di Baire.

2.42 Teorema. (*Teorema delle sfere incluse*). Sia (X, d) uno spazio metrico. Le seguenti affermazioni sono equivalenti:

- (1) X è completo;

(2) Sia B_1, B_2, \dots una successione di palle chiuse i cui raggi tendono a zero e tali che $B_1 \supset B_2 \supset \dots$. Allora esiste $x \in X$ tale che

$$\bigcap_{i=1}^{\infty} B_i = \{x\}.$$

Dimostrazione. (1) \Rightarrow (2). Sia (X, d) completo, siano x_n e r_n rispettivamente il centro e il raggio della palla B_n . Per ipotesi $r_n \rightarrow 0$. Affermo che la successione (x_n) è di Cauchy. Infatti sia $\varepsilon > 0$. Siccome $r_n \rightarrow 0$, esiste N tale $r_N < \varepsilon/2$. Ora, se $n, k \geq N$, si ha

$$x_n \in B_n \subset B_N \qquad x_k \in B_k \subset B_N,$$

quindi x_n e x_k appartengono entrambi a B_N , di conseguenza $d(x_n, x_k) \leq \text{diam } B_N = 2r_N < \varepsilon$. Quindi (x_n) è una successione di Cauchy. Poiché X è completo esiste $x \in X$ tale che $x_n \rightarrow x$. Affermo che

$$(2.17) \qquad x \in B_n \qquad \forall n \in \mathbb{N}^*.$$

Se n è un intero positivo, la successione

$$x_n, x_{n+1}, x_{n+2}, \dots$$

è contenuta nella palla B_n e inoltre converge a x . Ma B_n è chiuso, quindi la Proposizione 2.29 ci assicura che $x \in B_n$. Dunque la (2.17) è dimostrata, vale a dire $x \in \bigcap_{n=1}^{\infty} B_n$. Per far vedere che x è l'unico punto contenuto nell'intersezione di tutti i B_n , sia $y \in \bigcap_{n=1}^{\infty} B_n$. Ma allora x e y appartengono entrambi a B_n , quindi

$$d(x, y) \leq \text{diam } B_n = 2r_n \qquad \forall n \in \mathbb{N}^*.$$

Facendo il limite $n \rightarrow \infty$ nella precedente disuguaglianza ottengo $d(x, y) = 0$, ovvero $x = y$.

(2) \Rightarrow (1). Supponiamo ora che valga (2) e sia $(x_n)_{n=1}^{\infty}$ una successione di Cauchy in X . Voglio far vedere che essa ammette un limite. L'idea è di usare il risultato del Problema 2.38, cioè quella di costruire una sottosuccessione $(x_{j(k)})$ che sia convergente. La costruzione è la seguente. Siccome (x_n) è di Cauchy posso trovare una successione di interi positivi $j(1), j(2), j(3), \dots$ con la seguente proprietà

$$(2.18) \quad \begin{array}{ll} j(1) > 0 & d(x_{j(1)}, x_n) < 1/2^1 \quad \forall n \geq j(1) \\ j(2) > j(1) & d(x_{j(2)}, x_n) < 1/2^2 \quad \forall n \geq j(2) \\ \dots & \\ j(k) > j(k-1) & d(x_{j(k)}, x_n) < 1/2^k \quad \forall n \geq j(k) \\ \dots & \end{array}$$

A questo punto voglio costruire una successione di palle chiuse per sfruttare l'ipotesi (2) Poniamo

$$(2.19) \qquad B_k := \overline{B}_{r_k}(x_{j(k)}) \qquad r_k := \frac{1}{2^{k-1}}.$$

È ovvio che $r_k \rightarrow 0$. Devo far vedere che per ogni k si ha $B_{k+1} \subset B_k$. Sia $y \in B_{k+1}$. Si ha

$$\begin{aligned} d(y, x_{j(k)}) &\leq d(y, x_{j(k+1)}) + d(x_{j(k+1)}, x_{j(k)}) \\ &\leq r_{k+1} + \frac{1}{2^k} && \text{[uso: (2.18)]} \\ &\leq \frac{1}{2^k} + \frac{1}{2^k} = \frac{1}{2^{k-1}} = r_k. \end{aligned}$$

Quindi $y \in B_k$. Abbiamo così dimostrato che $B_{k+1} \subset B_k$. Questo ci permette di sfruttare l'ipotesi (2) e affermare che esiste $x \in X$ tale che

$$\{x\} = \bigcap_{k=1}^{\infty} B_k.$$

Poiché $x \in B_k$, per ogni k si ha $d(x, x_{j(k)}) < r_k$, che implica

$$x_{j(k)} \rightarrow x.$$

Bene, abbiamo fatto vedere che la successione originale (x_n) ammette una sottosuccessione $x_{j(k)}$ convergente a x . Grazie all'affermazione del Problema 2.38 possiamo concludere che $x_n \rightarrow x$. Quindi X è completo. \square

Per enunciare il risultato principale di questa sezione abbiamo bisogno della seguente definizione.

2.43 Definizione. Un insieme A in uno spazio metrico (X, d) è detto *mai denso* se la sua chiusura \overline{A} non ha punti interni.

2.44 Proposizione. Se l'insieme A nello spazio metrico (X, d) è mai denso, allora per ogni palla aperta $B \subset X$, e per ogni $\gamma > 0$ esiste un'altra palla aperta $B_1 \subset B$ tale che il raggio di B_1 è minore o uguale a γ e $\overline{B_1} \cap A = \emptyset$.

Dimostrazione. Sia B una palla aperta in X . B non può essere interamente contenuta in \overline{A} , altrimenti il centro di B sarebbe interno a \overline{A} . Sia quindi $x \in B \setminus \overline{A}$. Poiché $B \setminus \overline{A}$ è un aperto, esiste $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset B \setminus \overline{A}$. Sia $\delta = \min\{\varepsilon/2, \gamma\}$. Poniamo $B_1 = B_\delta(x)$. \square

2.45 Teorema. (Teorema della categoria di Baire). Sia (X, d) uno spazio metrico e siano $(A_n)_{n=1}^{\infty}$ sottoinsiemi di X mai densi in X tali che $X = \bigcup_{n=1}^{\infty} A_n$. Allora X non è completo.

Dimostrazione. Sia B_0 una palla di raggio 1. Per la Proposizione 2.44 esiste una palla aperta $B_1 \subset B_0$ di raggio $r_1 \leq 1/2$ tale che $\overline{B_1} \cap A_1 = \emptyset$. Analogamente possiamo trovare una palla aperta $B_2 \subset B_1$ di raggio $r_2 \leq 1/2^2$ tale che $\overline{B_2} \cap A_2 = \emptyset$. Proseguendo in questo modo, per ogni $n \in \mathbb{N}^*$ troviamo una palla aperta B_n di raggio $r_n \leq 1/2^n$ tale che $\overline{B_n} \cap A_n = \emptyset$. Supponiamo ora per assurdo che X sia completo. Per il Teorema 2.42 possiamo affermare che esiste $x \in X$ tale che

$$\bigcap_{n=1}^{\infty} \overline{B_n} = \{x\}.$$

Quindi x è contenuto in ciascuna palla chiusa $\overline{B_n}$. Siccome $\overline{B_n} \cap A_n = \emptyset$, questo implica che $x \notin A_n$. Ma per ipotesi X è l'unione di tutti gli A_n , quindi $x \notin X$. Abbiamo ottenuto una contraddizione. \square

2.4 Compattezza

2.46 Definizione. Uno spazio metrico è detto *compatto* o *sequenzialmente compatto*² se ogni successione ammette una sottosuccessione convergente.

2.47 Proposizione. Uno spazio metrico compatto è completo.

²per gli spazi metrici queste due nozioni coincidono

Dimostrazione. Sia (X, d) uno spazio metrico compatto. Voglio far vedere che (X, d) è completo, vale a dire che ogni successione di Cauchy è convergente. Sia dunque $(x_n)_{n=1}^{\infty}$ una successione di Cauchy di elementi di X . Dato che X è compatto so che esiste una sottosuccessione (x_{n_k}) convergente. Ma, come afferma il Problema 2.38, se una successione di Cauchy ammette una sottosuccessione convergente, allora la successione stessa è necessariamente convergente. Quindi, nel nostro caso, la successione (x_n) è convergente, dunque X è completo. \square

2.48 Definizione. Uno spazio metrico (X, d) è detto *totalmente limitato* se per ogni $\varepsilon > 0$ esiste un sottoinsieme finito di X , $\{x_1, x_2, \dots, x_n\}$ tale che

$$X = \bigcup_{i=1}^n B_{\varepsilon}(x_i).$$

In altre parole uno spazio metrico è totalmente limitato se per ogni $\varepsilon > 0$ posso trovare un insieme finito di punti $\{x_1, x_2, \dots, x_n\}$ con la proprietà che per ogni $x \in X$ esiste i tale che $d(x, x_i) < \varepsilon$.

2.49 Proposizione. *Uno spazio metrico totalmente limitato è limitato.*

Dimostrazione. (Facile)

2.50 Teorema. *Uno spazio metrico è compatto se e solo se è completo e totalmente limitato.*

Dimostrazione. Sia (X, d) uno spazio metrico compatto. Per la Proposizione 2.47 X è completo. Voglio dimostrare che X è totalmente limitato. Suppongo che non sia totalmente limitato e faccio vedere che questo conduce ad una contraddizione. Se X non è totalmente limitato esiste $\varepsilon_0 > 0$ tale che X non può essere espresso come unione di un numero finito di palle di raggio ε_0 . Scelgo allora un punto arbitrario $x_1 \in X$. Poiché $X \neq B_{\varepsilon_0}(x_1)$ posso trovare $x_2 \in X$ tale che $d(x_1, x_2) \geq \varepsilon_0$. Questa procedura può essere iterata nel modo seguente. Siccome

$$X \neq B_{\varepsilon_0}(x_1) \cup B_{\varepsilon_0}(x_2),$$

esiste $x_3 \in X$ tale che $d(x_3, x_1) \geq \varepsilon_0$ e $d(x_3, x_2) \geq \varepsilon_0$. In questo modo costruisco una successione $(x_n)_{n=1}^{\infty}$ di punti di X con la proprietà che se $i \neq j$ si ha $d(x_i, x_j) \geq \varepsilon_0$. Questa successione non può ammettere sottosuccessioni convergenti per cui X non è compatto in contraddizione con l'ipotesi.

Suppongo ora che X sia completo e totalmente limitato e dimostro che X è compatto. Sia quindi $(x_n)_{n=1}^{\infty}$ una successione di elementi di X . Voglio far vedere che essa ammette una sottosuccessione convergente. La strategia è simile a quella che si usa per far vedere che ogni successione limitata in \mathbb{R} ammette una sottosuccessione convergente.

Grazie all'ipotesi di totale limitatezza X può essere scritto come l'unione di un numero finito di palle di raggio 1. Almeno una di queste palle, chiamiamola B_1 , deve contenere un numero infinito di elementi della successione (x_n) . Sia

$$(2.20) \quad x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots$$

la sottosuccessione di tutti gli elementi (x_n) contenuti in B_1 . Sempre grazie alla totale limitatezza possiamo scrivere X come unione di un numero finito di palle di raggio $1/2$. Una di queste deve necessariamente contenere infiniti termini della successione (2.20). Sia

$$(2.21) \quad x_1^{(2)}, x_2^{(2)}, x_3^{(2)}, \dots$$

la sottosuccessione di tutti gli elementi $(x_n^{(1)})$ contenuti in B_2 . Continuando in questo modo, per ogni m intero positivo, avrò costruito una palla B_m di raggio $1/2^m$ e una sottosuccessione

$$(2.22) \quad x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \dots \in B_m$$

definita ricorsivamente come la sottosuccessione di tutti gli elementi $(x_n^{(m-1)})$ che appartengono a B_m . A questo punto vorrei applicare in teorema delle sfere incluse 2.42. Il problema è che le palle B_1, B_2, \dots , (a parte che non sono chiuse, ma questo non è un problema perché posso considerare $\overline{B_m}$) non sono necessariamente contenute l'una nell'altra. Sappiamo solo che B_m contiene termini della successione $(x_n^{(m-1)})$ che è a sua volta tutta contenuta in B_{m-1} . Quindi sappiamo che $B_m \cap B_{m-1} \neq \emptyset$. Poco male, triplichiamo i raggi. Sia y_m il centro della palla B_m e sia $r_m = 1/2^m$ il raggio corrispondente, per cui si ha

$$B_m = B_{r_m}(y_m).$$

Definisco

$$C_m = B_{3r_m}(y_m)$$

e affermo che $C_{m-1} \supset C_m$. Infatti, poiché B_{m-1} e B_m hanno intersezione non vuota, la distanza fra i loro centri soddisfa

$$d(y_{m-1}, y_m) \leq r_{m-1} + r_m = \frac{1}{2^{m-1}} + \frac{1}{2^m} = \frac{3}{2^m}.$$

Supponiamo ora che $x \in C_m$. Allora

$$\begin{aligned} d(x, y_{m-1}) &\leq d(x, y_m) + d(y_m, y_{m-1}) \leq 3r_m + \frac{3}{2^m} \\ &= \frac{3}{2^m} + \frac{3}{2^m} = \frac{3}{2^{m-1}} = 3r_{m-1}, \end{aligned}$$

quindi $x \in C_{m-1}$. Possiamo quindi applicare il teorema 2.42 alla successione $\overline{C_1} \supset \overline{C_2} \supset \dots$ e ottenere che esiste $x \in X$ tale che

$$\bigcap_{m=1}^{\infty} \overline{C_m} = \{x\}.$$

A questo punto non resta che affidarsi al trucco della sottosuccessione “diagonale” costituita dagli elementi

$$x_1^{(1)}, x_2^{(2)}, x_3^{(3)}, \dots, x_m^{(m)}, \dots$$

Siccome x appartiene a ogni C_k e $x_m^{(m)} \in B_m \subset C_m$ si ha

$$d(x, x_m^{(m)}) \leq \text{diam } C_m = 6r_m = \frac{6}{2^m},$$

per cui $\lim_{m \rightarrow \infty} d(x_m^{(m)}, x) = 0$, vale a dire $x_m^{(m)} \rightarrow x$. Abbiamo fatto vedere che una successione arbitraria (x_n) ammette una sottosuccessione convergente, quindi X è compatto. \square

2.51 Corollario. (Heine–Borel). *Un sottoinsieme di \mathbb{R}^n è compatto se e solo se esso è chiuso e limitato.*

Idea della dimostrazione. Sia A un sottoinsieme di \mathbb{R}^n . Grazie alle Proposizioni 2.35 e 2.37 sappiamo che A è completo se e solo se è chiuso. Poi si dimostra che per un sottoinsieme di \mathbb{R}^n limitatezza e totale limitatezza sono la stessa cosa. A questo punto possiamo usare il teorema 2.50. \square

2.52 Osservazione. In generale la totale limitatezza è un concetto più forte della semplice limitatezza. Vedi ad esempio al punto 3.22 dove c'è un esempio di un insieme chiuso e limitato ma non compatto.

2.5 Funzioni continue

In questa sezione (X, d) e (Y, ρ) sono due spazi metrici.

2.53 Definizione. Il *supporto* della funzione $f : X \rightarrow Y$ si denota con il simbolo $\text{supp } f$ ed è definito come

$$\text{supp } f := \overline{\{x \in X : f(x) \neq 0\}}.$$

2.54 Esempi. (1) Sia $X = Y = \mathbb{R}$ e

$$f(x) := \begin{cases} x & \text{se } x \geq 0 \\ 0 & \text{se } x < 0. \end{cases}$$

Allora $\{x \in X : f(x) \neq 0\} = (0, +\infty)$, quindi $\text{supp } f = [0, +\infty)$.

(2) Sia $f(x) = \sin(x)$ con $X = Y = \mathbb{R}$. Allora

$$\{x \in X : f(x) \neq 0\} = \{x \in \mathbb{R} : x \neq k\pi, k \in \mathbb{Z}\}$$

quindi $\text{supp } f = \mathbb{R}$.

2.55 Definizione. Una funzione $f : X \rightarrow Y$ si dice *continua nel punto* $x \in X$ se vale l'implicazione

$$x_n \rightarrow x \Rightarrow f(x_n) \rightarrow f(x).$$

La funzione f è detta *continua* se è continua in ogni punto.

2.56 Proposizione. Sia $f : X \rightarrow Y$. Le seguenti affermazioni sono equivalenti:

- (1) f è continua;
- (2) Per ogni $x \in X$ e per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $d(y, x) < \delta$ allora $\rho(f(y), f(x)) < \varepsilon$;
- (3) Se A è aperto in Y allora $f^{-1}(A)$ è aperto in X ;
- (4) Se A è chiuso in Y allora $f^{-1}(A)$ è chiuso in X .

2.57 Proposizione. Sia $f : X \rightarrow Y$ è un'isometria, allora f è continua.

Dimostrazione. Se $(x_n)_{n=1}^{\infty}$ è una successione in X tale che $x_n \rightarrow x \in X$, allora

$$\rho(f(x_n), f(x)) = d(x_n, x) \rightarrow 0$$

vale a dire $f(x_n) \rightarrow f(x)$. \square

2.58 Problema. Dimostrare che la funzione distanza $d : X \times X \rightarrow \mathbb{R}$ è continua, vale a dire se $x_n \rightarrow x$ e $y_n \rightarrow y$ allora si ha $d(x_n, y_n) \rightarrow d(x, y)$.

2.59 Definizione. La funzione $f : X \rightarrow Y$ è detta *uniformemente continua* se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $d(x, y) < \delta$ allora $\rho(f(x), f(y)) < \varepsilon$. La funzione f è detta *Lipschitz* se esiste $M > 0$ tale che per ogni $x, y \in X$ si ha

$$\rho(f(x), f(y)) \leq M d(x, y).$$

2.60 Proposizione. Siano f, g funzioni continue dallo spazio metrico (X, d) a valori in \mathbb{R} (\mathbb{C}) e siano a, b due numeri reali (complessi). Allora le funzioni $af + bg$ e f/g sono continue. La funzione f/g è continua in $x \in X$ se $g(x) \neq 0$.

2.61 Problema. Dimostrare con un controesempio che il prodotto di due funzioni uniformemente continue non è necessariamente una funzione uniformemente continua.

2.62 Proposizione. Siano (X, d) , (Y, ρ) e (Z, σ) tre spazi metrici. Se $f : X \rightarrow Y$ e $g : Y \rightarrow Z$ sono due funzioni continue allora $g \circ f$ è una funzione continua da X in Z .

2.63 Proposizione. Sia K un sottoinsieme compatto di uno spazio metrico (X, d) e sia $f : K \rightarrow \mathbb{R}$ una funzione continua. Allora f assume in K un massimo e un minimo assoluto, vale a dire esistono $x_0, x_1 \in K$ tali che per ogni $x \in K$ si ha $f(x_0) \leq f(x) \leq f(x_1)$.

Dimostrazione. La dimostrazione è praticamente identica al caso in cui K è un intervallo chiuso e limitato di \mathbb{R} . \square

2.64 Proposizione. Se $f : X \rightarrow Y$ è una funzione continua e X è compatto, allora f è uniformemente continua.

Dimostrazione. La dimostrazione è praticamente identica al caso in cui X è un intervallo chiuso e limitato di \mathbb{R} . \square

2.65 Definizione. Una successione di funzioni $f_n : X \rightarrow Y$ si dice *convergente puntualmente* alla funzione $f : X \rightarrow Y$ se

$$\forall x \in X \text{ si ha } \lim_{n \rightarrow \infty} f_n(x) = f(x).$$

2.66 Definizione. Una successione di funzioni $f_n : X \rightarrow Y$ si dice *uniformemente convergente* alla funzione $f : X \rightarrow Y$ se per ogni $\varepsilon > 0$ esiste un intero positivo N tale che

$$\text{per ogni } n \geq N \text{ e per ogni } x \in X \text{ si ha } \rho(f_n(x), f(x)) < \varepsilon.$$

2.67 Proposizione. Se una successione di funzioni continue $f_n : X \rightarrow Y$ converge uniformemente alla funzione $f : X \rightarrow Y$, allora f è continua.

Dimostrazione. Identica al caso delle funzioni reali.

2.68 Definizione. Sia $f_n : X \rightarrow \mathbb{C}$. La serie di funzioni $\sum_{k=1}^{\infty} f_k(x)$ è detta *puntualmente (uniformemente) convergente* alla funzione $f : X \rightarrow \mathbb{C}$, se la successione associata delle somme parziali

$$F_n(x) = \sum_{k=1}^n f_k(x)$$

converge puntualmente (uniformemente) a $f(x)$. In questo caso si scrive

$$\sum_{k=1}^{\infty} f_k(x) = f(x) \quad (\text{uniformemente}).$$

2.69 Proposizione. (*Test M di Weierstrass*). Sia (X, d) uno spazio metrico e sia $f_n : X \rightarrow \mathbb{C}$ una successione di funzioni tali che

(a) $|f_n(x)| \leq M_n$ per ogni $x \in X$

(b) $\sum_{n=1}^{\infty} M_n < \infty$.

Allora le serie $\sum_{n=0}^{\infty} f_n(x)$ è uniformemente convergente.

Dimostrazione. La dimostrazione è identica a quella valida per funzioni $f_n : \mathbb{R} \rightarrow \mathbb{R}$.

2.6 Completamento di uno spazio metrico

Avere a che fare con uno spazio metrico non completo può essere a volte molto secante, a causa della presenza di successioni che “sembrano convergenti” (perché sono di Cauchy), ma che non hanno limite, o il cui limite esiste “al di fuori” dello spazio metrico considerato. Immaginiamo di conoscere soltanto i numeri razionali, ma non i reali. La successione di numeri razionali

$$x_n := \left(1 + \frac{1}{n}\right)^n$$

sembra convergente in quanto per grandi valori di n , gli x_n si addensano vicino allo stesso punto. Tuttavia la successione *non* è convergente in \mathbb{Q} perché, col senno del poi, sappiamo che x_n tende alla costante e che è irrazionale (a proposito, siete sicuri che e è irrazionale? Come si dimostra?).

Tornando al caso generale di uno spazio metrico (X, d) , come spesso accade, ci sono due notizie: una buona e una cattiva. Quella buona è che esiste una procedura generale per “completare” uno spazio metrico non completo. Così come \mathbb{R} costituisce il completamento di \mathbb{Q} , analogamente, uno spazio metrico (X, d) non completo può essere sempre “immerso” in uno spazio metrico più grande (\tilde{X}, \tilde{d}) completo. Non solo, ma così come avviene per \mathbb{Q} e \mathbb{R} , questa procedura di completamento può essere fatta in modo tale che X è denso in \tilde{X} , quindi ogni elemento di \tilde{X} può essere approssimato da un elemento di X a meno di un ε arbitrariamente piccolo, in simboli:

$$\forall y \in \tilde{X} \quad \forall \varepsilon > 0 \quad \exists x \in X \quad \tilde{d}(x, y) < \varepsilon.$$

Inoltre questo spazio metrico \tilde{X} , che viene definito il *completamento* di X è “essenzialmente” unico,³ per cui possiamo affermare che \tilde{X} è il completamento di X . Sarebbe, a questo punto, che il problema della non completezza di alcuni spazi metrici, venga risolto una volta per tutte da questo teorema, per cui noi non ce ne dovremo più preoccupare. Se non fosse che... non abbiamo ancora parlato della notizia cattiva. La notizia cattiva è la seguente: questa procedura generale, che ci permette costruire il completamento di uno spazio metrico, è di natura astratta e non dà alcuna informazione pratica per “fare i calcoli” nello spazio \tilde{X} . In un capitolo successivo accenneremo ad un caso importante: quello degli spazi L_p , che costituiscono il completamento di spazi di funzioni continue. La comprensione della struttura di questi spazi ha richiesto la costruzione di una nuova teoria dell’integrazione (di Lebesgue) che generalizza il concetto di integrale di Riemann.

³Volendo essere precisi: unico a meno di isometrie suriettive che lasciano fissi i punti di X .

2.6.1 I numeri p -adici

Abbiamo visto come l'insieme dei numeri reali costituisce il completamento dell'insieme dei razionali. Vogliamo ora descrivere un completamento diverso di \mathbb{Q} che non produce \mathbb{R} , ma, fissato un numero primo arbitrario p , dà origine al campo \mathbb{Q}_p dei cosiddetti numeri p -adici. Ma come? Non abbiamo appena detto che il completamento di uno spazio metrico è unico? Giusto, infatti per costruire i numeri p -adici, partiamo dall'insieme \mathbb{Q} , ma su di esso definiamo una distanza diversa da quella usuale, che denotiamo con d_p . Quindi \mathbb{R} è il completamento dello spazio metrico (\mathbb{Q}, d) in cui d è la distanza usuale, mentre \mathbb{Q}_p è il completamento di (\mathbb{Q}, d_p) .

La rappresentazione dei reali in base arbitraria

Prima di parlare dei numeri p -adici, ricordiamo alcuni fatti elementari che riguardano la rappresentazione in base s (s -naria) dei numeri razionali e, più in generale, reali. Dato un intero $s \geq 2$, ogni numero reale x può essere scritto nella forma

$$(2.23) \quad x = \pm \sum_{k=-\infty}^n x_k s^k \quad \text{in cui } x_k \in \{0, 1, \dots, s-1\}.$$

Quindi x viene rappresentato dalla stringa di simboli

$$(2.24) \quad x = \pm [x_n x_{n-1} \dots x_1 x_0 . x_{-1} x_{-2} x_{-3} \dots]_s \quad x_k \in \{0, 1, \dots, s-1\}$$

in cui l'ultimo pedice s denota la base che stiamo usando per la rappresentazione. Nella (2.24) il numero di cifre a sinistra del *punto* è sempre finito, mentre a destra del punto possiamo avere un numero finito o infinito di cifre. Diciamo che la cifra x_k occupa la *posizione* k , quindi, ad esempio, la posizione 0 è quella occupata dalla prima cifra a sinistra del punto e la posizione -1 è quella occupata dalla prima cifra a destra del punto. Assumiamo inoltre che non ci siano “zeri inutili” che non modificano il valore di x . Quindi

$$00123.4567000_{10} \quad \text{viene scritto} \quad 123.4567.$$

Se x è razionale la sua espressione in base s risulta particolarmente semplice. Ogni $x \in \mathbb{Q}$ può essere scritto come

$$x = \frac{a}{b} \quad a \in \mathbb{Z}, b \in \mathbb{N}^*, a \text{ e } b \text{ senza fattori comuni.}$$

Definiamo $\text{num}(x) = a$ e $\text{den}(x) = b$. Ad esempio:⁴

$$x := \frac{33}{-9} = \frac{-11}{3} \implies \text{num}(x) = -11 \quad \text{den}(x) = 3.$$

Fissata la base s , i razionali si possono dividere in due gruppi $\mathbb{Q} = \mathbb{Q}_{\text{fn}}^{(s)} \cup \mathbb{Q}_{\text{per}}^{(s)}$ (attenzione la suddivisione *dipende da* s) nel modo seguente:

- (a) se ogni fattore primo di $\text{den}(x)$ è anche un fattore primo di s allora x ha una rappresentazione *finita*, cioè con un numero finito di cifre. Diciamo che $x \in \mathbb{Q}_{\text{fn}}^{(s)}$.
Ad esempio:

$$\begin{array}{ll} \frac{1}{5} = 0.2_{10} = 0.1_5 & \frac{17}{16} = 1.0625_{10} = 1.04_8 \\ \frac{1}{3} = 0.2_6 = 0.3_9 & \frac{41}{36} = 1.05_6 = 1.18_{12}. \end{array}$$

⁴Se in un'espressione compaiono cifre senza un pedice che indica esplicitamente la base, si intende base 10.

(b) se, al contrario, $\text{den}(x)$ contiene fattori primi che non dividono s , allora x ha una rappresentazione *periodica*, vale a dire, una rappresentazione che, dopo un numero finito di cifre, diventa periodica⁵ (come si dimostra?). Diciamo che $x \in \mathbb{Q}_{\text{per}}^{(s)}$. La parte periodica della rappresentazione viene comunemente indicata con una barra posta al di sopra:

$$1.234\overline{5678} = 1.23456785678567856785678 \dots$$

Esempi:

$$\frac{5}{3} = 1.\overline{6}_{10} = 1.\overline{31}_5 \qquad \frac{1}{14} = 0.\overline{0714285}_{10} = 0.0\overline{3}_7.$$

I razionali in $\mathbb{Q}_{\text{fin}}^{(s)}$ ammettono *anche* una seconda rappresentazione periodica. Ad esempio:

$$(2.25) \qquad 1_{10} = 0.\overline{9}_{10} \qquad 1_5 = 0.\overline{4}_5 \qquad 13.451_7 = 13.450\overline{6}_7.$$

Il motivo è il seguente: prendiamo in esame la prima identità $1_{10} = 0.\overline{9}_{10}$. Ponendo

$$y_j := 0.\underbrace{9999 \dots 9}_{j \text{ cifre}}_{10}$$

si ha che la distanza di y_j da 1 è data da

$$d(y_j, 1) = |1 - y_j| = 0.\underbrace{0000 \dots 0}_{j-1 \text{ zeri}}1_{10} = 10^{-j}$$

quindi, nel limite $j \rightarrow \infty$, questa distanza tende a zero. Di conseguenza

$$(2.26) \qquad 0.\overline{9}_{10} = \lim_{j \rightarrow \infty} y_j = 1.$$

Ogni numero reale x è il limite di una successione di elementi di $\mathbb{Q}_{\text{fin}}^{(s)}$, ottenuta troncando la rappresentazione di x nella base scelta. In base 10, si ha

$$\sqrt{2} = 1.414213562373095 \dots_{10}.$$

Se definisco la successione

$$(2.27) \qquad y_1 = 1_{10} \qquad y_2 = 1.4_{10} \qquad y_3 = 1.41_{10} \qquad y_4 = 1.414_{10} \qquad \dots$$

ottengo una successione di elementi di $\mathbb{Q}_{\text{fin}}^{(s)}$ che si verifica facilmente essere *di Cauchy* perché aggiungere cifre a destra dalla posizione $-n$ ha un effetto sul numero rappresentato che non supera s^{-n} . Questa procedura di *aggiungere infinite cifre verso destra* produce sempre una successione di Cauchy e, poiché \mathbb{R} è il completamento di \mathbb{Q} , identifica di conseguenza un numero reale, come nel caso di $\sqrt{2}$.

Le regole del calcolo aritmetico in base s sono identiche a quelle in base 10, ma il riporto va fatto quando si arriva a s . Ad esempio:

$$2_7 + 4_7 = 6_7 \qquad 2_7 + 5_7 = 0_7 \text{ col riporto di } 1 \qquad 2_7 \cdot 5_7 = 3_7 \text{ col riporto di } 1.$$

⁵Bisognerebbe più propriamente definirla *rappresentazione eventualmente periodica* o qualcosa del genere.

I numeri p -adici

Passiamo ora ai numeri p -adici in cui con p indichiamo un numero *primo* fissato. Detto in due (poche) parole e molto malamente: i numeri p -adici si ottengono dai razionali aggiungendo infinite cifre, invece che a destra, *a sinistra del punto* (oooohh!!!). Chi volesse provare a capire perché può continuare a leggere.

Come abbiamo già detto, i p -adici si ottengono sempre come completamento dei razionali, a patto di usare una distanza appropriata, diversa dalla distanza usuale. Così come la distanza usuale fra x e y si ottiene come valore assoluto della differenza fra x e y , analogamente la distanza p -adica è data da

$$(2.28) \quad d_p(x, y) := |x - y|_p,$$

in cui $|\cdot|_p$ è il valore assoluto p -adico, definito nel modo seguente: se $x = 0$, $|x|_p = 0$, mentre se $0 \neq x \in \mathbb{Q}$, posso sempre scrivere x nella forma

$$(2.29) \quad x = p^n \frac{a}{b},$$

in cui a e b sono interi non divisibili per p e n è un intero arbitrario. Allora definisco

$$(2.30) \quad |x|_p := p^{-n}.$$

Ad esempio, se $p = 3$, si ha

$$\frac{231}{270} = \frac{3 \cdot 7 \cdot 11}{3^3 \cdot 2 \cdot 5} = 3^{-2} \frac{7 \cdot 11}{2 \cdot 5} \implies \left| \frac{231}{270} \right|_3 = 3^2 = 9.$$

Utilizzando la fattorizzazione mostrata sopra si vede subito che

$$\left| \frac{231}{270} \right|_2 = 2 \quad \left| \frac{231}{270} \right|_5 = 7 \quad \left| \frac{231}{270} \right|_7 = 1/7 \quad \left| \frac{231}{270} \right|_{11} = 1/11,$$

mentre, per tutti gli altri primi p si ha $|231/270|_p = 1$. Si verifica facilmente che la funzione $|\cdot|_p : \mathbb{Q} \rightarrow \mathbb{R}$ è un *valore assoluto* sul campo \mathbb{Q} , valse a dire soddisfa le proprietà⁶

- (a) $|x|_p \geq 0$;
- (b) $|x|_p = 0$ se e solo se $x = 0$;
- (c) $|xy|_p = |x|_p |y|_p$;
- (d) $|x + y|_p \leq |x|_p + |y|_p$.

Come conseguenza si dimostra che la funzione d_p definita nella (2.28) è effettivamente una distanza su \mathbb{Q} . A questo punto il teorema generale sull'esistenza e l'unicità del completamento, a cui abbiamo accennato in precedenza, ci permette di dare la seguente bellissima

2.70 Definizione. Per ogni numero primo p , si definisce l'insieme dei *numeri p -adici* come

$$\mathbb{Q}_p = \text{completamento di } \mathbb{Q} \text{ rispetto alla distanza } d_p.$$

⁶Più precisamente $|\cdot|_p$ soddisfa una disuguaglianza più forte della disuguaglianza triangolare (d), data da (d') $|x + y|_p \leq \max\{|x|_p, |y|_p\}$. Inoltre, se $|x|_p \neq |y|_p$, allora $|x + y|_p = \max\{|x|_p, |y|_p\}$. I valori assoluti che soddisfano (d') sono detti non-Archimedei, rispetto ai quali "tutti i triangoli sono isosceli".

Bene, abbiamo finito, possiamo andare a surfare! Purtroppo no. Perché, insomma, ma come sono fatti questi numeri p -adici? Vi avevo avvertito che, su questo punto, il teorema generale non è di grande aiuto. Capire la struttura del completamento richiede un lavoro addizionale, che varia da caso a caso. A priori sappiamo che \mathbb{Q}_p in quanto completamento di \mathbb{Q} , contiene \mathbb{Q} più tutti i limiti delle successioni di Cauchy dei razionali *rispetto alla distanza* d_p . Per avere un'idea intuitiva su cos'è una successione di Cauchy rispetto a d_p bisogna iniziare a capire quand'è che 2 numeri sono vicini e, di conseguenza, bisogna acquisire dimestichezza con il valore assoluto p -adico. Iniziamo ad occuparci dei numeri razionali *positivi* che in base p , hanno una rappresentazione finita. Notate che, essendo p primo, un numero razionale x ha rappresentazione finita in base p se e solo se $\text{den}(x)$ è una potenza di p (inclusa $p^0 = 1$). Se $0 < x \in \mathbb{Q}_{\text{fin}}$, allora

$$x = \sum_{k=m}^n x_k p^k = \begin{cases} [x_n x_{n-1} \dots x_m 0 \dots 0]_p & \text{se } 0 \leq m \leq n \\ [x_n x_{n-1} \dots x_1 x_0 . x_{-1} x_{-2} \dots x_m]_p & \text{se } m < 0 \leq n \\ [0 . 0 0 \dots x_n x_{n-1} \dots x_m]_p & \text{se } m \leq n < 0. \end{cases}$$

Quindi

$$x = p^m (x_m + x_{m+1}p + \dots + x_n p^{n-m}) = p^m (x_m + pa)$$

in cui

$$a := x_{m+1} + \dots + x_n p^{n-m-1}.$$

Siccome x_m è un intero compreso fra 0 e $p-1$, il numero intero $x_m + pa$ non è divisibile per p (x_m è il resto di tale divisione). Quindi, dalla definizione segue una regola semplicissima per calcolare il valore assoluto p -adico di un numero razionale positivo in \mathbb{Q}_{fin} , *a patto di scrivere questo numero in base* p :

$$(2.31) \quad |x|_p = p^{-m} = p^{-(\text{posizione dell'ultima cifra diversa da zero})}.$$

Ad esempio:

$$|1|_7 = 7^0 = 1 \quad |1110|_7 = 7^{-1} \quad |12300|_7 = 7^{-2} \quad |12300.1|_7 = 7 \quad |123.12|_7 = 7^2.$$

Se conosciamo x in un'altra base, lo convertiamo in base p e se la sua rappresentazione in base p è finita, possiamo calcolare $|x|_p$:

$$|63.9936|_{10} = |223.4441|_5 = 5^4.$$

Adesso possiamo calcolare anche le distanze:

$$\begin{aligned} d_7(100_7, 300_7) &= |200|_7 = 7^{-2} \\ d_7(123_7, 4123_7) &= |4000|_7 = 7^{-3} \\ d_7(100.1_7, 100.2_7) &= |0.1|_7 = 7 \\ d_7(100.1234_7, 100.12345_7) &= |0.00005|_7 = 7^5. \end{aligned}$$

È chiaro quindi che: *se* y *si ottiene da* x *aggiungendo una cifra alla sinistra di* x *nella rappresentazione in base* p , *la distanza di* y *da* x *è uguale a* p^{-n} *in cui* n *è la posizione della cifra aggiunta.* Ad esempio, se parto dal numero 123_7 e aggiungo via via altre cifre arbitrarie a sinistra ottengo

$$\begin{aligned} d_7(123_7, 5123_7) &= |5000|_7 = 7^{-3} && \text{ho aggiunto un } 5 \text{ nella pos. } n = 3 \\ d_7(5123_7, 45123_7) &= |40000|_7 = 7^{-4} && \text{ho aggiunto un } 4 \text{ nella pos. } n = 4 \\ d_7(45123_7, 245123_7) &= |200000|_7 = 7^{-5} && \text{ho aggiunto un } 2 \text{ nella pos. } n = 5. \end{aligned}$$

Di conseguenza si dimostra facilmente che, se continuo ad aggiungere cifre a sinistra, ottengo una successione di Cauchy rispetto alla distanza d_p . Poiché \mathbb{Q}_p è il completamento di \mathbb{Q} rispetto a questa distanza, questa successione di Cauchy identifica un elemento di \mathbb{Q}_p . Il limite di questa successione sarà dunque rappresentato da un numero che possiede, in generale, infinite cifre *a sinistra del punto!*

$$(2.32) \quad [\dots x_1 x_0 \cdot x_{-1} x_{-2} \dots x_m]_p \quad x_k \in \{0, 1, \dots, p-1\}.$$

Viceversa se provo ad aggiungere cifre a destra, come facevo nella (2.27) per approssimare $\sqrt{2}$, ottengo una successione in cui la distanza di un elemento dal precedente diverge esponenzialmente. La situazione è dunque (vagamente) speculare a quanto accade nella costruzione dei reali.

Domanda: ok, è chiaro che ogni stringa di simboli nel formato che appare nella (2.32) identifica un numero p -adico, ma è vero anche il contrario? In altre parole, posso ottenere in questo modo *tutti* gli elementi di \mathbb{Q}_p ? La risposta è affermativa.

2.71 Teorema. *Ogni numero p -adico $x \in \mathbb{Q}_p$ ha un espansione unica (a parte zeri "irrilevanti") della forma*

$$(2.33) \quad x = \sum_{k=m}^{\infty} x_k p^k \quad \text{in cui } m \in \mathbb{Z} \text{ e } x_k \in \{0, 1, \dots, p-1\}$$

e può essere quindi rappresentato come nella (2.32).

Bene, quindi ogni numero p -adico può essere rappresentato nella forma (2.32). Iniziamo da un caso semplice: se $x \in \mathbb{Q}_{\text{fin}}^{(p)}$ ed è positivo allora la rappresentazione p -adica coincide con la rappresentazione in base p . Ad esempio

$$(2.34) \quad \frac{2236}{125} = 32.421_5 \quad \text{rappresentazione sia in base 5 che 5-adica.}$$

2.72 Attenzione! Alcuni autori usano una notazione diversa per rappresentare i numeri p -adici, nella quale la stringa (2.32) viene scritta al contrario. Ad esempio il numero 22, che in base 7 si scrive 31_7 , seguendo questa convenzione si scriverebbe 0.13_7 e il numero 32.421_5 che compare nella (2.34) si scriverebbe 124.23_5 .⁷

Per i razionali negativi sembrerebbe esserci un problema! Infatti, a differenza della (2.24), nella (2.32) non c'è la possibilità di inserire un eventuale segno \pm , consistentemente con quanto affermato nel Teorema 2.71. Niente segno neanche lì. Ma come faccio allora, a rappresentare i razionali negativi o anche soltanto gli interi negativi? Tranquilli, tutto sotto controllo. La rappresentazione p -adica di -1 , ad esempio, è

$$(2.35) \quad -1 = \dots q q q q q_p = \bar{q}_p \quad \text{in cui } q = p-1.$$

$$(2.36) \quad -1 = \dots 6 6 6 6 6_7 = \bar{6}_7 \quad (p=7).$$

Osservate che, in questa notazione, una barra posta al di sopra di un gruppo di cifre indica che quel gruppo di cifre deve essere ripetuto periodicamente infinite volte verso sinistra. Verificare le (2.35), (2.36) è immediato. Analogamente a come abbiamo

⁷Se scoprite chi ha inventato questa notazione e, magari, dove parcheggia la macchina, casomai fatemelo sapere.

dimostrato che $0.\overline{9}_{10} = 1$ nella (2.26), si ottiene

$$\begin{aligned} d_7(-1, \underbrace{666 \dots 666}_n) &= \underbrace{|666 \dots 666 + 1|}_n \\ &= \underbrace{|1000 \dots 000|}_{n+1} \quad [\text{uso: } 6_7 + 1 = 0_7 \text{ col riporto di } 1] \\ &= 7^{-n}. \end{aligned}$$

Di conseguenza

$$d_7(-1, \overline{6}_7) = \lim_{n \rightarrow \infty} d_7(-1, \underbrace{666 \dots 666}_n) = \lim_{n \rightarrow \infty} 7^{-n} = 0.$$

Da notare che facendo l'usuale addizione col riporto si verifica che

$$\begin{array}{r} \dots 666666_7 + \\ 1_7 = \\ \hline \dots 00000_7 \end{array}$$

Quindi $\overline{6}_7 + 1_7 = 0$. La regola generale per cambiare segno ad un numero p -adico è la seguente:

- (1) se il numero ha una rappresentazione finita, aggiungere sempre infiniti zeri (irrilevanti) a sinistra della prima cifra;
- (2) sia x_m l'ultima cifra diversa da zero (la più a destra diversa da zero). Essa viene sostituita con la regola $x_m \rightarrow p - x_m$;
- (3) alle cifre a sinistra di x_m viene applicata la trasformazione $x_k \rightarrow p - x_k - 1$;
- (4) gli eventuali zeri a destra di x_m rimangono invariati.

Ad esempio:

$$\begin{aligned} 105_{10} &= 210_7 = \dots 000210_7 \\ -105_{10} &= \dots 666460_7 = \overline{6}460_7. \end{aligned}$$

C'è un'altra piccola difficoltà da risolvere: non potendo usare un numero infinito di cifre a destra del punto, come faccio a rappresentare, nella notazione p -adica, un numero $x \in \mathbb{Q}_{\text{per}}^{(p)}$? Ad esempio, come si rappresenta in notazione 7-adica il numero

$$\frac{5}{24} = 0.131313 \dots_7 = 0.\overline{13}_7?$$

Si dimostra che ogni razionale in $\mathbb{Q}_{\text{per}}^{(p)}$ ha una rappresentazione p -adica anch'essa periodica, ma invece che essere periodica "verso destra" è periodica "verso sinistra". Nel caso di $5/24$ si ha

$$(2.37) \quad \frac{5}{24} = \dots 3535354_7 = \overline{35}4_7.$$

Verifichiamolo. Per definizione

$$\begin{aligned} (2.38) \quad \overline{35}4_7 &= 4 + 5 \cdot 7 + 3 \cdot 7^2 + 5 \cdot 7^3 + 3 \cdot 7^4 + \dots \\ &= 4 + 5 \sum_{k=0}^{\infty} 7^{2k+1} + 3 \sum_{k=1}^{\infty} 7^{2k}. \end{aligned}$$

Attenzione, le due serie sono convergenti perché nel mondo p -adico si ha

$$(2.39) \quad \sum_{k=0}^{\infty} p^{\ell k} = \frac{1}{1-p^\ell} \quad \ell = 1, 2, \dots$$

Per verificare la (2.39), uso la definizione di somma di una serie come limite delle somme parziali. Poiché

$$S_n := \sum_{k=0}^n p^{\ell k} = \sum_{k=0}^n (p^\ell)^k = \frac{1-p^{\ell(n+1)}}{1-p^\ell}$$

trovo

$$S_n - \frac{1}{1-p^\ell} = -\frac{p^{\ell(n+1)}}{1-p^\ell}$$

e, siccome $1-p^\ell$ non è divisibile per p , la regola (2.30) per calcolare il valore assoluto p -adico implica

$$\left| S_n - \frac{1}{1-p^\ell} \right|_p = p^{-\ell(n+1)} \xrightarrow{n \rightarrow \infty} 0,$$

che è equivalente alla (2.39). Tornando alla (2.38), otteniamo

$$\overline{354}_7 = 4 + 5 \cdot 7 \sum_{k=0}^{\infty} 7^{2k} + 3 \left(\sum_{k=0}^{\infty} 7^{2k} - 1 \right) = 1 + 38 \frac{1}{1-7^2} = \frac{5}{24}.$$

Se qualcuno fosse curioso di capire come si trova la (2.37) può guardare, ad esempio, l'articolo *Koc, A tutorial on p -adic Arithmetic*, facilmente reperibile.

2.73. *Gli interi p -adici:* \mathbb{Z}_p . Si definiscono *interi p -adici* quei numeri p -adici $x \in \mathbb{Q}_p$ che non hanno cifre diverse da zero a destra del punto. L'insieme degli interi p -adici si denota con \mathbb{Z}_p .⁸ Quindi

$$(2.40) \quad \mathbb{Z}_p := \left\{ x = \sum_{k=0}^{\infty} x_k p^k : x_k \in \{0, 1, \dots, p-1\} \right\}.$$

Così come ogni razionale è il rapporto fra 2 interi, analogamente, ogni numero p -adico è il rapporto fra due interi p -adici. Il motivo è banale e si capisce con un esempio:

$$\dots 123.45_7 = \frac{\dots 12345_7}{100_7}.$$

Osservate che, grazie alla (2.31) è possibile dare una definizione *metrica* di \mathbb{Z}_p

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}.$$

A differenza di \mathbb{Z} , \mathbb{Z}_p non è un insieme numerabile, ma ha la potenza del continuo, cioè esiste una corrispondenza biunivoca fra \mathbb{Q}_p e \mathbb{R} . Un'altra differenza è la seguente: gli unici interi il cui inverso è ancora un numero intero sono -1 e $+1$, mentre ogni intero p -adico la cui ultima cifra è diversa da zero ha un inverso che è ancora un intero

⁸C'è un altro caso di notazione non universale. Qualcuno usa il simbolo \mathbb{Z}_p per indicare l'insieme degli *interi modulo p* , che invece noi indichiamo con $\mathbb{Z}/(p\mathbb{Z})$.

p -adico. In particolare, se k è un intero, allora $1/k$ è un intero p -adico se e solo se k non è multiplo di p . Ad esempio:

$$(2.41) \quad \frac{1}{3_{10}} = \frac{1}{3_7} = \bar{4}5_7 \qquad \frac{1}{14_{10}} = \frac{1}{20_7} = \bar{3}.4_7$$

Queste identità si verificano banalmente, facendo vedere con la regola elementare della moltiplicazione col riporto, che si ottiene

$$\bar{4}5_7 \cdot 3_7 = 1_7 \qquad \bar{3}.4_7 \cdot 20_7 = 1_7.$$

2.74. Radici. Esiste la radice quadrata di un numero p -adico? La risposta è: dipende! Supponiamo di voler risolvere l'equazione

$$(2.42) \quad x^2 = 2 \qquad x \in \mathbb{Q}_7.$$

Poiché $|2|_7 = 1$, abbiamo $|x^2|_7 = |x|_7^2 = 1$, Quindi $|x|_7 = 1$, vale a dire x è un intero p -adico in cui l'ultima cifra è diversa da zero

$$x = \dots x_2 x_1 x_0 \qquad x_0 \neq 0.$$

Facendo la moltiplicazione col riporto di x per sé stesso si ottiene che l'ultima cifra di x^2 è $x_0^2 \pmod{7}$, dunque una condizione *necessaria* affinché si abbia $x^2 = 2$ è

$$(2.43) \quad x_0^2 \equiv 2 \pmod{7}.$$

Esiste un tale x_0 ? In generale si può usare il cosiddetto criterio di Eulero per stabilirlo (cercate "Euler's criterion" su Wikipedia), ma nel nostro caso possiamo facilmente trovare per tentativi tutte le soluzioni della (2.43), facendo una tabella.

n	1	2	3	4	5	6
$n^2 \pmod{7}$	1	4	2	2	4	1

Quindi l'ultima cifra di x deve essere $x_0 = 3$ o $x_0 = 4$. Ma l'esistenza di una soluzione della (2.43) è anche una condizione sufficiente per l'esistenza di una soluzione della (2.42)? Se $p \neq 2$, un teorema, noto come Lemma di Hensel, permette di rispondere affermativamente. Più precisamente la situazione è la seguente: *se $p \neq 2$ e y è un intero p -adico, allora \sqrt{y} esiste in \mathbb{Q}_p se e solo se la rappresentazione p -adica di y termina con un numero pari di zeri e l'ultima cifra diversa da zero, y_0 , è tale che l'equazione $x_0^2 \equiv y_0 \pmod{p}$ ha soluzioni.* Tornando al caso $p = 7$ dalla tabella precedente si vede che \sqrt{y} esiste in \mathbb{Q}_7 se y termina in 1, 2 o 4 seguiti, eventualmente, da un numero pari di zeri. Gli interi fra 0 e 20 che ammettono radice quadrata in \mathbb{Q}_7 sono quindi:

$$\begin{array}{lllll} 1_{10} = 1_7 & 2_{10} = 2_7 & 4_{10} = 4_7 & 8_{10} = 11_7 & 9_{10} = 12_7 \\ 11_{10} = 14_7 & 15_{10} = 21_7 & 16_{10} = 22_7 & 18_{10} = 24_7 & \end{array}$$

La stessa regola si applica anche agli interi negativi. Poiché, ricordando la (2.36),

$$-1_{10} = \bar{6}_7 \qquad -2_{10} = \bar{6}5_7 \qquad -3_{10} = \bar{6}4_7.$$

otteniamo che \mathbb{Q}_7 contiene $\sqrt{-3}$, ma non $\sqrt{-1}$ e $\sqrt{-2}$.

Osservate che se y è un intero p -adico, allora $|y|_p \leq 1$, quindi anche \sqrt{y} , se esiste, è un intero p -adico perché $|\sqrt{y}|_p = \sqrt{|y|_p} \leq 1$.

Il caso $p = 2$ è diverso. Gli interi che ammettono radice quadrata in \mathbb{Q}_2 sono quelli della forma

$$k = 2^{2m}(8n + 1) \qquad m \in \mathbb{N}, n \in \mathbb{Z}.$$

3. Spazi vettoriali e di Banach

*D: They don't advertise for killers in the newspaper.
That was my profession. Ex-cop. Ex-blade runner.
Ex-killer.*

3.1 Definizioni

3.1 Definizione. Uno spazio vettoriale (o spazio lineare) sul campo \mathbb{F} ($\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p^n, \dots$) è un insieme (non vuoto¹) V tale che

(A) Per tutti gli $u, v \in V$ esiste il vettore $u + v \in V$ detto somma di u e v

(B) Per ogni $u \in V, c \in \mathbb{F}$ esiste il vettore $cu \in V$ detto prodotto di c per u .

Inoltre le seguenti proprietà sono soddisfatte, per ogni $u, v, w \in V$ e per ogni $c, d \in \mathbb{F}$:

(SV1) $u + v = v + u$

(SV2) $(u + v) + w = u + (v + w)$

(SV3) Esiste $0 \in V$ tale che $u + 0 = u$ per ogni $u \in V$

(SV4) Per ogni $u \in V$ esiste $v \in V$ tale che $u + v = 0$

(SV5) $1u = u$

(SV6) $(cd)u = c(du)$

(SV7) $(c + d)u = cu + du$

(SV8) $c(u + v) = cu + cv$

3.2 Esempi

(1) $\mathbb{R}^n, \mathbb{C}^n$.

(2) L'insieme di tutti gli $x \in \mathbb{R}^4$ tali che $2x_1 + 3x_2 + x_4 = 0$.

(3) (non-esempio). L'insieme di tutti gli $x \in \mathbb{R}^4$ tali che $x_2 - x_4 = 1$. Non è uno spazio vettoriale perché, ad esempio, il "vettore zero", cioè $0 = (0, 0, 0, 0)$ non appartiene all'insieme.

(4) (non-esempio). L'insieme di tutte le matrici. Non è uno spazio vettoriale perché, per dirne una, non è possibile sommare una matrice 3×2 ad una matrice 6×19 .

(5) $\text{Mat}_{n \times m}(\mathbb{F})$ = l'insieme di tutte le matrici $n \times m$ a coefficienti nel campo \mathbb{F} .

(6) L'insieme di tutte le matrici $A \in \text{Mat}_{n \times n}(\mathbb{C})$ tali che $\text{tr } A = 0$.

¹qualcuno tende ad innervosirsi quando legge la precisazione "non vuoto". Io lo metto fra parentesi sperando di alleviare la sofferenza

- (7) (non-esempio). L'insieme di tutte le matrici $A \in \text{Mat}_{n \times n}(\mathbb{C})$ tali che $\det A = 0$. Non va bene perché, ad esempio, le matrici

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

hanno entrambe determinante uguale 0, ma la loro somma $A + B$ ha determinante uguale ad 1, quindi $A + B$ non appartiene all'insieme.

- (8) $\mathbb{R}[x]$ e $\mathbb{C}[x]$, l'insieme di tutti i polinomi a coefficienti rispettivamente reali o complessi sono spazi vettoriali rispetto alle usuali operazioni di somma e prodotto per un numero. A volte useremo i simboli $P(\mathbb{R})$ e $P(\mathbb{C})$.
- (9) \mathbb{R}^∞ . Così come \mathbb{R}^n indica l'insieme di tutte le n -ple, \mathbb{R}^∞ indica l'insieme di tutti i vettori con infinite componenti, vale a dire l'insieme di tutte le successioni. Gli elementi di \mathbb{R}^∞ sono della forma

$$x = (x_0, x_1, x_2, x_3, \dots)$$

\mathbb{R}^∞ è anche indicato con $\mathbb{R}^{\mathbb{N}}$.

- (10) L'insieme di tutti i polinomi di grado non superiore a 100 è ancora uno spazio vettoriale.
- (11) (non-esempio). L'insieme di tutti i polinomi di grado superiore a 100 non è uno spazio vettoriale perché il polinomio nullo, di grado zero, non appartiene all'insieme.
- (12) (non-esempio). Sia

$$V := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x \text{ ha almeno una componente nulla}\}$$

V non è uno spazio vettoriale, perché la proprietà (A) di chiusura rispetto alla somma è violata. Infatti, sia

$$u = (0, 1, 1) \quad v = (1, 0, 0).$$

Sia u che v sono elementi di V , ma $u + v = (1, 1, 1)$ non appartiene a V .

- (13) (non-esempio). Sia

$$V := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1 \leq x_2\}.$$

Il fatto che V sia definito tramite una *disuguaglianza* suggerisce che non sia uno spazio vettoriale. Infatti $u = (0, 1) \in V$. Però se moltiplichiamo u per -1 otteniamo $-1 \cdot u = (0, -1)$ che non appartiene a V . Geometricamente V è un semipiano, mentre uno spazio vettoriale è sempre un (iper)piano.

- (14) (non-esempio). Sia

$$V := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0\}.$$

Il fatto che V sia definito tramite un'equazione non lineare (quadratica in questo caso) suggerisce che non sia uno spazio vettoriale, e infatti geometricamente V non è un iperpiano, bensì... Per dimostrare che V non è uno spazio vettoriale faccio vedere che viola la proprietà (A) e cerco due vettori u, v tali che $u, v \in V$ ma $u + v \notin V$. Basta prendere ad esempio

$$u = (1, 0, 1) \quad v = (0, 1, 1) \quad u + v = (1, 1, 2).$$

3.2 Attenzione! Non è detto che se c'è di mezzo una disuguaglianza o una relazione non lineare allora sicuramente non si tratta di uno spazio vettoriale. Ci possono essere dei camuffamenti più o meno espliciti. Siano ad esempio

$$V := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1 \geq x_2 \text{ e } x_1 \leq x_2\} \quad (\text{apparenti disuguaglianze})$$

$$W := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^3 = x_2^3\} \quad (\text{relazione apparentemente cubica}).$$

Questi sono entrambi spazi vettoriali nonostante le apparenze, perché

$$\begin{aligned} x_1 \geq x_2 \text{ e } x_1 \leq x_2 & \Leftrightarrow x_1 = x_2 \\ x_1^3 = x_2^3 & \Leftrightarrow x_1 = x_2 \end{aligned}$$

3.3 Dritta. (Come dimostrare che X è uno spazio vettoriale in 2 mosse). Generalmente per dimostrare che X è uno spazio vettoriale bisogna verificare le 10 proprietà che appaiono nella definizione, vale a dire (A), (B), (1), (2), ..., (8). Però se uno va di fretta e in più il caso vuole che X sia un sottoinsieme di uno spazio più grande Y e già si sa che Y è uno spazio vettoriale, allora per dimostrare che anche X è uno spazio vettoriale è sufficiente dimostrare la sua chiusura rispetto alle operazioni di somma e moltiplicazione per uno scalare, vale a dire le 2 proprietà (A) e (B). perché è vero questo? (Dimostrarlo). Attenzione però, se pensate che X non sia uno spazio vettoriale e volete dimostrarlo, allora potrebbe tornare comodo far vedere che una delle proprietà (1), (2), ..., (8) non è verificata.

3.4 Problema. Dire se l'insieme²

$$X = \{p \in \mathbb{R}[x] : p(2) + p'(3) = 0\}.$$

è uno spazio vettoriale dimostrandolo.

Soluzione. L'insieme X è fortunatamente un sottoinsieme dell'insieme di tutti i polinomi reali che abbiamo già detto essere uno spazio vettoriale. Quindi, per quanto detto al punto 3.3, se dimostro che le proprietà (A) e (B) sono verificate, posso concludere che X è uno spazio vettoriale.

Dimostrazione di (A). Siano p e q due elementi di X . Devo far vedere che anche la loro somma $p + q$ appartiene a X . Dato che $p, q \in X$ so che

$$p(2) + p'(3) = 0 \quad q(2) + q'(3) = 0.$$

D'altra parte, siccome l'addizione di due polinomi è definita come $(p + q)(x) := p(x) + q(x)$, ottengo

$$(p + q)(2) + (p + q)'(3) = p(2) + q(2) + p'(3) + q'(3) = 0.$$

Quindi $p + q \in X$.

Dimostrazione di (B). Sia $p \in X$ e sia $c \in \mathbb{R}$. Allora

$$(cp)(2) + (cp)'(3) = cp(2) + cp'(3) = c(p(2) + p'(3)) = 0.$$

Dunque $cp \in X$. Segue che X è uno spazio vettoriale.

3.5 Osservazioni.

²ricordo che $\mathbb{R}[x]$ è l'insieme di tutti i polinomi reali

- (a) Lo zero³ in uno spazio vettoriale è unico. Infatti, se ci fossero due zeri distinti 0 e $0'$, allora si avrebbe $0 = 0 + 0' = 0'$.
- (b) Anche l'opposto è unico. Infatti supponiamo che w e w' siano entrambi opposti di u . Allora

$$w = w + 0 = w + (u + w') = (w + u) + w' = 0 + w' = w'.$$

- (c) Se $u + v = u$ allora $v = 0$, infatti

$$v = v + 0 = v + [u + (-u)] = (v + u) + (-u) = u + (-u) = 0.$$

- (d) Si ha $0u = 0$. Infatti

$$0 = 0u - 0u = (0 + 0)u - 0u = 0u + 0u - 0u = 0u.$$

- (e) L'opposto di u , la cui esistenza è garantita dalla proprietà (4) è dato $(-1)u$. Infatti

$$u + (-1)u = 1u + (-1)u = (1 - 1)u = 0u = 0.$$

Quindi $v = (-1)u$ è l'unico opposto di u , che si denota con il simbolo $-u$.

3.6 Definizione. Una *norma* su uno spazio vettoriale V è un'applicazione $\|\cdot\| : V \rightarrow \mathbb{R}$ tale che

- (N1) $\|v\| \geq 0$, per ogni $v \in V$
 (N2) $\|v\| = 0$ se e solo se $v = 0$.
 (N3) $\|cu\| = |c| \|u\|$ per ogni $c \in \mathbb{R}$ (\mathbb{C}), $u \in V$.
 (N4) $\|u + v\| \leq \|u\| + \|v\|$ per ogni $u, v \in V$.

Uno spazio vettoriale V dotato di una norma $\|\cdot\|$ è detto *spazio vettoriale normato* (ma no!).

3.7 Proposizione. Uno SVN^4 $(V, \|\cdot\|)$ diventa uno spazio metrico se si sceglie come distanza la funzione $d(x, y) := \|x - y\|$, per $x, y \in V$.

Dimostrazione. (facile, lasciata al lettore).

3.8 Definizione. Un sottoinsieme K di uno spazio vettoriale V è chiamato un *sottospazio* di V se K è uno spazio vettoriale, vale a dire se (ricorda quanto detto al punto 3.3)

- (A) $u, v \in K$ implica $u + v \in K$
 (B) $u \in K$ e $c \in \mathbb{R}$ implica $cu \in K$

3.9 Proposizione. Se $(V, \|\cdot\|)$ è uno spazio vettoriale normato e W è un sottospazio di V allora $\|\cdot\|$ è una norma su W , vale a dire $(W, \|\cdot\|)$ è anch'esso uno SVN .

Dimostrazione. Se le proprietà che definiscono una norma valgono per tutti i vettori di V , varranno automaticamente per gli elementi di W .

3.10 Problema. Dire se l'insieme indicato è o meno uno spazio vettoriale sul campo reale.

³usiamo lo stesso simbolo 0 sia per indicare il numero zero che il vettore zero, sperando che non sorga confusione

⁴spazio vettoriale normato

- (a) $\{x \in \mathbb{R}^3 : x_1 = x_2 - x_3\}$.
 (b) $\{x \in \mathbb{R}^3 : x_3^2 = x_1^2\}$.
 (c) $\{x \in \mathbb{R}^3 : x_3^3 = x_1^3\}$.
 (d) L'insieme di tutte le matrici 3×3 con determinante nullo.
 (e) L'insieme di tutte le funzioni reali derivabili.
 (f) L'insieme di tutte le funzioni reali che si annullano almeno in un punto.

Soluzione. (a) S. (b) N. Infatti i vettori $x = (1, 0, 1)$ e $y = (1, 0, -1)$ appartengono allo spazio ma la loro somma $x + y = (2, 0, 0)$ non vi appartiene. (c) S, perché la condizione è equivalente a $x_3 = x_1$. (d) N. (e) S. (f) N. Infatti le funzioni $f(x) = x$ e $g(x) = 1 - x$ appartengono allo spazio ma la loro somma $f + g = 1$ non vi appartiene.

3.3 Esempi: Spazi finito dimensionali

Nello spazio vettoriale \mathbb{R}^n delle ennuple di numeri reali (o, analogamente, in \mathbb{C}^n) è possibile introdurre una norma in molti (infiniti) modi diversi

3.11 Esempi. Sia $x = (x_1, \dots, x_3) \in \mathbb{R}^3$. Esempi di norme possibili sono

- (1) $\|x\| := |x_1| + |x_2| + |x_3|$.
 (2) $\|x\| := 3|x_1| + 5|x_2| + 13|x_3|$.
 (3) $\|x\| := (|x_1|^2 + |x_2|^2 + |x_3|^2)^{1/2}$. Questa è la cosiddetta *norma euclidea*.

I seguenti sono invece esempi di quantità che non sono norme perché violano almeno una della proprietà della Definizione 3.7.

- (a) $\|x\| := |x_1| + |x_2| - |x_3|$. Se scelgo $x := (0, 0, 1)$, ottengo $\|x\| = -1$, mentre la norma non può mai essere negativa.
 (b) $\|x\| := x_1 + x_2 + x_3$. Se scelgo $x = (-1, -1, -1)$, ho $\|x\| = -3$, violando di nuovo la positività.
 (c) $\|x\| := |x_1|^3 + |x_2|^3 + |x_3|^3$. Questa viola l'omogeneità. Infatti, se c è un numero reale,

$$\|cx\| = |cx_1|^3 + |cx_2|^3 + |cx_3|^3 = |c|^3 (|x_1|^3 + |x_2|^3 + |x_3|^3) = |c|^3 \|x\|,$$

mentre dovrebbe essere $\|cx\| = |c| \|x\|$.

- (d) $\|x\| = |x_1| + |x_3|$. Anche questa non può essere una norma perché è possibile trovare un vettore non nullo x tale che $\|x\| = 0$. Si prenda, ad esempio, $x = (0, 1, 0)$.

3.3.1 Le norme $\|\cdot\|_p$

Una classe molto importanti di norme che si possono introdurre in \mathbb{R}^n , \mathbb{C}^n e (vedremo) anche in spazi infinito dimensionali è dato dalle cosiddette *norme p*. Si sceglie un numero $p \in [1, \infty]$ (il valore ∞ è incluso) e si pone

$$(3.1) \quad \|x\|_p := \left[\sum_{i=1}^n |x_i|^p \right]^{1/p} \quad \|x\|_\infty := \max_{i=1, \dots, n} |x_i| \quad p \in [1, \infty].$$

Quando $p = 2$ ritroviamo la usuale *norma euclidea*. Iniziamo a far vedere che in questo caso si ha effettivamente una norma.

3.12 Proposizione. L'applicazione $\|\cdot\|_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ data da

$$\|x\|_2 := \left[\sum_{i=1}^n |x_i|^2 \right]^{1/2}$$

è una norma su \mathbb{R}^n .

Dimostrazione. È ovvio che $\|x\|_2$ è sempre positivo o nullo. Inoltre la quantità $\|x\|_2$ si annulla se e solo se $x = 0$. L'omogeneità segue da

$$\|cx\|_2 = \left[\sum_{i=1}^n |cx_i|^2 \right]^{1/2} = \left[|c|^2 \sum_{i=1}^n |x_i|^2 \right]^{1/2} = |c| \left[\sum_{i=1}^n |x_i|^2 \right]^{1/2} = |c| \|x\|_2.$$

Rimane dunque da dimostrare la disuguaglianza triangolare $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$. Siano quindi x e y due vettori in \mathbb{R}^n e sia t un numero reale arbitrario. Partendo dalla strabiliante disuguaglianza

$$\sum_{i=1}^n (tx_i + y_i)^2 \geq 0 \quad \forall t \in \mathbb{R}$$

si ottiene

$$(3.2) \quad t^2 \sum_{i=1}^n x_i^2 + 2t \sum_{i=1}^n x_i y_i + \sum_{i=1}^n y_i^2 \geq 0 \quad \forall t \in \mathbb{R}.$$

Ponendo

$$a := \|x\|_2^2 \quad b := \sum_{i=1}^n x_i y_i \quad c := \|y\|_2^2,$$

la (3.2) può essere riscritta come

$$at^2 + 2bt + c \geq 0 \quad \forall t \in \mathbb{R}.$$

Ma se questo trinomio di secondo grado è non negativo per *tutti* i valori di t , necessariamente si ha

$$b^2 - ac \leq 0,$$

quindi

$$\left[\sum_{i=1}^n x_i y_i \right]^2 \leq \|x\|_2^2 \|y\|_2^2,$$

e, di conseguenza,

$$(3.3) \quad \left| \sum_{i=1}^n x_i y_i \right| \leq \|x\|_2 \|y\|_2.$$

Possiamo quindi scrivere

$$\begin{aligned} \|x+y\|_2^2 &= \sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n (x_i^2 + y_i^2 + 2x_i y_i) \\ &\leq \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2 \|y\|_2 = (\|x\|_2 + \|y\|_2)^2. \end{aligned}$$

Prendendo infine la radice quadrata si ottiene

$$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2,$$

che è ciò che dovevamo dimostrare. \square

3.3.2 Disuguaglianze di Hölder e di Minkowski

La dimostrazione che $\|\cdot\|_p$ è una norma anche quando $p \neq 2$ è più complicata ed è basata sulla seguente disuguaglianza che è una generalizzazione della (3.3):

3.13 Proposizione. (*Disuguaglianza di Hölder*). Sia $p \geq 1$ e sia q tale che

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (\text{ovviamente anche } q \geq 1).$$

Allora si ha, per ogni coppia x, y di elementi di \mathbb{R}^n

$$(3.4) \quad \sum_{k=1}^n |x_k y_k| \leq \|x\|_p \|y\|_q$$

Dimostrazione. Il caso più semplice è quando $p = 1$ e, di conseguenza $q = \infty$. Infatti abbiamo

$$(3.5) \quad \sum_{k=1}^n |x_k y_k| \leq \max_k |y_k| \sum_{k=1}^n |x_k| = \|y\|_\infty \|x\|_1.$$

Passiamo quindi a considerare il caso $p > 1$. La disuguaglianza (3.4) è una conseguenza delle disuguaglianza numerica

$$(3.6) \quad ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \forall a, b \geq 0.$$

Faccio prima vedere come da questo segua la (3.4) e infine dimostro la (3.6). Poniamo

$$s_k = \frac{x_k}{\|x\|_p} \quad t_k = \frac{y_k}{\|y\|_q},$$

in modo tale da avere

$$(3.7) \quad \sum_{k=1}^n |s_k|^p = \frac{1}{\|x\|_p^p} \sum_{k=1}^n |x_k|^p = 1 \quad \sum_{k=1}^n |t_k|^q = \frac{1}{\|y\|_q^q} \sum_{k=1}^n |y_k|^q = 1.$$

Grazie alle (3.6), (3.7) posso scrivere

$$\sum_{k=1}^n |s_k t_k| \leq \frac{1}{p} \sum_{k=1}^n |s_k|^p + \frac{1}{q} \sum_{k=1}^n |t_k|^q = \frac{1}{p} + \frac{1}{q} = 1.$$

Di conseguenza

$$\sum_{k=1}^n |x_k y_k| = \|x\|_p \|y\|_q \sum_{k=1}^n |s_k t_k| \leq \|x\|_p \|y\|_q.$$

Abbiamo dunque dimostrato la disuguaglianza di Hölder. Rimane da dimostrare la (3.6). Consideriamo la funzione

$$f(u) := \frac{u^p}{p} + \frac{1}{q} - u \quad u \geq 0.$$

Voglio far vedere che f ha un minimo assoluto nel punto $u = 1$. Infatti, se derivo, ottengo

$$f'(u) = u^{p-1} - 1$$

Quindi $f' = 0$ se e solo se $u = 1$. Per concludere che si tratta di un minimo assoluto basta osservare che, poiché $p > 1$ e dunque $p - 1 > 0$, f' è negativa in $(0, 1)$ e positiva in $(1, \infty)$. Siccome $u = 1$ è il minimo assoluto della f possiamo scrivere

$$f(1) \leq f(u) \quad \forall u \geq 0,$$

cioè

$$u \leq \frac{u^p}{p} + \frac{1}{q} \quad \forall u \geq 0.$$

Pongo a questo punto $u = ab^{-q/p}$ e ottengo

$$ab^{-q/p} \leq \frac{a^p b^{-q}}{p} + \frac{1}{q}.$$

Moltiplicando ambo i membri per b^q , poiché $q - q/p = 1$, abbiamo

$$ab^{-q/p+q} = ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

□

Grazie alla disuguaglianza di Hölder si dimostra la disuguaglianza triangolare per $\|\cdot\|_p$ e quindi il fatto che $\|\cdot\|_p$ è una norma (le altre proprietà sono banali):

3.14 Proposizione. (*Disuguaglianza di Minkowski*). *Sia $p \geq 1$. Allora per ogni $x, y \in \mathbb{R}^n$ si ha*

$$(3.8) \quad \|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Dimostrazione. Se x, y sono due n -ple arbitrarie posso scrivere

$$(3.9) \quad \begin{aligned} \sum_{k=1}^n (|x_k| + |y_k|)^p &= \sum_{k=1}^n \left[(|x_k| + |y_k|)^{p-1} (|x_k| + |y_k|) \right] \\ &= \sum_{k=1}^n (|x_k| + |y_k|)^{p-1} |x_k| + \sum_{k=1}^n (|x_k| + |y_k|)^{p-1} |y_k| \\ &= \sum_{k=1}^n u_k |x_k| + \sum_{k=1}^n u_k |y_k|. \end{aligned}$$

in cui ho posto $u_k := (|x_k| + |y_k|)^{p-1}$. Sia q tale che $1/p + 1/q = 1$. Allora $q(p-1) = p$. Usando la disuguaglianza di Hölder ottengo

$$\begin{aligned} \sum_{k=1}^n |x_k| u_k &\leq \|x\|_p \left[\sum_{k=1}^n u_k^q \right]^{1/q} = \|x\|_p \left[\sum_{k=1}^n (|x_k| + |y_k|)^{q(p-1)} \right]^{1/q} \\ &= \|x\|_p \left[\sum_{k=1}^n (|x_k| + |y_k|)^p \right]^{1/q} \end{aligned}$$

e, in modo identico,

$$\sum_{k=1}^n u_k |y_k| \leq \|y\|_p \left[\sum_{k=1}^n (|x_k| + |y_k|)^p \right]^{1/q}.$$

Utilizzando queste due disuguaglianze nella (3.9) si arriva a

$$\sum_{k=1}^n (|x_k| + |y_k|)^p \leq (\|x\|_p + \|y\|_p) \left[\sum_{k=1}^n (|x_k| + |y_k|)^p \right]^{1/q}.$$

Dividendo ambo i membri per la quantità $\left[\sum_{k=1}^n (|x_k| + |y_k|)^p\right]^{1/q}$ si ottiene

$$\left[\sum_{k=1}^n (|x_k| + |y_k|)^p\right]^{1-1/q} \leq \|x\|_p + \|y\|_p,$$

vale a dire

$$(3.10) \quad \left[\sum_{k=1}^n (|x_k| + |y_k|)^p\right]^{1/p} \leq \|x\|_p + \|y\|_p.$$

D'altra parte la quantità che appare a sinistra nella (3.8) può essere facilmente stimata come

$$(3.11) \quad \|x + y\|_p = \left[\sum_{k=1}^n (|x_k + y_k|)^p\right]^{1/p} \leq \left[\sum_{k=1}^n (|x_k| + |y_k|)^p\right]^{1/p}.$$

Dalle (3.10), (3.11) segue la disuguaglianza di Minkowski (3.8) \square

Tanto la disuguaglianza di Hölder che quella di Minkowski possono essere estese al caso in cui, al posto di due ennuple, x e y abbiamo due successioni infinite oppure due integrali di funzioni continue. Valgono dunque le seguenti disuguaglianze

$$(3.12) \quad \sum_{i=1}^{\infty} |x_i y_i| \leq \left[\sum_{i=1}^{\infty} |x_i|^p\right]^{1/p} \left[\sum_{i=1}^{\infty} |y_i|^q\right]^{1/q}$$

$$(3.13) \quad \left[\sum_{i=1}^{\infty} |x_i + y_i|^p\right]^{1/p} \leq \left[\sum_{i=1}^{\infty} |x_i|^p\right]^{1/p} + \left[\sum_{i=1}^{\infty} |y_i|^p\right]^{1/p}$$

$$(3.14) \quad \int_{\mathbb{R}} |f(x)g(x)| dx \leq \left[\int_{\mathbb{R}} |f(x)|^p dx\right]^{1/p} \left[\int_{\mathbb{R}} |g(x)|^q dx\right]^{1/q}$$

$$(3.15) \quad \left[\int_{\mathbb{R}} |f(x) + g(x)|^p dx\right]^{1/p} \leq \left[\int_{\mathbb{R}} |f(x)|^p dx\right]^{1/p} + \left[\int_{\mathbb{R}} |g(x)|^p dx\right]^{1/p}$$

3.4 Concetti “metrici” negli spazi vettoriali normati

Uno spazio vettoriale normato $(V, \|\cdot\|)$ è un esempio particolare (molto particolare) di spazio metrico, in cui la distanza fra due punti $x, y \in V$ viene definita come

$$d(x, y) := \|x - y\| \quad x, y \in V.$$

Per questo motivo negli SVN si possono introdurre tutti i concetti che sono propri degli spazi metrici. Ne ricordiamo alcuni.

Convergenza, Cauchy. Si dice che successione $(x_n)_{n=1}^{\infty}$ di elementi di V è *convergente* a $x \in V$ se la successione numerica $\|x_n - x\|$ converge a zero, cioè se

$$\text{per ogni } \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che per ogni } n \geq N \text{ si ha } \|x_n - x\| < \varepsilon.$$

La successione $(x_n)_{n=1}^{\infty}$ si dice *di Cauchy* se

$$\text{per ogni } \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che per ogni } n, k \geq N \text{ si ha } \|x_n - x_k\| < \varepsilon.$$

Palle. Indico con $B_r^V(x)$ la *palla aperta* di raggio r con centro in x nello spazio V

$$B_r^V(x) := \{y \in V : \|y - x\| < r\}.$$

Se non c'è ambiguità rispetto allo spazio ambiente ometterò l'apice V scrivendo semplicemente $B_r(x)$.

Aperto. Un sottoinsieme A è detto *aperto* in X se per ogni $x \in A$ esiste $\varepsilon > 0$ tale che $B_\varepsilon^X(x) \subset A$.

Chiusura, chiuso. Ricordo che, se A è un sottoinsieme di uno SVN $(V, \|\cdot\|)$ allora la *chiusura di A* , indicata con \bar{A} , è l'insieme di tutti quegli elementi di V che si possono ottenere come limite di una successione di elementi di A , vale a dire

$$u \in \bar{A} \quad \Leftrightarrow \quad \exists \text{ una successione } (u_n) \text{ di elementi di } A \text{ tale che } u_n \rightarrow u$$

Quindi \bar{A} contiene *tutti* i punti di A più, eventualmente, l'insieme di tutti quei punti che non appartengono a A , ma che si possono ottenere come limite di punti di A . Chiamo l'insieme di questi punti A_{ple} (punti limite esterni). Quindi posso scrivere

$$\bar{A} = A \cup A_{ple}$$

3.15 Esempio. Nello spazio \mathbb{R}^2 con la usuale norma euclidea $\|\cdot\|_2$, sia

$$(3.16) \quad A := \{x = (x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 < 1\}.$$

In questo caso si ha evidentemente che i punti limite esterni sono quelli che si trovano sul bordo superiore del quadrato, cioè

$$A_{ple} = \{x \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 = 1\}.$$

quindi

$$\bar{A} = A \cup A_{ple} = \{x \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}.$$

Un insieme A è detto *chiuso* se $\bar{A} = A$, ovvero se non ci sono punti limite esterni, ovvero se tutti i punti limite appartengono all'insieme A . L'insieme A definito nella (3.16) non è dunque chiuso. In altre parole per dimostrare che un insieme è chiuso bisogna far vedere che se ho una successione convergente (x_1, x_2, \dots) di elementi di A allora anche il suo limite è un elemento di A . Insomma un insieme chiuso non permette di “uscire” dall'insieme con una successione convergente.

Occhio! Per qualche sottile motivo psicologico o a causa di quelle che Perek chiama “le fallaci seduzioni del ragionamento analogico”⁵, molti studenti sono indotti a pensare che un insieme è come un bar: o è aperto, o è chiuso. NON È COSÌ! Esistono tonnellate e tonnellate di insiemi che non sono né aperti né chiusi. Esempio: nello spazio \mathbb{R} l'insieme $A = [0, 1)$. Non è aperto perché se scelgo $x = 0$, non esiste alcun $\varepsilon > 0$ tale che $B_\varepsilon(0) \subset [0, 1)$. E non è neanche chiuso perché la successione $x_n = 1 - 1/n$ è una successione di elementi di A che però converge al punto 1 che *non* appartiene ad A . Un esempio più eclatante di insieme che non è né aperto né chiuso in \mathbb{R} è l'insieme \mathbb{Q} dei razionali. Esistono poi insiemi che sono sia aperti che chiusi. Nel caso in cui lo spazio ambiente sia uno spazio vettoriale normato V questi sono soltanto l'insieme vuoto e lo spazio stesso V . Se invece lo spazio ambiente è uno spazio metrico X non connesso ogni “componente connessa” di X è sia aperta che chiusa. Ad esempio nello spazio $X = [0, 1] \cup (2, 3)$ gli insiemi $A = [0, 1]$ e $B = (2, 3)$ sono sia aperti che chiusi (pensateci).

⁵George Perek, *La vita. Istruzioni per l'uso*, Torino, Einaudi, 1998.

Occhio! “Aperto” e “chiuso” sono concetti relativi che dipendono dallo spazio ambiente. Può accadere che lo stesso insieme A sia aperto (o chiuso) in uno spazio ambiente X ma non lo sia in uno spazio ambiente diverso Y . Consideriamo ad esempio

$$\begin{aligned} X &= \mathbb{R}^3 \\ Y &= \{x \in \mathbb{R}^3 : x_3 = 0\} \\ A &= \{x \in \mathbb{R}^3 : x_3 = 0 \text{ e } x_1^2 + x_2^2 < 1\} \end{aligned}$$

In questo caso A è aperto in Y , ma nello spazio ambiente X non è né aperto né chiuso. Nello spazio Y infatti le palle sono bidimensionali, cioè $B_\varepsilon^Y(x)$ è un disco bidimensionale, mentre nello spazio X , $B_\varepsilon^X(x)$ è una palla tridimensionale. È chiaro quindi che se $x = (x_1, x_2, 0)$ è un punto appartenente ad A allora esiste un $\varepsilon > 0$ abbastanza piccolo tale che $B_\varepsilon^Y(x) \subset A$, ma non sarà mai possibile avere $B_\varepsilon^X(x) \subset A$. Quindi A è aperto in Y ma non in X . Ovviamente A non è neanche chiuso in X perché la successione

$$x_n = \left(1 - \frac{1}{n}, 0, 0\right)$$

è una successione di punti di A convergente in X ma con un limite $(1, 0, 0)$ che non appartiene ad A .

3.5 Spazi infinito dimensionali: spazi di successioni

Così come una n -pla è una successione di n numeri reali (o complessi)

$$\mathbb{R}^n \ni x = (x_1, x_2, \dots, x_n) = (x_i)_{i=1}^n$$

una successione può essere considerata come un’infinitupla, un elemento⁶ cioè dello spazio \mathbb{R}^∞

$$\mathbb{R}^\infty \ni x = (x_0, x_1, x_2, x_3, \dots) = (x_i)_{i=0}^\infty$$

Quindi \mathbb{R}^∞ (\mathbb{C}^∞) definisco lo spazio di tutte le successioni reali (complesse).

Non ti far confondere dalla notazione! Ogni elemento di \mathbb{R}^∞ , vale a dire ogni *punto dello spazio* \mathbb{R}^∞ , è una successione, quindi una successione di elementi di \mathbb{R}^∞ è una *successione di successioni*. Di conseguenza servono due indici, uno che ci dice quale successione e l’altro che ci dice quale elemento della successione. Nel seguito useremo la seguente notazione:⁷ una successione in \mathbb{R}^∞ si denota con

$$(3.17) \quad (x^{(n)}) = (x^{(0)}, x^{(1)}, x^{(2)}, x^{(3)}, \dots) \quad x^{(n)} \in \mathbb{R}^\infty.$$

Ogni $x^{(n)}$ è a sua volta una successione di numeri reali, cioè

$$(3.18) \quad x^{(n)} = (x_0^{(n)}, x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \dots) \quad x_i^{(n)} \in \mathbb{R}.$$

⁶gli indici di una successione partiranno a volte da zero, altre volte da uno, senza una logica apparente

⁷tranne quando ce ne dimenticheremo

Ad esempio $x_7^{(4)}$ è il settimo elemento della quarta successione. Conviene a volte visualizzare una successione di successioni come una matrice infinita

$$(3.19) \quad \begin{array}{l} x^{(0)} = (x_0^{(0)}, x_1^{(0)}, x_2^{(0)}, \dots) \\ x^{(1)} = (x_0^{(1)}, x_1^{(1)}, x_2^{(1)}, \dots) \\ x^{(2)} = (x_0^{(2)}, x_1^{(2)}, x_2^{(2)}, \dots) \\ \vdots \\ x^{(n)} = (x_0^{(n)}, x_1^{(n)}, x_2^{(n)}, \dots) \\ \vdots \end{array}$$

Ogni *riga* di questa matrice è un punto nello spazio \mathbb{R}^∞ .

Uno spazio di successioni in cui avremo introdotto una norma è uno spazio vettoriale normato, quindi affermare che la successione (di successioni) $(x^{(0)}, x^{(1)}, x^{(2)}, \dots)$ è convergente a x significa, come al solito in uno SVN, che

$$(3.20) \quad \lim_{n \rightarrow \infty} \|x - x^{(n)}\| = 0.$$

Per cui quando si parla di convergenza senza specificare ulteriormente si intende la convergenza in norma data dalla (3.20). Si può però introdurre, accanto alla convergenza in norma, un'altra nozione di convergenza che è l'analogo della *convergenza puntuale* per le funzioni.

3.16 Definizione. Una successione di successioni $(x^{(n)})$ si dice *convergente puntualmente* alla successione x se converge ad x componente per componente, cioè se

$$(3.21) \quad \text{per ogni } k \in \mathbb{N}^* \text{ si ha } \lim_{k \rightarrow \infty} x_k^{(n)} = x_k.$$

Con riferimento alla rappresentazione (3.19) tramite una matrice infinita, la convergenza puntuale equivale alla convergenza di ciascuna *colonna* della matrice. Attenzione, questo vuol dire che bisogna mandare all'infinito l'indice di riga che appare in alto fra parentesi nella nostra notazione, come appare nelle (3.21).

3.17 Esempio. Consideriamo la successione nello spazio \mathbb{R}^∞ (successione di successioni) data da

$$x_k^{(n)} = k/n.$$

Nella rappresentazione di matrice infinita questa successione può essere scritta come

$$\begin{array}{l} x^{(1)} = (1, \quad 2, \quad 3, \quad \dots) \\ x^{(2)} = (1/2, 1, \quad 3/2, \dots) \\ \vdots \\ x^{(n)} = (1/n, 2/n, 3/n, \dots) \\ \vdots \end{array}$$

Si vede che il limite di ciascuna colonna è zero. E infatti, per ogni k fissato, se facciamo tendere n all'infinito otteniamo

$$\lim_{n \rightarrow \infty} x_k^{(n)} = \lim_{n \rightarrow \infty} k/n = 0.$$

Quindi la successione $(x^{(n)})_{n=1}^{\infty}$ nello spazio \mathbb{R}^{∞} converge puntualmente al vettore nullo di questo spazio cioè alla successione $y = (0, 0, \dots)$.

Sarebbe questo il momento perfetto per introdurre una norma nello spazio \mathbb{R}^{∞} , in modo da avere uno spazio vettoriale normato nel quale si possa parlare, ad esempio, di distanza fra due successioni. Sarebbe, dicevo, il momento di abbandonare ogni titubanza e tirare fuori questa norma, se non fosse che (ahimé) costruire, o meglio dimostrare che esiste una norma su \mathbb{R}^{∞} è un esercizio (possibile, ma) non esattamente elementare che, come spesso si legge nei testi seri, “esula dallo scopo di questo corso”. Comunque potete provarci:

✱ **3.18 Problema.** Trovare una norma sullo spazio \mathbb{R}^{∞} .

Se invece di una norma, ci accontentiamo di una distanza, allora le cose diventano molto più semplici. Abbiamo visto nell'Esempio 2.7 che se definiamo, ad esempio, per $x, y \in \mathbb{R}^{\infty}$,

$$(3.22) \quad d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k - y_k|}{1 + |x_k - y_k|},$$

allora d è una distanza su \mathbb{R}^{∞} . È facile verificare inoltre che questa distanza corrisponde proprio alla “convergenza puntuale” nello spazio delle successioni, vale a dire l'affermazione $d(x^{(n)}, x) \rightarrow 0$ è equivalente alla (3.21).

La distanza (3.22) non è una norma in quanto viola l'omogeneità, vale a dire la proprietà (N3) della norma. Allo scopo di studiare spazi vettoriali normati infinito dimensionali, quello che faremo noi sarà considerare dei sottospazi dello spazio \mathbb{R}^{∞} costruiti imponendo alcune restrizioni sulle successioni e su ciascuno di questi sottospazi introdurremo una o più norme opportune. Cominciamo.

3.5.1 Lo spazio delle successioni limitate: ℓ_{∞}

Indichiamo con il simbolo ℓ_{∞} lo spazio di tutte le successioni *limitate*. Una successione

$$x = (x_1, x_2, x_3, \dots)$$

appartiene a ℓ_{∞} se $\sup_i |x_i| < \infty$, ovvero se

$$(3.23) \quad \text{esiste } M \in \mathbb{R} \text{ tale che } |x_i| \leq M \text{ per ogni } i \in \mathbb{N}.$$

ℓ_{∞} è chiaramente uno spazio vettoriale (verifica gli assiomi, ricordando la dritta 3.3) e grazie alla condizione di limitatezza (3.23), è possibile definire, per ogni elemento x di ℓ_{∞} , la quantità

$$(3.24) \quad \|x\|_{\infty} = \sup_i |x_i|.$$

3.19 Osservazione. La norma (3.24) è definita con l'estremo superiore e non con il massimo perché il massimo dei moduli degli elementi della successione potrebbe non esistere. Consideriamo la successione limitata

$$x = (2/3, 3/4, 4/5, 5/6, \dots, (k-1)/k, \dots)$$

In questo caso si ha $\sup_k |x_k| = 1$, ma la quantità $\max_k |x_k|$ non esiste.

3.20 Proposizione. L'applicazione $\|\cdot\|_{\infty}$ è una norma su ℓ_{∞} .

Dimostrazione. Per far vedere che un qualcosa è una norma devo dimostrare 5 affermazioni

- (0) L'applicazione $\|\cdot\|_\infty$ è ben definita su tutto lo spazio ℓ_∞ (ad esempio non dà mai ∞ come risultato).
- (1) $\|x\|_\infty \geq 0$.
- (2) $\|x\|_\infty = 0$ se e solo se $x = 0$.
- (3) $\|cx\|_\infty = |c| \|x\|_\infty$ per ogni $c \in \mathbb{R}$, $x \in \ell_\infty$.
- (4) $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$ per ogni $x, y \in \ell_\infty$.

Punto (0). Grazie alla (3.23) sappiamo che se $x \in \ell_\infty$ allora $\sup_i |x_i|$ è necessariamente un numero finito, quindi $\|\cdot\|_\infty$ è ben definita.

Punto (1). Banale

Punto (2). Se $x = 0$ allora banalmente $\|x\|_\infty = 0$. D'altra parte se $\|x\|_\infty = 0$ allora $\sup_i |x_i| = 0$, il che implica che tutti gli elementi x_i sono nulli, che è equivalente a dire che $x = 0$.

Punto (3). Possiamo scrivere

$$\|cx\|_\infty = \sup_i |cx_i| = \sup_i (|c||x_i|) = |c| (\sup_i |x_i|) = |c| \|x\|_\infty.$$

Punto (4). Si ha

$$\|x + y\|_\infty = \sup_i (|x_i + y_i|) \leq \sup_i (|x_i| + |y_i|) \leq (\sup_i |x_i|) + (\sup_i |y_i|) = \|x\|_\infty + \|y\|_\infty.$$

Ho quindi dimostrato che $\|\cdot\|_\infty$ è una norma su ℓ_∞ .

3.21 Osservazione. Nello spazio ℓ_∞ non sarebbe stato possibile definire come norma la quantità

$$(3.25) \quad \|x\| := \sum_{i=1}^{\infty} |x_i|$$

perché avrebbe violato la proprietà (0) della Proposizione 3.20. Infatti posso scegliere la successione di “tutti uni”

$$x = (1, 1, 1, 1, \dots)$$

che, essendo limitata, appartiene a ℓ_∞ , ma se provo a calcolare $\|x\|$ come nella (3.25) ottengo ∞ . Analogamente non posso definire su ℓ_∞

$$\|x\| := \sup_i x_i$$

senza i moduli, perché violerei la proprietà (1'). Un altro “non esempio” è

$$\|x\| := \sup_{i:\text{pari}} |x_i| = \sup\{|x_2|, |x_4|, \dots\}.$$

In questo caso infatti avrei elementi x non nulli, come

$$x = (1, 0, 1, 0, 1, 0, 1, 0, 1, \dots)$$

che hanno norma uguale a zero.⁸

⁸e questo non è carino

3.22 Osservazione. Una delle prime novità che presentano gli spazi vettoriali infinito dimensionali è che non è più vero che un insieme è compatto se e solo se è chiuso e limitato. Vale ancora l'implicazione

$$\text{compatto} \Rightarrow \text{chiuso e limitato}$$

ma non quella contraria. Un esempio di insieme chiuso e limitato ma non compatto è la palla unitaria in ℓ_∞

$$(3.26) \quad B_1 := \{x \in \ell_\infty : \|x\|_\infty \leq 1\}.$$

Questo insieme è chiaramente chiuso e limitato (il suo diametro è uguale a 2), ma non è compatto, vale a dire posso trovare in B_1 successioni che non ammettono sottosuccessioni convergenti. Consideriamo la successione di elementi di B_1 data dai “vettori di base canonici”

$$(3.27) \quad \begin{aligned} e^{(1)} &= (1, 0, 0, 0, 0 \dots) \\ e^{(2)} &= (0, 1, 0, 0, 0 \dots) \\ e^{(3)} &= (0, 0, 1, 0, 0 \dots) \\ &\vdots \\ e^{(k)} &= (0, 0, \dots, 0, 1, 0, 0 \dots) \\ &\vdots \end{aligned}$$

Se $n \neq m$ otteniamo

$$(3.28) \quad \|e^{(n)} - e^{(m)}\|_\infty := \sup_i |e_i^{(n)} - e_i^{(m)}| = 1$$

Quindi la distanza fra due elementi qualsiasi di questa successione è uguale ad 1. Di conseguenza non può esistere alcuna sottosuccessione convergente.

La convergenza puntuale di una successione di successioni $(x^{(n)})$ è una nozione di convergenza piuttosto debole. È facile verificare che se una successione converge in norma $\|\cdot\|_\infty$ allora converge anche puntualmente. A maggior ragione converge puntualmente se converge rispetto a $\|\cdot\|_p$. Viceversa non è difficile far vedere che esistono successioni di successioni convergenti puntualmente ma non in $\|\cdot\|_\infty$.

3.23 Problema. Fare un esempio di una successione $(x^{(n)})$ di elementi di ℓ_∞ che converge puntualmente (componente per componente) a zero, ma non converge in norma, cioè

$$\lim_{n \rightarrow \infty} \|x^{(n)}\|_\infty \neq 0.$$

Potrebbe venire il sospetto che la convergenza puntuale sia necessariamente più debole di una *qualsiasi* convergenza in norma. Non è così.

✱ **3.24 Problema.** Trovare una norma $\|\cdot\|$ sullo spazio vettoriale ℓ_∞ tale che esista una successione $(x^{(n)})$ di elementi di ℓ_∞ che converge a zero in norma, ma non puntualmente. Esibire esplicitamente questa successione.

Schema di soluzione. Se $x \in \ell_\infty$ definisco

$$\|x\| := \limsup_k |x_k| + \sum_{k=0}^{\infty} \frac{1}{k^2} |x_k - x_{k+1}|.$$

Usando le proprietà del limsup non è difficile far vedere che questa è una norma. A questo punto consideriamo la successione di successioni

$$\begin{aligned} x^{(1)} &:= (1, 0, 0, 0, 0, 0, \dots) \\ x^{(2)} &:= (1, 1, 0, 0, 0, 0, \dots) \\ &\dots \\ x^{(n)} &:= (\underbrace{1, 1, \dots, 1}_{n \text{ termini}}, 0, 0, 0, \dots). \end{aligned}$$

Questa successione non converge puntualmente a zero, in quanto converge puntualmente alla successione

$$y = (1, 1, 1, 1, 1, 1, \dots).$$

Tuttavia se calcoliamo la norma di $x^{(n)}$ otteniamo

$$\|x^{(n)}\| = 1/n^2,$$

quindi $\|x^{(n)}\| \rightarrow 0$.

3.5.2 Lo spazio delle successioni convergenti a zero: ℓ_0

Sia ℓ_0 lo spazio di tutte le successioni reali $x = (x_0, x_1, x_2, \dots) = (x_i)_{i=0}^{\infty}$ che convergono a zero

$$(3.29) \quad x \in \ell_0 \Leftrightarrow \lim_{i \rightarrow \infty} x_i = 0.$$

Analogamente si definisce $\ell_0(\mathbb{C})$ come lo spazio delle successioni complesse che tendono a zero.

3.25 Problema. Dimostrare che ℓ_0 è uno spazio vettoriale (ricorda quanto detto al punto 3.3).

Siccome ogni successione convergente è necessariamente limitata (come si dimostra?), abbiamo l'inclusione

$$(3.30) \quad \ell_0 \subset \ell_{\infty}.$$

Quindi ℓ_0 è un sottospazio di ℓ_{∞} . Di conseguenza, come spiegato al punto 3.9, possiamo⁹ introdurre in ℓ_0 la stessa norma che abbiamo introdotto in ℓ_{∞} , vale a dire la norma $\|\cdot\|_{\infty}$. In questo modo $(\ell_0, \|\cdot\|_{\infty})$ diventa uno spazio vettoriale normato.

3.26 Problema. È possibile, nello spazio ℓ_0 , scegliere la seguente norma?

$$(3.31) \quad \|x\|_2 := \left[\sum_{i=0}^{\infty} |x_i|^2 \right]^{1/2}$$

⁹non siamo costretti

3.5.3 Gli spazi ℓ_p con $p \in [1, \infty)$

Dato $p \geq 1$ definisco ℓ_p come l'insieme di tutte le successioni $x = (x_0, x_1, x_2, \dots) = (x_i)_{i=0}^\infty$ tali che

$$(3.32) \quad \sum_{i=0}^{\infty} |x_i|^p < \infty.$$

Questa condizione mi permette di definire un'applicazione

$$\|\cdot\|_p : \ell_p \rightarrow \mathbb{R}$$

come

$$(3.33) \quad \|x\|_p := \left[\sum_{i=0}^{\infty} |x_i|^p \right]^{1/p}$$

3.27 Proposizione. *L'insieme ℓ_p è uno spazio vettoriale e l'applicazione $\|\cdot\|_p$ è una norma su ℓ_p .*

Dimostrazione. Poiché ℓ_p è un sottoinsieme di \mathbb{R}^∞ che è uno spazio vettoriale, per far vedere che ℓ_p è anch'esso uno spazio vettoriale è sufficiente far vedere che è chiuso rispetto alle somme e alla moltiplicazione per uno scalare (come spiegato al punto 3.3).

Assumo $x, y \in \ell_p$. Grazie alla definizione (3.32) so che

$$A := \sum_{i=0}^{\infty} |x_i|^p < \infty \quad \text{e} \quad B := \sum_{i=0}^{\infty} |y_i|^p < \infty.$$

Dalla disuguaglianza di Minkowski per le successioni (3.13) ottengo che

$$\left[\sum_{i=1}^{\infty} |x_i + y_i|^p \right]^{1/p} \leq \left[\sum_{i=1}^{\infty} |x_i|^p \right]^{1/p} + \left[\sum_{i=1}^{\infty} |y_i|^p \right]^{1/p} = A^{1/p} + B^{1/p}.$$

Di conseguenza

$$\sum_{i=1}^{\infty} |x_i + y_i|^p < \infty,$$

quindi il vettore $x + y$ appartiene a ℓ_p .

Si assuma ora $x \in \ell_p$ e $c \in \mathbb{R}$. Allora, usando di nuovo il fatto che $x \in \ell_p$, ottengo

$$\sum_{i=1}^{\infty} |cx_i|^p = |c|^p \sum_{i=1}^{\infty} |x_i|^p < \infty.$$

Dunque anche il prodotto cx appartiene a ℓ_p . Abbiamo così dimostrato che ℓ_p è uno spazio vettoriale.

Ora faccio vedere che $\|\cdot\|_p$ è una norma. Innanzitutto osservo che la quantità $\|x\|_p$ è ben definita per ogni $x \in \ell_p$. Inoltre è ovvio che $\|x\|_p \geq 0$ e che $\|x\|_p = 0$ se e solo se $x = 0$. La disuguaglianza triangolare (proprietà (N4) della definizione di norma 3.6) in questo caso altri non è che la disuguaglianza di Minkowski (3.13). Rimane da verificare che $\|cx\|_p = |c| \|x\|_p$, cosa che si fa in quattro microsecondi

$$\|cx\|_p := \left[\sum_{i=1}^{\infty} |cx_i|^p \right]^{1/p} = |c| \left[\sum_{i=1}^{\infty} |x_i|^p \right]^{1/p} = |c| \|x\|_p \quad \square$$

3.28 Osservazione. Se x è un elemento di un qualche spazio ℓ_p con $p \geq 1$, allora sappiamo che

$$\sum_{i=0}^{\infty} |x_i|^p < \infty.$$

Se poniamo $y_i := |x_i|^p$, otteniamo quindi che la sommatoria delle y_i è convergente. Di conseguenza la successione (y_0, y_1, y_2, \dots) deve tendere a zero

$$\lim_{i \rightarrow \infty} y_i = \lim_{i \rightarrow \infty} |x_i|^p = 0.$$

Ma allora si ha anche

$$\lim_{i \rightarrow \infty} x_i = 0, \text{ cioè } x \text{ è un elemento di } \ell_0.$$

Abbiamo quindi mostrato che ogni elemento di ℓ_p è anche un elemento di ℓ_0 , vale a dire

$$\ell_p \subset \ell_0.$$

3.29 Problema. Fare un esempio di una successione x che appartiene a ℓ_4 ma non a ℓ_3 .

A questo punto ci si può chiedere se esiste un rapporto di inclusione fra i vari spazi ℓ_p con p diversi. La risposta è affermativa: più grande è il valore di $p \in (1, \infty)$ e più grande è lo spazio ℓ_p ,¹⁰ come dimostriamo di seguito. Si ha dunque una catena di inclusioni (ricordate che p non è necessariamente un numero intero)

$$(3.34) \quad \ell_1 \subset \ell_2 \subset \ell_{2.1} \subset \ell_3 \subset \ell_{\sqrt{301}} \subset \dots \subset \ell_{1000} \subset \dots \subset \ell_0 \subset \ell_\infty$$

3.30 Proposizione. Se $1 \leq p < q$ allora si ha $\ell_p \subset \ell_q$.

Dimostrazione. Sia $x = (x_1, x_2, \dots)$ un elemento di ℓ_p . Voglio dimostrare che $x \in \ell_q$, vale a dire che

$$\sum_{k=1}^{\infty} |x_k|^q < \infty.$$

Sappiamo che $\ell_p \subset \ell_\infty$, quindi esiste $M \geq 0$ tale che

$$|x_k| \leq M \quad \forall k = 1, 2, \dots$$

A questo punto possiamo concludere nel modo seguente

$$(3.35) \quad \sum_{k=1}^{\infty} |x_k|^q = \sum_{k=1}^{\infty} |x_k|^p |x_k|^{q-p} \leq \sum_{k=1}^{\infty} |x_k|^p M^{q-p} = M^{q-p} \sum_{k=1}^{\infty} |x_k|^p.$$

Poiché M^{q-p} è una costante e la serie $\sum_{k=1}^{\infty} |x_k|^p$ è convergente per ipotesi, dalla (3.35) si deduce che

$$\text{la serie } \sum_{k=1}^{\infty} |x_k|^q \text{ è convergente. } \quad \square$$

¹⁰non confondetevi con il fatto che ℓ_0 è più grande di tutti gli ℓ_p . Il pedice 0 che appare in ℓ_0 ha un significato diverso

3.5.4 Lo spazio delle successioni finite: ℓ_f

ℓ_f è definito come l'insieme di tutte le *successioni finite*, vale a dire quelle successioni $x = (x_0, x_1, x_2, \dots) = (x_i)_{i=0}^{\infty}$, per le quali esiste K tale che $x_i = 0$ se $i > K$.

ℓ_f è uno spazio vettoriale (dimostrarlo!) e, poiché si ha l'inclusione ovvia

$$(3.36) \quad \ell_f \subset \ell_p \subset \ell_\infty \quad p \geq 1$$

possiamo trasformare ℓ_f in uno spazio vettoriale normato scegliendo la norma $\|\cdot\|_\infty$ oppure la norma $\|\cdot\|_p$ con $p \geq 1$.

3.5.5 Problemi

3.31 Problema. Data la seguente successione di successioni, calcolare le quantità $\|x^{(2)}\|_\infty$ e $\|x^{(5)}\|_3$

$$x_k^{(n)} := \frac{n^2}{2^{n+k}} \quad k = 1, 2, 3, \dots$$

Risp: $1/2$; $25/(32\sqrt[3]{7})$.

3.32 Problema. Consideriamo la successione di successioni $(x^{(n)})_{n=1}^{\infty}$ data da

$$x_k^{(n)} := \begin{cases} 1 & \text{se } k \leq n \\ 0 & \text{se } k > n \end{cases}$$

Calcolare $\|x^{(1)}\|_\infty$, $\|x^{(4)}\|_\infty$, $\|x^{(8)}\|_1$ e $\|x^{(n)}\|_4$.

3.33 Problema. Sia $(x^{(n)})$ una successione di elementi di ℓ_1 . Dimostrare che

$$\lim_{n \rightarrow \infty} \|x^{(n)}\|_1 = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \|x^{(n)}\|_\infty = 0$$

Far vedere che l'inverso è falso, trovando un esempio di una successione $(x^{(n)})$ tale che

$$\lim_{n \rightarrow \infty} \|x^{(n)}\|_\infty = 0 \quad \text{ma} \quad \lim_{n \rightarrow \infty} \|x^{(n)}\|_1 \neq 0.$$

3.34 Problema. Per quali valori di $\alpha \in \mathbb{R}$ la successione

$$x_k = \frac{1}{k^\alpha}$$

appartiene agli spazi ℓ_f ? ℓ_p ? ℓ_0 ? ℓ_∞ ?

Risp: mai; $\alpha > 1/p$; $\alpha > 0$; $\alpha \geq 0$.

3.35 Problema. Consideriamo la successione di successioni $(x^{(n)})_{n=1}^{\infty}$ data da

$$(a) \quad x_k^{(n)} := \begin{cases} 1 & \text{se } k \leq n \\ 0 & \text{se } k > n \end{cases} \quad (b) \quad x_k^{(n)} := \begin{cases} 2 & \text{se } k \leq n^2 \\ 0 & \text{se } k > n^2 \end{cases}$$

Calcolare $\|x^{(1)}\|_\infty$, $\|x^{(4)}\|_\infty$, $\|x^{(8)}\|_1$ e $\|x^{(n)}\|_4$.

Risp: (a) 1, 1, 8, $\sqrt[4]{n}$. (b) 2, 2, 128, $2\sqrt{n}$.

3.36 Problema. Siano α e β due numeri reali positivi. Sotto quale condizione (su α, β) la seguente serie è convergente per ogni $x \in \ell_5$?

$$\textcircled{1} \quad \sum_{k=1}^{\infty} \frac{|x_k|^\alpha}{k^\beta}$$

Soluzione. Se $\alpha \geq 5$ si ha $\ell_5 \subset \ell_\alpha$, quindi $x \in \ell_5$ implica $x \in \ell_\alpha$, per cui

$$\sum_{k=1}^{\infty} \frac{|x_k|^\alpha}{k^\beta} \leq \sum_{k=1}^{\infty} |x_k|^\alpha < \infty.$$

Quindi se $\alpha \geq 5$, la serie $\textcircled{1}$ è convergente se $x \in \ell_5$ per ogni $\beta > 0$.

Consideriamo ora il caso $\alpha < 5$. Grazie alla disuguaglianza di Hölder si ha, per ogni $p > 1$, ponendo $q = p/(p-1)$

$$\sum_{k=1}^{\infty} \frac{|x_k|^\alpha}{k^\beta} \leq \left[\sum_{k=1}^{\infty} |x_k|^{\alpha p} \right]^{1/p} \left[\sum_{k=1}^{\infty} \frac{1}{k^{\beta q}} \right]^{1/q}.$$

Scelgo $p = \alpha/5$. Di conseguenza $\alpha p = 5$ e quindi la prima serie al secondo membro è convergente per ogni $x \in \ell_5$. In questo caso si ottiene

$$q = \frac{p}{p-1} = \frac{5}{5-\alpha}.$$

Se $\beta q > 1$ anche la seconda serie al membro di destra è convergente. Sostituendo il valore trovato in precedenza per q , la condizione $\beta q > 1$ diventa

$$\textcircled{2} \quad \frac{\alpha}{5} + \beta > 1.$$

Abbiamo quindi dimostrato che la $\textcircled{2}$ è sufficiente a garantire la convergenza della serie $\textcircled{1}$ per ogni $x \in \ell_5$.

Facciamo ora vedere che la $\textcircled{2}$ è anche necessaria. Supponiamo infatti che non sia soddisfatta, vale a dire che

$$\textcircled{3} \quad \frac{\alpha}{5} + \beta \leq 1.$$

Consideriamo la successione

$$x_k = \frac{1}{k^{1/5} (\log(1+k))^{1/\alpha}}.$$

Si verifica facilmente che $x \in \ell_5$. Infatti

$$\sum_{k=1}^{\infty} |x_k|^5 = \sum_{k=1}^{\infty} \frac{1}{k (\log(1+k))^{5/\alpha}},$$

che risulta convergente in quanto dalla $\textcircled{3}$ e dal fatto che per ipotesi $\beta > 0$ segue che $\alpha < 5$. D'altra parte

$$\sum_{k=1}^{\infty} \frac{|x_k|^\alpha}{k^\beta} = \sum_{k=1}^{\infty} \frac{1}{k^{\alpha/5+\beta} (\log(1+k))}$$

e, poiché $\alpha/5 + \beta \leq 1$, questa serie diverge.

Abbiamo quindi dimostrato che la condizione richiesta è proprio la $\textcircled{2}$

3.6 Spazi infinito dimensionali: spazi di funzioni

3.37 Notazione. La stragrande maggioranza delle affermazioni che faremo sugli spazi di funzioni sono valide sia per spazi di funzioni da \mathbb{R} a \mathbb{R} che per spazi di funzioni da \mathbb{R} a \mathbb{C} . Quindi, ad esempio, indicheremo genericamente con $C(\mathbb{R})$ lo spazio delle funzioni continue su \mathbb{R} a valori reali oppure a valori complessi. In quei casi in cui vogliamo essere più specifici useremo i simboli $C(\mathbb{R}; \mathbb{R})$ e $C(\mathbb{R}; \mathbb{C})$ e analoghi, in cui viene indicato esplicitamente il codominio delle funzioni considerate.

3.6.1 Gli spazi $C[a, b]$, $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, $C_c(\mathbb{R})$

Si considerino i seguenti spazi di funzioni:

- (1) $C[a, b]$, l'insieme delle funzioni continue su $[a, b]$;
- (2) $C_b(\mathbb{R})$, l'insieme delle funzioni continue e limitate su \mathbb{R} ;
- (3) $C_0(\mathbb{R})$, l'insieme delle funzioni continue che convergono a zero all'infinito;
- (4) $C_c(\mathbb{R})$, l'insieme delle funzioni continue che hanno *supporto compatto*, vale a dire che sono nulle al di fuori di un intervallo finito.

3.38 Proposizione. *Gli insiemi $C[a, b]$, $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, $C_c(\mathbb{R})$ sono spazi vettoriali. Si ha inoltre $C_c(\mathbb{R}) \subset C_0(\mathbb{R}) \subset C_b(\mathbb{R})$. Infine l'applicazione*

$$\|f\|_u := \sup_{x \in [a, b]} |f(x)|$$

è una norma su $C[a, b]$, mentre l'applicazione

$$\|f\|_u := \sup_{x \in \mathbb{R}} |f(x)|$$

è una norma su ciascuno degli spazi $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, $C_c(\mathbb{R})$.

Dimostrazione. La somma di funzioni continue è continua. Il prodotto di una funzione continua per un numero (reale o complesso) è continua. Questo dimostra che $C[a, b]$ è uno spazio vettoriale. Con lo stesso ragionamento si dimostra che $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, $C_c(\mathbb{R})$ sono spazi vettoriali. Il fatto che $\|\cdot\|_u$ è una norma si dimostra esattamente come si dimostra nella Proposizione 3.20 che $\|\cdot\|_\infty$ è una norma su ℓ_∞ , modificando opportunamente la notazione. \square

3.39 Osservazione. Osserviamo che la convergenza rispetto alla norma $\|\cdot\|_u$ non è nient'altro che la *convergenza uniforme* di una successione di funzioni. Sia infatti f_n una successione di funzioni nello spazio $C_b(\mathbb{R})$ che converge uniformemente alla funzione $f \in C_b(\mathbb{R})$. Per definizione di convergenza uniforme, si ha

$$\forall \varepsilon > 0 \exists N \text{ tale che } \forall n \geq N |f(x) - f_n(x)| \leq \varepsilon \forall x \in \mathbb{R}.$$

Ma questa affermazione equivale a dire

$$\forall \varepsilon > 0 \exists N \text{ tale che } \forall n \geq N \|f - f_n\|_u \leq \varepsilon,$$

che a sua volta, per definizione di limite, può essere scritta come

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

il che vuol dire appunto che $f_n \rightarrow f$ rispetto alla norma $\|\cdot\|_u$.

3.40 Proposizione. *Se $f \in C_0(\mathbb{R})$, allora f è uniformemente continua.*

Dimostrazione. Sia $f \in C_0(\mathbb{R})$. Poiché $f(x)$ tende a zero per $x \rightarrow \pm\infty$, posso scrivere

$$(3.37) \quad \forall \varepsilon > 0 \exists L \text{ tale che se } |x| \geq L, \text{ allora } |f(x)| < \varepsilon/2.$$

Siccome ogni funzione continua su un intervallo chiuso e limitato è uniformemente continua, scegliendo come intervallo $[-L-1, L+1]$, si ottiene

$$(3.38) \quad \forall \varepsilon > 0 \exists \delta \text{ t.c. se } x, y \in [-L-1, L+1] \text{ e } |x-y| < \delta, \text{ allora } |f(x) - f(y)| < \varepsilon.$$

Possiamo ovviamente supporre che la quantità δ nella (3.38) sia sempre minore di 1. In questo modo se $|x-y| < \delta$ avrà (a) $|x| \geq L$ e $|y| \geq L$, oppure (b) $x, y \in [-L-1, L+1]$. Nel caso (a) usiamo la (3.37) e otteniamo

$$|f(x) - f(y)| \leq |f(x)| + |f(y)| < \varepsilon,$$

mentre nel caso (b) usiamo la (3.38). Possiamo quindi concludere che

$$(3.39) \quad \forall \varepsilon > 0 \exists \delta \text{ t.c. se } |x-y| < \delta, \text{ allora } |f(x) - f(y)| < \varepsilon,$$

che significa proprio che f è uniformemente continua su \mathbb{R} . □

3.41 Controesempio. Non è vero che ogni funzione in $C_b(\mathbb{R})$ è uniformemente continua. Un esempio di funzione continua e limitata sull'asse reale, ma non uniformemente continua è

$$f(x) = \cos(\pi x^2).$$

Sia infatti $\varepsilon_0 = 1$. Comunque si scelga $\delta > 0$ poniamo

$$x = \lceil \delta^{-1} \rceil \qquad y = \sqrt{x^2 + 1}.$$

Osserviamo che, se $x > 0$, si ha

$$x^2 + 1 \leq x^2 + 1 + \frac{1}{4x^2} = \left(x + \frac{1}{2x}\right)^2 \implies \sqrt{x^2 + 1} \leq x + \frac{1}{2x},$$

per cui

$$|y - x| = \sqrt{x^2 + 1} - x \leq \frac{1}{2x} = \frac{1}{2\lceil \delta^{-1} \rceil} \leq \frac{\delta}{2} < \delta.$$

D'altra parte

$$\begin{aligned} |f(x) - f(y)| &= |\cos(\pi x^2) - \cos(\pi y^2)| = |\cos(\pi x^2) - \cos(\pi(1 + x^2))| \\ &= 2|\cos(\pi x^2)| = 2|\cos(\pi \lceil \delta^{-1} \rceil)| = 2. \end{aligned}$$

Abbiamo quindi fatto vedere che esiste $\varepsilon_0 > 0$ tale che per ogni $\delta > 0$ posso trovare x, y con $|x-y| < \delta$ e $|f(x) - f(y)| \geq \varepsilon_0$. Vale a dire f non è uniformemente continua.

3.6.2 Gli spazi $C_p[a, b]$, $C_p(\mathbb{R})$ in cui $p \in [1, \infty)$

Se $p \geq 1$, $C_p[a, b]$ è l'insieme di tutte le funzioni continue sull'intervallo $[a, b]$ tali che $\int_a^b |f(x)|^p dx < \infty$. Su questo insieme definisco

$$(3.40) \quad \|f\|_p := \left(\int_a^b |f(x)|^p dx \right)^{1/p}.$$

Analogamente $C_p(\mathbb{R})$ è l'insieme di tutte le funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ continue tali che $\int_{\mathbb{R}} |f(x)|^p dx < \infty$ e definisco

$$(3.41) \quad \|f\|_p := \left(\int_{\mathbb{R}} |f(x)|^p dx \right)^{1/p}$$

3.42 Proposizione. Per ogni $p \in [1, \infty)$ gli insiemi $C_p[a, b]$ e $C_p(\mathbb{R})$ sono spazi vettoriali e le funzioni definite nelle (3.40) e (3.41) sono norme rispettivamente su $C_p[a, b]$ e $C_p(\mathbb{R})$.

Dimostrazione. La dimostrazione è banale in quanto tutte le complicazioni sono risolte dalla disuguaglianza di Minkowski (3.15). Supponiamo di voler far vedere, ad esempio, che $C_p(\mathbb{R})$ è uno spazio vettoriale. Poiché esso è un sottoinsieme di $C(\mathbb{R})$ è sufficiente dimostrare l'implicazione

$$(3.42) \quad f, g \in C_p(\mathbb{R}) \implies \alpha f + \beta g \in C_p(\mathbb{R}) \quad \forall \alpha, \beta \in \mathbb{C}.$$

Grazie alla disuguaglianza di Minkowski

$$\begin{aligned} \left[\int_{\mathbb{R}} |\alpha f(x) + \beta g(x)|^p dx \right]^{1/p} &\leq \left[\int_{\mathbb{R}} |\alpha f(x)|^p dx \right]^{1/p} + \left[\int_{\mathbb{R}} |\beta g(x)|^p dx \right]^{1/p} \\ &= |\alpha| \left[\int_{\mathbb{R}} |f(x)|^p dx \right]^{1/p} + |\beta| \left[\int_{\mathbb{R}} |g(x)|^p dx \right]^{1/p}. \end{aligned}$$

Grazie all'ipotesi che $f, g \in C_p(\mathbb{R})$, i due integrali che compaiono al membro a destra sono finiti, quindi

$$\int_{\mathbb{R}} |\alpha f(x) + \beta g(x)|^p dx < \infty,$$

vale a dire $\alpha f + \beta g \in C_p(\mathbb{R})$. L'implicazione (3.42) è dimostrata, quindi $C_p(\mathbb{R})$ è uno spazio vettoriale. Le altre proprietà si dimostrano in modo analogo. \square

3.43 Problema. Determinare per quali valori di $\alpha \in \mathbb{R}$ la funzione

$$(3.43) \quad f(x) := \frac{1}{(1+x^2)^\alpha}$$

appartiene a $C_3(\mathbb{R})$.

Soluzione. Per definizione $f \in C_3(\mathbb{R})$ se e solo se $\int_{\mathbb{R}} |f(x)|^3 dx < \infty$. Poiché

$$|f(x)|^3 \sim \frac{1}{|x|^{6\alpha}} \quad t \rightarrow \pm\infty$$

usando il criterio del confronto asintotico 1.17 e la (1.33), otteniamo che $f \in C_3(\mathbb{R})$ se e solo se $6\alpha > 1$, cioè $\alpha > 1/6$.

3.44 Problema. Dimostrare direttamente, senza usare la disuguaglianza di Cauchy-Schwarz, che se $f, g \in C_2(\mathbb{R})$ il loro prodotto fg appartiene a $C_1(\mathbb{R})$.

Soluzione. Se $u, v \in \mathbb{C}$, si ha

$$0 \leq (|u| - |v|)^2 = |u|^2 + |v|^2 - 2|uv|,$$

vale a dire

$$|uv| = \frac{1}{2}(|u|^2 + |v|^2).$$

Di conseguenza, se f e g appartengono a $C_2(\mathbb{R})$, otteniamo

$$\int_{\mathbb{R}} |f(x)g(x)| dx \leq \frac{1}{2} \left(\int_{\mathbb{R}} |f(x)|^2 dx + \int_{\mathbb{R}} |g(x)|^2 dx \right) < \infty$$

che implica $fg \in C_1(\mathbb{R})$.

3.45 Osservazione. Abbiamo visto al punto 3.28 che $\ell_p \subset \ell_0$ e nella Proposizione 3.30 che, se $p < q$, allora $\ell_p \subset \ell_q$. È possibile estendere queste affermazioni agli spazi di funzioni? È vero che $C_p(\mathbb{R}) \subset C_0(\mathbb{R})$? In altre parole, se assumo che

$$(3.44) \quad \int_{\mathbb{R}} |f(x)|^p dx < \infty$$

ne segue che $\lim_{x \rightarrow \pm\infty} f(x) = 0$? La risposta è no. Addirittura è possibile trovare funzioni f tali che vale la (3.44) ma f non è neppure limitata! Una volta capito come deve essere fatta una tale funzione si capisce anche come trovare un controesempio anche all'inclusione $C_p(\mathbb{R}) \subset C_q(\mathbb{R})$ per $p < q$. Ne discutiamo nel seguente esempio.

3.46 Esempio. (Una funzione $f \in C_1(\mathbb{R})$ non limitata). Definiamo una funzione continua $f : \mathbb{R} \rightarrow \mathbb{R}$ nel modo seguente. f è nulla sul semiasse negativo, mentre sul semiasse positivo il grafico di f presenta una serie di picchi a forma triangolare centrati in corrispondenza di ciascun intero positivo k , come mostrato nella figura 3.1. Sia ora b_k la lunghezza della base del triangolo corrispondente all'intero k e sia h_k la relativa altezza e scegliamo

$$b_k = \frac{1}{k^3} \quad h_k = k.$$

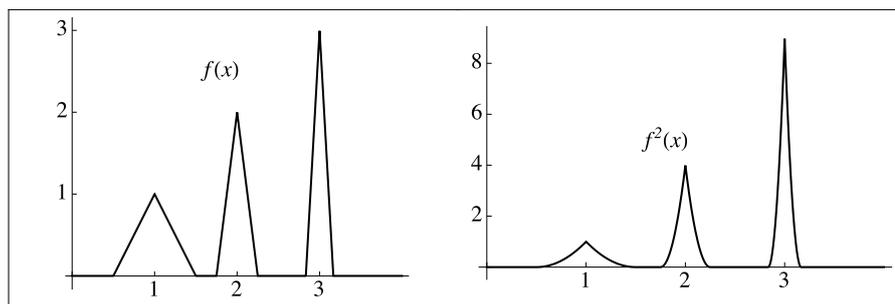
La funzione f è ovviamente illimitata in quanto $h_k \rightarrow +\infty$. D'altra parte si ha

$$\int_{\mathbb{R}} |f(x)| dx = \text{Area sottesa dal grafico di } f = \sum_{k=1}^{\infty} \frac{b_k h_k}{2} = \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty,$$

quindi $f \in C_1(\mathbb{R})$. Osservo anche che $f \notin C_2(\mathbb{R})$. Infatti il grafico di f^2 , mostrato in figura 3.1, è simile a quello di f e presenta dei "triangoli parabolici" di base b_k e altezza h_k^2 , in corrispondenza degli interi positivi. Siccome l'area di un triangolo parabolico è uguale a un terzo del prodotto base per altezza, otteniamo

$$\int_{\mathbb{R}} |f(x)|^2 dx = \text{Area sottesa dal grafico di } f^2 = \sum_{k=1}^{\infty} \frac{b_k h_k^2}{3} = \frac{1}{3} \sum_{k=1}^{\infty} \frac{1}{k} = \infty,$$

quindi $f \notin C_2(\mathbb{R})$.

Figura 3.1: Grafico di f e f^2 .

Riassumendo non c'è alcuna relazione di inclusione fra gli spazi $C_p(\mathbb{R})$ e $C_b(\mathbb{R})$ e fra i diversi spazi $C_p(\mathbb{R})$ per diversi p . Se però ci limitiamo a considerare funzioni *limitate*, allora è ancora vero che $f \in C_p(\mathbb{R})$ implica $f \in C_q(\mathbb{R})$ per $p \leq q$. Infatti abbiamo la seguente

3.47 Proposizione. *Se p e q sono due numeri reali tali che $1 \leq p \leq q$, allora vale l'inclusione*

$$C_p(\mathbb{R}) \cap C_b(\mathbb{R}) \subset C_q(\mathbb{R}).$$

Dimostrazione. Sia $f \in C_p(\mathbb{R}) \cap C_b(\mathbb{R})$. Poiché f è limitata, $M := \sup_{x \in \mathbb{R}} |f(x)|$ è finito. Allora

$$\int_{\mathbb{R}} |f(x)|^q dx = \int_{\mathbb{R}} |f(x)|^p |f(x)|^{q-p} dx \leq M^{q-p} \int_{\mathbb{R}} |f(x)|^p dx.$$

Poiché $f \in C_p(\mathbb{R})$ l'ultimo integrale è finito, quindi $f \in C_q(\mathbb{R})$. \square

3.6.3 Spazi di funzioni su \mathbb{R}^n . Multiindici

Abbiamo finora trattato spazi di funzioni definite su \mathbb{R} (o su un sottoinsieme di \mathbb{R}) a valori reali (o complessi). Le considerazioni fatte si possono più o meno facilmente (a seconda del tipo di considerazione) generalizzare al caso di funzioni di più variabili. Per trattare queste funzioni è comodo usare la notazione dei multiindici. Un multiindice n -dimensionale α è un'ennupla di interi non negativi

$$\alpha = (\alpha_1, \dots, \alpha_n) \quad \alpha_i \in \mathbb{N}$$

vale a dire $\alpha \in \mathbb{N}^n$. Il *modulo* di α è definito come

$$|\alpha| = \alpha_1 + \dots + \alpha_n.$$

Il *fattoriale* di α è definito come

$$\alpha! = \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n!.$$

Se $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, indichiamo inoltre con x^α il monomio

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

e con ∂^α la derivata parziale

$$\partial^\alpha = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

In analogia col caso unidimensionale, definiamo, ad esempio per ogni sottoinsieme aperto $\Omega \subset \mathbb{R}^n$ (quindi, in particolare si può avere $\Omega = \mathbb{R}^n$)

- (1) $C(\Omega)$ come l'insieme delle funzioni continue su Ω ;
- (2) $C_0(\Omega)$ come l'insieme delle funzioni continue f su Ω tali che per ogni $y \in \partial\Omega$ si ha $\lim_{x \rightarrow y} f(x) = 0$;
- (3) $C_c(\Omega)$ come l'insieme delle funzioni $f \in C(\Omega)$, tali che esiste un compatto $K \subset \Omega$ al di fuori del quale f è nulla;
- (4) $C^k(\Omega)$ come l'insieme delle funzioni $f : \Omega \rightarrow \mathbb{R}$ tale che per ogni multiindice α con $|\alpha| \leq k$, la derivata $\partial^\alpha f$ esiste ed è continua su Ω .

Il teorema di Taylor col resto sotto forma di integrale, trattato nella sezione 1.3.6 può essere generalizzato al caso di funzioni definite su (un sottoinsieme di) \mathbb{R}^n , la cui dimostrazione è molto simile al caso unidimensionale.

3.48 Teorema. (*Formula di Taylor in \mathbb{R}^n col resto sotto forma di integrale*). Sia Ω un aperto convesso in \mathbb{R}^n , sia $f \in C^k(\Omega)$ e $x_0 \in \Omega$. Allora, per ogni m tale che $0 \leq m < k$ e per ogni $x \in \Omega$ si ha

$$f(x) = \sum_{|\alpha| \leq m} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha + R_m(x)$$

con

$$R_m(x) = (m+1) \sum_{|\alpha|=m+1} \frac{(x-x_0)^\alpha}{\alpha!} \int_0^1 (1-t)^m \partial^\alpha f(x_0 + t(x-x_0)) dt.$$

3.49 Notazione. Se $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (oppure $x \in \mathbb{C}^n$) per maggior economia notazionale definiamo

$$|x| := \|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$

3.6.4 Lo spazio $C_c^\infty(\mathbb{R}^n)$ e lo spazio di Schwartz $\mathcal{S}(\mathbb{R}^n)$

Due spazi di funzioni che giocano un ruolo importante in molti campi dell'Analisi, ad esempio per quanto riguarda le distribuzioni e le trasformate di Fourier, sono lo spazio $C_c^\infty(\mathbb{R}^n)$ delle funzioni infinitamente derivabili a supporto compatto e lo spazio di Schwartz.

Lo spazio $\mathcal{S}(\mathbb{R}^n)$

Lo spazio di Schwartz, anche detto *spazio delle funzioni lisce rapidamente decrescenti*, viene usualmente denotato con $\mathcal{S}(\mathbb{R}^n)$ ed è definito come

$$(3.45) \quad \mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) : \text{per ogni coppia di multiindici } \alpha, \beta \text{ si ha } \|f\|_{\alpha, \beta} < \infty\},$$

in cui

$$(3.46) \quad \|f\|_{\alpha,\beta} := \|x^\alpha \partial^\beta f\|_u = \sup_{x \in \mathbb{R}} |x^\alpha \partial^\beta f(x)|.$$

Attenzione che, nonostante l'aspetto grafico della doppia barretta, le $\|\cdot\|_{\alpha,\beta}$ non sono norme ma *seminorme*, vale a dire può accadere che si abbia $\|f\|_{\alpha,\beta} = 0$ con $f \neq 0$. Ad esempio in una dimensione, se f è costante si ha $f' = 0$, quindi $\|f\|_{n,k} = 0$ per ogni $k > 0$. La (3.45) ci dice che una funzione di Schwartz f ha due proprietà:

- (a) è infinitamente differenziabile;
- (b) f e tutte le sue derivate parziali tendono a zero, per $|x| \rightarrow \infty$, più velocemente dell'inverso di una qualsiasi potenza.

Un'importante classe di esempi di funzioni $f \in \mathcal{S}(\mathbb{R}^n)$ è data da

$$(3.47) \quad f(x) = p(x) \exp(-a|x|^{2n}) \quad a > 0, n \in \mathbb{N}^*,$$

in cui p è un polinomio. Ad esempio, nel caso unidimensionale, le funzioni

$$e^{-x^2}, xe^{-x^2}, x^2e^{-x^2}, \dots, x^k e^{-x^2}, \dots$$

sono tutte funzioni di Schwartz.

A volte è più comodo usare una caratterizzazione delle funzioni di Schwartz leggermente diversa da quella che appare nella definizione (3.45).

3.50 Proposizione. *La funzione $f \in C^\infty(\mathbb{R}^n)$ appartiene a $\mathcal{S}(\mathbb{R}^n)$ se e solo se per ogni multiindice β e per ogni $a > 0$ esiste una costante $C_{a,\beta}$ tale che*

$$(3.48) \quad |\partial^\beta f(x)| \leq \frac{C_{a,\beta}}{(1+|x|^2)^{a/2}} \quad \forall x \in \mathbb{R}^n.$$

Dimostrazione. Per semplicità consideriamo il caso unidimensionale, $n = 1$. Sia $f \in C^\infty(\mathbb{R})$ e supponiamo che valga la (3.48). Se $j, k \in \mathbb{N}$, scegliendo $a = j$ nella (3.48), si ottiene

$$|x^j D^k f(x)| = |x^j| |D^k f(x)| \leq \frac{C_{j,k} |x|^j}{(1+x^2)^{j/2}} \leq C_{j,k} \quad \forall x \in \mathbb{R},$$

quindi

$$\|f\|_{j,k} \leq C_{j,k} < \infty.$$

Poiché j e k sono interi non negativi arbitrari, abbiamo dimostrato che $f \in \mathcal{S}(\mathbb{R})$.

Viceversa, sia $f \in \mathcal{S}(\mathbb{R})$. Devo dimostrare che per ogni $k \in \mathbb{N}$ e ogni $a > 0$ esiste $C_{a,k}$ tale che

$$(3.49) \quad |D^k f(x)| \leq \frac{C_{a,k}}{(1+x^2)^{a/2}} \quad \forall x \in \mathbb{R}.$$

Se $|x| \geq 1$ si ha

$$(1+x^2)^{a/2} \leq (x^2+x^2)^{a/2} = (2x^2)^{a/2} = 2^{a/2}|x|^a,$$

quindi, se $j = \lceil a \rceil$, vale a dire il più piccolo intero maggiore o uguale a a , ottengo

$$(3.50) \quad |D^k f(x)| = \frac{|x^j D^k f(x)|}{|x|^j} \leq \frac{\|f\|_{j,k}}{|x|^a} \leq \frac{2^{a/2} \|f\|_{j,k}}{(1+x^2)^{a/2}} \quad \text{se } |x| \geq 1.$$

Se invece $|x| \leq 1$, allora $(1+x^2)^{a/2} \leq 2^{a/2}$, quindi posso scrivere

$$(3.51) \quad |D^k f(x)| \leq \frac{2^{a/2} |D^k f(x)|}{(1+x^2)^{a/2}} \leq \frac{2^{a/2} \|f\|_{0,k}}{(1+x^2)^{a/2}} \quad \text{se } |x| < 1.$$

Se definisco $C_{a,k} := 2^{a/2} \max\{\|f\|_{[a],k}, \|f\|_{0,k}\}$ posso combinare le disuguaglianze (3.50) e (3.51) e ottenere così la (3.49). \square

Lo spazio $C_c^\infty(\mathbb{R}^n)$

Se una funzione ha supporto compatto, vale a dire è identicamente nulla al di fuori di una palla, la condizione di tendere a zero all'infinito più velocemente dell'inverso di una qualsiasi potenza è chiaramente soddisfatta, quindi si ha l'inclusione banale $C_c^\infty(\mathbb{R}^n) \subset S(\mathbb{R}^n)$. Inoltre, il fatto che una funzione di Schwartz f tende rapidamente a zero all'infinito garantisce che l'integrale

$$\int_{\mathbb{R}^n} |f(x)|^p dx$$

è finito per ogni $p \geq 1$, quindi $S(\mathbb{R}^n) \subset C_p(\mathbb{R}^n)$ per ogni $p \in [1, \infty)$.

Abbiamo visto nella (3.47) una classe di funzioni di Schwartz che hanno un'espressione esplicita molto semplice in termini di funzioni elementari. Scrivere un'espressione esplicita di una funzione $f \in C_c^\infty(\mathbb{R}^n)$ è leggermente più complicato e richiede un minimo di riflessione. Ok abbiamo riflettuto, eccola qua:

$$(3.52) \quad \varphi(x) := \begin{cases} c \exp\left[-\frac{1}{1-|x|^2}\right] & \text{se } |x| < 1 \\ 0 & \text{se } |x| \geq 1, \end{cases}$$

in cui la costante c è scelta in modo tale che valga $\int_{\mathbb{R}^n} \varphi(x) dx = 1$. φ è a supporto compatto per definizione. Per convincersi che è anche C^∞ il lettore è invitato a pensare al caso unidimensionale in cui gli unici punti in cui potrebbe esserci una discontinuità in una qualche derivata di φ sono i punti di raccordo $x = \pm 1$. Per simmetria è sufficiente occuparsi del punto $x = 1$. In sostanza, poiché a destra di $x = 1$ la funzione è identicamente nulla, quello che bisogna far vedere è che vale

$$\lim_{x \rightarrow 1} D^k \exp\left[-\frac{1}{1-x^2}\right] = 0 \quad \forall k \in \mathbb{N}.$$

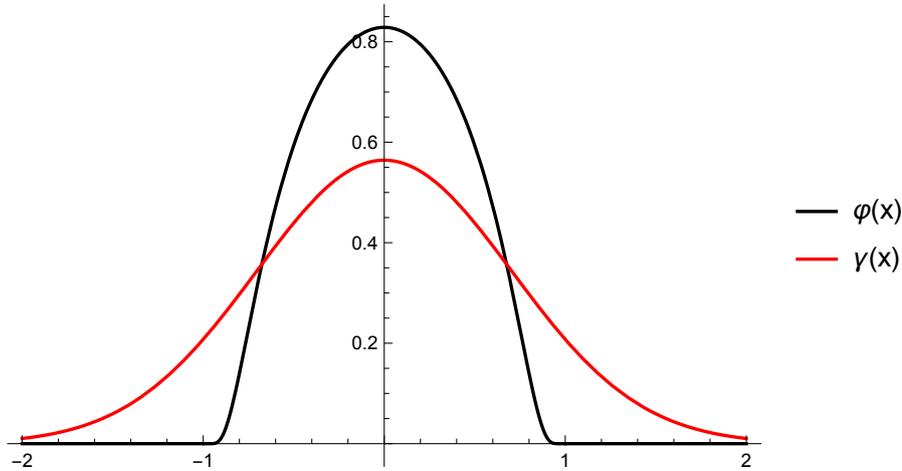
La figura 3.2 mostra il grafico di φ , confrontato con quello di una gaussiana, anch'essa normalizzata ad avere integrale uguale a 1

$$\gamma(x) := \frac{e^{-x^2}}{\sqrt{\pi}}.$$

Nel seguito in questa sezione ci limitiamo a considerare il caso $n = 1$, ma i risultati presentati si possono generalizzare a n arbitrario.

Bene, ora che abbiamo faticosamente costruito una funzione φ nello spazio $C_c^\infty(\mathbb{R})$, possiamo capitalizzare sul lavoro fatto e, utilizzando φ come “mollificatore” o “smussatore”, generare (quasi aggratis) un'infinità di altre funzioni lisce a supporto compatto. Ma c'è di meglio: data una funzione f continua o anche, più in generale, continua a tratti, a supporto compatto, possiamo costruire una funzione $\tilde{f}_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ che “assomiglia” ad f a meno di ε arbitrariamente piccolo (fra un po' saremo più precisi). Iniziamo col definire, per $\varepsilon > 0$, la funzione φ riscalata

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon} \varphi(x/\varepsilon),$$

Figura 3.2: Grafico di φ e γ .

che rappresenta una funzione simile alla φ , ma con il picco più stretto e più alto di un fattore ε , in modo tale che continui a valere $\int_{\mathbb{R}} \varphi_{\varepsilon}(x) dx = 1$. Definisco ora

$$(3.53) \quad \tilde{f}_{\varepsilon}(x) = (f * \varphi_{\varepsilon})(x) = \int_{\mathbb{R}} f(y) \varphi_{\varepsilon}(x - y) dy$$

che rappresenta il *prodotto di convoluzione* fra f e φ_{ε} ¹¹ e viene denotato con un asterisco $*$. Accade che, al tendere di ε a 0, \tilde{f}_{ε} si avvicina a f , ma rimane sempre una funzione liscia anche se f non lo è. In altre parole, operare la convoluzione con φ_{ε} ha l'effetto di "smussare gli spigoli" del grafico di f . Per questo motivo φ viene detto un mollificatore. Inoltre se f è a supporto compatto anche \tilde{f}_{ε} è a supporto compatto. Più precisamente si deduce facilmente dalla (3.53) che

$$\text{supp } \tilde{f}_{\varepsilon} \subset \{x \in \mathbb{R} : \text{dist}(x, \text{supp } f) \leq \varepsilon\}.$$

La figura 3.3 mostra come la funzione caratteristica dell'intervallo $[-1, 1]$

$$f(x) = \mathbb{1}_{[-1,1]} := \begin{cases} 1 & \text{se } x \in [-1, 1] \\ 0 & \text{se } |x| > 1 \end{cases}$$

viene approssimata sempre meglio, al decrescere di ε , dalle funzioni \tilde{f}_{ε} .

Fin qui abbiamo chiacchierato, ora veniamo al teorema che diamo senza dimostrazione.

3.51 Teorema. *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione continua a tratti a supporto compatto. Se \tilde{f}_{ε} è definita dalla (3.53), allora:*

- (1) per ogni $\varepsilon > 0$, $\tilde{f}_{\varepsilon} \in C_c^{\infty}(\mathbb{R})$;
- (2) per ogni $p \in [1, \infty)$ si ha $\lim_{\varepsilon \rightarrow 0^+} \|f - \tilde{f}_{\varepsilon}\|_p = 0$.

Grazie a questo teorema si ottiene:

¹¹il prodotto di convoluzione verrà discusso in modo più approfondito nella sezione 9.6.

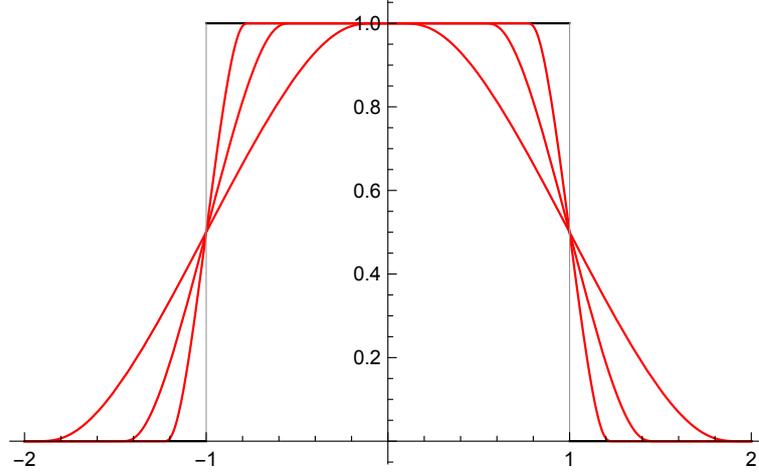


Figura 3.3: Approssimare una funzione a gradino f con funzioni lisce a supporto compatto \tilde{f}_ε con $\varepsilon = 1, 1/2, 1/4$.

3.52 Corollario. *Gli spazi $C_c^\infty(\mathbb{R})$ e $\mathcal{S}(\mathbb{R})$ sono entrambi densi in $(C_p(\mathbb{R}), \|\cdot\|_p)$ per ogni $p \in [1, \infty)$.*

Dimostrazione. Sia $p \in [1, \infty)$, sia $f \in C_p(\mathbb{R})$ e sia $\varepsilon > 0$. Siccome

$$\int_{\mathbb{R}} |f(x)|^p dx < +\infty$$

per il Corollario 1.15, esiste $L > 0$ tale che

$$\int_{|x| \geq L} |f(x)|^p dx < (\varepsilon/2)^p.$$

Se moltiplico f per la funzione caratteristica dell'intervallo $[-L, L]$, ottengo una funzione $g(x) := f(x)\mathbb{I}_{[-L, L]}(x)$ che non è più continua ma è comunque continua a tratti. Si ha

(3.54)

$$\|f - g\|_p = \left(\int_{\mathbb{R}} |f(x) - f(x)\mathbb{I}_{[-L, L]}(x)|^p dx \right)^{1/p} = \left(\int_{|x| \geq L} |f(x)|^p dx \right)^{1/p} < \varepsilon/2.$$

Poiché g è continua a tratti e a supporto compatto posso applicare il Teorema 3.51 a g e ottengo

$$\lim_{t \rightarrow 0^+} \|g - \tilde{g}_t\|_p = 0.$$

Dalla definizione di limite segue che esiste $\delta > 0$ tale che $\|g - \tilde{g}_\delta\|_p < \varepsilon/2$, che, insieme alla (3.54) implica

$$\|f - \tilde{g}_\delta\|_p \leq \|f - g\|_p + \|g - \tilde{g}_\delta\|_p < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Abbiamo quindi fatto vedere che per ogni $f \in C_p(\mathbb{R})$ e per ogni $\varepsilon > 0$ esiste $\tilde{g}_\delta \in C_c^\infty(\mathbb{R})$ tale che $\|f - \tilde{g}_\delta\|_p < \varepsilon$. Questo significa appunto che $C_c^\infty(\mathbb{R})$ è denso in $(C_p(\mathbb{R}), \|\cdot\|_p)$. Poiché $C_c^\infty(\mathbb{R})$ è un sottoinsieme di $\mathcal{S}(\mathbb{R})$, anche $\mathcal{S}(\mathbb{R})$ è necessariamente denso in $(C_p(\mathbb{R}), \|\cdot\|_p)$. \square

Un risultato analogo vale per lo spazio delle funzioni continue che tendono a zero all'infinito.

3.53 Teorema. *Gli spazi $C_c^\infty(\mathbb{R})$ e $\mathcal{S}(\mathbb{R})$ sono entrambi densi in $(C_0(\mathbb{R}), \|\cdot\|_u)$.*

3.6.5 Altri spazi vettoriali di funzioni

- (1) $C(\mathbb{R})$ l'insieme delle funzioni continue
- (2) $\mathbb{R}[x]$ l'insieme dei polinomi a coefficienti reali
- (3) $C^k(\mathbb{R})$ l'insieme delle funzioni derivabili k volte con la derivata k -sima continua
- (4) $C^k[a, b]$ l'insieme delle funzioni f derivabili k volte con derivata k -sima continua su (a, b) tali che per ogni $i = 0, 1, \dots, k$ esistono e sono finiti i limiti

$$\lim_{x \rightarrow a^+} f^{(i)}(x) \quad \lim_{x \rightarrow b^-} f^{(i)}(x)$$

- (5) $C_c^k(\mathbb{R})$ l'insieme delle funzioni $f \in C^k(\mathbb{R})$ con supporto compatto.
- (6) $C^\infty(\mathbb{R})$ l'insieme delle funzioni infinitamente differenziabili.
- (7) $C_0^\infty(\mathbb{R})$ l'insieme delle funzioni $f \in C^\infty(\mathbb{R})$ tali che $\lim_{x \rightarrow \pm\infty} f(x) = 0$.
- (8) $C_c^\infty(\mathbb{R})$ l'insieme delle funzioni infinitamente differenziabili con supporto compatto.

Riassunto inclusioni fra spazi vettoriali:

$$\begin{aligned} \ell_f &\subset \ell_1 \subset \ell_2 \subset \ell_3 \subset \ell_{3,1} \subset \ell_9 \subset \dots \subset \ell_0 \subset \ell_\infty \subset \mathbb{R}^\infty \\ C_c(\mathbb{R}) &\subset C_0(\mathbb{R}) \subset C_b(\mathbb{R}) \subset C(\mathbb{R}) \\ C_c(\mathbb{R}) &\subset C_2(\mathbb{R}) \subset C(\mathbb{R}) \\ \mathbb{R}[x] &\subset C^\infty(\mathbb{R}) \subset \dots \subset C^k(\mathbb{R}) \subset \dots \subset C^2(\mathbb{R}) \subset C^1(\mathbb{R}) \subset C(\mathbb{R}) \\ C_c^\infty(\mathbb{R}) &\subset C_0^\infty(\mathbb{R}) \subset C^\infty(\mathbb{R}) \end{aligned}$$

3.6.6 Problemi

3.54 Problema. Nei casi seguenti: se $\|\cdot\|$ è una norma dire semplicemente che è una norma, mentre se non lo è dimostrare esplicitamente che viola almeno una delle proprietà della norma.

- (a) $V = \mathbb{R}^3$ e $\|x\| = |x_1| + |x_3| + |x_2 - x_3|$
- (b) $V = C^3[a, b]$ e $\|f\| = f(0) + \sup_{x \in [a, b]} |f'(x)|$
- (c) $V = C_b(\mathbb{R})$ e $\|f\| = \int_{\mathbb{R}} \frac{|f(x)|}{1+|x|} dx$
- (d) $V = C_0(\mathbb{R})$ e $\|f\| = \int_{-\infty}^{\infty} |f(x)| dx$
- (e) $V = \mathbb{R}^3$ e $\|x\| = |x_1 + x_2 + x_3|$
- (f) $V = \ell_2$ e $\|x\| = \sum_{k=1}^{\infty} |x_k|$
- (g) $V = \ell_1$ e $\|x\| = \sum_{k=1}^{\infty} \frac{|x_k|}{k}$

Soluzione.

- (a) È una norma.
- (b) Non è una norma. Infatti si prenda $f(x) = -1$. Allora $\|f\| = -1$, mentre la norma deve essere non negativa.

(c) Non è una norma. Infatti si prenda $f(x) = 1$. Allora si ha

$$\|f\| = \int_{\mathbb{R}} \frac{dx}{1+|x|} = \infty$$

mentre la norma deve essere finita.

(d) Non è una norma. Infatti si prenda $f(x) = \frac{1}{1+|x|}$. Questa funzione appartiene a $C_0(\mathbb{R})$, però

$$\int_{\mathbb{R}} |f(x)| dx = \int_{\mathbb{R}} \frac{dx}{1+|x|} = \infty.$$

(e) No. Viola la positività stretta. Infatti se $x := (1, -1, 0)$ ottengo $\|x\| = 0$.

(f) No. La sommatoria può essere divergente. Sia infatti $x_k = 1/k$. Allora $(x_k) \in \ell_2$, ma $\sum_{k=1}^{\infty} 1/k = \infty$.

(g) Sì.

3.55 Problema. Determinare i valori di $\alpha \in \mathbb{R}$ per i quali la funzione f appartiene a $C_1(\mathbb{R})$ (lo spazio delle funzioni continue f tali che $\|f\|_1 < \infty$).

$$(a) f(x) := \frac{1}{(2+x^4)^\alpha} \quad (b) f(x) := \frac{(1+|x|)^\alpha}{e^{|x|}} \quad (c) f(x) := \frac{\exp\left[\left[\log(1+x^2)\right]^2\right]}{1+|x|^\alpha}$$

Risp: (a) $\alpha > 1/4$. (b) $\alpha \in \mathbb{R}$. (c) Mai. Il numeratore cresce più velocemente di qualsiasi potenza.

3.56 Problema. Determinare i valori di $\alpha \in \mathbb{R}$ per i quali la funzione indicata appartiene a $C_p(\mathbb{R})$ (lo spazio delle funzioni continue f tali che $\int_{\mathbb{R}} |f(x)|^p dx < \infty$).

$$(a) \frac{x \sin x}{(1+x^2+x^4)^\alpha} \quad (b) \frac{[\log(1+x^2)]^\alpha}{1+x^2} \quad (c) e^{-|x|^\alpha} \quad (d) \frac{|x|^\alpha}{1+x^8}$$

Risp: (a) $\alpha > \frac{1}{4} + \frac{1}{4p}$. (b) $\alpha \in \mathbb{R}$. (c) $\alpha > 0$. (d) $0 \leq \alpha < 8 - \frac{1}{p}$.

3.57 Problema. Determinare i valori di $\alpha \in \mathbb{R}$ per i quali la funzione f appartiene a $C_1(\mathbb{R})$ (lo spazio delle funzioni continue f tali che $\int_{\mathbb{R}} |f(x)| dx < \infty$).

$$(a) f(x) := \frac{1}{(2+x^6)^\alpha} \quad (b) f(x) := \frac{1+|x|}{5+e^{\alpha|x|}} \quad (c) f(x) := \frac{x^2 \cos x}{1+|x|^\alpha}$$

Soluzione. (a) $\alpha > 1/6$.

(b) $\alpha > 0$.

(c) $\alpha > 3$. Infatti se $\alpha > 3$ ottengo:

$$\int_{\mathbb{R}} |f(x)| dx \leq \int_{\mathbb{R}} \frac{x^2}{1+|x|^\alpha}$$

e quindi l'integrale è convergente. Far vedere che l'integrale $\int_{\mathbb{R}} |f(x)| dx$ è divergente per $\alpha \leq 3$ è leggermente più complicato a causa della presenza del coseno. Procedo nel modo seguente:

$$\begin{aligned} \int_{\mathbb{R}} |f(x)| dx &\geq \int_0^\infty |f(x)| dx = \int_0^{2\pi} |f(x)| dx + \int_{2\pi}^{4\pi} |f(x)| dx + \int_{4\pi}^{6\pi} |f(x)| dx + \dots \\ &\geq \int_0^{\pi/3} |f(x)| dx + \int_{2\pi}^{2\pi+\pi/3} |f(x)| dx + \int_{4\pi}^{4\pi+\pi/3} |f(x)| dx + \dots \\ &= \sum_{n=1}^{\infty} \int_{2n\pi}^{2n\pi+\pi/3} \frac{x^2 |\cos x|}{1+|x|^\alpha} dx. \end{aligned}$$

In ciascuno di questi intervalli $[2n\pi, 2n\pi + \pi/3]$ si ha $|\cos x| \geq 1/2$, mentre¹²

$$\frac{x^2}{1 + |x|^\alpha} \geq \frac{(2n\pi)^2}{1 + (2n\pi + \pi/3)^\alpha} \quad \forall x \in [2n\pi, 2n\pi + \pi/3]$$

Abbiamo quindi ottenuto

$$|f(x)| \geq \frac{1}{2} \frac{(2n\pi)^2}{1 + (2n\pi + \pi/3)^\alpha} \quad \forall x \in [2n\pi, 2n\pi + \pi/3]$$

Per cui

$$\int_{\mathbb{R}} |f(x)| dx \geq \frac{1}{2} \sum_{n=1}^{\infty} \frac{(2n\pi)^2}{1 + (2n\pi + \pi/3)^\alpha}$$

Per n grandi il termine n -simo della sommatoria va come $n^{2-\alpha}$, quindi la serie è divergente se $\alpha \leq 3$. A maggior ragione diverge l'integrale.

3.58 Problema. Fare un esempio di una funzione f che appartiene a $C_2(\mathbb{R})$ ma non a $C_1(\mathbb{R})$.

☆ **3.59 Problema.** Sia $f \in C_1(\mathbb{R})$. Dimostrare che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che comunque si scelga un intervallo $[a, b]$ di lunghezza $|b - a| < \delta$ allora si ha

$$\textcircled{1} \quad \int_a^b |f(x)| dx < \varepsilon.$$

Soluzione. Se le funzioni di $C_1(\mathbb{R})$ fossero tutte limitate allora sarebbe banale. Supponiamo infatti che sia $|f(x)| \leq M$ per tutti gli $x \in \mathbb{R}$. Allora basterebbe prendere $\delta := \varepsilon/M$ e la $\textcircled{1}$ sarebbe soddisfatta. Ma la vita è piena di amarezze e, in particolare, le funzioni di C_1 non sono necessariamente limitate.

Se potessimo considerare, invece che tutto l'asse reale, un intervallo chiuso e limitato, allora avremmo risolto, perché una funzione continua su tale intervallo ha un massimo e un minimo, quindi si avrebbe $|f(x)| \leq M$ su tutto l'intervallo. L'idea è di pensare \mathbb{R} come la somma di 3 pezzi

$$\mathbb{R} = (-\infty, -K) \cup [-K, +K] \cup (K, \infty)$$

in cui il pezzo importante è quello centrale, mentre i due pezzi all'esterno "contano poco". Per trasformare questi vagheggiamenti in matematica mi ricordo che se f è una funzione continua, allora, per il Corollario 1.15,

$$f \in C_1(\mathbb{R}) \iff \int_{\mathbb{R}} |f(x)| dx < \infty \iff \lim_{L \rightarrow +\infty} \int_{|x| \geq L} |f(x)| dx = 0.$$

Di conseguenza

$$\forall \varepsilon > 0 \exists L > 0 \text{ tale che } \int_{-\infty}^{-L} |f(x)| dx + \int_{+L}^{+\infty} |f(x)| dx < \frac{\varepsilon}{2}.$$

Bene, siamo quasi arrivati. Dato che f è continua in $[-L, +L]$ esisterà $M > 0$ tale che

$$|f(x)| \leq M \quad \forall x \in [-L, L].$$

¹²possiamo supporre $\alpha > 0$ altrimenti è ovvio che l'integrale diverge

Poniamo

$$\delta := \varepsilon/(2M).$$

In questo modo ottengo che se $[a, b]$ è un qualsiasi intervallo di lunghezza non superiore a δ , allora, ponendo $[a', b'] := [a, b] \cap [-L, L]$, posso scrivere

$$\begin{aligned} \int_a^b |f(x)| dx &\leq \int_{a'}^{b'} |f(x)| dx + \int_{-\infty}^{-L} |f(x)| dx + \int_{+L}^{+\infty} |f(x)| dx \leq \\ &\leq (b' - a') \max_{x \in [a', b']} |f(x)| + \frac{\varepsilon}{2} \leq \delta M + \frac{\varepsilon}{2} = \varepsilon. \quad \square \end{aligned}$$

3.7 Indipendenza lineare

3.60 Definizione. Un insieme di vettori u_1, \dots, u_k in uno spazio vettoriale reale (o complesso) V è detto *linearmente dipendente* se esistono k numeri reali (o complessi) c_1, \dots, c_k non tutti uguali a zero tali che

$$(3.55) \quad c_1 u_1 + \dots + c_k u_k = 0$$

u_1, \dots, u_k si dicono *linearmente indipendenti* se (sorpresa) non sono linearmente dipendenti. Un insieme infinito di vettori $\mathcal{A} := \{u_\alpha : \alpha \in I\}$ è detto *linearmente indipendente* se ogni sottoinsieme finito di \mathcal{A} è linearmente indipendente.

3.61 Proposizione. Un insieme di vettori u_1, \dots, u_k è linearmente dipendente se e solo se uno di questi vettori può essere espresso come combinazione lineare degli altri.

3.62 Definizione. La *dimensione* di uno spazio vettoriale V è il massimo numero di vettori linearmente indipendenti in V (questo numero può essere infinito, nel qual caso V è detto infinito dimensionale).

3.63 Esempio. Sia $V = P[0, 1]$. Allora i vettori

$$(3.56) \quad u_0 = 1, \quad u_1 = x, \quad u_2 = x^2, \quad u_3 = x^3, \quad u_4 = x^4$$

sono linearmente indipendenti. (Dimostrarlo. Suggerimento: si assuma $c_0 u_0 + \dots + c_4 u_4 = 0$ e si derivi alcune volte la precedente identità).

3.64 Problema. Dimostrare che le funzioni $1, e^x, e^{2x}, e^{3x}$ sono linearmente indipendenti in $C[0, 1]$.

3.65 Problema. Dimostrare che $1, \sin x, \sin(2x), \sin(3x)$ sono linearmente indipendenti in $C[0, 2\pi]$.

Soluzione. Assumiamo

$$(3.57) \quad c_0 1 + c_1 \sin x + c_2 \sin(2x) + c_3 \sin(3x) = 0$$

Attenzione: Lo “0” che appare a destra nella precedente uguaglianza è l’elemento zero in $C[0, 2\pi]$, cioè la funzione zero. Quindi la precedente uguaglianza significa

$$(3.58) \quad c_0 + c_1 \sin x + c_2 \sin(2x) + c_3 \sin(3x) = 0 \quad \text{per ogni } x \in [0, 2\pi].$$

Dobbiamo far vedere che i coefficienti c_i sono tutti uguali a zero. Il trucco consiste nello scegliere valori appropriati di x nella (3.58), in modo da trasformare la (3.58) in

un sistema di equazioni lineari la cui unica soluzione è $c_0 = \dots = c_3 = 0$. Infatti si ha

$$\begin{aligned} (x = 0) & \quad c_0 & = 0 \\ (x = \pi/2) & \quad c_0 + c_1 - c_3 & = 0 \\ (x = \pi/3) & \quad c_0 + \frac{\sqrt{3}}{2}(c_1 + c_2) & = 0 \\ (x = 2\pi/3) & \quad c_0 + \frac{\sqrt{3}}{2}(c_1 - c_2) & = 0 \end{aligned}$$

Questo sistema si risolve facilmente e l'unica soluzione è infatti $c_0 = \dots = c_3 = 0$ ¹³ \square

3.7.1 Come si dimostra che un insieme di vettori è linearmente indipendente?

Consideriamo diversi casi. V è il nostro spazio vettoriale.

- (1) $V = \mathbb{R}^n$ e avete n vettori. In questo caso con gli n vettori costruite una matrice quadrata $n \times n$ e calcolate il determinante. I vettori sono linearmente indipendenti se e solo se il determinante è diverso da zero.
- (2) $V = \mathbb{R}^n$ e avete m vettori con $m > n$. Uhm... questo è un caso segreto e non sono autorizzato a rivelarne la soluzione.
- (3) $V = \mathbb{R}^n$ e avete m vettori con $m < n$. Facciamo finta che i vettori siano $u_1 = (2, 3, 1, 0)$, $u_2 = (1, 5, 2, 1)$ e $u_3 = (-1, 2, 2, 1)$. Costruite la matrice 4×3 usando i 3 vettori come righe della matrice

$$(3.59) \quad \begin{bmatrix} 2 & 3 & 1 & 0 \\ 1 & 5 & 2 & 1 \\ -1 & 2 & 2 & 1 \end{bmatrix}$$

e riducete questa matrice a forma triangolare usando il vostro metodo preferito (metodo di Gauss descritto nella sezione 1.2.1, oppure usando un computer e scrivendo un programmetto, o meglio ancora convincendo vostro zio a farlo a posto vostro). Se alla fine ottenete una matrice con una o più righe di tutti zeri i vettori sono linearmente dipendenti, altrimenti sono linearmente indipendenti.

- (4) V è uno spazio di funzioni e avete n funzioni f_1, \dots, f_n . Bisogna considerare due sottocasi. Se pensate che le funzioni siano linearmente dipendenti allora bisogna provare a "intuire" una qualche "identità notevole" che potrebbe creare questa dipendenza lineare. Ad esempio le funzioni

$$3, \cos^2(x), \cos(2x)$$

sono linearmente dipendenti. Infatti, poiché

$$\cos(2x) = \cos^2(x) - \sin^2(x)$$

otteniamo

$$6 \cos^2(x) - 3 \cos(2x) = 3 \cos^2(x) + 3 \sin^2(x) = 3.$$

¹³vedremo in futuro un metodo migliore per dimostrare che le funzioni trigonometriche sono linearmente indipendenti (ma allora perché non abbiamo aspettato di imparare questo metodo per risolvere l'esercizio?)

Nel caso in cui si vuole invece dimostrare che le funzioni sono linearmente indipendenti ci sono due strategie che potrebbero funzionare. La prima è di considerare l'identità

$$(3.60) \quad c_1 f_1(x) + \cdots + c_n f_n(x) = 0$$

e derivarla un certo numero di volte (Esempio 3.63). Un'altra possibilità è quella di usare la (3.60) con n valori specifici x_1, \dots, x_n della variabile x scelti "opportunamente". In questo modo otteniamo un sistema di n equazioni lineari in n incognite e possiamo procedere con nel caso (1) (Problema 3.65). *Attenzione però*: se ottenente un determinante diverso da zero allora avete dimostrato che le funzioni sono linearmente indipendenti, *ma se il determinante è nullo non si può raggiungere alcuna conclusione!* Riprovate con valori diversi degli x_i .

3.8 Insiemi completi di vettori e basi

3.66 Definizione. Il *sottospazio generato* da un insieme di vettori $(u_\alpha)_{\alpha \in I}$ è il più piccolo sottospazio K tale che $u_\alpha \in K$ per ogni $\alpha \in I$. Questo sottospazio si denota con il simbolo

$$(3.61) \quad \text{span}\{u_\alpha : \alpha \in I\}$$

L'insieme dei vettori (u_α) può essere finito, numerabile o non numerabile.

In altre parole $\text{span}\{u_\alpha : \alpha \in I\}$ è l'insieme di tutti i vettori che possono essere espressi come *combinazione lineare finita* degli u_α , vale a dire quei vettori che si possono scrivere come

$$(3.62) \quad u = c_1 u_{\alpha_1} + \cdots + c_n u_{\alpha_n} \quad n \in \mathbb{N}, \alpha_i \in I$$

3.67 Esempio. Nello spazio vettoriale $C(\mathbb{R})$ consideriamo i vettori

$$u_0(x) = 1 \quad u_1(x) = x \quad u_2(x) = x^2 \quad \cdots \quad u_n(x) = x^n \quad \cdots$$

In questo caso si ha

$$\text{span}\{u_n : n \in \mathbb{N}\} = \mathbb{R}[x],$$

vale a dire l'insieme delle combinazioni lineari finite di $1, x, x^2, \dots$ non è nient'altro che l'insieme dei polinomi su \mathbb{R} .

3.68 Problema. Consideriamo nello spazio ℓ_0 i vettori

$$e^{(k)} := (0, 0, \dots, 0, \underset{(k)}{1}, 0, 0, \dots)$$

Identificare lo spazio generato da tutti questi vettori

$$(3.63) \quad \text{span}\{e^{(k)} : k = 1, 2, 3, \dots\} = ?$$

Nel caso di uno spazio vettoriale finito dimensionale V una base viene definita come una collezione di vettori u_1, u_2, \dots, u_n con due proprietà: innanzi tutto devono essere linearmente indipendenti, in secondo luogo deve valere

$$\text{span}\{u_1, \dots, u_n\} = V,$$

come a dire che ogni elemento di V si deve poter scrivere come combinazione lineare (necessariamente finita) degli u_n . Nel caso infinito dimensionale questa definizione è

troppo restrittiva per il seguente motivo. Prendiamo lo spazio ℓ_0 e i vettori della “base canonica”¹⁴

$$e^{(k)} := (0, 0, \dots, 0, \underset{(k)}{1}, 0, 0, \dots)$$

Poiché si ha $\text{span}\{e^{(k)} : k = 1, 2, \dots\} = \ell_f$ se utilizzassimo questa definizione potremmo dire che i vettori $e^{(k)}$ sono una base per ℓ_f ma non per ℓ_0 . Per questo motivo conviene modificare la definizione di base.

Ricordo la seguente definizione:

3.69 Definizione. Se Y e Z sono due sottoinsiemi di uno spazio metrico (X, d) si dice che Y è *denso in* Z se $\overline{Y} \supset Z$, vale a dire se per ogni $z \in Z$ e per ogni $\varepsilon > 0$ esiste $y \in Y$ tale che $d(z, y) < \varepsilon$.¹⁵ Un modo equivalente di dire la stessa cosa è il seguente: per ogni $z \in Z$ esiste una successione (y_n) di elementi di Y tale che $y_n \rightarrow z$.

Nel caso in cui Z coincida con l'intero spazio metrico X , dire che Y è denso in X equivale ovviamente alla condizione $\overline{Y} = X$.

3.70 Definizione. Un insieme di vettori $(u_\alpha)_{\alpha \in I}$ in uno spazio vettoriale normato $(V, \|\cdot\|)$ è detto *completo*¹⁶ se $\text{span}\{u_\alpha : \alpha \in I\}$ è denso in V , vale a dire se

$$\overline{\text{span}\{u_\alpha : \alpha \in I\}} = V$$

In altre parole dire che un insieme di vettori (u_α) è completo equivale a dire che ogni vettore $v \in V$, può essere *approssimato con precisione arbitraria* con una combinazione lineare finita degli u_α . Più precisamente deve valere la seguente condizione: per ogni $v \in V$ e per ogni $\varepsilon > 0$ esiste una combinazione lineare finita degli u_α

$$(3.64) \quad u = c_1 u_{\alpha_1} + \dots + c_n u_{\alpha_n}$$

tale che $\|u - v\| < \varepsilon$.

3.71 Definizione. Un insieme di vettori $(u_\alpha)_{\alpha \in I}$ in uno SVN $(V, \|\cdot\|)$ è detto una *base* se è un insieme di vettori linearmente indipendente e completo.

3.72 Problema. Dimostrare che ℓ_f è denso in $(\ell_0, \|\cdot\|_\infty)$ (di conseguenza i vettori unitari canonici $e^{(1)}, e^{(2)}, e^{(3)}, \dots$ sono una base in questo spazio).

Soluzione. Sia $x = (x_1, x_2, \dots)$ un elemento arbitrario in ℓ_0 . Faccio vedere che esiste una successione $(x^{(n)})$ di elementi di ℓ_f tale che $x^{(n)} \rightarrow x$, cioè tale che

$$\lim_{n \rightarrow \infty} \|x^{(n)} - x\|_\infty = 0.$$

Definisco la successione approssimante $x^{(n)}$ semplicemente mediante troncamenti della successione x che voglio ottenere come limite.

$$\begin{aligned} x^{(1)} &:= (x_1, 0, 0, 0, \dots) \\ x^{(2)} &:= (x_1, x_2, 0, 0, \dots) \\ &\vdots \\ x^{(n)} &:= (x_1, x_2, \dots, x_n, 0, 0, \dots) \\ &\vdots \end{aligned}$$

¹⁴in realtà ancora non ho detto cos'è una base

¹⁵nel caso di spazi vettoriali normati dovrà essere $\|z - y\| < \varepsilon$.

¹⁶da non confondere con la completezza di uno spazio metrico

Ora non resta che calcolare $\|x - x^{(n)}\|_\infty$ e far vedere che tende a zero. Infatti abbiamo

$$x - x^{(n)} = (0, 0, \dots, 0, 0, x_{n+1}, x_{n+2}, x_{n+3}, \dots),$$

per cui

$$\|x - x^{(n)}\|_\infty := \sup\{|x_{n+1}|, |x_{n+2}|, \dots\} = \sup_{k>n} |x_k| =: M_n.$$

Bisogna far vedere che

$$\lim_{n \rightarrow \infty} M_n = \lim_{n \rightarrow \infty} \sup_{k>n} |x_k| = \lim_{n \rightarrow \infty} \sup |x_n| = 0.$$

Ma poiché $x \in \ell_0$, per definizione si ha che x_n tende a zero, che, come è noto, implica che $\limsup_{n \rightarrow \infty} |x_n| = 0$. \square

3.73 Problema. Dimostrare che ℓ_f non è denso in $(\ell_\infty, \|\cdot\|_\infty)$.

3.74 Problema. Nello spazio vettoriale normato $(C_0(\mathbb{R}), \|\cdot\|_u)$ si consideri la funzione $f(x) := 1/(1+x^2)$. Determinare una funzione $g \in C_c(\mathbb{R})$ tale che $\|f - g\|_u \leq 1/10$. Disegnare il grafico di f e g .

3.75 Problema. Dimostrare che lo spazio $C_c(\mathbb{R})$ è denso in $(C_0(\mathbb{R}), \|\cdot\|_u)$, ma non in $(C_b(\mathbb{R}), \|\cdot\|_u)$.

Schema di soluzione. (1). Dimostrazione che $C_c(\mathbb{R})$ è denso in $C_0(\mathbb{R})$. Sia $f \in C_0(\mathbb{R})$. L'idea è di definire una successione di funzioni $f_n \in C_b(\mathbb{R})$ mediante un "troncamento" di f , cioè definendo $f_n = 0$ al di fuori dell'intervallo $[-n, n]$. Proprio così purtroppo la cosa non funziona, perché f_n in generale non sarà una funzione continua, quindi serve una costruzione leggermente più complicata. . .

(2). Dimostrazione che $C_c(\mathbb{R})$ non è denso in $C_b(\mathbb{R})$. Sia $f(x) = 1$. Questa è una funzione appartenente a $C_b(\mathbb{R})$, ma non c'è speranza alcuna di poterla approssimare con una funzione $g \in C_c(\mathbb{R})$, perché si ha necessariamente

$$\|f - g\|_u = \dots \quad \forall g \in C_c(\mathbb{R}).$$

È noto che in uno spazio vettoriale finito-dimensionale tutti i sottospazi (gli iperpiani) sono chiusi. Questa proprietà non è più vera nel caso infinito-dimensionale.

3.76 Problema. Dimostrare che ℓ_f non è chiuso in $(\ell_\infty, \|\cdot\|_\infty)$.

Soluzione. Consideriamo la seguente successione di elementi di ℓ_f

$$\begin{aligned} x^{(1)} &= (1, 0, 0, 0, 0, 0 \dots) \\ x^{(2)} &= (1, \frac{1}{2}, 0, 0, 0, 0 \dots) \\ x^{(3)} &= (1, \frac{1}{2}, \frac{1}{3}, 0, 0, 0 \dots) \end{aligned}$$

etc. È facile dimostrare che $x^{(n)}$ converge, nella norma $\|\cdot\|_\infty$, alla successione

$$(3.65) \quad x = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{n}, \dots).$$

Ma x non appartiene a ℓ_f , quindi x è un punto limite esterno di ℓ_f . Di conseguenza ℓ_f non è chiuso in $(\ell_\infty, \|\cdot\|_\infty)$.

3.77 Proposizione. ℓ_0 è un sottospazio chiuso di $(\ell_\infty, \|\cdot\|_\infty)$.

Dimostrazione. Dobbiamo far vedere che se

- (i) $x^{(n)}$ è una successione di elementi di ℓ_0 , e
- (ii) $\lim_{n \rightarrow \infty} \|x^{(n)} - x\|_\infty = 0$

allora $x \in \ell_0$, cioè x è una successione che converge a zero. Sia allora:

$$\begin{aligned} x^{(1)} &= (x_0^{(1)}, x_1^{(1)}, x_2^{(1)}, \dots) \\ x^{(2)} &= (x_0^{(2)}, x_1^{(2)}, x_2^{(2)}, \dots) \\ &\vdots \\ x^{(n)} &= (x_0^{(n)}, x_1^{(n)}, x_2^{(n)}, \dots) \\ &\vdots \\ x &= (x_0, x_1, x_2, \dots) \end{aligned}$$

Dobbiamo far vedere che $\lim_{i \rightarrow \infty} x_i = 0$, vale a dire che (per definizione)

$$(3.66) \quad \forall \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che } \forall i \geq N \quad |x_i| < \varepsilon.$$

Possiamo scrivere

$$(3.67) \quad |x_i| = |x_i - x_i^{(n)} + x_i^{(n)}| \leq |x_i - x_i^{(n)}| + |x_i^{(n)}| \leq \|x - x^{(n)}\|_\infty + |x_i^{(n)}|$$

Grazie all'ipotesi (ii) sappiamo che per ogni $\varepsilon > 0$ esiste N tale che se $n \geq N$ allora $\|x - x^{(n)}\|_\infty < \varepsilon$. Quindi, in particolare, possiamo scegliere $n = N$ e otteniamo

$$(3.68) \quad \|x - x^{(N)}\|_\infty < \varepsilon.$$

Ponendo $n = N$ nella (3.67) e usando la (3.68) troviamo

$$(3.69) \quad |x_i| \leq \|x - x^{(N)}\|_\infty + |x_i^{(N)}| \leq \varepsilon + |x_i^{(N)}|.$$

Di conseguenza

$$(3.70) \quad \lim_{i \rightarrow \infty} |x_i| \leq \varepsilon + \lim_{i \rightarrow \infty} |x_i^{(N)}|.$$

D'altra parte sappiamo che $x^{(N)}$ è un elemento di ℓ_0 , per cui $\lim_{i \rightarrow \infty} |x_i^{(N)}| = 0$. Questo implica che

$$(3.71) \quad \lim_{i \rightarrow \infty} |x_i| \leq \varepsilon.$$

Poiché ε è arbitrario, questo implica che $\lim_{i \rightarrow \infty} |x_i| = 0$ vale a dire che $x \in \ell_0$. \square

Seconda dimostrazione. Diamo una seconda dimostrazione meno diretta ma più corta. Vogliamo far vedere che $\ell_\infty \setminus \ell_0$ è aperto. Questo implica che ℓ_0 è chiuso. Al fine di dimostrare che $\ell_\infty \setminus \ell_0$ è aperto, facciamo vedere che per ogni $x \in \ell_\infty \setminus \ell_0$ esiste $\varepsilon > 0$ tale che la distanza fra x e ℓ_0 è almeno ε , vale a dire

$$(3.72) \quad \|x - y\|_\infty \geq \varepsilon \quad \forall y \in \ell_0.$$

Ciò significa che per ogni $x \in \ell_\infty \setminus \ell_0$ esiste $\varepsilon > 0$ tale che la palla $B_{\varepsilon/2}(x)$ è interamente contenuta in $\ell_\infty \setminus \ell_0$, quindi $\ell_\infty \setminus \ell_0$ è aperto e la dimostrazione è finita.

Sia allora $x \in \ell_\infty \setminus \ell_0$. Dobbiamo trovare $\varepsilon > 0$ tale che valga la (3.72). Siccome x non converge a zero sappiamo che

esiste $\varepsilon_0 > 0$ e una successione $x_{n(k)}$ tale che $|x_{n(k)}| \geq \varepsilon_0$ per ogni k .

Ma se $y \in \ell_0$, y converge a zero, quindi esiste N tale che se $i \geq N$ allora $|y_i| < \varepsilon_0/2$. Quindi abbiamo che se $n(k) \geq N$ allora

$$|x_{n(k)} - y_{n(k)}| \geq \varepsilon_0 - \varepsilon_0/2 = \varepsilon_0/2$$

e questo implica

$$(3.73) \quad \|x - y\|_\infty = \sup_{i \in \mathbb{N}} |x_i - y_i| \geq \varepsilon_0/2 \quad \forall y \in \ell_0$$

Dunque la (3.72) è stata dimostrata con $\varepsilon = \varepsilon_0/2$. \square

3.9 Separabilità

3.78 Definizione. Un insieme X è detto *numerabile* se esiste una corrispondenza biunivoca fra X e \mathbb{N} .¹⁷

In altre parole l'insieme X è numerabile se esiste un modo per “enumerare” tutti gli elementi di X .

3.79 Proposizione. (a) Il prodotto cartesiano di due insiemi numerabili è numerabile. (b) Un'unione numerabile di insiemi finiti o numerabili è numerabile, vale a dire se A_1, A_2, A_3, \dots sono numerabili allora la loro unione $\cup_{i=1}^\infty A_i$ è anch'essa numerabile.

Dimostrazione. (a) Sia X un insieme numerabile, quindi possiamo scrivere

$$X = \{x_1, x_2, x_3, \dots\}.$$

Il prodotto cartesiano $(X \times X)$ è l'insieme di tutte le coppie ordinate (x_i, x_j) . Voglio assegnare ad ogni coppia un intero positivo in maniera biunivoca. Per fare questo il trucco è quello di contare “in diagonale”. Chiamo, per brevità $x_{ij} := (x_i, x_j)$ e rappresento $X \times X$ come una matrice infinita

$$\begin{array}{cccc} x_{11} & x_{12} & x_{13} & \cdots \\ x_{21} & x_{22} & x_{23} & \cdots \\ x_{31} & x_{32} & x_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Devo ora assegnare un numero intero positivo a ogni elemento di questa matrice infinita. Procedo nel modo seguente: immagino di ruotare la matrice di 45 gradi in senso orario e di numerare poi i suoi elementi dall'alto verso il basso e da destra verso sinistra. In questo modo ottengo:

$$\begin{array}{cccccccc} x_{11} & x_{21} & x_{12} & x_{31} & x_{22} & x_{13} & x_{41} & x_{32} & \cdots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots \end{array}$$

¹⁷oppure fra X e l'insieme degli interi strettamente positivi $\{1, 2, 3, \dots\}$.

È chiaro che, con questo procedimento, prima o poi, avrò assegnato un intero positivo a ciascun elemento di $X \times X$ in modo biunivoco.

(b) Simile al caso precedente. \square

Grazie a questo risultato possiamo affermare che gli insiemi \mathbb{Z} e \mathbb{Z}^n sono numerabili. L'insieme \mathbb{Q} dei razionali è anch'esso numerabile perché ogni razionale q è un quoziente fra interi $q = n/k$, quindi ad ogni razionale possiamo associare un elemento di $(n, k) \in \mathbb{Z}^2$. Siccome \mathbb{Z}^2 è numerabile anche \mathbb{Q} è numerabile. C'è una piccola imprecisione in questo argomento poiché la corrispondenza fra \mathbb{Q} e \mathbb{Z}^2 non è biunivoca. Ad esempio al numero razionale $-3/4$ corrispondono le coppie $(-3, 4)$, $(3, -4)$, $(-6, 8)$, eccetera. Possiamo però stabilire una regola e ottenere una corrispondenza biunivoca fra \mathbb{Q} e un sottoinsieme di $\mathbb{Z} \times \mathbb{Z}$. Ora è facile vedere che un sottoinsieme di un insieme numerabile è finito o numerabile. Siccome \mathbb{Q} non è finito, deve essere numerabile.

Una volta stabilito che \mathbb{Q} è numerabile otteniamo che anche \mathbb{Q}^n è numerabile.

Se a qualcuno fosse a questo punto venuto il sospetto che tutti gli insiemi infiniti siano numerabili lo riportiamo subito alla dura realtà della matematica.

3.80 Proposizione. *Sia $X = \{0, 1\}^\infty$ l'insieme di tutte le successioni di zeri e uni. X non è numerabile.*

Dimostrazione. Supponiamo che X sia numerabile e troveremo una contraddizione. Se X è numerabile posso scrivere i suoi elementi come

$$X = (x^{(1)}, x^{(2)}, x^{(3)}, \dots)$$

in cui ciascun elemento è a sua volta una successione di zeri e uni (che chiamo “successioni binarie”). Abbiamo cioè una successione di successioni che rappresentiamo come una matrice infinita

$$(3.74) \quad \begin{array}{l} x^{(1)} = (x_1^{(1)}, x_2^{(1)}, \dots) \\ x^{(2)} = (x_1^{(2)}, x_2^{(2)}, \dots) \\ \vdots \\ x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots) \\ \vdots \end{array}$$

Ogni riga è una successione binaria e stiamo ipotizzando che in questo modo otteniamo tutte le successioni binarie possibili. Quindi se m'invento una successione binaria, ad esempio la successione

$$y = (1, 0, 1, 0, 1, 0, \dots),$$

ci dovrà essere una riga della matrice infinita in cui compare esattamente questa successione. Per cui se riuscissi a costruire una successione binaria che per qualche motivo non può essere uguale ad alcuna riga avrei trovato la contraddizione che sto cercando e avrei dimostrato il Teorema. La “successione impossibile” si costruisce nel modo seguente: si prende la “successione diagonale” e si cambiano tutti gli zeri in uno e tutti gli uni in zero. Per dirlo con una formula definisco

$$y = (1 - x_1^{(1)}, 1 - x_2^{(2)}, 1 - x_3^{(3)}, 1 - x_4^{(4)}, \dots).$$

perché affermo che questa successione non può essere uguale ad alcuna riga della matrice? Allora, y non può essere uguale alla prima riga $x^{(1)}$ perché y e $x^{(1)}$ differiscono sicuramente per quanto riguarda il loro primo elemento. y non può essere uguale alla

seconda riga perché il secondo elemento di y è sicuramente diverso dal secondo elemento di $x^{(2)}$. Insomma y non può essere uguale alla generica n -sima riga perché l' n -simo elemento di y differisce dall' n -simo elemento di $x^{(n)}$. Quindi y non appare come riga nella matrice infinita. Ma allora non è vero che X è l'insieme di tutte le successioni binarie. \square

3.81 Proposizione. *L'intervallo $[0, 1]$ non è numerabile.*

Schema di dimostrazione. Ogni numero reale $x \in [0, 1]$ può essere scritto in binario come una successione di zeri e uni, quindi l'enunciato discende dalla Proposizione 3.80 (Uhm... perché non è una dimostrazione vera e propria?)

Poiché $[0, 1]$ non è numerabile, non lo è neanche \mathbb{R} (e neanche \mathbb{R}^n). Tuttavia la non numerabilità contratta da \mathbb{R} è (fortunatamente) della forma “meno grave”. Il motivo è che \mathbb{R} , pur essendo non numerabile, ammette però un sottoinsieme numerabile \mathbb{Q} tale che ogni numero reale x può essere approssimato con precisione arbitraria da un elemento di \mathbb{Q} . Questa nozione è chiamata separabilità. Più precisamente:

3.82 Definizione. Uno spazio metrico (X, d) è detto *separabile* se esiste un insieme numerabile $Y \subset X$ tale che Y è denso in X , vale a dire¹⁸ per ogni $x \in X$ e per ogni $\varepsilon > 0$ esiste $y \in Y$ tale che $d(x, y) < \varepsilon$.

3.83 Esempio. \mathbb{R}^n e \mathbb{C}^n sono separabili (considera il sottoinsieme di tutti gli elementi a coordinate razionali).

3.84 Problema. Dimostrare che $\ell_f(\mathbb{Q})$ è numerabile (Sugg: scrivere $\ell_f(\mathbb{Q})$ come unione numerabile di insiemi numerabili...).

3.85 Problema. Dimostrare che $\ell_f(\mathbb{Q})$ è denso in $(\ell_0, \|\cdot\|_\infty)$. (Sugg: dimostrare che $\ell_f(\mathbb{Q})$ è denso in $\ell_f(\mathbb{R})$ ed usare il fatto noto che $\ell_f(\mathbb{R})$ è denso in ℓ_0).

Soluzione. Suppongo di aver già dimostrato che $\ell_f(\mathbb{R})$ è denso in $(\ell_0, \|\cdot\|_\infty)$. Quindi basta dimostrare che $\ell_f(\mathbb{Q})$ è denso in $\ell_0(\mathbb{R})$ rispetto alla norma $\|\cdot\|_\infty$ e usare il risultato che afferma che la proprietà di “essere denso in” è transitiva se riferita alla stessa norma.

Per far vedere che $\ell_f(\mathbb{Q})$ è denso in $\ell_f(\mathbb{R})$ dimostro che ogni elemento di $\ell_f(\mathbb{R})$ può essere approssimato con precisione arbitraria da un elemento di $\ell_f(\mathbb{Q})$. Sia dunque $x \in \ell_f(\mathbb{R})$

$$x = (x_1, x_2, \dots, x_N, 0, 0, \dots),$$

in cui x_N è l'ultimo elemento di x non nullo. Dato $\varepsilon > 0$ arbitrario, devo trovare $y \in \ell_f(\mathbb{Q})$ tale che $\|x - y\|_\infty < \varepsilon$. Poiché \mathbb{Q} è denso in \mathbb{R} esiste un razionale y_1 tale che $|y_1 - x_1| < \varepsilon$. Analogamente esiste $y_2 \in \mathbb{Q}$ tale che $|y_2 - x_2| < \varepsilon$. In generale si ha

$$\text{per ogni } n = 1, 2, \dots, N \text{ esiste } y_n \in \mathbb{Q} \text{ tale } |y_n - x_n| < \varepsilon.$$

Se $n > N$ si ha $x_n = 0$ per cui possiamo semplicemente porre

$$y_n = x_n = 0 \quad n = N + 1, N + 2, \dots$$

Ho quindi costruito una successione finita di razionali. Calcolo la distanza da x .

$$\|y - x\|_\infty = \sup_k |y_k - x_k| < \varepsilon. \quad \square$$

¹⁸ripeto per la millequattrocentoventisesima (si scrive con due q?) volta

3.86 Problema. Dimostrare che ℓ_p è separabile. (Sugg: usa il fatto che $\ell_f(\mathbb{Q})$ è numerabile).

3.87 Problema. Dimostrare la seguente affermazione: sia (X, d) uno spazio metrico e supponiamo che X contenga un insieme non numerabile di elementi $\{x_\alpha : \alpha \in I\}$ tale che per ogni $\alpha \neq \beta$ si ha $d(x_\alpha, x_\beta) \geq \frac{1}{10}$. Allora X non è separabile.

3.88 Problema. Dimostrare che $(\ell_\infty, \|\cdot\|_\infty)$ non è separabile. (Sugg: utilizzare il risultato del problema precedente).

3.10 Completezza. Spazi di Banach

In questa sezione discutiamo il problema della completezza di alcuni spazi vettoriali infinito dimensionali. Nel caso finito dimensionale il problema non si pone, perché tutti gli spazi vettoriali finito dimensionali su \mathbb{R} (su \mathbb{C}) sono isomorfi a \mathbb{R}^n (\mathbb{C}^n), quindi la loro completezza è assicurata dalla Proposizione 2.35. Nel caso infinito dimensionale le cose sono più complicate. Iniziamo con una definizione.

3.89 Definizione. Uno *spazio di Banach* è uno spazio vettoriale normato completo.

3.10.1 Spazi di successioni

3.90 Proposizione. $(\ell_\infty, \|\cdot\|_\infty)$ è completo.

Dimostrazione. Sia $x^{(n)}$ una successione di Cauchy in ℓ_∞ . Allora $\forall \varepsilon > 0$ esiste N tale che

$$(3.75) \quad \|x^{(n)} - x^{(m)}\|_\infty < \varepsilon \quad \forall n, m \geq N$$

il che implica

$$(3.76) \quad |x_i^{(n)} - x_i^{(m)}| < \varepsilon \quad \forall i \quad \forall n, m \geq N$$

Quindi, fissato l'indice i , si ha che la successione reale $(x_i^{(n)})_{n=1}^\infty$ è di Cauchy. Poiché \mathbb{R} è completo questa successione è convergente, vale a dire per ogni i esiste $x_i \in \mathbb{R}$ tale che

$$(3.77) \quad \lim_{n \rightarrow \infty} x_i^{(n)} = x_i$$

Consideriamo ora la successione $x = (x_1, x_2, x_3, \dots)$. Dobbiamo far vedere che:

(a) $x \in \ell_\infty$

(b) $x^{(n)}$ converge a x nello spazio ℓ_∞ , vale a dire $\lim_{n \rightarrow \infty} \|x^{(n)} - x\|_\infty = 0$

Prendendo il limite $m \rightarrow \infty$ nella (3.76) ottengo che $\forall \varepsilon > 0$ esiste N tale che

$$(3.78) \quad |x_i^{(n)} - x_i| \leq \varepsilon \quad \forall i \quad \forall n \geq N$$

Quindi se $n \geq N$ si ha

$$(3.79) \quad x^{(n)} - x \in \ell_\infty \quad \text{e} \quad \|x^{(n)} - x\|_\infty \leq \varepsilon$$

Ma allora x può essere scritta come somma $x = x^{(n)} + (x - x^{(n)})$ di due elementi di ℓ_∞ . Di conseguenza $x \in \ell_\infty$. D'altra parte la (3.79), per definizione, ci dice che $x^{(n)} \rightarrow x$. \square

3.91 Proposizione. Per ogni $p \geq 1$, $(\ell_p, \|\cdot\|_p)$ è completo.

Dimostrazione. Sia $(x^{(n)})$ una successione di Cauchy in ℓ_p . Per definizione di successione di Cauchy si ha

$$\forall \varepsilon > 0 \exists N > 0 \text{ tale che } \forall n, k \geq N \text{ si ha } \|x^{(n)} - x^{(k)}\|_p < \varepsilon.$$

Quindi se $n, k \geq N$

$$(3.80) \quad \sum_{i=1}^{\infty} |x_i^{(n)} - x_i^{(k)}|^p < \varepsilon^p.$$

Di conseguenza

$$(3.81) \quad |x_i^{(n)} - x_i^{(k)}| < \varepsilon \quad \forall i, \forall n, k \geq N.$$

Questo significa che per ogni i la successione di numeri reali $x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, \dots$ è una successione di Cauchy in \mathbb{R} . Ma fortunatamente \mathbb{R} è completo, quindi ognuna di queste successioni è convergente ad un qualche numero reale che chiamo x_i . Ho ottenuto dunque

$$\lim_{n \rightarrow \infty} x_i^{(n)} = x_i.$$

Consideriamo ora la successione

$$x := (x_1, x_2, x_3, \dots).$$

Voglio far vedere che

- (a) $x \in \ell_p$
- (b) $x^{(n)}$ converge a x in ℓ_p , vale a dire $\lim_{n \rightarrow \infty} \|x^{(n)} - x\|_p = 0$.

Se dimostro queste due affermazioni ho dimostrato che ℓ_p è completo.

► *Dimostrazione quasi giusta dei punti (a) e (b).* Scelgo $n \geq N$, poi uso la (3.80) facendo tendere k all'infinito. Ottengo così

$$(3.82) \quad \sum_{i=1}^{\infty} |x_i^{(n)} - x_i|^p \leq \varepsilon^p.$$

Questa disuguaglianza mi dice che se $n \geq N$ il vettore $x^{(n)} - x$ appartiene allo spazio ℓ_p e che la sua norma è $\|x^{(n)} - x\|_p \leq \varepsilon$. Ma il vettore x può essere scritto come somma

$$x = x^{(n)} + (x - x^{(n)})$$

di due vettori che appartengono entrambi a ℓ_p , per cui anche x appartiene ad ℓ_p . Contemporaneamente abbiamo fatto vedere che per ogni $\varepsilon > 0$ esiste un N tale che se $n \geq N$ allora $\|x - x^{(n)}\|_p \leq \varepsilon$, che è proprio la definizione del fatto che

$$\lim_{n \rightarrow \infty} \|x^{(n)} - x\|_p = 0, \text{ ovvero che } x^{(n)} \rightarrow x \text{ in } \ell_p.$$

Quindi abbiamo dimostrato che $x^{(n)}$ è una successione convergente. Di conseguenza ℓ_p è completo. *Fine della quasi dimostrazione* ◀

Al precedente argomento va fatta una piccola modifica, perché nell'espressione

$$\lim_{k \rightarrow \infty} \sum_{i=1}^{\infty} |x_i^{(n)} - x_i^{(k)}|^p \leq \varepsilon^p$$

al fine di ottenere la (3.82), abbiamo scambiato il limite con la serie, operazione che porta in qualche occasione a fare degli errori. Ma non in questo caso. Infatti, procedo ora in modo corretto e ottengo lo stesso risultato. Innanzitutto dalla (3.80) si deduce che

$$\forall L > 0 \quad \sum_{i=1}^L |x_i^{(n)} - x_i^{(k)}|^p \leq \varepsilon^p.$$

Ora ho una somma finita e posso effettivamente scambiare il limite per $k \rightarrow \infty$ con la somma

$$\lim_{k \rightarrow \infty} \sum_{i=1}^L |x_i^{(n)} - x_i^{(k)}|^p = \sum_{i=1}^L |x_i^{(n)} - x_i|^p \leq \varepsilon^p.$$

Ma questa disuguaglianza è valida *per ogni* L , quindi posso far tendere L all'infinito e ottengo

$$\lim_{L \rightarrow \infty} \sum_{i=1}^L |x_i^{(n)} - x_i|^p \leq \varepsilon^p$$

vale a dire

$$\sum_{i=1}^{\infty} |x_i^{(n)} - x_i|^p \leq \varepsilon^p.$$

Ora che ho di nuovo ottenuto la (3.82) *per benino* concludo come in precedenza. \square

3.92 Rimembranza. Ricordo quanto dimostrato nella Proposizione 2.35. Se $(V, \|\cdot\|)$ è uno spazio vettoriale normato *completo* e K è un sottospazio di V , allora $(K, \|\cdot\|)$ è completo se e solo se K è chiuso in V .

3.93 Proposizione. $(\ell_0, \|\cdot\|_{\infty})$ è completo.

Dimostrazione. La Proposizione 3.90 ci dice che lo SVN $(\ell_{\infty}, \|\cdot\|_{\infty})$ è completo. La Proposizione 3.77 afferma che $(\ell_0, \|\cdot\|)$ è chiuso in $(\ell_{\infty}, \|\cdot\|_{\infty})$. Quindi, per quanto osservato al punto 3.92, $(\ell_0, \|\cdot\|_{\infty})$ è completo. \square

3.94 Proposizione. $(\ell_f, \|\cdot\|_{\infty})$ non è completo.

Dimostrazione. Segue immediatamente da quanto osservato al punto 3.92 e dal risultato del Problema 3.76. \square

3.10.2 Spazi di funzioni

3.95 Proposizione. Lo spazio vettoriale normato $(C_b(\mathbb{R}), \|\cdot\|_u)$ è completo.

Dimostrazione. Sia $(f_n)_{n=1}^{\infty}$ una successione di Cauchy in $(C_b(\mathbb{R}), \|\cdot\|_u)$. Devo dimostrare che esiste una funzione f appartenente a $C_b(\mathbb{R})$ tale che

$$(3.83) \quad \lim_{n \rightarrow \infty} \|f - f_n\|_u = 0.$$

Dire che la successione (f_n) è di Cauchy nello spazio $(C_b(\mathbb{R}), \|\cdot\|_u)$ significa, per definizione, che

$$(3.84) \quad \forall \varepsilon > 0 \exists N \text{ tale che } \forall n, k \geq N \text{ si ha } \|f_n - f_k\|_u < \varepsilon.$$

Usando la definizione di $\|\cdot\|_u$ la precedente affermazione implica

$$(3.85) \quad \forall \varepsilon > 0 \exists N \text{ tale che } \forall n, k \geq N \text{ e } \forall x \in \mathbb{R} \text{ si ha } |f_n(x) - f_k(x)| < \varepsilon.$$

A questo punto conviene portare il “quantificatore” $\forall x \in \mathbb{R}$ all’inizio, ottenendo così

$$\forall x \in \mathbb{R} [\forall \varepsilon > 0 \exists N \text{ tale che } \forall n, k \geq N \text{ si ha } |f_n(x) - f_k(x)| < \varepsilon] .$$

Ora mi accorgo che il contenuto della parentesi [...] significa né più né meno che la *successione numerica* $(f_n(x))$ è di Cauchy. Questa deve essere la nostra giornata fortunata, perché \mathbb{R} è completo, quindi una successione *numerica* di Cauchy è necessariamente convergente. Dunque per ogni $x \in \mathbb{R}$ la successione $(f_n(x))$ deve avere un limite che chiamo subdolamente $f(x)$. Siamo arrivati a dire che

$$(3.86) \quad \forall x \in \mathbb{R} \text{ si ha } \lim_{n \rightarrow \infty} f_n(x) = f(x) .$$

La (3.86) ci dice che la successione di funzioni (f_n) converge *puntualmente* alla funzione f . Questo è un passo avanti, ma non è quello che ci serve. Quello che dobbiamo dimostrare infatti è che

- (a) $f \in C_b(\mathbb{R})$, cioè f è continua e limitata
- (b) vale la (3.83), cioè la successione (f_n) converge ad f *uniformemente*.

In realtà la (b) è sufficiente perché se abbiamo la convergenza uniforme, allora è noto che se una successione di funzioni continue e limitate converge uniformemente a qualcosa, il “qualcosa” è per forza continuo e limitato. Il problema è dunque quello di trasformare la convergenza puntuale in convergenza uniforme. Per far questo ripartiamo dalla (3.85) e facciamo il limite per $k \rightarrow \infty$. In questo modo otteniamo¹⁹

$$(3.87) \quad \forall \varepsilon > 0 \exists N \text{ tale che } \forall n \geq N \text{ e } \forall x \in \mathbb{R} \text{ si ha } |f_n(x) - f(x)| \leq \varepsilon ,$$

affermazione che può essere riscritta più semplicemente come

$$(3.88) \quad \forall \varepsilon > 0 \exists N \text{ tale che } \forall n \geq N \|f_n - f\|_u \leq \varepsilon .$$

Ma la (3.88) significa proprio che $\lim_{n \rightarrow \infty} \|f_n - f\|_u = 0$, esattamente quello che dovevo dimostrare. \square

3.96 Proposizione. *Lo spazio vettoriale normato $(C_p(\mathbb{R}), \|\cdot\|_p)$ non è completo.*

Schema di dimostrazione. Prendiamo ad esempio lo spazio $C_1[-1, 1]$, tanto non cambia niente. Per far vedere che $C_1[-1, 1]$ non è completo bisogna costruire una successione di funzioni (f_n) che sia di Cauchy ma non convergente. L’idea in questo caso è di prendere una successione di Cauchy che “converga a qualcosa che sta fuori dallo spazio $C_1[-1, 1]$ ”. In particolare si può pensare ad una successione di funzioni continue che converge ad una funzione non continua. Ovviamente bisognerà far vedere che questa successione è di Cauchy. Consideriamo la seguente successione di funzioni

$$f_n(x) = \begin{cases} 0 & \text{se } x \in [-1, 0] \\ nx & \text{se } x \in (0, 1/n) \\ 1 & \text{se } x \in [1/n, 1] \end{cases}$$

¹⁹attenzione al fatto che il $<$ facendo il limite diventa \leq

Faccio vedere che è una successione di Cauchy. Sia N un intero positivo e siano $n, k \geq N$. Supponiamo (ad esempio) che sia $n \geq k$. Allora

$$\begin{aligned} \|f_n - f_k\|_1 &= \int_{-1}^1 |f_n(x) - f_k(x)| dx \\ &= \int_{-1}^0 0 dx + \int_0^{1/n} (n-k)x dx + \int_{1/n}^{1/k} (1-kx) dx + \int_{1/k}^1 0 dx \\ &\leq \int_0^{1/k} 1 dx = \frac{1}{k} \leq \frac{1}{N}. \end{aligned}$$

Abbiamo così dimostrato che

$$\text{per ogni } n, k \geq N \text{ si ha } \|f_n - f_k\|_1 \leq 1/N,$$

che significa che la successione (f_n) è di Cauchy.

Ora è evidente che la successione (f_n) converge puntualmente alla funzione *discontinua*

$$f(x) = \begin{cases} 0 & \text{se } x \in [-1, 0] \\ 1 & \text{se } x \in (0, 1]. \end{cases}$$

Di conseguenza è impossibile che (f_n) possa convergere in norma $\|\cdot\|_1$ ad una qualche funzione continua g (dimostrarlo non è difficile). \square

3.10.3 Riassunto: strategia per dimostrare la completezza

Come si dimostra che uno spazio vettoriale normato è completo?

- (1) In generale quello che si fa è prendere una successione (u_n) di elementi di V , assumere che sia di Cauchy e dimostrare che esiste $u \in V$ tale che $u_n \rightarrow u$. In molti casi per indovinare chi è u si sfrutta il fatto che \mathbb{R} è completo (vedi la dimostrazione della Proposizione 3.90).
- (2) Una situazione in qualche senso più semplice è quando V è un sottospazio di uno spazio vettoriale normato più grande $V \subset W$ (V e W devono essere considerati con la stessa norma) e sappiamo già che W è completo. In questo caso per dimostrare che V è completo è sufficiente far vedere che V è chiuso in W . In pratica si fa così: si prende una successione (u_n) in V e si assume che questa successione sia convergente ad un qualche $u \in W$. A questo punto si dimostra che in realtà il punto limite u è un elemento di V (vedi l'Esempio 3.77).

Come si dimostra che uno spazio vettoriale normato non è completo?

Dimostrare la non completezza è di solito più facile che dimostrare la completezza. L'idea è quella di trovare una successione che sembri convergente, ma in realtà converge ad un punto che si trova "al di fuori" dello spazio V (vedi il Problema 3.76 o la Proposizione 3.96).

Ok, va bene, ma perché noi dovremmo sprecare il nostro tempo a dimostrare che uno spazio è (o non è) completo???

Ehm ...

3.97 Attenzione!. Se $(V, \|\cdot\|)$ è completo e K è un sottospazio chiuso di V allora K è automaticamente completo. Ad esempio abbiamo visto che ℓ_0 è un sottospazio chiuso dello SVN $(\ell_\infty, \|\cdot\|_\infty)$, e quindi $(\ell_0, \|\cdot\|_\infty)$ è anch'esso completo. Ma se lo spazio ambiente $(V, \|\cdot\|)$ non è completo allora un sottospazio chiuso di V potrebbe essere sia completo che non, mentre un sottospazio non chiuso è sicuramente non completo. Consideriamo lo SVN non completo $(\ell_f, \|\cdot\|_\infty)$ e consideriamo i seguenti tre sottospazi di ℓ_f :

- (1) K è l'insieme di tutte le successioni $x = (x_i)_{i=1}^\infty \in \ell_f$ tali che $x_i = 0$ per ogni $i > 3$. Quindi K è l'insieme di tutte le successioni tali che solo i primi tre elementi possono essere diversi da zero. È facile vedere che K è chiuso in $(\ell_f, \|\cdot\|_\infty)$ ed inoltre è completo;
- (2) K è l'insieme delle successioni x in ℓ_f in cui tutti gli elementi pari x_{2k} sono nulli. Questo spazio è chiuso in $(\ell_f, \|\cdot\|_\infty)$, ma non è completo. Il suo completamento è dato da ...
- (3) $K = \ell_f(\mathbb{Q})$, vale a dire l'insieme di tutte le successioni $x = (x_i)_{i=1}^\infty \in \ell_f$ tali che ogni elemento x_i è razionale. K non è chiuso in $(\ell_f, \|\cdot\|_\infty)$, infatti la chiusura di K è tutto lo spazio ℓ_f . Non essendo chiuso K non è neanche completo.

Possiamo fare un parallelo fra gli spazi ℓ_f , ℓ_0 , ℓ_∞ e gli spazi \mathbb{Q} , \mathbb{R} , \mathbb{C} .

$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$	$\ell_f \subset \ell_0 \subset \ell_\infty$
\mathbb{C} è completo	ℓ_∞ è completo
\mathbb{R} è chiuso in \mathbb{C} , quindi \mathbb{R} è completo	ℓ_0 è chiuso in ℓ_∞ , quindi ℓ_0 è completo
\mathbb{Q} è denso in \mathbb{R} , quindi $\overline{\mathbb{Q}} = \mathbb{R}$	ℓ_f è denso in ℓ_0 , quindi $\overline{\ell_f} = \ell_0$
\mathbb{Q} non è né chiuso né completo	ℓ_f non è né chiuso né completo

3.10.4 Lo spazio di Schwartz come spazio metrico completo

Nella sezione 3.6.4 abbiamo introdotto lo spazio di Schwartz $\mathcal{S}(\mathbb{R}^n)$ e abbiamo visto come questo sia un sottospazio denso in $(C_p(\mathbb{R}^n), \|\cdot\|_p)$ per ogni $p \in [1, \infty)$, quindi la chiusura di $\mathcal{S}(\mathbb{R}^n)$ in $C_p(\mathbb{R}^n)$ rispetto alla norma $\|\cdot\|_p$ è costituita da tutto lo spazio $C_p(\mathbb{R}^n)$. Essendo dunque $\mathcal{S}(\mathbb{R}^n)$ un sottospazio non chiuso, possiamo affermare che $(\mathcal{S}(\mathbb{R}^n), \|\cdot\|_p)$ non è completo. Lo stesso discorso si applica alla norma del sup $\|\cdot\|_u$: siccome $\mathcal{S}(\mathbb{R}^n)$ è denso in $(C_0, \|\cdot\|_u)$, come affermato nel Teorema 3.53, segue che $(\mathcal{S}(\mathbb{R}^n), \|\cdot\|_u)$ non è completo.

In futuro, quando tratteremo di trasformate di Fourier e distribuzioni, risulterà a volte conveniente introdurre una metrica in $\mathcal{S}(\mathbb{R}^n)$ rispetto alla quale questo spazio sia completo. Ho detto *metrica* e non norma perché allo scopo di indurre una topologia che sia "utile" nei contesti appena menzionati, non è possibile usare una norma, ma bisogna accontentarsi di una distanza.

Ci limitiamo, per semplicità di notazione, al caso unidimensionale. Ricordiamo che $f \in \mathcal{S}(\mathbb{R})$ se $f \in C^\infty$ e se, per ogni $j, k \in \mathbb{N}$, si ha

$$(3.89) \quad \|f\|_{j,k} := \|x^j D^k f\|_u = \sup_{x \in \mathbb{R}} |x^j D^k f(x)| < \infty.$$

Definiamo allora la seguente distanza

$$(3.90) \quad d(f, g) = \sum_{j,k=0}^{\infty} \frac{1}{2^{j+k}} \frac{\|f - g\|_{j,k}}{1 + \|f - g\|_{j,k}}.$$

Il fatto che si tratta effettivamente di una distanza di dimostra facilmente in modo molto simile al caso trattato nell'esempio 2.7. Questa distanza non è associabile ad una norma: infatti se lo fosse dovrebbe essere

$$\|f\| = d(f, 0) = \sum_{j,k=0}^{\infty} \frac{1}{2^{j+k}} \frac{\|f\|_{j,k}}{1 + \|f\|_{j,k}},$$

ma questa definizione viola la proprietà di omogeneità $\|cf\| = |c|\|f\|$ se $c \in \mathbb{C}$. Nel teorema seguente riassumiamo, senza dimostrazione, alcuni fatti importanti riguardanti lo spazio di Schwartz equipaggiato con la metrica (3.90).

3.98 Teorema. *Sia d la metrica definita nella (3.90). Allora:*

- (1) $(\mathcal{S}(\mathbb{R}), d)$ è uno spazio metrico completo;
- (2) la successione di funzioni di Schwartz $(f_n)_{n=1}^{\infty}$ converge a $f \in \mathcal{S}(\mathbb{R})$ se e solo se per ogni $j, k \in \mathbb{N}$ si ha

$$\lim_{n \rightarrow \infty} \|f - f_n\|_{j,k} = 0;$$

- (3) il sottospazio $C_c^{\infty}(\mathbb{R})$ è denso in $(\mathcal{S}(\mathbb{R}), d)$.

3.11 Bagliori di teoria della misura e spazi L_p

Abbiamo visto nella Proposizione 3.96 che gli spazi di funzioni $(C_p(\mathbb{R}), \|\cdot\|_p)$ non sono completi. Questi spazi sono molto importanti sia in matematica che in fisica. Vedremo ad esempio che se vogliamo introdurre la nozione di *prodotto scalare* in uno spazio di funzioni, la scelta “naturale” è quella di definire

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x) dx$$

nello spazio $(C_2(\mathbb{R}), \|\cdot\|_2)$. Il fatto che questi spazi non sono completi costituisce, come detto in precedenza, una notevole seccatura a causa della presenza di “buchi”, analogamente a quanto accade nei razionali, dove i buchi sono occupati dagli irrazionali. Da un punto di vista astratto, come accennato nella sezione 2.6, il problema viene risolto da un teorema che ci assicura che il completamento di $C_p(\mathbb{R})$ esiste ed è (essenzialmente) unico. Questo teorema non ci dice però “come è fatto” questo completamento. Che tipo di funzioni contiene? Nella dimostrazione della Proposizione 3.96, per far vedere che $C_1(\mathbb{R})$ non è completo abbiamo esibito una successione di Cauchy di funzioni continue che converge in $\|\cdot\|_1$ ad una funzione che non è continua, più precisamente ad una funzione continua a tratti. Si potrebbe dunque sperare che per completare $C_p(\mathbb{R})$ sia sufficiente aggiungere alle funzioni continue quelle continue a tratti. Ma non è così! La classe di funzioni che bisogna aggiungere per arrivare ad avere uno spazio completo è molto più ampia di quella delle funzioni continue a tratti. Si tratta delle funzioni *misurabili secondo Lebesgue*. Questo insieme di funzioni è talmente grande che per dimostrare che esistono funzioni non misurabili è necessario utilizzare l'assioma della scelta. Se denotiamo, in generale, con $\mathbb{1}_A$ la *funzione indicatrice* (o *funzione caratteristica*) dell'insieme A , definita come

$$(3.91) \quad \mathbb{1}_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \in A^c, \end{cases}$$

allora un esempio di funzione non continua in alcun punto, ma misurabile secondo Lebesgue è $\mathbb{1}_{\mathbb{Q}}$. A questo punto si capisce che ci aspetta un problema: dovendo costruire il completamento dello spazio C_p è chiaro che sarà necessario definire la norma $\|\cdot\|_p$ per questo tipo di funzioni, quindi sarà necessario definire l'integrale di funzioni misurabili secondo Lebesgue. Ma quanto vale, ad esempio,

$$(3.92) \quad \int_0^1 \mathbb{1}_{\mathbb{Q}}(x) dx ?$$

Se usiamo l'integrale di Riemann la risposta è che l'integrale non esiste, in quanto la funzione $\mathbb{1}_{\mathbb{Q}}$ semplicemente non è integrabile secondo Riemann. Infatti, data una partizione arbitraria di $[0, 1]$

$$0 = x_0 < x_1 < \dots < x_n = 1$$

siccome ogni intervallo di lunghezza positiva contiene sia numeri razionali che irrazionali, otteniamo che la somma di Riemann superiore e quella inferiore sono rispettivamente date da

$$U(\mathbb{1}_{\mathbb{Q}}, x) := \sum_{k=1}^n \left(\sup_{t \in [x_{k-1}, x_k]} \mathbb{1}_{\mathbb{Q}}(t) \right) (x_k - x_{k-1}) = \sum_{k=1}^n 1 \cdot (x_k - x_{k-1}) = 1$$

$$L(\mathbb{1}_{\mathbb{Q}}, x) := \sum_{k=1}^n \left(\inf_{t \in [x_{k-1}, x_k]} \mathbb{1}_{\mathbb{Q}}(t) \right) (x_k - x_{k-1}) = \sum_{k=1}^n 0 \cdot (x_k - x_{k-1}) = 0,$$

indipendentemente dalla *norma* della partizione x , definita come $\max_k |x_k - x_{k-1}|$.

Per poter integrare funzioni come $\mathbb{1}_{\mathbb{Q}}$ (e molto altro!) è quindi necessario costruire una teoria dell'integrazione che estenda l'integrale di Riemann a funzioni più complicate. Nel seguito diamo una descrizione piuttosto informale di questa costruzione.

3.11.1 La misura di Jordan

L'abilità di integrare una funzione $f : X \rightarrow \mathbb{R}$ è strettamente legata alla capacità di attribuire una *misura* ai sottoinsiemi del dominio X della funzione. Questa misura, che indichiamo con μ , va pensata come una sorta di generalizzazione del concetto di volume. In una generica teoria dell'integrazione che indichiamo con la lettera K , denotiamo con $I_K(f)$ il valore dell'integrale di f e con $\mu_K(A)$ la misura dell'insieme $A \subset X$. Il legame fra l'integrale e la misura è visibile nella seguente relazione che dovrà essere soddisfatta da una classe di sottoinsiemi A di X , affinché la teoria K sia minimamente sensata

$$(3.93) \quad I_K(\mathbb{1}_A) := \int_X \mathbb{1}_A(x) dx = \mu_K(A).$$

Ad esempio nella teoria dell'integrazione di Riemann se $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ è la funzione indicatrice di un rettangolo

$$f(x_1, x_2) = \mathbb{1}_{[a,b] \times [c,d]}(x_1, x_2) = \begin{cases} 1 & \text{se } x_1 \in [a, b] \text{ e } x_2 \in [c, d] \\ 0 & \text{altrimenti.} \end{cases}$$

si ha

$$I_R(\mathbb{1}_{[a,b] \times [c,d]}) = \mu_R([a, b] \times [c, d]) = (b - a) \cdot (d - c).$$

La teoria della misura associata all'integrale di Riemann si chiama in realtà *misura di Jordan*. Per questo motivo, nel seguito useremo il simbolo μ_J invece che μ_R . In precedenza abbiamo detto che la (3.93) deve essere soddisfatta da una classe di sottoinsiemi. Perché non da tutti? Il motivo è che, in generale, è impossibile attribuire una misura a tutti i sottoinsiemi, se lo scopo è quello di costruire una teoria dell'integrazione interessante. I sottoinsiemi ai quali è associata una misura μ_K sono detti *misurabili (secondo K)*. Il motivo per cui l'integrale di Riemann (3.92) non esiste è che l'insieme $\mathbb{Q} \cap [0, 1]$ non è misurabile secondo Jordan. Prima di definire gli insiemi misurabili secondo Jordan introduciamo un concetto astratto.

3.99 Definizione. Dato un insieme non vuoto X , si definisce *algebra* su X una collezione \mathcal{A} di sottoinsiemi di X tale che

- (a) $X \in \mathcal{A}$
- (b) se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$
- (c) se $A_1, \dots, A_n \in \mathcal{A}$ allora $\cup_{i=1}^n A_i \in \mathcal{A}$

Grazie alle leggi di De Morgan, segue dalla definizione precedente che

- (d) $\emptyset \in \mathcal{A}$
- (e) $A_1, \dots, A_n \in \mathcal{A}$ allora $\cap_{i=1}^n A_i \in \mathcal{A}$

Infatti $\emptyset = X^c$, quindi dalla (a) e dalla (b) segue che $\emptyset \in \mathcal{A}$. Inoltre, se $A_1, \dots, A_n \in \mathcal{A}$ abbiamo

$$\left(\cap_{i=1}^n A_i\right)^c = \left(\cup_{i=1}^n A_i^c\right)$$

quindi, usando la (b) (due volte) e la (c) si ottiene che l'intersezione di un numero finito di elementi di \mathcal{A} appartiene ad \mathcal{A} .

3.100 Esempio. (L'algebra degli insiemi semplici). Sia $X = \mathbb{R}^2$ e sia \mathcal{R} l'insieme di tutti i rettangoli finiti o infiniti, aperti o chiusi su ciascun lato, in cui un rettangolo finito è un insieme del tipo

$$(a, b) \times (c, d) \quad \text{oppure} \quad (a, b] \times (c, d) \quad \text{oppure} \quad \dots \quad \text{oppure} \quad [a, b] \times [c, d]$$

mentre in un rettangolo infinito a e/o c possono essere uguali a $-\infty$ e c e/o d possono essere uguali a $+\infty$. È chiaro che la condizione (a) è soddisfatta perché $\mathbb{R}^2 = (-\infty, +\infty) \times (-\infty, +\infty)$ è un rettangolo infinito e quindi appartiene ad \mathcal{R} . La condizione (b) è però violata perché il complemento del rettangolo $[0, 1]^2$ non è un rettangolo. Quindi \mathcal{R} non è un'algebra. Se vogliamo definire un'algebra che contenga tutti i rettangoli serve una classe più ampia di insiemi. Sia \mathcal{S} l'insieme di tutti i sottoinsiemi A di \mathbb{R}^2 che si possono scrivere come

$$A = R_1 \cup \dots \cup R_n \quad R_i \in \mathcal{R}.$$

In altre parole A appartiene ad \mathcal{S} se A si può scrivere come unione finita di rettangoli. È facile convincersi che \mathcal{S} è un'algebra.

La misura di Jordan

Per semplicità consideriamo \mathbb{R}^2 , ma quanto segue può essere esteso ad \mathbb{R}^n . Iniziamo ad attribuire una misura ai rettangoli nell'insieme \mathcal{R} dell'esempio 3.100

$$\mu_J([a, b] \times [c, d]) = \text{area}([a, b] \times [c, d]) = (b - a) \cdot (d - c).$$

Se un rettangolo è aperto su uno o più lati, la sua misura è la stessa del corrispondente rettangolo chiuso. Se uno o più lati del rettangolo hanno lunghezza infinita la sua misura è ∞ . Poi vogliamo definire la misura per gli insiemi dell'algebra \mathcal{S} dell'esempio 3.100, che si possono rappresentare come unione di un numero finito di rettangoli

$$(3.94) \quad A = R_1 \cup \dots \cup R_n \quad R_i \in \mathcal{R}.$$

Gli insiemi di questo tipo vengono chiamati *insiemi semplici*. Poiché i rettangoli R_i possono avere sovrapposizioni, sarebbe "errato" definire la misura di A come la somma delle misure degli R_i . Fortunatamente se A ha la forma (3.94), allora A può anche essere rappresentato come unione di rettangoli a 2 a 2 disgiunti

$$A = R'_1 \sqcup \dots \sqcup R'_m \quad R'_i \in \mathcal{R},$$

in cui il simbolo \sqcup viene utilizzato appunto per indicare che $R'_i \cap R'_j = \emptyset$ se $i \neq j$. Una volta scritto A in questo modo definiamo ovviamente la sua misura come

$$(3.95) \quad \mu_J(A) = \sum_{j=1}^m \mu_J(R'_j).$$

L'idea, a questo punto è di usare (3.95) come "base di partenza" per definire la misura di insiemi più complicati. Dato un insieme limitato B la sua *misura di Jordan esterna* è definita come

$$(3.96) \quad (\mu_J)^*(B) = \inf\{\mu_J(A) : A \text{ è semplice e } A \supset B\}.$$

Analogamente la *misura di Jordan interna* è definita come

$$(3.97) \quad (\mu_J)_*(B) = \sup\{\mu_J(A) : A \text{ è semplice e } A \subset B\}.$$

Finalmente diciamo che un insieme limitato B è *misurabile secondo Jordan* se la sua misura esterna coincide con quella interna. In quel caso la sua misura è data da

$$\mu_J(B) = (\mu_J)^*(B) = (\mu_J)_*(B).$$

Gli insiemi misurabili secondo Jordan costituiscono una collezione molto più ricca rispetto a quella degli insiemi semplici. Sono misurabili tutti quegli insiemi il cui bordo non è troppo "ciccione".²⁰ Non tutti gli insiemi aperti tuttavia sono misurabili.

Tornando all'insieme $B = \mathbb{Q} \cap [0, 1]$, a cui abbiamo accennato in precedenza, si vede facilmente che questo non è misurabile secondo Jordan. Infatti, siccome siamo in una dimensione, i "rettangoli" ora sono segmenti e gli insiemi semplici sono unioni di un numero finito di segmenti. Dalla definizione (3.96) segue che la misura di Jordan esterna di B è uguale a 1, perché una collezione finita di segmenti che contiene B deve necessariamente contenere tutto l'intervallo $[0, 1]$. Ma la misura di Jordan interna è ovviamente 0, quindi B non è misurabile secondo Jordan. Per superare questa difficoltà l'idea chiave, che porta al concetto di misura di Lebesgue, è la seguente: per definire la misura esterna, invece di ricoprire l'insieme B con un numero finito di segmenti, usiamo insiemi che sono *unione di un'infinità numerabile di segmenti!* Vediamo come mai questa cosa funziona. Sappiamo che i razionali sono numerabili, quindi, a maggior ragione, l'insieme $B = \mathbb{Q} \cap [0, 1]$ è numerabile. Possiamo scrivere dunque

$$B = \{b_1, b_2, b_3, \dots\}.$$

²⁰Possiamo anticipare che sono misurabili tutti gli insiemi il cui bordo ha misura di Lebesgue nulla.

Sia $\varepsilon > 0$ e consideriamo il seguente insieme

$$A = \bigcup_{n=1}^{\infty} \left[b_n - \frac{\varepsilon}{2^n}, b_n + \frac{\varepsilon}{2^n} \right].$$

La misura di A (non userò il pedice J perché non è più la misura di Jordan) è minore o uguale alla somma delle lunghezze di tutti questi segmenti

$$\mu(A) \leq \sum_{n=1}^{\infty} 2 \frac{\varepsilon}{2^n} = 2\varepsilon.$$

D'altra parte, poiché $B \subset A$, si ha $\mu^*(B) \leq \mu(A) \leq 2\varepsilon$, che, essendo ε arbitrario, implica $\mu^*(B) = 0$. Quindi, in questo approccio, B è misurabile e la sua misura vale 0. Nella prossima sezione diamo una descrizione più sistematica di questa costruzione.

3.11.2 σ -algebre e la misura di Lebesgue

3.101 Definizione. Dato un insieme non vuoto X , si definisce σ -algebra su X una collezione \mathcal{F} di sottoinsiemi di X tale che valgono le proprietà (a), (b) della definizione 3.99 e inoltre valga

(c') se A_1, A_2, A_3, \dots è una collezione infinita numerabile di elementi di \mathcal{F} allora $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

La coppia (X, \mathcal{F}) si chiama *spazio misurabile*.

La differenza, in apparenza minuscola, che intercorre fra la proprietà (c) della 3.99 e la (c') è in realtà enorme. Nell'esempio 3.100 abbiamo visto come, partendo dei rettangoli si arriva a costruire l'algebra \mathcal{S} degli insiemi semplici, che sono unione di un numero finito di rettangoli. Se invece di un'algebra siamo interessati a costruire una σ -algebra che contenga l'insieme \mathcal{R} di tutti i rettangoli, quello che si ottiene è una collezione di insiemi molto più ricca di \mathcal{S} . Se \mathcal{F} è una qualsiasi σ -algebra che contiene \mathcal{R} , allora \mathcal{F} deve necessariamente contenere tutti gli insiemi aperti e chiusi di \mathbb{R}^2 di forma qualsiasi (ma non solo!). Infatti sia A un insieme aperto arbitrario di \mathbb{R}^2 . Voglio far vedere che A si può scrivere come

$$(3.98) \quad A = \bigcup_{i=1}^{\infty} R_i$$

in cui ogni R_i è un rettangolo, per cui, grazie a (c'), A deve necessariamente appartenere ad \mathcal{F} . Per dimostrare la (3.98) sia \mathcal{D} la collezione di tutti i rettangoli R tali che $R \subset A$ e tali che le coordinate dei quattro vertici di R sono razionali. È chiaro allora che $\bigcup_{R \in \mathcal{D}} R \subset A$ in quanto ogni $R \in \mathcal{D}$ è contenuto in A . Voglio far vedere che $A \subset \bigcup_{R \in \mathcal{D}} R$. A questo proposito, sia $x \in A$. Dato che A è aperto esiste $\varepsilon > 0$ tale che $B_\varepsilon(x) \subset A$. È chiaro che esisterà dunque un quadratino R con vertici razionali contenuto nel cerchio $B_\varepsilon(x)$, quindi contenuto in A . Di conseguenza x è contenuto in un rettangolo della collezione \mathcal{D} , e poiché questo vale per ogni $x \in A$ si ottiene $A \subset \bigcup_{R \in \mathcal{D}} R$. Dunque deve valere l'uguaglianza $A = \bigcup_{R \in \mathcal{D}} R$. Ma l'insieme \mathcal{D} è numerabile, per cui, una volta numerato, possiamo scrivere A nella forma (3.98). Abbiamo quindi dimostrato che \mathcal{F} contiene tutti gli insiemi aperti, e, grazie alla proprietà (b) tutti gli insiemi chiusi. \mathcal{F} in realtà contiene molti altri insiemi. La *più piccola* σ -algebra \mathcal{F} che contiene \mathcal{R} è chiamata σ -algebra di Borel e verrà indicata con \mathcal{B} . A questo punto definiamo il concetto di misura su una σ -algebra.

3.102 Definizione. Dato lo spazio misurabile (X, \mathcal{F}) , si definisce *misura*²¹ su (X, \mathcal{F}) un'applicazione $\mu : \mathcal{F} \rightarrow [0, \infty]$ tale che

- (a) se $(A_i)_{i \in I}$ è una collezione finita o numerabile di elementi di \mathcal{F} tali che $A_n \cap A_k = \emptyset$ se $n \neq k$, allora

$$\mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i);$$

- (b) X può essere scritto come $X = \bigcup_{i=1}^{\infty} Y_i$ in cui ogni Y_i ha misura finita.

La proprietà (a) è detta *additività numerabile* o σ -*additività*.

3.103 Definizione. Uno *spazio di misura* è una tripla (X, \mathcal{F}, μ) in cui (X, \mathcal{F}) è uno spazio misurabile e μ è una misura su (X, \mathcal{F}) .

Osservo che l'insieme vuoto ha necessariamente misura nulla in quanto, grazie alla (a), si ha

$$\mu(\emptyset) = \mu(\emptyset \cup \emptyset) = \mu(\emptyset) + \mu(\emptyset).$$

La misura di Lebesgue

Continuiamo a considerare \mathbb{R}^2 per semplicità. Come dicevamo nella sezione 3.11.1, l'idea di base è quella di definire la misura esterna usando unioni numerabili (invece che finite) di rettangoli. La *misura di Lebesgue esterna* di un insieme $B \subset \mathbb{R}^2$ è dunque definita come

$$\mu_L^*(B) := \inf \left\{ \sum_{i=1}^{\infty} \text{area}(R_i) : R_i \in \mathcal{R}, \bigcup_{i=1}^{\infty} R_i \supset B \right\}.$$

A questo punto un insieme B è detto *misurabile secondo Lebesgue* se per ogni $C \subset \mathbb{R}^2$ si ha

$$\mu_L^*(C) = \mu_L^*(C \cap B) + \mu_L^*(C \setminus B).$$

Se B è misurabile la sua misura di Lebesgue è definita come $\mu_L(B) = \mu_L^*(B)$. Si dimostra che questa definizione di misurabilità è equivalente alla condizione che B possa essere approssimato dall'esterno con un insieme aperto e dall'interno con un insieme chiuso. Più precisamente B è misurabile secondo Lebesgue se e solo se per ogni $\varepsilon > 0$ esistono A aperto e C chiuso tali che

$$C \subset B \subset A \quad \text{e} \quad \mu_L^*(A \setminus C) < \varepsilon.$$

Si dimostra che l'insieme degli insiemi misurabili secondo Lebesgue è una σ -algebra che denotiamo con \mathcal{L} . In particolare tutti i rettangoli appartengono a \mathcal{L} e vale $\mu_L(R) = \text{area}(R)$ per ogni $R \in \mathcal{R}$. Poiché \mathcal{L} contiene tutti i rettangoli e siccome \mathcal{B} , la σ -algebra di Borel, è per definizione la più piccola σ -algebra che contiene \mathcal{R} , si ottiene che $\mathcal{B} \subset \mathcal{L}$. Quindi \mathcal{L} è più grande di \mathcal{B} , ma la differenza è in un certo senso inessenziale. Quello che accade è la cosa seguente: esistono insiemi in \mathcal{B} che hanno misura nulla, ma che ammettono sottoinsiemi che *non* sono elementi di \mathcal{B} . Se aggiungiamo a \mathcal{B} questi sottoinsiemi di insiemi di misura nulla, otteniamo esattamente \mathcal{L} .

²¹ci restringiamo al caso di misure σ -finite

3.104. Insiemi di misura nulla. Ogni insieme che contiene un solo punto $\{x\}$ ha ovviamente misura di Lebesgue uguale a 0, in quanto può essere considerato un rettangolo di area nulla. Grazie alla σ -additività della misura, ogni insieme numerabile ha misura nulla, perché

$$\mu_L(\{a_1, a_2, \dots\}) = \mu_L(\cup_{i=1}^{\infty} \{a_i\}) = \sum_{i=1}^{\infty} \mu_L(a_i) = 0.$$

Ad esempio \mathbb{Q}^2 , vale a dire l'insieme di tutti i punti di \mathbb{R}^2 con coordinate razionali ha misura nulla. Non tutti gli insiemi di misura nulla sono numerabili. Ad esempio il segmento

$$[0, 1] \times \{0\} = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 = 0\}$$

ha misura nulla perché è un rettangolo di base 1 e altezza 0.²² È facile convincersi che una qualsiasi retta in \mathbb{R}^2 ha misura nulla, e di conseguenza, anche una collezione numerabile di rette.

3.105. Insiemi non misurabili. Esistono insiemi che non sono misurabili secondo Lebesgue, ma la loro costruzione richiede l'assioma della scelta e risulta dunque non troppo esplicita. C'è una costruzione, in particolare, che vale la pena menzionare perché si tratta di un risultato a dir poco controintuitivo. È il famoso *paradosso di Banach-Tarski*. Dunque, reggetevi forte e, come si suol dire, don't try this at home:

è possibile suddividere una palla in \mathbb{R}^3 in 5 pezzi (pezzo = sottoinsieme), smontare i 5 pezzi e rimontarli diversamente, in modo da ottenere 2 palle identiche alla palla originaria.

Sì, avete capito bene:

- (1) i pezzi sono rigidi, non si possono deformare in alcun modo;
- (2) smontare e rimontare vuol dire che possiamo applicare traslazioni e rotazioni a ciascuno dei 5 pezzi;
- (3) ciascuna delle 2 palle ottenute è identica alla palla originale, ha lo stesso raggio.

Ora, se questi pezzi fossero misurabili, nella loro disposizione iniziale la somma delle loro misure (volumi) sarebbe uguale al volume di una palla di raggio R . Ma nella disposizione finale la somma delle loro misure deve essere uguale al volume di 2 palle di raggio R , e quindi si arriverebbe a concludere che $1 = 2$, che ha tutta l'aria di una contraddizione. Dunque questi pezzi non sono misurabili.

3.11.3 L'integrale di Lebesgue e gli spazi L_p

L'integrale di Lebesgue

Una volta definita la misura di Lebesgue, possiamo costruire il corrispondente integrale di Lebesgue. Iniziamo a definire l'integrale in uno spazio misurabile arbitrario (X, \mathcal{F}) sul quale è definita una misura (σ -additiva) μ . In seguito ritorneremo al caso particolare della misura di Lebesgue in \mathbb{R}^n .

²²Ovviamente rispetto alla misura di Lebesgue unidimensionale il segmento $[0, 1]$, come sottoinsieme di \mathbb{R} , ha misura 1. Costruire insiemi non numerabili di misura nulla in una dimensione è un po' più complicato.

3.106 Definizione. Dato uno spazio misurabile (X, \mathcal{F}) , la funzione $f : X \rightarrow \mathbb{R}$ si dice *misurabile* se per ogni $t \in \mathbb{R}$ si ha che l'insieme $\{x \in X : f(x) < t\}$ è misurabile, cioè appartiene a \mathcal{F} .

Se A è un insieme misurabile, allora la la funzione indicatrice dell'insieme A

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

è misurabile, poiché

$$\{x \in X : \mathbb{1}_A(x) < t\} = \begin{cases} A & \text{se } t > 1 \\ \emptyset & \text{se } t \leq 1. \end{cases}$$

Si inizia col definire l'integrale per le *funzioni misurabili semplici* vale a dire per le funzioni $f : X \rightarrow \mathbb{R}$ che si possono scrivere come

$$(3.99) \quad f = c_1 \mathbb{1}_{A_1} + \cdots + c_n \mathbb{1}_{A_n}$$

in cui c_i sono numeri reali e gli A_i sono insiemi misurabili. Per una funzione semplice f data dalla (3.99) l'integrale rispetto alla misura μ si definisce come

$$\int_X f(x) \mu(dx) = \sum_{i=1}^n c_i \mu(A_i).$$

Il passo successivo è definire l'integrale per funzioni a valori non negativi. Se $f : X \rightarrow [0, \infty)$ è misurabile, definiamo

$$(3.100) \quad \int_X f(x) \mu(dx) = \sup \left\{ \int_X g(x) \mu(dx) : g \text{ è misurabile semplice e } 0 \leq g \leq f \right\}.$$

Se f assume valori sia positivi che negativi, possiamo sempre scrivere

$$f = f^+ - f^- \quad \text{in cui } f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\}.$$

Poiché sia f^+ che f^- sono a valori non negativi, possiamo definire, riconducendoci alla (3.100),

$$\int_X f(x) \mu(dx) = \int_X f^+(x) \mu(dx) - \int_X f^-(x) \mu(dx).$$

Possiamo ulteriormente generalizzare, definendo l'integrale per una funzione a valori complessi $f : X \rightarrow \mathbb{C}$ come

$$\int_X f(x) \mu(dx) = \int_X \operatorname{Re} f(x) \mu(dx) + \int_X \operatorname{Im} f(x) \mu(dx).$$

Finora l'unica restrizione che abbiamo imposto sulla funzione integranda è quello di essere misurabile rispetto alla σ -algebra \mathcal{F} , ma nulla ci garantisce che il valore dell'integrale sia finito. Definiamo quindi

$$\mathcal{L}_1(X, \mathcal{F}, \mu) = \left\{ f : X \rightarrow \mathbb{C} : f \text{ è misurabile e } \int_X |f(x)| \mu(dx) < \infty \right\},$$

e, più in generale, per ogni $p \in [1, \infty)$

$$\mathcal{L}_p(X, \mathcal{F}, \mu) = \left\{ f : X \rightarrow \mathbb{C} : f \text{ è misurabile e } \int_X |f(x)|^p \mu(dx) < \infty \right\}.$$

Per semplicità si scrive spesso $\mathcal{L}_p(\mu)$ al posto di $\mathcal{L}_p(X, \mathcal{F}, \mu)$. Possiamo ora considerare l'integrale rispetto alla misura μ come un'applicazione, che denotiamo con I_μ , dallo spazio $\mathcal{L}_1(\mu)$ a valori complessi, definita come

$$I_\mu(f) = \int_X f(x) \mu(dx) \quad f \in \mathcal{L}_1(\mu).$$

Si verifica che $\mathcal{L}_1(\mu)$ è uno spazio vettoriale (lo sono tutti gli $\mathcal{L}_p(\mu)$) e che questa applicazione è *lineare*, come l'integrale di Riemann, vale a dire

$$I_\mu(\alpha f + \beta g) = \alpha I_\mu(f) + \beta I_\mu(g) \quad f, g \in \mathcal{L}_1(\mu), \alpha, \beta \in \mathbb{C}.$$

Non è un integrale diverso, ma è un'estensione! Nel caso in cui $X = \mathbb{R}^n$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (oppure $f : \mathbb{R}^n \rightarrow \mathbb{C}$) possiamo chiederci che relazione c'è fra l'integrale di Riemann e quello di Lebesgue della funzione f . La risposta è semplice: se f è integrabile secondo Riemann, allora f è anche integrabile secondo Lebesgue e i due integrali hanno lo stesso valore. In questo caso quindi, l'integrale di Lebesgue costituisce un'estensione dell'integrale di Riemann. Per questo motivo l'integrale rispetto alla misura di Lebesgue di funzioni definite su \mathbb{R}^n viene spesso denotato con lo stesso simbolo dx o $d^n x$ che si usa per l'integrale di Riemann.

Classi di equivalenza e spazi $L_p(\mu)$

Un piccola fastidiosa seccatura. Tornerebbe comodo poter considerare gli insiemi $\mathcal{L}_p(\mu)$ definiti precedentemente, come spazi vettoriali normati rispetto alla norma

$$(3.101) \quad \|f\|_p := \left(\int_X |f(x)|^p \mu(dx) \right)^{1/p},$$

analogamente a quanto abbiamo fatto nel caso degli spazi di funzioni continue $C_p(\mathbb{R}^n)$. C'è però una piccola difficoltà da appianare: l'applicazione (3.101), in generale, non è una norma su $\mathcal{L}_p(\mu)$. Il motivo è il seguente: abbiamo visto nel caso della misura di Lebesgue che esistono insiemi non vuoti di misura nulla. Questa non è una prerogativa della misura di Lebesgue. Qualsiasi spazio di misura può contenere questi insiemi. Ma se $A \neq \emptyset$ ha misura nulla, allora

$$\|\mathbb{1}_A\|_p^p = \int_X |\mathbb{1}_A(x)|^p \mu(dx) = \int_X \mathbb{1}_A(x) \mu(dx) = \mu(A) = 0.$$

Poiché $\mathbb{1}_A$ non è la funzione identicamente nulla, questo implica che $\|\cdot\|_p$ non è una norma.

Appianamento della difficoltà. Per superare questo impedimento alla definizione di una norma sugli spazi $\mathcal{L}_p(\mu)$ si introduce la seguente relazione di equivalenza: due funzioni f e g appartenenti a $\mathcal{L}_p(\mu)$ si dicono equivalenti se differiscono su un insieme di misura nulla

$$f \sim g \iff \mu\{x \in X : f(x) \neq g(x)\} = 0.$$

In generale un'affermazione $A(x)$ che dipende dal punto $x \in X$ si dice che è vera *quasi ovunque* e si abbrevia *q.o.* se l'insieme degli x per i quali $A(x)$ è falsa ha misura nulla. Quindi la condizione $f \sim g$ può essere anche scritta come

$$f(x) = g(x) \quad \text{q.o.}$$

La relazione di equivalenza $f \sim g$ è compatibile con la struttura di spazio vettoriale, vale a dire se $f \sim f'$ e $g \sim g'$, allora $f + g \sim f' + g'$ e, se $\alpha \in \mathbb{C}$, $\alpha f \sim \alpha f'$. Questo ci permette di definire lo spazio $L_p(\mu)$ come lo spazio vettoriale *quoziente* di $\mathcal{L}_p(\mu)$ rispetto a questa relazione di equivalenza. in simboli

$$L_p(\mu) = \mathcal{L}_p(\mu) / \sim .$$

Gli elementi di $L_p(\mu)$ quindi non sono funzioni ma classi di equivalenza di funzioni. Indichiamo con $[f]$ l'insieme di tutte le funzioni di $\mathcal{L}_p(\mu)$ equivalenti ad f .

3.107 Notazione. Nel caso particolare in cui $X = \mathbb{R}^n$, $\mathcal{F} = \mathcal{L}$ è la σ -algebra di Lebesgue su \mathbb{R}^n e $\mu = \mu_L$ è la corrispondente misura di Lebesgue, si definisce

$$L_p(\mathbb{R}^n) = L_p(\mu_L) = L_p(\mathbb{R}^n, \mathcal{L}, \mu_L) .$$

Prendiamo in esame ora, per fare un esempio, il caso della misura di Lebesgue sulla retta, dunque lo spazio $L_p(\mathbb{R})$. Consideriamo 3 funzioni

$$(3.102) \quad f(x) = e^{-x^2} \quad g(x) = \begin{cases} e^{-x^2} & x \in \mathbb{R} \setminus \mathbb{Q} \\ x & x \in \mathbb{Q} \end{cases} \quad h(x) = \begin{cases} e^{-x^2} & x \in \mathbb{R} \setminus \mathbb{Q} \\ \sin x & x \in \mathbb{Q} . \end{cases}$$

Poiché queste funzioni differiscono solo su \mathbb{Q} e $\mu_L(\mathbb{Q}) = 0$, esse sono equivalenti, quindi rappresentano lo stesso elemento in $L_p(\mathbb{R})$

$$(3.103) \quad f \sim g \sim h \quad \iff \quad [f] = [g] = [h] .$$

Nella notazione usuale tuttavia il fatto che gli elementi di $L_p(\mu)$ siano classi di equivalenza di funzioni viene sottinteso e si scrive $f \in L_p(\mu)$ per indicare $[f] \in L_p(\mu)$. Molte proprietà delle funzioni vengono impropriamente attribuite alla corrispondente classe di equivalenza. Tornando al caso della misura di Lebesgue sulla retta, affermare, ad esempio, che $f \in L_p(\mathbb{R})$ è continua significa che nella classe $[f]$ esiste un rappresentante continuo. La funzione h che appare nella (3.102), come elemento di $L_p(\mathbb{R})$, è continua perché nella stessa classe di equivalenza c'è f che è continua. In questo modo è possibile identificare lo spazio $C_p(\mathbb{R})$ come un sottospazio di $L_p(\mathbb{R})$.

Veniamo ora ad enunciare due risultati fondamentali riguardanti gli spazi L_p . Il primo riguarda spazi di misura arbitrari.

3.108 Teorema. *Sia (X, \mathcal{F}, μ) uno spazio di misura. Per ogni $p \in [1, \infty)$ lo spazio vettoriale normato $L_p(X, \mathcal{F}, \mu)$ è completo, quindi è uno spazio di Banach.*

Il secondo risultato riguarda il caso in cui X è uno spazio metrico (nel quale ha senso parlare di funzioni continue) e μ è una misura che soddisfa una certa proprietà di regolarità. Per semplicità enunciamo il teorema nel caso particolare della misura di Lebesgue su \mathbb{R}^n in cui questa proprietà si dimostra essere soddisfatta.

3.109 Teorema. *Per ogni $p \in [1, \infty)$, $C_p(\mathbb{R}^n)$ è denso in $L_p(\mathbb{R}^n)$, quindi $L_p(\mathbb{R}^n)$ è il completamento di $(C_p(\mathbb{R}^n), \|\cdot\|_p)$.*

Dato che $C_p(\mathbb{R}^n)$ è denso in $L_p(\mathbb{R}^n)$, per ogni $f \in L_p(\mathbb{R}^n)$ e per ogni $\varepsilon > 0$, esiste $g \in C_p(\mathbb{R}^n)$ tale che $\|f - g\|_p < \varepsilon$, vale a dire ogni elemento di $L_p(\mathbb{R}^n)$ può essere approssimato, rispetto alla norma $\|\cdot\|_p$, con precisione arbitraria, da un elemento di $C_p(\mathbb{R}^n)$. Per questo motivo, anche se gli spazi L_p contengono funzioni molto "strane", sarà sempre possibile trovare una base costituita da funzioni continue (addirittura C^∞), come vedremo nel caso di L_2 .

4. Spazi di Hilbert

F: Butch, whose motorcycle is this?

B: It's a chopper.

F: Whose chopper is this?

B: Zed's.

F: Who's Zed?

B: Zed's dead, baby, Zed's dead.

4.1 Definizioni

4.1 Definizione. Sia V uno spazio vettoriale su \mathbb{R} . Si definisce *prodotto scalare in V* un'applicazione

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

tale che, per ogni $u, v, z \in V$ e per ogni $c \in \mathbb{R}$ si ha:

- (1) $\langle v, v \rangle \geq 0$
- (2) $\langle v, v \rangle = 0$ se e solo se $v = 0$
- (3) $\langle u, v \rangle = \langle v, u \rangle$
- (4) $\langle u + v, z \rangle = \langle u, z \rangle + \langle v, z \rangle$
- (5) $\langle cu, v \rangle = c \langle u, v \rangle$

Se V è uno spazio vettoriale *complesso* questa definizione deve essere modificata.

4.2 Definizione. Sia V uno spazio vettoriale su \mathbb{C} . Si definisce *prodotto scalare in V* un'applicazione

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$$

tale che, per ogni $u, v, z \in V$ e per ogni $c \in \mathbb{C}$ si ha:

- (1) $\langle v, v \rangle \geq 0$
- (2) $\langle v, v \rangle = 0$ se e solo se $v = 0$
- (3) $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- (4) $\langle u + v, z \rangle = \langle u, z \rangle + \langle v, z \rangle$
- (5) $\langle cu, v \rangle = c \langle u, v \rangle$

4.3 Definizione. Uno spazio vettoriale reale (o complesso) V munito di un prodotto scalare $\langle \cdot, \cdot \rangle$ è detto *spazio euclideo* reale (o complesso). Uno spazio euclideo completo si dice *spazio di Hilbert*.¹

Le seguenti proprietà discendono in modo evidente dalle definizioni appena date.

4.4 Proposizione. *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo reale (complesso). Allora per ogni $u, v, z \in V$ e per ogni $c \in \mathbb{R}$ ($c \in \mathbb{C}$) si ha:*

¹per qualcuno spazio di Hilbert è necessariamente infinito dimensionale

- (6) $\langle z, u + v \rangle = \langle z, u \rangle + \langle z, v \rangle$
 (7a) $\langle u, cv \rangle = c \langle u, v \rangle$ (nel caso reale)
 (7b) $\langle u, cv \rangle = \bar{c} \langle u, v \rangle$ (nel caso complesso).

Le identità (4), (5) e (7b) ci dicono che, nel caso complesso, il prodotto scalare, non è una forma bilineare, ma *sesquilineare*, vale a dire lineare nel primo argomento e *coniugata lineare* nel secondo argomento.

4.5 Attenzione! In molti libri di fisica si trova la convenzione opposta, coniugata lineare nel primo argomento e lineare nel secondo, vale a dire

$$\langle cu, v \rangle = \bar{c} \langle u, v \rangle \quad \langle u, cv \rangle = c \langle u, v \rangle.$$

Basta saperlo...

4.6 Proposizione. (*Disuguaglianza di Cauchy-Schwarz*). Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo reale (complesso). Allora

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \langle v, v \rangle \quad \forall u, v \in V$$

Dimostrazione. Faccio la dimostrazione nel caso complesso che è quello più... complesso. Sia $t \in \mathbb{C}$ e

$$(4.1) \quad p(t) := \langle tu + v, tu + v \rangle.$$

Usando le proprietà del prodotto scalare ottengo

$$p(t) = |t|^2 \langle u, u \rangle + t \langle u, v \rangle + \bar{t} \langle v, u \rangle + \langle v, v \rangle.$$

Scrivo il numero complesso $\langle u, v \rangle$ in forma polare come $\langle u, v \rangle = \rho e^{i\vartheta}$. Poi scelgo il numero complesso t della forma $se^{-i\vartheta}$ con s reale. La precedente uguaglianza diventa

$$p(se^{-i\vartheta}) = s^2 \langle u, u \rangle + 2s\rho + \langle v, v \rangle.$$

Ma, per come è stata definita nella (4.1), la quantità $p(t)$ è non negativa, qualunque sia il numero complesso t . In particolare dovrà essere non negativa per tutti i t della forma $se^{-i\vartheta}$ in cui s può essere scelto in modo arbitrario. Dunque ottengo

$$s^2 \langle u, u \rangle + 2s\rho + \langle v, v \rangle \geq 0 \quad \forall s \in \mathbb{R},$$

che implica

$$\rho^2 \leq \langle u, u \rangle \langle v, v \rangle.$$

Ma ρ è esattamente il modulo di $\langle u, v \rangle$, per cui la Proposizione è dimostrata. \square

Uno dei vantaggi di avere un prodotto scalare è che, tramite il prodotto scalare, si può facilissimamente definire una norma (il contrario in generale non si può fare, come spiegato nella sezione 4.1.2).

4.7 Proposizione. Sia V uno spazio euclideo con prodotto scalare $\langle \cdot, \cdot \rangle$. Allora la funzione $\|\cdot\| : V \rightarrow \mathbb{R}$ definita come

$$(4.2) \quad \|u\| := \sqrt{\langle u, u \rangle} \quad u \in V$$

è una norma su V , quindi ogni spazio euclideo è anche uno spazio vettoriale normato.

Dimostrazione. Molto semplice, lasciata al lettore.

4.8 Proposizione. *L'addizione, la moltiplicazione per uno scalare e il prodotto scalare sono continui in uno spazio euclideo, vale a dire se $u_n \rightarrow u$ e $v_n \rightarrow v$ allora:*

- (1) $u_n + v_n \rightarrow u + v$
- (2) se $c \in \mathbb{C}$ allora $cu_n \rightarrow cu$
- (3) $\langle u_n, v_n \rangle \rightarrow \langle u, v \rangle$

Dimostrazione. Dimostro la continuità del prodotto scalare. Le altre affermazioni si dimostrano analogamente. Grazie alla disuguaglianza di Cauchy-Schwarz ottengo:

$$\begin{aligned} |\langle u, v \rangle - \langle u_n, v_n \rangle| &= |\langle u, v \rangle - \langle u, v_n \rangle + \langle u, v_n \rangle - \langle u_n, v_n \rangle| \\ &\leq |\langle u, v \rangle - \langle u, v_n \rangle| + |\langle u, v_n \rangle - \langle u_n, v_n \rangle| \\ &= |\langle u, v - v_n \rangle| + |\langle u - u_n, v_n \rangle| \\ &\leq \|u\| \|v - v_n\| + \|u - u_n\| \|v_n\|. \end{aligned}$$

Poiché la successione (v_n) è convergente, essa è anche limitata in norma,² quindi esiste $M > 0$ tale che $\|v_n\| \leq M$ per ogni n . La precedente disuguaglianza può quindi essere scritta come

$$|\langle u, v \rangle - \langle u_n, v_n \rangle| \leq \max\{\|u\|, M\} [\|v - v_n\| + \|u - u_n\|].$$

A questo punto non resta che prendere il limite $n \rightarrow \infty$ di ambo i membri. □

4.1.1 Esempi di spazi euclidei

- (1) Definisco, per $x, y \in \ell_2(\mathbb{R})$ (oppure $x, y \in \ell_2(\mathbb{C})$)

$$\langle x, y \rangle := \sum_{i=1}^{\infty} x_i y_i \quad (\text{caso reale}) \quad \langle x, y \rangle := \sum_{i=1}^{\infty} x_i \bar{y}_i \quad (\text{caso complesso})$$

Allora $\langle \cdot, \cdot \rangle$ è un prodotto scalare in ℓ_2 e la norma associata $\|x\| := \sqrt{\langle x, x \rangle}$ è proprio la norma $\|\cdot\|_2$ che abbiamo precedentemente introdotto in questo spazio.

- (2) (Spazi di successioni pesati). Sia $\mu = (\mu_1, \mu_2, \dots)$ una successione di numeri reali strettamente positivi, $\mu_i > 0$ e si definisca

$$\ell_2(\mathbb{C}, \mu) := \left\{ x \in \mathbb{C}^\infty : \sum_{i=1}^{\infty} \mu_i |x_i|^2 < \infty \right\}$$

Allora $\ell_2(\mu)$ è uno spazio vettoriale. Infatti se $x \in \ell_2(\mu)$ e $c \in \mathbb{R}$ allora cx è ancora un elemento di $\ell_2(\mu)$ poiché

$$\sum_{i=1}^{\infty} \mu_i |cx_i|^2 = |c|^2 \sum_{i=1}^{\infty} \mu_i |x_i|^2 < \infty$$

Inoltre se $x, y \in \ell_2(\mu)$ allora $x + y \in \ell_2(\mu)$. Infatti sappiamo che $2|ab| \leq (a^2 + b^2)$, quindi

$$\begin{aligned} \sum_{i=1}^{\infty} \mu_i |x_i + y_i|^2 &\leq \sum_{i=1}^{\infty} \mu_i (|x_i|^2 + |y_i|^2 + 2|x_i y_i|) \leq 2 \sum_{i=1}^{\infty} \mu_i (|x_i|^2 + |y_i|^2) \\ &= 2 \sum_{i=1}^{\infty} \mu_i |x_i|^2 + 2 \sum_{i=1}^{\infty} \mu_i |y_i|^2 < \infty. \end{aligned}$$

²la dimostrazione è identica al caso dei reali

Definiamo un prodotto scalare in $\ell_2(\mu)$ come

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \mu_i x_i \bar{y}_i.$$

- (3) Consideriamo lo spazio $C_2(\mathbb{R})$ delle funzioni reali a valori reali o complessi a quadrato sommabile. Date $f, g \in C_2(\mathbb{R})$ definisco

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x)g(x) dx \quad (\text{caso reale}) \quad \langle f, g \rangle := \int_{\mathbb{R}} f(x)\overline{g(x)} dx \quad (\text{caso complesso})$$

Si verifica facilmente che questo è un prodotto scalare e la norma associata è la norma $\|\cdot\|_2$.

- (4) (Spazi di funzioni pesati). Sia $p : \mathbb{R} \rightarrow [0, \infty)$ una funzione continua. Definisco:

$$C_2(\mathbb{R}, p(x)dx) = \left\{ f \in C(\mathbb{R}) : \int_{\mathbb{R}} |f(x)|^2 p(x)dx < \infty \right\}.$$

Questo è uno spazio vettoriale sul quale posso definire il prodotto scalare

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x)\overline{g(x)}p(x)dx.$$

- (5) Gli spazi $C_2(\mathbb{R})$ e $C_2(\mathbb{R}, p(x)dx)$ definiti precedentemente sono spazi euclidei, ma non sono spazi di Hilbert perché non sono completi. Accanto a questi possiamo definire i corrispondenti spazi L_2 che costituiscono il loro completamento, descritto nella sezione 3.11. Quindi $L_2(\mathbb{R})$ e $L_2(\mathbb{R}, p(x)dx)$ sono spazi di Hilbert. Il prodotto scalare in questi spazi è lo stesso di quello introdotto negli spazi C_2 , ma in generale, se l'integrando non è continuo, l'integrale va inteso come integrale di Lebesgue.

4.1.2 La regola del parallelogramma

4.9 Domanda. La Proposizione 4.7 ci dice che da un prodotto scalare definito su uno spazio vettoriale V possiamo sempre ricavare una norma mediante la definizione

$$(4.3) \quad \|v\| = \sqrt{\langle v, v \rangle} \quad v \in V,$$

quindi ogni spazio euclideo è anche uno spazio vettoriale normato. È possibile fare la strada inversa? Sia $(V, \|\cdot\|)$ uno SVN, è possibile definire un prodotto scalare su V in modo tale che valga la (4.3)? Abbiamo visto dei casi in cui questo è possibile. Ad esempio, dato lo SVN $(\ell_2, \|\cdot\|_2)$, se definiamo

$$\langle x, y \rangle = \sum_{k=1}^{\infty} x_k \bar{y}_k \quad x, y \in \ell_2.$$

allora la (4.3) è soddisfatta. Bene, si può fare qualcosa di analogo per gli altri ℓ_p , quando $p \neq 2$? La risposta è no. L'esistenza di un prodotto scalare compatibile con una norma data è legata ad una proprietà che deve soddisfare la norma, detta *regola del parallelogramma*. Negli spazi $(\ell_p, \|\cdot\|_p)$ questa regola è soddisfatta se e solo se $p = 2$.

Prima di enunciare la regola del parallelogramma, vediamo come ricostruire il prodotto scalare a partire dalla norma, in quei casi in cui questo sarà possibile.

4.10 Proposizione. (*Identità di polarizzazione*). Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $\| \cdot \|$ la norma associata al prodotto scalare (tramite la (4.3)). Allora, per ogni $u, v \in V$,

$$\begin{aligned}\langle u, v \rangle &= \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2 + i(\|u+iv\|^2 - \|u-iv\|^2)) && (V \text{ complesso}) \\ \langle u, v \rangle &= \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2) && (V \text{ reale}).\end{aligned}$$

Dimostrazione. È una semplice verifica. \square

4.11 Proposizione. (*Regola del parallelogramma*). Sia $(V, \| \cdot \|)$ uno SVN. Una condizione necessaria e sufficiente per l'esistenza di un prodotto scalare $\langle \cdot, \cdot \rangle$ in V tale che

$$(4.4) \quad \|u\| := \sqrt{\langle u, u \rangle} \quad u \in V$$

è la seguente identità

$$(4.5) \quad \|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

❶ *Dimostrazione della necessità.* Supponiamo che la norma $\| \cdot \|$ sia associata ad un prodotto scalare tramite la (4.4). Allora otteniamo

$$(4.6) \quad \begin{aligned}\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v \rangle + \langle u-v, u-v \rangle \\ &= \|u\|^2 + \langle u, v \rangle + \langle v, u \rangle + \|v\|^2 + \|u\|^2 - \langle u, v \rangle - \langle v, u \rangle + \|v\|^2 \\ &= 2(\|u\|^2 + \|v\|^2)\end{aligned}$$

❷ La dimostrazione della sufficienza è più complicata e viene omessa. \square

4.1.3 Problemi

4.12 Problema. Dire se la seguente espressione è un prodotto scalare in $\ell_2(\mathbb{R})$ o meno. Dimostrare ciò che si afferma.

$$\langle x, y \rangle = \sum_{i=1}^{\infty} (x_i y_{i+1} + x_{i+1} y_i)$$

4.13 Problema. Dire se la seguente espressione è un prodotto scalare in $\ell_2(\mathbb{R})$ o meno. Dimostrare ciò che si afferma.

$$\langle x, y \rangle = \sum_{i=1}^{\infty} (x_i y_i + x_i y_{i+1} + x_{i+1} y_i)$$

Soluzione. Non è un prodotto scalare, perché non è definito positivo. Infatti, scegliendo

$$x := (1, -1, 0, 0, 0, 0, \dots) \in \ell_2.$$

si ottiene

$$\langle x, x \rangle = x_1^2 + x_2^2 + 2x_1 x_2 = 0.$$

pur essendo $x \neq 0$.

4.14 Problema. Dimostrare che la seguente espressione è un prodotto scalare in $\ell_2(\mathbb{R})$

$$\langle x, y \rangle = \sum_{i=1}^{\infty} (5x_i y_i + x_i y_{i+1} + x_{i+1} y_i).$$

4.15 Problema. Quale condizione deve soddisfare una matrice $n \times n$ A , affinché la seguente espressione sia un prodotto scalare in \mathbb{R}^n ?

$$\langle u, v \rangle := \sum_{i,j=1}^n A_{ij} u_i v_j \quad u, v \in \mathbb{R}^n$$

4.16 Problema. Dire se la seguente espressione è un prodotto scalare nello spazio $M_{n \times n}(\mathbb{R})$. Dimostrare ciò che si afferma (A^t è la trasposta di A).

$$(a) \quad \langle A, B \rangle := \operatorname{tr}(AB) \qquad (b) \quad \langle A, B \rangle := \operatorname{tr}(A^t B)$$

4.17 Problema. Sia $p \geq 1$. Dimostrare che se $p \neq 2$ non esiste alcun prodotto scalare su ℓ_p compatibile con la norma $\|\cdot\|_p$.

Soluzione. Si prenda $u = (1, 0, 0, \dots, 0)$, $v = (0, 1, 0, 0, \dots, 0)$ e si imponga la legge del parallelogramma.

4.18 Problema. Dimostrare che non esiste alcun prodotto scalare in $(C_b(\mathbb{R}), \|\cdot\|_u)$ (Sugg: si cerchino 2 funzioni che violano la regola del parallelogramma).

4.2 Complemento ortogonale

4.19 Definizione. Dati due vettori u, v in uno spazio euclideo reale $(V, \langle \cdot, \cdot \rangle)$ si definisce *angolo formato da u e v* la quantità

$$\arccos \left[\frac{\langle u, v \rangle}{\|u\| \|v\|} \right].$$

u e v sono detti *ortogonali* o *perpendicolari* se $\langle u, v \rangle = 0$ e si scrive $u \perp v$. Il concetto di ortogonalità è valido anche per uno spazio euclideo complesso. Nel caso in cui u sia ortogonale a tutti i vettori appartenenti ad un sottoinsieme $X \subset V$ si scrive $u \perp X$ oppure $u \in X^\perp$. Definiamo infatti:

4.20 Definizione. Dato un qualsiasi sottoinsieme X di uno spazio euclideo $(V, \langle \cdot, \cdot \rangle)$, si definisce *complemento ortogonale di X* l'insieme

$$X^\perp := \{v \in V : \langle v, w \rangle = 0 \quad \forall w \in X\}$$

4.21 Esempio. In \mathbb{R}^3 , sia $v_1 = (1, 0, 0)$ e $v_2 = (0, 1, 0)$. Allora

$$\{v_1, v_2\}^\perp = \text{l'asse } z.$$

4.22 Proposizione. Se X è un insieme di vettori nello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$, X^\perp è un sottospazio chiuso di V .

Dimostrazione. (1) X^\perp è un sottospazio. Infatti, se assumo che v_1 e v_2 appartengono a X^\perp e se $c_1, c_2 \in \mathbb{C}$, allora

$$\langle c_1 v_1 + c_2 v_2, w \rangle = c_1 \langle v_1, w \rangle + c_2 \langle v_2, w \rangle = 0 \quad \forall w \in X,$$

quindi $c_1 v_1 + c_2 v_2 \in X^\perp$, e dunque X^\perp è un sottospazio.

(2) X^\perp è chiuso. Infatti sia (v_n) una successione di elementi di X^\perp tale che $v_n \rightarrow v \in V$. Faccio vedere che $v \in X^\perp$. Fisso in vettore $w \in X$ arbitrario. Siccome $v_n \in X^\perp$ posso scrivere

$$\langle v_n, w \rangle = 0 \quad \forall n.$$

Poiché il prodotto scalare è continuo posso far passare il limite dentro il prodotto scalare e ottengo

$$0 = \lim_{n \rightarrow \infty} \langle v_n, w \rangle = \langle v, w \rangle.$$

Quindi $v \perp w$. Ma w è un vettore arbitrario di X , da cui concludo che $v \in X^\perp$. \square

4.23 Proposizione. *Se X, Y sono due insiemi di vettori nello uno spazio euclideo $(V, \langle \cdot, \cdot \rangle)$, allora*

(1) *se $X \subset Y$ allora $X^\perp \supset Y^\perp$*

(2) $X^\perp = \overline{X}^\perp = (\overline{\text{span}(X)})^\perp$.

Dimostrazione.

(1). Assumo $v \in Y^\perp$ e faccio vedere che $v \in X^\perp$. Infatti, se v è ortogonale a tutti i vettori di Y , a maggior ragione sarà ortogonale a tutti i vettori di X , poiché $X \subset Y$, quindi $v \in X^\perp$.

(2). Siccome si ha $X \subset \overline{X} \subset \overline{\text{span}(X)}$, grazie all'affermazione precedente ottengo

$$X^\perp \supset \overline{X}^\perp \supset (\overline{\text{span}(X)})^\perp.$$

Per concludere la dimostrazione basta far vedere che

$$(4.7) \quad X^\perp \subset (\overline{\text{span}(X)})^\perp.$$

Sia dunque $v \in X^\perp$, quindi $\langle v, w \rangle = 0$ per ogni $w \in X$. Voglio far vedere innanzi tutto che

$$(4.8) \quad v \in \text{span}(X)^\perp.$$

Infatti se $w \in \text{span}(X)$ allora w si può scrivere come combinazione lineare finita di elementi di X , vale a dire

$$w = \sum_{i=1}^n c_i w_i \quad w_i \in X.$$

Quindi, grazie alla bilinearità del prodotto scalare

$$\langle v, w \rangle = \sum_{i=1}^n c_i \langle v, w_i \rangle = 0,$$

come a dire che v è ortogonale a tutti i vettori che appartengono a $\text{span}(X)$. La (4.8) è dimostrata. Sia infine $w \in \overline{\text{span}(X)}$. Dimostro che $v \perp w$. Sia (w_n) una successione di elementi di $\text{span}(X)$ tale che $w_n \rightarrow w$. Allora, siccome il prodotto scalare è continuo,

$$\langle v, w \rangle = \langle v, \lim_{n \rightarrow \infty} w_n \rangle = \lim_{n \rightarrow \infty} \langle v, w_n \rangle = 0,$$

grazie alla (4.8) e al fatto che $w_n \in \text{span}(X)$.

Avendo ipotizzato $v \in X^\perp$ siamo arrivati alla conclusione che v è ortogonale a tutti gli elementi di $\overline{\text{span}(X)}$. La (4.7) è quindi dimostrata. \square

4.3 Il teorema della proiezione

In questa sezione vogliamo dimostrare che, sotto un'opportuna ipotesi (che nel caso finito dimensionale è sempre soddisfatta), dato un vettore v e un "iperpiano" W è sempre possibile scomporre v come somma di 2 termini $v = w + z$ in cui w è la proiezione di v sull'iperpiano W e z è ortogonale all'iperpiano.

4.24 Lemma. *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio di Hilbert, sia W un sottospazio chiuso di V e sia $v \in V$. Allora esiste nel sottospazio W un unico vettore w_0 che è il più vicino a v , cioè tale che*

$$(4.9) \quad d := \|v - w_0\| \leq \|v - w\| \quad \forall w \in W.$$

Dimostrazione. Sia

$$d := \inf_{w \in W} \|v - w\|.$$

Per definizione di inf posso trovare una successione $w_n \in W$ tale che

$$(4.10) \quad \lim_{n \rightarrow \infty} \|v - w_n\| = d.$$

Affermo che (w_n) è una successione di Cauchy. Infatti, usando la regola del parallelogramma, si ha

$$(4.11) \quad \begin{aligned} \|w_n - w_k\|^2 &= \|(w_n - v) + (v - w_k)\|^2 \\ &= 2\|v - w_n\|^2 + 2\|v - w_k\|^2 - \|w_n + w_k - 2v\|^2 \\ &= 2\|v - w_n\|^2 + 2\|v - w_k\|^2 - 4\|\tfrac{1}{2}(w_n + w_k) - v\|^2. \end{aligned}$$

Poiché $\tfrac{1}{2}(w_n + w_k)$ appartiene a W , deve valere, per definizione di d ,

$$\|\tfrac{1}{2}(w_n + w_k) - v\| \geq d,$$

per cui la (4.11) può essere riscritta come

$$(4.12) \quad \|w_n - w_k\|^2 \leq 2\|v - w_n\|^2 + 2\|v - w_k\|^2 - 4d^2.$$

Usando ora la (4.10) ottengo che per ogni $\varepsilon > 0$ esiste N tale che per ogni $n \geq N$ si ha $\|v - w_n\| < d + \varepsilon$. Di conseguenza ottengo dalla (4.12)

$$(4.13) \quad \forall n, k \geq N \quad \|w_n - w_k\|^2 \leq 4(d + \varepsilon)^2 - 4d^2 = 8\varepsilon + 4\varepsilon^2.$$

Quindi la successione (w_n) è di Cauchy. Siccome V è completo e W è chiuso in V , W è anch'esso completo. Dunque esiste $w_0 \in W$ tale che $w_n \rightarrow w_0$. A questo punto, nella (4.10), faccio tendere $n \rightarrow \infty$, e, poiché la norma è continua, ottengo

$$\|v - w_0\| = d.$$

Ho dunque dimostrato l'esistenza di un vettore w_0 nel sottospazio W che minimizza la distanza da v .

Dimostrazione dell'unicità. Supponiamo per assurdo che esistano due vettori distinti w_1 e w_2 appartenenti a W che minimizzano entrambi, in W , la distanza da v , vale a dire tali che

$$(4.14) \quad \|v - w_1\| = \|v - w_2\| \leq \|v - w\| \quad \forall w \in W.$$

Se si prova a fare un disegno in cui w_1 e w_2 sono equidistanti da v si intuisce che il vettore $w_3 := (w_1 + w_2)/2$ è più vicino a v di quanto non lo siano w_1 e w_2 . Più precisamente dovrebbe essere:

$$\|v - w_3\|^2 = \|v - w_1\|^2 - \frac{1}{4}\|w_2 - w_1\|^2.$$

Dimostriamolo. Sia

$$d := \|v - w_1\| = \|v - w_2\|,$$

e poniamo

$$z_1 = v - w_3 = v - \frac{1}{2}(w_1 + w_2) \qquad z_2 = \frac{1}{2}(w_1 - w_2).$$

Applicando la regola del parallelogramma (4.5) ai vettori z_1 e z_2 otteniamo

$$\|z_1 + z_2\|^2 + \|z_1 - z_2\|^2 = 2(\|z_1\|^2 + \|z_2\|^2),$$

vale a dire

$$\|v - w_1\|^2 + \|v - w_2\|^2 = 2\|v - w_3\|^2 + \frac{1}{2}\|w_1 - w_2\|^2,$$

da cui si ottiene

$$d^2 = \frac{1}{2} [\|v - w_1\|^2 + \|v - w_2\|^2] = \|v - w_3\|^2 + \frac{1}{4}\|w_1 - w_2\|^2.$$

Abbiamo ottenuto che, se $w_1 \neq w_2$, la distanza di w_3 da v è minore di d , in contraddizione con l'ipotesi. \square

4.25 Teorema. (*Teorema della proiezione*). *Se W è un sottospazio chiuso di uno spazio di Hilbert $(V, \langle \cdot, \cdot \rangle)$, allora ogni vettore $v \in V$ può essere univocamente scritto come*

$$(4.15) \qquad v = w + z \qquad \text{in cui } w \in W \text{ e } z \in W^\perp.$$

Dimostrazione.

❶ Esistenza della rappresentazione (4.15).

Sia $v \in V$. Il Lemma 4.24 ci dice che esiste $w \in W$ tale che

$$(4.16) \qquad d := \|v - w\| \leq \|v - w'\| \qquad \forall w' \in W.$$

Scrivo allora

$$v = w + (v - w) = w + z$$

e faccio vedere che $z := v - w$ è ortogonale a W , cioè che

$$\langle v - w, x \rangle = 0 \qquad x \in W.$$

Sia x un elemento arbitrario di W . Per far vedere che $v - w \perp x$ costruisco il vettore $w + tx$, in cui $t \in \mathbb{C}$. Il vettore $w + tx$ è ancora un elemento di W quindi, per la (4.16), si ha, $\forall t \in \mathbb{C}$,

$$(4.17) \qquad \begin{aligned} d^2 &\leq \|v - (w + tx)\|^2 = \|(v - w) - tx\|^2 \\ &= \|v - w\|^2 - \bar{t}\langle v - w, x \rangle - t\langle x, v - w \rangle + |t|^2 \|x\|^2. \end{aligned}$$

Come nella dimostrazione della disuguaglianza di Cauchy-Schwarz, scriviamo

$$\langle x, v - w \rangle = |\langle x, v - w \rangle| e^{i\vartheta}$$

e scegliamo t della forma $t = se^{-i\vartheta}$, con $s \in \mathbb{R}$ arbitrario. Le (4.17) diventa

$$-2s|\langle v - w, x \rangle| + s^2 \|x\|^2 \geq 0 \quad \forall s \in \mathbb{R},$$

che implica

$$\langle v - w, x \rangle = 0.$$

Quindi ho dimostrato che $v - w$ è ortogonale ad un qualsiasi elemento $x \in W$, cioè $z := v - w \in W^\perp$, per cui la rappresentazione (4.15) esiste.

② Unicità della rappresentazione (4.15).

Assumo che ci sia un altro modo di scrivere v come una somma di un elemento di W e un elemento di W^\perp

$$v = w' + z' \quad \text{in cui } w' \in W \text{ e } z' \in W^\perp.$$

Avrei allora

$$0 = (w - w') + (z - z') \quad \text{in cui } w - w' \in W \text{ e } z - z' \in W^\perp.$$

Quindi

$$\begin{aligned} 0 &= \|(w - w') + (z - z')\|^2 = \langle (w - w') + (z - z'), (w - w') + (z - z') \rangle \\ &= \|w - w'\|^2 + \|z - z'\|^2 \end{aligned}$$

Questo significa che $w = w'$ e $z = z'$. □

4.26 Osservazione. La novità infinito dimensionale è rappresentata dal dover assumere che W sia un sottospazio chiuso. Se W non è chiuso allora la conclusione del precedente teorema in generale è falsa. Consideriamo, ad esempio, ℓ_f come sottospazio di ℓ_2 . ℓ_f non è chiuso, infatti la sua chiusura (in ℓ_2) coincide proprio con ℓ_2 , in quanto ℓ_f è denso in ℓ_2 . Ora è chiaro che nessun vettore con infiniti elementi non nulli può essere rappresentato come nella (4.15) semplicemente perché $(\ell_f)^\perp = (\ell_2)^\perp = \{0\}$.

4.27 Proposizione. *Se W è un sottospazio dello spazio di Hilbert V , allora si ha*

$$(W^\perp)^\perp = \overline{W}$$

quindi nel caso in cui W sia chiuso si ottiene

$$(W^\perp)^\perp = W.$$

Dimostrazione. (1). Dimostrazione che $\overline{W} \subset (W^\perp)^\perp$.

Se $w \in W$ allora w è ortogonale a tutti i vettori di W^\perp , quindi $w \in (W^\perp)^\perp$. Abbiamo dunque

$$W \subset (W^\perp)^\perp.$$

D'altro canto possiamo prendere la chiusura di ambo i membri e, dato che $(W^\perp)^\perp$ è chiuso, si ottiene

$$\overline{W} \subset \overline{(W^\perp)^\perp} = (W^\perp)^\perp.$$

(2). Dimostrazione che $\overline{W} \supset (W^\perp)^\perp$.

Sia $w \in (W^\perp)^\perp$. Dato che \overline{W} è un sottospazio chiuso, per il Teorema 4.25 e la Proposizione 4.23, posso scrivere

$$(4.18) \quad w = v + z \quad \text{in cui } v \in \overline{W} \text{ e } z \in \overline{W}^\perp = W^\perp.$$

Siccome $w \in (W^\perp)^\perp$, si ha $\langle w, z \rangle = 0$, quindi

$$\|z\|^2 = \langle z, w - v \rangle = \langle z, w \rangle - \langle z, v \rangle = 0$$

quindi $z = 0$. Questo significa che $w = v \in \overline{W}$. \square

4.28 Definizione. (Proiezioni ortogonali). Il Teorema 4.25, garantendo l'esistenza e l'unicità della rappresentazione (4.15), ci permette di definire, fissato un sottospazio chiuso W di V , una funzione che indichiamo con π_W che associa ad ogni vettore $v \in V$ la sua proiezione ortogonale su W . Questa funzione $\pi_W : V \rightarrow W$ è univocamente definita dalle due condizioni

$$(4.19) \quad \pi_W v = w \iff (w \in W \text{ e } v - w \in W^\perp).$$

È facile vedere che π_W è un operatore lineare, vale a dire che

$$\pi_W(\alpha u + \beta v) = \alpha \pi_W(u) + \beta \pi_W(v) \quad \forall \alpha, \beta \in \mathbb{C}, \quad \forall u, v \in V.$$

4.29 Problema. Sia π_W la proiezione ortogonale sul sottospazio chiuso W di uno spazio di Hilbert V . Dimostrare che

- (a) $\pi_W^2 = \pi_W$, cioè $\pi_W(\pi_W(v)) = \pi_W(v)$
- (b) π_W è simmetrico, cioè $\langle \pi_W v, z \rangle = \langle v, \pi_W z \rangle$ per ogni $v, z \in V$.
- (c) Se W e Z sono due sottospazi chiusi di V mutuamente ortogonali, allora $\pi_W \circ \pi_Z = 0$.
- (d) Se W e Z sono due sottospazi chiusi di V tali che $W \subset Z$, allora $\pi_W \circ \pi_Z = \pi_Z \circ \pi_W = \pi_W$.

4.30 Problema. Sia π_W il proiettore ortogonale sul sottospazio chiuso W dello spazio di Hilbert V . Dimostrare che $\langle u, \pi_W u \rangle \geq 0$ per ogni $u \in V$.

4.31 Definizione. Lo spazio euclideo V si dice *somma diretta* dei sottospazi chiusi V_1, V_2, V_3, \dots e si scrive

$$V = \bigoplus_{k=1}^{\infty} V_k = V_1 \oplus V_2 \oplus \dots$$

se

- (1) i sottospazi V_k sono a due a due ortogonali
- (2) ogni elemento v di V si può univocamente rappresentare come

$$v = v_1 + v_2 + \dots \quad \text{in cui } v_k \in V_k$$

4.32 Corollario. Se W è un sottospazio chiuso dello spazio di Hilbert V allora

$$V = W \oplus W^\perp$$

4.4 Sistemi ortogonali, completezza, basi

4.33 Definizione. Un insieme di vettori non nulli $X \subset V$ è detto *sistema ortogonale* (SOG) se per ogni coppia di vettori $u, v \in X$ con $u \neq v$ si ha $u \perp v$. X è detto *sistema ortonormale* (SON) se inoltre per ogni $u \in X$ si ha $\|u\| = 1$.

4.34 Proposizione. Sia X un sistema ortogonale nello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$. Allora X è un sistema di vettori linearmente indipendenti.

Dimostrazione. Identica al caso finito dimensionale.

4.35 Definizione. Un sistema di vettori $(v_\alpha)_{\alpha \in I}$ ortogonale e completo in uno spazio euclideo $(V, \langle \cdot, \cdot \rangle)$ è detto *base ortogonale* (BOG) di V . Se si ha inoltre $\|v_\alpha\| = 1$ per ogni $\alpha \in I$ allora $(v_\alpha)_{\alpha \in I}$ è detto *base ortonormale* (BON).

4.36 Lemma. Sia v_1, v_2, \dots, v_n un SOG e sia u un vettore arbitrario. Allora il vettore z definito come

$$z := u - \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u, v_2 \rangle}{\|v_2\|^2} v_2 - \dots - \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n$$

è ortogonale a ciascuno dei v_i , di conseguenza si ha $z \perp \text{span}\{v_1, \dots, v_n\}$.

Dimostrazione. Si verifica immediatamente che per ogni $i = 1, \dots, n$, si ha $\langle z, v_i \rangle = 0$. Quindi

$$z \in \{v_1, \dots, v_n\}^\perp.$$

Grazie alla Proposizione 4.23 ottengo

$$z \in \{v_1, \dots, v_n\}^\perp = \overline{\text{span}\{v_1, \dots, v_n\}^\perp} = \text{span}\{v_1, \dots, v_n\}^\perp,$$

in cui l'ultima uguaglianza deriva dal fatto che il sottospazio generato da un numero finito di vettori è necessariamente chiuso. \square

4.37 Teorema. (Gram-Schmidt). Sia $(v_n)_{n=1}^\infty$ un insieme di vettori linearmente indipendenti nello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$. Allora esiste un SON $(u_n)_{n=1}^\infty$ tale che

(1) Ciascun u_n è una combinazione lineare di v_1, \dots, v_n

$$u_n = a_{n1}v_1 + \dots + a_{nn}v_n$$

con $a_{nn} \neq 0$

(2) Ciascun v_n può essere scritto come

$$v_n = b_{n1}u_1 + \dots + b_{nn}u_n$$

con $b_{nn} \neq 0$.

Di conseguenza, per ogni $n = 1, 2, \dots$, si ha

$$(4.20) \quad \text{span}\{v_1, \dots, v_n\} = \text{span}\{u_1, \dots, u_n\}.$$

Dimostrazione. Per costruire il SON utilizziamo la procedura di ortogonalizzazione di Gram–Schmidt, definendo:

$$(4.21) \quad \begin{array}{ll} w_1 = v_1 & u_1 = w_1/\|w_1\| \\ w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 & u_2 = w_2/\|w_2\| \\ w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2 & u_3 = w_3/\|w_3\| \\ \vdots & \vdots \\ w_n = v_n - \sum_{k=1}^{n-1} \frac{\langle v_n, w_k \rangle}{\|w_k\|^2} w_k & u_n = w_n/\|w_n\| \\ \vdots & \vdots \end{array}$$

Grazie al Lemma 4.36 possiamo affermare che, per ogni $n = 1, 2, \dots$ il vettore w_n è ortogonale ai vettori w_1, \dots, w_{n-1} . Poiché n è arbitrario il sistema $(w_n)_{n=1}^\infty$ è ortogonale e quindi il sistema $(u_n)_{n=1}^\infty$ è ortonormale.

La formula (4.21) ci dice che v_n si può scrivere come

$$v_n = w_n + \beta_{n1}w_1 + \dots + \beta_{n,k-1}w_{k-1},$$

quindi v_n è una combinazione lineare di w_1, \dots, w_n . Viceversa, poiché $w_1 = v_1$, sostituendo nella formula per w_2 otteniamo

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1.$$

Quindi w_2 è una combinazione lineare di v_1 e v_2 . Sostituendo questa espressione nella formula per w_3 otteniamo che w_3 è una combinazione lineare di v_1, v_2, v_3 . Mediante sostituzioni successive otteniamo che, per ogni intero positivo n , w_n è una combinazione lineare di v_1, \dots, v_n . Questo fatto, unitamente a ciò che abbiamo affermato in precedenza, vale a dire che ogni v_n è una combinazione lineare di w_1, \dots, w_n equivale a dire che per ogni n si ha

$$\text{span}\{v_1, \dots, v_n\} = \text{span}\{w_1, \dots, w_n\} = \text{span}\{u_1, \dots, u_n\},$$

in cui l'ultima uguaglianza è banale. \square

4.38 Problema. (Polinomi di Legendre). Usando la procedura di Gram–Schmidt, trovare un sistema ortonormale p_0, p_1, p_2, p_3 in $C_2[-1, 1]$, a partire dal seguente sistema di vettori linearmente indipendenti:

$$v_0(x) = 1, \quad v_1(x) = x, \quad v_2(x) = x^2, \quad v_3(x) = x^3, \quad v_4(x) = x^4$$

Soluzione. Uso il procedimento di ortogonalizzazione di Gram–Schmidt. Osservo innanzitutto che

$$\int_{-1}^1 x^n = \begin{cases} \frac{2}{n+1} & \text{se } n \text{ è un intero pari} \\ 0 & \text{se } n \text{ è un intero dispari.} \end{cases}$$

Quindi

$$w_0 = 1 \quad \|w_0\|^2 = \int_{-1}^1 1 \, dx = 2.$$

Normalizzo

$$p_0 = w_0 / \|w_0\| = 1/\sqrt{2}.$$

Procedo

$$w_1(x) = v_1(x) - \frac{\langle v_1, w_0 \rangle}{\|w_0\|^2} w_0(x) = x - \frac{1}{2} \int_{-1}^1 x \, 1 \, dx = x$$

e

$$\|w_1\|^2 = \int_{-1}^1 x^2 \, dx = 2/3$$

quindi

$$p_1(x) = w_1 / \|w_1\| = \sqrt{\frac{3}{2}} x$$

Il prossimo

$$w_2(x) = v_2(x) - \frac{\langle v_2, w_0 \rangle}{\|w_0\|^2} w_0(x) - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1(x).$$

Si ha

$$\langle v_2, w_0 \rangle = \int_{-1}^1 x^2 \, 1 \, dx = \frac{2}{3}$$

e

$$\langle v_2, w_1 \rangle = 0$$

Per cui

$$w_2(x) = x^2 - \frac{1}{3}.$$

Devo normalizzare w_2

$$\|w_2\|^2 = \int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 \, dx = \frac{8}{45}$$

quindi

$$p_2(x) = \sqrt{\frac{45}{8}} \left(x^2 - \frac{1}{3}\right) = \frac{1}{2} \sqrt{\frac{5}{2}} (3x^2 - 1).$$

w_3, p_3 si trovano proseguendo nello stesso modo.

4.39 Problema. (Polinomi di Hermite). Nello spazio euclideo pesato

$$C_2(\mathbb{R}, e^{-x^2} dx)$$

ortogonalizzare (senza normalizzare) i polinomi: $1, x, x^2, x^3, x^4$. (Sugg: ricorda la formula degli integrali Gaussiani del Problema 1.21).

Risp: $1, x, x^2 - \frac{1}{2}, x^3 - \frac{3x}{2}, x^4 - 3x^2 + \frac{3}{4}$.

4.40 Problema. (Polinomi trigonometrici). Dimostrare che il sistema di funzioni

$$1, \sin x, \cos x, \sin(2x), \cos(2x), \sin(3x), \cos(3x), \dots$$

è ortogonale in $C_2[0, 2\pi]$. (Sugg: si utilizzino le formule (di prostaferesi?) per scrivere il prodotto di due funzione trigonometriche come somma (o differenza) di funzioni trigonometriche).

4.41 Proposizione. (Esistenza di una base ortonormale). In uno spazio euclideo completo o separabile esiste una base ortonormale.

Dimostrazione nel caso separabile. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo separabile e sia $(v_n)_{n=1}^\infty$ un insieme numerabile denso in V . Da questo insieme di vettori si può estrarre un sistema linearmente indipendente semplicemente scartando quei v_n che si possono ottenere come combinazione lineare di v_1, \dots, v_{n-1} . Supponiamo di aver effettuato questa selezione. Chiamo w_1, w_2, \dots la collezione di vettori linearmente indipendenti così ottenuta. È chiaro che

$$\text{span}\{(w_n)_{n=1}^\infty\} = \text{span}\{(v_n)_{n=1}^\infty\}$$

Ma sappiamo che i v_n costituiscono un insieme denso in V , quindi

$$\overline{\text{span}\{(w_n)_{n=1}^\infty\}} = \overline{\text{span}\{(v_n)_{n=1}^\infty\}} = V,$$

vale a dire (w_n) è un insieme completo di vettori linearmente indipendenti. A questo punto applico la procedura di ortogonalizzazione di Gram–Schmidt a (w_n) e ottengo un sistema ortonormale (u_n) . Grazie alla (4.20) ottengo

$$\text{span}\{u_1, \dots, u_n\} = \text{span}\{w_1, \dots, w_n\}$$

e quindi, di nuovo,

$$\overline{\text{span}\{(u_n)_{n=1}^\infty\}} = \overline{\text{span}\{(w_n)_{n=1}^\infty\}} = V.$$

Il sistema ortonormale (u_n) è completo e, per definizione, una base. \square

4.42 Osservazione. (Sulla separabilità). Abbiamo visto che se V è uno spazio euclideo separabile esiste una base ortonormale numerabile $(u_k)_{k=1}^\infty$ che può essere “costruita” a partire da un sistema completo di vettori linearmente indipendenti, mediante la procedura di Gram–Schmidt. Nel caso in cui V sia completo ma non separabile è ancora vero che esiste una base ortonormale (si prende un sistema ortonormale massimale, la cui esistenza è garantita dal Lemma di Zorn), però:

- (a) Non esiste (in generale) un metodo “costruttivo” per trovare questa base ortonormale.
- (b) La base $(u_\alpha)_{\alpha \in I}$ è necessariamente non numerabile. Infatti se esistesse una base numerabile $(u_k)_{k=1}^\infty$ allora, dalla completezza di questo insieme di vettori seguirebbe che $V = \overline{\text{span}\{(u_k)\}}$. Sia ora $X := \text{span}_{\mathbb{Q}}\{(u_k)\}$ l’insieme delle combinazioni lineari finite a coefficienti razionali dei vettori u_k . È facile far vedere che X è numerabile e che X è denso in $\overline{\text{span}\{(u_k)\}}$, quindi X è denso in V . Di conseguenza V è separabile in contraddizione con la nostra ipotesi.

4.5 Disuguaglianza di Bessel

Supponiamo che

$$\mathcal{U} = (u_1, u_2, u_3, \dots)$$

costituisca una base ortonormale in uno spazio euclideo infinito dimensionale separabile V . Dato un vettore $v \in V$, possiamo associare a v la successione di numeri reali (o complessi, se V è uno spazio euclideo complesso)

$$c = (c_1, c_2, c_3, \dots) \quad \text{in cui } c_k := \langle v, u_k \rangle$$

I coefficienti c_k vengono chiamati *coefficienti di Fourier del vettore v rispetto alla base \mathcal{U}* . Dunque, fissata la base \mathcal{U} posso associare ad un elemento v di V un elemento c di \mathbb{R}^∞ (o \mathbb{C}^∞). Posso definire quindi un’applicazione

$$\begin{aligned} \mathcal{F}_{\mathcal{U}} : V &\rightarrow \mathbb{C}^\infty \\ &: v \rightarrow c \end{aligned}$$

4.43 Problema. Dimostrare che \mathcal{F}_U è un operatore lineare³

Sorgono a questo punto un certo numero di questioni riguardanti i coefficienti di Fourier:

4.44 Domande.

- (1) In analogia col caso finito-dimensionale, è vero che ogni vettore può essere scritto come somma dei suoi coefficienti di Fourier moltiplicati per i vettori di base corrispondenti? Vale a dire, si può scrivere

$$(4.22) \quad v = \sum_{k=1}^{\infty} c_k u_k \quad \text{in cui } c_k = \langle v, u_k \rangle ?$$

Per far vedere che vale la precedente uguaglianza bisogna dimostrare che

$$(4.23) \quad \lim_{n \rightarrow \infty} \left\| v - \sum_{k=1}^n c_k u_k \right\| = 0$$

- (2) Si può dire qualcosa di più sulle successioni $c = (c_1, c_2, \dots)$ dei coefficienti di Fourier? Potrebbe essere che c appartenga necessariamente a qualche spazio più piccolo di \mathbb{C}^∞ , come $\ell_\infty(\mathbb{C})$ o $\ell_0(\mathbb{C})$ o $\ell_p(\mathbb{C})$ per un qualche p ?
- (3) L'operatore \mathcal{F}_U è iniettivo? Vale a dire se $\mathcal{F}_U(v) = \mathcal{F}_U(w)$, quindi se v e w hanno gli stessi coefficienti di Fourier, allora deve necessariamente essere $v = w$?

4.45 Teorema. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, sia $(u_k)_{k=1}^\infty$ un sistema ortonormale in V e sia $v \in V$. Poniamo

$$c_k := \langle v, u_k \rangle \quad S_n := \sum_{k=1}^n c_k u_k .$$

Dati dei coefficienti arbitrari $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ definiamo anche

$$S_n^\alpha := \sum_{k=1}^n \alpha_k u_k .$$

Allora:

- (1) $\|v - S_n\| \leq \|v - S_n^\alpha\|$;
- (2) $\sum_{k=1}^\infty |c_k|^2 \leq \|v\|^2$ (disuguaglianza di Bessel);
- (3) si ha $v = \sum_{k=1}^\infty c_k u_k$ se e solo se $\sum_{k=1}^\infty c_k^2 = \|v\|^2$ (uguaglianza di Parseval).

Dimostrazione. Possiamo scrivere

$$(4.24) \quad \begin{aligned} \|v - S_n^\alpha\|^2 &= \langle v - S_n^\alpha, v - S_n^\alpha \rangle = \langle v, v \rangle - 2 \operatorname{Re} \langle v, S_n^\alpha \rangle + \langle S_n^\alpha, S_n^\alpha \rangle \\ &= \langle v, v \rangle - 2 \operatorname{Re} \sum_{k=1}^n c_k \overline{\alpha_k} + \sum_{k=1}^n |\alpha_k|^2 \\ &= \|v\|^2 - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |\alpha_k - c_k|^2 . \end{aligned}$$

³anche se "operatore lineare" non è stato ancora definito

Notare che la quantità (4.24) è minima quando l'ultima sommatoria è uguale a zero, quindi quando si ha $\alpha_k = c_k$. Di conseguenza, fra tutte le approssimazioni di v fatte di combinazioni lineari di u_1, \dots, u_n la migliore è quella ottenuta proprio con i coefficienti di Fourier, cioè

$$(4.25) \quad \|v - S_n\| \leq \|v - S_n^\alpha\|.$$

Nel caso di $\alpha_k = c_k$ la (4.24) diventa

$$(4.26) \quad \|v - S_n\|^2 = \|v\|^2 - \sum_{k=1}^n |c_k|^2.$$

Dato che il lato sinistro è non negativo, si ottiene

$$\sum_{k=1}^n |c_k|^2 \leq \|v\|^2$$

che implica, facendo tendere n all'infinito,

$$(4.27) \quad \sum_{k=1}^{\infty} |c_k|^2 \leq \|v\|^2 \quad (\text{disuguaglianza di Bessel}).$$

Questa disuguaglianza ci dice che la somma dei quadrati dei coefficienti di Fourier di un vettore v non supera il quadrato della norma di v , quindi, in particolare è finita! Possiamo dunque rispondere alla domanda posta in precedenza: se $c = \mathcal{F}_{\mathcal{U}}(v)$ è la successione dei coefficienti di Fourier rispetto ad un qualsiasi sistema ortonormale, allora si ha $c \in \ell_2$.

Infine, facendo tendere $n \rightarrow \infty$ nell'uguaglianza (4.26) si vede che

$$(4.28) \quad v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k \quad \text{se e solo se} \quad \|v\|^2 = \sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2 \quad \square$$

Il teorema precedente ci dice che i sono perciò due possibilità:

- (i) o la (4.27) è un'uguaglianza, vale, cioè la cosiddetta *uguaglianza di Parseval*

$$(4.29) \quad \sum_{k=1}^{\infty} |c_k|^2 = \|v\|^2 \quad (\text{uguaglianza di Parseval}).$$

e quindi v può effettivamente essere scritto come combinazione lineare infinita dei vettori u_k

$$(4.30) \quad v = \sum_{k=1}^{\infty} c_k u_k$$

- (ii) oppure le (4.27) è una disuguaglianza stretta, nel qual caso la (4.30) non è valida.

Notare che, dato che $c = \mathcal{F}_{\mathcal{U}}(v)$, l'uguaglianza di Parseval può essere simbolicamente riscritta come

$$(4.31) \quad \|\mathcal{F}_{\mathcal{U}}(v)\|_{\ell_2} = \|v\|_V \quad (\text{ancora l'uguaglianza di Parseval}).$$

Si capisce quindi che hanno importanza speciale quei sistemi ortonormali \mathcal{U} per i quali si verifica sempre l'uguaglianza di Parseval. Questo giustifica la seguente

4.46 Definizione. Il sistema ortonormale $(u_k)_{k=1}^{\infty}$ nello spazio euclideo V è detto *chiuso* se per ogni $v \in V$ si ha

$$\sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2 = \|v\|^2 \quad \text{e quindi anche} \quad v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k .$$

4.47 Proposizione. Sia $\mathcal{U} = (u_n)_{n=1}^{\infty}$ un sistema ortonormale nello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$. Allora \mathcal{U} è un sistema completo di vettori (quindi \mathcal{U} è una base ortonormale) se e solo se \mathcal{U} è chiuso. Inoltre, se \mathcal{U} è completo (o chiuso), V è necessariamente separabile.

Dimostrazione. ❶. Completo \Rightarrow chiuso. Sia $(u_k)_{k=1}^{\infty}$ un base ortonormale in V . Devo far vedere che il sistema di vettori (u_k) è chiuso, cioè che per ogni $v \in V$ si ha

$$v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k = \lim_{n \rightarrow \infty} S_n ,$$

dove ho indicato con S_n la somma parziale n -sima costruita con i coefficienti di Fourier di v , cioè

$$S_n = \sum_{k=1}^n \langle v, u_k \rangle u_k .$$

Equivalentemente devo far vedere che

$$(4.32) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che } \forall n \geq N \text{ si ha } \|v - S_n\| < \varepsilon .$$

Per definizione di completezza del sistema (u_k) si ha che

$$\overline{\text{span}(u_k)_{k=1}^{\infty}} = V .$$

In altre parole per ogni $v \in V$ e per ogni $\varepsilon > 0$ posso trovare un intero positivo N e dei coefficienti $\alpha_1, \dots, \alpha_N$ tali che

$$\left\| v - \sum_{k=1}^N \alpha_k u_k \right\| < \varepsilon .$$

Ma grazie alla disuguaglianza (4.25) ottengo

$$(4.33) \quad \|v - S_N\| \leq \left\| v - \sum_{k=1}^N \alpha_k u_k \right\| < \varepsilon .$$

Per dimostrare la (4.32) e quindi concludere mi basta osservare che, grazie all'identità (4.26), la quantità $\|v - S_n\|$ è non crescente in n , per cui dalla (4.33) segue che $\|v - S_n\| < \varepsilon$ per tutti gli n maggiori di N .

❷. Chiuso \Rightarrow completo. Se il sistema (u_k) è chiuso, per definizione, si ha che, per ogni $v \in V$,

$$v = \lim_{n \rightarrow \infty} \sum_{k=1}^n \langle v, u_k \rangle u_k = \lim_{n \rightarrow \infty} S_n .$$

Poiché $S_n \in \text{span}(u_k)$, la precedente equazione ci dice che v è un punto di aderenza di $\text{span}(u_k)$, cioè

$$v \in \overline{\text{span}(u_k)} .$$

Ma v è un vettore arbitrario in V , quindi

$$V = \overline{\text{span}(u_k)},$$

vale a dire il sistema (u_k) è completo.

③ La dimostrazione che la completezza di \mathcal{U} implica la separabilità di V segue dalle considerazioni fatte nell'Osservazione 4.42. \square

Il teorema 4.45 ci dice che, fissato un sistema ortonormale $(u_k)_{k=1}^{\infty}$ in uno spazio euclideo V , possiamo associare ad ogni vettore $v \in V$ la successione $c = (c_k)_{k=1}^{\infty}$ dei suoi coefficienti di Fourier. Questa successione appartiene allo spazio ℓ_2 . Possiamo procedere al contrario? Data una successione $c \in \ell_2$ possiamo ricostruire il vettore v ? Se V è completo la risposta è affermativa.

4.48 Teorema. (*Riesz-Fisher*). *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio di Hilbert, e sia*

$$\mathcal{U} = (u_1, u_2, u_3, \dots)$$

un sistema ortonormale⁴ in V . Allora per ogni $c = (c_1, c_2, \dots) \in \ell_2$ esiste $v \in V$ tale che $c = \mathcal{F}_{\mathcal{U}}(v)$, vale a dire

$$(4.34) \quad c_k = \langle v, u_k \rangle \quad \forall k = 1, 2, 3, \dots$$

Inoltre si ha

$$(4.35) \quad \|v\|^2 = \|c\|_2^2 = \sum_{k=1}^{\infty} |c_k|^2, \quad v = \sum_{k=1}^{\infty} c_k u_k.$$

Dimostrazione. Definisco una successione di vettori v_1, v_2, \dots , in cui

$$v_n := \sum_{k=1}^n c_k u_k.$$

Idea: far vedere che la successione (v_n) è una successione di Cauchy in V . Dato che V è completo, per ipotesi, in questo modo si ottiene che la successione (v_n) è convergente a un qualche elemento v di V .

Procedo a far vedere che (v_n) è di Cauchy: siano n, m sono due interi positivi con $m \geq n$. Si ha

$$(4.36) \quad \begin{aligned} \|v_n - v_m\|^2 &= \left\| \sum_{k=n+1}^m c_k u_k \right\|^2 = \left\langle \sum_{k=n+1}^m c_k u_k, \sum_{k=n+1}^m c_k u_k \right\rangle \\ &= \sum_{k=n+1}^m \sum_{j=n+1}^m c_j \overline{c_k} \langle u_j, u_k \rangle = \sum_{k=n+1}^m |c_k|^2 \end{aligned}$$

per cui

$$(4.37) \quad \|v_n - v_m\|^2 \leq \sum_{k=n+1}^{\infty} |c_k|^2.$$

Dato che, per ipotesi, $c \in \ell_2$, la serie $\sum_{k=1}^{\infty} |c_k|^2$ è convergente. Quindi il resto n -simo tende a zero, vale a dire

$$\lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} |c_k|^2 = 0$$

⁴non necessariamente una base

che, tradotto, significa

$$(4.38) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } N \text{ tale che se } n \geq N \text{ allora } \sum_{k=n+1}^{\infty} |c_k|^2 < \varepsilon.$$

Dalle (4.37) e (4.38) si ottiene

$$(4.39) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } N \text{ tale che se } n, m \geq N \text{ allora } \|v_n - v_m\| < \varepsilon$$

che, significa esattamente che la successione (v_n) è di Cauchy.

Dato che V è completo, esiste $v \in V$ tale che v_n converge a v , cioè

$$(4.40) \quad \lim_{n \rightarrow \infty} \|v - v_n\| = 0$$

ovvero, usando la definizione di v_n

$$(4.41) \quad v = \sum_{k=1}^{\infty} c_k u_k.$$

Ora devo far vedere che valgono le (4.34) e (4.35).

Iniziamo con la prima. Posso scrivere

$$(4.42) \quad \langle v, u_k \rangle = \langle v_n, u_k \rangle + \langle v - v_n, u_k \rangle$$

Se $n \geq k$ si ha che il primo termine è uguale a

$$\langle v_n, u_k \rangle = \sum_{i=1}^n c_i \langle u_i, u_k \rangle = c_k \quad \forall n \geq k$$

D'altra parte si ha

$$|\langle v - v_n, u_k \rangle| \leq \|v - v_n\| \|u_k\| = \|v - v_n\|$$

Grazie alla (4.40)

$$\lim_{n \rightarrow \infty} |\langle v - v_n, u_k \rangle| = 0$$

per cui, prendendo il limite $n \rightarrow \infty$ nella (4.42) ottengo

$$(4.43) \quad \langle v, u_k \rangle = c_k$$

Rimane da dimostrare la (4.35).

Ma questa è quasi immediata. Infatti dalle (4.41) e (4.43) segue

$$v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k$$

che è equivalente, come osservato nella (4.28) alla (4.35) □

4.49 Proposizione. *Sia $\mathcal{U} = (u_n)_{n=1}^{\infty}$ un sistema ortonormale nello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$. Si considerino le seguenti affermazioni:*

- (1) \mathcal{U} è un sistema completo di vettori (quindi \mathcal{U} è una base ortonormale).
- (2) \mathcal{U} è totale, vale a dire: non esiste alcun vettore $v \in V$ non nullo che sia ortogonale a tutti gli u_k .

Allora valgono le seguenti implicazioni:

$$\begin{array}{l} \text{se } V \text{ è completo} \\ (1) \Leftrightarrow (2) \end{array}$$

$$\begin{array}{l} \text{se } V \text{ non è completo} \\ (1) \Rightarrow (2) \end{array}$$

Dimostrazione che (1) \Rightarrow (2). Sia $(u_n)_{n=1}^\infty$ un sistema completo di vettori. Per la Proposizione 4.47 $(u_n)_{n=1}^\infty$ è un sistema chiuso, vale a dire per ogni $v \in V$ vale l'uguaglianza di Parseval

$$\|v\|^2 = \sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2.$$

Se v è ortogonale a tutti i vettori u_k , questa uguaglianza ci dice che v è necessariamente il vettore nullo.

Dimostrazione che (2) \Rightarrow (1) se V è completo. Voglio ora far vedere che, sotto l'ipotesi che V è completo (quindi uno spazio di Hilbert), vale l'implicazione contraria: se $(u_n)_{n=1}^\infty$ è totale allora è completo. Per dimostrarlo assumo che $(u_n)_{n=1}^\infty$ non sia un sistema completo di vettori in V e farò vedere che non è totale dimostrando che esiste un vettore $w \neq 0$ ortogonale a tutti gli u_n . Supponiamo quindi che $(u_n)_{n=1}^\infty$ non sia completo e, di conseguenza, non sia chiuso. Per definizione esiste (almeno) un vettore $v \in V$ per il quale non vale l'uguaglianza di Parseval, ma si ha

$$(4.44) \quad \|v\|^2 > \sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2.$$

Pongo

$$(4.45) \quad c_k := \langle v, u_k \rangle.$$

Poiché V è completo, per il Teorema di Riesz–Fisher, esiste $w \in V$ tale che

$$(4.46) \quad \langle w, u_k \rangle = c_k \quad \forall k = 1, 2, 3, \dots \text{ e, inoltre vale } \|w\|^2 = \sum_{k=1}^{\infty} |c_k|^2.$$

Dalle (4.44), (4.46) si ottiene che $\|v\|^2 > \|w\|^2$ quindi, se definisco $z := v - w$, il vettore z è necessariamente diverso da zero. D'altra parte dalle (4.45), (4.46) si ottiene anche

$$\langle z, u_k \rangle = \langle v - w, u_k \rangle = \langle v, u_k \rangle - \langle w, u_k \rangle = c_k - c_k = 0.$$

Abbiamo quindi trovato un vettore non nullo z ortogonale a tutti gli u_k . □

Nel seguente problema si capisce che l'assunzione di completezza dello spazio euclideo nella precedente proposizione è necessaria al fine di dimostrare l'implicazione (2) \Rightarrow (1).

- ★ **4.50 Problema.** Trovare uno spazio euclideo V e un sistema ortonormale $(u_k)_{k=1}^\infty$ non completo in V , tale che non esista alcun vettore $v \neq 0$ ortogonale a tutti gli u_k . (Sugg: la cosa è possibile solo se V non è completo. Si prenda come V il sottospazio di ℓ_2 definito come $V := \text{span}\{x, e^{(2)}, e^{(3)}, \dots\}$, in cui $x = (1, 1/2, 1/3, 1/4, \dots)$ e gli $e^{(n)}$ sono i vettori di base “canonici”. Si consideri poi il sistema ortonormale $\{e^{(2)}, e^{(3)}, e^{(4)}, \dots\}$. È ovvio che non è completo. Nonostante ciò si assuma che y sia un elemento di V ortogonale a tutti gli $e^{(n)}$ per $n = 2, 3, 4, \dots$ e si dimostri che deve essere necessariamente $y = (0, 0, 0, \dots)$).

4.5.1 C'è un unico spazio di Hilbert separabile

Una conseguenza importante del Teorema di Riesz–Fisher è il fatto che c'è un unico spazio di Hilbert (infinito dimensionale) separabile. Più precisamente se V e W sono due spazi di Hilbert infinito dimensionali separabili, allora esiste un *isomorfismo* $\varphi : V \rightarrow W$, cioè V e W sono *isomorfi*. L'applicazione φ è detta isomorfismo se è biunivoca e “preserva la struttura dello spazio” V . Nel caso di uno spazio di Hilbert, preservare la struttura significa che φ deve essere lineare e deve preservare il prodotto scalare, vale a dire

$$\langle \varphi u, \varphi v \rangle_W = \langle u, v \rangle_V \quad \forall u, v \in V.$$

Grazie all'identità di polarizzazione della Proposizione 4.10, per far vedere che φ preserva il prodotto scalare è sufficiente far vedere che preserva la norma associata, cioè che vale

$$(4.47) \quad \|\varphi u\|_W = \|u\|_V \quad \forall u \in V.$$

Infatti, assumendo che valga la (4.47), si ottiene (consideriamo per risparmiare inchiostro il caso reale, ma il caso complesso è identico)

$$\begin{aligned} \langle \varphi u, \varphi v \rangle_W &= \frac{1}{4} (\|\varphi u + \varphi v\|_W^2 - \|\varphi u - \varphi v\|_W^2) \\ &= \frac{1}{4} (\|\varphi(u+v)\|_W^2 - \|\varphi(u-v)\|_W^2) \\ &= \frac{1}{4} (\|u+v\|_V^2 - \|u-v\|_V^2) = \langle u, v \rangle_V. \end{aligned}$$

Osserviamo inoltre che se $\varphi : V \rightarrow W$ è lineare e preserva la norma, allora è necessariamente iniettiva. Infatti

$$\begin{aligned} \varphi u = \varphi v &\implies \varphi(u-v) = 0 \implies \|\varphi(u-v)\|_W = 0 \implies \|u-v\|_V = 0 \\ &\implies u-v = 0 \implies u = v. \end{aligned}$$

Riassumendo per far vedere che φ è un isomorfismo fra spazi di Hilbert è sufficiente dimostrare che

- (1) φ è lineare,
- (2) φ preserva la norma,
- (3) φ è suriettiva.

4.51 Teorema. *Siano V e W due spazi di Hilbert infinito dimensionali separabili. Allora V e W sono isomorfi.*

Dimostrazione. Dimostriamo che sia V che W sono isomorfi a ℓ_2 dotato del prodotto scalare canonico

$$\langle x, y \rangle = \sum_{k=1}^{\infty} x_k \bar{y}_k \quad x, y \in \ell_2.$$

Dalla Proposizione 4.41 e dall'Osservazione 4.42 segue che V ha una base ortonormale numerabile $\mathcal{U} := (u_k)_{k=1}^{\infty}$. Consideriamo la funzione $\mathcal{F}_{\mathcal{U}} : V \rightarrow \ell_2$ che associa ad ogni vettore $v \in V$ la successione $c = (c_k)_{k=1}^{\infty}$ dei suoi coefficienti di Fourier $c_k := \langle v, u_k \rangle$. È banale verificare che $\mathcal{F}_{\mathcal{U}}$ è lineare. Poiché \mathcal{U} è una base ortonormale vale l'uguaglianza di Parseval che, come osservato nella (4.31), equivale a dire che $\mathcal{F}_{\mathcal{U}}$ preserva la norma. Infine, il Teorema di Riesz–Fisher ci dice che per ogni $c \in \ell_2$ esiste $v \in V$ tale che $c = \mathcal{F}_{\mathcal{U}}(v)$ che vuol dire proprio che $\mathcal{F}_{\mathcal{U}}$ è suriettiva. In virtù di quanto detto prima dell'enunciato del teorema, possiamo concludere che $\mathcal{F}_{\mathcal{U}}$ è un isomorfismo da V a ℓ_2 . Quindi V è isomorfo a ℓ_2 . In simboli $V \cong \ell_2$. Analogamente si ottiene $W \cong \ell_2$ e dunque $V \cong W$. \square

4.6 Riassunto delle precedenti puntate

Sia dunque V uno spazio euclideo infinito dimensionale separabile e sia

$$\mathcal{U} = (u_1, u_2, u_3, \dots)$$

un sistema ortonormale in V . Allora, per ogni $v \in V$ vale l'identità

$$\|v - \sum_{k=1}^n \langle v, u_k \rangle u_k\|^2 = \|v\|^2 - \sum_{k=1}^n |\langle v, u_k \rangle|^2$$

Se \mathcal{U} è completo (e quindi chiuso) (e quindi totale) allora ambo i membri di questa identità tendono a zero quando $n \rightarrow \infty$, quindi

$$(4.48) \quad v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k \quad \text{e} \quad \|v\|^2 = \sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2$$

Se il sistema ortonormale \mathcal{U} non è completo, le cose si complicano. In questo caso, mentre ci saranno dei vettori v per i quali la (4.48) è ancora valida, ce ne sono anche degli altri per i quali

$$(4.49) \quad v \neq \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k \quad \text{e} \quad \|v\|^2 > \sum_{k=1}^{\infty} |\langle v, u_k \rangle|^2$$

Bene, ma che significa dire che

$$(4.50) \quad v \neq \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k ?$$

Può significare due cose:

- (1) o che la serie $\sum_{k=1}^{\infty} \langle v, u_k \rangle u_k$ converge a un qualche w , ma $w \neq v$. Questo significa che il vettore $v - w$ è ortogonale a tutti gli u_k , quindi il sistema \mathcal{U} non è totale
- (2) oppure che la serie $\sum_{k=1}^{\infty} \langle v, u_k \rangle u_k$ non converge a nessun elemento di V . Questo può accadere solo se V non è completo.

4.7 Proiezioni ortogonali

Nel caso in cui W sia un sottospazio chiuso separabile, possiamo scrivere una formula esplicita per l'operatore di proiezione ortogonale su W .

4.52 Teorema. *Sia W è un sottospazio chiuso separabile dello uno spazio di Hilbert $(V, \langle \cdot, \cdot \rangle)$ e sia $(u_k)_{k=1}^{\infty}$ una base ortonormale di W . Allora la proiezione ortogonale $\pi_W v$ di v su W , definita nella (4.19), si può scrivere come*

$$(4.51) \quad w = \pi_W v = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k.$$

4.53 Osservazione. (Ancora sulla separabilità). Se W non è separabile, la formula (4.51) diventa

$$w = \sum_{\alpha \in I} \langle v, u_{\alpha} \rangle u_{\alpha}$$

in cui I è un insieme non numerabile e (u_α) è una base ortonormale di W . Ora è facile far vedere che la somma di una infinità non numerabile di termini positivi è necessariamente infinita. E infatti si dimostra che, fissato w , solo una quantità finita o numerabile dei termini $\langle v, u_\alpha \rangle$ sono diversi da zero e quindi la sommatoria è di fatto una somma di una quantità al più numerabile di termini. Il sottoinsieme finito o numerabile degli u_α che danno un contributo non nullo alla somma non è però fissato a priori, ma varia al variare di $w \in W$.

Dimostrazione. Poiché W è un sottospazio chiuso di V e V è completo, per la Proposizione 2.37, si ha che W è anch'esso completo. Quindi W è uno spazio di Hilbert separabile. Se $(u_k)_{k=1}^\infty$ è una BON di W , essa sarà anche un SON in V . Dato $v \in V$, posso quindi definire la successione $(c_k)_{k=1}^\infty$ dei suoi coefficienti di Fourier, dati da

$$c_k := \langle v, u_k \rangle \quad k = 1, 2, 3, \dots$$

Grazie alla disuguaglianza di Bessel so che la successione $c = (c_k)_{k=1}^\infty$ appartiene ad ℓ_2 . Per il Teorema di Riesz-Fisher applicato a W esiste $w \in W$ tale che

$$(4.52) \quad w = \sum_{k=1}^{\infty} \langle v, u_k \rangle u_k.$$

Affermo che $w = \pi_W v$. Essendo ovviamente $w \in W$, per dimostrarlo è sufficiente far vedere che il vettore $z = v - w$ è ortogonale a W . Verifichiamo:

$$\begin{aligned} \langle z, u_j \rangle &= \langle v - w, u_j \rangle = \langle v, u_j \rangle - \langle w, u_j \rangle \\ &= c_j - \left\langle \lim_{n \rightarrow \infty} \sum_{k=1}^n c_k u_k, u_j \right\rangle \\ &= c_j - \lim_{n \rightarrow \infty} \left\langle \sum_{k=1}^n c_k u_k, u_j \right\rangle \quad [\text{uso la continuità del prod. scalare}] \\ &= c_j - \lim_{n \rightarrow \infty} \sum_{k=1}^n c_k \langle u_k, u_j \rangle \\ &= c_j - c_j = 0. \end{aligned}$$

Questo mostra che $z \perp u_k$, per ogni k . Grazie alla Proposizione 4.23 e usando il fatto che $(u_k)_{k=1}^\infty$ è una base di W otteniamo

$$z \in \overline{\text{span}\{u_1, u_2, \dots\}}^\perp = W^\perp.$$

Quindi $w \in W$ e $z = v - w \in W^\perp$, vale a dire w è la proiezione ortogonale di v su W . \square

4.7.1 Problemi

4.54 Problema. Sia $V = \mathbb{R}^n$ e sia $u \in \mathbb{R}^n$. Poniamo $\pi_u := \pi_{\text{span}\{u\}}$, per cui abbiamo

$$\pi_u v = \frac{\langle v, u \rangle}{\|u\|^2} u$$

Dimostrare che l'operatore π_u è rappresentato, nella base canonica di \mathbb{R}^n dalla matrice A con elementi

$$A_{ij} = \frac{u_i u_j}{\|u\|^2} \quad i, j = 1, \dots, n$$

4.55 Problema. Sia $u := (1, 2, 3, 4) \in \mathbb{R}^4$. Determinare la matrice che rappresenta l'operatore π_u (la proiezione ortogonale su $\text{span}\{u\}$) nella base canonica.

4.56 Problema. Sia $L = \mathbb{R}^n$ e sia u_1, \dots, u_p un sistema di p vettori ortogonali in \mathbb{R}^n . Sia anche $W = \text{span}\{u_1, \dots, u_p\}$. Poniamo

$$\pi_{u_k} := \pi_{\text{span}\{u_k\}}$$

Sia $A^{(k)}$ la matrice che rappresenta l'operatore π_{u_k} nella base canonica di \mathbb{R}^n

$$A_{ij}^{(k)} = \frac{u_{k,i} u_{k,j}}{\|u\|^2} \quad k = 1, \dots, p, \quad i, j = 1, \dots, n$$

Dimostrare che

$$\pi_W = \pi_{u_1} + \dots + \pi_{u_p},$$

quindi π_W è rappresentato nella stessa base dalla matrice

$$A = A^{(1)} + \dots + A^{(p)}$$

4.57 Problema. Sia $V = \mathbb{R}^4$ e poniamo

$$v_1 = (1, 1, 0, 0) \quad v_2 = (0, 1, 1, 0) \quad W = \text{span}\{v_1, v_2\}$$

Determinare la matrice 4×4 A che rappresenta π_W nella base canonica. Verificare che $A^2 = A$.

Soluzione. Ortogonalizzando i vettori v_1, v_2 si ottengono

$$w_1 = v_1 = (1, 1, 0, 0) \quad w_2 = (-1/2, 1/2, 1, 0)$$

Siano $A^{(1)}$ e $A^{(2)}$ le matrici che rappresentano π_{w_1} e π_{w_2} rispettivamente. Allora dalla formula generale

$$A_{ij}^{(k)} = \frac{u_{k,i} u_{k,j}}{\|u\|^2}$$

si ottiene

$$A^{(1)} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^{(2)} = \frac{1}{6} \begin{pmatrix} 1 & -1 & -2 & 0 \\ -1 & 1 & 2 & 0 \\ -2 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

e

$$A = A^{(1)} + A^{(2)} = \frac{1}{3} \begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 2 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

4.58 Problema. Sia $V = C_2[0, 1]$ e $W = \text{span}\{x, x^3\}$. Determinare il nucleo integrale $K(x, y)$ dell'operatore π_W . Verificare che

$$\int_0^1 K(x, y)K(y, z) dy = K(x, z).$$

Calcolare $\pi_W x^2$ e verificare che $x^2 - \pi_W(x^2)$ è ortogonale a W .

Soluzione. Ortogonalizzo i vettori $v_1 = x, v_2 = x^3$, ottenendo

$$\begin{aligned}
w_1(x) &= v_1(x) = x \\
\|w_1\|^2 &= \int_0^1 w_1(y)^2 dy = \int_0^1 y^2 dy = \frac{1}{3} \\
w_2(x) &= v_2(x) - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1(x) = x^3 - 3 \int_0^1 y^3 y dy x = x^3 - \frac{3}{5}x \\
\|w_2\|^2 &= \int_0^1 w_2(y)^2 dy = \int_0^1 \left(y^3 - \frac{3}{5}y\right)^2 dy = \frac{1}{7} - \frac{6}{5} \frac{1}{5} + \frac{9}{25} \frac{1}{3} = \frac{4}{175}
\end{aligned}$$

A questo punto posso scrivere

$$K_{w_i}(x, y) = \frac{w_i(x) w_i(y)}{\|w_i\|^2} \quad \text{e} \quad K(x, y) = K_{w_1}(x, y) + K_{w_2}(x, y)$$

quindi

$$\begin{aligned}
K_{w_1}(x, y) &= 3xy \\
K_{w_2}(x, y) &= \frac{175}{4} \left(x^3 - \frac{3}{5}x\right) \left(y^3 - \frac{3}{5}y\right) = \frac{175}{4} \left(x^3 y^3 - \frac{3}{5}(xy^3 + x^3 y) + \frac{9}{25}xy\right)
\end{aligned}$$

Finalmente

$$K(x, y) = 3xy + \frac{175}{4} \left(x^3 y^3 - \frac{3}{5}(xy^3 + x^3 y) + \frac{9}{25}xy\right) = \frac{5}{4} (35x^3 y^3 - 21(xy^3 + x^3 y) + 15xy)$$

L'ultima parte dell'esercizio è solo un calcolo

4.59 Problema. Sia $V = C_2[-1, 1]$ e $W = \text{span}\{x, x^3\}$. Determinare il nucleo integrale dell'operatore π_W , vale a dire determinare $K(x, y)$ tale che

$$(\pi_W f)(x) = \int_{-1}^1 K(x, y) f(y) dy$$

Calcolare $\pi_W x^5$.

Soluzione. Inanzitutto ortogonalizzo i vettori x, x^3 .

$$\begin{aligned}
w_1(x) &= x \\
\|w_1\|^2 &= \int_{-1}^1 x^2 dx = \frac{2}{3} \\
w_2(x) &= x^3 - \frac{\langle x^3, w_1 \rangle}{\|w_1\|^2} w_1(x) = x^3 - \frac{3}{2} \int_{-1}^1 x^4 dx x = x^3 - \frac{3}{5}x \\
\|w_2\|^2 &= \int_{-1}^1 \left(x^3 - \frac{3}{5}x\right)^2 dx = \frac{8}{175}
\end{aligned}$$

Il nucleo integrale di π_W è dato da

$$K(x, y) = \frac{w_1(x) w_1(y)}{\|w_1\|^2} + \frac{w_2(x) w_2(y)}{\|w_2\|^2} = \frac{3}{2} xy + \frac{175}{8} \left(x^3 - \frac{3}{5}x\right) \left(y^3 - \frac{3}{5}y\right)$$

Infine per calcolare la proiezione del vettore x^5 utilizzo la definizione di nucleo integrale che appare nel testo dell'esercizio

$$\begin{aligned}\pi_W x^5 &= \int_{-1}^1 K(x, y) y^5 dy = \int_{-1}^1 \left[\frac{3}{2} xy + \frac{175}{8} \left(x^3 - \frac{3}{5} x \right) \left(y^3 - \frac{3}{5} y \right) \right] y^5 dy \\ &= \frac{3}{2} x \int_{-1}^1 y^6 dy + \frac{175}{8} \left(x^3 - \frac{3}{5} x \right) \int_{-1}^1 \left(y^8 - \frac{3}{5} y^6 \right) dy \\ &= \frac{3}{7} x + \frac{175}{8} \left(x^3 - \frac{3}{5} x \right) \left(\frac{2}{9} - \frac{6}{35} \right) = \frac{3}{7} x + \frac{175}{8} \left(x^3 - \frac{3}{5} x \right) \frac{16}{315} \\ &= \frac{3}{7} x + \frac{10}{9} \left(x^3 - \frac{3}{5} x \right) = \frac{10}{9} x^3 - \frac{5}{21} x.\end{aligned}$$

(Facoltativo). Controllo che il risultato sia giusto. Il vettore $x^5 - \pi_W x^5$ deve essere ortogonale a x e a x^3 .

$$\begin{aligned}v(x) &:= x^5 - \pi_W x^5 = x^5 - \frac{10}{9} x^3 + \frac{5}{21} x \\ \langle v, x \rangle &= \int_{-1}^1 \left(x^5 - \frac{10}{9} x^3 + \frac{5}{21} x \right) x dx = 2 \left[\frac{1}{7} - \frac{2}{9} + \frac{5}{63} \right] = 0 \\ \langle v, x^3 \rangle &= \int_{-1}^1 \left(x^5 - \frac{10}{9} x^3 + \frac{5}{21} x \right) x^3 dx = 2 \left[\frac{1}{9} - \frac{10}{63} + \frac{1}{21} \right] = 0\end{aligned}$$

Ottimo!

4.8 Basi ortonormali negli spazi L_2

4.8.1 $L_2(\mathbb{R})$ e i polinomi di Hermite

4.60 Problema. (Polinomi di Hermite). Si consideri lo spazio euclideo $C_2(\mathbb{R}, e^{-x^2} dx)$ con prodotto scalare

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x) \overline{g(x)} e^{-x^2} dx.$$

- (a) Ortogonalizzare (senza normalizzare) i polinomi: $1, x, x^2, x^3$, utilizzando il procedimento di Gram-Schmidt.
 (b) Si definiscano i polinomi di Hermite come

$$H_n(x) := (-1)^n e^{x^2} D^n(e^{-x^2})$$

e si dimostri che

$$H'_n(x) = 2xH_n(x) - H_{n+1}(x).$$

- (c) Utilizzando l'identità ottenuta al punto precedente, dimostrare per induzione l'identità

$$H'_n(x) = 2nH_{n-1}(x)$$

che implica quindi la formula di ricorrenza

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x).$$

Dopo aver calcolato H_0 e H_1 a partire dalla definizione, calcolare con la formula di ricorrenza H_2, H_3, H_4, H_5 . Verificare che H_0, H_1, H_2 e H_3 coincidono, a parte la normalizzazione, con i polinomi ottenuti al punto (a).

(d) Dimostrare che H_n soddisfa l'equazione differenziale:

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

(e) Osservare che

$$\textcircled{1} \quad e^{-x^2} H_n(x) = -D(e^{-x^2} H_{n-1}(x))$$

e dimostrare che

$$\langle H_n, H_m \rangle = \sqrt{\pi} 2^n n! \delta_{n,m}.$$

In particolare quindi $(H_n)_{n=1}^\infty$ è un sistema ortogonale.

(Sugg: sia $I_{n,m} := \langle H_n, H_m \rangle$. Se $m = 0$ e $n > 0$ (o viceversa), l'ortogonalità segue dalla $\textcircled{1}$. Se m e n sono entrambi positivi si può integrare per parti usando la $\textcircled{1}$ ottenendo una relazione fra $I_{n,m}$ e $I_{n-1,m-1}$. Questa relazione implica $I_{n,m} = 0$ se $n \neq m$, mentre, se $n = m$ si ottiene, iterando, $I_{n,n} = \sqrt{\pi} 2^n n!$).

(f) Osservare che, se γ è un cammino chiuso nel piano complesso che gira una volta in senso antiorario attorno al punto x , grazie alla formula integrale di Cauchy posso scrivere

$$H_n(x) = (-1)^n e^{x^2} \frac{n!}{2\pi i} \int_\gamma \frac{e^{-z^2}}{(z-x)^{n+1}} dz.$$

Dimostrare quindi che la funzione generatrice dei polinomi di Hermite è data da

$$F(t, x) := \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{n!} = e^{2xt-t^2}.$$

(Sugg: scambiare la serie con l'integrale ed usare nuovamente la formula integrale di Cauchy per calcolare l'integrale risultante).

(g) Le funzioni di Hermite sono definite come

$$\psi_n(x) := A_n e^{-x^2/2} H_n(x) \quad \text{in cui } A_n = (\sqrt{\pi} 2^n n!)^{-1/2}.$$

Dimostrare che $(\psi_n)_{n=0}^\infty$ è un sistema ortonormale nello spazio euclideo $C_2(\mathbb{R})$ (o nel suo completamento $L_2(\mathbb{R})$). Dimostrare che ψ_n soddisfa l'equazione di Schrödinger dell'oscillatore armonico, data da (accatagliati a parte)

$$\left[-\frac{1}{2} \frac{d^2}{dx^2} + \frac{x^2}{2} \right] \psi_n(x) = \lambda \psi_n(x). \quad \text{con } \lambda = n + \frac{1}{2}.$$

5. Funzionali lineari e distribuzioni

*J: We should have shotguns for this kind of deal.
V: How many up there?
J: Three or four.
V: Counting our guy?
J: I'm not sure.
V: So there could be five guys up there?
J: It's possible.
V: We should have fuckin' shotguns.*

5.1 Funzionali lineari continui

5.1 Definizione. Un *funzionale lineare* F sullo spazio vettoriale V è un'applicazione $F : V \rightarrow \mathbb{R}$ tale che

- (1) $F(u + v) = F(u) + F(v)$ per ogni $u, v \in V$
- (2) $F(cu) = cF(u)$ per ogni $c \in \mathbb{R}, u \in V$

5.2 Esempi.

- (a) $V = \mathbb{R}$. Dato $a \in \mathbb{R}$ definisco $\varphi_a(x) := ax$. φ_a è un funzionale lineare su \mathbb{R} . Tutti i funzionali lineari su \mathbb{R} sono di questo tipo (perché?).
- (b) In \mathbb{R}^n . Se $a \in \mathbb{R}^n$ possiamo costruire il funzionale lineare φ_a definito come $\varphi_a(x) := \sum_{i=1}^n a_i x_i$.
- (c) Nello spazio ℓ_1 : se $a = (a_i)_{i=1}^\infty \in \ell_\infty$, posso definire il funzionale lineare $\varphi_a : \ell_1 \rightarrow \mathbb{R}$ come

$$(5.1) \quad \varphi_a(x) := \sum_{i=1}^{\infty} a_i x_i \quad x \in \ell_1$$

Questa sommatoria è finita, infatti

$$(5.2) \quad \left| \sum_{i=1}^{\infty} a_i x_i \right| \leq \sum_{i=1}^{\infty} |a_i x_i| \leq \sup_i |a_i| \sum_{i=1}^{\infty} |x_i| = \|a\|_\infty \|x\|_1$$

- (d) Nello spazio $(C_b(\mathbb{R}), \|\cdot\|_u)$: se $g \in C_1(\mathbb{R})$ allora posso definire il funzionale lineare $\varphi_g : C_b(\mathbb{R}) \rightarrow \mathbb{R}$ come

$$(5.3) \quad \varphi_g(f) := \int_{\mathbb{R}} g(x) f(x) dx \quad f \in C_b(\mathbb{R})$$

Osservo che l'integrale è ben definito, infatti

$$(5.4) \quad |\varphi_g(f)| \leq \int_{\mathbb{R}} |g(x) f(x)| dx \leq \sup_{x \in \mathbb{R}} |f(x)| \int_{\mathbb{R}} |g(x)| dx = \|g\|_1 \|f\|_u$$

(e) Nello spazio $(C_0(\mathbb{R}), \|\cdot\|_u)$: se $y \in \mathbb{R}$ definisco il funzionale lineare *delta di Dirac* $\delta_y : C_0(\mathbb{R}) \rightarrow \mathbb{R}$ come

$$(5.5) \quad \delta_y(f) := f(y)$$

Si ha ovviamente

$$|\delta_y(f)| = |f(y)| \leq \|f\|_u$$

5.3 Definizione. (Funzionali lineari continui). Dato lo spazio vettoriale normato $(V, \|\cdot\|)$ indichiamo con V^* l'insieme dei funzionali lineari *continui* su V . V^* è detto spazio *duale* o *coniugato* dello spazio V .

5.4 Definizione. Un funzionale lineare F sullo spazio vettoriale V è detto *limitato* se la quantità

$$\|F\| := \sup_{x \in V: \|x\| \leq 1} |F(x)|$$

è finita. La quantità $\|F\|$ è chiamata *norma* del funzionale limitato F .¹

5.5 Proposizione. Se F è un funzionale lineare limitato sullo spazio vettoriale normato V , la norma di F può anche essere scritta come

$$(5.6) \quad \|F\| = \sup_{x \neq 0} \frac{|F(x)|}{\|x\|}.$$

Di conseguenza vale la disuguaglianza

$$(5.7) \quad |F(x)| \leq \|F\| \|x\| \quad \forall x \in V.$$

Dimostrazione. Definisco

$$A := \sup_{x \in V: \|x\| \leq 1} |F(x)| \quad B := \sup_{x \neq 0} \frac{|F(x)|}{\|x\|}$$

e faccio vedere che $A = B$.

❶ Dimostrazione che $B \leq A$. Grazie all'omogeneità di F , posso scrivere

$$B = \sup_{x \neq 0} \frac{|F(x)|}{\|x\|} = \sup_{x \neq 0} \left| F\left(\frac{x}{\|x\|}\right) \right| = \sup_{x \in V: \|x\|=1} |F(x)| \leq \sup_{x \in V: \|x\| \leq 1} |F(x)| = A.$$

❷ Dimostrazione che $A \leq B$.

$$\begin{aligned} A &:= \sup_{x \in V: \|x\| \leq 1} |F(x)| = \sup_{x \in V: \|x\| \leq 1, x \neq 0} |F(x)| \\ &\leq \sup_{x \in V: \|x\| \leq 1, x \neq 0} \frac{|F(x)|}{\|x\|} \leq \sup_{x \in V: x \neq 0} \frac{|F(x)|}{\|x\|}. \quad \square \end{aligned}$$

5.6 Proposizione. (Condizioni equivalenti). Sia F un funzionale lineare sullo spazio vettoriale normato V . Le seguenti condizioni sono equivalenti

- (1) F è continuo
- (2) F è continuo nell'origine
- (3) F è limitato

¹dimostriamo fra poco che è una vera norma

Dimostrazione di (1) \Rightarrow (2). Banale

Dimostrazione di (2) \Rightarrow (3). Assumo F continuo in 0. Allora, per definizione di continuità, dato che $F(0) = 0$, si ha

$$(5.8) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } \delta > 0 \text{ tale che se } \|x\| \leq \delta \text{ allora } |F(x)| < \varepsilon$$

Ma allora se x è un vettore di V con $\|x\| \leq 1$ poniamo

$$u := \delta x$$

per cui

$$\|u\| = \|\delta x\| = |\delta| \|x\| \leq \delta$$

Quindi, sfruttando la (5.8) si ottiene

$$|F(x)| = |F(u/\delta)| = \frac{1}{\delta} |F(u)| \leq \frac{\varepsilon}{\delta} \quad \forall x \in V \text{ tali che } \|x\| \leq 1,$$

da cui

$$\|F\| = \sup_{x \in V: \|x\| \leq 1} |F(x)| \leq \frac{\varepsilon}{\delta}.$$

Quindi F è un funzionale lineare limitato.

Dimostrazione di (3) \Rightarrow (1). Assumo F limitato e voglio dimostrare che F è continuo. Devo dunque far vedere che

$$(5.9) \quad \forall x \in V, \forall \varepsilon > 0, \text{ esiste } \delta > 0 \text{ tale che se } \|y - x\| < \delta \text{ allora } |F(y) - F(x)| < \varepsilon$$

Scelgo $x \in V$ e $\varepsilon > 0$. Allora, grazie alla (5.7), posso scrivere

$$(5.10) \quad |F(y) - F(x)| = |F(y - x)| \leq \|F\| \|y - x\|$$

Scegliamo ora $\delta := \varepsilon / \|F\|$. In questo modo si ha che, se $\|y - x\| < \delta$ allora

$$|F(y) - F(x)| \leq \delta \|F\| = \frac{\varepsilon}{\|F\|} \|F\| = \varepsilon.$$

Abbiamo quindi dimostrato che F è continuo □

5.7 Proposizione.

(1) *Se esiste $C > 0$ tale che*

$$(5.11) \quad |F(x)| \leq C \|x\| \quad \forall x \in V$$

allora si ha che $F \in V^$ e $\|F\| \leq C$.*

(2) *Se invece $F \in V^*$ e*

$$(5.12) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } x \in V \text{ tale che } |F(x)| \geq (C - \varepsilon) \|x\|$$

allora si ha $\|F\| \geq C$.

5.8 Osservazione. Quindi per dimostrare che $\|F\| = C$ bisogna far vedere che valgono la (5.11) e la (5.12). In casi particolarmente fortunati la (5.12) può essere sostituita da una condizione più semplice, vale a dire che

$$(5.13) \quad \text{esiste } x \in V \text{ tale che } |F(x)| = C \|x\|$$

Dimostrazione. Se vale la (5.11), allora

$$\|F\| := \sup_{x \in V: \|x\| \leq 1} |F(x)| \leq \sup_{x \in V: \|x\| \leq 1} (C\|x\|) = C \sup_{x \in V: \|x\| \leq 1} \|x\| = C$$

Per dimostrare la seconda affermazione, assumo, per assurdo, che sia $\|F\| < C$ e scelgo $\varepsilon = (C - \|F\|)/2 > 0$ in modo tale che

$$C = \|F\| + 2\varepsilon$$

Per la (5.12) esisterà $x \in V$ tale che

$$|F(x)| \geq (C - \varepsilon)\|x\| = (\|F\| + \varepsilon)\|x\|,$$

che è un assurdo perché vale sempre $|F(x)| \leq \|F\|\|x\|$. \square

5.9 Problema. Nei casi seguenti dire se F è un funzionale lineare continuo su L . Nel caso di risposta negativa giustificare la propria affermazione

$$\begin{aligned} (a) \quad L = \ell_2, \quad F(x) &= \sum_{i=1}^{\infty} x_i & (b) \quad L = \ell_2, \quad F(x) &= \sum_{i=1}^{\infty} \frac{x_i}{i} \\ (c) \quad L = \ell_0, \quad F(x) &= \sum_{i=1}^{\infty} \frac{x_i}{\sqrt{i}} & (d) \quad L = C_4(\mathbb{R}), \quad F(f) &= \int_{\mathbb{R}} \frac{f(x)}{1 + \sqrt{|x|}} dx \\ (e) \quad L = C_{4/3}(\mathbb{R}), \quad F(f) &= \int_{\mathbb{R}} \frac{f(x)}{1 + \sqrt{|x|}} dx & (f) \quad L = C_0(\mathbb{R}), \quad F(f) &= \int_{\mathbb{R}} \frac{f(x)}{1 + |x|} dx \end{aligned}$$

Soluzione. (a) No. Si prenda $x_i = 1/i$. Allora $x \in \ell_2$, ma $\sum_{i=1}^{\infty} x_i = \infty$.

(b) Sì. Infatti, per la disuguaglianza di Cauchy-Schwarz,

$$\left| \sum_{i=1}^{\infty} \frac{x_i}{i} \right| \leq \|x\|_2 \left[\sum_{i=1}^{\infty} \frac{1}{i^2} \right]^{1/2}$$

e la seconda serie è convergente.

(c) No. Si prenda $x_i = 1/\sqrt{i}$. Allora $x \in \ell_0$, ma $\sum_{i=1}^{\infty} x_i/\sqrt{i} = \sum_{i=1}^{\infty} 1/i = \infty$.

(d) No. Si prenda $f(x) := (1 + |x|^{1/3})^{-1}$. Allora $f \in C_4(\mathbb{R})$, ma

$$\int_{\mathbb{R}} \frac{f(x)}{1 + \sqrt{|x|}} dx = \int_{\mathbb{R}} \frac{1}{(1 + |x|^{1/2})(1 + |x|^{1/3})} dx$$

L'integrando, quando $x \rightarrow \infty$, converge a zero come $|x|^{-5/6}$, quindi l'integrale è divergente.

(e) Sì. Infatti, per la disuguaglianza di Hölder,

$$\int_{\mathbb{R}} \frac{|f(x)|}{1 + \sqrt{|x|}} dx \leq \|f\|_{4/3} \left[\int_{\mathbb{R}} \frac{1}{(1 + \sqrt{|x|})^4} dx \right]^{1/4}$$

in cui l'integrale a destra è chiaramente convergente.

(f) No. Infatti si prenda $f(x) := [\log(2 + |x|)]^{-1}$. In questo caso

$$\int_{\mathbb{R}} \frac{f(x)}{1 + |x|} dx = \int_{\mathbb{R}} \frac{1}{(1 + |x|) \log(2 + |x|)} dx$$

L'integrando, quando $x \rightarrow \infty$, converge a zero come $(|x| \log |x|)^{-1}$, quindi l'integrale è divergente.

5.10 Problema. Calcolare la norma del funzionale lineare δ_a , con $a \in \mathbb{R}$ che agisce sullo spazio normato $(C_0(\mathbb{R}), \|\cdot\|_u)$.

5.11 Problema. Dimostrare che la delta di Dirac δ_a , nello spazio $(C_1(\mathbb{R}), \|\cdot\|_1)$, è un funzionale lineare *non* limitato.

5.12 Problema. Nello spazio ℓ_p consideriamo il funzionale lineare φ_a definito come

$$\varphi_a(x) := \sum_{i=1}^{\infty} a_i x_i.$$

Sia $q := p/(p-1)$, in modo tale che $\frac{1}{p} + \frac{1}{q} = 1$, e sia $a \in \ell_q$.

- (1) Dimostrare che il funzionale lineare φ_a è ben definito su ℓ_p , vale a dire la sommatoria è convergente per ogni $x \in \ell_p$
- (2) Dimostrare che il funzionale lineare φ_a è limitato su ℓ_p e trovare un limite superiore alla norma di φ_a
- (3) Dimostrare che $\|\varphi_a\| = \|a\|_q$.

Soluzione. Punti 1 e 2. Sia $a \in \ell_q$. Allora, per ogni $x \in \ell_p$ si ha, per la disuguaglianza di Hölder,

$$(5.14) \quad |\varphi_a(x)| = \left| \sum_{i=1}^{\infty} a_i x_i \right| \leq \|a\|_q \|x\|_p$$

Questa disuguaglianza dimostra che la sommatoria è finita e al tempo stesso che $\|\varphi_a\| \leq \|a\|_q$.

Punto 3. Sapendo già che $\|\varphi_a\| \leq \|a\|_q$, mi basta dimostrare che vale la disuguaglianza opposta $\|\varphi_a\| \geq \|a\|_q$. A questo scopo, dato $a \in \ell_q$, cerco un particolare vettore $x \in \ell_p$ tale che valga

$$(5.15) \quad |\varphi_a(x)| = \|a\|_q \|x\|_p$$

Non sempre è possibile trovarlo (vedi il Problema 5.18 in cui è necessario usare un metodo più complicato). In questo caso però con la scelta

$$x_i = \operatorname{sgn}(a_i) |a_i|^{q-1}$$

la (5.15) è soddisfatta. La verifica è un semplice calcolo □

5.13 Problema. Sia $g \in C_1(\mathbb{R})$ e sia φ_g il funzionale lineare nello spazio $(C_b(\mathbb{R}), \|\cdot\|_u)$ definito come

$$\varphi_g(f) := \int_{\mathbb{R}} g(x) f(x) dx \quad f \in C_b(\mathbb{R}).$$

Dimostrare che $\|\varphi_g\| = \|g\|_1$.

Soluzione. Dimostro prima che $\|\varphi_g\| \leq \|g\|_1$ e poi che $\|\varphi_g\| \geq \|g\|_1$. Osservo che per ogni $f \in C_b(\mathbb{R})$ si ha

$$|\varphi_g(f)| \leq \int_{\mathbb{R}} |g(x) f(x)| dx \leq \sup_{x \in \mathbb{R}} |f(x)| \int_{\mathbb{R}} |g(x)| dx = \|g\|_1 \|f\|_u$$

e questo implica, come osservato nella Proposizione 5.7, che $\|\varphi_g\| \leq \|g\|_1$.

Passiamo a dimostrare la disuguaglianza opposta. Questo è un po' più complicato perché, dato un $\varepsilon > 0$, devo trovare funzione $f \in C_b(\mathbb{R})$ (che in generale dipende dal valore di ε) tale che

$$\heartsuit \quad |\varphi_g(f)| := \left| \int_{\mathbb{R}} g(x) f(x) dx \right| > (\|g\|_1 - \varepsilon) \|f\|_u.$$

► *Divagazione.* Il problema deriva dal fatto che la funzione f deve essere continua. Se potessi scegliere una funzione f limitata ma non necessariamente continua, allora risolverei facilmente senza neppure dover usare l'epsilon (!). Sia infatti

$$\clubsuit \quad f(x) := \operatorname{sgn}(g(x)) : \begin{cases} +1 & \text{se } g(x) \geq 0 \\ -1 & \text{se } g(x) < 0 \end{cases}.$$

Il motivo di questa scelta di f è che posso scrivere

$$f(x) g(x) = |g(x)| \quad \forall x \in \mathbb{R}$$

e a questo punto ho svoltato. Infatti ottengo

$$|\varphi_g(f)| := \left| \int_{\mathbb{R}} g(x) f(x) dx \right| = \left| \int_{\mathbb{R}} |g(x)| dx \right| = \int_{\mathbb{R}} |g(x)| dx =: \|g\|_1,$$

che mi dice che la (\heartsuit) è soddisfatta, dato che $\|f\|_u = 1$. Purtroppo questa funzione f non è continua. *Fine Divagazione.* ◀

La divagazione precedente ci ha fatto perdere del tempo ma non è stata completamente inutile, perché l'idea (\clubsuit) è più o meno quella giusta. Siccome però f deve essere continua non può saltare istantaneamente da -1 a $+1$ e viceversa, quindi dovremo creare dei “raccordi obliqui” fra i tratti nei quali f vale $+1$ e quelli in cui vale -1 . Per di più² il numero di questi raccordi può essere infinito. Infatti la funzione continua g potrebbe cambiare segno un numero infinito di volte. Fortunatamente si tratta di un'infinità numerabile, perché ogni volta che g assume un valore (diciamo) positivo, allora sarà positiva in tutto un intervallino che contiene sicuramente un razionale. Un altro modo di dirlo è che, siccome g è continua, gli insiemi

$$\begin{aligned} A_+ &:= \{x \in \mathbb{R} : g(x) > 0\} \\ A_- &:= \{x \in \mathbb{R} : g(x) < 0\} \end{aligned}$$

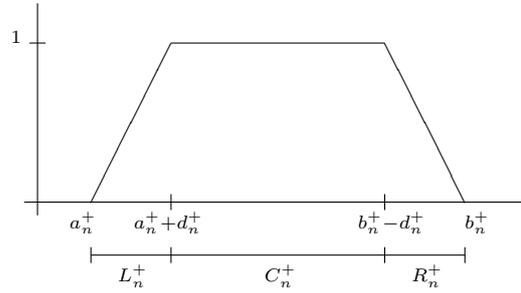
sono entrambi aperti. Ma è noto [KF p.64] che ogni insieme aperto di \mathbb{R} si può rappresentare come unione finita o numerabile di intervalli aperti disgiunti. Quindi³ esistono 4 successioni di numeri reali a_n^+ , b_n^+ , a_n^- , b_n^- tali che

$$\begin{aligned} A_+ &:= \cup_{n=1}^{\infty} (a_n^+, b_n^+) \\ A_- &:= \cup_{n=1}^{\infty} (a_n^-, b_n^-) \end{aligned}$$

Su ciascun intervallo (a_n^+, b_n^+) voglio definire la funzione f in modo che valga sempre 1 eccetto nelle vicinanze dei 2 estremi dell'intervallo (analogamente sugli intervalli (a_n^-, b_n^-)). Scelgo due successioni di numeri reali positivi (d_n^+) e (d_n^-) e definisco la funzione f sull'intervallo (a_n^+, b_n^+) nel modo seguente

²come se non avessimo di meglio da fare che creare “raccordi obliqui”

³facciamo il caso in cui il numero di questi intervalli è infinito numerabile



vale a dire

$$f(x) := \begin{cases} (x - a_n^+)/d_n^+ & \text{se } x \in L_n^+ := (a_n^+, a_n^+ + d_n^+) \\ 1 & \text{se } x \in C_n^+ := [a_n^+ + d_n^+, b_n^+ - d_n^+] \\ (b_n^+ - x)/d_n^+ & \text{se } x \in R_n^+ := (b_n^+ - d_n^+, b_n^+) \end{cases}$$

La definizione di f sugli intervalli (a_n^-, b_n^-) è analoga, con il segno cambiato. Infine definisco $f(x) = 0$ se $g(x) = 0$. Non ho ancora detto come devo scegliere le quantità d_n^\pm . Per il momento osservo soltanto che, affinché la definizione della funzione f sia univoca deve essere

$$d_n^+ \leq |b_n^+ - a_n^+|/2 \qquad d_n^- \leq |b_n^- - a_n^-|/2.$$

Una scelta precisa per queste quantità la farò più tardi. Ricordiamoci che il nostro scopo è dimostrare la (\heartsuit). Osservo che con la scelta fatta di f , le due funzioni f e g hanno sempre lo stesso segno, quindi il prodotto fg è sempre non negativo. Allora posso scrivere

$$\begin{aligned} \left| \int_{\mathbb{R}} g(x) f(x) dx \right| &= \int_{A_+} g(x) f(x) dx + \int_{A_-} g(x) f(x) dx \\ &= \sum_{n=1}^{\infty} \int_{a_n^+}^{b_n^+} g(x) f(x) dx + \sum_{n=1}^{\infty} \int_{a_n^-}^{b_n^-} g(x) f(x) dx. \end{aligned}$$

D'altra parte si ha

$$\begin{aligned} \int_{a_n^+}^{b_n^+} g(x) f(x) dx &\geq \int_{C_n^+} g(x) f(x) dx = \int_{C_n^+} g(x) dx \\ &= \int_{a_n^+}^{b_n^+} g(x) dx - \int_{L_n^+} g(x) dx - \int_{R_n^+} g(x) dx \end{aligned}$$

Sommando su tutti gli n ottengo

$$\int_{A_+} g(x) f(x) dx \geq \int_{A_+} g(x) dx - \sum_{n=1}^{\infty} \left[\int_{L_n^+} g(x) dx + \int_{R_n^+} g(x) dx \right].$$

Analogamente

$$\int_{A_-} g(x) f(x) dx \geq - \int_{A_-} g(x) dx + \sum_{n=1}^{\infty} \left[\int_{L_n^-} g(x) dx + \int_{R_n^-} g(x) dx \right],$$

che, poiché su questi intervalli g è negativa, si può riscrivere come

$$\int_{A_-} g(x) f(x) dx \geq \int_{A_-} |g(x)| dx - \sum_{n=1}^{\infty} \left[\int_{L_n^-} |g(x)| dx + \int_{R_n^-} |g(x)| dx \right].$$

Mettendo tutto insieme ho

$$|\varphi_g(f)| = \int_{\mathbb{R}} g(x) f(x) dx \geq \int_{A_+ \cup A_-} |g(x)| dx - \sum_{n=1}^{\infty} \left[\int_{L_n^+} |g(x)| dx + \int_{R_n^+} |g(x)| dx + \int_{L_n^-} |g(x)| dx + \int_{R_n^-} |g(x)| dx \right],$$

Dato che

$$\int_{A_+ \cup A_-} |g(x)| dx = \|g\|_1 \quad \text{e} \quad \|f\|_u = 1,$$

al fine di dimostrare la (\heartsuit) tutto ciò che manca è far vedere che, con una scelta opportuna delle quantità d_n^\pm , si ha

$$\sum_{n=1}^{\infty} \left[\int_{L_n^+} |g(x)| dx + \int_{R_n^+} |g(x)| dx + \int_{L_n^-} |g(x)| dx + \int_{R_n^-} |g(x)| dx \right] < \varepsilon.$$

A questo scopo uso il risultato del Problema 3.59 che afferma che se $g \in C_1(\mathbb{R})$ allora

per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che se $[a, b]$ è un qualsiasi intervallo di lunghezza non superiore a δ allora $\int_a^b |g(x)| dx < \varepsilon$.

Nel mio caso devo stare leggermente più attento perché ho una somma infinita di questi integrali, ma nulla di preoccupante. Utilizzo l'affermazione appena citata prendendo al posto di ε i valori $\varepsilon/(4 \cdot 2^n)$. Dunque ottengo una successione di quantità corrispondenti δ_n tali che

se $[a, b]$ è un intervallo di lunghezza non superiore a δ_n allora $\int_a^b |g(x)| dx < \frac{\varepsilon}{4 \cdot 2^n}$.

Una volta che ho i δ_n posso finalmente scegliere la larghezza d_n^\pm degli intervallini sui quali f è obliqua, larghezza che avevo lasciato in sospeso. Mi ricordo che d_n^\pm può essere al massimo uguale a $|b_n^\pm - a_n^\pm|/2$, dunque pongo

$$d_n^+ := \min\{\delta_n, |b_n^+ - a_n^+|/2\}$$

$$d_n^- := \min\{\delta_n, |b_n^- - a_n^-|/2\}.$$

A questo punto ho ottenuto

$$\sum_{n=1}^{\infty} \left[\int_{L_n^+} |g(x)| dx + \int_{R_n^+} |g(x)| dx + \int_{L_n^-} |g(x)| dx + \int_{R_n^-} |g(x)| dx \right]$$

$$< \sum_{n=1}^{\infty} \left[\frac{\varepsilon}{4 \cdot 2^n} + \frac{\varepsilon}{4 \cdot 2^n} + \frac{\varepsilon}{4 \cdot 2^n} + \frac{\varepsilon}{4 \cdot 2^n} \right] = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$$

che è esattamente ciò che dovevo dimostrare □

5.2 Lo spazio duale

5.2.1 Lo spazio duale è uno spazio vettoriale normato

5.14 Definizione. Dati due funzionali lineari F e G sullo spazio vettoriale normato $(V, \|\cdot\|)$ e dato $c \in \mathbb{R}$, posso definire le operazioni di

(1) *somma* fra due funzionali lineari come

$$(F + G)(x) := F(x) + G(x) \quad x \in V$$

(2) *prodotto* di un funzionale lineare per un numero come

$$(cF)(x) := cF(x) \quad x \in V.$$

5.15 Proposizione. *Lo spazio duale V^* di uno spazio vettoriale normato $(V, \|\cdot\|)$, con le due operazioni introdotte nella Definizione 5.14, è uno spazio vettoriale. La quantità*

$$\|F\| := \sup_{x \in V: \|x\| \leq 1} |F(x)|$$

è una norma su V^ .*

Dimostrazione. Vanno dimostrate le proprietà (A), (B), (SV1), ..., (SV8) che definiscono uno spazio vettoriale (Definizione 3.1). Tutte queste proprietà si dimostrano in modo completamente meccanico, non è richiesta attività cerebrale oltre a quella necessaria per l'uso della penna. Ad esempio, per quanto riguarda la proprietà (A). Devo far vedere che se $F, G \in V^*$ allora la loro somma $F + G$ appartiene a V^* . Sia $H := F + G$. Dire che $H \in V^*$ significa due cose: che H è un funzionale lineare su V e che H è limitato. Vediamo che H è un funzionale lineare. Siano $x, y \in V$ e $c \in \mathbb{R}$. Allora

$$\begin{aligned} H(x + y) &:= F(x + y) + G(x + y) = F(x) + F(y) + G(x) + G(y) = H(x) + H(y) \\ H(cx) &:= F(cx) + G(cx) = cF(x) + cG(x) = cH(x). \end{aligned}$$

Quindi H è lineare. Per quanto riguarda la limitatezza osservo che

$$\begin{aligned} \sup_{x \in V} |H(x)| &= \sup_{x \in V} |F(x) + G(x)| \leq \sup_{x \in V} [|F(x)| + |G(x)|] \\ &\leq \sup_{x \in V} |F(x)| + \sup_{x \in V} |G(x)| = \|F\| + \|G\|. \end{aligned}$$

Ho quindi dimostrato non solo che H è limitato, ma anche che

$$\|F + G\| \leq \|F\| + \|G\| \quad \forall F, G \in V^*,$$

che ci torna utile per affermare che $\|\cdot\|$ è una norma su V^* .

In modo pressoché identico si dimostra che se $F \in V^*$ e $c \in \mathbb{R}$ allora $cF \in V^*$ (proprietà (B)).

Così, tanto per evidenziare la banalità della cosa, dimostro una delle altre proprietà degli spazi vettoriali, diciamo la (SV4) che afferma

$$\text{per ogni } F \in V^* \text{ esiste } G \in V^* \text{ tale che } F + G = 0.$$

Sia dunque $F \in V^*$. Allora definisco un'applicazione $G : V \rightarrow \mathbb{R}$ come

$$G(x) := -F(x) \quad x \in V.$$

È immediato verificare che G è lineare e che G è limitato con $\|G\| = \|F\|$. Quindi $G \in V^*$ e, ovviamente $F + G = 0$.

Le altre proprietà sono altrettanto ovvie. □

5.16 Problema. Sia F il funzionale lineare definito come $F(f) := \int_{-1}^1 x f(x) dx$. Calcolare $\|F\|$ quando F agisce: (a) su $(C[-1, 1], \|\cdot\|_1)$, (b) su $(C[-1, 1], \|\cdot\|_2)$, (c) su $(C[-1, 1], \|\cdot\|_u)$.

5.17 Problema. Sia F il funzionale lineare definito come

$$F(f) := \int_0^\pi \cos x f(x) dx.$$

Calcolare $\|F\|$ quando F agisce: (a) sullo spazio $(C[0, \pi], \|\cdot\|_1)$; (b) sullo spazio $(C[0, \pi], \|\cdot\|_2)$ (in questo caso dare semplicemente la risposta senza dimostrazione).

Soluzione. (a) Se f è una funzione continua su $[0, \pi]$ si ha

$$|F(f)| = \left| \int_0^\pi \cos x f(x) dx \right| \leq \|\cos x\|_u \|f\|_1.$$

Questa disuguaglianza ci dice che

$$\|F\| \leq \|\cos x\|_u = 1.$$

D'altra parte consideriamo una successioni di funzioni continue f_n che hanno un picco sempre più stretto in corrispondenza di uno dei massimi del $|\cos x|$, ad esempio in $x = 0$. Scegliamo, ad esempio,

$$f_n(x) := \begin{cases} 1 - nx & \text{se } x \in [0, 1/n] \\ 0 & \text{se } x \in (1/n, \pi) \end{cases}$$

In questo modo si ottiene

$$R(f_n) := \frac{|F(f_n)|}{\|f_n\|_1} = \frac{|\int_0^1 \cos x f_n(x) dx|}{\int_0^1 |f_n(x)| dx} = \frac{\int_0^{1/n} \cos x f_n(x) dx}{\int_0^{1/n} f_n(x) dx}.$$

Poiché il coseno è decrescente nell'intervallo $[0, 1/n]$ si avrà

$$\int_0^{1/n} \cos x f_n(x) dx \geq \cos(1/n) \int_0^{1/n} f_n(x) dx.$$

Di conseguenza

$$R(f_n) \geq \frac{\cos(1/n) \int_0^{1/n} f_n(x) dx}{\int_0^{1/n} f_n(x) dx} = \cos(1/n).$$

Poiché sappiamo già che $\|F\| \leq 1$ otteniamo

$$\cos(1/n) \leq R(f_n) \leq 1,$$

di conseguenza, siccome $\lim_{n \rightarrow \infty} \cos(1/n) = 1$, si ha

$$\lim_{n \rightarrow \infty} R(f_n) = 1,$$

che dimostra $\|F\| = 1$.

(b) Utilizzo il seguente fatto noto: se $p > 1$, $q = p/(p-1)$ e $g \in C_q(\mathbb{R})$ allora il funzionale lineare $\varphi_g : C_p(\mathbb{R}) \mapsto \mathbb{R}$ definito come

$$\varphi_g(f) := \int_{\mathbb{R}} g(x) f(x) dx \quad f \in C_p(\mathbb{R})$$

è limitato e vale $\|\varphi_g\| = \|g\|_q$. Nel caso considerato ottengo:

$$\|F\|^2 = \|\cos x\|_2^2 = \int_0^\pi \cos^2 x dx = \int_0^\pi \frac{1 + \cos(2x)}{2} dx = \frac{\pi}{2},$$

quindi $\|F\| = \sqrt{\pi/2}$.

5.2.2 Identificazione di alcuni spazi duali importanti

5.18 Problema. Nello spazio ℓ_1 consideriamo il funzionale lineare φ_a definito come

$$\varphi_a(x) := \sum_{i=1}^{\infty} a_i x_i.$$

Allora

- (1) Dimostrare che se $a \in \ell_\infty$, il funzionale lineare φ_a è ben definito su ℓ_1 , vale a dire la sommatoria è convergente per ogni $x \in \ell_1$
- (2) Dimostrare che se $a \in \ell_\infty$, il funzionale lineare φ_a è limitato su ℓ_1
- (3) Dimostrare che vale $\|\varphi_a\| = \|a\|_\infty$.
- (4) Dimostrare inoltre che φ è un isomorfismo di ℓ_∞ su ℓ_1^* , vale a dire far vedere che ogni funzionale lineare continuo F su ℓ_1 si può scrivere come $F = \varphi_a$ per un qualche $a \in \ell_\infty$.

Soluzione. Punti 1 e 2. Sia $a \in \ell_\infty$. Allora, per ogni $x \in \ell_1$ si ha

$$(5.16) \quad |\varphi_a(x)| = \left| \sum_{i=1}^{\infty} a_i x_i \right| \leq \sum_{i=1}^{\infty} |a_i| |x_i| \leq \sup_i |a_i| \sum_{i=1}^{\infty} |x_i| = \|a\|_\infty \|x\|_1$$

Questa disuguaglianza dimostra che la sommatoria è finita e al tempo stesso che $\|\varphi_a\| \leq \|a\|_\infty$.

Punto 3. Sappiamo che $\|\varphi_a\| \leq \|a\|_\infty$. Per dimostrare che $\|\varphi_a\| \geq \|a\|_\infty$, usiamo la condizione (5.12). Sia dunque $\varepsilon > 0$. Voglio trovare $x \in \ell_1$ tale che

$$(5.17) \quad |\varphi_a(x)| \geq (\|a\|_\infty - \varepsilon) \|x\|_1$$

Dato che $\|a\|_\infty := \sup_i |a_i|$, per definizione di estremo superiore, esiste un intero j tale che $|a_j| > \|a\|_\infty - \varepsilon$. Scelgo a questo punto la mia successione $x \in \ell_1$ come

$$x = (0, 0, \dots, 0, 1, 0, 0, 0, \dots)$$

in cui l'elemento 1 appare alla posizione j -sima. Ovviamente si ha

$$\|x\|_1 := \sum_{i=1}^{\infty} |x_i| = 1.$$

Con questa scelta di x ottengo.

$$|\varphi_a(x)| = \left| \sum_{i=1}^{\infty} a_i x_i \right| = |a_j x_j| = |a_j| > \|a\|_\infty - \varepsilon = (\|a\|_\infty - \varepsilon) \|x\|_1$$

e la (5.17) è dimostrata. Quindi $\|\varphi_a\| = \|a\|_\infty$

Punto 4. Ora voglio far vedere che se F è un funzionale lineare continuo su ℓ_1 allora esiste $a \in \ell_\infty$ tale che $F = \varphi_a$. Sia dunque $F \in \ell_1^*$ e $x \in \ell_1$. Se $(e^{(n)})_{n=1}^\infty$ è la base canonica di ℓ_1 si ha

$$\lim_{n \rightarrow \infty} \left\| x - \sum_{i=1}^n x_i e^{(i)} \right\|_1 = \lim_{n \rightarrow \infty} \sum_{i=n+1}^{\infty} |x_i| = 0,$$

quindi posso scrivere

$$x = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_i e^{(i)} = \sum_{i=1}^{\infty} x_i e^{(i)}$$

Dato che F è continuo si avrà

$$F(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_i F(e^{(i)}) = \sum_{i=1}^{\infty} x_i F(e^{(i)})$$

Ponendo $a_i = F(e^{(i)})$ ottengo proprio che $F = \varphi_a$. Devo ancora far vedere che $a \in \ell_{\infty}$, cioè che a è una successione limitata. Ma dato che F è continuo e dunque limitato si ha

$$|a_i| = |F(e^{(i)})| \leq \|F\| \|e^{(i)}\|_1 = \|F\|$$

quindi $\|a\|_{\infty} \leq \|F\|$, quindi $a \in \ell_{\infty}$ □

5.19 Proposizione. *Nello spazio normato $(\ell_0, \|\cdot\|_{\infty})$ sia $\varphi_a : \ell_0 \rightarrow \mathbb{R}$ definito come*

$$(5.18) \quad \varphi_a(x) := \sum_{i=1}^{\infty} a_i x_i \quad x \in \ell_0$$

Se $a = (a_k)_{k=1}^{\infty} \in \ell_1$, allora

- (1) φ_a è un funzionale lineare continuo su ℓ_0 .
- (2) $\|\varphi_a\| = \|a\|_1 := \sum_{k=1}^{\infty} |a_k|$
- (3) Se F è funzionale lineare continuo su ℓ_0 allora esiste $a \in \ell_1$ tale che $F = \varphi_a$, vale a dire tutti gli elementi di ℓ_0^* si possono scrivere nella forma (5.18) con $a \in \ell_1$.

Dimostrazione di (1).

$$(5.19) \quad |\varphi_a(x)| = \left| \sum_{i=1}^{\infty} a_i x_i \right| \leq \sum_{i=1}^{\infty} |a_i x_i| \leq \sup_i |x_i| \sum_{i=1}^{\infty} |a_i| = \|a\|_1 \|x\|_{\infty}$$

Dimostrazione di (2). La (5.19) dimostra che $\|\varphi_a\| \leq \|a\|_1$. Per dimostrare che invece $\|\varphi_a\| \geq \|a\|_1$, sia $\alpha \in \ell_1$. Faremo vedere che

$$(5.20) \quad \text{per ogni } \varepsilon > 0 \text{ esiste } x \in \ell_0 \text{ tale che } |\varphi_a(x)| \geq (\|a\|_1 - \varepsilon) \|x\|_{\infty}$$

da cui, grazie alla (5.12), potremo concludere $\|\varphi_a\| \geq \|a\|_1$. Sia quindi $\varepsilon > 0$. Siccome $a \in \ell_1$ esiste n tale che

$$(5.21) \quad \sum_{i=n+1}^{\infty} |a_i| < \varepsilon$$

Poniamo quindi

$$x_i = \begin{cases} \operatorname{sgn}(a_i) & \text{se } i \leq n \\ 0 & \text{se } i > n \end{cases}$$

in cui $\operatorname{sgn}(a_i)$ rappresenta il *segno* del numero reale a_i , vale a dire $+1$ se a_i è positivo o nullo e -1 se a_i è negativo. Quindi la successione x può essere scritta come

$$x = \left(\frac{a_1}{|a_1|}, \frac{a_2}{|a_2|}, \dots, \frac{a_n}{|a_n|}, 0, 0, 0, \dots \right)$$

in cui, convenzionalmente, se $a_i = 0$, intendiamo $a_i/|a_i| = 0$. Chiaramente $x \in \ell_0$ e la sua norma è

$$\|x\|_\infty := \sup_i |x_i| = 1.$$

Il valore di φ_a su questo vettore x allora si può stimare dal basso nel modo seguente:

$$|\varphi_a(x)| = \left| \sum_{i=1}^{\infty} a_i x_i \right| = \left| \sum_{i=1}^n a_i \frac{a_i}{|a_i|} \right| = \left| \sum_{i=1}^n a_i \frac{a_i}{|a_i|} \right| = \sum_{i=1}^n |a_i| = \|a\|_1 - \sum_{i=n+1}^{\infty} |a_i| > \|a\|_1 - \varepsilon$$

Dato che $\|x\|_1 = 1$ si ha dunque

$$|\varphi_a(x)| \geq (\|a\|_1 - \varepsilon) \|x\|_1.$$

Abbiamo dimostrato in questo modo la (5.20), quindi $\|\varphi_a\| \geq \|a\|_1$. Insieme alla disuguaglianza inversa che già conoscevamo grazie alla (5.19) questo ci permette di concludere che $\|\varphi_a\| = \|a\|_1$.

Dimostrazione di (3). Sia F un funzionale lineare continuo su ℓ_0 . Voglio far vedere che esiste $a \in \ell_1$ tale che $F = \varphi_a$. Sia quindi $x \in \ell_0$, cioè $x = (x_i)_{i=1}^{\infty}$ e vale

$$(5.22) \quad \lim_{i \rightarrow \infty} x_i = 0.$$

Poniamo, come al solito,

$$e^{(i)} := (0, 0, \dots, 0, 1, 0, 0, \dots)_{(i)}$$

Affermo che

$$(5.23) \quad x = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_i e^{(i)}.$$

Infatti

$$D_n := \left\| x - \sum_{i=1}^n x_i e^{(i)} \right\|_\infty = \sup_{k \geq n} |x_k|$$

Grazie alla (5.22) si ha che D_n tende a zero quando $n \rightarrow \infty$ quindi vale la (5.23). Di conseguenza, dato che F è continuo posso scambiare il limite con $F(\cdot)$ e ottenere

$$F(x) = F\left(\lim_{n \rightarrow \infty} \sum_{i=1}^n x_i e^{(i)}\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_i F(e^{(i)})$$

Ponendo ora

$$a_i := F(e^{(i)})$$

otteniamo

$$F(x) = \sum_{i=1}^{\infty} a_i x_i = \varphi_a(x).$$

Resta da verificare che $a \in \ell_1$, vale a dire che

$$(5.24) \quad \sum_{i=1}^{\infty} |a_i| < \infty.$$

Dimostreremo un'affermazione più forte della (5.24), vale a dire che

$$\text{per ogni } n \text{ si ha } \sum_{i=1}^n |a_i| \leq \|F\|.$$

A questo scopo poniamo

$$x^{(n)} = \left(\frac{a_1}{|a_1|}, \frac{a_2}{|a_2|}, \dots, \frac{a_n}{|a_n|}, 0, 0, 0, \dots \right) = \sum_{i=1}^n \frac{a_i}{|a_i|} e^{(i)}$$

Ovviamente abbiamo che

$$x^{(n)} \in \ell_0 \quad \text{e} \quad \|x^{(n)}\|_\infty = 1$$

di conseguenza si ha

$$|F(x^{(n)})| \leq \|F\| \|x\|_\infty = \|F\|.$$

Ma, d'altra parte,

$$|F(x^{(n)})| = \left| F \left(\sum_{i=1}^n \frac{a_i}{|a_i|} e^{(i)} \right) \right| = \left| \sum_{i=1}^n \frac{a_i}{|a_i|} F(e^{(i)}) \right| = \left| \sum_{i=1}^n \frac{a_i}{|a_i|} a_i \right| = \sum_{i=1}^n |a_i|$$

per cui

$$\sum_{i=1}^n |a_i| \leq \|F\| \quad \forall n$$

che implica

$$\sum_{i=1}^{\infty} |a_i| \leq \|F\|.$$

Di conseguenza $a \in \ell_1$ e la proposizione è dimostrata \square

5.20 Problema. Per $a \in \mathbb{R}$, calcolare la norma di δ_a che agisce sullo spazio $(C_c^\infty(\mathbb{R}), \|\cdot\|_u)$.

Soluzione. Possiamo scrivere

$$|\delta_a(f)| = |f(a)| \leq \sup_{y \in \mathbb{R}} |f(y)| \leq \|f\|_u \quad \forall f \in C_c^\infty(\mathbb{R})$$

Questo ci dice che $\|\delta_a\| \leq 1$. Per dimostrare che $\|\delta_a\| = 1$ è quindi sufficiente trovare una funzione particolare $h \in C_c^\infty(\mathbb{R})$ tale che

$$|\delta_a(h)| = \|h\|_u$$

Sia dunque h una funzione C^∞ a supporto compatto il cui modulo abbia il massimo assoluto nel punto $x = a$, ad esempio

$$h(x) := \begin{cases} e^{-\frac{1}{1-(x-a)^2}} & \text{se } |x-a| \leq 1 \\ 0 & \text{se } |x-a| > 1 \end{cases}$$

È immediato verificare che il massimo di h si ha appunto per $x = a$ e vale $h(a) = 1/e$. Quindi

$$|\delta_a(h)| = h(a) = \sup_{x \in \mathbb{R}} |h(x)| = \|h\|_u.$$

Questo implica

$$\|\delta_a\| = \sup_{f \neq 0} \frac{|\delta_a(f)|}{\|f\|_u} \geq \frac{|\delta_a(h)|}{\|h\|_u} = 1$$

Quindi $\|\delta_a\| \geq 1$. Ma già sapevamo che $\|\delta_a\| \leq 1$, quindi deve essere $\|\delta_a\| = 1$ \square

5.21 Problema. Fare un esempio di un funzionale lineare non continuo nello spazio normato $(C_1(\mathbb{R}), \|\cdot\|_1)$.

5.2.3 Lo spazio duale di uno spazio di Hilbert

Uno spazio di Hilbert è un caso particolare di uno spazio di Banach in cui abbiamo, in più, un prodotto scalare. Se $(V, \langle \cdot, \cdot \rangle)$ è uno spazio di Hilbert possiamo dunque ancora parlare di spazio duale, come l'insieme di tutti i funzionali lineari continui su V . Nel caso hilbertiano però c'è un modo "naturale" per trovare alcuni funzionali lineari che (si dimostra facilmente) sono continui. Basta porre per un generico $v \in V$

$$(5.25) \quad \varphi_v(z) := \langle z, v \rangle \quad z \in V.$$

È facile vedere che φ_v è un funzionale lineare continuo.

5.22 Domanda. Tutti i funzionali lineari continui su V sono della forma (5.25)?

La risposta è affermativa. In altre parole il duale di uno spazio di Hilbert è isomorfo a se stesso $V^* \cong V$.

5.23 Teorema. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio di Hilbert reale e definiamo per ogni $v \in V$, la funzione $\varphi_v : V \rightarrow \mathbb{R}$ come

$$\varphi_v(z) := \langle z, v \rangle \quad \forall z \in V.$$

Allora

- (1) per ogni $v \in V$ si ha $\varphi_v \in V^*$ e, inoltre, $\|\varphi_v\| = \|v\|$
- (2) se $c_1, c_2 \in \mathbb{R}$ e $v_1, v_2 \in V$ allora $\varphi_{c_1 v_1 + c_2 v_2} = c_1 \varphi_{v_1} + c_2 \varphi_{v_2}$
- (3) se F è un funzionale lineare continuo su V , allora esiste $v \in V$ tale che $F = \varphi_v$.

Dimostrazione di (1). Faccio vedere innanzi tutto che φ_v è un funzionale lineare. Siano $w, z \in V$ e sia $c \in \mathbb{R}$. Allora, grazie alle proprietà del prodotto scalare, abbiamo

$$\begin{aligned} \varphi_v(w + z) &:= \langle w + z, v \rangle = \langle w, v \rangle + \langle z, v \rangle = \varphi_v(w) + \varphi_v(z) \\ \varphi_v(cw) &:= \langle cw, v \rangle = c \langle w, v \rangle = c \varphi_v(w) \end{aligned}$$

dunque φ_v è lineare. Passo ad occuparmi della sua norma. Dalla disuguaglianza di Cauchy-Schwarz ottengo

$$|\varphi_v(z)| = |\langle z, v \rangle| \leq \|z\| \|v\|,$$

che implica $\|\varphi_v\| \leq \|v\|$. D'altra parte scegliendo $z = v$ si ha

$$|\varphi_v(v)| = |\langle v, v \rangle| = \|v\|^2.$$

Di conseguenza $\|\varphi_v\| = \|v\|$.

Dimostrazione di (2). L'affermazione si dimostra immediatamente usando la bilinearità del prodotto scalare. Infatti, poiché V è uno spazio di Hilbert *reale*, possiamo scrivere

$$\begin{aligned} \varphi_{c_1 v_1 + c_2 v_2}(z) &= \langle z, c_1 v_1 + c_2 v_2 \rangle = c_1 \langle z, v_1 \rangle + c_2 \langle z, v_2 \rangle \\ &= c_1 \varphi_{v_1}(z) + c_2 \varphi_{v_2}(z) = (c_1 \varphi_{v_1} + c_2 \varphi_{v_2})(z). \end{aligned}$$

Dimostrazione di (3). Sia $F \in V^*$ e consideriamo il nucleo di F

$$N_F := \{u \in V : F(u) = 0\}.$$

Dato che F è continuo N_F è un sottospazio chiuso di V . Se $N_F = V$, allora si ha

$$F(u) = 0 \quad \forall u \in V,$$

vale a dire F è il funzionale nullo, $F = 0$. In questo caso scelgo $v = 0$ e ottengo

$$\langle v, u \rangle = \langle 0, u \rangle = 0 = F(u) \quad \forall u \in V,$$

per cui abbiamo $F = \varphi_v$ e l'affermazione (3), in questo caso banale, è dimostrata.

Consideriamo dunque l'altro caso in cui $N_F \neq V$. Per il teorema 4.25 sulle proiezioni ortogonali il complemento ortogonale N_F^\perp contiene almeno un elemento $w \neq 0$. Scegliamo ora

$$v = \frac{F(w)}{\|w\|^2} w$$

e dimostriamo che

$$F = \varphi_v \quad \text{vale a dire che} \quad F(u) = \varphi_v(u) = \langle v, u \rangle \quad \forall u \in V.$$

Sia quindi u un qualsiasi vettore in V . Siccome $w \notin N_F$, $F(w) \neq 0$, quindi u può essere scritto come

$$u = u_1 + \frac{F(u)}{F(w)} w \quad \text{in cui } u_1 := \left(u - \frac{F(u)}{F(w)} w \right)$$

Notare che

$$F(u_1) = F\left(u - \frac{F(u)}{F(w)} w\right) = F(u) - F(u) = 0,$$

per cui $u_1 \in N_F$. Poiché $w \in N_F^\perp$ avremo

$$\langle u_1, w \rangle = 0.$$

Di conseguenza

$$\begin{aligned} \varphi_v(u) &= \langle v, u_1 \rangle + \langle v, \frac{F(u)}{F(w)} w \rangle = \left\langle \frac{F(w)}{\|w\|^2} w, u_1 \right\rangle + \left\langle \frac{F(w)}{\|w\|^2} w, \frac{F(u)}{F(w)} w \right\rangle \\ &= \frac{F(w)}{\|w\|^2} \langle w, u_1 \rangle + \frac{F(w)}{\|w\|^2} \frac{F(u)}{F(w)} \langle w, w \rangle = 0 + F(u) = F(u) \end{aligned}$$

Abbiamo quindi mostrato che

$$F(u) = \varphi_v(u) \quad \forall u \in V$$

il che significa che $F = \varphi_v$ che è esattamente ciò che affermava il punto (3). \square

5.24 Osservazione. Se V è uno spazio di Hilbert *complesso* le conclusioni del Teorema 5.23 rimangono inalterate, ad eccezione del punto 2 che va modificato come segue:

$$\text{se } c_1, c_2 \in \mathbb{C} \text{ e } v_1, v_2 \in V \text{ allora } \varphi_{c_1 v_1 + c_2 v_2} = \bar{c}_1 \varphi_{v_1} + \bar{c}_2 \varphi_{v_2}.$$

5.3 Convergenza debole

5.25 Definizione. (Convergenza debole). Una successione $(v_k)_{k=1}^\infty$ in uno spazio vettoriale normato $(V, \|\cdot\|)$ si dice *debolmente convergente al vettore* $v \in V$ e si scrive $v_k \xrightarrow{w} v$ se

$$\text{per ogni } F \in V^* \text{ si ha } \lim_{k \rightarrow \infty} F(v_k) = F(v).$$

Il concetto di convergenza debole è più debole⁴ della convergenza usuale che si esprime tramite la norma come

$$\lim_{k \rightarrow \infty} \|v - v_k\| = 0,$$

nel senso che tutte le successioni convergenti sono anche debolmente convergenti,⁵ mentre non vale (in generale) il viceversa. Considerate infatti il seguente esempio.

5.26 Esempio. Nello spazio $(\ell_2, \|\cdot\|_2)$ consideriamo la successione

$$(5.26) \quad e^{(1)}, e^{(2)}, e^{(3)}, \dots$$

in cui

$$e^{(i)} := (0, 0, \dots, 0, \underset{[i]}{1}, 0, 0, \dots)$$

Questa successione è chiaramente non convergente in norma. Infatti si ha

$$\|e^{(i)} - e^{(j)}\| = \sqrt{2} \quad \forall i \neq j$$

che implica che la successione (5.26) non può essere di Cauchy. Di conseguenza non è neanche convergente. Però la (5.26) è una successione debolmente convergente a zero. Per dimostrarlo bisogna far vedere che

$$\text{per ogni } F \in \ell_2^* \text{ si ha } \lim_{n \rightarrow \infty} F(e^{(n)}) = 0$$

Ma dato che ℓ_2 è uno spazio di Hilbert sappiamo che tutti i funzionali lineari continui F su ℓ_2 si possono rappresentare come $F = \varphi_a = \langle \cdot, a \rangle$ per un qualche $a \in \ell_2$. Quindi quello che dobbiamo dimostrare è

$$\text{per ogni } a \in \ell_2^* \text{ si ha } \lim_{n \rightarrow \infty} \langle e^{(n)}, a \rangle = 0.$$

Ma, d'altra parte, se $a \in \ell_2$ si ha

$$\langle e^{(n)}, a \rangle = \sum_{i=1}^{\infty} a_i e_i^{(n)} = a_n$$

e, dato che $a \in \ell_2$, abbiamo

$$\sum_{i=1}^{\infty} |a_i|^2 < \infty,$$

per cui

$$\lim_{n \rightarrow \infty} a_n = 0.$$

Abbiamo quindi fatto vedere che

$$\lim_{n \rightarrow \infty} \langle e^{(n)}, a \rangle = \lim_{n \rightarrow \infty} a_n = 0 \quad \forall a \in \ell_2$$

dunque $e^{(n)} \xrightarrow{w} 0$.

⁴ma guarda un po'

⁵perché?

5.27 Esempio. Se f_n converge debolmente a f in $(C_b(\mathbb{R}), \|\cdot\|_u)$ allora f_n converge puntualmente a f ,⁶ vale a dire

$$(5.27) \quad \text{per ogni } x \in \mathbb{R} \text{ si ha } \lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Infatti supponiamo $f_n \xrightarrow{w} f$. Ciò significa, ricordiamo, che

$$\text{per ogni } F \in C_n(\mathbb{R})^* \text{ si ha } \lim_{n \rightarrow \infty} F(f_n) = F(f).$$

In particolare, dato che le delta di Dirac sono funzionali lineari continui su $C_b(\mathbb{R})$ dovrà valere

$$\text{per ogni } x \in \mathbb{R} \text{ si ha } \lim_{n \rightarrow \infty} \delta_x(f_n) = \delta_x(f).$$

che, per definizione di δ_x , è equivalente alla (5.27) la quale afferma appunto che f_n converge a f puntualmente.

5.28 Problema. Trovare una successione in $(\ell_0, \|\cdot\|_\infty)$ debolmente convergente ma non convergente. (Sugg: sfruttare il fatto che tutti i funzionali lineari continui su ℓ_0 sono della forma ...)

⁶ricorda che invece la convergenza in norma $\|\cdot\|_u$ è equivalente alla convergenza uniforme

6. Distribuzioni

In questo capitolo vogliamo introdurre degli oggetti chiamati *distribuzioni* che in un qualche senso generalizzano il concetto di funzione ordinaria.¹ L'idea (vaga) è quella di costruire uno spazio molto più grande dello spazio delle funzioni, sul quale sia ancora possibile eseguire (quasi) tutte le operazioni che si eseguono sulle funzioni. Questo spazio include, ad esempio, tutte le funzioni continue $g \in C(\mathbb{R}; \mathbb{C})$ o, più precisamente, è possibile associare ad ogni funzione continua g una distribuzione che indichiamo con φ_g . Una delle novità è che indipendentemente dal fatto che g sia derivabile o meno, la distribuzione associata è sempre infinitamente derivabile! Potremo definire, ad esempio, $D^n|x|$ in cui n è un intero positivo arbitrario, a patto di interpretare la derivata *nel senso delle distribuzioni*.

Si pensi lo spazio delle funzioni C^{k-1} come l'insieme delle funzioni ottenute applicando l'operatore derivata allo spazio C^k

$$D : C^k \longrightarrow C^{k-1} .$$

L'idea è di continuare ad operare con D anche quando k è uguale a zero, ottenendo in questo modo spazi C^k in cui k diventa negativo! In un qualche senso le distribuzioni costituiscono l'unione di tutti questi spazi²

$$\{\text{distribuzioni}\} = \bigcup_{k=0}^{\infty} C^{-k}(\mathbb{R}) .$$

Il passo fondamentale per definire matematicamente le distribuzioni sarà quello di pensare alle funzioni ordinarie in modo diverso, vale a dire come a *funzionali lineari* che agiscono su un opportuno spazio di funzioni. La distribuzione associata alla funzione g è il funzionale lineare φ_g , definito come

$$(6.1) \quad \varphi_g(f) = \int_{\mathbb{R}} g(x) f(x) dx .$$

Il secondo passo sarà rendersi conto che a questo punto è possibile arricchire notevolmente l'insieme delle distribuzioni con funzionali lineari che *non* possono essere espressi nella forma (6.1), come ad esempio la delta di Dirac, definita semplicemente come

$$\delta_0(f) = f(0) .$$

Iniziamo a questo punto a capire quali ipotesi è necessario fare su g e su f affinché la (6.1) abbia senso.

6.1 Notazione. In questo capitolo le funzioni saranno a valori complessi a meno che non sia esplicitamente detto il contrario, quindi, ad esempio, con $C(\mathbb{R})$ intendiamo $C(\mathbb{R}; \mathbb{C})$ (analogamente per altri spazi di funzioni su \mathbb{R}).

¹infatti le distribuzioni vengono anche chiamate *funzioni generalizzate*.

²la seguente identità va interpretata "fra virgolette".

6.1 Funzioni continue o differenziabili a tratti

6.2 Definizione. Una funzione $f : \mathbb{R} \rightarrow \mathbb{C}$ è detta avere *discontinuità isolate* se l'insieme

$$D_f := \{x \in \mathbb{R} : f \text{ non è continua in } x\}$$

non ha punti di accumulazione.

In altre parole una funzione f ha discontinuità isolate se e solo se, comunque si scelga un intervallo finito $[a, b]$, il numero dei punti di discontinuità di f nell'intervallo $[a, b]$ è finito. Si osservi che l'insieme D_f di tutti i punti di discontinuità è un insieme finito o numerabile. Infatti

$$D_f = \cup_{n \in \mathbb{Z}} [D_f \cap [n, n+1)].$$

Poiché ciascuno degli insiemi $D_f \cap [n, n+1)$ ha cardinalità finita, D_f essendo unione numerabile di insiemi di cardinalità finita è numerabile.

6.3 Definizione. Siano $a, b \in \mathbb{R}$ con $a < b$. Una funzione $f : [a, b] \rightarrow \mathbb{C}$ si dice *continua a tratti* (CAT) se

- (a) f ha discontinuità isolate
- (b) per ogni punto di discontinuità $u \in D_f$, con $u \in (a, b)$, esistono e sono finiti i limiti

$$(6.2) \quad f(u^-) := \lim_{x \rightarrow u^-} f(x) \quad f(u^+) := \lim_{x \rightarrow u^+} f(x)$$

- (c) esistono e sono finiti i limiti

$$f(a^+) := \lim_{x \rightarrow a^+} f(x) \quad f(b^-) := \lim_{x \rightarrow b^-} f(x).$$

La funzione $f : \mathbb{R} \rightarrow \mathbb{C}$ è detta *continua a tratti* se valgono le proprietà (a) e (b).

6.4 Definizione. Una funzione $f : [a, b] \rightarrow \mathbb{C}$ è detta *differenziabile a tratti* (DAT) se f ed f' sono continue a tratti in $[a, b]$.

In altre parole f è differenziabile a tratti in $[a, b]$ se esiste un insieme finito di punti $\alpha_i \in [a, b]$, $i = 0, \dots, n$ tali che

- (a) $a = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_n = b$.
- (b) f è differenziabile con derivata continua in ciascun intervallo aperto (α_{i-1}, α_i) .
- (c) Esistono e sono finiti i limiti

$$\begin{array}{ccccccc} \lim_{x \rightarrow \alpha_0^+} f(x), & \lim_{x \rightarrow \alpha_1^-} f(x), & \lim_{x \rightarrow \alpha_1^+} f(x), & \dots, & \lim_{x \rightarrow \alpha_n^-} f(x) \\ \lim_{x \rightarrow \alpha_0^+} f'(x), & \lim_{x \rightarrow \alpha_1^-} f'(x), & \lim_{x \rightarrow \alpha_1^+} f'(x), & \dots, & \lim_{x \rightarrow \alpha_n^-} f'(x) \end{array}$$

6.5 Definizione. Una funzione $f : \mathbb{R} \rightarrow \mathbb{C}$ è detta *differenziabile a tratti* (DAT) se la restrizione di f ad un qualsiasi intervallo chiuso e limitato $[a, b]$ è differenziabile a tratti.

6.6 Esempio. (Una funzione DAT). Consideriamo la funzione

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ x^2 & \text{se } 0 \leq x < 1 \\ 3 - x & \text{se } 1 \leq x \leq 2 \\ 1 & \text{se } x > 2 \end{cases}$$

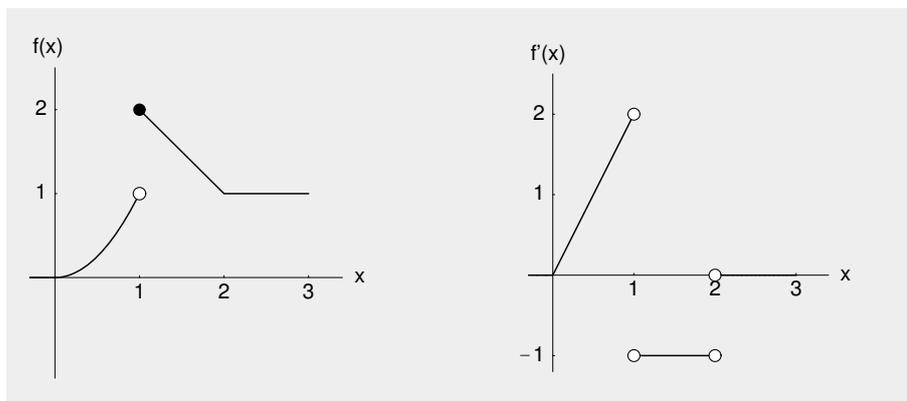


Figura 6.1: Esempio di funzione DAT

La funzione f ha un solo punto di discontinuità in $x = 1$. In questo punto abbiamo

$$f(1^-) = 1 \quad f(1^+) = f(1) = 2,$$

Quindi f è continua a tratti. Per quanto riguarda la derivata di f osserviamo subito che ci sono due punti, $x = 1$ e $x = 2$ in cui la derivata *non esiste*. Nonostante ciò la derivata ha limite sia destro che sinistro finito in ciascun punto di discontinuità

$$f'(1^-) = 2 \quad f'(1^+) = -1 \quad f'(2^-) = -1 \quad f'(2^+) = 0,$$

quindi f' è continua a tratti e, di conseguenza, f è differenziabile a tratti.

6.7 Esempio. (Una funzione CAT ma non DAT). Consideriamo adesso la funzione

$$f(x) = \sqrt{|x|}.$$

Questa funzione è addirittura continua, quindi, a maggior ragione continua a tratti, ma la sua derivata, per $x \neq 0$, è data da

$$f'(x) = \frac{\operatorname{sgn} x}{2\sqrt{|x|}}.$$

Sia il limite destro che quello sinistro di $f'(x)$ sono infiniti, quando x tende a zero, quindi f' non è continua a tratti, di conseguenza f non è differenziabile a tratti.

Integrazione per parti di funzioni continue con derivata CAT

È nota la formula di integrazione per parti

$$(6.3) \quad \int_a^b f(x) g(x) dx = [fG]_a^b - \int_a^b f'(x) G(x) dx$$

in cui G è una primitiva di g

$$G(x) = \int_a^x g(u) du + C$$

e $[fG]_a^b = f(b)G(b) - f(a)G(a)$ è l'incremento del prodotto fG lungo l'intervallo $[a, b]$. La formula (6.3) viene usualmente dimostrata sotto le ipotesi $f \in C^1[a, b]$ e $g \in C[a, b]$.

Vogliamo far vedere che la formula di integrazione per parti è valida sotto le seguenti ipotesi più generali:

- (1) f continua con derivata CAT
 (2) g CAT.

In altre parole la funzione f che volete derivare deve essere continua, ma la sua derivata basta che sia continua a tratti, mentre la funzione g “da integrare” è sufficiente che sia continua a tratti. Per dimostrarlo iniziamo a dare un risultato più generale.

6.8 Proposizione. *Siano F e G due funzioni reali, definite sull'intervallo $[a, b]$ differenziabili a tratti. Allora vale la seguente identità*

$$(6.4) \quad [FG]_{a^+}^{b^-} = \int_a^b [F'(x)G(x) + F(x)G'(x)] dx + \sum_{x \in (a,b)} [F(x^-)\Delta G(x) + G(x^-)\Delta F(x) + \Delta F(x)\Delta G(x)],$$

in cui la sommatoria è estesa a tutti i punti di discontinuità di F e di G e $\Delta F(x) := F(x^+) - F(x^-)$ rappresenta il valore del salto di F nel punto x (analogamente per G).

Osserviamo che la prima riga della (6.4) è l'usuale formula di integrazione per parti, a parte il fatto che, poiché F e G non sono continue, nel fattore che tiene conto dell'incremento del prodotto FG è necessario specificare che i valori di FG vanno calcolati dalla “parte interna” dell'intervallo $[a, b]$, per cui in a^+ e in b^- . La seconda riga invece tiene conto del fatto che F e G possono avere dei salti all'interno dell'intervallo (a, b) e questi salti vanno “conteggiati” secondo la (6.4).

Nel caso in cui F e G siano entrambe continue (con derivata CAT) tutti i termini ΔF e ΔG sono nulli, di conseguenza ritroviamo la formula usuale. Possiamo a questo punto dimostrare il risultato che abbiamo annunciato in precedenza.

6.9 Corollario. *Siano f, g due funzioni reali, definite sull'intervallo $[a, b]$. Se f è continua con derivata CAT e g è continua a tratti, allora vale la formula di integrazione per parti “classica” (6.3).*

Dimostrazione. Poniamo $F = f$ e sia inoltre G una primitiva di g . La Proposizione 6.8 ci dice che vale l'uguaglianza (6.4). Poiché la primitiva di una funzione CAT è continua, si ha che F e G sono entrambe continue, di conseguenza tutti i termini di salto contenenti ΔF o ΔG sono nulli. Otteniamo quindi

$$(6.5) \quad [FG]_a^b = \int_a^b [F'(x)G(x) + F(x)G'(x)] dx,$$

che, sostituendo $F = f$, diventa proprio la (6.3). □

Rimane quindi da dimostrare la (6.4).

Dimostrazione della Proposizione 6.8 Se $H : [a, b] \rightarrow \mathbb{C}$ è una funzione DAT allora il teorema fondamentale del calcolo integrale va riscritto sotto la forma

$$H(b^-) - H(a^+) = \int_a^b H'(x) dx + \sum_{x \in (a,b)} \Delta H(x),$$

in cui la sommatoria tiene conto dei salti di H . Applicando questa identità alla funzione $H = FG$, otteniamo

$$[FG]_{a^+}^{b^-} = \int_a^b [F'(x)G(x) + F(x)G'(x)] dx + \sum_{x \in (a,b)} \Delta(FG)(x).$$

A questo punto scriviamo il salto $\Delta(FG)(x)$ nel modo seguente

$$\begin{aligned}\Delta(FG)(x) &= F(x^+)G(x^+) - F(x^-)G(x^-) \\ &= F(x^+)G(x^+) - F(x^+)G(x^-) + F(x^+)G(x^-) - F(x^-)G(x^-) \\ &= F(x^+)\Delta G(x) + \Delta F(x)G(x^-) \\ &= [F(x^-) + \Delta F(x)]\Delta G(x) + \Delta F(x)G(x^-) \\ &= F(x^-)\Delta G(x) + \Delta F(x)G(x^-) + \Delta F(x)\Delta G(x).\end{aligned}$$

La proposizione è così dimostrata. \square

6.2 Lo spazio delle distribuzioni

6.10 Definizione. Ricordiamo che con $C_c^\infty(\mathbb{R}; \mathbb{C})$ denotiamo lo spazio di tutte le funzioni $f : \mathbb{R} \rightarrow \mathbb{C}$ tali che

- (1) f ha supporto compatto, cioè esiste un intervallo finito al di fuori del quale f è nulla
- (2) f ha derivate continue di tutti gli ordini.

Nel seguito ometteremo (a volte) di enfatizzare il fatto che la funzione f è, a priori, a valori complessi, scrivendo semplicemente $f \in C_c^\infty(\mathbb{R})$, invece che $f \in C_c^\infty(\mathbb{R}; \mathbb{C})$. Lo spazio $C_c^\infty(\mathbb{R})$ è anche detto spazio delle *funzioni fondamentali*. Questo spazio, per brevità viene anche indicato con il simbolo \mathcal{K} . Se $f \in C_c^\infty(\mathbb{R})$ indichiamo, come al solito, con $f^{(k)}$ la derivata k -sima di f . Vogliamo ora introdurre nello spazio \mathcal{K} una nozione di convergenza particolare.

6.11 Definizione. La successione $(f_n)_{n=1}^\infty$ di elementi di \mathcal{K} si dice *convergente ad* $f \in \mathcal{K}$ e si scrive $f_n \xrightarrow{\mathcal{K}} f$ se

- (a) esiste un intervallo I al di fuori del quale tutte le f_n sono nulle
- (b) per ogni intero non negativo k si ha³

$$f_n^{(k)} \rightarrow f^{(k)} \text{ uniformemente su } I.$$

6.12 Definizione. Una *distribuzione su* \mathbb{R} è un funzionale lineare continuo $F : \mathcal{K} \rightarrow \mathbb{C}$, vale a dire un funzionale lineare tale che

$$(6.6) \quad \text{se } f_n \xrightarrow{\mathcal{K}} f \text{ allora } F(f_n) \rightarrow F(f).$$

L'insieme di tutte le distribuzioni si indica con \mathcal{K}^* .⁴

Le distribuzioni generalizzano in un certo senso il concetto di funzione. Come vedremo ad ogni funzione “buona” è sempre possibile associare una distribuzione, mentre esistono distribuzioni che non “provengono” da nessuna funzione. L’idea è quella di costruire degli oggetti più generali delle funzioni, ma tali che su di essi si possa operare con molte delle “operazioni” permesse per le funzioni ordinarie, come somma, prodotto per un numero, derivata ecc. Iniziamo a far vedere che ad ogni funzione che soddisfa una condizione chiamata “integrabilità locale” corrisponde una distribuzione.

³per definizione poniamo $f^{(0)} := f$

⁴spesso vengono usati, al posto di \mathcal{K} e \mathcal{K}^* , i simboli \mathcal{D} e \mathcal{D}'

6.13 Definizione. Una funzione $f : \mathbb{R} \rightarrow \mathbb{C}$ si dice *localmente integrabile* se:

- (a) f ha discontinuità isolate.
- (b) per ogni $a, b \in \mathbb{R}$ con $a < b$, si ha che l'integrale (possibilmente improprio) $\int_a^b |f(x)| dx$ è finito.

6.14 Esempi. Vediamo alcuni esempi di funzioni localmente integrabili.

- (1) Ogni funzione continua f è localmente integrabile. Infatti, se f è continua anche $|f|$ è continua, quindi l'integrale di $|f|$ su ogni intervallo chiuso e limitato $[a, b]$ è finito.
- (2) Sia $\alpha \in \mathbb{R}$. Allora la funzione $1/|x|^\alpha$ è localmente integrabile se e solo se $\alpha < 1$. Infatti, se $b > 0$,

$$\int_0^b \frac{dx}{x^\alpha} = \lim_{\varepsilon \rightarrow 0^+} \int_\varepsilon^b \frac{dx}{x^\alpha}.$$

L'ultimo integrale si calcola facilmente (considerando separatamente il caso $\alpha = 1$) e si trova che il limite $\varepsilon \rightarrow 0^+$ è finito se e solo se $\alpha < 1$.

- (3) La funzione $f(x) := (\log |x|)^n$ è localmente integrabile per ogni intero n . Si tratta infatti di una funzione continua su tutto l'asse reale ad eccezione del punto $x = 0$. Ci dobbiamo quindi preoccupare soltanto di quanto f diverge velocemente per $x \rightarrow 0$. Facendo il confronto asintotico con la funzione $g(x) = 1/\sqrt{|x|}$, si ottiene

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0} \sqrt{|x|} (\log |x|)^n = 0,$$

vale a dire

$$(\log |x|)^n < \frac{1}{\sqrt{|x|}} \quad \text{asintoticamente per } x \rightarrow 0.$$

Per quanto detto in precedenza, la funzione $1/\sqrt{|x|}$ è localmente integrabile e dunque anche $(\log |x|)^n$ è localmente integrabile.

La proposizione seguente afferma che una classe di funzioni che sarà particolarmente rilevante nello studio delle distribuzioni è composta di funzioni localmente integrabili.

6.15 Proposizione. Se $f : \mathbb{R} \rightarrow \mathbb{C}$ è continua a tratti allora è localmente integrabile.

Dimostrazione. Sia f continua a tratti e siano $a, b \in \mathbb{R}$ con $a < b$. Esistono quindi $u_0, \dots, u_n \in \mathbb{R}$, con

$$a = u_0 < u_1 < \dots < u_{n-1} < u_n = b,$$

tali che f è continua in ciascun intervallo (u_{i-1}, u_i) , con limite destro e sinistro finito rispettivamente in u_{i-1} e u_i . Per ogni $i = 1, \dots, n$ definiamo una funzione $\tilde{f}_i : [u_{i-1}, u_i] \rightarrow \mathbb{C}$ come

$$\tilde{f}_i(x) := \begin{cases} f(x) & \text{se } x \in (u_{i-1}, u_i) \\ f(u_{i-1}^+) & \text{se } x = u_{i-1} \\ f(u_i^-) & \text{se } x = u_i. \end{cases}$$

La funzione \tilde{f}_i è continua per costruzione su l'intervallo chiuso $[u_{i-1}, u_i]$. Di conseguenza anche $|\tilde{f}_i|$ è continua sullo stesso intervallo. Si ottiene in questo modo

$$\int_a^b |f(x)| dx = \sum_{i=1}^n \int_{u_{i-1}}^{u_i} |f(x)| dx = \sum_{i=1}^n \int_{u_{i-1}}^{u_i} |\tilde{f}_i(x)| dx.$$

La quantità $\int_a^b |f(x)| dx$ è dunque finita, in quanto somma finita di integrali di funzioni continue su intervalli chiusi e limitati. Di conseguenza f è localmente integrabile. \square

6.16 Proposizione. *Sia $g : \mathbb{R} \rightarrow \mathbb{C}$ localmente integrabile e sia $\varphi_g : \mathcal{K} \rightarrow \mathbb{C}$ definita come*

$$(6.7) \quad \varphi_g(f) := \int_{\mathbb{R}} g(x)f(x) dx \quad f \in \mathcal{K}.$$

Allora φ_g è una distribuzione.

Dimostrazione.

❶ L'applicazione φ_g è ben definita sullo spazio \mathcal{K} .

Infatti se $f \in \mathcal{K}$ allora esiste $a > 0$ tale che f è nulla al di fuori dell'intervallo $[-a, a]$. Siccome f è continua esiste M tale che $|f(x)| \leq M$ per ogni $x \in [-a, a]$. Quindi

$$\int_{\mathbb{R}} |g(x)f(x)| dx = \int_{-a}^a |g(x)f(x)| dx \leq M \int_{-a}^a |g(x)| dx.$$

Poiché g è localmente integrabile l'ultimo integrale è finito. Abbiamo così mostrato che l'integrale che definisce $\varphi_g(f)$ è assolutamente convergente.

❷ L'applicazione φ_g è un funzionale lineare.

Infatti, se $\alpha, \beta \in \mathbb{R}$ e $f, h \in \mathcal{K}$, si ha

$$\varphi_g(\alpha f + \beta h) = \int_{\mathbb{R}} g(x) [\alpha f(x) + \beta h(x)] dx = \alpha \varphi_g(f) + \beta \varphi_g(h).$$

❸ L'applicazione φ_g è continua.

Devo far vedere che vale l'implicazione (6.6). Assumo quindi che $f_n \xrightarrow{\mathcal{K}} f$. Questo implica che

- (a) esiste $a > 0$ tale che tutte le funzioni f_n si annullano al di fuori dell'intervallo $[-a, a]$;
- (b) la successione (f_n) converge a f uniformemente, cioè $\lim_{n \rightarrow \infty} \|f - f_n\|_u = 0$.

Poiché g è localmente integrabile si ha

$$A := \int_{-a}^a |g(x)| dx < \infty.$$

Di conseguenza ottengo

$$\begin{aligned} |\varphi_g(f) - \varphi_g(f_n)| &= \left| \int_{-a}^a g(x) [f(x) - f_n(x)] dx \right| \\ &\leq \|f - f_n\|_u \int_{-a}^a |g(x)| dx = A \|f - f_n\|_u. \end{aligned}$$

Quindi, grazie alla proprietà (b) citata in precedenza,

$$\lim_{n \rightarrow \infty} |\varphi_g(f) - \varphi_g(f_n)| = 0.$$

Abbiamo dunque dimostrato che se $f_n \xrightarrow{\mathcal{K}} f$, allora si ha $\varphi_g(f_n) \rightarrow \varphi_g(f)$, vale a dire φ_g è continua. \square

6.17 Definizione. La distribuzione $F \in \mathcal{K}^*$ è detta *regolare* se esiste una funzione $g : \mathbb{R} \rightarrow \mathbb{C}$ localmente integrabile tale che $F = \varphi_g$.

6.18 Esempi. Vediamo alcuni esempi di funzioni localmente integrabili alle quali è possibile associare una distribuzione.

- (1) Se g è una funzione *continua* allora è automaticamente localmente integrabile e quindi φ_g è una distribuzione. Notare che g può divergere con velocità arbitraria quando $x \rightarrow \infty$.⁵ Ad esempio $g(x) = e^{x^{100}}$ dà luogo ad una buona distribuzione φ_g su \mathcal{K} .
- (2) Analogamente, se g è una funzione continua a tratti, φ_g è una distribuzione. Esempio: $g(x) := \lfloor x \rfloor$ (la “parte intera” di x). Un altro esempio semplice ma importante è quello della funzione “a gradino” di Heaviside

$$(6.8) \quad H(x) := \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

H è chiaramente una funzione localmente integrabile, alla quale possiamo associare la distribuzione *theta*

$$(6.9) \quad \vartheta(f) := \varphi_H(f) := \int_{\mathbb{R}} H(x)f(x) dx = \int_0^{\infty} f(x) dx \quad f \in \mathcal{K}.$$

- (3) Se invece g ha dei punti y_1, y_2, \dots in cui diverge allora g è localmente integrabile se ciascuna di queste divergenze è assolutamente integrabile. Ad esempio, come già detto $\frac{1}{\sqrt{|x|}}$ e $\log|x|$ danno luogo a distribuzioni, ma $1/x, 1/x^2, 1/x^3, \dots$ no.⁶

6.19 Osservazione. Siano g e h due funzioni localmente integrabili che differiscono in un numero finito di punti u_1, u_2, \dots, u_n . Allora g e h producono la stessa distribuzione, vale a dire $\varphi_g = \varphi_h$ in quanto, se $f \in \mathcal{K}$, si ha

$$\begin{aligned} \varphi_g(f) &= \int_{\mathbb{R}} g(x)f(x) dx = \\ &= \int_{-\infty}^{u_1} g(x)f(x) dx + \int_{u_1}^{u_2} g(x)f(x) dx + \dots + \int_{u_n}^{\infty} g(x)f(x) dx \\ &= \int_{-\infty}^{u_1} h(x)f(x) dx + \int_{u_1}^{u_2} h(x)f(x) dx + \dots + \int_{u_n}^{\infty} h(x)f(x) dx \\ &= \int_{\mathbb{R}} h(x)f(x) dx = \varphi_h(f). \end{aligned}$$

Sia ad esempio

$$\tilde{H}(x) := \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x \leq 0. \end{cases}$$

Le funzioni H e \tilde{H} differiscono solo nel punto 0, quindi possiamo affermare che sono *identiche come distribuzioni*, vale a dire $\varphi_H = \varphi_{\tilde{H}} = \vartheta$.

⁵ovviamente questo dipende da come abbiamo scelto lo spazio \mathcal{K} delle funzioni di prova.

⁶in seguito verrà definita una distribuzione $P(1/x^n)$ associata alla funzione $1/x$ con un procedimento diverso.

6.3 La delta di Dirac

Se ad ogni funzione localmente integrabile g corrisponde una distribuzione φ_g , l'insieme delle distribuzioni è però molto più ampio. Un esempio fondamentale di una distribuzione che non può essere scritta come φ_g è costituito dalla delta di Dirac. Se $a \in \mathbb{R}$, la delta di Dirac nel punto a è definita come

$$(6.10) \quad \delta_a(f) := f(a) \quad f \in \mathcal{K}.$$

È banale far vedere che si tratta di un funzionale lineare continuo su \mathcal{K} , quindi una distribuzione. Voglio far vedere che non può essere espressa nella forma (6.7). Supponiamo per assurdo che esista una funzione localmente integrabile, che indichiamo con il simbolo $\delta(x)$, tale che valga (scelgo ora $a = 0$)

$$(6.11) \quad \delta_0(f) = \int_{\mathbb{R}} \delta(x)f(x) dx.$$

Allora affermo che

6.20. La funzione $\delta(x)$ può essere diversa da zero solo nell'origine o in un punto di discontinuità.

Se così non fosse, infatti, sia $b \in \mathbb{R}$, $b \neq 0$, tale che $\delta(x)$ è continua in b e $\delta(b) = A > 0$ (il caso in cui $\delta(b) < 0$ è analogo). Vediamo che questa ipotesi conduce ad un assurdo. Poiché $\delta(x)$ è continua in b , esiste $\varepsilon > 0$ tale che $\delta(x) \geq A/2$ per ogni $x \in [b - \varepsilon, b + \varepsilon]$. Possiamo sempre supporre $\varepsilon < b$. Sia $f \in \mathcal{K}$ una funzione non negativa tale che $f(b) = 1$ e $\text{supp } f \subset [b - \varepsilon, b + \varepsilon]$. Poiché f è continua esiste $\varepsilon' \in (0, \varepsilon)$ tale che $f(x) \geq 1/2$ per ogni $x \in [b - \varepsilon', b + \varepsilon']$. Otteniamo quindi

$$\delta_0(f) = \int_{\mathbb{R}} \delta(x)f(x) dx = \int_{-\varepsilon}^{\varepsilon} \delta(x)f(x) dx \geq \int_{-\varepsilon'}^{\varepsilon'} \delta(x)f(x) dx \geq \int_{-\varepsilon'}^{\varepsilon'} \frac{A}{2} \frac{1}{2} dx = \frac{\varepsilon' A}{2}.$$

D'altra parte, poiché $\text{supp } f \subset [b - \varepsilon, b + \varepsilon]$ con $b - \varepsilon > 0$, abbiamo $f(0) = 0$, quindi $\delta_0(f) = 0$, giungendo così ad un assurdo. La 6.20 è dimostrata.

Faccio ora vedere che la 6.20 è a sua volta in contraddizione con la definizione della delta di Dirac (6.10). Infatti, se $f \in \mathcal{K}$, esiste un intervallo chiuso e limitato $[a, b]$ tale che $\text{supp } f \subset [a, b]$. Grazie alla 6.20 il prodotto $\delta(x)f(x)$ è sempre nullo in $[a, b]$ ad eccezione, al più, di un numero finito di punti. Quindi

$$\int_{\mathbb{R}} \delta(x)f(x) dx = \int_a^b \delta(x)f(x) dx = 0 \quad \forall f \in \mathcal{K},$$

quindi $\delta_0(f) = 0$ per ogni $f \in \mathcal{K}$ e dunque δ_0 è il funzionale nullo, in contraddizione con la (6.10). Abbiamo dimostrato che la delta di Dirac non può essere scritta nella forma (6.11).

Nonostante quello che abbiamo appena mostrato, è *prassi comune usare per la delta di Dirac la stessa notazione che si usa per le distribuzioni associate a funzioni localmente integrabili*, per cui l'identità $\delta_a(f) = f(a)$ viene solitamente (e impropriamente) scritta come

$$(6.12) \quad \int_{-\infty}^{\infty} f(x) \delta(x - a) dx = f(a),$$

in cui la "funzione inesistente" $\delta(x)$ viene "pensata" come soddisfacente le condizioni

$$(6.13) \quad \delta(x) = \begin{cases} +\infty & \text{se } x = 0 \\ 0 & \text{se } x \neq 0 \end{cases} \quad \int_{\mathbb{R}} \delta(x) dx = 1.$$

L'utilizzo della notazione (6.12) è innocuo (e addirittura, a volte, comodo) se fatto con la dovuta cautela, tenendo sempre presente che la funzione $\delta(x - c)$ non esiste ed evitando quindi di farsi prendere la mano e scrivere espressioni come

$$\int_{-\infty}^{\infty} f(x) (\delta(x - a))^2 dx$$

che non hanno alcun significato.

Le condizioni (6.13), pur non avendo significato matematico, suggeriscono un metodo per ottenere la delta di Dirac come *limite* di distribuzioni associate a funzioni localmente integrabili g_n che soddisfano, nel limite $n \rightarrow \infty$, le (6.13), vale a dire

$$\lim_{n \rightarrow \infty} g_n(x) = \begin{cases} \infty & \text{se } x = 0 \\ 0 & \text{se } x \neq 0 \end{cases} \quad \int_{\mathbb{R}} g_n(x) dx = 1.$$

Più precisamente si ha che

6.21 Proposizione. *Sia $g_n : \mathbb{R} \rightarrow \mathbb{R}$ una funzione non negativa, continua (a tratti) tale che*

$$(a) \int_{\mathbb{R}} g_n(x) dx = 1$$

(b) g_n è nulla al di fuori dell'intervallo $[-a_n, a_n]$, in cui (a_n) è una successione di numeri positivi che tende a zero.

Allora

$$(6.14) \quad \lim_{n \rightarrow \infty} \varphi_{g_n}(f) := \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) f(x) dx = \delta_0(f) \quad \forall f \in \mathcal{K}.$$

Dimostrazione. Dato che $\int g_n = 1$,

$$\begin{aligned} \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \delta_0(f) \right| &= \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \int_{\mathbb{R}} g_n(x) f(0) dx \right| \\ &= \left| \int_{-a_n}^{a_n} g_n(x) [f(x) - f(0)] dx \right| \\ &\leq \int_{-a_n}^{a_n} g_n(x) |f(x) - f(0)| dx \\ &\leq \sup_{|x| \leq a_n} |f(x) - f(0)| \int_{\mathbb{R}} g_n(x) dx = \sup_{|x| \leq a_n} |f(x) - f(0)| \end{aligned}$$

Ma se $f \in \mathcal{K}$, f è continua in 0, quindi, poiché $a_n \rightarrow 0$ si ha

$$\lim_{n \rightarrow \infty} \sup_{|x| \leq a_n} |f(x) - f(0)| = 0,$$

Il che dimostra la (6.14) □

Nella precedente proposizione abbiamo visto come la delta di Dirac si possa approssimare con funzioni non negative g_n il cui integrale è uguale a uno e il cui supporto tende a concentrarsi nell'origine. È anche possibile approssimare la delta con funzioni a supporto illimitato. Una ricetta utile è la seguente.

6.22 Proposizione. Sia $g \in C_1(\mathbb{R})$ non negativa, tale che $\int_{\mathbb{R}} g = 1$. Poniamo

$$g_n(x) := n g(nx) \quad x \in \mathbb{R}$$

Allora

$$(6.15) \quad \lim_{n \rightarrow \infty} \varphi_{g_n}(f) := \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) f(x) dx = \delta_0(f) \quad \forall f \in \mathcal{K}.$$

Dimostrazione. Notiamo innanzi tutto che ciascuna g_n ha integrale uguale ad 1. Infatti, ponendo $y = nx$, si ha

$$\int_{\mathbb{R}} g_n(x) dx = \int_{\mathbb{R}} n g(nx) dx = \int_{\mathbb{R}} g(y) dy = 1.$$

Possiamo quindi procedere nel modo seguente:

$$(6.16) \quad \begin{aligned} \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \delta_0(f) \right| &= \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \int_{\mathbb{R}} g_n(x) f(0) dx \right| \\ &= \left| \int_{\mathbb{R}} n g(nx) [f(x) - f(0)] dx \right| \\ &\leq \int_{\mathbb{R}} n g(nx) |f(x) - f(0)| dx \\ &\leq \int_{\mathbb{R}} g(y) |f(y/n) - f(0)| dy \end{aligned}$$

Dato che $g \geq 0$ e $\int_{\mathbb{R}} g = 1$, fissato un $\varepsilon > 0$ arbitrario è possibile trovare $L > 0$ tale che

$$1 - \varepsilon \leq \int_{-L}^L g(y) dy \leq 1 \quad \text{e} \quad \int_{|y| > L} g(y) dy \leq \varepsilon.$$

Di conseguenza

$$(6.17) \quad \begin{aligned} &\int_{\mathbb{R}} g(y) |f(y/n) - f(0)| dy \\ &= \int_{-L}^L g(y) |f(y/n) - f(0)| dy + \int_{|y| \geq L} g(y) |f(y/n) - f(0)| dy \\ &\leq \sup_{|y| \leq L} |f(y/n) - f(0)| + \sup_{|y| \geq L} |f(y/n) - f(0)| \varepsilon \\ &\leq \sup_{|x| \leq L/n} |f(x) - f(0)| + (\sup_{y \in \mathbb{R}} |f(y/n)| + |f(0)|) \varepsilon \\ &\leq \sup_{|x| \leq L/n} |f(x) - f(0)| + 2\|f\|_u \varepsilon \end{aligned}$$

Ma dato che f è continua in 0 si ha

$$\lim_{n \rightarrow \infty} \sup_{|x| \leq L/n} |f(x) - f(0)| = 0.$$

Quindi dalla (6.17) ricaviamo

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} g(y) |f(y/n) - f(0)| dy \leq 2\|f\|_u \varepsilon,$$

che, insieme alle (6.16) ci dice che

$$\lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \delta_0(f) \right| \leq 2 \|f\|_u \varepsilon.$$

Ma ε è arbitrario, quindi

$$\lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \delta_0(f) \right| = 0,$$

e la (6.14) è dimostrata \square

6.23 Esempio. Due esempi comunemente usati, per quanto riguarda la scelta della funzione g nella Proposizione 6.22 sono

$$g(x) := \frac{1}{\pi(1+x^2)} \quad \text{oppure} \quad g(x) := \frac{1}{\sqrt{\pi}} e^{-x^2}$$

Riscalando queste funzioni si ottengono 2 successioni che “tendono” a δ_0 , vale a dire, ponendo

$$g_n(x) = \frac{n}{\pi(1+n^2x^2)} \quad \text{oppure} \quad g_n(x) := \frac{n}{\sqrt{\pi}} e^{-n^2x^2}$$

si ha

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) f(x) dx = f(0) \quad \forall f \in \mathcal{K}.$$

6.24 Osservazione. La Proposizione 6.22 può essere rafforzata nel modo seguente. Sia $g \in C_1(\mathbb{R})$ non negativa, tale che $\int_{\mathbb{R}} g = 1$ e sia $g_n(x) = ng(nx)$. Allora

$$(6.18) \quad \lim_{n \rightarrow \infty} \int_a^b g_n(x) f(x) dx = \begin{cases} f(0) & \text{se } 0 \in (a, b) \\ 0 & \text{se } 0 \notin [a, b] \end{cases} \quad \forall f \in \mathcal{K}.$$

Da notare che la (6.18) non dice nulla nel caso in cui 0 coincide con un dei due estremi dell'intervallo $[a, b]$.

Caso (i): $0 \in (a, b)$. Siano dunque $a < 0$ e $b > 0$. Voglio far vedere che

$$\lim_{n \rightarrow \infty} \int_a^b g_n(x) f(x) dx = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} g_n(x) f(x) dx = f(0),$$

dove, nella seconda uguaglianza, ho usato la (6.15). Verifico quindi che la porzione dell'integrale al di fuori dell'intervallo $[a, b]$ dà un contributo che tende a zero nel limite $n \rightarrow \infty$. Infatti

$$\begin{aligned} \int_b^{\infty} |g_n(x) f(x)| dx &\leq \|f\|_u \int_b^{\infty} |g_n(x)| dx \\ &= \|f\|_u \int_b^{\infty} |ng(nx)| dx \\ &= \|f\|_u \int_{nb}^{\infty} g(y) dy. \end{aligned}$$

Poiché g è integrabile, si ha

$$\lim_{n \rightarrow \infty} \int_{nb}^{\infty} g(y) dy = 0,$$

quindi

$$(6.19) \quad \lim_{n \rightarrow \infty} \int_b^{\infty} |g_n(x)f(x)| dx = 0.$$

Analogamente

$$(6.20) \quad \lim_{n \rightarrow \infty} \int_{-\infty}^a |g_n(x)f(x)| dx = 0.$$

Si ha dunque

$$\lim_{n \rightarrow \infty} \int_a^b g_n(x)f(x) dx = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} g_n(x)f(x) dx = f(0).$$

Caso (ii): $0 \notin [a, b]$. Possiamo supporre che valga, ad esempio, $0 < a < b$. Dalle (6.19) ottengo

$$(6.21) \quad \lim_{n \rightarrow \infty} \int_a^b |g_n(x)f(x)| dx \leq \lim_{n \rightarrow \infty} \int_a^{\infty} |g_n(x)f(x)| dx = 0.$$

6.4 La parte principale di $1/x^n$

Se n è un intero positivo la funzione $1/x^n$ non è localmente integrabile. Non è quindi possibile definire la distribuzione associata a questa funzione semplicemente come

$$\varphi_{1/x^n}(f) := \int_{\mathbb{R}} \frac{f(x)}{x^n} dx \quad f \in \mathcal{K},$$

in quanto la singolarità nell'origine rende questo integrale mal definito (a meno che la funzione di prova f non abbia uno zero di molteplicità n in 0). Per ovviare a questo problema si definisce la distribuzione *parte principale di $1/x^n$* come

$$(6.22) \quad P(1/x^n)(f) := \lim_{K \rightarrow +\infty} \int_{-K}^{+K} \frac{1}{x^n} \left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right] dx.$$

Anche se a questo punto può apparire una scelta arbitraria, vedremo in seguito che questa è la definizione "corretta", nel senso che $P(1/x^n)$, nell'ambito delle distribuzioni, obbedisce alle stesse regole di calcolo alle quali obbedisce la funzione $1/x^n$ nell'ambito delle funzioni ordinarie. Per il momento vediamo che si tratta proprio di una distribuzione.

6.25 Proposizione. *L'espressione (6.22) definisce una distribuzione.*

Dimostrazione. Sia $f \in \mathcal{K}$ e sia

$$I_K^n(f) := \int_{-K}^{+K} \frac{1}{x^n} \left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right] dx \quad n \in \mathbb{N}^*, K > 0.$$

Bisogna dimostrare che

- (a) per ogni $K > 0$ l'integrale $I_K^n(f)$ è ben definito;
- (b) il limite $\lim_{K \rightarrow +\infty} I_K^n(f)$ esiste ed è finito;
- (c) la mappa $f \rightarrow P(1/x^n)(f)$ è lineare;

(d) la mappa $f \rightarrow P(1/x^n)(f)$ è continua rispetto alla nozione di convergenza introdotta dello spazio \mathcal{K} .

Iniziamo con il punto (a). Dalla formula di Taylor col resto sotto forma di integrale 1.30 segue che possiamo scrivere il resto $(n-1)$ -simo dello sviluppo di Taylor di f nell'origine come

$$R_{n-1}(x) := f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k = x^n r_{n-1}(x),$$

in cui $r_{n-1} \in C^\infty(\mathbb{R})$. Quindi l'integrale

$$(6.23) \quad I_K^n(f) := \int_{-K}^{+K} \frac{1}{x^n} \left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right] dx = \int_{-K}^{+K} r_{n-1}(x) dx$$

è ben definito in quanto integrale di una funzione continua su un intervallo limitato. Passiamo al punto (b). Qui bisogna fare attenzione perché R_{n-1} e r_{n-1} sono funzioni C^∞ ma non sono a supporto compatto! C'è dunque il pericolo che il $\lim_{K \rightarrow +\infty} I_K^n(f)$ possa non esistere o divergere. Ma così non è. Infatti, poiché f è a supporto compatto esiste $b > 0$ tale che $\text{supp } f \subset [-b, b]$. Allora, per ogni $K > b$, si ha

$$(6.24) \quad \begin{aligned} I_K^n(f) &= \int_{-K}^{+K} r_{n-1}(x) dx = \int_{-b}^b r_{n-1}(x) dx + \int_{b \leq |x| \leq K} r_{n-1}(x) dx \\ &= I_b^n(f) + \int_{b \leq |x| \leq K} \frac{1}{x^n} \left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right] dx \\ &= I_b^n(f) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} \int_{b \leq |x| \leq K} \frac{dx}{x^{n-k}} \end{aligned}$$

L'ultimo integrale a destra è nullo se $n-k$ è dispari in quanto integrale di una funzione dispari su un insieme simmetrico rispetto all'origine. Se $n-k$ è pari, allora $n-k \geq 2$ e quindi l'integrale è convergente nel limite $K \rightarrow +\infty$. Più precisamente, se $n-k$ è pari, si ha

$$\int_{b \leq |x| \leq K} \frac{dx}{x^{n-k}} = 2 \int_b^K \frac{dx}{x^{n-k}} = \frac{2}{n-k-1} \left[\frac{1}{b^{n-k-1}} - \frac{1}{K^{n-k-1}} \right].$$

Sostituendo nella (6.24) si ottiene la seguente identità valida se $\text{supp } f \subset [-b, b]$

$$(6.25) \quad P(1/x^n)(f) = \lim_{K \rightarrow +\infty} I_K^n(f) = I_b^n(f) - \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{f^{(k)}(0)}{k!} \frac{2}{(n-k-1)b^{n-k-1}}.$$

Dunque $\lim_{K \rightarrow +\infty} I_K^n(f)$ esiste ed è finito quindi la quantità $P(1/x^n)(f)$ è ben definita.

Il punto (c) è banale.

Per quanto riguarda (d), osserviamo innanzitutto che la (6.25) produce un limite superiore su $|P(1/x^n)|$. Infatti la formula di Taylor col resto di Lagrange afferma che

$$r_{n-1}(x) = \frac{R_{n-1}(x)}{x^n} = \frac{f^{(n)}(\xi(x))}{n!}$$

in cui $\xi(x) \in (0, x)$. Di conseguenza dalla (6.23) segue che

$$\begin{aligned} |I_K^n(f)| &\leq \int_{-K}^K |r_{n-1}(x)| dx \leq 2K \sup_{x \in [-K, K]} |r_{n-1}(x)| \\ &\leq 2K \sup_{x \in [-K, K]} \frac{|f^{(n)}(x)|}{n!} \leq \frac{2K}{n!} \|f^{(n)}\|_u. \end{aligned}$$

Usando questa stima nella (6.25) si ottiene, per ogni $f \in \mathcal{K}$ tale che $\text{supp } f \subset [-b, b]$,

$$\begin{aligned} (6.26) \quad |P(1/x^n)(f)| &\leq |I_b^n(f)| + \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{|f^{(k)}(0)|}{k!} \frac{2}{(n-k-1)b^{n-k-1}} \\ &\leq \frac{2b}{n!} \|f^{(n)}\|_u + \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{|f^{(k)}(0)|}{k!} \frac{2}{(n-k-1)b^{n-k-1}}. \end{aligned}$$

Ovviamente posso sostituire b con $b_1 := \max\{b, 1\}$, in quanto se il supporto di f è contenuto in $[-b, b]$, a maggior ragione sarà contenuto in $[-b_1, b_1]$. Poiché $b_1 \geq 1$ e $n-k-1 \geq 1$, si ha

$$\frac{2}{(n-k-1)b_1^{n-k-1}} \leq 2,$$

per cui dalla (6.26) si ricava $\forall f \in \mathcal{K}$ tale che $\text{supp } f \subset [-b, b]$,

$$(6.27) \quad |P(1/x^n)(f)| \leq 2 \max\{b, 1\} \frac{\|f^{(n)}\|_u}{n!} + 2 \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{|f^{(k)}(0)|}{k!}.$$

Grazie a questa disuguaglianza è facile far vedere che $P(1/x^n)$ è un funzionale continuo. Supponiamo infatti che $(f_j)_{j=1}^\infty$ sia una successione di elementi di \mathcal{K} tale che $f_j \xrightarrow{\mathcal{K}} f$. Dobbiamo dimostrare che

$$(6.28) \quad \lim_{j \rightarrow \infty} P(1/x^n)(f_j) = P(1/x^n)(f).$$

Siccome $f_j \xrightarrow{\mathcal{K}} f$, sappiamo che esiste $b > 0$ tale che $\text{supp } f_j \subset [-b, b]$ per ogni j e, analogamente $\text{supp } f \subset [-b, b]$. Sia

$$h_j(x) := f(x) - f_j(x).$$

Allora h_j è una funzione C^∞ il cui supporto è contenuto in $[-b, b]$. Di conseguenza, applicando la (6.27) a h_j si ottiene

$$\begin{aligned} (6.29) \quad |P(1/x^n)(f) - P(1/x^n)(f_j)| &= |P(1/x^n)(h_j)| \\ &\leq 2 \max\{b, 1\} \frac{\|h_j^{(n)}\|_u}{n!} + 2 \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{|h_j^{(k)}(0)|}{k!} \end{aligned}$$

Siccome $f_j \xrightarrow{\mathcal{K}} f$, si ha che

$$\text{per ogni } k \in \mathbb{N} \quad h_j^{(k)} \rightarrow 0 \text{ uniformemente}$$

Dunque, per ogni $k \in \mathbb{N}$ e per ogni $\varepsilon > 0$ esiste un intero positivo N_k tale che $\forall j \geq N_k$ si ha

$$|h_j^{(k)}(x)| \leq \frac{\varepsilon}{2(b+1+e)}.$$

Scegliamo ora

$$\bar{N} := \max\{N_0, N_1, \dots, N_{n-1}\}.$$

Di conseguenza, dalla (6.29), segue che per ogni $j \geq \bar{N}$,

$$|P(1/x^n)(f) - P(1/x^n)(f_j)| \leq \frac{\varepsilon}{2(b+1+e)} \left[2 \max\{b, 1\} + 2 \sum_{\substack{k=0 \\ n-k \text{ pari}}}^{n-1} \frac{1}{k!} \right] \leq \varepsilon.$$

Abbiamo quindi dimostrato la (6.28) e, con essa il punto (d) della proposizione. \square

Nel caso $n = 1$ a volte è conveniente usare un'espressione alternativa per $P(1/x)$.

6.26 Proposizione. Per ogni $f \in \mathcal{K}$ si ha

$$(6.30) \quad P(1/x)(f) = \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon} \frac{f(x)}{x} dx = \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon}^{\infty} \frac{f(x) - f(-x)}{x} dx.$$

Dimostrazione. Sia $b > 0$ tale che $\text{supp } f \subset [-b, b]$. Allora per ogni $L > b$ si ha

$$\begin{aligned} \int_{-L}^L \frac{f(x) - f(0)}{x} dx &= \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon \leq |x| \leq L} \frac{f(x) - f(0)}{x} dx \\ &= \lim_{\varepsilon \rightarrow 0^+} \left[\int_{\varepsilon \leq |x| \leq L} \frac{f(x)}{x} dx - f(0) \int_{\varepsilon \leq |x| \leq L} \frac{dx}{x} \right] \\ &= \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon \leq |x| \leq L} \frac{f(x)}{x} dx && \left[\frac{1}{x} \text{ è dispari} \right] \\ &= \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon \leq |x|} \frac{f(x)}{x} dx && [\text{supp } f \subset [-L, L]] \end{aligned}$$

Poiché la precedente identità vale per ogni $L > b$ possiamo fare il limite per $L \rightarrow +\infty$ e otteniamo

$$P(1/x)(f) = \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon \leq |x|} \frac{f(x)}{x} dx.$$

Questo dimostra la prima uguaglianza della (6.30). Per ottenere la seconda uguaglianza osserviamo che

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon} \frac{f(x)}{x} dx &= \lim_{\varepsilon \rightarrow 0^+} \left(\int_{-\infty}^{-\varepsilon} \frac{f(x)}{x} dx + \int_{\varepsilon}^{\infty} \frac{f(x)}{x} dx \right) \\ &= \lim_{\varepsilon \rightarrow 0^+} \left(\int_{\infty}^{\varepsilon} \frac{f(-x)}{x} dx + \int_{\varepsilon}^{\infty} \frac{f(x)}{x} dx \right) \\ &= \int_0^{\infty} \frac{f(x) - f(-x)}{x} dx \end{aligned}$$

\square

6.5 Operazioni sulle distribuzioni

Vogliamo ora introdurre nello spazio \mathcal{K}^* delle distribuzioni alcune operazioni che hanno un analogo negli spazi di funzioni. Il principio che ci guida nel definire queste operazioni è che esse debbano essere “compatibili” con le corrispondenti operazioni esistenti negli spazi di funzioni ordinarie. Supponiamo ad esempio di voler definire la derivata F' di una distribuzione F . La definizione che daremo dovrà essere tale che, nel caso in cui sia associabile ad una funzione g differenziabile, vale a dire $F = \varphi_g$, si dovrà avere $F' = \varphi_{g'}$. In altre parole l'operazione di derivazione deve rispettare la corrispondenza fra funzioni e distribuzioni, in quei casi in cui questa corrispondenza esiste.

6.27 Definizione. Siano $F, G \in \mathcal{K}^*$, sia $f \in \mathcal{K}$ e sia $c \in \mathbb{C}$. Si definisce

- (1) la *somma fra due distribuzioni* come $(F + G)(f) := F(f) + G(f)$
- (2) il *prodotto di una distribuzione per uno scalare* come $(cF)(f) := cF(f)$
- (3) il *prodotto di una distribuzione F per una funzione $h \in C^\infty(\mathbb{R})$* come $(hF)(f) := F(hf)$.
- (4) la *derivata di una distribuzione F* come: $DF(f) = F'(f) = -F(f')$.

6.28 Osservazione. Il prodotto fra una distribuzione e una funzione C^∞ è ben definito, perché se $f \in \mathcal{K}$ e $h \in C^\infty(\mathbb{R})$ il prodotto fh è sia C^∞ che a supporto compatto, quindi appartiene allo spazio \mathcal{K} . Per cui ha senso calcolare $F(fh)$.

6.29 Osservazione. È facile verificare che le operazioni introdotte in questo modo obbediscono al principio di compatibilità descritto prima. Infatti, per quanto riguarda ad esempio il prodotto di una distribuzione F per una funzione $h \in C^\infty(\mathbb{R})$ si ha, nel caso in cui $F = \varphi_g$,

$$(h\varphi_g)(f) := \varphi_g(hg) = \int_{\mathbb{R}} g(x) [h(x)f(x)] dx = \int_{\mathbb{R}} [g(x)h(x)] f(x) dx = \varphi_{hg}(f),$$

e dunque

$$h\varphi_g = \varphi_{hg}.$$

Analogamente, per quanto riguarda la definizione di F' , integrando per parti e tenendo presente che f è a supporto compatto, si ottiene

$$\begin{aligned} F'(f) &= -F(f') = -\varphi_g(f') = -\int_{\mathbb{R}} g(x)f'(x) dx \\ &= -g(x)f(x) \Big|_{-\infty}^{+\infty} + \int_{\mathbb{R}} g'(x)f(x) dx = \varphi_{g'}(f), \end{aligned}$$

vale a dire

$$\varphi'_g = \varphi_{g'}.$$

Ancora più facilmente si verifica che

$$\varphi_f + \varphi_g = \varphi_{f+g} \qquad c\varphi_f = \varphi_{cf}.$$

In conclusione le operazioni introdotte nello spazio delle distribuzioni costituiscono un “estensione” delle corrispondenti operazioni esistenti negli spazi di funzioni.

6.30 Esempio. Usando la definizione di prodotto di una funzione C^∞ per una distribuzione e quella di derivata, voglio verificare le seguenti identità fra distribuzioni

$$(a) \quad x\delta_0 = 0 \qquad (b) \quad xP(1/x) = \varphi_1 \qquad (c) \quad x\delta'_0 = -\delta_0.$$

Inizio da (a): sia $f \in \mathcal{K}$. Allora, per definizione

$$(6.31) \qquad (x\delta_0)(f) = \delta_0(xf) = [xf(x)]_{x=0} = 0 \cdot f(0) = 0.$$

Poiché questa uguaglianza vale *per ogni* $f \in \mathcal{K}$, allora posso concludere che

$$x\delta_0 = 0.$$

L'identità (b) si verifica in modo analogo. Se $f \in \mathcal{K}$, dalla (6.30) segue che

$$\begin{aligned} [xP(1/x)](f) &= P(1/x)(xf) = \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon} \frac{x f(x)}{x} dx \\ &= \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon} 1 \cdot f(x) dx = \varphi_1(f). \end{aligned}$$

Quindi $[xP(1/x)](f) = \varphi_1(f)$ per ogni $f \in \mathcal{K}$ e dunque posso scrivere semplicemente

$$xP(1/x) = \varphi_1.$$

Per quanto riguarda (c) si ottiene

$$\begin{aligned} (x\delta'_0)(f) &= \delta'_0(xf) = -\delta_0((xf)') = -\delta_0(xf' + f) \\ &= -0 \cdot f'(0) - f(0) = -f(0) = -\delta_0(f). \end{aligned}$$

6.31 Attenzione! (La notazione può essere ambigua). Sia φ_g la distribuzione associata alla funzione g . È una prassi comune quella di usare il simbolo g per denotare indifferentemente la *funzione* g oppure la *distribuzione* φ_g . Quindi, ad esempio, l'identità $xP(1/x) = \varphi_1$ viene scritta di solito come $xP(1/x) = 1$. Quando si vuole sottolineare il fatto che una qualche relazione va intesa come relazione fra distribuzioni, a volte si aggiunge la postilla “in senso debole” oppure “nel senso delle distribuzioni”. Possiamo quindi scrivere

$$xP(1/x) = 1 \quad \text{nel senso delle distribuzioni.}$$

In molti testi,⁷ al posto di $\varphi_g(f)$ si scrive $g(f)$ oppure (g, f) . Quindi

$$(g, f) = g(f) := \int_{\mathbb{R}} g(x)f(x) dx.$$

Fate attenzione al fatto che (g, f) *non* è un prodotto scalare in quanto è lineare in entrambi gli argomenti, mentre un prodotto scalare è antilineare nel secondo argomento (Proposizione 4.4). Il lettore saprà evitare di farsi confondere da queste ambiguità nella notazione e, col tempo, imparerà addirittura ad apprezzarle (forse).

6.32 Osservazione. (perché non c'è il prodotto fra distribuzioni). Non vi sarà certamente sfuggito che nella lista delle operazioni introdotte sopra manca il *prodotto fra distribuzioni*.⁸ Questa non è una dimenticanza, ma piuttosto un inevitabile prezzo da pagare per poter trattare oggetti (le distribuzioni appunto) molto più generali

⁷potrebbe accadere anche in questo

⁸a Beavis and Butthead potrebbe forse essere sfuggito.

delle funzioni ordinarie. Non è possibile definire il prodotto fra distribuzioni in modo da rispettare la corrispondenza con le funzioni. Se ciò fosse possibile, infatti, questo prodotto dovrebbe essere commutativo, associativo e si dovrebbe avere

$$0 \cdot F = 0 \qquad \varphi_1 \cdot F = 1 \cdot F = F \qquad \forall F \in \mathcal{K}^* .$$

D'altra parte, dalle identità dell'Esempio 6.30 segue che

$$0 = P(1/x) \cdot 0 = P(1/x) (x\delta_0) = [P(1/x)x] \delta_0 = 1 \cdot \delta_0 = \delta_0 ,$$

che è un assurdo.

Introduciamo ora una nozione di convergenza nello spazio delle distribuzioni.

6.33 Definizione. Una successione $(F_n)_{n=1}^{\infty}$ di distribuzioni si dice *convergente* alla distribuzione $F \in \mathcal{K}^*$ e si scrive $F_n \xrightarrow{\mathcal{K}^*} F$ se si ha $F_n(f) \rightarrow F(f)$ per ogni $f \in \mathcal{K}$.

6.34 Esempio. Se

$$g_n(x) := \frac{n e^{-n^2 x^2}}{\sqrt{\pi}} ,$$

dalla Proposizione 6.22 segue che

$$\lim_{n \rightarrow \infty} \varphi_{g_n}(f) = \delta_0(f) \qquad \forall f \in \mathcal{K} .$$

Dunque possiamo scrivere $\varphi_{g_n} \xrightarrow{\mathcal{K}^*} \delta_0$, oppure, equivalentemente,

$$\frac{n e^{-n^2 x^2}}{\sqrt{\pi}} \rightarrow \delta_0 \quad (\text{nel senso delle distribuzioni}).$$

6.35 Proposizione. *Ogni distribuzione ha derivate di tutti gli ordini.*

Dimostrazione. Basta verificare che, iterando la definizione di derivata, si ottiene

$$(6.32) \qquad F^{(n)}(f) = (-1)^n F(f^{(n)}) . \quad \square$$

6.36 Esempio. (La derivata n -sima della delta di Dirac). Per definizione, la derivata della delta di Dirac è data da

$$\delta_0'(f) = -\delta_0(f') = -f'(0) .$$

Derivando una seconda volta si ha

$$\delta_0''(f) = [D(\delta_0')](f) = -\delta_0'(f') = \delta_0(f'') = f''(0) .$$

Posso iterare la procedura e calcolare le derivate successive, ottenendo

$$\delta_0^{(n)}(f) = (-1)^n f^{(n)}(0) .$$

Ad esempio

$$\delta_0^{(n)}(e^x) = (-1)^n D^n(e^x)|_{x=0} = (-1)^n$$

$$\delta_0^{(n)}(\sin x) = (-1)^n D^n(\sin x)|_{x=0} = \begin{cases} 0 & \text{se } n \text{ è pari} \\ (-1)^{n+(n-1)/2} & \text{se } n \text{ è dispari.} \end{cases}$$

6.37 Proposizione. Se $h \in C^\infty(\mathbb{R})$ e $F \in \mathcal{K}^*$ allora $(hF)' = h'F + hF'$.

Dimostrazione. Si tratta di una semplice verifica che si ottiene utilizzando le due definizioni

$$\begin{aligned} F'(f) &:= -F(f') & f \in \mathcal{K} & & \text{regola della derivata} \\ (hF)(f) &:= F(hf) & f \in \mathcal{K}, h \in C^\infty(\mathbb{R}) & & \text{regola del prodotto} \end{aligned}$$

insieme alla linearità delle distribuzioni. Infatti, se $f \in \mathcal{K}$, si ha

$$\begin{aligned} [D(hF)](f) &= -(hF)(f') = -F(hf') = -F((hf)' - h'f) \\ &= -F((hf)') + F(h'f) = F'(hf) + F(h'f) \\ &= (hF')(f) + (h'F)(f) = [hF' + h'F](f). \end{aligned}$$

Poiché la precedente identità vale per ogni $f \in \mathcal{K}$, abbiamo

$$D(hF) = h'F + hF'. \quad \square$$

6.38 Attenzione! Abbiamo visto quali sono le regole per calcolare la derivata di una distribuzione e il prodotto di una distribuzione per una funzione C^∞ . Spesso queste due operazioni compaiono entrambi in una certa espressione e bisogna fare attenzione a procedere nell'ordine giusto. Supponiamo, ad esempio, di voler "sviluppare" l'espressione

$$(x^2 D[\sin x D(xF)])(f).$$

Si procede nel modo seguente:

$$\begin{aligned} (x^2 D[\sin x D(xF)])(f) &= D[\sin x D(xF)](x^2 f) = -(\sin x D(xF))((x^2 f)') \\ &= -(\sin x D(xF))(2xf + x^2 f') = -D(xF)(2x \sin xf + x^2 \sin xf') \\ &= (xF)([2x \sin xf + x^2 \sin xf']') \\ &= (xF)(2 \sin xf + 2x \cos xf + 4x \sin xf' + x^2 \cos xf' + x^2 \sin xf'') \\ &= F(2x \sin xf + 2x^2 \cos xf + 4x^2 \sin xf' + x^3 \cos xf' + x^3 \sin xf''). \end{aligned}$$

6.39 Proposizione. Sia (g_n) una successione di funzioni continue su \mathbb{R} che converge uniformemente alla funzione continua g . Allora

$$\varphi_{g_n} \xrightarrow{\mathcal{K}^*} \varphi_g.$$

Dimostrazione. Devo far vedere che

$$\varphi_{g_n}(f) \rightarrow \varphi_g(f) \quad \forall f \in \mathcal{K},$$

vale a dire che

$$(6.33) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) f(x) dx = \int_{\mathbb{R}} g(x) f(x) dx \quad \forall f \in \mathcal{K}.$$

Sia $f \in \mathcal{K}$ e sia $a > 0$ tale che $\text{supp } f \subset [-a, a]$. Allora

$$(6.34) \quad \begin{aligned} \left| \int_{\mathbb{R}} g_n(x) f(x) dx - \int_{\mathbb{R}} g(x) f(x) dx \right| &\leq \int_{\mathbb{R}} |g(x) - g_n(x)| |f(x)| dx \\ &= \int_{-a}^a |g(x) - g_n(x)| |f(x)| dx \leq 2a \|g - g_n\|_u \|f\|_u, \end{aligned}$$

da cui la (6.33) segue. □

6.40 Proposizione. Siano $(F_n)_{n=1}^{\infty}$ e F distribuzioni. Se $F_n \xrightarrow{\mathcal{K}^*} F$ allora $F'_n \xrightarrow{\mathcal{K}^*} F'$.

Dimostrazione. Se $f \in \mathcal{K}$ allora $f' \in \mathcal{K}$. Quindi, se $F_n \xrightarrow{\mathcal{K}^*} F$, si ha

$$\lim_{n \rightarrow \infty} F'_n(f) = \lim_{n \rightarrow \infty} -F_n(f') = -F(f') = F'(f).$$

Quindi $F'_n \xrightarrow{\mathcal{K}^*} F'$. □

6.41 Esempio. Consideriamo la distribuzione $\varphi_{|x|}$, definita, come al solito, da

$$\varphi_{|x|}(f) := \int_{\mathbb{R}} |x| f(x) dx \quad f \in \mathcal{K}.$$

La funzione $|x|$ non è derivabile nell'origine in senso ordinario. Tuttavia la distribuzione associata $\varphi_{|x|}$ è perfettamente derivabile e vale, così come uno si aspetta

$$(6.35) \quad D\varphi_{|x|} = \varphi_{\text{sgn}(x)},$$

o, usando la notazione più comune,

$$|x|' = \text{sgn}(x) \quad (\text{nel senso delle distribuzioni}).$$

La dimostrazione della (6.35) è una semplice verifica:

$$\begin{aligned} D\varphi_{|x|}(f) &:= -\varphi_{|x|}(f') = -\int_{\mathbb{R}} |x| f'(x) dx \\ &= \int_{-\infty}^0 x f'(x) dx - \int_0^{\infty} x f'(x) dx. \end{aligned}$$

Integrando per parti e tenendo conto del fatto che f è a supporto compatto ottengo

$$\int_{-\infty}^0 x f'(x) dx = [xf(x)]_{-\infty}^0 - \int_{-\infty}^0 f(x) dx = -\int_{-\infty}^0 f(x) dx.$$

Analogamente

$$\int_0^{\infty} x f'(x) dx = [xf(x)]_0^{\infty} - \int_0^{\infty} f(x) dx = -\int_0^{\infty} f(x) dx.$$

Mettendo insieme i due pezzi ottengo

$$D\varphi_{|x|}(f) = -\int_{-\infty}^0 f(x) dx + \int_0^{\infty} f(x) dx = \int_{\mathbb{R}} \text{sgn}(x) f(x) dx = \varphi_{\text{sgn}(x)}(f).$$

Quindi si ha che $D\varphi_{|x|}(f) = \varphi_{\text{sgn}(x)}(f)$ per ogni $f \in \mathcal{K}$, vale a dire $D\varphi_{|x|} = \varphi_{\text{sgn}(x)}$.

L'identità (6.35) può essere generalizzata, come traspare dal problema seguente.

6.42 Problema. Sia $g \in C^1(\mathbb{R})$. Dimostrare che la derivata nel senso delle distribuzioni di $g(|x|)$ è data da $g'(|x|) \text{sgn}(x)$.

6.43 Esempio. ($\vartheta' = \delta_0$). Vogliamo far vedere che vale (ricorda la definizione della distribuzione ϑ , (6.9))

$$(6.36) \quad \vartheta' = \delta_0.$$

Infatti, per ogni $f \in \mathcal{K}$, dal teorema fondamentale del calcolo, segue che

$$\vartheta'(f) = -\vartheta(f') = -\int_0^{\infty} f'(x) dx = -(f(\infty) - f(0)) = f(0) = \delta_0(f).$$

La ϑ rappresenta il caso più semplice di una di una distribuzione associata ad una funzione con dei punti di discontinuità. Mostriamo ora una formula che generalizza la (6.36) al caso di una funzione con molti punti di salto.

6.44 Notazione. Sia $g : \mathbb{R} \rightarrow \mathbb{C}$. Definisco

$$(D_{\circ}g)(x) = \begin{cases} g'(x) & \text{se } g \text{ è derivabile in } x \\ 0 & \text{altrimenti.} \end{cases}$$

6.45 Proposizione. Sia g una funzione differenziabile a tratti con discontinuità nei punti u_i , $i = 1, 2, 3, \dots$. Ponendo $\Delta g(u_i) := g(u_i^+) - g(u_i^-)$ uguale al valore della discontinuità di g in u_i , si ha la seguente identità, nel senso delle distribuzioni

$$Dg := D_{\circ}g + \sum_{i=1}^{\infty} \Delta g(u_i) \delta_{u_i}$$

Dimostrazione. Considero il caso in cui ci sia un solo punto di discontinuità u . Abbiamo dunque

$$\lim_{x \rightarrow u^-} g(x) = g(u^-) \quad \lim_{x \rightarrow u^+} g(x) = g(u^+) \quad \Delta g(u) := g(u^+) - g(u^-).$$

Posso scrivere, per ogni $f \in \mathcal{K}$,

$$\begin{aligned} \varphi'_g(f) &= -\varphi_g(f') = -\int_{\mathbb{R}} g(x)f'(x) dx = -\int_{-\infty}^u g(x)f'(x) dx - \int_u^{\infty} g(x)f'(x) dx \\ &= -g(x)f(x) \Big|_{-\infty}^u + \int_{-\infty}^u g'(x)f(x) dx - g(x)f(x) \Big|_u^{\infty} + \int_u^{\infty} g'(x)f(x) dx \\ &= -f(u)g(u^-) + f(u)g(u^+) + \int_{\mathbb{R}} D_{\circ}g(x)f(x) dx \\ &= f(u)[g(u^+) - g(u^-)] + \varphi_{D_{\circ}g}(f) \\ &= \Delta g(u) \delta_u(f) + \varphi_{D_{\circ}g}(f). \quad \square \end{aligned}$$

6.46 Esempi. Vediamo alcune semplici applicazioni della Proposizione 6.45.

- (1) Sia $a \in \mathbb{R}$. Allora $\text{sgn}(x - a)$ ha derivata nulla ovunque tranne nel punto a in cui la funzione ha un salto pari a 2. Quindi

$$D \text{sgn}(x - a) = 2 \delta_a.$$

Analogamente

$$D \text{sgn}(a - x) = -2 \delta_a.$$

- (2) Sia

$$g(x) := \begin{cases} 1 & x < 0 \\ x - 2 & x \geq 0 \end{cases}$$

La derivata “ordinaria” di g è data da

$$g'(x) := \begin{cases} 0 & x < 0 \\ \text{non esiste} & x = 0 \\ 1 & x > 0 \end{cases} \quad \text{nel senso ordinario,}$$

dunque, ricordando l'Osservazione 6.19, possiamo scrivere

$$D_{\circ}g = H.$$

Inoltre g ha un'unica discontinuità in $x = 0$ e vale

$$(\Delta g)(0) = g(0^+) - g(0^-) = -2 - 1 = -3,$$

quindi, usando la Proposizione 6.45, otteniamo

$$Dg = H - 3\delta_0. \quad \text{nel senso delle distribuzioni.}$$

- (3) Sia $[x]$ la *parte intera* di x , vale a dire il più grande intero minore o uguale a x . Questa funzione è derivabile, con derivata nulla, per ogni $x \in \mathbb{R} \setminus \mathbb{Z}$. In corrispondenza di ogni intero k si ha una discontinuità $\Delta[x](k) = 1$. Quindi

$$D[x] = \sum_{k \in \mathbb{Z}} \delta_k.$$

Osserviamo (se qualcuno si fosse posto il problema) che la serie $\sum_{k \in \mathbb{Z}} \delta_k$ è sempre convergente come distribuzione. Infatti se $f \in \mathcal{K}$ esiste $L > 0$ tale che $\text{supp } f \subset [-L, L]$, quindi

$$\left(\sum_{k \in \mathbb{Z}} \delta_k \right) (f) = \sum_{k \in \mathbb{Z}} f(k) = \sum_{k=-L}^L f(k),$$

e dunque si tratta in realtà di una somma di un numero finito di termini.

- (4) La funzione $g(x) := [x^2]$, come la precedente è costante a tratti, vale a dire è derivabile con derivata nulla ad eccezione dei punti $\pm\sqrt{k}$ con $k \in \mathbb{N}^*$ (vedi figura 6.2). Quindi $D_{\circ}g = 0$. Inoltre la funzione nel punto $\pm\sqrt{k}$ ha una discontinuità pari rispettivamente a ± 1 . Quindi

$$D[x^2] = \sum_{k=1}^{\infty} (\delta_{\sqrt{k}} - \delta_{-\sqrt{k}}).$$

- (5) La funzione

$$\text{sgn}(\sin x)$$

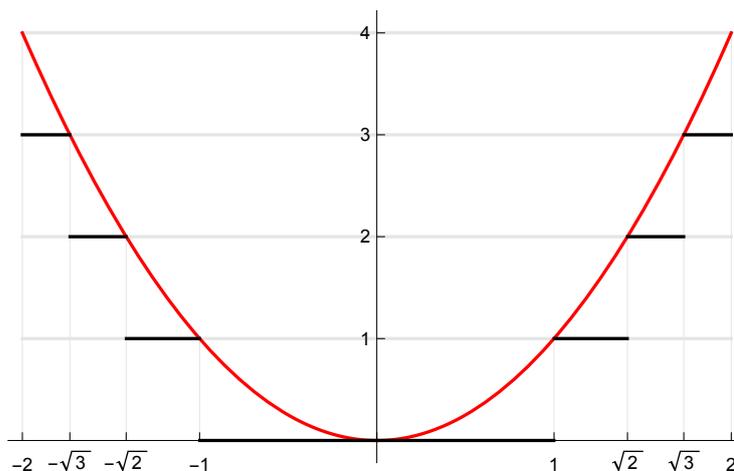
ha derivata ordinaria ovunque nulla tranne che nei punti $k\pi$ con $k \in \mathbb{Z}$, in cui essa non esiste, quindi $D_{\circ} \text{sgn}(\sin x) = 0$. Nei punti $x = k\pi$ la funzione g ha un salto pari a ± 2 a seconda della parità di k . Otteniamo quindi, nel senso delle distribuzioni,

$$D \text{sgn}(\sin x) = 2 \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi}$$

La formula $x\delta'_0 = -\delta_0$, ottenuta nell'esempio 6.30 è un caso particolare di un risultato più generale che afferma che il prodotto di una funzione C^∞ per la derivata di una delta di Dirac è una combinazione lineare (a coefficienti costanti) di derivate di delta di Dirac.

6.47 Proposizione. *Se $h \in C^\infty(\mathbb{R})$ e $a \in \mathbb{R}$, si ha*

$$(6.37) \quad h(x) \delta_a^{(n)} = \sum_{k=0}^n (-1)^k \binom{n}{k} h^{(k)}(a) \delta_a^{(n-k)}.$$

Figura 6.2: Grafico della funzione $[x^2]$.

Dimostrazione. Sia $f \in \mathcal{K}$. Allora

$$\left(h(x) \delta_a^{(n)} \right) (f) = \delta_a^{(n)} (hf) = (-1)^n D^n (hf)(a).$$

A questo punto posso usare la regola di Leibniz 1.12 per la derivata n -sima del prodotto e ottengo

$$\begin{aligned} \left(h(x) \delta_a^{(n)} \right) (f) &= (-1)^n \sum_{k=0}^n \binom{n}{k} h^{(k)}(a) f^{(n-k)}(a) \\ &= (-1)^n \sum_{k=0}^n \binom{n}{k} h^{(k)}(a) (-1)^{n-k} \delta_a^{(n-k)}(f) \\ (6.38) \quad &= \sum_{k=0}^n (-1)^k \binom{n}{k} h^{(k)}(a) \delta_a^{(n-k)}(f) \\ &= \left[\sum_{k=0}^n (-1)^k \binom{n}{k} h^{(k)}(a) \delta_a^{(n-k)} \right] (f) \end{aligned}$$

Poiché la (6.38) è valida per ogni $f \in \mathcal{K}$, la proposizione è dimostrata. \square

6.48 Esempio. Per apprezzare l'utilità della Proposizione 6.47, supponiamo di dover calcolare

$$D^7 (x e^x \delta_0'').$$

Si capisce che se uno inizia a calcolare subito la derivata settima rischia di passarci mezzo pomeriggio. Infatti il conto inizierebbe così

$$D^7 (x e^x \delta_0'') = D^6 [D (x e^x \delta_0'')] = D^6 [e^x \delta_0'' + x e^x \delta_0'' + x e^x \delta_0''']$$

bene, la prima derivata l'abbiamo fatta. Ne mancano altre sei. Invece di proseguire con questo metodo che richiede una certa dose di pazienza, vediamo qual è l'approccio

corretto: la prima cosa da fare è *semplificare l'espressione* $xe^x\delta_0''$ usando la 6.47. Sia $h(x) = xe^x$. Dunque

$$\begin{aligned} h'(x) &= e^x(1+x) & h'(0) &= 1 \\ h''(x) &= e^x(2+x) & h''(0) &= 2. \end{aligned}$$

Grazie alla 6.47 trovo

$$\begin{aligned} xe^x\delta_0'' &= h(x)\delta_0'' = \binom{2}{0}h(0)\delta_0'' - \binom{2}{1}h'(0)\delta_0' + \binom{2}{2}h''(0)\delta_0 \\ &= \delta_0'' - 2\delta_0' + 2\delta_0. \end{aligned}$$

A questo punto, essendo “scomparsa la x ”, fare la derivata settima è banale

$$D^7(xe^x\delta_0'') = D^7(\delta_0'' - 2\delta_0' + 2\delta_0) = \delta_0^{(9)} - 2\delta_0^{(8)} + 2\delta_0^{(7)}.$$

6.49 Esempio. Vogliamo calcolare le seguenti distribuzioni, semplificando, se possibile il risultato:

$$(a) \quad \cos x \delta_0'' \qquad (b) \quad D^3[e^{2x}D^4(|x|)].$$

Soluzione. (a) Usando la (6.37), si ottiene

$$\begin{aligned} \cos x \delta_0'' &= \binom{2}{0}\cos(0)\delta_0'' - \binom{2}{1}D[\cos x]_{x=0}\delta_0' + \binom{2}{2}D^2[\cos x]_{x=0}\delta_0 \\ &= \cos(0)\delta_0'' - 2\sin(0)\delta_0' - \cos(0)\delta_0 = \delta_0'' - \delta_0. \end{aligned}$$

(b) Conviene innanzitutto semplificare l'espressione all'interno della derivata terza. Poiché $D|x| = \operatorname{sgn}(x)$, otteniamo

$$D^2(|x|) = D\operatorname{sgn}(x) = 2\delta_0,$$

quindi, usando ancora la (6.37),

$$\begin{aligned} e^{2x}D^4(|x|) &= 2e^{2x}\delta_0'' \\ &= 2[e^0\delta_0'' - 2D(e^{2x})(0)\delta_0' + D^2(e^{2x})(0)\delta_0] \\ &= 2\delta_0'' - 8\delta_0' + 8\delta_0. \quad \blacktriangleleft \end{aligned}$$

Un caso importante da considerare è il prodotto di un monomio x^m per la derivata di una delta nell'origine.

6.50 Problema. Siano m, n due interi non negativi. Dimostrare che

$$x^m\delta_0^{(n)} = \begin{cases} (-1)^m \frac{n!}{(n-m)!} \delta_0^{(n-m)} & \text{se } n \geq m \\ 0 & \text{se } n < m \end{cases}$$

Suggerimento: il risultato segue facilmente dal Lemma 6.51.

Può essere utile tenere a mente alcuni casi particolari:

$$(6.39) \quad x^n\delta_0^{(n)} = (-1)^n n! \delta_0 \qquad x^m\delta_0^{(n)} = 0 \quad \text{se } m > n.$$

6.51 Lemma. Siano $k, m \in \mathbb{N}$. Allora, se $f \in C^k(\mathbb{R})$, si ha

$$D^k(x^m f)(0) = \begin{cases} 0 & \text{se } k < m \\ \frac{k!}{(k-m)!} f^{(k-m)}(0) & \text{se } k \geq m. \end{cases}$$

Dimostrazione. Osserviamo innanzitutto che

$$D^j(x^m)(0) = m! \delta_{jm} = \begin{cases} m! & \text{se } j = m \\ 0 & \text{se } j \neq m. \end{cases}$$

Di conseguenza, usando la regola di Leibniz 1.12, posso scrivere

$$D^k(x^m f)(0) = \sum_{j=0}^k \binom{k}{j} D^j(x^m)(0) f^{(k-j)}(0) = m! \sum_{j=0}^k \binom{k}{j} \delta_{jm} f^{(k-j)}(0).$$

Se $k < m$ tutti i termini dell'ultima somma sono nulli. Viceversa, se $k \geq m$, nell'ultima somma c'è un unico termine diverso da zero, quello con $j = m$. \square

Mentre in generale se voglio definire il prodotto di una funzione h per una distribuzione arbitraria F devo assumere che h sia C^∞ , nel caso in cui la distribuzione sia una delta di Dirac o una derivata della delta di Dirac, la Proposizione 6.47 mi permette di *definire* tale prodotto sotto ipotesi più deboli su h .

6.52 Definizione. Sia $a \in \mathbb{R}$ e sia h una funzione di classe C^m in un intorno del punto a . Definisco, per ogni $n \leq m$,

$$(6.40) \quad h(x) \delta_a^{(n)} = \sum_{k=0}^n (-1)^k \binom{n}{k} h^{(k)}(a) \delta_a^{(n-k)}.$$

6.53 Esempio. La definizione precedente ci dice che per definire $h(x) \delta_a^{(n)}$ ho bisogno che h sia *almeno* di classe C^n in un intorno di a . Ad esempio se h è continua in un intorno di a posso scrivere $h(x) \delta_a = h(a) \delta_a$. Se h è C^1 in un intorno di a posso definire anche $h(x) \delta'_a = h(a) \delta'_a - h'(a) \delta_a$. Esempi:

$$\begin{aligned} \operatorname{sgn}(x) \delta_1 &= \operatorname{sgn}(1) \delta_1 = \delta_1 & \operatorname{sgn}(x-1) \delta_0 &= \operatorname{sgn}(-1) \delta_0 = -\delta_0 \\ H(x) \delta'_2 &= H(2) \delta'_2 - H'(2) \delta_2 = \delta'_2 & H(x) \delta'_{-3} &= H(-3) \delta'_{-3} - H'(-3) \delta_{-3} = 0. \end{aligned}$$

Attenzione a non scrivere espressioni senza senso

$$\begin{aligned} H(x) \delta_0 &= \text{non è definita perché } H \text{ è discontinua in } 0 \\ |x| \delta'_0 &= \text{non è definita perché } |x| \text{ non è } C^1 \text{ in un intorno di } 0. \end{aligned}$$

6.54 Proposizione. Siano g e h due funzioni differenziabili a tratti, tali che nessun punto di discontinuità di g coincide con un punto di discontinuità di h . Allora vale la seguente identità nel senso delle distribuzioni

$$D(gh) = (Dg)h + g(Dh).$$

Dimostrazione. Segue facilmente dalla 6.45. \square

6.55 Osservazione. Nella Proposizione 6.54, l'ipotesi che g e h siano discontinue in punti diversi è fondamentale per evitare di scrivere espressioni che non hanno senso. Supponiamo infatti di volerlo applicare ad H (la funzione di Heaviside) e sgn . In questo caso otterremmo

$$(6.41) \quad \begin{aligned} D(H(x) \operatorname{sgn}(x)) &= H'(x) \operatorname{sgn}(x) + H(x) D \operatorname{sgn}(x) \\ &= \operatorname{sgn}(x) \delta_0 + 2H(x) \delta_0 = (\operatorname{sgn}(x) + 2H(x)) \delta_0. \end{aligned}$$

Ma l'espressione a destra non ha alcun significato perché $H(x)\delta_0$ e $\operatorname{sgn}(x)\delta_0$ non sono ben definite. D'altra parte la derivata di $H(x)\operatorname{sgn}(x)$ si può facilmente calcolare senza bisogno di usare la formula della derivata del prodotto. Infatti, poiché

$$H(x)\operatorname{sgn}(x) = H(x),$$

ottengo banalmente

$$D(H(x)\operatorname{sgn}(x)) = \delta_0.$$

6.56 Esempio. Per calcolare

$$D(\operatorname{sgn}(x-1)\operatorname{sgn}(x+1)e^{-x^2}),$$

possiamo usare la Proposizione 6.54 ottenendo

$$\begin{aligned} D(\operatorname{sgn}(x-1)\operatorname{sgn}(x+1)e^{-x^2}) &= 2\delta_1\operatorname{sgn}(x+1)e^{-x^2} + 2\delta_{-1}\operatorname{sgn}(x-1)e^{-x^2} - 2x\operatorname{sgn}(x-1)\operatorname{sgn}(x+1)e^{-x^2} \\ &= (2\delta_1 - 2\delta_{-1})e^{-1} - 2x\operatorname{sgn}(x-1)\operatorname{sgn}(x+1)e^{-x^2}. \end{aligned}$$

6.6 Alcune proprietà di $P(1/x^n)$

Quando abbiamo introdotto la distribuzione $P(1/x^n)$ avevamo preannunciato che è naturale pensare che essa sia la corretta distribuzione da associare alla funzione $1/x^n$, perché soddisfa le stesse regole di calcolo. Le seguenti tre proposizioni precisano questo concetto.

6.57 Proposizione. *Si ha*

$$D \log |x| = P(1/x) \quad \text{nel senso delle distribuzioni.}$$

Dimostrazione. Data la funzione localmente integrabile $x \rightarrow \log |x|$, consideriamo la distribuzione associata $\varphi_{\log|\cdot|}$. Osserviamo che la derivata della funzione $\log|\cdot|$ è la funzione $1/x$ che non è localmente integrabile. Per capire chi è $\varphi'_{\log|x|}$ partiamo perciò dalla definizione di derivata di una distribuzione

$$\begin{aligned} \varphi'_{\log|x|}(f) &= -\varphi_{\log|x|}(f') = -\int_{\mathbb{R}} \log|x| f'(x) dx \\ &= -\lim_{\varepsilon \rightarrow 0^+} \left[\int_{-\infty}^{-\varepsilon} \log(-x) f'(x) dx + \int_{\varepsilon}^{\infty} \log x f'(x) dx \right]. \end{aligned}$$

Integrando per parti si trova

$$-\int_{\varepsilon}^{\infty} \log x f'(x) dx = -[\log x f(x)]_{\varepsilon}^{\infty} + \int_{\varepsilon}^{\infty} \frac{f(x)}{x} dx = \log \varepsilon f(\varepsilon) + \int_{\varepsilon}^{\infty} \frac{f(x)}{x} dx$$

Analogamente si ottiene

$$-\int_{-\infty}^{-\varepsilon} \log(-x) f'(x) dx = -\log \varepsilon f(-\varepsilon) + \int_{-\infty}^{-\varepsilon} \frac{f(x)}{x} dx,$$

da cui,

$$(6.42) \quad \varphi'_{\log|x|}(f) = \lim_{\varepsilon \rightarrow 0^+} \log \varepsilon [f(\varepsilon) - f(-\varepsilon)] + \int_{|x| \geq \varepsilon} \frac{f(x)}{x}.$$

Poiché f è derivabile

$$\lim_{\varepsilon \rightarrow 0^+} \log \varepsilon [f(\varepsilon) - f(-\varepsilon)] = \lim_{\varepsilon \rightarrow 0^+} \left[(\varepsilon \log \varepsilon) \frac{f(\varepsilon) - f(-\varepsilon)}{\varepsilon} \right] = \lim_{\varepsilon \rightarrow 0^+} (\varepsilon \log \varepsilon) 2f'(0) = 0.$$

Quindi, facendo tendere ε a zero nella (6.42) e ricordando la (6.30),

$$\varphi'_{\log|x|}(f) = \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon} \frac{f(x)}{x} =: P(1/x)(f).$$

Abbiamo quindi ottenuto

$$\varphi'_{\log|x|} = P(1/x),$$

che, tenendo presente quanto detto al punto 6.31, si può anche scrivere come

$$(\log|x|)' = P(1/x).$$

6.58 Proposizione. *Siano n, m due interi positivi. Allora*

$$x^m P(1/x^n) = \begin{cases} x^{m-n} & \text{se } m \geq n \\ P(1/x^{n-m}) & \text{se } m < n. \end{cases}$$

Dimostrazione. Sia $f \in \mathcal{K}$. Allora

$$(6.43) \quad [x^m P(1/x^n)](f) = P(1/x^n)(x^m f) = \lim_{L \rightarrow +\infty} I_L^n(x^m f),$$

in cui

$$(6.44) \quad I_L^n(x^m f) := \int_{-L}^L \frac{1}{x^n} \left(x^m f(x) - \sum_{k=0}^{n-1} \frac{D^k(x^m f)(0)}{k!} x^k \right) dx.$$

Caso (a): $m \geq n$. La quantità $D^k(x^m f)(0)$ è nulla se $k < m$. Poiché il massimo valore che può assumere k è $n-1$ e poiché $n-1 < m$, si ha che tutti i termini della sommatoria sono nulli. Quindi

$$I_L^n(x^m f) := \int_{-L}^L \frac{1}{x^n} x^m f(x) dx = \int_{-L}^L x^{m-n} f(x) dx.$$

Facendo tendere L a $+\infty$, dalla (6.43) otteniamo

$$[x^m P(1/x^n)](f) = \int_{-\infty}^{\infty} x^{m-n} f(x) dx \quad \forall f \in \mathcal{K},$$

vale a dire

$$x^m P(1/x^n) = x^{m-n} \quad \text{nel senso delle distribuzioni.}$$

Caso (b): $m < n$. In questo caso, grazie al Lemma 6.51, otteniamo

$$\begin{aligned} \sum_{k=0}^{n-1} \frac{x^k}{k!} D^k(x^m f)(0) &= \sum_{k=m}^{n-1} \frac{x^k}{k!} \binom{k}{m} m! f^{(k-m)}(0) \\ &= \sum_{k=m}^{n-1} \frac{x^k}{(k-m)!} f^{(k-m)}(0) = \sum_{k=0}^{n-m-1} \frac{f^{(k)}(0)}{k!} x^{m+k}. \end{aligned}$$

Sostituendo questa espressione nella (6.44) otteniamo

$$(6.45) \quad I_L^n(x^m f) = \int_{-L}^L \frac{1}{x^{n-m}} \left(f(x) - \sum_{k=0}^{n-m-1} \frac{f^{(k)}(0)}{k!} x^k \right) dx.$$

Sostituendo nella (6.43) troviamo infine

$$[x^m P(1/x^n)](f) = \int_{-\infty}^{\infty} \frac{1}{x^{n-m}} \left(f(x) - \sum_{k=0}^{n-m-1} \frac{f^{(k)}(0)}{k!} x^k \right) dx = P(1/x^{n-m})(f).$$

□

6.59 Proposizione. *Se n è un intero positivo si ha*

$$DP(1/x^n) = -nP(1/x^{n+1}).$$

Dimostrazione. Sia $b > 0$ tale che $\text{supp } f \subset [-b, b]$. Poniamo

$$I_L^n(f) := \int_{-L}^L \frac{1}{x^n} \left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right] dx.$$

Dalla definizione di derivata di una distribuzione si ottiene

$$DP(1/x^n)(f) := -P(1/x^n)(f') = -\lim_{L \rightarrow \infty} I_L^n(f').$$

D'altra parte, integrando per parti ed usando il fatto che $f(x) - f(0)$ è una primitiva di $f'(x)$, si ha

$$\begin{aligned} I_L^n(f') &= \int_{-L}^L \frac{1}{x^n} \left[f'(x) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(0)}{k!} x^k \right] dx \\ &= \left[\frac{1}{x^n} \left(f(x) - f(0) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(0)}{k!} \frac{x^{k+1}}{k+1} \right) \right]_{-L}^L \\ &\quad + \int_{-L}^L \frac{n}{x^{n+1}} \left(f(x) - f(0) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(0)}{k!} \frac{x^{k+1}}{k+1} \right) dx \\ &= \left[\frac{1}{x^n} \left(f(x) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right) \right]_{-L}^L + n \int_{-L}^L \frac{1}{x^{n+1}} \left(f(x) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right) dx \\ &= \frac{1}{L^n} \left(f(L) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} L^k \right) - \frac{1}{(-L)^n} \left(f(-L) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} (-L)^k \right) \\ &\quad + n I_L^{n+1}(f). \end{aligned}$$

Se $L > b$ si ha $f(L) = f(-L) = 0$, quindi

$$\begin{aligned} I_L^n(f') &= n I_L^{n+1}(f) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} \left(\frac{1}{(-L)^{n-k}} - \frac{1}{L^{n-k}} \right) \\ &= n I_L^{n+1}(f) - \sum_{\substack{k=0 \\ n-k \text{ dispari}}}^n \frac{f^{(k)}(0)}{k!} \frac{2}{L^{n-k}}. \end{aligned}$$

A questo punto possiamo fare il limite per $L \rightarrow \infty$, ottenendo

$$DP(1/x^n)(f) := -\lim_{L \rightarrow \infty} I_L^n(f') = -nP(1/x^{n+1})(f). \quad \square$$

6.60 Esempio. Vogliamo calcolare

$$(a) \quad D^3 [x^2 D^4 \log |x|] \qquad (b) \quad D^3 [x^4 D^2 \log |x|] .$$

(a) Poiché $(\log |x|)' = P(1/x)$, si ha

$$D^4 \log |x| = D^3 P(1/x) = (-1) \cdot (-2) \cdot (-3) P(1/x^4) = -6P(1/x^4) .$$

Quindi

$$\begin{aligned} D^3 [x^2 D^4 \log |x|] &= -6 D^3 [x^2 P(1/x^4)] = -6 D^3 [P(1/x^2)] \\ &= -6 \cdot (-24) P(1/x^5) = 144 P(1/x^5) . \end{aligned}$$

(b) Procedendo in modo analogo al caso precedente si ottiene

$$D^3 [x^4 D^2 \log |x|] = -D^3 [x^4 P(1/x^2)] = -D^3 (x^2) = 0 .$$

6.7 Problemi svolti

6.61 Problema. Calcolare le distribuzioni

$$\begin{array}{ll} (a) \quad D[\operatorname{sgn}(x) e^{-|x|}] & (b) \quad D[\operatorname{sgn}(x) \operatorname{sgn}(x-1)] \\ (c) \quad x(\operatorname{sgn}(x+2))' & (d) \quad \sin(x) |x|'' \end{array}$$

Soluzione. (a) Grazie alla Proposizione 6.54 e a quanto affermato nel Problema 6.42 ottengo

$$\begin{aligned} D[\operatorname{sgn}(x) e^{-|x|}] &= D(\operatorname{sgn}(x)) e^{-|x|} + \operatorname{sgn}(x) D(e^{-|x|}) \\ &= 2\delta_0 e^{-|x|} - (\operatorname{sgn}(x))^2 e^{-|x|} \\ &= 2\delta_0 e^{-|0|} - e^{-|x|} && \text{[uso: } (\operatorname{sgn}(x))^2 = 1] \\ &= 2\delta_0 - e^{-|x|} . \end{aligned}$$

(b) Analogamente al caso precedente si ottiene

$$\begin{aligned} D[\operatorname{sgn}(x) \operatorname{sgn}(x-1)] &= D(\operatorname{sgn}(x)) \operatorname{sgn}(x-1) + \operatorname{sgn}(x) D(\operatorname{sgn}(x-1)) \\ &= 2\delta_0 \operatorname{sgn}(x-1) + 2\operatorname{sgn}(x) \delta_1 \\ &= 2(\delta_0 \operatorname{sgn}(-1) + \operatorname{sgn}(1) \delta_1) \\ &= 2(\delta_1 - \delta_0) . \end{aligned}$$

(c)

$$x(\operatorname{sgn}(x+2))' = x 2\delta_{-2} = (-2) 2\delta_{-2} = -4\delta_{-2} .$$

(d)

$$\sin x |x|'' = \sin x \operatorname{sgn}(x)' = 2 \sin x \delta_0 = 2 \sin(0) \delta_0 = 0 .$$

6.62 Problema. Calcolare le seguenti distribuzioni, semplificando il più possibile il risultato

$$(a) \quad D^3 [\sin x D^4 (|x|)] \qquad (b) \quad D^3 (e^{x-x^2} \delta_0'') .$$

Soluzione. (a) Poiché $D^2(|x|) = D(\operatorname{sgn}(x)) = 2\delta_0$, ottengo

$$\sin x D^4(|x|) = 2 \sin x \delta_0''.$$

Usando l'identità $h(x) \delta_0'' = h(0) \delta_0'' - 2h'(0) \delta_0' + h''(0) \delta_0$, posso scrivere

$$2 \sin x \delta_0'' = 2 [\sin(0) \delta_0'' - 2 \cos(0) \delta_0' - \sin(0) \delta_0] = -4\delta_0'.$$

Quindi

$$D^3[\sin x D^4(|x|)] = -4\delta_0^{(4)}.$$

(b) Dall'identità

$$h(x) \delta_0'' = h(0) \delta_0'' - 2h'(0) \delta_0' + h''(0) \delta_0,$$

con

$$\begin{aligned} h(x) &= e^{x-x^2} & h(0) &= 1 \\ h'(x) &= h(x)(1-2x) & h'(0) &= 1 \\ h''(x) &= h'(x)(1-2x) - 2h(x) & h''(0) &= -1 \end{aligned}$$

si ottiene

$$e^{x-x^2} \delta_0'' = \delta_0'' - 2\delta_0' - \delta_0.$$

Dunque

$$D^3(e^{x-x^2} \delta_0'') = \delta_0^{(5)} - 2\delta_0^{(4)} - \delta_0^{(3)}.$$

6.63 Problema. Calcolare le seguenti distribuzioni, semplificando il più possibile il risultato

$$(a) D^4(\sin(|x|)) \qquad (b) D^2(|\sin x|)$$

Soluzione. (a) Uso il fatto che se una funzione f ha una parità definita (cioè è pari oppure dispari), allora

$$f(|x|) = \begin{cases} f(x) & \text{se } f \text{ è pari} \\ f(x) \operatorname{sgn} x & \text{se } f \text{ è dispari.} \end{cases}$$

Poiché il seno è dispari

$$D(\sin(|x|)) = D(\sin x \operatorname{sgn} x) = \cos x \operatorname{sgn} x + 2 \sin x \delta_0 = \cos x \operatorname{sgn} x.$$

Faccio la seconda derivata

$$D^2(\sin(|x|)) = D(\cos x \operatorname{sgn} x) = -\sin x \operatorname{sgn} x + 2 \cos x \delta_0 = -\sin x \operatorname{sgn} x + 2\delta_0.$$

Faccio la terza derivata

$$\begin{aligned} D^3(\sin(|x|)) &= D(-\sin x \operatorname{sgn} x + 2\delta_0) \\ &= -\cos x \operatorname{sgn} x - 2 \sin x \delta_0 + 2\delta_0' = -\cos x \operatorname{sgn} x + 2\delta_0'. \end{aligned}$$

Ottengo infine

$$\begin{aligned} D^4(\sin(|x|)) &= D(-\cos x \operatorname{sgn} x + 2\delta'_0) \\ &= \sin x \operatorname{sgn} x - 2 \cos x \delta_0 + 2\delta''_0 = \sin |x| - 2\delta_0 + 2\delta''_0. \end{aligned}$$

(b) Data una qualsiasi quantità reale X posso scrivere $|X| = X \operatorname{sgn} X$, quindi, ricordando che

$$D(\operatorname{sgn}(\sin x)) = 2 \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi},$$

ottengo

$$\begin{aligned} D^2(|\sin x|) &= D^2(\operatorname{sgn}(\sin x) \sin x) \\ &= D\left(2 \sin x \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi} + \operatorname{sgn}(\sin x) \cos x\right) \\ &= D\left(2 \sum_{k \in \mathbb{Z}} (-1)^k \sin(k\pi) \delta_{k\pi} + \operatorname{sgn}(\sin x) \cos x\right) \\ &= D(\operatorname{sgn}(\sin x) \cos x) \\ &= -\sin x \operatorname{sgn}(\sin x) + 2 \cos x \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi} \\ &= -\sin x \operatorname{sgn}(\sin x) + 2 \sum_{k \in \mathbb{Z}} (-1)^k \cos(k\pi) \delta_{k\pi} \\ &= -|\sin x| + 2 \sum_{k \in \mathbb{Z}} \delta_{k\pi} \quad \text{[uso: } \cos(k\pi) = (-1)^k \text{].} \end{aligned}$$

Per capire intuitivamente questo risultato si osservi il grafico della funzione $f(x) = |\sin x|$ in Fig. 6.3. *Localmente* questa funzione è uguale a $\sin x$ oppure a $-\sin x$, quindi in ogni caso deve valere $f'' = -f$. Fanno eccezione i punti $x_k = k\pi$ con k intero, in corrispondenza dei quali il grafico mostra degli “spigoli” dove f non è derivabile. Se comunque sviluppiamo f in serie di Taylor nell’intorno di ciascuno di questi punti usando il valore della derivata destra e sinistra di f otteniamo

$$\begin{aligned} f(x) &= f(x_k) + f'(x_k)(x - x_k) + \mathcal{O}((x - x_k)^2) \\ &= |\sin(k\pi)| + |x - k\pi| + \mathcal{O}((x - k\pi)^2) = |x - k\pi| + \mathcal{O}((x - k\pi)^2). \end{aligned}$$

Ma la derivata seconda di $|x - k\pi|$ è proprio $2\delta_{k\pi}$, il che spiega l’origine del termine $2 \sum_{k \in \mathbb{Z}} \delta_{k\pi}$ nella derivata seconda di $|\sin x|$.

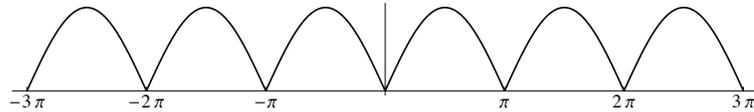


Figura 6.3: Grafico della funzione $|\sin x|$.

6.64 Problema. Siano n e k due interi positivi arbitrati e sia

$$g_n(x) := H(x) \frac{x^n}{n!}.$$

Calcolare, nel senso delle distribuzioni, $D^k g_n$ (Sugg: considerare separatamente i casi (1) $k < n$, (2) $k = n$ e (3) $k > n$).

Soluzione. Poiché $n > 0$, $x^n \delta_0 = 0$, quindi

$$Dg_n = \frac{x^n}{n!} \delta_0 + H(x) \frac{x^{n-1}}{(n-1)!} = H(x) \frac{x^{n-1}}{(n-1)!} = g_{n-1}.$$

Iterando questa identità ottengo che se $k \leq n$ si ha

$$D^k g_n = g_{n-k}.$$

Nel caso $k = n$ posso scrivere

$$D^n g_n = D(D^{n-1} g_n) = Dg_1 = D(xH(x)) = H(x) + x\delta_0 = H(x).$$

Grazie a questo risultato posso risolvere il caso $k > n$ nel modo seguente:

$$D^k g_n = D^{k-n} D^n g_n = D^{k-n} H = \delta_0^{k-n-1}.$$

6.65 Problema. Sia k un intero positivo. Calcolare $D^k(H(x)e^{-x})$ nel senso delle distribuzioni.

Soluzione. Sia $F(x) = H(x)e^{-x}$. Parto col caso più semplice, $k = 1$, e ottengo

$$DF(x) = D(H(x)e^{-x}) = \delta_0 e^{-x} - H(x)e^{-x} = \delta_0 - F(x).$$

Derivando ancora e riciclando il risultato sopra ottenuto si ha

$$D^2 F(x) = \delta_0' - F'(x) = \delta_0' - \delta_0 + F(x).$$

Ok, facciamone un'altra, ma ormai è chiaro qual è il risultato generale.

$$D^3 F(x) = D(\delta_0' - \delta_0 + F(x)) = \delta_0'' - \delta_0' + F'(x) = \delta_0'' - \delta_0' + \delta_0 - F(x).$$

È il momento di azzardare la congettura

$$D^k(H(x)e^{-x}) = (-1)^k \left[H(x)e^{-x} - \sum_{j=0}^{k-1} (-1)^j \delta_0^{(j)} \right]$$

che può essere facilmente dimostrata per induzione.

6.8 La distribuzione $\delta(b(x))$ (da non confondere con $\delta_0(b)!$)

Abbiamo visto che la delta di Dirac non è una distribuzione regolare, vale a dire non può essere associata ad una funzione localmente integrabile $\delta(x)$. Nonostante ciò viene comunemente usata la stessa notazione che si usa per le distribuzioni regolari, scrivendo

$$(6.46) \quad \int_{\mathbb{R}} f(x) \delta(x-a) dx = f(a).$$

In questa sezione vogliamo dare un senso matematico preciso ad un'altra notazione impropria che si incontra comunemente, in cui appare la *funzione composta* della "funzione" delta di Dirac che già di suo non esiste, per cui figuriamoci se esiste la funzione

composta. Comunque alla fine, come si vedrà, tutto ha un senso ben preciso. Voglio quindi attribuire un significato, ad esempio, all'espressione

$$(6.47) \quad \int_{\mathbb{R}} f(x) \delta(x^2 - 1) dx,$$

in cui compare appunto la composizione fra $\delta(x)$ e $b(x) = x^2 - 1$.

6.66 Attenzione! Non confondete $\delta(x^2 - 1)$ con $\delta_a(x^2 - 1)$: se scrivo $\delta_a(x^2 - 1)$ intendo la *distribuzione* δ_a applicata alla funzione di prova $f(x) = x^2 - 1$. Tralasciando il fatto (qui irrilevante) che $x^2 - 1$ non appartiene allo spazio \mathcal{K} perché non è a supporto compatto si ha quindi $\delta_a(x^2 - 1) = a^2 - 1$. Con il simbolo $\delta(x^2 - 1)$, senza pedice sulla δ , intendo invece la funzione composta della delta, che definiremo fra poco.

Per attribuire un significato alla (6.47) parto dalla Proposizione 6.22 che ci dice che la distribuzione δ_0 , anche se non è una distribuzione regolare, può tuttavia essere ottenuta (in molti modi) come *limite di una distribuzione regolare*. Ad esempio possiamo usare la gaussiana normalizzata

$$\gamma(x) := \frac{1}{\sqrt{\pi}} e^{-x^2}$$

ed ottenere la delta come limite della funzione γ opportunamente riscalata. Vale a dire, se definisco

$$\gamma_n(x) := n\gamma(nx) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}$$

allora si ha

$$(6.48) \quad \delta(x) = \lim_{n \rightarrow \infty} \gamma_n(x).$$

La (6.48) è una notazione simbolica che significa

$$(6.49) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(x) f(x) dx = f(0) \quad \forall f \in \mathcal{K}.$$

Usando la notazione impropria (6.46) possiamo riscrivere la (6.49) come

$$(6.50) \quad \int_{\mathbb{R}} f(x) \delta(x) dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(x) f(x) dx \quad \forall f \in \mathcal{K}.$$

Questa identità ci suggerisce un modo per definire una distribuzione $\delta(b(x))$ che, in un qualche senso, può essere pensata come la *composizione della delta per un'altra funzione* b .

6.67 Definizione. Sia $b \in C^1(\mathbb{R})$ una funzione con zeri semplici e isolati. Definisco

$$\int_{\mathbb{R}} f(x) \delta(b(x)) dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(b(x)) f(x) dx \quad f \in \mathcal{K}.$$

A questo punto faccio vedere che questa definizione rappresenta infatti una distribuzione, più precisamente una combinazione lineare di delta di Dirac centrate sugli zeri di b con coefficienti opportuni.

6.68 Proposizione. Sia $b \in C^1(\mathbb{R})$ e supponiamo che b abbia zeri semplici e isolati nei punti x_1, x_2, x_3, \dots , allora si ha, per ogni $f \in \mathcal{K}$,

$$(6.51) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(b(x)) f(x) dx = \sum_k \frac{\delta_{x_k}(f)}{|b'(x_k)|},$$

di conseguenza possiamo scrivere

$$(6.52) \quad \int_{\mathbb{R}} f(x) \delta(b(x)) dx = \sum_k \frac{\delta_{x_k}(f)}{|b'(x_k)|}.$$

Dimostrazione. Dimostriamo la Proposizione in un caso particolare in cui b è una funzione crescente ed ha uno zero che è quindi necessariamente unico. Sia x_0 il punto in cui b si annulla. Per ipotesi x_0 è uno zero semplice, quindi $b'(x_0) \neq 0$. Poiché b è crescente esistono i limiti

$$c = \lim_{x \rightarrow -\infty} b(x) < 0 \qquad d = \lim_{x \rightarrow +\infty} b(x) > 0$$

a patto che ammettiamo $-\infty$ come possibile valore di c e $+\infty$ come possibile valore di d . Usando il cambio di variabile

$$\begin{aligned} y &= b(x) \\ x &= b^{-1}(y) \\ dx &= Db^{-1}(y) dy = \frac{dy}{b'(b^{-1}(y))} \end{aligned}$$

e ricordando la (6.18), si ottiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(b(x)) f(x) dx &= \lim_{n \rightarrow \infty} \int_c^d \gamma_n(y) \frac{f(b^{-1}(y))}{b'(b^{-1}(y))} dy \\ &= \frac{f(b^{-1}(0))}{b'(b^{-1}(0))} = \frac{f(x_0)}{b'(x_0)} = \frac{\delta_{x_0}(f)}{|b'(x_0)|}. \quad \square \end{aligned}$$

L'identità (6.52) spesso viene scritta in maniera simbolica come

$$(6.53) \quad \delta(b(x)) = \sum_k \frac{1}{|b'(x_k)|} \delta(x - x_k) = \sum_k \frac{\delta_{x_k}}{|b'(x_k)|}.$$

6.69 Esempi. Vediamo con alcuni esempi come applicare la Proposizione precedente

- (1) Sia $a \neq 0$. Allora $\delta(ax) = \delta(b(x))$ in cui $b(x) = ax$ si annulla solo in $x = 0$ e vale $b'(0) = a$. La (6.53) diventa quindi

$$\delta(ax) = \frac{1}{|a|} \delta(x) \qquad a \neq 0.$$

- (2) $b(x) = \arctan x$. Anche in questo caso b si annulla soltanto nell'origine. Inoltre

$$b'(x) = \frac{1}{1+x^2} \qquad b'(0) = 1,$$

quindi

$$\delta(\arctan x) = \delta(x).$$

- (3) $b(x) = (x^2 - a^2)$, con $a \neq 0$. La funzione b ha 2 zeri in $\pm a$. inoltre

$$b'(x) = 2x \qquad |b'(\pm a)| = 2|a|$$

quindi

$$\delta(x^2 - a^2) = \frac{1}{2|a|} (\delta(x - a) + \delta(x + a)).$$

(4) $b(x) = e^x$. La funzione b non si annulla mai, quindi $\delta(e^x) = 0$.

(5) $b(x) = \cos x$. La funzione b si annulla in $x_k := \pi/2 + k\pi$ con $k \in \mathbb{Z}$. Inoltre

$$b'(x) = -\sin x \qquad |b'(x_k)| = |\sin(x_k)| = 1,$$

quindi

$$\delta(\cos x) = \sum_{k \in \mathbb{Z}} \delta\left(x - \frac{\pi}{2} - k\pi\right) = \sum_{k \in \mathbb{Z}} \delta\left(x - \frac{\pi}{2} + k\pi\right).$$

(6) $b(x) = \sin x$. Analogamente al caso precedente si ottiene

$$(6.54) \qquad \delta(\sin x) = \sum_{k \in \mathbb{Z}} \delta(x - k\pi) = \sum_{k \in \mathbb{Z}} \delta_{k\pi}.$$

6.70 Osservazione. (Alla fine tutto ha un senso). Pensare all'espressione

$$\int_{\mathbb{R}} f(x) \delta(b(x)) dx$$

come se coinvolgesse una “funzione composta” della delta di Dirac, pur essendo improprio dal punto di vista notazionale, riflette tuttavia un aspetto concettuale corretto. Supponiamo ad esempio di voler calcolare la derivata di $g(x) := \operatorname{sgn}(\sin x)$. Abbiamo già fatto questo calcolo nel Problema 6.63 e il risultato è

$$D(\operatorname{sgn}(\sin x)) = 2 \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi}.$$

Ora facciamo lo stesso conto in un altro modo: chiudiamo gli occhi, facciamo gli scongiuri e applichiamo la regola di derivazione delle funzioni composte sperando che vada tutto bene. Siccome $D \operatorname{sgn}(x) = 2\delta(x)$, applicando la (6.54), otteniamo

$$\begin{aligned} D(\operatorname{sgn}(\sin x)) &= \operatorname{sgn}'(\sin x) \cos x = 2\delta(\sin x) \cos x \\ &= 2 \cos x \sum_{k \in \mathbb{Z}} \delta_{k\pi} = 2 \sum_{k \in \mathbb{Z}} \cos x \delta_{k\pi} \\ &= 2 \sum_{k \in \mathbb{Z}} \cos(k\pi) \delta_{k\pi} = 2 \sum_{k \in \mathbb{Z}} (-1)^k \delta_{k\pi}. \end{aligned}$$

Sembra incredibile, ma tutto torna!

6.8.1 Problemi

6.71 Problema. Calcolare il seguente integrale scritto nella notazione “dei fisici”, semplificando il più possibile il risultato.

$$\int_{-\infty}^{\infty} e^{-|x|} \delta(\sin(2x)) dx$$

Soluzione. La funzione $b(x) := \sin(2x)$ si annulla nei punti

$$\sin(2x) = 0 \iff x_k = \frac{k\pi}{2} \text{ con } k \in \mathbb{Z}.$$

Poiché

$$|b'(x_k)| = |2 \cos(2x_k)| = |2 \cos(k\pi)| = |2(-1)^k| = 2,$$

si ottiene

$$\delta(\sin(2x)) = \frac{1}{2} \sum_{k \in \mathbb{Z}} \delta_{\kappa\pi/2}.$$

Quindi

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-|x|} \delta(\sin(2x)) dx &= \frac{1}{2} \sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} e^{-|x|} \delta(x - \kappa\pi/2) dx \\ &= \frac{1}{2} \sum_{k \in \mathbb{Z}} e^{-|k\pi/2|} = \sum_{k=0}^{\infty} e^{-k\pi/2} - \frac{1}{2} = \frac{1}{1 - e^{-\pi/2}} - \frac{1}{2}. \end{aligned}$$

6.72 Problema. Calcolare il seguente integrale scritto nella notazione “dei fisici”, in cui compare la “funzione composta della delta di Dirac”, semplificando il più possibile il risultato.

$$\int_{-\infty}^{\infty} \frac{\delta(\cos(\pi x/2))}{x^2} dx$$

Soluzione. La funzione $b(x) := \cos(\pi x/2)$ si annulla nei punti

$$\cos(\pi x/2) = 0 \iff x_k := 2k + 1 \text{ con } k \in \mathbb{Z}.$$

Poiché

$$|b'(x_k)| = \left| -\frac{\pi}{2} \sin\left(\frac{\pi x_k}{2}\right) \right| = \left| \frac{\pi}{2} \sin(k\pi + \pi/2) \right| = \frac{\pi}{2},$$

si ottiene

$$\delta(\cos(\pi x/2)) = \frac{2}{\pi} \sum_{k \in \mathbb{Z}} \delta_{x_k} = \frac{2}{\pi} \sum_{k \in \mathbb{Z}} \delta_{2k+1}.$$

Quindi

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\delta(\cos(\pi x/2))}{x^2} dx &= \frac{2}{\pi} \sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} \frac{\delta(x - x_k)}{x^2} dx \\ &= \frac{2}{\pi} \sum_{k \in \mathbb{Z}} \frac{1}{x_k^2} = \frac{2}{\pi} \sum_{k \in \mathbb{Z}} \frac{1}{(2k+1)^2} = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi}{2}, \end{aligned}$$

in cui, nell’ultima uguaglianza, ho usato il fatto noto che la somma dei quadrati degli interi positivi dispari è uguale a $\pi^2/8$.

6.73 Problema. Calcolare il seguente integrale scritto nella notazione “dei fisici”, in cui compare la “funzione composta della delta di Dirac”, semplificando il più possibile il risultato (fornire, se possibile, una risposta numerica)

$$\int_{-\infty}^{\infty} 2^{-|x|} \delta\left(\cos\left(\frac{\pi x}{3}\right) - \frac{1}{2}\right) dx.$$

Soluzione. La funzione $b(x) := \cos(\pi x/3) - 1/2$ si annulla nei punti

$$x_k = \pm 1 + 6k \quad \text{con } k \in \mathbb{Z}.$$

Poiché

$$|b'(x_k)| = |\pi/3 \sin(\pi x_k/3)| = \frac{\pi\sqrt{3}}{6} = \frac{\pi}{2\sqrt{3}}$$

si ottiene

$$\delta(\cos(\pi x/3) - 1/2) = \frac{2\sqrt{3}}{\pi} \sum_{k \in \mathbb{Z}} [\delta_{1+6k} + \delta_{-1+6k}].$$

Quindi

$$\begin{aligned}
 & \int_{-\infty}^{\infty} 2^{-|x|} \delta(\cos(\pi x/3) - 1/2) dx \\
 &= \frac{2\sqrt{3}}{\pi} \sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} 2^{-|x|} [\delta(x-1+6k) + \delta(x+1+6k)] dx \\
 &= \frac{2\sqrt{3}}{\pi} \sum_{k \in \mathbb{Z}} [2^{-|1+6k|} + 2^{-|-1+6k|}] = \frac{4\sqrt{3}}{\pi} \sum_{k \in \mathbb{Z}} 2^{-|1+6k|} \\
 &= \frac{4\sqrt{3}}{\pi} \left[\sum_{k=0}^{\infty} 2^{-(1+6k)} + \sum_{k=1}^{\infty} 2^{-(6k-1)} \right] \\
 &= \frac{4\sqrt{3}}{\pi} \left[\frac{2^{-1}}{1-2^{-6}} + \frac{2^1 2^{-6}}{1-2^{-6}} \right] = \frac{4\sqrt{3}}{\pi} \frac{2^5 + 2}{2^6 - 1} = \frac{136\sqrt{3}}{63\pi}.
 \end{aligned}$$

6.74 Problema. Sia $a \in \mathbb{R}$, $a > 0$ e sia

$$I(a) := \int_{-\infty}^{\infty} e^{-x^2+ix} \delta((x^2 - a^2) e^{-x^2}) dx.$$

Calcolare il massimo della funzione $f(a) := |aI(a)|$ per $a \in (0, \infty)$.

Soluzione. La funzione $b(x) := (x^2 - a^2) e^{-x^2}$ si annulla nei punti

$$x = \pm a.$$

Inoltre si ha

$$b'(x) = e^{-x^2} (2x - 2x(x^2 - a^2)) \quad |b'(\pm a)| = 2ae^{-a^2},$$

da cui si ottiene

$$\delta(b(x)) = \frac{e^{a^2}}{2a} (\delta_a + \delta_{-a}).$$

Quindi

$$\int_{-\infty}^{\infty} e^{-x^2+ix} \delta((x^2 - a^2) e^{-x^2}) dx = \frac{e^{a^2}}{2a} [e^{-a^2+ia} + e^{-a^2-ia}] = \frac{\cos(a)}{a}.$$

Otteniamo così che la funzione $f(a) = |aI(a)| = |\cos(a)|$ ha come valore massimo 1.

6.9 Alcune identità “notevoli” fra distribuzioni

6.75 Problema. Sia $\frac{1}{x-(x_0 \pm i0)}$ la distribuzione definita come

$$\frac{1}{x-(x_0 \pm i0)}(f) = \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}} \frac{f(x)}{x-(x_0 \pm i\varepsilon)} dx$$

Dimostrare che

$$\frac{1}{x-(x_0 \pm i0)} = P\left(\frac{1}{x-x_0}\right) \pm i\pi\delta_{x_0}$$

Soluzione. È sufficiente considerare il caso $x_0 = 0$. Dato $\alpha > 0$, posso scrivere

$$\begin{aligned}
 \int_{\mathbb{R}} \frac{f(x)}{x - i\varepsilon} dx &= \int_{-\alpha}^{\alpha} \frac{f(x)}{x - i\varepsilon} dx + \int_{|x| \geq \alpha} \frac{f(x)}{x - i\varepsilon} dx \\
 (6.55) \qquad &= \int_{-\alpha}^{\alpha} \frac{f(0)}{x - i\varepsilon} dx + \int_{-\alpha}^{\alpha} \frac{f(x) - f(0)}{x - i\varepsilon} dx \\
 (6.56) \qquad &+ \int_{|x| \geq \alpha} \frac{f(x)}{x} dx + \int_{|x| \geq \alpha} \left[\frac{f(x)}{x - i\varepsilon} - \frac{f(x)}{x} \right] dx \\
 &= A(\alpha) + B(\alpha) + C(\alpha) + D(\alpha),
 \end{aligned}$$

In cui $A(\alpha)$, $B(\alpha)$, $C(\alpha)$ e $D(\alpha)$ rappresentano i quattro integrali che compaiono nelle (6.55), (6.56). Iniziamo dall'ultimo.⁹

$$\begin{aligned}
 |D(\alpha)| &\leq \int_{|x| \geq \alpha} \left| \frac{f(x)}{x - i\varepsilon} - \frac{f(x)}{x} \right| dx \\
 &\leq \int_{|x| \geq \alpha} \frac{\varepsilon |f(x)|}{|x| |x - i\varepsilon|} dx \leq \int_{|x| \geq \alpha} \frac{\varepsilon |f(x)|}{|x|^2} dx,
 \end{aligned}$$

dove, nell'ultima disuguaglianza, abbiamo usato il fatto che $|x - i\varepsilon| = \sqrt{x^2 + \varepsilon^2} \geq |x|$. Di conseguenza, ponendo

$$\alpha := \varepsilon^{1/3}, \quad I := \int_{\mathbb{R}} |f(x)| dx$$

otteniamo

$$\begin{aligned}
 |D(\varepsilon^{1/3})| &\leq \int_{|x| \geq \varepsilon^{1/3}} \frac{\varepsilon |f(x)|}{|x|^2} dx \leq \int_{|x| \geq \varepsilon^{1/3}} \frac{\varepsilon |f(x)|}{\varepsilon^{2/3}} dx \\
 (6.57) \qquad &\leq \varepsilon^{1/3} \int_{\mathbb{R}} |f(x)| dx = \varepsilon^{1/3} I,
 \end{aligned}$$

per cui

$$\lim_{\varepsilon \rightarrow 0^+} D(\varepsilon^{1/3}) = 0.$$

Passiamo ad occuparci del termine B . Sviluppando f al prim'ordine,

$$f(x) = f(0) + x f'(\vartheta x) \quad \vartheta \in (0, 1)$$

e ricordando che $|x - i\varepsilon| \geq |x|$, si ottiene

$$\begin{aligned}
 |B(\varepsilon^{1/3})| &\leq \int_{-\varepsilon^{1/3}}^{\varepsilon^{1/3}} \frac{|f(x) - f(0)|}{|x - i\varepsilon|} dx \leq \int_{-\varepsilon^{1/3}}^{\varepsilon^{1/3}} \frac{|f(x) - f(0)|}{|x|} dx \\
 &\leq \|f'\|_u \int_{-\varepsilon^{1/3}}^{\varepsilon^{1/3}} 1 dx = 2 \|f'\|_u \varepsilon^{1/3}.
 \end{aligned}$$

Di conseguenza

$$(6.58) \qquad \lim_{\varepsilon \rightarrow 0^+} B(\varepsilon^{1/3}) = 0.$$

⁹uno di quei casi in cui “gli ultimi saranno i primi”

Per quanto riguarda il termine $C(\varepsilon^{1/3})$, si ottiene, ricordando che $\alpha(\varepsilon) = \varepsilon^{1/3}$,

$$\lim_{\varepsilon \rightarrow 0^+} C(\varepsilon^{1/3}) = \lim_{\varepsilon \rightarrow 0^+} \int_{|x| \geq \varepsilon^{1/3}} \frac{f(x)}{x} dx = \lim_{\alpha \rightarrow 0^+} \int_{|x| \geq \alpha} \frac{f(x)}{x} dx = P(1/x)(f).$$

Infine

$$\begin{aligned} A(\varepsilon^{1/3}) &= \int_{|x| \leq \varepsilon^{1/3}} \frac{f(0) dx}{x - i\varepsilon} \\ (6.59) \quad &= f(0) \int_{|x| \leq \varepsilon^{1/3}} \frac{x}{x^2 + \varepsilon^2} dx + i\varepsilon f(0) \int_{|x| \leq \varepsilon^{1/3}} \frac{dx}{x^2 + \varepsilon^2} \\ &= i\varepsilon f(0) \int_{|x| \leq \varepsilon^{1/3}} \frac{dx}{x^2 + \varepsilon^2} = 2i f(0) \arctan(\varepsilon^{1/3}/\varepsilon). \end{aligned}$$

Quindi

$$(6.60) \quad \lim_{\varepsilon \rightarrow 0^+} A(\varepsilon^{1/3}) = i\pi f(0).$$

Mettendo tutti i pezzi insieme ottengo

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}} \frac{f(x)}{x - i\varepsilon} dx &= \lim_{\varepsilon \rightarrow 0^+} [A(\varepsilon^{1/3}) + B(\varepsilon^{1/3}) + C(\varepsilon^{1/3}) + D(\varepsilon^{1/3})] \\ &= P(1/x)(f) + i\pi f(0) = [P(1/x) + i\pi\delta_0](f), \end{aligned}$$

che è esattamente quello che dovevo dimostrare. \square

6.76 Proposizione. *Sia*

$$h(x) := \frac{\sin x}{\pi x} \quad e \quad g_n(x) := nh(nx) = \frac{\sin(nx)}{\pi x}$$

Allora

$$\varphi_{g_n} \xrightarrow{\mathcal{K}^*} \delta_0.$$

Prima di dimostrare questo risultato, osserviamo che non può essere considerato un caso particolare della Proposizione 6.22. Infatti si può far vedere che la funzione h non appartiene a $C_1(\mathbb{R})$ vale a dire l'integrale $\int_{\mathbb{R}} |h(x)| dx$ è divergente. Nonostante ciò con metodi di analisi complessa si dimostra che l'integrale improprio esiste. Più precisamente si ha che

$$\lim_{K \rightarrow \infty} \int_{-K}^{+K} h(x) dx = 1.$$

Dimostrazione. Devo mostrare che

$$\varphi_{g_n}(f) \rightarrow \delta_0(f) \quad \forall f \in \mathcal{K}$$

ovvero che

$$(6.61) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) f(x) dx = f(0).$$

Possiamo scrivere

$$\begin{aligned} I(n) &:= \int_{\mathbb{R}} g_n(x) f(x) dx \\ &= \int_{|x| \leq \sqrt{n}} g_n(x) f(x) dx + \int_{|x| \geq \sqrt{n}} g_n(x) f(x) dx =: I_1(n) + I_2(n). \end{aligned}$$

Voglio far vedere che

$$I_1(n) \rightarrow f(0) \quad \text{e} \quad I_2(n) \rightarrow 0$$

da cui segue la (6.61). Il secondo limite è banale. Infatti, dato che $f \in \mathcal{K}$, esiste $a > 0$ tale che $\text{supp } f \subset [-a, a]$. Quindi abbiamo

$$\text{se } \sqrt{n} > a \text{ allora } I_2(n) = 0$$

per cui $I_2(n) \rightarrow 0$.

Rimane dunque da dimostrare che $I_1(n)$ converge a $f(0)$. Osserviamo innanzitutto che

$$\lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} g_n(x) dx = \lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} \frac{\sin(nx)}{\pi x} dx = \lim_{n \rightarrow \infty} \int_{|x| \leq n^{3/2}} \frac{\sin y}{\pi y} dy$$

D'altra parte è noto che

$$\lim_{K \rightarrow \infty} \int_{-K}^{+K} \frac{\sin y}{\pi y} dy = 1$$

quindi possiamo scrivere

$$f(0) = \lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} g_n(x) f(0) dx$$

Per dimostrare la Proposizione, è quindi sufficiente far vedere che

$$(6.62) \quad \lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} g_n(x) [f(x) - f(0)] dx = 0.$$

perché questo è equivalente a dire che $I_1(n) \rightarrow f(0)$. Se poniamo

$$h(x) := \frac{f(x) - f(0)}{x}$$

la (6.62) (moltiplicata per π) si può riscrivere come

$$(6.63) \quad \lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} \sin(nx) h(x) dx = 0.$$

Per dimostrare questo limite integriamo per parti

$$\int_{|x| \leq \sqrt{n}} \sin(nx) h(x) dx = -\frac{\cos(nx)}{n} h(x) \Big|_{-\sqrt{n}}^{+\sqrt{n}} + \int_{|x| \leq \sqrt{n}} \frac{\cos(nx)}{n} h'(x) dx,$$

da cui si ottiene

$$(6.64) \quad \left| \int_{|x| \leq \sqrt{n}} \sin(nx) h(x) dx \right| \leq \frac{2}{n} \|h\|_u + \int_{|x| \leq \sqrt{n}} \frac{|\cos(nx)|}{n} |h'(x)| dx \\ \leq \frac{2}{n} \|h\|_u + \frac{1}{n} \|h'\|_u 2\sqrt{n}$$

Dobbiamo ora preoccuparci del fatto che $\|h\|_u$ e $\|h'\|_u$ siano quantità finite. Dallo sviluppo di Taylor di f e di f' otteniamo che esistono $\vartheta_1(x)$, $\vartheta_2(x)$ e $\vartheta_3(x)$ tali che

$$f(x) = f(0) + x f'(\vartheta_1(x)) \\ f'(x) = f'(0) + x f''(\vartheta_2(x)) \\ f(x) = f(0) + x f'(0) + \frac{1}{2} x^2 f''(\vartheta_3(x)).$$

Di conseguenza

$$\begin{aligned} h(x) &= f'(\vartheta_1(x)) \\ h'(x) &= \frac{d}{dx} \frac{f(x) - f(0)}{x} = \frac{f'(x)x + f(0) - f(x)}{x^2} \\ &= \frac{f'(0)x + f(0) - f(x)}{x^2} + f''(\vartheta_2(x)) = \frac{1}{2}f''(\vartheta_3(x)) + f''(\vartheta_2(x)). \end{aligned}$$

Valgono quindi le seguenti stime

$$\|h\|_u \leq \|f'\|_u \quad \text{e} \quad \|h'\|_u \leq \frac{3}{2} \|f''\|_u.$$

Le grandezze $\|f'\|_u$ e $\|f''\|_u$ sono finite in virtù del fatto che $f \in \mathcal{K}$. Dalla (6.64) si ha quindi

$$\left| \int_{|x| \leq \sqrt{n}} \sin(nx) h(x) dx \right| \leq \frac{2}{n} \|f'\|_u + \frac{3}{\sqrt{n}} \|f''\|_u,$$

che implica

$$\lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} \sin(nx) h(x) dx = 0,$$

vale a dire

$$\lim_{n \rightarrow \infty} \int_{|x| \leq \sqrt{n}} \frac{\sin(nx)}{x} [f(x) - f(0)] dx = 0.$$

Abbiamo quindi dimostrato la (6.62) e, di conseguenza, la Proposizione □

Dalla Proposizione 6.76, mediante una traslazione, si ottiene che se

$$g_n(x) := \frac{\sin(n(x-y))}{\pi(x-y)} = \frac{\sin(n(y-x))}{\pi(y-x)}$$

allora

$$\varphi_{g_n} \xrightarrow{\mathcal{K}^*} \delta_y.$$

Ne segue che (simbolicamente)

$$\delta(x-y) = \lim_{n \rightarrow \infty} g_n(x) = \lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_{-n}^n dk e^{ik(y-x)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(y-x)}$$

È bene ricordare che il significato dell'identità simbolica

$$\delta(x-y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(y-x)}$$

è il seguente:

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{-n}^n dk e^{ik(y-x)} \right] f(x) dx = \delta_y(f) = f(y) \quad \forall f \in \mathcal{K}.$$

Si può far vedere che in realtà l'ordine di integrazione può essere invertito, per cui si ha anche

$$\frac{1}{2\pi} \int_{\mathbb{R}} e^{iky} \left[\int_{\mathbb{R}} e^{-ikx} f(x) dx \right] dk = \delta_y(f) = f(y) \quad \forall f \in \mathcal{K}.$$

Passiamo ora a mostrare una rappresentazione integrale della funzione ϑ .

6.77 Proposizione. Per ogni $x \in \mathbb{R}$, $x \neq 0$ si ha

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{-n}^{+n} \frac{e^{ikx}}{k - i\varepsilon} dk = \vartheta(x).$$

Dimostrazione.

Caso 1. $x > 0$. Se $x > 0$ allora l'integrale può essere "chiuso nel semipiano complesso superiore". In altre parole, sia γ_n^+ la curva nel piano complesso formata dal segmento $[-n, n]$ e dalla semicirconferenza $\{ne^{it} : t \in [0, \pi]\}$. Allora, grazie al Lemma di Jordan, il contributo all'integrale dovuto alla semicirconferenza tende a zero quando $n \rightarrow \infty$, per cui

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{-n}^{+n} \frac{e^{ikx}}{k - i\varepsilon} dk = \lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{\gamma_n^+} \frac{e^{izx}}{z - i\varepsilon} dz$$

D'altra parte¹⁰ se $\varepsilon > 0$ allora la funzione integranda ha un polo semplice all'interno della curva, quindi

$$\frac{1}{2\pi i} \int_{\gamma_n^+} \frac{e^{izx}}{z - i\varepsilon} dz = \frac{1}{2\pi i} 2\pi i \operatorname{Res} \left[\frac{e^{izx}}{z - i\varepsilon}, i\varepsilon \right] = e^{i(i\varepsilon)x} = e^{-\varepsilon x}.$$

Otteniamo in questo modo

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{-n}^{+n} \frac{e^{ikx}}{k - i\varepsilon} dk &= \lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{\gamma_n^+} \frac{e^{izx}}{z - i\varepsilon} dz \\ &= \lim_{\varepsilon \rightarrow 0^+} e^{-\varepsilon x} = 1 = \vartheta(x), \end{aligned}$$

in cui, nell'ultima uguaglianza, abbiamo usato ancora il fatto che $x > 0$.

Caso 2. $x < 0$. Si procede analogamente al caso 1, con la differenza che, poiché $x < 0$, l'integrale può essere chiuso nel semipiano complesso inferiore. Di conseguenza, siccome la singolarità dell'integrando si trova nel punto $z = i\varepsilon$, succede che non ci stanno singolarità all'interno della curva γ_n^- . Conseguentemente si avrà

$$\int_{\gamma_n^-} \frac{e^{izx}}{z - i\varepsilon} dz = 0 \quad \text{se } x < 0,$$

per cui

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{-n}^{+n} \frac{e^{ikx}}{k - i\varepsilon} dk = 0$$

in accordo col fatto che, se $x < 0$, $\vartheta(x) = 0$ □

6.10 Distribuzioni su \mathbb{R}^n

La maggior parte dei risultati ottenuti riguardanti le distribuzioni sull'asse reale possono essere estesi, con opportune modifiche, alle distribuzioni in n dimensioni. Se $\alpha = (\alpha_1, \dots, \alpha_p)$ è un multiindice in \mathbb{R}^n , cioè se $\alpha_i \in \{1, \dots, n\}$ per ogni $i = 1, \dots, p$, indico con $|\alpha|$ la somma di tutti gli indici $\sum_{i=1}^p \alpha_i$ e definisco la derivata parziale

$$\partial^\alpha f := \frac{\partial^{\alpha_1} \dots \partial^{\alpha_p} f}{\partial x_1^{\alpha_1} \dots \partial x_p^{\alpha_p}}.$$

¹⁰ricorda che vogliamo calcolare il limite $\varepsilon \rightarrow 0^+$

Indichiamo quindi con $\mathcal{K}(\mathbb{R}^n)$ lo spazio delle funzioni $f \in C^\infty(\mathbb{R}^n)$ a supporto compatto nel quale introduciamo una nozione di convergenza che è l'ovvia generalizzazione della Definizione 6.11, vale a dire:

6.78 Definizione. La successione $(f_n)_{n=1}^\infty$ di elementi di $\mathcal{K}(\mathbb{R}^n)$ si dice *convergente* ad $f \in \mathcal{K}(\mathbb{R}^n)$ e si scrive $f_n \xrightarrow{\mathcal{K}(\mathbb{R}^n)} f$ se

- (a) esiste un compatto $A \subset \mathbb{R}^n$ al di fuori del quale tutte le f_n sono nulle;
- (b) per ogni multiindice α si ha che $\partial^\alpha f_n \rightarrow \partial^\alpha f$ uniformemente su A .

A questo punto possiamo definire le distribuzioni multidimensionali.

6.79 Definizione. Una *distribuzione* su \mathbb{R}^n è un funzionale lineare continuo $F : \mathcal{K}(\mathbb{R}^n) \rightarrow \mathbb{C}$, vale a dire un funzionale lineare tale che

$$(6.65) \quad \text{se } f_n \xrightarrow{\mathcal{K}(\mathbb{R}^n)} f \text{ allora } F(f_n) \rightarrow F(f).$$

La somma fra distribuzioni e il prodotto fra una funzione C^∞ e una distribuzione si definiscono in modo identico al caso unidimensionale. Le derivate parziali sono definite in modo analogo, vale a dire, se α è un multiindice e F è una distribuzione su \mathbb{R}^n si ha

$$(\partial^\alpha F)(f) := (-1)^{|\alpha|} F(\partial^\alpha f).$$

Ad esempio se Δ è l'operatore Laplaciano

$$\Delta = \sum_{i=1}^n \partial_i^2$$

si ha

$$(6.66) \quad (\Delta F)(f) = \sum_{i=1}^n (\partial_i^2 F)(f) = \sum_{i=1}^n (-1)^2 F(\partial_i^2 f) = F(\Delta f).$$

Una differenza che vale la pena citare rispetto al caso unidimensionale riguarda la condizione di integrabilità locale. Nel caso $n = 1$ abbiamo detto che una singolarità del tipo $1/|x|^a$ è localmente integrabile se e solo se $a < 1$. Nel caso generale, l'elemento di volume n -dimensionale può essere espresso in coordinate sferiche come

$$d^n x = r^{n-1} dS_{n-1} dr$$

in cui dS_{n-1} è l'elemento di volume di una sfera $(n-1)$ -dimensionale, per cui

$$\int_{\mathbb{R}^n} \frac{f(x)}{|x|^a} d^n x = \int_0^\infty \left[\int_{S_{n-1}} f(x) dS_{n-1} \right] \frac{r^{n-1}}{r^a} dr.$$

Di conseguenza $1/|x|^a$ è localmente integrabile se e solo se $a < n$. In particolare $1/x$ e $1/|x|$ sono localmente integrabili e quindi rappresentano distribuzioni regolari su \mathbb{R}^n per ogni $n \geq 2$.

6.11 Applicazioni

6.11.1 Il potenziale elettrostatico

Consideriamo una carica puntiforme q posta nell'origine che genera quindi un potenziale¹¹

$$V(x) = \frac{q}{|x|} \quad x = (x_1, x_2, x_3) \in \mathbb{R}^3 \setminus \{0\}.$$

Possiamo parlare, in questa situazione, di densità di carica? Proviamo: se ρ è la densità di carica devono valere le seguenti relazioni

$$\rho(x) = \begin{cases} +\infty & \text{se } x = 0 \\ 0 & \text{se } x \neq 0. \end{cases} \quad \int_{\mathbb{R}^3} \rho(x) d^3x = q.$$

Viene quindi naturale congetturare $\rho(x) = q\delta(x)$ in cui δ è la delta di Dirac in \mathbb{R}^3 . D'altra parte sappiamo che il potenziale deve soddisfare l'equazione di Poisson

$$\Delta V = \nabla^2 V = -4\pi\rho.$$

Sostituendo $V = q/|x|$ e $\rho = q\delta$ nell'equazione di Poisson, si ottiene

$$(6.67) \quad \Delta \frac{1}{|x|} = -4\pi\delta(x).$$

Voglio dare una dimostrazione diretta che la (6.67) è corretta, purché sia intesa nel senso delle distribuzioni. In altre parole, grazie alla (6.66), bisogna far vedere che

$$(6.68) \quad \int_{\mathbb{R}^3} \frac{1}{|x|} \Delta f(x) dx = -4\pi f(0) \quad \forall f \in \mathcal{K}(\mathbb{R}^3).$$

Per produrre la (6.68) abbiamo bisogno dei seguenti ingredienti che potete facilmente reperire in un qualsiasi negozio di Analisi Matematica elementare:

- (1) l'elemento di volume in coordinate sferiche: $d^3x = dV = r^2 \sin \vartheta dr d\vartheta d\varphi$;
- (2) il Laplaciano in coordinate sferiche:

$$\Delta = \frac{1}{r^2} \partial_r (r^2 \partial_r) + \frac{1}{r^2 \sin \vartheta} \partial_\vartheta (\sin \vartheta \partial_\vartheta) + \frac{1}{r^2 \sin^2 \vartheta} \partial_\varphi^2.$$

- (3) la seconda identità di Green, che afferma, sotto opportune condizioni di regolarità che nel nostro caso sono ampiamente soddisfatte

$$\int_V [f\Delta g - (\Delta f)g] d^3x = \int_{\partial V} [f\nabla g - (\nabla f)g] \cdot d_n S$$

in cui ∂V denota la superficie che delimita il volume V e $d_n S$ è l'elemento di superficie, normale alla superficie e diretto verso l'esterno di V .

Ok, a questo punto la versione breve è semplicemente buttare tutti gli ingredienti nel frullatore, frullare per 10 minuti circa, e più o meno viene fuori la (6.68). La versione lunga, se qualcuno fosse insoddisfatto di quella breve, la trovate di seguito.

¹¹Usiamo ovviamente il sistema CGS

La prima cosa da fare è eliminare l'origine che crea problemi a causa di $1/|x|$. Siccome f è a supporto compatto, esiste $R > 0$ tale che f è nulla al di fuori della palla $B_{R/2}(0)$. Allora si ha

$$(6.69) \quad \int_{\mathbb{R}^3} \frac{1}{|x|} \Delta f(x) d^3x = \int_{|x| \leq R} \frac{1}{|x|} \Delta f(x) d^3x = \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon \leq |x| \leq R} \frac{1}{|x|} \Delta f(x) d^3x.$$

Sia $V(\varepsilon, R) = \{x \in \mathbb{R}^3 : \varepsilon \leq |x| \leq R\}$ il guscio sferico sul quale devo integrare. Il bordo di $V(\varepsilon, R)$ è costituito da due superfici sferiche di raggi rispettivamente ε e R . Dall'identità di Green ottengo

$$(6.70) \quad \int_{V(\varepsilon, R)} \frac{1}{|x|} \Delta f(x) d^3x = \int_{V(\varepsilon, R)} \Delta \frac{1}{|x|} f(x) d^3x + \int_{\partial V(\varepsilon, R)} \left[\frac{1}{|x|} \nabla f(x) - \nabla \frac{1}{|x|} f(x) \right] \cdot d_n S$$

Dall'espressione le Laplaciano in coordinate sferiche ottengo che, se $x \neq 0$ allora

$$\Delta \frac{1}{|x|} = \Delta \frac{1}{r} = \frac{1}{r^2} \partial_r \left(r^2 \partial_r \frac{1}{r} \right) = \frac{1}{r^2} \partial_r (-1) = 0,$$

quindi il primo addendo nel membro a destra della (6.70) è nullo. Il secondo addendo a sua volta è la somma dei due contributi costituiti dagli integrali sulle due superfici sferiche che compongono $\partial V(\varepsilon, R)$

$$(6.71) \quad \int_{\partial V(\varepsilon, R)} [\dots] \cdot d_n S = \int_{|x|=\varepsilon} [\dots] \cdot (-\hat{\mathbf{r}}) dS + \int_{|x|=R} [\dots] \cdot \hat{\mathbf{r}} dS,$$

in cui $\hat{\mathbf{r}}$ è il versore $x/|x|$ e il segno meno deriva dal fatto che $d_n S$ è diretto verso l'esterno del volume V per cui, quando integriamo sulla superficie sferica interna di raggio ε , $d_n S$ è diretto verso l'origine. Il secondo addendo della (6.71) è anch'esso nullo perché f è nulla al di fuori della palla di raggio $R/2$ e dunque sia f che ∇f sono nulli quando $|x| = R$. La (6.70) diventa quindi

$$(6.72) \quad \int_{V(\varepsilon, R)} \frac{1}{|x|} \Delta f(x) d^3x = - \int_{|x|=\varepsilon} \left[\frac{1}{|x|} \nabla f(x) - \nabla \frac{1}{|x|} f(x) \right] \cdot \hat{\mathbf{r}} dS.$$

Inoltre, poiché vale

$$\nabla \frac{1}{|x|} = \nabla \frac{1}{r} = -\frac{\mathbf{r}}{r^3} = -\frac{\hat{\mathbf{r}}}{r^2},$$

posso ulteriormente semplificare le (6.72). Denotando con $d\Omega = \sin \vartheta d\vartheta d\varphi$ l'elemento di angolo solido, posso scrivere

$$(6.73) \quad \begin{aligned} \int_{V(\varepsilon, R)} \frac{1}{|x|} \Delta f(x) d^3x &= - \int_{|x|=\varepsilon} \left[\frac{1}{r} \nabla f(x) + \frac{\hat{\mathbf{r}}}{r^2} f(x) \right] \cdot \hat{\mathbf{r}} r^2 d\Omega \\ &= - \int_{|x|=\varepsilon} \nabla f(x) \cdot \hat{\mathbf{r}} r d\Omega - \int_{|x|=\varepsilon} f(x) d\Omega. \end{aligned}$$

Il primo addendo nel membro a destra della (6.73) tende a 0 quando $\varepsilon \rightarrow 0^+$. Infatti

$$\begin{aligned}
 (6.74) \quad \left| \int_{|x|=\varepsilon} \nabla f(x) \cdot \hat{\mathbf{r}} r d\Omega \right| &= \left| \varepsilon \int_{|x|=\varepsilon} \nabla f(x) \cdot \hat{\mathbf{r}} d\Omega \right| \\
 &\leq \varepsilon \int_{|x|=\varepsilon} |\nabla f(x) \cdot \hat{\mathbf{r}}| d\Omega \\
 &\leq \varepsilon \int_{|x|=\varepsilon} |\nabla f(x)| d\Omega \\
 &\leq \varepsilon \sup_{x \in \mathbb{R}^3} |\nabla f(x)| \int_{|x|=\varepsilon} d\Omega \\
 &= \varepsilon \sup_{x \in \mathbb{R}^3} |\nabla f(x)| 4\pi \xrightarrow{\varepsilon \rightarrow 0^+} 0.
 \end{aligned}$$

Per quanto riguarda il secondo addendo nel membro a destra della (6.73) osservo che

$$\begin{aligned}
 (6.75) \quad \left| \int_{|x|=\varepsilon} f(x) d\Omega - 4\pi f(0) \right| &= \left| \int_{|x|=\varepsilon} [f(x) - f(0)] d\Omega \right| \\
 &\leq \int_{|x|=\varepsilon} |f(x) - f(0)| d\Omega \\
 &\leq \sup_{|x|=\varepsilon} |f(x) - f(0)| \int_{|x|=\varepsilon} d\Omega \\
 &= 4\pi \sup_{|x|=\varepsilon} |f(x) - f(0)|.
 \end{aligned}$$

Siccome f è continua si ha

$$\lim_{\varepsilon \rightarrow 0^+} \sup_{|x|=\varepsilon} |f(x) - f(0)| = 0,$$

per cui la (6.75) implica

$$(6.76) \quad \lim_{\varepsilon \rightarrow 0^+} \int_{|x|=\varepsilon} f(x) d\Omega = 4\pi f(0).$$

Finalmente, dalle relazioni (6.69), (6.73), (6.74), (6.76) si ottiene l'identità (6.68). \square

6.80 Problema. Sia $\rho \in C_c^\infty(\mathbb{R}^3)$ e sia V il potenziale generato dalla densità di carica ρ

$$V(x) := \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|} dy.$$

Dimostrare che

$$\Delta V(x) = -4\pi\rho(x).$$

Soluzione. Grazie a (6.67) si ha

$$\begin{aligned}
 \Delta V(x) &= \int_{\mathbb{R}^3} \Delta_x \frac{1}{|x-y|} \rho(y) dy \\
 &= \int_{\mathbb{R}^3} (-4\pi)\delta(y-x) \rho(y) dy = -4\pi\rho(x).
 \end{aligned}$$

6.81 Problema. Verificare che in \mathbb{R}^2 vale la seguente identità nel senso delle distribuzioni

$$\Delta \log|x| = 2\pi\delta(x).$$

6.11.2 Potenziali ritardati

6.82 Problema. (Potenziali ritardati). Sia $\rho \in C_c^\infty(\mathbb{R}^3)$ e sia

$$f(x, t) := \int_{\mathbb{R}^3} \frac{\rho(u, t - |x - u|)}{|x - u|} du \quad x \in \mathbb{R}^3, t \in \mathbb{R}$$

in cui $|x - u|$ è la usuale distanza in \mathbb{R}^3 data da $|x - u| := (\sum_{i=1}^3 |x_i - u_i|^2)^{1/2}$. Dimostrare che f soddisfa l'equazione

$$\textcircled{1} \quad [\partial_t^2 - \Delta] f(x, t) = 4\pi\rho(x, t)$$

in cui

$$\Delta = \sum_{i=1}^3 \partial_i^2 = \sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2}$$

Schema di soluzione. Si può scrivere

$$\begin{aligned} \Delta f(x, t) &= \int_{\mathbb{R}^3} \Delta \left[\frac{\rho(u, t - |x - u|)}{|x - u|} \right] du \\ &= \int_{\mathbb{R}^3} \Delta \left[\frac{1}{|x - u|} \right] \rho(u, t - |x - u|) du \\ &\quad + 2 \int_{\mathbb{R}^3} \nabla \left[\frac{1}{|x - u|} \right] \cdot \nabla \rho(u, t - |x - u|) du \\ &\quad + \int_{\mathbb{R}^3} \frac{1}{|x - u|} \Delta \rho(u, t - |x - u|) du \end{aligned}$$

L'identità $\textcircled{1}$ si ottiene (con uno zinzino di pazienza) ricordando che

(a) La funzione $\frac{1}{|x-u|}$ soddisfa (in 3 dimensioni!) la seguente identità in senso debole

$$\Delta \left[\frac{1}{|x - u|} \right] = -4\pi\delta(x - u)$$

(b) Dalle regole di derivazione delle funzioni composte segue che

$$\partial_i \rho(u, t - |x - u|) = -\partial_i \rho(u, t - |x - u|) \partial_i |x - u|$$

(c) Dalla definizione di $|x - u|$ e dalle usuali regole del calcolo si ottiene

$$\partial_i |x - u| = \frac{x_i - u_i}{|x - u|} \quad \Delta |x - u| = \frac{2}{|x - u|}$$

7. Operatori lineari

J: Now Yolanda, we're not gonna do anything stupid, are we?

Y: You don't hurt him.

J: Nobody's gonna hurt anybody. We're gonna be like three little Fonzie's here. And what's Fonzie like? Come on Yolanda what's Fonzie like?

Y: Cool?

J: What?

Y: He's cool.

Jules: Correctamundo. And that's what we're gonna be. We're gonna be cool. Now Ringo, I'm gonna count to three, and when I count three, you let go of your gun, and sit your ass down. But when you do it, you do it cool. Ready? One... two... three.

7.1 Definizione ed esempi

7.1 Definizione. Dati due spazi lineari normati $(V, \|\cdot\|_V)$ e $(Z, \|\cdot\|_Z)$, si definisce *operatore lineare* da V in Z un'applicazione $T : V \rightarrow Z$ tale che¹

$$T(c_1v_1 + c_2v_2) = c_1T(v_1) + c_2T(v_2) \quad \forall c_1, c_2 \in \mathbb{R} \quad \forall v_1, v_2 \in V$$

Nel caso in cui i due spazi V e Z coincidano T è detto un operatore lineare *su* V .

7.2 Definizione. L'operatore lineare T da V in Z è detto *continuo nel punto* $x \in V$ se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $\|x - y\|_V < \delta$ allora $\|Tx - Ty\|_Z < \varepsilon$.

7.3 Definizione. L'operatore lineare T da V in Z è detto *limitato* se

$$\sup_{x \in V: \|x\|_V \leq 1} \|Tx\|_Z < \infty.$$

La *norma* di T è definita come²

$$(7.1) \quad \|T\| := \sup_{x \in V: \|x\|_V \leq 1} \|Tx\|_Z = \sup_{x \neq 0} \frac{\|Tx\|_Z}{\|x\|_V}.$$

7.4 Proposizione. (*Condizioni equivalenti*). Sia T un operatore lineare da V in Z . Le seguenti condizioni sono equivalenti

- (1) T è continuo
- (2) T è continuo nell'origine
- (3) T è limitato

¹spesso è utile considerare operatori definiti non in tutto lo spazio V , ma in un sottospazio D_T detto dominio di T

²dimostriamo fra poco che è una vera norma

Dimostrazione. Dimostrazione identica al caso dei funzionali lineari (rimpiazzare $|F(x)|$ con $\|Tx\|_Z$). \square

7.5 Definizione. L'insieme di tutti gli operatori limitati da V in Z si denota con $\mathcal{L}(V, Z)$. L'insieme di tutti gli operatori limitati su V si denota con $\mathcal{L}(V)$.

7.6 Definizione. Sia $T \in \mathcal{L}(V, Z)$. L'insieme di tutti i vettori $v \in V$ tali che $Tv = 0$ si chiama *nucleo* di T e si denota con $\text{Ker } T$. L'insieme di tutti i vettori $z \in Z$ che si possono esprimere come $z = Tv$ per un qualche $v \in V$ si chiama *immagine* di T e si denota con $\text{Ran } T$.

7.7 Proposizione. Sia $T \in \mathcal{L}(V, Z)$. Allora

- (1) Il nucleo di T è un sottospazio chiuso di V .
- (2) L'immagine di T è un sottospazio di Z .

Dimostrazione. (Molto facile) \square

7.8 Problema. Consideriamo l'operatore lineare continuo T sullo spazio $(\ell_2, \|\cdot\|_2)$, definito come

$$T(x_1, x_2, x_3, \dots) := \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots\right).$$

Dimostrare che $\text{Ran } T$ non è chiuso in ℓ_2 .

Soluzione. Per dimostrare che $\text{Ran } T$ non è chiuso in ℓ_2 troverò una successione $(y^{(n)})$ di elementi di ℓ_2 tale che

- (a) $y^{(n)} \in \text{Ran } T$, vale a dire per ogni n esiste $x^{(n)} \in \ell_2$ tale che $Tx^{(n)} = y^{(n)}$
- (b) $(y^{(n)})$ converge, nella norma $\|\cdot\|_2$, ad un elemento y di ℓ_2 che *non* appartiene a $\text{Ran } T$.

Definisco dunque

$$y^{(n)} = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, 0, 0, 0, \dots\right).$$

Si ha

$$y^{(n)} = Tx^{(n)}$$

in cui

$$x^{(n)} = (1, 1, \dots, 1, 0, 0, \dots),$$

di conseguenza $y^{(n)} \in \text{Ran } T$. D'altra parte $(y^{(n)})$ è una successione convergente in ℓ_2 , infatti è facile verificare che

$$\lim_{n \rightarrow \infty} y^{(n)} = y := \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots\right).$$

Tuttavia il limite y della successione $(y^{(n)})$ non appartiene a $\text{Ran } T$, perché se si potesse scrivere $y = Tx$, allora dovrebbe essere

$$x = (1, 1, 1, 1, \dots)$$

che non appartiene a ℓ_2 . \square

7.1.1 Esempi

- (1) L'operatore identità $Ix = x$. $\|I\| = 1$. $\text{Ker } T = \{0\}$, $\text{Ran } T = V$.
- (2) Se V e W sono spazi vettoriali finito dimensionali tutti gli operatori lineari sono limitati. Inoltre se T è un operatore lineare da V in W , T è rappresentato da una matrice. Più precisamente se $(v_i)_{i=1}^n$ è una base di V e $(w_i)_{i=1}^m$ è una base di W , allora esiste un'unica matrice $m \times n$ A tale che

$$Tv_k = \sum_{i=1}^m w_i A_{ik}$$

Ricordo che vale l'uguaglianza

$$\dim V = \dim \text{Ker } T + \dim \text{Ran } T$$

- (3) Proiezioni ortogonali. Se W è un sottospazio chiuso di V possiamo definire l'operatore lineare $T = \pi_W$ che rappresenta la proiezione ortogonale su W . T è univocamente definito dalle due condizioni

$$Tv \in W \quad v - Tv \in W^\perp.$$

È facile far vedere che $\|\pi_W\| = 1$, $\text{Ker } T = W^\perp$, $\text{Ran } T = W$.

- (4) Dato uno spazio vettoriale normato di successioni (es: ℓ_p, ℓ_0, \dots) possiamo definire gli operatori di traslazione verso destra ϑ_+ e verso sinistra ϑ_- come

$$\begin{aligned} \vartheta_+(x_1, x_2, x_3, \dots) &:= (0, x_1, x_2, x_3, x_4, \dots) \\ \vartheta_-(x_1, x_2, x_3, \dots) &:= (x_2, x_3, x_4, \dots) \end{aligned}$$

e, più in generale, se n è un intero positivo

$$\begin{aligned} \vartheta_n(x_1, x_2, x_3, \dots) &:= (0, \dots, 0, x_1, x_2, x_3, \dots) \\ \vartheta_{-n}(x_1, x_2, x_3, \dots) &:= (x_{n+1}, x_{n+2}, x_{n+3}, \dots) \end{aligned}$$

Notare che valgono le identità

$$(\vartheta_+x)_i = \begin{cases} x_{i-1} & \text{se } i > 1 \\ 0 & \text{se } i = 1 \end{cases} \quad (\vartheta_-x)_i = x_{i+1}$$

È immediato mostrare che sono operatori limitati con norma uguale ad 1

- (5) Operatori integrali. Se $K \in C([a, b] \times [a, b])$, possiamo considerare l'operatore T , che agisce su $(C[a, b], \|\cdot\|_u)$, definito come

$$(Tf)(x) := \int_a^b K(x, y)f(y) dy.$$

Bisogna far vedere che Tf appartiene a $C[a, b]$ e che T è limitato. Dato che K è continua su un compatto, è anche uniformemente continua, quindi

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che se } |x - x'| + |y - y'| < \delta \text{ allora } |K(x, y) - K(x', y')| < \varepsilon.$$

Quindi se $|x - z| \leq \delta$,

$$\begin{aligned} |Tf(x) - Tf(z)| &\leq \int_a^b |K(x, y) - K(z, y)| |f(y)| dy \\ &\leq \sup_{y \in [a, b]} |K(x, y) - K(z, y)| \int_a^b |f(x)| dx \leq \varepsilon(b-a) \|f\|_u \end{aligned}$$

che implica la continuità di Tf . Per far vedere che T è limitato dobbiamo dimostrare che $\|Tf\|_u \leq C\|f\|_u$ per una qualche costante C . E infatti si ha per ogni $x \in [a, b]$

$$|Tf(x)| = \left| \int_a^b K(x, y) f(y) dy \right| \leq \int_a^b |K(x, y)| |f(y)| dy \leq \left[\int_a^b |K(x, y)| dy \right] \|f\|_u$$

Di conseguenza

$$\|Tf\|_u = \sup_{x \in [a, b]} |Tf(x)| \leq \left[\sup_{x \in [a, b]} \int_a^b |K(x, y)| dy \right] \|f\|_u$$

Abbiamo quindi dimostrato che T è limitato e abbiamo trovato anche un limite superiore esplicito per la sua norma

$$\|T\| \leq \sup_{x \in [a, b]} \int_a^b |K(x, y)| dy.$$

(6) Operatori differenziali. Sia D l'operatore derivata che agisce su $(C[a, b], \|\cdot\|_u)$. Osserviamo che:

- (a) D non è definito su tutto $C[a, b]$. Il suo dominio è costituito dalle funzioni $f \in C[a, b]$ che sono differenziabili
- (b) D non è continuo nell'origine. Consideriamo infatti la successione (f_n) , in cui $f_n(x) := \sin(nx)/n$. Allora si ha $Df_n(x) = \cos(nx)$, per cui $\|f_n\|_u$ tende a zero, ma $\|Df_n\|_u = 1 \not\rightarrow 0$.

(7) Operatori moltiplicativi. Nello spazio $\mathcal{S}(\mathbb{R})$ possiamo definire l'operatore "moltiplicativo" X come

$$X(f) := x f(x) \quad f \in \mathcal{S}(\mathbb{R}).$$

Dimostrare che, scegliendo ad esempio la norma uniforme $\|\cdot\|_u$, l'operatore X non è limitato. Più in generale se P è un polinomio possiamo considerare l'operatore

$$P(X)(f) := P(x) f(x) \quad f \in \mathcal{S}(\mathbb{R}).$$

L'operatore $P(X)$ viene generalmente denotato con $P(x)$ confondendolo con la funzione ad esso associata.

(8) Operatori di Schrödinger. In meccanica quantistica si studiano operatori del tipo

$$H := -\frac{\hbar^2}{2} \Delta + V(x)$$

che agiscono nello spazio $(C_2^\infty(\mathbb{R}^n), \|\cdot\|_2)^3$ come

$$(Hf)(x) = -\frac{\hbar^2}{2} \Delta f(x) + V(x) f(x) \quad f \in C_2^\infty(\mathbb{R}^n)$$

7.9 Problema. Sia K una funzione continua su $[a, b] \times [a, b]$. Sia T l'operatore su $C_2[a, b]$ definito come

$$Tf(x) := \int_a^b K(x, y) f(y) dy.$$

³più precisamente nel suo completamento $L_2(\mathbb{R}^n)$

Dimostrare che T è limitato e trovare un limite superiore alla norma di T .

Soluzione. Devo dimostrare che esiste una costante $C > 0$ tale che

$$\|Tf\|_2 \leq C \|f\|_2.$$

Questo mi dice che T è limitato e che $\|T\| \leq C$. Per stimare la quantità $\|Tf\|_2$ procedo nel modo seguente

$$\textcircled{1} \quad \|Tf\|_2^2 := \int_a^b |Tf(x)|^2 dx = \int_a^b \left| \int_a^b K(x, y) f(y) dy \right|^2 dx$$

Per un dato valore di x definisco la funzione

$$g_x(y) := K(x, y).$$

Grazie alla disuguaglianza di Cauchy-Schwarz posso scrivere

$$\begin{aligned} \textcircled{2} \quad \left| \int_a^b K(x, y) f(y) dy \right|^2 &= \left| \int_a^b g_x(y) f(y) dy \right|^2 = |\langle g_x, f \rangle|^2 \\ &\leq \|g_x\|_2^2 \|f\|_2^2 = \|f\|_2^2 \int_a^b K(x, y)^2 dy. \end{aligned}$$

Dalla $\textcircled{1}$ e dalla $\textcircled{2}$ ottengo

$$\|Tf\|_2^2 \leq \int_a^b \left[\|f\|_2^2 \int_a^b K(x, y)^2 dy \right] dx = \|f\|_2^2 \int_a^b \int_a^b K(x, y)^2 dx dy.$$

Abbiamo così ricavato la disuguaglianza

$$\|Tf\|_2 \leq C \|f\|_2,$$

in cui

$$C := \left[\int_a^b \int_a^b K(x, y)^2 dx dy \right]^{1/2}.$$

Di conseguenza T è limitato e vale

$$\|T\| \leq \left[\int_a^b \int_a^b K(x, y)^2 dx dy \right]^{1/2}. \quad \square$$

7.10 Problema. Nello spazio $(C_c(\mathbb{R}), \|\cdot\|_u)$ consideriamo l'operatore lineare T definito come $(Tf)(x) := xf(x)$.

- (a) Dimostrare che T non è limitato.
 (b) Determinare il nucleo e l'immagine di T .

Soluzione. (a). Definisco una successione di funzioni $f_n \in C_b(\mathbb{R})$ scelte in modo tale che

$$\textcircled{1} \quad \lim_{n \rightarrow \infty} \frac{\|Tf_n\|_u}{\|f_n\|_u} = \infty,$$

il che implica che T è illimitato. Definisco dunque f_n nel modo seguente

$$f_n(x) := \begin{cases} x - n + 1 & \text{se } x \in [n - 1, n] \\ -x + n + 1 & \text{se } x \in [n, n + 1] \\ 0 & \text{altrimenti} \end{cases}$$

Si vede allora che $\|f_n\|_u = 1$, e che

$$\|Tf_n\|_u = \sup_{x \in \mathbb{R}} |xf_n(x)| \geq |nf_n(n)| = |n \cdot 1| = n,$$

da cui segue la ① e il fatto che T è illimitato.

(b). Supponiamo che sia $Tf = 0$. Allora, per definizione, si ha

$$xf(x) = 0 \quad \forall x \in \mathbb{R},$$

il che implica che $f(x)$ è uguale a zero per tutti gli $x \neq 0$. Ma f è una funzione continua, quindi se è nulla per tutti gli $x \neq 0$ è necessariamente nulla anche per $x = 0$, vale a dire f è la funzione identicamente nulla. Per cui $\text{Ker } T = \{0\}$.

Per determinare $\text{Ran } T$ bisogna capire come deve essere fatta una funzione $g \in C_c(\mathbb{R})$ affinché possa essere ottenuta come $g = Tf$ per una qualche f . Scriviamo quindi l'equazione $Tf = g$ in modo esplicito come

$$\textcircled{2} \quad xf(x) = g(x) \quad \forall x \in \mathbb{R}$$

e cerchiamo di risolverla per f . È chiaro che, affinché valga la ②, deve essere

$$f(x) = \frac{g(x)}{x} \quad \forall x \neq 0.$$

Ma f deve essere continua su tutto l'asse reale, in particolare in $x = 0$, quindi g deve essere tale che

$$\textcircled{3} \quad \text{il limite } \lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} g(x)/x \text{ esiste ed è finito.}$$

Poiché g è continua la ③ implica che $g(0) = 0$. Di conseguenza

$$\lim_{x \rightarrow 0} \frac{g(x)}{x} = \lim_{x \rightarrow 0} \frac{g(x) - g(0)}{x} = g'(0).$$

La condizione ③ equivale quindi ad affermare che

$$g(0) = 0 \text{ e } g \text{ è derivabile nell'origine.}$$

Se questa condizione è soddisfatta possiamo definire $f \in C_c(\mathbb{R})$ come

$$f(x) := \begin{cases} g(x)/x & \text{se } x \neq 0 \\ g'(0) & \text{se } x = 0. \end{cases}$$

La f così definita risolve l'equazione $Tf = g$. Abbiamo quindi dimostrato che

$$\text{Ran } T := \{g \in C_c(\mathbb{R}) : g(0) = 0 \text{ e } g \text{ è derivabile nell'origine}\}. \quad \square$$

7.2 Somme e prodotti di operatori lineari

7.11 Definizione. Dati $S, T \in \mathcal{L}(V, Z)$ e dato $\alpha \in \mathbb{R}$ definisco gli operatori $S + T$ e αT da V in Z come

$$\begin{aligned} (S + T)(u) &:= Su + Tu & u \in V \\ (\alpha T)(u) &:= \alpha(Tu) & u \in V. \end{aligned}$$

7.12 Proposizione. *L'insieme $\mathcal{L}(V, Z)$ è uno spazio vettoriale. Inoltre la funzione $\|\cdot\| : \mathcal{L}(V, Z) \rightarrow \mathbb{R}$ introdotta nella (7.1) è una norma su $\mathcal{L}(V, Z)$.*

Dimostrazione. La prima affermazione è ovvia. Passiamo alla seconda. È banale mostrare che $\|T\| \geq 0$, che $\|T\| = 0$ se e solo se $T = 0$ e che $\|\alpha T\| = |\alpha| \|T\|$ per ogni $\alpha \in \mathbb{R}$. Infine mostriamo che vale la disuguaglianza triangolare:

$$\begin{aligned} \|S + T\| &:= \sup_{x \neq 0} \frac{\|(S + T)(x)\|_Z}{\|x\|_V} = \sup_{x \neq 0} \frac{\|Sx + Tx\|_Z}{\|x\|_V} \leq \sup_{x \neq 0} \frac{\|Sx\|_Z + \|Tx\|_Z}{\|x\|_V} \\ &\leq \sup_{x \neq 0} \frac{\|Sx\|_Z}{\|x\|_V} + \sup_{x \neq 0} \frac{\|Tx\|_Z}{\|x\|_V} = \|S\| + \|T\| \end{aligned}$$

quindi $\|S + T\| \leq \|S\| + \|T\|$.

7.13 Definizione. (Prodotto fra operatori). Se $S \in \mathcal{L}(V, W)$ e $T \in \mathcal{L}(W, Z)$ si definisce il *prodotto* o *composizione* TS come

$$\begin{aligned} TS &: V \rightarrow Z \\ TS(x) &:= T(Sx) \end{aligned}$$

7.14 Proposizione. *Se $S \in \mathcal{L}(V, W)$ e $T \in \mathcal{L}(W, Z)$ allora TS è un operatore lineare continuo da V in Z . Inoltre*

$$(7.2) \quad \|TS\| \leq \|T\| \|S\|$$

Dimostrazione. È sufficiente far vedere che vale la (7.2).

$$\|TS\| := \sup_{x \neq 0} \frac{\|TSx\|_Z}{\|x\|_V} \leq \sup_{\substack{x \neq 0 \\ Sx \neq 0}} \frac{\|TSx\|_Z}{\|x\|_V} + \sup_{\substack{x \neq 0 \\ Sx = 0}} \frac{\|TSx\|_Z}{\|x\|_V}.$$

Il secondo termine nella precedente somma è nullo, perché se $Sx = 0$ allora $TSx = 0$. Possiamo quindi scrivere

$$\begin{aligned} \|TS\| &:= \sup_{x \in V: Sx \neq 0} \frac{\|TSx\|_Z}{\|x\|_V} = \sup_{x \in V: Sx \neq 0} \frac{\|TSx\|_Z}{\|Sx\|_W} \frac{\|Sx\|_W}{\|x\|_V} \\ &\leq \sup_{x \in V: Sx \neq 0} \frac{\|TSx\|_Z}{\|Sx\|_W} \sup_{x \in V: Sx \neq 0} \frac{\|Sx\|_W}{\|x\|_V} \\ &\leq \sup_{y \in \text{Ran } S: y \neq 0} \frac{\|Ty\|_Z}{\|y\|_W} \sup_{x \in V: x \neq 0} \frac{\|Sx\|_W}{\|x\|_V} \\ &\leq \sup_{y \in W: y \neq 0} \frac{\|Ty\|_Z}{\|y\|_W} \sup_{x \in V: x \neq 0} \frac{\|Sx\|_W}{\|x\|_V} =: \|T\| \|S\| \quad \square \end{aligned}$$

7.15 Proposizione. *Se Z è completo allora $(\mathcal{L}(V, Z), \|\cdot\|)$ è completo.*

Dimostrazione. Sia $(T_k)_{k=1}^\infty$ una successione di Cauchy in $\mathcal{L}(V, Z)$ allora, per definizione

$$\text{per ogni } \varepsilon > 0 \text{ esiste } N > 0 \text{ tale che se } n, m \geq N \text{ si ha } \|T_n - T_m\| < \varepsilon.$$

Dato $x \in V$, otteniamo

$$(7.3) \quad \|T_n x - T_m x\|_Z \leq \|T_n - T_m\| \|x\|_V \leq \varepsilon \|x\|_V \quad \forall n, m \geq N$$

quindi, per ogni $x \in V$, la successione $(T_n x)$ è una successione di Cauchy di elementi di Z . Poiché Z è completo $T_n x$ è convergente. Poniamo

$$T(x) := \lim_{n \rightarrow \infty} T_n(x)$$

T è lineare, infatti

$$T(\alpha x + \beta y) = \lim_{n \rightarrow \infty} T_n(\alpha x + \beta y) = \alpha \lim_{n \rightarrow \infty} T_n(x) + \beta \lim_{n \rightarrow \infty} T_n(y) = \alpha T x + \beta T y.$$

Per dimostrare che T è limitato, passiamo al limite $m \rightarrow \infty$ nella (7.3) ottenendo

$$(7.4) \quad \|(T_n - T)x\|_Z = \|T_n x - T x\|_Z \leq \varepsilon \|x\|_V \quad \forall n \geq N$$

Quindi $T - T_n$ è un operatore limitato. Ma allora $T = T_n + (T - T_n)$ è anch'esso limitato. \square

7.16 Proposizione. *Sia V uno spazio di Banach, $T \in \mathcal{L}(V)$ e sia $(c_k)_{k=0}^\infty$ una successione di numeri reali (o complessi) tale che*

$$(7.5) \quad \sum_{k=0}^{\infty} |c_k| \|T\|^k < \infty.$$

Allora esiste $S \in \mathcal{L}(V)$ tale che

$$(7.6) \quad \sum_{k=0}^{\infty} c_k T^k = S$$

Dimostrazione. Consideriamo le somme parziali

$$S_n := \sum_{k=0}^n c_k T^k$$

La successione S_n è di Cauchy, infatti, se $n < m$, si ha

$$\|S_n - S_m\| = \left\| \sum_{k=n+1}^m c_k T^k \right\| \leq \sum_{k=n+1}^m |c_k| \|T\|^k \leq \sum_{k=n+1}^{\infty} |c_k| \|T\|^k$$

Grazie alla (7.5) sappiamo che

$$\lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} |c_k| \|T\|^k = 0,$$

quindi (S_n) è di Cauchy. Poiché $\mathcal{L}(V)$ è completo esiste un operatore $S \in \mathcal{L}(V)$ tale che $S_n \rightarrow S$. \square

7.3 Operatore inverso

7.17 Definizione. Un operatore lineare T da V in Z è detto *invertibile* se T è un'applicazione biunivoca da V su Z , vale a dire se

$$(7.7) \quad \text{per ogni } y \in Z \text{ esiste un } \underline{\text{unico}} \text{ elemento } x \in V \text{ tale che } y = T x.$$

Equivalentemente T è invertibile se $\text{Ran } T = Z$ e $\text{Ker } T = \{0\}$. Se T è invertibile possiamo definire quindi l'*operatore inverso* $T^{-1} : Z \rightarrow V$. Se $y \in Z$, $T^{-1}y$ è definito univocamente dalla proprietà

$$(7.8) \quad T^{-1}y = x \quad \Leftrightarrow \quad T x = y.$$

7.18 Proposizione. *Se T è un operatore lineare invertibile da V in Z allora T^{-1} è un operatore lineare invertibile da Z in V .*

Dimostrazione. Facciamo vedere che T^{-1} è lineare, vale a dire che

$$(7.9) \quad T^{-1}(\alpha u + \beta v) = \alpha T^{-1}u + \beta T^{-1}v \quad \forall \alpha, \beta \in \mathbb{R}, \forall u, v \in Z$$

Poniamo

$$u' := T^{-1}u \quad v' := T^{-1}v.$$

Per definizione di T^{-1} si ha

$$Tu' = u \quad Tv' = v$$

Siccome T è lineare

$$T(\alpha u' + \beta v') = \alpha Tu' + \beta Tv'$$

che implica

$$\alpha u' + \beta v' = T^{-1}(\alpha Tu' + \beta Tv')$$

che è proprio la (7.9) □

Dalla definizione (7.8) segue immediatamente che

$$TT^{-1}y = y \quad \forall y \in Z \quad T^{-1}Tx = x \quad \forall x \in V,$$

ovvero che

$$TT^{-1} = I_Z \quad T^{-1}T = I_V.$$

7.19 Problema. Trovare due operatori lineari limitati in uno spazio di Banach V tali che $AB = I$, ma $BA \neq I$ (quindi B non è l'inverso di A).

7.20 Problema. Sia T è un operatore lineare invertibile da V in Z . È vero che se T è limitato allora T^{-1} è limitato?

Soluzione. La risposta è no. Infatti consideriamo il caso $V = Z = \ell_f$ con la norma $\|\cdot\|_\infty$. Se $x = (x_i)_{i=1}^\infty$ definiamo

$$Tx := \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \frac{x_4}{4}, \dots\right)$$

T è chiaramente lineare. Inoltre T è limitato, perché $\|Tx\|_\infty \leq \|x\|_\infty$, quindi $\|T\| \leq 1$. Infine T è biunivoco, quindi invertibile e si ha

$$T^{-1}x = (x_1, 2x_2, 3x_3, 4x_4, \dots)$$

è ovvio che T^{-1} non è limitato infatti ponendo

$$e^{(n)} := (0, 0, \dots, 0, \underset{[n]}{1}, 0, 0, \dots)$$

si ha

$$\frac{\|T^{-1}e^{(n)}\|_\infty}{\|e^{(n)}\|_\infty} = n \rightarrow \infty \quad \square$$

Quindi l'inverso di un operatore limitato non è in generale limitato ... però, aggiungendo alle ipotesi l'ingrediente della completezza le cose cambiano

7.21 Teorema. (*Teorema sull'operatore inverso*). Siano $(V, \|\cdot\|_V)$ e $(Z, \|\cdot\|_Z)$ due spazi di Banach e sia $T \in \mathcal{L}(V, Z)$ invertibile. Allora T^{-1} è limitato, cioè $T^{-1} \in \mathcal{L}(Z, V)$.

(senza dimostrazione ... richiede Baire cat thm).

7.22 Proposizione. Sia $(V, \|\cdot\|)$ uno spazio di Banach e sia $A \in \mathcal{L}(V)$ con $\|A\| < 1$. Allora $(I - A)^{-1}$ esiste, è limitato e si può scrivere come

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

Dimostrazione. Poiché $\|A\| =: \rho < 1$ si ha

$$(7.10) \quad \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1 - \rho} < \infty.$$

Dalla (7.10) e dalla Proposizione 7.16 segue che posso definire un operatore $B \in \mathcal{L}(V)$ come

$$B := \sum_{k=0}^{\infty} A^k$$

Per dimostrare che B è l'inverso di $I - A$ devo far vedere che

$$B(I - A) = (I - A)B = I.$$

A questo scopo osservo che

$$B = \lim_{n \rightarrow \infty} \sum_{k=0}^n A^k,$$

quindi

$$(7.11) \quad B(I - A) = \lim_{n \rightarrow \infty} \sum_{k=0}^n (A^k - A^{k+1}) = \lim_{n \rightarrow \infty} (I - A^{n+1}) = I - \lim_{n \rightarrow \infty} A^{n+1}$$

L'ultimo limite nella precedente equazione è nullo, infatti

$$\|A^n\| \leq \|A\|^n = \rho^n,$$

per cui

$$(7.12) \quad \lim_{n \rightarrow \infty} \|A^n\| \leq \lim_{n \rightarrow \infty} \rho^n = 0.$$

Dalle (7.11), (7.12) ottengo

$$B(I - A) = I.$$

In modo analogo si dimostra che $(I - A)B = I$, da cui segue che $(I - A)$ è invertibile e che $B = (I - A)^{-1}$. \square

7.4 Operatore aggiunto (di Hilbert)

In questa sottosezione V è uno spazio euclideo con prodotto scalare $\langle \cdot, \cdot \rangle$.

7.23 Definizione. Dato un operatore lineare $T \in \mathcal{L}(V)$, l'operatore *aggiunto*⁴ T^* è definito dalle identità

$$\langle T^*u, v \rangle = \langle u, Tv \rangle \quad \forall u, v \in V.$$

⁴detto a volte *hermitiano coniugato*. Questo è l'aggiunto di Hilbert. C'è anche l'aggiunto di Banach che è una cosa leggermente diversa

7.24 Definizione. Un operatore lineare limitato T su V è detto *autoaggiunto*⁵ se $T^* = T$, vale a dire se

$$\langle Tu, v \rangle = \langle u, Tv \rangle \quad \forall u, v \in V.$$

7.25 Esempio. Sia M la matrice (finita o infinita) associata all'operatore T rispetto ad una base ortonormale $(e_k)_{k=1}^{\infty}$ di uno spazio euclideo *complesso* V , e sia M^* la matrice associata all'operatore aggiunto T^* . Allora si ha

$$Te_i = \sum_{k=1}^{\infty} e_k M_{ki} \qquad T^*e_i = \sum_{k=1}^{\infty} e_k M_{ki}^*.$$

per cui

$$\begin{aligned} \langle Te_i, e_j \rangle &= \sum_{k=1}^{\infty} M_{ki} \langle e_k, e_j \rangle = M_{ji} \\ \langle T^*e_i, e_j \rangle &= \sum_{k=1}^{\infty} M_{ki}^* \langle e_k, e_j \rangle = M_{ji}^*. \end{aligned}$$

Di conseguenza

$$(M^*)_{ij} = \langle T^*e_j, e_i \rangle = \langle e_j, Te_i \rangle = \overline{\langle Te_i, e_j \rangle} = \overline{M_{ji}}$$

quindi la matrice associata a T^* è la complessa coniugata della trasposta di M .

7.26 Problema. Dimostrare che in ℓ_2 vale $\vartheta_+^* = \vartheta_-$.

Soluzione. Per definizione abbiamo, per ogni $x, y \in \ell_2$,

$$\begin{aligned} \langle \vartheta_+^* x, y \rangle &= \langle x, \vartheta_+ y \rangle = \sum_{i=1}^{\infty} x_i (\vartheta_+ y)_i = \sum_{i=2}^{\infty} x_i y_{i-1} = \sum_{k=1}^{\infty} x_{k+1} y_k \\ &= \sum_{k=1}^{\infty} (\vartheta_- x)_k y_k = \langle \vartheta_- x, y \rangle \end{aligned}$$

Quindi

$$\langle \vartheta_+^* x, y \rangle = \langle \vartheta_- x, y \rangle \quad \forall x, y \in \ell_2,$$

che implica $\vartheta_+^* = \vartheta_-$ □

7.27 Problema. Nello spazio $C_2([a, b]; \mathbb{C})$ sia T l'operatore lineare integrale con nucleo $K(x, y)$. Determinare il nucleo integrale di T^* .

7.28 Proposizione. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e siano $S, T \in \mathcal{L}(V)$

- (1) $(S + T)^* = S^* + T^*$
- (2) $(\alpha T)^* = \bar{\alpha} T^*$
- (3) $(ST)^* = T^* S^*$
- (4) $T^{**} = T$
- (5) $\|T^*\| = \|T\|$
- (6) Se T ha un inverso limitato allora anche T^* ha un inverso limitato e $(T^*)^{-1} = (T^{-1})^*$.

⁵c'è qualcuno che dice *hermitiano*

Dimostrazione.

Punto (3).

$$\langle (ST)^*v, w \rangle = \langle v, STw \rangle = \langle S^*v, Tw \rangle = \langle T^*S^*v, w \rangle$$

Punto (5).

$$\|T^*v\|^2 = |\langle T^*v, T^*v \rangle| = |\langle v, TT^*v \rangle| \leq \|v\| \|TT^*v\| \leq \|v\| \|T\| \|T^*v\|$$

quindi

$$\|T^*v\| \leq \|T\| \|v\|$$

da cui

$$\|T^*\| \leq \|T\|.$$

D'altra parte

$$\|T\| = \|T^{**}\| \leq \|T^*\|$$

Punto (6). Siccome

$$TT^{-1} = T^{-1}T = I$$

possiamo prendere l'aggiunto

$$(TT^{-1})^* = (T^{-1}T)^* = I^* = I$$

e grazie al punto (3) otteniamo

$$(T^{-1})^* T^* = T^* (T^{-1})^* = I$$

che implica che $(T^{-1})^*$ è l'inverso di T^* □

7.29 Proposizione. *Sia $T \in \mathcal{L}(V)$. Allora*

$$\text{Ker } T^* = (\text{Ran } T)^\perp \quad e \quad \overline{\text{Ran } T^*} = (\text{Ker } T)^\perp$$

Dimostrazione. Sia $v \in \text{Ker } T^*$. Allora $T^*v = 0$, per cui

$$\langle T^*v, u \rangle = 0 \quad \forall u \in V,$$

per cui

$$\langle v, Tu \rangle = 0 \quad \forall u \in V,$$

che è equivalente a dire

$$\langle v, z \rangle = 0 \quad \forall z \in \text{Ran } T.$$

Ma questo significa appunto che $v \perp \text{Ran } T$. Abbiamo quindi mostrato che se $v \in \text{Ker } T^*$ allora $v \in (\text{Ran } T)^\perp$, vale a dire che $\text{Ker } T^* \subset (\text{Ran } T)^\perp$. L'inclusione inversa si dimostra facendo la stessa strada al contrario.

La seconda identità segue dalla prima. Infatti

$$\overline{\text{Ran } T^*} = ((\text{Ran } T^*)^\perp)^\perp = (\text{Ker } T^{**})^\perp = (\text{Ker } T)^\perp$$

□

7.30 Problema. Sia T l'operatore lineare su ℓ_2 definito come

$$Tx := \left(x_1 + \frac{x_2}{2}, x_3 + \frac{x_4}{4}, x_5 + \frac{x_6}{6}, \dots \right).$$

Determinare: (a) T^* , (b) $\text{Ker } T$, (c) $\text{Ran } T$.

Soluzione. (a). Per determinare l'aggiunto T^* partiamo dalla definizione

$$\langle T^*x, y \rangle = \langle x, Ty \rangle \quad \forall x, y \in \ell_2.$$

Quindi abbiamo

$$\begin{aligned} \langle T^*x, y \rangle &= \langle x, Ty \rangle \\ &= \left\langle \left(x_1, x_2, x_3, \dots \right), \left(y_1 + \frac{y_2}{2}, y_3 + \frac{y_4}{4}, y_5 + \frac{y_6}{6}, \dots \right) \right\rangle \\ &= x_1 \left(y_1 + \frac{y_2}{2} \right) + x_2 \left(y_3 + \frac{y_4}{4} \right) + x_3 \left(y_5 + \frac{y_6}{6} \right) + \dots \\ &= x_1 y_1 + \frac{x_1}{2} y_2 + x_2 y_3 + \frac{x_2}{4} y_4 + x_3 y_5 + \frac{x_3}{6} y_6 \\ &= \left\langle \left(x_1, \frac{x_1}{2}, x_2, \frac{x_2}{4}, x_3, \frac{x_3}{6}, \dots \right), \left(y_1, y_2, y_3, y_4, y_5, y_6, \dots \right) \right\rangle \end{aligned}$$

Poiché il prodotto scalare nell'ultima riga deve essere uguale a $\langle T^*x, y \rangle$, otteniamo

$$\textcircled{1} \quad T^*x = \left(x_1, \frac{x_1}{2}, x_2, \frac{x_2}{4}, x_3, \frac{x_3}{6}, \dots \right)$$

Un modo alternativo di calcolare T^* è quello di scrivere la matrice infinita A associata all'operatore T . La matrice associata a T^* sarà A^* , l'hermitiana coniugata (vale a dire la trasposta della complessa coniugata) di A . Dalla definizione di T si legge che la matrice A è data da

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 1/4 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 1/6 & \dots \\ \vdots & & & & & \vdots & \end{pmatrix}$$

Quindi

$$A^* = (\overline{A})^t = A^t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1/2 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1/4 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1/6 & 0 & 0 & 0 & \dots \\ \vdots & & & & & \vdots & \end{pmatrix}$$

In questo modo otteniamo nuovamente T^*x come nella $\textcircled{1}$.

(b). Per determinare il nucleo di T dobbiamo trovare la soluzione generale dell'equazione $Tx = 0$, che scritta esplicitamente diventa un sistema di infinite equazioni:

$$x_1 + x_2/2 = 0 \quad x_3 + x_4/4 = 0 \quad x_5 + x_6/6 = 0 \quad \dots$$

Queste equazioni sono fortunatamente (o per esagerata magnanimità di chi ha preparato l'esercizio) tutte indipendenti e si possono risolvere, ad esempio, ricavando le variabili pari in funzioni delle dispari:

$$x_{2k} = -2k x_{2k-1} \quad k = 1, 2, 3, \dots$$

Otteniamo quindi

$$\text{Ker } T := \{x \in \ell_2 : x_{2k} = -2k x_{2k-1} \quad k = 1, 2, 3, \dots\}$$

(c). L'immagine di T è costituita da tutti quelle successioni $y \in \ell_2$ tali che si possono ottenere come $y = Tx$ per un qualche $x \in \ell_2$. Affermo che $\text{Ran } T = \ell_2$. Infatti data una $y \in \ell_2$ arbitraria, posso sempre scrivere $y = Tx$, ponendo, ad esempio

$$x_{2k} = 0 \quad x_{2k-1} = y_k \quad k = 1, 2, 3, \dots$$

L'ultima cosa da verificare è che la x appartiene ad ℓ_2 . A questo proposito osservo che

$$\sum_{k=1}^{\infty} x_k^2 = \sum_{k=1}^{\infty} x_{2k-1}^2 = \sum_{k=1}^{\infty} y_k^2,$$

quindi, poiché $y \in \ell_2$, si ha che anche $x \in \ell_2$. Abbiamo dimostrato dunque che $\text{Ran } T = \ell_2$.

7.31 Problema. Sia T l'operatore lineare su ℓ_2 definito come

$$Tx := (0, x_1, 0, x_2, 0, x_3, 0, \dots).$$

- (a) Determinare T^* , vale a dire scrivere esplicitamente $T^*x = (?, ?, \dots)$
 (b) Determinare $\|T\|$, $\text{Ker } T$ e $\text{Ran } T$.

Soluzione. (a)

$$\begin{aligned} \langle T^*x, y \rangle &= \langle x, Ty \rangle = \sum_{k=1}^{\infty} x_k (Ty)_k = \sum_{k=1}^{\infty} x_{2k} y_k \\ &= x_2 y_1 + x_4 y_2 + x_6 y_3 + \dots \end{aligned}$$

Quindi

$$T^*x = (x_2, x_4, x_6, \dots).$$

(b) Poiché $\|Tx\|_2 = \|x\|_2$ per ogni $x \in \ell_2$, T è un'isometria, quindi $\|T\| = 1$. Inoltre se $Tx = 0$ allora deve essere $x = 0$ quindi $\text{Ker } T = \{0\}$. Infine

$$\text{Ran } T := \{x \in \ell_2 : x_{2k+1} = 0 \ \forall k \in \mathbb{N}\}.$$

7.4.1 Proiettori ortogonali

7.32 Definizione. $P \in \mathcal{L}(V)$ è detto una *proiezione ortogonale* se P è autoaggiunto e se $P^2 = P$.

7.33 Problema. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio di Hilbert. Dimostrare che se P è una proiezione ortogonale in V allora,

- (a) $(I - P)v$ è ortogonale a Pv
 (b) $\|Pv\| \leq \|v\|$
 (c) $\text{Ran } P = \text{Ker}(I - P)$

7.5 Spettro

7.5.1 Il caso finito dimensionale (\mathbb{C}^n).

Sia A una matrice complessa $n \times n$ e sia $z \in \mathbb{C}$. Allora ci sono solo due possibilità

- (1) se $\det(zI - A) = 0$, allora $zI - A$ non è iniettiva, vale a dire l'equazione

$$Ax = zx \quad x = (x_1, \dots, x_n) \in \mathbb{C}^n$$

ammette una soluzione x non banale ($x \neq 0$). In questo caso z è detto *autovalore* di A e x è il corrispondente *autovettore*

- (2) se $\det(zI - A) \neq 0$ allora $zI - A$ è invertibile, quindi è iniettiva e suriettiva. In questo caso, se $b \in \mathbb{C}^n$, l'equazione

$$(zI - A)x = b$$

ammette sempre un'unica soluzione data da

$$x = (zI - A)^{-1}b.$$

Nel caso infinito dimensionale, la situazione è più complicata, perché è possibile che $(zI - A)$ sia iniettiva ma non suriettiva.

7.5.2 Il caso generale

7.34 Definizione. (Spettro, insieme risolvente, risolvente). Sia $(V, \|\cdot\|)$ uno spazio di Banach complesso, sia $T \in \mathcal{L}(V)$ e sia $z \in \mathbb{C}$. Allora

- (1) se $(zI - T)$ non è iniettivo, ossia se l'equazione

$$Tx = zx \quad x \in V$$

ammette una soluzione $x \neq 0$, allora z è un *autovalore* di T . L'insieme degli autovalori di T è detto *spettro puntuale* e si indica con $\sigma_p(T)$

- (2) se $(zI - T)$ è iniettivo ma non suriettivo, allora si dice che z appartiene allo *spettro continuo* di T e si indica con $\sigma_c(T)$

- (3) se $(zI - T)$ è un operatore biunivoco su V (quindi con inverso limitato) allora z si dice appartenere all'*insieme risolvente* di T , che si denota con $\rho(T)$. L'operatore $R_z(T) := (zI - T)^{-1}$ è detto il *risolvente* di T in z .

Lo *spettro* di T è definito come

$$\text{spec } T = \sigma(T) = \mathbb{C} \setminus \rho(T) := \sigma_p(T) \cup \sigma_c(T)$$

7.35 Proposizione. $\rho(T)$ è aperto in \mathbb{C} , quindi $\sigma(T)$ è chiuso.

(senza dimostrazione?)

7.36 Proposizione. Sia $(V, \|\cdot\|)$ uno spazio di Banach. Se $T \in \mathcal{L}(V)$, allora lo spettro di T è contenuto nel disco chiuso di centro 0 e raggio $\|T\|$.

Dimostrazione. Sia $z \in \mathbb{C}$ tale che $|z| > \|T\|$. Allora

$$zI - T = z(I - z^{-1}T).$$

Ma dato che $\|z^{-1}T\| = |z^{-1}|\|T\| < 1$, grazie alla Proposizione 7.22, l'operatore $I - z^{-1}T$ è invertibile con inverso limitato, quindi anche $zI - T$ è invertibile con inverso limitato e si ha

$$R_z(T) = (zI - T)^{-1} = z^{-1}(I - z^{-1}T)^{-1}$$

Abbiamo quindi fatto vedere che se $|z| > \|T\|$ allora $z \in \rho(T)$ che è equivalente all'affermazione della Proposizione da dimostrare \square

7.37 Problema. In $(C([a, b]; \mathbb{C}), \|\cdot\|_u)$ consideriamo l'operatore

$$(Tf)(x) := x f(x)$$

Dimostrare che T non ha autovalori e che $\sigma(T) = [a, b]$.

7.38 Problema. Sia T l'operatore lineare sullo spazio $C_2([-1, 1]; \mathbb{C})$ definito come $Tf(x) = x^2 f(x)$. Calcolare $\|T\|$, gli autovalori e lo spettro continuo di T .

7.39 Problema. Nello spazio $(C([0, 2]; \mathbb{C}), \|\cdot\|_u)$ consideriamo l'operatore

$$(Tf)(x) := g(x) f(x) \quad \text{in cui} \quad g(x) := \begin{cases} x & \text{se } x \in [0, 1] \\ 1 & \text{se } x \in (1, 2] \end{cases}$$

Determinare gli autovalori e lo spettro continuo di T .

7.40 Problema. Sia ϑ_+ l'operatore di traslazione a destra che agisce in $\ell_2(\mathbb{C})$. Dimostrare che

- (1) ϑ_+ non ha autovalori
- (2) $\sigma(\vartheta_+) = \overline{B_1} := \{z \in \mathbb{C} : |z| \leq 1\}$

7.41 Problema. Sia ϑ_- l'operatore di traslazione a sinistra che agisce in $\ell_2(\mathbb{C})$. Determinare gli autovalori e lo spettro continuo di ϑ_- .

Risp: $\sigma_p(\vartheta_-) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$, $\sigma_c(\vartheta_-) := \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

7.6 Operatori autoaggiunti

7.42 Teorema. Se $(V, \langle \cdot, \cdot \rangle)$ è uno spazio di Hilbert complesso e $T \in \mathcal{L}(V)$ è autoaggiunto allora

- (a) gli autovalori di T sono reali
- (b) autovettori corrispondenti ad autovalori distinti sono ortogonali
- (c) $\sigma(T) \subset \mathbb{R}$ (generalizza (a)).

Dimostrazione. I punti (a) e (b) si dimostrano come nel caso finito dimensionale. Se λ è un autovalore di T allora esiste $v \in V$, $v \neq 0$ tale che $Tv = \lambda v$. Quindi

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle,$$

per cui si ha $\lambda = \bar{\lambda}$.

Se λ e μ sono due autovalori distinti con autovettori rispettivamente u e v , allora

$$\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, \mu v \rangle = \mu \langle u, v \rangle,$$

per cui

$$(\lambda - \mu) \langle u, v \rangle = 0.$$

Poiché $\lambda \neq \mu$ deve essere necessariamente $\langle u, v \rangle = 0$.

Punto (c). Sia $z = a + ib$ con $a, b \in \mathbb{R}$ e supponiamo $b \neq 0$. Voglio far vedere che $z \notin \sigma(T)$. Infatti, siccome $T^* = T$, e quindi $(\lambda I - T)^* = (\lambda I - T)$, abbiamo, per ogni $v \in V$,

$$\begin{aligned} \|(zI - T)v\|^2 &= \langle (aI - T)v + ibv, (aI - T)v + ibv \rangle \\ &= \|(aI - T)v\|^2 + \|bv\|^2 - ib \langle (aI - T)v, v \rangle + ib \langle v, (aI - T)v \rangle \\ &= \|(aI - T)v\|^2 + \|bv\|^2 - ib \langle (aI - T)v, v \rangle + ib \langle (aI - T)v, v \rangle \\ &= \|(aI - T)v\|^2 + \|bv\|^2 \\ &= \|(aI - T)v\|^2 + |b|^2 \|v\|^2. \end{aligned}$$

Di conseguenza

$$(7.13) \quad \|(zI - T)v\| \geq |b| \|v\|.$$

Con lo stesso procedimento si ottiene che

$$(7.14) \quad \|(zI - T)^*v\| = \|(\bar{z}I - T)v\| \geq |b| \|v\|.$$

Questo implica che $zI - T$ e $(zI - T)^*$ sono iniettivi.⁶ Infatti se $(zI - T)v = 0$ allora $\|(zI - T)v\| = 0$ e, grazie alla (7.13), $\|v\| = 0$, per cui $v = 0$. L'iniettività di $zI - T$ è dimostrata, mentre quella di $(zI - T)^*$ segue dalla (7.14).

Ora faccio vedere che $zI - T$ è suriettivo, cioè che $\text{Ran}(zI - T) = V$. Poiché $(zI - T)^*$ è iniettivo si ha

$$\text{Ker}(zI - T)^* = \{0\},$$

quindi, grazie alla Proposizione 7.29 abbiamo

$$(7.15) \quad \overline{\text{Ran}(zI - T)} = (\text{Ker}(zI - T)^*)^\perp = \{0\}^\perp = V.$$

Questa identità è quasi quello che dobbiamo dimostrare, a parte per l'operazione di chiusura. Per completare l'argomento faremo vedere che $\text{Ran}(zI - T)$ è chiuso. Sia (v_n) una successione di elementi di $\text{Ran}(zI - T)$ tale che $v_n \rightarrow v \in V$. Devo far vedere che anche v è un elemento di $\text{Ran}(zI - T)$. Infatti, dato che $v_n \in \text{Ran}(zI - T)$, esiste $u_n \in V$ tale che $v_n = (zI - T)u_n$. La successione (v_n) è convergente quindi è di Cauchy. In virtù della (7.13)

$$\|u_n - u_m\| \leq |b|^{-1} \|v_n - v_m\|$$

per cui la successione (u_n) è anch'essa di Cauchy. Ma V è completo per ipotesi, quindi

$$\text{esiste } u \in V \text{ tale che } u_n \rightarrow u$$

Infine, poiché $(zI - T)$ è continuo

$$(zI - T)u_n = v_n \rightarrow (zI - T)u.$$

Abbiamo quindi

$$v_n \rightarrow (zI - T)u \quad \text{e} \quad v_n \rightarrow v.$$

Per l'unicità del limite deve essere $v = (zI - T)u$, che implica $v \in \text{Ran}(zI - T)$. Abbiamo dunque dimostrato che

$$v_n \in \text{Ran}(zI - T) \text{ e } v_n \rightarrow v \in V \Rightarrow v \in \text{Ran}(zI - T)$$

che significa proprio che $\text{Ran}(zI - T)$ è chiuso.

Da questo fatto e dalla (7.15) segue che

$$\text{Ran}(zI - T) = V.$$

L'operatore $zI - T$ è quindi biunivoco. Per il teorema sull'operatore inverso, dato che V è completo, il suo inverso è limitato. Questo implica $z \in \rho(T)$, per cui z non appartiene allo spettro di T .

Osservo che la limitatezza di $(zI - T)^{-1}$ è una diretta conseguenza della (7.13), senza quindi bisogno di invocare il teorema sull'operatore inverso. \square

⁶questo, ovviamente, segue anche dal punto (a), ma a noi in ogni caso servirà la disuguaglianza (7.13)

7.43 Attenzione! A differenza del caso finito dimensionale, se T è un operatore limitato autoaggiunto su V , non è detto che esista una base ortonormale di V costituita da autovettori di T . Considerate ad esempio l'operatore di moltiplicazione per x definito nel Problema 7.37, il quale non ha autovalori pur essendo autoaggiunto (verificare). Una classe importante di operatori, per i quali esiste sempre una base ortonormale di autovettori è costituita dagli operatori autoaggiunti e compatti.

7.7 Operatori compatti

7.44 Definizione. Un operatore lineare limitato T su uno spazio di Banach $(L, \|\cdot\|)$ si dice *compatto* se vale l'implicazione

se $A \subset L$ è limitato, allora \overline{TA} è compatto

o, equivalentemente, se per ogni successione limitata $(x_n)_{n=1}^\infty$, la successione $(Tx_n)_{n=1}^\infty$ ammette una sottosuccessione convergente.

7.45 Esempio. L'operatore identico I in uno spazio infinito dimensionale non è compatto. Infatti se B_1 è la palla unitaria, si ha $\overline{TB_1} = B_1$. Ma la palla unitaria chiusa in infinite dimensioni non è compatta, come spiegato al punto 3.22.

7.46 Esempio. (Operatori integrali). Consideriamo due operatori integrali S, T in $C_2[a, b]$ associati al nucleo $K \in C([a, b] \times [a, b])$

$$(7.16) \quad Sf(x) := \int_a^b K(x, y) f(y) dy \quad f \in C_2[a, b]$$

$$(7.17) \quad Tf(x) := \int_a^x K(x, y) f(y) dy \quad (\text{operatore di Volterra}) \quad f \in C_2[a, b]$$

È possibile far vedere che S e T sono compatti.

Un criterio utile a volte a determinare la compattezza di un operatore è il seguente

7.47 Proposizione. Sia V uno spazio di Hilbert separabile e sia $T \in \mathcal{L}(V)$. Allora T è compatto se e solo se esiste una successione di operatori $T_n \in \mathcal{L}(V)$ tali che

- (a) Lo spazio $\text{Ran } T_n$ ha dimensione finita
- (b) $\lim_{n \rightarrow \infty} \|T - T_n\| = 0$.

7.48 Osservazione. Nelle ipotesi del risultato precedente, se $\text{Ran } T$ ha dimensione finita, T è compatto

7.49 Esempio. Ad esempio siano $(u_k)_{k=1}^n$ e $(w_k)_{k=1}^n$ 2 ennuple ortonormali in V e siano $c_1, \dots, c_n > 0$ Allora l'operatore

$$Tv := \sum_{k=1}^n c_k \langle v, u_k \rangle w_k$$

è compatto. Infatti $\text{Ran } T = \text{span}\{w_1, \dots, w_n\}$. Chi è il nucleo di T ?

7.50 Teorema. (Fredholm). Se T è un operatore compatto nello spazio di Hilbert V allora

- (a) Se $z \in \sigma(T)$ e $z \neq 0$ allora z è un autovalore di molteplicità finita, cioè lo spazio degli autovettori corrispondenti è finito dimensionale

(b) L'insieme degli autovalori può essere finito o infinito (o vuoto). Nel secondo caso l'origine è l'unico punto di accumulazione degli autovalori.

L'importanza di questo risultato si può vedere nel seguente esempio. Supponiamo di voler risolvere l'equazione integrale

$$(7.18) \quad \int_a^b K(x, y) f(y) dy - cf(x) = g(x) \quad c \neq 0$$

in cui g è data e f è la funzione l'incognita. Allora ci sono due problemi da risolvere: il problema dell'esistenza di una soluzione e il problema dell'unicità della soluzione. Accade spesso che il problema di esistenza sia (molto) più difficile del problema di unicITÀ. Ma in questo caso unicITÀ implica esistenza, quindi basta dimostrare l'unicITÀ. Infatti la (7.18) può essere scritta, usando la (7.16) come

$$(7.19) \quad (S - cI)f = g.$$

Ma se sappiamo per qualche motivo che la soluzione è unica, allora sappiamo che $S - cI$ è iniettivo, quindi c non è un autovalore. Siccome $c \neq 0$, grazie al Teorema di Fredholm possiamo affermare che $(S - cI)$ ha inverso limitato, quindi la (7.19) può essere risolta esplicitamente come

$$f = (S - cI)^{-1}g$$

7.51 Teorema. (Teorema di Hilbert-Schmidt). Se T è un operatore compatto e autoaggiunto nello spazio di Hilbert V , allora esiste una base ortonormale di V $(e_n)_{n=1}^\infty$ costituita da autovettori di T . Quindi si ha

$$Te_n = \lambda_n e_n \quad \forall n.$$

Inoltre $\lim_{n \rightarrow \infty} \lambda_n = 0$.

Nel caso di un operatore compatto e autoaggiunto su uno spazio di Hilbert si possono quindi verificare le seguenti possibilità

	$\sigma_p(T)$	$\sigma_c(T)$	Esempio banale in ℓ_2
1	$\{0, \lambda_1, \lambda_2, \dots, \lambda_n\}$	\emptyset	$Tx = (x_1, \frac{x_2}{2}, \dots, \frac{x_n}{n}, 0, 0, \dots)$
2	$\{0, \lambda_1, \lambda_2, \lambda_3, \dots\}$ con $\lambda_n \neq 0$ e $\lambda_n \rightarrow 0$	\emptyset	$Tx = (0, \frac{x_2}{2}, \frac{x_3}{3}, \frac{x_4}{4}, \dots)$
3	$\{\lambda_1, \lambda_2, \lambda_3, \dots\}$ con $\lambda_n \neq 0$ e $\lambda_n \rightarrow 0$	$\{0\}$	$Tx = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \frac{x_4}{4}, \dots)$

7.52 Problema. Sia $(a_n)_{n=1}^\infty$ una successione di numeri reali convergente a 0. Dimostrare che l'operatore T su ℓ_2 definito come

$$Tx = (a_1x_1, a_2x_2, a_3x_3, \dots)$$

è compatto.

7.53 Problema. Sia T l'operatore che agisce in ℓ_2 come

$$Tx = \left(0, \frac{x_1}{2}, \frac{x_2}{3}, \frac{x_3}{4}, \dots\right)$$

- (a) Determinare $\|T\|$
 (b) Dimostrare che T è compatto
 (c) Trovare gli autovalori e lo spettro continuo di T

7.54 Problema. Sia T l'operatore su ℓ_2 definito come

$$Tx = \left(x_1, x_1, \frac{x_2}{2}, \frac{x_2}{2}, \frac{x_3}{3}, \frac{x_3}{3}, \frac{x_4}{4}, \frac{x_4}{4}, \dots\right)$$

- (a) Determinare $\|T\|$
 (b) Determinare T^* . $T^*x = (?, ?, ?, ?, \dots)$
 (c) Dimostrare che T è compatto
 (d) Trovare gli autovalori e lo spettro continuo di T (sugg: ricorda il teorema di Fredholm)

7.55 Problema. Sia T l'operatore su ℓ_2 definito come

$$Tx = \left(\frac{x_2}{2}, 0, \frac{x_4}{4}, 0, \frac{x_6}{6}, 0, \frac{x_8}{8}, 0, \dots\right)$$

- (a) Determinare $\|T\|$
 (b) Determinare T^* . $T^*x = (?, ?, ?, ?, \dots)$
 (c) Trovare gli autovalori e lo spettro continuo di T (sugg: dato che T è compatto si può utilizzare il teorema di Fredholm)

Soluzione. (a) È facile mostrare che $\|T\| = 1/2$.

(b) Per determinare T^* procediamo come segue

$$\begin{aligned} \langle T^*x, y \rangle &= \langle x, Ty \rangle \\ &= \left\langle (x_1, x_2, x_3, \dots), \left(\frac{y_2}{2}, 0, \frac{y_4}{4}, 0, \frac{y_6}{6}, 0, \frac{y_8}{8}, 0, \dots\right) \right\rangle \\ &= \frac{x_1 y_2}{2} + \frac{x_3 y_4}{4} + \frac{x_5 y_6}{6} + \frac{x_7 y_8}{8} + \dots \\ &= \left\langle \left(0, \frac{x_1}{2}, 0, \frac{x_3}{4}, 0, \frac{x_5}{6}, 0, \frac{x_7}{8}, \dots\right), (y_1, y_2, y_3, \dots) \right\rangle \\ &= \langle T^*x, y \rangle, \end{aligned}$$

per cui si ha

$$T^*x = \left(0, \frac{x_1}{2}, 0, \frac{x_3}{4}, 0, \frac{x_5}{6}, \dots\right).$$

(c) Per trovare gli autovalori di T scrivo l'equazione $Tx = \lambda x$ che diventa

$$\begin{aligned} \frac{x_2}{2} &= \lambda x_1 & 0 &= \lambda x_2 \\ \frac{x_4}{4} &= \lambda x_3 & 0 &= \lambda x_4 \\ &\vdots & & \\ \frac{x_{2n}}{2n} &= \lambda x_{2n-1} & 0 &= \lambda x_{2n} \\ &\vdots & & \end{aligned}$$

Nel caso in cui λ sia diverso da zero, la colonna di destra ci dice che tutte le componenti pari di x sono nulle. Sostituendo nella colonna di sinistra ottengo che anche le componenti dispari sono nulle, cioè $x = 0$. Quindi se $\lambda \neq 0$ l'unica soluzione dell'equazione agli autovalori $Tx = \lambda x$ è la soluzione nulla, vale a dire λ non è un autovalore di T .

Rimane da esaminare il caso $\lambda = 0$. Si verifica immediatamente che $\hat{x} = (1, 0, 0, 0, \dots)$ verifica $T\hat{x} = 0$, quindi $\lambda = 0$ è un autovalore di T . Dunque si ha

$$\sigma_p(T) = \{0\}.$$

Spettro continuo. Poiché T è compatto (verificare) sappiamo che lo spettro continuo è vuoto o al massimo consiste solo del punto $\lambda = 0$. Ma sappiamo già che 0 è un autovalore, dunque $\sigma_c(T) = \emptyset$. \square

7.56 Problema. Sia T l'operatore su ℓ_2 definito come

$$Tx = \left(x_1, x_1, \frac{x_3}{3}, \frac{x_3}{3}, \frac{x_5}{5}, \frac{x_5}{5}, \dots\right)$$

- Determinare $\|T\|$
- Determinare T^* . $T^*x = (?, ?, ?, ?, \dots)$
- Trovare gli autovalori e lo spettro continuo di T (sugg: dato che T è compatto si può utilizzare il teorema di Fredholm)

Soluzione.

(a)

$$\|Tx\|^2 = 2 \sum_{k=0}^{\infty} \frac{|x_{2k+1}|^2}{(2k+1)^2} \leq 2 \sum_{k=0}^{\infty} |x_{2k+1}|^2 \leq 2\|x\|^2,$$

da cui si ha $\|T\| \leq \sqrt{2}$. D'altra parte se $y = (1, 0, 0, \dots)$ si ottiene $Ty = (1, 1, 0, 0, \dots)$, quindi

$$\frac{\|Ty\|}{\|y\|} = \sqrt{2}.$$

Quindi $\|T\| = \sqrt{2}$.

(b) Per determinare T^* procediamo come segue

$$\begin{aligned} \langle T^*x, y \rangle &= \langle x, Ty \rangle = \left\langle (x_1, x_2, x_3, \dots), \left(y_1, y_1, \frac{y_3}{3}, \frac{y_3}{3}, \frac{y_5}{5}, \frac{y_5}{5}, \dots\right) \right\rangle \\ &= x_1 y_1 + x_2 y_1 + \frac{x_3 y_3}{3} + \frac{x_4 y_3}{3} + \frac{x_5 y_5}{5} + \frac{x_6 y_5}{5} + \dots \\ &= (x_1 + x_2) y_1 + \frac{x_3 + x_4}{3} y_3 + \frac{x_5 + x_6}{5} y_5 + \dots = \langle T^*x, y \rangle, \end{aligned}$$

per cui si ha

$$T^*x = \left(x_1 + x_2, 0, \frac{x_3 + x_4}{3}, 0, \frac{x_5 + x_6}{5}, 0, \dots\right).$$

(c) Scrivo l'equazione agli autovalori $Tx = \lambda x$

$$\begin{array}{ll} x_1 = \lambda x_1 & x_1 = \lambda x_2 \\ \frac{x_3}{3} = \lambda x_3 & \frac{x_3}{3} = \lambda x_4 \\ \frac{x_5}{5} = \lambda x_5 & \frac{x_5}{5} = \lambda x_6 \\ \vdots & \end{array}$$

Se $\lambda = 0$ allora $x = (0, 1, 0, 0, 0, \dots)$ è chiaramente un autovettore, quindi 0 è un autovalore di T .

Se $\lambda \neq 0$ non è della forma $1/(2k - 1)$ con k intero positivo, allora, dalla colonna di sinistra si ottiene che tutte le componenti dispari di x sono nulle. Sostituendo nella colonna di destra si vede che anche le componenti pari sono nulle, cioè $x = 0$. In questo caso dunque non ci sono autovettori e λ non è un autovalore.

Se invece $\lambda = 1/(2k - 1)$ con k intero positivo allora c'è una soluzione non nulla data (ad esempio) da

$$x = (0, 0, 0, \dots, 0, 0, 1, 1, 0, 0, \dots)$$

in cui gli elementi diversi da zero sono il $2k - 1$ -simo e il $2k$ -simo. Abbiamo quindi ottenuto che gli autovalori sono

$$\sigma_p(T) = \{0\} \cup \{1/(2k - 1) : k \in \mathbb{Z}, k > 0\}.$$

Poiché lo zero è un autovalore di T per il teorema di Fredholm T non ha spettro continuo.

8. Serie di Fourier

*Lui balla e non la guarda,
lei balla e non lo vede,
ma quando i due si sfiorano
il pavimento cede (E.)*

8.1 $L_2[-\pi, \pi]$ e i polinomi trigonometrici

Data una qualsiasi base ortogonale $(u_k)_{k=1}^\infty$ nello spazio $(L_2[-\pi, \pi], \|\cdot\|_2)$ possiamo rappresentare una funzione arbitraria $f \in L_2[-\pi, \pi]$ come

$$(8.1) \quad f = \sum_{k=1}^{\infty} c_k u_k.$$

in cui i coefficienti c_k sono dati da

$$(8.2) \quad c_k := \frac{\langle f, u_k \rangle}{\|u_k\|_2^2}.$$

Ricordo che la convergenza della sommatoria va intesa rispetto alla norma $\|\cdot\|_2$, vale a dire la (8.1) è una scrittura simbolica che sta a significare

$$(8.3) \quad \lim_{n \rightarrow \infty} \left\| f - \sum_{k=1}^n c_k u_k \right\|_2 = 0.$$

Nel seguito, per brevità, indicheremo la convergenza rispetto alla norma $\|\cdot\|_2$ con la notazione

$$(8.4) \quad f(x) \sim \sum_{k=1}^{\infty} c_k u_k(x)$$

per distinguerla dalla scrittura

$$(8.5) \quad f(x) = \sum_{k=1}^{\infty} c_k u_k(x)$$

che denota la usuale convergenza *puntuale*. È importante ricordare che mentre la convergenza in norma $\|\cdot\|_2$ è sempre verificata grazie al fatto che (u_k) è una base ortogonale e grazie alla Proposizione 4.47, per quanto riguarda la convergenza puntuale (8.5) a priori non si può dire nulla ed è un problema che va affrontato separatamente.

Per quanto riguarda la scelta della base ortogonale (u_k) c'è ovviamente una arbitrarietà enorme. La base dei polinomi trigonometrici, associata alla serie di Fourier "classica", che appare nella proposizione seguente si dimostra però particolarmente conveniente nell'analisi di molti problemi.

8.1 Teorema. Nello spazio $L_2[-\pi, \pi]$ consideriamo la seguente collezione di funzioni

$$(8.6) \quad u_0(x) := 1, \quad u_k(x) := \cos(kx), \quad v_k(x) := \sin(kx) \quad k = 1, 2, \dots$$

Allora:

(1) Il sistema $\{(u_n)_{n=0}^\infty, (v_n)_{n=1}^\infty\}$ è ortogonale e completo (quindi una base ortogonale) e vale:

$$(8.7) \quad \|u_0\| = \|1\| = \sqrt{2\pi} \quad \|u_n\| = \|v_n\| = \sqrt{\pi} \quad \forall n \geq 1.$$

(2) per ogni $f \in L_2[-\pi, \pi]$ si ha

$$(8.8) \quad f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)]$$

in cui

$$(8.9) \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx$$

(3) per ogni $f \in L_2[-\pi, \pi]$ si ha

$$(8.10) \quad \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2).$$

Dimostrazione.

(1). L'ortogonalità del sistema (8.6) e le uguaglianze (8.7) riguardanti la loro norma sono semplici verifiche che si ottengono usando le formule:

$$\begin{aligned} \cos(nx) \cos(kx) &= \frac{1}{2} [\cos((n+k)x) + \cos((n-k)x)] \\ \sin(nx) \sin(kx) &= \frac{1}{2} [\cos((n-k)x) - \cos((n+k)x)] \\ \cos(nx) \sin(kx) &= \frac{1}{2} [\sin((n+k)x) - \sin((n-k)x)] \end{aligned}$$

Ad esempio per dimostrare che $\langle u_n, v_k \rangle = 0$ basta osservare che

$$\langle u_n, v_k \rangle = \int_{-\pi}^{\pi} \cos(nx) \sin(kx) dx = 0$$

perché la funzione integranda è dispari (prodotto di una funzione pari per una funzione dispari) e l'intervallo di integrazione è simmetrico rispetto all'origine.

Facciamo vedere che $\langle u_k, u_n \rangle = 0$ se $n \neq k$.

$$\begin{aligned} \langle u_n, u_k \rangle &= \int_{-\pi}^{\pi} \cos(nx) \cos(kx) dx \\ &= \frac{1}{2} \int_{-\pi}^{\pi} [\cos((n-k)x) + \cos((n+k)x)] dx. \end{aligned}$$

Sia $n+k$ che $n-k$ sono interi, quindi

$$\sin[(n-k)\pi] = \sin[(n+k)\pi] = \sin[-(n-k)\pi] = \sin[-(n+k)\pi] = 0.$$

Inoltre $n \neq k$, per cui $n - k \neq 0$, e dunque otteniamo

$$\langle u_n, u_k \rangle = \frac{1}{2(n-k)} [\sin[(n-k)x]]_{-\pi}^{\pi} + \frac{1}{2(n+k)} [\sin[(n+k)x]]_{-\pi}^{\pi} = 0.$$

La dimostrazione della completezza è (parecchio) più complicata e rimandata alla sezione 8.1.5.

(2). Una volta dimostrato che l'insieme delle funzioni (8.6) è ortogonale e completo, il punto (2) non è altro che un caso particolare delle uguaglianze (8.1) e (8.2).

(3). L'uguaglianza (8.10), infine, è un caso particolare dell'uguaglianza di Parseval

$$\|f\|_2^2 = \sum_k |c_k|^2 \|u_k\|^2. \quad \square$$

8.1.1 Serie di Fourier sull'intervallo $[a, b]$

Sappiamo già che le funzioni

$$(8.11) \quad u_0(x) := 1, \quad u_k(x) := \cos(kx), \quad v_k(x) := \sin(kx) \quad k = 1, 2, \dots$$

sono una base ortogonale in $L_2[-\pi, \pi]$, con la normalizzazione data dalla (8.7). Con il cambio di variabile

$$y(x) = \frac{\pi(2x - a - b)}{b - a}$$

l'intervallo $[a, b]$ si trasforma nell'intervallo $[-\pi, \pi]$. Consideriamo allora le funzioni

$$\begin{aligned} \hat{u}_k(x) &:= u_k(y(x)) = \cos\left(\frac{k\pi(2x - a - b)}{b - a}\right) & k = 0, 1, 2, \dots \\ \hat{v}_k(x) &:= v_k(y(x)) = \sin\left(\frac{k\pi(2x - a - b)}{b - a}\right) & k = 1, 2, \dots \end{aligned}$$

È facile verificare che queste funzioni sono una base ortogonale in $L_2[a, b]$. Tutto quello che bisogna fare è cambiare variabile nell'integrale che definisce il prodotto scalare. Infatti

$$\begin{aligned} \langle \hat{u}_k, \hat{u}_n \rangle &= \int_a^b \hat{u}_k(x) \hat{u}_n(x) dx = \int_a^b u_k(y(x)) u_n(y(x)) dx \\ &= \frac{b-a}{2\pi} \int_{-\pi}^{\pi} u_k(y) u_n(y) dy = \frac{b-a}{2\pi} \langle u_k, u_n \rangle. \end{aligned}$$

Analogamente si trova

$$\begin{aligned} \langle \hat{v}_k, \hat{v}_n \rangle &= \frac{b-a}{2\pi} \langle v_k, v_n \rangle \\ \langle \hat{u}_k, \hat{v}_n \rangle &= \frac{b-a}{2\pi} \langle u_k, v_n \rangle. \end{aligned}$$

Dall'ortogonalità del sistema $\{(u_n)_{n=0}^{\infty}, (v_n)_{n=1}^{\infty}\}$ deriva quindi l'ortogonalità in $L_2[a, b]$ del sistema $\{(\hat{u}_n)_{n=0}^{\infty}, (\hat{v}_n)_{n=1}^{\infty}\}$. L'unica cosa che cambia è la normalizzazione, per cui al posto della (8.7) abbiamo

$$(8.12) \quad \|\hat{u}_0\| = \sqrt{b-a} \quad \|\hat{u}_n\| = \|\hat{v}_n\| = \sqrt{(b-a)/2} \quad \forall n \geq 1.$$

Di conseguenza possiamo scrivere lo sviluppo di Fourier

$$(8.13) \quad f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[a_k \cos \left(\frac{k\pi(2x-a-b)}{b-a} \right) + b_k \sin \left(\frac{k\pi(2x-a-b)}{b-a} \right) \right]$$

in cui i coefficienti sono dati da

$$(8.14) \quad \begin{aligned} a_k &= \frac{2}{b-a} \int_a^b f(x) \cos \left(\frac{k\pi(2x-a-b)}{b-a} \right) dx \\ b_k &= \frac{2}{b-a} \int_a^b f(x) \sin \left(\frac{k\pi(2x-a-b)}{b-a} \right) dx. \end{aligned}$$

8.1.2 Serie di Fourier sull'intervallo $[-l, l]$

Questo è un caso particolare di quello trattato in precedenza, quindi basta porre $a = -l$ e $b = l$ nelle formule (8.13) e (8.14).

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[a_k \cos \left(\frac{k\pi x}{l} \right) + b_k \sin \left(\frac{k\pi x}{l} \right) \right],$$

in cui

$$a_k = \frac{1}{l} \int_{-l}^l f(x) \cos \left(\frac{k\pi x}{l} \right) dx \quad b_k = \frac{1}{l} \int_{-l}^l f(x) \sin \left(\frac{k\pi x}{l} \right) dx.$$

L'identità di Parseval diventa

$$\frac{1}{\ell} \int_{-\ell}^{\ell} |f(x)|^2 dx = \frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2).$$

8.1.3 Serie di Fourier con solo coseni o seni

8.2 Proposizione. *Ciascuno dei seguenti due insiemi di vettori*

$$(a) \quad \cos(nx) \quad n = 0, 1, 2, \dots \quad (b) \quad \sin(nx) \quad n = 1, 2, \dots$$

è una base ortogonale in $L_2[0, \pi]$. Inoltre, se $f \in L_2[0, \pi]$, si ha:

$$(8.15) \quad f(x) \sim \frac{\hat{a}_0}{2} + \sum_{k=1}^{\infty} \hat{a}_k \cos(kx)$$

$$(8.16) \quad f(x) \sim \sum_{k=1}^{\infty} \hat{b}_k \sin(kx)$$

in cui

$$(8.17) \quad \hat{a}_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(kx) dx \quad \hat{b}_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx) dx$$

Dimostrazione. L'ortogonalità dei due sistemi di vettori si verifica facilmente. Per dimostrarne la completezza faccio vedere che, per ogni $f \in L_2[0, \pi]$, valgono gli sviluppi (8.15) e (8.16).

Se $f \in L_2[0, \pi]$ posso definire $f_p, f_d \in L_2[-\pi, \pi]$ come i prolungamenti pari e dispari di f , vale a dire come

$$f_p(x) := \begin{cases} f(x) & \text{se } x \in [0, \pi] \\ f(-x) & \text{se } x \in [-\pi, 0) \end{cases} \quad f_d(x) := \begin{cases} f(x) & \text{se } x \in [0, \pi] \\ -f(-x) & \text{se } x \in [-\pi, 0) \end{cases}$$

Applicando la (8.8) alle funzioni f_p e f_d ottengo

$$(8.18) \quad f_p(x) \sim \frac{a_0^p}{2} + \sum_{k=1}^{\infty} [a_k^p \cos(kx) + b_k^p \sin(kx)]$$

$$(8.19) \quad f_d(x) \sim \frac{a_0^d}{2} + \sum_{k=1}^{\infty} [a_k^d \cos(kx) + b_k^d \sin(kx)]$$

Utilizzando le (8.9) ed il fatto che le funzioni $f_p(x) \cos(kx)$ e $f_d(x) \sin(kx)$ sono pari, mentre le funzioni $f_p(x) \sin(kx)$ e $f_d(x) \cos(kx)$ sono dispari ottengo

$$a_k^p = \frac{1}{\pi} \int_{-\pi}^{\pi} f_p(x) \cos(kx) dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(kx) dx = \hat{a}_k$$

$$b_k^p = \frac{1}{\pi} \int_{-\pi}^{\pi} f_p(x) \sin(kx) dx = 0$$

$$a_k^d = \frac{1}{\pi} \int_{-\pi}^{\pi} f_d(x) \cos(kx) dx = 0$$

$$b_k^d = \frac{1}{\pi} \int_{-\pi}^{\pi} f_d(x) \sin(kx) dx = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx) dx = \hat{b}_k,$$

e quindi, le (8.18), (8.19) diventano

$$(8.20) \quad f_p(x) \sim \frac{\hat{a}_0}{2} + \sum_{k=1}^{\infty} \hat{a}_k \cos(kx) \quad f_d(x) \sim \sum_{k=1}^{\infty} \hat{b}_k \sin(kx)$$

La prima uguaglianza significa

$$(8.21) \quad \lim_{n \rightarrow \infty} \|f_p - S_n\|_2 = 0,$$

in cui abbiamo posto

$$S_n(x) := \frac{\hat{a}_0}{2} + \sum_{k=1}^n \hat{a}_k \cos(kx).$$

Poiché la differenza $f_p - S_n$ è una funzione pari, abbiamo

$$(8.22) \quad \begin{aligned} \|f_p - S_n\|_2^2 &= \int_{-\pi}^{\pi} |f_p(x) - S_n(x)|^2 dx \\ &= 2 \int_0^{\pi} |f_p(x) - S_n(x)|^2 dx = 2 \|f - S_n\|_{L_2[0, \pi]}^2. \end{aligned}$$

Dalla (8.21) e (8.22) segue la (8.15). La (8.16) si dimostra in modo analogo. \square

8.1.4 Problemi

8.3 Problema. Verificare i seguenti sviluppi in serie di Fourier

$$(a) \quad \operatorname{sgn}(x) \sim \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{(2k+1)} \sin[(2k+1)x] \quad (b) \quad x \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin(kx)}{k}$$

$$(c) \quad |x| \sim \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\cos[(2k+1)x]}{(2k+1)^2}$$

Usando l'uguaglianza di Parseval, calcolare $\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}$, $\sum_{k=1}^{\infty} \frac{1}{k^2}$ e $\sum_{k=0}^{\infty} \frac{1}{(2k+1)^4}$.

8.4 Problema. Sviluppare in serie trigonometrica di Fourier nell'intervallo $[-\pi, \pi]$ la funzione

$$f(x) = \begin{cases} 0 & \text{se } x \in [-\pi, 0) \\ x & \text{se } x \in [0, \pi]. \end{cases}$$

Scrivere, oltre alla formula completa, lo sviluppo esplicito di f fino a $k = 4$, vale a dire

$$f(x) \sim \frac{a_0}{2} + a_1 \cos x + a_2 \cos(2x) + a_3 \cos(3x) + a_4 \cos(4x) \\ + b_1 \sin x + b_2 \sin(2x) + b_3 \sin(3x) + b_4 \sin(4x) + \dots$$

Soluzione. Iniziamo a calcolare i coefficienti a_0 e a_k .

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{\pi} \int_0^{\pi} x dx = \frac{\pi}{2}. \\ a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx = \frac{1}{\pi} \int_0^{\pi} x \cos(kx) dx \\ = \frac{1}{\pi} \left[\frac{x \sin(kx)}{k} \right]_0^{\pi} - \frac{1}{k\pi} \int_0^{\pi} \sin(kx) dx \\ = \frac{1}{k^2\pi} [\cos(kx)]_0^{\pi} = \frac{1}{k^2\pi} [\cos(k\pi) - 1] = \frac{1}{k^2\pi} [(-1)^k - 1].$$

Convien distinguere i k pari dai dispari. Si ottiene

$$a_{2k} = 0 \qquad a_{2k+1} = -\frac{2}{\pi(2k+1)^2}.$$

Per quanto riguarda i b_k

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx = \frac{1}{\pi} \int_0^{\pi} x \sin(kx) dx \\ = -\frac{1}{\pi} \left[\frac{x \cos(kx)}{k} \right]_0^{\pi} + \frac{1}{k\pi} \int_0^{\pi} \cos(kx) dx \\ = -\cos(k\pi) + \frac{1}{k^2\pi} [\sin(kx)]_0^{\pi} = -\frac{1}{k} \cos(k\pi) = \frac{(-1)^{k+1}}{k}.$$

Posso quindi scrivere

$$f(x) \sim \frac{\pi}{4} - \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{\cos[(2k+1)x]}{(2k+1)^2} + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sin(kx).$$

Lo sviluppo esplicito fino a $k = 4$ è

$$f(x) \sim \frac{\pi}{4} - \frac{2}{\pi} \cos x - \frac{2}{9\pi} \cos(3x) + \sin(x) - \frac{1}{2} \sin(2x) + \frac{1}{3} \sin(3x) - \frac{1}{4} \sin(4x) + \dots$$

8.5 Problema. Sviluppare in serie trigonometrica di Fourier nell'intervallo $[-\ell, \ell]$ la funzione $f(x) = x^2$, ed utilizzare il risultato per calcolare $\sum_{n=1}^{\infty} 1/n^4$.

Soluzione. Poiché la funzione x^2 è pari e l'intervallo $[-\ell, \ell]$ è simmetrico rispetto all'origine, i coefficienti b_k sono nulli. Dalle formule della serie di Fourier in $[-\ell, \ell]$ si

ottiene

$$\begin{aligned} a_0 &= \frac{1}{\ell} \int_{-\ell}^{\ell} f(x) dx = \frac{1}{\ell} \int_{-\ell}^{\ell} x^2 dx = \frac{2}{\ell} \int_0^{\ell} x^2 dx = \frac{2\ell^2}{3}. \\ a_k &= \frac{1}{\ell} \int_{-\ell}^{\ell} f(x) \cos\left(\frac{k\pi x}{\ell}\right) dx = \frac{2}{\ell} \int_0^{\ell} x^2 \cos\left(\frac{k\pi x}{\ell}\right) dx \\ &= \frac{2}{\ell} \underbrace{\left[x^2 \sin\left(\frac{k\pi x}{\ell}\right) \frac{\ell}{k\pi} \right]_0^{\ell}}_{=0} - \frac{2}{\ell} \int_0^{\ell} 2x \sin\left(\frac{k\pi x}{\ell}\right) \frac{\ell}{k\pi} dx \\ &= \frac{4}{k\pi} \left[x \cos\left(\frac{k\pi x}{\ell}\right) \frac{\ell}{k\pi} \right]_0^{\ell} - \frac{4}{k\pi} \int_0^{\ell} \cos\left(\frac{k\pi x}{\ell}\right) \frac{\ell}{k\pi} dx \\ &= \frac{4\ell^2}{(k\pi)^2} \cos(k\pi) - \frac{4\ell}{(k\pi)^2} \underbrace{\left[\sin\left(\frac{k\pi x}{\ell}\right) \frac{\ell}{k\pi} \right]_0^{\ell}}_{=0} = (-1)^k \frac{4\ell^2}{(k\pi)^2}. \end{aligned}$$

Otteniamo quindi lo sviluppo

$$x^2 \sim \frac{\ell^2}{3} + \frac{4\ell^2}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \cos\left(\frac{k\pi x}{\ell}\right).$$

L'identità di Parseval afferma che

$$\frac{1}{\ell} \int_{-\ell}^{\ell} x^4 dx = \frac{2\ell^4}{9} + \sum_{k=1}^{\infty} \frac{16\ell^4}{\pi^4 k^4}$$

da cui otteniamo

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{16} \left(\frac{2}{5} - \frac{2}{9} \right) = \frac{\pi^4}{90}.$$

8.1.5 Completezza dei polinomi trigonometrici (schema di dimostrazione)

Voglio far vedere che la collezione di funzioni

$$(8.23) \quad 1, \cos(x), \cos(2x), \dots, \sin(x), \sin(2x), \dots$$

costituisce un sistema completo di vettori nello spazio di Hilbert $L_2[-\pi, \pi]$. Chiamo *polinomio trigonometrico* una funzione che può essere rappresentata come combinazione lineare finita delle funzioni (8.23). In altre parole

$$\{\text{polinomi trigonometrici}\} = \text{span}\{1, \cos(x), \cos(2x), \dots, \sin(x), \sin(2x), \dots\},$$

o, equivalentemente, un polinomio trigonometrico è una funzione della forma

$$p(x) = \frac{\alpha_0}{2} + \sum_{k=1}^N [\alpha_k \cos(kx) + \beta_k \sin(kx)]$$

in cui la sommatoria contiene un numero finito di termini. Ricordando la definizione di completezza per un sistema di vettori (sezione 3.8) devo far vedere che

$$\forall f \in L_2[-\pi, \pi], \forall \varepsilon > 0 \text{ esiste un polinomio trigonometrico } p \text{ tale che } \|f - p\|_2 < \varepsilon.$$

Per dimostrare questa affermazione servono due risultati che non dimostriamo:

- (1) $C[-\pi, \pi]$ è denso in $L_2[-\pi, \pi]$;¹
 (2) un noto teorema di Weierstrass afferma: *se g è una funzione continua di periodo 2π , allora può essere scritta come il limite uniforme di una successione di polinomi trigonometrici.*

A questo punto procediamo. Sia f una funzione arbitraria in $L_2[-\pi, \pi]$ e sia $\varepsilon > 0$. Grazie al punto (1) esiste una funzione $g \in C[-\pi, \pi]$ tale che

$$(8.24) \quad \|f - g\|_2 < \varepsilon/3.$$

A questo punto c'è un piccolo problemino perché il teorema di Weierstrass si applica giustamente a funzioni continue e *periodiche*, mentre la nostra g potrebbe risultare discontinua quando la prolunghiamo periodicamente nel caso in cui sia $g(-\pi) \neq g(\pi)$. Ma, come già detto, si tratta solo di una fastidiosa distrazione. Infatti non è difficile far vedere che posso trovare un'altra funzione continua \hat{g} che è identica a g tranne nelle immediate vicinanze degli estremi dell'intervallo $[-\pi, \pi]$. In queste due regioni la g viene modificata in modo tale da avere, ad esempio, $\hat{g}(-\pi) = \hat{g}(\pi) = 0$, ma senza rovinare la continuità. Si vede che se quest'operazione di modifica viene fatta con opportuna delicatezza si ottiene

$$(8.25) \quad \|g - \hat{g}\|_2 < \varepsilon/3.$$

A questo punto possiamo applicare il teorema di Weierstrass alla funzione \hat{g} (o meglio al suo prolungamento periodico) e troviamo un polinomio trigonometrico p tale che

$$(8.26) \quad \|p - \hat{g}\|_u < \varepsilon/(3\sqrt{2\pi}).$$

Il fattore $3\sqrt{2\pi}$ al denominatore è stato introdotto, col senno di poi, per motivi estetici, ma se ne può fare a meno. Notare che qui compare la norma uniforme perché questo è esattamente quanto afferma W. Fortunatamente passare dalla norma uniforme a quella L_2 (su un intervallo finito!) è uno scherzo:

$$(8.27) \quad \|p - \hat{g}\|_2 = \left[\int_{-\pi}^{\pi} |p(x) - \hat{g}(x)|^2 dx \right]^{1/2} \leq \left[2\pi \frac{\varepsilon^2}{18\pi} \right]^{1/2} = \varepsilon/3.$$

Dalle (8.24), (8.25), (8.27) e dalla disuguaglianza triangolare della norma segue $\|f - p\|_2 < \varepsilon$ che è quello che volevamo dimostrare. \square

8.1.6 Relazione fra la serie di Fourier di f e quella di f'

8.6 Domanda. Supponiamo di conoscere la serie di Fourier di f . È lecito calcolare la serie di Fourier della derivata f' semplicemente derivando termine a termine? È lecito calcolare la serie di Fourier di una primitiva F semplicemente integrando termine a termine?

La risposta è: in generale no. Vediamo subito un esempio in cui la cosa non funziona. La serie di Fourier della funzione $f(x) = x$ su $[-\pi, \pi]$ è data da

$$(8.28) \quad x \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin(kx)}{k}$$

¹se $L_2[-\pi, \pi]$ viene *definito* come il completamento di $C_2[-\pi, \pi]$ questa affermazione è nient'altro che la definizione

Se derivo termine a termine ottengo che la serie di Fourier per la funzione $f(x) = 1$ è data da

$$1 \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \cos(kx),$$

risultato palesemente falso perché la serie di Fourier di 1 è la funzione stessa

$$1 \sim 1.$$

Se invece *integro* la (6.4) ottengo

$$\frac{x^2}{2} \sim -2 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \cos(kx) + C.$$

La costante additiva C non è altri che il termine costante $a_0/2$ che compare nello sviluppo di Fourier e va calcolato a parte.

$$C = \frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{x^2}{2} dx = \frac{1}{2\pi} \frac{\pi^3}{3} = \frac{\pi^2}{6}.$$

Ottingo quindi

$$(8.29) \quad x^2 \sim \frac{\pi^2}{3} - 4 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \cos(kx).$$

Questo sviluppo è giusto? Per verificare basta calcolare direttamente i coefficienti di Fourier della funzione x^2 tramite la (8.9). Si ottiene che lo sviluppo (8.29) è corretto! Osservo anche che se parto dalla (8.29) e derivo termine a termine ottengo la (6.4) che è giusta. Quindi l'operazione di derivazione termine a termine appare corretta se applicata alla (8.29), sbagliata se applicata alla (6.4).

Provo a vedere cosa succede se integro nuovamente la (8.29). Ottengo

$$(8.30) \quad \frac{x^3}{3} \sim \frac{\pi^2}{3} x - 4 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3} \sin(kx) + C.$$

Disastro! Che ci fa la funzione x nello sviluppo di Fourier? Chi l'ha invitata? La funzione x non fa parte della nostra base, quindi non può comparire in uno sviluppo di Fourier. Solo seni, coseni e costante. Però... forse si può sostituire la x che appare nella (8.30) con il suo sviluppo dato dalla (6.4). Provo.

$$\begin{aligned} x^3 &\sim \pi^2 x - 12 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^3} \sin(kx) + 3C \\ &= 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(k\pi)^2 - 6}{k^3} \sin(kx) + 3C. \end{aligned}$$

La costante additiva come al solito va calcolata a parte.

$$3C = \frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^3 dx = 0,$$

da cui concludo che lo sviluppo di x^3 ottenuto in questo modo poco ortodosso (integro termine a termine e sostituisco x con il suo sviluppo) verrebbe

$$(8.31) \quad x^3 \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(k\pi)^2 - 6}{k^3} \sin(kx).$$

Di nuovo, vado a verificare questa formula calcolandomi direttamente gli a_k e i b_k della funzione x^3 e trovo che la (8.31) è corretta. *Attenzione però*: se provo a tornare indietro derivando la (8.31) termine a termine trovo

$$(8.32) \quad 3x^2 \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(k\pi)^2 - 6}{k^2} \cos(kx).$$

che NON è lo sviluppo corretto di $3x^2$, come si vede confrontando con la (8.29).

Ok, è venuto il momento di dimostrare qualcosa, per avere, in questo mare di dubbi e incertezze, un appiglio solido e sicuro (ehm...).

8.7 Proposizione. *Sia $f \in C[-\pi, \pi]$ con f' continua a tratti. Siano a_k, b_k i coefficienti di Fourier di f e a'_k, b'_k i coefficienti di Fourier di f' . Allora*

$$(8.33) \quad a_k = -b'_k/k \quad b_k = (a'_k + (-1)^{k+1}a'_0)/k \quad k = 1, 2, 3, \dots$$

Se inoltre vale $f(\pi) = f(-\pi)$ si ha

$$(8.34) \quad a_k = -b'_k/k \quad b_k = a'_k/k \quad k = 1, 2, 3, \dots$$

Dimostrazione. La (8.33) si ottiene facilmente integrando per parti. Se $f(\pi) = f(-\pi)$ allora

$$a'_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) dx = \frac{1}{\pi} (f(\pi) - f(-\pi)) = 0,$$

e quindi la (8.34) segue dalla (8.33). \square

8.8 Osservazione. Le relazioni (8.33) ci dicono che se f è continua a tratti e F è una primitiva di f allora la serie di Fourier di F può essere ottenuta integrando formalmente la serie di Fourier di f , a patto di espandere, nella serie integrata, l'eventuale fattore lineare nuovamente in serie di Fourier. Più esplicitamente, sia

$$\begin{aligned} f(x) &\sim \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \\ F(x) &\sim \frac{A_0}{2} + \sum_{k=1}^{\infty} [A_k \cos(kx) + B_k \sin(kx)]. \end{aligned}$$

Integrando formalmente otterremmo

$$(8.35) \quad F(x) \sim \frac{a_0}{2} x + \sum_{k=1}^{\infty} \left[\frac{a_k}{k} \sin(kx) - \frac{b_k}{k} \cos(kx) \right] + C,$$

in cui $C = A_0/2$ è una costante da determinare tramite la relazione

$$A_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) dx.$$

Il termine lineare nella (8.35) va espanso in serie di Fourier. Sapendo che

$$x \sim 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin(kx)}{k},$$

otteniamo

$$F(x) \sim \frac{A_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k + (-1)^{k+1}a_0}{k} \sin(kx) - \frac{b_k}{k} \cos(kx) \right].$$

Possiamo a questo punto scrivere la relazione fra i coefficienti di Fourier di F e quelli di f

$$A_k = -b_k/k \quad B_k = (a_k + (-1)^{k+1}a_0)/k.$$

Ritroviamo in questo modo la (8.33) applicata alla coppia di funzioni f e F .

Se invece vogliamo *derivare* uno sviluppo di Fourier termine a termine, la condizione per poterlo fare è che la funzione f da derivare sia continua su $[-\pi, \pi]$, con derivata f' CAT e inoltre deve valere $f(-\pi) = f(\pi)$. In questo caso, infatti, valgono le relazioni (8.34) che dicono proprio che lo sviluppo di Fourier può essere derivato termine a termine. Negli esempi fatti in precedenza l'operazione di derivazione porta ad un risultato corretto nel caso di $f(x) = x^2$ che verifica la condizione $f(-\pi) = f(\pi)$, ma non nel caso di x o di x^3 .

8.9 Morale. Riassumo:

- per poter integrare lo sviluppo di Fourier di f termine a termine:
 - (1) f deve essere continua a tratti
 - (2) se il termine costante della f è non nullo, integrando compare la funzione x , che va sostituita con il suo sviluppo di Fourier
 - (3) la costante additiva di integrazione va calcolata esplicitamente
- per poter derivare lo sviluppo di Fourier di f termine a termine:
 - (1) f deve essere continua con derivata CAT
 - (2) deve valere $f(-\pi) = f(\pi)$.

Nel dubbio calcolare i coefficienti di Fourier usando le formule (8.9).

8.1.7 Serie di Fourier complessa

A volte può essere conveniente usare come base ortogonale nello spazio $L_2[-\pi, \pi]$, al posto di quella costituita da seni e coseni

$$(8.36) \quad 1, \cos(x), \cos(2x), \dots, \sin(x), \sin(2x), \dots$$

quella costituita dagli esponenziali con esponente immaginario di periodo sottomultiplo di 2π

$$\dots, e^{-i2x}, e^{-ix}, 1, e^{ix}, e^{i2x}, e^{i3x}, \dots$$

legata alla base precedente dalla relazione

$$e^{ikx} = \cos(kx) + i \sin(kx).$$

8.10 Proposizione. *L'insieme di vettori*

$$(8.37) \quad e_k(x) = e^{ikx} \quad k \in \mathbb{Z}$$

è una base ortogonale in $L_2[-\pi, \pi]$. Inoltre $\|e_k\|_2 = \sqrt{2\pi}$.

Dimostrazione. L'ortogonalità è immediata. Infatti, se $n \neq k$, si ha

$$\langle e_n, e_k \rangle = \int_{-\pi}^{\pi} e^{i(n-k)x} dx = 0.$$

Inoltre

$$\|e_k\|_2^2 = \int_{-\pi}^{\pi} |e^{-ikx}|^2 dx = \int_{-\pi}^{\pi} 1 dx = 2\pi.$$

Rimane da dimostrare la completezza della base (8.37). Questa discende immediatamente dalla completezza della base (8.36). Infatti, poiché le funzioni $\cos(kx)$ e $\sin(kx)$ sono una combinazione lineare delle funzioni e^{ikx} e e^{-ikx} e viceversa, questo vuol dire che

$$\text{span}\{e^{ikx}, e^{-ikx}\} = \text{span}\{\cos(kx), \sin(kx)\}.$$

visto che il sistema (8.36) è completo, anche il sistema (8.37) è completo. \square

La Proposizione 8.10 ci dice che vale lo sviluppo di Fourier

$$f(x) \sim \sum_{k \in \mathbb{Z}} c_k e^{ikx}$$

in cui

$$c_k = \frac{\langle f, e_k \rangle}{\|e_k\|_2^2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx.$$

Nel caso in cui f sia una funzione reale si ha

$$c_{-k} := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{ikx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(x) e^{ikx}} dx = \overline{c_k},$$

quindi possiamo scrivere

$$\begin{aligned} (8.38) \quad f(x) &\sim c_0 + \sum_{k=1}^{\infty} [c_k e^{ikx} + c_{-k} e^{-ikx}] \\ &= c_0 + \sum_{k=1}^{\infty} [c_k e^{ikx} + \overline{c_k e^{ikx}}] = c_0 + 2 \sum_{k=1}^{\infty} \text{Re}(c_k e^{ikx}) \quad (f \text{ reale}). \end{aligned}$$

8.11 Problema. Sviluppare in serie trigonometrica di Fourier nell'intervallo $[-\pi, \pi]$ la funzione $f(x) = e^x$. Dimostrare che

$$\sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{1}{2}(\pi \coth \pi - 1)$$

Soluzione. Usiamo la base degli esponenziali.

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{(1-ik)x} dx \\ &= \frac{1}{2\pi} \frac{1}{1-ik} [e^{(1-ik)x}]_{-\pi}^{\pi} = \frac{1}{2\pi} \frac{1}{1-ik} [e^{(1-ik)\pi} - e^{-(1-ik)\pi}] \end{aligned}$$

Osservo che

$$e^{-ik\pi} = e^{ik\pi} = (-1)^k,$$

per cui

$$\textcircled{1} \quad c_k = \frac{(-1)^k}{2\pi(1-ik)} [e^{\pi} - e^{-\pi}] = \frac{(-1)^k \sinh \pi}{\pi(1-ik)}.$$

Poiché f è reale il suo sviluppo di Fourier può essere scritto nella forma

$$f(x) \sim c_0 + 2 \sum_{k=1}^{\infty} \text{Re}(c_k e^{ikx}).$$

Quindi otteniamo

$$\begin{aligned} f(x) &\sim \frac{\sinh \pi}{\pi} \left[1 + 2 \sum_{k=1}^{\infty} (-1)^k \operatorname{Re} \left(\frac{e^{ikx}}{1-ik} \right) \right] \\ &= \frac{\sinh \pi}{\pi} \left[1 + 2 \sum_{k=1}^{\infty} (-1)^k \operatorname{Re} \left(\frac{(\cos(kx) + i \sin(kx))(1+ik)}{1+k^2} \right) \right] \\ &= \frac{\sinh \pi}{\pi} \left[1 + 2 \sum_{k=1}^{\infty} \frac{(-1)^k}{1+k^2} (\cos(kx) - k \sin(kx)) \right] \end{aligned}$$

Per dimostrare l'identità richiesta dal problema uso l'identità di Parseval

$$\textcircled{2} \quad \|f\|_2^2 = \sum_{k \in \mathbb{Z}} |c_k|^2 \|u_k\|_2^2 = 2\pi \sum_{k \in \mathbb{Z}} |c_k|^2.$$

Per quanto riguarda la norma di f ottengo

$$\|f\|_2^2 = \int_{-\pi}^{\pi} e^{2x} dx = \frac{1}{2} (e^{2\pi} - e^{-2\pi}) = \sinh(2\pi).$$

Per il modulo quadro di c_k uso l'espressione $\textcircled{1}$

$$|c_k|^2 = \frac{\sinh \pi^2}{\pi^2} \frac{1}{|1-ik|^2} = \frac{(\sinh \pi)^2}{\pi^2(1+k^2)}.$$

Dalla $\textcircled{2}$ ottengo quindi

$$\sinh(2\pi) = 2\pi \frac{(\sinh \pi)^2}{\pi^2} \left[1 + 2 \sum_{k=1}^{\infty} \frac{1}{1+k^2} \right],$$

da cui, siccome $\sinh(2\pi) = 2 \sinh(\pi) \cosh(\pi)$,

$$\sum_{k=1}^{\infty} \frac{1}{1+k^2} = \frac{1}{2} \left[\frac{\pi \sinh(2\pi)}{2(\sinh \pi)^2} - 1 \right] = \frac{1}{2} [\pi \coth \pi - 1].$$

8.2 Convergenza puntuale della serie di Fourier

Abbiamo visto che se f è una qualsiasi funzione appartenente ad $L_2[-\pi, \pi]$ allora possiamo scrivere

$$(8.39) \quad f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)]$$

in cui

$$(8.40) \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx.$$

Ricordo che la (8.39) è una scrittura simbolica che denota la convergenza della serie nel senso della norma $\|\cdot\|_2$. In sostanza la (8.39) significa che

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \left| f(x) - \left(\frac{a_0}{2} + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)] \right) \right|^2 dx = 0.$$

Vogliamo affrontare ora il problema della convergenza *puntuale* della serie di Fourier alla funzione f , capire cioè sotto quali condizioni è possibile scrivere

$$(8.41) \quad f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] .$$

Non è difficile vedere con un controesempio che, in generale, la convergenza rispetto alla norma $\|\cdot\|_2$ non implica la convergenza puntuale, per cui è ragionevole aspettarsi che la validità della (8.41) è subordinata ad ipotesi supplementari sulla funzione f .

8.2.1 Derivata destra e sinistra di una funzione DAT

8.12 Definizione. Sia $f : [a, b] \rightarrow \mathbb{C}$ una funzione continua a tratti e sia $x \in [a, b)$. f è detta *derivabile da destra* nel punto x se il limite

$$(8.42) \quad f'_+(x) := \lim_{\varepsilon \rightarrow 0^+} \frac{f(x + \varepsilon) - f(x^+)}{\varepsilon}$$

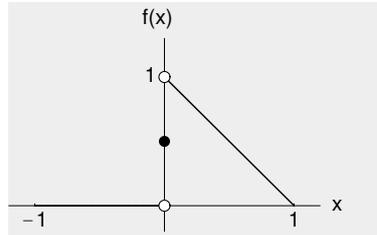
esiste ed è finito. La quantità $f'_+(x)$ è detta *derivata destra* della funzione f nel punto x . Se $x \in (a, b]$ si definisce, analogamente, la *derivata sinistra* di f in x come

$$(8.43) \quad f'_-(x) := \lim_{\varepsilon \rightarrow 0^+} \frac{f(x - \varepsilon) - f(x^-)}{-\varepsilon}$$

Si noti che la derivata destra (sinistra) è stata definita usando il valore $f(x^+)$ ($f(x^-)$) nel rapporto incrementale e *non* il valore $f(x)$. Il motivo è che in questo modo è possibile definire $f'_\pm(x)$ anche in casi in cui f non è continua in x .

8.13 Esempio. Consideriamo la funzione

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1/2 & \text{se } x = 0 \\ 1 - x & \text{se } x > 0 \end{cases}$$



In questo caso, usando le definizioni (8.42) e (8.43), otteniamo

$$\begin{aligned} f'_+(0) &= \lim_{\varepsilon \rightarrow 0^+} \frac{f(0 + \varepsilon) - f(0^+)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{1 - \varepsilon - 1}{\varepsilon} = -1 \\ f'_-(0) &= \lim_{\varepsilon \rightarrow 0^+} \frac{f(0 - \varepsilon) - f(0^-)}{-\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{0 - 0}{-\varepsilon} = 0. \end{aligned}$$

Però, se noi avessimo definito la derivata destra (ad esempio) nell'origine usando il valore $f(0)$ invece che il valore $f(0^+)$ nel rapporto incrementale avremmo ottenuto

$$f'_+(0) := \lim_{\varepsilon \rightarrow 0^+} \frac{f(0 + \varepsilon) - f(0)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{(1 - \varepsilon) - 1/2}{\varepsilon} = +\infty.$$

Quindi la funzione f non è continua nel punto $x = 0$ ma, in accordo con la nostra definizione, esistono sia la derivata destra che quella sinistra. Osserviamo che la derivata destra nell'origine coincide con il limite destro della derivata e, analogamente, la derivata sinistra coincide con il limite sinistro. Infatti, si ha

$$f'(x) = \begin{cases} 0 & \text{se } x < 0 \\ -1 & \text{se } x > 0 \end{cases}$$

per cui si ottiene (banalmente)

$$f'(0^-) = \lim_{x \rightarrow 0^-} f'(x) = 0 \quad f'(0^+) = \lim_{x \rightarrow 0^+} f'(x) = -1.$$

8.14 Domanda. È sempre vero che la derivata destra coincide con il limite destro della derivata e stessa cosa a sinistra? La risposta è: *quasi sempre*. Nel senso che, come spiegato nella Proposizione 8.15 e nell'osservazione 8.16, se la funzione f è differenziabile a tratti allora la risposta è affermativa, però può accadere che la derivata destra (ad esempio) $f'_+(x)$ esista ma il limite destro della derivata $f'(x^+)$ non esista.

8.15 Proposizione. Se $f : [a, b] \rightarrow \mathbb{C}$ è differenziabile a tratti e $x \in (a, b)$ allora esiste sia la derivata destra che quella sinistra di f in x . Inoltre si ha

$$\begin{aligned} f'_+(x) &= \lim_{\varepsilon \rightarrow 0^+} f'(x) =: f'(x^+) \\ f'_-(x) &= \lim_{\varepsilon \rightarrow 0^-} f'(x) =: f'(x^-). \end{aligned}$$

Dimostrazione. Dimostriamo solo l'affermazione riguardante la derivata destra, in quanto quella sulla derivata sinistra si dimostra in modo identico.

Se f è differenziabile a tratti in $[a, b]$ esiste $\gamma > 0$ tale che f è differenziabile nell'intervallo $(x, x + \gamma)$. Inoltre sappiamo che esiste il limite

$$(8.44) \quad f'(x^+) := \lim_{y \rightarrow x^+} f'(y).$$

Per definizione di limite destro la (8.44) ci dice che

$$(8.45) \quad \forall \alpha > 0 \exists \varepsilon_1 > 0 \text{ tale che se } y \in (x, x + \varepsilon_1) \text{ si ha } |f'(y) - f'(x^+)| < \alpha.$$

Per il Teorema di Lagrange (giusto?) sappiamo che, se $0 < \delta < \varepsilon$,

$$(8.46) \quad \frac{f(x + \varepsilon) - f(x + \delta)}{\varepsilon - \delta} = f'(\vartheta) \quad \text{in cui } \vartheta \in (x + \delta, x + \varepsilon).$$

Dalla (8.45) e (8.46) segue che, se $0 < \delta < \varepsilon \leq \varepsilon_1$ si ha

$$\left| \frac{f(x + \varepsilon) - f(x + \delta)}{\varepsilon - \delta} - f'(x^+) \right| < \alpha.$$

Di conseguenza

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{\delta \rightarrow 0^+} \left| \frac{f(x + \varepsilon) - f(x + \delta)}{\varepsilon - \delta} - f'(x^+) \right| \leq \alpha,$$

ovvero

$$\lim_{\varepsilon \rightarrow 0^+} \left| \frac{f(x + \varepsilon) - f(x^+)}{\varepsilon} - f'(x^+) \right| \leq \alpha.$$

Poiché α è arbitrario

$$\lim_{\varepsilon \rightarrow 0^+} \left| \frac{f(x + \varepsilon) - f(x^+)}{\varepsilon} - f'(x^+) \right| = 0,$$

ovvero la derivata destra $f'_+(x)$ esiste e vale

$$f'_+(x) := \lim_{\varepsilon \rightarrow 0^+} \frac{f(x + \varepsilon) - f(x^+)}{\varepsilon} = f'(x^+).$$

□

8.16 Osservazione. Può accadere che esista la derivata destra $f'_+(x)$, ma *non* il limite $\lim_{\varepsilon \rightarrow 0^+} f'(x)$. Consideriamo infatti la funzione

$$f(x) := x^2 \sin(1/x)$$

Abbiamo che

$$f(0^+) = \lim_{\varepsilon \rightarrow 0^+} f(x) = 0,$$

quindi

$$\lim_{\varepsilon \rightarrow 0^+} \frac{f(\varepsilon) - f(0^+)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{f(\varepsilon)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \varepsilon \sin(1/\varepsilon) = 0.$$

Questo ci dice che la derivata destra $f'_+(0)$ esiste e vale 0. D'altra parte

$$f'(x) = 2x \sin(1/x) - \cos(1/x),$$

per cui il limite

$$f'(0^+) := \lim_{\varepsilon \rightarrow 0^+} f'(\varepsilon)$$

non esiste a causa del termine $\cos(1/x)$.

8.2.2 Il teorema sulla convergenza puntuale

Per enunciare il teorema sulla convergenza puntuale della serie di Fourier conviene pensare ad un funzione definita sull'intervallo $[-\pi, \pi]$, come se fosse una funzione periodica definita su tutto l'asse reale.

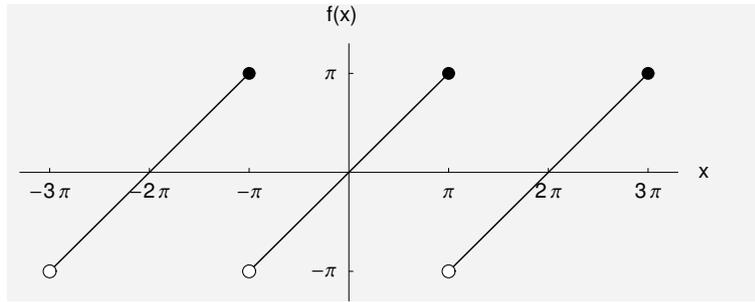
8.17 Definizione. Data una funzione $f : [-\pi, \pi] \rightarrow \mathbb{C}$, definiamo il *prolungamento periodico di f* la funzione $\tilde{f} : \mathbb{R} \rightarrow \mathbb{C}$ definita come

$$\tilde{f}(x) = f(x + 2k\pi)$$

in cui $k \in \mathbb{Z}$ è univocamente determinato dalla condizione $x + 2k\pi \in (-\pi, \pi]$.

8.18 Osservazione. Sia $f : [-\pi, \pi] \rightarrow \mathbb{C}$. Per come l'abbiamo definita, la funzione \tilde{f} coincide con f nell'intervallo $(-\pi, \pi]$, quindi se $f(-\pi) \neq f(\pi)$, allora \tilde{f} ha discontinuità di prima specie nei punti $(2k + 1)\pi$ con $k \in \mathbb{Z}$.

8.19 Esempio. Consideriamo la funzione $f(x) = x$ definita sull'intervallo $[-\pi, \pi]$. Il suo prolungamento periodico appare come in figura



8.20 Lemma. (Riemann–Lebesgue). Se $f \in L_1[a, b]$ allora per ogni $a, b \in \mathbb{R}$ con $a < b$ si ha

$$(8.47) \quad \lim_{s \rightarrow \infty} \int_a^b f(x) \sin(sx) dx = 0.$$

Dimostrazione. Dimostriamo questo Lemma nel caso particolare in cui f è continua con derivata continua a tratti. Per estendere il risultato a tutte le funzioni $f \in L_1[a, b]$ si usa il fatto che le funzioni continue con derivata CAT sono dense in $L_1[a, b]$ (in realtà anche le funzioni $C^1[a, b]$ sono dense in $L_1[a, b]$).

Sia quindi f continua con derivata CAT. Il Corollario 6.9 ci dice che possiamo integrare per parti, per cui

$$\int_a^b f(x) \sin(sx) dx = -\frac{1}{s} \left[f(x) \cos(sx) \right]_a^b + \frac{1}{s} \int_a^b f'(x) \cos(sx) dx$$

Poiché f è continua e f' è CAT, sia f che f' sono limitate in $[a, b]$. Possiamo porre

$$M := \sup_{x \in [a, b]} |f(x)| \quad M' := \sup_{x \in [a, b]} |f'(x)|.$$

Abbiamo quindi

$$\left| \int_a^b f(x) \sin(sx) dx \right| \leq \frac{1}{|s|} \left[|f(a)| + |f(b)| + |b - a| M' \right] \leq \frac{2M + |b - a| M'}{s}.$$

Facendo tendere s a ∞ otteniamo la (8.47). \square

8.21 Lemma. Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione continua a tratti e sia $x \in \mathbb{R}$ tale che la derivata destra di f nel punto x , $f'_+(x)$ esiste. Definiamo la funzione

$$(8.48) \quad D_n(z) := \frac{1}{2\pi} \frac{\sin \left[\frac{2n+1}{2} z \right]}{\sin(z/2)} = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^n \cos(kz)$$

Allora:

$$(8.49) \quad \lim_{n \rightarrow \infty} \int_0^\pi f(x+z) D_n(z) dz = \frac{f(x^+)}{2}.$$

Analogamente se esiste $f'_-(x)$ si ha

$$(8.50) \quad \lim_{n \rightarrow \infty} \int_{-\pi}^0 f(x+z) D_n(z) dz = \frac{f(x^-)}{2}.$$

Prima di dimostrare il Lemma 8.21 facciamo vedere come da questo segue facilmente il risultato che stiamo cercando sulla convergenza puntuale della serie di Fourier.

8.22 Teorema. *Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione 2π -periodica continua a tratti e sia S_n la somma parziale n -sima di Fourier di f . Se $x \in \mathbb{R}$ è un punto in cui esistono sia la derivata destra $f'_+(x)$ che quella sinistra $f'_-(x)$ di f , si ha:*

$$(8.51) \quad \lim_{n \rightarrow \infty} S_n(x) = \begin{cases} f(x) & \text{se } f \text{ è continua in } x \\ \frac{1}{2} [f(x^+) + f(x^-)] & \text{nel caso generale.} \end{cases}$$

Dimostrazione. Iniziamo a scrivere esplicitamente la somma parziale n -sima di Fourier associata ad f

$$\begin{aligned} S_n(x) &= \frac{a_0}{2} + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)] \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) dt + \sum_{k=1}^n \left[\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt \cos(kx) \right. \\ &\quad \left. + \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt \sin(kx) \right] \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{k=1}^n (\cos(kx) \cos(kt) + \sin(kx) \sin(kt)) \right] dt \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{k=1}^n \cos[k(t-x)] \right] dt. \end{aligned}$$

Scrivendo $\cos(ku) = \operatorname{Re} e^{iku}$ non è difficile far vedere che vale l'identità

$$\frac{1}{2} + \sum_{k=1}^n \cos[ku] = \frac{\sin \left[\frac{2n+1}{2} u \right]}{2 \sin \frac{u}{2}},$$

per cui, ricordando la definizione (8.48), otteniamo che la somma parziale n -sima può essere scritta come

$$S_n(x) = \int_{-\pi}^{\pi} f(t) D_n(t-x) dt.$$

A questo punto possiamo cambiare variabile di integrazione, ponendo $z = t - x$,

$$S_n(x) = \int_{-\pi-x}^{\pi-x} f(x+z) D_n(z) dz.$$

Sfruttando il fatto che f e D_n sono entrambi funzioni 2π -periodiche l'intervallo di integrazione, che equivale ad un periodo completo, può essere traslato di una quantità arbitraria, per cui

$$S_n(x) = \int_{-\pi}^{\pi} f(x+z) D_n(z) dz.$$

Per concludere la dimostrazione ed ottenere la (8.51) è sufficiente spezzare in due parti l'intervallo di integrazione $[-\pi, \pi]$ e ricordare quanto afferma il Lemma 8.21

$$\begin{aligned} \lim_{n \rightarrow \infty} S_n(x) &= \lim_{n \rightarrow \infty} \left[\int_{-\pi}^0 f(x+z) D_n(z) dz + \int_0^{\pi} f(x+z) D_n(z) dz \right] \\ &= \frac{1}{2} [f(x^+) + f(x^-)]. \quad \square \end{aligned}$$

Dimostrazione del Lemma 8.21. Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione continua a tratti e supponiamo che nel punto x esista la derivata destra $f'_+(x)$ della funzione f . Devo far vedere che si ha

$$(8.52) \quad \lim_{n \rightarrow \infty} \int_0^\pi f(x+z) D_n(z) dz = \frac{f(x^+)}{2}.$$

La dimostrazione della seconda affermazione (8.50) è identica. Inizio con l'osservare che, l'integrale di D_n esteso all'intervallo $[0, \pi]$ è uguale ad $1/2$. Infatti

$$(8.53) \quad \int_0^\pi D_n(z) dz = \frac{1}{\pi} \int_0^\pi \left[\frac{1}{2} + \sum_{k=1}^n \cos(kz) \right] dz.$$

Poiché si ha

$$\int_0^\pi \cos(kz) dz = \left[\frac{\sin(kz)}{k} \right]_0^\pi = 0,$$

otteniamo appunto

$$\int_0^\pi D_n(z) dz = \frac{1}{2}.$$

Questo mi istiga a pensare che se potessi sostituire la quantità $f(x+z)$ nell'integrale con $f(x^+)$ allora il valore dell'integrale sarebbe *esattamente* uguale a $f(x^+)/2$ senza neanche bisogno di dover fare il limite. Dunque, l'idea potrebbe essere quella di *aggiungere e togliere* $f(x^+)$ nell'integrando nel modo seguente:

$$(8.54) \quad \begin{aligned} & \int_0^\pi f(x+z) D_n(z) dz \\ &= \int_0^\pi f(x^+) D_n(z) dz + \int_0^\pi [f(x+z) - f(x^+)] D_n(z) dz \\ &= \frac{f(x^+)}{2} + \int_0^\pi [f(x+z) - f(x^+)] D_n(z) dz. \end{aligned}$$

A questo punto se riesco a dimostrare che

$$(8.55) \quad \lim_{n \rightarrow \infty} \int_0^\pi [f(x+z) - f(x^+)] D_n(z) dz = 0$$

avrei ottenuto proprio la (8.52) che è il risultato cercato. Per dimostrare la (8.55) mi viene in mente che forse potrei utilizzare il Lemma di Riemann–Lebesgue 8.20. Scrivo quindi

$$(8.56) \quad \begin{aligned} & \int_0^\pi [f(x+z) - f(x^+)] D_n(z) dz \\ &= \frac{1}{2\pi} \int_0^\pi [f(x+z) - f(x^+)] \frac{\sin \left[\frac{2n+1}{2} z \right]}{\sin(z/2)} dz \\ &= \frac{1}{2\pi} \int_0^\pi \frac{f(x+z) - f(x^+)}{\sin(z/2)} \sin \left[\frac{2n+1}{2} z \right] dz \\ &= \frac{1}{2\pi} \int_0^\pi \frac{f(x+z) - f(x^+)}{\sin(z/2)} \sin(sz) dz, \end{aligned}$$

in cui ho posto $s = (2n+1)/2$. Ora potrei applicare il Lemma di Riemann–Lebesgue e affermare che quando $n \rightarrow \infty$ (o, equivalentemente, $s \rightarrow \infty$) l'integrale tende a zero, a

patto di poter dire che la funzione che moltiplica il seno appartiene a $L_1[0, \pi]$. Rimane dunque da dimostrare che

$$(8.57) \quad \text{la funzione } g(z) := \frac{f(x+z) - f(x^+)}{\sin(z/2)} \quad \text{appartiene ad } L_1[0, \pi].$$

In realtà dimostrerò che g è limitata in $[0, \pi]$ che implica $g \in L_1[0, \pi]$. Per dimostrare che g è limitata osservo che g è continua a tratti in $(0, \pi]$ poiché il numeratore è CAT e il denominatore è continuo e non nullo. Se riesco a far vedere inoltre che il limite

$$(8.58) \quad \lim_{z \rightarrow 0^+} g(z)$$

esiste ed è finito, allora avrò dimostrato che g è CAT, e quindi limitata, su $[0, \pi]$. Ma il limite (8.58) si può calcolare esplicitamente. Infatti

$$\begin{aligned} \lim_{z \rightarrow 0^+} g(z) &= \lim_{z \rightarrow 0^+} \frac{f(x+z) - f(x^+)}{\sin(z/2)} \\ &= \lim_{z \rightarrow 0^+} \frac{f(x+z) - f(x^+)}{z} \frac{z}{\sin(z/2)} = 2f'_+(x). \end{aligned}$$

Ricapitolando: g è una funzione CAT in $(0, \pi]$, inoltre ho fatto vedere che il limite quando z tende a 0 da destra esiste ed è finito, quindi g è CAT su $[0, \pi]$. Di conseguenza g è limitata e, a maggior ragione, appartiene a $L_1[-\pi, \pi]$. Questo mi permette di usare il Lemma di Riemann–Lebesgue nella (8.56) e posso così dimostrare la (8.55). Finalmente dalla (8.54) e (8.55) segue la (8.52). \square

8.23 Corollario. Se $f : \mathbb{R} \rightarrow \mathbb{C}$ è una funzione 2π -periodica e differenziabile a tratti allora

$$(8.59) \quad \lim_{n \rightarrow \infty} S_n(x) = \begin{cases} f(x) & \text{se } f \text{ è continua in } x \\ \frac{1}{2} [f(x^+) + f(x^-)] & \text{nel caso generale.} \end{cases}$$

Dimostrazione. È una conseguenza diretta del Teorema 8.22 e della Proposizione 8.15. \square

8.24 Osservazione. Esistenza di funzioni continue con serie di Fourier non convergente.

8.25 Problema. Per ciascuna delle seguenti funzioni $f : [-\pi, \pi] \mapsto \mathbb{R}$ sia S_n la somma parziale n -esima della serie di Fourier associata. Calcolare il limite $S(x) := \lim_{n \rightarrow \infty} S_n(x)$ per ogni $x \in [-\pi, \pi]$.

$$(a) \quad f(x) = x^3 \qquad (b) \quad f(x) = \begin{cases} 0 & \text{se } x \in [-\pi, 0) \\ 1 - x^2 & \text{se } x \in [0, \pi] \end{cases}$$

Soluzione. (a). Si tratta di una funzione differenziabile a tratti, quindi vale il teorema che assicura la convergenza di $S_n(x)$ al valore $f(x)$ nei punti in cui f è continua, al valore $(f(x^-) + f(x^+))/2$ nei punti in cui f ha un salto.

Attenzione: per studiare la convergenza agli estremi dell'intervallo $[-\pi, \pi]$ bisogna sempre pensare al *prolungamento periodico* di f .

Mentre la funzione x^3 è continua in $[-\pi, \pi]$, il suo prolungamento periodico è continuo nei punti interni all'intervallo, cioè è continuo in $(-\pi, \pi)$ ma ha un salto nei punti

$x = -\pi$ e $x = \pi$. In entrambi questi punti il limite da sinistra vale π^3 , mentre il limite da destra vale $-\pi^3$. Quindi ottengo, per $x = \pm\pi$,

$$S(\pm\pi) = \frac{1}{2} (f(\pi^-) + f(\pi^+)) = \frac{1}{2} (\pi^3 - \pi^3) = 0.$$

Riassumendo

$$S(x) = \begin{cases} x^3 & \text{se } x \in (-\pi, \pi) \\ 0 & \text{se } x = -\pi \text{ o } x = \pi \end{cases}$$

(b). I punti di discontinuità del prolungamento periodico di f sono $x = -\pi$, $x = 0$ e $x = \pi$. Quindi

$$S(\pm\pi) = \frac{1}{2} (f(\pi^-) + f(\pi^+)) = \frac{1}{2} (1 - \pi^2 + 0) = \frac{1 - \pi^2}{2}$$

$$S(0) = \frac{1}{2} (f(0^-) + f(0^+)) = \frac{1}{2} (0 + 1) = \frac{1}{2}.$$

Riassumendo

$$S(x) = \begin{cases} 0 & \text{se } x \in (-\pi, 0) \\ 1/2 & \text{se } x = 0 \\ 1 - x^2 & \text{se } x \in (0, \pi) \\ (1 - \pi^2)/2 & \text{se } x = -\pi \text{ o } x = \pi \end{cases}$$

8.3 Quale sviluppo scegliere?

In generale è possibile sviluppare una funzione f in serie di Fourier in più di un modo. La scelta dello sviluppo da adottare va fatta in base alle esigenze o alla richiesta esplicita del problema che si sta risolvendo. Consideriamo il problema seguente: sviluppare in serie di Fourier nell'intervallo $[0, \pi]$ la funzione

$$f(x) = \frac{x^2}{2} - 2.$$

Sviluppo n.1. Possiamo usare le formule (8.13), (8.14) per lo sviluppo in serie di Fourier in un intervallo arbitrario $[a, b]$. Questo equivale considerare la funzione f definita in $[0, \pi]$ è prolungarla periodicamente a tutto l'asse reale, come si vede nella figura 8.1.

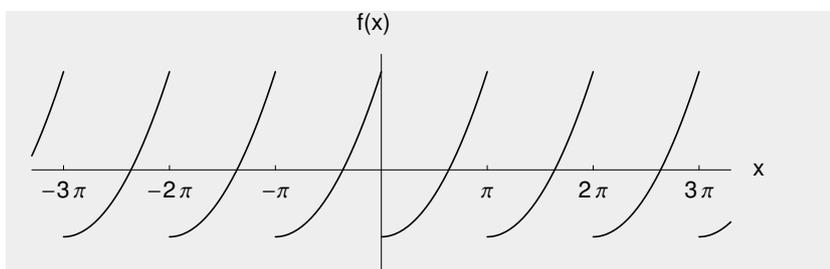


Figura 8.1: Sviluppo n.1

Sviluppo n.2. Un secondo sviluppo possibile è quello che usa solo i coseni, dato dalle formule (8.15), (8.17). Questo equivale a prendere la funzione f definita in $[0, \pi]$,

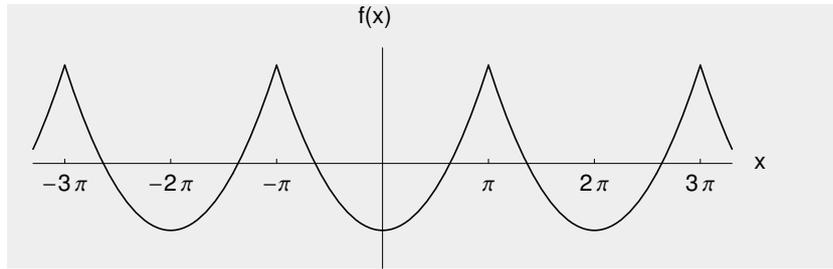


Figura 8.2: Sviluppo n.2

prolungarla prima in modo *pari* sull'intervallo $[-\pi, \pi]$ e poi prolungarla periodicamente a tutto l'asse reale come mostrato nella figura 8.2.

Sviluppo n.3. Possiamo infine usare lo sviluppo dei seni dato dalle formule (8.16), (8.17), che equivale a prendere la funzione f definita in $[0, \pi]$, prolungarla prima in modo *dispari* sull'intervallo $[-\pi, \pi]$ e poi prolungarla periodicamente a tutto l'asse reale come mostrato nella figura 8.3.

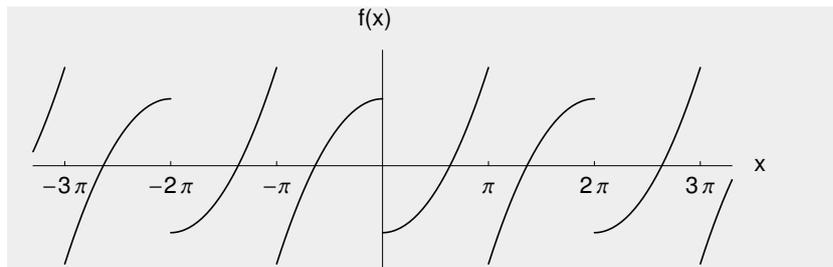


Figura 8.3: Sviluppo n.3

Dove convergono questi sviluppi? Chiamo $S^{(r)}(x)$ la somma della serie di Fourier associato allo sviluppo r -simo, in cui r può valere 1, 2, 3. In tutti e tre i casi ho a che fare con funzioni differenziabili a tratti, per cui vale il Corollario 8.23. Ottengo quindi,

$$S^{(1)}(x) = \begin{cases} f(x) & x \in (0, \pi) \\ \frac{\pi^2}{4} - 2 & x \in \{0, \pi\} \end{cases} \quad S^{(2)}(x) = f(x) \quad x \in [0, \pi]$$

$$S^{(3)}(x) = \begin{cases} f(x) & x \in (0, \pi) \\ 0 & x \in \{0, \pi\} \end{cases}$$

8.4 Equazione del calore su un intervallo finito

8.4.1 Condizioni al bordo di Neumann

Voglio risolvere la seguente equazione differenziale a derivate parziali (equazione del calore) utilizzando il metodo dello sviluppo in serie di Fourier:

$$(8.60) \quad u_t(x, t) = \lambda u_{xx}(x, t) \quad x \in [0, \ell], \quad t > 0$$

$$(8.61) \quad u(x, 0) = f(x) \quad x \in [0, \ell]$$

$$(8.62) \quad u_x(0, t) = u_x(\ell, t) = 0 \quad t > 0,$$

in cui λ è un numero reale positivo e le lettere x e t usate come pedici denotano, al solito, l'operazione di derivata rispetto a quel simbolo, cioè

$$u_t := \frac{\partial u}{\partial t} \quad u_x := \frac{\partial u}{\partial x}.$$

La funzione $u(x, t)$ è definita sulla semistriscia infinita $[0, \ell] \times [0, \infty)$. Per chi si trova più a suo agio pensando ad una situazione fisica concreta piuttosto che ad un'astrazione matematica, si pensi ad una barra unidimensionale² di lunghezza ℓ fatta di un qualche materiale che conduce calore. La funzione $u(x, t)$ rappresenta la temperatura della barra nel punto x al tempo t . Il parametro $\lambda > 0$ è una cosa tipo la *conducibilità termica* della barra. E infatti se potessimo $\lambda = 0$ otterremmo dalla (8.60) che la derivata rispetto al tempo di u è nulla, vale a dire la temperatura è costante nel tempo perché la barra non conduce calore. La funzione f che appare nella (8.61) è la temperatura iniziale, o meglio la distribuzione iniziale della temperatura della barra. Le equazioni (8.62) infine sono le *condizioni al bordo* di questo problema. In questo caso si chiamano condizioni Neumann perché riguardano la *derivata* della funzione u nei due punti di bordo ($x = 0$ e $x = \ell$) del problema. Imporre che la derivata spaziale della temperatura sia nulla corrisponde all'assunzione fisica che la barra sia termicamente isolata dall'esterno, senza cioè scambio di calore ai due estremi.

Bene, dobbiamo quindi trovare $u(x, t)$ che soddisfi le (8.60), (8.61), (8.62). L'idea è quella di supporre che per ogni valore fissato di $t \geq 0$, la funzione $u(x, t)$ si possa scrivere come serie di Fourier nel dominio delle x , con coefficienti di Fourier che ovviamente dipenderanno dal tempo. Ora possono nascere i primi dubbi e tentennamenti, perché abbiamo visto che c'è più di una modalità di sviluppare una funzione in serie di Fourier. Potremmo, per dirne una, usare gli esponenziali complessi e, considerando che siamo sull'intervallo $[0, \ell]$, scriviamo

$$u(x, t) = \sum_{k \in \mathbb{Z}} c_k(t) e^{i2k\pi x/\ell},$$

oppure potremmo usare la serie di Fourier dei seni

$$u(x, t) = \sum_{k=1}^{\infty} b_k(t) \sin(k\pi x/\ell)$$

o magari quella dei coseni. L'elemento del problema che ci indica la via giusta da seguire è la condizione al bordo (8.62). Poiché le quantità $u_x(0, t)$ e $u_x(\ell, t)$ devono essere entrambi nulle *per ogni* $t > 0$ sarebbe carino che la funzione $u_x(x, t)$ si potesse esprimere come una somma di $\sin(k\pi x/\ell)$, in modo tale che la condizione al bordo sia automaticamente soddisfatta. Di conseguenza la scelta che sembra appropriata è quella di espandere $u(x, t)$ come serie dei coseni, in modo tale che, derivando rispetto ad x otteniamo una serie di seni. Poniamo quindi

$$(8.63) \quad u(x, t) = \frac{a_0(t)}{2} + \sum_{k=1}^{\infty} a_k(t) \cos\left[\frac{k\pi x}{\ell}\right].$$

Assumendo di poter derivare la serie termine a termine otteniamo

$$\begin{aligned} u_t(x, t) &= \frac{\dot{a}_0(t)}{2} + \sum_{k=1}^{\infty} \dot{a}_k(t) \cos\left[\frac{k\pi x}{\ell}\right] \\ u_{xx}(x, t) &= - \sum_{k=1}^{\infty} \left(\frac{k\pi}{\ell}\right)^2 a_k(t) \cos\left[\frac{k\pi x}{\ell}\right]. \end{aligned}$$

²uhm... non che una barra unidimensionale sia esattamente una cosa "concreta"

Sostituendo queste espressioni nell'equazione $u_t = \lambda u_{xx}$ si ottiene un numero infinito di equazioni differenziali *ordinarie* per i coefficienti di Fourier

$$(8.64) \quad \dot{a}_0(t) = 0 \quad \dot{a}_k(t) = -\lambda \left(\frac{k\pi}{\ell}\right)^2 a_k(t) \quad k = 1, 2, \dots$$

che si risolvono facilmente

$$a_0(t) = a_0 \quad a_k(t) = a_k \exp\left[-\lambda \left(\frac{k\pi}{\ell}\right)^2 t\right],$$

in cui, per semplicità, abbiamo posto

$$a_n := a_n(0) \quad n = 0, 1, 2, \dots$$

Sostituendo nella (8.63) otteniamo infine la funzione incognita $u(x, t)$

$$(8.65) \quad u(x, t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \exp\left[-\lambda \left(\frac{k\pi}{\ell}\right)^2 t\right] \cos\left[\frac{k\pi x}{\ell}\right].$$

Rimangono da determinare i coefficienti a_k . A questo scopo usiamo la condizione iniziale del problema (8.61)

$$u(x, 0) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left[\frac{k\pi x}{\ell}\right] = f(x).$$

Questa uguaglianza ci dice che i coefficienti a_k sono nient'altro che i coefficienti associati alla funzione data f , quando questa viene sviluppata in serie di Fourier dei coseni nell'intervallo $[0, \ell]$. Di conseguenza abbiamo

$$(8.66) \quad a_k = \frac{2}{\ell} \int_0^{\ell} f(y) \cos\left[\frac{k\pi y}{\ell}\right] dy.$$

La (8.65) insieme alla (8.66) risolve il problema posto

Comportamento della soluzione quando $t \rightarrow \infty$

Qual è il comportamento della soluzione trovata quando il tempo tende all'infinito? Senza bisogno di essere dei grandi geni di equazioni differenziali a derivate parziali si intuisce che, se è vero che questa equazione rappresenta l'evoluzione della distribuzione della temperatura in una barra *isolata termicamente*, allora, per $t \rightarrow \infty$, la temperatura dovrà essere uniforme (indipendente da x) e (probabilmente) assumerà il valore medio della distribuzione iniziale della temperatura.

Verifico. Facendo tendere t all'infinito nella (8.65) tutti i termini tranne il primo tendono a zero. Inoltre la (8.66) ci dice che

$$a_0 = \frac{2}{\ell} \int_0^{\ell} f(y) dy.$$

Quindi abbiamo

$$\lim_{t \rightarrow \infty} u(x, t) = \frac{a_0}{2} = \frac{1}{\ell} \int_0^{\ell} f(y) dy,$$

che è proprio il valor medio della temperatura iniziale.

8.4.2 Condizioni al bordo di Dirichlet

Consideriamo ora lo stesso problema con diverse condizioni al bordo:

$$(8.67) \quad u_t(x, t) = \lambda u_{xx}(x, t) \quad x \in [0, \ell] \quad t > 0$$

$$(8.68) \quad u(x, 0) = f(x) \quad x \in [0, \ell]$$

$$(8.69) \quad u(0, t) = u(\ell, t) = 0 \quad t > 0.$$

La condizione (8.69) ci dice che la barra non è più isolata termicamente, come nel caso precedente, ma i suoi estremi vengono costantemente mantenuti a temperatura zero.

...

8.5 Convergenza uniforme della serie di Fourier

Il Corollario 8.23 ci dice che se f è periodica, differenziabile a tratti, allora la serie di Fourier di f converge puntualmente ad f in tutti i punti in cui f è continua. Sorge dunque spontanea la

8.26 Domanda. Non sarà che, niente niente, la convergenza della serie di Fourier è uniforme?

Senza qualche ipotesi addizionale la risposta è necessariamente negativa. Infatti la successione di funzioni $S_n(x)$, in cui, come al solito, abbiamo posto

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)],$$

è una successione di funzioni *continue*. Ora sappiamo che se una successione di funzioni continue converge uniformemente, il suo limite è una funzione continua, quindi affinché S_n abbia una qualche speranza di convergere uniformemente ad f bisogna assumere che f sia continua. Abbiamo infatti il seguente risultato:

8.27 Teorema. Sia $f : \mathbb{R} \rightarrow \mathbb{C}$ una funzione 2π -periodica, continua, con derivata continua a tratti. Allora la somma parziale n -sima di Fourier S_n di f converge ad f uniformemente, quando $n \rightarrow \infty$.

Dimostrazione. Le ipotesi del teorema insieme al Corollario 8.22 ci dicono che la serie di Fourier converge puntualmente ad f su tutto l'asse reale, per cui possiamo scrivere

$$(8.70) \quad f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \quad x \in \mathbb{R}.$$

Rimane da dimostrare che la serie di funzioni che appare al secondo membro converge uniformemente. L'idea è quella di usare il criterio di Weierstrass per la convergenza uniforme di una serie di funzioni. Devo cioè riuscire a maggiorare il modulo del termine k -simo della serie con una costante

$$|a_k \cos(kx) + b_k \sin(kx)| \leq M_k$$

e sperare che sia

$$\sum_{k=1}^{\infty} M_k < \infty.$$

Una possibilità che non richiede particolari doti di immaginazione³ è quella di scrivere

$$(8.71) \quad |a_k \cos(kx) + b_k \sin(kx)| \leq |a_k \cos(kx)| + |b_k \sin(kx)| \leq |a_k| + |b_k|.$$

Il problema è che i coefficienti di Fourier a_k e b_k sono, a priori, in ℓ_2 e non in ℓ_1 , per cui possiamo affermare che la *sommatoria dei quadrati* è convergente, ma non sappiamo dire nulla sulla convergenza di

$$\sum_{k=1}^{\infty} [|a_k| + |b_k|].$$

Fra le ipotesi del Teorema c'è però anche il fatto che f' è continua a tratti. Di conseguenza f' è limitata, quindi $f' \in L_2[-\pi, \pi]$. Questo implica che f' ha una serie di Fourier convergente in L_2 .

$$f'(x) \sim \frac{a'_0}{2} + \sum_{k=1}^{\infty} [a'_k \cos(kx) + b'_k \sin(kx)].$$

Non possiamo far affermazioni sulla convergenza puntuale di questo sviluppo, ma non è importante. Quello che ci interessa è il fatto che i coefficienti della serie di Fourier di f' sono in ℓ_2 , cioè

$$(8.72) \quad \sum_{k=1}^{\infty} |a'_k|^2 < \infty \quad \sum_{k=1}^{\infty} |b'_k|^2 < \infty.$$

Poiché f è continua e f' è continua a tratti e poiché f è periodica possiamo usare le relazioni (8.34) che legano i coefficienti di Fourier di f a quelli di f' . Abbiamo dunque

$$(8.73) \quad a_k = -b'_k/k \quad b_k = a'_k/k \quad k = 1, 2, 3, \dots$$

Possiamo quindi scrivere

$$(8.74) \quad \sum_{k=1}^{\infty} [|a_k| + |b_k|] = \sum_{k=1}^{\infty} [|a'_k|/k + |b'_k|/k].$$

Dalla disuguaglianza $ab \leq (|a|^2 + |b|^2)/2$ otteniamo

$$(8.75) \quad \sum_{k=1}^{\infty} [|a'_k|/k + |b'_k|/k] \leq \frac{1}{2} \sum_{k=1}^{\infty} [|a'_k|^2 + 1/k^2 + |b'_k|^2 + 1/k^2]$$

Dalle relazioni (8.74), (8.75) e (8.72) si ottiene infine che la sommatoria

$$\sum_{k=1}^{\infty} [|a_k| + |b_k|]$$

è convergente. Questo fatto, insieme alla disuguaglianza (8.71) ci permette di usare il criterio di Weierstrass e di concludere che la serie (8.70) è uniformemente convergente. \square

³seguendo il vecchio adagio romano che consiglia in questi casi "Nun te 'nventà gnente".

9. Trasformata di Fourier

V: Also, you know what they call a Quarter Pounder with Cheese in Paris?

J: They don't call it a Quarter Pounder with Cheese?

V: No, they got the metric system there, they wouldn't know what the fuck a Quarter Pounder is.

9.1 L'idea

9.1 Notazione. In questo capitolo le funzioni sono a priori a valori complessi, a meno che non sia detto esplicitamente il contrario, quindi, ad esempio

$$L_p(\mathbb{R}) := L_p(\mathbb{R}; \mathbb{C}) \quad \mathcal{S}(\mathbb{R}) := \mathcal{S}(\mathbb{R}; \mathbb{C}) \quad C^\infty(\mathbb{R}) := C^\infty(\mathbb{R}; \mathbb{C}).$$

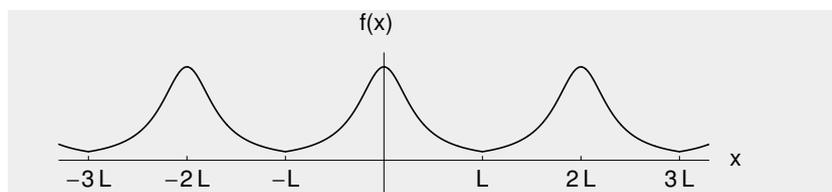
Il metodo dello sviluppo in serie di Fourier ci permette di scrivere una qualsiasi funzione periodica come combinazione lineare (infinita) di seni e coseni, oppure di esponenziali a esponente immaginario (funzioni periodiche elementari).

9.2 Domanda. È possibile estendere in qualche modo questo metodo allo scopo di poter trattare funzioni non periodiche definite su tutto l'asse reale?

Sia ad esempio

$$f(x) = \frac{1}{1+x^2} \quad x \in \mathbb{R}.$$

Si può scrivere f come combinazione lineare di funzioni periodiche elementari? L'idea è la seguente: invece di considerare f come una funzione definita su \mathbb{R} , la pensiamo come se fosse definita sull'intervallo $[-L, L]$, e poi ne facciamo il prolungamento periodico. A questo punto abbiamo una funzione periodica di periodo $2L$ che chiamiamo f_L .



Dunque possiamo usare lo sviluppo di Fourier, ad esempio quello in termini degli esponenziali immaginari. Si ottiene così

$$(9.1) \quad f_L(x) \sim \sum_{k \in \mathbb{Z}} c_k e^{ik\pi x/L} \quad c_k = \frac{1}{2L} \int_{-L}^L f_L(x) e^{-ik\pi x/L} dx.$$

L'idea della trasformata di Fourier a questo punto è la seguente: *faccio tendere nella (9.1) L all'infinito*. In questo modo la funzione f_L tende alla mia funzione originale f .

Come vedremo nel Teorema 9.27, quando $L \rightarrow \infty$ la somma su k diventa un integrale di Riemann e si ottiene (se f è continua) il risultato seguente:

$$(9.2) \quad f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} g(\lambda) e^{i\lambda x} d\lambda \quad g(\lambda) = \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx.$$

La formula a sinistra ci dice che, nel caso di una funzione non periodica, è ancora possibile scrivere f come “combinazione lineare infinita” di esponenziali immaginari, a patto di considerare come combinazione lineare infinita non una serie ma un integrale. Non avremo più, quindi, una successione di coefficienti di Fourier c_k , ma una *funzione di coefficienti* $g(\lambda)$. La funzione g è detta *trasformata di Fourier di f* .

9.2 Parità, traslazioni, dilatazioni

Sia $f: \mathbb{R} \rightarrow \mathbb{C}$. Denotiamo con Pf la funzione ottenuta applicando l'operatore *parità* P a f , definito come

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

Dato $a \in \mathbb{R}$, denotiamo con T_a l'operatore di *traslazione* definito come

$$(T_a f)(x) = f(x - a).$$

Osservate che se a è positivo, il grafico di $T_a f$ si ottiene a partire da quello di f mediante una traslazione nel verso positivo dell'asse x . Infine, se $a \in \mathbb{R}$ e $a \neq 0$, indichiamo con Φ_a l'operatore di *dilatazione/contrazione* dell'asse x , definito come

$$(9.3) \quad (\Phi_a f)(x) = f(ax).$$

Se $a > 1$, l'operatore Φ_a induce una *contrazione* (nella direzione x) del grafico della funzione a cui viene applicato, mentre se $a \in (0, 1)$ si tratta di una dilatazione. Se $a < 0$ la dilatazione/contrazione è composta con un'operazione di parità. In particolare se $a = -1$ si ha ovviamente $\Phi_a = P$. Traslazioni e dilatazioni non commutano. Si ha infatti

$$(9.4) \quad (\Phi_a T_b f)(x) = (T_b f)(ax) = f(ax - b)$$

$$(9.5) \quad (T_b \Phi_a f)(x) = (\Phi_a f)(x - b) = f(a(x - b)) = f(ax - ab),$$

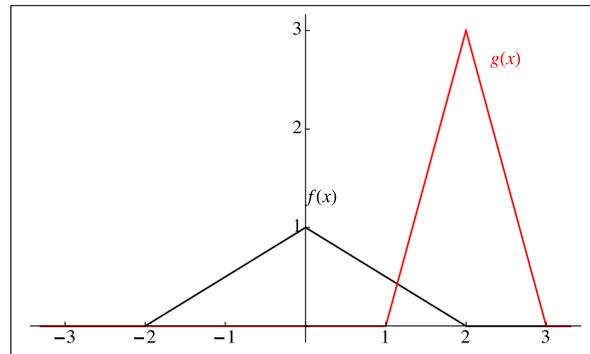
quindi

$$(9.6) \quad T_b \Phi_a = \Phi_a T_{ab}.$$

È importante dunque fare attenzione all'ordine con il quale vengono fatti agire gli operatori T_a e Φ_b . Supponiamo ad esempio di voler ottenere la funzione g della figura 9.1, applicando opportunamente gli operatori di traslazione e di dilatazione alla funzione f della figura stessa. Si vede il grafico di g si ottiene a partire da quello di f mediante: (a) una contrazione nella direzione x di un fattore 2; (b) una dilatazione nella direzione y di un fattore 3; (c) una traslazione verso destra di 2 unità. L'ordine corretto per ottenere g a partire da f è

$$(9.7) \quad g = 3 T_2 \Phi_2 f$$

vale a dire bisogna prima operare la contrazione lungo l'asse x di un fattore 2 e poi la traslazione verso destra di 2 unità (la moltiplicazione per 3 può essere fatta in qualsiasi momento perché T e Φ sono lineari). Se, al contrario, applicassimo prima la traslazione T_2 e poi la contrazione Φ_2 , siccome la contrazione modifica tutte le distanze sull'asse x , la funzione risultante $g' = 3 \Phi_2 T_2 f$ avrebbe un grafico centrato non in $x = 2$ ma in $x = 1$.

Figura 9.1: $g = 3 T_2 \Phi_2 f$.

9.3 Definizione e proprietà elementari

9.3 Definizione. Se $f \in L_1(\mathbb{R})$ si definisce *trasformata di Fourier* di f e si indica con $\mathcal{F}(f)$ oppure con \hat{f} , la funzione

$$\mathcal{F}(f)(\lambda) = \hat{f}(\lambda) := \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx \quad \lambda \in \mathbb{R}.$$

Osservo che, poiché $|f(x) e^{-i\lambda x}| = |f(x)|$, se $f \in L_1(\mathbb{R})$, anche la funzione $f(x) e^{-i\lambda x}$ appartiene a $L_1(\mathbb{R})$, quindi l'integrale è ben definito.

9.4 Proposizione. Siano $f, h \in L_1(\mathbb{R})$. Allora:

- (1) $\mathcal{F}[\alpha f + \beta h] = \alpha \mathcal{F}[f] + \beta \mathcal{F}[h]$ per ogni $\alpha, \beta \in \mathbb{C}$;
- (2) $\mathcal{F}[\bar{f}](\lambda) = \overline{\mathcal{F}[f](-\lambda)}$;
- (3) Se f è reale $\mathcal{F}[f](\lambda) = \overline{\mathcal{F}[f](-\lambda)}$;
- (4) Se f è reale pari $\mathcal{F}[f]$ è reale e pari;
- (5) Se f è reale dispari $\mathcal{F}[f]$ è immaginaria pura dispari.

Dimostrazione. Puramente meccanica, lasciata come esercizio.

9.5 Proposizione. Sia $f \in L_1(\mathbb{R})$ e $\hat{f} := \mathcal{F}[f]$. Allora, per ogni $a, b \in \mathbb{R}$ valgono le seguenti identità:

- (1) $\mathcal{F}[f(-x)](\lambda) = \hat{f}(-\lambda)$;
- (2) $\mathcal{F}[f(x-a)](\lambda) = e^{-i\lambda a} \hat{f}(\lambda)$;
- (3) $\mathcal{F}[f(x) e^{iax}](\lambda) = \hat{f}(\lambda - a)$;
- (4) $\mathcal{F}[f(x) \cos(ax)](\lambda) = \frac{1}{2} [\hat{f}(\lambda - a) + \hat{f}(\lambda + a)]$;
- (5) $\mathcal{F}[f(x) \sin(ax)](\lambda) = \frac{1}{2i} [\hat{f}(\lambda - a) - \hat{f}(\lambda + a)]$;
- (6) $\mathcal{F}[f(ax)](\lambda) = \frac{1}{|a|} \hat{f}(\lambda/a)$, se $a \neq 0$;
- (7) $\mathcal{F}[f(a(x-b))](\lambda) = \frac{1}{|a|} e^{-i\lambda b} \hat{f}(\lambda/a)$, se $a \neq 0$;

$$(8) \mathcal{F}[f(ax - b)](\lambda) = \frac{1}{|a|} e^{-i\lambda(b/a)} \hat{f}(\lambda/a), \text{ se } a \neq 0.$$

Dimostrazione. Anche queste identità derivano direttamente dalla definizione. Come esempio dimostriamo la (2). Le altre sono lasciate per esercizio.

$$\begin{aligned} \mathcal{F}[f(x - a)](\lambda) &= \int_{\mathbb{R}} f(x - a) e^{-i\lambda x} dx \\ &= \int_{\mathbb{R}} f(t) e^{-i\lambda(t+a)} dt && [t = x - a] \\ &= e^{-i\lambda a} \int_{\mathbb{R}} f(t) e^{-i\lambda t} dt = e^{-i\lambda a} \hat{f}(\lambda). \end{aligned}$$

□

9.6 Osservazione. Le proprietà enunciate nella Proposizione 9.5 possono essere espresse in termini degli operatori di parità, traslazione e dilatazione:

- (1) $\mathcal{F} \circ P = P \circ \mathcal{F}$;
- (2) $(\mathcal{F}T_a f)(\lambda) = e^{-i\lambda a} \hat{f}(\lambda)$;
- (3) $\mathcal{F}[f(x) e^{iax}] = T_a \hat{f}$;
- (4) $\mathcal{F}[f(x) \cos(ax)] = \frac{1}{2}(T_a \hat{f} + T_a)$;
- (5) $\mathcal{F}[f(x) \sin(ax)] = \frac{1}{2i}(T_a \hat{f} - T_{-a} \hat{f})$;
- (6) $\mathcal{F} \Phi_a f = \frac{1}{|a|} \Phi_{1/a} \mathcal{F}f$, se $a \neq 0$;
- (7) $(\mathcal{F}T_b \Phi_a f)(\lambda) = \frac{1}{|a|} e^{-i\lambda b} (\Phi_{1/a} \mathcal{F}f)(\lambda)$, se $a \neq 0$;
- (8) $(\mathcal{F} \Phi_a T_b f)(\lambda) = \frac{1}{|a|} e^{-i\lambda(b/a)} (\Phi_{1/a} \mathcal{F}f)(\lambda)$, se $a \neq 0$.

9.3.1 Funzioni che si incontrano spesso

Sulle tavole delle trasformate di Fourier appaiono spesso alcune funzioni che hanno un comportamento semplice sotto l'azione della trasformata di Fourier e che possono essere usate come punto di partenza per calcolare la trasformata di Fourier di funzioni più complicate. Vediamo alcuni esempi (si veda anche la figura 9.2):

- (1) la gaussiana normalizzata $\gamma(x) = e^{-x^2}/\sqrt{\pi}$;
- (2) la funzione indicatrice dell'intervallo $[a, b]$, che denotiamo con $\mathbb{I}_{[a,b]}$

$$\mathbb{I}_{[a,b]}(x) = \begin{cases} 1 & \text{se } x \in [a, b] \\ 0 & \text{se } x \in \mathbb{R} \setminus [a, b]; \end{cases}$$

- (3) la funzione *rettangolo* è un caso speciale del precedente

$$\text{rect}(x) := \mathbb{I}_{[-1/2, 1/2]}$$

e soddisfa $\int_{\mathbb{R}} \text{rect}(x) dx = 1$;

- (4) la *funzione a gradino di Heaviside* è anch'essa un caso particolare della funzione indicatrice. Viene usualmente denotata con H o anche con u ("unit step" function)

$$H(x) = u(x) := \mathbb{I}_{[0, \infty)}.$$

(5) la funzione *triangolo*

$$\text{tri}(x) := \max(1 - |x|, 0) = \text{rect}(x/2)(1 - |x|)$$

che soddisfa $\int_{\mathbb{R}} \text{tri}(x) dx = 1$;

(6) la funzione sinc

$$\text{sinc}(x) := \frac{\sin(\pi x)}{\pi x},$$

anch'essa normalizzata con la condizione $\int_{\mathbb{R}} \text{sinc}(x) dx = 1$, che si può dimostrare ad esempio usando il teorema dei residui per le funzioni analitiche nel campo complesso.

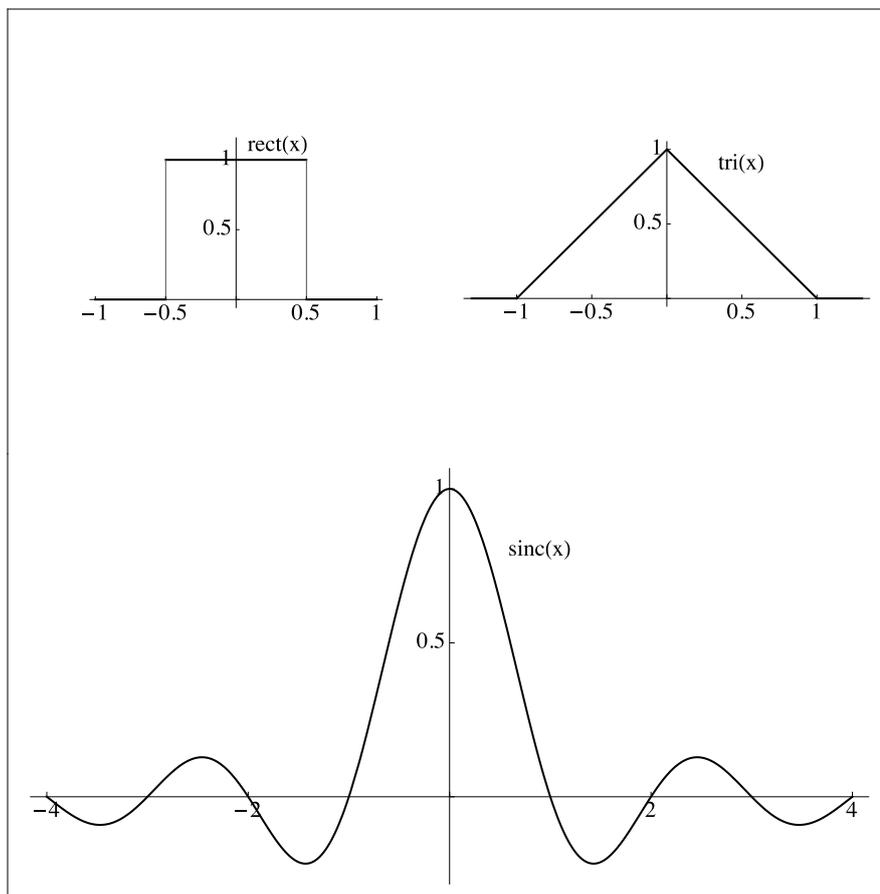


Figura 9.2: Grafici delle funzioni *rect*, *tri* e *sinc*.

Iniziamo a calcolare qualche trasformata di Fourier.

9.7 Problema. Calcolare la trasformata di Fourier di $f(x) = \mathbb{1}_{[-a,a]}$, con $a > 0$.

Soluzione. Si ha

$$\hat{f}(\lambda) = \int_{\mathbb{R}} \mathbb{1}_{[-a,a]}(x) e^{-i\lambda x} dx = \int_{-a}^a e^{-i\lambda x} dx = \left[\frac{e^{-i\lambda x}}{-i\lambda} \right]_{-a}^a = \frac{2}{\lambda} \sin(\lambda a).$$

Ponendo $a = 1/2$, otteniamo come caso particolare

$$(9.8) \quad \mathcal{F}[\text{rect}(x)](\lambda) = \frac{\sin(\lambda/2)}{\lambda/2} = \frac{\sin(\pi\lambda/(2\pi))}{\lambda/(2\pi)} = \text{sinc}(\lambda/(2\pi)).$$

9.8 Problema. Calcolare la trasformata di Fourier di $\text{tri}(x)$.

Soluzione. Si ha

$$\begin{aligned} \mathcal{F}[\text{tri}(x)](\lambda) &= \int_{\mathbb{R}} \text{tri}(x) e^{-i\lambda x} dx = \int_{-1}^1 (1 - |x|) e^{-i\lambda x} dx \\ &= \int_{-1}^0 (1 + x) e^{-i\lambda x} dx + \int_0^1 (1 - x) e^{-i\lambda x} dx. \end{aligned}$$

Con il cambio di variabile $x \rightarrow -x$ nel primo integrale si ottiene

$$\begin{aligned} \mathcal{F}[\text{tri}(x)](\lambda) &= \int_0^1 (1 - x) e^{i\lambda x} dx + \int_0^1 (1 - x) e^{-i\lambda x} dx \\ &= 2 \int_0^1 (1 - x) \cos(\lambda x) dx \\ &= 2 \underbrace{\left[(1 - x) \frac{\sin(\lambda x)}{\lambda} \right]_0^1}_{=0} + 2 \int_0^1 \frac{\sin(\lambda x)}{\lambda} dx. \end{aligned}$$

Il primo termine nel membro a destra è nullo, quindi

$$\begin{aligned} \mathcal{F}[\text{tri}(x)](\lambda) &= 2 \int_0^1 \frac{\sin(\lambda x)}{\lambda} dx = -\frac{2}{\lambda^2} [\cos(\lambda x)]_0^1 \\ &= \frac{2(1 - \cos(\lambda))}{\lambda^2} = \frac{4 \sin^2(\lambda/2)}{\lambda^2} = \left(\frac{\sin(\lambda/2)}{\lambda/2} \right)^2 = \text{sinc}^2(\lambda/(2\pi)). \end{aligned}$$

Vedremo in seguito che il fatto che la trasformata di Fourier della funzione triangolo sia il quadrato della trasformata di Fourier della funzione rettangolo non è un caso, ma è legato ad una proprietà che riguarda il cosiddetto *prodotto di convoluzione*.

9.9 Problema. Calcolare la trasformata di Fourier di $f(x) := H(x) e^{-ax}$, con $a > 0$.

Soluzione. Si ottiene

$$\hat{f}(\lambda) = \int_{\mathbb{R}} H(x) e^{-ax} e^{-i\lambda x} dx = \int_0^{+\infty} e^{-(a+i\lambda)x} dx = \left[\frac{e^{-(a+i\lambda)x}}{-(a+i\lambda)} \right]_0^{+\infty} = \frac{1}{a+i\lambda}.$$

Nell'ultima uguaglianza abbiamo usato il fatto che

$$\lim_{x \rightarrow +\infty} |e^{-(a+i\lambda)x}| = \lim_{x \rightarrow +\infty} e^{-ax} = 0,$$

e, di conseguenza

$$\lim_{x \rightarrow +\infty} e^{-(a+i\lambda)x} = 0.$$

9.10 Problema. Calcolare la trasformata di Fourier di $f(x) := 1/(1+x^2)$.

Soluzione. Dobbiamo calcolare

$$\hat{f}(\lambda) := \int_{\mathbb{R}} \frac{e^{-i\lambda x}}{1+x^2} dx.$$

Per calcolare questo integrale usiamo il teorema dei residui. Se $\lambda > 0$ il cammino di integrazione può essere chiuso nel semipiano inferiore del piano complesso. Se indichiamo con

$$g(z) := \frac{e^{-i\lambda z}}{1+z^2} \quad z \in \mathbb{C}$$

otteniamo

$$\hat{f}(\lambda) = -2\pi i \sum_{\text{Im } z < 0} \text{Res}(g, z).$$

Il segno meno deriva dal fatto che, chiudendo il cammino nel semipiano inferiore, si produce un cammino chiuso che gira in senso orario. L'unica singolarità di g che si trova nel semipiano inferiore è $z = -i$ e si ha

$$\text{Res}(g, -i) = \left[\frac{e^{-i\lambda z}}{z-i} \right]_{z=-i} = \frac{e^{-\lambda}}{-2i}.$$

Concludendo

$$(9.9) \quad \hat{f}(\lambda) = \pi e^{-\lambda} \quad \lambda > 0.$$

Il caso $\lambda < 0$ si risolve in modo analogo e si ottiene

$$(9.10) \quad \hat{f}(\lambda) = \pi e^{\lambda} \quad \lambda < 0.$$

Alternativamente si poteva osservare che f è reale pari, quindi, per la Proposizione 9.4, \hat{f} è anche reale pari. Questo fatto, insieme alla (9.9), ci dice che l'espressione di \hat{f} valida su tutto l'asse reale si ottiene, sostituendo nella (9.9) λ con $|\lambda|$, ottenendo

$$\hat{f}(\lambda) = \pi e^{-|\lambda|} \quad \lambda \in \mathbb{R},$$

che infatti sintetizza le (9.9) e (9.10).

9.11 Problema. Verificare che, se $a > 0$, si ha

$$\mathcal{F}[e^{-a|x|}](\lambda) = \frac{2a}{a^2 + \lambda^2}.$$

Suggerimento: si può semplicemente usare la definizione di trasformata di Fourier, oppure osservare che

$$e^{-a|x|} = H(x)e^{-ax} + H(-x)e^{ax} = H(x)e^{-ax} + P(H(x)e^{-ax}),$$

in cui P è l'operatore di parità. In questo modo ci si riconduce al Problema 9.9.

9.12 Problema. Verificare che

$$\mathcal{F}[e^{-x^2}](\lambda) = \sqrt{\pi} e^{-\lambda^2/4}.$$

Suggerimento: integrare la funzione analitica $f(z) = e^{-z^2}$ lungo il rettangolo $[-a, a, a+ib, -a-ib, a]$. Scegliere opportunamente b e fare il limite $a \rightarrow \infty$.

9.3.2 Il metodo grafico

Le regole di calcolo della trasformata di Fourier, elencate nella Proposizione 9.5, possono essere rappresentate graficamente. Prendiamo ad esempio la (1) della Proposizione 9.5: $F[f(-x)](\lambda) = \hat{f}(-\lambda)$. A questa identità associamo il diagramma

$$[x \rightarrow -x] \quad f(-x) \quad \xrightarrow{\mathcal{F}} \quad \hat{f}(-\lambda).$$

L'identità (2) si rappresenta come

$$[x \rightarrow x - a] \quad f(x - a) \quad \xrightarrow{\mathcal{F}} \quad e^{-i\lambda a} \hat{f}(\lambda).$$

Come ultimo esempio, l'identità (6) si rappresenta come

$$[x \rightarrow ax] \quad f(ax) \quad \xrightarrow{\mathcal{F}} \quad \frac{1}{|a|} \hat{f}(\lambda/a).$$

L'insieme delle regole di calcolo sono raccolte nella tabella 9.3, mentre la tabella 9.4 contiene la trasformata di Fourier di alcune funzioni elementari.

Vediamo ora come usare questa rappresentazione per calcolare la trasformata di Fourier di una funzione f sfruttando, quando è possibile, la conoscenza della trasformata di Fourier di un'altra funzione legata a f da opportune trasformazioni.

9.13 Problema. Calcolare la trasformata di Fourier delle seguenti funzioni, riconducendosi, se possibile, a casi noti

$$(a) \quad f(x) = \cos x e^{-4x^2+4x} \qquad (b) \quad f(x) = H(3-x) e^{2x-4}.$$

Soluzione. (a) Partiamo dalla gaussiana e operiamo opportune trasformazioni per arrivare alla funzione f

$$\begin{array}{lll} & e^{-x^2} & \xrightarrow{\mathcal{F}} \quad \sqrt{\pi} e^{-\lambda^2/4} \\ [x \rightarrow 2x] & e^{-4x^2} & \xrightarrow{\mathcal{F}} \quad \frac{1}{2} \sqrt{\pi} e^{-\lambda^2/16} \\ \left[x \rightarrow x - \frac{1}{2} \right] & e^{-4(x-1/2)^2} = e^{-4x^2+4x-1} & \xrightarrow{\mathcal{F}} \quad \frac{1}{2} \sqrt{\pi} e^{-\lambda^2/16-i\lambda/2} \\ [f \rightarrow e f] & e^{-4x^2+4x} & \xrightarrow{\mathcal{F}} \quad \frac{e}{2} \sqrt{\pi} e^{-\lambda^2/16-i\lambda/2} \\ [f \rightarrow \cos x f] & \cos x e^{-4x^2+4x} & \xrightarrow{\mathcal{F}} \quad \frac{e\sqrt{\pi}}{4} \left[e^{-(\lambda-1)^2/16-i(\lambda-1)/2} + e^{-(\lambda+1)^2/16-i(\lambda+1)/2} \right] \end{array}$$

L'espressione finale può essere leggermente semplificata:

$$\hat{f}(\lambda) = \frac{e\sqrt{\pi}}{4} e^{-(\lambda^2+1)/16-i\lambda/2} \left[e^{\lambda/8+i/2} + e^{-\lambda/8-i/2} \right].$$

(b) Partendo dalla trasformata di Fourier di $H(x)e^{-ax}$, calcolata nel Problema 9.9, si ottiene

$$\begin{array}{llll}
 & H(x) e^{-ax} & \xrightarrow{\mathcal{F}} & \frac{1}{a + i\lambda} \\
 [x \rightarrow -x] & H(-x) e^{ax} & \xrightarrow{\mathcal{F}} & \frac{1}{a - i\lambda} \\
 [x \rightarrow x - 3] & H(3 - x) e^{ax-3a} & \xrightarrow{\mathcal{F}} & \frac{e^{-i3\lambda}}{a - i\lambda} \\
 [a = 2] & H(3 - x) e^{2x-6} & \xrightarrow{\mathcal{F}} & \frac{e^{-i3\lambda}}{2 - i\lambda} \\
 [f \rightarrow e^2 f] & H(3 - x) e^{2x-4} & \xrightarrow{\mathcal{F}} & \frac{e^{-i3\lambda+2}}{2 - i\lambda}.
 \end{array}$$

$[x \rightarrow -x]$	$f(ax)$	$\xrightarrow{\mathcal{F}}$	$\frac{1}{ a } \hat{f}(\lambda/a)$
$[x \rightarrow x - a]$	$f(x - a)$	$\xrightarrow{\mathcal{F}}$	$e^{-i\lambda a} \hat{f}(\lambda)$
$[f \rightarrow e^{iax} f]$	$e^{iax} f(x)$	$\xrightarrow{\mathcal{F}}$	$\hat{f}(\lambda - a)$
$[f \rightarrow \cos(ax) f]$	$\cos(ax) f(x)$	$\xrightarrow{\mathcal{F}}$	$\frac{1}{2} (f(\lambda - a) + f(\lambda + a))$
$[f \rightarrow \sin(ax) f]$	$\sin(ax) f(x)$	$\xrightarrow{\mathcal{F}}$	$\frac{1}{2i} (f(\lambda - a) - f(\lambda + a))$
$[x \rightarrow ax]$	$f(ax)$	$\xrightarrow{\mathcal{F}}$	$\frac{1}{ a } \hat{f}(\lambda/a)$
$[f \rightarrow f']$	$f'(x)$	$\xrightarrow{\mathcal{F}}$	$i\lambda \hat{f}(\lambda)$
$[f \rightarrow x f]$	$x f(x)$	$\xrightarrow{\mathcal{F}}$	$iD \hat{f}(\lambda)$
$[f \rightarrow \hat{f}]$	$\hat{f}(x)$	$\xrightarrow{\mathcal{F}}$	$2\pi f(-\lambda)$

Figura 9.3: Regole di calcolo delle trasformate di Fourier

f	\hat{f}	Note
$\text{rect}(x)$	$\text{sinc}(\lambda/(2\pi)) = \frac{2 \sin(\lambda/2)}{\lambda}$	
$\text{tri}(x)$	$\text{sinc}^2(\lambda/(2\pi))$	
$\text{sinc}(x)$	$\text{rect}(x/2\pi)$	
$\mathbb{I}_{[-a,a]}$	$\frac{2 \sin(\lambda a)}{\lambda}$	
$\cos x \mathbb{I}_{[-a,a]}$	$\frac{\sin[(\lambda+1)a]}{\lambda+1} + \frac{\sin[(\lambda-1)a]}{\lambda-1}$	
$\cos x \mathbb{I}_{[-\pi/2,\pi/2]}$	$-\frac{2}{\lambda^2-1} \cos\left(\frac{\lambda\pi}{2}\right)$	
$e^{-a x }$	$\frac{2a}{\lambda^2+a^2}$	
$e^{-ax} H(x)$	$\frac{1}{a+i\lambda}$	$a > 0$
$e^{-ax} H(ax)$	$\frac{\text{sgn } a}{a+i\lambda}$	
$\frac{1}{x^2+a^2}$	$\frac{\pi}{a} e^{-a \lambda }$	
$\frac{1}{ax^2+bc+c}$	$\frac{\pi}{\beta a} e^{-\beta \lambda } e^{i\lambda b/2a}$	$\beta = \frac{\sqrt{D}}{2 a }, \underbrace{4ac - b^2}_D > 0$
$\frac{1}{(1+x^2)^2}$	$\frac{\pi}{2} e^{- \lambda } (\lambda + 1)$	
$\frac{1}{1+x^4}$	$\pi e^{- \lambda /\sqrt{2}} \sin\left[\frac{ \lambda }{\sqrt{2}} + \frac{\pi}{4}\right]$	
e^{-ax^2}	$\sqrt{\frac{\pi}{a}} e^{-\lambda^2/(4a)}$	
e^{ix^2}	$\sqrt{\pi} \frac{1+i}{\sqrt{2}} e^{-i\lambda^2/4}$	
e^{-ix^2}	$\sqrt{\pi} \frac{1-i}{\sqrt{2}} e^{i\lambda^2/4}$	
$\cos(x^2)$	$+\sqrt{\pi} \cos[(\lambda^2 - \pi)/4]$	
$\sin(x^2)$	$-\sqrt{\pi} \sin[(\lambda^2 - \pi)/4]$	
$\text{sech}(x)$	$\pi \text{sech}(\lambda\pi/2)$	
$\psi_n(x) = \frac{e^{-x^2/2} H_n(x)}{\sqrt{\sqrt{\pi} 2^n n!}}$	$\sqrt{2\pi} (-i)^n \psi_n(\lambda)$	$\psi_n = \text{Hermite func}$

Figura 9.4: Alcune trasformate di Fourier

9.4 Regolarità e andamento all'infinito

9.14 Proposizione. Se $f \in L_1(\mathbb{R})$ allora $\mathcal{F}(f) \in C_0(\mathbb{R})$.

Dimostrazione.

❶ Dimostrazione della continuità. Poniamo $g(\lambda) := \mathcal{F}(f)(\lambda)$. Allora

$$\begin{aligned} |g(\lambda + \varepsilon) - g(\lambda)| &= \left| \int_{\mathbb{R}} f(x) e^{-i\lambda x} (e^{-i\varepsilon x} - 1) dx \right| \\ &\leq \int_{\mathbb{R}} |f(x)| |e^{-i\varepsilon x} - 1| dx \\ &\leq \int_{|x| \leq a} |f(x)| |e^{-i\varepsilon x} - 1| dx + 2 \int_{|x| > a} |f(x)| dx \end{aligned}$$

Siccome

$$\begin{aligned} (9.11) \quad |e^{-i\varepsilon x} - 1|^2 &= (1 - \cos(\varepsilon x))^2 + \sin(\varepsilon x)^2 \\ &= 2(1 - \cos(\varepsilon x)) = 4 \sin^2(\varepsilon x/2) \leq (\varepsilon x)^2, \end{aligned}$$

nell'intervallo $|x| \leq a$ vale la disuguaglianza

$$|e^{-i\varepsilon x} - 1| \leq |\varepsilon x| \leq |\varepsilon| a$$

Otteniamo dunque

$$\begin{aligned} |g(\lambda + \varepsilon) - g(\lambda)| &\leq |\varepsilon| a \int_{|x| \leq a} |f(x)| dx + 2 \int_{|x| > a} |f(x)| dx \\ &\leq |\varepsilon| a \|f\|_1 + 2 \int_{|x| > a} |f(x)| dx. \end{aligned}$$

Poiché a è arbitrario possiamo porre $a := 1/\sqrt{|\varepsilon|}$. In questo modo otteniamo

$$(9.12) \quad |g(\lambda + \varepsilon) - g(\lambda)| \leq \sqrt{|\varepsilon|} \|f\|_1 + 2 \int_{|x| > 1/\sqrt{|\varepsilon|}} |f(x)| dx.$$

Siccome

$$(9.13) \quad \lim_{K \rightarrow \infty} \int_{-K}^K |f(x)| dx = \|f\|_1.$$

si ha

$$\lim_{K \rightarrow \infty} \int_{|x| > K} |f(x)| dx = 0.$$

Di conseguenza l'integrale che appare al secondo membro della (9.12) tende a zero, quando $\varepsilon \rightarrow 0$. Abbiamo quindi ottenuto che

$$\lim_{\varepsilon \rightarrow 0} |g(\lambda + \varepsilon) - g(\lambda)| = 0,$$

che equivale a dire che g è continua nel punto (arbitrario) λ .

❷ Dimostrazione che $\lim_{\lambda \rightarrow \infty} g(\lambda) = 0$. Analogamente a quanto fatto in precedenza scriviamo l'integrale come somma di 2 pezzi.

$$\begin{aligned} |g(\lambda)| &\leq \left| \int_{|x| \leq a} f(x) e^{-i\lambda x} dx \right| + \left| \int_{|x| > a} f(x) e^{-i\lambda x} dx \right| \\ &\leq \left| \int_{|x| \leq a} f(x) e^{-i\lambda x} dx \right| + \int_{|x| > a} |f(x)| dx \end{aligned}$$

Poiché il secondo integrale tende a zero quando $a \rightarrow \infty$, dato $\varepsilon > 0$ arbitrario, possiamo scegliere a in modo tale che questo integrale sia minore od uguale a ε . Quindi

$$|g(\lambda)| \leq \left| \int_{|x| \leq a} f(x) e^{-i\lambda x} dx \right| + \varepsilon.$$

Grazie al lemma di Riemann–Lebesgue, l'integrale va a zero quando $\lambda \rightarrow \infty$, quindi

$$\lim_{\lambda \rightarrow \infty} |g(\lambda)| \leq \varepsilon.$$

Siccome ε è arbitrario ottengo che $\lim_{\lambda \rightarrow \infty} g(\lambda) = 0$. \square

9.15 Proposizione. *Sia $f \in C^{p-1}(\mathbb{R})$ con $f^{(p)}$ continua a tratti e assumiamo che le funzioni $f, f', f'', \dots, f^{(p)}$ appartengano tutte a $L_1(\mathbb{R})$. Allora*

(1) *per ogni $k = 1, \dots, p$ si ha $\mathcal{F}(f^{(k)})(\lambda) = (i\lambda)^k \mathcal{F}(f)$.*

(2) *$\mathcal{F}(f)(\lambda)$ tende a zero, quando $\lambda \rightarrow \infty$ più velocemente di $1/\lambda^p$, vale a dire*

$$\lim_{\lambda \rightarrow \infty} \lambda^p \mathcal{F}(f)(\lambda) = 0.$$

Dimostrazione. Integrando per parti si ottiene

$$\begin{aligned} \mathcal{F}[f'](\lambda) &= \int_{\mathbb{R}} f'(x) e^{-i\lambda x} dx \\ (9.14) \quad &= f(x) e^{-i\lambda x} \Big|_{-\infty}^{\infty} + i\lambda \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx \\ &= [f(x) e^{-i\lambda x}]_{-\infty}^{+\infty} + i\lambda \mathcal{F}[f](\lambda). \end{aligned}$$

Nel caso in cui $\lim_{\pm\infty} f(x) = 0$ otteniamo $c\mathcal{F}[f'](\lambda) = i\lambda \mathcal{F}[f](\lambda)$ che, iterata, dimostra il punto (1). Sappiamo però (vedi l'esempio 3.46) dall'integrabilità assoluta di f non discende necessariamente che f tende a zero all'infinito. In questo caso però si può usare l'ipotesi che $f' \in L_1(\mathbb{R})$. Possiamo scrivere infatti

$$f(x) = f(0) + \int_0^x f'(u) du.$$

Poiché $f' \in L_1(\mathbb{R})$ esistono i limiti

$$\lim_{x \rightarrow \pm\infty} \int_0^x f'(u) du,$$

di conseguenza esistono i limiti $\lim_{x \rightarrow \pm\infty} f(x)$. Ma se f ha un limite quando x tende a $-\infty$ e a $+\infty$, il valore di entrambi i limiti deve essere necessariamente zero, altrimenti f non sarebbe integrabile. Abbiamo quindi dimostrato che $[f(x) e^{-i\lambda x}]_{-\infty}^{+\infty} = 0$. Dalla (9.14) segue che

$$\mathcal{F}[f'](\lambda) = i\lambda \mathcal{F}[f](\lambda),$$

che, iterata, dimostra il punto (1) della proposizione. Per dimostrare il punto (2) usiamo il punto (1) con $k = p$ e la Proposizione 9.14, ottenendo

$$\lim_{\lambda \rightarrow \infty} |\lambda^p \mathcal{F}[f](\lambda)| = \lim_{\lambda \rightarrow \infty} |\mathcal{F}[f^{(p)}](\lambda)| = 0. \quad \square$$

9.16 Proposizione. *Sia $f \in L_1(\mathbb{R})$ tale che le funzioni $xf, x^2f, \dots, x^p f$ appartengano tutte ad $L_1(\mathbb{R})$. Allora $\mathcal{F}(f) \in C^p(\mathbb{R})$. Inoltre, ponendo $g = \mathcal{F}(f)$, si ha*

$$g^{(k)} = \mathcal{F}[(-ix)^k f] \quad k = 1, \dots, p.$$

Dimostrazione. Sia

$$g(\lambda) := \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx.$$

Voglio far vedere che $g'(\lambda) = \mathcal{F}[(-ix)f](\lambda)$, vale a dire che

$$(9.15) \quad \lim_{\varepsilon \rightarrow 0} \frac{g(\lambda + \varepsilon) - g(\lambda)}{\varepsilon} = \int_{\mathbb{R}} (-ix) f(x) e^{-i\lambda x} dx.$$

Per far vedere che vale la (9.15) procedo nel modo seguente

$$(9.16) \quad \begin{aligned} & \left| \frac{g(\lambda + \varepsilon) - g(\lambda)}{\varepsilon} - \int_{\mathbb{R}} (-ix) f(x) e^{-i\lambda x} dx \right| \\ &= \left| \int_{\mathbb{R}} f(x) e^{-i\lambda x} \left[\frac{e^{-i\varepsilon x} - 1}{\varepsilon} + ix \right] dx \right| \\ &\leq \int_{\mathbb{R}} |f(x)| \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon} + ix \right| dx \\ &= \int_{|x| \leq a} |x f(x)| \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon x} + i \right| dx \\ &\quad + \int_{|x| > a} |x f(x)| \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon x} + i \right| dx \end{aligned}$$

Il termine $|(e^{-i\varepsilon x} - 1)/\varepsilon x + i|$ può essere stimato in due modi diversi. Un primo modo è il seguente: grazie alla (9.11) si ottiene

$$(9.17) \quad \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon x} + i \right| \leq \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon x} \right| + |i| \leq 2.$$

Alternativamente, dallo sviluppo di Taylor del seno e del coseno, si ottiene

$$|\cos \vartheta - 1| \leq \frac{1}{2} \vartheta^2 \quad |\sin \vartheta - \vartheta| \leq \frac{1}{3!} |\vartheta|^3.$$

Di conseguenza abbiamo

$$(9.18) \quad \begin{aligned} \left| \frac{e^{-i\varepsilon x} - 1}{\varepsilon x} + i \right| &= \left| \frac{\cos(\varepsilon x) - 1}{\varepsilon x} + i \frac{\varepsilon x - \sin(\varepsilon x)}{\varepsilon x} \right| \\ &\leq \frac{1 - \cos(\varepsilon x)}{|\varepsilon x|} + \frac{|\varepsilon x - \sin(\varepsilon x)|}{|\varepsilon x|} \\ &\leq \frac{1}{2} |\varepsilon x| + \frac{1}{6} |\varepsilon x|^2. \end{aligned}$$

Usando sia la (9.17) che la (9.18) si ottiene dalla (9.16)

$$(9.19) \quad \begin{aligned} & \left| \frac{g(\lambda + \varepsilon) - g(\lambda)}{\varepsilon} - \int_{\mathbb{R}} (-ix) f(x) e^{-i\lambda x} dx \right| \\ &\leq \int_{|x| \leq a} |x f(x)| \left(\frac{|\varepsilon x|}{2} + \frac{|\varepsilon x|^2}{6} \right) dx + 2 \int_{|x| > a} |x f(x)| dx \\ &\leq \left(\frac{|\varepsilon a|}{2} + \frac{|\varepsilon a|^2}{6} \right) \int_{|x| \leq a} |x f(x)| dx + 2 \int_{|x| > a} |x f(x)| dx \\ &\leq \left(\frac{|\varepsilon a|}{2} + \frac{|\varepsilon a|^2}{6} \right) \|x f\|_1 + 2 \int_{|x| > a} |x f(x)| dx \end{aligned}$$

Scegliendo ora $a := 1/\sqrt{|\varepsilon|}$ e facendo il limite $\varepsilon \rightarrow 0$ si ottiene

$$(9.20) \quad \lim_{\varepsilon \rightarrow 0} \left| \frac{g(\lambda + \varepsilon) - g(\lambda)}{\varepsilon} - \int_{\mathbb{R}} (-ix) f(x) e^{-i\lambda x} dx \right| \\ \leq \lim_{\varepsilon \rightarrow 0} \left(\frac{\sqrt{|\varepsilon|}}{2} + \frac{|\varepsilon|}{6} \right) \|x f\|_1 + 2 \int_{|x| > 1/\sqrt{|\varepsilon|}} |x f(x)| dx = 0.$$

Abbiamo quindi fatto vedere che vale la (9.15) che, iterata, dimostra la proposizione. \square

9.4.1 Un importante sottospazio invariante

Le Proposizioni 9.15 e 9.16 ci dicono che, qualitativamente, tanto più f tende velocemente a zero all'infinito, tanto più la sua trasformata di Fourier è differenziabile e viceversa. Questo ci consente di riconoscere un importante sottospazio di $L_1(\mathbb{R})$ che è *invariante* sotto l'azione della trasformata di Fourier. Si tratta dello spazio di Schwartz, introdotto nella sezione 3.6.4.

9.17 Proposizione. *Lo spazio $\mathcal{S}(\mathbb{R})$ è invariante sotto l'azione della trasformata di Fourier \mathcal{F} , vale a dire se $f \in \mathcal{S}(\mathbb{R})$ allora $\mathcal{F}[f] \in \mathcal{S}(\mathbb{R})$.*

Dimostrazione. Ricordiamo che, per definizione, $f \in \mathcal{S}(\mathbb{R})$ se $f \in C^\infty(\mathbb{R})$ e, per ogni coppia di interi non negativi j e k , si ha

$$(9.21) \quad \|f\|_{j,k} := \|x^j D^k f\|_\infty = \sup_{x \in \mathbb{R}} |x^j D^k f(x)| < \infty.$$

Per la Proposizione 3.50, questa condizione è equivalente a richiedere che per ogni $k \in \mathbb{N}$ e per ogni $a > 0$ si ha

$$(9.22) \quad |D^k f(x)| \leq \frac{C_{a,k}}{(1+x^2)^{a/2}} \quad \forall x \in \mathbb{R}.$$

Supponiamo dunque che $f \in \mathcal{S}(\mathbb{R})$ e siano $j, k \in \mathbb{N}$. Devo far vedere che la (9.21) è soddisfatta. Usando la (9.22) con $a = j + 2$ ottengo

$$\int_{\mathbb{R}} |x^j D^k f(x)| dx \leq \int_{\mathbb{R}} \frac{|x^j| C_{j+2,k}}{(1+x^2)^{j/2+1}} \leq \int_{\mathbb{R}} \frac{C_{j+2,k}}{1+x^2} < \infty.$$

Dunque $x^j D^k f \in L_1(\mathbb{R})$ per ogni $j, k \in \mathbb{N}$. Dalla regola di Leibniz della Proposizione 1.12 si ottiene

$$D^j(x^k f) = \sum_{m=0}^j \binom{j}{m} D^m(x^k) D^{j-m} f$$

quindi $D^j(x^k f)$ è una combinazione lineare di funzioni del tipo $x^r D^s f$. Visto che queste ultime appartengono a $L_1(\mathbb{R})$, otteniamo

$$x^j D^k f \in L_1(\mathbb{R}) \quad \text{e} \quad D^j(x^k f) \in L_1(\mathbb{R}) \quad \forall j, k \in \mathbb{N}.$$

Questo, unitamente al fatto che f è infinitamente derivabile, ci permette di usare le Proposizioni 9.15 e 9.16 ottenendo

$$(i\lambda)^j (D^k \hat{f})(\lambda) = (i\lambda)^j (-i)^k \mathcal{F}[x^k f](\lambda) \\ = (-i)^k (i\lambda)^j \mathcal{F}[x^k f](\lambda) = (-i)^k \mathcal{F}[D^j(x^k f)](\lambda).$$

La Proposizione 9.14 ci dice che la funzione $\mathcal{F}[D^j(x^k f)]$ appartiene a $C_0(\mathbb{R})$, quindi è limitata. Possiamo dunque concludere che

$$\|\hat{f}\|_{j,k} := \sup_{\lambda \in \mathbb{R}} |\lambda^j (D^k \hat{f})(\lambda)| = \sup_{\lambda \in \mathbb{R}} |\mathcal{F}[D^j(x^k f)](\lambda)| < \infty.$$

Siccome j e k sono arbitrari, possiamo affermare che $\hat{f} \in \mathcal{S}(\mathbb{R})$. \square

9.4.2 Problemi

Le Proposizioni 9.15 e 9.16 forniscono altre 2 regole di calcolo della trasformata di Fourier che possono essere scritte come:

$$\begin{aligned}\mathcal{F}(D^k f)(\lambda) &= (i\lambda)^k \mathcal{F}(f) \\ \mathcal{F}(x^k f)(\lambda) &= i^k (D^k \hat{f})(\lambda).\end{aligned}$$

Nella notazione del “metodo grafico” della sezione 9.3.2 queste regole verranno indicate, per $k = 1$, rispettivamente come

$$\begin{aligned}[f \rightarrow f'] & \quad f'(x) & \xrightarrow{\mathcal{F}} & \quad i\lambda \hat{f}(\lambda) \\ [f \rightarrow x f] & \quad x f(x) & \xrightarrow{\mathcal{F}} & \quad i D \hat{f}(\lambda).\end{aligned}$$

9.18 Problema. Calcolare la trasformata di Fourier di

$$f(x) = x e^{-x^2 + 5ix}.$$

Soluzione. Partendo da una coppia nota f, \hat{f} ed usando le regole di calcolo delle trasformate di Fourier, ottengo

$$\begin{aligned}e^{-x^2} & \xrightarrow{\mathcal{F}} \sqrt{\pi} e^{-s^2/4} \\ [f \rightarrow x f] & \quad x e^{-x^2} \xrightarrow{\mathcal{F}} i D \left[\sqrt{\pi} e^{-s^2/4} \right] = -\frac{i\sqrt{\pi}}{2} s e^{-s^2/4} \\ [f \rightarrow e^{5ix} f] & \quad x e^{-x^2 + 5ix} \xrightarrow{\mathcal{F}} -\frac{i\sqrt{\pi}}{2} (s - 5) e^{-(s-5)^2/4}.\end{aligned}$$

9.19 Problema. Calcolare la trasformata di Fourier di f , riconducendosi, se possibile, a casi noti

$$f(x) = \frac{x^2}{(1 + 4x^2)^2}.$$

Soluzione. Partendo da una coppia nota f, \hat{f} ed usando le regole di calcolo delle trasformate di Fourier, ottengo

$$\begin{aligned}[x \rightarrow 2x] & \quad \frac{1}{1 + x^2} \xrightarrow{\mathcal{F}} \pi e^{-|\lambda|} \\ [f \rightarrow f'] & \quad -\frac{8x}{(1 + 4x^2)^2} \xrightarrow{\mathcal{F}} \frac{i\pi\lambda}{2} e^{-|\lambda|/2} \\ [f \rightarrow -\frac{x}{8} f] & \quad \frac{x^2}{(1 + 4x^2)^2} \xrightarrow{\mathcal{F}} -\frac{1}{8} i D \left(\frac{i\pi\lambda}{2} e^{-|\lambda|/2} \right) \\ & \quad \frac{x^2}{(1 + 4x^2)^2} \xrightarrow{\mathcal{F}} \frac{\pi}{16} e^{-|\lambda|/2} \left(1 - \frac{|\lambda|}{2} \right).\end{aligned}$$

Nell'ultimo passaggio ho usato l'identità: $\lambda \operatorname{sgn}(\lambda) = |\lambda|$.

9.20 Problema. Sia \mathcal{F} l'operatore lineare che opera la trasformata di Fourier

$$\mathcal{F}(f)(\lambda) := \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx$$

Considerando \mathcal{F} come un operatore dallo spazio $(L_1(\mathbb{R}), \|\cdot\|_1)$ nello spazio $(C_0(\mathbb{R}), \|\cdot\|_\infty)$, dimostrare che $\|\mathcal{F}\| = 1$.

Soluzione. Se $f \in L_1(\mathbb{R})$ si ha

$$|\mathcal{F}(f)(\lambda)| \leq \int_{\mathbb{R}} |f(x) e^{-i\lambda x}| dx = \int_{\mathbb{R}} |f(x)| dx = \|f\|_1.$$

Questa disuguaglianza è valida per ogni valore di λ reale, quindi abbiamo $\|\mathcal{F}(f)\|_\infty \leq \|f\|_1$, il che implica

$$\|\mathcal{F}\| \leq 1.$$

Per dimostrare che vale l'uguaglianza si scelga una funzione $h \in L_1(\mathbb{R})$ che sia non negativa e che abbia un massimo in $x = 0$, ad esempio $h(x) = (1 + x^2)^{-1}$. Allora

$$\|\mathcal{F}(h)\|_\infty = \sup_{\lambda \in \mathbb{R}} |\mathcal{F}(h)(\lambda)| \geq |\mathcal{F}(h)(0)| = \left| \int_{\mathbb{R}} h(x) dx \right| = \int_{\mathbb{R}} |h(x)| dx = \|h\|_1. \quad \square$$

9.21 Problema. Calcolare la trasformata di Fourier di

$$f(x) := \frac{e^{-|x|} \sin(ax)}{x} \quad a > 0.$$

Soluzione. Pongo

$$g(x) := e^{-|x|} \quad h(x) := e^{-|x|} \sin(ax).$$

Allora, poiché $\hat{g}(\lambda) = 2/(1 + \lambda^2)$, si ha

$$\hat{h}(\lambda) = \frac{1}{2i} [\hat{g}(\lambda - a) - \hat{g}(\lambda + a)] = \frac{1}{i} \left[\frac{1}{(\lambda - a)^2 + 1} - \frac{1}{(\lambda + a)^2 + 1} \right].$$

D'altra parte $h(x) = xf(x)$, quindi

$$\hat{h}(\lambda) = iD\hat{f}(\lambda).$$

Ottengo dunque

$$D\hat{f}(\lambda) = \frac{1}{(\lambda + a)^2 + 1} - \frac{1}{(\lambda - a)^2 + 1},$$

e, integrando, ottengo

$$\hat{f}(\lambda) = \arctan(\lambda + a) - \arctan(\lambda - a) + c \quad c \in \mathbb{R}.$$

So che la trasformata di Fourier di una funzione $f \in L_1(\mathbb{R})$ tende a zero all'infinito. Questo mi permette di determinare la costante di integrazione come $c = 0$, quindi

$$\hat{f}(\lambda) = \arctan(\lambda + a) - \arctan(\lambda - a).$$

9.22 Problema. Sia \hat{f} la trasformata di Fourier della funzione f . Cosa si può affermare, nei casi seguenti, sul grado di derivabilità di \hat{f} senza doverla calcolare?

$$(a) f(x) = x^7 e^{-\sqrt{|x|}} \quad (b) f(x) = \frac{\sin x}{1+x^2} \quad (c) f(x) = \frac{1}{1+|x|^3}$$

Soluzione.

(a) Poiché $x^n f \in L_1(\mathbb{R})$ per ogni $n \in \mathbb{N}$, \hat{f} è infinitamente differenziabile.

(b) In questo caso $f \in L_1(\mathbb{R})$ ma $xf \notin L_1(\mathbb{R})$, quindi non possiamo affermare nulla sulla derivabilità di \hat{f} .

(c) Si vede che le funzioni f e xf appartengono a $L_1(\mathbb{R})$ ma $x^2 f \notin L_1(\mathbb{R})$, quindi \hat{f} è di classe C^1 .

9.23 Problema. Sia \hat{f} la trasformata di Fourier della funzione f . Senza dover calcolare \hat{f} , è possibile affermare che \hat{f} è di classe C^p e che $\lim_{\lambda \rightarrow \infty} \lambda^q \hat{f}(\lambda) = 0$, in cui p e q valgono ...

$$(a) f(x) = \frac{1}{1+x^6} \quad (b) f(x) = \begin{cases} \sin^2(x) & \text{se } x \in [0, 2\pi] \\ 0 & \text{altrimenti} \end{cases}$$

Soluzione.

(a) Poiché $x^k f \in L_1$, per $k = 0, 1, 2, 3, 4$ $\hat{f} \in C^4$.

Poiché $f \in C^\infty$ e $f^{(k)} \in L_1$ per tutti i k interi non negativi si ha che $\lim_{\lambda \rightarrow \infty} \lambda^k \hat{f}(\lambda) = 0$ per ogni $k \in \mathbb{N}$, vale a dire \hat{f} va a zero all'infinito più velocemente di qualsiasi potenza.

(b) La funzione f è continua a supporto compatto, quindi $x^k f \in L_1$ per ogni $k \in \mathbb{N}$. Di conseguenza $\hat{f} \in C^\infty$.

Voglio trovare il massimo valore di p tale che $f \in C^{p-1}$ con $f^{(p)}$ continua a tratti. Gli unici punti in cui f o le sue derivate possono essere discontinue sono i due punti di "racordo", $x = 0$ e $x = 2\pi$. Poiché f è simmetrica rispetto al punto $x = \pi$ è sufficiente limitarsi a studiare la continuità delle derivate di f in $x = 0$. Sia $h(x) := \sin^2(x)$. Allora

$$\begin{aligned} h'(x) &= 2 \sin(x) \cos(x) = \sin(2x) \\ h''(x) &= 2 \cos(2x) \end{aligned}$$

Per cui otteniamo:

$$\begin{array}{ll} \lim_{x \rightarrow 0^-} f(x) = 0 & \lim_{x \rightarrow 0^+} f(x) = \sin^2(0) = 0 \\ \lim_{x \rightarrow 0^-} f'(x) = 0 & \lim_{x \rightarrow 0^+} f'(x) = \sin(0) = 0 \\ \lim_{x \rightarrow 0^-} f''(x) = 0 & \lim_{x \rightarrow 0^+} f''(x) = 2 \cos(0) = 2. \end{array}$$

Quindi la funzione f è di classe C^1 con f'' continua a tratti. Di conseguenza \hat{f} tende a zero quando $\lambda \rightarrow \infty$ più velocemente di $1/\lambda^2$.

9.24 Problema. Sia \hat{f} la trasformata di Fourier della funzione f . Senza dover calcolare \hat{f} , è possibile affermare che \hat{f} è di classe C^p e che $\lim_{\lambda \rightarrow \infty} \lambda^q \hat{f}(\lambda) = 0$, in cui p e q valgono ...

$$f(x) = x^n e^{-|x|} \quad \text{con } n \text{ intero non negativo.}$$

Soluzione. Poiché $x^k f \in L_1(\mathbb{R})$ per ogni intero positivo k , $\hat{f} \in C^\infty(\mathbb{R})$.

Il valore di q è determinato dalla condizione

$$\begin{array}{ll} f, f', f'', \dots, f^{(q-1)} & \text{continue} \\ f^{(q)} & \text{continua a tratti.} \end{array}$$

L'unico punto possibile di discontinuità per le derivate di f è l'origine. Si ha per $x > 0$

$$f(x) = f_+(x) = x^n e^{-x},$$

quindi usando la regola di Leibniz,

$$f_+^{(k)}(x) = \sum_{j=0}^k \binom{k}{j} D^j(x^n) D^{k-j}(e^{-x}) = \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} D^j(x^n) e^{-x}.$$

Poiché

$$[D^j(x^n)]_{x=0} = \begin{cases} n! & \text{se } j = n \\ 0 & \text{se } j \neq n \end{cases}$$

ottengo

$$f_+^{(k)}(0) = \begin{cases} 0 & \text{se } k < n \\ (-1)^{k-n} \binom{k}{n} n! = (-1)^{k-n} \frac{k!}{(k-n)!} & \text{se } k \geq n. \end{cases}$$

Il caso $x < 0$ in cui

$$f(x) = f_-(x) = x^n e^x$$

si tratta in modo analogo al precedente. Si ottiene

$$f_-^{(k)}(0) = \begin{cases} 0 & \text{se } k < n \\ \frac{k!}{(k-n)!} & \text{se } k \geq n. \end{cases}$$

Possiamo così concludere che il più piccolo valore di k tale che $f_+^{(k)}(0) \neq f_-^{(k)}(0)$ è $k = n + 1$. Quindi $q = n + 1$.

☆ **9.25 Problema.** (Il principio di indeterminazione di Heisenberg). Si consideri lo spazio delle funzioni di Schwartz $\mathcal{S}(\mathbb{R}; \mathbb{C})$ con il prodotto scalare canonico $\langle f, g \rangle := \int_{\mathbb{R}} f(x) \overline{g(x)} dx$. Sia $\psi \in \mathcal{S}(\mathbb{R}; \mathbb{C})$ di norma unitaria

$$\|\psi\|^2 = \int_{\mathbb{R}} |\psi(x)|^2 dx = 1.$$

e sia $\varphi(\lambda) := \mathcal{F}(\psi)(\lambda)/\sqrt{2\pi}$. Anche φ ha norma unitaria (giusto?). Si assuma inoltre che valgano le condizioni

$$\textcircled{1} \quad \int_{\mathbb{R}} x |\psi(x)|^2 dx = 0 \quad \int_{\mathbb{R}} \lambda |\varphi(\lambda)|^2 d\lambda = 0.$$

(Data una ψ arbitraria di norma 1 è sempre possibile definire $\tilde{\psi}(x) = e^{iax} \psi(x - b)$ in modo tale che le condizioni ① siano verificate per $\tilde{\psi}$). Interpretando $|\psi(x)|^2$ e $|\varphi(\lambda)|^2$ come densità di probabilità, le ① affermano che le variabili x e λ hanno media nulla. Per questo motivo, le rispettive varianze sono date da

$$\sigma_x^2 = \int_{\mathbb{R}} x^2 |\psi(x)|^2 dx = \|x\psi\|^2 \quad \sigma_\lambda^2 = \int_{\mathbb{R}} \lambda^2 |\varphi(\lambda)|^2 d\lambda = \|\lambda\varphi\|^2.$$

Dimostrare che¹ vale il *principio di indeterminazione*: $\sigma_x \sigma_\lambda \geq 1/2$.

(Sugg: usando le proprietà della trasformata di Fourier, dimostrare che $\|\lambda\varphi\|^2 = \|\psi'\|^2$. Questo fatto, insieme alla disuguaglianza di Cauchy-Schwarz,² implica che $\sigma_x^2 \sigma_\lambda^2 \geq |\langle x\psi, \psi' \rangle|^2$. Per stimare quest'ultima quantità usare infine il fatto che $|z| \geq |\operatorname{Re} z| = |z + \bar{z}|/2$).

9.5 Formula di inversione e Teorema di Plancherel

Dimostriamo ora la formula (9.2) annunciata all'inizio del capitolo, che afferma che se \hat{f} è la trasformata di Fourier di f

$$(9.23) \quad \hat{f}(\lambda) = \mathcal{F}(f)(\lambda) = \int_{\mathbb{R}} f(x) e^{-i\lambda x} dx,$$

allora è possibile invertire la trasformazione e ritrovare f a partire da \hat{f} tramite la formula

$$(9.24) \quad f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda x} d\lambda.$$

9.26 Definizione. Se $f \in L_1(\mathbb{R})$ si definisce *antitrasformata di Fourier* di f e si indica con $\mathcal{F}^a f = \check{f}$, la funzione

$$\mathcal{F}^a(f)(x) = \check{f}(x) := \frac{1}{2\pi} \int_{\mathbb{R}} f(\lambda) e^{i\lambda x} d\lambda \quad x \in \mathbb{R}.$$

Si vede che l'antitrasformata di Fourier assomiglia molto alla trasformata di Fourier, le uniche differenze sono il fattore $1/(2\pi)$ e il segno dell'esponente nell'esponenziale. Usando il cambio di variabile $u = -\lambda$, posso scrivere

$$\mathcal{F}^a(f)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} f(\lambda) e^{i\lambda x} d\lambda = \frac{1}{2\pi} \int_{\mathbb{R}} f(-u) e^{-iux} du = \frac{1}{2\pi} \mathcal{F}(Pf)(x),$$

per cui, ricordando anche il punto (1) della Proposizione 9.6, otteniamo l'identità operatoriale

$$(9.25) \quad \mathcal{F}^a = \frac{1}{2\pi} (\mathcal{F} \circ P) = \frac{1}{2\pi} (P \circ \mathcal{F}).$$

Vogliamo ora dimostrare che l'antitrasformata di Fourier è proprio l'inverso della trasformata di Fourier, cioè che $\mathcal{F}^a = \mathcal{F}^{-1}$ o, in altre parole, che

$$\mathcal{F}^a \circ \mathcal{F} = \mathcal{F} \circ \mathcal{F}^a = I.$$

¹in un ipotetico universo nel quale $\hbar = 1 \dots$

²non è un errore di battitura, gli Schwarz(t)z sono 2, uno con la "t" e uno senza.

Inizialmente ci restringiamo alla classe delle funzioni di Schwartz. Che una qualche forma di restrizione sia necessaria è chiaro per il motivo seguente. Se $f \in L_1(\mathbb{R})$ possiamo calcolarne la sua trasformata di Fourier \hat{f} tramite la (9.23). Fin qui nessun intralcio. Il problema nasce nel caso in cui la funzione f non appartenga ad $L_1(\mathbb{R})$. Prendiamo, per dire,

$$f(x) = \begin{cases} 1 & \text{se } |x| \leq 1 \\ 0 & \text{se } |x| > 1. \end{cases}$$

La trasformata di Fourier \hat{f} in questo caso è

$$\hat{f}(\lambda) = \frac{2 \sin \lambda}{\lambda},$$

che *non* è una funzione in $L_1(\mathbb{R})$. Di conseguenza la formula di inversione (9.24), così com'è scritta non può essere valida semplicemente perché l'integrale non è ben definito. Queste complicazioni non esistono nel caso in cui f sia una funzione di Schwartz, perché in quel caso anche \hat{f} è rapidamente decrescente e dunque gli integrali che appaiono nelle formule (9.23) e (9.24) sono entrambi ben definiti.

Una volta stabilita la formula di inversione nel caso di f rapidamente decrescente, spiegheremo come si può estendere ad una classe più ampia di funzioni.

9.27 Teorema. (1) Se $f \in \mathcal{S}(\mathbb{R})$, allora

$$(9.26) \quad \mathcal{F}^a \mathcal{F} f = \mathcal{F} \mathcal{F}^a f \quad (\text{Formula di inversione}),$$

quindi \mathcal{F}^a rappresenta l'operatore inverso \mathcal{F}^{-1} della trasformata di Fourier definita su $\mathcal{S}(\mathbb{R})$. Le mappe \mathcal{F} e \mathcal{F}^a sono trasformazioni biunivoche da $\mathcal{S}(\mathbb{R})$ a $\mathcal{S}(\mathbb{R})$.

(2) Sia $\langle f, g \rangle := \int_{\mathbb{R}} f(x) g(x) dx$ il prodotto scalare canonico e sia $\| \cdot \|_2$ la norma corrispondente. Allora, per ogni $f, g \in \mathcal{S}(\mathbb{R})$, si ha

$$(9.27) \quad \|\mathcal{F}(f)\|_2 = \sqrt{2\pi} \|f\|_2 \quad (\text{Teorema di Plancherel})$$

$$(9.28) \quad \langle \mathcal{F}(f), \mathcal{F}(g) \rangle = 2\pi \langle f, g \rangle$$

Dimostrazione. Assumiamo inizialmente che $f \in C_c^\infty(\mathbb{R})$ e scegliamo $a > 0$ tale che $f(x) = 0$ se $|x| > a$. Sia poi $\varepsilon \in (0, 1/a)$ e sviluppiamo f in serie di Fourier nell'intervallo $[-1/\varepsilon, 1/\varepsilon]$.

$$f(x) = \sum_{k \in \mathbb{Z}} c_k e^{ik\pi\varepsilon x}$$

con

$$(9.29) \quad c_k = \frac{\varepsilon}{2} \int_{-1/\varepsilon}^{1/\varepsilon} f(x) e^{-ik\pi\varepsilon x} dx = \frac{\varepsilon}{2} \int_{\mathbb{R}} f(x) e^{-ik\pi\varepsilon x} dx = \frac{\varepsilon}{2} \hat{f}(k\pi\varepsilon).$$

Possiamo quindi scrivere

$$f(x) = \frac{\varepsilon}{2} \sum_{k \in \mathbb{Z}} \hat{f}(k\pi\varepsilon) e^{ik\pi\varepsilon x}.$$

Ponendo

$$\lambda_k^{(\varepsilon)} := k\pi\varepsilon,$$

abbiamo

$$(9.30) \quad \begin{aligned} f(x) &= \frac{\varepsilon}{2} \sum_{k \in \mathbb{Z}} \hat{f}(\lambda_k^{(\varepsilon)}) e^{i\lambda_k^{(\varepsilon)} x} = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \hat{f}(\lambda_k^{(\varepsilon)}) e^{i\lambda_k^{(\varepsilon)} x} (\lambda_k^{(\varepsilon)} - \lambda_{k-1}^{(\varepsilon)}) \\ &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \hat{f}(\lambda_k^{(\varepsilon)}) e^{i\lambda_k^{(\varepsilon)} x} \Delta\lambda_k^{(\varepsilon)} \end{aligned}$$

Poiché questa identità vale per ogni $\varepsilon < 1/a$ posso fare il limite per $\varepsilon \rightarrow 0^+$. Osservo che la somma su k è una somma di Riemann fatta sulla partizione $\lambda_k^{(\varepsilon)}$ di modulo $\pi\varepsilon$. Quindi, quando $\varepsilon \rightarrow 0^+$, la somma tende all'integrale di Riemann

$$(9.31) \quad f(x) = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \hat{f}(\lambda_k^{(\varepsilon)}) e^{i\lambda_k^{(\varepsilon)} x} \Delta\lambda_k^{(\varepsilon)} = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda x} d\lambda,$$

che è proprio la (9.26).

Per dimostrare la (9.27) partiamo dall'uguaglianza di Parseval. Usando la (9.29) otteniamo

$$\int_{\mathbb{R}} |f(x)|^2 dx = \frac{2}{\varepsilon} \sum_{k \in \mathbb{Z}} |c_k|^2 = \frac{\varepsilon}{2} \sum_{k \in \mathbb{Z}} |g(k\pi\varepsilon)|^2 = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} |g(\lambda_k^{(\varepsilon)})|^2 \Delta\lambda_k^{(\varepsilon)}.$$

Facendo tendere nuovamente ε a zero da destra si ha

$$\int_{\mathbb{R}} |f(x)|^2 dx = \frac{1}{2\pi} \int_{\mathbb{R}} |g(\lambda)|^2 d\lambda,$$

che è appunto la (9.27). La (9.28) segue dalla (9.27) e dall'identità di polarizzazione 4.10.

Estensione per continuità. Senza fornire una vera dimostrazione voglio dare un'idea di come si può estendere, senza troppo sforzo, la formula di inversione (9.26) a tutte le funzioni rapidamente decrescenti, avendolo dimostrato per le funzioni C^∞ a supporto compatto. Si tratta di un trucco molto utile che si usa spesso in matematica chiamato *estensione per continuità*. Faccio un'analogia: siano p e q due funzioni reali e supponiamo di sapere che

- (1) p e q sono entrambi continue
- (2) $p(x) = q(x)$ per ogni x razionale.

Allora segue immediatamente che le due funzioni coincidono su tutto l'asse reale. Infatti se x non è razionale trovo una successione (x_n) di razionali che converge ad x , usando la continuità di p e q , ottengo

$$p(x) = p\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} p(x_n) = \lim_{n \rightarrow \infty} q(x_n) = q\left(\lim_{n \rightarrow \infty} x_n\right) = q(x).$$

In questo modo abbiamo esteso "per continuità" la formula

$$p(x) = q(x) \quad \forall x \in \mathbb{Q}$$

a tutto l'asse reale. È chiaro che al posto di \mathbb{Q} può esserci un qualsiasi sottoinsieme denso di \mathbb{R} .

Torniamo ora al teorema di inversione. Abbiamo fin qui dimostrato che

$$\mathcal{F}^a \mathcal{F}(f) = \mathcal{F} \mathcal{F}^a(f) = f = I(f) \quad \forall f \in C_c^\infty(\mathbb{R}).$$

Per poter estendere questa identità a tutto lo spazio $\mathcal{S}(\mathbb{R})$ quello che bisogna fare è introdurre una metrica opportuna su $\mathcal{S}(\mathbb{R})$ e fare vedere che, rispetto a questa metrica,

- (a) gli operatori \mathcal{F} e \mathcal{F}^a sono continui (l'operatore identità è sempre continuo)
 (b) $C_c^\infty(\mathbb{R})$ è denso in $\mathcal{S}(\mathbb{R})$,

e a questo punto si procede come nel caso delle due funzioni p e q . La metrica che serve è questo scopo è proprio quella introdotta nella sezione 3.10.4

$$(9.32) \quad d(f, g) = \sum_{j,k=0}^{\infty} \frac{1}{2^{j+k}} \frac{\|f - g\|_{j,k}}{1 + \|f - g\|_{j,k}}.$$

Si dimostra infatti che \mathcal{F} e \mathcal{F}^a sono continui rispetto a questa metrica e che $C_c^\infty(\mathbb{R})$ è denso in $(\mathcal{S}(\mathbb{R}), d)$, come già affermato nel Teorema 3.98.

Una volta dimostrato che $\mathcal{F} \circ \mathcal{F}^a = \mathcal{F}^a \circ \mathcal{F} = I$, segue che \mathcal{F} e \mathcal{F}^a sono entrambi biunivoci. \square

La formula di inversione è valida per una classe molto più ampia di funzioni rispetto alle funzioni rapidamente decrescenti. Il teorema seguente fornisce un risultato simile al Corollario 8.23 che riguarda la convergenza della serie di Fourier. Diamo, senza dimostrazione, il seguente risultato.

9.28 Teorema. *Sia $f \in L_1(\mathbb{R})$ differenziabile a tratti e sia $\hat{f} := \mathcal{F}(f)$ la trasformata di Fourier di f . Allora abbiamo*

$$(9.33) \quad \frac{f(x^-) + f(x^+)}{2} = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda x} d\lambda \quad \forall x \in \mathbb{R}.$$

Di conseguenza, se f è continua in x , si ottiene

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda x} d\lambda.$$

Nel caso in cui \hat{f} non appartenga ad $L_1(\mathbb{R})$ l'integrale $\int_{\mathbb{R}}$ va inteso come

$$(9.34) \quad \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda x} d\lambda = \lim_{K \rightarrow +\infty} \int_{-K}^K \hat{f}(\lambda) e^{i\lambda x} d\lambda.$$

9.29 Esempio. Per illustrare il contenuto del Teorema 9.28, sia $f(x) = H(x)e^{-x}$. Nel Problema 9.9 abbiamo visto che $\hat{f}(\lambda) = 1/(1 + i\lambda)$. Proviamo ora a calcolare l'antitrasformata di \hat{f} . Poiché $\hat{f} \notin L_1(\mathbb{R})$ dobbiamo usare la (9.34), quindi

$$(\mathcal{F}^{-1}\hat{f})(x) = \frac{1}{2\pi} \lim_{K \rightarrow +\infty} \int_{-K}^K \frac{e^{i\lambda x}}{1 + i\lambda} d\lambda.$$

Dobbiamo considerare 3 casi.

Caso (1): se $x > 0$, grazie al Lemma di Jordan possiamo chiudere il cammino di integrazione nel semipiano superiore del piano complesso ed ottenere, col Teorema dei residui,

$$(\mathcal{F}^{-1}\hat{f})(x) = \frac{1}{2\pi} 2\pi i \operatorname{Res} \left(\frac{e^{izx}}{1 + iz}, i \right) = e^{-x}.$$

Caso (2): se $x < 0$, posso chiudere il cammino di integrazione nel semipiano inferiore del piano complesso. Siccome in questo semipiano non ci sono singolarità, l'integrale è nullo.

Caso (3): se $x = 0$, si ha

$$\begin{aligned}
 (\mathcal{F}^{-1}\hat{f})(0) &= \frac{1}{2\pi} \lim_{K \rightarrow +\infty} \int_{-K}^K \frac{1}{1+i\lambda} d\lambda = \frac{1}{2\pi} \lim_{K \rightarrow +\infty} \int_{-K}^K \frac{1-i\lambda}{1+\lambda^2} d\lambda \\
 &= \frac{1}{2\pi} \lim_{K \rightarrow +\infty} \left[\int_{-K}^K \frac{1}{1+\lambda^2} d\lambda + i \underbrace{\int_{-K}^K \frac{\lambda}{1+\lambda^2} d\lambda}_{=0 \text{ perché dispari}} \right] \\
 &= \frac{1}{2\pi} \lim_{K \rightarrow +\infty} (\arctan(K) - \arctan(-K)) = \frac{1}{2}.
 \end{aligned}$$

Riassumendo: abbiamo ottenuto, in accordo con la (9.33),

$$(\mathcal{F}^{-1} \circ \mathcal{F}f)(x) = \begin{cases} H(x) e^{-x} = f(x) & \text{se } x \neq 0 \\ \frac{1}{2} = (f(0^-) + f(0^+))/2 & \text{se } x = 0. \end{cases}$$

9.30. Regola di calcolo. Dal Teorema 9.27 e dalla definizione (9.25) seguono le identità operatoriali

$$(9.35) \quad \mathcal{F}^{-1} = \frac{1}{2\pi} (\mathcal{F} \circ P) = \frac{1}{2\pi} (P \circ \mathcal{F})$$

$$(9.36) \quad \mathcal{F} \circ \mathcal{F} = 2\pi I,$$

che rappresenteremo graficamente nel modo seguente

$$\begin{array}{llll}
 [f \rightarrow \hat{f}] & \hat{f}(x) & \xrightarrow{\mathcal{F}} & 2\pi f(-\lambda) \\
 [f \rightarrow \hat{f}(-x)] & \hat{f}(-x) & \xrightarrow{\mathcal{F}} & 2\pi f(\lambda).
 \end{array}$$

9.31 Problema. Sapendo che $\mathcal{F}[H(x) e^{-x}](\lambda) = 1/(1+i\lambda)$. Calcolare la trasformata di Fourier della funzione

$$\frac{1}{(x+i)^3}.$$

Soluzione. Partendo dalla trasformata di Fourier che ci viene fornita nel testo del problema, otteniamo

$$\begin{array}{llll}
 & H(x) e^{-x} & \xrightarrow{\mathcal{F}} & \frac{1}{1+i\lambda} \\
 [f \rightarrow \hat{f}] & \frac{1}{1+ix} & \xrightarrow{\mathcal{F}} & 2\pi H(-\lambda) e^\lambda \\
 [x \rightarrow (-x)] & \frac{1}{1-ix} & \xrightarrow{\mathcal{F}} & 2\pi H(\lambda) e^{-\lambda} \\
 [f \rightarrow (-i)f] & \frac{1}{x+i} & \xrightarrow{\mathcal{F}} & -2i\pi H(\lambda) e^{-\lambda} \\
 [f \rightarrow f''] & \frac{2}{(x+i)^3} & \xrightarrow{\mathcal{F}} & 2i\pi\lambda^2 H(\lambda) e^{-\lambda} \\
 [f \rightarrow f/2] & \frac{1}{(x+i)^3} & \xrightarrow{\mathcal{F}} & i\pi\lambda^2 H(\lambda) e^{-\lambda}
 \end{array}$$

9.6 Convoluzione

La formula $\mathcal{F}(f')(\lambda) = (i\lambda)\mathcal{F}(f)$ ci dice che l'operazione di derivazione si riflette sulla trasformata di Fourier come una semplice moltiplicazione per $i\lambda$. In un qualche senso la derivata è un'operazione che risulta "più semplice" se effettuata nello spazio di Fourier.

Voglio introdurre ora un'altra operazione che ha la stessa caratteristica, vale a dire appare più complicata nello "spazio x " e più semplice nello "spazio λ ".

9.32 Definizione. Date due funzioni $f, g \in L_1(\mathbb{R})$ limitate, si definisce *convoluzione* (o prodotto di convoluzione) fra f e g e si indica con il simbolo $f * g$, la funzione

$$(9.37) \quad (f * g)(x) := \int_{\mathbb{R}} f(y) g(x - y) dy \quad x \in \mathbb{R}.$$

9.33 Osservazione. Come al solito quando in una definizione appare un integrale bisogna cautelarsi che l'integrale esista e sia finito. Nel caso della (9.37) non ci sono problemi in quanto, sotto le ipotesi appena fatte per f e g , si ottiene

$$(9.38) \quad \begin{aligned} \left| \int_{\mathbb{R}} f(y) g(x - y) dy \right| &\leq \int_{\mathbb{R}} |f(y) g(x - y)| dy \\ &\leq \|f\|_u \int_{\mathbb{R}} |g(x - y)| dy = \|f\|_u \|g\|_1. \end{aligned}$$

9.34 Proposizione. Siano $f, g \in L_1(\mathbb{R})$ limitate. Allora anche la convoluzione $f * g$ è un elemento di $L_1(\mathbb{R})$ ed è limitato. Inoltre

$$\|f * g\|_u \leq \|f\|_u \|g\|_1 \quad \|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Dimostrazione. La disuguaglianza (9.38) ci dice che, per ogni $x \in \mathbb{R}$, $|(f * g)(x)| \leq \|f\|_u \|g\|_1$, per cui $f * g$ è limitata e vale

$$\|f * g\|_u \leq \|f\|_u \|g\|_1.$$

Inoltre, se $a, b \in \mathbb{R}$ con $a < b$, scambiando l'ordine degli integrali e cambiando variabile otteniamo

$$\begin{aligned} \int_a^b |(f * g)(x)| dx &= \int_a^b \left| \int_{\mathbb{R}} f(y) g(x - y) dy \right| dx \leq \int_a^b \left[\int_{\mathbb{R}} |f(y) g(x - y)| dy \right] dx \\ &= \int_{\mathbb{R}} \left[\int_a^b |f(y) g(x - y)| dx \right] dy = \int_{\mathbb{R}} |f(y)| \left[\int_a^b |g(x - y)| dx \right] dy \\ &= \int_{\mathbb{R}} |f(y)| \left[\int_{a-y}^{b-y} |g(u)| du \right] dy \leq \int_{\mathbb{R}} |f(y)| \left[\int_{-\infty}^{\infty} |g(u)| du \right] dy \\ &= \int_{\mathbb{R}} |f(y)| \|g\|_1 dy = \|g\|_1 \|f\|_1. \end{aligned}$$

Nella precedente disuguaglianza possiamo far tendere a a $-\infty$ e b a $+\infty$. In questo modo si ha

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Quindi $f * g \in L_1(\mathbb{R})$. □

9.35 Proposizione. *Il prodotto di convoluzione è commutativo e associativo, vale a dire, se f, g, h sono 3 funzioni limitate di $L_1(\mathbb{R})$, si ha*

$$f * g = g * f \qquad (f * g) * h = f * (g * h).$$

Dimostrazione. Ponendo $u = x - y$ nell'integrale che definisce la convoluzione si ottiene

$$(g * f)(x) := \int_{\mathbb{R}} g(y) f(x - y) dy = \int_{\mathbb{R}} g(x - u) f(u) du = (f * g)(x).$$

Inoltre

$$\begin{aligned} [(f * g) * h](x) &= \int_{\mathbb{R}} (f * g)(y) h(x - y) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t) g(y - t) dt \right) h(x - y) dy \\ &= \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} g(y - t) h(x - y) dy \right) dt \\ &= \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} g(u) h(x - t - u) dy \right) dt \\ &= \int_{\mathbb{R}} f(t) (g * h)(x - t) dt = [f * (g * h)](x). \quad \square \end{aligned}$$

Una caratteristica utile della convoluzione è costituita dal fatto che affinché il prodotto $f * g$ sia “liscio” è sufficiente che sia liscia una sola delle due funzioni f e g . Più precisamente si ha:

9.36 Proposizione. *Siano $f, g \in L_1(\mathbb{R})$ limitate. Se $g \in C^1(\mathbb{R})$ e se g' è limitata e appartiene a $L_1(\mathbb{R})$, allora $f * g$ è derivabile e si ha*

$$(f * g)' = f * g'.$$

Dimostrazione. Sia

$$A(t, x) = f(t) g(x - t)$$

per cui

$$(f * g)(x) = \int_{\mathbb{R}} A(t, x) dt.$$

Bisogna semplicemente far vedere che si può portare la derivata sotto il segno dell'integrale che definisce la convoluzione, vale a dire che

$$\frac{\partial}{\partial x} \int_{\mathbb{R}} A(t, x) dt = \int_{\mathbb{R}} \partial_x A(t, x) dt.$$

Poiché g' per ipotesi è limitata, esiste $M > 0$ tale che $|g'(x)| \leq M$, per ogni $x \in \mathbb{R}$, quindi

$$|\partial_x A(t, x)| = |f(t) g'(x - t)| \leq M |f(t)|$$

e, dato che $f \in L_1(\mathbb{R})$, si ha $\int_{\mathbb{R}} |f(t)| dt < \infty$. Questa è la condizione fondamentale che appare come ipotesi (c) nel Teorema 1.32 che ci permette appunto di derivare sotto il segno di integrale. Si verifica facilmente che l'argomento usato nella dimostrazione di quel teorema, si può applicare, essenzialmente senza modifiche, al caso in esame. \square

Veniamo ora al risultato più importante che riguarda il prodotto di convoluzione che afferma che “nello spazio di Fourier” il prodotto di convoluzione diventa un prodotto semplice.

9.37 Proposizione. Siano $f, g \in L_1(\mathbb{R})$ limitate. Allora

$$\mathcal{F}(f * g) = \mathcal{F}(f) \mathcal{F}(g).$$

Dimostrazione. Siano f e g due funzioni limitate nello spazio $L_1(\mathbb{R})$. Allora

$$\begin{aligned} \mathcal{F}(f * g)(\lambda) &:= \int_{\mathbb{R}} (f * g)(x) e^{-i\lambda x} dx \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(y) g(x - y) dy \right] e^{-i\lambda x} dx \end{aligned}$$

Le ipotesi su f e g ci consentono di scambiare l'ordine di integrazione, dunque

$$\begin{aligned} \mathcal{F}(f * g)(\lambda) &= \int_{\mathbb{R}} f(y) \left[\int_{\mathbb{R}} g(x - y) e^{-i\lambda x} dx \right] dy \\ &= \int_{\mathbb{R}} f(y) e^{-i\lambda y} \left[\int_{\mathbb{R}} g(x - y) e^{-i\lambda(x-y)} dx \right] dy \\ &= \int_{\mathbb{R}} f(y) e^{-i\lambda y} \left[\int_{\mathbb{R}} g(u) e^{-i\lambda u} du \right] dy \\ &= \int_{\mathbb{R}} f(y) e^{-i\lambda y} \mathcal{F}(g)(\lambda) dy = \mathcal{F}(f)(\lambda) \mathcal{F}(g)(\lambda). \quad \square \end{aligned}$$

9.38 Esempio. Sia $a > 0$. Vogliamo calcolare $\mathbb{1}_{[-a, a]} * \mathbb{1}_{[-a, a]}$. Osservo innanzitutto che, dati 2 insiemi qualsiasi A e B si ha $\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A \cap B}$. Inoltre

$$\mathbb{1}_{[-a, a]}(-x) = \mathbb{1}_{[-a, a]}(x) \quad \mathbb{1}_{[-a, a]}(x - u) = \mathbb{1}_{[-a+u, a+u]}(x),$$

quindi

$$\begin{aligned} (\mathbb{1}_{[-a, a]} * \mathbb{1}_{[-a, a]})(x) &= \int_{\mathbb{R}} \mathbb{1}_{[-a, a]}(y) \mathbb{1}_{[-a, a]}(x - y) dy \\ &= \int_{\mathbb{R}} \mathbb{1}_{[-a, a]}(y) \mathbb{1}_{[-a, a]}(y - x) dy \\ &= \int_{\mathbb{R}} \mathbb{1}_{[-a, a]}(y) \mathbb{1}_{[-a+x, a+x]}(y) dy \\ &= \int_{\mathbb{R}} \mathbb{1}_{[-a, a] \cap [-a+x, a+x]}(y) dy = |[-a, a] \cap [-a + x, a + x]|, \end{aligned}$$

dove, con $|[c, d]|$ indichiamo la lunghezza del segmento $[c, d]$. Si verifica facilmente che

$$|[-a, a] \cap [-a + x, a + x]| = \begin{cases} 0 & \text{se } x > 2a \text{ o } x < -2a \\ 2a - x & \text{se } x \in [0, 2a] \\ 2a + x & \text{se } x \in [-2a, 0]. \end{cases}$$

Di conseguenza otteniamo

$$(9.39) \quad (\mathbb{1}_{[-a, a]} * \mathbb{1}_{[-a, a]})(x) = \max\{2a - |x|, 0\} = 2a \operatorname{tri}(x/2a).$$

Nel caso particolare in cui $a = 1/2$ la (9.39) diventa

$$(9.40) \quad \operatorname{rect} * \operatorname{rect} = \operatorname{tri}.$$

Osserviamo che, applicando \mathcal{F} ad entrambi i membri della (9.40) e ricordando la (9.8), si trova

$$\mathcal{F}(\text{tri})(\lambda) = \mathcal{F}(\text{rect} * \text{rect})(\lambda) = \mathcal{F}(\text{rect})^2(\lambda) = \text{sinc}^2\left(\frac{\lambda}{2\pi}\right).$$

Abbiamo in questo modo ritrovato lo stesso risultato ottenuto nel Problema 9.8, senza dover calcolare esplicitamente la trasformata di Fourier.

9.39 Problema. Verificare che, se $a > 0$, allora

$$\mathbb{I}_{[-a,a]}^{*3}(x) := (\mathbb{I}_{[-a,a]} * \mathbb{I}_{[-a,a]} * \mathbb{I}_{[-a,a]})(x) = \begin{cases} 3a^2 - x^2 & \text{se } |x| \leq a \\ (3a - |x|)^2/2 & \text{se } a < |x| \leq 3a \\ 0 & \text{se } |x| > 3a. \end{cases}$$

9.40 Problema. Siano α e β due numeri reali positivi. Calcolare

$$e^{-\alpha x^2} * e^{-\beta x^2}.$$

Soluzione. Invece di calcolare direttamente la convoluzione, che comunque non sarebbe difficile, usiamo il Teorema 9.37. Quindi

$$\begin{aligned} e^{-\alpha x^2} * e^{-\beta x^2} &= \mathcal{F}^{-1} \left[\mathcal{F} \left(e^{-\alpha x^2} * e^{-\beta x^2} \right) \right] = \mathcal{F}^{-1} \left[\mathcal{F} \left(e^{-\alpha x^2} \right) \mathcal{F} \left(e^{-\beta x^2} \right) \right] \\ &= \mathcal{F}^{-1} \left[\sqrt{\frac{\pi}{\alpha}} e^{-\lambda^2/(4\alpha)} \sqrt{\frac{\pi}{\beta}} e^{-\lambda^2/(4\beta)} \right] \\ &= \pi \sqrt{\frac{1}{\alpha\beta}} \mathcal{F}^{-1} \left[\exp \left(-\frac{\lambda^2}{4} \frac{\alpha + \beta}{\alpha\beta} \right) \right]. \end{aligned}$$

Poiché $\mathcal{F}^{-1} e^{-\lambda^2/(4a)} = \sqrt{a/\pi} e^{-ax^2}$, otteniamo

$$\begin{aligned} e^{-\alpha x^2} * e^{-\beta x^2} &= \pi \sqrt{\frac{1}{\alpha\beta}} \sqrt{\frac{\alpha\beta}{(\alpha + \beta)\pi}} \exp \left(-x^2 \frac{\alpha\beta}{\alpha + \beta} \right) \\ &= \sqrt{\frac{\pi}{\alpha + \beta}} \exp \left(-x^2 \frac{\alpha\beta}{\alpha + \beta} \right). \end{aligned}$$

9.41 Problema. Dimostrare che, se P , T_a e Φ_a sono gli operatori di parità, traslazione e dilatazione introdotti nella sezione 9.2, si ha

$$\begin{aligned} P(f * g) &= (Pf) * (Pg) \\ (T_a f) * g &= f * (T_a g) = T_a(f * g) \\ (\Phi_a f) * (\Phi_a g) &= \frac{1}{|a|} (\Phi_a f) * (\Phi_a g) \quad a \neq 0. \end{aligned}$$

9.42 Problema. Dimostrare che, se $a > 0$, si ha

$$\int_{\mathbb{R}} \left(\frac{\sin(ax)}{x} \right)^3 dx = \frac{3\pi a^2}{4}.$$

(Sugg: usare il teorema di Plancherel).

Soluzione. Sia $f(x) = \sin(ax)/x$. Allora

$$f(x) = a \operatorname{sinc}(ax/\pi),$$

A partire dalla trasformate di Fourier della funzione rettangolo, otteniamo

$$\begin{array}{llll} & \operatorname{rect}(x) & \xrightarrow{\mathcal{F}} & \operatorname{sinc}(\lambda/(2\pi)) \\ [f \rightarrow \hat{f}] & \operatorname{sinc}(x/(2\pi)) & \xrightarrow{\mathcal{F}} & 2\pi \operatorname{rect}(-\lambda) = 2\pi \operatorname{rect}(\lambda) \\ [x \rightarrow 2ax] & \operatorname{sinc}(ax/\pi) & \xrightarrow{\mathcal{F}} & \frac{\pi}{a} \operatorname{rect}(\lambda/(2a)) \\ [f \rightarrow af] & f(x) & \xrightarrow{\mathcal{F}} & \pi \operatorname{rect}(\lambda/(2a)). \end{array}$$

Analogamente

$$\begin{array}{llll} & \operatorname{tri}(x) & \xrightarrow{\mathcal{F}} & \operatorname{sinc}^2(\lambda/(2\pi)) \\ [f \rightarrow \hat{f}] & \operatorname{sinc}^2(x/(2\pi)) & \xrightarrow{\mathcal{F}} & 2\pi \operatorname{tri}(-\lambda) = 2\pi \operatorname{tri}(\lambda) \\ [x \rightarrow 2ax] & \operatorname{sinc}^2(ax/\pi) & \xrightarrow{\mathcal{F}} & \frac{\pi}{a} \operatorname{tri}(\lambda/(2a)) \\ [f \rightarrow af] & f^2(x) & \xrightarrow{\mathcal{F}} & a\pi \operatorname{tri}(\lambda/(2a)). \end{array}$$

Quindi, usando il teorema di Plancherel, si ottiene

$$\begin{aligned} \int_{\mathbb{R}} \left(\frac{\sin(ax)}{x} \right)^3 dx &= \langle f^2, f \rangle_{L_2} = \frac{1}{2\pi} \langle \hat{f}^2, \hat{f} \rangle_{L_2} \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} a\pi^2 \operatorname{tri}(\lambda/(2a)) \operatorname{rect}(\lambda/(2a)) d\lambda \\ &= a^2\pi \int_{\mathbb{R}} \operatorname{tri}(t) \operatorname{rect}(t) dt = a^2\pi \int_{-1/2}^{1/2} \operatorname{tri}(t) dt \\ &= 2a^2\pi \int_0^{1/2} \operatorname{tri}(t) dt. \end{aligned}$$

L'ultimo integrale è uguale all'area sottesa dal grafico della funzione $\operatorname{tri}(\cdot)$ fra i punti 0 e 1/2, vale a dire l'area di un trapezio di base maggiore 1, base minore 1/2 e altezza 1/2, quindi

$$\int_{\mathbb{R}} \left(\frac{\sin(ax)}{x} \right)^3 dx = 2a^2\pi \frac{(1 + 1/2) \cdot 1/2}{2} = \frac{3a^2\pi}{4}.$$

9.7 Equazione del calore su un intervallo infinito

Consideriamo nuovamente l'equazione del calore già incontrata nella sezione 8.4, ma questa volta su tutto l'asse reale.

$$(9.41) \quad u_t(x, t) = K u_{xx}(x, t) \quad x \in \mathbb{R}, \quad t > 0$$

$$(9.42) \quad u(x, 0) = u_0(x) \quad x \in \mathbb{R}.$$

Immaginiamo quindi una barra infinita fatta di un qualche materiale conduttore di conducibilità termica $K > 0$, che inizialmente ha un profilo di temperatura dato dalla funzione $u_0(x) \geq 0$. La soluzione $u(x, t)$ di questa equazione fornisce la temperatura della barra nel punto x ad un tempo qualsiasi $t > 0$, successivo all'istante iniziale.

Sia $v(\lambda, t)$ la trasformata di Fourier nella variabile x della funzione $u(x, t)$

$$(9.43) \quad v(\lambda, t) = \mathcal{F}(u(x, t)) = \int_{\mathbb{R}} u(x, t) e^{-i\lambda x} dx.$$

La Proposizione 9.15 ci dice che

$$(9.44) \quad (i\lambda)^2 v(\lambda, t) = \mathcal{F}(u_{xx}(x, t)).$$

D'altra parte, supponendo di poter scambiare la derivata rispetto a t con l'integrale nella (9.43), possiamo scrivere

$$v_t(\lambda, t) = \int_{\mathbb{R}} u_t(x, t) e^{-i\lambda x} dx,$$

vale a dire

$$v_t(\lambda, t) = \mathcal{F}(u_t(x, t)).$$

Applicando quindi l'operazione trasformata di Fourier alle uguaglianze (9.41), (9.42), queste si diventano

$$\begin{aligned} v_t(\lambda, t) &= -K \lambda^2 v(\lambda, t) & \lambda \in \mathbb{R}, \quad t > 0 \\ v(\lambda, 0) &= \mathcal{F}(u_0) & \lambda \in \mathbb{R}. \end{aligned}$$

Si vedi quindi che, nello spazio di Fourier, l'equazione differenziale a derivate parziali (9.41), (9.42) si è trasformata in una collezione infinita di equazioni differenziali ordinarie, una per ogni, valore di λ la cui soluzione è data da

$$(9.45) \quad v(\lambda, t) = v(\lambda, 0) e^{-K\lambda^2 t} = \mathcal{F}(u_0)(\lambda) e^{-K\lambda^2 t}.$$

A questo punto conviene scrivere il fattore esponenziale come la trasformata di Fourier di qualcos'altro. Trattandosi di una gaussiana si ottiene, come è noto

$$e^{-K\lambda^2 t} = \mathcal{F}\left[\frac{1}{2\sqrt{\pi K t}} e^{-x^2/(4Kt)}\right].$$

La (9.45) diventa dunque

$$(9.46) \quad v(\lambda, t) = \mathcal{F}(u_0)(\lambda) \mathcal{F}\left[\frac{1}{2\sqrt{\pi K t}} e^{-x^2/(4Kt)}\right].$$

La Proposizione 9.37 ci dice che il prodotto delle trasformate di Fourier può essere scritto come la trasformata di Fourier della convoluzione, per cui otteniamo

$$\mathcal{F}(u(x, t)) = \mathcal{F}\left[u_0 * \frac{1}{2\sqrt{\pi K t}} e^{-x^2/(4Kt)}\right].$$

Applicando ad ambo i membri la trasformata di Fourier inversa³ possiamo finalmente scrivere la soluzione dell'equazione del calore esplicitamente in termini del dato iniziale u_0 come

$$(9.47) \quad u(x, t) = u_0(x) * \frac{1}{2\sqrt{\pi K t}} e^{-x^2/(4Kt)} = \frac{1}{2\sqrt{\pi K t}} \int_{\mathbb{R}} e^{-(x-y)^2/(4Kt)} u_0(y) dy.$$

³stiamo assumendo di poter usare il teorema di inversione

9.43 Osservazione. (L'operatore di evoluzione temporale P_t). Definiamo

$$K_t(x, y) := \frac{e^{-(x-y)^2/(4Kt)}}{2\sqrt{\pi Kt}}.$$

La quantità K_t può essere pensata come il nucleo integrale di un operatore P_t che agisce su un qualche spazio, ad esempio $L_1(\mathbb{R})$. La (9.47) può essere riscritta come

$$u(x, t) = \int_{\mathbb{R}} K_t(x, y) u_0(y) dy,$$

o, equivalentemente, in forma simbolica

$$u(\cdot, t) = P_t u_0.$$

L'operatore P_t determina quindi l'evoluzione temporale del profilo di temperatura.

9.44 Osservazione. La quantità

$$\int_{\mathbb{R}} u(x, t) dx$$

è costante nel tempo. Infatti si verifica immediatamente a partire dalla (9.47) che

$$\int_{\mathbb{R}} u(x, t) dx = \int_{\mathbb{R}} u_0(x) dx \quad \forall t \geq 0.$$

A quale velocità si propaga il calore?

Supponiamo di voler determinare a quale velocità si “propaga” il calore. Per farlo si può pensare di usare come condizione iniziale $u_0(x)$ una funzione che abbia un picco molto stretto ad esempio nell'origine e calcolare quanto tempo bisogna aspettare affinché l'effetto arrivi nel punto x . Poniamo ad esempio

$$u_0(x) := \begin{cases} 1 & \text{se } |x| \leq \varepsilon \\ 0 & \text{se } |x| > \varepsilon \end{cases}$$

in cui ε è un numero positivo “molto piccolo”.⁴ Con questa scelta, la soluzione data dalla (9.47) si può scrivere in modo approssimato come

$$(9.48) \quad u(x, t) = \frac{1}{2\sqrt{\pi Kt}} \int_{-\varepsilon}^{\varepsilon} e^{-(x-y)^2/(4Kt)} dy \approx \frac{\varepsilon e^{-x^2/(4Kt)}}{\sqrt{\pi Kt}}.$$

Notiamo subito una cosa: al tempo iniziale la temperatura è diversa da zero soltanto in un intervallo di larghezza 2ε centrato sull'origine. Nonostante ciò, comunque si scelga un valore di t positivo piccolo a piacere e un valore di x grande a piacere, la (9.48) ci dice⁵ che *comunque* si ha $u(x, t) > 0$. Quindi l'effetto del picco nell'origine viene percepito *immediatamente* (anche se in modo “fievole” a causa del fattore gaussiano) a distanze arbitrariamente grandi, cosa che non avviene, ad esempio, nel caso dell'equazione delle onde. In questo senso, quindi, il calore si propaga con *velocità infinita*.

Qualcuno potrebbe obiettare che, sì, va bene, c'è un effetto di propagazione a velocità infinita, ma non è interessante, proprio perché quest'effetto è molto piccolo a causa

⁴quanto piccolo lo dirò più in là

⁵si può usare l'espressione esatta se non ci si fida di quella approssimata

del fattore e^{-x^2} , e allora perché dargli tutta questa importanza? L'importanza sta nel comportamento *qualitativo* della soluzione dell'equazione del calore che per questa (e mille altre) ragioni è radicalmente diverso dal comportamento qualitativo, per dire, dell'equazione delle onde. Comunque, quasi a voler dar ragione (parzialmente) all'obbiettore, considero un secondo problema, in un qualche senso complementare al primo, sempre collegato alla velocità di propagazione del calore. Invece di accontentarmi di vedere un effetto "fievole", mi chiedo: quando devo aspettare per vedere nel punto $x = L$ un innalzamento "consistente" della temperatura, dovuto alla propagazione del calore che "origina nell'origine" (ehm...)? Innanzi tutto devo stabilire cosa è ragionevole definire "consistente". Un modo per stabilirlo è il seguente: fissiamo l'attenzione sulla temperatura in $x = L$. Al tempo zero la temperatura è nulla (supponendo $L > \varepsilon$). Poi, come abbiamo visto, diventa immediatamente positiva e ci si può aspettare che aumenti fino ad un certo valore massimo che si avrà in corrispondenza di un certo valore del tempo $t = \tau(L)$. Dopodiché scenderà di nuovo tendendo a zero quando $t \rightarrow \infty$. Potremmo quindi (forse) definire velocità di propagazione del calore come il rapporto $L/\tau(L)$ (forse). Provo. Prendiamo la formula (9.48). Se assumo $\varepsilon L \ll 4Kt$ la formula approssimata che appare nel membro a destra della (9.48) è una buona approssimazione (giusto?), per cui si ha

$$u(L, t) \approx \frac{\varepsilon e^{-L^2/(4Kt)}}{\sqrt{\pi Kt}}.$$

Per calcolare il massimo, come funzione del tempo, poniamo uguale a zero la derivata

$$0 = u_t(L, t) = \frac{\varepsilon}{\sqrt{\pi K}} \left[-\frac{1}{2t^{3/2}} e^{-L^2/(4Kt)} + \frac{1}{\sqrt{t}} e^{-L^2/(4Kt)} \frac{L^2}{4Kt^2} \right],$$

da cui si ottiene

$$-\frac{1}{2} + \frac{L^2}{4Kt} = 0,$$

ovvero, il valore di $t = \tau(L)$ per il quale si ha un massimo della temperatura nel punto L è dato da

$$\tau(L) = \frac{L^2}{2K}.$$

Conclusione: *il tempo necessario per la per propagazione del calore fino al punto L non è proporzionale ad L , ma ad L^2 !* In questo senso possiamo dire che la velocità di propagazione del calore *non esiste*. Un comportamento di questo tipo in cui il tempo necessario a percorrere una certa distanza è proporzionale al quadrato della distanza è chiamato comportamento *diffusivo* e si verifica, ad esempio, nel moto browniano che (non a caso) è matematicamente intimamente legato all'equazione del calore. Osserviamo infine che il tempo $\tau(L)$ è inversamente proporzionale alla conducibilità termica K , come è giusto che sia.

9.8 Brevi considerazioni probabilistiche

9.8.1 La funzione di distribuzione cumulativa

Sia X una variabile casuale a valori reali e sia $F_X : \mathbb{R} \rightarrow [0, 1]$ la funzione di distribuzione (cumulativa) associata ad X , definita come

$$F_X(x) = \text{Prob}\{X \leq x\}.$$

F_X è detta *assolutamente continua* (rispetto alla misura di Lebesgue sulla retta) se esiste $f_X \in L_1(\mathbb{R})$ tale che

$$(9.49) \quad F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

In questo caso f_X è chiamata la funzione densità di probabilità di X . Se, ad esempio, $F_X \in C^1(\mathbb{R})$, allora F_X è assolutamente continua. Infatti, dal teorema fondamentale del calcolo integrale, si ottiene

$$F_X(x) = \int_{-\infty}^x F'_X(t) dt,$$

quindi la (9.49) è soddisfatta scegliendo $f_X = F'_X$.

In generale ci sono 2 tipi di ostacoli che possono impedire di scrivere la (9.49):

- (1) la variabile X ha una componente discreta, vale a dire esistono uno o più valori a_1, a_2, \dots tali che $\text{Prob}\{X = a_i\} = p_i > 0$. In questo caso F_X è discontinua nei punti a_i , in quanto

$$\begin{aligned} F(a_i) - \lim_{\varepsilon \rightarrow 0^+} F(a_i - \varepsilon) &= \lim_{\varepsilon \rightarrow 0^+} (F(a_i) - F(a_i - \varepsilon)) \\ &= \lim_{\varepsilon \rightarrow 0^+} \text{Prob}\{a_i - \varepsilon < X \leq a_i\} = \text{Prob}\{X = a_i\} = p_i > 0. \end{aligned}$$

Di conseguenza la (9.49) non può essere valida, altrimenti avremmo

$$\lim_{\varepsilon \rightarrow 0^+} (F(a_i) - F(a_i - \varepsilon)) = \lim_{\varepsilon \rightarrow 0^+} \int_{a_i - \varepsilon}^{a_i} f_X(t) dt = 0.$$

- (2) la variabile X ha componente cosiddetta *singolarmente continua* (rispetto alla misura di Lebesgue). Può accadere che F_X sia continua ma

$$(9.50) \quad F_X(x) \neq \int_{-\infty}^x F'_X(t) dt.$$

Senza entrare nel dettaglio di questo fenomeno, ci limitiamo a citare l'esempio della funzione di Cantor o scala di Cantor: si tratta di una funzione F continua, monotona non decrescente, derivabile quasi ovunque (vale a dire a meno di un insieme di misura di Lebesgue nulla), tale che $F'(t) = 0$ per ogni t in cui la derivata esiste. Quindi l'integrale che appare al membro a destra della (9.50) è nullo per la funzione di Cantor, ma il membro a sinistra è diverso da zero se $x > 0$.

9.8.2 Somma di variabili casuali indipendenti

Dato un insieme di variabili casuali (X_1, \dots, X_n) , possiamo associagli la funzione di distribuzione congiunta definita come

$$(9.51) \quad F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \text{Prob}\{X_1 \leq x_1, \dots, X_n \leq x_n\}.$$

Queste variabili si dicono indipendenti se

$$(9.52) \quad F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

Nel caso in cui le X_i sono assolutamente continue, possiamo derivare la relazione precedente ottenendo

$$(9.53) \quad f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

Siano ora X e Y due variabili casuali indipendenti assolutamente continue. Vogliamo calcolare la funzione densità di $X + Y$. Partendo dalla funzione di distribuzione, otteniamo

$$\begin{aligned} F_{X+Y}(t) &= \iint_{\{(x,y):x+y \leq t\}} f_{X,Y}(x,y) dx dy = \iint_{\{(x,y):x+y \leq t\}} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} \left(f_X(x) \int_{-\infty}^{t-x} f_Y(y) dy \right) dx = \int_{-\infty}^{\infty} f_X(x) F_Y(t-x) dx. \end{aligned}$$

Derivando rispetto a t si ha

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) dx = (f_X * f_Y)(t),$$

vale a dire la densità di probabilità della somma di due variabili indipendenti è data dalla convoluzione delle rispettive densità di probabilità. Generalizzando al caso di n variabili indipendenti, possiamo scrivere

$$(9.54) \quad f_{X_1+\dots+X_n} = f_{X_1} * \dots * f_{X_n}.$$

9.45 Problema. Siano X e Y due variabili casuali gaussiane con valor medio rispettivamente m_x, m_y e varianza rispettivamente σ_x^2, σ_y^2 . Calcolare la densità di probabilità di $X + Y$.

Soluzione. Si ha

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right) \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{(y-m_y)^2}{2\sigma_y^2}\right)$$

Quindi, utilizzando gli operatori di traslazione, posso scrivere

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma_x} T_{m_x} e^{-t^2/(2\sigma_x^2)} \quad f_Y(t) = \frac{1}{\sqrt{2\pi}\sigma_y} T_{m_y} e^{-t^2/(2\sigma_y^2)}.$$

Grazie al risultato del Problema 9.41, otteniamo

$$(f_X * f_Y)(t) = \frac{1}{2\pi \sigma_x \sigma_y} T_{m_x} T_{m_y} \left(e^{-t^2/(2\sigma_x^2)} * e^{-t^2/(2\sigma_y^2)} \right).$$

A questo punto, usando $T_a T_b = T_{a+b}$ e il calcolo fatto nel Problema 9.40, otteniamo

$$\begin{aligned} (f_X * f_Y)(t) &= \frac{1}{2\pi \sigma_x \sigma_y} T_{m_x+m_y} \sqrt{\frac{\pi}{1/(2\sigma_x^2) + 1/(2\sigma_y^2)}} \exp\left(-\frac{t^2}{2(\sigma_x^2 + \sigma_y^2)}\right) \\ &= \frac{1}{2\pi(\sigma_x^2 + \sigma_y^2)} \exp\left(-\frac{(t-m_x-m_y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right). \end{aligned}$$

In conclusione, $X + Y$ è anch'essa una variabile gaussiana con valor medio $m_x + m_y$ e varianza $\sigma_x^2 + \sigma_y^2$.

9.46 Problema. Siano X_1, \dots, X_n variabili casuali indipendenti ciascuna delle quali ha una distribuzione esponenziale con parametro α (vale a dire valor medio $1/\alpha$). Calcolare la funzione densità di probabilità di $S_n := X_1 + \dots + X_n$.

Soluzione. La densità di probabilità di una variabile casuale X esponenziale con valor medio $1/\alpha$ è data da

$$f_X(x) = H(x) \alpha e^{-\alpha x},$$

in cui H è la funzione a gradino di Heaviside. Poiché le X_i sono indipendenti e identicamente distribuite, per la (9.54), si ha

$$f_{S_n} = f_{X_1} * \cdots * f_{X_n} = f_X * \cdots * f_X = (f_X)^{*n}.$$

Per calcolare questo prodotto di convoluzione passiamo nello spazio di Fourier e poi torniamo indietro, usando la Proposizione 9.37. Quindi, sapendo che $\mathcal{F}(H(x)e^{-\alpha x})(s) = (\alpha + is)^{-1}$, si ottiene

$$(9.55) \quad \begin{aligned} f_{S_n} &= \mathcal{F}^{-1} \mathcal{F}((f_X)^{*n}) = \mathcal{F}^{-1} [\mathcal{F}(f_X)]^n \\ &= \mathcal{F}^{-1} \left[\left(\frac{\alpha}{\alpha + is} \right)^n \right] = \frac{1}{2\pi} P \mathcal{F} \left[\left(\frac{\alpha}{\alpha + is} \right)^n \right]. \end{aligned}$$

Osservo che

$$\left(\frac{\alpha}{\alpha + is} \right)^n = (i\alpha)^n \frac{1}{(i\alpha - s)^n} = \frac{(i\alpha)^n}{(n-1)!} D^{n-1} \frac{1}{i\alpha - s},$$

per cui

$$(9.56) \quad \begin{aligned} \mathcal{F} \left(\left(\frac{\alpha}{\alpha + is} \right)^n \right) (t) &= \frac{(i\alpha)^n}{(n-1)!} \mathcal{F} \left(D^{n-1} \frac{1}{i\alpha - s} \right) (t) \\ &= \frac{(i\alpha)^n}{n-1!} (it)^{n-1} \mathcal{F} \left(\frac{1}{i\alpha - s} \right) (t). \end{aligned}$$

Dalle (9.55), (9.56) ottengo

$$(9.57) \quad \begin{aligned} f_{S_n}(t) &= \frac{1}{2\pi} P \left[\frac{(-\alpha t)^{n-1} (i\alpha)}{(n-1)!} \mathcal{F} \left(\frac{1}{i\alpha - s} \right) (t) \right] = \frac{(\alpha t)^{n-1} (i\alpha)}{(n-1)!} \frac{1}{2\pi} P \mathcal{F} \left(\frac{1}{i\alpha - s} \right) (t) \\ &= \frac{(\alpha t)^{n-1}}{(n-1)!} \mathcal{F}^{-1} \left(\frac{\alpha}{\alpha + is} \right) (t) = \frac{\alpha^n t^{n-1}}{(n-1)!} H(t) e^{-\alpha t}. \end{aligned}$$

9.8.3 Funzione caratteristica

La *funzione caratteristica*⁶ di una variabile casuale X è definita come

$$\varphi_X(t) = \mathbb{E}(e^{itX}),$$

in cui $\mathbb{E}(\cdot)$ denota il valore atteso. Se X è assolutamente continua e ammette dunque una densità di probabilità f_X , si ha

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x) dx.$$

Vediamo dunque che la funzione caratteristica è, a meno di un segno, la trasformata di Fourier della densità di probabilità

$$(9.58) \quad \varphi_X = P \mathcal{F} f_X.$$

⁶Questa funzione caratteristica non ha nulla a che fare con la funzione caratteristica o funzione indicatrice di un insieme, definita nella (3.91).

Osservo che $f_X \in L_1(\mathbb{R})$, siccome f_X è non negativa e $\int_{\mathbb{R}} f_X(x) dx = 1$. Quindi la sua trasformata di Fourier è sempre ben definita. Grazie alla Proposizione 9.14 sappiamo che $\varphi_X \in C_0(\mathbb{R})$, di conseguenza, per la Proposizione 3.40, è uniformemente continua su \mathbb{R} .

9.47 Esempio. Voglio calcolare la funzione caratteristica di una variabile gaussiana X di media m e varianza σ^2 . La funzione densità di probabilità di X è, per definizione,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

Calcolo la sua trasformata di Fourier, a partire da un caso noto:

$$\begin{array}{llll} e^{-x^2} & & \xrightarrow{\mathcal{F}} & \sqrt{\pi} e^{-t^2/4} \\ \left[x \rightarrow \frac{x}{\sqrt{2\sigma}} \right] & e^{-x^2/(2\sigma^2)} & \xrightarrow{\mathcal{F}} & \sqrt{2\pi}\sigma e^{-t^2\sigma^2/2} \\ [x \rightarrow x - m] & e^{-(x-m)^2/(2\sigma^2)} & \xrightarrow{\mathcal{F}} & \sqrt{2\pi}\sigma e^{-t^2\sigma^2/2 - imt} \\ \left[f \rightarrow f/(\sqrt{2\pi}\sigma) \right] & \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/(2\sigma^2)} & \xrightarrow{\mathcal{F}} & e^{-t^2\sigma^2/2 - imt} \end{array}$$

Grazie alla (9.58), ottengo

$$\varphi_X(t) = e^{imt - t^2\sigma^2/2}.$$

9.48 Definizione. Sia X_1, X_2, \dots una successione di variabili casuali. Si dice che la successione $(X_n)_{n=1}^{\infty}$ converge in distribuzione oppure converge debolmente e si scrive $X_n \xrightarrow{d} X$ oppure $X_n \xrightarrow{w} X$, se, per ogni $f \in C_b(\mathbb{R})$, si ha

$$\mathbb{E}(f(X_n)) \rightarrow \mathbb{E}(f(X)).$$

Il motivo del nome *convergenza in distribuzione* deriva dal fatto essa è equivalente a richiedere che

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

per ogni $x \in \mathbb{R}$ in cui F_X è continua. Un modo per dimostrare che una successione di variabili casuali converge in distribuzione è proprio attraverso la funzione caratteristica. Infatti si dimostra il seguente

9.49 Teorema. Sia $(X_n)_{n=1}^{\infty}$ una successione di variabili casuali con rispettive funzioni caratteristiche φ_{X_n} . Sia X una variabile casuale con funzione caratteristica φ_X . Allora $X_n \xrightarrow{d} X$ se e solo se

$$\lim_{n \rightarrow \infty} \varphi_{X_n}(t) = \varphi_X(t) \quad \forall t \in \mathbb{R}.$$

Nella prossima sezione vedremo come questo risultato conduce ad una semplice dimostrazione di un risultato fondamentale in teoria della probabilità.

9.8.4 Il teorema del limite centrale “classico”

9.50 Teorema. *Siano X_1, \dots, X_n sono variabili casuali indipendenti e identicamente distribuite con valor medio μ e varianza σ^2 . Sia $S_n = X_1 + \dots + X_n$, e sia Z una variabile gaussiana con media nulla e varianza 1. Allora*

$$(9.59) \quad Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} Z.$$

Quasi-dimostrazione. Sia $Y_k = X_k - \mu$, per cui

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n Y_k.$$

Si ha

$$(9.60) \quad \mathbb{E}Y_k = \mathbb{E}(X_k - \mu) = 0 \quad \mathbb{E}Y_k^2 = \mathbb{E}[(X_k - \mu)^2] = \text{Var } X_k = \sigma^2.$$

Poiché le Y_k sono identicamente distribuite, esse hanno tutte la stessa funzione caratteristica che chiamiamo

$$\varphi_Y(t) = \mathbb{E}(e^{itY_k}).$$

Dalla formula di Taylor col resto di Lagrange sappiamo che esiste $\vartheta(x) \in (0, 1)$ tale che

$$e^{ix} = 1 + ix + \frac{(ix)^2}{2} e^{i\vartheta x} = 1 + ix - \frac{x^2}{2} [1 + (e^{i\vartheta x} - 1)],$$

per cui, ricordando la (9.60),

$$(9.61) \quad \begin{aligned} \varphi_{Z_n}(t) &= \mathbb{E} \left(1 + itY_k - \frac{t^2 Y_k^2}{2} [1 + (e^{i\vartheta t Y_k} - 1)] \right) \\ &= \mathbb{E}(1) + it\mathbb{E}(Y_k) - \frac{t^2}{2} \mathbb{E}(Y_k^2) - \frac{t^2}{2} \mathbb{E}[Y_k^2 (e^{i\vartheta t Y_k} - 1)] \\ &= 1 - \frac{t^2 \sigma^2}{2} - \frac{t^2}{2} \varepsilon(t), \end{aligned}$$

in cui abbiamo posto

$$(9.62) \quad \varepsilon(t) = \mathbb{E}[Y_k^2 (e^{i\vartheta t Y_k} - 1)].$$

Veniamo ora al punto per cui questa è una quasi-dimostrazione. La quantità all'interno del valore atteso nella (9.62) tende puntualmente a zero per $t \rightarrow 0$, in quanto l'esponenziale tende a 1. Ma questa convergenza in generale non è uniforme (lo è se, ad esempio, le variabili Y_k sono limitate). Tuttavia, grazie ad un teorema chiamato “della convergenza dominata”, è comunque possibile scambiare il limite $t \rightarrow 0$ col valore atteso $\mathbb{E}(\cdot)$ e si ottiene

$$(9.63) \quad \lim_{t \rightarrow 0} \varepsilon(t) = \mathbb{E} \left[\lim_{t \rightarrow 0} Y_k^2 (e^{i\vartheta t Y_k} - 1) \right] = 0.$$

A questo punto ci ricordiamo che le variabili X_k , e di conseguenza anche le Y_k sono indipendenti, per cui

$$\begin{aligned} \varphi_{Z_n}(t) &= \mathbb{E}(e^{itZ_n}) = \mathbb{E} \left(\exp \left(\frac{it}{\sigma\sqrt{n}} \sum_{k=1}^n Y_k \right) \right) = \mathbb{E} \left(\prod_{k=1}^n \exp \left(\frac{itY_k}{\sigma\sqrt{n}} \right) \right) \\ &= \prod_{k=1}^n \mathbb{E} \left(\exp \left(\frac{itY_k}{\sigma\sqrt{n}} \right) \right) = \prod_{k=1}^n \varphi_{Y_k} \left(\frac{t}{\sigma\sqrt{n}} \right) = \left[\varphi_Y \left(\frac{t}{\sigma\sqrt{n}} \right) \right]^n. \end{aligned}$$

Dalla (9.61) segue

$$(9.64) \quad \varphi_{Z_n}(t) = \left(1 - \frac{t^2}{2n} - \frac{t^2}{2n\sigma^2} \varepsilon\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n = \left(1 - \frac{t^2 + \delta_n(t)}{2n}\right)^n,$$

in cui abbiamo posto

$$\delta_n(t) = \frac{t^2}{\sigma^2} \varepsilon\left(\frac{t}{\sigma\sqrt{n}}\right).$$

Grazie alla (9.63) abbiamo

$$\lim_{n \rightarrow \infty} \delta_n(t) = 0.$$

Finalmente possiamo fare il $\lim_{n \rightarrow \infty}$ nella (9.64). Poiché $\log(1+x) = x + \mathcal{O}(x^2)$, si ottiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \varphi_{Z_n}(t) &= \lim_{n \rightarrow \infty} \exp \left[n \log \left(1 - \frac{t^2 + \delta_n(t)}{2n} \right) \right] \\ &= \lim_{n \rightarrow \infty} \exp \left[n \left(-\frac{t^2 + \delta_n(t)}{2n} + \mathcal{O}(n^{-2}) \right) \right] \\ &= \lim_{n \rightarrow \infty} \exp \left(-\frac{t^2}{2} + \frac{\delta_n(t)}{2} + \mathcal{O}(n^{-1}) \right) = e^{-t^2/2}. \end{aligned}$$

Ma come abbiamo visto $e^{-t^2/2}$ è proprio la funzione caratteristica di una variabile gaussiana con media 0 e varianza 1. Possiamo dunque concludere che

$$\lim_{n \rightarrow \infty} \varphi_{Z_n}(t) = \varphi_Z(t) \quad \forall t \in \mathbb{R},$$

che, grazie al Teorema 9.49, implica la (9.59). \square

9.9 Problemi

9.51 Problema. Sia $f \in L_1(\mathbb{R})$. Dimostrare (se vero) o fornire un controesempio (se falso):

- (a) se f è reale, allora \hat{f} è reale;
- (b) se f è pari, allora \hat{f} è pari;
- (c) se f è a supporto compatto, allora \hat{f} è a supporto compatto.

9.52 Problema. Sia $(f_n)_{n=1}^\infty$ una successione di funzioni nello spazio $L_1(\mathbb{R}; \mathbb{C})$ che converge ad $f \in L_1(\mathbb{R}; \mathbb{C})$ rispetto alla norma $\|\cdot\|_1$. Sia $\hat{f}_n := \mathcal{F}[f_n]$ e $\hat{f} := \mathcal{F}[f]$. Dimostrare che \hat{f}_n converge uniformemente a \hat{f} .

9.53 Problema. Sapendo che

$$\mathcal{F} \left[x e^{-(x^2-x)} \right] (t) = \frac{\sqrt{\pi}}{2} e^{-(t+i)^2/4} (1-it)$$

calcolare

$$\mathcal{F} \left[\frac{\sqrt{\pi}}{2} e^{-(x+i)^2/4} (1-ix) \right] (t).$$

9.54 Problema. Calcolare

$$\int_{\mathbb{R}} \left(\frac{\sin(ax)}{x} \right)^4 dx. \quad \forall a \in \mathbb{R}$$

(Sugg: utilizzare opportunamente la trasformata di Fourier).

Risp: $2|a|^3\pi/3$.

★ **9.55 Problema.** Trovare una funzione $f \in L_1(\mathbb{R})$ la cui trasformata di Fourier \hat{f} tende a zero all'infinito più lentamente di $(\log x)^{-2}$. Più precisamente bisogna far vedere che esiste una successione (λ_k) di numeri reali tale che $\lambda_k \rightarrow +\infty$ e tale che

$$\textcircled{1} \quad \lim_{k \rightarrow \infty} [\log(\lambda_k)]^2 \hat{f}(\lambda_k) = +\infty.$$

(Sugg: cercare f della forma $f = \sum_{n=1}^{\infty} c_n \mathbb{I}_{[-a_n, a_n]}$, in cui $a_n \rightarrow 0$ e $c_n \rightarrow \infty$, e aggiustare i coefficienti a_n e c_n in modo tale che: (1) $f \in L_1(\mathbb{R})$ e (2) valga la $\textcircled{1}$ per un'opportuna successione λ_k).